

关于统计力学的基本原理

郑伟谋[†]

(中国科学院理论物理研究所 北京 100190)

2018-01-24收到

[†] email: zheng@itp.ac.cn

DOI: 10.7693/wl20181001

On the fundamental principles of statistical mechanics

ZHENG Wei-Mou[†]

(Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China)

摘要 文章简要讨论有关统计力学基本原理的几个问题, 包括正则系综理论、系综分布函数的支集、与热力学的对应、不可逆性及分布函数的时间演化。

关键词 统计力学, 热力学, 不可逆性

Abstract Several problems concerning the fundamental principles of statistical mechanics, including the canonical ensemble theory, the supporting set of an ensemble distribution function, the correspondence to thermodynamics, irreversibility, and time evolution, are briefly discussed.

Keywords statistical mechanics, thermodynamics, irreversibility

宏观系统有为数不多的几个可直接观测量, 如气体的压强 p 、体积 V 和温度 T 。热力学描述这些量之间的关系, 唯象刻画系统的整体行为。统计力学的目的是研究宏观物体的行为和性质所遵循的特殊一类规律性, 它的一个重要任务是解释作为唯象理论的热力学。统计力学可由分子微观性质计算热力学量。统计力学有双重意义: 由微观力学(如分子能级、谱学测量)知识计算热力学量, 由测量宏观热力学性质反推微观性质(如分子间相互作用)。统计力学可以突破热力学的局限, 将研究延伸至热力学不再成立的领域。非平衡态体系一般没有简单的热力学宏观量描述, 但分布函数描述仍是明确的。统计力学处理服从哈密顿动力学的微观系统, 但原则上微观对象也可以是经济学量、社会学量等, 它们并不满足哈密顿动力学。

1 统计规律性

考虑体积为 V 的空间里有遵从经典哈密顿动力学的 N 个粒子, 这个体系的状态由这些粒子的坐标和动量 $(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N; \mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N) \equiv (\mathbf{r}^N, \mathbf{p}^N)$ 给定, 这种状态也叫微观构象态或构象态。构象态对应于由 \mathbf{r}^N 和 \mathbf{p}^N 所张成的 $6N$ 维相空间中的一点。设体系哈密顿量为 $H(\mathbf{r}^N, \mathbf{p}^N) = K(\mathbf{p}^N) + U(\mathbf{r}^N)$, 则运动方程为

$$\dot{\mathbf{r}}_j = \frac{\partial H}{\partial \mathbf{p}_j}, \quad \dot{\mathbf{p}}_j = -\frac{\partial H}{\partial \mathbf{r}_j}. \quad (1)$$

体系构象态随时间的演化, 在相空间中描画出一条“相轨道”或分子轨道。这样的体系虽然遵从经典力学, 不难写下运动微分方程, 但其自由度巨大, 不可能对给定的初条件积分方程求解。巨大的自由度数目, 导致体系全新的规律性。作为热力学研究对象的宏观体系总是存在于某种环境

之中。内在的(混沌系统动力学不可预测性)和外在的(环境扰动噪声)原因,使得分子轨道之间不断混合。原先的分子轨道图像不复存在,精确求解动力学也不再必要。体系出现新的规律性即统计规律性,例如,体积 V 内任一足够大的体元中的粒子数相当恒定。这导致热力学中的观测结果:大系统表现出十分简单有序的行为,可仅用少数几个变量表征。这时分子轨道的语言为分布的语言所替代。统计规律性定义了体系的一种全新的状态即统计力学状态,它指定了支在相空间上的一个分布,刻画体系可在特定相空间点附近出现的概率。应该强调,这种统计规律性不依赖于微观规律的具体细节,无论粒子的运动是用经典力学还是用量子力学描述,统计力学的理论框架并不改变。热力学中的热力学平衡宏观状态即热力学态,由少数几个独立变量完全限定,相应地有热力学态空间。热力学态空间中的路径对应于热力学过程。热力学量,一些可借助统计力学分布通过求平均得到,另一些则并非平均量而须由分布直接导出,后者有必要特别指出。区分构象态、统计力学状态和热力学态并不难,自觉地准确运用这些概念思考和分析问题至关重要。

2 正则系综

体系的统计力学状态,是指支在相空间上的一个分布。一般而言,分子轨道的动力学演化时间尺度远小于分布演化的时间尺度。统计力学体系,总是处于环境中,且具有大自由度,使得精确求解分子轨道动力学既不可能也非必要。这也是体系平衡态存在简单的宏观热力学描述的原因。大自由度的微观体系,作为动力系统,除哈密顿量外不存在其他的独立运动积分。在理想的情形下,体系只有哈密顿量也只是在统计平均的意义上是守恒的。统计力学体系的分布演化的长时间行为,可只取决于其哈密顿量。分子轨道的时间尺度和分布的时间尺度彼此分离,这导致时间在平衡统计力学中不扮演举足轻重的角色。上述的统计力学原理给出了平衡统计分布为正则分

布。关于正则分布,值得稍深入地讨论。

1902年 Gibbs 引入了系综的概念,它是满足某种统计分布的物理实体的集合,物体的性质由取系综的统计平均来计算。关于系综概念,马上庚认为它“不必要而且不合事实”。的确,“系综”至多只能看作是“分布”的同义词,将之解释为体系的某种复本集合,是画蛇添足。其实,在对单个宏观体系作热力学测量的时间尺度内,感知的是统计力学状态即分布,物理体系也是依分布制备的。考虑历史上的原因,也不必取消“系综”一词,只当它是“分布”的同义词就足够了。Gibbs 的系综理论,是一种公理化表述,约定分布函数的写法,不问其从何而来。对于正则分布,可以给出如下的最大熵原理的说明。

设决定体系哈密顿量的所有外部参数如粒子数、体积等变量均已给定,体系在相空间上的分布记作 $P(\mathbf{r}^N, \mathbf{p}^N)$, 加在体系上的唯一约束是如下的 U 为定值:

$$U = \int d\mathbf{r}^N d\mathbf{p}^N E(\mathbf{r}^N, \mathbf{p}^N) P(\mathbf{r}^N, \mathbf{p}^N) \equiv \langle E(\mathbf{r}^N, \mathbf{p}^N) \rangle, \quad (2)$$

此处 E 为体系能量。根据最大熵原理,体系的分布应取

$$P(\mathbf{r}^N, \mathbf{p}^N) = Z(\beta, N, V)^{-1} \exp[-\beta E(\mathbf{r}^N, \mathbf{p}^N)], \quad (3)$$

此处 Z 定义为如下的归一因子即“状态和”或配分函数:

$$Z(\beta, N, V) = \int d\mathbf{r}^N d\mathbf{p}^N \exp[-\beta E(\mathbf{r}^N, \mathbf{p}^N)]. \quad (4)$$

此分布 $P(\mathbf{r}^N, \mathbf{p}^N)$ 在唯一约束 U 下(或者说 U 为唯一可验信息时)最大化熵

$$S = - \int d\mathbf{r}^N d\mathbf{p}^N P(\mathbf{r}^N, \mathbf{p}^N) \log P(\mathbf{r}^N, \mathbf{p}^N). \quad (5)$$

(对于均匀分布,熵等于状态数的对数,因而熵可理解为有效状态数的对数。)这里约定,已适当选取单位使得玻尔兹曼常数 $k_B = 1$ 。分布参数 $\beta = 1/T$, 为温度 T 的倒数或倒温度,只要不引起混淆也直接称 β 为温度(统计力学的表达式中永远只出现 $k_B T$ 的组合)。这个分布在统计力学中称为正则系综分布或麦克斯韦—玻尔兹曼分布。平均能量 U 是热力学中的内能,可用配分函数 Z 表出:

$$U = - \frac{\partial}{\partial \beta} [\log Z(\beta, N, V)]. \quad (6)$$

统计力学研究处在环境中的体系。环境可以

是状态预设的测量仪器、其他系统或热浴。环境的理想模型称为热库。可以想象它的自由度要多大就有多大，其具体的粒子组成及动力学并不重要。其本质在于，它处于平衡态，为处于其中的体系提供能量交换，在与体系交换能量的过程中热库的状态不发生改变，或者说其状态改变永远可略。因而，热库永远处于平衡态，其最重要的属性是它的温度 β 。简而言之，热库是恒温的能量池。

值得注意，通常教科书给出的配分函数定义(4)中，均不指明积分范围。它应为分布函数(3)的支集。支集的指定，是统计力学计算的第一步，取决于所研究的物理体系和问题。支集在统计力学中扮演的角色，未受到应有的重视，后面将进一步讨论。

可验信息限定了能量均值即内能，但正则分布的能量具有涨落。内能可用配分函数 Z 关于温度 β 的导数表出，能量的涨落同样可用配分函数 Z 表出，能量的方差为

$$\begin{aligned} \langle(\delta E)^2\rangle &= \langle(E - \langle E \rangle)^2\rangle = \langle E^2 \rangle - \langle E \rangle^2 \\ &= Z^{-1} \left(\frac{\partial^2 Z}{\partial \beta^2} \right)_{N, V} - \left[Z^{-1} \left(\frac{\partial Z}{\partial \beta} \right)_{N, V} \right]^2 \\ &= \left(\frac{\partial^2 \log Z}{\partial \beta^2} \right)_{N, V} = - \left(\frac{\partial U}{\partial \beta} \right)_{N, V}. \end{aligned} \quad (7)$$

注意到等容热容量的定义 $C_V = (\partial U / \partial T)_{N, V}$ ，有 $\langle(\delta E)^2\rangle = \beta^2 C_V$ 。能量涨落正比于温度的平方，表明它可取作温度的一种度量，反映分子轨道混合的程度。这个结果将能量的自发涨落与因温度变化引起的能量变化的响应率联系起来，也预示着线性响应理论和涨落—耗散定理。

3 微正则系综

正则系综直接描述的是 $T-V-N$ 固定的恒温、恒容和恒质的体系。由正则分布可导出 $T-P-N$ 固定的吉布斯系综分布及其他系综分布，证明不同系综间的等价性。系综等价性的推导中要求体系和热浴都很大。正则分布中体积固定，而

吉布斯系综分布中体积有涨落，不过根据中心极限定理体系很大时涨落趋于零。

微正则系综，又称 NVE 系综，考虑体系能量处于以 E 为中心的无限窄区的极限。特别是在原理上，微正则系综看似最为简单，但它不对应于实际体系，除了计算上的不便之外，对熵和温度等的定义还存在含糊性。微正则系综中，可借助于能量小于 E 的相体积函数 $v(E)$ ，定义出3种熵。

相体积函数的定义对于量子力学和经典力学有所不同。通常有必要引入宽度为 ω 的以 E 为中心的归一核函数即光滑函数 $f\left(\frac{H-E}{\omega}\right)$ 。在量子力学中，分布用密度矩阵 $\hat{\rho}$ ：

$$\hat{\rho}(E) = \frac{1}{W} \sum_i f\left(\frac{H_i - E}{\omega}\right) |\psi_i\rangle \langle \psi_i|, \quad W = \sum_i f\left(\frac{H_i - E}{\omega}\right),$$

此处 H_i 和 $|\psi_i\rangle$ 为哈密顿量的本征值和本征矢。如果取微正则系综的极限 $\omega \rightarrow 0$ ，原本的 $\delta(H-E)$ 会出问题，因为能量面宽度小于能级间隔时能级计数可能为零。复杂体系的能级简并几乎仅偶然发生，最终能级计数随能量值离散变化，导数只能取零或无穷。因而，有必要通过光滑函数 f 保持能量面的宽度， NVE 系综成为 $NVE-\omega$ 系综。

$$\rho = \frac{1}{h^n C} \frac{1}{W} f\left(\frac{H-E}{\omega}\right), \quad W = \int dq^N dp^N \frac{1}{h^n C} f\left(\frac{H-E}{\omega}\right),$$

$$v(E) = \int_{H < E} dq^N dp^N \frac{1}{h^n C},$$

此处 C 为(如粒子等同性的)重复计数修正因子， W 相当于展宽的能量面的有效体积，因而有 $W = \omega(dv/dE)$ 。

关于系综与热力学的对应，玻尔兹曼只考察了理想气体，详细的深入分析是吉布斯完成的。3种熵的定义分别为玻尔兹曼熵 S_B 、体积熵 S_V 和面积熵 S_S ：

$$S_B = k \log W = k \log(\omega dv/dE), \quad S_V = k \log v,$$

$$S_S = k \log(dv/dE) = S_B - k \log \omega,$$

相应的温度可定义为 $1/T_V = dS_V/dE$ 和 $1/T_S = 1/T_B = dS_S/dE = dS_B/dE$ 。由体积熵可以求得

$$dE = T_V dS_V - \langle P \rangle dV,$$

对应于热力学第一定律，虽然由面积熵也可得到类似公式，但压强复杂且不对应于相应量的平

均。微正则系综的 T_v 或 T_s 虽均与 ω 无关, 但仍未能完全担负热力学温度的角色, 例如, 不指示热的流向。(记体系 1、2 及其复合体系的能量分别为 E_1 、 E_2 和 $E_{12} = E_1 + E_2$, 一般而言, $dS_{v1}/dE_1 = dS_{v2}/dE_2$ 不保证 $dS_{v1}/dE_1 = dS_{v2}/dE_2 = dS_{v,12}/dE_{12}$ 。只有在复合体系微正则系综的平均意义下, $dS_{v,12}/dE_{12} = \langle dS_{v1}/dE_1 \rangle_{E_{12}} = \langle dS_{v2}/dE_2 \rangle_{E_{12}}$ 。) 它们在处理复合体系时有较严重困难, 还因体系态密度未必单调而出现负温度。

微正则系综到底是否可与其他系综如正则系综等价? 从正则系综出发看微正则系综, 当且仅当微正则系综的 E 满足 $E=U$ 即为体系的内能时, 忽略涨落而两系综等价。相关文献强调了微正则系综一般而言不等价于其他系综。从正则系综推导微正则系综, 容易发现问题所在。热力学体系的能量在内能附近涨落。分子动力学模拟哈密顿动力学, 是能量守恒的。此时体系的总能量应在内能附近才有意义。然而, 体系的内能一般事先未知, 不可控制, 因而往往通过标度平均动能来调节, 或者借助模拟恒温器热库的算法如 Nosé-Hoover (NH) 恒温器。

4 相变与各态历经破缺

伊辛模型在温度 $T=0$ 时存在所有自旋平行排列的最低能量态。因为铁磁相互作用参数 $J>0$, 相邻自旋取向一致对能量有利, 但这样的有序对熵不利。然而, 在温度很低时, 熵对自由能的贡献被压低, 自旋取向有可能在宏观距离上倾向于一致, 即出现自旋长程关联或长程序, 于是, 即使外场不存在时 $M \equiv \langle \sum_i \sigma_i \rangle$ 也不为零。这种现象是自发磁化。可出现自发磁化的最高温度, 是伊辛模型的临界温度。

无磁场的伊辛模型, 关于自旋的上下取向是对称的。看来, 精确计算 M 的结果只会为零, 因为对于每一个总自旋 $m = \sum_i \sigma_i$ 为正的组态, 必有另一个 m 为负的对称组态, 彼此相消。自发磁化的对称破缺如何发生呢? 一个机制是引入“辅助

场”, 由它得到对称破缺的初始分布, 之后即使辅助场趋于零, 对称破缺仍可留下。不过, 比较自然的一个机制是以下讨论的各态历经破缺。

限定对总自旋 $m = \sum_i \sigma_i$ 为给定值 μ 的所有组态求和, 定义如下的“拟组态和” $y(\mu)$ 和“拟自由能” $g(\mu)$:

$$y(\mu) = \sum_{m=\mu} e^{\beta E(\sigma_i)}, \quad g(\mu) \equiv -T \log y(\mu),$$

则配分函数显然为 $Y(T, h, N) = \sum_{\mu} y(\mu) = \sum_{\mu} e^{-\beta g(\mu)}$, 而 $y(\mu)/Y$ 是观察到总自旋为 μ 的组态的概率。从上述关于自发磁化的讨论可以设想, 函数 $g(\mu)$ 作为一维有效势应该有如下的行为: 在高温下它为单阱的, 但是, 当温度低于临界温度时, 它为双阱势。外场可使势出现不对称, 特别地, 双阱势的两阱深度不等。重要的是, 一旦出现势垒, 其高度依磁畴表面积估计应 $\sim N^{(d-1)/d}$, 只要是按 N 的正幂次标度, 均将在热力学极限下趋于无限。除热力学极限外, 平衡态还涉及时间趋于无限的动力学极限。无限高势垒将使以下两个极限顺序不等价:

$$\lim_{T \rightarrow \infty} \lim_{V \rightarrow \infty}, \quad \lim_{V \rightarrow \infty} \lim_{T \rightarrow \infty}.$$

在后一顺序下, 只要温度不是极低, 体系仍有机会访问势垒两侧, 而在前一顺序下则不然。对于实际体系, 时间和体积均有限, 到底取哪一种顺序, 应由体系的具体过程确定。简而言之, 相变伴随着各态历经破缺。

气相或液相的数值模拟中, 初态一般取某种均匀态。如果体系的参数对应于气—液相变区, 类似于伊辛模型的情形, 一般得不到相共存。如果要得到相共存, 如作界面行为的模拟时, 必须选择很不同的初条件, 例如让所有粒子处于容器一侧, 然后令体系弛豫。气相和液相间最显著的差别在于密度, 尤其是远离临界点时。选择适当的截断半径, 一个粒子归入气相或液相的判据可取作以该粒子的位置为中心的截断球内的粒子数。各态历经破缺表现为, 上述两种相对极端的初始构型有极其不同的演化特征, 标记作气相的粒子将长时间处于气相, 液相粒子亦然。当模拟的体系的尺度不大时, 各态历经破缺现象不会十

分显著,气相或液相粒子的寿命均不长,彼此频繁转换。随着体系尺度的增大,气相或液相粒子的寿命将有显著增长。是否及如何出现突变,值得观察并深入探讨。

5 热力学极限及分布支集

最终证明统计力学与热力学的对应,依赖于热力学极限的存在,而这个极限的存在是后验的,与体系哈氏量的本性密切相关(Bogolyubov (1946), van Hove (1949), Yang-Lee (1952), Fisher-Ruelle (1966))。以巨正则系综为例,需要证明极限 $\lim_{V \rightarrow \infty} V^{-1} \log \Xi$ 存在。李-杨考虑了一个比较一般的体系哈密顿量,证明了其热力学极限的存在。马上庚述及李-杨证明的可能推广,并指出李-杨二体势并不合乎物理上的最重要的作用力如电磁相互作用。

如同在标准的统计力学教科书中一样,在第2节的配分函数定义式(4)中,未明确写出积分区域,即未指明分布函数的支集。上节讨论的各态历经破缺出现时,物理体系所访问的相空间区域发生改变,表明分布函数的支集有相应的变化。如何选取分布函数的支集,取决于所讨论的物理问题。例如,如果讨论晶体,首先须考虑晶格对称性,限定了分布函数的支集。

Gibbs 引入巨正则系综处理化学反应体系。以 $\text{H} + \text{H} \rightarrow \text{H}_2$ 即氢原子结合为氢分子的反应为例,比较一下量子力学和统计力学。量子力学从二氢原子的哈密顿量出发,在 Born-Oppenheimer 近似下分离出原子核自由度,然后解电子的薛定谔方程求电子能级 $U_i(\mathbf{R})$ 作为原子核距离 \mathbf{R} 的函数,再将束缚态解释为氢分子。在统计力学处理中,一开始就必须将氢原子和氢分子同时包括在巨正则系综中。从纯粹的氢原子哈密顿量出发,不可能由统计力学得到氢分子。纯氢原子体系的相空间,区别于氢原子和氢分子混合体系的相空间,与各态历经破缺有联系。一级相变可以看作是最简单的化学反应 $\text{A} \rightarrow \text{B}$, 必须用相应的巨正则系综处理。

无论分子动力学模拟或是蒙特卡罗模拟,不改变分布函数的支集是必要的前提。运用各种加速手段或加强采样方法时,应该谨慎,文献中似未予充分重视。

6 热力学对应:热和自由能

热力学是物理学的一个分支,研究热和温度及其与能量和功的关系。它定义宏观变量如内能、熵和压强,描述这些变量之间的普适关系,不涉及特定物质的特定性质。这样的一般法则集中体现在热力学第零到第三的4条定律中。

热力学研究的最初动力在于提升热机的效率。热力学最基本的概念是体系和环境,而最基本的对象为热力学状态和热力学过程。热力学体系是宏观物理对象。体系与给定条件的环境处于热力学平衡时,状态完全由表明宏观性质的物理化学变量即状态变量描述。体系的热力学平衡态,由有限个彼此独立的状态变量确定,它们张成热力学状态空间,其中一点对应于一个状态。状态变量分作强度量和广延量。前者在增减体系尺度或质量时不变,而后者与体系尺度成正比,有可加性。限定状态时至少要有一个广延量,否则定不下体系大小。熵并不符合广延量的直接表述的定义。热力学量有三类。第一类是外参数,属于外部环境,与体系内部状态无关。第二类是体系微观构象动力学函数的平均。第三类是热力学独有的典型量,可以称为热力学量。它们没有微观意义,只可在宏观水平上把握,完全属于分布函数。状态方程,也称物态方程,描述热力学体系中不完全独立的多个状态变量之间的关系。热力学过程是热力学状态的变化,对应于热力学状态空间中的路径。(非平衡过程没有路径表示。)热力学过程须小心定义独立变量和因变量。如等压过程中压强选作独立变量,预设其不变,可由其变化推导出因变量体积的变化。热力学只是唯象理论。例如,状态方程满足某些热力学约束,但因为与体系的物质构成性有关,不可能由热力学导出。由分子知识推导热力学关系及计算热力学

学量，是统计力学的任务。

热力学第一定律定义了内能，将之分为功和热，但只有内能是态变量，热和功依赖于过程而非态变量。最常见的功，为机械功 $W_{\text{mech}} = -pdV$ ，此处 p 为压强，外部对体系做功时体积减小， W_{mech} 为正。功的一般形式为 $W = \sum_i f_i dX_i$ ，此处 f_i 为广义力， X_i 为广义位移。

统计力学中，内能 U 为平均能量：

$$U(\beta, N, V) = \int d\mathbf{r}^N d\mathbf{p}^N E(\mathbf{r}^N, \mathbf{p}^N) P(\mathbf{r}^N, \mathbf{p}^N), \quad (8)$$

此处概率密度对于正则分布为 $P = e^{-\beta E} / Z$ ，此时温度 β 和粒子数 N 为外参数，体积 V 为独立变量。因而，内能变化为

$$dU = \int d\mathbf{r}^N d\mathbf{p}^N [P(\mathbf{r}^N, \mathbf{p}^N) dE(\mathbf{r}^N, \mathbf{p}^N) + E(\mathbf{r}^N, \mathbf{p}^N) dP(\mathbf{r}^N, \mathbf{p}^N)],$$

器壁对体系的限制导致体系哈密顿量随体积的变化有 $dE/dV = -p$ ，此处 p 与微观构象或组态无关，所以 $dE = -pdV$ ，方括号中第一项给出做功项。由概率密度有

$$\beta E(\mathbf{r}^N, \mathbf{p}^N) = -\log Z(\beta, N, V) - \log P(\mathbf{r}^N, \mathbf{p}^N), \quad (9)$$

因而方括号中第二项对应于

$$-\beta^{-1} \int d\mathbf{r}^N d\mathbf{p}^N \log P(\mathbf{r}^N, \mathbf{p}^N) dP(\mathbf{r}^N, \mathbf{p}^N) = -\beta^{-1} \int d\mathbf{r}^N d\mathbf{p}^N d[P(\mathbf{r}^N, \mathbf{p}^N) \log P(\mathbf{r}^N, \mathbf{p}^N)] = TdS. \quad (10)$$

即吸热项。(此处已考虑到 $\log Z$ 为常量且概率 P 归一，因而 $\log Z$ 的相应项无贡献。) 至此推导出热力学第一定律

$$dU = -pdV + TdS. \quad (11)$$

此式可直接推广到一般形式的功。如果允许粒子数变化，引入 $\mu = (\partial U / \partial N)_{\beta, V}$ ，则

$$dU = -pdV + TdS + \mu dN, \quad (12)$$

此处 μ 称为化学势，度量体系添加一个粒子所增加的内能。总而言之，改变哈密顿量引起的内能变化是功，而对应于分布改变的内能变化是热。

热力学第二定律，是可用于任何与热有系统的基本假定，它有多种不同的表述，解释自然中的不可逆现象。克劳修斯 (Clausius) 的一种表述是，热不可能自发地从冷处传至热处。以统计力学的语言可将热力学第二定律表述成最小原理：处于热库中的体系，自发过程中自由能总是

减小。此处的“减小”，指明热力学过程的方向，说的是“趋于平衡”，超出平衡统计力学的解释范围。平衡统计力学只表述到最小原理：所有状态中以平衡态的自由能为最小。

为使记号简单，以下用求和取代积分。对于任意两个分布 $\{P_i\}$ 和 $\{Q_i\}$ ，相对熵定义为 $D(Q, P) = \sum_i Q_i \log(Q_i/P_i)$ ，利用 $\log x \leq 1 - x$ ，令 $x = P_i/Q_i$ ，容易证明 $D(Q, P) \geq 0$ ，而且等号只在 $\{P_i\}$ 和 $\{Q_i\}$ 等同时成立。现在将体系的正则分布记作 $\{P_i\} = Z^{-1} \exp(-\beta E_i)$ ，假定 $\{Q_i\}$ 为体系的另外一个任意分布。根据相对熵的非负性，有

$$\begin{aligned} \sum_i Q_i \log(Q_i/P_i) &= \log Z + \beta \sum_i E_i Q_i + \sum_i Q_i \log Q_i \\ &= \log Z + \beta \langle E \rangle_Q - S(Q) \geq 0, \end{aligned} \quad (13)$$

此处 $\langle X \rangle_Q$ 记随机变量 X 在分布 Q 下的平均。由 P 和 Q 等同时等号成立，有

$$\log Z = -\beta \langle E \rangle_P + S(P) = -\beta U + S,$$

于是，

$$\begin{aligned} F(\beta, N, V) &\equiv F_P \\ &\equiv -\beta^{-1} \log Z(\beta, N, V) = U - TS, \end{aligned} \quad (14)$$

就是热力学中的亥姆霍兹自由能。依此，一般分布 Q 的亥姆霍兹自由能可定义为

$$F_Q \equiv \langle E \rangle_Q + T \sum_i Q_i \log Q_i = U_Q - TS_Q. \quad (15)$$

根据 (13)，可写

$$F_Q \geq F(\beta, N, V), \quad (16)$$

即平衡态的自由能为最小。顺便指出，这里的不等式关系与正则分布有最大熵的最大熵原理是一致的。值得注意，自由能 F_Q 定义中出现的温度 T 是平衡态分布 P 的，或者说是热库的。也就是说，任意状态的自由能，只对处于热库中的体系有定义。再运用 (12)，可得

$$dF = -SdT - pdV + \mu dN. \quad (17)$$

上式是平衡态过程即可逆过程的自由能变化，可特别记作 $(dF)_{\text{rev}}$ 。考虑非平衡态到平衡态恒温恒容过程，此时 $(dF)_{\text{irrev}} < (dF)_{\text{rev}} = 0$ ，可一般地写 $dF \leq -SdT - pdV + \mu dN$ 。以上给出了体系一般状态即非平衡态的自由能的定义，但与热力学过程方向有关的趋于平衡，不可由平衡分布导出。

7 时间之箭:不可逆性

宏观不可逆性被形象地称作“时间之箭”，主宰一切宏观现象。微观可逆性与宏观不可逆性似乎水火不相容，让人十分纠结。统计力学从历史发展看，分子运动论为代表的非平衡研究在先，而平衡研究在后，前者涉及有关碰撞等诸多的微观过程细节。庞加莱 (Poincaré) 证明了著名的复归定理：一个有限力学系统将无限多次返回无限接近初始状态的某点。这曾是困扰玻尔兹曼一生的魔咒。吉布斯提出关于平衡统计分布的公理化表述，取得了极大的成功。(吉布斯著作发表后，便立即得到庞加莱的推崇。)在吉布斯理论中，魔咒踪影不见，但仍未驱除。其实，微观可逆性说的是分子轨道的演化，而宏观不可逆性说的是分布的演化，二者本来未必对立。

1953年夏，Fermi-Pasta-Ulam (FPU) 用当时刚有不久的计算机 MANIAC，模拟了耦合振子的一维非线性格子，期望看到能量在不同模之间均分，但看到的却是庞加莱复归现象，这被称作 FPU 佯谬。(现在许多人认为 FPU 佯谬应称作 FPUT 佯谬，以肯定希腊裔女学者 Tsingou 的贡献，她当时是 MANIAC 的程序员。)费米实验设计看的是分子轨道，当然只会看到庞加莱复归。如果要分布，则应考虑初态集合并允许轨道混合。

理解分子轨道和分布之间的联系，统计力学是一端，非线性动力学是另一端。在如何用分布语言描述非线性动力学方面，科尔莫戈罗夫提出，物理上有意义的不变分布，是系统加噪声但噪声强度趋于零后仍然存活的不变分布。揉面映射 (Baker map: 当 $0 \leq x < \frac{1}{2}$ 时, $(x', y') = (2x, \frac{1}{2}y)$; 当 $\frac{1}{2} \leq x < 1$ 时, $(x', y') = (2x - 1, \frac{1}{2}(y + 1))$ 。)属于最简单的可逆动力学的例子，值得在揉面映射这样的系统中演示科尔莫戈罗夫的思想。

8 分布函数的时间演化

体系的统计力学状态是支在微观相空间上的

分布函数。统计力学中处理的分布以及宏观热力学量的时间演化，考虑的时间尺度，比 ($6N$ 维相空间中的) 分子轨道微观动力学时间尺度大许多。吉布斯系综理论不含时间，约定了平衡分布函数的写法，只说明平衡态的自由能与非平衡分布相比最小，但并不指示趋于平衡。刘维尔方程，本质上等价于分子轨道动力学方程，并非合适的出发点。吉布斯理论中轨道演化和分布演化的时间尺度彼此分离的原则，也适用于一般分布或非平衡分布。贯彻吉布斯平衡系综理论的精神，需要放弃从分子轨道动力学方程导出分布演化动力学方程的企图，如同不企图从哈密顿力学方程导出量子力学薛定谔方程一样，有必要重新引入新的原理刻画分布函数的时间演化。

分布演化方程的一个必要条件是，平衡分布为其定态解。体系的演化动力学，以主方程描述最为自然而直接。考虑分布演化的主方程由转移概率 $T(\mathbf{z} \rightarrow \mathbf{z}')$ 限定：

$$P_{t+\tau}(\mathbf{z}) = \int d\mathbf{z}' P_t(\mathbf{z}') T(\mathbf{z}' \rightarrow \mathbf{z}), \quad (18)$$

此处转移概率 $T(\mathbf{z} \rightarrow \mathbf{z}')$ 完全决定于体系哈密顿量，满足如下细致均衡条件：

$$P_{\text{eq}}(\mathbf{z}') T(\mathbf{z}' \rightarrow \mathbf{z}) = P_{\text{eq}}(\mathbf{z}) T(\mathbf{z} \rightarrow \mathbf{z}'),$$
$$\frac{T(\mathbf{z} \rightarrow \mathbf{z}')}{T(\mathbf{z}' \rightarrow \mathbf{z})} = \frac{P_{\text{eq}}(\mathbf{z}')}{P_{\text{eq}}(\mathbf{z})} = \frac{e^{-\beta H(\mathbf{z}')}}{e^{-\beta H(\mathbf{z})}}. \quad (19)$$

上述分布动力学可保证体系趋于正则平衡分布 $P_{\text{eq}}(\mathbf{z})$ 。如果将之作为统计力学的基本原理，则可表述如下：

处于热库或环境中的体系，其分布函数演化的动力学由满足细致均衡条件 (19) 的主方程 (18) 描述。

值得指出，这里的细致均衡条件是对原始的 $6N$ 维变量而言的，虽然在约化变量空间里细致均衡条件未必成立。

转移概率矩阵 T 满足 $\int d\mathbf{z}' T_k(\mathbf{z}' \rightarrow \mathbf{z}) = 1$ ，有本征右矢 $\mathbf{1}$ 。如果 T 是不可约的，则由 Perron-Frobenius 定理知，其谱半径为 1，相应于本征右矢 $\mathbf{1}$ 的本征左矢为正矢量，对应于平衡分布。有限马尔科夫链的推广是紧致转移算符理论，包括 Krein-Rutman 定理。另外，离散时间也可推广为连续时间。

转移概率 $T(\mathbf{z} \rightarrow \mathbf{z}')$ 作为算符通常不是厄密的。引入如下的厄密化算符 t 比较方便:

$$t(\mathbf{z}, \mathbf{z}') = t(\mathbf{z}', \mathbf{z}) \equiv \sqrt{\frac{P_{\text{eq}}(\mathbf{z})}{P_{\text{eq}}(\mathbf{z}')}} T(\mathbf{z} \rightarrow \mathbf{z}').$$

将 T 的本征值为 λ 的本征左、右矢分别记作 $\Phi_\lambda(\mathbf{z})$ 和 $\Psi_\lambda(\mathbf{z})$, 令 $\Phi_1(\mathbf{z}) \equiv P_{\text{eq}}(\mathbf{z})$, 则 t 的本征矢为 $\phi_\lambda(\mathbf{z}) = \Phi_\lambda(\mathbf{z})/\phi_1(\mathbf{z})$, 此处 $\phi_1(\mathbf{z}) = \sqrt{\Phi_1(\mathbf{z})}$, 对应于本征值 1, 其余的所有本征值小于 1。写成 Dirac 记号形式, 记 $\Psi_\lambda(\mathbf{z}) \rightarrow |\Psi_\lambda(\mathbf{z})\rangle \equiv \phi_\lambda(\mathbf{z})|\phi_\lambda(\mathbf{z})\rangle$, $\Phi_\lambda(\mathbf{z}) \rightarrow \langle \Phi_\lambda(\mathbf{z})| \equiv \langle \phi_\lambda(\mathbf{z})|$, 则转移算符

$$T(\mathbf{z} \rightarrow \mathbf{z}') = \sum_\lambda \lambda |\Psi_\lambda(\mathbf{z}')\rangle \langle \Phi_\lambda(\mathbf{z})|,$$

$$t(\mathbf{z} \rightarrow \mathbf{z}') = \frac{\phi_1(\mathbf{z}')}{\phi_1(\mathbf{z})} T(\mathbf{z} \rightarrow \mathbf{z}') = \sum_\lambda \lambda |\phi_\lambda(\mathbf{z}')\rangle \langle \phi_\lambda(\mathbf{z})|.$$

此处已约定本征矢归一, 正交归一关系可表示为 $\langle \phi_\mu(\mathbf{z}) | \phi_\nu(\mathbf{z}) \rangle = \langle \Phi_\mu(\mathbf{z}) | \Psi_\nu(\mathbf{z}) \rangle = \delta_{\mu\nu}$ 。任意分布 $P(\mathbf{z})$ 对应于 $\langle P(\mathbf{z}) | = [\phi_\lambda(\mathbf{z})]^{-1} \langle p(\mathbf{z}) |$ 。

至于约化自由度的描述即物理现象的随机过程描述, 与粗粒化程度有关。热力学几乎只适于静态, 高频过程更多依赖微观描述。虽然朗之万方程可用于非马尔科夫随机力, 但不便处理非线性, 而福克尔—普朗克 (FP) 方程虽可用于非线性和非平稳情形, 但只适于白噪声。借助平衡解 P_{eq} , FP 方程可表作

$$\frac{\partial P(u, t)}{\partial t} = \frac{\partial}{\partial u} \left\{ D(u) \left[- \left(\frac{d}{du} \log P_{\text{eq}}(u) \right) + \frac{\partial}{\partial u} \right] \right\} P(u, t). \quad (20)$$

对于在势场 $U(x)$ 中的布朗粒子, $P_{\text{eq}}(x) = Ce^{-\beta U}$, 因而 $\partial_t P = \partial_x [D(x)(\beta \partial_x U + \partial_x)] P$ 。这样的分布演化方程, 满足趋于平衡, 漂移项依赖于平衡解, 驱动力将含平均场之类的有效场效应。考虑极性分子在非极性溶剂中的稀溶液, 电极化为 $\mathbf{P} = \mathbf{P}_d + \mathbf{P}_a + \mathbf{P}_c$, 分别来自偶极取向、距离和电荷分布, 后二者在红外和光学高频区有滞后, 在慢变场下可吸收入 ε_∞ , 仅须处理取向项。设密度 n 均匀, 平衡取向分布为

$$f_{\text{eq}}(\theta, \phi) = \frac{n e^{\beta \mu E \cos \theta}}{\int e^{\beta \mu E \cos \theta} d\Omega} \approx \frac{n}{4\pi} (1 + \beta \mu E \cos \theta),$$

$$\mathbf{P}_d = \langle \mu \cos \theta \mathbf{E} / E \rangle \approx \frac{1}{3} n \beta \mu^2 \mathbf{E}.$$

描述偶极矩的取向布朗运动的福克尔—普朗克方程为

$$\frac{\partial f}{\partial t} = D \left\{ \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left[\sin \theta \left(\frac{\partial}{\partial \theta} + \beta \mu E(t) \sin \theta \right) f \right] + \frac{1}{\sin^2 \theta} \frac{\partial^2 f}{\partial \phi^2} \right\}.$$

另一例子是核磁共振。响应函数等应可由统计力学导出, 并与关联函数联系, 表现为涨落—耗散定理。平衡统计力学提供静态响应函数, 非平衡统计力学着重讨论动态响应函数的推导。

统计力学必须以某种粗粒化将巨自由度的动力学演化转换为随机演化, 其基本问题是给出该转换的数学逻辑, 但尚未解决。久保指出, 平衡统计力学建立在各态历经上, 但进展有限。非平衡远为困难, 首先是概念宽泛, 须有所限制。两类限制之一是动理学方法, 考虑玻尔兹曼型方程, 但仅适用于平均自由程足够长, 外场频率足够低的情形, 不过不限于线性。另一类是近平衡过程, 将非平衡性质直接与平衡涨落相联系, 不依赖于上述的随机转换, 因而可用于随机转换不可行或不必要的情形。后者不假定马尔科夫性或高斯性。二者的适用性有所重叠。Van Kampen 一度曾严批线性响应理论, 认为轨道不稳定, 微扰计算无据, 微观线性与宏观线性完全不同, 后者须由玻尔兹曼方程处理。然而, 线性响应只微扰处理分布而非轨道, 刘维尔方程的确与哈密顿方程等价, 但噪声项或扩散项出现后不再等价。轨道不稳定性引起混合导致分布稳定性。动理学理论先随机化而后线性化, 线性响应理论则顺序相反。顺便指出, 多体微扰项非小量而为 $O(N)$, 宜用约化分布或累积量。

统计力学的核心概念是支在相空间上的分布。统计力学描述的体系, 其分布演化的时间尺度远大于相应微观体系的微观态演化或微观轨道的时间尺度。分布演化的动力学方程, 不可能由微观态演化方程导出。分布演化方程的一个必要条件是, 平衡分布为其长时间演化解。这样的演化方程中驱动项将与平衡分布相关, 理论处理的困难不见得来自非平衡, 更多地归结于平衡统

计力学方法本身的困难。前面曾指出,分布演化方程的一个简单候选,是转移概率满足细致

均衡条件的主方程,由之可导出约化的其他形式的方程。

参考文献

- [1] Brown L, Pippard B, Pais A (ed). Twentieth Century Physics. IOP Publishing Ltd, 1995. 该书第一卷第七章“平衡统计力学”由 C Domb 撰写,第八章“非平衡统计力学”由 M Dresden 撰写,两章文末均附了丰富的文献,篇幅不大,值得一读。该书的中文译本:20世纪物理学(第一卷)。Brown L, Pippard B, Pais A 编,刘寄星译。科学出版社,2014
- [2] 网页 <http://stp.clarku.edu/books/> 开列了统计力学和热力学的一些教科书的书目,可以看看
- [3] Gibbs J W. Elementary Principles of Statistical Mechanics. University Press, 1902. 该书也已有中文译本:统计力学的基本原理。吉布斯(毛俊雯译)。中国科技大学出版社,2016。吉布斯专著有很长的副标题,指明该书特别关注与热力学对应关系的基础
- [4] “微正则系综”一节,可参考英文维基百科“microcanonical

ensemble”词条,其中也引述了吉布斯 1902 年专著。较近的文
献有: Hilbert S, Hänggi P, Dunkel J. Phys. Rev. E, 2014, 90:
062116

- [5] 统计力学教科书通常不讨论热力学极限的存在问题,但马上庚的书(统计力学。环华出版,1982)和 R Balescu (Equilibrium and Nonequilibrium Statistical Mechanics. John Wiley & Sons, 1975) 均有专门的章节讨论。“热力学极限”一节引的文献为: Bogoliubov N N. J. Phys. USSR, 1946, 10: 257; Van Hove L. Physica, 1949, 15: 951; Yang C N, Lee T D. Phys. Rev., 1952, 87: 404; Fisher M E, Ruelle D. J. Math. Phys., 1966, 7: 260
- [6] 各态历经破缺”一节可参考 Chandler D. Introduction to Modern Statistical Mechanics. Oxford University Press, 1987 (中译本:现代统计力学导论。鞠国兴译。高等教育出版社,2013)

读者和编者

订阅《物理》得好礼

——超值回馈《岁月留痕》
—<物理>四十年集萃

为答谢广大读者长期以来的关爱和支持,《物理》编辑

部特推出优惠订阅活动:向编辑部连续订阅2年《物理》杂志,将获赠《岁月留痕—<物理>四十年集萃》一本。该书收录了1972年到2012年《物理》发表的40篇文章,476页精美印刷,定价68元,值得收藏。

希望读者们爱上《物理》!

订阅方式(编辑部直接订阅优惠价180元/年)

(1) 邮局汇款

收款人地址:北京市中关村南三街8号中科院物理所,100190

收款人姓名:《物理》编辑部

(2) 银行汇款

开户行:农行北京科院南路支行

户名:中国科学院物理研究所

帐号:112 501 010 400 056 99

(请注明《物理》编辑部)

咨询电话:010-82649470; 82649277

Email: physics@iphy.ac.cn

