

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	13
1.3.5	再来计算两能级系统	14
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	30
2.2.1	能量均分定理	31
2.2.2	熵和吉布斯悖论	34
2.2.3	吉布斯悖论的更深一步讨论	35
2.2.4	化学势	36
2.3	麦克斯韦分布	37
2.3.1	题外话：高斯积分的小技巧	37
2.3.2	题外话： N 维球的体积	38
2.4	双原子分子	39
2.5	相互作用的经典气体	41
2.5.1	梅耶 (Mayer) f 函数和第二位力系数	41
2.5.2	范德瓦耳斯状态方程	43
2.5.3	计算更多粒子的相互作用：集团展开	45
2.6	德拜-休克尔等离子体模型	49
3	量子气体	50
3.1	量子气体的配分函数	51
3.1.1	正则系综	52
3.1.2	巨正则系综	53
3.2	态密度	55
3.2.1	非相对论系统	55
3.2.2	相对论系统	55
3.3	光子气体	57
3.3.1	Stefan-Boltzmann 定律	59
3.3.2	辐射压强	59
3.3.3	宇宙微波背景辐射	60
3.3.4	量子力学的诞生	61
3.4	声子	62

3.4.1	德拜模型	62
3.5	回顾双原子分子热容	64
3.6	玻色统计	66
3.6.1	玻色-爱因斯坦统计	67
3.6.2	理想玻色气体	68
3.6.3	高温玻色气体	69
3.6.4	低温极限——玻色爱因斯坦凝聚	70
3.6.5	低温玻色气体的状态方程	72
3.6.6	玻色气体的热容——BEC 是一种相变	73
3.6.7	附录：多重对数函数	73
3.6.8	附录：使用正则系综推导玻色-爱因斯坦分布	74
3.7	费米统计	75
3.7.1	费米-狄拉克分布	75
3.7.2	理想费米气体	76
3.7.3	低温极限：简并费米气体和费米面	77
3.7.4	低温费米气体	79
3.7.5	索末菲展开	82
3.7.6	白矮星和钱德拉塞卡极限	85
3.7.7	泡利顺磁性	87
3.7.8	朗道抗磁性	87
4	相变初步	88

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。

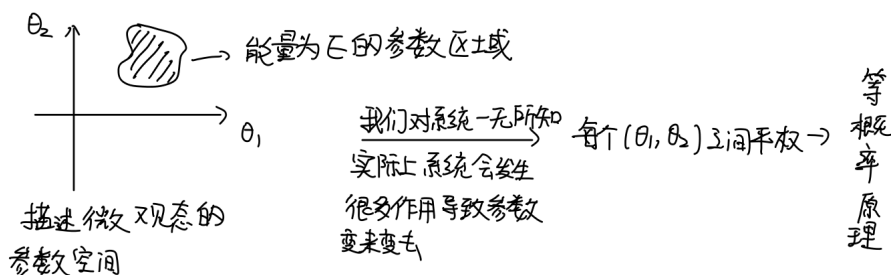


图 1: 等概率原理解释

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?

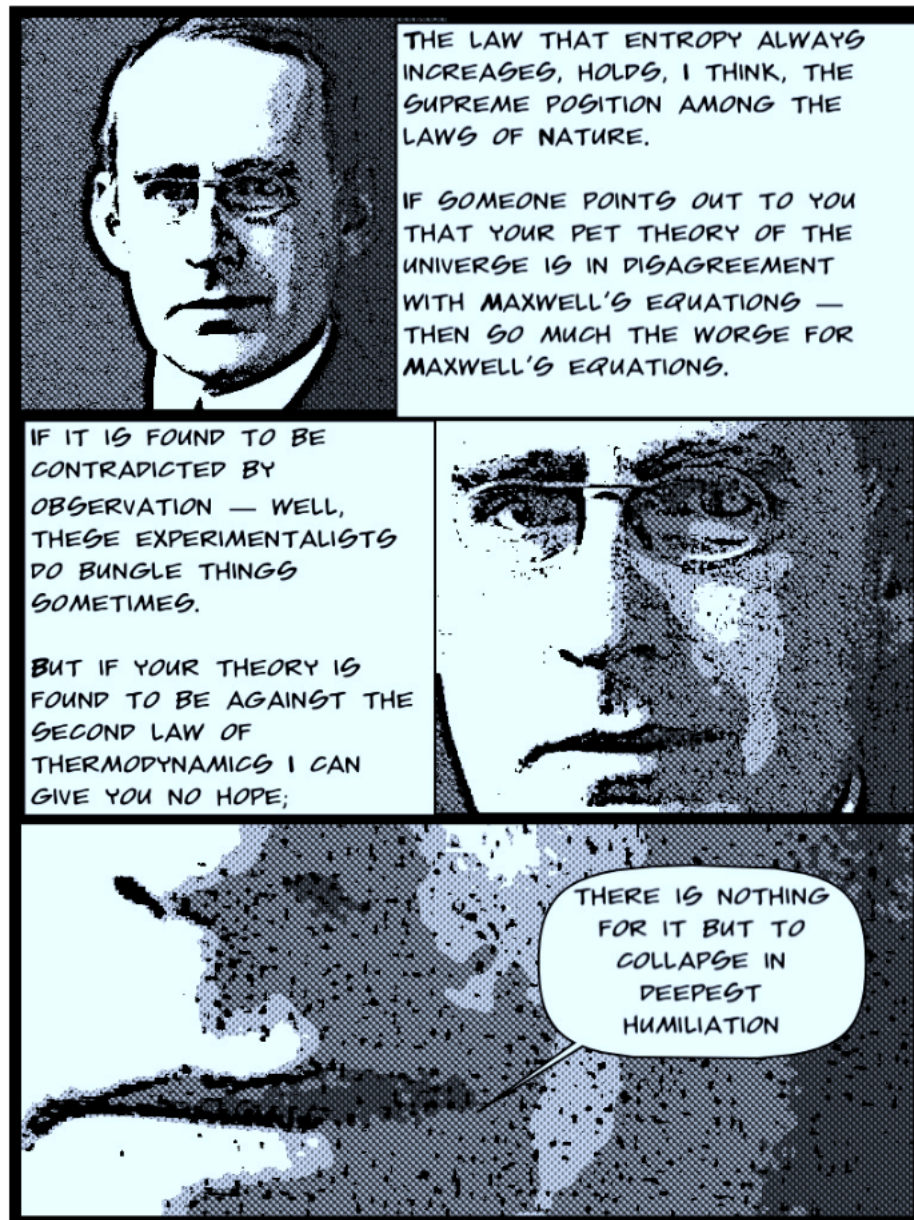


图 2: Arthur Eddington

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

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

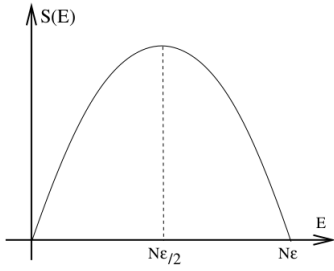


图 3: 熵和能量的关系

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

$$S = Nk_B \ln 2 \quad (1.22)$$

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。不过在实际应用中，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

$$Z = \sum_n e^{-\beta E_n} \quad (1.31)$$

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

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（每个粒子取的两能级不同，即你取你的两个能级，我取我的两个能级）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

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

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。

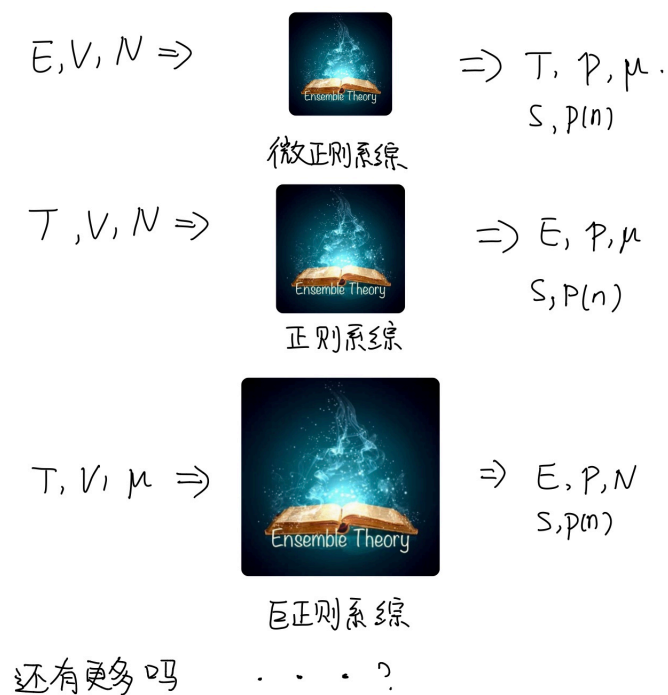


图 4: 三种系综

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

一片花飞减却春，风飘万点正愁人。
且看欲尽花经眼，莫厌伤多酒入唇。
江上小堂巢翡翠，苑边高冢卧麒麟。
细推物理须行乐，何用浮名绊此身。

杜甫·《曲江二首·其一》

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 $\hbar \rightarrow 0$ 时的极限。也就意味着 $[x, p] = i\hbar$ ，光凭对易关系可能不能看出这里近似的意义，更重要的是在量子力学中由对易关系推导的不确定性原理

$$\langle \Delta X^2 \rangle \langle \Delta P^2 \rangle \geq \left| \frac{1}{2} [x, p] \right| \quad (2.1)$$

如果令 $\hbar \rightarrow 0$ ，更重要的是意味着

$$\langle \Delta X^2 \rangle \rightarrow 0, \langle \Delta P^2 \rangle \rightarrow 0 \quad (2.2)$$

这将成为区分经典气体和量子气体的标志。如果粒子是空间上可以分辨的，那么则有（注意我们前面两能级系统也是空间上可以分辨的系统）

$$Z = \frac{Z_1^N}{N!} \quad (2.3)$$

这里 $N!$ 是因为组成气体的原子是完全相同的，我们将在 2.2.2 小节中详细论述。上面的式子意味着，每一个粒子可以看成是完全独立的系统。但是随着温度降低，量子效应的显现，上式便不能使用了，这是因为粒子之间发生交叠，你已经不能用经典的图像来理解粒子了！这时的“粒子”应该从量子化的角度去理解，即系统的微观态是一个能级上有多少个被激发的“粒子”。**这时你不能再经典的图像理解粒子了！**虽然这时你依然能计算单粒子配分函数 Z_1 ，但是总配分函数不再满足上式。在第三章我们会看到量子的配分函数如何变成这一配分函数。

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

$$Z = \sum_n e^{-\beta E_n} \quad (2.4)$$

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

$$Z = \frac{1}{N!} \frac{1}{h^{3N}} \int d^{3N}p d^{3N}q e^{-\beta H(p,q)} \quad (2.5)$$

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。

需要指出的是，整个系统是用 $6N$ 个坐标和动量描述的，每两个坐标可以一样也可以不一样，每个粒子之间互不干扰，所以我们才能对配分函数作出那样的拆分。如果考虑量子系统，上面的方法就是不适用的，因为我们在第三章会看到，量子系统中的能级不是属于粒子自己的，而是属于大家的，到那个时候我们会用另一种拆分方式。

非相对论情况下我们假设哈密顿量

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{q}) \quad (2.6)$$

所以单粒子的配分函数可以写成

$$Z_1 = \sum_n e^{-\beta E_n} = \sum_n \langle n | e^{-\beta \hat{H}} | n \rangle \quad (2.7)$$

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 $[q, p] = i\hbar$

$$e^{-\beta(\frac{p^2}{2m} + V(\hat{q}))} = e^{\frac{p^2}{2m}} e^{-\beta V(\hat{q})} + O(\hbar) \quad (2.8)$$

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

$$Z_1 = \int dq \langle q | e^{-\beta p^2/2m} e^{-\beta V(\hat{q})} | q \rangle \quad (2.9)$$

$$= \int dq e^{-\beta V(q)} \langle q | e^{-\beta p^2/2m} | q \rangle \quad (2.10)$$

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

$$Z_1 = \int dq dp dp' e^{-\beta V(q)} \langle q | p \rangle \langle p | e^{-\beta p'^2/2m} | p' \rangle \langle p' | q \rangle \quad (2.11)$$

在一维情况, 利用

$$\langle q | p \rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{ipq/\hbar} \quad (2.12)$$

所以

$$Z_1 = \frac{1}{h^3} \int dp dq e^{-\beta H(p,q)} \quad (2.13)$$

便得到了一维情况的经典配分函数, 和三维情况式 2.5 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。

还要注意的，上面的推导只是为了解释系数 h 如何来的，并不是经典配分函数的量子推导，因为量子气体中

$$Z = \frac{Z_1^N}{N!} \quad (2.14)$$

不再成立！如果大家想看量子气体如何变成经典气体，请看第三章。

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

$$\hat{H} = \frac{\hat{p}^2}{2m} \quad (2.15)$$

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

$$\lambda = \sqrt{\frac{2\pi\hbar^2}{mk_B T}} \quad (2.16)$$

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

$$Z_1(V, T) = \frac{V}{\lambda(T)^3} \quad (2.17)$$

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

$$Z(N, V, T) = Z_1^N / N! = \frac{V^N}{N! \lambda^{3N}} \quad (2.18)$$

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

$$pV = N k_B T \quad (2.19)$$

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！

练习 1 尝试用巨正则系综求解理想气体问题。

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

$$C_V = \left(\frac{\partial E}{\partial T} \right)_V = \frac{3}{2} N k_B \quad (2.20)$$

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

$$H = \frac{1}{2} m (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) \quad (2.21)$$

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, 不论是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

$$P = \frac{e^{-E(\theta, X)/k_B T}}{\int d\theta dX e^{-E(\theta, X)/k_B T}} = C \cdot e^{-E(\theta, X)/k_B T} \quad (2.22)$$

利用归一化条件

$$\int d\theta dX C \cdot e^{-E(\theta, X)/k_B T} = 1 \quad (2.23)$$

进行分部积分, 得到

$$\int dX \theta e^{-E(\theta, X)/k_B T} \Big|_{\theta_{\min}}^{\theta_{\max}} + \frac{1}{k_B T} \int \theta e^{-E(\theta, X)/k_B T} \frac{\partial E}{\partial \theta} = 1 \quad (2.24)$$

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

$$\left\langle \theta \frac{\partial E}{\partial \theta} \right\rangle = k_B T \quad (2.25)$$

我们代入 $E \sim \theta^n$, 得到

$$\langle E \rangle = \frac{1}{n} k_B T \quad (2.26)$$

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

$$H = \frac{p^2}{2m} + \lambda q^3 \quad (2.27)$$

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。

还应该注意的是, 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。

$$H = \frac{p^2}{2m} + \lambda q^4 \quad (2.28)$$

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

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx \quad (2.29)$$

它是阶乘的推广。从分部积分很容易证明它满足和阶乘类似的性质

$$\Gamma(\alpha + 1) = \alpha \Gamma(\alpha) \quad (2.30)$$

因此有

$$\int 4\pi q^6 e^{-\beta\lambda q^4} dq = \frac{\pi}{(\beta\lambda)^{7/4}} \Gamma\left(\frac{7}{4}\right) \quad (2.31)$$

同理，我们可以得到

$$\int 4\pi q^2 e^{-\beta\lambda q^4} dq = \frac{\pi(\beta\lambda)^{1/4}}{\beta\lambda} \Gamma\left(\frac{3}{4}\right) \quad (2.32)$$

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

$$\langle E \rangle = N \left(\frac{3}{2} k_B T + \lambda \frac{\Gamma(7/4)}{\Gamma(3/4)} \frac{1}{\beta\lambda} \right) \quad (2.33)$$

$$= \frac{9}{4} N k_B T \quad (2.34)$$

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

$$Z(N, V, T) = Z_1^N = \frac{V^N}{\lambda^{3N}} \quad (2.35)$$

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。所以如果考虑粒子的不可分辨性。

我们下面来更加清楚的说明这一点。我们将第 i 个粒子的坐标和动量记作 $\xi_i = (\vec{x}_i, \vec{p}_i)$ ，那么描述经典系统的坐标便是 (ξ_1, \dots, ξ_N) 。配分函数里面的对所有态求和实际上应该写作

$$\sum_n \rightarrow \sum_{(\xi_1, \dots, \xi_N)} \quad (2.36)$$

我们如果考虑交换 1 号粒子和 2 号粒子。从量子的观点来看，整个系统变了吗？没有变。所以交换 1 号和 2 号粒子的态其实是相同的。然而让 ξ_i 遍历所有可能的态的过程将同一个状态求和了两遍。所以实际的求和过程包括的这两项

$$e^{-\beta E}(\xi_1 = \eta_1, \xi_2 = \eta_2) + e^{-\beta E}(\xi_1 = \eta_2, \xi_2 = \eta_1) \quad (2.37)$$

只能求和一次，所以应该变为

$$\frac{1}{2!} \sum_{\xi_1, \xi_2} e^{-\beta E} \quad (2.38)$$

上面我们只考虑了两个粒子的交换过程，如果 N 个粒子都无法区分，随便交换的话，式 2.37 会有 $N!$ 项。所以正确的配分函数应该写为

$$Z(N, V, T) = \frac{1}{N!} Z_1^N = \frac{V^N}{N! \lambda^{3N}} \quad (2.39)$$

这样修改之后不会影响状态方程和能量，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。

吉布斯说，考虑两个粒子数均为 N ，体积均为 V ，完全相同的两团气体，中间放置了一个隔板，两团气体完全一样。现在将隔板抽去，那么两团气体会互相扩散而导致系统整体的熵变大。但是由于两个系统完全一样，所以熵应该不变。这就导致了吉布斯悖论。实际上，这个悖论正是由没有考虑上面的 $N!$ 所导致的，如果考虑了 $N!$ 的因子，熵就不会增加了。这可能是在量子力学发现之前，人们第一次意识到世界可能是由完全相同的原子所构成的。

2.2.3 吉布斯悖论的更深一步讨论

你也许会问，那如果各个粒子不同该怎么办？如果这样，那我们一开始的假设

$$S = S(E, V, N) \quad (2.40)$$

便有问题了，这是因为熵不仅仅依赖于总粒子数，可以分解地依赖于：第一种粒子的粒子数 N_1 ，第二种粒子的粒子数 N_2 ，等等。当每一种粒子都加倍，整个系统的熵才会加倍。（这就是我在第一章强调广延量的定义是 E, V, N 都加倍的道理。）如果只是粒子的种类加倍了，那显然不能保证熵确实是加倍的。

值得指出的是，在正则系综那里我们曾经考虑了简单的两态系统，我们在那里其实也应该加上 $N!$ 。

2.2.4 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

$$\begin{aligned}\mu &= -T \left(\frac{\partial S}{\partial N} \right)_{V,E} \\ &= -T \left[\left(\frac{\partial S}{\partial N} \right)_{V,T} - \left(\frac{\partial S}{\partial T} \right)_{N,V} \left(\frac{\partial T}{\partial N} \right)_{E,V} \right] \\ &= k_B T \ln \left(\frac{\lambda^3 N}{V} \right)\end{aligned}$$

其中我们运用到了切换固定量公式，其推导如下。我们知道熵是 E, V, N 的函数，当固定 V 时，有

$$(dS)_V = \left(\frac{\partial S}{\partial N} \right)_{V,E} dN + \left(\frac{\partial S}{\partial E} \right)_{V,N} dE \quad (2.41)$$

我们现在再加入固定 T 的条件

$$(dS)_{V,T} = \left(\frac{\partial S}{\partial N} \right)_{V,E} (dN)_T + \left(\frac{\partial S}{\partial E} \right)_{V,N} (dE)_T \quad (2.42)$$

两边同时除以 $(dN)_T$ ，得到

$$\left(\frac{\partial S}{\partial N} \right)_{V,T} = \left(\frac{\partial S}{\partial N} \right)_{V,E} + \left(\frac{\partial S}{\partial T} \right)_{N,V} \left(\frac{\partial T}{\partial N} \right)_{E,V} \quad (2.43)$$

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

$$\mu = \left(\frac{\partial E}{\partial N} \right)_{S,V} \quad (2.44)$$

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

$$f(v)dv = 4\pi v^2 e^{-mv^2/2k_B T} dv \quad (2.45)$$

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。

2.3.1 题外话：高斯积分的小技巧

这是题外话，可以跳过。

在麦克斯韦分布有关计算中，我们常常遇到高斯积分

$$G_n = \int_{-\infty}^{+\infty} |x|^n e^{-\frac{x^2}{2}} dx \quad (2.46)$$

下面我们介绍一种积分技巧。首先我们推导一个递推公式

$$\begin{aligned} G_n &= 2 \int_0^{\infty} x^n e^{-\frac{x^2}{2}} dx \\ &= 2 \int_0^{\infty} e^{-\frac{x^2}{2}} d\left(\frac{x^{n+1}}{n+1}\right) \\ &= x^{n+1} e^{-\frac{x^2}{2}} \Big|_0^{\infty} + \int \frac{x^{n+2}}{n+1} e^{-x^2/2} dx \\ &= \frac{1}{n+1} G_{n+2} \end{aligned}$$

所以我们得到

$$G_{n+2} = (n+1)G_n \quad (2.47)$$

这样，我们只要知道了 G_0 和 G_1 ，我们就得到了所有的 G_n 。 G_0 可以由很多种方法得到，我一般使用高斯分布来进行记忆

$$G_0 = \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi} \quad (2.48)$$

而 G_1 是一个简单积分

$$\begin{aligned} G_1 &= 2 \int_0^{\infty} x e^{-\frac{x^2}{2}} \\ &= 2 \int_0^{\infty} e^{-u} du \\ &= 2 \end{aligned}$$

所以根据递推公式,

$$G_{2n} = \sqrt{2\pi}(2n-1)!!, \quad G_{2n+1} = 2 \times (2n)!! \quad (2.49)$$

知道上面的技巧就可以快速计算高斯积分了。

2.3.2 题外话: N 维球的体积

高斯分布的归一化关系还能给我们 N 维球的体积公式。

我们考虑 N 个变量的各向同性的高斯分布, 由于各向同性, 所以分布函数可以转化为径向分布函数, 意味着求 $r \rightarrow r + dr$ 的概率, 那么根据原来的概率分布, 又因为这是一个小球壳, 所以小球壳上的概率可认为是常数。这样的一个小球壳的概率等于 n 维球的表面积乘上 dr , 我们记 n 维球的表面积为 $S_n r^{n-1}$, 则有

$$dV_n = S_n r^{n-1} dr \quad (2.50)$$

进而径向概率分布函数为

$$f(r) = \frac{S_n}{(2\pi)^{n/2}} r^{n-1} e^{-r^2/2} \equiv C_n r^{n-1} e^{-r^2/2} \quad (2.51)$$

我们记 $C_n = \frac{S_n}{(2\pi)^{n/2}}$, 我们利用归一化条件, 便可以求出 C_n , 进而求出 S_n , 归一化条件为

$$C_n \int_0^\infty r^{n-1} e^{-r^2/2} dr = 1 \quad (2.52)$$

利用上一小节的结论

$$C_n \int_0^\infty r^{n-1} e^{-r^2/2} dr = C_n G_{n-1} \quad (2.53)$$

所以 $C_n = 1/G_{n-1}$ 。从上式我们可以得到四维空间球的表面积为 $2\pi^2 r^3$, 体积为 $\pi^2 r^4/2$, 实际上, 你可以求任意维的。

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

$$Z_1 = Z_{\text{trans}} Z_{\text{rot}} Z_{\text{vib}} \quad (2.54)$$

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

$$\mathcal{L}_{\text{rot}} = \frac{1}{2} I (\dot{\theta}^2 + \sin^2 \theta \dot{\phi}^2) \quad (2.55)$$

进而可以求得转动的哈密顿量为

$$H_{\text{rot}} = \frac{p_{\theta}^2}{2I} + \frac{p_{\phi}^2}{2I \sin^2 \theta} \quad (2.56)$$

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，不能振动，则系统应该有能量

$$E = \frac{5}{2} N k_B T \quad (2.57)$$

热容

$$C_V = \frac{5}{2} N k_B \quad (2.58)$$

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

$$C_V = \frac{7}{2} N k_B \quad (2.59)$$

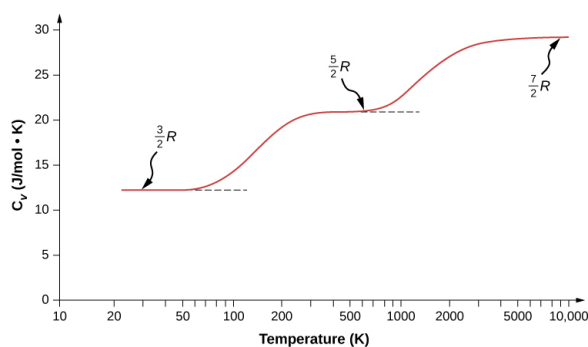


图 6: 双原子分子热容随温度变化的实验结果

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 $5/2 Nk_B$ ，当温度升高到 $\sim 5000K$ 左右，振动自由度被激活，热容变为 $7/2 Nk_B$ 。

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

$$\frac{p}{k_B T} = \frac{N}{V} + B_2(T) \frac{N^2}{V^2} + B_3(T) \frac{N^3}{V^3} + \cdots \quad (2.60)$$

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

$$U(r) = U_0 \left[\left(\frac{r_0}{r} \right)^{12} - \left(\frac{r_0}{r} \right)^6 \right] \quad (2.61)$$

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

$$U(r) = \begin{cases} +\infty & , r < r_0 \\ -U_0 \left(\frac{r_0}{r} \right)^6 & , r > r_0 \end{cases} \quad (2.62)$$

2.5.1 梅耶 (Mayer) f 函数和第二位力系数

气体作用的哈密顿量

$$H = \sum_{i=1}^N \frac{p_i^2}{2m} + \sum_{i<j} U(r_{ij}) \quad (2.63)$$

其中 $r_{ij} = |\vec{r}_i - \vec{r}_j|$ 。因此整个系统的配分函数

$$\begin{aligned} Z(N, V, T) &= \frac{1}{N!} \frac{1}{(2\pi\hbar)^{3N}} \int \prod_{i=1}^N d^3p_i d^3r_i e^{-\beta H} \\ &= \frac{1}{N!} \frac{1}{(2\pi\hbar)^{3N}} \left(\int \prod_{i=1}^N d^3p_i e^{-\beta p_i^2/2m} \right) \times \left(\int \prod_{i=1}^N d^3r_i e^{-\beta \sum_{j<k} U(r_{jk})} \right) \\ &= \frac{1}{N! \lambda^{3N}} \int \prod_{i=1}^N d^3r_i e^{-\beta \sum_{j<k} U(r_{jk})} \end{aligned}$$

我们只考虑弱相互作用的情况，即 $U_0 \ll k_B T$ 很小。所以我们很自然地想到把积分按照 U 展开

$$\exp \left(-\beta \sum_{j<k} U(r_{jk}) \right) = 1 - \beta \sum_{j<k} U(r_{jk}) + \frac{\beta^2}{2} \left(\sum_{j<k} U(r_{jk}) \right)^2 + \cdots \quad (2.64)$$

但是上面的展开系数并不是一个很好的展开系数。因为对积分 $U(r)$ 积分是发散的。显然一个发散的东​​西不能拿来当展开系数。取而代之，我们采用梅耶 (Mayer) f 函数来当作展开系数，Mayer f 函数定义为

$$f(r) = e^{-\beta U(r)} - 1 \quad (2.65)$$

这个 Mayer f 函数对于坐标积分就不会出现发散的情况。而当 $\beta U_0 \ll 1$ 时， $f(r)$ 也具有 βU_0 的量级。所以， $f(r)$ 是一个很好的展开系数。所以我们有

$$\begin{aligned} Z(N, V, T) &= \frac{1}{N! \lambda^{3N}} \int \prod_i d^3r_i \prod_{j<k} (1 + f_{jk}) \\ &= \frac{1}{N! \lambda^{3N}} \int \prod_i d^3r_i \left(1 + \sum_{j>k} f_{jk} + \sum_{j>k, l>m} f_{jk} f_{lm} + \cdots \right) \end{aligned}$$

我们先只考虑两个粒子的相互作用，因此我们把第二项后面的都扔掉。第一项积分结果为 V^N ，对于第二项，由于任何两个粒子之间的 f 函数都一样，所以对于任何的 j 和 k ，结果都为

$$\int \prod_{i=1}^N d^3r_i f_{12} = V^{N-2} \int d^3r_1 d^3r_2 f(r_{12}) \quad (2.66)$$

我们作如下的换元

$$\begin{aligned} \vec{R} &= \frac{1}{2}(\vec{r}_1 + \vec{r}_2) \\ \vec{r} &= \vec{r}_1 - \vec{r}_2 \end{aligned}$$

注意换元时, 对于体积元要乘上 Jacobi 行列式的绝对值。得到

$$\begin{aligned}
 \int \prod_{i=1}^N d^3 r_i f_{12} &= V^{N-2} \int d^3 r_1 d^3 r_2 f(r_{12}) \\
 &= V^{N-2} \int \left| \frac{D(r_1, r_2)}{D(R, r)} \right|^3 d^3 R d^3 r f(r) \\
 &= V^{N-1} \int d^3 r f(r)
 \end{aligned}$$

这时随便抓了两个 j, k , 因此对于全部的 $j < k$ 求和, 求和的个数为 $\frac{1}{2}N(N-1) \simeq \frac{1}{2}N^2$ 个。因此配分函数变为

$$Z(N, V, T) = \frac{V^N}{N! \lambda^{3N}} \left(1 + \frac{N^2}{2V} \int d^3 r f(r) + \cdots \right) \quad (2.67)$$

我们现在只考虑 $f(r)$ 的一阶项, 这意味着我们可以为所欲为: 只要前两项和上面一样的式子都可以看作是在这个近似下正确的。所以我们有

$$\begin{aligned}
 Z(N, V, T) &= \frac{V^N}{N! \lambda^{3N}} \left(1 + \frac{N^2}{2V} \int d^3 r f(r) + \cdots \right) \\
 &= \frac{V^N}{N! \lambda^{3N}} \left(1 + \frac{N}{2V} \int d^3 r f(r) + \cdots \right)^N
 \end{aligned}$$

上面的操作只是为了确保广延量 $F = -k_B T \ln Z$ 正比于 N 。实际上, 最后我们要得到的状态方程也不依赖于是否进行这一步操作。我们可以计算自由能

$$F = -k_B T \ln Z = F_{\text{ideal}} - N k_B T \ln \left(1 + \frac{N}{2V} \int d^3 r f(r) \right) \quad (2.68)$$

我们计算压强

$$p = -\frac{\partial F}{\partial V} = \frac{N k_B T}{V} \left(1 + \frac{N}{2V} \int d^3 r f(r) \right) \quad (2.69)$$

这可以给我们第二位力系数 $B_2(T)$ 。

2.5.2 范德瓦耳斯状态方程

现在我们代入势能的具体形式。简单起见, 我们在这里只计算钢球势。

$$\begin{aligned}
 \int d^3 r f(r) &= \int_0^{r_0} d^3 r (-1) + \int_{r_0}^{+\infty} d^3 r \left(e^{+\beta U_0 (r_0/r)^6} - 1 \right) \\
 &\simeq -\frac{4}{3} \pi r_0^3 + \int_{r_0}^{\infty} 4\pi r^2 dr \beta U_0 r_0^6 / r^6 \\
 &= \frac{4\pi r_0^3}{3} \left(\frac{U_0}{k_B T} - 1 \right)
 \end{aligned}$$

代入，整理得到状态方程

$$\frac{pV}{Nk_B T} = 1 - \frac{N}{V} \left(\frac{2\pi r_0^3}{3} \frac{U_0}{k_B T} - \frac{2\pi r_0^3}{3} \right) \quad (2.70)$$

我们定义

$$a = \frac{2\pi r_0^3 U_0}{3}, \quad b = \frac{2}{3} \pi r_0^3 \quad (2.71)$$

进一步整理得到

$$\left(p + \frac{N^2 a}{V^2} \right) (V - Nb) = Nk_B T \quad (2.72)$$

上式就是大家喜闻乐见的范德瓦耳斯状态方程。注意这个方程只在密度很小、温度很高的情况下适用。

范德瓦耳斯方程中， a 可以解释为气体分子之间的吸引力，它贡献在压强的位置使压强减小。 b 正比于分子体积，由于分子占据一部分体积，所以实际得到的体积应该比理想气体更大。值得注意 $b = 4 \times$ 分子体积，这可以做如下的解释。注意从吸引变到排斥的结点是 r_0 ，因此每个钢球的半径是 $r_0/2$ ，每个钢球排斥周围的 $\Omega = \frac{4}{3}\pi r_0^3 = 2b$ 的空间那为什么在方程中排斥的空间是 b 而不是 $2b$ 呢？

我们一个一个来放入原子，第一个原子可占据的体积是 V ，第二个原子是 $V - \Omega$ ，第三个原子是 $V - 2\Omega$ 等等。在相空间中占据的总体积为

$$\begin{aligned} \frac{1}{N!} \prod_{m=1}^N (V - m\Omega) &= \frac{V^N}{N!} \prod_{m=1}^N \left(1 - \frac{m\Omega}{V} \right) \\ &= \frac{V^N}{N!} \left(1 - \frac{N(N-1)}{2} \frac{\Omega}{V} + \dots \right) \\ &\simeq \frac{V^N}{N!} \left(1 - \frac{N\Omega}{2V} \right)^N \\ &= \frac{1}{N!} \left(V - \frac{N\Omega}{2} \right)^N \end{aligned}$$

这便解释了为什么方程里面是 $N\Omega/2$ 。

要注意的是，似乎公式 2.69 可以计算任何势能，但是其实也有一些限制。从计算中可以发现，在三维空间中，只有当 $n \geq 4$ 时，积分才不会发散。对于发散的积分，不适合拿来当展开系数，这其中就包括了库伦势 $\sim 1/r$ 。

2.5.3 计算更多粒子的相互作用：集团展开

我们回到最初的弱相互作用展开

$$Z(N, V, T) = \frac{1}{N! \lambda^{3N}} \int \prod_i d^3 r_i \left(1 + \sum_{j>k} f_{jk} + \sum_{j>k, l>m} f_{jk} f_{lm} + \cdots \right) \quad (2.73)$$

现在我们要计算更多粒子的相互作用了。你也许会认为 $f_{jk} f_{lm}$ 这样的项就会贡献第三位力系数。但是实际上不是这样的，从之前的计算我们可以体会到，为什么第二项会贡献第二位力系数？这是因为在积分时给了我们 V^{N-1} ，那么我们要得到第三位力系数自然期望积分给我们 V^{N-2} ，例如形如

$$\int \prod_i d^3 r_i f_{12} f_{23} \quad (2.74)$$

这样的积分确实占用了两个 \vec{r}_i ，最终给我们 V^{N-2} ，但是下面这样的式子

$$\int \prod_i d^3 r_i f_{12} f_{23} f_{13} \quad (2.75)$$

计算这样的式子，也只需要积分掉两个坐标，所以我们看到，三个 f 的乘积也贡献第三位力系数。所以我们不能采取只看 f 的方法来计算位力系数，而是要采取下面一种方法——集团展开 (cluster expansion)。实际上可以证明，具有 l 个粒子的团 (后面的 f 要保证每一个粒子都有和至少另外一个粒子的相互作用) 所贡献的很多 f 的乘积，一定只和 $l-1$ 个相对坐标有关，所以最终会在积分中贡献 V^{N-l+1} ，从而给我们第 l 位力系数。所以我们应该按照成团的方式对配分函数进行计算。

我们可以用画图的方法直观地表达很多个 f 相乘，用点来表示粒子，用直线表示他们之间有相互作用。例如三粒子集团的积分可以表示为

$$\begin{aligned} U_3 &\equiv \int d^3 r_1 d^3 r_2 d^3 r_3 (f_{12} f_{23} + f_{13} f_{12} + f_{23} f_{13} + f_{12} f_{23} f_{13}) \\ &= \int d^3 r_1 d^3 r_2 d^3 r_3 \left(\begin{array}{c} \text{3} \\ \diagup \quad \diagdown \\ \text{1} \text{---} \text{2} \end{array} + \begin{array}{c} \text{3} \\ \diagdown \quad \diagup \\ \text{1} \text{---} \text{2} \end{array} + \begin{array}{c} \text{3} \\ \diagup \quad \diagdown \\ \text{1} \text{---} \text{2} \end{array} + \begin{array}{c} \text{3} \\ \diagup \quad \diagdown \\ \text{1} \text{---} \text{2} \end{array} \right) \end{aligned}$$

我们亲切的把上面的图案称之为“图”，用字母 G 表示。我们定义 l 个原子的集团的积分 (不要混淆这个 U_l 和之前的势 $U(r)$)

$$U_l \equiv \int \prod_i d^3 r_i \sum_{G \in l\text{-cluster}} G \quad (2.76)$$

现在，我们的野心可以大一点，我们直接求出 N 个粒子的配分函数。我们把 N 个粒子分成 m_l 个 l -cluster(具有 l 个粒子的团)。再对所有可能的 $\{m_l\}$ 求和。 m_l 和 l 满足

$$N = \sum_{l=1}^N m_l \cdot l \quad (2.77)$$

因此

$$Z(N, V, T) = \frac{1}{N! \lambda^{3N}} \sum_{\{m_l\}} \prod_l \frac{N!}{(l!)^{m_l} m_l!} U_l^{m_l} \quad (2.78)$$

读到这里，读者可能会有点晕了。没错，我第一次学的时候也晕了，读者这时候最好出去喝杯水，溜达一下，我们一定要把这里的每个式子的含义和来源搞清楚。

式中 $N!/(l!)^{m_l}$ 是典型地分堆问题的结果，额外的 $m_l!$ 是因为每个堆之间是全同的。所以 $\frac{N!}{(l!)^{m_l} m_l!} U_l^{m_l}$ 统计了 m_l 个 l -cluster 的积分。再把不同 l 乘起来，便是一个 $\{m_l\}$ 组合下的积分，再对于所有的 $\{m_l\}$ 求和便统计了所有可能的情况。

对上面的求和是十分困难的事情，因为你要满足归一化条件式 2.77。实际上，对于那么大的 N ，想要穷举 m_l 都是一件十分困难的事情。这时候怎么办？

是时候请出我们的老朋友——巨正则系综了，我们在这个讲义里面还没用过它。注意我在 1.5 节提到的，不同的系综理论只是以不同方式从三个量给出另外对偶的三个量。在巨正则系综中，我们最后求出的关系也可以假定 N 是固定的，所以使用什么系综和研究对象无关。在这里用巨正则系综纯粹是为了计算方便。

巨正则系综有什么好处呢？我们马上会看到。巨正则系综的配分函数为

$$\mathcal{Z}(\mu, V, T) = \sum_N e^{\beta \mu N} Z(N, V, T) \quad (2.79)$$

我们定义逸度 (fugacity) $z = e^{\beta \mu}$ ，于是有

$$\begin{aligned} \mathcal{Z}(\mu, V, T) &= \sum_N z^N Z(N, V, T) \\ &= \sum_N \frac{1}{\lambda^{3N}} \sum_{\{m_l\}} \prod_l \frac{N!}{(l!)^{m_l} m_l!} U_l^{m_l} \\ &= \sum_{m_l=0}^{\infty} \prod_{l=1}^{\infty} \left(\frac{z}{\lambda^3} \right)^{m_l l} \frac{1}{m_l!} \left(\frac{U_l}{l!} \right)^{m_l} \end{aligned}$$

将求和和积分的顺序交换，得到

$$\mathcal{Z} = \prod_{l=1}^{\infty} \exp \left(\frac{U_l z^l}{\lambda^{3l} l!} \right) \quad (2.80)$$

我们定义

$$b_l = \frac{\lambda^3}{V} \frac{U_l}{l! \lambda^{3l}} \quad (2.81)$$

注意这个 b_l 大致代表 l -cluster 对于配分函数的贡献，而 $b_1 = 1$, $U_1 = V$ 。所以我们有

$$\mathcal{Z}(\mu, V, T) = \prod_{l=1}^{+\infty} \exp \left(\frac{V}{\lambda^3} b_l z^l \right) = \exp \left(\frac{V}{\lambda^3} \sum_{l=1}^{\infty} b_l z^l \right) \quad (2.82)$$

我们可以看到，通过上面的操作，我们成功地把 \mathcal{Z} 写成了集团展开的形式。

我们来计算压强

$$\frac{pV}{k_B T} = \ln \mathcal{Z} = \frac{V}{\lambda^3} \sum_{l=1}^{\infty} b_l z^l \quad (2.83)$$

粒子数

$$\frac{N}{V} = -\frac{1}{V} \frac{\partial}{\partial \mu} \ln \mathcal{Z} = \frac{1}{\lambda^3} \sum_{l=1}^{\infty} l b_l z^l \quad (2.84)$$

注意上面计算过程和正则系综中的区别，在正则系综中我们直接就得到了压强，但是在这里我们却得到了压强是化学势的函数，而还要通过粒子数于化学势的关系推导出状态方程。可以看到，虽然我们在上面的求和过程中使用了巨正则系综来避免复杂的求和，但是求出来之后也是要面对复杂的关系。

所以我们得到状态方程

$$\frac{pV}{Nk_B T} = \frac{\sum_l b_l z^l}{\sum_l l b_l z^l} \quad (2.85)$$

现在，我们就是要代入 N 和 μ 的关系，注意， p 也依赖于 N (通过位力展开而依赖)。所以我们将 N 消掉，通过比较 z 的系数的方式可以求得位力系数。

整理上式得到

$$\begin{aligned}
 \left(\sum_{l=1}^{\infty} b_l z^l \right) &= \frac{p}{k_B T} \frac{V}{N} \left(\sum_{m=1}^{\infty} m b_m z^m \right) \\
 &= \left(\sum_{l=1}^{\infty} B_l \left(\frac{N}{V} \right)^l \right) \frac{V}{N} \left(\sum_{m=1}^{\infty} m b_m z^m \right) \\
 &= \left(\sum_{l=1}^{\infty} B_l \left(\frac{N}{V} \right)^{l-1} \right) \left(\sum_{m=1}^{\infty} m b_m z^m \right) \\
 &= \left[\sum_{l=1}^{\infty} \frac{B_l}{\lambda^{3(\lambda-1)}} \left(\sum_{n=1}^{\infty} n b_n z^n \right)^{l-1} \right] \sum_{m=1}^{\infty} m b_m z^m \\
 &= \left[1 + \frac{B_2}{\lambda^3} (z + 2b_2 z^2 + 3b_3 z^3 + \dots) + \frac{B_3}{\lambda^6} (z + 2b_2 z^2 + 3b_3 z^3 + \dots)^2 + \dots \right] \\
 &\quad \times [z + 2b_2 z^2 + 3b_3 z^3 + \dots]
 \end{aligned}$$

左右打开进行比较系数

$$z + b_2 z^2 + b_3 z^3 + \dots = z + \left(\frac{B_2}{\lambda^3} + 2b_2 \right) z^2 + \left(3b_3 + \frac{4b_2 B_2}{\lambda^3} + \frac{B_3}{\lambda^6} \right) z^3 + \dots$$

得到

$$\begin{aligned}
 B_2 = -\lambda^3 b_2 &= -\lambda^3 \frac{\lambda^3}{V} \frac{U_2}{2! \lambda^6} \\
 &= -\frac{1}{2V} \int d^3 r_1 d^3 r_2 f(\vec{r}_1 - \vec{r}_2) \\
 &= -\frac{1}{2} \int d^3 r f(\vec{r})
 \end{aligned}$$

这就是我们之前只考虑两个粒子相互作用的结果。同理可以得到

$$B_3 = \lambda^6 (4b_2^2 - 2b_3) \quad (2.86)$$

更高阶的计算，我们交给计算机。但是即使这样，系数 b_l 还需要很多对于图的积分。这些我们可能都无能为力，交给计算机就好。

更重要的是，我们从这里学到了一个系统计算相互作用微扰的方法——集团展开，而且我们能够做到任意价数。这种通过将相互作用展开，并且画图来追踪粒子之间相互作用的方法在量子场论中还有很重要的应用，那里的图被称作费曼图。

2.6 德拜-休克尔等离子体模型

我们再来看最后一个经典气体的例子。我们考虑单一成分的等离子体。等离子体每个电子可以自由运动，带有 $-q$ 的电量，还存在一背景电荷 $+q\rho$ ，它使得整体呈电中性。电子之间有相互作用能

$$U(r) = \frac{q^2}{r} \quad (2.87)$$

我们采取了一个奇怪的单位制 $4\pi\epsilon_0 = 1$ 。现在问题是，如果一个电子放在原点，要求解任意一点的电势，电势显然满足下面的泊松方程

$$\nabla^2\phi(\vec{r}) = -4\pi(-q\delta(\vec{r}) + q\rho - q\rho g(\vec{r})) \quad (2.88)$$

其中第一项是放在原点电子的贡献，第二项是均匀正电荷背景的贡献，第三项是其他电子的贡献，求解的一个很大的问题就是不知道 $g(\vec{r})$ 。如果电子气体的温度为 0，那么由于经典排斥作用，他们会互相离开，但是如果温度不是 0，内能会使他们彼此接近。一个电子接近原点需要消耗能量 $E(\vec{r}) = -q\phi(\vec{r})$ 。这个模型的精妙之处就在于，在原点附近，电子的分布可以用热力学来近似

$$g(\vec{r}) = e^{\beta q\phi(\vec{r})} \quad (2.89)$$

所以我们代入求解得到

$$\phi(\vec{r}) = -\frac{qe^{-r/\lambda_D}}{r} \quad (2.90)$$

上面这个势能很好地描述了等离子体中电势的分布情况，可见，相比于正常的库伦势能，多了一个指数衰减因子。着一个效应被称作 screening 效应， λ_D 被称作德拜 screening 长度。注意，上述推导的条件是在 λ_D 附近有大量的电子，即 $\rho\lambda_D^3 \gg 1$ 。

我希望你们不要只死记结果，适用于特殊例子的公式，你们要好好研究这些公式与原则赖以成立的条件，没有事实作为根据，公式只不过是精神垃圾。我们了解人的精神倾向是活动胜于思考，但精神上的劳作并不是思想，只有那些付出了巨大劳动的人们，才会获得应用的习惯。发现理解一个原则要比写出一个公式难得多。

麦克斯韦

3 量子气体

先告诉大家一个令人振奋的消息，本章我们完全不考虑粒子之间的相互作用！又因为量子体系都是由一系列离散的能级组成，而不考虑相互作用意味着加入一个新的粒子对于原来的能级无影响。在第二章的时候我们曾经提到，在量子力学中，原子的德布罗意波长很长，因此经典关于粒子的图像便失效了，取而代之的是量子化的粒子，即系统的微观态是能级上的“粒子”数⁶。

另外，值得指出的是，我们考虑的能级是离散的。在经典气体里面我们曾有一个不成文的规定，那就是坐标和动量都可以连续取值。实际上，他们并不是连续取值的。考虑气体被放在一个 L^3 的正方体盒子里面，可以用一个波函数来描述一个本征态⁷

$$\psi_{\vec{k}} = \frac{1}{\sqrt{V}} e^{i\vec{k} \cdot \vec{x}} \quad (3.1)$$

那么由于周期性边界条件 $\psi_{\vec{k}}(\vec{x}) = \psi_{\vec{k}}(\vec{x} + \vec{L})$ ， k 必须取如下离散的值

$$k_i = \frac{2\pi n_i}{L}, \quad n_i \in Z \quad (3.2)$$

当我们考虑经典极限 $L \gg \lambda_D$ 时（其中 λ_D 为德布罗意波长）， k_i 的间隔很小，可以近似认为它是连续的。在我们本章的大部分情况，我们也考虑这种近似，但是在低温的情况下， $\lambda_D \gtrsim L$ ，这个近似产生的问题会找上门来。

⁶注意！经典的粒子和量子的粒子虽然都是一个名词，但是本质并不一样！

⁷这里这样描述只是为了表述方便，实际上应该使用量子场论中的描述。



图 7: 量子系统的一个比喻，每一层书架相当于一个能级，每个方格相当于能级中的一个简并态，每个方格中的书就好像每个态上的粒子，每个方格自己形成一个独立系统。

3.1 量子气体的配分函数

毫不夸张的说，这一节是本章最重要的一节，甚至是全书最重要的一节，把这节弄明白，你可以说真的明白了系综理论。

前面说过，量子系统的能级是分立的，而又因为我们不考虑粒子之间的相互作用，所以加进来一些粒子不会使得系统的能级发生微扰（或者说微扰可以忽略）。我们反复强调也是最重要的一点是，经典的粒子和量子的粒子含义不同，因此计算配分函数的方式自然不同。

所以，我们应该这样看我们的系统：系统有很多个分立的能级，说一个系统处于什么微观态，就是说有多少个粒子在特定的能级 $|r\rangle$ 上。我们把系统以能级分成很多个子系统，如果一个粒子分布在这个系统上，则这个粒子具有能量 E_r 。我们之前推导过，整个系统的配分函数就是每个子系统配分函数的乘积。一个形象的比喻如图 7 所示。

3.1.1 正则系综

我们先考虑正则系综

$$\begin{aligned}
 Z &= \sum_{|\psi\rangle} e^{-\beta E_\psi} \\
 &= \sum_{\{n_r\}} e^{-\sum_r n_r \beta E_r} \\
 &= \prod_r \sum_{\{n_r\}} e^{-n_r \beta E_r}
 \end{aligned}$$

其中 n_r 表示在能级 $|r\rangle$ 上分布的粒子数有多少, 又被称作布居数。考虑到可能会有很多个 $|r\rangle$ 对应相同的能量, 而对于每个能量系统的求和方式又相同, 所以上式可以化成对能量求乘积

$$\begin{aligned}
 Z &= \sum_{|\psi\rangle} e^{-\beta E_\psi} \\
 &= \prod_E \left(\sum_{\{n_r\}} e^{-n_r \beta E} \right)^{g(E)}
 \end{aligned}$$

其中 $\sum_{\{n_r\}}$ 是对粒子数固定的情况使得 $\sum_r n_r = N$ 的 $\{n_r\}$ 的组合求和; $g(E)$ 被称作能级的简并度 (即一个能量对应于多少微观的能级), 或者叫做态密度, 是指一个确定的能量下有多少个微观态 (简并态)⁸⁹。或者按照上面的比喻说, $g(E)$ 就是 E 那么高的层有多少个格子。注意, 这里的能级 E_r 中的 r 不写了, 因为它们是有一定简并的。所以这个记号乱用行为经常引起人们困惑, 因为通常总能量也用 E 表示。

我们经常要使用配分函数的对数

$$\ln Z = \sum_E g(E) \ln \left(\sum_{\{n_r\}} e^{-n_r \beta E} \right) \quad (3.3)$$

一般来说, 我们只考虑能级非常密的情况, 因此上面的积分应该转换成求和。 $g(E)$ 变成从 $E \rightarrow E + dE$ 区间内的简并度, 即

$$\sum_E g(E) \rightarrow \int g(E) dE \quad (3.4)$$

⁸ “简并” 在英文中是 “degeneracy”, 他们是完全不同的态, 但是能量恰好相同。

⁹ 再次强调!! E 是能级!!

因此配分函数通常写作

$$\ln Z = \int dE g(E) \ln \left(\sum_{\{n_r\}} e^{-n_r \beta E} \right) \quad (3.5)$$

系统的能量

$$\begin{aligned} \langle E \rangle &= -\partial_\beta \ln Z \\ &= \int dE g(E) \frac{n_r E}{\sum_{\{n_r\}} e^{-n_r \beta E}} \\ &= \int dE E g(E) \langle n_r \rangle \end{aligned}$$

其中 $\langle n_r \rangle$ 是在一个子系统 $|r\rangle$ 中的平均粒子数。上式是符合物理直觉的，平均能量就等于每个能级上的平均能量再对能级求和。在很多教材中，会在求出 $\langle n_r \rangle$ 后直接使用上面的式子，我认为会使初学者感到迷惑。我们非常详细地验证了它的自洽性。

3.1.2 巨正则系综

从上面正则系综的推导中就感受得到，对 $\sum_{\{n_r\}}$ 求和基本是一件不可能的事情。读者也可以从生活经验发现，整理书架是一件困难的事情，怪不得我们的图书馆需要那么多学生助理！因此，不固定总粒子数，让每一个能级上面的粒子数无限制，对于计算来说是很方便的事情。

巨正则系综的配分函数

$$\begin{aligned} \mathcal{Z} &= \sum_{|\psi\rangle} e^{-\beta(E_\psi - \mu n_\psi)} \\ &= \sum_{n_r} e^{-\sum_r n_r \beta(E_r - \mu)} \\ &= \prod_r \sum_{n_r} e^{-n_r \beta(E_r - \mu)} \\ &= \prod_E \left(\sum_{n_r} e^{-n_r \beta(E - \mu)} \right)^{g(E)} \end{aligned}$$

求配分函数的对数

$$\ln \mathcal{Z} = \sum_E g(E) \ln \left(\sum_{n_r} e^{-n_r \beta(E - \mu)} \right) \quad (3.6)$$

引入能级连续化假设

$$\sum_E g(E) \rightarrow \int g(E) dE \quad (3.7)$$

配分函数写作

$$\ln \mathcal{Z} = \int dE g(E) \ln \left(\sum_{n_r} e^{-n_r \beta (E - \mu)} \right) \quad (3.8)$$

系统的能量

$$\begin{aligned} \langle E \rangle &= -\partial_\beta \ln \mathcal{Z} \\ &= \int dE g(E) \frac{n_r E}{\sum_{n_r} e^{-n_r \beta (E - \mu)}} \\ &= \int dE E g(E) \langle n_r \rangle \end{aligned}$$

系统的粒子数

$$\begin{aligned} \langle N \rangle &= \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln \mathcal{Z} \\ &= \int dE n_r \frac{1}{\sum_{n_r} e^{-n_r \beta (E - \mu)}} \\ &= \int dE g(E) \langle n_r \rangle \end{aligned}$$

正如我们所期望的。

3.2 态密度

我们下面来求一些常见的简单系统的态密度。实际上，清楚了态密度的定义，下面的计算过程就十分简单轻松了，这一节的重要性其实不如上一节。

假如微观系统告诉我们 $|\vec{k}\rangle$ 就是一个态¹⁰，那么要求出能级 $E + dE$ 的微观态数目，就是要求出能级 $E \rightarrow E + dE$ 所对应的 k 空间的体积所对应的态数目，故有

$$\sum_E g(E) \simeq \frac{L^3}{(2\pi)^2} \int d^3k = \int g(E) dE \quad (3.9)$$

注意 k 的取值实际上是离散的，而在 k 空间平均 $(2\pi)^3/L^3$ 的体积中有一个 k ，所以应该除去 k 的密度，从而得到前面的系数。

在正则系综中，上述求和还要满足总粒子数 N 固定，即每个能级上的粒子数加起来要等于 N ，即 $\sum_r n_r = N$ ，所以我们一般在巨正则系综中考虑问题比较方便。

3.2.1 非相对论系统

我们现在来把态密度表示成对于能量的求和，我们先考虑一个最简单的情况

$$E = \frac{\hbar^2 k^2}{2m} \quad (3.10)$$

这个时候等能面就是 k 相同的面，因此

$$\begin{aligned} \sum_E g(E) &\simeq \frac{V}{(2\pi)^3} \int_0^\infty 4\pi k^2 dk \\ &= \int_0^\infty \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \sqrt{E} dE \end{aligned}$$

从 $E \rightarrow E + dE$ 的态的数目为

$$g(E) = \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \sqrt{E} \quad (3.11)$$

3.2.2 相对论系统

我们再来计算相对论的系统

$$E = \sqrt{\hbar^2 k^2 c^2 + m^2 c^4} \quad (3.12)$$

¹⁰并不总是这样，例如在这里我们就没有考虑自旋带来的简并

读者可以计算得到

$$\sum_n = \frac{V}{(2\pi)^3} \int 4\pi k^2 dk = \int g(E) dE \quad (3.13)$$

其中

$$g(E) = \frac{VE}{2\pi^2\hbar^3c^3} \sqrt{E^2 - m^2c^4} \quad (3.14)$$

对于静止质量为 0 的粒子，态密度为

$$g(E) = \frac{VE^2}{2\pi^2\hbar^3c^3} \quad (3.15)$$

我们上面考虑的 E 只和 $|\vec{k}|$ 有关，因此 k 空间中的等能面是球面。在一般情况下， E 和 \vec{k} 有复杂的依赖关系，等能面可能是奇形怪状的。

有些文献中会将体积除去，而将剩下的记作 $g(E)$ ，是单位体积的态密度。

3.3 光子气体

我们来分析第一个例子，光子气体。光子气体可以取的能级为

$$E = \hbar\omega \quad (3.16)$$

它的态密度为

$$g(E) = 2 \times \frac{VE^2}{2\pi^2\hbar^3c^3} \quad (3.17)$$

其中 2 的系数是考虑了光子的自旋 (或偏振)，对于相同 \vec{k} ，可以有两种偏振模式¹¹。另外我们要注意，光子数是不守恒的。如果你在一个盒子里放几个原子，那么过了几个月这几个原子可能不会消失，但是光子就不一定了，容器壁会吸收或者发射光子，吸收数和发射数也不一定相等。或者，一个最简单在我们的世界验证光子是不守恒的方法就是：关灯。有的时候也说，因为光子数不守恒，所以产生或者湮灭一个光子不会消耗能量，所以光子的化学势为 0，所以对于光子气体，在正则系综和巨正则系综计算都一样。因为你根本不用考虑粒子数恒定的事情了。

所以，任何能级 (特定波矢和偏振) 上能容纳的粒子可以有无限多个，这些无限多个态就是正则系综中可以达到的态了。所以我们有配分函数 (这里我们采用了正则系综)

$$\ln Z = \ln \left(\sum_{\psi} e^{-\beta E_{\psi}} \right) = \int dE g(E) \ln \left(\sum_n e^{-\beta n \hbar \omega} \right) \quad (3.18)$$

我们作一下换元

$$g(E)dE = g(\omega)d\omega = \frac{V\omega^2}{\pi^2c^3}d\omega \quad (3.19)$$

代入配分函数，我们得到

$$\begin{aligned} \ln Z &= \int_0^{\infty} d\omega g(\omega) \ln \left(\sum_n e^{-\beta n \hbar \omega} \right) \\ &= \int_0^{\infty} d\omega g(\omega) \ln (1 + e^{-\beta \hbar \omega} + e^{-2\beta \hbar \omega} + \dots) \\ &= -\frac{V}{\pi^2c^3} \int_0^{\infty} d\omega \omega^2 \ln(1 - e^{-\beta \hbar \omega}) \end{aligned}$$

¹¹这里严谨的表述应该从光场量子化出发，要引入两个产生湮灭算符。

进一步求得能量

$$\begin{aligned} E &= -\frac{\partial}{\partial \beta} \ln Z = \frac{V}{\pi^2 c^3} \int_0^\infty d\omega \, \omega^2 \frac{e^{-\beta \hbar \omega}}{1 - e^{-\beta \hbar \omega}} \hbar \omega \\ &= \frac{V \hbar}{\pi^2 c^3} \int_0^\infty d\omega \frac{\omega^3}{e^{\beta \hbar \omega} - 1} \end{aligned}$$

上述能量由每个能级 ω 上的能量加起来贡献的，每个能级上的能量为

$$f(\omega) d\omega = \frac{V \hbar}{\pi^2 c^3} \frac{\omega^3}{e^{\beta \hbar \omega} - 1} d\omega \quad (3.20)$$

上面这个能量分布被称作普朗克分布。这曾经是物理学大厦上的一朵乌云，普朗克大胆的提出假设 $E = \hbar \omega$ ，解决了这个问题，也开启了人类探索量子力学的征程。

通过

$$\frac{df(\omega)}{d\omega} = 0 \quad (3.21)$$

求得发出能量最大的频率为

$$\omega_{\max} = \zeta \frac{k_B T}{\hbar} \quad (3.22)$$

上面被称作韦恩位移定律，其中 $3 - \zeta = 3e^{-\zeta}$ ，得到 $\zeta = 2.822$ 。

我们来对所有的频率积分，得到总能量，我们令 $x = \beta \hbar \omega$ ，有

$$\begin{aligned} E &= \frac{V \hbar}{\pi^2 c^3} \int_0^\infty \frac{dx}{\beta \hbar} \frac{(x/\beta \hbar)^3}{e^x - 1} \\ &= \frac{V}{\pi^2 c^3} \frac{(k_B T)^4}{\hbar^3} \int_0^\infty \frac{x^3 dx}{e^x - 1} \end{aligned}$$

我们以后会证明

$$\int_0^\infty \frac{x^3 dx}{e^x - 1} = \Gamma(4) \zeta(4) = \frac{\pi^4}{15} \quad (3.23)$$

所以我们得到能量密度

$$\mathcal{E} = \frac{E}{V} = \frac{\pi^2 k_B^4}{15 \hbar^3 c^3} T^4 \quad (3.24)$$

我们可以看到总能量密度正比于温度的四次方。这和下面的 Stefan-Boltzmann 定律有着联系：

$$\text{Energy Flux} = \frac{\mathcal{E} c}{4} = \sigma T^4 \quad (3.25)$$

其中 $\sigma = \pi^2 k_B^4 / 60 \hbar^3 c^2 = 5.67 \times 10^{-8} \text{ J} \cdot \text{s}^{-1} \cdot \text{m}^{-2} \cdot \text{K}^{-4}$ ，称作 Stefan-Boltzmann 常数。

3.3.1 Stefan-Boltzmann 定律

那么, 系数 $1/4$ 是哪来的呢? 这是因为我们只考虑一个方向接受到的辐射通量。对于一般具有各项同性分布 $f(\vec{v})$ 的气体, 我们考虑在容器壁上开一个小孔, 我们将小孔的法线指向外的方向定义为 z 方向。则从小孔溢出的能量为

$$\text{溢出的能量} = \langle v_z^+ \rangle \times \text{面积} \times \text{时间} \times \mathcal{E} \quad (3.26)$$

我们现在要求 $\langle v_z^+ \rangle$, 它是 v_z 在 $v_z > 0$ 空间的平均值, 这被称作泻流速率。计算得到

$$\begin{aligned} \langle v_z^+ \rangle &= \frac{1}{4\pi} \int_{v_z > 0} 4\pi v^2 dv d\Omega f(\vec{v}) v \cos \theta \\ &= \frac{1}{4\pi} \int_0^{2\pi} d\varphi \int_0^{\pi/2} \sin \theta \cos \theta \int_0^\infty 4\pi v^2 f(\vec{v}) v \cos \theta \\ &= \frac{1}{4} \langle v \rangle \end{aligned}$$

其中 $\langle v \rangle$ 是气体分子的平均速率。对于我们的光子气体, 显然有

$$\langle v_z^+ \rangle = \frac{c}{4} \quad (3.27)$$

所以得到的能流 (单位面积、单位时间内通过小孔的能量) 应该是 $\mathcal{E}c/4$ 。

3.3.2 辐射压强

我们来计算光子的辐射压强。光子气体的自由能

$$\begin{aligned} F &= -k_B T \ln Z \\ &= \frac{V k_B T}{\pi^2 c^3} \int_0^\infty d\omega \omega^2 \ln(1 - e^{-\beta \hbar \omega}) \end{aligned}$$

这个积分不是很好计算, 我们进行分部积分

$$\begin{aligned} \int_0^\infty d\omega \omega^2 \ln(1 - e^{-\beta \hbar \omega}) &= \int_0^\infty d\left(\frac{1}{3}\omega^3\right) \ln(1 - e^{-\beta \hbar \omega}) \\ &= \frac{1}{3}\omega^3 \ln(1 - e^{-\beta \hbar \omega}) \Big|_0^\infty - \int_0^{+\infty} \frac{1}{3}\omega^3 \frac{e^{-\beta \hbar \omega} \beta \hbar}{1 - e^{-\beta \hbar \omega}} d\omega \end{aligned}$$

这又出现了我们熟悉的积分 (Be patient! 我们后面会证明的), 我们进行换元 $x = \hbar\omega/k_B T$, 求得自由能

$$\begin{aligned} F &= -\frac{V k_B T}{3\pi^2 c^3} \frac{\hbar}{k_B T} \int_0^\infty \frac{\omega^3 d\omega}{e^{\beta\hbar\omega} - 1} \\ &= -\frac{V}{3\pi^2 c^3 \hbar^3} (k_B T)^4 \int_0^\infty \frac{x^3 dx}{e^x - 1} \\ &= -\frac{V \pi^2}{45 c^3 \hbar^3} (k_B T)^4 \end{aligned}$$

所以我们可以求解得到由于电磁辐射所产生的压强

$$p = -\left(\frac{\partial F}{\partial V}\right)_T = \frac{\pi^2 (k_B T)^4}{45 c^3 \hbar^3} = \frac{4\sigma}{3c} T^4 = \frac{E}{3V} \quad (3.28)$$

所以我们得到了一个十分重要的结果

$$p = \frac{1}{3} \mathcal{E} \quad (3.29)$$

上式又被称作光子的状态方程, 这个光子的状态方程在宇宙学中非常重要。最后的最后, 我们还可以顺带求一下熵和热容。光子气体的熵为

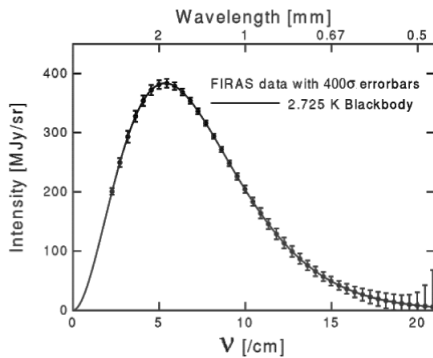
$$S = \left(\frac{\partial F}{\partial T}\right)_V = \frac{16V\sigma}{3c} T^3 \quad (3.30)$$

热容为

$$C_V = \left(\frac{\partial E}{\partial T}\right)_V = \frac{16V\sigma}{c} T^3 \quad (3.31)$$

3.3.3 宇宙微波背景辐射

宇宙微波背景辐射 (Cosmic Microwave Background, or CMB), 是在“宇宙大爆炸”后, 光子和其他物质脱离相互作用后, 传播到今天被人们所看到的微波辐射。在那种老式电视机没有信号的“雪花”画面中, 竟有相当一部分是宇宙微波背景辐射贡献的。在 1999 年人们使用了 COBE 卫星来测量宇宙大爆炸的频谱, 发现和我们上面的普朗克分布几乎完全一致! 如图 8 所示



在图中, 误差棒被放大了 400 倍! 否则你根本看不到那个误差棒! 令人感觉十分激动的是, 光几乎传播了 137 亿年, 而它的规律和我们上面简单的推导一样。

图 8: COBE 卫星测量宇宙微波背景辐射的数据

3.3.4 量子力学的诞生

如果取高温极限, 即 $\beta\hbar\omega \ll 1$, 则普朗克分布的公式近似为

$$f(\omega) = \frac{V}{\pi^2 c^3} k_B T \quad (3.32)$$

这被称作瑞利-金斯公式, 可以从经典的电磁波理论推导出来。但是我们看到, 瑞利-金斯公式给出的总能量是发散的, 在历史上这被称作紫外灾难。而实际上, 高频模式不是有一点点能量就可以激发的, 而是有一个激发高频模式的最低能量 $\hbar\omega$, 这完全是量子力学的性质, 而就是这个黑体辐射问题, 使得人们开始逐渐意识到量子力学的存在。

3.4 声子

现在我们来研究晶体。晶体是由很多原子规则地排列在周期性排列的格点上。简单起见，我们考虑一维单原子链，在平衡位置，每个原子之间的间距为 a ，我们只考虑原子之间有二阶相互作用势，哈密顿量写为

$$H = \frac{1}{2}m \sum_i \dot{u}_i^2 + \frac{\alpha}{2} \sum_i (u_i - u_{i+1})^2 \quad (3.33)$$

利用哈密顿正则方程

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \quad (3.34)$$

得到运动方程

$$\ddot{u}_i = \frac{\alpha}{m} (2u_i - u_{i+1} - u_{i-1}) \quad (3.35)$$

我们使用离散傅里叶变换

$$u_l = \frac{1}{\sqrt{N}} \sum_k \tilde{u}_k e^{i(kla - \omega_k t)} \quad (3.36)$$

代入上面的运动方程，得到色散关系

$$\omega_k = 2\sqrt{\frac{\alpha}{m}} \left| \sin\left(\frac{ka}{2}\right) \right| \quad (3.37)$$

我们把固体系统分成以 ω 区分的无数个子系统，则我们有

$$E(k) = \hbar\omega_k \quad (3.38)$$

则配分函数

$$\ln Z = \frac{L^3}{(2\pi)^3} \int d^3k \ln \left(\sum_n e^{-n\beta\hbar\omega} \right) \quad (3.39)$$

3.4.1 德拜模型

我们考虑三维的固体，代入上面的色散关系会使得计算变得十分繁琐，而且和我们下面的计算差别很小，所以我们对色散关系采用下面的近似

$$\omega_k = c|k| \quad (3.40)$$

即原来的 \sin 的色散关系用直线来近似。其中 c 是固体中的声速。我们将 k_i 的取值范围取在 $[-\frac{\pi}{a}, \frac{\pi}{a}]$ 的范围才是有意义的。因为若做变换 $k \rightarrow k + 2\pi/a$ 从离散傅里叶变换可以发现，得到的解是完全等价的。上述范围又被叫做第一

布里渊区。对于三维情况，布里渊区就是一个三维的立方体， k 积分区域就是第一布里渊区。这里我们做一个近似，因为 ω 只依赖于 k 的大小而不依赖于 k 的方向，所以将积分区域近似为球形是十分方便的。我们把积分频率改为半径为 ω_D 的球，并保证球内的 \vec{k} 的数目和第一布里渊区的 \vec{k} 的数目相同。由于我们采用的声子的色散关系和光子的相同，所以声子的态密度为

$$g(\omega) = \frac{3V}{2\pi^2 c^3} \omega^2 \quad (3.41)$$

所以半径 ω_D 内的 \vec{k} 的数目为

$$\int_0^{\omega_D} g(\omega) d\omega = \frac{V \omega_D^3}{2\pi^2 c^3} \quad (3.42)$$

第一布里渊区的 \vec{k} 的数目为 $3N$ ，所以我们得到德拜频率

$$\omega_D = c \left(\frac{6\pi^2 N}{V} \right)^{1/3} \quad (3.43)$$

注意到我们做了如此多的近似，所以上式对于实际的固体一般是不成立的。但是我们将会看到，德拜模型的意义在于，可以大致的描绘固体热容的行为。如果分析实际的固体，在计算机科学发达的 21 世纪，我们应该用严格的公式，并诉诸于数值计算。

德拜频率 ω_D 在我们的模型里面就是固体中可以传播的最大频率。我们定义这个频率所对应的温度，称之为德拜温度

$$k_B T_D = \hbar \omega_D \quad (3.44)$$

对于软的材料，如铅， $T_D \sim 100$ K。对于硬的材料， $T_D \sim 2000$ K。对于大部分材料，德拜温度都在 100 K 左右。

我们来计算配分函数

$$\ln Z_{\text{phonon}} = \int_0^{\omega_D} d\omega g(\omega) \ln \left(\sum_n e^{-n\beta\hbar\omega} \right) \quad (3.45)$$

代入得到

$$\ln Z_{\text{phonon}} = \int_0^{\omega_D} \frac{3V}{2\pi^2 c^3} \omega^2 d\omega \ln(1 - e^{-\beta\hbar\omega}) \quad (3.46)$$

我们计算能量为

$$\begin{aligned} E &= -\partial_\beta \ln Z_{\text{phonon}} \\ &= \frac{3V\hbar}{2\pi^2 c^3} \frac{(k_B T)^4}{\hbar^3} \int_0^{T_D/T} \frac{x^3}{e^x - 1} dx \end{aligned}$$

在上面的计算过程中我们进行了换元 $x = \beta\hbar\omega$ 。上面的积分公式是没有解析解的。我们可以通过研究两个重要极限来研究总能量的性质。当温度很低的时候 $T \ll T_D$ 时, 我们有

$$\int_0^\infty \frac{x^3}{e^x - 1} dx = \frac{\pi^4}{15} \quad (3.47)$$

所以在低温区域

$$E = \frac{\pi^4}{15} \frac{3V}{2\pi^2 c^3} \frac{(k_B T)^4}{\hbar^2} \quad (3.48)$$

热容为

$$C_V = \left(\frac{\partial E}{\partial T} \right)_V \sim T^3 \quad (3.49)$$

这一性质与实验符合的非常好。我们再来看高温的情况, 由于 $T \gg T_D$, 所以我们有 $x \ll 1$, 所以可以对积分做展开

$$\frac{x^3}{e^x - 1} = x^2 + \dots \quad (3.50)$$

我们求得热容

$$C_V = 3Nk_B \quad (3.51)$$

这一结果回到了经典的能均分定理, 这个结果被称作 Dulong-Petit 定律。一个固体热容的实验数据如图所示。其中德拜模型是实线, 而高温部分的差异是因为这个实验测量的是定压比热。

3.5 回顾双原子分子热容

我们这一小节来解决我们经典气体中没有解决的问题, 即双原子分子气体的热容问题。注意, 我们这一节计算配分函数的策略仍然是

$$Z = \frac{Z_1^N}{N!}$$

这是因为我们考虑的高温极限, 每个分子的能级是彼此独立的。所以我们来计算单粒子的转动配分函数。在量子力学中, 转动部分使用哈密顿量

$$H = \frac{\hbar^2}{2I} j(j+1) \quad (3.52)$$

考虑到对于角动量 j 有 $2j+1$ 个简并态, 所以配分函数是

$$Z_{\text{rot}} = \sum_{j=0}^{+\infty} (2j+1) e^{-\beta \hbar^2 j(j+1)/2I} \quad (3.53)$$

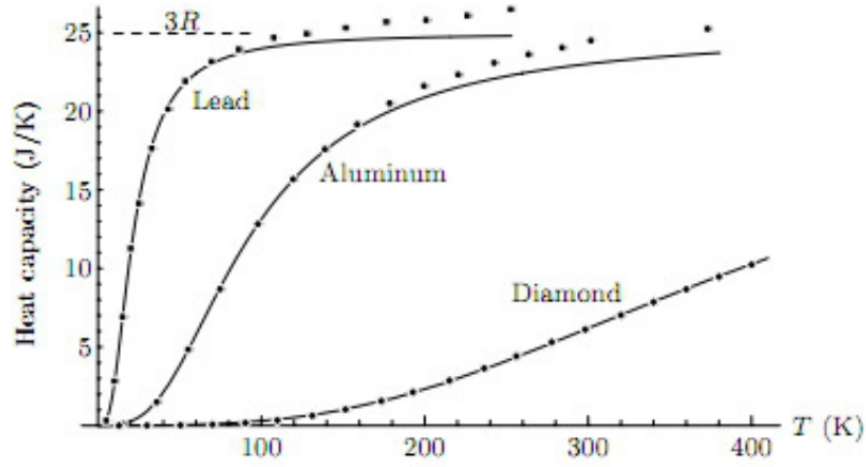


图 9: 固体热容的实验结果与德拜模型的比较

我们考虑两个极限，第一个是高温极限 $T \gg \hbar^2/2Ik_B$ ，这时能级十分密集，求和可以用积分代替

$$Z_{\text{rot}} \simeq \int_0^\infty dx (2x+1) e^{-\beta \hbar^2 x(x+1)/2I} = \frac{2I}{\beta \hbar^2} \quad (3.54)$$

和经典的情况是一样的。而当 $T \ll \hbar^2/2Ik_B$ 时，这些能级的间距变得非常大，因此对于所有的 $j \neq 0$ 项，对于配分函数的贡献都可以忽略不计，只有 $j = 0$ 有主要的贡献，这时

$$Z_{\text{rot}} \simeq 1 \quad (3.55)$$

这时去求能量则几乎不会贡献热容。也就是说因为量子的能级是离散的，从而能量不能够使得其激发到激发态，只能待在基态。我们把这种现象形象地称作“自由度冻结”。

同样的，我们也可以用这种方法来分析转动自由度，转动自由度量量子化之后能级可以表述为

$$E = \left(n + \frac{1}{2}\right) \hbar \omega \quad (3.56)$$

我们来计算一个粒子的转动配分函数

$$\begin{aligned} Z_{\text{vib}} &= \sum_n e^{-\beta\hbar\omega(n+1/2)} \\ &= e^{-\beta\hbar\omega/2} \sum_n e^{-\beta\hbar\omega n} \\ &= \frac{1}{2 \sinh(\beta\hbar\omega/2)} \end{aligned}$$

当高温情况时 $\beta\hbar\omega \gg 1$, 我们有

$$Z_{\text{vib}} = \frac{1}{\beta\hbar\omega} \quad (3.57)$$

这和经典的结果一致。对于低温情况, 则有

$$Z_{\text{vib}} = e^{-\beta\hbar\omega/2} \quad (3.58)$$

这一项只贡献零点能

$$E = -\partial_{\beta} \ln Z = \frac{\hbar\omega}{2} \quad (3.59)$$

而这一项为常数, 不会对热容产生影响。所以可以看到在低温下振动自由度也被冻结。

3.6 玻色统计

从这一节开始我们来研究玻色子和费米子的统计。其实在前面光子气体的过程中, 光子其实就是一种玻色子。从理论上, 这些粒子都是由量子场论中“场的激发态”来描述的。我们可以不必知道具体描述的细节, 而知道下面的事实。

1. 玻色子和费米子都是量子化意义的粒子。当我们说“能级上的粒子数”时, 我们的含义是能级上的“占据数态”¹² $|n\rangle$ 所对应的 n 是多少。
2. 玻色子是具有整数自旋的粒子, 而费米子是具有半整数自旋的粒子。
3. 在一个能级上可以存在多个玻色子。就像前面光子的粒子一样, 一个 ω 能级上可以存在任意多个能量为 $\hbar\omega$ 的粒子。
4. 费米子服从泡利不相容原理, 即一个能级上只能容纳一个粒子。正如读者从原子物理中学到的那样。

¹²是一个量子态

我们即将看到，在高温极限，玻色统计和费米统计都将回到经典的波尔兹曼统计。而在低温极限，玻色子和费米子表现的就很不同了。玻色子所有的粒子都会趋向于基态，而费米子由于泡利不相容原理，则会在相空间填满一个面，叫做费米面。

组成我们自然界的粒子，如电子、质子和中子都是费米子。但是多个费米子的行为表现也可以像一个玻色子。如果能量不够高，能量便不足以赶出里面的费米子。实际上我们知道，质子、中子、电子等也不是基本粒子，而是由更基本的夸克组成的，夸克是费米子。如果构成我们世界的物理规律不同，如果4个夸克组成一个束缚态的话，那么质子和中子就会成为玻色子，它们之间没有泡利不相容原理的排斥，原子核将不复存在，整个世界也不复存在了。

3.6.1 玻色-爱因斯坦统计

我们先不说明我们研究的能级是什么样子的，我们来做一个一般的讨论。根据前面的论述，我们使用巨正则系综，把每一个能级看作一个子系统，对于能级 $|r\rangle$ ，我们有配分函数

$$\mathcal{Z}_r = \sum_{n_r=0}^{+\infty} e^{-\beta n_r (E_r - \mu)} = \frac{1}{1 - e^{-\beta(E_r - \mu)}} \quad (3.60)$$

其中我们用到了一个能级上可以有多个玻色子的性质。注意到如果上面的级数求和是收敛的，我们必须有 $E_r - \mu > 0$ 。如果我们取基态 $E_0 = 0$ ，那么如果我们要求级数收敛，我们必须取

$$\mu < 0 \quad (3.61)$$

总的配分函数是所有能级的乘积

$$\mathcal{Z} = \prod_r \frac{1}{1 - e^{-\beta(E_r - \mu)}} \quad (3.62)$$

我们可以求得每个能级上的平均粒子数

$$\langle n_r \rangle = \frac{1}{\beta} \ln \mathcal{Z} = \frac{1}{e^{\beta(E_r - \mu)} - 1} \quad (3.63)$$

这便是玻色-爱因斯坦分布，它描述了玻色分布能级上的平均粒子数。总粒子数

$$N = \sum_r \frac{1}{e^{\beta(E_r - \mu)} - 1} \quad (3.64)$$

为了方便，我们引入符号

$$z = e^{\beta\mu} \quad (3.65)$$

被称作逸度 (fugacity)，因为 $\mu < 0$ ，所以我们有 $0 < z < 1$ 。

3.6.2 理想玻色气体

对于理想玻色气体，我们取色散关系

$$E = \frac{\hbar^2 k^2}{2m} \quad (3.66)$$

我们曾经求得过态密度

$$g(E) = \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} E^{1/2} \quad (3.67)$$

回忆态密度是对求和做转换 $\sum_r \rightarrow \int g(E) dE$ 。因此，总粒子数应该改写为

$$N = \int dE \frac{Eg(E)}{z^{-1}e^{\beta E} - 1} \quad (3.68)$$

可以看出 $N = N(\mu, T)$ ，但是在第一章曾经说过，通常我们使用巨正则系综，但是我们还是希望粒子数固定，而使用巨正则系综只是为了计算简单。所以最后，为了将最后的结果表示成粒子数和温度的形式，我们的代价是将上式求反函数，得到

$$\mu = \mu(N, T) \quad (3.69)$$

本来在巨正则系综中 μ 和 T 是独立的变量。但是现在因为要强制固定粒子数，所以他们不再是独立的变量了，而现在把 N 和 T 当作独立的变量。因为化学势对人们来说还是太抽象和陌生。你也许已经在理想气体的一个练习中尝试过一个例子了。由之前的论述，平均能量

$$E = \int dE g(E) \langle n_r \rangle = \int dE \frac{Eg(E)}{z^{-1}e^{\beta E} - 1} \quad (3.70)$$

如果读者忘了这个式子怎么来的，最好看一下前面。再次指出的是，不要把等式左边的 E 和右边被积分掉的 E 弄混了，左边的 E 是系统的总平均能量，而右面是系统的能级。我们进而可以计算压强

$$pV = \frac{1}{\beta} \ln \mathcal{Z} = -\frac{1}{\beta} \int dE g(E) \ln (1 - ze^{-\beta E}) \quad (3.71)$$

利用分部积分且 $g(E) \sim E^{1/2}$, 我们有

$$\begin{aligned}
 -\frac{1}{\beta} \int dE g(E) \ln(1 - ze^{-\beta E}) &= -\frac{1}{\beta} \int d \left(\frac{2}{3} g(E) dE \right) \ln(1 - ze^{-\beta E}) \\
 &= -\frac{1}{\beta} \left[\frac{2}{3} g(E) E \ln(1 - ze^{-\beta E}) \right]_0^{+\infty} - \int_0^{+\infty} \frac{2}{3} g(E) E \frac{-z\beta e^{-\beta E}}{1 - ze^{-\beta E}} dE \\
 &= \int_0^{\infty} \frac{2}{3} E g(E) \frac{1}{z^{-1} e^{-\beta E} - 1} dE \\
 &= \frac{2}{3} E
 \end{aligned}$$

我们没有直接将状态方程算出来, 要获得状态方程的具体关系, 我们首先要将 3.70 积分出来, 得到 $E = E(\mu, V, T)$, 再将反解得到的 $\mu = \mu(N, T, V)$ 联立, 才能得到 $p = p(N, V, T)$ 。

3.6.3 高温玻色气体

实际上上面那些积分都很难计算, 为了简单, 我们先考虑高温极限, 即 $\beta \rightarrow 0$, 即 $e^{\beta\mu} \rightarrow 0$ 。我们先试图反解 μ , 因此我们计算 N

$$\begin{aligned}
 N &= \int dE \frac{g(E)}{z^{-1} e^{\beta E} - 1} \\
 &= \int dE \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} E^{1/2} \frac{1}{z^{-1} e^{\beta E} - 1}
 \end{aligned}$$

我们进行换元 $x = \beta E$, 进行一下计算, 得到

$$\frac{N}{V} = \frac{z}{4\pi^2} \left(\frac{2m}{\beta \hbar^2} \right)^{3/2} \int_0^{\infty} dx \sqrt{x} e^{-x} \frac{1}{1 - ze^{-x}}$$

由于我们考虑的是高温极限, $x \ll 1$, 所以我们将上式展开

$$\frac{N}{V} = \frac{z}{4\pi^2} \left(\frac{2m}{\beta \hbar^2} \right)^{3/2} \int_0^{+\infty} dx \sqrt{x} e^{-x} (1 + ze^{-x} + \dots)$$

做替换 $x = u^2$, 上面的积分就变成了高斯积分, 保留前几项得到

$$\frac{N}{V} = \frac{z}{\lambda^3} \left(1 + \frac{z}{2\sqrt{2}} + \dots \right) \quad (3.72)$$

因为 $z \ll 1$, 所以从上面的式子可以得到 $\lambda^3 N/V \ll 1$, 这便符合了经典近似。因此, 我们有

$$z(1 + \frac{z}{2\sqrt{2}} + \dots) = \frac{\lambda^3 N}{V} \quad (3.73)$$

我们希望从 $N = N(\mu, V, T)$ 中解出 $\mu = \mu(N, V, T)$, 我们知道 $\alpha = N\lambda^3 N/V$ 是小量, 所以我们想到把 z 展开

$$z = z(N=0) + \frac{\partial z}{\partial \alpha} \alpha + \frac{1}{2!} \frac{\partial^2 z}{\partial \alpha^2} \alpha^2 + \dots \quad (3.74)$$

其中导数项由式 3.73 给出。我们得到

$$z = \frac{\lambda^3 N}{V} \left(1 - \frac{1}{2\sqrt{2}} \frac{\lambda^3}{N} + \dots \right) \quad (3.75)$$

如此一来, 在这个高温极限下, 我们就反解出了 z 。

有了上面这个关系, 带入到系统的能量和压强中, 就能得到我们希望看到的结果。能量为

$$\frac{E}{V} = \int_0^{+\infty} \frac{g(E)E}{z^{-1}e^{\beta E} - 1} \quad (3.76)$$

作同样的近似和展开, 得到

$$\frac{E}{V} = \frac{3z}{2\lambda^3 \beta} \left(1 + \frac{z}{4\sqrt{2}} + \dots \right) \quad (3.77)$$

我们将反解好的 z 代入, 得到

$$E = \frac{3}{2} \frac{N}{\beta} \left(1 - \frac{1}{2\sqrt{2}} \frac{\lambda^3 N}{V} + \dots \right) \left(1 + \frac{1}{4\sqrt{2}} \frac{\lambda^3 N}{V} + \dots \right) \quad (3.78)$$

从而我们有

$$pV = Nk_B T \left(1 - \frac{\lambda^3 N}{4\sqrt{2}V} + \dots \right) \quad (3.79)$$

从而玻色气体在高温极限下回到了经典气体。

3.6.4 低温极限——玻色爱因斯坦凝聚

我们同样是做上面的工作, 即: 反解出 μ , 然后代入能量表达式求出能量, 进而求出状态方程。在低温极限 $z \rightarrow 1$, 粒子数密度为

$$\frac{N}{V} = \frac{1}{4\pi^2} \left(2mk_B T / \hbar^2 \right)^{3/2} \int_0^{+\infty} dx \frac{x^{1/2}}{z^{-1}e^x - 1} \equiv \frac{1}{\lambda^3} g_{3/2}(z) \quad (3.80)$$

其中函数 $g_n(z)$ 的定义为

$$g_n(z) = \frac{1}{\Gamma(n)} \int_0^{\infty} dx \frac{x^{n-1}}{z^{-1}e^x - 1} \quad (3.81)$$

这是一个特殊函数，被称作多重对数函数。在附录中我们将证明

$$g_n(z) = \sum_{m=1}^{+\infty} \frac{z^m}{m^n} \quad (3.82)$$

当 $z = 1$ 时

$$g_n(1) = \zeta(n) \quad (3.83)$$

其中 $\zeta(n)$ 是黎曼 ζ 函数，其一个粗糙的定义为

$$\zeta(n) = \sum_{m=1}^{+\infty} \frac{1}{m^n} \quad (3.84)$$

在我们这里常用 $\zeta(3/2) \simeq 2.612$ 。

因此我们有

$$\frac{N}{V} = \frac{1}{\lambda^3} g_{3/2}(z) \quad (3.85)$$

我们现在来考虑低温的情况，我们拼命地降低温度，且保持粒子数不变。这意味着化学势随着温度自我调整而使得总粒子数不变。我们将 $g_{3/2}(z)$ 的图像画出来如图所示，我们看到 $g_{3/2}(z)$ 是一个单调增加的函数。当分母变大的时候， $z = e^{\beta\mu}$ 也应该跟着变大，所以化学势 μ 也跟着变大。当温度降低到一定程度时， μ 取到了 0。我们将这时的温度称作临界温度。当温度到达临界温度时，有

$$\frac{N}{V} = \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2} \zeta\left(\frac{3}{2}\right) \quad (3.86)$$

我们解得

$$T_C = \left(\frac{N}{\zeta(3/2)V} \right)^{2/3} \frac{2\pi\hbar^2}{mk_B} \quad (3.87)$$

当我们继续降低温度，这时似乎我们无法再维持系统的总粒子数不变了。

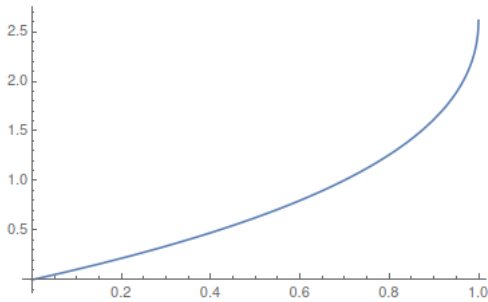


图 10: $g_{3/2}(z)$ 的图像。

到底出了什么错呢？在前面的计算中，我们将系统的态连续化了，即

$$\sum_r \simeq \int g(E) dE \sim \int dE E^{1/2} \quad (3.88)$$

当我们转换成积分的时候，我们忽略了 $E = 0$ 那里是有能级存在的，按照玻色爱因斯坦分布， $E = 0$ 时有

$$\langle n_0 \rangle = \frac{1}{z^{-1} - 1} \quad (3.89)$$

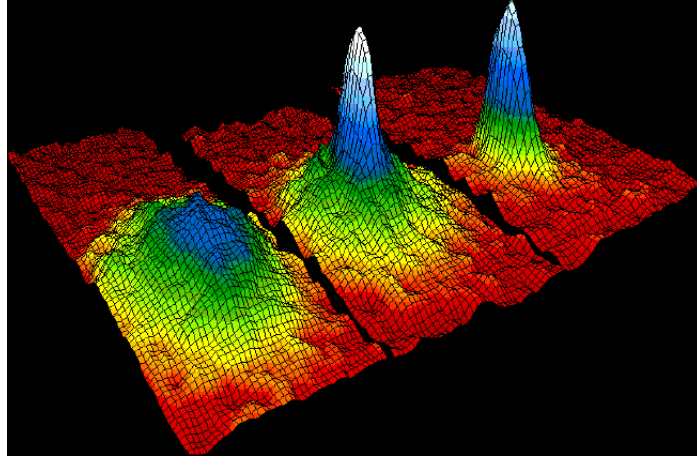


图 11: 于 1995 年得到的 Rb 原子的玻色爱因斯坦凝聚的实验结果, 最左边的一张图是在凝聚之前, 中间一张图是凝聚现象发生, 最右面一张图是进一步冷却得到的更纯的凝聚态。

在大多数高温情况, $z \rightarrow 0$, 所以这时候 $\langle n_0 \rangle \rightarrow 0$, 确实可以忽略不计。可是当 $T \rightarrow 0$ 的时候, $\langle n_0 \rangle \rightarrow +\infty$, 系统的粒子将全部待在基态上。因此我们改变一点我们上面的连续态的近似

$$N = \frac{V}{\lambda^3} g_{3/2}(z) + \frac{z}{1-z} \quad (3.90)$$

所以当 $T < T_C$ 时, 后面那一项开始变得重要起来, 于是我们有

$$\frac{n_0}{N} = 1 - \frac{V}{N\lambda^3} \zeta(3/2) = 1 - \left(\frac{T}{T_C} \right)^{3/2} \quad (3.91)$$

可见当 $T \rightarrow 0$ 时, $n_0/N \rightarrow 1$ 。这种在低温情况下大量粒子跑到基态上的现象被叫做玻色-爱因斯坦凝聚 (Bose Einstein Condensation, BEC)。

通常, 利用原子实现玻色爱因斯坦凝聚的温度都很低 $\sim 10^{-7}$ K, 玻色-爱因斯坦凝聚在 1995 年被观察到, 如图 11。

3.6.5 低温玻色气体的状态方程

我们可以求得系统的能量

$$E = \int_0^{+\infty} E g(E) \frac{1}{e^{\beta(E-\mu)} - 1} dE$$

进而我们可以得到压强

$$p = \frac{2}{3} \frac{E}{V} = \frac{k_B T}{\lambda^3} g_{5/2}(z) \quad (3.92)$$

在低温极限下，我们可以取 $z \rightarrow 1$ (可以当成温度为 0)，故有

$$p = \frac{k_B T}{\lambda^3} \zeta(5/2) \quad (3.93)$$

我们可以看到低温玻色气体的压强和高温的情况很不同， $p \sim T^{5/2}$ 。而当温度为 0 时，压强为 0，能量也为 0。

3.6.6 玻色气体的热容——BEC 是一种相变

严格来说，首先要把能量表示成为 N, V, T 的函数，首先要从 3.90 中将 z 求解出来，然后代入能量的表达式，再求出热容。我们可以看到这很困难，手算几乎不可能，所以我们就不算了。

最后如果我们将 $C_V(T)$ 的曲线画出来，会发现曲线的一阶导数不连续。当 $T \rightarrow T_C$ 时，如果粒子数 $N \rightarrow$ 趋于无穷，我们还会看到热容是发散的。这种不连续性被解释为一种相变，我们将会在下一章学习相变的更多性质。

3.6.7 附录：多重对数函数

$$\begin{aligned} g_n(z) &= \frac{1}{\Gamma(n)} \int_0^{+\infty} \frac{x^{n-1}}{z^{-1}e^x - 1} dx \\ &= \frac{1}{\Gamma(n)} \int_0^{\infty} dx \frac{ze^{-x} x^{n-1}}{1 - ze^{-x}} \\ &= \frac{1}{\Gamma(n)} \int_0^{\infty} dx x^{n-1} e^{-x} \frac{z}{1 - ze^{-x}} \\ &= \frac{z}{\Gamma(n)} \int_0^{\infty} dx x^{n-1} e^{-x} \sum_{m=0}^{+\infty} z^m e^{-mx} \\ &= \frac{1}{\Gamma(n)} \sum_{m=1}^{\infty} z^m \int_0^{\infty} dx x^{n-1} e^{-mx} \\ &= \frac{1}{\Gamma(n)} \sum_{m=1}^{\infty} \frac{z^m}{m^n} \int_0^{\infty} du u^{n-1} e^{-u}, (u = mx) \\ &= \sum_{m=1}^{\infty} \frac{z^m}{m^n} \end{aligned}$$

当 $z = 1$ 时，则有

$$g_n(1) = \zeta(n) \quad (3.94)$$

3.6.8 附录：使用正则系综推导玻色-爱因斯坦分布

为了证实我们第一章的论述：研究什么问题和使用什么系综无关，我们在这一小节中使用正则系综来推导玻色-爱因斯坦分布。

同样地，把系统分成很多个能级，能级 $|r\rangle$ 上的平均粒子数写作

$$\langle n_r \rangle = \sum_{\psi} n_r(\psi) p(\psi) \quad (3.95)$$

正则系综取态 ψ 的概率为

$$p(\psi) = \frac{e^{-\beta E(\psi)}}{Z(N)} \quad (3.96)$$

其中 $Z(N)$ 表示 N 个粒子的配分函数。于是有

$$\langle n_r \rangle = \frac{1}{Z(N)} \left(\sum_{n_1=0}^{+\infty} \sum_{n_2=0}^{+\infty} \cdots \right) n_r e^{-\beta(n_1 \epsilon_1 + n_2 \epsilon_2 + \cdots)} \quad (3.97)$$

注意上面的每个能级的粒子数应该满足

$$n_1 + n_2 + \cdots n_r + \cdots = N \quad (3.98)$$

下一步，取 $n'_r \equiv n_r - 1$ ，所以我们有

$$\begin{aligned} \langle n_r \rangle &= \frac{1}{Z(N)} \left(\sum_{n_1=0}^{+\infty} \sum_{n_2=0}^{+\infty} \cdots \sum_{n'_r=0}^{+\infty} \cdots \right) (n'_r + 1) e^{-\beta \epsilon_r (1 + n'_r) - \beta(n_1 \epsilon_1 + \cdots)} \\ &= \frac{1}{Z(N)} \left(\sum_{n_1=0} e^{-\beta n_1 \epsilon_1} \right) \left(\sum_{n_2=0} e^{-\beta n_2 \epsilon_2} \right) \cdots \left(\sum_{n'_r=0} (n'_r + 1) e^{-\beta (n'_r + 1) \epsilon_r} \right) \cdots \end{aligned}$$

那么上面的求和则改为

$$n_1 + n_2 + \cdots n'_r + \cdots = N - 1 \quad (3.99)$$

所以我们有了一个递推关系

$$\begin{aligned} \langle n_r \rangle(N) &= \frac{e^{-\beta \epsilon_r}}{Z(N)} (Z(N-1) (1 + \langle n_r \rangle(N-1))) \\ &= \frac{Z(N-1)}{Z(N)} e^{-\beta \epsilon_r} (1 - \langle n_r \rangle(N-1)) \end{aligned}$$

其中 $n_r(N)$ 表示系统有 N 个粒子时 r 能级上的平均粒子数。现在我们引入化学势的定义

$$\mu = \left(\frac{\partial F}{\partial N} \right)_{T,V} \quad (3.100)$$

但是我们应该注意到, N 是离散的, 所以在 N 很大的情况下

$$\mu = F(T, V, N) - F(T, V, N-1) = -k_B T (\ln Z(N) - \ln Z(N-1)) \quad (3.101)$$

所以我们有

$$\boxed{\frac{Z(N-1)}{Z(N)} = e^{\mu/k_B T}} \quad (3.102)$$

代入上面的递推关系, 我们有

$$\langle n_r \rangle(N) = e^{\mu/k_B T} e^{-\beta \epsilon_r} (1 + \langle n_r \rangle(N-1)) \quad (3.103)$$

由于 N 非常大, 所以多一个粒子和少一个粒子并不会影响 r 上的平均粒子数, 所以我们有

$$\langle n_r \rangle(N) \simeq \langle n_r \rangle(N-1) \quad (3.104)$$

代入, 可以得到

$$\langle n_r \rangle = \frac{1}{e^{(\epsilon_r - \mu)/k_B T} - 1} \quad (3.105)$$

这正是我们的玻色爱因斯坦分布。

3.7 费米统计

下面我们来考虑费米气体, 注意我们在本章所考虑的气体都是无相互作用的。费米子服从泡利不相容原理, 如果你试图把两个相同的费米子放在一个能级上, 波函数将会消失。也就是说, 一个能级上只能容纳一个费米子。

3.7.1 费米-狄拉克分布

尽管我们也能从正则系综推导出费米气体的分布 (和上一小节的步骤类似), 我们在这里采用更加方便的巨正则系综。由于每个能级上要么只有一个粒子, 要么没有粒子, 每个能级上的配分函数为

$$\begin{aligned} \mathcal{Z}_r &= \sum_{n=0,1} e^{-\beta n(E_r - \mu)} \\ &= 1 + e^{-\beta(E_r - \mu)} \end{aligned}$$

由此可以得到每个能级上的平均粒子数

$$\begin{aligned}\langle n_r \rangle &= \frac{1}{\beta} \ln \mathcal{Z}_r \\ &= \frac{1}{e^{\beta(E_r - \mu)} + 1}\end{aligned}$$

上式被称作费米-狄拉克分布。和在推导玻色气体的时候一样，我们也可以从整体系统的配分函数出发来验证这一点。整个系统的配分函数为

$$\mathcal{Z} = \prod_r \mathcal{Z}_r = \prod_r (1 + e^{-\beta(E_r - \mu)}) \quad (3.106)$$

可以求得总粒子数

$$N = \frac{1}{\beta} \ln \mathcal{Z} = \sum_r \frac{1}{e^{\beta(E_r - \mu)} + 1} = \sum_r \langle n_r \rangle \quad (3.107)$$

和期望的一样。关于上面的过程读者如果比较困惑，建议看一下玻色气体的论述。

注意有一点是我们在推导玻色气体的时候为了保证级数收敛，所以必须有 $\mu < 0$ ，但是费米子这里没有出现这个要求，所以费米子的化学势可正可负。

3.7.2 理想费米气体

同样，上面我们只是对于费米气体做了一个一般性的讨论，现在我们来研究理想费米气体，理想费米气体的能量为

$$E = \frac{p^2}{2m} \quad (3.108)$$

我们还要考虑费米子的自旋简并度。一般来说，费米子的自旋简并度为

$$g_s = 2s + 1 \quad (3.109)$$

其中 s 是半整数。考虑到上面的哈密顿量和简并度，我们可以求出态密度

$$\begin{aligned}\sum_r \rightarrow \int g(E) dE &= \frac{V}{(2\pi)^3} \int d^3k g_s \\ &= \frac{V}{(2\pi)^3} \int 4\pi k^2 dk g_s \\ &= (2s + 1) \int \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \sqrt{E} dE\end{aligned}$$

所以态密度为

$$g(E) = (2s+1) \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \sqrt{E} \quad (3.110)$$

同样地，我们引入逸度 $z = e^{\beta\mu}$ ，总平均粒子数为

$$N = \int dE \frac{g(E)}{z^{-1} \exp(\beta E) + 1} \quad (3.111)$$

注意，有一天，我们会从上式解出 μ ，最终将我们的能量换成和 N, T, V 有关的函数。

总平均能量为

$$E = \int dE \frac{E g(E)}{z^{-1} e^{\beta E} + 1} \quad (3.112)$$

压强为

$$\begin{aligned} pV &= \frac{1}{\beta} \ln \mathcal{Z} \\ &= \frac{1}{\beta} \int g(E) dE \ln (1 + z e^{-\beta E}) \\ &= \frac{2}{3} \int E g(E) \frac{1}{z^{-1} e^{\beta E} + 1} \\ &= \frac{2}{3} E \end{aligned}$$

再次强调，等式右边被积分掉的 E 是能级，而最后的 E 实际上是平均能量 $\langle E \rangle$ 。

上面我们写出了费米气体的所有公式，我们先考虑高温极限，按照 $z \ll 1$ 展开，重复我们在玻色气体的过程，最终我们可以得到状态方程

$$pV = N k_B T \left(1 + \frac{\lambda^3 N}{4\sqrt{2} g_s V} + \dots \right) \quad (3.113)$$

我们看到，一阶近似增加了压强。物理上，这是因为费米子之间的泡利不相容原理，排斥了粒子。

3.7.3 低温极限：简并费米气体和费米面

下面我们来研究当温度 $T \rightarrow 0$ 时，费米气体的行为。我们之前讲过，当 $T \rightarrow 0$ 时，玻色粒子全部分布在基态。而费米子由于泡利不相容原理，所以便会不断地填满能级。具体地，我们可以从每个能级的平均粒子数看出来，当 $T \rightarrow 0$ 时，有

$$\langle n_r \rangle = \frac{1}{e^{\beta(E-\mu)} + 1} \rightarrow \begin{cases} 1, & E < \mu \\ 0, & E > \mu \end{cases} \quad (3.114)$$

由于我们考虑的粒子数是有限的，它们最多只能填到一个特定的能级上，这个能级我们记作 E_F ，这样一个能级在 \vec{k} 空间上便是满足 $E_{\vec{k}} = E_F$ 的曲面，这个曲面我们把他叫做费米面，而能量 E_F 被称作费米能。在数值上

$$E_F = \mu(T = 0) \quad (3.115)$$

我们在这里又要唠叨一遍。正常来说，在巨正则系综中 (T, V, μ) 是三个独立变量，巨正则系综就是输入这三个参数，给出 (E, p, N) 三个参数。但是我们往往考虑 N 固定的情况，因此 μ 便是 N, V, T 的函数。所以才有了“ $T = 0$ 时候的化学势”这一说，否则按照巨正则系综 μ 应该是不依赖于任何变量的独立参数。所以实际上，化学势，或者说费米能，应该由粒子数来确定，即（零温时）

$$N = \int_0^{E_F} dE g(E) = \frac{g_s V}{6\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} E_F^{3/2} \quad (3.116)$$

从中我们可以解出

$$E_F = \frac{\hbar^2}{2m} \left(\frac{6\pi^2 N}{g_s V} \right)^{2/3} \quad (3.117)$$

在我们考虑的自由费米气体的能量 $E = p^2/2m$ ，等能面是 $|\vec{k}|$ 相同的球面，但是在一些其他的情况，尤其是固体中，能量的表达式可能不仅仅依赖于 \vec{k} 的大小，还有可能依赖于 \vec{k} 的方向，那时候，等能面会变得十分复杂。

在自由费米气体的情况，我们可以定义对应的费米动量

$$\hbar k_F = (2mE_F)^{1/2} \quad (3.118)$$

$|\vec{k}| < k_F$ 的能态都被填满了，我们把它们形象地乘坐费米海或者费米球。我们还可以定义费米温度，它是费米能量对应的温度

$$k_B T_D = E_F \quad (3.119)$$

我们用费米温度来界定我们的近似范围，当 $T > T_D$ 时，可以认为高温近似是有效的，而当 $T < T_D$ 时，这是“相对低”的一个温度。但是对于金属中的电子，费米温度 $T_F \sim 10^4$ K，对于白矮星中的电子， $T_F > 10^7$ K，所以可想而知，我们日常的室温对于费米气体来说是一个非常低的温度，所以在室温下大多电子都填满基态而形成费米海。

在零温时，我们也可以将能量和压强用费米能表示

$$E = \int_0^{E_F} dE E g(E) = \frac{3}{5} N E_F \quad (3.120)$$

压强为

$$pV = \frac{2}{3}E = \frac{2}{5}NE_F \quad (3.121)$$

可以看到在零温时，费米气体也有压强，这是因为泡利不相容原理导致的，被称作简并压。

3.7.4 低温费米气体

这里所说的低温，并不是日常生活中的低温，而是指 $T \ll T_F$ ，实际上，大多数室温的费米气体就已经满足这个条件了。我们重新列出关键的式子

$$N = \int_0^{+\infty} dE g(E) \frac{1}{z^{-1}e^{\beta E} + 1} \quad (3.122)$$

$$E = \int_0^{+\infty} dE \frac{Eg(E)}{z^{-1}e^{\beta E} + 1} \quad (3.123)$$

其中态密度为

$$g(E) = \frac{(2s+1)V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} E^{1/2} \quad (3.124)$$

我们先采用一个粗糙的近似

Claim 1 如果保持粒子数不动 $\left(\frac{\partial N}{\partial T} \right)_{V,\mu} = 0$ ，将会导致

$$\left(\frac{\partial \mu}{\partial T} \right)_{N,V} \Big|_{T=0} = 0 \quad (3.125)$$

下面我们来说明这个近似。因为

$$\begin{aligned} \frac{dN}{dT} &= \frac{d}{dT} \int_0^{+\infty} dE \frac{g(E)}{e^{\beta(E-\mu)} + 1} \\ &= \int_0^{+\infty} dE g(E) \frac{d}{dT} \left(\frac{1}{e^{\beta(E-\mu)} + 1} \right) \end{aligned}$$

我们知道括号内的函数在零温时是一个阶跃函数，偏离零温的费米狄拉克分布如图 12 所示，这张图其实是偏离零温很多（大于 T_F 好几个数量级）的情况，由于我们只考虑零温附近的情况，所以上面的分布其实仍然近似于零温的分布。所以我们可以做近似 $\mu \simeq E_F$ ，我们将这个代入上面的式子，得到

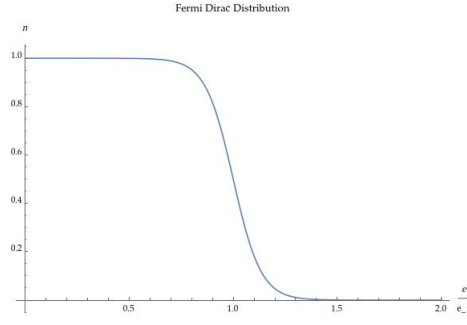


图 12: 偏离零温很多的费米-狄拉克分布

$$\begin{aligned}
 \frac{dN}{dT} &= \int_0^\infty dE g(E) \frac{\partial}{\partial T} \left(\frac{1}{e^{\beta(E-E_F)} + 1} \right) \\
 &\simeq g(E_F) \int_0^\infty dE \frac{E - E_F}{k_B T^2} \frac{1}{4 \cosh^2(\beta(E - E_F))/2} \\
 &\simeq 0
 \end{aligned}$$

上面将 $g(E_F)$ 提出来是因为整个积分只在 E_F 附近有值，而近似等于 0 是因为被积函数在 E_F 附近是奇函数。由此我们便证明了上面的声明。

于是我们使用上面的近似来计算热容

$$C_V = \left(\frac{\partial E}{\partial T} \right)_{N,V} = \int_0^\infty dE E g(E) \frac{\partial}{\partial T} \left(\frac{1}{e^{\beta(E-\mu)} + 1} \right) \quad (3.126)$$

我们泰勒展开上式的第一项

$$E g(E) = E_F g(E_F) + \frac{3}{2} g(E_F) (E - E_F) \quad (3.127)$$

我们来看能级上的平均粒子数的这一项，它求导之后是在 E_F 附近的一个非常尖锐的峰。

$$\frac{\partial}{\partial T} \left(\frac{1}{e^{\beta(E-E_F)} + 1} \right) = \frac{1}{k_B T^2} \frac{(E - E_F) e^{\beta(E-E_F)}}{(e^{\beta(E-E_F)} + 1)^2} \quad (3.128)$$

我们做变量替换

$$x = \beta(E - E_F) \quad (3.129)$$

我们得到

$$\frac{\partial}{\partial T} \left(\frac{1}{e^{\beta(E-E_F)} + 1} \right) = \frac{x e^x / \beta}{e^{2x} + 2e^x + 1} \frac{1}{k_B T^2} \quad (3.130)$$

所以我们得到热容的第一项

$$\int_0^\infty dE E_F g(E_F) \frac{\partial}{\partial T} \left(\frac{1}{e^{\beta(E-\mu)} + 1} \right) \simeq \frac{1}{2} k_B E_F g(E_F) \int_{-\infty}^\infty x dx \frac{1}{1 + \cosh x} = 0$$

其中我们利用能级平均粒子数的特性，将积分 $\int_{-E_F}^\infty$ 近似为 $\int_{-\infty}^\infty$ ，这是因为当 $x < -E_F$ 时，被积函数几乎为 0，于是忽略。上式等于 0 是因为

$$\begin{aligned} \cosh(x) &\equiv \frac{1}{2} (e^x + e^{-x}) \\ &= \cos(ix) \\ &= 1 - \frac{1}{2!}(ix)^2 + \frac{1}{4!}(ix)^4 + \dots \\ &= 1 + \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \dots \end{aligned}$$

由此可见， $\cosh(x)$ 是一个偶函数。所以上式积分为 0。用同样的技巧计算第二项，由此得到热容为

$$C_V \simeq \frac{3}{2} g(E_F) k_B^2 T \int_{-\infty}^\infty \frac{x^2 dx}{4 \cosh^2(x/2)} = \frac{\pi^2}{2} g(E_F) k_B^2 T \quad (3.131)$$

我们来看一下上面结果的物理含义。这意味着，当温度升高的时候，在 $E < E_F$ 的区域的粒子会受到热的扰动，从而能量变为 $E_F \pm k_B T$ ，从而使得系统总能量增加了 $C_V T$ 那么多，因此可以求出受到热扰动的粒子数目大概为 $\sim g(E_F) k_B T$ 。

最后，我们利用粒子数和费米能的关系，有

$$C_V \sim N k_B \left(\frac{T}{T_F} \right) \quad (3.132)$$

这是一个很重要的结果。金属中的电子就是费米子，而金属的热容可以看成是晶格振动（声子）的贡献加上自由电子气体的贡献，因此实验上用下面的式子来拟合热容-温度曲线

$$C_V = \gamma T + \alpha T^3 \quad (3.133)$$

实验上，我们常常改写为

$$C_V/T = \gamma + \alpha T^2 \quad (3.134)$$

这意味着如果画 $C_V/T, T$ 的曲线，将得到一条直线。通过仔细比较声子的贡献和电子的贡献，我们可以得到当

$$T^2 \sim T_D^3/50 T_F \quad (3.135)$$

的时候两项是差不多的，其中 T_D 是我们在玻色气体那一小节里面的德拜温度。

说实话，如果有我特别想做的实验的话，我会无比想去实验室验证这个关系。可惜我在写讲义的时候没有找到太好的实验数据来说明上面的推导和实验吻合的怎么样，如果真的能去实验室一点点把热容测出来，那真是一件特别有成就感的事情。你真的可以从一小块固体里面，知道整个世界是什么样子的！

3.7.5 索末菲展开

上面我们使用的近似十分粗糙，现在我们使用严谨的方法，这个严谨的方法会帮助你计算到你想要的任何一阶的效应，这种方法被叫做索末菲展开。

在计算之前，清楚我们的目标是至关重要的。我们知道要想精确计算热容无非就是精确求解

$$N(z, \beta, V) = \int_0^{+\infty} dE g(E) \frac{1}{z^{-1}e^{\beta E} + 1} \quad (3.136)$$

$$= \frac{g_s}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^{\infty} \frac{E^{1/2}}{z^{-1}e^{\beta E} + 1} \quad (3.137)$$

在这里，我已经把原来的 $N(T, V, \mu)$ 转换成了三个变量 (z, T, V) ，它和原来的表述是等价的。之前我们要解出 μ ，而现在我们要求解 z ，它的本质没有变，只是数学上更加方便了。

我们做换元 $x = \beta E$ ，我们得到

$$\frac{N}{V} = \frac{g_s}{\lambda^3} f_{3/2}(z) \quad (3.138)$$

其中定义一个和在玻色气体见到的类似的特殊函数

$$f_n(z) = \frac{1}{\Gamma(n)} \int_0^{\infty} dE \frac{x^{n-1}}{z^{-1}e^x + 1} \quad (3.139)$$

因此能量可以写作

$$\frac{E}{V} = \frac{3}{2} \frac{g_s}{\lambda^3} k_B T f_{5/2}(z) \quad (3.140)$$

索末菲展开的方法就是将特殊函数 $f_n(z)$ 展开成 $\ln z$ 的级数，我们同样考虑低温近似 $z \rightarrow \infty$ 。我们将上面的特殊函数拆成两部分

$$\Gamma(n) f_n(z) = \int_0^{\beta\mu} dx \frac{x^{n-1}}{z^{-1}e^x + 1} + \int_{\beta\mu}^{+\infty} dx \frac{x^{n-1}}{z^{-1}e^x + 1} \quad (3.141)$$

因为 ze^{-x} 是一个小量, 我们将第一项表示为

$$\begin{aligned}\int_0^{\beta\mu} dx \frac{x^{n-1}}{z^{-1}e^x + 1} &= \int_0^{\beta\mu} dx x^{n-1} \left(1 - \frac{1}{1 + ze^{-x}}\right) \\ &= \frac{1}{n} (\ln z) - \int_0^{\beta\mu} \frac{x^{n-1}}{1 + ze^{-x}}\end{aligned}$$

我们做变量替换 $\eta_1 = \beta\mu - x$, $\eta_2 = x - \beta\mu$, 则有

$$\Gamma(n)f_n(z) = \frac{1}{n} (\ln z) - \int_0^{\beta\mu} \frac{(\beta\mu - \eta_1)^{n-1}}{1 + e^{\eta_1}} d\eta_1 + \int_0^\infty d\eta_2 \frac{(\beta\mu + \eta_2)^{n-1}}{1 + e^{\eta_2}} \quad (3.142)$$

我们取 $z \rightarrow \infty$ 的近似, 于是我们将中间的积分上限改为 ∞ 并且把两个积分合在一起

$$\Gamma(n)f_n(z) = \frac{1}{n} (\ln z) + \int_0^{+\infty} d\eta \frac{(\beta\mu + \eta)^{n-1} - (\beta\mu - \eta)^{n-1}}{1 + e^\eta} \quad (3.143)$$

将分子按照 μ 展开, 得到

$$\begin{aligned}(\beta\mu + \eta)^{n-1} - (\beta\mu - \eta)^{n-1} &= (\beta\mu)^{n-1} \left[(1 + \eta/\beta\mu)^{n-1} - (1 - \eta/\beta\mu)^{n-1} \right] \\ &\simeq 2(n-1)(\beta\mu)^{n-2}\eta\end{aligned}$$

所以 $\Gamma(n)f_n(z)$ 最终可以表达为

$$\Gamma(n)f_n(z) = \frac{1}{n} (\ln z)^n + 2(n-1)(\ln z)^{n-2} \int_0^\infty \frac{\eta d\eta}{1 + e^\eta} \quad (3.144)$$

其中的积分可以按如下方法计算

$$\begin{aligned}\int_0^\infty \frac{\eta d\eta}{1 + e^\eta} &= \int_0^\infty d\eta \eta \sum_{m=1}^{+\infty} (-1)^{m+1} e^{-m\eta} \\ &= \sum_{m=1}^{+\infty} \frac{(-1)^{m+1}}{m^2} \int_0^\infty u e^{-u} du, \quad (u = m\eta) \\ &= \sum_{m=1}^{+\infty} \frac{(-1)^{m+1}}{m^2} \\ &= 1 - \frac{1}{2^2} + \frac{1}{3^2} - \frac{1}{4^2} + \cdots \\ &= 1 + \left(\frac{1}{2^2} - 2 \times \frac{1}{2^2} \right) + \left(\frac{1}{3^2} \right) + \left(\frac{1}{4^2} - 2 \times \frac{1}{4^2} \right) \cdots \\ &= \left(1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots \right) - 2 \times \left(\frac{1}{2^2} + \frac{1}{4^2} + \frac{1}{6^2} + \cdots \right) \\ &= \left(1 - \frac{1}{2} \right) \left(1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots \right) \\ &= \frac{1}{2} \zeta(2) \\ &= \frac{\pi^2}{12}\end{aligned}$$

$\zeta(2)$ 有很多种求法, 下面这种求法归功于欧拉。考虑函数 $\sin x$, 它具有零点

$$x = \pm m\pi, m = 0, 1, 2, \dots \quad (3.145)$$

将 $\sin x$ 进行多项式近似, 并保证它在原点的行为和原函数类似, 则有

$$\sin x = \lim_{n \rightarrow \infty} x \left(1 - \frac{x^2}{\pi^2}\right) \left(1 - \frac{x^2}{4\pi^2}\right) \cdots \left(1 - \frac{x^2}{n^2\pi^2}\right) \quad (3.146)$$

$$= 1 - \left(1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots\right) x^2/\pi^2 + \cdots \quad (3.147)$$

和 $\sin x$ 的二阶展开式做比较, 有

$$\zeta(2) = \left(1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots\right) = \frac{\pi^2}{6} \quad (3.148)$$

最终我们得到 $f_n(z)$ 的展开

$$f_n(z) = \frac{(\ln z)}{\Gamma(n+1)} \left(1 + \frac{\pi^2}{6} \frac{n(n-1)}{(\ln z)^2}\right) \quad (3.149)$$

注意因为我们上面使用了亿些近似, 所以上式只在低温情况下近似成立。通过这样一个近似, 我们便可以实现我们的梦想: 解出 $\mu(T, V, N)$, 利用

$$\frac{N}{V} = \frac{g_s}{\lambda^3} f_{3/2}(z) \quad (3.150)$$

代入上面的展开式, 并利用费米能的定义式

$$E_F = \frac{\hbar^2}{2m} \left(\frac{6\pi^2}{g_s} \frac{N}{V}\right)^{2/3} \quad (3.151)$$

则有

$$E_F = \mu \left(1 + \frac{\pi^2}{12} \left(\frac{k_B T}{\mu}\right)^2\right) \quad (3.152)$$

从上式可以解得

$$\mu(T=0) = E_F \quad (3.153)$$

对上式两边对 T 求导数, 可以求得 μ 在 $T=0$ 处的一阶导数

$$\left(\frac{\partial \mu}{\partial T}\right)_{T=0} = -\frac{\pi^2}{6} \frac{k_B^2 T}{E_F} \quad (3.154)$$

因此可以求得化学势的一阶近似

$$\mu \simeq E_F \left(1 - \frac{\pi^2}{6} \left(\frac{k_B T}{E_F}\right) + \cdots\right) \quad (3.155)$$

实际上，你想求到几阶就求到几阶。但是注意这个近似还是只在 $T = 0$ 处成立，因为我们在得到式 3.152 之前还做了一些近似。

从上式我们可以看出 μ 的变化趋势， μ 在 $T = 0$ 取最大，而在经典极限， $\mu < 0$ 。最后，我们终于可以更加严谨地计算能量

$$\frac{E}{V} = \frac{g_s}{10\pi^2\hbar^3} (2m)^{3/2} \mu^{5/2} \left(1 + \frac{5\pi^2}{8} \left(\frac{k_B T}{\mu} \right)^2 + \dots \right) \quad (3.156)$$

于是

$$\frac{E}{N} = \frac{3E_F}{5} \left(1 + \frac{5\pi^2}{12} \left(\frac{k_B T}{E_F} \right)^2 + \dots \right) \quad (3.157)$$

则热容为 (保留到 T 的最低阶)

$$C_V = \left(\frac{\partial E}{\partial T} \right)_{N,V} = N k_B \frac{\pi^2}{2} \frac{T}{T_F} \quad (3.158)$$

这正是我们之前的结果。

3.7.6 白矮星和钱德拉塞卡极限

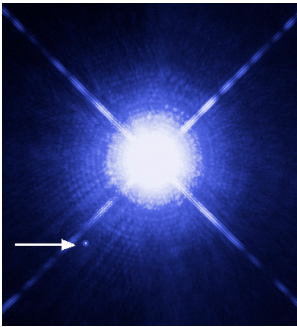


图 13: 哈勃望远镜拍摄的天狼星 A 和天狼星 B 的照片，其中天狼星 B 是一颗白矮星，是如图箭头所指的小点。

费米气体还有一个重要的应用就是白矮星。白矮星描述的是当恒星的能量耗尽之后，行星由于自身引力塌缩，最终达到一个平衡态的状态。这时引力造成的压强和电子的简并压 (这简并压正是我们之前提到的) 相同。电子简并压和费米能之间的关系为

$$pV = \frac{2}{5} N E_F \quad (3.159)$$

白矮星整体的能量除了电子的动能之外还有引力势能，我们假设质量均匀分布，白矮星的密度

$$\rho = \frac{M}{\frac{4}{3}\pi R^3} \quad (3.160)$$

因此可以求出白矮星整体的引力势能为

$$E_{\text{grav}} = - \int_0^R \frac{G \rho^2 4}{r} \pi r^3 \cdot (4\pi r^2) dr = - \frac{3}{5} \frac{GM^2}{R} \quad (3.161)$$

其中 G 为牛顿引力常数。注意，尽管整体的引力很大，但是引力对于每一个电子的能级的影响我们是把他忽略了。电子的行为我们仍然认为是一团自由

的费米气体。而且我们在这里使用了牛顿的引力描述，但是这并不影响我们讨论其中的物理。

我们还有一点要做修正。就是我们之前的非相对论近似是不对的，应该采用相对论的能量-动量关系

$$E = \sqrt{p^2 c^2 + m^2 c^4} \quad (3.162)$$

实际上，我们考虑相对论的能量-动量关系是因为白矮星的温度很高（表面温度 8000 K），但实际上，它又远远低于白矮星的费米温度 T_F ，所以 $T = 0$ 的近似在分析的时候仍然好用。

使用这个能量-动量关系，我们可以计算得到态密度，考虑到电子有两个自旋分量 ($s = 1/2$)

$$g(E) = \frac{2 \times V E}{2\pi^2 \hbar^3 c^3} \sqrt{E^2 - m^2 c^4} \quad (3.163)$$

$$= \frac{2 \times V E^2}{2\pi^2 \hbar^3 c^3} \left(1 - \left(\frac{m c^2}{E} \right)^2 \right)^{1/2} \quad (3.164)$$

$$= \frac{V}{\pi^2 \hbar^3 c^3} \left(E^2 - \frac{1}{2} m^2 c^4 + \dots \right), \text{ (利用近似 } E \gg m c^2 \text{)} \quad (3.165)$$

现在我们考虑零温的情况作为我们的近似，(因为温度相对于费米温度是个小量，所以这其实算零阶近似)，每个能级上的平均粒子数为

$$\langle n_r \rangle = \begin{cases} 1, & E < \mu \\ 0, & E > \mu \end{cases} \quad (3.166)$$

其中化学势在数值上等于费米能

$$\mu(T = 0) = E_F \quad (3.167)$$

我们忽略引力势能对于单个电子的影响，得到动能为

$$E_{\text{kinetic}} = \int_0^{E_F} E g(E) dE = \frac{V}{\pi^2 \hbar^3 c^3} \left(\frac{1}{4} E_F^4 - \frac{m^2 c^4}{4} E_F^2 + \dots \right) \quad (3.168)$$

其中的费米能 (零温化学势) 由白矮星的总粒子数决定

$$\begin{aligned} N &= \int_0^{E_F} dE g(E) \\ &= \frac{V}{\pi^2 \hbar^3 c^3} \left(\frac{1}{3} E_F^3 - \frac{1}{2} m^2 c^4 E_F + \dots \right) \end{aligned}$$

我们只保留到 mc^2 的零阶，有

$$E_F = \left(\frac{3N\pi^2\hbar^3 c^3}{V} \right)^{1/3} = \hbar c \left(\frac{3N\pi^2}{V} \right)^{1/3} \quad (3.169)$$

所以我们代入动能的表达式，并且利用对于密度的近似

$$\frac{N}{V} = \frac{M}{m_p} \frac{3}{4\pi R^3} \quad (3.170)$$

其中 m_p 为质子质量，我们把白矮星的总质量看成全部由质子提供。读者需要经过几步推导才能得到

$$E_{\text{kinetic}} = \frac{3}{4} \hbar c \left(\frac{9\pi M^4}{4m_p^4} \right)^{1/3} \frac{1}{R} + O(R) \quad (3.171)$$

其中正比于 R 的第二项的前面的系数可以计算是正的。我们将动能和势能合并起来

$$\begin{aligned} E = E_{\text{grav}} + E_{\text{kinetic}} &= C_1 \frac{1}{R} + C_2 R \\ &= \left[\frac{3}{4} \hbar c \left(\frac{9\pi M^4}{4m_p^4} \right)^{1/3} - \frac{3}{5} GM^2 \right] \frac{1}{R} + O(R) \end{aligned}$$

通过对函数 $C_1 \frac{1}{R} + C_2 R$ 的判断我们知道，在 $C_2 > 0$ 的情况，若想使得能量有稳定解，那么必须有 $C_1 > 0$ ，如果 $C_1 \leq 0$ ，那么就不存在稳定解了。也就是说，如果当白矮星的质量足够大，则会导致上面的 $C_1 \leq 0$ ，这时电子的简并压就不能阻止引力的效应，就不能形成稳定的星体了，恒星就会进一步塌缩，形成黑洞或者中子星（以中子简并压对抗引力的星体）。我们将这个临界质量记作 M_c ，当 $M < M_c$ 时，恒星就会塌缩形成白矮星，而不会形成黑洞或者中子星

恰好能取稳定解的条件为

$$M_c \sim \left(\frac{\hbar c}{G} \right)^{3/2} \frac{1}{m_p^2} \sim 1.4 M_\odot \quad (3.172)$$

其中 M_\odot 是太阳质量。因此我们的太阳当燃料耗尽的时候会变成一颗白矮星。这个极限 M_c 被称作钱德拉塞卡极限。

3.7.7 泡利顺磁性

3.7.8 朗道抗磁性

To see a world in a grain of sand
And a heaven in a wild flower,
Hold infinity in the palm of your hand
And eternity in an hour.

William Blake, *Auguries of Innocence*

4 相变初步