

УДК 517.393

МАТЕМАТИКА

Г. М. ПИГОЛКИН

ЭНТРОПИЙНОЕ НЕРАВЕНСТВО И КРИТЕРИЙ СУЩЕСТВОВАНИЯ СТАТИСТИЧЕСКОГО РАВНОВЕСИЯ

(Представлено академиком М. А. Лаврентьевым 22 VI 1971)

Из статистической механики известно, что энтропия замкнутой системы может разве лишь возрастать (широко известная H -теорема Больцмана). При этом возникает вопрос, при каких условиях энтропия не может возрастать неограниченно. Такие условия (необходимые и достаточные) назовем критерием существования статистического равновесия. В литературе по статистической механике при выводе условий существования равновесия обычно считается, что внешний потенциал обладает некоторыми дифференциальными свойствами. Такое требование является слишком сильным для того, чтобы получить необходимые и достаточные условия, т. е. критерий существования равновесия. Мы получим этот критерий (теорема 3) с помощью одного информационного неравенства ((²), стр. 49, теорема 2.1 и ссылки там же), которое в требуемой форме устанавливает

Теорема 1. Пусть на множестве $\mathcal{D} \subset R^n$ задана измеримая функция g .

Тогда для любой неотрицательной суммируемой на $\mathcal{D}$ функции $f \not\equiv 0$ справедливо неравенство

$$-\int_{\mathcal{D}} f \ln f \, dx \leq \int_{\mathcal{D}} g f \, dx + \int_{\mathcal{D}} f \, dx \ln \left(\int_{\mathcal{D}} e^{-g} \, dx \Big/ \int_{\mathcal{D}} f \, dx \right); \quad (1)$$

$$f(x) \geq 0, \quad x \in \mathcal{D}.$$

Знак равенства в (1) достигается, разве лишь если $f(x) = C e^{-g(x)}$, где константа $C > 0$.

Неравенство (1) имеет смысл, если обе его части существуют. При этом функцию g можно выбрать так, что знак равенства в (1) не будет достигаться ни на какой $f \not\equiv 0$.

Это несколько отличает нашу формулировку от указанной теоремы из теории информации. Приведем коротко независимое от (²) доказательство сформулированной теоремы.

Формально можем записать неравенство Гёльдера в виде (¹)

$$\left(\int_{\mathcal{D}} f^{1/p} \, dx \right)^p \equiv \left[\int_{\mathcal{D}} \frac{f^{1/p} (1 + \gamma g)^{1/p}}{(1 + \gamma g)^{1/p}} \, dx \right]^p \leq \left[\int_{\mathcal{D}} (1 + \gamma g)^{-1/(p-1)} \, dx \right] \times$$

$$\times \left(\int_{\mathcal{D}} f \, dx + \gamma \int_{\mathcal{D}} g f \, dx \right), \quad (2)$$

$$p > 1; \quad 1 + \gamma g(x) > 0, \quad x \in \mathcal{D}.$$

Предположим сначала, что на $\mathcal{D}$ выполнены условия

$$|g(x)| \leq b, \quad 1/a \leq f(x) \leq a, \quad a > 1. \quad (3)$$

Запишем (2) при

$$p = 1/(1 - \gamma); \quad 0 < \gamma < \min \{1; 1/(2b)\}.$$

Обе части полученного неравенства разделим на величину $0 < \int_{\mathcal{D}} f dx < \infty$

и затем возведем в степень, равную $(1 - \gamma) / \gamma$. После этого перейдем к пределу при $\gamma \rightarrow +0$ и прологарифмируем предельное неравенство. В результате получим (1). Здесь, при выполнении (3), предельный переход не вызывает затруднений.

При фиксированной (временно) f условия (3) определяют множество $\mathcal{D}(a, b) \subset \mathcal{D}$, причем при $f \not\equiv 0$ всегда найдутся $a_0 > 1$, $b_0 > 0$ такие, что

$$\int_{\mathcal{D}(a, b)} f dx \geq C_0 \geq 0 \quad \text{при } a \geq a_0, b \geq b_0.$$

По доказанному, можем записать (1), подставив $\mathcal{D}(a, b)$ на место $\mathcal{D}$. Переходя последовательно к пределу по a и b — сначала $a \rightarrow \infty$ (b фиксировано), затем $b \rightarrow \infty$, — найдем (1) в окончательном виде.

Для установления всех случаев достижения равенства, не нарушив общности, положим

$$\int_{\mathcal{D}} f dx = 1. \quad (4)$$

Величина $f \ln f + gf$ есть выпуклая функция от f . Отсюда равенство в (1) достигается разве лишь при $f = ce^{-g}$ (если $f \not\equiv 0$), и теорема доказана.

Измеримую неотрицательную функцию, удовлетворяющую нормировочному равенству (4), называем плотностью на $\mathcal{D}$ (или просто плотностью). Если f — плотность, то левая часть (1) называется энтропией в соответствии с терминологией статистической механики. В связи с этим, называем (1) энтропийным неравенством.

Пусть фиксирован набор измеримых на $\mathcal{D}$ функций: $\rho = \rho(x) = \{\rho_1(x), \dots, \rho_m(x)\}$, $x \in \mathcal{D}$. В совокупности эти функции образуют вектор-вес. Пишем $Z \equiv \{\langle \rho \rangle(f)\}$ и говорим, что Z есть среднее, если найдется плотность f такая, что

$$Z = \langle \rho \rangle(f) \equiv \int_{\mathcal{D}} \rho f dx, \quad f \geq 0, \quad \int_{\mathcal{D}} f dx = 1.$$

Легко видеть, что множество $\{\langle \rho \rangle(f)\}$ выпукло в R^m . Его замыкание $\overline{\{\langle \rho \rangle(f)\}}$ есть выпуклое тело в R^m , если и только если вектор-вес вполне нетривиален. Последнее означает выполнение условия

$$\sum_{i=1}^m \lambda_i \rho_i(x) \not\equiv \text{const}, \quad x \in \mathcal{D}, \quad |\lambda| \neq 0, \quad (5)$$

при произвольно фиксированном ненулевом векторе $\lambda \in R^m$.

Теорема 2.

Если вектор-вес ρ вполне нетривиален (5) и выполнены условия

$$0 < \int_{\mathcal{D}} dx < \infty, \quad |\rho(x)| \leq c < \infty, \quad x \in \mathcal{D}, \quad (6)$$

то канонические плотности

$$f_\lambda \equiv \frac{e^{-(\lambda, \rho(x))}}{\int_{\mathcal{D}} e^{-(\lambda, \rho)} dx}, \quad x \in \mathcal{D}; \quad (\lambda, \rho) \equiv \sum_{i=1}^m \lambda_i \rho_i, \quad \lambda \in R^m, \quad (7)$$

обладают тем свойством, что средние $\langle \rho \rangle_\lambda \equiv \langle \rho \rangle(f_\lambda)$ пробегают все внутренние точки выпуклого тела $\overline{\{\langle \rho \rangle(f)\}}$, когда вектор-параметр λ пробегает всё R^m . При этом зависимость между λ и $\langle \rho \rangle_\lambda$ взаимно однозначна.

При доказательстве первого утверждения теоремы в основе лежит простая оценка, которая имеет место при условии (6) и указывает на то, что при $\lambda = \lambda^{(0)}t$, $|\lambda^{(0)}| = 1$, $t \rightarrow +\infty$ предельное значение $\langle \rho \rangle_\lambda$ лежит на

опорной к $\{\langle \rho \rangle(f)\}$ гиперплоскости. Взаимная однозначность зависимости между λ и $\langle \rho \rangle_\lambda$ следует из строгой выпуклости функции

$$\varphi(t) \equiv \ln \int_{\mathcal{Q}} e^{-(\lambda_t \rho)} dx, \quad \lambda_t \equiv \lambda' + t(\lambda'' - \lambda'), \quad 0 \leq t \leq 1,$$

при любых $\lambda' \in R^m$, $\lambda'' \in R^m$, $\lambda' \neq \lambda''$.

Плотность (7), если она существует, всегда называем канонической.

Пусть среди плотностей, удовлетворяющих условию $Z = \int_{\mathcal{Q}} \rho f dx$, где $Z \in R^m$ фиксировано, найдется плотность, доставляющая максимум энтропии $S(f) \equiv - \int_{\mathcal{Q}} f \ln f dx$. Такую плотность называем равновесной и говорим, что равновесие существует.

Теорема 3. Если $Z = \int_{\mathcal{Q}} \rho f dx$ фиксировано во внутренней точке выпуклого тела $\{\langle \rho \rangle(f)\}$ (вес ρ вполне нетривиален (5)), то равновесие существует тогда и только тогда, когда уравнение относительно $\lambda \in R^m$

$$\int_{\mathcal{Q}} (\rho(x) - Z) e^{-(\lambda \rho(x))} dx = 0 \quad (8)$$

разрешимо.

В этом случае равновесная плотность единственна и совпадает с канонической f_λ , где λ есть решение (8).

Если уравнение (8) не имеет решения, то найдется последовательность плотностей $f^{(k)}$ таких, что $Z = \langle \rho \rangle(f^{(k)})$, а последовательность значений энтропии $S(f^{(k)})$ при $k \rightarrow \infty$ неограниченно возрастает. При этом $f^{(k)}$ могут быть выбраны так, что каждая из них является квазиканонической, т. е. на своем носителе пропорциональна величине $\exp\{-(\lambda \rho(x))\}$ при некотором $\lambda \in R^m$.

Доказательство следует из предыдущих теорем.

Последнюю теорему можно интерпретировать как критерий существования статистического равновесия механической системы при заданных макроскопических характеристиках этой системы. Случай, когда Z есть граничная точка множества средних, сводится к рассмотренному, с понижением размерности вектор-веса.

Институт гидродинамики
Сибирского отделения Академии наук СССР
Новосибирск

Поступило
17 VI 1971

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

- ¹ Г. Г. Харди, Д. Е. Литтлвуд, Г. Полиа, Неравенства, ИЛ, 1948.
- ² С. К у л ь б а к, Теория информации и статистики, «Наука», 1967.