

目录

1	牛顿力学	2
1.1	质点力学	2
1.2	粒子系力学	5
2	分析力学	7
2.1	广义坐标	7
2.2	虚功原理和达朗伯 (d'Alembert) 原理	8
2.3	拉格朗日 (Lagrange) 方程	9
2.4	哈密顿 (Hamilton) 方程	11
2.5	相空间	12
2.6	泊松 (Poisson) 括号和刘维尔算符	13
2.7	刘维尔定理	13
3	热力学	14
3.1	基本概念	14
3.2	热力学第一定律	16
3.3	热力学第二定律	18
3.4	热力学函数和关系	20
4	经典统计力学	25
4.1	微正则系综理论	25
4.2	正则系综理论	28
4.3	经典概率密度函数	31
4.4	巨正则系综理论	37

摘要

《分子动力学模拟入门》: 物理基础

在学习分子动力学模拟之前, 需要熟悉经典力学、热力学和统计物理。本章将简要地回顾这方面的主要知识。

1 牛顿力学

1.1 质点力学

理论上，牛顿力学里面最基本的研究对象是质点，它是指一个有特定的质量但其大小对所研究的问题不重要的物体。有人也称它为一个粒子。这也将是分子动力学模拟中最重要的研究对象之一。一个粒子在三维空间的运动由一个位置函数 $\vec{x}(t)$ 完全描述。也就是说，在一个时刻 t ，粒子的位置由一个有三个分量的矢量

$$\vec{x}(t) = x(t)\vec{e}_x + y(t)\vec{e}_y + z(t)\vec{e}_z$$

给出。当然，如果不指定一个参考系，我们将无法确定这三个分量的值。参考系的数学本质是坐标系。我们说粒子的位置函数完全地描述了粒子的运动性质，是因为其它的运动性质，比如速度和加速度，都可以由位置函数导出。

速度函数定义为位置函数对时间的一阶导数：

$$\vec{v}(t) = \frac{d\vec{x}(t)}{dt} = \dot{\vec{x}}(t) = \frac{dx}{dt}\vec{e}_x + \frac{dy}{dt}\vec{e}_y + \frac{dz}{dt}\vec{e}_z. \quad (1)$$

加速度函数定义为位置函数对时间的二阶导数，或者速度函数对时间的一阶导数：

$$\vec{a}(t) = \frac{d^2\vec{x}(t)}{dt^2} = \frac{d\vec{v}(t)}{dt} = \ddot{\vec{x}}(t) = \dot{\vec{v}}(t). \quad (2)$$

以上对粒子运动的描述就是所谓的运动学。下面我们讨论质点的动力学。

动力学研究物体运动及其改变的原因。牛顿运动定律让我们可以在一定的条件下确定一个粒子的位置函数 $\vec{x}(t)$ 。

牛顿第一定律。一个粒子在施加于它的合外力 \vec{F} 为零时保持静止或者一个恒定的速度 \vec{v} ：

$$\vec{F} = \vec{0} \Rightarrow \vec{v} = \text{const.} \quad (3)$$

该定律也叫惯性定律，或者惰性定律，意思是在合外力为零的情况下，一个粒子懒得改变它原来的运动状态：原来静止的话就继续静止；原来运动的话就继续原来的运动。

牛顿第二定律。一个粒子的动量 \vec{p} 的时间变化率 $d\vec{p}/dt$ 正比于作用于它的合外力 \vec{F} ：

$$\vec{F} = \frac{d\vec{p}}{dt}. \quad (4)$$

一个粒子的动量定义为其惯性质量 m 与速度 \vec{v} 的乘积：

$$\vec{p} = m\vec{v}$$

如果粒子的惯性质量是个常数的话，那么上式就变成

$$\vec{F} = m \frac{d\vec{v}}{dt} = m\vec{a}. \quad (5)$$

我们强调了这里的 m 是惯性质量，原因是牛顿力学里面还有另外一个质量，叫做引力质量，虽然它们是等价的。

牛顿第三定律。每一个作用力都有一个与之大小相等、方向相反的反作用力。例如，如果有一个由粒子 2 作用在粒子 1 上的作用力 \vec{F}_{12} ，那么一定同时存在一个由粒子 1 作用在粒子 2 上的反作用力 \vec{F}_{21} ，它们大小相等，方向相反：

$$\vec{F}_{12} = -\vec{F}_{21}. \quad (6)$$

当然，作用力和反作用力是相对的；把一对作用力中的哪一个叫做作用力，哪一个叫做反作用力是随意的。上式表达的是牛顿第三定律的弱形式。还有一个强形式的牛顿第三定律：

$$\vec{F}_{12} \propto \vec{r}_{12} \equiv \vec{r}_2 - \vec{r}_1 \quad (7)$$

显然，满足了牛顿第三定律的强形式，就一定满足了牛顿第三定律的弱形式，但反之不然。

我们在陈述牛顿第二定律时已经定义了动量。根据牛顿第二定律，如果作用在一个粒子上的合力为零，那么该粒子的动量将保持为一个常数：

$$\vec{F} = \vec{0} \Rightarrow \vec{p} = \text{常数}. \quad (8)$$

这就叫做动量守恒定律。

一个粒子的角动量被定义为：

$$\vec{L} = \vec{x} \times \vec{p}. \quad (9)$$

因为位置 \vec{x} 依赖于坐标系原点的选取，角动量也依赖于坐标原点的选择。类似地，我们定义作用在粒子上的力矩 $\vec{\tau}$ 为：

$$\vec{\tau} = \vec{x} \times \vec{F}. \quad (10)$$

其中， \vec{F} 是作用在粒子上的力。让我们来考察角动量的时间变化率：

$$\frac{d\vec{L}}{dt} = \frac{d(\vec{x} \times \vec{p})}{dt} = \vec{v} \times \vec{p} + \vec{x} \times \vec{F} = \vec{x} \times \vec{F} = \vec{\tau}. \quad (11)$$

这就是说，角动量的变化率等于力矩。显然，如果作用在粒子上的力矩为零，则粒子的角动量是常数：

$$\vec{\tau} = \vec{0} \Rightarrow \vec{L} = \text{常数}. \quad (12)$$

这叫做角动量守恒定律。

除了动量，还可以定义一个仅仅与质量和速度有关的物理量，叫做动能：

$$T = \frac{1}{2}m\vec{v}^2 = \frac{1}{2}m\vec{v} \cdot \vec{v} = \frac{1}{2}m(v_x^2 + v_y^2 + v_z^2). \quad (13)$$

动能如何得以改变呢？让我们来看看一个力对一个粒子做的功。在一个力 \vec{F} 的作用下，如果粒子运动了一个微分位移

$$d\vec{x} = dx\vec{e}_1 + dy\vec{e}_2 + dz\vec{e}_3$$

那么我们定义该力对该粒子做的微功为：

$$dW = \vec{F} \cdot d\vec{x} = F_x dx + F_y dy + F_z dz. \quad (14)$$

一般情况下，力是随时间变化的，可以写成时间的函数。如果在一个过程中，粒子受到一个力 $\vec{F}(t)$ ，而其位置的变化由轨道 $\vec{x}(t)$ ($t_1 < t < t_2$) 描述，那么我们说该过程中力 \vec{F} 对粒子做的总功为如下积分：

$$W = \int_{\vec{x}_1}^{\vec{x}_2} \vec{F} \cdot d\vec{x}. \quad (15)$$

其中，我们定义了 $\vec{x}_1 \equiv \vec{x}(t_1)$ 和 $\vec{x}_2 \equiv \vec{x}(t_2)$ 。力与微分位移的标量积可以变成如下形式：

$$\vec{F} \cdot d\vec{x} = \vec{F} \cdot \vec{v} dt = (\vec{F} dt) \cdot \vec{v} = d\vec{p} \cdot \vec{v} = m d\vec{v} \cdot \vec{v} = d\left(\frac{1}{2}m\vec{v}^2\right).$$

于是，我们有

$$W = \frac{1}{2}m(\vec{v}_2^2 - \vec{v}_1^2) \equiv T_2 - T_1. \quad (16)$$

其中，我们定义了 $\vec{v}_i \equiv \vec{v}(t_i)$ 和 $T_i \equiv \frac{1}{2}m\vec{v}_i^2$ ($i = 1, 2$)。

以上推导告诉我们：在一个过程中，外力对一个粒子做的功等于粒子动能的改变量。

在上面的讨论中，我们假定了一条连接起点和终点的具体的轨迹（即路径）。外力 \vec{F} 作的功一般来说依赖于路径的选取。然而，如果外力做的功与具体的路径无关，而只与起点和终点有关，那么由矢量分析的定理可知，该外力可以写成一个标量场 $U(\vec{x})$ 的梯度的负值（这个负号是一种约定）：

$$\vec{F} = -\nabla U(\vec{x}). \quad (17)$$

该标量场 $U(\vec{x})$ 叫做粒子的势能场，简称为势能，或者进一步简称为势。与此对应的力称为保守力。该保守力沿任意路径 $\vec{x}(t)$ ($t_1 < t < t_2$) 对粒子做的总功为：

$$W = - \int_{\vec{x}_1}^{\vec{x}_2} \nabla U(\vec{x}) \cdot d\vec{x} = - \int_{U_1}^{U_2} dU(\vec{x}) = U_1 - U_2. \quad (18)$$

其中，我们定义了 $U_1 = U(\vec{x}_1)$ 和 $U_2 = U(\vec{x}_2)$ 。将该式与前面得到的 $W = T_2 - T_1$ 比较可得：

$$T_1 + U_1 = T_2 + U_2. \quad (19)$$

这就是说，在一个保守力的作用下，任意过程中粒子的动能与势能的和都不改变。这个不改变的量称为机械能：

$$E = T + U. \quad (20)$$

所以，在保守力的作用下，粒子的机械能是守恒的。这就是机械能守恒原理。虽然称为原理，其实它并不是一个真正的原理，而是牛顿运动定律的推论。牛顿力学中真正的原理只有牛顿三大定律，其它的都是推论。

1.2 粒子系力学

在具体讨论粒子系的动力学行为之前必须先弄清楚内力和外力的区别。内力是系统中某个粒子作用于另一个粒子的，而外力来自于系统之外。内力满足牛顿第三定律。虽然外界施与粒子 i 一个力 \vec{F}_i^{ext} 的同时，粒子 i 同时也施与外界一个大小相等、方向相反的反作用力，但由于我们的系统不包含外界，我们通常不会对外力利用第三定律。以上就是内力与外力的区别。

对任意一个粒子 i ，我们可以写下它的动力学方程：

$$m_i \ddot{\vec{x}}_i = \sum_{j \neq i} \vec{F}_{ij} + \vec{F}_i^{\text{ext}}. \quad (21)$$

将上式左右两边都对指标 i 求和，得

$$\sum_i m_i \ddot{\vec{x}}_i = \sum_i \sum_{j \neq i} \vec{F}_{ij} + \sum_i \vec{F}_i^{\text{ext}}. \quad (22)$$

根据牛顿第三定律，等号右边的第一项等于零 (练习题)。系统的总质量定义为 $m = \sum_i m_i$ 。如果定义一个平均坐标

$$\vec{x} = \frac{\sum_i m_i \vec{x}_i}{m}, \quad (23)$$

那么我们有

$$m \ddot{\vec{x}} = \sum_i \vec{F}_i^{\text{ext}}. \quad (24)$$

这个式子看上去很像一个质量为系统总质量 m ，坐标为平均坐标 \vec{x} 的粒子的动力学方程；该粒子所受的合外力为整个粒子系所受的合外力。我们称这个等效的粒子为粒子系的质心。质心的质量就是整个系统的质量；质心坐标就是上述平均坐标；质心的运动满足上述等效的牛顿第二定律。

由质心坐标可以定义质心速度

$$\dot{\vec{x}} = \frac{\sum_i m_i \dot{\vec{x}}_i}{m} \quad (25)$$

和质心动量

$$\vec{p} = m\dot{\vec{x}} = \sum_i m_i \dot{\vec{x}}_i. \quad (26)$$

于是，质心的牛顿第二定律可用质心动量表达为：

$$\frac{d\vec{p}}{dt} = \sum_i \vec{F}_i^{\text{ext}}. \quad (27)$$

如果系统受到的合外力为零，那么系统的质心动量（即系统的总动量）是守恒的。这就是质点系的动量守恒定律。

类似地，由牛顿第三定律可以证明内力对系统的总力矩的贡献也是零，即系统所受的总力矩等于外力的总力矩 $\vec{\tau}^{\text{ext}}$ ：

$$\vec{\tau}^{\text{ext}} = \sum_i^N \vec{x}_i \times \vec{F}_i^{\text{ext}}. \quad (28)$$

如果定义系统的总角动量为

$$\vec{L} = \sum_i^N \vec{L}_i = \sum_i^N \vec{x}_i \times \vec{p}_i, \quad (29)$$

那么可以证明如下的角动量定理：

$$\frac{d\vec{L}}{dt} = \vec{\tau}^{\text{ext}}. \quad (30)$$

这就是说，外力产生的总力矩等于系统总角动量的时间变化率。如果外力产生的总力矩等于零，则有系统的总角动量守恒。这就是质点系的角动量守恒定理。

不同于力和力矩，内力对系统中的粒子做的总功并不一定为零。例如，两个粒子在相互的排斥力的作用下从静止开始做反向的加速运动，它们之间的内力做了正功。在一个微小过程中，如果第 i 个粒子的微分位移为 $d\vec{x}_i$ ，那么内力对系统中的粒子做的功为

$$dW^{\text{int}} = \sum_i^N \vec{F}_i^{\text{int}} \cdot d\vec{x}_i, \quad (31)$$

外力对系统中的粒子做的功为

$$dW^{\text{ext}} = \sum_i^N \vec{F}_i^{\text{ext}} \cdot d\vec{x}_i. \quad (32)$$

系统中的粒子得到的总功为

$$dW = dW^{\text{ext}} + dW^{\text{int}}. \quad (33)$$

因为对每一个粒子，有

$$(\vec{F}_i^{\text{int}} + \vec{F}_i^{\text{ext}}) \cdot d\vec{x}_i = dT_i, \quad (34)$$

所以，综合起来有

$$dW = dT. \quad (35)$$

其中， T 定义为系统的总动能：

$$T = \sum_i^N T_i = \sum_i^N \left(\frac{1}{2} m_i \vec{v}_i^2 \right). \quad (36)$$

这就是说，内力和外力做的功的总和等于系统总动能的改变。只有内力和外力做的总功为零时，系统的动能才守恒。

对于有相互作用的多粒子系统，如果内力和外力都是保守力，那么系统的总势能可以表达为

$$U = \sum_i U_i + \frac{1}{2} \sum_i^N \sum_{j \neq i}^N U_{ij}, \quad (37)$$

其中， U_i 是第 i 个粒子在外力场中的势能， U_{ij} 是系统中由 i 与 j 的相互作用导致的势能。

2 分析力学

2.1 广义坐标

我们知道，一个由 N 个质点构成的力学系统需要用 $3N$ 个坐标分量及其相应的速度分量来描述其运动状态。我们说，该系统的力学自由度为 $3N$ 。但如果是一个由这 N 个质点构成的刚体，则自由度只有 6：3 个平动自由度和 3 个转动自由度。刚体的自由度之所以小于自由质点系统的自由度，是因为刚体系统中有约束。

约束的分类很复杂，但我们只关注最简单的一种：稳定的几何约束。考虑一个由 N 个质点构成的力学系统，其中第 i 个质点的位置为 \vec{x}_i 。一个稳定的几何约束一般可以表达为：

$$f(\vec{x}_1, \vec{x}_2, \dots) = 0. \quad (38)$$

其中， f 是个一般的函数。

容易看出，一个原本有 M 个自由度的体系，若受到 m 个几何约束，则其自由度为 $s = M - m$ 。要描述该体系，我们不一定需要使用原来的 M 个坐标和速度，而可以构造 s 个新的坐标和速度。这 s 个新的坐标不一定限于在原来的 M 个老坐标中挑选，而完全可以另行选取。这样选取的新坐标叫做广义坐标。广义坐标不一定能三个一组地构成矢量，而且某个广义坐标也不一定具有长度的量纲。相应地，某个广义速度也不一定具有速度的量纲。

2.2 虚功原理和达朗伯 (d'Alembert) 原理

在有约束的情况下，力学系统各个部分虽然不能随心所欲地运动，但总能作某种运动。我们把考虑约束之前的某个自由度在有约束的情况下所具有的一切可能的位移叫做虚位移，记为 $\delta \vec{x}_i$ 。

系统中某个质点 i 除了可能受到普通的力 \vec{F}_i （叫做主动力）之外，还可能受到约束反力 \vec{N}_i 。若该质点保持力学平衡，则有

$$\vec{F}_i + \vec{N}_i = 0. \quad (39)$$

正如力与位移的标量积定义为功，力与虚位移的标量积定义为虚功。显然，质点 i 平衡时主动力与约束反力的虚功之和 δW_i 为零：

$$\delta W_i = (\vec{F}_i + \vec{N}_i) \cdot \delta \vec{x}_i = 0. \quad (40)$$

对所有质点进行求和，则有

$$\delta W = \sum_{i=1}^N \delta W_i = \sum_{i=1}^N (\vec{F}_i + \vec{N}_i) \cdot \delta \vec{x}_i = 0. \quad (41)$$

在很多重要的情况下，作为内力的各个约束反力做的总虚功为零：

$$\sum_{i=1}^N \vec{N}_i \cdot \delta \vec{x}_i = 0. \quad (42)$$

这种情况下的约束叫做理想约束。在理想约束作用下，我们有

$$\sum_{i=1}^N \vec{F}_i \cdot \delta \vec{x}_i = 0. \quad (43)$$

该式表达的就是虚功原理。

以上考虑的是静力学问题。在动力学问题中，质量为 m_i 、坐标为 \vec{x}_i 的质点的运动方程可以写为：

$$\vec{F}_i + \vec{N}_i = m_i \ddot{\vec{x}}_i. \quad (44)$$

若选取跟随质点一同运动的参考系，则上式可改写为

$$\vec{F}_i - m_i \ddot{\vec{x}}_i + \vec{N}_i = 0. \quad (45)$$

其中， $-m_i \ddot{\vec{x}}_i$ 可以理解为惯性力。类似于对虚功原理的推导，我们有如下等式：

$$\sum_{i=1}^N (\vec{F}_i - m_i \ddot{\vec{x}}_i + \vec{N}_i) \cdot \delta \vec{x}_i = 0. \quad (46)$$

在理想约束的作用下，我们有

$$\sum_{i=1}^N \vec{F}_i \cdot \delta \vec{x}_i - \sum_{i=1}^N m_i \ddot{\vec{x}}_i \cdot \delta \vec{x}_i = 0. \quad (47)$$

该式表达的就是达朗伯原理。它是下一小节讨论的出发点。

2.3 拉格朗日 (Lagrange) 方程

考虑 N 个质点组成的力学系统。质点 i 的质量和坐标分别为 m_i 和 \vec{x}_i ，其中 $i = 1, 2, \dots, N$ 。若有 m 个理想的几何约束，则可以定义 $s = 3N - m$ 个广义坐标 q_1, q_2, \dots, q_s 。通常将这些坐标的集合简记为 q 。原来的坐标可以表达为广义坐标的函数：

$$\vec{x}_i = \vec{x}_i(q) \quad (i = 1, 2, \dots, N). \quad (48)$$

运用上述坐标变换，可以将达朗伯原理的表达式写成

$$\vec{F}_i \cdot \frac{\partial \vec{x}_i}{\partial q_\alpha} \delta q_\alpha - m_i \ddot{\vec{x}}_i \cdot \frac{\partial \vec{x}_i}{\partial q_\alpha} \delta q_\alpha = 0. \quad (49)$$

从上式开始，我们用了爱因斯坦求和约定。另外要注意的是，虚位移不是原来的坐标所特有的，新的广义坐标也可以有虚位移，因为一个坐标如果没有虚位移那就没有让它成为坐标的必要了。

因为新的广义坐标是相互独立（因为广义坐标个数等于自由度个数，我们总可以做到这一点）的，所以根据上式我们有下式：

$$\vec{F}_i \cdot \frac{\partial \vec{x}_i}{\partial q_\alpha} - m_i \ddot{\vec{x}}_i \cdot \frac{\partial \vec{x}_i}{\partial q_\alpha} = 0. \quad (\alpha = 1, 2, \dots, s). \quad (50)$$

下面的任务就是从该式出发推导拉格朗日方程。

上式中的加速度 $\ddot{\vec{x}}_i$ 是比较讨厌的，我们可以用求导的乘积率换掉它：

$$m_i \ddot{\vec{x}}_i \cdot \frac{\partial \vec{x}_i}{\partial q_\alpha} = m_i \frac{d}{dt} \left(\dot{\vec{x}}_i \cdot \frac{\partial \vec{x}_i}{\partial q_\alpha} \right) - m_i \dot{\vec{x}}_i \cdot \frac{d}{dt} \left(\frac{\partial \vec{x}_i}{\partial q_\alpha} \right) \quad (51)$$

上面的等式中出现了速度 $\dot{\vec{x}}_i$ 和位置 \vec{x}_i 的混合，还是比较讨厌。你是不是想：要是

$$\frac{\partial \vec{x}_i}{\partial q_\alpha} = \frac{\partial \dot{\vec{x}}_i}{\partial \dot{q}_\alpha}, \quad (52)$$

$$\frac{d}{dt} \left(\frac{\partial \vec{x}_i}{\partial q_\alpha} \right) = \frac{\partial \dot{\vec{x}}_i}{\partial q_\alpha} \quad (53)$$

就好了？这两个式子确实是成立的。对于第一个，我们证明右边等于左边：

$$\frac{\partial \dot{\vec{x}}_i}{\partial \dot{q}_\alpha} = \frac{\partial}{\partial \dot{q}_\alpha} \left(\frac{\partial \vec{x}_i}{\partial q_\beta} \dot{q}_\beta \right) = \frac{\partial \vec{x}_i}{\partial q_\beta} \delta_{\alpha\beta} = \frac{\partial \vec{x}_i}{\partial q_\alpha}. \quad (54)$$

对于第二个，我们证明左边等于右边：

$$\frac{d}{dt} \left(\frac{\partial \vec{x}_i}{\partial q_\alpha} \right) = \frac{\partial}{\partial q_\beta} \left(\frac{\partial \vec{x}_i}{\partial q_\alpha} \right) \dot{q}_\beta = \frac{\partial}{\partial q_\alpha} \left(\frac{\partial \vec{x}_i}{\partial q_\beta} \right) \dot{q}_\beta = \frac{\partial \dot{\vec{x}}_i}{\partial q_\alpha} \quad (55)$$

利用它们，我们马上有：

$$\begin{aligned} m_i \ddot{\vec{x}}_i \cdot \frac{\partial \vec{x}_i}{\partial q_\alpha} &= m_i \frac{d}{dt} \left(\dot{\vec{x}}_i \cdot \frac{\partial \vec{x}_i}{\partial \dot{q}_\alpha} \right) - m_i \dot{\vec{x}}_i \cdot \frac{\partial \dot{\vec{x}}_i}{\partial q_\alpha} \\ &= \frac{d}{dt} \left[\frac{\partial}{\partial \dot{q}_\alpha} \left(\frac{m_i \dot{\vec{x}}_i \cdot \dot{\vec{x}}_i}{2} \right) \right] - \frac{\partial}{\partial q_\alpha} \left(\frac{m_i \dot{\vec{x}}_i \cdot \dot{\vec{x}}_i}{2} \right) \end{aligned} \quad (56)$$

上式中第二个等号右边的两对小括号中的表达式都是系统的动能 T 。于是，我们最终可以将达朗伯原理的表达式写成：

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_\alpha} \right) - \frac{\partial T}{\partial q_\alpha} = \vec{F}_i \cdot \frac{\partial \vec{x}_i}{\partial q_\alpha}. \quad (\alpha = 1, 2, \dots, s) \quad (57)$$

这就是我们要推导的拉格朗日方程。右边的表达式叫做广义力。对于保守力，主动力的虚功可以写成

$$\vec{F}_i \cdot \frac{\partial \vec{x}_i}{\partial q_\alpha} \delta q_\alpha = \vec{F}_i \cdot \delta \vec{x}_i = -\delta V. \quad (58)$$

其中， V 是体系的势能函数。于是，

$$\vec{F}_i \cdot \frac{\partial \vec{x}_i}{\partial q_\alpha} = -\frac{\delta V}{\delta q_\alpha} = -\frac{\partial V}{\partial q_\alpha}. \quad (59)$$

所以，对于保守力体系，拉格朗日方程可写为：

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_\alpha} \right) - \frac{\partial T}{\partial q_\alpha} = -\frac{\partial V}{\partial q_\alpha}. \quad (\alpha = 1, 2, \dots, s) \quad (60)$$

势能 V 仅仅是广义坐标的函数，而不是广义速度 \dot{q}_α 的函数，故有

$$\frac{d}{dt} \left(\frac{\partial(T-V)}{\partial \dot{q}_\alpha} \right) - \frac{\partial(T-V)}{\partial q_\alpha} = 0. \quad (\alpha = 1, 2, \dots, s) \quad (61)$$

这里出现的动能与势能的差是一个很重要的量，叫做拉格朗日量，记为

$$L(q, \dot{q}) = T(\dot{q}) - V(q). \quad (62)$$

用拉格朗日量可以将拉格朗日方程写成更加简洁的形式：

$$\boxed{\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\alpha} \right) - \frac{\partial L}{\partial q_\alpha} = 0. \quad (\alpha = 1, 2, \dots, s)} \quad (63)$$

例子：简谐振子的拉格朗日量为

$$L = \frac{1}{2} m \dot{x}^2 - \frac{1}{2} k x^2 \quad (64)$$

由此可以推导出简谐振子的运动方程。

2.4 哈密顿 (Hamilton) 方程

由于拉格朗日量具有能量的量纲，故当广义坐标具有长度的量纲时，偏导数 $\frac{\partial L}{\partial \dot{q}_\alpha}$ 具有动量的量纲。我们称这个偏导数为广义动量，记为

$$\boxed{p_\alpha = \frac{\partial L}{\partial \dot{q}_\alpha}}. \quad (65)$$

注意，当广义坐标不具有长度的量纲时，相应的广义动量也不具有动量的量纲。运用广义动量可以将拉格朗日方程写成更加简洁的形式：

$$\boxed{\dot{p}_\alpha - \frac{\partial L}{\partial q_\alpha} = 0. \quad (\alpha = 1, 2, \dots, s)} \quad (66)$$

虽然广义动量出现在拉格朗日方程中，但我们要清楚的是拉格朗日量是广义坐标和广义速度的函数，而不是广义动量的函数。用勒让德变换可以改变一个多元函数的独立变量。为此，我们先写下拉格朗日量的全微分：

$$dL(q, \dot{q}) = \frac{\partial L}{\partial q_\alpha} dq_\alpha + \frac{\partial L}{\partial \dot{q}_\alpha} d\dot{q}_\alpha = \dot{p}_\alpha dq_\alpha + p_\alpha d\dot{q}_\alpha. \quad (67)$$

下面的勒让德变换能将独立变量从广义速度变成广义动量：

$$L \rightarrow p_\alpha \dot{q}_\alpha - L. \quad (68)$$

该式箭头右边的表达式是一个很重要的量，记为

$$H = p_\alpha \dot{q}_\alpha - L. \quad (69)$$

该量称为哈密顿量。我们来看看它是不是真的是广义动量的函数了。为此，我们计算它的全微分：

$$dH = p_\alpha d\dot{q}_\alpha + \dot{q}_\alpha dp_\alpha - (\dot{p}_\alpha dq_\alpha + p_\alpha d\dot{q}_\alpha) = \dot{q}_\alpha dp_\alpha - \dot{p}_\alpha dq_\alpha. \quad (70)$$

看来哈密顿量确实是广义动量以及广义坐标的函数，而不是广义速度的函数了。

既然哈密顿量是广义动量和广义坐标的函数，那么根据全微分的定义，我们又有

$$dH = \frac{\partial H}{\partial q_\alpha} dq_\alpha + \frac{\partial H}{\partial p_\alpha} dp_\alpha. \quad (71)$$

对比以上两个式子，我们得到如下两组重要的方程：

$$\dot{q}_\alpha = \frac{\partial H}{\partial p_\alpha}, \quad (\alpha = 1, 2, \dots, s) \quad (72)$$

$$\dot{p}_\alpha = -\frac{\partial H}{\partial q_\alpha}. \quad (\alpha = 1, 2, \dots, s) \quad (73)$$

这 $2s$ 个一阶微分方程组称为哈密顿正则方程。这里的“正则”意为“简单且对称”。

例子：简谐振子的哈密顿量。

2.5 相空间

广义坐标和广义动量是相互独立的变量。我们可以将 s 个广义坐标 $\{q_\alpha\}_{\alpha=1}^s$ 和 s 个广义动量 $\{p_\alpha\}_{\alpha=1}^s$ 看成一个 $2s$ 维“空间”的“坐标”。这个抽象的“空间”叫做相空间。一组给定的广义坐标和广义动量叫做相空间的一个相点。另外，根据哈密顿正则方程，只要给定一个初始条件，即初始时刻的各个广义坐标和广义动量，就可以唯一地确定任意时刻的各个广义坐标和广义动量，即如下 $2s$ 个函数：

$$q_\alpha = q_\alpha(t), \quad p_\alpha = p_\alpha(t). \quad (\alpha = 1, 2, \dots, s) \quad (74)$$

这 $2s$ 个函数将在 $2s$ 维的相空间给出一条轨迹，叫做相轨迹。随着时间的推移，一条相轨迹会在相空间跑动，历经很多不同的相点。一个系统中不同的初始条件会给出不同的相轨迹，而两条不同的相轨迹绝对不会相交于某一个相点（请读者自行证明）。

例子：简谐振子的相空间。

2.6 泊松 (Poisson) 括号和刘维尔算符

一个一般的物理量可以表达为相空间坐标的函数

$$A = A(q_\alpha, p_\alpha)$$

我们求它的时间导数

$$\frac{dA}{dt} = \frac{\partial A}{\partial q_\alpha} \dot{q}_\alpha + \frac{\partial A}{\partial p_\alpha} \dot{p}_\alpha$$

利用哈密顿正则方程，可得

$$\frac{dA}{dt} = \frac{\partial A}{\partial q_\alpha} \frac{\partial H}{\partial p_\alpha} - \frac{\partial A}{\partial p_\alpha} \frac{\partial H}{\partial q_\alpha}$$

定义任意两个物理量之间的泊松括号

$$\{A, B\} = \frac{\partial A}{\partial q_\alpha} \frac{\partial B}{\partial p_\alpha} - \frac{\partial A}{\partial p_\alpha} \frac{\partial B}{\partial q_\alpha}$$

我们有

$$\frac{dA}{dt} = \{A, H\}$$

一个物理量 A 和哈密顿量之间的泊松括号运算也常用刘维尔算符表示

$$\frac{dA}{dt} = \{A, H\} \equiv iLA$$

该方程的形式解为

$$A(t) = e^{iL} A(t=0)$$

可以证明，刘维尔算符是厄米算符，故上述指数算符是么正算符。我们以后会大量使用刘维尔算符进行推导。

2.7 刘维尔定理

就像可以定义质量密度和电荷密度一样，也可以定义相空间的相点密度。首先定义相空间的体积元

$$d\Gamma = dq_1 dq_2 \cdots dq_s dp_1 dp_2 \cdots dp_s = dq dp \quad (75)$$

体积元并不是数学上严格的微分，其大小的选取满足“宏观小”和“微观大”的要求。上式中第二个等号右边是常用的简写形式。如果在该体积元中相点的个数为 dN ，那么就可以定义该体积元所在之处的相点密度：

$$\rho(q, p) = \frac{dN}{d\Gamma} \quad (76)$$

上式中的相点 (q, p) 可以是体积元 $d\Gamma$ 中的任意一点。如果将相空间类比于普通三维空间，将相点类比为三维空间中流体的质点，那么相点密度就对应于流体的质量密度。这个类比对下面的讨论是非常有用的。

下面我们来证明在统计物理中非常重要刘维尔定理：

$$\frac{d\rho}{dt} = 0. \quad (77)$$

这个定理是说：相点密度不随时间的变化而变化。首先，我们要清楚的是，虽然我们上面将相点密度写成了广义坐标和广义动量的函数，并不能说相点密度对时间的导数显然是零，因为广义坐标和广义动量可以是时间的函数。确切地讲，刘维尔定理说的是对任意的相轨迹，相点密度是一个常量。

正如上面提到的，我们将相空间和其中的相点比作流体。对于这样的“流体”，我们有如下守恒定律：

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial q_\alpha}(\rho \dot{q}_\alpha) + \frac{\partial}{\partial p_\alpha}(\rho \dot{p}_\alpha) = 0. \quad (78)$$

类似的公式在电磁学中叫做电荷守恒定律。因为相点是不会凭空产生和消失的，那么这个“流体”还不是一般的流体，而是不可压缩流体。对于不可压缩流体，上式中的“散度”处处为零：

$$\frac{\partial}{\partial q_\alpha}(\rho \dot{q}_\alpha) + \frac{\partial}{\partial p_\alpha}(\rho \dot{p}_\alpha) = 0. \quad (79)$$

于是，我们有

$$\frac{\partial \rho}{\partial t} = 0. \quad (80)$$

这就是说，相点密度不显含时间。这也是我们将相点密度写成 $\rho(q, p)$ ，而不是 $\rho(q, p, t)$ 的原因。

要证明刘维尔定理，就是要证明下式：

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial q_\alpha} \dot{q}_\alpha + \frac{\partial \rho}{\partial p_\alpha} \dot{p}_\alpha = 0. \quad (81)$$

将此式与上面“散度”为零的式子比较可知，只要能证明下式就大功告成了：

$$\frac{\partial \dot{q}_\alpha}{\partial q_\alpha} + \frac{\partial \dot{p}_\alpha}{\partial p_\alpha} = 0. \quad (82)$$

根据哈密顿正则方程，该式显然是恒成立的。刘维尔定理证毕。

3 热力学

3.1 基本概念

一个热力学系统是大量粒子的集合，常简称为系统。系统之外的所有物质称为环境。

如果系统与环境既不能交换能量也不能交换物质，则称其为一个孤立系统；若可交换能量但不能交换物质，则称为闭合系统；若即可交换能量也可交换物质，则称为开放系统。交换的能量可以是热能，称为热量，也可以是机械能，称为功。

如果我们说一个系统达到了热力学平衡，那么它的各个部分 (叫做子系统) 之间一定同时达到了热平衡、力学平衡和扩散平衡。

热力学中一共有四条基本经验定律，其中在时间上最后提出但在逻辑上排在第一个的叫热力学第零定律。该定律说如果两个系统分别与第三个系统能达到热平衡，那么这两个系统也一定能达到热平衡。这说明处于热平衡的两个系统之间有一个相等的量。这个量就是热力学中最重要的物理量之一：温度。这就是温度的定义之一。

通过规定两个特定系统的温度值，可以建立温标。一般选择一个大气压下水的冰点和沸点作为这两个特定系统。以下是两个常用温标：

(1) 摄氏 (Celsius) 温标规定冰点的温度为 0 摄氏度，沸点的温度为 100 摄氏度，而且将这中间的温度均分为 100 份。这是我们生活中常用的温标。

(2) 开尔文 (Kelvin) 温标规定冰点的温度为 273.15 K，沸点的温度为 373.15 K，而且将这中间的温度均分为 100 份。这是本书后面使用的温标。

考虑一个简单的孤立或者闭合系统，达到热力学平衡时，可由三个物理量描述。它们是：压强 p 、体积 V 和温度 T 。给定这三个量，就确定了系统的一个状态。事实上，这三个量并非完全独立，而是由状态方程联系着的。该方程可写为

$$f(p, V, T) = 0.$$

其中， f 是一个特定的三元函数。我们可以将其中任何一个量写成另外两个量的两元函数，比如 $V = V(p, T)$ 。这样的两元函数在由 p 、 V 、 T 这三个参数构成的三维空间中可以表示为一个曲面。系统的某一个状态就对应于该曲面上的一个点。

如果系统的状态发生了变化，我们称系统经历了一个过程。如果系统在其状态发生变化时始终无限接近平衡态，那么系统经历的过程为准静态过程。显然，准静态过程对应于状态方程曲面上的一条曲线。

理想气体是严格满足如下定律的系统：(1) 玻意耳-马略特 (Boyle-Mariotte) 定律：等温时压强反比于体积。(2) 查理 (Charles) 定律：等容时压强正比于温度。(3) 盖吕萨克 (Gay-Lussac) 定律：等压时体积正比于温度。

通过这些定律以及标准状态 (压强为一个大气压，温度为 0 摄氏度) 下的摩尔体积 (约 $22.4 \times 10^{-3} \text{ m}^3/\text{mol}$) 可以定出理想气体的状态方程：

$$pV = \nu RT.$$

这里， ν 是物质的量， $R \approx 8.31 \text{ J/mol/K}$ 是理想气体普适量。它与玻尔兹曼 (Boltzmann) 常量 $k_B \approx 1.38 \times 10^{-23} \text{ J/K}$ 由下式联系：

$$R \equiv N_A k_B. \quad (83)$$

这里， $N_A \approx 6.02 \times 10^{23} \text{ mol}^{-1}$ 是阿伏伽德罗 (Avogadro) 常量。注意到 $\nu N_A = N$ 就是系统的粒子数，我们可将理想气体的状态方程改写成

$$pV = Nk_B T. \quad (84)$$

再定义

$$n = N/V \quad (85)$$

为系统的数密度，我们又可以将理想气体的状态方程改写成

$$p = nk_B T. \quad (86)$$

对于单原子理想气体，可以证明：

$$p = \frac{1}{3} nm \langle \bar{v}^2 \rangle. \quad (87)$$

其中， m 是一个原子的质量， $\langle \bar{v}^2 \rangle$ 是原子速度平方的平均值。于是，我们有

$$\frac{3}{2} k_B T = \frac{1}{2} m \langle \bar{v}^2 \rangle. \quad (88)$$

这就是温度的微观意义之一：温度是原子平均平动动能的量度。

3.2 热力学第一定律

热力学第一定律是说，在一个过程中，如果环境对系统做了功，也传了热，那么系统的内能的增加量 ΔE 就等于环境对系统做的功 W 和传给系统的热 Q 的和：

$$\Delta E = Q + W. \quad (89)$$

如果系统对环境做功，则约定 $W < 0$ ；如果系统传给环境热量，则约定 $Q < 0$ 。这个定律是能量守恒定律在热力学系统中的应用。

如果考虑的是无限小过程，那么可以将热力学第一定律写成

$$dE = \delta Q + \delta W. \quad (90)$$

功和热量都不是状态量，而是过程量，即它们依赖于具体的过程。或者说，功和热量不能描述系统的状态。这点与内能不同。从数学的角度来说，状态量的微分是恰当微分，而过

程量的微分则不是。所以，为了区分，我们用 dE 表示内能的微分，而用 δW 和 δQ 表示微小的功和热量。

系统在吸热时一般温度会升高 (有反例)。因为热量与过程有关，所以将一个系统的温度升高一定的值所需要的热量依赖于系统所经历的过程。指定一个过程，就可以定义一个叫做热容的物理量：

$$C = \frac{\delta Q}{dT}. \quad (91)$$

常见的两个过程是等容过程和等压过程，对应的热容分别为等容热容和等压热容。

如果体积固定，系统与外界互不做功，热力学第一定律告诉我们 $\delta Q = dE$ 。从而，等容热容可表达为

$$C_V = \left(\frac{\delta Q}{dT} \right)_V = \left(\frac{\partial E}{\partial T} \right)_V. \quad (92)$$

如果压强固定 (但体积不固定)，系统要对环境做功 $p dV$ ，此时热力学第一定律告诉我们 $\delta Q = dE + p dV = d(E + pV)$ 。从而，等压热容可表达为

$$C_p = \left(\frac{\delta Q}{dT} \right)_p = \left(\frac{\partial H}{\partial T} \right)_p. \quad (93)$$

其中，我定义了一个新的类似于内能的热力学函数，焓：

$$H = E + pV. \quad (94)$$

综上所述：等容过程中系统吸收的热量等于其内能的增加量；等压过程中系统吸收的热量等于其焓的增加量。

若是一个过程中，系统与环境没有热交换，那么该过程叫做绝热过程。对理想气体来说，容易证明，绝热过程可以由下式描述：

$$pV^\gamma = \text{constant}. \quad (95)$$

其中，

$$\gamma \equiv \frac{C_p}{C_V}. \quad (96)$$

是等压热容和等容热容的比值，叫做绝热指数。

循环过程是系统终态等于初态的过程。在 $p - V$ 图中，循环过程对应于一个闭合路径。若闭合路径为顺时针方向，则系统对环境做净功并从环境中吸净热，对应于热机；反之，环境对系统做净功并从系统中吸净热，对应于制冷机。

理论上最重要的循环过程为理想气体的卡诺 (Carnot) 循环 (这里需要一个图，但是我先不画)。图中的 $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 1$ 为正向 (顺时针) 卡诺循环，对应卡诺热机，而 $1 \rightarrow 4 \rightarrow 3 \rightarrow 2 \rightarrow 1$ 为逆向 (逆时针) 卡诺循环，对应卡诺制冷机。其中 $1 \rightarrow 2$ 和 $3 \rightarrow 4$

都是等温过程，温度分别为较高温度 T_1 和较低温度 T_2 。这两个温度由环境中的热浴来保持。另外两个部分 ($2 \rightarrow 3$ 和 $4 \rightarrow 1$) 都是绝热过程。

根据热力学第一定律，对于卡诺热机，系统将在 $1 \rightarrow 2$ 的等温过程中从温度为 T_1 的高温热源吸热，并在 $3 \rightarrow 4$ 的等温过程中向温度为 T_2 的低温热源放热。计系统吸入和放出的热量为 Q_1 和 $|Q_2|$ ($Q_2 < 0$)，并定义热机的效率 η 为系统所做净功与从高温热源所吸热量的比值：

$$\eta = \frac{Q_1 - |Q_2|}{Q_1}. \quad (97)$$

很容易证明，这个效率其实只与温度有关而且总小于 1：

$$\eta = \frac{T_1 - T_2}{T_1} = 1 - \frac{T_2}{T_1}. \quad (98)$$

为什么热机的效率总小于 1 呢？即：为什么系统不能把吸收的热量全部转化为功呢？这是热力学第一定律回答不了的问题。要回答这个问题，我们需要学习热力学第二定律。

3.3 热力学第二定律

比较上面关于卡诺热机效率的两个公式可知 (注意 $Q_2 < 0$)

$$\frac{Q_1}{T_1} + \frac{Q_2}{T_2} = 0. \quad (99)$$

该式用微积分的语言可以写成

$$\oint \frac{dQ}{T} = 0. \quad (100)$$

克劳修斯证明了该式不光对卡诺循环成立，而且对任何可逆循环过程都成立。如果该是不成立，对应的循环就是不可逆的。对不可逆循环，克劳修斯证明了下述不等式：

$$\oint \frac{dQ}{T} < 0. \quad (101)$$

以上两式分别称为克劳修斯等式和不等式。这两个式子表述的内容被称为克劳修斯定理。

根据微积分知识，适用于可逆过程的克劳修斯等式意味着 $\frac{\delta Q}{T}$ 是一个恰当微分。或者说，积分 $\int_A^B \frac{\delta Q}{T}$ 与从初态 A 到终态 B 的路径无关。一个恰当微分总可以写成某个函数 (或者说变量；变量与函数的概念是相对的) 的全微分。我们用 dS 来表示这个全微分：

$$dS \equiv \frac{\delta Q}{T}. \quad (102)$$

中国老辈物理学家为这个新的热力学函数翻译了个不错的名字：熵。我们可以这样理解这个名字：第一，它是一个“商”，即热量除以温度；第二，它有个“火”字旁，表明它与热量有关。

用这个记号，就可以把热力学第一定律写成一个只涉及热力学变量的形式：

$$dE = TdS - pdV. \quad (103)$$

这个式子在热力学中特别重要，常被称为热力学基本方程。对热力学理论的探索常常从这个方程出发。

从这个方程可以看出，熵是一个广延量，或则说可加量。更重要的是，由于熵是一个状态量，除了一个积分常数（类似于势能的零点）它的值不依赖于系统所经历的过程。

考虑单原子理想气体，其内能为体系的动能

$$E = \frac{3}{2}Nk_B T.$$

证明从温度为 T_1 、体积为 V_1 的状态变化到温度为 T_2 、体积为 V_2 的状态的过程中，系统的熵增加量为

$$S_2 - S_1 = Nk_B \ln \left[\frac{V_2}{V_1} \left(\frac{T_2}{T_1} \right)^{3/2} \right]. \quad (104)$$

克劳修斯定理（即克劳修斯等式和不等式）可以说是热力学第二定律的一种数学表述。然而，该数学表述是基于一个循环过程的。很多情况下，考虑一个普通的过程或者一个无限小过程是更方便的。通过适用于可逆过程的克劳修斯等式，我们确立了一个恰当微分 $dS = \frac{\delta Q}{T}$ 。对此积分，便可得对应于一个从态 A 到态 B 的有限的可逆过程中的熵变：

$$S_B - S_A = \int_A^B \frac{\delta Q}{T}. \quad (105)$$

现在，我们假设通过一个任意的（可逆的或者不可逆的）过程让系统从态 B 回到态 A 。根据克劳修斯定理，我们有

$$S_B - S_A + \int_B^A \frac{\delta Q}{T} = \int_A^B \frac{\delta Q}{T} + \int_B^A \frac{\delta Q}{T} = \oint \frac{\delta Q}{T} \leq 0, \quad (106)$$

即

$$S_A - S_B \geq \int_B^A \frac{\delta Q}{T}. \quad (107)$$

令 A 状态无限接近 B 状态，即得

$$dS \geq \frac{\delta Q}{T}. \quad (108)$$

从推导的过程可知，这里的等号和不等号分别对应于克劳修斯定理中的等号和不等号，从而分别对应于可逆与不可逆过程。由于与克劳修斯定理等价，这个不等式也可以当做热力学第二定律的数学表述。

特别地，如果过程是绝热的，即 $\delta Q = 0$ ，那么有

$$dS \geq 0. \quad (109)$$

该式被称为熵增加原理。一个更特殊的情况是孤立系统中的过程。由于孤立系统不与环境交换物质与能量（当然，包括热能），无论它的内部发生何种变化，它的熵只能增加或不变，不能减少。

在结束这一节之前，我们考察一个不可逆过程：气体的自由膨胀。为简单起见，我们考虑单原子理想气体的绝热自由膨胀（即假定理想气体被装在绝热容器里面）。如果气体的体积从 V_1 自由地膨胀至 V_2 ，那么，由于其温度不变（气体与环境无热交换，自由膨胀的理想气体与环境也无功的交换，故内能不变，从而温度不变），系统熵的增量为（ N 为气体分子数）

$$\Delta S = Nk_B \ln \frac{V_2}{V_1} > 0. \quad (110)$$

于是，该绝热过程是不可逆的，与我们的直觉一致。

3.4 热力学函数和关系

到目前为止，我们只研究了与环境没有物质交换的孤立和闭合系统。对于开放系统，其粒子数 N 是允许变化的。为了能处理粒子数变化的过程，我们将 N 也视作一个热力学变量并将热力学基本方程推广为

$$\boxed{dE = TdS - PdV + \mu dN}. \quad (111)$$

这里，我们引入了一个新的强度量， μ ，叫做化学势。从上述内能的全微分可以看出，

$$\mu = \left(\frac{\partial E}{\partial N} \right)_{S,V}. \quad (112)$$

这个式子可以看作是化学势的一个定义。可以证明，理想气体的化学势是负的。理想气体的熵是随着 N 、 E 和 V 这些变量的增大而增大的。因此，如果要在增大 N 的同时将 S 和 V 固定，必须将 E 减小。因此，根据化学势的定义式，理想气体的化学势必然是负的。

热力学基本方程告诉我们内能是熵、体积以及粒子数的一个函数：

$$E = E(S, V, N). \quad (113)$$

该式中所有的变量都是广延量，意味着内能是这些变量的齐次函数，即

$$E(\lambda S, \lambda V, \lambda N) = \lambda E(S, V, N). \quad (114)$$

其中， λ 是一个任意的参数。将该式对参数 λ 求导可得

$$\begin{aligned} & \frac{\partial E(\lambda S, \lambda V, \lambda N)}{\partial(\lambda S)} \frac{\partial(\lambda S)}{\partial \lambda} + \frac{\partial E(\lambda S, \lambda V, \lambda N)}{\partial(\lambda V)} \frac{\partial(\lambda V)}{\partial \lambda} + \frac{\partial E(\lambda S, \lambda V, \lambda N)}{\partial(\lambda N)} \frac{\partial(\lambda N)}{\partial \lambda} \\ & = E(S, V, N). \end{aligned} \quad (115)$$

再将 λ 取为 1，可得

$$E = \frac{\partial E}{\partial S} S + \frac{\partial E}{\partial V} V + \frac{\partial E}{\partial N} N. \quad (116)$$

另外，从推广的热力学基本方程可知如下关系：

$$T = \left(\frac{\partial E}{\partial S} \right)_{V, N}$$

$$p = - \left(\frac{\partial E}{\partial V} \right)_{S, N}$$

以及

$$\mu = \left(\frac{\partial E}{\partial N} \right)_{S, V}$$

于是，我们得到一个重要的方程：

$$\boxed{E = TS - pV + \mu N}. \quad (117)$$

该方程被称为欧拉 (Euler) 方程。

由欧拉方程可以推导另一个重要的关系式。对欧拉方程两边作全微分，可得

$$dE = TdS + SdT - pdV - Vdp + \mu dN + Nd\mu. \quad (118)$$

将此与推广的热力学基本方程对照便知

$$\boxed{SdT - Vdp + Nd\mu = 0}. \quad (119)$$

该方程被称为吉布斯-杜安 (Gibbs-Duhem) 关系。这个关系告诉我们：三个强度量 (压强，温度，化学势) 中只有两个是独立的。因此，我们可以说：这样的系统的热力学自由度为 2。

根据欧拉方程，我们可以将熵表达为能量、体积以及粒子数的函数：

$$S = S(E, V, N) = \frac{E + pV - \mu N}{T}. \quad (120)$$

其全微分可由热力学基本方程得到：

$$dS = \frac{1}{T}dE + \frac{p}{T}dV - \frac{\mu}{T}dN. \quad (121)$$

于是，我们从能量函数过渡到了熵函数。从熵函数的全微分出发，可以更方便地讨论热力学第二定律，因为热力学第二定律更多地是与熵，而不是能量相联系。

我们在本讲第一节就提到了三个热力学平衡：热平衡，力平衡，以及扩散平衡。热力学第零定律告诉我们一个系统达到热平衡时内部温度处处相等，而基于力学的直觉告诉我们系统达到力平衡时应该处处压强相等。在这一小节，我们将从热力学第二定律的数学表述之一，即熵增加原理来更为严格地推导热平衡与力平衡的条件，以及我们还没有讨论过的扩散平衡条件。

一个孤立系统是不受外界干扰的，无论它的初始条件如何，在等待足够长的时间之后，它总会达到一个固定的状态，宏观上表现为达到热力学平衡。这与上一节得到的孤立系统的熵总是趋于最大化的结论不谋而合。所以，对热力学平衡的向往是孤立系统中熵增加的动力。如果孤立系统的熵在增加，那么它就还未达到热力学平衡；如果孤立系统的熵达到了一个最大值，不再增加了，它就达到了热力学平衡；在达到热力学平衡之后，孤立系统的熵就会保持最大值，不再变化。

为简单起见，我们考虑一个孤立系统，并用一个假想的边界将该系统分为两个子系统： A 和 B 。显然，这两个子系统都是开放系统，它们之间能够交换物质与能量。我们用下标 A 和 B 表示两个子系统的热力学量。由于熵是可加量，故整个孤立系统的熵的全微分可写为

$$dS = dS_A + dS_B = \frac{1}{T_A} dE_A + \frac{p_A}{T_A} dV_A - \frac{\mu_A}{T_A} dN_A + \frac{1}{T_B} dE_B + \frac{p_B}{T_B} dV_B - \frac{\mu_B}{T_B} dN_B. \quad (122)$$

因为整个系统的粒子数、体积、以及内能都是守恒的，我们有

$$dS = \left(\frac{1}{T_A} - \frac{1}{T_B} \right) dE_A + \left(\frac{p_A}{T_A} - \frac{p_B}{T_B} \right) dV_A - \left(\frac{\mu_A}{T_A} - \frac{\mu_B}{T_B} \right) dN_A. \quad (123)$$

若整个孤立系统达到了热力学平衡，则有 $dS = 0$ ，从而有如下平衡条件：

$$T_A = T_B \quad (\text{热平衡}), \quad (124)$$

$$p_A = p_B \quad (\text{力平衡}), \quad (125)$$

$$\mu_A = \mu_B \quad (\text{扩散平衡}). \quad (126)$$

前两个平衡条件与我们的之前的结果是一致的。最后一个平衡条件是说孤立系统在达到扩散平衡时，其内部的化学势处处相等。

如果系统仍未达到平衡态，熵增加原理告诉我们整个系统的广延量（即内能、体积及粒子数）会在子系统之间按照如下规则重新分配：

$$T_A > T_B \implies dE_A < 0, \quad (127)$$

$$T_A = T_B \text{ 且 } p_A > p_B \implies dV_A > 0, \quad (128)$$

$$T_A = T_B \text{ 且 } \mu_A > \mu_B \implies dN_A < 0. \quad (129)$$

也就是说，在趋向平衡的过程中，具有较高温度的子系统会失去内能以降低温度，具有较高压强的子系统会扩展体积以降低压强，具有较高化学势的子系统会失去物质（粒子）以降低化学势。所以，化学势是对粒子扩散趋势的一种量度。

无论是能量的全微分还是熵的全微分，其独立变量（也叫自然变量）都是广延量。然而，实验上更容易控制的是强度量。有没有一种方法将部分或全部独立变量用强度量替代呢？勒让德变换可以帮助我们做到这一点。勒让德变换可以将一个函数变换为另外相关的一个函数。具体地，新的函数是将原来的函数减去该函数的一个独立变量与对应的偏导数的乘积而得到的。

考虑能量函数 $E = E(S, V, N)$ 并取 V 为需要变换的独立变量，对应的勒让德变换将能量函数变换为焓函数：

$$H = H(S, p, N) = E - \left(\frac{\partial E}{\partial V} \right)_{S, N} V = E + pV = TS + \mu N. \quad (130)$$

其中，我们在最后一步用了欧拉公式。你可能也注意到上式暗示了焓函数的独立变量是 S 、 p 和 N 。这一论断能由下式证实（第二个等号运用了热力学基本方程）：

$$dH = dE + pdV + Vdp = TdS + Vdp + \mu dN. \quad (131)$$

取 S 为需要变换的独立变量，对能量函数进行勒让德变换，就可以得到亥姆霍兹 (Helmholtz) 自由能函数：

$$F = F(T, V, N) = E - \left(\frac{\partial E}{\partial S} \right)_{V, N} S = E - TS = -pV + \mu N. \quad (132)$$

它的独立变量为 T 、 V 和 N ，因为

$$dF = dE - TdS - SdT = -SdT - pdV + \mu dN. \quad (133)$$

如果同时针对 V 和 S 两个变量进行勒让德变换，则可得到吉布斯 (Gibbs) 函数：

$$G = G(T, p, N) = E - TS + pV = \mu N. \quad (134)$$

它的独立变量为 T 、 p 和 N ，因为

$$dG = -SdT + Vdp + \mu dN. \quad (135)$$

最后，如果取 N 为需要变换的独立变量并对亥姆霍兹自由能进行勒让德变换，我们将得到巨正则函数：

$$\Phi = \Phi(T, V, \mu) = F - \mu N = -pV. \quad (136)$$

它的独立变量为 T 、 V 和 μ ，因为

$$d\Phi = -SdT - pdV - Nd\mu. \quad (137)$$

内能和上述四个通过勒让德变换得到的函数统称为热力学势。我们能用勒让德变换定义更多的函数，但我们以后只会用到这几个。

因为热力学势都是系统的状态变量，它们的全微分都是恰当的。这意味着它们对任何两个独立变量的二阶偏导数与求偏导的次序无关。例如，根据

$$dE = TdS - pdV + \mu dN$$

我们有

$$\frac{\partial}{\partial V} \left(\frac{\partial E}{\partial S} \right) = \frac{\partial}{\partial S} \left(\frac{\partial E}{\partial V} \right). \quad (138)$$

将等式两边括号内的一阶偏导数用合适的变量替代，可得

$$\left(\frac{\partial T}{\partial V} \right)_{S,N} = - \left(\frac{\partial p}{\partial S} \right)_{V,N}. \quad (139)$$

这样得到的关系称为麦克斯韦 (Maxwell) 关系。显然，可以有很多麦克斯韦关系，我们就不一一列举了。

前面，我们根据熵增加原理讨论了孤立系统趋于平衡的过程。然而，孤立系统的模型往往不是讨论一个特定问题时的最佳选择。这里，我们来考察其它类型的系统 (闭合系统和开放系统) 趋于平衡的过程。

一个常见的例子是体积和粒子数固定的闭合系统，它与一个热浴接触而保持一个恒定的温度。对这样的非孤立系统，熵增加原理不再适用，因为系统与环境 (主要是热源) 可能会交换热量。为了研究此时系统的热力学演化行为，我们必须从热力学第二定律的更为一般的表达式，即 $TdS \geq \delta Q$ 出发。一方面，由于温度是恒定的，我们有 $TdS = d(TS)$ ；另一方面，由于粒子数和体积固定，热力学第一定律告诉我们 $\delta Q = dE$ 。结合这三个式子，并注意到亥姆霍兹自由能的定义，我们得到如下不等式：

$$dF \leq 0. \quad (140)$$

该式被称为自由能最小原理。它是说，一个粒子数、体积和温度固定的系统总是朝着自由能减小的方向演化的。如果系统的自由能还未达到最小值，那么它还未达到热力学平衡；如果系统的自由能达到了最小值，那么它就达到了热力学平衡。

另一个常见的例子是粒子数、压强和温度固定的系统。类似地，我们可以根据热力学第一和第二定律得到不等式 $d(TS) \geq dE + d(pV)$ 。利用吉布斯函数的定义可以将该不等式表达为

$$dG \leq 0. \quad (141)$$

这就是吉布斯函数最小原理，即一个粒子数、压强和温度固定的系统总是朝着吉布斯函数减小的方向演化。吉布斯函数减小时体系还未达到热力学平衡，吉布斯函数不变时体系便达到了热力学平衡。

以上两个原理统称为最小能量原理，它是力学中的最小势能原理在热力学中的推广。由于我们是从热力学第一和第二定律出发推导该原理的，它也可以被当做是热力学第二定律的数学表述之一。它与熵增加原理在本质上是等价的，而且在应用中是互为补充的。

4 经典统计力学

4.1 微正则系综理论

我们考虑一个孤立系统。这个系统具有固定的粒子数 N （系统与外界无粒子交换）、体积 V （系统与外界无体积功的交换）以及能量 E （系统与外界也不交换热量，故能量不变）。为了简化讨论，我们还假设系统中的粒子之间没有相互作用，即考虑理想气体系统。还是为了讨论方便，我们假设粒子的能量可以取一系列离散的值 E_i ($i = 1, 2, \dots$)，并设能量等于 E_i 的粒子的数目为 n_i 。于是，我们有：

$$N = \sum_i n_i, \quad (142)$$

$$E = \sum_i n_i E_i. \quad (143)$$

我们知道，从热力学的角度来看，系统的状态就由 N 、 V 和 E 来描述。这些量叫做宏观量，而一组宏观量就确定了一个宏观态。

然而，从微观的角度来说，即使一个系统处于一个确定的宏观态，系统中各个粒子的运动状态也可能是不确定的。系统中所有粒子的运动状态的组合构成一个微观态。一个 N 、 V 和 E 确定的宏观态可能具有多个微观态，而且微观态的个数一般来说是 N 、 V 和 E 的函数。我们将这个函数记为

$$\Omega = \Omega(N, V, E). \quad (144)$$

上面只考虑了一个孤立系统。现在考虑一个由子系统 1 和子系统 2 构成的复合系统。设这两子个系统之间可以相互传热，但不能相互传递粒子，也不能相互做体积功。于是，

每个子系统的粒子数和体积都不变，但能量可能随时间的推移而改变。设两个子系统之间的相互作用能可以忽略不计，则复合系统的总能量 E 就是每个子系统的能量的和：

$$E_1 + E_2 = E. \quad (145)$$

我们假定复合系统是孤立的，其总能量 E 是一个常数。

因为子系统 1 有 $\Omega_1(N_1, V_1, E_1)$ 个微观状态，子系统 2 有 $\Omega_2(N_2, V_2, E_2)$ 个微观状态，而子系统 1 的任何一个微观状态与子系统 2 的任何一个微观状态一起构成了复合系统的一个可能的微观状态，故由乘法原理可知，复合系统有

$$\Omega(N_1, V_1, E_1, N_2, V_2, E_2) = \Omega_1(N_1, V_1, E_1)\Omega_2(N_2, V_2, E_2) \quad (146)$$

个微观状态。我们现在的的问题是：总能量 E 如何在两个子系统之间分配，才能让两个子系统之间达到热力学平衡？也就是说，如果一开始两个子系统之间没有达到热力学平衡，那么能量 E_1 （或者能量 E_2 ）要如何改变才能使得两个子系统趋向热力学平衡？

要回答上面的问题，必须对宏观态与微观态的对应作出一个假设。这个假设就是等概率原理。这个原理是说，一个孤立系统所有可能的微观态出现的概率都是相等的。等概率原理是统计力学中唯一的原理。这个原理正确与否只有实验能够判决。迄今为止的所有科学事实证明，这个原理是非常合理的。

我们对等概率原理并不陌生。例如，当你说抛掷一个色子得到一点的概率是 $1/6$ 时你就运用了等概率原理。

等概率原理如何运用到统计力学和热力学中呢？我们知道，一个系统在一定的宏观约束下具有多个可能的宏观态，每个宏观态对应一定的微观状态数。既然每一个微观状态——不管它属于哪一个宏观态——出现的概率都相等，那么那个具有最大微观状态数的宏观态出现的概率自然是最大了。这个具有最大微观状态数的宏观态叫做最可几态。所以，随着时间的推移，系统将演化到最可几态。这个最可几态自然就是平衡态。

根据上面的等概率原理，当复合系统达到热力学平衡时，微观状态数 $\Omega(N_1, V_1, E_1, N_2, V_2, E_2)$ 应该具有最大值。因为这个问题中唯一可以变化的是 E_1 （或者 E_2 ），这个最大值发生的条件是：

$$\frac{\partial}{\partial E_1} \Omega(N_1, V_1, E_1, N_2, V_2, E_2) = \frac{\partial}{\partial E_1} [\Omega_1(N_1, V_1, E_1)\Omega_2(N_2, V_2, E_2)] = 0. \quad (147)$$

在我们的问题中， N_1 、 N_2 、 V_1 、 V_2 都是常数，故可将上式简写为

$$\frac{\partial}{\partial E_1} [\Omega_1(E_1)\Omega_2(E_2)] = 0. \quad (148)$$

注意到 $E_1 + E_2 = E$ ，我们有

$$\left(\frac{\partial \Omega_1(E_1)}{\partial E_1} \right) \Omega_2(E_2) - \Omega_1(E_1) \left(\frac{\partial \Omega_2(E_2)}{\partial E_2} \right) = 0, \quad (149)$$

即

$$\frac{1}{\Omega_1(E_1)} \left(\frac{\partial \Omega_1(E_1)}{\partial E_1} \right) = \frac{1}{\Omega_2(E_2)} \left(\frac{\partial \Omega_2(E_2)}{\partial E_2} \right), \quad (150)$$

亦即

$$\left(\frac{\partial \ln[\Omega_1(E_1)]}{\partial E_1} \right)_{N_1 V_1} = \left(\frac{\partial \ln[\Omega_2(E_2)]}{\partial E_2} \right)_{N_2 V_2}. \quad (151)$$

注意：我们在两边的求导操作中加上了之前省略不写的条件。

上式就是两个子系统达到热力学平衡的条件。因为子系统之间不能传粒子也不能做功，这个条件具体地说就是热平衡的条件。这不就是热力学第零定律吗？因此，我们从统计力学推导出了热力学第零定律：两个子系统达到热平衡时必有某个量相等，而这个量就是温度。

热平衡时子系统之间的温度相等，即 $T_1 = T_2$ ，或者 $1/T_1 = 1/T_2$ 。运用热力学基本方程，我有

$$\left(\frac{\partial S_1}{\partial E_1} \right)_{N_1 V_1} = \left(\frac{\partial S_2}{\partial E_2} \right)_{N_2 V_2}. \quad (152)$$

比较以上两式，完全可以作出如下猜测：

$$S \propto \ln[\Omega(E)]. \quad (153)$$

也就是说，一个系统的某个宏观态的熵正比于该宏观态的微观状态数的对数。这里，我们去掉了表示系统的下标，因为这个公式对任何系统都是成立的。这就是熵的微观解释，是玻尔兹曼 (Boltzmann) 最先得到这个关系的。

如何确定上式中的比例常数呢？我们注意到玻尔兹曼常数 k_B 与熵具有相同的量纲，而微观状态数的对数的量纲是 1，故可猜测这个比例常数就是 k_B ：

$$\boxed{S = k_B \ln[\Omega(E)]}. \quad (154)$$

历史上，是普朗克 (Planck) 最先写出这个等式的。这正是刻在玻尔兹曼墓碑上的公式。

下面，我们以理想气体为例来验证这个比例常数确实就是玻尔兹曼常数。首先，可以确定，理想气体的微观状态数与气体的体积有如下关系：

$$\Omega(N, V, E) \propto V^N. \quad (155)$$

再由热力学关系

$$\frac{p}{T} = \left(\frac{\partial S}{\partial V} \right)_{NE}$$

可得

$$pV = Nk_B T. \quad (156)$$

这正是理想气体状态方程。于是可以确定， k_B 就是以前在热力学中引入的玻尔兹曼常数。

最后，我们注意到，熵的微观解释和等概率原理也包含了热力学第二和第三定律。根据等概率原理，一个孤立系统总是向微观状态数最大的宏观态演化。这就是用来表述热力学第二定律的熵增加原理。根据熵的公式 $S = k_B \ln \Omega$ ，熵有一个绝对的最小值 0，当且仅当 $\Omega = 1$ 时取得。这为热力学第三定律提供了一个理论基础。

4.2 正则系综理论

上一讲研究的具有固定的粒子数、体积和能量的孤立系统对应于微正则系综中的一个系统。一个系综就是由若干个假想的与所研究的系统具有相同的宏观态的系统的集合。微正则系综在很多情况下都不方便使用。本讲研究一个更有用的系综：正则系综。

正则系综考虑的是 M 个相同的具有一定粒子数、体积和温度的热力学系统。可以想象 M 个相同的系统排成一个圈；相邻的系统之间有微弱的相互作用，使得所有的系统最终能处于同一温度。用 M_i 表示在微观状态 i 上的系统数目， E_i 表示第 i 个态的能量（需要假设能级不简并吗？我不是很清楚）。则总系统数和总能量为

$$M = \sum M_i. \quad (157)$$

$$E_M = \sum M_i E_i. \quad (158)$$

知道了分布 $\{M_i\}$ ，并没有确定各个系统的状态。可以证明：对于一个给定的分布 $\{M_i\}$ ，系综的微观状态数为

$$\Omega = \frac{M!}{\prod_i M_i!}. \quad (159)$$

从而，这个系综（作为一个很大的孤立系统）的熵为

$$S_M = k_B \ln \Omega = k_B \ln \left(\frac{M!}{\prod_i M_i!} \right). \quad (160)$$

下面我们要问：哪个分布 $\{M_i\}$ 的概率最大？那个具有最大概率的分布就对应于平衡态。根据等概率原理，肯定是微观状态数最大的分布概率最大。因为微观状态数往往是巨大的，所以一般不直接求 Ω 的最大值，而是求 $\ln \Omega$ 的最大值。

对于一个多元函数 $\Omega(M_1, M_2, \dots)$ ，要使其取最大值，自然的想法是令其对所有变量的导数都等于零：

$$\frac{\partial \ln \Omega}{\partial M_i} = 0. \quad (161)$$

然而，这个式子是错的，因为我们忽略了上面的约束条件 $M = \sum M_i$ 和 $E_M = \sum M_i E_i$ 。有约束条件的极值问题一般用拉格朗日乘子法来解决。引入两个拉格朗日乘子 α 和 β 。极

值条件可以写成

$$\frac{\partial \ln \Omega}{\partial M_i} - \alpha \frac{\partial \sum_j M_j}{\partial M_i} - \beta \frac{\partial \sum_j M_j E_j}{\partial M_i} = 0. \quad (162)$$

在 M 趋近于无穷大时，所有的 M_i 也趋近于无穷大。利用斯特林 (Stirling) 公式，有

$$\ln \Omega \approx M(\ln M - 1) - \sum_i [M_i(\ln M_i - 1)]. \quad (163)$$

因为 $\frac{\partial M}{\partial M_i} = 1$ 、 $\frac{\partial E_M}{\partial M_i} = E_i$ 、 $\frac{\partial \ln \Omega}{\partial M_i} = -\ln M_i$ ，于是我们可以将极值条件写成

$$\ln M_i = -\alpha - \beta E_i. \quad (164)$$

对上式两边取指数，可得

$$M_i = \exp[-\alpha - \beta E_i]. \quad (165)$$

既然 M_i 是处于微观 i 的系统的个数，那么自然地，一个系统处于微观态 i 的概率为

$$w_i = \frac{M_i}{M} = \frac{\exp[-\beta E_i]}{\sum_i \exp[-\beta E_i]}. \quad (166)$$

可见，常数 α 没有什么物理意义，在求概率的时候就被消掉了。上式中的分母叫做正则配分函数，记为 Z ：

$$Z = \sum_i \exp[-\beta E_i]. \quad (167)$$

表面上看，配分函数只不过是一个归一化因子而已。然而实际上，配分函数包含了体系所有的热力学性质。后面马上会验证此论断。

那么，常数 β 有什么物理意义呢？一方面，我们知道，正则系综中的温度是一个常数，所以可以猜测 β 与温度有关。另一方面，从量纲的角度来看，可以进一步猜测：

$$\beta = \frac{1}{k_B T}. \quad (168)$$

于是，正则配分函数可以写成

$$Z = \sum_i \exp \left[-\frac{E_i}{k_B T} \right]. \quad (169)$$

怎么确定这里引入的 T 就是绝对温度呢？等到讲完后面的位力定理时我们就会明白。在那之前，先假设 T 就是绝对温度，并考察几个热力学量的计算。

可以证明，整个正则系综的熵可以化为如下形式：

$$S_M = -k_B M \sum_i w_i \ln w_i. \quad (170)$$

于是，由熵的可加性得到单个系统的熵为

$$S = -k_B \sum_i w_i \ln w_i. \quad (171)$$

这个公式叫做熵的吉布斯公式。

将概率函数 w_i 的表达式代入熵的吉布斯公式可得

$$S = k_B \ln Z + \frac{E}{T}. \quad (172)$$

其中，

$$E = \sum_i w_i E_i \quad (173)$$

是系统能量的平均值。于是，

$$E - TS = -k_B T \ln Z. \quad (174)$$

上式右边就是亥姆霍兹自由能：

$$\boxed{F = -k_B T \ln Z}. \quad (175)$$

这样就将正则配分函数与系统的亥姆霍兹自由能联系起来了。这个联系的意义是深远的，因为我们知道，在粒子数、体积、温度固定的系统中，亥姆霍兹自由能函数包含了系统所有的热力学性质。例如，正则系综中系统的能量和压强的平均值可以表示为

$$E = -\frac{\partial}{\partial \beta} \ln Z \quad (176)$$

$$p = \frac{1}{\beta} \frac{\partial}{\partial V} \ln Z \quad (177)$$

当然，如果系统的粒子数不是无穷大的话，计算出的热力学量应该有涨落，即标准偏差不等于零。正则系综中能量平方的平均值

$$\langle E^2 \rangle = \sum_i w_i E_i^2 \quad (178)$$

可以写成

$$\langle E^2 \rangle = \langle E \rangle^2 - \frac{\partial \langle E \rangle}{\partial \beta}. \quad (179)$$

其中，为了明确起见，我们把之前定义的平均能量 E 写成了 $\langle E \rangle$ 。进而可求出能量的方差

$$(\Delta E)^2 = \langle E^2 \rangle - \langle E \rangle^2 = -\frac{\partial \langle E \rangle}{\partial \beta} = \frac{1}{k_B \beta^2} \frac{\partial \langle E \rangle}{\partial T} = k_B T^2 C_V. \quad (180)$$

其中, C_V 是系统的等容热容。这就证明了等容热容一定是非负的。因为能量和等容热容都是广延量, 能量的相对偏差在热力学极限 (保持粒子数密度不变时让粒子数趋近于无穷大) 下趋近于零:

$$\frac{\Delta E}{\langle E \rangle} = \frac{\sqrt{k_B T^2 C_V}}{\langle E \rangle} \rightarrow \frac{1}{\sqrt{N}}. \quad (181)$$

这个结论的数学根源就是中心极限定理的结论: 即不管每一个粒子的能量如何分布, 在粒子数趋近于无穷大时, 系统总能量的相对偏差一定趋近于零。

4.3 经典概率密度函数

在得到正则配分函数之后, 让我们从离散能量的情形回到连续能量的情形。为此, 只要将概率函数 w_i 换成概率密度函数 $\rho(H(q, p))$ 即可:

$$\rho(H(q, p)) = \frac{e^{-\beta H(q, p)}}{\int dq dp e^{-\beta H(q, p)}}. \quad (182)$$

这个概率密度函数显然是归一化的。这里的分母

$$Z = \int dq dp e^{-\beta H(q, p)} \quad (183)$$

就是连续情形中的配分函数。这个配分函数的定义不是最终的正确形式 (一个明显的问题就是 Z 的量纲不为 1), 但在遇到问题之前, 我们不知道该对它做怎样的修正。其实, 在很多的问題中, 我们用不到配分函数, 不妨先简单地将它看做一个归一化常数。

对经典粒子系统, 总能量 $H(q, p)$ 是动能 $K(p)$ 和势能 $U(q)$ 的和:

$$H(q, p) = K(p) + U(q). \quad (184)$$

于是, 系统处于 $dq dp$ 的概率 $\rho(H(q, p)) dq dp$ 可以分解动量部分 $f_p(p) dp$ 和坐标部 $f_q(q) dq$ 的乘积:

$$\frac{e^{-\beta H(q, p)}}{Z} dq dp = [A e^{-\beta K(p)} dp] [B e^{-\beta U(q)} dq] \equiv [f_p(p) dp] [f_q(q) dq]. \quad (185)$$

其中, 常数 A 和 B 分别是动量和坐标部分的归一化因子, 满足 $AB = 1/Z$ 。这就是说, 我们可以将动量分布函数与坐标分布函数分开来研究。这其实就是关于独立随机变量的乘法原理的体现。

对一般的系统, 如果我们不知道体系的势能函数, 是无法研究坐标分布函数的。然而, 无论哪种系统 (现在只讨论非相对论的经典系统; 该论断在量子统计中不正确), 其动量分布函数都是一样的。因为任何系统的动能都可以写成单个粒子的动能的和, 所以我们可以进一步将整个系统的动量分布函数 $f_p(p) dp$ (这里的 p 指代所有 $3N$ 个动量分量)

分解为单个粒子的动量分布函数 $f_{\vec{p}}(\vec{p})$ (这里的 \vec{p} 指代某个原子的动量矢量) 的乘积。其中, $f_{\vec{p}}(\vec{p})$ 为

$$f_{\vec{p}}(\vec{p})dp_xdp_ydp_z = ae^{-\beta\frac{p_x^2+p_y^2+p_z^2}{2m}}dp_xdp_ydp_z = ae^{-\frac{p_x^2+p_y^2+p_z^2}{2mk_BT}}dp_xdp_ydp_z. \quad (186)$$

注意, 此处的归一化因子 a 与前面的归一化因 A 的关系为 $A = a^N$ 。

容易计算, 上述单个粒子的动量分布函数的归一化因子为

$$a = \frac{1}{(2\pi mk_BT)^{3/2}}. \quad (187)$$

于是, 我们可以将单粒子动量分布函数完整地写出来:

$$f_{\vec{p}}(\vec{p})dp_xdp_ydp_z = \frac{1}{(2\pi mk_BT)^{3/2}}e^{-\frac{p_x^2+p_y^2+p_z^2}{2mk_BT}}dp_xdp_ydp_z. \quad (188)$$

如果将随机变量从动量换成速度 \vec{v} , 我们可以立刻写出速度分布函数:

$$f_{\vec{v}}(\vec{v})dv_xdv_ydv_z = \left(\frac{m}{2\pi k_BT}\right)^{3/2} e^{-\frac{m(v_x^2+v_y^2+v_z^2)}{2k_BT}}dv_xdv_ydv_z. \quad (189)$$

这就是麦克斯韦于 1860 年通过分子运动论推导出的速度分布函数。

我们可以继续将各个方向的速度分量的分布函数分离出来。例如, x 方向的速度分布函数为

$$f_{v_x}(v_x)dv_x = \left(\frac{m}{2\pi k_BT}\right)^{1/2} e^{-\frac{mv_x^2}{2k_BT}}dv_x. \quad (190)$$

如果从速度的直角坐标换到速度的球坐标 (v 、 θ 、 ϕ), 则有

$$f_{\vec{v}}(\vec{v})dv_xdv_ydv_z = \left(\frac{m}{2\pi k_BT}\right)^{3/2} e^{-\frac{mv^2}{2k_BT}}v^2dv \sin\theta d\theta d\phi. \quad (191)$$

若定义 $f_v(v)dv$ 为速率间隔 dv 内的概率, 则有

$$\begin{aligned} f_v(v)dv &= \left(\frac{m}{2\pi k_BT}\right)^{3/2} e^{-\frac{mv^2}{2k_BT}}v^2dv \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi \\ &= 4\pi \left(\frac{m}{2\pi k_BT}\right)^{3/2} e^{-\frac{mv^2}{2k_BT}}v^2dv. \end{aligned} \quad (192)$$

利 $f_{v_x}(v_x)$ 的表达式, 可以证明:

$$\langle v_x^2 \rangle = \frac{k_BT}{m} \quad (193)$$

利用 $f_v(v)$ 的表达式, 可以证明:

$$\langle v^2 \rangle = \frac{3k_BT}{m} \quad (194)$$

$$\langle v \rangle^2 = \frac{8k_B T}{\pi m} \quad (195)$$

由上述结果可知，每一个平动自由度的平均能量为 $k_B T/2$ 。下面我们马上会知道，这是能量均分定理的体现。

考虑一个经典的多粒子系统，我们用 q 代表系统的 s 个广义坐标，用 p 代表 s 个广义动量，而 x_i 或者 x_j 代表任意一个广义坐标或者广义动量。我们现在利用上述概率密度函数来计算经典正则系综中的一个平均值

$$\left\langle x_i \frac{\partial H}{\partial x_j} \right\rangle = \frac{\int dq dp e^{-\beta H(q,p)} x_i \frac{\partial H}{\partial x_j}}{\int dq dp e^{-\beta H(q,p)}}. \quad (196)$$

为了计算这个平均值，让我们先关注分母。注意到分母中的乘积 $e^{-\beta H(q,p)} \frac{\partial H}{\partial x_j}$ 可以写成 $-\frac{1}{\beta} \frac{\partial e^{-\beta H(q,p)}}{\partial x_j}$ 。于是，分母中的被积函数可以写成

$$-\frac{1}{\beta} x_i \frac{\partial e^{-\beta H(q,p)}}{\partial x_j} = -\frac{1}{\beta} \frac{\partial}{\partial x_j} (x_i e^{-\beta H(q,p)}) + \frac{1}{\beta} \delta_{ij} e^{-\beta H(q,p)}. \quad (197)$$

上式等号右边的第一项的积分等于

$$-\frac{1}{\beta} \int \frac{dq dp}{dx_j} \int dx_j \frac{\partial}{\partial x_j} (x_i e^{-\beta H(q,p)}) = -\frac{1}{\beta} \int \frac{dq dp}{dx_j} (x_i e^{-\beta H(q,p)})_{x_j^{min}}^{x_j^{max}} = 0 \quad (198)$$

上式等于零的理由如下。当 x_j 等于某个广义坐标时， x_j^{min} 和 x_j^{max} 一定对应于系统的边界处（容器壁），那里的势能无穷大，使得因子 $x_i e^{-\beta H(q,p)}$ 等于零。当 x_j 等于某个广义动量时， x_j^{min} 和 x_j^{max} 就分别等 $-\infty$ 和 $+\infty$ ，使得系统的动能为无穷大，从而依然使得因子 $x_i e^{-\beta H(q,p)}$ 等于零。于是，我们要求的平均值为

$$\left\langle x_i \frac{\partial H}{\partial x_j} \right\rangle = \frac{\int dq dp \frac{1}{\beta} \delta_{ij} e^{-\beta H(q,p)}}{\int dq dp e^{-\beta H(q,p)}} = \frac{1}{\beta} \delta_{ij} = k_B T \delta_{ij}. \quad (199)$$

现在，假设系统的哈密顿量是某个广义坐标或者广义动量的二次函数，即假设由自由度 x_i 贡献的哈密顿量为

$$H_i(x_i) = a x_i^2. \quad (200)$$

其中， a 是常数系数。对于这样的哈密顿量，显然有

$$H_i(x_i) = \frac{1}{2} x_i \frac{\partial H_i(x_i)}{\partial x_i}. \quad (201)$$

于是，与自由度 x_i 相关的能量平均值为

$$\langle H_i(x_i) \rangle = \frac{1}{2} k_B T. \quad (202)$$

这就是能量均分定理，即哈密顿量的每个具有平方形式的自由度对总能量的平均贡献都是 $\frac{1}{2}k_B T$ 。

将能量均分定理应用与单原子理想气体，因为每个原子只有三个平动自由度，故可得内能

$$E = \frac{3}{2} N k_B T. \quad (203)$$

于是，单原子理想气体的等容热容为 $C_V = \frac{3}{2} N k_B$ 。该预言与实验结果是符合得很好的。再考虑双原子分子的理想气体。因为每个分子有三个平动动能项，两个转动动能项，一个振动动能项与一个振动势能项，故由能量均分定理，内能应该为

$$E = \frac{7}{2} N k_B T. \quad (204)$$

相应的等容热容为 $C_V = \frac{7}{2} N k_B$ 。然而，实验结果显示，室温下大部分的双原子气体的等容热容都接近于 $C_V = \frac{5}{2} N k_B$ 。这是经典统计力学遭遇的困难之一。经典统计力学在固体比热和黑体辐射的问题上也遭遇了巨大的困难，但这都需要由量子统计解决。

很容易写出上面所定义的平均值的一个 $x_i = x_j = q_\alpha$ 的特例：

$$\left\langle q_\alpha \frac{\partial H}{\partial q_\alpha} \right\rangle = k_B T. \quad (205)$$

考虑由 N 个粒子组成的经典系统，将上式对 α 求和可得

$$\sum_{\alpha=1}^{3N} \left\langle q_\alpha \frac{\partial H}{\partial q_\alpha} \right\rangle = 3N k_B T. \quad (206)$$

因为 $-\frac{\partial H}{\partial q_\alpha} = F_\alpha$ 是与坐标 q_α 对应的力，故有

$$\sum_{\alpha=1}^{3N} \langle q_\alpha F_\alpha \rangle = -3N k_B T. \quad (207)$$

克劳修斯于 1870 年为上式左边的量起了一个名字：Virial。常见的中文翻译有两个：一个是位力，一个是维里。我们将用第一个，因为这个量的物理意义就是位置乘以力的和的平均值，翻译成位力是非常棒的。我们将用符号 W 表示这个量，即

$$W = \sum_{\alpha=1}^{3N} \langle q_\alpha F_\alpha \rangle = -3N k_B T. \quad (208)$$

这个公式叫做位力定理。

现在将位力定理应用于经典理想气体，即无相互作用的经典多粒子系统。因为粒子之间没有相互作用，故所有的力来自于粒子与容器的碰撞。取一个坐标系，记容器壁 S 上的

任意一点的坐标为 \vec{x} 。假设系统的压强为 p ，则容器壁上位于 \vec{x} 处的面积元 $d\vec{a}$ （朝外的方向为正方向）施加给系统的力为 $d\vec{F} = -pd\vec{a}$ 。于是，系统的位力可表示为

$$W = \oint_S \vec{x} \cdot d\vec{F} = \oint_S \vec{x} \cdot (-pd\vec{a}) = -p \oint_S \vec{x} \cdot d\vec{a}. \quad (209)$$

记容器所在区域为 Ω ，包围的体积为 V ，并应用高斯定理，可得

$$W = -p \int_{\Omega} (\nabla \cdot \vec{x}) dv = -3pV. \quad (210)$$

将上式与位力定理对比可知：

$$pV = Nk_B T. \quad (211)$$

这就是理想气体的状态方程。这就证明了之前利用 $\beta = \frac{1}{k_B T}$ 所定义的 T 就是绝对温度。以后我们就可以放心地将 $\beta = \frac{1}{k_B T}$ 中的 T 当做绝对温度了。

虽然之前用位力定理推导了理想气体状态方程，但我们并没有系统计算单原子分子理想气体的各个热力学量。现在是时候考虑这些问题了。首先，我们明显地写出 N 个无相互作用的原子组成的理想气体系统的能量函数：

$$H(q, p) = \sum_{i=1}^N \frac{\vec{p}_i^2}{2m} \quad (212)$$

对这样的系统，其配分函数是

$$Z = \int \exp \left[- \sum_{i=1}^N \frac{\vec{p}_i^2}{2mk_B T} \right] \prod_{i=1}^N d\vec{x}_i d\vec{p}_i \quad (213)$$

然而，正如之前就指出过的，这个配分函数的量纲都不对。要使配分函数的量纲等于 1，我们必须将上式除以一个量纲为 $([\text{长度}] \times [\text{动量}])^{3N}$ 的量。这样做其实就是定义一个量纲为 $[\text{长度}] \times [\text{动量}]$ 的“最小”的相空间体积 ω_0 ，使得

$$\int \frac{\prod_{i=1}^N d\vec{x}_i d\vec{p}_i}{\omega_0^{3N}} \quad (214)$$

等于系统中总的“相点个数”。我们期望这个最小相空间体积 ω_0 应该是一个小量。

普朗克曾引入一个常数 h ，它的量纲与 ω_0 的量纲相同。所以， ω_0 一定正比于 h 。那个比例常数是多少呢？在将由此导出的结果与其它理论结果或者实验结果对比之前，我们是无法知道答案的。暂且就假设 $\omega_0 = h$ 吧。于是，配分函数变为

$$Z = \frac{1}{h^{3N}} \int \exp \left[- \sum_{i=1}^N \frac{\vec{p}_i^2}{2mk_B T} \right] \prod_{i=1}^N d\vec{x}_i d\vec{p}_i \quad (215)$$

容易证明，上述配分函数可以写成

$$Z = Z_1^N. \quad (216)$$

$$Z_1 = \frac{V}{\lambda^3}. \quad (217)$$

$$\lambda = \frac{h}{\sqrt{2\pi m k_B T}}. \quad (218)$$

于是，体系的自由能为

$$F = -k_B T \ln Z = -N k_B T \ln \left(\frac{V}{\lambda^3} \right). \quad (219)$$

通过上述自由能，可以计算体系的压强，从而导出理想气体状态方程：

$$pV = N k_B T. \quad (220)$$

正如所期望的，我们推导出了理想气体应该满足的状态方程。

下面再计算熵。体系的熵为：

$$S = N k_B \ln \left(\frac{V}{\lambda^3} \right) + \frac{3}{2} N k_B \quad (221)$$

虽然上述熵和自由能以及内能满足关系 $F = E - TS$ ，但值得注意的是熵和自由能都不是广延量。这是不可接受的。这说明我们的理论还是有不完美的地方。这个问题能由所谓的吉布斯佯谬更生动地展现出来。

考虑由两个温度和数密度都相同的理想气体系统构成的孤立系统，粒子数分别为 N_1 和 N_2 。可以证明，两个系统混合后与混合前总熵的差为

$$\Delta S = k_B (N \ln N - N_1 \ln N_1 - N_2 \ln N_2) \approx k_B (\ln N! - \ln N_1! - \ln N_2!). \quad (222)$$

如果两个子系统的气体是相同种类的气体，这个结果是很荒谬的。这就是吉布斯佯谬。同时，上述结果暗示我们，如果重新定义配分函数，使得熵的值为原来的值减去 $k_B \ln N!$ ，也许就能消除这个佯谬。根据熵与配分函数的关系可以猜测，应该重新定义如下的配分函数

$$Z = \frac{1}{h^{3N} N!} \int \exp \left[- \sum_{i=1}^N \frac{\vec{p}_i^2}{2m k_B T} \right] \prod_{i=1}^N d\vec{x}_i d\vec{p}_i \quad (223)$$

这样定义的结果就是将系统中总的相点（状态数）个数减小 $N!$ 倍。重复之前的推导可得

$$Z = \frac{Z_1^N}{N!}. \quad (224)$$

从这个新的配分函数出发，可以证明理想气体的自由能和熵的正确表达式应该是

$$F = -k_B T \ln Z = -Nk_B T \left[\ln \left(\frac{v}{\lambda^3} \right) + 1 \right]. \quad (225)$$

$$S = Nk_B \ln \left(\frac{v}{\lambda^3} \right) + \frac{5}{2} Nk_B \quad (226)$$

其中，

$$v = V/N. \quad (227)$$

上述熵的公式叫做 Sackur-Tetrode 公式。正是 Sackur 在 1911 年首次假设 $\omega_0 = h$ ，并由 Tetrode 于 1912 年最终通过将理论与实验比较确定该表达式的。

在对配分函数作出修正之后，上面计算出的自由能和熵都是广延量了。而且可以验证，吉布斯佯谬不复存在。这说明这个修正是合理的。那么，这个修正究竟代表什么意思呢？答案是：它反映了全同粒子的不可分辨性。这是量子力学中的一个结果。由于全同粒子之间不可分辨，对体系的所有粒子做一个重排并不改变系统的微观状态，故需要将体系的微观状态数在原来（没有考虑粒子的不可分辨性时）的基础上除以所有可能的重排数目，即 $N!$ 。

最后，容易计算，理想气体的化学势为

$$\mu = -k_B T \ln \left(\frac{v}{\lambda^3} \right). \quad (228)$$

可见它是负的，与我们在热力学中得到结果一致。

4.4 巨正则系综理论

巨正则系综考虑的是 M 个相同的具有一定化学势、体积和温度的热力学系统。可以想象 M 个相同的系统排成一排，相邻的系统之间有微弱的相互作用，使得所有的系统最终能处于同一温度。相邻的系统之间还可以交换粒子，从而每个系统的粒子数是不固定的，但根据热力学理论我们知道平衡时每个系统的化学势会相等。我们的目的是发展一套能够描述这个系综中的某一个系统的热力学性质的统计理论。大家会看到，本讲的内容与《正则系综》那一讲的内容是类似的。

用 M_{Ni} 表示粒子数 N 且状态为 i 的系统数目，而用 E_{Ni} 表示对应的能量，则整个系综的总系统数、总粒子数和总能量为

$$M = \sum_N \sum_{i(N)} M_{Ni}. \quad (229)$$

$$N_M = \sum_N \sum_{i(N)} M_{Ni} N. \quad (230)$$

$$E_M = \sum_N \sum_{i(N)} M_{Ni} E_{Ni}. \quad (231)$$

对于一个给定的分布 $\{M_{Ni}\}$ ，系综的微观状态数为

$$\Omega = \frac{M!}{\prod_N \prod_{i(N)} M_{Ni}!}. \quad (232)$$

从而，这个系综（作为一个很大的孤立系统）的熵为

$$S_M = k_B \ln \Omega = k_B \ln \left(\frac{M!}{\prod_N \prod_{i(N)} M_{Ni}!} \right). \quad (233)$$

下面我们要问：哪个分布 $\{M_{Ni}\}$ 的概率最大？那个具有最大概率的分布就对应于平衡态。根据等概率原理，肯定是微观状态数最大的分布概率最大。类比正则系综的讨论，我们知道，我们需要求 $\ln \Omega$ 的最大值。方法是引入三个拉格朗日乘子 α 、 β 和 γ ，并将极值条件写成

$$\frac{\partial \ln \Omega}{\partial M_{Ni}} - \alpha \frac{\partial \sum_N \sum_{j(N)} M_{Nj}}{\partial M_{Ni}} - \beta \frac{\partial \sum_N \sum_{j(N)} M_{Nj} E_{Nj}}{\partial M_{Ni}} - \gamma \frac{\partial \sum_N \sum_{j(N)} M_{Nj} N}{\partial M_{Ni}} = 0. \quad (234)$$

在 M 趋近于无穷大时，所有的 M_{Ni} 也趋近于无穷大。利用斯特林公式，可将极值条件化为

$$\ln M_{Ni} = -\alpha - \beta E_{Ni} - \gamma N. \quad (235)$$

对上式两边取指数，可得

$$M_{Ni} = \exp[-\alpha - \beta E_{Ni} - \gamma N]. \quad (236)$$

既然 M_{Ni} 是处于态 Ni 的系统的个数，那么自然地，一个系统处于态 Ni 的概率为

$$w_{Ni} = \frac{M_{Ni}}{M} = \frac{\exp[-\beta E_{Ni} - \gamma N]}{\sum_N \sum_{i(N)} \exp[-\beta E_{Ni} - \gamma N]}. \quad (237)$$

可见，常数 α 没有什么物理意义，在求概率的时候就被消掉了。上式中的分母叫做巨正则配分函数，记为 Ξ ：

$$\Xi = \sum_N \sum_{i(N)} \exp[-\beta E_{Ni} - \gamma N]. \quad (238)$$

巨配分函数包含了体系所有的热力学性质。

根据我们对正则系综的讨论，我们可以继续认为

$$\beta = \frac{1}{k_B T}. \quad (239)$$

下面的问题是确定系数 γ 的物理意义。突破口还是熵和其它热力学函数。将概率函数的表达式代入熵的吉布斯公式可得

$$S = k_B \ln \Xi + \frac{E}{T} + k_B \gamma N. \quad (240)$$

其中，

$$E = \sum_N \sum_{i(N)} w_{Ni} E_{Ni} \quad (241)$$

是系统能量的平均值，

$$N = \sum_N \sum_{i(N)} w_{Ni} N \quad (242)$$

是系统粒子数的平均值。于是，

$$E - TS + k_B T \gamma N = -k_B T \ln \Xi. \quad (243)$$

上式右边等同于巨热力学势

$$\boxed{\Phi = -k_B T \ln \Xi}, \quad (244)$$

所以， γ 与化学式有如下联系：

$$\boxed{\gamma = -\frac{\mu}{k_B T} = -\beta \mu}. \quad (245)$$

于是，巨正则配分函数可以写成

$$\boxed{\Xi = \sum_N \sum_{i(N)} \exp \left[-\frac{E_{Ni}}{k_B T} + \frac{\mu N}{k_B T} \right]}. \quad (246)$$

容易证明，巨正则系综中系统的粒子数、能量和压强的平均值可以表示为

$$\langle N \rangle = -\frac{\partial}{\partial \gamma} \ln \Xi \quad (247)$$

$$\langle E \rangle = -\frac{\partial}{\partial \beta} \ln \Xi \quad (248)$$

$$\langle p \rangle = \frac{1}{\beta} \frac{\partial}{\partial V} \ln \Xi \quad (249)$$

如果系统的粒子数不是无穷大的话，计算出的热力学量应该有涨落，即标准偏差不等于零。可以证明：巨正则系综中能量的方差为

$$(\Delta E)^2 = \langle E^2 \rangle - \langle E \rangle^2 = -\frac{\partial \langle E \rangle}{\partial \beta}, \quad (250)$$

粒子数的方差为

$$(\Delta N)^2 = \langle N^2 \rangle - \langle N \rangle^2 = -\frac{\partial \langle N \rangle}{\gamma}. \quad (251)$$

于是，可以判断，在热力学极限下（即保持粒子数密度不变的条件下将粒子数增加到无穷大），能量和粒子数的相对偏差都趋近于零：

$$\frac{\Delta E}{\langle E \rangle} \rightarrow \frac{1}{\sqrt{\langle N \rangle}}. \quad (252)$$

$$\frac{\Delta N}{\langle N \rangle} \rightarrow \frac{1}{\sqrt{\langle N \rangle}}, \quad (253)$$