· 388 · 化学通报 2013 年 第 76 卷 第 5 期 http://www.hxtb.org

专家论坛

热力学的公理体系
——论“凭什么相信计算”之一

陈敏伯
（中国科学院上海有机化学研究所 上海 200032）

摘 要 实验结果拟合得到的规律即唯象理论，它们的例外总是容易找到的。可是，为什么唯象理论中唯独热力学几乎找不到例外呢？显然，要从理论的“基础”来考察，理论的“基础”通常指理论是否具有公理化的形式体系。具有公理结构的理论就几乎不可能找到例外。100 年多前物理学通过演绎法摆脱了“Russell 鸡”单一归纳的逻辑，现在应该是将化学建立在完整逻辑（即归纳加演绎的逻辑）基础上的时候了。本文介绍热力学的公理结构，以此提醒理论或概念的“基础”是很重要的，以实验现象的归纳为唯一依据得到的唯象理论很有可能最后被发现是建立在沙滩上的。

关键词 热力学 公理形式体系 基础

Axiomatic Thermodynamics
——On “What do We Believe Calculations on?” - I.

Chen Minbo
(Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 200032)

Abstract It is always easy to find exceptions of the phenomenological theories, which are the rules or laws obtained directly from experiments. Although thermodynamics belongs phenomenological theory, it is impossible to find its exception. Obviously, the “foundation” of thermodynamics should be examined, namely its axiomatic formalism. Physicists had get rid of pure inductive way of thinking, i. e. the way of Russell's chicken, by introduction of deduction many years ago. It is now the turn for chemists. This article describes the axiomatic formalism of the thermodynamics, and reminds the importance of the “foundation” of any theory, and otherwise some theories on which our thinking is based will eventually be found to be built on sand.

Keywords Thermodynamics, Axiomatic formalism, Foundation.

1 唯象理论和热力学

从 19 世纪中叶到 20 世纪初建立的热力学三大定律，发现了能量和熵两个重要概念，于是成为整个物理学、化学的重要基础理论之一。直至今日，化学家每天的思考都离不开热力学。

热力学是从大量宏观经验中总结出来的规律，本质上是唯象的，属于唯象理论（phenomenological theory）。什么是唯象理论呢？自然科学首先是实验科学，自 Francis Bacon 以来，把从受控制的实验中通过归纳法（induction）得到的规律称为唯象理论[1,2]①。例如，反应速度的 Arrhenius 定律、反应速度常数或平衡常数中的 Hammett 规则、Pauling 电负性公式、电学中的 Ohm 定律、扩散中的 Fick 定律、导热中的 Fourier 定律、流体粘度的 Newton 定律、……都是唯象理论[3]，它们都是从实验结果的拟合归纳而得到的。唯象理论构成了自然科学理论的起点，但必须强调它仅仅是起点，通常只适用于那些实验涉及的

陈敏伯 男，研究员，博士生导师，长期从事理论和计算化学研究。E-mail：mbchen@mail.sioc.ac.cn

2012-12-04 收稿，2013-01-14 再修回，2013-01-15 接受

① 笔者极力避免使用疑似来自哲学的词汇，但为节省篇幅有时还得使用。

范围,所以每个唯象理论都存在例外,换言之,它们都有适用的范围。唯象理论是探索真理途中的中间站,但却往往不是真理本身,甚至距离真理还相当远。

虽然热力学本质上是唯象的,但是,热力学预言的两种永动机(创造能量的第一类永动机和从单一热源把热量无偿转化为功的第二类永动机)在宏观范围内是绝对无法实现的。只有当延拓到宇宙尺度或原子核尺度之下之类非常极端的情况,不排除有热力学例外的可能。当然,此时已经不在化学家关心的范围内了。总之,可以说在宏观实验的条件下,同样属于唯象理论的热力学是找不到例外的,而其他唯象理论只要在它赖以形成的实验范围之外是可以轻而易举地找到例外的。于是,人们就会想到热力学的"基础(foundation)"①应该要远比其他的唯象理论严谨。

为什么唯象理论中唯独热力学的基础有如此质的不同呢?根源要归结于热力学内蕴的公理化结构。尽管各个理论的公理化程度有异,但还是可以说热力学有公理体系,而其他唯象理论没有。很多化学理论在建立的时候或直到运用的时候,就没有公理化的意识,没有想到与统一的第一原理接轨。

2 公理化

什么是公理化呢?公理就是不言自明的公认道理。公理化是将知识重组的过程,从尽可能少的原始概念和公理出发,按照逻辑规则推导出所有其他知识,从而形成一个演绎说理的知识体系。最早的公理体系是公元前 3 世纪古希腊 Euclid 写的《几何原本》一书中的平面几何理论。Euclid 选定的公理包括 5 条,即熟知的"两点之间能且只能连一条直线;两条相交的直线只能有一个交点;……"。由此公理集出发,演绎出平面几何的所有定理,成为一个公理体系。所有推翻其中任意一个结论的企图,都必将归结为对 Euclid 公理集中某个公理的否定,绝无其他途径可寻。

19 世纪末,数学家 D Hilbert(1862 ~ 1943)把 Euclid 朴素的公理化思想进一步提高到现代数学的高度。他认为,Euclid 公理体系中关于点、直线和平面的具体定义并不重要,可以抛弃其字面内容,只剩下纯粹的逻辑形式,而从推广意义上来理解。Hilbert 构造了更高层次的公理体系,称为形式公理学。他指出,公理体系需要满足三点要求:协调性、独立性和完备性。协调性是指公理集中的公理之间及它们的推论之间都不存在矛盾;独立性是指公理之间无变换关系即无重叠关系;完备性则是指由此公理集推得的演绎体系能够完整地描述该学科领域的所有问题[4]。

Hilbert 公理体系的意义远不止于数学领域,而是在科学方法论上给整个自然科学起到了示范作用。Hilbert 建议用数学公理化的思想改造整个物理学,将整个物理学理论公理化②。于是,从 19 世纪末到 20 世纪 30 年代世界科学界出现了一场公理化运动,浩浩荡荡。尽管,1930 年 K Gödel 证明了"不完备定理",打破了 Hilbert 使物理学绝对公理化的美梦。但是,务实的自然科学家们并不在乎绝对真理无法达到,对绝对真理的趋近就会使他们心满意足。卅年公理化运动硕果累累,达到了前所未有的高度,公理化运动的车轮没有因此而停下来。Hilbert 的学生 von Neumann 在 1927 年完成了量子力学的公理化[6]。苏联的 Kolmogoroff 在 1933 年完成概率论的公理化。统计力学也建立了它的公理体系[7,8]。"基础(Foundation)"已成为理论的重要组成部分。

Hilbert 的学生 C Carathéodory 在 1909 年首先着手建立公理化热力学[9,10]。后来经 H Callen、L Tisza 等的不断努力[11,12],将宏观热力学理论表述为极为简单而紧凑的公理化形式,以致于各种热力学势之间的关系特别明显。以下作简要介绍。

这里先假设读者已经了解热力学中各种物理概念,如平衡态、内能、熵、压强、体积、化学势、……等。不管你以前是如何理解整体热力学的,现在开始,把这些概念当构件按照公理化的逻辑重新构建新的理论框架。由此展现的理论框架称为公理化的热力学。

---

① 目前化学界很少特意关心理论的基础,基础是比理论更深入的部分。热力学理论的基础是它内蕴的公理结构。

② 正如李政道先生所言"只有一个物理学"[5],是指物理科学只有一个统一原理。化学是电子、原子、分子直至分子聚集体、纳米聚集体层次的物理科学,在此意义上,其实化学也只有一个。

3 公理化的热力学

公理化热力学从由下面 4 条公理假设组成的公理集出发,然后经演绎最后得到热力学的全部结果[11,12]。

公理 1(平衡态公理):宏观均匀体系的平衡态总可以用体系的内能 U 和一组广延量($X_1,\cdots,X_c,X_{c+1},X_{c+2}$) $\equiv\{X_j\}$ 来作体系的完备描述,其中 c 为体系组份数,平衡态的温度必为正值。

实际上,这里的广延量 $X_1,\cdots,X_c,X_{c+1},X_{c+2}$ 分别就是各组分的摩尔数 $N_1,\cdots,N_c$、内能 U 和体积 V。其中广延量定义为:若体系 A 由两个相互独立的它们 $A^{(1)}$、$A^{(2)}$ 组成,子体系分别可由性质 $\{X_j^{(1)}\}$ 与 $\{X_j^{(2)}\}$ 作完备描述,而对于任意 j 必有描述体系性质 X_j 且 $X_j = X_j^{(1)} + X_j^{(2)}$;则称性质 X_j 为该体系的广延量。除了 U(除非子体系之间有强烈的互作用)、V 和 N 之外,广延量还有体系的磁矩 $\boldsymbol{M}$、电极化强度 $\boldsymbol{P}$ 等。

公理 2:所有处于平衡态的宏观体系都存在一个称为熵(S)的量,它是体系广延量的函数。在体系的各种可能的、有内部约束的各种宏观状态的诸 S 值中,只有当所处状态没有内部约束时,体系的 S 值达到最大。

公理 3:S 属于广延量,是其自变量的连续可微函数,且 S 是 U 的单调递增函数。

公理 4:热力学第三定律(不一定总是成立的;限于篇幅本文不予讨论):

$$\text{当}\left(\frac{\partial S}{\partial U}\right)_{\{X_j\}}\to 0\text{ 时},S\to 0。\tag{3.1}$$

下面对此公理集进行讨论。

(1)热力学性质的本性涉及数学中的 Euler 齐次函数[7]。若函数 $f(x_1,x_2,\cdots,x_N)$ 满足

$$f(\lambda x_1,\ \lambda x_2,\cdots,\ \lambda x_N) = \lambda^m f(x_1,x_2,\cdots,x_N),\tag{3.2}$$

其中,λ 为实数,m 为整数,则 f 称为 m 次 Euler 齐次函数。将式(3.2)对 λ 求导再令 $\lambda=1$,可证得

$$\sum_{i=1}^{N}\frac{\partial f}{\partial x_i}x_i = mf(x_1,x_2,\cdots,x_N)。\tag{3.3}$$

广延量在数学上就是 1 次 Euler 齐次函数,$m=1$,所以对于广延量 f 有

$$\sum_{i=1}^{N}\frac{\partial f}{\partial x_i}x_i = f(x_1,x_2,\cdots,x_N)。\tag{3.4}$$

(2)从公理 1 与 2 可得 S 与各广延量之间的函数关系为

$$S = S(X_1,\cdots,X_c,U,X_{c+2}),\tag{3.5}$$

此关系被称为体系的**熵基本方程**。既然广延量 $\{X_j\}$ 能够完备描述体系,所以基本方程必含有体系的全部宏观信息。变量 $S,X_1,\cdots,X_c,U,V$ 张成的空间称为**热力学构型空间**。

(3)可从公理 3 直接给出下列推论:

(Ⅰ)既然 S 是广延量,且它的自变量均是广延量,故熵必为其变量的 1 次 Euler 齐次函数,即

$$S(\lambda X_1,\cdots,\ \lambda X_c,\ \lambda U,\ \lambda X_{c+2}) = \lambda S(X_1,\cdots,X_c,U,X_{c+2}),\tag{3.6}$$

其中实数 λ 为对体系物质量设想的缩放倍数。

(Ⅱ)S 是 U 的单调递增函数,又必须满足平衡态的温度为正值,即

$$\left(\frac{\partial S}{\partial U}\right)_{\{X_j\}} > 0。\tag{3.7}$$

(Ⅲ)U 可以从 S 的反演得到,于是从熵基本方程可以等价地写出

$$U = U(X_1,\cdots,X_c,S,X_{c+2})。\tag{3.8}$$

这个方程称为**能量基本方程**,是最常用的基本方程形式。下面将看到,由能量基本方程就可给出常用的所有热力学势。既然 U 是广延量,故它也是其自变量的 1 次 Euler 齐次函数,即

$$U(\lambda X_1,\cdots,\ \lambda X_c,\ \lambda S,\ \lambda X_{c+2}) = \lambda U(X_1,\cdots,X_c,S,X_{c+2})。\tag{3.9}$$

(4)在熵基本方程 $S = S(X_1,\cdots,X_c,U,X_{c+2})$ 中,把 U 看作自变量,S 看作因变量,此时讨论问题称为**熵表象**。在能量基本方程 $U = U(X_1,\cdots,X_c,S,X_{c+2})$ 中,把 S 看作自变量,U 看作因变量,用能量基本方程

讨论问题则称为**能量表象**。

4 强度量

4.1 能量表象中的强度量

对于多组份体系,能量表象中作为自变量的广延量为 $N_1,N_2,\cdots,N_c,S,V$,其中$(N_1,N_2,\cdots,N_c)\equiv\{N\}$ 为各组份的摩尔数(其他类型的体系,如存在电场或磁场时只要加上适当的广延量就可以得到对应的结果)。将能量基本方程[式(3.8)]求全微分,得到

$$\mathrm{d}U=\left(\frac{\partial U}{\partial S}\right)_{V,\{N\}}\mathrm{d}S+\left(\frac{\partial U}{\partial V}\right)_{S,\{N\}}\mathrm{d}V+\sum_{k=1}^{c}\left(\frac{\partial U}{\partial N_k}\right)_{S,V,\{N_{j(\neq k)}\}}\mathrm{d}N_k。\tag{4.1}$$

定义能量表象中的下列各种强度量:

温度 $T\equiv\left(\frac{\partial U}{\partial S}\right)_{V,\{N\}}$; (4.2)

压强 $p\equiv-\left(\frac{\partial U}{\partial V}\right)_{S,\{N\}}$; (4.3)

化学势 $\mu_k\equiv\left(\frac{\partial U}{\partial N_k}\right)_{S,V,\{N_{j(\neq k)}\}}$。 (4.4)

于是在公理理论中引入全部强度量,即引入温度、压强和化学势概念,就可以进一步将式(4.2)~(4.4)纳入下面的普遍定义中(为强调它们是能量表象中的强度量,将上述广延量 X_j 加上标,改记为 $X_j^{(U)}$),即将能量表象中的**第 m 个强度量**定义为:

$$p_m^{(U)}\equiv\left(\frac{\partial U}{\partial X_m^{(U)}}\right)_{\{X_{j(\neq m)}\}}\quad\forall m=1,\cdots,(c+1)。\tag{4.5}$$

再将 S、V 作为第 $c+1$、$c+2$ 个广延量 $X_{c+1}^{(U)}$ 和 $X_{c+2}^{(U)}$;将对应的强度量温度、$-p$ 改称为第 $c+1$、$c+2$ 个强度量 $p_{c+1}^{(U)}$ 和 $p_{c+2}^{(U)}$。于是,能量表象的广延量有

$$(X_1^{(U)},\cdots,X_c^{(U)},X_{c+1}^{(U)},X_{c+2}^{(U)})=(N_1,\cdots,N_c,S,V)\equiv\{X^{(U)}\};\tag{4.6}$$

它们对应的(即共轭的)强度量为

$$(p_1^{(U)},\cdots,p_c^{(U)},p_{c+1}^{(U)},p_{c+2}^{(U)})=(\mu_1,\cdots,\mu_c,T,-p)\equiv\{p^{(U)}\}。\tag{4.7}$$

于是,U 的全微分[式(4.1)]又可写为

$$\mathrm{d}U=\sum_{j=1}^{c}\mu_j\mathrm{d}N_j+T\mathrm{d}S-p\mathrm{d}V=\sum_{m=1}^{c+2}p_m^{(U)}\mathrm{d}X_m^{(U)},\tag{4.8}$$

表现为 $\mathrm{d}U=\sum_{m=1}^{c+2}(\text{强度量})_m^{(U)}d(\text{共轭广延量})_m^{(U)}$ 的形式。

此时强度量是广延量 $\{N_j\}$,S,V 的函数,即 $T=T(N_1,\cdots,N_c,S,V)$,$p=p(N_1,\cdots,N_c,S,V)$ 和 $\mu_j=\mu_j(N_1,\cdots,N_c,S,V)(j=1,\cdots,c)$。这些函数关系均称**为状态方程**。所以称它们为强度量是因为这些量与体系的物质量无关;显然,强度量都是 0 次 Euler 齐次函数。可以证明式(4.2)定义的温度是 0 次 Euler 齐次函数(见文后附录 A)。类似地,也可以证明其他强度量均为 0 次 Euler 齐次函数。

4.2 熵表象中的强度量

有时不知道函数 $z(x,y)$ 的具体形式,只知道 x,y,z 三者之间存在依赖关系 $f(x,y,z)=0$,这就是隐函数。可以求证这三个互相依赖的变量 x,y,z 之间存在**乘积 -1 规则**[7附录D],即

$$-1=\left(\frac{\partial x}{\partial z}\right)_y\left(\frac{\partial y}{\partial x}\right)_z\left(\frac{\partial z}{\partial y}\right)_x。\tag{4.9}$$

对于多组份体系,根据熵表象中的基本方程[式(3.5)],求得 S 的全微分为

$$\mathrm{d}S=\sum_{k=1}^{c}\left(\frac{\partial S}{\partial N_k}\right)_{U,V,\{N_{j(\neq k)}\}}\mathrm{d}N_k+\left(\frac{\partial S}{\partial U}\right)_{V,\{N\}}\mathrm{d}U+\left(\frac{\partial S}{\partial V}\right)_{U,\{N\}}\mathrm{d}V\equiv\sum_{k=1}^{c+2}p_k^{(S)}\mathrm{d}X_k^{(S)},\tag{4.10}$$

其中,为强调是在熵表象讨论问题,将广延量 $N_1,\cdots,N_c,U,V$ 改记为 $X_1^{(S)},\cdots,X_c^{(S)},X_{c+1}^{(S)},X_{c+2}^{(S)}$,即 $\{X_j^{(S)}\}$。

对应地,熵表象中的强度量为:

$$p_k^{(S)} \equiv \left(\frac{\partial S}{\partial N_k}\right)_{U,V,\{N_{j(\neq k)}\}} = -\left(\frac{\partial U}{\partial N_k}\right)_{V,S,\{N_{j(\neq k)}\}} \Big/ \left(\frac{\partial U}{\partial S}\right)_{V,N_k,\{N_{j(\neq k)}\}} = -\mu_k T^{-1}; \tag{4.11}$$

上面第二步演绎运用了式(4.9)。同理得到

$$p_{c+1}^{(S)} \equiv \left(\frac{\partial S}{\partial U}\right)_{V,\{N\}} = \left\{\left(\frac{\partial U}{\partial S}\right)_{V,\{N\}}\right\}^{-1} = T^{-1}; \tag{4.12}$$

$$p_{c+2}^{(S)} \equiv \left(\frac{\partial S}{\partial V}\right)_{U,\{N\}} = -\left(\frac{\partial U}{\partial V}\right)_{S,\{N\}} \Big/ \left(\frac{\partial U}{\partial S}\right)_{V,\{N\}} = pT^{-1}。 \tag{4.13}$$

将式(4.11)~(4.13)纳入如下普遍的定义,即将 U,V 两个广延量分别记为 $X_{c+1}^{(S)}$、$X_{c+2}^{(S)}$,于是对应的两个强度量 T^{-1}、pT^{-1} 就可分别记为 $p_{c+1}^{(S)}$、$p_{c+2}^{(S)}$。总之,熵表象的**第 m 个强度量**可定义为:

$$p_m^{(S)} \equiv \left(\frac{\partial S}{\partial X_m^{(S)}}\right)_{\{X_{j(\neq m)}\}} \quad \forall m = 1,\cdots,(c+2)。 \tag{4.14}$$

这样,现在熵表象中的广延量与强度量分别为:

$$\{X^{(S)}\} \equiv (X_1^{(S)},\cdots,X_c^{(S)},X_{c+1}^{(S)},X_{c+2}^{(S)}) = (N_1,\cdots,N_c,U,V), \tag{4.15}$$

和

$$\{p^{(S)}\} \equiv (p_1^{(S)},\cdots,p_c^{(S)},p_{c+1}^{(S)},p_{c+2}^{(S)}) = (-\mu_1 T^{-1},\cdots,-\mu_c T^{-1},T^{-1},pT^{-1})。 \tag{4.16}$$

5 平衡条件

导出强度量之后,现在可以从上述热力学公理体系出发求得导热、膨胀、扩散现象等各种现象的平衡条件。先以热传导为例试求其平衡条件。据式(4.8),不受外力体系的熵的全微分为

$$\mathrm{d}S = -\sum_{j=1}^{c}\frac{\mu_j}{T}\mathrm{d}N_j + \frac{1}{T}\mathrm{d}U + \frac{p}{T}\mathrm{d}V。 \tag{5.1}$$

式(5.1)右边各系数项 $\{\mu_j/T\}$,$1/T$,p/T 各自均为变量 $\{N_j\}$,U,V 的函数,即

$$\frac{1}{T} = \frac{1}{T}(N_1,\cdots,N_c,U,V), \tag{5.2}$$

$$\frac{p}{T} = \frac{p}{T}(N_1,\cdots,N_c,U,V), \tag{5.3}$$

$$\frac{\mu_j}{T} = \frac{\mu_j}{T}(N_1,\cdots,N_c,U,V) \quad \forall j = 1,\cdots,c。 \tag{5.4}$$

这组方程被称为**熵表象中的状态方程**。

从公理 2 得知,体系平衡时 $\mathrm{d}S=0$。据此,可以得到复合体系的平衡条件(可以划为几个子体系的体系称为**复合体系**)。例如,设一由两个子体系组成的且其子体系之间由刚性、粒子不可渗透的导热墙隔开(即两个子体系之间只能通过热传导交换能量)的孤立体系。于是 $\mathrm{d}U^{(2)} = -\mathrm{d}U^{(1)}$,且 $\mathrm{d}V^{(2)} = \mathrm{d}V^{(1)} = \mathrm{d}N_k^{(2)} = \mathrm{d}N_k^{(1)} = 0$。因此,从式(5.1)得到 $0 = \mathrm{d}S = \frac{1}{T}\mathrm{d}U = \frac{1}{T^{(1)}}\mathrm{d}U^{(1)} + \frac{1}{T^{(2)}}\mathrm{d}U^{(2)}$。即该平衡的复合体系满足 $\left(\frac{1}{T^{(1)}} - \frac{1}{T^{(2)}}\right)\mathrm{d}U^{(1)} = 0$。于是,当且仅当 $T^{(1)} = T^{(2)}$ 时,这种孤立的复合体系才能达到平衡。这就从热力学公理体系(而不是从常识)导出温度相等是两个子体系间导热现象的平衡条件。

还可以从热力学公理体系(而不是从常识)出发,通过下面的演绎来证明“热流从高温区流向低温区”。还是讨论上述含两个子体系的复合体系,设初始时隔墙是绝热、刚性的不透墙,且 $T^{(1)} < T^{(2)}$。接着撤去隔墙的绝热约束,即将隔墙换成刚性、不透的导热墙。于是,根据公理 2,此时任意使体系趋向未来平衡态的自发过程的熵应该满足 $\Delta S > 0$。即 $0 < \mathrm{d}S = \frac{1}{T}\mathrm{d}U = \frac{1}{T^{(1)}}\mathrm{d}U^{(1)} + \frac{1}{T^{(2)}}\mathrm{d}U^{(2)} = \left(\frac{1}{T^{(1)}} - \frac{1}{T^{(2)}}\right)\mathrm{d}U^{(1)}$。由于此时 $T^{(1)} < T^{(2)}$,故 $\mathrm{d}U^{(1)} > 0$,即低温区的能量将上升,热流从高温区流入低温区。

类似地,我们还可从热力学公理体系(而不是从常识)出发演绎证明(1)若两个子体系之间由弹性、不透热的绝热膜隔开,且两边初始压强不等,则压强高的子体系将向压强低的子体系膨胀;(2)若两个子体系之间由刚性、可透粒子的绝热膜隔开,且两边初始化学势不等,则化学势高的子体系中粒子将流向化学势低的子体系。

总之,我们从公理集出发就可以演绎得到全部直观的经验现象或常识(如上述的导热、膨胀、扩散现象等)。

6 Euler 方程、Gibbs-Duhem 方程、相律

内能 U 的广延量属性将带来一些重要性质。广延量属于 1 次 Euler 齐次函数,根据式(3.4)得到多组份体系的内能为 $U = U(N_1,\cdots,N_c,S,V) = \sum_{j=1}^{c}\left(\frac{\partial U}{\partial N_j}\right)_{S,V,\{N_{j(\neq k)}\}} N_j + \left(\frac{\partial U}{\partial S}\right)_{V,\{N_j\}} S + \left(\frac{\partial U}{\partial V}\right)_{S,\{N_j\}} V$。再根据式(4.2)、(4.3)和(4.4)得到对于无外场条件下的多组份体系的**能量表象中的 Euler 方程**[式(6.1)]。

$$U = \sum_{j=1}^{c}\mu_j N_j + TS - pV, \tag{6.1}$$

同理,根据熵基本方程[式(3.5)]和广延量熵是 1 次 Euler 齐次函数,运用式(4.11)~(4.13),在熵表象中证得**熵表象中的 Euler 方程**[式(6.2)]。

$$S = -\sum_{j=1}^{c}\frac{\mu_j}{T}N_j + \frac{1}{T}U + \frac{p}{T}V, \tag{6.2}$$

当然式(6.1)、(6.2)都是指不受外力的体系,对于更普遍的存在外力的体系也可以同理证得类似的 Euler 方程。

将能量表象中的 Euler 方程[式(6.1)]求全微分,再减去能量的全微分[式(4.8)],就可得到 Gibbs-Duhem 方程①。

$$\sum_{j=1}^{c} N_j \mathrm{d}\mu_j + S\mathrm{d}T - V\mathrm{d}p = 0, \tag{6.3}$$

由此可见,对于含 c 个组份的单相体系,必须受此 Gibbs-Duhem 方程的约束,于是体系的独立变量个数为 $f = (c+2) - 1$,此即热力学自由度的数目。若体系平衡时有 p 个相,则应该受 p 个 Gibbs-Duhem 方程的约束。因此,体系自由度 $f = (c+2) - p$。这就是著名的 Gibbs 相律。

7 最大熵原理与能量最低原理

公理 2 等价于如下表述的**最大熵原理:在给定内能的条件下,体系无约束的内部状态变量在平衡态时将使体系熵值达到极大**。我们下面来证明,最大熵原理等价于如下表述的**能量最低原理:在给定体系熵值的条件下,体系无约束的内部状态变量在平衡态时将使体系能量达到极小**[11,12]。

Tisza[11]假设体系始态按能量判据看是不稳定的,则在保持该体系体积 V 和其余广延量 X_j 不变的约束之下,对体系施以扰动,先后经历两个过程。第一个过程是放出能量的等熵过程,于是 $\Delta U^{(1)} < 0$ 而同时 $\Delta S^{(1)} = 0$(上标$^{(1)}$ 表示第一个过程)。接着,将体系与热浴接触作第二次扰动,让体系经历第二个过程:吸入能量的熵增过程。使之将第一次扰动时失去的能量正好等量吸入,即 $\Delta U^{(2)} = |\Delta U^{(1)}| > 0$,可是因为第二次过程中必有 $\left(\frac{\partial S}{\partial U}\right)_{V,\{X\}} = T^{-1} > 0$,故必有 $\Delta S^{(2)} > 0$。经过这两次过程后,其总的结果为 $\Delta S = \Delta S^{(1)} + \Delta S^{(2)} > 0$ 和 $\Delta U = \Delta U^{(1)} + \Delta U^{(2)} = 0$,按照熵判据可见,该体系的始态也是不稳定的。

同理,若体系处于的始态按照熵判据来说是不稳定的,则可以类似地证明按照能量判据其始态也是不稳定的。从两个方向证得同样的含义,可见它们是等价的。

---

① 法国物理学家 Pierre M. Duhem(1861-1916)强烈反对单一的归纳法。中文书中常把 Duhem 译为"杜亥姆",其实不对,按读音[dyǎ]应如少数书中采用的译名"迪昂"。

Callen[12]从另一个角度同样证明了平衡态时最大熵原理与最低能量原理的等价性,见文后附录 B。等价关系也可以从图 1 看出。基本方程在热力学构型空间中是一个超曲面。由此可以得到图 1 右上部分的最大熵原理,以及图 1 右下部分的最低能量原理[12]。

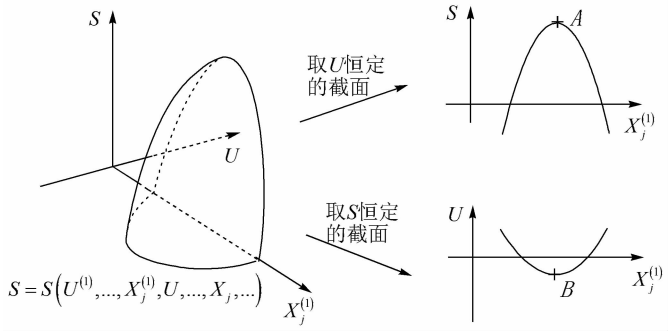

图 1 最大熵原理与最低能量原理之间等价关系证明的图示[12]

Fig. 1 The proof of the equivalence between entropy maximum principle and energy minimum principle [12]

8 Legendre 变换和热力学势

Legendre 变换是一种改换函数中部分自变量的数学方法。对于不同的实验条件,我们可以从公理 2 的能量表象或熵表象开始,用 Legendre 变换得到适当的热力学势[8,13],如焓 H、Helmholtz 自由能 F、Gibbs 自由能 G 和巨势 Φ(grant potential)。最常见的是从能量表象开始。

设 $Y(X_0,\cdots,X_k)$ 为广延量$\{X_j\}$ 的函数,令 $p_j \equiv \left(\frac{\partial Y}{\partial X_j}\right)_{\{X_{i(\neq j)}\}}$ 为 **Y-表象**中对应的强度量;定义一个有关函数 Y 关于$\{X_{j(\leqslant i)}\}$ 的 **Legendre 变换**的新函数 Λ[式(8.1)]

$$\Lambda \equiv Y[p_0,\cdots,p_i]_{i\leqslant k} \equiv Y - \sum_{j=0}^{i} p_j X_j \tag{8.1}$$

○将式(8.1)求微分,得到

$$\mathrm{d}\Lambda = -\sum_{j=0}^{i} X_j \mathrm{d}p_j + \sum_{j=i+1}^{k} p_j \mathrm{d}X_j。 \tag{8.2}$$

于是,Y 的自变量被变换成为新函数 Λ 的自变量 $p_0,\cdots,p_i,X_{i+1},\cdots,X_k$。这就如经典力学中,可用 Legendre 变换方法来选用实验可观测的或受控的变量作为独立变量。

对于无外场作用的体系,从内能 U 开始得到如下最常见的4个热力学势 F,H,G,Φ:(1) 恒温恒容的体系(正则系综) 中的 Helmholtz 自由能:$F(\{N_k\},T,V) \equiv U - TS = \sum_k \mu_k N_k - pV$,或 $\mathrm{d}F = \sum_k \mu_k \mathrm{d}N_k - S\mathrm{d}T - p\mathrm{d}V$;(2) 恒压恒熵体系中的焓:$H(\{N_k\},S,p) \equiv U + pV = \sum_k \mu_k N_k + TS$,或 $\mathrm{d}H = \sum_k \mu_k \mathrm{d}N_k + T\mathrm{d}S + V\mathrm{d}p$;(3) 恒温恒压体系中的 Gibbs 自由能:$G(\{N_k\},T,p) \equiv U + pV - TS = \sum_k \mu_k N_k$,或 $\mathrm{d}G = \sum_k \mu_k \mathrm{d}N_k - S\mathrm{d}T + V\mathrm{d}p$;(4) 恒温恒容恒化学势体系(巨正则系综) 中的巨势:$\Phi(\{\mu_k\},T,V) \equiv U - TS - \sum_k \mu_k N_k = -pV$,和 $\mathrm{d}\Phi = \sum_k N_k \mathrm{d}\mu_k - S\mathrm{d}T - p\mathrm{d}V$。这几个热力学势,在其独立变量固定的情况下,平衡态时相应的热力学势将处于极小值。它们各自也都与最大熵原理或最低能量原理呈等价关系。

9 热力学响应函数

在指定情况下,一个热力学变量相对于其余变量中任意一个变量的导数称为**热力学响应函数**。例如,对于无外场作用的体系,有等压热膨胀率 $\alpha \equiv \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_{p,\{N_k\}}$ 和等温压缩率 $\kappa_T \equiv -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_{T,\{N_k\}}$ 等响应函数。

在恒定 $\{N_k\}$ 的条件下,内能 U 的全微分为 $dU = TdS - pdV$。因为 $-pdV$(负号来自对压强约定的方向)是与扰动有关的无穷小的对外作功,所以 TdS 必定是以热方式的能量传入,即无穷小吸热。与温度改变对应的响应函数称为热容。依据取哪个变量固定的不同,热容有恒容热容 C_v 和恒压热容 C_p,它们分别为:

$$C_V \equiv \left(\frac{\partial U}{\partial T}\right)_{V,\{N_k\}} = T\left(\frac{\partial S}{\partial T}\right)_{V,\{N_k\}} \text{ 和 } C_p \equiv \left(\frac{\partial H}{\partial T}\right)_{p,\{N_k\}} = T\left(\frac{\partial S}{\partial T}\right)_{p,\{N_k\}} \text{。}$$

热力学势的全微分,用它的二阶混合导数可以得出许多有用的表式,称为 Maxwell 关系式。例如从 U 的全微分可得到式(9.1)。

$$\left(\frac{\partial T}{\partial V}\right)_{S,\{N_k\}} = -\left(\frac{\partial p}{\partial S}\right)_{V,\{N_k\}} \tag{9.1}$$

本文上述介绍的公理化热力学方案可总结为图 2。

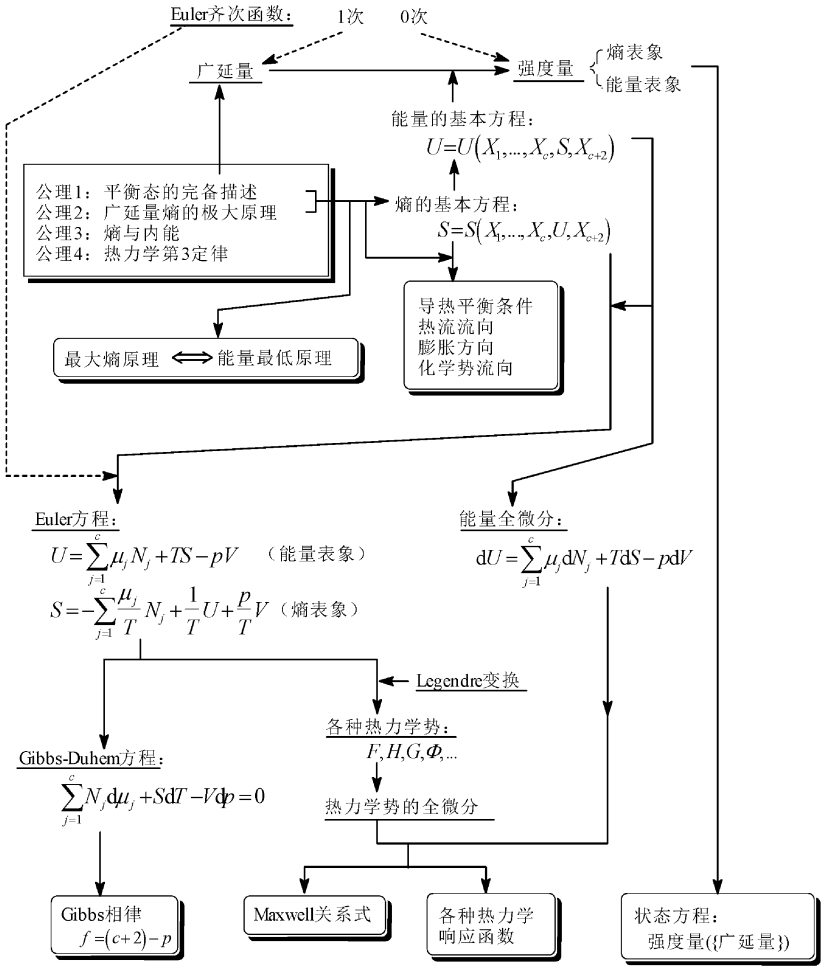

图 2 Callen 的热力学公理化方案的思路

Fig. 2 Schematic of the axiomatic formalism of thermodynamics suggested by Callen

综上所述,热力学公理化本质上是,先根据基于经验常识的传统热力学来假设公理集,然后就可以重整思路,只从公理集出发,导出全部经验常识和全部热力学关系。表面上看,似乎是从常识出发"拐个弯,又回来了",其实不然。经过公理化思考,热力学理论的内部结构更为清晰了。热力学成为了高度逻辑性的理论,它以类似于 Euclid 几何的公理化形式出现在整个科学理论的大框架中。在公理化热力学中,Carnot 循环不见了。Carnot 循环犹如婴儿学步车一样,在引进熵和帮助理解熵概念的历史作用完成后,就可以被抛弃不用了。M Planck 第一个认识到这一点。

虽然人们常说 Euclid 几何来自丈量土地、航海等世界各处都有的生产活动,但是为何唯独 Euclid

从丈量土地出发“拐个弯又回来”创造的平面几何能够成为后来现代自然科学的奠基石呢,其深刻意义值得深思。这个“拐弯”导致欧洲超出了原本相当先进的中国的科学技术至少 200 年。活跃在 1850 年间的数学家李善兰是把解析几何和微积分介绍到中国的第一人。而此时,欧洲已经掌握了微积分、变分法、复变函数,正在研究微分方程、抽象代数、非欧几何和经典微分几何了。自然科学家们看到,Euclid 几何公理化这种对经验常识的逆向、拐弯和依据演绎重组的思维方法对整个现代自然科学基础的建立起着决定性的作用,这样的体会要远比说一句“螺旋式上升”深刻得多、丰富得多。

16 世纪末,英国哲学家 Francis Bacon 提出,科学真理来自实验的归纳,从而奠定了实验科学的思想基础。在相当长时间内,人们把归纳法当作发现真理的唯一方法。但是,20 世纪初的英国数学家、哲学家 Betrand Russell 给沉湎于归纳法的实验科学家们一个警示。Russell[14] 说:“每天喂小鸡而喂了它一辈子的那个人,临了却绞断了这只小鸡的脖子,……某一件事已发生过若干次,就只这一点就使动物和人预料(这件事)它还会发生。……我们(人)所处的地位并不比脖子出乎预料被绞断的小鸡更好些”。Russell 以“Russell 鸡”的故事说明,“实践是检验真理的唯一标准”的说法不等同于能无端地把过去的实践也当作将来的实践。Russell 鸡临终才明白它的求真判据有问题。

把思维的基础抽象为公理结构。有了公理化的理论之后,就可以在包含全部归纳法取得的知识之上建立演绎法。这也是所有关于自然的科学理论曾经走过或将要走过的共同道路,而且是唯一的道路。不这样做,就无法超越“Russell 鸡”的求真判据。那是人类与其他动物之间在探求真理上的实质差别所在。

公理化实际上是跨越经验归纳、通过演绎法发现隐藏其内的内禀规律。上述热力学的公理化使我们透过热力学的唯象理论外表,发现它内在蕴含着的公理化的严密逻辑体系。正因为如此,杜绝了出现例外的可能,其他唯象理论就远没有如此严密(exact)的基础。

唯象理论特别可贵,它们是真正科学的起点,指出了通向真理的线索,是通向真理的中间站,尤其可贵的是那最早的前一、两种唯象理论。但是,必须清楚的是,在唯象理论的脚下还可能是一片沙滩或沼泽,不是久留之地;如果不依靠演绎法的帮助,即使补充新的实验数据、改进一千次归纳也必将无济于事。我们追求的是物理模型,不得已时才采用数学模型。同一个客观过程,可以轻易地设想出一千种难分高低的数学模型。甚至得到一条几乎通过所有实验数据点的拟合曲线也是可能的,不过那是无意义的过拟合。

10 结束语:“凭什么相信计算”

1998 年以前,“化学是实验科学”和“化学有化学的规律”的说法是无可质疑的、压倒一切的,谁质疑谁就被认为不具备化学家起码的学术素养。可是,1998 年 10 月颁布的诺贝尔化学奖之后,形势有了划时代的变化。当年的诺贝尔化学奖公告的补充材料中明确宣称“化学不再是单纯的实验科学了”[15]。至今,形形色色的计算已经遍布化学中的所有领域。面对新形势,人们自然不禁要发问:“凭什么相信计算?”其实,那是在追问当今人们心目中关于自然规律认识的可靠性,所以不只是对理论化学家、计算化学家发问,而是对所有化学家的发问。100 年前物理学家通过演绎法摆脱了“Russell 鸡”的单纯归纳思维的逻辑,现在应该是化学家清理思想、建立完整思维逻辑意图的时候了。要看到以实验现象的归纳为唯一依据得到的唯象理论很有可能最后被发现是建立在沙滩上的。

在当今化学界模拟计算相当盛行的背景下,笔者希望通过热力学公理化的介绍引起读者对“凭什么相信计算”问题作深入的思考。思考在计算背后的理论其基础是否坚实、可靠;思考计算是否应该独立于实验(作为计算起点的若干物理常数必须由实验提供除外)。尽管理论起源于实验,但是,高度依赖于实验的“理论计算值”是否有足够的资格值得被实验去检验呢?思考是否真有脱离第一原理的“人工智能”?以期化学学科得到更健康、更踏实地发展。

参 考 文 献

[1] J Ziman. 可靠的知识:对科学信仰中的原因的探索,赵振江译,北京:商务印书馆,2003.

[2] 王泽农. 现代物理知识, 2008, 20(3): 52~54;http://en.wikipedia.org/wiki/Phenomenology_(science).

[3] L Pauling. The Nature of Chemical Bond, 3rd ed., Cornell University Press, Ithaca, 1960; 陈念贻. 键参数函数及其应用,北京:科学出版社,1976;温元凯,邵俊. 离子极化导论,合肥:安徽教育出版社,1985;L P Hammett. J. Am. Chem. Soc., 1937, 59: 96~103.
[4] 高隆昌. 数学及其认识,北京:高等教育出版社,2001,第 2 章第 3 节.
[5] 甘子钊为北京大学物理学丛书写的序,见彭桓武、徐锡申著 理论物理基础,北京:北京大学出版社,1998, 8.
[6] J C von Neumann. Mathematical Foundations of Quantum Theory, New Jersey: Princeton Univ. Press, 1955.
[7] 沈惠川. 统计力学,合肥:中国科学技术大学出版社,2011.
[8] 陈敏伯. 统计力学:理论化学用书,北京:科学出版社,2012 年.
[9] C Carathéodory. Math. Ann. ,1909, 67: 355.
[10] 介绍 Carathéodory 理论的书籍和文章还有:王竹溪. 热力学,第 2 版,北京:北京大学出版社,2005;H Margenau, G M Murphy. The Mathematics of Physics and Chemistry, 2nd ed., D. van Nostrand Co., New Jersey, 1956,26~31;C S Helrich. Modern Thermodynamics with Statistical Mechanics, Springer, 2009, 372~378; H A Buchdahl. Am. J. Phys., 1949, 17: 41, 44, 212.
[11] L Tisza. Generalized Thermodynamics, MIT Press, 1966.
[12] H B Callen. Thermodynamics and an Introduction to Thermostatistics, 2nd ed., John Wiley, New York, 1985.
[13] 陈敏伯. 计算化学:从理论化学到分子模拟,北京:科学出版社,2009.
[14] B Russell. 哲学问题,北京商务印书馆,1959,第 6 章"论归纳法".
[15] The Royal Swedish Academy of Sciences. Additional Background Material on the Nobel Prize in Chemistry 1998, Oct. 12, 1998;陈敏伯. 科学 1999, 50 (1) :58;科学时报 1999 年 6 月 23、24 日第二版转载.

[附录 A]

求证:式(4.2)定义的温度是 0 次 Euler 齐次函数。

证明:依次根据式(4.2)和(3.6),可得

$$T(N_1,\cdots,N_c,S,V) \equiv \left(\frac{\partial U(N_1,\cdots,N_c,S,V)}{\partial S}\right)_{V,\{N_j\}} = \left(\frac{\partial \lambda U(N_1,\cdots,N_c,S,V)}{\partial(\lambda S)}\right)_{V,\{N_j\}}$$

$$= \left(\frac{\partial U(\lambda N_1,\cdots,\lambda N_c,\lambda S,\lambda V)}{\partial(\lambda S)}\right)_{V,\{N_j\}} = T(\lambda N_1,\cdots,\lambda N_c,\lambda S,\lambda V)。$$

其中第 3 步利用内能 U 是其变量的 1 次 Euler 齐次函数;第 4 步利用温度的定义式(4.2)。于是,可见温度是 0 次 Euler 齐次函数。

[附录 B]

求证:平衡体系的最大熵原理与最低能量原理是等价的(Callen[11]方法)。

证明:对于含两个子体系的复合体系。其熵表象的自变量为 $(U^{(1)},\{X_j^{(1)}\},U,\{X_j\})$。当复合体系达到最大熵时,熵 S 关于子体系 1 的任一广延量 $X_j^{(1)}$ 的偏导数应当满足一阶导数为零和二阶导数小于零,即

$$\left(\frac{\partial S}{\partial X_j^{(1)}}\right)_{U,X} = 0,\text{和}\left(\frac{\partial^2 S}{\partial X_j^{(1)2}}\right)_{U,X} < 0 \quad \forall j, \tag{B.1}$$

其中约定对 $X_j^{(1)}$ 作偏导数时固定变量下标 X 指不包含 $X_j^{(1)}$ 的全部$\{X_i^{(1)}\}$ 和$\{X_j\}$。

此时,体系的能量 U 又是如何变化的呢?能量表象的自变量为$(S^{(1)},\{X_j^{(1)}\},S,\{X_j\})$。令体系能量对子体系任一广延量 $X_j^{(1)}$ 的一阶偏导数为

$$P \equiv \left(\frac{\partial U}{\partial X_j^{(1)}}\right)_{S,X}, \tag{B.2}$$

其中约定固定变量下标 S 指 $S^{(1)}$,S。利用乘积 -1 规则,且考虑到$\left(\frac{\partial S}{\partial U}\right)_{X_j^{(1)},X} = \frac{1}{T}$,得到

$$P \equiv \left(\frac{\partial U}{\partial X_j^{(1)}}\right)_{S,X} = -\left(\frac{\partial S}{\partial X_j^{(1)}}\right)_{U,X} \Big/ \left(\frac{\partial S}{\partial U}\right)_{X_j^{(1)},X} = -T\left(\frac{\partial S}{\partial X_j^{(1)}}\right)_{U,X} = 0, \tag{B.3}$$

其中最后一步演绎的理由是此时熵已达到极大。再看体系内能 U 对该 $X_j^{(1)}$ 的二阶偏导数:

$$\left(\frac{\partial^2 U}{\partial X_j^{(1)2}}\right)_{S,X} = \left(\frac{\partial P}{\partial X_j^{(1)}}\right)_{S,X} = \left(\frac{\partial P}{\partial U}\right)_{S,X}\left(\frac{\partial U}{\partial X_j^{(1)}}\right)_{S,X} + \left(\frac{\partial P}{\partial X_j^{(1)}}\right)_{U,X} = \left(\frac{\partial P}{\partial U}\right)_{S,X} P + \left(\frac{\partial P}{\partial X_j^{(1)}}\right)_{U,X} = \left(\frac{\partial P}{\partial X_j^{(1)}}\right)_{U,X}, \tag{B.4}$$

其中第 2 步演绎把 P 看成 U、$X_j^{(1)}$、$X(\neq X_j^{(1)})$、S 的函数,求 S、$X(\neq X_j^{(1)})$ 固定时 P 对 $X_j^{(1)}$ 的导数,于是有 $X_j^{(1)}$ 通过 U 间接造成 P 变化的贡献(第 1 项),和 $X_j^{(1)}$ 直接造成 P 变化的贡献(第 2 项)。第 3 步和最后一步演绎的依据分别是式(B.2)和(B.3)。再将式(B.3) 中的 P 代入式(B.4),得到

$$\left(\frac{\partial^2 U}{\partial X_j^{(1)2}}\right)_{S,X} = \left(\frac{\partial}{\partial X_j^{(1)}}\left[-\frac{\left(\frac{\partial S}{\partial X_j^{(1)}}\right)_{U,X}}{\left(\frac{\partial S}{\partial U}\right)_{X_j^{(1)},X}}\right]\right)_{U,X} = -\frac{\left(\frac{\partial^2 S}{\partial X_j^{(1)2}}\right)_{U,X}}{\left(\frac{\partial S}{\partial U}\right)_{X_j^{(1)},X}} + \left(\frac{\partial S}{\partial X_j^{(1)}}\right)_{U,X}\frac{1}{\left(\frac{\partial S}{\partial U}\right)_{X_j^{(1)},X}^{2}}\frac{\partial S}{\partial X_j^{(1)}\partial U} \tag{B.5}$$

熵极大时,右边第 2 项为零,再考虑到式(4.2),故式(B.5)变为

$$\left(\frac{\partial^2 U}{\partial X_j^{(1)2}}\right)_{S,X} = -T\left(\frac{\partial^2 S}{\partial X_j^{(1)2}}\right)_{U,X} > 0 \tag{B.6}$$

因为温度总是正值(负温度不是平衡态),可见当复合体系达到最大熵的时候,体系能量也恰好达到极小值;这表明我们从最大熵原理证得最低能量原理。反过来,也可类似地从最低能量原理证得最大熵原理。因此,平衡体系的最大熵原理与最低能量原理是等价的。