Глава 6
Метод средних значений

Вернемся теперь к задаче, поставленной во второй главе, и рассмотрим ее, пользуясь новым методом. Мы делаем это по ряду причин. Во-первых, потому, что рассуждения в пятой главе показали, что «метод наиболее вероятных значений» не является вполне строгим; настоящий метод, которым мы обязаны Дарвину и Фаулеру, многим представляется более убедительным, может быть, даже совершенно точным. Во-вторых, всегда заманчиво и поучительно видеть, что в точности тот же результат может быть получен путем совершенно иных рассуждений, в особенности, если дело идет о весьма общей теореме, имеющей фундаментальное значение. В-третьих, развиваемый здесь математический метод окажется очень полезным также и в других приложениях.

Нашей целью является вычисление средних значений a_l в ансамбле Гиббса, определяемых соотношением (5.1). Воспользуемся приемом, который мы применили в (5.2), (5.3) и (5.4) и в силу которого все, что мы хотим знать, может быть получено из величины

$$\sum P = \sum_{(a_l)} \frac{N!}{a_1!\, a_2! \ldots a_l! \ldots} \, \omega_1^{a_1} \omega_2^{a_2} \ldots \omega_l^{a_l} \ldots , \tag{6.1}$$

где суммирование производится по всем совокупностям величин a_l, удовлетворяющим (2.3). Таким образом, нам достаточно вычислить эту сумму.

Если бы единственным ограничением для a_l являлось условие $\sum a_l = N$, то эта задача решалась бы немедленно с помощью формулы для полинома, и сумма, по крайней мере формально, была бы равна

$$(\omega_1 + \omega_2 + \omega_3 + \ldots + \omega_l + \ldots)^N.$$

(Чтобы результат был конечным, следует обрывать ряды уровней на каком-либо очень высоком уровне.) Второе условие $\sum a_l \varepsilon_l = E$ автоматически ограничивает число членов в (6.1), в силу невозможности существования уровня $\varepsilon_l > E - (N-1)\varepsilon_1$; однако в то же самое время оно же и определяет трудность задачи, состоящую в необходимости выбора лишь тех членов, которые удовлетворяют этому условию.

Чтобы преодолеть эту трудность, мы воспользуемся следующим ухищрением. Не принимая во внимание второго из указанных ограничений, вычислим такую сумму:

$$
\begin{aligned}
&\sum P z^{a_1\varepsilon_1 + a_2\varepsilon_2 + \ldots + a_l\varepsilon_l + \ldots} = \\
&= \sum \frac{N!}{a_1!\, a_2! \ldots a_l! \ldots} (\omega_1 z^{\varepsilon_1})^{a_1} (\omega_2 z^{\varepsilon_2})^{a_2} \ldots (\omega_l z^{\varepsilon_l})^{a_l} \ldots = \qquad (6.2)\\
&= (\omega_1 z^{\varepsilon_1} + \omega_2 z^{\varepsilon_2} + \ldots \omega_l z^{\varepsilon_l} + \ldots)^N = f(z)^N,
\end{aligned}
$$

где

$$
f(z) = \omega_1 z^{\varepsilon_1} + \omega_2 z^{\varepsilon_2} + \ldots \omega_l z^{\varepsilon_l} + \ldots \qquad (6.3)
$$

Если бы все ε_l и E были целыми числами, то искомая сумма $\sum P$ была бы, очевидно, коэффициентом при z^E в функции от z в (6.2); она может быть вычислена методом теории вычетов в комплексной плоскости z.

Чтобы осуществить этот план, мы должны — в этом и заключается наше ухищрение — объявить с самого начала выбранную нами единицу энергии столь малой, что можно считать с любой желаемой степенью точности все уровни ε_l, и заданную полную энергию E целыми кратными этой единицы или, если угодно, даже заменить их целыми кратными этой единицы, не меняя сколько-нибудь существенно самой физической задачи. Имеются, конечно, случаи, в которых это оказывается невозможным, в частности, когда густота уровней ε_l становится бесконечной вблизи некоторой конечной энергии ε; это, например, имеет место для электронных уровней свободного атома водорода. Мы исключаем такие случаи, которые, как можно показать, вообще недоступны статистическому исследованию без специальных предосторожностей (например, атом водорода должен был бы быть заключен в большой, но конечный ящик, препятствующий удалению электрона в бесконечность).

Представляется удобным ввести два дальнейших ограничения, касающиеся ε_l. Во-первых, если $\varepsilon_1 \neq 0$, то мы пользуемся вместо уровней $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_l, \ldots$ уровнями $0, \varepsilon_2 - \varepsilon_1, \varepsilon_3 - \varepsilon_1, \ldots, \varepsilon_l - \varepsilon_1, \ldots$, заменяя одновременно E на $E - N\varepsilon_1$. Рассмотрение (6.3) и следующей формулы (6.4) убеждает нас в том, что это не вносит никаких изменений, являясь лишь более удобным математическим способом выражения.

Предположим, для простоты, что $\varepsilon_1 = 0$. Во-вторых, мы предполагаем, что величины ε_l не имеют общего делителя. (Этого всегда можно достигнуть, так как в противном случае E также должно было бы его содержать, чтобы условие $\sum a_l\varepsilon_l = E$ могло бы в точности быть выполнено.) Таким образом, если общий наибольший делитель равен τ, то мы выбираем единицу энергии в τ раз большей, что уничтожает делитель, оставляя, однако, все значения целыми числами.

Приняв это условие, получаем простое и очевидное решение

$$\sum P = \frac{1}{2\pi i}\oint z^{-E-1} f(z)^N\, dz. \tag{6.4}$$

Здесь интегрирование ведется вдоль любого замкнутого контура (рис. 2), охватывающего начало координат в комплексной плоскости z и, кроме того, лежащего внутри круга сходимости $f(z)$, что позволяет избежать аналитического продолжения. Интеграл оценивается методом наиболее крутого спуска (метод седловидной точки). Рассмотрим поведение подынтегрального выражения, двигаясь от нуля до бесконечности вдоль действительной положительной оси и учитывая, что все ω в (6.3) виртуально равны единице и что $0 = \varepsilon_1 \leqslant \varepsilon_2 \leqslant \varepsilon_3$. Первый сомножитель в подынтегральном выражении z^{-E-1} убывает быстро и монотонно, начиная с бесконечного положительного значения. Второй множитель $f(z)^N$, начиная со значения 1 при $z = 0$, монотонно возрастает, стремясь к бесконечности по мере приближения z к границе круга сходимости $f(z)$, где бы она ни была расположена. Кроме того, относительное уменьшение первого сомножителя

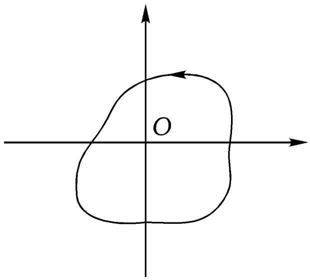

Рис. 2

$$\frac{E+1}{z}$$

само убывает монотонно от $+\infty$ при $z = 0$ до нуля при $z = \infty$; относительное возрастание второго сомножителя

$$N\frac{f'(z)}{f(z)} = N\frac{\sum\limits_l \varepsilon_l z^{\varepsilon_l}}{\sum\limits_l z^{\varepsilon_l}}$$

обнаруживает противоположное поведение. Эта величина равна нулю при $z = 0$ и монотонно возрастает. В самом деле,

$$\frac{d}{dz}\left(N\frac{f'(z)}{f(z)}\right) = N\frac{\sum\limits_l \varepsilon_l^2 z^{\varepsilon_l}\sum\limits_k z^{\varepsilon_k} - \left(\sum\limits_m \varepsilon_m z^{\varepsilon_m}\right)^2}{\left(\sum\limits_s z^{\varepsilon_s}\right)^2}.$$

Числитель этого выражения может быть переписан в виде

$$\sum_l z^{\varepsilon_l}\left(\varepsilon_l \sqrt[+]{\left(\sum_k z^{\varepsilon_k}\right)} - \frac{\sum\limits_m \varepsilon_m z^{\varepsilon_m}}{\sqrt[+]{\left(\sum\limits_k z^{\varepsilon_k}\right)}}\right)^2 > 0,$$

откуда видно, что он положителен.

При этих обстоятельствах подынтегральная функция обнаруживает один и только один минимум (не принимая других экстремальных значений) внутри круга сходимости $f(z)$. Этот минимум, как можно ожидать и как будет показано в дальнейшем, является очень острым, что вытекает из того, что обе экспоненты, т.е. $E+1$ и N — весьма большие числа; действительно, не следует забывать, что нас интересует переход к пределам $N \to \infty$, $E \to \infty$ при сохранении постоянства отношения $\frac{E}{N}$, поскольку последнее выражает среднюю энергию, которой обладает одна система, входящая в ансамбль.

Другими словами, в этой точке на действительной положительной оси (обозначим ее пока через z_0, опустив впоследствии индекс 0) первая производная подынтегральной функции обращается в нуль, вторая же производная должна быть положительной, и можно думать, что она очень велика. Следовательно, если мы будем двигаться от этой точки в направлении, перпендикулярном действительной оси, давая чисто мнимые приращения, подынтегральная функция будет проходить (оставаясь вначале действительной) через очень острый максимум. Возьмем в качестве контура интегрирования в (6.4) окружность с центром в точке O, проходящую через точку $z = z_0$, рассчитывая, что основную долю в значение интеграла будет вносить только непосредственная окрестность этого чрезвычайно острого максимума. В свое время мы это докажем.

Определим прежде всего значение z_0, полагая первую производную равной нулю, и определим значение второй производной при $z = z_0$. При этом будет удобно воспользоваться логарифмическими производными. Пусть на действительной положительной оси

$$z^{-E-1} f(z)^N = e^{g(z)} \tag{6.5}$$

[мы берем, разумеется, главную ветвь, т. е. действительную часть $g(z)$]. Тогда z_0 будет определяться выражением

$$g'(z_0) = -\frac{E+1}{z_0} + N\frac{f'(z_0)}{f(z_0)} = 0; \tag{6.6}$$

и, далее,

$$g''(z_0) = \frac{E+1}{z_0^2} + N\left(\frac{f''(z_0)}{f(z_0)} - \frac{f'(z_0)^2}{f(z_0)^2}\right). \tag{6.7}$$

Это показывает, во-первых, что при достаточно больших E и N z_0 не будет изменяться при пропорциональном возрастании E и N; во-вторых, поскольку величина $g''(z_0)$ в этом случае изменяется пропорционально E и N, она может быть сделана, таким образом, сколь угодно большой (то, что она положительна, не нуждается в доказательстве, так как это вытекает из общих соображений).

Следовательно, для очень малого, чисто мнимого приращения iy, сообщаемого z вблизи $z = z_0$, подынтегральное выражение может быть написано следующим образом:

$$z_0^{-E-1} f(z_0)^N e^{-\frac{1}{2}y^2 g''(z_0)+\dots}, \tag{6.8}$$

и ближайшая окрестность окружности, по которой производится интегрирование, приведет [с любой желательной точностью, если с увеличением N величина $g''(z_0)$ делается достаточно большой] к результату:

$$\begin{aligned}\left[\sum P\right] &= \frac{1}{2\pi i} z_0^{-E-1} f(z_0)^N \int\limits_{-\infty}^{+\infty} e^{-\frac{1}{2}y^2 g''(z_0)} i\, dy = \\ &= z_0^{-E-1} f(z_0)^N \frac{1}{\sqrt{\{2\pi g''(z_0)\}}}.\end{aligned} \tag{6.9}$$

Мы заключили $\sum P$ в скобки, ибо нам предстоит еще доказать, что вычисление можно считать оконченным, т. е. что долей, вносимой остальной частью окружности, можно пренебречь при больших N.

Образно говоря[1], это происходит потому, что отдельные члены ряда (6.3), «усиливающие» друг друга на действительной оси, будут, по мере перемещения z по окружности, «поворачиваться» вокруг начала координат с различными скоростями, определяемыми различными целыми числами ε_l; в результате $|f(z)|$ будет, вообще говоря, значительно меньше, чем $f(z_0)$ (если не рассматривать область, лежащую непосредственно вблизи $z = z_0$, что мы уже оговорили). Отношение абсолютного значения подынтегральной функции, взятой в произвольной точке z на окружности, к значению ее в $z = z_0$ равно

$$\left\{\frac{|f(z)|}{f(z_0)}\right\}^N, \tag{6.10}$$

и становится для больших N сколь угодно малым по сравнению с последним (тоже малым) множителем в (6.9), т. е. с $[2\pi g''(z_0)]^{-\frac{1}{2}}$, порядок малости которого равен $N^{-\frac{1}{2}}$. Чтобы сделать это заключение строгим, мы должны показать, что максимальное значение $|f(z)|$, скажем, M существенно меньше, чем $f(z_0)$:

$$M < f(z_0).$$

Действительно, при этом доля, вносимая в интеграл остальной частью окружности, наверное, не превышает величины

$$\frac{1}{2\pi} z_0^{-E} M^N 2\pi = z_0^{-E} f(z_0)^N \left(\frac{M}{f(z_0)}\right)^N, \tag{6.12}$$

которой, при $N \to \infty$ можно пренебречь по сравнению с (6.9).

Чтобы доказать (6.11), обратим внимание на то, что равенство $M = f(z_0)$ могло бы иметь место только в том случае, если в какой-нибудь точке окружности z, существенно отличной от z_0, все члены (6.3) снова складывались бы наиболее благоприятным образом. Поскольку первый член действителен и положителен ($\varepsilon_1 = 0$), то и все остальные должны были бы быть действительными и положительными в этой точке. Пусть $\varphi(\leqslant 2\pi)$ — фазовый угол в этой точке. Тогда все произведения

$$\varepsilon_1\varphi,\ \varepsilon_2\varphi,\ \ldots,\ \varepsilon_l\varphi,\ \ldots$$

должны были бы быть целыми кратными 2π, а все целые числа ε_l — целыми кратными $\frac{2\pi}{\varphi}$; т. е.

$$\varepsilon_l = n_l \frac{2\pi}{\varphi}.$$

[1] Читатель при желании может полностью не знакомиться с этим растянутым доказательством (набранным петитом).

Однако это возможно только, если $\varphi = 2\pi$ (т. е. при $z = z_0$), ибо если $\frac{2\pi}{\varphi}$ больше единицы, то оно должно быть рациональной дробью $\frac{p}{q}$ с числителем больше единицы, даже если эта дробь образована наименьшими целыми числами p и q. Тогда p было бы общим делителем всех ε_l что противоречит нашему предположению об отсутствии такового.

Это доказательство довольно искусственно и не очень импонирует физику, которому трудно поверить, что отдельный уровень ε_l может испортить все дело. Действительно, мы можем себе представить, что все уровни, кроме одного, имеют довольно большой общий делитель p, который не может быть устранен вследствие того, что один уровень им не обладает. Приходится удовлетвориться тем, что даже один единственный «нарушитель» может воспрепятствовать неограниченному приближению максимума M к $f(z_0)$. Действительно, поскольку не все ε_l, имеют общий делитель, они должны приобретать это свойство (не иметь делителя) в некоторой конечной точке ряда, например ε_m. Предполагаемый «нарушитель» может тогда оказаться лишь членом ряда с $\varepsilon_l \leqslant \varepsilon_m$, что, очевидно, устанавливает также верхний предел величины предполагаемого общего делителя p остальной части. Не полностью действительный (частично мнимый) член ряда будет обладать фазовым углом, равным, по крайней мере, $\frac{2\pi}{p}$, и будет выражаться в виде

$$\omega_l z_0^{\varepsilon_l} e^{\frac{2\pi i}{p}}.$$

Это, очевидно, вызывает конечные отклонения $|f(z)|$ от $|f(z_0)|$, хотя при достаточно больших ε_l и p отклонение может быть достаточно малым; остальное может быть учтено путем перехода к пределу $N \to \infty$ в (6.10) или (6.12).

Вернемся теперь к нашим основным результатам (6.6), (6.7) и (6.9). Перепишем их, отбросив для краткости индекс в z_0, так как нас будет интересовать именно это действительное положительное значение z. В уравнении (6.3) мы также будем подразумевать под z эту величину. Таким образом, подводя итог нашим результатам, получаем:

$$f(z) = \omega_1 z^{\varepsilon_1} + \omega_2 z^{\varepsilon_2} + \ldots + \omega_l z^{\varepsilon_l} + \ldots, \tag{6.13}$$

$$g'(z) = -\frac{E+1}{z} + N\frac{f'(z)}{f(z)} = 0, \tag{6.14}$$

$$g''(z) = \frac{E+1}{z^2} + N\left(\frac{f''(z)}{f(z)} - \frac{f'(z)^2}{f(z)^2}\right), \tag{6.15}$$

$$\sum P = z^{-E-1} f(z)^N \frac{1}{\sqrt{\{2\pi g''(z)\}}}, \tag{6.16}$$

$$\log \sum P = -(E+1)\log z + N \log f(z) - \frac{1}{2}\log(2\pi g''(z)). \tag{6.17}$$

Даже последний член в последней формуле оказывается пренебрежимо малым, и мы могли бы опустить его на том основании, что он имеет лишь порядок $\log N$. Однако для осторожности мы все же удержим его на некоторое время.

Теперь мы получим из (5.3) средние числа заполнения

$$\overline{a}_l = \omega_l \frac{\partial \log \sum P}{\partial \omega_l} = \omega_l g'(z) \frac{\partial z}{\partial \omega_l} + \frac{\omega_l N z^{\varepsilon_l}}{f(z)} - \frac{1}{2} \frac{\partial}{\partial \omega_l} \log g''(z). \tag{6.18}$$

Первый член равен нулю в силу (6.14). (Нам следовало, конечно, принять во внимание неявную зависимость z от ω_l.) Рассматривая последний член, введем среднюю энергию

$$\frac{E}{N} = U, \tag{6.19}$$

которая не меняется при предельном переходе $N \to \infty$, $E \to \infty$. Тогда (6.15) принимает вид:

$$g''(z) = N\left(\frac{U}{z^2} + \frac{f''}{f} - \frac{f'^2}{f^2}\right). \tag{6.20}$$

Следовательно, последний член в (6.18) при предельном переходе является также постоянным, и мы получаем (полагая, согласно намеченному плану, все ω_l равными единице):

$$\overline{a}_l = N \frac{z^{\varepsilon_l}}{z^{\varepsilon_1} + z^{\varepsilon_2} + \ldots + z^{\varepsilon_l}}. \tag{6.21}$$

Если учесть (6.19) и положить все ω равными 1, то уравнение (6.14), определяющее z, может быть написано следующим образом:

$$U = \frac{\varepsilon_1 z^{\varepsilon_1} + \varepsilon_2 z^{\varepsilon_2} + \ldots + \varepsilon_l z^{\varepsilon_l} + \ldots}{z^{\varepsilon_1} + z^{\varepsilon_2} + \ldots + z^{\varepsilon_l} + \ldots} \tag{6.22}$$

Если мы положим

$$\log z = -\mu, \tag{6.23}$$

то последние два уравнения будут тождественными копиями основного соотношения (2.6), исходя из которого мы начали построение термодинамической теории. Отличие состоит лишь в том, что средние значения $\overline{a}_l$ заменены теперь наиболее вероятными значениями. Наша $f(z)$

играет роль статистической суммы, таким образом, мы можем теперь утверждать, что мы обосновали теорию новым, независимым способом. Посмотрим теперь, что дает (5.4) для флуктуаций. Используя (6.18), образуем

$$\overline{a_l^2} - (\overline{a}_l)^2 = \omega_l \frac{\partial}{\partial \omega_l} \left\{ \omega_l g'(z) \frac{\partial z}{\partial \omega_l} + \frac{\omega_l N z^{\varepsilon_l}}{f(z)} - \frac{1}{2} \frac{\partial}{\partial \omega_l} \log g''(z) \right\}.$$

Первый член не дает здесь ничего, так как $g'(z) = 0$. Уравнение (6.14) следует считать выполняющимся тождественно относительно ω_l. Последний член может быть опущен, так как, согласно (6.20), он имеет «нулевой порядок» относительно N, и члены, имеющие порядок N, окажутся преобладающими. Дифференцируя средний член, мы снова должны принять во внимание, что z зависит от ω_l (хотя, как правило, не очень сильно и, согласно (6.14) и (6.19), независящим от N образом). Мы получаем:

$$\overline{a_l^2} - (\overline{a}_l)^2 = \omega_l N \frac{z^{\varepsilon_l}}{f(z)} + \omega_l^2 N \left\{ \frac{\partial z}{\partial \omega_l} \left(-\frac{z^{\varepsilon_l} f'(z)}{f(z)^2} + \frac{\varepsilon_l z^{\varepsilon_l - 1}}{f(z)} \right) - \frac{z^{2\varepsilon_l}}{f(z)^2} \right\}.$$

Полагая все ω равными единице и пользуясь (6.14), (6.19) и (6.21), без труда получаем:

$$\overline{a_l^2} - (\overline{a}_l)^2 = \overline{a}_l \left[1 + (\varepsilon_l - U) \frac{\partial \log z}{\partial \omega_l} - \frac{\overline{a}_l}{N} \right]. \tag{6.24}$$

Поскольку в квадратных скобках несомненно нет членов, имеющих порядок N, среднее квадратичное отклонение если и не является в точности «нормальным», то во всяком случае имеет «нормальный порядок», т. е. порядок $\overline{a}_l$. Таким образом, относительная флуктуация стремится к нулю, когда N и все $\overline{a}_l$ стремятся к бесконечности. Распределение становится бесконечно острым. Средние значения, наиболее вероятные значения и вообще любые значения с неисчезающей вероятностью все становятся одними и теми же.

Можно, впрочем, дать среднему члену в (6.24) точную оценку. Это довольно поучительно, хотя и не имеет особенно важного значения. Этот член оказывается всегда отрицательным. Для этой цели представляется несколько более удобным перейти от величин z к величинам μ или T, пользуясь соотношением (6.23):

$$\log z = -\mu = -\frac{1}{kT}.$$

Тогда

$$\frac{\partial \log z}{\partial \omega_l} = -\frac{\partial \mu}{\partial \omega_l} \left(= \frac{1}{kT} \frac{\partial \log T}{\partial \omega_l}\right).$$

Эта зависимость μ от одного из ω_l вычисляется из (6.14), которое может быть переписано в виде:

$$U = \frac{\sum \varepsilon_l \omega_l e^{-\mu\varepsilon_l}}{\sum \omega_l e^{-\mu\varepsilon_l}}.$$

При этом подразумевается, что $U = \text{const}$. Следовательно,

$$d \log U = \frac{ds_1}{s_1} - \frac{ds_0}{s_0} = 0,$$

где мы полагаем для краткости,

$$s_k = \sum \varepsilon_l^k \omega_l e^{-\mu\varepsilon_l}.$$

Далее, варьируя только μ и одну из величин ω_l:

$$ds_1 = \varepsilon_l e^{-\mu\varepsilon_l}\, d\omega_l - s_2\, d\mu, \qquad ds_0 = e^{-\mu\varepsilon_l}\, d\omega_l - s_1\, d\mu.$$

Тогда

$$\frac{\varepsilon_l e^{-\mu\varepsilon_l}\, d\omega_l - s_2\, d\mu}{s_1} - \frac{e^{-\mu\varepsilon_l}\, d\omega_l - s_1\, d\mu}{s_0} = 0.$$

Отсюда мы легко получаем:

$$\frac{\partial \mu}{\partial \omega_l} = \frac{\left(\varepsilon_l - \dfrac{s_1}{s_0}\right) e^{-\mu\varepsilon_l}}{\left(\dfrac{s_2}{s_0} - \dfrac{s_1^2}{s_0^2}\right) s_0}.$$

Принимая во внимание смысл входящих сюда величин, можно написать[1]:

$$\frac{\partial \mu}{\partial \omega_l} = \frac{\varepsilon_l - U}{\widetilde{(\varepsilon_l - U)^2}} \frac{\overline{a}_l}{N}.$$

[1]Волнистая черта указывает на усреднение по членам ансамбля: об этом уже говорилось в конце предыдущей главы.

Таким образом из (6.24) получаем:

$$\overline{a_l^2} - (\overline{a}_l)^2 = \overline{a}_l \Big[1 - \frac{(\varepsilon_l - U)^2}{\widetilde{(\varepsilon_l - U)^2}} \frac{\overline{a}_l}{N} - \frac{a_l}{N} \Big].$$

Будем называть дисперсию нормальной, если средний член равен нулю. Это имеет место для уровней, достигающих средней энергии $(\varepsilon_l - U)$. Во всех прочих случаях дисперсия является, субнормальной[1].

Одной из привлекательных сторон статистической термодинамики является то, что величины и функции, введенные первоначально как чисто математические понятия, почти неизменно приобретают фундаментальный физический смысл. Примером этого являлись множитель Лагранжа μ, максимум z, статистическая сумма (или функция распределения). Каков же смысл $\sum P$? Мы установим его, рассмотрев выражение (6.17). Откинем последний член этого выражения как малый, используем обозначения (6.19) и (6.23) и, вспомнив, что $\mu = \frac{1}{kT}$, получим

$$\frac{1}{N} \log \sum P = \frac{U}{kT} + \log f(z) = \frac{U}{kT} + \frac{1}{k} \left(S - \frac{U}{T} \right) = \frac{1}{k} S.$$

Тогда

$$\frac{k}{N} \log \sum P = S \tag{6.25}$$

будет энтропией отдельной системы. Этот результат уже сам по себе достаточно интересен, однако он оказывается еще более заслуживающим внимания, если вернуться на некоторое время ко второй главе и вычислить логарифм максимума P (мы могли сделать это там же, но не воспользовались этой возможностью). Из (2.2) и формулы Стирлинга имеем:

$$\log P = N(\log N - 1) - \sum_l a_l(\log a_l - 1) = N \log N - \sum_l a_l \log a_l.$$

Пользуясь величинами a_l, соответствующими «максимуму», —

$$a_l = N \frac{e^{-\mu\varepsilon_l}}{\sum\limits_l e^{-\mu\varepsilon_l}}$$

[1] Напомним случай отдельного осциллятора Ферми, где флуктуации $\overline{a}_l$ оказались в точности равными нулю.

и логарифмируя их:

$$\log a_l = \log N - \mu\varepsilon_l - \log \sum_l e^{-\mu\varepsilon_l},$$

получим

$$\log P_{\max} = N\log N - \sum_l a_l \log N + \mu \sum a_l\varepsilon_l + \sum a_l \log \sum e^{-\mu\varepsilon_l} =$$
$$= \mu E + N\log \sum e^{-\mu\varepsilon_l}$$

и, далее,

$$\log P_{\max} = \frac{E}{kT} + N\frac{1}{k}\left(S - \frac{U}{T}\right).$$

Умножая на $\frac{k}{N}$ и вспоминая, что $\frac{E}{N} = U$, получим:

$$\frac{k}{N}\log P_{\max} = \frac{U}{T} + S - \frac{U}{T} = S. \tag{6.26}$$

Сравнивая это с (6.25), видим, что энтропия может быть вычислена с равным успехом либо как $\log\sum P$, либо как $\log P_{\max}$. Дело в том, что хотя число тех P, которые сравнимы с $P_{\max}$, и очень велико, оно, тем не менее, исчезающе мало по сравнению с величиной самого $P_{\max}$. Следовательно, разница логарифмов исчезающе мала. На это обстоятельство указал Г. А. Лорентц в своем знаменитом мемуаре «Нечувствительность термодинамических функций».

Существуют и другие статистические аналогии энтропии, не имеющие, однако, столь общей применимости. Понятие энтропии, выведенное из статистической суммы, приложимо к любой системе как большой, так и малой, как к отдельному осциллятору, так и к газу, твердому телу или гетерофазной системе.

Один из аналогов, на котором следует остановиться (указанный В. Гиббсом), предполагает, что система, будучи помещена в тепловой резервуар, обнаруживает лишь малые флуктуации энергии, что, как мы знаем, свойственно любой большой системе. Заняты лишь уровни, весьма близкие к средней энергии U. Но что же показывает статистическая сумма

$$e^{-\mu\varepsilon_1} + e^{-\mu\varepsilon_2} + \ldots + e^{-\mu\varepsilon_l} + \ldots?$$

Поскольку ε расположены в арифметическом порядке, экспоненты непрерывно убывают. А ведь они являются мерой частоты заполнения!

На первый взгляд кажется удивительным — как может возникнуть острый максимум, да и почему он вообще возникает?

Объяснение этого кроется в характере увеличения ε_l, при увеличении l, а именно — в том, что по мере передвижения по ряду это увеличение становится все более и более медленным, притом с чудовищно возрастающим замедлением. Другими словами, число уровней, приходящихся на единичное приращение ε_l, т. е. плотность уровней колоссально возрастает. Максимум возникает в результате компромисса между возрастающей плотностью уровней и убыванием экспонент.

Обсудим результат с этой точки зрения. Мы можем рассматривать ε_l как функцию индекса l, т. е. как $\varepsilon(l)$, и, следовательно, также и l как функцию ε, $l(\varepsilon)$, выражающую число уровней, достигших предела ε. Рассмотрим теперь

$$l(U),$$

где U — средняя энергия (отклонения от которой очень малы). Тогда

$$k \log l(U)$$

является энтропией.

Нетрудно разобраться в том, что это значит; однако приведем сначала еще одно определение энтропии. Выберем какое-нибудь подходящее малое приращение $\Delta\varepsilon$ и соберем вместе все уровни, лежащие внутри этого приращения, обозначив их число через Δl. Тогда статистическая сумма может быть написана в виде:

$$\sum e^{-\mu\varepsilon} \Delta l,$$

где ε — значение энергии внутри интервала Δl. Мы можем также написать

$$\sum e^{-\mu\varepsilon} \frac{\Delta l}{\Delta\varepsilon} \Delta\varepsilon.$$

Область максимального заполнения, т. е. область, в которой $\varepsilon \sim U$, определяется максимумом подынтегральной функции или, если угодно, ее логарифма

$$-\varepsilon\mu + \log \frac{\Delta l}{\Delta\varepsilon}.$$

Итак,

$$-\mu + \left(\frac{d}{d\varepsilon}\log\frac{\Delta l}{\Delta\varepsilon}\right)_{\varepsilon=U} = 0, \qquad \frac{1}{T} = \left(k\frac{d}{d\varepsilon}\log\frac{\Delta l}{\Delta\varepsilon}\right)_{\varepsilon=U},$$

$$\frac{1}{T} = \left(\frac{dk\log\dfrac{\Delta l}{\Delta\varepsilon}}{d\varepsilon}\right)_{\varepsilon=U}.$$

Отсюда, разумеется, следует, что

$$S = \left(k\log\frac{\Delta l}{\Delta\varepsilon}\right)_{\varepsilon=U}$$

играет роль энтропии.

Причина, по которой мы можем брать саму функцию $l(U)$ вместо

$$\frac{\Delta l}{\Delta\varepsilon} \qquad \text{или} \qquad \frac{dl(U)}{dU},$$

заключается в том, что $l(U)$ практически возрастает всегда как чрезвычайно высокая степень U:

$$l(U) = CU^n, \qquad \frac{dl}{dU} = nCU^{n-1}, \qquad \log l = \log C + n\log U,$$

$$\log\frac{dl}{dU} = \log(nC) + (n-1)\log U.$$

Как видно, практически разницы нет никакой.

Мне хотелось бы указать путь к интуитивному пониманию причины экспоненциальной зависимости частоты заполнения от ε в условиях теплового резервуара.

Пусть $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_l$ и т. д. суть энергетические уровни системы, а $b_1, b_2, \ldots, b_k$ и т. д. — уровни теплового резервуара. Тогда полная энергия (E) (сумма энергии системы и энергии резервуара) будет постоянной, а уровни объединенного целого (т. е. системы и резервуара) будут равны $\varepsilon_l + b_k$.

Поскольку полная энергия постоянна, обмен происходит только между вырожденными уровнями, т. е. в случае

$$\varepsilon_l + b_k = (\text{почти})\,\mathrm{const} = E$$

(«почти» — ввиду наличия энергии связи!). Все эти уровни $\varepsilon_l + b_k$ для всех комбинаций (l, k) имеют, конечно, равную частоту заполнения, что попросту означает равную априорную вероятность для любого отдельного уровня. Причина уменьшения частоты заполнения для более высоких ε_l заключается в том, что число уровней теплового резервуара, обозначаемых через

$$b_k = E - \varepsilon_l,$$

экспоненциально убывает с уменьшением $E - \varepsilon_l$, т. е. с уменьшением энергии, приходящейся на долю теплового резервуара. Это становится совершенно ясным в случае, когда число уровней резервуара равно, например,

$$C(E - \varepsilon_l)^n = CE^n\left(1 - \frac{\varepsilon_l}{E}\right)^n \approx CE^n e^{-\frac{n\varepsilon_l}{E}} \qquad (n - \text{очень велико}).$$

Этот вывод не претендует на строгость (строгий вывод был дан ранее), однако он поясняет суть дела: чем большая часть общей энергии (E) приходится на долю самой системы (ε_l), тем меньше $(E - \varepsilon_l)$ приходится на долю теплового резервуара. А это экспоненциальным образом уменьшает число используемых уровней теплового резервуара (даже в случае бесконечного теплового резервуара или, вернее, именно в этом случае). Характеризующий это уменьшение экспоненциальный множитель является не чем иным, как уже знакомым нам соответствующим членом статистической суммы, другими словами, относительной вероятностью нахождения нашей системы в состоянии ε_l условиях теплового резервуара.