
ГЛАВА 2

Метод наиболее вероятного распределения

Рассмотрим ансамбль из N идентичных систем. Мы описываем природу любой из этих систем, перечисляя ее возможные состояния, обозначаемые нами через $1, 2, 3, 4, \dots, l, \dots$. В принципе мы всегда имеем в виду квантово-механическую систему, состояния которой должны описываться собственными значениями замкнутой системы коммутирующих переменных. Собственные значения энергии на этих состояниях мы будем обозначать через $\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_l, \dots$, располагая их так, чтобы $\varepsilon_{l+1} \geq \varepsilon_l$. В случае необходимости эта схема может быть применена также и к «классической системе», состояния которой следует описывать как ячейки фазового пространства (p_k, q_k) , имеющие равный объем, причем независимо от того, являются они бесконечно малыми во всех направлениях или нет, они должны быть во всяком случае такими, чтобы энергия не менялась существенно в пределах ячейки. Впрочем, это применение является более или менее второстепенным, а более существенно следующее.

При рассмотрении состояния ансамбля мы будем всегда задавать его, указывая, что система №1 находится в состоянии, например l_1 , система №2 — в состоянии l_2, \dots , система № N — в состоянии l_N . Мы будем поступать так в дальнейшем, хотя в сущности делать этого нельзя, ибо квантово-механическая система не находится в том или ином состоянии, которое должно описываться замкнутой системой некоммутирующих переменных, выбранных раз и навсегда. Становясь на эту точку зрения, мы мыслим в строго «классическом» духе. При выбранной совокупности состояний можно считать, что индивидуальная система имеет лишь определенную амплитуду вероятности и, следовательно, определенную вероятность быть найденной при наблюдении в состояниях №1, №2, №3 и т. д. Я подчеркиваю: лишь амплитуду вероятности. Даже такая детерминированность отдельной системы является здесь излишней. Действительно, для приписывания отдельной системе «чистого состояния» вообще нет достаточных оснований.

Если бы мы углубились в рассуждения этого рода, то это завело бы нас в область очень тонких квантово-механических вопросов. Это было

сделано Нейманом, Вигнером и другими, однако их результаты не отличаются существенно от тех, которые получаются при указанных нами выше более простых и более наивных рассуждениях, которыми мы и воспользуемся. В таком случае некоторый класс состояний ансамбля будет характеризоваться тем, что $a_1, a_2, a_3, \dots, a_l$ систем из общего числа N систем находятся соответственно в состояниях $1, 2, 3, \dots, l, \dots$ и все состояния ансамбля охватываются — без перекрытия — классами, описываемыми всеми различными допустимыми совокупностями чисел a_l :

| | | | | | | | |
|------------------|-----------------|-----------------|-----------------|-----|-----------------|-----|-------|
| № состояния | 1 | 2 | 3 | ... | l | ... | |
| Энергия | ε_1 | ε_2 | ε_3 | ... | ε_l | ... | (2.1) |
| Число заполнения | a_1 | a_2 | a_3 | ... | a_l | ... | |

Число отдельных состояний, относящихся к этому классу, равно, очевидно

$$P = \frac{N!}{a_1! a_2! a_3! \dots a_l! \dots}. \quad (2.2)$$

Совокупность чисел a_l должна, разумеется, удовлетворять условиям

$$\sum_l a_l = N, \quad \sum_l \varepsilon_l a_l = E. \quad (2.3)$$

Выражениями (2.2) и (2.3) наш расчет в сущности заканчивается. Однако в этой форме результаты совершенно недоступны для исследования.

В настоящем методе предполагается, что, поскольку N чрезвычайно велико, общее число распределений (т.е. сумма всех P) почти полностью исчерпывается суммой тех P , для которых совокупности чисел a_l не отличаются существенным образом от совокупности, дающей максимальное значение P [разумеется, из числа тех, которые удовлетворяют соотношениям (2.3)]. Другими словами, считая, что такая совокупность чисел заполнения может быть получена всегда, мы пренебрегаем лишь очень малой долей всех возможных распределений, которая имеет «исчезающе малые шансы быть когда-либо реализованной».

Это предположение строго выполняется в пределе $N \rightarrow \infty$ (т.е. в применении к «мысленному» или «виртуальному» ансамблю, в сомнительных случаях мы всегда подразумеваем этот предельный переход, физически соответствующий «бесконечному тепловому резервуару»; мы еще раз убеждаемся в крупных преимуществах точки зрения Гиббса). Мы принимаем здесь это положение без доказательства; последнее

будет дано позднее при рассмотрении другого метода — «метода средних значений» Дарвина–Фаулера.

При большом, но конечном N это положение справедливо только приблизительно. Действительно, в больцмановском случае мы уже не можем пренебрегать распределениями, в которых «числа заполнения» образуют совокупности, отклоняющиеся от совокупности, соответствующей максимуму. Они дают сведения о термодинамических флуктуациях больцмановской системы, поддерживаемой при постоянной энергии E , находящейся в условиях идеальной тепловой изоляции.

Мы не будем, однако, останавливаться на этом подробно — отчасти ввиду весьма ограниченной применимости самой больцмановской точки зрения, а также по следующей причине: поскольку условие идеальной тепловой изоляции практически не может быть осуществлено, результаты, полученные для термодинамических флуктуаций при этом нереализуемом условии, применимы к действительности лишь отчасти, а именно в той мере, в какой можно доказать или принять, что эти результаты таковы же, как и для случая теплового резервуара. Флуктуации системы, помещенной в тепловой резервуар при постоянной температуре, могут быть получены значительно проще — непосредственно из концепции Гиббса. Таким образом, нам незачем рассматривать более сложный случай, из которого могут быть получены данные, применимые лишь к идеальным, нереализуемым условиям.

Возвращаясь к (2.2) и (2.3), выберем в качестве функции, максимум которой мы ищем, логарифм P , учитывая дополнительные условия обычным методом множителей Лагранжа λ и μ , т. е. будем искать абсолютный максимум выражения

$$\log P - \lambda \sum_l a_l - \mu \sum_l \varepsilon_l a_l. \quad (2.4)$$

Для преобразования логарифмов факториалов воспользуемся формулой Стирлинга в виде:

$$\log(n!) = n(\log n - 1). \quad (2.5)$$

Разумеется, мы рассматриваем величины a_l как непрерывные переменные. Для вариации (2.4) имеем:

$$-\sum_l \log a_l \delta a_l - \lambda \sum_l \delta a_l - \mu \sum_l \varepsilon_l \delta a_l = 0,$$

причем мы должны приравнять нулю коэффициенты при каждом δa_l , так что (при любом l):

$$\log a_l + \lambda + \mu \varepsilon_l = 0$$

или

$$a_l = a^{-\lambda - \mu \varepsilon_l},$$

λ и μ должны быть определены из добавочных условий:

$$\sum_l e^{-\lambda - \mu \varepsilon_l} = N, \quad \sum_i \varepsilon_i e^{-\lambda - \mu \varepsilon_i} = E.$$

Путем почленного деления мы исключаем λ , хотя можно получить $e^{-\lambda}$ также и непосредственно из первой формулы. Обозначив через $E/N = U$ среднюю долю энергии, приходящуюся на одну систему, мы можем выразить полученный нами результат следующим образом:

$$\left. \begin{aligned} \frac{E}{N} = U &= \frac{\sum \varepsilon_l e^{-\mu \varepsilon_l}}{\sum e^{-\mu \varepsilon_l}} = -\frac{\partial}{\partial \mu} \log \sum e^{-\mu \varepsilon_l}, \\ a_l &= N \frac{e^{-\mu \varepsilon_l}}{\sum e^{-\mu \varepsilon_l}} = -\frac{N}{\mu} \frac{\partial}{\partial \varepsilon_l} \log \sum e^{-\mu \varepsilon_l}. \end{aligned} \right\} \quad (2.6)$$

Совокупность уравнений второй строки выражает распределение наших N систем по их энергетическим уровням. Можно сказать, что эти уравнения содержат в зародыше всю термодинамику, которая целиком зиждется на этом основном распределении. Само выражение совершенно прозрачно — экспонента $e^{-\mu \varepsilon_l}$ выражает число заполнения в виде доли общего числа N систем, а сумма в знаменателе является лишь «нормирующим множителем». Разумеется, μ следовало бы определять из первого уравнения в зависимости от средней энергии и от характера системы (т. е. величин ε_l). Естественно, что решить это уравнение в общем виде относительно μ невозможно. В самом деле, очевидно, что функциональная зависимость между μ и U не является универсальной, а целиком зависит от характера системы.

К счастью, однако, мы можем дать нашим соотношениям вполне удовлетворительную общую физическую интерпретацию, не решая это уравнение относительно μ , так как последняя величина (введенная первоначально лишь как множитель Лагранжа, т. е. как математическое вспомогательное средство) оказывается значительно более фундаментальной величиной, чем само U ; поэтому физик предпочитает получать

в каждом частном случае U как функцию μ , а не наоборот, что было бы совершенно нецелесообразным.

Чтобы объяснить это без излишних околичностей, мы решительно станем на точку зрения Гиббса, т. е. будем считать, что мы имеем дело с виртуальным ансамблем, отдельный элемент которого представляет собой самую рассматриваемую систему. Поскольку все отдельные элементы равноправны, мы можем, переходя к физической интерпретации, представлять себе a_l , или, вернее, $\frac{a_l}{N}$, как частоты, с которыми отдельная система, погруженная в большой тепловой резервуар, будет встречаться в состоянии ε_l ; U является в этих условиях средней энергией отдельной системы.

Применим наши результаты (2.6) к трем различным ансамблям систем:

$$\begin{array}{ccc} A & B & A + B \\ \text{уровни:} & \alpha_k & \beta_m & \varepsilon_l = \alpha_k + \beta_m. \end{array} \quad (2.7)$$

При этом мы подразумеваем, что в первом и втором случаях отдельные элементы будут являться любыми двумя рассмотренными ранее различными системами общего типа. В третьем же случае отдельный элемент будет состоять из одной системы A и одной системы B , находящихся в слабом энергетическом взаимодействии, так что общий энергетический уровень в третьем случае будет суммой любого α_k и любого β_m , (индекс l заменяет два индекса k и m). Нетрудно видеть, что в третьем случае сумма распадается на произведение двух сумм:

$$\sum_l e^{-\mu\varepsilon_l} = \sum_k \sum_m e^{-\mu(\alpha_k + \beta_m)} = \sum_k e^{-\mu\alpha_k} \sum_m e^{-\mu\beta_m}. \quad (2.8)$$

Тогда, на основании (2.6), — речь идет о третьем случае, — общее число заполнения a_l (которое может быть обозначено и как $a_{k,m}$) равняется

$$a_l \equiv a_{(k,m)} = N \frac{e^{-\mu(\alpha_k + \beta_m)}}{\sum_k e^{-\mu\alpha_k} \sum_m e^{-\mu\beta_m}}. \quad (2.9)$$

Продолжим рассмотрение третьего случая и будем искать число систем $A+B$, в которых A находится на определенном уровне α_k . Оно, очевидно, находится суммированием (2.9) по всем m . В результате $\sum e^{-\mu\beta}$

исчезает в числителе и в знаменателе, и мы получаем

$$\sum_m a_{(k, m)} = N \frac{e^{-\mu\alpha_k}}{\sum_k e^{-\mu\alpha_k}}.$$

Таким образом, мы видим, что общее статистическое распределение систем A в третьем случае (включающее среди многого другого среднее значение их энергии) в точности такое же, как и в ансамбле A (первый случай), при условии, что мы обеспечиваем (надлежащим выбором E/N в случае A) одинаковость значений μ в обоих случаях.

Поскольку те же рассуждения применимы и к системе B , то, согласно вышеизложенному, получается следующее: если привести системы A и B в слабое взаимодействие друг с другом и поместить их в тепловой резервуар, то обе системы будут вести себя в точности так, как если бы каждая из них в отдельности была помещена в тепловой резервуар, при условии, что все три тепловых резервуара выбираются такими, чтобы значения μ были во всех трех случаях одинаковыми. Иначе говоря, если это осуществлено, установившаяся энергетическая связь оказывается достаточно слабой, вследствие чего взаимное влияние или обмен энергией в среднем отсутствуют.

Равенство значений μ означает равенство температур. Вряд ли можно интерпретировать это иначе. А так как мы можем выбрать раз и навсегда стандартную систему A («термометр») и привести во взаимодействие с любой другой системой B , то μ должно быть универсальной функцией температуры T .

Это заключение может быть сделано значительно более убедительным, если нам удастся определить явный вид зависимости μ от температуры.

Для этой цели полезно обратить внимание на самоочевидный и в то же время весьма существенный побочный результат наших предшествующих рассуждений. Мы видели, что в случае $A + B$

$$\sum_l e^{-\mu\varepsilon_l} = \sum_k e^{-\mu\alpha_k} \sum_m e^{-\mu\beta_m}.$$

Следовательно, функция от μ , играющая, как мы увидим, весьма важную роль, а именно функция

$$\log \sum_l e^{-\mu\varepsilon_l} \tag{2.10}$$

[ее значение явствует из рассмотрения последних членов (2.6)] является аддитивной для двух систем, находящихся в слабом энергетическом взаимодействии. В этом и состоит то очевидное, но существенное положение, о котором я упоминал выше.

Каково же функциональное соотношение между μ и T ? Чтобы выбрать из множества монотонных функций $f(T)$ ту, которая является «истинной» температурой T , имеется, как известно, лишь один критерий: $1/T$ является универсальным интегрирующим множителем бесконечно малого приращения тепла dQ в состоянии термодинамического равновесия — универсальным в том смысле, что он пригоден для любой системы. Других функций от T , обладающих этим свойством, не существует — в этом заключается определение T (Кельвин). Наша модель является еще недостаточной, чтобы мы могли воспользоваться этим определением, ибо если «природа» каждой системы (т. е. ее уровень ε_l) фиксирована раз и навсегда, все зависит от одного параметра: μ или U , или T . В случае одной переменной рушится понятие «интегрирующего множителя», так как при некотором dx любое $\varphi(x) dx$ также будет «полным дифференциалом». Поэтому, чтобы установить T , мы должны ввести в рассмотрение другие параметры или, что то же, ввести понятие о механической работе, совершаемой нашей системой.

Положим, для краткости,

$$\log \sum_l e^{-\mu \varepsilon_l} = F, \quad (2.11)$$

где F должно рассматриваться как функция μ и всех ε_l , и напишем, пользуясь (2.6), бесспорно правильное математическое соотношение, приводящее в дальнейшем к физическим приложениям:

$$dF = \frac{\partial F}{\partial \mu} d\mu + \sum_l \frac{\partial F}{\partial \varepsilon_l} d\varepsilon_l = -U d\mu - \frac{\mu}{N} \sum_l a_l d\varepsilon_l, \quad (2.12)$$

и, таким образом,

$$d(F + U\mu) = \mu \left(dU - \frac{1}{N} \sum_l a_l d\varepsilon_l \right). \quad (2.13)$$

Применим это к следующему физическому процессу, которому мы подвергаем наш ансамбль из N систем.

Пусть с каждой из этих систем связан в точности одинаковый «механизм» — цилиндры, поршни и прочие приспособления, с помо-

щью которых мы можем управлять системой и изменять ее характер (т. е. уровни ε_l).

При этом, разумеется, мы изменяем ε_l одинаковым образом во всех системах так, чтобы основное условие тождественности N систем (условие, на котором покоятся наши рассуждения) было выполнено. Кроме того, мы обеспечиваем также непосредственное «изменение температуры», связывая наш ансамбль с большим тепловым резервуаром (имеющим ту же температуру), изменяя общую температуру очень медленно и затем снова изолируя ансамбль от теплового резервуара.

При применении (2.13) к этому процессу $a_l d\varepsilon_l$ является работой, которую мы должны совершить над поршнями и т. д. этих a_l систем, чтобы «поднять» эти системы со старого уровня ε_l на измененный уровень $\varepsilon_l + d\varepsilon_l$. $\sum a_l d\varepsilon_l$ есть работа, совершаемая при этом над ансамблем; $-\sum a_l d\varepsilon_l$ — работа, производимая ансамблем, и $-\frac{1}{N} \sum a_l d\varepsilon_l$ — средняя работа, производимая одним из членов ансамбля. Следовательно, поскольку dU является средним приращением энергии, выражение, стоящее в скобках в правой части (2.13), должно быть средним приращением тепла dQ , сообщаемого ансамблю. μ должно при этом рассматриваться как интегрирующий множитель. Одного этого уже вполне достаточно, чтобы заключить, что μ неизбежно должно равняться $1/T$, ибо не существует другой функции от T , обладающей этим свойством для всех систем. Таким образом, $F + U\mu$ должно представлять собой энтропию.

Для более прямого доказательства обозначим

$$F + U\mu = G. \quad (2.14)$$

Тогда, согласно общей математической теореме, отношение двух интегрирующих множителей $1/T$ и μ является функцией G :

$$\frac{1}{T\mu} = \varphi(G). \quad (2.15)$$

Тогда из (2.13)

$$\varphi(G) dG = \frac{dQ}{T} = dS, \quad (2.16)$$

где S — энтропия. Это дает, после интегрирования, G в виде некоторой функции от S :

$$G = \chi(S). \quad (2.17)$$

Далее, из (2.14), (2.11) и (2.6), функция

$$G = \log \sum_l e^{-\mu \varepsilon_l} - \mu \frac{\partial}{\partial \mu} \log \sum_l e^{-\mu \varepsilon_l}$$

оказывается «аддитивной» при комбинировании двух систем (вследствие аддитивности $\log \sum_l e^{-\mu \varepsilon_l}$). Обозначая функцию χ из (2.17) через χ_A в случае системы A , через χ_B в случае системы B и через χ_{AB} в случае системы $A+B$ и обозначая энтропии в этих трех случаях соответственно через S_A , S_B и S_{AB} , получим

$$\chi_A(S_A) + \chi_B(S_B) = \chi_{AB}(S_{AB}).$$

С другой стороны, энтропия также является аддитивной функцией или, во всяком случае,

$$S_{AB} = S_A + S_B + C,$$

где C не зависит от S_A и S_B . Следовательно,

$$\chi_A(S_A) + \chi_B(S_B) = \chi_{AB}(S_A + S_B + C).$$

Если продифференцировать это уравнение один раз по S_A и второй раз по S_B и сравнить результаты, то получим

$$\chi'_A(S_A) = \chi'_B(S_B) = \text{универсальная постоянная} = \frac{1}{k}.$$

Таким образом, из (2.17), (2.16) и (2.15).

$$\mu = \frac{1}{kT} \tag{2.18}$$

и

$$S = kG + \text{const} = k \log \sum_l e^{-\frac{\varepsilon_l}{kT}} + \frac{U}{T} + \text{const}, \tag{2.19}$$

где const во всяком случае не зависит от T и от «параметров» (объем и т. д.), от которых зависят ε_l .

Мы отбрасываем const , не вдаваясь пока в анализ того, что означает такое приравнение нулю постоянной. Тогда

$$k \log \sum_l e^{-\frac{\varepsilon_l}{kT}} = S - \frac{U}{T} = \Psi. \tag{2.20}$$

Мы получили, таким образом, общее правило вывода термодинамики системы, исходя из ее механики. Это правило применимо ко всем случаям (включая так называемую «новую» статистику).

Образуем «функцию распределения» (называемую также «суммой по состояниям» или «статистической суммой»)

$$Z = \sum_l e^{-\frac{\varepsilon_l}{kT}}. \quad (2.21)$$

Тогда $k \log Z$ (где k — постоянная Больцмана) есть свободная энергия, взятая с обратным знаком и деленная на T . (Следуя Планку, мы обозначаем эту функцию буквой Ψ). Из замечаний, сделанных по поводу (2.13), легко видеть, что $k \log Z$ является термодинамической Ψ -функцией не только по отношению к изменению температуры, но также и по отношению к параметрам (например, объему V), от которых могут зависеть ε_l . Таким образом, средние силы, с которыми система стремится увеличить эти параметры (например, давление p в случае, когда в качестве параметра берется объем), находятся с помощью известных из термодинамики формул, прототипом которых является соотношение

$$p = T \frac{\partial \Psi}{\partial V} = kT \frac{\partial}{\partial V} \log \sum_l e^{-\frac{\varepsilon_l}{kT}}, \quad (2.22)$$

тогда как первое уравнение (2.6)

$$U = -\frac{\partial}{\partial \mu} \log \sum_l e^{-\mu \varepsilon_l} = T^2 \frac{\partial \Psi}{\partial T} \quad (2.23)$$

является столь же хорошо известной формулой общей термодинамики. (Во всем вышесказанном Ψ следует рассматривать как функцию от T и от параметров типа V , т. е. от которых могут зависеть ε_l ; макроскопически эти параметры должны удовлетворять тому требованию, что при поддержании их постоянными система не совершает механической работы¹). Таким образом, статистическая трактовка, давая в принципе Ψ как функцию T и параметров, подобных V , дает исчерпывающую

¹Против этой трактовки может быть выдвинуто то возражение, что экспериментально затруднительно поддерживать постоянство объема V и реализовать полную тепловую изоляцию. Статистическая трактовка, предпочитающая вводить в качестве параметров вместо величин V такие величины, как p , вполне возможна, быть может, даже предпочтительна, однако значительно более сложна.

информацию о термодинамическом поведении объекта. [Как известно, термодинамическая функция дает полную информацию только в том случае, если она известна как функция некоторых определенных переменных. Этим свойством не обладают, например, функции $\Psi(T, p, \dots)$ или $S(T, V, \dots)$, в то время как $S(U, V, \dots)$, например, дает исчерпывающую информацию.]