中国科学 G辑: 物理学 力学 天文学 2008年 第38卷 第6期: 781 ~ 784
www.scichina.com phys.scichina.com
《中国科学》杂志社
SCIENCE IN CHINA PRESS

热力学第二定律的一个普遍的信息论证明

张启仁

北京大学技术物理系, 北京 100871
E-mail: zhangqr@pku.edu.cn

收稿日期: 2008-01-07; 接受日期: 2008-03-25
国家自然科学基金项目资助(批准号: 10305001)

摘要 指出信息量的守恒及其不可加性以及熵的可加性导致孤立系统的熵增加. 量子纠缠态及其塌缩提供了信息量不可加性的一个例子. 这种不可加性在相互作用信息重要的其他领域, 例如在经典统计和社会统计中, 也是真的. 这便在极普遍的情形中证明了热力学第二定律. 它不仅在量子统计和经典统计中, 而且在信息量守恒和不可加的其他过程中, 是精确成立的.

关键词
信息量守恒
信息量的不可加性
熵增加

了解热力学第二定律的理论基础是长期以来物理学要解决的一个重大问题. 从教科书[1~3]中得知, 熵越高的宏观系统实现的可能性越大. 然而, 系统是否总是从可能性较低的状态过渡到可能性较高的状态仍是一个未解决的问题. 玻尔兹曼的H-定理是关于趋向平衡的一个经典证明. 它基于一个将宏观物质当作互相碰撞的经典粒子系统的模型, 因此, 即使从经典统计物理看也是不够普遍的. 20 世纪 40 年代末, Shannon等人[4]提出和发展了强有力的信息论, 可用于所有统计过程的分析. 下面会看到, 用信息论的基本概念可简单地普遍证明热力学第二定律. 在量子统计力学中, 这基于随时间演化过程中系统各部分之间的态纠缠和测量中的态塌缩, 因此是普遍的. 孤立系统密度算符$\rho(t)$的时间演化遵从von Neumann方程. 它的解为

$$\rho(t) = U(t,t_0)\rho(t_0)U(t_0,t), \tag{1}$$

其中$U(t,t_0)$为态从时刻t_0到时刻t的演化算符. 定义t时刻的信息量

$$I(t) \equiv \mathrm{tr}\,[\rho(t)\ln\rho(t)]. \tag{2}$$

鉴于$U(t_0,t)U(t,t_0) = 1$, 由(1)式可见

$$\begin{aligned} I(t) &= \mathrm{tr}[U(t,t_0)\rho(t_0)\ln\rho(t_0)U(t_0,\mathrm{t})] \\ &= \mathrm{tr}[\rho(t_0)\ln\rho(t_0)U(t_0,t)U(t,t_0)] \\ &= I(t_0), \end{aligned} \tag{3}$$

此为量子力学中的信息量守恒[5]. 这是物理学中一条精确的守恒律. 下面以此律为基础证明热力学第二定律.

为测一个系统的熵，要将此系统分成宏观看来无穷小的诸部分，其中第i部分的熵定义为$S_i \equiv -k\mathrm{tr}(\rho_i \ln\rho_i)$, ρ_i是第i部分的约化密度算符. 整个系统的熵，作为广延量，定义为和

$$S = \Sigma_i S_i = -k\Sigma_i \,\mathrm{tr}(\rho_i \ln\rho_i). \tag{4}$$

当t_0时刻测量系统的熵时，会破坏系统各部分间态的纠缠(破坏各部分间的关联). 系统的态和密度算符都因此而因子化. 在此条件下，系统的熵为

$$\begin{aligned} S(t_0) &= -k\Sigma_i \,\mathrm{tr}[\rho_i(t_0)\ln \rho_i(t_0)] \\ &= -k\,\mathrm{tr}[\rho(t_0)\ln\rho(t_0)] = -k\, I\,(t_0). \end{aligned} \tag{5}$$

对一个孤立系统，信息量守恒(3)式成立. 系统在$t > t_0$时刻的信息量因此为

$$I\,(t) = -S(t_0)/k. \tag{6}$$

在从时刻t_0到t的期间，系统不同部分间的作用使它们的态再纠缠起来. 即不同部分的态互相关联. 若在t时刻测系统的熵，就要在此时刻测系统各部分的熵，然后将它们加起来，这就又一次破坏了这种纠缠. 这便是态的塌缩. 它导致关联信息的丢失. 系统诸部分各自并不孤立，它们各自的信息量并不守恒. 这使t时刻的熵

$$S(t) = -\,k\Sigma_i \,\mathrm{tr}[\rho_i(t)\ln\rho_i(t)]. \tag{7}$$

一般不等于$S(t_0)$. 凭直观判断，系统各部分信息量之和不应超过系统的信息量，因其中不含各部分间的关联信息，即

$$\Sigma_i \,\mathrm{tr}[\rho_i(t)\ln\rho_i(t)] \leqslant I\,(t), \tag{8}$$

若果如此，则按(6)~(8)式，对孤立系统有

$$S(t) \geqslant S(t_0). \tag{9}$$

为证明关系式(8)，要引用一些数学不等式，它们可在教科书[5,6]中找到. 约定$0\ln0 \equiv \lim_{\xi\to 0}\xi\ln\xi = 0$，它们是下面 4 个引理.

引理 1　对非负变量 x 有

$$x \ln x \geqslant x - 1, \tag{10}$$

等式对且仅对 $x = 1$ 成立.

引理 2　对非负数集$[w_i]$和满足$\Sigma_i x_i = 1$ 的非负数集$[x_i]$，有

$$\Sigma_i x_i w_i \ln\Sigma_j x_j w_j \leqslant \Sigma_i x_i w_i \ln w_i. \tag{11}$$

引理 3　对满足

$$\Sigma_i W_i = 1 \quad 和 \quad \Sigma_i T_{ij} = \Sigma_j T_{ij} = 1 \tag{12}$$

的非负数集$[W_i]$和$[T_{ij}]$有

$$W'_j \equiv \Sigma_i W_i T_{ij} \geqslant 0, \quad (对每一j) \tag{13}$$

$$\Sigma_j W'_j = 1 \tag{14}$$

和

$$\Sigma_j W'_j \ln W'_j \leqslant \Sigma_i W_i \ln W_i. \tag{15}$$

引理 4　对满足$\Sigma_{ij} W_{ij} = 1$ 的正数$[W_{ij}]$，及正数$W_i = \Sigma_j W_{ij}$和$W'_j = \Sigma_i W_{ij}$，有

$$\Sigma_i W_i = 1, \quad \Sigma_j W'_j = 1 \tag{16}$$

和

$$\Sigma_{ij} W_{ij} \ln W_{ij} \geqslant \Sigma_i W_i \ln W_i + \Sigma_j W_j \ln W'_j, \tag{17}$$

末式中等号当且仅当对所有ij均可作因式分解 $W_{ij} = W_i W'_j$时成立.

设$[L]$为系统的一组完备的能同时确定的力学量. 其完备正交归一的本征态系为$[|n\rangle]$. 密度算符ρ的$[L]$象表象为一矩阵, 矩阵元为$\rho_{n,n'} = \langle n|\rho|n'\rangle$. 若$\rho$自身也包含在完备集$[L]$中, 则$\rho$的$[L]$表象就称为自然表象. 在自然表象中, 密度矩阵是对角的: $\rho_{n,n'} = W_n\delta_{n,n'}$, 其中$W_n$是$\rho$的第$n$本征值, 表示发现系统处于态$|n\rangle$的概率. 信息量(2)式可用一非负数集$[W_n]$表为

$$I = \Sigma_n W_n \ln W_n. \tag{18}$$

也可考虑关于一组特定的完备力学量$[L']$的信息量, 其完备正交归一的本征态系记为$[|m\rangle]$. 对系统的一个密度算符为ρ的系综, 发现系统处于态$|m\rangle$的概率为

$$W'_m = \Sigma_n\langle m|n\rangle W_n\langle n|m\rangle. \tag{19}$$

关于力学量组$[L']$的信息量的定义为

$$I_{[L']} \equiv \Sigma_m W'_m \ln W'_m. \tag{20}$$

由于$|\langle n|m\rangle|^2$非负, 且$\Sigma_n |\langle n|m\rangle|^2 = \Sigma_m |\langle n|m\rangle|^2 = 1$, 按引理 3 和(18)式得

$$I_{[L']} \leqslant I. \tag{21}$$

现在将系统分成两部分a和b. 令$i = a$或b, 设$[L_i]$为i部分的一组完备力学量, $|n_i\rangle$ 为它们的第n_i本征态, 而$[|n_i\rangle]$为i部分的一组完备正交归一的态. 因此$[|n_a n_b\rangle] \equiv [|n_a\rangle|n_b\rangle]$ 是系统的一组完备正交归一的态. 在$[L_\mathrm{a}, L_\mathrm{b}]$表象中, 系统的密度算符为一矩阵, 矩阵元为

$$\rho_{n_a n_b,\, n'_a n'_b} \equiv \langle n_a n_b|\rho|n'_a\, n'_b\rangle. \tag{22}$$

由(19)式知, 发现a部分处于态$|n_a\rangle$且b部分处于态$|n_b\rangle$的概率为

$$W_{n_a n_b} = \Sigma_n\langle n_a n_b|n\rangle W_n\langle n|n_a n_b\rangle, \tag{23}$$

归一化为

$$\Sigma_{n_a n_b} W_{n_a n_b} = 1. \tag{24}$$

关于力学量$[L_a, L_b]$的信息量为

$$I_{[L_a, L_b]} = \Sigma_{n_a n_b} W_{n_a n_b} \ln W_{n_a n_b} \leqslant I. \tag{25}$$

发现a部分处于态$|n_a\rangle$的概率和发现b部分处于态$|n_b\rangle$的概率分别为

$$W_{n_a} = \Sigma_{n_b} W_{n_a n_b} \quad \text{和} \quad W'_{n_b} = \Sigma_{n_a} W_{n_a n_b} \tag{26}$$

在(24)~(26)式中求和可只对$W_{n_a n_b}>0$ 的n_a和n_b进行.

a部分的密度算符ρ_a由系统的密度算符ρ约化而来. 在$[L_a]$表象中, 它是一个元为

$$(\rho_a)_{n_a n'_a} = \Sigma_{n_b}\rho_{n_a n_b,\, n'_a n_b} = \Sigma_{n_b}\langle n_a n_b|\rho|n'_a\, n_b\rangle \tag{27}$$

的矩阵. 可将它紧凑地写成

$$\rho_a = \mathrm{tr}_b\, \rho, \tag{28}$$

下标b表示阵迹中求和遍及对b部分量子数对角的矩阵元. 类似地, $\rho_b = \mathrm{tr}_a\, \rho$. 设$\rho_i$包含在集$[L_i]$中, 发现$i$部分处于态$|n_i\rangle$中的概率为它的本征值$W_{n_i}$, 且可由(26)式表示. 关于$i$部分的信息量为

$$I_i = \mathrm{tr}\,\rho_i \ln \rho_i = \Sigma_{n_i} W_{n_i} \ln W_{n_i}. \tag{29}$$

由引理 4 和方程(25)和(26)知

$$\mathrm{tr}\rho_a \ln\rho_a + \mathrm{tr}\rho_b \ln\rho_b \leqslant \mathrm{tr}\rho\ln\rho. \tag{30}$$

等号当且仅当系统的密度算符可分解为其各部分的密度算符的直积时成立, 即当且仅当其各部分彼此互不关联时成立. 进一步细分系统的各部分并将(30)式一再地用于它们, 结果便是(8)

式. 如前所述, 此结果导致(9)式, 即

定理: 一个孤立系统熵的变化只能增加不能减少.

这正是热力学第二定律. 此定律因而得证. 按前面所引宏观态出现的概率与熵的关系[1~3], 这便转而证明了: 一个孤立的宏观系统总是从可能性较低的状态过渡到可能性较高的状态.

这里的证明是普遍的. 它似乎依赖于量子力学的态纠缠和态塌缩, 实际上比这更普遍. 这是一个信息论证明, 仅依赖于信息量的守恒(3)式和它的不可加性(30)式. 一个量的不可加性即系统的总量一般不等于其各部分量之和的性质. 熵的广延性(4)式, 即其可加性, 在证明中也是重要的. 信息量守恒是动力学的一种特征. 量子动力学和经典动力学都具此特征. 其他(包括尚未发现的)动力学也可具此特征. 信息量的不可加性是纯数学的, 可用前述数学不等式从概率的普遍关系(26)式导得. 态纠缠及其塌缩只是实现以上诸点的特殊方式. 它们也可在经典力学或某种未知力学中实现. 热力学第二定律因此是极其普遍,几乎与动力学无关的, 只要信息量守恒成立. 即使有一天人们发现量子力学也只是近似的, 必须由某种新力学代替, 只要新力学还遵从信息量守恒, 热力学第二定律便依然严格成立. 此定律也是广泛适用的, 不仅适用于热力学, 且适用于一切信息量守恒的统计科学. 考虑将它用于社会学和经济学的可能性是有趣的.

从证明中还可领悟到: 孤立系统熵增加仅仅由于丢失了系统各部分间的关联信息. 这提示发展重视关联信息的统计学科, 其中相当于热力学第二定律的内容将直接表现为信息量守恒及信息形态的转化.

参考文献

1 Landau L D, Lifshitz E M. Statistical Physics, 3rd edition. Oxford: Butterworth Heinemann, 1980

2 Greiner W, Neise L, Stöcker H. Thermodynamics and Statistical Mechanics. Heidelberg: Springer-Verlag, Inc, 1995

3 张启仁. 统计力学. 北京: 科学出版社, 2004

4 Shannon C E, Weaver W. The Mathematical Theory of Communication. Urbana: University of Illinois Press, 1949. I11

5 De Witt B S, Graham N. The Many-Worlds Interpretation of Quantum Mechanics. Princeton: Princeton University Press, 1973

6 Jaynes E T. Probability Theory: The Logic of Science. Cambridge: Cambridge University Press, 2003