

《热力学与统计物理》讲义

作者：吕京涛

组织：华中科技大学物理学院

时间：2022 年 1 月

版本：3.14159

目录

1 前言	2
1.1 与其他课程及学科关系	2
1.2 研究对象、研究方法	2
1.3 研究必要性 ($1 \neq 10^{23}$)	3
2 概率与排列组合	5
2.1 概率	5
2.2 排列组合	12
3 第零定律与温度	16
3.1 热力学基本名词	16
3.2 状态量、过程量和状态方程	17
3.3 热量、热容	18
3.4 第零定律和温度	19
3.5 温度的统计定义	21
3.6 系综 (Ensemble)/Gibbs 系综	25
3.7 Boltzmann 因子和配分函数	26
4 热力学第一定律	28
4.1 上部分回顾：压强的统计定义 *	28
4.2 热力学第一定律	29
4.3 常见热力学过程	31
4.4 恰当微分 (exact differential)	35
5 热力学第二定律	37
5.1 Carnot 循环	37
5.2 热力学第二定律	39

5.3 Carnot 定理	39
5.4 Clausius 定理	39
6 熵	41
6.1 热力学熵	41
6.2 Boltzmann 熵	43
6.3 Gibbs 熵	45
6.4 Shannon 熵	47
6.5 von Neumann 熵	48
6.6 最大熵原理	50
6.7 信息论: Maxwell-Szilard-Landauer-Bennett	52
7 热力学势	57
7.1 Legendre 变换	57
7.2 焓	58
7.3 Hemholtz 势	59
7.4 Gibbs 势	60
7.5 热力学第二定律	61
7.6 稳定性条件	62
8 化学势	63
8.1 化学势的定义	63
8.2 从自由能的角度考虑化学势	65
8.3 巨正则系统	65
8.4 Euler 方程、Gibbs-Duhem 关系	68
8.5 粒子数守恒与化学势	69
8.6 化学势: 内部贡献和外部贡献	69
8.7 应用: 化学反应	70
9 统计物理的简单框架	73
9.1 理论原理	73

9.2 最大熵原理	76
9.3 实际应用	78
9.4 涨落, 不同系综的“等价性”	81
9.5 举例 1: spin 1/2	84
9.6 举例 2: 理想气体	86
10 量子统计	89
10.1 密度矩阵	89
10.2 von Neumann 熵	95
10.3 Wigner 变换和 Wigner 分布函数	96
10.4 举例	97
11 理想气体	101
11.1 系统	101
11.2 单粒子配分函数	101
11.3 N 粒子配分函数: 第一步尝试	103
11.4 N 粒子配分函数: 粒子的不可分辨性	104
11.5 Sackur-Tetrode 方程	106
11.6 Gibbs 佯谬	108
12 光子和声子气体	110
12.1 光子气体	110
12.2 声子气体	115
13 量子气体	119
13.1 回顾: 经典理想气体	119
13.2 自旋统计定理	120
13.3 占据数表象	120
13.4 费米子	121
13.5 玻色子	122
13.6 经典极限	123

13.7 Fermi 和 Bose 气体	125
13.8 应用举例: Fermi 气体	127
13.9 应用举例 2: Bose 气体	135
14 玻色-爱因斯坦凝聚	136
14.1 问题引入	136
14.2 宏观系统	137
14.3 解决方案	138
14.4 玻色-爱因斯坦凝聚	138
15 van der Waals 方程	142
15.1 vdW 方程	142
15.2 vdW 方程的应用	147
15.3 位力展开	150
16 相变	151
16.1 相与相变	151
16.2 气-液-固相变	153
16.3 超导相变	156
16.4 Ising 模型	159
16.5 Landau 相变理论	169
16.6 临界指数	172
17 布朗运动	176
17.1 导论: 为什么研究涨落?	176
17.2 布朗运动	176
17.3 朗之万方程的微观理论	189
17.4 Fokker-Planck 方程	195
18 主方程方法	199
18.1 经典主方程	199
18.2 Lindblad 量子主方程	204

19 分子动理论学	210
19.1 概率密度分布	210
19.2 Liouville 定理, 方程	210
19.3 BBGKY 链式方程组	212
19.4 Boltzmann 方程	214
19.5 Boltzmann H-定理	218
19.6 流体动力学方程	220
19.7 扩展阅读	226
20 线性响应理论	228
20.1 唯象理论	228
20.2 线性响应理论: 绝热近似	233
20.3 响应函数的本征表示	236
20.4 涨落-耗散定理	237
20.5 线性响应理论的应用: 自由电子气对电磁场的响应	238
20.6 阅读材料	251
21 练习 & 作业	253
21.1 系综理论	253
22 附录 I: 二次量子化	255
22.1 单粒子量子系统	255
22.2 多粒子量子系统的希尔伯特空间	256
22.3 全同粒子系统的交换对称性	258
22.4 产生/湮灭算符	260
22.5 占据数表象	263
22.6 算符的二次量子化	265
22.7 举例	268
22.8 总结	269
23 附录 II: 概率论基础	271

24 附录 III: 常用积分	289
24.1 Γ 函数	289
24.2 高斯积分	289
24.3 Stirling 公式	290
24.4 n 维球的体积和表面积	291
24.5 Riemann ζ 函数	292
24.6 多对数函数	293
25 附录:	295
25.1 简谐振子	295
26 授课时间表	297

说明

本讲义包含大三《热力学与统计物理》以及研一《高等统计物理》的部分内容。

第1章 前言

1.1 与其他课程及学科关系

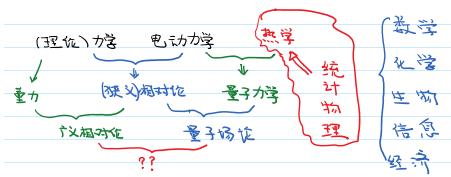


图 1.1: 与其他课程及学科的关系。

1.2 研究对象、研究方法

《热力学与统计物理》这门课程分两个主要部分。这两个部分的研究对象是相同的，都是有大量微观粒子所组成的宏观系统，但研究方法不同。热力学部分与我们学习过的《热学》课程有很大的重叠，侧重于从大量宏观的实验现象来总结出系统所满足的规律，所以热力学从事的是现象学的研究。这也决定了热力学定律是非常普适的物理规律。这一点在我们学习《热学》课程的时候已经深有体会，这里不再赘述。统计物理的主要任务是从已知的描述微观粒子的运动方程出发利用统计的方法来推导出大量微观粒子组成的宏观系统所满足的基本规律，也就是建立微观和宏观物理量之间的联系。所以，统计物理采用的是一个“自下而上”的方法来构建热力学规律的微观理论。这是我们这门课程需要侧重的内容，也是这门课与《热学》课程的主要区别。

要进行统计物理的研究我们需要两个准备条件：(1) 理解微观粒子的运动规律；(2) 基本的统计方法。其中，条件(1)是我们之前的课程学习到的。比如，要研究微观粒子的运动以及它们之间的相互作用通常我们需要去求解薛定谔方程或者牛顿运动方程。这些基础知识我们在量子力学和理论力学中已经学到。而要描述电磁波的传播就需要用到 Maxwell 方程，这是我们在电动力学中学习的内容。条件(2)是统计学的基本知识，这部分内容我们会在本课

程中给大家简单介绍一下。

1.3 研究必要性 ($1 \neq 10^{23}$)

从上面的介绍我们可以知道，热统作为物理学院本科生“四大力学”专业基础课程之一，与其他的三大力学（理论力学、量子力学、电动力学）是完全不同的。其他三大力学的研究对象都是个体或少体，而热统研究的是由大量个体聚集在一起之后所表现出来的规律。至少有下面几个原因使得热统成为独立于其他三大力学的一门重要的专业基础课程。

1.3.1 涌现现象 (Emergent Phenomena)

传统的还原论观点认为，只要我们理解了微观粒子之间的相互作用或运动规律，就可以推断出大量微观粒子聚集起来之后的集体行为规律。因此，人们去研究组成世界的基本粒子。通过这门课程的学习我们会发现，事实远非这样简单。当粒子数量大到一定程度就会衍生出单个或少体粒子所不具备的新现象，这就是所谓的涌现现象 (Emergent Phenomena)。这一现象在凝聚态物理以及生物物理中非常常见。可以用著名凝聚态物理学家 P. W. Anderson 的一句话概括：“More is different.”¹ 比如，固体中电子与电子、电子与声子的相互作用的微观机制我们很清楚，但宏观材料所表现出来的多种多样的电学性质却完全不同。超导现象就可以理解为一种涌现现象，它是由大量电子在特定晶格中聚集在一起所表现出来的宏观电学特性。我们可以对其进行现象学的描述，但有关高温超导的微观机制我们还不清楚。再比如，我们人脑是有大量的神经元所组成的。人们认为每个神经元的作用就类似于计算机的一个元件。大量神经元集合在一起形成了人脑，产生了意识。人的意识产生的微观机制人们也不清楚。所有这些例子表明，如果我们要构建一个所谓的万物理论 (“Theory of Everything”) 的话，统计物理必不可少。

1.3.2 统计平均行为

退一步说，即使不存在我们上面提到的涌现现象。一个宏观体系往往包含了大量的微观粒子。我们虽然可以把每个粒子的微观运动方程写出来，但用如此之多的方程来描述系统是不现实的。我们可以举一个简单的例子：我们暂时不考虑描述系统的运动，而只是希望使用

¹P. W. Anderson, Science, 177, 393 (1972)

计算机来数一下一摩尔分子。我们知道一摩尔包含了 $N_A = 6.022 \times 10^{23}$ 个分子, N_A 为阿佛加德罗常数 (Avogadro number)。如果计算机 cpu 的主频为 3 GHz, 那么每秒钟可以数 3×10^9 个分子。数完一摩尔的话, 需要 $\sim 2 \times 10^{14}$ 秒。而我们知道一年有 $365 \times 24 \times 60 \times 60 = 31536000$ 秒。所以, 我们一共需要 ~ 6 百多万年来数完这一摩尔的分子。仅仅数一下这些分子就需要这么长时间, 更不要说去跟踪每个分子的运动了。

事实上, 我们在基本所有情况下都对每个微观粒子的运动情况不感兴趣。我们真正感兴趣的是这些大量粒子表现出来的宏观特性。比如对于气体来说, 我们希望知道他们的体积、压强、温度等有限几个物理量。这些宏观量是大量粒子所表现出来的平均行为, 与具体每个粒子的运动情况没有关系。**而如何从微观的动力学来得到这些宏观特性正是我们统计物理要研究的问题。** Feynman 曾经说过这样一段话, 可以对我们上面的讨论给一个很好的总结: “Anyone who wants to analyze the properties of matter in a real problem might want to start by writing down the fundamental equations and then try to solve them mathematically. Although there are people who try to use such an approach, these people are the failures in this field...”

1.3.3 与其他学科的关系

历史上来看, 对热现象的研究是始终伴随着人类文明的发展。从人类对火的认识和使用开始, 到由蒸汽机的发明带来的第一次工业革命, 到第二次工业革命, 再到信息技术的发展, 热力学和统计物理都起到了举足轻重的作用。蒸汽机的使用自然不必说。信息技术的发展也是和热力学统计物理紧密联系在一起的。从热力学发展的初期, 人们就意识到信息在热力学中的重要地位。其中, 最为著名的例子就是 Maxwell 小妖。后来, Shannon 有关通信原理的阐释中最为重要的概念信息熵, 也和热力学和统计物理中的熵密切相关。

热力学与统计物理的研究对很多学科的发展也起到了重要的推动作用。比如, 原子论的逐步建立是和统计物理的发展密不可分的。此外, 人们关于黑体辐射的研究直接导致了量子力学的诞生。我们有理由相信, 在未来科学的发展过程中, 热力学与统计物理也必将发挥同样重要的作用。不仅如此, 统计物理原理如此普适, 以至于其对于化学、生物甚至经济、社会学科的发展都起到了举足轻重的作用。

第 2 章 概率与排列组合

2.1 概率

统计物理的基本语言是概率论。概率论是数学的一个重要分支。它是由法国人 Pierre de Fermat 和 Blaise Pascal 为了解决赌博方面的问题发展起来的。下面我们简单介绍我们课程中需要了解的概率论基本知识。

2.1.1 随机变量

随机变量的取值不确定，但是我们可以对它的可能取值的范围，以及在这个取值范围内取不同值的概率有一个确定的认识。根据随机变量取值范围的特点，我们可以把它分为离散随机变量、连续随机变量和混合随机变量。它们的意义显而易见。

对于物理问题来说，概率的起源至少有两种：

- 首先，对于某些问题我们可能已知的条件有限，所以也就不能得出确定的结论，而只能谈论不同可能性的概率。比如，对于一个容器内的气体来说。随机挑选其中一个分子，这个分子的速度是随机的。给定一个小的空间范围，其中的分子数也是随机的。这是因为我们对这些气体分子运动的初始条件不清楚。理论上讲，如果我们知道了某一个时刻所有分子的状态，同时也知道分子之间相互散射的微分散射截面，那么在后面任意时刻，每个分子的状态就变为确定的了。
- 概率可能还来自于理论本身。我们知道在量子力学的理论框架下，即使我们确定的知道某一个粒子处于一个能量本征态，它在空间的分布也是不确定的，而对应的本征态波函数的模方可以用来描述它在不同空间位置的分布概率。

随机变量不同取值的概率可以从两种方法得到：

- 客观方法：我们对随机变量进行多次测量，根据测量的结果来确定其取不同值的概率。比如，如果我们对随机变量 X 进行 N 次测量，它取值为 x 的次数为 N_x ，那么它取值为 x 的概率为

$$P_X(x) = \lim_{N \rightarrow \infty} \frac{N_x}{N}.$$

- 主观方法：我们根据对随机变量的认识来得出其取值概率。我们的认识程度不同，相应

的概率也就不一样。比如，我们知道掷骰子时，骰子的可能取值为 $\{1, 2, 3, 4, 5, 6\}$ ，这是一个离散随机变量。而如果这个骰子是正常的话，那么每个取值的概率都是一样的，都是 $1/6$ 。

如果我们知道了与随机变量所有可能取值相对应的概率，那么这个随机变量就确定下来了。

2.1.1.1 离散随机变量

离散随机变量的可能取值是离散的。比如，掷骰子得到的数字、我们这个课堂的人数、每个家庭的孩子数等。对于离散随机变量，我们需要知道的是它取每一个数值的概率。我们用 $P_X(x_i)$ 表示离散随机变量 X 取值为 x_i 的概率。显然， $P_X(x_i)$ 需要满足下面几个特征：

- 概率要非负: $P_X(x_i) \geq 0$
- 概率要满足归一化条件: $\sum_i P_X(x_i) = 1$
- 概率要可以叠加: $P_X(x_1 \mid\mid x_2) = P_X(x_1) + P_X(x_2)$

已知概率分布，我们可以：

- 计算离散随机变量的平均值：

$$\langle X \rangle = \sum_i x_i P_X(x_i).$$

值得说明的是离散随机变量的平均值不一定是它的可能取值。

- 计算任意 X 的函数 $f(X)$ 的平均

$$\langle f(X) \rangle = \sum_i f(x_i) P_X(x_i).$$

- 为了表征随机变量相对于平均值的分布范围，我们定义其方差 (variance)：

$$\sigma_X^2 = \langle (X - \langle X \rangle)^2 \rangle.$$

而 σ_X 为标准差。根据这个定义我们可以证明：

$$\sigma_X^2 = \langle X^2 \rangle - \langle X \rangle^2.$$

2.1.1.2 连续随机变量

对于连续随机变量，由于其可取的值无限多，我们可以定义其概率密度分布 (Probability Density Distribution - PDF) $P(x)$ 使得 $P_X(x)dx$ 为随机变量 X 取值在 $[x, x + dx]$ 范围内的概率。这样定义的 $P_X(x)$ 也满足

- 概率非负: $P_X(x) \geq 0$
- 可归一化: $\int P_X(x)dx = 1$
- 可以叠加: $P_X(x_1 \text{ or } x_2) = P_X(x_1) + P_X(x_2)$

我们可以定义类似的平均值和方差:

$$\begin{aligned}\langle X \rangle &= \int P_X(x)x dx, \\ \langle f(X) \rangle &= \int P_X(x)f(x) dx, \\ \sigma_X^2 &= \langle (X - \langle X \rangle)^2 \rangle = \langle X^2 \rangle - \langle X \rangle^2.\end{aligned}$$

此外, 对于离散随机变量, 我们也可以通过使用 Dirac delta 函数来写出其概率密度分布函数

$$P_X(x) = \sum_i \delta(x - x_i)P_i.$$

2.1.1.3 随机变量的线性变换

对随机变量 X 进行线性变换可以生成新的随机变量 Y , 比如 $Y = aX + b$ 。相应 Y 的平均值和方差也可以由 X 得到

- $\langle Y \rangle = a\langle X \rangle + b$
- $\langle Y^2 \rangle = \langle (aX + b)^2 \rangle = \langle a^2X^2 + b^2 + 2abX \rangle = a^2\langle X^2 \rangle + b^2 + 2ab\langle X \rangle$
- $\sigma_Y^2 = \langle Y^2 \rangle - \langle Y \rangle^2 = a^2\sigma_X^2$

举例: 如果一个随机变量只有两个可能取值, 比如我们用 1 和 0 表示, 对应的概率分别为 p 和 $1 - p$ 。很容易计算:

$$\langle X \rangle = p,$$

$$\langle X^2 \rangle = p,$$

$$\sigma_X^2 = p(1 - p).$$

现在取 n 个这样的随机变量, 其中值为 1 的变量数为 m 的概率满足二项式分布 $P(n, m) = C_n^m p^m (1 - p)^{n-m}$ 。

举例: 高斯分布对应的概率密度分布可以表示为

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}},$$

(1) 验证它是归一化的; (2) 计算其均值和方差。

高斯积分:

$$\int_{-\infty}^{+\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}}$$

$$\int_{-\infty}^{+\infty} x^2 e^{-ax^2} dx = \frac{1}{2} \sqrt{\frac{\pi}{a^3}}$$

$$\int_{-\infty}^{+\infty} x^4 e^{-ax^2} dx = \frac{3}{4} \sqrt{\frac{\pi}{a^5}}$$

$$\int_{-\infty}^{+\infty} x e^{-ax^2} dx = 0$$

$$\int_0^{+\infty} x e^{-ax^2} dx = \frac{1}{2a}$$

$$\int_0^{+\infty} x^3 e^{-ax^2} dx = \frac{1}{2a^2}$$

举例: 泊松分布 考虑这样的一种情况: 一个随机事件沿着 X 轴 (可以是时间也可以是空间) 发生且满足下面两个条件:

- 当 $\Delta X \rightarrow 0$ 时, 这个事件在从 X 到 $X + \Delta X$ 这个范围内发生且只发生一次的概率为 $r\Delta X$ 。这里 r 为一个不依赖于 X 的常数。
- 在 X 轴不同位置发生的事件之间是相互独立的, 没有任何关系。

这种情况下, 在长度为 L 的范围内发生 n 次事件的概率为泊松分布, 可以表示为:

$$P(n) = \frac{1}{n!} \frac{(rL)^n}{e^{rL}}.$$

比如, 一块儿放射性物质在时间 T 范围内发生衰变的次数满足泊松分布。当然这里时间 T 必须小于放射性物质的衰变寿命。

练习:验证泊松分布的归一性, 计算其均值和方差。

2.1.2 多个随机变量

上面的很多考虑可以推广到多个随机变量。比如, 对于两个随机变量, 我们可以定义类似的联合概率密度分。比如, 对于 X 和 Y , 我们定义 X 取值范围在 $[x, x + dx]$, Y 取值范围在 $[y, y + dy]$ 的概率为 $P_{X,Y}(x, y)dx dy$ 。对于平均值和方差的计算都是类似的。

根据联合概率密度分布函数，我们可以得到相对于某一个变量的约化概率密度分布

$$P_X(x) = \int P_{X,Y}(x,y)dy.$$

这个约化之后的概率密度分布满足单变量概率密度分布的所有特征。

此外，我们还可以定义一个条件概率密度分布 $P_X(x|y)$ 表示 Y 取 y 值时， X 取 x 的概率密度。

根据这一定义，我们不难理解

$$P_X(x|Y=y) = P_{X,Y}(x,y)/P_Y(y),$$

或者

$$P_{X,Y}(x,y) = P_X(x|y)P_Y(y).$$

同理有下面的关系式

$$P_{X,Y}(x,y) = P_Y(y|x)P_X(x).$$

这个关系式为 Bayes 定律。

如果两个随机变量是相互独立的，它们的联合概率密度分布变为每个变量概率密度分布的乘积

$$P_{X,Y}(x,y) = P_X(x)P_Y(y).$$

这时，我们有

$$P_X(x|y) = P_X(x).$$

物理上，两个相互独立的随机变量的取值之间没有任何关联。 Y 的取值不依赖于 X 的取值，反之亦然。

举例: [B&B,15.5] 已知参加运动会的运动员有 1% 的概率服用兴奋剂。而进行兴奋剂检查时，检查仪器给出正确结果的概率为 95%。现在有一位运动员仪器查出他/她服用兴奋剂。请问他/她真的服用了兴奋剂的概率是多少？

举例: [B&B,15.6] 已知一对夫妇生有两个小孩。其中一个为男孩，那么另一个小孩为女孩的概率为多大？如果告诉你，较高的那个是男孩，那么另一个小孩为女孩的概率为多大？

2.1.3 随机变量的函数

一个随机变量 X 通过一个函数变换可以生成一个新的随机变量 Y , 即 $Y = f(X)$ 。如果我们已知 X 的概率密度分布函数 $P_X(x)$, 那么如何得到 Y 的概率密度分布函数呢? 下面我们介绍一种比较普适的方法, 但开始之前, 我们需要定义个新的函数累积概率密度分布

$$CP_X(x) = \int_{-\infty}^x P_X(x') dx'.$$

显然如果知道了 $CP_X(x)$ 我们可以通过求微分的方式得到 $P_X(x)$ 。

我们可以采用下面几个步骤来计算 $P_Y(y)$:

- 根据 $y = f(x)$ 的曲线确定满足 $y \leq \eta$ 的 x 取值范围
- 在确定的范围内对 $P_X(x)$ 进行积分, 确定 $CP_Y(\eta)$
- 把上面得到的 $CP_Y(\eta)$ 相对于 η 做微分得出 $P_Y(\eta)$

注意, 在第二部我们得到一个积分形式的表达式, 但我们没有必要去计算这个积分。在第三步时我们只需要对第二步得到的积分表达式进行微分。这里需要用到下面的数学关系:

$$\frac{d}{dy} \int_{l(y)}^{h(y)} g(x, y) dx = g(h(y), y) \frac{d}{dy} h(y) - g(l(y), y) \frac{d}{dy} l(y) + \int_{l(y)}^{h(y)} \frac{d}{dy} g(x, y) dx.$$

举例: 已知电场振幅 E 的概率密度分布为

$$P(E) = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{E^2}{2\sigma^2}},$$

求其场强 $I = aE^2$ 的概率密度分布。

举例: 已知

$$P(\theta) = \begin{cases} \frac{1}{2\pi} & 0 \leq \theta < 2\pi, \\ 0 & otherwise, \end{cases}$$

计算 $x = x_0 \sin(\theta)$ 的概率密度分布。

上面的方法还可以用来计算两个或多个随机变量的情况。比如, 如果已知 $f(x, y)$ 和 $P_{X,Y}(x, y)$, 我们可以计算 $P(f)$ 。

举例: 计算两个随机变量之积 $z = xy$ 的概率密度分布。

2.1.4 随机变量的和

我们经常会需要把一些随机变量进行求和得到一个新的随机变量, 比如 $X = \sum_i X_i$, 其中, 每个 X_i 都是一个随机变量。进行求和的这些随机变量之间往往不是相互独立的。如果我

们已知每个随机变量 X_i 的均值和方差，就可以来计算求和得到的随机变量的均值和方差。

均值的计算很简单我们不再赘述，我们可以得到

$$\langle X \rangle = \sum_i \langle X_i \rangle.$$

对于方差，我们有

$$\sigma_X^2 = \langle (X - \langle X \rangle)^2 \rangle = \langle \left(\sum_i (X_i - \langle X_i \rangle)^2 \right) \rangle.$$

把上式右边展开后我们会得到 $\langle (X_i - \langle X_i \rangle)(X_j - \langle X_j \rangle) \rangle$ 这样的表达式。它们可以这样计算

$$\langle (X_i - \langle X_i \rangle)(X_j - \langle X_j \rangle) \rangle = \langle X_i X_j \rangle - \langle X_i \rangle \langle X_j \rangle = \begin{cases} 0 & i \neq j \\ \sigma_{X_i}^2 & i = j \end{cases}.$$

把所有非零项考虑进去，我们有

$$\sigma_X^2 = \sum_i \sigma_{X_i}^2.$$

所以相互独立的随机变量之和对应的随机变量的均值和方差分别是各个独立随机变量均值和方差之和。

在特殊情况下，如果有 n 个随机变量，而且它们具有相同的均值和方差，那么

$$\langle X \rangle = n \langle X_i \rangle$$

$$\sigma_X^2 = n \sigma_{X_i}^2$$

举例：(1) 多次实验测量；(2) 醉汉随机行走。

2.1.4.1 中心极限定理

不严谨版本的中心极限定理可以描述如下：如果有 n 个相互独立的随机变量，它们满足相同的分布，均值为 E_x ，方差为 σ_x^2 。 n 个这样的随机变量之和得到的随机变量 S_n 的均值为 nE_x ，方差为 $n\sigma_x^2$ 。当 n 足够大时，如果 nE_x 和 $n\sigma_x^2$ 都不发散，那么 S_n 通常近似满足均值和方差相同的高斯分布

$$P_{S_n}(s) = \frac{1}{\sqrt{2\pi n\sigma_x^2}} e^{-\frac{(S_n - nE_x)^2}{2n\sigma_x^2}}. \quad (2.1)$$

对于中心极限定理这里我们不做证明，只做几点说明：

- σ_x 不能发散：比如，Lorentzian 分布 $P(x) = \frac{1}{\pi} \frac{\delta}{x^2 + \delta^2}$ 的标准差是发散的，所以不在此列。

但我们也知道，满足 Lorentzian 分布的随机变量之和仍然是 Lorentzian 分布。

- n 多大算足够大？对于具体应用来说，一般不需要太大的数值。有时甚至 $n < 10$ 就可以。
- 通常近似：当然有特殊情况。比如离散随机变量之和的概率密度分布，不管 n 多大，永远是离散的，它不可能等于高斯分布。但这时，这一离散概率密度分布的包络满足高斯分布。

此外，中心极限定理不只适用于我们描述的情况。我们有关相互独立随机变量的假设可以放松，而且有时这些随机变量也不一定满足相同的概率密度分布。总之，在很多很多随机过程中的随机变量都可以近似为高斯分布。

2.2 排列组合

在统计物理中我们经常需要考虑多粒子体系中的组合问题，这可以通过下面的例子进行说明。考虑一个由 10 个原子组成的系统，每个原子可以处于两能级之一：低能级对应能量为零，高能级对应一个能量 ε_0 。如果已知体系的宏观总能量为 $10\varepsilon_0$ ，那么可以确定每个原子都处于高能级。但如果已知总能量为 $4\varepsilon_0$ ，那么这四个能量量子分别分布在哪些原子上呢？首先，如果我们可以区分每一个原子，我们可以把它们编号为 1 到 10 号。这是，可以有 $10 \times 9 \times 8 \times 7 = \frac{10!}{6!}$ 中可能的能量分布状态满足总能量为 $4\varepsilon_0$ 。如果每个粒子是不可区分的，相应的分布状态变为 $C_{10}^4 = \frac{10!}{4! \times 6!}$ 。

上面的例子中我们只考虑了 10 个原子，假如我们考虑 10^{23} 个量级的原子的话，我们就会遇到 $10^{23}!$ 这么大的数了。在处理这么大的数字的时候我们可以考虑它的对数。这是我们反复用到下面的关系式

$$\ln n! \approx \frac{1}{2} \ln 2\pi + (n + \frac{1}{2}) \ln n - n + \frac{1}{12n} + O(\frac{1}{n^2}).$$

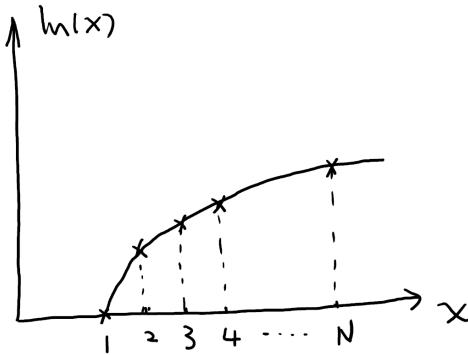
这个公式叫做 Stirling 近似。它的证明可以参考教材的附录。很多时候我们还可以做进一步近似，写成最为常用的形式：

$$\boxed{\ln n! \approx n \ln n - n + \frac{1}{2} \ln 2\pi n \approx n \ln n - n.}$$

这个近似可以用对数函数的积分来理解。当 n 足够大的时候，

$$\ln n! = \sum_{m=1}^n \ln m \approx \int_1^n \ln x dx \approx n \ln n - n,$$

见下图。



下面我们应用这个结果来验证一下中心极限定理。

例 1：二项式分布

$$P(n, m) = \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m}$$

当 $n \rightarrow \infty$, np 有限大小时趋于均值为 np 泊松分布

$$P(m) = \frac{(np)^m}{m!} \cdot e^{-(np)},$$

证明：由

$$P(m) = \frac{(np)^m}{m!} \times \frac{n!}{(n-m)! \cdot n^m} (1-p)^{n-m}.$$

利用 Stirling 公式 $\ln n! \approx n \ln n - n$, 我们有

$$\begin{aligned} & \ln \frac{n!}{(n-m)! n^m} (1-p)^{n-m} \\ &= n \ln n - n - (n-m) \ln(n-m) + (n-m) - m \ln n + (n-m) \ln(1-p) \\ &= (n-m) \ln \frac{n(1-p)}{(n-m)} - m \\ &= (n-m) \ln \left[1 + \frac{m-np}{n-m} \right] - m \\ &\approx (n-m) \times \frac{m-np}{n-m} - m \\ &= -np \end{aligned}$$

这里，我们利用了当 n 足够大，但 np 为有限大小时， $\left| \frac{m-np}{n-m} \right| \rightarrow 0$ ，对 \ln 函数做了泰勒展开。

例 2：同样为二项式分布，当 n 足够大， $np(1-p)$ 有限时，可以近似为均值为 np ，方差为 $np(1-p)$ 的高斯分布。

证明：

① 写出高斯分布：

均值: np ; 方差: $np(1-p) \gg 1$.

对应高斯分布:

$$P_g(m) = \frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{(m-np)^2}{2np(1-p)}}$$

取对数:

$$\ln P_g(m) = -\frac{(m-np)^2}{2np(1-p)} - \frac{1}{2} \ln[2\pi np(1-p)] \quad (2.2)$$

②对应二项式分布:

$$P_b(m) = \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m}.$$

取对数:

$$\ln P_b(m) = \ln n! - \ln m! - \ln(n-m)! + m \ln p + (n-m) \ln(1-p). \quad (2.3)$$

我们需要比较式 (2.2) 和 (2.3), 证明它们在 n 足够大时是相等的。我们同样利用 Stirling 近似:

$\ln n! \approx n \ln n - n + \frac{1}{2} \ln 2\pi n$, 从式 (2.3) 得到

$$\begin{aligned} \ln P_b &= n \ln n - n - m \ln m + m - (n-m) \ln(n-m) + (n-m) \\ &\quad + \frac{1}{2} \ln \frac{2\pi n}{2\pi m 2\pi(n-m)} + m \ln p + (n-m) \ln(1-p) \\ &= m \ln \frac{np}{m} + (n-m) \ln \frac{n-np}{n-m} + \frac{1}{2} \ln \frac{2\pi n}{2\pi m 2\pi(n-m)} \end{aligned} \quad (2.4)$$

下面, 我们 $np - m$ 是一个小量, 对其进行展开。(通过前面的分析我们知道, 当 n 增加的时候, 随机变量的取值偏离其均值的概率会越来越小。把 $np - m$ 看成一个小量的意思是说, 我们在随机变量均值附近对其进行近似。这个近似在 $n \rightarrow \infty$ 的时候应该会足够好, 因为那时随机变量偏离其均值的概率非常小。)

$$\begin{aligned} m \ln \frac{np}{m} &= m \ln \left(1 + \frac{np-m}{m}\right) \\ &\approx (np-m) - \frac{(np-m)^2}{2m}, \end{aligned} \quad (2.5)$$

$$(n-m) \ln \frac{n-np}{n-m} \approx (m-np) - \frac{(m-np)^2}{2(n-m)}, \quad (2.6)$$

$$\frac{1}{2} \ln \frac{1}{2\pi m \left(1 - \frac{m}{n}\right)} \stackrel{\frac{m}{n} \approx p}{\approx} -\frac{1}{2} \ln [2\pi np(1-p)]. \quad (2.7)$$

式 (18.27) 和 (18.28) 相加得：

$$\begin{aligned} m \ln \frac{np}{m} + (n-m) \ln \frac{n(1-p)}{n-m} &\approx -\frac{1}{2}(m-np)^2 \cdot \left(\frac{1}{m} + \frac{1}{n-m}\right) \\ &= -\frac{1}{2}(m-np)^2 \cdot \frac{n}{m(n-m)} \\ &\approx -\frac{(m-np)^2}{2np(1-p)} \end{aligned}$$

于是，我们发现， $\ln P_b \approx \ln P_g$.

第3章 第零定律与温度

“Thermodynamics is a funny subject. The first time you go through it, you don't understand it at all. The second time you go through it, you think you understand it, except for one or two small points. The third time you go through it, you know you don't understand it, but by that time you are used to it, so it doesn't bother you any more.”

Arnold Sommerfeld

3.1 热力学基本名词

具备了基本数学概念后，我们这一部分开始热力学的学习。首先，我们要统一一下语言，介绍几个大家热学中已经接触到的名词。热力学中我们通常把整个宇宙中我们关心的部分称为系统，而把宇宙中所有其他部分成为环境。我们关心的系统主要可以分为以下三种：

- 孤立系统 (isolated system): 系统通过绝热壁与外界环境隔离，它们之间既没有能量，又没有粒子交换。
- 封闭系统 (closed system): 系统与外界有能量交换，但没有粒子交换。
- 开放系统 (open system): 系统与外界既有能量交换，又有粒子交换。

我们假设，对于一个孤立系统，只要我们等足够长的时间，它就会最终达到一个平衡态。达到平衡态之后，系统内部不同部分之间达到了力、热、电、化学等各个方面的平衡。这时我们去测量系统的某些宏观参量的话，它们的均值不再随时间发生变化。或者说，系统的宏观状态不再发生变化。这里我们强调系统的宏观状态不再变化，是因为其微观状态一直在改变。比如，在平衡态下如果去跟踪系统中某一个粒子的话，它的位置和速度微观参量等都在随时间发生变化。从这个角度讲，平衡态是一个动态的平衡¹。

这里我们提到了宏观参量和微观参量。我们说，热力学中我们关心的是系统的宏观参量，比如体积 (V)、压强 (p) 以及我们马上要讨论的温度 (T)。对于电介质的话，我们可能还需要考虑极化强度、磁化强度、电场强度、磁场强度等。这些都是可以测量得到的。而微观参量则是描述组成系统的微观粒子状态的量，比如，每个粒子的动量、位移等。我们很难对它们进行实际测量。

¹这里，我们需要区别平衡态和稳态。稳态一般存在系统与外界之间的能量或者粒子交换。

对于宏观参量，根据它们随系统尺度的变化是否改变，我们可以把它们分为两类。第一类为广延量 (extensive variable)，随着系统大小的增加而增加，这包括体积、粒子数 (N)、电极化、磁化、熵、能量等。第二类为强度量 (intensive variable)，不随系统大小的变化而变化，比如温度、化学势、电场强度、磁场强度、粒子数密度等。在粒子数密度 $n = N/V$ 保持不变的情况下，粒子数 $N \rightarrow \infty$ ，体积 $V \rightarrow \infty$ 的极限状态成为热力学极限(thermodynamic limit)。

对于一个给定的处于平衡态的系统，我们可以用几个相互独立的状态参量来对它进行描述。具体参量的个数，不同的系统是不同的。比如，对于理想气体，我们只需要两个独立的参量，比如 p 和 V 。不同的状态参量之间，可能存在不同的关系，这个关系可以写成状态方程 (equation of state)的形式。比如理想气体的状态方程为

$$PV = \nu RT = Nk_B T.$$

这里， ν 为物质的量， $R = 8.314 \text{ J/mol/K}$ 为理想气体常数， $k_B = 1.3807^{-23} \text{ J/K}$ 为 Boltzmann 常数。

理想气体状态方程的由来： 我们《热学》课程学习过，理想气体状态方程是人们根据几个不同的实验规律而总结出来的。这包括：

1. T 固定时， $p \propto 1/V$ (Boyle's Law)
2. P 固定时， $V \propto T$ (Charles' Law)
3. V 固定时， $P \propto T$ (Gay-Lussac's Law)
4. ν, P, T 相同时， V 相同 (Avogadro's Law)

根据 1–3 我们可以知道 $pV \propto T$ ，再根据 4，我们有 $pV \propto \nu T$ 。具体的比例系数 R 可以利用实验来进行测量。

3.2 状态量、过程量和状态方程

- 热力学平衡态、热平衡
- 状态参量：处于平衡态的系统可以用几个状态参量来描述。顾名思义，状态参量的取值恰当由系统所处的状态来决定，与系统如何达到这一状态没有关系（与历史无关）。我们熟悉的状态参量有体积 V 、压强 P 、温度 T 等。
- 过程参量：与状态参量相对应的是过程量。过程量只有指定了具体过程后才能确定下来。过程不同，相应的过程量取值也可能不同。比如，热量和功都是过程量。

- 状态方程：不同类型的系统需要的状态参量的个数也不一样。对于我们熟悉的理想气体，我们需要两个独立的状态参量来描述。比如我们可以用 p, T 或者 V, T 。不同的状态参量之间存在一定的关系，这个关系可以用状态方程来描述。我们最熟悉的状态方程就是理想气体状态方程 $PV = Nk_B T$ 。
- 系统从一个状态 i 变化到另外一个状态 f ，其状态参量也会做相应的变化，比如从 x_i 变化到 x_f 。相应状态参量的函数也会从一个值变化到另外一个值，比如： $f(x_i) \rightarrow f(x_f)$ 。状态参量或者状态函数的这个性质在数学上可以用恰当微分的语言来描述。

3.3 热量、热容

3.3.1 热量 (heat)

热量是热力学中最为重要的概念之一。它表示的是两个系统在达到热平衡过程中进行的能量传递或转移。我们上节介绍的参量都是状态量。给定系统所处的平衡态，这些量就确定了下来，与系统如何达到这个平衡态无关，即与具体过程无关。而现在我们讨论的热量是不同的，它是过程量。它是描述能量在两个系统之间相互转移这个过程的量。我们只能针对某个过程来讨论热量，只给定系统状态是不行的。另外一个重要的过程量就是功，我们暂且不讨论。因为热量描述的是能量的转移，所以它的国际单位和能量相同为焦耳。

3.3.2 热容 (heat capacity)

热容描述的是一个物体存储能量的能力，所以成为能容可能更合适。但由于早期人们对热的认识不够，所以把它叫做热容，这个说法也就延续了下来。我们可以把它定义为

$$C \equiv \frac{\partial Q}{\partial T},$$

即使物体温度升高单位温度需要传递给它的热量。

但我们刚刚说过，热量是一个过程量。升高相同的温度，过程不同，吸收的热量可能也不一样，所以热容应该不是一个固定的值。我们接触最多的有两个：定容热容和定压热容。它们的定义为

$$C_{V/p} \equiv \left(\frac{\partial Q}{\partial T} \right)_{V/p}.$$

分别表示在保持体积和压强不变时，是系统升高单位温度所需热量。根据定义我们知道，热

容是一个广沿量。我们还可以定义相应的强度量。比如，摩尔热容表示一摩尔的物质的热容，比热容 (specific heat capacity) 表示单位质量的物质的热容。

对于我们熟悉的理想气体来说，我们知道 $C_p > C_v$ ，因为在定压条件下，气体温度升高的过程中会发生膨胀，相应气体对外做功消耗能量。气体吸收的热量一部分用来提高温度，一部分以对外做功的形式消耗掉；而对于定容过程，不需要对外做功。

3.4 第零定律和温度

3.4.1 第零定律

首先，我们介绍热平衡的概念。想象把两个原来相互独立的孤立系统相互实现热接触，使之可以相互传递热量。这两个系统一起构成了一个新的孤立系统。只要我们等足够长时间，系统会达到热平衡，这时两个子系统之间不再有宏观的热量交换。热力学第零定律描述的是这种热平衡的可传递性：

热力学第零定律：如果两个系统同时和第三个系统达到了热平衡，那么它们之间也已经达到热平衡。

3.4.2 温度的引入

第零定律最为重要的作用就是用来引入温度的概念。假设我们有三个系统 1, 2, 3，考虑下面的情况：

- 首先，我们把系统 1 (P_1, V_1) 与系统 3 (P_3, V_3) 进行热接触。一般情况下，这两个系统不会即时达到热平衡。至少一个系统的状态要做调整才能使二者达到平衡。比如，保持系统状态 1 不变，我们来调节系统 3 的状态。假设保持 P_1, V_1, V_3 不变，那么要使两个系统达到热平衡的话，我们需要对 P_3 进行调整。最终的结果是 P_3 不能任意取值，而是由 P_1, V_1, V_3 来决定的。也就是说，热平衡状态下，这四个变量间应该存在一定的关系：

$$F_{31}(P_3, V_3, P_1, V_1) = 0,$$

或者

$$P_3 = f_{31}(V_3, P_1, V_1).$$

- 同样道理，系统 2 和系统 3 达到热平衡的话，也有类似的关系

$$F_{32}(P_3, V_3, P_2, V_2) = 0,$$

或者

$$P_3 = f_{32}(V_3, P_2, V_2).$$

- 综合以上，我们有

$$f_{32}(V_3, P_2, V_2) = f_{31}(V_3, P_1, V_1). \quad (3.1)$$

- 下面我们利用第零定律知道，系统 1 和 2 也处于热平衡状态，即

$$F_{21}(P_2, V_2, P_1, V_1) = 0. \quad (3.2)$$

也就是说，利用式 (3.1) 我们可以得到式 (3.2)。我们看到，前者包含 V_3 ，但后者没有。这说明在前者当中的 V_3 可以被消去，即它可以写成

$$f_2(P_2, V_2) = f_1(P_1, V_1). \quad (3.3)$$

方程的左边和右边分别是系统 2 和 3 的状态参量的函数。我们可以引入状态参量“温度” T ，并把它写成

$$g(T_2) = f_2(P_2, V_2). \quad (3.4)$$

这样，当两个系统达到热平衡后， $g(T_2) = g(T_3)$ 。函数 g 可以有不同的形式。最为简单的就是 $g(T) = T$ 。

这样引入温度后，我们就可以说，两个系统如果达到了热平衡，那么它们的温度相同。如果两个系统同时和第三个系统的温度相同，那么这两个系统的温度也相同。

3.4.3 温度计

利用第零定律我们可以制作温度计，作为测温的参考系统，去测量不同系统的温度。当然，第零定律只是给出了温度的概念，但如果要定量的描述温度，还需要规定温标，用来确定温度的绝对数值。常用的温标有摄氏温标、华氏温标、理想气体温标、热力学温标等。

- 摄氏温标 (degree C): 把一个标准大气压下，水的沸点定位 100 度，水的冰点定位 0 度。
- 华氏温标 (degree F): 它与摄氏温标换算关系： $F = 32 + \frac{9}{5}C$
- 理想气体温标 (K): 利用理想气体状态方程 $PV = nRT$ 定义温度，这样的话，水的三相点温度为 273.16 K。与摄氏度换算关系为： $K = C - 273.15$

- 热力学温标 (K): 利用卡诺循环来定义

使用温度计测量温度是一种间接测量。比如，我们需要选定测温物质以及对应的测温属性。这些属性可以是电学、力学、几何属性。而且我们需要知道这些测温属性与温度之间的函数关系。通过测量测温属性，利用函数关系，得出温度。

3.5 温度的统计定义

3.5.1 微观和宏观状态

我们前面说过，系统的宏观状态确定后，其微观状态是在不断变化的。这说明同一个宏观状态可以对应不同的微观状态。这里可以举一个简单的例子：假设一个盒子里有两个硬币 ($N = 2$)，把盒子晃动后停下来。我们现在关心的是硬币的正面是朝上还是朝下。我们认为两个硬币之间是独立的，互不影响。另外，硬币正面朝上和朝下的概率是一样的。这个两硬币系统的微观状态有四种：上上，上下，下上，下下。我们定义宏观状态可以用正面朝上和朝下的硬币数 ($N_{\uparrow}, N_{\downarrow}$) 来描述 ($N = N_{\uparrow} + N_{\downarrow}$)。比如，一个正面朝上，一个正面朝下。因为我们只有两个硬币，所以宏观和微观状态对应比较简单：

宏观状态 ($N_{\uparrow}, N_{\downarrow}$)	对应微观状态数 (Ω)
(2,0)	1
(1,1)	2
(0,2)	1

这个两硬币系统有三种不同的宏观状态，四种微观状态。出现这四种微观状态的概率都是相同的，但出现这三种宏观状态的概率不同。因为同一个宏观状态可能对应不同种微观状态。

如果我们把硬币的数目增加到 10，那么情况就稍微复杂一些：

宏观态 ($N_{\uparrow}, N_{\downarrow}$)	(10,0)	(9,1)	(8,2)	(7,3)	(6,4)	(5,5)	(4,6)	(3,7)	(2,8)	(1,9)	(0,10)
微观态数 (Ω)	1	10	45	120	210	252	210	120	45	10	1

对应的分布如图??左图。

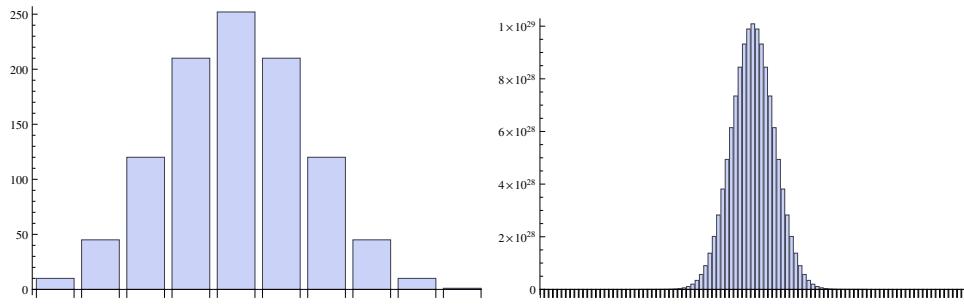


图 3.1: 左图: $N = 10$ 时不同宏观状态对应的微观状态数。横轴为正面向上的硬币数。右图: $N = 100$ 的情况。

如果我们把盒子里的硬币进一步增加到 100 个, 情况就大不相同了。一个宏观状态对应的微观状态数会差别很大:

宏观状态	(100,0)	(99,1)	(90,10)	(50,50)
对应微观状态数	1	100	$\sim 10^{13}$	$\sim 10^{29}$

所有可能的微观状态数为 2^{100} 。根据这些信息我们知道, 摆完盒子后, 出现 (50,50) 这一宏观状态的几率最大, 这是一个最为可能的分布, 具体的概率为 $C_{100}^{50}/2^{100}$ 。而出现 (100,0) 的概率最小, 只有 $1/2^{100}$ 。

依此类推, 随着硬币个数的增加, 各个宏观状态的概率分布差别会越来越大。根据中心极限定理, 概率分布越来越倾向于高斯分布。因为高斯分布是一个对称的分布, 系统最可能出现的宏观态 N_\uparrow (使得 $P(N_\uparrow)$ 最大化) 和系统平均可能出现的状态 $\langle N_\uparrow \rangle$ 是相同的。对于现在的硬币系统来说, 我们可以得到

$$P(N_\uparrow) = \sqrt{\frac{2}{N\pi}} \exp\left[-\frac{(N_\uparrow - N/2)^2}{N/2}\right]. \quad (3.5)$$

如果系统均值 $\langle N_\uparrow \rangle = N/2 \sim 10^{22}$, 那么考虑 N_\uparrow 稍微偏离均值, 比如 $\delta = N_\uparrow - N/2 \sim 10^{12}$, 对应的相对偏差为 $\sim 10^{-10}$, 其概率会降为均值概率的 10^{-44} 。也就是说, N_\uparrow 相对其均值偏差 10^{-10} 的概率为 10^{-44} 。

通过以上分析, 我们得出如下结论: (1) 我们通常关心的, 可以测量的, 是系统的宏观状态, 比如 $(N_\uparrow, N_\downarrow)$ 。(2) 尽管系统各个微观状态出现的概率相同, 但不同的宏观状态出现的概率远远不同, 因为它们对应的微观状态数不同。最可能出现的宏观状态对应的微观状态数最多。(3) 对于宏观尺度的系统, 基本可以确定系统处于对应微观状态数最多的宏观态, 对这一状态的偏离基本可以忽略。

我们考虑的热力学系统和盒子中的硬币是类似的。比如，把每个硬币换成一个自旋为 $1/2$ 的粒子，两个状态分别对应自旋的两种取向。那么给定一个宏观状态 $(N_\uparrow, N_\downarrow)$ ，系统的总自旋、能量等就可以确定。

举例：一个小组有 N 个人， q 分钱。现在需要把这 q 分钱分给这 N 个人，可以多少种分配方案？这里我们的宏观状态由 (N, q) 给出，每个分配方案对应一个微观状态。（答案： C_{q+N-1}^{N-1} ）

3.5.2 统计物理基本假设

统计物理中，我们对系统宏观和微观状态的描述和上面的硬币系统是非常类似的。统计物理理论是建立在下面的基本假设基础上的：

对于一个孤立热力学系统，在给定限制条件下（比如给定能量或粒子数或其他量），系统可以访问的每个可能微观状态出现的概率是相同的（the principle of equal a priori probability）。系统的内部动力学决定了在足够长时间内系统会遍历所有这些等概率的微观状态，而且在每一个状态所处的时间是相同的。

3.5.3 温度的统计定义

根据统计物理的基本假设，我们就从统计物理的角度来定义温度。我们考虑两个孤立系统1和2。开始时，它们分别处于平衡状态。现在将两个系统实现热接触，使得它们之间可以交换热量，但其他物理量都保持不变，比如体积、粒子数等。它们组成一个新的孤立系统，总系统的能量 $E = E_1 + E_2$ 固定不变，其中 E_1 和 E_2 分别为两个子系统的能量。总系统的宏观状态可以用 E_1 来描述。其初始宏观态为 E_1^0 ，而它可以访问的宏观态取 $0 \leq E_1 \leq E$ 。如果每个子系统对应的微观状态数分别用 $\Omega_1(E_1)$ 和 $\Omega_2(E_2)$ 表示的话，每个宏观态对应的微观状态数为

$$\Omega(E_1) = \Omega_1(E_1)\Omega_2(E_2),$$

而总的微观状态数为 $\sum_{0 \leq E_1 \leq E} \Omega(E_1)$ 。注意，这里我们忽略了1与2之间的相互作用对总能量以及可能微观状态数的贡献²。按照上面的假设，当整个系统达到热平衡后所处的宏观状态对应的微观状态数取最大值。这时

$$\frac{\partial}{\partial E_1} \Omega(E_1) = \frac{\partial}{\partial E_1} (\Omega_1(E_1)\Omega_2(E_2)) = 0. \quad (3.6)$$

²是否合理？

这样得到：

$$\Omega_2(E_2) \frac{\partial}{\partial E_1} \Omega_1(E_1) + \Omega_1(E_1) \frac{\partial}{\partial E_1} \Omega_2(E_2) = 0. \quad (3.7)$$

因为

$$dE_1 + dE_2 = 0,$$

所以当 1 和 2 达到热平衡的条件是：

$$\frac{\partial \ln \Omega_1(E_1)}{\partial E_1} = \frac{\partial \ln \Omega_2(E_2)}{\partial E_2}. \quad (3.8)$$

上式等号两边分别为描述系统 1 和 2 的量，它们具有相同的形式。系统达到热平衡后，两个系统的对应量相等。和热力学的分析类似，我们可以用这个量来定义温度

$$g(T_i) = \frac{\partial \ln \Omega_i(E_i)}{\partial E_i}. \quad (3.9)$$

这样，我们就从统计物理的角度引入了温度的概念，但 $g(T)$ 的具体形式无法确定。这里我们具有选择不同形式 $g(T)$ 的自由度。我们可以利用这个自由度，建立起与热力学温度的联系。

问题：我们根据条件 (4.4) 找到的 \hat{E}_1 对应的极值是极大值吗？请自己分析。

下面我们来分析 $g(T)$ 随 T 的变化关系。我们的出发点是 1 和 2 组成的整个孤立系统在演化至热平衡的过程中其微观状态数是逐渐增加的，因为根据我们的基本假设，热平衡时整个系统的微观状态数最多。假设系统 1 和 2 接触后，在向平衡态演化的过程中，一小部分热量 $0 < \Delta \ll E_1$ 从系统 1 传到了系统 2。我们有

$$\begin{aligned} \ln \Omega(E_1 - \Delta) &= \ln \Omega_1(E_1 - \Delta) + \ln \Omega_2(E_2 + \Delta) \\ &\approx \ln \Omega_1(E_1) + \ln \Omega_2(E_2) + \Delta \left(\frac{\partial \ln \Omega_2(E_2)}{\partial E_2} - \frac{\partial \ln \Omega_1(E_1)}{\partial E_1} \right). \end{aligned} \quad (3.10)$$

因为 $\ln \Omega(E_1 - \Delta) > \ln \Omega(E_1)$ ，所以

$$\Delta(g(T_2) - g(T_1)) > 0. \quad (3.11)$$

因为 $\Delta > 0$ ，要与热力学温度自洽就要求 $T_1 > T_2$ 。但上式同时给出 $g(T_1) < g(T_2)$ 。据此，我们知道 $g(T)$ 随 T 的增加而降低。

有关 $g(T)$ 的其他信息我们很难通过这样的普适性的分析得到了。人们发现 $g(T)$ 的如下形式

$$g(T) = \frac{1}{k_B T} = \frac{\partial \ln \Omega(E)}{\partial E}. \quad (3.12)$$

可以建立起统计物理与热力学的自洽联系。上式就是温度的统计物理定义，对应的温度 T 为

绝对温度，而 k_B 为 Boltzmann 常数。它把宏观状态参数 T 与系统给定能量下的微观状态数联系起来。我们以后会看到很多类似的定义。实际上，熵可以定义为

$$S = k_B \ln \Omega(E). \quad (3.13)$$

这就是 Boltzmann 熵的定义。这个定义镌刻在了 Boltzmann 的墓碑上。³ 利用这个定义，式 (3.12) 又可以写成

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right)_X, \quad (3.14)$$

与热力学中的结果相一致。这里需要说明两点：

- 我们可以看到，Boltzmann 常数的出现完全出于量纲的考虑。因为温度的概念和定义比统计物理出现的早很多，我们只能去与热力学定义看齐。如果从统计物理的角度出发的话，完全可以把温度定义为能量量纲，这样就不必出现 k_B 了。
- 我们在式 (3.14) 中，加入了下标 X 。这里 X 代表了所有在达到热平衡过程中保持不变的那些量，比如粒子数 N 和体积 V 等。

3.6 系综 (Ensemble)/Gibbs 系综

系综的概念是 Gibbs 首先引入到统计物理中来的，因此又称为 Gibbs 系综。引入系综的概念是因为一个给定的系统宏观状态对应很多微观状态。我们可以使用概率的方式对它们进行描述。比如系统处于某一个微观状态的概率。我们所测量的宏观物理量是这个量相对于所有这些微观状态的概率平均。为了方便，我们可以想象对把系统复制成很多相同的拷贝，每个拷贝对应一个微观状态，然后去分别考察每一个拷贝，最后对所有这些拷贝做平均。这一系列的拷贝就构成了一个系综 (Ensemble)。

统计物理中我们最为常用的系综有三种：

- 微正则系综 (microcanonical ensemble)：体系的能量固定为 E ，或者在 E 附近很小的范围 $[E, E + \Delta E]$ 。
- 正则系综 (canonical ensemble)：体系与一个大的热库相连接，用来固定其温度 T ，但其能量可以有涨落。
- 巨正则系综 (grand canonical ensemble)：与体系连接的热库还是一个粒子库，用于固定

³事实上，刻在 Boltzmann 墓碑上的公式为 $S = k \log W$ 。这里 W 代表德语 Wahrscheinlichkeit，也就是 probability 的意思。而这个表达式实际上是由 Planck 在 1900 年左右写下的。他写下这个式子时指出：“the logarithmic connection between entropy and probability was first stated by L. Boltzmann in his kinetic theory of gases.”

体系的温度和化学势，但能量仍然有涨落。

我们这里主要考虑正则系综。

3.7 Boltzmann 因子和配分函数

对于正则系综，因为和热库相连，所以系统温度固定，但能量不定。一个自然的问题就是系统处于不同状态的概率。我们这里的主要任务就是来确定这个概率。我们可以考虑一个大的系统，总能量为 E （或者为 $[E, E + \Delta E]$ ）。现在我们将它分为两个部分，一部分为系统，另外一部分作为热库。热库要比系统大很多，以致于系统能量的变化对热库的状态影响非常非常小。假设我们已知系统处于某一个微观状态 k ，对应的能量为 ε_k 。这时，热库的能量为 $E - \varepsilon(k)$ ，对应的微观状态数为 $\Omega(E - \varepsilon(k))$ 。根据等概率假设，系统处于状态 k 的概率正比于整个体系的微观状态数

$$P(\varepsilon(k)) \propto 1 \times \Omega(E - \varepsilon(k)).$$

因为 $E \gg \varepsilon(k)$ ，我们可以对 $\ln\Omega(E - \varepsilon(k))$ 做展开：

$$\begin{aligned} \ln\Omega(E - \varepsilon(k)) &\approx \ln\Omega(E) - \frac{\partial \ln\Omega(E)}{\partial E} \varepsilon(k) \\ &= \ln\Omega(E) - \frac{\varepsilon(k)}{k_B T}. \end{aligned} \quad (3.15)$$

于是有

$$P(k) \propto e^{-\beta\varepsilon(k)}, \quad (3.16)$$

这里 $\beta = 1/(k_B T)$ 。这就是我们的核心结果，指数因子 $e^{-\beta\varepsilon(k)}$ 称为 Boltzmann 因子。所以，系统处于状态 k 的概率正比于 Boltzmann 因子。

上面我们只是得到了系统处于状态 k 的概率正比于 Boltzmann 因子。要想得到正确的概率分布，我们还需要把它归一化，这需要对所有系统可能所处的状态求和。相应归一化之后的分布为

$$P(k) = \frac{1}{Z} e^{-\beta\varepsilon(k)}. \quad (3.17)$$

这里的归一化因子

$$Z = \sum_k e^{-\beta\varepsilon(k)}. \quad (3.18)$$

这个看似简单的归一化因子在统计物理中有一个自己专门的名字叫配分函数（Partition func-

tion)。顾名思义，它包含了系统被“分配”到每一个可能微观状态的概率。配分函数是统计物理中最为重要的函数。知道了配分函数，基本所有的热力学量都可以得到。

举例：自旋 $1/2$ 体系。在介绍微观状态和宏观状态时我们考虑的硬币的例子对应在物理学中就是自旋 $1/2$ 体系。这里我们继续考虑这个体系，只不过现在我们考虑两个状态的能量不再相同。考虑 N 个相互独立的自旋 $1/2$ 粒子处在某一个方向 (z 向) 的磁场 B 中。自旋倾向于平行磁场排布。当它与磁场平行时能量较低为 $-\mu B (\uparrow)$ ，当与磁场反向是能量较高 $\mu B (\downarrow)$ ，这里的 μ 为自旋磁矩，而这两个状态对应自旋 $1/2$ 在磁场中的两个本征态。下面我们考虑这一自旋体系的正则分布及相关性质。

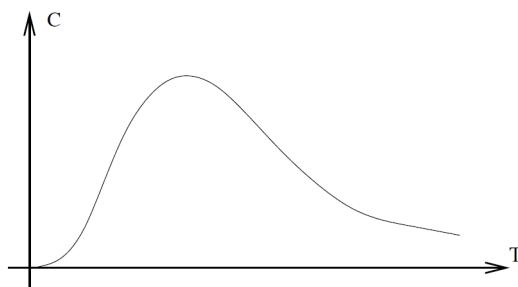
1. 如果只考虑一个自旋 $1/2$ ，那么它只有两种可能的微观状态。所以其配分函数为 $Z_1 = e^{\beta\mu B} + e^{-\beta\mu B} = 2 \cosh(\beta\mu B)$ 。其平均能量为 $E_1 = \mu B(e^{-\beta\mu B} - e^{\beta\mu B})/Z_1 = -\mu B \tanh(\beta\mu B)$ 。
2. 因为每个自旋之间是相互独立的，所以整个 N 自旋体系的平均能量为

$$E = \sum_i E_i = -N\mu B \tanh(\beta\mu B). \quad (3.19)$$

据此，我们还可以求出体系的热容：

$$C_N = \left(\frac{\partial E}{\partial T} \right)_N = Nk_B \frac{(\beta\mu B)^2}{\cosh^2(\beta\mu B)}. \quad (3.20)$$

可以看出，当 $T \rightarrow 0$ 时， C 随 T 以指数形式趋向于 0。当 $T \gg \mu B/k_B$ 时， C 随 T 以 T^{-2} 衰减向 0。在两者之间存在一个极大值，如下图所示。这种行为与电子和声子的热容有很大区别，这个极大值的出现称为 Schottky 反常。



第4章 热力学第一定律

4.1 上部分回顾：压强的统计定义 *

我们前面考虑了两个系统 1 和 2 通过热接触实现热平衡的情况。根据热平衡时对应的微观状态数最多的条件给出了温度的统计定义。采用类似的思路，我们可以给出压强的统计定义。这时，我们不光要考虑热平衡，还要考虑力学平衡。

我们考虑被活塞一分为二的一个容器，并定义活塞的左右两边为系统 1 和 2。假设活塞和器壁间没有任何摩擦，并且 1 和 2 之间可以通过活塞传递热量。而 1 和 2 组成的整个系统是孤立的。这样从某一个初始状态开始，1 和 2 组成的整个系统最终会达到平衡。这是 1 和 2 之间会同时实现热和力学平衡，即 $T_1 = T_2$ 而且 $P_1 = P_2$ 。

我们可以用 1 的能量 E_1 和体积 V_1 来表征整个系统 1+2 的宏观状态，因为对于整个系统能量和体积是守恒的 $E_1 + E_2 = E$, $V_1 + V_2 = V$ 。整个系统达到平衡后 $\Omega(E_1, V_1)$ 取最大值。这时考虑 E_1 和 V_1 的微小扰动，我们有

$$d\Omega(E_1, V_1) = \left(\left(\frac{\partial \Omega_1}{\partial E_1} \right)_{V_1} \Omega_2 - \left(\frac{\partial \Omega_2}{\partial E_1} \right)_{V_2} \Omega_1 \right) dE_1 \quad (4.1)$$

$$+ \left(\left(\frac{\partial \Omega_1}{\partial V_1} \right)_{E_1} \Omega_2 - \left(\frac{\partial \Omega_2}{\partial V_1} \right)_{E_2} \Omega_1 \right) dV_1. \quad (4.2)$$

根据平衡条件我们得到下面的关系式：

$$\left(\frac{\partial \ln \Omega_1}{\partial E_1} \right)_{V_1} = \left(\frac{\partial \ln \Omega_2}{\partial E_1} \right)_{V_2} \quad (4.3)$$

$$\left(\frac{\partial \ln \Omega_1}{\partial V_1} \right)_{E_1} = \left(\frac{\partial \ln \Omega_2}{\partial V_1} \right)_{E_2} \quad (4.4)$$

根据式 (4.4) 我们定义了温度，按照同样的思路我们可以根据式 (16.34) 来定义压强：

$$f(P, T) \equiv \left(\frac{\partial \ln \Omega}{\partial V} \right)_E \quad (4.5)$$

可以验证，如果取

$$f(P, T) = \frac{P}{k_B T},$$

我们可以得到和热力学相自洽的结果。例如，考虑到熵的定义 $S = k_B \ln \Omega$ 。我们可以把上面的

结果概括为：

$$dS(E, V) = \frac{1}{T} dE + \frac{P}{T} dV. \quad (4.6)$$

将上式稍作整理我们可以得出热力学第一定律形式的关系式：

$$dE = TdS - PdV. \quad (4.7)$$

4.2 热力学第一定律

前面我们介绍热力学第零定律说明了热平衡的可传递性。以此为基础我们可以引入温度的概念，并设计温度计来测量系统的温度。简言之，热力学第零定律引入了温度的概念。与此类似，我们现在介绍的热力学第一定律引入了能量的概念，尽管我们在这门课中早就开始使用这个概念了。同时，**热力学第一定律告诉我们能量有不同种的表现形式，它们之间可以相互转换，但总的能量是守恒的。**

考虑一个热力学系统，用 U 表示该系统的内能。它是系统的一个状态量。给定一个系统的平衡状态，其内能就是确定的。要使系统达到这个状态，可以有多种方法。比如我们可以通过传递热量的方式，也可以通过做功的方式。这两种方式都可以达到相同的效果，也就是说，在达到我们的目的方面，它们两个是等价的。一旦系统达到了我们制定的平衡态，单纯研究系统所处的这个平衡态，我们无法得知系统是如何达到这个状态的。这里，热量和功都是过程量，与系统到达平衡态的过程有关，只有得知了具体过程后我们才可以谈论这两个量。如果系统内能经历一个非常小的变化，上面的描述可以概括为下面的等式：

$$dU = dQ + dW. \quad (4.8)$$

这里，内能是状态量，所以 dU 是恰当微分；而热量和功都是过程量，所以 dQ 和 dW 都不是恰当微分，用 d 来表示。这里我们把外界对系统做功和系统从外界吸收热量定义为正。

外界对系统做功的形式可以有多种。对于我们熟悉的，用活塞封闭在容器中的理想气体模型，最常见的就是机械功。如果活塞和容器间无摩擦，而系统的体积元变化又非常缓慢的话，这个机械功可以写为：

$$dW = -PdV. \quad (4.9)$$

4.2.1 热容

前面我们定义了热容，并说明热容是一个依赖具体过程的量。下面我们根据热力学第一定律，再来分析一下。如果式(4.9)适用的话，

$$dQ = dU + PdV.$$

考虑内能为温度 T 和体积 V 的函数，我们可以得到：

$$dU = \left(\frac{\partial U}{\partial T}\right)_V dT + \left(\frac{\partial U}{\partial V}\right)_T dV.$$

综合上面的两个式子，我们得到：

$$dQ = \left(\frac{\partial U}{\partial T}\right)_V dT + \left(\left(\frac{\partial U}{\partial V}\right)_T + P\right) dV. \quad (4.10)$$

两边都除以 dT 我们有：

$$\frac{dQ}{dT} = \left(\frac{\partial U}{\partial T}\right)_V + \left(\left(\frac{\partial U}{\partial V}\right)_T + P\right) \frac{dV}{dT}. \quad (4.11)$$

据此，定容热容为

$$C_V = \left(\frac{\partial Q}{\partial T}\right)_V = \left(\frac{\partial U}{\partial T}\right)_V.$$

而定压热容为

$$C_P = \left(\frac{\partial U}{\partial T}\right)_V + \left(\left(\frac{\partial U}{\partial V}\right)_T + P\right) \left(\frac{\partial V}{\partial T}\right)_P.$$

可见，

$$C_P - C_V = \left(\left(\frac{\partial U}{\partial V}\right)_T + P\right) \left(\frac{\partial V}{\partial T}\right)_P \neq 0.$$

对于理想气体来说，其内能只与温度有关

$$U = \frac{3}{2}Nk_B T = \frac{3}{2}\nu RT.$$

利用其状态方程，考虑 1 摩尔的气体，我们有

$$C_P - C_V = R.$$

利用状态方程，我们可以得出 $C_V = \frac{3}{2}$, $C_P = \frac{5}{2}R$ 。我们可以定义一个绝热指数 $\gamma = \frac{C_P}{C_V} = \frac{5}{3}$ 。它的意义会在后面讲到。

4.3 常见热力学过程

系统处于平衡态时，我们可以用几个有限的热力学状态参量来对其进行描述。在热力学中，我们不光要关注系统所处的状态，也要关心系统外界驱动下在不同状态之间的演化。但是，系统在演化过程中，很可能不再处于平衡状态。这时，我们再不能用状态参量来描述系统在演化过程中的某一时刻的状态了。对于非平衡系统的描述是一件非常困难的事情。为了在平衡态的限制下来描述热力学系统的演化，我们只能对演化的过程进行理想化近似。这里我们考虑几种常见的理想演化过程。

4.3.1 准静态过程

我们假设系统的演化进行的非常缓慢。这里的缓慢是和系统的弛豫时间进行对比的。这时，假如我们对于系统参数做微小变化，系统可以快速做出响应，在我们再次对系统参数进行变化前重新达到新的平衡态。我们就可以假设系统在演化过程中的任何一个时刻都处于平衡状态。我们也就可以用热力学状态参量来描述演化过程中的任何一点，把这些点连起来就可以这个过程。这个过程就可以表示在以系统状态量为坐标轴的空间下表示出来。对于单元系统，我们只需要指定两个状态参量就可以确定其状态。这样的话，我们就可以在二维的坐标空间来描述准静态过程。

4.3.2 可逆过程

热力学中我们最经常考虑的过程是可逆过程。这里需要稍微区别一下热力学的可逆过程和组成系统的微观自由度运动方程的可逆性。我们知道的微观运动方程，不管是经典的牛顿运动方程还是量子的薛定谔方程都是可逆的。也就是说，如果我们把时间轴反向($t \rightarrow -t$)，运动方程的形式不会发生变化。所以我们常说微观运动方程是可逆的。另一方面，对于一个宏观的热力学系统的演化往往是不可逆的。比如，如果我们把容器内被压缩到角落里的空气在释放对其限制后会自发扩散到整个容器空间，但反向的过程却很难发生。那么，为什么微观的可逆运动方程会带来宏观的不可逆过程呢？这其实是统计物理需要回答的重要问题之一。我们学习热力学第二定律的时候再回到这个问题上来。

这里我们需要讨论的是理想的宏观可逆过程。顾名思义，可逆过程的意思就是我们可以在这个过程的任何时刻通过系统参数的微小变化把系统演化的方向改变，从而使系统回到最

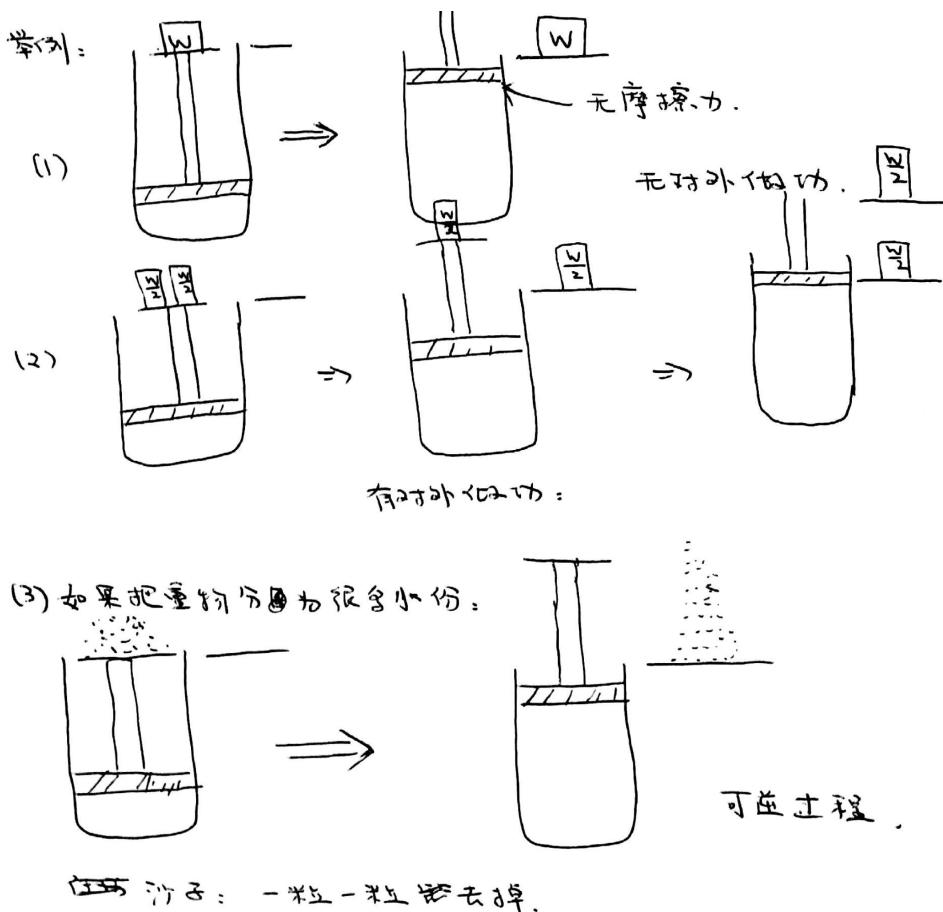


图 4.1: 例子 1

开始的状态，而且对外界环境没有任何影响。可逆过程首先需要时准静态过程，也就是说它要进行的无线慢。其次，在进行过程中不能有功转变为热的过程或者说不能有耗散。比如，我们如果通过活塞对容器内的理想气体做功的话，活塞和器壁之间不能有摩擦力。一旦有摩擦力，即使对于准静态过程，系统在压缩或膨胀过程中，需要施加的外力就不再由系统的压强直接决定，还要取决于摩擦系数。如果我们改变系统的演化方向，使之回到初始的状态的话，虽然系统恢复了初态，但部分体积功由于摩擦力的存在转化为了热量。这个过程就不是可逆的。

问题：下面的说法正确吗？

- 根据上面的描述，可逆过程不允许有功转化为热的过程。
- 可逆过程不允许有热量的交换。

为方面我们对上面描述的理解，下面考虑两个具体的例子：

- 见下图 1
- 将系统温度从 T_1 升高至 T_2 (图 2)

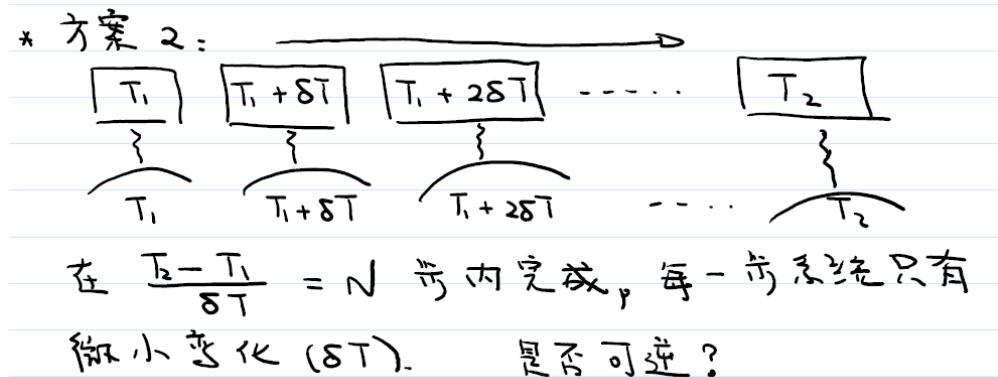
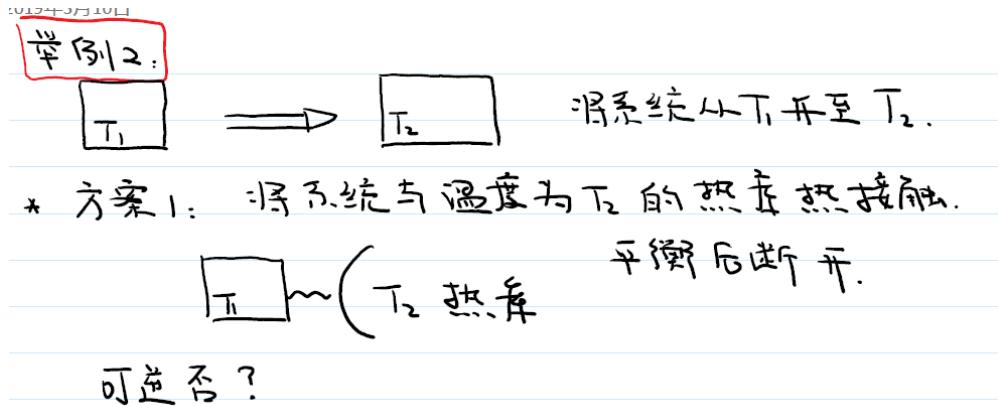


图 4.2: 例子 2

我们前面介绍热力学第一定律时说到, 式 (4.8) 右边的传热和做功两项都是过程量。给定系统的初态和末态, 尽管传热和做功之和可以确定, 但要分别确定它们每一项, 需要给出具体的过程。也就是说, 要把系统从一个状态变到另外一个状态, 可以有无限中可能的过程来实现。每个过程中做功和传热的数值可能都不一样。更糟的是, 这个过程可能是一个非平衡系统, 我们根本无法简单的通过系统的状态参量来描述。但是, 如果我们对这些过程做一些限制的话, 是有可能使用系统的状态量来描述做功和热传递过程的。比如, 对于体积功, 我们给出了式 4.9。可见, 对于可逆过程, 我们就可以用系统的状态参量来描述做功。如果我们将可逆过程做更进一步的限制的话, 这个功就可能完全确定下来。如果 dU 和 dW 都确定了下来, 那么 dQ 也就确定了。下面我们以理想气体为例, 来分析两个常见的可逆过程。

4.3.3 等温过程

等温过程指系统温度保持不变的可逆过程。我们考虑理想气体的等温膨胀。理想气体的内能只与温度有关，所以等温膨胀过程中内能不变。由于膨胀对外做功，所以它必然需要从外界吸收热量，即

$$dU = dQ + dW = 0.$$

我们可以计算 1 摩尔理想气体从体积 V_1 膨胀为 V_2 所吸收热量的多少：

$$Q = \int_1^2 dQ = \int_1^2 PdV = RT \int_1^2 \frac{1}{V} dV = RT \ln \frac{V_2}{V_1}.$$

4.3.4 绝热过程

绝热过程是与外界无热量交换的可逆过程，即 $dQ = 0$ 。这时的热力学第一定律为

$$dU = dW = -PdV.$$

对于理想气体，根据 $dU = C_V dT$ 我们有

$$C_V dT = -PdV,$$

再根据状态方程我们得到：

$$\frac{C_V}{T} dT = -\frac{R}{V} dV.$$

据此，我们可以得出理想气体绝热过程的过程方程：

$$TV^{\gamma-1} = \text{Constant}$$

或者

$$PV^\gamma = \text{Constant}$$

这里我们就明白为什么我们把 γ 叫绝热指数了。

举例：绝热大气 举例：Carnot 循环

4.4 恰当微分 (exact differential)

4.4.1 定义

对于一个给定函数 $z = f(x, y)$, 其微分形式可以写成:

$$dz = u(x, y)dx + v(x, y)dy. \quad (4.12)$$

如果满足条件

$$\frac{\partial u(x, y)}{\partial y} = \frac{\partial v(x, y)}{\partial x}, \quad (4.13)$$

那么我们称方程 (4.12) 为一个恰当微分。比如, 如果我们有

$$dz = \left(\frac{\partial f}{\partial x} \right) \Big|_y dx + \left(\frac{\partial f}{\partial y} \right) \Big|_x dy, \quad (4.14)$$

那么, 只要 $f(x, y)$ 是连续可微的, 那么

$$\frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \Big|_y \right) \Big|_x = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \Big|_x \right) \Big|_y. \quad (4.15)$$

方程 (4.14) 就是一个恰当微分。

4.4.2 性质

恰当微分具有以下性质:

1. 对于恰当微分,

$$\int_{(x_1, y_1)}^{(x_2, y_2)} dz = \int_{(x_1, y_1)}^{(x_2, y_2)} u(x, y)dx + v(x, y)dy = f(x_2, y_2) - f(x_1, y_1),$$

恰当由积分起始和终止的位置决定, 与积分路径无关。

2. 环路积分为零,

$$\oint_{\text{closed}} dz = \oint_{\text{closed}} u(x, y)dx + v(x, y)dy = 0.$$

3. 已知 dz , 除了一个常数外, 我们可以确定 $z(x, y)$ 。

4.4.3 拓展

上面的定义可以拓展到多个变量的情况。比如对于一个函数 $f(x_i)$, $i = 1 \dots N$, 其微分形式写为

$$df(x_i) = \sum_m g_m(x_i)dx_m. \quad (4.16)$$

如果对于任何 m, n , 条件

$$\frac{\partial g_n}{\partial x_m} = \frac{\partial g_m}{\partial x_n}, \quad (4.17)$$

都满足, 那么 $df(x_i)$ 为一个恰当微分。

举例:根据上面的定义, 我们很容易验证 (1). $df = d(xy) = ydx + xdy$ 是一个恰当微分。 (2). $df = ydx$ 不是一个恰当微分。 (3). 对于 (1), 验证从 i 到 f 的两个路径①和②+③得出的积分是相等的。

第5章 热力学第二定律

5.1 Carnot 循环

利用我们前面介绍的两种可逆过程，我们可以构成一个热力学中最为重要的可逆循环过程：Carnot 循环。它是由两个等温和两个绝热过程组成，如下图所示。整个循环分四个部分：

- $A \rightarrow B$ ：工作物质从高温热源（设温度为 T_h ）吸收热量 Q_h 并做等温膨胀
- $B \rightarrow C$ ：切断与热源的热接触，工作物质做绝热膨胀，对外做功
- $C \rightarrow D$ ：工作物质与低温热源（温度为 T_l ）接触，被等温压缩，向低温热源传递热量 Q_l
- $D \rightarrow A$ ：切断与热源接触，做绝热压缩，回到初态 A ，完成一个循环

Carnot 循环可以用做热机，把热量转换为功。其效率可以定义为总输出功 W 与从高温热源吸收的热量之比

$$\eta = \frac{W}{Q_h}.$$

利用第一定律我们有 $W + Q_l = Q_h$ ，所以效率又可以写为

$$\eta = \frac{W}{Q_h} = 1 - \frac{Q_l}{Q_h}.$$

如果我们考虑以理想气体作为工作物质的话，很容易计算出，Carnot 热机的效率

$$\boxed{\eta_C = 1 - \frac{Q_l}{Q_h} = 1 - \frac{T_l}{T_h}.} \quad (5.1)$$

练习：试以理想气体为工作物质，证明上式 (5.1)。

这里我们需要说明的是， $A \rightarrow B$ 和 $C \rightarrow D$ 为等温可逆过程。在这两个过程中工作物质的温度需要与热源温度无限接近，否则这个过程就不再是准静态过程，也就不再是可逆过程。

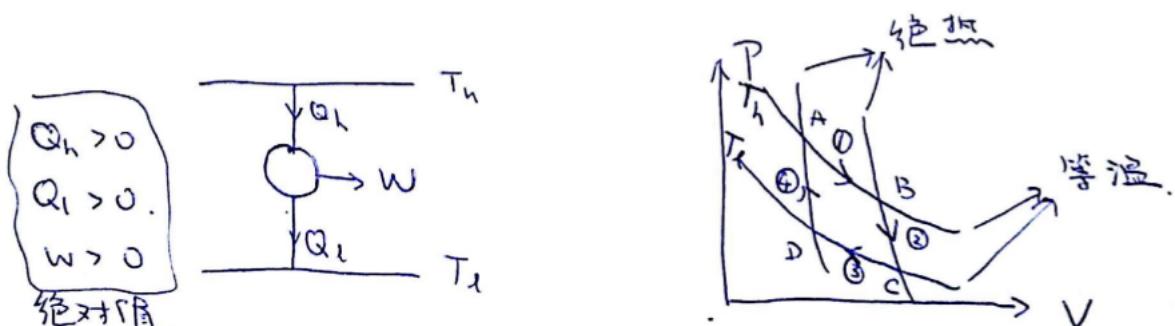


图 5.1: Carnot 循环。

由于这个要求，工作物质和热源之间的热量交换非常缓慢，因此完成一个 Carnot 循环的时间会非常长。这也是可逆过程所要求的。如果我们定义热机的功率为 $P = W/\tau$ (τ 为完成一个循环所需时间) 的话，不难想象，Carnot 热机的功率 $P_C \rightarrow 0$ 。现实生活中，我们更关心的是在最大输出功率时的效率。这个效率由 Curzon&Ahlborn 给出¹

$$\eta_{CA} = 1 - \sqrt{\frac{T_l}{T_h}}. \quad (5.2)$$

这也是一个只依赖热库温度的量。显然，功率输出必须要牺牲效率。详细的分析可以参考下面的联系以及 Curzon 和 Ahlborn 的原始论文。

练习：(from Schroeder)

Problem 4.6. To get more than an infinitesimal amount of work out of a Carnot engine, we would have to keep the temperature of its working substance below that of the hot reservoir and above that of the cold reservoir by non-infinitesimal amounts. Consider, then, a Carnot cycle in which the working substance is at temperature T_{hw} as it absorbs heat from the hot reservoir, and at temperature T_{cw} as it expels heat to the cold reservoir. Under most circumstances the rates of heat transfer will be directly proportional to the temperature differences:

$$\frac{Q_h}{\Delta t} = K(T_h - T_{hw}) \quad \text{and} \quad \frac{Q_c}{\Delta t} = K(T_{cw} - T_c).$$

I've assumed here for simplicity that the constants of proportionality (K) are the same for both of these processes. Let us also assume that both processes take the same amount of time, so the Δt 's are the same in both of these equations.*

- (a) Assuming that no new entropy is created during the cycle except during the two heat transfer processes, derive an equation that relates the four temperatures T_h , T_c , T_{hw} , and T_{cw} .
- (b) Assuming that the time required for the two adiabatic steps is negligible, write down an expression for the power (work per unit time) output of this engine. Use the first and second laws to write the power entirely in terms of the four temperatures (and the constant K), then eliminate T_{cw} using the result of part (a).
- (c) When the cost of building an engine is much greater than the cost of fuel (as is often the case), it is desirable to optimize the engine for maximum power output, not maximum efficiency. Show that, for fixed T_h and T_c , the expression you found in part (b) has a maximum value at $T_{hw} = \frac{1}{2}(T_h + \sqrt{T_h T_c})$. (Hint: You'll have to solve a quadratic equation.) Find the corresponding expression for T_{cw} .
- (d) Show that the efficiency of this engine is $1 - \sqrt{T_c/T_h}$. Evaluate this efficiency numerically for a typical coal-fired steam turbine with $T_h = 600^\circ\text{C}$ and $T_c = 25^\circ\text{C}$, and compare to the ideal Carnot efficiency for this temperature range. Which value is closer to the actual efficiency, about 40% of a real coal-burning power plant?

上面，我们利用理想气体为工作物质得到了 Carnot 热机的效率。如果更换工作物质的话，Carnot 热机的效率是否会改变呢？要回答这个问题，我们需要用到热力学第二定律。

¹F. L. Curzon and B. Bhlborn, American Journal of Physics, 41, 22 (1975).

5.1.1 冰箱：

如果把热机反向运行，就可以实现制冷。这就是冰箱的原理。冰箱的工作效率改如何定义？它的效率可以大于 1 吗？

5.1.2 热泵：

如果把热机反向运行，还可以将热量泵浦到高温的热源，实现热泵。类似制热空调。这时，我们应该如何定义空调的效率？它可以大于 1 吗？为什么空调比电暖的效率高？

练习：你还知道哪些类型的循环？

5.2 热力学第二定律

- Clausius 描述：不可能有这样的过程，其唯一结果是热量从低温物体传到高温物体。
- Kelvin 描述：不可能有这样的过程，其唯一结果是将热全部转化为功。

5.3 Carnot 定理

- 工作于相同两个热源之间的所有可逆热机的效率相同，都等于 η_C
- 工作于相同两个热源之间的不可逆热机的效率低于可逆热机的效率
- 利用 Carnot 循环可以证明 Clausius 和 Kelvin 描述的等价性
- Carnot 定理可以看做是热力学第二定律的另外一种等价描述

5.4 Clausius 定理

对于任意一个循环

$$\oint dQ/T \leq 0,$$

等号对应可逆循环。Clausius 定理可以作为可逆循环的一个定义。练习：试证明 Clausius 定理。

Carnot 定理、Clausius 定理可以作为热力学第二定律的其他表述形式。

扩展阅读：我们这里考虑的热机、热泵或冰箱的工作物质都是传统的热力学系统，即包含了无穷多的自由度。最近一些年，人们对由量子系统作为工作物质的热机产生了很大的兴趣。一个很显然的问题是，如果引入量子力学的一些特征，热机的效率是否可以高于 Carnot 效率？这里列出几篇相关的综述论文供大家阅读：

- 孙昌璞，全海涛，董辉，量子信息启发的量子统计和热力学若干问题研究，《量子力学新进展（第五辑）》，第 1 章。
- Sai Vinjanampathy and Janet Anders, Quantum Thermodynamics, Contemporary Physics, 57, 545 (2016).
- J. Roßnagel, et al., A single atom heat engine, Science, 352, 325 (2016).

第6章 熵

“Entropy is an anthropomorphic concept.”

from E. P. Wigner told by E. T. Jaynes.

6.1 热力学熵

6.1.1 定义

根据 Clausius 定理，对于可逆循环

$$\oint dQ_{\text{rev}}/T = 0,$$

也就是对于所有的可逆过程，积分

$$\int_A^B dQ_{\text{rev}}/T$$

都相同，与具体可逆过程的路径无关。据此，我们可以定义一个描述系统状态的量 S 使得

$$S_B - S_A \equiv \int_A^B dQ_{\text{rev}}/T.$$

我们把它叫做熵 (entropy)。它是一个状态量。

“... I propose to call the magnitude S the entropy of the body... I have intentionally formed the word entropy so as to be as similar as possible to the word energy; for the two magnitudes to be denoted by these words are so nearly allied in their physical meanings, that a certain similarity in designation appears to be desirable.”

R. Clausius

有了熵的定义之后，我们可以证明，对于任意一个连接状态 A 和 B 的系统

$$\Delta S_{BA} = S_B - S_A \geq \int_A^B \frac{dQ}{T}. \quad (6.1)$$

等号对应可逆过程。我们从可逆过程出发得出了这个关系。反过来，我们也可以把可逆过程做如下定义：如果一个过程保持整个宇宙（系统 + 环境）的熵不变，那么它就是一个可逆过程。

以上结果的一个推论就是：对于一个孤立系统 $dQ = 0$ ，所以 $\Delta S \geq 0$ 。这是我们更加熟悉

的热力学第二定律的一种描述：孤立系统的熵不会减少。

6.1.2 热力学定律

根据第一定律我们有

$$dU = dQ + dW, \quad (6.2)$$

而对于可逆过程我们根据上节的定义有

$$dQ_{\text{rev}} = SdT,$$

同时，如果只考虑体积功，对于可逆过程我们还有

$$dW = -PdV.$$

把这些结果重新整理，我们可以得到

$$dU = TdS - PdV. \quad (6.3)$$

这是一个最基本的热力学关系式。与式 (6.2) 不同的是，这个式子中所有量都是状态量，不涉及具体的过程。

把式 (6.3) 稍作整理我们可以得到

$$dS = \frac{1}{T}dU + \frac{P}{T}dV.$$

我们把熵 S 作为 U 和 V 的函数。于是我们可以得到

$$\begin{aligned} \left(\frac{\partial S}{\partial U}\right)_V &= \frac{1}{T}, \\ \left(\frac{\partial S}{\partial V}\right)_U &= \frac{P}{T}. \end{aligned}$$

这和我们前面给出的 T 和 P 的统计定义相同。这也说明我们之前的统计定义与热力学定律相符合。

举例： 谱本 14.1

6.1.3 焦耳膨胀

下面我们来计算焦耳膨胀过程中的熵变。焦耳膨胀前，理想气体的体积为 V ，压强为 P ，温度为 T 。当气体从体积 V 膨胀到 $2V$ 后，其压强变为 $P/2$ ，因为膨胀过程中没有与外界的热量交换，外界也没有对系统做功，所以其内能和温度都没有变化。

因为焦耳膨胀不是一个可逆过程，我们无法使用熵的定义来计算。我们需要设计一个可逆过程来把气体的初态和末态连接起来。等温可逆膨胀可以实现这个目的。对于等温可逆膨胀，系统吸收的热量等于它对外所做的功。这样我们就可以计算系统的熵变：

$$\Delta S_{\text{sys}} = \int \frac{dQ_{\text{rev}}}{T} = \int_V^{2V} P dV = R \ln 2.$$

因为焦耳膨胀过程中外界没有任何变化，所以环境的熵变为零 $\Delta S_{\text{env}} = 0$ 。于是整个宇宙的熵变为

$$\Delta S_{\text{universe}} = \Delta S_{\text{sys}} + \Delta S_{\text{env}} = R \ln 2 > 0.$$

可见，焦耳膨胀是一个不可逆过程，因为整个宇宙的熵增加了。

如果希望把焦耳膨胀之后的系统还原到原来的状态，可以对其他进行一个等温可逆压缩的过程。这时，环境对系统做功，系统对环境释放热量，所以系统的熵变 $\Delta S_{\text{sys}} = -R \ln 2$ 。因为环境吸收了热量，所以它的熵增加 $\Delta S_{\text{env}} = R \ln 2$ 。整个压缩过程中，宇宙的熵保持不变。整个循环完成之后，系统回到了原来的状态，但一部分 $R \ln 2$ 的功被转化为热量。这是一个不可逆循环。相反，如果等温膨胀和等温压缩组成一个循环，整个过程宇宙的熵变就是零。

6.2 Boltzmann 熵

我们前面已经介绍了 Boltzmann 熵的定义：

$$S_B(X) \equiv k_B \ln \Omega(X). \quad (6.4)$$

这里， k_B 是 Boltzmann 常数，而 $\Omega(X)$ 是给定宏观状态（用一组状态参数 X 来描述）对应的微观状态数。在前面讲温度和压强的统计定义时，我们其实已经使用了这个定义。

这里需要再次强调一下微观状态和宏观状态。在前面硬币的粒子中，给定 N 硬币组成的系统，我们把包含每个硬币信息（正/反，或者 0/1）的状态称为一个微观态；而把用正面朝上硬币的数目 N_\uparrow 表示的状态称为一个宏观态。定义了这两种状态之后，我们才可以讨论 Boltzmann 熵。比如，

$$S_B(N_\uparrow) = k_B \ln \frac{N!}{N_\uparrow!(N - N_\uparrow)!}.$$

这里需要强调的是，它们的选取是人为的、相对的。一种情况下我们认定的微观态在另外一种情况下可能对应一个宏观态。如果我选取硬币的势能作为一个参数来确定宏观态的话，那么不管硬币是正面向上还是反面向上，其势能都是一样的。所以不管硬币处于什么样的状态，

其势能是不变的。这时，Boltzmann 熵变成了

$$S_B(V) = k_B \ln 2^N.$$

这个由势能确定的宏观态就包含了所有可能的微观状态。对于我们关心的物理系统也是这样的，通常我们在讨论系统在某一状态的熵时，需要给出定义这个宏观状态的热力学参量。微观状态的选取也是类似的。比如，我们在定义硬币系统的微观状态时并没有考虑组成硬币的原子的状态。其实一个向上的硬币，如果我们去考虑组成硬币的更小尺度的原子的话，可以对应非常多的更加微观的状态。我们在定义 N 硬币系统时，是否需要考虑这些更加微观的状态呢？更进一步，即使我们考虑了组成硬币的每个原子的状态，我们是否需要再考虑原子中原子核和电子的状态？是否还需要考虑组成原子核的粒子的状态？显然，考虑与否直接影响到熵的大小。但如果在系统变换前后这些更加微观的状态不发生变化的话，不会影响到变换过程中的熵变（为什么？参见例题 14.5）。因此，不考虑它们不会给我们带来什么麻烦。

现在我们利用 Boltzmann 熵来重新计算一下焦耳膨胀过程中 1 摆尔理想气体的熵变。我们选取 U 和 V 作为我们的宏观参量。气体在膨胀前后内能不变，但体积变为原来的 2 倍。根据 Boltzmann 熵的定义，熵变可以写为

$$\Delta S = k_B (\ln \Omega(U, 2V) - \ln \Omega(U, V)) \quad (6.5)$$

$$= k_B \ln \Omega(U, 2V) / \Omega(U, V) \quad (6.6)$$

$$= k_B N_A \ln 2 = R \ln 2. \quad (6.7)$$

因为，在气体膨胀之后，每个分子可取的空间状态变为了原来的 2 倍，所以对应的微观状态数是膨胀前的 2^{N_A} 倍。这个结果和我们利用热力学得到的结果是一样的。

6.2.1 混合熵

使用同样的办法，我们可以考虑两种不同的气体混合所带来的熵变。加入初始状态为两种物质的量分别为 xN_A 和 $(1-x)N_A$ 的气体分别装在左右两个体积为 xV 和 $(1-x)V$ 的容器内。它们的温度相同。现在把隔在两者之间的挡板去掉，它们会发生混合。混合之后的熵变为

$$\Delta S = -R [(1-x) \ln(1-x) + x \ln x] > 0. \quad (6.8)$$

显然，当 $x = 1/2$ 时，熵变最大为 $R \ln 2$ 。

6.2.2 Gibbs 佯谬

前面我们考虑混合熵时，两个容器所装气体不同。如果我们考虑两边装有相同气体的话，单独考虑每一部分气体熵变的话，我们会得到与不同气体时一样的结果，即式 (6.8)。但事实是，因为两个容器中的气体相同，挡板是否存在好像并不影响其状态。虽然气体分子之间有混合，但我们并没有办法区分每个分子，因为它们都是相同的。所以直觉上系统的熵应该没有变化。这就产生了一个矛盾。这就是著名的 Gibbs 佯谬。[可见，我们是否可以区分分子对熵变非常重要。](#) 后面第 10 章我们会详细讨论这个问题。

6.3 Gibbs 熵

6.3.1 概率表示

Gibbs 在系综理论的基础上，给出了熵的另外一种表示形式。下面我们做一个简单介绍。根据系综理论，我们想象把关心的系统复制出完全相同的 N 个，组成一个系综。我们可以使 N 足够大，远大于系统所有可取的微观态数目。在平衡状态下，系统处于第 i 个微观状态的概率如果是 P_i 的话，那么 N 个系统中就会有 NP_i 个系统处于第 i 个微观态。由这个 N 个系统中每个微观态出现的次数 NP_i 来给定的这个宏观状态对应的微观状态数是

$$\Omega(P_i) = \frac{N!}{\prod_i (NP_i)!} \quad (6.9)$$

对应的熵就是

$$S_G = \frac{1}{N} k_B \ln \Omega(P_i) \quad (6.10)$$

$$= \frac{1}{N} k_B (\ln N! - \sum_i \ln (NP_i)!) \quad (6.11)$$

$$\approx -k_B \sum_i P_i \ln P_i. \quad (6.12)$$

系统的 Gibbs 熵就定义为：

$$S_G = -k_B \sum_i P_i \ln P_i.$$

讨论：

- Gibbs 熵使用系统处于每一个微观状态的概率来表示的。它的适用范围更广，不只适用于宏观热力学系统，也适用于其他小系统；不只适用于孤立体系，也适用与其他体系。

微正则系综：系统处于每个微观态的概率都相同 $P_i = 1/\Omega$ ，其中 Ω 为总的微观状态数。

那么 $S_G = k_B \ln \Omega = S_B$ 。看来 S_B 可以看作 S_G 的一种特殊情况。

- 正则系综：这时我们知道

$$P_i = \frac{1}{Z} e^{-\beta E_i}.$$

代入到定义中我们有

$$S_G = -k_B \sum_i \frac{1}{Z} e^{-\beta E_i} (-\beta E_i - \ln Z) \quad (6.13)$$

$$= k_B (\beta \langle E \rangle + \ln Z) \quad (6.14)$$

$$= k_B \frac{\partial T \ln Z}{\partial T}. \quad (6.15)$$

6.3.2 另一种考虑方式

我们还可以从另外一个角度来考虑这个问题。就像从微正则系统得到正则系统的情况一样，我们可以把我们关心的系统以及与系统耦合的热库作为一个大的孤立系统，其熵为

$$S_T = \ln \Omega_T(E).$$

如果我们将热库的状态不关心，就可以考虑定义系统的熵。这时，我们要求 S_T 是系统的熵和热库的熵的总和

$$S_T = S_S + S_B$$

也就是说，熵满足叠加性。两个弱耦合系统的总熵由每个系统熵的总和来得到。此外，由于热库的状态不确定，我们把它写成一个统计平均的形式

$$S_B = \sum_{\epsilon} P(\epsilon) \ln \Omega_B(E - \epsilon).$$

根据上面的条件，我们可以得到系统熵的表达式

$$\begin{aligned} S_S &= \ln \Omega_T(E) - \sum_{\epsilon} P(\epsilon) \ln \Omega_B(E - \epsilon) \\ &= \sum_{\epsilon} P(\epsilon) \ln \frac{\Omega_B(E - \epsilon)}{\Omega_T(E)} \\ &= - \sum_k P(k) \ln P(k). \end{aligned} \quad (6.16)$$

这里 k 表示系统的微观状态，而 ϵ 表示系统的能量。我们有

$$P(\epsilon) = \Omega_S(\epsilon) P(k_{\epsilon})$$

这里 k_ϵ 表示能量为 ϵ 的微观态 k 。

6.4 Shannon 熵

Shannon 在研究信息学时，给出了定量描述一条信息的信息量大小的量度¹。现在我们称为 Shannon 熵，它的表达形式和 Gibbs 熵几乎相同

$$S_{Sh} = - \sum_i P_i \log_2 P_i. \quad (6.17)$$

只不过这里的对数函数是以 2 为底的。

为什么熵也可以用来定量描述信息量的大小呢？我们书上举了一个描述牛顿生日这件事情的几条信息：牛顿的生日（1）是全年中的某一天；（2）是下半年；（3）是某月的 xx 号。这三条信息正确的概率分别为 $1, 1/2, 1/12$ 。那么哪条信息的信息量最大呢？显然是最后一条。所以一条信息正确的概率越小，它包含的信息量就越大。另外，信息（2）和（3）同时成立的概率为 $1/2 * 1/12 = 1/24$ 。我们希望相互独立的两条信息的信息量可以相加，所以我们可以自然的联想到使用对数函数，即 $S_i \propto -\log P_i$ 。为了处理二进制的信息，选择以 2 为底的对数会比自然对数来得方便。比如，如果一条信息正确的概率是 $1/2$ ，那么它包含的信息量是最大的，就是 1 个比特。

Shannon 给出式 (6.17) 是出于下面的考虑。给定一组概率分布 $\{P_i\}$ ，他要求给出的表达式 $H(\{P_i\})$ 满足几个条件：

1. H 是 P_i 的连续函数。如果 $\Omega_i > \Omega_{i'}$ ，那么 $H(\Omega_i) > H(\Omega_{i'})$ 。
2. 如果改变 $\{P_i\}$ 的顺序， H 需要保持不变
3. 对于任何一组 $\{P_i\}$ ，式 $H(\{P_i\}) \leq H(\frac{1}{Q}, \dots, \frac{1}{Q})$ 总成立。也就是说，当所有的 P_i 都相等时， H 最大。
4. H 具有可加性 (additive)，并且与如何组合微观态无关。

以上几点中，最后一点需要一些解释。blahblah... Shannon 证明了，同时满足以上几个条件的函数只能具有式 $H = -k \sum_i P_i \ln P_i$ 的形式。我们没有其他选择。可以说，Shannon 的理论建立了两个原来毫不相干的学科，即热力学和信息论，之间的关系。后人把这个联系发扬光大，发展了所谓的最大熵原理。

注：Shannon 并没有学过热统，所以他写出上式的时候也不知道热统里面已经有了和他的

¹The Bell System Technical Journal, 27, 379–423; 27, 623–656 (1948).

式子相同的表达式，即 Gibbs 熵。Shannon 当时是 Princeton 的一个博士后。他得出这个结果后，去咨询 von Neumann。是 von Neumann 建议 Shannon 把他的表达式叫做熵。von Neumann 给了他两个理由：(1) 统计物理里面已经有这个名字了；(2) 熵的名字听起来高大上一些，而且没人真正知道熵是什么东西。所以别人要和你争论的话，你可以占据主动。

6.5 von Neumann 熵

将 Shannon 熵推广到量子力学的情况下我们就得到了 von Neumann 熵

$$S_v = -\text{Tr}\hat{\rho}\ln\hat{\rho}. \quad (6.18)$$

其中， $\hat{\rho}$ 为系统的密度矩阵。它可以表示为：

$$\hat{\rho} = \sum_i P_i |i\rangle\langle i|.$$

这里， P_i 表示系统处于状态 $|i\rangle$ 的概率。

- 如果 $|i\rangle$ 为系统的本征态，那么 $\hat{\rho}$ 就可以用对角矩阵表示，其第 i 个对角元对应 P_i 。于是我们知道 $\text{Tr}(\hat{\rho}) = 1$ 。在对角表象下，von Neumann 熵就可以写成和经典的 Gibbs 熵相同的形式：

$$S_v = -\sum_i P_i \ln P_i.$$

- 如果 $P_m = 1, P_n = 0 (n \neq m)$ ，那么系统处于纯态；否则系统处于混态。
- 我们很容易证明 $1 \geq \rho_{ii} \geq 0$ 。所以， $\text{Tr}(\hat{\rho}) \geq \text{Tr}(\hat{\rho}^2)$ 。
- 如果系统与一个热库达到热平衡，它必然处于混态。对应的密度矩阵称为热密度矩阵：

$$\hat{\rho}_{\text{th}} = \sum_i e^{-\beta E_i} |i\rangle\langle i| = e^{-\beta \hat{H}}$$
。
- 算符 \hat{O} 的期待值可以表示为 $\langle \hat{O} \rangle = \text{Tr}(\hat{\rho}\hat{O})$ 。

6.5.1 用密度矩阵来计算算符的期待值

任意一个算符的期待值可以表示为：

$$\langle O \rangle = \sum_i P_i \langle i | \hat{O} | i \rangle.$$

我们可以把算符 \hat{O} 在它的本征表象下表示为

$$\hat{O} = \sum_{\alpha} O_{\alpha} |\alpha\rangle\langle\alpha|$$

代入上式我们有：

$$\langle O \rangle = \sum_i P_i \langle i | \sum_{\alpha} O_{\alpha} |\alpha\rangle \langle \alpha | i \rangle = \sum_{\alpha} O_{\alpha} (\sum_i P_i \langle i | \alpha \rangle \langle \alpha | i \rangle) = \sum_{\alpha} O_{\alpha} \rho_{\alpha\alpha} = \text{Tr}(\hat{O} \hat{\rho})$$

上式第二个等号后面小括号内包含了系统处于第 i 个态的概率 P_i ，以及第 i 个态投影到 \hat{O} 的本征态 α 的概率。第一个概率是一个统计概率，而第二个概率是量子概率。

6.5.2 用 von Neumann 熵表示量子纠缠

量子比特之间的纠缠可以用熵来度量。一个量子比特可以处于两个状态，我们可以用 $| \uparrow \rangle$ 和 $| \downarrow \rangle$ 来表示。我们考虑由两个量子比特组成的系统。如果系统处于一个纯态 $|\psi_{\pm}\rangle$ 表示为

$$|\psi_{\pm}\rangle = \frac{1}{\sqrt{2}} (| \uparrow \downarrow \rangle \pm | \downarrow \uparrow \rangle).$$

这时系统的密度矩阵为

$$\hat{\rho}_t = |\psi_{\pm}\rangle \langle \psi_{\pm}|.$$

显然，这时系统的 von Neumann 熵为 0。

现在，如果我们不考虑第 2 个量子比特，而只专心第一个量子比特。我们可以对第 2 个比特求迹得到约化密度矩阵

$$\rho_1 = \text{Tr}_2(\hat{\rho}_t) = \frac{1}{2} (| \uparrow \rangle \langle \uparrow | + | \downarrow \rangle \langle \downarrow |).$$

这时，因为我们失去了第 2 个量子比特的相关信息，约化密度矩阵的 von Neumann 熵变为了 1。这是一个量子比特可取熵的最大值。同样的道理，我们可以计算出

$$\rho_2 = \text{Tr}_1(\hat{\rho}_t) = \frac{1}{2} (| \uparrow \rangle \langle \uparrow | + | \downarrow \rangle \langle \downarrow |).$$

这时，我们说量子比特 1 和 2 之间的纠缠度最高。它们之间的纠缠可以体现在如下方面：对于 $|\psi_{\pm}\rangle$ ，如果我们已知第 1 个量子比特的状态为 \uparrow ，那么不需要测量，我们可以立刻知道第 2 个量子比特处于 \downarrow 状态。同理，如果第 1 个量子比特处于 \downarrow 状态，那么第 2 个必然处于 \uparrow 状态。

注：利用量子力学中 Schmidt 分解的方法，我们可以把由 A 和 B 两部分组成的一个系统的纯态 $|\Psi_{AB}\rangle$ 写成下面的形式

$$|\Psi_{AB}\rangle = \sum_i \lambda_i |i_A\rangle |i_B\rangle.$$

这里的系数 $\sum_i \lambda_i^2 = 1$ 。根据这个结论我们知道，给定一个纯态，子系统 A 和 B 的熵相同

$$S_A = \text{Tr}(\hat{\rho}_A \ln \hat{\rho}_A) = S_B = \text{Tr}(\hat{\rho}_B \ln \hat{\rho}_B),$$

with

$$\hat{\rho}_A = \text{Tr}_B(|\Psi_{AB}\rangle\langle\Psi_{AB}|),$$

$$\hat{\rho}_B = \text{Tr}_A(|\Psi_{AB}\rangle\langle\Psi_{AB}|).$$

这个结论不适用与混态。

6.6 最大熵原理

最大熵原理是人们从信息论的角度来建立统计物理理论框架的结果。这是 E. T. Jaynes 的开创性工作²。通过这个原理，我们可以给出在给定限制条件下，系统处于某一个状态的概率。这些给定限制条件就是我们对系统了解的信息。我们了解的信息不同，给出的概率就不同。

比如，如果我们对系统一无所知，那么最为公平的方式就是假设系统处于每个可能状态的概率都是一样的。这是最公平的，每个状态的地位都一样，没有哪个状态比其他状态更“优越”。这显然就是我们前面讲的统计物理的基本假设：对于孤立系统的等概率假设。

我们还希望能够更进一步。如果我们通过某种途径，得知了系统的一些信息。比如我们通过测量发现，系统的平均能量为 $U = \sum_i P_i E_i$ 。这时，我们可能无法再给每一个微观状态相同的概率了。现在怎么办呢？信息论给出一个方案。我们的任务是根据我们已知的条件（平均能量为 U ），给出系统处于每一个状态的概率。假如系统一共有 Ω 个微观状态，要给出其中每一个状态的概率 P_i ，我们分以下几步：

- 选择 $N(\gg \Omega)$ 个“几率量子”，每个几率量子代表 $1/N$ 的几率。然后把这 N 个几率量子分给 Ω 个微观态。这样，数一下每个微观态“得到”的几率量子数，我们就可以判断系统处于这个状态的概率 $P_i = N_i/N$ 。
- 检查一下上一步得到的一组概率 $\{P_i\}$ 是否满足我们已知的信息，比如 $\sum_i P_i E_i$ 是否等于 U 。
- 如果满足的话，我们就找到了一组满足条件的概率分布。如果没有的话，就重复上面两步。

如果我们按照上面的操作进行的话，最有可能得到的结果是什么？因为出现每一种情况

²E. T. Jaynes, Phys. Rev. 106, 620 (1957); 108, 171 (1957)

的概率都是一样的。最容易出现的结果应该是，在满足已知信息的前提下，使

$$W = \frac{N!}{\prod_{i=1}^{\Omega} N_i!}$$

取最大值的结果。将上式去对数，我们可以得到，只要 N 足够大，

$$\ln W \approx N \sum_i P_i \ln P_i.$$

所以，最容易出现的结果就是，在满足已知信息的情况下，使 S_G 取最大值的那组 P_i 。这就是最大熵原理。

6.6.1 具体实现

数学上如何实现在给定限制条件下最大化一个函数呢？大家肯定知道是拉格朗日乘子法。

6.6.2 应用 1：微正则系综

对于微正则系综，我们需要保证的只是

$$\sum_i P_i = 1 \tag{6.19}$$

这个条件。所以我们需要最大化下面的函数：

$$\tilde{S} = S_G + \alpha \left(\sum_i P_i - 1 \right).$$

由

$$\frac{\partial \tilde{S}}{\partial P_i} = 0$$

我们不难得出

$$P_i = e^{\alpha-1}.$$

这个概率不依赖与具体状态 i 。根据归一化条件，我们可以得到

$$P_i = 1/\Omega.$$

这里， Ω 为总的状态数。也就是说，我们对系统信息一无所知的话，我们得出系统处于每个状态的概率都相同。这正是孤立系统处于各个微观态的等概率假设。

6.6.3 应用 2：正则系综

对于正则系综，给定系统的平均能量，除了需要保证式 (6.19) 外，我们还需要保证

$$\langle E \rangle = \sum_i P_i E_i = U.$$

因此，我们需要最大化

$$\tilde{S} = S_G + \alpha \left(\sum_i P_i - 1 \right) - \beta \left(\sum_i P_i E_i - U \right).$$

据此，我们得到

$$\ln P_i = \frac{\alpha + 1 - \beta E_i}{k_B}.$$

这里的拉格朗日乘子 α 和 β 可以由两个已知条件获得。我们可以把结果写成比较熟悉的形

$$P_i \propto e^{-\beta E_i}.$$

这里我们把 β/k_B 重新定义为 β ，把与 E_i 无关的量忽略了。他们可以通过归一化条件给出

$$Z = \sum_i e^{-\beta E_i}.$$

$$P_i = \frac{1}{Z} e^{-\beta E_i}.$$

6.7 信息论：Maxwell-Szilard-Landauer-Bennett

6.7.1 Maxwell

前面我们看到，热力学与信息论可以通过熵来联系起来。Shannon 熵和 Gibbs 熵的形式非常相似，虽然它们最初分别用来表征不同的量。热力学与信息论之间的密切联系还可以从一个物理学中著名的“妖怪”联系起来，即 Maxwell 小妖。这是 Maxwell 在 1867 年所设计的一个思想实验。小妖有很多版本，我们只提一种。在 Joule 膨胀中，我们计算了气体的空间体积从 V 自由膨胀到 $2V$ 的过程。这个过程中气体与外界没有热量和功的交换，所以是一个孤立系统。它的熵增为 $\Delta S = k_B N \ln 2$ 。Maxwell 通过引入一个巧手小妖 (neat-fingered being)，来实现 Joule 膨胀的逆过程。这个小妖知道每个分子的状态，当遇到从右半部分过来的分子时，ta 打开开关，允许分子进入左半部分。当 ta 遇到从左边过来的分子时，关闭开关，分子被反弹回去。这样下来，可以让所有右半部分的气体分子重新回到左半部分，从而使气体的熵变小，

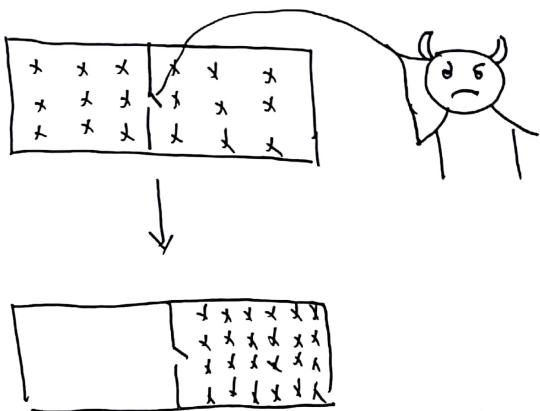


图 6.1: Maxwell 妖实现 Joule 膨胀的反过程。

“违背”热力学第二定律。要实现这个过程，最关键的因素是小妖知道每个分子的状态。只有 ta 知道了这些信息，我们上面描述的过程才会发生。

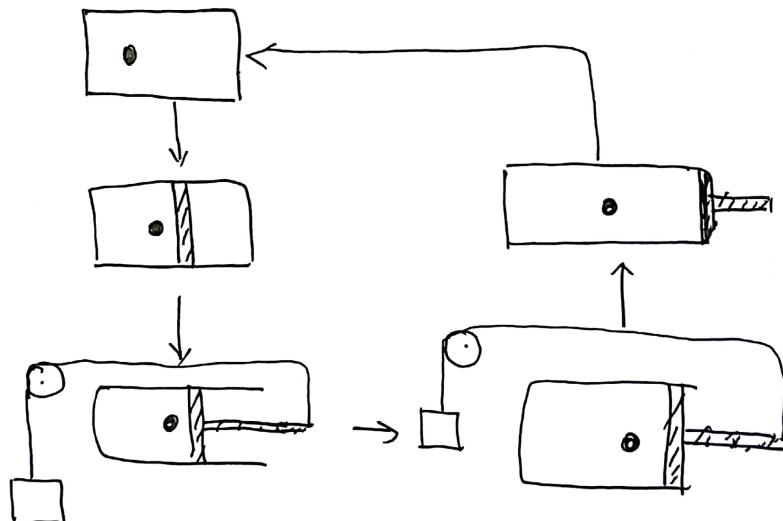
6.7.2 Szilard

当然，我们现在知道，虽然只看气体本身的话，好像违背了热力学第二定律，但如果我们将小妖也考虑在内的话，热力学第二定律就不会被破坏。后人对 Maxwell 的这个思想实验做了仔细研究。最具有代表性的是 Szilard 在 1929 年提出的一个简化版本的热机³。Szilard 热机的工作物质只有一个分子，它位于一个容器内。这个容器与一个温度为 T 的热库接触。热机的工作流程如下图所示。

- 在容器中间插入一个活塞，将分子限制在一半的范围内
- 判断分子在左边还是右边部分，并将空的部分的器壁打开，将活塞与一个重物相连
- 由于气体分子的压强，推动活塞做等温膨胀，从热库吸收热量，同时将重物提起。即，将热转换为功。
- 将活塞移走，并装上容器器壁，完成一个循环。

这里，我们假设插入活塞，除去器壁这些工作只需消耗无限小的功。Szilard 热机的工作看上去只依赖一个热库，并且热机将从热源吸收的热量全部转化为功，“违背”了热力学第二定律。这里最为关键的是我们需要知道分子在容器的左边还是右边，以便把活塞安装到正确的位置，

³L. Szilard, “On the Decrease of Entropy in a Thermodynamic System by the Intervention of Intelligent Beings”, Zeitschrift fur Physik 53, 840–856 (1929). English translation in The Collected Works of Leo Szilard: Scientific Papers, B.T. Feld and G. Weiss Szilard (eds.), Cambridge, Massachusetts: MIT Press, 1972, pp. 103–129.



Szilard 热机

图 6.2: Szilard 热机“实现”从单一热源吸收热量并把它转换为功，用于提升重物。注：这里的容器始终与一个温度为 T 的热源做热接触。

才能将热转化为功。也就是说，热机的工作需要对系统进行“测量”，然后根据“测量”的结果做出正确的“反馈”。⁴ Szilard 对这个热机的分析指出了信息在热机工作过程中的重要性。我们首先要知道分子处于左边还是右边，才能在接下来的操作中对外做功。后人的分析都是基于以上思路进行的。比如，Brillouin。他们的观点可以概括为：要获得分子的位置信息，需要对它进行测量，而测量的过程需要消耗功，这就会引起外界环境的变化。因此，虽然我们将热完全转化为了功，但因为这一过程对外界产生了影响，所以并没有违背热力学第二定律。

6.7.3 Landauer

1961 年，IBM 科学家 Landauer 对计算过程能量消耗及热力学进行了分析⁵。他把计算过程中的一些基本逻辑操作分成逻辑可逆和逻辑不可逆两种。逻辑可逆过程是指根据操作的结果可以反推出操作前的状态。也就是说操作的初态和末态之间存在一个 1:1 的映射。而逻辑不可逆过程无法末态退出初态。比如，相同的末态可能对应不同的初态。Landauer 指出，所有的逻辑运算都是在实际的物理系统中进行的。逻辑可逆的运算前后系统的状态空间没有变

⁴ 我们这里的测量都是指经典的测量。

⁵ R. Landauer, “Irreversibility and heat generation in the computing process”, IBM Journal of Research and Development, 5, 183–191 (1961)

化，但逻辑不可逆的过程发生后系统的状态空间变小了。因此，他认为，逻辑不可逆的操作都伴随着系统向环境的热量耗散，而逻辑可逆过程可以避免这一耗散⁶。

一个典型的逻辑不可逆过程就是重置操作。一个比特不管它初始状态是 0 还是 1，我们都把它置位一个确定的状态（可以是 0，也可以是 1）。重置前后，这个比特的熵降低了 $k_B \ln 2$ 。这个过程中，这个比特位向外界传递热量至少 $k_B T \ln 2$ 。这被称为 Landauer 原理。通俗地讲，存储信息不需要耗散热量，而擦除信息需要向外散热。利用这个原理，我们就可以分析实际计算过程中有哪些不可逆操作，对应每种操作会释放多少热量。这就给出了计算过程中能量消耗的下限。

Landauer 本人并没有分析测量过程是否一定是不可逆的。后来，Landauer 在 IBM 的同事 Bennett 基于 Landauer 的分析指出，可以通过可逆过程来完成测量⁷。甚至整个计算过程也可以用可逆操作来实现。这就意味着至少可以设计不需要消耗能量的测量过程，来得到 Szilard 热机中分子的位置。Maxwell 妖也可以不耗散能量的前提下得知每个分子的状态。Bennett 的分析说明 Szilard 当初有关测量需要消耗能量的论断是站不住脚的。

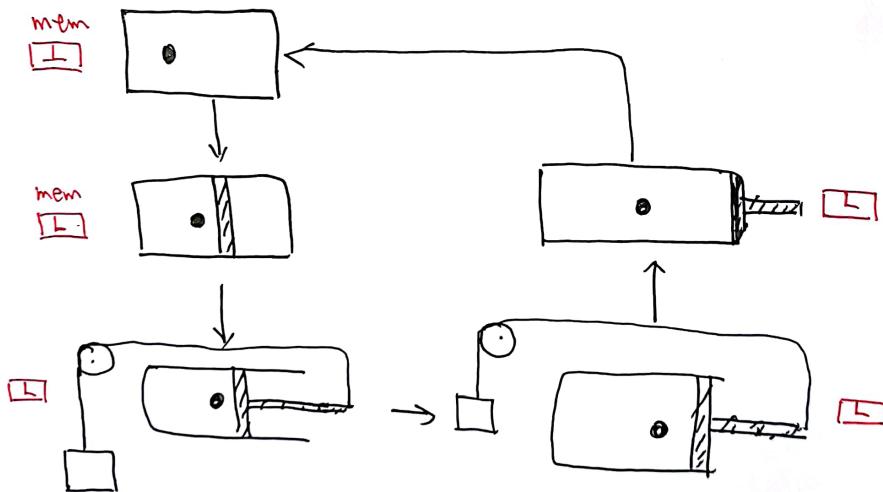
6.7.4 Szilard again

现在我们利用 Landauer 原理来重新分析一下 Szilard 热机的循环过程，寻找新的可以解救第二定律的地方。这个关键的地方在于，我们对于热机中分子的状态进行测量之后需要把它记录在一个比特中。这个过程虽然不消耗能量。但要完成一个热机的循环，我们在最后一步需要把这个比特重置。根据 Landauer 原理，如果这个比特所处的温度与热机所处的温度相同，都是 T 的话，那么重置操作需要向外界释放热量 $k_B T \ln 2$ 。如果重置前后这个比特的能量不变的话，我们需要对比特做相同的功才可以完成这个重置操作。这部分功的大小正好和热机等温膨胀做的功相同。于是，一个循环完成之后，我们并没有真正向外输出任何功。自然，这个过程也就没有违背热力学第二定律。我们得救了！

思考：如果系统所处温度为 T_S ，比特位所处温度为 T_B 。系统对位做功和擦除比特位所需能量不再相等。试分析说明：如果 $T_S > T_B$ ，Szilard 热机相当于 Carnot 热机，它可以对外输出功，效率为 $\eta = 1 - T_B/T_S$ 。如果 $T_S < T_B$ ，相当于反向工作的热机，即利用输出的功将热量从

⁶思考：逻辑可逆和热力学可逆是相同的概念吗？

⁷C. H. Bennett, “Logical reversibility of computation”, IBM Journal of Research and Development, 17, 525–532 (1973). see also: C. H. Bennett, “The thermodynamics of computation—a review”, International Journal of Theoretical Physics, 21, 905–940 (1982)



Szilard 热机

图 6.3: 实际上，Szilard 热机在工作过程中需要记录气体分子的位置，因此需要一个比特的存储单元。在循环完成的最后一步，需要将存储单元中的信息擦除掉。这个过程需要消耗 $k_B T \ln 2$ 的功。这正好等于热机输出的功相同。所以，Szilard 热机无法从单一热源吸热，并把它全部转化为功。

低温热库转移到高温热库。

通过以上分析我们发现，信息是和热力学紧密联系在一起的。利用信息，我们可以对外做功。而信息的处理过程也伴随着热量传递以及做功。所以，Landauer 曾经有一个论断：**Information is physical**⁸。我们需要像对待其他物理量一样来对待信息。看来，要完整的描述热力学定律必须把信息以一定的方式包含进来。这里，我们考虑的只是经典的信息。如果再和量子信息相结合，我们可以考虑更加有趣的问题。事实上，在最近几年越来越多的人开始从热力学的角度来分析量子信息中的问题。这使得信息热力学成为一个很受关注的研究领域⁹。

⁸R. Landauer, “Information is physical”, Physics Today, 44, 23 (1991); Physica Letter A, 217, 188 (1996).

⁹(1)Sai Vinjanampathy, Janet Anders, Contemporary Physics, 57, 545 (2016). (2) Juan M. R. Parrondo, Jordan M. Horowitz and Takahiro Sagawa, Nature Physics, 11, 131 (2015).

第7章 热力学势

前面通过介绍热力学第零、第一和第二定律我们分别引入了温度、能量、熵等热力学量。同时，要描述一个宏观热力学系统，我们也使用了体积、压强等力学量。此外，还包括粒子数以及我们后面要介绍的化学势。这些量当中，温度、压强和化学势为强度量，相应的熵、体积和粒子数为广延量。温度和熵，压强和体积，化学势和粒子数各自互为共轭变量，两者相乘的量纲为能量。

一般来说，要描述系统的一个状态需要指定几个相互独立的物理量，比如要描述给定理想气体 (N 固定) 的状态只需要两个独立变量 (比如 P, V)。其他的物理量可以通过状态方程用这些独立变量表示出来。具体选择哪些物理量作为独立变量，可能要根据具体问题来确定。根据具体问题不同，我们经常可以方便的使用除了内能之外的其他能量来描述系统，统称为热力学势。这部分主要就是要介绍一些在不同情况下常用的热力学势。

7.1 Legendre 变换

利用 Legendre 变换，我们可以把独立变量从某个量转换为其对应的共轭变量。这个变换我们在理论力学中已经遇到过了。当时，我们用它来实现从 Lagrangian 向 Hamiltonian 变换： $\mathcal{H}(p, q) = p\dot{q} - \mathcal{L}(q, \dot{q})$ 。具体做法大家都非常熟悉。这里我们着重介绍一下 Legendre 变换的具体含义。

一般来说，我们做 Legendre 变换的函数 $f(x)$ 是一个平滑的凸函数 (convex function)，也就是说其二次微分为正。这样，我们就可以方便的对其做微分，令 $p(x) = \frac{\partial f(x)}{\partial x}$ 为其一次微分，或者表示为 $f(x) = p(x)dx$ 。对于凸函数，给定 x 就可以确定相应的 p ，反过来，给定 p 也可以唯一确定 x 。所以， x 和 p 之间是一一对应的。

有时，我们可能使用 p 来表示函数更方便。这时，我们就可以对 $f(x)$ 做 Legendre 变换，把变量从 x 变为 p 。变换的过程可以从图形上来理解 (图7.1)：

- 取任意一点 x_0 ，函数 $f(x)$ 在这一点的斜率可以用其一次微分表示即 $p(x_0)$
- 沿着 x_0 位置的切线做直线。这条直线可以表示为 $T(x) = p(x_0)(x - x_0) + f(x_0)$
- $T(x)$ 与 y 轴的交点为 $g(x_0) = f(x_0) - x_0 p(x_0)$

这里，给定 x_0 就可以得到 $p(x_0)$ ，进而就可以得到 $g(x_0)$ ，反之亦然。我们定义 Legendre 变换

图 7.1: Legendre 变换 (from Am. J. Phys., 77, 614 (2009).)

图 7.2: From D. V. Schroeder, An Introduction to Thermal Physics.

为 $g(x(p)) = f(x(p)) - x(p)p$ 。这里定义的 $g(x(p))$ 其实只是 p 的函数，因为

$$dg = pdx - xdp - pdx = -xdp. \quad (7.1)$$

我们只需要把 $p(x) = f'(x)$ 求逆就可以把 x 用 p 来表示 $x = (f')^{-1}(p)$ ，然后把它带入到 $g = f - xp$ 中就最终得到了 Legendre 变换的最后形式。

这里 (f, x) 和 (g, p) 是对相同函数的不同种表示形式。如果我们进一步对 $g(p)$ 做 Legendre 变换，就又可以得到 $f(x)$ ，即：做两次 Legendre 变换就可以得到原函数。下面，我们就利用 Legendre 变换来定义几个常用的热力学势。

回顾： Lagrangian 向 Hamiltonian 的变换

7.2 焓

我们介绍的第一个热力学势是焓。它的定义如下：

$$H = U + PV = TS + \mu N.$$

焓在描述等压环境下的热力学过程比较方便。数学上，增加 PV 这一项后，我们可以把 H 的自然变量换为 S 和 P ，即

$$dH = dU + d(PV) = TdS + VdP + \mu dN.$$

物理上讲，增加的 PV 可以理解为在等压环境下，要腾出放置系统的空间 V ，需要克服外界压强做功，因此需要额外付出能量 PV 。所以，在等压环境下要从无到有准备系统的话，除了需要付出系统的内能对应的能量外，还要额外付出腾出空间需要的能量。同样的道理，如果从有到无的话，我们不光可以获得系统的内能，还可以得到额外的 PV 大小的能量，如下图所示。

我们前面介绍了等容热容和等压热容的概念。根据热力学第一定律，如果只考虑体积功，那么，定容热容可以表示为

$$C_V = \left(\frac{dQ}{dT} \right)_V = \left(\frac{\partial U}{\partial T} \right)_V.$$

对于定压热容，我们可以做类似的考虑（仍然只考虑体积功）

$$dU = dQ - PdV$$

在等压条件下，我们有

$$dQ = dU + PdV = d(U + PV) = dH.$$

同时，我们可以利用焓的定义，选取 T 和 P 作为变量

$$dH = \left(\frac{\partial H}{\partial T}\right)_P dT + \left(\frac{\partial H}{\partial P}\right)_T dP.$$

所以，定压热容可以表示为

$$C_P = \left(\frac{\partial H}{\partial T}\right)_P.$$

引入焓的概念后， C_P 可以用简洁的形式表示出来，其物理意义也更加清楚。

7.3 Hemholtz 势

Hemholtz 势通过以下变换定义：

$$F = U - TS = -PV + \mu N.$$

其微分形式可以表示为：

$$dF = -SdT - PdV + \mu dN.$$

可见，其自然变量为 T 、 V 、 N 。在实际实验中， T 比 S 要容易控制，因此 Hemholtz 势使用起来更为方便，特别是在等温条件下。我们知道对于正则系综，系统与热库耦合，其温度固定，使用 F 比较方便。

物理上，Hemholtz 势又被称为 Hemholtz 自由能。后面我们详细分析，其实 Hemholtz 势表示系统内能中可以用来“自由支配”对外做功的那部分能量，所以被称为自由能。从内能中减掉的那部分 TS 是系统熵所贡献的能量，这部分是无法转换为有效输出功的。它是与温度为 T 的热库耦合之后所必然携带的能量，无法被“加以利用”。

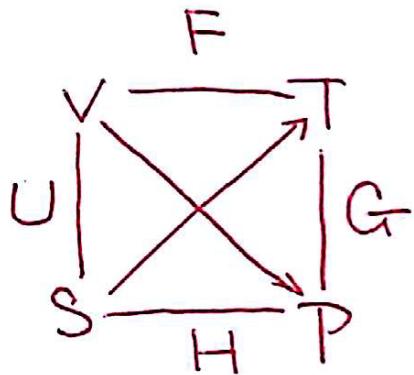


图 7.3: 热力学势和相应自然变量。

表 7.1: 常用热力学势

热力学势	微分形式	自然变量
U	$dU = TdS - PdV$	S, V
H	$dH = TdS + VdP$	S, P
F	$dF = -SdT - PdV$	T, V
G	$dG = -SdT + VdP$	T, P

7.4 Gibbs 势

Gibbs 势是在 Hemholtz 势的基础上进一步做变换：

$$G = F + PV = U - TS + PV + \mu N = \mu N.$$

微分形式为：

$$dG = -SdT + VdP + \mu dN.$$

其自然变量为 T 、 P 、 N 。Gibbs 势在等温、等压的情况下使用起来比较方便。很多化学反应都是在这样的条件下使用的，所以 Gibbs 势在化学中使用非常频繁。相反，物理学中人们更多的使用 Hemholtz 势。与 F 类似， G 表示的是在等温、等压条件下，系统可以用来对外做功的那部分能量，因此又被称为自由能。

表??总结了我们定义的热力学势及相应的基本关系和自然变量。它们之间的关系还可以用图7.3来表示出来。

7.5 热力学第二定律

7.5.1 孤立系统

对于孤立系统，我们前面得到热力学第二定律可以表示为：

$$\Delta S \geq 0$$

也就是说，孤立系统经历自发过程后熵不会减少。系统在趋于平衡的过程中熵会不断增加，直到平衡时取最大值。

如果我们关心的系统与一个温度为 T 的热库耦合，使得它们之间可以有热量的交换；那么系统不再是孤立系统，所以我们不能再用熵增原理来描述第二定律。这时，如何单纯通过系统的状态量来描述第二定律呢？

7.5.2 Helmholtz 自由能降低

如果系统与热库耦合，并且体积固定。那么把系统与环境作为一个大的孤立系统，对于任意过程，我们有

$$\Delta S + \Delta S_B \geq 0$$

这里， ΔS 和 ΔS_B 分别为系统和环境的熵变。对于环境，我们有

$$\Delta S_B = -Q/T_B$$

这里， Q 为系统从环境吸收的热量， T_B 为热库的温度。综合起来，我们有

$$\Delta S - Q/T_B \geq 0$$

而利用第一定律我们有

$$\Delta U = Q + W$$

于是我们有：

$$W \geq \Delta U - T_B \Delta S = \Delta F$$

上式说明，系统从一个状态转为另外一个状态时，外界对它做的功大于等于其自由能的变化。上式第二个等式是因为系统的初态和末态都是平衡态，所以其温度都和热库温度相同。尽管系统在状态转换过程中可能处于非平衡态，所以温度可能与热库不同。这里的等号对应可逆

过程，外界对系统做的功全部转化为自由能 F 。这里可以分两种情况讨论：(1) $W > 0$ ，即外界对系统做功，系统自由能 F 增加。这时，最为经济的方式就是通过可逆过程，否则必须付出更多的功。(2) $W < 0$ ，系统对外界做功。这时，如果是可逆过程， F 的降低全部转化为功；如果是不可逆过程，只有部分 F 转化为功。通过上面分析我们看到，自由能其实也就是系统可以对外做功的那部分能量。

以上结果告诉我们另外一个事实，如果系统与外界没有功的交换，那么 $\Delta F \leq 0$ 。也就是说，系统 Helmholtz 自由能倾向于降低，直到达到平衡后，自由能不再变化，取极小值。所以，热力学第二定律在这种情况下可以表示为 Helmholtz 自由能降低。

7.5.3 Gibbs 自由能降低

如果系统与一个温度为 T_B ，压强为 P_B 的环境耦合，那么系统如果从一个状态转化为另一个状态，进行类似的分析，我们可以得到 $\Delta G \leq 0$ 。等号对应的是可逆过程。这时，第二定律可以表示为 Gibbs 自由能的降低。

7.6 稳定性条件

第8章 化学势

8.1 化学势的定义

化学势是热力学与统计物理中另外一个重要的核心概念。与温度类似，在其他的课程都没有深究它的定义和物理意义，所以也是我们需要重点学习和理解的一个概念。我们之所以这么晚才介绍它，是因为它的作用只有在考虑开放系统时才显现出来。也就是说，它和系统粒子数的变化密切相关。而我们先前介绍的系统基本上都是孤立或封闭系统，粒子数保持恒定。因此，化学势就隐藏在我们的表达式中，并不显山露水，我们也不必考虑。

8.1.1 热力学

前面我们提到了共轭变量的概念。比如，温度和熵为一对共轭变量，压强和体积为另外一对共轭变量。每对共轭变量的乘积为能量量纲，一个为强度量，一个为广延量。这里介绍的化学势 μ 和粒子数 N 是一对共轭变量，所以化学势的量纲为能量。与其它两对共轭变量类似， μ 和 N 也对系统的能量有贡献。引入化学势的概念后，我们需要把基本热力学关系式表示为：

$$dU = TdS - PdV + \mu dN. \quad (8.1)$$

我们前面已经学习到，温度可以用来表征一个系统向与之热接触的其他系统传递热量的难易程度。温度越高，系统越倾向于向外输出热量。两个系统热接触达到平衡后，两个系统的温度相同。同理，压强可以用来表征系统向外膨胀的难易程度；可以自由交换体积的两个系统达到力学平衡后，压强相同。以此类推，化学势可以用来表示系统向外输出粒子的难易程度。化学势越高，越容易向外输出粒子。如果把两个系统放到一起，允许它们交换粒子，粒子会从化学势高的系统流向化学势低的系统，就像水往低处流一样。

8.1.2 化学势的统计定义

我们可以把式(8.1)写成以熵为基本量的形式

$$dS = \frac{1}{T}dU + \frac{P}{T}dV - \frac{\mu}{T}dN. \quad (8.2)$$

仍然与温度和压强的定义作类比，我们可以通过下面的一个过程从统计物理的角度来定义化学势。假设原来两个孤立的系统分别处于平衡态，现在将两个系统连通起来，它们可以在 U 和 V 不变时，通过交换粒子数来实现平衡，最终达到平衡后，复合系统的熵达到最大值。我们仍然忽略两个子系统间相互作用对复合系统熵的贡献

$$dS = dS_1 + dS_2.$$

这里，下标 1 和 2 分别代表子系统。根据上述假设条件，我们可以进一步得到

$$dS = \left(\frac{\partial S_1}{\partial N_1}\right)_{U_1, V_1} dN_1 + \left(\frac{\partial S_2}{\partial N_2}\right)_{U_2, V_2} dN_2 \quad (8.3)$$

$$= \left(\left(\frac{\partial S_1}{\partial N_1}\right)_{U_1, V_1} - \left(\frac{\partial S_2}{\partial N_2}\right)_{U_2, V_2}\right) dN_1. \quad (8.4)$$

当复合系统达到平衡后， $dS = 0$ ，于是：

$$\left(\frac{\partial S_1}{\partial N_1}\right)_{U_1, V_1} = \left(\frac{\partial S_2}{\partial N_2}\right)_{U_2, V_2}.$$

这里我们引入化学势的定义：

$$-\frac{\mu}{T} \equiv \left(\frac{\partial S}{\partial N}\right)_{U, V}. \quad (8.5)$$

或者

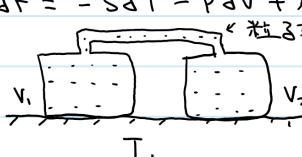
$$\mu \equiv -T \left(\frac{\partial S}{\partial N}\right)_{U, V}. \quad (8.6)$$

这个定义与温度和压强的统计定义形式相同。之所以选择负号，是为了保证在复合系统熵增过程中，粒子从化学势高的子系统流向化学势低的子系统。而分母中温度的存在可以保证式子的量纲正确，也可以保证我们的定义和热力学关系式(8.2)是自洽的。

验证：根据上面的定义，如果子系统 1 和 2 的温度相同，我们得到：

$$dS = -\frac{1}{T} (\mu_1 - \mu_2) dN_1 \geq 0$$

可见，如果 $\mu_1 \geq \mu_2$ ，那么 $dN_1 \leq 0$ ，即粒子从化学势高的子系统流向化学势低的子系统。

$$dF = -SdT - PdV + \mu dN \Rightarrow \mu = \left(\frac{\partial F}{\partial N} \right)_{T, V}$$


平衡条件: $dF = 0 \Rightarrow F = F_1 + F_2$. 弱相互作用
 $dF_1 + dF_2 = 0 \Rightarrow \left(\frac{\partial F_1}{\partial N_1} \right)_{T, V_1} dN_1 + \left(\frac{\partial F_2}{\partial N_2} \right)_{T, V_2} dN_2 = 0$.
而 $dN_1 = -dN_2$ (总粒子数守恒)
 $\Rightarrow \left(\frac{\partial F_1}{\partial N_1} \right)_{T, V_1} = \left(\frac{\partial F_2}{\partial N_2} \right)_{T, V_2} \Rightarrow \mu_1 = \mu_2$

图 8.1: 通过自由能来定义化学势。

8.2 从自由能的角度考虑化学势

上面我们从内能和熵的角度对化学势进行了介绍。在引入化学势的统计定义时，我们假设系统的 U 和 V 不变，来考虑 S 随 N 的变化。理论上我们当然可以这么考虑，但如果要实际上实现这样操作，我们会遇到麻烦。我们在改变系统的粒子数 N 时，不可避免的会改变其 U 或 S ，因为它们都是随 N 而变的。这样的困难可以通过对自由能的分析来避免。

考虑到 μ 和 N 后，Helmholtz 自由能可以表示为：

$$dF = d(U - TS) = -SdT - PdV + \mu dN.$$

根据上面的关系，我们可以把化学势定义为：

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

固定温度和体积显然要容易得多。因此，通过自由能对化学势的定义从操作上讲就容易得多，如图8.1。同理，我们也可以定义

$$\mu = \left(\frac{\partial G}{\partial N} \right)_{T, P}$$

8.3 巨正则系统

粒子数可以变化的系统属于开放系统，相应的系综为巨正则系综。这时，要指定系统所处的微观状态需要同时给定粒子数 N 以及在 N 粒子情况下的某一个微观状态 k_N 。按照与正则系统类似的分析，我们可以得出系统处于某一个微观态 (N, k_N) 的概率

$$P(N, k_N) \propto e^{-\beta(\varepsilon_{k_N} - \mu N)}$$

称为 Gibbs 因子。它和正则系统的 Boltzmann 因子相对应。据此，我们可以定义系统的巨正则配分函数：

$$\mathcal{Z} = \sum_N \sum_{k_N} e^{-\beta(\varepsilon_{k_N} - \mu N)}$$

与正则系统类似，所有的热力学量都可以通过巨正则配分函数来得到。比如，

$$\langle N \rangle = \bar{N} = \beta^{-1} \left(\frac{\partial \ln \mathcal{Z}}{\partial \mu} \right)_T. \quad (8.7)$$

$$U = \mu \bar{N} - \frac{\partial \ln \mathcal{Z}}{\partial \beta} \quad (8.8)$$

$$\begin{aligned} S &= -k_B \sum_i P_i \ln P_i \\ &= -k_B \sum_i P_i (-\beta(\varepsilon_i - \mu N_i) - \ln \mathcal{Z}) \\ &= \frac{1}{T} (U - \mu N + k_B T \ln \mathcal{Z}). \end{aligned} \quad (8.9)$$

这里注意，巨正则系统内能表达式和正则系统是不一样的。对比正则系统，我们可以定义一个新的热力学势：巨势。它的地位和正则系统中的 F 相同

$$\Phi_G \equiv -k_B T \ln \mathcal{Z}. \quad (8.10)$$

根据式 (9.6) 和 (9.7)，我们有

$$\Phi_G = U - ST - \mu N = F - \mu N. \quad (8.11)$$

或

$$d\Phi_G = -SdT - PdV - Nd\mu$$

据此，我们可以得到以下关系式：

$$S = - \left(\frac{\partial \Phi_G}{\partial T} \right)_{V\mu}, \quad (8.12)$$

$$P = - \left(\frac{\partial \Phi_G}{\partial V} \right)_{T\mu}, \quad (8.13)$$

$$N = - \left(\frac{\partial \Phi_G}{\partial \mu} \right)_{TV}. \quad (8.14)$$

8.3.1 举例：理想气体（这部分内容需要学完理想气体一章后才能明白）

8.3.1.1 化学势

我们在介绍理想气体的时候考虑的是粒子数 N 固定，即正则系统的情况。如果要计算理想气体的化学势，理论上我们需要考虑巨正则系统，即允许粒子数 N 的变化，所以我们的出发点应该是巨配分函数

$$\mathcal{Z} = \sum_N e^{\beta\mu N} \sum_{k_N} e^{-\beta\varepsilon_{k_N}}$$

利用上一章的内容，我们可以得到

$$\mathcal{Z} = \sum_N e^{\beta\mu N} \frac{(Vn_Q)^N}{N!} \quad (8.15)$$

$$= \sum_N \frac{1}{N!} (Vn_Q e^{\beta\mu})^N \quad (8.16)$$

$$= e^{Vn_Q e^{\beta\mu}}. \quad (8.17)$$

或

$$\ln \mathcal{Z} = Vn_Q e^{\beta\mu}$$

把上面的结果代入到式 (9.4) 中

$$\bar{N} = Vn_Q e^{\beta\mu}$$

于是化学势可以表示为

$$\mu = -k_B T \ln \left(\frac{n_Q}{\bar{n}} \right). \quad (8.18)$$

我们已经得到的理想气体的结果在 $n \ll n_Q$ 时才成立，这时理想气体的化学势是负值 $\mu < 0$ 。从定义 (??) 出发， $\mu < 0$ 意味着给系统增加一个粒子，系统的内能降低。如何理解这一结果呢？这里式 (??) 中下标的重要性就体现出来了。我们需要维持 S 和 V 不变，如果系统粒子数增加 1 的话，系统的熵也会随之增加，要维持熵不变的话，系统需要向外传递热量来降低它的熵。 $\mu < 0$ 的意思就是说，要维持熵不变的话，系统需要向外输出的热量比这个粒子所携带能量更多。这样，最终的结果是系统内能降低。

练习：根据这个结果，大家可以估计一下标准条件下（室温，一个标准大气压）空气中空气分子的化学势。同时验证一下， $n \ll n_Q$ 。

8.3.1.2 巨势

理想气体的巨势可以写为：

$$\Phi_G = -\beta^{-1} V n_Q e^{\beta \mu} \quad (8.19)$$

如果把 μ 替换掉的话

$$\Phi_G = -\beta^{-1} V n_Q e^{\beta \mu} = -N k_B T = -PV. \quad (8.20)$$

据此，我们可以验证 (8.12-8.14) 中的各个关系式。

8.4 Euler 方程、Gibbs-Duhem 关系

我们的基本热力学关系式都是通过微分的形式给出来的。我们利用广延量的特征，可以将这些关系式表示为非微分的形式。以熵为例，根据广延量的特征¹，我们有

$$S(\lambda U, \lambda V, \lambda N) = \lambda S(U, V, N)$$

现在取 $\lambda = 1 + \epsilon$, $\epsilon \rightarrow 0^+$, 对上式做泰勒展开取一阶项，我们可以得到下面的关系式：

$$\begin{aligned} S(U, V, N) + \epsilon \left(\left(\frac{\partial S}{\partial U} \right)_{VN} U + \left(\frac{\partial S}{\partial V} \right)_{UN} V + \left(\frac{\partial S}{\partial N} \right)_{UV} N \right) \\ = S(U, V, N) + \epsilon S(U, V, N) \end{aligned} \quad (8.21)$$

即：

$$U = TS - PV + \mu N. \quad (8.22)$$

上式称为 Euler 方程。结合 (8.1)，我们还可以得到 Gibbs-Duhem 关系：

$$SdT - VdP + Nd\mu = 0$$

它说明尽管我们可以独立控制 S, V, N 这三个广延量，实际上我们无法独立控制对应的强度量 T, P 和 μ 。这三个量之间必须要满足 Gibbs-Duhem 关系。

¹这里需要注意，个别系统不具有这样的广延性。比如，(1) 如果系统内部的相互作用势为长程势（比如引力），(2) 系统尺度较小，导致组合时候，不同子系统间的相互作用势能不可忽略。如果系统不满足广延性，那么这一节介绍的 Gibbs-Duhem 关系就不保证成立。我们这里只考虑广延系统。

我们可以写出其他热力学势的表达式：

$$H = U + PV = TS + \mu N \quad (8.23)$$

$$F = U - TS = -PV + \mu N \quad (8.24)$$

$$G = F + PV = \mu N \quad (8.25)$$

$$\Phi_G = F - \mu N = -PV. \quad (8.26)$$

讨论：

- 式 (8.25) 告诉我们化学势的另外一个意义：它表示单粒子的 Gibbs 自由能 $\mu = G/N$ 。
- 式 (9.7) 是一个普适关系式，不只适用于理想气体，但它成立的前提是熵具有广延量的特性。

8.5 粒子数守恒与化学势

如果一个系统的粒子数可以产生和湮灭，它的能量会随着粒子数的变化而改变。这个过程中可能伴随着能量在外界环境与系统之间的交换。当系统达到平衡时，我们有

$$dF = -SdT - PdV + \mu dN$$

在等温等容条件下，系统达到平衡后， $(\partial F / \partial N)_{TV} = 0$ 意味着 $\mu = 0$ 。也就是说，粒子数不守恒的系统，粒子化学势为零。²

8.6 化学势：内部贡献和外部贡献

我们前面讨论的化学势是系统本身的性质，可以称为系统的内部化学势 (intrinsic)。这部分我们想说明，我们可以通过施加外部的势来改变系统的总化学势。

我们可以举例说明。考虑两个系统 1 和 2 同时和热库耦合达到平衡温度 T ，同时它们的体积 V_1 和 V_2 也不发生变化。我们将两个系统连通，如果初始时它们的化学势不同，就会导致粒子从化学势高的系统流向化学势低的系统。我们不妨假设 $\mu_1 > \mu_2$ 。这时，如果我们把系统 2 的位置抬高 h ，那么系统 2 会获取额外的重力势能 $Nmgh$ 。系统的自由能也相应的增加：

$$dF_2 = mghN_2.$$

²如果系统与一个非平衡的库耦合达到稳态，这时系统并不处于热平衡。即使系统粒子数不守恒，化学势也不一定为零。详见：P. Wurfel, The chemical potential of radiation, Journal of Physics C: Solid State Physics, 15, 3967 (1982).

这时，我们可以定义一个总的化学势为：

$$\mu_t = \mu_2 + m_2 gh.$$

这样我们可以把重力势能看成是外部化学势，而把系统本身的化学势看作其本征的内部化学势。如果重力势能的贡献正好保证

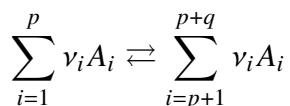
$$m_2 gh = \mu_1 - \mu_2,$$

这样，把系统 2 抬高后可以正好抵消它与系统 1 的化学势之差，从而避免粒子在两个系统之间的交换。

因此，我们可以把总的化学势分成内部 (intrinsic) 和外部 (extrinsic) 的贡献之和。从这个角度讲，化学势相当于我们熟悉的势能，比如电势、重力势等。比如，我们还可以通过施加外加电场来改变带电粒子的化学势，这时人们通常把它称为电化学势 (electrochemical potential)。

8.7 应用：化学反应

我们可以利用化学势的概念来分析化学反应的平衡态如何与反应涉及到的反应物和产物的化学势建立联系。化学反应一般通过化学反应方程式来表示：



式子左边为反应物，右边为产物。其中， ν_i 表示反应物或产物在反应中的占比。我们可以把所有反应物移到右边，但把对应的比例系数表示为负值，而产物的系数表示为正值。

如果我们在一个等温、等压的环境下把任意反应物和产物混合到一起，混合系统经过足够长时间后会最终达到一个平衡的状态。这时各种组分的比例可以通过反应平衡的条件来确定。我们知道，Gibbs 自由能在平衡状态下取极值，而在温度和压强不变的情况下，我们有

$$dG = \sum_i \mu_i dN_i = 0. \quad (8.27)$$

根据化学反应方程式，我们知道：

$$\frac{dN_i}{dN_1} = \frac{\nu_i}{\nu_1},$$

所以我们得到：

$$\sum_i \mu_i \nu_i = 0. \quad (8.28)$$

这正是我们想要的反应平衡的条件。

为了表征反应达到平衡的状态，我们定义反应平衡常数

$$K = \prod_{i=1}^{p+q} \left(\frac{p_i}{p_0} \right)_i^{\nu_i}. \quad (8.29)$$

这里， p_0 为标准大气压。如果把反应物和产物近似为理想气体，我们可以把它们的化学势表示为：

$$\mu(p) = \mu_0 + k_B T \ln \frac{p}{p_0}.$$

代入到式 (8.28) 中，我们有：

$$\sum_i \nu_i \mu_0 = -k_B T \ln K. \quad (8.30)$$

从而得到：

$$\boxed{\ln K = -\frac{\Delta G_0}{RT}}. \quad (8.31)$$

我们定义了 $\Delta G_0 = N_A \sum_i \nu_i \mu_0$ 。利用 Gibbs-Helmholtz 关系，

$$H = G + TS = G - T \left(\frac{\partial G}{\partial T} \right)_P = -T^2 \left(\frac{\partial (G/T)}{\partial T} \right)_P,$$

我们可以把上式改写为：

$$\frac{d \ln K}{dT} = -\frac{1}{R} \frac{d(\Delta G_0/T)}{dT} = \frac{\Delta H_0}{RT^2}. \quad (8.32)$$

或者

$$\frac{d \ln K}{d(1/T)} = -\frac{\Delta H_0}{R}. \quad (8.33)$$

- 这个结果与 Le Chatelier 原理相符：在系统达到化学平衡后，如果给它一个扰动，系统会沿减小外界扰动的方向来对扰动进行响应。
- 对于一个化学反应，根据 $\ln K$ 和 $1/T$ 的直线关系可以拟合出 ΔH_0 的大小。

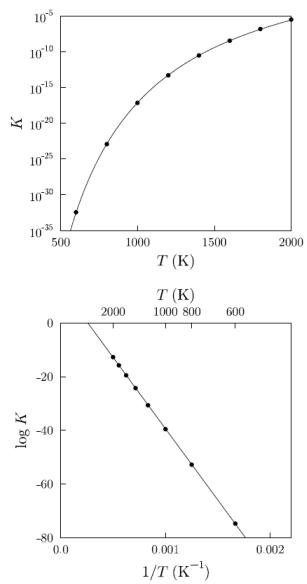


Fig. 22.4 The equilibrium constant for the reaction $H_2 \rightarrow H\cdot + H\cdot$, as a function of temperature. The same data are plotted in two different ways.

第9章 统计物理的简单框架

到目前为止，我们已经基本把热力学部分的内容介绍完了。我们的讲法和《热学》的区别在于，在介绍热力学基本定律的同时我们引入了常见热力学状态参量的统计定义，从而和统计物理的基本理论建立了联系。在开始统计物理的内容之前，我们有必要简单回顾一下我们已经引入的统计物理的基本假设和基本理论框架。

9.1 理论原理

9.1.1 基本假设

等概率假设 (Principle of equal apriori probability)：对于一个孤立热力学系统，在给定限制条件下（比如给定能量、粒子数等），系统处于每个可能微观状态的概率相同。也就是说，系统的内部动力学决定了在足够长时间内系统会遍历所有允许的微观状态，而且在每一个状态所处的时间相同。但是，由于每个宏观状态对应的微观状态数不同，不同宏观状态出现的概率不同。最可能出现的宏观状态对应的微观状态数最多。

根据系统所处宏观状态对应的微观状态数 $\Omega(E, N, V)$ ，我们可以定义 Boltzmann 熵：

$$S_B(E, N, V) = k_B \ln \Omega(E, N, V).$$

熵是最抽象也最难实验测量的物理量，但它却是统计物理中最为核心的一个物理量。从熵出发，我们可以定义一系列的其他热力学参量。

9.1.2 温度、压强、化学势的统计定义

我们考虑子系统 1 和 2 组成的复合系统。每个子系统的状态可以用能量 E 、体积 V 和粒子数 N 来表征。假设两个子系统之间可以存在热量、体积和粒子数的交换，但整个复合系统的能量、体积和粒子数恒定。通过这些物理量的交换，整个复合系统最终达到平衡。这时，复合系统的熵达到最大值。利用最大熵条件，我们可以得到温度、压强、化学势的统计定义。

取复合系统的熵为两个子系统熵之和 $S = S_1 + S_2$ ，从而忽略两个系统耦合带来的微小修

正¹。那么，我们有

$$dS = \left(\frac{\partial S}{\partial E_1} \right)_{V_1, N_1} dE_1 + \left(\frac{\partial S}{\partial V_1} \right)_{E_1, N_1} dV_1 + \left(\frac{\partial S}{\partial N_1} \right)_{E_1, V_1} dN_1 = 0$$

因为 E, V 和 N 为相互独立的三个参量，上式中的每一项都取零。从

$$\left(\frac{\partial S}{\partial E_1} \right)_{V_1, N_1} = \left(\frac{\partial S_1}{\partial E_1} \right)_{V_1, N_1} + \left(\frac{\partial S_2}{\partial E_1} \right)_{V_1, N_1} = 0$$

利用总能量恒定的条件，我们得到

$$\left(\frac{\partial S_1}{\partial E_1} \right)_{V_1, N_1} = \left(\frac{\partial S_2}{\partial E_2} \right)_{V_1, N_1}$$

上式表示两个系统的热量交换达到平衡，而且等式左右两边分别为表征系统 1 和 2 的物理量。由此，我们可以把它和温度联系起来。定义温度为

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right)_{V, N}.$$

同样的道理，我们可以用下面的两个式子

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

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

分别表征体积和粒子数的平衡。据此，我们可以得到压强和化学势的定义：

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

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

9.1.3 热力学基本定律

9.1.3.1 基本关系

把熵作为 E, V 和 N 的函数，我们可以写出

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

根据我们上面的定义，我们可以得到：

$$dE = T dS - P dV + \mu dN$$

于是我们就从统计的角度得到了基本的热力学表达式。

¹对于子系统间通过短程相互作用耦合的宏观系统，这个近似足够好，因为宏观的三维系统之间最多以二维的面互相接触。但对于小系统或者存在长程相互作用的系统，这个近似不一定成立。

9.1.3.2 热力学第二定律

根据我们对熵的定义，很容易说明热力学第二定律是成立的。比如，我们把由相同种类分子构成的，原来分别处于平衡的两个子系统 1 和 2 放到一起，形成一个新的复合系统。子系统 1 和 2 的初始状态分别用 (E_1, V_1, N_1) 和 (E_2, V_2, N_2) 来表示，对应的新复合系统的状态用 $(E = E_1 + E_2, V = V_1 + V_2, N = N_1 + N_2)$ 来表示。我们知道初态的熵为 $S_i(E, V, N) = S_1(E_1, V_1, N_1) + S_2(E_2, V_2, N_2)$ 。两个系统耦合在一起后会发生能量、体积和粒子数的交换，最终复合系统演化到新的平衡态。这时，复合系统的熵可以表示为 $S_f(E, V, N) = \sum'_{E_1, V_1, N_1} (S_1(E_1, V_1, N_1) + S_2(E_2, V_2, N_2))$ 。这里的求和表示在给定总能量 N 、总体积 V 和总粒子数 N 的前提下，系统所有可能访问到的状态的熵之和。显然 S_f 包含了 S_i ，而 $S \geq 0$ ，所以 $S_f > S_i$ 。

9.1.3.3 第三定律

$T = 0$ 时，系统处于基态，所以熵为

$$S = k_B \ln g$$

这里 g 为基态的简并度。可见，零温时，系统的熵取一个很小的常数。如果简并度为 1，则熵为零。

9.1.4 从孤立系统到封闭和开放系统

9.1.4.1 封闭/正则系统

如果系统和一个热库耦合，并且可以和热库之间交换热量。通过把系统加上热库当成一个孤立复合系统，通过类似的分析，我们可以得出系统处于任何一个微观态 k 的概率 $P(k)$

$$P(k) \propto e^{-\beta E_k}$$

这里 $\beta = 1/k_B T$ ， E_k 为微观态 k 对应的能量。为了得到绝对概率我们需要对上面的 Boltzmann 因子做归一化，归一化因子定义为系统的配分函数

$$Z = \sum_k e^{-\beta E_k}$$

正则系统的熵可以用概率 $P(k)$ 来表示为

$$S = -k_B \sum_k P(k) \ln P(k)$$

称为 Gibbs 熵。我们发现，Boltzmann 熵是 Gibbs 熵的特例。因为对于孤立系统，它处在每一个微观态的概率相同，都是 $P(k) = 1/\Omega$ 。

9.1.4.2 开放/巨正则系综

我们可以同时考虑热量和粒子数的交换，也就是考虑巨正则系统。采用与正则系统类似分析方法，我们可以得到

$$P(N, k_N) \propto e^{-\beta(E_{k_N} - N\mu)}$$

其中， $P(N, k_N)$ 为系统处于状态 (N, k_N) 的概率。这时，我们指定系统的微观态时需要指定粒子数 N 和给定 N 的情况下系统可以访问的微观状态 k_N 。这时的归一化函数我们称为巨配分函数

$$\mathcal{Z} = \sum_{N, k_N} e^{-\beta(E_{k_N} - N\mu)}$$

9.2 最大熵原理

基于 Shannon 的信息熵，E. T. Jaynes 在上世纪 50 年代的工作把信息论与统计物理建立了一个联系，发展了最大熵原理²。利用最大熵原理，我们也可以得到微正则、正则和巨正则系统的概率分布。这样，统计物理就可以看做是数学上统计推断的一个应用：在给定限制条件（已知的宏观量）下，对系统性质进行统计推断。推断的办法就是最大化目标函数--系统的熵

$$S = -k_B \sum_{\lambda} P_{\lambda} \ln P_{\lambda}. \quad (9.1)$$

9.2.1 微正则 (E, V, N) 系综

微正则系综下给定系统的宏观量有能量 E 、粒子数 N 和体积 V 。在这些固定条件下的平衡态，系统处于给定微观状态 λ 的概率由最大熵原理给出。

- 限制条件： $\sum_{\lambda} P_{\lambda} = 1$
- 最大化目标函数： $\tilde{S} = S + \alpha(1 - \sum_{\lambda} P_{\lambda}) = -k_B \sum_{\lambda} P_{\lambda} \ln P_{\lambda} + \alpha(1 - \sum_{\lambda} P_{\lambda})$. 这里 α 为拉格朗日乘子。

²E. T. Jaynes, Phys. Rev., 106, 620 (1957); Phys. Rev., 108, 171 (1957).

- \tilde{S} 取最大值的必要条件:

$$\frac{\partial \tilde{S}}{\partial P_\mu} = 0$$

由此可以得到:

$$-k_B(\ln P_\mu + 1) - \alpha = 0 \rightarrow P_\mu = \text{Const.}$$

- 由限制条件可以得到微正则系综下, 系统处于微观态 λ 的概率:

$$P_\lambda = \frac{1}{\Omega(E, N, V)}.$$

这里, $\Omega(E, N, V)$ 为给定宏观态下对应的微观态的数目。它与给定的状态 λ 无关。也就是说, 系统处于任何一个状态的概率都相同。这也就是统计物理中的等概率假设, 但这里是根据最大熵原理得到的。

- 微正则系综下, 系统平衡态的熵又可以表示为:

$$S = -k_B \ln \Omega.$$

这是我们熟悉的 Boltzmann 熵的表达式。可见, Boltzmann 熵可以看作是 Gibbs-Shannon-von Neumann 熵的一种特殊形式。

9.2.2 正则 (T, V, N) 系综

9.2.2.1 概率分布

正则系综下, 系统与一个温度为 T 的热库耦合, 从而保持温度恒定, 但系统可以与热库交换热量, 所以其能量可以变化。

- 这时, 我们有两个限制条件: (1) $\sum_\lambda P_\lambda = 1$; (2) $\sum_\lambda E_\lambda P_\lambda = U$. 目标函数为: $\tilde{S} = -k_B \sum_\lambda P_\lambda \ln P_\lambda + \alpha(1 - \sum_\lambda P_\lambda) + \beta(U - \sum_\lambda P_\lambda E_\lambda)$. 这里 α, β 都是拉格朗日乘子。
- 最大化 \tilde{S} 给出正则系综下的概率分布:

$$P_\lambda \propto e^{-\beta E_\lambda}.$$

- 归一化因子为配分函数:

$$Z = \sum_\lambda e^{-\beta E_\lambda}$$

处于能量本征态 E_λ 的概率为:

$$P_\lambda = Z^{-1} e^{-\beta E_\lambda}.$$

- 拉格朗日因子可以由给定的条件

$$U = \frac{\sum_{\lambda} P_{\lambda} E_{\lambda}}{\sum_{\lambda} P_{\lambda}}$$

来确定。

9.2.3 巨正则 (T, V, μ) 系综

9.2.3.1 概率分布

巨正则系综下，系统与外界既可以交换能量，又可以交换粒子，所以其能量和粒子数可以变化，但温度和化学势固定。

- 这时，我们有三个限制条件：(1) $\sum_{\lambda} P_{\lambda} = 1$; (2) $\sum_{\lambda} E_{\lambda} P_{\lambda} = U$; (3) $\sum_{\lambda} n_{\lambda} P_{\lambda} = \langle N \rangle = N$. 目标函数为： $\tilde{S} = -k_B \sum_{\lambda} P_{\lambda} \ln P_{\lambda} + \alpha(1 - \sum_{\lambda} P_{\lambda}) + \beta(U - \sum_{\lambda} P_{\lambda} E_{\lambda}) + \gamma(N - \sum_{\lambda} P_{\lambda} n_{\lambda})$. 这里 α, β, γ 都是拉格朗日乘子。
- 最大化 \tilde{S} 给出正则系综下的概率分布：

$$P_{\lambda} \propto e^{-\beta E_{\lambda} - \gamma n_{\lambda}} = e^{-\beta(E_{\lambda} - \mu n_{\lambda})}.$$

- 归一化因子为巨配分函数：

$$Z = \sum_{\lambda} e^{-\beta(E_{\lambda} - \mu n_{\lambda})} = \text{Tr}(e^{-\beta(H - \mu n)})$$

处于能量本征态 E_{λ} 的概率为：

$$P_{\lambda} = Z^{-1} e^{-\beta(E_{\lambda} - \mu n_{\lambda})}.$$

- 拉格朗日因子可以由给定的条件

$$U = \frac{\sum_{\lambda} P_{\lambda} E_{\lambda}}{\sum_{\lambda} P_{\lambda}},$$

和

$$N = \frac{\sum_{\lambda} P_{\lambda} n_{\lambda}}{\sum_{\lambda} P_{\lambda}}$$

来确定。

9.3 实际应用

上节我们介绍了统计物理的基本理论。我们发现，熵这个概念在统计物理的理论中占据了核心的地位。我们很多热力学量的定义都是基于熵做出的，不同系统的概率分布也是在熵

的基础上得到的。

在实际应用中，我们往往会“反其道而行之”。我们首先去计算系统的配分函数；从配分函数出发，我们可以逐一的研究系统的热力学量和性质。配分函数在基本理论部分只是作为一个归一化因子而出现的，但它在实际应用中却担任着最为重要的角色。原则上，只要我们写出了系统的配分函数，我们的统计物理分析就结束了。下面我们要做的只是一些计算而已。我们后半部分的内容就是要把这样的“套路”反复利用到不同的系统。我们会考虑理想气体、光子气体、声子气体等几个可以严格求解的模型系统。大家在以后的科研工作中可能也会遇到其他的系统，那时往往我们不能像课上举得这些例子这样，对它们严格求解。这就要求我们发展实用的近似方法。

我们课本中介绍了几个简单的例子来帮助大家熟悉这一个套路。其中包括：二能级系统、谐振子、N 能级系统以及角动量等。基本过程如下：

- 利用其他知识，比如量子力学，得出系统的能谱
- 根据能谱写出配分函数
- 根据配分函数计算热力学量和状态方程

我们的统计物理理论既适用与经典又适用于量子情况。从量子力学得到的能谱出发我们可以得到在低温下仍然适用的一些结果。比如，对于上面的几个例子，我们得到的热容在 $T \rightarrow 0$ 时都趋于 0，与第三定律相符。如果我们从经典的结果出发的话，只能得出在经典、高温极限下适用的结果。下面我们分别介绍正则和巨正则系综下如果从配分函数得到其他热力学量。

9.3.0.1 正则系综

内能: 内能可以用配分函数表示为：

$$U = \langle E \rangle = \frac{\sum_n E_n e^{-\beta E_n}}{\sum_n e^{-\beta E_n}} = -\frac{\partial \ln Z}{\partial \beta}.$$

熵: 对于熵，我们从 Gibbs 的表达式出发更为方便。因为往往我们研究的系统不是一个孤立系统。

$$S = -k_B \sum_n P_n \ln P_n \quad (9.2)$$

$$= k_B (\beta U + Z). \quad (9.3)$$

Helmholtz 自由能: Helmholtz 自由能与配分函数的关系最为简单，从上面熵的表达式我

们可以直接得出：

$$F = -\beta^{-1} \ln Z = -k_B T \ln Z.$$

也正是由于这个原因，Helmholtz 自由能往往使我们得到配分函数后首先考虑计算的物理量。

物态方程 基于 Helmholtz 自由能，我们可以得到压强的表达式：

$$P = - \left(\frac{\partial F}{\partial V} \right)_T.$$

由此，可以得到系统的物态方程。

根据这些结构，我们可以进一步计算引入的所有其他热力学势或热容等响应函数。

9.3.0.2 巨正则系综

与正则系统类似，所有的热力学量都可以通过巨正则配分函数来得到。

平均粒子数：

$$\langle N \rangle = \bar{N} = \left(\frac{\partial \ln \mathcal{Z}}{\partial \alpha} \right)_{T,V}. \quad (9.4)$$

这里， $\alpha = \beta\mu$, μ 为系统的化学势。

内能：

$$U = \mu \langle N \rangle - \left(\frac{\partial \ln \mathcal{Z}}{\partial \beta} \right)_{V,\mu} \quad (9.5)$$

如果利用前面引入的 α ，我们还可以表示成：

$$U = F - \mu N = - \left(\frac{\partial \ln \mathcal{Z}}{\partial \beta} \right)_{V,\alpha}$$

熵：

$$\begin{aligned} S &= -k_B \sum_{\lambda=\{N,k_N\}} P_\lambda \ln P_\lambda \\ &= -k_B \sum_{\lambda} P_\lambda (-\beta(\varepsilon_{k_N} - \mu N) - \ln \mathcal{Z}) \\ &= \frac{1}{T} (U - \mu \langle N \rangle + k_B T \ln \mathcal{Z}). \end{aligned} \quad (9.6)$$

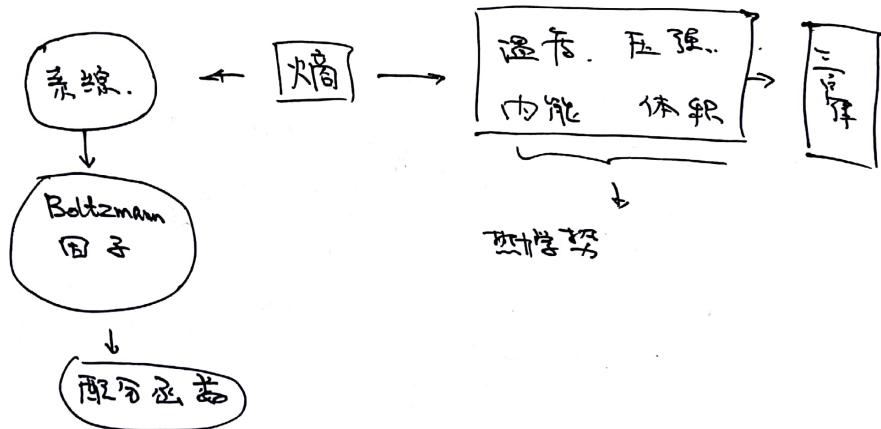
这里注意，巨正则系统内能表达式和正则系统是不一样的。

巨势：

$$\Phi_G \equiv -k_B T \ln \mathcal{Z}. \quad (9.7)$$

✎ **练习 9.1** 压强系综：假设系统与外部环境的耦合导致系统的温度和压强恒定。这时其能量和体积可以发生变化。这个系综称为压强系综。根据最大熵原理，请给出压强系综下，系统微

* 理论框架.



* 实际应用: 反其道而行之.

配分函数 \rightarrow 一切热力学量.

图 9.1: Basic Framework

观态的占据概率。

9.4 涨落, 不同系综的“等价性”

9.4.1 正则系综中的能量涨落

从正则系综下的配分函数出发, 我们已知系统的平均能量可以写成:

$$U = \langle E \rangle_{V,N} = - \left(\frac{\partial \ln Z}{\partial \beta} \right)_{V,N}. \quad (9.8)$$

再次微分, 我们可以得到:

$$\left(\frac{\partial^2 \ln Z}{\partial \beta^2} \right)_{V,N} = - \left(\frac{\partial U}{\partial \beta} \right)_{V,N} = \langle E^2 \rangle_{V,N} - \langle E \rangle_{V,N}^2 \equiv \langle \Delta E^2 \rangle_{V,N}. \quad (9.9)$$

上面式子中, $\langle \Delta E^2 \rangle_{V,N}$ 称为方差 (variance), 表示在 V, N 不变的情况下, 能量偏离其平均值的程度。本节后半部分为了避免麻烦, 我们省掉下标 V, N

$$\langle \Delta E^2 \rangle = \langle (E - \langle E \rangle)^2 \rangle = \langle E^2 \rangle - \langle E \rangle^2. \quad (9.10)$$

我们定义标准差 (standard deviation) 为方差的平方根, 并用它来表示系统偏离平均值的大小。系统能量的相对涨落可以表示为

$$\frac{\sqrt{\langle \Delta E^2 \rangle}}{\langle E \rangle} \propto \frac{1}{\sqrt{N}}. \quad (9.11)$$

在热力学极限下, $N \rightarrow \infty$, 而相对涨落趋于零。也就是说, 虽然在正则系综中我们允许系统的能量发生变化, 但在热力学极限下其涨落相对于均值来说是无限小的, 即系统与热库之间的热量交换是无限小的。我们考察这样的系统与考察完全孤立的系统得到的结果将不会有什区别。在这个意义上, 我们可以认为在热力学极限下, 正则系统与微正则系综是“等价”的。

虽然热力学极限下, 正则系统的相对能量涨落非常小, 但涨落并不是没有用处的。实际上, 我们可以把系统的能量涨落用一些描述系统的可观测的物理量联系起来。通过对系统涨落性质的研究就可以得到这些可观测量。我们知道:

$$-\left(\frac{\partial U}{\partial \beta}\right)_V = k_B T^2 \left(\frac{\partial U}{\partial T}\right)_V = k_B T^2 C_V. \quad (9.12)$$

其中, C_V 为系统的定容热容。之所以是定容热容, 是因为我们在对 β 或 T 做微分的时候, 假定系统体积不变, 即能量 E_i 不变。于是, 我们得到:

$$\langle \Delta E^2 \rangle = k_B T^2 C_V. \quad (9.13)$$

讨论:

- 上式说明系统的能量涨落和温度相关, 与系统热容相关。实际上, 上面的关系是平衡态统计物理中一个非常重要的涨落-耗散关系的特例。热容是系统的一个响应函数, 它表征的是系统对热量传递的温度响应。我们这里看到, 系统的能量涨落是和它对吸收、释放热量的温度响应之间联系在一起的。关于涨落-耗散关系我们后面会做详细介绍。
- 系统与热库达到平衡后, 根据我们的定义, 其能量涨落是非负的, 因此根据式 (9.13) 我们也需要 C_V 非负, $C_V \geq 0$ 。后面我们在学习相变是会发现, 系统发生类似液-气相变的时候会持续吸热, 但其温度不会发生变化。这时系统的热容为无穷大。对应其能量涨落就不能再被忽略。也就是说, 系统在发生相变的过程中, 它的某些物理量的涨落可以非常大。

9.4.2 巨正则系综中的能量涨落

从巨正则系综的配分函数出发, 我们得到粒子数平均值为:

$$\langle N \rangle = \bar{N} = \frac{\sum_{N,k_N} N e^{\alpha N - \beta E_{k_N}}}{\sum_{N,k_N} e^{\alpha N - \beta E_{k_N}}} = \left(\frac{\partial \ln Z}{\partial \alpha} \right)_{T,V}. \quad (9.14)$$

使用上节相同的方法, 我们得到:

$$\langle \Delta N^2 \rangle = \left(\frac{\partial \bar{N}}{\partial \alpha} \right)_{T,V} = k_B T \left(\frac{\partial \bar{N}}{\partial \mu} \right)_{T,V} = \frac{N^2 k_B T \kappa_T}{V}. \quad (9.15)$$

这里, $\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_{T,N}$ 为等温压缩率。我们看到, 粒子数的相对涨落同样反比与粒子数的 $1/2$ 次方

$$\frac{\sqrt{\langle \Delta N^2 \rangle}}{\bar{N}} \sim \frac{1}{\sqrt{\bar{N}}}. \quad (9.16)$$

对式 (9.15) 的证明: 第二个等号等价于:

$$\frac{N^2}{V^2} \left(\frac{\partial V}{\partial P} \right)_{T,N} = - \left(\frac{\partial N}{\partial \mu} \right)_{T,V}.$$

我们需要证明上式。利用

$$\Phi_G = -PV, \quad d\Phi_G = -SdT - PdV - Nd\mu,$$

可得:

$$\left(\frac{\partial P}{\partial \mu} \right)_{T,V} = \left(\frac{\partial N}{\partial V} \right)_{T,\mu} = \frac{N}{V}.$$

代入需证:

$$-\left(\frac{\partial P}{\partial \mu} \right)_{T,V} \left(\frac{\partial V}{\partial P} \right)_{T,N} = \left(\frac{\partial N}{\partial \mu} \right)_{T,V} \left(\frac{\partial \mu}{\partial P} \right)_{T,V} = \left(\frac{\partial N}{\partial P} \right)_{T,V}.$$

下面我们利用:

$$\left(\frac{\partial N}{\partial P} \right)_{T,V} \left(\frac{\partial P}{\partial V} \right)_{T,N} \left(\frac{\partial V}{\partial N} \right)_{T,P} = -1.$$

即需证明:

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

上式成立吗? 是的!

下面考虑其能量涨落

$$\langle \Delta E^2 \rangle = - \left(\frac{\partial U}{\partial \beta} \right)_{\alpha,V} = k_B T^2 \left(\frac{\partial U}{\partial T} \right)_{\alpha,V}, \quad (9.17)$$

为了具体分析这一能量涨落, 我们把它分成两部分:

$$\left(\frac{\partial U}{\partial T} \right)_{\alpha,V} = \left(\frac{\partial U}{\partial T} \right)_{N,V} + \left(\frac{\partial U}{\partial N} \right)_{T,V} \left(\frac{\partial N}{\partial T} \right)_{\alpha,V} \quad (9.18)$$

为了得到上式, 我们使用了链式法则, 对 $U = U(T, N(\alpha, T, V), V)$ 在保持 α 和 V 不变的情况下求偏导。上面式子的第一部分是保持粒子数不变的情况, 也就是和正则系统类似的情况; 第二部分是由于粒子数变化带来的涨落。对于第二部分, 我们还可以做进一步整理。根据

$$\langle N \rangle = \left(\frac{\partial \ln Z}{\partial \alpha} \right)_{\beta,V}, \quad U = - \left(\frac{\partial \ln Z}{\partial \beta} \right)_{\alpha,V}, \quad (9.19)$$

我们得到

$$\left(\frac{\partial \langle N \rangle}{\partial \beta}\right)_{\alpha,V} = - \left(\frac{\partial U}{\partial \alpha}\right)_{\beta,V}. \quad (9.20)$$

根据这个关系还可以写成:

$$\left(\frac{\partial \langle N \rangle}{\partial T}\right)_{\alpha,V} = \frac{1}{k_B T^2} \left(\frac{\partial U}{\partial \langle N \rangle}\right)_{\beta,V} \left(\frac{\partial \langle N \rangle}{\partial \alpha}\right)_{\beta,V}. \quad (9.21)$$

最后, 利用这些关系式, 我们可以把巨正则系综下的能量涨落写为:

$$\boxed{\langle \Delta E^2 \rangle_{GCE} = \langle \Delta E^2 \rangle_{CE} + \left(\frac{\partial U}{\partial \langle N \rangle}\right)_{T,V}^2 \langle \Delta N^2 \rangle}. \quad (9.22)$$

我们可以看到, 在巨正则系综下, 系统的能量涨落分两个部分。第一部分与正则系综下的能量涨落类似, 第二部分对应由于粒子数涨落而导致的能量涨落。而且, 不管是粒子数还是能量, 它们的相对涨落都正比于 $1/\sqrt{N}$ 。与正则系综类似, 我们也可以说, 在热力学极限下, 巨正则系综、正则系综和微正则系综是“等价”的。

总结:

- 以上分析表明, 在热力学极限下, 虽然我们允许正则和巨正则系综中能量和粒子数发生变化, 但实际上它们的涨落相对与均值来说趋于零, 因此微正则、正则和巨正则系综在热力学极限下是等价的。当然, 如果系统的尺度不是足够大, 三种系综之间会有明显的区别。
- 不同系综下可变物理量的涨落是和系统的对应响应函数相联系的。而这些响应函数是很方便测量的。这其实是涨落-耗散定理的一个体现。

9.5 举例 1: spin 1/2

我们考虑 N 个格点组成的系统。每个格点上有一个自旋 $1/2$ 粒子, 其自旋态可以取 \uparrow 和 \downarrow 两种状态。在外加磁场 B 的作用下, 两个状态的能量分别取 0 和 $\Delta > 0$ 。下面我们考虑这一自旋体系的热力学性质。

9.5.1 正则系综

在正则系综下, 我们考虑一个格点, 那么它只有两种可能的微观状态。所以其配分函数为

$$Z_1 = 1 + e^{-\beta \Delta}.$$

由于 N 个格点相互独立，所以总的配分函数为：

$$Z_N = Z_1^N.$$

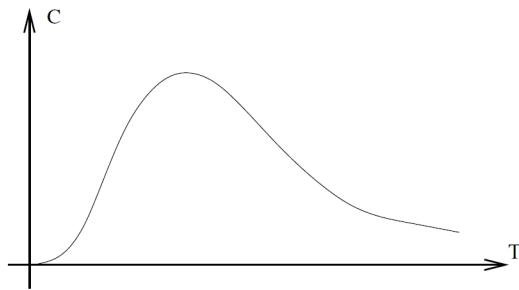
内能：据此我们得到系统的内能为

$$U = -\frac{\partial \ln Z_N}{\partial \beta} = \frac{N\Delta}{e^{\beta\Delta} + 1}.$$

据此，我们还可以求出体系的热容：

$$C_N = \left(\frac{\partial U}{\partial T} \right)_N = Nk_B(\beta\Delta/2)^2 \cosh^{-2}(\beta\Delta/2). \quad (9.23)$$

可以看出，当 $T \rightarrow 0$ 时， C 随 T 以指数形式趋向于 0。当 $T \gg \Delta/k_B$ 时， C 随 T 以 T^{-2} 衰减向 0。在两者之间存在一个极大值，如下图所示。这种行为与电子和声子的热容有很大区别，这个极大值的出现称为 Schottky 反常。



自由能：Helmholtz 自由能为：

$$F = -k_B T \ln Z_N = -Nk_B T \ln(1 + e^{-\beta\Delta}).$$

熵：由于系统的平均能量又可以表示为 $U = N_\uparrow \Delta$ ，所以有

$$\frac{N_\uparrow}{N_\downarrow} = e^{-\beta\Delta}.$$

根据 $S = (U - F)/T$ 我们可以得到：

$$S = -k_B N(P_\uparrow \ln P_\uparrow + P_\downarrow \ln P_\downarrow), \quad (9.24)$$

这里 $P_x = N_x/N$ 。

9.5.2 微正则系综

下面我们考虑微正则系综，给定系统能量为 $E = N_\uparrow \Delta$ 。这一宏观状态对应的总微观状态数为

$$\Omega(N_\uparrow) = \frac{N!}{N_\uparrow!(N - N_\uparrow)!}. \quad (9.25)$$

相应的熵为：

$$S = k_B \ln \Omega(N_\uparrow).$$

可以验证，这与式 (9.24) 相同。根据这些结果，我们可以来计算整个体系的温度与 N_\uparrow 之间的关系：

$$\frac{1}{k_B T} = \frac{\partial \ln \Omega}{\partial E} = \frac{\partial \ln \Omega}{\partial N_\uparrow} \frac{\partial N_\uparrow}{\partial E} \quad (9.26)$$

$$= -\frac{1}{\Delta} \ln \frac{N_\uparrow}{N_\downarrow} \quad (9.27)$$

我们可以给这个孤立系统定义一个温度：

$$T(N_\uparrow) = -\frac{\Delta}{k_B} \left(\ln \frac{N_\uparrow}{N_\downarrow} \right)^{-1}. \quad (9.28)$$

利用上式我们可以得到

$$N_\uparrow = \frac{N}{1 + e^{-\beta \Delta}}. \quad (9.29)$$

把上式带入到式 (??) 中，我们可以得到式 (3.19)。这样我们就直接计算得到了系统的总能量表达式，但过程要复杂很多。

从温度的表达式 (9.28) 中我们可以发现一个很有意思的现象：当 $N_\uparrow > N_\downarrow$ 时， $T > 0$ ；但当 $N_\uparrow < N_\downarrow$ 时， $T < 0$ ；在临界的状态 $N_\uparrow = N_\downarrow$ ，系统的温度为无穷大。这说明，根据温度的统计学定义，它可以为正，也可以为负，而且 $E(T < 0) > E(T = +\infty)$ 。当我们把 $T < 0$ 的系统与 $T = +\infty$ 的另一个相同系统进行热接触时，热量将会从负温度的系统流向正温度的系统。所以，我们可以说，负温度系统比正无穷大温度的系统还要热，而 $T = 0$ 仍然是最冷的温度。这一点和我们的直观印象不同。这里我们考虑的自旋体系实际上是实现负温度的理想体系，而且已经有人实现了这一温度状态³。虽然人们可以在不同的系统中实现这一状态，但并不是所有的理论学家接受负温度的概念⁴。

9.6 举例 2：理想气体

理想气体是我们下一章介绍的内容。这里我们从完全经典的角度先做一个分析。

³Purcell&Pound, 81, 279 (1951), Hakonen&Lounasmaa, Science, 265, 1821 (1994), Braun, Ronzheimer, Schreiber, Hodgman, Rom, Bloch, Schneider, Science, 339, 52 (2013)

⁴Dunkel&Hilbert, Nature Physics, 10, 67 (2014), Swendsen&Wang, arXiv:1506.0691, Wang, arXiv:1507.02022

9.6.1 微正则系综

考虑孤立理想气体系统，系统状态可以由 (E, V, N) 来确定。由于 E 为连续取值，完全固定 E 很困难，我们可以允许 E 在小范围内变化， $[E, E + \Delta]$ 。后面的分析会告诉我们，这个 Δ 的具体取值并不重要。给定这个宏观状态后，系统的微观状态数由下式得到⁵：

$$\Omega(E, V, N) \propto \Delta \int d^{3N}q d^{3N}p \delta \left(E - \sum_i \frac{p_i^2}{2m} \right) \quad (9.30)$$

因为我们忽略气体粒子间的相互作用势能，能量并不依赖于粒子的位置，所以对于每个粒子位置的积分都给出系统体积 V 。而对于动量的积分其实是在 $3N$ 维空间，半径为 R 的超球面面积。利用 d 维空间超球面的面积公式：

$$S_d(R) = \frac{2\pi^{d/2}}{(d/2 - 1)!} R^{d-1} = \frac{2\pi^{d/2}}{\Gamma(d/2)} R^{d-1}, \quad (9.31)$$

可以得到：

$$\Omega(E, V, N) \propto \Delta V^N \frac{2\pi^{3N/2}}{(3N/2 - 1)!} (2mE)^{(3N-1)/2}. \quad (9.32)$$

对应的 Boltzmann 熵为：

$$\begin{aligned} S/k_B &= N \ln V + (3N/2) \ln \pi + \frac{1}{2} (3N - 1) \ln (2mE) + \ln 2 - \ln (3N/2 - 1)! \\ &\quad + \ln \Delta + \text{some constant} \end{aligned}$$

在热力学极限下，所有与 N 无关的常量都可以忽略。我们有

$$S \approx Nk_B \left[\ln V + \frac{3}{2} \ln \left(\frac{4\pi emE}{3N} \right) \right]. \quad (9.33)$$

我们看到，尽管我们允许系统的能量可以在一个很小的范围 Δ 内变化，在热力学极限下，这个变化范围由于与系统大小无关，它的贡献可以忽略。

内能：利用温度的定义，我们可以得到：

$$\frac{1}{T} = \frac{\partial S}{\partial E} = \frac{3N}{2E} \rightarrow E = \frac{3}{2} N k_B T.$$

把以上能量的表达式代入到熵的表达式，我们又可以把熵表示为：

$$S \approx Nk_B \left[\ln V + \frac{3}{2} \ln (2\pi emk_B T) \right].$$

压强：利用压强的定义，我们可以得到理想气体的物态方程：

$$\frac{P}{T} = \frac{\partial S}{\partial V} = \frac{N}{V} \rightarrow PV = Nk_B T.$$

⁵显然下式的两边量纲不同。我们只知道它们成正比，但目前我们还无法给出 Ω 的绝对值。

化学势：

$$\frac{\mu}{T} = - \left(\frac{\partial S}{\partial N} \right) = -k_B \left(\ln V + \frac{3}{2} \ln \frac{4\pi emE}{3N} \right) = -k_B \left(\ln V - \frac{3}{2} \ln (2\pi emk_B T) \right). \quad (9.34)$$

这里，大家可能注意到一个问题：化学势应该是一个强度量，但我们得到的结果却不是这样 ($\propto \ln V$)。如果回头看熵的表达式，我们发现它也不满足广延量的性质。这个问题我们将在下一章再讨论。

练习： Volume entropy of Gibbs!

9.6.2 正则系综

如果考虑正则系综，系统的配分函数可以表示为：

$$Z \propto \prod_i \int d^3 q_i d^3 p_i \exp \left(-\beta \sum_i \frac{p_i^2}{2m} \right) \quad (9.35)$$

由于我们忽略了粒子之间的相互作用势能，我们可以针对每一个粒子对位置和动量分别进行积分。于是，

$$= V^N (2\pi m k_B T)^{3N/2}. \quad (9.36)$$

据此，我们也可以得到与微正则系综下完全相同的物态方程、内能和熵的表达式。我们得到的熵同样存在不满足广延性的问题。

第 10 章 量子统计

10.1 密度矩阵

10.1.1 密度算符

统计物理的基本理论既适用于经典体系又适用于量子体系，但经典和量子理论描述系统状态的方式不同。在经典系统中，系统的微观状态由相空间的一个点来确定 ($\{\vec{q}_i\}, \{\vec{p}_i\}$)，即我们指定每个粒子的空间位置 \vec{q}_i 和它的动量 \vec{p}_i 就可以指定系统所处的微观状态。而在量子系统中，系统所处的状态由一个波函数 $|\psi\rangle$ 来描述，所有这些波函数的集合构成一个希尔伯特空间。与经典系统不同，在给定状态下，粒子的位置和动量无法同时确定。量子系统的观测量由厄密算符来表示，给定状态 $|\psi\rangle$ 下观测量 \hat{O} 的期待值为

$$\langle \hat{O} \rangle = \langle \psi | \hat{O} | \psi \rangle$$

系统的宏观状态可以通过少数几个宏观量来确定，比如内能 U ，体积 V ，粒子数 N 等。一个宏观状态可以对应非常多的微观态。处于平衡态的系统宏观量保持不变，但其微观力学导致系统会在对应这一宏观态的多个微观态之间来回跃迁。在经典系统中，系统的动力学由哈密顿运动方程或者牛顿运动方程来描述，而在量子系统中则是由 Schrödinger 运动方程来描述。从对系统进行统计描述的角度来看，具体微观运动方程的形式不是很重要。但我们都假设系统在足够长的时间内会遍历它所有可以访问的，符合给定宏观条件的微观态，而且系统处于任何一个微观态的概率都是相同的。在这样的各态历经假设下，平衡态量子系统中某一个观测量 \hat{O} 的平均值可以表示为：

$$\overline{\langle \hat{O} \rangle}_t = \frac{1}{T} \int_0^T dt \langle \psi(t) | \hat{O} | \psi(t) \rangle. \quad (10.1)$$

这里的 $\langle \cdot \rangle_t$ 表示 t 时刻的量子平均，而上面的横线表示时间平均。只要时间 T 足够长，就可以保证系统遍历所有可能的微观状态。根据系综理论，我们可以用系综平均来代替上面的时间平均。我们假想一个系综中存在 N 个处于相同宏观态的系统，而每个系统可能处于不同的微观态。系统数 N 远远大于微观态个数 $\Omega(E, V, N, \dots)$ 。这样，物理观测量的平均可以表示为

系综平均：

$$\overline{\langle \hat{O} \rangle} = \frac{1}{N} \sum_{i=1}^N \langle \psi^i | \hat{O} | \psi^i \rangle \quad (10.2)$$

由于 $N \gg \Omega$, 微观态 λ 出现的概率可以表示为

$$P_\lambda = \frac{N_\lambda}{N}.$$

于是, 上面的平均又可以写成:

$$\overline{\langle \hat{O} \rangle} = \sum_{\lambda=1}^{\Omega} P_\lambda \langle \psi^\lambda | \hat{O} | \psi^\lambda \rangle. \quad (10.3)$$

式 (10.2) 中的求和是对系综中的所有组员, 而 (10.3) 中的求和是对所有的微观态。为了简化, 我们后面会省略系综平均的符号, 而把 $\overline{\langle \hat{O} \rangle}$ 直接写为 $\langle \hat{O} \rangle$ 。式 (10.3) 可以改写为:

$$\begin{aligned} &= \sum_{m,n,\lambda} P_\lambda \langle \psi^\lambda | \phi_m \rangle \langle \phi_m | \hat{O} | \phi_n \rangle \langle \phi_n | \psi^\lambda \rangle \\ &= \sum_{m,n,\lambda} P_\lambda \langle \phi_n | \psi^\lambda \rangle \langle \psi^\lambda | \phi_m \rangle \langle \phi_m | \hat{O} | \phi_n \rangle \\ &= \sum_{m,n} \rho_{n,m} O_{m,n} \\ &\equiv \text{Tr}(\hat{\rho} \hat{O}). \end{aligned}$$

这里, $\{|\phi_m\rangle\}$ 为一组正交完备的基矢。我们定义了密度算符:

$$\boxed{\hat{\rho} = \sum_{\lambda} P_\lambda |\psi^\lambda\rangle \langle \psi^\lambda|}. \quad (10.4)$$

它在特定表象下可以表示为矩阵的形式称为密度矩阵。算符 \hat{O} 的系综平均可以表示为:

$$\boxed{\langle \hat{O} \rangle = \text{Tr}(\hat{\rho} \hat{O})}. \quad (10.5)$$

10.1.2 密度矩阵的性质

首先, 我们有以下几点说明:

- 式 (10.4) 中不同状态下的 $|\psi^\lambda\rangle$ 不一定正交。而 λ 的个数 Ω 是微观态的个数, 它不一定与系统所处希尔伯特空间的大小相同。在统计物理中, Ω 往往非常大, 但在一些其他量子系统中, Ω 可能只有有限几个。
- P_λ 为系综平均的权重, 这是一个经典权重系数。它的出现是由于我们对系统的微观信息缺失而导致的。不同状态 $|\psi^{\lambda_1}\rangle$ 和 $|\psi^{\lambda_2}\rangle$ 之间只是简单的加和, 没有相互关联。而任一给定状态 $|\psi^\lambda\rangle$ 可以是不同波函数的线性叠加形式 $|\psi^\lambda\rangle = \sum_m c_m^\lambda |\phi_m\rangle$ 。这里的权重因子

c_m^λ 为波函数线性叠加的‘量子权重’(概率福)， ϕ_i 和 ϕ_j 之间存在相干叠加的可能。

- 给定一组正交完备的基矢，我们可以将波函数表示为：

$$|\psi^\lambda\rangle = \sum_n c_n^\lambda |\phi_n\rangle.$$

密度算符的矩阵元可以表示为：

$$\rho_{mn} = \langle\phi_m|\hat{\rho}|\phi_n\rangle = \sum_\lambda P_\lambda \langle\phi_m|\psi^\lambda\rangle \langle\psi^\lambda|\phi_n\rangle = \sum_\lambda P_\lambda c_m^\lambda c_n^{\lambda*}. \quad (10.6)$$

其对角元 ρ_{mm} 为系统处于状态 ϕ_m 的概率。这个概率包括了量子概率 $|c_m^\lambda|^2$ 与统计概率 P_λ 。

- 一方面，同一个状态，在不同的表象下可以有不同的密度矩阵。另一方面，同一个表象下的同一个密度矩阵也可以对应不同的状态。也就是说，系统和密度矩阵之间不是一一对应的关系。

可以证明密度矩阵具有如下性质：

- **归一性**: $\text{Tr}\rho = 1$.
- **厄密性**: 由 (10.6) 可见，密度矩阵为厄密矩阵， $\rho_{mn} = \rho_{nm}^*$ 或者 $\hat{\rho} = \hat{\rho}^\dagger$ 。
- **正定性**: 密度矩阵的对角元 $1 \geq \rho_{ii} \geq 0$ 。由于其厄密性，我们总可以把密度矩阵对角化得到一组实本征值，每个本征值也满足 $1 \geq \rho_{ii} \geq 0$ 。
- $\text{Tr}\rho^2 \leq 1$ 。等号对应纯系综 (pure ensemble)，不等号对应混系综 (mixed ensemble)。

10.1.3 纯系综和混系综

在统计物理中，我们通常称纯系综 (pure ensemble) 和混系综 (mixed ensemble)。系综中的具体系统数通常非常大，比如远大于其微观状态数。在量子信息等领域，我们还会遇到纯态 (pure state) 和混态 (mixed state) 的说法。这时，涉及到的微观态数目往往不一定很多。两者其实没有本质的区别。

纯系综/纯态: 整个系综只能处于一个状态

$$P_\lambda = \begin{cases} 1 & \lambda = i \\ 0 & \text{otherwise} \end{cases}$$

这里需要注意的是，在给定基组下，状态 i 可以是不同状态的线性叠加

$$|\psi^i\rangle = \sum_m c_m^i |\phi_m\rangle$$

但它仍然是一个纯态。这时的密度矩阵元

$$\rho_{mn} = c_m^i c_n^{i*}$$

也就是说，纯系综的密度矩阵可以是非对角的，但它满足

$$\rho^2 = \rho.$$

它也是一个投影算符。纯系综密度矩阵的对角元 ρ_{mm} 给出了系统处于 $|\phi_m\rangle$ 这个状态的概率。而非对角元 ρ_{mn} 描述的是不同状态 $|\phi_m\rangle$ 和 $|\phi_n\rangle$ 之间的相干性。比如，系统处于位置 r 的概率为：

$$\rho(r) = \langle r | \rho | r \rangle = |\langle r | \psi \rangle|^2 = \sum_m |c_m^i \phi_m(r)|^2 + \sum_{m \neq n} 2 \operatorname{Re}(\rho_{mn} \phi_m(r) \phi_n^*(r)).$$

$\rho(r)$ 不仅取决于密度矩阵的对角元，而且可以来自于非对角元的相干贡献。

混系综/混态：系统可能处于不同的状态，也就是系综中包含了经典的统计概率分布。混系综中，系统可能所处的不同状态之间是相互独立的，每个状态 ψ^i 对应一个经典的概率 P_i 。它们之间只是一个简单加和，没有关联。

10.1.4 约化密度矩阵

对于给定的密度矩阵，我们可以针对部分自由度进行求迹操作，于是可以得到剩余自由度的约化密度矩阵。约化密度矩阵可以用来描述剩余自由度的性质。

10.1.5 密度算符的运动方程

对于一个不含时的量子系统，密度算符的运动方程为

$$i\hbar \frac{\partial \hat{\rho}}{\partial t} = [\hat{H}, \hat{\rho}]. \quad (10.7)$$

它被称为 von Neumann 方程或者量子 Liouville 方程。我们可以从 Schrödinger 方程和密度算符的定义得到。

- 当系统处于平衡态时，上式左边为零，所以 $\hat{\rho}$ 与 \hat{H} 对易，它们拥有共同的本征态。所以在热平衡状态下，正则密度矩阵在能量本征态表象下为对角矩阵 $\rho_{mn} = \delta_{m,n} P_n$ ，其中 P_n 为系统处于能量本征态 $|n\rangle$ 的概率。利用与经典系统类似的论证，我们可以得到：

$$P_n \propto e^{-\beta \varepsilon_n}.$$

配分函数为

$$Z = \sum_n e^{-\beta \varepsilon_n} = \text{Tr} \left(e^{-\beta \hat{H}} \right).$$

- 从另外一个角度考虑，我们可以在能量本征态 $|\varepsilon_n\rangle$ 下把密度算符写成矩阵的形式：

$$\rho_{mn}(t) = \sum_{\lambda} P_{\lambda} c_m^{\lambda} c_n^{\lambda*} e^{-i(\varepsilon_m - \varepsilon_n)t/\hbar}. \quad (10.8)$$

若使 $\rho_{mn}(t)$ 不依赖 t ，则要求当 $\varepsilon_m \neq \varepsilon_n$ 时，上式为零。也就是要求其非简并的两个状态之间的非对角元为零。若存在简并的情况，我们仍然可以选择合适的基矢使非对角元为零。于是，我们得出了与上一点相同的结论。密度矩阵在平衡态下总是可以在能量本征态表象下写成对角矩阵的形式。

- 定义系统演化算符为：

$$\hat{U}(t, t_0) = e^{-\frac{i}{\hbar} H(t-t_0)}.$$

我们可以把系统状态的演化表示为：

$$|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle.$$

所以，已知 t_0 时刻的密度算符 $\hat{\rho}(t_0)$ ， t 时刻的密度矩阵由上式可以得出：

$$\hat{\rho}(t) = \hat{U}(t, t_0)\hat{\rho}(t_0)\hat{U}^{\dagger}(t, t_0) = \hat{U}(t, t_0)\hat{\rho}(t_0)\hat{U}(t_0, t). \quad (10.9)$$

10.1.5.1 经典运动方程

对于一个不含时的经典体系，系统自由度的运动方程可以用 Hamilton 运动方程来描述：

$$\dot{q} = \frac{dq}{dt} = \frac{\partial H}{\partial p}, \quad \dot{p} = \frac{dp}{dt} = -\frac{\partial H}{\partial q}. \quad (10.10)$$

这里的 q 和 p 都是系统所有自由度的位移和动量构成的向量。这样的话，对于任意一个依赖于 q 和 p 的物理量 $O(p, q)$ ，其运动方程可以表示为：

$$\dot{O} = \frac{dO}{dt} = \frac{\partial O}{\partial q} \frac{\partial H}{\partial p} - \frac{\partial O}{\partial p} \frac{\partial H}{\partial q} \equiv -\{H, O\}. \quad (10.11)$$

这里我们定义了 Poisson 括号 $\{H, O\} = \frac{\partial H}{\partial q} \frac{\partial O}{\partial p} - \frac{\partial H}{\partial p} \frac{\partial O}{\partial q}$ 。

对于描述系统粒子分布的经典概率密度函数 $\rho(q, p, t)$ ，我们利用 Liouville 定理，

$$\dot{\rho} = \frac{d\rho}{dt} = 0, \quad (10.12)$$

也可以写出其运动方程:

$$\frac{\partial \rho}{\partial t} = -\frac{\partial \rho}{\partial q} \frac{\partial H}{\partial p} + \frac{\partial \rho}{\partial p} \frac{\partial H}{\partial q} \equiv \{H, \rho\}. \quad (10.13)$$

同样, 我们可以看到, 在系统处于热力学平衡态时, ρ 不依赖时间 $\partial \rho / \partial t = 0$, 即 $\{H, \rho\} = 0$ 。据此, 我们可以把 ρ 用 H 来表示出来 $\rho(H)$ 。

10.1.5.2 Liouville (超) 算符

上面经典和量子的情形的运动方程可以利用 Liouville 算符来统一起来:

$$\mathcal{L}\{O\} \equiv \begin{cases} i\{H, O\} & \text{经典} \\ \frac{1}{\hbar}[H, O] & \text{量子} \end{cases} \quad (10.14)$$

有了这个定义, 我们可以把经典和量子情形下的运动方程统一写为下面的形式:

$$\frac{\partial}{\partial t} O(t) = i\mathcal{L}\{O(t)\}. \quad (10.15)$$

于是我们可以得出一个 $O(t)$ 的形式解:

$$O(t) = e^{i\mathcal{L}t} O(0). \quad (10.16)$$

我们下面会用到这个形式解. 同理, 我们有:

$$\frac{\partial}{\partial t} \rho = -i\mathcal{L}\rho, \quad (10.17)$$

从而得到:

$$\rho(t) = e^{-i\mathcal{L}t} \rho(0). \quad (10.18)$$

通过 Liouville 算符我们可以把经典和量子的情形写成统一的形式。

10.1.5.3 Liouville 算符的厄密性

下面我们来说明 Liouville 算符是厄密算符。也就是说, 对于任何两个算符 A 和 B , 我们有

$$(A, \mathcal{L}B) = (\mathcal{L}A, B). \quad (10.19)$$

这里我们定义内积为

$$(A, B) = \begin{cases} \int dq \int dp A^*(q, p) B(q, p) & \text{经典} \\ \text{Tr}[A^\dagger B] & \text{量子} \end{cases} \quad (10.20)$$

经典情况

$$\begin{aligned}
(A, \mathcal{L}B) &= i \int dq \int dp A^* \left(\frac{\partial H}{\partial q} \frac{\partial B}{\partial p} - \frac{\partial H}{\partial p} \frac{\partial B}{\partial q} \right) \\
&= -i \int dq \int dp \left(\frac{\partial}{\partial p} \left(A^* \frac{\partial H}{\partial q} \right) B - \frac{\partial}{\partial q} \left(A^* \frac{\partial H}{\partial p} \right) B \right) + \text{boundary terms} \\
&= -i \int dq \int dp \left(\frac{\partial A^*}{\partial p} \frac{\partial H}{\partial q} B - \frac{\partial A^*}{\partial q} \frac{\partial H}{\partial p} B \right) \\
&= \int dq \int dp (\mathcal{L}A)^* B \\
&= (\mathcal{L}A, B).
\end{aligned}$$

量子情况

$$\begin{aligned}
(A, \mathcal{L}B) &= \text{Tr}[A^\dagger \mathcal{L}B] = \frac{1}{\hbar} \text{Tr}[A^\dagger [H, B]] = \frac{1}{\hbar} \text{Tr}[A^\dagger HB - A^\dagger BH] \\
&= \frac{1}{\hbar} \text{Tr}[A^\dagger HB - HA^\dagger B] = \frac{1}{\hbar} \text{Tr}[([H, A])^\dagger B] = \text{Tr}[(\mathcal{L}A)^\dagger B] = (\mathcal{L}A, B).
\end{aligned}$$

可见，不管是经典还是量子的情况，Liouville 算符都是厄密算符。

10.2 von Neumann 熵

利用密度矩阵我们可以定义量子系统的 von Neumann 熵，它与经典系统的 Gibbs-Shannon 熵对应：

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

密度矩阵为厄密矩阵（见下面讨论），因此我们总是可以找到一个将其对角化的表示。在对角表象下，它的形式和 Gibbs-Shannon 熵完全相同。

我们可以通过下面的例子来进一步说明。已知一个两粒子系统处于纯态

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow_1 \downarrow_2\rangle + |\downarrow_1 \uparrow_2\rangle).$$

对应的密度矩阵为

$$\begin{aligned}\rho &= |\psi\rangle\langle\psi| \\ &= \frac{1}{2}(|\uparrow_1\downarrow_2\rangle\langle\uparrow_1\downarrow_2|\psi| + |\uparrow_1\downarrow_2\rangle\langle\downarrow_1\uparrow_2|\psi| + |\downarrow_1\uparrow_2\rangle\langle\uparrow_1\downarrow_2|\psi| + |\downarrow_1\uparrow_2\rangle\langle\downarrow_1\uparrow_2|\psi|) \\ &= \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\end{aligned}$$

因为其为纯态，可以验证 $\rho^2 = \rho$ 。将其转换到对角表象，可得其熵为 $S = 0$ 。

下面，我们相对粒子 1 求迹，得到粒子 2 的约化密度矩阵

$$\rho_2 = \text{Tr}_1(|\psi\rangle\langle\psi|) = \frac{1}{2}(|\uparrow_2\rangle\langle\uparrow_2| + |\downarrow_2\rangle\langle\downarrow_2|).$$

这时子系统 2 处于混态，对应的熵为

$$S_2 = k_B \ln 2.$$

可见，在粒子 1 和 2 组成的空间中，系统处于纯态，熵为 0；但约化到粒子 2 所处的子系统后，熵增加为 $k_B \ln 2$ 。因为通过部分求迹操作后，我们丧失对粒子 1 所处状态的了解，丢失了部分信息。

10.3 Wigner 变换和 Wigner 分布函数

物理量 \hat{O} 的期待值利用密度矩阵可以表示为：

$$\langle O \rangle = \text{Tr}(\hat{A}\hat{\rho}).$$

我们可以把上式在坐标表象下表示出来：

$$\langle O \rangle = \int dq \int dq' A(q, q') \rho(q', q).$$

其中， $\rho(q', q) = \langle q'|\rho|q\rangle$ 。Wigner 定义了如下的变换：

$$f(\bar{q}, p) = \int d\Delta q f(\bar{q} + \Delta q/2, \bar{q} - \Delta q/2) e^{ip\Delta q/\hbar}, \quad (10.21)$$

$$f(\bar{q} + \Delta q/2, \bar{q} - \Delta q/2) = \int \frac{dp}{2\pi\hbar} f(\bar{q}, p) e^{-ip\Delta q/\hbar}. \quad (10.22)$$

这样的话， \hat{O} 的期待值可以表示为：

$$\langle O \rangle = h^{-1} \int dq \int dp \rho(q, p) O(q, p).$$

这个形式与经典的统计平均相同。这里的 $\rho(q, p)$ 称为 Wigner(分布) 函数，对应经典系统的概率密度分布。虽然 Wigner 函数取实数值，但它可正可负。因此，Wigner 函数不是经典意义的概率密度分布，其负值恰恰对应了系统的量子特征。

10.4 举例

10.4.1 二能级系统

下面以一个简单的二能级系统 $|\uparrow\rangle, |\downarrow\rangle$ 为例来说明以上内容。

- (1) 系统处于状态 $|\uparrow\rangle$ 的概率为 $1/2$ ，处于 $|\downarrow\rangle$ 的概率为 $1/2$ 。取

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (10.23)$$

这时的密度矩阵为：

$$\rho = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (10.24)$$

(2) 系统各有 $1/2$ 的概率处于 $|+\rangle$ 和 $|-\rangle$ 态，这里 $|\pm\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle \pm |\downarrow\rangle)$ 。可以验证，这时的密度矩阵仍然由 (10.24) 给出。尽管这是两个不同的混态系综。

- 纯态 $|\pm\rangle$ 的密度矩阵

$$\rho_{\pm} = \frac{1}{2} \begin{pmatrix} 1 & \pm 1 \\ \pm 1 & 1 \end{pmatrix} \quad (10.25)$$

可以验证 $\rho^2 = \rho$ 。

- 注意

$$|+\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$$

与

$$\rho = \frac{1}{2}|\uparrow\rangle\langle\uparrow| + \frac{1}{2}|\downarrow\rangle\langle\downarrow|$$

的区别。

10.4.2 理想气体

密度矩阵本征表象下可以表示为：

$$\rho = \sum_k P_k |k\rangle\langle k|.$$

其中，

$$\phi_k(q) = \langle q|k\rangle = \frac{1}{\sqrt{V}} e^{i\vec{k}\cdot\vec{q}}.$$

在坐标表象下的表示为：

$$\begin{aligned} \rho(q, q') &= \langle q|\rho|q'\rangle \\ &= \frac{1}{ZV} \sum_k e^{-\beta\hbar^2 k^2/2m} e^{i\vec{k}\cdot(\vec{q}-\vec{q}')} \\ &= \frac{1}{V} \exp\left\{-\frac{\pi(q-q')^2}{\lambda_{th}^2}\right\}. \end{aligned} \quad (10.26)$$

这里，热波长 λ_{th} 由式 (11.2) 给出。

- 坐标表象下，密度矩阵对角元给出粒子处于坐标 q 的概率 V^{-1} 。由于系统具有平移对称性，这个概率与具体坐标无关。
- 密度矩阵的非对角元表征了两个不同坐标位置之间的关联，它随着两个粒子之间的距离呈指数衰减，对应的特征长度为热波长 λ_{th} 。

进行 Wigner 变换，我们可以得到粒子的 Wigner 分布函数为：

$$\rho(\vec{q}, \vec{p}) = \frac{\lambda_{th}^3}{V} \exp\left\{-\frac{p^2}{4\pi(\hbar/\lambda_{th})^2}\right\} = \lambda_{th}^3 V^{-1} \exp\left\{-\frac{p^2}{2mk_B T}\right\}. \quad (10.27)$$

可见，在经典极限下，Wigner 函数退化为 Maxwell-Boltzmann 分布

$$\rho(\vec{v}) d\vec{v} = \left(\frac{m}{2\pi k_B T}\right)^{\frac{3}{2}} \exp\left\{-\frac{mv^2}{2k_B T}\right\} d\vec{v}.$$

10.4.3 谐振子

在热平衡状态下，量子谐振子在能量本征态下的密度矩阵可以表示为：

$$\rho_{nm} = \delta_{n,m} e^{-(n+1/2)x}/Z,$$

配分函数

$$Z = \left(2 \sinh \frac{1}{2}x\right)^{-1}, \quad x = \beta\hbar\omega_0.$$

在坐标表象下，密度矩阵可以表示为：

$$\begin{aligned}\rho(q, q') &= \langle q | \rho | q' \rangle = \sum_n \langle q | n \rangle \langle n | \rho | n \rangle \langle n | q' \rangle \\ &= 2 \sinh \frac{1}{2}x \sum_n e^{-(n+1/2)x} \phi_n(q) \phi_n(q') \\ &= \left(\frac{m\omega_0}{\pi\hbar} \tanh \frac{x}{2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{m\omega_0}{4\hbar} [(q + q')^2 \tanh(x/2) + (q - q')^2 \coth(x/2)] \right\} \quad (10.28)\end{aligned}$$

$$= \left(\frac{1}{2\pi l^2} \tanh \frac{x}{2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{1}{8l^2} [(q + q')^2 \tanh(x/2) + (q - q')^2 \coth(x/2)] \right\}. \quad (10.29)$$

密度矩阵对角元给出谐振子处于位置 q 的概率：

$$\rho(q) = \left(\frac{\tanh(x/2)}{2\pi l^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{q^2}{2l^2} \tanh(x/2) \right\}.$$

这是一个高斯分布。

- 低温极限下 $T \rightarrow 0$: $\tanh(x/2) \rightarrow 1$ 高斯分布的展宽由谐振子的特征长度决定：

$$\rho(q) = \frac{1}{\sqrt{2\pi l^2}} \exp \left\{ -\frac{q^2}{2l^2} \right\}.$$

因为这时，谐振子处于基态。

- 高温极限下 $T \rightarrow \infty$: $\tanh(x/2) \rightarrow x/2$ 高斯分布的展宽由温度决定：

$$\rho(q) = \sqrt{\frac{m\omega_0^2}{2\pi k_B T}} \exp \left\{ -\frac{m\omega_0 q^2}{2k_B T} \right\}.$$

试证明上式。

10.4.3.1 Wigner 分布函数

(1) 角频率为 ω_0 , 质量为 m 的量子谐振子基态的 Wigner 函数为：

$$\rho_0(q, p) = 2 \exp \left\{ -\frac{q^2}{2l^2} - \frac{p^2}{\hbar^2/2l^2} \right\}. \quad (10.30)$$

第一激发态的 Wigner 函数为：

$$\rho_1(q, p) = 2 \left(\frac{q^2}{2l^2} + \frac{p^2}{\hbar^2/2l^2} - \frac{1}{2} \right) \exp \left\{ -\frac{q^2}{2l^2} - \frac{p^2}{\hbar^2/2l^2} \right\}. \quad (10.31)$$

基态的 Wigner 函数在相空间为高斯函数的形式，因此 $\rho_0(q, p) \geq 0$ ；但 $\rho_1(q = 0, p = 0) < 0$ 。

(2) 将式 (10.29) 做 Wigner 变换，我们可以得到谐振子在热平衡状态下的 Wigner 分布函数：

$$\rho(q, p) = 2 \tanh \left(\frac{x}{2} \right) \exp \left\{ -\frac{q^2 \tanh(x/2)}{2l^2} \right\} \exp \left\{ -\frac{p^2 \tanh(x/2)}{\hbar^2/2l^2} \right\} \quad (10.32)$$

我们可以验证，低温极限下 $T \rightarrow 0$ ，密度矩阵趋于式 (21.1)；在高温极限下，密度矩阵趋于

Maxwell-Boltzmann 分布。

第 11 章 理想气体



本章我们利用前面介绍的统计物理基本方法来继续分析理想气体。我们可以得到理想气体的内能、自由能、熵等热力学量的表示及其状态方程。同时，我们也会去继续讨论并解决上一章留下来的一个问题：用经典统计的方法得到的熵、自由能不是广延量。

11.1 系统

我们考虑放置于边长为 L 的立方体中的单原子理想气体。设分子数为 N ，那么理想气体的密度就是 $n = N/V$, $V = L^3$ 。气体处于温度为 T 的平衡态下。这可以通过把立方气与温度为 T 的热库做热接触，如图11.1左图所示。在上一章中我们已经从经典统计的角度对这个问题进行了分析。但我们没有给出配分函数的完整表达式。这里我们更进一步，考虑粒子的量子状态，从而得到其完整形式。

11.2 单粒子配分函数

我们首先考虑单粒子配分函数，而要得到配分函数，我们需要知道粒子的能谱。这里，我们需要利用到一点量子力学的知识。

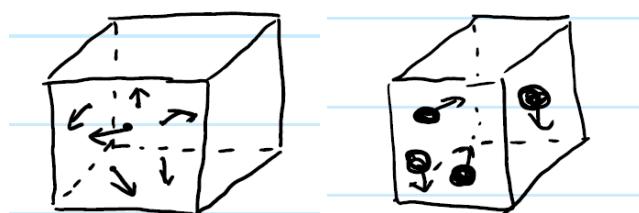


图 11.1: 理想气体模型

11.2.1 态密度

因为考虑理想气体，我们可以忽略分子间的相互作用。这时，单分子的哈密顿量就只包括其动能项

$$\varepsilon = \frac{\hbar k^2}{2m}.$$

我们还需要选择一个边界条件。这里有两种选择：(1). 周期性边界条件；(2). 固定边界条件。对于热力学系统，两种边界条件应该给出相同的结果¹。我们这里以周期性边界条件为例来讨论。

考虑平面波形式的波函数 $\psi(\vec{r}) = e^{i\vec{k}\cdot\vec{r}}$ 。要满足周期性边界条件

$$\psi(x = L) = \psi(x = 0), \quad \psi(y = L) = \psi(y = 0), \quad \psi(z = L) = \psi(z = 0),$$

就要求波矢 \vec{k} 满足

$$k_x L = 2\pi n, \quad k_y L = 2\pi n, \quad k_z L = 2\pi n, \quad n = 0, \pm 1, \pm 2, \dots$$

所以，波矢不再像自由空间那样取连续值，只能取离散值，即

$$k_{\alpha,n} = \frac{2\pi n}{L_\alpha}, \quad \alpha = x, y, z.$$

根据以上结果，我们可以得到在波矢空间，每个单粒子态所占的体积元

$$\Delta k = \frac{(2\pi)^3}{V}.$$

可以看到，实空间体积越大，波矢空间中每个状态所占的体积元就越小。

对于宏观的热力学系统，我们可以认为波矢的取值是连续的，从而把配分函数中对所有状态求和转化为积分。比如

$$\sum_{\vec{k}} (\dots) = \int_0^\infty dk g(k) (\dots)$$

上式中的 $g(k)$ 称为态密度。这里 $k = |\vec{k}|$ 表示波矢的大小。不难验证，它可以表示为

$$g(k) = \frac{V k^2}{2\pi^2}.$$

这样，在范围 $[k, k + dk]$ 内的状态数就可以用 $g(k)dk$ 表示。

练习 11.1 (1) 如何把态密度用能量为参数表示出来？ $g(\varepsilon)d\varepsilon = \dots$ ？(2) 如何表示二维和一维系统的态密度？(3) 如果考虑极端相对论气体呢？

¹两者的细微差别可以参考 Pathria, Appendixes A.

11.2.2 单粒子配分函数

有了态密度，我们就可以很方便的来计算单粒子的配分函数

$$\begin{aligned}
 Z_1 &= \sum_{\vec{k}} e^{-\beta \varepsilon(\vec{k})} \\
 &= \int_0^{\infty} dk g(k) e^{-\beta \varepsilon(\vec{k})} \\
 &= \frac{V}{2\pi^2} \int_0^{\infty} dk k^2 e^{-\beta \varepsilon(\vec{k})} \\
 &= V \left(\frac{2\pi m k_B T}{h^2} \right)^{\frac{3}{2}} \\
 &\equiv V n_Q.
 \end{aligned} \tag{11.1}$$

我们在最后一步定义了量子浓度 n_Q 。同时，我们还可以定义粒子的热波长

$$\lambda_{\text{th}} = \sqrt{\frac{2\pi\hbar^2}{mk_B T}} = \frac{h}{\sqrt{2\pi m k_B T}}. \tag{11.2}$$

它们之间满足

$$n_Q = \frac{1}{\lambda_{\text{th}}^3}.$$

λ_{th} 的物理意义是温度为 T 的理想气体中粒子的平均 de Broglie 波长。粒子的平均平动动能为 $U = \frac{3}{2}k_B T$ ，对应的动量大小为 $p = \sqrt{2mU} = \sqrt{3mk_B T}$ 。由于波粒二象性，动量为 p 的粒子对应的波矢为 $k = p/\hbar$ ，所以波长为 $\lambda = 2\pi/k = h/\sqrt{3mk_B T}$ 。这与我们定义的热波长相差不多。

11.3 N 粒子配分函数：第一步尝试

得到单粒子配分函数后，我们可以根据上一章的讨论，把 N 粒子配分函数写为

$$Z_N = (Z_1)^N$$

由此，我们可以去计算其他的热力学量。我们出发点仍然是：

$$\ln Z_N = N \ln Z_1 = N \left(\ln V + \frac{3}{2} \ln (2\pi m k_B T) - 3 \ln h \right).$$

上式与上一章的式 (9.36) 相比只是多出了最后与 Planck 常数相关的一项 $-3N \ln h$ 。尽管我们的出发点是量子理论，但我们只考虑了单个粒子的能级量子化，更重要的是对于宏观系统，我

们把对量子态的求和变成了积分。这一步很大程度上把我们又带回了经典理论，但它保留了一个重要的量子特征 h 。正是这个 h 解决了我们经典理论中无法确定配分函数完整形式的问题。通过对比，我们知道需要在式 (9.30) 和 (9.35) 中加入额外的因子 $1/h^{3N}$ 来得到正确的量纲，而 h^3 其实是每个粒子在相空间的“量子化”体积元。这是 Heisenberg 不确定性原理的直接结果。在相空间中我们无法同时确定粒子的位置和动量，因此相空间体积元不能无限细分。它的最小单位为 h^3 。

下面我们来计算对应的热力学量。**内能：**

$$U = -\frac{\partial \ln Z_N}{\partial \beta} = \frac{3}{2} N k_B T. \quad (11.3)$$

这个结果与能量均分定理给出的结果一致。

自由能：

$$F = -N k_B T \ln Z_1 = -N k_B T \left(\ln V + \frac{3}{2} \ln \frac{2\pi m k_B T}{h^2} \right). \quad (11.4)$$

压强：

$$P = - \left(\frac{\partial F}{\partial V} \right)_T = -N k_B T / V. \quad (11.5)$$

熵：

$$S = \frac{U - F}{T} = N k_B \left(\frac{3}{2} + \ln(V n_Q) \right). \quad (11.6)$$

我们看到，和上章使用经典理论得到的结果相比，我们最重要的进展就是得到了自由能和熵的正确量纲。但它们仍然不是广延量。所以，以上结果还有问题。

11.4 N 粒子配分函数：粒子的不可分辨性

对上面这一问题的解决方案是把 Z_N 写为：

$$Z_N = \frac{1}{N!} Z_1^N. \quad (11.7)$$

这里引入的因子 $1/N!$ 用来表示粒子的**不可分辨性**。写出上节的 N 粒子配分函数时，我们默认粒子是可以分辨的，但实际上我们并没有办法区分系统中的每个粒子。如果任意调换其中两个粒子的位置，我们完全无法区分调换前后的两个状态，因此应该认为是一个微观态。可见，粒子是否可以分辨和系统微观状态数的多少有直接关联。可分辨粒子可以组成更多的状态。

比如，以图11.2所示的 3 个粒子为例。如果这三个粒子是可以分辨的（分别用三角、正方形和圆形表示），那么三个粒子分别处于 1,2,3 三个位置的微观状态数有 $3! = 6$ 个，因为我们

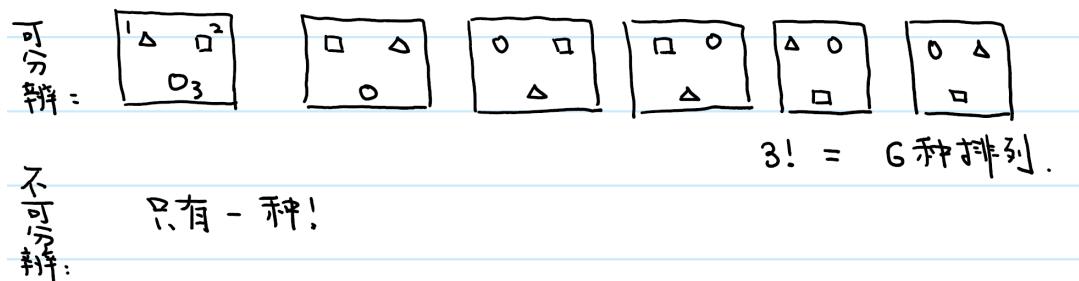


图 11.2: 可分辨和不可分辨粒子

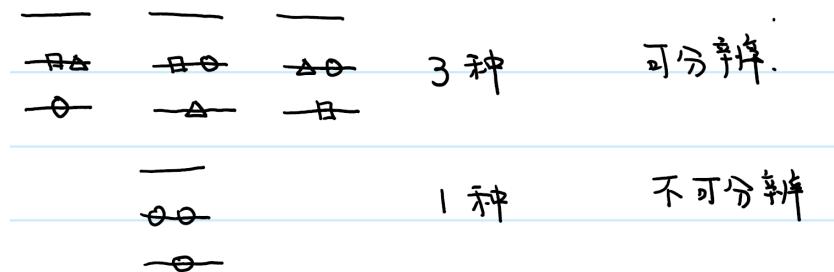


图 11.3: 可分辨和不可分辨粒子

任意交换两个粒子的位置都可以得到一个新的微观态。但如果这三个粒子是不可分辨的，那么就只剩下一个微观态了。它们正好相差粒子数的阶乘倍。这正是增加因子 $1/N!$ 的原因。

这时，细心的同学可能会提出异议：并不是所有情况下，可分辨和不可分辨情况微观状态数之比正好是 $N!$ ，比如图11.3所示的情况。同样有 3 个粒子，可分辨和不可分辨的状态数之比为 3 而不是 6。这两种情况的区别在于：图11.2中，每个粒子所处的单粒子态都不同；而图11.3中，有两个粒子处于相同的单粒子态。所以，除了粒子是否可以分辨外，是否有粒子处于相同的状态也会影响微观状态数。要考虑后者的话，问题会变得非常复杂。这里，我们暂时忽略这一因素²，假设每个粒子都处于不同的单粒子态；所以只要考虑因子 $1/N!$ 就可以了。粒子数密度 n 越小，两个粒子处于相同单粒子态的情况就越少，我们的这个近似就越准确。所以，对于稀薄气体，我们这么考虑问题不大。如果 n 较大的话（相对于 n_Q ），我们的近似就不再适用。**总结一下，考虑理想气体模型，在 n 较小的情况下，我们需要在上节的 N 粒子配分函数基础上乘以因子 $1/N!$ 用于考虑粒子的不可分辨性。**

利用式 (11.7) 我们可以重复上一节的计算，得到气体的热力学量。很容易验证，内能和压强不会发生变化，但自由能和熵的结果发生了变化。新的表达式都与 N 成正比，也就是满足它们是广延量的条件。我们上一节的问题就得到了解决。

²对这一问题的全面考虑是量子统计的内容。

自由能:

$$F = -Nk_B T \left(\ln \frac{n_Q}{n} + 1 \right). \quad (11.8)$$

熵:

$$S = \frac{U - F}{T} = Nk_B \left(\frac{5}{2} + \ln \frac{n_Q}{n} \right). \quad (11.9)$$

结果讨论:

- 我们考虑单粒子态时是从量子力学出发的, 但我们最后得到的其实是经典的结果。比如, 内能的结果与经典的能量均分定理结果相同; 熵在温度趋于零时并不会趋于零, 不符合第三定律。这主要是因为我们考虑宏观的热力学系统, 从而把离散的单粒子态求和转化为了积分。
- 尽管我们做了经典近似, 我们最后的结果中仍然包含 \hbar 或者 h , 它隐含在 n_Q 中。这说明, 即使我们考虑经典体系, 从统计物理的角度来分析问题时, 我们也需要考虑一定的量子效应。这里的量子效应主要体现在以 h 表示的相空间体积元量子化。
- 我们上面对于粒子不可分辨性的讨论都是定性的。这里我们把以上结果稍微具体一下。我们引入了量子浓度 $n_Q = 1/\lambda_{\text{th}}^3$ 。在熵和自由能的表达式中都包含了 $\ln(n_Q/n)$ 。我们所说的粒子数密度足够小, 实际上是要求 $n \ll n_Q$, 即 $n_Q/n \gg 1$ 。这时, 每个粒子所占据的体积 $V/N \gg \lambda_{\text{th}}^3$ 。我们可以把热波长 λ_{th} 理解为粒子在温度为 T 时的 de Broglie 波长, 而 λ_{th}^3 就是它所占据的空间。如果 $V/N \gg \lambda_{\text{th}}^3$, 那么两个粒子处于同一位置的概率就很低。

11.5 Sackur-Tetrode 方程

我们得到的式 (11.9) 又被称为 Sackur-Tetrode 方程。它是由 Sackur 和 Tetrode 分别独立得到的³。当时他们采用的并不是我们这里的方法, 而是基于微正则系综, 直接写出了熵的表达式。

³<https://www.aps.org/publications/apsnews/200908/physicshistory.cfm>

11.5.1 单个粒子的微观状态数

单个粒子对应微观状态数 Ω_1 可以表示为：

$$\Omega_1 \propto VV_p$$

这里， V_p 为粒子在动量空间所占的“体积”。为了得到一个等式，我们可以假设每个粒子在实空间每个方向所占的长度为 Δx 、 Δy 、 Δz 。相应的，在动量空间的长度为 Δp_x 、 Δp_y 、 Δp_z 。这样我们有

$$\Omega_1 = \left(\frac{L_x}{\Delta x} \right) \left(\frac{L_y}{\Delta y} \right) \left(\frac{L_z}{\Delta z} \right) \left(\frac{L_{p_x}}{\Delta p_x} \right) \left(\frac{L_{p_y}}{\Delta p_y} \right) \left(\frac{L_{p_z}}{\Delta p_z} \right)$$

在经典理论中，我们可以同时指定粒子的位置和动量。如果我们忽略其大小的话， $\Delta x/y/z$ 和 $\Delta p_{x/y/z}$ 可以无限小，而 Ω_1 显然取决于我们对位置和动量描述的“精度”。这给我们带来很大的麻烦。幸运的是，这个麻烦可以通过量子理论来解决。在量子力学中，Heisenberg 测不准原理决定了我们无法无限精确的同时确定粒子的位置和动量。它们之间的不确定度

$$\Delta x \Delta p_x \sim h$$

我们这里不妨取等号，把 Ω_1 表示为

$$\Omega_1 = \frac{1}{h^3} VV_p, \quad (11.10)$$

从而把普朗克常数 h 引入到我们的理论中。如果要考虑 N 个粒子，我们还需要考虑粒子的不可分辨性。于是

$$\Omega_N = \frac{1}{N!} \frac{V^N V_{Np}}{h^{3N}}, \quad (11.11)$$

上式中每一项的物理意义非常明显： V^N 表示了粒子实空间位置的不确定性； V_{Np} 表示了动量空间的不确定性； h^{3N} 表示了量子不确定性； $N!$ 表示粒子的不可分辨性。

11.5.2 动量空间“体积” V_p

如果粒子的内能为 E ，那么其平均动量大小为 $p = \sqrt{2mE}$ 。对于单个粒子，在 p_x 、 p_y 和 p_z 三维空间中，它位于半径为 $\sqrt{2mE}$ 的球面上。这时， V_p 对应球面的面积。像上一章所描述的那样，我们可以上面的结果推广到 N 粒子系统。这时，我们发现 V_p 正比于 $3N - 1$ 维的超球面面积（式9.31）。

11.5.3 Sackur-Tetrode (1911)

根据以上结果，我们把 N 粒子系统对应的微观状态数表示为：

$$\Omega_N = \frac{V^N}{N!h^{3N}} \frac{2\pi^{3N/2} (2mE)^{(3N-1)/2}}{(3N/2 - 1)!}. \quad (11.12)$$

这个结果和上一章我们得到的式 (9.32) 非常类似。只不过多了两项：一项是 $N!$ 用来表示粒子的不可分辨性，一项是 h^{3N} 用来表示 $3N$ 个自由度的相空间体积元。这两项也正是我们本章相对于上一章的重要进步。把上式带入到熵的表达式中，利用 $E = \frac{3}{2}Nk_B T$ ，我们得到 Sackur-Tetrode 方程。

11.6 Gibbs 佯谬

考虑了粒子的不可分辨性后我们在第 6.2.2 节提到的 Gibbs 佯谬就可以得到解决。首先，如果是两种不同气体 A 和 B 的混合。那么，混合前后 A 和 B 的熵变之和为：

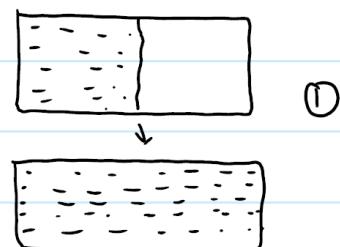
$$\Delta S = \Delta S_A + \Delta S_B = \sum_{X=A,B} N_X k_B \ln \frac{V}{V_X}.$$

但如果是同种气体，由于气体分子的不可区分性，熵变为零。

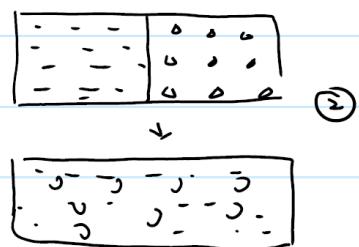
三. Gibbs 佯謬:

$$\Delta S^{\text{①}} = Nk_B \ln\left(\frac{n_a}{n/2}\right) - Nk_B \ln\left(\frac{n_a}{n}\right)$$

$$= Nk_B \ln 2.$$



$$\Delta S^{\text{②}} = 2Nk_B \ln 2.$$



$\Delta S = 0$. 不可分辨性!!

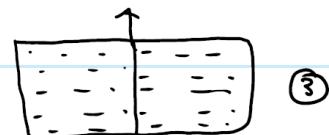
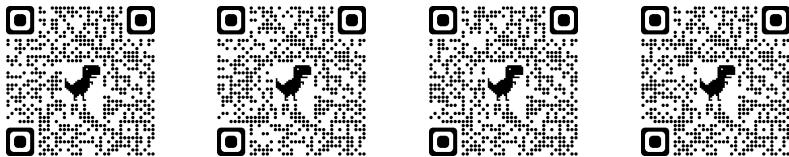


图 11.4: Gibbs paradox

第 12 章 光子和声子气体



12.1 光子气体

12.1.1 系统

我们考虑的系统为一个光腔内的电磁波模式。这些模式的空间和时间分布可以通过在给定边界条件下，求解经典的 Maxwell 方程来得到。这样得到的每个模式之间是相互独立的。对于给定一个角频率为 $\omega_k = ck$ 的模式，其能量按照经典电磁学的结果应该是连续的。我们可以把它看成是具有相同角频率的谐振子。如果按照经典统计物理的方法来处理光子气体，就会出现大家熟知的紫外灾难的问题。这是因为光子的角频率没有上限，按照经典的统计物理结果，每个模式在温度为 T 的环境下平局能量为 $k_B T$ 。为了避免这样的问题，历史上，Planck 做出了能量量子化的假设。也就是说，假设每个电磁波模式的能级是量子化、不连续的 $E_k = n\hbar\omega_k$, $n = 1, 2, 3, \dots$ 。下面我们将从 Planck 的量子化假设出发，来考虑光子气体的统计性质。

12.1.2 态密度

与理想气体的情况类似，我们需要知道光子气体的态密度

$$g(k)dk = 2 \times \frac{V}{2\pi^2} k^2 dk.$$

这里的因子 2 表示电磁波的两种极化模式¹。光子气体的色散关系与经典的自由粒子不同，

$$\omega_k = ck$$

因此，其态密度用频率来表示的话可以写为

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

¹光子自旋为 1，而自旋为 1 的粒子的简并度本来应该是 3，对应 $S_z = -1, 0, 1$ 。但因为光子没有质量，其 $S_z = 0$ 的状态不允许。因此其简并度为 2。

12.1.3 光子气体的统计力学

有了这些准备知识，我们就可以来考虑光子气体的统计力学。我们仍然是首先写出其配分函数

$$\begin{aligned}
 \ln Z &= \sum_k \ln Z_k = \int_0^{+\infty} dk g(k) \ln(1 - e^{-\beta \hbar \omega_k})^{-1} \\
 &= -\frac{V}{\pi^2 c^3} \int_0^{+\infty} d\omega \omega^2 \ln(1 - e^{-\beta \hbar \omega}) \\
 &= -\frac{V}{\hbar^3 \beta^3 \pi^2 c^3} \int_0^{+\infty} dx x^2 \ln(1 - e^{-x}) \\
 &= \frac{V}{\hbar^3 \beta^3 \pi^2 c^3} \int_0^{+\infty} \frac{x^3}{3} \frac{1}{e^x - 1} \\
 &= \frac{V \pi^2}{45 \hbar^3 \beta^3 c^3}.
 \end{aligned} \tag{12.1}$$

上式中，从第二行到第三行，我们定义了无量纲的量 $x = \beta \hbar \omega$ ；从第三行到第四行，我们做了分部积分。得到光子气体的配分函数后，我们就可以方便的对其热力学量进行计算了。

12.1.3.1 自由能

$$F = -k_B T \ln Z = -\frac{V \pi^2 k_B^4 T^4}{45 \hbar^3 c^3}. \tag{12.2}$$

12.1.3.2 内能

$$U = -\frac{\partial \ln Z}{\partial \beta} = \frac{V \pi^2 k_B^4 T^4}{15 \hbar^3 c^3} = -3F. \tag{12.3}$$

或者，我们可以写成内能密度的形式

$$u = \frac{\pi^2 k_B^4 T^4}{15 \hbar^3 c^3}. \tag{12.4}$$

这里，我们可以看出，光子气体的内能（密度）与 T^4 成正比。

如果我们在光腔腔壁凿一个小孔，根据我们气体动理论中得到的泻流公式，我们可以得到腔体向外辐射的能流：

$$J = \frac{1}{4} c u = \sigma T^4, \tag{12.5}$$

这里

$$\sigma = \frac{\pi^2 k_B^4}{60 \hbar^3 c^2}$$

为 Stefan-Boltzmann 常数。这正是黑体辐射的 Stefan-Boltzmann 定律，它只依赖于腔的温度，与其他性质无关。

12.1.3.3 压强

$$P = \frac{\partial F}{\partial V} = \frac{\pi^2 k_B^4 T^4}{45 \hbar^3 c^3} = -\frac{F}{V} = \frac{U}{3V}. \quad (12.6)$$

所以，光子气体的状态方程可以写成

$$PV = \frac{1}{3}U.$$

零点能、Casimir 效应：我们现在知道每个电磁波模式可以看做一个谐振子，我们把谐振子量子化后，它们基态的能量非零，对应零点涨落的能量 $\hbar\omega/2$ 。Casimir 考虑了零点能的可观测效应，并预测：把两块儿金属板靠得很近时，由于金属板之间和金属板外部电磁波模式的状态数目不同，会产生一个辐射压之差，导致两块儿板倾向于互相吸引。这被称为 Casimir 效应，对应的吸引力称为 Casimir 力，而且已经被实验证实。

估算：请估算一下太阳光照到你身上后，给你施加的辐射压力有多大。

12.1.3.4 熵

$$S = \frac{U - F}{T} = \frac{4V\pi^2 k_B^4 T^3}{45 \hbar^3 c^3} = \frac{4U}{3T}. \quad (12.7)$$

对于等熵过程，我们可得：

$$TV^{1/3} = \text{Constant.}$$

平均光子数：

$$\langle N \rangle = \int_0^{+\infty} d\omega g(\omega) n(\omega) = \frac{V\zeta(3)\Gamma(3)}{\pi^2(\hbar c\beta)^3}. \quad (12.8)$$

这样单位光子的熵可以表示为：

$$s = S/\langle N \rangle = \frac{2\pi^4}{45\zeta(3)} k_B \approx 3.6k_B. \quad (12.9)$$

我们可以把这个结果与经典非相对论理想气体的结果进行对比。哪个更大，为什么？

- ✍ **练习 12.1 光子气体 Carnot 热机：**请利用本节得到的结果来分析工作物质为光子气体的 Carnot 热机的工作流程：(1) 画出其 P - V 图， P - S 图；(2) 验证其效率仍然是 $\eta_c = 1 - T_l/T_h$ 。
- ✍ **练习 12.2** 计算一维光子气体的热容。

12.1.4 能谱密度

上面我们考虑光子气体的时候关心的是气体的总能量。我们还可以进一步考察气体的能谱密度，即给定频率的电磁波对系统能量的贡献。这时，我们不需要做频率积分，得到：

$$u(\omega)d\omega = \frac{V}{\pi^2 c^3} \frac{\hbar\omega^3}{e^{\beta\hbar\omega} - 1} d\omega. \quad (12.10)$$

上式给出了温度为 T 的光子气体的能谱密度，也是著名的**黑体辐射能谱密度分布**。我们还可以把它写成一个更容易理解的形式

$$u(\omega)d\omega = g(\omega)n(\omega)\hbar\omega d\omega. \quad (12.11)$$

其中，我们定义了 Planck 分布²：

$$n(\omega) = \frac{1}{e^{\beta\hbar\omega} - 1}. \quad (12.12)$$

- **Wien 位移定律：**我们可以分析上式的能谱密度在高频和低频极限的情况。在高频极限下 ($\hbar\omega \gg k_B T$)， $u(\omega)$ 以指数形式衰减为零；在低频极限下 ($\hbar\omega \ll k_B T$)， $u(\omega)$ 以 ω^2 的形式趋于零。这说明， $u(\omega)$ 在给定温度下必然存在一个极大值，对应的频率可以通过

$$\frac{\partial u(\omega)}{\partial \omega} = 0$$

得到 ω_{max} 。它满足：

$$xe^x = 3, \quad x = \beta\hbar\omega_{max}.$$

可见， $\beta\hbar\omega_{max}$ 为一常数。这正是 Wien 位移定律。据此，假设辐射物体为黑体，我们可以通过测量其辐射谱的 ω_{max} 来得到它的大致温度。我们也可以把式 (12.10) 写成波长的

²这里，Planck 分布形式上和我们后面介绍的 $\mu = 0$ 时的 Bose-Einstein 分布形式相同。我们知道，光子实际是玻色子，而又由于光子数不守恒，可以产生和湮灭，在热平衡的情况下其化学势为 0，所以它的分布函数必然是这种形式。但需要注意的是，这里，我们并没有应用光子是玻色子来得出其分布。我们唯一的假设是每个电磁波模式的能量是量子化的 $E_n = n\hbar\omega$ 。

形式：

$$u(\lambda)d\lambda = \frac{8\pi hc/\lambda^5 d\lambda}{e^{\beta hc/\lambda} - 1}. \quad (12.13)$$

由 $\partial u(\lambda)/\partial\lambda = 0$ ，我们做类似分析，从而得到：黑体辐射谱中最高功率密度的波长与温度的乘积为一个常数。

- **紫外灾难：**我们可以对式 (12.10) 取经典极限，取 $\beta\hbar\omega$ 为小量，我们可以把 Planck 分布近似为 $n(\omega) \approx k_B T / \hbar\omega$ ，即 Rayleigh-Jeans 分布。这时，能谱密度就进一步写为：

$$u(\omega)d\omega = k_B T g(\omega)d\omega. \quad (12.14)$$

这正是经典统计物理给出的结果。每一个电磁模式贡献平均能量 $k_B T$ ，能谱密度随着 ω 的增加以 ω^2 的形式发散。这就是所谓的紫外灾难。我们上面基于能级量子化得出的结果可以很好的解决紫外发散的问题。其原因就是，在高频区域，光子模式满足 Planck 分布，它随着 ω 以指数形式衰减。

- 我们这里得出光子气体统计性质的出发点是假设每个模式光子的能量是量子化的。这也正是 Planck 为了解决经典统计物理导致的黑体辐射的紫外灾难而引入的假设，也是量子力学的开端。因此，我们可以说，量子力学的诞生原因之一是人们对黑体辐射能谱密度的研究。

12.1.4.1 宇宙微波背景辐射

1965 年，美国 Bell 实验室的 A. Penzias 和 R. W. Wilson 在用他们设计的大型天线进行射电天文学研究，但他们收到的信号中存在一个无法消除的背景噪声，使得实验无法继续进行。那个噪声的波长为 7.35 厘米，相当于大约 3 K 的黑体辐射谱，其各向同性的程度极高，而且与季节变化无关（见下图）。他们随后花了一年时间来想办法去除这个背景噪声。他们尝试了各种办法来追踪这个背景噪声的来源，比如附近城市的无线电信号影响。他们甚至连天线里的鸽子窝都给清除了，但那个噪声还是无法消除。在万般无奈的情况下，他们打电话给普林斯顿大学的 Robert Dicke，希望他能帮忙作出一种解释。Dicke 马上意识到他们想要除去的东西正是自己研究组正在设法寻找的东西——宇宙大爆炸残留下来的某种宇宙背景辐射。Penzias 和 Wilson 因此获得了 1978 年诺贝尔物理学奖³。

³他们分享了 1/2 的奖金，另外一半被 P. L. Kapitsa 获得 “for his basic inventions and discoveries in the area of low-temperature physics.”

练习 12.3 光子气体的绝热膨胀：考虑处于体积为 V 的腔体内温度为 T 的光子气体。在绝热条件下，光子气体发生膨胀，导致腔体的体积增加。在等熵条件下 $TV^{1/3}$ 保持不变。(1) 假设宇宙背景辐射在温度为 3000K 时与其他物质解耦。比较当时和现在的宇宙大小。如果宇宙的线度随时间线性增加，以宇宙目前寿命为单位，这个解耦发生在什么时间？(2) 证明在这个过程中，光子气体膨胀做功为

$$W = (\pi^2/15\hbar^3 c^3) V_i k_B^4 T^3 (T_i - T_f).$$

这里的 i 和 f 分别代表初态和末态。

12.2 声子气体

12.2.1 系统

晶体中的原子按照一定的周期在实空间排列。在热平衡的情况下，这些原子并不是完全固定不动的。由于热涨落或零点能的影响，它们在平衡位置附近做振动。通常振动的振幅相对于原子之间的距离来说非常小，因此我们可以把它们的振动近似为简谐振动，作为处理晶格中原子动力学的第一步。这相当于对原子之间的相互作用势在平衡位置附近做 Taylor 展开，然后取至二阶：

$$V(\{r_i\}) = V(\{r_i^{(0)}\}) + \frac{\partial V}{\partial r_i} \Big|_0 u_i + \frac{1}{2} \frac{\partial^2 V}{\partial r_i \partial r_j} \Big|_0 u_i u_j + \dots$$

这里的 r_i 和 $u_i = r_i - r_i^{(0)}$ 分别为系统第 i 个自由度的位置和位移。这个展开式的第一项为所有原子都处于平衡位置时的势能，它是一个常数，我们选为我们的能量零点。式子的第二项为线性项。由于原子处于平衡位置对应势能的极小值，所以这个位置的一阶微分为零。所以，我们只剩下了二阶项。这就是我们所说的简谐近似。在简谐近似下，原子之间的相对运动还是相互耦合在一起的。因为一个原子的位移会影响它周围一个范围内原子的受力。对应的运动方程可以表示为：

$$m_i \frac{\partial^2 u_i}{\partial t^2} = -K_{ij} u_j.$$

这里的力常数矩阵 $K_{ij} = \frac{\partial^2 V}{\partial r_i \partial r_j} \Big|_0$ 。我们可以利用原子排布的周期性和 Bloch 定理，将运动方程转化到 k 空间，然后对角化每个给定波矢 k 下的运动方程，从而得到对应的声子简正运动模

式。而波矢 k 的取值在周期性边界条件下依然是：

$$k_{\alpha,n} = \frac{2\pi}{L_\alpha} n, \quad \alpha = x, y, z.$$

简正坐标下，原子的振动可以表示为相互独立的简谐振子的运动。每个简谐模式有一个确定的角频率，对应一种形式的晶格集体振动，称为格波。对格波进行量子化后称为声子。这样，我们就可以把固体中原子的集体振动行为用无相互作用的声子气体来描述。其定量描述是《固体物理》课的内容。这里，我们只分析声子气体在低温和高温极限下的情形。特别是，分析其热容的温度依赖关系。

12.2.2 Dulong-Petit 定律

早期对固体热力学性质的研究主要考虑的是温度比较高的范围，声子的热容总结为 Dulong-Petit 定律

$$C = 3Nk_B$$

其中， N 为系统中的原子数。由于固体的热膨胀系数很小，其定容热容和定压热容的差别很小。我们这里不做区分。这个结果我们前面考虑经典能量均分定理时已经得出。因为每个原子可以在三个方向运动，在简谐近似下，每个方向贡献 $k_B T$ 的能量。所以系统的总内能为 $U = 3Nk_B T$ 。经典 Dulong-Petit 定律的主要特点是系统热容不依赖于温度。当然，它不满足热力学第三定律，在低温极限下不成立。实验测量发现，(三维) 固体的热容在低温极限下正比于 T^3 。为了解释低温的实验数据，Einstein 和 Debye 分别提出自己的模型，下面我们分别介绍。

12.2.3 爱因斯坦模型

为了解释固体材料低温下的热力学性质，爱因斯坦假设固体中每个原子在平衡位置附近的振动具有相同的角频率 ω_E 。这样，声子气体的热容就可以由单频的谐振子热力学性质得到

$$\ln Z = -3N \left(\frac{1}{2}\beta\hbar\omega_E + \ln(1 - \exp(-\beta\hbar\omega_0)) \right)$$

从此配分函数出发，我们可以得到其热力学性质。因为我们前面已经介绍过单个谐振子的热力学性质。爱因斯坦模型的结果和单个谐振子模型类似，只是相差一个因子 $3N$ 。所以我们这

里直接给出结果。内能可以表示为：

$$U = 3N\hbar\omega_E \left(\frac{1}{2} + \frac{1}{\exp(\beta\hbar\omega_E) - 1} \right)$$

对温度微分可以得出其热容

$$C = 3Nk_B \frac{x_E^2 e^{x_E}}{(e^{x_E} - 1)^2}, \quad x = \beta\hbar\omega, x_E = \beta\hbar\omega_E.$$

热容对温度的依赖关系和单个谐振子完全相同。

- 低温：低温极限下 β 和 x_E 都趋于 ∞ ，所以 $C \sim x_E^2 e^{-x_E}$ 。热容随温度指数衰减。这是具有能隙的系统具有的共同特征。
- 高温：高温极限下 β 和 x_E 都趋于 0，所以 $C \rightarrow 3Nk_B$ ，与经典的 Dulong-Petit 定律结果一致。

我们可以看出，Einstein 模型保证了声子气体的热容在低温极限下满足热力学第三定律。但与实验测量的 T^3 温度依赖关系还有差别。

12.2.4 Debye 模型

So I decided to calculate the spectral distribution of the possible free vibrations for a continuous solid and to consider this distribution as a good enough approximation to the actual distribution. The sonic spectrum of a lattice must, of course, deviate from this as soon as the wavelength becomes comparable to the distances of the atoms. ... The only thing which had to be done was to adjust to the fact that every solid of finite dimensions contains a finite number of atoms and therefore has a finite number of free vibrations. ... At low enough temperatures, and in perfect analogy to the radiation law of Stefan-Boltzmann ..., the vibrational energy content of a solid will be proportional to T^4 . — P. Debye

Einstein 模型虽然可以保证固体热容在零温极限下趋于零，但它无法成功解释固体热容在低温极限下的温度依赖关系。Debye 模型则可以很好的弥补这一缺陷。Debye 模型假设声子具有与光子气体类似的线性色散关系，并且具有三重简并度，对应一个纵向两个横向声学模式。相应的声子群速度为 v_s 对应固体中的声速。通过固体物理的学习我们知道实际材料的声子色散关系往往很复杂。Debye 模型的近似抓住了低温极限下声子的特性，因此可以在不考虑材料具体细节的前提下，给出低温极限下正确的热容温度依赖关系。因为我们知道，三维固体

材料在长波极限下都具有三个模式的声学声子，且都具有线性的色散关系，尽管它们的群速度可能不同。

声子气体和光子气体的另外一个重要区别是声子的角频率有上限，这个上限在 Debye 模型下由 Debye 频率 ω_D 给出。存在 Debye 频率的物理原因是固体中的声子模式数目是由系统中的原子个数决定的。参考光子气体的结果，我们可以得出 ω_D 需要满足下面的关系

$$\int_0^{\omega_D} d\omega g(\omega) = 3N.$$

这里声子的态密度与光子几乎相同（简并度为 3 而不是光子的 2）

$$g(\omega)d\omega = \frac{3V}{2\pi^2 v_s^3} \omega^2 d\omega$$

于是我们得到⁴

$$\omega_D = k_D v_s,$$

对应的 Debye 波矢为

$$k_D = (6\pi^2)^{1/3} \left(\frac{N}{V} \right)^{1/3}$$

Debye 波矢给出了固体系统中声子的最短波长，对应为原子间距的大小 $k_D \sim (N/V)^{1/3}$ 。

确定了声子的色散关系后，我们就可以使用和光子气体相同的方法来得到声子气体的热力学性质。这里我们不再重复，直接给出结果。首先，

$$\ln Z = - \int_0^{\omega_D} g(\omega) \left(\frac{1}{2} \beta \hbar \omega + \ln(1 - e^{-\beta \hbar \omega}) \right) \quad (12.15)$$

$$= - \frac{9N}{\omega_D^3 (\beta \hbar)^3} \int_0^{x_D} dx x^3 \left(\frac{1}{2} x + \ln(1 - e^{-x}) \right) \quad (12.16)$$

这里 $x_D = \beta \hbar \omega_D$ 。这个结果和光子气体的唯一区别在于积分的上限是 x_D 而非 $+\infty$ 。但在低温极限下，我们可以近似 $x_D \rightarrow \infty$ 。声子和光子气体没有区别。据此，我们知道声子气体在低温极限下的热容必然与光子气体相同，满足 $C \sim T^3$ 。热容的具体表达式为：

$$C = \frac{9Nk_B^4 T^3}{\omega_D^3 \hbar^3} \int_0^{x_D} dx \frac{x^4 e^x}{(e^x - 1)^2}$$

- $T \rightarrow 0$ 时，上式的积分上限可以近似为 ∞ ，所以我们有 $C \sim T^3$ 。

- $T \rightarrow \infty$ 时， $x_D \rightarrow 0$ 。我们可以对 e^x 关于 x 做展开，得到 $C \rightarrow 3Nk_B$ 。

可见，Debye 模型可以很好的解释声子气体在高温和低温极限下的热力学性质。

⁴如果 V 是整个固体系统的体积，那么 N 就是系统中的所有原子数。如果 V 是原胞体积，那么 N 就是原胞中的原子数。

第 13 章 量子气体



My own undergraduate experience in thermodynamics and statistical mechanics, a half-century ago at MIT, consisted of a single semester of Sears' Thermodynamics (skillfully taught by the man himself.) But it was a subject that seemed as distant from “real” physics as did poetry or French literature. Graduate study at the University of Illinois in Urbana-Champaign was not that different, except that the course in statistical mechanics was taught by the brilliant lecturer Francis Low the year before he departed for... MIT. I asked my classmate J.R. Schrieffer, who presciently had enrolled in that class, whether I should chance it later with a different instructor. He said not to bother — that he could explain all I needed to know about this topic over lunch. On a paper napkin, Bob wrote $e^{-\beta H}$. “That’s it in a nutshell!” “Surely you must be kidding, Mr Schrieffer,” I replied (or words to that effect.) “How could you get the Fermi-Dirac distribution out of THAT?” “Easy as pie,” was the reply... and I was hooked.

- D. C. Mattis in preface of his book “Statistical mechanics made simple”

13.1 回顾：经典理想气体

我们介绍统计物理方法的第一个应用就是经典理想气体。我们主要按照下面的顺序进行了分析，并得出了其热力学状态方程等性质：

- 计算得出单粒子态的分布及态密度
- 考虑一个粒子的配分函数 Z_1
- 由于不同粒子间的相互作用可以忽略， N 粒子配分可以从单粒子结果得到 $Z_N = Z_1^N$
- 进一步考虑粒子的不可区分性，引入因子 $1/N!$
- 从配分函数出发得出其热力学性质

这里需要说明的是，虽然我们是从单粒子的量子态出发来得出其配分函数的，而且我们得到

的配分函数中包含 \hbar , 但最后我们得出的热力学性质都是经典的结果。也就是说, 即使考虑经典的理想气体, 我们的统计物理理论也要包含一定的量子性。通过我们前面的学习, 我们知道这部分量子性是指每个单粒子在相空间所占的体积元是量子化的, 要满足量子不确定性原理的限制 $\Delta x \Delta p \sim \hbar$ 。此外, 为了使我们得到的自由能、熵具有广延性, 我们不得不引入 $1/N!$ 因子来考虑粒子的不可区分性。这些相对“别扭”的过程告诉我们应该有更好的方法来处理这个问题。这是我们这一讲的主要内容。我们将转换一个视角, 从开始就考虑粒子的不可区分性, 同时考虑不同种类粒子之间满足的交换对称性。我们会发现不同交换对称性会导致不同的统计分布。

13.2 自旋统计定理

对于量子气体系统, 我们首先需要考虑的是气体粒子的交换对称性。根据粒子的不可区分性, 如果我们任意交换系统中的两个粒子的话, 系统的状态不会发生变化。也就是说, 如果系统的状态可以用某个波函数 $\psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$ 来描述的话, 交换其中任意两个粒子的坐标后得到的波函数仍然可以用来描述系统的状态, 即交换后波函数不变, 它们之间最多存在一个相位差而已

$$\psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = e^{i\phi} \psi(\vec{r}_2, \vec{r}_1, \dots, \vec{r}_N)$$

对于玻色子, $\phi = 2n\pi$, 波函数完全不变; 而对于费米子, $\phi = (2n+1)\pi$, 波函数变号¹。自旋统计定理告诉我们, 自旋为整数的粒子为玻色子, 而自旋为半整数的粒子为费米子。费米子的交换反对称性决定了一个单粒子态最多有一个粒子占据。这就是 Pauli 不相容原理。我们下面就分别考虑玻色子和费米子统计性质的区别。

13.3 占据数表象

在考虑不可区分的多粒子系统时, 采用占据数表象会比较方便²。这就要求我们稍微转换一下视角。原来我们是分别考虑每一个粒子, 比如写出每个粒子的配分函数, 然后用它们来构建多粒子系统的配分函数。最后, 我们需要再回过头来考虑粒子的不可区分性。现在我们

¹除了玻色子或费米子外, 还可以存在任意子, 其相位 ϕ 可以去任意值。但任意子只存在于低维空间。我们这里不做考虑。

²这是量子力学的内容, 附录22给出了具体的介绍。

可以用另一个角度来做分析。既然粒子的无法区分的，我们干脆不去考虑这个粒子再去考虑那个粒子，我们只去计数(因为我们要计数，自然要使用巨正则系综来考虑)。比如，知道了系统允许的单粒子态后，我们只关心每个态上的粒子数 n_i ，而不去关心到底是这个粒子还是那个粒子。要指定一个单粒子态我们可能需要给出描述这个状态的所有量子数，比如波矢、轨道、自旋等。

我们可以举一个简单的粒子来说明这一点。我们在上课的时候可能会点名。点名的过程是跟踪个体，因为每个人是可以区分的，而且我需要知道每个人来了还是没来，他或她坐在什么位置。但假如我不关心谁来了谁没来，我关心的只是来了多少人。我就可以直接按照座位顺序去数，看每个座位有几个人(当然只有两个可能 0 或 1)。这个角度就是从座位的占据数表象来看了。我只需要知道每个座位的人数，而不去区分是谁坐在那里。

在占据数表象下，我要做的是对每个单粒子态进行统计(我们这里忽略粒子间的相互作用)，得出对应配分函数；然后把所有的单粒子态综合起来，得到整个系统的配分函数。从单粒子态到多粒子态的过程与我们前面从单个粒子到多个粒子的过程相同，因为我们仍然假设不同态之间相互独立。即，总的配分函数为每个单粒子态配分函数的乘积

$$Z_N = \prod_k Z_k.$$

下面我们从这个角度来分别考虑费米子和玻色子。

13.4 费米子

13.4.1 单个状态

我们考虑给定一个单粒子态，其能量为 ε_1 。在这个状态上的粒子数只有两种情况：0 和 1，所以其配分函数非常简单

$$Z_1 = 1 + e^{-\beta(\varepsilon_1 - \mu)}.$$

有了配分函数，我们可以计算这个态上的平均占据数

$$\langle n_1 \rangle = \frac{e^{-\beta(\varepsilon_1 - \mu)}}{Z_1} = \frac{1}{e^{\beta(\varepsilon_1 - \mu)} + 1} \equiv f_{FD}(\varepsilon_1).$$

我们用 $f_{FD}(\varepsilon_1)$ 来表示这个占据数，称为 Fermi-Dirac 分布，见图13.1。

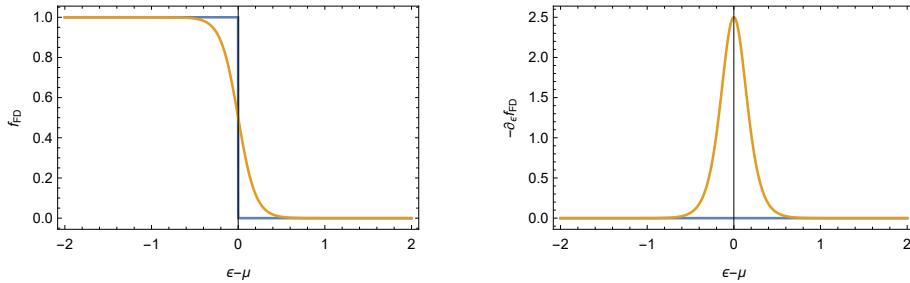


图 13.1: 左边: Fermi-Dirac 分布 $f_{FD}(\epsilon)$; 右边: $-\partial_\epsilon f_{FD}(\epsilon)$ 。蓝色: $k_B T \rightarrow 0$; 橙色: $k_B T = 0.1$ 。

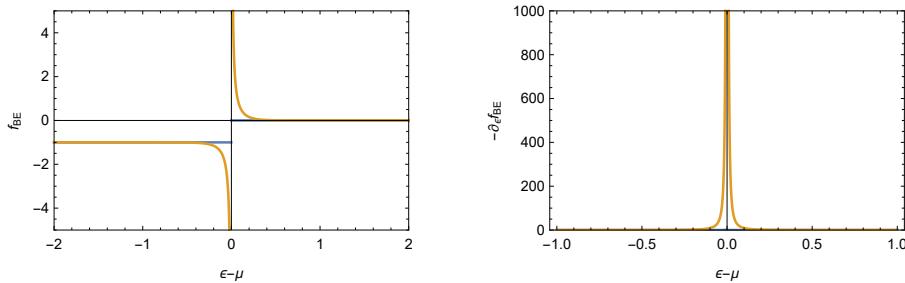


图 13.2: 左边: Bose-Einstein 分布 $f_{BE}(\epsilon)$; 右边: $-\partial_\epsilon f_{BE}(\epsilon)$ 。蓝色: $k_B T \rightarrow 0$; 橙色: $k_B T = 0.1$ 。

13.4.2 多个状态

对于多个状态的情况我们可以类似考虑。由于各个状态之间相互独立，系统总的平均占据数由各个单粒子态的平局占据数之和给出

$$\langle n \rangle = \sum_k \langle n_k \rangle.$$

13.5 玻色子

13.5.1 单个状态

对于玻色子，由于每个状态可以有任意个粒子，配分函数为

$$Z_1 = (1 - e^{-\beta(\epsilon_1 - \mu)})^{-1}.$$

我们可以用与费米子类似的方法计算平均占据数。这里我们介绍一个简单的方法。为此，我们引入逸度 λ 的定义

$$\lambda = e^{\beta\mu}.$$

这样配分函数可以重写为：

$$Z_1 = (1 - \lambda e^{-\beta \varepsilon_1})^{-1}.$$

而平均占据数可以表示为：

$$\langle n_1 \rangle = \lambda \frac{\partial \ln Z_1}{\partial \lambda}.$$

据此我们得到

$$\langle n_1 \rangle = \frac{1}{e^{\beta(\varepsilon_1 - \mu)} - 1} = f_{BE}(\varepsilon_1).$$

上式为玻色子所满足的 Bose-Einstein 分布，见图13.2。

13.5.2 多个状态

如果系统包含多个状态，那么整个系统的配分函数就是每个单粒子态配分函数的乘积。相应的平均占据数就是每个状态平均占据数之和

$$\langle n \rangle = \sum_i \langle n_i \rangle = \sum_i \frac{1}{e^{\beta(\varepsilon_i - \mu)} - 1}.$$

13.6 经典极限

通过上面的分析我们看出，粒子所满足的交换对称性不同，相应它们的统计性质，如每个状态的平均占据数也不同。费米子和玻色子分别对应 Fermi-Dirac 和 Bose-Einstein 分布。我们可以把它们统一写成下面的形式：

$$f = \frac{1}{e^{\beta(\varepsilon - \mu)} \pm 1},$$

其中 + 号对应 Fermi-Dirac，"-" 对应 Bose-Einstein 分布。我们可以看出，如果忽略上式中的 ± 1 ，那么玻色子和费米子的区别也就不复存在了，平均占据数可以写成 Boltzmann 分布的形式

$$f = e^{-\beta(\varepsilon - \mu)}. \quad (13.1)$$

这对应经典极限的情况。上面的经典近似要成立，需要满足下面的条件：对于系统允许的任意状态对应的能量 ε ，必须满足

$$e^{\beta(\varepsilon - \mu)} \gg 1.$$

如果把系统的基态能量定为 0 的话，这个条件就变为：

$$\lambda \ll 1. \quad (13.2)$$

我们前面已经考虑了经典理想气体的情况，并得出一系列结果。这里，我们推测从 Bose-Einstein 或者 Fermi-Dirac 分布出发取经典极限后得出的结果应该和前面经典理想气体的结果相符。为了检查这一推测，我们需要重新来回顾一下理想气体中的一些结论。

首先，在经典理想气体中，某个粒子处于状态 k 的概率为

$$P_k = \frac{1}{Z_1} e^{-\beta \varepsilon_k} = \frac{1}{V n_Q} e^{-\beta \varepsilon_k}.$$

由于粒子之间相互独立，所以给定能级上的平均占据数为

$$\langle n_k \rangle = N P_k = \frac{n}{n_Q} e^{-\beta \varepsilon_k}.$$

又因为理想气体的化学势可以表示为

$$\mu = -\beta^{-1} \ln(n_Q/n).$$

上式正好可以写成 Boltzmann 分布的形式

$$\langle n_k \rangle = e^{-\beta(\varepsilon_k - \mu)}.$$

这就是我们得到的 Boltzmann 分布。所以的确理想玻色和费米气体在经典极限下给出了经典理想气体的结果。此外，我们得出的经典理想气体的结果的适用条件是

$$n \ll n_Q,$$

如果用逸度来表示的话，恰恰就是式 (13.2). 可见，理想玻色和费米气体取经典极限的话，的确可以得到我们前面经典理想气体的相应结果。

13.6.1 Boltzmann 因子和 Boltzmann 分布

这里有一点需要说明一下。我们前面考虑巨正则系综时得出一个结论：系统处于某一个状态的概率

$$P(N, k_N) \propto e^{-\beta(\varepsilon_{k_N} - N\mu)},$$

这里， N 为系统中包含的粒子数， ε_{k_N} 为给定 N 情况下的一个微观态的能量，这个因子为 Gibbs-Boltzmann 因子。把所有状态求和我们就得到了系统的配分函数。需要说明的是，Gibbs-Boltzmann 因子既适用于经典系统，又适用于量子系统。而且，我们得到的只是一个正比关系。

相反，我们上面得到的 Boltzmann 分布（式13.1）是给定系统的一个单粒子态后，该单粒子态上的平均占据数，是一个绝对值。两者的形式虽然看起来非常类似，但物理意义是不一样的。我们需要分清楚。

练习 13.1

证明给定能级 ε 上的占据数相对涨落可以表示为

$$\frac{\langle \Delta n_\varepsilon^2 \rangle}{\langle n_\varepsilon \rangle^2} = \lambda^{-1} e^{\beta\varepsilon} = \frac{1}{\langle n_\varepsilon \rangle} \mp 1.$$

其中，负号对应 Fermi 粒子，正号对应 Bose 粒子。如何从物理上理解 Fermi, Bose 粒子数涨落的区别。

证明对于 Bose 粒子，能级 ε 被 n 个粒子占据的概率为：

$$P(n) = \frac{\langle n_\varepsilon \rangle^n}{(\langle n_\varepsilon \rangle + 1)^{n+1}}$$

而对于 Fermi 子，对应概率为：

$$P(n) = \begin{cases} 1 - \langle n_\varepsilon \rangle & n = 0 \\ \langle n_\varepsilon \rangle & n = 1 \end{cases}$$

对于经典的 Maxwell-Boltzmann 粒子

$$P(n) = \frac{(\lambda e^{-\beta\varepsilon})^n}{e^{\lambda e^{-\beta\varepsilon}}}$$

可见，对于 Bose 粒子， $P(n+1)/P(n) = \langle n_\varepsilon \rangle / (\langle n_\varepsilon \rangle + 1)$ ，与 n 无关。这说明，对于 Bose 粒子，再往能级上增加一个粒子的概率与目前已经占据的粒子数 n 无关。这时，我们通常说玻色子之间有聚束 (bunching) 效应。但费米子却有反聚束效应 (anti-bunching)。对于经典粒子来说，在能级上增加一个粒子的概率与前面能级是否有占据，具体占据情况如何是没有关系的。

13.7 Fermi 和 Bose 气体

根据上一节的介绍，我们现在可以考虑量子气体的统计和热力学性质。我们所采用的思路和原来一样。主要区别是：我们从原来的跟踪粒子转换为跟踪单粒子状态，也就是转到占据数表象来考虑问题。

首先，我们仍然要对所有的单粒子态求和，所以需要用到态密度的表达式把求和转化为积分：

$$\sum_k(..) = \int_0^{+\infty} dk g(k)(..) = \int_0^{+\infty} d\varepsilon g(\varepsilon)(..). \quad (13.3)$$

如果考虑具有抛物线色散关系的三维气体的话，我们可以得到态密度的表达式：

$$g(k)dk = g_s \frac{V}{2\pi^2} k^2 dk, \quad (13.4)$$

以及

$$g(\varepsilon)d\varepsilon = \frac{g_s V m^{3/2}}{\sqrt{2\pi^2 \hbar^3}} \sqrt{\varepsilon} d\varepsilon = \Delta \sqrt{\varepsilon} d\varepsilon. \quad (13.5)$$

上式中， $g_s = (2S+1)$ 表示每个粒子的自旋简并度。我们的出发点是每个单粒子态的配分函数

$$Z_k = (1 \pm \lambda e^{-\beta \varepsilon_k})^{\pm 1}. \quad (13.6)$$

整个系统的配分函数可以写成

$$\ln Z = \sum_k \ln Z_k \quad (13.7)$$

$$= \pm \int_0^{+\infty} d\varepsilon g(\varepsilon) \ln(1 \pm \lambda e^{-\beta \varepsilon}) \quad (13.8)$$

$$= \pm \Delta \int_0^{+\infty} d\varepsilon \sqrt{\varepsilon} \ln(1 \pm \lambda e^{-\beta \varepsilon}) \quad (13.9)$$

$$= \frac{2}{3} \beta \Delta \int_0^{+\infty} d\varepsilon \frac{\varepsilon^{3/2}}{\lambda^{-1} e^{\beta \varepsilon} \pm 1}. \quad (13.10)$$

下面我们需要使用一个常用积分³

$$\int_0^{+\infty} d\varepsilon \frac{\varepsilon^{n-1}}{\lambda^{-1} e^{\beta \varepsilon} \pm 1} = \beta^{-n} \Gamma(n) [\mp \text{Li}_n(\mp \lambda)]. \quad (13.11)$$

其中， $\text{Li}_n(\lambda)$ 为多对数函数，而 $\Gamma(x)$ 为 Euler Gamma 函数。利用这个积分我们就可以得到

$$\ln Z = g_s V n_Q [\mp \text{Li}_{5/2}(\mp \lambda)]. \quad (13.12)$$

同时，我们得到系统的巨势

$$\Phi_G = -k_B T \ln Z = -g_s k_B T V n_Q [\mp \text{Li}_{5/2}(\mp \lambda)]. \quad (13.13)$$

这里 n_Q 是量子浓度。利用相同的方法，我们可以得到系统的平均粒子数

$$N = \lambda \left(\frac{\partial \ln Z}{\partial \lambda} \right)_{T,V} = \mp g_s V n_Q \text{Li}_{3/2}(\mp \lambda) \quad (13.14)$$

$$U = - \left(\frac{\partial \ln Z}{\partial \beta} \right)_{\lambda,V} = \frac{3}{2} k_B T N \frac{\text{Li}_{5/2}(\mp \lambda)}{\text{Li}_{3/2}(\mp \lambda)} \quad (13.15)$$

从上面的结果我们不难验证

$$\Phi_G = -\frac{2}{3} U. \quad (13.16)$$

³详细讨论见附录24.6。

这个结果对玻色和费米气体同时适用。我们可以进一步把上面的结果写成简化的形式：

$$U = \frac{3}{2} N k_B T \frac{\text{Li}_{5/2}(\mp\lambda)}{\text{Li}_{3/2}(\mp\lambda)} \quad (13.17)$$

$$\Phi_G = -N k_B T \frac{\text{Li}_{5/2}(\mp\lambda)}{\text{Li}_{3/2}(\mp\lambda)} \quad (13.18)$$

$$P = n k_B T \frac{\text{Li}_{5/2}(\mp\lambda)}{\text{Li}_{3/2}(\mp\lambda)} \quad (13.19)$$

- 在经典极限下 $\lambda \ll 1$, 我们玻色子和费米子的区别消失, 我们可以得到经典理想气体的熟悉结果

$$U = \frac{3}{2} N k_B T, \quad (13.20)$$

$$\Phi_G = -N k_B T, \quad (13.21)$$

$$P = -\Phi_G/V = n k_B T, \quad (13.22)$$

$$N = g_s V n_Q \lambda \rightarrow \lambda = \frac{n}{g_s n_Q}. \quad (13.23)$$

可见, $\lambda \ll 1$ 对应 $n \ll g_s n_Q$ 。

- 在量子极限下 $\lambda \gg 1$, 玻色和费米气体的行为会有很大的不同。我们下面会详细论述。

13.8 应用举例：Fermi 气体

13.8.1 零温极限

我们上面给出了 Fermi 和 Bose 气体的一些普适结果。下面我们来看一下, 考虑粒子的量子属性后, 其性质有哪些变化。我们首先考虑零温极限。对于费米气体, 由于 Pauli 不相容原理, 每个单粒子态只能有一个粒子占据; 一个状态被占据后, 其他粒子只能去占据能量更高的状态; 最终所有粒子找到各自的状态。最高占据态的能量对应系统的化学势。化学势以上的所有状态都是空的, 而其以下的所有状态都被占满。所以, Fermi 气体的化学势和系统中的粒子数直接相关, 理论上可以取从基态到最高能量态之间的任意值。这就像往容器中倒水一样, 先倒进去的水占据容器的底部, 其他水只能占据较高的位置, 导致水面上升。由于在零温下, Fermi-Dirac 分布变成一个台阶函数, 所以零温下 Fermi 气体的性质可以通过非常简单的计算得到。

13.8.1.1 费米能级

首先，我们可以在给定系统体积 V 和粒子数 N 的情况下计算零温的化学势，又被称为 Fermi 能级 $E_F = \mu(T = 0)$ ，由下式给出

$$\int_0^{E_F} d\varepsilon g(\varepsilon) = N. \quad (13.24)$$

我们得到

$$E_F \equiv \frac{\hbar^2 k_F^2}{2m} = \frac{\hbar^2}{2m} \left(\frac{6\pi^2 n}{g_s} \right)^{\frac{2}{3}}. \quad (13.25)$$

我们还定义了费米波矢 k_F 。最为常见的 Fermi 气体就是金属中的自由电子气。由于金属中的电子密度非常高，其费米能级也很高，一般在几个 eV。这个能量和室温对应的能量 $k_B T \sim 25$ meV 相比要高很多，所以即使在室温下，金属中的自由电子气的量子属性也非常重要。对于电子来说仍然处于低温极限，或者成为简并极限。

13.8.1.2 内能

零温下系统的内能

$$U = \int_0^{E_F} d\varepsilon \varepsilon g(\varepsilon) = \frac{3}{5} N E_F. \quad (13.26)$$

13.8.1.3 简并压

费米气体与经典理想气体以及玻色气体的一个重要区别是：即使在零温下，它对外的压强也不为零，称为费米气体的简并压。产生简并压的原因是 Pauli 不相容原理导致每个状态最多只有一个粒子占据，由于系统中存在多个电子，必然有电子占据其它动能非零的状态，这些电子的运动会产生压强。这是费米气体和玻色气体的重要区别。我们可以计算气体的巨势，进而得到其压强。根据上节的结果我们有 $\Phi_G = -\frac{2}{3}U$ ，于是

$$P = -\frac{\Phi_G}{V} = \frac{2}{3} u. \quad (13.27)$$

13.8.1.4 体弹性模量

费米气体的体弹性模量为

$$B \equiv -V \frac{\partial P}{\partial V} = -\frac{2}{3} \frac{\partial U}{\partial V} + \frac{2}{3} \frac{U}{V} \frac{10U}{9V} = \frac{2}{3} n E_F. \quad (13.28)$$

练习 13.2 用 E_F 来表示 (1) 系统的总粒子数；(2) 粒子的平均能量。

13.8.2 有限温度：热容、Sommerfeld 展开

如果要考虑 Fermi 气体热容的话，我们零温极限下的分析就不适用了。一个常用的方法是利用 Sommerfeld 展开来考虑有限温度对零温结果的修正

$$\begin{aligned} \int_0^{+\infty} d\varepsilon \phi(\varepsilon) f_{FD}(\varepsilon) &= \sum_{n=0, \text{neven}}^{+\infty} 2 \left(\frac{d^n \Psi}{dx^n} \right)_{x=0} (1 - 2^{1-n}) \zeta(n) \\ &= \int_{-\infty}^{\mu} \phi(\varepsilon) d\varepsilon + \frac{\pi^2}{6} (k_B T)^2 \left(\frac{d\phi}{d\varepsilon} \right)_{\varepsilon=\mu} + \frac{7\pi^4}{360} \left(\frac{d^3 \psi}{d\varepsilon^3} \right)_{\varepsilon=\mu} + \dots \quad (13.29) \end{aligned}$$

其中， $\psi(x) = \int_0^x \phi(x') dx'$ 。从上面的结果我们不难得出，费米气体的热容在低温下随温度呈线性变化 $C_X \propto T$ 。

Sommerfeld 展开 (13.29) 的证明如下：引入函数 $\psi(\varepsilon) = \int_0^\varepsilon d\varepsilon' \phi(\varepsilon')$ ，使得 $\phi(\varepsilon) = d\psi/d\varepsilon$ 。

于是，

$$\int_0^{+\infty} d\varepsilon \phi(\varepsilon) f_{FD}(\varepsilon) = \psi(\varepsilon) f_{FD}(\varepsilon)|_0^\infty - \int_0^\infty d\varepsilon \psi(\varepsilon) \frac{df_{FD}(\varepsilon)}{d\varepsilon}$$

这里，我们如果 $\psi(\varepsilon = 0) = 0$ ，我们就可以忽略边界项。另 $x = (\varepsilon - \mu)\beta$ ，利用

$$\frac{df_{FD}}{d\varepsilon} = -\frac{1}{k_B T} \frac{e^x}{(e^x + 1)^2},$$

然后把 $\psi(\varepsilon)$ 展开

$$\psi(\varepsilon) = \sum_{s=0}^{\infty} \frac{x^s}{n!} \left(\frac{d^s \psi}{dx^s} \right)_{x=0},$$

我们得到

$$= \sum_{s=0}^{\infty} \frac{1}{s!} \left(\frac{d^s \psi}{dx^s} \right)_{x=0} \int_{-\mu\beta}^{\infty} \frac{x^s e^x dx}{(e^x + 1)^2}$$

如果 $k_B T \ll \mu$ ，我们可以进一步把上式积分的下限改为 $-\infty$ 。这个条件在考虑金属中的电子时是都会满足的。这样的话，所有的 s 为奇数的积分都为零。我们只需要考虑 s 为偶数

的情况。于是，我们可以进一步得到：

$$\begin{aligned}
 \int_{-\infty}^{+\infty} \frac{x^s e^x dx}{(e^x + 1)^2} &= 2 \int_0^{+\infty} \frac{x^s e^x dx}{(e^x + 1)^2} \\
 &= 2 \int_0^{+\infty} dx \sum_{n=0}^{+\infty} e^x x^n [(n+1)(-1)^{n+1} e^{-nx}] \\
 &= 2 \sum_{n=1}^{\infty} (-1)^{n+1} n \int_0^{\infty} x^s e^{-nx} dx \\
 &= 2(s!) \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^s} \\
 &= 2(s!)(1 - 2^{-1s})\zeta(s),
 \end{aligned}$$

这里 $\zeta(s)$ 为 Riemann zeta 函数。代入到我们需要计算的积分，

$$\begin{aligned}
 &= \sum_{s=0, even}^{\infty} 2 \left(\frac{d^s \psi}{dx^s} \right)_{x=0} (1 - 2^{1-s})\zeta(s) \\
 &= \psi + \frac{\pi^2}{6} \left(\frac{d^2 \psi}{dx^2} \right)_{x=0} + \frac{7\pi^4}{360} \left(\frac{d^4 \psi}{dx^4} \right)_{x=0} + \dots \\
 &= \int_{-\infty}^{\mu} \phi(\varepsilon) d\varepsilon + \frac{\pi^2}{6} \left(\frac{d\phi}{d\varepsilon} \right)_{\varepsilon=\mu} + \frac{7\pi^4}{360} \left(\frac{d^3 \phi}{d\varepsilon^3} \right)_{\varepsilon=\mu} + \dots
 \end{aligned}$$

这就是 Sommerfeld 展开。

下面我们做一个稍微定量的分析。首先，考虑粒子数

$$\begin{aligned}
 N &= \Delta \int_0^{+\infty} d\varepsilon \varepsilon^{1/2} f_{FD}(\varepsilon) \\
 &= \frac{2}{3} \Delta \mu^{3/2}(T) \left(1 + \frac{\pi^2}{8} (k_B T / \mu)^2 + \dots \right).
 \end{aligned} \tag{13.30}$$

根据零温下的结果

$$N = \frac{2}{3} \Delta E_F^{3/2},$$

我们可以得到化学势的温度依赖：

$$\mu(T) = E_F \left[1 - \frac{\pi^2}{12} \left(\frac{k_B T}{E_F} \right)^2 + \dots \right] \tag{13.31}$$

可见，对于三维抛物线型色散的非相对论费米气体，其化学势随着温度的升高而降低。

练习 13.3 分析二维和一维情况下，具有非相对论抛物线色散关系的粒子化学势如何依赖于温度。

同理，我们可以计算内能：

$$\begin{aligned} U &= \Delta \int_0^{+\infty} d\varepsilon \varepsilon^{3/2} f_{FD}(\varepsilon) \\ &= \frac{2}{5} \Delta \mu^{5/2}(T) \left(1 + \frac{5\pi^2}{8} (k_B T / \mu)^2 + \dots \right) \\ &= \frac{3}{5} N E_F \left[1 + \frac{5\pi^2}{12} \left(\frac{k_B T}{E_F} \right)^2 + \dots \right] \end{aligned} \quad (13.32)$$

因此，其热容可以表示为：

$$C_V = \frac{3}{2} N k_B \frac{\pi^2}{3} \left(\frac{k_B T}{E_F} \right) + O(T^3). \quad (13.33)$$

正如我们上面的分析，费米气体的热容在低温极限下与温度成正比，对应的斜率反比与 E_F 。

☞ **练习 13.4** 请总结一下固体中的电子、声子以及自旋 1/2 系统的热容在低温极限下对温度的依赖关系。

13.8.3 Fermi 气体的磁性

13.8.3.1 Pauli 顺磁性

Fermi 气体的顺磁性是因为在外加磁场作用下，粒子的自旋倾向于与沿磁场方向排列。考虑自旋为 1/2 的粒子，在外场作用下它只能与磁场平行或反平行排列。这时，粒子的能量为：

$$\varepsilon(k) = \frac{\hbar^2 k^2}{2m} + \mu^* \cdot \mathbf{B}$$

这里的第二项为 Zeeman 相互作用能量， μ^* 为粒子的内禀磁矩。

先考虑零温的情况。在外加磁场的作用下，粒子不同自旋的粒子分成了两组。第一组粒子数记为 N^+ ，填充到费米能级 $\varepsilon_F + \mu^* B$ ；第二组粒子数记为 N^- ，填充至 $E_F - \mu^* B$ 。它们满足：

$$N^\pm = \frac{4\pi V}{3h^3} [2m(E_F \pm \mu^* B)]^{3/2}$$

其净磁矩为：

$$M = \mu^* (N^+ - N^-)$$

对应的磁化率为：

$$\chi_0 = \lim_{B \rightarrow 0} \frac{1}{V} \frac{\partial M}{\partial B} = \frac{4\pi \mu^{*2} (2m)^{3/2} E_F^{1/2}}{h^3} = \frac{3}{2} n \mu^{*2} / E_F = D(E_F) \mu^{*2}. \quad (13.34)$$

这里 n 为没有外加磁场时的粒子数密度 (13.25, $g_s = 2$)， $E(E_F) = g(E_F)/V$ 为费米能级位置的

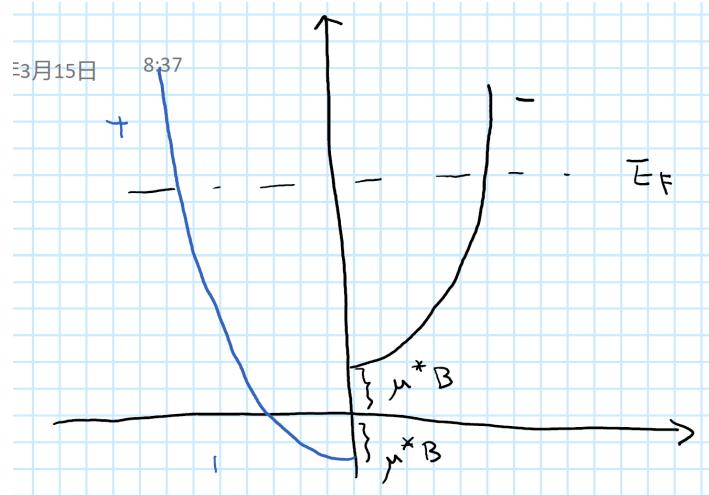


图 13.3：由于外加磁场的作用，不同自旋的费米粒子的能量发生变化，从而导致平衡态后不同自旋取向的粒子数目不同，产生一个宏观的磁矩。从图中可以看出，当外加磁场很小的时候，这个磁矩与费米能级 E_F 位置的态密度成正比。

态密度。可见，磁化率与费米能级的态密度直接相关。这可以通过下面的图像来简单理解。

基于以上简单分析的结果，我们也可以直接从式 (13.14) 出发，来得到不同自旋取向的粒子的数目：

$$N^\pm = -Vn_Q \text{Li}_{3/2}(-\lambda^\pm)$$

这里，不同自旋的有效逸度不同 $\lambda^\pm = e^{\beta(\mu \pm \mu^* B)}$ 。因此，对应的磁化率为：

$$\chi = \lim_{B \rightarrow 0} 2\mu^{*2} \frac{1}{V} \frac{\partial N}{\partial \mu} = -2\mu^{*2} n_Q \text{Li}_{1/2}(-\lambda)/k_B T.$$

基于这个结果，我们可以分析高温和低温极限下的情况：

(1) **高温极限：** $\lambda \ll 1$ ，这时 $\text{Li}_{1/2}(-\lambda) \approx -\lambda$ ，于是：

$$\chi_\infty = n\mu^{*2}/k_B T.$$

这里 n 为所有自旋粒子的总粒子浓度。这是大家熟悉的 Curie 定律形式。

(2) **低温极限：** 这时 $\lambda \rightarrow \infty$ ，利用多对数函数渐进展开形式 (24.13)，如果只保留第一项，我们可以得到式 (13.34)。为了考虑温度效应，我们可以考虑下一项⁴：

$$\text{Li}_{1/2}(-\lambda) \approx -\frac{2\sqrt{\beta\mu}}{\sqrt{\pi}} \left(1 - \frac{\pi^2}{24} (\beta\mu)^{-2} + \dots \right)$$

⁴也可以使用 Sommerfeld 展开

同时考虑到 $\mu(T)$ 的温度依赖，我们得到：

$$\chi(T) = \chi_0 \left(1 - \frac{\pi^2}{12} \left(\frac{k_B T}{E_F} \right)^2 + \dots \right). \quad (13.35)$$

✎ **练习 13.5** 从式 (13.30-13.31) 出发，尝试得到式 (13.35)。

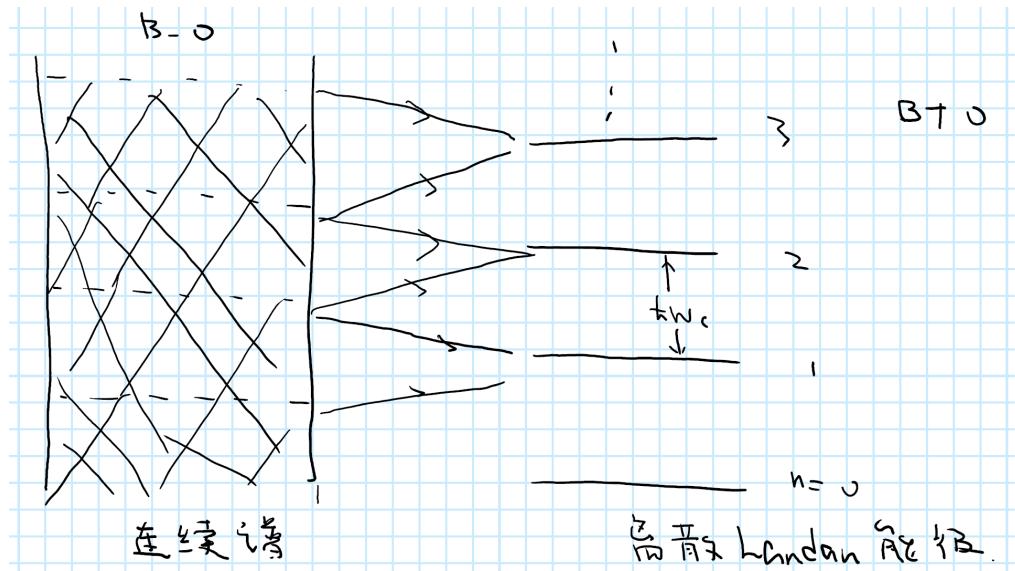
13.8.3.2 Landau 抗磁性

如果费米粒子携带电荷，那么它们在外加电磁场作用下的轨道会发生变化。由于这些微观态的变化，其统计热力学性质也会发生变化。Landau 抗磁性来源于外加磁场对电子轨道的影响。我们知道经典电荷在外磁场作用下如果其初始速度非零会进行圆周运动，对应的角频率为回旋共振频率

$$\omega_c = \frac{eB}{m}$$

通过量子理论分析，我们会发现在垂直于磁场的方向电子的能级会发生变化，形成所谓的 Landau 能级。原本准连续的能谱会变成离散的高度简并的 Landau 能级。这些 Landau 能级高度简并。而在平行磁场的方向，粒子的状态不发生变化，仍然为连续色散。如果考虑三维、非相对论费米气体，假设外加磁场方向沿 z 向，那么其能谱会变为：

$$\varepsilon(n, k_z) = \left(n + \frac{1}{2} \right) \hbar \omega_c + \frac{\hbar^2 k_z^2}{2m}.$$



为了分析其统计热力学性质。我们首先考虑每个 Landau 能级的简并度。它对应 $B = 0$ 时，二维空间能量范围在 $\hbar \omega_c$ 范围内的粒子态的数目。我们知道，二维空间非相对论粒子的态密

度为：

$$g(\varepsilon)d\varepsilon = \frac{Am}{2\pi\hbar^2}d\varepsilon$$

所以对应 Landau 能级简并度为：

$$g_n = \frac{eBA}{h}$$

这里， e 为粒子所带电荷， A 为二维空间的面积。

这样，我们就可以直接写出对应的配分函数：

$$\ln Z = \sum_{k_z} \sum_n g_n \ln(1 + \lambda \exp\{-\beta((n + 1/2)\hbar\omega_c + \hbar^2 k_z^2 / 2m)\})$$

我们按照相同的方法，将对 k_z 的积分转化为求和，利用 Euler 求和规则：

$$\sum_n f(n + 1/2) \approx \int_0^\infty dx f(x) + \frac{1}{24} f'(0)$$

可以得到：

$$= \frac{eBV}{2\pi h} \left\{ \int_{-\infty}^{+\infty} dk_z \int_0^\infty dx \ln \left[1 + \lambda e^{-\beta(\hbar\omega_c x + \hbar^2 k_z^2 / 2m)} \right] - \frac{1}{24} \int_{-\infty}^{+\infty} dk_z \frac{\beta \hbar \omega_c}{\lambda^{-1} e^{\beta \hbar^2 k_z^2 / 2m} + 1} \right\}$$

我们知道 $\omega_c \sim B$ ，所以上式大括号中第一项在积分完之后与 B 无关，因此对系统的磁化率没有贡献。下面我们主要分析第二项的贡献

$$\frac{e^2 B^2 V \text{Li}_{1/2}(-\lambda)}{24h\sqrt{2\pi m k_B T}} = \frac{e^2 B^2 V \text{Li}_{1/2}(-\lambda) \lambda_{\text{th}}}{24h^2}$$

对应贡献的磁化率为：

$$\chi = \frac{1}{\beta VB} \left(\frac{\partial}{\partial B} \ln Z \right)_{\lambda, V, T} = \frac{e^2 \text{Li}_{1/2}(-\lambda)}{12h} \sqrt{\frac{k_B T}{2\pi m}} = \frac{e^2 \lambda_{\text{th}} k_B T \text{Li}_{1/2}(-\lambda)}{12h^2} < 0.$$

因此，电子轨道磁矩贡献的磁化率为负值，对应抗磁性。

(1) **低温极限：**当 $z \gg 1$ 时，和考虑顺磁情况类似，我们有：

$$\begin{aligned} \chi_0 &= -\frac{e^2}{6\pi h} \sqrt{\frac{E_F}{2m}} = -\frac{1}{2} n \mu_{\text{eff}}^2 / E_F = -\frac{1}{3} D(E_F) \mu_{\text{eff}}^2, \\ \chi(T) &= \chi_0 \left(1 - \frac{\pi^2}{12} \left(\frac{k_B T}{E_F} \right)^2 + \dots \right). \end{aligned}$$

(2) **高温极限：**当 $z \ll 1$ 时，利用 $\text{Li}_{1/2}(-\lambda) \approx -\lambda = -n\lambda_{\text{th}}^2$ ，得到高温极限下的结果：

$$\chi_\infty \approx -\frac{e^2 h^2}{48\pi^2 m^2 k_B T} = -n_{\text{eff}}^2 / 3k_B T.$$

我们定义了

$$\mu_{\text{eff}} = \frac{e h}{4\pi m}$$

如果 $m = m_e$ ，那么 $\mu_{\text{eff}} = \mu_B$ 就是大家熟悉的 Bohr 磁子。所以，可以看到，如果是自由电子

气的情况，Landau 抗磁性总是弱于 Pauli 顺磁性，前者为后者的 $1/3$ 。

13.9 应用举例 2: Bose 气体

我们知道光子为玻色子，因此可以直接将上面的结果应用到光子气体。光子自旋为 1，但由于它没有质量，只能有两个极化模式；而且光子由于可以产生和湮灭，在热平衡的状态下化学势 $\mu = 0$ 。代入到式 (13.13-??) 我们可以得到光子气体的热力学性质。可以验证，与我们前面得到的结果相同。

第 14 章 玻色-爱因斯坦凝聚

14.1 问题引入

回顾量子气体部分的内容，我们知道 Bose 和 Fermi 气体的粒子数密度可以表示为：

$$n = g_s n_Q [\mp \text{Li}_{3/2}(\mp \lambda)]. \quad (14.1)$$

其中， $g_s = 2S + 1$ 为其自旋简并度， S 为粒子自旋； $\text{Li}_n(\lambda)$ 为多对数函数； $\lambda = e^{\beta\mu}$ 为逸度；而

$$n_Q = \frac{(2\pi m k_B T)^{3/2}}{h^3}$$

为量子浓度。

对于 Fermi 和 Bose 气体两种情况，我们可以画出相应的多对数函数随 λ 的变化情况，如图14.1所示。这里的一个最重要的信息是，对于玻色子，因为 $0 < \lambda < 1$ ，所以 $0 < \text{Li}_{3/2}(\lambda) < \text{Li}_{3/2}(1) \approx 2.612$ 。多对数函数的这个有限取值范围会带来一个问题。这可以从以下两个方面进行分析。

- 情况 1：我们给定固定浓度的玻色气体，然后降低系统的温度，这时 n_Q 会逐渐降低；式(14.1)的右边会逐步减小，为了保证等式仍然成立， λ 会相应增加，即系统的化学势 μ 逐渐增加。但是，因为 $\text{Li}_{3/2}(\lambda)$ 有一个最大值，达到这个最大值后，进一步降低温度，式(14.1) 将不再成立。
- 情况 2：我们在保持温度不变的情况下，逐步提高系统的粒子浓度。这时，为保证式(14.1)成立， λ 和 μ 同样会增加。式子右边存在一个最大值，当系统浓度增加到一定程度时，式(14.1) 同样会不再成立。

我们前面提到过，当系统粒子数密度增加或者系统温度降低的过程中，系统的量子属性将逐步增强。也就是说，我们上面提到的问题在量子玻色气体中会出现。根据上面的分析，对于玻色气体，我们根据

$$n = g_s n_Q(T_c) \text{Li}_{3/2}(1).$$

定义一个临界温度 T_c

$$T_c = \frac{2\pi\hbar^2}{mk_B} \left(\frac{n}{g_s \text{Li}_{3/2}(1)} \right)^{\frac{2}{3}} \quad (14.2)$$

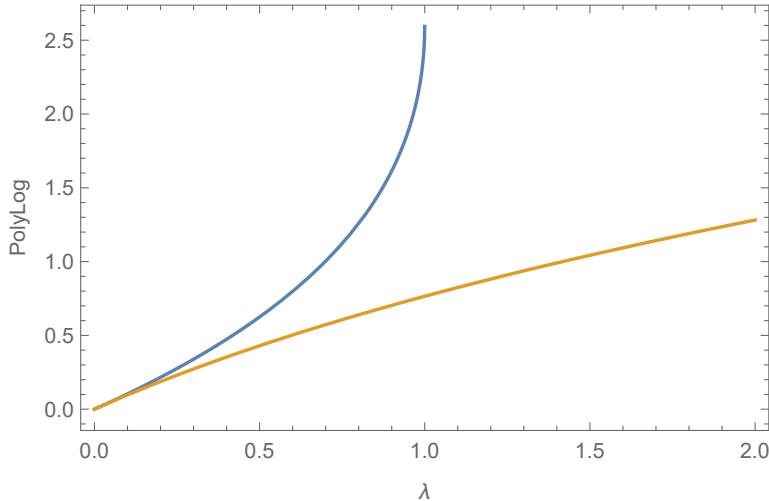


图 14.1: PolyLog 函数, 蓝色对应玻色子, 橙色对费米子。

我们上面的问题出现在 T_c 以下。那么问题的来源是什么呢? 要回答这个问题, 我们先来看一个简单系统。

14.2 宏观系统

对于一个宏观尺度的热力学系统, 我们可以定义一个宏观占据数来表征每个能级上粒子数的多少

$$n_s = \lim_{V \rightarrow \infty} g_s \frac{f_s}{V} = \lim_{V \rightarrow \infty} \frac{1}{V} \frac{g_s}{\lambda^{-1} e^{\beta \varepsilon_s} - 1}. \quad (14.3)$$

我们知道, 对于一个边长为 L 的立方体系统,

$$\varepsilon(n_x, n_y, n_z) = \frac{\hbar^2}{2m} \left(\frac{2\pi}{L} \right)^2 (n_x^2 + n_y^2 + n_z^2).$$

系统的最低能量状态 (ε_0) 和次低能量状态 (ε_1) 的能量差为 $\Delta\varepsilon' \sim \frac{\hbar^2}{2m} \left(\frac{2\pi}{L} \right)^2$ 。当 $L \rightarrow \infty$ 时, 系统的能谱趋于连续, 而且 $\varepsilon_1 \rightarrow 0$, $\Delta\varepsilon' \rightarrow 0$, 所以我们可以将对 \vec{k} 的求和转换为积分的形式进行计算。但是, 在低温极限下这会出现问题。我们仍然把能量参考点取为 $\varepsilon_0 = 0$ 。通过上面对有限系统的分析我们知道, 在低温极限下, $\lambda \sim 1 - 1/N$, 即 $\beta\mu \approx \ln(1 - 1/N) \approx -1/N$ 。这时,

$$\mu \approx -\frac{k_B T}{N} \rightarrow 0$$

这样的话, 我们有 (取 $T \sim 1 \text{ mK}$, $N \sim 10^{23}$)

$$\Delta\varepsilon = \varepsilon_0 - \mu \sim \frac{k_B T}{N} \sim k_B \times 10^{-26}.$$

图 14.2: 能级关系图。

如果对比 ($L = 1 \text{ m}$)

$$\Delta\varepsilon' \sim k_B \times 10^{-18}$$

的话，我们可以看到虽然两者都趋于 0，但它们的相对值显然相差非常大

$$\frac{\Delta\varepsilon'}{\Delta\varepsilon} \sim 10^{-8}$$

其后果就是，次低能级的占据粒子数相对于最低能级要低很多。

$$\frac{n_1}{n_0} \sim 10^{-8},$$

这说明，在低温极限下，最低能级的粒子占据数要远远大于其他能级的占据数。也就是说，

$$n_0 = \lim_{V \rightarrow \infty} \frac{1}{V} \frac{g_s}{\lambda^{-1} - 1}$$

不会趋近于 0. 在这样的情况下，我们计算粒子数时，将对 k 的求和转换为积分的过程就不合适了。直接转换为积分的话，由于 $g(\varepsilon) \rightarrow 0$ ，最低能级的宏观占据数被忽略掉了。如果我们把它加回去，就可以解决上面提到的问题。

14.3 解决方案

考虑到以上因素，我们把基态的粒子数单独拿出来考虑：

$$N = \sum_{\vec{k}} \frac{1}{\lambda^{-1} e^{\beta\varepsilon_{\vec{k}}} \pm 1} = N_0 + \int_0^{+\infty} d\varepsilon \frac{g(\varepsilon)}{\lambda^{-1} e^{\beta\varepsilon} \pm 1}. \quad (14.4)$$

即：

$$n = n_0 + g_s n_Q \text{Li}_{3/2}(\lambda) \quad (14.5)$$

这样的话，当 $\lambda \rightarrow 1$ 时， $n_0 \rightarrow n$ 。不管 n 如何增加，或者 T 如何降低，上式总是可以成立。

14.4 玻色-爱因斯坦凝聚

通过对这个问题的解决，我们得到了一个非常重要的结果。当逐步降低玻色气体的温度时，会有越来越多的粒子聚集到系统的基态（能量最低态）上，最终当温度足够低时，绝大部分粒子都处于基态。这个现象称为玻色-爱因斯坦凝聚 (BEC)。BEC 其实是一个相变过程。下

下面我们具体分析几个物理量随温度的变化行为。

14.4.1 基态宏观占据数 n_0

在 $T > T_c$ 时, $n_0 \rightarrow 0$, 但当 $T < T_c$ 时, 我们有

$$n_0 = n \left[1 - \text{Li}_{3/2}(1) g_s \frac{n_Q}{n} \right] \quad (14.6)$$

$$= n \left[1 - \left(\frac{T}{T_c} \right)^{\frac{3}{2}} \right] \quad (14.7)$$

这时 n_0 就有一个非零值, 而且随着温度趋近于零, $n_0 \rightarrow n$ 。这时, 所有的粒子都处于能量最低的轨道上。 n_0 随温度的变化显示在图14.3上。

14.4.2 内能

根据 n_0 的行为, 我们还可以计算系统的内能随温度的变化。在 T_c 以上, 我们已知

$$U = \frac{3}{2} N k_B T \frac{\text{Li}_{5/2}(\lambda)}{\text{Li}_{3/2}(\lambda)}$$

当 $T < T_c$ 时, 我们需要去除位于能量为零的基态的粒子数

$$U = \frac{3}{2} (N - N_0) k_B T \frac{\text{Li}_{5/2}(\lambda)}{\text{Li}_{3/2}(\lambda)} = \frac{3}{2} N k_B T \frac{\text{Li}_{5/2}(\lambda)}{\text{Li}_{3/2}(\lambda)} \left(\frac{T}{T_c} \right)^{3/2}.$$

更进一步, 从内能的结果我们可以得到其定容热容 C_V 随温度的变化。显然在 T_c 以上和以下, 玻色气体热容的温度依赖发生变化, 如图14.4所示。

14.4.3 逸度 λ

当 $T < T_c$ 时, $\lambda \approx 1$ 。在高温极限下, $\lambda \ll 1$, 我们有 $\text{Li}_{3/2}(\lambda) \approx \lambda$, 这时 $\lambda \propto T^{-3/2}$ 。

14.4.4 等温线

当 $T < T_c$ 时,

$$P = g_s n_Q k_B T \zeta(5/2)$$

与体积没有关系, 设为 P_0 。我们有 $P_0 \propto T^{5/2}$, 可以转化为临界体积 $v_c = 1/n_c$ (n_c 由临界温度来确定, 见14.2) 得到 $P_0 v_c^{5/3} = C$ 。这给出了 $P - v$ 图上的一条转变曲线。在高温极限下, 回归经典理想气体的结果 $P \propto V^{-1}$ (图14.5)。

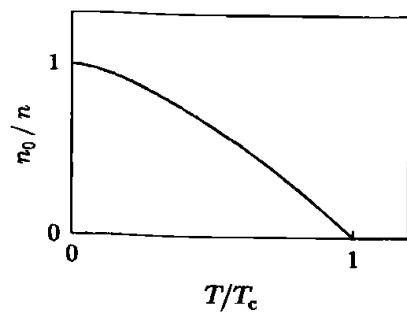
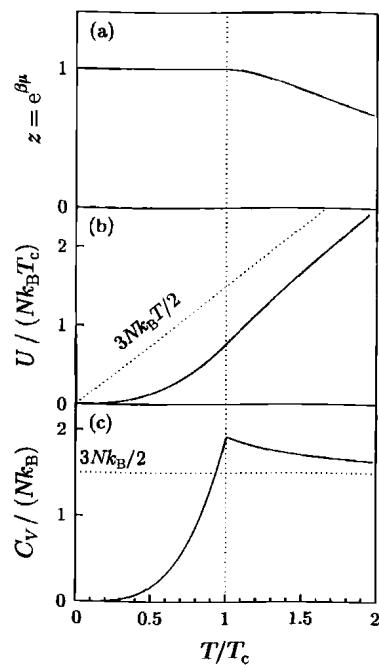
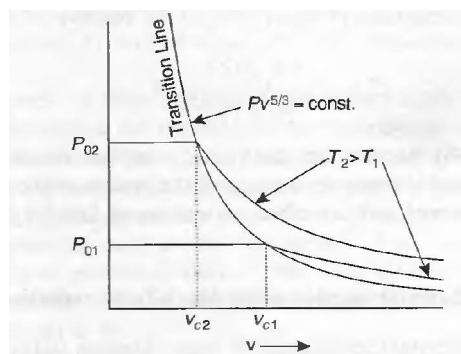
图 14.3: n_0 随温度的变化。图 14.4: λ , U 和 C_v 随温度的变化。

FIG. 7.6. The isotherms of an ideal Bose gas.

图 14.5: 理想玻色气体的等温 P - V 曲线。

✎ **练习 14.1** 理想玻色气体的等熵过程：利用热力学关系式

$$U = TS - PV + N\mu$$

我们可以得到理想玻色气体的熵

$$\frac{S}{Nk_B} = \frac{U + PV}{Nk_B T} - \frac{\mu}{k_B T} = \begin{cases} \frac{5}{2} \frac{\text{Li}_{5/2}(\lambda)}{\text{Li}_{3/2}(\lambda)} - \ln \lambda, & T > T_c \\ \frac{5}{2} \zeta(5/2) \frac{g_s n_Q}{n}, & T \leq T_c \end{cases} \quad (14.8)$$

(1) 证明其绝热线可以表示为：

$$Pv^{5/3} = \text{Constant}$$

与经典理想气体的情况完全相同。(2) 证明理想玻色气体的绝热系数

$$\gamma = \frac{C_P}{C_V} = \frac{5}{3} \frac{\text{Li}_{5/2}(\lambda) \text{Li}_{1/2}(\lambda)}{(\text{Li}_{3/2}(\lambda))^2}.$$

只有当 $T \gg T_c$ 时才有 $\gamma \approx 5/3$ 趋于经典理想气体的结果；而当 $T \rightarrow T_c$ 时， $\gamma \rightarrow \infty$ 。

第 15 章 van der Waals 方程



理想气体模型是热力学与统计物理中最为有用的模型，可以帮助我们理解很多基本概念和理论，也可以用于解释诸如黑体辐射、声子热容等重要的实验结果。但理想气体模型忽略了粒子的实际大小，也忽略了它们之间的相互作用势，因此在理解分子浓度比较高时的现象就无能为力。为了考虑这些因素，我们可以从理想气体方程出发对其进行修正。可以想象，根据我们研究的具体系统，我们提出不同的修正，从而可以得到非常多形式的修正。这里，我们将重点介绍 van der Waals (vdW) 对理想气体方程做了修正，以及以他的名字命名的方程。我们首先使用统计物理的理论来给出对它的简单推导。这样一来我们可以更好的从微观的角度去理解 vdW 方程，另一方面我们也可以了解当包含了粒子间的相互作用势后，如何去近似的计算系统的配分函数分析其热力学性质。然后，我们基于 vdW 方程来分析实际气体的一些性质。

15.1 vdW 方程

15.1.1 相互作用势

分子之间距离很近的时候，排斥势占主导，距离较远时，吸引势占主导。在 vdW 势中，分子之间的相互吸引是由于中性分子中的随机偶极振荡导致的。假设一个中性分子中产生一个随机偶极子 p_1 ，那么它会在距离 r 处的另一个分子中诱导处另外一个偶极子 $p_2 \sim p_1/r^3$ 。两个偶极子之间的相互作用势

$$U(r) \propto \frac{p_1 p_2}{r^3} \propto \frac{1}{r^6}$$

对于排斥势，我们可以采用最简单的模型，认为两个分子间距小于一个临界值后，排斥势为 ∞ (硬球势)。这个距离通常对应分子的直径。于是，我们可以把我们势函数模型写为(见图15.1):

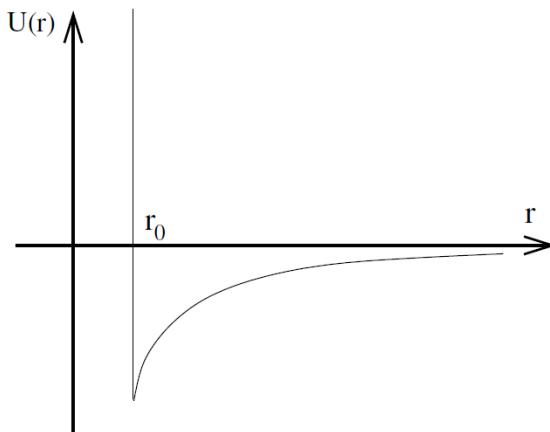


图 15.1: Hard-core and vdW potential

$$U(r) = \begin{cases} +\infty, & r \leq r_0 \\ -U_0 \left(\frac{r_0}{r}\right)^6, & r > r_0 \end{cases} \quad (15.1)$$

另外一种更为准确、常用的势函数是 Lennard-Jones (LJ) 势或者叫 6-12 势：

$$U(r) = \frac{A}{r^{12}} - \frac{B}{r^6}. \quad (15.2)$$

这里我们使用式 (15.1)。包含了相互作用势后的 N 粒子系统哈密顿量为

$$H = \sum_{i=1}^N \frac{p_i^2}{2m} + \sum_{i < j} U(r_{ij}). \quad (15.3)$$

这里的求和条件 $i < j$ 是为了避免重复计数。因为我们这里采用的是一对分子间的相互作用势，如果没有这个条件的话，每个势能项会被统计两次。

15.1.2 配分函数

有了哈密顿量后，我们可以写出 N 粒子系统的配分函数（参考理想气体配分函数）

$$Z(N, V, T) = \frac{1}{N! h^{3N}} \int d^{3N} \vec{p}_i d^{3N} \vec{r}_i e^{-\beta H}. \quad (15.4)$$

这个形式大家目前应该已经非常熟悉了，不需要再做过多介绍。下面我们需要考虑的是引入相互作用势能后的系统配分函数如何计算。因为哈密顿量中的动能和势能是分开的，当忽略 H 中的势能项后，我们会得到理想气体的配分函数。所以，我们可以把上式重写为

$$Z = \frac{n_Q^N}{N!} \int d^{3N} \vec{r}_i e^{-\beta U} \quad (15.5)$$

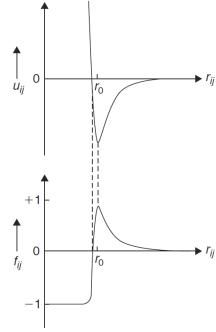


图 15.2: From Pathria, Chap. 10: A typical two-body potential u_{ij} and the corresponding Mayer function f_{ij} .

这里动能部分的计算和理想气体相同，而总的势能我们用 U 来表示，为了简化我们把粒子坐标省去了。为了处理势能项，我们能想到的最简单的办法就是对它做 Taylor 展开

$$e^{-\beta U} = 1 - \beta U + \frac{1}{2}(\beta U)^2 + \dots \quad (15.6)$$

但不幸的是，这个展开是不收敛了，因为在我们的模型中 U 不是一个小量。比如，当两个粒子距离很近时， U 会非常大。这个问题可以通过引入 Mayer f -函数来解决：

$$\boxed{f(r) \equiv e^{-\beta U(r)} - 1.} \quad (15.7)$$

分析一下 $f(r)$ 我们发现， $f(r)$ 的取值范围要小很多，在高温极限下它基本局限在 $[-1, 1]$ 的范围内（图15.2），因此我们可以对 $f(r)$ 函数做展开。于是，

$$Z = \frac{n_Q^N}{N!} \int d^{3N} \vec{r}_i \prod_{j>k} (f_{jk} + 1). \quad (15.8)$$

把上式中的乘积展开我们得到

$$Z = \frac{n_Q^N}{N!} \int d^{3N} \vec{r}_i \left(1 + \sum_{j>k} f_{jk} + \dots \right) \quad (15.9)$$

我们这里只保留了 f_{jk} 的一阶项，忽略所有其它高阶项。上式第一项给出 V^N ，对应理想气体的结果。对于第二项我们可以做进一步的简化：

$$\int d^{3N} \vec{r}_i f_{jk} = V^{N-2} \int d\vec{r}_j d\vec{r}_k f_{jk} = V^{N-1} \int d\vec{r} f(r).$$

这里，我们利用了势函数只依赖于两个粒子间的距离。式 (15.9) 中的求和中包含了 $N(N-1)/2$ 个类似的项。因为 N 足够大，我们可以近似为 $N^2/2$ 。综上，我们有

$$Z \approx Z_{\text{ideal}} \left(1 + \frac{N^2}{2V} \int d\vec{r} f(r) \right) \quad (15.10)$$

于是, Helmholtz 自由能为:

$$\begin{aligned} F &= -k_B T \ln Z \\ &\approx F_{\text{ideal}} - k_B T \ln \left(1 + \frac{N^2}{2V} \int d\vec{r} f(r) \right) \\ &= F_{\text{ideal}} - N k_B T \left(\frac{N}{2V} \int d\vec{r} f(r) \right). \end{aligned} \quad (15.11)$$

这里的下标 ideal 表示理想气体。从自由能我们可以得到相应的物态方程:

$$P = - \left(\frac{\partial F}{\partial V} \right)_T = \frac{N k_B T}{V} \left(1 - \frac{N}{2V} \int d\vec{r} f(r) \right). \quad (15.12)$$

在这个结果中, 粒子间的相互作用势全部集中到了括号中对 $f(r)$ 的积分了。只要我们知道了相互作用势函数的具体形式, 就可以利用上式得出对应的状态方程。

以上对相互作用势的处理是一个非常简化版本的集团展开方法 (the method of cluster expansions)。它是由 Mayer 及合作者从 1937 年开始发展的一个处理经典体系相互作用势的系统方法。之后, Kahn 和 Uhlenbeck (1938) 尝试将其推广到量子系统。最后, 由李政道和杨振宁在他们 1959-1960 年的 5 篇系列论文中彻底完成这一任务¹。关于这些内容的详细讲述请可以参考一些研究生教材, 比如 Pathria。

15.1.3 van der Waals 方程

下面我们利用式 (15.1) 来计算

$$\begin{aligned} \int d\vec{r} f(r) &= \int_0^{r_0} (-1) d\vec{r} + \int_{r_0}^{+\infty} \left(e^{\beta U_0 (\frac{r_0}{r})^6} - 1 \right) d\vec{r} \\ &\approx -\frac{4}{3} \pi r_0^3 + \beta U_0 \int_{r_0}^{+\infty} \left(\frac{r_0}{r} \right)^6 d\vec{r} \\ &= \frac{4}{3} \pi r_0^3 (\beta U_0 - 1) \\ &= 2(\beta a - b). \end{aligned} \quad (15.13)$$

我们定义了两个常数:

$$a = bU_0, \quad (15.14)$$

和

$$b = \frac{2\pi r_0^3}{3}. \quad (15.15)$$

¹T. D. Lee, C. N. Yang, “Many-Body Problem in Quantum Statistical Mechanics”, Phys. Rev., 113, 1165; 116, 25; 117, 12; 117, 22; 117, 897.

常数 a 和吸引势能大小有关，而 b 和分子所占体积有关²。于是我们得到

$$\ln Z = \ln Z_{\text{ideal}} + \ln \left(1 + \frac{N^2}{V} (\beta a - b) \right) \approx \ln Z_{\text{ideal}} + \frac{N^2}{V} (\beta a - b). \quad (15.16)$$

基于这个结果我们就可以写下其相应的热力学量的表达式

- 内能：

$$U = \frac{3}{2} N k_B T - \frac{aN^2}{V}.$$

可见，相互吸引势对内能有一个修正，但分子的有限大小不会影响内能。

- Helmholtz 自由能：

$$\begin{aligned} F &= F_{\text{ideal}} - N k_B T \left(\frac{N}{V} (\beta a - b) \right) \\ &\approx -N k_B T \left(\ln \frac{n_Q (V - Nb)}{N} + 1 \right) - \frac{aN^2}{V}. \end{aligned} \quad (15.17)$$

- 熵：

$$S = \frac{U - F}{T} = N k_B \left[\ln \frac{n_Q (V - Nb)}{N} + \frac{5}{2} \right]$$

可见，分子的有限大小因为影响到分子的空间占据，所以对熵有影响，但相互作用势不会影响熵。

- 状态方程为：

$$P = \frac{N k_B T}{V - Nb} - \frac{aN^2}{V^2}. \quad (15.18)$$

重新整理一下，我们得到：

$$\left(P + a \frac{N^2}{V^2} \right) (V - bN) = N k_B T. \quad (15.19)$$

取 $N = N_A$ ，以上方程可以等价的写为：

$$\left(P + \frac{a_m}{V_m^2} \right) (V_m - b_m) = RT. \quad (15.20)$$

这里， $a_m = aN_A^2$, $b_m = bN_A$ 。

²为什么不是 $4\pi r_0^3/3$ 而只是其一半？

15.2 vdW 方程的应用

15.2.1 vdW 气体的 Joule 膨胀

我们知道理想气体发生 Joule 膨胀的过程中其内能不变，温度也不变。对于 vdW 气体，由于分子之间的相互作用，气体在膨胀过程中势能发生变化，而总的内能守恒，所以我们可以推测其动能降低，温度也会随之降低。对于实际的气体我们可以定义

$$\mu_J = \left(\frac{\partial T}{\partial V} \right)_U \quad (15.21)$$

来表征 Joule 膨胀过程中其温度的变化。

利用我们的热力学知识，我们可以把它稍作整理写成一个方便使用物态方程的形式：

$$\begin{aligned} \mu_J &= -\frac{1}{C_V} \left(\frac{\partial U}{\partial V} \right)_T \\ &= -\frac{1}{C_V} \left[T \left(\frac{\partial P}{\partial T} \right)_V - P \right]. \end{aligned}$$

利用 vdW 物态方程，我们可以得到

$$\mu_J = -\frac{aN^2}{C_V V^2} < 0.$$

这个结果与 a 有关，表明确实是由于粒子间的相互吸引导致实际气体的温度在 Joule 膨胀过程中降低。

15.2.2 vdW 气体的熵

我们也可以从热力学的角度来考虑 vdW 气体的熵。把熵表示为温度和体积的变量，可以得到：

$$\begin{aligned} dS &= \left(\frac{\partial S}{\partial T} \right)_V dT + \left(\frac{\partial S}{\partial V} \right)_T dV \\ &= \frac{C_V}{T} dT + \left(\frac{\partial P}{\partial T} \right)_V dV. \end{aligned} \quad (15.22)$$

对于 vdW 气体来说，

$$\left(\frac{\partial P}{\partial T} \right)_V = \frac{Nk_B}{V - Nb},$$

所以

$$S = C_V \ln T + Nk_B \ln(V - Nb) + \text{constant}.$$

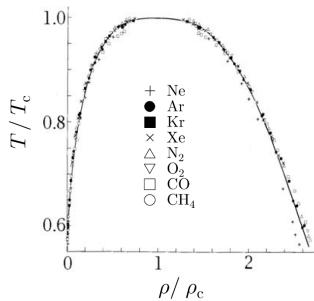


图 15.3: 不同种类气体的气液共存线满足相同的约化形式。from: Blundell quoted from Guggenheim.

只与 b 有关! 这与上一节得出的结果相似。

15.2.3 临界点及约化方程

与理想气体方程相比, vdW 方程最大的特点是在低温下, 给定一个压强, 可以对应多个体积。因为它是关于体积的三次方程。对应的等温线会与理想气体有很大的区别, 见图 (15.4)。从等温线我们发现, vdW 气体存在一个临界温度 T_c , 这个温度以下, 给定 T 和 P , 会存在三个不同的体积都满足 vdW 方程。通过条件

$$\frac{\partial P}{\partial V} = \frac{\partial^2 P}{\partial V^2} = 0$$

我们可以确定一个临界点 (critical point):

$$P_c = a/27b^2, \quad T_c = 8a/27k_B b, \quad V_c = 3Nb.$$

利用临界点的物理量我们可以把 vdW 方程写成一个约化的形式

$$\left(\tilde{P} + \frac{3}{\tilde{V}^2} \right) (3\tilde{V} - 1) = 8\tilde{T}. \quad (15.23)$$

这里, 我们定义了无量纲的约化量定义约化量:

$$\tilde{P} = P/P_c, \quad \tilde{V} = V/V_c, \quad \tilde{T} = T/T_c.$$

这个方程不依赖于分子间相互作用势的强弱和分子的几何尺寸, 只与相互作用势的函数形式有关。不同种类的气体可以用相同的约化方程来描述。下图给出了不同种类气体的气液共存线, 可以看到它们满足相同的函数形式

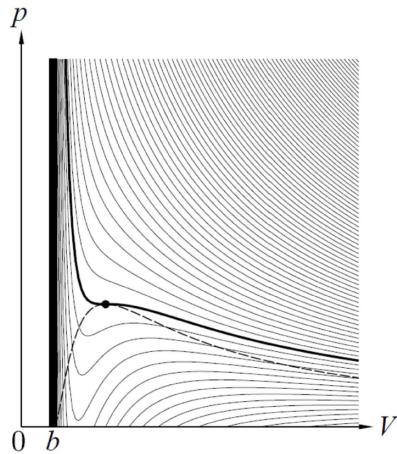


图 15.4: vdW 气体的等温线。

15.2.4 气液共存线

上面我们说到，在临界点以下，vdW 方程存在多个解。实际上，这三个解中体积最大的为气态解，体积最小的为液态解，而中间的解是不稳定的。从 $T < T_c$ 时的等温线我们可以看到，中间解对应位置的 $\partial P / \partial V > 0$ 。这是一个不稳定的解，对应的等温压缩率 κ_T 为负。在这个位置对系统进行微小扰动会使它马上偏离这个状态。在给定温度和压强的情况下，我们可以通过系统的 Gibbs 势来判断其稳定性。所以，我们可以画出 vdW 气体 Gibbs 势与压强的对应关系图，如 (15.5) 所示³。从这个图我们可以看出，系统在气态 B_2 和液态 B_1 两个点的 Gibbs 势是相同的。事实上，系统会在与 B_1/B_2 对应的压强下发生气液相变，相变过程中系统的 Gibbs 函数、温度、压强等都保持不变。这种情况下的等温线实际上是连接 B_1 和 B_2 的一条直线。vdW 方程给出的线中， $B_1 - X$ 和 $Y - B_2$ 这两段状态的 Gibbs 函数高于平衡态的值，为亚稳态；而 $X - Y$ 这段为非稳态。从 vdW 等温线如何找到对应的发生气液相变的压强呢？我们可以根据 $G(B_1) = G(B_2)$ 来得到。在温度和粒子数都不变的情况下， $dG = VdP$ ，所以我们要求沿着路径 $p_1 : B_1 \rightarrow X \rightarrow Y \rightarrow B_2$ 的积分保持 Gibbs 函数不变：

$$\int_{p_1} V dP = 0.$$

这个条件要求，等温线中的气液共存线段 $B_1 - B_2$ ，把范德耳斯等温线 p_1 分成两块儿相等的面积。这就是 Maxwell 等面积法则。根据这一法则，我们很容易从 vdW 等温线得到对应温度下的气液共存线。

³这个图如何画出来？尝试自己画一下。

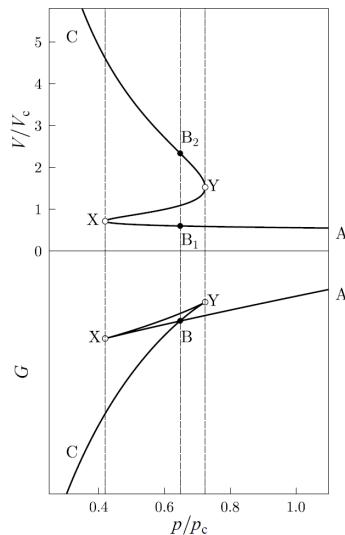


图 15.5: Maxwell 等面积法则。

15.3 位力展开

vdW 方程只是对实际气体物态方程的一种修正形式。在研究具体物理系统的时候，我们还可以考虑其它类型的修正。一个比较系统的修正方式是采用位力展开的方式。我们把一般形式的物态方程写成如下关于体积的展开式形式：

$$\frac{PV_m}{RT} = 1 + \frac{B}{V_m} + \frac{C}{V_m^2} + \dots$$

这里 B, C 等为位力展开系数，这些系数是温度的函数。不同的位力系数给出不同的物态方程。

第 16 章 相变

16.1 相与相变

我们知道同一种物质可以处于不同的相。不同的温度、压强等热力学条件下，系统所处的相可能不同。尽管组成不同相的原子或分子完全相同，但由于它们的微观结构不同，导致不同的相表现出不同的宏观性质。我们可以从这些性质的变化来知道系统是否发生了相变。大家最为熟悉的就是同一物质的气、液、固三种相。气体分子之间距离比较远，分子间相互作用力较弱，因此相同物质的量的气体体积比液体和固体要大得多。固体中原子或分子之间的距离要小得多，它们之间的相互作用要强得多，而且原子往往只能在平衡位置附近振动，无法像在气体或液体中那样改变空间位置。液体与固体的体积相似，却和气体一样具有流动性。

以上是我们对相的最为直接的理解。实际上，相的含义要更加丰富得多。比如，同一种元素组成的固体可以处于多种不同的相。最为常见的就是碳元素。由于碳原子的空间排布不同，可以存在金刚石、石墨、石墨烯、碳纳米管等不同的结构，它们的物理、化学性质也截然不同。金刚石是绝缘体，而石墨具有较好的导电性，碳纳米管既可以是导体也可以是半导体或绝缘体。此外，即使相同结构的系统，其物理性质也可能随着温度的变化而发生大的改变。比如，很多金属甚至绝缘体在温度低于一个临界温度后电阻率会突然降到零，进入到超导状态。在这前后，系统的晶格结构并没有发生变化，但它也经历了一个相变。所以，相变不仅仅体现在结构的变化，还可以体现在很多其他电学、磁学等物理性质。以上，我们只讨论了单元系，实际上我们研究的系统还可以包含多种类型的物质，形成多元系。多元系也可以有更为丰富的物相。

在热力学参量发生变化时，系统可能从一种相转变为另外一个相，这个过程称为相变。几种不同的相在特定的热力学条件下也可以共存达到平衡。不同相处于热力学平衡时它们的温度、压强、化学势等都相同，这保证了热学、力学和化学平衡的条件可以得到满足。我们可以在 $P-T$ 图中来描述不同相存在的范围，不同相之间通过它们的共存线分割开来。这称为相图。

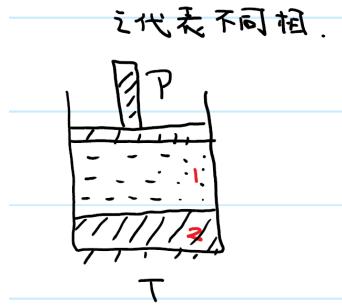


图 16.1: 在给定温度的情况下，通过改变系统的压强可以实现气态与液态之间的相变。

16.1.1 Gibbs 相律

考虑由不同组元组成的多元系在多种相共存的情况下达到平衡。不同组元之间的占比表示为 x_i ，我们有

$$\sum_{i=1}^C x_i = 1.$$

系统的总变量数为

$$2 + P(C - 1)$$

其中，2 中包含 T, P 两个变量，而每一相中不同组元的占比共有 $C - 1$ 个变量，总变量数需要再乘以总的相数 P 。当系统达到平衡之后，不同相之间的同种组元会达到化学平衡，所以一共有

$$C(P - 1)$$

个平衡限制条件。这样的话，系统总的自由度数目为：

$$2 + P(C - 1) - C(P - 1) = 2 - P + C \geq 0.$$

我们需要这是一个非负的整数。所以，我们得到如下限制条件：

$$P \leq 2 + C.$$

这就是 Gibbs 相律。它给出了一个系统中可以同时存在的相的最大数目。比如，对于单元系， $C = 1$, $P \leq 3$ 。也就是，最多有三相共存。

16.1.2 相变的分类

Ehrenfest 根据热力学势(比如 Gibbs 或 Helmholtz 势)相对于温度或压强¹的微分是否连续将相变分为不连续相变和连续相变。连续相变中我们可以进一步把它们分为 n 级相变, 这里的 n 表示热力学势对温度或压强的第 n 阶微分不连续。选择温度和压强作为外界控制参量的话, 我们使用 Gibbs 自由能作为热力学势最为方便。根据 $dG = VdP - SdT$ 我们可以知道, 一级相变前后系统的体积 V 或者熵 S 会发生不连续变化。其中, 熵的不连续变化表现为系统在相变时需要吸收或释放热量, 称为相变潜热。比如气-液、气-固、固-液相变都属于一级相变, 都伴随着体积的不连续变化以及吸热和放热现象。二级相变没有体积跃变或相变潜热, 但其热容可以有不连续的跳变, 因为热容可以表示为 $C_P = T(\partial S / \partial T)_P$, 对应 Gibbs 势对温度的二阶微分。

① 一级相变: $\frac{\partial G}{\partial x} \Big|_{T_c}$ 不连续; 气-液.

$$S = -\left(\frac{\partial G}{\partial T}\right)_P, \quad V = \left(\frac{\partial G}{\partial P}\right)_T.$$



② 二级相变: $\frac{\partial^2 G}{\partial x^2}$ 不连续.

$$\text{比如: } C_x = T \left(\frac{\partial S}{\partial T}\right)_x, \quad S = -\left(\frac{\partial G}{\partial T}\right)_P.$$

如果 C_x 不连续, 为二级相变.



16.2 气-液-固相变

16.2.1 Clausius-Clapeyron 方程

我们可以根据两相共存达到平衡的条件来得到描述两相界面的方程。两相之间达到热力学平衡后, 它们有相同的 T, P 和 μ 分别来表征它们之间的热、力和化学平衡。用 l 和 g 来分别表示两相, 我们有:

$$T_l = T_g, \quad P_l = P_g, \quad \mu_l = \mu_g.$$

¹ 我们这里采用 PVT 系统来说明, 因此对应的强度量为压强。这些讨论可以推广到其它变量。

沿着相界我们有：

$$\mu_l(T, P) = \mu_g(T, P), \quad \mu_l(T + dT, P + dP) = \mu_g(T + dT, P + dP).$$

这里， (T, P) 和 $(T + dT, P + dP)$ 分别是相界上两个紧邻的两个点 ($dT \ll T, dP \ll P$)。我们对第二个式子做 Taylor 展开至一阶，并利用第一个式子，可以得到：

$$-S_l dT + V_l dP = -S_g dT + V_g dP, \quad (16.1)$$

整理之后我们可以得到所谓的 Clausius-Clapeyron 方程：

$$\boxed{\frac{dP}{dT} = \frac{L}{T(V_g - V_l)}}. \quad (16.2)$$

其中， L 为相变潜热。通常我们可以取体积 V_g 和 V_l 为摩尔体积，那么 L 也对应每摩尔相变吸收的热量。当然，如果两个体积取为每个粒子的平均体积，那么 L 也要做相应的改变。但不管取什么样的单位，它们的比值不变。Clausius-Clapeyron 方程可用于描述一级相变的相界，它之所以重要是因为方程右边的所有量都可以进行实验观测。

如果遇到连续相变，我们就无法使用 Clausius-Clapeyron 方程了。对于 n 阶相变，我们需要回到式 (16.1)，将系数扩展到 n 阶来得到相应的关系式。比如，对于二级相变，我们可以得到：

$$\frac{dP}{dT} = \frac{\alpha_g - \alpha_l}{\kappa_{T,g} - \kappa_{T,l}} = \frac{C_{P,g} - C_{P,l}}{T(\alpha_g - \alpha_l)}. \quad (16.3)$$

1. 若 L 不依赖温度，气体满足理想气体方程：

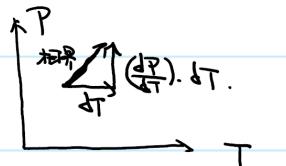
$$\Rightarrow \frac{dP}{dT} = \frac{LP}{T^2 R} \Rightarrow \frac{dP}{P} = \frac{L}{R} \frac{dT}{T^2}.$$

$$\Rightarrow \ln P = -\frac{L}{RT} + C. \Rightarrow P = P_0 e^{-\frac{L}{RT}}$$

举例：得出气液相变的近似 $P-T$ 关系。

2. 依熱溫系.

$$\text{沿着相界: } \frac{d}{dT} \left(\frac{L}{T} \right) = \left(\frac{\partial}{\partial T} \right)_P + \left(\frac{\partial}{\partial P} \right)_T \cdot \left(\frac{dP}{dT} \right)$$



$$\text{由 } \frac{L}{T} = \Delta S = S_v - S_L.$$

$$\Rightarrow \frac{d}{dT} \left(\frac{L}{T} \right) = \left(\frac{\partial \Delta S}{\partial T} \right)_P + \left(\frac{\partial \Delta S}{\partial P} \right)_T \cdot \left(\frac{dP}{dT} \right)$$

$$= \frac{1}{T} (C_{PV} - C_{PL}) + \left[\left(\frac{\partial S_v}{\partial P} \right)_T - \left(\frac{\partial S_L}{\partial P} \right)_T \right] \cdot \left(\frac{dP}{dT} \right)$$

$$= \frac{1}{T} (C_{PV} - C_{PL}) - \left(\left(\frac{\partial V_v}{\partial T} \right)_P - \left(\frac{\partial V_L}{\partial T} \right)_P \right) \cdot \left(\frac{dP}{dT} \right)$$

$$\text{由 } V_v \gg V_L, \quad PV_v = RT$$

$$\Rightarrow \frac{d}{dT} \left(\frac{L}{T} \right) = \frac{C_{PV} - C_{PL}}{T} - \frac{R}{P} \cdot \frac{LP}{RT^2}.$$

$$\downarrow -\frac{L}{T^2} + \frac{1}{T} \left(\frac{dL}{dT} \right)$$

$$\Rightarrow dL = (C_{PV} - C_{PL}) dT.$$

$$\Rightarrow L = L_0 + (C_{PV} - C_{PL}) T.$$

$$\text{得: } P = P_0 \exp \left[-\frac{L_0}{RT} + \frac{(C_{PV} - C_{PL}) \ln T}{R} \right].$$

16.2.2 液滴/气泡的形成

通过对 vdW 气体等温线的分析我们知道，在特定的压强和温度范围内可能存在过冷气体或者过热液体。这里我们进一步分析它们出现的原因。在气液相变发生时，首先要在气体环境中形成小的液体；而在液气相变发生时，需要在液体环境中形成小的气泡。这对应着新相析出的过程。在这个过程发生的开始，往往需要额外的能量供给来克服液体表面能。如果没有足够的能量，那么即使达到了相变的热力学条件，相变过程也不一定会发生。这时就会形成过冷气体或者过热液体。

我们以气相中液滴的形成来说明这一问题。我们可以对比在液滴形成前后的 Gibbs 自由能的变化：

$$\Delta G(r) = G_l - G_g = \frac{4}{3}\pi r^3 n_l \Delta \mu + 4\pi r^2 \gamma.$$

这里 γ 为表面张力系数。当液滴的半径 r 很小时，尽管按照我们前面的热力学考虑，达到了

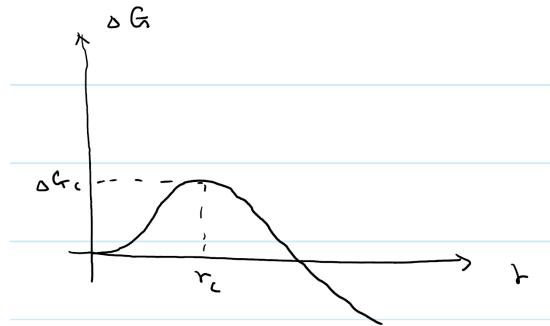


图 16.2: 气液 Gibbs 自由能随液滴半径大小的变化。

气液相变的条件 $\Delta\mu \leq 0$ ，但是由于表面能的限制， $\Delta G(r) > 0$ ，相变并不会发生。这时，即使系统中出现小的液滴，也会由于表面张力的作用而变小消失。因为考虑了表面张力的影响后，气态仍然是 Gibbs 自由能比较小的状态。只有当 r 达到一个临界值时，系统才会跨过这个自由能势垒，于是液滴逐渐长大，发生气液相变。数学上，当 r 非常小时， r^2 项总是大于 r^3 项；只有当 r 超过一个临界值时， r^3 项才会占主导。对应的势函数曲线如图 (16.2) 所示。这个临界半径可以由下式得到：

$$\frac{\partial \Delta G(r)}{\partial r} = 0.$$

$$\rightarrow r_c = \frac{2\gamma}{n_l |\Delta\mu|}.$$

当 r 取 r_c 时，对应最高的势垒

$$\Delta G_c = \frac{16\pi}{3} \frac{\gamma^3}{(n_l \Delta\mu)^2}.$$

16.3 超导相变

超导的发现得益于低温技术的发展。Onnes 成功将 He 液化后发现（1911 年），在液氦温度附近，水银的电阻率随着温度降低会发生一个跳变。低于一个临界温度后，其电阻率接近于零。这就是超导现象。这一电阻率的跳变也对应着水银的金属-超导相变。之后人们发现很多其他金属都会发生类似的相变，只不过相变的温度有所不同。Onnes 也因为这一发现获得了 1913 年的诺贝尔物理学奖。大约 60 年之后，Bardeen, Cooper, Schrieffer (BCS) 三人因为构建了超导现象的微观理论再次获得诺贝尔物理学奖。再之后的 1986 年 Bednorz 和 Müller 因为发现了铜基氧化物的超导现象再次获得诺贝尔物理学奖。由于铜基氧化物在超导温度之上为绝缘体，而且它们的超导转变温度往往高于普通的金属，因此这类超导体被称为高温超导

体；而之前的 BCS 超导机制又被称为常规超导。之后，人们又发现了铁基氧化物中的高温超导现象。目前，人们对高温超导的微观机理尚不清楚。高温超导的机理也被称为凝聚态物理中的一颗皇冠，对其机理的成功解释必然会再次带来诺贝尔物理学奖。

这里我们只介绍常规超导体中的超导相变。常规超导体根据它们在磁场下的不同行为有可以被分为第一类和第二类超导体。其中，第一类超导体表现出完美的抗磁性。如果在超导转变温度以下对样品施加磁场 H 的话，超导体样品会产生对磁场的完全排斥。也就是说，超导体内部的磁感应强度 $B = 0$ 。这是因为超导体会产生与外加磁场强度大小相同，方向相反的磁矩，从而抵消外场带来的有效磁矩导致 $B = 0$ 。刚才说介绍的超导体的抗磁性只适用于外加磁场不大的情况下，实验发现当外加磁场超过一个临界值时，超导会被破坏；破坏之后，磁场就可以进入导体的内部。第二类超导体存在两个临界磁场。介于两个临界磁场之间时，超导体并不表现出完全的抗磁性，它可以允许磁场进入超导体内部部分区域，从而产生非零的磁通。这里我们只关注第一类超导体。我们可以根据样品是否在超导状态来确定样品内部的磁感应强度：

$$B = \begin{cases} 0, & H < H_c \\ \mu_0 H, & H > H_c. \end{cases}$$

我们首先来考虑金属-超导相变的相界。这个问题分析类似我们考虑气-液相变的相界从而得到 Clausius-Clapeyron 方程的过程。只不过这里的压强变成了磁场强度 H ，而体积变成了系统的磁感应强度 B 。于是，在相边界，超导 (SC) 和正常 (N) 相之间满足：

$$-s_{SC}dT - B_{SC}dH_{co} = -s_NdT - B_NdH_{co}. \quad (16.4)$$

因为超导体中 $B_{SC} = 0$ ，所以我们得到：

$$\frac{dH_{co}}{dT} = \frac{s_{SC} - s_N}{\mu_0 H_{co}}. \quad (16.5)$$

这是正常相和超导相的边界所满足的微分关系式，对应超导相变的 Clausius-Clapeyron 方程， H_{co} 为对应温度下的两相共存时的磁场。随着温度的增加， H_{co} 应该会减小，因此 $s_{SC} < s_N$ 。超导相的熵要小于正常相，因为超导相更有序。

事实上， H_{co} 近似满足如下的函数关系（图16.3）：

$$H_{co} = H_c \left(1 - \frac{T^2}{T_c^2}\right).$$

这里的 H_c 为零温下的临界磁场强度。于是，在相界的位置，

$$s_{SC} - s_N = -2\mu_0 H_c^2 \frac{T}{T_c^2} \left(1 - \frac{T^2}{T_c^2}\right)$$

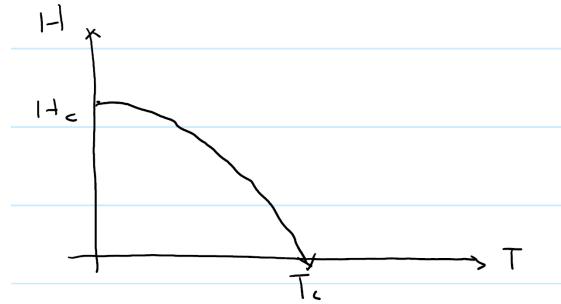


图 16.3: 临界磁场随温度的变化。

可见，

- 在 $T \rightarrow 0$ 时，两相的熵都趋于零，满足第三定律的要求；而在 $T = T_c$ 时两者也相等。在中间的温度范围内， $S_{SC} < S_N$ 。对应超导相的熵在温度趋于零时，以更快的速度趋于零。
- 我们知道正常金属相的热容主要来源于电子的贡献 $C_N = \gamma T$ 。实际上，在低温极限下 $T \rightarrow 0$ 时， $C_{SC} - C_N \approx -C_N$ 主要来源于正常相的贡献。据此，我们可以推测 $\gamma = 2\mu_0 H_c^2 / T_c^2$ 。
- 在零磁场， T_c 处的超导相变为连续相变；在 T_c 以下，由于磁场诱导的超导-正常相变为不连续相变。
- 我们可以进一步考察 $T = T_c$ 时，两相的热容之差：

$$C_{SC} - C_N = T \frac{\partial(s_{SC} - s_N)}{\partial T} = \mu_0 T \left(\frac{\partial H_{co}}{\partial T} \right)^2 + \mu_0 T H_{co} \frac{\partial^2 H_{co}}{\partial T^2} = \frac{1}{2} \mu_0 T \frac{\partial^2 (H_{co})^2}{\partial T^2}.$$

在临界温度 T_c 位置， $H_{co} = 0$ ，所以 $C_{SC} > C_N$ 。事实上，根据我们上面对电子热容的粗略估计，我们可以得到，在 $T = T_c$ 时， $C_{SC}/C_N \approx 3$ 。

- 我们还可以进一步比较超导相和正常相之间的 Gibbs 自由能。我们知道，在相边界的地方

$$G_S(H_{co}) = G_N(H_{co}),$$

而根据式 (16.4) 我们还有：

$$G_N(H_{co}) = G_N(H = 0) - \frac{1}{2} \mu_0 H_{co}^2, \quad G_S(H = 0) = G_S(H_{co})$$

所以，我们可以得到：

$$G_N(H = 0) - G_S(H = 0) = \frac{1}{2} \mu_0 H_{co}^2.$$

这说明，当外加磁场对应的能量大于超导配对能后，超导就会被破坏。

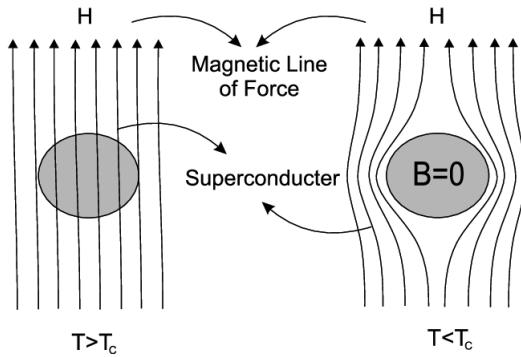


图 16.4: 超导体的抗磁性, 即 Meissner 效应。图片来自网络。

16.4 Ising 模型

Ising 模型是统计物理中最为重要的格点模型, 它描述了处于一定类型格点中的自旋系统。

Ising 模型哈密顿量可以表示为:

$$H = -J \sum_{\langle i,j \rangle} S_i S_j - B \sum_i S_i. \quad (16.6)$$

这里, S_i 为格点 i 上的自旋, 可以取向上或向下 (用 +1 和 -1 来表示), J 为交换耦合常数, 表示近邻自旋之间的交换相互作用。这里对 i 和 j 的求和只考虑最近邻自旋对, 而且每对自旋只取一次 (比如 $i < j$)。我们还可以考虑外加磁场对系统的影响, 因此加入了第二项。外加磁场使得 S_i 倾向于与与其平行排列。系统的维度以及格点的类型对系统的性质往往有重要的影响。如果耦合常数 $J > 0$, 那么近邻自旋倾向于平行排列; 如果 $J < 0$, 近邻自旋则倾向于反向排列。这里, 我们只考虑 $J > 0$ 的经典情况, 不考虑自旋之间的量子对易关系。

对于这样一个模型系统, 最自然的一个问题就是自旋的平均磁矩如何随温度发生变换。如果暂时不考虑外加磁场, 我们知道, 在零温下, 系统处于基态, 所有自旋倾向于沿相同方向排列。这时系统具有最低的能量。如果我们翻转所有自旋, 系统的能量不会发生变化。所以, 系统存在两个简并的基态。外加磁场可以导致简并的基态去简并。另外一个极限情况是 $T \rightarrow +\infty$, 这是自旋之间的交换相互作用不再重要 $k_B T \gg J$, 磁矩趋于随机分布, 平均磁矩趋于零。但仅仅知道这两个极限的情况还不够, 更重要的是系统如果从一个极限状态转变为另一个状态。是随温度变化平滑过渡, 还是通过相变过程? 这正是我们关心的问题。

对于这样的格点模型不涉及自旋格点数目的变化, 所以我们可以使用正则系综来考虑这

个问题。因此其配分函数可以表示为：

$$Z = \sum_{S_1} \cdots \sum_{S_N} = \sum_{\{S_i\}} \exp(-\beta H).$$

这里 N 为自旋数目。为了简化后面的式子，我们可以进一步定义

$$K = \beta J, \quad h = \beta B,$$

从而得到：

$$Z = \sum_{\{S_i\}} \exp \left(K \sum_{\langle i,j \rangle} S_i S_j + h \sum_i S_i \right).$$

16.4.1 独立自旋系统

我们前面考虑了由 N 个相互独立自旋组成的系统。每个自旋可以取向上和向下两个状态，分别用 $S_i = \pm 1$ 来表示。在外加磁场的作用下，与磁场方向平行的自旋能量为 $-B$ ，相反方向的能量为 B 。每个自旋的配分函数为

$$Z_1 = e^{-h} + e^h = 2 \cosh(h).$$

所以， N 个相互独立自旋组成的系统的配分函数就是

$$Z_N = Z_1^N = (2 \cosh(h))^N.$$

于是，其内能就是

$$U = -\frac{\partial \ln Z_N}{\partial \beta} = -NB \tanh(h).$$

我们还可以定义系统的总磁矩为

$$M = N_\uparrow - N_\downarrow.$$

这样的话，系统的内能可以用总磁矩来表示为

$$U = -BM.$$

比较以上两个内能表达式，我们得到系统的总磁矩为

$$M = N \tanh(h).$$

把每个自旋的平均磁矩定义为 m 我们有

$m = \tanh(h).$

我们在下面的讨论中将用到这个结果。

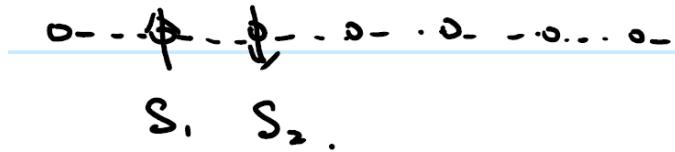


图 16.5: 一维 Ising 模型示意图。

16.4.2 一维 Ising 模型: 转移矩阵方法

我们详细考虑 $B = 0$ 的一维 Ising 模型, 详细验证边界条件不会影响我们的结果。考虑格点数目为 N 的一维模型, 我们分别采用开放和周期性边界条件来分析。

开边界条件: 若取开放边界条件, 则 S_1 和 S_N 无相互作用。

$$H_\Omega[J, H] = -J \sum_{i=1}^{N-1} S_i S_{i+1} \quad (16.7)$$

于是,

$$\begin{aligned} Z_\Omega[K, 0, N] &= \sum_{S_1=\pm 1} \cdots \sum_{S_N=\pm 1} \exp \left(K \sum_{i=1}^{N-1} S_i S_{i+1} \right) \\ &= \sum_{S_1=\pm 1} \sum_{\sigma_1=\pm 1} \cdots \sum_{\sigma_{N-1}=\pm 1} \exp \left(K \sum_{i=1}^{N-1} \sigma_i \right) \\ &= \sum_{S_1=\pm 1} \prod_{i=1}^{N-1} \sum_{\sigma_i=\pm 1} \exp(K\sigma_i) \\ &= 2(2 \cosh K)^{N-1} \end{aligned} \quad (16.8)$$

周期边界条件: 若取周期性边界条件, 则 S_1 和 S_N 有相互作用, 定义 $S_{N+1} = S_1$; 只要确定了 $S_1, \sigma_1, \dots, \sigma_{N-1}$, 系统的微观状态就完全确定了。于是,

$$\begin{aligned}
Z_{\Omega}[K, 0, N] &= \sum_{S_1=\pm 1} \cdots \sum_{S_N=\pm 1} \exp \left(K \sum_{i=1}^{N-1} S_i S_{i+1} + K S_N S_1 \right) \\
&= \sum_{S_1=\pm 1} \sum_{\sigma_1=\pm 1} \cdots \sum_{\sigma_{N-1}=\pm 1} \exp \left(K \sum_{i=1}^{N-1} \sigma_i \right) \\
&\times \exp(K \sigma_1 \sigma_2 \cdots \sigma_{N-1}) \\
&= 2 \sum_{\sigma_1=\pm 1} \cdots \sum_{\sigma_{N-1}=\pm 1} \exp \left(K \sum_{i=1}^{N-1} \sigma_i \right) \\
&\times \sum_{n=0}^{\infty} \frac{(K \sigma_1 \sigma_2 \cdots \sigma_{N-1})^n}{n!} \\
&= 2 \sum_{n=0}^{\infty} \frac{K^n}{n!} \prod_{i=1}^{N-1} \sum_{\sigma_i=\pm 1} \sigma_i^n \exp(K \sigma_i) \\
&= 2 \sum_{n=0}^{\infty} \frac{K^n}{n!} [e^K + (-1)^n e^{-K}]^{N-1} \\
&= (2 \cosh K)^N + (2 \sinh K)^N
\end{aligned} \tag{16.9}$$

我们期待，当 N 足够大时，具体采用哪种边界条件应该不会影响我们的结果。下面我们做一个验证。先考虑开放边界条件，系统的自由能为

$$\begin{aligned}
F_{\Omega}^O[K, 0, N] &= -k_B T \ln [2(2 \cosh K)^{N_1}] \\
&= -k_B T [N \ln 2 + (N-1) \ln(\cosh K)].
\end{aligned} \tag{16.10}$$

当 N 足够大时，

$$\lim_{N \rightarrow \infty} \frac{F_{\Omega}^O[K, 0, N]}{N} = -k_B T [\ln 2 + \ln(\cosh K)] \tag{16.11}$$

考虑周期性边界条件，则得到

$$\begin{aligned}
F_{\Omega}^P[K, 0, N] &= -k_B T \ln [(2 \cosh K)^N + (2 \sinh K)^N] \\
&= -k_B T \left\{ \ln(2 \cosh K)^N + \ln \left[1 + \left(\frac{\sinh K}{\cosh K} \right)^N \right] \right\}
\end{aligned}$$

当 N 足够大时，因为

$$\frac{\sinh K}{\cosh K} = \frac{e^K - e^{-K}}{e^K + e^{-K}} \in (0, 1), \tag{16.12}$$

所以

$$\lim_{N \rightarrow \infty} \frac{\ln \left[1 + \left(\frac{\sinh K}{\cosh K} \right)^N \right]}{N} = 0. \tag{16.13}$$

于是,

$$\lim_{N \rightarrow \infty} \frac{F_\Omega^P[K, 0, N]}{N} = -k_B T [\ln 2 + \ln(\cosh K)] \quad (16.14)$$

比较 (16.11) 和 (16.14) 证明了不同的边界条件在 $N \rightarrow \infty$ 时对热力学性质是没有影响的。

我们也可以使用转移矩阵方法来解决一维 Ising 模型。使用周期边界条件, N 格点模型的配分函数可以表示为

$$\begin{aligned} Z_\Omega[K, h, N] &= \sum_{S_1=\pm 1} \cdots \sum_{S_N=\pm 1} \exp \left(K \sum_{i=1}^N S_i S_{i+1} + h \sum_{i=1}^N S_i \right) \\ &= \sum_{S_1=\pm 1} \cdots \sum_{S_N=\pm 1} \exp \left[K S_1 S_2 + \frac{h}{2} (S_1 + S_2) \right] \\ &\quad \times \exp \left[K S_2 S_3 + \frac{h}{2} (S_2 + S_3) \right] \\ &\quad \cdots \exp \left[K S_N S_1 + \frac{h}{2} (S_N + S_1) \right] \end{aligned} \quad (16.15)$$

这里的关键是, 我们在第二个等式中把配分函数写成了 N 个自旋对 (S_i, S_{i+1}) 的贡献, 这里 $i = 1, \dots, N$ 。第 i 个自旋对的贡献为:

$$\Gamma_{S_i, S_{i+1}} = \exp \left[K S_i S_{i+1} + \frac{h}{2} (S_i + S_{i+1}) \right].$$

我们可以把上式的 Γ 理解为一个矩阵, 而 $\Gamma_{S_i, S_{i+1}}$ 为其矩阵元。因为每个自旋可以取 ± 1 , 所以 Γ 为 2×2 的矩阵

$$\Gamma = \begin{bmatrix} \Gamma_{1,1} & \Gamma_{1,-1} \\ \Gamma_{-1,1} & \Gamma_{-1,-1} \end{bmatrix} = \begin{bmatrix} e^{K+h} & e^{-K} \\ e^{-K} & e^{K-h} \end{bmatrix}. \quad (16.16)$$

这样的话, 配分函数就可以表示为 Γ 矩阵的乘积形式

$$Z_\Omega[K, h, N] = \sum_{S_1=\pm 1} \cdots \sum_{S_N=\pm 1} \Gamma_{S_1, S_2} \Gamma_{S_2, S_3} \cdots \Gamma_{S_N, S_1} = \text{Tr}(\Gamma^N).$$

Γ 是一个实对称矩阵, 一定可以对角化, 其本征值为

$$\lambda_{1,2} = e^K \left(\cosh h \pm \sqrt{\sinh^2 h + e^{-4K}} \right). \quad (16.17)$$

我们可以规定 $\lambda_1 > \lambda_2$ 。对角化之后再计算迹就十分容易

$$Z_\Omega[K, h, N] = \text{Tr}(\Gamma^N) = \lambda_1^N + \lambda_2^N. \quad (16.18)$$

得到配分函数后，我们就可以利用它来得到系统的热力学性质

$$F_\Omega[K, h, N] = -k_B T \ln Z = -k_B T \left\{ \ln \lambda_1^N + \ln [1 + (\lambda_2/\lambda_1)^N] \right\}.$$

在热力学极限 $N \rightarrow \infty$ 下，

$$\lambda_2/\lambda_1 < 1 \quad \Rightarrow \quad \lim_{N \rightarrow \infty} \frac{\ln [1 + (\lambda_2/\lambda_1)^N]}{N} = 0, \quad (16.19)$$

所以每个自旋的平均自由能为

$$\begin{aligned} f_b[K, h, N] &= \lim_{N \rightarrow \infty} \frac{F_\Omega[K, h, N]}{N} \\ &= -k_B T \ln \lambda_1 \\ &= -k_B T \ln \left[e^K \left(\cosh h + \sqrt{\sinh^2 h + e^{-4K}} \right) \right] \\ &= -J - k_B T \ln \left[\cosh h + \sqrt{\sinh^2 h + e^{-4K}} \right]. \end{aligned} \quad (16.20)$$

当 $T > 0$ 时， $f_b[K, h, N]$ 总是解析的，所以一维 Ising 模型中不存在有限温度相变。这一点也可以通过其平均磁矩体现出来

$$m = \lim_{N \rightarrow \infty} -\frac{1}{N} \frac{\partial F}{\partial B} = \frac{\sinh(\beta B)}{\sqrt{\sinh^2(\beta B) + e^{-4K}}} \quad (16.21)$$

m 在 $B \rightarrow 0$ 时总是为零：

$$\lim_{B \rightarrow 0} m = 0.$$

系统没有自发磁化。外磁场下的磁化率为：

$$\chi_T = \lim_{B \rightarrow 0} \frac{\partial m}{\partial B} = \beta \exp(2\beta J). \quad (16.22)$$

相反，如果考虑 $T = 0$ 的情况，我们有：

$$\lim_{T \rightarrow 0} F_\Omega[K, h, N] = \lim_{T \rightarrow 0} -k_B T \ln \left(\lambda_1^N + \lambda_2^N \right) = -N(J + |B|) \quad (16.23)$$

先考虑热力学极限 $N \rightarrow \infty$ ，再考虑 $B \rightarrow 0$ 的极限，得到：

$$m = \lim_{N \rightarrow \infty} -\frac{1}{N} \frac{\partial F}{\partial B} = \begin{cases} 1, & B > 0 \\ -1, & B < 0 \end{cases}. \quad (16.24)$$

所以，系统中出现宏观磁性。

16.4.3 二维 Ising 模型

二维 Ising 模型的严格求解过程十分复杂，这里我们无法给出详细过程。历史上，1941 年，Kramers 和 Wannier 在不直接求解的情况下，利用对称性确定了相变的温度。1944 年，Onsager 给出了严格解。Onsager 的解在统计物理学中占据重要的地位。1949 年，Onsager 在学术会议上报告了磁化强度的临界指数，但是他没有发表计算细节。1952 年，杨振宁给出了详细的计算过程。

为了进一步理解 Onsager 结果的重要性。我们不妨考虑一个有限大小的二维 Ising 模型。系统的配分函数为有限数目的指数函数之和。因为指数函数 $e^{-\beta E}$ 是一个解析函数，它本身及它的各阶微分都是连续的。有限数目的指数函数之和得到的函数一定也是解析函数。因为相变表现为热力学量随外界参数的不连续变化。如果配分函数是解析的，那么可以推断系统的相应热力学量随外界参数是连续变化的，因此不会发生相变。也就是说，**从严格意义上讲，有限大小的系统不会发生相变**。当系统从有限大小过渡到无限大小时情况如何呢？如果没有 Onsager 的严格解是很难判断的。

16.4.4 平均场理论

16.4.4.1 1. $B = 0$

考虑高维 Ising 模型的一个简单的方式是平均场理论。它的基本思想是考虑系统中某一个自旋时，把周围其他所有自旋对这个自旋的作用用一个平均场来描述。为此，我们把每个自旋写为其平均值与偏离之和的形式：

$$S_i = \bar{S}_i + \Delta S_i$$

把它带入 Ising 模型的哈密顿量中，并且忽略 ΔS_i 的二次项，我们得到平均场哈密顿量

$$H_{mf} = \frac{1}{2} N n J \bar{S}^2 - J \bar{S} n \sum_i S_i$$

上式第一项为一个常量，而第二项的形式与独立自旋模型类似。这样，对于第 i 个自旋，我们可以用下面的平均场哈密顿量来描述

$$H_{i,mf} = -J n \bar{S} S_i.$$

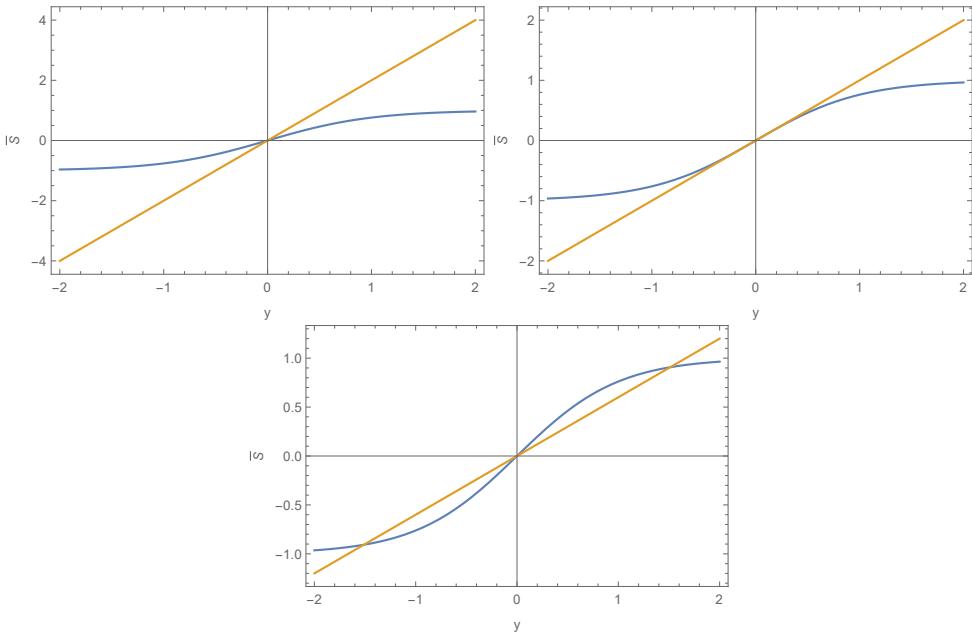


图 16.6: 不同温度下, 方程 16.26 左右两边函数曲线。

这里, \bar{S} 为与 i 近邻自旋的平均磁矩, 而 n 为最近邻自旋的个数。经过平均场近似后, Ising 模型又被简化为独立自旋的形式。所以, 每个自旋的平均磁矩可以表示为:

$$\bar{S}_i = \tanh(\beta n J \bar{S}).$$

因为 S_i 是我们在系统中随机选择的一个自旋, 所以上式应该对任何一个自旋都成立。也就是说, 我们有

$$\bar{S}_i = \bar{S}. \quad (16.25)$$

这样, 我们就最终得到了平均场的结果

$$\boxed{\bar{S} = \tanh(\beta n J \bar{S})}. \quad (16.26)$$

这个方程的两边都有 \bar{S} , 我们需要对其进行自洽迭代求解, 具体流程是: 首先给定一个初始的 \bar{S} , 将它代入到上式等号右边, 我们就会得到一个新的 \bar{S} ; 再将其代入方程右边, 再次得到一个新的 \bar{S} 。如此反复, 直到 \bar{S} 不再变化为止。

式 (??) 的定性行为可以通过做图来进行分析。几种不同情况如图 16.6 所示。从图形中我们可以看到, 系统存在一个临界温度

$$k_B T_c = nJ. \quad (16.27)$$

这时, 两条曲线在 $y = 0$ 的位置正好相切。我们可以分下面几种情况讨论:

- $T > T_c$: 两条曲线只在 $y = 0$, 即 $\bar{S} = 0$ 的位置相交。这时, 平均磁矩为零。

- $T < T_c$: 除了零点外，两条曲线还在另外两个点相交。这两个点对应的平均磁矩非零，而且正好相反。这样的解必然成对出现，这是模型的对称性所决定的：如果我们把系统的所有自旋方向反向，那么系统的哈密顿量不会发生变化。如果我们从 $T > T_c$ 开始逐步降低系统温度的话，系统在 T_c 会发生相变，选择其中一个解。系统选择哪个解是不确定的。因为这两个解完全等价。确定的是，系统一旦选择了其中一个解，其状态的对称性就会降低，低于系统哈密顿量的对称性。这个过程称为**自发对称性破缺** (spontaneous symmetry breaking)。
- 这里需要注意的是，在 $T < T_c$ 时， $\bar{S} = 0$ 的解不是一个稳定解。因为如果 \bar{S} 稍微偏离零点，将 \bar{S} 带入式16.26，我们得到 $\tanh \beta n J \bar{S} > \bar{S}$ （不妨取 $\bar{S} > 0$, $\bar{S} < 0$ 的情况可做类似分析）。这说明，我们得到的新 \bar{S} 比原来的数值更大。再做一次迭代的话，我们会得到更大的 \bar{S} 。这样，系统就会越来越偏离 $\bar{S} = 0$ 的解，最后会稳定到 $\bar{S} > 0$ 的解。把类似的分析应用到其他两个解上，我们发现它们对于 \bar{S} 的微小涨落是稳定的。
- 我们也可以从自由能的角度来理解这个解的不稳定性：它对应自由能的一个局域极大值。为说明这个结论，我们把系统的自由能

$$F = -k_B T \ln Z = \frac{1}{2} N n J \bar{S}^2 - N k_B T \ln(2 \cosh \beta n J \bar{S})$$

在 $\bar{S} = 0$ 附近做展开

$$\begin{aligned} F &= \frac{1}{2} N n J \bar{S}^2 - N k_B T \ln(2 \cosh \beta n J \bar{S}) \\ &\approx \frac{1}{2} N n J \bar{S}^2 - N k_B T \left\{ \ln \left[2 \left(1 + \frac{1}{2} (\beta n J \bar{S})^2 + \frac{1}{4!} (\beta n J \bar{S})^4 \right) \right] \right\} \\ &\approx -N k_B T \ln 2 + \frac{1}{2} N k_B T_c \left(1 - \frac{T_c}{T} \right) \bar{S}^2 + \frac{1}{12} N k_B T_c \left(\frac{T_c}{T} \right)^3 \bar{S}^4. \end{aligned} \quad (16.28)$$

从上式可以看出，当 $T < T_c$ 时， \bar{S}^2 项的系数为负；当 $T > T_c$ 时变为正。因此， $m = 0$ 这个解在低于 T_c 时对应自由能的局域极大值，不稳定；高于 T_c 时变为极小值，稳定，如下图16.7所示。

如果我们考察自旋平均值或平均磁矩随温度的变化的话，可以得到图16.8所示的结果。可见，在临界温度以上，系统平均磁矩为零。当温度缓慢降低到 T_c 以下时，系统会发生对称性自发破缺，选择图中所示的两个状态之一。这是一个连续相变的过程，因为系统的平均磁矩（自由能的一阶微分）随温度变化是连续的。

平均场理论给出的结果是，任意维度的 Ising 模型都会发生相变。相变温度为

$$T_c = nJ/k_B.$$

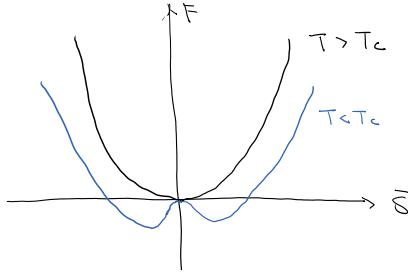


图 16.7: 在高于和低于 T_c 时，自由能 F 随平均磁矩 \bar{S} 的变化曲线。可以看出 $m = 0$ 这个解在两种情况下分别对应自由能的最低和局部最高值。因此， $T < T_c$ 时， $m = 0$ 为非稳定解，系统发生自发磁化。这时， F 有两个能量最低值， \bar{S} 取值大小相同，方向相反。它们对应图16.6和16.8中的两个非零解。

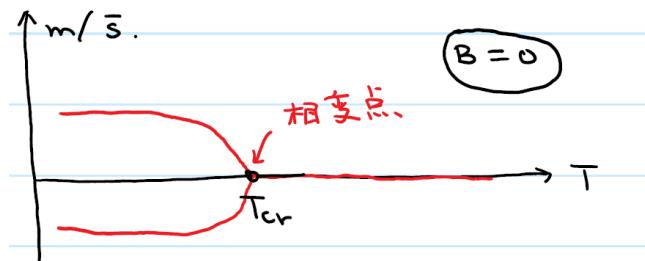


图 16.8: 自旋平均磁矩随温度的变化关系。

可见相变温度与最近邻自旋数目和交换能大小成正比。但通过上面的分析我们发现，一维的 Ising 模型是不会发生相变的。因此，对于一维体系，平均场理论的预言是错误的。平均场理论失效的原因是最近邻自旋数太少。对于一维系统 $n = 2$ 。这时，自旋的涨落是比较重要的，铁磁序被涨落压制，所以不会发生相变。对于二维系统，平均场理论能够正确预言相变的存在，但给出的临界温度不对。随着系统维度的提高，最近邻自旋的数目越来越多，平均场理论的预言就会越来越准确。

16.4.4.2 2. $B \neq 0$

前面的分析中我们没有考虑外加磁场的影响。如果系统处于某一个外加磁场 B 中，那么两种自旋取向对应的能量就会发生变化。通过相同的分析我们可以得到：

$$\bar{S} = \tanh(\beta n J \bar{S} + \beta B) \quad (16.29)$$

图16.9给出了存在外加磁场时，在 T 大于，小于或等于 T_c 时的情况。当 $T \gg T_c$ 时，系统的平均磁矩变为零。但当 T 不是远大于 T_c 时，由于外场的作用，系统会有一个非零磁矩。随着温度的降低，这个磁矩越来越大。与 $B = 0$ 的情况不同的是，这个变化过程是平滑的，没有相变，如图16.10所示。

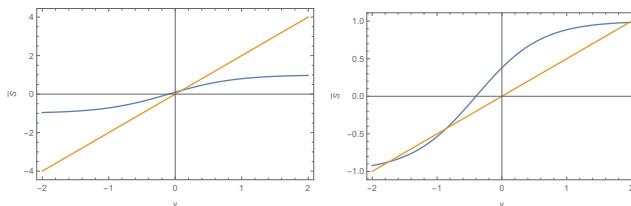


图 16.9: 外加磁场 $B = 0.2$ 时, 不同温度下, 方程 (16.29) 左右两边函数曲线。

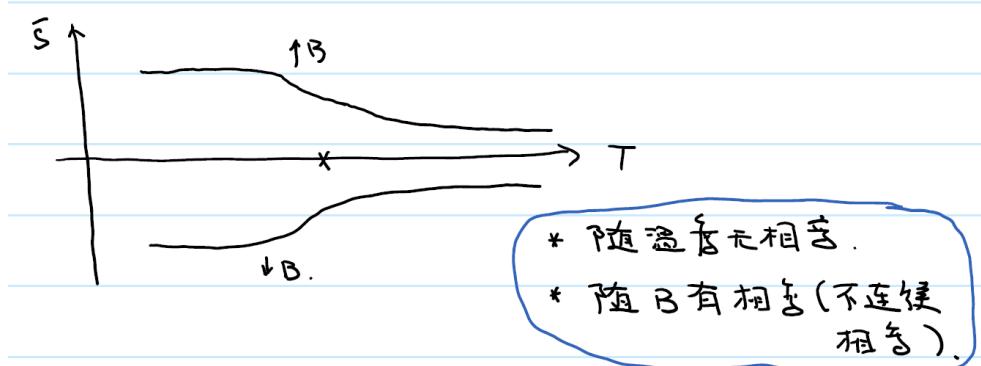


图 16.10: 外加磁场 $B = 0.2$ 下, 系统平均磁矩随温度的变化。

需要注意的是, 当存在外加磁场时, 在临界温度以下, 原本等价的两个解不再等价, 如图 16.9 右图所示。这时, m 与磁场方向平行的解自由能更低, 更稳定。而 m 与磁场方向相反的解是一个亚稳态, 具有相对较高的自由能。另外一个由 $m = 0$ 演化来的解仍是一个非稳解。这样, 当我们固定温度, 改变磁场 B , 从 $B > 0$ 向 $B < 0$ 过渡时系统平均磁矩会发生跳变, 因此会发生磁场诱导的相变, 这个相变是一级相变。

16.5 Landau 相变理论

前面在分析 Ising 模型的时候, 我们基于平均场近似得到了系统的自由能近似表达式。然后我们对自由能相对于系统的磁化 m 做展开, 得到了式 (16.28)。我们可以分析在临界温度上下, 系统自由能对 m 的依赖关系发生如何变化, 从而可以对相变的过程有一个理解。

Landau 相变理论正是对上述针对 Ising 模型分析的推广, 它是一个完全唯象的平均场理论。我们考虑正则系综, 所以选择 T 和 V 为系统参量。Landau 理论的核心物理量为 Landau 自由能 $F_L(T, V; m)$, 其中的参量除了 T, V 外, 还包含所谓的序参量 (**order parameter**) m 。序参量在不同的相取值不同, 相变的过程可以用序参量的变化来描述。不同类型的相变具有不同的序参量。在有的相变中, 序参量非常显而易见。比如, 在 Ising 模型中, 它对应系统的磁矩 m ; 在气液相变中, 两相的体积差为其序参量。但在有的相变中, 系统的序参量却不是那

么明显。比如，在金属-超导相变、Bose-Einstein 凝聚中，相变的序参量为其基态的宏观波函数，或者更准确的说为其单粒子密度矩阵的非对角长程序。Landau 自由能是系统序参量的函数，我们可以通过改变序参量来使 Landau 自由能最小化，从而找到系统的平衡态。平衡态时的序参量用 m_0 表示，这时的 Landau 自由能就对应系统的 Helmholtz 自由能

$$F(T, V) = F_L(T, V; m_0) \leq F_L(T, V; m).$$

为了确定 m_0 ，通常将 Landau 自由能 $F_L(T, V; m)$ 相对于 m 做展开。这个展开要求 F_L 相对于 m 的变化是连续平滑的；而且要求 m 是一个小量。做展开时，我们需要考虑系统本身所具有的对称性。这些对称性决定了展开式中 m 对 F_L 的依赖形式。比如，没有外加磁场时，Ising 模型具有的自旋对称性决定了 F_L 只依赖 m 的偶次幂。这代表了很大的一类系统。下面我们以这类系统为例，把 F_L 做展开至 4 阶得到²：

$$F_L(T, V; m) = F_L(T, V; 0) + a(T)m^2 + b(T)m^4 + \dots \quad (16.30)$$

这些展开系数都依赖于系统的温度（暂时不考虑体积的变化）。当系统处于平衡状态时，我们要求

$$\frac{\partial F_L}{\partial m} = 0.$$

为了考虑相应的位置为极大还是极小值，我们还需要进一步考虑其二次微分。由此，我们可以得到 m 的可能取值。这里，需要分两种情况讨论。

- 如果 $a(T) > 0$ ，我们得到唯一的一个解为 $m = 0$ ，即序参量为零。这时， $\partial^2 F / \partial m^2 = 2a > 0$ ， $m = 0$ 为一个稳定解。
- 如果 $a(T) < 0$ ，我们得到三个解 $m = 0$ 以及 $m = \pm\sqrt{-a/2b}$ 。将这三个解代入到 F_L 的表达式，我们得到相应的 Landau 自由能。我们可以验证， $m = 0$ 对应 Landau 自由能的局部最大值，因此是一个不稳定的解；而另外两个解为稳定解，它们给出相同的 F_L 。这些结果可以看做是对 Ising 模型讨论的一个推广。 $a(T)$ 从正到负的转变对应一个相变的过程，由于在此过程中 m 是连续变化的，所以它是一个连续相变。这些正好对应图 (16.7) 和 (16.8) 的情况。
- 不同温度下的 Helmholtz 自由能可以表示为：

$$F(T, V) = \begin{cases} F(T, V; m = 0) & T > T_c \\ F(T, V; m = 0) - \frac{a^2}{4b} & T < T_c. \end{cases}$$

²具体展开的几阶依赖于我们考虑的问题。

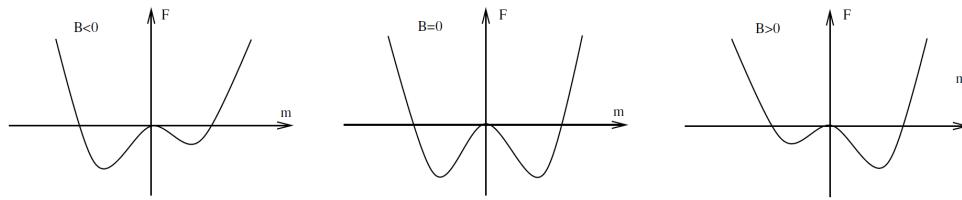


图 16.11: 在外加磁场下, Landau 自由能随 m 的变化。这时, 原来简并的两个解去简并。一个变成亚稳态解, 一个变成稳态解。

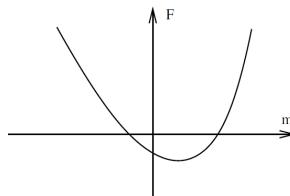


图 16.12: 进一步提高温度, 原来的三个解变成了一个稳态解。

可以验证, 在相变点附近, $\partial F / \partial T$ 是连续的, 但 $\partial^2 F / \partial T^2$ 不是连续的。因此, 这里我们描述的是一个二级相变。

当存在外加磁场时, Ising 模型的哈密顿量不再具有自旋翻转对称性。这时, 在对 Landau 自由能进行展开时我们就需要考虑 m 的奇次幂, 做如下展开:

$$F_L(T, V; m) = F_L(T, V; 0) + \alpha(B, T)m + a(T)m^2 + \beta(B, T)m^3 + b(T)m^4 + \dots \quad (16.31)$$

我们可以再次重复上面对 $B = 0$ 情况的分析。当 B 较小时, 我们得到的三个 $B = 0$ 的解会发生小的变化, 但它们都还存在。这时, 从 $m = 0$ 演化过来的解仍然是个非稳解; 而刚才简并的两个解由于磁场作用不再简并。其中一个 Landau 自由能最小的解为稳定解, 另外一个变成了亚稳态的解。它和稳态解之间被一个势垒隔开(图16.11)。在给定磁场, 增加温度的话, 解的个数会进一步减小, 最后只剩下一个稳态解(图16.12)。

可以进一步验证, 基于以上的 Landau 理论, 我们也可以得到平均场近似下得到的结果: 在低温下, m 随外加磁场的方向变化会发生跳变, 即发生不连续相变。但是, 在给定外加磁场时, m 随温度的变化是平滑连续的, 不再发生 $B = 0$ 时的相变了。

通过以上分析, 我们看到, 基于唯象的 Landau 相变理论分析得到的结果和平均场近似基本一致。实际上, Landau 的理论是对我们前面介绍的平均场理论的一个推广。Ginzburg 之后对 Landau 理论做了进一步的扩展, 将 m 的空间变化和涨落效应包含了进来。后人将这一套理论统称为 Landau-Ginzburg 相变理论。当然, 作为平均场理论, Landau-Ginzburg 理论也存在

	严格解	平均场
临界温度	$T_c = \frac{2J}{k_B \ln(1+\sqrt{2})}$	$\frac{4J}{k_B}$
热容 C	对数发散	有限跃变
磁矩 $m \sim (T_c - T)^\beta$	1/8	1/2
磁化率 $\chi_T \sim T_c - T ^{-\gamma}$	7/4	1
磁矩 $m \sim B^{1/\delta}$	尚未直接解出	3

表 16.1: 二维 Ising 模型的严格解和平均场结果对比。

类似的问题。它们无法准确的描述很多系统的相变。

16.6 临界指数

相变临界点附近物理量随外界条件，比如温度、磁场，的变化可以用临界指数来描述。这些指数可以用来表征相变的基本特征。很多看似毫无关联的系统可能具有相同的临界指数，这表现出相变行为的普适性（universality）。下面我们用平均场理论来分析 Ising 模型和气液相变。

16.6.1 Ising 模型

对于 Ising 模型，利用平均场方法，我们可以考虑磁矩、热容、磁化率等量随温度的变化规律。我们分析的出发点为平均场方程：

$$\tanh \beta(k_B T_c m + B) = m,$$

或者

$$k_B T \text{ArcTanh}(m) = k_B T_c m + B.$$

在临界点附近，我们可以利用 $\tanh(x) = x - \frac{1}{3}x^3 + \dots$ 或者 $\text{ArcTanh}(x) = x + \frac{1}{3}x^3 + \dots$ 来进行分析。表 (16.1) 给出了平均场理论与二维 Ising 模型严格解的结果对比。

16.6.1.1 磁矩

首先，当 $B = 0$, $T < T_c$ 时，系统的平均磁矩随温度的变化关系为

$$\bar{S} \propto \left(1 - \frac{T}{T_c}\right)^{1/2}.$$

这个指数通常用 β 表示（注意与 $k_B T = \beta^{-1}$ 中的 β 相区分）。可见，平均场理论给出的临界指数与系统维度无关， $\beta = 1/2$ 。

其次，在临界温度 T_c ，当磁场逐渐趋于零时，系统的磁矩满足

$$\bar{S} \sim B^{1/\delta}.$$

平均场理论给出指数 $\delta = 3$ 。

16.6.1.2 热容

临界温度以下 $T < T_c$ ，每个自旋的平均内能可以表示为：

$$U = -\frac{1}{2}nJ\bar{S}^2 = -\frac{1}{2}k_B T_c \bar{S}^2.$$

于是，我们得到

$$C = \frac{3}{2}Nk_B, \quad (T < T_c).$$

在临界温度以上， $\bar{S} = 0$ ，所以 $C = 0$ 。因此，平均场理论预言在临界温度，系统热容存在一个跳变。一个约定的习惯是把这种情况下的临界指数表示为 $\alpha = 0, \text{jump}$ 。

16.6.1.3 磁化率

系统的磁化率

$$\chi = \left. \frac{\partial \bar{S}}{\partial B} \right|_{B=0}$$

随温度的依赖关系可以表示为

$$\chi(T, 0) \propto A_{\pm} |T - T_c|^{-\gamma_{\pm}}$$

可以验证，在平均场理论下， $\gamma_{\pm} = 1$ ， $A_+ = 2A_-$ 。这里的 + 和 - 分别表示从 T_c 以上和以下趋近 T_c 。

16.6.2 van der Waals

除了 Ising 模型外，我们前面还使用 van der Waals 方程考虑了气液相变。在约化变量表示下，我们可以把 van der Waals 方程表示为

$$\tilde{P} + \frac{3}{\tilde{V}^2} = \frac{8\tilde{T}}{3\tilde{V} - 1}. \quad (16.32)$$

其中, $\tilde{T} = T/T_c$, $\tilde{P} = P/P_c$, $\tilde{V} = V/V_c$ 。与 Ising 模型类似, 我们也可以问, 在临界温度附近, 气体的不同物理量以何种方式趋于临界点, 它们的温度依赖如何? 我们可以定义类似的临界指数。

16.6.2.1 体积

首先, 我们可以考虑在临界点附近气相和液相两种状态的体积差如何随温度发生变化。在 $T < T_c$ 或 $\tilde{T} < 1$ 时, 我们有

$$\tilde{P} = \frac{8\tilde{T}}{3\tilde{V}_l - 1} - \frac{3}{\tilde{V}_l^2} == \frac{8\tilde{T}}{3\tilde{V}_g - 1} - \frac{3}{\tilde{V}_g^2}.$$

根据上式我们可以得出

$$\tilde{T} = \frac{(3\tilde{V}_l - 1)(3\tilde{V}_g - 1)(\tilde{V}_l + \tilde{V}_g)}{8\tilde{V}_g^2\tilde{V}_l^2}.$$

可见, 当 $\tilde{V}_g \rightarrow 1$, $\tilde{V}_l \rightarrow 1$ 时, $\tilde{T} \rightarrow 1$, 这符合我们的预期。在临界点附近, 我们可以把 $m = \tilde{V}_g - \tilde{V}_l$ 当成一个小量, 假设 $\tilde{V}_g = 1 + m/2$, $\tilde{V}_l = 1 - m/2$, 我们有

$$\tilde{T} \approx 1 - \frac{1}{16}(\tilde{V}_g - \tilde{V}_l)^2.$$

整理后得

$$m = \tilde{V}_g - \tilde{V}_l \sim (T_c - T)^{1/2}. \quad (16.33)$$

这里的临界指数为 $\beta = 1/2$ 。

与 Ising 模型类似, 我们也可以考虑在临界温度 T_c 系统的压强对体积的依赖关系。我们知道, 在临界点,

$$\frac{\partial P}{\partial V} = \frac{\partial^2 P}{\partial V^2} = 0.$$

所以, 如果要做展开的话, 最低阶项为

$$P - P_c \sim (V - V_c)^3. \quad (16.34)$$

或者

$$m \sim (P - P_c)^{1/3} \quad (16.35)$$

所以临界指数 $\delta = 3$ 。

16.6.2.2 等温压缩率

下面我们考虑气体的等温压缩率

$$\kappa_T = \frac{1}{V} \frac{\partial V}{\partial P} \Big|_T.$$

在临界点附近， κ_T 的温度依赖关系。我们知道，在临界点的位置，

$$\frac{\partial P}{\partial V} \Big|_{T_c} = 0,$$

在临界点附近展开有

$$\frac{\partial P}{\partial V} \Big|_{T, V=V_c} = -a(T - T_c) + \dots,$$

所以，等温压缩率

$$\kappa_T \sim (T - T_c)^{-1}. \quad (16.36)$$

所以， $\gamma = 1$ 。

以上是 van der Waals 方程给出的气体在临界点附近的临界指数。这些指数可以和平均场理论给出的 Ising 模型的临界指数一一对应起来。其根本原因是，van der Waals 方程也是基于平均场理论得到的。

平均场理论，特别是 Landau 相变理论，给出了描述相变的一个理论框架。但基于平均场理论得到的结果只有在高维度下才相对准确。其预言与实际相变系统的测量结果往往有比较大的偏差。这就要求我们进一步寻找能够更加准确描述相变过程的理论。标度假设、重整化群的理论方法正是在这样的背景下诞生的。

第 17 章 布朗运动

17.1 导论: 为什么研究涨落?

我们知道处于平衡态的系统中, 物理量的涨落正比于 $1/\sqrt{N}$, 其中 N 为系统中的粒子数。因此, 在宏观热力学系统 ($N \sim 10^{23}$) 中, 涨落可以忽略。基于这一原因, 我们前面课程中重点讨论的是各个物理量的统计平均性质。在下面一部分的学习中, 我们需要把涨落考虑进来。原因至少有以下两个方面:

- 有的时候我们关心的并不是整个宏观系统, 而是宏观系统的某一小部分。当我们关心的自由度数目不是足够大时, 系统与库之间的相互作用导致的系统物理量涨落可以很容易被观测到。布朗运动就是这样一个例子。这将是本章的主要内容。
- 物理量的涨落很多时候是和它的不可逆或者非平衡态性质联系在一起的。这是我们研究涨落的另一个重要原因。后面, 我们会学习到涨落耗散定理 (fluctuation-dissipation theorem)。它告诉我们系统的涨落与耗散之间是密切联系的。知道了其一就可以获取另一个的信息。而研究系统的非平衡态性质是非常困难的一件事情, 需要使用一些更加复杂的理论方法。通过对平衡态系统涨落的研究, 我们可以得到其相关的非平衡态性质。这是非常有价值的。

17.2 布朗运动

布朗运动是苏格兰的植物学家布朗在 1827 年通过显微镜观察悬浮在水中的花粉颗粒的运动时发现的。我们现在都知道由于溶液分子对花粉颗粒的随机碰撞而导致其在溶液中的无规则运动。但布朗的发现在当时是很出乎意料也令人费解的, 因为在那个年代, 人们对物质的结构还没有微观的认识。比如对于物质是否由原子组成, 在许多著名科学家之间还存在很大的争议。有关布朗运动的定量理论是由爱因斯坦在他 1905 年的一篇著名论文中提出来的。后来, Langevin, Smoluchowski 等人又对布朗运动的理论做了进一步的发展。

爱因斯坦 1905 年的 5 篇论文的题目如下

- Investigations on the Theory of the Brownian Movement
- On the Motion of Small Particles Suspended in Liquids at Rest Required by the Molecular-Kinetic Theory of Heat
- On the Electrodynamics of Moving Bodies
- Does the Inertia of a Body Depend on Its Energy Content?
- On a Heuristic Point of View Concerning the Production and Transformation of Light

17.2.1 随机行走模型

我们首先从随机行走的角度去考虑这个问题。为了简化分析，考虑一维的情况。假设花粉颗粒沿着 x 方向运动，每次移动的距离为 δx ，时间步长为 δt 。在某一时刻，它可能向左 ($-x$ 方向) 移动，也可能向右 (x 方向) 移动，两者的概率相同，都是 $1/2$ 。经过 N 步之后，颗粒的总位移可以写成每一步位移之和 $Y_N = \sum_{i=1}^N X_i$ 。粒子在 N 步中，向右移动 n 步的概率为：

$$P_{Y_N}(n) = \frac{N!}{n!(N-n)!} \left(\frac{1}{2}\right)^N. \quad (17.1)$$

我们可以根据这个概率分布来计算 Y_N 的均值和方差¹，得到：

$$\langle Y_N \rangle = 0, \quad \langle \Delta Y_N^2 \rangle = N\delta l^2. \quad (17.2)$$

二项式分布的均值和方差：考虑二项式分布

$$P_{Y_N}(n) = \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n}. \quad (17.3)$$

为计算其均值和方差，我们首先考虑：

$$\begin{aligned} \langle n \rangle &= \sum_{n=0}^N n P_{Y_N}(n) = p \sum_{n=1}^N \frac{N!}{(n-1)!(N-n)!} p^{n-1} (1-p)^{N-n} \\ &= pN \sum_{n=1}^N \frac{(N-1)!}{(n-1)!(N-1-(n-1))!} p^{n-1} (1-p)^{N-1-(n-1)} \\ &= pN. \end{aligned} \quad (17.4)$$

¹ 我们还可以采用附录23中的方法，考虑其特征函数，然后由特征函数确定均值和方差。

$$\begin{aligned}
\langle n(n-1) \rangle &= \sum_{n=0}^N n(n-1) P_{Y_N}(n) \\
&= p^2 N(N-1) \sum_{n=2}^N \frac{(N-2)!}{(n-2)!(N-2-(n-2))!} \\
&\quad \times p^{n-2} (1-p)^{N-2-(n-2)} \\
&= N(N-1)p^2.
\end{aligned} \tag{17.5}$$

根据上面的两个结果很容易得到：

$$\langle Y_N \rangle = (2p-1)N\delta x, \tag{17.6}$$

$$\langle \Delta Y_N^2 \rangle = 4(\langle n^2 \rangle - \langle n \rangle^2) = 4p(1-p)N\delta x^2. \tag{17.7}$$

根据**中心极限定理**，当 N 足够大的时候，二项式分布可以近似为具有相同平均值和方差的高斯分布

$$P_{Y_N}(y, t) = \frac{1}{\sqrt{2\pi N\delta l^2}} \exp\left[-\frac{y^2}{2N\delta l^2}\right] \equiv \frac{1}{\sqrt{4\pi D t}} \exp\left[-\frac{y^2}{4Dt}\right]. \tag{17.8}$$

通过上式，我们定义了扩散系数

$$D \equiv \delta l^2 / 2\delta t, \tag{17.9}$$

这里 $t = N\delta t$, δt 可以理解为花粉颗粒的平均自由运动时间或者叫驰豫时间，而 δl 可以认为是其平均自由运动的距离或者叫平均自由程。根据这个定义，我们可以把

$$\langle \Delta Y_N^2 \rangle = 2Dt.$$

我们发现，布朗运动粒子的位移服从高斯分布，而分布的方差与时间和扩散系数成正比。相应的，其标准差正比于总时间 t 的平方根。这是布朗运动与普通力学体系下的运动不同的地方。我们知道服从牛顿力学的粒子平均运动距离与时间 t 是成正比的，通常称为“弹道”过程。相应的，布朗粒子的运动称为**扩散过程**。

17.2.2 粒子扩散

为了更好的理解布朗粒子的扩散行为，我们下面从另一个角度分析其运动。想象我们在某一位置放置了大量的粒子（系综），每个粒子都开始做布朗运动。这些粒子从初始位置出发会向空间各个方向扩散。如果从扩散的角度来分析这个问题，最后我们会发现这一大量粒子

组成的系综满足的分布也是高斯分布。

我们的出发点是 Fick 定律：

$$\vec{j}(\vec{r}, t) = -D \nabla n(\vec{r}, t). \quad (17.10)$$

这里， $n(\vec{r}, t)$ 是粒子满足的密度分布， $\vec{j}(\vec{r}, t)$ 为由于扩散导致的粒子流密度，而 D 为扩散系数。另外，由于粒子数守恒，我们还有连续性方程：

$$\frac{\partial n(\vec{r}, t)}{\partial t} + \nabla \cdot \vec{j}(\vec{r}, t) = 0. \quad (17.11)$$

这两个方程联立可以得到密度分布 $n(\vec{r}, t)$ 所满足的扩散方程：

$$\boxed{\frac{\partial n(\vec{r}, t)}{\partial t} = D \nabla^2 n(\vec{r}, t)}. \quad (17.12)$$

为了和上面有关布朗运动的分析对比，我们下面只考虑一维的情况。扩散方程的解可以通过 Fourier 变换的方式来得到：

$$n(y, t) = \frac{N}{\sqrt{4\pi Dt}} \exp\left(-\frac{y^2}{4Dt}\right), \quad (17.13)$$

其中 N 为总粒子数。如果我们考虑一个粒子的情况，上式中的 $n(y, t)$ 就可以理解为概率密度分布，它满足 $\int dy n(y, t) = 1$ 。这和我们从随机行走的角度得到的布朗粒子满足的概率密度分布是一样的，同时我们上面定义的 $D = \delta l^2 / 2\delta t$ 的确是扩散系数。

17.2.3 朗之万方程

布朗运动粒子的动力学行为也可以用朗之万方程来描述。单粒子、单自由度的朗之万方程可以写为：

$$\boxed{m\ddot{x} = -V'(x) - m\gamma\dot{x} + f(t)}. \quad (17.14)$$

这里， m 为粒子的质量， γ 为摩擦系数， $V'(x)$ 为粒子所处的势场贡献的力，而 $f(t)$ 为噪声或者随机力 $\langle f(t) \rangle = 0$ 。随机力的特征可以用其关联函数来表征。关联函数 $\langle f(0)f(t) \rangle$ 具有下面的特征：

- 随机力来自于布朗粒子周围很多小分子的无规则碰撞。根据中心极限定理的思想，我们可以把它近似成满足高斯分布。
- 稳态下，关联函数只依赖于两个时间之差： $\langle f(t)f(t') \rangle = \langle f(0)f(t'-t) \rangle$ 。
- $\langle f^2(0) \rangle = \langle f(0)f(0) \rangle \geq 0$ ，且为关联函数的最大值。 $\langle f(0)f(t) \rangle = \langle f(0)f(-t) \rangle$ 随时间衰减。当 t 大于一个特征时间的时候， $\langle f(0)f(t) \rangle = \langle f(0) \rangle \langle f(t) \rangle = 0$ 。

- 这里的 $\langle \cdot \rangle$ 表示对噪声大量的具体实现进行系综平均。我们进一步假设噪声满足“各态遍历”的条件(详见附录23)。这样，这个系综平均就等价于针对一个实现的长时平均：

$$\lim_{T \rightarrow +\infty} \overline{f(t)f(0)}^T = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_{(t-T)/2}^{(t+T)/2} f(t+t')f(t') dt' = \langle f(t)f(0) \rangle.$$

我们首先考虑 $V'(x) = 0$ 的情况，并假设噪声关联函数为 δ 函数：

$$\langle f(0)f(t) \rangle = C\delta(t). \quad (17.15)$$

在时域 δ 关联的噪声其关联函数频谱不依赖于频率，所以被称为白噪声 (white noise)。相应的，如果频谱依赖于频率，则称为色噪声 (colored noise)。

17.2.3.1 噪声和耗散同时总是出现

在开始分析朗之万方程之前我们先强调一点：朗之万方程中的噪声项和耗散项总是同时出现。这是因为它们是同一个物理过程(溶液分子对布朗粒子的随机碰撞)的不同方面反映。

如果我们假设只有噪声，没有耗散，那么

$$m\dot{v} = f(t). \quad (17.16)$$

对方程两边积分得到

$$v(t) = v(0) + \int_0^t f(t')/mdt'.$$

对所有可能的噪声实现做系综平均，我们有

$$\langle v(t) \rangle = v(0).$$

进一步考虑

$$\begin{aligned} \langle v(t)^2 \rangle &= v^2(0) + \frac{C}{m^2} \int_0^t dt' \int_0^t dt'' \delta(t' - t'') \\ &= v^2(0) + \frac{C}{m^2} t. \end{aligned}$$

所以，速度的涨落为

$$\langle \delta v^2(t) \rangle = \frac{C}{m^2} t.$$

我们看到，如果没有耗散项，粒子平均速度会维持在初始速度，但其速度涨落(对应其动能)会随着时间线性增大，粒子无法达到一个稳定的状态。

同理，如果我们忽略噪声而只考虑耗散项，布朗粒子将由于耗散而最终停止运动

$$v(t) = v(0)e^{-\gamma t}. \quad (17.17)$$

下面通过求解分析这个最简单的 Langevin 方程，我们可以看到，只有噪声和耗散共存时，它才能与环境最终达到热平衡状态。

17.2.3.2 平均速度 $\langle v \rangle$

下面我们求解朗之万方程来得到布朗粒子的速度 $v(t)$ 。我们从式 (17.17) 出发。当考虑噪声项后，我们假设方程的解为

$$v(t) = C(t)e^{-\gamma t}$$

将上式代入到朗之万方程中，我们可以得到关于 $C(t)$ 的方程：

$$\frac{dC(t)}{dt} = \frac{1}{m} f(t) e^{\gamma t}$$

于是，我们得到：

$$v(t) = v(0)e^{-\gamma t} + \frac{1}{m} \int_0^t e^{-\gamma(t-t')} f(t') dt'. \quad (17.18)$$

如果对所有可能实现的随机力做平均，我们有

$$\langle v(t) \rangle = v(0)e^{-\gamma t}, \quad (17.19)$$

可见布朗粒子的平均速度随时间指数衰减，其稳态平均速度为零，平均位移也是零。

17.2.3.3 速度关联 $\langle v(t)v(t') \rangle$

我们将式 (17.18) 相乘，可以得到速度关联函数

$$\langle v(t)v(t') \rangle = e^{-\gamma(t+t')} \left(\left(v(0) + \frac{1}{m} \int_0^t e^{\gamma t_1} f(t_1) dt_1 \right) \left(v(0) + \frac{1}{m} \int_0^{t'} e^{\gamma t_2} f(t_2) dt_2 \right) \right)$$

将等式右边展开我们得到四项，其中两项与随机力 $f(t)$ 线性相关，系综平均后为零。因此，我们只得到两项：

$$= e^{-\gamma(t+t')} \left(v^2(0) + \frac{1}{m^2} \int_0^t dt_1 \int_0^{t'} dt_2 e^{\gamma(t_1+t_2)} \langle f(t_1)f(t_2) \rangle \right).$$

首先考虑 $t > t'$ 的情况。由于噪声的 δ 关联， t_1 的积分上限可以变为 t' 。于是

$$\begin{aligned} &= e^{-\gamma(t+t')} \left(v^2(0) + \frac{C}{2\gamma m^2} (e^{2\gamma t'} - 1) \right) \\ &= v^2(0)e^{-\gamma(t+t')} + \frac{C}{2\gamma m^2} (e^{-\gamma(t-t')} - e^{-\gamma(t+t')}). \end{aligned}$$

这里我们得到了给定初始速度 $v(0)$ 时，在后续 t 和 t' 时刻 $t > t'$ 的速度关联。我们可以看到，如果 t 和 t' 都远大于 $\tau_\gamma = \gamma^{-1}$ ，那么上式的第一项就接近 0。也就是说，当我们等的时间足够长时，初始条件将不再重要。这时，如果进一步令 $t = t'$ ，根据能量均分定理，我们得到

$$C = 2m\gamma k_B T.$$

于是我们得到了下面的重要关系

$$\langle f(t)f(t') \rangle = 2m\gamma k_B T \delta(t - t'). \quad (17.20)$$

这个式子把噪声关联函数和摩擦系数联系起来，是耗散-涨落定理在布朗运动中的具体体现。我们还可以把它写成积分的形式：

$$\gamma = \frac{1}{2mk_B T} \int dt \langle f(t)f(0) \rangle.$$

最后，速度关联函数可以写成

$$\langle v(t)v(t') \rangle = v^2(0)e^{-\gamma(t+t')} + k_B T/m(e^{-\gamma(t-t')} - e^{-\gamma(t+t')}). \quad (17.21)$$

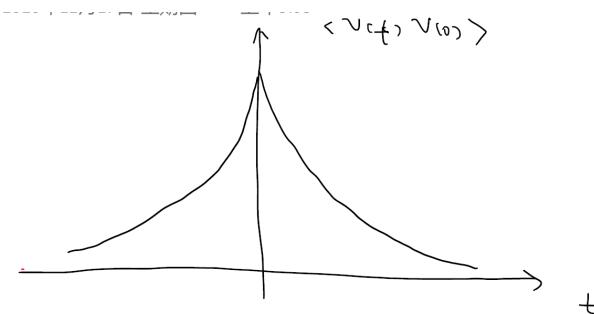
上面的结果是在假设 $t > t'$ 的情况下得到的。如果 $t' > t$ ，那么上式中的 $t - t'$ 需要被 $t' - t$ 来代替。所以我们可以把这两种情况统一写为

$$\langle v(t)v(t') \rangle = v^2(0)e^{-\gamma(t+t')} + \frac{k_B T}{m}(e^{-\gamma|t-t'|} - e^{-\gamma(t+t')}). \quad (17.22)$$

讨论：

- 虽然随机力 $f(t)$ 是 δ 关联的，但布朗粒子的速度关联函数随时间呈指数衰减，对应的特征衰减时间为 τ_γ 。当 $t + t' \gg \tau_\gamma$ 时（如下图所示），

$$\langle v(t)v(t') \rangle = \frac{k_B T}{m} e^{-\gamma|t-t'|}. \quad (17.23)$$



- 如果 $t = t'$ ，则

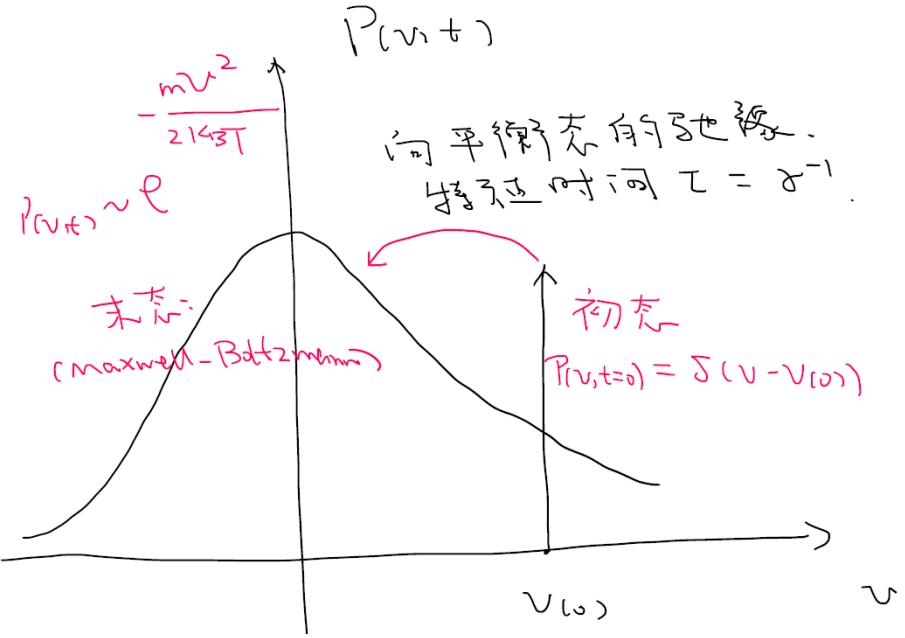
$$\langle \delta v^2(t) \rangle = \frac{k_B T}{m} (1 - e^{-2\gamma t}). \quad (17.24)$$

- $t = 0$ 时，粒子的速度确定，因此涨落为零；

- $t \ll \tau_\gamma$ 时, $\delta v^2(t) \sim 2t\gamma k_B T/m$ 。这时, 速度 $v(t)$ 也表现出一定程度的扩散现象, 对应的速度扩散系数为 $\mathcal{D}_v = \gamma k_B T/m$ 。这可以与下面得到的位移扩散系数(17.30)对比。这时, 耗散项还没有起到作用。
- $t \gg \tau_\gamma$ 时, 耗散和涨落项的贡献都体现了出来, 粒子与周围环境达到热平衡, 其速度涨落由环境温度确定为 $k_B T/m$ 。
- 由于随机力为一个稳态高斯随机过程, 而我们现在分析的朗之万方程为线性方程。高斯过程的线性变换仍然为高斯过程。根据 $\langle v(t) \rangle$ 和 $\langle \delta v^2(t) \rangle$, 我们可以写出布朗粒子的概率密度分布

$$P(v, t) = \sqrt{\frac{m}{2\pi k_B T(1 - e^{-2\gamma t})}} \exp\left(-\frac{m(v(t) - v(0)e^{-\gamma t})^2}{2k_B T(1 - e^{-2\gamma t})}\right). \quad (17.25)$$

上式被称为 Ornstein-Uhlenbeck 分布。可见, 虽然布朗粒子初始时刻速度为 $v(0)$, 但随着时间的推移, 其分布趋向于平衡态的麦克斯韦分布。而这一过程正是由于周围液体分子对它的随机碰撞导致的。这是一个从非平衡分布向平衡分布弛豫的过程, 所需的特征时



间为 τ_γ (如下图)。

17.2.3.4 位移 $\langle x(t) \rangle$

把速度表达式相对时间做积分我们可以得到 Langevin 方程的位移解

$$x(t) = x(0) + \frac{v(0)}{\gamma} (1 - e^{-\gamma t}) + \frac{1}{m} \int_0^t dt' \frac{1 - e^{-\gamma(t-t')}}{\gamma} f(t'). \quad (17.26)$$

如果对随机力做系综平均，我们有

$$\langle x(t) \rangle = x(0) + \frac{v(0)}{\gamma} (1 - e^{-\gamma t}).$$

当 $t \gg \tau_\gamma$ 时，上式还可以写成

$$\langle x(t) - x(0) \rangle = v(0)\tau. \quad (17.27)$$

所以，布朗粒子平均偏离初始位置的距离由它的初始速度和驰豫时间的乘积给出。

根据式 (17.26)，我们可以进一步计算位移的关联函数。这里我们主要考虑偏离初始位置的位移 $\delta x(t) = x(t) - x(0)$

$$\boxed{\langle \delta x^2(t) \rangle = \frac{v^2(0)}{\gamma^2} (1 - e^{-\gamma t})^2 - \frac{k_B T}{m\gamma^2} (1 - e^{-\gamma t})(3 - e^{-\gamma t}) + \frac{2k_B T t}{m\gamma}.} \quad (17.28)$$

讨论：

- 这个结果依赖于初始速度。如果我们认为粒子的初始速度由周围环境的统计性质随机给出的话，我们有

$$\overline{v^2(0)} = \frac{k_B T}{m}.$$

对初始速度做平均，我们得到

$$\boxed{\langle \delta x^2(t) \rangle = \frac{2k_B T}{m\gamma^2} (\gamma t - 1 + e^{-\gamma t})}. \quad (17.29)$$

- 当 $t \ll \tau_\gamma$ 时， $\overline{\langle x^2(t) \rangle} \approx \frac{k_B T}{m} t^2 = \overline{\langle v^2 \rangle} t^2$ 。这说明，布朗粒子的运动近似满足牛顿定律 $\sqrt{\overline{\langle x^2(t) \rangle}} \propto t$ 。
- 当 $t \gg \tau$ 的时候， $\overline{\langle x^2(t) \rangle} \approx \frac{2k_B T}{m\gamma} t$ ，布朗粒子表现出随机运动的特性 $\sqrt{\overline{\langle x^2(t) \rangle}} \propto \sqrt{t}$ 。
- 如果把上面的结果与随机行走理论的结果 $\overline{\langle \delta x^2 \rangle} = 2Dt$ 进行对比，我们发现位移扩散系数

$$\boxed{D = \frac{k_B T}{m\gamma}} \quad (17.30)$$

这就是著名的爱因斯坦关系³。它把布朗粒子在液体中的扩散系数和它所感受的摩擦力联系了起来。这说明，两者具有相同的来源，即液体分子的随机、无规则运动。或者说，布朗粒子的扩散行为是由于液体分子的随机碰撞所导致的。对于球形颗粒，进一步利用 Stokes 定理

$$m\gamma = 6\pi\eta a$$

²我们也可以不求解 $x(t)$ 直接根据 $\overline{\langle \delta x^2(t) \rangle} = \left(\int_0^t dt_1 v(t_1)\right) \left(\int_0^t dt_1 v(t_1)\right)$ 求得。

³Einstein, Investigations on the theory of the Brownian movement (1905).

我们得到

$$D = \frac{k_B T}{6\pi\eta a}$$

可见，如果知道了液体的粘性系数 η 和温度，通过测量扩散系数，我们可以推测出布朗粒子的半径大小 a 。此外，如果知道布朗粒子的大小，我们可以利用此关系来确定 Boltzmann 常数。实际上，这正是法国人 Jean Baptiste Perrin 所做的。⁴

17.2.4 外力作用下的布朗运动

我们可以进一步考虑在一个恒定外力 F 作用下的布朗粒子的运动。这可以对应很多种具体的物理系统。比如，金属中电子在外加静电场中的运动⁵，物体在重力场作用下的运动等。这时的运动方程为：

$$m\dot{v}(t) = -m\gamma v(t) + F + f(t). \quad (17.31)$$

在外场作用下，粒子会获得一个平均速度，在统计平均后得到：

$$\langle v \rangle = \frac{F}{m\gamma}.$$

我们可以定义对应的粒子迁移率：

$$\mu = \frac{\langle v \rangle}{F} = \frac{1}{m\gamma}.$$

我们可以进而把 Einstein 关系写成：

$$D = \frac{k_B T}{m\gamma} = k_B T \mu.$$

17.2.5 谐振子布朗运动

下面我们考虑布朗粒子处于简谐势下的运动情况。单个谐振子的朗之万方程可以写为：

$$m\ddot{x}(t) = -m\omega_0^2 x(t) - m\gamma\dot{x}(t) + f(t). \quad (17.32)$$

因为简谐势是线性势，上面的 Langevin 方程仍然是一个线性方程。我们可以利用傅里叶分析的方法来得到朗之万方程的稳态解。下面我们具体介绍这个方法的应用。

首先我们假设系统在稳态随机力 $f(t)$ 及其它驱动力的作用下本身已经达到稳态。这时，系统的状态不依赖于绝对时间，而只和相对时间有关。我们仍然假设随机力服从高斯分布，并

⁴这里可以设计一个物理实验！

⁵我们可以基于本节的模型来得到金属导电的 Drude 模型，详见 Ashcroft&Mermin, Solid State Physics, 第一章。

且 $\langle f(t)f(t') \rangle = C\delta(t-t')$, 常数 C 待定。

17.2.5.1 稳态位移

系统达到稳态之后, 我们可以利用傅立叶变换

$$x[\omega] = \int x(t)e^{i\omega t} dt, \quad x(t) = \int x[\omega]e^{-i\omega t} \frac{d\omega}{2\pi}. \quad (17.33)$$

代入到式 (17.32) 中,

$$-m\omega^2 x[\omega] = -m\omega_0^2 x[\omega] + im\gamma\omega x[\omega] + f[\omega]. \quad (17.34)$$

于是我们可以得到谐振子位移在频率空间的解

$$x[\omega] = -\frac{f[\omega]/m}{\omega^2 - \omega_0^2 + i\gamma\omega} = -\frac{f[\omega]/m}{(\omega + i\gamma/2)^2 - (\omega_0^2 - \gamma^2/4)}. \quad (17.35)$$

定义 $\tilde{\omega}_0^2 = \omega_0^2 - \gamma^2/4$ 和推迟格林函数

$$D^r[\omega] = \frac{1}{(\omega + i\gamma/2)^2 - \tilde{\omega}_0^2}, \quad (17.36)$$

我们得到简化的结果

$$x[\omega] = -D^r[\omega]f[\omega]/m. \quad (17.37)$$

下面我们要做的是从频率空间转换到时间空间。

首先看推迟格林函数

$$\begin{aligned} D^r(t) &= \int \frac{e^{-i\omega t}}{(\omega + i\gamma/2)^2 - \tilde{\omega}_0^2} \frac{d\omega}{2\pi} \\ &= \frac{1}{2\tilde{\omega}_0} \int \left(\frac{e^{-i\omega t}}{\omega + i\gamma/2 - \tilde{\omega}_0} - \frac{e^{-i\omega t}}{\omega + i\gamma/2 + \tilde{\omega}_0} \right) \frac{d\omega}{2\pi} \\ &= -\frac{i}{2\tilde{\omega}_0} \Theta(t) \left[e^{-i(\tilde{\omega}_0 - i\gamma/2)t} - e^{-i(-\tilde{\omega}_0 - i\gamma/2)t} \right] \\ &= -\Theta(t) \frac{\sin \tilde{\omega}_0 t}{\tilde{\omega}_0} e^{-\gamma t/2}. \end{aligned} \quad (17.38)$$

这里, 如果 $t > 0$, 阶跃函数 $\Theta(t) = 1$; 如果 $t < 0$, $\Theta(t) = 0$ 。它的积分形式可以写成:

$$\boxed{\Theta(t) = \lim_{\delta \rightarrow 0^+} \frac{e^{-i\omega t}}{\omega + i\delta} \frac{d\omega}{(-2\pi i)}}. \quad (17.39)$$

利用傅立叶变换的性质, 频率空间的直接相乘变为在时间空间的卷积, 所以我们得到时域解为

$$x(t) = -\frac{1}{m\tilde{\omega}_0} \int_{-\infty}^t \sin \tilde{\omega}_0(t-t') e^{-\gamma(t-t')/2} f(t') dt'. \quad (17.40)$$

从上面的分析我们可以看到, 由于摩擦系数 γ 的存在导致我们在时域得到一个阶跃函数 $\Theta(t)$ 。而我们知道摩擦系数导致能量损耗/耗散, 破坏时间反演对称, 对应在时域台阶函数起到相同的

作用。在时域可以这样理解, $D^r(t)$ 其实是一个响应函数, 它描述的是谐振子在 t 时刻对 0 时刻扰动的响应, 参考式 (17.40)。我们知道谐振子只会对 t 时刻之前的扰动产生响应, 而不会对 t 时刻之后的扰动响应, 而 $D^r(t)$ 中的阶跃函数正好保证了这一点。

17.2.5.2 平均动能

下面我们用简谐分析的方法来计算谐振子的平均动能。当振子达到平衡后, 其平均动能不再依赖时间

$$\begin{aligned}\frac{m}{2}\langle\dot{x}^2\rangle &= -\frac{m}{2}\int\frac{d\omega}{2\pi}\int\frac{d\omega'}{2\pi}\omega\omega'e^{-i\omega t}e^{-i\omega't}\langle x(\omega)x(\omega')\rangle \\ &= -\frac{1}{2m}\int\frac{d\omega}{2\pi}\int\frac{d\omega'}{2\pi}e^{-i(\omega+\omega')t}\frac{\omega\omega'\langle f(\omega)f(\omega')\rangle}{[(\omega+i\gamma/2)^2-\tilde{\omega}_0^2][(\omega'+i\gamma/2)^2-\tilde{\omega}_0^2]}\end{aligned}$$

频域的噪声关联函数为:

$$\langle f(\omega)f(\omega')\rangle = 2\pi C\delta(\omega+\omega'),$$

这就要求 $\omega+\omega'=0$ 。于是

$$\begin{aligned}&= \frac{C}{2m}\int\frac{\omega^2}{[(\omega+i\gamma/2)^2-\tilde{\omega}_0^2][(\omega-i\gamma/2)^2-\tilde{\omega}_0^2]}\frac{d\omega}{2\pi} \\ &= \frac{C}{4m\gamma}\end{aligned}$$

这里的常数 C 可以同样通过经典的能量均分定理来得到

$$C = 2\gamma k_B T m.$$

与自由粒子类似, 随机力的关联函数可以写成

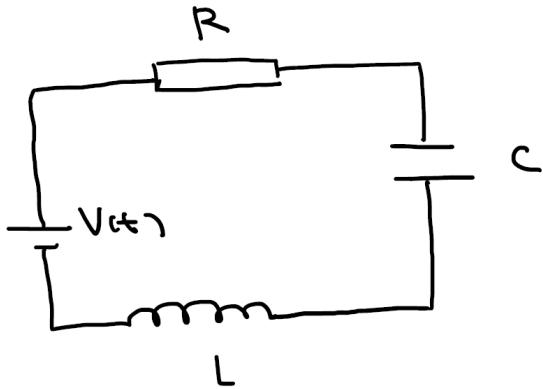
$$\langle f(t)f(t')\rangle = 2m\gamma k_B T \delta(t-t').$$

例题 17.1 对于 RCL 电路 (如下图), 其偏压关系可以表示为:

$$V(t) = I(t)R + \frac{Q(t)}{C} + L\frac{dI(t)}{dt}$$

它可以看作是一个谐振子的 Langevin 方程:

$$L\frac{d^2Q(t)}{dt^2} = -\frac{Q}{C} - RI(t) + \bar{V}(t) - \delta V(t).$$



笔记 Wiener-Khinchin 定理：考虑一个稳态，满足各态遍历特征的随机过程，用变量 $f(t)$ 来描述⁶。其均值 $\langle f(t) \rangle$ 及时间关联函数 $\langle f(t+t')f(t) \rangle$ 不随绝对时间的变化而变化。

我们把时间关联函数的 Fourier 变换表示为：

$$F[\omega] = \lim_{T \rightarrow +\infty} \int_{-T}^{+T} dt e^{i\omega t} \langle f(t) f(0) \rangle.$$

这里，我们引入了随机过程的时间长度 $[-T, +T]$ ，最终我们需要让 $T \rightarrow +\infty$ 。也就是说，随机过程的时间长度足够长。这里由于各态遍历特征，关联函数满足：

$$\langle f(t) f(t') \rangle = \lim_{T \rightarrow \infty} \int_{t+t_1, t'+t_1 \in [-T, T]} dt_1 f(t+t_1) f(t'+t_1).$$

同理，我们还可以把随机变量 $f(t)$ 的 Fourier 变换表示为：

$$f[\omega] = \int_{-T}^{+T} dt e^{i\omega t} f(t).$$

定义随机变量 $f(t)$ 的功率谱为

$$S[\omega, -\omega] = \frac{1}{2T} f[\omega] f[-\omega] = \frac{1}{2T} |f[\omega]|^2. \quad (17.41)$$

Wiener-Khinchin 定理告诉我们，功率谱正好是关联函数的 Fourier 变换

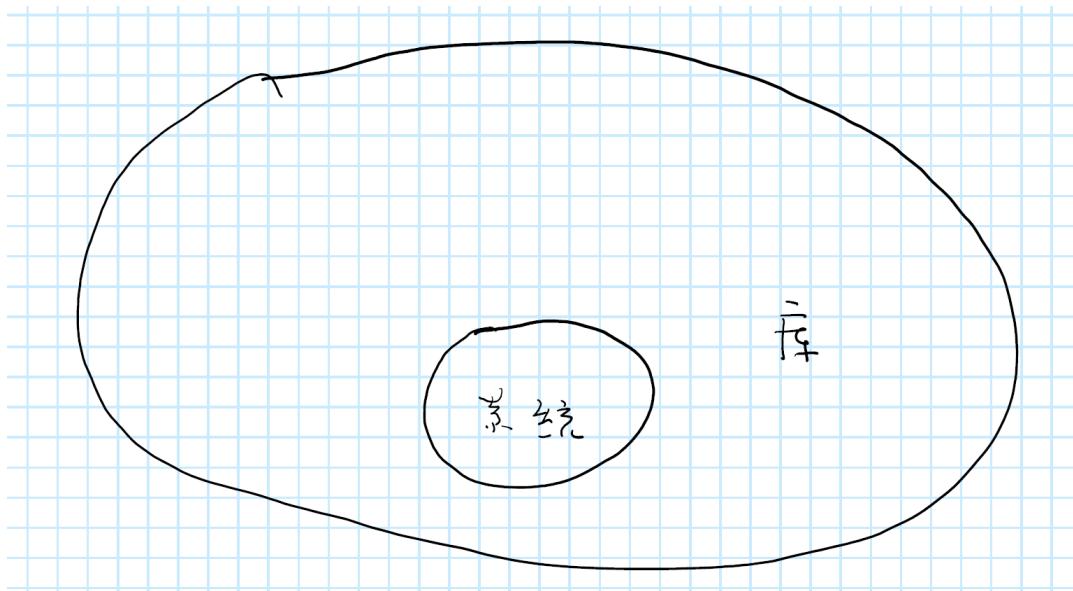
$$S[\omega, -\omega] = F[\omega]. \quad (17.42)$$

或者推广为：

$$f[\omega] f[\omega'] = 2\pi \delta(\omega + \omega') F[\omega].$$

⁶ 我们这里只考虑 $f(t)$ 为实数的情况。

17.3 朗之万方程的微观理论



我们前面是通过物理上的考虑来直接写出了朗之万方程。实际上，我们可以通过给定的微观的物理模型将其推导出来。现在我们就以谐振子模型为例，来推导出量子版本的朗之万方程。其基本思路如下：我们从一个包含很多自由度的总系统出发。我们选择其中某一个或有限个自由度作为研究的系统，把其他所有自由度作为环境。我们希望研究系统的动力学行为，对于环境我们不太关心。这时，我们可以从所有自由度的运动方程出发，把环境自由度消去，只留下我们关心的系统部分。这样，我们就可以得到一组所谓的约化运动方程，而环境对系统自由度的影响还可以保留在约化运动方程中。我们会看到，它们会对系统自由度带来涨落和耗散，类似花粉颗粒在溶液环境中所做的布朗运动。在这样的模型下，我们的哈密顿量可以表示为

$$H = H_s + H_b + V \quad (17.43)$$

其中， H_s , H_b 和 V 分别代表系统、环境以及它们之间的耦合哈密顿量。

耦合谐振子哈密顿量：我们考虑简谐近似下固体中的晶格振动。体系总的哈密顿量为：

$$H = \sum_i \frac{1}{2m_i} p_i^2 + \frac{1}{2} \sum_{i,j} x_i K_{ij} x_j. \quad (17.44)$$

其中， x_i 和 p_i 分别为每个自由度偏离平衡位置的位移和相应的动量， m_i 为其质量。定义

$$u_i = \sqrt{m_i} x_i,$$

我们可以隐去质量，并写成矩阵的形式：

$$H = \frac{1}{2}p^T p + \frac{1}{2}u^T Ku. \quad (17.45)$$

这里 p 和 u 为两个列向量，每个向量包含了所有的自由度。按照上面的讨论，我们可以把整个体系分成两部分 (s 表示系统， b 表示环境/热库)

$$\begin{aligned} H &= H_s + H_b + V = H_0 + V \\ &= \sum_{\alpha=s,b} \left(\frac{1}{2}p_\alpha^T p_\alpha + \frac{1}{2}u_\alpha^T K_\alpha u_\alpha \right) + \frac{1}{2}u_s^T V^{sb} u_b + \frac{1}{2}u_b^T V^{bs} u_s. \end{aligned} \quad (17.46)$$

我们把系统与库的耦合 V 分成了两部分。以上哈密顿量是在实空间下的描述。省略掉算符符号，它可以同时描述经典和量子系统。我们可以将其转化到模式空间，并用二次量子化的形式表示出来。根据标准的定义

$$\begin{aligned} u_q(t) &= \sqrt{\frac{1}{2\omega}}(a_q^\dagger(t) + a_q(t)), \\ p_q(t) &= i\sqrt{\frac{\omega}{2}}(a_q^\dagger(t) - a_q(t)), \end{aligned}$$

我们得到：

$$H_\alpha = \sum_{q \in \alpha} \left(a_q^\dagger a_q + \frac{1}{2} \right) \hbar\omega_q, \quad \alpha = s, b \quad (17.47)$$

$$V = \sum_{q \in s, p \in b} (a_q + a_q^\dagger) g_{qp} (a_p + a_p^\dagger). \quad (17.48)$$

我们将主要以此模型为例，来介绍本部分的内容。

17.3.1 运动方程

量子：

如果我们考虑量子系统，就要从海森堡运动方程出发。在海森堡表象下，算符的运动方程为 ($\hbar = 1$)

$$\frac{d}{dt} \hat{A}(t) = -i[\hat{A}(t), \hat{H}(t)], \quad (17.49)$$

而 \hat{p} 和 \hat{u} 的对易关系为 ($\hbar = 1$)

$$[\hat{u}_k, \hat{p}_j] = i\delta_{k,j}. \quad (17.50)$$

于是我们得到：

$$\dot{\hat{p}}_s = \ddot{\hat{u}}_s = -K_s \hat{u}_s - V_{sb} \hat{u}_b, \quad (17.51)$$

$$\dot{\hat{p}}_b = \ddot{\hat{u}}_b = -K_b \hat{u}_b - V_{bs} \hat{u}_s. \quad (17.52)$$

这里， \hat{p}_s 和 \hat{u}_s 为系统的动量和位移算符，而 \hat{p}_b 和 \hat{u}_b 为热库的动量和位移算符。

经典：

如果我们考虑经典的谐振子系统，那么我们就可以从哈密顿运动方程出发

$$\begin{aligned}\dot{u} &= \frac{\partial H}{\partial p}, \\ \dot{p} &= -\frac{\partial H}{\partial u}.\end{aligned}$$

很容易验证，对于谐振子来说，其量子和经典的运动方程形式是一样的。只不过经典方程中 u 和 p 不再是算符。

17.3.2 系统的约化运动方程

因为我们对环境/热库的微观细节并不关心，所以我们下面的目的是从式(17.51)和(17.52)出发，消去热库自由度来得到一个描述系统自由度的约化运动方程。

首先，定义热库自由度的响应函数如下：

$$-\frac{\partial^2}{\partial t'^2} D_b^r(t, t') = D_b^r(t, t') K_b + \delta(t, t'). \quad (17.53)$$

把方程(17.52)中的 t 换成 t' ，左乘 $D_b^r(t, t')$ ，然后对 t' 积分，同时对方程(17.53)右乘 $u_b(t')$ 再对 t' 积分，最后把两个式子相加，我们得到

$$\begin{aligned}\hat{u}_b(t) &= [D_b^r(t, t') \dot{u}(t') - \dot{D}_b^r(t, t') \hat{u}(t')] \Big|_{t_0}^{+\infty} + \int_{t_0}^{+\infty} D_b^r(t, t') V_{bs} \hat{u}_s(t') dt' \\ &= -D_b^r(t, t_0) \dot{\hat{u}}(t_0) + \dot{D}_b^r(t, t_0) \hat{u}(t_0) + \int_{t_0}^t D_b^r(t, t') V_{bs} \hat{u}_s(t') dt'.\end{aligned} \quad (17.54)$$

这里， $D_b^r(t)$ 为响应函数，它的物理意义是系统在 t 时刻对 0 时刻扰动的响应。这就要求 $D_b^r(t) = 0$ if $t < 0$ ，因为系统只能对过去有响应，无法对未来的扰动有响应。而且要注意 $\dot{D}_b^r(t, t') = \frac{\partial}{\partial t'} D_b^r(t, t')$ 。把方程(17.54)代入到(17.51)，我们得到

$$\ddot{\hat{u}}_s(t) = -K_s \hat{u}_s(t) - \int_{t_0}^t \Pi^r(t - t') \hat{u}_s(t') dt' + \hat{\xi}(t) \quad (17.55)$$

或者

$$\ddot{\hat{u}}_s(t) = -K_s \hat{u}_s(t) - \int_0^{t-t_0} \Pi^r(\tau) \hat{u}_s(t - \tau) d\tau + \hat{\xi}(t), \quad (17.56)$$

再或者

$$\ddot{u}_s(t) = -K'_s \hat{u}_s(t) - \int_0^{t-t_0} \Gamma^r(\tau) \dot{\hat{u}}_s(t-\tau) d\tau + \xi(t), \quad (17.57)$$

在方程 (17.57) 中，我们对 (17.56) 做了分部积分，把右边第二项写成了耗散/摩擦系数与速度的卷积形式。这样做的话，分部积分的边界项会对 K_s 有个修正，包含修正之后我们把它表示为 K'_s 。上面的几个方程统称为广义朗之万方程。之所以叫广义，是因为它的耗散项存在记忆效应。为了把方程写成上面的简化形式，我们定义了

$$\hat{\xi}(t) \equiv V_{sb} (D_b^r(t-t_0) \dot{\hat{u}}_b(t_0) - \dot{D}_b^r(t-t_0) \hat{u}_b(t_0)), \quad (17.58)$$

$$\Gamma^r(t) \equiv \int_{-\infty}^t \Pi^r(t') dt', \quad (17.59)$$

其中，

$$\Pi^r(t-t') = V_{sb} D_b^r(t-t') V_{bs}$$

是热库贡献的声子自能。如果考虑经典系统，我们会得到相同形式的方程，只不过算符就变成我们熟悉的位移、动量等经典物理量。

耗散项：对于单个谐振子模式 q ，我们可以计算其响应函数得到

$$D_b^r(t) = -\Theta(t) \frac{1}{\omega_q} e^{-\gamma t} \sin(\omega_q t) \quad (17.60)$$

我们引入小量 γ 的目的是，保证响应函数在 $t \rightarrow +\infty$ 时趋于 0。于是，要得到自能的表达式，我们可以先对热库自由度做一个变换到简正坐标下，这样的话利用 (17.60) 就可以把自能表示为：

$$\Pi^r(t) = -\Theta(t) \sum_q V_{sq} \frac{\sin(\omega_q t)}{\omega_q} V_{qs}, \quad (17.61)$$

为了对上式进行进一步简化，我们引入库的谱密度函数

$$J(\omega) = \pi \sum_q V_{sq} \delta(\omega - \omega_q) V_{qs},$$

于是

$$\Pi^r(t) = -\Theta(t) \int \frac{d\omega}{\pi} \omega^{-1} J(\omega) \sin(\omega t), \quad (17.62)$$

耗散项对于经典和量子的情况取相同的表达式。我们可以看到，这里得到的耗散矩阵是具有记忆效应的。也就是说，当前系统的运动规律依赖于它之前的运动情况。

噪声项：

- **量子：**对于噪声项，量子情况下， ξ 为算符。如果热库处于热平衡态，我们可以利用热

平衡态下的密度矩阵来计算它的均值和方差：

$$\langle \hat{\xi}_\alpha \rangle_b = 0, \quad (17.63)$$

$$\langle \hat{\xi}(t) \hat{\xi}(0) \rangle_b = \int \frac{d\omega}{\pi} \omega^{-1} J(\omega) \left(\left[n_B(\omega) + \frac{1}{2} \right] \cos[\omega t] - \frac{i}{2} \sin[\omega t] \right). \quad (17.64)$$

我们假设热库温度为 T , 其中的平均 $\langle \cdots \rangle_b$ 为相对于热库平衡态正则密度矩阵的量子平均。比如,

$$\langle \hat{\xi}(t) \hat{\xi}(t') \rangle_b \neq \langle \hat{\xi}(t') \hat{\xi}(t) \rangle_b.$$

也就是说, 不同时间的噪声算符不对易, 这是量子力学的直接结果.

- **经典:** 如果我们考虑经典系统, 热库的统计性质由 Maxwell-Boltzmann 分布给出。这时我们可以得到:

$$\langle p_q p_q \rangle_b = k_B T, \quad \langle u_q u_q \rangle_b = k_B T / \omega_q^2,$$

$$\langle p_q u_q \rangle_b = 0, \quad \langle u_q p_q \rangle_b = 0.$$

可以验证, 这正是量子力学的结果取经典极限。这里的经典极限包含两重意思:

- 忽略动量和位移的不对易性 $\langle p_q u_q \rangle_b = \langle u_q p_q \rangle_b = 0$.
- 将 Bose-Einstein 分布取高温极限 $n_B(\omega) + \frac{1}{2} \approx \frac{k_B T}{\omega}$.

于是, 噪声关联函数变为:

$$\langle \xi_\alpha \rangle_b = 0, \quad (17.65)$$

$$\langle \xi(t) \xi(0) \rangle_b = k_B T \int \frac{d\omega}{\pi} \omega^{-2} J(\omega) \cos(\omega t). \quad (17.66)$$

✿ **练习 17.1** 证明噪声关联函数满足式子 (17.64). 步骤如下:

1. 在本征模式下, 热库原子的不同振动模式相互独立, 我们只需要单独考虑每个模式.
2. 对于模式 q 来说,

$$\begin{aligned} \langle \xi(t) \xi(t') \rangle_b &= V_{sq} \langle [D_q^r(t-t_0) p_q(t_0) - \dot{D}_q^r(t-t_0) u_q(t_0)] \\ &\quad \times [D_q^r(t'-t_0) p_q(t_0) - \dot{D}_q^r(t'-t_0) u_q(t_0)] \rangle_b V_{qs} \end{aligned}$$

对于每一个模式来说, 我们知道其格林函数为

$$D_q^r(t) = - \lim_{\gamma \rightarrow 0} \Theta(t) \frac{\sin \omega_q t}{\omega_q} e^{-\gamma t}.$$

3. 展开之后我们得到类似 $\langle p_q(t_0) p_q(t_0) \rangle_b$ 这样的四项, 我们可以利用??, 和粒子数算符

$$\langle a_q^\dagger a_q \rangle_b = n_B(\omega_q)$$

$$\langle a_q a_q^\dagger \rangle_b = 1 + n_B(\omega_q)$$

来计算有：

$$\begin{aligned}\langle p_q p_q \rangle &= \omega_q \left(n_B(\omega_q) + \frac{1}{2} \right), \\ \langle u_q u_q \rangle &= \frac{1}{\omega_q} \left(n_B(\omega_q) + \frac{1}{2} \right), \\ \langle u_q p_q \rangle &= \frac{i}{2}, \quad \langle p_q u_q \rangle = -\frac{i}{2}.\end{aligned}$$

其中

$$n_B(\omega) = \frac{1}{\exp\left(\frac{\omega}{k_B T}\right) - 1}$$

为玻色-爱因斯坦分布。

4. 最后整理可以得到式 (17.64)。

✎ 练习 17.2 从 Maxwell-Boltzmann 分布出发证明 (??-??)。

17.3.3 Markov 近似

以上我们得到的广义 Langevin 方程仍然存在记忆效应。下面我们说明，当库自由度的时间尺度要远快于系统自由度时，它会退化为大家熟悉的 Langevin 方程的形式。基于 (17.59) 和 (17.62) 我们可以得到：

$$\Gamma^r(t) = \int \frac{d\omega}{\pi} \omega^{-2} J(\omega) \cos(\omega t)$$

如果我们忽略 $D(\omega)/\omega^2$ 的频率依赖，令 $J(\omega) = D\omega^2$ 。这样我们就得到了 Markov 形式的结果：

$$\Gamma^r(t) = 2D\delta(t)$$

相应的，噪声关联函数也会发生变化。对于经典的情况，我们得到熟悉的结果：

$$\langle \xi(t)\xi(0) \rangle_b = 2Dk_B T\delta(t).$$

✎ 练习 17.3 证明：对于三维、线性色散的声子 $\omega(q) = cq$ 。对应的谱密度函数满足 $J(\omega) \propto \omega^2$ 。

17.3.4 半经典近似

介于经典和量子之间，我们还可以考虑一种半经典近似。我们只忽略 u_q 和 p_q 的不对易性，但保留 $\langle u_q u_q \rangle_b$ 和 $\langle p_q p_q \rangle$ 的量子平均。这样的话，噪声关联函数不再包含虚部，我们就可以用经典随机变量来描述噪声项。把 $\langle \xi(t)\xi(t') \rangle_S$ 和 $\Pi^r(t-t')$ 转换到频域，我们可以得到下

面的关系：

$$\begin{aligned}\langle \xi(\omega)\xi(\omega') \rangle_S &= -2 \left(n_B(\omega) + \frac{1}{2} \right) \text{Im}\Pi^r(\omega)\delta(\omega + \omega') \\ &= 2\omega \left(n_B(\omega) + \frac{1}{2} \right) \text{Re}\Gamma^r(\omega)\delta(\omega + \omega').\end{aligned}\quad (17.67)$$

这是量子版本的耗散-涨落定理。其形式和我们前面讨论的经典结果非常类似。随机力中既包含热涨落，也包含零点振动带来的涨落，或者称为量子涨落。

- ❖ **练习 17.4** 由式 (17.64) 出发，忽略其虚部后，通过傅里叶变换证明式 (17.67)。
- ❖ **练习 17.5** 参考 17.2.5 节证明 (17.68) 的结果，并说明在高温极限下，17.2.5 节的结果可以由 (17.68) 给出。
- ❖ **练习 17.6** 证明如果把一个角频率为 ω_0 的简谐振子用半经典的朗之万方程描述的话，当它与热库达到平衡之后，谐振子的平均能量为

$$E = \left(n_B(\omega_0) + \frac{1}{2} \right) \omega_0, \quad (17.68)$$

也就是说我们可以得到量子力学的结果。

- ❖ **练习 17.7** 我们考虑二次量子化表象下来描述一个谐振模式与热库耦合时，产生和湮灭算符所满足的 Langevin 方程。考虑下面的哈密顿量：

$$\hat{H} = \hat{H}_s + \hat{H}_b + \hat{V} = \hat{H}_0 + \hat{V}, \quad (17.69)$$

$$\hat{H}_s = \hbar\Omega(\hat{a}^\dagger \hat{a}), \quad (17.70)$$

$$\hat{H}_b = \sum_k \hbar\omega_k (\hat{a}_k^\dagger \hat{a}_k + \frac{1}{2}), \quad (17.71)$$

$$\hat{V} = \sum_k (g_k \hat{a}_k^\dagger \hat{a} + g_k^* \hat{a}^\dagger \hat{a}_k). \quad (17.72)$$

- (1). 写出 \hat{a} 和 \hat{a}_k 的运动方程，然后使用类似的方法得到描述 \hat{a} 的 Langevin 方程，指出耗散和噪声项，并分析噪声算符的关联函数。
- (2). 根据以上结果，写出 \hat{a}^\dagger 的 Langevin 方程。

17.4 Fokker-Planck 方程

Langevin 方程就像牛顿运动方程或者哈密顿运动方程一样，描述的是系统的状态参量（坐标、速度）随时间的演化。它们的主要区别是 Langevin 方程中包含了环境对系统的影响，即耗散和涨落项。前面我们已经论述过，对于经典多粒子系统，我们通常采用概率的方式来描述其统计性质，从而引入概率密度分布函数，并得到了与哈密顿运动方程对应的 Liouville 运动方程来描述概率密度分布函数的演化。类似的，也存在与 Langevin 方程相对应的系统概率

密度分布演化方程。这个方程为 Fokker-Planck (F-P) 方程。这是本节的主要内容。

为了得到 F-P 方程⁷，我们的出发点仍然是概率密度分布函数 $\rho(X, t)$ 在整个相空间是守恒的，因此它满足一个对应的连续性方程：

$$\frac{\partial \rho(\vec{X}, t)}{\partial t} = -\frac{\partial}{\partial \vec{X}} \cdot (\dot{\vec{X}} \rho(X, t)). \quad (17.73)$$

这里 X 为相空间的一点，比如 $\vec{X} = (q_i, p_i)$ 或者 (q_i, v_i) 。矢量 X 所满足的 Langevin 方程为：

$$\dot{\vec{X}} = \vec{v}(\vec{X}) + \vec{f}$$

这里 $\vec{f}(t)$ 为随机力，而 \vec{v} 为其它的力。 \vec{v} 的具体形式现在并不重要，但我们要求 $\vec{f}(t)$ 是稳态的高斯白噪声

$$\langle \vec{f}(t) \rangle = 0, \quad \langle f_i(t) f_j(t') \rangle = D_{ij} \delta_{i,j} \delta(t - t'), \quad \text{or} \quad \langle \vec{f}(t) \vec{f}(t') \rangle = \vec{D} \delta(t - t').$$

显然 \vec{D} 为对角矩阵。将上式代入到 (17.73) 中得到：

$$\frac{\partial \rho(\vec{X}, t)}{\partial t} = -\frac{\partial}{\partial \vec{X}} \cdot \left[\vec{v}(\vec{X}) \rho(X, t) + \vec{f} \rho(X, t) \right]. \quad (17.74)$$

因为包含了随机力 f ，因此这是一个随机微分方程。下面我们希望能够对噪声进行平均从而得到一个确定性的微分方程。为了简化记号，我们引入一个符号 \mathcal{L} 定义为：

$$\mathcal{L}\rho \equiv \frac{\partial}{\partial \vec{X}} \cdot (\vec{v}(\vec{X}) \rho)$$

下面的分析和求解 Langevin 方程类似。如果不考虑随机力的影响，我们有：

$$\rho(\vec{X}, t) = e^{-\mathcal{L}t} \rho(\vec{X}, 0)$$

考虑随机力的话，我们可以得到：

$$\rho(\vec{X}, t) = \rho(\vec{X}, 0) e^{-\mathcal{L}t} - \int_0^t ds e^{-\mathcal{L}(t-s)} \frac{\partial}{\partial \vec{X}} \cdot [\vec{f}(s) \rho(\vec{X}, s)]$$

这里 $t \geq s$ 表示系统对 t 时刻之前的状态有依赖。将上式代入 (17.74) 得到：

$$\begin{aligned} \frac{\partial \rho(\vec{X}, t)}{\partial t} &= -\mathcal{L}\rho(\vec{X}, t) - \frac{\partial}{\partial \vec{X}} \cdot \left[\vec{f} \rho(\vec{X}, 0) e^{-\mathcal{L}t} \right] \\ &\quad + \frac{\partial}{\partial \vec{X}} \cdot \left\{ \vec{f}(t) \int_0^t ds e^{-\mathcal{L}(t-s)} \frac{\partial}{\partial \vec{X}} \cdot [\vec{f}(s) \rho(\vec{X}, s)] \right\} \end{aligned}$$

这个迭代过程可以一直持续下去，但最终我们需要对随机力 \vec{f} 进行系综平均。由于我们假设 $f(t)$ 是高斯白噪声，所以更高阶的关联函数都可以被表示为二阶关联函数。最终，我们可以

⁷Robert Zwanzig, Nonequilibrium Statistical Mechanics.

得到⁸:

$$\frac{\partial \langle \rho(\vec{X}, t) \rangle}{\partial t} = -\mathcal{L} \langle \rho(\vec{X}, t) \rangle + \frac{1}{2} \frac{\partial}{\partial \vec{X}} \cdot \vec{D} \frac{\partial}{\partial \vec{X}} \langle \rho(\vec{X}, t) \rangle$$

把 $\langle \rho \rangle$ 记为 P , 并把 \mathcal{L} 展开, 我们就得到了最后的 F-P 方程:

$$\frac{\partial P(\vec{X}, t)}{\partial t} = -\frac{\partial}{\partial \vec{X}} \cdot \vec{v}(\vec{X}) + \frac{1}{2} \frac{\partial}{\partial \vec{X}} \cdot \vec{D} \frac{\partial}{\partial \vec{X}} P(\vec{X}, t)$$

上式 $\vec{D} \partial_{\vec{X}}$ 可以理解为对角矩阵 \vec{D} 与梯度算符的乘积, 最后得到一个矢量再与前面的梯度算符做点乘。上式也可以写成

$$\frac{\partial P(\vec{X}, t)}{\partial t} = -\nabla_{\vec{X}} \cdot \vec{v}(\vec{X}) + \frac{1}{2} \nabla_{\vec{X}} \cdot \vec{D} \nabla_{\vec{X}} P(\vec{X}, t)$$

和 Liouville 方程进行对比, F-P 方程等号右边第一项与之相同, 第二项来源于对随机力的平均。我们分别称为漂移项和扩散项。

说明: 在上面的分析中, 我们要求 Langevin 方程满足 Markov 形式, 并且噪声项为高斯、白噪声。这是得到 F-P 方程的前提条件。但在分析过程中我们并没有要求 Langevin 方程为线性方程。

举例 1: X 对应速度 v : 根据 Langevin 方程, 可以写出对应的 F-P 方程:

$$\frac{\partial}{\partial t} P(v, t) = \gamma \frac{\partial}{\partial v} \left(v + \frac{k_B T}{m} \frac{\partial}{\partial v} \right) P(v, t)$$

• **稳态解:** 在稳态下, $\frac{\partial}{\partial t} P(v, t) = 0$, 所以

$$\left(v + \frac{k_B T}{m} \frac{\partial}{\partial v} \right) P(v, t) = \text{const.} \quad (17.75)$$

这个常数只能是零, 因为我们要求上式对任意的 v 都成立, 而当 $v \rightarrow +\infty$ 时, $P \rightarrow 0$ 。

所以我们得到平衡态的时候布朗粒子服从 Maxwell-Boltzmann 分布:

$$P(v, t) \propto \exp \left[-\frac{mv^2}{2k_B T} \right]. \quad (17.76)$$

稳态解与初态无关。

• **瞬态解:** 如果给定初始条件为

$$P(v, 0) = \delta(v - v_0), \quad (17.77)$$

那么我们可以求解 Fokker-Planck 方程, 解正好就是我们前面得到的 Ornstein-Uhlenbeck 分布

$$P(v, t) = \left(\frac{m}{2\pi k_B T (1 - e^{-2\gamma t})} \right)^{\frac{1}{2}} \exp \left\{ -\frac{m(v - v_0 e^{-\gamma t})^2}{2k_B T (1 - e^{-2\gamma t})} \right\}. \quad (17.78)$$

⁸细节可以参考: Zwanzig 第 2 章或 Reichl 第 5 章

举例 2: 对于一个在束缚势限制下作一维布朗运动的粒子, 我们可以令 $\vec{X} = (x, v)$, 于是得到:

$$\begin{aligned}\frac{d}{dt}x &= v \\ \frac{d}{dt}v &= -\frac{1}{m} \frac{V(x)}{x} - \gamma v + f(t)/m,\end{aligned}$$

对应的 F-P 方程又称为 Kramers 方程

$$\frac{\partial}{\partial t}P(x, p, t) = \left[-\frac{\partial}{\partial x} \frac{p}{m} + \frac{\partial}{\partial p} \frac{\partial V}{\partial x} + \gamma \frac{\partial}{\partial p} \left(p + m k_B T \frac{\partial}{\partial p} \right) \right] P(x, p, t). \quad (17.79)$$

对应的稳态解我们可以猜测出来为 Boltzmann 分布:

$$P_{eq}(x, p) \propto \exp \left[-\frac{V(x) + p^2/2m}{k_B T} \right]. \quad (17.80)$$

第 18 章 主方程方法

18.1 经典主方程

18.1.1 随机过程

- 随机过程可以用一系列的随机变量来表示。当我们有无穷多的随机变量 X_1, X_2, \dots 时，使用离散的指标 i 就会很不方便。这时我们可以引入一个连续变量 t ，把这一系列随机变量表示为 t 的函数 $X(t)$ 。每给定一个参数 t ，就有一个对应的随机变量 $X(t)$ 。
- 对于任意给定参数 t , $X(t)$ 都有一个对应的取值 $x(t)$ 。这些 $x(t)$ 构成了随机过程 $X(t)$ 的一个实现，对应对这个随机过程的一次取样。
- 假设我们产生 N 个 $X(t)$ 的具体实现 $x^i(t), i = 1, \dots, N$ 。 $X(t)$ 的均值还可以表示为：

$$\langle X(t) \rangle = \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_i x^i(t).$$

- 给定 N 个具体实现，我们可以按照通常的方法定义随机变量 $X(t)$ 满足的概率密度分布 $P_X(x, t) = \sum_i \delta(x_i(t) - x)/N$ 。该时刻 $X(t)$ 的均值还可以表示为：

$$\langle X(t) \rangle = \int dx P_X(x, t) x.$$

以上两个均值的表达式是等价的，我们统称为 $X(t)$ 的系综平均或统计平均。

- 以上的平均是针对任意给定的一个时刻 t 进行的。类似地，我们还可以定义 $\langle X(t_1)X(t_2) \rangle$, $\langle X(t_1)X(t_2)X(t_3) \rangle$ 等多个时刻随机变量之间的平均，表示不同时刻的关联。与单时刻类似：

$$\langle X(t_1)X(t_2) \rangle = \int dx_1 dx_2 P_X(x_1, t_1; x_2, t_2) x_1 x_2.$$

或

$$= \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_i x^{(i)}(t_1) x^{(i)}(t_2).$$

理论上，我们需要给出任意一组时刻上随机变量之间的关联才能确定这个随机过程的统计性质。

- 对于**高斯**随机过程，我们只需要指定其单时刻和双时刻的关联就可以把整个过程确定下来。它的所有更高阶关联都可以用这两个关联函数表示出来。通常来说，我们最常用的

就是高斯随机过程。

- 稳态随机过程的关联函数具有时间平移不变性，也就是说它不依赖于时间的绝对值，比如，

$$\langle X(t) \rangle = \langle X(t') \rangle = \langle X \rangle, \quad \langle X(t_1)X(t_2) \rangle = \langle X(t_1 - t_2)X(0) \rangle.$$

此外，对于随机过程 $X(t)$ 的任意一组实现 $x^{(i)}(t)$ ，我们可以对其进行时间平均：

$$\lim_{T \rightarrow +\infty} \int_{t-T/2}^{t+T/2} dt' x^{(i)}(t') = \overline{X(t)}^T.$$

如果随机过程满足各态遍历的特征，那么

$$\begin{aligned}\overline{X(t)}^T &= \langle X(t) \rangle \\ \overline{X(t_1) \cdots X(t_n)}^T &= \langle X(t_1) \cdots X(t_n) \rangle.\end{aligned}$$

我们这个课程的讨论仅限于最简单的稳态、高斯随机过程，并且我们假设它满足各态遍历的特征。

下面我们进一步引入一些数学上的定义来描述一个随机过程 $X(t)$ 的概率密度分布。

- $P_1(x_1, t_1)$ 表示时刻 t_1 随机变量 X 取值为 x_1 的概率。
- $P_2(x_1, t_1; x_2, t_2)$ 表示随机变量 X 在 t_1 时刻取值 x_1 ，在 t_2 时刻取值 x_2 的联合概率密度。
- $P_n(x_1, t_1; \dots; x_n, t_n)$ 的意义以此类推。

根据这些定义我们可以得到下面的性质：

- 归一化： $\int P_1(x_1, t_1) dx_1 = 1$
- $\int P_n(x_1, t_1; \dots; x_n, t_n) dx_n = P_{n-1}(x_1, t_1; \dots; x_{n-1}, t_{n-1})$
- 平均值： $\langle x_1(t_1)x_2(t_2) \cdots x_n(t_n) \rangle = \int \cdots \int x_1 x_2 \cdots x_n P(x_1, t_1; \dots; x_n, t_n) dx_1 \cdots dx_n$

稳态随机过程满足：

$$P_n(x_1, t_1; x_2, t_2; \dots; x_n, t_n) = P_n(x_1, t_1 + \tau; x_2, t_2 + \tau; \dots; x_n, t_n + \tau). \quad (18.1)$$

下面需要定义的是条件概率分布 $P_{1|1}(x_1, t_1 | x_2, t_2)$ 。它表示的是，在已知一个随机变量 X 在 t_1 时刻的取值为 x_1 的情况下，它在 t_2 时刻的取值为 x_2 的概率¹。根据定义，我们有

$$P_{1|1}(x_1, t_1 | x_2, t_2) = \frac{P_2(x_1, t_1; x_2, t_2)}{P_1(x_1, t_1)}. \quad (18.2)$$

其归一化条件为：

$$\int P_{1|1}(x_1, t_1 | x_2, t_2) dx_2 = 1. \quad (18.3)$$

¹这里参数 1 和 2 的顺序我们采用 Reichl 书中的规则。很多其他教材采用的相反的规则，记为 $P_1(x_2, t_2 | x_1, t_1)$

类似的，我们还可以定义联合条件概率密度：

$$P_{m|n}(x_1, t_1; \dots; x_m, t_m | x_{m+1}, t_{m+1}; \dots; x_{m+n}, t_{m+n}) = \frac{P_{m+n}(x_1, t_1; \dots; x_{m+n}, t_{m+n})}{P_m(x_1, t_1; \dots; x_m, t_m)}. \quad (18.4)$$

18.1.2 Markov 过程和 Chapman-Kolmogorov 方程

Markov 过程满足：

$$P_{n-1|1}(t_1, x_1; \dots; t_{n-1}, x_{n-1} | t_n, x_n) = P_{1|1}(t_{n-1}, x_{n-1} | t_n, x_n). \quad (18.5)$$

其中，我们把 $P_{1|1}(t_{n-1}, x_{n-1} | t_n, x_n)$ 叫作跃迁几率。这说明，马尔柯夫过程在 t_n 时刻的取值概率只由前一个时刻 t_{n-1} 完全决定，与 t_{n-1} 之前时刻的历史没有任何关系，也就是没有任何记忆效应。这样的话，一个马尔柯夫过程可以由概率密度分布 $P_1(x_1)$ 和跃迁几率 $P_{1|1}(t_1, x_1 | t_2, x_2)$ 来完全确定。比如：

$$\begin{aligned} P_3(x_1, t_1; x_2, t_2; x_3, t_3) &= P_2(x_1, t_1; x_2, t_2) P_{2|1}(x_1, t_1; x_2, t_2 | x_3, t_3) \\ &= P_2(x_1, t_1; x_2, t_2) P_{1|1}(x_2, t_2 | x_3, t_3) \\ &= P_1(x_1, t_1) P_{1|1}(x_1, t_1 | x_2, t_2) P_{1|1}(x_2, t_2 | x_3, t_3). \end{aligned}$$

对于 P_n 可以类似证明。

在式子 (23.82) 两边同时对 x_2 做积分，我们可以得到：

$$P(x_1, t_1 | x_3, t_3) = \int P(x_1, t_1 | x_2, t_2) P(x_2, t_2 | x_3, t_3) dx_2. \quad (18.6)$$

这个方程叫做 **Chapman-Kolmogorov (C-K) 方程**。这里我们把符号做了简化，把 P 的下标 $1|1$ 省略了。这个方程理解起来非常简单。方程的右边表示粒子从 t_1 时刻的位置 x_1 到达 t_3 时刻的位置 x_3 必然要在中间时刻 t_2 处于某一个位置 x_2 。学过费曼路径积分的同学马上就意识到路径积分的基本出发点即此（图18.1）。另外，随机过程的马尔柯夫特性表现在这些跃迁几率只依赖于起始和终止时刻的位置，与粒子如何到达起始时刻的没有关系。

18.1.3 主方程

C-K 方程是一个非线性的积分方程。在特定的条件下，我们可以从它出发来得到一组描述 Markov 过程跃迁概率的线性方程组，称为主方程。我们的出发点是假设 Markov 过程在时间微元 Δt 内可以写成如下形式：

$$P(q', t | q, t + \Delta t) = \delta(q' - q)(1 - \gamma(q')\Delta t) + w_{q'q}\Delta t + O(\Delta t^2). \quad (18.7)$$

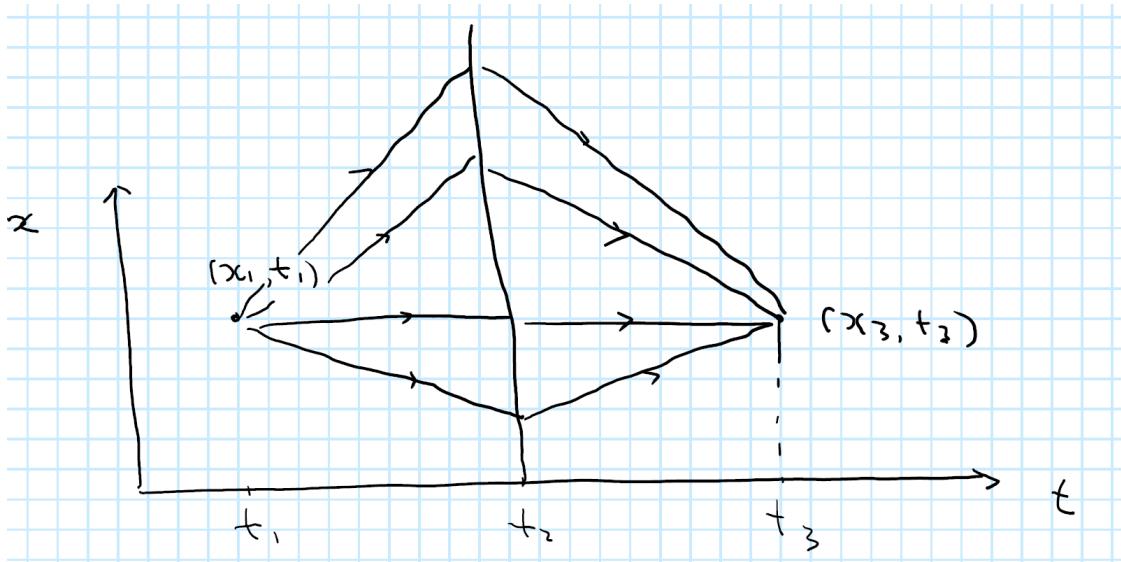


图 18.1: Chapman-Kolmogorov 方程的图解。

这个方程的左边是一个条件概率, 即已知粒子在 t 时刻位于位置 q' 的话, 其在 $t + \Delta t$ 时刻位于位置 q 的概率。方程的右边两项描述了粒子在 $t + \Delta t$ 时刻到达 q 的两种方式: (1) 第二项表示粒子在 Δt 的时间内从 q' 跃迁到 q 的概率. 其中 $w_{q'q} = \langle q'|w|q\rangle$ 表示跃迁速率, 而 $w_{q'q}\Delta t$ 为跃迁概率。这里我们假设它不依赖时间. (2) 第一项对应 $q = q'$ 的情况, 那么只需要粒子保持在原来的位置就可以了, 这个概率为 $(1 - \gamma(q')\Delta t)$, 而 $\gamma(q') = \int w_{q'q}dq$, $\gamma(q')\Delta t$ 表示粒子从 q' 跃迁到其他任何状态的概率。

我们把上面的式子带入到 C-K 方程

$$P(q', t_1 | q, t) = \int P(q', t_1 | q'', t - \Delta t) P(q'', t - \Delta t | q, t) dq''. \quad (18.8)$$

可以得到:

$$\begin{aligned} &= \int P(q', t_1 | q'', t - \Delta t) [\delta(q'' - q)(1 - \gamma(q'')\Delta t) + w_{q''q}\Delta t] dq'' \\ &= P(q', t_1 | q, t - \Delta t)(1 - \gamma(q)\Delta t) + \Delta t \int P(q', t_1 | q'', t - \Delta t) w_{q''q} dq''. \end{aligned}$$

整理得到微分方程的形式:

$$\begin{aligned} \frac{\partial}{\partial t} P(q', t_1 | q, t) &= -\gamma(q)P(q', t_1 | q, t) + \int w_{q''q} P(q', t_1 | q'', t) dq'' \\ &= - \int w(q, q'') P(q', t_1 | q, t) dq'' + \int w_{q''q} P(q', t_1 | q'', t) dq''. \end{aligned}$$

这是 Markov 过程中的跃迁概率 $P(q', t_1 | q, t)$ 所满足的主方程。这个式子的物理意义比较容易理解: 等式右边第一项表示系统离开 q 状态的速率, 第二项表示从其他态跃迁到 q 态的速率。

我们也可以对初始条件做积分

$$\begin{aligned}\frac{\partial}{\partial t}P(q,t) &= -\gamma(q)P(q,t) + \int w_{q''q}P(q'',t)dq'' \\ &= -\int w_{qq''}P(q,t)dq'' + \int w_{q''q}P(q'',t)dq''.\end{aligned}$$

如果状态空间是离散的，则主方程具有下面的离散形式：

$$\frac{\partial}{\partial t}P(q,t) = -\sum_{q''}W(q,q'';t)P(q,t) + \sum_{q''}W(q'',q;t)P(q'',t) \quad (18.9)$$

当系统达到稳定的状态后，应该满足**细致平衡**的条件：

$$W(q,q'';t)P(q,t) = W(q'',q;t)P(q'',t).$$

18.1.4 Fokker-Planck 方程

我们上一章得到的 F-P 方程也可以通过主方程来得到。为此，我们考虑系统在态空间的一个小的位移，定义

$$w_{q''q} = w(q'', q - q'') \equiv \Delta q$$

代入到主方程得到：

$$\frac{\partial}{\partial t}P(q',t_1|q,t) = -\left(\int w(q,\Delta q)d\Delta q\right)P(q',t_1|q,t) + \int w(q-\Delta q,\Delta q)P(q',t_1|q-\Delta q,t)d\Delta q. \quad (18.10)$$

考虑 $\Delta q \ll q$ ，我们可以对 Δq 做展开：

$$\sum_{n=0}^{\infty} \frac{(-r)^n}{n!} \frac{\partial^n}{\partial q^n} f(q) = f(q-r), \quad (18.11)$$

这被称为 Kramers-Moyal 展开。代入到主方程可以得到：

$$\frac{\partial}{\partial t}P(q',t_1|q,t) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial q^n} \left[P(q',t_1|q,t) \int (\Delta q)^n w(q,\Delta q) dr \right]. \quad (18.12)$$

定义

$$\alpha_n(q) \equiv \int d\Delta q (\Delta q)^n w(q,\Delta q) \quad (18.13)$$

$$= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int d\Delta q (\Delta q)^n P(q,t|q+\Delta q,t+\Delta t) \quad (18.14)$$

$$= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \langle \Delta q^n \rangle_q, \quad (18.15)$$

我们最后得到 Kramers-Moyal 展开形式的主方程:

$$\frac{\partial}{\partial t} P(q', t_1 | q, t) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial q^n} [P(q', t_1 | q, t) \alpha_n(q)]. \quad (18.16)$$

如果当 $n \geq 3$ 时 $\alpha_n = 0$, 那么上面的方程就是 F-P。稳态的高斯 Markov 过程满足以上条件。可见, F-P 方程是一种特殊形式的主方程。定义概率流:

$$j(q', t_1 | q, t) = P(q', t_1 | q, t) \alpha_1(q) - \frac{1}{2} \partial_q [P(q', t_1 | q, t) \alpha_2].$$

这里, 第一项为漂移项, 第二项为扩散项。我们得到相应的连续性方程:

$$\frac{\partial}{\partial t} P(q', t | q, t) = -\partial_q j(q', t | q, t).$$

18.2 Lindblad 量子主方程

在量子系统中, 与概率密度分布对应的是密度矩阵。我们可以基于描述密度矩阵演化的量子 Liouville 方程, 通过对库自由度求迹的方式来得到描述系统约化密度矩阵的运动方程, 即量子主方程。这可以与经典主方程或 F-P 方程相对应。下面我们介绍在研究量子开放系统动力学时被广泛使用的具有 Markov 性质的 Lindblad 量子主方程。它具有非常好的数学性质, 被广泛应用于统计物理、凝聚态物理、量子光学等研究领域²。

18.2.1 Lindblad 主方程的推导

与前面的讨论类似, 我们将整个体系分为系统和环境或库两部分。这样我们就可以把整个哈密顿量拆分成下面的形式:

$$\hat{H} = \hat{H}_s + \hat{H}_b + \hat{V} = \hat{H}_0 + \hat{V}. \quad (18.17)$$

与 \hat{H} 对应的密度矩阵我们表示为 $\hat{\rho}^S(t)$, 它满足量子 Liouville 方程或者 von Neumann 方程(式10.7)。在相互作用绘景下, 运动方程具有更加简化的形式 ($\hbar = 1$):

$$\dot{\hat{\rho}}(t) = -i [\hat{V}(t), \hat{\rho}(t)], \quad (18.18)$$

这里, $\hat{V}(t) = e^{i\hat{H}_0 t} \hat{V} e^{-i\hat{H}_0 t}$, $\hat{\rho}(t) = e^{i\hat{H}_0 t} \hat{\rho}^S(t) e^{-i\hat{H}_0 t}$, $\hat{\rho}^S$ 是整个宇宙 Schrödinger 绘景下的密度矩阵。我们这里只考虑不含时的哈密顿量。上式有一个形式解:

$$\hat{\rho}(t) = \hat{\rho}(0) - i \int_0^t ds [\hat{V}(s), \hat{\rho}(s)]. \quad (18.19)$$

²Göran Lindblad, “On the Generators of Quantum dynamical semigroups”, Commun. Math. Phys. 48, 119 (1976).

我们可以把上式右边的 $\hat{\rho}(s)$ 再做替换，就可以得到：

$$\dot{\rho}(t) = -i[\hat{V}(t), \hat{\rho}_0] - i \int_0^t ds [\hat{V}(t), [\hat{V}(s), \hat{\rho}(s)]]]. \quad (18.20)$$

为了简化，下面我们省略算符符号 $\hat{\cdot}$ 。如果我们选择系统与库的相互作用满足：

$$\text{Tr}_b \{ [V(t), \rho(0)] \} = 0, \quad (18.21)$$

通过对库自由度求迹的办法我们可以得到描述系统的约化密度矩阵 $\rho_s(t) = \text{Tr}_b \{ \rho(t) \}$ 所满足的方程：

$$\dot{\rho}_s(t) = -i \int_0^t ds \text{Tr}_b \{ [V(t), [V(s), \rho(s)]] \} \quad (18.22)$$

为了更进一步，我们假设系统与库的耦合非常弱。这样，我们就可以将相互作用保留到二阶。注意，虽然 (18.22) 形式上看起来已经是 $V(t)$ 的二阶形式了，但实际上 $\rho(s)$ 中还隐含着更高阶的项。为此，我们可以对 $\rho(t)$ 做更进一步近似

$$\rho(t) \approx \rho_s(t) \otimes \rho_b. \quad (18.23)$$

也就是说，我们忽略系统和库之间的关联，而将 $\rho(t)$ 直接写成两者直积的形式。这个近似通常称为 **Born 近似**。代入到式 (18.22)，我们就可以得到一个封闭形式的运动方程

$$\dot{\rho}_s(t) = -i \int_0^t ds \text{Tr}_b \{ [V(t), [V(s), \rho_s(s) \otimes \rho_b]] \}. \quad (18.24)$$

下面我们要做的是 **Markov 近似**。上式右边是一个卷积的形式，具有记忆效应，因此我们需要知道 ρ_s 从初始时刻到 t 时刻的所有信息。为了将其转化为 Markov 形式，我们假设库自由度的时间尺度要远快于系统自由度³，库自由度的时间关联长度较系统演化时间尺度要短得多。我们将 $\rho_s(s)$ 替换为 $\rho_s(t)$

$$\dot{\rho}_s(t) = -i \int_0^t ds \text{Tr}_b \{ [V(t), [V(s), \rho_s(t) \otimes \rho_b]] \}. \quad (18.25)$$

上式称为 **Redfield 主方程**。下面，我们做一个变量替换 $s \rightarrow t - s$

$$\int_0^t ds \cdots V(s) \rightarrow \int_0^t ds \cdots V(t - s)$$

同样基于库和系统自由度之间时间尺度的分离，我们可以将积分上限扩展到 $+\infty$ ，从而得到：

$$\dot{\rho}_s(t) = -i \int_0^\infty ds \text{Tr}_b \{ [V(t), [V(t - s), \rho_s(t) \otimes \rho_b]] \}, \quad (18.26)$$

从而得到 Markov 形式的主方程。

为了得到 Lindblad 形式的主方程，我们需要假设系统与库的耦合哈密顿量可以写成两者

³与推导 Langevin 方程的情况类似。

算符直积的形式：

$$V = \sum_i S_i \otimes B_i, \quad (18.27)$$

这里 i 表示不同的耦合机制。我们同时要求

$$[H_s, S_i] = -\epsilon_i S_i. \quad (18.28)$$

在以上条件下，我们有

$$V(t) \equiv e^{iH_0 t} V e^{-iH_0 t} = \sum_i e^{-i\epsilon_i t} S_i \otimes B_i(t) = V^\dagger(t) = \sum_j e^{i\epsilon_j t} S_j^\dagger \otimes B_j^\dagger(t), \quad (18.29)$$

这里 $B_i(t) = e^{iH_b t} B_i e^{-iH_b t}$. 将 $V(t)$ 代入到主方程中：

$$\begin{aligned} & \dot{\rho}_s(t) \\ &= - \int_0^\infty ds \text{Tr}_b \{ [V(t), [V(t-s), \rho_s(t) \otimes \rho_b]] \} \\ &= \int_0^\infty ds \text{Tr}_b \{ V(t-s) \rho_s(t) \otimes \rho_b V(t) - V(t) V(t-s) \rho_s \otimes \rho_b \} + \text{h.c.} \\ &= \int_0^\infty ds \text{Tr}_b \{ V(t-s) \rho_s(t) \otimes \rho_b V^\dagger(t) - V^\dagger(t) V(t-s) \rho_s \otimes \rho_b \} + \text{h.c.} \\ &= \int_0^\infty ds \text{Tr}_b \left\{ \left(\sum_i e^{-i\epsilon_i(t-s)} S_i \otimes B_i(t-s) \right) \rho_s(t) \otimes \rho_b \left(\sum_j e^{i\epsilon_j t} S_j^\dagger \otimes B_j^\dagger(t) \right) \right. \\ &\quad \left. - \left(\sum_j e^{i\epsilon_j t} S_j^\dagger \otimes B_j^\dagger(t) \right) \left(\sum_i e^{-i\epsilon_i(t-s)} S_i \otimes B_i(t-s) \right) \rho_s(t) \otimes \rho_b \right\} + \text{h.c.} \quad (18.30) \\ &= \sum_{i,j} \int_0^\infty ds e^{i\epsilon_i s} \langle B_j^\dagger(t) B_i(t-s) \rangle e^{-i(\epsilon_i - \epsilon_j)t} \left(S_i \rho_s(t) S_j^\dagger - S_j^\dagger S_i \rho_s(t) \right) + \text{h.c.} \\ &\stackrel{\text{RWA}}{=} \sum_i \int_0^\infty ds e^{i\epsilon_i s} \langle B_i^\dagger(t) B_i(t-s) \rangle \left(S_i \rho_s(t) S_i^\dagger - S_i^\dagger S_i \rho_s(t) \right) + \text{h.c.} \\ &= \sum_i \int_0^\infty ds e^{i\epsilon_i s} \langle B_i^\dagger(s) B_i(0) \rangle \left(S_i \rho_s(t) S_i^\dagger - S_i^\dagger S_i \rho_s(t) \right) + \text{h.c.} \end{aligned}$$

在倒数第二行，我们使用了旋波近似 (Rotating Wave Approximation, RWA) 从而忽略了 $\epsilon_i \neq \epsilon_j$ 的情况。定义

$$\Gamma(\epsilon_i) \equiv \int_0^\infty ds e^{i\epsilon_i s} \langle B_i^\dagger(s) B_i(0) \rangle \equiv \frac{1}{2} \gamma(\epsilon_i) + i\Delta(\epsilon_i), \quad (18.31)$$

这里 $\gamma(\epsilon_i) = 2 \operatorname{Re}\{\Gamma(\epsilon_i)\}$ and $\Delta(\epsilon_i) = \operatorname{Im}\{\Gamma(\epsilon_i)\}$ 。于是，

$$\begin{aligned}
\dot{\rho}_s(t) &= \sum_i \Gamma(\epsilon_i) \left(S_i \rho_s(t) S_i^\dagger - S_i^\dagger S_i \rho_s(t) \right) + \text{h.c.} \\
&= \sum_i \Gamma(\epsilon_i) \left(S_i \rho_s(t) S_i^\dagger - S_i^\dagger S_i \rho_s(t) \right) + \sum_i \Gamma^*(\epsilon_i) \left(S_i \rho_s(t) S_i^\dagger - \rho_s(t) S_i^\dagger S_i \right) \\
&= \sum_i \Gamma(\epsilon_i) \left(S_i \rho_s(t) S_i^\dagger - S_i^\dagger S_i \rho_s(t) \right) + \sum_i \Gamma^*(\epsilon_i) \left(S_i \rho_s(t) S_i^\dagger - \rho_s(t) S_i^\dagger S_i \right) \\
&= \sum_i \left\{ (\Gamma(\epsilon_i) + \Gamma^*(\epsilon_i)) S_i \rho_s S_i^\dagger - \left(\frac{1}{2} \gamma(\epsilon_i) \right. \right. \\
&\quad \left. \left. + i \Delta(\epsilon_i) \right) S_i^\dagger S_i \rho_s(t) - \left(\frac{1}{2} \gamma(\epsilon_i) - i \Delta(\epsilon_i) \right) \rho_s(t) S_i^\dagger S_i \right\} \\
&= -i \left[\sum_i \Delta(\epsilon_i) S_i^\dagger S_i, \rho_s(t) \right] + \sum_i \gamma(\epsilon_i) \left(S_i \rho_s(t) S_i^\dagger - \frac{1}{2} \{ S_i^\dagger S_i, \rho_s(t) \} \right) \\
&= -i [H_{LS}, \rho_s(t)] + \sum_i \gamma(\epsilon_i) \left(S_i \rho_s(t) S_i^\dagger - \frac{1}{2} \{ S_i^\dagger S_i, \rho_s(t) \} \right),
\end{aligned} \tag{18.32}$$

这里

$$H_{LS} = \sum_i \Delta(\epsilon_i) S_i^\dagger S_i$$

表示由于系统与库的耦合而导致的对系统哈密顿量的修正。在弱耦合情况下，这个数值是很小的，因此我们可以将其忽略。重新转换到 Schrödinger 绘景下，我们得到 Lindblad 形式的主方程：

$$\dot{\rho}_s(t) = -i [H_s, \rho_s(t)] + \sum_i \gamma(\epsilon_i) \left(S_i \rho_s(t) S_i^\dagger - \frac{1}{2} \{ S_i^\dagger S_i, \rho_s(t) \} \right). \tag{18.33}$$

通常我们还定义耗散算符：

$$\mathcal{D}_S[\rho] = S \rho(t) S^\dagger - \frac{1}{2} \{ S^\dagger S, \rho \}$$

将上式简化为：

$$\dot{\rho}_s(t) = -i [H_s, \rho_s(t)] + \sum_i \gamma(\epsilon_i) \mathcal{D}_{S_i}[\rho_s(t)]. \tag{18.34}$$

上式右边第一项表示系统本身的演化，第二项表示由于与库的耦合而导致的耗散。

18.2.2 一个具体实例

谐振子 + 谐振子库：考虑式 (17.48) 所描述的模型，假设系统只包含一个谐振子。我们把

哈密顿量重新写出来 ($\hbar = 1$):

$$\begin{aligned} H &= H_s + H_b + V, \\ H_s &= \Omega a^\dagger a \\ H_b &= \sum_l \omega_l b_l^\dagger b_l \\ V &= a \cdot \sum_l t_l (b_l^\dagger + b_l) + \text{h.c..} \end{aligned} \tag{18.35}$$

我们定义 $S_0 = a$, $B_0 = \sum_l t_l (b_l^\dagger + b_l)$, $\epsilon_0 = \Omega$, $S_1 = a^\dagger$, $B_1 = \sum_l t_l^* (b_l^\dagger + b_l)$, $\epsilon_1 = -\Omega$ 。Lindblad 主方程为:

$$\dot{\rho}_S(t) = -i \left[\sum_{i=0,1} \Delta(\epsilon_i) S_i^\dagger S_i, \rho_S(t) \right] + \sum_i \gamma(\epsilon_i) \left(S_i \rho_S(t) S_i^\dagger - \frac{1}{2} \{ S_i^\dagger S_i, \rho_S(t) \} \right), \tag{18.36}$$

这里 $\gamma(\epsilon_i) = 2 \operatorname{Re}\{\Gamma(\epsilon_i)\}$, $\Delta(\epsilon_i) = \operatorname{Im}\{\Gamma(\epsilon_i)\}$

$$\Gamma(\epsilon_i) = \int_0^\infty ds e^{i\epsilon_i s} \langle B_i^\dagger(s) B_i(0) \rangle.$$

考虑库处于热平衡的状态，我们有:

$$\begin{aligned} \Gamma(\epsilon_i) &= J(\epsilon_i)(1 + n_B(\epsilon_i)) + J(-\epsilon_i)n_B(-\epsilon_i) \\ &\quad + iP \int_{-\infty}^\infty \frac{d\omega}{\pi} J(\omega) \frac{1 + n_B(\omega)}{\epsilon_i - \omega} + iP \int_{-\infty}^\infty \frac{d\omega}{\pi} J(\omega) \frac{n_B(\omega)}{\epsilon_i + \omega}. \end{aligned}$$

我们定义了谱函数:

$$J(\omega) = \pi \sum_l |t_l|^2 \delta(\omega - \omega_l).$$

忽略 H_{LS} , 我们可以得到 Schrödinger 绘景下的主方程:

$$\dot{\rho}_S(t) = -i[H_S, \rho(t)] + \kappa(1 + n_B(\Omega)) \mathcal{D}_a[\rho_S(t)] + \kappa n_B(\Omega) \mathcal{D}_{a^\dagger}[\rho_S(t)], \tag{18.37}$$

我们利用了 $n_B(-\omega) = -(1 + n_B(\omega))$, 定义了 $\kappa = 2(J(\omega) - J(-\omega))$

练习 18.1 验证, 平衡态密度矩阵 $\rho_{eq} \propto \sum_n e^{-\beta(n+1/2)\hbar\Omega} |n\rangle\langle n|$ 为稳态 Lindblad 主方程的解。

练习 18.2 考虑一个二能级系统与一个温度为 T 的热库耦合。热库用一系列谐振子来描述。对应的哈密顿量为:

$$\begin{aligned} H_s &= \frac{1}{2} \Omega \sigma_z, \\ V &= \sum_l (t_l b_l^\dagger \sigma_- + h.c.) \end{aligned}$$

H_b 为式 (18.35) 的形式。其中, $\sigma_- = (\sigma_x - i\sigma_y)/2$, $\sigma_{x/y/z}$ 为 Pauli 矩阵。写出对应的 Lindblad 主方程，并分析密度矩阵如何从任意初态演化到热平衡态。

18.2.3 对主方程的分析

下面我们回到式 (18.33-18.34) 进一步理解一下 Lindblad 主方程。为了简化形式，我们省略下标 s，而且只考虑一种耦合。将其写成如下离散的形式：

$$\rho(t+dt) = \rho(t) - iH_{\text{eff}}\rho(t)dt + i\rho(t)H_{\text{eff}}^\dagger dt + \gamma S\rho(t)S^\dagger dt$$

我们定义了一个非厄密的有效哈密顿量：

$$H_{\text{eff}} = H - \frac{i}{2}\gamma S^\dagger S$$

这样我们可以把主方程的演化分为两部分：(1) 基于 H_{eff} 的确定性“可逆”演化；(2) 不可逆的量子跳跃 (quantum jump) 过程。它们分别对应上式的前三项和最后一项。首先，我们定义量子跳跃部分：

$$\rho(t) \rightarrow \frac{S\rho(t)S^\dagger}{\text{Tr}(S^\dagger S\rho(t))} \equiv \rho_{\text{irrev}}(t)$$

对应的概率定义为：

$$p(t)dt = \text{Tr}(S^\dagger S\rho(t))\gamma dt.$$

这样的话，上式最后一项可以表示为：

$$\rho_{\text{irrev}} p(t)dt$$

下面我们看前面三项。当 $dt \rightarrow 0$ 时，我们可以将其近似写为：

$$\rho(t) \rightarrow \frac{e^{-iH_{\text{eff}}dt}\rho(t)e^{iH_{\text{eff}}dt}}{\text{Tr}(e^{-iH_{\text{eff}}dt}\rho(t)e^{iH_{\text{eff}}dt})} \equiv \rho_{\text{rev}}(t)$$

于是前三项可以表示为：

$$\rho_{\text{rev}}(1 - p(t)dt)$$

综合起来，密度矩阵通过 Lindblad 主方程的演化可以表示为：

$$\rho(t+dt) = \rho_{\text{rev}}(t)(1 - p(t)dt) + \rho_{\text{irrev}}p(t)dt \quad (18.38)$$

这个式子的物理意义非常清晰。它表示在每个时间间隔 dt ，系统每步以 $(1 - p(t)dt)$ 的概率根据 H_{eff} 发生确定性演化，以 $p(t)dt$ 的概率发生不可逆的量子跳跃过程。与经典的 F-P 方程对比，第一项相当于 F-P 方程的漂移项，第二项相当于其扩散项。

第 19 章 分子动理论

动理论从组成宏观体系的大量粒子的微观运动方程出发来研究体系的宏观物理量性质。我们以 Boltzmann 最初考虑的稀薄气体为原型介绍这一理论。具体来说，我们从描述系统每个自由度的经典 Hamilton 运动方程出发，来得到 Boltzmann 方程，再从 Boltzmann 方程出发只考虑几个宏观守恒量，来得到描述这些守恒量的流体动力学方程。

19.1 概率密度分布

要描述 N 个粒子组成经典体系，我们可以跟踪每一个粒子在相空间 (\vec{r}_i, \vec{p}_i) 内随时间的变化。具体来说，我们可以去求解每个粒子或者自由度的运动方程。我们也可以使用我们上一章引入的概率密度分布 $P(\vec{r}_1, \vec{p}_1, \dots, \vec{r}_N, \vec{p}_N, t)$ 来描述，它表示在时刻 t ，粒子 i 在 (\vec{r}_i, \vec{p}_i) 的概率密度。有了概率密度分布，我们就可以计算物理量的系综平均

$$\langle A \rangle = \int d\Gamma P(\Gamma, t) A. \quad (19.1)$$

19.2 Liouville 定理, 方程

Liouville 定理是说，对于 Hamilton 体系，相空间密度就像一个不可压缩的液体。也就是说，如果我们在时刻 t 选取相空间的一个小的体积元 $d\vec{\Gamma} = \Pi_i d\vec{\Gamma}_i = \Pi_i d\vec{r}_i d\vec{p}_i$ ，那么由于相空间坐标 (\vec{r}, \vec{p}) 满足 Hamilton 运动方程，在时刻 t' 这个体积元会移动到了相空间的另一个位置，可能它的形状也发生了变化，但它的体积是不变的。

根据 Liouville 定理，我们有：

$$P(\vec{r}_1, \vec{p}_1, \dots, \vec{r}_N, \vec{p}_N, t) = P(\vec{r}_1 + \Delta\vec{r}_1, \vec{p}_1 + \Delta\vec{p}_1, \dots, \vec{r}_N + \Delta\vec{r}_N, \vec{p}_N + \Delta\vec{p}_N, t + \Delta t). \quad (19.2)$$

把上式右边做 Taylor 展开：

$$\begin{aligned} P(\vec{r} + \Delta\vec{r}, \vec{p} + \Delta\vec{p}, t + \Delta t) &= P(\vec{r}, \vec{p}, t) + \left(\frac{\partial P}{\partial t} + \sum_i \left(\frac{\partial P}{\partial \vec{r}_i} \dot{\vec{r}}_i + \frac{\partial P}{\partial \vec{p}_i} \dot{\vec{p}}_i \right) \right) \Delta t \\ &\equiv P(\vec{r}, \vec{p}, t) + \frac{DP}{Dt} \Delta t. \end{aligned} \quad (19.3)$$

再根据 Hamilton 运动方程:

$$\dot{\vec{r}}_i = \frac{\partial H}{\partial \vec{p}_i}, \quad (19.4)$$

$$\dot{\vec{p}}_i = -\frac{\partial H}{\partial \vec{r}_i}, \quad (19.5)$$

我们得到:

$$\boxed{\frac{\partial P}{\partial t} = \sum_i \left(\frac{\partial H}{\partial \vec{r}_i} \frac{\partial P}{\partial \vec{p}_i} - \frac{\partial H}{\partial \vec{p}_i} \frac{\partial P}{\partial \vec{r}_i} \right) \equiv \{H, P\}.} \quad (19.6)$$

这里, 我们定义了 Poisson 括号。此外, 我们还得到

$$\frac{DP}{Dt} = 0. \quad (19.7)$$

这里, D/Dt 表示全微分. 我们需要把它和 $\partial P/\partial t$ 区分. 具体来说, $\partial P/\partial t$ 表示在相空间某一个固定位置的概率密度随时间的变化. 当系统处于平衡态 (或者非平衡稳态) 的时候, $\partial P/\partial t = 0$. 而全微分 DP/Dt 可以理解为以运动的相空间为参考系, 再去看概率密度分布的变化. Liouville 定理保证了 $DP/Dt = 0$, 不管系统是否处于平衡态.

根据 Liouville 方程我们可以得到:

(1). 物理量 $A(\vec{\Gamma}) = A(\vec{r}, \vec{p})$ 的期待值随时间的变化 根据式 (19.1), 我们有:

$$\frac{d}{dt} \langle A(\vec{\Gamma}) \rangle = \int \frac{\partial}{\partial t} P(\vec{\Gamma}, t) A(\vec{\Gamma}) d\vec{\Gamma} \quad (19.8)$$

$$= \int \{H, P\} A(\vec{\Gamma}) d\vec{\Gamma} = \sum_i \int \left(\frac{\partial H}{\partial \vec{r}_i} \frac{\partial P}{\partial \vec{p}_i} - \frac{\partial H}{\partial \vec{p}_i} \frac{\partial P}{\partial \vec{r}_i} \right) A(\vec{\Gamma}) d\vec{\Gamma}. \quad (19.9)$$

因为我们要对整个相空间做积分, 所以我们可以利用分部积分. 而 $P(\vec{\Gamma}, t)$ 当 $\vec{r}_i \rightarrow \infty$ 或者 $\vec{p}_i \rightarrow \infty$ 的时候等于零, 所以分部积分的边界项为零. 所以

$$\frac{d}{dt} \langle A(\vec{\Gamma}) \rangle = \sum_i \int \left(\frac{\partial H}{\partial \vec{r}_i} \frac{\partial P}{\partial \vec{p}_i} - \frac{\partial H}{\partial \vec{p}_i} \frac{\partial P}{\partial \vec{r}_i} \right) A(\vec{\Gamma}) \Pi_i d\vec{r}_i d\vec{p}_i \quad (19.10)$$

$$= - \sum_i \int P(\vec{\Gamma}) \left[\left(\frac{\partial H}{\partial \vec{r}_i \partial \vec{p}_i} - \frac{\partial H}{\partial \vec{p}_i \partial \vec{r}_i} \right) A(\vec{\Gamma}) + \left(\frac{\partial H}{\partial \vec{r}_i} \frac{\partial A}{\partial \vec{p}_i} - \frac{\partial H}{\partial \vec{p}_i} \frac{\partial A}{\partial \vec{r}_i} \right) \right] d\vec{\Gamma} \quad (19.11)$$

$$= \int P(\vec{\Gamma}) \{A(\vec{\Gamma}), H\} d\vec{\Gamma} = \langle \{A(\vec{\Gamma}), H\} \rangle. \quad (19.12)$$

即, 如果 $\{A, H\} = 0$, 那么 A 的平均值不随时间变化. 但 $\{A, H\} = 0$ 正说明 A 是一个守恒量, 所以这个结果很容易理解.

(2). 如果 P 可以表示成 H 或者其他守恒量的函数, 那么 $\partial P/\partial t = 0$. 这是因为

$$\frac{\partial P(H)}{\partial t} = \{H, P(H)\} = \frac{\partial P}{\partial H} \left(\frac{\partial H}{\partial \vec{r}_i} \frac{\partial H}{\partial \vec{p}_i} - \frac{\partial H}{\partial \vec{p}_i} \frac{\partial H}{\partial \vec{r}_i} \right) = 0. \quad (19.13)$$

19.3 BBGKY 链式方程组

19.3.1 约化概率密度分布及粒子密度分布

我们上面定义的概率密度分布包含了 $6N + 1$ 个变量, 所以它是一个非常复杂的函数。首先, 要得到这个分布不太容易; 其次, 即使我们已经得到了这个分布, 使用起来也不太方便。比如, 如果我们的体系中有 10^{23} 个粒子, 那么概率密度分布函数就有对应这么多变量。即使存储这样一个函数, 也需要耗费巨大的空间, 更不要说根据它来得到我们关心的物理量的信息。我们希望能够通过合理的近似来使用稍微简单的分布函数来描述我们的系统. 就像上一章我们推导 Langevin 方程的方法一样, 我们可以设法消除一些自由度。为此, 我们可以定义约化概率密度分布 $P_s(\vec{r}_1, \vec{p}_1, \dots, \vec{r}_s, \vec{p}_s, t)$:

$$P_s(\vec{r}_1, \vec{p}_1, \dots, \vec{r}_s, \vec{p}_s, t) = \int P(\vec{r}_1, \vec{p}_1, \dots, \vec{r}_N, \vec{p}_N, t) \Pi_{i=s+1}^N d\Gamma_i, \quad (19.14)$$

即, 我们把第 $s + 1$ 到 N 粒子的自由度积分掉, 只保留 s 个粒子的自由度. 特别是当粒子浓度比较低的时候, 我们希望能够用 $P_1(\vec{r}_1, \vec{p}_1, t)$ 来描述整个 N 粒子体系, 这样可以把参数空间大大减小为 7 个自由度. 同时, 我们还引入粒子密度分布:

$$f_s(\vec{r}_1, \vec{p}_1, \dots, \vec{r}_s, \vec{p}_s, t) = \frac{N!}{(N-s)!} P_s(\vec{r}_1, \vec{p}_1, \dots, \vec{r}_s, \vec{p}_s, t). \quad (19.15)$$

这样 $f_1(\vec{r}_1, \vec{p}_1, t)$ 就表示粒子数密度

$$\int f_1(\vec{r}_1, \vec{p}_1, t) d\vec{r}_1 d\vec{p}_1 = N. \quad (19.16)$$

当然, 要使用我们定义的这些约化概率密度分布, 我们需要知道它们满足的运动方程. 如果我们可以求解他们的运动方程得到它们的解, 那么问题就简单多了.

19.3.2 Bogoliubov-Born-Green-Kirkwood-Yvan (BBGKY) 方程链

首先, 我们需要定义系统的 Hamilton 量

$$H = \sum_{i=1}^N \frac{\vec{p}_i^2}{2m} + \sum_i U_i(\vec{r}_i) + \sum_{i < j} V_{ij}(|\vec{r}_i - \vec{r}_j|) = \sum_i H_i + \sum_{i < j} V_{ij}(|\vec{r}_i - \vec{r}_j|). \quad (19.17)$$

第一项为动能项, 第二项为单粒子的势能项, 第三项为两体相互作用项. 这里我们忽略三体及以上的多体相互作用。这个近似对于稀薄粒子体系是合理的。为了计算 P_s 满足的运动方程,

我们还可以把 Hamilton 量写成:

$$H = H_s + H_{N-s} + H_{s,N-s}, \quad (19.18)$$

$$H_s = \sum_{i=1}^s H_i + \sum_{i < j \leq s} V_{ij}, \quad (19.19)$$

$$H_{N-s} = \sum_{i=s+1}^N H_i + \sum_{s < i < j \leq N} V_{ij}, \quad (19.20)$$

$$H_{s,N-s} = \sum_{i \leq s < j} V_{ij}. \quad (19.21)$$

这一对体系自由度的划分仍然和推导 Langevin 方程的思路类似。从 Liouville 方程出发, 我们对第 $s+1$ 到第 N 个粒子自由度积分得到

$$\int \frac{\partial}{\partial t} P(\Gamma, t) (\Pi_{i=s+1}^N d\Gamma_i) = \frac{\partial}{\partial t} P_s = \int \{H_s + H_{N-s} + H_{s,N-s}, P\} (\Pi_{i=s+1}^N d\Gamma_i). \quad (19.22)$$

上式的 Poisson 括号里包括三项, 我们分别计算:

19.3.2.1 $\{H_s, P\}$

因为 H_s 里面不包含 $s+1 \leq i \leq N$ 中粒子的自由度, 所以我们可以直接进行积分得到

$$\int \{H_s, P\} (\Pi_{i=s+1}^N d\Gamma_i) = \{H_s, P_s\}. \quad (19.23)$$

19.3.2.2 $\{H_{N-s}, P\}$

对于第二项, 因为我们要对 $s+1 \leq i \leq N$ 的自由度进行积分, 所以我们可以利用分部积分:

$$\int \{H_{N-s}, P\} (\Pi_{i=s+1}^N d\Gamma_i) = \int \sum_{j=s+1}^N \sum_{k>j} \{H_j + V_{jk}, P\} (\Pi_{i=s+1}^N d\Gamma_i) = 0. \quad (19.24)$$

19.3.2.3 $\{H_{s,N-s}, P\}$

$$\int \{H_{s,N-s}, P\} (\Pi_{i=s+1}^N d\Gamma_i) = \int \sum_{j=1}^s \sum_{k=s+1}^N \{V_{jk}(|\vec{r}_j - \vec{r}_k|), P\} (\Pi_{i=s+1}^N d\Gamma_i) \quad (19.25)$$

首先, 粒子 j 与 $s+1 \leq k \leq N$ 中的任何一个粒子的相互作用我们认为是相同的, 所以我们可以只考虑 j 与其中一个粒子的相互作用, 比如第 $s+1$ 个, 于是:

$$\int \{H_{s,N-s}, P\} (\Pi_{i=s+1}^N d\Gamma_i) = (N-s) \int \sum_{j=1}^s \{V_{j,s+1}(|\vec{r}_j - \vec{r}_{s+1}|), P\} (\Pi_{i=s+1}^N d\Gamma_i). \quad (19.26)$$

这样我们就可以对 $s+2 \leq s \leq N$ 个粒子的自由度进行积分得到 P_{s+1}

$$\int \{H_{s,N-s}, P\} (\Pi_{i=s+1}^N d\Gamma_i) = (N-s) \int \sum_{j=1}^s \{V_{j,s+1}(|\vec{r}_j - \vec{r}_{s+1}|), P_{s+1}\} d\Gamma_{s+1} \quad (19.27)$$

$$= (N-s) \int \sum_{i=1}^N \sum_{j=1}^s \left(\frac{\partial V_{j,s+1}}{\partial \vec{r}_i} \frac{\partial P_{s+1}}{\partial \vec{p}_i} \right) d\Gamma_{s+1} \quad (19.28)$$

$$= (N-s) \int \sum_{j=1}^s \left(\frac{\partial V_{j,s+1}}{\partial \vec{r}_j} \left(\frac{\partial}{\partial \vec{p}_j} - \frac{\partial}{\partial \vec{p}_{s+1}} \right) P_{s+1} \right) d\Gamma_{s+1} \quad (19.29)$$

$$= (N-s) \int \sum_{j=1}^s \frac{\partial V_{j,s+1}}{\partial \vec{r}_j} \frac{\partial P_{s+1}}{\partial \vec{p}_j} d\Gamma_{s+1} \quad (19.30)$$

综合以上结果我们可以得到 P_s 满足的方程:

$$\frac{\partial}{\partial t} P_s + \{P_s, H_s\} = (N-s) \int \sum_{j=1}^s \frac{\partial V_{j,s+1}}{\partial \vec{r}_j} \frac{\partial P_{s+1}}{\partial \vec{p}_j} d\Gamma_{s+1}, \quad (19.31)$$

写成约化粒子密度函数的形式为:

$$\frac{\partial}{\partial t} f_s + \{f_s, H_s\} = \int \sum_{j=1}^s \frac{\partial V_{j,s+1}}{\partial \vec{r}_j} \frac{\partial f_{s+1}}{\partial \vec{p}_j} d\Gamma_{s+1}. \quad (19.32)$$

这个方程的左边描述的是 s 个粒子的运动方程, 而方程的右边是这 s 个粒子与其他 $N-s$ 个粒子的相互碰撞作用. 从这个方程我们看到, f_s 的运动方程包含了 f_{s+1} . 也就是说, 不同阶的密度分布 f_i 是互相耦合的. 我们最后得到的是由 N 个方程组成的链, 它们互相依赖. 这就是 BBGKY 方程链. 我们把问题从求解包含 N 粒子自由度的一个分布函数, 转化为了求解 N 个不同阶数的分布函数. 所以, 我们的问题并没有得到简化. 下面我们考虑利用这一组链式方程组对问题进行简化.

19.4 Boltzmann 方程

我们首先写出 BBGKY 方程链的前两个:

$$\frac{\partial}{\partial t} f_1(\vec{r}_1, \vec{p}_1, t) + \frac{\vec{p}_1}{m} \cdot \frac{\partial f_1}{\partial \vec{r}_1} + \vec{F}_1 \cdot \frac{\partial f_1}{\partial \vec{p}_1} = \int \vec{F}_{12} \frac{\partial f_2}{\partial \vec{p}_1} d\vec{r}_2 d\vec{p}_2. \quad (19.33)$$

$$\begin{aligned} \frac{\partial}{\partial t} f_2(\vec{r}_1, \vec{p}_1, \vec{r}_2, \vec{p}_2, t) &+ \sum_{i=1}^2 \left(\frac{\vec{p}_i}{m} \cdot \frac{\partial f_2}{\partial \vec{r}_i} + \vec{F}_i \cdot \frac{\partial f_2}{\partial \vec{p}_i} \right) - \vec{F}_{12} \left(\frac{\partial}{\partial \vec{p}_1} - \frac{\partial}{\partial \vec{p}_2} \right) f_2(\vec{r}_1, \vec{p}_1, \vec{r}_2, \vec{p}_2, t) \\ &= \int \sum_{i=1}^2 \vec{F}_{i3} \frac{\partial f_3}{\partial \vec{p}_i} d\vec{r}_3 d\vec{p}_3. \end{aligned} \quad (19.34)$$

with

$$\vec{F}_i = -\frac{\partial U_i}{\partial \vec{r}_i}, \quad (19.35)$$

$$\vec{F}_{ij} = \frac{\partial V_{ij}}{\partial \vec{r}_i}. \quad (19.36)$$

从这两个方程我们可以看出, 方程的每一项对应的量纲为 $1/t$. 我们需要分析一下它们的大小. 以一个容器内的气体分子为例:

- $\frac{\vec{p}_i}{m} \cdot \frac{\partial}{\partial \vec{r}_i}$ 中, 粒子的速度 \vec{p}_i/m 一般在 $\sim 10^2$ m/s 的量级, 而容器的边长一般在 1 m 的量级, 所以这个量一般在 $\sim 10^2$ s⁻¹ 的量级.
- 对于容器里的气体分子来说, F_i 比较小, 因为只有在它们碰撞器壁的时候才感受到受力, 所以这一项比前一项还要小.
- \vec{F}_{ij} 由于粒子之间的短程相互作用 ($10^{-9} - 10^{-10}$ m 的量级), 所以它要比其他项大很多个数量级 $1/\tau_c \sim v/d \sim 10^{12}$ s⁻¹ 的量级. 这里 τ_c 碰撞时间.
- 方程 (19.34) 的右边一项大概在 $1/\tau_d \sim \frac{d^3 v}{V d} N \sim d^2 n v \sim 10^{-20} \times 10^{26} \times 10^2 \sim 10^8$ s⁻¹ 的量级, 比对应的左边包含 F_{12} 的项要小很多数量级.

根据上面的分析, 我们可以作下面的近似:

- 对 BBGKY 链作截断, 忽略方程 (19.34) 右边的项.
- 比较 (19.34) 和 (19.33), 我们可以认为 f_2 要比 f_1 更容易达到平衡, 所以我们可以让 $\frac{\partial f_2}{\partial t} = 0$.

于是, 我们有

$$\sum_{i=1}^2 \left(\frac{\vec{p}_i}{m} \cdot \frac{\partial f_2}{\partial \vec{r}_i} + \vec{F}_i \cdot \frac{\partial f_2}{\partial \vec{p}_i} \right) = \vec{F}_{12} \left(\frac{\partial}{\partial \vec{p}_1} - \frac{\partial}{\partial \vec{p}_2} \right) f_2(\vec{r}_1, \vec{p}_1, \vec{r}_2, \vec{p}_2, t) \quad (19.37)$$

方程 (19.33) 的右边可以写为 (利用分部积分可以证明添加的一项为零)

$$\int \vec{F}_{12} \frac{\partial f_2}{\partial \vec{p}_1} d\vec{r}_2 d\vec{p}_2 = \int \vec{F}_{12} \left(\frac{\partial}{\partial \vec{p}_1} - \frac{\partial}{\partial \vec{p}_2} \right) f_2 d\vec{r}_2 d\vec{p}_2 \approx \sum_{i=1}^2 \int \frac{\vec{p}_i}{m} \cdot \frac{\partial f_2}{\partial \vec{r}_i} d\vec{r}_2 d\vec{p}_2. \quad (19.38)$$

这里我们忽略了 F_i 项的贡献, 因为对于容器内的分子来说, F_i 只在边界的地方有作用. 这样, 我们得到

$$\frac{\partial}{\partial t} f_1(\vec{r}_1, \vec{p}_1, t) + \frac{\vec{p}_1}{m} \cdot \frac{\partial f_1}{\partial \vec{r}_1} + \vec{F}_1 \cdot \frac{\partial f_1}{\partial \vec{p}_1} = \int \left(\frac{\vec{p}_1}{m} \cdot \frac{\partial f_2}{\partial \vec{r}_1} + \frac{\vec{p}_2}{m} \cdot \frac{\partial f_2}{\partial \vec{r}_2} \right) d\vec{r}_2 d\vec{p}_2 \equiv \frac{\partial f_1}{\partial t} \Big|_{coll}. \quad (19.39)$$

显然, 方程的左边描述了单个粒子的运动, 而方程的右边描述的是两个粒子之间的碰撞相互作用.

对于两粒子碰撞, 我们可以在质心坐标系下看. 考虑质量相同的两个粒子, 我们做一个

变换 $\vec{r} = \vec{r}_1 - \vec{r}_2$, $\vec{Q} = (\vec{r}_1 + \vec{r}_2)/2$, 对动量做类似变换 $\vec{p} = \vec{p}_1 - \vec{p}_2$, $\vec{P} = (\vec{p}_1 + \vec{p}_2)/2$. 于是,

$\frac{\partial}{\partial \vec{r}_1} = -\frac{\partial}{\partial \vec{r}_2} \approx \frac{\partial}{\partial \vec{r}}$ (忽略随质心的变化),

$$\begin{aligned} \left. \frac{\partial f_1}{\partial t} \right|_{coll} &= \int \frac{1}{m} (\vec{p}_1 - \vec{p}_2) \frac{\partial f_2}{\partial a} da d\vec{b} d\vec{p}_2 \\ &= \int \frac{1}{m} |\vec{p}_1 - \vec{p}_2| (f_2(\vec{p}_1, \vec{p}_2, \vec{r}_1, \vec{r}_2, +) - f_2(\vec{p}_1, \vec{p}_2, \vec{r}_1, \vec{r}_2, -)) d\vec{b} d\vec{p}_2 \\ &= \int \frac{1}{m} |\vec{p}_1 - \vec{p}_2| (f_2(\vec{p}'_1, \vec{p}'_2, \vec{r}'_1, \vec{r}'_2, -) - f_2(\vec{p}_1, \vec{p}_2, \vec{r}_1, \vec{r}_2, -)) d\vec{b} d\vec{p}_2 \\ &\approx \int \frac{1}{m} |\vec{p}_1 - \vec{p}_2| (f_2(\vec{p}'_1, \vec{p}'_2, \vec{r}_1) - f_2(\vec{p}_1, \vec{p}_2, \vec{r}_1)) d\vec{b} d\vec{p}_2 \quad (19.40) \end{aligned}$$

$$= \int' w(\vec{p}_1, \vec{p}_2, \vec{p}'_1, \vec{p}'_2) (f_2(\vec{p}'_1, \vec{p}'_2, \vec{r}_1) - f_2(\vec{p}_1, \vec{p}_2, \vec{r}_1)) d\vec{p}_2 d\vec{p}'_1 d\vec{p}'_2. \quad (19.41)$$

这里有几点需要说明:

- 首先, $d\vec{r}_2 = dad\vec{b}_2$. 这里, \vec{b} 为垂直与两个粒子连线方向 ($\vec{r} = \vec{r}_1 - \vec{r}_2$) 的二维矢量.
- 第二行中的, + 和 - 分别表示碰撞后和碰撞前的分布函数.
- 从第二行到第三行, 我们利用碰撞过程的时间反演不变性做了替换. 把碰撞后的分布函数 $f_2(\vec{p}_1, \vec{p}_2, \vec{r}_1, \vec{r}_2, +)$ 换成了 $f_2(\vec{p}'_1, \vec{p}'_2, \vec{r}'_1, \vec{r}'_2, -)$. 注意这个替换看起来像是一个个人选择, 但这在整个理论中起决定作用。
- 从第三行到第四行, 我们做了一个近似: 认为粒子间的碰撞发生在空间的一个位置 \vec{r}_1 . 但我们知道实际上两个粒子的相互作用根据它们的作用强弱不同在空间是有一个分布的. 这里我们忽略了这一空间分布, 也就是对空间做了一个粗粒化近似 (coarse-graining).
- $\int' = \int \delta(\vec{p}_{1\alpha} + \vec{p}_{2\alpha} - \vec{p}'_{1\alpha} - \vec{p}'_{2\alpha}) \delta(E_1 + E_2 - E'_1 - E'_2)$, 即在保证能量和三个方向动量守恒的前提下积分. $d\vec{b}_2$ 是一个二维矢量, 而 $d\vec{p}'_1 d\vec{p}'_2$ 是六维的, 动量和能量守恒保证了从第三行到第四行的独立变量数目的一致性. 事实上, 根据散射理论, 它们之间存在一一对应的关系. 而这里 $w(\vec{p}_1, \vec{p}_2, \vec{p}'_1, \vec{p}'_2)$ 的引入是为了确保它们两个定量相等, 当然不同的相互作用势对应 w 表达式也不一样, 但在这里并不重要. 我们这里只假设它满足时间和空间反演对称性.

现在, 我们的方程还没有封闭, 因为还包含了 f_2 . 为了封闭我们的方程, 我们作最后一个近似

$$f_2(\vec{r}_1, \vec{r}_2, \vec{p}_1, \vec{p}_2) \approx f_1(\vec{r}_1, \vec{p}_1) f_1(\vec{r}_2, \vec{p}_2). \quad (19.42)$$

即, 认为粒子 1 和粒子 2 在碰撞前的分布是相互没有关联的. 这个近似非常关键, 它破坏了运动方程的时间反演对称性, 为整个体系的动力学行为设定了一个时间箭头. (见下节的 Boltzmann

H-定理)这样,我们就得到了 Boltzmann 方程:

$$\begin{aligned} \frac{\partial}{\partial t} f_1(\vec{r}_1, \vec{p}_1, t) + \frac{\vec{p}_1}{m} \cdot \frac{\partial f_1}{\partial \vec{r}_1} + \vec{F}_1 \cdot \frac{\partial f_1}{\partial \vec{p}_1} &= \left. \frac{\partial f_1}{\partial t} \right|_{coll} \\ &= \int' \frac{1}{m} |\vec{p}_1 - \vec{p}_2| [f_1(\vec{r}_1, \vec{p}'_1, t) f_1(\vec{r}_1, \vec{p}'_2, t) - f_1(\vec{r}_1, \vec{p}_1, t) f_1(\vec{r}_1, \vec{p}_2, t)] d\vec{p}_2 d\vec{p}'_1 d\vec{p}'_2 \quad (19.43) \\ &= \int w(\vec{p}_1, \vec{p}_2, \vec{p}'_1, \vec{p}'_2) [f_1(\vec{r}_1, \vec{p}'_1, t) f_1(\vec{r}_1, \vec{p}'_2, t) - f_1(\vec{r}_1, \vec{p}_1, t) f_1(\vec{r}_1, \vec{p}_2, t)] d\vec{p}_2 d\vec{p}'_1 d\vec{p}'_2 \end{aligned}$$

或者简写为:

$$\mathcal{L}f_1 = C(f_1, f_1). \quad (19.45)$$

到这里,我们对我们的问题已经做了重要的简化. 我们通过截断 BBGKY 方程链得到了单粒子密度分布 f_1 满足的方程. 如果我们知道 f_1 就可以得到很多我们关心的有关系统的信息. 比如, 我们可以去求解重要物理量的平均值. 但 Boltzmann 方程是一个微分-积分方程, 其求解也不容易, 很多时候我们需要把它线性化.

我们可以看到, Boltzmann 方程中最复杂的一项就是两个粒子的碰撞积分。我们通过一系列的考虑才最终得到这个碰撞积分 (19.44)。但最后我们得到的形式还是比较容易理解的。其中, $w(\vec{p}_1, \vec{p}_2, \vec{p}'_1, \vec{p}'_2)$ 由二体碰撞的性质来决定。而中括号中的第一项表示初态为 \vec{p}_1, \vec{p}_2 的两个粒子碰撞后变为 \vec{p}'_1, \vec{p}'_2 , 而第二项表示反过程, 最终由于碰撞而导致的分布函数的变化由两个过程之差来决定。

19.4.1 局域平衡

当 $C(f_1, f_1) = 0$ 的时候, 我们说系统在 \vec{r} 位置处于局域平衡. 很显然,

$$f_1(\vec{r}_1, \vec{p}'_1, t) f_1(\vec{r}_1, \vec{p}'_2, t) = f_1(\vec{r}_1, \vec{p}_1, t) f_1(\vec{r}_1, \vec{p}_2, t) \quad (19.46)$$

或者

$$\ln f_1(\vec{r}_1, \vec{p}'_1, t) + \ln f_1(\vec{r}_1, \vec{p}'_2, t) = \ln f_1(\vec{r}_1, \vec{p}_1, t) + \ln f_1(\vec{r}_1, \vec{p}_2, t). \quad (19.47)$$

是 $C(f_1, f_2) = 0$ 的充分条件。可以证明, 它也是必要条件。上面的式子说明, $\ln f_1$ 在碰撞前后是不变的。而我们知道在碰撞过程中粒子数, 能量和动量守恒, 我们可以根据这些构建 f_1

$$\ln f_1(\vec{r}, \vec{p}, t) = \vec{\alpha} \cdot \vec{p} - \beta \frac{\vec{p}^2}{2m} + \gamma, \quad (19.48)$$

于是我们可以看到

$$f_1(\vec{r}, \vec{p}, t) \propto \exp \left\{ -\frac{\beta(\vec{r}, t) [\vec{p}(\vec{r}, t)/m - \alpha'(\vec{r}, t)\vec{u}(\vec{r}, t)]^2}{2m} \right\}. \quad (19.49)$$

满足局域平衡条件.

19.4.2 量子版本

对于量子全同粒子, 局域平衡条件变为:

$$f_1(\vec{r}_1, \vec{p}'_1, t) f_1(\vec{r}_1, \vec{p}'_2, t) (1 \mp f_1(\vec{r}_1, \vec{p}_1, t)) (1 \mp f_2(\vec{r}_1, \vec{p}_2, t)) = f_1(\vec{r}_1, \vec{p}_1, t) f_1(\vec{r}_1, \vec{p}_2, t) (1 \mp f_1(\vec{r}_1, \vec{p}'_1, t)) (1 \mp f_2(\vec{r}_1, \vec{p}'_2, t)) \quad (19.50)$$

这里的负号表示费米子, 正号表示玻色子. 相同的思路给出分布函数:

$$f_1(\vec{r}, \vec{p}, t) \propto \left[\exp \left\{ -\frac{\beta(\vec{r}, t) (\vec{p}(\vec{r}, t)/m - \alpha'(\vec{r}, t) \vec{u}(\vec{r}, t))^2}{2m} \right\} \pm 1 \right]^{-1}, \quad (19.51)$$

对应费米和玻色分布.

19.4.3 半经典 Boltzmann 方程

半经典的 Boltzmann 方程被广泛的应用于半导体材料中的电子输运. 从我们这里得到的经典 Boltzmann 方程出发, 我们需要做如下修正:

- 电子能级利用量子力学的 Bloch 定理得出; $\frac{\partial f_1}{\partial t} \Big|_{coll}$ 需要用量子力学的微扰理论计算.
- 半导体中的 Bloch 电子满足半经典的运动方程:

$$m^* \dot{\vec{r}} = \hbar \vec{k},$$

$$\hbar \dot{\vec{k}} = \vec{F}.$$

这里的 $\hbar \vec{k}$ 与经典版本的动量对应, 而且 m^* 为电子的有效质量, 由能带结构决定.

19.5 Boltzmann H-定理

Boltzmann 定义了 H 函数 (注: Boltzmann 最早采用的并不是符号 H)

$$H(t) = \int f_1(\vec{r}, \vec{p}, t) \ln f_1(\vec{r}, \vec{p}, t) d\vec{r} d\vec{p}, \quad (19.52)$$

并且证明了

$$\frac{dH(t)}{dt} \leq 0. \quad (19.53)$$

如果我们定义熵为:

$$S_B(t) = -k_B H(t), \quad (19.54)$$

这可以认为是特定版本的热力学第二定律。这样，我们从一组严格满足时间反演对称的 Hamilton 运动方程出发得到的 Boltzmann 方程已经破坏了时间反演对称性。而且 H 定理告诉我们，满足 Boltzmann 方程的体系会倾向于向平衡态的方向发展，与我们的实际经验符合。而这一过程得以实现最重要的处理有两个地方：

- 忽略三体和多体碰撞过程，从而对 BBGKY 链做截断，
- 碰撞过程在时间和空间上的粗粒化，认为碰撞发生在确定的时间和空间，而忽略其具体过程，
- 认为碰撞前两个粒子的分布没有关联，从而把碰撞前的分布函数 f_2 分解为两个 f_1 相乘的形式。这个选择破坏了碰撞前后的对称性，直接导致 H-定理。可以验证，如果我们认为碰撞后的分布函数没有关联，那么我们将得到完全相反的结果，与实际不符。在平衡态的时候我们很难确定那个选择更合适，但在非平衡态的时候，只有我们的选择与实际相符合。

19.5.1 H 定理的证明

下面我们证明 Boltzmann 的 H 定理

$$\frac{dH(t)}{dt} = \frac{d}{dt} \int f_1(\vec{r}, \vec{p}, t) \ln f_1(\vec{r}, \vec{p}, t) d\vec{r} d\vec{p} \quad (19.55)$$

$$= \int \frac{\partial f_1}{\partial t} (\ln f_1 + 1) d\vec{r} d\vec{p} = \int \frac{\partial f_1}{\partial t} \ln f_1 d\vec{r} d\vec{p}. \quad (19.56)$$

利用 Boltzmann 方程，我们知道 f_1 对 t 的依赖可以分为三项。其中前两项可以通过分部积分证明

$$-\int \left(\frac{\vec{p}}{m} \cdot \frac{\partial f_1}{\partial \vec{r}} + \vec{F}_1 \cdot \frac{\partial f_1}{\partial \vec{p}} \right) \ln f_1 d\vec{r} d\vec{p} = 0. \quad (19.57)$$

所以我们只需要考虑碰撞项：

$$\frac{dH(t)}{dt} = \int \frac{1}{m} |\vec{p}_1 - \vec{p}_2| (f_1(\vec{p}'_1) f_1(\vec{p}'_2) - f_1(\vec{p}_1) f_1(\vec{p}_2)) \ln f_1(\vec{p}_1) d\vec{p}'_1 d\vec{p}'_2 d\vec{p}_2 d\vec{p}_1 d\vec{r}_1 \quad (19.58)$$

$$(19.59)$$

由于我们需要对所有的动量进行积分, 所以它们都是中间变量, 我们可以对它们进行重命名. 于是,

$$\begin{aligned} \frac{dH(t)}{dt} &= \frac{1}{2} \int \frac{1}{m} |\vec{p}_1 - \vec{p}_2| (f_1(\vec{p}'_1) f_1(\vec{p}'_2) - f_1(\vec{p}_1) f_1(\vec{p}_2)) [\ln f_1(\vec{p}_1) + \ln f_1(\vec{p}_2)] d\vec{p}'_1 d\vec{p}'_2 d\vec{p}_2 d\vec{r}_1 d\theta \\ &= \frac{1}{4} \int \frac{1}{m} |\vec{p}_1 - \vec{p}_2| (f_1(\vec{p}'_1) f_1(\vec{p}'_2) - f_1(\vec{p}_1) f_1(\vec{p}_2)) \\ &\quad \times [\ln f_1(\vec{p}_1) + \ln f_1(\vec{p}_2) - \ln f_1(\vec{p}'_1) - \ln f_1(\vec{p}'_2)] d\vec{p}'_1 d\vec{p}'_2 d\vec{p}_2 d\vec{p}_1 d\vec{r}_1 \end{aligned} \quad (19.61)$$

$$\begin{aligned} &= \frac{1}{4} \int \frac{1}{m} |\vec{p}_1 - \vec{p}_2| (f_1(\vec{p}'_1) f_1(\vec{p}'_2) - f_1(\vec{p}_1) f_1(\vec{p}_2)) \\ &\quad \times [\ln(f_1(\vec{p}_1) f_1(\vec{p}_2)) - \ln(f_1(\vec{p}'_1) f_1(\vec{p}'_2))] d\vec{p}'_1 d\vec{p}'_2 d\vec{p}_2 d\vec{p}_1 d\vec{r}_1 \end{aligned} \quad (19.62)$$

$$\begin{aligned} &\leq 0. \end{aligned} \quad (19.63)$$

最后, 我们利用了数学关系 $(x - y)(\ln x - \ln y) \geq 0$.

19.5.2 H 函数和熵的关系

通常情况下, 我们把熵写为 $S = \ln \Omega$ 。这里, Ω 为给定宏观状态对应的微观状态数。我们可以把它与 H 函数通过下面的方式建立联系。考虑这样一个问题: 考虑把 N 个粒子放入到 M 个盒子中的问题。一个给定的“宏观状态”--第 i 个盒子里有 N_i 个粒子--对应的微观状态数为:

$$\Omega = \frac{N!}{\prod_i N_i!} \quad (19.64)$$

对应的 Boltzmann 熵在 $N \rightarrow \infty$ 的时候

$$\ln \Omega = \lim_{N \rightarrow \infty} \ln \frac{N!}{\prod_i N_i!} \approx -N \sum_i P_i \ln P_i = -\sum_i f_1(i) (\ln f_1(i) - \ln N) = -\sum_i f_1(i) \ln f_1(i) + Constant \quad (19.65)$$

这里 $P_i = N_i/N$ 。

19.6 流体动力学方程

通过上面 Boltzmann 方程和 H 定理的分析我们知道, 单粒子分布函数 $f_1(\vec{r}, \vec{p}, t)$ 满足局域平衡条件(19.49)的话可以使方程右边的碰撞项为零, 同时可以使得 $dH/dt = 0$. 但是, 很容易验证, 仅仅满足局域平衡的分布函数并不能满足 Boltzmann 方程, 因为它并不能使 Boltzmann 方程左边的 Louville 项 $\{H_1, f_1\}$ 为零(所以 $dH/dt = 0$ 不一定意味着体系达到平衡). 而讨论 Louville 方程的时候我们知道如果 f_1 可以写成哈密顿量的函数 $f_1(H)$, 那么它就满足 $\{H_1, f_1\} = 0$. 比较两者的形式, 我们发现两者的形式还是很接近的. 另外, 要达到平衡, 还要保证分布函数不显

式依赖时间.

此外, 我们知道 Boltzmann 方程中的碰撞项是时间尺度最快的项, 而方程右边的 Louville 项要慢很多. 所以体系要达到平衡的话, 很可能首先是通过粒子间的相互碰撞达到局域平衡, 然后通过 Louville 类型的动力学过程来达到整体的平衡. 这一过程的时间尺度比碰撞要慢很多. 比如整个体系在不同的空间部分达到了局域平衡可以用局域压强, 局域温度等宏观量来描述. 但由于不同空间位置的压强或者温度不同, 可能会导致气体的扩散行为.

当我们的时间分辨率没有那么高的时候, 或者我们不关心微观的碰撞过程的时候, 我们可以去考察这些宏观量在较慢的时间尺度下的行为. 它们的行为可以用一系列的流体动力学方程来描述. 我们这一节的主要目的就是从 Boltzmann 方程出发, 消除碰撞项, 来推导出流体动力学方程.

在很多时候我们只对实空间分布更感兴趣. 因此, 我们可以更进一步把相空间中的动量项积分掉. 比如

$$n(\vec{r}, t) = \int f_1(\vec{r}, \vec{p}, t) d\vec{p} \quad (19.66)$$

为空间粒子数密度. 在本节, 对于任意物理量 O 我们定义

$$\langle O \rangle = \frac{\int O f_1(\vec{p}, \vec{r}, t) d\vec{p}}{\int f_1(\vec{p}, \vec{r}, t) d\vec{p}} = \frac{1}{n} \int O f_1(\vec{p}, \vec{r}, t) d\vec{p}. \quad (19.67)$$

注意虽然 O 只依赖位置 \vec{r} 和动量 \vec{p} , 但 $\langle O \rangle(\vec{r}, t)$ 为时间和空间的函数.

19.6.1 碰撞守恒量

我们首先考虑碰撞守恒量. 我们从 Boltzmann 方程出发, 考虑物理量 $X(\vec{r}, \vec{p})$,

$$\int \frac{\partial f_1(\vec{r}_1, \vec{p}_1)}{\partial t} \Big|_{coll} X(\vec{r}_1, \vec{p}_2) d\vec{p}_1 = \int X(\vec{p}_1) w(\vec{p}_1, \vec{p}_2, \vec{p}'_1, \vec{p}'_2) (f_1(\vec{p}'_1) f_1(\vec{p}'_2) - f_1(\vec{p}_1) f_1(\vec{p}_2)) d\vec{p}_1 d\vec{p}_2 d\vec{p}'_1 d\vec{p}'_2 \quad (19.68)$$

由于上式中我们需要对所有的动量指标求积分, 所以它们都是内部变量, 我们可以对它们进行重命名. 类似证明 H 定理的方法, 我们发现, 如果在碰撞前后, X 满足

$$X(\vec{p}_1) + X(\vec{p}_2) = X(\vec{p}'_1) + X(\vec{p}'_2), \quad (19.69)$$

那么式 (19.68) 为零. 满足式子 (19.69) 条件的物理量我们把它叫做碰撞守恒量. 显然, 粒子数, 动量和能量都是碰撞守恒量. 对于这些量, 我们考虑 Boltzmann 方程的左边的项

$$\int X(\vec{r}, \vec{p}) \mathcal{L}\{f_1(\vec{r}, \vec{p})\} d\vec{p} = \int (\mathcal{L}\{X(\vec{r}, \vec{p})\} f_1(\vec{r}, \vec{p}) - f_1(\vec{r}, \vec{p}) \mathcal{L}\{X(\vec{r}, \vec{p})\}) d\vec{p}. \quad (19.70)$$

上式有两部分. 我们首先考虑第一部分

$$\int \mathcal{L}\{X(\vec{r}, \vec{p})f_1(\vec{r}, \vec{p})\}d\vec{p} = \int \left(\partial_t(Xf_1) + \frac{p_\alpha}{m}\partial_\alpha(Xf_1) + F_\alpha\partial_{p_\alpha}(Xf_1) \right) d\vec{p} \quad (19.71)$$

$$= \partial_t(\langle X \rangle n) + \partial_\alpha \left(\langle \frac{p_\alpha}{m} X \rangle n \right). \quad (19.72)$$

这里我们为了简化方程, 默认对所有重复的指标求和, $\partial_\alpha = \frac{\partial}{\partial r_\alpha}$, $\partial_{p_\alpha} = \frac{\partial}{\partial p_\alpha}$, $\partial_t = \frac{\partial}{\partial t}$. 并且, $\mathcal{L}\{f\} \equiv \partial_t f + p_\alpha/m\partial_\alpha f + F_\alpha\partial_{p_\alpha} f$. 第二部分为:

$$\int f_1(\vec{r}, \vec{p})\mathcal{L}\{X(\vec{r}, \vec{p})\}d\vec{p} = \langle \partial_t X \rangle n + \langle \frac{p_\alpha}{m}\partial_\alpha X \rangle n + \langle F_\alpha\partial_{p_\alpha} X \rangle n. \quad (19.73)$$

综合这两部分我们得到:

$$\boxed{\partial_t(\langle X \rangle n) + \partial_\alpha \left(\langle \frac{p_\alpha}{m} X \rangle n \right) - \langle \partial_t X \rangle n - \langle \frac{p_\alpha}{m}\partial_\alpha X \rangle n - \langle F_\alpha\partial_{p_\alpha} X \rangle n = 0.} \quad (19.74)$$

这是所有碰撞守恒量满足的方程, 也是我们下面所有讨论的基础. 下面我们分别考虑粒子数, 动量和动能.

19.6.2 粒子数

对于粒子数, 我们令 $X = 1$, 我们得到连续性方程:

$$\partial_t n + \partial_\alpha \left(\langle \frac{p_\alpha}{m} \rangle n \right) = 0. \quad (19.75)$$

或者

$$\boxed{\partial_t n + u_\alpha \partial_\alpha n \equiv \mathcal{D}n = -n\partial_\alpha u_\alpha.} \quad (19.76)$$

我们定义了平均速度 $u_\alpha = \langle \frac{p_\alpha}{m} \rangle$, 和全微分 $\mathcal{D} = \partial_t + u_\alpha \partial_\alpha$.

19.6.3 动量

令 $X = \vec{c} = (\vec{p}/m - \vec{u})$, 利用 $\langle c_\alpha \rangle = 0$, 我们得到

$$\frac{1}{n}\partial_\alpha \left(\langle \frac{p_\alpha}{m} c_\beta \rangle n \right) - \langle \partial_t c_\beta \rangle - \langle \frac{p_\alpha}{m} \partial_\alpha c_\beta \rangle - \frac{1}{m} \langle F_\beta \rangle = 0. \quad (19.77)$$

定义 $P_{\alpha\beta} = mn\langle c_\alpha c_\beta \rangle$, 整理得到:

$$\boxed{\mathcal{D}u_\beta = \frac{1}{m}F_\beta - \frac{1}{mn}\partial_\alpha P_{\alpha\beta}.} \quad (19.78)$$

19.6.4 动能

定义 $X = m(p - u)^2/2$, $\varepsilon = \langle m(p - u)^2/2 \rangle$, 利用连续性方程 (19.76), 我们有

$$\mathcal{D}\varepsilon = -\frac{1}{n}P_{\alpha\beta}u_{\alpha\beta} - \frac{1}{n}\partial_\alpha h_\alpha. \quad (19.79)$$

我们这里又定义了热流密度 $h_\alpha = \frac{1}{2}m\langle c_\alpha c^2 \rangle$, 形变率张量 $u_{\alpha\beta} = (\partial_\alpha u_\beta + \partial_\beta u_\alpha)/2$.

方程组 (19.76-19.79) 给出了一组有关粒子数密度 n , 动量 u_α 和能量 ε 的流体动力学方程. 但这组方程不是封闭的, 因为我们定义了新的量比如 $P_{\alpha\beta}$, h_α . 而要计算这些量, 我们需要知道分布函数 f_1 才可以, 这就要求我们求解 Boltzmann 方程.

19.6.5 零阶流体动力学方程

如何再进一步呢? 我们可以根据上面的分析, 先猜测一个分布函数的形式. 比如用局域平衡分布函数,

$$f_1^0(\vec{p}, \vec{r}, t) = \frac{n(\vec{r}, t)}{(2\pi k_B T/m)^{3/2}} \exp\left[-\frac{m(\vec{p}/m - \vec{u}(\vec{r}, t))^2}{2k_B T(\vec{r}, t)}\right] \quad (19.80)$$

保证 Boltzmann 方程中最大的项碰撞项为零. 从 f_1^0 我们可以计算得到 $h_\alpha = 0$, $P_{\alpha\beta} = nk_B T \delta_{\alpha\beta}$, $\varepsilon = 3k_B T/2$. 这种情况下, 我们的流体动力学方程变为:

$$\partial_t n + u_\alpha \partial_\alpha n = -n \partial_\alpha u_\alpha. \quad (19.81)$$

$$\partial_t u_\beta + u_\alpha \partial_\alpha u_\beta = \frac{F_\beta}{m} - \frac{1}{mn} \partial_\beta (nk_B T). \quad (19.82)$$

$$\partial_t T + u_\alpha \partial_\alpha T = -\frac{2}{3}T u_{\alpha\alpha}. \quad (19.83)$$

为了说明这组方程不能保证体系达到平衡态, 我们可以做一个线性分析. 考虑粒子数密度, 速度和温度对平均值的微小偏离, 并假设各向同性, 以及 $F_\beta = 0$. 比如

$$n = \bar{n} + \delta n e^{i(\vec{k} \cdot \vec{r} - \omega t)}, \quad (19.84)$$

$$T = \bar{T} + \delta T e^{i(\vec{k} \cdot \vec{r} - \omega t)}, \quad (19.85)$$

$$u = 0 + \delta u e^{i(\vec{k} \cdot \vec{r} - \omega t)}. \quad (19.86)$$

把它们带入到方程 (19.81-19.103) 中, 我们得到一组方程组来描述系统的线性响应:

$$\begin{pmatrix} 0 & \bar{n}k_\alpha & 0 \\ \frac{k_B \bar{T}}{m \bar{n}} k_\alpha \delta_{\alpha\beta} & 0 & \frac{k_B}{m} k_\alpha \delta_{\beta\alpha} \\ 0 & \frac{2}{3} \bar{T} k_\alpha & 0 \end{pmatrix} \begin{pmatrix} \delta n \\ \delta u_\alpha \\ \delta T \end{pmatrix} = \omega \begin{pmatrix} \delta n \\ \delta u_\beta \\ \delta T \end{pmatrix} \quad (19.87)$$

求解这组线性方程组, 我们可以得到系统的本征模式.

19.6.6 一阶流体动力学方程

19.6.6.1 $\mathcal{L}f_1^0$

我们假设的满足局域平衡的分布函数 f_1^0 并不是 Boltzmann 方程的解, 因为如果我们把它带入到 Boltzmann 方程 (19.44) 的左边的话, 我们可以得到:

$$\mathcal{L}\ln f_1^0 = (\partial_t + \frac{p_\alpha}{m}\partial_\alpha + F_\alpha\partial_{p_\alpha})\ln f_1^0 = (\mathcal{D} + c_\alpha\partial_\alpha + F_\alpha\partial_{p_\alpha})\ln f_1^0. \quad (19.88)$$

为方便起见, 我们这里考虑了

$$\ln f_1^0 = -\frac{mc^2}{2k_B T} + \ln(nT^{-3/2}) - \frac{3}{2}\ln(2\pi k_B m).$$

这个式子有三项, 我们分别计算. 第一项为:

$$\begin{aligned} \mathcal{D}\ln f_1^0 &= \left(\frac{mc^2}{2k_B T} - \frac{3}{2}\right)\frac{\mathcal{D}T}{T} + \frac{mc_\beta}{k_B T}\mathcal{D}u_\beta + \frac{1}{n}\mathcal{D}n \\ &= -\left(\frac{mc^2}{2k_B T} - \frac{3}{2}\right)\frac{2}{3nk_B T}(\partial_\alpha h_\alpha + P_{\alpha\beta}u_{\alpha\beta}) - \frac{c_\beta}{k_B T}\left(\frac{1}{n}\partial_\alpha P_{\alpha\beta} - F_\beta\right) - \partial_\alpha u_\alpha \end{aligned} \quad (19.89)$$

第二项为:

$$c_\alpha\partial_\alpha\ln f_1^0 = \left(\frac{mc^2}{2k_B T} - \frac{3}{2}\right)\frac{c_\alpha\partial_\alpha T}{T} + \frac{mc_\beta}{k_B T}c_\alpha\partial_\alpha u_\beta + \frac{1}{n}c_\alpha\partial_\alpha n. \quad (19.90)$$

第三项为:

$$F_\alpha\partial_{p_\alpha}\ln f_1^0 = \left(\frac{mc^2}{2k_B T} - \frac{3}{2}\right)\frac{F_\alpha\partial_{p_\alpha} T}{T} - \frac{mc_\beta}{k_B T}F_\alpha\partial_{p_\alpha} c_\beta = -\frac{c_\beta}{k_B T}F_\beta. \quad (19.91)$$

利用我们得到的零阶流体动力学方程组 (19.81-19.103), 我们可以得到:

$$\mathcal{L}\ln f_1^0 = \left(\frac{mc^2}{2k_B T} - \frac{5}{2}\right)\frac{c_\alpha}{T}\partial_\alpha T + (1 - \delta_{\alpha\beta})\frac{m}{k_B T}c_\alpha c_\beta\partial_\alpha u_\beta. \quad (19.92)$$

19.6.6.2 驰豫时间近似

为了考虑对局域平衡分布函数 f_1^0 的修正, 我们可以采用驰豫时间近似, 即引入唯一的驰豫时间 τ , 来描述粒子碰撞的频率。对应的 Boltzmann 方程可以写为

$$\mathcal{L}f_1(\vec{r}, \vec{p}, t) = -\frac{f_1 - f_1^0}{\tau}. \quad (19.93)$$

把分布函数写为 $f_1 = f_1^0(1 + g)$ 的形式，代入到上式中，考虑到上式右边部分比左边部分要大很多，我们可以得到

$$\mathcal{L}f_1^0 = -\frac{gf_1^0}{\tau}. \quad (19.94)$$

利用式子 (19.92)，我们可以得到对局域平衡分布函数的修正：

$$g = -\tau \left[\left(\frac{mc^2}{2k_B T} - \frac{5}{2} \right) \frac{c_\alpha}{T} \partial_\alpha T + (1 - \delta_{\alpha\beta}) \frac{m}{k_B T} c_\alpha c_\beta u_{\alpha\beta} \right]. \quad (19.95)$$

包含这一修正后，我们需要对 $P_{\alpha\beta}$, h_α 等量重新计算。我们注意到， h_α 的定义包含奇数个 c_α ，所以在 0 阶近似下为零。现在修正项 g 的第一项包含单个的 c_α ，所以这时的 h_α 不再为零

$$h_\alpha^{(1)} = -\frac{mn\tau}{2T} \left\langle \left(\frac{mc^2}{2k_B T} - \frac{5}{2} \right) c_\alpha^2 c^2 \right\rangle_0 \partial_\alpha T = -\frac{5}{2} \frac{n k_B^2 T \tau}{m} \partial_\alpha T \equiv -\kappa \partial_\alpha T. \quad (19.96)$$

最后一个等式定义了热导率 κ 。同样，式子 (19.95) 的第二项会对 $P_{\alpha\beta}$ 有一个修正，导致它不再仅对角元非零

$$P_{\alpha\beta}^{(1)} = -(1 - \delta_{\gamma\xi}) \frac{m^2 n \tau}{k_B T} u_{\gamma\xi} \langle c_\alpha c_\beta c_\gamma c_\xi \rangle = -2k_B T n \tau u_{\alpha\beta} (1 - \delta_{\alpha\beta}) \equiv -\mu (\partial_\beta u_\alpha + \partial_\alpha u_\beta) (1 - \delta_{\alpha\beta}). \quad (19.97)$$

最后一个式子定义了粘滞系数 $\mu = k_B T n \tau$ 。于是，一阶的流体动力学方程变为：

$$\partial_t n + u_\alpha \partial_\alpha n = -n \partial_\alpha u_\alpha. \quad (19.98)$$

$$\partial_t u_\beta + u_\alpha \partial_\alpha u_\beta = -\frac{k_B}{mn} \partial_\beta (nT) + \frac{\bar{\mu}}{mn} \partial_\alpha (\partial_\alpha u_\beta + \partial_\beta u_\alpha) (1 - \delta_{\alpha\beta}). \quad (19.99)$$

$$\partial_t T + u_\alpha \partial_\alpha T = -\frac{2}{3} k_B T u_{\alpha\alpha} + \frac{\bar{\kappa}}{n} \partial_\alpha (\partial_\alpha T). \quad (19.100)$$

其中蓝色的部分为一阶修正项。或者可以把它们写成矢量的形式：

$$\partial_t n + \vec{u} \cdot \nabla n = -n \nabla \cdot \vec{u}. \quad (19.101)$$

$$\rho (\partial_t \vec{u} + (\vec{u} \cdot \nabla) \vec{u}) = \frac{\rho \vec{F}}{m} - \nabla P + \bar{\mu} \nabla^2 \vec{u} + \frac{\bar{\mu}}{3} \nabla (\nabla \cdot \vec{u}). \quad (19.102)$$

$$\partial_t T + (\vec{u} \cdot \nabla) T + \frac{2}{3} k_B T \nabla \cdot \vec{u} - \frac{\kappa}{n} \nabla^2 T = 0. \quad (19.103)$$

我们这里忽略了 $u_{\alpha\beta}^2$ 项，以及 κ 和 μ 的空间依赖性，以它们空间平均 $\bar{\kappa}$ 和 $\bar{\mu}$ 来代替，另外 $\rho = mn$ 。其中的式子 (19.102) 是著名的 Navier-Stokes 方程。它被广泛应用于流体力学，用于研究飞机、汽车空气动力学、血液流动、天气预报、污染分析、视频游戏等等科学工程的方方面面。但有意思的是，在 19 世纪人们就建立了这个方程，但有关它的一些基本数学问题仍然没有得到解决。“Navier–Stokes existence and smoothness” 是 Clay 数学基金会 2000 年悬赏的 7 个千禧年悬而未决的数学问题之一。只要能够证明解的存在性就可以获得 100 万美元的奖励。

如果我们在此基础上再做一个线性分析的话，我们会发现原来的本征模式会由于耗散的引入而具有一个有限的寿命。

作业为扩展阅读部分：

19.7 扩展阅读

19.7.1 固体中的准粒子输运

Boltzmann 方程虽然最初用来描述稀薄气体的输运行为，但在凝聚态物理中，它也被广泛的应用于描述晶体材料中的电子和声子的半径典输运行为。与自由空间的连续平移不变性不同，晶体材料中的平移不变性是非连续的（平移晶格矢量）。因此，其中准粒子的动量要理解为准动量或者叫晶格动量。这时，准粒子之间的散射不一定满足准动量守恒，比如倒逆过程，散射前和散射后准粒子的晶格动量之和可以差一个倒格矢。同时，由于材料内部还存在各种各样的缺陷，准粒子和缺陷的散射会损失动量。准粒子与其他自由度的散射也会损失能量和动量。因此，在晶体材料中，我们这里所描述的流体动力学行为就不太容易发生。但并不是不可能。最近的一些工作表明，在高质量的晶体材料中，电子、声子等准粒子在特定的条件下也会发生流体动力学输运行为，表现出和经典流体类似的输运性质。下面列举的几篇最近的文献就是有关这方面的。有兴趣的同学可以试着阅读一下。

电子：

news&views: Quantum Hydrodynamic Transport in Graphene

Electrons go with the flow in exotic material systems, Science, 351, 1026 (2016).

research work:

Negative local resistance caused by viscous electron backflow in graphene, Science, 351, 1055 (2016).

Evidence for hydrodynamic electron flow in PdCoO₂, Science, 351, 1061 (2016).

Observation of the Dirac fluid and the breakdown of the Wiedemann-Franz law in graphene, Science, 351, 1058 (2016).

声子输运：

Phonon hydrodynamics in two-dimensional materials, Nature Communications, 6, 6400 (2015).

Hydrodynamic phonon transport in suspended graphene, Nature Communications, 6, 6290

(2015).

Thermal Transport in Crystals as a Kinetic Theory of Relaxons, Phys. Rev. X, 6, 041013 (2016),
with news& views here: [Viewpoint: Relaxons Heat Up Thermal Transport](#).

第 20 章 线性响应理论

本章主要讨论线性响应的理论。为了研究我们所关心系统的性质，我们可以对它施加一个很小的外部微扰，然后研究与系统相关的某一个物理量对这一微扰的响应。这一响应可以用响应函数(response function)来描述。施加的这个扰动一般称为“广义力”(generalized force)。首先，我们从唯象的角度讨论响应函数的一般性质；然后，我们介绍线性响应理论；最后，我们应用这一理论来考察电介质材料对外加电磁波的响应。

20.1 唯象理论

20.1.1 响应函数和广义极化率

我们考虑一个处于平衡态的系统，对系统施加一个很小的广义力，系统会对扰动做出响应，产生一个广义位移或广义速度。这里的广义力可以是机械力、电磁场力、温度梯度¹等；对应的广义位移或速度可以是机械位移、电极化、磁极化或电流、热流等。这里暂时不考虑具体细节，但我们可以写出一个普适的唯象关系：

$$\langle \delta A \rangle_t \equiv \langle A \rangle_t - \langle A \rangle_{eq} = \int_{-\infty}^t \chi(t-t') F(t') dt'. \quad (20.1)$$

上式中 $\chi(t-t')$ 是响应函数， $F(t')$ 是广义力， $\langle \delta A \rangle_t$ 表示物理量 A 对外力的响应，在 t 时刻其平均值 $\langle A \rangle_t$ 相对于平衡时的均值 $\langle A \rangle_{eq}$ 发生的变化。这个式子有下面几个特点：

- 系统的响应与外力呈线性关系 (linear)
- 系统的响应是非马柯夫的 (non-Markovian)
- 系统在时刻 t 只会对 t 时刻之前施加的外力做出响应，所以积分的上限为 t (causal)
- 系统的响应满足时间平移不变性，所以响应函数只和时间差有关 $\chi(t,t') = \chi(t-t')$ 。可以想象，如果 $t-t' \rightarrow +\infty$ 有 $\chi(t-t') \rightarrow 0$ 。也就是说，系统有一个“记忆时间”，它会忘记这个时间之前对它的扰动。
- 上面普适关系式可以推广到矩阵形式： $\delta A_i(t) = \sum_j \int \chi_{ij}(t-t') F_j(t') dt'$ 它表示系统的某些物理量可以对几种外力产生响应。这时系统对不同外力的响应组成一个响应矩阵。

¹ 温度场比较特殊，我们这里暂时不考虑。可以参考 Kubo, et al., J. Phys. Soc. Jpn. 12, pp. 1203-1211 (1957)

下面我们来考察响应函数在频率空间的性质。为此，我们定义傅里叶变换：

$$f[\omega] = \int f(t)e^{i\omega t} dt \quad (20.2)$$

$$f(t) = \int f[\omega]e^{-i\omega t} \frac{d\omega}{2\pi} \quad (20.3)$$

把定义带入到式20.1：

$$\int \delta A[\omega] e^{-i\omega t} \frac{d\omega}{2\pi} = \int_{-\infty}^t \chi(t-t') \int F[\omega] e^{-i\omega t'} \frac{d\omega}{2\pi} dt' \quad (20.4)$$

上式进一步可以表示为：

$$\int \frac{d\omega}{2\pi} (\delta A[\omega] - \chi[\omega]F[\omega]) = 0. \quad (20.5)$$

于是我们可以得到

$$\delta A[\omega] = \chi[\omega]F[\omega]. \quad (20.6)$$

其中，我们定义系统的广义极化率 (generalized susceptibility)：

$$\chi[\omega] = \int_{-\infty}^t dt' \chi(t-t') e^{i\omega(t-t')} \quad (20.7)$$

$$= \int_0^{+\infty} ds \chi(s) e^{i\omega s} \quad (20.8)$$

$$= \int_{-\infty}^{+\infty} ds \Theta(s) \chi(s) e^{i\omega s} \quad (20.9)$$

上式第二个等式中我们做了变量替换 $s = t - t'$ ，其中 $\Theta(t)$ 为台阶函数。如果 $F[t], \delta A, \chi(t-t')$ 都是实数，我们有 $\chi[-\omega] = \chi^\star[\omega]$ ，或者

$$\text{Re}\chi[-\omega] = \text{Re}\chi[\omega], \quad (20.10)$$

$$\text{Im}\chi[-\omega] = -\text{Im}\chi[\omega]. \quad (20.11)$$

20.1.2 广义极化率的性质

从广义极化率的定义

$$\chi[\omega] = \int \Theta(t) \chi(t) e^{i\omega t} dt \quad (20.12)$$

出发，我们可以得到它所满足的一个普适关系，称为色散关系 (dispersion relation) 或者 Kramers-Kronig (K-K) 关系。

从式 (20.12) 出发，利用 Fourier 变换的性质，我们有

$$\chi[\omega] = \int \frac{d\omega'}{2\pi} \Theta(\omega - \omega') \chi'[\omega']. \quad (20.13)$$

这里，我们引入了 $\chi'(t)$ ，它在 $t > 0$ 时与 $\chi(t)$ 相同。在 $t < 0$ 时，它的具体形式并不重要，因为 Θ 函数将其变为零。根据这个自由度，我们任意选取 $t < 0$ 是它的取值。我们不妨假设

$$\chi'(-t) = -\chi'(t).$$

下面，我们需要知道 $\Theta(t)$ 和 $\chi'(t)$ 的 Fourier 变换。对于 $\Theta(t)$ 我们前面已经讨论过

$$\Theta[\omega] = \frac{i}{\omega + i\gamma}. \quad (20.14)$$

于是，我们得到

$$\chi[\omega] = \int \frac{id\omega'}{2\pi} \frac{\chi'[\omega']}{\omega + i\gamma - \omega'}.$$

利用 $\chi'(t)$ 的对称性，我们可以把它写成

$$\chi'[\omega] = 2i\text{Im}\chi[\omega].$$

带入到式 (20.1.2) 中，我们得到：

$$\chi[\omega] = - \int \frac{d\omega'}{\pi} \frac{\text{Im}\chi[\omega']}{\omega + i\gamma - \omega'}. \quad (20.15)$$

或者写为

$$\boxed{\text{Re}\chi[\omega] = \mathcal{P} \int \frac{d\omega'}{\pi} \frac{\text{Im}\chi[\omega']}{\omega' - \omega}}. \quad (20.16)$$

如果我们假设

$$\chi'(-t) = \chi'(t),$$

那么通过类似的分析我们得到

$$\boxed{\text{Im}\chi[\omega] = -\mathcal{P} \int \frac{d\omega'}{\pi} \frac{\text{Re}\chi[\omega']}{\omega' - \omega}}. \quad (20.17)$$

式 (20.16) 和 (20.17) 称为色散关系或者 K-K 关系。色散关系给出了 $\chi[\omega]$ 的实部和虚部之间的关系，他们之间通过 Hilbert 变换联系在一起。只要知道其中一个，我们就可以利用色散关系得到另外一个。这在实际的应用中是非常方便的。

回顾上面的分析我们发现，**我们得到的色散关系是非常普遍的，它的成立只依赖于因果律**。当然，如果我们改变 Fourier 变换的定义，色散关系中的符号也需要做相应的修改。但我们也仍旧可以写出实部和虚部之间的关系。

作业： 如果我们定义 Fourier 变换为

$$f[\omega] = \int f(t)e^{-i\omega t} dt, \quad (20.18)$$

色散关系应该是什么形式?

举例: LRC 电路 考虑一个 LRC 电路, 我们外加一个偏压 $V(t)$ 来考察电路的电流响应。我们不难写出下面 $I(t)$ 与 $V(t)$ 的关系

$$L \frac{dI}{dt} + RI(t) + \frac{1}{C} \int_{-\infty}^t I(t') dt' = V(t) \quad (20.19)$$

这里电流的积分实际上是电荷

$$Q = \int_{-\infty}^t I(t') dt',$$

把上式用 Q 表示

$$\ddot{Q}(t) + \frac{R}{L}\dot{Q}(t) + \frac{1}{LC}Q(t) = V(t)/L.$$

我们定义 $\omega_0 = \frac{1}{\sqrt{LC}}$, $\gamma = \frac{R}{L}$, 就可以得到一个与谐振子 Langevin 方程很类似的方程

$$\ddot{Q}(t) + \gamma\dot{Q}(t) + \omega_0^2 Q(t) = V(t)/L.$$

这里的电荷对应谐振子的位移, 电流对应速度, 耗散系数对应 $\gamma = R/L$, 谐振子频率对应 $\omega_0 = 1/\sqrt{LC}$ 。我们先前对谐振子 Langevin 方程的分析都可以对应过来。

这里, 我们看电流对外加偏压的响应, 做 Fourier 变换, 在频率空间我们有

$$\left(-i\omega L + R + \frac{1}{(-i\omega)C} \right) I[\omega] = V[\omega]. \quad (20.20)$$

$$(20.21)$$

于是,

$$I[\omega] = Y[\omega]V[\omega] \quad (20.22)$$

$$Y[\omega] = \frac{1}{-i\omega L + R - \frac{1}{i\omega C}} \quad (20.23)$$

$$= \frac{i\omega}{L(\omega^2 + i\omega R/L - \frac{1}{LC})} \quad (20.24)$$

$$= \frac{i\omega/L}{\omega^2 - \omega_0^2 + i\gamma\omega} \quad (20.25)$$

$$= \frac{i\omega/L(\omega^2 - \omega_0^2 - i\gamma\omega)}{(\omega^2 - \omega_0^2)^2 + \gamma^2\omega^2}. \quad (20.26)$$

可以验证, $Y[\omega]$ 满足 K-K 关系, 而且 $Y[-\omega] = Y^*[\omega]$ 。

举例: Debye 弛豫 如果响应函数随时间指数衰减

$$\chi(t) = \chi_0 e^{-t/\tau} = \chi_0 e^{-\gamma t}$$

对应的广义极化率实部和虚部之间存在一个简单的关系。下面我们予以说明。

$$\chi[\omega] = \int_0^{+\infty} dt e^{i\omega t} \chi(t) \quad (20.27)$$

$$\begin{aligned} &= \chi_0 \int_0^{+\infty} e^{i(\omega+i\gamma)t} dt \\ &= \chi_0 \frac{i}{\omega + i\gamma} \\ &= \chi_0 \tau \left(\frac{1}{1 + \omega^2 \tau^2} + i \frac{\omega \tau}{1 + \omega^2 \tau^2} \right) \end{aligned} \quad (20.28)$$

令

$$u = \operatorname{Re} \chi, \quad (20.29)$$

$$v = \operatorname{Im} \chi, \quad (20.30)$$

那么它们满足这样的方程

$$\boxed{\left(u - \frac{1}{2} \right)^2 + v^2 = \frac{1}{4}} \quad (20.31)$$

实验上可以测量不同频率的极化率来拟合上式，从而验证系统是否满足 Debye 驰豫的特征。

20.1.3 单频响应

下面我们考虑系统对一个角频率为 ω_0 的外力的响应

$$F(t) = \frac{1}{2} (F_0 e^{-i\omega_0 t} + F_0^* e^{i\omega_0 t})$$

代入到 (20.1) 我们可以得到

$$\delta\langle A \rangle_t = \frac{1}{2} \left(F_0 \int_{-\infty}^t \chi(t-t') e^{-i\omega_0 t'} dt' + c.c. \right) \quad (20.32)$$

$$= \frac{1}{2} \left(F_0 \int_0^\infty \chi(\tau) e^{-i\omega_0(t-\tau)} d\tau + c.c. \right) \quad (20.33)$$

$$= \frac{1}{2} (F_0 \chi(\omega_0) e^{-i\omega_0 t} + c.c.). \quad (20.34)$$

这里 $c.c.$ 表示复共轭。通过这个结果我们可以看到，系统对频率为 ω_0 的单位外力的响应由对应频率的广义极化率 $\chi(\pm\omega_0)$ 给出。如果取 F_0 为实数，我们得到

$$\delta\langle A \rangle_t = F_0 (\cos \omega_0 t \operatorname{Re} \chi[\omega_0] + \sin \omega_0 t \operatorname{Im} \chi[\omega_0]). \quad (20.35)$$

20.1.4 能量耗散

由于系统和外界的耦合，外力可以对系统做功，当系统达到稳定状态后，系统的内能保持不变。这时，外力对系统的做功全部被耗散掉。于是，系统的耗散可以表示为：

$$P \propto \lim_{T \rightarrow +\infty} \frac{1}{2T} \int_{-T}^T F(t)(-\delta \dot{X}(t))dt \quad (20.36)$$

$$\begin{aligned} &= - \lim_{T \rightarrow +\infty} \frac{1}{2T} \int_{-T}^T \left(\int F[\omega] e^{-i\omega t} \frac{d\omega}{2\pi} \right) \left(\int (-i\omega') \delta X[\omega'] e^{-i\omega' t} \frac{d\omega'}{2\pi} \right) dt \\ &= i \int \omega F[-\omega] \delta X[\omega] \frac{d\omega}{(2\pi)^2} \\ &= i \int \omega F[-\omega] \chi[\omega] F[\omega] \frac{d\omega}{(2\pi)^2} \\ &= - \int \omega |F[\omega]|^2 \text{Im}\chi[\omega] \frac{d\omega}{(2\pi)^2} \end{aligned} \quad (20.37)$$

这里我们利用了 $\text{Re}\chi[\omega] = \text{Re}\chi[-\omega]$ ，所以 $\text{Re}\chi$ 积分部分为零。我们可以看到 $\text{Im}\chi_{xx}[\omega]$ 与体系的能量耗散相关联。这可以从另一个角度理解：通过前面分析我们知道， $-\text{Im}\chi_{xx}[\omega] = \text{Im}\chi_{xx}[-\omega]$ 。它经过傅里叶变换有： $-\chi[t] = \chi[-t]$ 对应时间反演改变符号，即不满足时间反应不变性。系统的耗散项必然不满足时间反演不变性，因为耗散是一种“单向”的行为。

20.2 线性响应理论：绝热近似

考虑一个量子体系用哈密顿量 \hat{H} 来描述。外力 $F(t)$ 与系统的某个量 \hat{B} 耦合。包含这个耦合之后总的哈密顿量为

$$\hat{H}_F(t) = \hat{H} + F(t)\hat{B} = \hat{H} + \hat{V}(t). \quad (20.38)$$

这里我们考虑一个绝热耦合过程 (adiabatic switch-on)，认为系统与外力的耦合是一个绝热的慢变过程。在开始时刻 (t_0)，耦合为 0，到了 $t = 0$ 的时刻耦合完全打开，中间经历了一个渐变的绝热过程，也就是说在这个过程中由于哈密顿量的变化，系统的本征能级可能会发生变化，但它们的相对排布不发生变化，也没有发生非弹性跃迁的过程。与外界耦合前每个能级的粒子占据数是多少，耦合打开之后仍然是多少。

为了考察系统对外界扰动的响应，我们考察在耦合打开之后系统另外一个量 \hat{A} 的期待值 $\langle \hat{A}_S(t_0) \rangle_{\rho(t)} = \langle \hat{A}_H(t) \rangle_{\rho(t_0)}$ ($t \gg 0$)。我们知道如果我们在 Schrodinger 表象下，算符 \hat{A} 不依赖时间，而密度矩阵 $\rho(t)$ 随时间变化；而如果在 Heisenberg 表象下，算符 $\hat{A}(t)$ 依赖时间，而密度矩阵 $\rho(t_0)$ 不随时间变化。但不管在哪种表象下，算符的期待值都是一样的。Schrodinger 表象

下，系统的本征波函数随时间的变化满足

$$|\Psi_n(t)\rangle = \hat{U}(t, t_0)|\Psi_n(t_0)\rangle, \quad (20.39)$$

这里 $\hat{U}(t, t_0)$ 为系统的演化算符 ($\hbar = 1$)

$$\hat{U}(t, t_0) = \exp \left\{ -i \int_{t_0}^t \hat{H}_F(t') dt' \right\}.$$

为了考虑外界扰动的影响，我们可以把演化算符分开写成两部分：

$$\hat{U}(t, t_0) = \exp \{-i\hat{H}(t - t_0)\} \hat{U}_I(t, t_0).$$

于是，

$$\hat{U}_I(t, t_0) = \exp \{i\hat{H}(t - t_0)\} \exp \left\{ -i \int_{t_0}^t \hat{H}_F(t') dt' \right\}$$

我们不难写出 $\hat{U}_I(t, t_0)$ 所满足的运动方程：

$$\frac{d}{dt} \hat{U}_I(t, t_0) = -iF(t)\hat{B}_I(t)\hat{U}_I(t, t_0). \quad (20.40)$$

这里我们定义了

$$\hat{B}_I(t) = \exp \{iH(t - t_0)\} \hat{B}(t_0) \exp \{-iH(t - t_0)\}$$

回忆一下高量中学习的内容，我们知道它是 \hat{B} 在相互作用表象下的表示。而且 t_0 前没有施加微扰，相互作用表象和海森堡表象在 t_0 及之前是相同的。同理， $U_I(t, t_0)$ 为相互作用表象下的演化算符。从方程 (20.40) 我们得

$$\hat{U}_I(t, t_0) = \exp \left\{ -i \int_{t_0}^t F(t') \hat{B}_I(t') dt' \right\} \quad (20.41)$$

$$= 1 - i \int_{t_0}^t F(t') \hat{B}_I(t') dt' + O(\hat{V}^2). \quad (20.42)$$

我们知道演化算符满足 $\hat{U}_I(t, t_0) = \hat{U}_I^\dagger(t_0, t)$ ， $\hat{U}_I(t, t_0)\hat{U}_I^\dagger(t, t_0) = I$ 。

现在，我们可以利用上面的结果去计算 t 时刻算符 \hat{A} 的平均值

$$\langle \hat{A}_S(t_0) \rangle(t) = \sum_n P_n \langle \Psi_n(t) | |\hat{A}_S(t_0)| |\Psi_n(t)\rangle \quad (20.43)$$

$$= \sum_n P_n \langle \Psi_n(t_0) | |\hat{U}^\dagger(t, t_0)\hat{A}(t_0)\hat{U}(t, t_0)| |\Psi_n(t_0)\rangle \quad (20.44)$$

$$\approx \sum_n P_n \langle \Psi_n(t_0) | |\hat{A}(t_0)| |\Psi_n(t_0)\rangle + \sum_n P_n \langle \Psi_n(t_0) | |\hat{A}_I(t)| \left(-i \int_{t_0}^t F(t') \hat{B}_I(t') dt' \right) | |\Psi_n(t_0)\rangle \quad (20.45)$$

$$+ \sum_n P_n \langle \Psi_n(t_0) | \left(i \int_{t_0}^t F(t') \hat{B}_I(t') dt' \hat{A}_I(t) \right) | |\Psi_n(t_0)\rangle \quad (20.45)$$

$$= \langle \hat{A}(t_0) \rangle - i \int_{t_0}^t \langle [\hat{A}_I(t), \hat{B}_I(t')] \rangle_0 F(t') dt'. \quad (20.46)$$

于是,

$$\delta\langle A(t) \rangle \equiv \langle \hat{A}(t) \rangle - \langle \hat{A}(t_0) \rangle = -i \int_{t_0}^t \langle [\hat{A}_I(t), \hat{B}_I(t')] \rangle_0 F(t') dt'. \quad (20.47)$$

这里, $[A, B] = AB - BA$, 同时 $\langle A(t_0) \rangle$ 表示相对于没有受到扰动的系统做量子平均. 因为没有外界扰动的时候系统处于平衡态, 哈密顿量不依赖时间, 所以我们有

$$\boxed{\chi_{AB}(t, t') = -i\Theta(t - t') \langle [\hat{A}_I(t), \hat{B}_I(t')] \rangle_0 = -i\Theta(t - t') \langle [\hat{A}_I(t - t'), \hat{B}_I(0)] \rangle_0.} \quad (20.48)$$

和以前一样, 这里的 Θ 为阶跃函数。我们可以看到, 系统的响应函数其实就是一个推迟格林函数。重要的一点是, 这里的算符平均是相对于没有微扰, 处于平衡态体系的。最后我们得到系统的响应为, $t_0 \rightarrow -\infty$,

$$\delta\langle A(t) \rangle = \int_{-\infty}^{+\infty} \chi_{AB}(t - t') F(t') dt'. \quad (20.49)$$

根据我们前面的讨论, 在频域我们有:

$$\delta\langle A \rangle[\omega] = \chi_{AB}[\omega] F[\omega]. \quad (20.50)$$

而且, $\chi[\omega]$ 满足我们前面得到的响应函数满足的普适关系。

以上分析没有考虑空间坐标, 我们可以把它推广到含空间坐标的形式:

$$\delta\langle A_\alpha \rangle(r, t) = \sum_\beta \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \chi_{A_\alpha B_\beta}(rt, r't') F_\beta(r't') dr' dt'. \quad (20.51)$$

如果系统具有空间平移不变性, 那么在傅里叶空间可以写为:

$$\delta\langle A_\alpha \rangle(q, \omega) = \sum_\beta \chi_{A_\alpha B_\beta}(q, \omega) F_\beta(q, \omega). \quad (20.52)$$

时间的傅里叶变换我们仍然按照前面的约定, 而空间的傅里叶变换我们使用下面的约定:

$$\begin{aligned} f[q] &= \int f(r) e^{-iqr} dr, \\ f(r) &= \int f[q] e^{iqr} \frac{dq}{2\pi}. \end{aligned} \quad (20.53)$$

20.2.1 响应函数举例：密度响应函数

考虑一个 N 电子系统, 电子的密度算符可以写成

$$\hat{n}(r) = \sum_{i=1}^N \delta(r - \hat{r}_i), \quad (20.54)$$

其傅里叶变换为

$$\hat{n}(q) = \sum_i e^{-iqr_i}. \quad (20.55)$$

20.3 响应函数的本征表示

我们可以把响应函数在 \hat{H} 的本征空间表示出来, 从而得出其一系列普适关系. 用 ε_n 表示第 n 个本征态的能量, 我们有

$$\begin{aligned} \langle [\hat{A}_I(t), \hat{B}_I(0)] \rangle_0 &= \sum_n P_n \left(\langle n | e^{iHt} \hat{A}_I(0) e^{-iHt} \hat{B}_I(0) | n \rangle \right. \\ &\quad \left. - \langle n | \hat{B}_I(0) e^{iHt} \hat{A}_I(0) e^{-iHt} | n \rangle \right) \end{aligned} \quad (20.56)$$

插入一组完备基矢, 我们得到

$$\langle [\hat{A}_I(t), \hat{B}_I(0)] \rangle = \sum_{m,n} P_n (e^{i\omega_{nm}t} A_{nm} B_{mn} - e^{i\omega_{mn}t} B_{nm} A_{mn}). \quad (20.57)$$

这样的话, 我们就可以做傅里叶变换

$$\chi_{AB}[\omega] = \sum_{m,n} P_n \left(\frac{A_{nm} B_{mn}}{\omega + i\delta - \omega_{mn}} - \frac{B_{nm} A_{mn}}{\omega + i\delta - \omega_{nm}} \right) \quad (20.58)$$

$$= \sum_{m,n} (P_n - P_m) \frac{A_{nm} B_{mn}}{\omega + i\delta - \omega_{mn}}. \quad (20.59)$$

这里, 为了保证指数函数的收敛性, 我们引入了小量 $\delta = 0^+$. 利用

$$\frac{1}{\omega + i\delta - \omega_{mn}} = \mathcal{P} \frac{1}{\omega - \omega_{mn}} - i\pi\delta(\omega - \omega_{mn}), \quad (20.60)$$

我们可以把它实部和虚部写出来

$$\text{Re}\chi_{AB}[\omega] = \sum_{m,n} (P_n - P_m) \mathcal{P} \frac{A_{nm} B_{mn}}{\omega - \omega_{mn}} \quad (20.61)$$

$$\text{Im}\chi_{AB}[\omega] = -\pi \sum_{m,n} (P_n - P_m) A_{nm} B_{mn} \delta(\omega - \omega_{mn}). \quad (20.62)$$

上面的结果满足我们前面讲过的响应函数的普遍性质. 比如:

- $\chi_{AB}[\omega]$ 在上半个复平面是解析的;
- $\chi_{AB}[-\omega] = \chi_{A^\dagger B^\dagger}^*[\omega]$. 如果 $A = A^\dagger$ 和 $B = B^\dagger$. 这也意味着 $\chi[0]$ 是实数. 如果 $A = B$, 那么 $\chi[0] < 0$.
- 如果系统具有时间反演对称性, 那么波函数可以用实数来表示, $A_{mn} = \langle \Psi_m | |\hat{A}| | \Psi_n \rangle = (\langle \Psi_n | |\hat{A}^\dagger| | \Psi_m \rangle)^* = \langle \Psi_n | |(\hat{A}^\dagger)^*| | \Psi_m \rangle = (A^T)_{nm}$, 于是 $\chi_{AB}[\omega] = \chi_{B^T A^T}[\omega]$. 如果存在外加磁场的话, 时间反演也要求磁场变号 $\chi_{AB}[\omega, B] = \chi_{B^T A^T}[\omega, -B]$. 这个是所谓的 Onsager

关系(1968年Nobel化学奖)。Onsager关系式给出了不同类型响应函数的关系,即物理量 \hat{A} 对外加与 \hat{B} 耦合扰动的响应,同与之相反的过程,即物理量 \hat{B} 对外加与 \hat{A} 耦合扰动的响应是相互关联的。

20.4 涨落-耗散定理

首先我们定义一个关联函数

$$S_{AA^\dagger}[\omega] = \int_{-\infty}^{+\infty} \langle \hat{A}(t)\hat{A}^\dagger \rangle_0 e^{i\omega t} dt. \quad (20.63)$$

叫做动态结构因子(dynamical structure factor)。下面我们说明它与量 A 的涨落相关

$$\langle \delta\hat{A}(t)\delta\hat{A}^\dagger \rangle_0 = \langle (\hat{A}(t) - \langle A(t) \rangle_0)(\hat{A}^\dagger - \langle \hat{A}^\dagger \rangle_0) \rangle_0 \quad (20.64)$$

$$= \langle \hat{A}(t)\hat{A}^\dagger \rangle_0 - \langle \hat{A}(t) \rangle_0 \langle \hat{A}^\dagger \rangle_0. \quad (20.65)$$

于是在频域我们有

$$S_{AA^\dagger}[\omega] = 2\pi|\langle A \rangle_0|^2\delta(\omega) + \int_{-\infty}^{+\infty} \langle \delta\hat{A}(t)\delta\hat{A}^\dagger \rangle_0 e^{i\omega t} dt. \quad (20.66)$$

所以我们定义的动态结构因子与涨落密切相关。

我们可以像上一节一样,在本征表象下来计算它

$$S_{AA^\dagger}[\omega] = \sum_{mn} \int_{-\infty}^{+\infty} P_n |A_{nm}|^2 e^{i(\omega+\omega_{nm})t} dt \quad (20.67)$$

$$= 2\pi \sum_{mn} P_n |A_{nm}|^2 \delta(\omega - \omega_{mn}). \quad (20.68)$$

从式子(20.62)我们得到

$$-2\text{Im}\chi_{AA^\dagger}[\omega] = 2\pi \sum_{m,n} (P_n - P_m) |A_{nm}|^2 \delta(\omega - \omega_{mn}) \quad (20.69)$$

$$= S_{AA^\dagger}[\omega] - S_{A^\dagger A}[-\omega]. \quad (20.70)$$

而

$$S_{AA^\dagger}[-\omega] = S_{A^\dagger A}[\omega]e^{-\beta\omega},$$

于是

$-2\text{Im}\chi_{AA^\dagger}[\omega] \left(n_B(\omega) + \frac{1}{2} \right) = \frac{1}{2} (S_{AA^\dagger}[\omega] + S_{A^\dagger A}[-\omega]). \quad (20.71)$

这就是我们以前遇到的涨落-耗散定理。式子左边广义极化率的虚部表示耗散,而式子右边的

动态结构因子表示涨落。我们还可以把它在时域表示出来

$$-2\text{Im}\chi_{AA^\dagger}[\omega] = \coth^{-1}(\beta\hbar\omega/2) \int dt e^{i\omega t} \langle \{A_I(t), A_I^\dagger(0)\} \rangle. \quad (20.72)$$

这里 $\{A, B\} = AB + BA$ 。

20.5 线性响应理论的应用: 自由电子气对电磁场的响应

这里以自由电子气与电磁场的耦合为例, 考虑线性响应理论的基本应用。

20.5.1 真空中的 Maxwell 方程

真空中描述电磁场的 Maxwell 方程可以写为:

$$\vec{\nabla} \cdot \vec{E} = \frac{1}{\epsilon_0} \rho \quad (20.73)$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad (20.74)$$

$$\vec{\nabla} \cdot \vec{B} = 0 \quad (20.75)$$

$$\vec{\nabla} \times \vec{B} = \mu_0 (\vec{J} + \epsilon_0 \frac{\partial \vec{E}}{\partial t}) \quad (20.76)$$

其中, 式 (20.75) 告诉我们, 磁场可以用某一个函数 \vec{A} 的旋度来表示 $\vec{B} = \vec{\nabla} \times \vec{A}$ 。这样就可以保证 (20.75) 成立。把这个表达式带入到 (20.74) 中, 我们可以得到 $\vec{E} = -\frac{\partial \vec{A}}{\partial t}$ 。事实证明, 这样写还不完整, 我们还需要一个标量场来描述纵向电场使其满足 (20.73)。综合起来, 电磁场可以用标势 $\phi(\vec{r}, t)$ 和矢势 $A(\vec{r}, t)$ 来描述:

$$\vec{E}(\vec{r}, t) = -\vec{\nabla} \phi(\vec{r}, t) - \frac{\partial \vec{A}(\vec{r}, t)}{\partial t}, \quad (20.77)$$

$$\vec{B}(\vec{r}, t) = \vec{\nabla} \times \vec{A}(\vec{r}, t). \quad (20.78)$$

20.5.2 规范不变性

描述电磁场的矢势和标势满足规范不变性。如果我们同时对矢势和标势做如下的变换

$$\vec{A}(\vec{r}, t) \rightarrow \vec{A}(\vec{r}, t) + \vec{\nabla} \Lambda(\vec{r}, t) \quad (20.79)$$

$$\phi(\vec{r}, t) \rightarrow \phi(\vec{r}, t) - \frac{\partial \Lambda(\vec{r}, t)}{\partial t} \quad (20.80)$$

并不影响电磁场 \vec{E} 和 \vec{B} 。

利用 (20.77) 和 (20.78) 两个式子，我们可以把其余两个方程表示为：

$$\frac{\partial}{\partial t} \vec{\nabla} \cdot \vec{A} + \nabla^2 \phi = -\frac{1}{\epsilon_0} \rho, \quad (20.81)$$

$$\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{A} - \nabla^2 \vec{A} + \nabla \left(\vec{\nabla} \cdot \vec{A} + \frac{1}{c^2} \frac{\partial}{\partial t} \phi \right) = \mu_0 \vec{J}. \quad (20.82)$$

这一组方程看起来很丑，我们可以利用规范选择的自由度来得到更为简洁或者更为对称的形式。

- 库仑规范：我们可以选择

$$\vec{\nabla} \cdot \vec{A} = 0$$

这时，方程 (20.81-20.82) 变为：

$$\nabla^2 \phi = -\frac{1}{\epsilon_0} \rho, \quad (20.83)$$

$$\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{A} - \nabla^2 \vec{A} + \frac{1}{c^2} \frac{\partial}{\partial t} \nabla \phi = \mu_0 \vec{J}. \quad (20.84)$$

这个选择的好处是，标势完全由电荷分布根据库仑定律即时决定，没有任何推迟效应。

- 洛伦兹规范：我们选择

$$\vec{\nabla} \cdot \vec{A} + \frac{1}{c^2} \frac{\partial}{\partial t} \phi = 0. \quad (20.85)$$

这样的话，方程 (20.81) 和 (20.82) 变为：

$$\square \phi = \frac{1}{\epsilon_0} \rho, \quad (20.86)$$

$$\square \vec{A} = \mu_0 \vec{J}. \quad (20.87)$$

这里，我们定义了 d'Alembertian 算符 $\square \phi = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \phi - \nabla^2 \phi$ 。这样做好处是关于标势和矢势的方程具有相同的形式。左边都是 d'Alembertian 算符，右边为他们对应电荷和电流源。

20.5.3 纵场、横场、电磁波方程

我们还可以把电磁场分为横向和纵向两部分。我们规定

$$\vec{E} = \vec{E}_\perp + \vec{E}_\parallel, \quad \nabla \cdot \vec{E}_\perp = 0, \quad \nabla \times \vec{E}_\parallel = 0. \quad (20.88)$$

对于 \vec{B} 和 \vec{A} 我们可以做类似规定。这样的话, 从式 (20.75) 我们发现 $\vec{B} = \vec{B}_\perp$ 为横场, 而 (20.73) 给出的电场为纵向部分, (20.74) 给出的电场为横向部分。于是我们得到:

$$\vec{\nabla} \cdot \vec{E}_\parallel = \frac{1}{\epsilon_0} \rho \quad (20.89)$$

$$\vec{\nabla} \times \vec{E}_\perp = -\frac{\partial \vec{B}_\perp}{\partial t} \quad (20.90)$$

$$\vec{\nabla} \cdot \vec{B}_\perp = 0 \quad (20.91)$$

$$\vec{\nabla} \times \vec{B}_\perp = \mu_0 (\vec{J}_\perp + \epsilon_0 \frac{\partial \vec{E}_\perp}{\partial t}) \quad (20.92)$$

其中, 根据电荷守恒

$$-\nabla \cdot \vec{J}_\parallel = \frac{\partial \rho}{\partial t}, \quad (20.93)$$

我们利用了

$$\vec{J}_\parallel + \frac{1}{\epsilon_0} \frac{\partial \vec{E}_\parallel}{\partial t} = 0. \quad (20.94)$$

它可以由 (20.92) 得到。这样的话, 我们成功的把 Maxwell 方程写成了横场和纵场所满足的方程。显而易见, 由电荷所产生的电场为纵场, 而在真空中传播的电磁场为横场。把式 (20.90) 和 (20.92) 联立, 我们还可以得到描述横场电磁波的波动方程:

$$\nabla^2 \vec{B}_\perp - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{B}_\perp = -\mu_0 \nabla \times \vec{J}_\perp, \quad (20.95)$$

$$\nabla^2 \vec{E}_\perp - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{E}_\perp = \mu_0 \frac{\partial}{\partial t} \vec{J}_\perp. \quad (20.96)$$

20.5.4 Fourier 变换形式

上述 Maxwell 方程, 特别是描述电磁场的运动方程 (20.81) 和 (20.82) 是在时间和空间上的微分方程。正是这样的微分方程的形式保证了电磁波在时空的传播。为了简化 Maxwell 方程的形式, 我们可以对之进行 Fourier 变换。按照我们的惯例, 我们有

$$\vec{E}(\vec{r}, t) = \frac{1}{(2\pi)^4} \int d\vec{k} \int d\omega \vec{E}(\vec{k}, \omega) e^{i(\vec{k} \cdot \vec{r} - \omega t)}. \quad (20.97)$$

其他量具有类似的变换形式。这样的话, 我们就有如下的变换:

$$\frac{\partial}{\partial t} \vec{E}(\vec{r}, t) \rightarrow -i\omega \vec{E}(\vec{k}, \omega), \quad (20.98)$$

$$\nabla \cdot \vec{E}(\vec{r}, t) \rightarrow i\vec{k} \cdot \vec{E}(\vec{k}, \omega). \quad (20.99)$$

在 Fourier 空间, 不同角频率和波矢的模式之间就变为相互独立的了

$$ikE_{\parallel} = \frac{1}{\epsilon_0}\rho \quad (20.100)$$

$$i\vec{k} \times \vec{E}_{\perp} = i\omega \vec{B}_{\perp} \quad (20.101)$$

$$i\vec{k} \cdot \vec{B}_{\perp} = 0 \quad (20.102)$$

$$i\vec{k} \times \vec{B}_{\perp} = \mu_0(\vec{J}_{\perp} - i\omega\epsilon_0\vec{E}_{\perp}) \quad (20.103)$$

电荷守恒变为了

$$\vec{k} \cdot \vec{J}_{\parallel} = \omega\rho. \quad (20.104)$$

20.5.5 耦合哈密顿量

电磁场与自由电子气的耦合, 可以表示为如下的“标准”哈密顿量

$$\hat{H}(t) = \sum_i -\frac{1}{2m} \left(\hat{\vec{p}}_i + e\vec{A}(\vec{r}_i, t) \right)^2 + (-e) \sum_i \phi(\vec{r}_i, t) \quad (20.105)$$

$$= \sum_i -\frac{1}{2m} \left(\hat{\vec{p}}_i + e\vec{A}(\vec{r}_i, t) \right)^2 + (-e) \int \hat{n}(\vec{r}, t) \phi(\vec{r}, t) d\vec{r}. \quad (20.106)$$

这里 $e > 0$ 为基本电荷单位, 电子密度算符为

$$\hat{n}(\vec{r}, t) = \sum_i \delta(\vec{r} - \vec{r}_i) \quad (20.107)$$

我们看到, 标势与密度算符耦合, 矢势与动量算符耦合。我们这里把电场和磁场作为经典的物理量来考虑它们和自由电子气的耦合。

20.5.6 与电场的耦合

20.5.6.1 与标势的耦合

首先考虑由于外界电荷所引起的纵向电场 $\vec{E}_{\parallel}(\vec{r}, t)$, 我们可以把它用标势表示为 $\vec{E}_{\parallel}(\vec{r}, t) = -\nabla\phi(\vec{r}, t)$ 。我们把耦合哈密顿量表示为:

$$H_{int} = \int \hat{n}(\vec{r}, t) V(\vec{r}, t) d\vec{r} \quad (20.108)$$

其中 $V(\vec{r}, t) = (-e)\phi(\vec{r}, t)$ 。对比前面的普适结果, 我们可以做下替换: $V(\vec{r}, t) \rightarrow F(\vec{r}, t), \vec{A} \rightarrow \hat{n}, \hat{B} \rightarrow \hat{n}$ 就可以得到

$$\langle \delta n(\vec{r}, t) \rangle = \iint \chi_{\hat{n}\hat{n}}(\vec{r}, t; \vec{r}', t') V(\vec{r}', t') dt' d\vec{r}' \quad (20.109)$$

同时密度响应函数为：

$$\chi_{\hat{n}\hat{n}}(\vec{r}, t; \vec{r}', t') = (i\hbar)^{-1} \Theta(t - t') \langle [\hat{n}(\vec{r}, t), \hat{n}(\vec{r}', t')] \rangle_0 \quad (20.110)$$

我们下面考虑密度响应函数在自由电子气本征表象下的表示。

20.5.6.2 自由电子气哈密顿量

自由电子气无耦合时的哈密顿量在波矢 \vec{k} 空间为：

$$\hat{H} = \varepsilon_k c_k^\dagger c_k.$$

其中

$$\varepsilon(k) = \frac{\hbar^2 k^2}{2m}$$

对应波矢 \vec{k} 的波函数为平面波：

$$\langle \vec{r} | \vec{k} \rangle = \psi_{\vec{k}}(\vec{r}) = \frac{1}{\sqrt{V}} e^{i\vec{k} \cdot \vec{r}},$$

这里暂时忽略自旋自由度。

下面我们得出二次量子化表象下的密度算符 $\hat{n}(\vec{r})$ 的表达式。在实空间下，

$$\hat{n}(\vec{r}) = \sum_i \delta(\vec{r} - \vec{r}_i) \quad (20.111)$$

傅里叶变换到动量空间中有

$$\hat{n}(\vec{q}) = \int \hat{n}(\vec{r}) e^{-i\vec{q} \cdot \vec{r}} d\vec{r} = \sum_i e^{-i\vec{q} \cdot \vec{r}_i} = \sum_i \vec{n}_i(\vec{q}) \quad (20.112)$$

一次量子化矩阵元：

$$\langle k' | \hat{n}_i(q) | k \rangle = \frac{1}{V} \int e^{i(\vec{k} - \vec{k}' - \vec{q}) \cdot \vec{r}_i} d\vec{r}_i = \delta_{k, k'+q} \quad (20.113)$$

其中 Kronecker δ 函数定义为

$$\delta_{k, k'} = \begin{cases} 1, & (\vec{k} = \vec{k}') \\ 0, & (\vec{k} \neq \vec{k}') \end{cases} \quad (20.114)$$

于是在二次量子化表象下，密度算符可以表示为：

$$\hat{n}(\vec{q}) = \sum_{k, k'} \langle k' | \hat{n}(q) | k \rangle c_{k'}^\dagger c_k = \sum_{k, k'} n_{k' k} c_{k'}^\dagger c_k = \sum_{k, k'} \delta_{k'+q, k} c_{k'}^\dagger c_k \quad (20.115)$$

后面我们会在没有歧义的情况下省略矢量标记。

20.5.6.3 密度关联函数的频率空间表示

利用上面的结果，我们就可以计算密度关联函数：

$$\chi_{\hat{n}\hat{n}}[\vec{q}, \omega] = \iint \chi_{\hat{n}\hat{n}}(r, t; r', t') e^{-i\vec{q} \cdot (\vec{r} - \vec{r}')} e^{i\omega(t-t')} d(t-t') d(\vec{r} - \vec{r}') \quad (20.116)$$

$$= -i \int_0^{+\infty} d(t-t') \int_{-\infty}^{+\infty} \langle [\hat{n}_I(\vec{r}, t), \hat{n}_I(\vec{r}', t')] \rangle e^{-i\vec{q} \cdot (\vec{r} - \vec{r}')} e^{i\omega(t-t')} d(\vec{r} - \vec{r}')$$

$$= \frac{-i}{L^d} \int_0^{+\infty} e^{i\omega(t-t')} d(t-t') \int_{-\infty}^{+\infty} \langle [\hat{n}_I(\vec{r}, t), \hat{n}_I(\vec{r}', t')] \rangle e^{-i\vec{q} \cdot (\vec{r} - \vec{r}')} d\vec{r} d\vec{r}'$$

$$= \frac{-i}{L^d} \int_0^{+\infty} \langle [\hat{n}_I(\vec{q}, t), \hat{n}_I(-\vec{q}, t')] \rangle e^{i\omega(t-t')} d(t-t')$$

(20.117)

以上我们完成了空间的 Fourier 变化。对于时间部分我们可以采用与前面相同的方法（见下框中内容）。

$$\begin{aligned} \chi_{nn}[\vec{q}, \omega] &= - \sum_k P_k \frac{i}{L^d} \int_0^{+\infty} [e^{i(\omega+i0^+ - (\omega_{k'} - \omega_k))(t-t')} n_{kk'}(q) n_{k'k}(-q) c_k^\dagger c_{k'} c_{k'}^\dagger c_k \\ &\quad - e^{i(\omega+i0^+ - (\omega_k - \omega_{k'}))(t-t')} n_{kk'}(-q) n_{k'k}(q) c_k^\dagger c_{k'} c_{k'}^\dagger c_k] \\ &= L^{-d} \sum_k P_k \frac{c_k^\dagger c_{k+q} c_{k+q}^\dagger c_k}{\omega + i0^+ - (\omega_{k+q} - \omega_k)} \\ &\quad - L^{-d} \sum_k P_{k+q} \frac{c_{k+q}^\dagger c_k c_k^\dagger c_{k+q}}{\omega + i0^+ - (\omega_{k+q} - \omega_k)} \\ &= \frac{1}{L^d} \sum_k \frac{\langle c_k^\dagger c_{k+q} c_{k+q}^\dagger c_k \rangle - \langle c_{k+q}^\dagger c_k c_k^\dagger c_{k+q} \rangle}{\omega + i0^+ - (\omega_{k+q} - \omega_k)} \end{aligned}$$

这里，我们采用另外一个方式。相互作用表象下的密度算符的时间依赖表示为：

$$\hat{n}_I(\vec{q}, t) = \sum_{k_1} c_{k_1}^\dagger c_{k_1+q} e^{-i(\omega_{k_1+q} - \omega_{k_1})t}.$$

$$\hat{n}_I(-\vec{q}, t) = \sum_{k_2} c_{k_2}^\dagger c_{k_2-q} e^{-i(\omega_{k_2-q} - \omega_{k_2})t}.$$

利用以上关系，我们可以把所有的时间依赖提取出来：

$$= \sum_{k_1, k_2} (iL^d)^{-1} \int_0^{+\infty} e^{-i(\omega_{k_1+q} - \omega_{k_1})t} e^{-i(\omega_{k_2-q} - \omega_{k_2})t'} e^{i\omega(t-t')} \langle [\hat{n}(\vec{q}), \hat{n}(-\vec{q})] \rangle d(t-t')$$

利用产生、湮灭算符的反对易关系，我们可以得到：

$$\begin{aligned}\langle [\hat{n}(\vec{q}), \hat{n}(-\vec{q})] \rangle &= \sum_{k_1, k_2} (\langle c_{k_1}^\dagger c_{k_1+q} c_{k_2}^\dagger c_{k_2-q} \rangle - \langle c_{k_2}^\dagger c_{k_2-q} c_{k_1}^\dagger c_{k_1+q} \rangle) \\ &= \delta_{k_1+q, k_2} (f(\varepsilon_{k_1}) - f(\varepsilon_{k_1} + q)).\end{aligned}$$

这里，

$$f(k) = \frac{1}{\exp[(\varepsilon_k - \mu)/k_B T] + 1}$$

为费米分布函数。于是，我们得到：

$$\begin{aligned}\chi_{nn}[\vec{q}, \omega] &= \sum_{k_1} (iL^d)^{-1} \int_0^{+\infty} e^{i(\omega - (\omega_{k_1+q} - \omega_{k_1})) (t - t')} d(t - t') (f(k_1) - f(k_1 + q)) \\ &= \frac{1}{L^d} \sum_k \frac{f(\varepsilon_k) - f(\varepsilon_{k+q})}{\varepsilon + i0^+ - (\varepsilon(k+q) - \varepsilon(k))}.\end{aligned}$$

最后我们把 \hbar 放了回来。上面的结果又被称为 Lindhard 函数，其三维形式是由丹麦物理学家 Lindhard 在 1954 年最早得出的。它还可以被写成以下几种不同的形式：

$$\chi_{nn}[\vec{q}, \omega] = \frac{1}{L^d} \sum_k \frac{f(\varepsilon_k) - f(\varepsilon_{k+q})}{\varepsilon + i0^+ + \varepsilon_k - \varepsilon_{k+q}} \quad (20.118)$$

$$= \frac{1}{L^d} \sum_k f(\varepsilon_k) \left(\frac{1}{\varepsilon + i0^+ + \varepsilon_k - \varepsilon_{k+q}} - \frac{1}{\varepsilon + i0^+ + \varepsilon_{k-q} - \varepsilon_k} \right) \quad (20.119)$$

$$= \frac{1}{L^d} \sum_k \left(\frac{f(\varepsilon_k)(1 - f(\varepsilon_{k+q}))}{\varepsilon + i0^+ + \varepsilon_k - \varepsilon_{k+q}} - \frac{f(\varepsilon_k)(1 - f(\varepsilon_{k-q}))}{\varepsilon + i0^+ + \varepsilon_{k-q} - \varepsilon_k} \right). \quad (20.120)$$

对于自由电子气来说， $\omega_k = \omega_{-k}$ ，所以我们还有

$$(k \rightarrow -k) = \frac{1}{L^d} \sum_k f(\varepsilon_k) \left(\frac{1}{\varepsilon + i0^+ + \varepsilon_k - \varepsilon_{k+q}} - \frac{1}{\varepsilon + i0^+ + \varepsilon_{k+q} - \varepsilon_k} \right) \quad (20.121)$$

$$= \frac{1}{L^d} \sum_k f(\varepsilon_k) [1 - f(\varepsilon_{k+q})] \left(\frac{1}{\varepsilon + i0^+ + \varepsilon_k - \varepsilon_{k+q}} - \frac{1}{\varepsilon + i0^+ + \varepsilon_{k+q} - \varepsilon_k} \right). \quad (20.122)$$

有意思的是，这里式 (20.121) 和 (20.122) 相差一个因子，但它们是等价的。

20.5.6.4 与矢势的耦合

矢势与自由电子气耦合的哈密顿量在线性近似下为：

$$\hat{H}_{int} \approx e \int \hat{j}_p(\vec{r}) \cdot \vec{A}(\vec{r}) d\vec{r}, \quad (20.123)$$

这里, 我们定义了顺磁粒子流

$$\hat{j}_p(\vec{r}) = \frac{1}{2m} \sum_i \left(\hat{\vec{p}}_i \delta(\vec{r} - \hat{\vec{r}}_i) + \delta(\vec{r} - \hat{\vec{r}}_i) \hat{\vec{p}}_i \right). \quad (20.124)$$

我们首先把 j_p 在动量空间表示出来:

$$\hat{j}_p(\vec{q}) = \frac{1}{m} \sum_i \left(\hat{\vec{p}}_i + \frac{\vec{q}}{2} \right) e^{-i\vec{q}\cdot\hat{\vec{r}}_i} \quad (20.125)$$

它在自由电子气本征表象下的矩阵元为:

$$\langle \vec{k}' | \hat{j}_p(\vec{q}) | \vec{k} \rangle = \frac{1}{m} \sum_i \left(\vec{k}' + \frac{\vec{q}}{2} \right) \delta_{\vec{k}'+\vec{q},\vec{k}}. \quad (20.126)$$

于是, 二次量子化表象下的顺磁流密度算符为

$$\hat{j}_p(\vec{q}) = \frac{1}{m} \sum_{\vec{k}\vec{k}'} \left(\vec{k}' + \frac{\vec{q}}{2} \right) \hat{c}_{\vec{k}}^\dagger \hat{c}_{\vec{k}} \delta_{\vec{k}'+\vec{q},\vec{k}}. \quad (20.127)$$

我们可以把这个耦合分解成纵向和横向两部分:

$$\hat{j}_p(\vec{r}) \cdot \vec{A}(\vec{r}) = \hat{j}_{p,\parallel}(\vec{r}) \cdot \vec{A}_{\parallel}(\vec{r}) + \hat{j}_{p,\perp}(\vec{r}) \cdot \vec{A}_{\perp}(\vec{r}) \quad (20.128)$$

相应的顺磁流密度关联函数

$$\chi_{j_p j_p, \parallel}(\vec{r}, t; \vec{r}', t') = -i\Theta(t - t') \langle [\hat{j}_{p,\parallel}(\vec{r}, t), \hat{j}_{p,\parallel}(\vec{r}', t')] \rangle_0, \quad (20.129)$$

$$\chi_{j_p j_p, \perp}(\vec{r}, t; \vec{r}', t') = -i\Theta(t - t') \langle [\hat{j}_{p,\perp}(\vec{r}, t), \hat{j}_{p,\perp}(\vec{r}', t')] \rangle_0 \quad (20.130)$$

代入到式 (20.59) 中, 我们可以直接得到响应函数:

$$\chi_{j_p j_p, \parallel}[\vec{q}, \omega] = \sum_{m,n} P_n |\langle n | \hat{j}_{p,\parallel}(\vec{q}) | m \rangle|^2 \left(\frac{1}{\omega + i\delta - \omega_{mn}} - \frac{1}{\omega + i\delta - \omega_{nm}} \right), \quad (20.131)$$

$$\chi_{j_p j_p, \perp}[\vec{q}, \omega] = \sum_{m,n} P_n |\langle n | \hat{j}_{p,\perp}(\vec{q}) | m \rangle|^2 \left(\frac{1}{\omega + i\delta - \omega_{mn}} - \frac{1}{\omega + i\delta - \omega_{nm}} \right). \quad (20.132)$$

20.5.7 密度响应: 介电函数

下面我们考虑如何使用密度关联函数来描述介电屏蔽效应。考虑电磁波辐照到材料上, 电磁波的交变电场会对材料内部的带电粒子施加一个交变的电场力。在电场力的作用下, 材料内部的带电粒子可能会偏离原来的平衡位置发生小的位移。这一位移会导致内部的电荷分布发生变化, 从而诱导出电场/电势。最后, 材料内部的电场/电势是外加电场/电势和诱导电场/电势之和。这一过程可以表示为:

$$\phi_{ext} \rightarrow \delta n \rightarrow \delta \phi \rightarrow \phi_{eff} = \phi_{ext} + \delta \phi, \vec{E}_{eff} = \vec{E}_{ext} + \delta \vec{E}. \quad (20.133)$$

其中 ϕ_{ext} 是外加电磁场对应的电势, δn 和 $\delta\phi$ 是外加电磁波扰动下系统中电子密度和电势的改变, 而 ϕ_{eff} 为最后的有效电势。各向同性材料的相对介电函数可以表示为外加电势和有效电势之比:

$$\varepsilon_r = \frac{\phi_{ext}}{\phi_{eff}} = \frac{\vec{E}_{ext}}{\vec{E}_{eff}} \quad (20.134)$$

本节中的 \vec{E} 实际上是 $\vec{E}_{||}$, 因为这是由于电荷导致的响应, 为了简化我们忽略了下标. 我们要解决的问题就是怎么从密度关联函数得到介电函数 ε_r (同理, 实际上应该为 $\varepsilon_{r,||}$)。

1. 首先, 我们可以利用泊松方程联系 δn 和 $\delta\phi$:

$$\nabla^2 \delta\phi = -\frac{1}{\varepsilon_0} \delta n(-e) = \frac{e}{\varepsilon_0} \delta n \quad (20.135)$$

如果系统具有空间和时间平移不变性, 我们可以写出在傅里叶变换后的关系式:

$$-q^2 \delta\phi[q, \omega] = \frac{e}{\varepsilon_0} \delta n[q, \omega] \quad (20.136)$$

2. 其次, 利用线性响应理论, 我们可以把 δn 和 $\delta\phi$ 联系起来:

$$\delta n[q, \omega] = \chi[q, \omega] \phi_{ext}[q, \omega](-e) \quad (20.137)$$

于是:

$$\delta\phi = -\frac{e}{\varepsilon_0 q^2} \delta n[q, \omega] = \frac{e^2}{\varepsilon_0 q^2} \chi[q, \omega] \phi_{ext}[q, \omega] \quad (20.138)$$

3. 最后我们可以得到介电函数:

$$\varepsilon_r[q, \omega] = \frac{\phi_{ext}}{\phi_{eff}} = \frac{1}{1 + \frac{e^2}{\varepsilon_0 q^2} \chi[q, \omega]} \quad (20.139)$$

由于 $\chi[q, 0] < 0$, 所以 $\varepsilon_r[q, 0] > 1$, 即有效电势 (电场) 比外加电势 (电场) 要弱。也就是说, 系统倾向于屏蔽外场的影响.

4. 上面的分析中得到的 $\chi[q, \omega]$ 实际上隐含着库仑, 因为它把外加电磁场和系统的密度响应联系了起来

$$\delta n = \chi \phi_{ext}(-e) \quad (20.140)$$

我们还可以考虑一个"自由电子" 版本的密度响应函数 χ_0 。它联系的是系统内的总电势和系统的密度响应

$$\delta n = -e \chi_0 \phi_{eff} \quad (20.141)$$

这样, 我们就把库仑相互作用转移到 ϕ_{eff} 中。利用上面的关系式, 我们可以得出

$$\frac{\chi_0[q, \omega]}{\chi[q, \omega]} = \frac{\phi_{ext}}{\phi_{eff}} = \varepsilon_r = \left[1 + \frac{e^2}{\varepsilon_0 q^2} \chi[q, \omega] \right]^{-1} \quad (20.142)$$

$$\chi[q, \omega] = \frac{\chi_0[q, \omega]}{1 - \frac{e^2}{\varepsilon_0 q^2} \chi_0[q, \omega]} \quad (20.143)$$

$$\varepsilon_r[q, \omega] = 1 - \frac{e^2}{\varepsilon_0 q^2} \chi_0[q, \omega] \quad (20.144)$$

其中的 $\frac{e^2}{\varepsilon_0 q^2}$ 是库仑相互作用的傅里叶变换。可以看出, 我们上一节得出的 Lindhard 函数是在自由电子模型下得出来的, 不包含库仑相互作用, 所以它应该对应我们这里的 χ_0 。

库仑势 $\phi(r) = 1/|r|$ 的傅里叶空间表示:

$$\begin{aligned} \phi(\vec{q}) &= \lim_{\delta \rightarrow 0^+} \int d^3 \vec{r} e^{-i\vec{r} \cdot \vec{q}} \frac{e^{-\delta r}}{|\vec{r}|} \\ &= \lim_{\delta \rightarrow 0^+} 2\pi \int_0^\pi d\theta \sin \theta \int_0^{+\infty} dr r^2 e^{-irq \cos \theta} \frac{1}{r} e^{-\delta r} \\ &= \lim_{\delta \rightarrow 0^+} 2\pi \int_{-1}^1 e^{-irqy} dy \int_0^{+\infty} dr r^2 \frac{1}{r} e^{-\delta r} \\ &= \lim_{\delta \rightarrow 0^+} \frac{4\pi}{q^2 + \delta^2}. \end{aligned} \quad (20.145)$$

介电函数是另一个联系体系的微观粒子运动、相互作用与宏观光学、电学性质的重要函数。比如材料的光学折射率、电导率、光学吸收、等离激元振荡等性质都与介电函数有关。

20.5.7.1 与电导率关系

我们前面考虑了在外加纵向电场作用下系统的密度响应。我们还可以考虑电流响应。比如, 考虑各向同性的体系, 我们根据欧姆定律有

$$J_{||}[q, \omega] = \sigma_{||}[q, \omega] E_{||}[q, \omega]. \quad (20.146)$$

根据电荷守恒 $\omega(-e)\delta n = \vec{q} \cdot \vec{J}_{||}$, 以及 $\vec{E}_{||} = -i\vec{q}\phi$, 我们可以建立一个电导率 $\sigma_{||}$ 和介电函数 $\varepsilon_{r,||}$ 之间的关系

$$\sigma_{||}(q, \omega) = \frac{ie^2 \omega}{q^2} \chi_0[q, \omega]. \quad (20.147)$$

$$\varepsilon_r(q, \omega) = 1 - \frac{1}{i\omega\varepsilon_0} \sigma_{||}(q, \omega). \quad (20.148)$$

20.5.7.2 Lindhard 介电函数

我们把上节得到的 Lindhard 函数带入到介电函数表达式中，就得到了 Lindhard 介电函数

$$\varepsilon_r[q, \omega] = 1 - \frac{e^2}{\varepsilon_0 q^2} \chi_0[q, \omega] = 1 - \frac{e^2}{\varepsilon_0 q^2 L^d} \sum_k \frac{f(\varepsilon_k) - f(\varepsilon_{k+q})}{\omega + i\delta - (\varepsilon_{k+q} - \varepsilon_k)} \quad (20.149)$$

下面对 Lindhard 介电函数做一些讨论：

- 只有 $f(\varepsilon_n) \neq f(\varepsilon_m)$ 才对 $\varepsilon_r[q, \omega]$ 有贡献。如果考虑 $\omega \rightarrow 0$ ，那么只有费米面附近的电子才会对介电函数有贡献

$$\chi_0[q, \omega \rightarrow 0] \approx \frac{1}{L^d} \sum_k \frac{f(\varepsilon_k) - f(\varepsilon_{k+q})}{\varepsilon_k - \varepsilon_{k+q}} \quad (20.150)$$

$$\approx -D(\varepsilon_F). \quad (20.151)$$

其中 $D(\varepsilon_F)$ 是费米能级态密度。于是，长波极限下

$$\varepsilon_r[q, \omega \rightarrow 0] \approx 1 + \frac{k_{TF}^2}{q^2} \quad (20.152)$$

其中 $k_{TF}^2 = \frac{e^2}{\varepsilon_0} D(\varepsilon_F)$ ，而 k_{TF} 是 Thomas-Fermi 屏蔽波矢。可见， $D(\varepsilon_F)$ 越大，有越多的电子对外界电荷进行响应，屏蔽效应越明显。Thomas-Fermi 对库仑屏蔽的近似可以看作是 Lindhard 近似的一个极限情形。对于一个点电荷 Ze 形成的库仑势，屏蔽势为

$$V_{sc}[q] = \frac{4\pi}{\varepsilon_r[q]} \frac{Ze}{\varepsilon_0 q^2} = \frac{4\pi Ze}{\varepsilon_0 (q^2 + k_{TF}^2)} \quad (20.153)$$

Fourier 变换到实空间之后，我们发现屏蔽波矢 k_{TF} 表征了库仑势由于屏蔽效应而衰减的快慢

$$V_{sc}(r) = \frac{Ze}{\varepsilon_0 r} e^{-k_{TF} r}. \quad (20.154)$$

而 k_{TF}^{-1} 为系统的特征屏蔽长度。

- 高频极限：重新把 $\chi_0[q, \omega]$ 写成：

$$\chi_0[q, \omega] = \frac{1}{L^d} \sum_k \left(\frac{f(\varepsilon_k)}{\omega + i\delta - (\varepsilon_{k+q} - \varepsilon_k)} - \frac{f(\varepsilon_k)}{\omega + i\delta - (\varepsilon_k - \varepsilon_{k-q})} \right) \quad (20.155)$$

$$= \frac{1}{L^d} \sum_k f(\varepsilon_k) \left(\frac{1}{\omega + i\delta - (\varepsilon_{k+q} - \varepsilon_k)} - \frac{1}{\omega + i\delta - (\varepsilon_k - \varepsilon_{k-q})} \right) \quad (20.156)$$

$$= \frac{1}{L^d} \sum_k f(\varepsilon_k) \frac{2(\varepsilon_{k+q} - \varepsilon_k)}{(\omega + i\delta)^2 - (\varepsilon_{k+q} - \varepsilon_k)^2} \quad (20.157)$$

在长波极限下， $q \rightarrow 0$ ，

$$\chi_0[0, \omega] \approx \frac{1}{L^d} \left(\sum_k f(\varepsilon_k) \right) \times \frac{2q^2}{2m} \cdot \frac{1}{\omega^2} \quad (20.158)$$

$$= \frac{nq^2}{m\omega^2} \quad (20.159)$$

这里电子密度 $n = \frac{1}{L^d} (\sum_n f(\varepsilon_n))$ 。于是，高频、长波极限下，介电函数可以简化为

$$\varepsilon[0, \omega] \approx 1 - \frac{e^2}{\varepsilon_0 q^2} \cdot \frac{nq^2}{m\omega^2} = 1 - \frac{\omega_p^2}{\omega^2} \quad (20.160)$$

其中 $\omega_p^2 = \frac{ne^2}{m\varepsilon_0}$ 为等离子体频率。等离子体频率可以用经典的电磁学角度来得到。可见，当 $\omega < \omega_p$ 的时候， $\varepsilon < 0$ ，这时电磁波无法在材料内部传播。只有当 $\omega > \omega_p$ 时，才可以传播。我们利用线性响应理论，得出了这些电磁学中的基本结论。

20.5.8 电流响应

下面我们考虑系统与外加电磁场中矢势的耦合。根据上面的讨论，电磁场矢势与自由电子气耦合的哈密顿量可以表示为：

$$\hat{H}_i = e \int \hat{j}_{p,\parallel}(\vec{r}) \cdot \vec{A}_{\parallel}(\vec{r}) d\vec{r}, \quad (20.161)$$

总的粒子流 $\hat{j}_{\parallel} = \hat{j}_{p,\parallel} + \hat{j}_{d,\parallel}$ 包括抗磁部分

$$\langle \delta \hat{j}_{d,\parallel}(\vec{r}) \rangle = \frac{n}{m} e \vec{A}_{\parallel}(\vec{r}). \quad (20.162)$$

和顺磁部分

$$\langle \delta \hat{j}_{\parallel} \rangle = e \chi_{j_{j0},\parallel} \vec{A}_{\parallel}. \quad (20.163)$$

根据式20.77，矢势贡献的电场强度在傅里叶空间可以表示为：

$$\vec{E}[\vec{q}, \omega] = i\omega \vec{A}[\vec{q}, \omega].$$

于是，我们可以得到电导率和电流响应函数之间的关系

$$\boxed{\sigma_{\parallel}[\vec{q}, \omega] = \frac{ie^2}{\omega} \left(\frac{n[\vec{q}, \omega]}{m} + \chi_{j_p j_{p0},\parallel}[\vec{q}, \omega] \right)}, \quad (20.164)$$

以及

$$\boxed{\chi_{j_p j_{p0},\parallel} = \frac{\omega^2}{q^2} \chi_0}. \quad (20.165)$$

我们可以把总的纵向电流响应函数写成

$$\chi_{j_{j0},\parallel}[\vec{q}, \omega] = \chi_{j_p j_{p0},\parallel}[\vec{q}, \omega] + \frac{n[\vec{q}, \omega]}{m}. \quad (20.166)$$

下面我们考虑零频极限。利用式(20.186)可以证明

$$\operatorname{Re}\chi_{j_p j_p 0, \parallel}[\vec{q}, 0] = \sum_{m,n} P_n |\langle n | \hat{j}_{p,\parallel}(\vec{q}) | m \rangle|^2 \frac{2}{\omega_{nm}} = -\frac{n}{m}. \quad (20.167)$$

于是，我们有：

$$\operatorname{Re}\chi_{j_j 0, \parallel}[\vec{q}, 0] = 0, \quad (20.168)$$

零频极限下的电导率可以表示为：

$$\boxed{\sigma_{\parallel}[\vec{q}, \omega = 0] = -e^2 \lim_{\omega \rightarrow 0} \frac{\operatorname{Im}\chi_{j_p j_p 0, \parallel}[\vec{q}, \omega]}{\omega}.} \quad (20.169)$$

20.5.8.1 电流响应函数验证：

利用式(20.131)，我们有

$$\chi_{j_p j_p 0, \parallel}[\vec{q}, \omega] = \sum_{m,n} P_n |\langle n | \hat{j}_{p,\parallel}(\vec{q}) | m \rangle|^2 \left(\frac{1}{\omega + i\delta - \omega_{mn}} + \frac{1}{\omega_{mn}} - \frac{1}{\omega + i\delta - \omega_{nm}} - \frac{1}{\omega_{nm}} \right) \quad (20.170)$$

$$= \sum_{m,n} P_n |\langle n | \hat{j}_{p,\parallel}(\vec{q}) | m \rangle|^2 \frac{\omega^2}{\omega_{mn}^2} \frac{2\omega_{mn}}{(\omega + i\eta)^2 - \omega_{mn}^2}. \quad (20.171)$$

根据式(20.179)，我们可以得到：

$$= \frac{\omega^2}{q^2} \sum_{m,n} P_n |\langle n | \hat{j}_{p,\parallel}(\vec{q}) | m \rangle|^2 \frac{2\omega_{mn}}{(\omega + i\eta)^2 - \omega_{mn}^2} \quad (20.172)$$

$$= \frac{\omega^2}{q^2} \sum_{m,n} P_n |\langle n | \hat{j}_{p,\parallel}(\vec{q}) | m \rangle|^2 \left(\frac{1}{\omega + i\delta - \omega_{mn}} - \frac{1}{\omega + i\delta - \omega_{nm}} \right) \quad (20.173)$$

$$= \frac{\omega^2}{q^2} \chi_0[\vec{q}, \omega]. \quad (20.174)$$

验证了式(20.165)的结果。

我们还可以利用相同的方法考虑系统对横向电磁场的响应。这是需要计算横向电流响应函数。

关系 1：连续性方程：

$$\frac{\partial \hat{n}}{\partial t} = -\nabla \cdot \hat{j} \quad (20.175)$$

在倒空间，可以写为：

$$\frac{\partial \hat{n}(\vec{q})}{\partial t} = -i\vec{q} \cdot \hat{j} = -iqj_{\parallel} \quad (20.176)$$

利用 Heisenberg 运动方程

$$\frac{\partial \hat{n}(\vec{q})}{\partial t} = i[\hat{H}, \hat{n}(\vec{q})], \quad (20.177)$$

我们得到：

$$[\hat{n}(\vec{q}), \hat{H}] = q j_{\parallel}. \quad (20.178)$$

相应的矩阵元可以写成：

$$(\varepsilon_n - \varepsilon_m) \langle m | \hat{n}(\vec{q}) | n \rangle = q \langle m | j_{\parallel} | n \rangle. \quad (20.179)$$

关系 2：利用方程 (20.178)，我们来计算

$$\langle n | [[\hat{n}(\vec{q}), \hat{H}], \hat{n}^{\dagger}(\vec{q})] | n \rangle = q \langle n | [j_{\parallel}, \hat{n}^{\dagger}(\vec{q})] | n \rangle. \quad (20.180)$$

上式左边可以写为

$$\begin{aligned} \langle [[\hat{n}(\vec{q}), \hat{H}(\vec{q})], \hat{n}^{\dagger}] \rangle &= \sum_n P_n \langle n | [[\hat{n}(\vec{q}), \hat{H}], \hat{n}^{\dagger}(\vec{q})] | n \rangle = \langle n | \hat{n} \hat{H} \hat{n}^{\dagger} - \hat{n}^{\dagger} \hat{n} \hat{H} - \hat{H} \hat{n} \hat{n}^{\dagger} + \hat{n}^{\dagger} \hat{H} \hat{n} | n \rangle \\ &= \sum_{m,n} P_n [(\varepsilon_m - \varepsilon_n) (|\langle n | \hat{n}(\vec{q}) | m \rangle|^2 - |\langle n | \hat{n}^{\dagger}(\vec{q}) | m \rangle|^2)] \end{aligned} \quad (20.181)$$

$$= \sum_{m,n} (P_n + P_m) (\varepsilon_m - \varepsilon_n) |\langle n | \hat{n}(\vec{q}) | m \rangle|^2 \quad (20.182)$$

$$= 2 \sum_{m,n} P_n (\varepsilon_m - \varepsilon_n) |\langle n | \hat{n}(\vec{q}) | m \rangle|^2. \quad (20.183)$$

上式右边可以写为：

$$q \langle n | [j_{\parallel}, \hat{n}^{\dagger}(\vec{q})] | n \rangle = \frac{q}{2m} \sum_i \langle n | [\hat{p}_i e^{-i\vec{q}\cdot\vec{r}_i} + e^{-i\vec{q}\cdot\vec{r}_i} \hat{p}_i, e^{i\vec{q}\cdot\vec{r}_i}] | n \rangle = \frac{nq^2}{m} \quad (20.184)$$

让左右两边相等我们得到：

$$\sum_{m,n} P_n (\varepsilon_m - \varepsilon_n) |\langle n | \hat{n}(\vec{q}) | m \rangle|^2 = \frac{nq^2}{2m}. \quad (20.185)$$

这个关系称为 f-sum rule。利用式 (20.179) 我们得到：

$$\sum_{m,n} P_n \frac{1}{\varepsilon_{mn}} |\langle n | \hat{j}_{\parallel}(\vec{q}) | m \rangle|^2 = \frac{n}{2m}. \quad (20.186)$$

20.6 阅读材料

- J. B. Johnson, Phys. Rev. 32, 97 (1928)
- H. Nyquist, Phys. Rev. 32, 110 (1928).
- L. Onsager, Phys. Rev. 37, 405 (1931); 38, 2265 (1931).
- H. B. Callen and T. A. Welton, Phys. Rev. 83, 34 (1951).
- R. Kubo, J. Phys. Soc. Japan 12, 570 (1957).

- R. Kubo, M. Yokota, and S. Nakajima, J. Phys. Soc. Jpn. 12, 1203-1211 (1957).
- J. M. Luttinger, Phys. Rev. 135, A1505 (1964).

第 21 章 练习 & 作业

21.1 系综理论

1. [20220228] 利用学习到的热力学知识证明以下关系式：

$$C_P = C_V + \frac{TV\alpha_P^2}{\kappa_T},$$

其中

$$\alpha_P = \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_P$$

为等压膨胀系数，

$$\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_T$$

为等温压缩率。

2. [20220307] 考虑 N 个爱因斯坦振子组成的系统，振子的角频率为 ω_0 。类似理想气体，它们之间可以存在微弱的相互作用导致它们可以交换能量。(1) 分别用微正则和正则系综计算其熵和内能。(2) 给出量子谐振子的熵和内能，证明经典极限下回归到 (1) 的结果。

3. [20220307] 证明质量为 m ，角频率为 ω_0 的量子谐振子在热平衡状态下，在坐标表象下的密度矩阵为

$$\rho(q, q') = \left(\frac{1}{2\pi l^2} \tanh \frac{x}{2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{1}{8l^2} [(q+q')^2 \tanh(x/2) + (q-q')^2 \coth(x/2)] \right\}.$$

注： $x = \beta \hbar \omega_0$, $l = \sqrt{\frac{\hbar}{2m\omega_0}}$.

4. [20220309] 请分别验证谐振子基态、第一激发态和热态对应的 Wigner 函数 [式 (21.1-21.1)]:

- 基态的 Wigner 函数为：

$$\rho_0(q, p) = 2 \exp \left\{ -\frac{q^2}{2l^2} - \frac{p^2}{\hbar^2/2l^2} \right\}.$$

- 第一激发态的 Wigner 函数为：

$$\rho_1(q, p) = 2 \left(\frac{q^2}{2l^2} + \frac{p^2}{\hbar^2/2l^2} - \frac{1}{2} \right) \exp \left\{ -\frac{q^2}{2l^2} - \frac{p^2}{\hbar^2/2l^2} \right\}.$$

基态的 Wigner 函数在相空间为高斯函数的形式，因此 $\rho_0(q, p) \geq 0$ ；但 $\rho_1(q=0, p=0) < 0$ 。

- 热平衡状态下的 Wigner 分布函数:

$$\rho(q, p) = 2 \tanh\left(\frac{x}{2}\right) \exp\left\{-\frac{q^2 \tanh(x/2)}{2l^2}\right\} \exp\left\{-\frac{p^2 \tanh(x/2)}{\hbar^2/2l^2}\right\}$$

5. [20220314] 光子气体的是具有线性色散关系, 化学势 $\mu = 0$ (为什么?) 的量子玻色气体。利用我们介绍的量子气体理论, 分析光子气体的热力学性质, 计算其内能、辐射压、熵等热力学量, 并与经典理想气体的结果进行比较。

第 22 章 附录 I：二次量子化

统计物理的主要任务是从描述系统微观自由度的理论出发通过统计平均的方法来构建描述系统宏观(平衡态)热力学性质的理论。宏观系统中往往包含大量的微观粒子($\sim 10^{23}$ 量级)。根据量子力学，这些大量粒子是不可区分的。这时，用二次量子化的语言来描述量子多体系统的状态是非常方便的。这是量子力学的课程内容，这里我们做简单介绍。

二次量子化语言实际上是我们考虑多体系统时“视角”的一个转变。从单粒子系统扩展到多粒子系统的一个最为直接的方法就是从追踪单个粒子的运动扩展到追踪多个粒子的运动。但这样有几个问题。首先，统计物理中要考虑的系统粒子数是非常多的($\sim 10^{23}$)，同时追踪这么多粒子的状态实际上无法实现。其次，对于量子多体系统，相同种类的粒子是不可分辨的。我们无法再去追踪每个粒子。这时，我们可以把视角从追踪粒子转换到统计粒子的分布上。这就是二次量子化语言的关键之处。举例来说，在期末考试时，我如果想了解同学们学得如何，一个办法是去了解每个同学的成绩，这就需要去记录每个同学的分数。另外一个办法是，我来统计每个分数段的人数。这时，每位同学对我来说仅仅是一位同学，同学的姓名、性别等标记信息都不再重要。在学生人数非常多的情况下，如果我仅仅想了解同学们的学习状况，后一种方法就显得方便很多。二次量子化的语言其实就是基于这样的一个认识。

22.1 单粒子量子系统

我们从单粒子量子系统出发。单粒子系统的状态可以用粒子的波函数 $|\psi\rangle$ 来描述。 $|\psi\rangle$ 为单粒子系统希尔伯特空间 \mathcal{H}_1 的一个复向量。量子力学假设 $|\psi\rangle$ 可归一化。而系统波函数可以表示为 $|\psi\rangle$ 在实空间的投影：

$$\psi(\vec{r}) = \langle \vec{r} | \psi \rangle.$$

而 $|\vec{r}\rangle$ 为位置算符的本征态：

$$\hat{\vec{r}}|\vec{r}\rangle = \vec{r}|\vec{r}\rangle$$

由于

$$\langle \vec{r}' | \vec{r} \rangle = \delta(\vec{r} - \vec{r}')$$

不满足归一化条件，因此不属于希尔伯特空间 \mathcal{H}_1 。但它具有完备性条件

$$\int d^3r |\vec{r}\rangle\langle\vec{r}| = 1$$

可以用来表示实空间波函数。

暂时只考虑不含时的情况，如果我们知道描述粒子的哈密顿量

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{r}).$$

理论上我可以求解描述系统状态的定态薛定谔方程

$$\hat{H}|\alpha\rangle = \varepsilon_\alpha|\alpha\rangle$$

来得到系统的本征状态 $|\alpha\rangle$ 和本征能级 ε_α 。我们可以使用这些本征态作为一组正交、完备的基，把 \mathcal{H}_1 中描述单粒子状态的波函数 $|\psi\rangle$ 表示为这组基的线性叠加的形式

$$|\psi\rangle = \sum_{\alpha} c_{\alpha} |\alpha\rangle.$$

基组的正交完备性体现在¹:

$$\langle\beta|\alpha\rangle = \delta_{\alpha,\beta}, \quad \sum_{\alpha} |\alpha\rangle\langle\alpha| = 1.$$

22.2 多粒子量子系统的希尔伯特空间

如果系统中包含多个粒子，我们需要在多粒子系统组成的希尔伯特空间 (\mathcal{H}_N) 来描述系统的状态。下面我们多粒子系统希尔伯特空间可以由单粒子空间直积的方式得到 ($\mathcal{H}_N = \mathcal{H}_1 \otimes \dots$)，对应的基组可以通过单粒子系统基组的直积来构建。

以两粒子系统为例。假如粒子 1 在所处希尔伯特空间 V 下的状态可以用基组 $|v_i\rangle$ 来描述。粒子 2 可以用基组 $|w_j\rangle$ ，对应的希尔伯特空间为 W 。那么在粒子 1 和 2 组成的复合系统对应的希尔伯特空间可以表示为 $V \otimes W$ ，对应的基组可以表示为 $|v_i\rangle \otimes |w_j\rangle$ 。其中 $|v_i\rangle$ 和 $|w_j\rangle$ 分别表示粒子 1 和 2 所处的状态。 $|v_i\rangle \otimes |w_j\rangle$ 为希尔伯特空间 $V \otimes W$ 中的一个向量。也就是说， $V \otimes W$ 可以定义为 V 中的向量 v 和 W 中的向量 w 构成的新向量 $v \otimes w$ 的线性组合所构成的空间。我们要求它满足下面的两个规则：

- $(av) \otimes w = v \otimes (aw) = a(v \otimes w), \quad a \in \text{Complex number}$
- $(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$, 而且 $v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2$

我们有：

¹ 我们用 $\delta(x)$ 表示 Dirac δ -函数，用 $\delta_{x,y}$ 表示 Kronecker δ -函数。

- 向量 $0 \otimes w = v \otimes 0 = 0(v \otimes w)$ 为 $V \otimes W$ 空间中的零向量。

- $V \otimes W$ 空间中的任意向量可以表示为下面线性叠加的形式：

$$|\psi\rangle = \sum_{i,j} c_{i,j} |v_i\rangle \otimes |w_j\rangle.$$

- $V \otimes W$ 的维度为 V 的维度 N_V 和 W 的维度 N_W 的乘积 $N_V N_W$ 。

我们可以看到，如果把两个单粒子态的直积用我们定义基组的线性组合来表示

$$\left(\sum_i c_i |v_i\rangle\right) \otimes \left(\sum_j c_j |w_j\rangle\right) = \sum_{i,j} c_i c_j |v_i\rangle \otimes |w_j\rangle.$$

对应的展开系数 $c_{i,j} = c_i c_j$ 是可以分解的。这样的状态称为直积态 (Product state)。但给定任意线性组合

$$|\psi\rangle = \sum_{i,j} c_{i,j} |v_i\rangle \otimes |w_j\rangle,$$

$|\psi\rangle$ 并不总是直积态。不能写成直积形式的态为纠缠态 (Entangled State)。纠缠态下两个粒子之间存在量子关联。这样得到的两个波函数的内积可以表示为：

$$\langle v_k | \otimes \langle w_l | v_i \rangle \otimes w_j \rangle = \langle v_k | v_i \rangle \langle w_l | w_j \rangle$$

采用类似方法，我们也可以构建直积空间的算符。假如算符 O_1 和 O_2 分别为粒子 1 和 2 对应空间的算符，它们的直积表示为 $O_1 \otimes O_2$ 。可以将个算符作用到对应波函数

$$(O_1 \otimes O_2) |v_i \otimes w_j\rangle = O_1 |v_i\rangle \otimes O_2 |w_j\rangle$$

所以

$$\langle v_k | \otimes \langle w_l | (O_1 \otimes O_2) |v_i \otimes w_j\rangle = \langle v_k | O_1 | v_i \rangle \otimes \langle w_l | O_2 | w_j \rangle$$

将以上介绍推广，我们可以从单粒子基组来构建描述一个 N 粒子系统的基组

$$|\alpha_1 \alpha_2 \cdots \alpha_N\rangle \equiv |\alpha_1\rangle \otimes |\alpha_2\rangle \otimes \cdots \otimes |\alpha_N\rangle. \quad (22.1)$$

它满足如下正交、归一、完备条件：

$$(\alpha'_1 \alpha'_2 \cdots \alpha'_N | \alpha_1 \alpha_2 \cdots \alpha_N) = \delta_{\alpha_1, \alpha'_1} \cdots \delta_{\alpha_N, \alpha'_N}, \quad \sum_{\alpha_1 \cdots \alpha_N} |\alpha_1 \alpha_2 \cdots \alpha_N\rangle (\alpha_1 \alpha_2 \cdots \alpha_N| = \mathbf{1}_N. \quad (22.2)$$

这样， N 粒子希尔伯特空间 \mathcal{H}_N 中的任意向量都可以表示为其线性叠加的形式：

$$|\psi\rangle = \sum_{\alpha_1, \dots, \alpha_N} c_{\alpha_1 \cdots \alpha_N} |\alpha_1\rangle \otimes |\alpha_2\rangle \otimes \cdots \otimes |\alpha_N\rangle.$$

22.3 全同粒子系统的交换对称性

全同粒子的不可区分性以及其自旋统计特征决定了多体波函数的交换对称性。由于粒子的不可区分性，如果我们任意交换两个粒子，系统的状态不会发生变化，即 $\psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$ 和 $\psi(\vec{r}_2, \vec{r}_1, \dots, \vec{r}_N)$ 表示相同的状态，它们之间最多只相差一个相位

$$\psi(\vec{r}_2, \vec{r}_1, \dots, \vec{r}_N) = e^{i\theta} \psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N).$$

如果再次交换 1, 2, 波函数就会和初始时完全相同，所以

$$\zeta^2 = (e^{i\theta})^2 = 1$$

这里 $\zeta = 1$ 对应玻色子， $\zeta = -1$ 对应费米子。

Pauli 的自旋统计定理告诉我们，自旋为半整数的粒子为费米子，而自旋为整数的粒子为玻色子。常见的费米子包括电子、质子、中子、He₃ 等，常见的玻色子包括光子、声子、He₄ 等。

上面在构建多粒子系统的基组时我们并没有考虑到粒子的交换对称性。为了满足交换对称性有两个办法，我们可以约束线性展开系数 $c_{\alpha_1 \dots \alpha_N}$ ，我们也可以进一步对构建的基组进行对称化操作。这可以通过引入置换操作 $\mathcal{P}_{B/F}$ 来实现

$$\mathcal{P}_\zeta |\alpha_1 \dots \alpha_N\rangle = \frac{1}{N!} \sum_P \zeta^P |\alpha_{P1} \dots \alpha_{PN}\rangle. \quad (22.3)$$

- 这里， $\zeta = \pm 1$ 分别对应玻色子和费米子， $(P1 \dots PN)$ 为对粒子状态进行置换的操作，每一个置换操作都可以通过多次的两粒子交换来实现。而这里的求和表示对置换所产生的所有排列形式进行求和。
- ζ^P 为相应排列对应的置换因子。对于费米子 ($\zeta = -1$)，如果 P 由奇数次的两粒子交换来实现，则 $\zeta^P = -1$ ；如果由偶数次两粒子交换实现则 $\zeta^P = 1$ 。对于玻色子， $\zeta = 1$ ，所以 ζ^P 总为 1。

可以验证，通过以上置换操作后得到的波函数满足相应的交换对称性。因此，我们引入新的定义：

$$|\alpha_1 \dots \alpha_N\rangle \equiv \sqrt{N!} \mathcal{P}_{B/F} |\alpha_1 \dots \alpha_N\rangle = \frac{1}{\sqrt{N!}} \sum_P \zeta^P |\alpha_{P1} \dots \alpha_{PN}\rangle. \quad (22.4)$$

我们引入的对称化操作 \mathcal{P}_ζ 实际上是一个投影算符，它满足

$$\mathcal{P}_\zeta^2 = \mathcal{P}_\zeta.$$

它的意义也很明显，对已经通过置换做了对称化的波函数，再次对称化不起任何作用。这一投影操作实际上把我们引入的 N 粒子希尔伯特空间 \mathcal{H}_N 投影到了满足特定对称性的子空间

$$\mathcal{B}_N = \mathcal{P}_B \mathcal{H}_N \mathcal{P}_B, \quad \mathcal{F}_N = \mathcal{P}_F \mathcal{H}_N \mathcal{P}_F.$$

分别对应 N 粒子玻色子和费米子的希尔伯特空间。相应 \mathcal{B}_N 或 \mathcal{F}_N 空间下的完备性关系可以表示为：

$$\sum_{\alpha_1 \cdots \alpha_N} \frac{1}{N!} |\alpha_1 \cdots \alpha_N\rangle \langle \alpha_1 \cdots \alpha_N| = \mathbf{1}_\zeta. \quad (22.5)$$

对称化之后的基矢满足如下归一化关系：

$$\{\alpha_1 \cdots \alpha_N | \alpha_1 \cdots \alpha_N\} = N! \left(\alpha_1 \cdots \alpha_N | \mathcal{P}_{B/F}^2 | \alpha_1 \cdots \alpha_N \right) \quad (22.6)$$

$$= \sum_P \zeta^P \langle \alpha'_1 | \alpha_{P1} \rangle \cdots \langle \alpha'_N | \alpha_{PN} \rangle \quad (22.7)$$

$$= \begin{cases} 1, & \text{Fermion} \\ \prod_\alpha n_\alpha!, & \text{Boson} \end{cases} \quad (22.8)$$

$$= \prod_\alpha n_\alpha!. \quad (22.9)$$

- 费米子：对于费米子来说， N 个粒子必须占据不同的状态，所以利用单粒子态的正交归一性，满足 $\langle \alpha'_1 | \alpha_{P1} \rangle = \cdots = \langle \alpha'_N | \alpha_{PN} \rangle = 1$ 的只有一种情况。于是我们得到结果 1。
- 玻色子：对于玻色子，同一个状态可以同时占据多个粒子。我们用 n_α 来表示处于单粒子态 $|\alpha\rangle$ 的粒子数。这时，对 P 的求和中就有多种情况非零。原因是，处于状态 $|\alpha\rangle$ 的相同状态粒子做内积可以有 $n_\alpha!$ 种组合。

据此，我们可以进一步对波函数进行归一化，定义：

$$|\alpha_1 \cdots \alpha_N\rangle \equiv \frac{1}{\sqrt{\prod_\alpha n_\alpha!}} |\alpha_1 \cdots \alpha_N\rangle = \frac{1}{\sqrt{N! \prod_\alpha n_\alpha!}} \sum_P \zeta^P |\alpha_{P1} \cdots \alpha_{PN}\rangle. \quad (22.10)$$

至此，我们找到了归一化的满足粒子交换对称性的一组基，来表示 N 粒子玻色或费米系统希尔伯特空间中的任意矢量。在 \mathcal{B}_N 或 \mathcal{F}_N 空间下，相应的完备性条件可以表示为：

$$\sum_{\alpha_1 \cdots \alpha_N} \frac{\prod_\alpha n_\alpha!}{N!} |\alpha_1 \cdots \alpha_N\rangle \langle \alpha_1 \cdots \alpha_N| = \mathbf{1}_{B/F}. \quad (22.11)$$

22.3.1 费米子的多体波函数

N 个费米子的反对称波函数 $\langle \vec{r}_1 \cdots \vec{r}_N | \alpha_1 \cdots \alpha_N \rangle = \frac{1}{\sqrt{N!}} \sum_P (-1)^P \psi_{P\alpha_1}(\vec{r}_1) \cdots \psi_{P\alpha_N}(\vec{r}_N)$ 可以用 Slater 行列式表示。对于两粒子系统：

$$\frac{1}{\sqrt{2!}} \begin{vmatrix} \psi_{\alpha_1}(\vec{r}_1) & \psi_{\alpha_1}(\vec{r}_2) \\ \psi_{\alpha_2}(\vec{r}_1) & \psi_{\alpha_2}(\vec{r}_2) \end{vmatrix} = \frac{1}{\sqrt{2}} (\psi_{\alpha_1}(\vec{r}_1)\psi_{\alpha_2}(\vec{r}_2) - \psi_{\alpha_2}(\vec{r}_1)\psi_{\alpha_1}(\vec{r}_2)) \quad (22.12)$$

同理，对于 N 粒子系统

$$\frac{1}{\sqrt{N!}} \begin{vmatrix} \Phi_{k_1}(x_1) & \Phi_{k_1}(x_2) & \cdots & \Phi_{k_1}(x_n) \\ \Phi_{k_2}(x_1) & \Phi_{k_2}(x_2) & \cdots & \Phi_{k_2}(x_n) \\ \vdots & \vdots & & \vdots \\ \Phi_{k_n}(x_1) & \Phi_{k_n}(x_2) & \cdots & \Phi_{k_n}(x_n) \end{vmatrix} \quad (22.13)$$

交换反对称性是行列式的基本性质。任意交换两行，行列式变号。若两行完全相同，则行列式为零（Pauli 不相容原理）。

22.4 产生/湮灭算符

22.4.1 产生算符

我们定义产生算符

$$a_\lambda^\dagger |\lambda_1 \lambda_2 \cdots \lambda_N\rangle \equiv |\lambda \lambda_1 \lambda_2 \cdots \lambda_N\rangle \quad (22.14a)$$

其作用是在状态 $|\lambda\rangle$ 产生一个新的粒子，系统中的粒子数增加 1。这里需要注意，上式右边新产生在粒子排在第一位。这个位置是一个约定。根据这个定义，我们很容易验证：

$$a_\lambda^\dagger |\lambda_1 \lambda_2 \cdots \lambda_N\rangle = \sqrt{n_\lambda + 1} |\lambda \lambda_1 \lambda_2 \cdots \lambda_N\rangle. \quad (22.14b)$$

讨论：

- 这里 n_λ 为产生算符作用前系统中处于状态 λ 的粒子数。对于费米子，如果 $n_\lambda = 1$ ，那么新生成的波函数中 λ 态同时存在两个粒子，对应的波函数为零，所以前面的系数不再重要。
- 我们可以定义一个真空态： $|0\rangle = |0\rangle$ ，表示系统中没有任何粒子。利用产生算符，我们

可以依次生成所有的状态：

$$a_\lambda^\dagger |0\rangle = |\lambda\rangle = |\lambda\rangle,$$

$$a_{\lambda_1}^\dagger a_{\lambda_2}^\dagger \cdots a_{\lambda_N}^\dagger |0\rangle = |\lambda_1 \cdots \lambda_N\rangle = \sqrt{\prod_\lambda n_\lambda!} |\lambda_1 \cdots \lambda_N\rangle.$$

- 粒子数为 N 的系统所处的状态处于相应的希尔伯特空间 \mathcal{H}_N 。由所有可能粒子数对应的希尔伯特空间直和构成的空间称为 Fock 空间： $\mathcal{F} = \mathcal{H}_0 \oplus \mathcal{H}_1 \oplus \cdots = \bigoplus_{n=0}^{\infty} \mathcal{H}_n$ 。对于 \mathcal{F}_N 和 \mathcal{B}_N 可以做相应的 Fock 空间的定义： $\mathcal{F}_{\mathcal{F}} = \bigoplus_{n=0}^{\infty} \mathcal{F}_n$, $\mathcal{F}_{\mathcal{B}} = \bigoplus_{n=0}^{\infty} \mathcal{B}_n$ 。

22.4.2 湮灭算符

湮灭算符可以定义为产生算符的伴算符

$$a_\lambda \equiv (a_\lambda^\dagger)^\dagger. \quad (22.15)$$

从这个定义出发，我们有：

$$\{\lambda_1 \cdots \lambda_N | a_\lambda = \{\lambda \lambda_1 \cdots \lambda_N |.$$

据此，

$$\begin{aligned} a_\lambda |\lambda_1 \cdots \lambda_N\rangle &= \sum_{M=0}^{\infty} \frac{1}{M!} \sum_{\alpha_1 \cdots \alpha_M} \{\alpha_1 \cdots \alpha_M | a_\lambda | \lambda_1 \cdots \lambda_N\} |\alpha_1 \cdots \alpha_M\rangle \\ &= \sum_{M=0}^{\infty} \frac{1}{M!} \sum_{\alpha_1 \cdots \alpha_M} \{\lambda \alpha_1 \cdots \alpha_M | \lambda_1 \cdots \lambda_N\} |\alpha_1 \cdots \alpha_M\rangle \end{aligned}$$

这里，只有 $M+1=N$ ，而且 $\lambda \alpha_1 \cdots \alpha_M$ 可以通过置换与 $\lambda_1 \cdots \lambda_N$ 一一对应时，内积才非零。如果 λ 与第 i 个粒子的状态相同，我们可以先把 λ_i 置换为第一个粒子的状态，同时产生一个置换因子 ζ_{i-1}

$$= \sum_i \frac{1}{M!} \sum_{\alpha_1 \cdots \alpha_M} \zeta^{i-1} \delta_{\lambda_i, \lambda} \{\lambda \alpha_1 \cdots \alpha_M | \lambda_i \lambda_1 \cdots \lambda_{i-1} \lambda_{i+1} \cdots \lambda_N\} |\alpha_1 \cdots \alpha_M\rangle$$

- 先考虑费米子。要使上式中两个多体波函数的内积非零，需要 $\alpha_1 \cdots \alpha_M$ 这 M 个不相同的状态为 $\lambda_1 \cdots \lambda_{i-1} \lambda_{i+1} \cdots \lambda_N$ 的特定排列，对应的置换因子为 ζ^P 。于是：

$$\begin{aligned} &= \sum_i \frac{1}{M!} \sum_{\alpha_1 \cdots \alpha_M} \zeta^{i-1} \delta_{\lambda, \lambda_i} \zeta^P |\alpha_1 \cdots \alpha_M\rangle \\ &= \sum_i \frac{1}{M!} \sum_{\alpha_1 \cdots \alpha_M} \zeta^{i-1} \delta_{\lambda, \lambda_i} |\lambda_1 \cdots \lambda_{i-1} \lambda_{i+1} \cdots \lambda_N\rangle \end{aligned}$$

这里 $\sum_{\alpha_1 \dots \alpha_M}$ 中共有 $M!$ 种排列方式，所以最后结果为：

$$= \sum_i \zeta^{i-1} \delta_{\lambda, \lambda_i} |\lambda_1 \dots \lambda_{i-1} \lambda_{i+1} \dots \lambda_N \rangle.$$

- 再考虑玻色子。对于玻色子，不需要考虑置换带来的因子。

$$= \sum_i \frac{1}{M!} \sum_{\alpha_1 \dots \alpha_M} \prod_{\alpha} n_{\alpha}! \delta_{\lambda, \lambda_i} |\lambda_1 \dots \lambda_{i-1} \lambda_{i+1} \dots \lambda_N \rangle$$

这时，不同粒子所处状态可以相同。上式第二个求和共有 $M!/\prod_{\alpha} n_{\alpha}!$ 项。这个因子正好与上式的阶乘因子消去，得到：

$$= \sum_i \delta_{\lambda, \lambda_i} |\lambda_1 \dots \lambda_{i-1} \lambda_{i+1} \dots \lambda_N \rangle$$

综合以上考虑，我们得到一个普适形式：

$a_{\lambda} |\lambda_1 \dots \lambda_N \rangle = \sum_{i=1}^N \zeta^{i-1} \delta_{\lambda, \lambda_i} |\lambda_1 \dots \lambda_{i-1} \lambda_{i+1} \dots \lambda_N \rangle = \sum_{i=1}^N \zeta^{i-1} \delta_{\lambda, \lambda_i} |\lambda_1 \dots \hat{\lambda}_i \dots \lambda_N \rangle. \quad (22.16a)$

或者

$a_{\lambda} |\lambda_1 \dots \lambda_N \rangle = \frac{1}{\sqrt{n_{\lambda}}} \sum_{i=1}^N \zeta^{i-1} \delta_{\lambda, \lambda_i} |\lambda_1 \dots \hat{\lambda}_i \dots \lambda_N \rangle. \quad (22.16b)$

这里的 n_{λ} 为 $|\lambda_1 \dots \lambda_N \rangle$ 中 λ 态的占据数，而 $\hat{\lambda}_i$ 表示不包含状态 λ_i 。若 $n_{\lambda} = 0$ ，则上式右边为零。所以，湮灭算符 a_{λ} 的作用是湮灭占据单粒子态 λ 的一个粒子，使系统的状态数减 1。

22.4.3 算符对易关系

根据产生和湮灭算符的定义以及 (22.14a-22.16b)，我们可以证明如下对易关系：

$$[a_{\mu}^{\dagger}, a_{\lambda}^{\dagger}]_{-\zeta} \equiv a_{\mu}^{\dagger} a_{\lambda}^{\dagger} - \zeta a_{\lambda}^{\dagger} a_{\mu}^{\dagger} = 0. \quad (22.17)$$

$$[a_{\mu}, a_{\lambda}]_{-\zeta} \equiv a_{\mu} a_{\lambda} - \zeta a_{\lambda} a_{\mu} = 0. \quad (22.18)$$

$$[a_{\mu}, a_{\lambda}^{\dagger}]_{-\zeta} \equiv a_{\mu} a_{\lambda}^{\dagger} - \zeta a_{\lambda}^{\dagger} a_{\mu} = \delta_{\mu, \lambda}. \quad (22.19)$$

请自行证明以上对易关系。

根据以上，我们还可以定义粒子数算符

$$\hat{n}_i = a_i^{\dagger} a_i. \quad (22.20)$$

我们可以验证

$$\hat{n}_i |\lambda_1 \cdots \lambda_i \cdots \rangle = n_i |\lambda_1 \cdots \lambda_i \cdots \rangle \quad (22.21a)$$

$$\hat{n}_i |\lambda_1 \cdots \lambda_i \cdots \} = n_i |\lambda_1 \cdots \lambda_i \cdots \}. \quad (22.21b)$$

22.5 占据数表象

利用全同粒子的特征，我们没有必要去追踪每个粒子的状态，只需要统计每个单粒子态所占据的粒子数。因此我们可以用下面的形式来表示一个 N 粒子系统的波函数：

$$|n_1 n_2 \cdots \rangle \equiv |\{n_i\}\rangle,$$

或者可以进一步标明系统中的总粒子数：

$$|n_1 n_2 \cdots ; N\rangle \equiv |\{n_i\}; N\rangle, \quad (22.22)$$

其中 n_i 表示处于 i 状态的粒子数， i 遍历所有的单粒子态，而 $\sum_i n_i = N$ 为系统的总粒子数。要指定一个状态，我们需要给出所有单粒子态的占据数。为了方便，我们还可以约定这些状态的顺序。比如，我们按照它们的能量来进行排序 $\varepsilon_1 \leq \varepsilon_2 \leq \cdots$ 。这样定义的任何状态都可以通过使用产生算符依次作用到真空态得到。如果所有粒子都处于相同的状态则

$$|n_i\rangle \equiv |0 \cdots n_i \cdots 0 \cdots ; n_i\rangle = \frac{1}{\sqrt{n_i!}} (a_i^\dagger)^{n_i} |0\rangle. \quad (22.23)$$

同理：

$$|n_1 n_2 \cdots ; N\rangle = \frac{1}{\sqrt{n_1! n_2! \cdots}} (a_1^\dagger)^{n_1} (a_2^\dagger)^{n_2} \cdots |0\rangle. \quad (22.24)$$

或者写成：

$$|\{n_i\}; N\rangle = \frac{1}{\sqrt{\prod_i n_i!}} \prod_i (a_i^\dagger)^{n_i} |0\rangle.$$

产生、湮灭算符作用到这个波函数导致增加或减少一个粒子：

$$\begin{aligned} a_i^\dagger |n_1 n_2 \cdots n_i \cdots \rangle &= \zeta^{\sum_{l=1}^{i-1} n_l} \sqrt{n_i + 1} |n_1 n_2 \cdots (n_i + 1) \cdots \rangle. \\ a_i |n_1 n_2 \cdots n_i \cdots \rangle &= \zeta^{\sum_{l=1}^{i-1} n_l} \sqrt{n_i} |n_1 n_2 \cdots (n_i - 1) \cdots \rangle. \\ a_i^\dagger a_i |n_1 n_2 \cdots n_i \cdots \rangle &= n_i |n_1 n_2 \cdots n_i \cdots \rangle. \end{aligned} \quad (22.25)$$

占据数表象下，波函数的正交、完备性条件表示为：

$$\langle n_1 n_2 \cdots; N | n'_1 n'_2 \cdots; N \rangle = \delta_{n_1, n'_1} \delta_{n_2, n'_2} \cdots, \quad (22.26)$$

$$\sum_{\{n_i\}} |n_1 n_2 \cdots; N\rangle \langle n_1 n_2 \cdots; N| = \mathbb{1}_{\mathcal{B}_N/\mathcal{F}_N}. \quad (22.27)$$

以上是我们的主要结果。我们看到，通过使用粒子数表象，我们可以方便地使用产生、湮灭算符来生成 Fock 空间的任意状态。

举例：我们可以把以上结果用到单个谐振子系统。这时只有一个单粒子态，但其占据数可以为任意非负整数：

$$a^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle,$$

$$a |n\rangle = \sqrt{n} |n-1\rangle.$$

其中， $|n\rangle$ 表示谐振子的振动量子数为 n 。

22.5.1 有关完备性关系

对于我们遇到的完备性关系，这里需要做一个说明。式 (22.2) 是在 \mathcal{H}_N 空间的完备性，而 (22.5) 和 (22.11) 是在做了对称化之后的 \mathcal{B}_N 或 \mathcal{F}_N 空间。 \mathcal{B}_N 和 \mathcal{F}_N 为 \mathcal{H}_N 的子空间，所以相应的完备性条件前面存在额外因子。这个因子正是粒子的不可分辨性导致的组合数。而在 (22.27) 中，我们做了如下替换：

$$\sum_{\lambda_1 \lambda_2 \cdots} \frac{\prod_i n_i!}{N!} \rightarrow \sum_{n_1 n_2 \cdots} \quad (22.28)$$

替换的原因是每给定一组占据数 $\{n_i\}$ 对应有 $N!/\prod_i n_i!$ 组 $\{\lambda_1 \lambda_2 \dots \lambda_N\}$ 与之对应。这是由于粒子的不可分辨性所导致的结果。下面的简单例子可以帮助我们来理解这一点。

举例：考虑两个粒子 1, 2 和两个状态 ϕ_1 和 ϕ_2 。首先，如果不考虑粒子的不可分辨性，我们可以构建下面的一组基矢来描述这个两体系统：

$$|\phi_1(r_1), \phi_2(r_2)\rangle, |\phi_2(r_1), \phi_1(r_2)\rangle, |\phi_1(r_1), \phi_1(r_2)\rangle, |\phi_2(r_1), \phi_2(r_2)\rangle$$

考慮玻色粒子及其不可分辨性，我們根据以上內容可以定义：

$$|\phi_1(r_1), \phi_2(r_2)\rangle = \frac{1}{\sqrt{2!}} [|\phi_1(r_1), \phi_2(r_2)\rangle + |\phi_2(r_1), \phi_1(r_2)\rangle],$$

$$|\phi_2(r_1), \phi_1(r_2)\rangle = \frac{1}{\sqrt{2!}} [|\phi_2(r_1), \phi_1(r_2)\rangle + |\phi_1(r_1), \phi_2(r_2)\rangle],$$

$$|\phi_1(r_1), \phi_1(r_2)\rangle = |\phi_1(r_1), \phi_1(r_2)\rangle,$$

$$|\phi_2(r_1), \phi_2(r_2)\rangle = |\phi_2(r_1), \phi_2(r_2)\rangle$$

可以看到，做了对称化之后，第一个和第二个波函数实际上是相同的，同时对应占据数表象 $|11\rangle$ 这个状态。由于在对粒子状态的求和中会考虑两次，所以需要 (22.28) 中左边的组合因子来消除这样的重复。如果我们针对不同状态的占据数进行求和，就不会出现重复，自然也不需要这个组合因子。

22.6 算符的二次量子化

以上我们只是讨论了波函数，下面我们考虑算符的二次量子化表示。我们所关心的观测量由对应的厄密算符来表示。比如，在一次量子化表象，状态 Ψ 下，算符 \hat{O} 的期待值为

$$\langle \hat{O} \rangle = \langle \Psi | \hat{O} | \Psi \rangle,$$

而给定一组基后，我们可以在这组基下将算符表示为：

$$\hat{O} = \sum_{i,j} \langle i | \hat{O} | j \rangle |i\rangle \langle j| = \sum_{i,j} O_{ij} |i\rangle \langle j|.$$

我们下面来讨论如何用产生湮灭算符来表示。

22.6.1 单体算符

单体算符具有如下特点：

$$\hat{U}|\alpha_1 \cdots \alpha_N\rangle = \sum_{i=1}^N \hat{U}_i |\alpha_1 \cdots \alpha_N\rangle.$$

这里， \hat{U}_i 只作用到第 i 个粒子。比如，动能算符 \hat{T} 在动量表象下 \vec{p} 可以表示为：

$$\hat{T}|\vec{p}_1 \cdots \vec{p}_N\rangle = \sum_{i=1}^N \frac{\hat{p}_i^2}{2m} |\vec{p}_1 \cdots \vec{p}_N\rangle.$$

这里 \hat{p}_i^2 只作用到 $|\vec{p}_i\rangle$ 上。N 粒子系统单体算符的任意矩阵元可以表示为：

$$(\alpha_1 \cdots \alpha_N | \hat{U} | \beta_1 \cdots \beta_N) = \sum_{i=1}^N (\alpha_1 \cdots \alpha_N | \hat{U}_i | \beta_1 \cdots \beta_N) = \sum_{i=1}^N \langle \alpha_1 | \beta_1 \rangle \cdots \langle \alpha_i | \hat{U}_i | \beta_i \rangle \cdots \langle \alpha_N | \beta_N \rangle$$

我们看到，只要知道了单粒子系统的矩阵元，我们就可以很方便的得到 N 粒子系统的结果。

为了把单体算符用产生、湮灭算符来表示。我们首先选择一组基矢，使得单体算符 \hat{U} 可以被对角化：

$$\hat{U}_i |\alpha\rangle = U_\alpha |\alpha\rangle,$$

$$U_\alpha = \langle \alpha | U_i | \alpha \rangle.$$

在这样的一组基下，我们考虑 N 粒子系统的任意矩阵元：

$$\begin{aligned} \{\alpha'_1 \cdots \alpha'_N | U | \alpha_1 \cdots \alpha_N\} &= \sum_i \frac{1}{N!} \sum_{P,Q} \zeta^P \zeta^Q (\alpha'_{Q1} \cdots \alpha'_{QN} | \hat{U}_i | \alpha_{P1} \cdots \alpha_{PN}) \\ &= \sum_i \sum_P \zeta^P (\alpha'_1 \cdots \alpha'_N | \hat{U}_i | \alpha_{P1} \cdots \alpha_{PN}) \\ &= \sum_i \sum_P \zeta^P \langle \alpha'_{P1} | \alpha_1 \rangle \cdots \langle \alpha'_{Pi} | \hat{U}_i | \alpha_i \rangle \cdots \langle \alpha'_{PN} | \alpha_N \rangle \\ &= \sum_i U_{\alpha_i} \{\alpha'_1 \cdots \alpha'_N | \alpha_1 \cdots \alpha_N\} \end{aligned} \quad (22.29)$$

$$= \sum_\alpha U_\alpha n_\alpha \{\alpha'_1 \cdots \alpha'_N | \alpha_1 \cdots \alpha_N\} \quad (22.30)$$

$$= \{\alpha'_1 \cdots \alpha'_N | \sum_\alpha U_\alpha a_\alpha^\dagger a_\alpha | \alpha_1 \cdots \alpha_N\}$$

这里 (22.29) 中 $\sum_i U_{\alpha_i}$ 要对所有粒子进行求和， $|\alpha_i\rangle$ 为第 i 个粒子所处的状态。从 (22.29) 到 (22.30) 我们把对粒子的求和换成了对粒子所在状态的求和 \sum_α 。这里的 α 遍历所有单粒子态。在这个求和中，我们需要考虑每个单粒子态的占据数 $a_\alpha^\dagger a_\alpha$ 。由于我们考虑的是任意矩阵元，所以我们可以得到在对角表象下单粒子算符的二次量子化表达式：

$$\hat{U} = \sum_\alpha U_\alpha a_\alpha^\dagger a_\alpha.$$

(22.31)

如果要考虑任意表象下的情况，我们需要使用产生、湮灭算符在两个不同表象下的变换关系（作业：请证明这两个关系！）：

$$a_\alpha^\dagger = \sum_\beta \langle \beta | \alpha \rangle a_\beta^\dagger, \quad (22.32)$$

$$a_\alpha = \sum_\beta \langle \alpha | \beta \rangle a_\beta. \quad (22.33)$$

得到：

$$\hat{U} = \sum_{\beta_i, \beta_j} U_{\beta_i \beta_j} a_{\beta_i}^\dagger a_{\beta_j}. \quad (22.34)$$

22.6.2 两体算符

两体算符可以表示为：

$$\hat{V} = \sum_{1 \leq i < j \leq N} \hat{V}_{ij} = \frac{1}{2} \sum_{i \neq j} \hat{V}_{ij}.$$

我们仍然考虑其矩阵元：

$$(\alpha'_1 \cdots \alpha'_N | \frac{1}{2} \sum_{i \neq j} \hat{V}_{ij} | \alpha_1 \cdots \alpha_N) = \frac{1}{2} \sum_{ij} (\alpha'_i \alpha'_j | \hat{V}_{ij} | \alpha_i \alpha_j) \prod_{k \neq i, k \neq j} \langle \alpha'_k | \alpha_k \rangle$$

与单体算符类似，我们只需要知道两体系统的矩阵元就可以得到 N 粒子系统的矩阵元。

要得到其二次量子化表示，我们仍然从对角化表象出发：

$$V_{ij} | \alpha_i \alpha_j \rangle = V_{\alpha_i \alpha_j} | \alpha_i \alpha_j \rangle,$$

$$V_{\alpha_i \alpha_j} = (\alpha_i \alpha_j | V_{ij} | \alpha_i \alpha_j).$$

N 粒子系统的两体矩阵元可以表示为：

$$\begin{aligned} \{\alpha'_1 \cdots \alpha'_N | \frac{1}{2} \sum_{i \neq j} \hat{V}_{ij} | \alpha_1 \cdots \alpha_N\} &= \frac{1}{2} \sum_{i \neq j} \sum_P \zeta^P \prod_{k \neq i, k \neq j} \langle \alpha'_i | \alpha_j \rangle \langle \alpha'_i \alpha'_j | \hat{V}_{ij} | \alpha_i \alpha_j \rangle \\ &= \frac{1}{2} \sum_{i \neq j} V_{\alpha_i \alpha_j} \sum_P \zeta^P \langle \alpha'_1 | \alpha_1 \rangle \cdots \langle \alpha'_N | \alpha_N \rangle \\ &= \{\alpha'_1 \cdots \alpha'_N | \frac{1}{2} \sum_{\alpha \beta} V_{\alpha \beta} a_\alpha^\dagger a_\beta^\dagger a_\beta a_\alpha | \alpha_1 \cdots \alpha_N\}. \end{aligned}$$

与单体算符类似，这里我们也做了从粒子对 (i, j) 求和到双粒子状态 $(\alpha \beta)$ 求和的转换。于是，

$$\hat{V} = \frac{1}{2} \sum_{\alpha \beta} (\alpha \beta | V | \alpha \beta) a_\alpha^\dagger a_\beta^\dagger a_\beta a_\alpha. \quad (22.35)$$

在非对角表象下为

$$\hat{V} = \frac{1}{2} \sum_{\alpha \beta \gamma \delta} (\alpha \beta | V | \gamma \delta) a_\alpha^\dagger a_\beta^\dagger a_\delta a_\gamma. \quad (22.36)$$

这里需要注意，矩阵元中 α, β 的顺序和湮灭算符的顺序相反。

22.7 举例

例 1：自由电子气哈密顿量的二次量子化

$$\begin{aligned}\hat{H} &= \hat{T} = \sum_i -\frac{\hbar^2}{2m} \nabla_i^2 \\ \langle \vec{r} | \vec{k} \rangle &= \frac{1}{\sqrt{V}} e^{i\vec{k} \cdot \vec{r}} \\ \hat{T} &= \sum_{k,k'} \langle k' | -\frac{\hbar^2}{2m} \nabla^2 | k \rangle a_{k'}^\dagger a_k = \sum_k \frac{\hbar^2 k^2}{2m} a_k^\dagger a_k\end{aligned}\quad (22.37)$$

例 2：场算符为坐标表象下的产生、湮灭算符。利用 (22.32-22.33) 我们可以得到：

$$a^\dagger(x) \equiv \psi^\dagger(r) = \sum_\lambda \langle \lambda | r \rangle a_\lambda^\dagger, \quad (22.38)$$

$$a(x) \equiv \psi(r) = \sum_\lambda \langle r | \lambda \rangle a_\lambda. \quad (22.39)$$

场算符满足的对易关系：

$$[\psi(r), \psi^\dagger(r')]_{-\zeta} = \delta(r - r'), \quad [\psi^\dagger(r), \psi^\dagger(r')]_{-\zeta} = 0, \quad [\psi(r), \psi(r')]_{-\zeta} = 0. \quad (22.40)$$

如果 $|\lambda\rangle$ 为动量算符的本征态，则：

$$\begin{aligned}\psi^\dagger(r) &= \sum_k \psi_k^*(r) a_k^\dagger = \frac{1}{\sqrt{V}} \sum_k e^{-ikr} a_k^\dagger \\ \psi(r) &= \frac{1}{\sqrt{V}} \sum_k e^{ikr} a_k\end{aligned}\quad (22.41)$$

粒子动能在动量本征态下的二次量子化表示为：

$$\hat{T} = \sum_k \frac{\hbar^2 k^2}{2m} a_k^\dagger a_k. \quad (22.42)$$

相应的场算符下的表示则为：

$$\hat{T} = - \int dr \psi^\dagger(r) \frac{\hbar^2}{2m} \nabla^2 \psi(r).$$

例 3：粒子数密度算符

$$\hat{n}(r) = \sum_{i=1}^N \delta(r - \hat{r}_i)$$

证明动量空间它的二次量子化表示为：

$$\hat{n}(r) = L^{-d} \sum_{k,q} e^{iqr} a_k^\dagger a_{k+q}.$$

这里， L^d 为系统的体积。场算符表示为：

$$\hat{n}(r) = \psi^\dagger(r) \psi(r).$$

例 4：单体势的二次量子化表示。势能

$$V_{ext} = \sum_i V_i(r_i)$$

对应的二次量子化表示为：

$$V_{ext} = L^{-d} \sum_{k,q} V(q) a_{k+q}^\dagger a_k.$$

这里 $V(q)$ 为 $V(r)$ 的傅里叶变换：

$$V(q) = \int dr e^{-iqr} V(r).$$

对应的逆变换为：

$$V(r) = L^{-d} \sum_q e^{iqr} V(q).$$

例 5：库仑相互作用的二次量子化表示。

$$\hat{V}_{int} = \frac{1}{2} \sum_{i \neq j} \hat{V}(i, j), \quad \hat{V}(i, j) = \frac{e^2}{|\vec{r}_i - \vec{r}_j|}.$$

其动量空间的二次量子化表示为：

$$\hat{V}_{int} = \frac{1}{2} L^{-d} \sum_{k, k', q} V_q a_{k-q}^\dagger a_{k'+q}^\dagger a_{k'} a_k.$$

其中， V_q 为库仑势的傅里叶变换。

22.8 总结

- N 个全同粒子组成的状态可以在占据数表象下表示为：

$$|n_1 n_2 \cdots; N\rangle \equiv |\{n_i\}; N\rangle.$$

这些状态可以用产生算符从真空态来生成：

$$|\{n_i\}; N\rangle = \frac{1}{\sqrt{\prod_i n_i!}} \prod_i (a_i^\dagger)^{n_i} |0\rangle.$$

- 产生、湮灭算符作用于这些状态：

$$a_i^\dagger |n_1 n_2 \cdots n_i \cdots\rangle = \zeta^{\sum_{l=1}^{i-1} n_l} \sqrt{n_i + 1} |n_1 n_2 \cdots (n_i + 1) \cdots\rangle.$$

$$a_i |n_1 n_2 \cdots n_i \cdots\rangle = \zeta^{\sum_{l=1}^{i-1} n_l} \sqrt{n_i} |n_1 n_2 \cdots (n_i - 1) \cdots\rangle.$$

$$a_i^\dagger a_i |n_1 n_2 \cdots n_i \cdots\rangle = n_i |n_1 n_2 \cdots n_i \cdots\rangle.$$

- 正交归一完备性关系:

$$\langle n_1 n_2 \cdots ; N | n'_1 n'_2 \cdots ; N \rangle = \delta_{n_1, n'_1} \delta_{n_2, n'_2} \cdots,$$

$$\sum_{n_1 n_2 \cdots} |n_1 n_2 \cdots ; N\rangle \langle n_1 n_2 \cdots ; N| = 1_{\mathcal{B}_N / \mathcal{F}_N}.$$

- 产生、湮灭算符的对易关系:

$$[a_\mu^\dagger, a_\lambda^\dagger]_{-\zeta} = 0.$$

$$[a_\mu, a_\lambda]_{-\zeta} = 0.$$

$$[a_\mu, a_\lambda^\dagger]_{-\zeta} = \delta_{\mu, \lambda}.$$

- 一个单体算符在 N 粒子系统下可以用产生、湮灭算符表示为:

$$\hat{U} = \sum_{\beta_i, \beta_j} U_{\beta_i \beta_j} a_{\beta_i}^\dagger a_{\beta_j}.$$

- 两体算符可以表示为:

$$\hat{V} = \frac{1}{2} \sum_{\alpha \beta \gamma \delta} (\alpha \beta | V | \gamma \delta) a_\alpha^\dagger a_\beta^\dagger a_\delta a_\gamma.$$

第 23 章 附录 II: 概率论基础

在这个附录中我们简答介绍一下概率论的基本知识和几种常用的概率分布。

随机变量

我们用大写的 X 表示一个随机变量。随机变量的可能取值构成一个空间。如果可取值是离散的，那么它就是离散随机变量；而如果是连续的，就是连续随机变量。我们用 x_i 来表示 X 的一个可能取值。要确定一个随机变量，需要同时给出其取值空间，以及相应的概率分布。

概率密度分布函数

概率密度分布 $P_X(x)$ 表示随机变量取值在 $[x, x + dx]$ 范围内的概率。它满足：

- $P_X(x) \geq 0$,
- $\int dx P_X(x) = 1$.

这里的积分范围为随机变量的取值空间。我们可以把它不能取值的空间范围概率密度定位 0，那么积分就可以扩展到整个空间。对于实变量，就是整个实数空间。如果 X 为离散随机变量，它取值 x_i 的概率为 $p(x_i)$ ，那么 $P_X(x)dx = \sum_i p(x_i)\delta(x - x_i)$ 。

概率分布函数

概率分布函数由概率密度分布函数来定义

$$F_X(x) = \int_{-\infty}^x P_X(x')dx'. \quad (23.1)$$

从其定义我们可以得到， $F_X(-\infty) = 0$, $F_X(+\infty) = 1$ 。

随机变量的矩 (moments)

给定概率密度分布后，我们可以去求解随机变量 X 的矩。 n 级矩的定义如下：

$$\langle x^n \rangle = \int_{-\infty}^{+\infty} x^n P_X(x)dx. \quad (23.2)$$

显然，一级矩对应随机变量的平均值或者叫期待值。

特征函数 (characteristic function)

与随机变量概率密度分布函数对应的特征函数定义为

$$\Phi_X(k) = \langle e^{ikx} \rangle = \int_{-\infty}^{+\infty} e^{ikx} P_X(x) dx = \sum_{n=0}^{+\infty} \frac{(ik)^n}{n!} \langle x^n \rangle. \quad (23.3)$$

根据定义，我们可以认为它是概率密度分布函数的傅立叶变换。

如果知道了特征函数，我们就知道了它的任意阶矩：

$$\langle x^n \rangle = \left. \frac{\partial^n \Phi_X(k)}{\partial (ik)^n} \right|_{k=0}. \quad (23.4)$$

值得一提的是，并不是所有的分布都可以用分布函数的形式写出来，但其对应的特征函数往往更容易得到。因此，很多情况下，特征函数比密度分布函数更实用。

累积量 (cumulants) 及其生成函数

累计量的生成函数定义为：

$$\Psi_X(k) = \ln \Phi_X(k) = \sum_{n=1}^{+\infty} \frac{(ik)^n}{n!} \langle x^n \rangle_c. \quad (23.5)$$

与特征函数类似，我们根据累积量产生函数可以生成任意阶的累积量：

$$\langle x^n \rangle_c = \left. \frac{\partial^n \Psi_X(k)}{\partial (ik)^n} \right|_{k=0}. \quad (23.6)$$

可以证明，零阶累积量为零；一阶累积量对应一阶矩；二阶累积量对应随机变量的方差；三阶累积量表示随机变量的偏度 (skewness)。顾名思义，它表征分布的非对称程度。偏度越大，概率密度分布越不对称。实际上， n 阶的累积量可以用不大于 n 的矩来表示出来。比如：

$$\begin{aligned} \langle x \rangle_c &= \langle x \rangle \\ \langle x^2 \rangle_c &= \langle \Delta x^2 \rangle = \langle (x - \langle x \rangle)^2 \rangle \\ \langle x^3 \rangle_c &= \langle x^3 \rangle - 3\langle x \rangle \langle x^2 \rangle + 2\langle x \rangle^3 \end{aligned}$$

用累积量来表示 n 阶矩

上面我们用直接计算的方法来求得了 1 至 3 阶累积量用 $n \leq 3$ 的矩来表示的结果。反过来，我们也可以用累积量来表示 n 阶矩。它们之间有一个简单实用的图形规则。

根据特征函数的定义我们有：

$$\Phi(k) = \sum_n \frac{(ik)^n}{n!} \langle x^n \rangle \quad (23.7)$$

根据累积量产生函数的定义我们也可以把特征函数写为:

$$\Phi(k) = e^{\Psi(k)} \quad (23.8)$$

$$= \sum_{q=0}^{\infty} \frac{1}{q!} \left(\sum_{m=0}^{\infty} \frac{(ik)^m}{m!} \langle x^m \rangle_c \right)^q \quad (23.9)$$

$$= \sum_n \frac{(ik)^n}{n!} \langle x^n \rangle. \quad (23.10)$$

根据上面的结果, 如果我们可以把 n 阶矩用不大于 n 阶的累积量表示出来. 我们只需要把上式第二行的展开式中所有正比于 $(ik)^n$ 的项提取出来就可以了, 但它可能来自不同类型的组合. 下面我们来仔细分析如何得到.

我们注意到 0 阶的累积量为零. 假如我们需要计算 n 阶矩, 上式第二行中的 q 只能取 $1 \leq q \leq n$.

(1) 最简单的情况是 $q = n$, 这时

$$\frac{1}{n!} \left(\sum_{m=0}^{\infty} \frac{(ik)^m}{m!} \langle x^m \rangle_c \right)^n \quad (23.11)$$

这是 n 个多项式相乘的形式, 把它展开我们可以得到无穷个项, 但其中 $(ik)^n$ 阶的项只有一个, 即从 n 个多项式中的每一个 ik 阶项提取出来再相乘

$$\frac{1}{n!} \left(\frac{(ik)^1}{1!} \langle x^1 \rangle_c \right)^n. \quad (23.12)$$

(2) 下面我们考虑 $q = n - 1$ 的情况

$$\frac{1}{(n-1)!} \left(\sum_{m=0}^{\infty} \frac{(ik)^m}{m!} \langle x^m \rangle_c \right)^{n-1}. \quad (23.13)$$

要从上面的 $n - 1$ 个多项式相乘的展开中得到 $(ik)^n$ 阶的项的话, 需要从某一个多项式内提出一个 $(ik)^2$ 的项, 从其他所有 $n - 2$ 个多项式中提取 (ik) 阶的项. 这样的话, 共有 $n - 1$ 种组合方式, 所以:

$$(n-1) \frac{1}{(n-1)!} \left(\frac{(ik)^1}{1!} \langle x^1 \rangle_c \right)^{n-2} \left(\frac{(ik)^2}{2!} \langle x^2 \rangle_c \right)^1 = \frac{(ik)^n}{(n-2)!(1!)^{n-2}(2!)^1} \langle x \rangle_c^{n-1} \langle x^2 \rangle_c^1. \quad (23.14)$$

(3) 一般情况下, 如果要从一个 q 阶展开

$$\frac{1}{q!} \left(\sum_{m=0}^{\infty} \frac{(ik)^m}{m!} \langle x^m \rangle_c \right)^q, \quad (23.15)$$

中得到对应 $(ik)^n$ 阶项的话, 我们是从这 q 个多项式里分别提取一项得到的. 如果我们提取了 n_{m_j} 个 m_j 阶累积量的话, 那么应该有 $\sum_{m_j=1}^n n_{m_j} = q$, 因为我们一共有 q 个多项式. 而且, 我们

图 23.1: 用累积量来表示 3 和 4 阶矩的图形规则。

还需要 $\sum_{m_j=1}^n n_{m_j} m_j = n$, 以保证我们得到的是 $(ik)^n$ 阶项. 而这样的组合方式有

$$\frac{q!}{\prod_{m_j=1}^n n_{m_j}!} \quad (23.16)$$

种. 所以我们得到

$$\frac{q!}{\prod_{m_j=1}^n n_{m_j}!} \frac{1}{q!} \prod_{m_j=1}^n \left(\frac{(ik)^{m_j} \langle x^{m_j} \rangle_c}{m_j!} \right)^{n_{m_j}} = (ik)^n \prod_{m_j=1}^n \left(\frac{\langle x^{m_j} \rangle_c^{n_{m_j}}}{n_{m_j}! (m_j!)^{n_{m_j}}} \right) \quad (23.17)$$

(4) 综上, 我们可以得到:

$$\langle x^n \rangle = n! \sum_{q=1}^n \delta \left(\sum_{m_j=1}^n m_j n_{m_j} - n \right) \delta \left(\sum_{m_j=1}^n n_{m_j} - q \right) \prod_{m_j=1}^n \left(\frac{\langle x^{m_j} \rangle_c^{n_{m_j}}}{n_{m_j}! (m_j!)^{n_{m_j}}} \right). \quad (23.18)$$

这个结果可以用图形表示出来, 见图 Fig. 1.

从单变量推广到多变量

上面所有的定义我们都可以从一个随机变量推广到多个随机变量的情况。比如, 多个随机变量的分布函数为

$$F_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) = \text{Prob}\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n\}. \quad (23.19)$$

相应的概率密度分布函数为:

$$P_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) = \frac{\partial}{\partial x_1 \dots \partial x_n} F_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n). \quad (23.20)$$

特征函数、累积量产生函数等可以类似定义

$$\Phi_{X_1, \dots, X_n}(k_1, \dots, k_n) = \int dx_1 \dots \int dx_n P_{X_1, X_2, \dots, X_n}(x_1, \dots, x_n) e^{i \sum_{i=1}^n k_i x_i}. \quad (23.21)$$

最后, 对于多个变量的情况, 我们可以定义约化密度分布函数:

$$P_{X_1}(x_1) = \int P_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) dx_2 \dots dx_n. \quad (23.22)$$

常用随机分布

下面我们讨论几种常用的分布, 包括二项式、高斯和泊松分布。后两个都可以在二项式分布特定的极限下得到。

二项式分布 (Binomial distribution)

设想我们对一个随机变量 X 的测量只能得到两个可能的结果 1 和 0，概率分别为 p 和 q ($p + q = 1$)。那么对这个随机变量的多次、独立测量结果之和就满足二项式分布。

对于单次测量， X 概率密度分布很简单：

$$P_X(x) = p\delta(x - 1) + q\delta(x). \quad (23.23)$$

对应的特征函数为：

$$\Phi_X(k) = pe^{ik} + q. \quad (23.24)$$

下面我们用 N 次独立测量的结果之和来定义一个新的随机变量 $Y_N = \sum_{n=1}^N X_n$ 。在这 N 次测量中，我们得到 n_1 次 1、 n_0 次 0 的概率为：

$$P_N(n_1) = \frac{N!}{n_1!n_0!} p^{n_1} q^{n_0}, \quad n_0 + n_1 = N. \quad (23.25)$$

很容易验证，这个分布是归一化的：

$$\sum_{n_1=0}^N P_N(n_1) = \sum_{n_1=0}^N \frac{N!}{n_1!n_0!} p^{n_1} q^{n_0} = (p + q)^N = 1. \quad (23.26)$$

这是一个二项式展开的形式，所以称为二项式分布。

下面我们来看 Y_N 的概率密度分布函数。一般来说，如果已知一个随机变量 X 的概率分布 $P_X(x)$ ，那么 X 的一个函数 $Y = f(X)$ 的概率分布密度可以写为：

$$P_Y(y) = \int \delta(y - f(x)) P_X(x) dx. \quad (23.27)$$

这样， Y_N 的概率密度分布就可以写为：

$$P_{Y_N}(y) = \int \delta(y - x_1 - x_2 - \cdots - x_N) P_X(x_1) \cdots P_X(x_N) dx_1 \cdots dx_N. \quad (23.28)$$

对应的特征函数为

$$\begin{aligned} \Phi_{Y_N}(k) &= \int \int e^{iky} \delta(y - x_1 - x_2 - \cdots - x_N) P_X(x_1) \cdots P_X(x_N) dx_1 \cdots dx_N dy \\ &= \int e^{ik(x_1 + \cdots + x_N)} P_X(x_1) \cdots P_X(x_N) dx_1 \cdots dx_N \\ &= (pe^{ik} + q)^N. \end{aligned} \quad (23.29)$$

可见， N 次测量结果之和对应的特征函数为单次测量特征函数的 N 次方。从特征函数我们可以得到其累积量产生函数

$$\Psi_{Y_N}(k) = \ln \Phi_{Y_N}(k) = N \ln(pe^{ik} + q) \quad (23.30)$$

为 N 次独立测量 Y_N 对应的特征函数为单个随机变量特征函数的 N 倍。

验证：我们可以把 $\Phi_{Y_N}(k)$ 做二项式展开得到：

$$\Phi_{Y_N}(k) = \sum_{n=0}^N \frac{N!}{n!(N-n)!} q^{N-n} p^n e^{ikn}. \quad (23.31)$$

做傅立叶变换得到：

$$P_{Y_N}(y) = \sum_{n=0}^N \frac{N!}{n!(N-n)!} q^{N-n} p^n \delta(y - n). \quad (23.32)$$

这与式 (23.25) 符合。

根据这些结果我们可以计算随机变量的各阶矩。我们可以直接使用概率密度分布函数或者使用特征函数。比如：

$$\langle y \rangle = \left. \frac{\partial \Phi_{Y_N}(k)}{\partial (ik)} \right|_{k=0} = pN. \quad (23.33)$$

$$\langle y^2 \rangle = \left. \frac{\partial^2 \Phi_{Y_N}(k)}{\partial (ik)^2} \right|_{k=0} = p^2 N^2 + pqN. \quad (23.34)$$

二项式分布的方差为：

$$\langle y^2 \rangle_c = \langle y^2 \rangle - \langle y \rangle^2 = p^2 N^2 + pqN - p^2 N^2 = pqN. \quad (23.35)$$

其标准差为 $\sigma_N = \sqrt{pqN}$. 这一结果也可以从累积量产生函数来得到。我们下面会看到，二项式分布可以用来描述布朗运动粒子的运动行为。

布朗运动

一维布朗运动粒子的概率分布为：

$$P_{Y_N}(n\delta l) = \frac{N!}{\left(\frac{N-n}{2}\right)! \left(\frac{N+n}{2}\right)!} \left(\frac{1}{2}\right)^N. \quad (23.36)$$

我们可以根据这个概率分布来计算 Y_N 的均值和方差。由于颗粒的每一步位移 X_i 都是随机且相互独立的。它们满足相同的概率分布：

$$P_{X_i}(x) = \begin{cases} \frac{1}{2} & (x = \delta l) \\ \frac{1}{2} & (x = -\delta l) \end{cases}. \quad (23.37)$$

对应的特征函数也很简单

$$\Psi_{X_i}(k) = \cos k\delta l. \quad (23.38)$$

于是, Y_N 的特征函数为

$$\Psi_{Y_N}(k) = \sum_y e^{iky} P_{Y_N}(y) = \sum_y e^{iky} \sum_{x_i} P_{X_1}(x_1) \cdots P_{X_N}(x_N) \delta(y - \sum_{i=1}^N x_i) \quad (23.39)$$

$$= \sum_{x_i} e^{ik \sum_i x_i} P_{X_1}(x_1) \cdots P_{X_N}(x_N) \quad (23.40)$$

$$= \prod_i \sum_{x_i} P_{X_i}(x_i) e^{ikx_i} = \prod_i \Psi_{X_i}(k) \quad (23.41)$$

$$= (\cos k\delta l)^N. \quad (23.42)$$

根据特征函数, 我们可以更方便的求其平均值和方差

$$\boxed{\langle Y_N \rangle = 0, \quad \langle \Delta Y_N^2 \rangle = N\delta l^2.} \quad (23.43)$$

高斯分布 (Gaussian distribution)

高斯分布可以从二项式分布在独立测量次数 N 足够大的时候得到。这也是中心极限定理 (Central Limit Theorem) 的一个表现形式。我们现在根据二项式分布再定义一个新的随机变量

$$Z_N = \frac{Y_N - \langle Y_N \rangle}{\sigma_N} = \frac{Y_N - pN}{\sqrt{pqN}}. \quad (23.44)$$

其概率密度分布为:

$$P_{Z_N}(z) = \int \delta \left(z - \frac{y - pN}{\sqrt{pqN}} \right) P_{Y_N}(y) dy. \quad (23.45)$$

其特征函数为:

$$\Phi_{Z_N}(k) = \int \int e^{ikz} \delta \left(z - \frac{y - pN}{\sqrt{pqN}} \right) P_{Y_N}(y) dy dz \quad (23.46)$$

$$= \int e^{\frac{ik(y-pN)}{\sqrt{pqN}}} P_{Y_N}(y) dy \quad (23.47)$$

$$= e^{-ikpN/\sqrt{pqN}} \Phi_{Y_N}(k/\sqrt{pqN}) \quad (23.48)$$

$$= e^{-ikpN/\sqrt{pqN}} (pe^{ik/\sqrt{pqN}} + q)^N \quad (23.49)$$

$$= (pe^{ikq/\sqrt{pq}} + qe^{-ikp/\sqrt{pq}})^N. \quad (23.50)$$

我们可以对括号里的两项做一个泰勒展开:

$$\Phi_{Z_N}(k) = (pe^{ikq/\sqrt{pq}} + qe^{-ikp/\sqrt{pq}})^N \quad (23.51)$$

$$= \left\{ \sum_n \frac{(ik)^n}{n!N^{n/2}} \left[q^{n/2} p^{1-n/2} + (-1)^n p^{n/2} q^{1-n/2} \right] \right\}^N. \quad (23.52)$$

当 N 足够大的时候，我们只保留到了展开的第二阶

$$\Phi_{Z_N}(k) \approx \left(1 - \frac{k^2}{2N}\right)^N. \quad (23.53)$$

最后，利用

$$\lim_{N \rightarrow \infty} \left(1 + \frac{x}{N}\right)^N = e^x. \quad (23.54)$$

我们得到

$$\Phi_{Z_N}(k) \approx \lim_{N \rightarrow \infty} \left(1 - \frac{k^2}{2N}\right)^N = e^{-\frac{1}{2}k^2}. \quad (23.55)$$

利用上面的结果，我们得到

$$P_{Z_N}(z) = \int e^{-\frac{1}{2}k^2} e^{-ikz} \frac{dk}{2\pi} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}. \quad (23.56)$$

由于 $y = \sigma_{Y_N} z + \langle y \rangle$ ，我们可以把 z 替换为 y 得到：

$$P_{Y_N}(y) = \int P_{Z_N}(z) \delta(y - \langle y \rangle - \sigma_{Y_N} z) dz = \frac{1}{\sqrt{2\pi}\sigma_{Y_N}} \exp\left[-\frac{(y - \langle y \rangle)^2}{2\sigma_{Y_N}^2}\right]. \quad (23.57)$$

上面的分析表明，当 N 足够大的时候，二项式分布变成了高斯分布。而高斯分布的平均值和标准差与对应的二项式分布相同。对于高斯分布的性质我们后面还会做详细讨论。

泊松分布 (Poisson distribution)

当 $N \rightarrow \infty$, $p \rightarrow 0$, 但是 $Np = a \ll N$ 的时候，二项式分布变成了泊松分布。对应的特征函数变为

$$\Phi_{Y_N}(k) = (q + pe^{ik})^N = \left(1 - \frac{a}{N} + \frac{a}{N}e^{ik}\right)^N = \left[1 + \frac{a}{N}(e^{ik} - 1)\right]^N. \quad (23.58)$$

当 N 无限大的时候

$$\Phi_{Y_N}(k) = \exp[(e^{ik} - 1)a] = e^{-a} \sum_{m=0}^{\infty} \frac{a^m}{m!} e^{ikm}. \quad (23.59)$$

于是，概率密度分布为：

$$P_{Y_N}(y) = e^{-a} \int \sum_{m=0}^{\infty} \frac{a^m}{m!} e^{ik(m-y)} \frac{dk}{2\pi} = \sum_{m=0}^{\infty} \frac{a^m}{m!} e^{-a} \delta(m - y). \quad (23.60)$$

这里的系数

$$P(m) = \frac{a^m}{m!} e^{-a}$$

就是泊松分布。它只依赖于一个参数，即泊松分布的平均值

$$\sum_m m P(m) = a$$

泊松分布的累积量产生函数为：

$$\Psi_{Y_N}(k) = (e^{ik} - 1)a = a \sum_{n=1}^{+\infty} \frac{(ik)^n}{n!}. \quad (23.61)$$

可见泊松分布的所有阶累积量都相同 $\langle x^n \rangle_c = a$ 。

高斯分布

下面我们讨论一些高斯分布的性质，方便以后使用。首先，高斯分布

$$f(x) = A \exp \left[-\frac{(x - x_0)^2}{2\sigma^2} \right]. \quad (23.62)$$

由两个参数来决定，平均值 x_0 和方差 σ^2 。

归一化

其归一化因子为 $A = \frac{1}{\sqrt{2\pi}\sigma}$ 。简单证明如下：

$$\begin{aligned} \int f(x) dx &= \left[\int \int f(x)f(y) dx dy \right]^{1/2} \\ &= \left[\int \int \exp \left(-\frac{(x - x_0)^2 + (y - y_0)^2}{2\sigma^2} \right) dx dy \right]^{1/2} \\ &= \left[\int 2\pi \exp \left[-\frac{r^2}{2\sigma^2} \right] r dr \right]^{1/2} = \sqrt{2\pi}\sigma. \end{aligned} \quad (23.63)$$

特征函数

事实上，我们可以把高斯分布的定义推广到整个复平面。上面的计算也同样适用于复平面上的高斯函数：

$$f(z) = A \exp \left[-\frac{(z - z_0)^2}{2\sigma^2} \right], \quad (23.64)$$

这里 z 和 z_0 为复数。利用这个性质，我们可以计算高斯分布的特征函数：

$$\begin{aligned}
 \Phi(k) &= \int f(x)e^{ikx}dx = A \int \exp\left[-\frac{(x-x_0)^2}{2\sigma^2} + ikx\right] dx \\
 &= A \int \exp\left[-\frac{(x-x_0)^2 - 2i\sigma^2 kx}{2\sigma^2}\right] dx \\
 &= A \int \exp\left[-\frac{[x-(x_0+i\sigma^2 k)]^2}{2\sigma^2}\right] dx \exp\left[\frac{ik(2x_0+i\sigma^2 k)}{2}\right] \\
 &= \exp[ikx_0] \exp\left[-\frac{1}{2}k^2\sigma^2\right] \\
 &= \exp\left[-\frac{x_0^2}{2\sigma^2}\right] \exp\left[-\frac{1}{2}\sigma^2\left(k-\frac{ix_0}{\sigma^2}\right)^2\right].
 \end{aligned}$$

我们看到，高斯分布的特征函数仍然是高斯分布。知道了特征函数，高斯分布的所有阶矩都可以求得。

累积量函数

其累积量产生函数为

$$\Psi(k) = \ln\Phi(k) = ikx_0 - \frac{1}{2}k^2\sigma^2. \quad (23.65)$$

由此可见，高斯分布的大于二阶的累积量都是零。

多维情况

以上的讨论都适用于 x 是一维的情况，我们可以把这些结果推广到二维和三维的情况。一般来说，其归一化因子 $A = 1/(2\pi\sigma^2)^{d/2}$ 。我们可以定义包含多个变量的高斯分布

$$f(x_i) = A \exp\left[-\frac{1}{2} \sum_{j,k} (x_j - \bar{x}_j)([\sigma^2]^{-1})_{kj}(x_k - \bar{x}_k)\right]. \quad (23.66)$$

现在 σ^2 变成了一个矩阵。这时要得到归一化因子，我们需要做一个线型变换。我们让 $[\sigma^2]^{-1} = B$ ，并把 B 对角化

$$B = M^T \Lambda M, \quad u = Mx. \quad (23.67)$$

这里， M 为变换矩阵， Λ 为一对角矩阵，其本征值用 λ_i 表示。作此变换之后的高斯分布可以写为

$$f(u_i) = A \exp\left[-\frac{1}{2} \sum_i \lambda_i (u_i - \bar{u}_i)^2\right]. \quad (23.68)$$

这样很容易求出归一化因子应该为 $A = \prod_i \lambda_i^{1/2} / (2\pi)^{N/2} = (2\pi)^{-N/2} \det(\sigma^2)^{-1/2}$. 其特征函数为

$$\Phi(k_i) = \exp \left[i \sum_j \bar{x}_j k_j - \frac{1}{2} \sum_{j,k} k_j \sigma_{jk}^2 k_k \right], \quad (23.69)$$

$$\Psi(k_i) = i \sum_j \bar{x}_j k_j - \frac{1}{2} \sum_{j,k} k_j \sigma_{jk}^2 k_k. \quad (23.70)$$

我们看到高斯分布只有第一和第二阶的累积量非零，它们分别是随机变量的平均值和方差。

最后，我们指出如果随机变量 X 服从高斯分布，那么其任意线型变换 $Y(X)$ 仍然满足高斯分布

$$\langle y(t) \rangle = \int_a^b C(t, t') \langle z(t') \rangle dt', \quad (23.71)$$

$$\langle y(t_1) y(t_2) \rangle_C = \int_a^b dt'_1 \int_a^b dt'_2 C(t_1, t'_1) C(t_2, t'_2) \langle z(t'_1) z(t'_2) \rangle_C. \quad (23.72)$$

从上面特征函数的计算我们已经发现，高斯分布的特征函数也服从高斯分布，因为它是分布函数的傅立叶变换。

练习：

- 对于单变量的情况，利用高斯分布的特征函数

$$\Phi(k) = \exp \left(-\frac{1}{2} k^2 \sigma^2 \right), \quad (23.73)$$

以及特征函数的性质 (23.4)，证明：

$$\langle x^{2n} \rangle = \frac{(2n)!}{2^n n!} (\langle x^2 \rangle)^n, \quad (23.74)$$

$$\langle x^{2n+1} \rangle = 0. \quad (23.75)$$

- 对于多变量的情况，仍然假设每个变量的平均值为零 (式 (23.69) 中 $\bar{x}_j = 0$)，证明类似的结论： $\langle x_1 x_2 x_3 \cdots x_{2n+1} \rangle = 0$ ，并且 $\langle x_1 x_2 x_3 \cdots x_{2n} \rangle$ 等于所有两两组合的平均乘积之和。比如：

$$\langle x_1 x_2 x_3 x_4 \rangle = \langle x_1 x_2 \rangle \langle x_3 x_4 \rangle + \langle x_1 x_3 \rangle \langle x_2 x_4 \rangle + \langle x_1 x_4 \rangle \langle x_2 x_3 \rangle. \quad (23.76)$$

在量子场论里面多个算符相对于一个二阶哈密顿量的量子平均都可以用路径积分的方法写成高斯积分的形式。上面的结果就是量子场论里面的一个非常重要的结果叫 Wick 定理。当然，对于量子算符来说我们在交换它们的位置的时候需要考虑其交换对称性。

随机过程

23.0.1 随机过程

- 随机过程可以用一系列的随机变量来表示。当我们有无穷多的随机变量 X_1, X_2, \dots 时，使用离散的指标 i 就会很不方便。这时我们可以引入一个连续变量 t ，把这一系列随机变量表示为 t 的函数 $X(t)$ 。每给定一个参数 t ，就有一个对应的随机变量 $X(t)$ 。
- 对于任意给定参数 t , $X(t)$ 都有一个对应的取值 $x(t)$ 。这些 $x(t)$ 构成了随机过程 $X(t)$ 的一个实现，对应对这个随机过程的一次取样。
- 假设我们产生 N 个 $X(t)$ 的具体实现 $x^i(t), i = 1, \dots, N$ 。 $X(t)$ 的均值还可以表示为：

$$\langle X(t) \rangle = \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_i x^i(t).$$

- 给定 N 个具体实现，我们可以按照通常的方法定义随机变量 $X(t)$ 满足的概率密度分布 $P_X(x, t) = \sum_i \delta(x_i(t) - x)/N$ 。该时刻 $X(t)$ 的均值还可以表示为：

$$\langle X(t) \rangle = \int dx P_X(x, t) x.$$

以上两个均值的表达式是等价的，我们统称为 $X(t)$ 的系综平均或统计平均。

- 以上的平均是针对任意给定的一个时刻 t 进行的。类似地，我们还可以定义 $\langle X(t_1)X(t_2) \rangle$, $\langle X(t_1)X(t_2)X(t_3) \rangle$ 等多个时刻随机变量之间的平均，表示不同时刻的关联。与单时刻类似：

$$\langle X(t_1)X(t_2) \rangle = \int dx_1 dx_2 P_X(x_1, t_1; x_2, t_2) x_1 x_2.$$

或

$$= \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_i x^{(i)}(t_1) x^{(i)}(t_2).$$

理论上，我们需要给出任意一组时刻上随机变量之间的关联才能确定这个随机过程的统计性质。

- 对于**高斯**随机过程，我们只需要指定其单时刻和双时刻的关联就可以把整个过程确定下来。它的所有更高阶关联都可以用这两个关联函数表示出来。通常来说，我们最常用的就是高斯随机过程。
- **稳态**随机过程的关联函数具有时间平移不变性，也就是说它不依赖于时间的绝对值，比

如,

$$\langle X(t) \rangle = \langle X(t') \rangle = \langle X \rangle, \quad \langle X(t_1)X(t_2) \rangle = \langle X(t_1 - t_2)X(0) \rangle.$$

此外, 对于随机过程 $X(t)$ 的任意一组实现 $x^{(i)}(t)$, 我们可以对其进行时间平均:

$$\lim_{T \rightarrow +\infty} \int_{t-T/2}^{t+T/2} dt' x^{(i)}(t') = \overline{X(t)}^T.$$

如果随机过程满足各态遍历的特征, 那么

$$\begin{aligned}\overline{X(t)}^T &= \langle X(t) \rangle \\ \overline{X(t_1) \cdots X(t_n)}^T &= \langle X(t_1) \cdots X(t_n) \rangle.\end{aligned}$$

我们这个课程的讨论仅限于最简单的稳态、高斯随机过程, 并且我们假设它满足各态遍历的特征。

23.0.2 Markov 过程和 Chapman-Kolmogorov 方程

下面我们进一步引入一些数学上的定义来描述一个随机过程 $X(t)$ 的概率密度分布。

- $P_1(x_1, t_1)$ 表示时刻 t_1 随机变量 X 取值为 x_1 的概率
- $P_2(x_1, t_1; x_2, t_2)$ 表示随机变量 X 在 t_1 时刻取值 x_1 , 在 t_2 时刻取值 x_2 的联合概率密度
- ...
- $P_n(x_1, t_1; \dots; x_n, t_n)$ 的意义以此类推。

根据这些定义我们可以得到下面的性质:

- 归一化: $\int P_1(x_1, t_1) dx_1 = 1$
- $\int P_n(x_1, t_1; \dots; x_n, t_n) dx_n = P_{n-1}(x_1, t_1; \dots; x_{n-1}, t_{n-1})$
- 平均值: $\langle x_1(t_1)x_2(t_2) \cdots x_n(t_n) \rangle = \int \cdots \int x_1 x_2 \cdots x_n P(x_1, t_1; \dots; x_n, t_n) dx_1 \cdots dx_n$

稳态随机过程满足:

$$P_n(x_1, t_1; x_2, t_2; \dots; x_n, t_n) = P_n(x_1, t_1 + \tau; x_2, t_2 + \tau; \dots; x_n, t_n + \tau). \quad (23.77)$$

下面需要定义的是**条件概率分布** $P_{1|1}(x_1, t_1 | x_2, t_2)$. 它表示的是, 在已知一个随机变量 X 在 t_1 时刻的取值为 x_1 的情况下, 它在 t_2 时刻的取值为 x_2 的概率. 根据定义, 我们有

$$P_{1|1}(x_1, t_1 | x_2, t_2) = \frac{P_2(x_1, t_1; x_2, t_2)}{P_1(x_1, t_1)}. \quad (23.78)$$

其归一化条件为:

$$\int P_{1|1}(x_1, t_1 | x_2, t_2) dx_2 = 1. \quad (23.79)$$

类似的，我们还可以定义联合条件概率密度：

$$P_{m|n}(x_1, t_1; \dots; x_m, t_m | x_{m+1}, t_{m+1}; \dots; x_{m+n}, t_{m+n}) = \frac{P_{m+n}(x_1, t_1; \dots; x_{m+n}, t_{m+n})}{P_m(x_1, t_1; \dots; x_m, t_m)}. \quad (23.80)$$

Markov 过程满足：

$$\boxed{P_{n-1|1}(t_1, x_1; \dots; t_{n-1}, x_{n-1} | t_n, x_n) = P_{1|1}(t_{n-1}, x_{n-1} | t_n, x_n)}. \quad (23.81)$$

其中，我们把 $P_{1|1}(t_{n-1}, x_{n-1} | t_n, x_n)$ 叫作跃迁几率。这说明，马尔柯夫过程在 t_n 时刻的取值概率只由前一个时刻 t_{n-1} 完全决定，与 t_{n-1} 之前时刻的历史没有任何关系，也就是没有任何记忆效应。这样的话，一个马尔柯夫过程可以由概率密度分布 $P_1(x_1)$ 和跃迁几率 $P_{1|1}(t_1, x_1 | t_2, x_2)$ 来完全确定。比如：

$$\begin{aligned} P_3(x_1, t_1; x_2, t_2; x_3, t_3) &= P_2(x_1, t_1; x_2, t_2) P_{2|1}(x_1, t_1; x_2, t_2 | x_3, t_3) \\ &= P_2(x_1, t_1; x_2, t_2) P_{1|1}(x_2, t_2 | x_3, t_3) \\ &= P_1(x_1, t_1) P_{1|1}(x_1, t_1 | x_2, t_2) P_{1|1}(x_2, t_2 | x_3, t_3). \end{aligned}$$

对于 P_n 可以类似证明。

在式子 (23.82) 两边同时对 x_2 做积分，我们可以得到：

$$\boxed{P(x_1, t_1 | x_3, t_3) = \int P(x_1, t_1 | x_2, t_2) P(x_2, t_2 | x_3, t_3) dx_2}. \quad (23.82)$$

这个方程叫做 Chapman-Kolmogorov (C-K) 方程。这里我们把符号做了简化，把 P 的下标 $1|1$ 省略了。这个方程理解起来非常简单。方程的右边表示粒子从 t_1 时刻的位置 x_1 到达 t_3 时刻的位置 x_3 必然要在中间时刻 t_2 处于某一个位置 x_2 。学过费曼路径积分的同学马上就意识到路径积分的基本出发点即此。另外，随机过程的马尔柯夫特性表现在这些跃迁几率只依赖于起始和终止时刻的位置，与粒子如何到达起始时刻的没有关系。

23.0.3 Fokker-Planck 方程

下面我们来推导 Fokker-Planck 方程。我们的出发点是 Markov 过程在时间微元 Δt 内所满足的一个方程：

$$P(q', t | q, t + \Delta t) = \delta(q' - q)(1 - \gamma(q')\Delta t) + w_{q'q}\Delta t + O(\Delta t^2). \quad (23.83)$$

这个方程的左边是一个条件概率，即已知粒子在 t 时刻位于位置 q' 的话，其在 $t + \Delta t$ 时刻位于位置 q 的概率。方程的右边两项描述了粒子在 $t + \Delta t$ 时刻到达 q 的两种方式：方程右边第二项表示粒子在 Δt 的时间内从 q' 跃迁到 q 的概率。其中 $w_{q'q} = \langle q' | w | q \rangle$ 表示跃迁速率，而 $w_{q'q}\Delta t$

为跃迁概率。这里我们假设它不依赖时间。方程右边第一项对应 $q = q'$ 的情况，那么只需要粒子保持在原来的位置就可以了，这个概率为 $(1 - \gamma(q')\Delta t)$ ，而 $\gamma(q') = \int w_{q'q} dq$ ， $\gamma(q')\Delta t$ 表示粒子从 q' 跃迁到其他任何状态的概率。

我们把上面的式子带入到 C-K 方程

$$P(q', t_1 | q, t) = \int P(q', t_1 | q'', t - \Delta t) P(q'', t - \Delta t | q, t) dq''. \quad (23.84)$$

可以得到：

$$\begin{aligned} &= \int P(q', t_1 | q'', t - \Delta t) [\delta(q'' - q)(1 - \gamma(q'')\Delta t) + w_{q''q}\Delta t] dq'' \\ &= P(q', t_1 | q, t - \Delta t)(1 - \gamma(q)\Delta t) + \Delta t \int P(q', t_1 | q'', t - \Delta t) w_{q''q} dq''. \end{aligned}$$

整理得到微分方程的形式：

$$\boxed{\frac{\partial}{\partial t} P(q', t_1 | q, t) = -\gamma(q)P(q', t_1 | q, t) + \int w_{q''q} P(q', t_1 | q'', t) dq''}. \quad (23.85)$$

这是 Markov 过程中的概率密度分布 $P(q, t)$ 所满足的主方程。等式第一项表示系统离开 q 状态的速率，第二项表示从其他态跃迁到 q 态的速率。

下面我们从主方程出发来得到 F-P 方程。为此，我们定义 $w_{q''q} = w(q'', q - q'' \equiv \Delta q)$ ，上式还可以写成：

$$\frac{\partial}{\partial t} P(q', t_1 | q, t) = - \left(\int w(q, \Delta q) d\Delta q \right) P(q', t_1 | q, t) + \int w(q - \Delta q, \Delta q) P(q', t_1 | q - \Delta q, t) d\Delta q. \quad (23.86)$$

考虑 $\Delta q \ll q$ ，我们可以对 Δq 做微扰展开：

$$\sum_{n=0}^{\infty} \frac{(-r)^n}{n!} \frac{\partial^n}{\partial q^n} f(q) = f(q - r), \quad (23.87)$$

式 (23.86) 还可以写成：

$$\frac{\partial}{\partial t} P(q', t_1 | q, t) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial q^n} \left[P(q', t_1 | q, t) \int (\Delta q)^n w(q, \Delta q) dr \right]. \quad (23.88)$$

定义

$$\alpha_n(q) \equiv \int d\Delta q (\Delta q)^n w(q, \Delta q) \quad (23.89)$$

$$= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int d\Delta q (\Delta q)^n P(q, t | q + \Delta q, t + \Delta t) \quad (23.90)$$

$$= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \langle \Delta q^n \rangle_q, \quad (23.91)$$

我们最后得到:

$$\boxed{\frac{\partial}{\partial t} P(q', t_1 | q, t) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial q^n} [P(q', t_1 | q, t) \alpha_n(q)]}. \quad (23.92)$$

如果 $\alpha_n = 0 (n \geq 3)$, 那么上面的方程就叫做 Fokker-Planck 方程。稳定的高斯 Markov 过程满足以上条件。

给定 Fokker-Planck 方程, 我们可以定义概率流:

$$j(q', t_1 | q, t) = P(q', t_1 | q, t) \alpha_1(q) - \frac{1}{2} \partial_q [P(q', t_1 | q, t) \alpha_2].$$

第一项为漂移项, 第二项为扩散项。这样我们有:

$$\frac{\partial}{\partial t} P(q', t | q, t) = -\partial_q j(q', t | q, t).$$

举例 1: $q \rightarrow v$: 对于布朗粒子来说,

$$\begin{aligned} \lim_{\Delta t \rightarrow 0} \frac{\langle \Delta v \rangle}{\Delta t} &= -\gamma v \equiv F(v), \\ \lim_{\Delta t \rightarrow 0} \frac{\langle \Delta v^2 \rangle}{\Delta t} &= \frac{2\gamma k_B T}{m} \equiv 2D_v, \\ \lim_{\Delta t \rightarrow 0} \frac{\langle \Delta v^n \rangle}{\Delta t} &= 0 \quad (n \geq 3), \end{aligned}$$

所以, 对应的方程为:

$$\begin{aligned} \frac{\partial}{\partial t} P(v_0, 0 | v, t) &= \gamma \frac{\partial}{\partial v} \left(v + \frac{k_B T}{m} \frac{\partial}{\partial v} \right) P(v_0, 0 | v, t) \\ &= -\frac{\partial}{\partial v} [F(v) P(v_0, 0 | v, t)] + \frac{\partial^2}{\partial v^2} [D_v P(v_0, 0 | v, t)]. \end{aligned}$$

- **稳态解:** 在稳态下, $\frac{\partial}{\partial t} P(v_0, 0 | v, t) = 0$, 所以

$$\left(v + \frac{k_B T}{m} \frac{\partial}{\partial v} \right) P(v_0, 0 | v, t) = \text{const} = 0. \quad (23.93)$$

因为 $P \rightarrow 0$ if $v \rightarrow +\infty$. 所以我们得到平衡态的时候布朗粒子服从麦克斯韦-玻尔兹曼分布:

$$P(v_0, 0 | v, t) \propto \exp \left[-\frac{mv^2}{2k_B T} \right]. \quad (23.94)$$

稳态解与初态无关。

- **瞬态解:** 如果给定初始条件为

$$P(v, 0) = \delta(v - v_0), \quad (23.95)$$

那么我们可以求解 Fokker-Planck 方程, 解正好就是我们前面得到的 Ornstein-Uhlenbeck

分布

$$P(v_0, 0|v, t) = \left(\frac{m}{2\pi k_B T (1 - e^{-2\gamma t})} \right)^{\frac{1}{2}} \exp \left\{ -\frac{m(v - v_0 e^{-\gamma t})^2}{2k_B T (1 - e^{-2\gamma t})} \right\}. \quad (23.96)$$

举例 2: q 为位移 x :

请说明对应的 F-P 方程为:

$$\frac{\partial}{\partial t} P(x, t) = -\partial_x(v(t)P(x, t)) + \partial_{x^2}(DP(x, t)).$$

我们可以定义概率流:

$$j_x(x, t) = v(t)P(x, t) - \partial_x(DP(x, t)).$$

上式第一项为漂移项，第二项为扩散项。这样 F-P 方程还可以写成概率守恒的形式:

$$\frac{\partial}{\partial t} P(x, t) = -\partial_x j(x, t).$$

举例 3: 对于一个在束缚势限制下作布朗运动的粒子，我们可以令 $q = (x, v)$ ，那么相空间的 F-P 方程为:

$$\begin{aligned} \frac{d}{dt}x &= v \\ \frac{d}{dt}v &= -V'(x) - \gamma v + f(t)/m, \end{aligned}$$

于是我们有:

$$\begin{aligned} \lim_{\Delta t \rightarrow 0} \frac{\langle \Delta x \rangle}{\Delta t} &= v, \\ \lim_{\Delta t \rightarrow 0} \frac{\langle \Delta v \rangle}{\Delta t} &= -V'(x) - \gamma v, \\ \lim_{\Delta t \rightarrow 0} \frac{\langle \Delta x^2 \rangle}{\Delta t} &= 0, \\ \lim_{\Delta t \rightarrow 0} \frac{\langle \Delta v^2 \rangle}{\Delta t} &= \frac{2\gamma k_B T}{m}, \\ \lim_{\Delta t \rightarrow 0} \frac{\langle \Delta v^n \rangle}{\Delta t} &= 0 \quad (n \geq 3), \end{aligned}$$

对应的 Fokker-Planck 方程又称为 Kramers 方程

$$\frac{\partial}{\partial t} P(x, p, t) = \left[-\frac{\partial}{\partial x} \frac{p}{m} + \frac{\partial}{\partial p} \frac{\partial V}{\partial x} + \gamma \frac{\partial}{\partial p} \left(p + m k_B T \frac{\partial}{\partial p} \right) \right] P(x, p, t). \quad (23.97)$$

对应的稳态解我们可以猜测出来为 Boltzmann 分布:

$$P_{eq}(x, p) \propto \exp \left[-\frac{V(x) + p^2/2m}{k_B T} \right]. \quad (23.98)$$

我们推导 F-P 方程的时候要求 Langevin 方程描述的为高斯、Markov 过程，但并不要

求 Langevin 方程是线性的，尽管我们这里考虑的大部分都是线性的问题。但对于非线性的 Langevin 方程处理起来稍微复杂，具体可以参考相关教材 [比如 Kubo, Statistical Physics II 的 2.4 节]，这里不做介绍。

第 24 章 附录 III: 常用积分

24.1 Γ 函数

整数 n 的阶乘可用下面的积分形式表示：

$$n! = \int_0^{+\infty} dx x^n e^{-x}. \quad (24.1)$$

上式可以用递推的方式进行证明。首先， $n = 0$ 时，我们可以证明其成立。其次，如果 $n = k$ 成立，那么在上式两边同乘 $k + 1$ ，然后对等式右边做分部积分，即可证明当 $n = k + 1$ 时它同样成立。

我们可以对式 (24.1) 进行延拓，作为非整数和负数阶乘的定义。这个函数还有一个特殊的名字称为 Γ 函数：

$$\Gamma(n+1) = \int_0^{+\infty} dx x^n e^{-x} \quad (24.2)$$

因此， $\Gamma(n+1) = n!$ 。表 (24.1) 给出常用的 Γ 函数的数值。

z	$-\frac{3}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{3}{2}$	2	$\frac{5}{2}$	3	4
$\Gamma(z)$	$\frac{4\sqrt{\pi}}{3}$	$-2\sqrt{\pi}$	$\sqrt{\pi}$	1	$\frac{\sqrt{\pi}}{2}$	1	$\frac{3\sqrt{\pi}}{4}$	2	6

表 24.1: Γ 函数的取值。利用 $\Gamma(z+1) = z\Gamma(z)$ 可以生成其他的结果。

24.2 高斯积分

高斯函数具有 e^{-ax^2} 的形式，是统计物理中最为常用的函数，也被称为正态分布函数。最为有用的高斯函数积分为；

$$\int_{-\infty}^{+\infty} dx e^{-ax^2} = \sqrt{\frac{\pi}{a}} \quad (24.3)$$

或者

$$\int_{-\infty}^{+\infty} dx e^{-\frac{x^2}{2\sigma^2}} = \sqrt{2\pi\sigma^2} \quad (24.4)$$

基于这一结果我们可以得到一系列的相关积分：

$$\int_{-\infty}^{+\infty} x^2 e^{-ax^2} dx = \frac{1}{2} \sqrt{\frac{\pi}{a^3}} \quad \int_{-\infty}^{+\infty} x^4 e^{-ax^2} dx = \frac{3}{4} \sqrt{\frac{\pi}{a^5}}$$

以上结果可以推广

$$\int_{-\infty}^{+\infty} x^{2n} e^{-ax^2} dx = \frac{(2n)!}{n! 2^n} \sqrt{\frac{\pi}{a^{2n+1}}}.$$

据此，我们又可以得到正半轴的相应积分：

$$\begin{aligned}\int_0^{+\infty} dx e^{-ax^2} &= \frac{1}{2} \sqrt{\frac{\pi}{a}}, \\ \int_0^{+\infty} x^2 e^{-ax^2} dx &= \frac{1}{4} \sqrt{\frac{\pi}{a^3}}, \\ \int_0^{+\infty} x^4 e^{-ax^2} dx &= \frac{3}{8} \sqrt{\frac{\pi}{a^5}}, \\ \int_0^{+\infty} x e^{-ax^2} dx &= \frac{1}{2a}, \\ \int_0^{+\infty} x^3 e^{-ax^2} dx &= \frac{1}{2a^2}, \\ \int_0^{+\infty} x^5 e^{-ax^2} dx &= \frac{1}{a^3}.\end{aligned}$$

将奇次幂的积分推广

$$\int_0^{+\infty} x^{2n+1} e^{-ax^2} dx = \frac{n!}{2a^{n+1}}.$$

24.3 Stirling 公式

Stirling 公式是统计物理中最为重要、最为常用的一个公式：

$$\boxed{\ln n! = n \ln n - n + \frac{1}{2} \ln 2\pi n + \dots} \quad (24.5)$$

当 n 足够大时，我们可以忽略其他的高阶项

$$\boxed{\ln n! \approx n \ln n - n.} \quad (24.6)$$

这个近似可以从 Γ 函数的定义出发进行证明：

$$n! = \int_0^{\infty} x^n e^{-x} dx$$

上式等号右边的被积函数中，第一项 x^n 随着 x 的增加单调递增，而 e^{-x} 单调递减。当 x 较小时， x^n 占主导，所以整个被积函数随 x 增加而增加。但当 x 足够大时， e^{-x} 占主导，被积函数随 x 增加而减小。因此，被积函数必然在中间某个位置有一个极大值。因为我们最后要取 \ln ，所以不妨取

$$e^{f(x)} = x^n e^{-x}$$

我们不难找到其取最大值时

$$\frac{df(x)}{dx} = 0 \rightarrow x = n.$$

再次微分

$$\frac{d^2f(x)}{dx^2} = -\frac{n}{x^2}.$$

对 $f(x)$ 在最大值附近做 Taylor 展开

$$\begin{aligned} f(x) &= f(n) + \left(\frac{df}{dx}\right)_{x=n}(x-n) + \frac{1}{2} \left(\frac{d^2f}{dx^2}\right)_{x=n}(x-n)^2 + \dots \\ &= n \ln n - n - \frac{n}{n^2}(x-n)^2 + \dots \end{aligned}$$

当 n 足够大时，积分的主要贡献将来自于极大值附近的位置，因此我们可以忽略展开式中的高阶项。将上式代入到最初的积分，我们得到：

$$n! = e^{n \ln n - n} \int_0^\infty dx e^{-\frac{(x-n)^2}{2n}} + \dots = e^{n \ln n - n} + \sqrt{2\pi n} + \dots$$

两边取对数，我们就可以得到 Stirling 公式。

24.4 n 维球的体积和表面积

考虑 n 维空间中半径为 R 的球，其体积为：

$$V_n(R) = \int_{-\infty}^{+\infty} dx_1 \cdots \int_{-\infty}^{+\infty} dx_n \Theta\left(R^2 - \sum x_i^2\right)$$

这里 $\Theta(x)$ 为阶跃函数。我们假设

$$V_n(R) = C_n R^n,$$

我们的任务是确定 C_n 。两边对 R 求导，我们可以得到：

$$dV_n(R) = S_n dR = n C_n R^{n-1} dR.$$

这里进一步引入了 n 维球面的面积 S_n ，也由 C_n 确定。

为了计算 C_n ，我们借助高斯积分

$$\sqrt{\pi} = \int_{-\infty}^{+\infty} dx e^{-x^2}.$$

将以上结果扩展到 n 维:

$$\begin{aligned}\pi^{n/2} &= \int_{-\infty}^{+\infty} dx_1 \cdots \int_{-\infty}^{+\infty} dx_n e^{-\sum_{i=1}^n x_i^2} \\&= \int_0^R dR S_n(R) e^{-R^2} \\&= \int_0^R dR n C_n R^{n-1} e^{-R^2} \\&= \frac{1}{2} n C_n \Gamma\left(\frac{n}{2}\right) \\&= C_n \left(\frac{n}{2}\right)!\end{aligned}$$

所以,

$$C_n = \frac{\pi^{n/2}}{(n/2)!}.$$

于是,

$$\begin{aligned}S_n(R) &= \frac{2\pi^{n/2}}{\Gamma(n/2)} R^{n-1} \\V_n(R) &= \frac{\pi^{n/2}}{(n/2)!} R^n\end{aligned}$$

24.5 Riemann ζ 函数

ζ 函数的定义为:

$$\zeta(n) = \sum_{k=1}^{\infty} \frac{1}{k^n} \quad (24.7)$$

常用的几个取值如表 (24.2) 所示。

n	$\zeta(n)$
1	∞
$\frac{3}{2}$	≈ 2.612
2	$\pi^2/6 \approx 1.645$
$\frac{5}{2}$	≈ 1.341
3	≈ 1.20206
4	$\pi^4/90 \approx 1.0823$
5	≈ 1.0369
6	$\pi^6/945 \approx 1.017$

表 24.2: ζ 函数的几个常用取值。

24.6 多对数函数

多对数函数的定义为：

$$\text{Li}_n(z) = \sum_{k=1}^{\infty} \frac{z^k}{k^n}. \quad (24.8)$$

这里， z 位于复平面的单位圆内部（可以将其解析延拓至复平面）。它具有下面的一些常用性质：

- 多对数函数的微分：

$$z \frac{\partial \text{Li}_n(z)}{\partial z} = \text{Li}_{n-1}(z).$$

•

$$\lim_{\operatorname{Re}(\mu) \rightarrow \infty} \text{Li}_s(\pm e^\mu) = -\frac{\mu^s}{\Gamma(s+1)}, \quad s \neq -1, -2, -3, \dots \quad (24.9)$$

$$\lim_{\operatorname{Re}(\mu) \rightarrow \infty} \text{Li}_{-n}(e^\mu) = -(-1)^n e^{-\mu}, \quad n = 1, 2, 3, \dots \quad (24.10)$$

•

$$\text{Li}_n(z) \approx z, \quad |z| \rightarrow 0 \quad (24.11)$$

•

$$\text{Li}_n(1) = \sum_{k=1}^{\infty} \frac{1}{k^n} = \zeta(n) \quad (24.12)$$

- 当 $n = 1$ 时，它退化为自然对数

$$\text{Li}_1(z) = -\ln(1-z)$$

- 当 $|z| \gg 1$ 时，我们有如下渐进展开形式：

$$\text{Li}_s(z) = \sum_{k=0}^{\infty} (-1)^k (1 - 2^{1-2k}) (2\pi)^{2k} \frac{B_{2k}}{(2k)!} \frac{[\ln(-z)]^{s-2k}}{\Gamma(s+1-2k)} \quad (24.13)$$

where B_{2k} are the Bernoulli numbers. Both versions hold for all s and for any $\arg(z)$. As usual, the summation should be terminated when the terms start growing in magnitude. For negative integer s , the expansions vanish entirely; for non-negative integer s , they break off after a finite number of terms. Wood (1992, § 11) describes a method for obtaining these series from the Bose-Einstein integral representation (his equation 11.2 for $\text{Li}_s(e^t)$ requires $2 < \operatorname{Im}(t) < 0$).¹

¹from Wikipedia

多对数函数在计算玻色和费米分布函数的积分时比较有用，主要体现在下面的公式：

$$\boxed{\int_0^{+\infty} \frac{x^{n-1} dx}{\lambda^{-1} e^x \pm 1} = \mp \Gamma(n) \text{Li}_n(\mp \lambda).} \quad (24.14)$$

下面我们来证明上式。利用：

$$\frac{1}{z^{-1} e^x \pm 1} = \mp \sum_{m=0}^{\infty} (\mp z e^{-x})^{m+1},$$

积分可以表示为：

$$\begin{aligned} \int_0^{\infty} dx \frac{x^{n-1}}{z^{-1} e^x - 1} &= \mp \sum_{m=0}^{\infty} \int_0^{\infty} dx x^{n-1} (\mp z e^{-x})^{m+1} \\ &= \mp \sum_{m=0}^{\infty} (\mp z)^{m+1} \int_0^{\infty} dx x^{n-1} e^{-(m+1)x} \\ &= \mp \sum_{m=0}^{\infty} \frac{(\mp z)^{m+1}}{(m+1)^n} \int_0^{\infty} dy y^{n-1} e^{-y} \\ &= \mp \Gamma(n) \sum_{m=0}^{\infty} \frac{(\mp z)^{m+1}}{(m+1)^n} \\ &= \mp \Gamma(n) \sum_{k=1}^{\infty} \frac{(\mp z)^k}{k^n} \\ &= \mp \Gamma(n) \text{Li}_n(\mp z). \end{aligned}$$

第 25 章 附录：

25.1 简谐振子

考虑一个简谐振子，其哈密顿量为：

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega_0^2 q^2.$$

其位移 q 和动量 p 满足的运动方程为：

$$\dot{q} = p/m, \quad \dot{p} = -m\omega_0^2 q. \quad (25.1)$$

这个形式既适用于经典又适用于量子情况。

25.1.1 本征能谱和波函数

谐振子的能级为等间隔的离散形式：

$$\varepsilon_n = \left(n + \frac{1}{2}\right) \hbar\omega_0, \quad n = 0, 1, 2, \dots$$

对应的本征波函数在坐标空间的表示为：

$$\phi_n(q) = \langle q | n \rangle = \frac{1}{\sqrt{2^n n!}} \frac{1}{(2\pi)^{1/4} l^{1/2}} \exp\left(-\frac{x^2}{4l^2}\right) H_n\left(\frac{x}{\sqrt{2}l}\right).$$

其中，谐振子的特征长度为：

$$l = \sqrt{\frac{\hbar}{2m\omega_0}}.$$

Hermite 多项式为：

$$H_n(z) = (-1)^n e^{z^2} \frac{d^n}{dz^n} \left(e^{-z^2}\right) \quad (25.2)$$

$$= \frac{e^{z^2}}{\sqrt{\pi}} \int_{-\infty}^{+\infty} (-2iu)^n \exp(-u^2 + 2izu) du. \quad (25.3)$$

比如，基态波函数为：

$$\phi_0(q) = \frac{1}{(2\pi)^{1/4} l^{1/2}} e^{-\frac{q^2}{4l^2}},$$

第一激发态波函数为：

$$\phi_1(q) = \frac{1}{(2\pi)^{1/4} l^{1/2}} \frac{q}{l} e^{-\frac{q^2}{4l^2}}$$

25.1.2 产生和湮灭算符

二次量子化下的产生、湮灭算符与位移和动量之间的关系为：

$$q = l(a^\dagger + a), \quad (25.4)$$

$$p = im\omega_0 l(a^\dagger - a). \quad (25.5)$$

或者我们可以写为：

$$a = \frac{1}{2m\omega_0 l}(m\omega_0 q + ip), \quad (25.6)$$

$$a^\dagger = \frac{1}{2m\omega_0 l}(m\omega_0 q - ip). \quad (25.7)$$

利用产生和湮灭算符，简谐振子的哈密顿量可以表示为：

$$H = (a^\dagger a + 1/2)\hbar\omega_0$$

对应的本征态波函数也可以从其基态波函数出发来构建：

$$\phi_n(q) = \langle q | n \rangle = \langle q | \frac{1}{\sqrt{n!}}(a^\dagger)^n | 0 \rangle.$$

第 26 章 授课时间表

20220301	课程简介和复习	-
20220303	最大熵原理、系综、经典 Liouville 运动方程	讲义第 9 章
20220307	经典理想气体、密度矩阵、von Neumann/量子 Liouville 方程	讲义第 10 章
20220309	理想气体: 半经典和量子描述	讲义第 11, 13 章

表 26.1: 2022 年春实际授课时间表