2.8.8. Принцип экстремума энтропии. Историческая справка

Понятие «энтропия» впервые было введено в 1865 г. немецким физиком Клаузиусом как мера необратимого рассеяния энергии. Для обратимых (квазиравновесных) процессов Клаузиус определил

$$\Delta S = \frac{\Delta Q}{T_0}, \tag{2.8.127}$$

где ΔS – изменение энтропии, ΔQ – изменение теплоты, T_0 – абсолютная термодинамическая температура. Оказалось, что неравновесные процессы в изолированных системах сопровождаются ростом энтропии S так, что при приближении к равновесному состоянию она достигает максимального значения S_{max}.

Австрийский ученый Л. Больцман (1844 – 1906), развивая молекулярно-кинетическую теорию газов, связал энтропию Клаузиуса с вероятностью осуществления макроскопического состояния системы с помощью соотношения

$$S = k \ln W, \tag{2.8.128}$$

где k – постоянная Больцмана.

W – вероятность осуществления данного состояния системы:

$$W = \frac{N!}{N_1! N_2! \ldots N_i!},$$

где N – полное число частиц в системе

N_i – число частиц в i-ой ячейке системы.

Говорят, что формула (2.8.128) выгравирована на надгробной плите Людвига Больцмана.

Один из создателей кибернетики американский ученый Клод Шеннон в 1948 г. предложил рассматривать усредненную энтропию системы, состоящую из i элементов

$$\overline{S} = -\sum P_i \log_2 P_i, \tag{2.8.129}$$

где P_i – вероятность i -го элемента системы, такая что

$$\sum P_i = 1 \,. \tag{2.8.130}$$

Позже выяснилось, что функция (2.8.129) обладает интересными экстремальными свойствами. Продемонстрируем эти свойства на простейшем примере. Пусть система состоит из двух взаимоисключающих элементов. Например, монета может выпадать «орлом» с вероятностью P_1 или «решкой» с вероятностью P_2, так что

$$P_1 + P_2 = 1. \tag{2.8.131}$$

При этом энтропия данной системы, согласно (2.8.129) и (2.8.131) имеет следующий вид

$$S = -\left(P_1 \log_2 P_1 + P_2 \log_2 P_2\right) = -\left[P_1 \log_2 P_1 + (1-P_1)\log_2(1-P_1)\right]. \tag{2.8.132}$$

График функции (2.8.132) представлен на рис. 2.8.4. Из этого графика видно, что максимум усредненной энтропии приходится на случай $P_1 = P_2 = ½$ [47]. Это соответствует наименее «определенному» и при этом наиболее вероятному случаю.

Усредненную энтропию вида (2.8.129) часто называют Шенноновской энтропией и рассматривают ее в качестве меры неопределенности.

Пусть, например, место нахождения хаотически блуждающей частицы определяется ПРВ $\rho(x)$. В современной статистической физике мера неопределенности (т. е. шеноновская энтропия) такой системы задается интегральным выражением вида

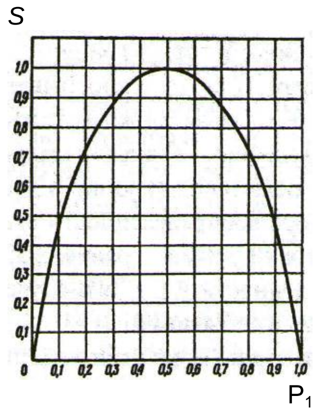

Рис. 2.8.4. Зависимость усредненной энтропии от вероятности P_1 [47]

$$\overline{S} = H(x) = -\int_{-\infty}^{\infty} \rho(x) \ln \rho(x) dx \,. \tag{2.8.133}$$

Найдем функцию ПРВ $\rho(x)$, при которой усредненная энтропия достигает экстремального значения. Поставленная задача сводится к нахождению экстремали функционала (2.8.133).

Для решения этой задачи введем условия общего характера [48]:

1) математическое ожидание случайного процесса x задается выражением

$$\int_{-\infty}^{\infty} x\rho(x)dx = a; \qquad (2.8.134)$$

2) дисперсия того же процесса задается выражением

$$\int_{-\infty}^{\infty} (x-a)^2 \rho(x)dx = \sigma_x^2; \qquad (2.8.135)$$

3) имеет место калибровочное соотношение

$$\int_{-\infty}^{\infty} \rho(x)dx = 1. \qquad (2.8.136)$$

Согласно теоремам вариационного исчисления, с учетом условий (2.8.134) – (2.8.136), будем искать экстремум функционала [48]

$$\int_{-\infty}^{\infty} F[x,\rho(x)]dx = \int_{-\infty}^{\infty} \left[-\rho(x)\ln\rho(x) + \lambda_1 (x-a)^2 \rho(x) + \lambda_2 x\rho(x) + \lambda_3 \rho(x)\right] dx, \qquad (2.8.137)$$

где λ_1, λ_2, λ_3 – неизвестные константы.

ПРВ $\rho\,(x)$, при которой функционал (2.8.137) достигает экстремального значения, определяется на основании уравнения Эйлера [48]

$$\frac{\partial F[x,\rho(x)]}{\partial\rho(x)} = -1 - \ln\rho(x) + \lambda_1 (x-a)^2 + \lambda_2 x + \lambda_3 = 0\,. \qquad (2.8.138)$$

Решение этого уравнения имеет вид [48]

$$\rho(x) = \exp\left\{\lambda_1 (x-a)^2 + \lambda_2 x + \lambda_3 - 1\right\}. \qquad (2.8.139)$$

Для определения констант λ_1, λ_2, λ_3 воспользуемся тремя условиями (2.8.134) – (2.8.136). Из условия (2.8.136) следует

$$\int_{-\infty}^{\infty} \rho(x)dx = \exp(\lambda_3 - 1)\int_{-\infty}^{\infty} \exp\left\{\lambda_1 (x-a)^2 + \lambda_2 x\right\}dx = 1\,. \qquad (2.8.140)$$

Для сходимости данного интеграла компонента λ_1 должна быть отрицательной. Делая замену переменной $y = (x - a)$, запишем [48]

$$\int_{-\infty}^{\infty} \exp\left\{\lambda_1 (x-a)^2 + \lambda_2 x\right\} dx = \exp(\lambda_2 a) \int_{-\infty}^{\infty} \exp\left\{\lambda_1 y^2 + \lambda_2 y\right\} dy = \sqrt{\frac{\pi}{-\lambda_1}} \exp\{\lambda_2 a\} \exp\left\{\frac{-\lambda_2^2}{4\lambda_1}\right\}. \tag{2.8.141}$$

Из (2.8.141) получаем

$$\exp(\lambda_3 - 1) = \sqrt{\frac{-\lambda_1}{\pi}} \exp\left(\frac{\lambda_2^2}{4\lambda_1} - \lambda_2 a\right). \tag{2.8.142}$$

Воспользуемся теперь условием (2.8.135), которое после несложных преобразований приводится к уравнению вида

$$\frac{1}{-\lambda_1 \sqrt{\pi}} \left[\frac{\sqrt{\pi}}{2} + \frac{\lambda_2^2 \sqrt{\pi}}{(-4\lambda_1)} + \frac{\lambda_2}{2\sqrt{-\lambda_1}} \int_{-\infty}^{\infty} e^{-z} dz \right] = \sigma_x^2. \tag{2.8.143}$$

Для конечности левой части уравнения (2.8.141) константа λ_2 должна быть равной нулю. Тогда

$$\lambda_1 = -\frac{1}{2} \sigma_x^{-2}. \tag{2.8.144}$$

При $\lambda_2 = 0$ выражение (2.8.142) приобретает следующий вид

$$\exp(\lambda_3 - 1) = \sqrt{\frac{-\lambda_1}{\pi}}. \tag{2.8.145}$$

После подстановки (2.8.144) и (2.8.145) в уравнение (2.8.139) обнаруживаем, что экстремалью функционала (2.8.133) при условиях (2.8.134) – (2.8.136) является гауссовская плотность распределения вероятности [48]

$$\rho(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left\{-\frac{(x-a)^2}{2\sigma_x^2}\right\}. \tag{2.8.146}$$

Таким образом, максимум усредненной энтропии рассматриваемой системы при выше указанных условиях достигается при гауссовской ПРВ, которая играет значительную роль во многих приложениях статистической физики.