А. Я. Хинчин

МАТЕМАТИЧЕСКИЕ ОСНОВАНИЯ КВАНТОВОЙ СТАТИСТИКИ
А. Я. ХИНЧИН

МАТЕМАТИЧЕСКИЕ ОСНОВАНИЯ КВАНТОВОЙ СТАТИСТИКИ

ГОСУДАРСТВЕННОЕ ИЗДАТЕЛЬСТВО ТЕХНИКО-ТЕОРЕТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА 1951 ЛЕНИНГРАД
ОГЛАВЛЕНИЕ

Предисловие ................................................. 5
Введение ....................................................... 7
 § 1. Важнейшие особенности математического аппарата квантовой статистики ........................................... 7
 § 2. Содержание книги .......................................... 12
Глава I. Предварительные сведения из теории вероятностей ...................................................... 23
 § 1. Целочисленные случайные величины ......................... 23
 § 2. Предельные теоремы ......................................... 29
 § 3. Метод характеристических функций .......................... 36
 § 4. Одномерная предельная теорема ............................ 47
 § 5. Двумерная предельная теорема ................................ 56
Глава II. Предварительные сведения из квантовой механики ......................................................... 69
 § 1. Описание состояния физической системы в квантовой механике .................................................. 69
 § 2. Механические величины и самосопряженные линейные операторы ................................................. 74
 § 3. Возможные значения механических величин ............... 81
 § 4. Изменение состояния системы во времени ................ 91
 § 5. Стационарные состояния. Закон сохранения энергии .... 96
Глава III. Общие начала квантовой статистики ................................................................. 104
 § 1. Основные идеи статистических методов в физике ........ 104
 § 2. Микрокаанонические средние ............................... 111
 § 3. Полная, симметрическая и антисимметрическая статистики .................................................... 117
 § 4. Построение основного линейного базиса .................. 124
 § 5. Числа заполнения. Первичные выражения структурных функций ................................................. 132
 § 6. О представительности микрокаанонических средних .... 139
Глава IV. Основы статистики фотонов ................................................................. 147
 § 1. Отличительные особенности статистики фотонов ........ 147
 § 2. Числа заполнения и их средние значения .................. 149
 § 3. Редукция к задаче теории вероятностей ................. 153
ГЛАВА V. ОСНОВЫ СТАТИСТИКИ МАТЕРИАЛЬНЫХ ЧАСТИЦ

1. Напоминание исходных положений ........................................ 177
2. Средние значения чисел заполнения ........................................ 179
3. Редукция к задаче теории вероятностей .................................. 187
4. Выбор значений параметров $\alpha$ и $\beta$ .................................. 194
5. Применение предельной теоремы теории вероятностей ............... 200
6. Средние значения суммарных величин ...................................... 204
7. Корреляционная связь между числами заполнения ................... 207
8. Дисперсии суммарных величин и представительность микроканонических средних .................................................. 211
9. Определение чисел $g_k$ для бесструктурных частиц в отсутствии внешних сил ...................................................... 214

ГЛАВА VI. ТЕРМОДИНАМИЧЕСКИЕ ВЫВОДЫ

1. Задачи статистической термодинамики ................................... 219
2. Внешние параметры, внешние силы и их средние значения .......... 221
3. Определение энтропии и вывод второго закона термодинамики ...... 226

ДОПОЛНЕНИЕ I. СТАТИСТИКА НЕОДНОРОДНЫХ СИСТЕМ .................... 231
ДОПОЛНЕНИЕ II. РАСПРЕДЕЛЕНИЕ КОМПОНЕНТЫ И ЕЕ ЭНЕРГИИ .......... 239
ДОПОЛНЕНИЕ III. ПРИНЦИП КАНОНИЧЕСКОГО ОСРЕДНЕНИЯ .......... 245
ДОПОЛНЕНИЕ IV. РЕДУКЦИЯ К ОДНОМЕРНОЙ ЗАДАЧЕ В СЛУЧАЕ ПОЛНОЙ СТАТИСТИКИ .................................................. 254
ПРЕДИСЛОВИЕ

В моей книге «Математические основания статистической механики» *, сданной в печать перед самым началом войны и посвящённой обоснованию классической статистической механики, мною было указано (стр. 11), что тот же метод в принципе может быть применён и к построению математических основ квантовой статистики. Учитывая, однако, что всё же при этом метод должен претерпеть некоторые видоизменения, я тогда же поставил себе задачей составление специальной монографии, посвящённой этому вопросу. Выполнение этого плана задержалось почти на десять лет — отчасти вследствие загруженности другой работой, отчасти же потому, что трудности, стоящие на намеченном мною пути, оказались более значительными, чем это мне представлялось первоначально. Необходимость включения в круг исследований "новых" статистик (симметрической и антисимметрической) потребовала довольно серьёзного изменения метода.

Однако, несмотря на более или менее значительные сдвиги технического характера, центральная идея метода остаётся неизменной. И в области квантовой статистики я ставлю себе целью показать, что строгое, логически планомерное математическое обоснование расчётных формул статистической физики не требует, как считалось обычно, создания специального громоздкого аналитического аппарата (метод Дарвина-Фаулера), но может быть достигнуто совершенно элементарной редукцией всех возникающих здесь задач к хорошо разработанным предельным теоремам теории вероятностей. Помимо своего чисто научного значения, которое очевидно и не

*) А. Я. Хинчин, Математические основания статистической механики, Гостехиздат, 1943.
требует комментариев, возможность такой редакции для нас, советских учёных, несёт с собою и некоторое особое удовлетворение. Учение о предельных теоремах теории вероятностей, основоположником которого был П. Л. Чебышев, в основном создано и продолжает развиваться русскими и советскими математиками. То, что это учение может составить собою аналитическую базу всех расчетных формул физической статистики, лишний раз показывает нам его прикладную ценность и позволяет ещё более гордиться этим фундаментальным созданием отечественной науки. Только недостаточным владением аналитическими методами теории вероятностей можно объяснить то, что зарубежная наука в течение тридцати лет не замечала этой связи.

Как и моя уже цитированная книга, эта монография целиком посвящена математическому методу излагаемой теории и ни в какой мере не может претендовать на полноту в физическом отношении; она вовсе не рассматривает никаких конкретных физических задач. В основном она обращена к читателю-математику. Я надеюсь, однако, что и физик, интересующийся математическим аппаратом своей науки, найдёт для себя в этой книге кое-что, что сможет его заинтересовать.

29 августа 1950 г.

А. Хинчин
ВВЕДЕНИЕ

§ 1. Важнейшие особенности математического аппарата квантовой статистики

Переход от классической механики к квантовой связан с существенной сменой основных идей и понятий этой науки. Нет поэтому ничего удивительного в том, что и математический аппарат статистической механики при переходе к концепциям квантовой физики должен быть подвергнут значительному изменению, которое в большинстве случаев выражается в его обобщении или уточнении, а иногда требует и введения существенно новых математических идей. Мы начнём с перечисления тех новых моментов квантовой физики, которые оказывают наибольшее преобразующее воздействие на её статистический аппарат.

Прежде всего упомянем два момента, которые, в значительной степени изменения внешний облик математического аппарата физической статистики, в то же время не затрагивают сколько-нибудь глубоко его содержания: 1) наличие у некоторых механических величин дискретного спектра (множества возможных значений) лишь внешним образом отражается на математическом аппарате, заставляя в известных случаях вместо обычных для классической механики интегралов пользоваться конечными суммами или бесконечными рядами, что ни в какой мере не касается руководящих идей статистической теории; 2) появление в составе механических величин, кроме обычных для классической механики гамильтоновых переменных, ещё новых, специфических для квантовой физики и не имеющих ничего аналогичного себе в классической теории, «спиновых» переменных — также ничего не меняет в руководящих идеях статистической физики, лишь
немного усложняя расчёты в тех или других случаях; чтобы не затемнять основных идей теории введением не имеющих (для математического метода) принципиального значения деталей подобного рода, мы в этой книге всюду, где это возможно, будем избегать упоминания о спине.

Следующим новым, чуждым классической физике, моментом квантовой механики является статистический характер её утверждений. В классической механике заданием состояния, в котором находится данная система, однозначно определяются значения всех связанных с нею механических величин, ибо всякая такая величина представляет как функция гамильтоновых переменных, задание значений которых и равносильно заданию состояния системы. В квантовой механике заданием состояния системы механические величины определяются лишь как случайные величины; задание состояния системы определяет собою не значения, а законы распределения связанных с нею механических величин. Эта принципиально статистическая черта квантовой механики отличает её совершенно независимо от применения специальных методов статистической физики; методы статистической физики всегда понимают под средними значениями тех или иных величин результаты осреднения их по различным состояниям системы; здесь же речь идёт о средних значениях величин в некотором определённом, фиксированном состоянии. Квантовая статистика является поэтому, в отличие от классической, статистической теорией в двояком значении этого слова. Очень важно тщательно различать между собою понятия и расчётные методы этих двух статистических концепций, вводя для каждой из них особую терминологию и особую систему обозначений; в настоящей книге мы будем твёрдо держаться этого правила, всерёз избегая смешения понятий, по сути дела не имеющих друг с другом ничего общего и сходных между собою лишь своей статистической природой.

Этот двояко-статистический характер квантовой статистики отражается на её математическом аппарате уже несколько больше, чем те два новых момента, о которых мы упоминали выше. Однако и здесь необходимые изменения не касаются основ самого метода. Принципиально-статистическая природа квантовой механики, именно потому, что она совершенно независима от специальных методов физической статистики, ничего
не меняет в сущности этих методов (а значит, и связанного с ними математического аппарата), требуя лишь некоторой надстройки, хотя и заметно изменяющей облик построенного здания, но ни в какой мере не затрагивающей его фундамента.

Наконец, мы должны более подробно рассмотреть две новые черты квантовой механики, которые оказывают на её статистический аппарат уже гораздо более глубокое преобразующее воздействие, требуя в некоторых случаях такого его расширения, которое знаменует собой уже качественное изменение.

Первая из этих черт вызывает необходимость охвата так называемых «новых» статистических схем (Бозе-Эйнштейна и Ферми-Дирака), не имеющих и не могущих иметь себе ничего аналогичного в классической статистической механике. Создающееся здесь положение, правда, по своей логической сущности принципиально возможно и в классической теории: речь идёт о необходимости формирования средних значений механических величин, произвольного осреднения не по всем состояниям, совместимым с данным значением полной энергии системы, а лишь по некоторой (небольшой) части этих состояний. В классической теории такая редукция осредняющего многообразия становится необходимой всякий раз, когда система уравнений движения имеет какой-либо однозначный интеграл, независимый от интеграла энергии; однако практически такая необходимость возникает редко, так как в обычных условиях, как правило, либо таких интегралов не бывает совсем, либо, если они имеются, средние по редуцированному многообразию оказываются практически неотличимы от средних по первоначальному полному многообразию.

Переход к «новым» статистикам означает именно такую редукцию многообразия «доступных» состояний системы, по которым должно производиться осреднение. Необходимость этой редукции и здесь вызывается наличием некоторого однозначного интеграла уравнения Шредингера, которое в квантовой механике заменяет собой систему «уравнений движения», описывающую эволюцию состояния системы с течением времени. Наличие этого интеграла (мы называем его в дальнейшем «индексом симметрии») является при этом правилом, а не исключением; и средние значения по редуцированным многообразиям отличаются от средних значений по первоначальному полному.
многообразиям в такой мере, которая делает учё́т этих разли́чий совершенно необходи́мым. В классической механике уравнения движения не могут иметь никаких интгралов, которые в какой-либо мере были бы аналогичны этому «индексу симметрии», описывающему таким образом некоторую специфи́ческую черту кванто́вой механики.

При построении общей статистической теории эта необхо́димость редукции осредняющих многообразий весьма существенно́м образом отражается на свойствах математического аппарата, заметно упроща́я его усложне́ния. Основной базой попрежнему оста́ются локальные предельные теоремы теории вероятностей. Однако даже в простейшем случае системы, состоящей из однотипных частиц, в условиях симметрической или антисимметрической статистики приходится пользоваться двумерными предельными теоремами вместо одномерных, которыми можно ограничиться в условиях полной статистики (т. е. в условиях классической «схемы Максвелла-Больцмана»). Редукция расчё́тных задач физической статистики к предельным теоремам теории вероятностей также претерпевает значительные изме́нения. Кроме того, сама необходимость вести все расчёты на весьма общей основе, охватывающей одновременно все три основные статистические схемы, естественным образом делает изложение более сложным.

Наконец, последняя специфическая черта квантовой меха́ники, также оказывающая на её статистические методы весьма существенное влияние, связана с проблемой «представи́тельности» *) даваемых этими методами средних значений, т. е. с проблемой обоснования реальной значимости, экспери́ментальной подтверде́нности этих средних. В классической статистической механике с этой целью обычно строятся так называемые эргодические гипотезы или теоремы, имеющие целью установить, что в своей эволюции, описываемой уравнениями движения, система в среднем пребывает в различных частях данного многообразия постоянной энергии такие доли общего промежутка времени, которые пропорциональны объё́м этих частей; поэтому, если мы станем через определённые промежутки времени измерять какую-либо связанную с данной системой механическую величину, то среднее арифме-

*) Или, как иногда говорят, «репрезентативности».
тическое полученных в достаточно большом числе результатов измерения будет, как правило, близко к статистическому среднему её значению, даваемому теорией. Известно, что все попытки такого «эргодического» пути обоснования представительности теоретических средних не привели до сих пор в классической статистической механике к сколько-нибудь законченным результатам (несмотря на ряд отдельных замечательных успехов). Однако независимо от этого, такой «эргодический» путь в квантовой механике с самого начала оказывается принципиально невозможным. Если классическая механическая система изменяет своё состояние согласно уравнениям движения, то с течением времени состояние её — по крайней мере в принципе — может как угодно близко подходить к любому наперёд заданному состоянию, соответствующему данному значению её полной энергии; на этом и основана возможность ставить вопрос о сравнении теоретических осреднений по всем таким состояниям с данными, полученными путём измерения той или иной величины на одной и той же системе в различные моменты её эволюции. В квантовой механике дело обстоит совершенно иначе. Если система имеет определённое (достоверное) значение полной энергии (находится, как говорят, в «стационарном» состоянии) и эволюционирует согласно уравнению Шредингера, то закон распределения любой связанной с ней механической величины остаётся неизменным с течением времени (мы докажем это в гл. II, § 5). Но в квантовой статистике задание состояния системы не даёт для связанных с нею механических величин ничего иного, кроме их законов распределения; мы должны поэтому считать, что состояние системы, имеющей некоторое определённое (достоверное) значение полной энергии, вообще с течением времени не меняется; поэтому серия измерений, произведённых над одной и той же системой (даже если бы такая серия была возможна без радикального нарушения состояния системы при каждом отдельном измерении) должна была бы давать результат, ничего общего не имеющий с теоретическим средним, представляющим собой результат осреднения величины по всем состояниям, совместимым с данным значением полной энергии.

Таким образом, как мы ни расценивали эффективность эргодических методов в классической статистической механике,
в квантовой статистике они принципиально ничего не могут дать для обоснования представительности теоретических средних значений механических величин, так как здесь «временные средние» таких величин в силу приведённых нами сообщений будут, как правило, во всех наблюдаемых случаях существенно отличны от теоретических средних. Квантовая статистика должна поэтому при выборе своего математического аппарата полностью учитывать необходимость создания других приёмов установления представительности средних значений. Как мы увидим, конкретно это означает, что точность оценки остаточных членов в используемых предельных теоремах теории вероятностей должна быть существенно повышена сравнительно с той, какая требуется для простой оценки средних значений.

Несмотря на все указанные нами, порою весьма значительные, расширения, дополнения и просто изменения, какие терпевает математический аппарат физической статистики при переходе от классической к квантовой физике, мы хотим ещё раз подчеркнуть, что центральная идея развиваемых нами методов при этом переходе остаётся незыблёной. Эта идея состоит в систематическом применении ко всем расчётам статистической физики асимптотических формул теории вероятностей, представляющей собою общее учение о массовых явлениях, и в построении таким образом для статистической физики строго обоснованного математического фундамента, без необходимости создания какого-либо специального аналитического аппарата.

§ 2. Содержание книги

Как было указано в предисловии, эта книга рассчитана на две категории читателей: физиков, интересующихся математическими основаниями своей науки, и математиков, которые хотят ознакомиться с физическими приложениями математики. Эти две группы приступают к чтению книги, как правило, с различным характером предварительной подготовки. Учитывая это различие и желая дать каждой из двух категорий читателей хотя бы самое необходимое для беспрепятственного овладения основными разделами книги, мы несколько расширили вводную часть её по сравнению с обычными в таких слу-
чаях размерами. Две большие главы (первая и вторая) целиком посвящены предварительным сведениям, и лишь с третьей главы начинается трактовка вопросов квантовой статистики.

Первая глава имеет целью изложение и полное доказательство тех предельных теорем теории вероятностей, которые затем будут использованы в основных разделах книги. Речь идёт здесь о предельных теоремах локального типа для сумм одинаково распределённых случайных величин, могущих принимать лишь целые неотрицательные значения. Как известно, общие условия применимости теорем такого рода были найдены лишь совсем недавно Б. В. Гнеденко и его учениками. В гл. I дано полное доказательство локальных теорем для одномерного и двумерного случаев, проведённое в основном методом Б. В. Гнеденко; однако в целях нужных нам приложений мы проводим расчёты несколько более детально, чтобы получить не только асимптотические формулы, но и достаточно точную оценку остаточных членов. Таким образом, эта глава содержит известный элемент новизны даже для математика, специальностью которого является теория вероятностей. Для математиков других специальностей, а также для физиков, она будет несомненно целиком новой. Тем из читателей, которые не интересуются деталями доказательств нужных в дальнейшем предельных теорем, мы можем рекомендовать не читать первой главы целиком, а ознакомиться лишь с формулировками доказываемых в ней предложений. Эти формулировки со всей отчётливостью выделены нами в конце §§ 4 и 5.

Вторая глава содержит необходимые предварительные сведения из квантовой механики; для образованного физика она, как правило, будет излишней, и мы можем рекомендовать ему лишь бегло проглядеть её, чтобы освоиться с принятой во всём дальнейшем терминологией и системой обозначений. Для математика она, напротив, как правило, окажется необходимой. Мы должны предупредить эту категорию читателей, что ознакомление с этой главой не может заменить собой предварительного усвоения основных идей квантовой физики, хотя бы по литературе более или менее популярного характера. Наша гл. II не может рассматриваться ни как краткий курс, ни даже как конспект курса квантовой механики; выбор
пизлагаемого в ней материала не претендует ни в какой мере на какую бы то ни было полноту, и целиком обусловлен интересами тех специальных приложений, которым посвящены последующие главы. В частности, эта глава говорит почти исключительно о математическом аппарате квантовой механики, оставляя, как правило, в стороне содержательно-физическую сторону вопроса. Таким образом, хотя с формальной стороны эта глава содержит всё необходимое для понимания последующих разделов книги, однако для читателя, совсем не знакомого с идеями квантовой физики, чтение её окажется почти бесполезным: знания его будут чисто формальными, висящими в воздухе (достаточно указать, что вся глава не содержит упоминания ни об одном эксперименте). Мы повторяем поэтому, что приступающий к изучению нашей книги математик обязательно должен иметь хотя бы весьма скромное знакомство с основным идейным багажом, с общим стилем квантовой физики; как мы уже говорили, знакомство это может быть достигнуто с помощью источников совсем элементарного характера; если ведущие физические идеи этого учения читателю хоть в скромной мере знакомы, то наша вторая глава уже без затруднений сможет поднять эти знания на нужный для целей прочного понимания последующих глав математический уровень.

Третья глава содержит изложение общих идей и основ расчётных методов квантовой статистики. В классической механике статистическая теория имеет основной целью исследование статистики различных механических величин при заданной полной энергии системы; подобным же образом квантовая статистика исследует законы распределения различных механических величин при условии, что полная энергия системы имеет некоторое достоверное значение; таким образом, в рассмотрение входят (по крайней мере, в первую очередь) только такие состояния системы, в которых полная энергия имеет достоверные значения (и которые описываются, следовательно, собственными функциями оператора $H$ полной энергии). В классической механике совокупность состояний, в которых полная энергия системы имеет заданное постоянное значение, образует некоторую «поверхность постоянной энергии» в фазовом пространстве. В квантовой механике аналогичная совокупность состояний представляет собою линейное многообразие $\mathcal{M}$, эле-
ментами которого являются собственные функции оператора $H$, принадлежащие к некоторому определённому собственному значению этого оператора. Это многообразие всегда имеет конечное, но для систем статистической физики очень большое число измерений (высокая степень выражения собственных значений). В классической механике средние значения величин получаются путём осреднения по данной поверхности постоянной энергии (или её части, если кроме интеграла энергии имеются другие однозначные интегралы); подобным же образом, в квантовой статистике осреднение должно производиться по многообразию $\mathcal{M}$ или его части. Как мы отметили в § 1, фактически для большинства систем статистической физики это многообразие приходится существенным образом редуцировать, вводя в рассмотрение только симметрические или только антисимметрические собственные функции.

Таким образом, в статистических задачах квантовой физики необходимо развивать расчётные приёмы для трёх основных статистических схем — полной, симметрической и антисимметрической. С этой целью мы прежде всего для каждой из этих трёх схем устанавливаем определённую полную ортогональную систему собственных функций; эти функции, имеющие существенное значение для всего дальнейшего, мы называем основными собственными функциями, а описываемые ими состояния — основными состояниями системы.

Далее вводится основное для квантовой статистики понятие «чисел заполнения», показывающих, сколько из составляющих данную систему частиц находится в том или другом определённом состоянии. Выбранные нами основные состояния оказываются особенно удобными для статистических расчётов, потому что в каждом из этих состояний числа заполнения получают определённые (достоверные) значения. Таким образом, каждому основному состоянию, в любой из трёх статистических схем, соответствует некоторый определённый набор чисел заполнения. Обратно, каждому заданному набору чисел заполнения соответствует одно или несколько основных состояний; это число основных состояний, отвечающих данному набору чисел заполнения, различно для трёх основных статистических схем, и именно в этом различии находится своё важнейшее выражение статистическая разнородность этих схем.
Значительная часть важнейших механических величин, изучаемых статистической физикой, имеет «сумматорный» характер, т. е. является суммой величин, каждая из которых зависит от состояния только одной из составляющих данную систему частиц. Для такой «сумматорной величины» знание средних значений чисел заполнения позволяет непосредственно написать выражение для её среднего значения. Если в дополнение к средним значениям чисел заполнения мы умеем найти и средние значения их попарных произведений, то это даёт нам возможность непосредственно написать и дисперсию любой сумматорной величины. Всеми этими соображениями объясняется, почему в любом систематическом изложении квантовой статистики авторы считают своей важнейшей первоначальной целью отыскание именно средних значений чисел заполнения. Однако, всё же отметим, что средними значениями чисел заполнения определяются средние значения только сумматорных величин; будучи важнейшими, сумматорные величины не исчерпывают собы, однако, всех величин, которыми может интересоваться статистическая физика; естественным объектом её исследований может стать любая величина, симметричным образом зависящая от состояний частиц, составляющих собы данную систему; сумматорные величины, будучи среди таких симметрических функций простейшими и наиболее часто встречающимися, вместе с тем не исчерпывают собы, очевидно, их совокупности (так, дисперсия сумматорной величины уже не есть сумматорная величина). Для математической теории охват более широкого круга задач был бы, несомненно, интересен и благодарной задачей; надо, однако, отметить, что в самой теории вероятностей предельные закономерности для симметрических функций большого числа случайных величин представляют собы ещё совершенно неразработанную область исследований.

В конце третьей главы мы показываем, что задача обоснования представительности микроканонических средних может быть сведена к оценке микроканонических дисперсий соответствующих механических величин, и даём выражение дисперсии сумматорной величины, имеющее силу для всех трёх статистических схем.

После того как таким образом фундамент статистических методов в квантовой физике заложен, мы в четвёртой и пятой
главах даём конкретное построение квантовой статистики. Четвёртая глава посвящена статистике фотонов, пятая — статистике материальных частиц (т. е. частиц с отличной от нуля «массой покоя»). Мы начинаем с фотонов исключительно из педагогических соображений. Как известно, то обстоятельство, что число составляющих данную систему фотонов не является постоянным, а может меняться с течением времени, делает статистику «фотонного газа» значительно более простой, чем статистика систем, состоящих из материальных частиц. Мы сочли поэтому целесообразным развить все рассчётные методы сначала на этом наиболее простом примере, где на первых порах приходится иметь дело только с одномерными задачами; мы надеемся, что читатель, усвоивший эту главу, при переходе к более сложному случаю материальных частиц будет уже настолько знаком с основными идеями метода, что те чисто технические усложнения, с которыми он здесь встретится, не создадут для него сколько-нибудь значительных затруднений.

Построение основных рассчётных формул протекает в этих двух главах в полном параллелизме. Число измерений линейного многообразия собственных функций оператора $H$, принадлежащих к данному собственному значению $E$, есть функция от $E$ (в случае материальных частиц зависящая ещё и от числа частиц, составляющих данную систему), которую мы называем структурной функцией данной системы. Первый этап состоит в точном выражении средних значений чисел заполнения и их попарных произведений через различные значения структурной функции; эти выражения очень просты, но различны для различных статистических схем. Найденные таким образом формулы полностью редуцируют асимптотическую оценку средних значений чисел заполнения и их попарных произведений к отысканию приближённых выражений для структурной функции. Следующий этап состоит в выражении структурной функции через закон распределения случайной величины, определяемой как сумма очень большого числа взаимно независимых и одинаково распределённых случайных величин. Эти законы распределения в общем случае многомерны; лишь в простейшей задаче для фотонов они являются одномерными. Наконец, на последнем, третьем этапе к асимптотической оценке этих законов распределения применяются
установленные в гл. 1 предельные теоремы теории вероятнос-
ств. Это даёт для структурной функции, а следовательно, и для средних значений чисел заполнения и их попарных произведений, удобные и вместе с тем очень точные приближённые выражения, с помощью которых мы, пользуясь методами гл. III, легко находим столь же точные приближённые выражения для средних значений и дисперсий сумма-
торных величин. Получаемая точность оказывается при этом вполне достаточной для обоснования представительности микроканонических средних сумматорных механических велик.

Последняя, шестая глава имеет целью наметить путь, по-
зволяющий построить с помощью достигнутых результатов понятие энтропии и обоснование второго закона термодина-
мики, а тем самым и закончить обоснование термодинамики статистическими методами.

В «дополнениях», завершающих книгу, мы собрали ряд вопросов, представляющих значительный интерес, но лежа-
щих несколько в стороне от основной линии развития тео-
рии; именно поэтому мы и предпочли предоставить им особое место в конце книги.

Нам остается сделать несколько замечаний общего харак-
тера, касающихся особенностей принятого в этой книге из-
ложения предмета и необходимых для правильного понима-
ния дальнейшего.

1. Со времени первых работ Дарвина и Фауэлера, т. е. уже в течение примерно 30 лет, статистическая физика в лучших своих изложениях оперирует со средними значениями физических величин вместо их наивероятнейших значений, безраздельно господствовавших до этой эпохи. За этот сдвиг говорят по меньшей мере два весьма основательных сообра-
жения: 1) в тех случаях, когда среднее и наивероятнейшее значения величины существенно различны между собою, для массовой практики всегда определяющую роль играет именно среднее значение, и 2) методы расчёта наивероятнейших зна-
чений, принятые в статистической физике, всегда заслужи-
вали вполне справедливого упрёка в математическом несовершенстве, в то время как метод расчёта средних значений, созданный Дарвином и Фаулером около 30 лет тому назад, безукоризненно строг в математическом отношении.
В предлагаемой книге речь будет идти неизменно о средних значениях физических величин. Правда, для большинства величин, интересующих статистическую физику, различие между их средними и наивероятнейшими значениями настолько мало, что им практически можно пренебречь; однако, чтобы строго обосновать возможность такого пренебрежения (а не просто принять её как очевидный факт, что обычно делается сторонниками наивероятнейших значений), всё равно приходится излагать достаточно развитую теорию средних значений, так что наивероятнейшие значения в конечном счёте оказываются ненужными.

Единственное преимущество (математически несостоятельного) расчёта наивероятнейших значений заключается в его неоспоримой сравнительной простоте. Несомненно, что метод Дарвина и Фаулера, основанный на специально создаваемом громоздком аналитическом аппарате, в математическом отношении очень сложен; этим и объясняется сравнительно малая популярность его среди физиков. Но, как мы уже кратко отметили в предисловии, основной целью настоящей книги является показать, что для строгого обоснования методов расчёта средних значений физических величин нет необходимости в создании какого-либо специального аналитического аппарата. Расчёт этот полностью и совершенно элементарно приводится к применению общих и известных предельных теорем теории вероятностей. Таким образом, и последнее, чисто практическое возражение против перехода от наивероятнейших значений к средним должно отпасть.

2. Знакомый с предметом читатель, вероятно, обратит внимание на то, что в нашей книге, в противоположность большинству современных изложений предмета, нигде не упоминается о так называемых «статистических ансамблях»; систематическое употребление этого термина обычно означает, что изучаемая физическая система на протяжении всего исследования мыслится как элемент некоторого множества физических систем, имеющих ту же структуру, что и данная, но находящихся в различных состояниях. Само собою разумеется, что всякая статистическая теория даёт для физических величин значения, осреднённые по различным состояниям изучаемой системы, и что в действительности такому
значению, вообще говоря, соответствует некоторое среднее значение результатов большого числа наблюдений или экспериментов; эта форма контакта статистической теории с действительностью на протяжении нашей книги не только часто подчёркивается, но и подвергается более детальному анализу. Нам представляется, однако, что проводимое многими авторами систематическое привязывание к каждому из возможных состояний данной системы некоторой особой системы той же структуры (что и создает «ансамбль») совершенно излишне и только обременяет усвоение теории; мы предпочитаем иметь дело с множеством самих возможных состояний (фазовым пространством в классической физике), а не с множеством систем, привязанных к этим состояниям и лишь усложняющих картину явления. Эта же точка зрения была проведена и в нашей книге, посвященной математическим основаниям классической статистической механики: вместо того чтобы размещать в фазовом пространстве целый «ансамбль» однотипных систем и затем следить за эволюцией этого «ансамбля», мы просто говорили о «естественном движении» самого фазового пространства, полагая, что образ непрерывно преобразующегося в самого себя пространства (связанный, кстати сказать, с простой гидродинамической моделью) в одинаковой мере прост и убедителен как для математика, так и для физика, и только теряет в этой своей простоте от не вызываемого никакой необходимостью привешивания к каждой точке фазового пространства некоторой физической системы.

В порядке терминологического примечания заметим ещё, что пользование неуклюжим и плохо звучащим на русском языке иностранным термином «ансамбль» решительно ничем не оправдано. В математике (теории множеств) лет 40—50 тому назад тоже привился этот термин, однако очень скоро он был вытеснен равнозначными русскими словами «совокупность» и «множество», из которых в настоящее время окончательно утвердилось последнее. Правда, французское слово «ensemble» семантически очень хорошо передаёт тот синтез единства и множества («многое как единое»), который здесь имеется в виду; однако, например, русское слово «совокупность» в этом отношении ровно ничем ему не уступает.
3. Учитывая методологическую цель книги, мы в качестве объекта исследования всюду избираем лишь простейшие случаи, сознательно отказываясь от рассмотрения более сложных ситуаций, чтобы читатель имел возможность сосредоточить всё внимание на математическом методе. В основном тексте мы всюду ограничиваемся рассмотрением однородных (т. е. состоящих из частиц одинаковой структуры) систем, и лишь в дополнении I показываем, как развиваемый нами метод может быть применён к неоднородным системам. Частицы, составляющие данную систему, предполагаются заключёнными в сосуд данного постоянного объёма, что, как известно, даёт им дискретный энергетический спектр

$$\varepsilon_1 \leq \varepsilon_2 \leq \ldots \leq \varepsilon_r \leq \ldots$$

Уровни $\varepsilon_r$ энергии частицы мы, как это обычно в такого рода исследованиях, предполагаем целыми числами, что всегда может быть выполнено с достаточно точным приближением, если выбрать достаточно малую единицу энергии. Далее, мы, как это тоже обычно, предполагаем, что энергия системы равна сумме энергий составляющих её частиц. Разумеется, это означает, что мы пренебрегаем взаимным потенциалом частиц, т. е., строго говоря, ограничиваемся рассмотрением многоатомных идеальных газов. Известно, что последовательное, до конца идущее проведение этой точки зрения невозможно, так как при отсутствии взаимодействия между частицами они не могут обмениваться энергией, и вся статистическая задача лишается смысла. Обычно из этого затруднения выходят, допуская (и, повидимому, во многих случаях с достаточным основанием), что взаимодействие, имеющееся между частичками, достаточно для обеспечения между ними свободного обмена энергией, но вместе с тем настолько слабо, что при всех энергетических расчётах практически можно полагать энергию системы равной сумме энергий составляющих её частиц, пренебрегая таким образом «смешанными» членами, выражающими собою взаимный потенциал частиц.

Развиваемый нами метод имеет конечной целью вывод всех нужных квантовой статистике асимптотических формул. Как обычно, формулы эти устанавливаются нами в предположении, что число частиц $N$ данной системы, её полная
энергия $E$ и занимаемый ею объём $V$ — бесконечно большие величины, сохраняющие между собой постоянные отношения (постоянная средняя энергия частицы и постоянная средняя плотность газа). Практически это означает, разумеется, такой выбор единиц энергии и объёма, при котором отношения $\frac{E}{N}$ и $\frac{V}{N}$ не были бы ни слишком малыми, ни слишком большими. Все характеризующие данную систему величины, зависящие лишь от этих отношений, должны поэтому рассматриваться в получаемых нами асимптотических формулах как постоянные.
ГЛАВА I
ПРЕДВАРИТЕЛЬНЫЕ СВЕДЕНИЯ ИЗ ТЕОРИИ ВЕРОЯТНОСТЕЙ

§ 1. Целочисленные случайные величины

Задачей этой книги является строгое и детальное математическое обоснование важнейших формул квантовой статистики. Это обоснование проводится с помощью предельных теорем теории вероятностей. При этом во всех случаях речь идёт о предельных теоремах несколько особого типа, давно уже интересующих специалистов и получивших, особенно в последние годы в СССР, значительное развитие, но тем не менее не вошедших ещё, как правило, в учебные руководства *), и потому мало известных широким кругам учёных. Это побуждает нас дать в настоящей первой главе как детальные формулировки, так и полные доказательства нужных нам предельных теорем, причём у читателя мы предполагаем знакомство лишь с руководствами общего характера (такими, как известные курсы С. Н. Бернштейна или В. И. Глиевенко).

Нужный нам тип предельных теорем отличается следующими важнейшими специфическими чертами:

1) речь идёт всегда о случайных величинах, все возможные значения которых — целые числа;

2) все интересующие нас предельные теоремы — локальной природы, т. е. речь всегда идёт об асимптотической оценке

вероятности того или иного определённого значения изучаемой суммы случайных величин;

3) мы имеем возможность ограничиться рассмотрением сумм взаимно независимых и одинаково распределённых случайных величин, имеющих конечные моменты до пятого порядка включительно;

4) во всех случаях нам придётся, не ограничиваясь отысканием асимптотической формулы, ставить вопрос о более или менее точной оценке погрешности;

5) наконец, нас наряду с одномерными будут в равной степени интересовать и многомерные (в особенности двумерные) предельные теоремы описанного типа.

Локальные предельные теоремы для целочисленных случайных величин уже сравнительно давно привлекали к себе внимание исследователей, хотя им удавалось и значительно меньше усилий, чем теоремам «интегрального» типа. Так, в уже цитированной нами книге Мизеса можно найти довольно далеко идущие результаты в этом направлении. Однако достаточно общая постановка встающих здесь задач принадлежит лишь последнему времени. Предельные теоремы нужного нам общего типа были впервые доказаны Б. В. Гнеденко *) и его учениками **). Работами этой школы охвачены и многомерные задачи; но всё же основное направление этих исследований одной чертой существенно отличается от того, что нужно нам здесь: в то время как Б. В. Гнеденко и его ученики всегда стремятся найти возможно более широкие условия, при соблюдении которых имеет место основное предельное соотношение, мы, как уже отмечено выше, имеем возможность ограничиться достаточно узким классом исходных распределений, но зато не можем довольствоваться установлением предельных соотношений, а вынуждены искать ещё (иногда довольно точную) оценку получаемой погрешности. Таким образом, хотя методы, созданные Б. В. Гнеденко, и вполне достаточны для наших целей, всё же нужные нам


**) Д. Г. Мейзер, О. С. Парасюк, Е. Л. Рвачева, О многомерной локальной предельной теореме теории вероятностей, Укр. мат. журнал, № 1, 1949.
формулировки предельных теорем несколько отличаются от даваемых им и его школой. Это обстоятельство послужило дополнительной причиной, заставшей нас предпослать настоящей книге главу, содержащую детальное доказательство предельных теорем в нужной нам форме.

Случайные величины, с которыми нам придётся иметь дело в этой книге, всегда будут, как уже сказано, иметь своими возможными значениями только целые числа. Мы будем называть такие случайные величины целочисленными. Очевидно, закон распределения целочисленной случайной величины $\xi$ полностью определяется заданием для каждого целого числа $n$ вероятности

$$P(\xi = n) = p_n$$

того, что величина $\xi$ получит значение $n$. При этом мы, разумеется, всегда имеем $p_n \geq 0$ ($-\infty < n < +\infty$) и $\sum_{n=-\infty}^{+\infty} p_n = 1$.

Мы будем в дальнейшем кратко говорить, что величина $\xi$ подчиняется закону $p_n$ или распределена по закону $p_n$.

Если ряд

$$\sum_{n=-\infty}^{+\infty} n p_n$$

абсолютно сходится, то сумма его называется математическим ожиданием $E \xi$ величины $\xi^*$. Вообще, в случае абсолютной сходимости ряда

$$\sum_{n=-\infty}^{+\infty} f(n) p_n,$$

где $f(n)$ — любая вещественная или комплексная функция целочисленного аргумента $n$, мы будем называть сумму этого ряда математическим ожиданием $E f(\xi)$ случайной величины $f(\xi)$. В частности, математическое ожидание $E \xi^k$ величины $\xi^k$ (если оно существует) называется моментом порядка $k$ величины $\xi$. Математическое ожидание $E (\xi - E \xi)^k$ величины

*) Иногда вместо математического ожидания пользуются термином «среднее значение» случайной величины $\xi$; мы в настоящей книге будем тщательно избегать этой терминологии, так как термин «среднее значение» будет у нас иметь совсем другой смысл.
(ξ - E ξ)^k (если оно существует) называется центральным моментом порядка k величины ξ. В частности, центральный момент второго порядка

\[ D \xi = E (\xi - E \xi)^2 = \sum_{n=-\infty}^{+\infty} (n - E \xi)^2 p_n \]

(если он существует) называется дисперсией величины ξ и представляет собою наряду с математическим ожиданием E ξ одну из важнейших характеристик этой величины. Все случайные величины, с которыми нам придететь иметь дело в дальнейшем, будут фактически обладать моментами любого порядка k \( \geq 0 \). Однако для доказательства нужных нам предельных теорем будет, как мы увидим, достаточно предположить существование моментов лишь сравнительно невысоких порядков.

Если ξ' и ξ'' — целочисленные случайные величины, подчиняющиеся соответственно законам p'_n и p''_n, то и сумма ξ' + ξ'' = ξ — целочисленная случайная величина; закон распределения p_n этой суммы зависит, кроме законов p'_n и p''_n, ещё и от вида взаимной зависимости между величинами ξ' и ξ''. В частности, если при этих двух величинах взаимно независимы, то числа p_n очень просто выражаются через числа p'_n и p''_n; в самом деле, для того чтобы было ξ = n, необходимо и достаточно иметь ξ' = k, ξ'' = n - k, где k — какое-либо целое число; поэтому

\[ P (\xi = n) = \sum_{k=-\infty}^{+\infty} P (\xi' = k, \xi'' = n - k), \]

откуда в силу взаимной независимости величин ξ' и ξ''

\[ p_n = \sum_{k=-\infty}^{+\infty} p'_k p''_{n-k}, \]

что иначе можно записать так:

\[ p_n = \sum_{k+l=n} p'_k p''_l. \]

Подобным же образом, если мы имеем s взаимно независимых целочисленных случайных величин

ξ(1), ξ(2), ..., ξ(s)
с соответственными законами распределения
\[ p^{(1)}_n, \ p^{(2)}_n, \ldots, \ p^{(s)}_n, \]
то закон распределения \( p_n \) суммы
\[ \xi = \sum_{i=1}^{s} \xi(i) \]
получает выражение
\[ p_n = \sum_{k_1=-\infty}^{+\infty} \sum_{k_2=-\infty}^{+\infty} \ldots \sum_{k_{s-1}=-\infty}^{+\infty} p^{(1)}_{k_1} p^{(2)}_{k_2} \ldots p^{(s-1)}_{k_{s-1}} p^{(s)}_{\sum_{i=1}^{s-1} k_i} (1) \]
или, что то же,
\[ p_n = \sum_{\sum_{k_i=n}^{s} k_i} p^{(1)}_{k_1} p^{(2)}_{k_2} \ldots p^{(s)}_{k_s}. \]

Выражение закона распределения суммы взаимно независимых случайных величин через законы распределения слагаемых называют правилом композиции этих законов. Формулы (1) и (2) выражают, таким образом, правило композиции законов распределения любого числа целочисленных случайных величин.

Из элементов теории вероятностей известно, что математическое ожидание суммы любого числа случайных величин во всех случаях равно сумме их математических ожиданий. В случае, когда слагаемые взаимно независимы, аналогичное правило имеет место и для произведений. Наконец, дисперсия суммы равна сумме дисперсий слагаемых при условии, что слагаемые попарно взаимно независимы.

Нам придётся неоднократно встречаться со случаем, когда исходным элементом исследования служит не одна случайная величина, а группа из нескольких (двух, трёх или более) взаимно зависимых целочисленных случайных величин \( \xi, \eta, \ldots \). Ради краткости записи рассмотрим случай пары \( (\xi, \eta) \) таких величин (всё, что будет сказано для этого случая, с соответствующими очевидными изменениями имеет место и для троек, четвёрок и т. д. величин); такую пару иногда называют (двумерным) случайным вектором. Вероятность \( P(\xi=l, \eta=m) \) совместной реализации равенств \( \xi=l \)
и \( \eta = m \) мы будем обозначать через \( p_{lm} \); совокупность чисел 
\( p_{lm} (-\infty < l, m < +\infty) \) образует собою закон распределения случайной пары (вектора) \( (\xi, \eta) \). Если \( p_n \) и \( q_n \) соответственно означают законы распределения величин \( \xi \) и \( \eta \), то, очевидно,

\[
p_l = \sum_{m=-\infty}^{+\infty} p_{lm}, \quad q_m = \sum_{l=-\infty}^{+\infty} p_{lm}.
\]

Отсюда

\[
E\xi = \sum_{l=-\infty}^{+\infty} l p_l = \sum_{l=-\infty}^{+\infty} l \sum_{m=-\infty}^{+\infty} p_{lm},
\]

и аналогично

\[
E\eta = \sum_{m=-\infty}^{+\infty} m \sum_{l=-\infty}^{+\infty} p_{lm},
\]

в предположении, разумеется, что все рассматриваемые ряды абсолютно сходятся.

Если \( f(\xi, \eta) \)— произвольная вещественная или комплексная функция величин \( \xi \) и \( \eta \), то её математическим ожиданием называется величина

\[
E f(\xi, \eta) = \sum_{l=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} f(l, m) p_{lm}
\]

при условии, что двойной ряд абсолютно сходится. В частности, из формул (3) или (4) мы получаем выражения для дисперсий величин \( \xi \) и \( \eta \) с помощью чисел \( p_{lm} \):

\[
D\xi = E(\xi - E\xi)^2 = \sum_{l=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} (l - E\xi)^2 p_{lm},
\]

\[
D\eta = E(\eta - E\eta)^2 = \sum_{l=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} (m - E\eta)^2 p_{lm}.
\]

Как известно, отношение

\[
R(\xi, \eta) = \frac{E(\xi \eta - E\xi E\eta)}{\sqrt{D\xi D\eta}} = \frac{E(\xi \eta) - E\xi E\eta}{\sqrt{D\xi D\eta}}
\]

называют коэффициентом корреляции величин \( \xi \) и \( \eta \). Числитель этого отношения представляется в виде

\[
E\{(\xi - E\xi)(\eta - E\eta)\} = \sum_{l=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} (l - E\xi)(m - E\eta)p_{lm}.
\]
§ 2. Предельные теоремы

Подобно всем математическим теориям точного естествознания — таким, как теоретическая механика, термодинамика и многие другие — теория вероятностей стремится в подведенственной ей области устанавливать в первую очередь закономерности наиболее общего характера, которые относились бы не к отдельным происходящим в природе и человеческой практике процессам, а охватывали бы собой возможно более широкий круг явлений. И подобно тому, как, например, фундаментальные теоремы механики — теорема живых сил, теорема площадей и др. — относятся не к какому-либо специальному виду механического движения, а к весьма широким классам таких движений, так и основные предложения теории вероятностей (такие, как, например, закон больших чисел) охватывают собой не какие-либо специальные виды массовых явлений, но весьма широкие классы их. Можно сказать, что в закономерностях этого рода в известной мере и в известном направлении раскрывается сущность массовых явлений как таковых, т. е. те черты этих явлений, которые обуслов-
ливаются в первую очередь именно их массовым характером и, напротив, сравнительно мало зависят от индивидуальной природы тех объектов, из которых состоят изучаемые массы. Закону больших чисел подчиняются суммы случайных величин, отдельные слагаемые которых могут быть распределены (в весьма широких пределах) по любым законам; ни применимость, ни содержание закона больших чисел не зависит от этих законов (которые должны подчиняться лишь некоторым требованиям весьма общего характера).

Этой же потребностью устанавливать закономерности общего типа, охватывающие собою возможно более широкий круг реальных явлений, была порождена в теории вероятностей и другая её важнейшая глава — теория предельных теорем. В очень большом числе случаев, в особенности в простейших, первично возникавших задачах, массовый характер изучаемого явления математически выражался в том, что исследовать приходилось суммы очень большого числа (более или менее равнoprвных между собою, зависимых или независимых) случайных величин. Так, в теории ошибок измерения (одна из первых областей приложения теории вероятностей) фактически получаемая ошибка обычно представляет собою сумму большого числа отдельных погрешностей, вызываемых самыми различными факторами. Уже закон больших чисел имеет дело именно с такими суммами большого числа случайных слагаемых. В XVIII столетии Муавром и Лапласом было установлено, что в некоторых простейших случаях суммы большого числа взаимно независимых случайных величин после надлежащей нормировки подчинялись законам распределения, которые при неограниченном увеличении числа слагаемых всё более и более приближаются к так называемому «нормальному» закону, «плотность» которого даётся функцией

\[
\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} x^2}.
\]

Это была первая из предельных теорем теории вероятностей — так называемая теорема Муавра-Лапласа, входящая в настоящее время во все курсы теории вероятностей. Эта теорема охватывала собою лишь весьма узкий класс случаев, так называемую схему Бернулли, когда каждое слагаемое имеет своими возможными значениями только числа 0 и 1,
а вероятности этих значений один и те же для всех слагаемых. Однако — и это было отмечено уже Лапласом — причины, в силу которых законы распределения сумм в случае схемы Бернулли имеют тенденцию приближаться к нормальному закону, носят характер настолько общий, что есть все основания предполагать теорему Муавра-Лапласа лишь частным случаем некоторой значительно более общей закономерности. Лаплас пытался обосновать это стремление к нормальному закону для более широкого класса случаев; однако ни ему, ни кому-либо из его современников не удалось сколько-нибудь заметно продвинуться в этом направлении — отчасти тому, что известные в ту эпоху методы математического анализа были ещё недостаточны для этой цели. Первый метод, позволивший доказать предельную теорему как некую общую закономерность в поведении сумм большого числа взаимно независимых случайных величин, был создан только в середине XIX столетия П. Л. Чебышевым — великим русским учёным, которому, как известно, наука обязана и первой общей концепцией закона больших чисел. Вообще, то общее всему точному естествознанию стремление к установлению закономерностей широкого значения, о котором мы говорили в начале настоящего параграфа, в теорию вероятностей вошло лишь после исследований Чебышева, да и тогда в течение ряда десятилетий оставалось главной движущей пружиной научных исследований только в России, лишь в двадцатом столетии заняв, наконец, ведущую роль в мировой науке.

Чебышев работал над методом обоснования общей предельной теоремы в продолжение почти всей своей научной жизни; этот метод был им создан, но завершить доказательства самой теоремы он не успел. Это было сделано вскоре после смерти Чебышева его учеником и последователем академиком А. А. Марковым. Однако за несколько лет до работы Маркова академик А. М. Ляпунов, также ученик и продолжатель дела Чебышева, в весьма общих предположениях доказал предельную теорему другим методом, более близким к современным приёмам её доказательства. Интересно отметить, что эти достижения русских учёных, основоположное значение которых в настоящее время полностью признаётся всем научным миром, в течение примерно двадцати лет оставались неизвестными зарубежной науке. На западе теория вероятностей
в эти годы ещё находилась в состоянии упадка; исследования русских учёных, посвящённые общим проблемам большого принципиального значения, не находили там никакого отклика; только этим можно объяснить то, что через двадцать лет после того, как Ляпунов доказал свою знаменитую теорему, она была вторично «открыта» на западе, и только после этого несомненный приоритет русских учёных был твёрдо установлен.

Центральная предельная теорема, доказанная впервые Ляпуновым и позднее уточнённая другими исследователями, утверждает, что при соблюдении некоторых условий общего характера законы распределения надлежащим образом нормированных сумм большого числа взвешенно независимых случайных величин должны быть близки к «нормальному» закону, приведённому нами выше. При этом речь идёт о так называемых «интегральных» законах: если обозначить через $S$ уплотнённую нами нормированную сумму, то вероятность неравенства $S < x$ должна быть близка к

$$
\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}u^2} du.
$$

Непосредственно очевидно, что в общем случае, когда нам неизвестен тип законов распределения самих слагаемых, иная постановка задачи и невозможна. Так, например, если слагаемые представляют собою целочисленные случайные величины, то и сумма их, очевидно, может принимать только целые значения. Ставить вопрос о предельном поведении плотности закона распределения суммы не имело бы в этом случае никакого смысла.

Однако, если в общем случае произвольно распределённых случайных слагаемых имеют смысл только предельные теоремы «интегрального» типа, то для специальных типов тех законов, которым подчиняются слагаемые данной суммы, могут иметь место и «локальные» предельные теоремы, дающие приближённое выражение либо вероятностей отдельных возможных значений этой суммы, либо плотности её закона распределения, в зависимости от того, подчиняются ли слагаемые дискретным (в частности, целочисленным) или непрерывным законом распределения. Локальные предельные теоремы того и другого типа оказались чрезвычайно нужными,
вследствие чего наука в последние десятилетия уделяла их обоснованию значительное внимание. В частности, для нашей цели — математического обоснования квантовой статистики — нужны исключительно локальные предельные теоремы (для случая целочисленных случайных величин). Мы коснёмся поэтому в дальнейшем только локальных предельных теорем для целочисленных случайных величин, оставляя в стороне (также в достаточной мере разработанный) случай непрерывных распределений.

Пусть мы имеем целочисленную случайную величину $\xi_i$, подчиняющуюся закону распределения $p_k$. Возможными значениями $\xi_i$ служат такие целые числа $k$, для которых $p_k > 0$. Совокупность всех попарных разностей таких возможных значений величины $\xi_i$ имеет наибольший общий делитель, который мы условимся обозначать через $d$; тогда, очевидно, если $a_0$ есть одно из возможных значений величины $\xi_i$, то любое его возможное значение может быть представлено в виде $a_0 + ld$, где $l$ — некоторое целое число. Разумеется, обратное, вообще говоря, неверно: число $a_0 + ld$ не при любом целом $l$ должно быть возможным значением величины $\xi_i$; однако те числа $l$, для которых $a_0 + ld$ есть такое возможное значение, очевидно, образуют (конечную или бесконечную) совокупность целых чисел с наибольшим общим делителем 1.

Пусть теперь мы имеем $n$ взаимно независимых случайных величин $\xi_1, \xi_2, \ldots, \xi_n$, каждая из которых подчиняется только что описанному закону. Тогда сумма $s_n$ этих величин может, очевидно, принимать только значения вида $na_0 + ld$, где $l$ — целое число. Задача локальной предельной теоремы в этом случае состоит, таким образом, в том, чтобы найти удобное приближённое выражение для вероятности

$$P(s_n = na_0 + ld)$$

при различных значениях $l$.

Современная теория вероятностей распространяет предельные теоремы (как интегрального, так и локального типа) на суммы случайных векторов, стремясь установить, что при некоторых ограничениях общего характера законы распределения таких сумм после надлежащей нормировки приближаются к нормальному закону соответствующего числа измерений, когда число слагаемых безгранично возрастает.
Пусть, например, \((\xi_i, \eta_i) \ (i = 1, 2, \ldots, n)\) — взаимно независимые целочисленные случайные векторы, подчиняющиеся одному и тому же закону

\[
p_{ab} = P(\xi_i = a, \eta_i = b).
\]

Каждая пара целых чисел \((a, b)\), для которой \(p_{ab} > 0\), изображается некоторой точкой плоскости, имеющей целочисленные декартовы координаты \(a, b\) (такие точки мы для краткости будем просто называть «целыми» точками). Эти «возможные» точки вектора \((\xi_i, \eta_i)\) образуют на плоскости некоторое множество целых точек \(M\). Это множество составляет часть «решётки» всех целых точек плоскости. Вообще говоря, возможно (и в нужных нам приложениях нередко встречается), что существуют на плоскости и другие, более редкие параллелограммографические решётки целых точек, охватывающие собою множество \(M\) возможных точек нашего вектора. Совокупность точек любой параллелограммографической решётки плоскости может быть представлена в виде

\[
\begin{align*}
  x &= a_0 + ka + lb, \\
  y &= b_0 + k\gamma + l\delta,
\end{align*}
\]

где \(a_0, b_0, a, \beta, \gamma, \delta\) — постоянные целые числа, \(d = ad - \beta\gamma \neq 0\), а \(k\) и \(l\) пробегают всевозможные целые числа.

Как известно, основным параллелограммом такой решётки называется всякий параллелограмм, вершины которого принадлежат этой решётке, но других точек решётки ни внутри, ни на границе его не имеется. Данная решётка может иметь основные параллелограммы различной формы, но площадь их всегда одна и та же и является поэтому основной характеристикой самой решётки; эта площадь равна, как легко подсчитать, \(|d| = |ad - \beta\gamma|\). Решётка, охватывающая множество \(M\), называется максимальной, если площадь её основного параллелограмма имеет наибольшее возможное значение. Эта максимальная площадь основного параллелограмма играет в теории случайных двумерных векторов такую же роль, какая свойственна введённому нами выше числу \(d\) в одномерном случае.

Мы будем называть целочисленный случайный вектор \((\xi, \eta)\) вырожденным, если все его возможные точки лежат на одной прямой, т. е. если между \(\xi\) и \(\eta\) имеется линейная
функциональная зависимость. Покажем теперь, что всякий невырожденный вектор имеет максимальную решётку. В самом деле, если вектор \((\xi, \eta)\) невырожден, то он имеет по меньшей мере три возможные точки, не лежащие на одной прямой; добавим к ним четвёртую точку так, чтобы получить параллелограмм \(P\), и пусть площадь \(P\) равна \(s\). Очевидно, что \(P\) будет служить параллелограммом любой решётки, охватывающей множество \(M\); следовательно, основной параллелограмм такой решётки не может иметь площади, большей, чем \(s\); отсюда непосредственно вытекает, что среди таких охватывающих решёток найдётся максимальная *).

Если закон распределения случайного вектора \((\xi_i, \eta_i)\) \((i = 1, 2, \ldots, n)\) имеет охватывающую решётку (5) и если мы положим

\[
\sum_{i=1}^{n} \xi_i = S_n \quad \sum_{i=1}^{n} \eta_i = T_n,
\]

tо для вектора \((S_n, T_n)\) возможными точками могут служить, очевидно, лишь точки \((x, y)\) вида

\[
x = na_0 + ka + l\beta, \\
y = nb_0 + k\gamma + l\delta,
\]

где \(k\) и \(l\) — целые числа. Задача локальной предельной теоремы будет здесь состоять в том, чтобы найти удобное приближённое выражение для вероятности

\[
P(S_n = na_0 + ka + l\beta, \quad T_n = nb_0 + k\gamma + l\delta)
\]

при заданных целых \(k\) и \(l\).

Вся дальнейшая часть главы будет посвящена установлению предельных теорем подобного рода. При этом, кроме отыскания упомянутых выше приближённых выражений, мы должны будем во всех случаях, в интересах дальнейших приложений, подвергать довольно точной оценке погрешность,

*) Подробнее об элементарных свойствах точечных решёток см. 1) Б. Н. Делоне, Н. Н. Падуров, А. Д. Александров, Математические основы структурного анализа кристаллов, ГТТИ, 1934, стр. 32—34; 2) Б. Н. Делоне, Геометрия положительных квадратичных форм, Успехи математических наук, вып. III, 1937, стр. 21—22, 44—45; 3) Д. Гильберт и С. Э. Кон-Фоссен, Наглядная геометрия, ОНТИ, 1936, стр. 36—39.
возникающую при замене искомой вероятности установленным нами асимптотическим выражением.

В заключение настоящего параграфа мы приведём ещё краткую историческую справку о дальнейшем развитии учения о предельных теоремах после первых открытий Ляпунова и Маркова. Развитие это продолжалось (и продолжается до сих пор) в следующих основных направлениях: 1) распространение предельных теорем на многомерные случайные величины (случайные векторы); 2) установление локальных предельных теорем (непрерывного и дискретного типа); 3) распространение предельных теорем на суммы взаимно зависимых случайных величин; 4) изыскание наиболее широких (необходимых и достаточных) условий применимости предельных теорем; 5) уточнение оценки остаточных членов. При этом перечисленных мы ограничиваемся проблематикой, в которой в качестве предельных фигурируют только нормальные законы; вопрос о том, какие иные законы и при каких условиях могут служить предельными для сумм большого числа случайных величин, вырос за последние десятилетия в особую широко развитую теорию, которой мы здесь совсем не можем касаться *

В решении большинства перечисленных нами задач ведущая роль неизменно принадлежала советской вероятностной школе, главным образом в лице С. Н. Бернштейна, Б. В. Гнеденко, А. Н. Колмогорова, И. Г. Петровского, Н. В. Смирнова и А. Я. Хинчина.

§ 3. Метод характеристикских функций

В настоящее время наиболее удобным методом доказательства предельных теорем теории вероятностей, в особенности для случая взаимной независимости слагаемых, служит так называемый метод характеристикских функций. Мы изложим теперь основы этого метода применительно к интересующему нас случаю целочисленных случайных величин.

*) Читателя, интересующегося проблематикой этого рода, мы можем отослать к недавно вышедшей книге Б. В. Гнеденко и А. Н. Колмогорова, Предельные распределения для сумм независимых случайных величин, Гостехиздат, 1949.
Пусть $\xi$ есть целочисленная случайная величина, подчинённая закону

$$p_n = P(\xi = n) \quad (-\infty < n < +\infty).$$

Если $f(\xi)$ — произвольная (вещественная или комплексная) функция величины $\xi$, и если ряд

$$\sum_{n=-\infty}^{+\infty} f(n) p_n = E f(\xi)$$

абсолютно сходится, то сумму этого ряда мы в § 1 условились называть математическим ожиданием случайной величины $f(\xi)$. Положим теперь, в частности,

$$f(\xi) = e^{it\xi},$$

где $i$ — мнимая единица, а $t$ — произвольное вещественное число. Так как ряд

$$\sum_{n=-\infty}^{+\infty} e^{int} p_n,$$

очевидно, абсолютно сходится при любом вещественном $t$, то математическое ожидание

$$E e^{it\xi} = \sum_{n=-\infty}^{+\infty} p_n e^{itn}.$$

комплексной случайной величины $e^{it\xi}$ существует при любом $t \geq 0$ и представляет собою функцию вещественного параметра $t$. Эту функцию мы будем в дальнейшем обозначать через $\varphi(t)$ и называть характеристической функцией величины $\xi$ или закона распределения $p_n$; таким образом,

$$\varphi(t) = \sum_{n=-\infty}^{+\infty} p_n e^{itn}. \quad (6)$$

Отметим теперь некоторые простейшие общие свойства характеристических функций.

1°. Очевидно, мы всегда имеем:

$$\varphi(0) = \sum_{n=-\infty}^{+\infty} p_n = 1,$$
и при любом вещественном $t$

$$| \varphi(t) | \leq \sum_{n=-\infty}^{+\infty} p_n = 1.$$  

2°. Если величина $\xi$ имеет математическое ожидание $E \xi$, то оно представляется рядом

$$\sum_{n=-\infty}^{+\infty} np_n,$$

который в этом случае абсолютно сходится; но в таком случае ряд

$$\sum_{n=-\infty}^{+\infty} inp_ne^{int},$$

получаемый из ряда (6) почленным дифференцированием по $t$, сходится, очевидно, абсолютно и равномерно на всей числовой прямой. Это показывает, что из существования математического ожидания $E \xi$ величины $\xi$ вытекает дифференцируемость характеристической функции $\varphi(t)$ при любом $t$, а также соотношение

$$\varphi'(0) = iE \xi.$$  

3°. Совершенно аналогичным образом мы легко убедимся, что существование у величины $\xi$ момента порядка $k$ (где $k$ — любое натуральное число) влечёт за собой $k$-кратную дифференцируемость функции $\varphi(t)$, а также соотношение

$$\varphi^{(k)}(0) = i^k E \xi^k.$$  

4°. Непосредственно из равенства (6) вытекает, что характеристическая функция целочисленной случайной величины всегда есть периодическая функция и что периодом её (не обязательно наименьшим) всегда служит число $2\pi$.

5°. Пусть $\xi'$ и $\xi''$ — две взаимно независимые целочисленные случайные величины, имеющие соответственно законы распределения $p'_n$ и $p''_n$ и характеристические функции $\varphi_1(t)$ и $\varphi_2(t)$, так что

$$\varphi_1(t) = \sum_{n=-\infty}^{+\infty} p'_ne^{int}, \quad \varphi_2(t) = \sum_{n=-\infty}^{+\infty} p''_ne^{int}.$$
Пусть \( p_n \) и \( \varphi(t) \) соответственно означают закон распределения и характеристическую функцию суммы \( \xi = \xi' + \xi'' \); тогда, как мы видели в § 1,

\[
p_n = \sum_{k=-\infty}^{+\infty} p_k p'_n - k,
\]

и, следовательно,

\[
\varphi(t) = \sum_{n=-\infty}^{+\infty} p_n e^{itn} = \sum_{n=-\infty}^{+\infty} e^{itn} \sum_{k=-\infty}^{+\infty} p_k p'_n - k = \]

\[
= \sum_{n=-\infty}^{+\infty} (\sum_{k=-\infty}^{+\infty} p_k e^{itk} p'_n - k e^{it(n-k)}) = \]

\[
= \sum_{k=-\infty}^{+\infty} p_k e^{itk} \sum_{n=-\infty}^{+\infty} p'_n - k e^{it(n-k)} = \]

\[
= \sum_{k=-\infty}^{+\infty} p_k e^{itk} \sum_{l=-\infty}^{+\infty} p'_l e^{itl} = \varphi_1(t) \varphi_2(t)^*.
\]

Таким образом, при сложении взаимно независимых случайных величин характеристические функции их просто перемножаются. Это «правило композиции» характеристических функций, доказанное нами для сложения двух случайных величин, с помощью индукции, очевидно, непосредственно распространяется на случай любого числа (взаимно независимых) слагаемых.

Эта исключительная простота правила композиции характеристических функций и делает их столь удовлетворительным инструментом для исследования сумм большого числа взаимно независимых случайных величин и, в частности, для установления предельных теорем. В то время как правило композиции самих законов распределения, даваемое формулами (1) или (2) § 1, по своей сложности, в особенности при большом числе слагаемых, совершенно необязательно, для характеристических функций правило это, как мы видим, отличается чрезвычайно.

*) Если считать установленной теорему умножения математических ожиданий для взаимно независимых случайных величин, то полученный нами результат следует просто из того, что

\[
\varphi(t) = E e^{it\xi} = E e^{it(\xi' + \xi'')} = E (e^{it\xi'} e^{it\xi''}) = \]

\[
= E e^{it\xi'} E e^{it\xi''} = \varphi_1(t) \varphi_2(t).
\]
ной простотой, так что, зная характеристические функции слагаемых, мы немедленно и непосредственно создаём себе представление о характеристической функции суммы.

Однако для того, чтобы характеристические функции могли стать полноценным орудием исследования сумм большого числа случайных величин, одного только простого правила композиции ещё недостаточно; найдя характеристическую функцию исследуемой суммы, мы должны иметь возможность установить с её помощью и закон распределения этой суммы. Между тем, до сих пор мы связывали закон распределения с соответствующей характеристической функцией только при помощи формулы (6), выражающей \( \varphi(t) \) через \( p_n \); мы не только не имеем обратного выражения \( p_n \) через \( \varphi(t) \), но даже в сущности не знаем, определяет ли характеристическая функция \( \varphi(t) \) однозначно соответствующий ей закон распределения \( p_n \). Все эти вопросы, как мы теперь покажем, просто разрешаются классическими формулами Фурье.

6°. Умножим обе части равенства (6) на \( e^{-imt} \), где \( m \) — произвольное целое число, и проинтегрируем полученное соотношение по \( t \) в пределах от \(-\pi\) до \( \pi\). При этом ряд в правой части, будучи равномерно сходящимся в указанных пределах, допускает почленную интеграцию, и мы находим:

\[
\int_{-\pi}^{+\pi} e^{-imt} \varphi(t) \, dt = \sum_{n=-\infty}^{+\infty} p_n \int_{-\pi}^{+\pi} e^{i(n-m)t} \, dt,
\]

или, так как в правой части интеграл равен, очевидно, \( 2\pi \) при \( n = m \) и нулю при \( n \neq m \),

\[
\int_{-\pi}^{+\pi} e^{-imt} \varphi(t) \, dt = 2\pi p_m,
\]

откуда

\[
p_m = \frac{1}{2\pi} \int_{-\pi}^{+\pi} e^{-imt} \varphi(t) \, dt. \tag{7}
\]

Эта «формула обращения» показывает, что закон распределения \( p_m \) целочисленной случайной величины однозначно определяется заданием её характеристической функции \( \varphi(t) \). Вместе с тем формула (7) даёт и весьма простое выражение этой зависимости.
70. В § 1 мы записывали всю совокупность возможных значений случайной величины \( \xi \) в виде \( a_0 + ld \), где \( d \) — наибольший общий делитель попарных разностей этих значений, \( a_0 \) — одно из возможных значений и \( l \) — целое число. В дальнейшем мы условимся называть число \( d \), однозначно определяемое законом распределения величины \( \xi \), шагом этой величины. Если мы попрежнему будем обозначать закон распределения величины \( \xi \) через \( p_n \), то, очевидно, мы можем иметь \( p_n > 0 \) только при условии, что

\[
n \equiv a_0 \pmod{d},
\]

t. е. что число \( n \) имеет форму \( a_0 + ld \). Поэтому нам будет теперь удобнее, в отличие от предыдущего, обозначать через \( p_l \) вероятность равенства \( \xi = a_0 + ld \). При таком обозначении характеристическая функция величины \( \xi \) выражается, очевидно, формулой

\[
\varphi (t) = \sum_{l = -\infty}^{+\infty} p_l e^{it(a_0 + ld)} t = e^{ita_0} t \sum_{l = -\infty}^{+\infty} p_l e^{itld} dt.
\]

Мы непосредственно видим, что функция

\[
\varphi (t) e^{-ita_0} = \sum_{l = -\infty}^{+\infty} e^{itld} p_l
\]

имеет период \( \frac{2\pi}{d} \); вычисление, аналогичное проведённому нами в п. 60, легко показывает, что мы имеем:

\[
p_l = \frac{d}{2\pi} \int_{-\frac{\pi}{d}}^{\frac{\pi}{d}} e^{-it(a_0 + ld)} t \varphi (t) dt. \tag{8}
\]

Важное значение введённого нами понятия шага \( d \) целочисленной случайной величины обусловлено следующей леммой:

Лемма 1. Пусть \( \xi \) — целочисленная случайная величина, имеющая шаг \( d \) *) и характеристическую функцию \( \varphi (t) \). Тогда \( |\varphi (t)| < 1 \) при \( 0 < |t| \leq \frac{\pi}{d} \).

*) Наличие шага означает, что случайная величина \( \xi \) не вырождена, т. е. имеет не менее двух возможных значений.
В самом деле, если при каком-либо значении \( t \) мы имеем
\[
|\varphi(t)| = 1,
\]
то и
\[
\left| \sum_{\ell = -\infty}^{+\infty} p_{\ell} e^{it\ell t} \right| = |\varphi(t) e^{-ia_{0}t}| = 1.
\]

В силу \( \sum_{\ell = -\infty}^{+\infty} p_{\ell} = 1 \) это возможно лишь тогда, если все члены суммы, в которых \( p_{\ell} > 0 \), имеют один и тот же аргумент \( \theta \), т. е. в случае \( p_{\ell} > 0 \)
\[
l \eta t = \theta + 2\pi u_{\ell},
\]
где \( u_{\ell} \) — целое число. Пусть \( p_{\nu} > 0 \) и \( p_{\nu'} > 0 \); тогда
\[
(l'' - l') d = (u_{\nu'} - u_{\nu}) \frac{2\pi}{t}.
\]

Но \( (l'' - l') d = (a_{0} + l''d) - (a_{0} + l'd) \) есть разность любых двух возможных значений величины \( \xi \); таким образом, все такие разности — целые кратные числа \( \frac{2\pi}{t} \), а значит, и наибольший общий делитель этих разностей, который по определению равен \( d \), есть целое кратное числа \( \frac{2\pi}{t} \),
\[
d = s \cdot \frac{2\pi}{t},
\]
где \( s \) — целое число; отсюда
\[
t = \frac{2\pi s}{d}.
\]
Так как область \( 0 < |t| \leq \frac{\pi}{d} \) таких значений \( t \) не содержит, то в этой области \( |\varphi(t)| < 1 \), что и требовалось доказать.
§ 3  МЕТОД ХАРАКТЕРISTICИЧЕСКИХ ФУНКЦИЙ

Метод характеристических функций естественным образом распространяется и на многомерные случайные величины (случайные векторы). Пусть \((\xi, \eta)\) — целочисленный случайный вектор, подчинённый закону распределения

\[ p_{ab} = P(\xi = a, \eta = b). \]

Математическое ожидание величины \(e^{it\xi + is\eta}\), равное

\[ E e^{i(t\xi + s\eta)} = \sum_{a, b} e^{i(ta + sb)} p_{ab}, \]

где \(t\) и \(s\) — вещественные переменные, мы будем называть характеристической функцией вектора \((\xi, \eta)\) или закона \(p_{ab}\), и обозначать через \(\varphi(t, s)\), так что

\[ \varphi(t, s) = \sum_{a, b} p_{ab} e^{i(ta + sb)}. \]

Как в одномерном случае, функция \(\varphi(t, s)\) — периодическая с периодом \(2\pi\) относительно каждой из двух переменных; \(\varphi(0, 0) = 1\), и при любых \(t, s\) \(|\varphi(t, s)| \leq 1\).

Далее, если существует момент \(E(\xi^h \eta^k)\), то существует и частная производная

\[ \frac{\partial^{h+k} \varphi}{\partial t^h \partial s^k}, \]

и значение этой производной при \(t = s = 0\) равно \(i^{h+k} E(\xi^h \eta^k)\).

Если \((\xi_1, \eta_1)\) и \((\xi_2, \eta_2)\) — два взаимно независимых целочисленных случайных векторов с характеристиками функциями \(\varphi_1(t, s)\) и \(\varphi_2(t, s)\) и если \(\varphi(t, s)\) — характеристическая функция вектора \((\xi_1 + \xi_2, \eta_1 + \eta_2)\), то

\[ \varphi(t, s) = \varphi_1(t, s) \varphi_2(t, s); \]

это правило композиции остаётся верным и при сложении любого числа взаимно независимых случайных векторов.

Далее, имеет место «формула обращения»

\[ p_{ab} = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{-i(ta + sb)} \varphi(t, s) \, dt \, ds. \]

Все эти результаты устанавливаются в полной аналогии с одномерным случаем и с той же лёгкостью.
Пусть теперь закон \( p_{ab} \) имеет максимальную решётку

\[
\begin{align*}
a &= a_0 + k\alpha + l\beta, \\
b &= b_0 + k\gamma + l\delta,
\end{align*}
\]

где \( d = a\delta - \beta\gamma \neq 0 \). Тогда \( p_{ab} \) может быть отличным от нуля лишь при условии, что числа \( a \) и \( b \) имеют форму (9) при некоторых целых \( k \) и \( l \). Поэтому мы и здесь, как в одномерном случае, несколько изменим обозначения, полагая

\[
p_{kl} = \mathcal{P}(\xi = a_0 + k\alpha + l\beta, \eta = b_0 + k\gamma + l\delta);
\]

для характеристической функции вектора \((\xi, \eta)\) мы таким образом получаем выражение

\[
\varphi(t, s) = \sum_{k, l = -\infty}^{+\infty} p_{kl} e^{i[t(a_0 + k\alpha + l\beta) + s(b_0 + k\gamma + l\delta)]} = \sum_{k, l = -\infty}^{+\infty} p_{kl} e^{i[t(k\alpha + l\beta) + s(k\gamma + l\delta)]}.
\]

Наша ближайшая задача — найти «формулу обращения», выражающую закон \( p_{kl} \) через функцию \( \varphi(t, s) \), аналогичную формуле (8) одномерного случая. Возьмём с этой целью произвольную пару целых чисел \((k', l')\), умножим обе части равенства (10) на

\[
e^{-i[t(a_0 + k'\alpha + l'\beta) + s(b_0 + k'\gamma + l'\delta)]}
\]

и возьмём интегралы обеих частей, распространяя их на область \( D \) плоскости \((t, s)\), характеризуемую неравенствами

\[
\begin{align*}
-\pi &\leq at + \gamma s \leq \pi, \\
-\pi &\leq \beta t + \delta s \leq \pi.
\end{align*}
\]

и представляющую собой, очевидно, параллелограмм площади \( \frac{4\pi^2}{d} \) с центром в начале координат. Мы получаем:

\[
\begin{align*}
\int_{D} \int_{D} e^{-i[t(a_0 + k'\alpha + l'\beta) + s(b_0 + k'\gamma + l'\delta)]} \varphi(t, s) \, dt \, ds &= \\
&= \sum_{k, l = -\infty}^{+\infty} p_{kl} \int_{D} \int_{D} e^{i \left\{ \left[ (k-k')\alpha + (l-l')\beta \right] t + \left[ (k-k')\gamma + (l-l')\delta \right] s \right\}} \, dt \, ds.
\end{align*}
\]
В правой части этого равенства интеграл, соответствующий значениям \( k = k', \ l = l' \), очевидно, равен площади области \( D \), т. е. \( \frac{4\pi^2}{d} \). Убедимся теперь, что все остальные интегралы правой части обращаются в нуль. Возьмём какой-либо из этих интегралов и преобразуем в нём переменные, полагая

\[
\alpha t + \gamma s = u, \\
\beta t + \delta s = v.
\]

Тогда интеграл получает вид

\[
\iint_{-\pi}^{\pi} \iint_{-\pi}^{\pi} e^{i \left( (k-k')u + (l-l')v \right)} \frac{du\,dv}{d} =
\]

\[
= \frac{1}{d} \int_{-\pi}^{\pi} e^{i (k-k')u} \, du \int_{-\pi}^{\pi} e^{i (l-l')v} \, dv,
\]

и, очевидно, обращается в нуль, если по крайней мере одна из двух разностей \( k - k' \), \( l - l' \) отлична от нуля. Таким образом, мы находим:

\[
\iint_{D} e^{-i \left[ t (a_0 + k'\alpha + l'\beta) + s (b_0 + k'\gamma + l'\delta) \right]} \varphi(t, s) \, dt \, ds = \frac{4\pi^2}{d} p_{kl},
\]

откуда, опуская теперь штрихи, получаем для любой пары целых чисел \( k, l \):

\[
p_{kl} = \frac{d}{4\pi^2} \iint_{D} e^{-i \left[ t (a_0 + k\alpha + l\beta) + s (b_0 + k\gamma + l\delta) \right]} \varphi(t, s) \, dt \, ds.
\]

Это и есть нужная нам формула обращения.

Наконец, для установления двумерной предельной теоремы имеет существенное значение предложение, аналогичное вышеустановленной лемме 1: если решётка (9) максимална, то, за исключением точки \( t = s = 0 \), во всюду внутри и на границе области \( D \) мы имеем \( |\varphi(t, s)| < 1 \).

В самом деле, пусть \( (t, s) \) — любая точка плоскости, в которой \( |\varphi(t, s)| = 1 \). Это в силу формулы (10) даёт:

\[
\sum_{k, l}^{+\infty} p_{kl} e^{i \left[ t (a_0 + k\alpha + l\beta) + s (b_0 + k\gamma + l\delta) \right]} = 1;
\]

\[
(12)
\]
а так как и

\[ +\infty \sum_{k, l = -\infty} p_{kl} = 1, \]

то в сумме (12) все члены, в которых \( p_{kl} > 0 \), должны иметь один и тот же аргумент (с точностью до кратных \( 2\pi \)); таким образом, если \( p_{kl} > 0 \), то

\[ (a_0 + ka + l\beta) t + (b_0 + kb + l\delta) s = \theta + 2m\pi, \]

где \( \theta \) — постоянная, а \( m \) — целое число (зависящее, конечно, от \( k \) и \( l \)). Таким образом, любая возможная точка \((a, b)\) вектора \((\xi, \eta)\) удовлетворяет соотношению

\[ at + bs = \theta + 2m\pi, \]

где \( m \) — целое число.

Рассмотрим теперь точки плоскости \((t, s)\)

\[ t_1 = \frac{2\pi}{|d|} \delta, \quad s_1 = -\frac{2\pi}{|d|} \beta \]

и

\[ t_2 = -\frac{2\pi}{|d|} \gamma, \quad s_2 = \frac{2\pi}{|d|} \alpha, \]

очевидно, не лежащие на одной прямой с началом координат. Легко непосредственно проверить, что

\[ |\varphi(t_1, s_1)| = |\varphi(t_2, s_2)| = 1. \]

Пусть теперь мы имеем \( |\varphi(t, s)| = 1 \) в некоторой точке \((t, s)\), отличной от начала координат; как мы только что установили, любая возможная точка вектора \((\xi, \eta)\) лежит тогда на одной из систем параллельных прямых

\[ xt + ys = \theta + 2m\pi, \]

где \( m \) — целое число. Так как точки \((t_1, s_1)\) и \((t_2, s_2)\) не лежат на одной прямой с началом, то прямая, проходящая через начало и точку \((t, s)\), не может содержать обеих точек \((t_1, s_1)\) и \((t_2, s_2)\); для определённости пусть точка \((t_1, s_1)\) лежит вне этой прямой; тогда \( t_1 s - s_1 t \neq 0 \). Так как \( |\varphi(t_1, s_1)| = 1 \), то все возможные точки вектора \((\xi, \eta)\) принадлежат, кроме системы (13), и системе прямых

\[ xt_1 + ys_1 = \theta_1 + 2\pi n, \]
где \( n \) — целое число. Системы прямых (13) и (14) образуют своими точками пересечения решётку, которая, как мы показали, охватывает множество \( M \). Основной параллелограмм этой решётки имеет, как легко подсчитать, площадь

\[
\frac{4\pi^2}{|s_1t - t_1s|}.
\]

Если точка \((t, s)\) лежит, как мы теперь допустим, внутри или на границе области \( D \), то по определению этой области для неё \(|\beta t + \delta s| \leq \pi\), а потому

\[
|s_1t - t_1s| = \frac{2\pi}{|d|} |\beta t + \delta s| \leq \frac{2\pi^2}{|d|}.
\]

Но в таком случае площадь (15) основного параллелограмма построенной нами решётки должна быть \( \geq 2|d| \); это же невозможно, так как эта решётка охватывает множество \( M \), для которого \(|d|\) есть площадь основного параллелограмма максимальной решётки (9). Итак, мы доказали следующее предложение:

Лемма 2. Пусть все возможные точки \((a, b)\) случайного вектора \((\xi, \eta)\) охватываются максимальной решёткой

\[
a = a_0 + ka + lb,
b = b_0 + k\gamma + l\delta.
\]

Тогда \(|\varphi(t, s)| < 1\) всюду внутри и на границе области \( D \), определяемой неравенствами

\[
|\alpha t + \gamma s| \leq \pi,
|\beta t + \delta s| \leq \pi,
\]

за исключением точки \( t = s = 0 \).

§ 4. Одномерная предельная теорема

Пусть \( \xi_1, \xi_2, \ldots, \xi_n \) — взаимно независимые целочисленные случайные величины, подчинённые одному и тому же закону распределения

\[
p_t = P(\xi_i = a_0 + ld) \quad (-\infty < l < +\infty),
\]

где \( a_0 \) — одно из возможных значений, а \( d \) — шаг величины \( \xi_t \). Положим

\[
\xi_1 + \xi_2 + \ldots + \xi_n = s_n;
\]
случайная величина $s_n$ может, очевидно, принимать только значения вида $na_0 + ld$ (где $l$ — целое число); положим для краткости

$$P(s_n = na_0 + ld) = P_n(l).$$

Пусть характеристическая функция величины $\xi_i$ есть

$$\varphi(t) = \sum_{l=-\infty}^{+\infty} p_t e^{it(a_0 + ld)};$$

tогда, согласно правилу композиции, характеристическая функция величины $s_n$ есть $\{\varphi(t)\}^n$, что мы будем писать короче в виде $\varphi^n(t)$. Таким образом,

$$\sum_{l=-\infty}^{+\infty} P_n(l) e^{it(na_0 + ld)} = \varphi^n(t).$$

Формула обращения (8) § 3 отсюда даёт:

$$P_n(l) = \frac{d}{2\pi} \int_{-\pi/d}^{\pi/d} e^{-it(na_0 + ld)} \varphi^n(t) \, dt \quad (-\infty < l < +\infty). \quad (16)$$

Нашей целью является отыскание удобных приближённых выражений закона распределения $P_n(l)$, с оценкой соответствующей погрешности; формула (16) будет нам при этом служить отправным пунктом.

Мы допустим, что закон распределения $p_l$ величины $\xi_1$ имеет конечные моменты до пятого порядка включительно, т. е. что сходится ряд

$$\sum_{l=-\infty}^{+\infty} |l|^5 p_l;$$

как мы видели в § 3, это влечёт за собою существование производных до пятого порядка включительно у функции $\varphi(t)$; в частности, существует математическое ожидание

$$E\xi_i = \sum_{l=-\infty}^{+\infty} (a_0 + ld) p_l = -i\varphi'(0) = a.$$
§ 4] ОДНОМЕРНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА

и дисперсия

$$D_{\xi_i} = E_{\xi_i^2} - (E_{\xi_i})^2 = \sum_{l = -\infty}^{+\infty} (a_0 + ld - a)^2 p_l =$$

$$= - \varphi''(0) + [\varphi'(0)]^2 = b$$

величины $\xi_i$.

Положим $na_0 + ld - na = u$ (при данном $l$ величина $u$ представляет собой уклонение значения $na_0 + ld$ случайной величины $s_n$ от её математического ожидания, очевидно, равного $na$). Целью настоящего параграфа является показать, что при $n \to \infty$

$$P_n(l) = \frac{d}{\sqrt{2\pi nb}} e^{-\frac{u^2}{2nb}} + \frac{m_0 + m_1u}{n^{\delta}} + O\left(\frac{\sqrt{n + |u|}}{n^{\delta}}\right), \quad (17)$$

где $m_0$ и $m_1$ — постоянные (не зависящие от $n$ и $u$) числа. Формула (17) и есть нужный нам вид одномерной локальной предельной теоремы. В случае не слишком больших уклонений $u$ эта формула даёт весьма точную и в то же время очень простую приближённую оценку вероятности $P_n(l)$.

Для доказательства формулы (17) нам будет удобно предварительно установить одно вспомогательное предложение, представляющее собой уточнение леммы 1 § 3.

Лемма 3. Для $-\frac{\pi}{d} \leq t \leq \frac{\pi}{d}$ мы имеем:

$$|\varphi(t)| \leq e^{-ct^2},$$

где $c$ — постоянное положительное число.

Доказательство. В силу наших предпосылок о существовании моментов величины $\xi_i$ функция $\varphi(t)$ в окрестности нуля может быть представлена в виде

$$\varphi(t) = 1 + iat - \frac{1}{2}qt^2 + O(|t|^3),$$

где $a = E\xi$, $q = E\xi^2$. Отсюда

$$|\varphi(t)|^2 = \left(1 - \frac{1}{2}qt^2\right)^2 + a^2t^2 + O(|t|^3) =$$

$$= 1 - t^2(q - a^2) + O(|t|^3) =$$

$$= 1 - bt^2 + O(|t|^3),$$
где \( b = q - a^2 > 0 \) — дисперсия величины \( \xi_i \); отсюда

\[
| \varphi(t) |^2 \leq 1 - b t^2 + c_1 |t|^2,
\]

где \( c_1 \) — положительная постоянная. Положим \( \frac{b}{2c_1} = \delta \); при \( |t| < \delta \) мы получаем:

\[
| \varphi(t) |^2 \leq 1 - b t^2 + c_1 \delta t^2 = 1 - \frac{b}{2} t^2 \leq e^{-\frac{b}{2} t^2},
\]

\[
| \varphi(t) | \leq e^{-\frac{b}{4} t^2}.
\]  

(18)

Но при \( \delta \leq |t| \leq \frac{\pi}{d} \) мы в силу леммы 1 имеем \( | \varphi(t) | < 1 \); а так как функция \( | \varphi(t) | \) непрерывна, то найдётся такое \( \mu < 1 \), что \( | \varphi(t) | < \mu \) при \( \delta \leq |t| \leq \frac{\pi}{d} \); определяя положительное число \( c_2 \) из равенства

\[
\mu = e^{-c_2 \left( \frac{\pi}{d} \right)^2},
\]

мы при \( \delta \leq |t| \leq \frac{\pi}{d} \) будем иметь:

\[
| \varphi(t) | < \mu = e^{-c_2 \left( \frac{\pi}{d} \right)^2} < e^{-c t^2}.
\]  

(19)

Если обозначить теперь через \( c \) наименьшее из чисел \( \frac{b}{4} \) и \( c_2 \), то в силу (18) и (19) мы будем иметь уже во всём отрезке \( \left( -\frac{\pi}{d}, +\frac{\pi}{d} \right) \):

\[
| \varphi(t) | \leq e^{-c t^2}.
\]

Этим лемма 3 доказана.

Теперь мы приступим к доказательству формулы (17). Для более лёгкой обозримости мы разобьём это доказательство на отдельные этапы.

1. Выберем произвольное число \( \lambda, \frac{3}{7} < \lambda < \frac{1}{2} \), которое в дальнейшем будем считать постоянным, и положим \( n^{-\lambda} = \delta \), так что величина \( \delta \) бесконечно мала при \( n \rightarrow \infty \).
Так как при постоянных \( r \geq 0, \ a > 0 \) и при \( n \to \infty \)
\[
\int_{b}^{\infty} t^{r}e^{-\lambda t}dt = \frac{1}{r+1} \int_{a}^{\infty} u^{r}e^{-\frac{u^2}{2}}du < \]
\[
\frac{1}{r+1} e^{-\frac{1}{2} \lambda n^{1-2\lambda}} \int_{0}^{\infty} u^{r}e^{-\frac{u^2}{2}}du = o(n^{-\gamma}), \quad (20)
\]
где \( \gamma \) — любое постоянное положительное число (поскольку в промежутке интегриации \( e^{-u^2} = e^{-\frac{u^2}{2}} \cdot \frac{e^{-\frac{u^2}{2}}}{e^{-\frac{u^2}{2}}} \), то в силу леммы 3
\[
\left| \int_{b}^{\infty} e^{-it(na_{0}+ld)} \varphi^{n}(t)dt \right| \leq \int_{b}^{\infty} |\varphi(t)|^{n}dt \leq \]
\[
\leq 2 \int_{b}^{\infty} e^{-ntc_{0}^{2}}dt = o(n^{-\gamma}),
\]
а потому в силу (16)
\[
P_{n}(l) = \frac{d}{2\pi} \int_{-\gamma}^{\gamma} e^{-it(na_{0}+ld)} \varphi^{n}(t)dt + o(n^{-\gamma})
\]
\[
(-\infty < l < +\infty). \quad (21)
\]

2. Формула (21) позволяет нам в дальнейшем ограничиться рассмотрением значений \( t \), заключённых между \( -\delta \) и \( +\delta \). В частности, произведение \( n|t|^{3} \leq n^{1-3\lambda} \) мы в силу \( \lambda > \frac{1}{3} \) можем всегда считать бесконечно малым.

Так как функция \( \varphi(t) \), согласно нашему предположению, при \( t = 0 \) имеет производную пятого порядка, то
\[
\varphi(t) = 1 + itE_{1} - \frac{t^{2}}{2}E_{2} + i\frac{t^{3}}{6}E_{3} + \frac{t^{4}}{24}E_{4} + O(|t|^{5}),
\]
\[
\ln \varphi(t) = ita - \frac{t^{2}}{2}b + ic_{2}t^{3} + c_{4}t^{4} + O(|t|^{5}),
\]
где $c_3$ и $c_4$ — вещественные постоянные. Отсюда

$$-it(na + u) + t n \ln \varphi(t) =$$

$$= -itu - \frac{t^2}{2} nb + ic_3 n t^3 + c_4 n t^4 + O(n |t|^5),$$  \(22\)

где в правой части все слагаемые, начиная с третьего, бесконечно малы при $|t| \ll \delta = n^{-\lambda}$.

Учитывая, что $\lambda > \frac{3}{7}$, мы имеем $n^4 t^{12} = o(n |t|^5)$, а потому

$$\cos(c_3 n t^8) = 1 - \frac{1}{2} c_3^2 n^2 t^6 + O(n |t|^5).$$

Поэтому в силу

$$e^{c_3 n t^4} = 1 + c_4 n t^4 + O(n^2 t^8) = 1 + c_4 n t^4 + O(n |t|^5)$$

мы получаем из (22):

$$e^{-it(na + u)} \varphi^n(t) =$$

$$= e^{-itu - \frac{tn}{2} nb} (\cos c_3 n t^3 + i \sin c_3 n t^3) \{1 + c_4 n t^4 + O(n |t|^5)\} =$$

$$= e^{-itu - \frac{nb}{2} t^2} \left\{1 - \frac{1}{2} c_3^2 n^2 t^6 + O(n |t|^5) + i \sin c_3 n t^8\right\} \times$$

$$\times \{1 + c_4 n t^4 + O(n |t|^5)\} =$$

$$= e^{-itu - \frac{nb}{2} t^2} \left\{1 + c_4 n t^4 - \frac{1}{2} c_3^2 n^2 t^6 + i \sin c_3 n t^8 - \frac{1}{2} c_3^2 c_4 n^3 t^{10} + i c_4 n t^4 \sin c_3 n t^8 + O(n |t|^5)\right\}.$$  

Но здесь в силу $\lambda > \frac{3}{7} > \frac{1}{3}$

$$\frac{1}{2} c_3^2 c_4 n^3 t^{10} = o(n^2 |t|^7),$$

и мы находим окончательно:

$$e^{-it(na + u)} \varphi^n(t) = e^{-itu - \frac{nb}{2} t^2} \left\{1 + c_4 n t^4 - \frac{1}{2} c_3^2 n^2 t^6 + i \sin c_3 n t^8 + O(n |t|^5) + O(n^2 |t|^7)\right\},$$  \(23\)

В левой части этого равенства вследствие $na + u = na_0 + ld$ стоит подинтегральная функция интеграла (21),
для которой мы таким образом получили нужное нам асимптотическое выражение.

3. Оценим сперва интеграл главного члена этой формулы

\[ \frac{d}{2\pi} \int_{-\infty}^{\infty} e^{-itu - t^2 \frac{nb}{2}} dt. \]

Так как в силу известной формулы Лапласа *

\[ \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itu - t^2 \frac{nb}{2}} dt = \frac{1}{\sqrt{2\pi nb}} e^{-\frac{u^2}{2nb}}, \quad (L) \]

то

\[ \frac{d}{2\pi} \int_{-\delta}^{\delta} e^{-itu - t^2 \frac{nb}{2}} dt = \frac{d}{\sqrt{2\pi nb}} e^{-\frac{u^2}{2nb}} - \frac{d}{2\pi} \int_{|t|>\delta} e^{-itu - t^2 \frac{nb}{2}} dt, \]

откуда в силу (20)

\[ \frac{d}{2\pi} \int_{-\delta}^{\delta} e^{-itu - t^2 \frac{nb}{2}} dt = \frac{d}{\sqrt{2\pi nb}} e^{-\frac{u^2}{2nb}} + o\left(n^{-\gamma}\right), \quad (24) \]

gде \( \gamma \) — любое вещественное число.

4. Теперь обратимся к оценке остаточных членов. С этой целью заменим в правой части формулы (23) \( e^{-itu} \) через \( \cos ut - i \sin ut \) и произведём почленное перемножение, полагая ещё \( \cos ut = 1 + O\left(a^2t^2\right) \). При этом в произведении мы сохраним только члены, чётные относительно \( t \), так как нечётные после интеграции уничтожаются. Это даёт для совокупности остаточных членов в правой части (23) выражение

\[ e^{-t^2 \frac{nb}{2}} \left\{ c_4 nt^4 - \frac{1}{2} c_3 n^2 t^6 + \sin ut \sin c_8 nt^8 + \right. \]
\[ + O(nu^2 t^6) + O(n^2 u^2 t^8) + O(n |t|^5) + O(n^2 |t|^7) = \]
\[ = e^{-t^2 \frac{nb}{2}} \left\{ c_4 nt^4 - \frac{1}{2} c_3 n^2 t^6 + uc_8 nt^4 + O(|u| n^3 t^{10}) + \right. \]
\[ + O(|u|^3 nt^6) + O(u^2 nt^6) + O(u^2 n^2 t^8) + \]
\[ \left. + O(n |t|^5) + O(n^2 |t|^7) \right\}. \quad (25) \]

*) См., например, Г. М. Фихтенгольц, Курс дифференциального и интегрального исчисления, т. II, Гостехиздат, 1948, стр. 720.
Переходим теперь к интегрированию. При этом мы во всех случаях будем вести интегрирование в пределах \((-\infty, +\infty)\); для первых трех членов это в силу (20) вносит лишь погрешность порядка \(o(n^{-1})\) (где \(\gamma\) — любое вещественное число), для остальных же только усиливает оценку. Так как при любом \(r \geqslant 0\)

\[
\int_{-\infty}^{+\infty} |t|^r e^{-t^2/2} dt = \frac{\beta_r}{n^{r+1}},
\]

где \(\beta_r\) — положительная постоянная, то интегрирование трех первых членов разложения (25) даёт нам выражение вида

\[
\frac{k_0 + k_1 u}{n^{3/2}},
\]

где числа \(k_0\) и \(k_1\) не зависят от \(n\) и \(u\).

Остальные шесть членов разложения (25) после интегрирования соответственно дают:

\[
O\left(|u| n^{-5/2}\right), \; O\left(|u|^3 n^{-5/2}\right), \; O\left(u^2 n^{-2}\right), \; O\left(u^3 n^{-2}\right),
\]

\[
O(n^{-2}), \; O(n^{-2}).
\]

Следовательно, сумма всех полученных интегралов может быть записана в виде

\[
\frac{k_0 + k_1 u}{n^{3/2}} + O\left(\frac{Vn + |u|^3}{n^{5/2}}\right),
\]

где \(k_0\) и \(k_1\) — постоянные, не зависящие от \(n\) и \(u\). От этого выражения интеграл совокупности остаточных членов, взятый по интервалу \((-\delta, +\delta)\), отличается, следовательно, на величину порядка \(o(n^{-1})\). А так как интеграл главного члена даётся формулой (24), то мы находим:

\[
\frac{d}{2\pi} \int_{-\delta}^{\delta} e^{-it(na_0 + t\delta)} \varphi^3(t) dt =
\]

\[
= \frac{d}{\sqrt{2\pi nb}} e^{-u^2/2nb} + \frac{m_0 + m_1 u}{n^{3/2}} + O\left(\frac{Vn + |u|^3}{n^{5/2}}\right),
\]
где $u = na_0 + ld - na$, а $m_0$ и $m_1$ не зависят от $n$ и $u$. В силу (21) это даёт:

$$P_n(l) = \frac{d}{V 2\pi n b} e^{-\frac{u^2}{2nb}} + \frac{m_0 + m_1 u}{n^{3/2}} + O\left(\frac{\sqrt{n} + |u|}{n^{3/2}}\right),$$

что мы и имели в виду установить.

Выражаемая формулой (27) одномерная предельная теорема даёт для вероятности $P_n(l)$ оценку, достаточно точную для всех приложений, нужных квантовой статистике. В подавляющем большинстве случаев, однако, достаточной оказывается и более грубая (но зато и более простая) оценка, которую мы теперь и установим. Вернёмся к выражению (23) для нашей подинтегральной функции, снова заменим $e^{-itu} = \cos tu - i \sin tu$, но на этот раз не будем полагать, как выше, $\cos tu = 1 + O(t^2u^2)$, а просто заметим, что $\cos tu$ — чётная функция от $t$, по абсолютной величине не превосходящая 1. Это даёт нам попрежнему главный член

$$e^{-itu - \frac{nb}{2} t^2},$$

а для совокупности остаточных членов мы получаем выражение (выписывая лишь члены чётной относительно $t$ степени)

$$e^{-\frac{t^2}{2} nb} \left\{ c_4 nt^4 \cos tu - \frac{1}{2} c_3^2 nt^6 \cos tu + \right.$$  

$$\left. + \sin c_3 nt^8 + O(n |t|^6) + O(n^2 |t|^7) \right\} =$$

$$= e^{-\frac{t^2}{2} nb} \left\{ O(n t^4) + O(n^2 t^6) + O(|u| n t^4) \right\};$$

после интеграции в пределах $(-\infty, +\infty)$ полученные члены дают в силу (26)

$$O\left(n^{-\frac{3}{2}}\right), \quad O\left(n^{-\frac{3}{2}}\right), \quad O\left(|u| n^{-\frac{3}{2}}\right),$$

вследствие чего мы находим аналогично предыдущему:

$$P_n(l) = \frac{d}{V 2\pi n b} e^{-\frac{u^2}{2nb}} + O\left(\frac{1 + |u|}{n^{3/2}}\right).$$
Это и есть та более грубая оценка вероятности \( P_n(l) \), которая нам будет нужна в дальнейшем.

В заключение мы дадим полную формулировку доказанной нами одномерной предельной теоремы.

**Теорема 1.** Пусть

\[ p_l = P(\xi = a_0 + ld) \]

—закон распределения целочисленной случайной величины \( \xi \) с возможным значением \( a_0 \), шагом \( d \), математическим ожиданием \( a \) и дисперсией \( b \). Пусть \( s_n \) есть сумма \( n \) взаимно независимых случайных величин, каждая из которых подчинена закону \( p_l \), и пусть

\[ P(s_n = na_0 + ld) = P_n(l); \]

если при этом закон \( p_l \) имеет конечные моменты до пятого порядка включительно, то при \( n \to \infty \) равномерно относительно \( l \) имеют место соотношения

\[
P_n(l) = \frac{d}{\sqrt{2\pi}nb} e^{-\frac{n^2}{2nb}} + O\left(\frac{1}{n^{\beta}}\right)
\]

и

\[
P_n(l) = \frac{d}{\sqrt{2\pi}nb} e^{-\frac{u^2}{2nb}} + m_0 + m_1u + O\left(\frac{\sqrt{n} + |u|^8}{n^{\beta}}\right),
\]

где \( u = na_0 + ld - na \), а \( m_0 \) и \( m_1 \) — постоянные.

§ 5. Двумерная предельная теорема

Пусть \((\xi_1, \eta_1), (\xi_2, \eta_2), \ldots, (\xi_n, \eta_n)\) — взаимно независимые целочисленные случайные векторы подчинённые одному и тому же закону распределения \( p_{kl} \) с максимальной решёткой

\[ a_0 + k\alpha + l\beta, \quad b_0 + k\gamma + l\delta, \]

так что

\[ a\delta - \beta\gamma = d \neq 0, \]

так что

\[ p_{kl} = P(\xi_i = a_0 + k\alpha + l\beta, \quad \eta_i = b_0 + k\gamma + l\delta) \]

\( (1 \leq i \leq n; \quad -\infty < k < +\infty, \quad -\infty < l < +\infty) \).
Положим

$$\sum_{i=1}^{n} \xi_i = S_n, \quad \sum_{i=1}^{n} \eta_i = T_n;$$

для случайного вектора \((S_n, T_n)\) возможными точками будут, очевидно, только пары чисел вида

$$na_0 + k\alpha + l\beta, \quad nb_0 + k\gamma + l\delta,$$

(29)

составляющие собой, следовательно, если заставить \(k\) и \(l\) пробегать все целые числа, решётку вектора \((S_n, T_n)\). Положим для краткости

$$P(S_n = na_0 + k\alpha + l\beta, \quad T_n = nb_0 + k\gamma + l\delta) = P_n(k, l).$$

Характеристическая функция вектора \((\xi_i, \eta_i)\) есть

$$\varphi(t, s) = \sum_{k, l=-\infty}^{+\infty} p_{kl} e^{i[(a_0 + k\alpha + l\beta)t + (b_0 + k\gamma + l\delta)s]},$$

согласно правилу композиции, характеристическая функция вектора \((S_n, T_n)\) есть \(\{\varphi(t, s)\}^n\), или, как мы будем писать ради краткости, \(\varphi^n(t, s)\). Таким образом,

$$\sum_{k, l=-\infty}^{+\infty} P_n(k, l) e^{i[(na_0 + k\alpha + l\beta)t + (nb_0 + k\gamma + l\delta)s]} = \varphi^n(t, s).$$

Отсюда формула обращения (11) § 3 даёт нам:

$$P_n(k, l) = \frac{|d|}{4\pi^2} \int \int_D e^{-i[(na_0 + k\alpha + l\beta)t + (nb_0 + k\gamma + l\delta)s]} \varphi^n(t, s) dt ds,$$

(30)

где область \(D\) определяется неравенствами

$$|\alpha t + \gamma s| \leq \pi, \quad |\beta t + \delta s| \leq \pi$$

(D)

(мы не исследовали вопроса о максимальности введённой нами для вектора \((S_n, T_n)\) решётки (29); в этом, однако, и нет надобности, так как формула обращения (11) имеет место для любой решётки, максимальной или нет, лишь бы эта решётка охватывала собой все возможные точки данного вектора).

Нашей целью и здесь является установление удовлетворительного приближённого выражения для вероятности \(P_n(k, l)\) с достаточно точной оценкой получаемой при этом погрешности.
Мы допустим, что исходный закон распределения \( p_{kl} \) не вырожден, откуда в силу § 2 следует, что он имеет максимумальную решётку, что мы, впрочем, уже предположили в начале настоящего параграфа. Далее мы, как в одномерном случае, допустим, что \( \xi_i \) и \( \eta_i \) имеют конечные моменты до пятого порядка включительно, что равносильно сходимости двойного ряда

\[
\sum_{k, l=-\infty}^{+\infty} (|k|^5 + |l|^5) p_{kl}.
\]

Это допущение имеет своим следствием существование у характеристической функции \( \varphi(t, s) \) всех частных производных до пятого порядка включительно в точке \( t = s = 0 \); в частности, существуют математические ожидания

\[
a_1 = E\xi = -i \left( \frac{\partial \varphi}{\partial t} \right)_{t=s=0}, \quad a_2 = E\eta = -i \left( \frac{\partial \varphi}{\partial s} \right)_{t=s=0}
\]

и дисперсии

\[
b_{11} = E \left\{ (\xi - a_1)^2 \right\} = \left\{ \left( \frac{\partial \varphi}{\partial t} \right)^2 - \frac{\partial^2 \varphi}{\partial t^2} \right\} _{t=s=0},
\]

\[
b_{22} = E \left\{ (\eta - a_2)^2 \right\} = \left\{ \left( \frac{\partial \varphi}{\partial s} \right)^2 - \frac{\partial^2 \varphi}{\partial s^2} \right\} _{t=s=0},
\]

а также «смешанный» момент второго порядка

\[
b_{12} = \sqrt{b_{11} b_{22}} R_{12} = E \left\{ (\xi - a_1)(\eta - a_2) \right\} = \left\{ \frac{\partial \varphi}{\partial t} \frac{\partial \varphi}{\partial s} - \frac{\partial^2 \varphi}{\partial t \partial s} \right\} _{t=s=0}.
\]

При этом в силу предположенной нами невырожденности вектора \( (\xi_i, \eta_i) \) коэффициент корреляции \( R_{12} \) величин \( (\xi_i, \eta_i) \) отличен от \( \pm 1^* \), т. е.

\[
|b_{12}| < \sqrt{b_{11} b_{22}},
\]

или

\[
\Delta = b_{11} b_{22} - b_{12}^2 > 0.
\]

Математические ожидания величин \( S_n \) и \( T_n \) соответствственно равны \( na_1 \) и \( na_2 \). Уклонения значений \( na_0 + ka + \beta \),

* ) См. Б. В. Гнеденко, Курс теории вероятностей, Гостехиздат, 1950, стр. 145.
§ 5] ДВУМЕРНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА

\[ nb_0 + k\gamma + l \beta \] этих величин от их математических ожиданий будут поэтому соответственно

\[ na_0 + k\alpha + l_\beta - na_1 = u_1, \quad nb_0 + k\gamma + l_\beta - na_2 = u_2. \]

Целью настоящего параграфа является установление имеющего место при \( n \to \infty \) асимптотического равенства

\[
P_n (k, l) = \frac{1}{2\pi n} \frac{1}{\sqrt{\Delta}} e^{-\frac{1}{2n^2}} \left\{ b_{21}u_1^2 - 2b_{13}u_1u_2 + b_{11}u_2^2 \right\} +
\frac{m_0 + m_1u_1 + m_2u_2}{n^2} + O\left( \frac{\sqrt{n} + |u_1|^3 + |u_2|^3}{n^3} \right),
\]

(31)

где \( m_0, m_1, m_2 \) — постоянные (не зависящие от \( n, u_1 \) и \( u_2 \)) числа. Формула (31) и представляет собой наиболее точную нужную нам форму двумерной предельной теоремы.

Прежде всего мы докажем вспомогательное предложение, имеющее целью распространить на двумерный случай лемму 3 предыдущего параграфа (и одновременно — уточнить лемму 2 § 3):

Лемма 4. Существует такое постоянное число \( c > 0 \), что для любой точки \( (t, s) \) области \( D \)

\[
|\varphi (t, s)| \leq e^{-c(t^2 + s^2)}.
\]

Доказательство. В окрестности начала координат функция \( \varphi (t, s) \) в силу наших предпосылок может быть представлена в виде

\[
\varphi (t, s) = 1 + i (a_1t + a_2s) - \frac{1}{2} (a_{11}t^2 + 2a_{12}ts + a_{22}s^2) +
+ O \left( |t|^3 + |s|^3 \right),
\]

где числа \( a_1, a_2, a_{11}, a_{12}, a_{22} \) — моменты первого и второго порядков случайного вектора \( (\xi_t, \tau_t) \). Отсюда

\[
|\varphi (t, s)|^2 = \left[ 1 - \frac{1}{2} (a_{11}t^2 + 2a_{12}ts + a_{22}s^2) \right]^2 + (a_1t + a_2s)^2 +
+ O \left( |t|^3 + |s|^3 \right) = 1 - (a_{11}t^2 + 2a_{12}ts + a_{22}s^2) +
+ (a_1t + a_2s)^2 + O \left( |t|^3 + |s|^3 \right) = 1 - [(a_{11} - a_{22})t^2 +
+ 2(a_{12} - a_1a_2) ts + (a_{22} - a_2^2) s^2] + O \left( |t|^3 + |s|^3 \right).
\]
Здесь \( a_{11} - a_1^2 = b_{11}, \ a_{12} - a_1 a_2 = b_{12}, \ a_{22} - a_2^2 = b_{22} \) — центральные моменты второго порядка вектора \((\xi_i, \eta_i)\), о которых мы уже говорили выше. Таким образом,
\[
|\varphi(t, s)|^2 = 1 - (b_{11}t^2 + 2b_{12}ts + b_{22}s^2) + O(|t|^3 + |s|^3).
\]
Но выше мы видели, что \( \Delta = b_{11}b_{22} - b_{12}^2 \geq 0 \) (невырожденность вектора \((\xi_i, \eta_i)\)); следовательно, форма \( b_{11}t^2 + 2b_{12}ts + b_{22}s^2 \) — положительно определённая (так как \( b_{11} \geq 0, \ b_{22} \geq 0 \)), и, значит, существует такое положительное число \( c_1 \), что тождественно
\[
b_{11}t^2 + 2b_{12}ts + b_{22}s^2 \geq c_1 (t^2 + s^2),
\]
и, следовательно, в окрестности начала координат
\[
|\varphi(t, s)|^2 \leq 1 - c_1 (t^2 + s^2) + c_2 (|t|^3 + |s|^3),
\]
где \( c_2 \) — другая положительная постоянная; если \( \delta < \frac{c_1}{2c_2} \) и \( |t| < \delta, \ |s| < \delta \), то
\[
c_2 (|t|^3 + |s|^3) \leq \frac{c_1}{2} (t^2 + s^2),
\]
и, следовательно,
\[
|\varphi(t, s)|^2 \leq 1 - \frac{c_1}{2} (t^2 + s^2) \leq e^{-\frac{c_1}{2} (t^2 + s^2)},
\]
или
\[
|\varphi(t, s)| \leq e^{-\frac{c_1}{4} (t^2 + s^2)} \quad (|t| < \delta, \ |s| < \delta). \quad (32)
\]
Обозначим через \( Q \) квадрат \( |t| < \delta, \ |s| < \delta \) плоскости \((t, s)\). В силу леммы 2 всюду в (замкнутой) области \( D - Q \) мы имеем \( |\varphi(t, s)| < 1 \); так как функция \(|\varphi(t, s)| \) непрерывна, то должно существовать такое число \( \lambda < 1 \), что всюду в области \( D - Q \) \( |\varphi(t, s)| < \lambda \). Если обозначить через \( A \) наибольшее значение суммы \( t^2 + s^2 \) в области \( D \) и выбрать \( c_3 > 0 \) настолько малым, чтобы было
\[
\lambda < e^{-c_3A},
\]
то мы будем иметь всюду в области \( D - Q \):
\[
|\varphi(t, s)| < \lambda < e^{-c_3A} \leq e^{-c_6(t^2 + s^2)}. \quad (33)
\]
Наконец, если обозначить через $c$ меньшее из чисел $\frac{c_1}{4}$ и $c_8$, то в силу (32) и (33) мы будем иметь:

$$|\varphi(t, s)| \leq e^{-c(t^2 + s^2)}$$

как в области $Q$, так и в области $D - Q$, и следовательно, во всей области $D$. Этим лемма 4 доказана.

Преиспупкая теперь к доказательству асимптотической формулы (31), мы снова разобьем наши рассуждения на этапы, соответствующие тем, какие мы имели в предыдущем параграфе при доказательстве одномерной предельной теоремы.

1. Мы снова обозначим через $\lambda$ произвольное постоянное число, заключенное между $\frac{3}{7}$ и $\frac{1}{2}$, и положим $n^{\lambda} = \varepsilon$.

Двойной интеграл формулы (30) распространяется на область $D$, определяемую неравенствами (D). Эту область мы разобьем на две части, одна из которых есть квадрат $Q$: $|t| \leq \varepsilon$, $|s| \leq \varepsilon$, а другая — дополнение $D - Q$ квадрата $Q$ до исходной области $D$. В области $D - Q$ по меньшей мере одно из чисел $|t|$, $|s|$ превосходит $\varepsilon$, и потому в силу леммы 4 $|\varphi(t, s)| \leq e^{-c\varepsilon^2}$. Таким образом, мы находим (учитывая, что площадь области $D$ равна $\frac{4\pi^2}{|d|}$):

$$\left| \iint_{D - Q} e^{-i[(na_0 + ka + l \beta) t + (nb_0 + k \tau + l \beta) s]} \varphi^n(t, s) \, dt \, ds \right| \leq \frac{4\pi^2}{|d|} e^{-c n^{1-2\lambda}} = o(n^{-\gamma}),$$

где $\gamma$ — любое вещественное число. Это в силу формулы (30) даёт нам для $-\infty < k, l < +\infty$:

$$P_n(k, l) = \frac{|d|}{4\pi^2} \iint_Q e^{-i[(na_0 + ka + l \beta) t + (nb_0 + k \tau + l \beta) s]} \varphi^n(t, s) \, dt \, ds + o(n^{-\gamma}),$$

в полной аналогии с формулой (21) предшествующего параграфа.
2. Так как функция $\varphi(t, s)$ в точке $t = s = 0$ имеет частные производные до пятого порядка включительно, то мы можем (предполагая $|t| \ll \delta, |s| \ll \delta$) написать разложение

$$\ln \varphi(t, s) = iP_1(t, s) - \frac{1}{2} P_2(t, s) + iP_3(t, s) + P_4(t, s) + O(\tau^5),$$

где $\tau = |s| + |t|$, и $P_r(t, s)$ вообще означает однородный вещественный полином степени $r$ относительно $t$ и $s$; в частности,

$$P_1(t, s) = a_1 t + a_2 s,$$
$$P_2(t, s) = b_{11} t^2 + 2b_{12} ts + b_{22} s^2;$$

для любого $r$

$$P_r(t, s) = O(\tau^r).$$

Для логарифма подинтегральной функции в формуле (34) мы отсюда получаем выражение

$$- i (na_1 + u_1) t - i (na_2 + u_2) s + n \ln \varphi(t, s) =$$

$$= - i (u_1 t + u_2 s) - \frac{n}{2} (b_{11} t^2 + 2b_{12} ts + b_{22} s^2) +$$

$$+ i nP_3 + nP_4 + O(n\tau^5).$$

Учитывая, что $\lambda > \frac{3}{7}$, мы имеем $n^4 \tau^{12} = o(n\tau^5)$, и потому

$$\cos nP_3 = 1 - \frac{n^2P_3^2}{2} + O(n^4\tau^{12}) = 1 - \frac{n^2P_3^2}{2} + O(n\tau^5).$$

Отсюда в силу

$$e^{nP_4} = 1 + nP_4 + O(n^2\tau^8) = 1 + nP_4 + O(n\tau^5),$$

мы находим:

$$e^{-i[(na_1 + u_1) t + (na_2 + u_2) s]} \varphi^n(t, s) = e^{-i(u_1 t + u_2 s) - \frac{n}{2} P_3 \left[ \cos nP_3 + i \sin nP_3 \right] \left[ 1 + nP_4 + O(n\tau^5) \right]} =$$

$$= e^{-i(u_1 t + u_2 s) - \frac{n}{2} P_3 \left[ 1 - \frac{n^2P_3^2}{2} + \right.$$

$$+ O(n^5) + i \sin nP_3 \left[ 1 + nP_4 + O(n\tau^5) \right]} =$$

$$= e^{-i(u_1 t + u_2 s) - \frac{n}{2} P_3 \left[ 1 + nP_4 - \frac{n^2P_3^2}{2} + \right.$$

$$+ i \sin nP_3 - \frac{n^2P_3^2}{2} nP_4 + nP_4 i \sin nP_3 + O(n\tau^5) \right].$$
Но здесь в силу $\lambda > \frac{3}{7} > \frac{2}{5}$

$$
\frac{1}{2} n^2 P_3^2 n P_4 = O(n^3 \tau^{10}) = O(n \tau^5),
$$

$$
n P_4 \sin n P_8 = O(n^2 \tau^7),
$$

и мы находим окончательно:

$$
e^{-i \left[ (n a_1 + u_1) t + (n a_2 + u_2) s \right]} \varphi^n(t, s) =
$$

$$
e^{-i (u_1 t + u_2 s) - \frac{n}{2} P_3} \left[ 1 + n P_4 - \frac{1}{2} n^2 P_3^2 - i \sin n P_8 +
\right.
$$

$$
\left. + O(n \tau^5) + O(n^2 \tau^7) \right].
$$

Это асимптотическое выражение подинтегральной функции формулы (34) в точности аналогично выражению (23) предшествующего параграфа.

3. Прежде всего оценим интеграл

$$
\frac{|d|}{4\pi^2} \int e^{-i (u_1 t + u_2 s) - \frac{n}{2} P_3} dt ds
$$

главного члена формулы (35). Так как

$$
u_1 t + u_2 s = u_1 \left( t + \frac{b_{12}}{b_{11}} s \right) + s \left( u_2 - \frac{b_{12}}{b_{11}} u_1 \right)
$$

и

$$
P_2(t, s) = b_{11} t^2 + 2 b_{12} t s + b_{22} s^2 = b_{11} \left( t + \frac{b_{12}}{b_{11}} s \right)^2 + \frac{\Lambda}{b_{11}} s^2,
$$

то

$$
\frac{|d|}{4\pi^2} \int e^{+\infty}_{-\infty} - e^{+\infty}_{-\infty} (u_1 t + u_2 s) - \frac{n}{2} P_3(t, s) dt ds =
$$

$$
\int e^{+\infty}_{-\infty} - \left( u_2 - \frac{b_{12}}{b_{11}} u_1 \right) - \frac{n}{2} \frac{\Lambda}{b_{11}} s^2 ds \times
$$

$$
\times \int e^{-\infty}_{-\infty} \left( t + \frac{b_{12}}{b_{11}} s \right) - \frac{n}{2} b_{11} \left( t + \frac{b_{12}}{b_{11}} s \right)^2 dt =
$$

$$
\int e^{-\infty}_{-\infty} - \left( u_2 - \frac{b_{12}}{b_{11}} u_1 \right) - \frac{n}{2} \frac{\Lambda}{b_{11}} s^2 ds \int_{-\infty}^{+\infty} - i u_1 v - \frac{n}{2} b_{11} v^2 dv.
$$
К обоим интегралам правой части применим теперь формулу (L) предыдущего параграфа. Это даёт для первого интеграла:

\[ \frac{|d|}{2\pi n} \sqrt{b_{11}} e^{-\frac{(u_{11} - b_{11} u_{1})}{2b_{11}}} \]

а для второго:

\[ \frac{1}{2\pi n b_{11}} e^{-\frac{u_{1}^2}{2nb_{11}}} \]

Таким образом, мы находим:

\[ \frac{|d|}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{\frac{i}{2} (u_{11} t + u_{22} s) - \frac{n}{2} P_{2}(t, s)} dt ds = \]

\[ = \frac{|d|}{2\pi n \sqrt{\Delta}} e^{-\frac{1}{2n\Delta} \left\{ b_{11} u_{2}^2 - 2b_{11} u_{1} u_{2} + b_{22} u_{1}^2 \right\}} \]

Это — как раз нужный нам главный член правой части формулы (31); но мы интегрировали по всей плоскости (\(-\infty < t, s << +\infty\)) вместо квадрата Q и должны поэтому оценить происходящую отсюда погрешность. Для этого заметим, что, например,

\[ \int_{0}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{n}{2} P_{2}} dt ds = \int_{0}^{\infty} e^{-\frac{n}{2} b_{11} s^2} ds \int_{-\infty}^{\infty} e^{-\frac{n}{2} b_{11} \left( t + \frac{b_{11}}{b_{n}} s \right)^2} dt = \]

\[ = \int_{0}^{\infty} e^{-\frac{n\Delta}{2b_{11}} s^2} ds \cdot \int_{-\infty}^{\infty} e^{-\frac{n}{2} b_{11} u^2} du = o (n^{\gamma}) \]

при любом вещественном \( \gamma \), так как эту оценку для первого множителя даёт формула (20) § 4, а второй множитель стремится к нулю при \( n \to \infty \). Таким образом, мы можем считать установленным, что

\[ \frac{|d|}{4\pi^2} \int_{D} e^{-i (u_{11} t + u_{22} s) - \frac{n}{2} P_{2}(t, s)} dt ds = \]

\[ = \frac{|d|}{2\pi n \sqrt{\Delta}} e^{-\frac{1}{2n\Delta} \left\{ b_{11} u_{2}^2 - 2b_{11} u_{1} u_{2} + b_{22} u_{1}^2 \right\}} + o (n^{\gamma}), \quad (36) \]

и перейти к оценке остаточных членов.
4. С этой целью мы в полной аналогии с расчётами предыдущего параграфа заменим в правой части формулы (35) \( e^{-i(u_1 t + u_2 s)} \) через \( \cos (u_1 t + u_2 s) - i \sin (u_1 t + u_2 s) \). Производя затем частное перемножение, мы учтём ещё, что \( \cos (u_1 t + u_2 s) = 1 + O(u^2 \tau^2) \), где \( u = |u_1| + |u_2| \), и, как в одномерном случае, сохраним лишь члены чётного порядка относительно пары переменных \( t, s \), так как нечётные члены после интеграции по квадрату \( Q \) дадут нули. Это даёт для совокупности остаточных членов выражение

\[
e^{-\frac{n}{2} P_2}\left\{ nP_4 - \frac{1}{2} n^2 P_3^2 - \sin nP_3 \sin (u_1 t + u_2 s) + O(n u^2 \tau^6) + O(n^2 u^2 \tau^8) + O(n^3 \tau^7) + O(n \tau^5) \right\} =
\]

\[
e^{-\frac{n}{2} P_2}\left\{ nP_4 - \frac{1}{2} n^2 P_3^2 + nP_3 (u_1 t + u_2 s) + O(n u^3 \tau^6) + O(n^2 u^2 \tau^8) + O(n^2 \tau^7) + O(n \tau^5) \right\}. (37)
\]

Теперь мы должны интегрировать выражение (37) по области \( Q \). При этом три первых члена мы можем интегрировать по всей плоскости \( -\infty < t, s < +\infty \) вместо области \( Q \), внося этим лишь погрешность порядка \( n^{-\gamma} \), где \( \gamma \) — любое вещественное число. Чтобы в этом убедиться, достаточно, очевидно, показать, что при любых целых \( p \geq 0, q \geq 0 \):

\[
I(p, q) = \int_0^\infty s^q d s \int_{-\infty}^{+\infty} t^p e^{-\frac{n}{2} P_2} dt = O(n^{-\gamma});
\]

но, полагая \( t + \frac{b_{12}}{b_{11}} s = z \), мы, очевидно, имеем:

\[
P_2 = \frac{\Delta}{b_{11}} s^2 + b_{11} z^2,
\]

и, следовательно,

\[
I(p, q) = \int_0^\infty s^q e^{-\frac{n}{2} \frac{\Delta}{b_{11}} s^2} ds \int_{-\infty}^{+\infty} \left(z - \frac{b_{12}}{b_{11}} s \right)^p e^{-\frac{n}{2} b_{11} z^2} dz.
\]

Разлагая \( \left(z - \frac{b_{12}}{b_{11}} s \right)^p \) по формуле бинома, мы (не производя никаких вычислений) видим, что интеграл \( I(p, q) \) пред-
Сведения из теории вероятностей

ставляется в виде суммы произведений вида

$$
\sum_{0}^{\infty} s^{w} e^{-\frac{n}{2} a s} - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z^{w} e^{-\frac{n}{2} b z} dz,
$$

где при возрастании $n$ первый множитель имеет вид $o(n^{-1})$
в силу формулы (20) предыдущего параграфа, а второй во
всем случае остаётся ограниченным. Это и показывает, что
$I(u, v) = o(n^{-1})$.

Что касается остальных членов (стоящих под знаком $O$)
то здесь, разумеется, расширение области интегрирования только
усиливает получаемые неравенства. Поэтому мы можем и эти
члены интегрировать по всей плоскости $(t, s)$ вместо обла-
сти $Q$.

Для однородного многочлена $P_{r}(t, s)$ степени $r$ преобра-
зование $t\sqrt{n} = t', s\sqrt{n} = s'$ даёт:

$$
\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P_{r}(t, s) e^{-\frac{n}{2} P_{r}(t, s)} dt ds = \frac{c}{n^{r+1}}
$$

где $c$ не зависит от $n$; поэтому интеграция первых трёх чле-
нов разложения (37), распространённая на всю плоскость $(t, s)$,
даёт выражение вида

$$
\frac{k_{0} + k_{1} u_{1} + k_{2} u_{2}}{n^{2}}
$$

где числа $k_{0}, k_{1}, k_{2}$ не зависят от $n, u_{1}$ и $u_{2}$.

Для последующих же шести членов разложения (37) мы
после интеграции получаем соответственно оценки

$$
O(u n^{-8}), O(u^{3} n^{-8}), O(u^{2} n^{-8}), O(u^{2} n^{-8}), O(n^{-5}), O(n^{-5}).
$$

Таким образом, вся совокупность остаточных членов после
интеграции даёт:

$$
\frac{k_{0} + k_{1} u_{1} + k_{2} u_{2}}{n^{2}} + O\left(\frac{\sqrt{n} + u}{n^{3}}\right).
$$
§ 5] ДВУМЕРНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА

Сопоставляя это с интегралом (36) главного члена, мы находим \((m_0, m_1, m_2)\) означают числа, не зависящие от \(n, u_1\) и \(u_2\):

\[
\frac{|d|}{4\pi^2} \int_{Q} e^{-i\left[(na_0 + \xi + 13)t + (nb_0 + \xi + 13)s\right]} \varphi^* (t, s) \, dt \, ds =
\]

\[
= \frac{|d|}{2\pi n \sqrt{\Delta}} e^{-\frac{1}{2n\Delta} \left\{b_{11}u_1^2 - 2b_{12}u_1u_2 + b_{22}u_2^2\right\}} + m_0 + m_1u_1 + m_2u_2 +
\]

\[
+ O \left(\frac{\sqrt{n} + u^3}{n^3}\right),
\]

и, значит, в силу (34)

\[
P_n (k, l) = \frac{|d|}{2\pi n \sqrt{\Delta}} e^{-\frac{1}{2n\Delta} \left\{b_{11}u_1^2 - 2b_{12}u_1u_2 + b_{22}u_2^2\right\}} +
\]

\[
+ m_0 + m_1u_1 + m_2u_2 + O \left(\frac{\sqrt{n} + u^3}{n^3}\right),
\]

(38)

т. е. формулу, которую мы и имели в виду установить.

Теперь нам остаётся ещё только по аналогии с формулой (28) предыдущего параграфа установить для \(P_n (k, l)\) более грубую оценку, которая во многих случаях оказывается вполне достаточной. С этой целью мы в полной аналогии с § 4 при замене в формуле (35) множителя \(e^{-i(u_1t + u_2s)}\) его тригонометрическим выражением не будем, как мы это делали в п. 4 только что проведённого вывода, представлять \(\cos (u_1t + u_2s)\) в виде \(1 + O(u^2 t^2)\), а просто учтем, что \(\cos (u_1t + u_2s)\) есть чётная функция от \(t\) и \(s\), по абсолютной величине не превосходящая единицы. Всё дальнейшее протекает в столь полной аналогии с одномерным случаем, что мы можем здесь не повторять этого вывода, а привести лишь окончательный результат, который гласит:

\[
P_n (k, l) = \frac{|d|}{2\pi n \sqrt{\Delta}} e^{-\frac{1}{2n\Delta} \left\{b_{11}u_1^2 - 2b_{12}u_1u_2 + b_{22}u_2^2\right\}} + O \left(\frac{1 + u}{n^2}\right).
\]

(39)

В заключение мы и здесь дадим полную формулировку доказанной нами двумерной предельной теоремы.

Теорема 2. Пусть

\[p_{kl} = P (\xi = a_0 + \xi + l\beta, \eta = b_0 + \xi + l\beta)\]
— закон распределения невырожденного целочисленного случайного вектора \((\xi, \eta)\) с максимальной решёткой \(a_0 + ka + l_3\), \(b_0 + k\gamma + l_5\), математическими ожиданиями \(a_1, a_2\), дисперсиями \(b_{11}, b_{22}\) и коэффициентом корреляции \(\frac{b_{12}}{\sqrt{b_{11}b_{22}}}\). Пусть \((S_n, T_n)\) есть сумма \(n\) взаимно независимых случайных векторов, каждый из которых подчиняется закону \(p_{kl}\), и пусть

\[
P_n(k, l) = P\left(S_n = na_0 + ka + l_3, T_n = nb_0 + k\gamma + l_5\right).
\]

Если при этом закон \(p_{kl}\) имеет конечные моменты до пятого порядка включительно, то при \(n \to \infty\) равномерно относительно \(k\) и \(l\) имеют место соотношения

\[
P_n(k, l) = \frac{|\Delta|}{2\pi n} e^{-\frac{1}{2\Delta} \left\{ b_{11}u_1^2 - 2b_{12}u_1u_2 + b_{22}u_2^2 \right\} + O\left(\frac{1}{n^2}\right)} \tag{39}
\]

и

\[
P_n(k, l) = \frac{|\Delta|}{2\pi n} e^{-\frac{1}{2\Delta} \left\{ b_{11}u_1^2 - 2b_{12}u_1u_2 + b_{22}u_2^2 \right\} + \frac{m_0 + m_1u_1 + m_2u_2}{n^2} + O\left(\frac{\sqrt{n}}{n^3}\right),} \tag{38}
\]

где положено \(d = a_\delta - b_\gamma\), \(\Delta = b_{11}b_{22} - b_{12}^2 > 0\), \(u_1 = na_0 + ka + l_3 - na_1\), \(u_2 = nb_0 + k\gamma + l_5 - na_2\), \(u = |u_1| + |u_2|\), и где \(m_0, m_1, m_2\) означают постоянные (не зависящие от \(n, u_1, u_2\)) числа.
ГЛАВА II
ПРЕДВАРИТЕЛЬНЫЕ СВЕДЕНИЯ
ИЗ КВАНТОВОЙ МЕХАНИКИ

§ 1. Описание состояния физической системы
в квантовой механике

В классической механике состояние системы с s степенями свободы удобно описывать, задавая значения всех 2s «гамильтоновых переменных»: «обобщённых координат» q₁, q₂, ..., qₛ и «обобщённых импульсов» p₁, p₂, ..., pₛ. Таким образом, в классической механике состояние системы может быть изображено точкой её «фазового пространства», представляющего собой евклидов пространство 2s измерений с декартовыми координатами q₁, q₂, ..., qₛ, p₁, p₂, ..., pₛ. Известно, что такой способ изображения состояния системы особенно удобен как раз для целей статистической механики. При этом любая механическая величина, связанная с данной системой, будучи некоторой функцией F(q₁, q₂, ..., qₛ, p₁, p₂, ..., pₛ) её гамильтоновых переменных, в каждом определённом состоянии системы получает единственное определённое значение; иначе говоря, такая величина есть функция точки фазового пространства, имеющая в каждой такой точке некоторое определённое значение.

В квантовой механике вопрос о характеристизации состояния данной физической системы решается совершенно иначе. Состояние системы описывается здесь заданием некоторой (в общем говоря, комплексной) функции

\[ U = U(q₁, q₂, \ldots, qₛ) \]

одних только обобщённых координат системы (такую функцию принято, впрочем без достаточных оснований, называть «вол-
новой функцией» данной системы). Знание функции \( U \), описывающей состояние системы, не позволяет, вообще говоря, однозначно определить ни одной из гамильтоновых переменных \( q_1, q_2, \ldots, q_s, p_1, p_2, \ldots, p_s \); в лучшем случае, зная функцию \( U \), мы можем определить значения некоторых из этих переменных, но никогда не можем быть однозначно определены значения всех \( 2s \) переменных, ибо в квантовой механике величины \( q_i \) и \( p_i \) (для одного и того же \( i \)) ни в одном состоянии не могут обе иметь определённых значений. Тем более, конечно, какая-либо механическая величина \( F(q_1, \ldots, q_s, p_1, \ldots, p_s) \) не будет, вообще говоря, в состоянии, описываемом функцией \( U \) (или, как мы будем говорить короче, в состоянии \( U \)), иметь единственного определённого значения.

Заданием функции \( U \) в квантовой механике определяются не значения механических величин, а их законы распределения. Это значит, что, зная функцию \( U \), мы для любой механической величины \( \mathbf{X} \) можем вычислить вероятность того, что она получит значение, заключённое в любом заданном промежутке \((a, b)\), т. е. вероятность неравенства \( a < \mathbf{X} < b \). Только в том исключительном случае, когда получаемый для величины \( \mathbf{X} \) закон распределения имеет лишь одно возможное значение, можно говорить, что величина \( \mathbf{X} \) в состоянии \( U \) получает определённое значение.

В частности, если система находится в состоянии \( U \), то вероятность того, что значения переменных \( q_1, q_2, \ldots, q_s \) окажутся принадлежащими данной области \( V \) конфигурационного пространства\(^*\)), определяется формулой

\[
\frac{\int_V |U|^2 dq_1 \ldots dq_s}{\int |U|^2 dq_1 \ldots dq_s},
\]

где интеграл в знаменателе распространяется на всё конфигурационное пространство. В ещё более частном случае, когда система представляет собою элементарную частицу с простейшими гамильтоновыми переменными \( x, y, z, p_x, p_y, p_z \), веро-

\(^{*)}\) То-есть евклидова пространства с измерений с декартовыми координатами \( q_1, q_2, \ldots, q_s \).
явность того, что эта частица будет принадлежать области $V$
обыкновенного трёхмерного пространства, выражается формулой

$$
\int_V |U(x, y, z)|^2 \, dx \, dy \, dz
$$

$$
\int |U(x, y, z)|^2 \, dx \, dy \, dz',
$$

где интеграл в знаменателе распространяется на всё трёхмерное пространство.

Позволяя установить закон распределения механической величины $\mathcal{A}$, функция $U$, очевидно, позволяет однозначно определить и математическое ожидание $E\mathcal{A}$ этой величины. В частности, если

$$\mathcal{A} = F(q_1, \ldots, q_s)$$

есть любая функция обобщённых координат $q_1, q_2, \ldots, q_s$, то

$$E\mathcal{A} = \frac{\int F(q_1, \ldots, q_s) |U|^2 \, dq_1 \ldots dq_s}{\int |U|^2 \, dq_1 \ldots dq_s},$$

(1)

где оба интеграла берутся по всему конфигурационному пространству.

Все приведённые нами формулы показывают, что замена функции $U$ функцией $\lambda U$, где $\lambda \neq 0$ — любое комплексное постоянное число, не вызывает никаких изменений в статистике по крайней мере таких механических величин, которые зависят только от обобщённых координат $q_1, q_2, \ldots, q_s$ (т. е. не зависят от обобщённых импульсов $p_1, p_2, \ldots, p_s$); ниже мы увидим, что это же правило сохраняет силу и для механических величин любого рода; мы должны поэтому считать, что функции $U$ и $\lambda U$, каково бы ни было постоянное комплексное число $\lambda \neq 0$, описывают одно и то же состояние системы. Пользуясь этим произволом в выборе $\lambda$, мы можем, очевидно, описать любое состояние системы такой функцией $U$, для которой

$$\int |U|^2 \, dq_1 \ldots dq_s = 1;$$

такую функцию $U$ мы будем называть нормированной; для нормированных функций $U$ все написанные нами формулы упрощаются, так как знаменатели их обращаются в единицу.
Если $U$ — описывающая какое-либо состояние системы нормированная функция и $\lambda = e^{i\varphi}$ — любое постоянное комплексное число с модулем 1, то функция $\lambda U$ также нормирована и описывает то же состояние системы.

Само собою разумеется, что описанное нами применение функции $U$ к установлению статистики механических величин требует сходимости всех тех интегралов, которыми мы при этом пользовались; в частности, функция $|U|^2$ должна быть интегрируема по всему конфигурационному пространству (инчае говоря, функция $U$ должна быть элементом комплексного гильбертового пространства). Если этому требованию удовлетворяют функции $U_1$ и $U_2$, то, как известно, интеграл

$$\int U_1 U_2^* \, dq_1 \ldots dq_s$$

(где звёздочка означает переход к комплексно сопряжённой величине), взятый по всему конфигурационному пространству, абсолютно сходится; мы будем называть этот интеграл скалярным произведением функций $U_1$ и $U_2$ и обозначать его для краткости через $(U_1, U_2)$; в частности,

$$(U, U) = \int |U|^2 \, dq_1 \ldots dq_s,$$

и следовательно, для нормированной функции $U$

$$(U, U) = 1.$$

Скалярные произведения, очевидно, обладают следующими элементарными свойствами:

1°.

$$(U_1 + U_2, U_3) = (U_1, U_3) + (U_2, U_3),$$

$$(U_1, U_2 + U_3) = (U_1, U_2) + (U_1, U_3).$$

2°. Если $\lambda$ — любое комплексное число, то

$$(\lambda U_1, U_2) = \lambda (U_1, U_2),$$

$$(U_1, \lambda U_2) = \lambda^* (U_1, U_2).$$

3°.

$$(U_2, U_1) = (U_1, U_2)^*.$$
§ 1. Описание состояния физической системы

$\lambda_1$ и $\lambda_2$ — произвольные (комплексные) постоянные числа и $|\lambda_1| + |\lambda_2| > 0$, описывает некоторое возможное состояние системы. Этот принцип, очевидно, автоматически распространяется на случай любого числа слагаемых. Он распространяется и на бесконечные ряды: если ряд

$$\sum_{k=1}^{\infty} \lambda_k U_k$$

сходится в среднем *) к функции $U$ и если каждая из функций $U_k$ описывает некоторое возможное состояние системы, то это же верно и для функции $U$. Во всех случаях исключение составляет только сумма, тождественно равная нулю: функция $U \equiv 0$ никакого состояния системы не описывает; она не может быть нормирована, и все приведённые нами выше формулы для вероятностей и математических ожиданий в случае $U \equiv 0$ теряют смысл.

В истории статистической физики можно указать целый ряд случаев, когда недостаточно отчётливо определённое понятие вероятности того или иного события вело к смещениям и взаимонепониманию и тем самым тормозило развитие науки. В физике, как и во всех прикладных науках, вероятность события всегда означает относительную частоту его наступления. Однако одно и то же событие может в различных условиях иметь весьма различные вероятности. Поэтому любое вероятностное утверждение получает точный смысл только лишь в том случае, когда исчерпывающим образом указано те условия, к которым оно относится.

В квантовой механике, как мы это отчасти уже видели, значительно шире увидим в дальнейшем, основное значение имеют утверждения вида: «если система находится в состоянии $U$, и $\mathcal{A}$ есть связанная с этой системой механическая величина, то вероятность того, что $a < \mathcal{A} < b$, равна такому-то числу $p$». В целом избегания смещений и неясностей важно поэтому с самого начала со всею отчётливостью выяснить смысл такого утверждения. Пусть мы имеем большое число

*) То-есть $(U - S_n, U - S_n) \to 0$ при $n \to \infty$, где $S_n = \sum_{k=1}^{n} \lambda_k U_k$. 
§ 2. Механические величины
и самосопряженные линейные операторы

В § 1 мы видели, каким образом квантовая механика позволяет для каждого состояния $U$ определить статистику величины $\mathfrak{A}$, зависящей только от обобщенных координат $q_1, q_2, \ldots, q_s$ данной системы. Но в общем случае механическая
величина $\mathfrak{A}$ зависит от всей совокупности гамильтоновых переменных $q_1, q_2, \ldots, q_s, p_1, p_2, \ldots, p_s$, и мы должны теперь установить правила, позволяющие определить статистику такой величины в любом состоянии $U$ данной системы. С этой целью квантовая механика обращается к понятию линейного самосопряжённого оператора, которое мы и должны теперь рассмотреть.

Оператор $A$ в комплексном гильбертовом пространстве мы называем всякое правило, позволяющее отнести любому элементу $U$ этого пространства некоторый определённый другой элемент $AU$ того же пространства (иногда, впрочем, $AU$ определяется не для всех $U$, а лишь для некоторых из них). Оператор $A$ называется линейным, если для любых элементов $U_1, U_2$ и любых постоянных комплексных чисел $\lambda_1, \lambda_2$

$$A(\lambda_1 U_1 + \lambda_2 U_2) = \lambda_1 AU_1 + \lambda_2 AU_2.$$ 

Каждому оператору $A$ соответствует определённый комплексно-сопряжённый оператор $A^*$, определяемый условием

$$A^* U = (AU)^*.$$ 

Оператор $A$ называется самосопряжённым (или эрмитовским), если для любых $U_1, U_2$

$$(AU_1, U_2) = (U_1, AU_2).$$ \hspace{1cm} (2)

Квантовая механика ставит в соответствие каждой механической величине $\mathfrak{A}$ некоторый линейный самосопряжённый оператор $A$, с помощью которого и устанавливается статистика этой величины в любом состоянии $U$ системы. О том, как это делается, мы скажем ниже. Сейчас заметим, в первую очередь, что если величина $\mathfrak{A}$ является функцией одних только обобщённых координат $q_1, q_2, \ldots, q_s$, то соответствующий ей оператор $A$ определяется просто как умножение на эту функцию: если

$$\mathfrak{A} = F(q_1, \ldots, q_s),$$

то

$$AU(q_1, \ldots, q_s) = F(q_1, \ldots, q_s) U(q_1, \ldots, q_s).$$
Линейность этого оператора самоочевидна. Далее, если функция $F$ вещественна, то

$$
(AU_1, U_2) = \int (FU_1) U_2^* dq_1 \ldots dq_s = \\
= \int U_1 (FU_2)^* dq_1 \ldots dq_s = (U_1, AU_2);
$$

это показывает, что вещественной механической величине вида $\mathfrak{H} = F(q_1, \ldots, q_s)$ всегда соответствует некоторый самосопряжённый оператор.

Перейдём теперь к общему случаю величин вида

$$
\mathfrak{H} = F(q_1, \ldots, q_s, p_1, \ldots, p_s).
$$

Способ определения соответствующего такой величине линейного самосопряжённого оператора нас в дальнейшем нигде интересовать не будет; поэтому мы дадим это направление лишь самые краткие указания. Рассмотрим прежде всего простейший случай $\mathfrak{H} = p_k$ ($1 \leqslant k \leqslant s$). Квантовая механика ставит в соответствие обобщённому импульсу $p_k$ оператор

$$
P_k = -i\hbar \frac{\partial}{\partial q_k},
$$

т. е. полагает

$$
P_k U = -i\hbar \frac{\partial U}{\partial q_k},
$$

где $\hbar$ — постоянная Планка. Линейность этого оператора очевидна. Чтобы показать его самосопряжённость, заметим, что

$$
(P_k U_1, U_2) = \left(-i\hbar \frac{\partial U_1}{\partial q_k}, U_2\right) = \\
= -i\hbar \int \frac{\partial U_1}{\partial q_k} U_2^* dq_1 \ldots dq_s.
$$

Но при фиксированных $q_1, \ldots, q_{k-1}, q_{k+1}, \ldots, q_s$

$$
\int_{-\infty}^{+\infty} \frac{\partial U_1}{\partial q_k} U_2^* dq_k = \left.U_1 U_2^*\right|_{-\infty}^{+\infty} - \int U_1 \frac{\partial U_2^*}{\partial q_k} dq_k = \\
= -\int U_1 \frac{\partial U_2^*}{\partial q_k} dq_k,
$$
если функции $U_1$ и $U_2$, как мы допустим, обращаются в нуль при $q_k = -\infty$ и $q_k = +\infty$. Поэтому

$$(P_k U_1, U_2) = i\hbar \int U_1 \frac{\partial U_2^*}{\partial q_k} dq_1 \ldots dq_s =$$

$$= \int U_1 \left( -i\hbar \frac{\partial U_2^*}{\partial q_k} \right)^* dq_1 \ldots dq_s = (U_1, P_k U_2),$$

что и требовалось доказать. Заслуживает быть отмеченным, что самосопряжённость оператора $P_k$ существенным образом связана с наличием мнимой единицы в качестве множителя в определении этого оператора; вещественный оператор $\hbar \frac{\partial}{\partial q_k}$, как легко проверить, не обладает свойством самосопряжённости.

Если теперь механическая величина $\mathcal{A}$ является произвольной функцией

$$\mathcal{A} = F(q_1, \ldots, q_s, p_1, \ldots, p_s)$$

гамильтоновых переменных данной системы, то естественно возникает мысль определить соответствующий ей оператор $A$ с помощью формулы

$$A = F(Q_1, \ldots, Q_s, P_1, \ldots, P_s),$$

где $Q_k$, $P_k$ означают соответственно операторы, отнесённые нами величинам $q_k$, $p_k$ $(1 \leq k \leq s)$. Этим путём обычно и идёт квантовая механика. Трудность реализации этой, в целом весьма плодотворной мысли заключается лишь в вопросе о том, какой смысл следует придать оператору

$$F(Q_1, \ldots, Q_s, P_1, \ldots, P_s). \quad (3)$$

Даже в самом простом случае, когда $F$ есть многочлен с постоянными коэффициентами, решение этого вопроса отнюдь не самоочевидно. Чтобы составить многочлен (3), надо определить действия сложения и умножения операторов, надо построить алгебру операторов. Вот как строится эта алгебра. Если $A$ и $B$ — операторы, $\alpha$ и $\beta$ — вещественные числа, то под оператором

$$C = \alpha A + \beta B$$
разумеют оператор, определяемый соотношением
\[ CU = aAU + \beta BU .\]

Если операторы A и B — линейные самосопряжённые, то это же самое имеет место и для оператора C. Произведением AB операторов A и B называют оператор C, определяемый соотношением
\[ CU = A (BU), \]
так что результат применения оператора AB совпадает с результатом последовательного применения к той же функции сперва оператора B, а затем оператора A. Умножение операторов, вообще говоря, не коммутативно: ABU не совпадает, вообще говоря, с BAU. Так,
\[ P_k Q_k U = -i h \frac{\partial (U q_k)}{\partial q_k} = -i h \left( U + q_k \frac{\partial U}{\partial q_k} \right), \]
в то время как
\[ Q_k P_k U = q_k \left( -i h \frac{\partial U}{\partial q_k} \right) = -i h q_k \frac{\partial U}{\partial q_k} . \]

Эта некоммутативность умножения операторов делает неоднозначным смысл выражения (3) даже в том случае, когда оно представляет собою многочлен, и тем самым затрудняет выбор оператора, который должен быть поставлен в соответствие данной механической величинае. Произведение двух самосопряжённых операторов будет самосопряжённым оператором лишь при том условии, что сомножители его коммутируют друг с другом. Отметим ещё, что так как каждый оператор, очевидно, коммутирует с самим собой, то целая положительная степень любого самосопряжённого оператора всегда является вполне определённым самосопряжённым оператором.

Мы видим, таким образом, что вопрос о том, какой оператор должен быть поставлен в соответствие данной механической величине A, не всегда может быть решён какими-либо простыми общими методами, и во многих случаях требует специального физического рассмотрения. Но для целей настоящей книги затруднения этого рода не имеют существенного значения; в основном для нас важен лишь самый факт возможности отнесения каждой механической величине некоторого
линейного самосопряжённого оператора, позволяющего найти единообразным и в принципе очень простым путём статистику этой величины в любом состоянии системы. Теперь мы переходим к вопросу о том, как устанавливается эта статистика, если оператор данной величины нам известен.

В случае, когда механическая величина

$$\mathfrak{A} = F(q_1, \ldots, q_s)$$

зависит только от обобщённых координат $q_1, \ldots, q_s$, её математическое ожидание $E_{\mathfrak{A}}$ в состоянии $U$ определяется, как мы знаем [§ 1, (1)], формулой

$$E_{\mathfrak{A}} U = \int F(q_1, \ldots, q_s) |U(q_1, \ldots, q_s)|^2 dq_1 \ldots dq_s,$$

где для простоты мы предполагаем функцию $U$ нормированной. Так как в этом случае величина $\mathfrak{A}$ соответствует оператор $A$, определяемый соотношением

$$AU = F(q_1, \ldots, q_s) U,$$

то мы можем написать

$$E_{\mathfrak{A}} U = \int FUU^* dq_1 \ldots dq_s =$$

$$= \int (AU) U^* dq_1 \ldots dq_s = (AU, U).$$

Этот результат, установленный нами для величина $\mathfrak{A}$ весьма специального вида, квантовая механика распространяет на общий случай в качестве одного из своих основных принципов: если величине $\mathfrak{A}$ поставлен в соответствие линейный самосопряжённый оператор $A$, то математическое ожидание этой величины в состоянии $U$ равно *)

$$E_{\mathfrak{A}} U = (AU, U).$$

(4)

В этом простом общем правиле и содержится, разумеется, реальный смысл того «отнесения» каждой механической величине определённого линейного самосопряжённого оператора,

*) Функция $U$ предполагается нормированной. В общем случае, очевидно, $E_{\mathfrak{A}} U = \frac{(AU, U)}{(U, U)}$. 
о котором всё время идёт речь. Мы видим, что, зная соответствующий данной величине оператор, мы действительно можем определить по единому простому принципу математическое ожидание этой величины в любом состоянии $U$.

Можно думать, что полученный нами результат определяет лишь математическое ожидание величины $\mathcal{A}$, оставляя открытым вопрос о её законе распределения. Если нам известен только оператор величины $\mathcal{A}$, то это действительно так. Но в принципе указанный нами метод позволяет определять и всю статистику механических величин, связанных с данной системой. В самом деле, если мы хотим, например, найти вероятность неравенств

$$a < \mathcal{A} < b,$$

где $a$ и $b$ — данные вещественные числа, и если

$$\mathcal{A} = F(q_1, \ldots, q_s, p_1, \ldots, p_s),$$

то достаточно ввести в рассмотрение другую механическую величину $\mathcal{B}$, определяемую так:

$$\mathcal{B} = \begin{cases} 1, & \text{если } a < F < b, \\ 0, & \text{во всех других случаях.} \end{cases}$$

Эта величина, очевидно, также есть функция переменных $q_1, \ldots, q_s, p_1, \ldots, p_s$ и потому может рассматриваться как механическая величина, связанная с данной системой. Очевидно, с другой стороны, что вероятность неравенств $a < \mathcal{A} < b$, когда система находится в состоянии $U$, равна

$$P_U(a < \mathcal{A} < b) = E_U\mathcal{B}.$$

Если величина $\mathcal{B}$ соответствует линейный самосопряжённый оператор $\mathcal{B}$, то мы поэтому имеем в силу формулы (4):

$$P_U(a < \mathcal{A} < b) = (BU, U)$$

(где снова для простоты мы предполагаем функцию $U$ нормированной).

Мы видим, таким образом, что если мы умеем определить соответствующий ей оператор не только для самой величины $\mathcal{A}$, но и для любой её функции, то формула (4) позволяет найти и весь закон распределения величины $\mathcal{A}$.
Отметим ещё, что в силу самосопряжённости оператора $A$ мы всегда имеем:

$$(AU, U) = (U, AU);$$

с другой стороны, в силу общего свойства скалярных произведений ($\S$ 1, свойство $3^0$)

$$(U, AU) = (AU, U)^*;$$

таким образом, всегда

$$(AU, U) = (AU, U)^*,$$

т. е. математическое ожидание механической величины всегда вещественно. Этот результат уже ясно показывает смысл требования самосопряжённости тех операторов, которые мы ставим в соответствие вещественным механическим величинам. В дальнейшем мы получим ещё более убедительные аргументы в пользу этого требования.

§ 3. Возможные значения механических величин

Пусть механической величине $\mathcal{A}$ соответствует линейный самосопряжённый оператор $A$. В предыдущем параграфе мы видели, как, зная оператор $A$ и операторы, соответствующие различным функциям величины $\mathcal{A}$, можно построить закон распределения этой величины в любом состоянии $U$ данной системы. Иногда этот закон может оказаться таким, что величина $\mathcal{A}$ принимает в состоянии $U$ только одно возможное значение $\alpha$ (с вероятностью 1). В таком случае говорят, что $\mathcal{A}$ имеет в состоянии $U$ достоверное значение. Разумеется, в этом случае и

$$E_U\mathcal{A} = \alpha.$$

Рассмотрим теперь, при каких условиях может наступить это положение вещей. Для того чтобы случайная величина имела единственное (достоверное) возможное значение, необходимо и достаточно, чтобы её дисперсия равнялась нулю. Дисперсия величины $\mathcal{A}$ в состоянии $U$ равна

$$D_U\mathcal{A} = E_U(\mathcal{A} - \alpha)^2.$$
Мы, естественно, считаем на основании принципов, изложенных в предыдущем параграфе, что величина \((\mathcal{A} - \alpha)^2\) соответствует (линейный самосопряжённый) оператор \((\mathcal{A} - \alpha)^2\), так что, если функция \(U\) нормирована,

\[
\mathbf{D}_{\mathcal{A}} = ([\alpha - \alpha]^2 U, U);
\]

в силу самосопряжённости оператора \(\mathcal{A} - \alpha\) это даёт:

\[
\mathbf{D}_{\mathcal{A}} = ([\alpha - \alpha] U, [\alpha - \alpha] U) = \\
= \int ([\alpha - \alpha] U) ([\alpha - \alpha] U)^* dq_1 \ldots dq_s = \\
= \int |[\alpha - \alpha] U|^2 dq_1 \ldots dq_s.
\]

Если \(\mathbf{D}_{\mathcal{A}} = 0\), то в силу неотрицательности подинтегральной функции мы должны иметь тождественно относительно \(q_1, q_2, \ldots, q_s\):

\[
|[\alpha - \alpha] U|^2 = 0,
\]

и следовательно,

\[
[\alpha - \alpha] U = 0,
\]

\[
AU = \alpha U \quad (5)
\]

при любых \(q_1, q_2, \ldots, q_s\). Таким образом, соотношение (5) необходимо для того, чтобы величина \(\mathcal{A}\) имела в состоянии \(U\) достоверное значение \(\alpha\); вместе с тем оно и достаточно для этой цели, в чём мы непосредственно убеждаемся, проводя все расчёты в обратном порядке.

Соотношение (5) показывает, что действие оператора \(\mathcal{A}\) на функцию \(U\) сводится к умножению этой функции на постоянное вещественное число \(\alpha\). В теории операторов называют (отличную от тождественного шума) функцию \(U\), связанную с оператором \(\mathcal{A}\) соотношением вида (5), собственной функцией этого оператора, принадлежащей к собственному значению \(\alpha\). Таким образом, числа, могущие быть достоверными значениями величины \(\mathcal{A}\), являются собственными значениями оператора \(\mathcal{A}\), а состояние, в котором величина \(\mathcal{A}\) имеет достоверное значение \(\alpha\), описывается одной из собственных функций \(U\) оператора \(\mathcal{A}\), принадлежащих к собственному значению \(\alpha\).
§ 3. Возможные значения механических величин

Если рассматривать соотношение (5) как уравнение, определяющее функцию \( U \), то в типичных для квантовой механики случаях это будет уравнение в частных производных, так как оператор \( A \) составляется из операторов \( Q_k, P_k \) примерно так, как величина \( \mathcal{A} \) составлена из переменных \( q_k, p_k \). Искомые решения \( U \) этого уравнения должны быть однозначно определены во всём пространстве переменных \( q_1, q_2, \ldots ; q_s \), и должен существовать интеграл

\[
\int |U|^2 dq_1 \ldots dq_s
\]

взятый по всему пространству; кроме того, они обязаны удовлетворять некоторым требованиям общего характера — быть непрерывными, а иногда и обладать частными производными до того или иного порядка. Несмотря на кажущуюся широту этих требований, вообще говоря, оказывается, что соответствующие им решения существуют далеко не для всех значений параметра \( \alpha \); те значения этого параметра, для которых такие решения существуют, и будут собственными значениями оператора \( A \); их совокупность называют спектром этого оператора. Спектр оператора может быть как дискретным (т. е. состоящим из отдельных изолированных чисел), так и сплошным (т. е. охватывающим целый отрезок или даже всю числовую прямую); возможны и случаи спектров комбинированных, т. е. местами сплошных, а местами дискретных. В соответствии с задачами настоящей книги мы будем в дальнейшем рассматривать только случай, когда (дискретный) спектр оператора представляет собою неограниченно возрастющую последовательность чисел

\[
a_1 < a_2 < \ldots < a_n < \ldots \quad (\lim_{n \to \infty} a_n = +\infty)\]

(6)

Допустим, что при некотором измерении величины \( \mathcal{A} \) мы получили для неё значение \( \alpha \); непосредственно после этого измерения система будет тогда, как учит квантовая механика, находиться в таком состоянии, в котором величина \( \mathcal{A} \) имеет \( \alpha \) достоверным значением, так что вторичное измерение величины \( \mathcal{A} \) на той же системе, если только оно непосредственно следует за первым, уже обязательно должно иметь своим результатом \( \alpha \). Таким образом, любое число \( \alpha \), которое вообще может оказаться результатом измерения величины \( \mathcal{A} \) (в каком
бы то ни было состоянии), может вместе с тем быть (в некотором определённом состоянии) и достоверным значением этой величины; и следовательно, должно принадлежать спектру оператора $A$; этот спектр содержит, таким образом, все числа, могущие служить результатами измерения величины $A$ в каких бы то ни было состояниях. В рассматриваемом нами случае дискретного спектра (6) эта дискретность нигде специально не постулируется, а вытекает как следствие из требований весьма общего характера, налагаемых на собственные функции, подобно тому как, например, в теории колеблющейся струны дискретность ряда возможных колебаний вытекает из налагаемых граничных условий. Определение всех возможных значений механической величины путём построения спектра соответствующего ей оператора и называют обычно «квантованием» этой величины.

Мы должны будем теперь рассмотреть некоторые основные свойства собственных функций линейного самосопряжённого оператора $A$, ограничиваясь случаем, когда спектр имеет вид (6). Разберёмся сперва в совокупности собственных функций, принадлежащих к одному и тому же собственному значению $a_k$. В силу линейности оператора $A$, если $U_{k1}$ и $U_{k2}$ — две различные собственные функции, принадлежащие к собственному значению $a_k$, то есть $\lambda_1 U_{k1} + \lambda_2 U_{k2}$, где $\lambda_1$ и $\lambda_2$ — любые комплексные числа, не обращающиеся одновременно в нуль, будет собственной функцией, принадлежащей к тому же собственному значению $a_k$. Это означает, что рассматриваемая нами совокупность собственных функций образует линейное многообразие $L_k$. В тех случаях, с которыми мы будем иметь дело в этой книге, многообразие $L_k$ всегда имеет конечное число измерений $m$; это значит, что многообразие $L_k$ содержит $m$ линейно независимых функций $U_{k1}, U_{k2}, \ldots, U_{km}$, но любые $m + 1$ функций этого многообразия линейно зависимы. Следовательно, это является то, что любая функция $U$ многообразия $L_k$ может быть единственным образом представлена в виде

$$U = \sum_{j=1}^{m} \lambda_j U_{kj},$$

где $\lambda_j$ — комплексные числа. Функции $U_{kj}$ ($j = 1, 2, \ldots, m$) образуют линейный базис многообразия $L_k$. Этот базис может
бывать выбран различными способами. В частности, хорошо известные процессы ортогонализации всегда позволяют построить базис, состоящий из взаимно ортогональных функций $U_k^j$ (функции $U_1^1$ и $U_2^2$ взаимно ортогональны, если $(U_1, U_2) = 0$).

Базис $U_k^j$ (1 = 1, 2, ..., m) называется нормированным, если $(U_k^j, U_k^j) = 1$ (1 = 1, m), т. е. если нормированы все функции, составляющие этот базис.

Пусть $V_{k1}, V_{k2}, ..., V_{km}$ — любые m собственных функций оператора $A$, принадлежащих к тому же собственному значению $a_k$ этого оператора, что и функции $U_k^j$. Тогда, как мы знаем,

$$ V_{kg} = \sum_{j=1}^{m} \lambda_g^j U_k^j \quad (1 \leq g \leq m), \quad (7) $$

где $\lambda_g^j$ — комплексные числа. Отсюда при 1 = 1, m, 1 = h = m

$$ (V_{kg}^h, V_{kh}) = \left( \sum_{j=1}^{m} \lambda_g^j U_k^j, \sum_{i=1}^{m} \lambda_h^i U_k^i \right) = $$

$$ = \sum_{j=1}^{m} \sum_{i=1}^{m} \lambda_g^j \lambda_h^i (U_k^j, U_k^i). $$

Так как базис $U_k^j$ — ортогональный и нормированный, то

$$(U_k^j, U_k^i) = \delta_{ji},$$

где

$$ \delta_{ji} = \begin{cases} 1 & (j = i), \\ 0 & (j \neq i). \end{cases} $$

Поэтому мы получаем:

$$ (V_{kg}^h, V_{kh}) = \sum_{j=1}^{m} \sum_{i=1}^{m} \delta_{ij} \lambda_g^j \lambda_h^i = \sum_{j=1}^{m} \lambda_g^j \lambda_h^j. $$

Для того чтобы функции $V_{kg}$, определяемые соотношением (7), были ортогональны и нормированы, необходимо и достаточно иметь:

$$ \sum_{j=1}^{m} \lambda_g^j \lambda_h^j = \begin{cases} 1 & (g = h), \\ 0 & (g \neq h), \end{cases} $$

$$ (V_{kg}^h, V_{kh}^h) = \sum_{j=1}^{m} \sum_{i=1}^{m} \delta_{ij} \lambda_g^j \lambda_h^i = \sum_{j=1}^{m} \lambda_g^j \lambda_h^j. $$
или, что то же,

\[ \sum_{j=1}^{m} \lambda_{gj} \delta_{jh} = \delta_{gh} \quad (1 \leq g \leq m, \ 1 \leq h \leq m). \]  

(8)

Матрица, состоящая из комплексных чисел \( \lambda_{gh} \) (\( 1 \leq g \leq m, \ 1 \leq h \leq m \)) (а также и осуществляемое этой матрицей линейное преобразование (7)), называется унитарной, если для неё выполняются соотношения (8). Но в случае, когда функции \( V_{kg} \) ортогональны и нормированы, они тем самым и линейно независимы *) и, следовательно, образуют ортогональный и нормированный базис многообразия \( L_k \). Таким образом, если функции \( U_{kj} \) образуют ортогональный нормированный базис многообразия \( L_k \), то все другие такие базисы \( V_{kg} \) (\( 1 \leq g \leq m \)) этого многообразия мы получим, подвергая систему функций \( U_{kj} \) (\( 1 \leq j \leq m \)) всевозможным унитарным преобразованиям (7).

Число измерений \( m \) многообразия \( L_k \) является важнейшей характеристикой собственного значения \( \alpha_k \) и называется его кратностью или степенью вырождения; собственное значение \( \alpha_k \) вырождено, если \( m > 1 \), и не вырождено, если \( m = 1 \). С физической стороны \( m \) означает число линейно независимых состояний системы, в которых величина \( \Omega \) имеет достоверное значение \( \alpha_k \).

Пусть теперь \( U_1 \) и \( U_2 \) — две собственные функции оператора \( A \), принадлежащие к различным собственным значениям \( \alpha_{k_1} \) и \( \alpha_{k_2} \), так что

\[ AU_1 = \alpha_{k_1} U_1, \ AU_2 = \alpha_{k_2} U_2. \]  

(9)

В силу самосопряженности оператора \( A \) мы имеем:

\[ (AU_1, U_2) = (U_1, AU_2), \]

*) В самом деле, если бы, например, мы имели

\[ V_{k1} = a_2 V_{k2} + \ldots + a_m V_{km}, \]

где \( a_l \) — числа, то мы нашли бы

\[ (V_{k1}, V_{k1}) = \sum_{l=2}^{m} a_l (V_{kl}, V_{k1}) = 0, \]

что невозможно.
откуда в силу (9)
\[(a_{k_1}U_1, U_2) = (U_1, a_{k_3}U_2);\]
tак как \(a_{k_1}\) и \(a_{k_3}\) вещественные, то отсюда в силу свойства 
2° (§ 1) скалярных произведений
\[a_{k_1}(U_1, U_2) = a_{k_3}(U_1, U_2);\]
а так как \(a_{k_1} \neq a_{k_3}\), то
\[(U_1, U_2) = 0.\]

Это показывает, что две собственные функции оператора \(A\), принадлежащие к различным собственным значениям, всегда взаимно ортогональны.

Если мы теперь для каждого собственного значения \(a_k\) оператора \(A\) выберем какой-нибудь ортогональный линейный базис, то вся совокупность полученных таким образом собственных функций оператора \(A\) образует, очевидно, бесконечную (счётную) ортогональную систему элементов \(U\) комплексного гильбертово пространства. В наиболее простых и важных случаях, и в частности — во всех тех случаях, с которыми нам придётся иметь дело, эта ортогональная система оказывается полной (или, как ещё говорят, замкнутой); это означает, что всякий элемент \(U\), ортогональный всем элементам полученной нами системы, должен быть тождественным нулю.
Полнота нашей системы собственных функций имеет очень большое значение для физики. Если обозначить через \(U_1, U_2, \ldots, U_n, \ldots\) элементы полной ортогональной системы, перенумерованные в каком угодно порядке, то любой элемент \(U\) комплексного гильбертово пространства может быть представлен в виде ряда
\[\sum_{n=1}^\infty c_n U_n,\]
который сходится к \(U\) «в среднем» (это означает, что, полагая \(\sum_{k=1}^\infty c_k U_k = s_n\), мы имеем:
\[(U - s_n, U - s_n) \to 0 \quad (n \to \infty)).\]
Здесь \( c_n \) означают постоянные комплексные числа; в силу взаимной ортогональности функций \( U_n \) мы легко находим:

\[
    c_n = \frac{(U, U_n)}{(U_n, U_n)};
\]

если функции \( U_n \) нормированы, то

\[
    c_n = (U, U_n);
\]

dалее,

\[
    (U, U) = \sum_{n=1}^{\infty} c_n (U_n, U) = \sum_{n=1}^{\infty} c_n (U, U_n)^* = \sum_{n=1}^{\infty} |c_n|^2;
\]

в частности, если функция \( U \) тоже нормирована, то

\[
    \sum_{n=1}^{\infty} |c_n|^2 = 1.
\]

Мы видели в предыдущем параграфе, что, умев сопоставить каждой механической величине определённый линейный самосопряжённый оператор, мы принципиально можем определить статистику (закон распределения) этой величины в любом состоянии \( U \); теперь мы можем уже показать конкретно, как это делается, по крайней мере для величин с дискретным спектром. Пусть величина \( \mathcal{A} \) соответствует оператор \( A \) с дискретным спектром (6); так как никаких других значений, кроме чисел \( a_k \), величина \( \mathcal{A} \) принимать не может, то статистика этой величины в каком-либо состоянии \( U \) будет полностью определена, если мы для каждого из чисел \( a_k \) найдём вероятность

\[
    P_U(\mathcal{A} = a_k)
\]

tого, что в состоянии \( U \) величина \( \mathcal{A} \) окажется равной \( a_k \).

Обозначим, как мы это делали в § 2, через \( \mathcal{B} \) величину, равную 1, если \( \mathcal{A} = a_k \), и равную нулью, если \( \mathcal{A} \neq a_k \) (так что \( \mathcal{B} \) есть функция от \( \mathcal{A} \)), и пусть \( \mathcal{B} \) — линейный самосопряжённый оператор, соответствующий величине \( \mathcal{B} \); тогда вероятность равенства \( \mathcal{A} = a_k \) совпадает с вероятностью равенства \( \mathcal{B} = 1 \), очевидно, равной математическому ожиданию величины \( \mathcal{B} \).

Таким образом (предполагая функцию \( U \) нормированной), мы имеем:

\[
    P_U(\mathcal{A} = a_k) = (BU, U).
\]
Пусть теперь \( U_1, U_2, \ldots, U_n, \ldots \) — построенная нами полная ортогональная система нормированных собственных функций оператора \( A \) и пусть

\[
U = \sum_{n=1}^{\infty} c_n U_n, \quad c_n = (U, U_n) \quad (n \geq 1).
\]

Тогда

\[
(BU, U) = \left( \sum_{n=1}^{\infty} c_n BU_n, \sum_{n=1}^{\infty} c_n U_n \right).
\]

Но каждая из функций \( U_n \) есть собственная функция оператора \( A \), т. е. величина \( A \) имеет в каждом из состояний \( U_n \) некоторое достоверное значение; отсюда следует, что и величина \( B \), будучи функцией от \( A \), должна иметь достоверное значение в каждом из состояний \( U_n \); это значение равно 1 в тех состояниях, где \( A = \alpha_k \), и нуль в остальных состояниях; таким образом, каждая из функций \( U_n \) есть собственная функция оператора \( B \); соответствующее собственное значение этого оператора есть 1, если \( AU_n = \alpha_k U_n \), и нуль во всех остальных случаях. Мы можем поэтому написать

\[
BU_n = \beta_n U_n,
\]

где \( \beta_n = 1 \), если \( AU_n = \alpha_k U_n \), и \( \beta_n = 0 \) в противном случае. Поэтому мы находим:

\[
(BU, U) = \left( \sum_{n=1}^{\infty} c_n \beta_n U_n, \sum_{n=1}^{\infty} c_n U_n \right).
\]

Пользуяя свойствами скалярных произведений, ортогональностью системы функций \( U_n \) и нормированностью этих функций, мы отсюда легко получаем:

\[
P_U(A = \alpha_k) = (BU, U) = \sum_{n=1}^{\infty} \beta_n \left| c_n \right|^2 =
\]

\[
= \sum^{(k)} \left| c_n \right|^2 = \sum^{(k)} \left| (U, U_n) \right|^2,
\]

где \( \sum^{(k)} \) означает, что суммирование распространяется на все значения \( n \), для которых собственная функция \( U_n \) принадлежит к собственному значению \( \alpha_k \). Отсюда теорема сложения вероятностей уже непосредственно даёт:

\[
P_U(a < A < b) = \sum_{(a, b)} \left| c_n \right|^2 = \sum_{(a, b)} \left| (U, U_n) \right|^2,
\]
где суммирование распространяется на все собственные функции $U_n$ оператора $A$, принадлежащие к собственным значениям, заключенным между $a$ и $b$. Мы получаем таким образом вполне определённое правило для нахождения закона распределения величины $A$ в состоянии $U$: для этого нужно только разложить функцию $U$ в ряд по полной ортогональной системе нормированных собственных функций $U_n$ оператора $A$; вероятность неравенств $a < A < b$ определяется тогда как сумма квадратов модулей коэффициентов этого разложения при всех тех функциях $U_n$, которые принадлежат к собственным значениям, заключенным между $a$ и $b$.

Примем ещё полную ортогональную нормированную систему собственных функций оператора $A$ к доказательству одного элементарного предложения, которое понадобится нам в дальнейшем: если оператор $A$ имеет спектр (6), то спектр оператора $A^2$ состоит из чисел

$$
\alpha_1^2, \alpha_2^2, \ldots, \alpha_k^2, \ldots,
$$

количество собственных функций оператора $A^2$, принадлежащих к собственному значению $\alpha_k^2$, совпадает с количеством собственных функций оператора $A$, принадлежащих к собственному значению $\alpha_k$.

В самом деле, если $U$ — собственная функция оператора $A$, принадлежащая к собственному значению $\alpha_k$, то

$$
A^2U = A(AU) = A(\alpha_kU) = \alpha_kAU = \alpha_k^2U,
$$

т. е. $U$ есть собственная функция оператора $A^2$, принадлежащая к собственному значению $\alpha_k^2$.

Пусть теперь, обратно, $U$ есть собственная функция оператора $A^2$, принадлежащая к некоторому собственному значению $\beta$ этого оператора.

Разложим функцию $U$ в ряд по полной ортогональной системе собственных функций оператора $A$ (которые, как мы только что убедились, служат вместе с тем и собственными функциями оператора $A^2$):

$$
U = \sum_{k=1}^{\infty} c_k U_k,
$$

где

$$
c_k = (U, U_k) \quad (k = 1, 2, \ldots).
$$
Так как $U$ и $U_k$ — собственные функции одного и того же оператора $A^2$, то мы можем иметь $c_k \neq 0$ только при $\beta = a_k^2$; это показывает, что собственное значение $\beta$ оператора $A^2$ необходимо должно совпадать с одним из чисел $a_k^2$, т. е. что спектр оператора $A^2$ действительно даётся совокупностью чисел $a_k^2$. Пусть теперь $\beta = a_k^2$ и пусть $U_{k1}, U_{k2}, \ldots, U_{km}$ — собственные функции оператора $A$ выбранной нами ортогональной системы, принадлежащие к собственному значению $a_k$; тогда в разложении (10) могут быть отличными от нуля только коэффициенты при функциях $U_{ki}$ ($1 \leq i \leq m$), так что

$$U = \sum_{i=1}^{m} c_{ki} U_{ki},$$

и следовательно,

$$AU = a_k U,$$

т. е. $U$ есть собственная функция оператора $A$, чем наше утверждение и доказано.

§ 4. Изменение состояния системы во времени

В классической механике состояние системы описывается значениями всех $2s$ гамильтоновых переменных $q_1, q_2, \ldots, q_s, p_1, p_2, \ldots, p_s$. Чтобы знать, как изменяется это состояние во времени, надо задать все $2s$ гамильтоновых переменных как функции времени. То обстоятельство, что значениями гамильтоновых переменных в какой-либо начальный момент времени $t_0$ однозначно определяются их значения и в любой другой (предыдущий или последующий) момент времени $t$, является выражением принципа причинности в классической механике.

В квантовой механике состояние системы описывается некоторой функцией $U(q_1, q_2, \ldots, q_s)$ обобщённых координат. Изменение состояния во времени будет известно, если будет задана зависимость функции $U$ от времени, т. е. если эта функция будет задана как функция $s + 1$ переменных

$$U = U(q_1, \ldots, q_s, t);$$
выражением принципа причинности в квантовой механике является требование, чтобы видом функции $U$ в некоторый начальный момент времени $t_0$ однозначно определялся её вид и в любой другой (предыдущий или последующий) момент времени $t$. Так как функция $U$ в квантовой механике определяет собою не значения, а статистику (законы распределения) механических величин, то физически это требование означает, что статистикой механических величин в начальный момент времени $t_0$ однозначно должна определяться их статистика и в любой другой момент времени $t$.

В классической механике закон движения (т. е. выражение гамильтоновых переменных в функциях времени) устанавливается решением системы «уравнений движения»; эти уравнения движения в случае гамильтоновых переменных принимают известную, замечательную своей простотой и симметрией «каноническую» форму, в которой основную роль играет так называемая функция Гамильтона

$$H(q_1, \ldots, q_s, p_1, \ldots, p_s),$$

выражение которой в случае, когда действующие на систему силы не зависят от времени, совпадает с выражением полной энергии системы:

$$H = T + V,$$

где $T$ — кинетическая, а $V$ — потенциальная энергия системы.

В квантовой механике функция Гамильтона, как и всякой механической величине, соответствует некоторый линейный самосопряжённый оператор $H$ — оператор полной энергии или, как его иначе называют, г а м и л ь т о н а н. Мы увидим, что этот оператор имеет в квантовой механике для определения эволюции состояния системы во времени столь же решающее значение, как функция Гамильтона в классической механике.

Выражение оператора $H$ в простейших случаях нетрудно правильно выбрать по аналогии с выражением функции Гамильтона в классической механике. Если наша система есть элементарная частица, то в качестве гамильтонов переменных проще всего выбрать три её декартовых координаты $x$, $y$, $z$ и соответствующие им компоненты импульса $p_x$, $p_y$, $p_z$. Кинетическая энергия в этом случае будет:

$$T = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2),$$
где $m$ — масса частицы, а потенциальная энергия $V(x, y, z)$ будет зависеть только от координат частицы. Для гамильтоновой функции это даёт выражение

$$H = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2) + V(x, y, z).$$

Переходя к построению оператора $H$, мы вспомним, что величинам $p_x$, $p_y$, $p_z$ соответствуют операторы $-i\hbar \frac{\partial}{\partial x}$, $-i\hbar \frac{\partial}{\partial y}$, $-i\hbar \frac{\partial}{\partial z}$; квадратам же этих величин мы, в силу принятой нами в § 2 алгебры операторов, должны поставить в соответствие операторы

$$-\hbar^2 \frac{\partial^2}{\partial x^2}, -\hbar^2 \frac{\partial^2}{\partial y^2}, -\hbar^2 \frac{\partial^2}{\partial z^2}.$$ 

Таким образом, кинетической энергии $T$ будет соответство- вать оператор

$$-\frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) = -\frac{\hbar^2}{2m} \nabla,$$

где $\nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ — так называемый оператор Лапласа.

А так как потенциальной энергии $V(x, y, z)$ соответствует, как мы знаем, просто оператор $V$ умножения на $V(x, y, z)$, то мы, естественно, принимаем в качестве гамильтониана оператор

$$H = -\frac{\hbar^2}{2m} \nabla + V.$$

(11)

Столь же просто строится оператор $H$ в несколько более сложном случае, когда наша система состоит из нескольких таких бесструктурных материальных частиц, в предположении отсутствия взаимодействия между частицами. В этом случае полная энергия системы есть просто сумма полных энергий составляющих эту систему частиц, и мы, естественно, принимаем, что и гамильтониан системы равен просто сумме гамильтонианов составляющих её частиц, построенных согласно формуле (11).

Однако столь несомненное определение оператора $H$ получает лишь в простейших случаях. Даже в предположении,
что система состоит из одной бесструктурной частицы, дело обстоит столь просто лишь при выборе декартовых координат в качестве гамильтоновских переменных. В других координатных системах кинетическая энергия получает вид квадратичной формы относительно обобщённых импульсов, коэффициенты которой зависят от обобщённых координат; а в этом случае, как мы знаем, переход к соответствующим операторам уже не является столь однозначно очевидным.

Мы переходим теперь к закономерностям, определяющим в квантовой механике изменение состояния системы во времени. Условимся обозначать для краткости через $U(q, t)$ функцию $U(q_1, \ldots, q_s, t)$, описывающую состояние системы в момент времени $t$. Если, как мы этого требуем, заданием функции $U(q, t_0)$ однозначно определяется функция $U(q, t)$ для любого $t$, то тем самым, в частности, и функция $\frac{\partial U}{\partial t}$ при $t = t_0$ является однозначно определённой (при условии, разумеется, что эта производная существует, что мы здесь должны постулировать). Таким образом, так как момент $t_0$ может быть выбран произвольно, мы можем сказать, что для любого момента $t$ каждой функции $U$ однозначно соответствует некоторая другая функция $\frac{\partial U}{\partial t}$: Но такое однозначное отнесение каждой функции $U$ некоторой другой функции мы условились называть оператором. Таким образом, функция $\frac{\partial U}{\partial t}$ есть результат воздействия на функцию $U$ некоторого оператора. Этот оператор может, вообще говоря, быть различным для различных моментов времени, вследствие чего нам будет удобно обозначить его через $A_t$, так что

$$\frac{\partial}{\partial t} U(q, t) = A_t U(q, t). \tag{12}$$

Остаётся определить оператор $A_t$. Эвристическим путём, анализируя ряд простых частных примеров и обобщая полученные при этом результаты (с привлечением и ряда теоретических соображений), удалось установить универсальный вид этого оператора; оказалось, что надо положить

$$A_t = -\frac{i}{\hbar} \hat{H},$$
где $h$ — постоянная Планка, а $H$ — гамильтониан системы. Это правило имеет совершенно общую значимость; в частности, оно остаётся верным и при том условии, что оператор $H$ зависит от времени (на систему действуют переменные сильы) и не является уже, вообще говоря, оператором полной энергии. Однако мы во всём дальнейшем будем предполагать его не зависящим от времени и, следовательно, соответствующим той механической величине, которую мы называем полной энергией системы. Уравнение (12) получает вид

$$ih \frac{\partial U(q, t)}{\partial t} = HU(q, t),$$

или, кратко,

$$ih \frac{\partial U}{\partial t} = HU. \quad (13)$$

Это важнейшее для квантовой механики уравнение, вполне аналогичное по своей роли системе уравнений движения классической механики, обычно называют уравнением Шредингера. Оно определяет функцию $U(q_1, \ldots, q_s, t)$ как решение некоторого уравнения в частных производных; мы видим, что относительно $t$ это всегда уравнение первого порядка, относительно же обобщённых координат $q_k$ оно будет иметь второй порядок, так как оператор кинетической энергии, входящий в состав $H$, содержит (как мы видели выше) операторы $\frac{\partial^2}{\partial q_k^2}$.

В качестве первого применения уравнения Шредингера мы докажем теперь одну важную вспомогательную теорему, которая понадобится нам в дальнейшем. Условимся называть нормой функции $U$ величину

$$(U, U) = \int |U|^2 \, dq_1 \ldots dq_s,$$

так что, в частности, для нормированных функций $U$ норма равна единице. Покажем, что для решений уравнения Шредингера норма остаётся неизменной во времени (т. е. функционал $(U, U)$ есть интеграл уравнения Шредингера); в частности, если функция $U$, описывающая состояние системы, в какой-либо момент времени $t_0$ нормирована ($(U, U) = 1$), то она будет нормированной и в любой другой (предшествующей или следующей) момент времени $t$. 
Чтобы в этом убедиться, будем исходить из уравнения (13), где \( U = U(q_1, q_2, \ldots, q_s, t) \); переход к комплексно-сопряжённым величинам даёт:

\[
-i\hbar \frac{\partial U^*}{\partial t} = (HU)^*.
\]  

(14)

Помножим уравнения (13) и (14) соответственно на \( U^* \) и \( U \) и вычитем второе из первого; мы находим:

\[
-i\hbar \left( U^* \frac{\partial U}{\partial t} + U \frac{\partial U^*}{\partial t} \right) = i\hbar \frac{\partial |U|^2}{\partial t} = (HU)U^* - U(HU)^*;
\]

интегрируя же это соотношение по всему конфигурационному пространству, мы получаем:

\[
i\hbar \frac{d}{dt} (U, U) = (HU, U) - (U, HU);
\]

но правая часть равна нулю в силу самосопряжённости оператора \( H \); поэтому

\[
\frac{d}{dt} (U, U) = 0,
\]

что и требовалось доказать.

§ 5. Стационарные состояния. Закон сохранения энергии

В интересующих нас случаях спектр оператора \( H \) полной энергии системы всегда будет, как мы уже говорили, представлять собою бесконечно возрастющую последовательность чисел вида (6) § 3, так что мы можем выбрать линейный ортогональный базис нормированных собственных функций оператора \( H \):

\[
U_1, U_2, \ldots, U_n, \ldots;
\]

пусть функция \( U_n \) принадлежит к собственному значению \( E_n \) оператора \( H \), так что

\[
HU_n = E_n U_n \quad (n = 1, 2, \ldots)
\]  

(15)

(при этом, вообще говоря, в силу возможности выражения среди чисел \( E_n \) могут быть и равны между собой).

Функция \( U(q, t) \), служащая решением уравнения Шредингера, может тогда быть разложена в ряд по функциям \( U_n \):
коэффициенты этого разложения будут, разумеется, различными для различных моментов времени $t$, так что мы можем написать

$$U(q, t) = \sum_{n=1}^{\infty} a_n(t) U_n(q),$$

где $a_n(t)$ — комплексные функции от $t$. Подставляя это выражение в уравнение Шредингера (13) и предполагая, как мы условились, оператор $H$ не зависящим от времени, мы находим в силу (15):

$$i\hbar \sum_{n=1}^{\infty} \frac{d}{dt} a_n U_n(q) = \sum_{n=1}^{\infty} a_n (t) H U_n(q) = \sum_{n=1}^{\infty} E_n a_n (t) U_n(q).$$

В силу единственности разложений по функциям $U_n(q)$ мы можем приравнять коэффициенты при $U_n(q)$ в левой и правой части:

$$i\hbar \frac{d}{dt} a_n = E_n a_n (t) \quad (n = 1, 2, \ldots),$$

откуда легко находим:

$$a_n(t) = c_n e^{-\frac{iE_n t}{\hbar}},$$

и следовательно,

$$U(q, t) = \sum_{n=1}^{\infty} c_n e^{-\frac{iE_n t}{\hbar}} U_n(q),$$

где $c_n$ — постоянные комплексные числа. Так как функции $U_n(q)$ мы предположили нормированными, то отсюда:

$$(U, U) = \sum_{n=1}^{\infty} |c_n|^2;$$

посому, если мы хотим, чтобы функция $U(q, t)$ была нормированной, мы должны иметь:

$$\sum_{n=1}^{\infty} |c_n|^2 = 1.$$  

Таким образом, общее нормированное решение уравнения Шредингера может быть представлено в виде (16), где ком-
плексные постоянные $c_n$ в целях нормировки должны быть связаны соотношением (17); обратно, если $c_n$ — произвольные комплексные числа, удовлетворяющие соотношению (17), то формула (16) даёт нам некоторое нормированное решение уравнения Шредингера.

В частности, нормированными решениями уравнения Шредингера служат функции

$$e^{-\frac{iE_nt}{\hbar}} U_n(q) \quad (n = 1, 2, \ldots);$$

(18)

состояния системы, описываемые этими функциями, называются обычно стационарными. Это наименование оправдывается тем, что если состояние системы описывается функцией вида (18), то состояния этой системы в моменты $t_1$ и $t_2$ описываются функциями, отличающимися друг от друга только постоянным множителем, а потому статистика любой механической величины будет, как мы видели в § 5, одной и той же в два любых момента времени; другими словами, в стационарном состоянии статистика механических величин не меняется с течением времени. А так как эта статистика есть всё, что мы в квантовой механике можем извлечь из знания состояния системы, то мы должны заключить, что в случае, когда эволюция системы описывается функцией вида (18), физическое состояние этой системы с течением времени остаётся неизменным.

Но функции вида (18) являются собственными функциями оператора $H$ полной энергии и, следовательно, описывают такие состояния системы, в которых её полная энергия имеет некоторое достоверное значение. Мы приходим, таким образом, к тому важному выводу, что если полная энергия системы в некоторый момент времени имеет достоверное значение (например, если в этот момент как раз произведено измерение полной энергии), то физическое состояние системы в её дальнейшей эволюции остаётся неизменным. Этот вывод особенно интересен в том отношении, что мы не имеем ничего ему аналогичного в области классической механики; в самом деле, в классической механике, конечно, полная энергия всегда имеет достоверное значение и с течением времени (если внешние силы остаются постоянными) не изменяется; но, эволюционируя с течением времени согласно уравнениям движения,
Гамильтоновы переменные непрестанно изменяют свои значения; а так как совокупностью этих значений (или, что то же, точкой фазового пространства) в классической механике определяется состояние системы, то и это состояние претерпевает непрестанное изменение. Различные «эргодические» теоремы или гипотезы стремятся даже установить, что изменение это носит весьма широкий характер, что состояние системы с течением времени подходит как угодно близко к любому состоянию, совместимому с данным значением полной энергии. Мы видим теперь, что в квантовой механике эргодические поступали или теоремы подобного рода совершенно невозможны, по крайней мере в таких состояниях, в которых полная энергия имеет некоторое достоверное значение; а между тем, статистическая термодинамика в первую очередь имеет дело именно с такими состояниями.

Теперь мы должны обратиться к вопросу о том, какие механические величины в квантовой механике играют роль интегралов движения. В классической механике мы называем интегралом уравнений движения такую функцию гамильтоновых переменных, значение которой в силу уравнений движения остаётся неизменным с течением времени. В квантовой механике уравнение движения служит уравнение Шредингера; определяемая им функция \( U(q, t) \) позволяет находить не значения механических величин, а их законы распределения (их статистику); поэтому уравнение Шредингера может иметь своим следствием неизменность во времени не значения механической величины, которого оно вообще не определяет, а самое большее — её закона распределения. В связи с этим мы в квантовой механике, естественно, должны назвать интегралом движения такую механическую величину, закон распределения которой не меняется с течением времени.

Математический аппарат квантовой механики даёт для интегралов движения характеристический признак замечательной принципиальной простоты, выражающий следующим предложением.

Теорема. Для того чтобы механическая величина \( \mathbf{A} \) была интегралом движения, необходимо и достаточно, чтобы соответствующий ей оператор \( \mathbf{A} \) коммутировал с оператором \( \mathbf{H} \) полной энергии системы, т. е. чтобы было \( \mathbf{AH} = \mathbf{HA} \).
Мы ограничиваем доказательством достаточности этого признака, так как необходимость его нам в дальнейшем не понадобится. С этой целью докажем следующее весьма общее вспомогательное предложение.

Лемма. Пусть $A$ и $B$ — два линейных самосопряженных оператора с дискретными спектрами, и пусть $AB = BA$; тогда существует такая полная ортогональная система собственных функций оператора $A$, все функции которой в то же время являются собственными функциями оператора $B$.

Доказательство леммы. Мы будем исходить из произвольной полной ортогональной системы собственных функций оператора $A$. Пусть $\alpha$ — любое собственное значение оператора $A$ и $U$ — любая принадлежащая к нему собственная функция этого оператора, так что

$$AU = \alpha U.$$  

Если $AB = BA$, то

$$ABU = BAU = B(\alpha U) = \alpha BU,$$

т. е. вместе с $U$ и функция $BU$ есть собственная функция оператора $A$, принадлежащая к тому же собственному значению $\alpha$.

Пусть теперь собственное значение $\alpha$ оператора $A$ имеет кратность (степень вырождения) $m$, и пусть в избранной нами полной ортогональной системе к этому значению принадлежат собственные функции $U_1, U_2, \ldots, U_m$. Мы только что показали, что функция $BU_k$ $(k = 1, 2, \ldots, m)$ также есть собственная функция оператора $A$, принадлежащая к собственному значению $\alpha$: следовательно, мы должны иметь:

$$BU_k = \sum_{i=1}^{m} a_{ik} U_i \quad (k = 1, 2, \ldots, m), \quad (19)$$

где $a_{ik}$ — некоторые комплексные числа.

Пусть теперь

$$V = \sum_{k=1}^{m} c_k U_k$$

— линейная комбинация функций $U_k$ с произвольными комплексными коэффициентами $c_k$, так что $V$ есть также собственная
функция оператора A, принадлежащая к собственному значению a. Попытаемся выбрать эти коэффициенты с таким расчетом, чтобы функция V оказалась вместе с тем собственной функцией оператора V. Для этого мы должны иметь:

\[ V \sum_{k=1}^{m} c_k U_k = \beta \sum_{k=1}^{m} c_k U_k, \quad (20) \]

gде \( \beta \) — некоторое комплексное число. Но в силу линейности оператора V и формул (19)

\[ V \sum_{k=1}^{m} c_k U_k = \sum_{k=1}^{m} c_k B U_k = \sum_{k=1}^{m} c_k \sum_{i=1}^{m} a_{ki} U_i, \]

или, переставляя наименование индексов суммирования,

\[ V \sum_{k=1}^{m} c_k U_k = \sum_{l=1}^{m} c_l \sum_{k=1}^{m} a_{kl} U_k = \sum_{k=1}^{m} \left( \sum_{l=1}^{m} c_l a_{kl} \right) U_k. \]

Поэтому соотношение (20) даёт:

\[ \sum_{k=1}^{m} \left\{ \sum_{l=1}^{m} c_l a_{kl} - \beta c_k \right\} U_k = 0, \]

откуда в силу линейной независимости функции \( U_1, \ldots, U_m \)

\[ \sum_{l=1}^{m} c_l a_{kl} - \beta c_k = 0 \quad (k = 1, 2, \ldots, m). \]

Полагая, как прежде,

\[ \delta_{ik} = \begin{cases} 1 & (i = k), \\ 0 & (i \neq k), \end{cases} \]

мы можем переписать полученную систему уравнений в виде

\[ \sum_{l=1}^{m} (a_{kl} - \beta \delta_{kl}) c_l = 0 \quad (k = 1, 2, \ldots, m). \]

Таким образом, для определения коэффициентов \( c_l \) мы получаем систему однородных линейных уравнений; чтобы эта система допускала нетривиальное решение, надо, чтобы её определитель обращался в нуль; это, очевидно, даёт уравнение степени \( m \) относительно параметра \( \beta \). Пусть \( \beta_1, \beta_2, \ldots, \beta_m \) —
корни этого уравнения; корню $\beta_k$ соответствует тогда система коэффициентов $c_{ik} (1 \leq i \leq m)$, для которых

$$V_k = \sum_{i=1}^{m} c_{ik} U_i \quad (k = 1, 2, \ldots, m)$$

есть собственная функция оператора $B$, принадлежащая к собственному значению $\beta_k$). Вместе с тем $V_k$ есть и собственная функция оператора $A$, принадлежащая к собственному значению $\alpha$. Система функций $V_1, V_2, \ldots, V_m$ может заменить собою первоначально выбранную нами систему $U_1, U_2, \ldots, U_m$ в качестве линейного базиса многообразия собственных функций оператора $A$, принадлежащих к собственному значению $\alpha$ (ортогональность этого базиса следует из того, что $V_k$ — собственные функции оператора $B$, принадлежащие к различным собственным значениям $\beta_k$ этого оператора). Проделав такую замену для всех собственных значений $\alpha$ оператора $A$, мы и получаем в лице функций $V$ такую полную ортогональную систему собственных функций оператора $A$, какая требуется утверждением доказываемой леммы.

Доказательство теоремы. Пусть теперь $U_1, U_2, \ldots$ — существующая в силу доказанной леммы полная ортогональная система функций, все члены которой являются одновременно собственными функциями операторов $A$ и $B$. Разлагающая по этой ортогональной системе решение $U(q, t)$ уравнения Шредингера, мы получаем:

$$U(q, t) = \sum_{k=1}^{\infty} a_k(t) U_k(q),$$

где $a_k(t)$ — комплексные числа, могущие зависеть от времени. Так как $U_k(q)$ — собственные функции оператора $A$, то в силу общих результатов § 3 вероятно того, что величина $\Psi$ примет то или иное из своих возможных значений, выражается однозначно через величины $|a_k(t)|^2$. Но так как, с другой стороны, функции $U_k(q)$ служат вместе с тем и собственными

*) Ради простоты мы не будем здесь рассматривать того случая, когда среди чисел $\beta_k$ встречаются равные между собою; все наши выводы сохраняют силу и в этом случае и требуют лишь небольшой детализации рассуждений (см. И. М. Гельфанд, Лекции по линейной алгебре, Гостехиздат, 1948, стр. 127—129).
Функциями оператора $H$, то, как мы видели в начале настоящего параграфа,
\[ a_k(t) = c_k e^{-iE_k t / \hbar}, \]
где $c_k$ — постоянные комплексные числа.

Отсюда $|a_k(t)|^2 = |c_k|^2$, и следовательно, числа $|a_k(t)|^2$, а вместе с ними и вышеупомянутые вероятности, не меняются с течением времени. Это, очевидно, и означает, что величина $\mathcal{A}$ есть интеграл движения. Тем самым доказательство теоремы завершено.

Так как оператор $H$ тривиальным образом коммутирует с самим собою, то из доказанной теоремы, в частности, следует, что для консервативных систем (т. е. когда полная энергия не зависит явно от времени) функция Гамильтона $H(q_1, \ldots, q_s, p_1, \ldots, p_s)$ всегда является интегралом движения. Таким образом, для консервативной системы закон распределения полной энергии не изменяется с течением времени.

Это предложение мы должны рассматривать как выражение закона сохранения энергии в квантовой механике.
ГЛАВА III

ОБЩИЕ НАЧАЛА КВАНТОВОЙ СТАТИСТИКИ

§ 1. Основные идеи статистических методов в физике

Во всём дальнейшем мы будем условно называть феноменологическими физические теории, возникшие независимо от атомистических представлений о строении вещества. В феноменологических теориях состояние системы описывается значениями небольшого числа характеризующих его величин. Так, состояние данной массы газа, не подверженного действию каких-либо внешних силовых полей, полностью описывается заданием его объёма и температуры, так что всякая другая величина, характеризующая его состояние (например, упругость газа), определяется как некоторая функция этих двух основных характеристик. Но с точки зрения статистической теории заданием объёма и температуры состояние газа ещё не определяется однозначно, так как существует бесчисленное множество различных между собою комбинаций положений и скоростей частиц газа, совместимых с заданными значениями объёма и температуры; с точки зрения статистической теории состояние газа определено однозначно лишь тогда, если заданы положения и скорости всех его частиц. Таким образом, каждое определённое состояние в смысле феноменологической теории объединяет в себе бесчисленное множество состояний, определённых в смысле статистической теории. Это положение вещей не вызывало бы существенных затруднений, если бы всякая физическая величина, интересующая феноменологическую теорию (и, следовательно, определяемая в этой теории как функция объёма и температуры газа), с точки зрения статистической теории имела в точности одно и то же значение во всех тех различных состояниях, которые
совместимы с данными объёмом и температурой. Однако это не так; статистически определяемое давление газа может быть, например, при данных объёме и температуре весьма различным в различных, совместимых с этими данными, комбинациях положений и скоростей частиц газа.

В точности так же обстоит дело и в общем случае. С точки зрения феноменологической теории состояние данной физической системы полностью описывается значениями невольшого числа физических величин $A_1, A_2, \ldots, A_k$. Всякая другая величина $B$, характеризующая собою состояние системы, определяется как некоторая функция величин $A_1, A_2, \ldots, A_k$. В статистической же теории обычно существует не одно, а бесчисленное множество состояний системы, совместимых с данными значениями величин $A_1, A_2, \ldots, A_k$, причём, как правило, величина $B$ будет в этих состояниях принимать различные значения*). Между тем статистическая теория, если она действительно претендует на свою обосновывающую роль по отношению к феноменологическим теориям, должна, конечно, уметь ответить на вопрос о том, какое значение принимает величина $B$ при данных значениях величин $A_1, A_2, \ldots, A_k$ и почему опыт (в согласии с феноменологическими теориями) даёт для величины $B$ при данных $A_1, A_2, \ldots, A_k$ неизмненно одно и то же значение.

Рассмотрим создающееся таким образом положение несколько более детально. Если известны значения величин $A_1, A_2, \ldots, A_k$, то с точки зрения феноменологической теории система находится в некотором вполне определённом состоянии $U$. С точки зрения статистической теории $U$ представляет собою не одно состояние, а целое семейство состояний $U_1, U_2, \ldots$**). Величина $B$, с точки зрения феноменологических

*) В квантовой физике к этой трудности присоединяется ещё новое специфическое затруднение, вызываемое тем, что данная величина в данном статистическом определённом состоянии, вообще говоря, не имеет определённого значения, а имеет лишь некоторый закон распределения. В данный момент мы не будем фиксировать внимания на этой (впрочем, не имеющей для нас существенного значения) специфической трудности

**) Это семейство, вообще говоря, имеет мощность континуума, так что перенумеровать его члены нельзя; только ради простоты записи мы отвлекаемся от этого обстоятельства, которое в данный момент не имеет для нас никакого принципиального значения.
теорий имеющая в состоянии $U$ некоторое вполне определённое значение, вообще говоря, получает различные значения $B_i$ в различных состояниях $U_i (i = 1, 2, \ldots)$. Что же имеет место в действительности?

Допустим, что мы много раз подряд измеряем величину $B$, причём система каждый раз с точки зрения феноменологической теории находится в состоянии $U$ (при этом безразлично, экспериментируем ли мы над одной системой в различные моменты времени или над целым множеством одинаковых систем одновременно). С точки зрения статистической теории мы в этих экспериментах будем иметь дело с различными состояниями системы $(U_1, U_2, \ldots)$, вследствие чего получаемые результаты $(B_1, B_2, \ldots)$ будут различны между собою. На вопрос, «каково значение величины $B$ в состоянии $U$?», статистическая теория может ответить только указанием на среднее арифметическое полученных результатов. Если мы произвели $n$ измерений и при этом $n_1$ раз система находилась в состоянии $U_1$, $n_2$ раз — в состоянии $U_2$ и т. д., то мы ответим, что в состоянии $U$

$$B = \frac{1}{n} (n_1 B_1 + n_2 B_2 + \ldots) = \lambda_1 B_1 + \lambda_2 B_2 + \ldots,$$

где $\lambda_i$ — относительная частота состояний $U_i$ в данной серии экспериментов.

Числа $B_1, B_2, \ldots$ даются, конечно, теорией. Но числа $\lambda_i$ — относительные частоты различных состояний $U_i$, входящих в семейство $U$, — никакой теорией, очевидно, определяться не могут; они зависят от условий экспериментирования и могут существенно меняться при переходе от одной серии экспериментов к другой.

Это же положение вещей может быть описано ещё следующим образом: так как с точки зрения статистической теории $U$ есть семейство состояний $U_i$ с соответствующими значениями $B_i$ величины $B$, то на вопрос о том, каково значение величины $B$ в состоянии $U$, эта теория может ответить только указанием некоторого среднего значения

$$\bar{B} = \sum_i \lambda_i B_i, \quad (\lambda_i > 0, \sum_i \lambda_i = 1).$$

При этом выбор статистических «весов» $\lambda_i$ означает выбор
принципа осреднения. Вообще говоря, различные принципы осреднения приводят, разумеется, к различным средним значениям величины \( B \). Для того чтобы предсказываемое теорией значение величины \( B \) совпадало со средним арифметическим результатов данной серии экспериментов, надо, чтобы статистический вес \( \lambda_i \) совпадал с относительным числом («долей») тех экспериментов, которые заставляют систему в состоянии \( U_i \) (и, следовательно, дают для величины \( B \) значение \( B_i \)).

Из сказанного ясно, что создающееся положение действительно затруднительно. Прежде всего, феноменологические теории определенно приписываются величине \( B \) в состоянии \( U \) (и, значит, во всех состояниях \( U_i \)) некоторое единственноное значение, которое эта величина должна, таким образом, принимать в каждом эксперименте; напротив, статистическая теория утверждает, что величина \( B \) будет в различных экспериментах принимать различные значения, и в качестве эквивалента для определяемого феноменологическими теориями значения величины \( B \) может поэтому в лучшем случае предложить некоторое среднее значение \( \overline{B} \) возможных результатов измерения \( B_i \).

Оказывается, что для расчета этого среднего значения \( \overline{B} \) статистическая теория не может даже предложить никакого определённого принципа осреднения (статистических весов \( \lambda_i \)), так как с её точки зрения этот принцип должен зависеть от условий экспериментирования и меняться от одной серии экспериментов к другой.

Имеется только одна возможность примирить эти угрожающие научной теории противоречия: допустить, что интересующая нас величина \( B \) во всех (или почти всех) состояниях \( U_i \), являющихся разновидностями семейства \( U \), принимает одно и то же (или почти одно и то же) значение. Действительно, если почти все числа \( B_i \) весьма близки между собою, то при любых условиях эксперимента можно ожидать, что почти все получаемые результаты будут приблизительно одинаковы; при этих условиях среднее значение \( \overline{B} \) будет в весьма широких границах независимо от выбираемого нами принципа осреднения и потому может выступать в качестве полноправного аналога того значения, которое дают феноменологические теории величине \( B \) в состоянии \( U \).
Статистическая теория идёт именно этим путём, и другого пути у неё быть не может. Она прежде всего строит метод, позволяющий строго определённым путём находить некоторое среднее значение $\bar{B}$ величины $B$ при данных значениях величин $A_1, A_2, \ldots, A_k$ (в качестве которых обычно фигурируют энергия системы и так называемые «внешние параметры»; среди последних наиболее важным является занимаемый системой объём). Затем она доказывает представительность этого среднего значения $\bar{B}$, состоящую в том, что в подавляющем большинстве состояний, совместимых с заданными значениями величин $A_1, A_2, \ldots, A_k$, величина $B$ принимает значения, весьма близкие к $\bar{B}$. После того как это сделано, все трудности отпадают. Среднее значение $\bar{B}$ есть то значение величины $B$, которое она, согласно феноменологическим теориям, должна принимать всякий раз, когда величины $A_1, A_2, \ldots, A_k$ получают заданную систему значений; статистическая теория предсказывает, что в подавляющем большинстве экспериментов величина $B$ будет получать значения, близкие к $\bar{B}$.

Однако это, кажущееся столь простым, решение задачи всё же требует ещё целого ряда оговорок. Прежде всего надо иметь в виду, что во всём предыдущем под величиной $B$ мы должны, конечно, понимать далеко не всюкую величину, получающую определённое значение в любом из задаваемых статистической теорией состояний системы. Так, например, если $B$ есть скорость некоторой определённой входящей в состав данной системы частицы, то всё сказанное выше к ней совершенно неприменимо; хотя эта величина и имеет определённое среднее значение $\bar{B}$, но о представительности этого среднего, разумеется, не может быть и речи, так как это означало бы, что скорость выбранной нами частицы при всех или почти всех измерениях сохраняет примерно одно и то же значение. Но дело в том, что по отношению к такой величине, как скорость отдельной частицы, вся описанная выше проблематика вовсе и не может возникнуть, так как с точки зрения феноменологических, не статистических, теорий не существует и самого понятия частицы. Непосредственно ясно, что всякая величина $B$, которая может быть определена в терминах феноменологических теорий, в свете статистической теории должна
симметрично зависит от состояний всех составляющих данную систему частиц. Поэтому только для таких величин имеет смысл ставить вопрос о представительности их средних значений, и для них он (по крайней мере в важнейших и наиболее часто встречающихся случаях), как мы увидим, получает благоприятное разрешение.

Во вторую очередь надо иметь в виду, что тот метод, который статистическая теория избирает для построения средних значений физических величин, представляется хоть и весьма естественным, но всё же в полной мере произвольным. Можно, конечно, считать, что доказанная post factum представительность получаемых этим методом средних значений оправдывает сам метод средних (ибо в силу этой представительности, очевидно, любой другой метод оправдывает, как правило, должен приводить к практически тем же средним значениям). Однако вопрос этот всё же требует более внимательного рассмотрения. Дело в том, что когда мы утверждаем, что значение величины \( B \) близко к \( \bar{B} \) в почти всех состояниях данного семейства, то мы всё же допускаем (и мы вынуждены это допустить, ибо это соответствует действительности), что существует небольшое исключительное множество таких состояний, в которых \( B \) значительно отличается от \( \bar{B} \). Что значит здесь слово «небольшое»? Как мы оцениваем размеры этого множества?

Как мы видели выше, выбирая тот или другой способ оценки, мы задаём «веса» \( \lambda_i \) состояний \( U_i \), входящих в данное семейство. Множество таких состояний признаётся «малым», если мала сумма весов всех состояний, входящих в это множество; эту сумму удобно называть весом данного множества; можно, таким образом, сказать, что каждый принцис оценки приписывает определённый вес любому множеству состояний \( U_i \). И вот, когда мы выше говорили, что \( B \approx \bar{B} \) *) в почти всех состояниях \( U_i \), то точный смысл этого утверждения состоял в том, что \( B \approx \bar{B} \) для всех состояний \( U_i \), за исключением некоторого множества \( M \) весьма малого веса. Пусть мы теперь переходим к некоторому новому принципу оценки. Тогда все множества состояний получают

*) \( \approx \) означает «приближенно равно».
новые веса. Если исключительное множество $M$, вес которого при старом способе осреднения был мал, получит весьма малый вес и при новом методе осреднения, то попрежнему подавляющий вес получат состояния, в которых $B \approx \overline{B}$, и даваемое новым методом среднее значение не будет существенно отличаться от $\overline{B}$. Но если при новом способе осреднения вес множества $M$ становится большим, то при вычислении среднего значения величины $B$ заметное влияние получают состояния, в которых $B$ значительно отличается от $\overline{B}$, и новое среднее значение может поэтому оказаться существенно отличным от $\overline{B}$.

Таким образом, из представительности средних значений $\overline{B}$, даваемых избраннным нами методом осреднения, мы можем заключить, что какой-либо новый метод будет давать примерно те же средние значения, если множества состояний, веса которых малы при старом методе осреднения, не получают большей весов при новом методе. В противном случае новый метод может приводить к средним значениям, весьма существенно отличным от старых. Если это положение вещей подвергнуть некоторой математической идеализации, то мы можем сказать, что выбор того или другого принципа осреднения равносилен установлению некоторого мероопределения в «пространстве» состояний $U_i$ и что условием практического совпадения средних значений какой-либо величины при двух различных принципах осреднения служит взаимная абсолютная непрерывность соответствующих двух мероопределений (множество, мера которого равна нулю в первом меропределении, должно иметь меру нуль и во втором меропределении).

Как раз квантовая статистика знает замечательный по своей поучительности и по своим последствиям случай, когда новый принцип осреднения, в силу некоторых принципиальных соображений ставший на место старого, привёл к меропределению, не абсолютно непрерывному относительно прежнего; новые средние значения действительно оказались существенно отличными от старых и во многих случаях, а иногда даже значительно лучше согласие с опытными данными. Это — переход от обычных статистических концепций к так называемым «новым статистикам» — симметрической и антисимметрической. В бли-
жайших параграфах мы будем иметь случай рассмотреть этот пример во всей подробности.

Наконец, при осуществлении намеченного нами пути в области квантовой физики нам придётся в полной мере считаться с тем обстоятельством, что здесь, как мы уже знаем из главы II, даже при наиболее точном определении состояния системы нам бывают, как правило, известны не точные значения физических величин, а лишь их законы распределения (и, в частности, их математические ожидания); это обстоятельство, как мы увидим, не препятствует установлению представительности средних значений для наиболее важных из этих величин, но несколько усложняет ведущие к этой цели расчёты.

§ 2. Микроканонические средние

Мы обращаемся теперь, в согласии с намеченной нами программой, к установлению того принципа осреднения, которым мы будем пользоваться всюду в дальнейшем.

Мы будем иметь дело только с такими состояниями, в которых полная энергия \( E \) нашей системы точно фиксирована («стационарные» состояния); совокупность таких состояний, как мы видели в гл. II, описывается в квантовой физике совокупностью \( \mathcal{M}_E \) собственных функций оператора энергии \( \mathcal{H} \) данной системы, соответствующих собственному значению \( E \) этого оператора. Пусть \( U_1, U_2, \ldots, U_m \) — полная ортогональная нормированная система функций этого семейства; тогда любая нормированная функция \( U \) этого семейства может быть (единственным образом) представлена в виде

\[
U = \sum_{k=1}^{m} a_k U_k, \tag{1}
\]

где \( a_1, a_2, \ldots, a_m \) — комплексные числа, удовлетворяющие соотношению

\[
\sum_{k=1}^{m} |a_k|^2 = 1, \tag{2}
\]

и обратно, всякая функция вида (1), где числа \( a_k \) подчинены условию (2), есть одна из (нормированных) собственных функций...
ций оператора $H$, соответствующих собственному значению $E$ этого оператора. Таким образом, семейство стационарных состояний с данным значением $E$ полной энергии системы приводится во взаимно однозначное соответствие с комплексной $m$-мерной сферой (2). Установить в этом пространстве некоторое мероопределение мы можем поэтому, выбирая какое-либо мероопределение на сфере (2). Наиболее естественным здесь, очевидно, представляется следующий выбор: положим

$$a_k = r_k e^{i\varphi_k} \quad (r_k \geq 0, \quad \sum_{k=1}^{m} r_k^2 = 1, \quad 0 \leq \varphi_k \leq 2\pi, \quad 1 \leq k \leq m)$$

и обозначим через $S$ вещественную сферу

$$\sum_{k=1}^{m} r_k^2 = 1;$$

тогда в пространстве модулей $r_k$ мы устанавливаем естественную евклидову метрику вещественной сферы $S$, а фазы $\varphi_k$ предполагаем равномерно распределёнными в отрезке $(0, 2\pi)$ независимо друг от друга и от модулей. Элемент объёма комплексной сферы (2) представляется, таким образом, в виде

$$dS \ d\varphi_1 \ ... \ d\varphi_m,$$

где $dS$ — элементарная "площадка" вещественной сферы $S$. Другими словами, если $M$ — любое измеримое множество точек $(a_1, a_2, \ldots, a_m)$ комплексной сферы (2), и $\Psi(a_1, \ldots, a_m)$ — характеристическая функция множества $M^*$, то мера множества $M$ принимается пропорциональной интегралу

$$\int_S dS \int_{0}^{2\pi} \int_{0}^{2\pi} \Psi(a_1, a_2, \ldots, a_m) \ d\varphi_1 \ d\varphi_2 \ ... \ d\varphi_m.$$

Установленное таким образом мероопределение мы нормируем так, чтобы мера всей комплексной сферы (2) (а следовательно, и мера всего рассматриваемого нами семейства

*) То-есть $\Psi(a_1, \ldots, a_m)$ равна 1 или 0 в зависимости от того, принадлежит ли точка $(a_1, \ldots, a_m)$ сфере $S$ множеству $M$ или нет.
стационарных состояний) равнялась единице \( * \); это необходимо для того, чтобы мы могли рассматривать определённую нами меру множеств как их статистический вес при дальнейших усреднениях.

Пусть теперь \( f(U) \) — величина, принимающая определённое значение в каждом из стационарных состояний \( U \) нашего семейства; мы можем рассматривать

\[
f(U) = f(a_1U_1 + \ldots + a_mU_m)
\]

как функцию переменных \( a_1, a_2, \ldots, a_m \). В силу принятого нами меропределения мы, естественно, назовём средним значением \( \bar{f}(U) \) величины \( f(U) \) интеграл

\[
\bar{f}(U) = \int_0^{2\pi} dS \int_0^{2\pi} \ldots \int_0^{2\pi} f(\sum_{k=1}^m r_k e^{i\varphi_k} U_k) \, d\varphi_1 \ldots d\varphi_m,
\]

который можно, конечно, записать и более кратко в виде

\[
\bar{f}(U) = \int_{S^*} f(\sum_{k=1}^m \alpha_k U_k) \, dS^*,
\]

где \( S^* \) — комплексная сфера (2), а \( dS^* \) — элементарная «площадка» этой сферы (\( dS^* = \mu \, dS \, d\varphi_1 \ldots d\varphi_m \), где \( \mu \) — вышепомянутый нормирующий множитель). Величину вида \( f(U) \) мы во всём дальнейшем будем называть фазовой функцией нашей системы, а определённое нами среднее значение \( \bar{f}(U) \) этой функции — её микроканоническим средним. Мы установили, таким образом, некоторый единообразный принцип построения средних значений фазовых функций — принцип «микроканонического» усреднения.

Мы знаем, однако, что характеризующие состояние данной системы механические величины не являются, вообще говоря, в квантовой физике фазовыми функциями. Такая величина

\*

(*) Легко видеть, что нормирующим множителем будет служить выражение \( (2\pi)^{-m} \Sigma^{-1} \), где \( \Sigma = \frac{2(V/\pi)^m}{\Gamma \left( \frac{m}{2} \right)} \) — евклидова «площадь» сферы \( S \).
$A$ может в данном состоянии $U$ принимать различные значения, и квантовая теория даёт нам закон распределения этих значений. Как мы знаем из гл. II, величине $A$ ставится в соответствие некоторый (линейный самосопряжённый) оператор $A$, действующий на функцию $U$; математическое ожидание $E_UA$ величины $A$ в состоянии $U$ равно скалярному произведению $(AU, U)$. Это правило, если мы примем его для всех величин $A$ и всех состояний $U$, даёт нам, как мы видели в гл. II, не только математические ожидания, но и законы распределения любых механических величин.

Итак, физическую величину $A$ мы, вообще говоря, не можем считать фазовой функцией, а потому не можем говорить и о микроканоническом среднем такой величины; но математическое ожидание $(AU, U)$ величины $A$ в состоянии $U$ всегда представляет собою фазовую функцию, так что мы можем говорить о микроканоническом среднем $(AU, U)$ этого математического ожидания:

$$\langle AU, U \rangle = E_UA = \int_S (AU, U) \, dS^*.$$

Мы условимся во всём дальнейшем называть это микроканоническое среднее математического ожидания величины $A$ — микроканоническим средним самой величины $A$ (такое расширение понятия микроканонического осреднения не может, разумеется, привести к противоречиям, так как в случае величины, не имеющей в каждом состоянии единственного определённого значения, понятие микроканонического среднего до сих пор вообще никак не было определено). Таким образом, мы полагаем

$$\bar{A} = E_UA = \int_S (AU, U) \, dS^*, \quad (3)$$

где $A$ — оператор, соответствующий величине $A$. Можно сказать, что это соотношение определяет собою микроканоническое среднее для таких фазовых функций, значениями которых служат не числа, а случайные величины.

Так как построение математического ожидания случайной величины всегда представляет собою некоторый процесс осреднения, то в квантовой статистике мы, естественно, на каждом
шагу будем иметь дело с двумя осредняющими процессами, которые не имеют между собой ничего общего: формированием математического ожидания случайной величины и микроканоническим осреднением. Чрезвычайно важно никогда не смешивать между собой этих двух процессов. Во всём дальнейшем черта сверху всегда будет означать микроканоническое осреднение; термин «среднее значение» всегда будет означать микроканоническое среднее и не должен смешиваться с математическим ожиданием случайной величины; математическое ожидание величины $\mathcal{X}$ в состоянии $U$ мы будем там, где это понадобится, обозначать символом $E_U \mathcal{X}$.

Мы приведём теперь выражение (3) для микроканонического среднего величины $\mathcal{X}$ к виду, значительно более удобному для наших дальнейших целей. Так как

$$U = \sum_{k=1}^{m} a_k U_k,$$

то в силу линейности оператора $A$

$$(AU, U) = \left( \sum_{k=1}^{m} a_k AU_k, \sum_{l=1}^{m} a_l U_l \right);$$

отсюда в силу известных свойств скалярных произведений (см. гл. II, § 1)

$$(AU, U) = \sum_{l,k=1}^{m} a_k a_l^* (AU_k, U_l).$$

Поэтому формула (3) даёт:

$$\mathcal{X} = \sum_{l,k=1}^{m} (AU_k, U_l) \int_{S^*} a_k a_l^* dS^*.$$

При $l \neq k$

$$\int_{S^*} a_k a_l^* dS^* = \mu (2\pi)^{m-2} \int_{S} \int_{0}^{2\pi} \int_{0}^{2\pi} e^{i(\varphi_k - \varphi_l)} d\varphi_k d\varphi_l = 0,$$

а при $l = k$

$$\int_{S^*} \left| a_k \right|^2 dS^* = \frac{1}{m} \int_{S^*} \left\{ \sum_{k=1}^{m} \left| a_k \right|^2 \right\} dS^* = \frac{1}{m} \int dS^* = \frac{1}{m}.$$
так как мера комплексной сферы $S^*$ в силу принятой нами нормировки равна единице. Таким образом, мы получаем:

$$\overline{\mathbf{A}} = \frac{1}{m} \sum_{k=1}^{m} (AU_k, U_k).$$

(4)

Это означает, что микроканоническое среднее любой физической величины просто равно среднему арифметическому математических ожиданий этой величины для всех состояний некоторой полной ортогональной нормированной системы. Таким образом, микроканоническое осреднение равносильно осреднению по ортогональному базису, причём всем членам этого базиса должны быть приписаны одинаковые веса.

Мы определили микроканоническое осреднение, опираясь при этом на некоторый определённый линейный ортогональный нормированный базис $U_1, U_2, \ldots, U_m$ многообразия $\mathcal{M}_E$. Очевидно, однако, что для состоятельности этого определения необходимо независимость его от выбираемого базиса. Докажем теперь эту независимость *). Если $V_1, V_2, \ldots, V_m$ — другой линейный базис многообразия $\mathcal{M}_E$, то (гл. II, § 3)

$$V_k = \sum_i \lambda_{ik} U_i \quad (1 \leq k \leq m),$$

где числа $\lambda_{ik}$ образуют унитарную матрицу, и суммирование здесь и во всём дальнейшем ведётся в пределах от 1 до $m$. Отсюда

$$(AV_k, V_k) = (\sum_i \lambda_{ik} AU_i, \sum_j \lambda_{jk} U_j) =$$

$$= \sum_{i,j} \lambda_{ik} \lambda_{jk}^* (AU_i, U_j),$$

и следовательно,

$$\sum_k (AV_k, V_k) = \sum_{i,j} (\sum_k \lambda_{ik} \lambda_{jk}^*) (AU_i, U_j);$$

но в силу унитарности матрицы $\lambda_{ik}$ мы имеем:

$$\sum_k \lambda_{ik} \lambda_{jk}^* = \delta_{ij} \quad (1 \leq i, j \leq m);$$

*) В сущности, эта независимость вытекает уже из инвариантности введённого нами на стр. 112 мероопределения по отношению к выбору ортогонального нормированного базиса; лишь для полноты мы приводим это доказательство.
§ 3. Полная, симметрическая и антисимметрическая статистики

Принцип микроканонического осреднения, определённый нами в § 2, очевидно, основывается на идее равноправного участия в формировании средних значений всех собственных функций оператора $H$, соответствующих данному собственному значению $E$ этого оператора. Эта идея была руководящей и в классической статистической механике: при микроканоническом осреднении и там принимались в качестве равноправных все точки фазового пространства, принадлежащие данной «поверхности постоянной энергии». В классической механике эта идея базировалась на допущении, что уравнения движения не имеют других однозначных интегралов, кроме интеграла энергии; в самом деле, если существует интеграл $I$ уравнений движения, независимый от интеграла энергии, и если в данном случае $I = C$, то в качестве точек фазового пространства, способных изображать собою возможные состояния данной системы, могут фигурировать уже не все точки «поверхности» $H = E$, а лишь часть этих точек, составляющая собою некоторое многообразие меньшего числа измерений, характеризуемое уравнениями $H = E$, $I = C$. В этом случае осреднение, в котором равноправно участвуют все точки «поверхности» $H = E$, будет, как правило, приводить к неверным результатам, так как при таком осреднении подавляющую роль имели бы состояния, вовсе недоступные для данной системы. И то, что методы классической статистической механики там, где они применимы (т. е. где не требуется перехода к представлениям квантовой физики), всегда приводят к результатам, хорошо согласующимся с данными опыта, с несомненностью показывает, что в задачах классической механики такие «мешающие» интегралы действительно отсутствуют.
В квантовой физике вопрос ставится совершенно аналогично. Если мы выбираем такой метод осреднения, в котором принимают равноправное участие все стационарные состояния системы, соответствующие данному значению $E$ её энергии, то мы должны исходить из допущения, что ничто не препятствует нашей системе действительно оказаться в любом из этих состояний, или что все эти состояния, как принято говорить, достижимы для данной системы.

Но как раз для наиболее важных и часто встречающихся систем квантовой физики это допущение оказывается неверным. Такая система, как правило, при каждом данном значении $E$ своей полной энергии может принимать не все состояния, описываемые собственными функциями $U$ оператора $H$, соответствующими его собственному значению $E$, а лишь весьма незначительную часть этих состояний. Правильное осреднение должно поэтому для таких систем существенно отличаться от микроканонического, определённого нами выше; и в целом ряде физически важных случаев это новое осреднение приводит к результатам, значительно отличающимся от микроканонических средних. Мы должны теперь рассмотреть детально, как это происходит.

Пусть наша система состоит из $n$ совершенно одинаковых и полностью равноправных частиц; обозначим через $(q_i, p_i)$ совокупность всех гамильтоновых переменных $i$-й частицы; гамильтонова функция системы, в силу равноправности частиц, должна быть симметрической функцией переменных $(q_1, p_1), (q_2, p_2), \ldots, (q_n, p_n)$; при этом мы допускаем возможность любого взаимодействия между частицами, и даже явную зависимость гамильтоновой функции от времени. Очевидно, такой же симметрией относительно любой пары частиц будет обладать и оператор $H$ данной системы. Пусть функция

$$U = U(q_1, q_2, \ldots, q_n, t)$$

удовлетворяет уравнению Шредингера

$$i\hbar \frac{\partial U}{\partial t} = HU.$$  \(5\)

Покажем, что в таком случае функция

$$\tilde{U} = U(q_2, q_1, \ldots, q_n, t),$$
получаемая из функции $U$ перестановкой первых двух частиц, будет решением уравнения (5).

Обозначим через $P$ оператор, переводящий любую функцию $f(q_1, q_2, \ldots, q_n, t)$ в функцию $f(q_2, q_1, \ldots, q_n, t)$ (оператор перестановки двух первых частиц), и разберёмся подробнее в том, что означает постулированная нами симметрия оператора $H$. С этой целью мы должны будем определить, какой смысл имеет вообще перестановка двух первых частиц в некотором операторе $A$; пусть эта перестановка преобразует $A$ в некоторый другой оператор $\tilde{A}$ (мы не можем, как это казалось бы естественным, обозначить этот новый оператор через $PA$, так как этот символ имеет у нас другой смысл). Пусть теперь мы хотим переставить две первые частицы в выражении $AU$; естественно считать, что для этого мы должны переставить их как в операторе $A$ (результатом будет $\tilde{A}$), так и в функции $U$ (результатом будет $PU$). Поэтому

$$PAU = \tilde{A}PU;$$

полагая $PU = V$, откуда $U = PV$, находим:

$$\tilde{A}V = PAPV$$

или, в терминах одних операторов,

$$\tilde{A} = PAP.$$

Это и даёт нам естественное определение оператора $\tilde{A}$. Симметрия же оператора $A$ выражается требованием $\tilde{A} = A$, или

$$PAP = A,$$

откуда

$$PPAP = AP = PA,$$

t. e. симметричный оператор должен коммутировать с оператором $P$ перестановки двух частиц (а следовательно, и с любой перестановкой частиц между собой).

В частности, симметрия оператора $H$ имеет своим следствием

$$PH = HP,$$
т. е. операторы \( P \) и \( H \) коммутируют друг с другом. Поэтому из (5) следует

\[
\text{i}h \frac{\partial \tilde{U}}{\partial t} = P \left( \text{i}h \frac{\partial U}{\partial t} \right) = PHU = HPU = H\tilde{U},
\]

что и доказывает наше утверждение.

Итак, вместе с \( U \) решением уравнения Шредингера будет и функция \( \tilde{U} = PU \), а следовательно, и функция

\[
\phi_t = \phi(q_1, \ldots, q_n, t) = U - \tilde{U}.
\]

Допустим теперь, что при \( t = 0 \) функция \( U \) симметрична относительно \( q_1 \) и \( q_2 \), т. е.

\[
U(q_1, q_2, \ldots, q_n, 0) = U(q_2, q_1, \ldots, q_n, 0)
\]

точственно относительно \( q_1, q_2, \ldots, q_n \); иначе говоря, мы имеем \( \phi_0 = 0 \) тождественно относительно \( q_1, q_2, \ldots, q_n \). Отсюда и норма функции \( \phi_0 \) равна нулю; а так как для функции, удовлетворяющей уравнению Шредингера, норма не может меняться с течением времени (см. гл. II, § 4), то функция \( \psi_t \) при любом \( t \) имеет норму, равную нулю, вследствие чего \( \phi = 0 \) при любых \( t, q_1, q_2, \ldots, q_n \). Но это означает, что при любых значениях тех же переменных

\[
U(q_1, q_2, \ldots, q_n, t) = U(q_2, q_1, \ldots, q_n, t).
\]

Итак, всякая функция, удовлетворяющая уравнению Шредингера и в какой-либо момент времени симметрична относительно некоторой пары частиц, сохраняет эту симметрию в любой другой (предшествующей или последующей) момент времени. Если поэтому состояние системы в момент \( t = 0 \) описывается собственной функцией, симметричной относительно какой-либо пары частиц, то все состояния, входящие в эту симметрию, вообще недоступны для данной системы ни в её прошлом, ни в её будущем.

Но отсюда следует, что если собственная функция \( U = U(q_1, \ldots, q_n, t) \), описывающая состояние системы, при \( t = 0 \) симметрична относительно любой пары частиц (т. е. просто есть симметрическая функция переменных \( q_1, q_2, \ldots, q_n \)), то она обладает этим свойством и в любой другой (предыдущей или последующей) момент времени \( t \). Для такой си-
стемы достижимы только те состояния, которые описываются собственными функциями, симметричными относительно всех переменных \( q_i \). Очевидно, для такого рода систем реальная значимость средних значений физических величин требует, чтобы осреднение велось по одним только симметрическим собственным функциям, так как все другие собственные функции описывают состояния, принципиально недоступные для данной системы, и учёт влияния таких функций при формировании средних значений мог бы поэтому привести к результатам, не имеющим ничего общего с физической реальностью.

Собственные функции, соответствующие данному собственному значению \( \mathcal{E} \) оператора \( \mathcal{H} \), образуют, как мы знаем, линейное многообразие \( \mathcal{M}_E \), число измерений которого мы будем обозначать через \( m \). Симметрические функции, входящие в состав многообразия \( \mathcal{M}_E \), очевидно, также образуют некоторое линейное многообразие \( \mathcal{S}_E \), число \( S \) измерений которого, вообще говоря, \( < m \). Если нам известно, что наша система может находиться лишь в таких состояниях, которые изображаются симметрическими собственными функциями, то для такой системы осреднение, очевидно, должно вестись по многообразию \( \mathcal{S}_E \), а не по многообразию \( \mathcal{M}_E \). Во всём остальном все соображения, приведшие нас к выбору естественного принципа осреднения по многообразию \( \mathcal{M}_E \), остаются в полной силе и в новом случае. Мы выбираем какой-либо ортогональный нормированный линейный базис \( S_1, S_2, \ldots, S_s \), многообразия \( \mathcal{S}_E \) и по аналогии с формулой (4) § 2 определяем микроканоническое среднее величины \( \langle \mathcal{H} \rangle \), которой соответствует оператор \( \mathcal{A} \), с помощью соотношения

\[
\langle \mathcal{H} \rangle = \frac{1}{s} \sum_{k=1}^{s} (\mathcal{A} S_k, S_k).
\] (6)

Про систему, все возможные стационарные состояния которой описываются симметрическими собственными функциями, мы будем говорить, что она подчиняется симметрической статистике; для такой системы микроканоническое среднее всех физических величин должны вычисляться по формуле (6). Наряду с системами такого рода мы будем рассматривать и системы, статистика которых носит антисимметрический характер. Функция

\[
U = U(q_1, q_2, \ldots, q_n)
\]
называется антисимметрической относительно переменных \( q_1 \) и \( q_2 \), если тождественно

\[
P\hat{U} = \hat{U}(q_2, q_1, \ldots, q_n) = -\hat{U},
\]

т. е. если перемена местами переменных (или групп переменных) \( q_1 \) и \( q_2 \) вызывает изменение знака функции. Если под \( q_i \) понимать совокупность обобщённых координат, определяющих положение \( i \)-й частицы, то легко, подобно предыдущему, убедиться, что антисимметрия относительно данной пары частиц, подобно симметрии, инвариантна по отношению к уравнению Шредингера. Отсюда следует, что если состояние системы описывается в некоторый момент времени собственной функцией, антисимметричной относительно любой пары частиц, то так же обстояло дело и во все предшествующие и будет обстоять во все последующие моменты времени; для такой системы «достижимы» только состояния, описываемые антисимметрическими собственными функциями. В полной аналогии с предыдущим мы должны заключить, что для системы такого рода имеет место антисимметричная статистика: все осреднения должны проводиться по линейному многообразию \( \mathfrak{H}_F \) антисимметрических собственных функций (составляющему, подобно \( \mathfrak{S}_F \), лишь часть многообразия \( \mathfrak{H}_F \)).

Всё сказанное до сих пор не отвечает ещё на вопрос о том, должны ли мы смотреть на ту статистику, которой подчиняется данная система, как на специфическую особенность составляющих её частиц, или как на случайную «судьбу» именно данной, выбранной нами системы. Пусть, например, данная система подчиняется симметрической статистике; будет ли тогда обязательно всякая другая система, состоящая из частиц той же природы, при любых условиях также подчиняться симметрической статистике?

Как опыт, так и целый ряд теоретических соображений, которых мы не можем здесь привести, говорят нам, что дело обстоит именно так. Тип статистики, которой подчиняется та или другая система, не зависит ни от условий, в которых она находится, ни тем более от каких-либо случайных причин, но целиком определяется природой составляющих эту систему частиц. Элементарные материальные частицы (электроны, протоны, нейтроны) подчиняются всегда антисимметри-
ческой статистике. Фотоны, напротив, всегда управляются симметрической статистикой. Что касается материальных частиц более сложного строения, то соответствующий им тип статистики определяется в зависимости от того, сколько элементарных частиц составляет собою одну частьцу данного сложного типа; статистика будет симметрической или антисимметрической, смотря по тому, чётно или нечётно это число. Наконец, опыт и теоретические соображения в полном согласии говорят о том, что для систем, состоящих из свободных «нелокализованных» (т. е. не привязанных к определённому месту) частиц никаких иных типов статистики, кроме симметрической и антисимметрической, не встречается. Если, напротив, частицы локализованы в пространстве, то они теряют тем самым свою полную равнoprивязанность (прикованность двух частиц к разным местам создаёт между ними разность, позволяющую их индивидуализировать и отличать друг от друга); в этом случае может иметь место и та первоначально определённая нами статистика, при которой осреднение производится по всему многообразию $\mathcal{M}_E$, и которую мы поэтому будем называть полной статистикой.

Таким образом, в квантовой статистической физике мы неизменно должны учитывать три различные статистические схемы — полную, симметрическую и антисимметрическую, выбиная всякий раз ту, какая свойственна мы там рассматриваемого типа. Общее же развитие теории должно идти параллельно для всех трёх схем.

Условимся приписывать каждой физической системе, состоящей из однотипных частиц, определённый «индекс симметрии» $\sigma$, полагая $\sigma$ равным 0, 1 или $-1$ в зависимости от того, подчиняются ли частицы этой системы полной, симметрической или антисимметрической статистике. Так как тип этой статистики, как мы знаем, остаётся неизменным во времени, то мы можем смотреть на $\sigma$ как на своеобразный «интеграл» уравнения Шредингера, описывающего эволюцию состояния системы во времени. И подобно тому как в классической механике каждый новый однозначный интеграл уравнений движения сужает то многообразие, по которому должно производиться осреднение, так и здесь наличие «интеграла» $\sigma$ заставляет нас (если $\sigma \neq 0$) в целях осреднения заменить многообразие $\mathcal{M}_E$ более узким многообразием $\mathcal{S}_E$, или $\mathcal{A}_E$. 
§ 4. Построение основного линейного базиса

Пусть мы имеем дело с системой, состоящей из $N$ одно-
типных частиц произвольной структуры и занимающей неко-
торый конечный объём $V$. Пусть

$$u_1(x), u_2(x), \ldots, u_n(x), \ldots (7)$$

— полная ортогональная нормированная система (линейный
базис) собственных функций оператора энергии отдельной
частицы; при этом через $x$ мы ради краткости обозначаем
совокупность координат, определяющих положение данной
частицы. Порядок нумерации мы выбираем так, чтобы собс-
твенные значения (уровни энергии), которым соответствуют
функции $u_i(x)$, образовали неубывающую последовательность:
если собственная функция $u_i(x)$ соответствует собственному
значению $\varepsilon_i$, то

$$\varepsilon_1 \leq \varepsilon_2 \leq \ldots \leq \varepsilon_i \leq \ldots;$$

в последовательности $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_i, \ldots$ каждый уровень
энергии повторяется столько раз, какова его кратность (степе-
нь вырождения), т. е. каково число функций последова-
тельности (7), соответствующих этому уровню.

Рассмотрим теперь функцию

$$U = u_{1\alpha}(q_1) u_{2\beta}(q_2) \ldots u_{N\gamma}(q_N), (8)$$

где $q_i$ означает совокупность координат положения $i$-й части-
цы, а $r_1, r_2, \ldots, r_N$ — произвольные натуральные числа,
так что $U$ есть функция координат положения всех $N$ со-
ставляющих данную систему частиц (или, что то же, — ко-
ординат положения самой данной системы).

Теорема 1. Функция $U$, определяемая соотношением
(8), есть собственная функция оператора $H$ энергии дан-
ной системы, соответствующая собственному значению

$$E = \varepsilon_{r_1} + \varepsilon_{r_2} + \ldots + \varepsilon_{r_N} (9)$$

этого оператора.

Доказательство. Обозначая через $H_i$ оператор энер-
гии $i$-й частицы, мы имеем:

$$H = \sum_{i=1}^{N} H_i,$$
и следовательно,

\[ HU = \sum_{i=1}^{N} H_i U. \]

Но так как оператор \( H_i \) действует только на функцию \( u_{r_i}(q_i) \), а все остальные множители функции \( U \) могут быть вынесены за знак этого оператора, то

\[ H_i U = u_{r_1}(q_1) \ldots u_{r_{i-1}}(q_{i-1}) [H_i u_{r_i}(q_i)] u_{r_{i+1}}(q_{i+1}) \ldots u_{r_N}(q_N). \]

А так как \( u_{r_i}(q_i) \) есть собственная функция оператора \( H_i \), принадлежащая к собственному значению \( \varepsilon_{r_i} \), то

\[ H_i u_{r_i}(q_i) = \varepsilon_{r_i} u_{r_i}(q_i), \]

и следовательно,

\[ H_i U = \varepsilon_{r_i} U, \]

откуда

\[ HU = \sum_{i=1}^{N} H_i U = (\sum_{i=1}^{N} \varepsilon_{r_i}) U = EU, \]

что и требовалось доказать.

Если уровень энергии \( E \) системы задан заранее, то любой выбор натуральных чисел \( r_i (i = 1, 2, \ldots, N) \), удовлетворяющих соотношению (9), даёт нам, таким образом, некоторую собственную функцию \( U \) оператора \( H \), соответствующую этому уровню.

Теорема 2. Совокупность всех функций вида (8), получаемых при любых значениях натуральных чисел \( r_1, r_2, \ldots, r_N \), удовлетворяющих соотношению (9), есть линейный ортогональный нормированный базис многообразия \( W_E \) собственных функций оператора \( H \) энергии данной системы, соответствующих его собственному значению \( E \).

Доказательство. Убедимся сначала, что совокупность всех функций вида (8), получаемых при всевозможных значениях натуральных чисел \( r_1, r_2, \ldots, r_N \), удовлетворяющих или нет соотношению (9), представляет собой полную ортогональную нормированную систему функций. Ортогональность и нормированность этой системы очевидным образом вытекают из соответствующих свойств системы (7), и доказательства требует лишь полнота системы (8). Это доказательство проще
всего проводится с помощью индукции по числу частиц \( N \).
В самом деле, при \( N = 1 \) полнота системы (8) совпадает с предположенной нами полнотой системы (7). Пусть поэтому \( N > 1 \), и наше утверждение верно для систем, состоящих из \( N - 1 \) частиц. Пусть \( \varphi (q_1, q_2, \ldots, q_N) \) — функция, ортогональная всем функциям вида (8), так что при любых значениях натуральных чисел \( r_1, r_2, \ldots, r_N \)

\[
I = \int \varphi^* (q_1, \ldots, q_N) u_{r_1} (q_1) \ldots u_{r_N} (q_N) \, dq_1 \ldots dq_N = 0.
\]

Полагая

\[
\int \varphi^* (q_1, \ldots, q_N) u_{r_1} (q_1) \ldots u_{r_{N-1}} (q_{N-1}) \, dq_1 \ldots dq_{N-1} = \psi_{r_1 \ldots r_{N-1}} (q_N),
\]

мы имеем:

\[
I = \int_{-\infty}^{+\infty} \psi_{r_1 \ldots r_{N-1}} (q_N) u_{r_N} (q_N) \, dq_N = 0.
\]

Так как это равенство имеет место при любом \( r_N \), то в силу полноты системы (7)

\[
\psi_{r_1 \ldots r_{N-1}} (q_N) = 0
\]

при любых \( r_1, r_2, \ldots, r_{N-1} \). В силу (10) это даёт при любых \( r_1, r_2, \ldots, r_{N-1} \) и любом фиксированном \( q_N \)

\[
\int \varphi^* (q_1, \ldots, q_N) u_{r_1} (q_1) \ldots u_{r_{N-1}} (q_{N-1}) \, dq_1 \ldots dq_{N-1} = 0.
\]

В силу же предположения о справедливости нашего утверждения для случая \( N - 1 \) частиц отсюда следует, что при любом \( q_N \) мы имеем тождественно относительно \( q_1, \ldots, q_{N-1} \):

\[
\varphi (q_1, \ldots, q_N) = 0;
\]

другими словами, это равенство выполняется тождественно относительно всех \( N \) переменных, что и доказывает наше утверждение.

Теперь уже легко доказать, что функции вида (8), в которых числа \( r_1, \ldots, r_N \) связаны соотношением (9), образуют линейный базис множества \( \mathcal{M}_E \). В самом деле, пусть \( V = V (q_1, q_2, \ldots, q_N) \) — любая собственная функция опера-
тора \( H \), соответствующая собственному значению \( E \). В силу установленной полноты системы функций вида (8) мы можем написать

\[ V = \sum c_i U_i, \]

где \( U_i \) пробегает всю совокупность функций вида (8); общее правило образования коэффициентов Фурье даёт (гл. II, § 3):

\[ c_i = (V, U_i) \quad (i = 1, 2, \ldots). \]

Функция \( U_i \) в силу теоремы 1 есть, как и \( V \), собственная функция оператора \( H \); если соответствующее \( U_i \) собственное значение \( \varepsilon_{r_1} + \varepsilon_{r_2} + \ldots + \varepsilon_{r_N} \) не равно \( E \), т. е. если не выполняется соотношение (9), то \( V \) и \( U_i \) принадлежат разным собственным значениям этого оператора и потому взаимно ортогональны (гл. II, § 3), так что \( c_i = 0 \). Таким образом, функция \( V \) представляется как линейная комбинация функций вида (8), у которых числа \( r_i \) связаны соотношением (9), что и требовалось доказать.

Таким образом, совокупность функций вида (8), у которых \( \varepsilon_{r_1} + \varepsilon_{r_2} + \ldots + \varepsilon_{r_N} = E \), служит линейным ортогональным нормированным базисом многообразия \( W_E \). Эти функции монированным базисом многообразия \( W_E \). Эти функции мы везде в дальнейшем будем называть основными собственными функциями, а описывающие ими состояния — основными состояниями данной системы. Совокупность основных состояний системы однозначно определяется выбором исходной полной ортогональной системы (7) и заданием числа частиц \( N \) и энергии \( E \) данной физической системы. В частности, число \( m \) основных состояний системы, если известна структура составляющих её частиц, является функцией от \( N \) и \( E \), которую мы будем обозначать через \( \Omega(N, E) \) и называть структурной функцией данной системы; этой функции предстоит, как мы увидим, ведущая роль во всём дальнейшем. Обозначая через \( U_1, U_2, \ldots, U_m \) основные собственные функции, соответствующие уровню энергии \( E \), мы в случае полной статистики в силу § 2 имеем для любой физической величины \( \mathcal{A} \):

\[ \mathcal{A} = \frac{1}{\Omega(N, E)} \sum_{k=1}^{m} (AU_k, U_k). \]
Мы вскоре убедимся, что выбранный нами линейный базис, состоящий из основных собственных функций, представляет особые удобства для дальнейшего развития теории. Теперь же мы обратимся к случаю симметрической статистики.

В выражении (8) основной функции \( U \) индексы переменных \( q_1, q_2, \ldots, q_N \) пробегают ряд натуральных чисел от 1 до \( N \); произведён над ними произвольную перестановку \( P \), т. е. перейдём от функции \( U \) к функции

\[
P_U = u_{r_1}(q_{i_1})u_{r_2}(q_{i_2})\ldots u_{r_N}(q_{i_N}),
\]

где \( i_1, i_2, \ldots, i_N \) — числа ряда 1, 2, \ldots, \( N \), в порядке, определяемом перестановкой \( P \). Число всех таких перестановок равно, очевидно, \( N! \). Положим теперь

\[
S = S(q_1, q_2, \ldots, q_N) = \sum_P P_U,
\]

где суммирование производится по всем перестановкам ряда 1, 2, \ldots, \( N \), дающим различные между собою функции \( PU \). Очевидно, что \( S \) есть симметрическая функция переменных \( q_1, q_2, \ldots, q_N \) и в то же время — собственная функция оператора \( H \), соответствующая собственному значению \( E \).

То построение, которое мы только что провели для одной основной функции \( U \), мы проделаем теперь для каждой из \( \Omega(N, E) \) основных функций выбранного нами основного линейного базиса. Каждый раз мы при этом будем получать некоторую симметрическую функцию \( S \), причём, очевидно, некоторые из этих функций могут совпадать друг с другом. Обозначим (в любом порядке) через \( S_1, S_2, \ldots, S_s \) совокупность получаемых таким путём различных между собою симметрических собственных функций и через \( \Sigma_E \) — совокупность всех симметрических функций многообразия \( \mathcal{M}_E \), т. е. всех симметрических собственных функций оператора \( H \), соответствующих собственному значению \( E \).

Теорема 3. Функции \( S_1, S_2, \ldots, S_s \) образуют ортогональный линейный базис многообразия \( \Sigma_E \).

Доказательство. 1. Пусть \( m \neq n \),

\[
S_m = \sum_P P_U, \quad S_n = \sum_P P'U.
\]

(*) Совокупность этих перестановок определяется, вообще говоря, не однозначно.
Тогда

\[(S_m, S_n) = \sum_{p, p'} (P U_m, P' U_n),\]

где суммирование производится по всем параметрам \( P, P' \) соответствующих перестановок ряда \( 1, 2, \ldots, N \), встречающихся в суммах \((S)\). Если бы какое-либо из скалярных произведений \((P U_m, P' U_n)\) было отлично от нуля, то мы бы должны были иметь \( P U_m = P' U_n \), так как \( P U_m \) и \( P' U_n \) — основные функции, т. е. члены ортогональной системы функций (8). Но из \( P U_m = P' U_n \) следует

\[U_m = P^{-1} P' U_n = QU_n,\]

где \( Q = P^{-1} P' \) — перестановка ряда \( 1, 2, \ldots, N \). Таким образом, функция \( U_m \) получается с помощью некоторой перестановки \( Q \) из функции \( U_n \), что, очевидно, приводит к заведомо неверному равенству \( S_m = S_n \). Отсюда следует, что \((S_m, S_n) = 0\) при \( m \neq n \), т. е. функции \( S_1, S_2, \ldots, S_s \) попарно ортогональны.

2. Пусть \( S \) есть любая функция многообразия \( \mathcal{E}_E \) (а значит, и многообразия \( \mathcal{M}_E \)). В силу теоремы 2 тогда

\[S = \sum_{k=1}^{m} \gamma_k U_k,\]  

где \( \gamma_k \) — некоторые комплексные числа, а \( U_k \) пробегает основной линейный базис многообразия \( \mathcal{M}_E \). Пусть какие-либо две функции \( U_k \) и \( U_l \) связаны соотношением \( U_l = PU_k \), где \( P \) — некоторая перестановка ряда \( 1, 2, \ldots, N \). Тогда

\[\gamma_l = (S, U_l) = (S, PU_k).\]

Но в силу симметрии функции \( S \) мы должны иметь \( PS = S \), и следовательно,

\[\gamma_l = (PS, PU_k).\]

Но если представить скалярные произведения в виде интегралов, то \((PS, PU_k)\), очевидно, отличается от \((S, U_k)\) лишь обозначением переменных интегрирований, так что

\[\gamma_l = (S, U_k) = \gamma_k.\]

Таким образом, в сумме (11) все функции \( U_l \) вида \( PU_k \).
имеют один и тот же коэффициент \( \gamma_k \). Сумма соответствующей группы членов правой части (11) равна поэтому

\[
\gamma_k \sum_P P U_k = \gamma_k S_j,
\]

где \( S_j \) — одна из функций ряда \( S_1, S_2, \ldots, S_s \). Так как, очевидно, правая часть равенства (11) полностью распадается на такие группы, то

\[
S = \sum_{j=1}^s \lambda_j S_j,
\]

где \( \lambda_j \) — комплексные числа. Следовательно, функции \( S_j \) образуют линейный базис многообразия \( \Xi_E \). Этим теорема 3 доказана.

В случае, когда составляющие данную систему частицы подчиняются симметрической статистике, мы будем называть функции \( S_1, S_2, \ldots, S_s \) (после нормировки их путем умножения на соответствующие нормирующие множители) основными собственными функциями, а описываемые ими состояния — основными состояниями системы для данного уровня энергии \( E \). Число \( s \) этих основных состояний мы попрежнему будем обозначать через \( \Omega (N, E) \) и называть структурной функцией данной системы. Микроканоническое среднее физической величины \( \bar{A} \) в этом случае, когда осреднение ведется лишь по симметрическим (нормированным) собственным функциям, выражается формулой

\[
\bar{A} = \frac{1}{\Omega (N, E)} \sum_{j=1}^s (A S_j, S_j).
\]

Обратимся теперь к случаю антисимметрической статистики. Для каждой функции \( U \) вида (8) составим функцию

\[
A = A (q_1, q_2, \ldots, q_N) = \sum_P \pm P U,
\]

где суммирование теперь производится по всем перестановкам ряда \( 1, 2, \ldots, N \), а знак \( + \) или \( - \) берется в зависимости от чётности или нечётности перестановки \( P \). Легко видеть,
что функция $A$ может быть записана в виде определителя

$$A = \begin{vmatrix} u_{r_1}(q_1) & u_{r_1}(q_2) & \ldots & u_{r_1}(q_N) \\ u_{r_2}(q_1) & u_{r_2}(q_2) & \ldots & u_{r_2}(q_N) \\ \ldots & \ldots & \ldots & \ldots \\ u_{r_N}(q_1) & u_{r_N}(q_2) & \ldots & u_{r_N}(q_N) \end{vmatrix}; \quad (12)$$

впрочем, в дальнейшем нам почти не встретится надобности в этом выражении. Мы проделаем указанное построение для каждой из функций $U$ вида (8) и обозначим (после нормировки их) через $A_1, A_2, \ldots, A_a$ совокупность получаемых таким путём различных между собой и отличных от нуля (очевидно, антисимметричных) функций; пусть $\mathfrak{M}_E$ — многообразие всех антисимметрических функций, входящих в многообразие $\mathfrak{M}_E$. Тогда имеет место

**Теорема 4.** Функции $A_1, A_2, \ldots, A_a$ образуют ортогональный линейный базис многообразия $\mathfrak{M}_E$.

Нет надобности проводить здесь доказательства этой теоремы, так как оно в общем сходно с доказательством теоремы 3. В случае, когда данная система состоит из частиц, подчиняющихся антисимметрической статистике, мы будем называть (нормированные) функции $A_1, A_2, \ldots, A_a$ основными собственными функциями, а описываемые ими состояния — основными состояниями системы для данного уровня энергии $E$. Число $a$ этих основных состояний мы снова условились обозначать через $\Omega(N, E)$ и называть структурной функцией данной системы. Для микроканонического среднего величины $\bar{A}$ мы в этом случае, провода осреднение по антисимметрическим собственным функциям, находим выражение

$$\bar{A} = \frac{1}{\Omega(N, E)} \sum_{j=1}^{a} (AA_f, A_f).$$

Основные состояния данной системы, описываемые антисимметрическими собственными функциями $A_f$, обладают одной чрезвычайно важной особенностью: если система находится в таком состоянии, то состояния любых двух её частиц должны быть различны между собою. В самом деле, если в выражении такой функции два из индексов $r_i$ равны между
собою, то определитель (12), имея две одинаковые строки, обращается в нуль, и мы получаем.

$$A(q_1, q_2, \ldots, q_N) = 0;$$

такая функция $A$ не может быть нормирована и не описывает никакого реального состояния системы. Доказанное правило обычно называют принципом Паули.

В случае полной статистики, если задано основное состояние $U$ системы, то известно, в каком состоянии находится каждая из входящих в данную систему частиц (состояние $i$-й частицы определяется индексом $r_i$ в выражении (8) основного состояния $U$). В случае же симметрической или антисимметрической статистики такая постановка вопроса не имела бы никакого смысла: в самом деле, пусть, например,

$$S = \sum_P PU$$

— одна из основных симметрических собственных функций. В различных состояниях $PU$ состояние, скажем, первой частицы описывается различными функциями $u_{r_1}(q_1)$, так что сумма $\sum_P PU$ не дает для этой первой частицы никакого определенного состояния. В случае симметрической или антисимметрической статистики реально стоит лишь вопрос о том, сколько частиц находится в таком-то определённом состоянии, если нам известно состояние всей системы. Этот вопрос мы подробно рассмотрим в следующем параграфе.

§ 5. Числа заполнения. Первичные выражения структурных функций

Пусть нам задано стационарное состояние $U$ системы, подчиняющейся любой из трёх статистик: $U$, как обычно, есть собственная функция оператора энергии $H$ и соответствует определённому собственному значению (уровню энергии) $E$ этого оператора. Мы можем постановить вопрос о том, сколько частиц при этом будет находиться в состоянии, описываемом собственной функцией $u_r(x)$ (соответствующей уровню $E_r$ энергии частицы). Число $a_r = a_r(U)$ таких частиц есть, во-первых, физическая величина, значение которой не опре-
делятся заданием состояния \( U \) однозначно; состоянием \( U \) определяется лишь закон распределения числа \( a_r \) и, в частности, его математическое ожидание \( E_U a_r \). Но в каждом из основных состояний системы значение числа \( a_r \) является однозначно определённым. Так, если мы имеем дело с полной статистикой, то каждая из основных собственных функций согласно формуле (8) § 4 представляется в виде

\[
U = u_{r_1} (q_1) u_{r_2} (q_2) \ldots u_{r_N} (q_N),
\]

прямая показывающим, в каком из возможных для неё состояний находится каждая частица. Число \( a_r \) частиц, находящихся в состоянии \( u_r (x) \), равно поэтому просто числу индексов \( r_i \), равных \( r \), в выражении (8); таким образом, число это, действительно, для каждого основного состояния имеет вполне определённое значение. К тому же результату мы легко приходим и в случае двух других статистик; здесь основными функциями служат линейные комбинации функций вида (8), причём во всех членах такой комбинации число индексов \( r_i \), равных данному числу, одно и то же, и следовательно, число \( a_r \) также определяется однозначно.

но если состояние \( U \) не является основным, то его можно представить в виде

\[
U = \sum_{i=1}^{k} a_i U_i,
\]

где \( a_i \) — комплексные числа (в целях нормировки мы допустим, что \( |a_1|^2 + |a_2|^2 + \ldots + |a_k|^2 = 1 \)), а \( U_i \) — основные состояния; среди этих состояний \( U_i \) будут встречаться основные состояния с различными значениями числа \( a_r \), так что нельзя говорить о каком-либо определённом значении этого числа в состоянии \( U \); можно только утверждать, что если \( a_r^{(i)} \) есть значение числа \( a_r \) в основном состоянии \( U_i \), то \( |a_i|^2 \) есть вероятность того, что в состоянии \( U \) мы будем иметь \( a_r = a_r^{(i)} \); таким образом, в частности,

\[
E_U a_r = \sum_{i=1}^{k} |a_i|^2 a_r^{(i)} \quad (*).
\]

\[\text{(*) Мы опускаем здесь доказательства этих утверждений, так как в дальнейшем не будем пользоваться ими.}\]
числа \( a_r \) (\( r = 1, 2, \ldots \)), показывающие, сколько частиц находится в состоянии, описываемом собственной функцией \( u_r(x) \), в статистической физике обычно называют «числами заполнения»; очевидно, что для системы, состоящей из \( N \) частиц и имеющей энергию \( E \), мы должны иметь, независимо от вида управляющей данной системой статистики,

\[
\sum_{r=1}^{\infty} a_r = N, \quad \sum_{r=1}^{\infty} a_r \varepsilon_r = E. \tag{K}
\]

Числа заполнения играют очень значительную роль при любом изложении физической статистики; то обстоятельство, что для основных состояний эти числа, как мы видели, получают однозначно определённые значения, делает систему основных собственных функций особенно удобным линейным базисом того многообразия, по которому (в зависимости от типа управляющей системою статистики) должно в каждом отдельном случае производиться осреднение.

Как мы видели, в любой из трёх статистик каждому основному состоянию системы соответствует определённый набор чисел заполнения \( a_1, a_2, \ldots, a_r, \ldots \), удовлетворяющий соотношениям (K). Поставим теперь обратную задачу: пусть нам дан набор неотрицательных целых чисел \( a_r \) (\( r = 1, 2, \ldots \)), удовлетворяющих соотношениям (K); существуют ли тогда основные состояния системы, для которых числа \( a_r \) служат числом заполнения, и сколько имеется таких основных состояний?

Эта задача, очень важная для развития статистической теории, решается для трёх рассматриваемых нами статистических схем по-разному.

1. Полная статистика. Пусть \( a_1, a_2, \ldots, a_r, \ldots \) — неотрицательные целые числа, связанные соотношениями (K). Для того чтобы основное состояние (8) имело числа \( a_r \) своими числами заполнения, необходимо и достаточно, чтобы среди множителей функции \( U \) функции \( u_r(x) \) встречалась в точности \( a_r \)

раз. Так как \( \sum_{r=1}^{\infty} a_r = N \), то осуществить это требование, очевидно, всегда возможно. Число способов, которыми оно может быть осуществлено, равно числу перестановок группы
из \( N \) элементов, распадающейся на подгруппы из \( a_1, a_2, \ldots \) тождественных между собой элементов, т. е. равно

\[
\frac{N!}{a_1! \ a_2! \ldots a_r! \ldots}.
\]

Таким образом, в случае полной статистики каждому решению (в неотрицательных целых \( a_r \)) системы уравнений (К) соответствует

\[
\frac{N!}{\prod_{r=1}^{\infty} \ a_r!}
\]

основных состояний, имеющих числа \( a_r \) своими числами заполнения.

2. Симметрическая статистика. В выражении

\[
S = \sum_P PU
\]

основных состояний для случая симметрической статистики слагаемые правой части отличаются друг от друга только порядком индексов \( i_k \) в произведениях

\[
u_r (q_{i_1}) \ u_{r_2} (q_{i_2}) \ldots u_{r_N} (q_{i_N});
\]

индексы \( r_k \) во всех слагаемых одни и те же, а потому и числа заполнения для всех слагаемых одни и те же. Но очевидно и обратное: все произведения вида (13), обладающие некоторым определённым набором чисел заполнения, могут отличаться друг от друга только порядком индексов \( i_k \) и, значит, получаются друг из друга соответствующими перестановками этих индексов; но это означает, что все такие произведения входят в состав одной и той же симметрической основной функции \( S \), которая, таким образом, данным набором чисел заполнения однозначно определяется: в случае симметрической статистики каждому решению (в неотрицательных целых \( a_r \)) системы уравнений (К) соответствует одно основное состояние, имеющее числа \( a_r \) своими числами заполнения.

3. Антисимметрическая статистика. Здесь мы имеем для основных функций выражение

\[
A = \sum_P \pm PU,
\]
где знак + имеет место для чётных, а знак — для нечётных перестановок Р. Здесь также слагаемые правой части составляют собою всю совокупность функций вида (13) с одним и тем же набором чисел заполнения, и потому каждому данному решению системы (K) может соответствовать не более одного основного антисимметрического состояния. Однако не для всякой набор чисел заполнения мы получаем нормируемую собственную функцию A — в некоторых случаях она обращается в тождественный нуль. Рассмотрим какой-либо набор чисел заполнения \( a_r \) (\( r = 1, 2, \ldots \)). Если каждое \( a_r \) равно 0 или 1, то в произведении (13) все индексы \( r_k \) различны между собою, а потому два таких произведения, полученные одно из другого любой перестановкой индексов \( i_k \), не могут совпадать между собою; поэтому в сумме \( \sum_P \pm PU \) никакое приведение подобных членов невозможно, и A будет нормируемой собственной функцией. Но если среди чисел \( a_r \) хотя бы одно больше единицы, то среди индексов \( r_k \) в произведении (13) имеются равные между собою; так как порядок нумерации безразличен, то допустим, что \( r_1 = r_2 \); тогда произведение (13) в точности совпадает с произведением

\[
u_{r_1}(q_{i_2}) u_{r_2}(q_{i_1}) \ldots u_{r_N}(q_{i_N}),
\]

получаемым из него перестановкой индексов \( i_1 \) и \( i_2 \); так как из этих двух произведений каждое получается из другого путём простой транспозиции, то в сумме, составляющей A, они будут иметь разные знаки и взаимно уничтожатся; а так как такой уничтожающий партнёр найдется для любого члена этой суммы, то \( A = 0 \).

Таким образом, в случае антисимметрической статистики каждому решению (в неотрицательных целых \( a_r \)) системы уравнений (K) соответствует: 1) одно основное состояние, имеющее числа \( a_r \) своими числом заполнения, если все \( a_r \leq 1 \), и 2) ни одного такого состояния, если хотя бы одно \( a_r > 1 \). Нам будет удобно иметь для числа основных состояний, соответствующих данному набору чисел \( a_r \), такое выражение, которое годилось бы для всех трёх статистических схем. Положим с этой целью для \( n = 0, 1, 2, \ldots \)

\[
C(n) = \begin{cases} 
  n! & \text{в случае полной статистики,} \\
  1 & \text{в случае симметрической и антисимметрической статистики}
\end{cases}
\]
числа заполнения

$$\gamma (n) = \begin{cases} \frac{1}{n!} & \text{в случае полной статистики}, \\ 1 & \text{в случае симметрической статистики}, \\ 1 & (n \leq 1), \\ 0 & (n > 1) \end{cases} \text{в случае антисимметрической статистики.}$$

Тогда число основных состояний, соответствующих данному набору чисел $a_r \geq 0$, удовлетворяющих системе уравнений $(K)$, для любой из трёх статистик равно

$$C (N) \prod_{r=1}^{\infty} \gamma (a_r).$$

Отсюда для числа всех основных состояний системы, соответствующих данному уровню энергии $E$, мы получаем выражение

$$\Omega (N, E) = C (N) \sum_{(K)} \prod_{r=1}^{\infty} \gamma (a_r),$$

где суммирование в правой части производится по всем системам целых чисел $a_1, a_2, \ldots, a_r, \ldots \ (a_r \geq 0)$, удовлетворяющим уравнениям $(K)$.

Как мы уже отметили, структурные функции $\Omega (N, E)$ предназначены играть во всём дальнейшем очень значительную роль, вследствие чего мы должны будем тщательно их исследовать; исходным пунктом всех этих исследований будет служить только что полученная нами формула (14), связывающая функцию $\Omega (N, E)$ с решениями системы уравнений $(K)$.

Числа заполнения $a_r$, при любой статистике однозначно определённые для любого из основных состояний $U_k$, могут рассматриваться как функции этого состояния, и потому для любого уровня энергии $E$ имеют определённые (микроканонические) средние значения $\bar{a}_r$. Как мы увидим дальше, эти средние значения чисел заполнения имеют фундаментальное значение для всех расчётных формул физической статистики, и нахождение их асимптотических выражений составляет поэтому одну из основных математических задач развиваемой нами теории. Этой задаче будут посвящены в значительной степени две последующие главы; здесь же мы останавливаемся ещё только на некоторых общих соображениях, до известной
степени объясняющих эту ведущую роль чисел \( \bar{a}_r \) в асимптомических расчетах физической статистики.

Значительная часть важнейших физических величин \( \mathcal{A} \), характеризующих состояние системы, состоящей из \( N \) однотипных частиц, имеет некоторую специальную форму: такая величина \( \mathcal{A} \) является суммой \( N \) величин \( a_1, a_2, \ldots, a_N \), каждая из которых зависит (и притом одинаковым образом) только от состояния частицы с соответствующим номером; так, в простейшем случае, когда состояние системы описывается функцией

\[
U = u_{r_1}(q_1) u_{r_2}(q_2) \ldots u_{r_N}(q_N),
\]

величина \( a_k \) зависит только от числа \( r_k \), т.е. номера той собственной функции \( u_{r_k}(q_k) \), которой описывается состояние \( k \)-й частицы. Разложению

\[
\mathcal{A} = a_1 + a_2 + \ldots + a_N
\]

величина \( \mathcal{A} \) соответствует, конечно, разложение

\[
A = A_1 + A_2 + \ldots + A_N
\]

соответствующего ей оператора; в этом последнем разложении оператор \( A_k \) \((k = 1, 2, \ldots, N)\) действует только на собственные функции \( k \)-й частицы, так что

\[
A_k U = u_{r_1}(q_1) \ldots u_{r_{k-1}}(q_{k-1}) [A_k u_{r_k}(q_k)] u_{r_{k+1}}(q_{k+1}) \ldots u_{r_N}(q_N).
\]

Величины \( \mathcal{A} \) только что описанного типа мы будем называть сумматорными величинами. Мы рассмотрим теперь, как может быть построено среднее значение такой сумматорной величины.

Если, как мы предположили, зависимость величины \( a_k \) от состояния \( k \)-й частицы одна и та же для всех \( k \), то, в частности, математическое ожидание величины \( a_k \) в каком-либо состоянии \( u_r(q_k) \) \( k \)-й частицы не зависит от \( k \), а зависит только от \( r \). Обозначим это математическое ожидание через \( \lambda_r \); числа \( \lambda_r \) зависят, таким образом, только от структуры частиц и выбора элементарной величины \( a_i \); эти числа никак не связаны с системой и сохраняют свой смысл и свою величину.
и в том случае, когда мы вместо системы рассматриваем отдельную частицу.

Пусть теперь наша система находится в основном состоянии $U$ с числами заполнения $a_1, a_2, \ldots, a_r, \ldots$. Тогда, прежде всего,

$$E_U = \sum_{k=1}^{N} E_U a_k,$$

но очевидно, что $E_U a_k = \lambda_k$, если $k$-я частица находится в состоянии $u_r (q_k)$; поэтому в сумме (15) слагаемое $\lambda_r$ встречается столько раз, сколько частиц находится в состоянии $u_r (q_k)$, т. е. $a_r$ раз; отсюда

$$E_U = \sum_{r=1}^{\infty} a_r \lambda_r,$$

и следовательно,

$$\bar{\lambda} = E_U = \sum_{r=1}^{\infty} a_r \lambda_r. \quad (16)$$

Это простое соотношение показывает, что знание средних значений чисел заполнения позволяет нам немедленно написать среднее значение любой сумматорной величины $\bar{\lambda}$ (необходимо вспомнить, что числа $\lambda_r$ не зависят от состояния системы и при данной структуре частиц и данном выборе величины $\bar{\lambda}$ вычисляются раз навсегда). Эта роль чисел заполнения и является главной причиной, заставляющей всюкую статистическую теорию в первую очередь искать удобное асимптотическое выражение для их средних значений.

§ 6. О представительности микроканонических средних

В § 1 мы подробно говорили о том, что физические величины, характеризующие состояние системы с точки зрения феноменологических теорий, должны в статистической теории симметрично зависеть от всех составляющих данную систему частиц. Средние значения таких величин должны рассматриваться как те их значения, которые даются статистической теорией в качестве реально существующих при данных условиях. Для этого, как мы это подробно установили в § 1, даваемой нашей теорией средние значения должны обладать
Представительность; это означает, что в подавляющем большинстве тех состояний, по которым производится осреднение, величина должна принимать значения, весьма близкие к ее среднему значению.

Средним значением величины $\mathcal{A}$ для данного значения $E$ ее полной энергии мы условились называть величину

$$\bar{\mathcal{A}} = \frac{1}{\Omega(N,E)} \sum_{k=1}^{m} E_{U_k} \mathcal{A},$$

где $m = \Omega(N,E)$ есть число основных состояний $U_k$ системы для данного уровня энергии $E$, и суммирование распространяется на все такие состояния. При этом система может подчиняться статистике любого из трех рассматриваемых нами типов.

Простейшим и вместе с тем важнейшим типом величин $\mathcal{A}$, симметрично зависящих от всех составляющих систему частиц, являются, очевидно, сумматорные величины, рассмотренные нами в конце предыдущего параграфа, т. е. суммы $N$ величин, каждая из которых зависит, и притом одинаковым для всех частиц образом, от состояния какой-либо одной частицы. В настоящем параграфе мы всегда будем считать

$$\mathcal{A} = \sum_{i=1}^{N} a_i$$

такой сумматорной величиной. Так как зависимость величины $a_i$ от состояния $i$-й частицы одинакова для всех $i$, то среднее значение $\bar{a_i} = a$ не зависит от $i$; очевидно, мы имеем:

$$\bar{\mathcal{A}} = Na,$$

tак что, если только $a \neq 0$, мы можем считать $\bar{\mathcal{A}}$ весьма большим числом (порядка $N$); значение величины $\mathcal{A}$ мы можем поэтому считать близким к $\bar{\mathcal{A}}$, если

$$|\mathcal{A} - Na| < \varepsilon N,$$

где $\varepsilon$ — малое положительное число, ибо при выполнении неравенства (18) относительная погрешность равенства $\mathcal{A} \approx \bar{\mathcal{A}}$ меньше, чем $\left|\frac{\varepsilon}{a}\right|$. 
Теперь вспомним, что если данная система находится в основном состоянии $U$, то это не даёт ещё ответа на вопрос, выполняется или нет неравенство (18), так как в состоянии $U$ величина $\mathcal{A}$, вообще говоря, не является однозначно определённой; таким образом, если система находится в состоянии $U$, то имеется лишь определённая вероятность

$$P_U \{ |\mathcal{A} - N\alpha| < \varepsilon N \}$$

того, что неравенство (18) окажется удовлетворённым.

Пусть теперь $M$ означает меру (согласно установленному нами в § 2 мероопределению) множества тех собственных функций $U$ многообразия $\mathcal{M}_E$, для которых

$$P_U \{ |\mathcal{A} - N\alpha| < \varepsilon N \} < 1 - \delta,$$

или, что то же,

$$P_U \{ |\mathcal{A} - N\alpha| \geq \varepsilon N \} > \delta,$$  \hspace{0.6cm} (19)

где $\delta$ — любое (малое) положительное число. Тогда, очевидно, мы имеем, полагая, как в § 3,

$$\int_{S^*} P_U \{ |\mathcal{A} - N\alpha| \geq \varepsilon N \} dS^* \geq \delta M_b,$$

где $S^*$ — комплексная сфера (2) § 2, откуда

$$M_b \leq \frac{1}{\delta} \int_{S^*} P_U \{ |\mathcal{A} - N\alpha| \geq \varepsilon N \} dS^*.$$  

А так как в силу неравенства Чебышева

$$P_U \{ |\mathcal{A} - N\alpha| \geq \varepsilon N \} \leq \frac{E_U \{ (\mathcal{A} - aN)^2 \}}{\varepsilon^2 N^2},$$

то

$$M_b \leq \frac{1}{\delta \varepsilon^2 N^2} \int_{S^*} E_U \{ (\mathcal{A} - aN)^2 \} dS^* \leq \frac{(\mathcal{A} - aN)^2}{\delta \varepsilon^2 N^2}. \hspace{0.6cm} (20)$$

$M_b$ означает меру множества тех стационарных состояний $U$ многообразия $\mathcal{M}_E$, для которых выполняется неравенство (19); но для всех других состояний мы с вероятностью, превышающей $1 - \delta$, можем ожидать выполнения неравенства (18).
т. е. приближённого равенства $\mathcal{A} \approx \bar{\mathcal{A}}$. Таким образом, если $M_\delta$ при малых $\varepsilon$ и $\delta$ будет малой, то в значительном большинстве состояний $U$ можно будет с подавляющей вероятностью ожидать значения величины $\mathcal{A}$, близкого к $\bar{\mathcal{A}}$. Это и есть то, что мы называем представительностью среднего значения $\bar{\mathcal{A}}$. Как показывает неравенство (20), эта представительность будет, следовательно, гарантирована, если при малых $\varepsilon$ и $\delta$ отношение

$$\frac{(\bar{\mathcal{A}} - a N)^2}{\delta s^2 N^2}$$

всё ещё окажется достаточно малым.

Обозначим среднее значение $(\bar{\mathcal{A}} - a N)^2$ через $D(\mathcal{A})$ и будем называть его микроканонической дисперсией величины $\mathcal{A}$. До сих пор мы в наших расчётах нигде не предполагали, что $\mathcal{A}$ есть сумматорная величина; теперь мы установим специальную форму микроканонической дисперсии, имеющую место для сумматорных величин. Очевидно, мы имеем:

$$D(\mathcal{A}) = (\bar{\mathcal{A}} - a N)^2 = \bar{\mathcal{A}}^2 - a^2 N^2 = (\sum_{i=1}^{N} a_i)^2 - a^2 N^2. \quad (21)$$

Но для любого состояния $U$ системы

$$E_U(\mathcal{A}^2) = \sum_{i=1}^{N} E_U(a_i^2) + 2 \sum_{i,j=1}^{N} E_U(a_i \ a_j). \quad (22)$$

Пусть основному состоянию $U_k$ соответствует набор чисел заполнения $a_r (U_k) = a_r \ (r = 1, 2, \ldots)$; тогда в состоянии $U_k$ система содержит $\frac{a_r (a_r - 1)}{2}$ пар частиц в состоянии $u_r (x)$ ($r = 1, 2, \ldots$) и при $r \neq s a_r a_s$ пар частиц, одна из которых находится в состоянии $u_r (x)$, а другая — в состоянии $u_s (x)$. Пусть теперь $\lambda_r$ и $\mu_r$ соответственно означают математические ожидания величин $a_r^2$ и $\bar{a}_r^2$, тогда $i$-я частица находится в состоянии $u_r (x)$ ($r = 1, 2, \ldots$); эти величины не за-
висят от \( i \) и определяются структурой частиц и выбором величин \( \lambda_i \). Формула (22), очевидно, даёт

\[
E_{U_k}(\mathcal{A}^2) = \sum_{r=1}^{\infty} a_r \mu_r + 2 \sum_{r=1}^{\infty} \frac{a_r(a_r-1)}{2} \lambda_r^2 + 2 \sum_{r<s} a_s a_r \lambda_r \lambda_s,
\]

где, разумеется, \( a_r = a_r(U_k) \), \( a_s = a_s(U_k) \). Для микроканонического среднего величины \( \mathcal{A}^2 \) мы поэтому находим в силу (17):

\[
\mathcal{A}^2 = \frac{1}{m} \sum_{k=1}^{m} E_{U_k}(\mathcal{A}^2) = \\
= \sum_{r=1}^{\infty} \overline{a_r \mu_r} + \sum_{r=1}^{\infty} \overline{a_r^2 \lambda_r^2} - \sum_{r=1}^{\infty} \overline{a_r \lambda_r^2} + 2 \sum_{r<s} \overline{a_r a_s \lambda_r \lambda_s} = \\
= \sum_{r=1}^{\infty} (\mu_r - \lambda_r^2) \overline{a_r} + \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} \lambda_r \lambda_s \overline{a_r a_s},
\]

(во всём этом мы предполагаем, разумеется, что все получаемые ряды абсолютно сходятся; для этого, очевидно, достаточно допустить, что числа \( \lambda_r \) и \( \mu_r \) не возрастают слишком быстро с ростом \( r \)). Так как, с другой стороны, в силу формулы (16) § 5

\[
\mathcal{A} = N \lambda = \sum_{r=1}^{\infty} \lambda_r \overline{a_r},
\]

и значит,

\[
N^2 \sigma^2 = \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} \lambda_r \lambda_s \overline{a_r a_s}.
\]

то формулы (23) и (21) дают:

\[
D(\mathcal{A}) = \sum_{r=1}^{\infty} (\mu_r - \lambda_r^2) \overline{a_r} + \sum_{r<s} \sum_{s=1}^{\infty} \lambda_r \lambda_s (\overline{a_r a_s} - \overline{a_r} \overline{a_s}).
\]

Это и есть нужное нам выражение микроканонической дисперсии для сумматорных величин. Для оценки величины \( D(\mathcal{A}) \), как показывает формула (24), нам необходимо, кроме средних значений \( \overline{a_r} \) самих чисел заполнения знать ещё средние значения \( \overline{a_r a_s} \) их попарных произведений. Мы видели,
что представительность среднего значения \( \bar{X} \) обеспечивается достаточной малостью отношения

\[
\frac{D(\bar{X})}{\delta \epsilon^2 N^2},
\]

что при малых \( \epsilon \) и \( \delta \) требует, чтобы отношение \( \frac{D(\bar{X})}{N^2} \)
было ничтожно малым. В гл. IV и V мы увидим, что в физически актуальных случаях величина \( D(\bar{X}) \) есть при \( N \to \infty \) бесконечно большая порядка \( N \), вследствие чего отношение

\[
\frac{D(\bar{X})}{N^2}
\]
имеет порядок малости \( \frac{1}{N} \) и, значит, действительно становится ничтожно малым.

Остановимся ещё несколько подробнее на вопросе о том, каким образом представительность средних значений связана с их экспериментальной проверкой. Пусть известно, что полная энергия данной системы равна \( E \), но никакой другой информации относительно её состояния мы не имеем. Мы производим измерение величины \( \bar{X} \) и хотим знать вероятность

\[
P\left\{ |\bar{X} - aN| < \epsilon N \right\}
\]

того, что полученное нами значение этой величины будет заключено между \( (a - \epsilon)N \) и \( (a + \epsilon)N \). Но в такой постановке наш вопрос не имеет никакого смысла, так как теория даёт закон распределения величины \( \bar{X} \) лишь при условии, что известно состояние \( U \), в котором находится система; мы же знаем лишь её уровень энергии \( E \), которому может соответствовать бесчисленное множество различных состояний \( U \), и у нас нет никаких указаний на то, в каком из них находится наша система. Для того чтобы поставленная нами задача получила определённый смысл, мы должны знать, с какой относительной частотой при выбранных нами условиях эксперимента встречается то или другое состояние \( U \) многообразия \( M_E \), т. е. знать закон распределения состояний \( U \), свойственный данному способу экспериментирования. Допустим для простоты, что этот закон имеет определённую плотность вероятности \( \rho(U) \). Непосредственно ясно, что это распределение \( \rho(U) \) может, вообще говоря, существенно зависеть от условий эксперимента, например от того, измеряем ли мы величину \( \bar{X} \) в одной системе через определённые промежутки
времени *), или в большом числе систем одновременно. Как бы то ни было, если функция \(\rho(U)\) нам известна, то по формуле полной вероятности

\[
P\left\{ |\mathcal{U} - aN| > \epsilon N \right\} = \int_{\mathcal{M}_E} \rho(U) P_U \left\{ |\mathcal{U} - aN| > \epsilon N \right\} d\omega,
\]

где \(d\omega\) — элемент объёма многообразия \(\mathcal{M}_E\) в смысле введённого нами в § 2 на этом многообразии меропределения. Пусть \(M_\delta\) означает, как прежде, меру множества тех \(U\), для которых

\[
P_U = P_U \left\{ |\mathcal{U} - aN| > \epsilon N \right\} > \delta;
\]

тогда, предполагая, что функция \(\rho(U)\) ограничена на \(\mathcal{M}_E\), и обозначая через \(R\) её верхнюю границу, мы из (25) заключаем, что

\[
P\left\{ |\mathcal{U} - aN| > \epsilon N \right\} = \int_{P_U \leq \delta} \rho(U) P_U d\omega + \int_{P_U > \delta} \rho(U) P_U d\omega \leq \delta \int_{\mathcal{M}_E} \rho(U) d\omega + R \int_{P_U > \delta} d\omega = \delta + RM_\delta,
\]

откуда в силу (20)

\[
P\left\{ |\mathcal{U} - aN| > \epsilon N \right\} \leq \delta + \frac{RD(\mathcal{U})}{\epsilon^2 N^2}.
\]

Если, как это имеет место в подавляющем большинстве физических задач, отношение

\[
\frac{D(\mathcal{U})}{N^2}
\]

стремится к нулю при \(N \rightarrow \infty\), то при достаточно малом \(\delta\), сколь угодно малом \(\epsilon\) и достаточно большом \(N\) правая часть неравенства (26) становится сколь угодно малой. Это означает, что при достаточно большом числе частиц приближенное равенство \(\mathcal{U} \approx Na\) с вероятностью, сколь угодно близкой к единице, будет подтверждено экспериментом с относительной погрешностью \(< \frac{\epsilon}{|a|}\).

*) Эти промежутки должны быть достаточно велики для того, чтобы предшествующее измерение не влияло на результат последующего.
Этот вывод, весьма удовлетворительно отвечающий на вопрос о реальной значимости принятого нами метода осреднения, возможен, однако, только при условии, что закон распределения стационарных состояний $U$ имеет ограниченную плотность $\rho(U)$ (в остальном этот закон может быть каким угодно, в этом и состоит главная ценность полученного результата); несколько усложняя проведённый нами расчёт, можно было бы доказать, что к тому же результату мы придём и при более широком допущении, что этот закон абсолютно непрерывен относительно введённого нами на множество $M_E$ меропределения. Однако, если это не так, то все наши расчёты теряют основание, и принятый нами метод осреднения может привести к результатам, резко расходящимся с данными опыта. Замечательный пример такого положения вещей мы уже имели случай исследовать в § 3 при переходе от полной статистики к симметрической (или антисимметрической); множество $S_E$ симметрических функций многообразия $M_E$ имеет в принятом нами меропределении, вообще говоря, меру, равную нулю (как линейное многообразие меньшего числа измерений); между тем, если мы имеем дело с системой, подчиняющейся симметрической статистике, то реально возможными являются лишь состояния $U$, входящие в $S_E$, и следовательно,

$$\int_{S_E} \rho(U) d\omega = 1.$$ 

Это показывает, что закон $\rho(U)$ не является абсолютно непрерывным; и мы, действительно, в целях согласия наших выводов с данными опыта должны были отказаться от прежнего метода осреднения и заменить его новым; это и было переходом от полной статистики к симметрической (или антисимметрической).
ГЛАВА IV
ОСНОВЫ СТАТИСТИКИ ФОТОНОВ

§ 1. Отличительные особенности статистики фотонов

В этой главе мы будем рассматривать однородные системы, состоящие из фотонов (световых квантов). Такие системы обладают некоторыми особыми свойствами, существенно отличающими их от систем, состоящих из материальных частей. Поэтому, хотя все излагаемые нами в дальнейшем методы исследования в равной степени применимы к системам обоих типов, статистика фотонов всё же несколько отличается от статистики материальных частиц и требует особого рассмотрения. Мы начинаем со статистики фотонов потому, что она в математическом отношении значительно проще статистики материальных частиц. Поэтому в применении к ней руководящие идеи нашего метода, не будучи обременены слишком тяжёлым формальным аппаратом, должны выступать рельефнее и легче усваиваться. После же прочного усвоения этих идей на сравнительно простом математическом материале читатель без труда сможет проследить их и на более сложных в формальном отношении задачах статистики материальных частиц.

Основная особенность, отличающая и упрощающая статистику фотонов сравнительно со статистикой материальных частиц, состоит в том, что число $N$ фотонов, составляющих собою данную систему, при данном фиксированном значении $E$ её полной энергии не может рассматриваться как фиксированное, а напротив, меняется от состояния к состоянию. Таким образом, совокупность всех возможных состояний системы при данном значении $E$ её полной энергии описывается совокупностью собственных функций (принадлежащих
к собственному значению $E$) не одного какого-либо оператора $\mathcal{H}$, а операторов энергии всевозможных систем, состоящих из любого числа $N$ фотонов.

С другой стороны, фотонь везде подчиняются симметрической статистике. Поэтому, если $e_1, e_2, \ldots, e_r, \ldots$ — последовательность возможных уровней энергии фотонов в данных условиях (как обычно, $0 \leq e_1 \leq e_2 \leq \ldots \leq e_r \leq \ldots$, и каждый уровень повторяется столько раз, какова его кратность), то число $\Omega(N, E)$ линейно независимых (симметрических) собственных функций оператора $\mathcal{H}$ для данного уровня энергии $E$ и данного числа фотонов $N$ равно (гл. III, § 3) числу решений в целых $a_r \geq 0$ системы уравнений

$$\sum_{r=1}^{\infty} a_r = N, \quad \sum_{r=1}^{\infty} a_r e_r = E. \tag{1}$$

Если мы для каждого $N$ построим такой линейный базис из $\Omega(N, E)$ собственных функций и затем соединим все эти базисы между собою, то мы получим систему из

$$\Omega(E) = \sum_{N=1}^{\infty} \Omega(N, E)$$

(симметрических) собственных функций (различных операторов), принадлежащих к одному и тому же собственному значению $E$; каждая из этих функций описывает одно из возможных состояний нашей системы при данном значении $E$ её полной энергии; обратно, каждое такое состояние описывается линейной комбинацией функций полученной нами системы. Очевидно, что общее число $\Omega(E)$ функций полученной системы просто равно числу решений уравнения

$$\sum_{r=1}^{\infty} a_r e_r = E$$

в целых $a_r \geq 0$, так как первое из уравнений (1) в силу произвольности числа $N$ отпадает. По аналогии с § 4 гл. III мы будем называть $\Omega(E)$ структурной функцией нашей системы, состоящей из фотонов.

В случае, когда число $N$ частиц было фиксировано, мы (§ 2 гл. III) условились называть микроканоническим средним
фазовой функции \( f(U) \) среднее арифметическое её значений для \( \Omega(N, E) \) функций выбранного нами линейного базиса. Иначе говоря, мы при осреднении условились снабжать эти \( \Omega(N, E) \) функций одинаковыми весами.

Теперь, когда число \( N \) фотонов не имеет фиксированного значения, мы, естественно, принимаем в качестве базиса для осреднения только что построенную нами систему из \( \Omega(E) \) функций. Микроканоническим средним фазовой функции \( f(U) \) (при данном значении \( E \)) мы будем называть среднее арифметическое её значений для этих \( \Omega(E) \) функций. Этот принцип осреднения является, конечно, новым произвольным согласием: функции, относящиеся к различным \( N \) (и потому зависящие от различного числа переменных) в наших прежних рассуждениях никогда не сравнивались между собой по своим статистическим весам; теперь мы приписываем всем им, независимо от значения \( N \), один и тот же вес. Правомерность этого нового произвольного согласования, как и нашего прежнего, мы в дальнейшем будем стремиться обосновать, доказав, что получаемые выбранным нами методом осреднения результаты в актуальных для физической статистики случаях оказываются независимыми от этого метода, так что любой (в весьма широких границах) другой принцип осреднения привёл бы нас к тем же выводам.

\[ \text{§ 2. Числа заполнения и их средние значения} \]

Симметрические собственные функции, входящие в состав выбранного нами линейного базиса, могут быть, как мы видели в § 4 гл. III, выбраны из числа так называемых «основных» функций системы. В состоянии системы, описываемом такой функцией, число \( a_r \) частиц, находящихся в состоянии с уровнем энергии \( \varepsilon_r \), является вполне определённым; другими словами, в каждом из \( \Omega(E) \) основных состояний, составляющих выбранный нами базис, мы имеем совершенно определённый набор «чисел заполнения» \( a_r \) (\( r = 1, 2, \ldots \)), всегда удовлетворяющий, конечно, требованию

\[ \sum_{r=1}^{\infty} a_r \varepsilon_r = E. \] (2)
Микроканонические средние чисел заполнения $a_r$ выражаются поэтому формулой

$$\bar{a}_r = \frac{1}{\Omega(E)} \sum_U a_r(U),$$

где суммирование производится по всем основным функциям $U$ нашего базиса, и $a_r(U)$ означает величину числа $a_r$ в состоянии, описываемом функцией $U$.

Основная задача статистической термодинамики любых систем всегда состоит в изучении распределения энергии системы между её составными частями. В нашей теории, где средним числом частиц с уровнем энергии $\varepsilon_r$ мы условились считать микроканоническое среднее $\bar{a}_r$ этого числа, оценка этих величин $\bar{a}_r$ является, таким образом, первоочередной задачей. Вместе с тем, как мы видели в § 5 гл. III, знание величины $\bar{a}_r$ позволяет нам непосредственно написать микроканоническое среднее любой суммарной величины; но суммарными величинами как раз преимущественно и характеризуется состояние системы в статистической физике. Далее, как мы видели в § 6 гл. III, для оценки представительности микроканонических средних (т. е. для обоснования самого метода микроканонического осреднения) в случае суммарных величин необходимо знать ещё величины $\bar{a}_r a_s$ ($r, s = 1, 2, \ldots$), т. е. микроканонические средние попарных произведений чисел заполнения. Поэтому любой математический аппарат, используемый в физической статистике, должен ставить себе первой целью отыскание удобных аналитических выражений для величин $\bar{a}_r$ и $\bar{a}_r a_s$; это в одинаковой мере относится к частям любого типа и к любой из трёх основных статистических схем. В ближайших параграфах мы увидим на простейшем примере фотонов, как это может быть сделано.

Нашим первым шагом будет элементарное выражение чисел $a_r$ и $a_r a_s$ через значения структурной функции $\Omega(E)$, выражающей собою число решений уравнения (2) в целых $a_r \geq 0$. С этой целью заметим прежде всего, что формула (3), очевидно, может быть переписана в виде

$$\bar{a}_r = \frac{1}{\Omega(E)} \sum_{k=1}^{\infty} k A_k,$$
где $\Lambda_k$ означает число тех основных состояний (т. е. основных функций нашего базиса), в которых $a_r = k \ (k = 1, 2, \ldots)$. Обозначим теперь через $M_k \ (k = 1, 2, \ldots)$ число тех основных состояний, в которых $a_r \geq k$, так что

$$\Lambda_k = M_k - M_{k+1} \quad (k = 1, 2, \ldots).$$

Но $M_k$ равно, очевидно, числу таких решений уравнения (2), в которых $a_r \geq k$, или, что то же, числу решений в целых $b_r \geq 0$ уравнения

$$b_1\varepsilon_1 + \ldots + b_{r-1}\varepsilon_{r-1} + (k - h_r)\varepsilon_r + b_{r+1}\varepsilon_{r+1} + \ldots = E,$$

или

$$\sum_{i=1}^{\infty} b_i\varepsilon_i = E - k\varepsilon_r;$$

но это означает, что $M_k = \Omega (E - k\varepsilon_r) \ (k = 1, 2, \ldots)$, и следовательно,

$$\Lambda_k = \Omega (E - k\varepsilon_r) - \Omega (E -(k+1)\varepsilon_r) \quad (k = 1, 2, \ldots);$$

поэтому формула (4) даёт:

$$\overline{a_r} = \frac{1}{\Omega (E)} \sum_{k=1}^{\infty} k \left\{ \Omega (E - k\varepsilon_r) - \Omega (E -(k+1)\varepsilon_r) \right\},$$

откуда элементарным преобразованием Абеля находим:

$$\overline{a_r} = \sum_{k=1}^{\infty} \frac{\Omega (E - k\varepsilon_r)}{\Omega (E)}. \quad (5)$$

Эта простая формула и будет исходным пунктом наших дальнейших расчётов.

Обращаемся теперь к выводу аналогичного выражения для $\overline{a_r} a_s$. Допустим сначала, что $r \neq s$. Пусть $\Lambda_{kl}$ означает число решений уравнения (2), в которых $a_r = k, \ a_s = l$. Тогда, очевидно,

$$\overline{a_r} a_s = \frac{1}{\Omega (E)} \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} kl\Lambda_{kl}; \quad (6)$$
с другой стороны, если мы обозначим через $M_{kl}$ число решений уравнения (2), в которых $a_r \geqslant k$, $a_s \geqslant l$, то, как легко подсчитать,

$$
\Lambda_{kl} = M_{kl} - M_{k+1, l} - M_{k, l+1} + M_{k+1, l+1}.
$$

(7)

Далее мы аналогично предыдущему замечаем, что $M_{kl}$ может быть представлено как число решений в целых $b_i \geqslant 0$ уравнения

$$
b_1 \varepsilon_1 + b_2 \varepsilon_2 + \ldots + (k + b_r) \varepsilon_r + \ldots + (l + b_s) \varepsilon_s + \ldots = E,
$$

или

$$
\sum_{i=1}^{\infty} b_i \varepsilon_i = E - k \varepsilon_r - l \varepsilon_s,
$$

т. е.

$$
M_{kl} = \Omega (E - k \varepsilon_r - l \varepsilon_s);
$$

поэтому формулы (6) и (7) дают:

$$
\overline{a_r a_s} = \frac{1}{\Omega (E)} \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} kl \left\{ \Omega (E - k \varepsilon_r - l \varepsilon_s) - \Omega (E - (k + 1) \varepsilon_r - l \varepsilon_s) - \Omega (E - k \varepsilon_r - (l + 1) \varepsilon_s) + \Omega (E - (k + 1) \varepsilon_r - (l + 1) \varepsilon_s) \right\},
$$

или, после двукратного применения преобразования Абеля,

$$
\overline{a_r a_s} = \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \frac{\Omega (E - k \varepsilon_r - l \varepsilon_s)}{\Omega (E)}.
$$

(8)

Наконец, в случае $r = s$ речь идёт о выражении нужным нам образом величины $\overline{a_r^2}$. Сохраняя прежние обозначения, мы имеем:

$$
\overline{a_r^2} = \frac{1}{\Omega (E)} \sum_{k=1}^{\infty} k^2 \Lambda_k = \frac{1}{\Omega (E)} \sum_{k=1}^{\infty} k^2 (M_k - M_{k+1}) =
$$

$$
= \frac{1}{\Omega (E)} \sum_{k=1}^{\infty} (2k - 1) M_k.
$$
где последнее выражение получено из предыдущего преобразования Абеля. Так как мы нашли, что $M_k = \Omega (E - k\epsilon_r)$ $(k = 1, 2, \ldots)$, то

$$\overline{a_r^2} = \sum_{k=1}^{\infty} (2k-1) \frac{\Omega (E - k\epsilon_r)}{\Omega (E)} .$$

(9)

Этим завершается первый этап нашего исследования. Формулы (5), (8) и (9) дают нам выражения чисел $\overline{a_r}$, $a_r a_s$ $(r \neq s)$ и $\overline{a_r^2}$ через различные значения структурной функции $\Omega$ и тем самым редуцируют нашу задачу к отысканию удобного аналитического выражения для этой последней.

§ 3. Редукция к задаче теории вероятностей

Вторым шагом в решении нашей задачи будет сведение исследования структурной функции $\Omega$ нашей системы к сравнительно несложною проблеме теории вероятностей. Этот шаг имеет большое методологическое значение, так как именно здесь создаётся мост между основной проблематикой статистической механики и предельными задачами теории вероятностей — тот мост, с помощью которого хорошо разработанные аналитические методы современной теории вероятностей становятся удобным орудием исследования в статистической физике.

Однако, прежде чем осуществить эту редукцию, мы должны будем несколько детальнее рассмотреть последовательность

$$\epsilon_1, \epsilon_2, \ldots, \epsilon_r, \ldots$$

(10)
возможных уровней энергии наших фотонов. Пусть сначала наш «фотонный газ» заключён в сосуде объёма 1. Квантовая физика учит, что число $g_k$ возможных уровней энергии, заключённых между $k$ и $k + 1$, в среднем безгранично возрастает вместе с $k$ (*); так как мы в силу общего соглашения (§ 2 Введения) считаем все $\epsilon_r$ в порядке приближения целыми числами, то мы должны допустить, что натуральное число $k$ встречается $g_k$ раз в последовательности (10). Так обстоит

*) Закон этого возрастания мы найдем в § 5 настоящей главы.
дело в случае, когда объём, занимаемый нашей системой, равен 1; если этот объём равен V (мы всегда будем считать V большим целым числом), то, как показывает квантовая физика, число возможных уровней энергии в любом промежутке увеличивается в V раз; мы должны, таким образом, считать, что каждое натуральное число k повторяется \( V g_k \) раз в последовательности (10); в частности, это означает, что последовательность (10), а следовательно, и структурная функция \( \Omega(E) \) нашей системы существенным образом зависит от занимаемого этой системой объёма V. Мы сейчас же заметим, что те асимптотические формулы, к выводу которых мы перейдём в ближайших параграфах, будут основаны на предположении, что \( E \) и V — бесконечно большие, растущие в некотором постоянном отношении (постоянная объёмная плотность энергии \( \frac{E}{V} \)). Практически это означает следующее: за единицу энергии мы принимаем примерно среднюю энергию отдельного фотона, так что \( E \) становится очень большой величиной (порядка числа N фотонов); для объёма же мы выбираем единицу столь малую, чтобы в таком единичном объёме в среднем помещалось лишь небольшое число фотонов; тогда число V также получает порядок числа N.

Сделаем, наконец, следующее важное для дальнейшего замечание. Пусть мы имеем какую-либо функцию \( f(x) \), для которой ряд

\[
\sum_{r=1}^{\infty} f(\varepsilon_r)
\]  

абсолютно сходится; тогда в силу предыдущего

\[
\sum_{r=1}^{\infty} f(\varepsilon_r) = V \sum_{k=1}^{\infty} g_k f(k);
\]

так как сумма, стоящая в правой части этого равенства, не зависит от V, то всякая величина вида (11) в наших асимптотических формулах должна пониматься как бесконечно большая, пропорциональная V.

Перейдём теперь к установлению вероятностного выражения для структурной функции \( \Omega \). Пусть \( \beta \) — положительное число, пока совершенно произвольное. Рассмотрим случайную величину, возможными значениями которой служат неотрицательные...
ные целые числа — кратные некоторого фиксированного натурального числа $k$, т. е. числа

$$0, k, 2k, \ldots, nk, \ldots,$$

причём вероятность значения $nk$ равна

$$\frac{e^{-\beta nk}}{\sum_{n=0}^{\infty} e^{-\beta nk}} = (1 - e^{-\beta k}) e^{-\beta nk} \quad (n = 0, 1, \ldots).$$

Обозначим этот закон распределения через $p_k(x)$ и рассмотрим сумму бесконечного ряда попарно независимых случайных величин, из которых $g_k$ величин распределены по закону $p_k(x)$ ($k = 1, 2, \ldots$). Докажем, что этот ряд сходится с вероятностью 1. В самом деле, для величины, распределённой по закону $p_k(x)$, вероятность быть отличной от нуля равна $e^{-\beta k}$. Поэтому вероятность того, что по крайней мере один из $g_k$ членов нашего ряда, распределённых по закону $p_k(x)$, будет отличным от нуля, не превосходит $g_k e^{-\beta k}$. Но даваемые квантовой физикой значения чисел $g_k$ таковы, что ряд

$$\sum_{k=1}^{\infty} g_k e^{-\beta k}$$

сходится при любом $\beta > 0$. Поэтому с вероятностью, сколь угодно близкой к единице, можно ожидать, что все члены нашего ряда, начиная с $n$-го, обратятся в нуль, если число $n$ достаточно велико. А так как в этом случае ряд содержит лишь конечное число отличных от нуля членов и потому тривиальным образом сходится, то вероятность сходимости этого ряда сколь угодно близка к единице и, следовательно, равна единице ч. и. т. д.

Таким образом, сумма нашего ряда есть случайная величина, закон распределения которой мы обозначим через $P(x)$. Пусть теперь $S_V$ означает сумму $V$ взаимно независимых случайных величин

$$S_V = \sum_{i=1}^{V} X_i,$$
каждая из которых распределена по закону $P(x)$. Постараемся
найти закон распределения величины $S_V$, т. е. вероятность

$$P(S_V = q)$$

того, что $S_V$ равна данному неотрицательному целому числу $q$.

Величина $S_V$ есть сумма $V$ взаимно независимых величин $X_i$, каждая из которых, будучи распределена по закону $P(x)$, может в свою очередь рассматриваться как сумма бесконечного ряда взаимно независимых случайных величин, среди которых имеется $g_k$ величин, распределённых по закону $p_k(x)$. Таким образом, мы можем в конечном счёте рассматривать величину $S_V$ как сумму бесконечного ряда взаимно независимых случайных величин, среди которых имеется $Vg_k$ величин, распределённых по закону $p_k(x)$; эти величины мы будем обозначать через $x_{kl}$ ($1 \leq l \leq Vg_k$), так что

$$S_V = \sum_{k=1}^{V} \sum_{l=1}^{Vg_k} x_{kl},$$

где $x_{kl}$ подчинено закону $p_k(x)$, и все $x_{kl}$ попарно независимы. Для того чтобы мы имели $S_V = q$, нужно, следовательно, чтобы было

$$\sum_{k=1}^{\infty} \sum_{l=1}^{Vg_k} x_{kl} = q. \quad (13)$$

Но случайная величина $x_{kl}$ подчинена закону распределения $p_k(x)$ и, следовательно, может принимать лишь значения вида $kn_{kl}$, где $n_{kl} \geq 0$ — целое число. Поэтому для реализации равенства (13), вероятность которого нам надо определить, система значений $kn_{kl}$, принимаемых величинами $x_{kl}$, должна быть такова, чтобы

$$\sum_{k=1}^{\infty} \sum_{l=1}^{Vg_k} kn_{kl} = q. \quad (K)$$

В силу взаимной независимости величин $x_{kl}$ вероятность определённой системы значений $x_{kl} = kn_{kl}$ ($k = 1, 2, \ldots$,
§ 3  РЕДУКЦИЯ К ЗАДАЧЕ ТЕОРИИ ВЕРОЯТНОСТЕЙ  

\[ l = 1, 2, \ldots, Vg_k \) равна

\[
\prod_{k=1}^{\infty} \prod_{l=1}^{Vg_k} P(x_{kl} = kn_{kl}) = \prod_{k=1}^{\infty} \prod_{l=1}^{Vg_k} (1 - e^{-\beta k}) e^{-\beta kn_{kl}} = \\
= \prod_{k=1}^{\infty} \left\{ (1 - e^{-\beta k})^{Vg_k} \prod_{l=1}^{Vg_k} e^{-\beta kn_{kl}} \right\} = \\
= \prod_{k=1}^{\infty} (1 - e^{-\beta k})^{Vg_k} e^{-\beta \sum_{k=1}^{\infty} \sum_{l=1}^{Vg_k} kn_{kl}}.
\]

Вероятность же соотношений (13) равна сумме вероятностей только что написанного вида, распространённой на все системы значений величин \( x_{kl} \), удовлетворяющие условию (K), что можно записать в виде

\[
\sum_{(K)} \left\{ \prod_{k=1}^{\infty} (1 - e^{-\beta k})^{Vg_k} e^{-\beta \sum_{k=1}^{\infty} \sum_{l=1}^{Vg_k} kn_{kl}} \right\} = \\
= \prod_{k=1}^{\infty} (1 - e^{-\beta k})^{Vg_k} e^{-\beta q} \sum_{(K)} 1.
\]

Здесь первый множитель

\[
\prod_{k=1}^{\infty} (1 - e^{-\beta k})^{Vg_k}
\]

есть вполне определённая функция параметров \( \beta \) и \( V \); полагая

\[
\prod_{k=1}^{\infty} (1 - e^{-\beta k})^{g_k} = \Phi (\beta),
\]

мы имеем:

\[
\prod_{k=1}^{\infty} (1 - e^{-\beta k})^{Vg_k} = \{ \Phi (\beta) \}^{-V}.
\]

Последний же множитель \( \sum_{(K)} 1 \) есть число решений уравнения (K) в целых \( n_{kl} \geq 0 \). Но легко видеть, что уравнение (K) только системой обозначений отличается от уравнения

\[
\sum_{r=1}^{\infty} a_r \varepsilon_r = q.
\]
В самом деле, число уровней \( \varepsilon_r \), равных \( k \), в уравнении (14) равно \( Vg_k \), и коэффициенты \( a_r \) при этих уровнях мы можем иначе обозначить через \( n_{kl} \) (\( 1 \leq l \leq Vg_k \)), после чего уравнение (14) переходит в уравнение (K). Таким образом,

\[ \sum_{(K)} 1 = \Omega(q), \]

и мы находим:

\[ P(S_V = q) = \{\Phi(\beta)\}^{-1} V e^{-\beta q} \Omega(q), \]

откуда

\[ \Omega(q) = \{\Phi(\beta)\} V e^{\beta q} P(S_V = q). \quad (15) \]

Это выражение структурной функции \( \Omega(q) \) и было нашей целью*. Мы видим, что \( \Omega(q) \) весьма просто выражается через закон распределения \( P(S_V = q) \); так как случайная величина \( S_V \) определена нами как сумма очень большого числа \( V \) одинаково распределённых и взаимно независимых случайных величин, то для приближённого вычисления вероятности \( P(S_V = q) \) мы располагаем весьма детально разработанным аналитическим аппаратом теории вероятностей. Наличие в формуле (15) множителя \( \{\Phi(\beta)\} V \) (независимого от \( q \)) не может осложнить нашей задачи уже потому, что формулы (5), (8) и (9) предыдущего параграфа содержат только отношения различных значений функции \( \Omega(q) \), в которых этот множитель полностью сокращается.

Заметим, наконец, что параметр \( \beta \) в формуле (15) может иметь любое положительное значение; в дальнейшем мы выберем это значение так, чтобы получаемые нами асимптотические формулы приобрели возможно более простую форму.

§ 4. Применение предельной теоремы теории вероятностей

Мы переходим к третьему, последнему шагу нашего исследования — применению доказанной в § 4 гл. I предельной теоремы теории вероятностей (в локальной форме).

*) Оно показывает, в частности, что возможными значениями случайной величины \( S_V \) служат возможные уровни энергии данной системы. Это легко было бы усмотреть и непосредственно из определения величины \( S_V \).
Прежде всего нам надо выбрать значение параметра $\beta$. С этой целью мы докажем сперва следующее вспомогательное предложение:

Лемма. Каково бы ни было положительное число $A$, уравнение

$$\frac{d \ln \Phi(\beta)}{d \beta} + A = 0$$

имеет единственный положительный корень.

В самом деле, рассмотрим функцию

$$\Phi_A(\beta) = e^{A\beta} \Phi(\beta).$$

Очевидно, что $\Phi(\beta)$ (а следовательно, и $\Phi_A(\beta)$) безгранично возрастает при $\beta \to 0$. С другой стороны, так как всегда $\Phi(\beta) > 1$, то

$$\Phi_A(\beta) > e^{A\beta} \to \infty$$

и при $\beta \to \infty$. Таким образом, функция $\Phi_A(\beta)$, а следовательно, и функция $\ln \Phi_A(\beta)$, безгранично возрастает как при $\beta \to 0$, так и при $\beta \to \infty$. С другой стороны,

$$\ln \Phi_A(\beta) = A\beta + \ln \Phi(\beta),$$

и следовательно,

$$\frac{d^2 \ln \Phi_A(\beta)}{d \beta^2} = \frac{d^2 \ln \Phi(\beta)}{d \beta^2} = \sum_{k=1}^{\infty} \frac{k^2 g_k e^{-\beta k}}{(1 - e^{-\beta k})^2} > 0,$$

так что функция $\ln \Phi_A(\beta)$ выпукла на всей полупрямой $(0, +\infty)$. Сопоставление этих свойств показывает, что

$$\frac{d \ln \Phi_A(\beta)}{d \beta} = A + \frac{d \ln \Phi(\beta)}{d \beta},$$

обращается в нуль в одной и только одной точке этой полупрямой, что и доказывает лемму.

Пусть наш «фотонный газ» имеет энергию $E$ и занимает объём $V$. Тогда мы положим $\beta$ равным (существующему и единственному в силу нашей леммы) корню уравнения

$$\frac{d \ln \Phi(\beta)}{d \beta} + \frac{E}{V} = 0. \quad (16)$$

Так как наши асимптотические формулы, как уже замечено, будут выводиться в предположении, что $E$ и $V$ безгранично.
возрастают, а отношение $\frac{E}{V}$ остаётся постоянным, то в этих формулах число $\beta$ может, таким образом, рассматриваться как постоянное.

Величина $S_V$, для закона распределения которой мы хотим найти асимптотическое выражение, есть сумма $V$ взаимно неизвестных случайных величин $x_i \ (1 \leq i \leq V)$, каждая из которых распределена по закону $P(x)$; этот закон (кстати сказать, не зависящий от $E$ и $V$) есть закон распределения суммы

$$X_i = \sum_{k=1}^{\infty} \sum_{l=1}^{g_k} x_{kl},$$

где величина $x_{kl}$ распределена по закону $p_k(x)$. Математическое ожидание и дисперсия величины $x_{kl}$ поэтому, как легко вычислить, соответственно равны

$$a_{kl} = \frac{ke^{-\beta k}}{1 - e^{-\beta k}},$$

и

$$b_{kl} = \frac{k^2e^{-\beta k}}{(1 - e^{-\beta k})^2}.$$  

Отсюда математическое ожидание и дисперсия каждой из величин $X_i$, как показывает столь же элементарный подсчёт, равны соответственно

$$a = \sum_{k=1}^{\infty} \sum_{l=1}^{g_k} a_{kl} = \sum_{k=1}^{\infty} \frac{kg_k e^{-\beta k}}{1 - e^{-\beta k}} = -\frac{d}{d\beta} \ln \Phi(\beta)$$

и

$$b = \sum_{k=1}^{\infty} \sum_{l=1}^{g_k} b_{kl} = \sum_{k=1}^{\infty} \frac{k^2g_k e^{-\beta k}}{(1 - e^{-\beta k})^2} = \frac{d^2}{d\beta^2} \ln \Phi(\beta).$$

Для математического ожидания $A$ и дисперсии $B$ величины $S_V$ мы поэтому находим выражения

$$A = Va = -V \frac{d \ln \Phi}{d\beta}, \quad B = Vb = V \frac{d^2 \ln \Phi}{d\beta^2}.$$  

В силу соотношения (16) (т. е. в силу выбранного нами значения параметра $\beta$) мы, очевидно, имеем $A = E$. Поэтому
Применение одномерной локальной предельной теоремы в её упрощённой форме (формула (28) § 4 гл. I) даёт:

\[ P(S_{V} = E + u) = \frac{d}{\sqrt{2\pi B}} e^{-\frac{u^2}{2B}} + O\left(\frac{1+|u|}{V^{q_3}}\right), \]  

где \( d \) означает шаг последовательности возможных уровней энергии системы, а \( u \) — любое целое число, для которого \( E + u \) служит одним из таких возможных уровней.

Но

\[ e^{-\frac{u^2}{2B}} = 1 + O\left(\frac{u^2}{V}\right); \]

поэтому

\[ P(S_{V} = E + u) = \frac{d}{\sqrt{2\pi B}} + O\left(\frac{1+u^2}{V^{q_3}}\right). \]

Полагая в формуле (15) \( q = E + u \), мы находим:

\[ \Omega(E + u) = \{\Phi(\beta)\} V e^{\beta E} \left(\frac{d}{\sqrt{2\pi B}} + O\left(\frac{1+u^2}{V^{q_3}}\right)\right), \]

и, в частности, при \( u = 0 \)

\[ \Omega(E) = \{\Phi(\beta)\} V e^{\beta E} \left(\frac{d}{\sqrt{2\pi B}} + O\left(\frac{1}{V^{q_3}}\right)\right). \]

Отсюда

\[ \frac{\Omega(E + u)}{\Omega(E)} = e^{\beta u} \frac{1 + O\left(\frac{1+u^2}{V}\right)}{1 + O\left(\frac{1}{V}\right)} = e^{\beta u} \left(1 + O\left(\frac{1+u^2}{V}\right)\right). \]  

Теперь мы легко получаем простые асимптотические выражения для средних значений чисел заполнения \( \bar{a}_r \). Формула (5) § 2 в силу (18) даёт:

\[ \bar{a}_r = \sum_{m=1}^{\infty} e^{-\beta m_{sr}} \left\{1 + O\left(\frac{1+m^2 e_{sr}^2}{V}\right)\right\} = \]

\[ = \sum_{m=1}^{\infty} e^{-\beta m_{sr}} + O\left(\frac{1}{V}\right) = \frac{1}{e^{\beta e_{sr}} - 1} + O\left(\frac{1}{V}\right). \]
Так как среди уровней \( \varepsilon_r \) число \( k \) повторяется \( V g_k \) раз, то среднее число фотонов с энергией \( k \) отсюда равно

\[
\frac{V g_k}{e^{3k} - 1} + O(1).
\]  

(20)

Если \( \mathcal{A} \) есть сумматорная величина

\[
\mathcal{A} = \sum_i a_i,
\]

где величина \( a_i \) зависит лишь от координат \( i \)-го фотона, то микроканоническое среднее математического ожидания величины \( \mathcal{A} \) в силу формулы (16) § 5 гл. III равно

\[
\overline{\mathcal{A}} = \sum_{r=1}^{\infty} a_r \lambda_r,
\]

где \( \lambda_r \) — математическое ожидание величины \( a_i \) (предполагаемое одинаковым для всех \( i \)) в том состоянии фотона, которое характеризуется уровнем энергии \( \varepsilon_r \) (напомним, что величины \( \lambda_r \) зависят лишь от вида функции \( a_i \) и определяются независимо от каких-либо статистических соображений). Первое из равенств (19) даёт нам:

\[
\overline{\mathcal{A}} = \sum_{r=1}^{\infty} \lambda_r \sum_{m=1}^{\infty} e^{-\beta m \varepsilon_r} \left\{ 1 + O \left( \frac{1 + m^2 \varepsilon_r^2}{V} \right) \right\} =
\]

\[
= \sum_{r=1}^{\infty} \frac{\lambda_r}{e^{\beta \varepsilon_r} - 1} + O \left( \frac{\sum_{r=1}^{\infty} \frac{1}{e^{\beta \varepsilon_r} - 1} + \sum_{r=1}^{\infty} \frac{\lambda_r}{\varepsilon_r^2} \sum_{m=1}^{\infty} m^2 e^{-\beta m \varepsilon_r}}{V} \right).
\]

Если мы допустим (как это бывает в большинстве реальных физических задач), что \( \lambda_r \) однозначно зависит от \( \varepsilon_r \) (т. е. что величина \( a_i \) имеет одно и то же математическое ожидание для всех состояний фотона, соответствующих одному и тому же уровню энергии), то все ряды в правой части этого равенства имеют вид (11) (§ 3), а потому в силу формулы (12) представляют собою величины, пропорциональные \( V \) (само собою разумеется, что эти ряды предполагаются сходящимися). Отсюда мы находим:

\[
\overline{\mathcal{A}} = \sum_{r=1}^{\infty} \frac{\lambda_r}{e^{\beta \varepsilon_r} - 1} + O(1) = V \sum_{k=1}^{\infty} \frac{g_k \Lambda_k}{e^{\beta k} - 1} + O(1),
\]  

(21)
где $\Lambda_k = \lambda_r$ при $\varepsilon_r = k$. Таким образом, для сумматорной величины микроканоническое среднее всегда асимптотически пропорционально $V$, что, конечно, непосредственно очевидно. Полученная нами асимптотическая формула (21) точно определяет множитель пропорциональности и показывает, что относительная погрешность полученной оценки чрезвычайно мала.

В случае фотонов одной из важнейших сумматорных функций является число фотонов $N$, которое, как было указано в самом начале этой главы, при данных $E$ и $V$ может принимать различные значения в зависимости от того, в каком состоянии находится система; более того, в данном определённом состоянии системы число $N$, вообще говоря, не имеет фиксированного значения, а получает лишь определённый закон распределения, так что имеет смысл говорить лишь о микроканоническом среднем $\overline{N}$ математического ожидания числа $N$.

Если $\mathfrak{N} = N$, то, очевидно, $a_r = 1$, а следовательно, и $\lambda_r = 1 (r = 1, 2, \ldots)$. Поэтому мы находим:

$$\overline{N} = \sum_{r=1}^{\infty} a_r = \sum_{r=1}^{\infty} \frac{1}{e^{\varepsilon_r} - 1} + O(1) =$$

$$= V \sum_{k=1}^{\infty} \frac{g_k}{e^{\beta k} - 1} + O(1);$$

среднее число фотонов, приходящееся на единицу объёма, отсюда равно

$$\frac{\overline{N}}{V} = \sum_{k=1}^{\infty} \frac{g_k}{e^{\beta k} - 1} + O \left( \frac{1}{V} \right).$$

§ 5. Формула Планка

Для того чтобы формулы, выведенные нами в предыдущем параграфе, приобрели реальное значение, мы должны, конечно, определить числа $g_k$ и установить физический смысл параметра $\beta$. Вторую из этих задач мы рассмотрим подробно дальше, так как универсальное значение, придаваемое теорией параметру $\beta$, естественно требует для своего обоснования
общирного конкретного материала, которого в данный момент у нас ещё слишком мало. Здесь же, предвосхищая результат этого более детального обоснования, мы просто укажем, что теория во всех случаях приписывает параметру $\beta$ значение $\frac{1}{kT}$, где $T$ — абсолютная температура системы, а $k$ — универсальная константа (так называемая постоянная Больцмана). Таким образом, для любых, сколь угодно сложных систем, состоящих из частиц любого типа, параметр $\beta$ всегда характеризует собою температуру.

Нахождение чисел $g_k$, к которому мы теперь переходим, может быть произведено несколькими физически различными способами. Можно при этом исходить либо из волновых представлений о природе света, либо частью из волновых и частью из корпускулярных, либо, наконец, из чисто корпускулярных представлений. Мы, естественно, выберем последний путь, так как только он полностью соответствует всему духу излагаемой теории. Надо, впрочем, отметить, что вся задача определения чисел $g_k$ никак не связана с какими-либо статистическими рассмотрениями, и мы приводим здесь её решение лишь с целью дать нашим статистическим выводам такие применения, в которых они непосредственно могли бы быть сближены с данными опыта.

По своему определению $Vg_r$ означает число таких линейно независимых стационарных состояний фотона, в которых его энергия заключена между $r$ и $r + 1$. Мы должны поэтому написать так называемое «свободное от времени уравнение Шредингера»

$$\text{HU} = E\text{U},$$

где $H$ — оператор энергии фотона, а $E$ — постоянное число, и найти число его линейно независимых решений для значений $E$, заключённых между $r$ и $r + 1$. Однако обычное выражение для оператора энергии частицы, свободной от воздействия внешних сил *),

$$H = -\frac{\hbar^2}{8\pi^2m^2} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right),$$

*) С целью придадь выводам более обычную форму мы в настоящем параграфе обозначаем постоянную Планка через $\frac{\hbar}{2\pi}$ вместо $\hbar$. В физике приняты обе системы обозначения.
в случае фотонов не может, конечно, быть применен, так как для фотона \( m = 0 \), и связь энергии \( \varepsilon \) с импульсом \( p \) не даётся уравнением

\[
\varepsilon = \frac{1}{2m} p^2,
\]

как для материальной частицы. Чтобы правильно определить оператор \( \mathbf{H} \), мы должны исходить из общей связи между энергией и импульсом частицы, даваемой теорией относительности:

\[
\varepsilon = c V \sqrt{m^2 c^2 + p^2} \quad (*)
\]

и имеющей место для частиц любого типа; в случае фотонов \( m = 0 \) и

\[
\varepsilon = cp = c V \sqrt{p_x^2 + p_y^2 + p_z^2}
\]

(в случае материальных частиц \( p \ll mc \) и \( \varepsilon \approx mc^2 + \frac{p^2}{2m} \)).

Таким образом, для фотонов \( p^2 \) пропорционально не \( \varepsilon \), а \( \varepsilon^2 \), и оператор величины \( c^2 p^2 \), имеющий вид

\[
- \frac{c^2 \hbar^2}{4\pi^2} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right),
\]

соответствует не энергии, а её квадрату. Из этого, как мы знаем (гл. II, § 3), следует, что собственными значениями оператора \( 22 \) служат квадраты собственных значений оператора \( \mathbf{H} \), а собственные функции у этих двух операторов одни и те же; таким образом, если мы напишем уравнение

\[
- \frac{c^2 \hbar^2}{4\pi^2} \left( \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} \right) = E^2 U,
\]

или, что то же,

\[
\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} = - \frac{4\pi^2 E^2}{\hbar^2} U,
\]

то решения его будут собственными функциями оператора энергии фотона, соответствующими собственному значению \( E \) этой энергии. Потенциальная энергия фотона, в предположении

ннии отсутствия воздействующих на систему внешних сил, редуцируется к потенциалу стенок сосуда, в котором помещается наш фотонный газ, и внутри сосуда есть постоянная величина, которую мы можем принять равной нулю. Известно, что линейным базисом решений уравнения (23) может служить выражение

$$U = C \sin (ax + \alpha) \sin (by + \beta) \sin (cz + \gamma),$$

где $a, b, c, \alpha, \beta, \gamma, C$ — постоянные, из которых первые три связаны соотношением

$$a^2 + b^2 + c^2 = \frac{4\pi^2 E^2}{h^2c^2}. \tag{25}$$

При этом, так как вероятность найти фотон в соседстве точки $(x, y, z)$, пропорциональная $|U(x, y, z)|^2$, должна обращаться в нуль вне сосуда (а следовательно, по непрерывности и на его стенках), то мы получаем совершенно определённые граничные условия, для выражений которых надо принять какое-либо фиксированное допущение о форме сосуда. Мы предположим, что этот сосуд имеет параллелепипед $0 \leq x \leq l_1$, $0 \leq y \leq l_2$, $0 \leq z \leq l_3$ объёма $l_1l_2l_3 = V$. Тогда при выполнении хотя бы одного из шести условий $x = 0, y = 0, z = 0, x = l_1, y = l_2, z = l_3$ мы должны иметь $U = 0$. Первые три условия, очевидно, требуют, чтобы в выражении (24) было $a = \beta = \gamma = 0$, так что мы получаем:

$$U = C \sin ax \sin by \sin cz, \quad a^2 + b^2 + c^2 = \frac{4\pi^2 E^2}{h^2c^2}. $$

Требование $U = 0$ при $x = l_1$ (независимо от значений $y$ и $z$) влечёт за собою $a = \frac{n_1\pi}{l_1}$, где $n_1$ — целое число; к аналогичным соотношениям приводят и два последних требования, так что мы должны иметь:

$$a = \frac{n_1\pi}{l_1}, \quad b = \frac{n_2\pi}{l_2}, \quad c = \frac{n_3\pi}{l_3},$$

где $n_1, n_2, n_3$ — целые числа, в силу (25) удовлетворяющие соотношению

$$\frac{n_1^2}{l_1^2} + \frac{n_2^2}{l_2^2} + \frac{n_3^2}{l_3^2} = \frac{4E^2}{h^2c^2}. $$
Таким образом, возможными значениями энергии стационарных состояний (т. е. энергетическими уровнями фотона) служат числа $E$, квадраты которых имеют форму

$$\frac{\hbar^2 c^2}{4} \left( \frac{n_1^2}{l_1^2} + \frac{n_2^2}{l_2^2} + \frac{n_3^2}{l_3^2} \right),$$

где $n_1, n_2, n_3$ — целые числа; каждая тройка целых чисел $n_1, n_2, n_3$ даёт нам одну из линейно независимых собственных функций; точнее, мы должны при этом фиксировать знаки чисел $n_1, n_2, n_3$, так как замена, например, $n_1$ на $-n_1$ меняет $\mathcal{U}$ на $-\mathcal{U}$ и не приводит, следовательно, к собственной функции, отличной от предшествующих; для определённости мы условимся поэтому в дальнейшем рассматривать только значения $n_1 \geqslant 0, n_2 \geqslant 0, n_3 \geqslant 0$.

Число линейно независимых собственных функций, соответствующих значениям энергии $E \leqslant r$, где $r$ — данное натуральное число, равно, таким образом, числу решений в целых $n_1 \geqslant 0, n_2 \geqslant 0, n_3 \geqslant 0$ неравенства

$$\frac{\hbar^2 c^2}{4} \left( \frac{n_1^2}{l_1^2} + \frac{n_2^2}{l_2^2} + \frac{n_3^2}{l_3^2} \right) \leqslant r^2,$$

или, что то же, числу точек первого октаэдра, обладающих целочисленными координатами и лежащих внутри или на границе эллипсоида

$$\frac{\hbar^2 c^2}{4} \left( \frac{x^2}{l_1^2} + \frac{y^2}{l_2^2} + \frac{z^2}{l_3^2} \right) = r^2. \quad (26)$$

Это число асимптотически (при больших $r$) равно одной восьмой части объёма эллипсоида (26), т. е. равно

$$Q(r) = \frac{4}{3} \pi \frac{r^3}{\hbar^3 c^3} l_1 l_2 l_3 = \frac{4\pi V}{3\hbar^3 c^3} r^3.$$

Отсюда число собственных функций, принадлежащих к уровням энергии, заключённым между $r$ и $r + dr$, приближённо равно

$$Q'(r) \, dr = \frac{4\pi V}{\hbar^3 c^3} r^2 \, dr;$$
при этом интервал $dr$ должен быть достаточно большим, чтобы приближённая формула имела реальное значение, и вместе с тем малым сравнительно с $r$. С другой стороны, это же число в наших обозначениях равно

$$V \sum_{k=r}^{r+dr} g_k,$$

так что

$$\sum_{k=r}^{r+dr} g_k = \frac{4\pi}{\hbar^2 c^8} r^2 dr.$$

Таким образом, для $r < k < r + dr$ значение числа $g_k$ в среднем равно

$$\frac{4\pi}{\hbar^2 c^8} r^2;$$

для большинства расчётов можно без существенного искажения результатов прямо положить

$$g_k = \frac{4\pi}{\hbar^2 c^8} k^2, \quad (27)$$

по крайней мере для больших значений $k$. Надо, однако, помнить, что это лишь приём приближённого вычисления и что на самом деле число $g_k$ зависит от арифметической природы числа $k$; в частности, могут быть сколь угодно большие числа $k$, не встречающиеся среди уровней энергии фотона и для которых поэтому $g_k = 0$.

Число (27) физики ещё удваивают в силу квантовомеханических соображений, которые с точки зрения волновой теории соответствуют возможности различной поляризации света. Дело в том, что функции $U$, кроме своих «классических» аргументов $x$, $y$, $z$, зависят ещё от особой квантовой переменной («спин»), способной принимать только два значения; каждая из собственных функций, которые мы до сих пор определяли, описывает поэтому не одно, а два различных состояния фотона, соответственно двум значениям спиновой переменной. Таким образом, окончательно мы имеем (приближённо, в среднем)

$$g_r = \frac{8\pi}{\hbar^2 c^8} r^2.$$
Представляя это выражение в формулу (20) предыдущего параграфа, мы для среднего числа фотонов с энергией, заключённой между \( r \) и \( r + 1 \), находим:

\[
\frac{8\pi V r^2}{h^3 c^3 \left( e^{\beta r} - 1 \right)}.
\]

(28)

Энергия \( \varepsilon \) фотона связана с частотою \( \nu \) соответствующего света универсальным соотношением \( \varepsilon = h\nu \), где \( h \) — постоянная Планка. Число фотонов с частотою, заключённой между \( \nu \) и \( \nu + d\nu \), есть поэтому число фотонов с энергией, заключённой между \( h\nu \) и \( h\nu + h\, d\nu \), которое в среднем согласно формуле (28) равно

\[
\frac{8\pi V (h\nu)^2}{h^3 c^3 \left( e^{\beta h\nu} - 1 \right)}
\]

(так как \( \beta = \frac{1}{kT} \)). Так как энергия каждого такого фотона равна \( h\nu \), то средняя энергия, приходящаяся на всю совокупность фотонов с частотами, лежащими между \( \nu \) и \( \nu + d\nu \), равна

\[
\frac{8\pi hv}{c^3 \left( e^{\frac{h\nu}{kT}} - 1 \right)}
\]

(29)

а на единицу объёма

\[
\frac{8\pi h}{c^3 \left( e^{\frac{h\nu}{kT}} - 1 \right)}
\]

Это — известная, сыгравшая большую роль в истории квантовой физики, формула Планка, дающая распределение энергии в спектре так называемого «чёрного излучения». 50 лет тому назад, после ряда неудач старой классической теории, неизменно дававшей для этого распределения законы, расходящиеся с опытными данными, Планк в результате смелой и рискованной по тому времени гипотезы пришёл к своей формуле (29), находящейся в прекрасном согласии с тем, что даёт эксперимент. Вывод этой формулы, предложенный Планком, не имел, конечно, ничего общего с тем, который мы привели здесь, так как Планк твёрдо стоял на позициях волновой теории света, что и не могло быть иначе в ту
Согласно общему принципу статистической физики микроканонические средние физических величин представляют собою те их значения, которые для этих величин в данных условиях предсказывает статистическая теория и которые в целях проверки этой теории подлежат сличению с данными опыта. Эксперимент подтверждает теорию, если даваемое им значение величинь близко к её микроканоническому среднему. Однако, если осредняемая нами величина такова, что возможные для неё значения сильно рассеяны, т. е. в большинстве случаев заметно отличаются друг от друга (а значит, и от микроканонического среднего), то мы отнюдь не можем ожидать, что единичный эксперимент даст нам для неё значение, близкое к вычисленному нами среднему; возможно даже такое положение, когда даваемое экспериментом значение наверняка должно сильно отличаться от среднего (игра, участник которой с вероятностью $\frac{1}{2}$ может выиграть и с такою же вероятностью — проиграть крупную сумму). В таких случаях среднее значение (независимо от метода осреднения) ничего не может предсказать нам относительно результата эксперимента. Не можем мы рассчитывать и на то, что среднее значение результатов длинной серии экспериментов будет близко к нашему микроканоническому среднему: для этого было бы необходимо, чтобы частоты, с которыми различные возможные состояния системы встречаются в нашей серии экспериментов, совпадали с теми весами, которые мы придаём им при микроканоническом осреднении. У нас нет никаких оснований ожидать этого; прежде всего, выбор микроканонического способа осреднения был нами сделан из чисто теоретических соображений, без какого-либо учёта реальных условий; с другой
же стороны очевидно, что реальные частоты, о которых здесь идёт речь, будут весьма различны в различных условиях экспериментирования, так что не может существовать такого способа осреднения, который во всех случаях соответствовал бы реальным условиям.

Но дело обстоит совершенно иначе, если изучаемая нами величина рассеяна слабо, т. е. если возможные её значения в подавляющем большинстве мало отличаются друг от друга. В этом случае, очевидно, при любом (в широких пределах) способе осреднения мы для среднего значения найдём примерно одно и то же число, к которому будет близко подавляющее большинство возможных значений изучаемой нами величины, а значит, и подавляющее большинство результатов нашей серии экспериментов; в этом случае мы и от единичного результата с полным основанием имеем возможность ожидать, что он даст нам значение, близкое к найденному среднему.

Все эти соображения были уже, в сущности, приведены нами в гл. III; мы считаем, однако, нужным вновь и вновь привлекать к ним внимание читателя, так как именно они опи-сывают связь между излагаемой теорией и реальностью.

Как мы видели в § 6 гл. III, при оценке представительности микроканонического среднего $\alpha$ физической величины $\mathcal{A}$ играет существенную роль «микроканоническая дисперсия»

$$E \{(\mathcal{A} - \alpha)^2\} = D (\mathcal{A})$$

этой величины. Если при безгранично возрастающем числе частиц $N$ (и пропорционально ему возрастающей энергии $E$ системы) величина $D (\mathcal{A})$ бесконечно мала сравнительно с $N^2$ (или, что то же, с $E^2$), то мы можем быть уверены, что разность $|\mathcal{A} - \alpha|$ с вероятностью, сколь угодно близкой к единице, будет как угодно малой сравнительно с $|\alpha|$ при любом (в широких пределах) распределении частот возможных состояний системы. Но это как раз и означает то, что мы назвали «представительностью» микроканонического среднего $\alpha$. Таким образом, нужна нам представительность, полностью оправдывающая произвольно выбранный нами метод микроканонического осреднения и вместе с тем позволяющая ожидать близкого совпадения микроканонических средних с опытными
данными, будет гарантирована всякий раз, как мы сможем доказать соотношение

\[ D(\mathcal{A}) = o(E^2). \]

Как мы видели в том же § 6 гл. III, в случае, когда \( \mathcal{A} \) есть сумматорная величина, \( D(\mathcal{A}) \) в любой из трёх основных статистических схем получает выражение

\[ D(\mathcal{A}) = \sum_{r=1}^{\infty} (\mu_r - \lambda_r^2) \bar{a}_r + \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} \lambda_r \lambda_s (\bar{a}_r \bar{a}_s - \bar{a}_r \bar{a}_s). \] (30)

Здесь предположено, что

\[ \mathcal{A} = \sum_{i=1}^{N} a_i, \]

где величина \( a_i \) связана с состоянием одной только \( i \)-й частицы; \( \lambda_r \) и \( \mu_r \) соответственно означают математические ожидания величин \( a_i \) и \( a_i^2 \) (предполагаемые одинаковыми для всех \( i \)) при условии, что \( i \)-я частица находится в стационарном состоянии с энергией \( e_r \). Для оценки микроканонической дисперсии \( D(\mathcal{A}) \) мы должны найти асимптотические выражения для микроканонических средних \( \bar{a}_r \) и \( a_r \bar{a}_s \) чисел заполнения и их попарных произведений. Мы можем при этом предвидеть, что эти асимптотические выражения нам придётся брать достаточно тонкими, так как естественно ожидать, что в разностях \( \bar{a}_r \bar{a}_s - \bar{a}_r \bar{a}_s \) ряд главных членов взаимно уничтожится. И действительно, например, та точность, с которой мы определяли числа \( \bar{a}_r \) в § 4 и которая нам там представлялась вполне удовлетворительной, теперь была бы уже совершенно недостаточной для нашей цели. Поэтому асимптотический расчёт этих чисел мы должны произвести заново, на более точной базе; известные формы локальной предельной теоремы представляют нам для этого все необходимые основания. Упрощённая формула (28) § 4 гл. I теперь уже недостаточна для нашей цели, и мы должны воспользоваться более точной формулой (27) того же параграфа. Это даёт вместо формулы (17) § 4 более точную формулу

\[ P(S_v = E + u) = \frac{d}{V^{2\pi B}} e^{-\frac{u^2}{2B}} + \frac{t_0 + t_1 u}{B^{9/2}} + O\left(\frac{V^{1/2} + |u|}{V^{9/2}}\right), \]
§ 6 | О ПРЕДСТАВИТЕЛЬНОСТИ МИКРОКОСМОГИЧЕСКИХ СРЕДНИХ

где \( d \) и \( u \) имеют прежнее значение, а \( l_0 \) и \( l_1 \) не зависят ни от \( E \), ни от \( u \). Пользуясь же тем, что

\[
    e^{-\frac{u^2}{2B}} = 1 - \frac{u^2}{2B} + O\left(\frac{u^4}{V^4}\right),
\]

мы находим:

\[
    \mathbf{P} (S_V = E + u) = \frac{d}{\sqrt{2\pi B}} + \frac{l_0 + l_1 u + l_2 u^2}{B^{3/2}} + O\left(\frac{\sqrt{V} + u^4}{V^{3/2}}\right), \quad (31)
\]

где \( l_2 = -\frac{1}{2\sqrt{2\pi}} \). Формула (15) § 3 даёт поэтому при \( q = E + u \) более точное, чем мы имели в § 4, выражение структурной функции

\[
    \Omega (E + u) = \{ \Phi (\beta) \} V e^{\beta (E + u)} \left\{ \frac{d}{\sqrt{2\pi B}} + \frac{l_0 + l_1 u + l_2 u^2}{B^{3/2}} + O\left(\frac{\sqrt{V} + u^4}{V^{3/2}}\right) \right\},
\]

и, в частности, при \( u = 0 \)

\[
    \Omega (E) = \{ \Phi (\beta) \} V e^{\beta E} \left\{ \frac{d}{\sqrt{2\pi B}} + \frac{l_0}{B^{3/2}} + O\left(\frac{1}{V^2}\right) \right\}.
\]

Отсюда элементарный подсчёт легко даёт:

\[
    \frac{\Omega (E + u)}{\Omega (E)} = e^{\beta u} \left\{ 1 + \frac{k_1 u - \frac{1}{2} u^2}{B} + O\left(\frac{\sqrt{V} + u^4}{V^{3/2}}\right) \right\}, \quad (32)
\]

где \( k_1 \) — постоянная (не зависящая ни от \( E \), ни от \( u \)). Эта более точная формула призвана заменить собою формулу (18) § 4. Переходя теперь к выводу асимптотических формул для \( a_r \) и \( a_r a_s \), мы для сокращения записи введём следующие обозначения. Положим

\[
    T_r = T_r (\beta) = \frac{1}{e^{\beta a_r} - 1} = \sum_{m=1}^{\infty} e^{-m e_r \beta};
\]

тогда, очевидно,

\[
    \epsilon_r \sum_{m=1}^{\infty} m^k e^{-m e_r \beta} = (-1)^k \frac{d^k T_r}{d \beta^k} = (-1)^k T_r^{(k)}.
\]
Формула (5) § 2 теперь даёт нам (по аналогии с (19) § 4 в силу (32)):

\[
\bar{a}_r = \sum_{m=1}^{\infty} \frac{\Omega(E - m \epsilon_r)}{\Omega(E)} \rightarrow \\
= \sum_{m=1}^{\infty} e^{-\beta m \epsilon_r} \left\{ 1 - \frac{k_1 m \epsilon_r + \frac{1}{2} m^2 \epsilon_r^2}{B} + O\left(\frac{\sqrt{V} + m^4 \epsilon_r^4}{V^2}\right) \right\} = \\
= T_r - \frac{k_1 T_r' - \frac{1}{2} T_r''}{B} + O\left(\frac{T_r\sqrt{V} + T_r^{(4)}}{V^2}\right).
\]

(33)

Это и есть более точная формула, заменяющая собою формулу (19) § 4, которая в наших новых обозначениях имела бы простой вид

\[
\bar{a}_r = T_r + O\left(\frac{1}{V}\right).
\]

Далее, формула (8) § 2 в силу (32) даёт нам при \( r \neq s \)

\[
\bar{a}_r \bar{a}_s = \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \frac{\Omega(E - k \epsilon_r - l \epsilon_s)}{\Omega(E)} = \\
= \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} e^{-\beta (k \epsilon_r + l \epsilon_s)} \left\{ 1 - \frac{1}{B} \left[ k_1 \left( k \epsilon_r + l \epsilon_s \right) + \right.ight.

\[
\left. + \frac{1}{2} (k \epsilon_r + l \epsilon_s)^2 \right] + O\left(\frac{\sqrt{V} + (k \epsilon_r + l \epsilon_s)^4}{V^2}\right) \right\} = \\
= T_r T_s + \frac{1}{B} \left[ k_1 \left( T_r T_s' + T_s T_r' \right) - \frac{1}{2} \left( T_r T_s'' + 2 T_r' T_s' + \\
+ T_s T_r'' \right) \right] + O\left(\frac{V \sqrt{V} T_r T_s + T_r T_s^{(4)} + T_s T_r^{(4)}}{V^2}\right);
\]

в то же время из формулы (33) вытекает

\[
\bar{a}_r \bar{a}_s = T_r T_s + \frac{1}{B} \left[ k_1 \left( T_r T_s' + T_s T_r' \right) - \frac{1}{2} \left( T_r T_s'' + T_s T_r'' \right) \right] + \\
+ O\left(\frac{V \sqrt{V} T_r T_s + T_r T_s^{(4)} + T_s T_r^{(4)}}{V^2}\right).
\]
Из двух последних формул мы находим:

\[
\frac{a_r a_s - a_r a_s}{B} = -\frac{\bar{T}_r T_s'}{B} + O\left(\frac{\sqrt{V^2 T_r T_s + T_r T_s' + T_r T_s'}}{V^2}ight). \tag{34}
\]

Таким образом, коэффициент микроканонической корреляции чисел \(a_r\) и \(a_s\) (\(r \neq s\)), как и естественно было ожидать, всегда отрицателен и по абсолютной величине является бесконечно малой порядка \(\frac{1}{V}\).

Формула (34) даёт нам уже всё, что необходимо для оценки микроканонической дисперсии \(D(\mathbf{v})\). С этой целью мы прежде всего должны заметить, что величины \(\bar{a_r}\) в силу (33) асимптотически равны величинам \(T_r\), которые являются функциями от \(\varepsilon_r\); вследствие этого первый член правой части формулы (30) в силу общей формулы (12) § 3 есть величина, асимптотически пропорциональная \(V\). Во второй (двойной) сумме мы должны различать члены, где \(r \neq s\), и такие, где \(r = s\). В членах первого типа в силу (34) числа \(\bar{a_r a_s} - \bar{a_r a_s}\) асимптотически равны \(\frac{\bar{T}_r T_s'}{B}\), а потому двойное суммирование таких членов на основании той же формулы (12) § 3 даёт величину, асимптотически пропорциональную \(\frac{V^2}{B}\), т. е. опять-таки асимптотически пропорциональную \(V\). Наконец, члены второго типа имеют вид \(\lambda_r^2 (\bar{a}_r^2 - (\bar{a}_r)^2)\) и суммируются по \(r\) от 1 до \(\infty\). Но мы могли бы в полной аналогии с предыдущим найти с помощью формул (9) (§ 2) и (32) асимптотическое выражение для \(\bar{a}_r^2 - (\bar{a}_r)^2\) и убедиться, что

\[
\bar{a}_r^2 - (\bar{a}_r)^2 \approx T_r (T_r + 1);
\]

поэтому сумма членов второго типа асимптотически равна

\[
\sum_{r=1}^{\infty} \lambda_r^2 T_r (T_r + 1),
\]

и следовательно, в силу формулы (12) § 3 асимптотически пропорциональна \(V^*\).

*) Во всех этих оценках, разумеется, предполагается, что \(|\lambda_r|\) и \(\nu_r\) с ростом \(r\) возрастают не слишком быстро, так что все получаемые нами ряды абсолютно сходятся.
Таким образом, формула (30) показывает, что \( D(\overline{A}) \) асимптоматически пропорциональна \( V \). Но, как мы видели выше, для представительности среднего значения \( \overline{A} \) достаточно иметь:

\[
D(\overline{A}) = o(E^2),
\]

или, что то же,

\[
D(\overline{A}) = o(V^2);
\]

таким образом, полученные нами оценки с избытком дают представительность микроканонических средних сумматорных величин. Как мы уже отмечали выше, этим с теоретической стороны оправдывается произвольный выбор микроканонического способа осреднения, а с практической — обусловливается полноценность этих микроканонических средних в качестве материала, подлежащего сличению с экспериментальными данными.
ГЛАВА V
ОСНОВЫ СТАТИСТИКИ МАТЕРИАЛЬНЫХ ЧАСТИЦ

§ 1. Напоминание исходных положений

Методы, развитые в предыдущей главе, позволяют, как мы теперь увидим, построить статистическую теорию и для систем, состоящих из материальных частиц. Основной новый момент, с которым приходится считаться при этом переходе, состоит, как уже было замечено в начале предыдущей главы, в том, что число $N$ частиц, составляющих данную систему, мы теперь должны считать твёрдо фиксированным (так же как и полную энергию $E$ системы). Отсюда следует, что возможные состояния системы при данном значении $E$ её полной энергии на этот раз описываются собственными функциями оператора Н полной энергии системы, состоящей из $N$ частиц данного типа, соответствующими собственному значению $E$ этого оператора.

При этом в случае полной статистики «достигимым» являются все такие собственные функции, в случае же симметрической (соответственно антисимметрической) статистики достижимыми признаются лишь симметрические (соответственно антисимметрические) собственные функции. Во всех случаях из семейства всех достижимых собственных функций, соответствующих уровню энергии $E$, мы выбираем (конечный) линейный базис взаимно ортогональных и нормированных функций. Число членов этого базиса зависит от $N$ и $E$; мы будем обозначать его через $\Omega (N, E)$ и называть структурной функцией данной системы. Эта необходимость рассматривать в качестве структурной функции число $\Omega (N, E)$, зависящее от двух аргументов, и является причиной того, что статистика материальных частиц в формальном отношении сложнее статистики фотонов.
Мы видели в § 4 гл. II, что в любой из трёх основных статистических схем функции, составляющие вышеупомянутый линейный базис, могут быть выбраны из числа так называемых "основных" собственных функций данной системы. В состоянии, описываемом такой функцией, число $a_r$ частиц, находящихся в состоянии с уровнем энергии $\varepsilon_r$, получает вполне определённое значение, так что каждой "основной" собственной функции, являющейся членом нашего базиса, соответствует однозначно определённый набор "чисел заполнения" $a_r$ ($r=1, 2, \ldots$); числа $a_r$ при этом должны удовлетворять соотношениям

$$\sum_{r=1}^{\infty} a_r = N, \quad \sum_{r=1}^{\infty} a_r \varepsilon_r = E.$$  \hspace{1cm} (K)

Обратно, каждому определённому набору чисел заполнения, удовлетворяющему уравнениям (K), соответствует определённое число членов линейного базиса. Это число различно для различных статистических схем. Как мы видели (§ 5 гл. III), оно может быть представлено в виде

$$C(N) \prod_{r=1}^{\infty} \gamma(a_r),$$  \hspace{1cm} (1)

где $C(N) = N!$ в случае полной статистики и $C(N) = 1$ в случае двух других статистик и где

$$\gamma(a) = \frac{1}{a!} \quad \text{(полная статистика)},$$

$$\gamma(a) = 1 \quad \text{(симметрическая статистика)},$$

$$\gamma(a) = \begin{cases} 
1 & (a \leq 1) \\
0 & (a > 1)
\end{cases} \quad \text{(антисимметрическая статистика)}.$$

Это даёт для всех трёх случаев исходные выражения структурной функции $\Omega(N, E)$. Мы находим в случае полной статистики

$$\Omega(N, E) = \sum_{(K)} \frac{N!}{\prod_{r=0}^{\infty} a_r!},$$  \hspace{1cm} (2)

в случае симметрической статистики

$$\Omega(N, E) = \sum_{(K)} 1$$  \hspace{1cm} (3)
и в случае антисимметрической статистики

$$\Omega (N, E) = \sum_{(K)} * 1,$$

причём в первых двух случаях суммирование производится по всем системам чисел $a_r \geqslant 0$, удовлетворяющим условиям $\text{(K)}$, а в третьем случае знак $*$ должен показать, что суммирование распространяется лишь на такие решения уравнений $\text{(K)}$, в которых $a_r \leqslant 1 \ (r = 1, 2, \ldots)$. Как мы условились в § 2 гл. III, мы будем называть микроканоническим средним любой фазовой функции (т. е. величины, принимающей определённое значение по меньшей мере в каждом из основных состояний системы) среднее арифметическое значений, принимаемых ею в тех $\Omega (N, E)$ состояниях, которые описываются собственными функциями выбранного нами линейного базиса. Основания для такого выбора метода осреднения могут быть даны лишь позднее, когда будет рассмотрен вопрос о представительности микроканонических средних.

Таковы известные нам отправные пункты дальнейшего исследования, к которому мы теперь переходим.

§ 2. Средние значения чисел заполнения

Мы уже неоднократно имели случай заметить, что первоочередной задачей статистической теории является определение средних значений чисел заполнения $a_r$, а также их попарных произведений $a_r a_s$. В случае фотонов мы в § 2 предыдущей главы выразили с этой целью средние значения чисел $a_r$ и $a_r a_s$ через различные значения структурной функции системы. Этот первый шаг мы должны провести теперь для случая материальных частиц, и притом для всех трёх основных статистических схем.

Пусть снова $a_r (U)$ представляет собою значение числа $a_r$ в состоянии системы, описываемом основной собственной функцией $U$; тогда

$$\bar{a}_r = \frac{1}{\Omega (N, E)} \sum_U a_r (U),$$

где суммирование производится по всем основным функциям выбранного нами линейного базиса.
1. Полная статистика

Каждому набору чисел заполнения \( a_r \) \((r = 1, 2, \ldots)\), удовлетворяющему условиям \((K)\) § 1, соответствует в случае полной статистики \[ \prod_{r=1}^{\infty} a_r! \] различных функций линейного базиса, и следовательно, столько же слагаемых в сумме правой части формулы (5). Поэтому в случае полной статистики

\[ a_r = \frac{1}{\Omega(N, E)} \sum_{(K)} \prod_{i=1}^{\infty} a_i! a_r, \]

где суммирование производится по всем наборам чисел \( a_i \), удовлетворяющим условиям \((K)\).

Так как в сумме, стоящей в правой части формулы, все слагаемые, в которых \( a_r = 0 \), обращаются в нуль, то условия \((K)\), имеющие вид

\[ a_i \geqslant 0 \quad (i = 1, 2, \ldots), \quad \sum_{i=1}^{\infty} a_i = N, \quad \sum_{i=1}^{\infty} a_i \varepsilon_i = E, \]

мы можем, не меняя результата, заменить условиями

\[ a_r > 0, 
\sum_{i=1}^{\infty} a_i = N, \quad \sum_{i=1}^{\infty} a_i \varepsilon_i = E, \quad (L) \]

так что

\[ a_r = \frac{N!}{\Omega(N, E)} \sum_{(L)} \prod_{i=1}^{\infty} a_i! a_r, \]

где сумма распространяется на все системы чисел \( a_i \), удовлетворяющие условиям \((L)\). Если мы теперь положим \( b_r = a_r - 1 \), \( b_i = a_i \) \((i \neq r)\), то условия \((L)\) равносильны условиям

\[ b_i \geqslant 0 \quad (i = 1, 2, \ldots), \quad \sum_{i=1}^{\infty} b_i = N - 1, \quad \sum_{i=1}^{\infty} b_i \varepsilon_i = E - \varepsilon_r, \quad (L') \]
и мы получаем:

$$a_r = \frac{N!}{\Omega(N, E)} \sum_{(L')} \frac{1}{\prod_{i=1}^{\infty} b_i} ;$$

а так как, очевидно, в силу (2)

$$\sum_{(L')} \frac{(N-1)!}{\prod_{i=1}^{\infty} b_i} = \Omega(N - 1, E - \varepsilon_r),$$

то

$$a_r = \frac{N \Omega(N - 1, E - \varepsilon_r)}{\Omega(N, E)} .$$ (6)

Средние значения произведений $a_r a_s$ находятся аналогичным путём. При $r \neq s$ мы имеем:

$$a_r a_s = \frac{N!}{\Omega(N, E)} \sum_{(K)} a_r a_s \frac{1}{\prod_{i=1}^{\infty} a_i} ;$$

среди членов суммы $\sum_{(K)}$ отличны от нуля лишь те, в которых $a_r > 0, a_s > 0$; поэтому

$$a_r a_s = \frac{N!}{\Omega(N, E)} \sum_{a_r > 0, a_s > 0} a_r a_s \frac{1}{\prod_{i=1}^{\infty} a_i},$$

или, полагая, $b_r = a_r - 1$, $b_s = a_s - 1$, $b_i = a_i (i \neq r, i \neq s)$,

$$a_r a_s = \frac{N!}{\Omega(N, E)} \sum_{(M)} \frac{1}{\prod_{i=1}^{\infty} b_i} ,$$

где (M) означает систему условий

$$b_i \geq 0 (i = 1, 2, \ldots), \sum_{i=1}^{8} b_i = N - 2, \sum_{i=1}^{\infty} b_i \varepsilon_i = E - \varepsilon_r - \varepsilon_s, (M)$$

так что в силу (2)

$$\sum_{(M)} \frac{(N-2)!}{\prod_{i=1}^{\infty} b_i} = \Omega(N - 2, E - \varepsilon_r - \varepsilon_s);$$
посому мы находим:

\[ a_r a_s = \frac{N(N - 1) \Omega(N - 2, E - \varepsilon_r - \varepsilon_s)}{\Omega(N, E)}. \] (7)

Наконец, мы находим:

\[ a_r (a_r - 1) = \frac{N!}{\Omega(N, E)} \sum_{(K)} a_r (a_r - 1) = \frac{N!}{\Omega(N, E)} \sum_{a_r \geq 2} \prod_{i=1}^{\infty} a_i! \]

откуда, полагая \( b_r = a_r - 2, \quad b_i = a_i \quad (i \neq r), \)

\[ a_r (a_r - 1) = \frac{N!}{\Omega(N, E)} \sum_{b_i = 0,}^{\infty} \frac{1}{\prod_{i=1}^{\infty} b_i!} \]

\[ \sum_{i=1}^{8} b_i = N - 2, \]

\[ \sum_{i=1}^{8} b_i \varepsilon_i = E - 2\varepsilon_r, \]

\[ = \frac{N(N - 1)}{\Omega(N, E)} \Omega(N - 2, E - 2\varepsilon_r), \] (8)

и следовательно, в силу (6)

\[ a_r^2 = a_r (a_r - 1) + a_r = \]

\[ = \frac{1}{\Omega(N, E)} \{ N(N - 1) \Omega(N - 2, E - 2\varepsilon_r) + N\Omega(N - 1, E - \varepsilon_r) \}. \] (9)

2. Симметрическая статистика

В случае симметрической статистики формула (3) § 1 показывает, что \( \Omega(N, E) \) есть число решений системы уравнений

\[ \sum_{i=1}^{\infty} a_i = N, \quad \sum_{i=1}^{\infty} a_i \varepsilon_i = E \]

в целых \( a_i \geq 0 \). Обозначим через \( \Gamma_k \) \((k = 1, 2, \ldots)\) число решений той же системы с дополнительным условием \( a_r \geq k \).
Тогда, очевидно, $\Gamma_k - \Gamma_{k+1}$ будет числом решений системы уравнений

$$a_r = k, \ a_i \geq 0 \ (i = 1, 2, \ldots), \ \sum_{i=1}^{\infty} a_i = N, \ \sum_{i=1}^{\infty} a_i \varepsilon_i = E,$$

так что

$$\bar{a}_r = \frac{1}{\Omega(N, E)} \sum_{k=1}^{\infty} k (\Gamma_k - \Gamma_{k+1}),$$

откуда преобразованием Абеля легко получаем:

$$\bar{a}_r = \frac{1}{\Omega(N, E)} \sum_{k=1}^{\infty} \Gamma_k,$$

но, с другой стороны, полагая $b_r = a_r - k$, $b_i = a_i \ (i \neq r)$, мы находим, что $\Gamma_k$ есть число решений в целых $b_i \geq 0$ системы

$$\sum_{i=1}^{\infty} b_i = N - k, \ \sum_{i=1}^{\infty} b_i \varepsilon_i = E - k \varepsilon_r,$$

t. е.

$$\Gamma_k = \Omega (N - k, E - k \varepsilon_r),$$

и следовательно,

$$\bar{a}_r = \frac{1}{\Omega(N, E)} \sum_{k=1}^{\infty} \Omega(N - k, E - k \varepsilon_r). \ \ \ (10)$$

Далее, обозначим при $r \neq s$ через $\Gamma_{kl}$ ($k = 1, 2, \ldots; \ l = 1, 2, \ldots$) число решений системы

$$\sum_{i=1}^{\infty} a_i = N, \ \sum_{i=1}^{\infty} a_i \varepsilon_i = E \ (a_i \geq 0, \ a_r \geq k, \ a_s \geq l);$$

тогда число решений системы

$$\sum_{i=1}^{\infty} a_i = N, \ \sum_{i=1}^{\infty} a_i \varepsilon_i = E \ (a_i \geq 0, \ a_r = k, \ a_s = l)$$

будет, как легко подсчитать, равно

$$\Gamma_{kl} = \Gamma_{k+1, l} - \Gamma_{k, l+1} + \Gamma_{k+1, l+1}.$$
так что мы находим:

\[ \overline{a_r a_s} = \frac{1}{\Omega(N, E)} \sum_{k, l = 1}^{\infty} k l \{ \Gamma_{kl} - \Gamma_{k+1, l} - \Gamma_{k, l+1} + \Gamma_{k+1, l+1} \}, \]

или, применяя двукратное преобразование Абеля,

\[ \overline{a_r a_s} = \frac{1}{\Omega(N, E)} \sum_{k, l = 1}^{\infty} \Gamma_{kl}. \]

Но, с другой стороны, полагая

\[ b_i = a_i \quad (i \neq r, \ i \neq s), \quad b_r = a_r - k, \quad b_s = a_s - l, \]

мы аналогично предыдущему легко находим:

\[ \Gamma_{kl} = \Omega(N - k - l, E - k\varepsilon_r - l\varepsilon_s), \]

вследствие чего

\[ \overline{a_r a_s} = \frac{1}{\Omega(N, E)} \sum_{k, l = 1}^{\infty} \Omega(N - k - l, E - k\varepsilon_r - l\varepsilon_s). \quad (11) \]

Наконец, вполне аналогичный расчёт даёт нам:

\[ \overline{a_r^2} = \frac{1}{\Omega(N, E)} \sum_{k = 1}^{\infty} k^2 \left( \Gamma_k - \Gamma_{k+1} \right) = \]

\[ = \frac{1}{\Omega(N, E)} \sum_{k = 1}^{\infty} (2k - 1) \Gamma_k, \]

откуда

\[ \overline{a_r^2} = \frac{1}{\Omega(N, E)} \sum_{k = 1}^{\infty} (2k - 1) \Omega(N - k, E - k\varepsilon_r). \quad (12) \]

3. Антисимметрическая статистика

В этом случае число \( a_r \) может принимать только значения 0 и 1. Если мы обозначим через \( \Omega_r(N, E) \) число тех решений системы

\[ \sum_{i=1}^{\infty} a_i = N, \quad \sum_{i=1}^{\infty} a_i \varepsilon_i = E, \quad 0 \leq a_i \leq 1, \quad (N) \]
в которых \( a_r = 1 \), то, очевидно, поэтому

\[
\bar{a}_r = \frac{\Omega_r(N, E)}{\Omega(N, E)};
\]

обозначая же через \( \Omega_{r'}(N, E) \) число тех решений системы (N), в которых \( a_r = 0 \), мы имеем, с одной стороны,

\[
\Omega_r(N, E) + \Omega_{r'}(N, E) = \Omega(N, E);
\]

с другой же стороны, полагая \( b_i = a_i \) (\( i \neq r \)), \( b_r = a_r - 1 \), мы видим, что \( \Omega_r(N, E) \) равно числу таких решений системы

\[
\sum_{i=1}^{\infty} b_i = N - 1, \quad \sum_{i=1}^{\infty} b_i \varepsilon_i = E - \varepsilon_r, \quad 0 \leq b_i \leq 1,
\]

в которых \( b_r = 0 \), т. е. равно \( \Omega_{r'}(N - 1, E - \varepsilon_r) \), вследствие чего

\[
\Omega_r(N, E) = \Omega(N - 1, E - \varepsilon_r) - \Omega_r(N - 1, E - \varepsilon_r).
\]

Последовательное применение этой рекуррентной формулы даёт нам:

\[
\Omega_r(N, E) = \sum_{k=1}^{\infty} (-1)^{k-1} \Omega(N - k, E - k\varepsilon_r),
\]

и следовательно,

\[
\bar{a}_r = \frac{1}{\Omega(N, E)} \sum_{k=1}^{\infty} (-1)^{k-1} \Omega(N - k, E - k\varepsilon_r). \quad (13)
\]

Пусть теперь \( r \neq s \) и 1) \( \Omega_{rs}(N, E) \), 2) \( \Omega_{rs'}(N, E) \), 3) \( \Omega_{r's}(N, E) \), и 4) \( \Omega_{r's'}(N, E) \) означают числа таких решений системы (N), в которых соответственно 1) \( a_r = a_s = 1 \), 2) \( a_r = 1, a_s = 0 \), 3) \( a_r = 0, a_s = 1 \), 4) \( a_r = a_s = 0 \). Тогда, прежде всего, очевидно,

\[
\Omega(N, E) = \Omega_{rs}(N, E) + \Omega_{rs'}(N, E) + \Omega_{r's}(N, E) + \Omega_{r's'}(N, E)
\]

и

\[
\bar{a}_r a_s = \frac{\Omega_{rs}(N, E)}{\Omega(N, E)}.
\]
С другой стороны, привычный уже нам приём перехода от чисел \( a_i \) к числам \( b_i \) легко даёт:

\[
\Omega_{rs}(N, E) = \Omega_{r's'}(N - 1, E - \varepsilon_r - \varepsilon_s) = \\
= \sum_{k, l=1}^{\infty} (-1)^{k+l} \Omega(N - k - l, E - k\varepsilon_r - l\varepsilon_s),
\]

а рекуррентным применением этой формулы мы находим:

\[
\Omega_{rs}(N, E) = \sum_{k, l=1}^{\infty} (-1)^{k+l} \Omega(N - k - l, E - k\varepsilon_r - l\varepsilon_s),
\]

и следовательно, \( r \neq s \)

\[
a_r a_s = \frac{1}{\Omega(N, E)} \sum_{k, l=1}^{\infty} (-1)^{k+l} \Omega(N - k - l, E - k\varepsilon_r - l\varepsilon_s). \tag{14}
\]

Наконец, в случае антисимметричной статистики всегда \( a_r^2 = a_r \), так что в силу (13)

\[
a_r = \frac{1}{\Omega(N, E)} \sum_{k=1}^{\infty} (-1)^{k-1} \Omega(N - k, E - k\varepsilon_r). \tag{15}
\]

Окidyвая взором полученную нами систему формул (6) – (15), мы видим теперь, что во всех трёх статистических схемах средние значения \( a_r, a_r a_s \) \( (r \neq s) \) и \( a_r^2 \) весьма просто выражаются через отношения вида

\[
\frac{\Omega(N - u, E - v)}{\Omega(N, E)}, \tag{16}
\]

где \( u \) и \( v \) — некоторые положительные числа. Если поэтому мы получим достаточно простые и точные оценки отношений вида (16), то тем самым мы будем иметь нужные нам асимптотические оценки для средних значений чисел заполнения, их квадратов и попарных произведений. Этим путём мы и пойдём в дальнейшем.

Здесь же мы выведем ещё одну простую формулу, которая пригодится нам позднее. Мы знаем, что в ряду уровней энергии \( \varepsilon_r \) частиц, составляющих данную систему, могут встречаться равные между собой числа, т. е. что возможно \( \varepsilon_r = \varepsilon_s \),
при \( r \neq s \). Условимся в этом случае полагать для краткости \( a_r a_s = \tau_r \). Выражение величина \( \tau_r \) мы получаем для наших трёх статистических схем, полагая \( \varepsilon_s = \varepsilon_r \) соответственно в формулах (7), (11) и (14).

В случае полной статистики формулы (7) и (8) легко дают нам:

\[
\tau_r = a_r (a_r - 1) = a_r^2 - a_r. \tag{17}
\]

В случае симметрической статистики мы находим в силу формулу (10), (12) и (11):

\[
\overline{a_r^2} - a_r = \frac{2}{\Omega (N, E)} \sum_{m=1}^{\infty} (m - 1) \Omega (N - m, E - m \varepsilon_r) =
\]

\[
= \frac{2}{\Omega (N, E)} \sum_{k, l=1}^{\infty} \Omega (N - k - l, E - k \varepsilon_r - l \varepsilon_r) =
\]

\[
= 2\tau_r. \tag{18}
\]

Наконец, в случае антисимметрической статистики \( a_r^2 - a_r \), и потому

\[
\overline{a_r^2} - a_r = 0. \tag{19}
\]

С помощью введённого нами в § 3 гл. III «индекса симметрии» системы \( \sigma \) формулы (17), (18) и (19) можно объединить в одну формулу

\[
\overline{a_r^2} - a_r = (\sigma + 1) \tau_r, \tag{20}
\]

которая, таким образом, имеет место для всех трёх основных статистических схем.

§ 3. Редукция к задаче теории вероятностей

Теперь мы должны провести рассуждение, имеющее целью свести асимптотическую оценку функции \( \Omega (N, E) \) к некоторой предельной задаче теории вероятностей. Путь, которым это можно сделать, во всём основном указывает нам § 3 предыдущей главы. Мы должны, однако, учесть два существенных момента, отличающих нашу новую задачу от той,
которую мы решили в только что указанном месте: 1) в случае фотонов мы имели дело только с симметрической статистикой, теперь же мы должны охватить все три основные схемы; 2) в случае фотонов структурная функция $\Omega (E)$ зависит только от одной переменной $E$, теперь же $\Omega (N, E)$ в любой из трёх основных схем зависит от двух переменных $N$ и $E$; последнее обстоятельство позволяет предвидеть, что для решения нашей новой задачи нам придётся опираться на двумерные предельные теоремы теории вероятностей.

Будем, как и в предыдущей главе, обозначать через $g_k$ число уровней энергии отдельной частицы, заключённых между $k$ и $k + 1$, в предположении, что занимаемый системой объём равен единице. Тогда и здесь, как там, это число равно $Vg_k$ в случае, когда система занимает объём $V$, так что структурная функция $\Omega (N, E)$ попрежнему существенным образом зависит от $V$. В дальнейшем при выводе наших асимптотических формул мы будем предполагать, что числа $N, E$ и $V$ безгранично возрастают, оставаясь между собой в постоянных отношениях. При этом в полной силе остаётся формула (12) § 3 предыдущей главы, так что и здесь сумма всякого абсолютно сходящегося ряда вида

$$\sum_{i=1}^{\infty} f (e_i)$$

есть величина, пропорциональная $V$ (а значит, пропорциональная и каждому из чисел $N, E$).

Введём для каждого натурального числа $k$ двумерный закон распределения $p_k (x, y)$, определяемый следующим образом:

1) случайная величина $x$ может принимать только целые неотрицательные значения;

2) если $x = n$ $(n = 0, 1, \ldots)$, то обязательно $y = nk$;

3) вероятность того, что $x = n$, $y = nk$, равна

$$P (x = n, y = nk) = \frac{\gamma (n) e^{-(\alpha + \beta k) n}}{\sum_{i=0}^{\infty} \gamma (i) e^{-(\alpha + \beta k) i}} (n = 0, 1, \ldots), \quad (21)$$

где $\alpha$ и $\beta$ — параметры, значения которых мы распорядимся ниже, а $\gamma (n)$ есть функция целочисленной неотрицательной
переменной \( n \), введённая нами в гл. III и определяемая для различных видов статистики следующим образом:

1) для полной статистики
\[
\gamma(n) = \frac{1}{n!},
\]

2) для симметрической статистики
\[
\gamma(n) = 1,
\]

3) для антисимметрической статистики
\[
\gamma(n) = \begin{cases} 
1 & \text{при } n = 0 \text{ и } n = 1, \\
0 & \text{при } n > 1.
\end{cases}
\]

\( (I) \)

Разумеется, значения параметров \( \alpha \) и \( \beta \) должны быть выбраны так, чтобы ряд в знаменателе правой части формулы (21) сходился при любом \( k \gg 1 \).

Определённый таким образом закон распределения \( p_k(x, y) \), очевидно, является вырожденным: все пары возможных значений подчинённых этому закону величин \( (x, y) \) расположены на одной прямой \( y = kx \). Это обстоятельство, однако, не будет во всём дальнейшем играть никакой роли.

Рассмотрим теперь бесконечную последовательность \( (x_i, y_i) \) \( (i = 1, 2, \ldots) \) пар случайных величин, среди которых имеется

- \( g_1 \) пар, подчинённых закону \( p_1(x, y) \),
- \( g_2 \) пар, подчинённых закону \( p_2(x, y) \)

и вообще

\( g_k \) пар, подчинённых закону \( p_k(x, y) \) \( (k = 1, 2, \ldots) \);

при этом условимся пары с различными индексами считать взаимно независимыми.

Вероятность того, что пара случайных величин \( (x, y) \), распределённая по закону \( p_k(x, y) \), получит отличные от нуля значения, равна в силу (21)

\[
\frac{\sum_{n=1}^{\infty} \gamma(n) e^{-(\alpha + \beta k)n}}{\sum_{n=0}^{\infty} \gamma(n) e^{-(\alpha + \beta k)n}} \cdot \frac{\sum_{n=0}^{\infty} \gamma(n+1) e^{-(\alpha + \beta k)n}}{\sum_{n=0}^{\infty} \gamma(n) e^{-(\alpha + \beta k)n}} \leq e^{-\alpha - \beta k},
\]

tак как для всех трёх видов статистики \( \gamma(n+1) \leq \gamma(n) \) \( (n = 0, 1, 2, \ldots) \). Поэтому вероятность того, что по меньшей
мере одна из \( g_{k} \) пар нашей последовательности, распределённых по закону \( p_{k}(x, y) \), получит отличные от нуля значения, меньше, чем

\[
g_{k}e^{-\alpha-\beta_{k}}.
\]

В случае материальных частиц квантовая теория даёт, как мы увидим ниже, иные значения для чисел \( g_{k} \), чем в случае фотонов; однако и здесь ряд

\[
\sum_{k=1}^{\infty} g_{k}e^{-\beta_{k}}
\]

всегда сходится при любом \( \beta > 0 \). Как и в случае фотонов, мы отсюда легко заключаем, что с вероятностью 1 сходятся оба ряда

\[
\sum_{i=1}^{\infty} x_{i} = X, \quad \sum_{i=1}^{\infty} y_{i} = Y.
\]

Закон распределения пары случайных величин \((X, Y)\), очевидно, уже не являющийся выраженным, мы будем обозначать через \( P(X, Y) \); очевидно, что вид этого закона зависит, кроме природы составляющих данную систему частиц, только от параметров \( \alpha \) и \( \beta \). Этот закон \( P(X, Y) \) и будет играть основную роль во всём дальнейшем.

Занимаемый нашей системой объём \( V \) мы будем, как и в случае фотонов, считать целым числом. Рассмотрим \( V \) пар случайных величин \((X_{i}, Y_{i})\) \((i = 1, 2, \ldots, V)\), распределённых по одному и тому же закону \( P(X, Y) \) и взаимно независимых. Положим

\[
\sum_{i=1}^{V} X_{i} = S_{V}, \quad \sum_{i=1}^{V} Y_{i} = T_{V},
\]

\[
C(p) = p! \text{ для полной статистики и }
\]

\[
C(p) = 1 \text{ для двух других статистик;}
\]

dопустим, что ряд

\[
\sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{e^{-\alpha p - \beta q}}{C(p)} \Omega(p, q) = \Phi(\alpha, \beta)
\]
сходимся при выбранных нами значениях параметров \( \alpha \) и \( \beta \). Тогда имеет место

Теорема. Для любой пары неотрицательных целых чисел \( p, q \)

\[
\Omega (p, q) = C(p) \Phi (\alpha, \beta) e^{\alpha p + \beta q} P(S_V = p, T_V = q),
\]

где последний множитель правой части означает вероятность одновременного выполнения равенств \( S_V = p, T_V = q \).

Доказательство. Положим для краткости

\[
\left\{ \sum_{n=0}^{\infty} y(n) e^{-nz} \right\}^{-1} = \Gamma(z);
\]

тогда закон \( p_k(x, y) \) может быть записан в более кратком виде

\[
P(x = n, y = nk) = \left( \alpha + \beta k \right) e^{-n(\alpha + \beta k)} \quad (n = 0, 1, \ldots).
\]

Пара случайных величин \((S_V, T_V)\) определена нами как сумма \( V \) взаимно независимых пар \((X_i, Y_j)\) \((i = 1, 2, \ldots, V)\), каждая из которых распределена по закону \( P(X, Y) \). Этот закон \( P(X, Y) \) есть в свою очередь закон распределения суммы бесконечного ряда взаимно независимых случайных пар, среди которых имеется \( g_k \) пар, распределенных по закону \( p_k(x, y) \) \((k = 1, 2, \ldots)\). Это показывает, что пара \((S_V, T_V)\) может рассматриваться как сумма бесконечного ряда взаимно независимых пар, среди которых имеется \( V g_k \) пар, распределенных по закону \( p_k(x, y) \) \((k = 1, 2, \ldots)\); эти \( V g_k \) случайных пар мы обозначим через \((x_{kl}, y_{kl})\) \((l = 1, 2, \ldots, V g_k)\), так что

\[
S_V = \sum_{k=1}^{\infty} \sum_{l=1}^{V g_k} x_{kl}, \quad T_V = \sum_{k=1}^{\infty} \sum_{l=1}^{V g_k} y_{kl},
\]

и пара \((x_{kl}, y_{kl})\) распределена по закону \( p_k(x, y) \).

Для того чтобы величины \( S_V, T_V \) получили соответственно значения \( p \) и \( q \), нужно, чтобы значения \( x_{kl} = n_{kl} \), \( y_{kl} = kn_{kl} \), получаемые случайными величинами \( x_{kl}, y_{kl} \), удовлетворяли соотношениям

\[
\sum_{k=1}^{\infty} \sum_{l=1}^{V g_k} n_{kl} = p, \quad \sum_{k=1}^{\infty} \sum_{l=1}^{V g_k} k n_{kl} = q.
\]
Пусть \( n_{kl} (1 \leq k < \infty, 1 \leq l \leq Vg_k) \) есть некоторая определённая система целых неотрицательных чисел, удовлетворяющих уравнениям (P); тогда вероятность того, что мы будем иметь \( x_{kl} = n_{kl}, y_{kl} = k n_{kl} (1 \leq k < \infty, 1 \leq l \leq Vg_k) \) в силу взаимной независимости пар \((x_{kl}, y_{kl})\), будет равна:

\[
\prod_{k=1}^{\infty} \prod_{l=1}^{Vg_k} P(x_{kl} = n_{kl}, y_{kl} = kn_{kl}) =
\]

\[
= \prod_{k=1}^{\infty} \prod_{l=1}^{Vg_k} \Gamma(a + \beta k) \gamma(n_{kl}) e^{-n_{kl}(a+\beta k)} =
\]

\[
= \left\{ \prod_{k=1}^{\infty} \left[ \Gamma(a + \beta k) \right]^{Vg_k} \right\} e^{-a \sum_{k=1}^{\infty} \sum_{l=1}^{Vg_k} n_{kl} - \beta \sum_{k=1}^{\infty} \sum_{l=1}^{Vg_k} k n_{kl}} \times
\]

\[
\times \prod_{k=1}^{\infty} \prod_{l=1}^{Vg_k} \gamma(n_{kl}) =
\]

\[
e^{-a p - \beta q} \prod_{k=1}^{\infty} \left[ \Gamma(a + \beta k) \right]^{Vg_k} \prod_{k=1}^{\infty} \prod_{l=1}^{Vg_k} \gamma(n_{kl}). \tag{23}
\]

Такова вероятность того, что случайные величины \( x_{kl}, y_{kl} \) получают соответственно некоторые определённые значения \( n_{kl}, k n_{kl} \), удовлетворяющие системе уравнений (P). Вероятность же того, что мы будем иметь \( S_V = p, T_V = q \), равна в силу вышесказанного сумме вероятностей типа (23), распространённой на все системы целых неотрицательных чисел \( n_{kl} \), удовлетворяющие уравнениям (P), т. е.

\[
P(S_V = p, T_V = q) =
\]

\[
e^{-a p - \beta q} \prod_{k=1}^{\infty} \left[ \Gamma(a + \beta k) \right]^{Vg_k} \sum_{(P)} \prod_{k=1}^{\infty} \prod_{l=1}^{Vg_k} \gamma(n_{kl}). \tag{24}
\]

Если мы теперь вспомним, что в ряду

\[
\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_r, \ldots
\]

уровней энергии частицы натуральное число \( k \) встречается \( Vg_k \) раз, то мы заметим, что сумма

\[
\sum_{(P)} \prod_{k=1}^{\infty} \prod_{l=1}^{Vg_k} \gamma(n_{kl}),
\]
стоящая в правой части соотношения (24), лишь системой обозначений отличается от той суммы
\[ \sum_{(K)} \prod_{r=1}^{\infty} \gamma(a_r), \]
с помощью которой (см. гл. III, § 5) определяется \( \Omega(p, q) \), если условия (K) имеют вид
\[ \sum_{r=1}^{\infty} a_r = p, \quad \sum_{r=1}^{\infty} a_r \varepsilon_r = q \] (K)
(в случае полной статистики эта сумма должна быть умножена ещё на \( p! \) для получения \( \Omega(p, q) \)). Таким образом, мы получаем:
\[ C(p) \sum_{(P)} \prod_{k=1}^{\infty} \prod_{i=1}^{V_{g_k}} \gamma(n_{ki}) = \Omega(p, q), \]
где \( C(p) = p! \) для полной статистики и \( C(p) = 1 \) для двух других статистик. Отсюда формула (24) даёт:
\[ \mathcal{P}(S_V = p, T_V = q) = e^{-\alpha p - \beta q} \prod_{k=1}^{\infty} [\Gamma(a + \beta k)]^{V_{g_k}} \frac{\Omega(p, q)}{C(p)} ; \] (25)
суммируя же это соотношение по \( p \) и \( q \) от 0 до \( \infty \), мы находим:
\[ 1 = \prod_{k=1}^{\infty} [\Gamma(a + \beta k)]^{V_{g_k}} \sum_{p, q = 0}^{\infty} e^{-\alpha p - \beta q} \frac{\Omega(p, q)}{C(p)} = \]
\[ = \prod_{k=1}^{\infty} [\Gamma(a + \beta k)]^{V_{g_k}} \Phi(a, \beta). \]
Отсюда
\[ \prod_{k=1}^{\infty} [\Gamma(a + \beta k)]^{V_{g_k}} = \{\Phi(a, \beta)\}^{-1}, \] (26)
и соотношение (25) даёт:
\[ \mathcal{P}(S_V = p, T_V = q) = \frac{e^{-\alpha p - \beta q} \Omega(p, q)}{\Phi(a, \beta) C(p)}, \]
что равносильно формуле (22); таким образом, наша теорема доказана.
Формула (22) полностью сводит изучение нужных нам свойств функции \( \Omega(p, q) \) к исследованию закона распределения пары случайных величин \( (S_V, T_V) \); а так как эта пара служит суммой безгранично возрастающего числа \( V \) случайных пар, распределённых по одному и тому же постоянному закону и независимым между собой, то мы стоим перед одной из наиболее разработанных предельных задач теории вероятностей, решение которой известно с большой степенью точности. Заметим ещё, что присутствие в правой части формулы (22) множителя \( \Phi(\alpha, \beta) \) не может нас затруднить уже потому, что, как мы видели, нужные нам выражения средних значений различных фазовых функций всегда содержат лишь отношения различных значений функции \( \Omega \), вследствие чего этот множитель в них всегда полностью сокращается.

Сделаем в целях дальнейшего ещё следующее замечание. Функция \( \Phi(\alpha, \beta) \), кроме параметров \( \alpha \) и \( \beta \), зависит ещё от вида функции \( \Omega(p, q) \), которая в свою очередь, как мы видели, зависит от занимаемого системой объёма \( V \); таким образом, функция \( \Phi(\alpha, \beta) \) также зависит от \( V \). Вид этой зависимости очень прост, как показывает соотношение (26): обозначая через \( \Phi_1(\alpha, \beta) \) выражение функции \( \Phi(\alpha, \beta) \) при \( V = 1 \), мы находим:

\[
\Phi(\alpha, \beta) = \{\Phi_1(\alpha, \beta)\}^V, \quad \ln \Phi(\alpha, \beta) = V \ln \Phi_1(\alpha, \beta); \quad (27)
\]

таким образом, функция \( \ln \Phi(\alpha, \beta) \) просто пропорциональна занимаемому данной системой объёму \( V \). Легко видеть, что зависимость структурной функции \( \Omega(p, q) \) от этого объёма имеет значительно более сложный характер.

§ 4. Выбор значений параметров \( \alpha \) и \( \beta \)

Значения введенных нами параметров \( \alpha \) и \( \beta \) до сих пор были подчинены лишь общему требованию сходимости тех рядов, с которыми нам приходилось иметь дело; в остальном эти значения оставались произвольными. Теперь, прежде чем приступить к применению предельных теорем теории вероятностей, нам будет целесообразно выбрать эти значения так, чтобы по возможности упростить последующие расчёты. Этому выбору и будет посвящён настоящий параграф.
В § 3 мы для всех трёх статистических схем определили функцию

\[ \Phi (\alpha, \beta) = \sum_{p, q=0}^{\infty} \frac{e^{-\alpha p - \beta q}}{C(p)} \Omega (p, q). \]

Далее мы показали (формула (26)), что

\[ \Phi (\alpha, \beta) = \prod_{k=1}^{\infty} \left[ \Gamma (\alpha + \beta k) \right]^{-1} v g_k, \]

или, что то же,

\[ \Phi (\alpha, \beta) = \prod_{r=1}^{\infty} \left[ \Gamma (\alpha + \beta \varepsilon_r) \right]^{-1} = \prod_{r=1}^{\infty} \left\{ \sum_{n=0}^{\infty} \gamma (n) e^{-n(\alpha + \beta \varepsilon_r)} \right\}, \]

где функция \( \gamma (n) \) для каждой из трёх статистических схем определяется согласно правилам (I) § 3.

Обозначим теперь через \( E_0 \) наименьшее возможное значение энергии системы, состоящей из \( N \) частиц данного типа. Очевидно, что в случае полной или симметрической статистики \( E_0 = N \varepsilon_1 \); в случае же антисимметрической статистики, где в состоянии, характеризуемом уровнем \( \varepsilon_r \), может находиться не более одной частицы, мы, очевидно, имеем \( E_0 = \sum_{r=1}^{N} \varepsilon_r \).

Докажем теперь следующее общее предложение:

**Теорема.** Пусть система, состоящая из \( N \) частиц, имеет энергию \( E \geq E_0 \). Тогда система уравнений

\[ \frac{\partial \ln \Phi}{\partial \alpha} = -N, \quad \frac{\partial \ln \Phi}{\partial \beta} = -E \]

имеет едиственное решение \( (\alpha, \beta) \), причём \( \beta > 0 \).

Для доказательства рассмотрим функцию

\[ F (\alpha, \beta) = e^{\alpha N + \beta E} \Phi (\alpha, \beta) \]

и исследуем её поведение в полуплоскости \( \beta > 0 \); при этом мы последовательно установим, что

1°. При \( \beta \to 0 \) \( F (\alpha, \beta) \to +\infty \) равномерно относительно \( \alpha (-\infty < \alpha < +\infty) \).

2°. При \( \beta \to +\infty \) \( F (\alpha, \beta) \to +\infty \) равномерно относительно \( \alpha (-\infty < \alpha < +\infty) \).
3°. При \(|a| \to \infty\) \(F(a, \beta) \to -\infty\) равномерно относительно \(\beta\) (\(0 < \beta < \beta_0\)), где \(\beta_0\) — любое положительное число. Мы имеем:

\[
F(a, \beta) = e^{aN + \beta E} \prod_{r=1}^{\infty} \left\{ \sum_{n=0}^{\infty} \gamma(n) e^{-n(a + \beta \sigma_r)} \right\}.
\]

Для доказательства 1° заметим, что при любой статистике \(\alpha = 1\), и следовательно,

\[
F(a, \beta) \geq e^{aN + \beta E} \prod_{r=1}^{\infty} \left\{ 1 + e^{-a - \beta \sigma_r} \right\}.
\]

Пусть \(A\) — сколь угодно большое натуральное число, и пусть \(\beta\) настолько мало, что \(e^{-\beta \sigma_r} > \frac{1}{2} (r \leq A)\). Тогда

\[
\ln F(a, \beta) > aN + A \ln \left\{ 1 + e^{-\frac{1}{2} e^{-a}} \right\} = f(a).
\]

Функция \(f(a)\), как легко даёт элементарное исследование, имеет наименьшее значение при \(a = \ln \frac{A - N}{2N} = a_0\); поэтому

\[
\ln F(a, \beta) > f(a) \geq f(a_0) \geq A a_0 = N \ln \frac{A - N}{2N}.
\]

Так как \(A\) сколь угодно велико при достаточно малом \(\beta\), то этим утверждение 1° доказано.

Переходим к доказательству утверждений 2° и 3°. Допустим сперва, что мы имеем дело с полной или симметричной статистикой, так что \(E_0 = N e_1\) и \(\gamma(n) > 0\) для любого \(n \geq 0\). Так как в силу \(\gamma(0) = 1\) при любом \(r \geq 1\)

\[
\sum_{k=0}^{\infty} \gamma(k) e^{-k(a + \beta \sigma_r)} > 1,
\]

то в силу (28)

\[
\Phi(a, \beta) \geq \sum_{k=0}^{\infty} \gamma(k) e^{-k(a + \beta \sigma_1)},
\]

а потому и подавно при любом \(m \geq 0\)

\[
\Phi(a, \beta) \geq \gamma(m) e^{-m(a + \beta \sigma_1)},
\]
откуда

\[ F(\alpha, \beta) \geq \gamma(m) e^{\alpha(N-m) + \beta(E-m\beta_1)}. \]  

(30)

Полагая \( m = N \), мы находим:

\[ F(\alpha, \beta) \geq \gamma(N) e^{\beta(E-E_0)}. \]

Так как в силу предпосылок теоремы \( E > E_0 \), то утверждение 2° этим доказано. Далее, полагая в (30) \( m = N + 1 \) и \( m = N - 1 \), мы находим:

\[ F(\alpha, \beta) \geq \gamma(N + 1) e^{-\alpha - \beta e_1}, \]

\[ F(\alpha, \beta) \geq \gamma(N - 1) e^{\alpha + \beta e_1}; \]

первое из этих неравенств показывает, что при \( \alpha \rightarrow -\infty \), \( 0 < \beta < \beta_0 \) равномерно \( F(\alpha, \beta) \rightarrow \infty \); второе показывает, что то же имеет место и при \( \alpha \rightarrow +\infty \), \( \beta > 0 \). Таким образом, утверждение 3° также доказано.

Обратимся теперь к случаю антисимметрической статистики.

Здесь \( E_0 = \sum_{r=1}^{N} e_r \) и

\[ F(\alpha, \beta) = e^{\alpha N + \beta E} \prod_{r=1}^{\infty} (1 + e^{-\alpha - \beta e_r}) = \]

\[ = e^{\alpha N + \beta E} \prod_{r=1}^{N} e^{\alpha + \beta e_r} + 1 \prod_{r=N+1}^{\infty} (1 + e^{-\alpha - \beta e_r}). \]  

(31)

Отсюда

\[ F(\alpha, \beta) \geq e^{\alpha N + \beta E} \frac{\prod_{r=1}^{N} (1 + e^{\alpha + \beta e_r})}{e^{\alpha N + \beta E_0}} = \]

\[ = e^{\beta (E-E_0)} \prod_{r=1}^{N} (1 + e^{\alpha + \beta e_r}). \]

Это неравенство в силу \( E > E_0 \) прежде всего доказывает утверждение 2°; но отсюда же мы, очевидно, получаем и утверждение 3° для случая \( \alpha \rightarrow +\infty \). Чтобы доказать его и для случая \( \alpha \rightarrow -\infty \), достаточно заметить, что из (31) следует, очевидно, аналогично предыдущему

\[ F(\alpha, \beta) \geq e^{\alpha(E-E_0)} (1 + e^{-\alpha - \beta e_1 + 1}), \]
откуда при $a \to -\infty$ мы находим $F(a, \beta) \to \infty$ равномерно в любой области $0 < \beta < \beta_0$. Таким образом, утверждения 1°, 2° и 3° доказаны для всех трёх видов статистики.

Но из трёх доказанных утверждений, очевидно, следует, что $F(a, \beta)$ равномерно безгранично возрастает при $\beta \to 0$ и при $a^2 + \beta^2 \to \infty$. Следовательно, внутри полуплоскости $\beta > 0$ она принимает в некоторой точке $(a, \beta)$ наименьшее значение; в этой точке мы имеем:

$$\frac{\partial \ln F(a, \beta)}{\partial a} = N + \frac{\partial \ln \Phi(a, \beta)}{\partial a} = 0,$$
$$\frac{\partial \ln F(a, \beta)}{\partial \beta} = E + \frac{\partial \ln \Phi(a, \beta)}{\partial \beta} = 0,$$

откуда

$$\frac{\partial \ln \Phi}{\partial a} = -N, \quad \frac{\partial \ln \Phi}{\partial \beta} = -E.$$

Таким образом, существование решения системы уравнений (29) в полуплоскости $\beta > 0$ нами доказано; остаётся показать его единственность. С этой целью допустим, что точка $(a', \beta')$, отличная от точки $(a, \beta)$, также удовлетворяет этой системе. Тогда *)

$$\ln F(a, \beta) - \ln F(a', \beta') =$$
$$= \frac{1}{2} \left\{ \frac{\partial^2 \ln \Phi}{\partial a^2} (a - a')^2 + 2 \frac{\partial \ln \Phi}{\partial a} (a - a') (\beta - \beta') +$$
$$+ \frac{\partial^2 \ln \Phi}{\partial \beta^2} (\beta - \beta')^2 \right\},$$

где все три производные второго порядка берутся в одной и той же точке

$$[a' + \theta (a - a'), \beta' + \theta (\beta - \beta')] \quad (0 < \theta < 1).$$

Мы немедленно придём к искомому противоречию, если покажем, что квадратичная форма, стоящая в правой части последнего равенства, положительная определённая, ибо тогда

$$\ln F(a, \beta) - \ln F(a', \beta') > 0,$$

*) Производные второго порядка функции $\ln F$, очевидно, совпадают с соответствующими производными функции $\ln \Phi$. 
что невозможно по определению точки \((\alpha, \beta)\). Итак, нам остаётся доказать, что в области \(\beta > 0\) тождественно

\[
\frac{\partial^2 \ln \Phi}{\partial \alpha^2} > 0, \quad \frac{\partial^2 \ln \Phi}{\partial \alpha^2} \frac{\partial^2 \ln \Phi}{\partial \beta^2} - \left(\frac{\partial^2 \ln \Phi}{\partial \alpha \partial \beta}\right)^2 > 0.
\]

Мы имеем:

\[
\ln \Phi(\alpha, \beta) = \sum_{r=1}^{\infty} \sum_{n=0}^{\infty} \gamma(n) e^{-n(\alpha + \beta r)} = \sum_{r=1}^{\infty} S_r(\alpha, \beta),
\]
где

\[
S_r(\alpha, \beta) = \sum_{n=0}^{\infty} \gamma(n) e^{-n(\alpha + \beta r)}.
\]

Отсюда

\[
\frac{\partial^2 \ln \Phi}{\partial \alpha^2} = \sum_{r=1}^{\infty} \frac{1}{S_r^2} \left\{ S_r \frac{\partial^2 S_r}{\partial \alpha^2} - \left(\frac{\partial S_r}{\partial \alpha}\right)^2 \right\} = \sum_{r=1}^{\infty} T_r, \tag{32}
\]
где

\[
T_r = \frac{1}{S_r^2} \left\{ S_r \frac{\partial^2 S_r}{\partial \alpha^2} - \left(\frac{\partial S_r}{\partial \alpha}\right)^2 \right\}.
\]

Так как

\[
\frac{\partial S_r}{\partial \alpha} = -\sum_{n=0}^{\infty} n\gamma(n) e^{-n(\alpha + \beta r)},
\]

\[
\frac{\partial^2 S_r}{\partial \alpha^2} = \sum_{n=0}^{\infty} n^2\gamma(n) e^{-n(\alpha + \beta r)},
\]
то неравенство Буняковского легко даёт нам:

\[
T_r > 0 \quad (r = 1, 2, \ldots),
\]
и следовательно, в силу (32)

\[
\frac{\partial^2 \ln \Phi}{\partial \alpha^2} > 0.
\]

Легко видеть, далее, что если в какой-либо частной производной суммы \(S_r\), одно дифференцирование по \(\alpha\) заменить...
дифференцированием по $\beta$, то частная производная умножается на $\varepsilon_r$; поэтому из (32) следует

$$\frac{\partial^2 \ln \Phi}{\partial \alpha \partial \beta} = \sum_{r=1}^{\infty} \varepsilon_r T_r, \quad \frac{\partial^2 \ln \Phi}{\partial \beta^2} = \sum_{r=1}^{\infty} \varepsilon_r^2 T_r.$$

Отсюда мы в силу неравенства Буняковского снова легко за-ключаем, что в полуплоскости $\beta > 0$

$$\frac{\partial^2 \ln \Phi}{\partial \alpha^2} \frac{\partial^2 \ln \Phi}{\partial \beta^2} - \left( \frac{\partial^2 \ln \Phi}{\partial \alpha \partial \beta} \right)^2 > 0,$$

чем наша теорема полностью доказана.

Во всём дальнейшем мы под $\alpha$ и $\beta$ всегда будем понимать числа, удовлетворяющие системе уравнений (29); существова-

ние и единственность этой пары чисел нами только что до-

оказаны.

Сделаем ещё следующее важное замечание. В силу соот-

ношений (27) § 3 система уравнений (29), определяющая зна-

чения параметров $\alpha$ и $\beta$, равносильна системе

$$\frac{\partial \ln \Phi_1}{\partial \alpha} = -\frac{N}{V}, \quad \frac{\partial \ln \Phi_1}{\partial \beta} = -\frac{E}{V};$$

так как мы в дальнейшем будем считать, как уже сказано в
начале § 3, числа $N$, $E$ и $V$ бесконечно большими, отноше-
ния которых сохраняют постоянные значения, то правые части уравнений (34) представляют собою постоянные величины; а так как и вид функции $\Phi_1 (\alpha, \beta)$ от $N$, $V$ и $E$ не зависит, то
не зависят от них и выбраные нами значения параметров $\alpha$ и $\beta$. Таким образом, во всём дальнейшем мы можем рассмотривать числа $\alpha$ и $\beta$ как постоянные.

§ 5. Применение предельной теоремы

теории вероятностей

Теперь мы обращаемся к установлению асимптометических
ф ormул для закона распределения случайной пары ($S_1$, $T_1$), по-
строенной нами в § 3. В конце § 3 мы указали, что это —
одна из наиболее разработанных предельных задач теории вероятностей, так как мы находимся здесь в условиях примени-
мости центральной предельной теоремы (в её локальной форме), и притом в простейшем случае одинаково распределённых и взаимно независимых слагаемых.

В первую очередь нам необходимо найти математические ожидания величин $S_V$ и $T_V$. Так как

$$
\Phi (\alpha, \beta) = \sum_{p, q=0}^{\infty} \frac{e^{-\alpha p - \beta q}}{C(p)} \Omega(p, q),
$$

то в силу (22) и (29)

$$
E S_V = \sum_{p, q=0}^{\infty} p \mathbb{P}(S_V = p, T_V = q) =
\frac{1}{\Phi(\alpha, \beta)} \sum_{p, q=0}^{\infty} \rho \frac{e^{-\alpha p - \beta q}}{C(p)} \Omega(p, q) =
\frac{1}{\Phi(\alpha, \beta)} \frac{\partial \Phi(\alpha, \beta)}{\partial \alpha} = \frac{\partial \ln \Phi(\alpha, \beta)}{\partial \alpha} = N,
$$

и аналогично

$$
E T_V = \frac{\partial \ln \Phi(\alpha, \beta)}{\partial \beta} = E.
$$

Таким образом, при выбранных нами значениях параметров $\alpha$ и $\beta$ математические ожидания величин $S_V$ и $T_V$ соответственно равны $N$ и $E$, что, как мы сейчас увидим, придаёт особую простоту получаемым асимптотическим формулам.

Теперь нам надо определить центральные моменты второго порядка для величин $S_V$ и $T_V$. Мы находим:

$$
B_{11} = E \{ (S_V - ES_V)^2 \} = E \{ S_V^2 \} - \left( \frac{1}{\Phi} \frac{\partial \Phi}{\partial \alpha} \right)^2 =
\sum_{p, q=0}^{\infty} p^2 \mathbb{P}(S_V = p, T_V = q) - \left( \frac{1}{\Phi} \frac{\partial \Phi}{\partial \alpha} \right)^2 =
\frac{1}{\Phi} \sum_{p, q=0}^{\infty} p^2 \frac{e^{-\alpha p - \beta q}}{C(p)} \Omega(p, q) - \left( \frac{1}{\Phi} \frac{\partial \Phi}{\partial \alpha} \right)^2 =
\frac{1}{\Phi} \frac{\partial^2 \Phi}{\partial \alpha^2} - \left( \frac{1}{\Phi} \frac{\partial \Phi}{\partial \alpha} \right)^2 = \frac{\partial^2 \ln \Phi(\alpha, \beta)}{\partial \alpha^2},
$$
и следовательно, в силу (27) (§ 3)

\[ B_{11} = V \frac{\partial^2 \ln \Phi_1 (a, \beta)}{\partial a^2} = Vb_{11}, \]

где \( b_{11} = \frac{\partial^2 \ln \Phi_1}{\partial a^2} \) — постоянная величина.

Совершенно analogичным образом мы находим:

\[ B_{12} = E \{ S_V T_V - E S_V \overline{E T_V} \} = \frac{\partial^2 \ln \Phi (a, \beta)}{\partial a \partial \beta} = Vb_{12}, \]

\[ B_{22} = E \{ |T_V - E T_V| \}^2 = \frac{\partial^2 \ln \Phi (a, \beta)}{\partial \beta^2} = Vb_{22}, \]

где \( b_{12} \) и \( b_{22} \) — постоянные величины. В силу (33) § 3 при этом

\[ \Delta = B_{11}B_{22} - B_{12}^2 = V^2 (b_{11}b_{22} - b_{12}^2) = V^2 \delta > 0. \]

Для оценки вероятности \( P (S_V = N + u_1, \ T_V = E + u_2) \) мы применени теорему 2 гл. 1. Векторы \( (X_i, Y_i) \), из которых слагается вектор \( (S_V, T_V) \), — невырожденные, распределённые по одному и тому же закону \( P (X, Y) \), взаимно независимые целочисленные случайные векторы (§ 3). В силу § 3 гл. 1 вектор \( (X_i, Y_i) \) имеет поэтому максимальную решетку \( a_0 \pm k \alpha \pm \lambda \beta, b_0 \pm k \gamma \pm \lambda \delta, \ |a\delta - \beta \gamma| = d > 0 \). Возможные значения вектора \( (S_V, T_V) \) принадлежат тогда решётке \( Va_0 + k \alpha \pm \lambda \beta, Vb_0 + k \gamma \pm \lambda \delta \). В силу формулы (22) § 3 эти возможные значения являются, одновременно, теми значениями аргументов \( N \) и \( E \) структурной функции \( \Omega (N, E) \), для которых она отлична от нуля; иначе говоря, этой решёткой охватываются все физически возможные пары значений числа частиц и полной энергии системы. Поэтому мы должны положить

\[ N + u_1 = Va_0 + k \alpha + \lambda \beta, \ E + u_2 = Vb_0 + k \gamma + \lambda \delta. \]

Так как \( N \) и \( E \) являются соответственно математическими ожиданиями величин \( S_V \) и \( T_V \), то это означает, что в формулах (38) и (39) теоремы 2 гл. 1 \( u_1 \) и \( u_2 \) имеют то же значение, что и у нас. Кроме того, мы должны писать, очевидно,
Вместо \( n \) и \( \delta = b_{11}b_{22} - b_{12}^2 \) вместо \( \Delta \). Таким образом, формула (39) гл. 1 даёт:

\[
P(S_V = N + u_1, \ T_V = E + u_2) = \frac{d}{2\pi V V' \delta} e^{-\frac{1}{2\delta V} \left\{ b_{ii}u_i^2 - 2b_{is}u_i u_s + b_{ss}u_s^2 \right\}} + O\left(\frac{1 + u}{V^2}\right),
\]

где положено \( u = |u_1| + |u_2| \). Формула (38) гл. 1 даёт подобным же образом более точную оценку

\[
P(S_V = N + u_1, \ T_V = E + u_2) = \frac{d}{2\pi V V' \delta} e^{-\frac{1}{2\delta V} \left\{ b_{ii}u_i^2 - 2b_{is}u_i u_s + b_{ss}u_s^2 \right\}} + \frac{m_0 + m_1 u_1 + m_2 u_2}{V^2} + O\left(\frac{V V' + u^3}{V^3}\right),
\]

где \( m_0, m_1, m_2 \) — постоянные, не зависящие ни от \( V \), ни от \( u_1 \) и \( u_2 \).

Если учесть, что

\[
e^{-\frac{1}{2\delta V} \left\{ b_{ii}u_i^2 - 2b_{is}u_i u_s + b_{ss}u_s^2 \right\}} = 1 + O\left(\frac{u^2}{V}\right),
\]

то соотношение (35) даёт:

\[
P(S_V = N + u_1, \ T_V = E + u_2) = \frac{d}{2\pi V V' \delta} + O\left(\frac{1 + u^2}{V^2}\right).
\]

Подобным же образом, если учесть более точную оценку

\[
e^{-\frac{1}{2\delta V} \left\{ b_{ii}u_i^2 - 2b_{is}u_i u_s + b_{ss}u_s^2 \right\}} = 1 - \frac{1}{2\delta V} \left\{ b_{11}u_1^2 - 2b_{12}u_1 u_2 + b_{22}u_2^2 \right\} + O\left(\frac{u^4}{V^2}\right),
\]

то соотношение (36) даёт:

\[
P(S_V = N + u_1, \ T_V = E + u_2) = \frac{d}{2\pi V V' \delta} + \frac{1}{V^2} \left\{ m_0 + m_1 u_1 + m_2 u_2 \right\} - \frac{1}{4\pi \delta^3/2} \left( b_{11}u_1^2 - 2b_{12}u_1 u_2 + b_{22}u_2^2 \right) + O\left(\frac{V V' + u^4}{V^3}\right).
\]
Формулы (37) и (38) и будут нами применяться во всём дальнейшем. Укажем ещё раз, что справедливость этих формул требует такого выбора значений \( u_1 \) и \( u_2 \), чтобы точка \((N + u_1, E + u_2)\) принадлежала решётке \( V_0 + k\alpha + l\beta, V_{0'} + k\gamma + l\delta\); это условие, в частности, всегда выполнено, если \( \Omega(N + u_1, E + u_2) > 0 \), т. е. если система, состоящая из \( N + u_1 \) частиц, может иметь полную энергию \( E + u_2 \).

§ 6. Средние значения сумматорных величин

Для целей настоящего параграфа мы можем ограничиться грубой оценкой, даваемой очень простой формулой (37); более точная формула (38) нам понадобится лишь позднее.

Прежде всего, теорема § 3 (формула (22)) в соединении с формулой (37) даёт нам:

\[
\Omega(N + u_1, E + u_2) = C(N + u_1) \Phi(\alpha, \beta) e^{\alpha N + \beta E} e^{\alpha u_1 + \beta u_2} \left\{ \frac{d}{2\pi V \sqrt{\delta}} + O\left(\frac{1}{V^2}\right) \right\},
\]

откуда, в частности, при \( u_1 = u_2 = 0 \)

\[
\Omega(N, E) = C(N) \Phi(\alpha, \beta) e^{\alpha N + \beta E} \left\{ \frac{d}{2\pi V \sqrt{\delta}} + O\left(\frac{1}{V^2}\right) \right\},
\]

и следовательно,

\[
\frac{\Omega(N + u_1, E + u_2)}{\Omega(N, E)} = \frac{C(N + u_1)}{C(N)} e^{\alpha u_1 + \beta u_2} \left\{ 1 + O\left(\frac{1}{V}\right) \right\}.
\]  

Имея, таким образом, асимптотическую оценку отношений вида \( \frac{\Omega(N + u_1, E + u_2)}{\Omega(N, E)} \), мы можем на основании формул § 2 легко найти очень простые приближённые выражения для средних значений чисел заполнения.

В случае полной статистики формула (6) § 2 в соединении с формулой (39) в силу \( C(p) = p! \) даёт:

\[
\bar{a}_r = e^{-(\alpha + \beta \epsilon_r)} \left\{ 1 + O\left(\frac{1 + \epsilon_r^2}{V}\right) \right\}.
\]
В случае симметрической статистики мы имеем $C(p) = 1$, и формула (10) § 2 в соединении с формулой (39) даёт:

$$\bar{a}_r = \sum_{k=1}^{\infty} e^{-k(a + \beta \varepsilon_r)} \left\{ 1 + O\left(\frac{k^2 \varepsilon_r^2}{V}\right) \right\} =$$

$$= \frac{1}{e^{a - \beta \varepsilon_r - 1}} \left\{ 1 + O\left(\frac{1}{V} \varepsilon_r^2 \sum_{k=1}^{\infty} k^2 e^{-k(a + \beta \varepsilon_r)}\right) \right\}; \quad (41)$$

в случае антисимметрической статистики также $C(p) = 1$, и формула (13) § 2 в соединении с формулой (39) даёт:

$$\bar{a}_r = \frac{1}{e^{a - \beta \varepsilon_r + 1}} \left\{ 1 + O\left(\frac{1}{V} \varepsilon_r^2 \sum_{k=1}^{\infty} k^2 e^{-k(a + \beta \varepsilon_r)}\right) \right\}. \quad (42)$$

Для любого закреплённого значения $r$ формулы (40), (41) и (42) могут быть, очевидно, объединены в одну очень простую формулу

$$\bar{a}_r = \frac{1}{e^{a + \beta \varepsilon_r - \sigma}} + O\left(\frac{1}{V}\right), \quad (43)$$

где $\sigma$ — индекс симметрии данной системы.

Во всех этих выводах мы опирались на формулу (39), которая в свою очередь основывалась на формуле (37) § 5. Мы должны поэтому ещё убедиться, что если $(N, E)$ представляет собой возможную комбинацию числа частиц и полной энергии системы и, следовательно, точка $(N, E)$ принадлежит решётке $Va_0 + k\alpha + l\beta$, $Vh_0 + k\gamma + i\delta$, то той же решётке принадлежит и точка $(N - k, E - k\varepsilon_r)$ при любом целом $k > 0$. Но это очевидно, так как $(1, \varepsilon_r)$ есть возможная комбинация числа частиц и полной энергии системы, и следовательно, точки $(N, E)$ и $(1, \varepsilon_r)$, а значит, и точка $(N - k, E - k\varepsilon_r)$ при любом целом $k$ принадлежат упомянутой решётке.

Пусть теперь $\mathfrak{A} — связанная с данной системой сумматорная величина, т. е.

$$\mathfrak{A} = \sum_{i=1}^{N} a_i.$$
где величина $a_i$ зависит только от гамильтоновых переменных $i$-й частицы. Если мы допустим, что математическое ожидание $E_i a_i = \lambda_i$ величины $a_i$ в том состоянии, которое соответствует уровню энергии $\varepsilon$, одно и то же для всех $i$, и зависит поэтому только от $r$, то для среднего значения величины $E_A$ мы согласно § 5 гл. III имеем выражение

$$E_A = \sum_{r=1}^{\infty} \lambda_r a_r.$$  

Формулы (40), (41) и (42) дают нам поэтому:

$$E_A = \sum_{r=1}^{\infty} \frac{\lambda_r}{\alpha + \beta_{sr} - \sigma} + O\left(\frac{1}{V} \sum_{r=1}^{\infty} \varepsilon_r^2 |\lambda_r| \sum_{k=1}^{\infty} k^2 e^{-k(\alpha + \beta_{sr})}\right).$$

Но в силу замечания, сделанного нами в начале § 3, всякая абсолютно сходящаяся сумма вида

$$\sum_{r=1}^{\infty} f(\varepsilon_r)$$

представляет собою величину, пропорциональную $V$; поэтому, если только все ряды в правой части последнего равенства абсолютно сходятся (что требует не слишком быстрого роста величин $|\lambda_r|$ и всегда имеет место в реальных задачах), то мы получаем:

$$E_A = \sum_{r=1}^{\infty} \frac{\lambda_r}{\alpha + \beta_{sr} - \sigma} + O(1).$$

В наиболее часто встречающемся случае, когда $\lambda_r$ имеет одно и то же значение для всех состояний с одним и тем же уровнем энергии $\varepsilon_r$, т. е. когда $\lambda_r$ есть однозначная функция $f(\varepsilon_r)$ от $\varepsilon_r$, мы можем, очевидно, записать последнюю формулу с помощью введённых нами в начале § 3 чисел $g_k$ ещё следующим образом:

$$E_A = V \sum_{k=1}^{\infty} \frac{g_k f(k)}{e^{\alpha + \beta k} - \sigma} + O(1).$$

Это соотношение показывает, что $E_A$ асимптотически пропорционально $V$ (что ясно и непосредственно, так как $V$ пропорционально числу частиц $N$); вместе с тем это соотно-
§ 7. Корреляционная связь между числами заполнения

Нашей ближайшей задачей является оценка дисперсий сумматорных функций, а для этой цели нам придётся в первую очередь оценить корреляционную связь между числами заполнения с различными индексами, так как формула (24) § 6 гл. III показывает, что величина дисперсии существенным образом зависит от порядка малости разностей \( a_r a_s - a_r a_s \). Можно предвидеть, что, как и в случае фотонов, связь эта окажется весьма слабой, так как число частиц очень велико, а их взаимная корреляция обусловлена только соотношениями \( \sum_{1}^{\infty} a_r = N, \sum_{1}^{\infty} a_r \varepsilon_r = E \).

Мы увидим, что и здесь при вычитании \( a_r a_s \) из \( a_r a_s \) главные члены взаимно уничтожаются; вследствие этого та степень точности, с которой были выведены формулы § 6, здесь оказывается уже недостаточной, и мы должны снова произвести все расчёты, основывая их вместо формулы (37) на более точной формуле (38). Прежде всего формулы (22) и (38) дают нам для всех трёх статистик

\[
\mathcal{Q} (N + u_1, E + u_2) =
\]

\[
= C (N + u_1) \Phi (a, \beta) e^{\alpha N + \beta E} e^{u_1} + \beta u_2 \left\{ \frac{d}{2 \pi V V_0^2} + \right. \\
+ \left. \frac{1}{V_2} \left[ (m_0 + m_1 u_1 + m_2 u_2) - \frac{1}{4 \pi \delta / a} (b_{12} u_1^2 + 2 b_{12} u_1 u_2 + b_{11} u_2^2) \right] + \\
+ O \left( \frac{V V_0 + u_1}{V_3} \right) \right\},
\]

и, в частности, при \( u_1 = u_2 = O \):

\[
\mathcal{Q} (N, E) = C (N) \Phi (a, \beta) e^{\alpha N + \beta E} \left\{ \frac{d}{2 \pi V V_0^2} + \frac{m_0}{V_2^2} + O \left( \frac{1}{V_3^3} \right) \right\}.
\]
Отсюда
\[
\frac{\Omega (N+u_1, E+u_2)}{\Omega (N, E)} = \frac{C(N+u_1)}{C(N)} e^{\alpha u_1 + \beta u_2} \left\{ 1 + \frac{1}{V} \left[ l_1 u_1 + l_2 u_2 - \frac{1}{2\delta} (b_{22} u_1^2 - 2b_{12} u_1 u_2 + b_{11} u_2^2) \right] + O \left( \frac{\sqrt{V} + u^4}{V^2} \right) \right\}, \quad (44)
\]
где \( l_1 \) и \( l_2 \) не зависят ни от \( V \), ни от \( u_1 \) и \( u_2 \).

Эта более точная формула должна заменить нам в последующих расчётах формулу (39) предыдущего параграфа. Для сокращения записи мы положим для всех трёх статистик
\[
T_r = T_r(\alpha, \beta) = \frac{1}{e^{\alpha + \beta r} - \sigma}.
\]

Прежде всего мы найдём уточнённые выражения для средних значений чисел заполнения.

В случае полной статистики формула (6) § 2 в соединении с формулой (44) даёт:
\[
\overline{a}_r = e^{-(\alpha + \beta r)} \left\{ 1 - \frac{1}{V} \left[ l_1 + l_2 \epsilon_r + \frac{1}{2\delta} (b_{22} - 2b_{12} \epsilon_r + b_{11} \epsilon_r^2) \right] + O \left( \frac{V \sqrt{V} + \epsilon_r^4}{V^2} \right) \right\};
\]
замечая, что в случае полной статистики
\[
T_r = e^{-(\alpha + \beta r)},
\]
что следовательно,
\[
\frac{\partial T_r}{\partial \alpha} = -e^{-(\alpha + \beta r)}, \quad \frac{\partial T_r}{\partial \beta} = -\epsilon_r e^{-(\alpha + \beta r)},
\]
\[
\frac{\partial^2 T_r}{\partial \alpha^2} = -e^{-(\alpha + \beta r)}, \quad \frac{\partial^2 T_r}{\partial \alpha \partial \beta} = \epsilon_r e^{-(\alpha + \beta r)},
\]
\[
\frac{\partial^2 T_r}{\partial \beta^2} = \epsilon_r e^{-(\alpha + \beta r)},
\]
мы получаем отсюда:
\[
\overline{a}_r = T_r - \frac{1}{V} \left[ -l_1 \frac{\partial T_r}{\partial \alpha} - l_2 \frac{\partial T_r}{\partial \beta} + \frac{1}{2\delta} \left( b_{22} \frac{\partial^2 T_r}{\partial \alpha^2} - 2b_{12} \frac{\partial^2 T_r}{\partial \alpha \partial \beta} + b_{11} \frac{\partial^2 T_r}{\partial \beta^2} \right) \right] + O \left( \frac{T_r V \sqrt{V} \partial^4 T_r}{V^2} \right). \quad (45)
\]
в случае симметрической статистики формула (10) § 2 в соединении с формулой (44) даёт:

\[ \bar{a}_r = \sum_{k=1}^{\infty} e^{-k(\alpha + \beta \varepsilon_r)} \left\{ 1 - \frac{1}{V} \left[ k (l_1 + l_2 \varepsilon_r) + \frac{k^2}{2\delta} (b_{22} - 2b_{12} \varepsilon_r + b_{11} \varepsilon_r^2) \right] + O \left( \frac{V^2 + k^4 \varepsilon_r^4}{V^4} \right) \right\}. \quad (46) \]

Но на этот раз мы имеем:

\[ T_r = \frac{1}{e^\alpha + \beta \varepsilon_r - 1} = \sum_{k=1}^{\infty} e^{-k(\alpha + \beta \varepsilon_r)}, \]

откуда

\[ \frac{\partial T_r}{\partial \alpha} = -\sum_{k=1}^{\infty} ke^{-k(\alpha + \beta \varepsilon_r)}, \quad \frac{\partial T_r}{\partial \beta} = -\sum_{k=1}^{\infty} k \varepsilon_r e^{-k(\alpha + \beta \varepsilon_r)}, \]

\[ \frac{\partial^2 T_r}{\partial \alpha^2} = \sum_{k=1}^{\infty} k^2 e^{-k(\alpha + \beta \varepsilon_r)} \varepsilon_r, \quad \frac{\partial^2 T_r}{\partial \alpha \partial \beta} = \sum_{k=1}^{\infty} k^2 \varepsilon_r^2 e^{-k(\alpha + \beta \varepsilon_r)}, \]

\[ \frac{\partial^2 T_r}{\partial \beta^2} = \sum_{k=1}^{\infty} k^2 \varepsilon_r e^{-k(\alpha + \beta \varepsilon_r)}. \]

Потому мы находим:

\[ \bar{a}_r = T_r - \frac{1}{V} \left[ -l_1 \frac{\partial T_r}{\partial \alpha} - l_2 \frac{\partial T_r}{\partial \beta} + \frac{1}{2\delta} \left( b_{22} \frac{\partial^2 T_r}{\partial \alpha^2} - 2b_{12} \frac{\partial^2 T_r}{\partial \alpha \partial \beta} + b_{11} \frac{\partial^2 T_r}{\partial \beta^2} \right) \right] + O \left( \frac{T_r V^2 + \frac{\partial^2 T_r}{\partial \beta^2}}{V^4} \right), \]

т. е. в точности ту же формулу (45), что и в случае полной статистики.

Наконец, в случае антисимметрической статистики формула (13) § 2 в соединении с формулой (44) даёт нам формулу, отличающуюся от (46) лишь тем, что к-й член суммы умно-жается на \((-1)^{k-1}(k = 1, 2, \ldots). \) Так как в этом случае

\[ T_r = \frac{1}{e^\alpha + \beta \varepsilon_r - 1} = \sum_{k=1}^{\infty} (-1)^{k-1} e^{-k(\alpha + \beta \varepsilon_r)}, \]
то мы легко убеждаемся, что формула (45) без всяких изменений остаётся в силе и в случае антисимметрической статистики. Таким образом, формула (45), которая и представляет собою нужное нам уточнённое асимптотическое выражение для $a_r$, имеет место в случае всех трёх статистик (привём, разумеется, $T_r$ имеет в этих трёх случаях различное значение).

Перейдём теперь к оценке величин $a_r, a_s$ ($r \neq s$). В случае полной статистики мы должны исходить из формулы (7) § 2, которая в соединении с формулой (44) даёт:

$$
\overline{a_r a_s} = e^{-(\alpha + \beta r)} e^{-(\alpha + \beta s)} \left\{ 1 - \frac{1}{V} \left[ (l_1 + l_2 \varepsilon_r) + (l_1 + l_2 \varepsilon_s) + \frac{1}{2\delta} \left( 4b_{22} - 4b_{12} [\varepsilon_r + \varepsilon_s] + b_{11} [\varepsilon_r + \varepsilon_s]^2 \right) \right] + O \left( \frac{\sqrt{V} + \varepsilon_r + \varepsilon_s}{V^2} \right) \right\} = T_r T_s - \frac{1}{V} \left[ -l_1 \frac{\partial (T_r T_s)}{\partial a} - l_2 \frac{\partial (T_r T_s)}{\partial \beta} + \frac{1}{2\delta} \left( b_{22} \frac{\partial^2 (T_r T_s)}{\partial a^2} - 2b_{12} \frac{\partial^2 (T_r T_s)}{\partial a \partial \beta} + b_{11} \frac{\partial^2 (T_r T_s)}{\partial \beta^2} + \right) \right] + O \left( \frac{T_r T_s \sqrt{V} + T_r \frac{\partial^4 T_r}{\partial \beta^4} + T_s \frac{\partial^4 T_s}{\partial \beta^4}}{V^2} \right). \tag{47}
$$

Так как метод получения формулы этого рода показан нами теперь уже на значительном числе примеров, то мы можем предоставить читателю самостоятельно показать, что формула (47), выведенная нами в предположении полной статистики, без всяких изменений остаётся верной, подобно формуле (45), и для двух других статистических схем.

Переходим теперь к оценке разности $|a_r a_s - \overline{a_r a_s}|$. Прежде всего, формула (45) даёт:

$$
|a_r a_s - \overline{a_r a_s}| = T_r T_s - \frac{1}{V} \left[ -l_1 \frac{\partial (T_r T_s)}{\partial a} - l_2 \frac{\partial (T_r T_s)}{\partial \beta} + \frac{1}{2\delta} \left( b_{22} \frac{\partial^2 T_s}{\partial a^2} + T_s \frac{\partial^2 T_r}{\partial a^2} \right) - 2b_{12} \frac{\partial^2 (T_r T_s)}{\partial a \partial \beta} + T_s \frac{\partial^2 T_r}{\partial a \partial \beta} + b_{11} \left( \frac{T_r \frac{\partial^2 T_s}{\partial \beta^2} + T_s \frac{\partial^2 T_r}{\partial \beta^2}}{\partial \beta^2} \right) \right] + O \left( \frac{T_r T_s \sqrt{V} + T_r \frac{\partial^4 T_s}{\partial \beta^4} + T_s \frac{\partial^4 T_r}{\partial \beta^4}}{V^2} \right). \tag{48}
$$
Далее, вычитая почленно (48) из (47), мы находим:

$$a_r a_s - a_r a_s = -\frac{1}{\delta V} \left\{ b_{22} \frac{\partial T_r}{\partial \alpha} \frac{\partial T_s}{\partial \alpha} - b_{12} \left( \frac{\partial T_r}{\partial \alpha} \frac{\partial T_s}{\partial \beta} + \right. \right.$$  

$$+ \left. \frac{\partial T_s}{\partial \alpha} \frac{\partial T_r}{\partial \beta} \right) + b_{11} \frac{\partial T_r}{\partial \beta} \frac{\partial T_s}{\partial \beta} \right\} +$$  

$$+ O \left( \frac{T_r T_s \sqrt{V} + T_r \frac{\partial^4 T_s}{\partial \alpha^4} + T_s \frac{\partial^4 T_r}{\partial \beta^4}}{V^2} \right).$$  

(49)

Эта формула, подобно формулам (45), (47) и (48), имеет место для любой из трёх основных статистических схем. Она показывает, что разность $a_r a_s - a_r a_s$, измеряющая собою корреляционную зависимость между числами $a_r$ и $a_s$ ($s \neq r$), есть в наших условиях бесконечно малая, асимптотически пропорциональная $V^{-1}$. Задача, которую мы себе поставили в настоящем параграфе, таким образом полностью разрешена.

§ 8. Дисперсии суммарных величин и представительность микроканонических средних

Мы переходим теперь к оценке дисперсий суммарных величин, необходимой, как мы знаем, для установления представительности их микроканонических средних. Кроме уже найденных нами асимптотических оценок чисел $a_r$ и $a_r a_s - a_r a_s$, нам понадобится для этой цели оценить и микроканонические дисперсии чисел $a_r$, т. е. числа $\bar{a}_r^2 - (\bar{a}_r)^2$; этим мы и займёмся в первую очередь.

В конце § 2 [формула (20)] мы убедились, что для всех трёх основных статистик

$$\bar{a}_r^2 = a_r + (1 + \sigma) \tau_r,$$

где

$$\tau_r = a_r a_s \quad (s \neq r, \varepsilon_s = \varepsilon_r);$$

отсюда

$$\bar{a}_r^2 -(\bar{a}_r)^2 = a_r \sigma (\bar{a}_r)^2 + (1 + \sigma) [a_r a_s - a_r a_s], r \neq s.$$  

(50)
Эта формула полностью решает поставленную задачу, так как асимптотические выражения чисел \( \bar{a}_r \) и \( a_r a_s - \bar{a}_r \bar{a}_s \) \( (r \neq s) \) даются соответственно формулами (45) и (49) предыдущего параграфа.

Согласно формуле (24) § 6 гл. III, мы имеем для микроканонической дисперсии сумматорной величины \( \mathcal{U} \) выражение

\[
D (\mathcal{U}) = \sum_{r=1}^{\infty} (\mu_r - \lambda_r^2) \bar{a}_r + \sum_{r,s=1}^{\infty} \lambda_r \lambda_s (a_r a_s - \bar{a}_r \bar{a}_s) = \]

\[
= \sum_{r=1}^{\infty} (\mu_r - \lambda_r^2) \bar{a}_r + \sum_{r \neq s} \lambda_r \lambda_s (a_r a_s - \bar{a}_r \bar{a}_s) + \]

\[
+ \sum_{r=1}^{\infty} \lambda_r^2 \left[ \bar{a}_r^2 - (\bar{a}_r)^2 \right]. \tag{51}
\]

Полагая для любых \( r \) и \( s \)

\[
c_{rs} = \begin{cases} 
  a_r a_s - \bar{a}_r \bar{a}_s & (r \neq s), \\
  [a_r a_s - \bar{a}_r \bar{a}_s]_{r \neq s} & (r = s),
\end{cases}
\]

мы можем переписать формулу (50) в виде

\[
\bar{a}_r^2 - (\bar{a}_r)^2 = \bar{a}_r + \sigma (\bar{a}_r)^2 + (1 + \sigma) c_{rr};
\]

вставляя же это выражение в правую часть формулы (51), находим:

\[
D (\mathcal{U}) = \sum_{r=1}^{\infty} (\mu_r - \lambda_r^2) \bar{a}_r + \sum_{r \neq s} \lambda_r \lambda_s c_{rs} + \]

\[
+ \sum_{r=1}^{\infty} \lambda_r^2 \left[ \bar{a}_r + \sigma (\bar{a}_r)^2 \right] + (1 + \sigma) \sum_{r=1}^{\infty} \lambda_r^2 c_{rr} = \]

\[
= \sum_{r \neq s} \lambda_r \lambda_s c_{rs} + \sum_{r=1}^{\infty} \lambda_r^2 c_{rr} + \sum_{r=1}^{\infty} \bar{a}_r + \sigma \lambda_r^2 \left[ c_{rr} + (\bar{a}_r)^2 \right] \} = \]

\[
= \sum_{r,s=1}^{\infty} \lambda_r \lambda_s c_{rs} + \sum_{r=1}^{\infty} \{ \mu_r \bar{a}_r + \sigma \lambda_r^2 [c_{rr} + (\bar{a}_r)^2] \}. \tag{52}
\]
Переходя теперь к асимптотической оценке величины $D(\mathfrak{H})$, заметим прежде всего, что правая часть формулы (49), очевидно, при любых $r$ и $s$ (равных или неравных между собою) выражает величину $c_{rs}$. Положим для краткости $\sum_{r=1}^{\infty} \lambda_r T_r = Q = Q(\alpha, \beta)$. Применяя формулу (49), мы получаем:

$$
\sum_{r, s=1}^{\infty} \lambda_r \lambda_s c_{rs} = -\frac{1}{\delta V} \left\{ b_{22} \left( \frac{\partial Q}{\partial \alpha} \right)^2 - 2b_{12} \frac{\partial Q}{\partial \alpha} \frac{\partial Q}{\partial \beta} + \right.

+ b_{11} \left( \frac{\partial Q}{\partial \beta} \right)^2 \left. \right\} + O \left( \frac{Q^2 V \bar{V} + Q \frac{\partial^4 Q}{\partial \beta^4}}{V^2} \right).
$$

Вспомним теперь, что, как мы уже неоднократно это отмечали (см., например, начало § 3), всякий абсолютно сходящийся ряд вида

$$
\sum_{r=1}^{\infty} f(\varepsilon_r)
$$

представляет собою величину, пропорциональную $V$. Но величина $Q$ определена нами именно как ряд такого вида, так как $\lambda_r$ и $T_r$ зависят от $\varepsilon_r$; очевидно, что и любая частная производная функции $Q(\alpha, \beta)$ представляется в виде такого ряда. Таким образом, в правой части последнего равенства величина $Q$ и все её частные производные представляют собою величины, пропорциональные $V$. Это даёт нам:

$$
\sum_{r, s=1}^{\infty} \lambda_r \lambda_s c_{rs} = C_1 V + O(V \sqrt{V}),
$$

где $C_1$ — постоянная.

Далее, в силу того же самого замечания и

$$
\sum_{r=1}^{\infty} \{ \mu_r \bar{a}_r + \sigma \lambda_r [c_{rr} + (\bar{a}_r)^2] \} = C_2 V,
$$

где $C_2$ — другая постоянная. Формула (52) даёт поэтому ввиду постоянства отношения $\frac{V}{N}$:

$$
D(\mathfrak{H}) = (C_1 + C_2) \sqrt{V} + O(V \sqrt{V}) = CN + O(V \sqrt{N}).
$$
Таким образом, микроканоническая дисперсия суммарной величины $\mathcal{W}$ оказывается асимптотически пропорциональной числу частиц $N$. В § 6 гл. III мы видели, что для представительности микроканонических средних достаточно выполнения соотношения

$$D(\mathcal{W}) = o(N^2);$$

мы видим теперь, что для суммарных величин это соотношение выполняется с избытком. Вместе с тем мы получили вполне реализуемый метод для асимптотической оценки микроканонических дисперсий, что имеет существенное значение для теории флуктуаций в физике.

§ 9. Определение чисел $g_k$ для бесструктурных частиц в отсутствии внешних сил

По аналогии с тем, что мы сделали для фотонов в § 5 предшествующей главы, мы определим теперь для материальных частиц число $Vg_k$ возможных уровней энергии, заключённых между $k$ и $k + 1$, при условии, что частицы заключены в сосуд объёма $V$. Мы снова предположим, что состояние частицы определяется гамильтоновыми переменными $x, y, z, p_x, p_y, p_z$ и что частица не подвержена действию внешних сил, так что её полная энергия состоит только из кинетической энергии и потенциала стенок сосуда, который мы можем предположить равным нулю внутри сосуда и бесконечно большим вне его. Мы вновь подчёркиваем, что занимающая нас сейчас проблема не имеет никакого отношения к статистике — в частности, мы можем представлять себе, что имеем дело с одной единственной частицей.

Как было уже замечено в § 5 предшествующей главы, для материальных частиц энергия $\varepsilon$ связана с импульсом $p$ соотношением

$$\varepsilon = mc^2 + \frac{p^2}{2m},$$

где $m$ — масса частицы, а $c$ — скорость света в пустоте. Выбирая надлежащим образом начало отсчёта энергии, мы можем заменить это соотношение более простым

$$\varepsilon = \frac{p^2}{2m};$$
Это даёт для оператора энергии $H$ выражение

$$H = -\frac{\hbar^2}{8\pi^2m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right).$$

Таким образом, свободное от времени уравнение Шредингера получает вид

$$-\frac{\hbar^2}{8\pi^2m} \left( \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} \right) = EU,$$

или

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} = -\frac{8\pi^2mE}{\hbar^2} U \tag{53}$$

(при этом мы учтываем лишь кинетическую энергию частицы, так как потенциал стенок обращается в нуль внутри сосуда). Решения этого уравнения и будут описывать собою стационарные состояния частицы, соответствующие уровню энергии $E$.

Мы уже заметили в § 5 предыдущей главы, что линейный базис решений уравнения (53) может быть составлен из функций вида

$$U = C \sin (ax + \alpha) \sin (by + \beta) \sin (cz + \gamma),$$

где только на этот раз

$$a^2 + b^2 + c^2 = \frac{8\pi^2mE}{\hbar^2}. \tag{54}$$

Мы снова допустим, что заключающий данную частицу сосуд есть параллелепипед $0 \leq x \leq l_1$, $0 \leq y \leq l_2$, $0 \leq z \leq l_3$, $l_1l_2l_3 = V$. Так как величина $|U(x, y, z)|^2$, пропорциональная вероятности найти частицу в точке $(x, y, z)$ сосуда, должна быть равна нулю у стенок сосуда, то любое из условий $x = 0$, $y = 0$, $z = 0$, $x = l_1$, $y = l_2$, $z = l_3$ должно иметь своим следствием $U = 0$. Первые три условия приводят, очевидно, к требованию $a = b = c = 0$, так что

$$U = C \sin ax \sin by \sin cz, \quad a^2 + b^2 + c^2 = \frac{8\pi^2mE}{\hbar^2}.$$

Из того, что мы должны иметь $U = 0$ при $x = l_1$, вытекает, очевидно, $a = \frac{n_1\pi}{l_1}$, где $n_1$ — целое число; подобным же образом два последних условия приводят к требованиям
\[ b = \frac{n_2 \pi}{l_2}, \quad c = \frac{n_3 \pi}{l_3}, \text{ где } n_2 \text{ и } n_3 \text{ — целые числа. В силу (54)} \]
числа \( n_1, n_2 \) и \( n_3 \) должны быть связаны соотношением
\[ \frac{n_1^2}{l_1^2} + \frac{n_2^2}{l_2^2} + \frac{n_3^2}{l_3^2} = \frac{8mE}{\hbar^2}. \]

Но это означает, что возможными уровнями энергии стационарных состояний частицы служат все те и только те числа \( E \), которые имеют вид
\[ \frac{\hbar^2}{8m} \left( \frac{n_1^2}{l_1^2} + \frac{n_2^2}{l_2^2} + \frac{n_3^2}{l_3^2} \right), \]
где \( n_1, n_2, n_3 \) — целые числа, которые можно предполагать неотрицательными; каждая такая тройка даёт нам одно из линейно независимых решений уравнения Шредингера (53), соответствующих данным граничным условиям (\( U = 0 \) при выполнении хотя бы одного из равенств \( x = 0, y = 0, z = 0, x = l_1, y = l_2, z = l_3 \)).

Пусть теперь \( k \) — любое натуральное число; тогда число линейно независимых собственных функций оператора энергии частицы, соответствующих собственным значениям \( E \leq k \), будет, в силу только что сказанного, равно числу решений неравенства
\[ \frac{\hbar^2}{8m} \left( \frac{n_1^2}{l_1^2} + \frac{n_2^2}{l_2^2} + \frac{n_3^2}{l_3^2} \right) \leq k \]
в целых \( n_1 \geq 0, n_2 \geq 0, n_3 \geq 0 \). Это же число асимптотически (при больших \( k \)) равно, очевидно, одной восьмой части объёма эллипсоида
\[ \frac{\hbar^2}{8m} \left( \frac{x^2}{l_1^2} + \frac{y^2}{l_2^2} + \frac{z^2}{l_3^2} \right) = k, \]
т. е. равно
\[ Q(k) = \frac{1}{8} \left\{ \frac{4}{3} \pi l_1 l_2 l_3 \left( \frac{8mk}{\hbar^2} \right)^{\frac{1}{2}} \right\} = \frac{8}{3} \pi V \left( \frac{m}{\hbar^2} \right)^{\frac{1}{2}} k^{\frac{3}{2}}. \]

Число же стационарных состояний, уровни энергии которых заключены между \( k \) и \( k + dk \), приближённо равно
\[ Q'(k) dk = 4 \sqrt{2} \pi V \left( \frac{m}{\hbar^2} \right)^{\frac{1}{2}} k^{\frac{3}{2}} dk; \]  \[ (55) \]
при этом интервал \( dk \) должен быть для реальной значимости полученной приближённой формулы не слишком малым (но, разумеется, весьма малым в сравнении с \( k \)). С другой стороны, это же число в своих обозначениях равно

\[
V \sum_{r=k}^{k+dk} g_r,
\]

так что

\[
\sum_{r=k}^{k+dk} g_r \approx \frac{4\pi \sqrt{2} m^\frac{3}{2}}{h^3} \sqrt{k} \, dk;
\]

«среднее» значение \( g_r \) в промежутке \( k < r < k + dk \) поэтому равно

\[
\frac{4\pi \sqrt{2} m^\frac{3}{2}}{h^3} \sqrt{k}.
\]

Как и в случае фотонов, для большинства расчётов можно прямо положить

\[
g_k = \frac{4\pi \sqrt{2} m^\frac{3}{2}}{h^3} \sqrt{k},
\]

в особенности при больших значениях \( k \). Как и в случае фотонов, это число должно быть удвоено, если рассматриваемые нами частицы обладают «спином».

Сделаем ещё следующее замечание. В классической физике гамильтонова функция рассматриваемой нами частицы внутри сосуда имеет вид

\[
H = \frac{1}{2m} \left(p_x^2 + p_y^2 + p_z^2\right).
\]

Объём \( V(k) \) части (шестимерного) фазового пространства частицы, в точках которой \( H < k \), равен произведению объёма сосуда \( V \) на объём сферы \( H = k \), равный, очевидно,

\[
\frac{4\pi}{3} \left(2mk\right)^\frac{3}{2},
\]

т. е.

\[
V(k) = \frac{8\sqrt{2\pi}}{3} V m^\frac{3}{2} k^\frac{3}{2},
\]
и, следовательно, объём той части фазового пространства, где \( k < H < k + 1 \), равен

\[
V(k + 1) - V(k) \approx V'(k) = 4 \sqrt{2} \pi V m^2 \sqrt{k};
\]

сопоставляя это с формулой (55), мы находим:

\[
V_{gr, k} \approx \frac{V(k + 1) - V(k)}{h^3}.
\]

Это означает, что в среднем число стационарных состояний с уровнями энергии, заключёнными между \( k \) и \( k + 1 \), равно объёму части \( k < H < k + 1 \) классического фазового пространства частицы, если принять величину \( h^3 \) за единицу объема; иначе говоря, одно стационарное состояние в известном смысле эквивалентно «ячейке» объёма \( h^3 \) классического фазового пространства.
ГЛАВА VI
ТЕРМОДИНАМИЧЕСКИЕ ВЫВОДЫ

§ 1. Задачи статистической термодинамики

Важнейшей задачей статистической физики всегда было объяснение законов термодинамики на базе представления об атомистическом строении вещества. Известно, что термодинамика достигла высокой степени теоретического развития независимо от этих представлений; логическая схема её была доведена до чисто дедуктивного (аксиоматического) построения, где некоторое небольшое число основных принципов играет роль аксиом, из которых затем все дальнейшие законо мерности выводятся уже чисто логическим путём. Основные принципы (аксиомы) понимаются при этом как законы природы, найденные опытным путём. Поэтому задача обоснования термодинамики на базе тех или иных представлений о строении материи всегда сводится к обоснованию этих «аксиом» и может считаться решённой, как только обоснование этих принципов закончено.

Как может быть проведено такое обоснование на базе развиваемых нами статистических теорий? Чтобы подойти к решению этого вопроса, мы должны вспомнить не раз уже упомянутое основное возникающее здесь принципиальное затруднение. В «классической» (т. е. феноменологической, не статистической) термодинамике характеризующие данные тела величины зависят от очень небольшого числа основных величин — таких, как энергия, объём, температура, давление, энтропия и т. п. Если значения этих величин известны, то термодинамическое состояние тела считается однозначно определённым. В противоположность этому, в статистической теории мы имеем такое положение, когда, например, при за-
данных значениях энергии и объёма существует огромное количество различных состояний, совместимых с этими значениями. Каждая собственная функция $U$ оператора энергии, принадлежащая собственному значению $E$ этого оператора, определяет одно из таких возможных состояний; но таких собственных функций всегда имеется целый континуум, и даже если мы ограничившимся линейным базисом этого континуума, мы всё же будем иметь конечное, но очень большое число $\Omega(V, E)$ таких состояний. Какая-нибудь величина $\bar{A}$ в классической термодинамике, однозначно определяемая значениями энергии $E$ и объёма $V$ данной системы, в статистической теории будет иметь в различных состояниях, совместимых с данными $E$ и $V$, весьма различные значения. А между тем статистическая теория, если она претендует на роль логической базы термодинамических принципов, должна, конечно, давать для такой величины совершенно однозначное определение, совпадающее с тем, какое даёт ей классическая термодинамика.

Из этого затруднения может быть, как мы уже неоднократно это видели, только один выход: в качестве того значения величины $\bar{A}$, которое подлежит сопоставлению с соответствующей величиной классической термодинамики, статистическая теория должна давать некоторое среднее значение $\bar{A}$ этой величины, причём осреднение должно вестись по всем доступным для системы состояниям $U$, соответствующим данному уровню энергии $E$ (и, если это нужно, ещё и данным значениям некоторых других параметров). Принцип осреднения при этом может быть выбран произвольно, но статистическая теория обязана показать представительность даваемых ею средних, т. е. показать, что в значительном большинстве состояний значение величины $\bar{A}$ весьма близко к $\bar{A}$; пока этого не сделано, единичная опытная проверка даваемого теорией значения, очевидно, лишена всякого смысла, так же как и сравнение этого значения с тем, какое даётся классической теорией. В гл. IV и V это доказательство представительности действительно было нами проведено для важнейших случаев (в предположении микроканонического принципа осреднения).

Таким образом, путь, которым мы должны идти, представляется ясным: для каждой из величин, с которыми имеет дело классическая термодинамика, нам надлежит установить
§ 2. Внешние параметры, внешние силы и их средние значения

Гамильтоновы функции изучаемых нами физических систем и отдельных частиц могут зависеть от того или иного числа параметров, определяющих собою положение или состояние внешних тел; изменение значений этих параметров вызывает собою изменение вида оператора энергии и тем самым заставляет изменяться и определяемые уравнением Шредингера уровни энергии и возможные состояния системы или частицы. Во всём предшествующем наличии такого рода параметров в выражении потенциальной энергии системы или частицы нами отнюдь не исключалось; более того, при рассмотрении системы частиц, заключённых внутри некоторого сосуда, мы фактически всегда считали потенциальную энергию зависящей от объёма \( V \) этого сосуда, представляющего собою как раз важнейший параметр этого рода; другими подобными параметрами могут быть, например, величины, определяющие собою положение или состояние источников, действующих на данную систему внешних полей. До сих пор мы, однако, оставляли возможное присутствие такого рода параметров без особого внимания; мы могли это делать потому, что значения этих параметров до сих пор предполагались строго постоянными; именно в этом смысле мы и называли нашу систему изолированной. Теперь же, напротив, мы должны будем сосредоточить внимание на таких взаимодействиях между данной системой и окружающими её телами, при которых этого рода параметры претерпевают изменения, связанные с определённой работой частиц нашей системы, а потому связанные и с изменением энергии как этих частиц, так и всей системы в целом. Так, заключённый в сосуде цилиндрической формы газ, расширяясь, производит силою своих частиц работу, передвигающую поршень и тем самым изменяющую объём сосуда.
Описанные нами параметры мы будем называть внешними параметрами данной системы. Гамильтонова функция каждой частицы зависит, таким образом, кроме обычных гамильтоновых переменных этой частицы, ещё от ряда внешних параметров \( \lambda_1, \lambda_2, \ldots, \lambda_n \); с математической стороны при этом внешние параметры характеризуются тем, что имеют для всех частиц одно и то же значение; в этом можно, если угодно, видеть даже формальное определение внешнего параметра [надо только помнить, что этим же свойством (одно и то же значение для всех частиц) обладают и введенные нами в предшествующем параметры \( \alpha \) и \( \beta \), обычно называемые внутренними параметрами системы; мы знаем, что при известном строении частиц и известных значениях внешних параметров числа \( \alpha \) и \( \beta \) становятся однозначно определёнными, если заданы число частиц \( N \) и полная энергия \( E \) данной системы].

Пусть одна из частиц нашей системы находится в \( r \)-м стационарном состоянии (уровень энергии \( \varepsilon_r \)) и пусть элементарная работа \( dw \), произведённая этой частицей, вызывает изменение \( d\lambda_i \) внешнего параметра \( \lambda_i \). Эта работа сопровождается изменением \( d\varepsilon_r \), энергии нашей частицы, которое равно

\[ d\varepsilon_r = \frac{\partial \varepsilon_r}{\partial \lambda_i} d\lambda_i; \]

а так как в силу закона сохранения энергии \( dw = -d\varepsilon_r \), то

\[ dw = - \frac{\partial \varepsilon_r}{\partial \lambda_i} d\lambda_i. \]

Коэффициент при \( d\lambda_i \) в этом выражении элементарной работы в механике называют обобщённой силой, с которой выбранной нами частицей действует на внешние тела «в направлении» параметра \( \lambda_i \); таким образом, для выбранной нами частицы эта обобщённая сила равна \( \frac{\partial \varepsilon_r}{\partial \lambda_i} \) (само собою разумеется, мы представляем себе дело так, что возможные уровни \( \varepsilon_r \), энергии частицы зависят от значений внешних параметров; в гл. IV и V мы, например, видели, что уровни эти существенным образом зависят от занимаемого системой объёма, который в термодинамике служит важнейшим внешним параметром).

Если изменение \( d\lambda_i \) параметра \( \lambda_i \) вызывается совокупной работой всех частиц данной системы, то обобщённая сила \( \Lambda_i \),
с которой «в направлении» параметра $\lambda_i$ действует на внешние тела вся данная система, равна сумме обобщённых сил действия всех частиц системы. Пусть система находится в основном состоянии $U$, которому соответствует определённый набор чисел заполнения $a_r(U)$; тогда мы имеем, следовательно,

$$
\Lambda_i = - \sum_{r=1}^{\infty} a_r(U) \frac{\partial \varepsilon_r}{\partial \lambda_i};
$$

это соотношение показывает, что обобщённые внешние силы $\Lambda_i$ в статистической теории являются фазовыми функциями (определёнными по меньшей мере на множестве основных собственных функций). Если элементарная работа $\delta W$ всей системы имеет своим следствием изменения $d\lambda_1, d\lambda_2, \ldots, d\lambda_s$ внешних параметров, то мы имеем:

$$
\delta W = \sum_{i=1}^{s} \Lambda_i d\lambda_i = - \sum_{i=1}^{s} \left\{ \sum_{r=1}^{\infty} a_r(U) \frac{\partial \varepsilon_r}{\partial \lambda_i} \right\} d\lambda_i.
$$

Но так как

$$
\sum_{r=1}^{\infty} a_r(U) \varepsilon_r = E,
$$

очевидно, есть полная энергия системы, то

$$
\Lambda_i = - \sum_{r=1}^{\infty} a_r(U) \frac{\partial \varepsilon_r}{\partial \lambda_i} = - \frac{\partial E}{\partial \lambda_i},
$$

(1)

и мы можем написать

$$
\delta W = - \sum_{i=1}^{s} \frac{\partial E}{\partial \lambda_i} d\lambda_i;
$$

при этом необходимо иметь в виду, что $\frac{\partial E}{\partial \lambda_i}$ представляет собой фазовую функцию, определяемую соотношением (1), так что при данных значениях полной энергии и внешних параметров $\frac{\partial E}{\partial \lambda_i}$ ещё не является определённым, но зависит от того, в каком состоянии $U$ находится система.

Обобщённые внешние силы

$$
\Lambda_i = - \frac{\partial E}{\partial \lambda_i}
$$
являются, таким образом, в нашей теории фазовыми функциями (и притом, очевидно, суммарными величинами). В силу приведённых в § 1 общих методологических соображений мы должны поэтому считать, что те обобщённые силы, с которыми имеет дело феноменологическая термодинамика, должны иметь своими аналогами в статистической теории микроканоническое среднее этих фазовых функций, т. е. величины

$$
\overline{\lambda}_i = - \frac{\partial E}{\partial \lambda_i} = - \sum_{r=1}^{\infty} a_r \frac{\partial \varepsilon_r}{\partial \lambda_i};
$$

(2)

подобным же образом и классически понимаемая элементарная работа системы над внешними телами должна интерпретироваться в нашей теории как микроканоническое среднее

$$
\overline{\delta W} = - \sum_{i=1}^{s} \frac{\partial E}{\partial \lambda_i} d\lambda_i = - \sum_{i=1}^{s} \left( \sum_{r=1}^{\infty} a_r \frac{\partial \varepsilon_r}{\partial \lambda_i} \right) d\lambda_i.
$$

Приближённые выражения для величин $\overline{\lambda}_i$, даваемых нашей теорией, мы получим, если внесём в формулу (2) вместо чисел $a_r$ найденные нами в § 5 гл. V приближённые выражения

$$
a_r \approx \frac{1}{e^{\alpha + \beta \varepsilon_r} - \sigma},
$$

где $\sigma$ — индекс симметрии нашей системы. Это даёт:

$$
\overline{\lambda}_i \approx - \sum_{r=1}^{\infty} \frac{\partial \varepsilon_r}{e^{\alpha + \beta \varepsilon_r} - \sigma} \quad (i = 1, 2, ..., s).
$$

(3)

Вспомним теперь, что в § 5 гл. V мы нашли для функции $\Phi (\alpha, \beta)$ в случае любой из трёх статистических схем выражение

$$
\Phi (\alpha, \beta) = \prod_{r=1}^{\infty} \left\{ \sum_{n=0}^{\infty} \gamma (n) e^{-n (\alpha + \beta \varepsilon_r)} \right\},
$$
где

\[
\gamma(n) = \begin{cases} 
\frac{1}{n!} & \text{(полная статистика),} \\
1 & \text{(симметрическая статистика),} \\
1 & \text{(n \leq 1),} \\
0 & \text{(n > 1).}
\end{cases}
\]

Таким образом, в случае полной статистики

\[
\Phi(\alpha, \beta) = \prod_{r=1}^{\infty} e^{-\left(\alpha + \beta e_r\right)},
\]

\[
\ln \Phi(\alpha, \beta) = \sum_{r=1}^{\infty} e^{-\left(\alpha + \beta e_r\right)};
\]

в случае симметрической статистики

\[
\Phi(\alpha, \beta) = \prod_{r=1}^{\infty} \frac{1}{1 - e^{-\left(\alpha + \beta e_r\right)}},
\]

\[
\ln \Phi(\alpha, \beta) = -\sum_{r=1}^{\infty} \ln \left(1 - e^{-\left(\alpha + \beta e_r\right)}\right);
\]

наконец, в случае антисимметрической статистики

\[
\Phi(\alpha, \beta) = \prod_{r=1}^{\infty} \left\{1 + e^{-\left(\alpha + \beta e_r\right)}\right\},
\]

\[
\ln \Phi(\alpha, \beta) = \sum_{r=1}^{\infty} \ln \left(1 + e^{-\left(\alpha + \beta e_r\right)}\right).
\]

Элементарный подсчёт на основе этих формул показывает, что во всех трёх случаях имеет место соотношение

\[
\frac{\partial \ln \Phi(\alpha, \beta)}{\partial \lambda_i} = -\beta \sum_{r=1}^{\infty} \frac{\partial e_r}{\partial \lambda_i} e^\alpha + \beta e_r - \sigma.
\]

Поэтому формула (3) даёт для всех трёх случаев

\[
\Lambda_i = \frac{1}{\beta} \frac{\partial \ln \Phi}{\partial \lambda_i} \quad (1 \leq i \leq s),
\]
и, следовательно,

$$\delta W = \sum_{i=1}^{s} \Lambda_i d\lambda_i = \frac{1}{\beta} \sum_{i=1}^{s} \frac{\partial \ln \Phi}{\partial \lambda_i} d\lambda_i.$$  (5)

В гл. V мы видели, что число частиц $N$ и полная энергия $E$ системы получаются дифференцированием функции $\ln \Phi(\alpha, \beta)$ по параметрам $\alpha$ и $\beta$; теперь же мы видим, что средние значения обобщённых сил и элементарной работы системы просто выражаются через частные производные той же функции $\ln \Phi$ по соответствующим внешним параметрам.

Формулы (4) и (5), как мы увидим, будут лежать в основе всех термодинамических выводов излагаемой нами статистической теории.

§ 3. Определение энтропии и вывод второго закона термодинамики

Занимающая столь видное место в нашей теории функция $\Phi(\alpha, \beta)$, выражение которой для трёх основных статистических схем установлено нами в предшествующем параграфе, зависит, как мы видели там же, кроме $\alpha$ и $\beta$ ещё и от внешних параметров $\lambda_1, \lambda_2, \ldots, \lambda_s$. Логарифмическое дифференцирование этой функции по её различным аргументам даёт нам выражение для таких важнейших характеризующих состояние системы величин, как число частиц $N$, энергия $E$ и внешние силы $\Lambda_i$ ($i = 1, 2, \ldots, s$). Аналогичными свойствами обладает в феноменологической термодинамике, как известно, так называемая «характеристическая функция Планка», иначе имеемая «термодинамическим потенциалом». В классической теории эта функция зависит от температуры системы и внешних параметров. В нашей теории «внутренних» параметров имеется два — $\alpha$ и $\beta$; однако ввиду предполагаемой нами неизменности числа частиц $N$ данной системы соотношение

$$\frac{\partial \ln \Phi(\alpha, \beta)}{\partial \alpha} = -N$$

позволяет исключить один из этих параметров, например $\alpha$, из выражения как самой функции $\ln \Phi(\alpha, \beta)$, так и из любой из её частных производных (по $\beta$, $\lambda_1$, $\lambda_2$, $\ldots$, $\lambda_s$). Таким об-
разом, мы можем считать, что функция $\ln \Phi(\alpha, \beta)$ и каждая из этих частных производных зависит от параметров $\beta, \lambda_1, ..., \lambda_s$. Проводимая нами аналогия естественно приводит к допущению, что параметр $\beta$ должен быть однозначно связан с температурой системы; всё, что мы знаем о роли этого параметра в нашей теории, подтверждает это допущение; так, во всех случаях, когда система составляеться из нескольких компонент, находящихся в тепловом контакте друг с другом, существует единственное определённое значение параметра $\beta$, общее для всех этих компонент; параметр $\beta$ физически должен выражать собой поэтому такую величину, которая в случае теплового равновесия между несколькими системами имеет для всех этих частей одно и то же значение; но именно такова, как известно, роль температуры в термодинамике.

Вспомним теперь, что значения параметров $\alpha$ и $\beta$ нами были выбраны (гл. V, § 4) так, чтобы функция

$$F(\alpha, \beta) = e^{\alpha N + \beta E} \Phi(\alpha, \beta)$$

имела в точке $(\alpha, \beta)$ своё наименьшее значение. $F(\alpha, \beta)$, подобно функции $\Phi(\alpha, \beta)$, кроме $\alpha$ и $\beta$ зависит ещё от внешних параметров $\lambda_1, \lambda_2, ..., \lambda_s$. Мы имеем:

$$\ln F(\alpha, \beta) = \alpha N + \beta E + \ln \Phi(\alpha, \beta),$$

и следовательно (ввиду неизменности $N$),

$$d \ln F(\alpha, \beta) = N d\alpha + E d\beta + \beta dE + d \ln \Phi(\alpha, \beta) =$$

$$= N d\alpha + E d\beta + \beta dE + \frac{\partial \ln \Phi}{\partial \alpha} d\alpha + \frac{\partial \ln \Phi}{\partial \beta} d\beta + \sum_{i=1}^{s} \frac{\partial \ln \Phi}{\partial \lambda_i} d\lambda_i.$$  

Учитывая соотношения

$$\frac{\partial \ln \Phi}{\partial \alpha} = -N, \quad \frac{\partial \ln \Phi}{\partial \beta} = -E$$

и формулу (5) § 2, мы отсюда находим:

$$d \ln F(\alpha, \beta) = \beta (dE + D\overline{W}).$$

(6)

Согласно представлениям классической термодинамики элементарное изменение $dE$ энергии системы, вызванное соответствующими изменениями температуры и внешних параметров, слагается из произведённой над системой внешними телами
работы, которая, очевидно, равна \( \delta W \), и полученного системой количества теплоты \( \delta Q \):

\[
dE = - \delta W + \delta Q.
\]

Само собою разумеется, что величина \( \delta Q \), так же как и \( \delta W \), не представляет собою полного дифференциала какой-либо функции параметров \( a, \beta, \lambda, \gamma \). Но соотношение (6), которое может быть переписано в виде

\[
d \ln F(a, \beta) = \beta \delta Q,
\]

очевидно, показывает, что произведение \( \beta \delta Q \) представляет собою такой дифференциал, т. е. что \( \beta \) служит для величины \( \delta Q \) интегрирующим множителем.

Одна из наиболее удобных и употребительных формулировок второго закона термодинамики гласит: существует такая функция \( S \) температуры и внешних параметров и такая другая функция \( \theta \), зависящая только от температуры системы, что

\[
\theta \delta Q = dS;
\]

ничто говоря: величина \( \delta Q \) допускает интегрировать множитель, зависящий только от температуры системы. Соотношение (7) совпадает с соотношением (8), если положить

\[
\theta = \beta, \quad S = \ln F(a, \beta).
\]

Для функции \( \theta \) термодинамика даёт выражение \( \frac{1}{kT} \), где \( T \) — абсолютная температура системы, а \( k \) — так называемая постоянная Больцмана. Поэтому в статистической физике всегда определяют физический смысл параметра \( \beta \) универсальной формулой

\[
\beta = \frac{1}{kT}.
\]

Функцию \( S \) называют энтропией системы; это — одна из важнейших величин термодинамической теории. Соотношение (7) получает вид

\[
\frac{\delta Q}{T} = k dS
\]

и становится прямым выражением второго закона термодинамики, который, таким образом, является непосредственным
следствием нашей статистической теории. В частности, функция \[ F(\alpha, \beta), \] введённая нами в гл. V в качестве вспомогательного математического инструмента, приобретает теперь важнейший физический смысл — величина \[ \ln \Phi(\alpha, \beta) \] выражает собою энтропию данной системы. Наш специальный выбор значений параметров \( \alpha \) и \( \beta \), как дающих функции \( F(\alpha, \beta) \) наименьшее значение, мы рассматривали в гл. V как чисто математический приём, имеющий целью придать возможно более простую форму получаемым асимптотическим выражениям. Теперь этот выбор ведёт к ряду важных физических следствий, одно из которых мы для примера сейчас приведём.

Пусть мы имеем две системы, которые первоначально полностью изолированы как друг от друга, так и от окружающего мира; будем отмечать соответственно индексами 1 и 2 величины, относящиеся к первой и второй системе, и оставлять без индексов величины, относящиеся к той системе, которая получится после соединения двух данных в одну и установления в этой последней теплового равновесия. Очевидно, мы будем иметь:

\[ N = N_1 + N_2, \quad F = E_1 + E_2, \quad \Phi(\alpha, \beta) = \Phi_1(\alpha, \beta) \Phi_2(\alpha, \beta), \]

откуда энтропия суммарной системы

\[ S = N \alpha + E \beta + \ln \Phi(\alpha, \beta) = \\
= [N_1 \alpha + E_1 \beta + \ln \Phi_1(\alpha, \beta)] + [N_2 \alpha + E_2 \beta + \ln \Phi_2(\alpha, \beta)]. \quad (9) \]

Но функция

\[ \ln F_1(\alpha, \beta) = N_1 \alpha + E_1 \beta + \ln \Phi_1(\alpha, \beta) \]

имеет наименьшее значение при \( \alpha = \alpha_1, \beta = \beta_1 \), а функция \( \ln F_2(\alpha, \beta) \) — при \( \alpha = \alpha_2, \beta = \beta_2 \) (ибо числа \( \alpha_1, \beta_1, \alpha_2, \beta_2 \) по нашему общему соглашению выбираются как раз в согласии с этим требованием). Поэтому соотношение (9) даёт:

\[ S \geq \ln F_1(\alpha_1, \beta_1) + \ln F_2(\alpha_2, \beta_2) = S_1 + S_2, \]

g. е. если две системы, первоначально изолированные друг от друга, затем привести в тепловой контакт, то после установления равновесия энтропия суммарной системы будет либо больше, либо в крайнем случае равна сумме тех энтропий, которыми обладали составляющие системы до своего соединения.
Наступление равновесия при этом характеризуется тем, что параметры $a$ и $\beta$ получают общие для обеих составляющих значения. Заметим ещё, что в доказанном нами соотношении $S \geq S_1 + S_2$ знак равенства, очевидно, будет иметь место тогда и только тогда, когда $\alpha_1 = \alpha_2 = \alpha$, $\beta_1 = \beta_2 = \beta$, т. е. если обе системы до соединения имели одинаковые температуры.

Вспомним теперь, что закон сохранения энергии, являющийся первым законом термодинамики, нами установлен еще в конце гл. II, в той естественной форме, какую он должен иметь в квантовой механике. Этот закон, в противоположность второму, не нуждается для своего обоснования в статистических методах, а является, как и в классической механике, прямым следствием общих законов эволюции физических систем (т. е. в случае квантовой физики следствием уравнения Шредингера).

Так как на основе двух основных законов, полностью нами теперь обоснованных, термодинамика может быть построена уже чисто дедуктивно, независимо от каких-либо специальных представлений о строении вещества, то статистическая теория может на этом считать свою обосновывающую задачу завершённой.
ДОПОЛНЕНИЕ I

СТАТИСТИКА НЕОДНОРОДНЫХ СИСТЕМ

Во всём тексте книги изучаемые системы предполагались однородными, т. е. состоящими из частиц одинаковой структуры. Это имело своей целью такое упрощение формального аппарата, которое дало бы читателю возможность сосредоточиться на идеальных основах метода. Развитый нами метод позволяет, однако, без каких-либо принципиальных изменений охватить и статистику неоднородных систем, состоящих из частиц нескольких различных типов. Вызвываемые этим переходом усложнения расчётных формул носят чисто технический характер; в основном различие состоит лишь в том, что вместо двумерных предельных теорем теории вероятностей приходится применять аналогичные многомерные теоремы, формулировки, доказательства и условия применения которых полностью соответствуют установленным нами в гл. I одно- и двумерным случаям. Как правило, для системы, состоящей из частиц $k$ различных типов, приходится при этом пользоваться $k+1$-мерной предельной теоремой. Во всём остальном метод ничем не отличается от того, который мы применяли в простейшем случае однородных систем. Частицы различной структуры, из которых состоит данная неоднородная система, могут при этом подчиняться различным типам статистики.

В настоящем дополнительном параграфе мы кратко наметим путь исследований для систем, состоящих из материальных частиц двух различных типов; переход от двух к трём и более типам уже совсем тривиален и связан лишь с удлинением выписываемых формул.
Итак, пусть изучаемая нами система состоит из $N$ частиц двух различных типов; как обычно, мы допускаем, что эти частицы (в частности, и разнотипные) могут свободно обмениваться энергией; вместе с тем энергия взаимодействия частиц предполагается настолько незначительной, чтобы мы имели возможность при всех произвольных расчетах принимать полную энергию системы равной сумме полных энергий всех составляющих её частиц.

Пусть числа частиц первого и второго типов равны соответственно $N_1$ и $N_2$ ($N_1 + N_2 = N$). Обозначим возможные уровни энергии для частиц первого типа (в обычном расположении) через $\varepsilon_1, \varepsilon_2, \ldots$, а для частиц второго типа — через $\eta_1, \eta_2, \ldots$ (так что оба спектра предполагаются имеющими тот дискретный тип, рассмотрением которого мы ограничивались всюду в тексте). Условимся для краткости называть совокупность входящих в данную систему частиц первого (соответственно второго) типа — первой (соответственно второй) компонентой этой системы. Пусть $U_1, U_2, \ldots$ — полная ортогональная система «достигаемых» (т. е. хорошо известным нам образом согласованных с типом статистики) основных собственных функций оператора энергии первой компоненты, а $V_1, V_2, \ldots$ — аналогичная система для второй компоненты. В силу общих результатов гл. III система функций $U_i V_j (i, j = 1, 2, \ldots$) будет тогда полной ортогональной системой основных собственных функций оператора энергии для всей данной системы; мы совершенно естественно расширяем на нашу неоднородную систему понятие структурной функции, называя этим именем число $\Omega(N_1, N_2, E)$ собственных функций вида $U_i V_j$, принадлежащих к собственному значению $E$ оператора энергии данной системы. Так как энергия системы равна сумме энергий её компонент, то, обозначая через $\Omega_1 (N_1, E)$ и $\Omega_2 (N_2, E)$ (определенные обычным образом) структурные функции этих компонент, мы, очевидно, имеем:

$$\Omega(N_1, N_2, E) = \sum_{x=0}^{\infty} \Omega_1 (N_1, x) \Omega_2 (N_2, E - x).$$

(1)

Пусть теперь известно, что система находится в некотором определённом состоянии $U_i V_j$: это означает, что первая компонента находится в состоянии $U_i$, а вторая — в состоян-
нии \( V_j \); так как состояния \( U_i \) и \( V_j \) — основные, то при этом получают определённые значения все \( \text{"числа заполнения"}, \) т. е. для любых \( r \) и \( s \) в точности определено число \( a_r \) частиц первой компоненты, находящихся в состоянии с уровнем энергии \( \varepsilon_r \), и число \( b_s \) частиц второй компоненты, связанных по-добным же образом с уровнем \( \eta_s \); числа \( a_r \) и \( b_s \) при этом, очевидно, всегда связаны соотношениями

\[
\sum_{r=1}^{\infty} a_r = N_1, \quad \sum_{s=1}^{\infty} b_s = N_2, \quad \sum_{r=1}^{\infty} a_r\varepsilon_r + \sum_{s=1}^{\infty} b_s\eta_s = E. \tag{K}
\]

Обратно, если задан определённый набор чисел заполнения \( a_r \gg 0, b_s \gg 0 \), удовлетворяющих соотношениям (K), то этому набору, вообще говоря, соответствует определённое число состояний вида \( U_iV_j \), данной системы. По аналогии с тем, что мы имели в § 5 гл. III, легко показать, что это число равно

\[
C_1(N_1)C_2(N_2)\prod_{r=1}^{\infty} \gamma_1(a_r)\prod_{s=1}^{\infty} \gamma_2(b_s);
\]

(2)

вводимые при этом обозначения легко понятны в силу полной аналогии с предыдущим: \( C_1(N_1) = N_1! \) или \( = 1 \) в зависимости от того, подчиняется ли первая компонента полной или какой-либо из двух других статистик; \( C_2(N_2) \) имеет аналогичное значение для второй компоненты; функции \( \gamma_1(a_r) \) и \( \gamma_2(b_s) \) определяются соответственно для первой и второй компонент в полной аналогии с нашей прежней функцией \( \gamma(a) \) (и следовательно, также в зависимости от типа статистики, которой подчиняется соответствующая компонента).

Для того чтобы получить общее число \( \Omega(N_1, N_2, E) \) собственных функций вида \( U_iV_j \), принадлежащих к собственному значению \( E \) оператора энергии системы, мы, очевидно, должны просуммировать выражение (2) по всем возможным наборам чисел \( a_r \gg 0, b_s \gg 0 \), удовлетворяющим соотношениям (K). Таким образом, мы находим:

\[
\Omega(N_1, N_2, E) = C_1(N_1)C_2(N_2)\sum_{(K)} \left\{ \prod_{r=1}^{\infty} \gamma_1(a_r) \prod_{s=1}^{\infty} \gamma_2(b_s) \right\}. \tag{3}
\]

Это первичное выражение структурной функции вполне аналогично формуле (14) § 5 гл. III (стр. 137). И здесь, как там, оно служит исходным пунктом для всех дальнейших
расчётов. В частности, оно позволяет, по аналогии с тем, что было сделано в § 2 гл. V, выразить средние значения чисел $a_r$ и $b_s$, а также их попарных произведений, через отношения различных значений структурной функции $Q(N_1, N_2, E)$. Мы не будем проделывать этого здесь, так как в случае необходимости читатель может без затруднений самостоятельно установить соответствующие формулы.

Напротив, мы остановимся несколько подробнее на втором этапе исследования (для однородных систем проведённом в § 3 гл. V), сводящем асимптотическую оценку структурной функции к предельной задаче теории вероятностей.

Нам нужно будет ввести два вырожденных двумерных закона распределения $u_k$ и $v_l$, где $k$ и $l$ — произвольные натуральные числа; при этом мы введём ещё три параметра $a_1$, $a_2$ и $b$, значения которых, точно определяемые в дальнейшем, должны быть таковы, чтобы все вводимые нами в рассмотрение ряды были абсолютно сходящимися. Случайный вектор $(x, y)$, распределённый по закону $u_k$, имеет своими возможными точками лишь точки вида $x = n, y = nk (n = 0, 1, \ldots)$, причём

$$P(x = n, y = nk) = \frac{\gamma_1(n) e^{-n(a_1 + 3k)}}{\sum_{m=0}^{\infty} \gamma_1(m) e^{-m(a_1 + 3k)}} (n = 0, 1, \ldots).$$

Аналогично, для закона $v_l$ возможными точками служат точки вида $x = n, y = nl (n = 0, 1, \ldots)$, причём

$$P(x = n, y = nl) = \frac{\gamma_2(n) e^{-n(a_2 + 3l)}}{\sum_{m=0}^{\infty} \gamma_2(m) e^{-m(a_2 + 3l)}} (n = 0, 1, \ldots).$$

Обозначим соответственно через $g_{1k}$ и $g_{2l}$ кратности (степени вырождения) уровней энергии $e_r = k$ и $e_{ls} = l$ для частиц первого и второго типов, в предположении, что занимаемый системой объём $V = 1$. Рассмотрим сумму бесконечного ряда взаимно независимых случайных векторов $(x_{1i}, y_{1i}) (i = 1, 2, \ldots)$, среди которых имеется $g_{1k}$ векторов, распределённых по закону $u_k (k = 1, 2, \ldots)$, и положим

$$\sum_{i=1}^{\infty} x_{1i} = X_1, \quad \sum_{i=1}^{\infty} y_{1i} =: Y_1;$$
как в § 3 гл. V, легко показать, что оба ряда сходятся с вероятностью 1. Подобным же образом пусть \((x_{2i}, y_{2i}) (i=1, 2, \ldots)\) — последовательность взаимно независимых случайных векторов, среди которых имеется \(g_{2i}\) векторов, распределённых по закону \(v_l (l=1, 2, \ldots)\), и пусть
\[
\sum_{i=1}^{\infty} x_{2i} = X_2, \quad \sum_{i=1}^{\infty} y_{2i} = Y_2.
\]

Положим теперь \(Y_1 + Y_2 = Y\) и обозначим через \(P\) закон распределения трёхмерного вектора \((X_1, X_2, Y)\). Очевидно, что этот закон зависит только от природы составляющих данную систему частиц и от параметров \(\alpha_1, \alpha_2\) и \(\beta\); напротив, он независим от чисел \(N_1, N_2, V, E\) (если эти числа, как мы и здесь будем предполагать, изменяются так, что отношения их сохраняют постоянные значения).

Пусть, наконец, мы имеем сумму
\[
(S_{1V}, S_{2V}, T_V)
\]
(4)

\(V\) взаимно независимых трёхмерных случайных векторов, каждый из которых распределён по закону \(P\), только что нами описанному. Тогда, очевидно, можно положить
\[
S_{1V} = \sum_{k=1}^{\infty} \sum_{i=1}^{k} x_{1ki}, \quad S_{2V} = \sum_{l=1}^{\infty} \sum_{i=1}^{l} x_{2li},
\]
\[
T_V = \sum_{k=1}^{\infty} \sum_{i=1}^{k} y_{1ki} + \sum_{l=1}^{\infty} \sum_{i=1}^{l} y_{2li};
\]
здесь вектор \((x_{1ki}, y_{1ki})\) подчинён закону \(u_k\), вектор \((x_{2li}, y_{2li})\) — закону \(v_l\), и все эти элементарные векторы между собой независимы. Мы находим поэтому для закона распределения трёхмерного вектора (4)
\[
P (S_{1V} = p_1, \quad S_{2V} = p_2, \quad T_V = q) =
\]
\[
= \sum_{(k, p, p, q)} \prod_{k=1}^{\infty} \prod_{i=1}^{k} P (x_{1ki} = a_{ki}, \quad y_{1ki} = ka_{ki}) \times
\]
\[
\times \prod_{l=1}^{\infty} \prod_{i=1}^{l} P (x_{2li} = b_{li}, \quad y_{2li} = lb_{li})
\]
(5)
где суммирование распространяется на всевозможные системы чисел \( a_{ki}, b_{li} \), удовлетворяющие соотношениям

\[
\sum_{k=1}^{\infty} \sum_{i=1}^{\infty} a_{ki} = p_1, \quad \sum_{l=1}^{\infty} \sum_{i=1}^{\infty} b_{li} = p_2, \\
\sum_{k=1}^{\infty} \sum_{i=1}^{\infty} k a_{ki} + \sum_{l=1}^{\infty} \sum_{i=1}^{\infty} l b_{li} = q.
\]

(\( K_{p,p,q} \))

Но по самому определению законов \( u_k \) и \( v_l \), мы имеем:

\[
P (x_{1ki} = a_{ki}, \quad y_{1ki} = ka_{bi}, \quad \Gamma_1 (k) \gamma_1 (a_{bi}) e^{-a_{ki} (a_1 + \beta k)}, \\
P (x_{2li} = b_{li}, \quad y_{2li} = lb_{li}, \quad \Gamma_2 (l) \gamma_2 (b_{li}) e^{-b_{li} (a_2 + \beta l)}),
\]

где

\[
\Gamma_1 (k) = \left\{ \sum_{m=0}^{\infty} \gamma_1 (m) e^{-m (a_1 + \beta k)} \right\}^{-1}, \\
\Gamma_2 (l) = \left\{ \sum_{m=0}^{\infty} \gamma_2 (m) e^{-m (a_2 + \beta l)} \right\}^{-1}.
\]

Вставляя эти выражения в правую часть соотношения (5), мы находим:

\[
P (S_{1V} = p_1, \quad S_{2V} = p_2, \quad T_V = q) = \\
= \prod_{k=1}^{\infty} \{ \Gamma_1 (k) \} V_{g_{1k}} \prod_{l=1}^{\infty} \{ \Gamma_2 (l) \} V_{g_{2l}} \sum_{(K_{p,p,q})} \prod_{k=1}^{\infty} \prod_{i=1}^{\infty} \gamma_1 (a_{ki}) \times \\
\times \prod_{l=1}^{\infty} \prod_{i=1}^{\infty} \gamma_2 (b_{li}) \exp \left\{ -a_1 \sum_{k=1}^{\infty} \sum_{i=1}^{\infty} a_{ki} - a_2 \sum_{l=1}^{\infty} \sum_{i=1}^{\infty} b_{li} - \\
\beta \left( \sum_{k=1}^{\infty} \sum_{i=1}^{\infty} k a_{ki} + \sum_{l=1}^{\infty} \sum_{i=1}^{\infty} l b_{li} \right) \right\} = \\
e^{-a_1 p_1} a_2 p_2 \beta q \prod_{k=1}^{\infty} \{ \Gamma_1 (k) \} V_{g_{1k}} \prod_{l=1}^{\infty} \{ \Gamma_2 (l) \} V_{g_{2l}} \times \\
\times \sum_{(K_{p,p,q})} \prod_{k=1}^{\infty} \prod_{i=1}^{\infty} \gamma_1 (a_{ki}) \prod_{l=1}^{\infty} \prod_{i=1}^{\infty} \gamma_2 (b_{li}).
\]

Последняя сумма в правой части этого соотношения только системой обозначений отличается от суммы

\[
\sum_{(K)} \left\{ \prod_{r=1}^{\infty} \gamma_1 (a_r) \prod_{s=1}^{\infty} \gamma_2 (b_s) \right\},
\]
стоящей в правой части формулы (3). Таким образом,

\[ P(S_{1V} = p_1, S_{2V} = p_2, T_V = q) = \]

\[ = \prod_{k=1}^{\infty} \{\Gamma_1(k)\} V_{g_{1k}} \prod_{l=1}^{\infty} \{\Gamma_2(l)\} V_{g_{2l}} e^{-a_1 p_1 - a_3 p_3 - \beta q} \frac{\Omega(p_1, p_2, q)}{C_1(p_1) C_2(p_2)}. \] (6)

Суммируя это равенство по всем целым \( p_1, p_2, q, \) мы находим:

\[ 1 = \prod_{k=1}^{\infty} \{\Gamma_1(k)\} V_{g_{1k}} \prod_{l=1}^{\infty} \{\Gamma_2(l)\} V_{g_{2l}} \sum_{p_1, p_2, q} \frac{e^{-a_1 p_1 - a_3 p_3 - \beta q}}{C_1(p_1) C_2(p_2)} \Omega(p_1, p_2, q). \]

Сумма в правой части этого равенства есть функция параметров \( a_1, a_2 \) и \( \beta, \) которую мы обозначим через \( \Phi(a_1, a_2, \beta), \) так что

\[ \prod_{k=1}^{\infty} \{\Gamma_1(k)\} V_{g_{1k}} \prod_{l=1}^{\infty} \{\Gamma_2(l)\} V_{g_{2l}} = \{\Phi(a_1, a_2, \beta)\}^{-1}, \]

и соотношение (6) даёт:

\[ P(S_{1V} = p_1, S_{2V} = p_2, T_V = q) = \]

\[ = \{\Phi(a_1, a_2, \beta)\}^{-1} e^{-a_1 p_1 - a_3 p_3 - \beta q} \frac{\Omega(p_1, p_2, q)}{C_1(p_1) C_2(p_2)}, \]

откуда

\[ \Omega(p_1, p_2, q) = \]

\[ = \Phi(a_1, a_2, \beta) C_1(p_1) C_2(p_2) e^{z_1 p_1 + z_3 p_3 + \beta q} P(S_{1V} = p_1, S_{2V} = p_2, T_V = q). \]

Эта формула, в точности аналогичная формуле (22) § 3 гл. V, полностью решает задачу сведения асимптотической оценки структурной функции \( \Omega(p_1, p_2, q) \) к предельным теоремам теории вероятностей, так как вектор \( (S_{1V}, S_{2V}, T_V) \) определяется как сумма \( V \) взаимно независимых, одинаково распределённых трёхмерных случайных векторов. Стоит отметить, что неоднородность нашей системы имеет, таким образом, своим следствием только увеличение числа измерений соответствующей задачи теории вероятностей, в то время как суммировать приходится попрежнему лишь одинаково распределённые случайные векторы.

Последнего этапа — применения (трёхмерной) локальной предельной теоремы — мы снова можем здесь не рассматри-
вать, так как проведение его по существу ничем не отличается от того, что мы имели в двумерном случае. Укажем только, что значения параметров $\alpha_1$, $\alpha_2$, $\beta$ определяются из соотношений

$$\frac{\partial \ln \Phi}{\partial \alpha_1} = -N_1, \quad \frac{\partial \ln \Phi}{\partial \alpha_2} = -N_2, \quad \frac{\partial \ln \Phi}{\partial \beta} = -E,$$

существование и единственность решения которых могут быть доказаны аналогично предыдущему. Так как $\ln \Phi (\alpha_1, \alpha_2, \beta)$ (а значит, и все производные этой функции) пропорционален $V$, а отношения $\frac{N_1}{V}$, $\frac{N_2}{V}$, $\frac{E}{V}$ предполагаются неизменными, то параметры $\alpha_1$, $\alpha_2$, $\beta$ получают при этом постоянные значения.

В качестве основных результатов теории отметим, что средние значения чисел заполнения получают выражения:

$$\bar{a}_r = \frac{1}{e^{\alpha_1 + \beta \varepsilon_r - \sigma_1}} + O\left(\frac{1}{V}\right),$$

$$\bar{b}_s = \frac{1}{e^{\alpha_2 + \beta \eta_s - \sigma_2}} + O\left(\frac{1}{V}\right),$$

где $\sigma_1$ и $\sigma_2$ соответственно означают индексы симметрии первой и второй компонент. Энергия $E_1$ первой компоненты здесь уже не сохраняет, разумеется, постоянного значения, а является фазовой функцией системы, и притом, очевидно, суммарной величиной, среднее значение которой может быть непосредственно написано, если известны средние значения чисел заполнения: мы находим с хорошим приближением

$$\bar{E}_1 \approx \sum_{r=1}^{\infty} \frac{\varepsilon_r}{e^{\alpha_1 + \beta \varepsilon_r - \sigma_1}},$$

и аналогично для второй компоненты

$$\bar{E}_2 \approx \sum_{s=1}^{\infty} \frac{\eta_s}{e^{\alpha_2 + \beta \eta_s - \sigma_2}}.$$
ДОПОЛНЕНИЕ II

РАСПРЕДЕЛЕНИЕ КОМПОНЕНТЫ И ЕЁ ЭНЕРГИИ

1

Пусть данная система состоит из двух компонент, для которых мы сохраним терминологию и всю систему обозначений предшествующего дополнения. Состояния такой системы, соответствующие уровню энергии $E$, образуют линейное многообразие, базисом которого может служить система собственных функций вида $U_i V_j$; здесь $U_i$ — одна из основных собственных функций первой компоненты, принадлежащая к некоторому уровню энергии $E_1$; аналогично $V_j$ — одна из основных собственных функций второй компоненты, принадлежащая к некоторому уровню энергии $E_2$; допустимыми комбинациями функций $U_i$ и $V_j$ являются такие, для которых $E_1 + E_2 = E$. Как мы видели в предшествующем дополнении [формула (1)], эти соображения позволяют непосредственно написать важное соотношение

$$Q(N_1, N_2, E) = \sum_{x=0}^{\infty} Q_1(N_1, x) Q_2(N_2, E - x),$$

(1)

которое будет служить исходной точкой наших дальнейших расчётов.

Прежде всего найдём число таких основных функций $U_i V_j$ нашей системы, принадлежащих к уровню $E$ её энергии, в которых первый индекс $i$ имеет некоторое определённое значение, т. е. первая компонента находится в некотором фиксированном основном состоянии $U_i$; если это состояние соответствует уровню энергии $E_1 = E_1(U_i)$ первой компоненты,
то искомое число есть, очевидно, число функций \( V_j \), соответствующих уровню энергии \( E - E_1 \), т. е.

\[
\Omega_2 (N_2, E - E_1) = \Omega_2 [N_2, E - E_1 (U_i)].
\]

Пусть теперь \( \Omega \) есть связанная с данной системой механическая величина, имеющая определённое значение в каждом из основных состояний \( U_i V_j \) этой системы (фазовая функция), и пусть эта величина такова, что значение её полностью определяется состоянием \( U_i \) первой компоненты, будучи независимым от состояния \( V_j \) второй компоненты, так что

\[
\Omega = f(U_i).
\]

Тогда эта величина сохраняет одно и то же значение \( f(U_i) \) для всех тех \( \Omega_2 [N_2, E - E_1 (U_i)] \) основных состояний \( U_i V_j \) данной системы, в которых первый множитель равен \( U_i \) (т. е. первая компонента находится в состоянии \( U_i \)). Поэтому микроканоническое среднее такой величины может быть записано в виде

\[
\overline{\Omega} = \frac{1}{\Omega (N_1, N_2, E)} \sum_{i=1}^{\infty} f(U_i) \Omega_2 [N_2, E - E_1 (U_i)]. \tag{2}
\]

Это соотношение, будучи почти самоочевидным, тем не менее имеет глубокий смысл и, как мы увидим дальше, важные следствия. Прежде всего оно показывает, что для величины, зависящей от состояния одной только первой компоненты, микроканоническое осреднение (которое всегда есть осреднение по всем основным функциям \( U_i V_j \) нашей полной системы при уровне энергии \( E \)) может быть заменено некоторым осреднением по основным собственным функциям \( U_i \) первой компоненты. Отметим сейчас же, что это «редуцированное» осреднение имеет черты, резко отличивающие его от микроканонического осреднения: в микроканоническом осреднении всегда участвуют только состояния с одним и тем же фиксированным уровнем энергии, в то время как осреднение (2) производится по всем основным собственным функциям \( U_i \) первой компоненты, к каким бы уровням энергии они ни принадлежали; далее, в микроканоническом осреднении все участвую-
щие в нём состояния получают равные веса; напротив, в осреднении (2) основное состояние \( U_i \) имеет вес

\[
\frac{\Omega_2 [N_2, E - E_1 (U_i)]}{\Omega (N_1, N_2, E)},
\]

зависящий от значения \( E_1 (U_i) \) энергии первой компоненты в состоянии \( U_i \), и потому различный для различных основных состояний \( U_i \) этой первой компоненты.

Рассмотрим теперь важный, ещё более частный случай, когда величина \( \mathcal{A} \) зависит только от энергии \( E_1 \) первой компоненты, т. е. для всех состояний \( U_i \), соответствующих одному и тому же уровню энергии \( E_1 \), получает одно и то же значение

\[
\mathcal{A} = \varphi (E_1) = \varphi [E_1 (U_i)].
\]

Тогда в формуле (2) \( f(U_i) = \varphi (x) \) для всех слагаемых, для которых \( E_1 (U_i) = x \); но число таких слагаемых есть число основных собственных функций \( U_i \) первой компоненты, соответствующих уровню энергии \( x \), т. е. \( \Omega_1 (N_1, x) \). Таким образом, формула (2) приводится к виду

\[
\mathcal{A} = \varphi (E_1) = \frac{1}{\Omega (N_1, N_2, E)} \sum_{x=0}^{\infty} \varphi (x) \Omega_1 (N_1, x) \Omega_2 (N_2, E - x). \quad (3)
\]

В частности, микроканоническое среднее энергии первой компоненты

\[
\overline{E_1} = \frac{1}{\Omega (N_1, N_2, E)} \sum_{x=0}^{\infty} x \Omega_1 (N_1, x) \Omega_2 (N_2, E - x). \quad (4)
\]

Полученные нами формулы (2), (3) и (4) являются вполне точными; однако для конкретных расчётов они совершенно недоступны ввиду сложности выражений функций \( \Omega_1 \), \( \Omega_2 \) и \( \Omega \). Поэтому мы должны для практических вычислений всякий раз заменять их более простыми приближёнными формулами, которые легко могут быть получены с помощью изложенного в этой книге асимптотического метода. Мы теперь рассмотрим несколько важнейших приложений этого метода. При этом мы в большинстве случаев ограничимся выводом асимптотических выражений, не останавливаясь на детальной оценке остаточных членов.
Для получения нужных нам асимптотических выражений мы прежде всего обратимся к формуле (39) § 6 гл. V. Если заменить в ней \( \Omega \) на \( \Omega_2 \), \( N \) на \( N_2 \) и положить \( u_1 = 0 \), \( u_2 = -x \), то мы найдём:

\[
\frac{\Omega_2(N_2, E-x)}{\Omega_2(N_2, E)} = e^{-\beta x}\left\{1 + O\left(\frac{x^2}{V}\right)\right\}.
\]

Применяя эту же оценку к весовой функции

\[
\frac{\Omega_2[N_2, E - E_1(U_i)]}{\Omega(N_1, N_2, E)} = \frac{\Omega_2[N_2, E - E_1(U_i)]}{\sum_{x=0}^{\infty} \Omega_1(N_1, x) \Omega_2(N_2, E-x)}
\]

формулы (2), мы легко находим для неё очень простое асимптотическое выражение

\[
e^{-\beta E_1(U_i)} \frac{\sum_{x=0}^{\infty} \Omega_1(N_1, x) e^{-\beta x}}{\sum_{x=0}^{\infty} \Omega_1(N_1, x) e^{-\beta x}} + O\left(\frac{1}{V}\right);
\]

таким образом, формула (2) может быть заменена приближённой формулой

\[
\bar{\Omega} = f(U_i) \approx \sum_{i=1}^{\infty} \frac{\sum_{x=0}^{\infty} \Omega_1(N_1, x) e^{-\beta x}}{f(U_i) e^{-\beta E_1(U_i)}}.
\]

Приведённый нами вывод этой формулы очень прост, но содержит одну неточность, которая в общем случае делает формулу (5) неверной. В самом деле, в формуле (39) § 6 гл. V, на которой был основан наш расчёт, \( E \) означает фиксированную полную энергию той системы, структурной функцией которой служит \( \Omega(N, E) \); у нас же \( E \) означает фиксированное значение энергии нашей полной системы, в то время как \( \Omega_2 \) есть структурная функция второй компоненты. Мы будем, однако, рассматривать только тот случай, когда первая компонента ничтожно мала сравнительно со второй; говоря точно, мы будем строить наши асимптотические формулы в предположении, что числа \( N_2, E \) и \( V \) безгранично возрастают, сохраняя постоянные отношения, в то время как число \( N_1 \)
остаётся постоянным. Для приложений как раз наиболее интересен случай, когда \( N_1 = 1 \), т. е. первая компонента представляет собой отдельную частицу. В этом случае энергия \( E_2 \) второй компоненты асимптотически совпадает с энергией \( E \) полной системы и как нетрудно показать более подробным расчётом, вызываемая упомянутой нами неточностью погрешность не уменьшает той степени точности, с которой выполняется равенство (5).

Условимся говорить в только что описанном случае, что первая компонента является малой компонентой данной системы. Таким образом, для величин, определяемых состоянием такой малой компоненты (в частности, отдельной частицы), микроканонические средние приближённо могут быть получены путём осреднения по всем основным функциям этой компоненты с весовой функцией

\[
\frac{e^{-\beta E_1(U_i)}}{\sum_{x=0}^{\infty} \Omega_1(N_1, x) e^{-\beta x}}. \tag{6}
\]

Этот результат в статистической физике называют законом Больцмана. При этом обычно трактуют весовую функцию (6) как вероятность найти малую компоненту в состоянии \( U_i \). Против такой терминологии ничего нельзя было бы возразить, если бы термин «вероятность» в дальнейшем развитии теории всегда сохранял то значение, которое он получил по своему определению. Опыт показал, однако, что на почве этой терминологии, как правило, быстро вырастает большое число смешений и неотчётливых формулировок. Поэтому мы и здесь, как на всём протяжении книги, будем сознательно избегать вероятностной терминологии, тем более, что в ней не имеется настоятельной надобности.

С помощью выражения (6) для весовой функции формулы (3) и (4) получают соответственно асимптотические выражения

\[
\overline{\Omega} = \varphi(E_1) \approx \frac{\sum_{x=0}^{\infty} \varphi(x) \Omega_1(N_1, x) e^{-\beta x}}{\sum_{x=0}^{\infty} \Omega_1(N_1, x) e^{-\beta x}}. \tag{7}
\]
и

$$\overline{E_1} \approx \frac{\sum_{x=0}^{\infty} x \Omega_1 (N_1, x) e^{-\beta x}}{\sum_{x=0}^{\infty} \Omega_1 (N_1, x) e^{-\beta x}}.$$  \hspace{1cm} (8)$$

Таким образом, микроканоническое среднее любой функции энергии малой компоненты (в частности, отдельной частицы) может быть приближённо получено осреднением по всем возможным значениям $x$ этой энергии с весовой функцией

$$\frac{\Omega_1 (N_1, x) e^{-\beta x}}{\sum_{x=0}^{\infty} \Omega_1 (N_1, x) e^{-\beta x}}.$$
ДОПОЛНЕНИЕ III

ПРИНЦИП КАНОНИЧЕСКОГО ОСРЕДНЕНИЯ

В этой книге мы везде предполагали исследуемую систему энергетически изолированной, т. е. не обменивающейся энергией с окружающими телами. Весь метод микроканонического осреднения построен на этой предпосылке, ибо в формировании микроканонических средних участвуют лишь состояния, в которых энергия системы имеет данное строго фиксированное значение. Это требование полной энергетической изоляции может, однако, в реальных условиях выполняться лишь приблизительно; более того, в очень многих практически важных случаях изучаемая система находится, наоборот, в более или менее интенсивном энергетическом (например, тепловом) контакте с окружающими телами, вследствие чего энергия её не сохраняет постоянного значения, и следовательно, принцип микроканонического осреднения лишается своего теоретического основания.

Результаты предшествующего дополнения II позволяют нам рационально подойти с точки зрения нашей статистической теории и к построению статистики таких неизолированных систем. Рассмотрим противоположный крайний случай, когда данная система может свободно обмениваться энергией с окружающим практически бесконечно большой величины (энергия которого во много раз превышает энергию самой системы). Такое окружение физики называют термостатом, а проданную систему говорят, что она «погружена в термостат». Теоретические рассмотрения приводят к выводу, что для определения режима такой погруженной в термостат системы имеет значение только температура термостата; остальные его свойства, включая и его материальный состав (т. е. при-
роду составляющих его частиц) никакой роли при этом не играют. В частности, ничто не мешает нам представлять себе терmostat в виде огромного числа физических систем, в точности подобных изучаемой нами и свободно обменивающихся энергией как между собою, так и с данной системой.

Но если мы станем на такую точку зрения, то мы можем рассматривать соединение данной системы S и термостата T как некую изолированную систему S + T, по отношению к которой наша система S и каждая из подобных ей систем, составляющих терmostat T, играют роль отдельных частиц. Так как система S + T изолирована, то для неё мы естественно принимаем оправданный всем предыдущим принцип микроканонического осреднения. А так как изучаемая нами система S, очевидно, может рассматриваться как малая компонента (отдельная частица) изолированной системы S + T, то микроканоническое осреднение по состояниям системы S + T равносильно, как мы видели, для системы S осреднению по формуле (2) дополнения II, где Ω₂ и Ω соответственно означают структурные функции систем T и S + T и где N₁ = 1, N₂ — число систем, подобных данной, составляющих терmostat, E — (постоянная) полная энергия суммарной системы S + T, E₁(U₁) — энергия данной системы в основном состоянии U₁, и суммирование ведётся по всем таким состояниям. Этот вывод является совершенно точным. Однако, как мы видели, формула (2) может быть с хорошим приближением заменена несравненно более простой и удобной формулой (5), где суммирование распространяется на ту же область и где Ω₁ — структурная функция данной системы; что касается параметра β, то, он, как мы знаем, универсально связан с абсолютной температурой системы соотношением

$$\beta = \frac{1}{kT},$$

где k — постоянная Больцмана (и где символ T абсолютной температуры, употребляемый нами здесь лишь мимоходом, не следует, конечно, смешивать с обозначением термостата). При этом, разумеется, температура данной системы и температура термостата одинаковы. Формула (5) очень ясно показывает, что выражаемый ею принцип осреднения для данной системы S действительно не носит никаких следов специальной при-
роды термостата и зависит (через посредство параметра $\beta$) лишь от той температуры, которая устанавливается как для данной системы, так и для термостата в результате энергетического контакта этих двух тел.

Мы видим, таким образом, что если мы принимаем для изолированной системы в качестве исходной базы статистических расчётов принцип микрокаанонического осреднения, то это уже с необходимостью влечёт за собой для системы, свободно обменивающейся энергией с большим окружением (термостатом), некоторый совершенно определённый принцип осреднения, приближённо выражаемый формулой (5) предшествующего дополнения:

$$\overline{X} = f(U_i) \approx \sum_{i=1}^{\infty} \sum_{x=0}^{\infty} \Omega(N, x) e^{-\beta x}$$

(где $N$ — число частиц, $\Omega$ — структурная функция, $E(U_i)$ — энергия в состоянии $U_i$ данной системы, и суммирование ведётся по всем основным состояниям системы; $\beta = \frac{1}{kT}$, $T$ — абсолютная температура системы).

Осреднение, проводимое согласно этому принципу, называют каноническим; основные отличия его от микрокаанонического осреднения состоят в том, что 1) в каноническом осреднении участвуют все основные состояния системы, а не только принадлежащие к определённому уровню энергии, как в микрокааноническом (разумеется, это соответствует реальному различию между системой, погруженной в термостат, и следовательно, способной менять свою энергию, и изолированной системой, энергия которой остаётся неизменной); 2) в каноническом осреднении, в противоположность микрокааноническому, веса различных основных состояний различные, вес того или другого основного состояния зависит, как мы непосредственно видим, от соответствующего уровня энергии, так что все основные состояния, принадлежащие к одному и тому же уровню энергии, получают одинаковые веса.

Само собою разумеется, что для системы, погруженной в термостат, имеют силу правила осреднения (7) и (8) предыдущего дополнения II, так как они являются непосредственными следствиями формулы (5) того же дополнения.
Принцип канонического осреднения имеет очень большие преимущества практического характера; проводимые с его помощью осреднения несравненно проще микроканонических. Поэтому многие авторы с самого начала в порядке постулат вводят этот принцип в качестве базы для всех статистических расчётов. При этом иногда делается ссылка на доказанную нами теорему о том, что микроканоническое осреднение для изолированных систем влечёт за собой каноническое осреднение, погружённых в термостат, и делается явное предупреждение, что в дальнейшем будет итти речь только о системах второго типа. В большинстве же случаев принцип канонического осреднения вводится в чисто постулативном порядке и в дальнейшем применяется ко всем изучаемым системам, независимо от того, будут ли они изолированными, погружёнными в термостат, или, как это чаще всего бывает в действительности, имеют некоторой промежуточный характер, лишь более или менее приближаясь к тому или другому из этих двух крайних типов. Именно так поступает, в частности, Гиббс, впервые введший в статистическую механику оба рассматриваемых нами принципа и являющийся автором их общепринятых наименований. Закономерна ли такая практика?
Чтобы ответить на этот вопрос, вспомним прежде всего, что и микроканонический принцип был нами с самого начала введен в качестве постулат, произвольность которого мы неоднократно подчеркивали. И хотя этот выбор в дальнейшем получил некоторое обоснование благодаря проведённому нами доказательству представительности даваемых им средних значений, всё же, как мы снова неоднократно подчеркивали, окончательную проверку его допустимости мог дать только опыт; и мы во всех подробностях должны были ознакомиться как раз с таким случаем (введение симметрического и антисимветрического принципов осреднения), когда этот выбор опытом был решительно опровергнут, и потребовалась замена его другим выбором. Но именно на такую точку зрения становятся Гиббс и его последователи, вводя принцип канонического осреднения. С помощью этого принципа Гиббс строит статистическую термодинамику, различные понятия которой затем отождествляются с соответствующими понятиями феноменологической термодинамики; если при этом оказывается,
что статистическая теория способна доказать основные положения феноменологической термодинамики, давно и хорошо проверенные на практике, то это — всё, что о нём требуется, и введённый постулат тем самым оправдан.

Между двумя основными принципами осреднения — каноническими и микроканоническими — имеются, однако, глубокие принципиальные различия. Пусть мы имеем дело с изолированной системой, энергия которой имеет точно фиксированное значение \( E \); это означает, что система с достоверностью находится в одном из состояний, принадлежащих к уровню энергии \( E \); привлекать при таких условиях к формированию средних значений механических величин такие состояния, уровни энергии которых отличны от \( E \), в которых поэтому система принципиально не может находиться, с теоретической точки зрения явно недопустимо, и практический успех принципа канонического осреднения по отношению к изолированным системам, если он имеет место, может возникнуть не благодаря такому привлечению, а лишь непосредственно на него. Другая особенность канонического принципа относится к системам любого типа. В то время как микроканоническое осреднение имеет в своей основе лишь допущение весьма общего характера (одинаковые веса для всех допустимых состояний), и потому бесспорно является наиболее простым и естественным из всех возможных, канонический принцип апеллирует к весовой функции весьма специального вида, и мы естественно хотим знать, почему здесь фигурирует именно эта функция, а не какая-либо другая из многих возможных, и не сохраняется ли в силе все выводы, если это специальное допущение заменить каким-либо более общим. При постулативном введении канонического принципа все эти естественно встающие вопросы просто отменяются; Гиббс говорит в пользу своего выбора только то, что с так выбранный весовой функцией очень удобно производить вычисления. Если же идти тем путём, который мы систематически провели в этой книге, приняв за основу для изолированных систем наиболее простой и естественный микроканонический принцип, вдохновляя его естественным образом там, где этого требуют теория и опыт (переход к «новым» статистикам!), и, наконец, строго доказав канонический принцип для систем, погруженных в термостат, то все сомнения и недоуменные вопросы,
отмеченные выше, полностью отпадают, и мы получаем полное теоретическое удовлетворение при тех же практических выводах. Мы полагаем, что ради этой цели стоит провести ту не-
сложную математическую работу, которую затратил читатель на овладение этой книгой.

Наконец, можно установить ещё и другую точку зрения на взаимоотношение между каноническим и микроканоническим осреднениями. Пользуясь микроканоническим принципом для изолированных физических систем, мы (в этом состояла как раз главная задача разливаемого метода) ввиду сложности выражений получаемых средних научились находить для них простые асимптотические выражения. Так как, с другой стороны, в большинстве случаев каноническое осреднение, как показало развитие теории, приводит к результатам, мало отличающимся от микроканонических средних, то можно смотреть на канонический принцип просто как на математический приём, позволяющий очень легко и единным методом находить приближённые значения для микроканонических средних. Как такой приём метод канонического осреднения, разумеется, вполне допустим, тем более, что мы всё равно, оперируя микроканоническими средними, вынуждены довольствоваться для них приближёнными выражениями. Однако, ставши на эту точку зрения, мы, естественно, должны поставить вопрос о том, при каких условиях канонические средние действительно могут служить приближёнными выражениями микроканонических и как велика может оказаться погрешность этих приближённых выражений. Мы кратко разберёмся теперь в этом вопросе для важнейшего случая сумматорных механических величин.

Обозначим через \( F(U) \) произвольную фазовую функцию данной системы и через \( \overline{F}_E \) микроканоническое среднее величины \( F(U) \) при значении \( E = x \) полной энергии системы (осреднение по многообразию \( \mathfrak{M}_x \)). Каноническое же среднее значение функции \( F(U) \) (при значении \( E \) энергии системы) согласно формуле (1) равно

\[
\overline{F}_E = \frac{\sum_{i=1}^{\infty} F(U_i) e^{-\beta E(U_i)}}{\sum_{x=0}^{\infty} \Omega(x) e^{-\beta x}},
\]
где мы ради краткости пишем $\Omega(x)$ вместо $\Omega(N, x)$ и где
значение параметра $\beta$ определяется хорошо известным нам
способом, когда даны число частиц $N$ и полная энергия $E$
системы. Но, очевидно, мы можем записать

$$\sum_{i=1}^{\infty} F(U_i) e^{-\beta E(U_i)} = \sum_{x=0}^{\infty} e^{-\beta x} \sum_{E(U_i) = x} F(U_i) =$$

$$= \sum_{x=0}^{\infty} \Omega(x) e^{-\beta x} \frac{\sum_{E(U_i) = x} F(U_i)}{\Omega(x)} = \sum_{x=0}^{\infty} \bar{F}_x \Omega(x) e^{-\beta x},$$

так как $\frac{1}{\Omega(x)} \sum_{E(U_i) = x} F(U_i)$ есть не что иное, как микрокано-
ническое среднее функции $F(U)$ при значении $x$ полной
энергии системы.

Таким образом, мы находим для канонического среднего
функции $F(U)$ выражение

$$\bar{F}_E = \frac{\sum_{x=0}^{\infty} \bar{F}_x \Omega(x) e^{-\beta x}}{\sum_{x=0}^{\infty} \Omega(x) e^{-\beta x}}.$$ 

Это означает, что каноническое среднее $\bar{F}_E$ фазовой функции
$F(U)$ представляет собой некоторое взвешенное среднее зна-
чение её микроканонических средних для всех возможных
уровней энергии, причём уровню $x$ приписывается вес

$$\nu(x) = \frac{\Omega(x) e^{-\beta x}}{\sum_{x=0}^{\infty} \Omega(x) e^{-\beta x}};$$

этот вес зависит, конечно, от того, какое значение $E$ имеет
полная энергия системы, так как от этого значения $E$ зави-
сит значение параметра $\beta$; нашей задачей будет теперь срав-
нить каноническое среднее $\bar{F}_E$ функции $F(U)$ с её микрокана-
оническим средним $\bar{F}_E$ при том же значении $E$ энергии
системы.
Весовая функция \( p(x) \) (где \( Q(x) \), как мы помним, означает \( Q(N, x) \)), в силу полученных нами в § 5 гл. V асимптотических формул приближённо выражается, как легко видеть, нормальным законом *)

\[
p(x) \approx \frac{1}{\sqrt{2\pi B}} e^{-\frac{(x - A)^2}{2B}},
\]

где, как в гл. V, мы имеем \( A = E \) в силу сделанного нами выбора значения параметра \( \beta \), а дисперсия \( B \), более детальный расчёт который нам здесь не нужен, есть бесконечно большая порядка величин \( E, V \) и \( N \). Таким образом, мы получаем:

\[
\overline{F}_E \approx \frac{1}{\sqrt{2\pi B}} \sum_{x = 0}^{\infty} F_x e^{-\frac{(x - E)^2}{2B}}.
\]  

(2)

Весовая функция имеет наибольшее значение при \( x = E \) и ничтожно мала при значениях \( x \), достаточно далёких от \( E \), так что в том взвешенном среднем (2) функции \( F_x \), с помощью которого мы выразили каноническое среднее \( \overline{F}_E \), заметные веса получают лишь те значения функции \( F_x \), в которых \( x \) достаточно близко к \( E \); мы имеем здесь типичный пример приближённого выражения величины \( \overline{F}_E \) с помощью «интегрального ядра», роль которого играет весовая функция \( p(x) \).

Чтобы дать пример расчёта погрешности равенства \( \overline{F}_E \approx \overline{F}_x \), допустим, что величина \( F_x \) с изменением \( x \) изменяется вблизи \( x = E \) не слишком быстро, например так, что (по крайней мере для \( |y| \) ) выполняется «условие Липшица»

\[
|\overline{F}_{E+y} - \overline{F}_E| < C|y|,
\]

где \( C > 0 \) — постоянная.

*) Двумерный нормальный закон § 5 гл. V редуцируется здесь к одномерному, так как \( N \) остаётся неизменным, и следовательно, параметр \( u_1 \) тождественно равен нулю.
Тогда мы получаем:

$$|\bar{\Omega}_E - \bar{\Omega}_E| \approx \frac{1}{\sqrt{2\pi B_{x_+}}} \sum_{x=-\infty}^{+\infty} (\bar{\Omega}_x - \bar{\Omega}_E) e^{-\frac{(x-E)^2}{2B}} =$$

$$= \frac{1}{\sqrt{2\pi B_{y_-}}} \sum_{y=-\infty}^{+\infty} (\bar{\Omega}_{E+y} - \bar{\Omega}_E) e^{-\frac{y^2}{2B}} < \frac{C}{\sqrt{2\pi B_{y_-}}} \sum_{y=-\infty}^{+\infty} y | e^{-\frac{y^2}{2B}} \approx$$

$$\approx \frac{C}{\sqrt{2\pi B_{y_-}}} \int_{-\infty}^{+\infty} y | e^{-\frac{y^2}{2B}} dy = \frac{C}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} z | e^{-\frac{z^2}{2}} dz = C^* \sqrt{N},$$

где $C^*$ — постоянная. Если функция $F(U)$ представляет собой сумматорную величину (или математическое ожидание такой величины), то величина $\bar{\Omega}_E$ будет, как мы много раз видели, бесконечно большой порядка $N$; последнее неравенство показывает поэтому, что, заменяя микроканоническое среднее сумматорной величины её каноническим средним, мы вводим лишь ничтожно малую относительную погрешность. Но как раз это мы и хотели установить.
ДОПОЛНЕНИЕ IV
РЕДУКЦИЯ К ОДНОМЕРНОЙ ЗАДАЧЕ В СЛУЧАЕ ПОЛНОЙ СТАТИСТИКИ

Проведённая нами в § 3 гл. V редукция оценки структурных функций к предельной задаче теории вероятностей может быть в случае полной статистики заменена значительно более просто и приводящей к одномерной предельной задаче. Причина этой возможности, как мы сейчас увидим, заключена в том, что в случае полной статистики удаётся обойтись введением одного параметра (β) вместо двух (α и β), которые необходимы в случае двух других статистик*).

Рассмотрим элементарный закон распределения

\[ P(z = k) = g_\beta e^{-\beta k} / \Phi(\beta), \quad k = 0, 1, \ldots, \] 

где

\[ \Phi(\beta) = \sum_{k=0}^{\infty} g_\beta e^{-\beta k} = \sum_{r=1}^{\infty} e^{-\beta r}; \]

пусть случайная величина

\[ R = \sum_{i=1}^{N} z_i \]

есть сумма \( N \) взаимно независимых случайных величин, каждая из которых распределена по только что определённому нами элементарному закону. Тогда для данного натурального числа \( E \)

\[ P(R = E) = \sum_{k_1}^{N} \prod_{i=1}^{N} P(z_i = k_i), \]

* Мы пытаемся случаем указать, что изложение этого вопроса, данное в § 11 книги автора «Об аналитическом аппарате физической статистики» (Труды Матем. ин-та им. В. А. Стеклова, т. XXXIII, 1950), хотя и не содержит ошибок, всё же неудовлетворительно, так как эта важная однопараметричность задачи в нём явно не подчёркнута.
где суммирование распространяется на все комбинации целых неотрицательных чисел $k_i$, удовлетворяющих условию

$$
\sum_{i=1}^{N} k_i = E, \quad (L)
$$

что мы кратко будем обозначать символом $\sum_{(L)}$. В силу (1)

$$
P(R = E) = \{\Phi(\beta)\}^{-N} e^{-\beta E} \sum_{(L)} \prod_{i=1}^{N} g_{k_i}. \quad (2)
$$

Если выбрана определённая система чисел $k_i (1 \leq i \leq N)$, удовлетворяющая уравнению (L), то мы можем считать, что этим фиксирована энергия $k_i$-й частицы; но этой энергии, как мы знаем, соответствует $g_{k_i}$ различных состояний частицы; а такое основное состояние системы в случае полной статистики взаимно однозначно определяется заданием основного состояния каждой из частиц, то данному набору чисел $k_i (1 \leq i \leq N)$ соответствует

$$
\prod_{i=1}^{N} g_{k_i}
$$

различных основных состояний системы. Число же всех основных состояний системы, соответствующих данному уровню $E$ её энергии, равно поэтому

$$
\sum_{(L)} \prod_{i=1}^{N} g_{k_i}.
$$

Но это число есть не что иное, как $\Omega(N, E)$; поэтому формула (2) даёт:

$$
P(R = E) = \{\Phi(\beta)\}^{-N} e^{-\beta E} \Omega(N, E),$$

откуда

$$
\Omega(N, E) = \{\Phi(\beta)\}^{N} e^{\beta E} P(R = E), \quad (3)
$$
Так как $R$ есть сумма большого числа $N$ взаимно независимых случайных величин, распределённых по одному и тому же постоянному (независимому от $N$ и $E$) закону, то соотношение (3) позволяет найти для $\Omega(N, E)$ нужные нам асимптотические выражения с помощью хорошо известных одномерных предельных теорем теории вероятностей. Легко убедиться, что полученные этим путём асимптотические выражения совпадают с теми, которые были нами найдены в гл. V.