Г. Рейхенбах

Направление Времени

Hans Reichenbach
The Direction of Time
Edited by *Maria Reichenbach*

Г. Рейхенбах

Направление Времени

Перевод с английского
Ю. Б. Молчанова и Ю. В. Сачкова

Издание второе, стереотипное

Москва • 2003 УРСС

ББК 22.3о, 22.3я44, 87.2

Рейхенбах Ганс

Направление времени: Пер с англ. Изд. 2-е, стереотипное. — М.: Едиториал УРСС, 2003. — 360 с.

ISBN 5–354–00275–3

Настоящее произведение известного логика и философа Ганса Рейхенбаха является одной из наиболее важных его работ по изучению проблемы времени. В книге приводятся исторический обзор и отчасти критическая оценка различных интерпретаций данной проблемы. Характер обсуждения проблемы времени у Г. Рейхенбаха значительно отличается от постановки вопроса в традиционной философии.

В этой последней из своих книг автор стремится подвести итог всем своим философским исследованиям и часто рассматривает ее как труд, завершающий и обобщающий все его работы, касающиеся вопросов теории вероятностей, квантовой механики, теории относительности и причинности — вопросов, которые играют важную роль в этой книге.

Данный труд является также примером применения аналитического метода научной философии, в развитие и совершенствование которого Г. Рейхенбах также внес свой вклад.

Издательство «Едиториал УРСС». 117312, г. Москва, пр-т 60-летия Октября, 9.
Лицензия ИД № 05175 от 25.06.2001 г. Подписано к печати 13.05.2003 г.
Формат 60×90/16. Тираж 960 экз. Печ. л. 22,5. Зак. № 2-981/180.

Отпечатано в типографии ООО «Рохос». 117312, г. Москва, пр-т 60-летия Октября, 9.

ИЗДАТЕЛЬСТВО УРСС
НАУЧНОЙ И УЧЕБНОЙ ЛИТЕРАТУРЫ
E-mail: URSS@URSS.ru
Каталог изданий
в *Internet*: http://URSS.ru
Тел./факс: 7 (095) 135–44–23
Тел./факс: 7 (095) 135–42–46

ISBN 5–354–00275–3

© Едиториал УРСС, 2003

ПРЕДИСЛОВИЕ

Среди различных рукописных работ, оставшихся после внезапной смерти моего мужа в апреле 1953 года, наиболее важной является работа «Направление времени», посмертную публикацию которой и представляет настоящее издание.

Проблема времени занимала философов различных эпох. Первая глава данной книги представляет собой исторический обзор и отчасти критическую оценку различных интерпретаций этой проблемы, которые предлагались традиционными философскими системами. Кроме того, проблема времени, безусловно, находится в центре современных физических и философских дискуссий. Однако характер обсуждения этой проблемы на базе достижений современной науки значительно отличается от постановки вопроса в традиционной философии. Так как научная философия придает особое значение неразрывной связи между науками и философской экспликацией их фундаментальных понятий, то быстрое изменение физических понятий, естественно, имело значительное влияние на философские исследования в этом направлении. В данной книге предлагается экспликация некоторых свойств физического времени, в которой учитываются достижения современной физики.

Интерес к анализу проблемы времени характерен и для прежних работ моего мужа. В своих главных работах по теории относительности, например в «Теории от-

носительности и априорном знании» («Relativitaetstheorie und Erkenntnis apriori»), а также в «Философских проблемах учения о пространстве — времени» («Philosophie der Raum-Zeit-Lehre»), он рассматривал главным образом количественные, или метрические, свойства времени, тогда как в настоящей работе анализирует качественные, или топологические, атрибуты, такие, как порядок и направление. Он был убежден, что достиг решения проблемы направления времени, которое в состоянии дать ответы на все правильно поставленные вопросы. При этом многие вопросы, например такие, почему мы не можем изменить прошлое, или вопрос о свободе воли, которые другие философы были склонны игнорировать как псевдопроблемы, он считает вполне закономерными.

В плане работы над книгой содержалась заключительная глава, которую мой муж, к сожалению, не успел написать. Однако, несмотря на отсутствие этой главы, анализ физического времени можно считать завершенным, и поэтому книгу нельзя рассматривать как собрание отдельных фрагментов. В последней главе автор намеревался рассмотреть проблему связи между субъективным переживанием времени человеком и объективными свойствами времени в природе. В этой главе, посвященной человеческому разуму, он намеревался возвратиться к анализу качественных свойств времени, рассмотренных в § 2 данной книги, но уже исходя из анализа свойств физического времени, проведенного в § 3—30. Совпадение направления субъективного времени с положительным направлением физического времени он намеревался объяснить тем, что человеческая память представляет собой естественный регистрирующий инструмент, который хранит записи о прошлом, подобно другим естественным или искусственным протоколам. Последняя глава, следовательно, представляла бы собой применение общей теории физического времени к спе-

циальному случаю, а не дальнейшее развитие теории. В дополнении к этой книге приводится краткая формулировка некоторых из тех идей, которые он намеревался развить в заключительной главе.

Незадолго до своей смерти мой муж получил приглашение от Гарвардского университета прочитать лекции в честь Уильяма Джемса в период осеннего семестра 1953 года. Он выбрал тему «Время и свобода воли» и намеревался исходить из рукописи данной книги, опуская при этом наиболее специальные главы, для того чтобы сделать изложение более доступным для широкой аудитории.

Краткое содержание результатов, полученных автором в исследовании направления времени, было изложено в серии лекций, прочитанных в июне 1952 года в Сорбонне, в Институте имени Анри Пуанкаре (Париж). Эти лекции предназначались для физиков и математиков, и в них рассматривались в основном вопросы квантовой механики. В дальнейшем они были опубликованы [1].

В этой последней из своих книг автор стремился подвести итог всем своим философским исследованиям и часто рассматривал ее как труд, завершающий и обобщающий его работы, касающиеся вопросов теории вероятностей, квантовой механики, теории относительности и причинности — вопросов, которые играют важную роль в этой книге. Кроме того, интересно отметить, что здесь получает новое применение его теория эквивалентных описаний. Эта книга является также примером применения аналитического метода научной философии, в развитие и совершенствование которого он также внес свой вклад.

Конечно, было бы лучше, если бы автор сам просмотрел текст перед публикацией. Однако та минималь-

---

[1] «Les Fondements logiques de la mécanique des quanta», Annales de l'Institut Henri Poincaré, tome XIII, Paris, 1953 Fasc. II, p. 109—158.

ная редакторская правка, которая оказалась необходимой, не внесла никаких существенных изменений. Было добавлено несколько примечаний, которые ясно можно отличить от примечаний автора, так как они помечены моими инициалами и заключены в скобки.

Мне хотелось бы выразить искреннюю признательность и благодарность проф. Альфреду Ланде, проф. Адольфу Грюнбауму, д-ру Весли Салмону, д-ру Норману Далки, д-ру Альберту Латтеру, Рут Анне Матерс и Роберту Л. Маттерсу за их подробные консультации и великодушную помощь в окончательной подготовке рукописи к печати.

Рисунки и диаграммы в данной книге выполнены Джеймсом Мюрреем и Томасом Станиэлем.

Я также глубоко обязана Елене Травис за терпеливое редактирование рукописи и издательству «Университет Калифорния пресс», которое сделало все, что было в его силах, для издания этой книги.

Лос-Анжелос, май, 1955. *Мария Рейхенбах*

Глава I

ВВЕДЕНИЕ

§ 1. Эмотивное (emotive) значение времени

Проблема времени всегда заводила в тупик человеческий разум. Не только события внешнего мира, но и все наши субъективные переживания происходят во времени. Время представляется таким образом, как будто бы его поток, который упорядочивает события физического мира, протекает через человеческое сознание и вынуждает его приспосабливаться к этому порядку. Наши наблюдения над физическими предметами, наши чувства и эмоции, а также процессы мышления имеют протяженность во времени, и нам нельзя избавиться от этого неумолимого потока, который непрерывно течет от прошлого через настоящее к будущему.

То, что мы испытываем в один момент времени, в следующий момент ускользает в прошлое. Прошлое никогда не возвращается, оно неизменно, недоступно никакому контролю со стороны того, что приносит нам будущее, и все же оно хранится в нашей памяти как нечто такое, что однажды наполняло наш опыт в качестве непосредственного настоящего. Неужели оно никогда не вернется? Почему с нами этого не может случиться вторично?

Оставаясь безразличным к нашему вопросу, поток времени продолжает свое движение. Уже наше настоящее наполнено другими переживаниями, которых раньше мы совершенно не могли предвидеть. Хотя опыт настоящего времени отчасти предсказуем, однако он содержит много неожиданных и неизвестных ранее свойств. Что раньше было неопределенным, теперь определено. Возможности, которых мы ранее опасались или на которые надеялись, стали теперь реальностями; появились другие возможности, о которых мы никогда не думали.

И даже повседневный опыт, хотя и в высшей степени предсказуемый, обнаруживает в действительности такие специфические черты, которые никогда нельзя было предвидеть. Что ожидает нас в будущем? Может быть, случится война или другая политическая катастрофа? Дождемся ли мы увеличения заработной платы, на которое так долго надеялись? Может быть, придет письмо, сообщающее нам о смерти друга, который, как мы полагали, был совершенно здоров? Или письмо известит нас о счастливом для нас завещании дальнего родственника? И каковы вообще будут те мелочи жизни, с которыми мы столкнемся? Сразу ли тронется этот автомобиль? Пересечем ли мы перекресток до того, как загорится красный свет? Что скажет Фред, когда я сообщу ему, что Дорис выходит замуж за Джона?

Все это находится в будущем. Но что такое будущее? Хранит ли оно события, так сказать, как вещи на складе, а затем распределяет их согласно какому-то плану? Или события возникают случайно? Возникновение означает становление. Что такое становление? Как может нечто нереальное стать реальным? Ведь как только оно становится реальным, то сразу же уходит в прошлое, вновь становясь нереальным, не оставляя ничего, кроме тени в нашей памяти. Единственной реальностью является настоящее. Когда оно уходит, мы вступаем в новое настоящее, всегда оставаясь, таким образом, в вечном «теперь». Что же такое время, если все, что мы получаем от него, является именно этим «теперь», тем единственным моментом, который уходит вместе с нами сквозь поток событий, текущий от неизменяемого прошлого к непознаваемому будущему?

Подобные вопросы раскрывают высоко эмоциональное содержание, связанное с переживанием времени. Они соблазняют искать ответы, удовлетворяющие нашим эмоциям, но не ведущие к установлению ясных значений. Я не хочу сказать, что такие вопросы бессмысленны. Однако ответы на них могут сильно отличаться от того, что мы ожидаем. И, может быть, мы просто не сможем найти на них ответы до тех пор, пока сначала не проанализируем эти вопросы и не уточним их, что на данном этапе является лишь началом поиска ясных значений. Процесс человеческого мышления не следует схемам вычислительных машин, которые имеют ответ на

любой вопрос, при условии, что этот вопрос поставлен правильно. Мы не можем ответить на любой правильно поставленный вопрос, однако часто можем давать ответы на такие вопросы, которые поставлены неверно, придавая им сперва такую форму, в которой они имеют значение. Часто процессы переформулирования вопросов и получение ответа на них тождественны. В поисках ответа мы раскрываем новые значения и находим то, что искали.

Это и есть научный подход. Не ожидайте ответа до тех пор, пока не нашли ясных значений. Но не отбрасывайте неясных вопросов. Имейте их в виду до тех пор, пока не получите средства разъяснить их и одновременно дать на них ответ. Часто эти средства получаются в результате развития науки в других областях знания, которые на первый взгляд, кажется, не имеют никакого отношения к данному вопросу.

В истории философии имеется немало примеров процесса выяснения значений. Фалес из Милета полагал, например, что вода является субстанцией, из которой состоит все вещи. Гераклит же утверждал, что этой таинственной субстанцией является огонь. Но никто из них не знал точно, что означает утверждение, что данный кусок вещества состоит из некоторого количества субстанции. Современная химия с помощью методов химического анализа придала этому утверждению точное значение и доказала, что ни вода, ни огонь не являются химическими элементами. Другая иллюстрация имеется в философии Платона. Платон полагал, что геометрические отношения познаются благодаря созерцанию или воспоминанию переживаний, которые наша душа получила в потустороннем мире, на небесах, задолго до того, как началась ее земная жизнь. Современная математика показала, однако, что акт созерцания геометрических фигур можно понять посюсторонним образом: это воспоминание о нашем повседневном опыте, имеющем дело с объектами нашего окружения. Данный ответ вносит ясность и в значение термина «созерцание». И только здесь выясняется, что поставленный вопрос предполагает другое значение.

Аналогична история исследования природы времени. Эта проблема, которая весьма озадачивала древних, оставалась нерешенной в течение двух тысяч лет. Ответ

на нее был найден с помощью достижений современной физики, непосредственно связанных не с проблемой времени, а с проблемой причинности. Прежде чем перейти к рассмотрению этих достижений, было бы неплохо более тщательно рассмотреть концепции времени, нашедшие отражение в прошлых философских системах, так как в них раскрываются те эмоциональные реакции и формулируются те логические трудности, с которыми сталкивается каждый из нас при рассмотрении времени.

Наша эмоциональная реакция на течение времени в значительной степени определяется его неотвратимостью. Течение времени находится вне нашего контроля. Мы не можем остановить его, повернуть вспять, чувствуем, как оно уносит нас, и мы столь же беспомощны, как щепка в речном потоке. Мы знаем прошлое, но не в силах изменить его. Наша деятельность может быть направлена только в будущее. Но наши знания о будущем неполны; могут произойти неожиданные события, которые расстроят все наши планы. Конечно, будущее может принести и благоприятные исходы событий. Однако мы знаем, что число их ограниченно и что наше приспосабливание к тому, что может принести нам будущее, по существу, бесполезно; впереди — только ограниченный промежуток времени, и в конечном счете наша борьба с новыми ситуациями завершается смертью. Смерть есть неизбежный результат необратимого течения времени. Если бы мы могли остановить время, тогда бы смогли избежать смерти. Тот факт, что мы не можем этого сделать, определяет в конечном счете наше бессилие, приравнивая нас к щепке, дрейфующей в речном потоке. Поэтому страх перед смертью превращается в страх перед временем, ибо поток времени представляется нам выражением сверхчеловеческих сил, от которых нет спасения. Выражение «преставиться», в котором уклончиво, не называя ее по имени, говорится о смерти, отражает наше эмоциональное отождествление потока времени со смертью.

Эмоциональное неудовлетворение часто переносится в логику. В различных учениях о мироздании оно иногда проявляется в виде логических вопросов и псевдологических построений. Философ утверждает, что обнаружил загадку бытия, которую нельзя решить логическим путем. С таким же успехом он мог бы сказать, что обна-

ружил факт, возбуждающий в нем эмоциональное неприятие. На логический анализ проблемы времени, который предпринимался философами, большое влияние оказывал страх перед смертью. Убеждение в том, что они обнаружили парадоксы в течении времени, можно назвать «проекцией», согласно современной психологической терминологии. Оно функционирует как защитный механизм. С помощью парадоксов стремятся дискредитировать физические законы, которые возбуждают глубокий эмоциональный антагонизм.

Представители религиозной философии утверждают, что происходящее во времени не составляет всей реальности. Они настаивают на том, что существует еще другая высшая реальность, свободная от потока времени, и только низшая реальность человеческого опыта связана с временем. Довольно странно, что предполагаемая высшая реальность называется *вечной*, то есть обозначается термином, связанным с временем. Однако на языке, который употребляют эти философы, данный термин обозначает уже не непрерывную продолжительность, а что-то существующее вне времени, не подверженное временно́му потоку. Ее противоположностью, согласно церковной терминологии, является *суетное*, *мирское* (secular) [1]. Этот термин первоначально обозначал время жизни человека (которое, согласно распространенному мнению, равно столетию). Однако в данном случае этот термин обозначает нечто подверженное течению времени, то есть нечто земное, а именно проявление такой реальности, которая обладает низшей, физической природой. Жажда бессмертия и вечной, в смысле бесконечного времени, жизни, обычно несовместимая с физическими фактами, привела, таким образом, к концепции, согласно которой вечная жизнь есть жизнь не во времени, а в другой реальности. Для того чтобы избежать «преставления», которое неизбежно с течением времени, была придумана вневременна́я реальность.

В древности такие концепции были выдвинуты Парменидом и Платоном. Парменид говорит, что высшая реальность не возникает и не исчезает. «Не возникает

[1] Слово «secular» в английском языке означает: 1) вековой; 2) происходящий раз в сто лет; 3) мирской, светский. — *Прим. перев.*

оно и не подчиняется смерти. Цельное все, без конца, не движется и однородно. Не было в прошлом оно, не будет, но все — в настоящем. Без перерыва, одно». А Платон объясняет, что «время есть движущийся образ вечности». Здесь «вечность» не означает «бесконечное время». Предполагается, что этот термин обозначает реальность, не контролируемую течением времени, которая, так сказать, отражается в потоке времени. Случающееся во времени, в лучшем случае, является низшей формой реальности. Для Парменида оно кажется вовсе нереальным, иллюзорным.

Взгляды этих философов свидетельствуют об эмоциональной неудовлетворенности. Они употребляют метафоры, придуманные для того, чтобы удовлетворить стремлению вырваться из потока времени и ослабить страх перед смертью. Эти метафоры нельзя выразить в логически непротиворечивой форме. Однако довольно странно, что их иногда представляют в виде результатов логического анализа. При их обосновании ссылаются на парадоксы становления. Парменид утверждал, что если бы имелось становление, то вещь должна была бы возникать из ничего, что он считал логически невозможным. Его преемник в школе элеатов, Зенон, сформулировал знаменитые апории, которые, как он полагал, доказывают невозможность движения и справедливость концепции Парменида о вневременно́м бытии.

Апории Зенона обсуждались неоднократно. Он доказывал, что если движение есть переход из одной точки в другую, то летящая стрела не может двигаться, раз она находится именно в одной точке. Как она может передвинуться в другую точку? Прыгает через безвременный интервал? Очевидно, нет. Следовательно, движение невозможно.

Рассмотрим другую апорию. Ахиллес должен догнать черепаху. Сперва он должен достичь точки, от которой черепаха начала свое движение, но за это время черепаха уже передвинется дальше, в другую точку. В таком случае Ахиллесу нужно достичь уже этой точки, однако за это время черепаха продвинется еще дальше, и так без конца. Чтобы догнать черепаху, Ахиллесу нужно пройти бесконечное число отрезков, не равных нулю, однако он не в силах проделать это. Следовательно, он не сможет догнать черепаху.

Относительно апории «стрела» мы теперь отвечаем, что можно провести различие между покоем и движением в одной точке. «Движение» определяется, говоря более точно, как «переход от одной точки к другой в конечный и неисчезающий промежуток времени». Точно так же «покой» определяется как «отсутствие перехода от одной точки к другой в конечный и неисчезающий промежуток времени». Утверждение «покоиться в данной точке в данный момент времени» не определяется предшествующими дефинициями. Для того чтобы определить его, мы определяем понятие «скорость» с помощью предельных процессов, подобных тем, которые используются в дифференциальном исчислении. В таком случае «покой в данной точке» определяется как нулевое значение скорости. Этот логический анализ приводит к выводу, что летящая стрела в каждой точке обладает скоростью, отличной от нуля, и, следовательно, не покоится. К тому же вопрос, каким образом стрела может достичь следующей точки, является неправильно поставленным, потому что в континууме следующей точки вообще не существует. Непрерывность ряда целых чисел, где за каждым целым числом следует другое целое число, нужно отличать от непрерывности ряда точек, поскольку между любыми двумя точками всегда имеется какая-то другая точка. Что касается второй апории, то мы утверждаем, что Ахиллес может догнать черепаху, потому что бесконечное число неисчезающих отрезков, стремящихся к нулю, может иметь конечную сумму и покрываться за конечный промежуток времени.

Получение точных ответов на эти вопросы потребовало разработки теории бесконечности и предельных процессов, что было выполнено только в XIX столетии [1]. Поэтому в истории логики и математики апории Зенона занимают важное место. Они привлекали внимание к тому факту, что логическая теория упорядоченной одномерной совокупности точек (континуума) не может быть разработана до тех пор, пока отказываются от предположения о существовании некоторых простых законо-

---

[1] Современное рассмотрение апорий Зенона см.: B. Russel, Our knowledge of External World, Chicago, Open. Court Publ. Co., 1914, Chap. VI; см. также: Adolf Grünbaum, Reality and the Atomicity of Becoming, Rev. Metaphysics, 1950, vol. 4, p. 176.

мерностей, свойственных рядам целых чисел. В ходе этих исследований математики обнаружили, что понятие бесконечности поддается последовательному логическому анализу, что бесконечность точек на линии отличается от бесконечности ряда целых чисел и что апории Зенона не относятся только к течению времени, поскольку их точно так же можно сформулировать и решить и для чисто пространственного континуума.

Однако апории Зенона с психологической точки зрения интересны потому, что они были открыты не в результате поисков математической теории континуума, а в результате процесса рационалистического объяснения, и в их основе лежит желание элеатов доказать нереальность времени. Если бы Зенон при формулировке своих апорий не находился под влиянием «метафизической» цели, то он, возможно, пришел бы к другому решению. Он рассуждал бы иначе: поскольку в действительности стрела летит, а бегун догоняет черепаху, то, значит, имеется какая-то ошибка в логике его рассуждений, а не в физической реальности. Однако он не пожелал сделать такой вывод. Он хотел доказать, что изменение и становление иллюзорны, что бытие существует вне времени и не зависит от преходящего, контролируемого временем человеческого опыта, не подвержено исчезновению и смерти.

Теория времени Парменида стала историческим символом отрицательного эмоционального отношения к потоку времени. Однако с действительной структурой времени совместимы различные эмоциональные реакции, поэтому всегда существовало положительное отношение к временнбму потоку, утвердительный эмоциональный ответ на изменение и становление, для которого будущее является неисчерпаемым источником нового опыта и вызовом нашим способностям использовать наилучшие из возникающих возможностей. Историческим символом положительного эмоционального отношения к течению времени является философия Гераклита, современника и противника Парменида.

«Все течет» — в этой форме можно суммировать философию Гераклита. Становление является для него сущностью жизни. Утверждение «Солнце новое каждый день» означает следующее: хорошо, что каждый день порождает нечто новое. Нам не нужно цепляться за

прошлое, ибо мы прекрасно можем жить в мире постоянных изменений. «В одну и ту же реку нельзя войти дважды, ибо воды в ней вечно новые». Этот кажущийся парадокс не так глубок, как полагал Гераклит, ибо мы прекрасно можем называть реку той же самой, несмотря на то, что вода в ней все время меняется. Однако афоризм Гераклита обращает наше внимание на логическую природу физических вещей как серию различных состояний во времени. Для физического тождества необходимо, чтобы эти состояния были совершенно одинаковы. Человек все время остается тождественной себе личностью, хотя его тело растет, а химический состав его клеток изменяется. Физика вещей не требует отрицания временного потока. Концепция Гераклита согласуется как со здравым смыслом, так и с наукой.

Однако пророк временного потока не так уж много смог сказать о логической структуре времени. Афоризмы Гераклита дают нам эмотивные комментарии, а не логический анализ времени. Для логики оказались полезными не его требования изменений, а вопросы Зенона о природе изменения. Попытка Гераклита доказать тождество противоположностей, очевидно вытекающая из признания того, что одна и та же вещь может состоять из различных во времени состояний, является одним из тех сверхупрощенных обобщений, которые отчасти представляют собой обычные трюизмы, а отчасти явно ошибочны. «Путь вверх и вниз — один и тот же», — в этом положении формулируется тот тривиальный факт, что отношение и его противоположность могут быть с одинаковым успехом использованы для эквивалентных утверждений; термин «выше» может выражать то же, что и термин «ниже», если изменить порядок их связи. Однако было бы противоречием говорить, что утверждение «Петр выше Павла» равнозначно утверждению «Петр ниже Павла». В других примерах, иллюстрирующих так называемое тождество противоположностей, Гераклит приводит просто такие случаи, когда крайности являются противоположными концами одной масштабной шкалы, например горячее и холодное. Далее он ссылается также на тот тривиальный факт, что одна вещь может находиться в противоположных отношениях к различным вещам, как в его примере с морской водой, полезной для рыб и вредной для человека. Упомянутая логика проти-

воположностей не может решить проблему времени и становления.

Подход Гераклита к проблеме временнóго порядка является наивным, он представляет собой попытку понять время только путем размышления над понятиями, полученными из повседневного опыта. К сожалению, подобный подход даже в наше время рассматривается многими как истинно философский, в особенности если полученные выводы сформулированы на темном языке оракулов, подобном языку Гераклита. Однако темный язык очень часто оказывается лишь внешним нарядом философии, оперирующей тривиальностями наряду с ложными и абсурдными положениями. Эта философия либо провозглашает учение о единстве противоположностей, защищая доктрину, согласно которой противоречие является источником движения и жизни, или же защищает концепцию о превращении ничто в нечто. Разъяснение сущности времени и становления можно ожидать только в том случае, когда на вопросы, выдвигаемые здравым смыслом, отвечают с помощью научного метода. В нем объединяются точность научных формулировок, возможность проверки с помощью наблюдения и логических методов и далеко идущая способность связывать факты из весьма различных областей, которая представляет инструмент научного исследования, способный пролить новый свет на проблемы, выдвигаемые повседневным опытом. Для того чтобы достигнуть логической ясности, нужно связать анализ свойств времени с анализом науки.

Даже при поверхностном рассмотрении становится ясным, что изучение времени является задачей физики. Эмотивные реакции на поток времени не могут ответить на вопрос: что такое время? Субъективное переживание времени, хотя и вызывает эмоциональную реакцию, не в состоянии дать достаточную информацию о временнóм порядке, который связывает физические события. Мы знаем, что субъективное мнение о скорости временнóго потока является обманчивым. Тот факт, что при одних обстоятельствах нам кажется, что время течет быстро, а при других — медленно, зависит, например, от того, увлечены ли мы чем-либо или скучаем от безделья. Психологи доказали, что, строго говоря, мы называем

настоящим временем не временну́ю *точку*, а короткий временной *интервал*, длина которого характеризуется психологическим порогом восприятия времени. Например, для «построения» зрительного образа требуется время; этот факт объясняет восприятие движения в кинофильме, которое состоит из статических картин, показываемых в быстрой последовательности. Однако нам представляется, что существование восприятия времени не имеет отношения к изучению времени как объективного процесса, так же как и существование порога восприятий не имеет отношения к исследованию длины или интенсивности звука. В самом деле, если критический анализ свойств континуума, данный Зеноном, справедлив и в отношении времени, тогда следует предположить, что физический процесс течения времени не зависит от психологического переживания времени, структура которого отличается от структуры математического континуума и «атомистична» по своей природе[1]. Какова структура того времени, которое контролирует физические события? Нам необходимо узнать, оправдана ли наша эмотивная реакция на время, существует ли, говоря объективно, поток времени, который заставляет события уходить в прошлое и не позволяет им вернуться в настоящее. В каком смысле будущее отличается от прошлого? Если мы хотим понять само время, а не только психологические реакции на него, то нам следует обратиться к физике.

Выше было сказано, что прошлое отличается от будущего так же, как неизменное от непознаваемого. Следует ли из этого различия, что будущее все же доступно изменению? Исходя из нашего повседневного опыта, мы склонны ответить на этот вопрос утвердительно. Наш контроль над будущим, хотя и ограниченный, конечно, по своим размерам, часто достаточен для удовлетворения наших потребностей. Планируемые действия, основанные на ожидании будущих результатов, дают нам возможность практически использовать многие из даров будущего. Мы сеем и пожинаем плоды, мы строим себе кров, организуем человеческое общество, создаем машины, облегчающие наш повседневный труд.

---

[1] Это было убедительно подчеркнуто Грюнбаумом (Grünbaum, op. cit., p. 167).

Возможно, однако, что ученый был бы склонен оспаривать веру в возможность изменения будущего. Он мог бы аргументировать, что непознанное еще не означает недетерминированное и, возможно, что будущее так же детерминировано, как и прошлое, а различие между прошлым и будущим сводится к различию между познанным и непознанным. Очевидная асимметрия времени была бы в таком случае лишь вопросом знания и незнания, само же время было бы симметричным, его объективная природа была бы тождественной в направлении как прошлого, так и будущего. Такие концепции выдвигались в рамках самой науки, поскольку она признавала универсальную справедливость принципа причинности.

Причинные законы управляют как прошлым, так и будущим. Мы видим их действие в фактах прошлого, но мы видим также, как они подтверждаются фактами будущего, которые мы правильно предсказали и которые позднее стали реальностью. Будущее не является совершенно непознаваемым. Можно предсказать немало событий, например такие, как движение планет, смена времен года, рост растений, животных, людей и, конечно, смерть. Философов к сомнению в реальности времени приводит возможность предсказания некоторых нежелательных будущих событий, например таких, как смерть. Почему не предположить, что будущее так же детерминировано, как и прошлое?

Подобную точку зрения вновь можно проиллюстрировать множеством примеров из античной философии. Идея причинного детерминизма была сформулирована еще в античной философии. Иногда она выдвигалась в такой форме фатализма, когда предопределены только конечные цели человеческих устремлений, в то время как пути их осуществления остаются открытыми для случайности. Этот антропоморфический вид детерминизма напоминает страх ребенка перед родительским наказанием, превратившийся в страх перед небесным отцом. Однако цель может не иметь никакого отношения к вопросу о детерминации; можно считать, что детерминированность определяется только причинными законами, теми же причинными законами, которые управляют движением планет и ростом живых организмов. Некоторые античные философы, подобно Демокриту, защищали именно такой детерминизм. Однако подлинный причинный детерми-

низм в современном смысле слова возник вместе с современной наукой.

Физика Галилея и Ньютона показала, что можно предсказать гораздо больше событий, чем это можно предвидеть с точки зрения здравого смысла, и что предсказание может достигнуть изумительно высокой степени точности. Этими успехами в области предсказания физическая наука была обязана применению математических методов. Последние не только в высшей степени усовершенствовали астрономию и навигацию, но и способствовали развитию техники, что было бы невозможно, если бы причинные законы не управляли физическими событиями с исключительной степенью точности. Паровая машина и самолет свидетельствуют в пользу детерминированности будущего. Кто бы отважился лететь на самолете, если бы не был убежден, что законы аэродинамики обеспечивают высокую надежность предсказаний.

Не удивительно, что с прогрессом современной науки детерминизм становился все более влиятельной доктриной. Ньютонова физика открыла физические законы, управляющие как небесными, так и земными телами. Предполагалось также, что эти же законы царствуют в области атома и контролируют движение микрочастиц. Математик Лаплас без колебаний предположил, что точность астрономических законов сохраняется также и в атомной абласти. Поскольку даже мышление и чувства человека являются только отражением движения совокупностей атомов в мозгу, он пришел к выводу, что любое будущее событие детерминировано точно так же, как и прошлое. Только наше невежество не позволяет предсказывать будущее. В своем известном замечании о логически возможном сверхчеловеке он сформулировал отношение полной симметрии между прошлым и будущим.

«...мы должны рассматривать настоящее состояние вселенной как следствие ее предыдущего состояния и как причину последующего. Ум, которому были бы известны для какого-либо данного момента все силы, обусловливающие природу и относительные положения всех ее составных частей, если бы вдобавок он оказался достаточно обширным, чтобы подчинить эти данные анализу, обнял бы в одной формуле движение величайших тел вселенной наравне с движением легчайших атомов, не

оставалось бы ничего, что было бы для него недостоверно, и будущее, так же как и прошлое, предстало бы перед его взором»[1].

Если бы это высказывание, ставшее классической формулировкой детерминизма, было верным, то это означало бы крушение реалистической интерпретации потока времени. Если будущее так же детерминировано, как и прошлое, тогда настоящее не может создать ничего нового. Структура причинных связей, составляющих физический мир, может тогда рассматриваться протяженной либо от бесконечного прошлого к бесконечному будущему, либо, если угодно, от положительной бесконечности к отрицательной. В таком случае человеческий разум выступает только как зритель, которому удалось наблюдать эту структуру с преимущественной точки зрения настоящего, которая скользит вдоль временнòй оси, не являясь, однако, ареной становления. Для причинного детерминизма, так же как и для философии Парменида, становления не существует.

Детерминизм механики Ньютона получил еще более ясное и точное выражение в четырехмерном пространственно-временнòм континууме Эйнштейна — Минковского. Три измерения пространства и одно измерение времени составляют четыре оси этого континуума, а физические события представлены в виде «мировых линий», подобно линиям на диаграммах. Настоящее время является только поперечным сечением этой диаграммы, и совершенно безразлично, где мы нанесем его. Оно является лишь точкой отсчета, подобно *году*, с которого мы ведем счет нашей эры. Структура пространственно-временнòго многообразия везде одинакова, в том числе и по отношению к обоим направлениям времени. Форма всех мировых линий в этом многообразии определяется математическими законами. Эта безвременнàя вселенная является четырехмерным бытием Парменида, в котором ничего не случается: «Цельное все, без конца, не движется и однородно... без перерыва одно». Течение времени является иллюзией. Становление — также иллюзия. Это способ переживания времени человеком, однако в природе ничто не соответствует этому переживанию.

---

[1] Лаплас, Опыт философии теории вероятностей, М., 1908, стр. 9.

Детерминистскую концепцию временного потока можно сравнить с событиями, которые мы наблюдаем в кино. Когда мы следим за увлекательной сценой, то ее будущий исход уже запечатлен на пленке фильма. Становление является иллюзией, поскольку не имеет никакого значения, из какой точки мы следим за событиями. То, что мы рассматриваем как становление, — только познание будущего, но оно не имеет никакого отношения к самим событиям. Следующий случай, который, как мне говорили, произошел в действительности, может проиллюстрировать эту концепцию. В кинематографической версии «Ромео и Джульетты» была показана драматическая сцена. В склепе лежит безжизненная Джульетта, а Ромео, который считает, что она умерла, поднимает кубок с ядом. В этот момент из публики раздается возглас: «Остановись!» Мы смеемся над тем, кто, увлекшись своими субъективными переживаниями, забыл, что время в кино нереально и является только развертыванием изображений, отпечатанных на киноленте. Но умнее ли мы этого человека, когда полагаем, что течение времени в нашей действительной жизни представляет собой что-то другое? Является ли настоящее время чем-то большим, нежели нашим познанием предопределенной сети событий, развертывающихся, подобно киноленте?

Верно, конечно, что зритель находится вне событий фильма, тогда как мы являемся частью событий в этом мире. Но если наше будущее определено так же, как то «будущее», которое запечатлено на киноленте, то какое значение имеет тот факт, что мы могли бы узнать о нем заранее? Мы не смогли бы изменить его, мы были бы только зрителями наших собственных действий, и что касается контроля над будущим, то мы находимся не в лучшем положении, чем тот человек, который хотел предостеречь Ромео от самоубийства.

Парадокс детерминизма и планирования действий — подлинный парадокс. В дальнейшем я проанализирую его в более точных терминах, здесь же достаточно указать, что если время не имеет направления, то планируемые действия представляются непостижимыми. Мы планируем завтра пойти в театр, но, очевидно, совершенно бессмысленно планировать поход в театр вчера. Почему мы проводим это разграничение? Мы отвечаем, что можем изменить не прошлое, а только будущее. Если мы

ходили в театр вчера, то мы это делали, а если не ходили, то и не делали, и безразлично, что бы мы ни думали об этом сегодня. Почему мы так не рассуждаем, когда речь идет о посещении театра завтра? В нашем поведении выражается уверенность, что время течет в одном направлении, потому что планирование предполагает течение времени. Может ли научный анализ подтвердить эту уверенность?

Критические замечания относительно потока времени, выдвинутые древними, можно преодолеть с помощью точного логического анализа понятия бесконечности. Возражения, вытекающие из детерминистской физики, опровергнуть гораздо труднее. Прежде чем вдаваться в детальный анализ этих проблем, рассмотрим другую историческую попытку построения теории времени, которая в отличие от более ранних попыток была предпринята с полным знанием достижений науки и намеревалась найти выход из дилеммы: «детерминизм или свобода». Эта попытка, в которой в современном виде возраждалась концепция Парменида относительно иллюзорности времени, осуществлена Кантом в его теории времени.

Историческая задача системы Канта состояла в логической формулировке тех положений, которые более ранние спекулятивные философы пытались выразить в сравнениях и в мистических афоризмах. Это верно не только по отношению к кантовской теории времени, но и по отношению к его теории познания вообще. Кант хотел охарактеризовать с помощью понятия *синтетического суждения a priori* тот вклад, который вносит в познание человеческий разум. Таким образом, он излагает в логической форме теорию, согласно которой разум правит физическим миром. Это — рационалистическая теория, представлявшая собой в платоновской теории идей, например, мистическую веру, а не логическое построение. Однако здесь не следует вдаваться в анализ всей теории Канта [1], поэтому мы ограничимся рассмотрением лишь его теории времени.

Различение, которое проводил Парменид, между безвременным бытием и иллюзорным течением времени вновь

---

[1] Общий анализ исторической роли Канта см.: H. Reichenbach, The Pise of Scientific Philosophy, Berkeley Calif. Univ. Calif. Press, 1951, p. 40.

появляется в философии Канта в виде различения между безвременностью вещей в себе и временем как субъективной формой феноменов (вещей являющихся). Конечно, Кант стал бы возражать против отождествления его субъективного времени с такой интерпретацией, в которой время рассматривается как иллюзия. Однако различие между этими двумя интерпретациями имеет только эмотивное значение. Если время субъективно, то логика рассуждений приводит к тому, что оно не описывает какого-либо свойства физического мира. Однако, возможно, термин «субъективное время» является более точным и более наглядным, нежели термин «иллюзорное время» [1].

Кант приводит нам обширную аргументацию в пользу своей концепции. Он говорит, что знание проистекает из взаимодействия между физической реальностью и человеческим разумом и что человеческий рассудок наклады-

---

[1] [При попытке проанализировать понимание природы времени в философии Канта сталкиваются с большими терминологическими трудностями, во-первых, потому, что Кант не был последовательным в употреблении таких главнейших терминов, как «субъективный» и «объективный», и, во-вторых, потому, что он употребляет эти термины в таком смысле, который существенно отличается от общепринятых значений. Автор придерживался общепринятых значений, хотя это иногда приводит его к таким формулировкам, против которых сам Кант, вероятно, стал бы возражать. Расхождение в употреблении терминов возникает главным образом из-за метафизики и теории познания Канта. Хорошо известно, что Кант различал две сферы: сферу ноуменов, или вещей в себе, и сферу феноменов, или вещей являющихся. Только последняя сфера познаваема, и Кант обычно ограничивал применимость терминов «субъективный» и «объективный» лишь этой сферой. Она является также и сферой «физической реальности» в кантовском смысле. Однако феномены представляют собой отчасти продукты человеческого сознания и, следовательно, субъективны в современном смысле. Вещи в себе, по Канту, непознаваемы и, согласно современному критерию верификации, высказывания о них, вообще не имеют никакого познавательного значения. Но в то же время в концепции Канта вещи в себе составляют реальный мир, который существует независимо от человеческого познания, и этот реальный мир, как бы мы сказали теперь, составляет объективную, или физическую, реальность. В этом смысле и употребляются автором термины «объективный», «субъективный» и «физическая реальность», когда он утверждает, что в философии Канта время и причинность субъективны и что познание подразумевает существование взаимодействия между разумом и физической реальностью. Таким образом, хотя автор и использует терминологию, которая отличается от терминологии Канта, это различие не влияет на точность характеристики точки зрения Канта. — *М. Р.*]

вает определенные формы на физические вещи, для того чтобы упорядочить их и сделать постигаемыми. Такими формами являются пространство и время. Они, в частности, называются *формами созерцания;* этот термин означает, что пространство и время есть способ, с помощью которого мы видим вещи, то есть они накладывают условия на наши восприятия вещей, не имея, однако, никакого значения для вещей вне их отношения к наблюдателю. «Поэтому... пространство и время суть лишь субъективные условия всех наших наглядных представлений», — пишет он в «Критике чистого разума» [1]. Это справедливо и для событий, происходящих в моем разуме, раскрываемых посредством «внутренних чувств». «Но если бы я сам или какое-нибудь другое существо могло бы наглядно представлять меня без этого условия чувственности, то те же самые ощущения дали бы знание, в котором не было бы представления времени, а следовательно, не было бы таких представлений и изменений» [2]. Течение времени, изменение не является свойством бытия, говорит Парменид, и Кант соглашается с ним, добавляя, что время — это только форма, в которой мы воспринимаем бытие.

Как пришел Кант к своей концепции? Несомненно, что мотивы, которыми он руководствовался, родственны мотивам Парменида. Подобно Пармениду, он хотел избежать потока времени. Из его отношения к бессмертию видно, что страх перед смертью определяет всю его философию, но это не единственный мотив. Он хочет также спасти свободу воли и возможность планируемых действий. Парменид не был заинтересован в действии, которое, по-видимому, не казалось для него чем-то важным. Но Кант желает свободы действий, и, возможно, даже не для того, чтобы дать человеку возможность для усовершенствования мира, а для того, чтобы сделать его ответственным за свои действия. Он хочет обосновать мораль, но как можно оценить действия человека с точки зрения морали, если он не свободен?

В отличие от древних Кант столкнулся с весьма совершенной детерминистской физикой. Ньютонова теория тяготения, воплощавшая в себе концепции пространства и времени, произвела на Канта громадное впечатление.

---

[1] Кант, Критика чистого разума, Петроград, 1915, стр. 57.
[2] Там же, стр. 52.

Он видел, что детерминизм создает трудности для идеи свободы человека, и искал выхода. Он полагал, что нашел путь спасения свободы, сделав время субъективным, то есть интерпретируя течение времени как форму, в которой мы, люди, переживаем реальность, которая сама не контролируется временем. Таким образом, он учил об *идеальности* времени. В своем отрицании реалистической интерпретации времени Кант повторяет позицию Парменида; его аргументация почерпнута из таких достижений науки, которых древние не могли даже представить себе. Действительно, если придерживаться детерминизма, как он сформулирован в лапласовском образе сверхчеловека, то время субъективно. Течения времени не существует, если будущее, так же как и прошлое, может предстать перед его взором. Кант правильно усмотрел, что детерминизм не оставляет другой альтернативы. Однако он полагал, что можно спасти свободу воли. Решение, которое он предложил, состоит в замене реальности временнóго потока реальностью свободы воли.

Для того чтобы понять решение проблемы свободы, предлагаемое Кантом, мы должны рассмотреть также и его интерпретацию причинности. Кант хорошо видел тесную связь между временем и причинностью. Он видел, что мы раскрываем временнóй порядок, рассматривая порядок причинных связей. Он аргументировал, что временнóй порядок физических событий не определяется только последовательностью восприятий. В самом деле, временнóй порядок событий, как мы их воспринимаем, может быть противоположен временнóму порядку их действительного существования. «...в одном лишь восприятии объективное отношение следующих друг за другом явлений остается еще не определенным»[1]. Например, хотя первый звуковой удар отдаленного взрыва предшествует световой вспышке, мы видим свет раньше, а звук слышим позднее. Для того чтобы установить временнóй порядок, мы должны знать причинный порядок.

Для Канта причинный порядок так же субъективен, как и время. Он является одной из тех категорий, с помощью которых мы упорядочиваем наши ощущения, и не выражает свойства вещей в себе. Поэтому физика может говорить нам только о том, как мир является нам, а не

---

[1] Кант, Критика чистого разума, стр. 145.

о том, как он существует независимо от наблюдателя. И детерминизм существует только для мира явлений, мир вещей в себе не подчиняется причинным законам.

Это довольно странная доктрина. Для того чтобы сделать течение времени субъективным, Кант должен был даже причинность сделать субъективной. Чтобы спасти свободу и мораль, он должен был принести в жертву физику как науку об объективных вещах, которая превращается для него в науку о вещах переживаемых. Он пошел дальше Парменида в освобождении бытия не только от потока времени, но также и от причинной детерминации.

Трудно понять, в каком смысле Кант утверждал, что его теория восстанавливает свободу. Когда мы действуем, то мы хотим изменить те события, которые существуют во времени, хотим изменить будущее. Но временные события, говорит Кант, подчинены причинности и детерминизму. Тогда, как мы можем повлиять на них? Предположение, что за воспринимаемыми вещами имеются другие вещи, не подчиняющиеся физическим законам, нам не поможет. Мы хотим изменять те вещи, которые *подчиняются* причинности и *постигаются* опытным путем, и мы используем причинные законы для того, чтобы контролировать их, прекрасно зная, что если вещи не подчиняются этим законам, то все наши попытки планировать будущее, безусловно, потерпят неудачу. Кому нужна такая философия, если она уклоняется от ответа на вопрос, что может сделать человек, ссылаясь при этом на другую сферу бытия, которую мы, однако, не можем контролировать? Кантовская философия субъективного времени и субъективной причинности стремится увильнуть от решения этих проблем. Она не решает парадокса свободы и детерминизма, не вносит ясности в наше переживание потока времени, она не может дать оценки, с помощью которой можно провести различие между прошлым и будущим, между сферой неизменяемого и тем, что, как мы полагаем, можно изменить.

Различие между прошлым и будущим, объективная интерпретация времени как процесса становления, а не только как формы человеческого опыта находят сильную поддержку со стороны здравого смысла. Не так уж легко согласиться с философией, которая рассматривает эти понятия как иллюзии. Однако философ не может разделять убеждения здравого смысла без критических огово-

рок. Проблему времени нельзя решить также и посредством апелляции к интуитивному знанию, которое говорит нам, что должен иметься процесс становления, допускающий возможность планируемых действий. Надежда на так называемую интуицию слишком часто вводила в заблуждение. Можно искать спасения как в сфере здравого смысла, так и в сфере метафизических спекуляций. Однако они не дают ответов, приемлемых для тех, кто стремится к непредвзятому подходу, обеспечиваемому комбинацией логического анализа с наблюдением.

Эти замечания справедливы также и в отношении попытки Бергсона, утверждавшего, что он создал философию, для которой становление является сущностью времени. Подобно своему древнему предшественнику Гераклиту, Бергсон является одним из тех, кто реагирует на поток времени с положительной эмоциональной позиции, кто ищет в изменении настоящий элемент жизни и чей оптимистический подход к проблеме течения времени не поколеблен угрозой смерти. Для Бергсона, как и для Гераклита, идея становления является эмоциональным вознаграждением.

Такой подход может оказать большую помощь в начале логического исследования, но он не может заменить его. Смелая апелляция Бергсона к интуиции, к «непосредственным данным сознания» не может дать обоснования теории времени. Бергсон утверждает, что физик неверно понимает время, что он «опространствует» его, трактуя как пространственное измерение. Действительное время, или «длительность», говорит он, можно понять лишь с помощью непосредственного знания, которое открывает нам, что время есть акт становления. Подобные аргументы немного могут дать тому, кто желает знать, может ли он доверять своей интуиции. Если человек находится под гипнозом, то хотя его интуиция и говорит ему, что он свободен, однако мы знаем, что это не так. Если интуиция говорит нам, что будущее порождается «творческой эволюцией», то мне хотелось бы знать не только то, насколько истинно это положение, но и что́ оно означает. Однако на этот вопрос нельзя отвечать при помощи новой апелляции и интуиции. Сущность становления может быть выяснена только при помощи логического анализа, и этот анализ должен основываться на всех наших знаниях о физическом мире, включая и нас самих.

Нельзя опираться только на непосредственное созерцание. Для решения проблемы времени не существует других способов, кроме методов физики. Физика гораздо более других наук связана с природой времени. Если время объективно, то физик должен установить этот факт, если имеется становление, то физик должен познать его, однако если время лишь субъективно и бытие безвременно, тогда физик должен иметь возможность игнорировать время в своем истолковании реальности и описывать мир без ссылок на время. Утверждение Парменида, что время является иллюзией, утверждение Канта, что время субъективно, и утверждение Гераклита и Бергсона, что все течет, недостаточно обоснованы. Они не учитывают того, что́ о времени говорит физика. Исследование природы времени без изучения физики — безнадежное предприятие. Если имеется решение философской проблемы времени, то оно зафиксировано в уравнениях математической физики.

Возможно, более точным было бы сказать, что решение следует искать между строк физических исследований. Физические уравнения формулируют специфические законы в наиболее общей форме, но философский анализ связан с высказываниями *об* уравнениях, а не с их содержанием. Как уже было сказано, вопрос о детерминизме тесно связан с проблемой времени, однако не существует такого физического уравнения, в котором формулировался бы принцип детерминизма. Вопрос о справедливости этого принципа есть вопрос о физических уравнениях, вопрос о том, можно ли на основе определенных уравнений делать строгие предсказания, охватывающие все возможные явления. По этим причинам философские исследования физики ведутся не на языке самой физики, а на метаязыке, который говорит *о* языке физики.

Хорошо известно, что современная квантовая физика отказалась от детерминизма классической физики. В связи с этим следует отметить, что наиболее важное значение приобретает исследование, какие последствия для решения проблемы времени будет иметь отказ от классического детерминизма. Однако существуют и другие проблемы, изучение которых должно предшествовать исследованию проблем квантовой механики. В процессах, которые изучает термодинамика, физика самым

непосредственным образом интересуется проблемой течения времени; поэтому утверждение, что физика якобы «опространствует» время, является следствием глубокого заблуждения. Специфическая природа времени, отличная от природы пространства, отражается в фундаментальных уравнениях физики. Мы увидим, что вклад квантовой механики в решение проблемы времени можно понять лишь после изучения термодинамического подхода к решению этой проблемы. Ниже будет показано, что физика может дать оценку потоку времени и становлению, что здравый смысл прав и что мы можем изменять будущее. Однако доказательство этих положений требует применения научного метода. Даже значение таких терминов, как «время» и «становление», может быть понято лишь таким здравым смыслом, который ассимилировал достижения научного мышления.

Глава II

ВРЕМЕННОЙ ПОРЯДОК В МЕХАНИКЕ

§ 2. Качественные свойства времени

Если мы желаем доказать, что время можно объяснить с точки зрения физики, то нужно прежде всего перечислить те свойства, которыми характеризуется специфическая структура времени в общей картине физического мира, определяющего нашу повседневную жизнь.

Начнем с того, что сразу разграничим *количественные* и *качественные* свойства времени. При измерении времени с помощью часов мы используем его количественные, или *метрические*, свойства. К таким измерениям относится определение временных отрезков равной длины, например определение продолжительности двух последовательных часов, а также определение одновременности, то есть доказательство того, что события, происшедшие в двух удаленных друг от друга точках пространства, характеризуются одним и тем же моментом времени. Теория метрических свойств времени подробно разработана в современной физике, в частности в теории относительности Эйнштейна и в логическом анализе этой теории, поэтому нам нет необходимости заниматься изучением данного вопроса в настоящем исследовании, так как мы можем отослать читателя к другим публикациям [1].

Нашей главной задачей будет исследование качественных, или *топологических*, свойств времени. По сравнению с метрическими свойствами это более фундаментальные свойства в том смысле, что они не зависят от

---

[1] См. H. Reichenbach, Bericht über eine Axiomatik der Einsteinschen Raum-Zeit-Lehre, в «Phys. Zeitschr.», 1921, Bd. 22, S. 683—686; «Axiomatik der relativistischen Raum-Zeit-Lehre», Braunschweig F. Vieweg Sohn, A. G., 1924 и «Philosophie der Raum-Zeit-Lehre», Berlin und Leipzig, W. de Gruyter Co., 1928 (рус.пер.: Рейхенбах Г. Философия пространства и времени. М.: УРСС, 2003).

способов измерения времени и при изменении этих способов остаются неизменными.

В них включаются те свойства времени, которые определяют его специфическую природу, отличную от природы пространства, и объясняют, как это было показано в предыдущем разделе, наше эмотивное отношение к нему. Поскольку мы находимся в самом начале нашего исследования, то пока не будем добиваться точности формулировок. Здесь мы ставим своей задачей собрать материал для дальнейшей работы, оставляя разработку более точных формулировок для последующих разделов.

Наиболее очевидные свойства времени можно сформулировать следующим образом:

Определение 1. Время движется от прошлого к будущему.

Это определение касается потока времени и выражает то, что мы называем *становлением.* Время не статично, оно движется. Когда мы говорим о потоке времени, то обычно рассматриваем его как движение некоторой объективной сущности, которое мы воспринимаем и непрерывное течение которой не можем предотвратить. Все-таки не исключена возможность, что поток времени каким-то образом связан со структурой человеческого сознания и, несмотря на свое возникновение из объективного источника, является совокупным продуктом объективных и субъективных факторов. Мы сможем дать ответ на этот вопрос только после более точного исследования физических истоков времени.

Определение 2. Момент «теперь» есть настоящее время, отделяющее прошлое от будущего.

Это определение довольно неточное. В одном смысле оно кажется тривиальной тавтологией, потому что термины «настоящее время» и «теперь» подразумевают одну и ту же вещь. В другом смысле оно представляется удивительным и загадочным. Какой смысл имеет слово «теперь»? Обозначает ли оно временну́ю точку, которая по своим физическим свойствам отличается от других точек времени? Или оно выражает только наш субъективный подход к времени, преимущественную точку, из которой мы рассматриваем время, сравнимую, так сказать, с точкой пространства, откуда мы видим окружающие нас вещи в определенной пространственной перспективе?

Однако мы можем выбирать центр пространственной перспективы, например обойти стол и смотреть на него с другой стороны. Можем ли мы выбрать «теперь»? Здесь наши возможности ограничены. Мы можем решить подождать до завтра, чтобы иметь другое «теперь», однако мы не можем сделать «вчера» нашим «теперь». И даже в отложенном, выбранном на завтра «теперь» мы не можем избежать того, что мы поступаем *так именно «теперь»*, что другое «теперь», так сказать, навязывается нам как момент, в который имеет место акт решения отложить предполагаемое «теперь».

Изложенные соображения ведут к следующим определениям, характеризующим прошлое.

Определение 3. Прошлое никогда не возвращается.

Это определение представляется тесно связанным с течением времени, то есть с определением 1. Оно выражает тот факт, что течение времени происходит в линейном направлении, которое нигде не пересекается с самим собой и может, таким образом, рассматриваться как прямая линия, простирающаяся от отрицательной бесконечности к положительной. Концепция времени как одномерного, или линейного, континуума вытекает из определения 3.

Хотя значение терминов «прошлое» и «будущее» представляется совершенно ясным, будет полезным поискать специфические различия между ними. Следующие три определения предназначаются для того, чтобы сформулировать эти различия.

Определение 4. Мы не можем изменять прошлое, но можем изменять будущее.

Первая часть определения представляется очевидной, однако вторая требует некоторого уточнения. Наша способность изменять будущее сильно ограничена. Нажав на тормоз, мы можем предотвратить столкновение нашего автомобиля с другим. Мы можем также превратить бесплодный участок земли в жилые кварталы. Но мы не в силах контролировать космические события, погоду или землетрясения. Мы также слабы в вопросах контролирования человеческого общества, которое продолжает дрейфовать от кризиса к кризису и от войны к войне. Точнее говоря, это определение означает, что имеются некоторые будущие события, которые мы мо-

жем контролировать, но нет таких событий в прошлом, которые мы могли бы изменить.

Можно было бы сформулировать и другое отличие: «будущее неизвестно, тогда как прошлое известно». Однако очевидно, что такое определение ошибочно. Некоторые будущие события, такие, как астрономические, или тот факт, что осенью состоятся общие выборы, нам хорошо известны. В то же время многие события прошлого неизвестны. Если бы они были известны, тогда задача геологов и историков была бы более простой. Различие между прошлым и будущим заключается не в различии между знанием и незнанием, а в различии методов, с помощью которых мы получаем это знание.

Наше знание о прошлом основывается на *протоколах*, будь то документы, написанные летописцем, окаменелости, содержащиеся в геологических слоях, или же следы крови на одежде. Напротив, говорить о протоколах будущего представляется абсурдным. Протоколы — это небольшие побочные продукты более значительных событий, общими следствиями которых является множество других событий, причем гораздо более важных. Летопись — это побочный продукт войны, во время которой погибли тысячи людей; окаменелости — продукт интенсивных геологических событий; следы крови на одежде — побочный продукт убийства. Если мы хотим предсказать будущее, то для этого недостаточно таких отрывочных данных.

События будущего могут быть предсказаны только на основании исчерпывающей информации, охватывающей данный случай в целом. Если мы хотим предсказать, будет ли война, то нам нужно знать состояние вооружения различных стран, их политические стремления и многое другое. Если мы хотим предсказать, скроется ли в будущем под водой определенная часть суши, то нам следует провести измерения изменений уровня гораздо более обширной области. Для того чтобы предсказать убийство, мы должны детально знать душевное состояние определенной личности.

По этой причине у нас есть регистрирующие инструменты только для событий прошлого и нет таковых для событий будущего. Мы видели, что действие этих инструментов подчиняется *частным* следствиям более значительных событий, однако таким следствиям, которые

позволяют нам сделать вывод, что произошло более крупное событие. Говорить же о регистрирующих инструментах будущих событий, по-видимому, бессмысленно. Рассмотрим, например, действие барометра-самописца. Он включает в себя сосуд, из которого выкачан воздух и стенки которого под давлением окружающей атмосферы поднимают или опускают иглу. Положение иглы отмечается на медленно вращающемся барабане. Возможно, что этот незначительный эффект был побочным продуктом урагана, валившего деревья и срывавшего крыши с домов. Несмотря на это, из отметки на барабане мы можем только сделать вывод, что давление воздуха в данном месте было таким-то. Напротив, если мы хотим предсказать, какое будет давление воздуха в данном месте завтра, то для этого недостаточно произвести измерения только в интересующем нас месте. Для того чтобы предсказать атмосферное давление на одном участке, нужно использовать метеорологические данные, полученные в обширной географической области, причем данные, содержащие гораздо больше информации, нежели простые показания барометра.

Предсказание будущего на основании отрывочных данных возможно только в исключительных ситуациях. Такие ситуации имеют место, когда определенные наблюдения дают возможность сделать вывод, что происходит более широкий процесс, следствия которого предсказуемы на основании имеющегося у нас прошлого опыта. Например, трещины в стене дома позволяют предсказать, что дом скоро разрушится. Правительственный приказ об эвакуации всех женщин и детей из определенной местности может послужить указанием на близость войны. В таких случаях общая причина, которая порождает предсказанное следствие, может быть выведена из нескольких отрывочных наблюдений. Эти данные сперва используются как протоколы, которые служат основанием для выводов относительно прошлого, относительно более широкой причины, а затем, используя знания об этой причине, делаются предсказания. Даже если такие предсказания и возможны, то их можно делать лишь в самых общих выражениях, и они никогда не дают подробной информации, которая достижима относительно запротоколированных событий. Может быть, мы в состоянии предсказать, что дом раз-

рушится, но мы не можем предсказать точных мест расположения обломков. Эту информацию позднее легко зафиксировать в протоколах. Возможно, мы в состоянии предсказать войну, однако мы не можем предсказать, кто будет убит. Эту информацию впоследствии нам также дадут протоколы. Даже те, кто настаивает на том, что такие предсказания, хотя и не достижимые на практике, в принципе возможны, согласятся с тем, что предвидение не может основываться на нескольких отрывочных данных. Это расхождение в деталях возникает из специфического различия между выводами, относящимися к будущему, и выводами, относящимися к прошлому. *Предсказания* требуют знания полной причины; *констатации* или утверждения о прошлых событиях могут быть основаны на частных следствиях, на протоколах.

Мы формулируем это различие следующим образом:

Определение 5. Мы можем иметь протоколы прошлого, но не будущего.

В конечном счете, мы можем сформулировать различие между прошлым и будущим, которое выступает теперь в виде суммы различий, выраженных в определениях 4 и 5.

Определение 6. Прошлое детерминировано, будущее недетерминировано.

Если бы нас спросили, что мы обозначаем терминами «детерминированный» и «недетерминированный», то мы, по-видимому, почувствовали бы некоторую неловкость. Мы могли бы аргументировать, что прошлое состоит из установленных фактов, тогда как будущее — нет, и объяснить, что установленный факт — это нечто такое, чего мы не можем изменить и что можно запротоколировать, ссылаясь, таким образом, на определения 4 и 5. В самом деле, представляется сомнительным, чтобы определение 6 шло дальше определений 4 и 5.

Это один из тех вопросов, которые мы должны будем исследовать. Однако имеются и другие вопросы, которые также следует рассмотреть. Мы уже сделали предварительный обзор качественных свойств времени и выяснили, что значение многих применявшихся нами терминов требует проверки.

Разве вполне ясно, что мы имеем в виду, когда формулируем тот очевидный факт, что мы можем изменять некоторые будущие события? Достаточно ли определено

слово «протокол»? И что мы подразумеваем под движением времени, на которое ссылались в определении 1?

Наш анализ времени связан, следовательно, с установлением точных значений. Это проблема *экспликации*, то есть мы должны сконструировать точные понятия с тем, чтобы поставить их вместо неопределенных понятий, использовавшихся до сих пор. В экспликации неточное понятие называется *экспликандом*, а точное — *экспликансом*.

Задача состоит в том, чтобы определить такой экспликанс, подстановка которого на место экспликанда оправдывала бы в контексте человеческого поведения употребление данного утверждения.

Проблема экспликации играет большую роль в исследованиях научной философии. Можно напомнить, что такие вопросы, как «Что такое вероятность?», «Что такое причинный закон?», «Что такое логическая необходимость?» и «Что такое разум?», являются проблемами экспликации, и к их философскому обсуждению следует подходить именно с этой точки зрения. Экспликация никогда не может быть строго корректной как раз по той причине, что экспликанд представляет собой неточное понятие, и мы никогда не можем сказать, полностью ли соответствует экспликанс его свойствам. Мы можем требовать только, чтобы экспликация была адекватной, то есть чтобы экспликация соответствовала, по крайней мере качественно, употреблению терминов в разговорном языке, и если экспликанс ставится на место экспликанда, то большинство предложений разговорного языка не должно менять своих значений истинности. Если это требование соблюдено, то мы можем рассматривать экспликацию как предложение использовать новые термины вместо старых. Такая замена поможет нам не только достигнуть точных значений, но также и сформулировать отношения, характеризующие понятия, которые в противном случае остались бы неизвестными. Поэтому экспликация есть метод, используя который, мы в конечном счете добиваемся успеха в понимании значения такого понятия, которое слишком неясно, чтобы быть доступным непосредственному пониманию.

Экспликация понятия времени и производных от него понятий, подобно направлению времени, течению вре-

мени и т. д., имеет особое значение в силу той важной роли, которую играет время как в повседневной жизни, так и в науках.

Мы убедимся, что экспликация времени требует изучения некоторых фундаментальных разделов физики и что она внесет ясность как в основы физики, так и в основы философии. Ибо до тех пор, пока не понята природа времени, философ не может утверждать, что он в состоянии дать оценку физической реальности или человеческому знанию о ней.

§ 3. Причинная теория времени

Каким образом мы можем найти соответствующую экспликацию времени? Ясно, что это можно сделать только при изучении причинных отношений. Тесная связь между временным порядком и причинными процессами очевидна из предшествующего рассмотрения. Целью данного исследования будет показать, что эта связь не является простым совпадением, что временно́й порядок *можно свести* к причинному порядку. Причинная связь есть отношение между физическими событиями, и ей можно дать объективную формулировку. Если мы *определяем* временно́й порядок с помощью причинных связей, то этим мы показываем, какие специфические свойства физической реальности отражаются в структуре времени, и тем самым даем экспликацию неясного понятия временно́го порядка. Мы покажем, в частности, что определениям 1—6 в § 2 можно, таким образом, придать точное значение, в котором они являются истинными. Однако эта процедура потребует отказа от некоторых обычных ассоциаций и приведет нас к вероятностной интерпретации причинности.

Идея сведения временно́го порядка к причинному впервые была выдвинута Лейбницем[1]. Схематический набросок его идеи отмечен печатью гения, предвосхитившего концепцию относительности пространства и времени. Однако причинную теорию времени нельзя было окончательно обосновать до тех пор, пока Эйнштейн не создал свою теорию относительности. Решающий аргу-

[1] См. H. Reichenbach, Die Bewegungslehre bei Newton, Leibniz und Huyghens, в «Kant-Studien», 1924, Bd. 29, S. 421.

мент в пользу определения временнòго порядка с помощью причинного следует из критики Эйнштейном понятия одновременности. Хорошо известно, что преобразования Лоренца, в которых выражается специальный принцип относительности Эйнштейна, позволяют производить обращение временнòго порядка некоторых событий, а именно тех, между которыми нельзя установить причинную связь. Временнòй порядок остается инвариантным при преобразованиях Лоренца только в том случае, если рассматриваемые события могут быть связаны посредством сигналов, то есть посредством причинных цепей. Из этого следует, что если временнòй порядок был бы чем-то большим, нежели причинным порядком, тогда нельзя было бы признавать справедливость преобразований Лоренца и теории относительности Эйнштейна. Это положение не нуждается в том, чтобы его особо выделяли. Кто отрицает причинную теорию времени, кто настаивает на том, что временнòй порядок имеет смысл независимо от причинного, тот вынужден отрицать физическое значение лоренцевых преобразований, и теория относительности становится для него игрой символов, большинство из которых не имеет никакого отношения к физической реальности [1].

В свете этих соображений автором в более ранних публикациях была разработана причинная теория времени [2]. Автор не намеревается здесь резюмировать полученные им выводы. В этих исследованиях отношение причины и следствия рассматривалось как наиболее простой термин и было показано, что временнòй и даже пространственный порядок можно свести к этому простому отношению.

В настоящей книге автор намеревается исследовать само отношение причины и следствия; то есть не будет рассматривать его больше как простое, а постарается свести к другим отношениям. В процессе нашего исследования будет показано, какие другие отношения потребуются для этого. Такой анализ даст нам возможность решить проблему направления времени, которая не может быть решена в рамках теории относительности Эйнштейна, потому что ее решение требует перехода от

---

[1] См. также стр. 34.
[2] См. сноску 1 на стр. 60—61.

строгих причинных отношений к вероятностным отношениям.

Следует добавить некоторые замечания логического характера. Дискуссия по проблеме времени во многом страдает от того, что путают два понятия, пренебрегая различием между порядком и направлением.

Точки, находящиеся на прямой линии, простирающейся в бесконечность в обе стороны, расположены в определенном порядке, однако сама линия не обладает направлением. Если точки расположены в линейном или последовательном порядке, то они подчиняются отношениям асимметрии и транзитивности. Относительно рис. 1

Рис. 1. Порядок точек на линии.

мы можем сказать, например, что A находится слева от B, тогда как B находится не слева, а справа от A. Этот факт выражает асимметрию отношения *налево от*. Отношение *направо от* называется конверсией отношения *налево от*. Кроме того, если A находится слева от B, а B — слева от C, тогда A находится слева от C; этот факт выражает транзитивность данного отношения. Из теорий отношений хорошо известно, что любое асимметричное, связное и транзитивное отношение устанавливает порядок последовательности. Когда мы говорим, что линия, хотя и упорядоченная в отношении последовательности, не имеет направления, то это означает, что не существует способа структурного различения между левым и правым, между одним отношением и его конверсией. Для того чтобы сказать, какое направление мы желаем называть «левым», мы должны указать на диаграмму или же дать наименования точкам и указать избранное направление, используя эти наименования. Если бы мы решили назвать «правым» то, что является «левым», или наоборот, то мы не заметили бы никакого различия в их структуре, то есть отношение *налево от* обладает теми же самыми структурными свойствами, что и отношение *направо от*.

Иначе обстоит дело для случая одномерного континуума отрицательных и положительных вещественных чисел, который мы можем нанести на прямую линию. Числа подчиняются отношению *меньше чем*, которое асимметрично, связано и транзитивно, подобно отношению *налево от*, следовательно последовательность чисел упорядочена. Однако, кроме того, отношение *меньше чем* обладает направлением, то есть его структура отличается от структуры обратного ему отношения *больше чем*. Это видно хотя бы из того, что мы можем провести различие между отрицательными и положительными числами следующим образом. Квадрат положительного числа есть число положительное, квадрат отрицательного числа также число положительное. Следовательно, мы можем сформулировать следующее утверждение относительно класса вещественных чисел: любое число, являющееся квадратом другого числа, больше, чем любое число, которое не является квадратом другого числа. Таким образом, мы определили отношение *больше чем*, а тем самым и отношение *меньше чем*, не прибегая к помощи диаграммы, а путем рассмотрения различий в их структуре.

Применяя полученные результаты при анализе проблемы времени, мы находим, что время, как нам обычно представляется, имеет не только порядок, но и направление. Отношение *раньше чем* рассматривается как отношение, идентичное отношению *меньше чем*, а не как ненаправленное отношение, подобно отношению *налево от*. Значит, мы считаем, что структура отношения *раньше чем* отличается от структуры своей конверсии, от структуры отношения *позже чем*. Главная задача нашего исследования состоит в формулировании этого структурного различия, ибо метод, используемый для чисел, конечно, не может оказать нам помощи в случае с временны́ми точками. Мы обозначаем числами временны́е точки лишь после того, как определили направление времени. С помощью различия между прошлым и будущим мы намереваемся выразить направление времени, и шесть определений, сформулированных в предыдущем параграфе, по-видимому, выражают не только порядок времени, но и направление. Остается только рассмотреть, насколько истинно это положение по отношению к каждому определению.

§ 4. Причинность в классической физике

Поскольку мы намереваемся определить время с помощью причинности, нам необходимо исследовать, обладает ли направлением или по крайней мере порядком отношение причина — следствие.

В повседневном опыте мы считаем доказанным, что отношение причина — следствие обладает направлением. Мы убеждены, что более позднее событие не может быть причиной более раннего. Но когда нас спрашивают, как отличить причину от следствия, мы обычно говорим, что из двух причинно связанных событий причиной является то событие, которое предшествует другому во времени. То есть мы определяем направление причинного отношения с помощью направления времени. Однако если нам необходимо свести время к причинности, то такая процедура недопустима, и мы, следовательно, должны изыскать методы для характеристики отношения причина — следствие без ссылок на направление времени. Давайте посмотрим, не найдем ли мы такой критерий в законах физики.

Точка зрения, согласно которой такой критерий существует, критиковалась при помощи следующих соображений. Законы физики устанавливают функциональные отношения. Они утверждают, что если определенные физические величины $x_1 \ldots x_n$ имеют некоторые значения, то этим определяется значение и другой величины x_{n+1}.

Таким образом, эти величины можно записать в следующем виде[1]:

$$x_{n+1} = f(x_1 \ldots x_n). \tag{1}$$

Вообще решение таких уравнений может быть найдено для любых переменных. Например, мы можем записать

$$x_1 = g(x_2 \ldots x_{n+1}), \tag{2}$$

где g есть функция от f, получаемая согласно правилам решения математических уравнений.

---

[1] Нумерация формул производится по параграфам. При ссылках на формулу, находящуюся в пределах данного параграфа, дается только номер формулы, но когда делается ссылка на формулу из другого параграфа, тогда номер параграфа предшествует номеру формулы. Например, ссылка (5.2) означает формулу (2) из § 5, а просто (2) означает формулу (2) из этого же параграфа.

Утверждалось, что в силу обратимости функциональных отношений с помощью физических законов нельзя определить направление или порядок. При этом говорилось, что функциональные отношения симметричны, то есть если y есть функция от x, то и x есть функция от y. В настоящее время нет сомнений в истинности этого утверждения; функциональные отношения действительно симметричны. Однако возникает вопрос, должны ли мы отсюда сделать вывод, что отношение причинности также симметрично. Возможно, что отношение причинности имеет такую природу, которую нельзя исчерпывающим образом охарактеризовать посредством более общего понятия функционального отношения.

Подобные концепции пользовались дурной репутацией, так как выдвигались в рамках метафизических философских систем. Философы, утверждающие, что наблюдение не может дать нам информацию о конечных законах физического мира, рассматривают причинность как отношение, имеющее априорную природу, то есть как отношение, которое обосновывается только путем теоретического мышления. Они утверждают, что функциональные отношения составляют только часть знаний, которые можно получить путем наблюдения, но разум добавляет к этому знанию априорный элемент, в котором причинность рассматривается как некая таинственная нить, связывающая следствие с причиной и делающая его продуктом последней. Для таких философских систем причинность является направленным отношением.

Я предоставляю другим ученым критику этой точки зрения, которая больше не принимается всерьез в научной философии. Однако я постараюсь разъяснить, что концепция направленной причинности не ведет с необходимостью к какой-то метафизике, что различие между причинными и просто функциональными отношениями вполне успешно можно провести и в рамках эмпирической философии, настаивающей на том, чтобы все утверждения, имеющие научный характер, были верифицируемы с помощью наблюдаемых. Точнее, причинность может представлять собой специфическое функциональное отношение, в котором учитывается определение направления. Выше мы использовали пример с соотношением *меньше чем* для вещественных чисел. В этом случае мы также определяли направленное отношение

с помощью функционального отношения, не прибегая к метафизическим представлениям.

Правда, существуют некоторые физические функциональные отношения, которые симметричны. Однако легко можно показать, что этот факт вполне совместим с утверждением, что причинность есть направленное отношение. Существует известная логическая теорема, согласно которой из отношения последовательности всегда можно вывести отношение симметрии. Подобным же образом мы определяем и отношения причинной связи.

Определение: Если A есть причина B, или B — причина A, или имеется событие C, которое является причиной A и B, то событие A *причинно связано* с событием B.

Это и есть отношение причинной связи, которое, очевидно, симметрично в точках A и B и которое мы используем в симметричных функциональных отношениях физики.

Рассмотрим, например, закон Бойля — Мариотта для идеальных газов:

$$p \cdot v = R \cdot T \frac{M}{m}. \qquad (3)$$

Он гласит, что давление (p), объем (v), молекулярный вес (m), масса (M) и температура (T) находятся в определенном функциональном отношении, для формулирования которого требуется константа R, не зависящая от природы газа. Справедливость этого закона сохраняется при изменении значения любой из этих величин, однако он не говорит нам о том, какое изменение является причиной других изменений, или о том, изменение какой величины вызывает изменение других величин. Когда мы экспериментируем с газами, мы очень хорошо знаем, что̀ является причиной наблюдаемых изменений, однако это известно нам не из отношений, сформулированных в формуле (3), а совсем из других соображений. Когда мы накачиваем автомобильную камеру, то мы знаем, что возрастание давления и массы воздуха является причиной увеличения объема. Когда поршень в цилиндре автомобильного мотора движется вверх, то уменьшение объема является причиной возрастания давления находящихся в цилиндре паров бензина, если при этом масса их остается постоянной.

В рассмотренных процессах численные значения этих величин контролируются формулой (3), однако с помощью этой формулы мы не можем определить, какой процесс здесь происходит. В этом уравнении формулируется только соотношение причинных связей, вопрос же о причине изменений остается открытым.

Многие физические законы, например закон сохранения энергии, закон Ома и т. д., представляют собой именно такие законы и выражают только отношения причинных связей, не определяя направление или порядок. Если нам потребуется найти направление отношения причина — следствие, то следует обратиться к другому типу законов.

Эти соображения вынуждают нас по-иному подойти к данной проблеме. Нам нужно найти такие физические законы, которые описывают физические процессы, а не только выражают наличие причинных связей. Такие законы существуют. Их можно подразделить на два различных вида: к первому виду относятся законы, описывающие механические процессы; ко второму — законы, описывающие термодинамические процессы. Различие между этими двумя видами хорошо известно; механические процессы обратимы, тогда как термодинамические — необратимы, за некоторыми исключениями. Попытаемся дать точное определение этого различия.

Происходящий во времени процесс Q представляет собой серию состояний, упорядоченных относительно направления времени. Когда мы описываем процесс в такой форме, то мы имеем в виду описание в положительном направлении времени. Мы можем сформулировать другое описание, используя отрицательное направление времени. В таком случае мы нумеруем временные точки в обратном порядке, или, другими словами, вводим новое время t', связанное с первоначальным временем t посредством отношения $t' = -t$. В описании процесса Q заменим теперь t на $-t'$, а полученное описание процесса будем называть обращенным описанием или описанием в отрицательном времени. Оно описывает тот же самый процесс Q, только другим языком. Очевидно, что обращенное описание мы можем сформулировать для любого процесса.

Теперь мы используем обращенное описание и придадим ему другое значение: мы рассмотрим его как

описание в положительном времени. Это легко сделать, поставив начальные обозначения. Возникает вопрос, существует ли процесс Q^*, который соответствует этому описанию, или, говоря точнее, совместим ли процесс Q^* с физическими законами. Если совместим, то мы можем сказать, что процесс Q обратим, и называем Q^* обратным по отношению к нему процессом. Если не совместим, тогда мы говорим, что процесс Q необратим, или что обратного процесса не существует. Обратные процессы являются, так сказать, зеркальным отражением первоначальных процессов во времени.

Рис. 2. Траектория шара, брошенного из A в C в направлении, указанном сплошной стрелкой. Пунктирная стрелка обозначает движение в обращенном времени.

Поясним наши рассуждения с помощью простого примера, для понимания которого требуется лишь элементарная математическая подготовка. Пусть процесс Q представляет собой движение шара, брошенного из точки A в точку C в направлении сплошной стрелки (рис. 2).

Движение шара описывается следующими уравнениями:

$$x_1 = a_1 t - \frac{g}{2} t^2, \quad x_2 = a_2 t, \tag{4}$$

где a_1 и a_2 — компоненты скорости, с которой шар брошен из A. Вводя отрицательное время t' при помощи соотношения $t = -t'$, мы находим:

$$x_1 = -a_1 t' - \frac{g}{2} t'^2, \quad x_2 = -a_2 t'. \tag{5}$$

Это описание процесса Q в отрицательном времени. Теперь сформулируем описание обратного процесса Q^*,

заменив в (5) t' на t:

$$x_1 = -a_1 t - \frac{g}{2} t^2, \quad x_2 = -a_2 t. \tag{6}$$

Здесь t, как и в (4), выражает положительное время. Очевидно, что процесс, описанный в (6), совместим с механическими законами и действительно существует; его можно проиллюстрировать с помощью примера, когда шар брошен из C в A, в направлении пунктирной стрелки. Таким образом, движение шара представляет собой обратимый процесс.

На основание обратимости механических процессов указывалось неоднократно: дифференциальные уравнения классической механики являются уравнениями второго порядка, в которых отсутствуют первые производные по времени. Если $f(t)$ есть решение такого уравнения, то $f(-t)$ также является решением, потому что при взятии второй производной происходит повторное умножение на $\frac{d(-t)}{dt} = -1$, вследствие чего результат остается неизменным. Например, как (4), так и (6) удовлетворяют дифференциальному уравнению

$$m \cdot \frac{d^2 x_1}{dt^2} = -mg; \tag{7}$$

в нем сформулирован известный закон Ньютона, согласно которому сила (правая часть уравнения) равна массе, умноженной на ускорение (левая часть уравнения). Знак минус в выражении силы появляется потому, что за положительное значение координаты x_1 мы принимаем то направление, которое противоположно действию силы тяготения.

Мы привели здесь столь подробное объяснение потому, что, имея дело с проблемой времени, необходимо пользоваться только очень точными понятиями. Обращенное описание можно сформулировать для любого процесса, однако это не означает, что в действительности для любого процесса существует процесс, обратный ему. Необратимыми процессами являются такие, например, как перемешивание газов или жидкостей, и процессы передачи тепла от нагретых тел к более холодным. Такие процессы характеризуются возрастанием энтропии, то есть тенденцией к исчезновению существующих различий. Об-

ратные им процессы запрещены, поскольку они представляли бы собой процессы, в которых энтропия уменьшается. Такие процессы, будут рассмотрены в § 7.

Обращенное описание, то есть описание в отрицательном времени, представляет собой удобное средство для анализа проблемы времени. Такое описание всегда можно сформулировать, поскольку описания в положительном и отрицательном времени эквивалентны [1], и бессмысленно спрашивать, какое из них истинно. Однако вопрос о существовании определенного обратного процесса не бессмысленный. Кроме того, утверждение, что положительным временем является время нашего субъективного опыта, истинно. Ниже мы проанализируем его.

Отрицательное время прекрасно иллюстрируется демонстрацией кинофильма в обратном направлении. Процессы, которые при таком способе демонстрации кажутся совместимыми с физическими законами, — обратимы, те же, которые кажутся несовместимыми, — необратимы. Механические процессы, подобные бросанию шара или подниманию стоящей на столе вазы, при демонстрации фильма в обратном направлении кажутся нам вполне естественными. Напротив, такие процессы, как горение сигареты, наливание сливок в кофе, битье посуды и т. д., покажутся неправдоподобными. Такие случаи представляют собой необратимые процессы.

При обсуждении проблем, связанных с временем, обычно утверждают, что только необратимые процессы придают отношению причинности асимметричность, тогда как обратимые приводят к симметричным причинным отношениям. Такие представления ошибочны. Направление времени могут определить только необратимые процессы, что же касается обратимых процессов, то они определяют скорее порядок времени и тем самым придают отношению причинности асимметричность. Отсылаем читателя к обсуждению отношения *налево от* (см. § 3). Правильным же является утверждение, что только необратимые процессы характеризуют однонаправленную причинность.

Наше утверждение, что законы механики, хотя и приводящие к обратимым процессам, *обусловливают*

---

[1] Обсуждение теории эквивалентных описаний см.: H. Reichenbach, The Rise of Scientific Philosophy, p. 133, 136, 147, 180, 266.

асимметричность причинности и тем самым определяют порядок времени, требует более подробного анализа, который мы и дадим в следующем параграфе.

§ 5. Определение временнóго порядка с помощью причинных отношений

Чтобы понять, какие свойства временнóго порядка можно определить при помощи обратимых процессов следует найти свойства, которые остаются инвариантными при обращении направления времени. Такие свойства действительно существуют. Они представлены теми отношениями порядка, которые выражаются словом

Рис. 3. Траектории двух шаров, брошенных из точки A в точки C и E в направлениях, которым соответствуют сплошные стрелки. Пунктирные стрелки обозначают движение в обращенном времени.

«между». Отношение «между» симметрично. Если B находится между A и C, то оно находится также между C и A. В силу этого обратимые процессы могут определять линейный порядок.

Мы видели, что движение шара от A к C, изображенное на рис. 2, не определяет направления времени, потому что движение в обратном направлении — от C к A — также совместимо с законами механики. Однако для обоих процессов момент прохождения шара через точку B следует искать между моментами прохождения его через точки A и C. Поэтому законы механики могут прекрасно информировать нас о временнýх отношениях «между». Они говорят нам, что существует причинная связь, идущая от A через B к C или от C через B к A, и, таким образом, оставляют неопределенным только направление этой связи.

Значение этого вывода станет более ясным, когда мы рассмотрим два процесса.

На рис. 3 изображены два шара, одновременно брошенные из точки A, один, как и раньше, — вправо, а дру-

гой — влево. Описание в положительном времени обозначено сплошными стрелками. В отрицательном времени оба шара будут лететь в противоположных направлениях, как указано пунктирными стрелками. Обе интерпретации совместимы с законами механики. Таким образом, мы знаем, что здесь имеются две линии причинных связей, которые обе либо начинаются, либо заканчиваются в *A*. Однако мы не можем интерпретировать дело таким образом, что один шар движется в направлении пунктирной стрелки, а другой — в направлении сплошной. Следовательно, хотя нам и не известны направления линий, но мы знаем все-таки отношение между направлениями. Мы знаем, что две линии нельзя свести в одну, ибо они *направлены в противоположные стороны*. Этот термин выражает свойство порядка, справедливое для обоих направлений.

Эти соображения требуют некоторых пояснений. Одновременное движение одного шара в направлении сплошной, а другого — в направлении пунктирной стрелки, конечно, совместимо с законами физики, однако оно не совместимо с теми наблюдениями, которые привели нас к двум предшествующим интерпретациям. В порядке пояснения этого вопроса допустим, что шары сталкиваются в *A*, тогда движение в точке *A* уподобится движению, изображенному на диаграмме рис. 4, которая дает нам, так сказать, крупный план. Предположим, что как полные, так и приближенные совпадения являются наблюдаемыми. В этом предположении выражается постулат непрерывности, согласно которому приближенные совпадения в пространственно-временной окрестности полного совпадения определяются путем наблюдения точно так же, как и полное совпадение. Относительно рис. 4 мы тогда можем сказать, что практические совпадения *1—4*, *2—5* и *3—6* являются наблюдаемыми, и остается нерешенной только проблема их временного порядка. Если, как показано на рис. 3, один шар движется в направлении пунктирной, а другой — в направлении сплошной стрелки, то мы будем наблюдать совпадения *1—6*, *2—5*, *3—4*, вопрос о временном порядке которых все еще остается открытым. Теперь посмотрим, как решается этот вопрос, когда с помощью наблюдения будет установлено, что произошли совпадения *1—4*, *2—5*, *3—6*, а

не какие-либо другие. Например, предположим, что в окрестности полного совпадения можно наблюдать направление линии, которая связывает центры шаров. Наблюдатель может смотреть на шары сбоку; если шар, находящийся в положении *1*, закрывает от него другой шар, то это означает, что другой шар находится в положении *4*, а не в положении *6*. Такие наблюдения столь

Рис. 4. Столкновение двух шаров, иллюстрирующее локальную сравнимость временного порядка.

же возможны, как и наблюдения совпадений. Действительно, определение координат x_1 и x_2 шара в определенном положении происходит подобным же образом, то есть на шар смотрят со стороны масштабной шкалы. Подобные наблюдения обязывают нас интерпретировать события либо в смысле сплошной стрелки, либо в смысле прерывной, однако они не допускают смешанной интерпретации.

Следует отметить, что данное различение проводится не с помощью законов, определяющих процесс механического столкновения. Если бы один шар двигался в направлении *1—2—3*, а другой — в направлении *6—5—4*, тогда законы столкновения привели бы к точно такой же диаграмме, как и та, которая показана на рис. 4. Однако в действительности процесс движения приводит к другим наблюдениям. Если наши рассуждения будут удовлетворять постулату непрерывности, это значит, что непрерывная протяженность от полного совпадения до практических (включая приближенные) совпадений будет неизбежной, в противном случае можно было

бы считать, что один и тот же элемент объема характеризуется двумя противоположными направлениями времени.

Другими словами, если два процесса происходят в непосредственной пространственно-временной близости друг от друга, то мы считаем, что сравнение их временных направлений возможно. Утверждение, что в данном примере оба процесса обладают одинаковым направлением во времени, рассматривается как верифицируемое с помощью наблюдения. Мы будем говорить здесь о *локальной сравнимости временного порядка*. Без такого предположения о непрерывности процессов понятию совпадение нельзя дать определение, поскольку мы всегда наблюдаем только практические совпадения и вводим полное совпадение лишь схематически.

Такие предположения обычно делаются при решении проблем механики. Одни дифференциальные уравнения не дают еще уравнения движения; кроме них, необходимо учитывать граничные условия. Применяя уравнения (4.4), мы бы сказали, исходя из предположения о равенстве начальных скоростей обоих шаров, что их движение выражается следующими уравнениями:

$$\begin{aligned} &\text{Шар } ABC: x_1 = a_1 t - \frac{g}{2} t^2, \quad x_2 = a_2 t; \\ &\text{Шар } ADE: x_1 = a_1 t - \frac{g}{2} t^2, \quad x_2 = -a_2 t. \end{aligned} \tag{1}$$

Если мы в обоих уравнениях заменим t на $-t$, то получим обратные уравнения:

$$\begin{aligned} &\text{Шар } CBA: x_1 = -a_1 t - \frac{g}{2} t^2, \quad x_2 = -a_2 t; \\ &\text{Шар } EDA: x_1 = -a_1 t - \frac{g}{2} t^2, \quad x_2 = a_2 t. \end{aligned} \tag{2}$$

Мы считаем, что оба решения совместимы с наблюдаемыми граничными условиями до тех пор, пока с помощью необратимого процесса мы не сможем выбрать одно из решений. Однако замена t на $-t$ только в одном уравнении приводит к решению, которое несовместимо с наблюдаемыми.

Эти соображения можно суммировать следующим образом. Ни законы механики, ни наблюдаемые не

дают нам направления времени, если это направление не определено заранее ссылкой на какой-нибудь необратимый процесс. Например, если рассматривать скорость тела как наблюдаемую, то ее направление следует устанавливать путем сравнения с некоторыми направленными во времени процессами, такими, как время психологического опыта, имеющее своим источником необратимые процессы, происходящие в человеческом организме. Однако если не используется никакого эталона, то мы не имеем права рассматривать скорость как наблюдаемую. Мы можем только вывести ее величину из других наблюдаемых, которые, однако, оставляют знак скорости неопределенным.

Так происходит в том случае, когда скорость вычисляется при помощи наблюдений за пространственными положениями. Например, траектория и начальная скорость пули могут быть вычислены, если даны по крайней мере три точки ее траектории, скажем пробоины в стенах. Скорость можно вывести также из наблюдения над центробежными силами. Но знак скорости остается неизвестным до тех пор, пока для интерпретации не привлекаются необратимые процессы. Смысл нашего утверждения состоит, таким образом, в том, что ни законы, ни наблюдаемые механики не позволяют нам отличать друг от друга противоположно направленные процессы. Однако можно установить относительный порядок двух механических процессов; в этом и состоит значение принципа локальной сравнимости временного порядка.

Определение противоположно направленных линий приводит к отрицательному ответу на вопрос о временном порядке событий, относящихся к различным линиям. Таким образом, мы делаем вывод, что момент прохождения через точку A не расположен между моментами прохождения через точки D и B. Прохождение через точку A совершается либо раньше прохождения через эти точки, либо позже. Мы видим, что с помощью законов механики можно установить наличие или отсутствие отношения временной промежуточности. Применяя подобные методы при рассмотрении более поздних событий, мы можем составить *причинную сеть*, которая в общем обладает *линейным порядком*. Этот термин означает следующее: если направление предписывается

одной линии, то этим определяется направление каждой линии. Если мы обратим направление для первой линии, то направления всех других линий также обратятся. На рис. 5 все направления указаны для одной интерпретации времени; для обратного времени все стрелки обратятся. Здесь важно то, что можно установить общий линейный порядок, используя обратимые механические процессы, если только придерживаться принципа локальной сравнимости временного порядка.

Рис. 5. Причинная сеть. Q и R обозначают события, для которых временной порядок не является определенным.

Мы подошли к формулированию очень важного свойства причинной сети, которое следует рассматривать как результат многочисленных опытов. Предположим, что мы имеем избранное направление для одной линии. Как нам известно, это направление определяет тогда все остальные. Теперь мы можем передвигаться вдоль линий, начиная от какой-нибудь точки и всегда следуя по направлению стрелок. Если из одной точки выходит несколько линий, то мы выбираем любую из них. Совокупность линий, проходимых подобным образом, можно назвать *причинной цепью*. Путешествуя вдоль причинных цепей, мы сделаем открытие, что никогда не сможем вернуться в исходную точку, или, другими словами, что *замкнутых причинных цепей не существует*.

Отсутствие замкнутых причинных цепей есть общее свойство сети; мы будем говорить, что сеть *незамкнута*. Представляется очевидным, что если это свойство справедливо для одного выбора направления линий, то оно точно так же будет справедливым, если обратить все направления. Это означает, что незамкнутость сети есть свойство, характеризующее порядок, а не направление.

Следует помнить, что свойство незамкнутости причинных цепей представляет собой эмпирический факт и не может рассматриваться как следствие логической необходимости. Если вообразить замкнутые причинные цепи, то в этом нет ничего противоречивого, хотя существование таких цепей привело бы к весьма странным событиям. Например, тогда могло бы получиться так, что человек может встретить свое прошлое *Я* и беседовать с ним, замыкая, таким образом, причинную цепь при помощи звуковых волн. Когда это событие происходит первый раз, то в этом случае данный человек будет моложе своего *Я*, когда же оно происходит вторично, то здесь человек встречается со своим *Я*, которое старше его[1].

Возможно, что более старшему *Я* будет трудно убедить свое более молодое *Я* в их тождестве, однако более старшее *Я* вспомнит, что когда-то давным-давно подобное событие произошло и с ним. А когда это молодое *Я* постареет и вторично испытает подобное столкновение, то оно со своей стороны будет пытаться убедить некое «третье» *Я* в их физическом тождестве. Такая ситуация представляется нам парадоксальной, но в ней нет ничего нелогичного. Однако, если бы такие события происходили, тогда бы не существовало и временного порядка в обычном смысле.

Более того, не существовало бы уникального тождества физического объекта во времени. Это *физическое тождество* вещи, именуемое также *генетическим тождеством* (genidentity), следует отличать от логического тождества. Любое событие логически тождественно с самим собой, однако когда мы говорим, что различные события являются состояниями одной и той же вещи, то

---

[1] См. H. Reichenbach, Philosophie der Raum-Zeit-Lehre, S. 167.

мы используем отношение генетического тождества, имеющее силу для этих событий. Физическая вещь представляет собой, таким образом, серию событий. Любые два события, принадлежащие к этой серии, называются генетически тождественными. Отношение генетического тождества является, следовательно, двухместной пропозициональной функцией, *которая симметрична, транзитивна и рефлексивна.*

Мы приписываем отношение генетического тождества только таким событиям, которые связаны между собой причинной цепью. Если причинные цепи незамкнуты, то отсюда следует, что любые два различных события, которые генетически тождественны, не являются одновременными. Этот вывод придает генетическому тождеству некоторое подобие с логическим тождеством: среди событий временного поперечного сечения состояния вселенной, которое представлено $t = \text{constant}$, нет двух таких событий, которые были бы генетически тождественны; генетическое тождество равноценно в этом случае логическому тождеству. Другими словами, если x и y генетически тождественны и одновременны, то они тождественны и логически. Этот вывод, согласно которому генетическое тождество среди одновременных событий является *единственным* отношением, кажется для нас настолько привычным, что мы склонны рассматривать его как логическую необходимость, однако он справедлив только для незамкнутых причинных цепей и будет *ошибочным в примере с человеком, встретившимся* со своим прошлым *Я*.

К счастью, наш мир не таков. Когда мы рассматриваем разные по времени состояния человека как состояния одной и той же личности, то в этом случае правомерность использования отношения генетического тождества гарантируется законом, согласно которому в данное время t личность может существовать только в одном виде. Однако не логика гарантирует этот закон. Отсутствие замкнутых причинных цепей, позволяющее нам дать это определение тождества, не прибегая к раздвоению личности, является физическим законом. Эти соображения показывают, какое влияние оказывает физическая природа причинных цепей на наши фундаментальные понятия. Говорить, что вещи и личности остаются тождественными и уникальными в потоке

времени, можно только в силу причинной структуры вселенной, в силу того, что причинная сеть незамкнута.

Имеем ли мы убедительное доказательство того, что эта сеть незамкнута? Нельзя сказать, чтобы данное положение вытекало из уравнений ньютоновой механики; оно является лишь обобщением нашего опыта в нашем пространственно-временно̀м окружении. Однако в течение долгого времени незамкнутый характер причинной сети не ставился под сомнение и обычно рассматривался как самоочевидное предположение физики. Такой некритический подход был поставлен под сомнение в общей теории относительности, допускающей космологические структуры с замкнутой причинной сетью. Тем не менее эта теория приводит к выводу, что для ограниченной пространственно-временно̀й области причинная сеть незамкнута. Здесь термин «ограниченный» может обозначать область космических размеров, и поэтому мы можем позволить себе рассматривать причинную сеть человеческой истории, а может быть, и истории нашей планеты, как незамкнутую.

Существование линейного временно̀го порядка[1] в нашем физическом мире имеет своим основанием незамкнутость причинной сети. Допустим, что направление линий определено путем предписания, тогда любые два события, которые связаны причинной цепью и *упорядочены* в *причинном отношении*, упорядочены также и во *временно̀м отношении*. Если A находится в начале цепи, а B — в конце, тогда A называется *более ранним* событием, чем B, а B — более поздним, чем A. Это определение однозначно. Если бы имелась другая причинная цепь, согласно которой A произошло позже, чем B, тогда эти два события составляли бы замкнутую причинную цепь (рис. 6), однако в силу незамкнутости причинной сети существование такой причинной связи исключено.

---

[1] Мы проводим различие между *линейным* порядком сети и *линейным* порядком событий, расположенных на одной линии. Когда мы говорим о *линейном* порядке времени, то полагаем, что события во *всех* линиях могут быть пронумерованы при помощи соответствующего определения одновременности. Таким образом, эти номера моментов времени составляют одномерные последовательности, простирающиеся из отрицательной бесконечности в положительную.

Остаются события, для которых временнòй порядок не определен, такие, как события Q и R на рис. 5. Если можно показать, что причинная цепь, которая могла бы связать эти события, не только случайно отсутствует, но и физически *невозможна*, тогда такие события можно рассматривать как одновременные. Они могут быть названы событиями, *временнóй порядок которых является неопределенным*[1].

Рис. 6. Две причинные линии, составляющие замкнутую цепь. Такой конфигурации событий в действительности не существует.

Описанный временнòй порядок является линейным, но не направленным. Мы можем обратить направление причинно упорядоченных событий, тогда A будет рассматриваться как более позднее событие, чем B. Но в этом случае взаимные отношения между событиями, временнòй порядок которых является неопределенным, не меняются. Поэтому порядок незамкнутой сети выражает наиболее существенные свойства нашего линейного временнòго порядка, кроме направления.

Таким образом, мы пришли к важному выводу относительно определения 3 § 2. Тот факт, что прошлое никогда не возвращается, есть следствие незамкнутости причинной сети, ибо если бы прошлое возвращалось, тогда бы мы имели замкнутую причинную сеть. Однако если обратить направление времени, то ситуация не изменится. То, что было бы прошлым для этого направления времени, также никогда бы не вернулось. Мы видим, что определение 3 § 2 характеризует не свойство направления времени, а только свойство порядка, именно свойство незамкнутости причинной сети.

Существование событий, между которыми невозможна никакая причинная связь, предполагалось уже в классической физике. Это предположение составляет смысл понятия одновременности еще в физике XIX столетия: одновременные события свободны от причинного взаимодействия, поскольку причинное воздействие

---

[1] Подробное обсуждение этого термина и проблемы одновременности см. H. Reichenbach, op. cit., § 19, 22.

требует времени для распространения от одной точки к другой. Вводя принцип, согласно которому свет представляет собой сигнал, распространяющийся с наибольшей скоростью, Эйнштейн показал, что нельзя дать строгое определение термину «одновременный» путем одного лишь исключения возможности причинного взаимодействия. Пусть P и P' будут двумя пространственными точками, мировые линии которых обозначаются вертикальными линиями (рис. 7). В данном случае эти линии обозначают и временные координаты. В момент t_1 световой сигнал покидает P и достигает P' в момент t_2. Отраженный в P', он возвращается в точку P, которой достигает в момент t_3. Поскольку распространение света является наиболее быстрым процессом установления причинной цепи, то для всех событий, или временных точек, происходящих в P в промежутке между t_1 и t_3, возможность существования причинной связи с событием, происходящим в момент t_2 в точке P', исключена. Следовательно, любое из этих событий может быть названо одновременным с последним событием. Это и есть формулировка принципа *относительности одновременности*. Данный принцип приводит к заключению, что t_2 может быть любой величиной в пределах между t_1 и t_3, то есть что величина t_2 определится следующим соотношением:

Рис. 7. Световой сигнал, выходящий из точки P в момент времени t_1, достигает точки P' в момент времени t_2, где он отражается и вновь возвращается в точку P в момент времени t_3. События, расположенные во временном интервале между t_1 и t_3 в точке P, не могут вступить в причинное взаимодействие с событием, происшедшим в точке P' в момент t_2.

$$t_2 = t_1 + \varepsilon(t_3 - t_1), \qquad 0 < \varepsilon < 1. \tag{3}$$

Любой выбор величины ε в пределах интервала между 0 и 1 обеспечивает допустимое определение одно-

временности. Значение $\varepsilon = 1/2$ имеет преимущество в некоторой системе координат и используется в преобразованиях Лоренца для такой системы, в которой P и P' находятся в состоянии покоя. Однако данная величина ε не определяет какой-то «истинной» одновременности.

Эти отношения часто изображаются при помощи диаграммы, подобной той, которая изображена на рис. 8.

Рис. 8. Временной порядок в мире Эйнштейна — Минковского.

От события A конус его возможных причин простирается в прошлое, а конус возможных следствий — в будущее. Пространственный объем V_1 содержит в себе все события, которые в более ранний момент t_1 могут причинно воздействовать на событие A; точно так же объем V_3 включает все возможные следствия A в момент t_3. Слева и справа расположены области событий, которые не могут вступить в причинное взаимодействие с A и, следовательно, не обладают определенным временным порядком относительно A. Настоящее время события A представляет собой любое поперечное сечение, проходящее через точку A и лежащее в пределах этих областей, как показано горизонтальной и наклонной пунктирной линиями, проходящими через точку A. Событиям, распо-

ложенным на одной из этих линий, приписывается один и тот же момент времени t_2, зависящий от того значения величины ε, которое используется в формуле (3). Горизонтальная линия соответствует значению $\varepsilon = 1/2$. События A и B составляют *времени-подобный* интервал, а события A и C — *пространственно-подобный* интервал.

Структура причинных отношений, которая предполагается в теории относительности, тождественна структуре причинной сети, то есть она определяет порядок, а не направление времени. Описание мира в отрицательном времени, то есть взаимная перемена конусов прошлого и будущего на рис. 8, не имеет никакого значения для законов, сформулированных в теории относительности. Это можно показать с помощью преобразований Лоренца:

$$x = \frac{x' - vt'}{\sqrt{1 - v^2/c^2}}, \quad t = \frac{t' - \frac{v}{c^2}x'}{\sqrt{1 - v^2/c^2}}. \tag{4}$$

Если мы заменим в этих уравнениях t на $-t$, а t' на $-t'$, то получим точно такие уравнения, как и прежде, за исключением того, что в числителе минус изменится на плюс. Это означает, что мы получим процесс, в котором скорость v имеет противоположное направление. Обращение времени изменяет только направление движения и приводит, таким образом, к процессам, которые подчиняются тем же самым преобразованиям.

Выше было показано[1], что теория относительности поставила вопрос о необходимости причинного определения времени. Это верно, однако она не требует направленного времени. Она основывается только на предположении об упорядоченном времени. Но если бы временной порядок представлял собой нечто большее по сравнению с причинным порядком, то временной порядок этой теории был бы несостоятельным. Различие между времени-подобными и пространственно-подобными интервалами в мире Эйнштейна — Минковского идентично с различием между причинно упорядоченными событиями и событиями, временной порядок которых является неопределенным. Поэтому замечание, что теория

---

[1] См. стр. 60.

построен таким образом, что оставляет путь AB неизменным; следовательно, изменяется только вторая часть пути. Можем ли мы придумать другой эксперимент, в котором неизменной окажется только вторая часть пути, идущая от B к C? Это легко осуществить (рис. 11). Мы бросаем мяч от E к B, не помещая на его пути ракетку, тогда мяч летит прямо к C. Здесь мы осуществляем второй вариант опыта без применения ракетки, однако в

Рис. 9 (слева). Траектория теннисного мяча, отраженного ракеткой.

Рис. 10 (в центре) и рис. 11 (справа). Различные траектории теннисного мяча, не отраженного ракеткой.

этом варианте вторая часть пути остается неизменной, тогда как первая изменена.

Мы утверждаем, что последний эксперимент не может быть использован для изучения вмешательства. Здесь мы при помощи *альтернативной причины* воспроизвели то же самое следствие, а именно — полет мяча от B к C; следовательно, этот эксперимент доказывает только существование альтернативной причины.

Эти соображения приводят, однако, к важным заключениям. Процесс, показанный на рис. 9, сам по себе не говорит нам о том, функционирует ли здесь ракетка как вмешательство или же как альтернативная причина. Эти два термина характеризуют отношения между двумя экспериментами. Когда мы сравниваем процесс, изображенный на рис. 9, с процессом, изображенным на рис. 10, мы говорим, что теннисная ракетка совершает акт вмешательства; когда же мы сравниваем его с процессом, изображенным на рис. 11, то говорим, что теннисная ракетка действует как альтернативная причина.

Кроме того, мы видим, что на вопрос: «Что случится, если не будет пущена в ход теннисная ракетка?» — нельзя дать ответа до тех пор, пока не будут сделаны

дальнейшие уточнения. Могут произойти процессы, изображенные либо на рис. 10, либо на рис. 11.

Условное противодействие факту: «Если теннисная ракетка не ударит по мячу, то...» — двусмысленно; протокол не может быть заполнен до тех пор, пока мы не скажем, какой из отрезков траектории полета мяча мы желаем сохранить неизменным — AB или BC. Однако если рассматривать эти отрезки с точки зрения события B, то они будут представлять соответственно прошлое и будущее. Когда мы говорим, что удар ракетки изменяет будущую траекторию мяча, то в нашем утверждении имеется молчаливое предположение: «Если прошлое остается неизменным». Не удивительно, что акт вмешательства изменяет только будущее и не изменяет прошлого; термин «вмешательство» определяется условием, что прошлое неизменяемо. Утверждение, что акты вмешательства не могут изменить прошлого, есть не что иное, как тавтология. Эти соображения приводят нас к выводу, что с помощью актов вмешательства нельзя определить направление времени. Термину «вмешательство» можно дать определение только после того, как дано определение направлению времени. Актами вмешательства являются те, которые оставляют прошлое неизменным.

Этот вывод становится более очевидным, когда мы описываем процессы в отрицательном времени. Тогда термины «вмешательство» и «альтернативная причина» меняются своими местами. То, что в положительном времени называлось вмешательством, становится в отрицательном времени альтернативной причиной, и наоборот. Это показано на рис. 9—11. Если рис. 9 интерпретируется с помощью понятия отрицательного времени, тогда мяч летит от C через B к A. Сравнивая рис. 9 с рис. 11, мы увидим, что удар ракеткой следует назвать вмешательством, а сравнивая рис. 9 с рис. 10, его следует назвать альтернативной причиной. В отрицательном времени мы бы также сказали, что можем изменить не прошлое, а будущее: вместо того чтобы продолжать свой полет в E, мяч из-за вмешательства ракетки вынужден лететь от B к A.

Этот результат приводит к интересным критическим замечаниям относительно определения 4, § 2, согласно которому мы можем изменить будущее, но не прошлое. Поскольку вмешательство осуществляется в обратимые

процессы, постольку это определение, подобно определению 3, § 2, не выражает направления времени. Если термин «прошлое» определен в терминах данного направления времени и, следовательно, является *относительным*, то определение 4, § 2 истинно, как всякая тривиальная тавтология, и справедливо для обоих направлений времени. Однако то, что мы называем прошлым в положительном времени, может быть изменено в отрицательном, потому что для отрицательного времени оно является будущим. Таким образом, если значение слова «прошлое» ограничивается тем интервалом, который оно обозначает в положительном времени, и поэтому употребляется в абсолютном смысле, тогда мы можем изменять прошлое, используя отрицательное время. Следовательно, в этом смысле определение 4, § 2 ошибочно.

Определение 4, § 2 можно интерпретировать еще и в третьем смысле. Будем считать, что оно имеет следующее значение: продвигаясь вдоль причинной цепи, мы не можем оказывать влияние на уже пройденные участки, поскольку такое влияние представляло бы собой замкнутую причинную цепь. В данной интерпретации определение 4, § 2 истинно; однако оно снова определяет не направление, а порядок, поскольку оно истинно, независимо от того, двигаемся ли мы вдоль причинной цепи в положительном или отрицательном направлении.

Для иллюстрации этого вывода дадим описание игры в теннис в отрицательном времени. Предположим, что мы берем киноленту, на которой зафиксирована игра в теннис, и прокручиваем ее в обратном направлении. Мяч летает туда и сюда, как и в обычном времени. Единственным отличием будет лишь то, что каждый раз мяч принимается слета, то есть до удара ракеткой он не касается земли, и в свою очередь перед тем, как перелететь обратно через сетку, он может либо удариться о землю, либо нет. Движение ракетки представляется причиной возвращения мяча. Мы бы сказали так: если бы ракетка не ударила по мячу, он продолжал бы полет в прежнем направлении. Однако это утверждение имеет другой смысл, чем в положительном времени. Пусть диаграмма на рис. 12 представляет собой вертикальный план траектории полета мяча. В положительном времени мяч летит в направлении *ABCD*, в отрицательном — в направлении *DCBA*. Для положительного времени мы говорим: если бы

ракетка не ударила мяч, он полетел бы от C к E. Напротив, для отрицательного времени мы говорим, что он полетел бы от C к F. В обоих случаях вмешательство с помощью ракетки изменяет тот участок траектории, который мы называем будущим.

Подача мяча в обратном направлении будет выглядеть еще более странно. Партнер посылает ему мяч, который, перед тем как пролететь над сеткой, ударяется о

Рис. 12. Траектория теннисного мяча, перелетевшего через сетку и отраженного ракеткой.

землю. Подающий игрок отводит руку с ракеткой назад до тех пор, пока о нее не ударяется мяч. Удар наносит рука, все время движущаяся в обратном по сравнению с ударом направлении. Мяч, отскочив от ракетки, движется прямо вверх и затем падает в левую руку игрока. Несмотря на необычное течение событий, движение мяча точно так же, как и в обычной интерпретации, управляется законами механики. Само механическое движение теннисного мяча не говорит нам о том, что фильм крутится в обратном направлении.

Конечно, странно наблюдать, как игрок бегает спиной вперед. Такое движение, хотя и совместимое с законами механики, является непривычным, поскольку безопасность движения обеспечивается визуальным контролем. Таким образом, для решения нашей проблемы требуется рассмотрение необратимых процессов, чем мы и займемся в следующей главе.

Глава III

НАПРАВЛЕНИЕ ВРЕМЕНИ В ТЕРМОДИНАМИКЕ И МИКРОСТАТИСТИКЕ

§ 7. Второй закон термодинамики

Хотя необратимые процессы являются обычными для нашего повседневного опыта (напомним, что к ним относятся такие процессы, как горение, рост живых организмов, перемешивание газов или жидкостей), однако до XIX века они не подвергались научному анализу. На основании открытых Карно (1824) законов, управляющих тепловыми процессами циклической формы, именуемыми также циклами Карно, Р. Э. Клаузиус (1850) и У. Томсон (1851 лорд Кельвин) сформулировали второй закон термодинамики — принцип, который дает возможность выразить в математической форме направление течения физических событий. В самом названии принципа проводится аналогия с первым законом термодинамики, который более известен как закон сохранения энергии. Оба эти закона определяют изменение физических величин в течении процессов, охватывающих тепловые явления. Первый закон постулирует, что во всех изменениях существует определенная величина, именуемая *энергией*, которая остается постоянной. Второй закон в его классической формулировке гласит, что существует другая величина, именуемая *энтропией*, которая при одних изменениях остается постоянной, а при других — возрастает, но никогда не уменьшается. Необратимыми процессами являются те, в которых происходит возрастание энтропии.

Эта формулировка делает очевидным параллелизм обоих принципов, однако она объясняет нам также, почему только второй закон позволяет сделать вывод о направлении физических процессов. Если состояние A переходит в состояние B и при этом энтропия остается постоянной, то при обратном переходе от B к A она также останется постоянной. Однако если во время перехода от

A к *B* энтропия возрастает, то есть энтропия в состоянии *B* больше энтропии в состоянии *A*, тогда обратный переход характеризовался бы уменьшением энтропии. А если утверждение, что энтропия никогда не уменьшается, является законом физики, то это значит, что обратного перехода не существует. Направление физических процессов выражается при помощи функции состояния *S* — энтропии, которая имеет определенное численное значение для каждого данного состояния и располагает физические состояния в порядке возрастания этой величины.

Конечно, при формулировке этих принципов следует учитывать, что все процессы, о которых идет речь, замкнуты, то есть все изменения происходят в пределах рассматриваемой системы. Энергия может передаваться из одной системы в другую, и в этом случае одна система теряет энергию, а другая получает ее. Однако общая сумма всех изменений равняется нулю. Энтропия точно так же может передаваться из одной системы в другую, при этом одна система теряет энтропию, а другая приобретает ее. Однако сумма всех изменений в уровнях энтропии равна нулю только при обратимых процессах. При необратимых процессах она всегда имеет положительную величину. Это и есть формулировка второго закона термодинамики.

Энтропия представляет собой величину, определяющуюся с помощью термодинамических параметров, таких, как температура и давление. Механические параметры, например координаты и скорость, не входят в определение ее величины, то есть понятия механической энтропии не существует. Для этого имеются следующие основания: если в системе происходит чисто механический процесс, то есть процесс, в котором не изменяется ни температура, ни количество тепла и т. д., то такой процесс обратим, следовательно, энтропия системы должна оставаться постоянной. Другими словами, изменение механических параметров не оказывает влияния на величину энтропии. Энтропия системы изменится только в том случае, если, кроме механических, происходят еще и термодинамические изменения, такие, как порождение теплоты в процессе трения. Понятие механической энтропии излишне в силу обратимости механических процессов.

Однако, несмотря на то, что бо́льшая часть термодинамических процессов относится к необратимым, суще-

ствуют также и некоторые обратимые процессы, которые сопровождаются температурными изменениями. По крайней мере такие процессы можно рассматривать как предельные случаи необратимых процессов. Если две системы вступают во взаимодействие, а разность в температурах этих систем очень мала, тогда тепло из более теплой системы будет переходить в более холодную довольно медленно. Этот процесс представляет собой один из тех случаев, когда система проходит через непрерывную серию состояний, которые можно рассматривать как состояния равновесия. В таких процессах энтропия одной из систем увеличивается ровно настолько, насколько уменьшается энтропия другой системы, а общая сумма изменений энтропии равна нулю.

Эти соображения приводят к двум значениям слова «обратимый». В более узком смысле оно распространяется в целом на замкнутую систему, энтропия которой остается постоянной. В этом смысле обратимость определяется только для замкнутых систем: процесс, происходящий в такой системе, является обратимым, если он может совершаться также и в противоположном направлении, в то время как вся система остается замкнутой. В более широком смысле процесс, происходящий в части такой системы, также можно называть обратимым. Если термин «обратимость» употреблять в этом смысле, то его можно применять и по отношению к незамкнутым системам. В этом случае мы получаем парадоксальный вывод, что при обратимых процессах энтропия системы может либо увеличиваться, либо уменьшаться. Парадокс, который озадачивал многих читателей учебников по термодинамике, решается довольно просто, когда мы узнаем, что здесь слово «обратимый» употребляется в широком смысле. Однако представляется целесообразным избегать употребления этого термина в его широком значении и говорить об обратимости только по отношению ко всему процессу, который протекает в замкнутой системе.

Обратимые процессы, при которых происходит изменение термодинамических условий, представлены циклом Карно. Если такой процесс происходит в одном направлении, то теплота переносится от нагревателя к холодильнику при помощи газа, приводящего в движение поршень в цилиндре и производящего механическую энергию. Если процесс протекает в обратном направлении, то расходуется

механическая энергия и теплота переносится из холодильника в нагреватель. В первой форме процесс используется в паровой машине, во второй — в холодильных установках. Энтропия всей системы в целом, включая и нагреватель, и холодильник, и газ-теплоноситель, в течение цикла остается постоянной. Изменения в величине энтропии в составных частях систем происходят периодически таким образом, что когда увеличивается энтропия газа-теплоносителя, то уменьшается энтропия нагревателя и холодильника, и наоборот.

Строго говоря, обратимых процессов не существует. Если мы хотим обратить цикл Карно, нам следует затратить несколько больше механической энергии, чем та, которая была получена, когда цикл осуществлялся в первоначальном направлении. Однако чем медленнее протекает процесс, тем меньше разница между этими двумя количествами механической энергии. В этом и состоит смысл нашего утверждения, что обратимые процессы можно определить как предельные случаи реальных процессов. Это положение остается в силе также и для обратимых процессов механики. Во всех реальных механических процессах происходит порождение теплоты в результате трения. Полностью обратимых процессов не существует.

Если энтропия S изменяется в течение процесса, представляющего собой часть обратимого процесса (то есть частный процесс, являющийся обратимым в широком смысле), и если dQ обозначает количество теплоты, подводимой в систему или отводимой из нее, а T — абсолютную температуру, то изменение энтропии удовлетворяет следующему уравнению:

$$dS = \frac{dQ}{T}. \qquad (1)$$

Здесь $dS \lesseqgtr 0$ в зависимости от $dQ \lesseqgtr 0$. Если для рассматриваемого процесса $dS > 0$, то для остального процесса в целом мы имеем $dS < 0$. Можно показать, что интеграл выражения (1) не зависит от пути интегрирования, и, следовательно, энтропия обладает свойствами термодинамического потенциала. Это обстоятельство можно использовать при вычислении энтропии, или, вернее, разности в величинах энтропии, для таких явлений, которые нельзя рассматривать как элементы обратимых процессов.

Если в термос, наполненный теплой водой, бросить кусок льда, то мы положим тем самым начало процессу движения замкнутой системы к состоянию теплового равновесия. Сам этот процесс необратим, кроме того, он не является также частью какого-то обратимого процесса. Энтропия всей системы в целом будет медленно возрастать. Здесь вся система в целом не удовлетворяет соотношению (1), поскольку $dS > 0$, а $dQ = 0$. Однако можно связать первоначальное и конечное состояние при помощи другого процесса, который является незамкнутым и представляет собой часть обратимого процесса. Для этого процесса формула (1) остается справедливой, и с ее помощью можно, интегрируя по dS, подсчитать разность в величинах энтропии. Так, окольным путем можно определить значение энтропии для необратимых процессов, не являющихся составными частями обратимых процессов, с точностью до аддитивной постоянной. Все необратимые процессы, происходящие в замкнутых системах, относятся к этой категории.

С помощью подобных соображений Р. Э. Клаузиус и У. Томсон получили выражение энтропии S в термодинамических параметрах. Например, объем газа v, масса которого равняется единице, имеет энтропию (с точностью до аддитивной постоянной)

$$S = c_v \ln T + \frac{R}{m} \ln v, \tag{2}$$

где c_v — удельная теплоемкость при постоянном объеме, T — абсолютная температура, m — молекулярный вес газа и R — универсальная газовая постоянная. Если даны значения этих параметров, то мы можем высчитать величину S.

Если данная система получена в результате слияния двух других систем, то ее энтропия равна сумме энтропий этих систем:

$$S = S_1 + S_2. \tag{3}$$

Рассмотрим, например, тепловое взаимодействие двух систем, которые представляют собой равные количества одного и того же газа, обладающие разной температурой и находящиеся в двух смежных сосудах. Тепло от газа с высокой температурой переходит через стенку к газу с низкой температурой. Согласно уравнению (2), энтро-

пия первого газа будет уменьшаться, а второго — возрастать. Это происходит потому, что в формулу (2) температура включается в виде логарифмической функции возрастающей с увеличением аргумента. Температуры обеих систем в конце концов достигает среднего значения, однако логарифм среднего больше среднего логарифмов, и суммарное значение (3), подсчитанное после увеличения температуры, будет больше соответствующего суммарного значения в начале процесса.

Рассмотрим другой пример. Предположим, что газ находится в двух сосудах, сообщающихся между собой. Мы перекрываем кран и откачиваем газ из одного сосуда. Если мы теперь откроем кран, то газ потечет в пустой сосуд, заполняя, таким образом, больший объем. После достижения состояния равновесия температура не меняется, поскольку не было проделано никакой внешней работы. Из уравнения (2) мы видим, что если объем V увеличивается, то энтропия S возрастает. Это означает, что уже простое расширение газа, заполнение им большего объема, является необратимым процессом. Если мы хотим вернуть газ в состояние с прежним объемом, что, конечно, можно осуществить, то мы должны подвергнуть систему внешнему воздействию, то есть в данном случае эту систему уже нельзя будет рассматривать как замкнутую.

Поскольку не существует обратимых процессов, второй закон термодинамики можно сформулировать в следующем виде: энтропия замкнутой системы возрастает до тех пор, пока в этой системе происходят какие-нибудь процессы, или, другими словами, пока не достигнуто состояние равновесия. Это не означает, однако, что всякий процесс является направленным. Если система незамкнута, то направление процесса зависит от взаимодействия этой системы с другими. Мы не можем сказать, что тепло всегда течет от теплого тела к холодному. Противоположное утверждение иллюстрируется обратным циклом Карно и принципом работы любого рефрижератора. Однако процесс охлаждения в рефрижераторе происходит только потому, что здесь одновременно электрическая энергия преобразуется в тепловую, которая через радиатор отводится в окружающую атмосферу. Электричество представляет собой высокоупорядоченную форму энергии. Ее превращение в тепловую энергию вызывает возрастание

энтропии, превышающее уменьшение энтропии, возникающее в результате процесса охлаждения внутри холодильника. Следовательно, энтропия всей системы в целом увеличивается, что является следствием уменьшения энтропии одной из ее подсистем. Таким образом, принцип, согласно которому энтропия в замкнутой системе не может уменьшаться, остается справедливым.

Эти простые иллюстрации показывают, что направление физических процессов можно выразить с помощью функции состояния — энтропии, которая, так сказать, является мерой *степени уравнивания*, достигнутого системой. Все процессы направлены к возрастанию уравнивания. Этот вывод справедлив по отношению не только к отдельным процессам, но и ко всей вселенной в целом.

Допустим, что мгновенное поперечное сечение через вселенную, то есть одно из ее мгновенных состояний, представляет собой физическое состояние, которому можно приписать конечную величину энтропии. Обсуждение возражений относительно этого предположения мы проведем ниже (см. стр. 159), где будет показано, что предполагаемая нами аргументация не зависит от данного предположения. Если вселенная в целом обладает в каждый данный момент определенной величиной энтропии, то эта величина подчиняется общему закону возрастания энтропии. Это означает, что вселенная развивается в направлении все более уравновешенных состояний. Хотя этот принцип приводит к нежелательному выводу, — а именно, что когда-нибудь наша вселенная достигнет состояния полного покоя и не представит больше возможностей для существования таких неравновесных систем, как живые организмы, все же он указывает нам на направление времени: положительное время направлено в сторону возрастания энтропии.

Открытие венского физика Людвига Больцмана (1872) внезапно положило конец этим грезам классической физики. Он пришел к выводу, что принцип возрастания энтропии следует рассматривать не как строгий, а как статистический закон. Мы не можем утверждать, что «энтропия должна возрастать», с таким же правом, как утверждаем, что «энергия должна оставаться постоянной». Формулировать это положение необходимо следующим образом: «Весьма вероятно, что энтропия возрастет». Иначе говоря, обратимый термодинамический

процесс не физически невозможен, а только весьма неправдоподобен.

Опираясь на выводы кинетической теории газов, согласно которой тепловые свойства газа определяются хаотическим движением его молекул, Больцман показал, что столкновения молекул подчиняются статистическим законам, приводящим к усреднению по скоростям. Может случиться так, что молекула, обладающая небольшой скоростью, при столкновении с более быстрой потеряет почти всю свою скорость и сообщит ее последней, скорость которой после этого возрастет. Однако такие события являются исключениями. В подавляющем большинстве столкновений более быстрые молекулы будут терять свою скорость, а более медленные — приобретать ее. Переход от высокой температуры к низкой нужно понимать, таким образом, как статистическое уравнивание различий между молекулами по скоростям. Закон возрастания энтропии гарантируется законом больших чисел, хорошо известным во всех видах статистик. Однако этот закон не относится к строгим физическим законам, таким, как законы механики, которые не допускают исключений.

На более ранних стадиях эволюции кинетической теории газов вероятностные законы применялись при рассмотрении конфигураций молекул. В частности, Д. К. Максвелл показал, что распределение молекул по скоростям соответствует нормальной кривой. Однако Больцман показал, что вероятность W состояния связана с его энтропией S, и сформулировал известное соотношение между ними с точностью до аддитивной постоянной:

$$S = k \cdot \ln W. \qquad (4)$$

Возрастание энтропии в физических процессах означает в данной интерпретации, что упорядоченные конфигурации молекул превращаются в неупорядоченные. Легко убедиться, что вероятность последующего состояния значительно превосходит вероятность предыдущего. Порядок представляет собой исключение в мире случая. Аддитивность энтропии, согласно уравнению (3), обусловлена тем, что вероятности можно умножать, а с помощью логарифмической функции (4) умножение заменяется сложением. Это нечто большее, чем простое количественное соответствие. В больцмановской теории газов формулы, подобные формуле (2), получаются при

помощи строгих выводов, и, таким образом, термодинамика полностью сводится к статистической механике.

Направление физических процессов, а значит, и направление времени, объясняется, следовательно, как статистическая тенденция: акт становления представляет собой переход от менее вероятных конфигураций молекул к более вероятным. Эту интерпретацию направления времени мы рассмотрим в последующих разделах. Мы увидим, что она на самом деле представляет собой центральное звено теории потока времени.

В этой связи можно было бы рассмотреть другое важное следствие интерпретации Больцмана. Это следствие касается вышеупомянутого различия между двумя видами физических законов. Оба вида законов выражаются с помощью импликаций, то есть утверждений типа: «если, то». Однако если строгий или причинный закон выражается *строгой* или *номологической импликацией*, то вероятностный закон выражается в виде *вероятностной импликации*. Первый имеет форму: «если..., то всегда...»; второй: «если..., то в некотором числе случаев...». При низкой вероятности статистический характер закона представляется очевидным. Однако если вероятность высока, то легко по ошибке принять вероятностный закон за строгий. В самом деле, так произошло в случае со вторым законом термодинамики, и Больцман совершил великое открытие, показав, что этот закон относится к разряду статистических.

После того как мы ясно осознали различие между двумя видами законов, возникает важный вопрос. Который из этих двух видов более фундаментален? Можно ли вывести один вид из другого? Во времена Больцмана, согласно общему мнению, законы строгого типа имели логическое преимущество. Использование в термодинамике статистики представлялось лишь средством, которое навязывается нам в силу ограниченности человеческих способностей. Выдвигалось предположение, что если бы мы смогли наблюдать траекторию каждой молекулы, то, подобно лапласовскому гению, смогли бы строго предсказать будущее состояние газа, не прибегая к статистике. Статистические законы представлялись, таким образом, как законы *«за неимением лучшего»*, то есть как формулировки, которые установлены эмпирическим путем, за неимением ничего другого. Согласно этой точке зрения,

статистические законы в конечном счете можно свести к причинным законам. (Эту проблему мы рассмотрим в § 10 и 11). В противоречие же с этой точкой зрения выдвигалось другое мнение, согласно которому статистические законы являются более фундаментальными законами. У нас нет никаких доказательств, что движение атомов подчиняется строгим законам. Возможно, что все причинные законы макрофизики можно объяснить как результат действия закона больших чисел, который превращает ограниченную вероятность элементарных событий в высокую вероятность процессов, происходящих в больших ансамблях. Судьба, постигшая второй закон термодинамики, была бы тогда прототипом аналогичных интерпретаций всех других так называемых строгих законов физики.

Развитие современной квантовой физики вынесло решение в пользу второй точки зрения. Однако рассмотрение этого вопроса мы отложим до V главы. Для того чтобы можно было понять переворот, осуществленный квантовой физикой, нам нужно сначала проанализировать возможные интерпретации классической физики. Необходимо разобраться в таком вопросе, верно ли, что классическая физика может обойтись без понятия вероятности? Мы увидим, что классическая физика нуждается в этом понятии так же, как и квантовая, хотя в ней оно используется в несколько ином смысле.

§ 8. Статистическое определение энтропии

Больцмановская статистика газов в своей первоначальной форме представляла собой сочетание причинных и вероятностных методов. Он предположил, что все элементарные процессы определяются строгими причинными законами, и хотел доказать, что совокупность таких процессов подчиняется статистическим законам. Для вывода данного положения он ввел определенное вероятностное предположение, которое служило дополнением к причинным законам физики.

Методы и предположения, применявшиеся при построении этой теории, можно проиллюстрировать на очень простом примере. Проанализируем распределение газа, заполняющего какой-нибудь сосуд. Мы хотим доказать, что равномерное распределение газа в пространстве

является наиболее вероятным, то есть что его энтропия стремится к максимальной величине.

Разделим сосуд на одинаковые кубические ячейки. Пусть число этих ячеек равно m. *Распределение* задано в том случае, если мы знаем число молекул n_i, заполняющих каждую ячейку. При неравномерном распределении, например, когда сосуд еще только начинает заполняться газом, эти *числа заполнения* n_i могут быть весьма различными. Вследствие движения молекул и их столкновений происходит определенная перетасовка, и процесс развития в направлении равномерного распределения можно рассматривать как переход от маловероятного распределения к более вероятному. В состоянии равновесия числа заполнений n_i будут приблизительно равны и, несмотря на непрерывные перемещения отдельных молекул из ячейки в ячейку, будут оставаться неизменными. Сумма всех n_i всегда должна быть равной числу n всех молекул, находящихся в сосуде:

$$\sum_{i=1}^{m} n_i = n. \qquad (1)$$

Распределение молекул следует отличать от *конфигурации* (arrangement). Для определения понятия «конфигурация» мы предположим, что каждая молекула представляет собой индивидуальную сущность, которая в принципе отлична от каждой другой молекулы, и что нам известно, какую ячейку она занимает. *Распределение представляет собой класс конфигураций.* Например, если какая-то молекула, находившаяся в ячейке i, поменялась местами с другой молекулой, находившейся до этого в ячейке k, то мы получим другую конфигурацию, или «размещение», тогда как распределение останется прежним. Однако если две молекулы поменялись местами в пределах одной и той же ячейки, то конфигурация не меняется, поскольку размеры ячейки определяют тот предел точности, в рамках которого мы даем точное определение конфигурации.

Для того чтобы высчитать вероятность конфигурации, Больцман предположил, что молекула с одинаковой вероятностью может занимать любую ячейку. Эта *вероятностная метрика*, на которой основываются все его расчеты, может быть записана в следующей форме:

$$\varphi(x_1, x_2, x_3) = \text{const}, \qquad (2)$$

где x_1, x_2, x_3 суть пространственные координаты и $\varphi(dx_1, dx_2, dx_3)$ представляет собой вероятность нахождения определенной молекулы в элементе объема $dx_1\, dx_2\, dx_3$. Поскольку число ячеек равно m, то это означает, что вероятность нахождения какой-нибудь молекулы в любой из ячеек равна $\frac{1}{m}$. Следовательно, вероятность любой данной конфигурации равна $\left(\frac{1}{m}\right)^n$, если предположить, что конфигурацию можно рассматривать как совокупность n независимых событий, причем вероятность каждого из них равна $\frac{1}{m}$. Напротив, полное число N конфигураций задается формулой

$$N = m^n. \tag{3}$$

Значение этой величины легко вычислить следующим образом. Одна молекула может занимать любую из m ячеек, и уже это даст m возможных конфигураций. Добавляя вторую молекулу, мы видим, что каждая из ее m возможных конфигураций может комбинироваться с каждой из предшествующих конфигураций, что дает m^2 конфигураций, и т. д.

Поскольку распределение D является классом конфигураций, то вероятность распределения определяется вычислением числа относящихся к нему конфигураций N_D. При таком методе опираются на упомянутое предположение, а именно, что молекулы различимы в индивидуальном порядке, и, следовательно, если молекулы, находившиеся в двух различных ячейках, поменяются местами, то это приведет к двум различным конфигурациям. Тогда вероятность распределения D задается формулой

$$W = \frac{N_D}{N} = N_D \cdot \frac{1}{m^n}. \tag{4}$$

Значение N_D находится при помощи перестановки всех молекул в ячейках при постоянных значениях n_i. Перестановки этого типа рассматриваются в теореме Бернулли [1], являющейся одной из фундаментальных тео-

[1] См., например: H. Reichenbach, The Theory of Probability, Berkeley Calif., Univ. Calif. Press, 1949, p. 264. В дальнейших ссылках эта книга будет обозначаться через ThP.

рем теории вероятностей. Доказано, что число таких перестановок равно

$$N_D = \frac{n!}{n_1! \cdots n_m!}. \tag{5}$$

Удобнее вычислить $\ln N_D$, а не саму величину N_D. Кроме того, достаточно рассматривать только такие случаи, в которых значения n_i не слишком малы. При этом мы можем использовать известную формулу Стирлинга для приближенного вычисления факториала:

$$\ln n! = n \ln n - n + \frac{1}{2} \ln 2\pi n. \tag{6}$$

При большом n третий член по сравнению с первыми двумя обладает меньшим порядком. То есть если его разделить на n, то с возрастанием n частное стремится к нулю. Предположим теперь, что велико не только число n, но и числа заполнения n_i, пренебрегая при этом крайними случаями неравномерных распределений, тогда мы можем пренебречь и третьим членом. Применяя полученное выражение для определения n и n_i, мы находим:

$$\ln N_D = n \ln n - \sum_{i=1}^{m} n_i \ln n_i. \tag{7}$$

Эти выражения получаются из первого члена формулы (6), когда с ее помощью определяют числитель и знаменатель формулы (5). Второй член формулы (6) сокращается на основании формулы (1).

Использовав это удобное упрощение для вычисления $\ln N_D$, мы можем теперь перейти к определению $\ln W$. Комбинируя формулу (7) с формулами (4) и (3), мы находим, пользуясь больцмановским отождествлением энтропии и логарифма вероятности состояния, что

$$S = \ln W = n \ln n - \sum_{i=1}^{m} n_i \ln n_i - n \ln m. \tag{8}$$

Это выражение представляет собой *нормированную* форму энтропии в том смысле, что в нем не содержится никаких неопределенных постоянных. Кроме того, эта форма энтропии (8) является нормированной относительно вероятности, поскольку она непосредственно вытекает из логарифма вероятности состояния газа, несмотря

на то, что мы пренебрегаем последним членом формулы Стирлинга (6). Однако, поскольку этот член мал, соотношение (8) можно рассматривать как вполне допустимое приближение. Оно определяет энтропию как функцию чисел заполнения, а замена вероятности на ее логарифм гарантирует аддитивность энтропии. Больцман сформулировал это отношение несколько иначе. Для того чтобы получить соответствие со значением энтропии S, полученным Клаузиусом и Томсоном, он ввел постоянную k из соотношения (7. 4). Кроме того, он отказался от точного формулирования последнего члена в формуле (8), заменив его неопределенной постоянной, допуская, таким образом, возможность различных интерпретаций этой формулы.

Поэтому он написал выражение энтропии в *ненормированной* форме:

$$S = k\left[n \ln n - \sum_{i=1} n_i \ln n_i\right] + \text{const}. \tag{9}$$

Полезно рассмотреть и третью форму выражения энтропии, которая, подобно (8), не содержит никакой неопределенной постоянной и отличается от формулы (8) только коэффициентом.

Введем новую величину

$$p = \frac{n_i}{n}. \tag{10}$$

Величину p_i можно рассматривать как условную вероятность нахождения молекулы в ячейке i, когда газ характеризуется рассмотренным выше распределением D. Разделив формулу (8) на n, мы получаем:

$$\begin{aligned} S = \frac{1}{n} \ln W &= \ln n \left(\sum_{i=1}^{m} p_i\right) - \sum_{i=1}^{m} p_i \ln n_i - \ln m = \\ &= -\sum_{i=1}^{m} [p_i \ln n_i - p_i \ln n] - \ln m = \\ &= -\sum_{i=1}^{m} p_i \ln p_i - \ln m, \end{aligned} \tag{11}$$

учитывая, что

$$\sum_{i=1}^{m} p_i = 1. \tag{12}$$

Величина S, полученная в формуле (11), отличается от величины S формулы (8) наличием множителя $\frac{1}{n}$ и может быть названа *удельной энтропией*, поскольку она не зависит от числа молекул n. Величины S как в (8), так и в (11) являются нормированными относительно вероятности. Таким образом, они представляют собой отрицательные числа, поскольку W — дробное число, а так как $W \leqq 1$, то и ее логарифм $\leqq 0$. По сравнению с первоначальной формой энтропии нестатистической термодинамики *статистические формы* (8), (9) и (11) являются некоторым обобщением. В нестатистической термодинамике энтропия определяется только для состояний равновесия. Напротив, статистические формы определяют величину S для всех условий, которые действуют в то время, когда молекулы газа распределены в пространстве неравномерно. Определение статистической формы энтропии для неравновесных состояний позволяет объяснить движение в направлении состояния равновесия. Состояния равновесия обладают более высокой вероятностью, и эволюция газа из состояний, характеризуемых неравномерным распределением, к состояниям, характеризуемым равномерным распределением, представляется в виде тенденции перейти из состояний с низкой вероятностью в состояния с более высокой вероятностью.

Рассмотрим вопрос, как определить распределение, которое выражает состояние равновесия, то есть распределение, соответствующее максимальному значению W. Поскольку множитель $\frac{1}{N}$ в (4) является постоянным, то это распределение можно найти, определяя максимальное значение N_D по формуле (5). При этом удобнее пользоваться более простой формулой (7), поскольку она делает возможным использование упрощения, даваемого формулой Стирлинга. Определение значений n_i, соответствующих максимальному значению, полученному из формулы (7), несколько усложняется, поскольку n_i подчиняются условию (1). Для нахождения максимальных значений функций при некоторых ограничительных условиях Лагранж разработал вариационный метод множителей, применение которого приводит нас к следующему результату:

$$\sum_{i=1}^{m} [\delta(n_i \ln n_i) + \alpha \cdot \delta(n_i)] = 0. \qquad (13)$$

Первая часть получается путем вариации функции (7), а вторая — функции (1). Когда мы проводим дифференцирование, то множитель Лагранжа α позволяет нам рассматривать δn_i как независимые величины. Таким образом, каждый член этой суммы должен обращаться в нуль, и мы получаем соотношение

$$\ln n_i + 1 + \alpha = 0, \quad n_i = e^{-1-\alpha} = \text{const} = \frac{n}{m} = q. \quad (14)$$

Величина константы получается из формулы (1), и мы обозначаем ее через q.

Это означает, что максимальная вероятность ассоциируется с тем случаем, когда все ячейки заполнены одинаковым количеством молекул. Комбинируя формулу (14) с (10), мы видим, что в этом случае все p_i равны между собой, в частности они равны $\frac{1}{m}$. В таком случае это и есть то условие, когда энтропия достигает максимального значения.

Поскольку максимальное значение больцмановской энтропии S зависит еще и от значения постоянной в формуле (9), легко убедиться, что в нормированных значениях максимальной энтропии в формулах (8) и (11) предполагается, что величина $S = 0$. В формуле (8) этот результат получается в том случае, если мы заменим n_i на n/m, а для формулы (11), если заменим p_i на $\frac{1}{m}$. Это означает, что вероятность W здесь равна 1. Полученный вывод не совсем точен, вероятность W может быть лишь весьма близкой к 1. Неточность обусловлена тем, что мы пренебрегаем последним членом в формуле Стирлинга (6). Однако для всех практических задач мы можем приписать равномерному распределению молекул вероятность, равную 1.

В своих трудах Больцман вместо S пользуется функцией H, которая представляет собой отрицательную величину S и задается, следовательно, в случае удельной энтропии (11) следующей формулой:

$$H = \sum_{i=1}^{m} p_i \ln p_i + \ln m. \quad (15)$$

Для теории газов это просто более удобная форма записи, поскольку $H > 0$. Это означает, что H стремится к минимуму, а S — к максимуму. Ниже мы увидим, что H

можно приписывать особое значение и применять в другой области науки, а именно в теории информации.

Разберемся более подробно, каким образом в статистической интерпретации энтропии описывается переход к состояниям равновесия. Рассмотрим простой пример. Имеются два сообщающихся между собой сосуда одинаковых размеров: один заполненный газом, второй — пустой. Мы проанализируем процесс заполнения газом пустого сосуда после того, как был открыт кран, соединяющий оба сосуда. Используя выражение энтропии из формулы (8), нормированное относительно вероятности, мы можем дать следующее описание процесса. Пока кран закрыт, газ находится в пространстве, разделенном на m ячеек, в состоянии равновесия. Таким образом, числа заполнения n_i имеют постоянное значение $\frac{n}{m}$, заданное в формуле (14). Если мы через S_1 обозначаем энтропию этого состояния, то из формулы (8) легко убедиться, что величина $S_1 = 0$. Вероятность состояния газа, таким образом, равна 1. Затем, когда открываем кран, число ячеек увеличивается до $2m$, однако вначале все новые ячейки пусты. Хотя верхний предел суммирования в формуле (8) и увеличивается до $2m$, в итоге величина суммы не меняется, тогда как из третьего члена этой формулы, где m заменяется на $2m$, получается еще добавочный член $-n\ln 2$. Сразу же после того, как был открыт кран, энтропия уменьшается до величины S', которая равна

$$S' = S_1 - n\ln 2. \tag{16}$$

Поскольку $S_1 = 0$, то мы имеем здесь вероятность $W = \frac{1}{2^n}$. Такая малая вероятность получается потому, что вероятность всегда относительна, и с увеличением числа ячеек изменяется и соответствующая им вероятность. Однако состояние, характеризуемое малой вероятностью, недолговечно; газ врывается в пустой сосуд, заполняя его ячейки. В конечном состоянии равновесия каждая из $2m$ ячеек будет заполнена $\frac{n}{2m}$ молекулами. Когда мы по формуле (8) вычисляем величину энтропии, подставляя $\frac{n}{2m}$ вместо n_i и увеличивая верхний

предел суммирования до $2m$, мы вновь получаем $S = 0$. Таким образом, газ достигает состояния равновесия, которое характеризуется величиной S_2. Его энтропия такова, что $S_2 = S_1$.

Эта вероятностная интерпретация процесса в одном существенном пункте отличается от термодинамической. Термодинамическая энтропия есть функция состояния, которая не претерпевает изменений, если открывается кран, соединяющий оба сосуда. В самом деле, уменьшение энтропии, отраженное в формуле (16), представляется здесь совершенно невозможным. Газ сохраняет свою энтропию S_1, если открыть кран, однако с началом процесса расширения энтропия начинает возрастать, и когда, наконец, заполняется второй сосуд, вся система в целом, состоящая из двух сосудов, достигает энтропии S_2, такой, что $S_2 > S_1$.

Эти соображения показывают, что термодинамическую энтропию нельзя отождествлять с величиной S из формулы (8), которая нормирована относительно вероятности. Необходимо такое значение функции, которое оставалось бы постоянным после того, как открывается кран, но стало бы возрастать с началом процесса заполнения газом пустого сосуда. Это означает, что для термодинамики требуется *абсолютная энтропия*, и при ее интерпретации нельзя пользоваться *относительной энтропией*, получающейся при использовании вероятностных методов. Мы можем найти соответствующую величину, отождествив S не с $\ln W$, а с $\ln N_D$, то есть когда мы будем рассматривать как меру энтропии абсолютное, а не относительное число конфигураций. Из формулы (7) следует, что в таком случае при определении S нам следовало бы опустить третий член в формуле (8). Мы могли бы использовать также и формулу, в которой содержится аддитивная постоянная, если мы полагаем, что она не зависит от m, а имеет следующий вид: $n \cdot \mathrm{const}$. Зависимость от n необходима для того, чтобы с удвоением числа молекул энтропия также увеличивалась бы в 2 раза. В данной интерпретации больцмановская энтропия (9) является абсолютной [1].

---

[1] Несколько отличную нормировку постоянной, входящей в выражение абсолютной энтропии, см. § 9, стр. 97, прим. 1.

Хотя на первый взгляд кажется, что использование понятия абсолютной энтропии не должно встретить каких-либо возражений, оно все же приводит к некоторым трудностям, на которые впервые указал Гиббс. Предположим, что мы снова имеем два смежных сосуда, однако на этот раз второй сосуд наполнен газом, находящимся в точно таких же физических условиях, как и газ, заполняющий первый сосуд. До тех пор пока кран, соединяющий оба сосуда, закрыт, каждая система обладает энтропией, равной S_1, а вся система в целом, следовательно, обладает энтропией, равной $2S_1$. Теперь, если мы открываем кран, то с макроскопической точки зрения ничего не происходит. В самом деле, вся система в целом остается в состоянии равновесия. Величина ее энтропии S по-прежнему должна быть равна $2S_1$. Легко убедиться, что, когда мы используем понятие относительной энтропии (8), это требование выполняется. Если в формуле (8) мы заменим n на $2n$ и m на $2m$ при условии, что величины n_i останутся без изменений, то величина всего выражения просто удвоится. Дополнительный член $2n \ln 2$, получающийся из первого члена, сокращается здесь с $-2n \ln 2$, получающимся из третьего члена. Однако если мы опустим третий член и используем понятие абсолютной энтропии, то получим следующее соотношение:

$$S = 2S_1 + 2n \ln 2. \qquad (17)$$

Открывание крана влечет за собой возрастание энтропии, хотя в системе не произошло никаких макроскопических изменений. Энтропия всей системы в целом больше суммарной энтропии двух индивидуальных систем.

Этот *парадокс нормировки*, именуемый также парадоксом Гиббса, является неизбежным для статистики Максвелла — Больцмана [1], если пользоваться понятием абсолютной энтропии.

Здесь невозможна удовлетворительная интерпретация термодинамической энтропии с помощью абсолютного

---

[1] Парадокс исчезнет в том случае, если мы опускаем в формуле (8) не только третий, но и первый член. Однако тогда, при удвоении числа молекул n, возникает другой парадокс. Вместо увеличения от S до искомой величины $2S$ энтропия увеличивается от S до $2S - 2n\ln 2$.

числа конфигураций, потому что нормировку абсолютной энтропии нельзя провести удовлетворительным образом. Энтропия целого не равна сумме энтропий его частей. Таким образом, нарушается *постулат аддитивности*. Физик прибегает к такой интерпретации, которую лишь *с большой натяжкой* можно назвать вероятностной. Если термодинамическая система переходит от энтропии S_1 к энтропии S_2, а затем от S_2 к S_3, то физик относит одно возрастание вероятности к переходу от S_1 к S_2, другое — к переходу от S_2 к S_3. Однако эти отдельные возрастания вероятностей не объединяются. Здесь нужно произвести двойной пересчет: во-первых, S_1 нужно заменить на меньшую величину S_1', от которой система переходит к большей величине S_2', соответствующей S_2; во-вторых, S_2' нужно заменить на меньшую величину S_2'', от которой система переходит к большей величине S_3'', соответствующей S_3, и т. д. В математическом аппарате этот пересчет вероятностей вызывает появление неопределенных постоянных. Правильные формулы определения энтропии получаются после соответствующей корректировки с помощью аддитивных постоянных. Эта *аддитивная незамкнутость* является общим свойством энтропии в статистике Максвелла — Больцмана. В дальнейшем будет показано, что квантовая статистика имеет некоторое преимущество в этом отношении и не нуждается, чтобы прибегать к таким выражениям энтропии, которые незамкнуты в отношении аддитивности.

Выше было показано, что парадокс Гиббса не относится к относительным формам энтропии, сформулированным в соотношениях (8) и (11). Выполнение постулата аддитивности было проверено с помощью непосредственных вычислений. Следует иметь ясное представление, что эти вычисления лишь приблизительно верны, поскольку в них пренебрегают последним членом в формуле Стирлинга (6). Если при расчетах учитывать третий член, то в этом случае мы также придем в противоречие с постулатом аддитивности. Например, мы обнаружим, что если $W(n, m)$ есть вероятность распределения n частиц в m ячейках, которое задано формулами (4) и (5), то соотношение

$$W(2n, 2m) = W(n, m)^2, \tag{18}$$

строго говоря, не является истинным. Левая часть меньше правой, и их коэффициент пропорциональности будет равен $\frac{1}{\sqrt{\pi n}}$. В самом деле, даже относительная энтропия не удовлетворяет принципу, согласно которому энтропия целого равна сумме энтропий его частей, потому что вероятность целого, строго говоря, не равна произведению вероятностей его частей. Уменьшение вероятности с открыванием крана с точки зрения физики объясняется тем, что в этом случае число молекул в каждом из двух сосудов может измениться. Молекулы могут сосредоточиться в одном сосуде. Однако принимать во внимание подобную возможность нет необходимости, и, следовательно, пренебрежение третьим членом в формуле Стирлинга представляется оправданным [1].

§ 9. Распространение статистики на случай газа с различными уровнями энергии

Теперь мы перейдем к обобщению определения энтропии для такой системы газа, в которой ячейки или отделения связаны с различными уровнями энергии. Это положение иллюстрируется следующим образом: молекула может иметь различные значения кинетической энергии, которая представляет собой функцию от ее скорости. Поскольку скорости молекул зависят от температуры, то изучение статистики скоростей молекул и определение соответствующей им энтропии позволяет нам дать описание процесса уравнивания температуры в газе, который вначале характеризовался неравномерным распределением температуры. Мы будем следовать представлениям Больцмана.

Рассматривая три компоненты скорости молекулы v_1, v_2 и v_3 как декартовы координаты, Больцман постулировал пространство скоростей, подразделяемое им на

[1] Может возникнуть предположение, что мы рассматриваем пренебрежение последним членом в формуле Стирлинга не как апроксимацию, а как вопрос определения математического выражения энтропии. Тогда относительная энтропия в точности удовлетворяла бы постулату аддитивности. Однако такой выход из положения представляется неприемлемым, поскольку тогда мы не могли бы сказать, что вероятность действительных событий строго удовлетворяет понятию энтропии.

громадное число равносторонних кубических ячеек m, соответствующих тем ячейкам, которые были рассмотрены в предыдущем примере. Каждую ячейку можно представлять себе как маленький кубик $d\omega = dv_1 dv_2 dv_3$. Предполагается, что молекула, которая в определенный момент времени t обладает определенной скоростью, занимает одну из ячеек в этом воображаемом пространстве. В ячейке может находиться много молекул, а именно все те, которые в момент t обладают одинаковой скоростью в пределах степени точности, определяемой размерами ячейки. Если для каждой ячейки задано число заполнения n_i, то нам известно *распределение молекул по скоростям* в момент t. В другой момент t' может существовать другое распределение молекул по скоростям. Любое распределение по скоростям и любая конфигурация должны удовлетворять условию (8.1), которое мы вновь приведем в этом параграфе:

$$\sum_{i=1}^{m} n_i = n. \tag{1}$$

Однако, помимо этого, оно должно удовлетворять условию, согласно которому теплосодержание газа равно полной сумме механических энергий составляющих его молекул, поскольку тепловая энергия является не чем иным, как полной суммой кинетических энергий его молекул. Кинетическая энергия молекулы ε массы μ определяется формулой

$$\varepsilon = \frac{\mu}{2}(v_1^2 + v_2^2 + v_3^2). \tag{2}$$

Эта величина непрерывно изменяется с изменением v_1, v_2, v_3. Однако, поскольку мы предположили, что диапазоны этих переменных подразделены на отрезки таким образом, что составляют небольшие кубические ячейки, то мы соответственно разделяем диапазон ε на отрезки $d\varepsilon$. Каждая ячейка относится к некоторому уровню энергии, а каждый уровень энергии содержит определенное количество ячеек. Обозначая через ε_i уровень энергии, ассоциируемый с ячейкой i (отметим, что разные ячейки i и j могут относиться к одному и тому же уровню энергии, так что $\varepsilon_i = \varepsilon_j$), мы можем записать

энергетическое условие в следующем виде:

$$\sum_{i=1}^{m} n_i \varepsilon_i = E, \tag{3}$$

где E — полная энергия газа.

Если конфигурация удовлетворяет условиям (1), (3), то ее можно назвать *возможной конфигурацией*. Число возможных конфигураций N_E является конечным, поскольку скорости молекул ограничиваются энергетическим условием. Молекула обладает максимальной скоростью v_{max} в том случае, если в ней сосредоточена вся тепловая энергия газа, тогда как скорости всех остальных молекул равны нулю. Теперь, как и в предыдущем примере, нужно ввести вероятностную метрику. Это можно осуществить следующим образом. Наибольшая скорость v_{max} ограничивает существование ячеек в пространстве скоростей внутренней областью сферы K с радиусом v_{max}, проведенным из начала координат. Предположим, что в сфере содержится m ячеек. Как и прежде, здесь имеется $N = m^n$ конфигураций. Большинство из них не будет соответствовать энергетическому условию, записанному в формуле (3). Следовательно, все эти конфигурации аннулируются. Оставшиеся N_E конфигураций, удовлетворяющие энергетическому условию, рассматриваются теперь как равновероятные; их вероятность равна $\frac{1}{N_E}$. На этом вероятностном предположении основывается теория Больцмана. Мы видим, что она исходит из предположения о постоянстве вероятностной метрики пространства скоростей, которая подчиняется ограничениям, налагаемым энергетическим условием.

Число конфигураций в пределах распределения N_D вновь задается формулой (8.5). Таким образом, мы имеем

$$W = \frac{N_D}{N_E}. \tag{4}$$

Величину N_E нельзя получить при помощи простых вычислений. Поэтому, используя формулу (8.7), мы запишем значение энтропии, по аналогии с (8.8), следующим образом:

$$S = \ln W = n \ln n - \sum_{i=1}^{m} n_i \ln n_i - \ln N_E. \tag{5}$$

Поскольку нас, по сути дела, интересует только максимальное значение W, то есть распределение, которое для данной величины E соответствует состоянию равновесия, то мы можем не принимать во внимание, что нам неизвестна постоянная N_E в формуле (5). Мы можем интерпретировать формулу (5) как выражение абсолютной энтропии, в котором пренебрегают этой постоянной. Максимальное значение S в формуле (5) находится, как и прежде, путем максимизации членов суммирования, налагая, однако, на этот раз два ограничивающих условия (1) и (3). Используя метод Лагранжа, мы приходим к следующему результату:

$$\sum_{i=1}^{m} [\delta(n_i \ln n_i) + \delta(\alpha n_i) + \delta(\beta \varepsilon_i n_i)] = 0,$$

$$\ln n_i + 1 + \alpha + \beta \varepsilon_i = 0, \qquad (6)$$

$$n_i = a e^{-\beta \varepsilon_i}, \quad a = e^{-1-\alpha}.$$

Поскольку α и β постоянны, то эта формула показывает, что число заполнения n_i зависит только от уровня энергии, ассоциируемого с данной ячейкой. В применении к пространству скоростей это означает, что n_i зависит от кинетической энергии, соответствующей величинам v_1, v_2, v_3, характеризующим эту ячейку. Пусть $m(\varepsilon)$ будет числом ячеек, ассоциируемых с одним и тем же уровнем энергии ε, а $n(\varepsilon)$ — полным числом частиц, расположенных в этих ячейках, когда газ находится в состоянии равновесия. В таком случае мы имеем, поскольку n_i для этих ячеек постоянно:

$$n_i = \frac{n(\varepsilon)}{m(\varepsilon)} = q(\varepsilon) = a e^{-\beta \varepsilon}. \qquad (7)$$

Здесь для системы, находящейся в газообразном состоянии, $m(\varepsilon)$ постоянно. Величина $q(\varepsilon)$ представляет собой число частиц на ячейку, а следовательно, число заполнения любой ячейки в состоянии равновесия для уровня энергии ε. Нижний индекс i у буквы ε опущен, поскольку ε относится теперь не к отдельным ячейкам, а к уровню энергии.

Отложив на будущее рассмотрение постоянной a, мы сейчас подсчитаем постоянную β, то есть выразим ее в термодинамических параметрах. Во-первых, мы можем придать формуле (8.9) новый вид, заменяя $\ln n_i$ на

$\ln a - \beta\varepsilon_i$ на основании формулы (6). Используя формулы (1) и (3), мы находим:

$$S = k[n \ln n - n \ln a + \beta E]. \tag{8}$$

Во-вторых, мы можем комбинировать формулу (8), которая выражает зависимость S от полной энергии E, с фундаментальным соотношением (7.1) нестатистической термодинамики, в котором количество тепла Q представляет энергию E. Дифференцируя (8) и используя (7.1), мы находим

$$\frac{dS}{dE} = k\beta = \frac{1}{T}; \tag{9}$$

следовательно,

$$\beta = \frac{1}{kT}. \tag{10}$$

Подставляя это значение в формулу (7), мы получаем

$$n_i = \frac{n(\varepsilon)}{m(\varepsilon)} = q(\varepsilon) = ae^{-\varepsilon/kT}. \tag{11}$$

Таким образом, показано, что величина $q(\varepsilon)$, которая определяет число частиц на ячейку в состоянии равновесия, находится в экспоненциальной зависимости (11) от уровня энергии ε и температуры T. Поскольку T представляет собой свойство газообразной системы в целом, то отношение (11) говорит о том, что различие между числами заполнения разных уровней энергии зависит от условий, характеризующих газ как единую систему. Если температура газа T повышается, то число заполнения каждой ячейки определенного уровня энергии возрастает по сравнению с числами заполнения ячеек, относящихся к более низким уровням энергии. Молекулы покидают низкие уровни и переходят на более высокие уровни энергии.

Хотя при выводе формулы (11) мы рассматривали такую газообразную систему, в которой различие между уровнями энергии обусловливается только различиями между молекулами по скоростям, однако данное соотношение (11) не зависит от этого специального случая. В самом деле, соотношение (11) можно получить всякий раз, когда (8.5) выражает меру вероятности статистического состояния и когда система является зам-

кнутой, то есть если предполагается, что энергетическое условие имеет вид, представленный в формуле (3). Утверждение о том, что соотношение (8.5) определяет меру вероятности, следует из двух предположений: о равновероятности конфигураций и индивидуальной различимости элементарных частиц. Статистику, удовлетворяющую этим двум предположениям, обычно именуют *статистикой Максвелла — Больцмана.* Для этой статистики экспоненциальная зависимость стационарных чисел заполнения $q(\varepsilon)$ от уровня энергии и от температуры выражает фундаментальный закон, который справедлив независимо от того, определяются ли уровни энергии только скоростями молекул, или же скоростями и полем тяготения (как обстоит дело в случае с земной атмосферой), или другими условиями.

Мы рассматриваем аргумент ε в формуле (11) как такую переменную, которая изменяется непрерывным образом, тогда как величина n_l относится к ячейкам $d\omega$ определенных размеров. Для того чтобы элиминировать эту непоследовательность, введем *предположение непрерывности,* согласно которому, если $d\omega$ стремится к нулю, то существуют следующие пределы:

$$\frac{n_l}{d\omega} \to n'_l, \quad \frac{n(\varepsilon)}{d\omega} \to n'(\varepsilon), \quad \frac{q(\varepsilon)}{d\omega} \to q'(\varepsilon), \quad \frac{a}{d\omega} \to a'. \qquad (12)$$

Напротив, когда $d\omega$ стремится к нулю, величина $m(\varepsilon)$ остается постоянной. Постулирование существования пределов (12) порождает некоторые трудности, поскольку при уменьшении $d\omega$ величины n_l будут становиться равными либо 0, либо 1. Эту трудность можно преодолеть в том случае, если рассматривать n_l как среднее во времени число заполнения. Иначе мы можем сказать, как это обычно и делается в физике, что здесь мы имеем в виду не строгий предел, который получается, когда $d\omega$ неограниченно стремится к 0, а используем практический предел, получающийся в том случае, когда рассматриваются изменения $d\omega$ в рамках определенного порядка малости. Ассоциирование перехода к меньшим $d\omega$ с возрастанием числа частиц n допустимо только в том случае, если предполагается одновременное возрастание объема, и может, таким образом, относиться лишь к пространству скоростей, а не к физическому пространству. В противном случае возрастание

давления газа привело бы к изменению физических условий [1].

Теперь, если мы разделим n_l, $n(\varepsilon)$, $q(\varepsilon)$ и a в соотношении (11) на $d\omega$, то получим такие же соотношения, как и те, которые справедливы для штрихованных величин. Полагая

$$q'(\varepsilon) = n \cdot p(\varepsilon), \quad a' = n \cdot b \tag{13}$$

и выражая (ε) через формулу (2), мы в данном случае находим, рассматривая p как функцию от v_1, v_2, v_3:

$$p(v_1, v_2, v_3) = b \cdot e^{-(\mu/2kT)\cdot\left(v_1^2+v_2^2+v_3^2\right)}. \tag{14}$$

Из этого соотношения, в котором выражается *максвелловское распределение по скоростям*, вытекает, что вероятность $pd\omega$ нахождения молекулы в ячейке $d\omega$ для состояния равновесия задается нормальной кривой, или, вернее, нормальной поверхностью в трех измерениях, а именно в координатах v_1, v_2, v_3. Теперь можно определить постоянную b, вводя условие, что интегрирование соотношения (14) по dv_1, dv_2, dv_3 должно давать в итоге 1. При интегрировании получается следующий результат [2]:

$$b = \left(\frac{\mu}{2\pi kT}\right)^{3/2}. \tag{15}$$

---

[1] Используя пределы (12), мы можем определить значение абсолютной энтропии в форме, которая не зависит от размеров стороны ячейки $d\omega$. Для относительной энтропии это условие выполняется автоматически. Когда мы подсчитывали величину энтропии для ячеек меньших размеров, применяя упомянутое выше предположение непрерывности, то величина энтропии оставалась одной и той же, несмотря на то, что число ячеек возрастает. Например, если сторона ячейки уменьшается вдвое, то число их возрастает до $2m$, при этом мы обнаружим, что относительная энтропия (8.8) не меняется. Однако если пренебречь третьим членом в формуле (8.8), то энтропия изменится. Это изменение можно возместить, добавляя член $n \ln d\omega$ к постоянной и включая ее в процесс суммирования. Тогда абсолютную энтропию можно записать в следующем виде:

$$S = n \ln n - \sum_{l=1}^{m} \left[n_l' \ln n_l'\right] d\omega + \text{const},$$

где n_l имеет предельное значение, которое определяется по формуле (12). Эта формула позволяет заменить суммирование интегрированием.

[2] См., например, Richard C. Tolman, The Principles of Statistical Mechanics, Oxford, The Clareton Press, 1938, p. 90.

В отличие от постоянной β в формуле (18), которой можно дать общую оценку, мы видим, что постоянную $b = na'$ можно определить только после того, как энергия ε задана в виде функции от некоторых параметров, как это сделано в формуле (2). Вершина этой колоколообразной поверхности имеет координаты $v_1 = v_2 = v_3 = 0$. Это означает равновероятность всех направлений скоростей, а тем самым и то, что их среднее значение равно 0. В отношении средней, то есть абсолютной величины скорости v, дело обстоит по-другому. Из соотношения (14) получается следующее среднеквадратичное значение скорости v_0[1]:

$$v_0 = \sqrt{\overline{v^2}} = \sqrt{\frac{3kT}{\mu}}. \qquad (16)$$

Отсюда следует, что наиболее вероятным распределением является не то, которое характеризуется равными числами n_i для всех ячеек пространства скоростей, а такое распределение, при котором большинство молекул обладает средней скоростью v_0, в то время как остальные обладают большими или меньшими скоростями. Согласно формуле (16), средняя скорость v_0 возрастает пропорционально квадратному корню из температуры газа в целом. Значительные отклонения в обе стороны от среднего значения гораздо реже незначительных. Однако, как и прежде, это распределение в пространстве скоростей характеризуется наличием крайне большого числа конфигураций и, таким образом, показывает, что равновероятность всех возможных конфигураций ведет к неравным вероятностям различных распределений и к выделению одного из них как наиболее вероятного.

Вычисление вероятностей вновь приводит нас к выводу относительно изменения состояния системы со временем: если первоначальное распределение молекул характеризовалось низкой вероятностью и если они после этого находились в процессе движения и перемешивались, то система будет развиваться в направлении все более вероятных распределений. Механизм перемешивания обеспечивается столкновениями молекул; и главной

---

[1] Вывод формулы см. там же, стр. 92.

заслугой Больцмана является исследование того, какое влияние оказывают столкновения молекул на взаимный обмен скоростями между ними. Если вначале газ подразделялся на сферы, обладающие высокой и низкой температурами, то столкновение молекул ведет к выравниванию их скоростей, и в конечном счете газ достигает состояния термодинамического равновесия, то есть максимально вероятного состояния, в котором распределение молекул по скоростям происходит согласно нормальной кривой. Это распределение зафиксировано в уравнении (14). После того как система достигла этого распределения, оно становится стационарным, то есть, хотя молекулы и в дальнейшем при столкновениях обмениваются скоростями, распределение в целом остается неизменным. Таким образом, вероятностная интерпретация вновь позволяет объяснить эволюцию в направлении равновесия, в данном случае в направлении уравнивания температуры.

Мы рассмотрели понятие энтропии в двух отношениях. Во-первых, относительно ячеек физического пространства, заполненного газом, и, во-вторых, относительно ячеек в пространстве скоростей. Объединение этих двух видов статистики в один не представляет особых затруднений. Объединение трех измерений физического пространства с тремя измерениями пространства скоростей представляет собой шестимерное пространство параметров. Когда мы говорим, что молекула находится в некоторой ячейке этого пространства, то этим самым мы подразумеваем, что она находится в определенном месте физического пространства и что она обладает определенной скоростью. Определение энтропии в данном случае точно такое же, как и раньше. В самом деле, мы можем применять формулы (8.8), (8.9) и (8.11) без каких-либо изменений. Здесь верхним пределом суммирования будет число ячеек m в пространстве параметров. В данном случае метод максимизации, который раньше мы использовали в формулах (8.13) и (6) по отдельности, применялся бы сразу для определения энтропии в пространстве параметров и привел бы к тем же самым результатам. Эти соображения показывают, что энтропию можно рассматривать как функцию, которая дает наиболее полную характеристику состояния газа в целом.

Хотелось бы добавить несколько замечаний относительно термина «состояние». Мы проводим различие между *микросостоянием* и *макросостоянием*. Микросостояние — это то же, что и конфигурация. Макросостояние является классом, который либо объединяет, либо разграничивает распределения, получаемые при допущении существования небольших изменений числа молекул n_i, заполняющих ячейки, если только существование этих изменений нельзя установить с помощью макронаблюдений. Макросостояние A является, следовательно, классом *макроскопически неразличимых* микросостояний a. Например, если два газа настолько хорошо перемешаны, что в одинаковых небольших субобъемах содержится приблизительно одинаковое число молекул каждого газа, то данное макросостояние A может представлять различные микросостояния a. В самом деле, это есть макросостояние равновесия, которое остается неизменным в то время, как его микросостояния меняются непрерывно. Если A не является состоянием равновесия, то оно также может представлять множество состояний a. Например, если мы на мгновение приоткроем кран газовой плиты, то вначале газ будет занимать только ту часть пространства, которая находится в непосредственной близости от плиты. Можно прекрасно установить макроскопическое различие между данным макросостоянием A и таким состоянием, в котором газ полностью смешался с воздухом, заполняющим комнату. Однако состояние A может быть представлено различными микросостояниями a.

Различение между микросостояниями и макросостояниями, которое всегда последовательно проводится в литературе по статистической механике, приводит к естественной классификации микросостояний. Если дано микросостояние a_1, то этим уже определяется макросостояние A, к которому оно относится: A включает все те микросостояния a_i, которые макроскопически нельзя отличить от a_1. Мы называем макросостояние A *соответствующим* микросостоянию a_1, или, короче, — *соответствующим макросостоянием*. Эта классификация позволяет нам иногда говорить об энтропии микросостояния a. В этом случае мы имеем в виду не вероятность микросостояния a, а вероятность соответствующего макросостояния A. Если придерживаться данной интер-

претации, то можно смело пользоваться выражением *энтропия микросостояния*. В том же самом смысле мы будем говорить также и о макровероятности состояния *a*, подразумевая при этом вероятность соответствующего макросостояния *A*.

Состояние, в котором газы хорошо перемешаны, иногда называют *неупорядоченным* в отличие от *упорядоченного*, примером которого может служить такое состояние, когда все молекулы азота находятся в одной половине сосуда, а все молекулы кислорода — в другой. В этом случае смысл термина «неупорядоченный» не всегда соответствует тому смыслу, который подразумевается при использовании этого слова в нашей повседневной жизни. Могут возразить, что в примере с газовой плитой, когда газ заполняет пространство в непосредственной близости от нее, нам бы следовало назвать это состояние беспорядочным, тогда как равномерное распределение газа по всей комнате представляется скорее как состояние, обладающее высокой степенью порядка. Однако употребление термина «порядок» в таком значении не принято в статистике. Под «порядком» здесь понимается *степень разделения*. Следовательно, распределение, в котором газы хорошо перемешаны, обладает более низкой степенью порядка, тогда как уже частичное разделение обладает порядком более высокой степени. Термин «разделение» может относиться к разобщению молекул различных веществ, или молекул, обладающих разными скоростями, или по другим свойствам, которые благодаря этому разделению становятся наблюдаемыми макроскопическим образом.

Применяя термин «порядок» в этом значении, мы можем сказать, что макроскопически неразличимые конфигурации обладают *одинаковой степенью порядка*. То, что выше было названо соответствующим макросостоянием конфигураций, можно, следовательно, назвать также классом конфигураций, *обладающих одинаковым порядком*. Степень беспорядка можно тогда измерять числом конфигураций, относящихся к этому классу, то есть если для данной конфигурации *a* существует много конфигураций, обладающих такой же степенью порядка, то это значит, что степень порядка данной конфигурации *a* низка. Наоборот, высокая степень порядка означает, что класс конфигураций, обладающих одина-

ковой степенью порядка, невысок. Следовательно, понятию *порядок* можно дать определение с помощью понятия *одинакового порядка*, которое в свою очередь можно свести к понятию *макроскопически неразличимого*, а степень порядка задается объемом класса конфигураций, обладающих одинаковым порядком. На этих представлениях основывается интерпретация энтропии как меры беспорядка или нечто противоположного мере порядка, поскольку энтропия и вероятность состояния возрастают с увеличением числа конфигураций, которые относятся к данному состоянию.

§ 10. Детерминистская интерпретация термодинамической статистики

Содержание предыдущего параграфа показывает, что вероятности в теории газов определяются путем подсчета конфигураций, соответствующих одному и тому же классу упорядоченности. Было указано, что данный метод предполагает, как уже объяснялось, равновероятность всех конфигураций. Следовательно, вероятностные расчеты статистической механики основываются на предположении относительно *метрики начальных вероятностей*. Данное предположение носит эмпирический характер, и это следует из того, что изменения в состоянии газа мы рассматриваем как контролируемые этой вероятностью и надеемся, что поведение газа отражает предполагаемую метрику вероятностей: большую часть времени газ находится в более вероятных состояниях и редко в менее вероятных. Только благодаря этому переходу к величинам временных интервалов мы можем утверждать, что вероятности макросостояний представляют собой определяющий фактор, который объясняет переход от более упорядоченных состояний к менее упорядоченным состояниям.

Если хотят получить статистические выводы, использование метрики начальных вероятностей неизбежно. И все же вопрос о том, как объяснить эту метрику, всегда волновал статистиков. Математики, склонные к рационалистическому образу мышления, подобно Лапласу, полагали, что начальную метрику можно обосновать априорным образом. Хотя некоторые современные логики еще придерживаются подобных представлений,

однако большинство исследователей статистиков и физиков-теоретиков отвергают любое априорное введение равных вероятностей. Если предположить, что ячейки равных размеров в пространстве скоростей равновероятны, то физики станут утверждать, что такое предположение не должно основываться на каком-либо логическом аргументе, как, например, принцип неразличимости, но может быть обосновано только путем экспериментальной проверки его следствий. И они могут сослаться на тот факт, что в области очень низких температур эта метрика оказывается недействительной и заменяется другой метрикой. То есть метрика вероятностей может быть введена только как физическая гипотеза и нуждается в подтверждении с помощью наблюдений. Однако следует отметить, что в этом случае логическая теория подтверждения в свою очередь не должна основываться на предположении о метрике вероятностей. Некоторые современные логики забывают сделать этот вывод.

И все же для физика остаются другие, неаприористские способы объяснения метрики вероятностей еще до того, как будут проведены специальные испытания. Он может пытаться вывести эту метрику вероятностей из причинных законов физики. Это был бы вполне оправданный метод, потому что он просто переносит результаты одной области опыта в другую. Такие выводы могут быть сделаны двумя способами. В случае первого способа исходят из ряда предположений, которые включают, в дополнение к причинным законам, некоторые общие вероятностные предположения, не содержащие в себе метрику вероятностей. Таким, например, является вывод, предложенный Анри Пуанкаре для равновероятности, имеющей место в азартных играх[1]. В случае второго способа исходят только из причинных законов и не включают никаких вероятностных предположений. Вывод, полученный для метрики вероятностей кинетической теории газов, первоначально был выводом, осуществленным первым способом. Но физики-теоретики стремились полностью исключить вероятностные предположения. Мы скоро увидим, насколько близко они сумели подойти к этой цели. Рассмотрим эти выводы

---

[1] См. H. Reichenbach, The theory of Probability, § 69.

более подробно, начиная с рассмотренного вывода, осуществляемого первым способом. Такие исследования были предприняты Л. Больцманом, Д. К. Максвеллом, Д. У. Гиббсом и другими. В этих математических исследованиях обычно прибегают к более общим формам описания состояния молекул. Молекула рассматривается как система, характеризуемая m параметрами положения $q_1 \ldots q_m$ и также m параметрами импульса $p_1 \ldots p_m$. Эти параметры включают не только скорость и пространственную локализацию, но, возможно, и другие величины, характеризующие вращение молекулы. Для некоторого объема газа общее число параметров будет $N = 2mn$, где n — число всех молекул в газе, и это число может быть представлено через $x_1 \ldots x_N$. Мгновенное состояние газа, его микросостояние, можно изобразить геометрически как точку в N-мерном *фазовом пространстве*, образованном обобщенными координатами $x_1 \ldots x_N$. Так как микросостояния молекул постоянно изменяются во времени, то фазовая точка перемещается в фазовом пространстве. Допустим, что газ образует замкнутую систему, то есть систему, не испытывающую никаких взаимодействий со своим окружением. Тогда траектория фазовой точки, то есть переход системы от одного микросостояния к другому, точно определяется законами механики. Это означает: для некоторой заданной точки в фазовом пространстве существует одна определенная траектория, начинающаяся в этой точке; эта траектория зигзагообразно простирается в пространстве.

Причинными законами, определяющими траекторию, являются уравнения движения; они обычно записываются в *канонической* форме, введенной У. Р. Гамильтоном:

$$\frac{\partial H}{\partial p_i} = \frac{dq_i}{dt}, \quad \frac{\partial H}{\partial q_i} = -\frac{dp_i}{dt}. \qquad (1)$$

Эти дифференциальные уравнения в частных производных, где H — гамильтониан, представляющий полную энергию системы, являются обобщенной формой основных законов механики Ньютона.

Одним из следствий уравнений (1) является сохранение энергии. Поскольку это условие выражается соотношением вида

$$E(x_1 \ldots x_N) = \text{const}, \qquad (2)$$

то оно определяет N-1-мерную гиперповерхность в фазовом пространстве, а именно энергетическую поверхность, перемещением по которой и ограничено движение фазовой точки.

Метод, примененный Больцманом и другими, можно ныне охарактеризовать следующим образом. Хотя траектория, начинающаяся с некоторой заданной точки, строго определяется уравнениями (1) для любого промежутка времени, она рассматривается как такая последовательность состояний, которая обладает свойствами вероятностной последовательности. При своем продолжении она с равной вероятностью проходит через все возможные состояния и бо́льшую часть времени находится в неупорядоченных состояниях. Таким образом, мы встречаем здесь сочетание детерминистской и вероятностной интерпретации физических явлений: траектория фазовой точки в одно и то же время образует строго предсказываемую причинную цепь и, кроме того, вероятностную последовательность. Фазовая точка перемещается зигзагообразно по энергетической поверхности, подобно игрушечному автомобилю, который, будучи заведенным, бегает по определенным кривым по полу; и, подобно этой маленькой игрушке, эта фазовая точка обязана изменением своего положения скрытым законам, которые неизбежно перемещают ее с места на место.

Рассмотрим более детально эту вероятностную последовательность, пробегаемую фазовой точкой. Здесь мы имеем непрерывную вероятностную последовательность. Такие последовательности обладают последействием [1], то есть если фазовая точка находится в точке x, то в следующий момент времени она должна находиться в окрестности точки x. Но постепенно фазовая точка перемещается в более отдаленные области энергетической поверхности; и если нам неизвестно ее прошлое, мы можем надеяться обнаружить точку в любом месте на поверхности. Поэтому для такой последовательности существует *полная вероятность*, имеющая определенное значение для любого участка энергетической поверхности, то есть по отношению к любому участку A поверхности мы можем говорить о вероятности обнаружения

[1] См. § 12 данной книги.

фазовой точки внутри A, безотносительно к ее предыдущим положениям.

Используя терминологию теории вероятностей, мы назовем энергетическую поверхность *атрибутивным пространством* последовательности. Если описывать эту поверхность ортогональными координатами $y_1 \ldots y_{N-1}$, которые расположены на поверхности, в атрибутивном пространстве мы получим функцию распределения

$$\varphi(y_1 \cdots y_{N-1}), \tag{3}$$

интеграл от которой по некоторой площади A дает собою полную вероятность обнаружения фазовой точки на этом участке. Если A есть макросостояние, как определено выше, тогда мы можем определить его энтропию интегрированием (3). То, что мы называем микросостоянием a, или конфигурацией, представляется очень малой областью $dy = dy_1 \ldots dy_{N-1}$, его вероятность получается из (3) просто перемножением с dy, и поэтому она весьма мала. Однако микросостояния, имеющие одинаковую степень упорядоченности, могут занимать большие участки. То, что неупорядоченные состояния обладают большой вероятностью, означает, что они занимают очень большую область энергетической поверхности.

Вероятностные последовательности, которые задаются правилом, определяющим атрибут для каждого элемента, не представляют трудностей для теории вероятностей [1]. Примеры дискретных последовательностей этого вида анализировались неоднократно. Например, для значений квадратных корней из целых положительных чисел с точностью до 5-го десятичного знака показано [2], что последняя цифра с одинаковой вероятностью может быть четной или нечетной, или, другими словами, что полная вероятность для четных цифр равна $1/2$. Однако последовательности, образуемые из этих цифр, имеют последействие, то есть для больших целых чисел мы по-

[1] ThP, p. 132, 165, 343.

[2] См. György Pólya, Gábor Szegő, Aufgaben und Lehrsätze aus der Analysis, Bd. I, Berlin, I. Springer, 1925, S. 72 и 238. Анри Пуанкаре сделал аналогичное замечание относительно таблицы логарифмов, но Франель показал, что данное Пуанкаре доказательство не верно (см. Pólya, Szegő, op. cit., S. 73 и 239). В свое время, не зная этого, я приводил пример Пуанкаре в ThP, p. 343. Этот пример нужно заменить рассмотренным выше примером с квадратными корнями.

лучаем все более длинные ряды четных чисел и такие же ряды для нечетных чисел, потому что начальные выражения десятичных знаков начинают изменяться довольно медленно при переходе от одного целого числа к другому.

Если вероятностная последовательность задается некоторым правилом, то, конечно, нужно требовать, чтобы распределение вероятностей выводилось из этого определяющего правила, как это было сделано в приведенном выше примере. Поэтому физики предпринимали много попыток доказать, что уравнения движения (1) приводят к метрике вероятностей, принятой в кинетической теории.

Ядро всех этих доказательств образует теорема, названная в честь французского математика Лиувилля,

Рис. 13. Отображение в фазовом пространстве области A на область A', осуществляемое посредством фазовых траекторий.

который в 1838 году исследовал условие несжимаемости для жидкости. Допустим, что мы имеем область A в фазовом пространстве; из каждой точки области A мы строим траекторию фазовой точки для интервала времени Δt, одного и того же для всех точек. Тогда концы траектории будут находиться внутри области A' (рис. 13). Мы можем сказать: A отобразилось в A' посредством фазовой траектории за интервал времени Δt. Так как существует только одна траектория, начинающаяся в некоторой заданной точке области A, то отображение полностью определяется математическими соотношениями, содержащимися в уравнениях (1). Мы можем

также дать отображению физическую иллюстрацию: мы считаем, что каждая точка в A представляет отдельную изолированную систему в момент времени t; после того как каждая система претерпит ряд изменений в конфигурационных расположениях молекул к моменту времени $t + \Delta t$, все системы заполнят область A'. Используя уравнения (1), можно будет показать, что области A и A' имеют равные объемы[1]. Уподобляя движение всех этих фазовых точек потоку жидкости, мы можем сформулировать теорему Лиувилля как утверждение о том, что эта жидкость несжимаема и что при движении ее части сохраняют свои объемы. В данном физическом примере это означает, что в своих математических взаимосвязях ход изменений в совокупности отдельных изолированных систем имеет сходство с потоком в несжимаемой жидкости.

Эта теорема приводит к важному заключению относительно траектории только одной фазовой точки. Предположим, что траектория входит в область A и остается в ней в течение периода времени t_A. Далее, мы знаем, что если точка находится в A, то по прошествии промежутка времени Δt она окажется в A'; она будет затем находиться в A' в течение такого же времени, как и в A. То есть мы имеем $t_A = t_{A'}$. Если фазовая точка входит в область A и выходит из нее повторно, то то же самое будет иметь силу для области A'; за каждым пребыванием в A следует пребывание в A' по прошествии времени Δt. Следовательно, сумма всех интервалов времени t_A будет равна сумме интервалов времени $t_{A'}$. Так как, согласно теореме Лиувилля, A и A' имеют равные объемы, мы заключаем, что в равных объемах фазовая точка находится в течение равных периодов времени.

Это равенство *времен пребывания* дает основания для определения метрики вероятностей в теории газов. Например, оно приводит к равновероятности ячеек в пространстве скоростей, что в свою очередь приводит к малым значениям вероятностей для упорядоченных состояний, которые занимают только небольшой объем в фазовом пространстве. Таким образом, показано, что

[1] Доказательство можно найти в любом учебнике по статистической механике. См., например, Richard C. Tolman, The Principles of Statistical Mechanics, Oxford, the Clarendon Press, 1938, p. 49.

принятое Больцманом условие равновероятности является следствием уравнений движения.

Однако полученный результат не может быть использован без некоторых важных дополнений. Во-первых, мы должны предположить, что фазовая точка в конечном счете проходит через каждый элемент объема, каким бы малым мы его ни выбрали. Иначе могли бы существовать элементы объема, равные по размерам A, которые никогда не будут достигнуты эргодической траекторией, выходящей из A, и для которых нельзя утверждать равенство времен пребывания. Кроме того, мы должны предположить, что полное время пребывания Σt_A в A сохраняет определенное отношение к t, то есть, что

$$\lim_{t \to \infty} \frac{\sum t_A}{t} = \text{const}. \tag{4}$$

Это предположение позволяет нам высказаться о вероятности того, что фазовая точка окажется внутри области A; оно определяет эту вероятность, по аналогии с частотной интерпретацией вероятности, как предел отношения временных периодов [1]. Постоянная в (4) должна быть одной и той же для каждой траектории, проходящей через A. Такое предположение необходимо, потому что отдельные отрезки траекторий, пересекающие A, могут относиться к различным траекториям.

Это предположение было введено Больцманом под названием эргодической гипотезы (ergos — энергия, odos — траектория). Больцман следующим образом сформулировал эту гипотезу: фазовая точка в конечном счете проходит через каждую точку энергетической поверхности. Легко показать, что эта формулировка несостоятельна. П. и Т. Эренфесты заменили ее другой формулировкой, заключающейся в том, что траектория проходит близко к каждой точке в пределах любого выбранного нами небольшого расстояния ε, которое больше нуля.

Остается еще вопрос о том, должна ли эргодическая гипотеза рассматриваться как независимое предположение или же ее можно вывести из канонических уравнений (1), как это имеет место с теоремой Лиувилля. Если справедливо первое альтернативное предположение, мы

[1] Относительно этого определения вероятности для непрерывных последовательностей см. ThP, p. 240.

будем тогда иметь здесь вывод первого вида, как упоминалось выше, в котором вероятностное предположение, эргодическая гипотеза, добавляется к причинным законам, а именно к теореме Лиувилля, для того чтобы вывести метрику вероятностей. Если справедливо второе предположение, мы будем иметь вывод метрики вероятностей только из причинных законов.

Эта проблема долгое время занимала математиков. В конце концов она была блестяще решена Джоном

Рис. 14 Траектория биллиардного шара, отражаемого от стенок, которая покрывает в конечном счете всю поверхность биллиардного стола.

Рис. 15 Особый случай траектории биллиардного шара, представляющей замкнутый цикл.

Нейманом и Георгом Биркгофом [1], которые сумели доказать, что справедливо второе альтернативное предположение. Идею этого вывода эргодической гипотезы из уравнений движения можно проиллюстрировать следующими соображениями. Предположим, что биллиардный шар толкнули из точки A (рис. 14) в направлении к точке B. Допустим, что трение отсутствует, тогда шар, отразившись от стенки в точке B, продолжает свой путь без остановки. Вследствие дальнейших отражений от

[1] См. особенно шесть статей, опубликованных в «Proceedings of the National Academy of Sciences», vol. 17 (1931) и vol. 18 (1932).

стенок шар будет двигаться по зигзагообразной траектории, которая в конце концов покроет всю поверхность биллиардного стола. Это явилось бы частной и весьма простой иллюстрацией теоремы Неймана и Биркгофа, где фазовой точке соответствует шар, а энергетическая поверхность — биллиардному столу.

Только в случае специфических начальных условий может оказаться, что траектория шара покрывает не всю поверхность, а некоторую ее часть, или субповерхность (рис. 15). Однако Нейман показал, что в этом случае его теорема применяется к этой отдельно взятой субповерхности, которая снова будет покрываться с равной плотностью фазовой точкой [1].

После теоремы Неймана и Биркгофа детерминистическая физика достигла высшей степени совершенства: было показано, что строгий детерминизм в элементарных процессах ведет к статистическим законам для макроскопических явлений, метрика вероятностей которых представляет отражение причинных законов, управляющих движением элементарной частицы. К сожалению, этот триумф классической механики был достигнут в то время, когда она уже была заменена квантовой механикой, которая не признавала детерминизма элементарных процессов. Но отложим обсуждение квантовой физики и продолжим рассмотрение классического решения статистических проблем в теории газов. Мы увидим, что ясные представления о классической форме статистики

---

[1] Иногда предположение (4) называют *эргодической теоремой*, и, таким образом, его отличают от *эргодической гипотезы*, согласно которой фазовая точка проходит через каждый элемент объема. Справедливость эргодической гипотезы для всего заданного пространства зависит от начальных условий. В этом случае данное Нейманом доказательство можно сформулировать следующим образом: всегда существует некоторое субпространство, для которого справедлива эргодическая гипотеза, и для этого субпространства всегда удовлетворяется эргодическая теорема. Поскольку эргодическая теорема удовлетворяется тривиальным образом для элементов объема, через которые не проходит фазовая точка, и поскольку постоянная в (4) будет равна 0, то эту теорему можно не ограничивать условиями. Заметим, что если бы даже метрика вероятностей в фазовом пространстве была постоянной, она выражалась бы для энергетической поверхности другой функцией — эргодической плотностью σ. Причина этого состоит в том, что расстояние между соседними энергетическими поверхностями не является постоянным. Можно показать, что σ представляет величину, обратную градиенту энергии.

газов необходимы для того, чтобы оценить вклад квантовой механики в решение проблемы времени.

Однако даже в классической физике можно подвергнуть сомнению детерминистическое решение статистики газа. Доказано только то, что статистические законы, управляющие макрокосмом, *совместимы* со строгими законами, управляющими атомарными явлениями; но это не означает, что детерминистическое предположение дает *единственное* решение проблем статистики газов. Вполне можно допустить, что атомарные явления строго не предсказуемы на основе причинных законов, в то время как макроскопические явления, охватывающие большое число молекул, контролируются статистическими законами с высокой степенью точности. Можно последовательно провести такую концепцию, не вступая в противоречие с законами классической физики; фактически мы обращаемся к этой концепции всякий раз, когда интересуемся индивидуальными явлениями термодинамики.

Рассмотрим, например, применение эргодической теоремы к некоторой замкнутой физической системе, исследуемой в лаборатории. Мы знаем, что нельзя получить совершенную тепловую изоляцию. Поэтому поведение фазовой точки системы не будет определяться только уравнениями движения (1); она будет постоянно подвергаться беспорядочным небольшим воздействиям извне. Хотя эти случайные воздействия дают вначале слабый эффект, но со временем они заметно изменяют траекторию точки; практически они будут выталкивать точку из ограниченной субповерхности всякий раз, когда окажется, что фазовая точка займет исключительное положение p_0. По этим причинам, например, процесс, изображенный на рис. 15, никогда не может продолжаться в течение длительного времени; даже если бы мы смогли в точности создать необходимые начальные условия, воздействия извне вынудят шар покинуть свою специфическую траекторию, которая затем примет общий вид, представленный на рис. 14.

Поэтому именно в результате случайных воздействий извне последовательность положений, пробегаемая фазовой точкой, не является строго детерминированной; и все же последовательность может еще быть вероятностной последовательностью с последействием, удовлетворяющей всем условиям, требуемым для результатов

теории газов. Небольшие возмущения со стороны окружения, подчиненные только законам случая, создают трудности для использования уравнений Гамильтона; но они не нарушают статистических результатов, выводимых из этих уравнений, потому что вероятностная последовательность, или вероятностная цепь, может обладать всеми желательными статистическими свойствами, не имея никаких детерминистических соединительных звеньев между своими последовательными элементами. Все вероятностные последовательности, представленные индивидуальными физическими системами, оказываются последовательностями этого индетерминистического вида, потому что не существует полностью изолированных систем.

Можно доказывать, что, если существуют воздействия извне, их можно включить в рассматриваемую систему путем расширения ее внешних границ. Тогда возмущения будут вызываться изменениями параметров, контролируемыми законами движения, и, следовательно, правилами, управляющими эргодической траекторией. Единственное различие будет состоять в том, что наряду с новыми параметрами вводится некоторое добавочное число измерений для фазового пространства; но траектория системы снова будет строго определенной. Если будут дальнейшие возмущения извне, то это расширение может быть продолжено. В конце концов мы придем, таким образом, к системе, которая окажется идентичной со всей вселенной. Наиболее объемлющая система полностью изолирована и, следовательно, контролируется строгими причинными законами, не нарушаемыми воздействиями извне.

Такова гипотеза, выдвигаемая детерминизмом. Мы должны теперь исследовать, какие аргументы можно привести за и против детерминизма.

§ 11. Детерминизм versus индетерминизм в классической физике

Детерминизм не есть наблюдаемый факт. Он представляет собой теорию, которая выводится из наблюдений путем экстраполяции. Наблюдения приводят к физическим законам, делающим возможным научное предвидение. Поэтому предполагается, что, если бы мы смогли проводить наблюдения более тщательно, наши

предсказания оправдывались бы без исключений. Допускается, что за наблюдаемыми взаимоотношениями скрывается система первичных причинных связей. Таким образом, детерминизм основывается на распространении представлений о наблюдаемых регулярностях на области ненаблюдавшегося, и при этом предполагается, что наблюдающееся несовершенство предсказаний исчезло бы, если бы мы только смогли раскрыть первичные причинные структуры.

Проанализируем более подробно этот вывод. Предположим, что мы хотим предсказать место падения снаряда, выпущенного из орудия. Мы определяем направление ствола орудия, величину порохового заряда, массу снаряда и т. д.; и, используя эти величины, как заданные *параметры*, или *начальные условия*, процесса, мы вычисляем на основе физических законов точку падения снаряда. Хорошо известно, что такое предсказание не является достаточно надежным. Почему же оно ненадежно?

Потому, говорим мы, что в наших вычислениях не учитываются все факторы, оказывающие влияние на полет снаряда. Мы только можем сказать: если учитываемые параметры обладают принятыми при расчете значениями и если никакие другие параметры не оказывают никакого влияния на этот процесс, снаряд попадет в предсказанное место, то есть мы используем предположение о причинности, выраженное в *условной* форме. Но мы очень хорошо знаем, что эта форма недостаточна для того, чтобы гарантировать правильность предсказания. Условия «если» полностью не будут удовлетворяться, измеренные значения параметров будут верны лишь с некоторой степенью точности, и будут существовать другие параметры, опущенные при вычислениях, которые могут оказывать влияние на данное явление. Некоторые из них можно было бы учесть при более точных вычислениях. Например, мы могли бы принять во внимание вращение Земли и действие ветра. Многие параметры нельзя учесть, так как человеческие способности к наблюдению ограничены. Например, из глубин вселенной мог бы прилететь метеор и столкнуться со снарядом на его пути, тогда наше предсказание стало бы абсолютно неверным.

Не помогло бы нам и ограничение нашего предсказания выдвижением условия: «Если не произойдет никаких

нарушений». Предсказания, ограниченные таким образом, не имеют практической пользы, если мы не знаем, что нарушения и другие отклонения представляют исключительные явления. Мы хотим быть уверенными, что большинство предсказаний, основанных на учитываемых параметрах, правильны, то есть что любое такое предсказание весьма и весьма вероятно. Поэтому мы должны дополнить условную форму причинности вероятностной гипотезой, согласно которой мы можем таким образом произвести учет параметров, что они позволят сделать предсказание с высокой степенью вероятности даже и в том случае, когда условие «если» будет выполняться не строгим образом. Только в своем сочетании с вероятностной гипотезой условная форма причинности представляет собой содержательное утверждение, которое можно проверить путем наблюдений. В этом смысле идея вероятности необходима даже для классической физики.

На основе большого опыта мы знаем, что эта вероятностная гипотеза верна, и мы также знаем, что мы можем повысить вероятностную оценку предсказания, включая в вычисление все большее и большее число параметров. Кроме того, часто приходится заменять используемые физические законы на более точные. Например, законы, которым подчиняются процессы трения, испытываемые снарядом в воздухе, носят приближенный характер и допускают уточнения. Исходя из таких соображений, мы убеждаемся, что причинность можно сформулировать в *безусловной форме*. Мы предполагаем, что физически возможно узнать все параметры процесса и первичные физические законы, хотя это технически и невозможно. Затем, используя условную форму, мы делаем вывод, что при этом предположении мы можем сделать предсказание с вероятностью 1, то есть практически достоверное. Но именно безусловная форма причинности выражает *детерминизм*, и мы должны ее проанализировать. При этом мы будем исходить из предположения, что условная форма по крайней мере правильна и остается таковой в рамках классической физики. В квантовой физике отказались даже от условной формы, но мы отложим рассмотрение этой стороны проблемы до главы V.

Вследствие того что при определении детерминизма необходима ссылка на первичные начальные условия и

первичные законы, детерминизм представляется предельным утверждением. Как таковое он имеет смысл тогда, когда его можно преобразовать в конвергентные утверждения, относящиеся к действительным наблюдаемым и действительно известным законам. Уже отмечалось, что эти конвергентные утверждения включают понятие вероятности, которое необходимо точно сформулировать.

Если $D^{(1)}$ описывает физическое состояние пространственного объема v в момент времени t_1, мы можем предсказать с определенной вероятностью $p^{(1)}$, что объем v окажется в некотором состоянии, описываемом E в более поздний момент времени t_2. Для того чтобы эта вероятность приняла большее значение $p^{(2)}$, мы заменяем $D^{(1)}$ более точным описанием $D^{(2)}$, которое также относится к моменту времени t_1. Это новое описание может отличаться от $D^{(1)}$ в любом или во всех из следующих отношений:

1. Новое описание включает начальные граничные условия объема v, то есть оно относится к расширенному пространственному объему V. Однако описание E все еще относится к первоначальному объему v.

2. Новое описание относится к более точным измерениям параметров и использует внутренние параметры системы, входящей в v, которыми раньше пренебрегали; например, параметры, относящиеся к внутреннему состоянию молекул.

3. Новое описание использует более совершенные причинные законы.

Эту процедуру можно продолжить. Тогда описания $D^{(i)}$ будут изменяться от описаний макросостояния к описаниям микросостояния. В классической физике обычно предполагается, что в отношении п. 2 и 3 мы быстро достигаем конечной стадии, которая приводит к предельной точности, и что только п. 1 содержит трудности фундаментального характера. Полагали, что первичные законы найдены в уравнениях движения, а первичные параметры — в механической модели атома. Интересно, что еще в 1898 году Больцман подверг сомнению эту веру в отношении п. 2 и 3[1].

[1] Л. Больцман, Лекции по теории газов, М., 1956, стр. 529. Приводим соответствующий отрывок: «Так как у нас любят сейчас представлять время, когда наши воззрения на природу станут совершенно иными, мне хочется еще упомянуть, что основные урав-

В наше время едва ли кто согласится с предположением, что найдены первичные законы или установлена окончательная модель атома. Развитие как небесной механики, так и механики атома настроило нас скептически. Поэтому, даже не обращая внимания на статистический характер квантовой физики, мы должны рассмотреть все три возможности улучшения заданного описания.

В процессе постоянного совершенствования описания, переходя от $D^{(1)}$ к $D^{(2)}$ и далее ко все более и более точным описаниям, мы получаем последовательность конвергентных описаний. Детерминизм теперь можно сформулировать как гипотезу о том, что в этой последовательности вероятность предсказания более позднего состояния E возрастает, приближаясь к 1, и что существует предельно (ultimate) точное описание D, к которому сходятся приближенные описания. Эта гипотеза выражается схемой:

$$\begin{array}{l} D^{(1)}\ D^{(2)}\ D^{(3)}\ \ldots\ldots\ldots \to D \\ p^{(1)}\ p^{(2)}\ p^{(3)}\ \ldots\ldots\ldots \to 1. \end{array} \qquad (1)$$

В этой схеме описание E объема v в момент времени t_2 остается неизменным. Но, каким бы точным ни хотели мы сделать описание E, детерминизм требует, чтобы существовала сформулированная в (1) конвергентная схема.

Рассмотрим более подробно эту гипотезу. Даже в классической физике, в которой п. 2 и 3 считались решенными, п. 1 всегда вызывал трудности. Если вселенная пространственно бесконечна, то объем V будет увеличиваться с каждым новым описанием $D^{(i)}$, и предельное начальное описание D должно выражать состояние всего бесконечного пространства в момент времени t_1. Таким образом, нам понадобится бесконечное число параметров, для того чтобы сформулировать D. (В данном случае безразлично, к какому моменту времени t относится это поперечное сечение вселенной; «начальное»

нения движения для отдельных молекул могут оказаться лишь приближенными формами, дающими средние значения, которые вытекают, согласно исчислению вероятностей, из совместного действия очень большого количества отдельных движущихся частиц, составляющих окружающую среду...»

просто означает как соответствующее некоторой «выбранной исходной точке в причинной цепи».) Но тогда в высказываниях о предельном описании D не содержится много смысла, так как бесконечное число независимых параметров не может быть объединено в едином описании состояния.

Теория относительности Эйнштейна видоизменила это логическое затруднение в двух направлениях. Во-первых, космология Эйнштейна допускает, что вселенная может быть пространственно ограниченной; этот результат позволяет делать утверждения о предельном описании D, насколько это касается п. 1. Но, даже если бы предположение Эйнштейна об ограниченной вселенной оказалось несостоятельным, его теория дает и другие основания для того, чтобы считать рассматриваемый объем V конечным. При рассмотрении рис. 8 отмечалось, что световой конус, имеющий вершиной событие A, включает все события, которые когда-либо могли причинно воздействовать на A. Задав небольшой объем v и момент времени t_2, мы можем, следовательно, вычислить наибольший объем V_1 таким образом, чтобы события в V_1 в момент времени t_1 могли причинно воздействовать на состояние v в момент времени t_2. Событиями вне объема V_1 можно пренебречь, потому что потребовалось бы слишком много времени, для того чтобы причинное воздействие, распространяющееся от них, достигло v: такие воздействия достигли бы v после момента времени t_2. Поэтому достаточно, если предельное описание D относится к V_1; и нам не нужно знать всех параметров, определяющих все необъятное пространство.

Однако такая постановка проблемы также имеет свои недостатки. Если наблюдатель находится в v, то он не сможет раньше момента времени t_2 получить исчерпывающее знание состояния V_1 в момент времени t_1. Передача любой информации, идущей от границы V_1 к наблюдателю, ограничена скоростью света. Таким образом, релятивистский наблюдатель, по существу, находится в таком же затруднительном положении, как и наблюдатель в классической физике: тогда как последний не в состоянии в момент времени t_1 получить сведения о бесконечном пространстве, наблюдатель в релятивистской физике не может в момент времени t_1 получить сведения о той ограниченной части пространства,

которые необходимы для предсказаний в отношении некоторого события в момент времени t_2.

Последний вывод остается в силе по крайней мере для бесконечной релятивистской вселенной. Можно добавить, что для такой вселенной величина объема V_1 зависит от разности времен $t_1 - t_2$; если эта разность увеличивается, то потребуются сведения о большем объеме V_1. Напротив, если верно предположение Эйнштейна о пространственно ограниченной вселенной, то полный объем пространства представляет собой тот максимум, который является пределом для V_1. Тогда физически было бы возможным в момент времени t_1 иметь сведения о состоянии всей вселенной в некоторый более ранний момент времени t_0, хотя этот момент времени t_0 и мог бы быть миллиарды лет назад. Если станет известным это раннее состояние, то в детерминистском мире могло бы быть предсказано любое будущее состояние. Следовательно, для пространственно ограниченной вселенной ни в классическом, ни в релятивистском случае п. 1 не будет представлять логических затруднений по вопросу о существовании предельного описания, так как для этого описания не требуется бесконечного числа параметров.

Остается рассмотреть вопрос о физических предположениях, которые содержатся в том допущении, что возможно предельное описание состояния вселенной. Если наблюдатель находится в некоторой точке x в момент времени t_1, мы должны допустить, что каким-то образом из всех точек вселенной в момент t_0 испускаются сигналы, достигающие x в момент t_1. Точечное событие $\{x, t_1\}$ будет тогда содержать характеристику всей вселенной в момент времени t_0. Такая возможность представляется в высшей степени невероятной. Но нам и не нужно анализировать этот вопрос, потому что пространственная ограниченность вселенной проблематична. Мы лучше попытаемся дать формулировку детерминизма таким образом, чтобы ее можно было применять и к пространственно неограниченному миру. Тогда будет включена и возможность конечного пространства огромной протяженности.

Напрашивается следующая формулировка, которая уже неоднократно обсуждалась. Начальные условия предельного описания D относятся только к конечному

пространственному объему v, но в дополнение мы требуем, чтобы условия на границе объема v были известны в течение всего промежутка времени от t_1 до t_2. Эту формулировку можно проиллюстрировать рис. 16, на котором вертикальные линии изображают время, а горизонтальные линии представляют пространственные объемы. Объем v в момент времени t_1 простирается от точки *1* до точки *2*; в момент времени t_2 он представлен линией от точки *3* до точки *4*. Границы v в течение периода времени от t_1 до t_2 даны линиями *1—3* и *2—4*.

Рис. 16. Определение состояния пространственного объема v в момент времени t_2 на основе данных о состоянии пространственного объема V в момент времени t_1 и условий на границе V в течение времени от t_1 до t_2.

Так как все влияния со стороны пространственного окружения должны проходить через эти границы, все остающееся бесконечное пространство здесь заменяется конечной областью и, таким образом, становится возможным его описание посредством конечного числа параметров, насколько это касается рассматриваемой системы. Таким же образом устраняется трудность, с которой сталкивается теория относительности, а именно получение информации относительно большого объема V_1. Это происходит потому, что объем v можно выбрать достаточно малым. Несомненно, что описание D в новой интерпретации не позволяет нам в момент времени t_1 делать предсказания, потому что сведения о том, что произойдет позднее на границах объема, могут быть получены только в более поздние моменты времени. Но новый метод позволяет нам по крайней мере сформулировать гипотезу детерминизма даже для бесконечной

в пространстве вселенной, безразлично какой: классической или релятивистской.

Однако, используя эту формулировку, мы сталкиваемся со следующей трудностью. Когда мы рассматриваем достаточно малый пространственный объем v, будущие его состояния легко можно предсказать, исходя из одновременного знания граничных условий, но не из более ранних. Допустим, что мы хотим предсказать погоду в некотором городе. Мы выбираем занимаемую городом площадь как основание нашего начального объема v и над этим основанием мысленно представляем полусферическую поверхность. Это будет границей нашей системы. Было бы легко предсказать, какая погода будет через десять лет, если бы мы знали метеорологические условия на нашей граничной поверхности для момента предсказания. Наоборот, начальные условия, выражающие данные о сегодняшней погоде в городе, не будут иметь значения. Однако наше так называемое предсказание погоды не было бы долгосрочным прогнозом, а явилось бы просто интерполяцией для города одновременных метеорологических данных на граничной поверхности. Чтобы произвести эту интерполяцию, нам не нужны представления о детерминистском мире.

Чтобы избежать такого заключения, мы сформулируем следующее дополнительное требование. Задав пространственный объем v, мы окружим его большим объемом V, который опишем посредством $D^{(1)}$ в момент времени t_1 и который на рис. 16 обозначен линией *5—6*. Для того чтобы предсказать состояние E малого объема v в более поздний момент времени t_2, мы включаем в $D^{(1)}$ граничные условия объема V в течение времени от t_1 до t_2. На рис. 16 эти границы представлены линиями *5—7* и *6—8*. Когда мы говорим, что состояние E объема v можно предсказать с вероятностью $p^{(1)}$, то при этом подразумеваем, что такое предсказание можно получить, каким бы большим ни был выбран нами объем V. Например, когда мы говорим, что можно предсказать погоду в городе (объем v) на следующий день с вероятностью 80 процентов, мы имеем в виду, что это можно сделать, если известны данные о сегодняшней погоде на всем континенте и сверх того на воображаемой полусферической поверхности (граница объема V) и если данные о погоде на границе этой полусферической

оболочки известны вплоть до завтрашнего дня. Метеоролог, конечно, скажет, что это не прогноз погоды, ибо дополнительные данные, относящиеся к границам континента, не будут доступны раньше завтрашнего дня; и он добавит, что прекрасно может обойтись без них, потому что фронт шторма передвигается медленно, и что состояние погоды в отдаленных местах не окажет большого влияния на завтрашнюю погоду в городе. К тому же воображаемая полусфера заходит в такие области атмосферы, которые не оказывают влияния на погоду. Все это верно, но дальнейшее показывает, почему мы пользуемся этой формулировкой. Предсказание, основанное на данных о большем объеме V, практически эквивалентно предсказанию, основанному на данных о бесконечном объеме, или предсказанию, исходящему из данных об объеме V_1, характерного для релятивистской вселенной, даже если объем V намного меньше объема V_1. Таким образом, влияние граничных условий, в сущности, исчезает. Новая формулировка имеет то преимущество, что она не требует описания бесконечного или необъятного объема V_1. Вместо этого она использует описание не слишком большого объема V и условий, существующих в течение ограниченного времени на его границе. По этой причине она позволяет нам дать формулировку детерминизма [1].

Применяя эти результаты к схеме (1) и включая в $D^{(t)}$ граничные условия объема V между t_1 и t_2, мы следующим образом определяем детерминизм: каким бы большим мы ни выбрали объем V для данного объема v, если V и v сохраняют постоянные значения в последовательности описаний схемы (1), существует предельное описание D, которое позволяет предсказывать состояние E объема v с вероятностью 1. Другими словами, схема

---

[1] Формулировка детерминизма с использованием представлений об ограниченном объеме v и граничных условий обсуждалась неоднократно. Она была дана, например, Moritz Schlick в «Causality in Everyday Life and in Recent Science», в «Univ. Calif. Publ. Philos.», vol. 15 (1932), p. 99—126, перепечатанной Herbert Feigl, Wilfrid Sellars (eds.), Readings in Philosophical Analysis, New York, Appleton-Century-Croffs, 1949, p. 515—533. Однако Шлик здесь не упоминает о дополнительном требовании относительно большего объема V, без которого эта формулировка практически бессодержательна.

(1) символизирует класс схем, каждая из которых относится к определенному объему V.

Сформулируем теперь индетерминизм, по-прежнему не обращая внимания на квантовую физику; мы можем здесь говорить о *классическом индетерминизме*. В этой концепции мы по-прежнему предполагаем, что вероятность предсказания может увеличиваться, стремясь к 1. Но мы отрицаем возможность предельного описания D, то есть мы не предполагаем, что последовательность все более и более точных описаний сходится к пределу. Эта концепция выражается схемой:

$$\begin{matrix} D^{(1)} & D^{(2)} & D^{(3)} & \ldots\ldots \\ p^{(1)} & p^{(2)} & p^{(3)} & \ldots\ldots \rightarrow 1. \end{matrix} \qquad (2)$$

Если описание $D^{(t)}$ ограничено условиями в момент времени t_1, то, как уже объяснялось, бесконечность пространства была бы достаточной причиной для отказа от существования предельного описания D. Но схема (2) применима также и к пространственно ограниченному миру, или к описаниям $D^{(t)}$, которые содержат граничные условия объема V за время от t_1 до t_2. По отношению ко второму примеру мы утверждаем, что схема (2) применима при условии, что мы выберем объем V достаточно большим в сравнении с объемом v, к которому относится предсказание. Это означает, что существует такой объем V вокруг v, что для всех больших, также окружающих v объемов схема (2) остается в силе.

Обсуждая схему (2), следует отметить, что она логически непротиворечива. Из факта существования предела последовательности $p^{(t)}$ невозможно прийти к заключению, что последовательность $D^{(t)}$ также должна иметь предел. В качестве примера рассмотрим две последовательности:

$$\begin{matrix} 1, & 2, & 3, & 4, & \ldots \\ \frac{1}{1}, & \frac{1}{2}, & \frac{1}{3}, & \frac{1}{4}, & \ldots \end{matrix} \qquad (3)$$

Хотя эти две последовательности и находятся во взаимно-однозначном соответствии, только вторая имеет предел, равный 0, тогда как первая не имеет конечного предела.

Конечно, это не доказывает, что последовательность $D^{(t)}$ не имеет предела. Такое доказательство не может быть дано чисто логическим путем, мы должны обратиться к данным физики. Однако, отрицая существование предельного описания D, мы можем обратиться к п. 2 и подвергнуть сомнению предположение о том, что число внутренних параметров системы ограничено. А в отношении п. 3 мы можем сказать, что простые математические соотношения, известные под названием уравнений движения, справедливы только приближенно и что при дальнейших исследованиях они заменятся гораздо более сложными соотношениями, которые в этом процессе совершенствования не обязательно сходятся к предельной формуле. Таким образом, мы сформулировали концепцию, отрицающую детерминизм; и что касается граничных условий, то отрицание предельного описания D выводится не из бесконечности пространства, а базируется на другой основе.

Очевидно, что и детерминизм и индетерминизм представляют собою гипотезы, экстраполяции, выходящие за область наблюдаемого. Какие доказательства можно привести в пользу того или другого?

Так как мы здесь интересуемся гипотезами, относящимися к пределу последовательности, а сам предел никогда недостижим, мы должны пытаться построить доказательство, исходя из конвергентных свойств последовательности. С этой точки зрения напрашиваются следующие соображения. Сходимость вероятностей предсказания очевидна и получается путем индуктивного вывода, что можно рассматривать как достаточно хорошее ее подтверждение. Но обычно забывают, что такого же рода очевидность показывает расходимость вероятностей, как только начинают рассматривать более длинные интервалы времени. Исследуем сначала эту расходимость, используя формулировку, не учитывающую граничные условия. Задав описание $D_1^{(1)}$ начальных условий в v в момент времени t_1, мы сможем предсказать будущее состояние E_2 объема v в момент времени t_2 с большой вероятностью $p_2^{(1)}$; но если мы захотим узнать состояние E_3 объема v в более поздний момент времени t_3, тогда то же самое описание позволит сделать это, но только с более низкой вероятностью $p_3^{(1)}$. Конечно,

$p_3^{(1)}$ можно улучшить с помощью более точного описания $D_1^{(2)}$ для момента времени t_1. Но, с другой стороны, и для $D^{(2)}$ вероятность предсказания уменьшится, когда захотим сделать предсказание для времени t_4, более позднего, чем t_3. Таким образом, мы имеем двойной процесс сходимости и расходимости, который можно выразить следующей схемой, где нижние индексы обозначают соответствующие моменты времени:

$$\begin{array}{ccccccc}
D_1^{(1)} & D_1^{(2)} & D_1^{(3)} & \dots & D_1^{(l)} & \dots & \\
p_2^{(1)} & p_2^{(2)} & p_2^{(3)} & \dots & p_2^{(l)} & \dots & \to 1 \\
p_3^{(1)} & p_3^{(2)} & p_3^{(3)} & \dots & p_3^{(l)} & \dots & \to 1 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
p_k^{(1)} & p_k^{(2)} & p_k^{(3)} & \dots & p_k^{(l)} & \dots & \to 1 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\downarrow & \downarrow & \downarrow & & \downarrow & & \\
p & p & p & \dots & p & \dots &
\end{array} \qquad (4)$$

В таком расположении вероятностей в виде структуры во всех строках вероятности сходятся к 1, а в столбцах — к полной вероятности (или априорной вероятности) p состояния E. Это означает, что если мы вообще сможем сделать какое-нибудь утверждение о вероятности явления E по прошествии длительного времени, мы должны использовать наши знания о вероятности E в целом и не можем основывать наши предсказания на начальном состоянии $D_1^{(1)}$; следовательно, эта вероятность p одинакова для всех $D_1^{(l)}$ и, таким образом, для всех столбцов. Например, когда нас просят предсказать максимальную температуру в один из дней июля в каком-нибудь городе через 100 лет, мы используем среднюю вероятность, вытекающую из общих данных статистики для июля, взятых для этого города; наши знания о погоде в сегодняшний день тогда не будут иметь никакого значения. Так как для значительно удаленных вправо столбцов первые вероятности $p_2^{(l)}$, $p_3^{(l)}$ и т. д. довольно велики, значения вероятностей в этих столбцах уменьшаются, стремясь к p. Именно это мы подразумеваем под расходимостью этих вероятностей.

Теперь нам нужно включить в формулировку о расходимости граничные условия. Вспомним, что в данном

случае схема (4) выражает класс таких схем, каждая из которых характеризуется определенным объемом V, тогда как объем v, к которому относится предсказание, остается постоянным. Каждое $D_1^{(i)}$ включает описание условий на границах V в течение периода времени от t_1 до t_k. Когда мы увеличиваем отрезок времени, на который распространяется предсказание, от t_k до момента времени t_m, то описание $D_1^{(i)}$ следует дополнить условиями на границе V в период от t_k до t_m; мы предполагаем, что эти дополнительные данные будут выражены с помощью тех же параметров и с той же степенью точности, как и те, которые использовались для описания таких данных до момента времени t_k. Поэтому, когда мы движемся вдоль столбца, член «$D_1^{(i)}$» непрерывно меняет свое значение, но вполне определенным образом; фактически он означает систему $\sigma^{(i)}$ отдельных описаний, каждое из которых относится к определенному t_k или t_m, согласно определенным правилам. Следовательно, мы можем содержательным образом говорить о вероятностях $p_k^{(i)}$ и $p_m^{(i)}$.

Если же V очень мало или даже равно v, то столбцы не должны приводить к расходимостям вероятностей, то есть $p_m^{(i)}$ должны быть не меньше, чем $p_k^{(i)}$. Мы можем обратиться к примеру с предсказанием погоды в городе, где граничная поверхность была определена полусферой над городом. Однако если V достаточно велико, то мы обнаружим расходимость, выраженную в схеме (4), потому что условия на границе V практически не имеют отношения к явлениям внутри v. Это означает, что существует такой объем V, что для него и любого большего объема получают схему (4). Возвращаясь к предсказанию того, какая будет максимальная температура в один из июльских дней в некотором городе через 100 лет, мы можем выбрать в качестве граничных условий, как это делалось выше, метеорологические условия на полусферической поверхности радиусом 5000 миль. Эти сведения мало нам помогут, и всё, что мы можем сделать, — это предсказать температуру с помощью полной вероятности p. Таким образом, снова можно говорить о сходимости к p для столбцов. Причины уменьшения ценности предсказания указаны выше, в п. 2 и 3.

Схема (4) представляет структуру (lattice) с *неравномерной сходимостью*. Для данной степени точности ε и момента времени t_k мы можем найти такое описание $D_1^{(l)}$ для момента времени t_1, что оно предсказывает состояние v в момент времени t_k с вероятностью $p_k^{(l)} \geqq 1 - \varepsilon$. Но тогда можно найти такой момент времени $t_m > t_k$, благодаря которому то же самое описание $D_1^{(l)}$ позволяет сделать предсказание состояния v в момент времени t_m лишь с меньшей вероятностью $p_m^{(l)}$. Для малых значений ε не существует описания $D_1^{(l)}$, которое позволяло бы предсказывать состояние v для любого момента времени с вероятностью, равной или большей, чем $1 - \varepsilon$.

Используем эти результаты для рассмотрения вопроса о существовании предельного описания. Здесь мы сталкиваемся со специфической трудностью. Если последовательность описаний задана некоторым правилом, мы, возможно, сможем узнать, имеет ли она предел, то есть сходится ли она к предельному описанию. Например, бесконечная последовательность небольших и все уменьшающихся кругов, сходящихся к точке на плоскости, может рассматриваться как последовательность приближенных описаний этой точки; эта последовательность имеет предельное описание, а именно: координаты точки. Однако последовательность описаний физических состояний не задается некоторым определяющим правилом, и мы встречаемся с проблемой: как индуктивным путем определить, имеет ли последовательность предел. Поэтому мы должны осведомиться, какой вид индуктивного доказательства может считаться подтверждающим гипотезу о том, что последовательность имеет предел.

Итак, кажется правдоподобным утверждение о том, что если бы схема (4) обнаруживала равномерную сходимость, то это можно было бы рассматривать как доказательство, подтверждающее существование предела. Каждый столбец в (4) сходился бы тогда к некоторой такой величине $p^{(l)}$, так что последняя строка, которая содержит все $p^{(l)}$, сходилась бы к 1. В таком случае мы могли бы прийти к выводу, что существует предельное описание D, для которого вероятность предсказания повсюду равна 1. Такой последний столбец фактически был бы однозначно определен всеми остальными столбцами как их предельный случай. Для пространственно

бесконечной вселенной, когда мы должны упоминать о граничных условиях, определение D несколько усложняется. В этом случае D представляет бесконечную систему σ отдельных описаний, каждое из которых относится к определенному моменту времени t_k и обеспечивает значение $p_k = 1$; так как система описаний σ хорошо определена, как объяснялось выше, то предположение о существовании этой системы представляет плодотворную гипотезу.

И наоборот, если схема (4) обнаруживает неравномерную сходимость, то невозможно получить однозначным образом индуктивный вывод относительно последнего столбца. Прослеживая выводы вдоль каждой строки, мы проставим значения вероятностей, равные 1, на всех местах предельного столбца, как указано на схеме (4). Но следуя вдоль столбцов как целых, мы сделаем вывод, что вероятности в предельном столбце должны сходиться к p и, таким образом, не могут быть везде равны 1. Другими словами, правый нижний угол схемы нельзя заполнить непротиворечивым образом: если мы следуем вдоль последнего столбца, то мы должны записать на этом месте значение 1; если же мы следуем слева вдоль последней строки, мы запишем значение p. Такая двусмысленность хорошо известна в теории пределов; двойной переход к пределу часто ведет к различным результатам, так как результат зависит от той последовательности, в которой совершались эти переходы. Это означает, что для неравномерной сходимости схема не определяет предельный столбец последовательным образом, и, следовательно, она не даст нам индуктивного доказательства существования такого столбца.

Мы заключаем, что в случае неравномерной сходимости схема (4) не может служить доказательством, подтверждающим существование предельного описания. Это не означает, что данная схема является доказательством против предположения о пределе: она просто не подтверждает этого предположения. Отсюда вытекает, что предположение о предельном описании оказывается пустым дополнением, которое не обнаруживается в наблюдаемых отношениях.

Изучим эту проблему на примере, в котором неравномерная сходимость сочетается с существованием пре-

дельного описания и который, следовательно, показывает, что это сочетание логически непротиворечиво. Такой пример дает эргодическая гипотеза классической кинетической теории. С первого взгляда можно было бы заключить, что теорема Лиувилля, согласно которой любая область в фазовом пространстве отображается посредством фазовой траектории на область равных размеров (см. рис. 13), приводит к тому результату, что неточность описания должна навсегда остаться в пределах небольшого интервала ϵ. Но этот вывод был бы неверен. Теорема Лиувилля не требует, чтобы области сохраняли

Рис. 17. Если в фазовом пространстве сферическая область A отображается посредством траекторий фазовых точек на область A', то последняя может принять форму, напоминающую морскую звезду, лучи которой простираются на обширные участки пространства.

свои формы. Если A образует сферическую (или, более точно, гиперсферическую) область, тогда область A', на которую она отобразилась траекторией, может иметь форму, подобную морской звезде (рис. 17), хотя и будет занимать площадь такой же величины, как и A[1]. Тонкие ветви этой крайне разветвленной области A' могут в конце концов заполнить все пространство. Поэтому, если известно, что начальное положение фазовой точки находится внутри малого интервала ϵ, мы не можем заключить, что знание этого определяет положение точки в любой момент времени с точностью δ,

[1] В эргодическом масштабе при рассмотрении энергетической поверхности; см. объяснение в § 10, прим. на стр. 111.

имеющей небольшое значение и являющейся функцией ε. Фактически с увеличением времени точка может оказаться на любом, большем заданного δ, расстоянии от предсказанного места P', если мы определяем P' как точку, в которую отображается точка P, находящаяся в центре A. Следовательно, эргодический детерминизм допускает неравномерную сходимость: хотя и существует предельное описание, что позволяет делать строгие предсказания, любое приближение к нему ограничивает возможности предсказания с увеличением времени, и в итоге приходим к схеме, которая обнаруживает неравномерную сходимость.

В этом анализе эргодической гипотезы мы исходили из предположения о детерминизме и затем рассматривали конвергентные свойства. Если мы, используя этот же пример, поступим обратным образом и вначале рассмотрим конвергентные свойства, то мы не получим никакого индуктивного доказательства детерминизма. Если звездообразная часть области A' распространяется на большую часть фазового пространства, то путем измерений мы не сможем различить точки, лежащие внутри одной из тонких ветвей A', от точек, лежащих вне ее, потому что точность измерений ограничена. И если бы только интервал во времени t между областями A и A' был достаточно велик, то звездообразная область A' стала бы обладать всеми свойствами округленной области, огибающей A', по крайней мере насколько это касается эмпирических наблюдений. Эргодический детерминизм выступает как совершенная имитация индетерминизма, настолько совершенная, что ее невозможно отличить от оригинала.

Именно это мы подразумевали, когда выше говорили о том, что схема неоднородной сходимости приводит к заключению об отсутствии индуктивного доказательства существования предельного описания. А двойная сходимость схемы (4) представляет результат всех данных наблюдения в классической физике. Следовательно, если гипотезе о существовании предельного описания дать оценку с помощью наблюдаемых конвергентных свойств, то она выступает как пустое дополнение, без которого вообще можно обойтись. Поиски другого эмпирического доказательства детерминизма, помимо доказательства, полученного из конвергентных свойств приближенных

описаний, представляются довольно затруднительными. Если утверждать, что классическая физика подтверждает детерминизм, то это утверждение вытекает скорее из преклонения перед успехами математических методов в физике, чем из трезвого рассмотрения данных наблюдения. Относительно утверждений, что классическая физика предполагает детерминизм, мы должны заявить, что нет ни одного раздела физики, где детерминизм используется как необходимое предположение. Предполагаются только конвергентные свойства, но они вполне допустимы и в том случае, когда нет никакого предельного описания. Детерминизм логически совместим с классической физикой, но детерминизм никоим образом не доказывается ею, и к нему нельзя прийти путем индуктивного вывода. Даже для классической физики детерминизм представляет излишнее дополнение к системе гипотез, которые выражают основные законы физического мира [1].

Таков ответ на вопрос о детерминизме в кинетической теории, который вытекает из анализа значений. Эргодическая траектория может рассматриваться как хорошее приближение к наблюдаемым взаимоотношениям, однако не следует ее интерпретировать как нечто большее. Тот факт, что из причинных законов механики можно вывести метрику вероятностей для эволюции замкнутых систем, представляет результат огромного значения, который можно включить в концепцию индетерминизма, поскольку выводы из эргодической гипотезы здесь носят приближенный характер. Даже если физические законы не рассматриваются как абсолютно точные, из них можно вывести метрику вероятностей; то есть законы с малыми вероятностями можно вывести из других законов, выполняющихся с очень большой вероятностью. Но, однако, вывод о том, что эволюция замкнутых систем подчинена строгим правилам, отвер-

---

[1] Критика детерминизма в классической физике, представленная в этом параграфе, изложена в некоторых моих ранних статьях, частично опубликованных до открытия квантовомеханического индетерминизма. См. Hans Reichenbach, Die Kausalstruktur der Welt und der Unterschied von Vergangenheit und Zukunft, в «Sitzungsber. d. Bayer. Akad. d. Wiss.», math.-naturwiss. Abt., 1925, S. 135—138; «Die Kausalbehauptung und die Möglichkeit ihrer empirischen Nachprüfung», написанную в 1923 году и опубликованную в «Erkenntnis», Bd. 3 (1932), S. 32—64; и «Stetige Wahrscheinlichkeitsfolgen», в «Zeitschr. f. Physik», Bd. 53 (1929), S. 274—307.

гается. Для принятия этой детерминистской интерпретации термодинамической статистики нет оснований. Даже эволюцию вселенной мы можем рассматривать как действительную вероятностную цепь, где каждый последующий шаг представляет дело случая, в то время как последовательность в целом подчинена статистическим законам, которые так же надежны, как и так называемые строгие причинные законы физики.

§ 12. Вероятностная структура (lattice)

Математические и логические методы анализа вероятностных отношений развиты в теории вероятностей. Можно привести краткое изложение этих методов и используемой символики, прежде чем мы перейдем к применению этих методов к изучению процессов смешивания [1].

Вероятность представляет собою отношение между двумя классами, рассматриваемыми под углом зрения последовательности. Например, вероятность того, что человек 21 года в течение года умрет от туберкулеза, относится к двум классам: к *обусловливающему* (reference) *классу A*, то есть к классу людей в возрасте 21 года, и к *атрибутивному классу B*, то есть к классу людей, умирающих от туберкулеза. Такая вероятность записывается в форме

$$P(A, B) = p. \tag{1}$$

Значение этой вероятности определяется путем подсчета относительной частоты B в пределах A; при этом рассматривается определенная последовательность смертей в некоторой стране в течение определенного периода времени.

Однако относительную частоту наблюденных событий мы не рассматриваем непосредственно в качестве интерпретации вероятности p. Мы допускаем, что эта относительная частота при дальнейших наблюдениях сходится к пределу; и p измеряется этим пределом частоты. Пусть $N^n(A)$ выражает число элементов, относящихся к классу A, и $N^n(A, B)$ — число элементов, относящихся и к классу A и к классу B, где подсчет произ-

[1] Я использую символику, принятую в ThP. Анализ значения (1) см. в ThP, p. 47 и 50.

водится до n-го элемента последовательности. Тогда предел относительной частоты дается формулой

$$\lim_{n \to \infty} \frac{N^n(A.B)}{N^n(A)}. \qquad (2)$$

Эта формула выражает частотную интерпретацию (1).

Существование такого предела и возможность его регистрации при продолжении таких наблюдений в будущем и представляют наше предположение, которое делается при утверждениях о существовании вероятности (1). Конечно, достаточно и того, что последовательность имеет *практический предел,* так как все практически встречающиеся вероятностные последовательности ограничены и, таким образом, не могут удовлетворять требованию сходимости в математическом смысле термина «предел». Под практическим пределом мы понимаем то конечное значение относительной частоты, к которому она сходится довольно быстро и с достаточной степенью точности, так что ее отклонения от этого значения остаются в границах этого интервала точности в дальнейшем. Этими краткими замечаниями здесь можно ограничиться; если читателя интересует более детальное рассмотрение предположения о пределе, я отсылаю его к другой публикации [1].

Часто мы имеем дело с вероятностями, в которых общий обусловливающий класс A идентичен с классом событий последовательности; тогда в выражении вероятности удобно опустить A и вместо этого использовать абсолютные обозначения [2]. Например, вероятность погожих дней обычно подсчитывается в последовательности дней, так что мы можем опустить символ обусловливающего класса и записать эту вероятность в форме

$$P(B), \qquad (3)$$

где B означает «погожий день». Только когда используются особые обусловливающие классы, это будет выражено в символике. Например, вероятность того, что холодный день будет погожим днем, запишется в форме (1), где A представит класс холодных дней.

---

[1] Ibid., § 16, 64—66, 91.
[2] Ibid., p. 106.

Если мы хотим охарактеризовать внутреннюю структуру вероятностной последовательности, мы используем наряду с другими методами влияние предшествующих событий на вероятность наступления последующих событий, то есть мы учитываем наличие последействия. Случайная последовательность не имеет последействия; например, вероятность выпадения шестерки при бросании кости не зависит от того, выпадала или нет шестерка в предыдущем броске. Есть и другие свойства, которым удовлетворяют случайные последовательности. Однако общая теория вероятности не ограничивается рассмотрением случайных последовательностей. Она включает главу о *теории порядка*, которая имеет дело с разного вида упорядоченными последовательностями и определяет специальные виды. Случайная последовательность представляет просто один из этих видов, тогда как другие виды последовательностей могут быть так же важны для применения в физике. Выше указывалось, что эргодическая последовательность обладает последействием. Как уже упоминалось, другой пример дают наблюдения над погодой. Эта последовательность имеет последействие, потому что за хорошей погодой обычно следует такая же, а за плохой погодой — ненастная. Но последействие убывает: для второго дня, после дня с хорошей погодой, вероятность хорошей погоды становится меньше, и если мы хотим вычислить вероятность хорошей погоды через 20 дней после погожего дня, мы можем не учитывать последействия и использовать для вычислений наступления хорошей погоды полную вероятность.

Вероятность с учетом последействия записывается посредством *фазового индекса*. Например, если снова B представляет класс погожих дней, то выражение

$$P(B,\ B^1) \tag{4}$$

обозначает вероятность того, что в последовательности дней за днем, когда имеет место хорошая погода, снова следует день, когда имеет место хорошая погода. Общий обусловливающий класс дней здесь опущен как подразумевающийся. Эта вероятность выше, чем (3), то есть выше, чем полная вероятность хорошей погоды, вследствие вероятностного последействия. Частотная интерпре-

тация (4) записывается выражением

$$\lim_{n\to\infty} \frac{N^n(B.B^1)}{N^n(B)}. \qquad (5)$$

Последнее означает, что мы подсчитываем количество пар погожих дней среди общего числа таких дней.

Последействие играет важную роль в диффузии газов. Предположим, что мы имеем контейнер, разделенный перегородкой на две части (рис. 18); левую часть,

Рис. 18. Два газа, разделенные перегородкой. Когда перегородка убирается, начинается процесс перемешивания.

обозначенную B, мы заполним азотом, а правую, обозначенную $\bar{B}$, — кислородом. Когда мы удаляем перегородку, начинается процесс смешивания. Каждая молекула хаотически движется по всему контейнеру, и через некоторое время молекулы азота и кислорода будут находиться в равном числе в объемах B и $\bar{B}$. Но траектория молекулы обладает последействием. Если она находится в B, то вполне вероятно, что в течение небольшого последующего отрезка времени она будет оставаться в B; то же самое справедливо, если молекула находится в $\bar{B}$. Когда мы запишем эти последовательности для всех молекул, участвующих в процессе диффузии, мы получим вероятностную структуру:

$$\begin{array}{ccccc} x_{11} & x_{12} & \dots & x_{1l} & \dots \to p \\ x_{21} & x_{22} & \dots & x_{2l} & \dots \to p \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{k1} & x_{k2} & \dots & x_{kl} & \dots \to p \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \downarrow & \downarrow & & \downarrow & \\ p_1 & p_2 & \dots & p_l & \dots \to p \end{array} \qquad (6)$$

Каждая горизонтальная строка представляет историю некоторой молекулы, разделенную на отдельные события

небольшими интервалами времени Δt. Этим мы избегаем того, чтобы рассматривать непрерывные вероятностные последовательности[1].

В целях упрощения рассмотрим сначала структуру, относящуюся только к молекулам азота. Так как в начале процесса смешивания все молекулы азота находятся в B, каждый элемент первого столбца выражает событие B[2]. Полная, или абсолютная вероятность B в строке[3] равна $p = 1/2$; она записывается в форме

$$P(B^{kl})^i = p. \tag{7}$$

Это значение справедливо для любой строки k; мы отмечаем этот факт тем, что рассматриваем k как независимую переменную величину. Индексы указывают положение события в структуре. Повторение индекса за скобками указывает на изменяющийся индекс или на связанную переменную величину; другими словами, это повторение показывает индекс, по которому ведется суммирование в частотной интерпретации, и, таким образом, направление подсчета. Поэтому выражение (7) представляет *горизонтальную вероятность*.

Мы можем также образовать *вертикальную вероятность* и написать для абсолютной вероятности B в столбце i значение

$$P(B^{kl})^k = p_l. \tag{8}$$

Для первого столбца мы получим $p_1 = 1$, потому что он состоит только из B-членов. Второй столбец будет содержать несколько $\overline{B}$-членов; и постепенно p_l будет, таким

---

[1] Однако такие последовательности также можно выразить в структурном представлении; см. Ibid., p. 247.

[2] «B» и $\overline{B}$ здесь используются для обозначения того события, что в данный момент времени данная молекула находится в объеме B или в объеме $\overline{B}$. — *М. Р.*

[3] Фраза «вероятность B (или $\overline{B}$) в строке (или столбце)» представляет собой безобидное сокращение. Она может пониматься либо как означающая «вероятность B (или $\overline{B}$) в строке (или столбце)», либо как означающая «вероятность события B (или $\overline{B}$) в последовательности событий, представленной строкой (или столбцом)». Эти же самые результаты получаются при выборе любой интерпретации. — *М. Р.*

образом, становиться все меньше и стремиться к p. То есть мы имеем соотношения

$$P(B^{kl})^k > P(B^{kl})^l, \tag{9}$$

$$\lim_{l\to\infty} P(B^{kl})^k = P(B^{kl})^l. \tag{10}$$

Предельное соотношение (10) между вертикальной и горизонтальной вероятностями есть математическое выражение процессов смешивания.

Последействие, существующее для строк, можно записать в следующем виде, где разность фаз обозначена добавлением —1 в индексе соответствующего члена:

$$\begin{aligned} P(B^{k,\,l-1},\ B^{kl})^l &> P(B^{kl})^l. \\ P(\bar{B}^{k,\,l-1},\ B^{kl})^l &< P(B^{kl})^l. \end{aligned} \tag{11}$$

Эти соотношения выражают положительное последействие, то есть тенденцию к сохранению. Отрицательное последействие, то есть тенденция к изменению, существует в последовательностях других типов, но не встречается в строках рассматриваемых структур. Существуют упрощающие правила оперирования с индексами: в индексе можно прибавить или вычесть постоянное число при условии, что это будет сделано одинаково во всех членах[1]. Например, левую часть первого соотношения (11) можно записать в виде $P(B^{kl},\ B^{k,\,l+1})^l$. Последействие, выраженное в (11), уменьшается с увеличением расстояния; для любого целого числа $r > 0$ мы имеем

$$P(B^{k,\,l-r-1},\ B^{kl})^l < P(B^{k,\,l-r},\ B^{kl})^l. \tag{12}$$

При переходе к бесконечным расстояниям последействие исчезает полностью, а фазовая вероятность сходится к абсолютной вероятности:

$$\begin{aligned} \lim_{r\to\infty} P(B^{k,\,l-r},\ B^{kl})^l &= P(B^{kl})^l = p, \\ \lim_{r\to\infty} P(\bar{B}^{k,\,l-r},\ B^{kl})^l &= P(B^{kl})^l = p. \end{aligned} \tag{13}$$

Практически равенство будет достигнуто уже при больших r.

---

[1] См. ThP, p. 137.

В противоположность этому в столбцах нет последействия, потому что строки независимы друг от друга, то есть мы имеем:

$$P(B^{k-1,\,i},\ B^{ki})^i = P(B^{ki})^k. \tag{14}$$

Соотношение (14) относится к вертикальным вероятностям; соотношение (10) связывает вертикальные и горизонтальные вероятности. Мы можем спросить, можно ли вывести эти соотношения из соотношений (11)—(13), которые относятся только к горизонтальным вероятностям. Такой вывод не может быть дан целиком на основе общих законов теории вероятностей, потому что горизонтальные и вертикальные вероятности в общем являются взаимно независимыми математическими величинами; но он может быть сделан, если структура удовлетворяет двум ограничивающим условиям, которые мы теперь и должны рассмотреть.

Первое ограничивающее условие выражает *независимость* строк и относится к горизонтальным вероятностям. Оно устанавливает, что вероятность B в некоторой строке и также любая фазовая вероятность остаются неизменными в подпоследовательности, выбранной путем ссылки на события B в некоторой другой строке. Например, мы утверждаем, что

$$P(B^{k,\,i-1}\cdot B^{k-1,\,i},\ B^{ki})^i = P(B^{k,\,i-1},\ B^{ki})^i. \tag{15}$$

Это означает, что в случае последействия на вероятность B, отнесенную к непосредственно предшествующему члену (правая часть соотношения), не влияет наличие члена B на этом же месте в предшествующей строке (члена $B^{k-1,i}$ в левой части). Если мы хотим сформулировать это условие для любой группы предшествующих элементов и любых членов, встречающихся в других рядах, удобно использовать дизъюнкцию $B_1 \vee \ldots \vee B_m$; тогда нам не нужно записывать отдельных условий для B и $\overline{B}$. Таким образом, мы получим следующее выражение для условия независимости:

$$P\left(B_{m_r}^{k_1\,i-a_r} \ldots B_{m_1}^{k,\,i-a_1}\cdot B_{n_1}^{k+b_1,\,i+c_1} \ldots B_{n_s}^{k+b_s,\,i+c_s},\ B_{m_0}^{ki}\right)^i =$$
$$= P\left(B_{m_r}^{k,\,i-a_r} \ldots B_{m_1}^{k,\,i-a_1},\ B_{m_0}^{ki}\right)^i, \tag{16}$$
$$a_1 \ldots a_r > 0,\ b_1 \ldots b_s \neq 0.$$

Применение индексов у индексов позволяет нам указать на связь между нижними и верхними индексами. Нижний индекс, имеющий еще индекс внизу, представляет переменную величину, подобную любому простому нижнему индексу. Например, нижний индекс m_r может равняться 1, тогда член B_{m_r} принимает значение B_1. Условие $b_1 \ldots b_s \neq 0$ показывает, что соответствующие B-члены относятся к строкам, отличным от k-строки, к которой относится вероятность в правой части соотношения. Значения $c_1 \ldots c_s$ могут равняться 0. Условие $a_1 \ldots a_r > 0$ означает, что члены в обусловливающем классе предшествуют члену в атрибутивном классе. Если $r = 0$, все предшествующие члены исчезают и соотношение (16) тогда относится к абсолютным вероятностям. Точно так же, если $s = 0$, исчезают все члены в левой части соотношения, принадлежащие различным строкам, и соотношение тогда становится тождеством.

Второе ограничивающее условие выражает свойство, называемое *инвариантностью структуры*[1]. Оно связывает горизонтальные и вертикальные вероятности. Например, мы имеем

$$P\left(B^{k,\,i-1},\ B^{ki}\right)^{i} = P\left(B^{k,\,i-1},\ B^{ki}\right)^{k}. \qquad (17)$$

Это означает, что горизонтальная вероятность обнаружить B после B в k-й строке (левая часть соотношения) выступает еще раз как вертикальная вероятность, относящаяся к двум последовательным столбцам: она измеряет число B-членов, выбранных в i-ом столбце согласно условию, что предшествующий столбец также имеет B на том же самом месте (правая часть уравнения). Это условие нужно распространить на любую фазовую вероятность; однако оно не применимо к абсолютным вероятностям, так как вероятности для B в строках и столбцах различны. Более того, это условие должно содержать указание на члены в других строках, как в (16). Общая формулировка свойства инвариантности структуры тогда принимает следующую форму, в которой левая часть идентична с левой частью из (16), а правая часть получается просто путем замены за скобками верхнего повто-

---

[1] См. ThP, p. 172.

ряющегося индекса:

$$P\left(B_{m_r}^{k,\,i-a_r}\ldots B_{m_1}^{k,\,i-a_1}\cdot B_{n_1}^{k+b_1,\,i+c_1}\ldots B_{n_s}^{k+b_s,\,i+c_s},\ B_{m_0}^{ki}\right)^i=$$
$$=P\left(B_{m_r}^{k,\,i-a_r}\ldots B_{m_1}^{k,\,i-a_1}\cdot B_{n_1}^{k+b_1,\,i+c_1}\ldots B_{n_s}^{k+b_s,\,i+c_s},\ B_{m_0}^{ki}\right)^k, \tag{18}$$
$$a_1\ldots a_r>0,\ r>0,\ b_1\ldots b_s\neq 0.$$

В качестве иллюстрации значения этого выражения можно указать, что особая форма (17) получается из (18) при $r=1$, $a_1=1$, $s=0$, $m_1=1$, $m_0=1$, если мы заменим B_1 на B. Условие $r>0$ исключает абсолютные вероятности из (18). Следовательно, соотношение, выражающее инвариантные свойства структуры и сформулированное в (18), можно перефразировать следующим образом: каждая фазовая вероятность, вычисленная по горизонтали, равна соответствующей вертикальной фазовой вероятности; это справедливо, если учитываются и другие строки.

Соотношение (14) теперь можно вывести следующим образом. В (18) мы положим $s=1$, $b_1=-1$, $c_1=0$; $r=1$, $a_1=j-1$. В правой части мы положим $i=j$. Соотношение (18) тогда принимает вид

$$P\left(B_{m_1}^{k,\,i-j+1}\cdot B_{n_1}^{k-1,\,i},\ B_{m_0}^{ki}\right)^i=P\left(B_{m_1}^{k1}\cdot B_{n_1}^{k-1,\,j},\ B_{m_0}^{k,\,j}\right)^k. \tag{19}$$

Вследствие условия независимости (16) член $B_{n_1}^{k-1,\,i}$ в левой части можно опустить. Левую часть соотношения (19) можно рассматривать как получающуюся из левой части (18) при $s=0$, $r=1$; и, согласно (18), она равна выражению, получающемуся из правой части (19) после сокращения члена $B_{n_1}^{k-1,\,j}$. Таким образом, мы находим, что сокращение последнего упомянутого члена не изменяет правой части соотношения, и мы получаем:

$$P\left(B_{m_1}^{k1}\cdot B_{n_1}^{k-1,\,j},\ B_{m_0}^{kj}\right)^k=P\left(B_{m_1}^{k1},\ B_{m_0}^{kj}\right)^k. \tag{20}$$

Так как первый столбец состоит только из B-членов, в выражениях обеих частей соотношения (20) первые члены допускают запись в виде B^{k1} для любой строки и, таким образом, могут быть опущены. Соотношение (20) тогда будет тождественно (14), если мы положим $j=i$.

Соотношение (10) можно вывести следующим образом. Мы положим в (18) $s=0$, $r=1$, $a_1=j-1$;

и, как в (19), примем $i = j$ в правой части соотношения. Опуская нижние индексы, то есть записывая результат только для B-членов, мы находим

$$P(B^{k,\, i-j+1},\ B^{ki})^i = P(B^{k1},\ B^{k,\, j})^k = p_j. \qquad (21)$$

Так как каждая строка начинается с B-члена, последнее выражение представляет абсолютную вероятность B в столбце j. То есть величина p_j в (21) соответствует определению (8). Таким образом, соотношение (21) говорит о том, что абсолютная вероятность B в j-ом столбце равна горизонтальной вероятности с учетом последействия, что за B последует B через $j-1$ элементов. Согласно (13), первое выражение (21) стремится к p с ростом j. То же самое должно быть справедливо и для второго выражения; и, таким образом, (10) доказано.

Эти соображения показывают, что структура, в которой строки обладают последействием и первый столбец упорядочен, то есть не смешан, может характеризовать процесс смешивания, если она удовлетворяет условиям независимости и инвариантности. Структура вида, которая характеризуется соотношениями (11)—(13), (16) и (18), будет называться *структурой процесса смешивания.*

Данные выводы показывают, в частности, важность условия инвариантности структуры. Это условие позволяет нам преобразовывать горизонтальные вероятности в вертикальные вероятности. Так как строка относится к временной последовательности состояний одной молекулы, ее часто называют *временным ансамблем;* в противоположность этому столбец называется *пространственным ансамблем*, потому что он относится к одновременным состояниям многих молекул. Следовательно, соотношение (18) формулирует условие, которое позволяет нам *по временному ансамблю сделать заключение о пространственном ансамбле.* Этот вывод говорит, что поперечное сечение мировых линий многих составляющих отражает распределение в пределах мировой линии одной составляющей. В структуре процесса смешивания этот вывод непосредственно применяется только к фазовым вероятностям; по отношению к абсолютным вероятностям он справедлив только для достаточно удаленных от первого столбцов, то есть после того, как процесс смешивания продолжался уже в течение некоторого

времени. В этих столбцах частота B практически такая же, как и в строках. Однако наличие последействия в строках не передается столбцам, так как в столбцах нет последействия, как уже отмечалось ранее (14). Структуре процесса смешивания могут быть приданы другие формы путем выбора различных распределений для первого столбца, тогда как общие условия (11)—(13) и (18) по-прежнему будут выполняться. Например, мы можем включить в структурное представление молекулы и азота и кислорода, считая, что чередующиеся строки представляют мировые линии молекул азота и кислорода. Тогда мы будем иметь равное число элементов B и $\overline{B}$ в каждом столбце и, таким образом, всюду найдем

$$P(B^{kl})^i = P(B^{kl})^k. \tag{22}$$

Это означает, что вывод о пространственном ансамбле на основе временного ансамбля в применении к абсолютным вероятностям справедлив для каждого столбца, включая и начальные.

И наоборот, соотношение (14) несправедливо для этой структуры, и мы наблюдаем отрицательное последействие в столбцах. Первый столбец состоит из строго чередующихся членов B и $\overline{B}$. Как следствие этого, в последующих столбцах еще будет заметна тенденция к чередованию, и, таким образом, они будут содержать большое число пар $B.\overline{B}$ и меньшее число пар $B.B$. Этот факт записывается в виде следующих соотношений:

$$P(\overline{B}^{k-1,\,l},\ B^{kl})^k > P(B^{kl})^k,\ P(B^{k-1,\,l},\ \overline{B}^{kl})^k > P(\overline{B}^{kl})^k, \tag{23}$$

$$\lim_{l\to\infty} P(\overline{B}^{k-1,\,l}, B^{kl})^k = P(B^{kl})^k,\ \lim_{l\to\infty} P(B^{k-1,\,l}, \overline{B}^{kl})^k = P(\overline{B}^{kl})^k. \tag{24}$$

Отметим, что эти соотношения не противоречат соотношению (20); последняя формула выражает независимость строк при подсчете по вертикали и применима также и к данной структуре, так как она выводится из (16) и (18). Знак равенства в (20) получается благодаря члену $B^{kl}_{m_1}$ в обусловливающем классе; по отношению к первым элементам B или $\overline{B}$ строки элемент B не зависит от того, какой элемент в столбце ему предшествует. Последействие встречается только в тех случаях,

когда первый элемент строки не входит в обусловливающий класс. Это показывает, что последействие в столбцах порождается просто наличием строгого чередования в первом столбце, но не физической зависимостью между элементами.

Эту структуру можно рассматривать как наложение двух структур предыдущего типа или, скорее, как введение строк одной такой структуры между строками другой. Таким образом, в соответствии с (21) мы получаем

$$P(\overline{B}^{k,\, l-j+1},\ B^{kl})^{l} = P(\overline{B}^{k1},\ B^{kj})^{k} = q_j. \qquad (25)$$

Используя инвариантные свойства структуры и предельные соотношения (13) для горизонтальных вероятностей, мы получаем

$$\lim_{j\to\infty} p_j = p = \lim_{j\to\infty} q_j. \qquad (26)$$

Каждый столбец j в полной структуре представляет последовательность Пуассона[1], контролируемую чередующимися вероятностями p_j и q_j. Существование последействия, выраженного в (23) и (24), доказано в теории таких последовательностей. С ростом j — для достаточно удаленных столбцов — вероятности p_j и q_j становятся приближенно равными вследствие выраженной в (26) сходимости, и представленная столбцом последовательность все сильнее и сильнее становится похожей на последовательность, контролируемую одной вероятностью p. Однако предельное соотношение (10) здесь заменяется более строгим соотношением (22), так как мы принимаем, что $p = 1/2$. Это показано в теории последовательностей Пуассона, но также ясно и интуитивно.

Интересно, что в такой структуре процесс смешивания обнаруживается не в изменении значений вероятностей $P(B^{kl})^{k}$ от столбца к столбцу, а в наличии последействия во всех начальных столбцах. Это последействие исчезает только для столбцов, соответствующих более поздним временам. Это и выражено в соотношениях (23) и (24). Таким образом, здесь мы имеем пространственный ансамбль, в котором соотношения, отражающие

[1] ThP, p. 296.

внутренние зависимости, показывают, какую стадию проходит процесс смешивания.

Рассматривая другие виды распределения B в первом столбце структуры, можно задать вопрос, какие соотношения можно вывести для общего случая. Соотношения (21) и (25) всегда справедливы, и, поскольку мы всегда можем рассматривать данную структуру как наложение двух структур, заданных соответственно первыми элементами B или $\overline{B}$, соотношение (26) снова приводит к обоснованию соотношения (10). Так, мы находим, что любая структура процесса смешивания удовлетворяет соотношению (10). Вероятности для B в столбцах всегда сходятся к полной вероятности B в строках. Более строгое условие (22), согласно которому эта полная вероятность непосредственно контролирует каждый столбец, имеет силу только для частного случая, когда распределение B в первом столбце равно распределению B в строках. Доказательство (22) для этого случая легко получается, и нет надобности приводить его здесь.

Сходимость вероятностей в столбцах, выраженная в соотношении (10), представляет, следовательно, общую характеристику всех процессов смешивания, если процесс смешивания мы определим соотношениями (11)—(13), (16) и (18). Для столбцов, достаточно удаленных от начального, выводы о пространственном ансамбле на основе временнòго ансамбля справедливы, таким образом, для любой структуры. Единственное различие между строками и удаленными столбцами заключается в том, что эти строки обладают последействием, в то время как указанные столбцы последействием не обладают. Эти результаты, полученные здесь для дизъюнкции двух членов B и $\overline{B}$, легко распространяются на случай дизъюнкции m членов $B_1 \ldots B_m$.

В истории развития кинетической теории, в частности в работе Гиббса, вывод о пространственном ансамбле на основе временнòго ансамбля играет важную роль, и поэтому можно добавить некоторые замечания о его логической природе. Часто в курсах статистической механики можно прочесть, что этот вывод делается на основе теории вероятностей. Это неверно. Условие инвариантности структуры не выводится из аксиом теории вероятностей, но представляет специфическое условие,

которое справедливо только для определенных видов структур. Вопрос о том, удовлетворяет ли этому условию рассматриваемый физический процесс, относится к области эксперимента[1].

Чтобы изучить эти взаимоотношения, исследуем упрощенный вывод для структуры, в которой строки не обладают последействием. Легко заметить, что в этом случае условие инвариантности структуры, соотношение (18), приводит к заключению, что все столбцы, за исключением, может быть, первого, имеют вероятность q для B, равную вероятности B для строк. Другими словами, в этом случае справедливо соотношение (22), и мы получаем $q = p$.

Посмотрим, можно ли получить этот результат из условия независимости. Мы рассмотрим только n первых строк. Если строки независимы, то возможные конфигурации B и $\overline{B}$ в столбце из n членов можно вычислить методами, применяемыми в теореме Бернулли. Эти методы приводят к заключению, что в большинстве столбцов частота B соответствует его горизонтальной вероятности p. Таким образом, доказывается, что чем длиннее столбцы, тем больше горизонтальная вероятность w_n для вертикальной частоты B, соответствующей p. Если n стремится к бесконечности, то w_n стремится к 1.

Мы обычно склонны заключить, что существует горизонтальная вероятность $w_\infty = 1$, чтобы определить столбцы, для которых q равно p. И все же довольно удивительно, но мы не можем доказать, что существует хотя бы один столбец, для которого действительно выполняется $q = p$, даже если это равенство будет выполняться приближенно, то есть предел w_n не обязательно должен иметь значения горизонтальной вероятности столбцов. Это следует из построения следующей структуры. Слева от диагонали структуры, определенной элементами x_{11}, x_{22}, x_{33}, ..., мы везде поставим B. Вправо от этой линии мы продолжим все строки как случайные последовательности. Эта структура отвечает условиям, необходимым для применения теоремы Бернулли, так как удовлетворяет условию независимости (16); и

[1] Обсуждение этой проблемы см. там же, pp. 174, 280.

горизонтальная вероятность w_n имеет то же самое значение, что и раньше. Но в каждом столбце вертикальная вероятность q равна 1 и, таким образом, отличается от p. Причина этого состоит в том, что не выполняется условие инвариантности структуры (18).

Можно возразить, что построенная структура представляет вырожденный случай, возможный только для бесконечных столбцов, и что во всех практических применениях столбцы имеют конечную длину. Это верно, но неправомерно считать, что в случае конечных столбцов мы можем обойтись без условия инвариантности структуры. Использование бесконечных последовательностей представляет целесообразную идеализацию вероятностных проблем, и перенесение полученных результатов на конечные последовательности требует предосторожности. В действительности наши структуры конечны как в отношении столбцов, так и в отношении строк. Понятие предела для них заменяется понятием практического предела, которое просто означает достаточно быструю сходимость к конечному значению частоты, то есть частоты, вычисленной по последним элементам последовательности. Таким образом, это условие требует, чтобы частота довольно долгое время оставалась в пределах интервала сходимости ε. Для структуры мы также имеем подобные ограничения сходимости. Пусть n — длина столбца, а m — такая длина строк, что среди первых m столбцов процент, определяемый w_n, показывает частоту B, соответствующую p. Затем мы требуем, чтобы m было мало в сравнении с n; в противном случае структура во всех практических случаях имела бы свойства описанной вырожденной структуры. По этой причине условие (18) инвариантности структуры, переписанное в соответствующей форме, необходимо перенести в финитизированную теорию вероятностей.

Физическую иллюстрацию описанной вырожденной структуры может следующим образом дать процесс диффузии, представленный на рис. 18. Мы записываем структуру только для молекул азота, то есть в первом столбце мы имеем только члены B. Предположим, что разделяющая стенка вытаскивается очень медленно через щель в дне контейнера. Тогда если в верхних строках структуры (6) начинается процесс смешивания, то нижние строки остаются в B в течение довольно дол-

того времени. Эта структура не удовлетворяет условиям (18), и ее столбцы не отражают последействия в вероятностях, наблюдаемых вдоль строк, хотя строки взаимно независимы и удовлетворяют (16). Даже если столбцы имеют конечную длину, мы, таким образом, можем провести различие между нормальным процессом смешивания и вырожденным процессом смешивания; это различие указано в соотношении (18).

Попытка исключить условие инвариантности структуры показывает определенную тенденцию придавать большее значение вероятностям во временнòм ансамбле в сравнении с вероятностями в пространственном ансамбле. Эта тенденция проистекает из убеждения, что замена вертикальных вероятностей горизонтальными, относящихся к столбцам ограниченных размеров, дает некоторое преимущество. Однако трудно понять, каким образом можно оправдать такое мнение. Если задан столбец конечной длины, который не удовлетворяет ожидаемым отношениям частот, мы можем сказать: переходите к временнòму ансамблю, тогда вы получите столбцы, которые удовлетворяют ожидаемым отношениям. Но мы вправе также заявить: увеличьте длину столбца, тогда вы найдете, что ожидаемые отношения оправдываются. Если принимается вторая интерпретация фразы «вероятно, что», мы вынуждены рассматривать условие инвариантности структуры как особое условие, которое не является просто переходом к вероятностям временнòго ансамбля. Это легко увидеть из данного выше описания вырожденной структуры, для которой вторая интерпретация не является обоснованной и справедлива первая. И все же в отношении к ограниченным сериям наблюдений вероятностные утверждения всегда имеют значение предсказаний о будущих наблюдениях; и если этот взгляд принят, то нет основания для того, чтобы отдавать предпочтение временнòму ансамблю в сравнении с пространственным. Последовательная частотная интерпретация вероятностей ведет к такой трактовке вероятностных структур, при которой допускаются предельные утверждения и в горизонтальном и в вертикальном направлении; и в рамках такой теории условие инвариантности структуры просто необходимо.

Сравнение с условием независимости вскрывает логическую природу условия инвариантности структуры.

Это условие представляет расширенную форму предположения о независимости. Для того чтобы перейти в суждениях от временнòго ансамбля к пространственному, недостаточно знать, что строки взаимно независимы; они должны удовлетворять более строгому условию инвариантности структуры.

Структура процесса смешивания представляет математическую модель процесса смешивания. Она формулирует математические отношения, физическим эквивалентом которых является процесс смешивания молекул или других объектов, которые подчинены действию механизма размешивания. Она показывает, что процесс смешивания регулируется тремя основными отношениями: последействием, независимостью и инвариантностью структуры. Последействие относится к зависимостям в пределах строки; независимость — к зависимостям между различными строками; а инвариантность структуры — к зависимостям между горизонтальными и вертикальными вероятностями. В дальнейшем мы увидим, что эти зависимости могут также соблюдаться в физических системах, которые не подвержены действию механизма размешивания; и что использование структур при математической формулировке задач позволяет нам перенести полученные результаты на процессы, которые в более общем смысле представляют аспект смешивания.

§ 13. Парадокс обратимости

Теперь нам необходимо исследовать вопрос о том, позволяет ли статистическое решение проблемы энтропии определить *направление* времени. Можно напомнить, что ранее (§ 5) *порядок* времени был определен на основе причинных законов. Принимая этот порядок, то есть предполагая, что в физическом мире события упорядочены в причинные цепи, мы хотим найти способы различения одного направления этих цепей, которое соответствует положительному времени, от противоположного направления. То есть мы хотим показать, что время не только *упорядоченно*, но и *однонаправленно*. Данная проблема будет рассматриваться в этом и в следующих параграфах.

Вначале можно ответить на одно простое возражение против применения статистического определения напра-

вления времени. Выше упоминалось (§ 7), что, согласно статистической интерпретации закона энтропии, изменение направления термодинамических процессов не может уже считаться невозможным, но можно только сказать, что это весьма маловероятно. Возможно, что кто-нибудь стал бы доказывать, что в зависимости от вида наблюдаемых процессов могут вытекать противоречивые заключения о направлении времени. Если бы мы случайно встретились с одним из маловероятных процессов, ему приписали бы направление времени, противоположное направлению времени в других процессах. Но это возражение не обосновано. Направление времени представляет собою свойство причинной сети в целом, которое передается индивидуальным процессам. Положительное время мы определяем как такое направление, в котором протекает большинство термодинамических процессов. Если бы индивидуальные процессы стали протекать в направлении отрицательного времени, мы отметили бы эти исключительные случаи на основе общего определения. Мы также припишем направление обратимым процессам, которые сами по себе не обнаруживают однонаправленности, соотнося их с общим направлением причинной сети.

Кроме того, можно отметить, что статистический вариант определения направления времени не приводит к таким утверждениям, как: «Просто весьма маловероятно, чтобы прошлое возвратилось». Выше (§ 5) вы видели, что возврат к прошлому образовал бы замкнутую причинную цепь, отрицание которой есть проявление упорядоченности и не имеет никакого отношения к направлению времени. Статистическое истолкование направления времени вполне совместимо с физическим законом, что причинные цепи никогда не замыкаются. Критический разбор этого закона относится к другому контексту, к дискуссии о строгих законах физики.

Однако при статистическом определении направления времени возникают иные трудности. Эти трудности касаются вопроса, действительно ли изменение в состояниях систем в направлении возрастания энтропии, или увеличения вероятности, более вероятны, чем изменения в противоположном направлении. Существуют, хотя это довольно удивительно, серьезные возражения против того самого аргумента, на котором основывается

статистическая интерпретация энтропии, а именно против того принципа, что физические системы имеют тенденцию эволюционировать в направлении более высокой вероятности.

Проблему, с которой сталкивается физик, можно сформулировать следующим образом. Предполагается, что элементарные процессы статистической термодинамики, процессы движения и столкновения молекул, подчиняются законам классической механики и, следовательно, обратимы. Макропроцессы же необратимы, как нам известно. Каким образом можно примирить эту необратимость макропроцессов с *обратимостью микропроцессов*? Именно этот парадокс и должен разрешить физик, если он хочет объяснить направление термодинамических процессов и направление времени. Кажется, что статистическая интерпретация энтропии дает решение, потому что, согласно этой интерпретации энтропии, термодинамические процессы имеют тенденцию развиваться от порядка к беспорядку. Несмотря на это, при более внимательном изучении мы обнаруживаем, что это кажущееся решение можно поставить под сомнение и что обратимость элементарных процессов противоречит указанной тенденции развития от порядка к беспорядку. *Этот парадокс статистического определения направления* необходимо теперь исследовать более подробно.

Рассмотрим процесс диффузии, которого мы уже касались при обсуждении рис. 18. Как и раньше, мы заполним одну часть контейнера азотом, а другую — кислородом. Контейнер полностью изолирован от своего окружения. После этого удалим разделяющую перегородку, тогда начнется процесс смешивания газов. При термодинамическом истолковании энтропия в таком случае возрастет, потому что, как можно показать, сумма значений энтропий двух разделенных газов меньше, чем энтропия их общей смеси. При статистической интерпретации этот процесс идет в направлении от упорядоченного макросостояния к неупорядоченному и, таким образом, от состояния с низкой вероятностью к более высокой; этот результат был нами получен в § 12.

Произведем теперь мысленный эксперимент. После того как газы смешались, мы остановим каждую молекулу в отдельности и сообщим ей движение в строго

противоположном направлении, так что она начнет перемещаться обратно с той же самой по величине скоростью, какой она достигла к этому моменту. Такое изменение направления скорости можно получить, например, заставив каждую молекулу отразиться от небольшого зеркала, помещенного перпендикулярно к ее траектории; однако все эти отражения должны произойти в один и тот же момент времени. Что будет наблюдаться в этом случае? Очевидно, вследствие обратимости элементарных процессов должна произойти полная обратимость предшествовавшего процесса; газы должны разделиться, и спустя некоторое время мы должны будем обнаружить, что все молекулы кислорода сосредоточены в одной части контейнера, а молекулы азота — в другой его части.

Здесь мы обычно заявляем: но вы не сможете совершить этот эксперимент, который полностью вымышлен. Вы не сможете искусственно создать конфигурацию молекул с обратными направлениями скоростей; такая конфигурация явилась бы крайне маловероятной, потому что она была бы чрезвычайно упорядоченной. Следовательно, практически процесс разделения никогда не может произойти.

Именно здесь возникают действительные трудности. В теории газов Больцмана — и то же самое справедливо для вариантов, развитых Максвеллом и Гиббсом, — вероятность того, что молекула обладает определенной скоростью, не зависит от знака скорости; это видно из (9.14) для состояния равновесия. Следовательно, при изменении направления на противоположное скорость обладает той же самой вероятностью, как и исходная. Это не означает, что следует ожидать того, что все молекулы самопроизвольно и внезапно изменят направление своих скоростей. Это означает, что состояние, при котором молекулы движутся в строго противоположных направлениях и с теми же самыми значениями скоростей, как и в первом состоянии, может явиться естественным следствием непрерывного изменения скоростей при столкновениях и что такое состояние обладает во времени той же самой вероятностью, как и первое состояние. Но отсюда тогда следует, что процессы разделения должны встречаться так же часто, как и процессы смешивания.

Приведенный аргумент называется *парадоксом обратимости*. Он был выдвинут в первоначальный период развития кинетической теории коллегой Больцмана по Вене Й. Лошмидтом (1876)[1] и с тех пор часто обсуждался. Больцман дал на него свой ответ, который мы и рассмотрим сейчас.

Представим сперва указанный парадокс в несколько ином виде. Изменения энтропии S, или вероятности W, в замкнутой системе, которая начинает свое развитие с упорядоченного состояния, подобно упоминавшемуся

Рис. 19. Кривая энтропии замкнутой системы.

газу в контейнере, могут быть представлены кривой, изображенной на рис. 19. Начальное макросостояние A, которое, как мы предполагаем, соответствует положению *1*, имеет низкое значение энтропии. От этого значения энтропия возрастает к более высокому, соответствующему положению *2*, и затем остается вблизи этого значения в течение продолжительного времени. Однако это значение не будет строго сохраняться; будут наблюдаться хаотические флуктуации, которые приводят к колебаниям кривой. Большие флуктуации значительно менее вероятны, чем небольшие; но их вероятность выше нуля, и, следовательно, они время от времени должны происходить. Это представляет существенную черту вероятностных последовательностей: каждая комбинация

[1] J. Loschmidt, Über den Zustand des Wärmegleichgewichtes eines Systems von Körpern mit Rücksicht auf die Schwerkraft, в «Sitzungsber. d. Kaiserl. Akad. d. Wiss. in Wien», math.-naturwiss. Cl., Bd. 73, 1876, 2. Abt., S. 139.

атрибутов, вероятность которой не равна нулю, должна встречаться с этой не нулевой частотою. Если бросать игральную кость достаточно много раз, то в конце концов должна встретиться серия из тысячи бросаний, в которой подряд выпадет 6, потому что вероятность такой серии хотя и очень мала, но все же больше нуля. Когда мы тасуем колоду карт сразу после того, как мы уложили все красные карты поверх черных, мы преобразуем это упорядоченное состояние в перемешанное; но если тасовать достаточно долго, то мы должны в конце концов совершенно случайно вернуться к первоначальному состоянию, потому что вероятность получить такое расположение больше нуля. С точки зрения подобных вероятностных представлений можно заключить, что иногда газ должен подвергаться большим флуктуациям, настолько большим, что даже будет происходить разделение газов[1]. Конечно, вероятность такого случая чрезвычайно мала, и наблюдение в течение миллиона лет дает нам мало шансов увидеть такой процесс. Процитируем Больцмана: «Нельзя предполагать... что два диффундирующих газа смешаются и разделятся несколько раз в течение немногих дней. Время, в течение которого можно надеяться наблюдать разделение, настолько велико, что исключается любая возможность наблюдать такой процесс»[2]. Однако небольшие флуктуации действительно иногда происходят. Броуновское движение небольших частиц, наблюдаемое под микроскопом, представляет результат таких случайных флуктуаций; оно происходит вследствие того, что число мо-

---

[1] Этот же самый результат был получен Эрнестом Цермело и Анри Пуанкаре, показавшими, что механическая система, изменения которой подчиняются уравнениям Гамильтона (10.1), должна в конце концов вернуться к состоянию, весьма близкому к своему первоначальному, то есть что эргодическая траектория должна быть квазипериодической. См. E. Zermelo, Über einen Satz der Dynamik und die mechanische Wärmetheorie, в «Ann. d. Physik und Chemie», N. F., ed. by G. und E. Wiedemann, Bd. 57, 1896, S. 485—494; Henri Poincaré, Sur le problème des trois corps et les équations de la dynamique, в «Acta Math.», Tome 13, 1890, p. 1—270, esp. p. 67—72. Этот результат называется *парадоксом периодичности*, и его значение подобно значению парадокса обратимости.

[2] Переведено из Людвига Больцмана и Д. Набла: „Kinetische Theorie der Materie", в „Encyklopädie der mathematischen Wissenschaften", Bd. V, Teil I, Art. 8, Leipzig, 1907, S. 521.

лекул, ударяющихся о частицу с одной стороны, строго не равно числу молекул, ударяющихся о нее с другой стороны, и, таким образом, может рассматриваться как указание на наличие флуктуаций давления газа в очень малых объемах.

Принципы вероятности убеждают нас, что большую часть времени рассматриваемая нами кривая будет оставаться при высоких значениях. И это все, что мы можем сказать, однако случайные падения кривой к меньшим значениям нельзя исключить; более того, согласно этим же самым принципам, они должны происходить. Рассматривая диаграмму (рис. 19), мы теперь видим, что каждому участку подъема кривой, исключая начальный, предшествует соответствующий участок падения кривой. Поскольку на большом отрезке кривой влияние на относительную частоту незначительно, мы заключаем, что для кривой бесконечной длины число падений и подъемов будет одинаковым. Это означает, что переходы к низким значениям энтропии происходят столь же часто, сколь и переходы к высоким значениям; в этом и состоит формулировка парадокса обратимости.

Парадокс обратимости можно выдвинуть вне зависимости от того, используем ли мы детерминистическую или индетерминистическую интерпретацию процесса эволюции замкнутых систем. Причина этого состоит в том, что данный парадокс проистекает из самой природы вероятностей и не связан с предположением о наличии причинных законов, лежащих в основе статистики молекул. Вместе с тем следует признать, что первоначально, когда допускалась обратимость скоростей молекул, при формулировке парадокса опирались на строгость законов механики и обратимость механических процессов. Но если согласиться с тем, что мы можем приписать определенную вероятность появлению во времени каждой молекулярной конфигурации замкнутой системы, тогда мы должны получить кривую, подобную изображенной на рис. 19, потому что каждое значение вероятности должно проявляться в определенной частоте событий.

Чтобы более детально исследовать этот результат, рассмотрим процесс перемены направления скоростей молекул для не полностью изолированной системы. Пусть a обозначает начальное микросостояние, в котором газы

разделены, b — последующее микросостояние системы, когда газы перемешаны. Под b^* мы понимаем состояние, возникающее из b, когда все скорости молекул изменяют свое направление на обратное и из которого газы возвращаются в состояние a^*, соответствующее состоянию a, за исключением того, что все скорости молекул имеют обратное направление. Состояния a и a^*, которые относятся к тому же самому макросостоянию A, представляют упорядоченные состояния. Поскольку их порядок легко наблюдаем, мы говорим здесь о *явном порядке*. Состояния b и b^* являются неупорядоченными. Эти состояния возникли в результате процесса смешивания, макроскопически они неразличимы и относятся к одному и тому же макросостоянию B. Но так как b^* представляют собою особый класс состояний, который приводит к упорядоченному состоянию a^*, мы можем сказать, что b^* содержит в себе *неявный порядок*. Если мы теперь предположим, что система извне подвергается небольшим хаотическим воздействиям, мы обнаружим существенное различие между этими двумя видами порядка. Явный порядок является *устойчивым*, то есть небольшие воздействия не нарушают его. Но неявный порядок является *неустойчивым*; небольшие воздействия нарушают его в том смысле, что приведут к полному беспорядку. Например, если при изменении направления скоростей некоторые молекулы не изменяют своего направления, неявный порядок будет разрушен и данная система не перейдет в упорядоченное состояние. Однако небольшие подобные изменения в a не разрушают явного порядка, и система, как и прежде, перейдет в неупорядоченное состояние.

Можно было бы попытаться использовать это различие для ответа на парадокс обратимости. Однако этим путем можно доказать только то, что обратный процесс окажется еще более маловероятным, но не то, что он исключается. Хаотические воздействия могут происходить во всех видах и сочетаниях; и могут оказаться сочетания, которые не нарушают неявный порядок. Вероятность таких сочетаний случайных воздействий выше нуля, и, следовательно, они должны время от времени происходить. Возмущения извне действуют просто как перемешивающий механизм; они приводят к нарушениям последовательности в состояниях газа, но не могут

воспрепятствовать тому, чтобы она изредка развивалась в направлении к порядку.

Эти результаты теперь можно применить к рассмотрению вопроса об эволюции вселенной. Если вселенная пространственно ограничена, мы можем говорить о ее состоянии в некоторый момент времени t. Такое состояние иногда называют временны́м *сечением* четырехмерного пространственно-временно́го мира; оно определяется только после того, как дано определение одновременности (см. рис. 8). Мы можем допустить, что это состояние задано с точностью до микросостояния; но так как каждое микросостояние определяет соответствующее ему макросостояние (см. § 9), мы можем говорить о макровероятности, или об энтропии, *мгновенного состояния* вселенной. Следовательно, для пространственно-ограниченной вселенной кривая на рис. 19 выражает не только развитие изолированной системы газа, но также и эволюцию вселенной. Если же вселенная пространственно не ограничена, вероятности ее состояний невозможно определить, и тогда нельзя говорить об энтропии вселенной. Но мы пока отложим рассмотрение этой возможности и ограничимся исследованием вселенной, для которой можно определить полную энтропию. Однако мы предположим, что время бесконечно, то есть что причинная сеть открыта (см. § 5).

Для такой вселенной кривую на рис. 19 необходимо продолжить влево, потому что время бесконечно также и в отрицательном направлении. Это означает, что влево от точки *1* мы должны будем изобразить участок кривой, идущей вверх, что соответствует возрастанию энтропии в направлении отрицательного времени. Таким образом, вся часть кривой, идущая влево от точки *1*, напоминает по своей форме всю правую часть кривой, хотя, конечно, левая часть не будет представлять строгое зеркальное отображение правой части, поскольку кривая иррегулярна по форме. Однако очевидно, что вдоль кривой подъемы и падения встречаются одинаково часто и поэтому кривая не выражает определенного направления.

Поставим теперь два различных по характеру вопроса. Первый касается *вывода о величине энтропии на основе знания направления времени*, второй — *вывода о направлении времени на основе данных об энтропии*. Начнем с рассмотрения первого вопроса.

Первый вопрос можно сформулировать следующим образом. Пусть задано направление времени и энтропия S_A начального неравновесного макросостояния A; будет ли энтропия S_B макросостояния B иметь большее значение, если B следует позднее во времени, чем A, на интервал времени Δt? Ответ будет таким: если B следует за A через интервал времени Δt, который не слишком мал, то более вероятно, что $S_B > S_A$, чем $S_B < S_A$. Этот результат следующим образом можно получить из кривой, изображенной на рис. 19. Предположим, что A расположено на участке *5—6* в точке пересечения с пунктирной линией постоянной высоты. Тогда несколько более позднее состояние обладает более низким значением энтропии. Но если выбрать больший интервал Δt, состояние B изобразится точкой в верхней части участка *6—7* кривой, и, таким образом, его энтропия будет больше, чем энтропия A. Тот же результат будет справедлив и для еще большего Δt, потому что B тогда окажется расположенным на горизонтальном участке *7—8* кривой. Только если Δt станет еще больше, B сместится на участок *8—9* с меньшим значением энтропии. Но это будет выступать как исключение, потому что горизонтальные расстояния между участками спада кривой нерегулярны. Данная вероятность относится к классу точек A, лежащих на пересечениях пунктирной линии с кривой; и если все эти точки рассматриваются в пределах одного и того же интервала Δt, то в большинстве случаев B будет расположено на более высоком уровне. Этот результат справедлив для любого интервала Δt, при условии, что он не слишком мал, и при условии, что мы сохраняем постоянным интервал Δt между A и B, в то время как A скользит вдоль горизонтальной линии. В этом состоит смысл утверждения, что если A имеет низкое значение энтропии, то $S_B > S_A$ более вероятно, чем $S_B < S_A$.

Хотя это достаточно странно, но тот же самый результат можно получить, если B предшествует во времени A. Если A скользит вдоль пунктирной горизонтальной линии, в то время как B расположено слева от A на расстоянии временного интервала Δt, мы обнаружим тот же самый результат, что и выше: при постоянном Δt гораздо чаще B принимает более высокие значения энтропии, чем более низкие. Мы находим, что если между

наступлением B и наступлением A имеет место интервал, равный Δt, который не слишком мал, то более вероятно, что $S_B > S_A$, чем $S_B < S_A$. И мы видим: вывод об энтропии на основе времени ведет к тем же самым результатам вне зависимости от того, относится ли он к последующим или предыдущим событиям. Этот факт является выражением того, что кривая энтропии симметрична относительно направления времени.

Данные результаты были получены Больцманом и затем детально проанализированы П. и Т. Эренфестами[1]. Больцман рассматривал это соображение как ответ на парадокс обратимости[2].

Так как полученный результат дает нам основание надеяться, что упорядоченное состояние A будет переходить в неупорядоченное состояние B, он действительно является ответом, когда мы рассматриваем вывод об энтропии на основе времени. Например, когда мы удалим перегородку в контейнере, изображенном на рис. 18, можно ожидать с большой вероятностью, что спустя некоторое время газы будут хорошо перемешаны. Для контейнера обычных размеров временной интервал Δt даже в несколько минут приводит с громадной вероятностью к возрастанию энтропии. Вероятность обратного процесса не уменьшает заметным образом эту вероятность.

---

[1] P. und T. Ehrenfest, Begriffliche Grundlagen der statistischen Affassung in der Mechanik, в „Encyklopädie der mathematischen Wissenschaften", Bd. IV, Teil 4, D, Leipzig, 1911, Art. 32, S. 43.

[2] См., например, трактовку Больцманом парадокса обратимости, данную в письме к редактору журнала „Nature", Vol. 51, 1895, p. 413—415. Более подробно изложение этого взгляда можно найти в первом ответе Больцмана Й. Лошмидту: „Bemerkungen über einige Probleme der mechanischen Wärmetheorie", в „Sitzungsber. d. Kaiserl. Akad. d. Wiss. in Wien", math.-naturwiss. Cl., Bd. 75, 1877, 2. Abt., S. 62—100 (см. особенно S. 72) — и в его последней публикации о теории газов, написанной совместно с Дж. Наблом: „Kinetische Theorie der Materie", в „Encyklopädie der mathematischen Wissenschaften", Bd. V, Teil 1, Art. 8, 1907, No. 14, S. 519—522. В этих публикациях Больцман неоднократно говорит, что «второй закон термодинамики никогда не удастся доказать математически на основе одних лишь уравнений движения» (см. его письмо в „Nature", S. 414). Это замечание нельзя истолковать в том смысле, что Больцман полагал невозможным доказательство, подобное доказательству, данному Джоном Нейманом и Георгом Биркгофом. Больцман просто хотел сказать, что нельзя доказать, что энтропия всегда возрастает. Мы можем доказать только, что такое возрастание весьма вероятно.

Однако, так как тот же самый вывод можно получить для состояния B, которое во времени предшествует A, мы видим, что этот вывод об энтропии на основе времени не предполагает временной направленности кривой энтропии, он просто выражает тот факт, что эта кривая обладает относительно высокими и низкими значениями и остается при высоких значениях намного дольше, чем при низких. Кривую этого вида можно использовать для того, чтобы определить различие между высокими и низкими уровнями, но не для того, чтобы определить различие между левым и правым, то есть положительное и отрицательное время.

Рассмотренные выводы можно проиллюстрировать на следующей модели. Представим себе плато средней высоты в 4000 футов, изрезанное местами каньонами различной глубины, которые тянутся с севера на юг, и пусть при этом вертикальное поперечное сечение в направлении с запада на восток через эту возвышенность имеет форму кривой, представленной на рис. 19. Предположим, что некто пересекает плато в направлении строго на восток. При подходе к каньону он спускается и затем подымается с той же самой скоростью на другую сторону, всегда придерживаясь своего основного направления на восток. Теперь можно сделать следующее утверждение. Если этот человек находится на высоте 2000 футов (в каньоне), тогда через 2 часа он, вероятно, будет находиться еще выше. Но можно также сказать: если этот человек находится на высоте 2000 футов, тогда за 2 часа до этого он, вероятно, находился на более высоком уровне. Оба утверждения справедливы. Эта симметричность выражает тот факт, что поперечное сечение через плато не определяет направления, но показывает, что более вероятно оказаться на высоком уровне, чем на низком.

Согласно этим результатам, мы могли бы предположить, что система гораздо чаще будет находиться вблизи седловой точки. Это заключение неверно. Оно означало бы, что в пределах небольшого интервала времени Δt состояние с более высокой энтропией и предшествовало, и следовало бы за состоянием A. Но для небольших Δt вероятность каждого из этих случаев не очень велика, а вероятность их совместного наступления

меньше, чем вероятность каждого случая в отдельности. Можно показать, что она даже меньше, чем произведение этих двух вероятностей, то есть эти события не независимы, но до некоторой степени несовместимы. Следовательно, существует лишь малая вероятность того, что мы окажемся вблизи седловой точки.

Рассмотрим теперь второй вопрос, который касается вывода о времени на основе энтропии. Он формулируется следующим образом. Пусть два состояния A и B заданы таким образом, что энтропия S_B состояния B больше, чем энтропия S_A состояния A. Какое из этих двух состояний является более поздним?

Из предыдущего следует, что на этот вопрос нельзя ответить, используя кривую энтропии. Вероятность того, что A предшествует B, равна вероятности того, что B предшествует A. Такой вывод следует также из нашего рассмотрения парадокса обратимости (см. стр. 152—156). В таком случае очевидно, что для определения направления времени мы должны суметь ответить на второй вопрос, касающийся вывода о времени на основе энтропии. Наши ответы на первый вопрос относительно вывода об энтропии на основе времени не могут нам помочь, если мы еще не знаем направления времени. Отсюда следует, что, если мы хотим использовать кривую энтропии для определения направления времени, парадокс обратимости сохраняет свою силу, потому что эта кривая не дает возможности определить такое направление. Парадокс статистического подхода к определению направления остается нерешенным: обратимость элементарных процессов переносится на макропроцессы; следовательно, нельзя доказать, что статистически определенное направление изменения изолированных систем, точно так же как и направление самих элементарных процессов, является уникальным.

Если мы хотим разрешить этот парадокс и ответить на вопросы второго вида, то мы должны, следовательно, рассмотреть другие методы. Кривая энтропии и вместе с нею парадокс обратимости относятся к временно́му ансамблю, а именно к изменениям во времени одной системы. Мы сумеем лучше ответить на наш вопрос, если мы оставим временно́й ансамбль и обратимся к изучению пространственного ансамбля.

§ 14. Направление времени в пространственном ансамбле

Если мы хотим найти способы определения направления времени, то целесообразно изучить те реальные методы, которые используются для получения выводов о направлении времени.

Мы видели, что если $S_B > S_A$, то теория временнóго ансамбля приводит к утверждению, что A может быть как раньше, так и позднее, чем B. Однако во всех практических случаях нам весьма желательно прийти к заключению, что если $S_B > S_A$, то A раньше B. Например, если один наблюдатель говорит нам, что он видел газы в контейнере довольно строго разделенными, хотя в нем и не имелось перегородки, а второй наблюдатель информирует нас, что он видел газы достаточно перемешанными, мы сделаем вывод, что второе наблюдение было произведено позднее первого. Мы также добавим, что первоначально газы были разделены перегородкой и что некто должен был удалить перегородку незадолго до того, как было произведено первое наблюдение. Это означает, что мы исходим из предположения, что газ первоначально находился во взаимодействии со своим окружением, а не из предположения, что газ все время образовывал изолированную систему. Мы также делаем вывод, что маловероятное состояние скорее является результатом этого взаимодействия, чем результатом совершенно случайного процесса разделения, которое произошло в ходе развития замкнутой системы. Так как при этом выводе делается ссылка не только на наблюдаемые системы, то он не характеризуется кривой на рис. 19 и, следовательно, должен быть рассмотрен более внимательно.

Наше окружение изобилует процессами, которые либо в результате естественного развития событий, либо вследствие вмешательства человека приводят к образованию упорядоченных подсистем, которые после этого остаются изолированными и развиваются в направлении к беспорядку, как изображено на начальном участке *1—2* кривой на рис. 19. Под «изолированными» мы не подразумеваем системы, находящиеся в полной изоляции; достаточно, если по величине энергетического обмена процесс внутри подсистемы значительно превосходит взаимодействие с окружением. В § 7 упоминалось о рефрижераторе как

примере подсистем с низкой энтропией. Если мы возьмем кусок льда из рефрижератора и опустим его в стакан с водой, мы получим изолированную систему, хотя тепловая изоляция и не будет достаточно совершенной; таяние льда изображается участком *1—2* кривой энтропии на рис. 19. Камень, лежащий на снегу, нагревается солнцем; в ходе этого процесса система, состоящая из солнца и камня, увеличивает свою энтропию вследствие поглощения излучения камнем. Ночью камень и снег совместно образуют систему, которая вследствие своего внутреннего различия температур имеет низкую энтропию; охлаждение камня и таяние снега образуют процесс, ведущий к увеличению энтропии. Химические процессы, такие, как горение или процесс дыхания живых организмов, создают температурные различия со своим окружением, что приводит к процессу уравновешивания, идущему в направлении от порядка к беспорядку. В этих случаях создание упорядоченных начальных состояний — куска льда, плавающего в стакане с водой, температурных различий вследствие химических процессов — достигается не в изолированных системах, в которых происходит процесс уменьшения энтропии; эти начальные состояния создаются в подсистемах, входящих в более объемлющие системы, полное значение энтропии которых возрастает, в то время как подсистема приводится в состояние с относительно низкой энтропией. Природа изобилует *ответвившимися системами* подобного вида, то есть системами, которые ответвились от более обширной системы и после этого остаются изолированными в течение определенного времени. Их развитие начинается с упорядоченного состояния, то есть с состояния с относительно низкой энтропией, и происходит в направлении к беспорядку, то есть к относительно высокой энтропии. Мы здесь используем слово «относительно», чтобы указать, что эта энтропия относится к подсистеме, а не ко вселенной или основной системе.

Таковы наблюдаемые факты, на которые должно опираться статистическое определение направления времени. Когда мы из неравенства $S_B > S_A$ делаем вывод, что, по всей вероятности, A произошло раньше B, то это не та вероятность, которая рассматривалась в связи с парадоксом обратимости: она относится не к последовательности состояний изолированной системы, в которую входят

состояния A и B, а к серии подобных систем, которую можно охарактеризовать как ансамбль ответвившихся систем. Другими словами, она относится не к временно́му ансамблю, а к пространственному.

Чтобы изучить эти вероятности, сделаем сначала некоторые упрощения. Рассмотрим участок продолжительного подъема кривой энтропии вселенной и ограничим

Рис. 20. Начальный участок подъема кривой энтропии вселенной. Некоторые изолированные системы ответвляются, оставаясь изолированными в течение бесконечного времени. (Упрощенное схематическое представление.)

нашу статистику системами, ответвившимися на этом участке подъема, предполагая, что эти ответвившиеся системы остаются изолированными в течение бесконечного времени. На рис. 20 дано схематическое изображение этой проблемы. Мы также допускаем, что кривые ответвившихся систем изображены в большем масштабе, чем кривая полной энтропии вселенной.

Теперь мы представим мировые линии ответвившихся систем посредством вероятностной структуры, введенной в (12.6). В противоположность рассмотренной там структуре, описывающей поведение молекул, в данном случае структура относится к системам; другими словами, каждая строка представляет мировую линию целой системы, и мы записываем переход этой системы из одного состояния в следующее и т. д., причем каждое состояние

характеризуется через его энтропию. Для упрощения мы не станем рассматривать непрерывные вероятностные последовательности, для чего представим эволюцию системы как последовательность дискретных событий, которые следуют друг за другом через небольшие интервалы времени Δt, точно так же как это было сделано при рассмотрении поведения молекул. Обозначая индивидуальные события в эволюции системы через y, мы получим следующую вероятностную структуру:

$$\begin{matrix} y_{11} & y_{12} & y_{13} & y_{14} & \cdots & y_{1l} & \cdots \\ y_{21} & y_{22} & y_{23} & y_{24} & \cdots & y_{2l} & \cdots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ y_{k1} & y_{k2} & y_{k3} & y_{k4} & \cdots & y_{kl} & \cdots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{matrix} \qquad (1)$$

Атрибутивное пространство каждой строки задается энергетической поверхностностью в фазовом пространстве отвествившейся системы. Если A — некоторое значение энтропии, определенное достаточно точно, тогда A представляет определенную область в атрибутивном пространстве. Если состояние y_{kl} системы обладает этим значением энтропии, мы говорим, что элемент y_{kl} принадлежит к классу A. Первый элемент строки y_{k1} характеризует весьма упорядоченное состояние; в последующих элементах порядок постепенно исчезает, пока не будет достигнуто состояние беспорядка. То есть каждая строка обладает последействием и изменяет свои атрибуты постепенно. Если продолжить строки соответственно увеличению периода времени до значительных масштабов, то на кривых энтропии встретятся участки, где будут уменьшаться значения энтропии, как показано на рис. 20.

Таким образом, мы имеем здесь структуру процесса смешивания и можем применить символику и теоремы, изложенные в § 12. Если A и B — состояния с различными значениями энтропии, то существует горизонтальная вероятность того, что B следует за A через m элементов; эта вероятность записывается в виде

$$P\left(A^{kl},\ B^{k,\ l+m}\right)^{l}. \qquad (2)$$

Если A — состояние с небольшой энтропией, а B — с высоким ее значением, и если m не слишком мало, то вероятность (2) весьма велика.

Можно также определить вероятность того, что B предшествует A на m элементов. Эта вероятность получается из выражения (2) при изменении знака m, предполагая, что $m < l$. Вследствие того, что строка симметрична относительно направления, как это было показано в связи с парадоксом обратимости, мы получим здесь соотношение

$$P(A^{kl},\ B^{k,\ l+m})^l = P(A^{kl},\ B^{k,\ l-m})^l. \qquad (3)$$

Именно это равенство не позволяет нам использовать строку для определения направления времени.

Мы можем также выразить символически вертикальную вероятность, когда частота подсчитывается вдоль столбца:

$$P(A^{kl},\ B^{k,\ l+m})^k. \qquad (4)$$

Это выражение означает, отметим еще раз, вероятность того, что состояние B следует за A через m элементов в той же самой строке, но здесь эта вероятность подсчитывается в вертикальном направлении, то есть в ансамбле различных ответвившихся систем. Если A и B имеют тот же самый смысл, что и в (2), мы получаем соотношение

$$P(A^{kl},\ B^{k,\ l+m})^k = P(A^{kl},\ B^{k,\ l+m})^l,\ m > 0. \qquad (5)$$

Это означает, что мы исходим из допущения, что структура удовлетворяет условию инвариантности, изложенному в (12.18). Данное допущение является эмпирическим и подтверждается многими экспериментами.

В этом и состоит, на наш взгляд, физическое значение условия инвариантности структуры для пространственного ансамбля ответвившихся систем. Для пространственного ансамбля трудно получить приемлемое определение вероятностей. Но нам довольно хорошо известна метрика вероятностей временнóго ансамбля, потому что она определяется теоремой Лиувилля (см. § 10). Допущение об инвариантности структуры позволяет нам использовать эту метрику временнóго ансамбля для пространственного ансамбля.

Однако условие инвариантности структуры *не* применяется к вероятностям *предшествующих* событий, ибо мы исходим из предположения, что все ответвившиеся системы начинают свое развитие с состояний, характеризуе-

мых низкими значениями энтропии. *Уравнение (5) становится неверным при $m < 0$.* С другой стороны, если A — состояние с низкой энтропией, а B — состояние с высокой энтропией, то при вертикальном подсчете мы обнаруживаем, что в большинстве случаев энтропия возрастает при переходе вправо, пока мы еще не слишком далеко ушли от начала строк. Следовательно, для небольших i получаем неравенство

$$P(A^{ki},\ B^{k,\ i+m})^k > P(A^{ki},\ B^{k,\ i-m})^k,\quad 0 < m < i. \quad (6)$$

В противоположность соотношению (3) в данном случае мы получаем: вероятность того, что за состоянием с низким значением энтропии следует состояние с более высоким значением энтропии, больше вероятности, что тому же самому состоянию с низким значением энтропии предшествует состояние с высоким значением энтропии. Парадокс обратимости *не* имеет отношения к данному случаю, потому что мы рассматриваем вертикальные вероятности, то есть статистику ансамбля ответвившихся систем, а не эволюцию одной-единственной системы. Назовем горизонтальную вероятность *вероятностью по одной системе*, а вертикальную вероятность — *вероятностью по множеству систем*. Тогда решение парадокса обратимости основывается на различии между вероятностями по одной системе и по множеству систем, или, что то же самое, на различии между вероятностями, относящимися к временному ансамблю, и вероятностями, относящимися к пространственному ансамблю.

Если все ответвившиеся системы, представленные в структуре, имеют одну и ту же физическую природу и одинаковые начальные состояния A и если A, как и раньше, есть состояние со сравнительно низким значением энтропии, то, используя (5), мы получим

$$P(A^{k,\ m+1},\ B^{k,\ m+i})^m = P(A^{ki},\ B^{ki})^k = P(B^{ki})^k. \quad (7)$$

Это означает, что абсолютная вертикальная вероятность обнаружить B в i-ом столбце (последний член соотношения) непосредственно измеряет относительную горизонтальную вероятность того, что состояние B следует за высокоупорядоченным состоянием A через $i - 1$ элементов (первый член). Высокое значение вероятности по одной системе того, что состояние B с высоким значением эн-

тропии следует за состоянием A с низкой энтропией, здесь выражается через вероятность по множеству систем B.

Если энтропия B имеет среднее значение, промежуточное между низким и высоким значениями, то состояние B не будет часто встречаться в ходе развития индивидуальной системы. Таким образом, мы имеем

$$P(A^{k,\,m+1},\ B^{k,\,m+i})^m > P(B^{k,\,m+i})^m = P(B^{ki})^i. \qquad (8)$$

Однако такое состояние B с промежуточным значением энтропии довольно часто может встретиться в ансамбле ответвившихся систем; это вытекает из соотношения (7), которое в соединении с (8) приводит к выражению

$$P(B^{ki})^k > P(B^{ki})^i. \qquad (9)$$

Это неравенство аналогично соотношению (12.9) и выражает наличие влияния высокоупорядоченного начального состояния A на начальные элементы каждой строки. Но это влияние исчезает, и мы получаем

$$\lim_{i\to\infty} P(B^{ki})^k = P(B^{ki})^i. \qquad (10)$$

Это соотношение аналогично соотношению (12.10). Кроме того, вероятность в правой части соотношения (6) возрастает и сходится к значению вероятности слева, которая остается постоянной при этом переходе. Следовательно, при переходе к пределу $i \to \infty$ соотношение (6) превращается в равенство, соответствующее (3).

Данная вероятностная структура представляет структуру процесса смешивания, в которой все элементы первого столбца относятся к одному и тому же состоянию, описанному соотношениями (12.9—14). Структура этого вида определяет направление через отношения частот в своих начальных столбцах, а пространственный ансамбль определяет направление времени как перпендикулярное этому направлению, если последнее содержится в структуре процесса смешивания.

Эти соображения показывают, что структура для систем может рассматриваться как выражающая процесс смешивания в том же самом смысле, как и структура для молекул. Это справедливо, несмотря на значительное физическое различие между этими двумя случаями: молекулы постоянно сталкиваются между собой, и это их

взаимодействие выступает в качестве перемешивающего механизма, в то время как системы не взаимодействуют одна с другой и не перемешиваются посредством некоторого физического механизма. Функцию такого механизма здесь берут на себя законы случая, управляющие независимыми рядами событий. Эти законы, выраженные в условиях независимости и инвариантности структуры, в совокупности постепенно разрушают порядок, заключенный в конфигурации первого столбца, который содержит только состояния с низкими значениями энтропии, а именно состояния, соответствующие моменту ответвления систем от основной.

Когда мы изображаем процесс смешивания посредством вероятностной структуры, подчиняющейся математическим соотношениям, изложенным в § 12, мы можем сказать, что в случае ансамбля ответвившихся систем направление времени вытекает из процесса смешивания систем, в том же самом смысле, как и в случае процесса диффузии, где направление времени проистекает из смешивания молекул. Оба вида ансамблей представляют процесс смешивания, потому что они удовлетворяют условиям структуры смешивания, в которой первый столбец состоит из избранных состояний. Из этого следует, что процессы смешивания, в самом общем смысле этого термина, представляют собой инструменты, указывающие направление времени. Они могут выполнять эту роль, потому что они переводят симметрию горизонтальных вероятностей в асимметрию вертикальных вероятностей, или симметрию относительно направления во временны́м ансамбле в асимметрию пространственного ансамбля. Это приводит к наличию двух разных смыслов в выражении: «Вероятность того, что состоянию с низкой энтропией предшествует состояние с высокой энтропией», — и, таким образом, дает возможность объяснить вывод относительно времени на основе знаний об энтропии.

Можно сделать одно замечание для логиков. В исследованиях, представленных в этой книге, всегда предполагается, что термин «вероятность» означает предел относительной частоты, как уже объяснялось в (12.2). Эта частотная интерпретация вероятности, ныне широко принятая среди ученых, может объяснить все случаи использования термина «вероятно», которые могут быть проверены посредством предсказаний и, следовательно, обла-

дают прорицательной ценностью. Это означает, что если вероятностное утверждение в этой интерпретации истинно, то можно показать, что целесообразно его использование для предсказаний. Существует законченная математическая и логическая теория вероятностей, основанная на частотной интерпретации.

Однако когда мы спрашиваем о вероятности p того, что данному состоянию A с низким значением энтропии предшествует состояние с высокой энтропией, то вопрос относится к единичному случаю. Из частотной интерпретации следует, что вероятностное утверждение относительно единичного случая не имеет смысла. Чтобы оправдать применение вероятностного утверждения к единичным случаям, эти утверждения нужно заменить утверждениями, относящимися к частотам. Это требует, во-первых, выбора соответствующего обусловливающего класса [1], и, во-вторых, мы также должны выбрать соответствующую последовательность [2]. Пример с вероятностью p показывает необходимость второго требования: обусловливающий класс A и атрибутивный класс B не определяют эту вероятность; используя те же самые классы, мы получим высокое или низкое значение для p, соответственно тому, что мы выберем в качестве нашей последовательности — строку или столбец. Эта двусмысленность, которая является необходимым следствием рассматриваемых структурных конфигураций, показывает, что нельзя говорить содержательным образом о вероятности единичного случая. Логики, которые отвергают частотную интерпретацию и пытаются дать особое определение вероятности для единичного случая безотносительно к частотам, должны проверить свои теории на проблемах, подобных парадоксу обратимости; тогда они быстро обнаружат, что их системы не охватывают всех случаев использования вероятностей в практических применениях.

Встречающиеся в (6) вертикальные вероятности еще не тождественны вероятностям, используемым при выводе о времени на основе энтропии, потому что вероятности в (6) предполагают знание направления времени. Однако вероятность v, выражающая переход от энтропии

---

[1] См. ThP, p. 374.
[2] См. ibid., p. 376.

ко времени, может быть получена[1] из вероятностей, используемых в (6). Чтобы показать, что эта вероятность v также относится к столбцам, а не к строкам структуры, дадим здесь строгое символическое представление этой вероятности.

Сначала мы определим следующим образом вероятность v_{im}, относящуюся к i-му столбцу. Обусловливающий класс задается теми состояниями y_{ki}, которые принадлежат к A *и* за которыми следует *или* которым предшествует в m-ом элементе состояние B. Атрибутивный класс задается только случаями первого вида, то есть состояниями y_{ki}, принадлежащими A, за которыми в m-ом элементе *следует* состояние B. Другими словами, из всех случаев, когда состояние B следует или предшествует на m-ом месте состоянию A, мы принимаем во внимание только случаи первого вида. Эта вероятность v_{im} символически записывается в следующей форме:

$$v_{im} = P(A^{ki} . B^{k,\, i+m} \vee A^{ki} . B^{k,\, i-m},\ A^{ki} . B^{k,\, i+m})^k. \qquad (11)$$

Знак $\vee$ означает «или», а точка означает «и». Вероятность этого вида, в которой обусловливающий класс имеет форму дизъюнкции, вычисляется на основе правила редукции[2] следующим образом:

$$v_{im} = \frac{P(A^{ki} . B^{k,\, i+m})^k}{P(A^{ki} . B^{k,\, i+m})^k + P(A^{ki} . B^{k,\, i-m})^k - P(A^{ki} . B^{k,\, i+m} . B^{k,\, i-m})^k}. \qquad (12)$$

Каждый из этих членов имеет общую форму $P(A.B)$ и может быть записан, согласно общей теореме умножения, в виде $P(A) \cdot P(A, B)$. Получаемый таким образом первый множитель $P(A^{ki})^k$ является одним и тем же для всех членов и опускается (он больше 0, потому что мы предполагаем, что состояние A встречается в i-ом столбце). Мы получаем

$$v_{im} = \frac{P(A^{ki},\ B^{k,\, i+m})^k}{P(A^{ki},\ B^{k,\, i+m})^k + P(A^{ki},\ B^{k,\, i-m})^k - P(A^{ki},\ B^{k,\, i+m} . B^{k,\, i-m})^k}. \qquad (13)$$

---

[1] Это заключение, по-видимому, предполагает определение, данное в § 15 этой книги (стр. 174). — *М. Р.*

[2] ThP, p. 100, формула (17). В данном случае я опускаю ссылку на общий обусловливающий класс и использую абсолютную символику, введенную в ThP, p. 106.

Далее, для всех практических целей вероятность, стоящая в правой части (6), будет пренебрежимо мала, если $S_B > S_A$ и i не слишком велико. Поэтому второй и третий члены в знаменателе опускаются, так как легко показать, что третий член даже меньше второго. Отсюда мы находим

$$v_{im} \sim 1, \tag{14}$$

где знак $\sim$ означает «приближенно равно». Если i и m точно не определены, вероятность v получается как среднее число, или среднее значение[1], различных v_{im} и мы имеем

$$v = M_I(v_{im})_{im} \sim 1. \tag{15}$$

Индекс I, добавленный к символу среднего, может указывать, что значения i не должны быть слишком велики; то есть подобно неравенству (6), соотношение (15) применяется только к начальным столбцам вероятностной структуры.

Уже из первоначального рассмотрения структуры (1) в этом параграфе результат (15) кажется очевидным. Но все же полезно получить его на языке символов, потому что это раскрывает логическую структуру применявшихся здесь выводов. Такие выводы используются во многих ситуациях, благодаря чему мы находим, что упорядоченные состояния предшествуют неупорядоченным. Например, если на одной фотографии с воздуха, сделанной во время войны, мы видим город неразрушенным, тогда как на второй видим поверженные в руины дома, то мы знаем, которая из них сделана раньше.

Исходя из данного состояния B, которое имеет среднее значение энтропии, можно сделать иное по характеру заключение относительно прошлого, состоящее в том, что состоянию B предшествует состояние A с более низким значением энтропии. Такой вывод прямо основывается на соотношении

$$P(B^{ki},\ A^{k,\,i+m})^k \sim 1, \tag{16}$$

которое справедливо для начальных столбцов структуры (1), если $S_B > S_A$. Например, когда мы видим горящую сигарету в пепельнице, мы заключаем, что незадолго

[1] Обозначения см. в ThP, p. 180.

до этого она не горела и была целой; или же, если мы находим вечером камень, который теплее окружающего воздуха, то мы делаем вывод, что днем он был еще теплее и был нагрет солнцем. Заметим, что когда мы говорим о том, что более раннее состояние имело низкое значение энтропии, то мы имеем в виду полную систему, включающую все происшедшие изменения. Например, в данном случае следует учесть излучение остывающего камня. Соотношение (16) показывает, что вероятности по множеству систем позволяют нам сделать вывод, что в прошлом энтропия имела низкое значение. Мы видели, что такое заключение нельзя было вывести для вероятностей по одной системе. Только теория *пространственного* ансамбля позволяет разрешить парадокс обратимости.

Мы проводили выкладки в предположении, что *все* строки схемы (1) начинаются с состояний с относительно низкой энтропией. Если это не вполне справедливо в том смысле, что только подавляющее большинство строк удовлетворяет этому предположению, то можно получить те же самые результаты. Нет необходимости приводить доказательство этого более общего случая, поскольку вероятностные отношения остаются здесь неизменными.

§ 15. Секционный характер направления времени

Решение парадокса обратимости на основе введения вероятностей, относящихся к ответвившимся системам, указывает нам способ определения направления времени. Однако наше обсуждение до сих пор было ограничено схемой ответвившихся систем, приведенной на рис. 20. Эта схема неадекватна в двух отношениях. Во-первых, на ней представлен один только участок подъема кривой энтропии, а во-вторых, ответвившиеся системы рассматриваются как бесконечные во времени.

Прежде всего мы знаем, что продолжение основной кривой, представленной на рис. 20, приведет к переходу от подъема к горизонтальному участку и далее к падениям и подъемам, чередующимся нерегулярным образом, как указано на рис. 19. Затем нам известно, что ответвившиеся системы обычно находятся в состоянии изоляции только в течение ограниченного периода времени, а затем вновь приходят в соединение со своим окружением. Если мы нальем кофе и сливки в термос, то обычно

полученная смесь не остается в сосуде в течение слишком долгого времени; если нагретый солнцем камень вызывает таяние снега, то вода просачивается в землю и входит в общую систему окружающей почвы.

Объединяя первое соображение со вторым, мы получим кривую, подобную той, которая представлена на рис. 21; однако на этой диаграмме изображены только один подъем и одно падение, в то время как мы можем предположить, что эта кривая продолжается в обе стороны одинаково. На каждом участке показаны немногие

Рис. 21. Участок подъема и падения кривой энтропии вселенной. Изолированные системы ответвляются от основной системы и возвращаются к ней.

ответвившиеся системы, возвращающиеся к основной кривой.

Если мы хотим применить наши предыдущие вероятностные суждения к этому примеру, то возникают трудности. Горизонтальные последовательности здесь не являются бесконечными, что не должно нас смущать, так как для всех практических случаев вполне достаточны последовательности ограниченной длины, при условии, что они достаточно длинны. Столбцы в схеме (14.1) также конечны, потому что конечно число систем, ответвившихся от главной на участке ее подъема, но, как уже указывалось, это число чрезвычайно велико. Однако последовательности, соответствующие ответвившимся системам и представленные на рис. 21, не являются достаточно длинными, чтобы служить основанием для реализации горизонтальных вероятностей; данный рисунок содержит только начальные участки подъема ветвей, изображенные на рис. 20, далее следуют небольшие гори-

зонтальные участки, а затем возвращение к основной кривой. В действительно наблюдаемых нами ответвившихся системах разделение их составляющих никогда не встречается. Горизонтальные вероятности могут быть реализованы только для основной кривой, так как она имеет неограниченную длину.

Наоборот, вертикальные вероятности реализуются посредством частот, но здесь оказываются не удовлетворенными условия схемы (14.1). На рис. 21 только для участка подъема основной кривой ответвившиеся системы начинают развитие с состояния с низкой энтропией, что отмечено на рисунке точками *1*, *3* и *5*. Когда мы переходим к участку спуска кривой, всегда происходящего в том же самом направлении, ответвления начинаются при состояниях с высокими значениями энтропии, а именно в точках *7*, *9* и *11*, и заканчиваются в точках с низкой энтропией, а именно в точках *8*, *10* и *12*, соответственно. Если теперь мы продолжим структуру (14.1), используя системы, ответвившиеся на участке спуска и на дальнейших участках основной кривой, то мы встретимся с равными частотами перехода к высоким и низким значениям энтропии, и соотношение (14.6) заменится равенством этих двух вероятностей. И, таким образом, станет невозможным определение направления времени.

Отсюда следует, что направление времени можно определить только для отрезков кривой полной энтропии. В этих случаях применяется неравенство (14.6), что позволяет нам дать следующее определение.

Определение. То направление, в котором протекает большинство термодинамических процессов в изолированных системах, и представляет направление положительного времени.

Предположим, например, что состояние вселенной в настоящее время соответствует участку подъема кривой энтропии. Тогда доступная нам статистика всегда будет относиться к вертикальным вероятностям, или вероятностям по множеству систем, ограниченным этим участком подъема основной кривой. Горизонтальные вероятности, подобные тем, которые используются в (14.3), в форме частот нами не могут наблюдаться, потому что временные периоды возможных наблюдений довольно коротки. Следовательно, наблюдаемые вероятности всегда будут удовлетворять неравенству (14.6) и позволят определить

положительное направление времени. Однако мы знаем, что эта статистика выражает просто общую тенденцию на данном участке кривой энтропии, на котором нам довелось жить. Полная энтропия мира в его нынешнем состоянии не слишком высока: вселенная обладает, так сказать, большими резервами в отношении упорядоченных состояний, которые расходуются при создании ответвляющихся систем и, таким образом, позволяют нам определить направление времени.

Это соображение приводит к своеобразным следствиям. Определение направления времени, прослеженное для участка подъема основной кривой, с равным успехом можно применить и для участка спада кривой. Однако в этом случае будет определено направление, обратное направлению положительного времени: если мы определим соответствующим образом положительное время, то, исходя из событий на этом участке общей кривой, можно еще раз сказать, что вселенная развивается по направлению к более высокой энтропии. Отсюда следует, что мы не можем говорить о направлении времени, взятом в целом; только определенные отрезки времени обладают направлениями, и эти направления не одинаковы.

Первым, кто нашел в себе смелость сделать это заключение, был Людвиг Больцман[1]. Его идея об изменениях направления времени, которые только секционно определены статистическими процессами, представляет одно из наиболее глубоких проникновений в проблему времени. Философы пытались вывести свойства времени из общих рассуждений, но ни одна из их концепций не сравнима с результатом, который получил физик из *рассуждения* относительно выводов математической физики. Как и во многих других случаях, здесь проявила свое превосходство философия, основанная на данных науки. Нет никакой логической необходимости в том, что существует единое направление времени; вид кривой энтропии, которую вычерчивает вселенная, определяет собой то, существует ли только одно направление времени или же время меняет направление.

Больцман очень ясно показал, что представления об изменении в направлениях времени не являются абсурд-

---

[1] Л. Больцман, Лекции по теории газов, М., 1956, стр. 525–527.

ными. Он относит направление нашего времени к тому участку кривой энтропии, на котором мы живем. Предположим, что вселенная может достигнуть состояния с высокой энтропией и оставаться в этом состоянии в течение продолжительного времени. И если бы произошло, что «позже» кривая энтропии вселенной стала бы длительно спадать, тогда для этого участка кривой время имело бы противоположное направление: люди, которые могли бы жить в этот период развития вселенной, рассматривали бы как положительное время переход к высокой энтропии, и, таким образом, их время текло бы в направлении, обратном нашему. Так как эти два участка, где время имеет противоположные направления, были бы разделены эпохами с высокими энтропийными состояниями, при которых живые организмы не могут существовать, то для обитателей второго временного участка навсегда осталось бы неизвестным, что их направление времени отличается от нашего. Подобным образом два участка, начинающиеся в седловой точке и идущие в противоположных направлениях, были бы разделены продолжительными периодами, когда не могла бы поддерживаться жизнь, потому что при состояниях с низкой энтропией существовали бы высокие температуры, локализованные в местах сосредоточения материи, напоминающих раскаленные газовые шары. Жизнь ограничивается умеренными поясами на участках перехода в значениях кривой энтропии. Таким образом, чередования в направлениях времени не приводили бы ни к каким противоречиям в нашем опыте. Возможно, что мы являемся на самом деле жителями второго участка, когда энтропия «реально» уменьшается, но мы об этом не знаем.

Что означают такие утверждения? Очевидно, не имеет смысла говорить об одном из участков, что его энтропия «реально» растет или что его направление времени «реально» положительно. Больцман сравнивал эту проблему с вопросом, кто — мы или наши антиподы — ходят «реально» вверх головой. Однако имеется смысл в выражении, что вселенная состоит из отдельных нитей времени, соединяемых вместе разными концами, потому что временной *порядок* можно определить в классической механике (см. § 5) и без предположений об энтропии. Следовательно, мы можем говорить о *сверхвремени*, которое

упорядочивает кривую даже на участках равновесия, где энтропия практически постоянна, или в седловых точках, где градиент энтропии меняет свое направление. Сверхвремя не имеет направления, но только порядок, в то же время оно содержит индивидуальные участки, которые обладают направлением, хотя эти направления и изменяются от участка к участку. Однако для горизонтальных участков кривой энтропии направление времени нельзя определить содержательным образом [1].

Как мы узнаем, что наша вселенная в настоящее время находится в таком длительном состоянии, которое мы называем состоянием возрастания энтропии? Это заключение нельзя основывать только на том факте, что в настоящее время энтропия вселенной не слишком велика. Исходя только из вероятностей по одной системе, мы могли бы сделать вывод, что в будущем энтропия вселенной будет больше; но мы видели, что этот же вывод можно сделать в отношении прошлого, и, следовательно, мы могли бы с равным правом заключить, что в прошлом энтропия была еще выше. Даже оба заключения могли бы быть справедливыми, и мы оказались бы в седловой точке. Когда мы настаиваем на том, что в прошлом энтропия была более низкой, мы приходим к заключению иными путями.

Указанный вывод можно получить двумя способами. Во-первых, мы можем использовать причинные законы, то есть следовать детерминистским методам, и, во-вторых, можно исходить из вероятностей ответвившихся систем. Рассмотрим сначала детерминистский метод.

Мы уже знаем, что законы механики определяют порядок времени и также позволяют нам по данному состоянию системы сделать заключение о ее предыдущем и будущем состояниях. Законы планетарного движения можно использовать не только для предсказания будущего состояния планет, но также и для выводов о состояниях этой системы в предыдущие времена. Распространяющееся в пространстве солнечное излучение подчиняется дифференциальным уравнениям Максвелла и,

---

[1] Больцман также говорил (там же, стр. 526—527) о возможности того, что вселенная может состоять из отдельных областей, разделенных обширными пустыми пространствами, которые одновременно имеют противоположные направления времени. См. § 16 данной книги, особенно стр. 191—192.

подобно движению планет, представляет причинно упорядоченный процесс. Мы знаем, что материя во вселенной встречается весьма редко, и, следовательно, солнце излучает в практически пустое пространство. Распространение излучения в вакууме представляет процесс увеличения энтропии, но это такой процесс, направление которого не может измениться чисто случайным образом, а для перемены направления необходимо вмешательство извне. Только стенки из материи, отражающей излучение, могут изменить на противоположное направление его распространения. Следовательно, вывод относительно прошлого нашей вселенной может быть сделан практически в рамках детерминистских методов и основывается на детальном знании состояния вселенной в настоящее время. Простого знания ее энтропии в данный момент недостаточно для такого вывода. Вспомним детерминистскую интерпретацию эргодической траектории (§ 10); там мы видели, что если известно местоположение фазовой точки, то ее траекторию в прошлом и будущем можно рассчитать. Применяя подобные методы к макрокосму, мы можем рассчитать прошлое и будущее солнечной системы и даже галактик.

Применяя эти результаты к кривой энтропии вселенной, можно сказать, что они устанавливают наличие сильного последействия в вероятностной последовательности ее состояний. Именно это последействие вызывает продолжительные подъемы и спуски этой кривой, и ему же мы обязаны наличием секционных направлений времени. Но действительные заключения о прошлом не основываются просто на таких общих соображениях о вероятности; в этих случаях используются подробные описания состояний, когда выходят далеко за рамки характеристики состояний по энтропии, и применяются причинные законы, относящиеся к временному порядку этих состояний.

Природу этого вывода можно проиллюстрировать ссылкой на процесс диффузии в газах. Когда мы наблюдаем систему из двух газов в частично смешанном, частично разделенном состоянии, то есть в состоянии со средним значением энтропии, и нам известно, что система всегда будет оставаться в состоянии изоляции, мы заключаем, что в прошлом энтропия системы была выше по той же причине, по которой мы заключаем, что она

будет выше в будущем. Но если бы мы были в состоянии наблюдать положение и скорость каждой молекулы в отдельности, то мы смогли бы рассчитать прошлое системы, используя законы механики, и не привлекали бы выводы, основанные на понятии энтропии. На этом пути мы могли бы обнаружить, что система раньше находилась в состоянии, при котором ее компоненты были полностью разделены, и, следовательно, ее энтропия имела даже более низкое значение. Таким образом, детальное познание существующих условий делает не обязательным вывод, основанный на представлении об энтропии, и открывает пути использования причинных законов для получения знания о прошлом. Хотя мы и не можем использовать такие методы для газов, мы можем применить их для установления прошлого вселенной, потому что мы обладаем довольно детальным знанием ее настоящего состояния и в общих выводах не связаны представлением об энтропии.

Рассмотрим теперь второй способ вывода заключений о прошлом вселенной. Здесь мы используем соотношение (14.16) справедливое для начальных столбцов структуры ответвившихся систем. Подобные выводы делаются, например, в геологии, когда земная кора рассматривается как продукт процесса охлаждения, или в астрономии, когда считают, что вращение Луны замедляется вследствие наличия некоторого внутреннего трения, называемого «приливным», которое обнаруживается в движении Луны вокруг Земли. По существу, наше знание прошлого основывается на существовании множества ответвившихся систем, в каждой из которых энтропия прошла стадию начального подъема своей кривой или еще соответствует этому состоянию. В § 14 было показано, что, в противоположность выводам, основанным на вероятностях по одной системе, выводы от времени к энтропии, основывающиеся на вероятностях по множеству систем, не симметричны относительно прошлого и будущего. Это может привести к еще более низким значениям энтропии в прошлом.

Мы снова видим, что, для того чтобы объяснить методы умозаключений, используемые физиками, требуются вероятности по многим системам. Кинетическая теория не смогла бы информировать нас о прошлом нашей вселенной, если бы она ограничивалась вероятностями

по одной системе, относящимися к изолированным процессам. Для получения достаточного знания о прошлом детерминистские методы оказываются слишком ограниченными в своих возможностях. Например, на основе детерминистских методов, используемых физиком, нельзя было бы сделать вывод об эволюции видов, что представляет одно из величайших открытий в биологии. Это основывается исключительно на вероятностях, относящихся к ответвившимся системам.

Возвращаясь к проблеме энтропии вселенной и используя рассмотренные методы, мы заключаем, что в том направлении времени, которое мы называем будущим, энтропия вселенной возрастает и что она имела меньшее значение в том направлении, которое мы называем прошлым; другими словами, что мы живем в такой области вселенной, которая характеризуется длительным периодом возрастания энтропии. Однако возрастание энтропии всей вселенной как таковое не дает основания тому, чтобы называть соответствующее направление времени положительным. В нашем повседневном опыте нас мало беспокоит энтропия вселенной. Направление времени для нас выражается в направлениях процессов, происходящих в ответвившихся системах, которыми в изобилии снабжает нас наше окружение. Все эти процессы идут в одном и том же направлении, а именно в направлении возрастания энтропии. Мы знаем, что этот факт тесно связан с увеличением общей энтропии вселенной, и, следовательно, благодаря его повторению в ответвившихся системах рост энтропии вселенной определяет для нас направление времени. В поведении ответвившихся систем, так сказать, отражается общее увеличение энтропии мира, и только это отражение общей тенденции во множестве индивидуальных проявлений доступно нашему наблюдению и представляется нам как направление времени.

Наличие длительного периода увеличения энтропии не является, следовательно, *достаточным* условием для решения вопроса о направлении времени, хотя и представляет *необходимое* условие. Направление времени для нас выявляется только потому, что участки подъема характеризуются множеством состояний, в которых происходит ответвление подсистем, раскрывающих в своем дальнейшем развитии всеобщий характер возрастания

энтропии. Это возрастание энтропии не стало бы очевидным, если бы первоначально энтропия подсистемы не обладала низким значением в сравнении с другими возможными состояниями подсистемы, пока она остается изолированной. Наша вселенная удовлетворяет этому условию. Следовательно, она обязана своим направлением времени ветвящейся структуре того участка подъема кривой своей энтропии, на котором мы живем.

Мне хотелось бы надеяться, что подобные рассуждения выражают взгляды Больцмана. Он говорил о «небольших, изолированных от вселенной областях», которые «сначала» всегда находятся в маловероятном состоянии[1]. Но философские суждения Больцмана слишком лаконичны, и это делает невозможным их однозначное толкование. По-видимому, он не сознавал, что предположение о ветвящейся структуре, как описано в этом параграфе, выходит за рамки предположения о том, что наша вселенная в настоящее время находится на наклонной части своей энтропийной кривой. Больцман пишет: «То, что в природе переход от вероятного к маловероятному состоянию происходит не так же часто, как обратный переход, можно было бы удовлетворительно объяснить, предположив, что вся окружающая нас вселенная находилась в очень маловероятном состоянии, в результате которого любая система вступающих во взаимодействие тел также первоначально находится, вообще говоря, в маловероятном состоянии»[2]. Этот аргумент говорит в пользу наличия описанной выше ветвящейся структуры. Но нужно исследовать, какие предположения скрыты за этим правдоподобным выводом. Этот вопрос будет рассмотрен в § 16.

Кроме того, хотя процитированный отрывок, несомненно, относится к частотам в ответвившихся системах, в трудах Больцмана нет явной формулировки о вероятностях, относящихся к этому ансамблю. Насколько это касается перехода от энтропии ко времени, на парадокс обратимости нельзя ответить, исходя из вероятностей по одной системе, но только на основе использования вероятностей по многим системам, относящимся к ответвив-

---

[1] Л. Больцман, там же, стр. 526.
[2] Там же.

шимся системам. Ограничившись подсчетом горизонтальных вероятностей для участка подъема кривой, мы не можем получить для этих вероятностей неравенства, аналогичного (14.6), потому что на одном участке в изолированной системе состояние A с определенным значением энтропии встречается только однажды. Чтобы получить выражение вероятностей, подобных (14.2), через частоты, мы должны, следовательно, произвести подсчет по многим участкам и, таким образом, прийти к равенству (14.3). Так как для горизонтальных вероятностей такой секционный подсчет невозможен, то такие вероятности не могут определить направления времени для соответствующих участков [1]. Окончательный ответ на парадокс обратимости нельзя дать, только ссылаясь на отдельную изолированную систему, для этого требуется ссылка на множество систем, которые при своем возникновении не являются изолированными, так как ответвляются от общей системы.

Добавим одно замечание относительно пространственно бесконечной вселенной. Для конфигурации бесконечного числа молекул вероятность, а следовательно, и энтропию определить невозможно. Можно было бы попытаться следующим образом определить среднюю энтропию s: мы вычисляем энтропию S конфигурации некото-

---

[1] Можно было бы попытаться следующим образом представить секциональную статистику для вероятностей по одной системе. Пусть в класс Q входит множество состояний A, характеризуемых низкими значениями энтропии. Класс R может включать отчасти более высокие значения энтропии, хотя R может перекрываться с Q. Под вероятностью $P(Q^{kl}, R^{k,\, l+m})^l$ мы теперь будем понимать вероятность того, что за состоянием y_{kl}, принадлежащим Q, через m элементов следует состояние $y_{k,\, l+m}$, принадлежащее R и такое, что его энтропия выше энтропии состояния y_{kl}. Такие вероятности могут быть представлены частотами, относящимися к одному участку подъема кривой. Например, когда цемент медленно перемешивается с песком, мы видим процесс длящимся и наблюдаем, что за каждым частично упорядоченным состоянием следует до некоторой степени менее упорядоченное состояние. Но такие вероятности представляют результат усредняющего процесса; для их строгого определения необходима ссылка на вероятности, подобные (14.2), которые относятся к определенным состояниям A и которые суммируются по правилу редукции (ThP, p. 97 и 234) при учете предшествующих вероятностей A (которые зависят от способа выбора наших наблюдений внутри Q). Вероятности, подобные (14.2), секционно не могут быть определены.

рого трехмерного объема пространства V и находим удельную энтропию s из соотношения

$$s = \lim_{v \to \infty} \frac{S}{V}. \qquad (1)$$

Однако нам неизвестно, существует ли такой предел, и если существует, то он может оказаться равным нулю или некоторой константе и не зависеть от времени. Он может также колебаться нерегулярным образом. Следовательно, невозможно доказать, что s все время возрастает или следует кривой, изображенной на рис. 19.

Однако направление времени можно определить другим путем. Система, которая пространственно отделена от других систем, подобно звезде или галактике, не может считаться изолированной, потому что она теряет теплоту благодаря излучению в пространство. Если пренебрегать этим излучением, то энтропия системы будет убывать вследствие охлаждения системы. Энтропия системы будет возрастать только в случае, если мы учтем рассеянное в пространстве излучение. Поэтому можно было бы сказать: эволюция любой системы, как можно показать, представляет возрастание энтропии, если эта система включена в более обширную систему, которая достаточно велика [1]. Если этот принцип верен, то для бесконечной вселенной было бы однозначным образом определено направление времени, хотя мы и не могли бы говорить об энтропии вселенной.

При современном состоянии космологии весьма трудно прийти к заключению относительно времени в целом. Более того, нет необходимости предполагать, что время повсеместно имеет одно направление, для того чтобы объяснить направление времени в том участке мира, который доступен нашему наблюдению. В этом состоит один из важнейших результатов исследований Больцмана: вопрос о направлении времени в целом должен быть отделен от вопроса о направлении времени, доступного нашему наблюдению. Даже если упорядоченная структура времени не является открытой, что обычно предполагается, то есть даже если время в целом замкнуто,

---

[1] Такая формулировка была дана Максом Планком в его „Vorlesungen über die Thermodynamik", Leipzig, 1911, S. 104.

о чем упомянуто в § 5, еще остается возможность того, что отдельные секции времени обладают направлениями, которые статистически выражаются точно таким же образом, как и для незамкнутой во времени вселенной.

Своим статистическим определением направления времени, которое ограничено рамками одной секции из всего времени вселенной, Больцман указал путь разрешения парадокса статистического направления — проблемы согласования однонаправленности макровремени с обратимостью микропроцессов. В этом состоит великий вклад Больцмана в физику и философию. И значение этого результата не уменьшается от того, что его рассуждения основывались на механической модели атома, которая в настоящее время не является адекватной. Наоборот, мы увидим, что современная квантовая физика нуждается в идеях Больцмана не в меньшей мере, чем физика, основанная на механике Ньютона. Причина этого та же самая — современная физика тоже не раскрыла необратимости в своих элементарных процессах. Основным результатом всех исследований природы времени, по-видимому, является установление статистического характера направления времени. Этот факт не умаляет того значения, которое имеет для наших представлений о времени направленность зависящих от времени соотношений. Наоборот, в дальнейшем увидим, что наши повседневные опыты оказывают поддержку статистическому истолкованию направления времени, что эти опыты, так сказать, непосредственно говорят нам о том, к чему пришел Больцман абстрактным путем при изучении термодинамических процессов. В своем статистическом анализе Больцман раскрыл логическое ядро понятия времени.

Но если исследование направления времени относится к компетенции теории вероятностей и статистики, необходима ссылка на ансамбль ответвившихся систем. История одной изолированной системы, взятая в целом, не выражает направления времени. И даже отрезок этой истории, соответствующий подъему кривой энтропии, не позволяет разрешить парадокс обратимости, если структура этого отрезка не подобна структуре, образуемой системами, ответвляющимися от основной. Неоднократно выдвигалась идея ответить на парадокс обратимости путем ссылки на системы, начальные состояния которых с

низкой энтропией вызываются внешними причинами. Однако такие попытки обычно не рассматривались как плодотворные вследствие широко распространенного мнения, что вполне возможно определить направление времени на основе использования вероятностей по одной системе, то есть для изолированных систем. Ныне достаточно ясно, что этой цели невозможно достигнуть и что она накладывает неоправданное ограничение.

При статистическом определении направления времени предполагается множество систем, которые при своем возникновении не являются изолированными, но приобретают свои начальные крайне маловероятные состояния в результате взаимодействия с другими системами и после этого остаются изолированными в течение определенного времени. Наличие направления времени в развитии нашей вселенной, которая образует изолированную систему, обусловлено не просто возрастанием ее общей кривой энтропии, но тем фактом, что это возрастание включает множество ответвляющихся систем, подобных описанным выше. Направление времени определяется направлением энтропии, поскольку последнее проявляется в статистическом поведении большого числа отдельных систем, возникающих индивидуально, в общем движении ко все более и более вероятным состояниям.

§ 16. Гипотеза ветвящейся структуры

Теперь рассмотрим, из каких предположений исходят, когда утверждают, что существуют описанные выше ответвившиеся системы. Такие предположения можно сформулировать тремя способами: во-первых, на языке L_1, в котором употребляется обычное время, положительное направление этого времени определено направлением

---

[1] Этот взгляд излагается, например, Ван-дер-Ваальсом младшим в „Über die Erklärung der Naturgesetze auf statistisch — mechanischer Grundlage", в „Phys. Zeitschr.", Bd. 12, 1911, S. 547—549, и Герцем в его статье „Statistische Mechanik", в „Ergebnisse der exakten Naturwissenschaften", Bd. I, Berlin, 1922, S. 76. Подобные соображения были рассмотрены: Hans Reichenbach, Ziele und Wege der physikalischen Erkenntnis, в „Handbuch der Physik", ed. by Hans Geiger und Karl Scheel, Bd. IV, Berlin, Julius Springer, 1929, S. 63. Однако в этих публикациях еще не использовалась символика, основанная на вероятностной структуре.

возрастания энтропии основной системы; во-вторых, на языке L_2, в котором направление положительного времени, по определению, противоположно направлению положительного времени в L_1, и, в-третьих, на нейтральном языке L, который не отражает направления времени.

Используя язык L_1, мы рассматриваем участок *1—6* кривой на рис. 21 как идущий слева направо, то есть как участок подъема, и, таким образом, мы проводим различие между точками *ответвления*, такими, как *1, 3* и *5*, и точками *возврата 2, 4* и *6*. Точки ответвления являются точками с относительно низкой энтропией, то есть точками, в которых энтропия подсистемы мала. Назовем их *нижними точками*. Точки возврата являются точками с относительно высокой энтропией; здесь энтропия подсистемы имеет высокое значение, но энтропия основной системы еще мала, хотя и несколько выше энтропии соответствующей точки ответвления. Мы будем говорить об этих точках возврата как о *верхних точках*.

На языке L_2 эта терминология изменится следующим образом. Точки ответвления L_1 станут точками возврата, а точки возврата L_1 — точками ответвления. Если иметь в виду участок *6—1* на рис. 21, то тогда мы назовем точки *6, 4* и *2* точками ответвления, а точки *5, 3* и *1* будут называться точками возврата. Это означает, что в L_2 мы рассматриваем участок *6—1* как идущий справа налево и, следовательно, как участок спуска. В отношении участка *7—12*, или *12—7*, мы получим обратные результаты. Этот участок для L_1 идет справа налево, а для L_2 — слева направо. Следовательно, язык L_i определяется только в данной секции, и мы не распространяем его за рамки его секции.

Однако точки, которые мы называем нижними и верхними точками, в обоих языках являются теми же самыми. Например, точки *1, 3* и *5* являются нижними точками и для L_1, и для L_2; а точки *2, 4* и *6* — верхними точками в обоих языках. Следовательно, эти термины принадлежат *нейтральному* языку L. На языке L мы не говорим о времени, но только о значениях энтропии; и здесь мы имеем фразы, подобные «низкая энтропия» и «направление увеличения энтропии». В L можно также использовать выражение «ответвившаяся система», когда не касаются различия между точками ответвления и возврата. В этом случае мы просто говорим о двух границах, или

конечных пунктах, ответвившейся системы и различаем их как *нижнюю* и *верхнюю границы*.

Если мы хотим сформулировать предположения, из которых исходят при допущении ветвящейся структуры кривой энтропии, то неправомерно пользоваться языком, в котором определено направление времени. Когда Больцман писал, что вступающая во взаимодействие система тел находится, в общем, в весьма маловероятном состоянии, его утверждение кажется нам правдоподобным; но во фразе «вступающая во взаимодействие» допускается направление времени языка L_1, и нам необходимо узнать, не выводится ли правдоподобие данного аргумента из молчаливо принятого допущения о времени.

По этой причине мы на нейтральном языке L сформулируем предположения относительно ветвящейся структуры, на которых основано статистическое определение направления времени. Эти предположения можно объединить под названием *гипотезы о ветвящейся структуре*, и они формулируются следующим образом:

Предположение 1. Энтропия вселенной в настоящее время низка и соответствует участку подъема кривой энтропии.

Предположение 2. Существует множество ответвившихся систем, которые изолированы от основной системы в течение определенного периода, но которые связаны с основной системой на обеих своих границах.

Предположение 3. Структура ответвившихся систем представляет структуру процесса смешивания, как указано в (12.11) — (12.13), (12.16) и (12.18).

Предположение 4. В громадном большинстве ответвившихся систем одна граница ее энтропийной кривой представляет *нижнюю точку*, *другая* — верхнюю точку.

Предположение 5. В громадном большинстве ответвившихся систем направления возрастания энтропии параллельны друг другу и направлению основной системы.

Эти пять предположений могут рассматриваться как эмпирические гипотезы, тщательно проверенные. Проверка предположения 1 была рассмотрена в § 15. Единственное возражение, которое можно сделать в отношении предположения 1, заключается в том, что мы не уверены, можно ли определить содержательным образом энтропию вселенной. Это возражение анализировалось в

§ 15. Если такая космическая энтропия окажется неприемлемой, предположение 1 нужно будет переформулировать, как указано в § 15, например, следующим образом: каждая система представляет часть более общей системы, которая приближенно может рассматриваться как изолированная и которая обладает сравнительно низкой энтропией. Предположения 2—5 останутся без изменений, лишь в предположении 5 нужно опустить фразу «и направлению основной системы».

Больцман, по-видимому, полагал, что предположения 2—5 можно вывести из предположения 1. Но очевидно, что это не верно. В частности, предположение 5 не может быть выведено из предположения 1. Если градиент энтропии некоторой ответвившейся системы или даже всех их направлен прямо противоположно градиенту энтропии основной системы, то такая конфигурация не противоречит законам механики и, следовательно, вполне допустима с точки зрения причинного определения порядка времени. Это определение говорит нам, например, что на рис. 21 направление времени от точки *1* к точке *2* одно и то же как для основной, так и для ответвившейся системы: для обеих систем точка *1* расположена или раньше, чем *2*, или позднее. Но отсюда не следует, что энтропия должна изменяться одинаковым образом вдоль обеих линий. Фактически для весьма малых систем, таких, как системы, представленные во флюктуационных явлениях, действительно происходят противоположно направленные изменения. Другими словами, утверждение «определение направления времени через энтропию совместимо с причинным определением порядка времени» представляет собой эмпирическое утверждение. Предположение 5 можно назвать *принципом параллелизма в возрастании энтропии.*

Аргумент Больцмана можно истолковать в том смысле, что низкая энтропия вселенной делает вероятным существование небольших систем с относительно низкой энтропией. Это предполагает существование некоторого перемешивающего процесса, и утверждается, что неоднородности равномерно распределены во всех небольших частях вселенной. Но это соображение приводит просто к предположению 4, которое, даже если оно истинно, не означает, что нижняя граница ответвившейся системы расположена таким образом, что возрастание энтропии

подсистемы параллельно возрастанию энтропии основной системы. Нижняя точка может быть расположена на другом конце ветви. Когда мы исключаем эту возможность, добавляя предположение 5, мы вводим особую гипотезу. Именно эта гипотеза о параллелизме возрастания энтропии, скрытая во фразе Больцмана «вступающие во взаимодействие», делает нижнюю точку началом взаимодействия между частями изолированной системы. Но только благодаря предположению 5 мы вправе считать нижнюю точку ветви *началом* взаимодействия, при условии, что слово «начало» относится к одному и тому же направлению времени для всех систем, включая основную систему и все подсистемы. Если мы не желаем ссылаться на направление времени, то мы нижнюю точку назовем просто *точкой взаимодействия,* оставляя открытым вопрос о том, начинается ли здесь развитие системы как изолированной или оно заканчивается. Верхняя точка не выражает взаимодействия, так как это точка равновесия. Кроме того, этот термин может относиться к взаимодействию между ответившейся системой и ее окружением, и в этом случае также предполагается бо́льшая интенсивность взаимодействия только в нижней точке.

Может быть, и возможно показать, что в развитии такой системы, как вселенная, наклонные участки кривой энтропии, удовлетворяющие предположениям 2—5, встречаются значительно более часто, чем наклонные участки, не удовлетворяющие этим предположениям. Следовательно, мы имели бы высокую вероятность (по одной системе) того, что если предположение 1 верно, то справедливы и предположения 2—5. Но эта вероятность не окажет нам существенной помощи, так как она относится к частотам, которые реализуются в *истории вселенной* в течение весьма продолжительных периодов времени. А предположение о том, что мы действительно находимся на соответствующем наклонном участке кривой энтропии, нуждалось бы еще в эмпирическом доказательстве.

Однако можно дать другое доказательство. Если принимается предположение 4, то предположение 5 можно вывести из предположения 3, то есть из условий, выполняющихся для структуры смешивания. Следовательно, становится возможным опустить предположение 5; гипотеза о ветвящейся структуре полностью формулируется предположениями 1—4.

Прежде всего поясним этот результат неформальным соображением. Вероятностную структуру ответвившихся систем мы записываем для направления времени, употребляемого L_1, то есть начинаем с характеристики тех граничных точек, которые в L_1 являются точками ответвления. Термин «точка ответвления» вполне определен, поскольку *порядок* времени каждой ответвившейся системы определяется через причинные взаимоотношения, независимо от направления возрастания ее энтропии. Допустим теперь, что энтропия около половины ответвившихся систем возрастает в противоположном направлении, то есть, что их точки ответвления в L_1 будут верхними точками, в то время как их точки возврата, также в L_1, будут нижними точками. Первый столбец структуры (14.1) будет тогда иметь равное число состояний с низкой и высокой энтропией. Соотношение (14.5), которое выражает инвариантность структуры, будет тогда еще оставаться верным в том случае, если A есть состояние с низкой энтропией, а B — состояние с высокой энтропией; но это соотношение окажется несправедливым, если A — состояние с высокой энтропией, а B — с низкой энтропией, по крайней мере для последующих столбцов. Причина этого состоит в том, что для данного последнего случая вероятность в правой части (14.5) весьма низка, в то время как в структуре, первый столбец которой содержит равное число состояний с низкой и высокой энтропией, соответствующие вертикальные вероятности, то есть выражение в левой части (14.5), были бы высоки, потому что практически все системы, начинающиеся (в L_1) с состояний с высокой энтропией, в этом случае будут развиваться в направлении состояний с низкой энтропией.

Этим соображениям можно придать форму строгого вывода предположения 5 из предположений 3 и 4. В § 11 было показано, что вне зависимости от распределения в первом столбце более поздние столбцы вероятностной структуры процесса смешивания принимают распределения, которые сходятся к распределению, существующему в строках. Так как состояние A с низкой энтропией обладает в строке низкой вероятностью, оно должно встречаться с соответствующей низкой частотой в последующих столбцах. Отсюда следует, что только начальные столбцы структуры могут обладать большим числом со-

стояний с низкой энтропией; для последующих столбцов это невозможно. Если ответившиеся системы можно разделить на такие две группы, что системы в первой группе вначале имеют состояния с низкой энтропией, в то время как системы во второй группе имеют конечные состояния с низкой энтропией, то тогда невозможно эти две группы объединить в структуре процесса смешивания, если число систем в одной из них практически не приближается к нулю. Так как предположение 4 говорит о том, что большое число систем имеет в качестве одного из граничных состояние с низкой энтропией, мы заключаем, что практически все эти граничные состояния должны находиться на одной стороне; в противном случае результирующая вероятностная структура не могла бы быть структурой процесса смешивания, удовлетворяющей условию инвариантности структуры. Таким образом, мы получили вывод предположения 5 из предположений 3 и 4.

В процессе вывода мы предполагали, что строки конечной длины, сопоставляемые всему ходу изменений ответвившихся систем, достаточно длинны и реализуют предельные условия, которые теоретически устанавливаются для бесконечных последовательностей, и подобным образом предполагали, что достаточно длинны и столбцы, то есть, что мы имеем достаточно большое число ответвившихся систем. Но таких предположений нельзя избежать всякий раз, когда статистическая теория применяется к физическим событиям. Получаемое заключение не представляет строгого закона для данного конечного числа составляющих вследствие этих предположений о численном соответствии; оно допускает возможные исключения. Чтобы пояснить эту особенность, предположим, что мы наблюдаем исключительный случай. Пусть среди многих галактик существует одна такая, внутри которой время протекает в направлении, обратном направлению времени нашей галактики. Тогда мы имели бы ситуацию, рассмотренную Больцманом[1]. В этом случае некоторая удаленная часть вселенной находилась бы на таком участке кривой энтропии, который по отношению к нашему участку характеризуется уменьшением энтропии; однако если бы в этой части вселен-

[1] См. § 15, прим. на стр. 177.

ной были бы живые существа, то для них их окружение обладало бы всеми свойствами, характеризующими участок подъема.

Развитие такой системы в противоположном направлении времени можно было бы обнаружить по дошедшему от системы до нас излучению и, возможно, по наличию смещения спектральных линий. Вследствие громадного расстояния это сообщение пришло бы к нам так поздно, что оно просто информировало бы нас о том направлении времени, которое было присуще системе целую вечность тому назад. Таким образом, мы не сможем получить сведения о системе, направление времени которой в данный момент противоположно нашему, и такое ограничение не было известно Больцману. Это справедливо по крайней мере в тех случаях, когда пользуются обычным определением одновременности, согласно которому свет распространяется с одинаковой скоростью во всех направлениях. Если отказаться от этого определения, то нашу вечность можно было бы преобразовать в весьма короткий период времени. Неоднозначность в определении одновременности приводит к значительной неоднозначности при сравнении времени между двумя системами (см. 5.3). Кроме того, излучение, идущее от этой системы в нашу систему, распространялось бы с точки зрения этой системы в обратном направлении времени, то есть не покидало бы эту систему, а прибывало бы в нее. Возможно, этот сигнал мог бы быть истолкован обитателями этой системы как сообщение из нашей системы, говорящее им о том, что наша система развивается в обратном направлении времени. Здесь мы имеем связывающий световой луч, который с точки зрения одной системы прибывает на нее и исчезает в некотором абсорбционном процессе; а с точки зрения другой системы он излучается источником при необратимом процессе, идущем в противоположном направлении.

Такие возможные исключения из принципа параллелизма энтропии не влияют на данный вывод предположения 5 из предположений 3 и 4, потому что предположения 3 и 4 относятся к большому числу систем. Отдельные исключения не опровергают статистических гипотез. Можно добавить, что до настоящего времени не обнаружено никаких астрономических данных, которые указывали бы на существование таких исключений.

Следует отметить, что приведенный вывод предположения 5 не вынуждает нас использовать язык L_1. Используя L_2, мы можем построить вероятностную структуру, удовлетворяющую предположению 5, если первый столбец составим из всех верхних точек ответвившихся систем. Системы стали бы тогда развиваться по направлению к состояниям с низкой энтропией. Для этой структуры условие инвариантности удовлетворялось бы, так сказать, в обратном; мы просто заменяем $+m$ на $-m$ в левой части выражения (14.5). Но для L_2 также имел бы силу тот физический факт, что энтропия всех ответвившихся систем изменяется в одном и том же направлении, так как он не зависит от выбора языка. Действительно, построенная таким образом структура, если ее рассматривать справа налево, совпадает со структурой, построенной для L_1. С точки зрения логики наше предпочтение использовать L_1 и тем самым отождествлять направление положительного времени с направлением возрастания энтропии является делом определения. Позднее мы укажем психологические основания этого определения.

Так как условие инвариантности структуры выражает вывод о пространственном ансамбле на основе временно́го, то отсюда следует, если этот вывод справедлив, что практически все ответвившиеся системы должны проявлять возрастание энтропии в одном и том же направлении. Мир, в котором процессы смешивания управляются соотношениями, установленными для вероятностной структуры процесса смешивания и представленные формулами (12.11) — (12.13), (12.16) и (12.18), проявляет определенное единообразие; мы назовем его *статистической изотропностью*. Этот термин означает, что распределение временно́го ансамбля повторяется в пространственном ансамбле всякий раз, когда процесс смешивания происходит в течение определенного времени, соответственно формуле (12.10). Направление времени и пространственное направление проявляют здесь определенную эквивалентность, которую, по аналогии с используемой в механике и оптике терминологии, можно назвать *изотропностью*. Таким образом, мы находим, что ответвившиеся системы определяют однозначным образом направление времени, потому что мир статистически изотропен.

Таким образом, хотя мы и вывели предположение 5 из предположений 3 и 4, мы не должны забывать, что

оба предположения 3 и 4 выражают физические свойства мира и никоим образом не представляют логической необходимости. В отношении предположения 4 это совершенно очевидно, но это также справедливо и по отношению к предположению 3. Условие инвариантности структуры нельзя вывести из аксиом теории вероятностей, оно представляет самостоятельное дополнение, которое требует эмпирической проверки (см. § 12). Все имеющиеся доказательства, конечно, вполне определенно подтверждают предположение 3. Поэтому можно сказать, что мир своим направлением времени обязан определенному единообразию, проявляемому всеми его процессами смешивания, — своей статистической изотропности. Направление времени отражает общую статистическую тенденцию — тенденцию описывать временной ансамбль через пространственный. Этому подчиняются как ансамбли ответвившихся систем, так и любой ансамбль молекул, который образует отдельную ответвившуюся систему. Параллелизм в возрастании энтропии выражает общий статистический закон, имеющий силу для нашей вселенной, и этому же самому закону подчиняются процессы смешивания молекул в газах.

Физический закон, выраженный в статистической изотропности вселенной и, следовательно, в параллелизме возрастания энтропии, в логическом отношении отличен от причинного закона. Его можно назвать *законом поперечного сечения* вселенной, то есть законом, проявляющимся в структуре моментального сечения, для которого t постоянно, и проходящего через вселенную, или, скорее, законом, проявляющимся при сравнении структур нескольких таких поперечных сечений. Подобные законы играют важную роль в физике, хотя число их, по-видимому, мало [1].

Параллелизм ответвившихся систем можно было бы попытаться объяснить как причинно обусловленный тем, что они не полностью обособлены от вселенной; в таком

---

[1] Другим законом этого вида является соотношение неопределенностей Гейзенберга, которое вытекает из того закона поперечного сечения, что все данные наблюдения можно объединить посредством ψ-функции; см. Hans Reichenbach, Philosophic Foundations of Quantum Mechanics, Berkeley and Los Angeles, Calif., Univ. Calif. Press, 1944, p. 4. При дальнейших ссылках эта книга будет обозначаться как PhF.

случае небольшая оставшаяся связь со вселенной рассматривалась бы как передатчик причинного влияния, который приводит все ответвившиеся системы в согласование с основной системой. Но такая интерпретация не была бы неуязвимой, поскольку статистическая изотропность имеет место и при полной изоляции систем друг от друга.

Чтобы показать это, предположим, что каждая горизонтальная линия вероятностной структуры (14.1) представляет развитие полностью изолированной ответвившейся системы, совершающееся от бесконечного отрицательного времени к бесконечному положительному времени. Тогда структура была бы бесконечной как вправо, так и влево. Предположим далее, что каждую линию можно сместить в горизонтальном направлении и что мы располагаем линии таким образом, что один вертикальный столбец будет содержать только состояния с низкой энтропией. Тогда этот столбец разделит вероятностную структуру на две половины, каждая из которых проявляла бы, согласно нашим физическим представлениям, свойства структуры процесса смешивания и, таким образом, статистическую изотропность. Для проверки этого результата нам следует произвести вычисления в правой половине, идя вправо, а в левой — идя влево. Структура, таким образом, будет симметричной относительно направления, подобно кривой энтропии замкнутой системы. Но если отбросить одну половину этой структуры, то оставшаяся структура в своих начальных столбцах обнаружила бы направление. Тот факт, что ответвившиеся системы в своих нижних точках связаны с основной системой и не существуют, согласно языку L_1, до этой точки, представляет физическую аналогию с математической операцией обрезания структуры. Заметим, что соответствующие половины горизонтальных линий определены упорядоченным образом, но не предполагают направления, то есть мы не можем изменить на обратное направление одной из линий вероятностной структуры, оставляя другие без изменений, потому что эта перемена направления нарушила бы упорядоченные отношения.

Таким образом, статистическая изотропность представляет закон, управляющий причинно независимыми системами; он выражает тот факт, что такие системы в

своем пространственном распределении удовлетворяют условиям неупорядоченности. Если бы мы бросили одновременно 10 000 игральных костей, то мы бы обнаружили, что выпадение каждой из шести граней произойдет примерно в равном числе случаев. Это выполнение условия изотропности не является следствием слабой связи событий при молекулярных столкновениях со своим окружением. Хотя совершенно независимых событий мы никогда не сможем наблюдать, такой вывод можно получить как предельное утверждение, устанавливаемое из ряда экспериментальных конфигураций, для которых непрерывно совершенствуется причинная изоляция. А случайное распределение не представляет просто определения причинной изоляции, поскольку изоляцию можно проверить прямыми методами наблюдения.

Когда мы называем статистическую изотропность законом поперечного сечения, то этим мы хотим выразить отсутствие причинной связи. Поэтому закон поперечного сечения следует отличать от причинных соотношений внутри поперечного сечения, которые образуются благодаря обычным причинам и изучением которых мы займемся в § 19. Например, если рассмотрим поперечное сечение населения США, то обнаружим, что преобладающее большинство людей разговаривает на английском языке; это причинное соотношение, а не закон поперечного сечения. То, что наша вселенная, в дополнение к причинным законам, управляется определенными законами поперечного сечения, представляет собой реальный факт, и мы должны правильно истолковывать этот факт и не считать априорной необходимостью то, что независимые системы подчиняются законам случая. Все, что мы можем сделать, — это сформулировать существующий закон и проверить его наблюдениями.

Подытожим рассмотренные соображения. Статистическое определение направления времени основывается на гипотезе о ветвящейся структуре, сформулированной в предположениях 1—5. И утверждение о том, что вселенная в настоящее время находится на участке подъема своей энтропийной кривой, представляет просто *одно* из указанных предположений; и это необходимо уточнить через некоторые другие предположения, относящиеся к существованию ответвившихся систем. К таким предположениям относится предположение, выражающее на-

раллелизм в возрастании энтропии (предположение 5), но это предположение выводится из четырех других предположений и, таким образом, представляет собой логическое следствие общего свойства вселенной, называемого статистической изотропностью и сформулированного в предположении 3. Поэтому система предположений 1—5 эквивалентна системе предположений 1—4. Эту систему предположений — гипотезу о ветвящейся структуре — нельзя вывести ни из аксиом теории вероятностей, ни из эргодической гипотезы (рассмотренной в § 10). Последняя просто дает возможность получить метрику вероятностей вселенной. Чтобы вывести свойства направленности времени, нам необходимо включить в рассмотрение предположения *о фактах* 1—5. Современное состояние нашей вселенной представлено некоторым участком кривой ее энтропии, удовлетворяющим предположениям 1—5; это тот факт, которому мы обязаны существованием направления времени.

Логическую функцию предположений 1—5 в вероятностной теории времени можно выразить следующим образом. Среди ансамбля систем в нашем окружении значительное число изолированных систем обладает такими значениями энтропии и находится в таких переходных состояниях, которые в эволюции отдельной системы соответствуют малым вероятностям. Каким образом можно объяснить, относящуюся к множеству систем высокую частоту появления таких упорядоченных изолированных состояний, если учитывать их малую вероятность по отношению к одной системе? Как можно объяснить тот факт, что таким состояниям часто предшествуют состояния с еще меньшей энтропией? Ответы на эти вопросы содержатся в предположениях 1—5. Выводы о частоте появления упорядоченных изолированных состояний мы получаем, исходя из частоты появления точек взаимодействия, о которых в свою очередь сказано в предположениях 1—5. Объяснить — это значит логически вывести из фундаментальных свойств физического мира; следовательно, взаимодействие объясняет порядок, а не наоборот.

Эти логические взаимосвязи чрезвычайно важны для понимания проблемы, которую необходимо теперь исследовать: различие между причиной и следствием. Исследование этого приводит нас к новой главе.

Глава IV

НАПРАВЛЕНИЕ ВРЕМЕНИ В МАКРОСТАТИСТИКЕ

§ 17. Макроконфигурации и макроэнтропия

Теперь мы применим наши статистические соображения к иным процессам, нежели процессы смешивания молекул, а именно к процессам, элементарными «частицами» которых являются макроскопические объекты (такие, как крупинки песка или игральные карты), а элементарными конфигурациями — макросостояния, а не микросостояния. Для этих процессов также можно дать определение классов конфигураций; такой *класс конфигураций* можно отличать от самих *конфигураций*, обозначив его как *состояние*. При подсчете числа конфигураций в пределах такого класса к этим состояниям можно применять предыдущие результаты относительно вероятностей, и тогда вероятность состояния можно характеризовать функцией, которая аналогична энтропии и имеет малые значения для упорядоченных состояний и большие — для неупорядоченных. В этом случае мы можем говорить о статистике *макроконфигураций*, а также применять термины *макровероятность* и *макроэнтропия*. Так мы приходим к статистическим взаимосвязям, которые относятся исключительно к макрокосму и которые можно назвать *макростатистикой*, в противоположность *микростатистике* кинетической теории материи. В этой статистике существует направление, аналогичное направлению термодинамических процессов.

Начнем с определения понятия порядка в макростатистике. Выше было показано (§ 8), что в микростатистике понятие порядка допускает сведение к понятию *одинакового порядка*, которое в свою очередь можно определить через понятие макроскопически неразличимого. К макроконфигурациям последнее из упомянутых понятий нельзя применить, поскольку все такие конфигурации можно различить. Например, при тасовке колоды

карт можно отличить каждую их конфигурацию, образующуюся в некоторый момент, от любой другой конфигурации; при желании мы можем опознать ее, записав по порядку названия карт. В каком смысле тогда можно говорить об одинаковом порядке макроконфигураций?

Ссылаясь на пример с колодой карт, можно обратить внимание на тот факт, что мы проводим различие между упорядоченными и неупорядоченными конфигурациями. Если все карты красной масти находятся сверху, а все карты черной масти — под ними, то мы говорим об упорядоченной конфигурации, тогда как ту конфигурацию, которая обычно получается после тасовки карт, мы рассматриваем как неупорядоченную. Основанием для этого служит то, что первую конфигурацию можно охарактеризовать с помощью простого правила, тогда как в отношении последней этого сделать нельзя.

Было бы ошибочным думать, что неупорядоченная конфигурация не удовлетворяет никаким правилам. Так как число карт ограничено, мы всегда можем найти арифметические правила, соответствующие данной конфигурации. Например, будем обозначать карту красной масти цифрой 1, а карту черной масти — цифрой 2; получающиеся серии чисел будем рассматривать как дробную часть десятичной дроби. Тогда мы получим десятичную дробь с 52 десятичными знаками, которая легко превращается в дробное отношение двух целых чисел $\frac{m}{n}$. Следовательно, конфигурация карт описывается правилом, соответствующим десятичной записи дроби $\frac{m}{n}$. Но это не простое правило; вычисление дроби $\frac{m}{n}$ требует выполнения соответствующих логических операций, что, конечно, не может быть сделано без пера и бумаги.

Правило можно назвать простым, если с первого взгляда можно определить, удовлетворяет ли ему данная конфигурация. Правила: «Все карты красной масти лежат сверху» и «Каждая вторая карта — карта красной масти» — относятся к этому виду правил. Конечно, это весьма не строгое определение термина «простой»; и если правило не очень простое и в то же время не очень сложное, то можно разойтись во мнении, надо ли его включать в число простых правил. Однако для многих практических целей этого определения вполне достаточно.

Оно относится к числу тех определений, которые вполне удовлетворительно характеризуют особенности функционирования языка и не претендует на точность в той области, где она неуместна. Это делает фразу «простое правило» психологическим термином. Следовательно, когда мы говорим о конфигурациях, которые можно описать посредством простых правил, и о конфигурациях, которые нельзя описать таким образом, то мы применяем классификацию по реакциям наблюдателей.

При желании мы можем даже подразделить наблюдателей на различные категории. При игре в бридж такие правила, как «карты, дающие шансы на выигрыш» или «хороший ход», знаток будет применять иначе, чем непосвященный. Торговец зерном классифицирует образцы зерна по качеству на группы, которые неразличимы для неопытного глаза. Следовательно, можно говорить о простых правилах по отношению к определенной категории наблюдателей. Однако для общих статистических соображений не имеет большого значения, какой вид наблюдателей мы имеем в виду.

Назовем две конфигурации, a_1 и a_2, *неразличимыми по правилу*, если каждому простому правилу, которому удовлетворяет a_1, также удовлетворяет и a_2, и наоборот. В результате тасовки колоды карт обычно получаются конфигурации, которые неразличимы по правилу в следующем смысле: они удовлетворяют только одному простому правилу — «являются конфигурациями одной колоды карт» — и никакому другому простому правилу. Напротив, конфигурация, удовлетворяющая определению: «Все карты красной масти лежат сверху», отличима по правилу от случайных конфигураций, потому что она удовлетворяет этому простому правилу, помимо общего правила, справедливого для всех конфигураций. Число конфигураций, удовлетворяющих этому специфическому правилу, остается еще большим, поскольку можно тасовать по отдельности карты красной и черной масти. Сюда же входят конфигурации, удовлетворяющие более ограниченным правилам, таким, как правило: «Вначале все бубны, затем все черви, затем все пики, затем все трефи». Однако об этих конфигурациях можно сказать, что они удовлетворяют правилу: «Карты красной масти лежат сверху» — и не удовлетворяют другому правилу; эти конфигурации неразличимы по правилу друг от друга.

О двух конфигурациях, которые неразличимы по правилу, говорят, что они имеют *одинаковый порядок*. Таким образом, для данной конфигурации можно определить класс с одинаковым порядком, который назовем *состоянием* конфигурации, и методы микростатистики можно перенести в нашу макростатистику. Степень порядка конфигурации определяется расширением соответствующего класса с одинаковым порядком (ср. § 9).

Это определение содержит в себе предположение, что все конфигурации равновероятны. Для перемешивающихся при тасовке карт это, несомненно, справедливо. В других случаях это предположение о метрике вероятностей необходимо эмпирически проверить, так как для них мы не обладаем аналогом теоремы Лиувилля. Иногда может оказаться, что нам необходимо ввести другую метрику путем приписывания конфигурациям разных статистических весов. Часто оказывается достаточно довольно грубой оценки метрики.

Согласно этим определениям, тасование колоды карт представляет макромеханизм, в котором проявляются все статистические свойства микромеханизмов. Упорядоченная конфигурация характеризуется низкой макровероятностью, логарифм которой может быть назван макроэнтропией этой конфигурации. В процессе тасования карт первоначально упорядоченные конфигурации преобразуются в неупорядоченные, которые имеют высокие значения макровероятностей и отсюда высокие значения макроэнтропий. Таким образом, процесс тасования карт хорошо соответствует первому участку кривой на рис. 19. При дальнейшем тасовании кривая будет оставаться при высоких значениях, подвергаясь незначительным флюктуациям. Однако случай длительного спада крайне маловероятен. Таким образом, процесс тасования карт обладает свойствами необратимого процесса; его можно назвать *квазинеобратимым*. Он отличается от собственно необратимых процессов, потому что мы можем рассортировать перетасованные карты и, таким образом, вновь расположить в первоначальном порядке путем индивидуального выбора — действие, невозможное для микростатистики[1].

[1] Эта проблема неоднократно обсуждалась в связи с демоном Максвелла, который сортирует молекулы по скоростям, открывая нечто вроде окошка при приближении быстрой молекулы и закрывая его при приближении медленной молекулы. Было показано,

Однако весьма маловероятно, что перемешиванием мы сможем вернуть их к первоначальному порядку.

Процесс тасования карт — довольно сложный механический процесс, в котором результат каждой отдельной стадии можно предсказать лишь с некоторой вероятностью. Поэтому интересно рассмотреть такой квазинеобратимый процесс, в котором перемешивающий механизм состоял бы из простых механических процессов явно детерминистского характера.

Рис. 22. Биллиардные шары, расположенные в ряд. Если со стороны толкнуть шар так, чтобы он столкнулся с одним из шаров ряда, то это и последующие столкновения разрушат порядок в расположении шаров. В процессе движения шаров по столу восстановление этого порядка не может произойти случайным образом.

Преположим, что мы расположили биллиардные шары в ряд строго вдоль средней линии биллиардного стола (рис. 22), оставив между шарами небольшие промежутки. Затем еще один шар, расположенный вне ряда, мы толкнем в таком направлении, чтобы он ударился в один из шаров ряда под косым углом к ряду. После столкновения эти два шара ударятся о стенки и, отразившись, столкнутся с другими шарами. Если этот процесс будет продолжаться в течение нескольких минут, то упорядоченная конфигурация шаров будет нарушена и все шары в беспорядке расположатся по столу. Этот процесс, в котором порядок преобразуется в беспорядок, является квазинеобратимым, то есть его нельзя обратить.

что процесс открывания и закрывания окна связан с возрастанием энтропии и, таким образом, что даже демон, который по своим размерам не превосходит атомных, не может нарушать второй закон термодинамики. См. Лео Сциллард (Leo Szilard, Über die Entrepieninderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen, в „Zeitschr. f. Physik", Bd. 53, 1929, S. 840—856).

Конечно, если бы мы могли одновременно остановить все шары в неупорядоченном состоянии и сообщить им те же значения скоростей, но в строго противоположных направлениях, процесс пошел бы в обратном направлении и его результатом была бы первоначальная упорядоченная конфигурация. Но для нас технически невозможно это выполнить.

Этот мысленный эксперимент представляет точную аналогию процесса обращения направления скоростей молекул, рассмотренному в § 13. Неупорядоченная конфигурация, в которой мы изменяем на обратные направления скоростей, обладает *неявным порядком*, который ускользает от нашего внимания; эта конфигурация была бы не отличима по правилу от других неупорядоченных конфигураций. Этот пример показывает, что макростатистика контролируется соотношениями, подобными тем, которые справедливы для микростатистики, и может быть использована для определения направления времени точно так же, как это возможно в статистике Больцмана для молекул. Например, когда мы обнаруживаем, что биллиардные шары аккуратно расположены в ряд или что колода карт упорядочена таким образом, что все карты красной масти расположены сверху, мы заключаем, что это состояние было получено путем вмешательства. Слишком невероятно, чтобы такое упорядоченное состояние явилось бы как случайный результат в движении шаров или тасовании карт.

§ 18. Причина и следствие: порождение и протоколирование

Между микропроцессами и макропроцессами имеется существенное различие. Первые обладают естественным механизмом перемешивания в виде столкновения молекул. Вторые зачастую не обладают каким-нибудь естественным механизмом перемешивания и поэтому должны быть перемешаны искусственно, как это делается в случае с колодой карт; в других случаях естественный механизм перемешивания действует настолько медленно, что в некоторый заданный момент времени система остается практически неизменной, как в примере с песком, который медленно перемешивается в результате действия ветра. Эта характерная черта макростатистики приводит к осо-

бым следствиям: высокоупорядоченные состояния в этом случае могут легко наблюдаться и сохраняться в течение продолжительного времени. В этом заключена причина того, почему макростатистика поставляет нам так называемые *протоколы*, и почему в то же самое время она предоставляет нам ключ к пониманию *причинного объяснения*. Эту особенность теперь нужно изучить более подробно.

Предположим, что мы обнаружили на песке следы, которые несколько сглажены в результате действия ветра, но которые еще можно признать за отпечатки ног человека. Из этого «протокола» мы делаем вывод, что несколько ранее по песку прошел человек, в результате чего и появились отпечатки. Какова логическая схема этого вывода?

Различные конфигурации крупинок песка можно классифицировать по состояниям на основе использования простых правил. Одно состояние задается ровной поверхностью песка, другое — описанным условием, когда поверхность песка несет на себе отпечатки ног человека. Если мы хотим рассчитать вероятности таких состояний, то было бы неверно предполагать, что все конфигурации песчинок равновероятны. Конфигурации самого песка, подверженного только перемешивающему действию ветра, изменяются медленно; но, развевая небольшие холмы песка и занося ямки, ветер оказывает определенное воздействие, в результате чего ровная или волнообразная поверхность песка оказывается более вероятной, чем состояние, когда на песке видны отчетливые ямки. Такова метрика вероятностей, которая, как тому учат нас наблюдения, справедлива для изолированной системы песка, если под изоляцией понимать исключение всех влияний, кроме перемешивания ветром или волнами океана.

Согласно этой метрике, состояние, заданное отпечатками ног, представляет высокоупорядоченное состояние, в то время как ровная поверхность образует неупорядоченное состояние; фактически, последнее является состоянием равновесия для песка. Здесь мы имеем случай, когда высокоупорядоченное состояние сохраняется в течение некоторого времени; и мы спрашиваем, как можно объяснить наличие этого упорядоченного состояния?

Этот вопрос относится к тому же самому классу, который в отношении микропроцессов обсуждался выше (§ 14). Мы наблюдаем некоторое состояние, которое в истории развития изолированной системы крайне невероятно; как можно объяснить его? Как и раньше, ответ состоит в том, что наблюдаемая система, по предположению, представляет продукт взаимодействия и что мы предпочитаем интерпретацию, согласно которой система в прошлом не была изолирована. Как и для термодинамической статистики, этот вывод основан на вероятности по многим системам, которую можно свести к вероятности по одной системе. Следовательно, он предполагает гипотезу о ветвящейся структуре, относящуюся, однако, к статистике макроконфигураций.

По этой причине следы играют функцию протокола. Они позволяют нам сделать вывод, что несколько ранее имело место взаимодействие, что шаги человека явились причиной появления упорядоченного состояния песка; поскольку это состояние не «исчезло при перемешивании», оно представляет протокол взаимодействия. Протоколы являются упорядоченными макроконфигурациями, сохраняющими свой порядок; они образуют, так сказать, застывший порядок.

Признание статического порядка не противоречит предположению 3 § 16, так как это предположение относится к конфигурациям молекул, а система остается неизменной только в макроскопическом отношении. Тепловое состояние системы подвержено своему неизбежному развитию. Когда нога расталкивает песчинки, они приобретают кинетическую энергию, которая преобразуется в тепло, когда песчинки приходят в состояние покоя. История эволюции отпечатка ноги представляет ветвь, идущую от нижней точки к верхней точке и заканчивающуюся состоянием термодинамического равновесия. Но макроскопическое равновесие еще не достигнуто; и с точки зрения макростатистики и макроэнтропии ответвившаяся система не завершила своего пути и еще обнаруживает порядок. Однако перемешивающее действие ветра медленно направляет систему к состоянию, которое даже в макростатистике представляет равновесие, или к состоянию, в котором стерты следы макропорядка и отпечатки ног полностью сглажены.

Помимо выяснения вопроса о природе протоколов, пример с отпечатками ног помогает нам также проанализировать смысл *причинного объяснения*. Объяснение посредством причин необходимо в тех случаях, когда мы встречаемся с изолированной системой, находящейся в упорядоченном состоянии, крайне маловероятном в истории этой системы. Тогда мы предполагаем, что в более раннее время система не была изолирована; объяснение характеризует порядок в настоящем как следствие взаимодействия в прошлом. В других отношениях маловероятное событие представляется, таким образом, как вероятное благодаря замене низкой вероятности по одной системе высокой вероятностью по многим системам. Когда мы желаем объяснить маловероятный порядок посредством причин, мы руководствуемся правилом, что объяснение заключается в том, чтобы вывести маловероятное из ветвящейся структуры вселенной (см. предположения 1—5, § 16).

Эти результаты были получены на основе применения языка L_1. Переводя их на нейтральный язык L, мы получим следующее толкование терминов «причина» и «следствие»: *Причина есть взаимодействие на нижней границе ветви, пробегаемой изолированной системой, проявляющей порядок; а состояние порядка является следствием* [1].

Существует много примеров такой интерпретации причинного объяснения, в котором, если вернуться обратно к языку L_1, слово «причина» рассматривается как термин, относящийся к прошлому. Считается, что шрам на коже вызван порезом ножа; лужи воды, встречаемые засушливым летом высоко в горах, обусловлены таянием снега; если мы находим два листка бумаги с неправильными краями, но которые можно сложить вместе, мы делаем вывод, что они получились в результате разрыва одного листка бумаги надвое; появление чернильных знаков на бумаге объясняется путем взаимодействия — процесса написания, который обусловливает появление упорядоченных линий на бумаге, и так

[1] [Эта экспликация «причины» и «следствия» не может рассматриваться как всеобщая; обычно эти термины находят более широкое применение. Объяснение того, как из этого узкого значения возникает более широкое значение терминов, см. на стр. 212. — *М. Р.*]

далее. Во всех этих примерах наблюдаемый порядок является порядком макроконфигурации, который сохранился после того, как система достигла своего термодинамического равновесия после взаимодействия. В этом состоит важное значение макростатистики: она дает нам критерии порядка, которые не может дать микростатистика, и это позволяет нам раскрывать причинные цепи в макрокосме. Фактически источник происхождения понятия причины, по-видимому, связан с такими случаями, в которых человек сталкивался с неожиданным состоянием макропорядка, что побуждало его воображение допустить причины в прошлом. Но, конечно, это понятие аналогичным образом относится к упорядоченным состояниям микроконфигураций, так как эти состояния также требуют объяснения на основе взаимодействия, о чем уже говорилось в § 14.

Иногда нелегко установить наличие порядка в макроконфигурации, и для его обнаружения может потребоваться особый навык. Подобные протоколы, которые сумеет прочитать только опытный наблюдатель, играют важную роль в геологии и антропологии. Можно упомянуть об окаменелостях, которые являются результатом захоронения в песке живых организмов; о каменных орудиях, высеченных доисторическим человеком; об обнаруженных гранитных породах, расположенных в относительно упорядоченные линии и залегающих на оттаявших в ранние геологические периоды краях ледников. Потребовались годы, чтобы обнаружить, что конфигурации пород носят упорядоченный характер и нуждаются в объяснении. Иногда медленное изменение в макроконфигурации можно использовать для того, чтобы определить период времени, прошедший со времени взаимодействия. Подобную ситуацию представляет определение возраста горных пород и других тел по содержанию в них продуктов радиоактивного распада. В других случаях изменение состоит просто в дальнейшем приращении вещества, как в случае с кольцами деревьев, которые обусловлены ежегодным отложением слоев клеток и, таким образом, протоколируют возраст дерева.

Можно вспомнить, что до сих пор наш анализ проводился на языке L_1. Только определение причины и следствия было дано на ином языке, а именно — на нейтральном языке L. Мы нашли, что при использовании этого

определения объяснение в терминах L_1 всегда должно относиться к прошлому. Проверим теперь наши результаты на языке L_2, в котором употребляется противоположное направление времени.

С этой целью рассмотрим на языке L_2 пример с отпечатками следов на песке. Мысленно представив историю с отпечатками в обратном времени, мы получим следующее описание. Вначале на песке будут наблюдаться сглаженные отпечатки, отчасти напоминающие следы человека. Затем начинает дуть ветер, перенося песчинки туда и сюда, в результате чего отпечатки на песке становятся более отчетливыми и в конце концов принимают строгую форму следов человека. Наконец, появляется человек, идущий спиной вперед. Он последовательно вступает ногами в следы на песке, которые точно соответствуют его ноге. Когда он подымает свою ногу, след со всех сторон засыпается песком и полностью исчезает, так что это место на песке уже не отличается от своего окружения.

Согласно этой интерпретации, упорядоченное расположение сглаженных отпечатков не рассматривается как протокол; оно предшествует взаимодействию, а не следует за ним, и выражает состояние *до взаимодействия*, а не состояние *после взаимодействия*. Сами невероятные события, к которым этот порядок приводит, здесь нельзя объяснить на основе прошлых явлений. На языке L_2, при рассмотрении взаимодействий воздуха в прошлом, можно было бы утверждать, что сложились такие особые условия для воздуха, которые привели к необычным сериям воздействий, в результате чего сглаженные отпечатки преобразовались в строгие формы следов человека. Но в таком случае одно состояние порядка мы объясняем через другое, заданное расположением частиц воздуха. Объяснение предполагает сведение порядка к взаимодействию; в рамках L_2 в прошлом никакого такого объяснения нельзя найти.

Следовательно, мы пришли бы к весьма необычному заключению. Мы стали бы объяснять невероятные совпадения скорее их *целью*, чем их *причиной*. Ветер преобразует углубления в песке в следы человека *для того, чтобы* они соответствовали ноге человека, когда он будет проходить. Человек точно вступает в следы *для того, чтобы* они исчезали, когда он подымает свою ногу.

Когда мы пользуемся языком L_2, в котором направление времени противоположно обычному, объяснение приводит нас вместо *причинности* к *телеологизму*. Основанием для этого служит то, что, как и в случае L_1, объяснение заключается в выводе маловероятного порядка из ветвящейся структуры вселенной, гипотеза о которой была сформулирована на нейтральном языке в предположениях 1—5, § 16. Следовательно, объяснение на языке L_2 требует ссылки на будущее взаимодействие, которое представляет промежуточную стадию при получении вывода на основе предположений 1—5. Определение причины, данное выше в этом параграфе, могло бы быть перенесено в словарь L_2 просто путем замены слова «причина» на слово «цель»: взаимодействие в будущем, которое объясняет маловероятный порядок настоящего, следовало бы назвать целью явления.

Такое объяснение для нас выглядит неправдоподобным; тем не менее оно просто переводит наш собственный опыт на язык с противоположным направлением времени. Мы видим, что наше понятие *причинности*, понятие о прошлом, которое определяет настоящее и будущее, тесно связано с нашим определением положительного времени через возрастание энтропии. В случае противоположного направления времени мы находим его эквивалент в понятии *телеологизма*, согласно которому будущее определяет настоящее и прошлое.

Какой из языков верен: язык причинности или язык телеологизма? Это, несомненно, бессмысленный вопрос. Два языка, L_1 и L_2, представляют *эквивалентные описания* (см. § 4); каждый из них так же верен, как и другой. То, что язык L_1 кажется нам более естественным и что мы так сильно склонны к отождествлению направления от взаимодействия к порядку с положительным временем, имеет свое основание в природе человеческого организма. Обсуждение этой проблемы можно отложить[1].

Мы заключаем: Если мы определим направление времени обычным образом, для телеологизма не останется места и только причинность будет служить основанием

---

[1] [Эту проблему предполагалось обсудить в главе о человеческом разуме, которая должна была быть заключительной главой. — *М. Р.*]

для плодотворного объяснения. Эту точку зрения можно пояснить следующим образом. Если кто-то доказывает, что выбор направления возрастания энтропии в качестве направления времени является делом соглашения, его понимание нельзя назвать ложным. Но он не должен совершать ошибки, часто допускаемой при других формах конвенционализма: ошибки, состоящей в том, что не замечают эмпирического содержания, связанного с использованием этого соглашения. Эмпирический факт состоит в том, что во всех ответвившихся системах энтропия возрастает в том же самом направлении. На таком эмпирическом основании соглашение о том, что положительное время определяется через возрастание энтропии, невозможно отделить от признания причинности в качестве общего метода объяснения. Те, кто предпочитает давать объяснение на основе телеологизма, были бы вынуждены использовать противоположное направление времени и считать, что время идет от состояний с высокой энтропией к состояниям с низкой энтропией. Однако тогда телеологизм является просто другим языком, описывающим ту же самую вселенную, поскольку сформулированная на этом языке физика не отличалась бы от существующей. Конечно, пользоваться таким языком было бы чрезвычайно неудобно, потому что он противоречит направлению времени психологического опыта.

Этот анализ показывает, что использование возрастания энтропии для определения направления времени приводит к выводу, что причины, а не цели определяют события настоящего. Когда ученый не признает телеологию в качестве способа объяснения, он для этого имеет веские основания: именно второй закон термодинамики исключает телеологизм для мира, подобного нашему, в котором возрастание энтропии определяет положительное время. Если предполагается обычное направление времени, то вопрос о том, должны ли мы описывать мир на языке причин или цели, не зависит от личного предпочтения и от способа рассмотрения: то, что причинность, а не целенаправленность управляет вселенной, представляет физический закон. Связь между возрастанием энтропии и причиной как детерминирующим фактором всех событий дается статистической изотропностью вселенной, которая ведет к тому, что развитие

всех ответвившихся систем идет в одном и том же направлении и помещает точку взаимодействия в том конце ветви, который соответствует состоянию вселенной с более низким значением энтропии. Следовательно, для направления времени, соответствующего возрастанию энтропии, точка взаимодействия представляет начало, а не конец развития ответвившейся системы. Статистические взаимосвязи, содержащиеся в структуре смешивания, объясняют нашу концепцию, согласно которой прошлое *порождает (produces)* будущее, а не наоборот.

Слово «порождает» представляет собой статистическое понятие. Оно указывает направление от взаимодействия к изолированным состояниям, которые еще проявляют порядок, и оно означает, что наличие большого числа изолированных упорядоченных систем представляет логическое следствие частого наличия состояний взаимодействия, при которых происходит ответвление систем с относительно низкими значениями энтропии. Таким образом, состояние взаимодействия рассматривается как причина упорядоченных состояний. Этим показано, что различие между причиной и следствием обусловливается энтропией и соответствует различию между прошлым и будущим. Тем не менее оно предполагает не просто возрастание энтропии вселенной, но и наличие ветвящейся структуры, удовлетворяющей предположениями 2—5, § 16.

Причина порождает следствие, а так как причина оставляет следы в своем следствии, о ней можно сделать вывод, исходя из этого следствия. В этом лежит источник различия между прошлым и будущим, на который указывалось в определении 5, § 2. Это определение, по которому будущее не может быть запротоколировано, хотя прошлое можно запротоколировать, можно преобразовать в статистическое определение: *Упорядоченные изолированные состояния всегда есть состояния после взаимодействия и никогда не предшествуют ему.* Это определение вытекает из предположения 5, если положительное время определяется как направление возрастающей энтропии. В предыдущих иллюстрациях были даны примеры следов — протоколов прошлого.

Настоящее является посредником между прошлым и будущим; в нем содержатся активный агент, который производит будущее, и протоколы прошлого, которые

окончены и которые невозможно изменить. Когда мы освобождаем эти термины от сопутствующих им эмоций и переводим их на статистический язык, мы обнаруживаем, что они относятся к переходам от состояний взаимодействия к упорядоченным изолированным состояниям. Отсюда видно, что состояния взаимодействия обладают приоритетом в логическом отношении над упорядоченными изолированными состояниями, потому что их частое появление сформулировано в предположениях 1—5 и, таким образом, рассматриваются как фундаментальное свойство вселенной, из которого логически выводимо частое появление упорядоченных изолированных систем. Этот логический приоритет можно рассматривать как источник, из которого проистекает представление о порождении как некоторой активности; логически предшествующее выступает как определяющий фактор. Вместе со значением слова «порождение» дается значение слова «протоколирование», так как оно вытекает из приведенных здесь статистических соображений, согласно которым отношение *протоколирования* является обратным отношению *порождения*. Причина порождает следствие, следствие протоколирует причину. Если порождение рассматривается как активность, то протоколирование, следовательно, обозначает пассивность. Таким образом, эти эмоциональные термины получают толкование при статистическом определении направления времени.

Данное определение причины относится к особой связи, характерной для ответвившейся системы. Понятие причины, первоначально имевшее это узкое значение, затем получило более широкое употребление. В § 5 мы видели, что порядок времени можно определить (безотносительно к направлению) посредством причинных цепей и, кроме того, что причинная сеть как целое является упорядоченной. Отсюда следует, что, если мы припишем направление положительного времени одной причинной цепи, такое направление будет приписано всем причинным цепям. Этим отношением определяется направление времени для любых двух событий, которые связаны причинной цепью, в том числе и для обратимого процесса.

Следовательно, наиболее широкое определение причины и следствия дается путем ссылки на направление времени, что в свою очередь можно определить на

основе более узкого значения слова «порождение». Так, можно говорить о причине и следствии даже в отношении обратимых процессов, называя более раннее состояние причиной более позднего. Например, мы говорим, что положение и скорость Земли по ее орбите вокруг Солнца в настоящее время являются причиной ее положения и скорости, которые будут наблюдаться спустя три дня; и мы не в состоянии обратить эту связь. Если бы используемое здесь слово «причина» относилось просто к движению Земли вокруг Солнца, то было бы бессмысленным проводить различие между причиной и следствием, так как ни о том, ни о другом нельзя сказать, что одно «порождает» другое. Но с точки зрения взаимоотношений, справедливых для необратимых процессов в ответвившихся системах, использование слова «причина» в применении к макропроцессам, всецело управляемым законами механики, вполне обоснованно.

Можно было бы также заявить, что никакой макропроцесс не является обратимым, поскольку каждый такой процесс связан с некоторым тепловым изменением и, таким образом, с возрастанием энтропии. Например, Земля замедляется в своем вращении вокруг оси. В этом смысле каждый макропроцесс per se[1] обладает направлением времени. Но в таких случаях возрастание энтропии незаметно; с точки зрения психологии различие между причиной и следствием в их применении к механическим процессам необходимо рассматривать как возникающее в результате использования взаимосвязей, справедливых для необратимых процессов в ответвившихся системах.

§ 19. Принцип общей причины

Теперь мы рассмотрим некоторые применения макростатистики, которые еще раз приведут к различию между причиной и следствием и, таким образом, к определению направления времени. Результаты будут сформулированы в виде определенного принципа, действию которого подчиняются макроскопические конфигурации. Однако будет показано, что этот принцип не представляет собой нового предположения, а выводится из вто-

[1] Per se (*лат.*) — сам по себе.

рого закона термодинамики, если этот закон дополнить гипотезой о ветвящейся структуре.

Допустим, что от молнии вспыхнуло пламя и что сильный ветер раздувает и разносит огонь, что приводит к большому несчастью. Совпадение действия огня и ветра в данном случае имеет общее следствие — большой пожар. Но если мы спросим, почему произошло это совпадение, то мы обращаем внимание не на это общее следствие, а ищем общую причину. Гроза, которая вызвала молнию, вызвала также и ветер, и таким образом объясняется это маловероятное совпадение.

Схема этого рассуждения иллюстрирует правило, согласно которому маловероятное следует объяснять на основе причин, а не на основе следствий. Отложим исследование связи между этими выводами и гипотезой ветвящейся структуры и изучим отдельно сначала содержание этого вывода. Логическую схему, которой он подчиняется, можно назвать *принципом общей причины*. Этот принцип можно сформулировать следующим образом: *Для маловероятного совпадения должна существовать общая причина.*

В нашей повседневной жизни мы часто используем такие заключения. Допустим, что две лампы в комнате внезапно погасли. Мы считаем маловероятным, что обе лампы случайно перегорели в одно и то же время, и ищем перегоревший предохранитель или какой-нибудь другой разрыв в общей цепи питания. Маловероятное совпадение, таким образом, объясняется как результат действия общей причины. Общее следствие, то есть тот факт, что в комнате стало совершенно темно, не может объяснить такого совпадения. Или, допустим, несколько актеров заболели во время игры, и у всех появились симптомы пищевого отравления. Мы предполагаем, что недоброкачественная пища принималась одновременно, например что они вместе обедали, и, таким образом, совпадение стремимся объяснить на основе общей причины. Одновременное заболевание актеров имеет также и общее следствие: спектакль придется отменить, так как замена довольно большого числа актеров невозможна. Но это общее следствие не объясняет совпадения.

Случайные совпадения, конечно, возможны: лампы могут перегореть одновременно, актеры одновременно заболеть по разным причинам. Следовательно, суще-

ствование общей причины в таких случаях определено не абсолютно, а только с некоторой вероятностью. При повторяющихся совпадениях эта вероятность сильно увеличивается. Предположим, что два гейзера, расположенные недалеко один от другого, фонтанируют не регулярным образом, но всегда одновременно. Тогда практически несомненно, что существует подземная связь этих двух гейзеров с общим источником горячей воды. Измерительные приборы, например барометры, всегда дают одинаковые показания, если они не весьма далеко удалены один от другого; и этот факт является следствием существования общей причины — в данном случае давления воздуха. Легко найти и другие примеры.

Целесообразно рассмотреть принцип общей причины с точки зрения статистики. Для этого мы допустим, что наблюдалось частое появление A и B, так что можно говорить о вероятностях $P(A)$, $P(B)$ и $P(A.B)$ в отношении к определенным отрезкам времени. Например, мы можем подсчитывать количество извержений гейзеров по дням или по часам и, таким образом, получим вероятность того, будут ли гейзеры действовать в заданный день или в заданный час. Одновременно мы обобщим проблему, распространив ее на случаи статистической связи между причиной и следствием, то есть на такие случаи, в которых причина вызывает следствие лишь с некоторой вероятностью. Тогда в это общее рассмотрение будут включены и те особые случаи, когда вероятность практически равна единице, как в приведенных примерах.

Статистические связи двух одновременных событий с общей причиной, с одной стороны, и с общим следствием, с другой стороны, можно представить схемой, изображенной на рис. 23. Схема имеет вид сдвоенной вилки, в которой A и B представляют два события, одновременное осуществление которых маловероятно. C представляет их общую причину, E — их общее следствие. Мы говорим, что C, а не E объясняет одновременное осуществление A и B. Поэтому опустим верхнюю часть диаграммы и исследуем отдельную вилку, изображенную на рис. 24, в которой указана только общая причина C.

Если события A и B одновременно происходят чаще, чем этого можно ожидать для случайных совпадений, то

этот факт можно записать следующим образом:

$$P(A.\ B) > P(A) \cdot P(B). \qquad (1)$$

Применяя к левой части этого соотношения общую теорему умножения вероятностей

$$P(A.\ B) = P(A) \cdot P(A,\ B) = P(B) \cdot P(B,\ A), \qquad (2)$$

мы из него выведем два других соотношения:

$$P(A,\ B) > P(B), \qquad (3)$$

$$P(B,\ A) > P(A). \qquad (4)$$

И наоборот, каждое из этих соотношений можно использовать для вывода (1). Поэтому (3) или (4) эквива-

Рис. 23. Двойная вилка, образованная общей причиной и общим следствием.

Рис. 24. Открытая в будущее вилка, образованная общей причиной.

Рис. 25. Открытая в прошлое вилка, образованная общим следствием.

лентны (1). При этих выводах всегда предполагается, что ни одна из этих вероятностей не обращается в нуль.

Чтобы объяснить такое совпадение событий A и B, вероятность которого больше вероятности случайного совпадения, мы предполагаем, что существует общая причина C. Если существует не одна, а большее число возможных общих причин, то C может представлять дизъюнкцию этих причин. Теперь мы предположим, что вилка ACB удовлетворяет следующим соотношениям:

$$P(C,\ A.\ B) = P(C,\ A) \cdot P(C,\ B), \qquad (5)$$

$$P(\bar{C},\ A.\ B) = P(\bar{C},\ A) \cdot P(\bar{C},\ B), \qquad (6)$$

$$P(C,\ A) > P(\bar{C},\ A), \qquad (7)$$

$$P(C,\ B) > P(\bar{C},\ B). \qquad (8)$$

В дальнейшем будет показано, что соотношение (1) можно вывести из этих соотношений. На этом основании мы будем говорить, что соотношения (5) — (8) определяют *вилку конъюнкции*, то есть вилку, согласно которой конъюнкция двух событий A и B обладает большей частотой, чем это было бы для независимых событий. Когда мы говорим о том, что общая причина C *объясняет* частое совпадение, то, помимо самой возможности получить вывод соотношения (1), мы имеем в виду также и то, что по отношению к причине C события A и B взаимно независимы: *статистическая зависимость* здесь выводится из независимости. Общая причина представляет связующее звено, которое преобразует независимость в зависимость. Поэтому конъюнктивная вилка представляет статистическую модель связи, выраженной в принципе общей причины.

Для математического рассмотрения вопроса введем следующие обозначения:

$$P(A)=a,\ P(B)=b,\ P(C)=c,\ P(A.\ B)=w,$$
$$P(C,\ A)=u,\ P(C,\ B)=v,\ P(\overline{C},\ A)=r,\ P(\overline{C},\ B)=s. \tag{9}$$

Правило исключения[1] дает нам следующие уравнения:

$$w=cuv+(1-c)rs, \tag{10}$$
$$a=cu+(1-c)r, \qquad b=cv+(1-c)s. \tag{11}$$

Из этих соотношений с помощью простых выкладок получим выражение в виде произведения множителей

$$w-ab=c(1-c)(u-r)(v-s). \tag{12}$$

Мы примем, что $0<c<1$. При учете соотношений (7) — (8) правая часть (12) больше нуля, следовательно, $w>ab$. Этим доказывается соотношение (1).

Кроме того, мы видим, что $w=ab$, если выполняются равенства $u=r$, или $v=s$, или оба сразу. Это означает, что статистическая зависимость A и B получается из того, что, как сформулировано в (7)—(8), по отношению к C и $\overline{C}$ вероятности A и B неравны. Если $u<r$ и также $v<s$, мы еще раз получаем $w>ab$; в этом случае C и $\overline{C}$ просто меняются местами. Если $u>r$, а $v<s$, мы имеем $w<ab$. Этот случай можно свести к первому,

[1] ThP, p. 76.

заменив B на B. Следовательно, нам нужно рассмотреть только первый случай.

Используя соотношения (7)—(8), мы получим из (11)

$$u > a > r, \quad v > b > s. \tag{13}$$

Первая часть первого соотношения получается из первого уравнения (11) следующим образом. Если в (11) заменить r на u, мы получим $a = u$; и если $r < u$, то значение a уменьшается. Подобным же образом мы докажем вторую часть неравенства, а также и второе соотношение в (13).

Аналогичным образом, используя (7) — (8), из (10) мы получаем

$$uv > w. \tag{14}$$

Для доказательства заменим в (10) rs на uv; тогда найдем $w = uv$. Отсюда при $rs < uv$ значение w будет меньше, чем uv.

Соотношения (13) и (14) в вероятностных обозначениях принимают вид

$$P(C, A) > P(A) > P(\bar{C}, A), \quad P(C, B) > P(B) > P(\bar{C}, B), \tag{15}$$

$$P(C, A. B) > P(A, B). \tag{16}$$

Для обратных вероятностей мы имеем подобные же соотношения. Используя правило умножения в виде, аналогичном (2), из (15)—(16) мы получаем

$$P(A, C) > P(C), \quad P(B, C) > P(C), \tag{17}$$

$$P(A. B, C) > P(C). \tag{18}$$

Эти обратные вероятности можно вычислить следующим образом:

$$P(A, C) = \frac{u}{a} \cdot c, \quad P(B, C) = \frac{v}{b} \cdot c, \tag{19}$$

$$P(A. B, C) = \frac{P(A. B. C)}{P(A. B)} = \frac{uv}{w} \cdot c. \tag{20}$$

Привлекая (13), мы видим, что в каждом из этих соотношений множитель перед c больше единицы. Это означает, что вероятность осуществления события C увеличивается, если произошло событие A, или событие B,

или их сочетание. Получив эти результаты, мы полностью определили статистические свойства вилки конъюнкции[1].

До сих пор мы предполагали, что для вилки, изображенной на рис. 24, направление времени известно. Теперь допустим, что известен только *порядок* времени. Тогда мы с таким же успехом можем начертить вилку, заменив на рис. 25 E на C. Можно ли теперь использовать условия (5)—(8) для определения *направления* времени вилки, то есть чтобы провести различие между рис. 24 и 25?

Легко видеть, что непосредственно этого сделать нельзя. Общее следствие E, то есть верхняя вилка рис. 23, вполне может удовлетворять соотношениям (5)—(8) с заменой E вместо C. В действительности обычно это справедливо. Например, извержение гейзеров может иметь следствием образование двух облаков, которые сливаются в одно большое облако. Тогда появление этого большого облака представляет такое следствие E, которое удовлетворяет условиям (5)—(8).

Однако существует опосредованный путь использования соотношений (5)—(8) для определения направления времени. Для общего следствия E эти соотношения могут выполняться только при условии, что существует также и общая причина C, удовлетворяющая тем же соотношениям. То есть тогда мы должны иметь сдвоенную вилку, изображенную на рис. 23, а не одну, изображенную на рис. 25. Это видно из следующего. Неравенство (1) точно с таким же успехом получается из (5)—(8) и в том случае, если эти соотношения содержат E на месте C. Поэтому если существует вилка конъюнкции по отношению к общему следствию E, то одновременное осуществление A и B более вероятно, чем просто их случайное совпадение. Следовательно, если бы

---

[1] Иногда принцип общей причины используется в более слабой форме, в которой отказываются от соотношений (5) — (6). В этом случае считается достаточным, если причина C удовлетворяет неравенствам (1), (7)—(8) и (16). Неравенства (15) и (17) — (18) тогда могут быть выведены. Однако обычно утверждается, что при использовании более детального описания причина C настолько строго определяется, что также удовлетворяются (5)—(6). См. замечания по поводу отношения «между» на стр. 254, § 22.

не существовало общей причины C, то общее следствие установило бы существование статистической зависимости между A и B, — и объяснение было бы дано посредством «конечной цели». Учитывая содержание § 18, мы рассматриваем конечные цели как несовместимые со вторым законом термодинамики и считаем такие вилки невозможными. В данном случае это означает: принцип общей причины не всегда исключает статистическую зависимость в отношении к общему следствию, но он исключает такую зависимость, если общая причина не существует.

Наоборот, вполне возможно, что при наличии общей причины, удовлетворяющей (5)—(8), общего следствия не бывает. Это может быть использовано следующим образом для определения направления времени.

Если вилка ACB удовлетворяет условию конъюнктивности, то мы говорим, что вилка *замкнута* в C. Если по другую сторону вилки, удовлетворяющей (5)—(8), не существует события E, то мы говорим, что вилка открыта с этой стороны. Теперь дадим определение.

Определение. В конъюнктивной вилке ACB, открытой с одной стороны, C происходит раньше, чем A или B.

Это и есть макростатистическое определение направления времени[1].

---

[1] Определение направления времени, основанное на представлениях о вилках статистически связанных событий, для которых статистические корреляции могут устанавливаться только общей причиной, а не общим следствием, было дано ранее; см. Hans Reichenbach, Die Kausalstruktur der Welt und der Unterschied von Vergangenheit und Zukunft, в „Sitzungsber. d. Bayer. Akad. d. Wiss.", math.-naturwiss, Abt., 1925, S. 133—175. Однако приведенные в этой статье математические и логические соотношения, касающиеся вероятностных импликаций, являются ошибочными. Она была написана до того, как автором была разработана общая теория вероятностей, и содержащиеся в ней результаты теперь необходимо заменить соотношениями, представленными в данном разделе книги. Кроме того, я полагал, что, исходя из принципа энтропии, невозможно получить вывод о направлении времени. Эту концепцию в настоящее время, на основе приведенного в данной работе доказательства, я заменил другой, согласно которой такой вывод возможен, если принцип энтропии дополнить гипотезой о ветвящейся структуре. Представления о направлении времени в макростатистике аналогичны представлениям о направлении времени в микростатистике. Дальнейшее исправление взглядов, выраженных в упомянутой статье, см. § 23, прим. на стр. 265.

Эти результаты с точки зрения *принципа общей причины* можно подытожить следующим образом: если совпадения двух событий A и B встречаются более часто, чем это соответствовало бы их независимому осуществлению, то есть если события удовлетворяют соотношению (1), то для этих событий существует такая общая причина C, что вилка ACB будет удовлетворять условию конъюнктивности, то есть соотношениям (5)—(8).

Этот принцип означает, что статистическую зависимость одновременно происходящих событий необходимо объяснять на основе общей причины. Общее следствие нельзя рассматривать как объяснение, и, таким образом, нет необходимости в его существовании, хотя оно и может существовать и даже может устанавливать наличие статистической зависимости.

Конечно, существуют также вилки, открытые в прошлое и, таким образом, заостренные в будущее, такие, как вилка, изображенная на рис. 25; но такие вилки никогда не удовлетворяют соотношениям (5)—(8). Например, когда сближаются два грузовика, идущие по шоссе навстречу друг другу, их водители обычно обмениваются приветствиями, временами включая и выключая фары. Здесь мы имеем вилку AEB, где E — обмен приветствиями, который представляет общее следствие «совпадения» грузовиков, то есть событий A и B. Здесь мы получаем

$$P(E,\ A.\ B) > P(A) \cdot P(B). \qquad (21)$$

Это означает: из следствия E мы можем с большой вероятностью заключить, что два грузовика проследовали мимо друг друга. Но соотношение (1) не выполняется, вместо него мы имеем здесь

$$P(A.\ B) = P(A) \cdot P(B). \qquad (22)$$

Это соотношение устанавливает, что общее следствие, обмен приветствиями, не делает более частой встречу грузовиков. Следовательно, вилка не конъюнктивна.

Кроме того, существуют вилки ACB, в которых общая причина не устанавливает статистической зависимости между A и B. Например, пусть бросается несколько раз игральная кость; в ряду проистекающих событий выпадение шестерки, что мы будем называть

событием A, имеет вероятность $P(A) = 1/6$. Временами той же рукой может бросаться вторая кость; если здесь выпадет шестерка, мы будем говорить о наступлении события B. Если события A и B происходят вместе, то они берут начало от одной и той же причины C — бросания двух костей. Хотя здесь соотношения (5)—(6) удовлетворяются, соотношение (7) нарушается, так как оно заменяется равенством. Тогда из (12) мы выводим, что $w = ab$; следовательно, (1) заменяется равенством, и вилка становится не конъюнктивной.

Теперь мы рассмотрим связь между принципом общей причины и вторым законом термодинамики. На первый взгляд кажется, что этот принцип не имеет никакого отношения к концепции Больцмана, так как явления, к которым он относится, явно не похожи на процессы перемешивания. Но если вспомним описание процесса перемешивания при помощи вероятностной структуры, рассмотренное в § 12, то мы обнаружим, что этот взгляд неверен. Фактически эта структура процессов перемешивания содержит те же самые соотношения, на которых строится принцип общей причины.

Предположим, что имеется ансамбль ответвившихся систем, состоящий из систем двух классов. В классе T_A одним из возможных состояний будет состояние A, в классе T_B среди возможных состояний мы найдем состояние B. Последовательности состояний, через которые с течением времени проходит каждая система, строго не определены, за исключением только тех случаев, когда система не полностью изолирована от своего окружения; но линию развития каждой системы можно рассматривать как вероятностную цепь с последействием. Кроме того, предположим, что иногда, хотя и не всегда, система класса T_A и система класса T_B берут начало только от одного состояния взаимодействия C, то есть, что две системы ответвляются от одного причинного источника. И это происхождение из одного и того же причинного события приводит к частому совместному появлению состояний A и B и выражается статистической зависимостью.

Так как здесь мы имеем системы двух различных классов, то существуют две вероятностные структуры x_{ki} и y_{ki}; представим, что последняя находится непосредственно над первой. Первые столбцы структур обна-

руживают связь между собой, обусловленную событием C; всякий раз, когда нижняя структура в этом столбце имеет C, прямо над ним стоит C и в верхней структуре. Следовательно, то же самое справедливо и для $\bar{C}$, хотя случай $\bar{C}$ может представлять более широкий класс событий, чем C; он просто означает отсутствие C. Однако в остальном, если не считать этой связи, структуры независимы. Это условие можно записать в виде

$$\begin{aligned} P(C^{kl},\ A^{kl}.\ B^{kl})^k &= P(C^{kl},\ A^{kl})^k \cdot P(C^{kl},\ B^{kl})^k, \\ P(\bar{C}^{kl},\ A^{kl}.\ B^{kl})^k &= P(\bar{C}^{kl},\ A^{kl})^k \cdot P(\bar{C}^{kl},\ B^{kl})^k. \end{aligned} \tag{23}$$

Эти соотношения представляют условия (5)—(6), записанные в форме, принятой для вероятностных структур. Условия (7)—(8) принимают вид

$$P(C^{kl},\ A^{kl})^k > P(\bar{C}^{kl},\ A^{kl})^k,\ P(C^{kl},\ B^{kl})^k > P(\bar{C}^{kl},\ B^{kl})^k. \tag{24}$$

Теперь, как это делалось раньше, мы можем получить соотношение (12) и вместе с ним соотношение (1). Следовательно, мы имеем

$$P(A^{kl}.\ B^{kl})^k > P(A^{kl})^k \cdot P(B^{kl})^k. \tag{25}$$

Если мы предположим, что неравенства (24) выполняются заметным образом только при не очень большом удалении от причины C, то есть что с возрастанием индекса l соотношения (24) стремятся перейти в равенства, то мы найдем

$$\lim_{l\to\infty} P(A^{kl}.\ B^{kl})^k = P(A^{kl})^k \cdot P(B^{kl})^k. \tag{26}$$

Соотношения (25) и (26), преобразованные для относительных вероятностей, принимают вид

$$P(A^{kl},\ B^{kl})^k > P(B^{kl})^k,\ P(B^{kl},\ A^{kl}) > P(A^{kl})^k, \tag{27}$$

$$\lim_{l\to\infty} P(A^{kl},\ B^{kl})^k = P(B^{kl})^k,\ \lim_{l\to\infty} P(B^{kl},\ A^{kl})^k = P(A^{kl})^k. \tag{28}$$

Соотношения (27)—(28) соответствуют соотношениям (12.23)—(12.24), если A заменить на B. Единственное различие между ними состоит в том, что в (27)—(28)

мы не используем индекс в виде $k-1$. Это различие проистекает из того, что в вероятностной структуре в § 12 (12.6) мы использовали попеременно строки двух различных классов, тогда как в данном случае структура построена как суперпозиция двух структур, так что два события A и B, образующие пару, принадлежат к одной строке.

Существует и дальнейшее различие. По отношению к вероятностной структуре событий A и B мы не говорим о соотношении, аналогичном (12.22), которое сравнивает вертикальные вероятности с горизонтальными. Фактически, если мы продолжим линии причинности, ведущие от C или $\overline{C}$ к A или B, то последние события обычно не повторяются на этих линиях. Структуру событий A и B мы можем сравнить со структурой энтропии (14.1) и рассматривать A и B как состояния с очень низкой энтропией, которые в истории развития системы не повторяются в течение исключительно больших промежутков времени.

Наша структура общей причины представляет только начальные участки строк, в которых энтропия увеличивается, она не распространяется на участки строк, соответствующие более позднему времени. Фактически структуру, используемую для выражения принципа общей причины, можно рассматривать как рудиментарную форму вероятностной структуры процессов смешивания, в которой реализованы только первые столбцы, и соотношения, выражающие процесс смешивания, относятся только к вертикальным вероятностям.

Рассмотрим с этой точки зрения двойную структуру, которая подчиняется соотношениям (23) — (28). Первый столбец содержит C и $\overline{C}$. Для строк, начинающихся с C, вертикальная вероятность осуществления пар $A.B$ в некотором частном столбце i выше, чем для других строк. Действительно, A и B встречаются только в этом столбце: идеализируя проблему, мы допустим, что эти события могут также встретиться и в поздних столбцах, хотя, может быть, никогда и не встречаются дважды в одной строке. Влияние конфигурации первого столбца сказывается в наличии статистической зависимости A и B в i-ом столбце; это выражено в соотношениях (25) и (27). При идеализации проблемы эта статистическая

зависимость A и B пропадает, как сформулировано в (26) и (28).

Спаривание событий в столбце i служит признаком структуры процесса смешивания. Оно относится к процессу смешивания в математическом смысле, то есть к процессу смешивания, в котором элементы «перемешиваются» не посредством физического механизма, а отношениями независимости и структурной инвариантности.

В рассматриваемом случае структура характеризуется теми же самыми статистическими соотношениями, которые справедливы для процесса смешивания молекул или «смешивания» ответвившихся систем; в данной структуре они очевидны благодаря наличию внутренней зависимости, или последействия, внутри столбцов, что обсуждалось при рассмотрении структуры отношений (12.23). Таким образом, пространственный ансамбль здесь характеризует смешивание, несмотря даже на то, что не могут использоваться вероятности, относящиеся к временно̀му ансамблю. Поэтому мы будем считать, что рудиментарная вероятностная структура, формулирующая принцип общей причины, подчиняется тем же самым физическим правилам, что и обычная.

При этом условии можно гипотезу о ветвящейся структуре применить к данной вероятностной структуре, если эта гипотеза расширена так, чтобы включать как макростатистику, так и микростатистику. Эта гипотеза говорит нам, что если некоторое состояние в пространственном ансамбле встречается более часто, чем это соответствует определенному стандарту, а именно, его вероятности во временно̀м ансамбле, то в прошлом должно было осуществиться состояние взаимодействия.

У нашей рудиментарной структуры нет временно̀го ансамбля. Поэтому этот стандарт мы заменяем некоторой вероятностью, относящейся к пространственному ансамблю. Эта вероятность получается следующим образом.

Если бы две линии причинности, ведущие к A и B, не были связаны своими первыми элементами, то вероятность совместного осуществления определялась бы произведением

$$P(A) \cdot P(B) \tag{29}$$

как для временного, так и для пространственного ансамбля. Существует только последний ансамбль, так что мы берем значение (29) из пространственного ансамбля. Если мы теперь будем считать, что эти две линии образуют одну систему, а совместное осуществление A и B будем рассматривать как состояние этой системы, тогда величину (29) можно трактовать как вероятность временного ансамбля этой системы, то есть как стандартную вероятность. Так как в пространственном ансамбле конъюнкция $A.B$ наблюдается более часто, чем это соответствует (29), то отсюда можно сделать вывод, что это отклонение от стандарта требует наличия взаимодействия в прошлом. Это означает, что две линии должны брать начало от процесса взаимодействия, который содержит в себе их обеих.

Следует отметить, что эти выводы точно соответствуют тем выводам, которые мы делаем для обычных вероятностных структур процессов смешивания. Например, в структуре, данной соотношениями (12.22)—(12.24), мы имеем в точности такие же условия. Здесь вероятность B для временного ансамбля можно подсчитать в строке, но, подсчитывая по любому столбцу, мы найдем ту же самую величину для пространственного ансамбля. Тот факт, что при подсчете по пространственному ансамблю мы получаем значение частоты, превышающее норму для совместного осуществления B и $\bar{B}$, выражен в соотношениях (12.23)—(12.24). Эта более высокая вертикальная вероятность возможна только благодаря тому, что строки этой структуры отсчитываются, начиная с состояния взаимодействия в прошлом, состояния, в котором два газа начинают взаимодействовать, то есть вступают в процесс перемешивания.

Таким образом, статистическая зависимость в пространственном ансамбле нуждается в объяснении на основе общей причины. Это правило вытекает из принципа, что встречающееся маловероятное в ходе развития системы состояние необходимо объяснять на основе предшествующего состояния взаимодействия. Правило выполняется, когда выражение «маловероятное в ходе развития системы» заменяется указанием на стандартную вероятность, которую можно оценить по пространственному ансамблю, если временной ансамбль не существует, и которая тождественна с вероятностью по

временному ансамблю, если таковой существует. Таким образом, определение направления времени, даваемое принципом общей причины, идентично с правилом, согласно которому наблюдение упорядоченного состояния изолированной системы требует состояния взаимодействия в прошлом, то есть с правилом, что упорядоченное состояние возникает после взаимодействия. Одновременное осуществление A и B рассматривается как указание на наличие порядка, потому что оно происходит чаще, чем это было бы при случайном совпадении.

Если бы вилка конъюнкции была открыта в сторону прошлого, то это означало бы, что существуют процессы смешивания, идущие в противоположном направлении от состояния, которое в пространственном ансамбле встречается чаще, чем это определяется стандартной вероятностью, в направлении к состоянию взаимодействия. То, что такие процессы не происходят и что упорядоченные состояния не предшествуют взаимодействию, представляет содержание принципа параллелизма в возрастании энтропии, сформулированного в предположении 5 гипотезы о ветвящейся структуре (§ 16). Было показано, что этот принцип выражает общее статистическое свойство вселенной, свойство ее статистической изотропности. Таким образом, принцип общей причины повторяет тот самый принцип, который составляет ядро гипотезы о ветвящейся структуре.

§ 20. Энтропия и информация

Поскольку второй закон термодинамики приводит к заключению о существовании протоколов о прошлом, а протоколы хранят информацию, то следует ожидать, что между энтропией и информацией существует тесная связь. Недавно эта зависимость стала объектом интересных исследований по теории информации, которая возникла при изучении вопросов, относящихся к передаче информации по каналам связи и хранению информации в вычислительных машинах. Было показано, что для этих процессов можно определить меру информации, которая в действительности имеет свойства энтропии и которая оказалась чрезвычайно ценной при решении практических задач. Поскольку применяемые здесь понятия относятся к вероятностям, установленным для статистики

макроскопических объектов, то, следовательно, необходимо определить меру энтропии для макростатистики, то есть понятие макроэнтропии.

Термин «информация» относится к *вероятностной дизъюнкции*, или *вероятностной ситуации*. Дизъюнкция этого рода дается полной и единственной дизъюнкцией событий $B_1 \vee \ldots \vee B_m$, для которых известна система вероятностей $p_1 \ldots p_m$. Отсюда следует, что pi удовлетворяют условию

$$\sum_{i=1}^{m} p_i = 1. \qquad (1)$$

Спрашивается, как можно измерить информацию, выраженную такой дизъюнкцией.

Сначала допустим, что дизъюнкция имеет только два члена и что мы, таким образом, имеем только вероятности p и $1-p$. Мы скажем, что дизъюнкция содержит большую информацию, если, используя ее, мы сможем сделать большое число предсказаний, которые сбудутся. Мы обладаем выбором в предсказании; следовательно, мы предскажем либо B, либо $\bar{B}$, в зависимости от того, что больше $^1/_2$, значение p или $1-p$. Тогда мера информации выразилась бы просто через p или через $1-p$, в зависимости от того, что больше. Максимальное значение этой меры информации было бы равно единице; минимальное — равно $^1/_2$. Информация этого рода будет называться информацией о событии, или *интенсиональной информацией*. Она возрастает с увеличением процента правильных предсказаний и, таким образом, выражает *степень предсказуемости*, связанную с дизъюнкцией.

Существует, однако, информация другого вида. Когда наше предсказание оказывается неверным, то мы тоже кое-что узнаем из наблюдения, то есть событие сообщает нам информацию. Никакой подобной информации мы не получаем, когда предсказание верно, потому что в таком случае событие просто свидетельствует нам о том, что мы и ожидали. Информация второго вида будет называться информацией *из* события или *экстенсиональной информацией*. Ее мера возрастает с увеличением числа ложных предсказаний. Экстенсиональная информация обратна интенсиональной информации. Термин «информация» в его обычном смысле означает «интенсиональную информацию», но иногда для практических целей

удобнее использовать «экстенсиональную информацию». Очевидно, что если определен один вид информации, то другой вид также определен; эти два вида представляют просто различные способы выражения. Это различие подобно различию между теплом и холодом или между большим и малым. В дальнейшем мы будем применять слово *информация* в смысле интенсиональной информации, когда мы будем говорить об экстенсиональной информации, то будем это специально оговаривать.

Значение p, или $1-p$ было бы возможной мерой информации для дизъюнкции двух членов, но оно не подходит для большего их числа. Предположим, что мы имеем три вероятности p_1, p_2, p_3 и пусть $p_1 > p_2 > p_3$. Тогда мы предскажем событие B_1 и окажемся правы в p_1 процентах всех случаев. Если мы теперь используем p_1 в качестве меры информации, то не будет отражено различие между величинами p_2 и p_3. Поэтому будем искать другую меру. Мы допустим возможность двойного предсказания, вторая часть которого выражена в условной форме: сначала мы предсказываем B_1, а для случая, если B_1 не наступит, мы предсказываем B_2. Это предсказание сбудется в $p_1 + p_2$ случаях. Если мы используем $p_1 + p_2$ в качестве меры информации, то в этом найдет отражение тот факт, что как p_1, так и p_2 больше, чем p_3, но мы не отразим различия между p_1 и p_2. Действительно, предсказывая сначала B_2 и, если оно не наступит, B_1, мы приходим к тому же значению.

Поэтому мы изменим способ вычисления меры следующим образом. Мы будем приписывать предсказанию больший вес w_1, если сбывается его первая часть, и меньший вес w_2, если сбывается вторая часть. Представим себе человека, неоднократно делающего двойные предсказания: каждый раз, когда сбывается первая часть предсказания, ему засчитывается величина w_1 и каждый раз, когда сбывается вторая часть, на его счет заносится величина w_2. Затем сумма всех w_i делится на число наблюдений n. Применим этот способ подсчета к интересующему нас случаю, когда мы сначала предсказываем B_1, а затем B_2; тогда мы получим выражение

$$p_1 w_1 + p_2 w_2. \qquad (2)$$

Если бы кто-нибудь стал предсказывать сначала B_2, то получил бы выражение, в котором p_1 и p_2 поменялись

бы местами и которое, таким образом, обладало бы меньшим значением.

Поскольку мы уже ввели веса, мы можем с таким же успехом допустить тройное предсказание, добавив к предыдущему новую условную форму: если не сбудутся ни B_1, ни B_2, то наступит B_3. Конечно, это тройное предсказание всегда будет сбываться; но, закрепляя за третьим методом наименьший вес w_3, мы приходим к выражению

$$p_1w_1+p_2w_2+p_3w_3;\quad p_1>p_2>p_3,\ w_1>w_2>w_3, \tag{3}$$

в котором отражено различие между p_i и в силу которого мы вынуждены располагать предсказания в порядке, зависящем от величины их вероятностей. Добавление третьего члена имеет еще и формальное преимущество, заключающееся в том, что эта мера информации относится ко всей дизъюнкции, а не только к ее части. И, как отмечалось выше, информация представляет, таким образом, меру степени предсказуемости, связанную с дизъюнкцией.

Меру (3) можно распространить на дизъюнкцию любого числа членов m, включая и случай $m=2$. В этом случае, очевидно, мера (3) возрастает с увеличением значения большей из двух вероятностей p и $1-p$ и, таким образом, идентична мере, рассмотренной ранее, до того как мы стали приписывать веса отдельным предсказаниям, хотя ее численное значение и может быть другим. В исследованиях по теории информации[1] показано, что w_i удобно использовать в следующей форме:

$$w_i=\log p_i. \tag{4}$$

Эта форма имеет то преимущество, что величина w_i автоматически увеличивается с ростом значения p_i; таким образом, нам не нужно специально оговаривать условия, выраженные неравенствами в (3). Поскольку

[1] См. Claude E. Shannon, Warren Weaver, The Mathematical Theory of Communication, Urbana, Ill., Univ. Ill. Press, 1949, p. 19. Однако функция H, которую используют Шеннон и Уивер, представляет отрицание функции, данной в уравнении (5) этого параграфа, и, таким образом, выражает экстенсиональную информацию; кроме того, не включена константа $\log m$. См. также Норберт Винер, Кибернетика, М., 1958, гл. 3.

логарифм дроби меньшей единицы величина отрицательная, то к выражению (3) целесообразно добавить положительную константу, чтобы сделать полное значение меры положительным. Таким образом, мы приходим к следующему определению меры интенсиональной информации:

$$H(p)=\sum_{i=1}^{m} p_i \log p_i + \log m. \tag{5}$$

Можно показать, что $H \geqslant 0$ для всех возможных значений p_i. Функция H достигает своего наименьшего значения, равного 0, если все p_i равны между собой, и, таким образом, $p_i = \frac{1}{m}$. Она достигает своего максимального значения, если одно из p_i равно 1, а все остальные равны 0. В этом случае выражения под знаком суммы для всех p_i равны 0, и мы получаем $H = \log m$ для максимального H. В действительности это образует ситуацию с наибольшей степенью предсказуемости, то есть ситуацию с наибольшей информацией о событии. Минимум $H = 0$ соответствует неопределенной ситуации, когда ни одно B_i не заслуживает предпочтения перед другими. Выше мы видели, что при $m = 2$ в таком случае в действительности имеет место самая низкая степень предсказуемости. На рис. 26 представлена кривая $H(p)$ как функция от p_1 для $m = 2$.

Если мы имеем две вероятностные дизъюнкции $p_1 \ldots p_m$ и $p'_1 \ldots p'_m$, то нас может заинтересовать информация, которую дает нам их объединение. Это означает, что мы хотим знать информацию, относящуюся к возможным сочетаниям событий, к которым относятся эти вероятности. Предположим, что события независимы, тогда мы можем перемножить вероятности и получим

$$\begin{aligned} H(p \cdot p') &= \sum_{i=1}^{m} \sum_{k=1}^{m} p_i p'_k \log p_i p'_k + \log m^2 \\ &= \sum_{i=1}^{m} p_i \log p_i \sum_{k=1}^{m} p'_k + \sum_{k=1}^{m} p'_k \log p'_k \sum_{i=1}^{m} p_i + 2 \log m. \end{aligned} \tag{6}$$

Используя (1), мы находим

$$H(p \cdot p') = H(p) + H(p'). \tag{7}$$

Это означает, что для независимых событий информация аддитивна.

Иногда удобно ввести обобщенную меру информации. Пусть мы считаем, что дизъюнкция $B_1 \vee \ldots \vee B_m$ связана с системой вероятностей $q_1 \ldots q_m$, тогда как на самом деле она подчиняется вероятностям $p_1 \ldots p_m$. Когда мы теперь делаем предсказание, то для каждого

Рис. 26. Кривая информации $H(p)$ для двухчленной дизъюнкции, связываемой с вероятностями p_1 и $p_2 = 1 - p_1$, согласно формуле (5). Минимальное значение $H = 0$ достигается при $p_1 = p_2 = 1/2$; максимальное значение $H = \log 2$ при $p_1 = 0$ и при $p_1 = 1$.

наблюдаемого события B_i будем записывать вес $\log q_i$, согласно нашему знанию о ситуации. Как обычно, полная сумма всех записанных величин делится на число n наблюденных случаев. Поскольку частоты событий B_i в действительности контролируются p_i, то мы получаем меру

$$H(p;\ q) = \sum_{i=1}^{m} p_i \log q_i + \log m. \qquad (8)$$

Мы называем это информацией, содержащейся в $q_1 \ldots q_m$ по отношению к реальной ситуации $p_1 \ldots p_m$. Так как наша статистика наблюдений вскоре покажет нам, что q_i не применимы в этой ситуации, то, в конце концов, мы их заменим на p_i, которые мы можем взять из статистики наблюдений. Поэтому (8) мы будем также

называть *априорной информацией* относительно ситуации $p_1 \ldots p_m$, содержащейся в $q_1 \ldots q_m$.

Если теперь мы перейдем к использованию правильных весов $\log p_i$, то мы придем к значению (5). Следовательно, мы улучшим значение нашей меры на величину разности между (5) и (8), то есть на

$$H(p/q) = \sum_{i=1}^{m} p_i \log p_i - \sum_{i=1}^{m} p_i \log q_i. \tag{9}$$

Мы назовем это выражение *дифференциальной информацией*, информацией относительно $q_1 \ldots q_m$, содержащейся в $p_1 \ldots p_m$. Константа $\log m$ здесь исчезает. Можно показать, что (9) всегда $\geqslant 0$. Эта величина достигает своего минимального значения, равного 0, если $p_i = q_i$. Если q_1 — наименьшая величина среди q_i, то (9) достигает своего максимума при $p_1 = 1$; этот максимум равен $\log q_1$.

Если все q_i равны $\frac{1}{m}$, то выражение (9) становится тождественным (5). Следовательно, обычная информация (5) может рассматриваться как дифференциальная информация по отношению к неопределенной ситуации, то есть мы имеем

$$H(p) = H\left(p \Big/ \frac{1}{m}\right). \tag{10}$$

Это показывает, что на самом деле мы всегда измеряем разности в значениях информации и что, следовательно, для выражения информации наиболее подходит общая формула (9).

График формулы (9) для $m = 2$ изображен на рис. 27. Здесь мы принимаем

$$\begin{aligned} H(p/q) &= H_1 - H_2, \\ H_1 &= p \log p + (1-p) \log (1-p), \\ H_2 &= p \log q + (1-p) \log (1-q). \end{aligned} \tag{11}$$

В нижней части графика проведены две кривые H_1 и H_2 как функции p, которые везде $\geqslant 0$. H_2 — прямая линия, касающаяся H_1 в точке $p = q$. Легко видеть, что их разность $H_1 - H_2$ везде больше или равна 0. Эта кривая, представляющая $H(p/q)$, проведена в верхней части графика. Например, ордината X_2X_1 равна расстоянию Y_2Y_1.

Формула (5) представляет особый случай, когда прямая проходит горизонтально и касается кривой H_1 в самой нижней ее точке. Для $m > 2$ график, по существу, такой

Рис. 27. Дифференциальная информация $H(p/q)$ для двухчленной дизъюнкции, связываемой с вероятностями p и $1-p$; q и $1-q$ априорные вероятности. См. формулы (9) и (11). Минимальное значение $H=0$ достигается при $p=q$.

же, только H_2 в этом случае представляет касательную плоскость к поверхности H_1.

По аналогии с (7) мы можем доказать, что $H(p/q)$ аддитивна, если вероятности можно перемножать. Доказательство дается по аналогии с (6), если мы напи-

шем (9) для сочетаний $p_i p'_k$ и $q_i q'_k$ и соответствующим образом изменим суммирование. Таким образом, мы находим

$$H(p.\ p'/q.\ q') = H(p/q) + H(p'/q'). \qquad (12)$$

Это соотношение показывает, что дифференциальная информация аддитивна для независимых событий.

Для тех, кто знаком с кинетической теорией газов, формула (5) выглядит знакомой. Действительно, мы уже выводили эту формулу в (8.15), применяя идеи Больцмана к вопросу о распределении молекул газа в контейнере. Там мы видели, что H представляет просто энтропию, взятую с отрицательным знаком, и равна отрицательному значению удельной энтропии S, вычисляемой по формуле (8.11). То есть мы имеем

$$H = -S. \qquad (13)$$

И наоборот, мы можем, следовательно, называть энтропию обратной информацией. Действительно, энтропию можно рассматривать как экстенсиональную информацию. И именно принимая во внимание эту зависимость, выбрали форму (4) для весов, определяющих меру информации. Чтобы более подробно исследовать связь между энтропией и информацией, рассмотрим следующий пример.

Предположим, что газ находится в частично упорядоченном состоянии, то есть в неравновесном состоянии. Тогда энтропия газа имеет некоторое среднее значение. Распределение определено, если мы знаем, сколько молекул содержится в каждой ячейке, то есть когда известны вероятности p_i, входящие в (8.10). Будем считать, что молекулы пронумерованы по отдельности, хотя порядок этих номеров, или индивидуальных обозначений молекул, не имеет никакого отношения к существующей конфигурации.

Теперь представим человека, которому известно данное распределение и которого попросили, не глядя в контейнер, сказать, в какой ячейке находится в настоящий момент молекула № 1. Его информация тогда состоит в том, что ему известна вероятностная дизъюнкция $p_1 \ldots p_m$, и он сделает условное предсказание в описанном выше виде. Если предсказание повторяется для молекул № 2, № 3 и т. д., то его мера будет определяться (5). Следовательно, если p_i сильно различаются между собою и энтропия низка, тогда он добьется большого

успеха, то есть его информация велика. Наоборот, если p_i практически равны между собой и энтропия высока, то его успехи будут малыми, его информация незначительна. Энтропия увеличивается с уменьшением знаний о предсказываемом событии и находится в обратном отношении к интенсиональной информации. Однако если мы ничего не знаем *о* событии, то мы многое узнаем *из* события; следовательно, энтропия измеряет информацию *из* события, то есть экстенсиональную информацию.

Однако рассмотренная нами связь между энтропией и информацией подсказывает другую интерпретацию меры информации. Идентичность формул (5) и (8.15) показывает, что соотношению (5) может быть придано иное значение, чем то, которым мы пользовались при его выводе. Допустим, что вероятностная ситуация $p_1 \ldots p_m$ повторяется n раз так, что каждый раз происходит некоторое событие. В результате получается последовательность из n событий, для которых частоты событий $B_1 \ldots B_m$ определяются p_i; такова информация, которую нам дает вероятностная дизъюнкция $p_1 \ldots p_m$. Но мы не знаем, в каком порядке произойдут эти n событий. Существует целый класс конфигураций из n элементов последовательности, такой, что каждая из этих конфигураций удовлетворяет частотным условиям, налагаемым на последовательность вероятностями $p_1 \ldots p_m$. Число конфигураций N_D в этом классе находится перестановкой элементов, по теореме Бернулли; следовательно, это число N_D вычисляется по формуле (8.5), а его логарифм, даваемый формулой (8.7), тождественен с отрицательным значением суммируемого члена в (5). При возрастании меры информации число N_D уменьшается, тогда как при увеличении отрицательной информации, или энтропии, число N_D увеличивается. Следовательно, энтропия представляет собой меру экстенсивности класса возможных конфигураций; в форме абсолютной энтропии она представляет абсолютную меру, а в форме относительной энтропии — относительную меру.

Вероятностная дизъюнкция сообщает нам, что будет иметь место одна из конфигураций этого класса, но не говорит, какая из них; следовательно, она сообщает нам о многом, когда класс невелик, и говорит о немногом, когда класс большой, оставляя открытыми многочислен-

ные возможности. Таким образом, интенсиональность информации велика, когда экстенсиональность класса конфигураций мала. Когда мы определяем информацию как отрицательную энтропию, мы просто следуем закону взаимно обратного соотношения между объемом и содержанием, известному из традиционной логики. Формула (5) выглядит приемлемой мерой интенсиональной информации, даваемой вероятностной дизъюнкцией, так как ее отрицательное значение определяет экстенсиональность возможных конфигураций, совместимых с этой дизъюнкцией; и, таким образом, это отрицательное значение — энтропия — справедливо рассматривается как экстенсиональная информация.

Эта связь между энтропией и информацией справедлива всегда, даже в тех случаях, когда энтропии нельзя придать простую форму (8.11). Энтропия определяет экстенсиональность класса с определенным порядком, как разъяснялось в конце § 9. Низкая энтропия всегда определяет малый класс конфигурацией, а высокая энтропия — большой класс. Используя закон взаимосвязи объема и содержания, мы, следовательно, будем приписывать высокое значение интенсиональной информации состоянию с низкой энтропией. Соответствующая вероятностная дизъюнкция, то есть система вероятностей с этой мерой информации, задается не теми вероятностями, которые определяют энтропию, то есть не вероятностью W из (8.11), а частотами, с которыми молекулы встречаются в различных ячейках рассматриваемого пространства, то есть вероятностями p_i в (8.11).

Общей мере дифференциальной информации (9) может быть дано истолкование, подобное истолкованию (5). Формула (9) получается, когда ячейки в пространстве газа не обладают равными априорными вероятностями, но когда эти априорные вероятности равны q_i. Формула (8,4) тогда заменяется соотношением

$$W = N_D \cdot q_1^{n_1} \ldots q_m^{n_m}. \tag{14}$$

Применяя (8.7) к n_i из этой формулы и используя (8.10), мы легко выводим

$$S = \frac{1}{n} \log W = -\sum_{i=1}^{m} p_i \log p_i + \sum_{i=1}^{m} p_i \log q_i. \tag{15}$$

Это выражение равно (9), взятому с отрицательным знаком. Мера дифференциальной информации соответствует обобщенной форме энтропии, когда ячейки пространства параметров не обладают равными априорными вероятностями. Хотя такие представления об энтропии и не используются в микростатистике, они могут играть важную роль в макростатистике.

Различие между относительной и абсолютной энтропией, рассмотренное в § 8, находит свою аналогию в различии между относительной и абсолютной информацией. Мера информации (5) представляет относительную меру, так как она тождественна с отрицательным значением (8.11). Когда опускают константу $\log m$ в (5), то это соответствует опусканию константы $-\log m$ в (8. 11) и, таким образом, использованию абсолютной энтропии, которая определяется непосредственно логарифмом числа конфигураций, то есть выражением (8.7). Эта абсолютная форма экстенсиональной информации использовалась К. Шенноном [1]. И именно вследствие опускания этой константы энтропия для равновероятной дизъюнкции возрастает с увеличением числа членов дизъюнкции m. Тогда соответствующая интенсиональная информация будет уменьшаться. Наоборот, если используется форма (5), то интенсиональная информация одинакова для любого значения m равновероятной дизъюнкции, а именно равна 0. Максимум интенсиональной информации, получающийся, когда одна из вероятностей равна 1, изменяется противоположным образом: он не изменяет своего значения при любых m, когда эта константа опущена, но возрастает с ростом m, когда константа учитывается. Не существует априорного основания для предпочтения одного способа определения другому. Уменьшение интенсиональной информации при увеличении m в равновероятных дизъюнкциях, связанное с абсолютной информацией, оказывается адекватным, когда дизъюнкция $B_1 \vee \ldots \vee B_m$ расширяется добавлением новых членов, например добавлением ячеек равных размеров в рассматриваемом пространстве. Однако когда увеличение дизъюнкции вызывается переходом к меньшим по размеру ячейкам при сохранении полного объема пространства, то имеются все основания считать, что

[1] См. Shannon and Weaver, ibid., p. 18—22.

информация для равновероятных ячеек остается неизменной. Последнему условию удовлетворяет относительная информация. И тогда вполне допустимо, что возрастает максимальное значение интенсиональной информации.

Можно добавить, что в теории информации не существует аналогии парадоксу Гиббса (см. § 8). Формулы (7) и (12), в которых информация аддитивна, соответствуют случаю, когда энтропия вычисляется для двух контейнеров, расположенных рядом, но при условии, что между контейнерами не открывается никакого окошка. В таких условиях энтропия также аддитивна. Открытие окошка представляет физический процесс, который приводит к смешиванию газов и для одинаковых газов приводит к парадоксу Гиббса. Этот процесс не имеет аналогии в теории информации. По этой причине представление об абсолютно нормированой мере информации может быть проведено последовательно до конца. Однако для целей настоящего исследования оказывается более удобной относительно нормированная информация (5).

Формулирование определений, рассмотренных в этом параграфе, представляет собой решение типичной проблемы экспликации: проблемы замены неопределенного в повседневном языке термина «информация» точным термином, который соответствует его использованию. Когда мы здесь говорим о соответствии употреблению, мы очень хорошо знаем, что можно ожидать только качественного соответствия. В конце § 2 указывалось, что большего нельзя достигнуть по той самой причине, что понятие повседневного языка, или экспликанд, неопределенно и мы никогда не можем быть уверенными в том, что нашли его точное значение. Строго определенное понятие, или экспликанс, следует рассматривать в качестве программы для нового использования, при котором исчезает неопределенность, присущая экспликанду, пока строго, насколько это возможно, придерживаются его значения.

Надо считать замечательным достижением, что удалось дать такое истолкование настолько абстрактному и неопределенному понятию, как информация. Я вполне могу представить, что многие философы старой школы стали бы настаивать, что невозможно измерить инфор-

мацию, что это одно из многих понятий, которые не поддаются измерению и которыми весьма гордится склонная к мистицизму философия. Однако математики, не сдерживаемые философскими сомнениями, нашли смелость приступить к работе и предложили меру информации, которая удовлетворяет нуждам техники связи и которая в то же время дает философам инструмент для логического исследования.

Установление связи между теорией информации и статистической механикой — главный результат этих исследований; вместе с тем установление самой этой связи ставит перед нами новый и важный вопрос. Обратная связь между энтропией и информацией правильно подчеркивалась теми математиками, которым мы обязаны основанием и развитием теории информации и среди которых выдающееся место занимают имена Неймана, Винера и Шеннона. Тем не менее было доказано только то, что энтропия в теории газов соответствует природе отрицательной информации. Можно ли заключить, что, наоборот, информация теории информации соответствует природе отрицательной энтропии? То есть, можно ли макроэнтропию, определенную как отрицательное значение интенсиональной информации (5) и, таким образом, как экстенсиональную информацию, рассматривать в качестве индикатора направления времени? Несомненно, в теории информации учитывается, что определенная в ней энтропия тесно связана со вторым законом термодинамики. Но доказательство того, что это верно и что энтропия теории информации действительно определяет направление времени, может быть дано только при условии использования результатов, полученных для протоколирующих процессов из гипотезы о ветвящейся структуре. В следующем параграфе и будет дано такое доказательство.

§ 21. Направление времени в учении об информации и теория регистрирующих приборов

Чтобы показать, что информационные процессы указывают направление времени, мы покажем, что такие процессы основываются на принципе общей причины.

Сначала покажем обратное, что принцип общей причины можно перенести в теорию информации. Мы ис-

пользуем следующие обозначения:

$$P(A.\ B)=p_1,\ P(A.\ \bar{B})=p_2,\ P(\bar{A}.\ B)=p_3,$$
$$P(\bar{A}.\ \bar{B})=p_4;\ P(A)\cdot P(B)=q_1,\ P(A)\cdot P(\bar{B})=q_2, \qquad (1)$$
$$P(\bar{A})\cdot P(B)=q_3,\ P(\bar{A})\cdot P(\bar{B})=q_4.$$

Соотношение (19.1), которое означает, что в случае конъюнкции $A.B$ частота обладает бо́льшим значением, чем это следовало бы ожидать для случайных совпадений, можно в данном случае, если мы используем формулу дифференциальной информации (20.9) для $m=4$, выразить следующим образом:

$$H(p/q)=\sum_{i=1}^{m} p_i \log p_i-\sum_{i=1}^{m} p_i \log q_i>0. \qquad (2)$$

Это означает: отклонение наблюдаемых частот от случайного распределения дает нам информацию, которая больше нуля. Отметим, что эта формулировка носит нейтральный характер, при котором конъюнкция $A.B$ занимает равноправное положение наряду с тремя другими возможными конъюнкциями.

Используя связь между информацией и энтропией, сформулированную в (20.13), можно также сказать, что (2) определяет отрицательную макроэнтропию процесса, при котором происходят совпадения A и B. Большое количество информации означает, что макроэнтропия низка. Следовательно, необычные совпадения указывают на низкую энтропию.

Мы знаем, что состояние с низкой энтропией возникает в результате взаимодействия в прошлом; и мы видели, что то же самое справедливо и для соотношения (19.1). Мы делаем заключение, что информация возникает в результате прошедшего взаимодействия, то есть, что только состояния после взаимодействия могут доставлять информацию.

Такой вывод можно сделать по крайней мере в тех случаях, когда информация заключается в отклонении частоты событий от закона случайного распределения. В § 18 мы рассмотрели протоколы в природе, такие, как отпечатки на песке и окаменелости в почве. Все они одинаковой природы: случай бы не создал этих явлений,

которые мы справедливо истолковываем как результат действия причин в прошлом. Так же обстоит дело и в отношении искусственно созданных протоколов. Когда мы пишем на бумаге, записываем музыку на граммофонную пластинку или храним информацию в «памяти» электронной вычислительной машины, в качестве результата возникают особые конфигурации элементов, как чернильные следы на бумаге, бороздки на восковых дисках, электрические заряды в миниатюрных конденсаторах, которые настолько высокоупорядочены, что не представляется возможным их интерпретировать как возникшие в результате действия случая. И именно поэтому они являются протоколами и доставляют информацию. Вся информация есть протокол о прошлом. Она также может информировать нас и о будущем, но это возможно только в том случае, если отчет о прошлом мы используем в качестве базиса для индуктивных выводов. Однако, поскольку в данном контексте мы интересуемся вопросом об источнике информации, мы не будем входить в рассмотрение вопроса об использовании информации. В любом случае источник информации всегда лежит в прошлом.

Чтобы изучить соответствующий пример, рассмотрим логическую схему действия регистрирующих приборов. Они используются для регистрации наблюдаемых, которые изменяются нерегулярным образом и которые не могут быть предсказаны. Например, барограф записывает на вращающийся барабан нерегулярно изменяющиеся значения атмосферного давления. Зигзагообразная линия, которую вычерчивает перо этого прибора на полоске бумаги, содержит запись этих нерегулярных изменений.

Мы никогда не смогли бы предсказать изменения величины атмосферного давления, но записать их легко. Мы сталкиваемся здесь с ярким различием между прошлым и будущим о чем упоминалось ранее в § 2. В отношении нерегулярно изменяющихся физических величин *предсказание* оказывается технически невозможным, в то время как *послесказание* легко достигается путем использования регистрирующих приборов. Различие можно объяснить на основе соотношений, установленных для конъюнктивных вилок; и, таким образом, теперь можно объяснить определение 5, § 2.

если существует общая причина, связывающая A и B. На это указывалось в §§ 2 и 20. Теперь мы дадим более подробную иллюстрацию. Если в некотором месте измерение показало низкое атмосферное давление, мы можем сделать вывод с некоторой вероятностью, что на большой площади пройдет дождь. Этот вывод представлен схемой, приведенной на рис. 28. На этой схеме A — показание барографа, C — атмосферное давление в том месте,

Рис. 28. Сочетание открытой вилки и вилки, закрытой с обеих сторон. Наблюдение протокола A о событии C позволяет предсказать более позднее событие E.

где находится прибор, B — некоторые другие следствия этого местного давления, C_1 — атмосферное давление на более обширной площади, B_1 — другие следствия этого давления и E — дождь. Вилка ACB открыта в будущее, но вилка CEB_1 закрыта с обеих сторон, и, следовательно, частную причину C можно использовать, чтобы получить вывод относительно полного следствия E. Такой вывод в отношении будущего становится возможным, так сказать, благодаря опосредованию через прошлое[1].

[1] Это рассмотрение регистрирующих процессов и вывода полной причины на основе частного следствия, вывода, который не имеет аналогии в выводе полного следствия на основе частной причины, было дано в моей ранней статье, упомянутой в примечании к § 19; см. Hans Reichenbach, Die Kausalstruktur der Welt und der Unterschied von Vergangenheit und Zukunft, S. 133—175.

Так как вилка ACB на рис. 28 открыта с одной стороны, она должна быть направлена в прошлое. Сама запись A — чернильные знаки на бумаге — не приводит, конечно, к дождю. Но она позволяет нам сделать вывод в отношении причины C, хотя и представляет только ее частное следствие.

Напротив, в случае, когда вилка направлена в будущее и открыта в прошлое, нам нужно знать оба события A и B, если мы хотим предсказать общее следствие E с высокой вероятностью. Это можно проиллюстрировать примером, рассмотренным в § 19: когда мы видим приближающиеся друг к другу два грузовика, мы можем с высокой вероятностью предсказать, что водители обменяются приветствиями посредством световых сигналов; но мы не сможем сделать такое предсказание, если мы видим только один движущийся грузовик. И именно на основании подобных причин обычно требуется большое число данных наблюдения, чтобы делать предсказания. Вообще говоря, мы имеем здесь многократные вилки, в которых общее следствие E определяется рядом одновременных событий $A_1 \ldots A_n$. Конечно, результаты, полученные для двойных вилок, легко можно распространить на случай многократных вилок. Например, в § 2 отмечалось, что для предсказания атмосферного давления необходимы данные, полученные в обширной географической области. Это необходимо потому, что по отношению к метеорологическим событиям мы не имеем вилок, замкнутых в прошлое. То же самое относится к предсказаниям политических событий и войны. В этом случае мы должны делать предсказания о следствии, исходя из полной причины; это объясняет, почему столь трудно делать предсказания в метеорологии и в политике.

Математическое выражение информации на основе представлений об измерительном приборе можно дать следующим образом. Для упрощения проблемы допустим, что показания прибора изменяются не непрерывно, а принимают только r различных значений $A_1 \ldots A_r$. Соответственно этому наблюдаемая величина, например давление воздуха, может принимать r возможных состояний $B_1 \ldots B_r$. Мы можем отождествить значение наблюдаемой величины, полученное в акте регистрации, с ее значением вскоре после этого акта так, чтобы

можно было интерпретировать B_i как события, совпадающие с A_i в смысле принципа общей причины. Мы полагаем

$$P(A_i)=a_i,\ P(B_i)=b_i,\ P(A_i \cdot B_k)=p_{ik}. \tag{3}$$

Если прибор совершенен, то есть его показания всегда правильны, то мы наблюдаем только пары $A_i \cdot B_i$; следовательно, мы имеем $p_{ik}=0$ для $i \neq k$ и $p_{ii}=a_i=b_i$. Значения a_i, b_i и p_{ii} можно установить из статистически обработанных показаний прибора в течение продолжительного периода времени, когда мы подсчитываем относительные частоты состояний A_i, B_i и A_iB_i для каждого i. Следовательно, совершенный регистрирующий прибор можно охарактеризовать следующими соотношениями, написанными по аналогии с (1):

$$\begin{gathered} p_{ik}=0 \text{ для } i \neq k,\ p_{ii}=a_i=b_i; \\ q_{ik}=a_ib_k,\ q_{ii}=a_i^2. \end{gathered} \tag{4}$$

Если мы хотим найти дифференциальную информацию, определенную посредством (2), можно использовать двойное суммирование, что практически эквивалентно подсчету всех значений p_{ik} в сериях, состоящих из $m=r^2$ членов. Таким образом, мы находим

$$\begin{aligned} H(p/q) &= \sum_{i=1}^{r}\sum_{k=1}^{r} p_{ik}\log p_{ik} - \sum_{i=1}^{r}\sum_{k=1}^{r} p_{ik}\log q_{ik} \\ &= \sum_{i=1}^{r} a_i \log a_i - \sum_{i=1}^{r} a_i \log a_i^2 \\ &= -\sum_{i}^{r} a_i \log a_i = H^*(a), \end{aligned} \tag{5}$$

$$H^*(a) = -[H(a) - \log r]. \tag{6}$$

Выражение в последней строчке соотношения (5) получается, если мы положим $2\log a_i$ вместо $\log a_i^2$. Формула (6) означает, что *интенсиональная* информация H, выраженная через p_{ik} в отношении q_{ik}, равна *экстенсиональной* информации H^*, выраженной через a_i, если последняя измеряется абсолютно, то есть без аддитивной постоянной $\log r$, содержащейся в $H(a)$. Этот результат справедлив только для совершенного регистрирующего прибора, определенного соотношениями (4).

Он объясняет, почему в качестве меры информации в записывающих устройствах, подобных тем, которые изучал К. Шеннон, используется выражение вида $H^*(a)$ [1]. Однако выражение вида $H^*(a)$ явным образом не говорит о том, что оно выражает собой степень возрастания частоты совпадений A_i и B_i в сравнении с той, которая соответствует их независимому наступлению.

Если процесс регистрации повторяется n раз, то мы можем, согласно (20.12), найти сумму отдельных значений H, определяемых выражением (5), и, таким образом, будем иметь

$$H_n = n \cdot H \tag{7}$$

для полной информации H_n, которую сообщают n последовательных актов регистрации. Согласно этому результату, не имеет значения содержание индивидуального зарегистрированного показания прибора, так как информация определяется нейтральным путем для всей дизъюнкции в целом. Эта аддитивность имеет место только для независимых значений наблюдаемой величины, но обычно это условие удовлетворяется. И, таким образом, возрастание информации с увеличением числа зарегистрированных данных находит свое естественное выражение.

В статистической интерпретации эта аддитивность информации означает: чем больше число зарегистрированных показаний прибора, тем меньше вероятность того, что эти показания носят случайный характер, и выше вероятность того, что они представляют протокол прошлого. Довольно удивительно, что это приводит к заключению, согласно которому увеличение порядка указывает положительное направление времени, в противоположность заключению, основанному на представлениях об энтропии, и согласно которому положительное направление времени порождает тенденцию к беспорядку. Но это кажущееся противоречие легко разрешается. Порядок в зарегистрированных показаниях прибора не представляет последовательность состояний изолированной системы, но вытекает из пространственного

---

[1] См. Claude E. Shannon, Warren Weaver, The Mathematical Theory of Communication, p. 19—21. Шеннон записывает ее как $H(a)$.

ансамбля индивидуальных состояний взаимодействия, каждое из которых производит особое зарегистрированное показание прибора. И порядок в пространственном ансамбле указывает направление времени, перпендикулярное ансамблю. Таким образом, мы имеем здесь временну́ю последовательность, порядок которой или информационное содержание возрастает в направлении положительного времени, как это выражено в (7).

И именно как следствие этого факта вместе с ростом информации все более четко проявляет себя направление времени. Эта особенная черта регистрирующих приборов делается весьма ясной, если мы опишем процесс регистрации в противоположном направлении времени. На языке L_2 (см. § 16) действие барографа описывалось бы следующим образом. Барабан, на котором вычерчена зигзагообразная кривая, вращается в обратном направлении, в то время как рычаг с пером движется вверх и вниз под действием атмосферного давления. Это движение обусловливается теми состояниями давления воздуха, которые, в L_1, следуют непосредственно за некоторым определенным положением пера; так как в этом случае давление воздуха практически одинаково с тем, которое в L_1 предшествует положениям рычага, то здесь мы не имеем ничего необычного. Однако мы наблюдаем сверхъестественные совпадения в изменении давления воздуха и формы зигзагообразной кривой: перо всегда оказывается строго на том месте бумаги, где проходит кривая. Следовательно, кривая выражает собой предсказание давления воздуха. И наконец чернила с бумаги втекают в перо и чернильные знаки исчезают.

Последний процесс необычен, потому что вытекание чернил на бумагу и их высыхание образуют необратимые процессы. Но не только этот факт делает необычной всю ситуацию. Мы должны изучить постоянное совпадение кривой и положения пера. Фактически мы могли бы заменить сам процесс написания пером некоторым обратимым процессом, тогда осталось бы еще неправдоподобным постоянное совпадение кривой и положения пера.

Регистрирующий прибор, использующий обратимый процесс для записи наблюдений, можно создать следующим образом. Для описания мы используем язык L_1. На движущейся ленте установлены небольшие валики,

вокруг которых могут вращаться небольшие зубчатые колесики (рис. 29). Установленный за пределами ленты барометр-анероид управляет остроконечным рычагом, который может двигаться вверх и вниз, и таким путем он поднимается или опускается непосредственно

Рис. 29. Самопишущий барометр, основанный на использовании обратимых процессов.

Анероид подымает или опускает остроконечный рычаг. Столбики движутся вместе с лентой, и колесики, приходя в соприкосновение с остроконечным рычагом, приводятся во вращение. Распределение вращающихся колесиков регистрирует атмосферное давление. Будучи истолкованным на основе понятия обратного направления времени, этот процесс регистрации не противоречит физическим законам, но приводит к невероятным совпадениям.

под действием давления воздуха. Когда этот указательный рычаг достигает уровня одного из зубчатых колесиков, оно начинает вращаться вследствие движения ленты. Колесико будет вращаться неограниченно долго, если мы сделаем трение минимальным. Хотя лента движется медленно влево, на каждом стержне все время вращается одно из колесиков, и, таким образом, по высоте вращающегося колесика записывается значение

атмосферного давления. Другими словами, иррегулярные изменения атмосферного давления регистрируются упорядоченными состояниями колесиков, и это распределение вращающихся колесиков предполагает ту же самую функцию, что и зигзагообразная кривая обычного барографа.

Это распределение хранит информацию. Она может быть измерена посредством (5), если под $A_1 \ldots A_r$ мы понимаем различные положения вращающегося колесика на столбце и определяем каждое a_i путем подсчета относительной частоты состояний $A_1 \ldots A_r$, при условии, что прибор использовался в течение продолжительного периода времени. Если мы рассматриваем вращающееся колесико и соответствующее давление воздуха как одну совместную физическую систему, то макроэнтропия этой системы задается отрицательным значением (5) с точностью до аддитивной постоянной. Каждый акт регистрации представляет такую систему. Мы можем также рассматривать ряд таких систем как образующий одну обширную систему. Ее макроэнтропия определяется взятой с отрицательным знаком суммой величин H, вычисленных из (5), то есть величиной $-H_a$, как она определена в (7). Самое низкое значение энтропии этой системы показывает, что система находится в весьма упорядоченном состоянии, то есть, что она выражает большую информацию. Таким образом, в данном случае мы имеем всецело построенный при помощи обратимых процессов механизм, который указывает направление времени посредством низкого значения его макроэнтропии, точно так же как низкое значение микроэнтропии указывает на существование в предыдущем взаимодействия. Порядок вращающихся колесиков позволяет раскрыть направление времени благодаря принципу, посредством которого мы делаем вывод, что горящая сигарета незадолго до этого должна была быть зажжена.

Наличие направления в этом механизме становится особенно ясным, если мы описываем эти процессы на языке L_2. Говоря на языке, соответствующем отрицательному направлению времени, мы наблюдаем следующие совпадения. Лента движется слева направо, на каждом из расположенных на ленте столбцов одно из колесиков вращается. Высота этих вращающихся колесиков изменяется иррегулярно. Когда столбец проходит мимо

рычага анероида, этот рычаг всегда оказывается строго на высоте вращающегося колесика и останавливает его вращение.

Это прекращение вращательного движения представляет изменение направления процесса, который в случае L_1 мы назвали воздействием, вызывающим движение колесика. Так как это воздействие представляет механический процесс, его можно считать обратимым, если мы пренебрегаем трением и рассматриваем воздействие как абсолютно упругое. Следовательно, описание этого процесса в L_2 не противоречит физическим законам. Однако в L_2 мы постоянно наблюдаем совпадение в положениях рычага и приближающегося вращающегося колесика. И именно этот ряд крайне маловероятных совпадений нуждается в своем объяснении. В случае L_2 такое объяснение могло бы быть дано только на основе учёта общего следствия: цель этих событий, по-видимому, состоит в том, чтобы прекратить вращательное движение колесиков.

Напротив, объяснение в L_1 обычно учитывает общую причину. То, что высота вращающегося колесика вскоре после взаимодействия с рычагом еще соответствует высоте рычага, является совпадением, которое объясняется тем фактом, что рычаг вызывает вращение колесика. Хотя при измерениях на этом регистрирующем приборе и используются обратимые процессы, он приводит к определению направления времени на основе принципа общей причины и, таким образом, определяет направление времени с точки зрения макростатистики.

Эти рассуждения показывают, почему регистрирующие приборы могут регистрировать только прошлое, а не будущее и почему информация определяет то же самое направление времени, что и возрастающая энтропия. В процессе регистрации информации обнаруживается однонаправленный характер времени потому, что только причина в прошлом может установить соответствие между знаками и физическими данными, и потому, что возрастание количества информации характеризует порядок в пространственном ансамбле, что требует объяснения на основе причины в прошлом. Направление времени, определяемое информацией, есть направление, которое характеризует вселенную благодаря ее статистической изотропности.

§ 22. Чисто макростатистическое определение направления времени

Подведем итоги, к которым мы пришли к настоящему моменту в наших исследованиях времени. Сначала мы определили на основе механики порядок времени. Затем мы определили направление времени на основе термодинамики, или микростатистики. Наконец, мы применили методы микростатистики к макростатистике и показали, что направление времени можно определить на основе макростатистики. Это последнее определение просто повторяет соотношения, сформулированные нами в виде гипотезы о ветвящейся структуре и относящиеся как к микростатистике, так и к макростатистике, включая соотношения, контролирующие накопление информации.

Теперь мы покажем, что, используя макростатистику, можно определить также порядок времени. Кроме того, мы также покажем, что эта статистика может заменить собой классическую механику при построении причинной сети, обладающей линейным порядком. Затем уже легко сделать переход от линейной упорядоченности времени к направлению времени, используя принцип общей причины, как указывалось в § 19. Прежде чем обратиться к построению причинной сети, которую можно проиллюстрировать рис. 5, необходимо сделать некоторые предварительные замечания относительно наблюдаемых величин, предполагающихся при этом построении.

Мы предположим, что события x, y, ... наблюдаются как отдельные пространственно-временные элементы. Чтобы составить для них причинную сеть, исходя только из статистических соотношений, мы предположим, что возможна классификация событий x, y, ... по определенным классам A, B, ... Эти классы, конечно, определяем мы сами, но мы наложим определенные условия на эти определения.

Мы скажем, что класс A *соопределен*, если принадлежность события x к классу A можно определить непосредственно при наступлении события x. Другими словами, наблюдая x, мы можем определить, принадлежит ли x к A, и для этого нам нет необходимости знать, произошли ли некоторые другие события y, z,.. раньше, или позже, или одновременно при значительном удале-

нии. Например, термин «зеленый», характеризующий принадлежность к определенному классу, соопределен, потому что, взглянув на x, мы можем сказать, зелен ли он. Однако выражение «день, следующий за дождливым днем», характеризующее принадлежность к классу, не соопределено; чтобы узнать, относится ли некоторый день к этому классу, необходимо знать, предшествовал ли ему дождливый день.

Теперь мы введем по отношению к рассматриваемой нами статистике требование, чтобы классы A, B, ... были соопределенными. Использование соопределенных классов позволяет делать сообщения о наблюдениях в выражениях о том, что непосредственно наблюдалось по отношению к некоторому отдельному пространственно-временно́му элементу.

Говоря на языке логики, понятие соопределенного класса есть предикат с одним переменным, который не является сокращенным предикатом многих переменных. Например, слово «женатый» представляет сокращенный термин, потому что выражение «x — женат» означает, что «существует лицо y, на котором x женат»[1]. В разговорном языке иногда не ясно, является ли термин элементарным или сокращенным. Такой неопределенности нет в нашем исследовании. Данные наблюдения достаточно определенны, чтобы позволить нам сказать, можно ли проверить на основе наблюдения одного только x утверждение «x принадлежит к A».

Теперь нам необходимо ввести требование по отношению к вероятностям, характеризующим определенные классы. Вероятностное выражение обладает характерным численным значением только в том случае, если его обусловливающий класс не является пустым[2]. В общем рассматриваемые классы будут образованы таким образом, что они не будут пустыми. Однако изредка еще допустимо использовать пустые обусловливающие классы при следующем условии: значения вероятностей, которые будут использоваться, получаются из системы d утверждений о данных наблюдения посредством системы r правил, включающих индуктивные и дедуктив-

---

[1] См. Hans Reichenbach, Elements of Symbolic Logic, New York, Macmillan Co., 1947, p. 122.

[2] ThP, p. 55.

ные выводы. Если к d мы добавим утверждение о непустоте обусловливающего класса, то мы потребуем, что образованная таким образом система d' будет совместима с физическими законами и не будет им противоречить и что в случае, когда правила r применяются к d', они приводят к тем же самым значениям вероятности. Когда это требование выполняется, мы говорим о *подлинных* вероятностях. Из самого определения этого понятия следует, что оно относится к вероятностям с непустыми обусловливающими классами[1]. Причинная сеть, связывающая события x, y, ... на основе представлений о классах A, B, ..., строится посредством причинного отношения «*между*». В статистике это отношение определяется как отношение между классами. По отношению к событиям это выглядит следующим образом. Предположим, что неоднократно встречаются тройки событий x, y, z и мы хотим определить причинное отношение «между» внутри каждой тройки. Это также предполагает, что мы можем сказать, какие события принадлежат к одной и той же тройке. События в тройке не слишком отдалены друг от друга, и мы можем грубо оценить пространственно-временные расстояния. События из разных троек непосредственно сравнивать мы не будем. Такое сравнение можно проводить только с помощью промежуточных событий, относящихся к другим классам. Теперь представим себе, что мы рассматриваем некоторую тройку и находим, что x принадлежит классу A_1, y — классу A_2, z — классу A_3. Если мы нашли, что A_2 в причинном отношении находится между A_1 и A_3, то тогда мы говорим, что y в причинном отношении находится между x и z. Это объяснение необходимо иметь в виду при рассмотрении дальнейших формулировок, в которых мы будем говорить только о классах.

Предположим, что, говоря на обычном языке, в котором используется представление о положительном направлении времени, мы имеем причинную цепь, представленную на рис. 30, в которой положительное направление времени указано стрелками. Если мы знаем, что x принадлежит A_1, то тогда с вероятностью $P(A_1, A_2)$ мы

[1] Ср. Hans Reichenbach, Nomological Statements and Admissible Operations, Amsterdam, North-Holland Publ. Co., 1954, Appendix, p. 128.

можем предсказать, что y будет принадлежать A_2. Сокращенно мы говорим: если нам известно, что наступило A_1, то с вероятностью $P(A_1, A_2)$ мы можем предсказать, что наступит A_2. Подобным же образом мы можем предсказать, что наступит A_3. Мы предположим, что вероятность предсказания не равна 1, но увеличивается, стремясь к 1, по мере того как мы приближаемся к рассматриваемому событию; например, исходя из A_2, мы предсказываем A_3 с большей вероятностью, чем это достижимо из A_1. Эта зависимость справедлива весьма широко для непрерывных вероятностных последовательностей [1].

Рис. 30. Причинная цепь, образованная тремя событиями.

Однако, как только мы узнаем, что A_2 наступило, то уже нет необходимости знать, что ему предшествовало A_1; событие A_1 уже не имеет отношения к предсказанию A_3. Вклад A_1 в предсказание A_3, так сказать, поглощен A_2; и можно сказать, что A_2 *экранирует* A_1 от A_3. Это будет справедливо по крайней мере, если определения классов достаточно подробны и учитывают наблюдаемые траектории. Предположение о том, что такие определения могут быть даны, можно рассматривать как формулировку принципа близкодействия. Теперь мы определяем:

Определение. Событие A_2 находится в отношении «причинно между» событиями A_1 и A_3, если выполняются соотношения

$$1 > P(A_2, A_3) > P(A_1, A_3) > P(A_3) > 0, \qquad (1)$$

$$1 > P(A_2, A_1) > P(A_3, A_1) > P(A_1) > 0, \qquad (2)$$

$$P(A_1 . A_2, A_3) = P(A_2, A_3). \qquad (3)$$

[1] ThP, p. 237.

Символически эту связь мы запишем в виде

$$\text{мжд}\ (A_1,\ A_2,\ A_3). \tag{4}$$

Выражение (4) представляет утверждение; оно говорит, что события A_1, A_2, A_3 удовлетворяют соотношениям (1)—(3).

Можно показать, что отношение *причинно между* (causally between) *симметрично* по отношению к A_1 и A_3, то есть, что исходя из (4) мы можем прийти к выражению

$$\text{мжд}\ A_3,\ A_2,\ A_1. \tag{5}$$

Во-первых, соотношения (1) и (2) просто переходят друг в друга, когда мы меняем местами A_1 и A_3. Во-вторых, мы используем общее соотношение, доказанное в теории вероятностей[1],

$$\frac{P(A_i.\ A_k,\ A_m)}{P(A_i.\ A_m,\ A_k)} = \frac{P(A_i,\ A_m)}{P(A_i,\ A_k)}. \tag{6}$$

Положив $i = 2$, $k = 1$, $m = 3$, мы из (3), используя (6), выводим

$$P(A_3.\ A_2,\ A_1) = P(A_2,\ A_1). \tag{7}$$

Этим симметрия доказывается. Отсюда следует, что соотношения (1)—(3) не определяют направления стрелок на рис. 30. Отношение «между» просто упорядочивает три события.

Первое неравенство в (1) и первое неравенство в (2) необходимы, потому что иначе соотношения (7) и (3) стали бы тривиальными[2]. Например, из $P(A_2, A_1) = 1$ мы могли бы немедленно вывести (3), потому что добавление A_1 в первый член вероятностного выражения $P(A_2, A_3)$ не изменяет значения $P(A_2, A_3)$, если $P(A_2, A_1) = 1$. Можно показать также, что $P(A_1, A_2) < 1$ и $P(A_3, A_2) < 1$. Если бы мы имели $P(A_1, A_2) = 1$, то мы могли бы, используя ту же теорему, сделать вывод, что $P(A_1, A_3) = P(A_1.\ A_2, A_3)$; если мы затем применим (3), мы обнаружим противоречие второму неравенству в (1). Подобный же вывод доказывает и второе соотношение. Следовательно, не требуется, чтобы эти два соотношения были строго сформулированы. Можно добавить еще одно замечание, что третье неравенство в (2) следует из

[1] Там же, стр. 112, уравнение (33).
[2] См. там же, стр. 117, уравнение (5).

третьего неравенства в (1) благодаря следующему соотношению, доказанному в теории вероятностей[1]:

$$\frac{P(A_i, A_k)}{P(A_k, A_i)} = \frac{P(A_k)}{P(A_i)}. \tag{8}$$

Если мы положим, что $i = 1$ и $k = 3$, то заключаем, что если левый числитель больше правого числителя, то то же самое должно быть справедливо и для знаменателей.

Кроме того, можно доказать, что для трех событий отношение *причинно «между» является уникальным*. Это означает, что среди трех событий только одно может находиться в причинном отношении между двумя другими. Для доказательства мы должны показать, что если справедливо (4), то два соотношения

$$\text{мжд}\,(A_3, A_1, A_2), \quad \text{мжд}\,(A_1, A_3, A_2), \tag{9a, 9b}$$

ложны. Ложность (9*a*) следует из (6), если мы положим $i = 1$, $k = 2$, $m = 3$, потому что тогда, согласно (3) и (1), числитель слева больше, чем числитель справа; следовательно, знаменатель слева должен быть больше знаменателя справа, и, таким образом, мы получаем

$$P(A_3, A_1, A_2) > P(A_1, A_2). \tag{10}$$

Это показывает, что уравнение, аналогичное (3), для последовательности A_3, A_1, A_2 не удовлетворяется. Ложность (9*b*) подобным же образом следует из (6), если положим $i = 3$, $k = 2$, $m = 1$ и используем (7) и (2); таким образом, мы получаем

$$P(A_1, A_3, A_2) > P(A_3, A_2). \tag{11}$$

Это соотношение противоречит аналогичному соотношению (3) для последовательности A_1, A_3, A_2.

Благодаря этим свойствам отношение *причинно между* можно использовать для построения причинной сети. Строгая детерминация следствия A_3 его причиной A_1, или A_2, которая сформулирована в уравнениях механики и которую мы использовали в § 5 для определения отношения «между», заменяется здесь статистической детерминацией. Статистическая детерминация ослабевает с возрастанием расстояния от причины. Этот факт выражен в соотношениях (1)—(2). Промежуточное событие экранирует своих предшественников от следствия, что

[1] Там же, стр. 112, уравнение (32).

сформулировано в соотношении (3). Следовательно, совместное использование соотношений (1)—(3) приводит к упорядоченным отношениям, выполняющимся для причинной сети.

Следует заметить, что применимость соотношения экранирования (3) зависит от использования соопределенных классов. Соотношение (3) просто удовлетворяется и при использовании других способов определения классов. Например, если A_2 мы определим как класс событий, за которым практически всегда следует A_3, мы получаем $P(A_2, A_3) \sim 1$. Добавление A_1 в обусловливающий член вероятности тогда не внесло бы заметного отличия и (3) легко выполнялось бы. Если статистические соотношения служат тому, чтобы информировать нас о причинном порядке, они должны относиться к классификации событий, в которой события рассматриваются как индивидуумы и их отношения к другим событиям не отражаются в самой формулировке классов.

Если нам даны четыре события, то устанавливаемые для них отношения «между» позволяют провести различие между такими конфигурациями, какие показаны на рис. 31 и 32. Например, если выполняются следующие соотношения:

$$\text{межд}\ (A_1, A_2, A_3),\quad \text{межд}\ (A_1, A_2, A_4),\quad \text{межд}\ (A_4, A_2, A_3), \tag{12}$$

то мы заключаем, что эти четыре события обладают порядком, изображенным на рис. 32, и не обладают порядком, соответствующим рис. 31.

Теперь нам необходимо подчеркнуть ограничения, свойственные рассмотренному отношению «между». Первое ограничение состоит в том, что отношение «между», сформулированное в (1)—(3), не является транзитивным, то есть, что мы не можем доказать утверждение

$$\text{межд}\ (A_1, A_2, A_3) \cdot \text{межд}\ (A_2, A_3, A_4) \supset \text{межд}\ (A_1, A_2, A_4). \tag{13}$$

Подковообразный знак означает «содержит в себе». Это утверждение может быть справедливо в особых случаях, но в его справедливости нет необходимости. Следовательно, отношение «между» является не только не транзитивным, но антитранзитивным. Однако транзитивность едва ли необходима для рассматриваемого отношения

«между», которое используется для построения сети, а не направленного порядка в последовательности. Наличие разветвлений может привести к нарушению неравенств (1)—(2). Например, в случае представленной на рис. 23 конфигурации предсказание E на основе C более надежно, чем предсказание E на основе B,

Рис. 31 (слева) и рис. 32 (справа). Две конфигурации событий, отличающиеся своими отношениями причинно между.

хотя геометрически B расположено между C и E — по линии CBE. Причина этого в том, что E представляет общее следствие B и A; и вполне возможно предсказать A, исходя из C, с более высокой вероятностью, чем исходя из B.

Следовательно, отношение *причинно между* просто соединяет события в сеть, но не устанавливает направленного порядка в последовательности. Его можно назвать *отношением соседства;* но оно не является *отношением «между» в последовательности.*

Второе ограничение отношения *причинно между* касается относительного направления причинных линий. Рассмотрим рис. 33 и 34, на которых положительное время, как и на рис. 30, указано стрелками. Каждая из этих конфигураций может удовлетворять соотношениям (1)—(3). Общая причина, так же как и общее следствие, может экранировать событие A_1 от A_3. Например, если на рис. 33 A_2 представляет атмосферное давление, A_1 — показание барометра и A_3 — дождь, то для предсказания

дождя достаточно знать атмосферное давление; тогда знание того, что A_1 представляет запись этого давления, является излишним. Однако предсказание дождя, если мы исходим непосредственно из A_2, более надежно, чем это же предсказание, основанное только на записи A_1,

Рис. 33 (слева) и рис. 34 (справа). Как общая причина A_2, так и общее следствие A_2 могут находиться в причинном отношении между двумя событиями A_1 и A_3.

потому что барометр может оказаться неисправным. В качестве иллюстрации конфигурации, изображенной на рис. 34, можно использовать пример из § 19 с водителями грузовиков. Следовательно, отношение *причинно между* не позволяет нам отличить конфигурации, представленные на рис. 33 и 34, в которых причинные линии *противоположно направлены*, от конфигурации, представленной на рис. 30, в которой причинные линии обладают одинаковым направлением, то есть *однонаправлены*.

Этот результат напоминает подобную ситуацию, существующую для причинной сети в классической механике. В § 5 мы видели, что законы механики не могут сказать нам об относительном порядке двух различных механических движений, как пояснено на рис. 3. Когда мы говорим, что два шара на рис. 3 движутся либо в направлении сплошных стрелок, либо в направлении пунктирных стрелок, то этот результат основывается на особых наблюдениях совпадений и приближенных совпадений, таких, какие были изучены в отношении рис. 4. Следовательно, мы ввели принцип локальной сравнимости

временнóго порядка, который дает методы для выделения противоположно направленных от однонаправленных причинных линий. Данный принцип требуется и тогда, когда необходимо провести это различие для причинной сети, основывающейся на статистических законах вместо строгих причинных законов.

Рис. 35. События в окружении точки A_2, иллюстрирующие локальную сравнимость временнóго порядка.

Рассмотрим конфигурацию, подобную той, которая изображена на рис. 35. Пусть мы нашли, что A_2 в причинном отношении находится между A_3 и A_5 и что A_3 и A_5 находятся по соседству с A_2 и так близки друг к другу, что мы можем назвать A_3A_5 приближенным совпадением. Тогда мы заключаем, что причинные линии A_2A_3 и A_2A_5 противоположно направлены. Подобным же образом предположим, что A_2 в причинном отношении находится между A_1 и A_4, но что сочетание A_1A_4 представляет приближенное совпадение. Тогда мы заключаем, что причинные линии A_2A_1 и A_2A_4 противоположно направлены. Кроме того, пусть A_2 в причинном отношении находится между A_1 и A_3, но что A_1 и A_3 не представляют приближенного совпадения. Тогда мы заключаем, что причинная линия идет в направлении $A_1A_2A_3$ или в направлении $A_3A_2A_1$, то есть что линия A_1A_2 и A_2A_3 имеют одинаковое направление. Подобные выводы можно сделать и для линии $A_4A_2A_5$.

Эти выводы основываются на двух предположениях. Первое касается приближенных совпадений. Мы должны иметь средства наблюдения, чтобы выяснить, например, что A_3 и A_5 образуют практическое совпадение, в то время как A_1 и A_3 не образуют такого совпадения. Даже если A_1 и A_3 движутся более близко к A_2, мы предпола-

гаем, что, поскольку эти события можно отличить от A_2, мы можем говорить, что они не представляют практического совпадения. Другими словами, если мы используем весьма неточное определение «совпадения», например, означающее, что A_1 и A_3 объединяются в одно событие, то тогда это событие также будет включать и A_2. Иначе обстоит дело в отношении событий A_3 и A_5, которые, согласно некоторому неточному определению «совпадения», образуют одно событие, в то время как A_2 еще образовывает особое событие.

Второе предположение касается неравенств (1) и (2). Если события A_1, A_2, A_3 находятся весьма близко друг к другу, то значения первых двух вероятностных выражений в (1) и (2) почти равны 1, и трудно проверить наличие разницы в их значениях, что утверждается в (1) и (2). Однако наличие этой разницы существенно, потому что она была необходима для доказательства асимметрии отношения «между», выраженного в (10) и (11). Если бы мы не могли доказать эту асимметрию, то порядок $A_1A_2A_3$ невозможно было бы отличить от порядка $A_2A_1A_3$. По этим причинам мы должны предположить, что даже для соседних событий можно также проверить неравенство первых двух вероятностных выражений в (1) и (2).

Однако это предположение не вызывает каких-либо трудностей. Хотя мы можем предположить, что если наступает A_1, то наступление A_3 характеризуется высокой вероятностью, но ситуация резко изменится, если не наступит A_2; в этом случае A_3 будет весьма маловероятно. Это означает, что мы можем проверить соотношение

$$P(A_1 . \bar{A}_2, A_3) < P(A_1, A_3). \tag{14}$$

Из этого соотношения можно непосредственно получить [1], что

$$P(A_1 . A_2, A_3) > P(A_1, A_3). \tag{15}$$

---

[1] См. ThP, p. 79, замечания, следующие за уравнением (11в). Однако этот вывод оправдан только при наличии неравенства $P(A_1, A_2) < 1$, которое было получено выше из соотношений (1) — (3).

Так как, согласно (3), член слева в (15) равен $P(A_2, A_3)$, то теперь можно вывести и неравенство первых двух вероятностей в (1). Подобные соображения приводят и к первым двум неравенствам в (2). Таким образом, асимметрия отношения «между» в соседстве с A_2 сохраняется.

Если мы хотим проиллюстрировать применение принципа локальной сравнимости временного порядка, мы можем использовать пример с шарами, рассмотренный в связи с рис. 4. Так как для событий A_1, A_2, A_3, расположенных вблизи друг друга, относительные вероятности в (1) и (2) почти равны 1, то статистические соотношения между событиями практически невозможно отличить от причинных законов и можно считать, что они реализуются в механических процессах.

Фактически, использование строгих законов для движущихся шаров должно рассматриваться как схематизация, которая допустима при высоких значениях вероятностей, но которую в случае более точного анализа необходимо заменить вероятностными взаимоотношениями. Траектории движущихся шаров строго непредсказуемы, так как невозможно полностью исключить возмущения со стороны окружения.

Эта проблема была рассмотрена в § 10, и там было показано, что фактически достижимые описания механических процессов никогда не могут доставить абсолютно определенных предсказаний и ограничиваются предсказаниями с высокими значениями вероятностей. На этом основании мы можем считать, что неравенства (1)—(2) выполняются и в случае движущихся шаров, таким образом, предполагая, что относительные вероятности, связывающие последовательные положения шаров, никогда строго не равны 1. Неравенство (1) тогда можно рассматривать как доказываемое посредством ссылки на неравенство (14), которое можно проиллюстрировать примером с движущимися шарами: если шар вначале занимает положение A_1, но затем некоторое воздействие извне препятствует ему достигнуть положения A_2, он не займет и положения A_3. Подобным образом можно доказать неравенство (2).

Коль скоро показано, что соотношения (1)—(3) удовлетворяются в случае с движущимися шарами, обсу-

ждение вопроса о наблюдаемых совпадениях можно заимствовать из рассмотрения, данного в § 5[1].

С введением принципа локальной сравнимости временнóго порядка построенная на основе статистики причинная сеть приобретает некоторые свойства, характерные для причинной сети классической механики: она упорядочена как целое. Таким образом, эта сеть обладает линейным порядком. Это означает, что, если приписать некоторое направление времени одной из причинных линий, каждая линия станет характеризоваться направлением. Для придания направления мы можем использовать определение направления времени на основе принципа общей причины, рассмотренного в § 19. Таким образом, направление времени определяется для причинной сети в целом. Это определение направления времени основывается только на вероятностных методах и может быть последовательно проведено в рамках макростатистики.

Данное построение причинной сети и ее направления основываются на предположении, что имеются события, удовлетворяющие соотношениям (1)—(3) и принципу локальной сравнимости временнóго порядка. В следующем параграфе мы рассмотрим статистическое определение направления времени, которое не удовлетворяет этим условиям.

§ 23. Принцип маркировки и причинное соответствие

Можно построить причинную сеть и придать ей направление путем прямого использования необратимых процессов, в отношении которых не предполагается их предварительная упорядоченность, но которые обеспечивают и порядок и направление. Данное построение использует микростатистику (см. § 17), поскольку речь идет о необратимых процессах.

---

[1] Утверждение, что A_2 экранирует A_3 от A_5, здесь имеет следующий смысл. Если шар достигает A_3, мы прежде всего делаем вывод, что ранее он находился в A_2. Теперь предположим, что в A_2, если это событие вообще имело место, произошло столкновение с другим шаром. Тогда мы можем сделать заключение, что этот другой шар окажется в A_5. Однако если столкновение в A_2 наблюдалось, то, чтобы получить последнее заключение, нет необходимости знать, что первый шар достиг A_3.

Как и в § 22, мы примем, что макрособытия даны как пространственно-временные единицы и что их можно классифицировать по соопределенным классам. Однако отношения порядка между этими событиями будут определены не на основе макростатистики, а с помощью определенных необратимых процессов, накладываемых на существующие статистические соотношения между этими событиями. Это достигается посредством *маркировочного* устройства.

Маркировка представляет результат вмешательства необратимого процесса. В § 6 мы видели, что некоторый процесс можно истолковывать различным образом в зависимости от того, как рассматривается его возникновение: как результат вмешательства или как результат действия альтернативной причины. Если вмешательство осуществляется посредством обратимого процесса, то, следовательно, этот процесс мы не можем использовать для определения направления времени; используя понятие альтернативной причины, мы можем интерпретировать тот же самый процесс как идущий в противоположном направлении времени. По-другому обстоит дело, когда мы для вмешательства используем необратимый процесс. Так как обратного процесса не происходит, то мы не можем интерпретировать происходящее на основе альтернативной причины, и, таким образом, мы можем использовать вмешательство для определения положительного направления времени [1].

---

[1] В предыдущих публикациях автор неоднократно использовал принцип маркировки, для того чтобы различать причину и следствие: см. Hans Reichenbach, Bericht über eine Axiomatik der Einsteinschen Raum-Zeit-Lehre, в „Phys. Zeitschr., Bd. 22, 1921, S. 683—686; его же, Axiomatik der relativistischen Raum-Zeit-Lehre, Braunschweig, F. Viewegg Sohn, A. G., 1924; его же, Philosophie der Raum-Zeit-Lehre, Berlin und Leipzig, W. de Gruyter & Co. 1928. Некоторые из критиков истолковали мои представления в том смысле, что при использовании необратимых процессов нет необходимости в маркировке. Такая трактовка неверна. В своей книге „Axiomatik der relativistischen Raum-Zeit-Lehre" (стр. 22) я писал, что использование маркировки, возможно, предполагает использование необратимых процессов. Однако в начале 20-х годов проблема направления времени еще не была достаточно проанализирована, и я полагал, что для определения направления времени существуют иные статистические методы, чем те, которые представлены во втором законе термодинамики. Этот взгляд теперь мною исправлен; см. § 19, прим. на стр. 220.

Вмешательство посредством необратимых процессов можно совершить преднамеренно, но оно также может быть результатом естественных причин. Предположим, что луч белого света идет из A в B (рис. 36); мы совершим вмешательство, поместив на его пути красное стекло. Когда луч света достигнет B, то он окажется окрашенным в красный цвет, тогда как в A не произошло никаких изменений. Мы скажем: вмешательство посредством красного стекла изменяет луч света на пути между стеклом и точкой B, а между стеклом и точкой A оставляет свет неизменным.

Рис. 36. Пример маркирующего процесса: красное стекло на пути белого луча света.

Описывая эти два процесса в противоположном направлении времени, мы говорим: если нет никакого стекла, то из B в A распространяется белый свет; но то, что до A доходит белый свет, может быть вызвано также и альтернативной причиной: из B распространяется красный свет по направлению к красному стеклу, помещенному на пути света; это стекло испускает световые волны, которые соответствуют различным цветам и которые добавляются к пучку красного света таким образом, что в результате получается белый луч, распространяющийся по направлению к A. Однако такие процессы никогда не наблюдаются при положительном направлении времени, потому что поглощение света в красном стекле представляет необратимый процесс. Следовательно, мы делаем вывод, что второе описание было дано в отрицательном направлении времени. Таким образом, вмешательство посредством красного стекла определяет направление времени, характеризуя состояние в A как испускание светового луча, или причину, а состояние в B — как прибытие светового луча, или следствие.

В качестве других примеров маркировки можно привести отпечатки мела на биллиардном шаре, запах дыма в струе воздуха и снег на крыше железнодорожного вагона. Все такие процессы являются процессами того вида, который был изучен при рассмотрении гипотезы о ветвящейся структуре. Акт вмешательства снова представляет собой точку ответвления системы и обладает низкой энтропией, или вероятностью; затем система развивается по направлению к более высокой энтропии, или вероятности, и, таким образом, определяет направление времени. Например, когда луч белого света попадает на стекло, то это явление, как считал Больцман, представляет взаимодействие, с которого начинается необратимый процесс поглощения света в стекле. Стекло становится теплее, и его состояние изменяется в направлении увеличения энтропии. Красный луч света есть упорядоченное состояние, полученное путем выбора из неупорядоченной смеси лучей всех цветов; и, таким образом, он представляет результат взаимодействия. Когда кусок мела касается биллиардного шара, частицы мела переходят на шар; это есть необратимый процесс. Всякий раз, когда мы говорим: «Маркировка совершается при данном событии А», или «Процесс маркировки происходит в А», мы подразумеваем, что событие A представляет точку взаимодействия, в которой расположен нижний конец ветви. Выражение «точка начала маркировочного процесса» относится к этой точке взаимодействия. Следовательно, направление времени, устанавливаемое путем использования маркировочного процесса, представляет направление увеличения энтропии, выраженное в ветвящейся структуре вселенной.

Если маркировочный процесс, приложенный к событию A_i, проявляется в другом событии A_k, то мы знаем, что A_i предшествует A_k. Кроме того, мы знаем, что A_i является причиной, содействующей существованию A_k; это взаимоотношение может быть названо *причинным соответствием* (causal relevance). Следовательно, мы ввели следующее определение:

Определение 1. Если маркировка, произведенная при некотором событии A_i, проявляется в событии A_k, тогда A_i *причинно соответствует* A_k.

Если мы хотим установить порядок среди событий $A_1 \ldots A_n$, то недостаточно произвести маркировку в A_1.

Эта маркировка может проявляться во всех событиях $A_2 \ldots A_n$; но, таким образом, мы не узнаем порядок этих событий и не сможем даже сказать, располагаются ли события в одну причинную линию. Чтобы найти причинные линии, мы должны неоднократно повторить процесс маркировки, сделав его сначала в A_1, затем в A_2 и т. д. Предположим, что маркировка, произведенная в A_1, проявляется в A_2 и, возможно, также в A_3, и предположим, что маркировка в A_2 проявляется в A_3; тогда мы узнаем, что A_2 находится в причинном отношении между A_1 и A_3. Идя таким путем, мы можем события $A_1 \ldots A_n$ расположить упорядоченным образом в причинную линию, которая имеет направление, определяемое необратимостью маркирующих процессов.

Этот процесс не может привести обратно к событию A_1, а приводит к образованию незамкнутых линий. Это следует из предположения 5 гипотезы о ветвящейся структуре, сформулированной в § 16. Если бы линия, построенная процессом маркировки, оказалась замкнутой, то существовал бы по меньшей мере один необратимый процесс, направление которого не совпадало бы с направлением остальных; а это, в смысле вероятности по многим системам, настолько невероятно, что этим случаем можно пренебречь.

Допустим, что мы, таким образом, нашли, что события $A_1 \ldots A_n$ расположены в причинную линию. Тогда события $A_1 \ldots A_{n-1}$ мы назовем *причинными предками* A_n, а события $A_2 \ldots A_n$ будут называться *причинными потомками* A_1. Соответствующие отношения в родословном дереве носят в точности такой же характер. Процессы маркировки как бы проводят причинные линии между событиями. Между прочим, метод меченых атомов, применяемый в наше время для обнаружения причинной структуры обменных процессов в живых организмах, представляет процесс, указывающий на существование причинных линий.

Процесс маркировки достаточен, чтобы сделать возможным построение причинной сети, которая, таким образом, с самого начала характеризуется направлением. Однако возможно также соединить процесс маркировки и с другими статистическими методами. Например, мы можем сначала построить причинную сеть, используя отношение *причинно «между»*, определенное в (22.1)—

(22.3). Вместо применения принципа локальной сравнимости временнóго порядка мы можем в таком случае пользоваться процессами маркировки, для того чтобы отличить противоположно направленные линии от однонаправленных и, таким образом, найти причинные линии с постоянным направлением. Отношения «между» вдоль этих линий можно не проверять при помощи дальнейших маркировок, а брать из предыдущего построения.

Однако смешанная процедура этого вида предполагает определенное соответствие между порядком, установленным при помощи маркирующих процессов, и порядком, выведенным из соотношений (22.1)—(22.3). Это соответствие и необходимо теперь изучить. Оно базируется на некоторых предположениях, которые мы сформулируем по отдельности.

Ссылаясь на последнюю часть неравенства (22.1), мы вводим следующее предположение:

Предположение α. Если маркировка, произведенная в A_i, проявляется в A_k, то тогда

$$P(A_i,\ A_k) > P(A_k). \tag{1}$$

Таким образом, предполагается, что это неравенство представляет необходимое условие передачи результата маркировки, хотя и не является достаточным условием. Можно считать, что это предположение обосновано многими опытами. Неравенство (1) формулирует статистическое соответствие в самом широком смысле, и вполне возможно, что более узкое понятие причинного соответствия предполагает более широкое понятие статистического соответствия. Говоря более точно, (1) выражает *положительное соответствие* в статистическом смысле. Если знак «больше чем» заменить знаком «меньше чем», то мы получили бы *отрицательное соответствие*. Только если бы он был заменен знаком равенства, мы получили бы *несоответствие*. Следовательно, мы можем истолковывать предположение α как означающее, что причинное соответствие представляет частную форму положительного соответствия.

Обозначив через A' событие, получающееся в результате маркировки A, нам теперь необходимо ввести второе предположение, относящееся к процессу маркировки:

Предположение β. *Если маркировка производится в* A_l, *то тогда либо*

$$P(A_l', A_k') = P(A_l, A_k), \tag{2}$$

либо

$$P(A_l', A_k) = P(A_l, A_k). \tag{3}$$

Если справедливо (2), мы говорим, что произведенная при A_l маркировка проявляется в A_k. Если справедливо (3), мы скажем, что она не проявляется в A_k. Если A_k предшествует A_l, то всегда имеет место второй случай; но он также может произойти и тогда, когда A_k следует за A_l, потому что результат маркировки может исчезнуть в некотором месте на причинной линии, ведущей к A_k. Например, снег на крыше железнодорожного вагона может растаять. Уравнения (2)—(3) устанавливают, что процесс маркировки контролируется соотношениями, в которых практически не приходится сомневаться; он либо проявится, либо не проявится во втором событии в зависимости от примененного способа маркировки и, таким образом, не изменяет существующие вероятности.

Теперь нам необходимо исследовать связь между процессом маркировки и отношением экранирования, определенным в (22.3). Если процесс маркировки используется для того, чтобы указать причинные линии, результат маркировки проявляется во всех событиях вдоль линии; если этот знак исчезнет в какой-нибудь точке, то дальше он уже вновь не появится. Например, отметка мела на биллиардном шаре может стереться в некоторый момент времени; если это произойдет, то в дальнейшем эта отметка на шаре уже не появится. Это соображение приводит к следующему предположению, которое выражает непрерывность результата маркировки:

Предположение γ. *Если* A_2 *экранирует* A_1 *от* A_3 *и если произведенная в* A_1 *маркировка проявится в* A_3, *то тогда она также проявится и в* A_2.

Это предположение необходимо рассматривать как эмпирическую гипотезу, согласно которой процесс маркировки соответствует принципу близкодействия. Мы можем привести в его пользу следующий аргумент. Мы можем доказать, что, если предположение γ неверно, маркирующий процесс нарушает соотношение экранирова-

ния, то есть это отношение не выполняется после того, как произведена маркировка. Другими словами, предположение γ является необходимым условием, если мы предполагаем, что маркирующий процесс не нарушает соотношения экранирования.

Чтобы дать это доказательство, допустим, что произведенная в A_1 маркировка проявляется в A_3, но не проявляется в A_2. Нам нужно показать, что это допущение несовместимо с предположением, что соотношение экранирования выполняется для этих событий как при наличии маркировки, так и без нее. Последнее предположение выражается двумя соотношениями:

$$P(A_1 . A_2, A_3) = P(A_2, A_3), \tag{4}$$

$$P(A_1' . A_2, A_3') = P(A_2, A_3'). \tag{5}$$

Мы докажем, что эти соотношения противоречат соотношениям (2)—(3).

Доказательство начнем с использования правила исключения[1], теоремы из теории вероятностей, в следующем виде:

$$P(A_1', A_3') = P(A_1', A_2) \cdot P(A_1' . A_2, A_3') + \\ + P(A_1', \bar{A}_2) \cdot P(A_1' . \bar{A}_2, A_3'). \tag{6}$$

Применяя (2) к члену $P(A_1', A_3')$ и (3) к члену $P(A_1', A_2)$ и используя (5), мы находим

$$P(A_1, A_3) = P(A_1, A_2) \cdot P(A_2, A_3') + \\ + [1 - P(A_1, A_2)] \cdot P(A_1' . \bar{A}_2, A_3'). \tag{7}$$

Так как A_3' может произойти только в том случае, если не происходит A_3, то есть, если происходит $\bar{A}_3$, мы получаем

$$P(A_1, A_3') \leqq 1 - P(A_2, A_3). \tag{8}$$

Значение последней вероятности в (7) неизвестно; если мы заменим ее наиболее высоким возможным значением, равным 1, то этим мы можем только увеличить зна-

[1] ThP, p. 76.

чение выражения в правой части (7). Таким образом, мы находим при использовании (8)

$$P(A_1, A_3) \leqq P(A_1, A_2) \cdot [1 - P(A_2, A_3)] + 1 - P(A_1, A_2),$$
$$P(A_1, A_3) \leqq 1 - P(A_1, A_2) \cdot P(A_2, A_3). \quad (9)$$

Если мы теперь применим правило исключения к событиям A_1, A_2, A_3, мы получим уравнение, аналогичное (6), с единственным различием в том, что опущены знак «штрих». Используя (4), мы получаем

$$P(A_1, A_3) \leqq P(A_1, A_2) \cdot P(A_2, A_3). \quad (10)$$

Складывая два неравенства (9) и (10), мы находим, что они приводят к новому неравенству

$$P(P_1, A_3) \leqq \frac{1}{2}. \quad (11)$$

Теперь из (4)—(5) мы можем получить два соотношения:

$$P(A_1, A_2, \bar{A}_3) = P(A_2, \bar{A}_3), \quad (12)$$
$$P(A_1', A_2, \bar{A}_3') = P(A_2, \bar{A}_3'). \quad (13)$$

Кроме того, из (2)—(3) мы получаем два других соотношения:

$$\text{либо } P(A_i', \bar{A}_k') = P(A_i, \bar{A}_k), \quad (14)$$
$$\text{либо } P(A_i', \bar{A}_k') = P(A_i, \bar{A}_k). \quad (15)$$

Если мы теперь, исходя из соотношений (12)—(15), проведем доказательство, строго напоминающее предыдущее, мы найдем

$$P(A_1, \bar{A}_3) \leqq \frac{1}{2}. \quad (16)$$

Это неравенство противоречит (11), за исключением случая, когда $P(A_1, A_3) = 1/2$. Однако этот случай можно исключить, исходя из следующих соображений. Знак равенства в (9) возможен только при условии, что $P(A_1', \bar{A}_2, A_3') = 1$. Но так как в этом случае $P(A_1', \bar{A}_2, \bar{A}_3') = 0$, мы можем при этих условиях знак равенства в (16) опустить. Легко заметить, что вследствие этой зависимости только одно из соотношений (11) и (16) может иметь знак равенства. Но тогда эти неравенства несовме-

стимы. Последнее и доказывает, что предположение γ должно выполняться, если маркировочный процесс не нарушает соотношения экранирования.

Теперь мы несколько расширим трактовку отношения экранирования. В примере с конфигурацией, представленной на рис. 35, может случиться, что ни $A_2^{(1)}$, ни $A_2^{(2)}$ не экранируют A_1 от A_3, в то время как это делает их конъюнкция $A_2^{(1)} . A_2^{(2)}$. Это значит, что соотношение (22.3) будет удовлетворено, если вместо A_2 мы положим $A_2^{(1)} . A_2^{(2)}$. Следовательно, можно говорить о системе $A_2^{(1)} \ldots A_2^{(n)}$, экранирующей A_1 от A_3. Предположение γ теперь можно расширить следующим образом:

Предположение δ. Если система $A_2^{(1)} \ldots A_2^{(n)}$ экранирует A_1 от A_3 и если произведенная в A_1 маркировка проявляется в A_3, то тогда она по крайней мере проявляется в одном из событий $A_2^{(1)} \ldots A_2^{(n)}$.

Предположения α — δ формулируют соответствие между маркирующим процессом и отношением *причинно «между»*, определенным в (22.1)—(22.3). Используя это соответствие, мы теперь можем ввести второе определение причинного соответствия, которое эквивалентно первому и которое применимо, когда уже определено направление времени. Так как это предположение выполняется во многих практических применениях, второе определение причинного соответствия дает удобное орудие, обладающее большой практической ценностью.

Допустим, что ни одно из событий $A_2^{(1)} \ldots A_2^{(n)}$ не наступает позже, чем A_1; тогда некоторые из этих событий могут наступить раньше, чем A_1, другие могут произойти одновременно. Так как мы знаем, что для переноса результата маркировки всегда необходимо время, то произведенная в A_1 маркировка не может проявиться в каком-нибудь из событий $A_2^{(1)} \ldots A_2^{(n)}$. Пусть A_3 — событие, которое экранируется от A_1 набором $A_2^{(1)} \ldots A_2^{(n)}$; тогда из предложения δ мы выводим, что произведенная в A_1 маркировка не может проявиться в A_3. Объединяя этот результат с предположением α, мы приходим к следующему определению:

Определение 2. Событие A_1 *причинно соответствует* более позднему событию A_3, если

$$P(A_1, A_3) > P(A_3) \tag{17}$$

и если не существует такой системы событий $A_2^{(1)} \ldots A_2^{(n)}$, которые происходят раньше или одновременно с A_1, но которые экранируют A_1 от A_3.

Это определение причинного соответствия, которое предполагает направление времени, удовлетворяет следующему условию. Если A_1 причинно соответствует A_3, то тогда перенос результата маркировки от A_1 к A_3 не исключается предположением δ о непрерывности. Кроме того, во всех таких случаях причинного соответствия физически возможно найти или произвести маркировку, которая переходит от A_1 к A_3. И можно считать, что это предположение подтверждено многими экспериментами. Следовательно, экстенсиональная идентичность приведенных двух определений причинного соответствия доказывается эмпирическим путем.

Второе определение причинного соответствия можно использовать там, где по техническим причинам не представляется возможным применить маркирующий процесс, как, например, в случаях, когда необходимо исключить из рассмотрения экспериментальное вмешательство. Основную идею этого определения кратко можно изложить следующим образом: если некоторое событие A_1 экранируется от более позднего события A_3 посредством события A_2, которое наступает во времени не позднее, чем A_1, тогда A_1 причинно не соответствует A_3. Однако, если бы A_2 наступало позднее, чем A_1, оно могло бы экранировать A_1 и тем не менее оставляло бы A_1 причинно соответствующим A_3.

И определение 1, и определение 2 построены так, что они выражают направление и, таким образом, формулируют *причинное влияние* в отличие от *причинного указания*. Если A_1 причинно соответствует A_3, то тогда A_1 сказывается на существовании A_3. Неопределенные выражения «причинное влияние» и «сказывается на существовании события» получают строгое толкование, когда учитываются перенос результата маркировки или отсутствие события A_2, более раннего, чем A_1, но экранирующего A_1 от A_3. В противоположность этому причинное указание можно отождествить с положительным статистическим соответствием вообще, как сформулировано в (1).

Глава V

ВРЕМЯ В КВАНТОВОЙ ФИЗИКЕ

§ 24. Статистическая обратимость элементарных процессов в квантовой механике

Для классической статистики атомов проблема направления времени выступает, как было показано, в парадоксальной форме: необратимость макропроцессов нужно совместить с обратимостью микропроцессов [1]. Решение этого парадокса статистического направления было найдено в развитии идей Больцмана в результате признания секционной природы направления времени. В самом деле, развитие во времени большой изолированной системы в целом может определять направление времени секции, если эта секция богата ветвящимися системами, управляемыми законами статистической изотропии.

Проблема Больцмана возникла из представления, что атомные процессы подчиняются законам механики Ньютона, в которой одно направление времени не отличается от противоположного и предполагается обратимость всех элементарных процессов. Когда открытия квантовой физики привели к выводу, что атомные процессы выпадают из сферы действия классической механики, что они требуют создания специфической квантовой механики, которая в своих основных чертах существенным образом отличается от механики Ньютона, многие физики надеялись, что новая механика, возможно, позволит дать другое решение проблемы направления времени. Высказывалось предположение, что элементарные процессы в квантовой механике несимметричны относительно времени. Если бы это предположение было верным, тогда направление времени было бы имманентно наиболее фундаментальным законам физики и, чтобы дать ему

---

[1] См. стр. 150, 160 и 184.

оценку, нам не нужно было бы прибегать к статистическим соображениям.

Надежда на такое решение проблемы направления времени была вскоре оставлена физиками. Изучение закономерностей квантовой механики убедительно показало, что на основе квантовых процессов нельзя отличить одно направление времени от противоположного и что они так же обратимы, как и процессы классической механики. Это верно, даже несмотря на то, что квантовые процессы подчиняются не строгим причинным законам, а вероятностным. Отсюда следует вывод, что определение направления времени в квантовой физике, точно так же как и в классической, зависит от статистических свойств макроскопических процессов. Однако мы увидим, что сочетание квантовых законов, которым подчиняется микрокосм, с классическими статистическими законами, которым подчиняется макрокосм, позволяет выявить новые аспекты проблемы времени, поскольку квантовая физика построена на базе причинного индетерминизма. Действительно, однонаправленная природа времени получает более резко выраженную форму, когда она представляется в виде совокупного результата классических статистических свойств макропроцессов и индетерминированности микропроцессов. Развитие этого тезиса является предметом данной главы.

Прежде всего необходимо объяснить, почему квантовые процессы обратимы. В классической механике обратимость движения материальных точек вытекает из того факта, что дифференциальные уравнения движения являются уравнениями второго порядка, в них отсутствуют первые производные. Следовательно, если $f(t)$ есть решение уравнения, то $f(-t)$ также является решением (см. § 4). Напротив, зависящее от времени уравнение Шредингера, которое определяет изменение ф-функции во времени, является уравнением первого порядка относительно переменной t. Поэтому может показаться, что с помощью этого уравнения можно установить различие между противоположными направлениями времени.

Чтобы пояснить это соображение, отметим, что в квантовой механике физическая величина u характеризуется распределением ее вероятности, то есть функцией $d(u)$, которая говорит нам, с какой вероятностью будет наблюдаться определенное значение этой величины, если

произвести соответствующие измерения. Функция $\phi(q)$, которая занимает центральное положение в математическом аппарате квантовой механики, может рассматриваться как математическая схема, объединяющая в одной функции все эти функции распределения рассматриваемых величин. Величина q обозначает совокупность пространственных параметров рассматриваемой системы, для материальной точки q характеризует просто систему ее трех пространственных координат. Функция ϕ принимает комплексные значения, хотя ее аргументы q являются вещественными числами. Различные функции распределения $d(u)$, которые являются вещественными функциями, будучи плотностями вероятности, могут быть выведены из $\phi(q)$ при помощи определенных математических правил, рассматривать которые здесь нет необходимости [1].

Таким образом, функция $\psi(q)$ позволяет сформулировать описание D физического состояния системы, которое является настолько полным, насколько это возможно. Когда состояние меняется со временем, переменная t выступает в качестве другого аргумента функции, и функция тогда записывается в форме $\psi(q, t)$. Для задания изменения этой функции во времени Шредингер сформулировал дифференциальное уравнение, выражающее фундаментальный закон изменения в квантовой механике.

Это уравнение имеет форму

$$H_{op}\psi(q, t) = c\frac{\partial\psi(q, t)}{\partial t}; \quad c = \frac{ih}{2\pi}. \tag{1}$$

Символ H_{op} обозначает математический оператор, получаемый из классического выражения энергии H. Оператор — это правило, при помощи которого из функции, к которой применяется оператор, получается некоторая новая функция. Направление времени, в котором происходит изменение, раскрывается в знаке аргумента t.

При изучении этой проблемы мы должны были бы тщательно рассмотреть некоторые математические детали, которые понятны только читателю, знакомому с математическим аппаратом квантовой механики. Однако выводы все-таки можно понять и не будучи знакомым с этим аппаратом.

[1] В качестве руководства по квантовой механике, которое в состоянии дать более подробное рассмотрение соображений, упомянутых в данной главе, см. PhF.

Сначала мы отметим, что если $\phi(q, t)$ есть решение этого уравнения, то $\phi(q, -t)$ таковым не является. Последняя функция удовлетворяет уравнению

$$H_{op}\psi(q, -t) = -c\frac{\partial\psi(q, -t)}{\partial t}, \qquad (2)$$

которое отличается от (1) знаком минус в правой части уравнения.

Остается проблема, по каким признакам можно различать $\phi(q, t)$ и $\psi(q, -t)$. Для того чтобы различать эти две функции, нам сначала нужно узнать, какое из этих уравнений — (1) или (2) — правильно. Однако знак в правой части уравнения Шредингера может быть проверен путем наблюдения только в том случае, если предварительно определено направление времени. Мы используем здесь направление времени в макроскопических системах, с помощью которого сравниваем математические выводы из уравнения Шредингера с наблюдениями. Следовательно, попытка определить направление времени при помощи уравнения Шредингера представляла бы собой логический круг. Это уравнение дает только то направление времени, которое мы предварительно вводим с помощью макроскопических процессов.

Можно напомнить, что даже в классической физике направление движения молекулы относительно времени не может быть установлено из наблюдений за ее механическим движением, даже если такое непосредственное наблюдение было бы возможно. Оно определяется только путем сравнения с макропроцессами, для которых направление времени определяется статистикой. Точно так же направление времени элементарных квантово-механических процессов, подобных движению электрона, определяется только путем соотнесения со временем макропроцессов.

Эти соображения показывают, что даже с помощью фундаментального закона квантовой механики, которому подчиняется развитие во времени физических систем, нельзя провести различие между противоположными направлениями времени. Поскольку законы, управляющие наблюдаемыми в квантовой физике, являются не причинными, а вероятностными, то обратимость элементарных процессов предполагает здесь симметричную форму со-

отношений, связывающих распределения вероятностей. Эти соотношения представляют собой строгие законы, которые можно выразить с помощью дифференциального уравнения, а именно уравнения Шредингера (1); таким образом, статистические законы, управляющие наблюдаемыми, связаны между собой при помощи строгого математического соотношения, которое контролирует изменения распределений вероятностей во времени.

Существование этих связей истолковывалось иногда как возвращение к причинным законам. Однако между причинным законом в его традиционном смысле и законом в смысле математического соотношения между распределениями вероятностей имеется различие. В квантовой механике существуют только статистические законы, если под ними понимать отношения между наблюдаемыми. Когда мы говорим, что ψ-функция описывает физическое состояние, нельзя забывать, что мы называем здесь описанием состояния только статистическое описание наблюдаемых. Эта функция говорит нам, с какой вероятностью измерения приведут к некоторым частным результатам. И строгое соотношение между распределениями вероятностей еще не превращает отношения между наблюдаемыми в строгие отношения. Другими словами, если мы переходим от наблюдаемого A к распределению B с помощью вероятностной импликации, от B — к другому распределению C с помощью строгой импликации и от C — к наблюдаемому D с помощью вероятностной импликации, тогда отношение между A и D все же является вероятностной импликацией.

Однако справедливым является то, что изменение во времени физических состояний можно выразить с помощью соотношения, подчиняющегося строгому закону. Используя такое понимание, мы можем дать следующее описание временной симметрии квантовой механики. Функция $\psi(q, t_1)$, охватывающая некоторую область q-пространства, определяет систему s_1 распределений вероятностей всех физических величин, входящих в систему в момент t_1; функция $\psi(q_2, t_2)$ также определяет систему s_2 этих распределений вероятностей в момент t_2. Если физический закон допускает переход от s_1 к s_2, то обратный переход от s_2 к s_1 также не противоречит физическому закону. В этой версии обратимость элементарных процессов вновь появляется в квантовой физике.

Нельзя утверждать, что существует непосредственная обратимость каждого элементарного процесса, поскольку не все эти процессы управляются строгими законами. Если мы не можем гарантировать, что при точном повторении начальных условий данный процесс приведет к тем же результатам для того же направления времени, то мы не можем также говорить, что обратный процесс приведет к тем же самым начальным физическим условиям. Вместо этого мы можем говорить только о *статистической обратимости*. Эта обратимость выражается в форме временной симметрии строгого закона, определяющего изменение во времени распределений вероятностей. Таким образом, по сравнению с классической физикой, квантовая механика не внесла ничего нового в решение проблемы направления времени.

Теперь рассмотрим изменения, внесенные квантовой механикой в проблему детерминизма.

§ 25. Индетерминизм квантовой механики

Индетерминизм квантовой механики выражается в соотношении неопределенностей В. Гейзенберга. Система понятий квантовой механики приводит к выводу, что для любой физической величины u существует много других величин v, которые нельзя измерить одновременно с ней. Поскольку это отношение выражено в алгоритме квантовой механики посредством некоммутативности произведения матриц этих величин, такие величины называются *некоммутативными* или *некоммутирующими* [1]. Это не означает, что все величины просто делятся на два класса; в данном случае подразделение происходит на гораздо большее число классов, потому что имеются величины u, v, w, из которых ни одна не коммутирует с какой-либо другой. Специфическим примером некоммутирующих величин являются q и p координата и импульс частицы[2].

---

[1] Напротив, обычное арифметическое произведение всегда является коммутативным $x \cdot y = y \cdot x$.

[2] Обычно для обозначения импульса употребляют букву p, а для обозначения координат — букву q. Хотя автор использует те же самые буквы для обозначения вероятностей и подобных им величин, однако подразумеваемое значение всегда будет ясно из контекста.

Гейзенберг показал, что величины q и p могут быть измерены только до определенной степени точности Δq и Δp. При этом выполняется следующее неравенство:

$$\Delta q \cdot \Delta p \geqq \frac{h}{4\pi}. \qquad (1)$$

Величина h — квант действия Планка — очень мала, поэтому неточность проявляется только при очень тщательных измерениях. Так как для материальной точки $p = mv$, то соответствующее соотношение для координат q и скорости v принимает вид

$$\Delta q \cdot \Delta v \geqq \frac{h}{4\pi m}. \qquad (2)$$

Если m велико, тогда правая часть соотношения очень мала. Таким образом, для частиц большой массы, таких, как макроскопические частицы, или частицы пыли, видимые под микроскопом, практически предела измеримости не существует.

Однако для частиц, подобных электронам и другим микрочастицам, открытым в квантовой физике, ограничение (1) имеет существенное значение. Тем не менее это ограничение оставляет нам возможность выбора; мы можем измерить либо q, либо p настолько точно, насколько пожелаем, но при этом измерение второй величины будет весьма неточным. Подобные ограничения накладываются на любые пары некоммутирующих величин.

Из ограничения измеримости следует ограничение возможности предсказаний. В § 11 было показано, что точное предсказание будущего состояния физической системы требует точного знания ее состояния в настоящий момент. Если, согласно соотношению (1), это знание ограниченно, тогда точное предсказание невозможно. Следует отметить, что точное измерение одной величины вовсе не является достаточным условием для точного предсказания будущих значений этой же самой величины. Например, даже если мы точно измерили положение частицы, то не сможем предсказать ее будущее положение, если нам неизвестны ее скорость или импульс в настоящий момент.

Используя терминологию, разработанную в § 11, мы можем сказать, что соотношение Гейзенберга не исключает предсказаний в условной форме, но оно исключает

предсказания в безусловной форме, на которые претендует детерминизм. Если мы хотим предсказать будущее физическое состояние с вероятностью $p^{(l)}$, близкой к 1, то для этого нам нужно сформулировать точное описание D^l состояния в настоящий момент. Однако, согласно Гейзенбергу, мы не можем сформулировать это описание $D^{(l)}$ настолько точно, насколько это необходимо. Следовательно, как схема (11.1), так и схема (11.2) должны быть заменены в квантовой механике следующей схемой:

$$\begin{matrix} D^{(1)} & D^{(2)} & D^{(3)} \ldots D \\ p^{(1)} & p^{(2)} & p^{(3)} \ldots p \end{matrix} \qquad (3)$$

Таким образом, существует конечное описание D, наилучшее из тех, которые мы можем сформулировать. Оно выражается ψ-функцией заданного физического состояния. Однако это описание позволяет нам предсказывать будущее состояние только с вероятностью p, меньшей чем 1, и не существует никаких способов, чтобы когда-либо повысить эту вероятность. Схема (3) выражает индетерминизм квантовой механики и существование предела предсказуемости.

Такая трактовка выдвигает ряд важных вопросов. Может быть, подобное ограничение имеет силу только по отношению к наблюдателю-человеку? Может быть, существуют точные значения измеряемых величин, хотя и не известные нам? И не существует ли неизвестного нам точного описания D', которое позволяет делать предсказания с вероятностью p', более близкой к 1, чем p, или даже с вероятностью p', равной 1? Этот вопрос можно также сформулировать следующим образом: если мы должны отказаться от безусловной формы причинности, то не следует ли нам отказываться также и от ее условной формы? Или, можем ли мы в принципе допустить существование причинных законов следующего вида: если частица имеет в данный момент точное положение q_1 и импульс p_1, тогда в более поздний момент ее положение будет q_2, а импульс будет p_2?

Этот вопрос выдвигается иногда в более оптимистической версии. Аргументируют, что, может быть, ограничение измеримости справедливо только по отношению к состоянию знаний настоящего периода и что

позднее, возможно, будут найдены методы, при помощи которых можно будет преодолеть это ограничение и измерение некоммутирующих величин будет осуществляться с любой желаемой степенью точности. Тогда в будущем схема (3) не будет больше характеризовать положение дел в области предсказания. Таким образом, данная схема была бы справедливой только для физической науки нашего времени, но неправильно было бы считать, что она выражает одно из свойств физического мира.

Ответ на эти вопросы дать не легко. Он требует логического анализа смысла высказываний относительно физического мира, анализа природы умозаключений, которые ведут от утверждений о наблюдаемых к утверждениям о ненаблюдаемых. Это предполагает также тщательное изучение соображений, из которых следует неравенство Гейзенберга. Подобное исследование может помочь нам наметить границу сферы действия этого соотношения и выяснить, является ли оно следствием физического закона или же оно выражает только ограничение способностей современного наблюдателя-человека. Ниже будет показано, что последняя альтернатива несовместима с беспристрастным анализом следствий квантовой механики.

Рассмотрим сначала, как в рамках квантовой механики выводится соотношение неопределенностей (1). Напомним, что функция $\phi(q)$ представляет собой математическую схему, посредством которой в одной функции объединяются все распределения вероятностей $d(u)$ физических величин, характеризующих рассматриваемую систему (см. § 24). Индивидуальные распределения $d(u)$ выводятся из $\phi(q)$ при помощи определенных правил, сформулированных в математическом аппарате квантовой механики. В частности, мы можем вывести из $\phi(q)$ два распределения: $d(q)$ и $d(p)$ — для координат и импульса. С помощью символа d мы обозначаем «распределение». Математические функции $d(q)$ и $d(p)$ могут иметь самые разнообразные формы. Теперь можно показать, что в силу того, что эти две функции $d(q)$ и $d(p)$ выводятся из одной функции, их связь обратно пропорциональна: если $d(q)$ является крутой кривой, тогда $d(p)$ — пологая кривая, и наоборот (см. рис. 37) [1].

---

[1] См. PhF, p. 11.

Обратная пропорциональность получается при любой специфической форме функции $\phi(q)$. Если мы понимаем под Δq и Δp стандартные отклонения, относящиеся к кривым, то получаем соотношение неопределенностей Гейзенберга.

Следовательно, это соотношение прочно укоренилось в системе квантовой механики. Тем не менее оно обосновано не лучше самой системы. Оно вытекает, в частности,

Рис. 37. Обратная корреляция распределений координат $d(q)$ и импульса $d(p)$, которые выводятся из функции $\psi(q)$. Распределение координат q имеет форму кривой Гаусса. См. „Philosophical Foundations of Quantum Mechanics", p. 9, рис. 2 и 3, откуда заимствована эта иллюстрация. Кривая $d(p)$ выводится из разложения в ряд Фурье ψ-функции, откуда получается $d(q)$. Непосредственное разложение в ряд Фурье кривой $d(q)$ дано пунктирной линией на рис. 3 в книге „Philosophical Foundations of Quantum Mechanics".

из более общего предположения, на котором базируются все выводы квантовой механики, из предположения, что все знания, полученные путем наблюдения, можно выразить в ϕ-функции, другими словами, что описание D физического состояния с помощью ψ-функции является наиболее полным из возможных описаний. Назовем это предположение *синоптическим принципом* квантовой механики. Если этот принцип верен, тогда неизбежно и соотношение неопределенностей Гейзенберга, поскольку оно является математическим следствием этого принципа. Однако, если этот принцип окажется ложным, тогда можно будет отказаться и от соотношения неопределенностей.

Сомнения в справедливости этого принципа основываются на том, что в принципе неопределенности содержится отрицательное утверждение. Он утверждает, что не существует наблюдаемых, которые привели бы к лучшим измерениям, чем те наблюдаемые, которые совместимы с соотношением неопределенностей (1). Математический вывод этого соотношения относится только к величинам, управляемым ф-функцией. Две любые некоммутирующие величины, управляемые ф-функцией, удовлетворяют неравенству (1). Если утверждают, что это неравенство имеет силу всегда и везде, то этим предполагают, что не существует наблюдаемых, которые нельзя было бы охватить ф-функцией. Синоптический принцип предполагает, что соотношения неопределенностей Гейзенберга справедливы для всех наблюдаемых.

Важно запомнить это логическое соотношение. Мы можем говорить о проверке соотношения неопределенностей Гейзенберга только в том случае, если признаем справедливость синоптического принципа. Любая проверка соотношения неопределенностей, предложенная в рамках квантовой теории, зависит от признания синоптического принципа. Эта характеристика применима, в частности, к проверке, предпринятой Нейманом [1], который повторил вывод Гейзенберга в более общем с математической точки зрения виде. Однако эта проверка убедительна только в том случае, если верен синоптический принцип. Исследование правомерности квантово-механического индетерминизма должно поэтому сопровождаться доказательством справедливости синоптического принципа.

Имеется два вида доказательств этого принципа. Один негативный: не найдена ни одна физическая система, в которой не выполнялся бы синоптический принцип. Если бы это было единственным доказательством, которым мы располагали бы, тогда недостаточная обоснованность данного принципа была бы бесспорной. К счастью, существует еще позитивное доказательство. Система квантовой механики в целом говорит о справедливости этого принципа, потому что без него она привела бы к абсурдным выводам.

---

[1] John von Neuman, Mathematische Grundlagen der Quantenmechanik, Berlin, J. Springer, 1932, S. 170.

Перед тем как перейти к обсуждению этих выводов, рассмотрим логическую структуру аргументации на примере теории относительности. Принцип, согласно которому в природе не существует скоростей, превышающих скорость света, введенный Эйнштейном, основывается на позитивном и негативном доказательстве. Негативное доказательство этого принципа состоит в том, что никогда не наблюдалось сигналов, распространяющихся с большей скоростью, чем скорость света. Однако принцип Эйнштейна никогда не получил бы признания, если бы он не имел еще и позитивного доказательства. Специальная теория относительности привела бы к абсурдным выводам, если бы этот принцип не был правильным. Можно было бы обращать временной порядок причинных процессов, могли бы порождаться бесконечно большие энергии и т. д. Подобные выводы столь неправдоподобны, что обеспечивают доказательство противоположного утверждения. Представляется весьма маловероятным, чтобы принцип предельности скорости света был ложным. Невероятность выводов, следующих из отрицания справедливости этого принципа, превращается, таким образом, в его индуктивное доказательство.

Такой же ход рассуждений мы можем использовать и при доказательстве справедливости синоптического принципа квантовой механики. Конечно, такое доказательство является индуктивным. Другими словами, абсолютного доказательства этого принципа не существует. Он может быть только подтвержден ссылкой на опыт вообще и представляет собой эмпирическую гипотезу. Однако этот принцип получил весьма солидное экспериментальное подтверждение, поскольку он вытекает из наиболее общих свойств физического мира.

Для того чтобы понять доказательство справедливости синоптического принципа, следует рассмотреть другой фундаментальный принцип, который, подобно принципу Гейзенберга, следует из оснований квантовой механики. Однако этот принцип не зависит от справедливости синоптического принципа. Я предлагаю именовать его *принципом аномалии*. Он связан с принципом дополнительности Бора, однако идет дальше, поскольку из него следуют вполне определенные утверждения относительно ненаблюдаемых квантовой механики.

Наблюдаемые физики (это относится также и к квантовой механике) — всегда макроскопические объекты. Это показания приборов, треки на фотографической пленке и т. д. Все утверждения относительно физических величин меньшего масштаба выводятся при помощи умозаключений, основанием для которых служат макроскопические наблюдаемые.

Логическую структуру подобных умозаключений можно подразделить на несколько этапов. Сначала делаем вывод, что имело место элементарное совпадение, такое, как столкновение между двумя электронами, или между γ-лучами и электроном. Такие выводы основаны на довольно простой теории, в которой на самом деле не содержится квантово-механических закономерностей и которая может быть полностью выражена в понятиях классической физики. Элементарные совпадения этого рода я называю *феноменами*. Иногда их называют также *наблюдаемыми* в смысле несколько более широком, чем обычный. Чтобы избежать недоразумений, мы можем называть их *микронаблюдаемыми* в отличие от *макронаблюдаемых*. Их функция как наблюдаемых состоит в том, что они дают основания для выводов относительно событий другого вида.

Эти события происходят между феноменами, и поэтому я буду называть их *интерфеноменами*. Выводы относительно этих событий могут быть сделаны только в рамках квантовой механики. К тому же эти выводы предполагают существование некоторых определений, или *экстенсиональных правил* языка, позволяющих употреблять язык феноменов при описании интерфеноменов и, таким образом, осмысленно говорить о событиях, которые нельзя наблюдать даже в том случае, когда «наблюдение» понимается в широком смысле слова. Будучи определениями, эти правила не истинны и не ложны; однако мы можем спросить, какие выводы получаются при использовании языка, сконструированного на основании этих правил.

Термины «частица» и «волна» относятся к языку интерфеноменов; с их помощью формулируются утверждения о том, что происходит между локализованными феноменами. Говоря о частицах мы утверждаем, что интерфеномены также являются локализованными. Говоря

о волнах, мы утверждаем, что интерфеномены охватывают широкую область, хотя их конечные продукты — феномены — локализованы. Можно показать, что экстенсиональные правила, приводящие к этим двум видам сущностей, довольно просты. Результат измерения обычно интерпретируется как значение измеренной величины *после* измерения. Когда мы постулируем правило, что величина имела то же значение *до* измерения, то в этом случае мы становимся на точку зрения корпускулярной интерпретации. Когда же мы полагаем, что величина до измерения имела все возможные значения одновременно, то здесь мы переходим к волновой интерпретации. Обе интерпретации *исчерпывающие*, то есть они объясняют не только феномены, но также и интерфеномены.

Теперь мы можем задать вопрос о физических свойствах интерфеноменов, вытекающих из нашего определения. При этом мы получаем странный ответ, что поведение интерфеноменов не подчиняется принципу причинности. Говоря о таком неподчинении, я имею в виду нарушение принципа близкодействия, то есть действия путем непосредственного контакта. Траектория частицы после прохождения ее через щель зависит от того, закрыта или открыта другая щель. Вызывая локализованное совпадение — феномен в какой-либо точке, волна исчезает по всему своему фронту. Существует также много других аномалий. Некоторые из них будут рассмотрены в следующем параграфе. Под принципом аномалии автор подразумевает следующее: если мы используем исчерпывающую интерпретацию, то в результате интерфеномены не подчиняются принципу близкодействия. Точнее говоря, этот принцип может быть сформулирован следующим образом:

Принцип аномалии. Если мы на основании каких-либо правил будем утверждать, что некоммутирующие величины, распределение которых подчиняется ф-функции, могут одновременно обладать точными значениями, в таком случае интерфеномены будут подвержены каузальным аномалиям[1].

---

[1] PhF, p. 26—33. Доказательство принципа аномалии дано в той же работе на стр. 122—129. Упрощенное изложение этого принципа и обсуждение экстенсиональных правил см. H. Reichenbach, The Rise of Scientific Philosophy, Berkeley, Calif., Univ. Calif. Press, 1951, chap. XI.

Учитывая этот результат, можно принять решение не говорить об интерфеноменах и ограничить язык квантовой механики только феноменами. Такая *ограниченная интерпретация* выполнима. Интерпретация подобного вида была разработана Н. Бором и В. Гейзенбергом, которые рассматривают утверждения относительно интерфеноменов как бессмысленные. Интерпретация квантовой механики с помощью предложенной мной трехвалентной логики представляется предпочтительной, потому что она все же в некотором смысле разрешает включение интерфеноменов в язык физики, хотя она также ограничена, поскольку исключает утверждения относительно точного статуса интерфеноменов. Однако все эти соображения не имеют отношения к настоящему исследованию. Мы заинтересованы в исчерпывающей интерпретации, поскольку нас занимает проблема: индетерминизм versus детерминизм.

Поэтому, возвращаясь к вопросу об исчерпывающей интерпретации, отметим, что до сих пор мы еще не использовали синоптический принцип. Принцип аномалии был сформулирован для величин, распределение которых подчиняется ψ-функции; однако в нем не учитывается предположение о существовании величин другого рода. Рассмотрим теперь, что изменилось бы, если бы такие величины существовали.

Сначала мы будем исходить из предположения, что синоптический принцип верен и что все физические наблюдаемые могут быть суммированы в ψ-функции. В этом случае всякая непосредственная верификация интерфеноменов исключена, и не существует способа, с помощью которого можно было бы отличить друг от друга волновую и корпускулярную интерпретации. Эти интерпретации представляют, таким образом, *эквивалентные описания*; обе они допустимы и говорят об одном и том же, но только используя разную терминологию. В результате каузальные аномалии играют довольно безобидную роль. Они не влияют на наблюдаемые, а что касается ненаблюдаемых или интерфеноменов, то существование аномалий, так сказать, не локализуемо. Этим я хочу сказать, что всегда можно избежать каузальных аномалий при решении какой-нибудь специфической задачи, если использовать соответствующую интерпретацию. Однако в таком случае постулируется существо-

вание каузальных аномалий в каком-либо другом месте. Это позволяет решить задачу на основании принципа близкодействия. Затем, когда начинаем исследовать ту проблему, для которой постулировали существование каузальных аномалий, мы можем переменить интерпретацию и вновь рассматривать ее, не прибегая к каузальным аномалиям. Элиминируемость каузальных аномалий — важное свойство квантовой механики, признающей справедливость синоптического принципа. Такое переключение интерпретаций широко практикуется физиками. Оно является дозволенным и законным логическим приемом, потому что и волновая и корпускулярная интерпретации являются эквивалентными описаниями, каждое из которых так же истинно, как и другое.

Предположим теперь, что синоптический принцип неверен. Это означало бы, что существуют наблюдаемые, возможно сегодня и неизвестные, которые нельзя охватить ψ-функцией и распределение которых не подчиняется соотношению неопределенностей Гейзенберга (1). Используя такие наблюдаемые, можно было бы получить более полные знания об интерфеноменах; мы смогли бы узнать, существуют ли между феноменами локализированные события или же они охватывают широкую область. Другими словами, новые наблюдаемые позволили бы провести различие между волновой и корпускулярной интерпретацией.

Получили ли бы мы в этом случае возможность сформулировать строго причинные законы вместо вероятностных импликаций? Этот вопрос следует оставить открытым; мы не знаем, что можно было бы предсказывать на основании этих новых наблюдаемых. Предположим, однако, на мгновение, что они в самом деле дали бы нам возможность делать строгие предсказания, что они позволили бы сформулировать строго причинные законы. Какие бы это были законы? В таком случае мы столкнулись бы с тем, что каузальные аномалии превратились бы в свойства наблюдаемых. Эти аномалии больше нельзя было бы элиминировать, они вошли бы в формулировки причинных законов. Например, если предполагается, что можно измерить одновременно координаты и импульс частицы, проходящей через щель, тогда формулировку условного причинного закона: «Если частица имеет координаты q_1 и импульс p_1 в момент t_1, то соответ-

ствующие значения этих величин в момент t_2 будут q_2 и p_2», — следовало бы заменить другой формулировкой, а именно: «Если частица имеет координаты q_1 и импульс p_1 в момент t_1 и если другая щель закрыта, то соответствующие значения этих величин в момент t_2 будут q_2 и p_2, однако если эта щель открыта, то значения этих величин будут q_2' и p_2'». В таком условном причинном законе выражался бы принцип дальнодействия. Возникли бы и другие трудности. Можно показать, например, что статистическое распределение других частиц включалось бы в этот закон в качестве причинного фактора, даже если эти частицы начали свое движение в то время, когда рассматриваемая частица закончила свое движение[1]. Подобные не менее неожиданные выводы будут рассмотрены в следующем параграфе.

Подводя итоги, мы можем сказать, что существование наблюдаемых, которые не могут быть охвачены ψ-функцией и не подчиняются соотношению неопределенностей (если только существование таких наблюдаемых вообще приводит к строгим причинным законам), не устранит причинных аномалий, существующих для тех величин, которые охватываются ψ-функцией. Наоборот, в таком случае причинные аномалии войдут в сферу наблюдаемых. Если бы синоптический принцип оказался неверным и можно было бы распространить детерминизм на квантовую механику, то это был бы не детерминизм классической физики, а детерминизм, допускающий дальнодействие и каузальные аномалии.

Синоптический принцип, если он верен, действует как экран, предохраняющий нас от детерминизма подобного рода. Поэтому, я сказал бы, что с точки зрения физика следует радоваться тому, что мы не можем наблюдать интерфеномены и не можем вернуться к детерминизму, потому что в таком случае мы имели бы дело с детерминизмом весьма сомнительного вида. С философской точки зрения подход к данной проблеме должен быть менее эмоциональным. Однако и здесь нужно учитывать, что существование детерминизма, изобилующего каузальными аномалиями, является маловероятным, и, следовательно, приведенные соображения дают сильное

[1] См. PhF, § 7 и 8.

доказательство против предположения, что существуют физические величины, не контролируемые ф-функцией.

В самом деле, мы можем сделать следующий вывод: если бы синоптический принцип был неверным, то выводы из известных законов квантовой механики были бы настолько абсурдны, настолько несовместимы с известными свойствами физического мира, что есть основание утверждать о весьма вероятной справедливости синоптического принципа. Как и в случае с принципом, согласно которому в природе не существует скоростей, превышающих скорость света, вся теория в целом дает позитивное доказательство этого принципа. С подтверждением справедливости синоптического принципа индетерминизм квантовой механики представляется достаточно обоснованным не только как результат незнания противоречащих примеров, но и как результат последовательной теории, оказывающей сильную индуктивную поддержку тезису, согласно которому за ограниченностью предсказаний не существует никакого детерминизма. Синоптический принцип и индетерминизм квантовой механики можно, следовательно, рассматривать как физический закон.

Поэтому мы можем сказать, что квантовая механика, *по существу*, статистическая дисциплина, что она не является, подобно статистической версии классической термодинамики, статистической дисциплиной *за неимением лучшего*. Классическая физика дает основания для уверенности в том, что если бы мы могли произвести наблюдения элементарных частиц, то можно было бы предсказать будущее более подробно и с большей точностью. И хотя классическая статистика не предполагает детерминизма молекулярных процессов, она вполне совместима с таким детерминизмом. В § 9 было показано, что статистические законы термодинамики можно вывести из законов классической механики; из канонических уравнений движения на основе теоремы Лиувилля выводится вероятностная метрика фазового пространства, на основе теорем Неймана и Биркгоффа выводится эргодическая гипотеза. Когда мы говорили, что неоднозначная сходимость схемы (11.4) не дает почти никаких доказательств в пользу детерминизма, то этим вовсе не утверждали, что она с ним несовместима. Мы говорили только, что неоднозначная сходимость этой

схемы превращает детерминизм в бессодержательное дополнение. Иначе обстоит дело в квантовой физике. Статистические соотношения квантовой статистики являются соотношениями такого рода, что, если их комбинировать с детерминизмом элементарных событий, получаются абсурдные результаты. Они индуктивно несовместимы с детерминизмом.

Тем, кто хотел бы рассматривать квантовую механику как временное решение, нужно учитывать этот вывод. Им также не следует увлекаться ошибочными аналогиями, ссылаясь на развитие теории атомов в XIX столетии. Верно, что вначале знания об атомах были основаны только на статистических наблюдениях, и лишь позднее они смогли основываться на наблюдении непосредственных эффектов индивидуальных частиц. Однако здесь нельзя проводить параллель с квантовой механикой. Статистическая форма квантовой механики не может рассматриваться как прелюдия к будущей индивидуальной трактовке квантово-механических систем, потому что предположения, лежащие в основе такой трактовки, индуктивно несовместимы с квантовой статистикой, ф-функция квантовой механики несовместима с детерминистским фоном.

Рассмотрим этот приговор детерминизму более подробно с логической точки зрения. Мы не утверждаем, что говорить о детерминизме, лежащем за индетерминизмом квантовой механики, бессмысленно, мы скорее говорим, что такое высказывание ошибочно. Конечно, у нас есть только индуктивное основание, а не убедительное доказательство его ошибочности. Однако, поскольку мы можем проверить это утверждение с помощью индуктивной аргументации, оно не бессмысленно.

Утверждение, на котором основывается эта аргументация, имеет форму конъюнкции и гласит: «Имеются наблюдаемые физические величины, которые нельзя описать с помощью ф-функции и с помощью которых могут быть точно предсказаны все будущие значения физических величин». Это утверждение, как уже пояснялось, не бессмысленно, но, по всей вероятности, ошибочно. Мы можем рассмотреть также и другое утверждение, имеющее форму импликации: «Если бы мы смогли наблюдать физические величины, которые не подчиняются ограничениям, вытекающим из ф-функции,

тогда мы могли бы строго предсказывать будущие значения всех физических величин». Эта импликация выражает условное противодействие факту, в котором ложен антецедент, но если синоптический принцип является физическим законом, тогда эта импликация бессмысленна. Это не означает, что все условные противодействия фактам бессмысленны. Однако чтобы составить приемлемое условное противодействие факту, импликация должна удовлетворять определенным требованиям, сформулированным в теории номологических утверждений. Описанная импликация нарушает эти условия, поскольку претендует на описание закона природы, в то время как ее антецедент противоречит закону природы. Следовательно, эту импликацию невозможно верифицировать индуктивным путем, наблюдая случаи, в которых антецендент является истинным. В подобной импликации слово «если» используется в незаконном смысле, так как не учитывается значение этого слова по отношению к условному противодействию факту [1].

Возвращаясь к поставленному выше вопросу (стр. 282), совместима ли причинность в условной форме с соотношениями неопределенностей, мы получили в итоге, что причинные утверждения данной формы были бы бессмысленны.

---

[1] Относительно этой проблемы см. H. Reichenbach, Nomological Statements and Admissible Operations, chap. V. Если мы обозначаем описанную выше конъюнкцию через $a.b$, тогда $a.b$ имеет значение «ложно» второго порядка, согласно терминологии, разработанной в „Nomological Statements", где физический закон имеет значение «истинно» второго порядка. Теперь, если $a \supset b$ не имеет значения «истинно» третьего порядка, или значения аналитической истины, тогда ее T-случай $a.b$ является ошибкой того же порядка, что и истинная импликация, и, таким образом, данная импликация недопустима. Допустимую импликацию можно проиллюстрировать следующим условным противодействием факту: если бы человек коснулся провода высокого напряжения, он бы умер. Эта импликация имеет значение «истинно» второго порядка, поскольку его антецедент имеет значение «ложно» первого порядка. Напротив, недопустимую импликацию можно проиллюстрировать следующим условным противодействием факту. Если нагреть ледяной кубик до 86° С, он сожмется. Здесь антецедент имеет значение «ложно» второго порядка, тогда как импликация не имеет значения «истинно» третьего порядка, поскольку она выводится из закона природы, согласно которому лед при нагревании сжимается. Такая импликация бессмысленна. В самом деле, ссылаясь на свойства льда, мы можем с таким же успехом вывести импликацию: «Если ледяной кубик нагреть до 86° С, он расширится».

Утверждение: «Если физические параметры ситуации в момент t_1 имеют значения u_1, v_1, ..., тогда они в более поздний момент t_2 будут иметь значения u_2, v_2» — должно быть интерпретировано как эквивалентное следующей форме утверждения: «Если в момент t_1 мы смогли бы наблюдать все параметры точно и обнаружили бы значения u_1, v_1, ..., тогда в момент t_2 мы наблюдали бы значения u_2, v_2, ...,». А это и есть утверждение типа «если — тогда», которое мы только что обсуждали; следовательно, условная форма причинности бессмысленна.

Сравним полученный результат с обсуждением некоторых утверждений типа «если — тогда» на стр. 290—291. Эти утверждения «если — тогда» имеют смысл только потому, что они сделаны на основании предположения о возможности строгого предсказания будущих значений параметров; таким образом, эти утверждения имеют логическое значение. Однако они не имеют физического значения, поскольку предположение о строгой предсказуемости как физический закон ложно. Когда мы говорим, что условная форма причинности бессмысленна, мы имеем в виду физическое значение, то есть такие предполагаемые законы физически бессмысленны[1].

Эти соображения подводят нас к объективному значению неопределенности. Соотношение Гейзенберга не выражает только ограниченность познавательных способностей человека; в нем формулируется скорее физический закон, справедливый для всех физических величин.

Во многих работах по квантовой механике высказывалось мнение, что неопределенность проистекает из возмущения объекта наблюдателем, вмешательство которого изменяет наблюдаемую величину непредсказуемым образом. При этом одновременно выдвигалось утверждение, что квантовая механика бросает вызов эмпирической интерпретации и требует возвращения к идеалистической философии, согласно которой *Я* творит мир или мир по крайней мере не может существовать без наблюдающего *Я*. Я не думаю, что такие взгляды неуязвимы. Наблюдатель не имеет отношения к неопределен-

---

[1] Что касается терминов *логическое значение* и *физическое значение*, см. Hans Reichenbach, Experience and Prediction, Chicago, Univ. Chicago Press, 1938, p. 40.

ности, не он и не его наблюдательный акт порождают неопределенность. Неопределенность есть чисто физический вопрос и может быть сформулирована как объективное свойство физического мира, без ссылки на наблюдателя.

Во-первых, не всякое наблюдение обязательно возмущает объект. Измерительные инструменты, включая и те, которые применяются в квантовой физике, могут быть сконструированы таким образом, что будут давать свои показания в виде цифр на шкалах. Чтение этих цифр наблюдателем не возмущает объекта. Далее, верно, конечно, что когда мы проводим линию между объектом и наблюдателем таким образом, что измерительный прибор относится к стороне наблюдателя, то объект подвергается возмущению, однако нам нет надобности проводить линию именно так. Когда мы включаем прибор на сторону объекта, демаркационная линия проводится таким образом, что акт наблюдения не возмущает объекта. Поэтому не наблюдатель создает неопределенность. Неопределенность, которая все же существует в таких экспериментальных установках, проявляется скорее как отношение между измерительным инструментом и квантовым феноменом, следовательно, только как отношение между физическими объектами.

Поскольку измерительные инструменты являются макроскопическими объектами, мы можем сказать, что неопределенность возникает в том случае, когда затрагиваются отношения между макрокосмом и микрокосмом. Принцип Гейзенберга утверждает, что не существует способа определения микроскопических величин с помощью макроскопических величин с большей степенью точности, чем это допускается неравенством (1). Это утверждение выражает физический закон, отношение между макрокосмом и микрокосмом. Поскольку люди сами являются макроскопическими объектами и поскольку их органы чувств реагируют только на возбудители макроскопической сферы, постольку их выводы относительно микрокосма должны быть основаны на макроскопических наблюдениях. Вот почему человеческое знание о микрокосме ограничено принципом неопределенности. Таким образом, мы достигли чисто эмпирической интерпретации ограниченности человеческих знаний относительно квантовых феноменов. Если кван-

товая механика изучается под лупой логического анализа, то нет необходимости вновь становиться на точку зрения идеалистических философских систем.

§ 26. Генетическое тождество квантовых частиц

Отношение генетического тождества, связывающее различные состояния одной и той же вещи в различные моменты времени, рассматривалось в § 5. Вещь представляет собой серию событий, следующих друг за другом во времени; любые два события этой серии генетически тождественны. Понятие физического тождества индивидуальной вещи, которая остается той же самой на всем протяжении определенного отрезка времени, основывается на свойствах этого отношения.

Высказывания о вещах и высказывания о событиях представляют собой лишь различные виды языка. Если мы желаем перевести предложение с одного языка на другой, то нам следует позаботиться о том, чтобы не допустить смешения языка. Предложение «это дерево старое», которое сформулировано на *языке вещей*, было бы неправильно перевести на язык событий в следующем виде: «Все события, составляющие это дерево, стары». Правильный перевод этого предложения будет таким: «Первое событие серии, составляющей это дерево, отделено большим отрезком времени от событий настоящего момента». Точно так же предложение: «Петр выше на два дюйма, чем он был год тому назад» — переводится следующим образом: «Событие настоящего момента из серии событий, составляющих Петра, занимает пространственный объем, одно из измерений которого на два дюйма длиннее, чем было одно из измерений соответствующего объема два года тому назад». Если перевод выполнен тщательно, то мы не встретим никаких особых трудностей и сможем избежать многих псевдопроблем, которые пытались обнаружить философы при рассмотрении изменяющихся и все же продолжающих существовать вещей [1].

Однако мы должны исследовать проблему определения такого типа генетического тождества, которое

[1] Обсуждение взаимоотношений между двумя языками см. Hans Reichenbach, Elements of Simbolic Logic, New York, Macmillan Co., 1947, § 48.

позволяет нам рассматривать серию событий как состояния одной и той же вещи. Мы видели, что генетическое тождество ограничено событиями, являющимися членами причинной цепи. Очевидно, что это условие недостаточно для определения генетического тождества. Событие может быть расположено в различных причинных цепях, которые в нем перекрещиваются, и в таком случае возникает вопрос, как выделить цепь физического тождества.

Для макроскопических объектов мы определяем *субстанциальное генетическое тождество* (material genidentity) с помощью различных характеристик, которые можно разделить на три группы.

Во-первых, мы связываем субстанциальное генетическое тождество с определенной *непрерывностью* изменения. При движении биллиардного шара имеет место непрерывный переход из одного положения в другое. В случае его столкновения с другим шаром мы едва ли согласимся с такой интерпретацией этого явления, по которой происходит обмен всех характеристик, составляющих физическое тождество данных объектов, потому что это потребовало бы скачкообразный переход одного шара на место другого. Такой скачкообразный переход был бы перерывом непрерывности, центр одного шара переместился бы в центр другого. Очевидно, что этот вывод возможен только потому, что наши пространственные наблюдения являются достаточно точными, чтобы провести различие между положениями шаров, расстояние между которыми мало по сравнению с их диаметрами.

Во-вторых, одной из характеристик материальных объектов является свойство, согласно которому пространство, занимаемое одним из них, не может быть занято другим. Два биллиардных шара занимают различные объемы пространства. При этом может возникнуть неясность; так, можно сказать, что гвоздь, вбитый в кусок дерева, будет находиться внутри пространства, занимаемого этим куском. Однако мы можем изменить определение пространства, заполненного деревом, так, чтобы исключить из него полость, заполненную гвоздем. Таким образом, свойство *занимать исключительное место в пространстве* всегда может рассматриваться как

необходимая характеристика твердых материальных объектов.

В-третьих, мы нашли, что когда бы материальные объекты ни обменивались своими местами, этот факт может быть отмечен. Обычно мы распознаем это изменение позиции, пользуясь специфическими отметками на объектах. Биллиардные шары можно различать по их окраске, или по небольшим царапинам на их поверхности, или при помощи точного измерения веса и т. д. Эти отметки остаются на объекте в соответствии с критерием непрерывности и позволяют идентифицировать его даже в случае отсутствия наблюдений в то время, когда происходило изменение его положения в пространстве и, следовательно, критерий непрерывности применить нельзя. Другими словами, *взаимный обмен положениями в пространстве представляет собой верифицируемое изменение* даже в том случае, когда в нашем распоряжении не имеется никаких протоколов, говорящих об этом факте.

Эти три характеристики субстанциального генетического тождества являются необходимыми критериями, однако они недостаточны. Когда дом снесен, то груда камней и балок представляет собой ту же самую массу вещества, что и дом, однако нельзя сказать, что эта груда и есть та же самая вещь.

Логическое определение дома требует некоторых взаимоотношений между его частями, которым больше уже не удовлетворяет груда камней [1]. Однако для следующих рассуждений потребуются только те необходимые критерии, которые мы только что сформулировали.

Некоторые исследования показывают, что не все отметки переносятся в соответствии с критерием непрерывности. Если катящийся биллиардный шар сталкивается с другим шаром, который покоится, то он может передать всю свою скорость этому шару и остановиться. Здесь отметка, состоящая в движении, передается другому шару, в то время как субстанциальное генетическое тождество не передается. Это позволяет нам сформулировать другую цепь генетического тождества: мы говорим, что кинетическая энергия перешла от одного

---

[1] См Hans Reichenbach, Experience and Prediction, p. 105.

шара к другому, и указываем этим выражением на другую физическую сущность — энергию, генетическое тождество которой подчиняется другим правилам. Это не субстанциальное тождество; его можно назвать *функциональным генетическим тождеством*, генетическим тождеством в широком смысле. Таким же функциональным является генетическое тождество распространяющихся в воде волн, которое отличается от субстанциального генетического тождества частиц воды; эти частицы движутся вверх и вниз, на что указывает плавающий в воде кусок дерева, тогда как волна движется в горизонтальном направлении.

Вещи, обладающие только функциональным генетическим тождеством, легко могут обходиться без второй и третьей характеристик, тогда как первая остается для них в силе. Два сгустка энергии могут существовать в одном пространственном объеме; иными словами, энергия удовлетворяет принципу суперпозиции. Энергию нельзя также маркировать. Когда два биллиардных шара, обладающих одинаковой скоростью, испытывают центральное столкновение, они возвращаются по той же траектории с теми же скоростями. Обменялись ли они энергиями, или каждый из них сохраняет свою первоначальную энергию, или произошло только изменение скоростей на обратные? Это бессмысленный вопрос, потому что не существует способа идентификации энергии. Следует отметить, что здесь нельзя получить ответ с помощью критерия непрерывности, хотя он и выполняется, потому что оба истолкования приводят к непрерывному генетическому тождеству. Мы сталкиваемся здесь с такой ситуацией, когда пересекаются две мировые линии, и после пересечения мы не можем отличить одну линию от другой. Мы можем попытаться найти выход из этой неопределенности, рассматривая шары, обладающие разными скоростями. Верно, что если один из шаров передвигается первоначально с большей скоростью, то эту скорость он передаст после столкновения другому шару. Однако не следует делать вывод, что произошел полный обмен энергиями; возможно, что произошла только передача излишка энергии от одного шара к другому. Такая неопределенность типична для функционального генетического тождества. Если имеется ясное понимание, что генетическое тождество может быть определе-

но различными путями, то эта неопределенность не порождает трудностей.

Существует ли такая вещь, как «истинное» генетическое тождество? Этот вопрос представляется нелепым. Мы можем дать такое определение «генетического тождества», которое будет удовлетворять нашим целям. Однако мы можем проводить сравнения между разными видами генетического тождества, и есть смысл отличать субстанциальное генетическое тождество биллиардных шаров от функционального генетического тождества перемещающейся кинетической энергии или отличать субстанциальное тождество частиц воды от функционального тождества волн. Основание для этого отличия состоит в том, что субстанциальное тождество можно определить с помощью трех упомянутых характеристик: непрерывности изменения, исключительности занимаемого положения в пространстве и заметного различия между макроскопическими объектами, когда они меняются местами.

При более тщательном анализе мы обнаруживаем, что часто называем субстанциальным генетическим тождеством функциональное генетическое тождество. Человеческое тело непрерывно меняет свою вещественную субстанцию; известно, что после семи лет ни одна из его частей не является в вещественном отношении той же самой, что и раньше. Вследствие этого человеческое тело удовлетворяет критерию непрерывности только приблизительно; точная проверка показала бы, что части развиваются независимо от развития тела в целом. Однако количество вещества, которое находится в процессе замены, в любой момент мало по сравнению с количеством вещества, не подверженным изменению в данный момент. Когда парусная шлюпка ремонтируется несколько раз таким образом, что наконец заменяется каждая из ее деталей, можно совершенно правомерно назвать ее той же самой лодкой, хотя критерию непрерывности она удовлетворяет только приблизительным образом. Вода в реке меняется все время, и река, таким образом, генетически тождественна скорее в функциональном отношении, чем в субстанциальном. Когда Гераклит осознал этот факт, он сделал вывод, что это изменение противоречит условиям «истинного» генетического тождества, и, следовательно, нельзя дважды вой-

ти в одну и ту же реку. Вместо этого ему следовало бы признать правомерность понятия функционального генетического тождества, на основании которого можно сказать, что в одну и ту же реку можно войти дважды.

Однако, даже если мы соглашаемся с последним выводом, мы все-таки склонны настаивать на том, что всегда можно проследить субстанциальное генетическое тождество, если соответствующим образом определить понятие генетического тождества. Если в живом теле происходит непрерывный обмен веществ, то частицы материи, совершающие кругооборот в этом теле, казалось бы, определяют направление субстанциального генетического тождества. Несмотря на то что данные частицы вещества подвергаются химическому распаду и превращаются в частицы других веществ, которые в свою очередь могут быть разделены посредством окисления или иных химических процессов или каким-то другим образом, представляется справедливым утверждение, что элементарные частицы, из которых составлены эти частицы вещества, определяют строгую линию субстанциального генетического тождества. Теория атомов основывалась именно на этом положении. Если макроскопическое генетическое тождество во всех отношениях является функциональным, то в основе его, в микрокосме, существует субстанциальное генетическое тождество. Атомы являются теми последними единицами, неизменность которых определяет линию субстанциального генетического тождества в физическом мире. Данная концепция является предпосылкой античного атомизма; и хотя современные атомы не являются более неразложимыми частичками материи, все-таки предполагается, что их элементарные составляющие способны играть ту же роль конечных физических сущностей, что и атомы Демокрита.

В частности, атомы помогают нам понять материальную природу таких вещей, как газы и жидкости. Вторая характеристика субстанциального генетического тождества оказывается неприменимой к двум перемешанным газам, поскольку они занимают одно и то же пространство. Путем разделения газа на мельчайшие части, на атомы или молекулы, его генетическое тождество можно свести к субстанциальному генетическому тождеству атомов. Так как характеристики генетического тожде-

ства твердого макроскопического тела сформулированы как свойства конечных частиц материи, то газообразные и жидкие агрегатные состояния вещества можно интерпретировать, сводя их к движению твердых атомов.

Возникает вопрос, в каком смысле можно говорить о субстанциальном генетическом тождестве элементарных частиц. В сфере атома обычные способы идентификации неприменимы. Мы не можем путем наблюдения проверить непрерывность движения, и хотя классическая физика и оставляет нам надежду, что когда-нибудь мы сможем осуществить эту проверку, однако с точки зрения квантовой механики, в силу соотношения неопределенностей, данное предположение становится физически бессмысленным. Ссылаться на вторую характеристику генетического тождества, на исключительность занимаемого положения в пространстве также нельзя, поскольку не существует эксперимента, с помощью которого можно было бы решить, состоят ли интерфеномены из волн или частиц. Если они представляют собой волны, тогда они удовлетворяют принципу суперпозиции. Наличие третьей характеристики также не может быть установлено с помощью наблюдения, а если это и можно сделать, то, во всяком случае, не так, как это делается в случае макроскопических объектов. Мы не можем наносить на атомы такие метки, которые сохранялись бы во время их перемещения, как это делает зоолог, кольцуя птиц, чтобы проследить путь их миграции. Правда, мы можем наблюдать траектории индивидуальных частиц в камере Вильсона. В физиологии радиоактивные атомы используются в качестве меток при исследовании распределения химических веществ в организме и их динамики. Однако вопрос заключается в том, прослеживаем ли мы в данном случае субстанциальное или только функциональное генетическое тождество. Как можем мы провести различие между ними, если макроскопические методы идентификации неприменимы?

Удовлетворительный ответ на этот вопрос дается в статистической термодинамике классической физики. Предполагается, что исследования свойств отдельных частиц следует заменить выводами, основанными на статистических свойствах группы частиц. Можно пояснить эту программу, использовав статистические методы,

рассмотренные в § 8, где мы произвели вычисления вероятностей, характеризующих различные распределения молекул. Там мы рассмотрели конфигурации специфических индивидуальных частиц, находящихся в разных ячейках или отделениях либо физического пространства, либо пространства некоторых параметров. Напомним, что распределение является классом конфигураций и что подсчет конфигураций зависит от того, как мы понимаем различие между ними. Было подчеркнуто, что в классической физике считаются различными две такие конфигурации, которые получаются при взаимном обмене молекулами между двумя данными ячейками. Например, если мы имеем две частицы, которые именуем a_1 и a_2, и если мы желаем расположить их в трех ячейках, то это может быть сделано девятью различными способами (см. табл. 1). Третья характеристика субстанциального генетического тождества, а именно требование, чтобы изменение пространственного расположения вело к наблюдаемым различиям, переведено, таким образом, на язык статистических свойств. Она обеспечивает определение термина «конфигурация».

ТАБЛИЦА 1

Статистика Максвелла — Больцмана
(частицы можно отличить друг от друга)

	Конфигурация								
	1	2	3	4	5	6	7	8	9
Ячейка 1	a_1 a_2	—	—	a_1	a_2	—	—	a_1	a_2
Ячейка 2	—	a_1 a_2	—	a_2	a_1	a_1	a_2	—	—
Ячейка 3	—	—	a_1 a_2	—	—	a_2	a_1	a_2	a_1

Этот метод расчета является характерным для статистики Максвелла — Больцмана. В (8.3) было показано, что если мы имеем n частиц и m ячеек или состояний, то число возможных конфигураций задается следующей формулой:

Статистика Максвелла — Больцмана: $N(n, m) = m^n$. (1)

Подобным же образом были получены другие статистические выводы. Например, формула (8.5) получается из данного определения конфигурации. Поскольку эта формула в свою очередь входит в формулу значения энтропии, а законы энтропии доступны экспериментальной проверке, то эти эксперименты дают возможность проверить правильность данного определения «конфигурации», а следовательно, и предположение о субстанциальном генетическом тождестве молекул. То же самое справедливо и для других физических выводов, полученных из определения «конфигурации» с помощью исчисления вероятностей.

Эта проверка давала положительный ответ в той сфере событий, которая изучается классической физикой, однако всякий раз, когда нельзя было больше пренебрегать квантовой природой элементарных событий, ответ получался отрицательный, например в том случае когда рассматривалось поведение газов при очень низких температурах. Отметим сперва, что в квантовых статистиках ячейки, в которых располагаются частицы, имеют особое физическое значение. Они задаются различными квантовыми состояниями, в которых может находиться частица и которые представлены собственными значениями, получающимися при решении уравнения Шредингера [1]. Например, возможные орбиты электронов в модели атома Бора представляют собой такие состояния. Точно так же энергия частиц газа разделяется на дискретные состояния. В статистике газов квантовые состояния получают следующую естественную интерпретацию. В § 10 мы рассматривали фазовое пространство, которое характеризуется $2un$ координатами, получающимися из таких трех параметров, как координаты и импульс $q_1, \ldots q_u$; $p_1, \ldots, p_u$, n частиц. Здесь произведение $q_i p_i$ соответствующих параметров имеет размерность ml^2t^{-1} (m — масса, l — длина, t — время), которая тождественна с размерностью действия, то есть произведения энергии на время. Это размерность кванта действия Планка, h. Таким образом, фазовое пространство естественным образом подразделяется на ячейки со стороной h^{un}, которые и являются отделениями в квантовой статистике. Эти квантовые ячейки

[1] См., например, PhF, p. 73.

выполняют функцию тех ячеек, на которые Больцман разделил фазовое пространство; однако в отличие от ячеек фазового пространства, они получаются не в результате произвольного разделения пространства, а даны, так сказать, некоторым абсолютным образом. Они определяются фундаментальными законами квантовой физики.

Когда статистические свойства квантовых процессов определяются с помощью квантовых состояний, то оказывается, что обычные методы Максвелла — Больцмана не приводят к соответствию с наблюдаемыми результатами. Бозе и Эйнштейном было показано, что отклонения вытекают из способа подсчета конфигураций и для того, чтобы получить статистические результаты, совместимые с экспериментальными данными, не следует рассматривать как различные такие две конфигурации, когда частицы из разных ячеек взаимно меняются своими местами. Это означает, что мы должны сделать вывод, что частицы нельзя отличить друг от друга и для них, следовательно, третья характеристика субстанциального генетического тождества не является справедливой. В таком случае данные, приведенные в табл. 1, должны быть заменены данными табл. 2, где мы опускаем нижние индексы у a, и, таким образом, указываем, что частицы неразличимы. Здесь мы имеем только шесть различных конфигураций для двух частиц в трех отделениях.

ТАБЛИЦА 2

Статистика Бозе — Эйнштейна
(частицы неразличимы)

	Конфигурация					
	1	2	3	4	5	6
Ячейка 1	a a	—	—	a	—	a
Ячейка 2	—	a a	—	a	a	—
Ячейка 3	—	—	a a	—	a	a

Обобщение этого способа вычисления на случай n частиц и m состояний или ячеек проводится следующим образом. Расположим m отделений в один ряд, тогда

они будут иметь $m-1$ разделяющих их внутренних стенок, или перегородок. Теперь, если мы хотим расположить n частиц некоторым произвольным образом в m отделениях, то мы помещаем частицы внутри этих ячеек, из которых состоит ряд; n-частиц и $m-1$-перегородок составляют ряд из $n+m-1$ объектов. Число их перестановок равно $(n+m-1)!$ Поскольку, когда меняются местами или частицы, или перегородки, конфигурация остается неизменной, то полученную величину нужно разделить на $n!\ (m-1)!$. Например, когда мы рассматриваем первую конфигурацию или первую вертикальную колонку таблицы, где две частицы расположены в одной ячейке, а две остальные ячейки пусты, то мы можем переставить две частицы и две внутренние перегородки. Тогда мы получим не только все другие конфигурации, но каждую в четырех разновидностях, число 4 получается как результат произведения 2! 2!. Таким образом, мы находим:

$$\text{Статистика Бозе — Эйнштейна: } N(n,\ m)=\binom{n+m-1}{n} \quad (2)$$

Этот метод подсчета применялся при вычислении всех термодинамических величин. Результаты получили хорошее экспериментальное подтверждение. Среди подтверждений, которые можно проверить путем наблюдения, можно упомянуть вычисление химической константы газа; кроме того, жидкий гелий подчиняется статистике Бозе, как это видно из странного поведения его удельной теплоемкости, математическую теорию которой разработал Ф. Лондон.

Существуют также другие физические процессы, которые требуют третьей формы статистики. Здесь частицы также неразличимы, но, кроме того, они подчиняются принципу запрета, впервые сформулированному В. Паули и распространенному на газовую статистику Э. Ферми и П. Дираком. Принцип запрета гласит, что в данном состоянии не может находиться более одной частицы. Этот вид статистики представлен табл. 3. Мы имеем здесь три разные конфигурации для двух частиц и трех отделений. Для n частиц и m состояний, или ячеек, мы получаем:

$$\text{Статистика Ферми — Дирака: } N(n,\ m)=\binom{m}{n}. \quad (3)$$

ТАБЛИЦА 3

Статистика Ферми — Дирака
(частицы неразличимы и подчиняются принципу запрета)

	Конфигурация		
	1	2	3
Ячейка 1	*a*	—	*a*
Ячейка 2	*a*	*a*	—
Ячейка 3	—	*a*	*a*

Эта формула получается следующим образом. Для одной частицы мы имеем m возможных конфигураций. Когда мы добавляем другую частицу, то, в силу принципа запрета, имеется только $m-1$ конфигураций, которые доступны для нее. Таким образом, мы получаем $m(m-1)$ конфигураций. Это следует разделить на 2!, поскольку взаимная перемена частиц местами оставляет конфигурацию неизменной. Этот метод легко распространить на случай n частиц.

Величины $N(n, m)$ также находят применение в различных статистиках. Они применяются при исследовании определенных групп физических явлений, например при исследовании электронного «газа».

Бесспорно, что неклассические статистические формулы (2) и (3) дают результаты, которые соответствуют наблюдаемым фактам. Вопрос заключается в том, что дают эти факты для решения проблемы генетического тождества. Обычно соглашаются с тем, что они говорят о неразличимости частиц, и мы сейчас подробно рассмотрим это положение. Нам нужно узнать, является ли это положение верным или же оно нуждается в уточнении.

Выше мы говорили, что субстанциальное генетическое тождество частиц не доступно изучению, исходя из свойств отдельных частиц, и что вместо такого изучения мы должны делать выводы на основании статистических свойств группы частиц. Какого типа выводы мы делаем в данном случае? Очевидно, что мы делаем выводы относительно интерфеноменов, поскольку генетическое тождество относится к физическому тождеству, которое

связывает один наблюдаемый феномен с другим. Следовательно, эти выводы предполагают определенные экстенсиональные правила. В данном случае мы не можем спрашивать, являются ли эти правила истинными; мы можем только поставить вопрос, к каким выводам мы придем, если будем использовать некоторые экстенсиональные правила. Экстенсиональное правило, предполагаемое в данной интерпретации, обычно широко используется в физике. Оно гласит, что все различимые конфигурации должны быть равновероятными. Отсюда следует, например, что в статистике Бозе различимыми являются только те конфигурации, которые представлены в табл. 2. Однако вопрос заключается в том, является ли это экстенсиональное правило единственно возможным и нельзя ли с помощью какого-нибудь другого правила вновь постулировать физическое тождество индивидуальных частиц.

На этот вопрос можно ответить следующим образом. Мы используем в качестве экстенсионального правила не равновероятность различимых расположений, а постулат, согласно которому частицы следует рассматривать как различимые в индивидуальном порядке.

Поскольку вероятностная метрика, которой мы должны следовать, задается, скажем, статистикой Бозе, то это правило приводит к неравным вероятностям различимых конфигураций. Посмотрим, что это означает.

Когда мы бросаем две монеты одной рукой, мы имеем четыре возможные конфигурации. Если мы обозначим «герб» через букву H, а «надпись» — через T и, кроме того, обозначим каждую монету подстрочной цифрой, то эти возможные конфигурации можно записать следующим образом:

$$H_1H_2,\ H_1T_2,\ T_1H_2,\ T_1T_2. \qquad (4)$$

В классической статистике предполагается, что эти четыре конфигурации равновероятны, и действительные наблюдения подтверждают это предположение, если проделано достаточное количество бросков. Частота выпадения «только герб» оценивается в $^1/_4$, частота выпадения «только надпись» также оценивается в $^1/_4$, однако выпадение «герб и надпись» оценивается в $^1/_2$, потому что оно содержит две возможные конфигурации. Предположим теперь, что статистика приводит к другим

результатам и что каждый из трех названных случаев происходит в отношении, равном $1/3$ от всего числа бросаний. В таком случае мы имеем статистику неиндивидуализированных комбинаций, то есть статистику Бозе. Она будет символизироваться тремя случаями:

$$HH,\ HT,\ TT, \tag{5}$$

в которых мы опускаем подстрочные цифры. Если повторяющиеся наблюдения подтверждают эту статистику, для которой три конфигурации, перечисленные в (5), будут равновероятными, то не будем ли мы вынуждены сказать, что монеты неразличимы?

Такое заключение было бы неправомерным. Так как легко можно различить монеты, например при помощи видимых отметок, то можно сказать, что в данном случае существуют четыре возможные конфигурации (4), однако они не равновероятны. Каждая первая и последняя конфигурация в формуле (4) будут иметь вероятность, равную $1/3$, а каждая из двух средних — равную $1/6$. При данных предпосылках три конфигурации формулы (5) были бы равновероятными. Такая интерпретация означала бы, что бросания монет не являются независимыми событиями. Если одна монета падает гербом вверх, то для другой монеты будет существовать тенденция к выпадению того же самого результата, точно так же как и для другой стороны монеты.

Обозначая событие «герб» для первой и второй монеты через B_1 и B_2 соответственно, мы получим

$$\begin{gathered} P(B_1)=\frac{1}{2},\ P(B_2)=\frac{1}{2}, \\ P(B_1,\ B_2)=\frac{2}{3}>P(B_2),\ P(B_2,\ B_1)=\frac{2}{3}>P(B_1) \end{gathered} \tag{6}$$

и соответственно

$$P(\bar{B}_1,\ B_2)=\frac{1}{3}<P(B_2),\ P(\bar{B}_2,\ B_1)=\frac{1}{3}<P_1(B_1). \tag{7}$$

Это означает, что эти два события как бы находятся во взаимной причинной зависимости. Последовательность событий, состоящая из бросаний одной монеты, находится в статистической, а следовательно, и причинной зависимости от другой последовательности. Не представляет особых затруднений сформулировать соответ-

ствующие соотношения для дизъюнкции n событий и двух последовательностей[1]. Если рассматриваются три или более последовательностей, то отношения зависимости становятся более сложными. Например, для трех последовательностей, атрибуты которых обозначаются буквами B, C и D, мы имеем

$$P(B, C) \neq P(C), \; P(C, B) \neq P(B),$$
$$P(B. C, D) \neq P(D), \; P(C. D, B) \neq P(B), \tag{8}$$

и подобные им соотношения, когда буквы меняются местами. Величины с двумя буквами перед запятой равны в этом случае величинам с одной буквой перед запятой. Это означает, что одна последовательность находится в зависимости не только от каждой другой, но также и от обеих, вместе взятых[2].

Мы получили следующий результат. Предположим, что мы могли бы приписать субстанциальное генетическое тождество каждой частице ансамбля Бозе, тогда бы оказалось, что движение одной частицы находится в зависимости от движения других частиц. Если одна частица находится в определенном состоянии, то для других существует тенденция прийти в это же состояние; а для комбинаций частиц, подобных тем, которые имеют место в формуле (8), существуют даже более сложные отношения зависимости. Такие причинные соотношения представляли бы собой дальнодействие, поскольку частицы могут находиться на значительном расстоянии друг от друга, то есть отношения их зависимости были бы каузальными аномалиями. Другими словами, любое приписывание физического тождества частицам Бозе приводит к каузальным аномалиям.

Мы видим теперь, каким образом можно дать четкую формулировку тезису о неразличимости частиц. Если мы хотим быть точными, то не можем просто сказать, что частицы неразличимы. Мы должны говорить: либо частицы неразличимы, либо в их поведении обнаруживаются каузальные аномалии. Мы отказываемся от выбора между первой или второй интерпретацией. Ни одна из них не является «более правильной»; обе представляют собой эквивалентные описания.

---

[1] См. ThP, p. 156.
[2] См. там же, стр. 105.

Однако только одно из двух описаний, а именно то, согласно которому частицы неразличимы, обеспечивает нормальную систему, то есть систему, свободную от каузальных аномалий. Если мы следуем обычному правилу: всякий раз, когда это возможно, пользоваться нормальной системой, то мы можем, следовательно, без колебаний сказать, что частицы неразличимы.

В статистике Ферми ситуация несколько иная. Эта статистика не создает каких-либо трудностей для интерпретации, согласно которой частицы различимы. В этой интерпретации каждую конфигурацию статистики Ферми следовало бы умножить на постоянный множитель $n!$, поскольку это есть число перестановок n частиц. Число всех возможных конфигураций, выраженное в формуле (3), умножается, таким образом, на $n!$, однако отношение равновероятности этих конфигураций не изменится. Это означает, что любые две конфигурации в статистике Ферми, которые являются теперь классами индивидуальных конфигураций, будут, по-прежнему, равновероятны.

Каузальная аномалия статистики Ферми заключается скорее в принципе запрета, если этот принцип интерпретировать как силы, действующие между индивидуальными частицами. Это означает, что если мы приписываем физическое тождество частицам, то запрет означает сильную взаимную зависимость отрицательного типа. Если одна частица находится в определенном состоянии, то в этом состоянии не может находиться никакая другая частица.

Эти соображения делают значительный вклад в решение проблемы детерминизма. Вопрос индуктивного доказательства данной проблемы обсуждался в § 25. Мы сделали определенные выводы, которые последовали бы, если бы гипотеза детерминизма была бы истинной, а затем обратили неправдоподобность этих выводов в индуктивное доказательство против этой гипотезы. Мы можем сформулировать подобную аргументацию и на основании нашего анализа статистики Бозе.

Если бы синоптический принцип квантовой механики был неверен, то мы смогли бы каким-то образом одновременно наблюдать координаты и импульс частицы и проверить, таким образом, ее физическое тождество при помощи характеристики непрерывности. Поскольку пред-

полагается, что частицы обладают свойством пространственной исключительности, то эта идентификация привела бы к субстанциальному генетическому тождеству. Однако мы уже убедились, что субстанциальное генетическое тождество частиц в статистике Бозе приводит к каузальным аномалиям.

Таким образом, мы еще раз оказались перед ситуацией, в которой каузальные аномалии превратились в свойства наблюдаемых. Этот вывод дает еще одно индуктивное доказательство в пользу справедливости синоптического принципа и против гипотезы детерминизма.

Напротив, если мы признаем синоптический принцип, то его комбинирование со статистикой Бозе приводит к выводу, что для элементарных частиц не существует субстанциального генетического тождества, при условии, что мы описываем мир с помощью нормальной системы, то есть системы, свободной от каузальных аномалий. Из этого следует, что в физическом мире вообще не существует субстанциального генетического тождества в определенном выше смысле, а существует только функциональное генетическое тождество. Понятие субстанциального генетического тождества представляет собой идеализацию поведения некоторых макроскопических объектов, а именно твердых тел, которое, однако, соответствует этой идеализации лишь приблизительным образом. Надежды найти подтверждение этой идеализации в области атома, распространив ее на свойства объектов атомных масштабов, которые представляли собой программу классической теории атомов, как мы убедились, не оправдались. В сфере атома субстанциальное генетическое тождество заменяется функциональным.

§ 27. Понятие энтропии в квантовой статистике

Теперь мы должны перейти к рассмотрению новой меры энтропии, которая вытекает из предположения о неразличимости частиц. Наше исследование имеет целью определить, при каких условиях квантово-статистическая и классическая меры энтропии дают макроскопически наблюдаемые различия и при каких условиях они приводят к неразличимым результатам.

Классическая статистика, как было показано в § 8, основывается на *структуре для частиц*, которая соответствует структуре (12.6) и может быть проиллюстрирована табл. 4. Каждая строка или горизонтальный ряд представляет историю индивидуальной частицы. Полагая число частиц равным n, мы получаем n рядов. Подразделив время на небольшие интервалы Δt, мы запишем для каждого последовательного момента t_1, t_2, ... мгновенное состояние частицы C_l. Таким образом, состояния C_l используются в этой структуре в качестве характеристик. Предположим пока, что все состояния относятся к одному и тому же уровню энергии, тогда для вертикальных столбцов не существует никаких ограничений. Два столбца тождественны, если все их члены последовательно соответствуют друг другу по характеристикам C_l. В этом случае они представляют одну и ту же конфигурацию. Вероятностная метрика статистики Максвелла-Больцмана задается предположением, что все столбцы или конфигурации равновероятны или встречаются в горизонтальных строках с одинаковой частотой.

Если нам известно, как часто встречается состояние C_l в каждом вертикальном столбце, то этим уже определено число заполнения n_l данного состояния. Распределение D задано для каждого столбца в том случае, если определено число заполнения каждого состояния. Два столбца относятся к одному и тому же распределению, если для каждого состояния числа заполнения n_l в обоих столбцах одни и те же.

Таким образом, если мы рассматриваем два столбца, в которых числа заполнения n_l и n_k двух состояний C_l и C_k поменялись местами, то эти столбцы относятся к разным распределениям. Исключение следует сделать для случая, когда $n_l = n_k$. Следовательно, распределение D тождественно с упорядоченным множеством $\{n_l\}$ чисел заполнения n_l, $l = 1, 2, \ldots m$ данного столбца. Порядок множества определяется соответствующим правилом, согласно которому происходит возрастание индекса l состояний C_l.

Когда термин «распределение» определяется согласно предыдущему абзацу, число конфигураций N_D, относящихся к одному и тому же распределению, задается формулой (8.5).

ТАБЛИЦА 4

Структура для частиц

(Каждая горизонтальная строка представляет историю частицы. Характеристиками служат состояния C, в которых частицы находятся в последовательные моменты $t_1, t_2, \ldots$)

	t_1	t_2	t_3	....
Частица 1	C_3	C_1	C_3	...
Частица 2	C_1	C_4	C_6	...
Частица 3	C_6	C_5	C_1	...
.....	.....	.....	.....	...
Частица n	C_4	C_1	C_6	...

Вспомним выкладки, проделанные в (8.3)—(8.8), и запишем *формулы, характерные для статистики Максвелла — Больцмана:*

(1) $$N_D = \frac{n!}{n_1! \ldots n_m!}, \quad N = m^n, \tag{2}$$

$$\sum_{i=1}^{m} n_i = n, \tag{3}$$

$$W = \frac{N_D}{N}, \tag{4}$$

$$S = \ln W = n \ln n - \sum_{i=1}^{m} n_i \ln n_i - n \ln m. \tag{5}$$

Мы используем здесь относительно нормированную энтропию, которая в точности соответствует вероятности W.

Если для обоснования статистики Бозе — Эйнштейна мы попытаемся использовать структуру для частиц, то обнаружим, что гипотеза относительно вероятностной метрики не подтверждается. Какие бы мы ни строили предположения относительно генетического тождества частиц, различные типы столбцов, содержащих тождественные конфигурации (в определенном выше смысле) будут встречаться с неодинаковой частотой. Следовательно, мы должны будем сказать, что между частицами суще-

ствует взаимная зависимость, свидетельствующая о существовании каузальных аномалий. Эта зависимость рассматривалась в связи с (26,6)—(26,7). В силу этих соображений структура для частиц неприменима в квантовой статистике.

Таким образом, для квантовой статистики мы вынуждены построить другую структуру, которую назовем *структурой состояний*. Эта структура иллюстрируется табл. 5. Каждая горизонтальная строка представляет историю состояния C_i. Предположим вновь, что имеется m состояний, тогда мы получим m рядов. В качестве характеристики используем число заполнения состояния n_i в последовательные моменты t_1, t_2, ... Так как, используя это число, мы не указываем на какие-то индивидуальные свойства частиц, то структура состояний пригодна для истолкования квантовой статистики. Мы не говорим здесь о вероятности того, что какая-то частица перейдет в некоторое состояние C_i, ибо такое представление основывается на структуре для частиц и приводит к каузальным аномалиям. Мы скорее говорим о вероятности того, что состояние C_i будет заполнено числом частиц, равным r.

Назовем два вертикальных столбца тождественными или скажем, что они относятся к одной и той же конфигурации, если в соответствующих местах этих столбцов будут стоять одинаковые числа заполнения. Это означает, что состояния рассматриваются как различимые индивидуальности; перестановка двух различных чисел заполнения в одной и той же колонке приводит к новой конфигурации. Так как число частиц постоянно, то возможны не все вообразимые конфигурации. Таким образом, числа заполнения каждого столбца подчиняются условию (3). Мы вновь будем рассматривать пока только такую ситуацию, когда все состояния относятся к одному и тому же уровню энергии. В этом случае (3) будет ограничивать только число столбцов. Те столбцы, которые в замкнутой системе отличны от 0, всегда будут удовлетворять этому условию. Вероятностная метрика квантовой статистики задается теперь предположением о равновероятности всех возможных столбцов структуры состояний. Иначе говоря, мы будем считать, что они встречаются в горизонтальной строке с одинаковой частотой.

Следует отметить, что конфигурацией мы именуем здесь то, что в случае статистики Максвелла — Больцмана называлось распределением. Величина N_D статистики Максвелла — Больцмана заменяется единицей. Поэтому, чтобы получить статистику и определить распределение как класс конфигураций, следует сгруппировать конфигурации соответствующим образом. Это достигается так. Возможные числа заполнения являются целыми числами $r = 0, 1, 2, \ldots n$; когда известно число заполнения состояния C_i, мы пишем $n_i = r$. Теперь мы можем подсчитать, сколько состояний характеризуется данным числом заполнения. Например, мы можем подсчитать, сколько состояний имеет число заполнения $r = 3$. Количество этих состояний обозначим через m_3. Может случиться, что состояний, характеризуемых числом заполнения 3, не существует, тогда мы получим $m_3 = 0$. Таким же образом, для каждого столбца устанавливается множество *частот заполнения* m_r, $r = 0, 1, 2, \ldots n$. Распределение D задано, когда известно это множество $\{m_r\}$. Следовательно, два столбца или две конфигурации относятся к одному и тому же распределению D, если они дают одно и то же множество $\{m_r\}$.

ТАБЛИЦА 5

Структура состояний

(Каждая горизонтальная строка представляет историю состояния. Характеристиками служат числа заполнения, то есть число частиц, находящихся в данном состоянии в последовательные моменты $t_1, t_2, \ldots$)

	t_1	t_2	t_3	....
Состояние 1	4	0	0	...
Состояние 2	2	3	4	...
Состояние 3	0	1	2	...
.....	...	...	...	...
Состояние m	4	0	1	...

Отметим, что частоты m_r, составляющие множество, упорядочены только в сторону возрастания r. Этот порядок не имеет отношения к тому порядку состояний C_i,

который зафиксирован в структуре. Поэтому распределение может содержать макроскопически различимые конфигурации. Допустим, что мы определяем состояния, разделяя сосуд, наполненный газом, на некоторое число ячеек, и рассматриваем распределение, характеризуемое довольно значительным числом незаполненных ячеек. Если мы предположим, что все незаполненные ячейки расположены в одной половине сосуда, то получим конфигурацию, при которой газ концентрируется в другой половине. Если же мы предположим, что незаполненные ячейки распределены беспорядочным образом, то полученная конфигурация будет подобна однородному распределению. Каждое из этих двух макроскопически различимых распределений характеризуется одним и тем же множеством $\{m_r\}$. В этом состоит отличие квантово-статистического распределения от классического, которое никогда не содержит макроскопически различимых конфигураций.

Прежде всего применим к статистике Бозе — Эйнштейна новое определение распределения. Поскольку существует m состояний, то каждый столбец содержит m чисел заполнения и они могут быть расположены $m!$ способами. Однако взаимная перестановка одинаковых чисел заполнения оставляет конфигурацию неизменной; следовательно, мы должны разделить $m!$ на число таких перестановок для каждой группы m_r. Таким образом, мы определяем число N_D по аналогии с (1), заменяя n на m. Используя этот способ и применив (26.2), мы запишем следующие *формулы, характерные для статистики Бозе — Эйнштейна:*

$$N_D = \frac{m!}{m_0! \ldots m_n!}, \qquad N = \binom{n+m-1}{n}, \tag{6}$$

$$\sum_{r=0}^{n} m_r = m, \qquad \sum_{r=1}^{n} r m_r = n, \tag{8}$$

$$W = \frac{N_D}{N}, \tag{9}$$

$$S = m \ln m - \sum_{r=0}^{n} m_r \ln m_r - \ln N, \tag{10}$$

$$\ln N = (n+m-1)\ln(n+m-1) - \\ - n \ln n - (m-1)\ln(m-1). \tag{11}$$

Вместо соотношения (3) мы получили соотношение (9). Отметим, что если мы начнем суммирование в соотношении (9) с $r=0$, то от этого ничего не изменится. В соотношении (8) формулируется новое условие, согласно которому число возможных состояний является постоянным для всех распределений. В (11)—(12) мы использовали формулу Стирлинга, которая дает приближенное выражение факториала, пренебрегая последним членом в формуле (8.6).

Для статистики Ферми — Дирака соответствующие формулы сильно упрощаются. Они получаются в том случае, когда суммирование в предыдущих формулах ограничено значениями $r=0$ и $r=1$. Кроме того, поскольку ячейка в данном случае может быть заполнена только одной частицей, то мы имеем $m_1=n$. Произведя эти подстановки, мы увидим, что формула (6) становится тождественной формуле (26.3), то есть величина N_D тождественна величине N. В самом деле, мы имеем здесь только одно возможное распределение. Эти соотношения приводят к следующим *формулам, характерным для статистики Ферми — Дирака:*

(13) $$N_D=\frac{m!}{m_0!m_1!}=\binom{m}{n}, \quad N=\binom{m}{n}, \tag{14}$$

(15) $$m_0+m_1=m, \quad m_1=n, \tag{16}$$

$$W=\frac{N_D}{N}=1, \tag{17}$$

$$S=m\ln m-m_0\ln m_0-m_1\ln m_1+\text{const}. \tag{18}$$

Поскольку $W=1$, то в данном случае $S=0$. Для того чтобы избежать подобного вывода, мы записали S с точностью до неопределенной постоянной. Позднее мы будем пользоваться только абсолютной энтропией, которая задается величиной $\ln N_D$. Для статистики Ферми — Дирака понятие энтропии практически имеет смысл только в том случае, когда рассматриваются различные уровни энергии, что будет показано в § 28.

Преимущество квантово-статистической структуры состояний в том, что для ее построения можно использовать методы классической статистики, поскольку столбцы здесь равновероятны и их вероятности можно

определить путем подсчета столбцов. Правда, горизонтальные строки этой структуры взаимозависимы, так как они связаны ограничивающим условием (9), которое можно также записать в виде формулы (3). Когда мы записываем числа заполнения в виде матрицы n^{kl}, то отношение (3) может быть переписано в виде условия, справедливого для каждого столбца

$$\sum_{k=1}^{m} n^{kl} = n. \tag{19}$$

Это условие выражает статистическую зависимость. Например, если матрицу n^{kl} первых $m-1$ членов столбца i суммировать до $n-1$, то значение n^{ml} последнего члена должно быть равно 1. Однако статистическая зависимость этого типа не представляет собой каузальной аномалии. Она только является следствием сохранения числа частиц и получает, таким образом, причинное объяснение. Следует проводить тщательное разграничение между этим видом зависимости и физической зависимостью, остающейся в силе для горизонтальных строк структуры для *частиц*, которая подчиняется статистике Бозе, о чем мы уже говорили в связи с (26.5)—(26.7).

Отметим для ясности, что статистическая взаимная зависимость горизонтальных строк структуры состояний не является следствием влияния движения одной частицы на другую. Данное положение можно уточнить следующим образом. Предположим, что мы пренебрегаем условием (19) и что структура удовлетворяет двум следующим условиям:

Условие 1. Все величины $r = 0, 1, \ldots n$ являются равновероятными для n^{kl} в каждой горизонтальной строке k.

Условие 2. Горизонтальные строки совершенно не зависят друг от друга.

Следствием этих предположений будет равновероятность всех столбцов. Среди этих столбцов найдутся такие, которые случайно будут удовлетворять условию (19). Теперь, когда мы будем рассматривать подструктуру столбцов, которые случайно удовлетворяют условию (19), то обнаружим, что все ее статистические свойства могут быть выведены из двух данных предположений, включая и равновероятность ее столбцов

и статистическую взаимную зависимость ее горизонтальных строк[1].

В самом деле, только этой подструктуре можно дать физическую интерпретацию. Как уже объяснялось, в квантово-механической структуре состояний, которая тождественна такой подструктуре, причинное взаимодействие между частицами не проявляется. Ее статистические свойства можно вывести на основании предположения о независимости столбцов и ограничивающих условий. Структуру, столбцы которой подчиняются ограничивающему условию, можно интерпретировать с помощью методов классической статистики. Действительно, структура частиц статистики Максвелла—Больцмана, удовлетворяющая условию для энергии (9.3), является структурой этого типа.

Для структуры, подчиняющейся условию (19), большое значение имеет вопрос о вероятностях w, величин $r = 0, 1, \ldots, n$ матрицы n^{kl}. В силу этого ограничения w_r для различных r не одинаково. Когда мы пишем r^{kl} для случая $n^{kl} = r$, эти вероятности определяются следующим образом:

$$P(r^{kl})^i = w_r. \tag{20}$$

Если распределение стационарно и m достаточно велико, то мы можем предположить, что те же самые вероятности справедливы и для вертикального столбца. Таким образом, мы имеем

$$w_r = P(r^{kl})^i = P(r^{kl})^k = \frac{m_r}{m}. \tag{21}$$

---

[1] В этом выводе мы используем правило редукции (см. ThP, p. 97) и рассматриваем подструктуру как подмножество $\binom{n+m-1}{n}$ возможных конфигураций, удовлетворяющих условию (19). Эти соображения можно распространить на структуру с бесконечным числом столбцов, заменяя условие (19) на условие

$$\lim_{m\to\infty} \frac{1}{m} \sum_{k=1}^{m} n^{kl} = q. \tag{19a}$$

Предполагая равномерную сходимость всех столбцов, мы можем применить эти выводы к структуре с бесконечным числом столбцов, используя свойства предельного перехода.

Значения w_r могут быть найдены с помощью комбинаторных методов.

Прежде чем приступить к этим вычислениям, сделаем замечание относительно статистики Максвелла — Больцмана. Конечно, для статистики Максвелла — Больцмана также можно составить структуру состояний. Однако ее столбцы не будут равновероятны. Это можно показать следующим образом. То, что мы называем конфигурацией в квантовой статистике, соответствует N_D конфигурациям в статистике Максвелла — Больцмана. Поскольку в статистике Максвелла — Больцмана все N_D конфигураций равновероятны, частота, с которой в структуре состояний встречаются отдельные вертикальные столбцы, не одинакова, ибо умножается на коэффициенты N_D, а величины N_D для разных вертикальных столбцов различны. Неравные вероятности разных столбцов делают структуру состояний непригодной для статистики Максвелла — Больцмана. Однако вероятности w_r здесь также существуют и управляют горизонтальными строками точно так же, как и столбцами, а соотношение (21) остается в силе также и для стационарного распределения статистики Максвелла — Больцмана.

Теперь перейдем к вычислению значений w_r для трех видов статистики. Для классической статистики мы должны возвратиться к метрике, которая определяет структуру частиц и приводит к неравным вероятностям в структуре состояний.

Предположим, что некоторому состоянию C_i мы приписываем r частиц. Пусть N_r будет полным числом конфигураций, которые сходны по тому признаку, что в состоянии C_i существует ровно r частиц. Это число определяется размещением оставшихся $n-r$ частиц в остальных $m-1$ ячейках. В статистике Максвелла — Больцмана мы должны, кроме того, учитывать, что из общего числа частиц n число частиц r в ячейке C_i можно разместить $\binom{n}{r}$ способами. Таким образом, используя (26.1) — (26.3), мы находим:

Статистика Максвелла — Больцмана:

$$N_r = \binom{n}{r} \cdot N(n-r,\ m-1) = \binom{n}{r}(m-1)^{n-r}, \qquad (22)$$

Статистика Бозе — Эйнштейна:

$$N_r = N(n-r,\ m-1) = \binom{n-r+m-2}{n-r}. \qquad (23)$$

Статистика Ферми — Дирака:

$$\begin{cases} N_0 = N(n,\ m-1) = \binom{m-1}{n}, \\ N_1 = N(n-1,\ m-1) = \binom{m-1}{n-1}. \end{cases} \qquad (24)$$

В силу предположения, что структура имеет вероятностную метрику, величина w_r формулы (20) задается соотношением

$$w_r = \frac{N_r}{N}. \qquad (25)$$

После подстановки этих величин из формул (20)—(22) мы будем использовать методы апроксимации, которые применимы в том случае, если как m, так и n велики по сравнению с r. Мы используем здесь следующие формулы, которые справедливы, когда n велико по сравнению с x или соответственно r:

(26) $$\left(\frac{n+x}{n}\right)^n \sim e^x, \quad \binom{n}{r} \sim \frac{n_r}{r!}. \qquad (27)$$

Когда мы называем формулу (27) апроксимацией, то это означает, что при возрастании n частное с той и другой стороны отношения стремится к 1. Вывод формулы (27) также не составит затруднений, если мы записываем $\binom{n}{r}$ в форме $n(n-1)\ldots(n-r+1)/(r!)$ и пренебрегаем степенями, меньшими чем n^r. Кроме того, мы заменяем $m-1$ на m и $n-r$ на r, если они не встречаются в n-й степени.

Выполнив эти апроксимации [1], мы находим:

[1] Точные значения формул (28) и (29) можно записать соответственно следующим образом:

(28a) $$w_r = \binom{n}{r}\left(\frac{m-1}{m}\right)^n \cdot \frac{1}{(m-1)^r};$$

$$w_r = \frac{m-1}{u-r} \cdot \frac{\binom{n}{r}}{\binom{u}{r}}, \quad u = n+m-1. \qquad (29a)$$

Для формулы (30) приведено точное значение.

Статистика Максвелла — Больцмана:

$$w_r = \frac{1}{r!} e^{-q} q^r; \tag{28}$$

Статистика Бозе — Эйнштейна:

$$w_r = \frac{1}{1+q} \cdot \left(\frac{1}{1+q}\right)^r; \tag{29}$$

Статистика Ферми — Дирака:

$$\begin{cases} w_0 = 1 - q, \\ w_1 = q; \end{cases} \tag{30}$$

везде в этих формулах

$$q = \frac{n}{m}. \tag{31}$$

Величина q обозначает среднее число частиц на состояние. Для статистики Ферми — Дирака q всегда меньше или равно 1, для двух других статистик q может быть больше 1.

Что касается формулы (28), то в ней нас, в частности, интересует значение r_{max} величины r, для которой w_r является максимальной. Подставив в нее $\frac{dw_r}{d_r} = 0$, мы легко находим, что $r_{max} = q$. Это означает, что наиболее вероятное число заполнения определяется соотношением n/m.

Одинаковое распределение всех частиц по всем состояниям представляет, следовательно, согласно формуле (8.14), наиболее вероятное распределение.

Для того чтобы проанализировать формулу (29), положим:

$$\frac{q}{1+q} = w, \quad \frac{1}{1+q} = 1 - w, \tag{32}$$

тогда, используя (3), можно записать формулу (29) следующим образом:

Статистика Бозе — Эйнштейна:

$$w_r = (1 - w) w^r, \quad w = \frac{n}{n+m}. \tag{33}$$

Полагая r равным 0, мы находим, что $w_0 = 1 - w$. Это и есть вероятность обнаружения незаполненного со-

стояния, а следовательно, и максимальная вероятность любого числа заполнения, поскольку $w < 1$. Таким образом, мы имеем $r_{max} = 0$. Незаполненное состояние наиболее вероятно, и следствием статистики Бозе является то, что некоторые состояния заполнены большим количеством частиц. Однако все же, может быть, более вероятно, что состояние заполняется по крайней мере одной частицей, чем остается незаполненным. Вероятность того, что состояние заполнено по крайней мере одной частицей, равна w. Это вытекает из того, что $1 - w$ есть вероятность незаполненного состояния. Но это также согласуется с выражением суммы геометрической прогрессии, если мы используем формулы

$$\sum_{r=0}^{\infty} w^r = \frac{1}{1-w}, \quad \sum_{r=1}^{\infty} w^r = \frac{w}{1-w} \tag{34}$$

и при суммировании по всем возможным значениям w_r положим n равным бесконечности.

Для статистики Ферми — Дирака единственно возможным числом заполнения заполненного состояния является 1. Следовательно, здесь q есть вероятность того, что состояние заполнено. Вопрос, будет ли заполненное состояние более вероятным, чем незаполненное, и равно ли r_{max} 0 или 1, зависит от того, больше или меньше $^1/_2$ величина q.

При сравнении трех формул (28)—(30) между ними обнаруживается специфическое различие. Согласно формуле, определяющей максимальное значение w_r в статистике Максвелла — Больцмана, полученной из формулы (28), все состояния, или ячейки, прекрасно могут быть одновременно заполнены одинаковым числом частиц $r_{max} = n/m$; равное распределение всех частиц по всем отделениям является, таким образом, наиболее вероятным распределением. Однако для двух других статистик недопустимо, чтобы все отделения в одно и то же время имели число заполнения r_{max}. В статистике Бозе — Эйнштейна, где $r_{max} = 0$, недопустимо, чтобы все отделения были не заполнены, кроме случая, когда вообще не существует никаких частиц. Для статистики Ферми — Дирака ни значение $r_{max} = 0$, ни значение $r_{max} = 1$ не может быть справедливым для всех ячеек, если только не выполняются условия, когда $n = 0$ или $n = m$, что является

исключительным случаем. Следовательно, статистическое равновесие в квантовых статистиках задается не равным распределением всех частиц по всем отделениям, а характеристической вариацией чисел заполнения, распределение которых удовлетворяет вероятности w_r. Хотя числа заполнения состояний и меняются со временем, частоты заполнений при стационарном распределении остаются неизменными.

Если бы мы смогли когда-нибудь непосредственно вычислить числа заполнения состояний, то обнаружили бы заметные различия между тремя видами статистик. Можно показать, однако, что эти различия уменьшаются, когда уменьшается q — среднее число частиц на состояние.

Если $q \ll 1$, то мы находим, заменяя в формулах (28) и (29) $1 + q$ на 1, а e^{-q} на $1/(1 + q)$:

$$\frac{w_r\,(Б-Э)}{w_r\,(М-Б)} \sim r! \tag{35}$$

Теперь мы имеем $r! = 1$ для $r = 0$ и $r = 1$.

Из этого следует, что если q мало, то два числа заполнения 0 и 1 имеют одинаковые частоты в статистиках Бозе — Эйнштейна и Максвелла — Больцмана.

Подобные равенства получаются и при сравнении статистик Максвелла — Больцмана и Ферми — Дирака. Если q мало, то мы, подставив в формулу $1 - q$ вместо e^{-q}, находим для $r = 0$ и $r = 1$:

$$\frac{w_0\,(Ф-D)}{w_0\,(М-Б)} \sim 1, \quad \frac{w_1\,(Ф-D)}{w_1\,(М-Б)} \sim 1. \tag{36}$$

Только при числах заполнения $r > 1$ различия между тремя видами статистик становятся заметными. Однако, если q мало, частоты этих чисел заполнения зависят от q^2, q^3, ... и гораздо меньше частот чисел заполнения, равных 0 и 1. В частности, принцип запрета статистики Ферми — Дирака для низких вероятностей независимых событий фактически всегда выполняется, поскольку можно пренебречь величинами q^2, q^3 и т. д. по сравнению с q.

Три вида статистик могут быть проиллюстрированы числовым примером, представленным в табл. 6, где величины n и m малы.

ТАБЛИЦА 6

Сравнение статистик Максвелла — Больцмана (М — Б), Бозе — Эйнштейна (Б — Э) и Ферми — Дирака (Ф — Д)

Статистика	$n=2$			$m=3$				$n=2$			$m=10$			
	N	N_0	w_0	N_1	w_1	N_2	w_2	N	N_0	w_0	N_1	w_1	N_2	w_2
М — Б	9	4	0,44	4	0,44	1	0,11	100	81	0,81	18	0,18	1	0,01
Б — Э	6	3	0,50	2	0,33	1	0,13	55	45	0,818	9	0,164	1	0,02
Ф — Д	3	1	0,33	2	0,67	0	0	45	36	0,80	9	0,20	0	0

Обозначения:

n — общее число частиц;
m — число состояний, или ячеек;
N — общее число возможных конфигураций;
N_r — число конфигураций, при которых определенное отделение заполнено числом частиц, равным r;
w_r — вероятность нахождения r частиц в определенной ячейке $= N_r/N$.

Величины, приведенные для N и N_r, вычислены по формулам (2), (7), (14) и (22)—(24), а величина w_r — по формуле (25). Левая часть таблицы соответствует $n = 2$, $m = 3$, как и в табл. 1—3. Здесь различия величин w_r вполне наглядны. Правая часть таблицы соответствует той же самой величине $n = 2$, однако $m = 10$. Таким образом, переход из левой части таблицы в правую соответствует возрастанию числа возможных состояний при постоянстве числа частиц. Цифры в правой части таблицы показывают, что здесь величины w_r фактически одинаковы для всех трех видов статистик.

В результате мы получили, что переход к малой величине q и, следовательно, размешивание газа по всем возможным состояниям приводят к совпадению трех видов статистик. Физическую интерпретацию полученных результатов можно отложить до тех пор, пока мы не исследуем влияние изменений в уровне энергии на распределение (§ 28).

Сравнивая значение энтропии в классической статистике (5) со значениями энтропии (11) и (18) в квантовых статистиках, укажем на некоторые важные различия. В этом разделе уже упоминалось, что в квантовых статистиках мы называем конфигурацией то, что соответствует распределению в статистике Максвелла —

Больцмана. В последней статистике квантовая конфигурация умножалась бы на величину N_p, которая задается формулой (1). Используя частоты заполнения m_r, мы могли бы переписать формулу (1) в следующем виде:

Статистика Максвелла — Больцмана:

$$N_D = \frac{n!}{(0!)^{m_0}(1!)^{m_1}\ldots(r!)^{m_r}\ldots(n!)^{m_n}}. \qquad (37)$$

Мы видим, что все квантовые конфигурации, относящиеся к одному и тому же квантовому распределению, которое дано множеством $\{m_r\}$, умножаются на один множитель, а именно — на величину N_D, полученную в формуле (37). Однако для различных квантовых распределений, то есть для различных множеств $\{m_r\}$, этот множитель меняет свою величину.

Если бы мы заменили значение энтропии в квантовой статистике, как оно сформулировано в выражениях (11) и (18), на значение энтропии в статистике Максвелла — Больцмана, то мы должны были бы ввести в формулы (11) и (18) величину $\ln N_D$, вычисленную согласно формуле (37). Но, поскольку (37) меняется с изменением множества $\{m_r\}$, полученные максимальные значения энтропии были бы различными. В этом причина того, что два вида квантовой статистики приводят к различным формам стационарного распределения, то есть распределения, характеризуемого максимумом вероятности. Однако, несмотря на различие, касающееся статистической структуры состояния равновесия, эти два вида энтропии имеют одинаковую логическую природу. Эти два вида энтропии суть функции, характеризующие состояние макроскопической системы, функции, математическая форма которых выражается посредством вероятностей и которые характеризуют процесс развития в направлении состояния равновесия как тенденцию к переходу от состояний с меньшей вероятностью к состояниям с более высокой вероятностью.

Существует, однако, отличие, которое позволяет нам отдать предпочтение квантово-статистической форме энтропии по сравнению с классической формой энтропии. В § 8 мы рассматривали парадокс Гиббса и показали, что в классической статистике абсолютная энтропия не может быть нормирована удовлетворительным образом. Иначе обстоит с квантово-статистической энтропией;

здесь в математической формуле энтропии парадокс Гиббса автоматически элиминируется.

Напомним, что этот парадокс возникает в связи с рассмотрением общей энтропии двух сосудов, наполненных газом, который находится в одинаковых условиях. При открывании крана, соединяющего оба сосуда, энтропия, которая выражается суммой энтропий обоих сосудов, не должна меняться. Эта проблема уже обсуждалась при рассмотрении формулы (8.17).

В статистике Бозе эта проблема решается следующим образом. Если в уравнении энтропии (11) статистики Бозе — Эйнштейна мы заменим m на $2m$, а n — на $2n$, то частоты заполнения m_r также удвоятся. Таким образом, пренебрегая постоянной $-\ln N$, мы находим

$$S' = 2m \ln 2m - \sum_{r=1}^{2n} 2m_r \ln(2m_r) =$$

$$= 2m \ln m + 2m \ln 2 - 2\sum_{r=1}^{2n} m_r \ln m_r - 2m \ln 2 = 2S. \qquad (38)$$

Здесь мы не обращаем внимания на то, что суммирование до n заменено суммированием до $2n$, поскольку большие значения r не оказывают заметного влияния на сумму. Полученный результат $2S$ также будет справедливым, если энтропия (11) содержит аддитивную постоянную. При этом постоянная должна иметь форму $n \cdot \mathrm{const}\,A$, как и в статистике Максвелла — Больцмана. Мы видим, что здесь член $2m \ln 2$ можно сократить, комбинируя первое и второе слагаемые, в противоположность (8.17), где соответствующий член $2n \ln 2$ не сокращается, если пренебречь третьим членом в (8.8). Таким образом, мы нашли, что абсолютная энтропия удовлетворяет постулату аддитивности. То же справедливо и для статистики Ферми — Дирака. Для квантовой статистики парадокса нормировки не существует.

Физическое объяснение не представляет каких-либо затруднений. Когда мы открываем кран, соединяющий сосуды, наполненные газом, который находится в одинаковых условиях, то величина N_D возможных конфигураций для состояния равновесия в статистике Максвелла — Больцмана чудовищно возрастает, потому что молекулы газа из одного сосуда могут поменяться местами с молекулами газа из другого сосуда. Однако, если

частицы неразличимы, тогда такое изменение не имеет значения, и число конфигураций остается тем же самым, как и до открывания крана. При этом условии абсолютным числом конфигураций свободно можно пользоваться как мерой энтропии. Как и в случае относительной энтропии в статистике Максвелла — Больцмана, равенство энтропии до и после открывания крана сохраняется только приближенно, потому что после того, как кран открыт, частицы могут сосредоточиться в одном сосуде. Однако этой возможностью можно пренебречь. При выводе формулы (3) эта апроксимация вновь вытекает из того, что можно пренебречь последним членом в формуле Стирлинга (8.6).

Математическое требование определения абсолютной энтропии, удовлетворяющей постулату аддитивности, достигается только в квантовой физике[1]. Следовательно, только здесь может быть удовлетворительно интерпретирован третий закон Нернста, утверждающий, что при температуре, равной абсолютному нулю, энтропия также равна нулю, потому что в квантовой статистике можно удовлетворительно провести абсолютную нормировку энтропии.

Парадокс Гиббса остается в силе и для двух нетождественных газов, которые сначала были разделены, а затем слились в процессе диффузии. В этом случае абсолютная величина энтропии по статистике Максвелла — Больцмана возрастает точно до значения (8.17). В случае, если газы одинаковые, последует точно такой же результат, однако для предельного случая тождественных газов нормировка оставляет энтропию неизменной. В этом факте находит свое выражение прерывность явлений природы. В квантовой физике разрыв непрерывности можно обойти, используя понятие *частичного подобия*; будет существовать непрерывное изменение величины энтропии по мере того, как газы будут становиться все более и более подобными. Эти соображения были недавно использованы А. Ланде для вывода из посту-

---

[1] Этот вывод подчеркнул недавно Альфред Ланде в «Квантовой механике» (A. Landé, Quantum Mechanics, New York, Pitman Publ. Corp., 1951, p. 219, 227). Отсылаем читателя к этой книге, где блестяще изложена физическая сторона квантовой статистики и квантовой механики вообще.

лата непрерывности[1] необходимости квантовых скачков, подчиняющихся вероятностным закономерностям, и других квантово-механических принципов.

§ 28. Применение квантовой статистики в случае различных уровней энергии

Рассмотрим теперь расширение понятия энтропии на случай с газом, находящимся в состояниях с различными уровнями энергии. В § 9 данная проблема обсуждалась в рамках статистики Максвелла — Больцмана. Подведем кратко итоги этого обсуждения. Обозначая через ε_i величину энергии, соответствующую состоянию C_i, добавим энергетические условия к соотношениям (27.1)—(27.5) в следующей форме:

$$\sum_{i=1}^{n} \varepsilon_i n_i = E. \tag{1}$$

Тогда, как уже пояснялось в связи с рассмотрением формулы (9.4), постоянную N следует заменить меньшей, но неизвестной постоянной N_E.

Записав выражение энтропии с помощью неопределенной постоянной, мы заменим, таким образом, (27.5) на

$$S = n \ln n - \sum_{i=1}^{m} n_i \ln n_i + \text{const}. \tag{2}$$

Мы видим, что учет различных уровней энергии изменяет в значении S только постоянную. Кроме того, S подчиняется теперь ограничительному условию (1), которое добавляется к (27.3).

Если мы интересуемся максимумом S, то постоянная в (2) не существенна. Следовательно, безразлично, что мы используем относительную или абсолютную энтропию. Расчеты, выполненные в § 9, приводят к результату (9.11), который здесь можно привести:

$$q(\varepsilon) = \frac{n(\varepsilon)}{m(\varepsilon)} = ae^{-\varepsilon/kT}. \tag{3}$$

[1] См. A. Landé, Quantum Mechanics and Thermodynamic Continuity, в „Amer. Jour. Phys.", 1952, vol. 20, p. 353—358; его же, Thermodynamic Continuity and Quantum Principles, в „Phys. Rev.", 1952, vol. 87, p. 267—271; и его недавно вышедшую книгу „Foundations of Quantum Theory: A Study in Continuity and Symmetry", New Haven, Conn., Yale Univ. Press, 1955.

Эта формула выражает среднее число частиц на состояние для данного уровня энергии в виде экспоненциальной зависимости от температуры газа в целом. Другими словами, определенное число частиц $n(\varepsilon)$ из всего числа частиц n переходит в состояния с уровнем энергии ε; при низких температурах число частиц $n(\varepsilon)$ в состояниях с высоким уровнем энергии мало, однако с возрастанием температуры газа оно увеличивается. Постоянную a можно определить только в том случае, когда энергия задается как функция параметров. С ростом температуры она уменьшается, как показано для $b = a/n$ в (9.15).

Структура состояний, для которой мы определили величину w_r, должна быть построена теперь отдельно для каждого уровня энергии, и формула (27.28) применяется в том случае, если мы используем значение q, относящееся к тому уровню энергии, которое получено в формуле (3).

В квантовой статистике, где энтропия определяется относительно структуры состояний, а не относительно структуры для частиц, как в статистике Максвелла — Больцмана, нам следует начинать рассуждать с построения модели структуры состояний для каждого уровня энергии. Число конфигураций N_D, определяемое формулой (27.6), нужно теперь применять к каждому уровню энергии. Предположим, что возможные уровни энергии задаются величинами $\varepsilon^{(1)}$, $\varepsilon^{(2)}$, ... $\varepsilon^{(\alpha)}$. Все величины, ассоциируемые с некоторым уровнем энергии $\varepsilon^{(s)}$, могут быть обозначены при помощи соответствующих индексов. Тогда общее число конфигураций N_D задается произведением

$$N_D = N_D^{(1)} \ldots N_D^{(\alpha)}. \tag{4}$$

Соответственно выражения энтропии в (27.11) и (27.18) заменяются формулами суммирования. Поскольку здесь общее число возможных конфигураций N вновь следует заменить неопределенной величиной N_E, мы введем в меру энтропии неизвестную постоянную.

Хотя эти соображения применимы как к статистике Бозе — Эйнштейна, так и к статистике Ферми — Дирака, мы рассмотрим пока только первую. Запишем общее выражение энтропии газа с различными уровнями энергии

для статистики Бозе — Эйнштейна следующим образом:

$$S=\sum_{s=1}^{u} m^{(s)} \ln m^{(s)}-\sum_{s=1}^{u} \sum_{r=1}^{n^{(s)}} m^{(s)} \ln m_r^{(s)}+\text{const}. \qquad (5)$$

Если мы считаем, что аддитивная постоянная не зависит от m, то в этом случае мы имеем абсолютно нормированную энтропию, которая измеряется N_D, а не вероятностью W. Ограничительные условия для S имеют теперь следующую форму:

$$(6) \qquad \sum_{r=0}^{n} m_r^{(s)}=m^{(s)}, \qquad \sum_{s=1}^{u} \sum_{r=1}^{n} r m_r^{(s)}=n, \qquad (7)$$

$$\sum_{s=1}^{u} \varepsilon^{(s)} \sum_{r=1}^{n} r m_r^{(s)}=E. \qquad (8)$$

Условие (6) заменяет собой u аналогичных условий, каждое из которых выражает тот факт, что суммирование ячеек $m_r^{(s)}$ некоторого энергетического уровня $\varepsilon^{(s)}$ приводит к величине $m^{(s)}$, являющейся постоянной для данного уровня энергии. Каждое из двух других условий (7) и (8) выражает только одно отношение, хотя и состоит из двух сумм. Отметим, что суммирование по r в обоих соотношениях дает величину $n^{(s)}$ данного уровня энергии $\varepsilon^{(s)}$, но поскольку величина $n^{(s)}$ не постоянная для данного уровня энергии, то можно требовать только, чтобы суммирование $n^{(s)}$ по s приводило к величине n. Это условие сформулировано в соотношении (7).

Максимум S в формуле (5) находится, как и в § 9, при помощи вариационного метода множителей Лагранжа. Поскольку формула (6) представляет u уравнений, она содержит u множителей $\gamma^{(s)}$. Обозначая через α множитель для формулы (7), а через β — множитель для формулы (8), мы находим, дифференцируя (5)—(8) по $m_r^{(s)}$.

$$1+\ln m_r^{(s)}+\gamma^{(s)}+\alpha r+\beta \varepsilon^{(s)} r=0. \qquad (9)$$

Эти $u \cdot n$ уравнений для величин $m_r^{(s)}$ могут быть записаны, если перейти от логарифмической к экспоненциальной форме, следующим образом:

$$m_r^{(s)}=g^{(s)} w^r,$$
$$g^{(s)}=e^{-1-\gamma^{(s)}}, \quad a=e^{-\alpha}, \quad w=a \cdot e^{-\beta \varepsilon^{(s)}}. \qquad (10)$$

Для краткости мы пишем w вместо $w^{(s)}$. Для того чтобы определить постоянную $g^{(s)}$, мы используем условие (6), полагая при суммировании n равным бесконечности и применяя формулу (27.34):

$$\sum_{r=0}^{\infty} m_r^{(s)} = m^{(s)} = g^{(s)} \sum_{r=0}^{\infty} w^r = g^{(s)} \cdot \frac{1}{1-w}, \qquad (11)$$

$$g^{(s)} = m^{(s)}(1-w).$$

Используя соображения, применявшиеся при выводе формул (9.8) — (9.10), мы находим, что $\beta = \frac{1}{kT}$. Вычислим теперь величину, упомянутую выше, $n^{(s)}$, выражающую число частиц, энергия которых равна $\varepsilon^{(s)}$:

$$n^s = \sum_{r=1}^{\infty} r m_r^{(s)} = m^{(s)}(1-w) \sum_{r=1}^{\infty} r w^r = m^s \cdot \frac{w}{1-w}. \qquad (12)$$

В (12) мы используем формулы (10) и (11) и формулу суммы геометрической прогрессии; эта третья формула, которая легко выводится путем дифференцирования формулы (27.34) по переменному w, имеет следующий вид:

$$\sum_{r=0}^{\infty} r w^r = \sum_{r=1}^{\infty} r w^r = \frac{w}{(1-w)^2}, \quad 0 \leqq w < 1. \qquad (13)$$

Если мы теперь подставим (11) в (10) и решим (12) относительно w, то найдем, определяя $w_r^{(s)}$ по аналогии с (27.19):

$$w_r^s = \frac{m_r^{(s)}}{m^{(s)}} = (1-w) w^r, \quad w = \frac{n^{(s)}}{n^{(s)} + m^{(s)}}. \qquad (14)$$

Таким образом, мы нашли, что формула (27.33) справедлива в отдельности для каждого уровня энергии $\varepsilon^{(s)}$. Формула, которую мы вывели в виде (27.29) из чисто комбинаторных соображений, справедлива теперь только для максимума вероятности, то есть для стационарных условий. Поэтому только в стационарном состоянии $n^{(s)}$ будет иметь значение, заданное в (12).

Однако эти выкладки позволили нам получить важное дополнение к формулам (27.29) и (27.33). Они дают

нам величину w или q в виде функции от уровня энергии ε. Для того чтобы написать это уравнение, мы комбинируем величину w из формулы (10) с величиной w из формулы (14) и пишем $w(\varepsilon)$, заменяя везде дискретный индекс (s) на непрерывный аргумент (ε). Используя отношение (27.32) между w и q, мы, таким образом, находим:

$$\frac{n(\varepsilon)}{n(\varepsilon)+m(\varepsilon)}=\frac{q(\varepsilon)}{1+q(\varepsilon)}=w(\varepsilon)=ae^{-\varepsilon/kT}. \qquad (15)$$

В статистике Бозе — Эйнштейна это соотношение выражает зависимость числа частиц $n(\varepsilon)$ или связанных с ним чисел $q(\varepsilon)$ и $w(\varepsilon)$ от уровня энергии. Оно соответствует соотношению (3) в статистике Максвелла — Больцмана. Подобно последнему соотношению, оно выражается экспоненциальной зависимостью, хотя и в несколько иной форме. Однако если $q(\varepsilon)$ мало, то мы имеем $q(\varepsilon)\sim w(\varepsilon)$, и отношение (15) тогда фактически тождественно с (3). Постоянная a в таком случае также одна и та же для обоих соотношений, потому что она определяется суммированием экспоненциального выражения по всем уровням энергии. Отметим, что постоянная a содержит в себе множитель n. Кроме того, a зависит от температуры T, что иллюстрируется формулой (9.5), ибо $b=a/n$.

Соответствующие выкладки можно проделать и для статистики Ферми — Дирака. Используя (27.13)—(27.18) и суммируя по различным уровням энергии, мы находим:

$$S=\sum_{s=1}^{u} m^{(s)}\ln m^{(s)}-\sum_{s=1}^{u} m_0^{(s)}\ln m_0^{(s)}-$$

$$-\sum_{s=1}^{u} m_1^{(s)}\ln m_1^{(s)}+\text{const.}, \qquad (16)$$

(17) $$m_0^{(s)}+m_1^{(s)}=m^{(s)}, \quad m_1^{(s)}=n^{(s)}, \qquad (18)$$

(19) $$\sum_{s=1}^{u} m_1^{(s)}=n, \quad \sum_{s=1}^{u} \varepsilon^{(s)}m_1^{(s)}=E. \qquad (20)$$

Из условия максимальности значения S определим соотношения (9) и (10). Однако они представляют здесь только $2u$ соотношений.

Вместо (11) мы получим

$$m_0^{(s)} + m_1^{(s)} = m^{(s)} = g^{(s)}(1 + w),$$
$$g^{(s)} = \frac{m^{(s)}}{1 + w}. \tag{21}$$

Вместо формулы (12) мы теперь имеем

$$n^{(s)} = m_1^{(s)} = m^{(s)} \frac{w}{1 + w}. \tag{22}$$

Подставляя эти значения в (10) и полагая соответственно $r = 0$ и $r = 1$, используя при этом (18) и (27.31), находим

$$w_0^{(s)} = \frac{m_0^{(s)}}{m^{(s)}} = \frac{1}{1 + w} = 1 - q^{(s)}, \tag{23}$$

$$w_1^{(s)} = \frac{m_1^{(s)}}{m^{(s)}} = \frac{n^{(s)}}{m^{(s)}} = \frac{w}{1 + w} = q^{(s)}. \tag{24}$$

Таким образом, по аналогии с (14), мы нашли, что соотношения (27.30) справедливы по отдельности для каждого уровня энергии. Теперь, если мы употребляем вместо дискретного индекса (s) непрерывный аргумент ε и вводим величину w согласно формуле (10), то по аналогии с (15) имеем

$$\frac{n(\varepsilon)}{m(\varepsilon) - n(\varepsilon)} = \frac{q(\varepsilon)}{1 - q(\varepsilon)} = w(\varepsilon) = ae^{-\varepsilon/kT}. \tag{25}$$

Это соотношение устанавливает для статистики Ферми — Дирака экспоненциальную зависимость $n(\varepsilon)$ от уровня энергии. Оно отличается от (15) только знаками в знаменателях первого и второго выражений. Это обусловливает значительное различие при больших значениях $q(\varepsilon)$, однако для небольших значений $q(\varepsilon)$ формулы обоих видов квантовой статистики (Бозе — Эйнштейна и Ферми — Дирака) приближаются к экспоненциальной формуле (3) статистики Максвелла — Больцмана.

При решении уравнений (15) относительно q эти соотношения часто объединяются в следующем виде:

$$q(\varepsilon) = \frac{n(\varepsilon)}{m(\varepsilon)} = \frac{1}{a^{-1}e^{\varepsilon/kT} \pm 1}. \tag{26}$$

Знак минус стоит в формуле статистики Бозе — Эйнштейна; знак плюс — в формуле статистики Ферми — Дирака. Если постоянная a мала, то обратная ей величина $+a^{-1}$ велика. Тогда членом ± 1 можно пренебречь, и формула (26) становится тождественной формуле Максвелла — Больцмана (3).

Теперь мы можем дать ответ на вопрос относительно возможности постепенного перехода от квантовой статистики к статистике Максвелла — Больцмана, поставленный в начале § 27. Уравнения (3), (15) и (25) позволяют сделать вывод, что для небольших значений $q(\varepsilon)$ все три статистики дают фактически одинаковое число частиц $n(\varepsilon)$ на уровень энергии. Далее, рассмотрение соотношений (27.28)—(27.30) показывает, как это видно из (27.35) и (27.36), что для одинаковых $n(\varepsilon)$ и небольших $q(\varepsilon)$ вероятности w_r для трех статистик, по существу, одни и те же. Таким образом, небольшие значения $q(\varepsilon)$ приводят к равноправию трех видов статистик как относительно числа частиц на уровень энергии, так и относительно распределения частиц в его пределах. Следовательно, квантовые статистики и статистика Максвелла-Больцмана совпадают во всех отношениях, когда $q(\varepsilon)$ уменьшается, включая также и совпадение по величине энтропии.

Для низких значений q все числа заполнения фактически равны 0 или 1, и для $r > 1$ мы имеем $m_r \sim 0$. Это справедливо как для квантовых статистик, так и для статистики Максвелла—Больцмана. Таким образом, распределение максимума вероятности одинаково для всех статистик. Результаты различны только для численных значений энтропии. Используя статистику Максвелла-Больцмана, мы будем считать, что каждая конфигурация максимального распределения содержит N_D конфигураций Максвелла — Больцмана, где N_D задано формулой (27.37). Последняя величина для $r > 1$ сводится к $n!$, если $m_r = 0$. Общее число конфигураций равняется, следовательно, произведению квантово-статистического числа конфигураций на $n!$. В результате выражения энтропии (27.11) в статистике Бозе — Эйнштейна и (27.18) в статистике Ферми — Дирака отличаются от выражения энтропии (27.5) в статистике Максвелла — Больцмана только на аддитивную постоянную $\ln(n!)$. То же самое справедливо и для обобщенных выражений энтропии (5)

и (6) при сравнении с (2). Здесь аддитивная постоянная имеет форму $\sum_{s=1}^{a} \ln [n^{(s)}!]$. Поскольку $n^{(s)}$ практически одинаково для всех видов статистик, если q мало, то величина этой постоянной для них одинакова и значения энтропии в них также тождественны, если не принимать во внимание аддитивную постоянную. Для небольших значений q определение энтропии одинаково для всех трех видов статистики.

Далее, $q(\varepsilon)$ мало в том случае, если мала постоянная a в уравнениях (3), (15) и (25). Поскольку a содержит в себе n в качестве множителя, то общее число частиц должно быть мало по сравнению с числом допустимых состояний m. Кроме того, с возрастанием температуры величина a уменьшается. Инверсия этого соотношения зависит, однако, от природы взаимодействия атомов газа между собой, то есть от вида кривой энергии ε как функции параметров. Поскольку возрастание температуры означает к тому же и возрастание энергии, а следовательно, и возрастание допустимых состояний, мы можем сделать вывод, что для высоких температур три вида статистик дадут практически одинаковые наблюдаемые результаты.

Для обычного газа комнатная температура является высокой, но при температурах, приближающихся к абсолютному нулю, наблюдаются специфические отклонения в поведении газа, которые выражают различие между тремя видами статистик в наблюдаемых результатах. О вырождении жидкого гелия, который подчиняется статистике Бозе — Эйнштейна, упоминалось в § 26. Для электронного газа, подчиняющегося статистике Ферми — Дирака, о котором также упоминалось в § 26, комнатная температура является низкой. Таким образом, средства наблюдения позволяют нам показать, что квантовая статистика подтверждается, когда можно ожидать наблюдаемые различия, в то время как статистика Максвелла — Больцмана может быть справедливой только при таких условиях, при которых все три вида статистики приводят к одинаковым наблюдаемым результатам. Справедливость статистики Максвелла — Больцмана нужно, следовательно, понимать как приближение, которое при нормальных физических условиях нельзя отли-

чить от более точной формы статистики Бозе — Эйнштейна. Эти соображения объясняют, почему те же самые частицы можно при высоких температурах рассматривать как различимые, тогда как при низких температурах их следует рассматривать как неразличимые. Переход к более высоким температурам не ведет к появлению у частиц специфических отметок, по которым можно было бы различать отдельные частицы. Мы бы скорее сказали, что частицы остаются неразличимыми, несмотря на то, что температура возрастает, однако статистические эффекты в конечном счете таковы, что они кажутся различимыми, то есть в этом случае больше нельзя провести различие между неразличимостью и различимостью. Эти рассуждения раскрывают специфическую асимметрию между двумя концепциями: в то время как статистические результаты при низких температурах вынуждают нас говорить, что частицы неразличимы, статистические результаты при высоких температурах не дают доказательств того, что частицы различимы, а только позволяют трактовать их как различимые.

Может возникнуть вопрос, является ли число частиц q в данном состоянии при условиях нормальной температуры и давления столь незначительным, чтобы это позволило не проводить различия между тремя видами статистик. В кинетической теории газов классической физики обычно предполагается, что объем, занимаемый газом, подразделен на небольшие ячейки, которые, однако, достаточно велики для размещения в них большого числа молекул. Если бы эти ячейки рассматривались как отделения нашей статистики, то статистика Бозе привела бы к результатам, заметно отличающимся от результатов статистики Больцмана. Однако такие ячейки слишком велики, для того чтобы представлять собой состояния в смысле квантовой физики. Они являются конфигурациями квантовых состояний и должны рассматриваться как таковые. Когда мы переходим к ячейкам квантовых размеров, например при рассмотрении квантованной кинетической энергии, число частиц на ячейку настолько мало, насколько это необходимо для приближения квантовой статистики к классической.

Статистика совокупности квантовых состояний значительно отличается от статистики индивидуальных состояний. Когда мы разделяем комнату на ячейки размером

в 1 $мм^3$ каждая, то в каждой ячейке фактически находится одинаковое число молекул воздуха. В классической статистике был сделан вывод, что та же самая однородность наблюдалась бы и в том случае, если бы мы подразделили комнату на еще меньшие ячейки. Для каждой малой ячейки предполагаются флуктуации, однако считают, что они проявляются в незначительных отклонениях от среднего числа заполнения.

Применяя статистику Бозе, мы обычно по-другому объясняем результаты наблюдений. Мы рассматриваем нашу ячейку размером в 1 $мм^3$ как состоящую из громадного числа квантовых ячеек; для этих мельчайших ячеек мы имеем неоднородное распределение, задаваемое величиной w_r статистики Бозе, в которой молекулы сосредоточиваются в отдельных ячейках, а флуктуации гораздо значительнее, чем в статистике Максвелла — Больцмана. Поскольку ячейка размером в 1 $мм^3$ состоит из громадного числа квантовых ячеек, то мы находим число молекул n'_i в этой ячейке, применяя формулу (13) и используя $w_r^{(s)}$ из формулы (14) вместо w_r:

$$n'_i = m' \cdot \sum_{r=1}^{\infty} r w_r = m' \cdot \frac{w}{1-w} = m' \cdot q. \qquad (27)$$

Для различных ячеек размером в 1 $мм^3$ число n'_i фактически одинаково в силу закона больших чисел, согласно которому флуктуации n'_i уменьшаются пропорционально квадратному корню из m'. Таким образом, однородное распределение в больших ячейках представляется как результат статистической компенсации неоднородности в громадном числе более мелких ячеек, и, по-видимому, такое распределение будет наблюдаться, даже если число частиц q на квантовое состояние будет невелико. Следовательно, совокупности квантовых состояний ведут себя подобно состояниям Максвелла — Больцмана. Эти соображения можно проиллюстрировать на конкретном примере. Предположим, что мы имеем 10 больших ячеек и 100 частиц, тогда $q = n/m = 10$. Если мы применяем статистику Бозе — Эйнштейна, вероятность того, что большая ячейка незаполнена, равна $w_0 = 1/(1+q) = 1/11 \sim 1/10$. Если процесс перераспределения повторяется часто, что означает, что

в среднем не заполнено около одной большей ячейки. Теперь предположим, что каждая ячейка подразделяется еще на 100 более мелких ячеек. Относительно этих 1000 ячеек мы имеем $q = \frac{1}{10}$, а статистика Бозе говорит, что вероятность того, что одна из ячеек не заполнена $w_0 = \frac{10}{11}$. Для того чтобы одна из больших ячеек была незаполнена, нужно, чтобы все 100 маленьких ячеек подряд были незаполнены; следовательно, вероятность того, что это случится, равна теперь $(10/11)^{100}$, то есть величине порядка 10^{-5}, и, таким образом, этой величиной можно пренебречь. Это означает, что практически мы не имеем ни одной незаполненной большой ячейки.

Применяя статистику Масквелла — Больцмана в первой ситуации, когда большие ячейки не подразделены, и используя точную формулу, приведенную в § 27 на стр. 328, в сноске под индексом (28(а)), мы находим вероятность $w_0 = [(m-1)/m]^n = \left(\frac{9}{10}\right)^{100}$, которая практически тождественна с вероятностью, определенной для статистики Бозе — Эйнштейна. Если мы дальше подразделяем большие ячейки, то вероятность появления незаполненной малой ячейки будет равна $w_0 = (999/1000)^{100}$. Практически она будет равна $^9/_{10}$, если мы приближенно полагаем, что $(1-\delta)^n$ можно заменить на $1 - n\delta$. Вероятность появления незаполненной большой ячейки вновь равна вероятности появления 100 незаполненных малых ячеек подряд. Эта вероятность равна $(^9/_{10})^{100}$, то есть такая же, как и прежде.

Следовательно, для статистики Максвелла — Больцмана размеры ячеек или состояний не имеют значения. Напротив, статистика Бозе — Эйнштейна зависит от абсолютных размеров состояний. Выражаясь более точно, мы должны говорить о *статистике Бозе по отношению к определенной системе ячеек*. Такой вывод не выдвигает новых трудностей, поскольку в квантовой физике размеры состояний фактически даются по абсолютной величине (см. § 26). Совокупность состояний обладает свойствами, которые отличаются от свойств единичного состояния, поскольку ее статистика следует из комбинированного подсчета неразличимых частиц и различимых состояний. Если совокупность включает в себя большое

число состояний, то ее свойства, по существу, определяются статистикой различимых состояний и поэтому сходны со свойствами статистики Максвелла — Больцмана.

Кроме того, эти соображения объясняют, почему квантовая статистика становится тождественной статистике Максвелла — Больцмана для предельного случая, когда $h = 0$. Если размеры квантовой ячейки равны 0, то в этом случае статистика Бозе по отношению к этим ячейкам была бы тождественна статистике Максвелла — Больцмана для ячеек неисчезающих размеров. Точнее говоря, значение этого ограничивающего условия состоит в следующем: если имеется система ячеек C_i, тогда может быть найдена такая система ячеек меньших размеров $c_i = (1/k)C_i$, по отношению к которой статистика Бозе эквивалентна в пределах заданной степени точности $\delta \sim$ статистике Максвелла — Больцмана по отношению к системе C_i.

Мы видим, что из однородного распределения в больших ячейках нельзя делать вывод, что такая же однородность справедлива и для ячеек меньших размеров и что эти ячейки также удовлетворяют распределению Максвелла — Больцмана. Такой вывод был бы необоснованной экстраполяцией. Однородность в больших масштабах совместима с неоднородностью в малых масштабах, следовательно, макроскопическая однородность не позволяет нам обнаружить разницу между различимыми и неразличимыми частицами.

Поскольку наблюдения квантовых явлений в экспериментах, где особенности квантовых статистик проектируются в макрокосм, решают в пользу второй альтернативы, то мы должны рассматривать очевидную однородность макрокосма как результат статистического усреднения, а не подобной же однородности в микрокосме.

§ 29. Частицы, исчезающие в небытие

Математические выводы, полученные в §§ 27 и 28, можно использовать при ответе на вопрос, который имеет отношение к логическому анализу генетического тождества, проведенному в § 26. В процессе этого анализа было установлено, что мы не имеем, если говорить

точно, доказательства, что квантовые частицы неразличимы. Мы смогли только доказать, что данная интерпретация необходима для того, чтобы избежать каузальных аномалий. Возникает вопрос, существует ли интерпретация статистик Бозе и Ферми с помощью различимых частиц, в которой каузальные аномалии имеют вид не отношений взаимной зависимости, а принимают другие формы. В настоящем параграфе мы намереваемся показать, что на этот вопрос можно дать положительный ответ, то есть, что существует интерпретация, в которой сохраняется генетическое тождество частиц, сформулированное специфическим образом. Прежде всего, такую интерпретацию можно предложить для статистики Ферми. Предположим, что частица может существовать в двух формах; иногда она «активна», и ее можно обнаружить, а иногда «пассивна» (frozen), и ее нельзя обнаружить посредством физического взаимодействия с другими вещами. Предположим затем, что одна и та же частица остается в одном состоянии все время, меняя, однако, случайным образом условия своего существования. Иногда она находится в условиях пассивного, иногда — в условиях активного существования. Если частица находится в условиях пассивности, то она временно исчезает, так сказать, в небытие, и занимаемое ею состояние попадает в условия, которые, естественно, можно было бы назвать «пустотой». Вероятность того, что частица перейдет в условия активного существования, равна q; вероятность же того, что она перейдет в условия пассивности, равна $1-q$. Величина этих вероятностей определяется по формуле (27, 30), и в соответствии с (28.25) мы предполагаем, что q, которая постоянна для каждого состояния, находится в экспоненциальной зависимости от энергии ε. Мировая линия частицы, подразделенная на незначительные интервалы Δt, является тогда случайной вероятностной последовательностью, подчиняющейся вероятности q. Частицы совершенно не зависят друг от друга, и поэтому не требуется никакого принципа запрета. Вместо этого появляется аномалия; согласно этой аномалии частица существует в двух различных условиях, одно из которых является условием пассивного существования.

Подобную интерпретацию можно осуществить и для статистики Бозе. Здесь выражение (27.33) для w, при-

водит к предположению о возможности интерпретации с помощью случайной последовательности групп чисел. Мы рассмотрим эту интерпретацию, предположив сначала, что все состояния относятся к одному и тому же уровню энергии, поскольку формула (27.33) была выведена именно для этого случая.

Когда мы в случайной последовательности имеем дело с событием B, обладающим вероятностью w, и проводим разделяющую линию после каждого $\bar{B}$, то получаем следующую схему:

$$BB\bar{B}|BBBB\bar{B}|\bar{B}|B\bar{B}|B\bar{B}|BB\bar{B}|\ldots \qquad (1)$$

Каждая секция содержит группу, состоящую из точного числа «элементов» B, и оканчивается на $\bar{B}$. Вероятность выпадения такой группы определяется следующим образом:

$$P(\bar{B}, B^1 \ldots B^r \cdot \bar{B}^{r+1}) = (1 - w)\, w^r = w_r. \qquad (2)$$

При этом методе вычисления по формуле (2) мы опираемся на член $\bar{B}$ как на точку отсчета, потому что каждое $\bar{B}$ вводит следующую секцию. Если секция содержит только элемент $\bar{B}$, то мы имеем $r = 0$. Величина w_r, полученная из формулы (2), тождественна величине, получающейся из формулы (27.33) статистики Бозе.

Этот результат позволяет нам построить следующую макроскопическую модель статистики Бозе. Имеется мешок, наполненный шарами, и некоторое число ячеек, или состояний. Мы используем какой-нибудь механизм, с помощью которого определяем вероятность w выпадения события B, и начинаем игру. Если выпадет B, то мы кладем шар из мешка в первую ячейку и продолжаем игру дальше. Если второй раз снова выпадет B, мы кладем второй шар в эту же ячейку. Так продолжается до тех пор, пока не выпадет $\bar{B}$. Тогда мы уже не кладем шар в эту ячейку и прекращаем процесс ее заполнения. Продолжим игру дальше, используя механизм, определяющий вероятность w, заполняя теперь шарами вторую ячейку, следуя тем же правилам. Когда $\bar{B}$ следует за $\bar{B}$, шар в ячейку вообще не кладется. Мы продолжаем, таким образом, до тех пор, пока не заполним все

ячейки. В результате получаем ряд чисел заполнения, подчиняющийся статистике Бозе.

Существует и другой метод заполнения. Когда выпадает $\bar{B}$, мы все же кладем шар в ячейку, однако отмечаем его как «пассивный». Затем заполнение этой ячейки прекращается, и игра продолжается дальше. Таким образом, каждая ячейка получает только один «пассивный» шар вдобавок к «активным» шарам, составляющим модель статистики Бозе. То есть теперь мы получаем числа заполнения $r' = r + 1$, так что $w'_r = w_r$.

Для построения структур для частиц мы будем использовать второй метод, придерживаясь предположения, что все состояния относятся к одному и тому же уровню энергии. Предположим, что мы имеем m ячеек и n шаров. Если их число велико, мы вправе предположить, что при каждом повторении использованы все n шаров и заполнены все m ячеек. Кроме того, мы пронумеруем все шары, для того чтобы идентифицировать их, однако они будут извлекаться из мешка не в порядке их номеров, а случайным образом. Ячейки могут выбираться в некотором определенном порядке, который все время остается постоянным, или также случайным образом, меняющимся после каждой процедуры заполнения. Когда закончена первая процедура заполнения, мы записываем, в какой ячейке находится каждый шар, а также, находится ли он в условиях «активности» B. Затем вынимаем все шары, стираем метки на «пассивных» шарах, оставляя, однако, индивидуальные номера на них, и начинаем все сначала. Вновь шары отбираются в случайном порядке. Продолжая действовать таким образом, мы получаем структуру в том виде, как она представлена в табл. 7.

В этой таблице каждая строка представляет историю частицы. Относительно атрибутов B и $\bar{B}$ все строки и столбцы представляют собой случайные последовательности, которые подчиняются вероятностным законам и независимы друг от друга. Таким образом, мы в данном случае имеем структуру нормальных последовательностей в узком смысле[1]. Напротив, атрибуты C_i удовлетворяют статистике Бозе. Это означает, что когда мы

---

[1] ThP, p. 172.

ТАБЛИЦА 7

Структура для частиц в статистике Бозе. (В этой таблице приводятся атрибуты «активности» (B) и «пассивности» ($\bar{B}$), а кроме того, состояние C_i, в котором находится каждая частица. Выпадение события B как в строках, так и в столбцах подчиняется вероятности w)

	t_1	t_2	t_3	t_4	. . . . .
Частица 1	B, C_{12}	B, C_1	$\bar{B}$, C_4	B, C_4	. . . . .
Частица 2	B, C_2	$\bar{B}$, C_1	B, C_{21}	B, C_2	. . . . .
Частица 3	$\bar{B}$, C_{13}	$\bar{B}$, C_4	B, C_{67}	B, C_{14}	. . . . .
Частица 4	B, C_{19}	B, C_1	$\bar{B}$, C_{12}	B, C_4	. . . . .
. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .
Частица n	B, C_{20}	B, C_4	B, C_{67}	$\bar{B}$, C_1	. . . . .

подсчитываем в столбце число частиц $n = r$, характеризуемых одинаковыми атрибутами C_i, то частота этих чисел r, включающих одну «пассивную» частицу для каждой ячейки, контролируется величиной w_{r-1}, полученной из формулы (2). Точнее говоря, если m_r есть число ячеек в столбце, где все они характеризуются числом заполнения r, то мы находим $w_{r-1} = m_r/m$. Это положение справедливо для каждого столбца.

Данная структура допускает следующую физическую интерпретацию. История жизни частиц включает в себя условия пассивности и активности. Переход на каждую новую ступень происходит скачком: вероятность того, что она перейдет в условия активности, равна w, а вероятность того, что она перейдет в условия пассивности, равна $1 - w$. Во время всех этих скачков частицы независимы друг от друга. Предположим, что в какой-то момент заполнены все состояния. В следующий момент все частицы, включая и «пассивные», покидают свои состояния и переходят в новые. Когда частица переходит в новое состояние, вероятность того, что она будет находиться там в условиях активности, равна w; это справедливо также и в том случае, если частица первоначально находится в пассивном состоянии. Однако если частица при переходе в новое состояние попадает в условия пассивности, то доступ в это состояние оказы-

вается закрытым для всех остальных частиц. Такое перераспределение происходит каждое мгновение.

Благодаря данной интерпретации мы построили структуру для частиц, в которой не существует никаких каузальных аномалий, которые выражали бы причинную зависимость. Мировые линии частиц независимы друг от друга, поскольку это касается атрибута B, а статистика Бозе по отношению к атрибуту C_i обусловливается только процессом, который делает недоступными для остальных частиц состояния, заполненные пассивными частицами. Но это условие запрета представляет собой каузальную аномалию специфического вида. И хотя мы имеем здесь частицы, обладающие физической индивидуальностью, их генетическое тождество является аномальным, поскольку история жизни каждой частицы содержит периоды, в течение которых частица ненаблюдаема и «существует в форме небытия».

Расширение этих положений на случай состояния с различными уровнями энергии можно осуществить следующим образом. В данном случае вероятность w того, что частица активна, зависит, согласно (28.15), от уровня энергии, связанного с состоянием. В макроскопической модели процедуры заполнения мы используем для каждой ячейки соответствующую величину w. Строки в структуре (см. табл. 7) представляют тогда собой последовательности переменных вероятностей, специфическая постоянная величина w появляется только на отдельных участках, определяемых ячейками C_i, относящимися к одному уровню энергии. Такая же закономерность изменения w справедлива и для столбцов[1].

Физическая интерпретация такая же, как и раньше. Следует отметить, что данная интерпретация не претендует на роль физической гипотезы и не выдвигается здесь как предположение, чтобы физики изменили свое понимание статистики Бозе. Данная интерпретация представляет собой только один из способов языка. Все предположения, на которых она основана (переход частиц в условие активности или пассивности, предположение, что каждая пассивная частица изолирует данное состояние и не допускает в него другие частицы, а так-

---

[1] Последовательности, в которых вероятность меняется от элемента к элементу, анализируются довольно часто, см. ThP, p. 168.

же экспоненциальная зависимость вероятности нахождения частицы в условиях активности от уровня энергии), являются истинными по определению. Они представляют собой соглашения, на которых основывается данная интерпретация. Эти соглашения приемлемы потому, что они совместимы с наблюдаемыми статистиками. Эта интерпретация показывает, что когда мы желаем ввести субстанциальное генетическое тождество в статистику газов, то вытекающие отсюда каузальные аномалии все же зависят от нашего выбора.

§ 30. Частицы, движущиеся в обратном направлении времени

Обобщение понятия генетического тождества, которое приводит к функциональному генетическому тождеству неразличимых частиц, все еще сохраняет за элементарными частицами некоторые черты индивидуальности, поскольку при переходе к другим распределениям число частиц остается постоянным. Если частицы индивидуально неразличимы, то их по крайней мере можно сосчитать. Однако оказывается, что даже этот остаток индивидуальности может быть поставлен под сомнение. В теории фотонов, то есть в теории излучения, отказываются от сохранения числа частиц. Энергия фотона данной длины волны λ или частоты ν, согласно фундаментальному соотношению М. Планка, равна $h\nu$. Если замкнутый сосуд наполнен излучением, которое поглощается и вновь излучается его «черными» стенками, то происходит обмен длинами волн, поскольку фотоны, получающиеся при этом обмене, несут энергию, отличную от энергии первоначальных фотонов. Если полная суммарная энергия этих фотонов должна оставаться постоянной, то число фотонов уже не может быть постоянным. Действительно, если мы применяем в данном случае метод максимизации, описанный в § 28, и пренебрегаем условием (28.7), в котором сформулировано сохранение числа частиц, то константа α в (28.9) становится равной 0, и, следовательно, мы находим из формулы (28.10), что $a = 1 = a^{-1}$. Вводя это значение в (28.26) для статистики Бозе и заменяя ε на $h\nu$, мы получаем

$$q(\varepsilon) = \frac{n(\varepsilon)}{m(\varepsilon)} = \frac{1}{e^{-h\nu/kT} - 1}. \qquad (1)$$

Это знаменитый закон излучения Планка, с помощью которого квантовая теория проложила себе путь в физику. Сам Планк выводил свой закон по-другому, а именно при помощи классической статистики линейных осцилляторов. Но статистика Бозе, созданная спустя четверть века, приходит к такому же результату, если при рассмотрении излучения предполагается, что оно состоит из неразличимых частиц, число которых не сохраняется.

Однако фотоны не являются материальными частицами в обычном смысле. Они не обладают массой покоя, то есть если, скажем, их можно было бы остановить, то они вообще не имели бы массы. Они обязаны всей своей массой тому, что движутся со скоростью света. Следовательно, не удивительно, что их число может изменяться. Было бы совсем по-другому, если бы можно было доказать, что в пределах замкнутого сосуда может меняться число таких материальных частиц, как электроны.

Довольно неожиданно недавние исследования продемонстрировали, что генетическое тождество материальных частиц может быть поставлено под сомнение более серьезным образом, чем это делается в статистике Бозе. Можно показать, что различие между одной и двумя, даже тремя материальными частицами является вопросом интерпретации, то есть, что это различие не является объективным фактом, а зависит от языка, используемого при описании. Следовательно, число материальных частиц зависит от экстенсиональных правил языка. Однако такие интерпретации, допустимые с физической точки зрения, имеют существенные отличия от остальных: они требуют отказа от временнóго порядка.

Подобные концепции были выдвинуты Штюкельбергом и Фейнманом [1]. Их исследования показали, что позитрон, то есть частицу, обладающую массой, равной массе электрона, но несущую положительный заряд,

---

[1] См. E. C. G. Stückelberg, Remarque à propos de la création de paires de particules en théorie de relativité, в „Helv. phys. Acta", 1941, vol. 14, p. 588—594; его же, La mécanique du point matériel en théorie de relativité et en théorie des quanta, в „Helv. phys. Acta", 1942, vol. 15, p. 22—37. См. также: R. P. Feynman, The Theory of Positrons, в „Phys. Rev.", 1949, vol. 76, p. 749—759.

можно рассматривать как электрон, движущийся в обратном направлении времени. Отрицательно заряженный электрон, который движется в направлении времени, противоположном нашему, приводит к тем же самым физическим следствиям, что и положительно заряженный позитрон, движущийся в обычном направлении времени, и, следовательно, при помощи наблюдения нельзя провести различие между обеими интерпретациями.

Рис. 38. Порождение и аннигиляция пар элементарных частиц в камере Вильсона.

Фейнман показал, что эти соображения могут быть использованы при объяснении порождения и аннигиляции пар. Эти процессы можно наблюдать с помощью фотографий, сделанных в камере Вильсона, которые говорят нам, что при облучении этой камеры γ-лучами электрон и позитрон порождаются из «ничего» и движутся из одной точки по разным траекториям. Позитрон обычно живет недолго, он сталкивается с каким-нибудь другим электроном, свободно движущимся в пространстве, и сливается с ним в акте столкновения. Таким образом, эти две частицы исчезают полностью, порождая новые γ-лучи, начинающиеся из точки столкновения.

Эти процессы можно проиллюстрировать с помощью рис. 38. Положительное время здесь изображено вертикальной линией, идущей вверх, остальные непрерывные линии представляют собой мировые линии частиц. Прерывные линии изображают мировые линии γ-лучей. В событии A падающие γ-лучи порождают пару, состоящую из электрона и позитрона. В событии C позитрон

Рис. 39. Процесс, изображенный на рис. 38, в данном случае рассматривается как мировая линия одного электрона, который от точки C до точки A движется в обратном направлении времени.

сталкивается с электроном *2*; эта пара при столкновении аннигилирует и порождает только след γ-лучей, начинающийся в точке C. На фотографии траектории частиц являются видимыми и обладают пространственным расположением, подобным непрерывным линиям диаграммы (38); лучи на фотографии невидимы.

Согласно Фейнману, мы можем с таким же успехом интерпретировать процесс, изображенный на рис. 38, рассматривая цепочку линий $DCAB$ как мировую линию одного электрона, который движется от C к A в обратном направлении времени, как показано на рис. 39.

Таким образом, вместо трех частиц мы имеем здесь только одну. В этой интерпретации то преимущество, что нам не нужно говорить о порождении и аннигиляции пар. Здесь мы все время имеем дело с одной частицей. Элиминируются каузальные аномалии порождения частиц из ничего и исчезновение их в ничто; однако они заменяются другой каузальной аномалией, появляющейся в данном описании: часть своего пути электрон проделывает в обратном направлении времени.

Мы встречаемся здесь с новой иллюстрацией теории эквивалентных описаний, справедливой по отношению к интерфеноменам, то есть *по отношению к ненаблюдаемым*. На фотографии зафиксировано изображение ряда, состоящего из мельчайших капелек воды. Мы интерпретируем это явление следующим образом: частица на своем пути ионизирует другие частицы, вокруг которых конденсируются капельки водяного пара. Движущаяся частица только подразумевается и поэтому представляет собой интерфеномен. Применяя обычное описание, мы говорим, что интерфеномен представляет собой положительную частицу, движущуюся в обычном направлении времени. Применяя новое описание, мы говорим, что это отрицательная частица, движущаяся в обратном направлении времени. С макроскопической точки зрения серия столкновений, которую мы обнаружили с помощью капелек воды, сохраняет, таким образом, свой временнóй порядок. Однако, если рассматривать это явление, так сказать, с точки зрения частицы, то столкновения происходят в обратной последовательности. Этот вид языка, применяемый только по отношению к интерфеноменам, позволяет нам элиминировать аномалии порождения и аннигиляции пар, ибо в отношении частицы можно сказать, что не существует никаких пар частиц, а имеется только одна частица. Однако за элиминирование этой аномалии приходится уплатить появлением другой аномалии — движением в обратном направлении времени. Исчерпывающее описание интерфеноменов, как и всегда в квантовой механике, связано с каузальными аномалиями, и мы можем только сделать выбор, куда их поместить. Можно напомнить здесь выводы, полученные в § 29, где мы нашли, что иногда удобно предположить, что частицы временно исчезают в небытие. В данном примере мы отклоняемся от нормальной причинной ин-

терпретации интерфеноменов в другую сторону, применяя физическое описание, использующее обращение причинных мировых линий.

Хотя для нового описания не существует порождения и аннигиляции пар, однако еще остается два пучка γ-лучей, которые можно верифицировать путем наблюдения и нельзя «элиминировать из интерпретации». Эти лучи рассматриваются здесь как события, причинно связанные с обращением времени в двух точках — событиях A и C. Обращение здесь понимается как отклонение от направления времени окружающей среды. Можно дать математическое описание этих обратных процессов, если соответствующим образом сформулировать уравнения. Подобные соображения впервые были выдвинуты Штюкельбергом и явились исходным пунктом новой интерпретации.

Следует отметить, что данная интерпретация означает не только и не просто обращение направления времени, но и отказ от временнóго порядка. В нашем анализе направления времени мы уже рассматривали обращение всех причинных цепей и, вообразив киноленту, прокручиваемую в обратном направлении, пытались представить себе, какую картину представлял бы собой в этом случае мир. Однако при отождествлении позитрона с электроном, движущимся в обратном направлении времени, обращается лишь одна причинная цепь, тогда как остальные сохраняют свое направление. Вследствие этого отношение «между» (см. § 22) претерпевает изменение. На рис. 39 точка-событие C в причинном отношении находится между D и A, а точка A — между C и B. Что касается рис. 38, то здесь данное положение несправедливо. В этом и состоит то изменение отношения «между», на которое мы ссылаемся, когда говорим об отказе от временнóго порядка.

Отбрасывая временнóй порядок, мы открываем путь для возможного существования замкнутых причинных цепей. Применив язык обычной интерпретации (см. рис. 38), предположим, что в точке a электрон № *1* вступает в какой-то процесс столкновения, который сопровождается излучением светового луча; этот луч движется быстрее позитрона и сталкивается с электроном № *2* в точке d. Если мы предположим, что отрезок пути dC электрона *2* остается в этом случае неизменным, то

нам следует также предположить, что электрон вовсе не обязательно движется из точки D. Когда же мы вводим вторую интерпретацию (рис. 39), то мы уже не можем игнорировать световой луч ad; $dCAad$ теперь представляет собой замкнутую причинную цепь. Такие процессы еще не наблюдались; однако, хотя они и представляются крайне неправдоподобными, возможность их существования отрицать нельзя.

Замкнутые причинные линии данного типа не вызывают трудностей, возникающих в том случае, если замыкаются макроскопические причинные цепи. Мы видели в § 18, что понятие порождения есть статистическое понятие; следовательно, оно неприменимо к атомным цепям, и мы не можем сказать, что здесь следствие порождается своей собственной причиной. Понятие метки также является статистическим (см. § 23), поэтому замкнутая атомная цепь не может привести к парадоксам, возникшим при предположении о передаче метки.

Мы не можем применять здесь и понятие протоколирования, потому что оно также имеет статистическую природу. Рассматривая во второй интерпретации событие A как следствие события C, мы с точки зрения события A не можем сказать относительно события C чего-либо большего, чем это позволяет нам сделать обычная интерпретация. Замкнутая причинная линия атомного типа не изменяет нашего понятия причинности, являющегося макроскопическим, однако она противоречит понятию временнóго порядка, а тем самым и обычному понятию генетического тождества.

В самом деле, отождествление позитрона с электроном, движущимся в отрицательном времени, потрясает самые основания понятия физического тождества. Если одна и та же частица в одно и то же время может существовать более чем в одном экземпляре, тогда отношение генетического тождества для одновременных событий не является больше уникальным. Статистика Бозе потребовала от нас отказа от субстанциального генетического тождества. Однако заменившее его функциональное генетическое тождество все же удовлетворяло правилу, что среди одновременных состояний никакие два состояния не могут относиться к одной частице. В отношении генетического тождества позитронов и

электронов дело обстоит по-другому. Если мы с помощью макроскопического времени проведем одновременное поперечное сечение, как показано пунктирной горизонтальной линией на рис. 38 и 39, тогда в этот момент времени электрон, в интерпретации Фейнмана, существует в трех образцах, представленных тремя состояниями s_1, s_2 и s_3, в то время как в обычной интерпретации эти три состояния s_1, s_2, s_3 относятся к трем разным частицам, из которых одна является позитроном. Однако если рассматривать линию *DCAB на* рис. 39 с точки зрения причинности, то траектория, идущая от *C* к *A*, не находится в противоположном направлении времени. С точки зрения «собственного времени» электрона она также, подобно другим отрезкам цепи, направлена в одну сторону; и три состояния s_1, s_2 и s_3 являются не одновременными, а следуют одно за другим. Но то, что для самого электрона во времени направлено в одну сторону, для других вещей, в частности для такого макроскопического аппарата, как камера Вильсона, имеет противоположное направление во времени. Мы получаем здесь индивидуальные определения направления времени, которые нельзя объединить в последовательную сеть. Поэтому для таких событий направления времени не существует.

Это наиболее серьезный удар, который понятие времени получало когда-либо в физике. Классическая механика не может дать оценку направлению времени, но она может по крайней мере определить временнóй порядок. Статистическая механика может определить направление времени с помощью исчисления вероятностей, однако это определение предполагает существование временнóго порядка для тех атомных событий, статистическое поведение которых определяет направление времени. Если дальнейшие исследования подтвердят интерпретацию Фейнмана, которая сейчас дискутируется, то квантовая физика, по-видимому, не может говорить даже о единовременном порядке изучаемых в ней процессов. В настоящее время пока еще не известно, какое это будет иметь значение для квантовой физики. Возможно, это означает, что переменная *t*, которая часто играет в квантовой физике сомнительную роль, имеет только «местное» значение и не всегда может распространяться на более широкие области. Возможно, что

местное t следует отличать от общего t, представляющего собой время макроскопических процессов. Чтобы найти ответ на эти вопросы, следует подождать создания удовлетворительной теории атомного ядра.

Однако уже сейчас можно сделать один вывод. Даже временной порядок, а не только направление времени макрокосма следует оценивать с помощью статистических методов. Как это сделать, в принципе показано в § 22. Достижения квантовой физики открывают нам другую возможность для объяснения порядка времени. Время жизни позитронов крайне мало, и они составляют лишь незначительную группу среди всех частиц, существующих в момент времени макроскопических масштабов. Поэтому громадное большинство частиц подчиняется правилам упорядоченного времени, и порядок времени вытекает из атомного хаоса в силу статистического доминирования отрицательно заряженных частиц. Если бы дело обстояло иначе, если бы позитронов было столько же, сколько электронов, то существование упорядоченного времени представлялось бы сомнительным. Возможно, что в этом случае в макрокосме существовали бы замкнутые причинные линии. Отсюда мы должны сделать вывод, что существование линейного порядка времени связано с асимметрией отрицательного и положительного электричества. Согласно концепции Дирака, эта асимметрия обусловлена тем, что практически все состояния с отрицательной энергией заняты, тогда как состояния с положительной энергией являются вообще пустотой; позитрон может рассматриваться как дырка, оставляемая электроном, когда он покидает свое место в отрицательном уровне энергии. Это образный язык, однако он выражает идеи, которые можно сформулировать в виде математических уравнений. Крайне желателен логический анализ этих проблем. Но предположение о том, что существованием упорядоченного макроскопического времени мы обязаны специфическому различию между двумя формами электричества, представляется правдоподобным.

Из этого краткого обзора современных исследований можно сделать вывод, что с помощью квантовой физики нельзя определить направления времени; подобное же было показано при анализе атомных событий в рамках классической физики. Квантовая физика еще

больше обострила ситуацию, поскольку она показала, что и временнóй порядок, а не только направление времени является статистическим свойством. Время представляется всецело макроскопическим явлением, которое нельзя приписывать явлению микрокосма. Оно в каждый момент порождается из атомного хаоса в виде статистического соотношения. Довольно странно, что происхождение из беспорядка не уменьшает значения макроскопического времени. Напротив, мы увидим, что время статистического космоса, рожденное из атомного хаоса, обладает точно такими же свойствами, которые с точки зрения здравого смысла повседневного опыта представляются наиболее существенными характеристиками временнóго потока.

ДОБАВЛЕНИЕ

[В план книги была включена последняя глава, где должно было рассматриваться соотношение между субъективным переживанием времени человеком и объективными свойствами времени в природе. Для того чтобы указать на некоторые из идей автора по данной проблеме, я привожу здесь перевод небольшого параграфа из работы Ганса Рейхенбаха: „Les Fondements logiques de la méchanique des quanta", в „Annales de l'Institut Henri Poincaré", tome XIII, Fasc. II, Paris, 1953, p. 156—157. — *M. P.*]

Различие между индетерминизмом будущего и детерминизмом прошлого выражается, в конечном счете, в законах физики. Это важное следствие сочетания классической статистики и соотношения неопределенностей квантовой физики. Выводы из этого сочетания для времени нашего опыта, то есть для времени повседневной жизни, очевидны. Понятие *становления* приобретает физический смысл: настоящее время, отделяющее будущее от прошлого, является моментом, в который индетерминированное становится детерминированным, и «становление» означает то же самое, что и «становление детерминированным».

Остается рассмотреть один вопрос. Каково отношение между временем физики и временем нашего опыта? Почему поток психологического времени тождествен с направлением возрастания энтропии?

Ответ простой: человек — часть природы, и его память представляет собой регистрирующий инструмент, подчиняющийся законам теории информации. Возрастание информации определяет направление субъективного времени. Вчерашние переживания регистрируются в нашей памяти, а переживания завтрашнего дня — нет, и они не могут быть зарегистрированы до тех пор, пока *завтра* не станет *сегодня*. Время нашего опыта тождественно времени, которое фиксируется с помощью регистрирующих инструментов.

Определение течения времени не только человеческая прерогатива, любой регистрирующий инструмент делает то же самое. То, что мы называем направлением времени, направлением становления, представляет собой отношение между регистрирующим инструментом и его окружением. То, что это соотношение одинаково для всех инструментов, включая и человеческую память, гарантируется статистической изотропией вселенной.

Добавим замечание относительно термина «теперь». Символическая логика говорит нам, что мы должны проводить различие между индивидуальным знаком и классом знаков, символом. Относительно многих слов трудно согласиться, что они представляют собой индивидуальный знак, например слово «дом» имеет везде одинаковое значение. Иначе обстоит дело для слов, подобных «теперь», «здесь», «Я», значение которых меняется с каждым индивидуальным знаком. В этом случае я предпочел бы говорить об указательно-рефлективных знаках и указательно-рефлективных значениях, используя слово «знак» в смысле индивидуального знака.

Акт мышления является событием и, следовательно, определяет положение во времени. Если наши переживания всегда имеют место в пределах «теперь», то это означает, что каждый акт мышления определяет точку отсчета. Мы не можем избежать момента «теперь», потому что любая такая попытка представляет собой акт мышления и, следовательно, определяет «теперь». Любая мысль обладает точкой отсчета, потому что сама мысль определяет эту точку. Этот факт выражается в грамматике при помощи правила, что каждое предложение должно содержать глагол, то есть указательно-рефлективный знак, указывающий время события, о ко-

тором идет речь, ибо время глагола имеет указательно-рефлективное значение.

Понятиям *детерминированный* и *недетерминированный* также следует приписывать указательно-рефлективное значение. Слово «детерминация» обозначает отношение между двумя ситуациями *A* и *B*; ситуация *A* детерминирует или не детерминирует ситуацию *B*. Бессмысленно говорить, что ситуация *B*, когда она рассматривается изолированно, является детерминированной. Когда мы говорим, что прошлое детерминировано или что будущее недетерминировано, то это означает, что мы ссылаемся на ситуацию, существующую в настоящий момент времени. Только при наличии ссылки на момент «теперь» можно говорить, что прошлое детерминировано, а будущее — нет. Следовательно, эти слова, так же как и многие другие, имеют указательно-рефлективное значение. Тем не менее остается справедливым, что в этих словах выражается объективное отношение. Если *A* представляет собой ситуацию, определяемую актом высказывания, а ситуация, которая предшествует ей, является детерминированной относительно *A*, тогда как ситуация, следующая за ней, — нет, то это отношение является физическим фактом.

ОГЛАВЛЕНИЕ

Предисловие 7

Глава I. Введение 11

§ 1. Эмотивное значение времени 11

Глава II. Временной порядок в механике 34

§ 2. Качественные свойства времени 34
§ 3. Причинная теория времени 41
§ 4. Причинность в классической физике 45
§ 5. Определение временного порядка с помощью причинных отношений 52
§ 6. Вмешательство 65

Глава III. Направление времени в термодинамике и микростатистике 71

§ 7. Второй закон термодинамики 71
§ 8. Статистическое определение энтропии 80
§ 9. Распространение статистики на случай газа с различными уровнями энергии 91
§ 10. Детерминистская интерпретация термодинамической статистики 102
§ 11. Детерминизм versus индетерминизм в классической физике 113
§ 12. Вероятностная структура 132
§ 13. Парадокс обратимости 148
§ 14. Направление времени в пространственном ансамбле 161
§ 15. Секционный характер направления времени 172
§ 16. Гипотеза ветвящейся структуры 185

Глава IV. Направление времени в макростатистике . . 198

§ 17. Макроконфигурации и макроэнтропия 198
§ 18. Причина и следствие: порождение и протоколирование 203
§ 19. Принцип общей причины 213
§ 20. Энтропия и информация 227
§ 21. Направление времени в учении об информации и теория регистрирующих приборов 240
§ 22. Чисто макростатическое определение направления времени 252
§ 23. Принцип маркировки и причинное соответствие . . 264

Глава V. Время в квантовой физике 275

§ 24. Статистическая обратимость элементарных процессов в квантовой механике 275
§ 25. Индетерминизм квантовой механики 280
§ 26. Генетическое тождество квантовых частиц 297
§ 27. Понятие энтропии в квантовой статистике 313
§ 28. Применение квантовой статистики в случае различных уровней энергии 331
§ 29. Частицы, исчезающие в небытие 342
§ 30. Частицы, движущиеся в обратном направлении времени 348

Добавление 357

Издательство УРСС

специализируется на выпуске учебной и научной литературы, в том числе монографий, журналов, трудов ученых Российской Академии наук, научно-исследовательских институтов и учебных заведений.

Уважаемые читатели! Уважаемые авторы!

Основываясь на широком и плодотворном сотрудничестве с Российским фондом фундаментальных исследований и Российским гуманитарным научным фондом, мы предлагаем авторам свои услуги на выгодных экономических условиях. При этом мы берем на себя всю работу по подготовке издания — от набора, редактирования и верстки до тиражирования и распространения.

Среди вышедших и готовящихся к изданию книг мы предлагаем Вам следующие:

Сачков Ю. В. **Научный метод: вопросы и развитие.**

Грюнбаум А. **Философские проблемы пространства и времени.**

Карнап Р. **Философские основания физики. Введение в философию науки.**

Поппер К. Р. **Объективное знание. Эволюционный подход.**

Джеммер М. **Понятие массы в классической и современной физике.**

Овчинников Н. Ф. **Методологические принципы в истории научной мысли.**

Новиков А. С. **Научные открытия (опыт темпорального анализа).**

Майданов А. С. **Процесс научного творчества: Философско-методологический анализ.**

Канке В. А. **Формы времени.**

Аксенов Г. П. **Причина времени.**

Койре А. **Очерки истории философской мысли.**

Шульман М. Х. **Теория нарового расширяющейся Вселенной.** (Relata Refero.)

Табор М. **Хаос и интегрируемость в нелинейной динамике.**

Эбелинг В., Энгель А., Файстель Р. **Физика процессов эволюции.**

Олемской А. И., Кацнельсон А. А. **Синергетика конденсированной среды.**

Малинецкий Г. Г. **Хаос. Структуры. Вычислительный эксперимент.**

Пригожин И. **От существующего к возникающему.**

Серия «Синергетика: от прошлого к будущему»

Пригожин И., Стенгерс И. **Время. Хаос. Квант. К решению парадокса времени.**

Пригожин И., Стенгерс И. **Порядок из хаоса. Новый диалог человека с природой.**

Пригожин И., Николис Г. **Познание сложного. Введение.**

Пригожин И., Гленсдорф П. **Термодинамическая теория структуры, устойчивости и флуктуаций.**

Капица С. П., Курдюмов С. П., Малинецкий Г. Г. **Синергетика и прогнозы будущего.**

Малинецкий Г. Г., Потапов А. Б. **Современные проблемы нелинейной динамики.**

Баранцев Р. Г. **Методология современного естествознания.**

Баранцев Р. Г. и др. **Асимптотология — путь к целостной простоте.**

Чернавский Д. С. **Синергетика и информация (динамическая теория информации).**

Трубецков Д. И. **Введение в синергетику.** В 2 кн.

По всем вопросам Вы можете обратиться к нам:
тел./факс **(095) 135–44–23,** ***тел.*** **135–42–46**
или ***электронной почтой*** **urss@urss.ru.**
Полный каталог изданий представлен
в ***Интернет-магазине:*** **http://urss.ru**

Издательство УРСС

Научная и учебная литература

Издательство УРСС

Представляет Вам свои лучшие книги:

Рейхенбах Г.

Философия пространства и времени.

Книга известного логика и философа Г. Рейхенбаха посвящена рассмотрению оснований геометрии, теории времени и теории относительности Эйнштейна. В ней анализируются такие важные методологические проблемы современной науки и философии, как отношение между теорией и наблюдениями, между топологическими и метрическими свойствами пространства и времени, психологическая проблема возможности наглядного представления неевклидовых структур. Книга является одним из самых обстоятельных исследований философских проблем, поставленных неевклидовыми геометриями и их применением в теории относительности.

Эддингтон А.

Пространство, время и тяготение.

Эта книга написана выдающимся английским астрофизиком, профессором Кембриджского университета Артуром Эддингтоном (1882–1944), первым осуществившим экспериментальную проверку одного из эффектов, предсказанных общей теорией относительности Эйнштейна. Цель автора заключается в том, чтобы дать представление об этой теории, не пользуясь никакими специальными сведениями из области математики, физики или философии. Автор не только излагает вполне определившуюся теорию, но и беспристрастно обсуждает возможные возражения и указывает на слабые места.

Книга предназначена широкому кругу читателей, но может представлять интерес и для специалистов — математиков и физиков, а также философов и историков науки.

Уитроу Дж.

Естественная философия времени.

Настоящая книга, написанная выдающимся английским философом, первым президентом Международного общества по изучению времени Джеральдом Уитроу, представляет собой обобщающий труд, охватывающий проблему времени с разных сторон. Рассматривая проблему времени со стихийно-материалистических позиций, отвергая идеалистические попытки оторвать время от временных вещей, автор пытается дать анализ времени в его объективном отношении ко Вселенной, к пространству и человеку.

Издательство УРСС

Представляет Вам свои лучшие книги:

Брайан Грин

ЭЛЕГАНТНАЯ ВСЕЛЕННАЯ

Суперструны, скрытые размерности и поиски окончательной теории

В течение последнего полувека физики продолжали, основываясь на открытиях своих предшественников, добиваться все более полного понимания принципов устройства мироздания. И вот теперь, спустя много лет после того, как Эйнштейн объявил о своем походе на поиски единой теории, физики считают, что они смогли, наконец, выработать теорию, связывающую все эти прозрения в единое целое — единую теорию, которая в принципе способна объяснить все явления. Эта теория, *теория суперструн*, и является предметом данной книги.

Теория суперструн забрасывает очень широкий невод в пучины мироздания. Это обширная и глубокая теория, охватывающая многие важнейшие концепции, играющие центральную роль в современной физике. Она объединяет законы макромира и микромира, законы, действие которых распространяется в самые дальние дали космического пространства и на мельчайшие частицы материи; поэтому рассказать об этой теории можно по-разному. Автор выбрал подход, который базируется на эволюции наших представлений о пространстве и времени.

Роджер Пенроуз.

НОВЫЙ УМ КОРОЛЯ.

О компьютерах, мышлении и законах физики.

Монография известного физика и математика Роджера Пенроуза посвящена изучению проблемы искусственного интеллекта на основе всестороннего анализа достижений современных наук. Возможно ли моделирование разума? Чтобы найти ответ на этот вопрос, Пенроуз обсуждает широчайший круг явлений: алгоритмизацию математического мышления, машины Тьюринга, теорию сложности, теорему Геделя, телепортацию материи, парадоксы квантовой физики, энтропию, рождение вселенной, черные дыры, строение мозга и многое другое.

Книга вызовет несомненный интерес как у специалистов, так и у широкого круга читателей.

Издательство УРСС

(095) 135-42-46,
(095) 135-44-23,
URSS@URSS.ru

Наши книги можно приобрести в магазинах:

«Библио-Глобус» (м. Лубянка, ул. Мясницкая, 6. Тел. (095) 925-2457)
«Московский дом книги» (м. Арбатская, ул. Новый Арбат, 8. Тел. (095) 203-8242)
«Москва», (м. Охотный ряд, ул. Тверская, 8. Тел. (095) 229-7855)
«Молодая гвардия» (м. Полянка, ул. Б. Полянка, 28. Тел. (095) 238-5003, 238-7144)
«Дом деловой книги» (м. Проспект Мира, ул. Марьинская, 9. Тел. (095) 270-5421)
«Старый Свет» (м. Пушкинская, Тверской б-р, 23. Тел. (095) 202-3894)
«Гнозис» (м. Университет, 1 гум. корпус МГУ, комн. 141. Тел. (095) 939-4713)
«У Кентавра» (РГГУ) (м. Новослободская, ул. Чаянова, 15. Тел. (095) 973-4301)
«СПб. дом книги» (Невский пр., 28. Тел. (812) 311-2954)

Временной порядок в механике

Направление времени в термодинамике и микростатистике

Направление времени в макростатистике

Время в квантовой физике

Настоящее произведение известного логика и философа Ганса Рейхенбаха является одной из наиболее важных его работ по изучению проблемы времени. В книге приводятся исторический обзор и отчасти критическая оценка различных интерпретаций данной проблемы. Характер обсуждения проблемы времени у Г. Рейхенбаха значительно отличается от постановки вопроса в традиционной философии.

В этой последней из своих книг автор стремится подвести итог всем своим философским исследованиям и часто рассматривает ее как труд, завершающий и обобщающий все его работы, касающиеся вопросов теории вероятностей, квантовой механики, теории относительности и причинности — вопросов, которые играют важную роль в этой книге.

Данный труд является также примером применения аналитического метода научной философии, в развитие и совершенствование которого Г. Рейхенбах также внес свой вклад.

ИЗДАТЕЛЬСТВО УРСС
НАУЧНОЙ И УЧЕБНОЙ ЛИТЕРАТУРЫ
E-mail: URSS@URSS.ru
Каталог изданий
в *Internet*: http://URSS.ru
Тел./факс: 7 (095) 135-44-23
Тел./факс: 7 (095) 135-42-46

1813 ID 12017
9 785354 002757 >