

УСПЕХИ ФИЗИЧЕСКИХ НАУК**ОСНОВАНИЯ СТАТИСТИЧЕСКОЙ МЕХАНИКИ *)****Д. Тер-Хар**

СОДЕРЖАНИЕ

Введение	620
I. <i>H</i> -теорема в классической статистической механике	628
I.1. Кинетический аспект <i>H</i> -теоремы; гипотеза о числе столкновений	628
I.2. Парадокс обратимости и парадокс повторимости	633
I.3. Статистические аспекты <i>H</i> -теоремы	637
II. Классическая эргодическая теорема	647
II.1. Эргодическая и квазиэргодическая теоремы	647
II.2. Новейшие исследования	651
III. <i>H</i> -теорема в классической теории ансамблей	655
III.1. Мелкоструктурная и крупноструктурная плотности	655
III.2. Изображающие ансамбли	664
Резюме положения в классической термостатистике	669
IV. <i>H</i> -теорема в квантовой статистике. 1. <i>H</i> -теорема в элементарном методе рассмотрения. 2. <i>H</i> -теорема в квантовой теории ансамблей. 3. Изображающие ансамбли V. Эргодическая теорема в квантовой статистике. Резюме положения в квантовой термостатистике. Приложение I. Модель Лоренца. Приложение II. Доказательство существования квазиэргодических систем. Приложение III. Доказательство классической эргодической теоремы. Приложение IV. Доказательство квантово-механической эргодической теоремы. Приложение V. Переходы, пропорциональные времени. Приложение VI. Лемма Клейна. Приложение VII. Принцип детального равновесия. Цитированная литература	

*) Foundations of statistical Mechanics. Reviews of modern Physics. 27, 289, 1955. Перевод Г. Ф. Жаркова.

ВВЕДЕНИЕ

В последние годы делались многочисленные попытки ^{44K, 46B, 46G, 46K, 47B1, 47B2, 48B1, 48B2, 49B1, 49B2, 49B3, 49K1, 49K2, 52K, 52R, 52S2 *)} оправдать использование статистической механики для описания физических систем. Что касается классической, т. е. доквантовой статистической механики, то П. и Т. Эренфест в известном обзоре, опубликованном в Энциклопедии математических наук, проанализировали её основания и, как это будет видно из дальнейшего, их выводы, в большей или меньшей степени, остались справедливыми до настоящего времени**). Однако аналогичного обзора современного положения обоснования квантовой статистической механики до сих пор не существовало, и цель настоящей работы состоит в том, чтобы дать такой обзор. В недавно опубликованной монографии ^{54H1 ***)} автор затронул уже некоторые из вопросов, обсуждаемых ниже. Нужно, однако, учесть, что, с одной стороны, после выхода в свет этой монографии был получен ряд новых результатов, а, с другой стороны, то, что в настоящей работе упор делается скорее на изложение существующего положения с обоснованием статистической механики, чем на историческое развитие этих оснований, что было более подробно изложено ранее. Таким образом, настоящая работа в значительной степени дополняет то рассмотрение вопроса, которое было проведено в ESM. Это, конечно, не означает, что здесь мы будем полностью игнорировать исторические аспекты развития статистической механики.

Нам могут сказать, что проверить качество пирога можно, лишь съев его, и что раз статистическая механика оказывается в состоянии достаточно точно и с успехом предсказывать поведение физических систем при условиях равновесия и даже, при определённых обстоятельствах, в неравновесных состояниях ****), то этот факт должен в достаточной степени оправдывать используемые методы. Это тем более должно было быть справедливо в наши дни, когда использование квантовой механики позволило объяснить такие, казалось бы непреодолимые трудности, как парадокс с удельной тепло-

*) Ссылки помещены в конце статьи в хронологическом порядке. Каждый из номеров ссылки состоит из года публикации и первой буквы фамилии автора: например, ссылка на работу М. J. Klein, Phys. Rev. 87, 111 (1952) имеет вид ^{52K}.

**) Р. Курт ^{55K} также рассматривал основания классической термостатики с более аксиоматической точки зрения. Я должен выразить ему благодарность за возможность ознакомиться с его работой до опубликования.

***)) Мы будем использовать, насколько возможно, обозначения этой работы, которую будем цитировать как ESM.

****)) Для ознакомления со статистической механикой неравновесных процессов мы отсылаем читателя к недавно опубликованной обзорной статье Монтроля и Грина ^{55M} и имеющейся там библиографии.

ёмкостью и парадокс Гиббса. Однако логически вполне естественно, что многие физики не удовлетворились приведённым аргументом и попытались показать, что статистический формализм непосредственным образом следует из классической или квантовой механики. В настоящем обзоре будет изложено, в какой степени попытки этих авторов были успешными. Даже в тех случаях, когда успех был ограничен, эти попытки были чрезвычайно важны и полезны, поскольку они выявляли как возможности, так и ограничения статистического подхода.

Прежде чем начать обсуждение различных работ, посвящённых основным положениям статистической механики, необходимо отметить, что мы главным образом будем иметь дело со следующими двумя вопросами:

(А) Почему возможно описание поведения почти всех физических систем путём рассмотрения лишь равновесных состояний?

(В) Почему возможно описание поведения реальной физической системы путём рассмотрения большого числа одинаковых систем (в теории ансамблей) и отождествление усреднённого поведения этой группы систем с поведением интересующей нас физической системы?

Из последующего станет ясным, что эти два вопроса тесно связаны друг с другом. Первый из них возникает не только в статистической механике, но также и в термодинамике и кинетической теории и связанная с ним проблема может быть сформулирована в виде вопроса:

(А1) Каким образом можно определить равновесное состояние?

Второй вопрос имеет чисто статистическую природу и в свою очередь связан с вопросом:

(В1) Каким образом можно построить такой ансамбль, который «изображал» бы действительную, заданную физическую систему?

Этот вопрос об изображающих ансамблях (representative ansembles), который будет играть важную роль в нашем рассмотрении, был особо выделен Толменом ³⁸т. ⁴⁰т.

Мы достигли теперь того момента, когда необходимо ввести определение статистической механики. Мы сделаем это, используя определение, данное Крамерсом во вступительной речи на Международной конференции по статистической механике во Флоренции в 1949 г. ⁴⁹КЗ: «Для физика наших дней статистическая механика является как раз той отраслью физики, которая имеет дело с атомистической интерпретацией тепловых свойств вещества и излучения; по этой причине её можно назвать также «термостатистикой».

Наше изложение не будет касаться явным образом излучения, мы ограничимся системами, состоящими из частиц. Однако при обсуждении переходов из одного состояния в другое будет молчаливо предполагаться существование поля излучения, которое во многих случаях является тем механизмом, благодаря которому переходы могут осуществляться.

Статистическая механика возникла из кинетической теории газов, развитой в девятнадцатом веке. Отцом кинетической теории, повидимому, следует назвать Кронига, хотя более чем за сто лет до него Бернулли уже связывал свойства газа со свойствами отдельных частиц. Крониг^{56K} обратил внимание на тот факт, что хотя путь отдельного атома в газе настолько нерегулярен, что проследить за этим атомом совершенно невозможно, тем не менее можно, используя теорию вероятностей, свести такое совершенно хаотическое поведение к упорядоченному поведению газа. Кинетическая теория газов развивалась далее в работах Клаузиуса^{57C1, 57C2, 58C, 62C, 70C1, 70C2}, Максвелла^{60M, 67M, 68M1, 68M2} и Больцмана^{68B, 72B, 96B, 98B}*). В своей «Теории газов» Больцман подверг особому рассмотрению вопрос (A) и в 1872 году предложил^{72B} свою знаменитую *H*-теорему с целью показать, что любое неравновесное состояние будет изменяться таким образом, что оно будет приближаться к равновесному состоянию. Ответ на вопрос (A1) в этом случае состоит в том, что равновесное состояние является наиболее вероятным, точнее, наиболее вероятным состоянием, совместимым с некоторыми ограничительными условиями. В своей первоначальной неограниченной форме *H*-теорема, если её принять, доказывала, что любая система будет стремиться к равновесию, если равновесия в начальный момент не существовало**). Это означало бы, что если подождать достаточно долго, то система обязательно оказалась бы в состоянии равновесия и, более того, это состояние равновесия затем существовало бы вечно. Отсюда следовало бы, что равновесие представляет собой правило, а неравновесие — исключение, и таким образом был бы дан ответ на вопрос (A).

Однако скоро было осознано, что *H*-теорема в своей первоначальной неограниченной форме не может быть абсолютно строго доказана, а основывается на определённых предположениях относительно числа столкновений, которые данная частица претерпевает за некоторый промежуток времени. Более того, Лошмидт^{76L, 77L} и Цермело^{96Z} убедительно показали, что *H*-теорема в своей первоначальной форме не может быть справедливой. В своих последующих работах Больцман, ввиду этого тщательно подчёркивал статистические аспекты *H*-теоремы. Это означает, что *H*-теорема является утверждением о наиболее вероятном поведении системы. Однако флуктуации около равновесного состояния более не запрещены. Статистические аспекты *H*-теоремы особо подчёркивались Эренфе-

*) Для более подробного ознакомления с историческими деталями следует обратиться к библиографическим ссылкам в конце глав I и II и в конце приложения I в ESM.

**) Мы исключаем, разумеется, такие гипотетические системы, как полностью идеальные газы, находящиеся между идеальными стенками. В таких системах отсутствовал бы механизм, способный привести к изменению функции распределения и неравновесное состояние существовало бы неограниченно долго.

стом ^{07E, 11E2}, как исходя из общей точки зрения, так и при рассмотрении простых моделей *).

В первой части настоящей статьи будут обсуждены основания статистической механики. В разделе I.1 будет обсуждена неограниченная *H*-теорема. Аргументы Лошмидта и Цермело против первоначальной формулировки *H*-теоремы приводятся в разделе I. 2, в то время как в разделе I.3 обсуждаются статистические аспекты *H*-теоремы.

Кроме введения *H*-теоремы для доказательства (А), Больцман пытался также доказать, что усреднённое поведение системы совпадает с её равновесным поведением. Содержание этого утверждения состоит в том, что среднее по времени, взятое за бесконечный промежуток времени от фазовой функции, т. е. от функции, зависящей от значений координат и скоростей, которые полностью определяют физическое состояние или фазу рассматриваемой системы, должно быть равно значению фазовой функции в равновесии. Легко можно убедиться в том, что этот второй путь подхода к ответу на вопрос (А) эквивалентен первому пути. Во-первых, если *H*-теорема верна, то отсюда следует, что любое неравновесное состояние будет переходить в равновесное состояние, которое затем будет существовать вечно. Беря среднее по времени от фазовой функции за бесконечный промежуток времени, мы таким образом получим значение этой фазовой функции в равновесии. Во-вторых, если среднее по времени совпадает со значением в равновесии, это означает, что система должна находиться в состоянии равновесия большую часть времени и, следовательно, возвращаться к равновесию из любого неравновесного состояния **). Для доказательства того, что усреднённое поведение и поведение в равновесии совпадают, необходимо вычислить средние по времени; при этом встречаются серьёзные трудности. Больцман пытался преодолеть эти трудности, предположив, что большинство физических систем являются эргодическими ***). Эргодическая система — это такая система, изображающая точка которой в Γ -пространстве ****) проходит через каждую точку энерге-

*) Статистическая природа *H*-теоремы была недостаточно принята во внимание в недавней работе Сартре ^{54S}.

**) Легко убедиться, что эта эквивалентность останется в силе, если заменить *H*-теорему в её статистической форме на её неограниченную форму.

***) Эргодический — от греческих слов *ergon* (работа, используется здесь в смысле энергия) и *hodos* (путь): изображающая точка в Γ -пространстве проходит все точки энергетической поверхности. Термин «эргодический» был введён Больцманом в 1887 г. ^{87B}. Предположение эргодичности Максвелл называл предположением непрерывности пути.

****) Система с *S* степенями свободы может быть описана $\bar{S}$ (обобщёнными) координатами и $\bar{S}$ (обобщёнными) импульсами. Значения этих $2\bar{S}$ величин в заданный момент времени определяют точку (так называемую изображающую точку) в $2\bar{S}$ -мерном пространстве, называемом Γ -пространством (Γ от слова газ), или фазовым пространством. $2\bar{S}$ — 1-мерная гиперповерхность в Γ -пространстве, задаваемая уравнением $\epsilon = \text{const}$, где ϵ — энергия системы, называется энергетической поверхностью.

тической поверхности, соответствующей энергии системы. Вычисление среднего по времени в этом случае совпадает с вычислением среднего по энергетической поверхности, т. е. со средним по совокупности систем, обладающих той же самой энергией. Такая специфическая совокупность систем, или ансамбль, если использовать термин, введенный Гиббсом, представляет собой так называемый микроканонический ансамбль. Поскольку, согласно нашему предположению, изображающая точка любой системы ансамбля будет проходить через каждую точку энергетической поверхности, то орбиты, описываемые изображающими точками систем ансамбля, будут все совпадать друг с другом, с той единственной разницей между системами, что точный момент времени прохождения через данную точку орбиты для разных систем будет разным*). Из этого следует, что среднее по времени, представляющее собой среднее, взятое по одной и единственной орбите на энергетической поверхности, которую различные точки проходят в различные моменты времени, даст такой же результат, что и среднее, взятое по микроканоническому ансамблю и представляющее собой среднее, взятое по той же самой орбите, но где теперь все изображающие точки рассматриваются в один и тот же момент времени.

При рассмотрении средних по энергетической поверхности, или по микроканоническому ансамблю, Больцман отошёл от кинетической теории и вступил поистине в царство статистической механики, если использовать здесь определение статистической механики, данное Гиббсом в предисловии к его знаменитой монографии ^{02G **}).

*) Мы используем здесь тот факт, что направление орбиты в любой точке Г-пространства однозначно определяется из гамильтоновских уравнений движения. Из этого следует также, что орбита никогда не может вернуться в ту точку Г-пространства, через которую она уже прошла, если только мы не имеем дело со строго периодическими орбитами.

**) В этом месте мы могли бы посоветовать каждому изучающему статистическую механику прочесть и внимательно перечесть это предисловие, поскольку в нём наиболее ясно, чем где бы то ни было, сформулированы основные идеи рациональной термодинамики, как её называл Гиббс. Именно в этом предисловии был введен также и сам термин статистическая механика. Ввиду трудностей, вставших перед кинетической теорией во времена появления монографии Гиббса (парадокс с удельными теплоемкостями, трудность с пониманием закона излучения Планка и т. д.), Гиббс попытался создать рациональную систему, основанную на небольшом числе аксиом и не обязательно связанную с явлениями природы. Приведем здесь слова Гиббса:

«Хорошо известно, что хотя теория приписывает каждой молекуле газа шесть степеней свободы, наши опыты с теплоемкостью приводят к учёту не более чем пяти степеней. Конечно, тот, кто основывает свою работу на гипотезах, касающихся строения материи, стоит на ненадёжном фундаменте.

Затруднения этого рода удержали автора от попыток объяснения тайн природы и заставили его удовлетвориться более скромной задачей вывода некоторых более очевидных положений, относящихся к статистической отрасли механики. При этом здесь уже не может быть ошибки с точки зрения согласия гипотез с фактами природы, ибо в этом отношении ничего

Необходимо, однако, отметить, что теория ансамблей проникает, если можно так выразиться, через заднюю дверь, поскольку она вводится лишь как математический приём для вычисления поведения изолированной системы.

В начале этого столетия, в связи с развитием основ теории, было осознано, что эргодические системы никогда не осуществляются, однако выражалась надежда, что большинство физических систем является квазиэргодическими. Квазиэргодическая система представляет собой такую систему, которая имеет орбиту в Γ -пространстве, покрывающую энергетическую поверхность всюду плотно, хотя, в действительности, и не проходящую через каждую точку этой поверхности. Выражалась также надежда, что квазиэргодичности окажется достаточно для обеспечения равенства средних по времени и средних, взятых по соответствующим образом выбранному ансамблю. Такова была, например, точка зрения Эренфеста ^{11E2}.

В 1913 г. Розенталь ^{13R} и Планшерель ^{13P} независимо друг от друга доказали невозможность существования эргодических систем, а в 1923 г. Ферми ^{23F1} показал, что определённый класс систем является квазиэргодическим. Он, однако, не доказал равенства средних по времени и средних по ансамблю *). Это равенство доказывалось Розенталем ^{14R2}, но его доказательство оказалось недостаточно строгим **). В последнее время проблемой эквивалентности средних по времени и по ансамблю занимались, в основном, математики и её обычно называют эргодической теоремой. В 1931 и 1932 гг. Биркхоф ^{31B1, 31B2, 32B} и фон-Нейман ^{32N1, 32N2} установили, что при выполнении некоторых правдоподобных математических условий эти два типа средних дают одинаковый результат, а затем Окстоби и Улам ^{41O} недавно показали, что весьма широкий класс систем удовлетворяет этим математическим условиям. Эргодическая теорема будет обсуждаться в разделе II. В разделе II.1 будут обсуждены работы, опубликованные до 1930 г., в то время как позднейшие исследования рассматриваются в разделе II. 2.

Многие физики до сих пор считают эргодическую теорему наиболее удовлетворительным, если не единственным оправданием для использования статистических методов. Эти люди считают, что механика может иметь дело лишь с изолированными системами и что статистический подход представляет собой необходимый математический приём, не представляющий особого физического интереса.

и не предполагается. Единственной ошибкой, в которую можно впасть, является недостаточное согласие между предпосылками и выводами, а это, при некоторой осторожности, можно надеяться в основном избежать».

*) Ниже, однако, мы увидим, что это равенство в действительности имеет место для всех квазиэргодических систем.

**) В этом доказательстве был дважды произведён переход к пределу, на что было указано самим Розенталем в ответе Эпштейну (см. ^{36E}, стр. 478).

Автор настоящей статьи думает, что такая точка зрения не отвечает физическим основаниям для обращения к статистическому подходу *). Эти основания состоят в том, что огромное число частиц, образующих большинство физических систем, делает невозможным определение значений всех интегралов движения. Вообще говоря, мы можем определить лишь энергию, полный импульс и, может быть, ещё один-два интеграла. После этого нам необходимо использовать полученные весьма скудные сведения о системе для предсказания её поведения в будущем. Поскольку наших знаний совершенно недостаточно для предсказаний будущего системы с полной определённости, мы должны обратиться к статистическим методам. Именно здесь входят в рассмотрение так называемые *изображающие ансамбли* **).

Вместо того, чтобы рассматривать одну систему, будем рассматривать большое число систем, обладающих в точности такими же значениями тех величин, которые нам известны, а в остальном широко отличающихся. Для того чтобы удовлетворительным образом построить такой изображающий ансамбль, необходимо знать, какой вес приписать различным системам ансамбля, или, другими словами, необходимо сделать предположения относительно априорных вероятностей.

Если мы согласились с необходимостью изображающих ансамблей, то вопрос (B) разрешается в теории ансамблей при помощи доказательства того, что подавляющая часть систем рассматриваемых ансамблей ведёт себя практически так же, как это можно было бы ожидать от равновесной системы и, более того, приводит к тем же самым значениям фазовых функций ***). Однако вопрос (B1) остаётся ещё открытым. Всегда оказывается, что канонический ансамбль представляет собой изображающий ансамбль системы в температурном равновесии. Для доказательства этого вводятся обобщённая *H*-теорема и показывается, что если изображающий ансамбль в некоторый момент времени не является каноническим ансамблем, положение

*) Необходимо также добавить, что полностью изолированные системы не представляют для физиков интереса, поскольку с ними невозможно проводить эксперименты. Разумеется, если мы имеем дело с классическими системами, можно всегда рассмотреть идеализированные эксперименты и таким образом включить в рассмотрение и изолированные системы. Более того, Паули ^{49P} указал на то, что и в случае квантово-механических систем можно пренебречь влиянием акта наблюдения на систему постольку, поскольку мы имеем дело со статистической механикой, хотя для самой квантовой механики это представляет первостепенное значение.

**) См. разделы 24, 86, 112, 117 и 120 монографии Толмена ^{38T} и часть B работы ESM, где более подробно рассмотрен вопрос об изображающих ансамблях; см. также разделы III. 2 и IV.3 настоящей статьи.

***) Доказательство этого утверждения можно найти в любом учебнике по статистической механике, рассматривающем теорию ансамблей ^{02G, 38T, 54H1}.

ние в последующие моменты времени будет таково, что мы будем вынуждены использовать изображающий ансамбль, имеющий гораздо более близкое сходство с каноническим. Этот вопрос приближения к равновесию, описываемый ансамблями, и вопрос об изображающих ансамблях будут обсуждены в части III в той степени, в какой это касается классической теории ансамблей.

Переход от классической к квантовой статистике не вносит никаких фундаментальных изменений. В самом деле, можно параллельно рассмотреть оба случая, т. е. классический и квантово-механический, причём выявится их глубокая аналогия. Однако между ними имеются и определённые отличия, вследствие чего мы предпочитаем обсуждать каждый из этих случаев отдельно. В квантово-механическом случае мы снова можем подойти к проблеме либо посредством теории ансамблей, либо посредством эргодической теоремы. Эти два альтернативных метода подхода обсуждаются в частях IV и V соответственно.

В заключение вводных замечаний мне хотелось бы высказать несколько соображений относительно некоторых общих аспектов второго закона термодинамики и его места в статистической механике *). Тот факт, что во всех практических применениях термодинамики справедлив второй закон, т. е. энтропия возрастает, обязан комбинированному действию двух эффектов. Прежде всего, если энтропия в некоторый момент времени меньше её равновесного значения, то вероятность её возрастания подавляюще велика по сравнению с вероятностью того, что она будет уменьшаться. Во-вторых, наши наблюдения всегда происходят так, что, начав с заданной ситуации, мы наблюдаем дальнейшее развитие, однако для нас невозможно начать с заданной ситуации и наблюдать в предшествующие моменты времени. Тот факт, что мы можем делать одно и не можем делать другого, связан с иным фактом, а именно с тем, что у нас имеется память о прошлом и мы можем, следовательно, знать то, что происходило в предшествующие моменты времени, но не то, что произойдёт в дальнейшем. В этом смысле необратимость второго закона термодинамики психологически весьма обоснованна.

Другое замечание касается того, что, насколько можно судить в настоящее время, все наблюдения явлений во вселенной совместимы с представлением о развитии вселенной в целом из некоторого возможного, хотя термодинамически или статистически неправдоподобного состояния, удалённого в далёкое прошлое. Это возможно связать с космологическими идеями о более или менее сингулярном состоянии, имевшемся, допустим, три или пять миллиардов лет назад. С другой стороны, можно привести соображения, как это было сделано, например, Больцманом ^{95B}, о том, что представление о раз-

*) Я должен выразить благодарность проф. Р. Пайерлсу за его ценные критические замечания по этому вопросу.

витии нашего мира из менее вероятного в более вероятное состояние не обязательно находится в противоречии с представлением о вселенной, находящейся, как целое, в термодинамическом равновесии. Протицируем Больцмана: «Мы предполагаем, что вся вселенная находится, и всегда находилась, в тепловом равновесии. Вероятность того, что одна (и лишь одна) часть вселенной находится в некотором состоянии, будет тем меньше, чем больше удалено это состояние от состояния теплового равновесия, однако эта вероятность будет тем больше, чем больше сама вселенная. Если предположить, что вселенная достаточно велика, то вероятность нахождения её относительно малой части в любом заданном состоянии (удалённом, однако, от состояния равновесия), может быть как угодно велика. Мы можем получить большую вероятность также и того, что хотя вся вселенная находится в тепловом равновесии, наш мир настолько удалён от теплового равновесия, что невозможно даже вообразить чудовищно малую вероятность такого состояния. Однако, разве можем мы представить себе, с другой стороны, сколь малую часть всей вселенной занимает этот мир? Если же вселенная достаточно велика, то вероятность того, что её столь малая часть, как наш мир, находится в существующем состоянии, оказывается уже отнюдь не малой.

Если бы это предположение было правильно, всё постепенно возвращалось бы к тепловому равновесию; однако, так как вся вселенная настолько велика, то возможно, что когда-либо в будущем какой-нибудь другой мир столь же отклонится от равновесия, как это имеет место сейчас в нашем мире. Тогда упомянутая выше *H*-кривая *) изображала бы то, что происходит во вселенной. Максимумы кривой соответствовали бы мирам, в которых существовали бы видимые движения и жизнь».

Приводя эту цитату, мы не присоединяемся к точке зрения Больцмана, а лишь упоминаем о ней. При обсуждении этой точки зрения по существу необходимо исследовать сложную проблему, насколько возможно говорить о тепловом равновесии в незамкнутой системе, какой является наша вселенная.

1. *H*-ТЕОРЕМА В КЛАССИЧЕСКОЙ СТАТИСТИЧЕСКОЙ МЕХАНИКЕ

1.1. Кинетический аспект *H*-теоремы; гипотеза о числе столкновений

В 1872 г. Больцман ^{72B} предложил свою знаменитую *H*-теорему для доказательства того, что любое неравновесное распределение будет стремиться перейти в равновесное. Коротко изложим его

*) По существу, это кривая, дающая энтропию как функцию времени; ср. раздел 1.3.

рассуждения. Он рассматривал случай изолированной системы, состоящей из газа (в частности, из одноатомного газа в отсутствие внешних сил), хотя в дальнейшем он обобщил своё рассмотрение на случай многоатомных газов, находящихся под воздействием внешних сил ^{75В}.

Будем интересоваться вычислением скорости изменения функции распределения f , которая в нашем простейшем случае является функцией лишь трёх (декартовых) компонент u , v и w скорости атома c и времени t . Нам необходимо, стало быть, с одной стороны, учесть быстроту изменения скоростей атомов, находящихся в интервале c и $c + dc$ и, с другой стороны, учесть скорость, с которой при столкновениях образуются атомы со скоростями внутри заданного интервала. В формуле

$$\frac{\partial f}{\partial t} du dv dw dt = -A + B \quad (I.1,1)$$

величина $f du dv dw$ представляет собой число атомов в единице объёма, имеющих скорости с компонентами, лежащими в интервалах $(u, u + du)$, $(v, v + dv)$ и $(w, w + dw)$; A — число атомов в единице объёма со скоростями, лежащими внутри указанной области, которые за промежуток времени dt изменяют свои скорости; B — число атомов в единице объёма, которые за время dt изменяют свои скорости так, что после столкновения их скорости попадают в указанную выше область. Можно показать (см., например, ESM, раздел I.4), что A даётся выражением

$$A = f(u, v, w) du dv dw dt \int f(u_1, v_1, w_1) du_1 dv_1 dw_1 \int a d\omega. \quad (I.1,2)$$

В выражении (I.1,2) $d\omega$ представляет собой элемент телесного угла, охватывающего так называемую линию центров ω , представляющую собой вектор, направленный от одного из сталкивающихся атомов к другому в момент столкновения; a — величина, зависящая лишь от абсолютного значения относительной скорости $c_{отн}$ и от угла θ между вектором относительной скорости и ω ; интегрирование ведётся по всем значениям u_1 , v_1 и w_1 и по всем возможным направлениям линии центров. Выражение (I.1,2) получается путём умножения числа атомов со скоростями, лежащими внутри выбранного интервала, т. е. $f(u, v, w) du dv dw$, на число атомов, сталкивающихся с одним выбранным атомом за промежуток времени dt . Рассмотрим атомы, имеющие скорости между c_1 (компоненты u_1, v_1, w_1) и $c_1 + dc_1$. Если предположить, что отсутствует корреляция между скоростями и положениями различных атомов, то число столкновений между этими атомами и каким-либо одним выбранным атомом, в случае, когда линия центров лежит внутри телесного угла $d\omega$, будет равно $af(u_1, v_1, w_1) \times du_1 dv_1 dw_1 dt d\omega$, где величина a — нормировочная константа, зависящая от $c_{отн}$ и θ . Выражение (I.1,2) получается тогда после ин-

тегрирования по всем возможным значениям u_1, v_1, w_1 и по всем возможным направлениям ω .

Аналогично для B имеем выражение

$$B = dt \int du' dv' dw' f(u', v', w') du'_1 dv'_1 dw'_1 f(u'_1, v'_1, w'_1) \int a' d\omega'. \quad (I.1,3)$$

Штрихи у знаков интеграла указывают на то, что теперь интегрирование ведётся лишь по таким значениям c' и c'_1 и по таким направлениям ω' , чтобы скорости после столкновения, полностью определяемые значениями c', c'_1 и ω' , удовлетворяли бы условию, что одна из них попадёт в интервал c и $c + dc$. Выражая штрихованные величины через нештрихованные, можно вместо (I.1,3) (ср., например, ESM, раздел I.4) написать

$$B = du dv d\omega dt \int du_1 dv_1 d\omega_1 f(u', v', w') f(u'_1, v'_1, w'_1) \int a d\omega, \quad (I.1,4)$$

где u', v', w', u'_1, v'_1 и w'_1 являются функциями от u, v, w, u_1, v_1 и w_1 , ввиду наличия уравнений сохранения импульса и энергии. Комбинируя (I.1,1), (I.1,2) и (I.1,4), получим для быстроты изменения функции распределения из-за столкновений

$$\frac{\partial f}{\partial t} = - \int (ff_1 - f'f'_1) a du_1 dv_1 d\omega_1 d\omega, \quad (I.1,5)$$

где для краткости использованы обозначения

$$\left. \begin{aligned} f &= f(u, v, w), & f' &= f(u', v', w'), \\ f_1 &= f(u_1, v_1, w_1), & f'_1 &= f(u'_1, v'_1, w'_1). \end{aligned} \right\} \quad (I.1,6)$$

Вводя величину H посредством равенства *)

$$H = \int f \ln f du dv d\omega, \quad (I.1,7)$$

получим для скорости изменения H из (I.1,5)

$$\frac{dH}{dt} = \int (f'f'_1 - ff_1) (\ln f + 1) a du dv d\omega du_1 dv_1 d\omega_1 d\omega. \quad (I.1,8)$$

Из-за эквивалентности величин c, c_1, c', c'_1 в (I.1,8) это уравнение можно переписать в виде (см., например, ESM, раздел I.5)

$$\frac{dH}{dt} = \frac{1}{4} \int (f'f'_1 - ff_1) \ln \frac{ff_1}{f'f'_1} a du dv d\omega du_1 dv_1 d\omega_1 d\omega. \quad (I.1,9)$$

*) В своей первой работе, посвящённой H -теореме, Больцман использовал обозначение E (энтропия) вместо H .

Поскольку

$$(p - q) \ln\left(\frac{p}{q}\right) > 0, \text{ если } p \neq q; \text{ и } = 0, \text{ если } p = q \quad (\text{I.1,10})$$

и поскольку a — положительная величина, имеем из (I.1,9)

$$\frac{dH}{dt} \leq 0, \quad (\text{I.1,11})$$

причём равенство в (I.1,11) имеет место лишь в том случае, если для любых двух скоростей c и c_1 имеем $ff_1 = f'f'_1$.

Таким образом, из (I.1,11) следует, что при условии справедливости наших основных предположений величина H будет монотонно убывать, пока она не достигнет минимального значения, которое достигается, как только функция распределения удовлетворит условию

$$ff_1 = f'f'_1 \quad (\text{I.1,12})$$

для любой пары скоростей c и c_1 . Минимальное или равновесное значение H в большинстве случаев достигается быстро, если H заметно отличается от этого значения. Время релаксации в случае, например, газа при комнатной температуре и давлении в одну атмосферу, порядка 10^{-7} сек. (ESM, стр. 390).

Формула для равновесной функции распределения может быть выведена либо при использовании (I.1,12), либо если взять в качестве условия равновесия требование, чтобы в равновесии H имела минимум — в обоих случаях при дополнительных условиях постоянства полного числа частиц, постоянства полной энергии и постоянства полного импульса. В результате получается обобщённое максвелловское распределение.

К сожалению, случай реального газа трудно рассмотреть во всех деталях и хотя вывод H -теоремы всё же возможен, однако, представляется нелёгким делом подвергнуть критическому анализу различные допущения, делаемые для получения (I.1,11), и последствия отказа от дополнительных условий. Возможно, однако, ввести упрощённые модели ^{11E2, 53H, 54G, 55G, 55H2} и мы особо рассмотрим модель, которая по существу совпадает с той, которую использовал Лоренц ^{09L} в своих работах по электронной теории металлов. В этой модели имеется два типа частиц. Частицы первого типа (узлы решётки, или ионы металла) пространственно закреплены и мы предположим, что они беспорядочно распределены в пространстве с плотностью n на единицу объёма. Частицы второго типа (электроны) движутся сквозь решётку и их плотность равна N на единицу объёма. Будем пренебрегать столкновениями электронов с электронами и предположим, далее, что столкновения электронов с решёткой являются упругими и изотропными. Это означает, что (а) при таких столкновениях электрон не меняет абсолютной величины своей скорости и что (б) если $p(\theta, \varphi; \theta', \varphi')$ — вероятность

того, что электрон, движущийся в направлении, определяемом полярными углами θ и φ , меняет движение на направление, определяемое углами θ' и φ' , то значение $p(\theta, \varphi; \theta', \varphi')$ не зависит от θ' и φ' . Поскольку абсолютная величина скорости не меняется при столкновении, можно упростить нашу модель, предположив, что все электроны движутся с одинаковой скоростью c . Наконец, обозначим через σ сечение для столкновения электрона с решёткой.

Для рассмотрения приближения к равновесию введём функцию распределения. Пусть $f(\theta, \varphi) d\omega/4\pi$ ($d\omega = \sin \theta d\theta d\varphi$ — элемент телесного угла) равно числу электронов в единице объёма, имеющих скорости, лежащие внутри телесного угла $d\omega$. Условие нормировки тогда получается в виде

$$\int f(\theta, \varphi) d\omega = 4\pi N, \quad (I.1,13)$$

где интегрирование распространено по единичной сфере.

Согласно нашему предположению относительно модели, для числа электронов в единице объёма $N_{\omega, \omega'} dt d\omega d\omega'$, меняющих направления своих скоростей от элемента телесного угла $d\omega$ к телесному углу $d\omega'$ за время dt , имеем

$$N_{\omega, \omega'} dt d\omega d\omega' = f(\theta, \varphi) nc\sigma \frac{d\omega}{4\pi} \frac{d\omega'}{4\pi} dt, \quad (I.1,14)$$

причём мы использовали предположение об изотропии рассеяния.

Используя (I.1,13) и (I.1,14), можно легко вычислить скорость изменения $f(\theta, \varphi)$ и получить

$$\frac{df(\theta, \varphi)}{dt} = \alpha \left[\frac{1}{4\pi} \int f(\theta', \varphi') d\omega' - f(\theta, \varphi) \right] = \alpha [N - f], \quad (I.1,15)$$

где

$$\alpha = nc\sigma. \quad (I.1,16)$$

Из (I.1,15) находим

$$f(\theta, \varphi) = N - Ae^{-\alpha t}, \quad A = A(\theta, \varphi) = N - f_{t=0}(\theta, \varphi), \quad (I.1,17)$$

т. е. экспоненциальное приближение к равновесному значению, если при $t=0$ равновесие отсутствовало.

Опять, как и в случае (I.1,11) мы обнаруживаем монотонное приближение к равновесному состоянию. Однако, как и раньше, мы сделали фундаментальное предположение при написании (I.1,14) и опять это предположение касается числа столкновений; такое предположение часто называют гипотезой о числе столкновений (Stosszahlansatz). Так как столкновения играют решающую роль в этом рассмотрении, будем называть такой метод рассмотрения проблемы кинетическим. В следующем разделе будет показано, однако, что неограниченная гипотеза о числе столкновений приводит к противоречиям и что ввиду этого необходимо основывать своё рассмотрение скорее на так называемом статистическом методе, который в должной степени учитывает возможные флуктуации.

1.2. Парадокс обратимости и парадокс повторимости

Сформулируем H -теорему в несколько отличной форме. Мы изучали системы, составленные из большого числа тождественных частиц, и рассматривали функцию распределения, соответствующую такой системе, и быстроту её изменения. Другой способ изучения поведения системы состоит в том, чтобы описывать её состояние, или фазу, в каждый момент времени изображающей точкой в Γ -пространстве (см. ESM стр. 100, где дано определение Γ -пространства). Каждому состоянию, с одной стороны, соответствует некоторое распределение и, с другой стороны, соответствует некоторая точка в Γ -пространстве. Таким образом, имеется однозначное соответствие между точками в Γ -пространстве, с одной стороны, и значениями функции распределения, с другой*). Рассмотрим теперь последовательность времён $t_1, t_2, \dots, t_n, \dots$. Фаза нашей системы в момент t_n будет соответствовать точке P_n в Γ -пространстве и орбита, описываемая изображающей точкой системы, будет проходить через последовательность точек

$$\dots, P_1, P_2, \dots, P_{n-1}, P_n, P_{n+1}, \dots \quad (1.2,1)$$

Так как каждая точка соответствует функции распределения, мы можем вычислить значение H , соответствующее каждой из точек. Пусть H_n будет значением H в состоянии, соответствующем P_n . Если принять H -теорему в её неограниченной форме, то из этого вытекает, что должны выполняться следующие неравенства:

$$H_1 \geq H_2 \geq \dots \geq H_{n-1} \geq H_n \geq \dots, \quad (1.2,2)$$

где знак равенства имеет место лишь в случае достижения равновесия.

В общем случае H даётся вместо (1.1,7) выражением

$$H = \int f \ln f d\omega, \quad (1.2,3)$$

где интегрирование распространено на всё μ -пространство (ESM, стр. 30); здесь f — функция s обобщённых координат q_k и s обобщённых импульсов p_k (s — число степеней свободы одной частицы); $d\omega$ — элемент объёма μ -пространства:

$$d\omega = \prod_{k=1}^s dp_k dq_k. \quad (1.2,4)$$

*) Это справедливо, разумеется, лишь постольку, поскольку мы ограничиваемся рассмотрением сравнительно несущественных видовых фаз. Если использовать родовые фазы, то каждому распределению соответствует совокупность точек в Γ -пространстве [ср. раздел (1.3)].

Рассмотрим теперь два состояния, соответствующие изображающим точкам P_i и P'_i , которые отличаются лишь тем, что все p_k имеют одинаковые абсолютные значения, но противоположные знаки, в то время как q_k в обоих состояниях те же самые. Так как гамильтониан $\mathcal{H}(p_k, q_k)$ системы является однородным квадратичным полиномом по p_k , то он инвариантен относительно преобразования $q_k \rightarrow q'_k$, $p_k \rightarrow p'_k$, которое преобразует P_i в P'_i . Из канонических уравнений движения

$$\dot{q}_k = \frac{\partial \mathcal{H}}{\partial p_k}, \quad \dot{p}_k = -\frac{\partial \mathcal{H}}{\partial q_k} \quad (1.2,5)$$

следует тогда, что преобразование от P_i к P'_i соответствует обращению временной оси. Если рассмотреть теперь систему, проходящую последовательность (1.2,1), то после преобразования мы получим систему, проходящую последовательность

$$\dots, P'_{n+1}, P'_n, P'_{n-1}, \dots, P'_2, P'_1, \dots \quad (1.2,6)$$

Из определения H (1.2,3) следует, что

$$H_i = H'_i. \quad (1.2,7)$$

В этом можно убедиться, выполнив преобразование под знаком интеграла и затем переходя к новым переменным

$$\begin{aligned} H' &= \int f'(p', q') \ln f'(p', q') d\omega' = \int f(-p, q) \ln f(-p, q) d\omega = \\ &= \int f(p'', q'') \ln f(p'', q'') d\omega'' = H, \end{aligned} \quad (1.2,8)$$

где $p''_k = -p_k$, $q''_k = q_k$.

Из (1.2,7) и (1.2,2) следует, что последовательность (1.2,6) соответствует последовательности значений H , удовлетворяющих неравенствам

$$\dots \leq H'_{n+1} \leq H'_n \leq H'_{n-1} \leq \dots \leq H_2 \leq H_1 \leq \dots \quad (1.2,9)$$

Мы видим, таким образом, что каждой системе, для которой H монотонно убывает, можно сопоставить систему, для которой H монотонно растёт. Этот факт представляет собой так называемый парадокс обратимости Лошмидта^{76L, 77L} (Reversibility paradox).

Трудность, заключённая в нём, легко обнаруживается при рассмотрении простой модели, введённой в конце раздела 1.1. Обращая направление времени, получим монотонное удаление от равновесного распределения, что видно из (1.1,17).

Другая трудность была указана Цермело^{96Z}, который пользовался теоремой Пуанкаре^{90P} (см. также^{97B}). Пуанкаре показал, что если система, заключённая в замкнутый объём, проходит последовательность (1.2,1), скажем, от t_1 к t_{n+1} , то по истечении конеч-

ного промежутка времени система вновь будет проходить ту же последовательность с любой, наперед данной точностью *). Более точно, для любого конечного элемента длины Δs в Γ -пространстве можно найти такой промежуток времени T , что последовательность состояний в моменты $t_1 + T, t_2 + T, \dots, t_n + T, t_{n+1} + T$ будет даваться точками

$$\dots, P'_1, P'_2, \dots, P'_n, P'_{n+1}, \dots, \quad (I.2,10)$$

где

$$|P_i - P'_i| < \Delta s. \quad (I.2,11)$$

Промежуток времени может быть чрезвычайно большим. Для случая 10^{18} атомов в 10^{-6} м³, движущихся со средней скоростью $5 \cdot 10^6$ м/сек⁻¹, по оценке Больцмана^{96B} (см, также^{43C}) потребуется более 10^{10} лет для воспроизведения положений с точностью в $10 \text{ \AA}$ и скоростей с точностью в 10^4 м·сек⁻¹.

Если Δs достаточно мало, то для значений H , соответствующих последовательности (I.2,10), имеем

$$H'_i \approx H_i, \quad (I.2,12)$$

и, следовательно, имеем следующую последовательность во времена $t_1, t_2, \dots, t_n, \dots, t_1 + T, \dots, t_n + T$:

$$\dots H_1, H_2, \dots, H_n, \dots, H'_1, H'_2, \dots, H'_n, \dots \quad (I.2,13)$$

Используя (I.2,2) и (I.2,12), убеждаемся, что идя от t_n к $t_1 + T$, мы получаем возрастание H . Этот парадокс и представляет собой так называемый парадокс повторимости (Recurrence paradox).

В следующем разделе мы покажем, каким образом статистическое рассмотрение разъясняет создавшееся положение. Однако уже сейчас нам хотелось бы подчеркнуть, что в той степени, в какой парадокс обратимости и парадокс повторимости опровергают H -теорему в её неограниченной форме (т. е. утверждение, что H никогда не может убывать), они показывают, что гипотеза о числе столкновений не может быть справедлива при любых обстоятельствах. Поясним это на двух простых примерах.

Первый пример даёт нам рассмотрение простейшей модели, обсуждавшейся в предыдущем разделе. Допустим, что система изменяется со временем, пусть Δt — произвольный интервал в течение этого времени, а $f(\theta, \varphi)$ — функция распределения в начале этого интер-

*) За доказательством теоремы Пуанкаре мы отсылаем к литературе^{96Z, 43C}, а также ESM, стр. 341; последнее доказательство не вполне строго). Доказательство следует, по существу, из того факта, что конечная область в Γ -пространстве проходит сквозь всю достижимую часть фазового пространства (которое в случае замкнутой в объёме системы конечно) в течение конечного промежутка времени. Относительно связи теоремы Пуанкаре с эргодической теоремой Биркхофа см. монографию Винтера^{41W}, стр. 90—91.

вала. Обозначим снова через $N_{\omega, \omega'} \Delta t d\omega d\omega'$ число столкновений за этот интервал, приводящих к изменению скорости, находящейся внутри телесного угла $d\omega$, к телесному углу $d\omega'$. Рассмотрим теперь соответствующее изменение системы, обратив направление времени, и введём соответствующий интервал. Функцию распределения в начале интервала времени во втором случае обозначим через $f'(\theta, \varphi)$. Если $N'_{\omega', \omega} \Delta t d\omega' d\omega$ — число столкновений, приводящих к изменению скорости от ω' к ω за интервал Δt обращённого изменения, то имеем

$$N'_{\omega', \omega} = N_{\omega, \omega'} \quad (1.2,14)$$

для любой пары направлений ω и ω' . Из (1.1,14) следует тогда, что если бы гипотеза о числе столкновений была справедлива для прямого и обращённого изменения системы, то было бы

$$f(\theta, \varphi) = f'(\theta', \varphi'), \quad (1.2,15)$$

независимо от значений θ, φ, θ' и φ' , т. е.

$$f(\theta, \varphi) = \text{const.} \quad (1.2,16)$$

Мы видим, таким образом, что гипотеза о числе столкновений не может быть справедлива по крайней мере для одного из двух направлений изменения, если только мы не имеем дело с состоянием равновесия.

Второй пример состоит в следующем*). Рассмотрим поток точечных частиц, движущихся в одном и том же направлении, скажем, в положительном направлении оси x , и рассеиваемых системой непроницаемых сфер, беспорядочно распределённых в плоскости, перпендикулярной оси x . В результате рассеяния точечные частицы после их столкновения с плоскостью, в которой расположены сферы, образуют систему частиц с изотропно распределёнными скоростями. Рассмотрим теперь развитие системы, полученной из предыдущей путём обращения направления всех скоростей всех частиц. Ясно, что через некоторое время все частицы будут двигаться в отрицательном направлении оси x , т. е. система будет переходить от неупорядоченного состояния в весьма упорядоченное. Причину этого легко уяснить, рассмотрев столкновения частиц второй системы со сферами. Хотя эта система кажется распределённой полностью беспорядочно, все столкновения со сферами осуществляются в тех частях полусфер, которые обращены к отрицательному направлению оси x и ясно, что столкновения не происходят беспорядочным образом. Мы видим, таким образом, что для этой второй системы гипотеза о числе столкновений определённо не может осуществляться.

*) Я весьма признателен д-ру Х. М. Джеймсу, указавшему на этот пример.

1.3. Статистические аспекты H -теоремы

Уже упоминалось, что природа объектов, изучаемых статистической механикой, такова, что мы с необходимостью приходим к статистическому рассмотрению. Первым, кто использовал чисто статистический подход, был Больцман^{77В} и именно тогда кинетическая теория превратилась в статистическую механику, хотя прошло ещё двадцать с лишним лет, прежде чем Гиббс окончательно сформулировал её основы.

В своём статистическом рассмотрении Больцман прежде всего пытался дать ответ на вопрос (A1), в особенности для случая изолированной системы, и предложил для этого два возможных определения равновесного состояния:

(a) Равновесное состояние есть наиболее вероятное состояние при заданной энергии.

(b) Равновесное состояние есть среднее состояние, в котором система оказывается за бесконечно большой промежуток времени.

Рассмотрим в этом разделе сперва приближение к наиболее вероятному состоянию. Для доказательства того, что наиболее вероятное состояние соответствует равновесному состоянию, характеризующему (в случае систем независимых частиц, какие рассматривал Больцман) распределением Максвелла—Больцмана, мы поступим следующим образом.

Возьмём изолированную систему из N независимых частиц. Каждой частице соответствует изображающая точка $Q^{(k)}$ в μ -пространстве, где k нумерует частицы системы ($k = 1, 2, \dots, N$). Фаза всей системы, а стало быть, и изображающая точка в Γ -пространстве, определяется тогда N точками $Q^{(1)}, \dots, Q^{(N)}$. Разделим теперь μ -пространство на очень маленькие, но конечные ячейки равного объёма ω , которые пронумеруем последовательно $\omega_1, \omega_2, \dots, \omega_r, \dots$. Это подразделение представляет собой фундаментальный шаг в рассуждении. Величина ω такова, что, с одной стороны, размер ячеек мал по сравнению с мельчайшими макроскопически измеримыми размерами, но, с другой стороны, число изображающих точек $Q^{(k)}$, содержащихся в любой из них, велико^{*)}.

^{*)} Это подразделение на первый взгляд представляется искусственным в классической статистике, в то время как оно оказывается естественным для квантовой статистики. Уленбек^{27U} при обсуждении этого вопроса замечает: «Дело выглядит так, словно Больцман предчувствовал появление дискретных квантовых состояний в μ -пространстве». Однако оно имеет более глубокие основания, и в широком смысле можно сказать, что даже в классической статистике это подразделение естественно, так как оно соответствует ограниченности наших макроскопических измерений. В этой связи можно сравнить обсуждение вопроса о крупноструктурной плотности ансамбля в разделе III.1 и конец настоящего раздела.

Пусть N_i — число изображающих точек в i -й ячейке. Состояние Z^*) при этом полностью описывается заданием N_i .

Описание состояния при помощи чисел N_i вместо задания действительных положений $Q^{(k)}$ соответствует нашим экспериментальным возможностям, поскольку размер ячеек выбран таким, что мы не в состоянии различить на опыте различные точки внутри одной ячейки.

Соотношение между изображающей точкой системы в Γ -пространстве и состоянием Z описывается при помощи N_i следующим образом:

(а) Каждой точке в Γ -пространстве соответствует одно состояние Z .

(б) Для каждого Z имеется объём в Γ -пространстве, такой, что каждая точка в этом объёме соответствует тому же самому Z . Таковую область мы будем называть Z -звездой, а её объём даётся выражением

$$W(Z) = \frac{N!}{\prod_i N_i!} \omega^N. \quad (1.3,1)$$

Выражение (1.3,1) можно получить, если учесть, что под воздействием следующих двух операций совокупность чисел N_i не изменится и, следовательно, Z останется тем же:

(а) Каждая из N точек $Q^{(k)}$ движется по своей ячейке так, что происходит заполнение объёма $\Omega = \omega^N$ в Γ -пространстве.

(б) Любая из перестановок N точек, с условием, что переставляются лишь точки в различных ячейках.

Введём теперь величину $H(Z)$ при помощи выражения

$$H(Z) = -\ln W(Z). \quad (1.3,2)$$

Подставляя выражение (1.3,1) для $W(Z)$ в (1.3,2), используя формулу Стирлинга для факториалов

$$\ln x! = x \ln x - x \quad (1.3,3)$$

и пренебрегая аддитивными постоянными, получим

$$H(Z) = \sum_i N_i \ln N_i. \quad (1.3,4)$$

Сравнивая (1.2,3) и (1.3,4), обнаруживаем, что $H(Z)$ совпадает с больцмановской H при условии, что вместо интегрирования берётся сумма по ячейкам, введённым нами в μ -пространстве.

Легко показать (см. ESM, раздел 1.7), что если определить равновесное распределение N_i^e как такое распределение, для кото-

* Мы используем обозначение Эренфеста^{11E2}: Z от слова Zustandverteilung.

рого $W(Z)$ максимально, то в случае постоянства полной энергии E и полного числа частиц N , т. е.

$$N = \sum N_i, \quad (1.3,5)$$

$$E = \sum N_i E_i, \quad (1.3,6)$$

где E_i — энергия, соответствующая i -й ячейке, мы получим, что N_i^e удовлетворяют соотношению

$$N_i^e = \exp(\mu - \beta E_i), \quad (1.3,7)$$

представляющему собой так называемое распределение Максвелла — Больцмана.

При отождествлении равновесного распределения с распределением, для которого $W(Z)$ максимально, мы определили вероятность состояния, как соответствующий объём в Γ -пространстве. Для доказательства того, что равновесное распределение — это такое распределение, для которого $W(Z)$ максимально или $H(Z)$ минимально, нам необходимо ввести H -теорему (т. е. показать, что $\frac{dH}{dt}$ всегда отрицательно, за исключением случая $H = H_{\min}$) или, что-либо в этом роде. Однако, как только мы начинаем обсуждать поведение $H(Z)$ как функции времени, мы убеждаемся, что любая функция, зависящая от координат и импульсов частиц системы посредством N_i , будет разрывной функцией, ибо как только одна из изображающих точек в μ -пространстве переходит из одной ячейки в другую, два соответствующих N_i меняются на единицу. Для временной зависимости $H(Z)$ мы имеем, таким образом, ступенчатую функцию*), а её производная по времени будет иметь лишь три возможных значения: $-\infty$, 0 и $+\infty$ **).

Из ступенчатой функции мы получаем затем так называемую H -кривую, путём отбора дискретной последовательности точек, разделённых постоянными временными интервалами. Длина временного интервала τ выбирается так, чтобы он был мал по сравнению с промежутками времени, измеряемыми на опыте, но достаточно велик, чтобы за время τ произошло большое число столкновений***).

Теперь мы в состоянии использовать свойства H -кривой для того, чтобы высказать некоторые утверждения****). Будем при этом

*) Необходимо отметить, что этот факт не зависит от характера межмолекулярных сил и является следствием единственно лишь введения конечных ячеек ω .

**) Более строго: производная по времени всюду нуль, за исключением точек разрыва H , где она неопределённа.

***)) Сравни с выбором размеров ячеек в μ -пространстве.

****)) Нумерация утверждений и содержание некоторых из них отличны от имеющих у Эренфеста^{11Е2} или в ESM.

пользоваться простой моделью, введённой в разделе I.1.

Можно показать^{54G, 55G}, что эта модель является представительницей большого класса моделей, отличающихся тем же самым поведением*).

I. Можно найти функцию Δ , которая, с одной стороны, измеряет отклонение от равновесия и, с другой стороны, непосредственно связана с энтропией системы, или H^{**}).

II. Если следить за историей системы во времени, то равновесное распределение, для которого Δ равна своему минимальному значению, будет реализовываться гораздо более часто, чем любое другое распределение. Более того, шанс наблюдать другое распределение настолько ничтожен, что равновесное распределение совпадает со средним распределением.

III. Если система находится в состоянии, соответствующем $\Delta > \Delta_{\min}$, то развитие системы будет происходить так, что Δ , наиболее вероятно, будет убывать.

IV. Предыдущее утверждение будет справедливо вне зависимости от того, будем ли мы следовать по Δ -кривой — которая получается из ступенчатой функции $\Delta(t)$ так же, как H -кривая получается из $H(t)$ — от $t = -\infty$ к $t = +\infty$, или же от $t = +\infty$ к $t = -\infty$.

V. Δ -кривая практически всегда будет находиться вблизи $\Delta_{\min}$.

VI. Промежуток времени между повторениями значений Δ , отличных от $\Delta_{\min}$, резко возрастает с ростом Δ .

VII. Если D_n есть среднее из всех значений Δ в моменты $t_0 + n\tau$, начиная от Δ_0 при t_0^{***}), то D -кривая, т. е. последовательность значений $D_1, D_2, \dots, D_n, \dots$, будет монотонно убывать от значения Δ_0 и асимптотически приближаться к $\Delta_{\min}$.

VIII. Подавляющее большинство Δ -кривых будет следовать D -кривой в течение заметного промежутка времени, но практически ни одна из них не будет следовать D -кривой всё время.

Мы будем называть последовательность значений $\Delta'_1, \Delta'_2, \dots, \dots, \Delta'_n, \dots$, которые следуют из Δ_0 после применения неограниченной H -теоремы, основанной на гипотезе о числе столкновений, кривой чисел столкновений****).

*) См. также^{51B2} и^{52H}.

**) Мы не доказали, что, отвлекаясь от аддитивных и мультипликативных констант, H , даваемая выражением (1.2,3), является энтропией системы. За таким доказательством читатель может обратиться к литературе (см., например, ESM, стр. 22 и 43).

***) При рассмотрении утверждений VII и VIII (и частично утверждения III) следует иметь в виду, что H и, следовательно, Δ не определяют состояние полностью. Для заданного значения Δ имеется несколько возможных состояний Z .

****) Эренфест рассматривал H вместо Δ , поэтому он называл эту кривую кривой H -теоремы (см. также^{54G, 55G}).

IX. Кривая чисел столкновений и D -кривая совпадают.

Рассмотрим модель Лоренца. Функция Δ , упомянутая в утверждении I, будет определяться соотношением

$$\Delta = \sum_{v=-m}^{+m} (f_v - f_v^e)^2, \quad (1.3,8)$$

где f_v^e определены при помощи (1.3,11). В (1.3,8) мы ввели подразделение μ -пространства (которое в этом случае сводится к единичной сфере) на конечные ячейки. Каждая ячейка представляет собой элемент телесного угла размером $\delta\omega$, где

$$\delta\omega = \frac{4\pi}{2m+1}. \quad (1.3,9)$$

$2m+1$ выражение для f_v , которые являются функциями времени и в совокупности определяют состояние системы, задаются таким образом, что f_v представляет собой число электронов,двигающихся в направлении, характеризуемом v -м элементом телесного угла. Лишь $2m$ из величин f_v независимы, так как они удовлетворяют соотношению

$$\sum_{v=-m}^{+m} f_v(t) = N. \quad (1.3,10)$$

Равновесные значения f_v^e величин f_v равны между собой и даются выражением

$$f_v^e = \frac{N}{2m+1}. \quad (1.3,11)$$

Так как можно показать, что f_v находятся всегда вблизи f_v^e [ср. (1.3,16)] и что ввиду этого можно считать $f_v - f_v^e \ll f_v^e$, то с точностью до величин второго порядка по $(f_v - f_v^e)/f_v^e$ имеем

$$H = \sum_v f_v \ln f_v = \sum_v f_v^e \ln f_v^e + \frac{\Delta}{2N} = H_{\text{равн}} + \frac{\Delta}{2N}. \quad (1.3,12)$$

Из (1.3,8) и (1.3,12) следует, во-первых, что не будет никакой разницы, используем ли мы H или Δ для описания отклонения от равновесия, и, во-вторых, что Δ является неотрицательной функцией от f_v , обращающейся в нуль лишь в случае равенства всех f_v своим равновесным значениям.

Что касается утверждения II, то мы можем, прежде всего, вычислить относительную частоту появления состояния системы, соответствующего неравновесному значению Δ . При этом для

нормированной вероятности $w(\Delta) d\Delta$ того, что Δ принимает значение из интервала $(\Delta, \Delta + d\Delta)$, находим выражение *)

$$w(\Delta) d\Delta = \left[\frac{2m+1}{2N} \right]^m \left[\frac{\Delta^{m-1}}{(m-1)!} \right] \exp \left\{ -\frac{(2m+1)\Delta}{2N} \right\} d\Delta. \quad (I.3,13)$$

Функция $w(\Delta)$ играет роль функции $W(Z)$, введенной в начале раздела, и мы видим, что $w(\Delta)$ действительно достигает максимального значения при $\Delta=0$, т. е. для равновесного состояния.

Сделаем теперь допущение, что время $t(Z)$ нахождения в состоянии, описываемом Z , пропорционально $W(Z)$ **). В этом случае среднее по времени значение $\bar{G}$ фазовой функции G будет даваться выражением

$$\bar{G} = \frac{\sum_Z G(Z) W(Z)}{\sum_Z W(Z)}. \quad (I.3,14)$$

Только что сделанное нами допущение о пропорциональности $t(Z)$ и $W(Z)$ необходимо для продолжения нашего рассуждения. Это допущение тесно связано также с предположениями, делаемыми в теориях, изучающих исход событий в разного рода лотереях. В последнем случае, поскольку каждое, например, из бросаний кости по предположению происходит независимо от всех имевших место ранее, пропорциональность вероятности конкретного результата числу раз, когда такой результат осуществляется в бесконечной последовательности бросаний, имеет фундаментальное значение и в действительности используется для определения вероятности. Однако, когда мы имеем дело с процессами, которые фактически с необходимостью вытекают из предшествовавшей истории, пропорциональность времени нахождения в некотором состоянии вероятности этого состояния должна быть предположена. Более того, это обстоятельство нельзя использовать для определения вероятности, поскольку рассматриваемые процессы не являются даже марковскими. Трудность состоит в том, что $t(Z)$ связано с длительным поведением системы, а так как орбита полностью определяется, если точно известна одна из её точек в Γ -пространстве, то чем

*) Детальное обсуждение и вывод формул оставшейся части этого раздела см. в Приложении I.

**) Нам нет необходимости подчёркивать то обстоятельство, что это предположение в действительности состоит из двух. Сперва делается предположение о том, что область $W_E(Z)$, вырезаемая из энергетической поверхности, на которой расположена изображающая точка нашей системы, будет пропорциональна $W(Z)$. Затем делается допущение, что $t(Z)$ пропорционально $W_E(Z)$. Можно упомянуть, что Эйнштейн^{33E, 10E} делал такое же (комбинированное) предположение, какое делаем мы.

дольше мы будем двигаться вблизи орбиты, тем большее число данных, необходимых для полного определения орбиты, будет в нашем распоряжении и, следовательно, мы не будем более иметь дело с чисто статистическим случаем. В этой связи можно указать на рассмотрение аналогичного вопроса Толменом (^{38Т}, стр. 148; см. также ^{50В}). Таким образом сделанное нами допущение тесно связано с марковизацией процессов, соответствующих развитию системы (ср. ^{49S}, ^{54G}, ^{55G}, ^{55H2}).

Необходимо обратить внимание также и на тот факт, что пропорциональность $t(Z)$ и $W(Z)$ не доказана эргодической или квазиэргодической теоремой, как можно было бы надеяться, поскольку эта теорема доказывает лишь пропорциональность $t(Z)$ и $W_E(Z)$, но не пропорциональность $W_E(Z)$ и $W(Z)$. Таким образом, это остаётся одним из основных допущений при подходе к H -теореме в той её форме, в которой она обсуждается нами в данный момент. Более того, мы, повидимому, нисколько не приближаемся к доказательству этого положения, несмотря на огромный объём работы, проделанной как с упрощёнными моделями, так и на основе эргодической теоремы.

Продолжая рассмотрение утверждений II—IX, мы ограничимся в основном теми фазовыми функциями, которые зависят от p и q лишь посредством Δ . Для таких функций можно переписать (I.3,14) в виде

$$\bar{G} = \int \varpi(\Delta) G(\Delta) d\Delta, \quad (I.3,15)$$

поскольку $\varpi(\Delta)$ нормирована. Мы можем вычислить, например, среднее значение Δ и дисперсию около этого среднего значения

$$\Delta_{cp} = N, \quad \langle (\Delta - \Delta_{cp})^2 \rangle_{cp} = \frac{N^2}{m^2}. \quad (I.3,16)$$

Для того чтобы рассмотреть изменение Δ со временем, введем временной интервал τ и вычислим вероятность $\varpi(\Delta, \Delta')$ того, что Δ изменяет своё значение с Δ на Δ' за этот интервал времени τ . Найдя эти вероятности переходов, получаем из соотношения

$$\Delta'_{cp} = \int \varpi(\Delta, \Delta') \Delta' d\Delta' \quad (I.3,17)$$

среднее значение Δ' в конце интервала τ , в то время как в начале интервала было значение Δ ; при этом $\varpi(\Delta, \Delta')$ должно быть нормировано по отношению к Δ' . Для средней скорости изменения Δ (а стало быть и H , — ср. уравнение (I.3,12)) имеем

$$\left(\frac{d\Delta}{dt} \right)_{cp} = \frac{\Delta'_{cp} - \Delta}{\tau}. \quad (I.3,18)$$

В специальном случае модели Лоренца для $\varpi(\Delta, \Delta')$ находим

выражение (см. Приложение I)

$$w(\Delta, \Delta') = (16 \pi N A \Delta)^{-1/2} \exp \left\{ - \frac{[\Delta' - \Delta + A \Delta (2m + 1)]^2}{16 N A \Delta} \right\}, \quad (I.3,19)$$

где A даётся выражением

$$A = 2\alpha\tau, \quad (I.3,20)$$

а α определено посредством (I.1,16). Из (I.3,17) — (I.3,19) получаем

$$\Delta'_{cp} = (1 - A) \Delta, \quad (I.3,21)$$

$$\left(\frac{d\Delta}{dt} \right)_{cp} = - \frac{A \Delta}{\tau}. \quad (I.3,22)$$

Утверждение III непосредственно следует из (I.3,22). Здесь имеет место то же самое экспоненциальное приближение к равновесию, какое было найдено в разделе I.1 [уравнение (I.1,17)]. Вскоре мы убедимся в тесной связи, существующей между (I.1,15) и (I.3,22).

Для доказательства утверждения IV необходимо вычислить среднее значение Δ''_{st} функции Δ , которая после промежутка времени τ принимает значение Δ . Эта величина даётся выражением

$$\Delta''_{st} = \frac{\int \Delta'' w(\Delta'') w(\Delta'', \Delta) d\Delta''}{\int w(\Delta'') w(\Delta'', \Delta) d\Delta''}, \quad (I.3,23)$$

где знаменатель не равен единице, так как величина $w(\Delta'') w(\Delta'', \Delta)$ будет нормирована лишь после интегрирования по обоим переменным Δ и Δ'' . Из (I.3,23), (I.3,13) и (I.3,19) находим

$$\Delta''_{st} = (1 - A) \Delta, \quad (I.3,24)$$

что доказывает утверждение IV*).

Утверждение V следует из выражения, найденного для $w(\Delta)$ и предположения, что вероятность найти некоторое значение из последовательности Δ пропорциональна $w(\Delta)$.

Для доказательства утверждения VI используем формулу Чандрасекара^{43С} для $\Theta(\Delta)$ — времени повторения состояния, характеризуемого значением Δ^{**} :

$$\Theta(\Delta) = \frac{\tau [1 - w(\Delta)]}{w(\Delta) - w(\Delta) w(\Delta, \Delta')}, \quad (I.3,25)$$

*) Необходимо здесь отметить, что в более ранней работе^{53Н} мы использовали для доказательства эквивалентности последовательности с $t = -\infty \rightarrow t = +\infty$ и последовательности с $t = +\infty \rightarrow t = -\infty$ соотношение $w(\Delta) w(\Delta, \Delta') = w(\Delta') w(\Delta', \Delta)$, которое имеет место и в настоящей модели.

**) Возражения, выдвигавшиеся Бартлетом^{50В, 53В} против формулы Чандрасекара, не будут иметь места, если рассматривать последовательности, которые непрерывно марковизуются (см.^{50В}).

или

$$\Theta(\Delta) \approx \frac{\tau}{w(\Delta)}, \quad (1.3,26)$$

откуда и следует утверждение VI (см. 1.3,13)).

Для вычисления D_n , т. е. среднего значения величин Δ в момент $t_0 + n\tau$, начиная с Δ_0 при t_0 , необходимо прежде всего вычислить нормированную вероятность $w(\Delta_n; \Delta_0, \Delta_1, \Delta_2, \dots, \Delta_{n-1})$ последовательности значений Δ_0 при t_0 , Δ_1 при $t_0 + \tau$, Δ_2 при $t_0 + 2\tau, \dots, \Delta_{n-1}$ при $t_0 + (n-1)\tau$ и Δ_n при $t_0 + n\tau$. Эта величина даётся выражением (ср.⁵⁵ G) *)

$$w(\Delta_n; \Delta_0, \Delta_1, \dots, \Delta_{n-1}) = w(\Delta_0, \Delta_1) w(\Delta_1, \Delta_2) \dots w(\Delta_{n-1}, \Delta_n). \quad (1.3,27)$$

Для D_n тогда получаем

$$D_n = \int \Delta_n w(\Delta_n; \Delta_0, \Delta_1, \Delta_2, \dots, \Delta_{n-1}) d\Delta_1 d\Delta_2 \dots d\Delta_n, \quad (1.3,28)$$

откуда следует

$$D_n = (1 - A)^n \Delta_0, \quad (1.3,29)$$

и, полагая $t = n\tau$, имеем

$$D_n \approx \Delta_0 e^{-At/\tau} = \Delta_0 e^{-2\alpha t}. \quad (1.3,30)$$

Утверждение VII непосредственно следует из (1.3,29) или (1.3,30).

Можно вычислить также дисперсию σ_n величин Δ_n , для которой имеем

$$\sigma_n = \int (D_n - \Delta_n)^2 w(\Delta_n; \Delta_0, \Delta_1, \dots, \Delta_{n-1}) d\Delta_1 \dots d\Delta_n. \quad (1.3,31)$$

или

$$\sigma_n = 0. \quad (1.3,32)$$

Из (1.3,32) и того факта, что неравновесные значения Δ повторяются с частотой $\Theta(\Delta)^{-1}$, вытекает утверждение VIII.

Рассмотрим, наконец, кривую чисел столкновений. Заметим, что из (1.3,08) следует

$$\frac{d\Delta}{dt} = \sum_v 2(f_v - f_v^e) \frac{df_v}{dt}. \quad (1.3,33)$$

При условии справедливости гипотезы о числе столкновений скорость изменения f_v со временем даётся уравнением

$$\frac{df_v}{dt} = \frac{A}{\tau} (f_v^e - f_v) \quad (1.3,34)$$

— ср. уравнение (1.1,15); доказательство уравнения (1.3,34) приведено в Приложении I.

*) Можно отметить, что при написании (1.3,27) использовано предположение о непрерывной марковизации.

Используя уравнения (I.3,33) и (I.3,34), получаем для кривой чисел столкновений дифференциальное уравнение

$$\frac{d\Delta}{dt} = -\frac{A\Delta}{\tau}. \quad (\text{I.3,35})$$

Из сравнения (I.3,30) и (I.3,35) вытекает доказательство утверждения IX для рассматриваемой модели.

Подведём итоги рассмотрения существующего положения с *H*-теоремой в классической статистике. Отметим прежде всего, что хотя со времени весьма детального рассмотрения этой проблемы Эренфестом^{11Е2} положение улучшилось благодаря изучению упрощённых моделей, всё же предстоит сделать ещё многое. В той степени, в какой это касается проблем приближения к равновесию и флуктуаций около состояния равновесия, изучение упрощённых моделей показало, что положение во многом таково (утверждения II—VI), как этого можно было бы ожидать при условии возможности трактовать процессы как чисто марковские. Обсуждение поведения системы проводилось нами, в основном, с точки зрения флуктуаций, однако мы могли бы сделать то же самое и для возвращения к равновесию. Сигерт^{49S}, например, показал, что, начиная с любого Δ при $t=t_0$, вероятность найти значение Δ при $t=\infty$ всегда даётся величиной $\varphi(\Delta)$. Мы молчаливо полагали, что всегда имеем дело с системой, которая по существу «забыла» своё первоначальное состояние. Это можно делать в случае, когда время релаксации системы достаточно мало. Время релаксации $t_{\text{рел}}$ по существу определяется выражением τ/A , которое даёт величину $t_{\text{рел}} \sim 10^{-8}$ сек в предположении $n=10^{30}$ м⁻³, $c=10^4$ м·сек⁻¹, $\sigma=10^{-20}$ м², $m=10^6$ (ср. ESM, стр. 390). Такая быстрая релаксация совместно с большими интервалами времени между повторениями заметного отклонения от равновесия*) и являются теми причинами, благодаря которым имеется очень мало исключений из второго закона термодинамики, или, выражаясь несколько иначе, благодаря которым флуктуации очень редко играют существенную роль в поведении макроскопических физических систем. По Смолуховскому^{12S} система будет вести себя необратимо, если её начальное состояние характеризуется средним временем повторения, которое велико по сравнению с промежутком времени, доступным экспериментатору. В этой связи можно усмотреть подтверждение предположения Толмена^(38T, стр. 179) о том, что время релаксации будет весьма малым по сравнению с периодом цикла Пуанкаре (раздел I.2), при условии, чтобы размер элементарных ячеек не был бы слишком велик. Поскольку A обратно пропорционально числу ячеек $2m+1$, действи-

*) Для среднего отклонения от f_v^e порядка $10^{-8} f_v^e$ мы получаем для $\Theta(\Delta)$ величину порядка 10^{10^9} лет!

тельно видно, что $t_{\text{рел}}$ возрастает с m , а мы только что убедились, что $t_{\text{рел}}$ принимает приемлемое значение при приемлемом значении m .

Главной оставшейся проблемой является оправдание марковизации процессов [утверждение VII Эренфеста^{11E2} или утверждение IX ESM]. Эта проблема представляется весьма трудной, но при использовании доступных ныне мощных вычислительных машин можно вычислить H -кривую заданной системы для набора начальных условий, используя точные уравнения движения.

В заключение нам хотелось бы указать, что хотя эргодическая или, лучше сказать, квазиэргодическая теорема в некотором отношении решила вопрос об эквивалентности средних по времени и средних по ансамблю, всё же вследствие неполноты и неточности доступных экспериментальных данных мы можем определить состояние лишь с ограниченной точностью. Даже если мы имеем дело с изолированной системой, так что её изображающая точка будет оставаться на энергетической поверхности, мы всё равно не знаем, на какой именно поверхности, и единственным истинным изображающим ансамблем будет такой, который состоит из всех систем, изображающие точки которых в момент t_0 заполняют Z -звезду. Следовательно, H -теорема всё ещё остаётся связанной с той проблемой, которую мы сейчас обсуждаем.

II. КЛАССИЧЕСКАЯ ЭРГОДИЧЕСКАЯ ТЕОРЕМА

II.1. Эргодическая и квазиэргодическая теоремы

Как уже упоминалось во введении, многие авторы предпочитают подходить к эргодической теореме при помощи H -теоремы. В конце раздела I.3 были изложены возражения против такой точки зрения и к этому вопросу мы ещё вернёмся в части III. Однако эргодическая теорема играет столь важную роль в развитии предмета статистической механики, что представляется несомненно полезным обсудить эту проблему и в настоящей части.

В начале раздела I.3 упоминалось, что Больцман дал два различных определения равновесного состояния, а именно как наиболее вероятного состояния или же как среднего состояния. Первая возможность была рассмотрена нами выше; теперь мы обсудим вторую возможность. Больцман^{71B1, 71B2} (см. также^{68B}) и Максвелл^{79M} полагали, что более или менее последовательно можно доказать следующую теорему: среднее поведение системы независимых частиц соответствует распределению Максвелла — Больцмана. Среднее берётся в смысле усреднения по времени за бесконечный период.

Больцман полагал, что существуют так называемые эргодические системы, и в этом видел основание для надежд относительно справедливости указанной теоремы. Эргодическая система —

это такая система, изображающая точка которой в Γ -пространстве будет проходить через каждую точку соответствующей энергетической поверхности. Допустив существование эргодических систем, легко увидеть преимущества рассмотрения средних по времени. Прежде всего необходимо иметь в виду, что в любых физических экспериментах мы всегда измеряем средние по времени. Более того, поскольку времена релаксации обычно малы по сравнению с периодом, за который происходит измерение экспериментально определяемых величин, совершенно не существенно, берём ли мы среднее за конечный интервал времени или же за бесконечный период. Далее, если бы можно было доказать, что реальные физические системы являются эргодическими, то для любой фазовой функции $\varphi(p, q)$ имели бы место следующие равенства:

$$\langle \varphi \rangle_{cp} = \overline{\langle \varphi \rangle}_{cp} = \langle \overline{\varphi} \rangle_{cp} = \overline{\varphi}, \quad (\text{II.1,1})$$

где черта обозначает среднее по времени, а скобки $\langle \rangle_{cp}$ — среднее, взятое по микроканоническому ансамблю. Второй и третий члены в (II.1,1) дают соответственно среднее по времени от среднего по ансамблю и среднее по ансамблю от среднего по времени. Сейчас мы наметим доказательство равенств (II.1,1). Сделаем здесь лишь замечание, что из этих равенств вытекает возможность вычисления вместо средних по времени средних по ансамблю, а поскольку последние получаются гораздо более просто, мы имеем несомненную выгоду.

Среднее по времени $\overline{\varphi}$ от фазовой функции определяется равенством

$$\overline{\varphi} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{t-T}^{t+T} \varphi(p, q) dt, \quad (\text{II.1,2})$$

где p и q — значения координат изображающей точки системы в Γ -пространстве в момент времени t . Для того чтобы убедиться, соответствует или не соответствует среднее поведение какой-нибудь системы распределению Максвелла — Больцмана, выберем последовательность фазовых функций, которые в совокупности определяют распределение, и сравним их средние по времени со значениями, соответствующими распределению Максвелла — Больцмана.

Среднее по ансамблю $\langle \varphi \rangle_{cp}$ берётся по микроканоническому ансамблю. Это среднее дается выражением

$$\langle \varphi \rangle_{cp} = \frac{\int \varphi(p, q) \sigma(p, q) d\mathcal{E}}{\int \sigma(p, q) d\mathcal{E}}, \quad (\text{II.1,3})$$

где оба интеграла распространены на всю энергетическую поверхность $\mathcal{E}$; $\sigma(p, q)$ является поверхностной плотностью, соответствующей

щей микроканоническому ансамблю. Эта поверхностная плотность даётся выражением *)

$$\sigma(p, q) = \left\{ \sum_{k=1}^S \left(\frac{\partial \mathcal{H}}{\partial p_k} \right)^2 + \left(\frac{\partial \mathcal{H}}{\partial q_k} \right)^2 \right\}^{-1/2}, \quad (\text{II.1.4})$$

где S — число степеней свободы системы и, стало быть, число размерности Γ -пространства; $\mathcal{H}$ — гамильтониан системы.

Из теоремы Лиувилля (ESM, стр. 102) следует (см., например, ESM, стр. 355), что единственными стационарными ансамблями на энергетической поверхности будут те, для которых поверхностная плотность вдоль орбиты удовлетворяет (II.1.4).

Займёмся теперь равенствами (II.1.1). Первое равенство сразу следует из того факта, что микроканонический ансамбль является стационарным ансамблем. Второе равенство вытекает из возможности переставлять порядок усреднений. Остаётся последнее равенство, которое справедливо для эргодических систем, поскольку $\bar{\varphi}$ одинаково для всех систем ансамбля. Это видно из следующего. Так как для эргодической системы орбита изображающей точки проходит через каждую из точек $\mathfrak{S}$ и так как из канонических уравнений движения (I.2.5) следует, что в каждой точке направление орбиты однозначно определено, то, следовательно, на $\mathfrak{S}$ имеется лишь одна орбита. Единственное различие для разных систем ансамбля будет состоять в том, что точный момент времени прохождения через точку орбиты будет для них различным **). Поскольку все системы следуют по одной и той же орбите, то из (II.1.2) вытекает, что $\bar{\varphi}$ одинаково для всех систем и, таким образом, мы получаем последнее равенство из (II.1.1).

Для эргодической системы можно даже продвинуться ещё на шаг дальше и доказать, что имеет место следующее равенство ^{71B}:

$$\lim_{T \rightarrow \infty} \frac{dt}{T} = \int \frac{d\mathfrak{S}}{\sigma d\mathfrak{S}}, \quad (\text{II.1.5})$$

где dt — время, в течение которого изображающая точка находилась внутри элемента поверхности $d\mathfrak{S}$ за период T . Из (II.1.5), в принципе, можно вычислить частоту появления неравновесных состояний.

Доказательство теперь сводится к установлению эргодичности механических систем. Однако легко можно видеть, что, повидимому, никакая из систем не является эргодической (^{38T}, стр. 67). Вспомним, что система, для которой имеют место уравнения (I.2.5), обладает $2S$ интегралами движения. Одним из этих интегралов является энергия, другой фиксирует время вдоль орбиты. Выбирая

*) Выражение внутри скобок в (II.1.4) является абсолютным значением «скорости» изображающих точек в Γ -пространстве.

**) Мы исключаем строго периодические орбиты.

точку P на энергетической поверхности, мы фиксируем значения остальных $2S - 2$ интегралов движения. Взяв другие значения для этих констант, но то же самое значение энергии, получаем другую точку Q на той же самой энергетической поверхности, однако между точками P и Q не существует соединяющей их орбиты. Для завершения доказательства необходимо, стало быть, показать, что $2S - 2$ константы движения не являются константами почти всюду на энергетической поверхности, что представляется неправдоподобным.

Употребив выражение «почти всюду», мы вступили в область теории меры*) и именно при помощи этой теории можно доказать несуществование эргодических систем^{13R, 13P} (см. также^{52R}). Суть этого доказательства состоит в том, что точки орбиты образуют последовательность меры нуль на энергетической поверхности, мера которой отлична от нуля**). В этом можно убедиться следующим образом^{52R}. Рассмотрим точку P орбиты изображающей точки и возьмём малую область $\mathfrak{N}$ конечной меры вокруг P . Так как ясно, что орбита не может оставаться внутри $\mathfrak{N}$ неограниченное время, если она должна полностью покрыть всю $\mathfrak{S}$, то единственной частью орбиты внутри $\mathfrak{N}$ будет совокупность отдельных сегментов***), соответствующих последовательности конечных интервалов времени. Поскольку последовательность таких интервалов времени счётна, то сегменты орбиты внутри $\mathfrak{N}$ также могут быть перенумерованы и, стало быть, имеют меру нуль.

Хотя невозможность существования эргодических систем была показана Розенталем и Планишерелем лишь в 1913 г., П. и Т. Эренфест^(11E2 — особенно примечание 89a, — см. также^{91K}, стр. 484) были, повидимому, вполне уверены в неправдоподобности эргодических систем и в том факте, что точки орбиты и все точки энергетической поверхности обладают различными мерами. Для избежания этой трудности было предположено, что механические системы могут быть квазиэргодическими. Изображающая точка квазиэргодической системы****) будет подходить как угодно близко к лю-

*) Можно отметить здесь, что многие математики, занимающиеся статистикой (см., например,^{53D}), считают статистику лишь отделом теории меры. Для обсуждения теории меры отсылаем к недавно опубликованной книге Галмоша^{50H} (см. также^{53D}).

**) Мера точечной последовательности π на $\mathfrak{S}$ определяется следующим образом: пусть $f(P)$ является функцией, равной 1, если P принадлежит π , и равной 0 в противном случае. Интеграл Лебега $\int f dS$, распространённый на всю $\mathfrak{S}$, называется тогда лебеговской мерой π на $\mathfrak{S}$ и обозначается $\mathfrak{M}\pi$.

***). Их необходимо различать, так как для неперiodической орбиты никакая точка на $\mathfrak{S}$ никогда не проходится дважды.

****) Можно отметить здесь, что ввиду невозможности эргодических систем в первоначальном смысле этого слова, современные авторы часто называют квазиэргодические системы эргодическими.

бой точке энергетической поверхности, не проходя, однако, через каждую точку.

Использование представления о квазиэргодических системах не разрешило, однако, проблемы доказательства равенств (II.1,1), хотя Розенталь^{14R2} и дал не вполне строгое доказательство соотношения (II.1,5), основанного на квазиэргодичности механических систем, а также показал эквивалентность средних по времени и по ансамблю. Ферми^{23F1} доказал, что определённый класс систем, так называемые канонические нормальные системы*) являются квазиэргодическими. Кратко его доказательство изложено в Приложении II.

Для дальнейшего продвижения на этом пути необходимо было найти независимое доказательство эквивалентности средних по времени и по ансамблю, что явилось целью более современных исследований, посвящённых эргодической теореме, которые излагаются в следующем разделе. О доказательстве эквивалентности средних по времени и по ансамблю ныне говорят как об эргодической теореме, хотя с исторической точки зрения, повидимому, было бы более правильно использовать термин квазиэргодическая теорема.

II. 2. Новейшие исследования**)

Современная эргодическая теория начинается со статьи Биркхофа^{31B2}, который рассмотрел следующие средние по времени:

$$\bar{f}(P; t_0, T) = \frac{1}{T} \int_{t_0}^{t_0 + T} f(P_t) dt, \quad (\text{II.2,1})$$

где P_t — точка орбиты, проходящей через P в момент t_0 , в момент времени t . Он показал, прежде всего, что практически для всех точек P энергетической поверхности $\mathfrak{E}$ существует предел величины $\bar{f}(P; t_0, T)$ при $T \rightarrow \infty$. Более того, Биркхоф показал, что $\bar{f}(P; t_0, T)$ в пределе $T \rightarrow \infty$ не зависит от t_0 практически для всех точек P . Доказательство этого свойства средних по времени на энергетической поверхности дано в Приложении III. «Практически все» принимается здесь, как это было предложено Биркхофом в 1922 г.^{22B} (см. также^{26S2}), в смысле, что исключительные орбиты образуют на $\mathfrak{E}$ последовательность меры нуль.

Далее, Биркхоф показал, что $\bar{f}(P; \infty)$, т. е. среднее по времени, взятое за бесконечный промежуток, постоянно почти всюду на $\mathfrak{E}$, при условии, что группа преобразований $P \rightarrow P_t$ является

*) Интересно отметить, что Эйнштейн^{02E, 03E}, независимо от Гиббса предложивший теорию ансамблей, ограничивался сам такими системами, см. также^{11K2}).

**) См. также^{20M1, 20M2, 27H2}.

метрически транзитивной, или, другими словами, при условии, что $\mathcal{S}$ является метрически неразложимой *). Постоянство $\bar{f}(P; \infty)$ следует из того факта, что если бы оно не выполнялось, можно было бы найти такое значение F величины $\bar{f}(P; \infty)$, что условия $\bar{f}(P; \infty) < F$ и $\bar{f}(P; \infty) \geq F$ определяли бы на S две последовательности положительной меры, которые были бы обе инвариантны относительно преобразований $P \rightarrow P_t$ **).

Биркхоф продвинулся даже ещё дальше и доказал соотношение, эквивалентное уравнению (II.1,5), а именно, что практически для всех точек P на $\mathcal{S}$ имеем

$$\lim_{T \rightarrow \infty} \frac{dt(P; m)}{T} \frac{\mathfrak{M}m}{\mathfrak{M}\mathcal{S}}, \quad (\text{II.2,2})$$

где $dt(P; m)$ — время, которое изображающая точка орбиты, проходящей через P , проводит в области m в течение интервала T . Подчеркнём ещё раз, что (II.2,2) справедливо лишь в том случае, когда группа преобразований $P \rightarrow P_t$ является метрически транзитивной ***).

Доказательство Биркхофа несколько более сильное, чем то, которое дал Нейман ^{32N1, 32N2}, показавший, что сходимость $\bar{f}(P; t_0, T)$ имеет место в среднем, в то время как Биркхоф показал, что сходимость имеет место практически для всех P .

Проблема, связанная с эргодической теоремой, сведена теперь к проблеме доказательства, что энергетические поверхности, как правило, являются метрически неразложимыми. В 1941 г. Окстоби и Улам ^{41O} доказали это для довольно общего класса поверхностей, являющихся полиэдрами с размерностью, равной трём или больше. Они, таким образом, показали, говоря их собственными словами,

*) Группа преобразований $P \rightarrow P_t$ называется метрически транзитивной, а $\mathcal{S}$ называется метрически неразложимой, если $\mathcal{S}$ нельзя разбить на две части $\mathcal{S}_1$ и $\mathcal{S}_2$, обе положительной меры и обе инвариантные относительно каждого преобразования группы.

**) Мы не хотим здесь вдаваться в обсуждение усложнений, обязанных наличию других интегралов движения, таких, как импульс, угловой момент и т. д. Они приводят, по выражению Розенфельда, к несущественному разложению $\mathcal{S}$, в то время как нас интересует лишь существенная неразложимость $\mathcal{S}$. В нашем изложении мы будем предполагать, что других однородных интегралов движения, за исключением энергии, не существует и, более того, мы будем рассматривать метрическую неразложимость в физическом смысле. Для обсуждения избегаемых нами таким образом усложнений отсылаем к литературе (см., например, ^{52R}). Необходимо также указать на работу Града ^{52G1, 52G2}, подробно рассмотревшего статистическую механику систем, обладающих интегралами, отличными от энергии.

***)) Из (II.2,2) видно соотношение между эргодической теоремой Биркхофа и теоремой повторимости Пуанкаре (ср. ^{41W} стр. 90—91). Уравнение (II.2,2) выражает тот факт, что изображающая точка будет находиться в любой области положительной меры конечный промежуток времени.

«что эргодическая гипотеза в её современной форме метрической транзитивности, по крайней мере, свободна от каких-либо возражений, основанных на топологических соображениях». Согласно Гамову (^{49K1}, стр. 54; разрядка принадлежит Гамову) это внушает уверенность, «что в некотором смысле почти каждое непрерывное преобразование является метрически транзитивным». Гамов, к сожалению, не приводит доводов в пользу своего утверждения. Однако энергетические поверхности, которые рассматривали Окстби и Улам, не соответствуют вполне физическим системам и ввиду этого можно сказать, что хотя эргодическая теорема Биркхофа весьма приблизила проблему к её разрешению, всё же пока мы имеем дело с гипотезой. Биркхоф и Купман ^{32B} говорят: «Квази-эргодическая гипотеза заменена её современной версией: гипотезой метрической транзитивности». Однако было, повидимому, упущено из виду, что метрическая неразложимость энергетической поверхности доказана для определённого класса физических систем, некоторые из которых существуют. Для того чтобы это стало очевидным, заметим, во-первых, что для любой квазиэргодической системы энергетическая поверхность является метрически неразложимой; последнее оправдывает предположение Эренфеста о том, что для квазиэргодических систем имеет место соотношение (II.1,5). Действительно, квазиэргодичность предполагает, что каждая орбита на энергетической поверхности пройдёт через любую выбранную область $\mathcal{M}$ положительной меры. Поскольку орбита образует совокупность преобразований $P \rightarrow P_t$, где t изменяется от $-\infty$ до $+\infty$, то из этого следует, что условие метрической неразложимости удовлетворяется для энергетической поверхности любой квазиэргодической системы. Во-вторых, примем во внимание, что Ферми ^{23F1} доказал квазиэргодичность канонических нормальных систем и что Пуанкаре ^{92P} показал, что в упрощённой проблеме трёх тел мы имеем дело именно с такой системой*). Разумеется в этом случае число степеней свободы очень мало и представляло бы определённый интерес показать, что и другие физические системы также принадлежат тому же классу.

Канонические нормальные системы характеризуются следующими свойствами. После исключения однородных интегралов движения, таких как компоненты полного импульса, возможно ввести последовательность канонически сопряжённых переменных x_i и y_i , таких, что: 1) энергия E не зависит от времени, 2) в системе существует параметр α , такой, что гамильтониан $\mathcal{H}$ можно представить в виде ряда по степеням α

$$\mathcal{H} = \mathcal{H}_0 + \alpha \mathcal{H}_1 + \alpha^2 \mathcal{H}_2 + \dots, \quad (\text{II.2,3})$$

*) Параметр α , встречающийся в (II.2,3), в этом случае является отношением m_2/m_1 , где три массы m_1, m_2, m_3 удовлетворяют неравенствам $m_3 \ll m_2 < m_1$.

3) первый член $\mathcal{H}_0$ в (II.2,3) не зависит от x_i , но остальные члены могут зависеть как от x_i , так и от y_i , 4) все $\mathcal{H}_i$ — периодические величины с общим периодом по x_i .

Прежде, чем кратко изложить эргодическую теорему Хопфа, мы хотели бы обратить внимание на следующие два обстоятельства. Во-первых, Эйнштейн и другие авторы^{02E, 03E, 06P, 11K2} явно указывали на то, что они верят в принадлежность всех физических систем к классу канонических нормальных систем. Во-вторых, мы хотим указать, что в случае систем, удовлетворяющих условию метрической неразложимости своих энергетических поверхностей, различные интегралы движения являются почти*) постоянными на всей энергетической поверхности и что, таким образом, предположение Больцмана и Максвелла об эргодичности, которое, как мы видели, требует точного постоянства интегралов движения почти всюду на энергетической поверхности, не настолько устранено из современного рассмотрения, как это можно было бы думать.

Эргодическая теорема Биркхофа ограничивается рассмотрением орбит, лежащих на одной энергетической поверхности. Однако, как мы это подчёркивали выше, ближе к физической действительности было бы такое рассмотрение, которое принимает во внимание совокупность энергетических поверхностей, соответствующих энергиям внутри конечного интервала, например:

$$E_0 - \delta < E < E_0 + \delta \quad (\text{II.2,4})$$

(ср. ESM, стр. 99), где $\delta \ll E_0$, чтобы можно было бы всё же говорить о системе с более или менее определённой энергией. Для такого энергетического слоя Хопф^{37H} (см. также^{30H2, 32H1, 32H2, 32H3, 32H4, 34H, 52R}**) дал слегка видоизменённую эргодическую теорему, которая опять доказывает равенство средних по времени и по ансамблю. Необходимым условием для справедливости эргодической теоремы Хопфа является то, что не только каждая энергетическая поверхность в слое должна быть метрически неразложимой, но должна быть метрически неразложимой также и почти каждая энергетическая поверхность любого производного пространства, получаемого путём введения фазового пространства, координаты которого состоят из совокупности пар координат Г-пространства***). Если это условие выполнено, любое распределение в энергетическом слое станет, в конце концов, более или менее однородным (см. также часть III). Эргодическая теорема Хопфа не вносит чего-либо существенно нового. Наиболее важное, не говоря о её

*) Нам нежелательно делать это утверждение с большей строгостью, поскольку это привело бы к длинным рассуждениям.

**) Эти ссылки содержат обширную библиографию, относящуюся к математическим аспектам квазиэргодической теоремы.

***). Фазовое пространство, так сказать, удваивается. Более подробно этот вопрос изложен в работах Хопфа (см. также^{52R}).

ценности с математической точки зрения, заключается в том факте, что квантово-механический аналог эргодической теоремы более тесно связан с теоремой Хопфа, чем с теоремой Биркхофа, что будет видно в части V.

III. *H*-ТЕОРЕМА В КЛАССИЧЕСКОЙ ТЕОРИИ АНСАМБЛЕЙ

III.1. Мелкоструктурная и крупноструктурная плотности

До сих пор мы рассматривали лишь поведение изолированных систем. Как уже упоминалось несколько раз, такой подход, по нашему мнению, не отдаёт должного чисто статистическому рассмотрению. Он является подходом, характерным, скорее, для кинетической теории газов и ввиду этого неудивительно, что как *H*-теорема в её первоначальной форме — её мы обсуждали в части I, — так и классическая эргодическая теорема (часть II) ведут своё начало от работ Больцмана, вершиной которых была его «Теория газов»^{96B, 98B}. Когда Гиббс развил свою статистическую механику и ввёл теорию ансамблей, в рассмотрение был внесён новый элемент. Гиббс пытался показать (^{02G}, особенно глава XII; см. также ^{07L}), что ансамбль систем будет развиваться таким образом, что он станет приближаться к микро- или макроканоическому ансамблю*). Для этого необходимо ввести несколько видоизменённую *H*-теорему, что и будет обсуждено в настоящем разделе. Наше рассмотрение будет основываться на так называемых больших (*grand ansemlе*) ансамблях, частично из-за того, что, по нашему мнению, рассмотрение таких ансамблей является логическим следствием статистического подхода. Впервые такие ансамбли были введены Гиббсом в последней главе его монографии (^{02G}, гл. XV). Признав важность понятия изображающих ансамблей, мы автоматически приходим к большим ансамблям, как это будет видно из следующего раздела (см. также ESM, стр. 235 и ^{55H1}).

Для упрощения рассмотрения мы будем иметь в виду лишь системы, состоящие из частиц одного сорта. Обозначим через ν число частиц в системе. Система, состоящая из ν частиц, каждая из которых обладает s степенями свободы, может быть описана νs обобщёнными координатами (q) и νs обобщёнными импульсами (p). Её фазовое пространство, или Γ -пространство, имеет, следовательно, размерность, равную $2\nu s$. При рассмотрении большого ансамбля мы имеем дело с системой переменного числа степеней свободы. Пусть $D(\nu; p, q) d\Omega$, обозначает число систем в ансамбле с ν частицами и с изображающими точками, лежащими внутри объёма $d\Omega$, в

*) Мы часто используем термин макроканоический для ансамблей, которые Гиббс называл каноическими. Причины такого отклонения от работ классиков приведены в другом месте^{54H2}.

соответствующем фазовом пространстве (изображающей точкой называется точка фазового пространства системы с координатами, равными значениям обобщённых координат и импульсов частиц системы). Мы имеем здесь

$$d\Omega_v = \prod_{k=1}^{sv} dp_k dq_k. \quad (\text{III.1,1})$$

Зависимость D от p_k и q_k обозначена символами p и q . Пусть $N_{\text{анс}}$ — число систем ансамбля, так что

$$N_{\text{анс}} = \sum_v \int D(v; p, q) d\Omega_v, \quad (\text{III.1,2})$$

где интегрирование распространено каждый раз на всё фазовое пространство. Введём теперь плотность ансамбля $\rho(v; p, q)$ *):

$$\rho = \frac{D}{N_{\text{анс}}}; \quad (\text{III.1,3})$$

из (III.1,2) и (III.1,3) следует, что ρ удовлетворяет условию нормировки

$$\sum_v \int \rho(v; p, q) d\Omega_v = 1. \quad (\text{III.1,4})$$

Показатель вероятности η определяется соотношением

$$\eta = \ln \rho. \quad (\text{III.1,5})$$

Среднее значение $\langle G \rangle$ любой фазовой функции $G(v; p, q)$ даётся выражением

$$\langle G \rangle = \sum_v \int \rho G d\Omega_v. \quad (\text{III.1,6})$$

До сих пор мы не принимали ещё во внимание усложнение, вносимое тем фактом, что в системе, содержащей лишь частицы одного сорта, мы имеем дело с v тождественными величинами, ввиду чего необходимо решить, считать ли различными состояния, отличающиеся лишь перестановкой некоторых частиц, или же считать их одинаковыми. Если считать их различными, то мы имеем дело, так сказать, с видовыми фазами (specific phases); если же считать их одинаковыми, то мы имеем дело с родовыми фазами (generic phases). Легко видеть, что каждая родовая фаза содержит $v!$ видовых фаз. Хотя даже ещё до появления квантовой механики имелись веские основания предпочесть родовые фазы видовым фазам (см., например, ⁰²⁰ или ESM, стр. 142), решающий аргумент в пользу родовых фаз состоит в том, что лишь они приводят к формулам,

*) Гиббс называл D плотностью (фазовой), а ρ — коэффициентом вероятности.

являющимися предельными случаями квантово-механических формул. Ниже мы будем иметь дело лишь с родовыми фазами и, более того, примем, что интегрирования, подобные имеющимся в (III.1,2), (III.1,4) или (III.1,6), распространены лишь на все различные родовые фазы, т. е. из всех $\nu!$ различных видовых фаз будет приниматься во внимание лишь одна.

Рассмотрим теперь свойства величины σ , определяемой соотношением

$$\sigma = \sum_{\nu} \int \rho \ln \rho d\Omega_{\nu}. \quad (\text{III. 1,7})$$

Прежде всего видно, что σ является средним значением показателя вероятности η . Мы покажем далее, что σ имеет следующие свойства:

а) Если наш ансамбль таков, что все системы состоят из одинакового числа частиц N и энергии ϵ всех систем лежат в интервале $E \leq \epsilon \leq E + \delta E$, то σ будет минимальна тогда, когда ρ задается равенствами:

$$\left. \begin{aligned} \rho &= \text{const} \cdot \delta(\nu - N) \text{ при } E \leq \epsilon \leq E + \delta E; \\ \rho &= 0 \quad \quad \quad \text{при } E > \epsilon \text{ или } E + \delta E < \epsilon, \end{aligned} \right\} \quad (\text{III.1,8})$$

где $\delta(x)$ — дираковская δ -функция (^{35D}, разделы 20, 21).

б) Если наш ансамбль таков, что все системы состоят из одинакового числа частиц, но задана лишь средняя энергия, т. е. ρ подчиняется условию *)

$$\sum_{\nu} \int \epsilon \rho d\Omega_{\nu} = E, \quad (\text{III.1,9})$$

то σ будет минимальна тогда, когда ρ удовлетворяет соотношению

$$\rho = \delta(\nu - N) e^{\beta(\psi - \epsilon)}, \quad (\text{III.1,10})$$

где β и ψ — постоянные, которые можно найти из (III.1,4) и (III.1,9).

в) Если заданы лишь среднее число частиц и средняя энергия систем ансамбля, т. е. ρ подчинено условию (III.1,9) и условию

$$\sum_{\nu} \int \rho d\Omega_{\nu} = N, \quad (\text{III.1,11})$$

то σ будет минимальна тогда, когда ρ удовлетворяет соотношению

$$\rho = \exp[-q - \nu\mu - \beta\epsilon]. \quad (\text{III.1,12})$$

Опять β , μ и q — постоянные, т. е. они не зависят от p_k , q_k или ν и их можно найти из (III.1,4), (III.1,9) и (III.1,11).

*) Суммирование по ν проводится в этом случае тривиально, поскольку даёт вклад лишь один член из суммы.

Прежде чем доказывать эти свойства σ , напомним ещё раз, что значения плотности, приведённые в этих трёх случаях, относятся соответственно к ансамблю с энергетическим слоем*), макроканоническому ансамблю и каноническому большому ансамблю. Из обычного рассмотрения вытекает (см., например, ESM, разделы 5.3 и 6.1), что величина $1/\beta$, совпадающая с модулем ансамбля, равна kT (k — постоянная Больцмана, T — абсолютная температура), что ϕ является свободной энергией системы, изображаемой макроканоническим ансамблем, что μ равна умноженным на β парциальной свободной энергии или парциальному тепловому потенциалу, что q совпадает с q -потенциалом Крамерса, который для однородной системы равен βpV (p — давление, V — объём), и что в обоих случаях (а) и (в) $\langle \eta \rangle$ равно $-S/k$ (S — энтропия)**).

Доказательство свойств минимальности σ основывается главным образом на том, что функция y , задаваемая выражением

$$y = xe^x - e^x + 1, \quad (\text{III.1,13})$$

положительна при $x \neq 0$ и равна нулю при $x = 0$. Это свойство y следует наиболее просто из того, что 1) $y(0) = 0$ и 2) $\frac{dy}{dx} (= xe^x)$ имеет тот же знак, что и x .

Сравним теперь для трёх случаев (а), (б) и (в) две плотности ρ_1 и ρ_2 , где плотность ρ_1 в каждом из случаев соответствует минимальному значению σ , т. е. ρ_1 даётся соответственно соотношениями (III.1,8), (III.1,10) и (III.1,12), в то время как ρ_2 задаётся выражением

$$\rho_2 = \rho_1 e^{\Delta\eta}, \quad (\text{III. 1,14})$$

где $\Delta\eta$ может быть произвольной функцией p_k и q_k , а в случае (в) также и ν . Обе величины ρ_1 и ρ_2 удовлетворяют соотношению (III.1,4)

$$\sum \int \rho_1 d\Omega = \sum \int \rho_2 d\Omega = 1, \quad (\text{III.1,15})$$

а в случаях (б) и (в) они обе удовлетворяют (III.1,9):

$$\sum \int \varepsilon \rho_1 d\Omega = \sum \int \varepsilon \rho_2 d\Omega = E. \quad (\text{III.1,16})$$

В случае (в) они удовлетворяют (III.1,11)

$$\sum_{\nu} \int \rho_1 d\Omega = \sum_{\nu} \int \rho_2 d\Omega = N. \quad (\text{III.1,17})$$

*) Такой ансамбль иногда называют микроканоническим, хотя микроканонический ансамбль является в действительности лишь предельным случаем ансамбля типа энергетического слоя.

**) Это является ещё одним примером преимущества родовых плотностей по сравнению с видовыми плотностями, поскольку это соотношение между S и ρ не имеет места для видовой плотности.

Наконец, имеем

$$\sigma_2 - \sigma_1 = \sum \int (\rho_2 \ln \rho_2 - \rho_1 \ln \rho_1) d\Omega. \quad (\text{III.1,18})$$

Из (III.1,18) для всех трёх случаев можно получить соотношение*)

$$\sigma_2 - \sigma_1 = \sum \int \rho_1 [\Delta \eta e^{\Delta \eta} - e^{\Delta \eta} + 1] d\Omega \geq 0, \quad (\text{III.1,19})$$

последнее неравенство является следствием функции, определенной в (III.1,13).

Итак, мы доказали, что σ минимальна в случае энергетического слоя, макроканонического, или большого канонического ансамбля при выполнении некоторых условий (в следующем разделе будет обсуждено физическое значение этих условий). Сравнивая выражение (III.1,7) для σ с выражением (1.2,3) для H , можно подумать, что мы в состоянии доказать, что $\frac{d\sigma}{dt}$ отрицательна и таким путем установить тенденцию развития канонических ансамблей. Однако легко показать, что $\frac{d\sigma}{dt} = 0$. Это было выяснено Бебером^{ОЗВ} вскоре после опубликования монографии Гиббса. Для того чтобы убедиться в этом, напомним

$$\begin{aligned} \sigma(t'') &= \sum \int \rho'' \ln \rho'' d\Omega'' = \sum \int \rho' \ln \rho' J d\Omega' = \\ &= \sum \int \rho' \ln \rho' d\Omega' = \sigma(t'); \end{aligned} \quad (\text{III.1,20})$$

здесь ρ'' , ρ' — сокращённые обозначения величин $\rho(p'', q''; t'')$, $\rho(p', q'; t')$; связь между p'' , q'' и p' , q' состоит в том, что изображающая точка p' , q' в момент времени t' с течением времени (в момент t'')**) перейдет в точку p'' , q'' ; наконец, J — якобиан преобразования от p'' , q'' к p' , q' , который равен единице, в силу теоремы Лиувилля (см., например, ESM, стр. 102).

*) Для получения соотношения (III.1,19) в случае (а) прибавим к правой части (III.1,18) выражение $\sum \int (1 - \ln \rho_1) (\rho_1 - \rho_2) d\Omega$, которое равно нулю в силу (III.1,15) и того, что $\rho_1 = \text{const}$. В случае (б) прибавим к правой части (III.1,18) выражение $\sum \int \{\beta(\psi - \epsilon) + 1\} (\rho_1 - \rho_2) d\Omega$, которое равно нулю в силу (III.1,15) и (III.1,16). Наконец, в случае (в) прибавим выражение $\sum \int \{q - \gamma\epsilon + \beta\epsilon - 1\} (\rho_2 - \rho_1) d\Omega$, равное нулю в силу (III.1,15) — (III.1,17).

**) Поскольку мы имеем дело с системами, состоящими из одного компонента, изображающая точка системы будет оставаться в одном и том же Γ -пространстве и изменения в γ происходить не будет. Положение, однако, намного усложняется, если рассмотреть системы, в которых могут происходить химические реакции. В этом случае при рассмотрении можно использовать метод вторичного квантования (Шенберг 52S1, 53S1, 53S2).

Хотя мы убедились, что σ является постоянной, все же в некотором смысле мы имеем приближение к стационарному случаю. Поясним это на примере, указанном Гиббсом^{02G} (см. также^{03B, 04B1, 04B2, 06E2, 11E2}). Рассмотрим некоторый сосуд с жидкостью, например водой, к которой добавлена какая-либо нерастворимая краска, состоящая, например, из коллоидальных частиц. Из опыта хорошо известно, что если в первоначальном состоянии красящее вещество было распределено неоднородно по объёму, то практически любое перемешивание приведёт к состоянию, в котором краска распределена однородно, насколько это можно определить визуально. Это значит, что перемешивание приведёт к «равновесному» состоянию. Однако, если присмотреться внимательнее, то мы попрежнему обнаружим, что в микроскопических объемах часть пространства занята водой и часть занята коллоидальными частицами. Таким образом, хотя грубое распределение является однородным, мелкоструктурное распределение оказывается попрежнему неоднородным.

Из этого примера следует, что в ряде случаев может быть полезным наряду с мелкоструктурной плотностью ρ (fine-grained density) ввести крупноструктурную плотность P (coarse-grained density) определяемую следующим образом. Для каждого ν разделим соответствующее Γ -пространство на конечные, но малые ячейки *) $\Omega_i^{(\nu)}$ с объёмом $W(\Omega_i^{(\nu)})$ и пусть $P_i^{(\nu)}$ будет средним значением ρ в $\Omega_i^{(\nu)}$,

$$P_i^{(\nu)} = \frac{\int_{\Omega_i^{(\nu)}} \rho d\Omega}{W(\Omega_i^{(\nu)})}, \quad (\text{III.1,21})$$

где интегрирование распространено на ячейку $\Omega_i^{(\nu)}$ **).

Введём теперь крупноструктурную плотность $P(\nu; p, q)$, положив её постоянной в каждой ячейке $\Omega_i^{(\nu)}$ и равной $P_i^{(\nu)}$. Из (III.1,4) получаем, как легко убедиться,

$$\sum_{\nu} \sum_i P_i^{(\nu)} W(\Omega_i^{(\nu)}) = 1, \quad (\text{III.1,22})$$

или

$$\sum \int P d\Omega = 1; \quad (\text{III.1,23})$$

мы видим, что P нормирована.

*) Ср. введение ячеек в разделе I.3.

**) Мы можем несколько обобщить определение крупноструктурной плотности путем подразделения также и возможных значений ν на интервалы, содержащие некоторое число целых чисел каждый. Однако для того, чтобы не усложнять формулы, мы этого делать не будем, поскольку в случае систем, содержащих лишь один компонент, это обобщение не необходимо.

Введём, далее, вместо σ функцию Σ ,

$$\Sigma = \sum_v \sum_i W(\Omega_i^{(v)}) P_i^{(v)} \ln P_i^{(v)}, \quad (\text{III.1,24})$$

или, выражая через крупноструктурную плотность P ,

$$\Sigma = \sum \int P \ln P d\Omega. \quad (\text{III.1,25})$$

Если учесть, что $\ln P$ постоянен внутри любой ячейки $\Omega_i^{(v)}$ и что интегрирование P по $\Omega_i^{(v)}$ совпадает с интегрированием ρ по $\Omega_i^{(v)}$, получим

$$\Sigma = \langle \ln P \rangle. \quad (\text{III.1,26})$$

Свойства σ , указанные в случаях (а), (б) и (в), справедливы также и для Σ , если заменить всюду ρ на P . Однако теперь Σ не является больше константой; действительно, можно показать, что Σ убывает. Таким образом, можно найти тенденцию к установлению канонических ансамблей — во всяком случае постольку, поскольку это касается крупноструктурной плотности. Здесь можно отметить, что хотя в монографии Гиббса^{02G} введено различие между мелко-структурной и крупноструктурной плотностями, в самом изложении эти две плотности различаются не вполне четко. То же самое в большой степени относится к изложению Бербери^{03B},^{04B2} и Бумстеда^{04B1}; отметим также работы Эренфеста^{06E2},^{11E2}, Пуанкаре^{06P}, Лоренца^{07L} и Кроо^{11K2}.

Рассмотрим теперь изменение Σ со временем. Для этого необходимо использовать некоторые из соображений, приводимых в следующем разделе, где обсуждаются изображающие ансамбли. Предположим, что мы сделали некоторые наблюдения над физической системой в момент t' . Поскольку такие наблюдения никогда не дают максимально возможных сведений, мы построим ансамбль, средние свойства которого в момент t' соответствуют наблюдаемым свойствам рассматриваемой системы в тот же момент времени. Вследствие экспериментальных ограничений можно, в лучшем случае, определить $\rho(v; p, q)$, которая меняется от одной ячейки к другой, если только размеры ячеек выбраны в согласии с экспериментальными ограничениями (это будет предполагаться). Положим, таким образом, что мелкоструктурная плотность в каждой ячейке постоянна и при t' имеем:

$$P' = \rho' \quad (\text{III.1,27})$$

и для Σ

$$\Sigma' = \sum \int P' \ln P' d\Omega = \sum \int \rho' \ln \rho' d\Omega. \quad (\text{III.1,28})$$

Если состояние в момент t' уже соответствует равновесию, то Σ примет минимальное значение и каких-либо изменений в дальнейшем

ожидать будет нельзя, поскольку все три ансамбля, соответствующие минимуму Σ в случаях (а), (б) и (в), будут стационарными (см., например, ESM, стр. 105 и 137). Предположим поэтому, что ρ' не соответствует стационарному ансамблю. Через некоторое время*), в момент t'' , равенство (III.1,27) не будет более иметь место и мы получим

$$\rho'' \neq P''. \quad (\text{III.1,29})$$

Последнее вытекает из того, что хотя ρ остаётся постоянной при деформациях в фазовом пространстве ячейки, сохраняющей неизменный объём $W(\Omega_i^{(v)})$, характер этих деформаций будет изменяться так, что в последующие моменты времени каждая из ячеек будет заполнена точками, которые в момент t' принадлежали множеству различных ячеек.

Запишем соотношение (III.1,29) в виде

$$\rho'' = P'' e^{\Delta}, \quad (\text{III.1,30})$$

где Δ является функцией от v , ρ и q . Для Σ будем иметь

$$\Sigma'' = \sum \int P'' \ln P'' d\Omega; \quad (\text{III.1,31})$$

для изменения Σ получим**)

$$\Sigma' - \Sigma'' = \sum \int P'' [\Delta e^{\Delta} - e^{\Delta} + 1] d\Omega > 0. \quad (\text{III.1,32})$$

Отсюда видно, что Σ'' меньше, чем Σ' вследствие того, что ρ'' и P'' не равны теперь повсюду. Сравнивая полученный результат со случаем красящего вещества в жидкости, можно ожидать, что с течением времени ρ и P будут различаться всё более и более и что Σ будет убывать до тех пор, пока не установится равновесное распределение P , т. е.

случай (а)

$$\left. \begin{aligned} P &= \delta(v - N) \cdot \text{const}, & E \leq \varepsilon \leq E + \delta E; \\ P &= 0, & \varepsilon < E \text{ или } E + \delta E < \varepsilon; \end{aligned} \right\} \quad (\text{III.1,33})$$

*) Строго говоря, необходимо сказать «в другое время», не предпреляя вопроса о том, расположено ли t'' позднее или ранее, чем t' . Однако, поскольку мы построили наш ансамбль для изображения системы, наблюдаемой при t' , нас интересуют лишь предсказания, которые могут быть проверены на эксперименте, т. е. при более поздних временах t'' . В связи с этим интересно прочесть замечания Бербери и Эрэнфеста по этому вопросу ^{03B, 04B2, 06E2} (см. также ^{02G, 04B1}).

**) Соотношение (III.1,32) получено путём вычитания выражения (III.1,31) из (III.1,28), замены в подинтегральном выражении (III.1,28) ρ' на ρ'' , — что можно сделать, так как $\frac{d\rho}{dt} = 0$ [ср. (III.1,20)] — и добавления выражения $\sum \int (P'' - \rho'') d\Omega$, равного нулю в силу (III.1,4) и (III.1,23).

$$\text{случай (б)} \quad P = \delta(v - N) e^{\beta(\psi - \varepsilon)}, \quad (\text{III.1,34})$$

$$\text{случай (в)} \quad P = \exp[-q + \nu\mu - \beta\varepsilon]. \quad (\text{III.1,35})$$

Как только установятся распределения (III.1,33) — (III.1,35), Σ достигнет своего минимального значения и в дальнейшем, следовательно, уже не будет убывать. В этом случае, произведя наблюдение, мы придём к выводу, что состояние рассматриваемой системы лучше всего могло бы быть описано при помощи микроканонического или канонического большого ансамблей — которые являются стационарными. Отсюда можно извлечь дополнительное оправдание использования этих ансамблей для описания систем в термодинамическом равновесии *).

Время, необходимое для установления равновесного распределения, вообще говоря, порядка времени релаксации, а не порядка времён повторимости цикла Пуанкаре (раздел I.2), как предложили считать П. и Т. Эренфест^{11Е2}. Повидимому, они забыли принять во внимание неразличимость частиц, составляющих систему. Это значит, что период Пуанкаре следует разделить на $N!$, после чего получаются интервалы порядка времени релаксации.

В заключение этого раздела необходимо отметить еще два обстоятельства. Во-первых, мы доказали, что если в момент t' мы построили изображающий ансамбль, то величина Σ будет убывать от своего значения в момент t' к меньшему значению в момент t'' ; однако мы не доказали, а лишь пояснили на примере красящего вещества в воде, что Σ будет продолжать убывать. Положение здесь весьма напоминает то, с которым мы имели дело в разделе I.3, где обсуждались статистические аспекты H -теоремы. Заслуживало бы внимания более подробное исследование поведения Σ , как функции времени, аналогично тому, как это делалось в разделе I.3. Интересно отметить, что сам Гиббс^{02G} не утверждал, что Σ' будет монотонно убывать, а утверждал лишь, что

$$\lim_{t \rightarrow \infty} \Sigma(t) \leq \Sigma(t'); \quad (\text{III.1,36})$$

последнее соотношение сам он не доказал, но оно для специального случая было доказано Пуанкаре^{05P} и Кроо^{11K2}. Можно указать в связи с этим на Толмена^(38T), особенно разделы 49 и 51), а также на материал, имеющийся в следующем разделе.

Второе обстоятельство, на которое мы хотели бы обратить внимание, заключается в следующем: убывание Σ следует из того, что мы переходим от состояния, в котором ρ и P равны, к состоянию, в котором ρ не равно больше всюду P . Так как P получено путем экспериментальных наблюдений, можно сказать, что

*) Оправдания, которые обычно приводятся, можно найти, например, в монографии Гиббса^{02G} (особенно главы IV, X и XV) или ESM, разделы 5.3, 5.7 и 6.1 (ср. также ESM, разделы 5.6 и 6.3).

в момент t' мы «знаем» p . Однако в момент t'' мы уже не обладаем таким объёмом конкретных сведений. Поскольку Σ можно считать «крупноструктурной» σ и поскольку σ связано с энтропией соотношением

$$\sigma = \langle \eta \rangle = - \frac{S}{k}, \quad (\text{III.1,37})$$

то отсюда еще раз видно, каким образом увеличение степени отсутствия сведений соответствует возрастанию энтропии *).

III.2. Изображающие ансамбли

В предыдущем разделе уже рассматривалось поведение ансамблей и, таким образом, мы приступили к обсуждению поставленного во Введении вопроса (В), т. е. вопроса: почему возможно описание поведения системы путём рассмотрения среднего поведения в ансамбле? Это, повидимому, является наиболее важным вопросом в статистической механике. Однако по довольно неубедительным причинам его не рассматривали достаточно подробно, как он заслуживает. Частично это произошло из-за того, что сам Гиббс ввёл представление об ансамблях скорее для использования их при статистических рассмотрениях, чем для иллюстрации поведения физических систем, хотя сам он и не придерживался строго своих первоначальных намерений, как они изложены во введении к его монографии. В действительности имеются некоторые указания на то, что Гиббс имел в виду возможность, что среднее поведение системы в микро- или макроканоническом ансамбле может изображать действительное поведение физической системы в равновесии (см., например, замечание Бумстеда^{04B1}). В том же году, что и Гиббс, и независимо от него, Эйнштейн^{02E, 03E} предложил теорию ансамблей; он определённо имел в виду возможность эквивалентности среднего поведения в ансамбле и действительного поведения физической системы. Однако П. и Т. Эренфест^{11E2}, Орнштейн^{08O} и Уленбек^{27U} расценивали теорию ансамблей главным образом как математический приём для облегчения вычисления средних значений фазовых функций. Уленбек, например, определённо ставил теорию ансамблей в один ряд с методом Дарвина — Фаулера (22D1, 22D2, 22D3, 22F2, 23F3, 23D1, 23D2, 25F1, 26F2, 26F3; краткий обзор имеется в ESM, приложение IV, и в^{48S}, глава VI). Лишь после того, как Толмен ввёл понятие об изображающих ансамблях, теория ансамблей получила прочное физическое основание; однако

*) Мы не станем следить более подробно за соответствием между энтропией и отсутствием подробных сведений или объёмом незнания, а отошлём к литературе: см., например, ESM, стр. 160 и статьи Сциларда^{29S} и Бриллюэна^{51B4} (см. также^{48W, 49S1, 54M1, 54G1, 55G1}).

его точка зрения не всегда в должной мере принималась во внимание. В настоящем разделе будет довольно подробно изложен метод Толмена в применении к классической статистической механике; соответствующий материал, относящийся к квантово-механическому случаю, приводится в разделе IV.3.

Выше уже неоднократно упоминалось, что в виду сильной ограниченности наших возможностей мы не можем получить достаточно адекватные сведения, например, о начальном состоянии, изучаемых нами физических систем. Если вспомнить, что один моль газа содержит $6 \cdot 10^{23}$ атомов, то это значит, что вместо знания $36 \cdot 10^{23}$ значений координат и импульсов, необходимых для полного описания состояния (предполагая каждый атом точечной частицей с тремя степенями свободы), мы можем получить лишь несколько соотношений между ними, измеряя, например, давление, импульс и угловой момент системы. Более того, мы совершенно не в состоянии узнать полное число частиц в системе; напомним лишь, что число Авогадро известно с точностью всего в несколько десятисычных ^{51b1, 51b2}. Получив столь скудные сведения, мы хотим делать предсказания о поведении системы в будущем, основываясь на этих сведениях. Ясно, что единственным возможным путём для этого является использование теории вероятности и статистических методов. Если бы мы могли получить значения всех координат и импульсов в один момент времени — и если бы мы могли решить уравнения движения, — поведение системы было бы полностью определено, поскольку тогда мы имели бы дело с задачей классической механики и с достаточным числом граничных условий для определения орбиты изображающей точки системы в фазовом пространстве. Однако ввиду невозможности получить эти значения мы ограничиваемся утверждениями относительно наиболее вероятного поведения системы. Для определения этого наиболее вероятного поведения сравним различные системы, обладающие одинаковыми значениями тех величин, которые измеряются, а в остальном могущих различаться в широких пределах. Это значит, что мы строим ансамбль, и з о б р а ж а ю щ и й нашу систему. Остаётся вопрос, каким образом можно построить такой изображающий ансамбль; этот вопрос связан с проблемой априорных вероятностей. При построении классического ансамбля было предположено равенство априорных вероятностей для равных объёмов в Γ -пространстве*). При рассмотрении так называемых малых ансамблей (*petit anseuble*) это приводит к вполне определённым предписаниям. Однако в случае построения изображающих больших ансамблей необходимо помнить, что мы имеем дело с Γ -пространствами различного числа измерений, и что пространства с большим числом измерений имеют большие априорные

*) Это предположение молчаливо принималось, когда мы выписывали выражение (I.3,1) для вероятности состояния Z .

веса. Теперь мы постараемся оправдать, или, по крайней мере, показать правдоподобность нашего выбора априорных вероятностей.

Для того чтобы выявить тип ансамбля, наилучшим образом изображающего рассматриваемую систему, мы должны, прежде всего, иметь в виду, что практически при любых обстоятельствах мы измеряем лишь свойства малых подсистем. При измерении давления газа лишь часть газа, находящаяся вблизи стенки, принимает участие в образовании силы, действующей на стенку и измеряемой нашими приборами. При определении плотности системы при помощи оптического метода пучок света проходит лишь через часть системы (например, случай критической опалесценции). При измерении электрической проводимости полупроводника электрод прикладывается к части образца, что, в действительности, нередко используется для обнаружения различия между состояниями в различных местах образца и т. д. Из этого вытекает, что мы на самом деле интересуемся лишь предсказаниями относительно поведения подсистем. Для краткости будем предполагать, что полная система (которая может быть либо всей физической системой, заключенной в нашей аппаратуре, либо даже этой системой плюс часть внешнего окружения) является изолированной, так что она не может обмениваться ни энергией, ни частицами с внешним окружением. Пусть ρ будет плотность ансамбля, описывающего какую-нибудь подсистему, и ρ' — плотность ансамбля, описывающего «остальную систему», т. е. полную систему за вычетом подсистемы. Пусть, далее, ε и ν обозначают энергию и число частиц в подсистеме, а ε' и ν' — соответствующие величины в остальной системе. Плотности ρ и ρ' тогда будут удовлетворять соотношениям:

$$\sum \int \rho \, d\Omega = 1, \quad (\text{III.2,1})$$

$$\sum \int \rho' \, d\Omega = 1, \quad (\text{III.2,2})$$

$$\sum \int (\varepsilon + \varepsilon') \rho \rho' \, d\Omega \, d\Omega' = \text{const}, \quad (\text{III.2,3})$$

$$\sum \int (\nu + \nu') \rho \rho' \, d\Omega \, d\Omega' = \text{const}, \quad (\text{III.2,4})$$

где соотношения (III.2,3) и (III.2,4) выражают тот факт, что в изолированной системе полная энергия и полное число частиц постоянны. Допустим, далее, что установилось состояние равновесия, т. е. Σ минимальна для полной системы, или

$$\sum \int (\rho \rho') \ln (\rho \rho') \, d\Omega \, d\Omega' = \text{минимум}. \quad (\text{III.2,5})$$

При написании (III.2,3) и (III.2,5) было использовано то обстоятельство, что ансамбль полной системы представляет собой произве-

дение ансамблей подсистемы и остальной системы. Используя (III.2,1) — (III.2,4) можно переписать (III.2,3) и (III.2,4) в виде

$$\sum \int \varepsilon \rho d\Omega + \sum \int \varepsilon' \rho' d\Omega' = \text{const}, \quad (\text{III.2,6})$$

$$\sum \int \rho d\Omega + \sum \int \rho' d\Omega' = \text{const}. \quad (\text{III.2,7})$$

Равновесная плотность следует теперь из условия минимальности (III.2,5), которое должно удовлетворяться одновременно с (III.2,1), (III.2,2), (III.2,6) и (III.2,7). Это приводит обычным способом (метод множителей Лагранжа; см., например, ESM, стр. 20 или 444) к следующему выражению для равновесного значения ρ

$$\rho = \exp[-q + \nu\mu - \beta\varepsilon]. \quad (\text{III.2,8})$$

Можно здесь отметить, что поскольку (III.2,5) справедливо, в действительности, лишь для крупноструктурной плотности, то выражение (III.2,8) на самом деле относится к P .

Далее можно указать, что наше предположение относительно априорных вероятностей использовано при написании (III.2,1) — (III.2,4) без введения, наряду с ρ , функции плотности, которая давала бы априорные веса различных объёмов в фазовом пространстве.

Выше мы показали способом, несколько отличным от использованного в предыдущем разделе, что канонический большой ансамбль будет изображать любую подсистему в равновесии. Нам необходимо, однако, уделить некоторое внимание проблеме неравновесных систем; в этом вопросе мы будем придерживаться изложения Толмена (38^т, особенно разделы 79b, 51c и 112). В заключение этого раздела будет обсуждён выбор равных априорных вероятностей для равных объёмов в фазовом пространстве.

Допустим, что наша система находится в неравновесном состоянии, в чём мы убедились, проделав соответствующие наблюдения над системой в момент t' . Для предсказания дальнейшего её поведения построим ансамбль, изображающий систему, используя, с одной стороны, сведения, полученные из опыта, и, с другой стороны, предположение о равенстве априорных вероятностей (или равных весах) для равных объёмов в фазовом пространстве. Как указывалось в предыдущем разделе, мы можем, самое большее, определить крупноструктурную плотность P' и, используя предположение о равенстве весов для равных объёмов, выбрать мелкоструктурную плотность ρ' в соответствии с (III.2,27). В последующее время t'' мелкоструктурная плотность ρ'' , следующая из ρ' , не будет более постоянной в каждой из ячеек, на которые мы разделили фазовое пространство. Как было указано ранее, это вытекает из того факта, что P'' и ρ'' теперь уже не равны. Из (III.1,32) усматриваем, что

$\Sigma' - \Sigma''$ тем больше, чем больше отличаются между собой P'' и p'' и, стало быть, мы вправе ожидать, что Σ будет непрерывно убывать, пока не установятся распределения (III.1,33) — (III.1,35).

Для того, чтобы убедиться в приближении к равновесию, мы можем выполнить соответствующие наблюдения над системой. Из предыдущего рассмотрения можно было бы полагать, что эти наблюдения приведут к крупноструктурным плотностям, таким, что Σ будет монотонно убывающей функцией времени. Однако, делая такое допущение, мы прежде всего молчаливо полагаем, что перемешивание различных ячеек фазового пространства происходит столь же полно, как и перемешивание красящего вещества в воде. Далее, мы при этом забываем принять во внимание тот факт, что каждый раз, когда мы проводим наблюдение, количество доступных нам сведений увеличивается, и, стало быть, наш выбор изображающего ансамбля в момент t''' , например не будет уже столь простым, как в момент t' , ибо необходимо принять во внимание сведения, полученные и в момент t' и в момент t'' . В отношении перемешивания видно, что мы возвращаемся к вопросу, затронутому в эргодической теореме Хопфа. Действительно, исходя из одинаковых основных положений, мы убеждаемся, что как рассмотрение при помощи эргодической теоремы, так и рассмотрение при помощи теории ансамблей не столь сильно отличаются друг от друга, как это могло показаться с первого взгляда.

Рассмотрим теперь проблему априорных вероятностей. Ограничимся для простоты малыми ансамблями. Рассмотрим две области $\mathfrak{A}_0$ и $\mathfrak{B}_0$ в Γ -пространстве. Пусть эти две области в момент t_0 заполнены изображающими точками ансамбля и пусть $p(\mathfrak{A}_0)$ и $p(\mathfrak{B}_0)$ — априорные вероятности для двух областей. В моменты $t_1, t_2, \dots$ изображающие точки, заполнявшие при t_0 области $\mathfrak{A}_0$ и $\mathfrak{B}_0$, будут заполнять области $\mathfrak{A}_1$ и $\mathfrak{A}_1, \mathfrak{A}_2$ и $\mathfrak{B}_2 \dots$ с априорными вероятностями $p(\mathfrak{A}_1)$ и $p(\mathfrak{B}_1), p(\mathfrak{A}_2)$ и $p(\mathfrak{B}_2) \dots$. Очевидно, что допустим лишь такой выбор априорных вероятностей, при котором

$$\frac{p(\mathfrak{A}_0)}{p(\mathfrak{B}_0)} = \frac{p(\mathfrak{A}_1)}{p(\mathfrak{B}_1)} = \frac{p(\mathfrak{A}_2)}{p(\mathfrak{B}_2)} = \dots, \quad (\text{III.2,9})$$

а эти соотношения удовлетворяются, если взять

$$p(\mathfrak{A}) = W(\mathfrak{A}), \quad p(\mathfrak{B}) = W(\mathfrak{B}), \quad (\text{III.2,10})$$

где $W(\mathfrak{A})$ или $W(\mathfrak{B})$ — объёмы областей $\mathfrak{A}$ или $\mathfrak{B}$ в Γ -пространстве. Отметим здесь, что соотношения (III.2,10) будут удовлетворять (III.2,9) лишь в том случае, если мы имеем дело с фазовым пространством, состоящим из канонически сопряжённых p и q ; это одна из причин для выбора канонических координат для описания фазы системы.

Вместо (III.2,10) можно использовать более общую форму

$$p(\mathfrak{A}) = \int F(\varepsilon, \varphi_i) d\Omega, \quad (\text{III.2,11})$$

где интегрирование распространено на область $\mathfrak{M}$ и где $F(\epsilon, \varphi_i)$ — раз навсегда выбранная функция энергии ϵ и других интегралов движения φ_i .

Причина, по которой полагают $F(\epsilon, \varphi_i)$ равной единице (при этом выбирается (III.2,10) вместо более общего соотношения (III.2,11)), была первоначально исторической и основывалась на допущении о существовании эргодических систем. Поскольку ныне мы убеждены, что механические системы являются квазиэргодическими, значит все различные возможные значения φ_i встречаются практически одинаково часто в любой части фазового пространства. Следовательно, мы можем сразу ограничиться, по выражению Эренфеста, эргодическим выбором

$$\rho(\mathfrak{M}) = \int F(\epsilon) d\Omega. \quad (\text{III.2,12})$$

Однако, так как плотности ρ изображающих ансамблей равновесных систем являются функциями только энергии, выбор (III.2,12) вместо (III.2,10) соответствовал бы плотности

$$\rho'(\epsilon) = \frac{\rho(\epsilon)}{F(\epsilon)} \quad (\text{III.2,13})$$

вместо $\rho(\epsilon)$, заданной выражениями (III.1,8) или (III.1,10). Величина $F(\epsilon)$ фигурировала бы во всех интегралах, однако окончательные результаты были те же самые.

Можно здесь упомянуть, что замечание Эренфеста о том, что эргодический выбор часто нецелесообразен, не относится теперь к существу дела, ибо оно основывалось на трудностях, с которыми встретились термостатистика до появления квантовой механики*).

РЕЗЮМЕ ПОЛОЖЕНИЯ В КЛАССИЧЕСКОЙ ТЕРМОСТАТИКЕ

Прежде, чем начинать обсуждение роли и положения H -теоремы и эргодической теоремы в квантовой статистике, мы хотим кратко резюмировать содержание частей I—III. Имеются по существу три различных подхода к проблеме оправдания метода статистической механики: 1) утилитарный (прикладной) подход, 2) формалистический подход и 3) физический подход**), хотя последние два в некоторых отношениях могут перекрываться.

Утилитарный подход оправдывает статистический метод полученными результатами и не заботится о дальнейшем обосновании. При этом подходе приветствуются вычисления, проводимые любым методом, будь то метод Дарвина — Фаулера, использование микроканона-

*) Можно также сослаться на работу Эренфеста^{14E} при обсуждении возможного выбора функций $F(\epsilon, \varphi_i)$.

**) Читатель, надеюсь, извинит меня за такую классификацию, которая явно свидетельствует о моих собственных предубеждениях и вкусах.

нических ансамблей, использование канонических ансамблей или же больших ансамблей, — лишь бы они приводили к полезным результатам. Тот факт, что для физических систем со многими степенями свободы все эти вычисления приводят к одинаковым результатам, используется в том смысле, что выбор окончательного метода является делом лишь удобства *).

При формалистическом подходе обычно приводятся аргументы о том, что механическая система должна быть изолированной и, стало быть, должна обладать определённым значением энергии, а также числом частиц в ней, поскольку лишь такие системы рассматриваются классической механикой. Единственным стационарным ансамблем, допускаемым к рассмотрению, является, таким образом, микроканонический малый ансамбль, хотя в качестве математического приёма для упрощения вычислений можно использовать макроканонический, т. е. большой ансамбль. Более того, микроканонический ансамбль был допущен лишь после того, как убедились, что среднее по времени от любой фазовой функции, вычисленное для одной системы, совпадает, повидимому, со средним, взятым по микроканоническому ансамблю. Для доказательства этого необходимо ввести эргодическую, или, лучше сказать, квазиэргодическую теорему. При этом приверженцы формалистической школы мышления помещают эргодическую теорему в центр своего рассмотрения.

При физическом подходе подчёркивается тот факт, что мы в состоянии получить лишь ограниченное количество сведений об интересующей нас системе, и, стало быть, мы должны работать с изображающими ансамблями. В этом случае необходимо доказать, что развитие неравновесного состояния будет происходить таким образом, что ансамбль, изображающий наблюдаемую систему, будет всё более и более близко приближаться к каноническому ансамблю.

Попытки оправдания предположений, лежащих в основе последних двух методов подхода, разбиваются на три группы. Можно, например, попытаться использовать непосредственно *H*-теорему, но, как было видно в части I, при этом возникают трудности. Непосредственно *H*-теорема давала бы другую альтернативу эргодической теореме, однако *H*-теорема в её статистической форме не приводит к этому, если только мы не желаем вводить различного рода предположения относительно вероятностей перехода, которые мы не в состоянии подтвердить более строго. Тем не менее статистическая форма *H*-теоремы полезна в том отношении, что она указывает относительную частоту, с которой происходят флуктуации, и скорость, с какой неравновесное состояние может возвратиться к равновесию.

*) Интересно отметить, что даже Толмен использовал эту утилитарную точку зрения, когда он оправдывал предположение о равенстве объёмов в Γ -пространстве.

Таким образом, мы приходим к эргодической теореме и, как было видно в части II, проблема в настоящее время сводится к доказательству метрической неразложимости соответствующих фазовых пространств. Если бы можно было показать, что физические системы являются, вообще говоря, квазиэргодическими, то эргодическая теорема была бы доказана; этого, однако, не сделано. Интересно отметить, что как Эйнштейн^{02E}, так и Пуанкаре^{06P} открыто заявляли, что, по их мнению, энергия является единственным не зависящим от времени однозначным, аналитическим интегралом физической системы. Если ввести несколько большую степень правдоподобности, допустив изменение энергии в малой, но конечной области, то можно прибегнуть к помощи эргодической теоремы Хопфа, в которой имеет место смещение различных энергетических поверхностей. Отметим ещё раз, что до сих пор ещё не доказана метрическая неразложимость входящих в рассмотрение фазовых пространств.

Если предпочесть подход посредством изображающих ансамблей, можно использовать H -теорему в форме, которая соответствует теории ансамблей, т. е. отождествляя H со средним значением натурального логарифма крупноструктурной плотности. Приближение к равновесию при этом следует довольно просто, однако рассмотрение основывается на специальном выборе априорных вероятностей. Как мы видели в разделе III.2, этот выбор может быть оправдан лишь в том случае, если рассматриваемые системы являются квазиэргодическими.

Таким образом, мы видим, что во всех методах отсутствует доказательство квазиэргодичности физических систем. Если бы такое доказательство было дано, можно было бы использовать либо формалистический подход, поскольку (квази) эргодическая теорема в этом случае доказана, либо физический подход, поскольку выбор априорных вероятностей будет оправдан. В заключение можно сказать, что хотя между формалистическим и физическим подходами на первый взгляд имеются большие отличия, эти отличия оказываются на самом деле не столь уж значительными, а оба подхода оказываются узвизимыми в одном и том же пункте.

(Окончание в следующем выпуске)
