

An Introduction to

STATISTICAL PHYSICS

BY HAOTING

THE PUBLISHER YEAR

目录

1	统计物理基础	1
1.1	统计物理简介	1
1.2	微正则系综	2
1.2.1	等概率原理	2
1.2.2	等概率原理解释和合理性	2
1.2.3	微正则系综	3
1.2.4	熵增加原理	4
1.2.5	温度和热容	6
1.2.6	压强	7
1.2.7	热力学第一定律	7
1.2.8	一个例子——两态系统	8
1.3	正则系综	10
1.3.1	配分函数 (Partition Function)	11
1.3.2	密度矩阵的一种选择	12
1.3.3	平均能量和能量涨落	12
1.3.4	题外话：统计物理是理论物理中璀璨的宝石	14
1.3.5	再来计算两能级系统	15
1.3.6	正则系综中的熵	15
1.3.7	自由能	16
1.4	巨正则系综	18
1.4.1	化学势	18
1.4.2	巨正则系综	19
1.4.3	巨正则势	20
1.4.4	强度量和广延量	21
1.5	对于系综理论的思考	22
1.5.1	三种系综理论的总结	22
1.5.2	自己新造一种系综？最巨系综？	23
1.5.3	不同条件下使得熵最大	24

1.6	约西亚·威拉德·吉布斯 (Josiah Willard Gibbs)	25
2	经典气体	26
2.1	经典配分函数	26
2.2	理想气体	28
2.2.1	能量均分定理	29
2.2.2	熵和吉布斯悖论	32
2.2.3	化学势	33
2.3	麦克斯韦分布	33
2.4	双原子分子	33
2.5	相互作用的经典气体	35
2.5.1	梅耶 (Mayer) f 函数和第二位力系数	36
2.6	德拜-休克尔等离子体模型	36
3	量子气体	36
4	相变初步	36

Learn more and form your own opinion.

E. Witten

1 统计物理基础

1.1 统计物理简介

统计物理研究的对象是大量的粒子 ($N \simeq 10^{23}$) 的规律, 尽管我们知道了每个粒子的相互作用规律, 但是当大量粒子聚集在一起的时候, 每个粒子的规律就不是很重要。一是因为大量粒子聚集在一起的时候写出方程根本没法求解, 二是因为在宏观尺度上我们根本不关心每个粒子怎样。然而, 我们会问一些十分简单的问题, 比如说这系统的温度是多少 (尽管现在我们还没有定义温度)? 这个系统的压强是多少? 这个系统占有多少体积? 系统的能量如何? 这些问题与每个粒子的基本相互作用是有关的, 事实上, 正是统计物理上的结果与实验不符, 推动了量子力学的发展。如何从微观上每个粒子的作用推导到大量粒子宏观上的性质? 统计物理正是要解答这个问题。

1.2 微正则系综

1.2.1 等概率原理

我们先来考虑孤立系统，对于一个和外界完全孤立的系统，我们假设它有确定的能量 E 。这个系统已经扔在那里孤立了好长时间了，因此它的宏观统计性质不会随着时间的改变而改变，我们说假设系统处于态 $|\psi\rangle$ ，系统满足 $\hat{H}|\psi\rangle = E|\psi\rangle$ 。试图求解 $|\psi\rangle$ 十分困难。而且，一个粒子数巨大的系统一般来说都要受到一些小的扰动，这些小的扰动使得这些态之间来回转变。因此，一个巨大的系统微观上一定处于不断地变化之中。因此根本无法说，系统处于哪个态上面。因此我们采用一种统计的观点，我们研究每个态出现的概率。

那么这些态的概率如何？直觉使得我们猜想，取这些态的可能性相等（注意这些态具有相同的能量）。这就是统计物理学基本假设

假设 1 (等概率原理) 对于处于平衡态的孤立系统，所有可达到的态具有相同的可能性。

1.2.2 等概率原理解释和合理性

1. “概率”具体是什么意思？为什么我们的理论叫做系综理论？

我们这里说系统取一个状态的“概率”，可以这样操作：取 W 个系统，然后每个系统都看看它处于什么态，我们说当 $W \rightarrow \infty$ 时，处于 $|n\rangle$ 这个态的系统数目为 $p(n)W$ 。我们这里采取了很多个系统，所以这是一个系统所综合而成的系统的理论，故取名曰：系综理论。¹

2. 等概率原理在直觉上的理解。

等概率原理可作如下直观理解，通常，一个系统的状态是由一系列参数（通常为 N 的量级）所决定的（如图 1），一方面由于我们对于系统的无知（我们只知道系统的能量），另一方面，由于系统内部可能存在各种相互作用，会导致描述系统的参数从一点突变到另一点，这就导致在长时间看来，各个状态是平权的，即可以取的每个状态是等概率的。举一个简单的例子——理想气体，理想气体模型忽略了气体分子的相互作用，而只

¹这纯属个人猜测。

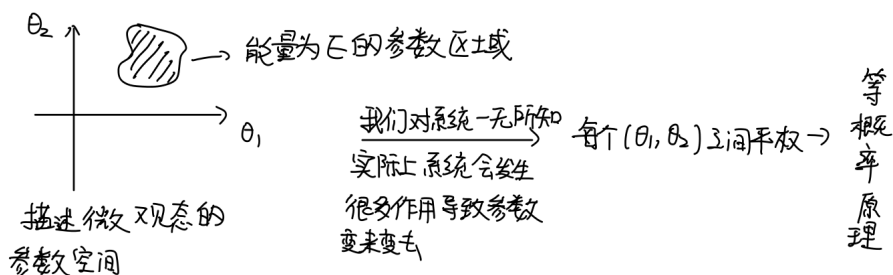


图 1: 等概率原理解释

保留了气体分子的弹性碰撞，即不损失能量。描述这个系统的参数就是每一个分子的位置和动量。而每次碰撞过程都会改变动量，过一段时间粒子也会运动到不同的位置，现在假设有多个相同的系统，这时候你拿 W 个能量都为 E 的系统，抽出一个来，问一问，系统状态取 (x_i, p_i) 的系统概率多大呀？假设比较除了 A 粒子之外其他都相同的系统，那么 A 粒子是往左运动的概率大呢，还是 A 粒子往右运动的概率大呢？直觉告诉我们，一样大。

1.2.3 微正则系综

基本假设中“可达到的”是指那些通过微小扰动可以达到的态。我们关心的只是宏观上的统计物理量，因此微观上的信息被大大缩减。因此，会有很多的微观态对应相同的宏观统计物理量。因此，在给定条件下，我们只要数一数有多少个微观态，我们就知道了系统取具体哪个态的概率，如果用 $\Omega(E)$ 表示微观态数目，那么根据等概率原理，取具体一个微观态的概率为

$$p = \frac{1}{\Omega(E)} \quad (1.1)$$

这个概率分布被叫做微正则系综。

要注意对于我们研究的系统， Ω 是一个大的离谱的数字。即使对于最简单的两能级系统， Ω 也大概为 2^N 个。我们研究 $N \sim 10^{23}$ ，因此它是 $2^{10^{23}}$ ，这个数字就离谱。一般来说，你不会在任何正确答案中看到这样的数字。这个数字无论代表什么、用什么单位计数，都会导致相同的结果。因此，玻尔兹曼老爷爷智慧地想到，定义一个新的物理量来解决这样一个指数上有指数的问题，这个物理量就是熵。

$$S = k_B \log \Omega \quad (1.2)$$

这里 k_B 是我们引入的一个常数, 叫做玻尔兹曼常数, 在国际单位制下为 $k_B = 1.38 \times 10^{-23} \text{J/K}$ 。根据上面的论断 $\Omega \sim e^N$, 因此 $S \sim N$ 。

1.2.4 熵增加原理

有了熵这样一个物理量, 接下来利用简单的论断, 就可以论证熵不会减小。考虑两个相互作用很弱的系统 (以至于对系统的能级无影响), 能量分别为 E_1, E_2 , 假设他们之间可以交换能量。两个系统的态的数目 $\Omega(E_1, E_2) = \Omega(E_1)\Omega(E_2)$ 。两个系统的熵是相加的关系。总系统的微观态数目

$$\Omega(E_{\text{total}}) = \sum_{\{E_i\}} \Omega_1(E_1)\Omega_2(E_2) \quad (1.3)$$

$$= \sum_{\{E_i\}} \exp\left(\frac{S_1(E_i)}{k_B} + \frac{S_2(E_{\text{total}} - E_i)}{k_B}\right) \quad (1.4)$$

其中的 $\sum_{\{E_i\}}$ 是对所有可能的能量求和。实际上, $\Omega(E_{\text{total}})$ 是由最大的 E_i 那一项主导。因为 $S \sim N$, 所以如果 E_i 所贡献的单粒子熵 $\left(\frac{S_1(E_i)}{k_B} + \frac{S_2(E_{\text{total}} - E_i)}{k_B}\right)/N$ 比另一个能量组合大两倍, 那么最后的微观态数目将大 e^{2N} 倍。所以几乎可以说, 微观态数目由使得总熵最大的能量组合主导²。根据等概率原理, 取每一个态的概率为

$$p = \frac{1}{\Omega(E_{\text{total}})}$$

因此具体是哪个能量组合概率最大, 而且比附近的概率明显大很多? 那就要看谁贡献的 $\left(\frac{S_1(E_i)}{k_B} + \frac{S_2(E_{\text{total}} - E_i)}{k_B}\right)$ 多, 贡献最多的是 E_* 所对应的项, 取这样的能量组合的概率最大。 E_* 满足

$$\frac{\partial}{\partial E} \left(\frac{S_1(E_i)}{k_B} + \frac{S_2(E_{\text{total}} - E_i)}{k_B} \right)_{E=E_*} = 0 \quad (1.5)$$

因此, 系统最有可能取 $\{E_*\}$, 即系统最有可能取总熵 $S \simeq S_1(E_i) + S_2(E_{\text{total}} - E_i)$ 最大的态! 这就是熵增加原理! 也被称作热力学第二定律。可以简洁地表述为

$$\delta S \geq 0 \quad (1.6)$$

注意, 当我们说概率大很多的时候, 读者应该清楚这个大很多是什么概念, 这概率大的相当于指数上的指数那么多倍。所以可能达到 E_* 以外的状态是十分不可能的! 有多不可能? 实际上, 它是“永不可能”的一个很好的操作性定义。

人们有时说, 热力学第二定律是物理学中最可怕的定律。如图 2 是爱丁顿对于热力学第二定律一个十分著名的赞美。

²怎么我一写东西就是绕口令?

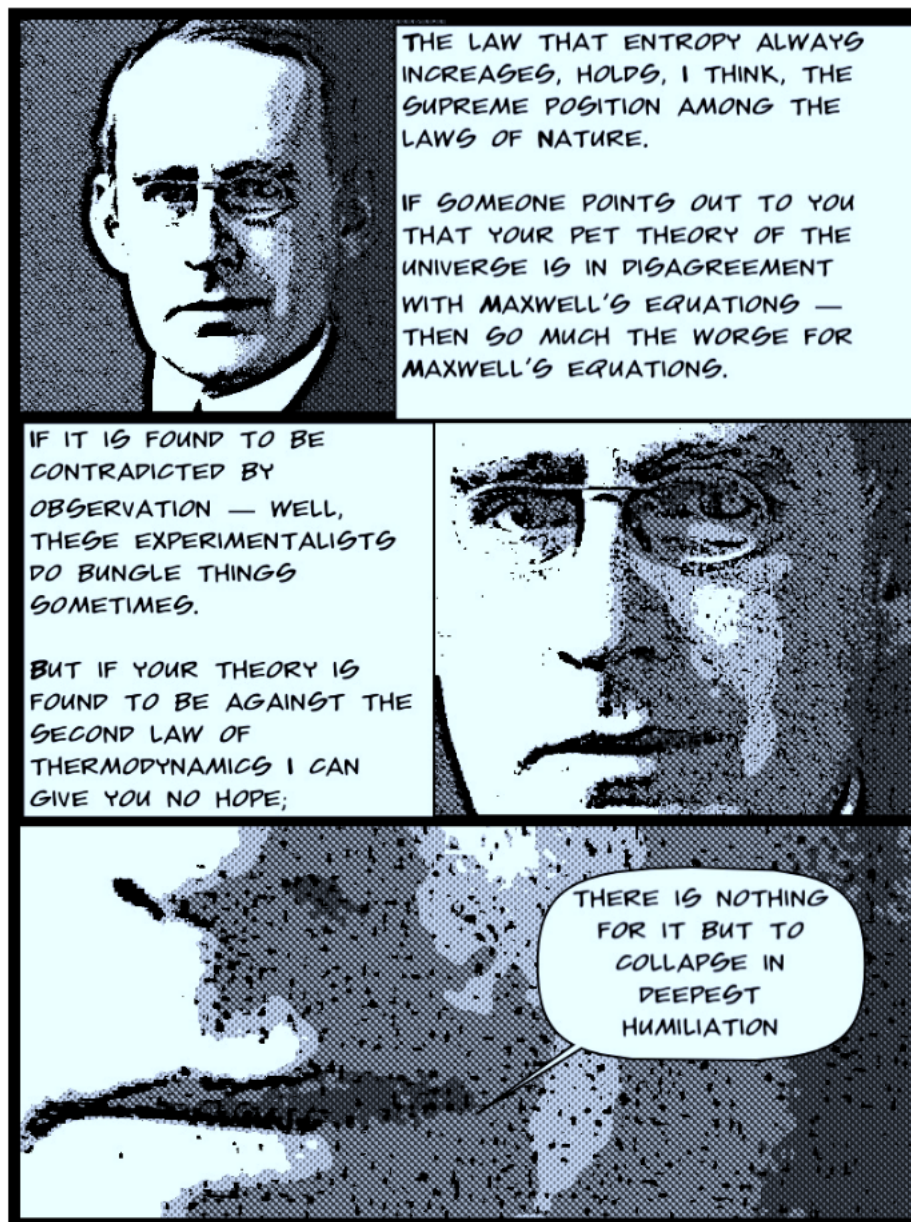


图 2: Arthur Eddington

1.2.5 温度和热容

我们定义温度为

$$\frac{1}{T} = \frac{\partial S}{\partial E} \quad (1.7)$$

这时 confusio 路过我们的课堂，说：“这是个什么奇怪的定义？你怎么保证这个定义和我妈妈说的温度一样？”在之后，我们将会看到更多的这个定义和日常生活符合的迹象。但是现在，这个定义也符合我们的直觉，下面来说明这一点。

找两个系统，他们的温度不同，现在将他们放在一起交换能量，熵的变化为

$$\delta S = \frac{\partial S_1(E_1)}{\partial E} \delta E_1 + \frac{\partial S_2(E_2)}{\partial E} \delta E_2 \quad (1.8)$$

因为能量守恒，故 $\delta E_1 = -\delta E_2$ ，并按照我们对温度的定义，有

$$\delta S = \left(\frac{1}{T_1} - \frac{1}{T_2} \right) \delta E_1 \quad (1.9)$$

由熵增加原理， $\delta S \geq 0$ ，因此我们看到，当 $T_1 > T_2$ ，为了保证熵增加原理，有 $\delta E_1 < 0$ ，1 系统能量减少，将能量传给 2 系统。同理，当 $T_1 < T_2$ 时，1 系统能量增加，能量从 2 系统传给 1 系统。这符合我们日常的直觉——热量只能从温度高的物体传给温度低的物体。因此这样定义温度从这个意义来说是合理的。

我们再定义一个物理量——热容，升高单位温度所获得的能量

$$C = \frac{\partial E}{\partial T} \quad (1.10)$$

热容还有另一个表示，利用 $S = S(E)$ ，则有

$$\frac{\partial S}{\partial T} = \frac{\partial S}{\partial E} \frac{\partial E}{\partial T} = \frac{C}{T} \quad (1.11)$$

因此我们有

$$\Delta S = \int_{T_1}^{T_2} \frac{C(T)}{T} dT \quad (1.12)$$

因此我们将熵转化成了一个可以实际测量的物理量——热容。从上式我们还可以感觉到，热容正比于粒子数，因此还可以将物质的多少除掉，来定义比热容。

最后，我们对温度的定义式进行微分，有

$$\frac{\partial^2 S}{\partial E^2} = \frac{1}{T^2} \frac{\partial T}{\partial E} = -\frac{1}{CT^2} \quad (1.13)$$

这个式子可以帮助我们确认熵是否是最大值，可见，若想使得熵确实是最大值而不是最小值，要求 $\frac{\partial^2 S}{\partial E^2} \leq 0$ ，因此有 $C > 0$ 。事实上，对于几乎所有系统，热容都是正数，但也有少数例外，比如黑洞。如果 $C > 0$ ，这个系统被称作热力学稳定的系统。

1.2.6 压强

实际上，一个系统的熵（即微观态数目的对数）往往还依赖于系统的体积 V 。例如对于经典系统，大量例子的微观态就是知道每一个粒子的位置和动量 (\vec{x}, \vec{p}) ，因此在位置和动量组成的相空间中往往要进行积分 $\int d^3x d^3p$ ，所以熵在我们分析的很多情况下很有可能依赖于体积。

在前面的讨论中，我们定义了温度——它的倒数是熵对于温度的导数。因为熵又依赖于体积，所以还可以定义一个新的量-压强。压强定义为

$$p = T \frac{\partial S}{\partial V} \quad (1.14)$$

至于为什么要这么定义，现在没有答案。重复以前的推导，考虑一个体积恒定的容器，中间被可以自由滑动的隔板分开成两个子系统，当一个系统的体积发生变化时，系统总熵变化

$$\delta S = \frac{\partial S_1}{\partial V} \delta V_1 + \frac{\partial S_2}{\partial V} \delta V_2 \quad (1.15)$$

利用 $\delta V_1 = -\delta V_2$ ，有

$$\delta S = \left(\frac{\partial S_1}{\partial V} - \frac{\partial S_2}{\partial V} \right) \delta V_1 = \left(\frac{p_1}{T_1} - \frac{p_2}{T_2} \right) \delta V_1 \quad (1.16)$$

之前说，两个系统的平衡态是温度相同，那么在温度相同的前提下，两个系统压强相等，则两个系统才达到了平衡。这符合我们对于压强的直觉：平衡时两系统压强相等。习惯上，压强相等被称作力学平衡，它意味着系统的力学量平均值不再演化。

1.2.7 热力学第一定律

利用前面的定义和导数关系式，就有

$$dS = \frac{\partial S}{\partial V} dV + \frac{\partial S}{\partial E} dE \quad (1.17)$$

代入上面的定义，稍加整理得到

$$dE = T dS - p dV \quad (1.18)$$

上面的关系式就是热力学第一定律。注意到我们是倒着来的，先有第二定律，后有第一定律。

1.2.8 一个例子——两态系统

让我们来应用微正则系综来解决一个最简单的问题，考虑一个有 N 个粒子的系统，每个粒子都有两个状态，自旋向上 $|\uparrow\rangle$ 和自旋向下 $|\downarrow\rangle$ 。假设自旋向上的能量为 ϵ ，而自旋向下的取成能量的零点。故取特定能量 $E = N_\uparrow \epsilon$ 的微观态数目为

$$\Omega(E) = \frac{N!}{N_\uparrow!(N - N_\uparrow)!} \quad (1.19)$$

计算熵

$$S(E) = k_B \ln \frac{N!}{N_\uparrow!(N - N_\uparrow)!} \quad (1.20)$$

这时使用斯特林公式 $\ln N! \simeq N \ln N - N + 1$ ，最终计算得到

$$S(E) = -k_B N \left[\left(1 - \frac{E}{N\epsilon}\right) \ln \left(1 - \frac{E}{N\epsilon}\right) + \frac{E}{N\epsilon} \ln \left(\frac{E}{N\epsilon}\right) \right] \quad (1.21)$$

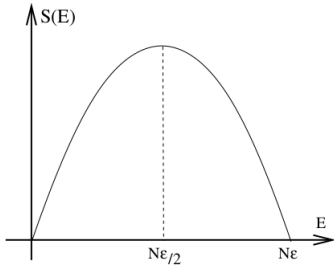


图 3: 熵和能量的关系

将熵和能量的曲线画出来如图 3 所示。从图中我们可以看到，当 $E = N\epsilon$ 和 $E = 0$ 的时候熵为 0，这是因为这两种情况只对应一个态。而当 $E = N\epsilon/2$ 时熵是最大的，此时的熵为

$$S = N k_B \ln 2 \quad (1.22)$$

计算温度

$$\frac{1}{T} = \frac{\partial S}{\partial E} = \frac{k_B}{E} \ln \left(\frac{N}{N_\uparrow} - 1 \right) \quad (1.23)$$

进一步整理得到具体温度下取上自旋的粒子数目

$$\frac{N_\uparrow}{N} = \frac{1}{\exp(\epsilon/k_B T) + 1} \quad (1.24)$$

从上面的式子我们可以看出，当 N_\uparrow/N 不断增加时，温度是从 0 不断增大到无穷大，而当 N_\uparrow/N 继续增大时，温度突然变成了负无穷大，最后趋近于 0。“负温度”似乎在正常情况下难以理解，它对应于增加了系统的能量，微观态数目反而减少。激光的原理正是利用了这一点，当两个能级的粒子数目反转 $N_\uparrow/N > 1/2$ 时，系统就会自发的向 N_\uparrow/N 发展，从而放出光子。

我们进一步可以计算热容

$$C = \frac{dE}{dT} = \frac{N\epsilon^2}{k_B T^2} \frac{e^{\epsilon/k_B T}}{(e^{\epsilon/k_B T} + 1)^2} \quad (1.25)$$

我们看到 C 正比于 N ，事实上对我们考虑的大部分系统，这一性质都成立。当 $T \rightarrow 0$ 时，我们可以看到热容很快地趋于 0，即升高一点温度，这时系统几乎不增加能量。这里的微观原因是因为能级是分立的。而当 $T \rightarrow \infty$ 时，由于更多的激发态被占据，因此升高温度难以使能量升高地更快。

最后，要指出的是，在实际固体中自旋对热容贡献很小，主要的部分是由固体中的电子或者光子气体（对于光子的系统）。但是，在一些特定的材料中，比如顺磁金属盐，自旋对于系统热容的贡献会导致在图样中有一个小偏离。这个小偏离被叫做 schottky anomaly。

上面这样一个两能级系统也可以用来描述晶格中原子的缺陷，其中 spin down 对应于正常的原子，而 spin up 对应于缺失的原子。

1.3 正则系综

在前面微正则系统的分析中，我们是假设整个系统的能量不变。不过在应用中，一个系统会与外界发生作用而导致能量不断扰动，比如放置在环境中的一杯水。说这杯水具有恒定的温度便更加合适。从温度可以来推导平均能量。

我们定一个热库（用字母 R 来表示），而系统用字母 S 来表示。热库的能量比系统的能量大很多，以至于系统可以无条件从热库拿走或者甩给热库任意多的能量而不导致热库的温度的变化。我们把系统 S 可以取的态用 $|n\rangle$ 来表示³，每个态的能量是 E_n ，注意不同的态 $|n\rangle$ 之间的能量可以相同。热库和系统整体考虑成一个孤立系统，应用微正则系综。

$$\Omega(E_{\text{total}}) = \sum_n \Omega_R(E_{\text{total}} - E_n) \equiv \sum_n \exp\left(\frac{S_R(E_{\text{total}} - E_n)}{k_B}\right) \quad (1.26)$$

为了精确理解上面的式子，需要注意以下几点⁴

- 评论 1**
1. 上面式子所有的等号都是严格的，不是近似结果。第二个等号是代入定义。而有人就会问为什么公式里只出现了 Ω_R 却没有出现 Ω_S ？这时要注意到我们前面是把整个系统 S 的态记为 $|n\rangle$ ，并且我们的求和是对所有的系统态 n 求和，而不是对于所有的能量求和。上面的式子相当于分情况讨论，对于所有的系统 S 的态，统计一下热库 R 的态数目，加起来就是整个系统的微观态数目。
 2. 在以后的计算中，我们经常先对相同的能量的态求和，希望学到后面读者不要混淆记号的含义。

利用系统的能量远远小于热库的能量，将式 1.26 进行泰勒展开得到

$$\Omega(E_{\text{total}}) \simeq \sum_n \exp\left(\frac{S_R(E_{\text{total}})}{k_B} - \frac{\partial S_R}{\partial E}(E = E_{\text{total}}) \frac{E_n}{k_B}\right) \quad (1.27)$$

³注意这里表示整个系统取的一个态。

⁴这里不禁要吐槽一下作者当时学统计物理时老师就把这个等号讲成了近似，还生动形象地举了大人和小孩取澳门玩的例子，最后作者的第一遍统计物理完全没有学懂。后来领悟到这些是 2020 年新型冠状病毒肺炎期间居家时。

利用温度的定义，有

$$\Omega(E_{\text{total}}) = \exp\left(\frac{S_R(E_{\text{total}})}{k_B}\right) \sum_n \exp(-E_n/k_B T) \quad (1.28)$$

这时使用等概率，热库和系统整体作为一个微正则系综，每个取 E_{total} 的概率相等，于是系统处于 $|n\rangle$ 的概率为

$$p(n) = \frac{e^{-E_n/k_B T}}{\sum_m e^{-E_m/k_B T}} \quad (1.29)$$

这就是正则系综的核心。每个态的概率按照每个态的能量以指数的方式分配。

1.3.1 配分函数 (Partition Function)

为了方便，我们引入一个新的记号

$$\beta \equiv \frac{1}{k_B T} \quad (1.30)$$

上面的归一化系数记作

$$Z = \sum_n e^{-E_n/k_B T} \quad (1.31)$$

上面的 Z 被叫做配分函数。我更愿意称之为“分配”函数，这个函数接收一个温度 β ，然后给每个态“分配”概率。这个配分函数是统计物理里面最重要的量。实际上，如果知道了一个系统的配分函数，我们就知道了在宏观上我们想知道的全部信息。

下面我们推导配分函数一条重要的性质。那就是：**对于两个独立的系统，它们整个系统的配分函数可以直接相乘**。假设系统整体处于 $|m\rangle|n\rangle$ ，他们的能量为 $E_m^{(1)}, E_n^{(2)}$ ，于是根据定义，系统整体的配分函数为

$$\begin{aligned} Z &= \sum_{m,n} e^{-\beta(E_m^{(1)} + E_n^{(2)})} \\ &= \sum_{m,n} e^{-\beta E_m^{(1)}} e^{-\beta E_n^{(2)}} \\ &= \left(\sum_n e^{-\beta E_n} \right) \left(\sum_m e^{-\beta E_m} \right) \\ &= Z_1 Z_2 \end{aligned}$$

在推导过程中我们用到了下面读者可能不太熟悉的数学：假如 $\{a_n\}, \{b_n\}$ 是

两个数列，则⁵

$$\sum_{m,n} a_m b_n = \left(\sum_m a_m \right) \left(\sum_n b_n \right) \quad (1.32)$$

1.3.2 密度矩阵的一种选择

在实际处理量子力学问题的时候，除了量子力学的本身导致的不确定性，我们还对这个东西到底是什么态而一无所知。即一般的态可以表示为 $|\psi\rangle = c_n |n\rangle$ ，其中 $|n\rangle$ 是完备的本征态。除了量子力学本身的不确定性，我们在很多情况下对前面的系数实际上还一无所知。我们通过定义密度矩阵来反应我们这种无知，密度矩阵算符定义为

$$\hat{\rho} \equiv |\psi\rangle\langle\psi| \quad (1.33)$$

那么有

$$\langle n|\hat{\rho}|n\rangle = |c_n|^2 \quad (1.34)$$

回想关于量子力学中测量的概念，经过一次测量，一个态 $|\psi\rangle$ 会塌缩到其中构成它的一个本征态 $|n\rangle$ ，而塌缩到 $|n\rangle$ 上的概率便是 $|c_n|^2$ 。那么，在我们不知道的情况下，如何选择（瞎猜）一个分布呢？其中最好的一个分布那就是统计物理告诉我们的分布。所以我们取

$$c_m^* c_n = \delta_{mn} e^{-\beta E_n} / Z \quad (1.35)$$

其实相当于取了

$$\hat{\rho} = \frac{e^{-\beta \hat{H}}}{Z} \quad (1.36)$$

1.3.3 平均能量和能量涨落

知道了每个态的概率，我们便可以计算系统的平均能量

$$\begin{aligned} \langle E \rangle &= \sum_n p(n) E_n \\ &= \sum_n \frac{1}{Z} E_n e^{-\beta E_n} \\ &= -\frac{\partial}{\partial \beta} \ln Z \end{aligned}$$

⁵读者在初学统计物理的时候，会碰到很多之前不太熟悉的数学知识，这并不是说读者的数学需要去额外补充，而需要读者向这些新的技巧发起挑战。

可见我们对于配分函数求导数就可以知道平均能量了。下面我们来计算能量的涨落，我们用能量的标准差来代表能量的涨落

$$\begin{aligned}
 \Delta E^2 &= \langle (E - \langle E \rangle)^2 \rangle \\
 &= \langle E^2 \rangle - \langle E \rangle^2 \\
 &= \partial_\beta^2 \ln Z \\
 &= -\frac{\partial \langle E \rangle}{\partial \beta}
 \end{aligned}$$

其中第三个等号需要读者从定义出发经过几步推导。利用热容的定义（注意现在能量不固定，热容定义便顺理成章修改为）

$$C_V = \frac{\partial \langle E \rangle}{\partial T} \Big|_V \quad (1.37)$$

利用 $\beta = 1/k_B T$ ，有

$$\begin{aligned}
 \Delta E^2 &= -\frac{\partial \langle E \rangle}{\partial T} \frac{dT}{d\beta} \\
 &= k_B T^2 C_V
 \end{aligned}$$

这个公式说明的道理并不是很显而易见：系统的能量分散程度正比于系统吸收热量的能力。上述结果只是在平衡态的时候推导的结果，在未来我们会见到上面式子的更一般形式，被叫做 fluctuation-dissipation 定理。

另外，由式 1.38，因为 $C_V \sim N$ ，而 $\Delta E^2 \sim N$ ， $E \sim N$ ，所以我们得到

$$\frac{\Delta E}{E} \sim \frac{1}{\sqrt{N}} \quad (1.38)$$

所以当 $N \rightarrow \infty$ 时， $\frac{\Delta E}{E} \rightarrow 0$ 。这说明随着粒子数的增加，概率密度的形状越来越窄，而当粒子数趋近于无穷大时，峰就会变得无限窄。这就意味着，当 $N \rightarrow \infty$ 时，我们回到了微正则系综——能量近乎于固定。在本书中我们几乎只考虑 N 很大的情况，所以我们今后我们对于 $\langle E \rangle$ 和 E 不加区分，为了简单，都写作 E ，但是读者应该知道每个语境下 E 的具体含义。

1.3.4 题外话：统计物理是理论物理中璀璨的宝石

Die ganzen Zahlen hat der liebe
Gott gemacht, alles andere ist
Menschenwerk.

Kronecker

正如我们即将看到的，正则系综会给计算带来很大的方便，微正则系综的求解过程是求解一道排列组合的数学题（数一数系统的微观态数目），而正则系综的求解过程只需要按部就班套公式就行了。但是哪个过程更能反映统计物理的本质呢？我觉得答案是微正则系综，它的基础是整个统计的基础：数数。我认为，这就是统计物理某种程度上不可能错的原因，正如著名德国科学家克罗内克所说“上帝创造了自然数，其它都是人的作品”。因此，我个人认为，无论描述世界的基础理论如何，牛顿力学、量子力学、量子场论、弦论等等，统计物理的框架和思想都不会发生太大的改变，因为他遵循着世界上最简单的逻辑，也是人类最先学会的数学技巧：数数。

有时候自己想想，统计物理有可能是最不依赖于实验的一门学科——尽管它的建立依赖于实验而且它的定义是尽量去满足早期实验学家的定义。学习统计物理，就像在学习数学，整个建立这个方法论的过程不依赖于任何“事实”和“经验”，你要是问统计物理的公理是什么呀，那可能就一直追溯到有关数的公理了。

统计物理一方面也许是物理学中最强大的武器，只要知道基础理论，采用上面的论证方式，你就可以知道大量粒子的行为！Powerful！而另一方面，我认为它也是理论物理中最优美的一门学科，真的是理论物理中最耀眼和最璀璨的宝石。尽管你以后会看到统计物理会有很多“说不清道不明”的近似，但是从基本原理来讲，统计物理的本质就是数数。而很难想象这样简单的 $1234, \dots$ 的数数是有问题的。

我以前学识浅薄，认为统计物理就是知道了单个粒子的运动去用一种技巧求解大量粒子的运动。研究单个粒子如何运动可能更令人激动。但是当我正式来学这门学科，我就被统计的强大所震撼——它只是基于数数。

也许再过四百年，我们现在所讲的现代物理学框架全被推翻了，统计物理的理论思想也是正确的。

1.3.5 再来计算两能级系统

我们下面用正则系综的方法演示以下如何使用正则系综进行计算。两能级系统由 N 个只能取两能级的粒子组成，而每两个能级之间是互不干涉的。因此我们直接把整个系统拆成 N 个系统，先求单粒子的配分函数

$$Z_1 = \sum_r e^{-\beta E_r} = 1 + e^{-\beta \epsilon} = 2e^{-\beta \epsilon/2} \cosh(\beta \epsilon/2) \quad (1.39)$$

根据我们前面证明的结论，整个系统的配分函数是每个互相独立的子系统的乘积，因此我们有

$$Z = Z_1^N = 2^N e^{-N\beta \epsilon/2} \cosh^N(\beta \epsilon/2) \quad (1.40)$$

接下来读者可以计算平均能量

$$\langle E \rangle = -\partial_\beta \ln Z = \frac{N\epsilon}{2} [1 - \tanh(\beta \epsilon/2)] \quad (1.41)$$

跟我们微正则系综得到的结果相同。

1.3.6 正则系综中的熵

在微正则系综中，我们定义了熵就是固定能量情况下系统可能的微观态数目取对数。但是在正则系综中情况很不一样，这时能量不再是定值，而是取不同的能量有不同的概率，因为我们简单的数数方法只能统计确定能量的微观态数目，那么在正则系综中怎么办呢？

我们假设有 W 个系统，如果 W 很大，那么取 $|n\rangle$ 态的系统数目为 $p(n)W$ ，因此这 W 个系统（整体组成一个孤立系统）的微观态数目为

$$\Omega = \frac{W!}{\prod_n (p(n)W)!} \quad (1.42)$$

运用斯特林公式（推导将来会有），整个 W 个系统的熵

$$\begin{aligned} S = k_B \log \Omega &= k_B (W \log W - W - \sum_n p(n)W \log(p(n)W) + p(n)W) \\ &= -k_B W \sum_n p(n) \ln p(n) \end{aligned}$$

所以每个系统的熵平均为

$$S = -k_B \sum_n p(n) \ln p(n) \quad (1.43)$$

这被叫做吉布斯熵公式。可以看到，正则系综中熵的平均值等于微正则系综中熵的精确值。几十年后，在经典信息理论中的香农熵，量子信息理论中的冯-诺依曼熵，都长成这个样子。另外，在量子力学中熵常用密度矩阵来表述

$$S = -k_B \text{Tr} \hat{\rho} \ln \hat{\rho} \quad (1.44)$$

最后，如果每个态的概率都相同，很容易证明上面熵的定义可以回到孤立系统熵的定义。我们将正则系综的概率代入式 1.43，得到

$$\begin{aligned} S &= -k_B \sum_n p(n) \ln p(n) \\ &= -k_B \sum_n \frac{e^{-\beta E_n}}{Z} \ln (e^{-\beta E_n} / Z) \\ &= -k_B \sum_n \frac{e^{-\beta E_n}}{Z} [(-\beta E_n - \ln Z)] \\ &= \frac{k_B \beta}{Z} \sum_n (E_n e^{-\beta E_n}) + k_B \ln Z \\ &= k_B \frac{\partial}{\partial T} (T \ln Z) \end{aligned}$$

上面的结果就是正则系综中一个很方便的求熵的公式。

1.3.7 自由能

我们定义自由能

$$F = \langle E \rangle - TS \quad (1.45)$$

如果为了和其他的量进行混淆，通常还叫做亥姆霍兹自由能。我们知道熵只是能量和体积的函数，即 $S = S(E, V)$ ，所以能量是表示成熵和体积的函数 $E = E(S, V)$ ，根据之前推导得到的热力学第一定律，有

$$dE = TdS - pdV \quad (1.46)$$

现在，经过上述的变换（又被称作勒让德变换），我们便可以得到 F 是温度和体积的函数，我们很容易得到

$$dF = d\langle E \rangle - d(TS) = -SdT - pdV \quad (1.47)$$

如果知道了一个系统的自由能，那么便可以很容易地求得熵和压强

$$S = - \left(\frac{\partial F}{\partial T} \right)_V \quad (1.48)$$

$$p = - \left(\frac{\partial F}{\partial V} \right)_T \quad (1.49)$$

在固定温度时，自由能是最重要的物理量，因为它和配分函数直接联系

$$\begin{aligned} F = E - TS &= k_B T^2 \frac{\partial}{\partial T} \ln Z - k_B T \frac{\partial}{\partial T} (T \ln Z) \\ &= -k_B T \ln Z \end{aligned}$$

上面的推导使用了 $S = k_B \frac{\partial}{\partial T} (T \ln Z)$ 和

$$\frac{\partial}{\partial \beta} = -k_B T^2 \frac{\partial}{\partial T} \quad (1.50)$$

上面这个式子有个快速得到的小技巧，根据量纲分析， β 有着能量的倒数的量纲，所以等号右边必然是能量的量纲，而我们知道 $k_B T$ 是能量的量纲，我们又知道这是换元，所以我们自信地写下 $k_B T^2 \frac{\partial}{\partial T}$ ，又因为 β 正比于 T 的倒数，换元肯定差个负号，所以填一个负号上去，这样就省的每次算一遍了。

1.4 巨正则系综

1.4.1 化学势

在我们之前的讨论中，我们只考虑了微观态依赖于 E, V 的系统，而一般地，微观态还常常依赖于粒子数 N ，实际上更严谨地，应该把前面的讨论都加上 N 。注意对于不同的系统，微观态所依赖的量是不同的，这个要具体情况具体分析，我们在这里考虑的都是通常的情况。这时有

$$S(E, V, N) = k_B \ln \Omega(E, V, N) \quad (1.51)$$

之前我们定一过温度和压强，他们分别与熵对能量和体积的偏导数有关，因此自然而然地，我们诞生了一个新的偏导数，它描述的是系统的化学平衡（一种粒子转换成另一种粒子，即粒子数的增减）的强度。它被称为化学势

$$\mu = -T \frac{\partial S}{\partial N} \quad (1.52)$$

同样根据上面的分析，我们可以预测两个系统当化学势相等的时候，它们的熵达到最大，即达到了化学平衡。这样，之前的推导应当扩充为

$$\begin{aligned} dS &= \frac{\partial S}{\partial E} dE + \frac{\partial S}{\partial V} dV + \frac{\partial S}{\partial N} dN \\ &= \frac{1}{T} dE + \frac{p}{T} dV - \frac{\mu}{T} dN \end{aligned}$$

挪一挪，得到

$$dE = T dS - p dV + \mu dN \quad (1.53)$$

从上面的式子我们可以看出化学势的含义了，它是放进单位粒子所花费的能量（当然，在熵和体积都固定的情况下）。这里可以和电荷与电势作一个类比，电势是放进单位电荷所花费的能量。

值得指出的是，化学势一般都是负数，即放进一个粒子会导致系统的总能量减小，这是因为我们固定熵的缘故。一般来说，温度 $\frac{1}{T} = \partial S / \partial E$ 为正数，所以能量越高，系统的熵越大。而这时加入一个粒子，重新和其他粒子争夺能量的分配，会导致熵增大，而这时只有总能量减小，没有那么多能量可以分配给每个粒子，从而维持熵不变。

另外，根据式 1.53 我们也可以写出化学势的另外一个定义

$$\mu = \left(\frac{\partial E}{\partial N} \right)_{S, V} \quad (1.54)$$

很容易证明两者相等。这里用到了下面的公式：如果三个变量 (x, y, z) 有依赖关系，那么

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1 \quad (1.55)$$

证明如下。利用

$$dz = \left(\frac{\partial z}{\partial x}\right)_y dx + \left(\frac{\partial z}{\partial y}\right)_x dy \quad (1.56)$$

令 $dz = 0$ ，即可求得当 z 不变时， y 对 x 的偏导数，进而可以证明上面的关系。

应用上面的推导出的关系，可以得到

$$\begin{aligned} \left(\frac{\partial E}{\partial N}\right)_{S,V} &= -\left(\frac{\partial E}{\partial S}\right)_{V,N} \left(\frac{\partial S}{\partial N}\right)_{V,E} \\ &= -T \left(\frac{\partial S}{\partial N}\right)_{V,E} \\ &= \mu \end{aligned}$$

很显然，验证了上面的两个定义是等价的。

1.4.2 巨正则系综

现在，我们考虑系统的粒子数不固定，而化学势 μ 固定的情况。同样地，我们考虑一个系统热库，系统和热库之间的化学势和温度相同，而热库的粒子数和能量远远大于系统的粒子数和能量。假设系统和热库总能量和粒子数分别为 $E_{\text{total}}, N_{\text{total}}$ ，那么系统和热库整体作为一个微正则系综，有

$$\begin{aligned} \Omega(E_{\text{total}}, N_{\text{total}}) &= \sum_n \Omega_R(E_{\text{total}} - E_n, N_{\text{total}} - N_n) \\ &= \sum_n \exp(S(E_{\text{total}} - E_n, N_{\text{total}} - N_n)/k_B) \\ &\simeq \sum_n \exp\left(S(E_{\text{total}}, N_{\text{total}})/k_B - \frac{\partial S}{\partial E} \cdot E_n/k_B - \frac{\partial S}{\partial N} \cdot N_n/k_B\right) \\ &= \sum_n \exp\left(S(E_{\text{total}}, N_{\text{total}})/k_B - \frac{E_n}{k_B T} + \frac{\mu}{k_B T} N_n\right) \\ &= \sum_n e^{S(E_{\text{total}}, N_{\text{total}})/k_B} e^{-\beta(E_n - \mu N_n)} \end{aligned}$$

上面推导过程中 \sum_n 和之前推导正则系综的意义相同，是对于系统的每个态求和。从上面的推导我们可以看到，一组能量和粒子数对应的态占据了的微观态数目正比于 $e^{-\beta(E_n - \mu N_n)}$ ，所以我们可以得到取态 n 的概率为

$$p(n) = \frac{1}{Z} e^{-\beta(E_n - \mu N_n)} \quad (1.57)$$

上面这个概率分布就被称作巨正则系综。其中 \mathcal{Z} 是概率的归一化系数，被称作巨配分函数，定义为

$$\mathcal{Z} = \sum_n e^{-\beta(E_n - \mu N_n)} \quad (1.58)$$

同样地，知道了巨配分函数，实际上就掌握了我们要知道的统计物理中的全部信息。重复和之前类似的计算，我们得到

$$\begin{aligned} S &= -k_B \sum_n p(n) \ln p(n) \\ &= k_B \frac{\partial}{\partial T} (T \ln \mathcal{Z}) \end{aligned}$$

和之前一样，也可以得到能量

$$\langle E \rangle - \mu \langle N \rangle = -\partial_\beta \ln \mathcal{Z} \quad (1.59)$$

和粒子数

$$\langle N \rangle = -\partial_\mu \ln \mathcal{Z} \quad (1.60)$$

和前面一样，我们也可以计算平均值的扰动，在这里我们给出 N 的扰动结果

$$\Delta N^2 = \frac{1}{\beta^2} \frac{\partial^2}{\partial \mu^2} \ln \mathcal{Z} = \frac{1}{\beta} \frac{\partial \langle N \rangle}{\partial \mu} \quad (1.61)$$

可以看到 $\Delta N / \langle N \rangle \sim 1/\sqrt{\langle N \rangle}$ ，当 $N \rightarrow \infty$ 时，三个系综的结果趋于一致。

1.4.3 巨正则势

受前面正则系综的启发，我们定义巨正则势，它是自由能的勒让德变换

$$\Phi = F - \mu N \quad (1.62)$$

进行变换后，使得

$$d\Phi = -SdT - pdV - Nd\mu \quad (1.63)$$

从而巨正则势是温度、压强和化学势的函数 $\Phi = \Phi(T, V, \mu)$ 。将上面的公式代入，可以得到

$$\Phi = -k_B T \ln \mathcal{Z} \quad (1.64)$$

是不是和想象的一样？我还是建议大家自己验证一下。

1.4.4 强度量和广延量

我们知道互相孤立的系统之间熵是可以直接相加的。我们现在把一个系统复制 λ 份，这就意味着它的能量、体积和粒子数同时被复制了 λ 份，所以 λ 个系统的总熵为

$$S(\lambda E, \lambda V, \lambda N) = \lambda S(N, V, E) \quad (1.65)$$

可见熵也变为 λ 倍，我们把这样的量叫做广延量 (extensive quantities)。而下面的量

$$\frac{1}{T} = \frac{\partial S}{\partial E}, p = T \frac{\partial S}{\partial V}, \mu = T \frac{\partial S}{\partial N} \quad (1.66)$$

在系统变为 λ 倍之后显然仍然是不变的。这样的量被称作强度量 (intensive quantities)。

强度量和广延量的概念不但能给上面这些量更好的理解，更可以给我们带来十分有用的结论，比如我们可以知道巨正则势 $\Phi = \Phi(T, V, \mu)$ 是广延量，因为

$$\Phi(T, \lambda V, \mu) = \lambda \Phi(T, V, \mu) \quad (1.67)$$

从这里我们可以得到 $\Phi \propto V$ ，又因为之前的全微分表达式 1.63，我们发现

$$\Phi = -p(T, \mu)V \quad (1.68)$$

上面这个是一个求得压强的一个非常方便的式子。

综上，我们介绍了巨正则系综的内容，巨正则系综是用来处理粒子数不固定而化学势固定的系统。但是在之后的内容中，我们将看到巨正则系综直接用来处理固定粒子数的系统。这只是因为先使用巨正则系综得到量之间的关系十分方便，但是我们在后续的处理中还要强制固定 $\langle N \rangle$ 不变，这时候巨正则系综量之间的依赖关系会发生改变，我们将在第三章量子气体中看到这一点。

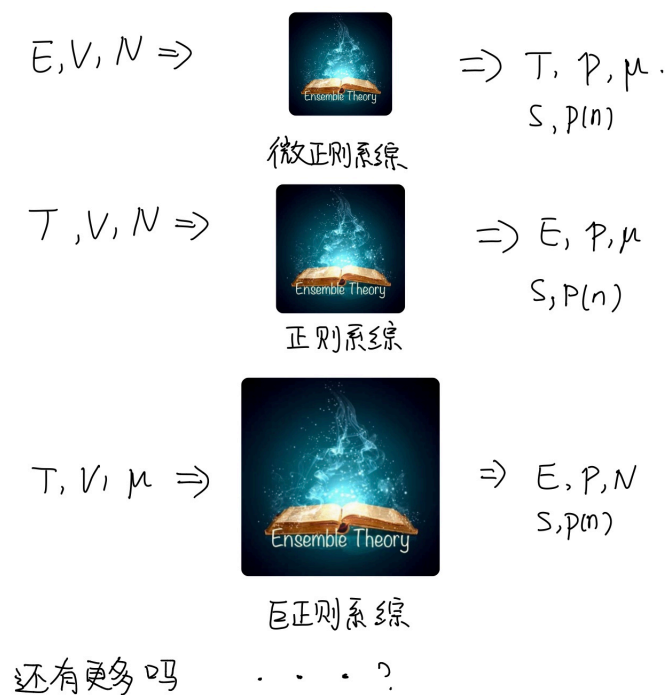


图 4: 三种系综

1.5 对于系综理论的思考

当读者第一次这样学完系综理论的时候,一定和我一样,有很多问号。系综理论到底是干嘛的,为什么三个系综叫这个名字?为什么正则系综是使得体系的能量不定,温度恒定,为什么不可以定义一个新的系综,使得体系的体积不定但是压强恒定?下面我来带大家梳理一下。

1.5.1 三种系综理论的总结

所以什么是系综?我们可以用一种“输入、输出”的观点来看待系综理论到底能够能为我们带来什么。三种系综的示意图如图 4,我们先从微正则系综开始。微正则系综是系统具有确定的能量、体积和粒子数(严谨的论证应该是同时知道这三个量,一开始为了使得论证简洁而清晰,所以没加),之后我们通过计算 $\Omega(E, V, N)$, 便得到了每个态出现的概率,进而可以计算系统的温度、压强和化学势(强度量),正如我们在二能级系统所做的那样。所以说微正则系综是给定系统的 E, V, N , 输出温度、压强、化学势。这中间的媒介就是微观

态数目随着 E, V, N 的变化关系。

我们再来看正则系综。正则系综是给定了系统的温度、体积和粒子数，通过与热库建立联系，从而知道了每个态的概率是 $p(n) = e^{-\beta E_n}/Z$ ，进而求得能量、压强和化学势。这中间的媒介就是 $p(n)$ ，或者等效的，配分函数就包含了所有的信息。

巨正则系综便是给定了系统的温度、体积和化学势，通过与热库建立联系，得到 $p(n)$ ，进而计算出剩下三个量的。

到这里我们便清楚了系综理论到底是干嘛的。能量与温度、体积与压强、粒子数与化学势，从这三组里面每组抽取一个，通过合适的系综理论，便可以得到另外三个（不能都抽取强度量，下面会解释）。所以显然，可以定义很多种系综，比如读者可以定义这样的系综 给出 $(E, p, N) \rightarrow (T, V, \mu)$ ，显然这种系综也能用。但是对于目前我们研究的问题，这三种系综已经够用了，再定义新的也只是玩玩数学游戏。那是不是对于研究什么样的问题，我们就用什么样的系综？不是的，读者可以证明，所有系综在 $N \rightarrow \infty$ 时所得到的结果都是一样的（这一点很关键）。所以在你得到的关系之后，你可以随意施加你想要的条件，尽管这条件和最初系综理论的条件不同，但是最后的结果也是对的。

1.5.2 自己新造一种系综？最巨系综？

上面说能量与温度、体积与压强、粒子数与化学势，从这三组里面每组抽取一个，通过合适的系综理论，便可以得到另外三个。或者说，可以定义一个新的系综理论。那么能不能都取成强度量，比如我们取一个系统温度、压强和化学势都固定，我能不能造一个系综理论，借助一个热库，求得系统的能量、体积和粒子数？这时候显然就不行了。因为当你取定了一个系统的温度、压强和化学势，这个系统本身就变成了什么？就变成了热库！所以对应的广延量可以等倍数放大而不改变强度量。这个时候定义不出对应的系综理论。

但是，上面的论证却可以帮助我们思考三种系综名字的由来！我们看，微正则系综是三个广延量固定，正则系综是两个广延量和一个强度量固定，而巨正则系综是两个强度量和一个广延量固定。我们可以看到，随着系综越来越“巨大”，固定的广延量越来越少，而固定的强度量越来越多。我们刚刚已经想过，三个强度量都固定是什么东西，是热库！而热库很大！所以三个广延量都固定了便是一个“最巨大”的系统，无论做什么它的强度量都不改变。这也许就是为什么三个系综前面的修饰词越来越大的原因了。

上面的讨论只是我们将要研究的一般情况，如果熵所依赖的变量不只三个

怎么办？那这时，我们就可以定义出新的系综了。

1.5.3 不同条件下使得熵最大

这部分就给读者以提示，大家按照提示自己推导。

还有一种普遍的方法来看待三种系综，将熵增加原理看作第一性原理，利用吉布斯熵公式（我们在 1.3.6 中也推导过）

$$S = -k_B \sum_n p(n) \ln p(n)$$

施加一个限制 $\sum_n p(n) = 1$ ，利用拉格朗日乘子法使得熵最大，尝试证明对于所有具有确定能量的态，概率都相等。

当施加的条件变为平均能量固定 $\langle E \rangle = \sum_n p(n) E_n$ ，同样利用拉格朗日乘子法，推导出正则系综的概率分布。

当固定平均能量 $\langle E \rangle$ 和平均粒子数 $\langle N \rangle$ 时，采用同样的方法，可以求出巨正则系综的概率分布。

1.6 约西亚·威拉德·吉布斯 (Josiah Willard Gibbs)



图 5: 约西亚·威拉德·吉布斯

这一章我们学到的好多东西都要归功于吉布斯 (1839 年 2 月 11 日 - 1903 年 4 月 28 日), 自由能、化学势, 还有系综的概念。甚至“统计物理”这个名词都是他发明的。

他的经历似乎很“平凡”, 本科在耶鲁大学, 博士在耶鲁大学, 教授也在耶鲁大学。他任耶鲁的前九年没有拿任何薪水, 只凭借着父母的积蓄生活。后来霍普金斯大学给他开出 3000 美元的年薪, 耶鲁为了挽留他给出了 2000 美元的年薪, 他最终还是留在了耶鲁。

这九年里他发表了三篇关于统计物理的论文, 并寄给世界各地一百四十七个物理、数学的科学家, 请他们提供意见。几乎所有人都读不懂他的理论, 也不知道吉布斯是何许人。吉布斯对各种学术团体发表研究的会议都没有参加, 不擅交际又爱深思的个性, 使他在现实的社会里, 几乎是不存在的边缘人。这种没有名气的文章, 有谁肯花脑筋去读呢? 有! 而且是当时最杰出的科学家, 电磁学大师麦克斯韦! 他深深地赞赏吉布斯的文章, 于是登高一呼: “这个人对于热的解释, 已经超过所有德国科学家的研究了。”这时大家才恍然大悟, 回头从纸屑堆中找出这三篇文章, 好好的研读。

吉布斯死于一九零三年四月二十八日。他的学生们遵照他的吩咐, 在耶鲁的小教堂里为他朗诵几段圣经, 唱几首他所喜爱的诗歌。他终身没有结婚, 全付心力去探索数学、热力学的美, 与教授耶鲁大学的学生。(资料来源: 百度百科)

2 经典气体

到现在我们就学会了统计物理的一般方法，之后的第二章和第三章我们就来应用我们学到的东西。我们将会学习大量的例子，希望我们能通过这些例子更加深刻地体会统计物理。

2.1 经典配分函数

在这一章我们将来计算和分析经典气体。经典气体就是量子气体在 $h \rightarrow 0$ 时的极限。由于经典气体的粒子数 N 固定，所以我们将正则系综中分析问题，那么第一步我们就要来求经典气体的配分函数。我们知道配分函数的定义为

$$Z = \sum_n e^{-\beta E_n} \quad (2.1)$$

那么对于经典气体，一个微观态就是知道每个粒子的位置和动量 (\vec{q}_i, \vec{p}_i) ，那么对于每个微观态的求和在经典极限下就相当于对于 (q, p) 空间积分。我们马上就会证明，配分函数为

$$Z = \frac{1}{h^3} \int d^{3N}p d^{3N}q e^{-\beta H(p,q)} \quad (2.2)$$

似乎符合我们的猜想。这时有人问了，我们研究的是经典气体，那么哪来的 h ？实际上，这个 h 并不影响我们求解所有的物理量，因为 h 为常数，我们通常会对配分函数取对数然后求导数。下面一个严格的证明，会帮助我们理解 h 的来源。为了描述简洁，我们考虑一维情况，并假设每个粒子相互独立，所以只考虑单个粒子的配分函数（回忆如果每个粒子相互独立，则系统可以分成 N 个子系统，总配分函数 $Z = Z_1^N$ ），三维情况读者可以推广。

非相对论情况下我们假设哈密顿量

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{q}) \quad (2.3)$$

所以单粒子的配分函数可以写成

$$Z_1 = \sum_n e^{-\beta E_n} = \sum_n \langle n | e^{-\beta \hat{H}} | n \rangle \quad (2.4)$$

我们插入量子力学中的单位算符 $\hat{I} = \int dq |q\rangle\langle q|$, 我们先把 $V(\hat{q})$ 的部分提出来

$$\begin{aligned}
 Z_1 = \sum_n e^{-\beta E_n} &= \sum_n \langle n | e^{-\beta \hat{H}} | n \rangle \\
 &= \sum_n \langle n | \int dq |q\rangle\langle q| e^{-\beta \hat{H}} \int dq' |q'\rangle\langle q'| n \rangle \\
 &= \int dq dq' \langle q | e^{-\beta \hat{H}} | q' \rangle \sum_n \langle n | q \rangle \langle q' | n \rangle \\
 &= \int dq dq' \langle q | e^{-\beta \hat{H}} | q' \rangle \sum_n \langle q' | n \rangle \langle n | q' \rangle \\
 &= \int dq dq' \langle q | e^{-\beta \hat{H}} | q' \rangle \langle q | q' \rangle \\
 &= \int dq \langle q | e^{-\beta \hat{H}} | q \rangle
 \end{aligned}$$

通过量子力学我们知道对于两个算符有 $e^{\hat{A}} e^{\hat{B}} = e^{\hat{A}+\hat{B}} e^{[\hat{A}, \hat{B}]/2}$, 那么对于我们的 $e^{-\beta \hat{H}}$, 利用 $[q, p] = i\hbar$

$$e^{-\beta(\frac{\hat{p}^2}{2m} + V(\hat{q}))} = e^{\frac{\hat{p}^2}{2m}} e^{-\beta V(\hat{q})} + O(\hbar) \quad (2.5)$$

因为在经典极限我们令 $\hbar \rightarrow 0$, 所以我们得到

$$Z_1 = \int dq \langle q | e^{-\beta \hat{p}^2/2m} e^{-\beta V(\hat{q})} | q \rangle \quad (2.6)$$

$$= \int dq e^{-\beta V(q)} \langle q | e^{-\beta \hat{p}^2/2m} | q \rangle \quad (2.7)$$

再插入单位算符 $\hat{I} = \int dp |p\rangle\langle p|$, 得到

$$Z_1 = \int dq dp dp' e^{-\beta V(q)} \langle q | p \rangle \langle p | e^{-\beta \hat{p}^2/2m} | p' \rangle \langle p' | q \rangle \quad (2.8)$$

在一维情况, 利用

$$\langle q | p \rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{ipq/\hbar} \quad (2.9)$$

所以

$$Z_1 = \frac{1}{h^3} \int dp dq e^{-\beta H(p,q)} \quad (2.10)$$

便得到了一维情况的经典配分函数, 和三维情况式 2.2 相似。

注意! 上面的 p, q 对于任何广义坐标和广义动量均成立, 而不只限于直角坐标分量。这是因为在量子力学中, $\langle q | p \rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{ipq/\hbar}$ 对于任何广义坐标和广义动量都成立。因为动量算符起源于空间的平移对称性, 所以可以知道任何动量算符在 q 表象下都是 $P = -i\frac{\partial}{\partial q}$, 从而任何 P 的本征态波函数都满足上面的形式。

2.2 理想气体

理想气体忽略了气体分子之间的相互作用，因此

$$\hat{H} = \frac{\hat{p}^2}{2m} \quad (2.11)$$

这里要指出的是，分子间的相互作用完全忽略，只是意味着哈密顿量取成了这个形式而已，而不是说所有气体分子就一直作匀速直线运动。别忘了，即使在微正则系综中，系统的态也会因为微扰而反复横跳，正是这种反复横跳才导致了统计物理的基本假设——处于相同能量的态出现概率相等。在理想气体的情况，可以认为这种反复横跳是由于分子之间碰撞所导致的。

在正式计算之前还要唠叨的一句就是，当我们学了系综理论之后，有好多种办法来求得最后的结果，比如可以先求得配分函数，然后通过求导得到各个量。也可以通过求得概率分布的方式，然后再由定义求出各个量。也可以通过自由能等函数，来求得想要的量。最后都会导致相同的结果，大家实际上爱用哪种用哪种。对于简单的例子，读者可能利用各种方式都能顺利地推导出结果。但是在之后的一些例子可能就会比较困难。实际上在理想气体的计算中，读者也可以感受到不同方法的难易。因此，有的时候当我们没有统一地采用先求配分函数的方法时，读者不要感觉到惊讶。读者如果忘记了用到的方法，应该多多回到第一章去看一看。

在这里，我们使用正则系综。我们来计算配分函数

$$\begin{aligned} Z_1(V, T) &= \frac{1}{h^3} \int d^3p \, d^3q \, e^{-\beta \vec{p}^2/2m} \\ &= \frac{V}{(2\pi\hbar)^3} \int d^3p \, e^{-\beta \vec{p}^2/2m} \\ &= \frac{V}{(2\pi\hbar)^3} \left(\sqrt{\frac{2\pi m}{\beta}} \right)^3 \\ &= V \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2} \end{aligned}$$

我们还可以从上式中发现一个令人惊奇的事实！右边这一项 $\left(\frac{mk_B T}{2\pi\hbar^2} \right)^{1/2}$ 恰好正比于气体分子的德布罗意波长。我们定义热德布罗意波长为

$$\lambda = \sqrt{\frac{2\pi\hbar^2}{mk_B T}} \quad (2.12)$$

因为我们马上会看到 $k_B T \sim mv^2$ 。所以 Z_1 可以表达为

$$Z_1(V, T) = \frac{V}{\lambda(T)^3} \quad (2.13)$$

而德布罗意波长代表着分子波函数的大小。所以这里配分函数有了个新的意义，那就是数一数在这么大的体积，最多能装下多少个这么大热德布罗意波长的分子。配分函数为

$$Z(N, V, T) = Z_1^N = \frac{V^N}{\lambda^{3N}} \quad (2.14)$$

我们来计算压强

$$\begin{aligned} p &= - \left(\frac{\partial F}{\partial V} \right)_T \\ &= \frac{\partial}{\partial V} (k_B T \ln Z) \\ &= \frac{\partial}{\partial V} \left[N k_B T \left(\ln V + \frac{3}{2} \ln \left(\frac{m k_B T}{2\pi \hbar^2} \right) \right) \right] \\ &= \frac{N k_B T}{V} \end{aligned}$$

所以我们得到了我们高中就喜闻乐见的理想气体状态方程

$$pV = N k_B T \quad (2.15)$$

怎么样？是不是特别震撼，似乎只能从实验和经验得到的理想气体状态方程，现在我们用简单的原理加上漂亮的数学推导就得到了。想想就觉得不可思议，我们以后会看到，世界上那么多现象，从宇宙微波背景辐射到日常的气体，从低温的现象到高温的现象，这套统计物理的方法竟然都好用！

2.2.1 能量均分定理

我们来计算能量

$$\begin{aligned} E = - \frac{\partial}{\partial \beta} \ln Z &= k_B T^2 \frac{\partial}{\partial T} \ln Z \\ &= k_B T^2 \frac{\partial}{\partial T} \left[\frac{3}{2} N \ln \left(\frac{m k_B T}{2\pi \hbar^2} \right) \right] \\ &= \frac{3}{2} N k_B T \end{aligned}$$

我们进而得到热容

$$C_V = \left(\frac{\partial E}{\partial T} \right)_V = \frac{3}{2} N k_B \quad (2.16)$$

对于经典气体，有下面一条一般的定理

定理 1 能量均分定理

当 $\theta e^{-E(\theta, X)/k_B T} \Big|_{\theta_{\min}}^{\theta_{\max}} = 0$ 时 (下面会说明这条式子的含义), 每一个独立的能量 n 次型贡献 $\frac{1}{n} k_B T$ 的能量。

例如我们考虑的理想气体

$$H = \frac{1}{2} m (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) \quad (2.17)$$

它的能量由三个独立的二次型组成。最终导致能量为 $3 \times \frac{1}{2} k_B T$ 。实际上, 不论对于经典气体还是以后要学到的量子气体, 能量均分定理在温度足够高的情况下都成立。

下面我们来严格证明这条定理。假设坐标 θ 在能量的表达式中贡献了一个正比于 θ^n 的能量项, 我们记其他的坐标为 X , 所以系统处于 (θ, X) 的概率为

$$P = \frac{e^{-E(\theta, X)/k_B T}}{\int d\theta dX e^{-E(\theta, X)/k_B T}} = C \cdot e^{-E(\theta, X)/k_B T} \quad (2.18)$$

利用归一化条件

$$\int d\theta dX C \cdot e^{-E(\theta, X)/k_B T} = 1 \quad (2.19)$$

进行分部积分, 得到

$$\int dX \theta e^{-E(\theta, X)/k_B T} \Big|_{\theta_{\min}}^{\theta_{\max}} + \frac{1}{k_B T} \int \theta e^{-E(\theta, X)/k_B T} \frac{\partial E}{\partial \theta} = 1 \quad (2.20)$$

如果 $\theta e^{-E(\theta, X)/k_B T} \Big|_{\theta_{\min}}^{\theta_{\max}} = 0$, 故有

$$\langle \theta \frac{\partial E}{\partial \theta} \rangle = k_B T \quad (2.21)$$

我们代入 $E \sim \theta^n$, 得到

$$\langle E \rangle = \frac{1}{n} k_B T \quad (2.22)$$

从而能量均分定理得到了证明。注意上面的条件 $\theta e^{-E(\theta, X)/k_B T} \Big|_{\theta_{\min}}^{\theta_{\max}} = 0$ 与能量的形式和广义坐标上下界的选取都有关。例如对于一维情况的哈密顿量

$$H = \frac{p^2}{2m} + \lambda q^3 \quad (2.23)$$

如果对于 q 不加合适的限制, 那么则不满足能量均分定理。实际上, 这样一个势能会导致所有粒子都滑向无穷远处势能无穷低的地方, 最后导致能量发散。再如, 对于一个实验室的盒子, 则不能使用能量均分定理。

还应该注意的，上面“独立的” n 次型的含义，我们一定要记住我们是在一组共轭的坐标和动量下积分。例如，对于下面的哈密顿量，这两个 p 和 q 具有迷惑性。在积分时，应该把他们都化成球坐标或者直角坐标。发现化成球坐标时，无法直接套用能量均分定理。

$$H = \frac{p^2}{2m} + \lambda q^4 \quad (2.24)$$

下面我们来计算三维空间上面哈密顿量对应的能量，读者可以从下面的推导中学到一些技巧，在下面的推导中我们在直角坐标中积分。

能量的平均值为

$$\begin{aligned} \langle E \rangle &= N \left\langle \frac{p^2}{2m} + \lambda q^4 \right\rangle \\ &= N \left(\frac{\int 4\pi p^2 dp (p^2/2m) e^{-\beta(p^2/2m)}}{\int 4\pi p^2 dp e^{-\beta(p^2/2m)}} + \frac{\lambda \int 4\pi q^6 e^{-\beta\lambda q^4} dq}{\int 4\pi q^2 e^{-\beta\lambda q^4} dq} \right) \end{aligned}$$

其中

$$\int 4\pi q^6 e^{-\beta\lambda q^4} dq = \int \frac{4\pi q^6}{4\beta\lambda q^3} e^{-\beta\lambda q^4} d(\beta\lambda q^4)$$

进行换元 $u = \beta\lambda q^4$, $q = (u/\beta\lambda)^{1/4}$, 得到

$$\begin{aligned} \int 4\pi q^6 e^{-\beta\lambda q^4} dq &= \int \frac{4\pi q^6}{4\beta\lambda q^3} e^{-\beta\lambda q^4} d(\beta\lambda q^4) \\ &= \frac{\pi}{\beta\lambda} \int_0^\infty (u/\beta\lambda)^{3/4} e^{-u} du \\ &= \frac{\pi}{(\beta\lambda)^{7/4}} \int_0^\infty u^{7/4-1} e^{-u} du \end{aligned}$$

我们用到 Gamma 函数的定义

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx \quad (2.25)$$

它是阶乘的推广。从分部积分很容易证明它和阶乘满足类似的性质

$$\Gamma(\alpha + 1) = \alpha \Gamma(\alpha) \quad (2.26)$$

因此有

$$\int 4\pi q^6 e^{-\beta\lambda q^4} dq = \frac{\pi}{(\beta\lambda)^{7/4}} \Gamma\left(\frac{7}{4}\right) \quad (2.27)$$

同理，我们可以得到

$$\int 4\pi q^2 e^{-\beta\lambda q^4} dq = \frac{\pi(\beta\lambda)^{1/4}}{\beta\lambda} \Gamma\left(\frac{3}{4}\right) \quad (2.28)$$

套用 Γ 函数的性质我们有 $\Gamma\left(\frac{7}{4}\right) = \frac{3}{4}\Gamma\left(\frac{3}{4}\right)$, 故有

$$\langle E \rangle = N \left(\frac{3}{2} k_B T + \lambda \frac{\Gamma(7/4)}{\Gamma(3/4)} \frac{1}{\beta \lambda} \right) \quad (2.29)$$

$$= \frac{9}{4} N k_B T \quad (2.30)$$

这个例子的作用就是要警示读者, 清楚能量均分定理的适用范围。另外, 更重要的是, 读者应该清楚一个哈密顿量下只能选一组广义坐标, 有时候简写, 读者应该注意辨别。

2.2.2 熵和吉布斯悖论

我们回到理想气体, 我们之前推导过

$$Z(N, V, T) = Z_1^N = \frac{V^N}{\lambda^{3N}} \quad (2.31)$$

回想这个式子的推导过程, 我们是把每个粒子单独当作了一个独立的系统。但是, 这实际上有一些问题, 因为这些在量子力学中, 这些粒子是不可分辨的, 所以应该把他们当作不可分辨的 N 个系统, 那么 \sum_n 所统计的微观态数目应该有所缩减, 按照量子力学, 正确的统计方式应该是

$$Z(N, V, T) = \frac{1}{N!} Z_1^N = \frac{V^N}{N! \lambda^{3N}} \quad (2.32)$$

这样修改之后不会影响熵和能量, 但是会改变熵, 我们下面来计算熵

$$\begin{aligned} S &= \frac{\partial}{\partial T} (k_B T \ln Z) \\ &= k_B \ln Z + \frac{3}{2} N k_B \\ &\simeq k_B [N \ln V - N \ln N + N - N \ln \lambda] + \frac{3}{2} N k_B \\ &= N k_B \left[\ln \left(\frac{V}{N \lambda^3} \right) + \frac{5}{2} \right] \end{aligned}$$

上面的近似过程用到了斯特林公式。如果没有配分函数的 $N!$, 则会导致中括号内第一项对数内没有分母的 N , 这就导致了如下的吉布斯悖论。

考虑一个系统, 粒子数为 N , 体积为 V 。现在考虑将这个系统加倍 $N \rightarrow 2N, V \rightarrow 2V$, 如果我们的配分函数中没有加入 $N!$, 那么我们不会得到熵加倍的结论, 而是多了一个奇怪的 $\ln 2$ 的系数。而我们知道, 熵是一个广延量, 所以系统的熵应该加倍。这就被称作吉布斯悖论。这可能是在量子力学发现之前, 第一次意识到量子力学的存在!

2.2.3 化学势

我们进而可以求出化学势，以便更好地理解化学势。根据定义

$$\mu = -T \frac{\partial S}{\partial N} = k_B T \ln \left(\frac{\lambda^3 N}{V} \right) \quad (2.33)$$

可以看到，当 $\lambda^3 \ll V/N$ 时，即量子效应不明显，化学势为负。回忆前面的另一个等价定义

$$\mu = \left(\frac{\partial E}{\partial N} \right)_{S,V} \quad (2.34)$$

这正如前面学习化学势提到的，当加入一个粒子而控制熵和体积不变，只能减小总能量。

但是化学势也有可能大于 0，这是因为粒子之间很强的排斥力。

2.3 麦克斯韦分布

由上面求得的概率公式，我们可以求得取到 $v \rightarrow v + dv$ 的概率为

$$f(v)dv = 4\pi v^2 e^{-mv^2/2k_B T} dv \quad (2.35)$$

上面的这个分部就被称作麦克斯韦分布。从这个分布可以求出特定温度下气体分子的最概然速率、平均速率等等。这里就不占篇幅求了。

2.4 双原子分子

我们下面来研究两个原子组成的分子。这时贡献的能量项数就变多了，我们即考虑质心的平动，也考虑双原子分子的转动，还有两个原子之间距离的振动。因为平动、转动和振动之间互不干涉，因此现在单粒子配分函数可表示为三部分相乘

$$Z_1 = Z_{\text{trans}} Z_{\text{rot}} Z_{\text{vib}} \quad (2.36)$$

我们先求转动部分的配分函数，我们选取坐标为 (θ, ϕ) ，转动部分的拉格朗日量为

$$\mathcal{L}_{\text{rot}} = \frac{1}{2} I (\dot{\theta}^2 + \sin^2 \theta \dot{\phi}^2) \quad (2.37)$$

进而可以求得转动的哈密顿量为

$$H_{\text{rot}} = \frac{p_\theta^2}{2I} + \frac{p_\phi^2}{2I \sin^2 \theta} \quad (2.38)$$

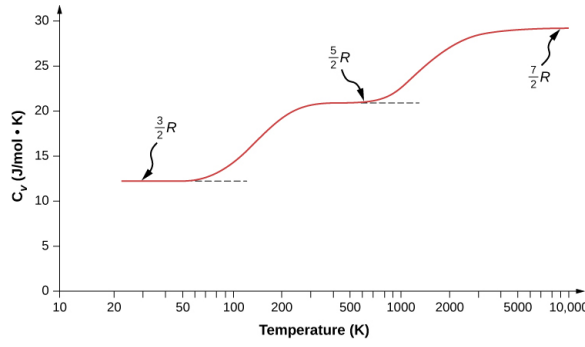


图 6: 双原子分子热容随温度变化的实验结果

所以我们可以求得配分函数

$$\begin{aligned} Z_{\text{rot}} &= \frac{1}{(2\pi\hbar)^2} \int d\theta d\phi dp_\theta dp_\phi e^{-\beta H_{\text{rot}}} \\ &= \frac{2I}{\beta\hbar^2} \end{aligned}$$

进而我们可以求得一个粒子转动的平均能量

$$\begin{aligned} E &= -\partial_\beta \ln Z_{\text{rot}} \\ &= k_B T \end{aligned}$$

进而验证了能量均分定理。

如果只能转动，不能转动，则系统应该有能量

$$E = \frac{5}{2} N k_B T \quad (2.39)$$

热容

$$C_V = \frac{5}{2} N k_B \quad (2.40)$$

如果再加上振动的自由度，直接根据能量均分定理，考虑到振动既有动能也有势能，因此给能量贡献了两个独立的二次型，故最后都加起来为

$$C_V = \frac{7}{2} N k_B \quad (2.41)$$

太好了！我们从经典的情况完全分析了双原子分子的热容，我们来看看实验结果吧！实验结果图 6 所示。从实验结果我们发现，结果并不是我们预测的那样。为什么会这样？事实上，在低温情况下，转动自由度和振动自由度被冻结，只

有平动部分有贡献，而当温度升高到室温附近，两个转动自由度被激活，热容变为 $5/2Nk_B$ ，当温度升高到 $\sim 5000K$ 左右，振动自由度被激活，热容变为 $7/2Nk_B$ 。

为什么会这样呢？这是因为量子力学中能级是离散的！所以当温度很低时，由温度引起的能量波动不足以使系统激发到激发态上，因此对于转动和振动自由度，只能呆在基态上。在第三章量子气体中，我们还会从量子力学的角度仔细解释一下这一点。

在实验上，测量双原子气体热容的实验也十分具有意义，它是人类第一次测到量子力学效应！

2.5 相互作用的经典气体

之前我们没有考虑气体分子之间的相互作用。我们从气体状态方程的测量可以知道，当气体非常稀薄，即 N/V 很小时，上面的理想气体模型是非常好的。如何描述非理想气体？人们便想到了按照 N/V 对状态方程展开

$$\frac{p}{k_B T} = \frac{N}{V} + B_2(T) \frac{N^2}{V^2} + B_3(T) \frac{N^3}{V^3} + \dots \quad (2.42)$$

其中 $B_j(T)$ 被称作位力系数，上面的展开又被称作位力展开。这一节我们的目的就是第一性原理计算上面的位力系数。

首先我们应该清楚气体分子之间的相互作用。气体分子之间的相互作用有以下两个特点

1. 正比于 $1/r^6$ 。这是因为两个中性原子之间电偶极矩的相互作用。由于量子扰动，其中一个分子会产生一个电偶极矩 p_1 ，这一电偶极矩产生的电场 $E \propto p_1/r^6$ ，这样一个电场会极化第二个分子产生电偶极矩 $p_2 \sim E \sim p_1/r^3$ ，所以可以得到两个电偶极矩的相互作用势为 $p_1 p_2 / r^3 \sim 1/r^6$ 。
2. 当两个气体分子靠的非常近时，会产生一个极强的排斥力，这个排斥力起源于泡利不相容原理。

为了给一个经典的近似，人们通常采用 Lennard-Jones 势来做近似

$$U(r) \sim \left(\frac{r_0}{r}\right)^{12} - \left(\frac{r_0}{r}\right)^6 \quad (2.43)$$

这个 12 没有什么物理意义，只是让它在趋近于 0 时候发散。另一个更简单的

近似是钢球势近似

$$U(r) = \begin{cases} +\infty & , r < r_0 \\ -U_0 \left(\frac{r_0}{r}\right)^6 & , r > r_0 \end{cases} \quad (2.44)$$

2.5.1 梅耶 (Mayer) f 函数和第二位力系数

简单起见，我们先只考虑两个分子之间的相互作用。

2.6 德拜-休克尔等离子体模型

3 量子气体

4 相变初步