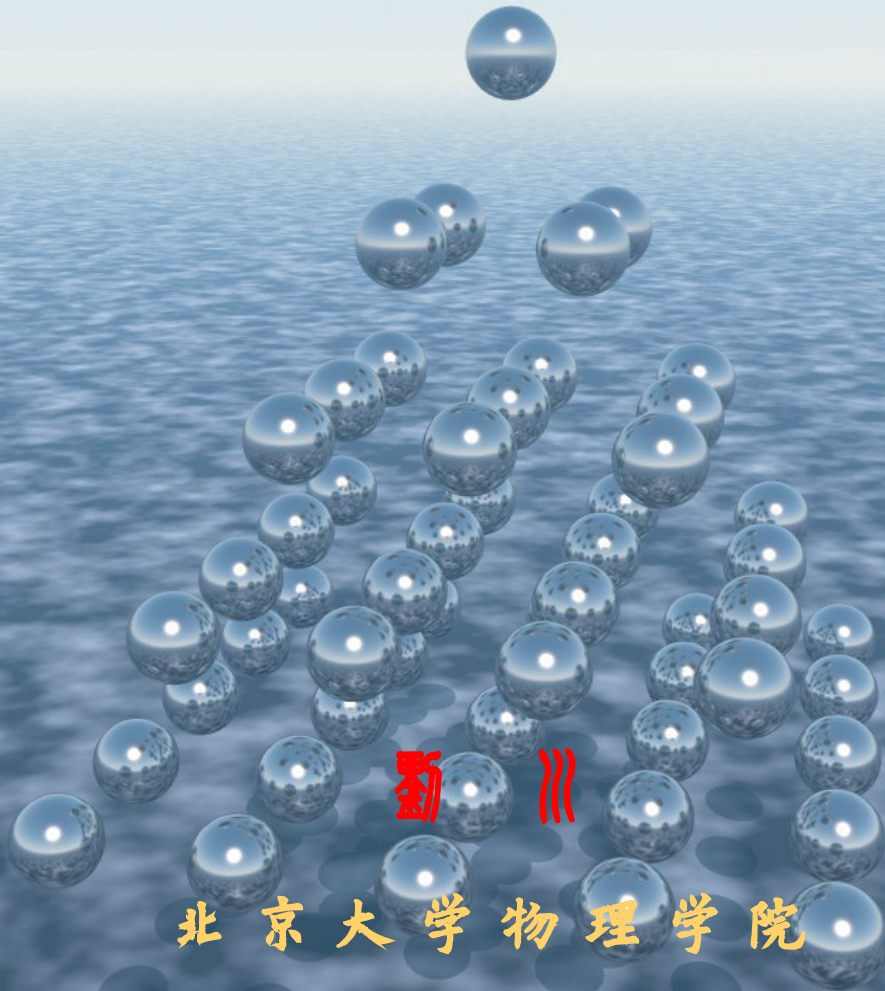


平衡态统计物理



VERSION 2.033

开篇说明

“不是我不明白，这世界变化快。”

这是崔健的歌词里面的两句。这种感叹不幸就发生在我热力学统计物理的讲稿上。从上世纪开始，我开始准备原物理系热力学统计物理课程的讲稿，前后横跨世纪。想当年曾宓芹先生报闻十载、增删五次已被传为佳话，我虽没有十载，但增删绝对大于五次。所以，还是投入了不少劳动。新世纪，新形象。2003年起，北大物理学院决定对原先的课程体系做比较大的改动。于是，我原先的热力学统计物理讲稿也必须重新安排，以便能扣上“与时俱进”的大帽子。于是，以我原先的《热力学与统计物理》课程的讲义草稿为蓝本，炮制了目前这个所谓《平衡态统计物理》的讲义。

这个讲义最为主要的调整是基本略去了传统的热力学部分，因为这些内容往往在普通物理热学中都有所涉及。另外，本课程中完全略去了非平衡态统计的部分。所有上述略去的内容，希望可以有的选修课（例如：热力学、非平衡态统计物理）来弥补。本课程主要讲述平衡态统计，同时将相应的热力学内容融会在其中，但是课程的主要方法是统计物理的方法。本讲义的第一章对于热力学有一个浓缩的总结，对于热力学方法特别有兴趣的同学建议参考其他专门的热力学书籍以扩充这些讨论。在课程的实际讲授的过程中，教师也可以适当进行调整。

本课程讲义的另外一个比较重大的改动是关于近独立子系统计的内容。在汪志诚先生的书中，这部分内容实际上是统计物理课程的开始，同时也是主体内容。我们现在改成从普遍的系综理论开始讲述，近独立子系的统计方法只是作为系综理论的一个特例来加以介绍。这种讲法是一个尝试，实际上以前也曾经尝试过。不过由于种种原因后来又放弃了。这次我们再进行一次尝试。希望能够成功。另外，增加的内容主要包括：液体的统计理论的介绍性讨论，这在以前的统计物理课程中往往一点也不涉及。考虑到软凝聚态物理的发展，我觉得适当增加一些基础的介绍是必要的；稀薄等离子体的平均场（Debye-Hückle）理论的介绍，因为等离子体和溶液都是其他物理分支或化学中经常遇到的体系。另外，最为主要的增加内容是关于自旋模型及其相变的统计理论（主要是平均场近似下的理论）。关于临界现象的统计理论是近代统计物理的重要组成部分，因此我认为有必要提前介绍。通过这些增加的内容，我试图加深同学们一个印象，那就是系综理论是统计物理的普遍理论，它可以用来处理各种各样的统计系统，特别是有相互作用的系统。这也是我选择从普遍的系综理论开始讲述统计物理的主要初衷。

在每一章的后面，我加入了一些习题。这些习题可以大致分为两类。一类是与课程中的内容的紧密联系；另一类则可能是课程中内容的推广或正文中没有讲述而又比较有意

义的问题。第一类的习题的难度适中，虽然不能完全说“容易”，但决不会难到无法完成的地步。我特别强调，第二类的习题中的一部分可能是相当难的，有些甚至不是简单地推导就可以得到结论的。比如，有些习题需要参考其他一些书籍或文献，有些甚至需要进行相当的数值计算。这一类的习题仅仅是作为例子，供有兴趣的同学在课外进行一些尝试。我一般并不要求所有同学都完成所有的习题。

原先《热力学统计物理》讲义中关于热力学和统计物理的简单历史已经重写并加以保留。这些内容可以使同学们了解这门学科的产生、发展的过程，从这些重要的历史发展中可以看到科学内在的一些规律。这一点其实很有必要，特别是考虑到现在的教科书或讲义（包括本讲义）都不可能按照原先历史的顺序来讲述。

本课程的光修课程是普物热学的 A 类课程，我们假设大家已经熟悉热力学的基本方法以及近独立子系中 Maxwell-Boltzmann 统计的基本内容，不管这些内容是通过课程或是自己学习到的。我们并不假设同学先修过量子力学课程，但是基本的量子的观念（例如：粒子的波粒二象性、物理量的量子化或者说分立性、自旋的概念等等）是重要的。同样重要的是分析力学中的一些基本概念（例如：广义坐标、广义动量、正则方程等等），这在引入系综时是十分有帮助的，但并不是必不可少的。

这个讲义可以供本课程的同学、教师免费下载。作者本人（也就是刘川！）对于讲义中的所有内容保有版权。未经作者本人许可，不允许将本讲义的全部或部分翻印后以任何方式销售给任何人；原则上也不允许将其外传给非本课程的学生和教师；不允许将本讲义的部分或全部放在网络上。一经发现，我保留诉诸法律的权利。

显然，本讲义是一个草稿。其中不乏各种错误，有些错误已经在课程的授课过程中改正了，但还有一些没有改正，欢迎大家指正。另外，讲义中的别字更是数不胜数（感谢计算机，它不会写错字）。对于优美的祖国语言的这种摧残，我本人深表歉意，力争在下次修改时尽量更正。所以，这个讲义的版本在不断地更新中，我无法在每次更改时都通知读者，对此我深表歉意。

本文利用中文 L^AT_EX 进行排版，讲义的源文件恕不予公布。短期内，作者还没有将本讲义正式出版的企图，主要是因为这个讲义还十分的简略，而且漏洞百出。考虑到正式出版会浪费大量宝贵的纸张和消费者的金钱，破坏森林，污染环境，得不偿失。所以，我觉得还是选择更为环保的方式，让它在虚拟时空中流通吧。希望这种虚伪的谦虚也能够使我进步，在听取广泛意见的基础上不断改善，希望这个讲义能够对于读者有所帮助，这就足以告慰我四只眼睛瞪着屏幕不知疲倦地码字所付出的劳动了。

在最后一段，一般总要对一些人表示致谢。如果读者觉得有些肉麻，请在您的眼睛瞥见我致谢谁之前，赶快翻过这一页吧。我想我首先要感谢我在北大理论物理所的所有同人，特别是林宗涵先生。他在课程的安排及讲授方面对我的影响是深刻的、多方面的。我要感谢我们理论物理所的马中水教授，他同时也是本课程的主讲人之一。结合他的授课，他提出了许多对于本讲义的宝贵修改意见，使我受益匪浅。同时，他通过 email 给我的有趣的图片，使我能够在每一天都保持一个较好的心情。我还要感谢每一届选平衡态统计课程的同学，他们仔细的阅读了这个讲义并且提出了很多的修改意见。在我的家人方面，我

首先要感谢我的妻子韦丹，她对我的支持一直是我内心最大的感动。同时，也应该感谢我的儿子，感谢他在我打字时没有经常地切断我计算机的电源，从而使这份讲义得以保存下来。

刘川，二零零四年六月



¶ 2005 年暑假，根据以往授课的体会以及同事们、同学们的建议对讲义进行了一些文字上的修改。从整体的结构来说，基本上没有大的改动，只是添加了少量内容（此版本号为：2.2）。对于各位同仁（特别是马中水教授）的建设性意见以及同学们提出的建议，在此一并致谢。

刘川，二零零五年八月



¶ 2006 年暑假，为了响应党中央关于建设节约型社会的号召，保护宝贵的森林资源，在这一版本（此版本号为：2.3）中加上了 `hyperref package`。它使得所有的交叉引用可以链接起来。这样一来，在电脑上阅读本讲义将更为方便（特别推荐使用 `Acrobat Reader` 并打开其书签功能）。因此，我诚恳地希望大家不要将这个讲义打印出来（没有电脑的同学除外，可是，没有电脑你又怎么打印呢??）。另外，根据以往教学中的发现，纠正了一些打错的公式。

刘川，二零零六年八月



¶ 根据教学中的发现和同学们的提醒，纠正了一些打错的公式（此版本号为：2.31）。特别要感谢物理学院 05 级的王朋博同学，他发来了原讲义中的一系列打印错误。

刘川，二零零七年一月



¶ 时隔七年年后又一次更改平衡态统计物理的讲义。添加了新的一小节（第 17.3 小节），讨论强简并电子气的磁性质。借更新的 `LATEX` 的版本之机，用 `XYLATEX` 重排本讲义并更正了一些错漏之处。目前版本号为 2.32。

刘川，二零一四年一月



¶ 感谢物理学院 12 级的郑鸣同学指出讲义中的一些错误与疏漏之处，我按照我的理解进行了修订。另外，第五章中进行了比较多的修订：第 28 节中加入了一维 `Ising` 模型的严格解的讨论；第 29 节中加入了有关 `Goldstone` 模式、`Mermin-Wagner` 定理等的讨论。目前版本号为 2.33。

刘川，二零一四年七月



目录

第一章	热力学基本规律回顾	1
1	热力学体系及平衡态的描述	1
2	热力学基本定律	3
3	热力学势、麦克斯韦关系和响应函数	7
4	相变热力学	10
4.1	单元复相系的热力学平衡条件与稳定条件	10
4.2	单元复相系中的相变	12
5	多元系的热力学	14
第二章	统计系综	19
6	经典系综理论的基本概念	20
7	量子统计与经典统计	23
7.1	量子力学对系统的描述及其对统计的影响	23
7.2	量子统计向经典统计的过渡	27
8	微正则系综	28
9	正则系综与巨正则系综	31
10	热力学公式	33
11	涨落的准热力学理论	36
12	近独立子系的统计分布	38
13	近独立子系中粒子数分布的涨落	43
第三章	量子理想气体	46
14	理想玻色气体	46
14.1	弱简并的非相对论性理想玻色气体	47
14.2	玻色爱因斯坦凝聚	49

15	黑体辐射	52
16	固体的热容量	56
17	理想费米气体	60
17.1	弱简并的理想费米气体	60
17.2	强简并的理想费米气体	61
17.3	磁场中的强简并理想费米气体	65
第四章	经典流体	69
18	经典理想气体	70
18.1	单原子分子理想气体	70
18.2	双原子分子理想气体及其热容量	71
19	混合理想气体及其化学反应	76
20	实际气体的状态方程	79
20.1	迈尔's集团展开理论	79
20.2	维里系数的计算	83
21	液体的热力学性质	84
21.1	对分布函数	84
21.2	流体中对分布函数的计算	87
22	稀薄等离子体的统计性质	88
第五章	二级相变的平均场理论	91
23	自旋模型的微观机制与处理方法	92
24	Ising 模型的平均场近似	97
25	相变的朗道理论	101
25.1	二级相变的朗道理论	101
25.2	一级相变的朗道理论	102
25.3	三临界点的朗道理论	103
26	Bragg-Williams 近似: 再看平均场近似	105
27	临界点附近的涨落与关联	106
28	易兴模型的严格解、高温展开和对偶性	110
28.1	一维易兴模型的严格解	110
28.2	二维易兴模型的高温展开与对偶性	113
29	具有连续对称性系统的相变	116

29.1	海森堡模型 Weiss 分子场理论	116
29.2	对称性的考虑与 Goldstone 模式	117
30	对称性破缺与临界现象中的普适性	120
附录 A 概率与随机过程		124
1	概率、随机变量与分布函数	124
2	常见的概率分布函数	125
2.1	二项式分布	125
2.2	Poisson 分布	125
2.3	高斯正态分布	126
3	随机过程	126
4	Markov 过程	127
5	Fokker-Planck 方程	127

第一章 热力学基本规律回顾

本章提要

- 热力学基本概念（1）和功
- 热力学第一、第二定律（2）
- 热力学势和响应函数（3）
- 热力学体系的平衡和稳定条件（4.1）
- 相变的热力学（4.2）
- 多元系的热力学（5）

1 热力学体系及平衡态的描写

热力学体系（或简称为**体系**或**系统**）是指一个宏观的系统，它一般由大量的微观粒子组成。在没有外界影响的条件下，若体系的各部分的宏观性质在长时间内不发生变化，我们称体系处于**平衡态**。热力学系统的平衡态一般可以用一些宏观（经典）的变数来加以描述，这些变数一般包括：**几何变数**（长度、面积、体积、形变等），**力学变数**（力、压强、胁强等），**电磁变数**（电场强度、电极化强度、磁场强度、磁化强度等），**化学变数**（各个组元的浓度、各个相的物质的摩尔数、化学势等）等。这些变数统称为热力学体系的**态变量**（state variables）。描述一个热力学体系的所有态变量组成的参数空间称为该热力学体系的**状态空间**，或简称**态空间**。态空间中的任意一个

点都对应于该热力学体系的一个平衡态。

热力学体系中物理性质均匀的一个宏观的部分称为一个相。热力学系统可以按照相的多少分为单相系，即仅有一个相的系统（例如液态的水）和复相系，即具有多于一个相的系统（例如冰水混合物）。还可以按照化学组元的个数来分，分为单元系，即仅含有一种化学组元的系统（例如化学纯的水）和多元系，即具有多个化学组元的系统（例如盐水、空气等）。¹

热力学体系处于平衡态需要满足一些所谓平衡条件。这些条件一般包括：热平衡条件，力学平衡条件，相平衡条件，化学平衡条件等等。热平衡条件确定了温度这一物理量的存在。这个事实在公理化热力学中被神圣化，²称为热力学第零定律。热平衡的这种可传递性说明处在热平衡的体系具有某种共同的物理量，这一点直观上是很容易理解的。从这一事实出发，Carathéodory首先严格地证明了：相互处于热平衡的体系具有一个共同的物理量，并定义这个共同的物理量为体系的温度，每一个体系的温度是该体系其他态变量的函数，也就是说：温度 θ ，以及体系的其他态变量 x_1, x_2, \dots, x_n 满足一个函数关系：[2]

$$F(\theta, x_1, x_2, \dots, x_n) = 0 \quad (1.1)$$

这个关系被称为该热力学体系的状态方程，或物态方程。我们所熟知的理想气体的状态方程，范德瓦耳斯（Van der Waals）气体状态方程，铁磁体的居里-外斯（Curie-Weiss）定律等等，都是热力学体系状态方程的具体例子。

热力学体系的态变量的函数称为态函数。按照定义，一个态函数只依赖于系统所处的状态而不依赖于系统如何到达该状态的过程。³所以热力学第零定律实际上是从宏观上说明了温度这一态函数（或者说，状态方程）的存在。从微观上讲，我们知道温度实际上是构成热力学体系的微观粒子的热运动剧烈程度的体现。

过程是状态随时间的改变。如果一个过程进行的无限缓慢，以至于在过程的每一时刻，系统都处于平衡态，这样的过程称为准静态过程（Quasi-static process）。准静态过程在热力学的讨论中起着非常重要的作用，它既可以是某个具体过程的近似，也可

¹如果在有外界影响下系统的性质长时间不发生改变，我们称这样的状态为一个稳态。例如：在恒定外电场下有电流通过的导体，这时导体的宏观性质也长时间不变，但是这只是一个稳态，不是一个平衡态。

²这个称呼首先是Fowler提出的。其实热力学第零定律的提出远远晚于热力学第一和第二定律。

³一个态函数也可以被取为态变量，这主要看以哪些态变量为独立变量。例如在体系状态方程中，可以把温度表达成其它态变量的函数，也可以把压强表达成温度和其他态变量的函数。

以是为了讨论上的方便所引用的一种工具。这主要是由于在准静态过程中系统始终处于平衡态，因此一个准静态过程可以用热力学系统状态空间中的一条曲线来描写。⁴

准静态过程的一个特性是：在一个无穷小的准静态过程中，外界对系统所做的微功 $\overset{\circ}{d}W$ 可以用系统的态变量及其微分来表达。我们用了符号 $\overset{\circ}{d}$ 而不是 d 来表示一个无穷小的准静态过程中的微功是要强调这个物理量与该无穷小的准静态过程有关。它不是一个与过程无关的微分量。例如，理想气体在一个无穷小的等温过程和一个无穷小的等压过程中的微功就不相同。与过程无关的微分量一定是系统的某个态函数的微分，我们将用符号 d 来表示，例如系统体积的变化 dV 仅仅依赖于系统初始和终止状态之间的体积的差，而与系统经历何种过程完成这一变化无关。在热力学中，我们可以将系统在无穷小准静态过程中的微功表达式写成：

$$\overset{\circ}{d}W = \sum_{i=1}^r Y_i dy_i, \quad (1.2)$$

其中 y_i 称为该系统的（热力学）广义坐标而 Y_i 称为相应的（或者说与 y_i 共轭的）广义力。

2 热力学基本定律

热力学第一定律是能量守恒定律在热力学中的具体体现。热力学第一定律实际上确定了系统的一个态函数——内能——的存在。热量则是由内能的变化及功的差决定的。

定律 1 在热力学体系经历的任何无限小过程中，

$$dU = \overset{\circ}{d}Q + \overset{\circ}{d}W, \quad (1.3)$$

上面表述中 $\overset{\circ}{d}W$ 为无限小过程中外界对体系所做的功， $\overset{\circ}{d}Q$ 为过程中体系从外界吸收的热量， dU 为体系内能的变化。当该过程是准静态过程时， $\overset{\circ}{d}W$ 可以用公式 (1.2) 计算。

¶ 卡诺（Carnot）循环由一个可逆热机工作于两个恒温热源之间的两个等温和两个绝热过程组成，参见示意图 1.1(a)。在 PV 曲线上，一个卡诺循环如图 1.1 (b) 所示。经过简单的计算（参见一般普通物理热力学教科书）就可以证明，以理想气体为工作物质的卡诺热机的效率由下式给出：

$$\eta \equiv \frac{W}{Q_1} = 1 - \frac{T_2}{T_1}. \quad (1.4)$$

⁴例如，理想气体的等温过程、等容过程、绝热过程等等就可以用 PV 图中的曲线来描写。

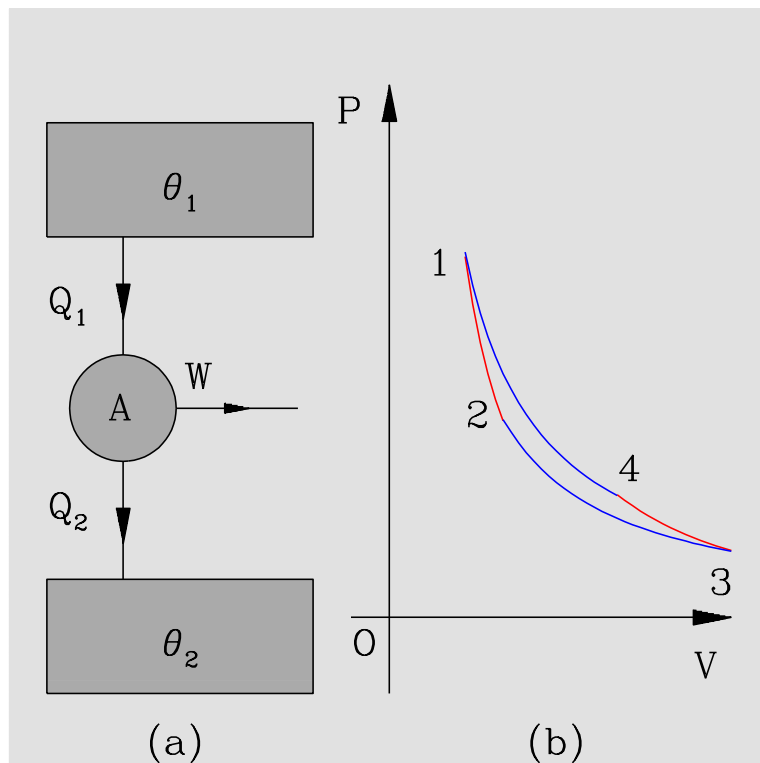


图 1.1: (a) 一个 Carnot 热机的示意图，它工作于两个恒温热源之间；(b) 卡诺循环在 PV 图上的表示，它由两条等温线和两条绝热线围成。

这个公式说明循环的效率只依赖于高温热源和低温热源的温度（这里取了理想气体温标）。这个结论实际上是更为一般的卡诺定理的特例，在下一节中，利用热力学第二定律，我们将证明该效率与工作物质也无关。历史上最先独立提出热力学第二定律的是德国物理学家克劳修斯（Rudolph Clausius）和英国物理学家开尔文（Lord Kelvin，即 William Thomson）。热力学第二定律实际上确定了体系的一个新的态函数—熵。

定律 2

- Clausius 表述：不可能把热从低温物体传到高温物体而不引起其他变化
- Kelvin 表述：不可能从单一热源吸热把它变为有用功而不产生其他影响，即第二类永动机是造不成的

¶ 可以证明，上述关于热力学第二定律的两种表述在逻辑上是完全等价的。承认了热力学第二定律，就可以来证明著名的 **卡诺定理**：

• • • •

定理 1.1 (Carnot, 1824) 所有工作于两个恒温热源之间的热机, 以可逆热机的效率为最大, 并且所有可逆热机的效率相等, 且只与两个恒温热源的温度有关, 与工作物质无关。

要证明这个定理, 我们设想有高温热源 θ_1 和低温热源 θ_2 。我们取两个卡诺热机 A 和 B , 其中 A 是可逆卡诺热机。这两个热机分别从高温热源吸热 Q_{1A} 和 Q_{1B} 而在低温热源放热 Q_{2A} 和 Q_{2B} 。它们的效率分别为:

$$\eta_A = \frac{W_A}{Q_{1A}} = 1 - \frac{Q_{2A}}{Q_{1A}}, \quad \eta_B = \frac{W_B}{Q_{1B}} = 1 - \frac{Q_{2B}}{Q_{1B}}. \quad (1.5)$$

我们现在要证明 $\eta_A \geq \eta_B$ 。

假定卡诺定理不对, 即 $\eta_A < \eta_B$, 为了方便起见, 我们假设 $Q_{1A} = Q_{1B}$ 。⁵由此可知 $W_B > W_A$ 。于是我们可以利用热机 B 输出的功来推动热机 A 来进行逆向循环。由于热机 A 是可逆热机, 它逆向循环必定从低温热源吸收热量 Q_{2A} 而在高温热源放出热量 $Q_{1A} = Q_{1B}$ 。同时, 由于 $W_B > W_A$, 所以热机 B 除了推动热机 A 逆向循环以外, 还可以净向外界输出功 $W_B - W_A > 0$ 。于是, 两个热机联合作用的结果是, 我们实现了从单一热源 θ_2 吸收热量并把它完全变成有用功, 同时没有带来其它变化, 这与热力学第二定律的开尔文表述矛盾。因此, 必定有 $\eta_A \geq \eta_B$ 。显然如果热机 B 也是可逆的, 我们可以类似地证明 $\eta_B \geq \eta_A$ 。所以, 所有可逆热机的效率必定相等, 它只与两个热源的温度有关, 与工作物质无关。这就证明了卡诺定理。

利用卡诺定理, 可以定义所谓的绝对热力学温标, 在绝对热力学温标中的温度称为绝对热力学温度, 或简称绝对温度, 一般用 T 来表示。假定一个可逆热机工作于绝对热力学温度为 T_1 和 T_2 两个热源之间, 卡诺定理可改写成:

$$\eta = 1 - \frac{T_2}{T_1}, \quad (1.6)$$

把卡诺定理的结论与上节的理想气体的卡诺循环的结论进行比较, 我们发现绝对热力学温标与理想气体温标是完全一致的。

现在我们将卡诺定理的结果推广到系统所经历的任意一个循环过程, 便可以推导出著名的克劳修斯 (Clausius) 不等式。

定理 1.2 (Clausius, 1854) 在任意循环过程中,

$$\oint \frac{dQ}{T} \leq 0, \quad (1.7)$$

其中等号仅对可逆循环过程成立。

⁵这并不影响普遍性。如果在一个循环中两个热机从高温热源的吸热不相等, 我们总可以调整两个热机做循环次数的比例使得在某一特定次循环后, 两者从高温热源吸收的热量相等

¶ 当过程是可逆的时候，克劳修斯不等式告诉我们 dQ/T 在两个态之间的积分与积分的路径（过程）无关，这说明可以定义一个态函数 **熵**（Entropie）：

$$S - S_0 = \int_{(P_0)}^{(P)} \frac{dU - dW}{T} . \quad (1.8)$$

其微分表达是：

$$TdS = dQ = dU - dW . \quad (1.9)$$

注意，虽然熵是通过一个可逆过程中 dQ/T 的积分定义的，但是它是态函数，与过程无关。⁶ 从数学意义上讲，态函数熵的存在性说明：虽然 $dU - dW$ 并不是一个全微分，但是乘上一个 **积分因子** 后就变成了全微分，这个积分因子就是 $1/T$ ，而相应的全微分就定义了态函数熵。

¶ 根据前面克劳修斯不等式的讨论，我们可以得到一个非常著名的原理——**熵增加原理**：

定理 1.3 当体系由一个平衡态经绝热过程到达另一个平衡态时，体系的熵永不减少。如果过程是可逆的，它的熵不变；如果过程是不可逆的，它的熵增加。

¶ 考虑热力学体系所经历的一个微小的可逆准静态过程，⁷ 热力学第一定律和热力学第二定律结合可以写成：

$$dU = TdS + dW = TdS + \sum_{i=1}^r Y_i dy_i , \quad (1.10)$$

其中 dU 是体系内能的微分改变， T 为体系的绝对温度， dS 为体系的熵的微分改变， dW 为外界对体系所做的功。在准静态过程中，这个功可以用体系的态变量（广义力和广义位移）表示。这就是热力学中最核心、最重要的方程，它被称为 **热力学基本微分方程**，它是平衡态热力学的基础。

¶ 下面我们对于熵的定义和热力学基本微分方程给出几点评论。

- 我们虽然是借助克劳修斯不等式中的等号情形，也就是对应于可逆过程的情形，定义了态函数熵。但是，千万不要忘记熵是一个 **态函数**。一旦系统的初态和末态给定，那么系统在两个状态间的熵的差也就完全给定了。这个熵差并不依赖于系统如何从初态变化到末态。具体到一个无穷小的过程，我们有：

$$dS \geq \frac{dQ}{T} .$$

⁶事实上，也可以通过其他的方法定义熵。有兴趣的同学可以参考王竹溪先生书中的相关讨论 [2]。

⁷在本课程中，我们将只讨论可逆过程。因此，以后提到的准静态过程都是指的可逆的准静态过程。

其中等号对应于可逆过程，大于号对应于不可逆过程。但是，在初态末态（只要都是平衡态）给定的情形下， dS 是固定的，不管这个无穷小过程是可逆还是不可逆。因此，上面这个式子的正确理解应当是：可逆过程中的 $(\delta Q)_{r.p.} = TdS$ ，而不可逆过程中的 $(\delta Q)_{i.p.} < (\delta Q)_{r.p.} = TdS$ 。也就是说，无论是可逆过程还是不可逆过程，它们的 dS 是相同的，是不依赖于过程的量。真正依赖于过程的量是 δQ ，不可逆过程中的 δQ 要小。

- 按照热力学第一定律，无论是可逆还是不可逆过程，我们都有： $\delta Q = dU - \delta W$ 。由于 dU 也是完全微分，也是不依赖于过程的量。因此，真正依赖于过程的量也可以等价地认为是 δW 。因此， $(\delta Q)_{i.p.} < (\delta Q)_{r.p.}$ 也可以表述为： $(\delta W)_{i.p.} > (\delta W)_{r.p.}$ 。
- 我们前面讨论元功的时候曾经提到过，原则上只有在无穷小的准静态过程（因而是可逆过程）中的元功可以用系统的状态参量及其微分的形式给出。一个一般的不可逆过程中的元功原则上是无法计算的，因为在过程中系统可能根本不处于平衡态，因此也就无法用系统的态参量描写。以一个简单的 PVT 系统为例，我们有： $(\delta W)_{r.p.} = -pdV$ ，和 $(\delta W)_{i.p.} > -pdV$ 。⁸
- 综合以上的讨论，对于一个 PVT 系统的无穷小过程，只要它的初态和终态都是平衡态，我们就一定有：

$$dU = TdS - pdV, \quad (1.11)$$

如果过程是一个准静态（可逆）过程， p 就是系统的压强，如果不是，那么 p 应当理解为一个“假想的”准静态过程中的压强。这就是一个 PVT 系统的热力学基本微分方程，它结合了热力学第一和第二定律。如前所述，由于不可逆过程中的功一般不能用 $-pdV$ 来表达，因此，对于一个无穷小的不可逆过程，我们不能写： $dU \leq TdS - pdV$ ；而应该写为： $dU \leq TdS + \delta W$ 。这一点在不少教材中都表述得不够清楚。

3 热力学势、麦克斯韦关系和响应函数

如果我们讨论由两个独立参数描写的简单 PVT 系统。这时热力学基本微分方程为：

$$dU = TdS - pdV. \quad (1.12)$$

⁸当然，这个时候的功可以写成 $(\delta W)_{i.p.} = -p_{\text{ext}}dV$ 其中 p_{ext} 表示外界施加的压强。但是如果系统甚至都可能不处于平衡态，压强 p_{ext} 与流体的其他状态参量可以没有任何关系。因此，从计算系统热力学性质的角度讲，这个表达式并没有什么太大的意义。

这个方程可以看成是内能以熵和体积为独立变数 (S, V) 时的标准微分表达式。因为内能是态函数，所以 dU 是一个全微分。这意味着，

$$\left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial p}{\partial S}\right)_V . \quad (1.13)$$

这种类型的关系称为 **麦克斯韦** (Maxwell) **关系**，或简称 **麦氏关系**。它体现了态函数的全微分应满足的性质。

¶ 有时有必要把热力学基本微分方程用其他独立参数表达。这可以方便地利用 **勒让德变换** (Legendre) 来实现。⁹ 比如说，如果想用 (S, p) 做独立变数，我们有：

$$d(U + pV) \equiv dH = TdS + Vdp , \quad H = U + pV , \quad (1.14)$$

其中 $H = U + pV$ 称为体系的 **焓**。类似地，我们还可以利用勒让德变换，将热力学基本微分方程换成以 (T, V) 或 (T, p) 为独立变数的微分方程：

$$d(U - TS) \equiv dF = -SdT - pdV , \quad F \equiv U - TS , \quad (1.15)$$

$$d(U + pV - TS) \equiv dG = -SdT + VdP , \quad G \equiv U - TS + pV . \quad (1.16)$$

这里我们引入了另外两个态函数，即体系的 **(亥姆霍兹) 自由能** (Helmholtz free energy) $F = U - TS$ 和体系的 **吉布斯函数** (Gibbs free energy) $G = U + pV - TS$ 。这些热力学函数被统称为 **热力学势**。由于这些热力学势全都是态函数，我们可以得到一组 Maxwell 关系：

$$\left(\frac{\partial T}{\partial p}\right)_S = \left(\frac{\partial V}{\partial S}\right)_p , \quad (1.17)$$

$$\left(\frac{\partial p}{\partial T}\right)_V = \left(\frac{\partial S}{\partial V}\right)_T , \quad (1.18)$$

$$\left(\frac{\partial V}{\partial T}\right)_p = -\left(\frac{\partial S}{\partial p}\right)_T . \quad (1.19)$$

Maxwell 关系的重要意义在于，它能够把一些不容易测量的偏微商与一些可以直接测量的偏微商（与物态方程和热容量联系着的偏微商）联系起来。

⁹熟悉分析力学的读者应当记得：当我们从一个力学体系的拉各朗日量变换到哈密顿量时，采用的就是勒让德变换。

¶ 前面的讨论仅仅涉及到封闭的（粒子数不变的）热力学系统。如果一个单元系由于相变，它的粒子数目（摩尔数）可以发生变化，它的热力学基本微分方程应当改为：

$$dU = TdS - pdV + \mu dn, \quad (1.20)$$

其中 $\mu = (U - Ts + pV)/n = G/n$ 是一摩尔物质的吉布斯自由能。对于一个单元系，一摩尔该物质的吉布斯自由能又被称为 **化学势**。这个名称的由来我们在后面讨论相变（第 4.1 节）和化学反应（第 5 节）时就会明白。

¶ 利用热力学研究一个系统的性质时，往往需要考察系统对于外界条件改变所做出的 **响应**。这些外界条件可以是温度、压强、化学组分、电场、磁场等等。热力学体系对于外界响应可以用响应函数来描写。最为常见的响应函数包括：热容量、压缩系数、电极化率、磁化率等等。

热容量 的定义是指热力学体系在某一特定过程中（因此它的数值与过程有关）升高单位温度所需要吸收的热量：

$$C_y = \lim_{\Delta T \rightarrow 0} \frac{Q_y}{\Delta T}, \quad (1.21)$$

其中 Q_y 是体系保持某一参量 y （例如压强、体积、磁场、电场等等）不变，而温度改变 ΔT 时所吸收的热量。最常用的是定容（体积不变）热容量和定压（压强不变）热容量。对于一个简单的 PVT 系统，当体积不变时，外界所做功为零（假定体系没有电磁功等的贡献），所以定容热容量 C_V 为：

$$C_V = \left(\frac{\partial U}{\partial \theta} \right)_V, \quad (1.22)$$

当压强不变时，外界所做功为 $-p\Delta V$ ，所以定压热容量 C_p 为：

$$C_p = \left(\frac{\partial H}{\partial \theta} \right)_p, \quad (1.23)$$

其中函数 H 为系统的焓。热容量实际上体现了系统的热力学势（内能、焓等）对温度的响应。

简单 PVT 系统中另一个重要的响应函数是 **压缩率**。等温压缩率 κ_T 和绝热压缩率 κ_S 的定义是：

$$\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_T. \quad (1.24)$$

它体现了 PVT 系统的体积对压强的响应。

¶ 利用麦克斯韦关系，我们可以在不同的响应函数之间建立一些普遍的联系，但是在热力学范畴内最基本的响应函数（热容量、压缩率或者说物态方程）需要通过实验来确定。

4 相变热力学

这一节中，我们简要讨论一下单元复相系中的相变热力学。我们将首先讨论单元复相系中各个相达成平衡的基本判据；随后我们讨论单元复相系中相变的基本特征。

4.1 单元复相系的热力学平衡条件与稳定条件

热力学体系的平衡是一种热动平衡。第2节中的熵增加原理告诉我们，当一个封闭热力学体系达到热动平衡时，它的熵达到其极大值。利用这个结论，可以得到系统达到热动平衡的熵判据：以一个简单的PVT系统为例，熵判据要求体系在其内能和总体积不变时，对于各种可能的虚变动来说，平衡态的熵最大。所谓虚变动是指描述系统的某些态变量的假想变动，这个概念是力学中虚位移的推广。为了区别系统的一个虚变动和一个真实的变动，我们用符号 δ 来标记某个热力学量的无穷小虚变动（或者称为变分），以区别于用 d 代表的一个真实的无穷小变动。

在数学上，熵判据可以表达为：在固定的内能和体积时，体系熵的一级变分为零，这称为热力学体系的平衡条件；同时，体系熵的二级变分大于零，这称为热力学体系的稳定条件。

一般的热力学平衡条件可以大致分为三类：热平衡条件、力学平衡条件、相平衡条件或化学平衡条件。下面我们利用平衡的熵判据，首先讨论一下一个单元复相系中达到平衡的条件。这些条件包括上述四类平衡条件中的前三类。至于化学平衡条件我们会在第5节中讨论。

我们假设一个单元复相系的几个相构成一个闭合系统。系统的总内能 $U = \sum_{\alpha} n_{\alpha} u_{\alpha}$ 、总体积 $V = \sum_{\alpha} n_{\alpha} v_{\alpha}$ 和总摩尔数 $n = \sum_{\alpha} n_{\alpha}$ 是固定的。于是，我们利用拉格朗日乘子法，考虑到上述约束条件，我们得到熵的一级变分为零的条件：

$$\bar{\delta}S \equiv \delta S - \frac{1}{T} \delta U - \frac{p}{T} \delta V + \frac{\mu}{T} \delta n = 0, \quad (1.25)$$

其中 T 、 (p/T) 和 (μ/T) 是引入的拉格朗日乘子； $\bar{\delta}S$ 只是包含了拉格朗日乘子和约束条件后熵的变分的一个简写。现在我们利用：

$$\delta s_{\alpha} = \frac{\delta u_{\alpha}}{T_{\alpha}} + \frac{p_{\alpha}}{T_{\alpha}} \delta v_{\alpha}, \quad (1.26)$$

其中 T_{α} 和 p_{α} 分别为 α 相的温度和压强。我们就得到体系的总的熵的一级变分为：

$$\begin{aligned} \delta S - \frac{1}{T} \delta U - \frac{p}{T} \delta V + \frac{\mu}{T} \delta n &= \sum_{\alpha} n_{\alpha} \left(\frac{1}{T_{\alpha}} - \frac{1}{T} \right) \delta u_{\alpha} + \sum_{\alpha} n_{\alpha} \left(\frac{p_{\alpha}}{T_{\alpha}} - \frac{p}{T} \right) \delta v_{\alpha} \\ &+ \sum_{\alpha} \left(s_{\alpha} - \frac{u_{\alpha}}{T} - \frac{p v_{\alpha}}{T} + \frac{\mu}{T} \right) \delta n_{\alpha}. \end{aligned} \quad (1.27)$$

由于引入了拉格朗日乘子, δu_α 、 δv_α 和 δn_α 现在可以看成独立的变量。于是, 利用系统达到平衡时一级变分为零 (熵判据中的平衡条件) 我们有:

$$\begin{aligned} \frac{1}{T_\alpha} - \frac{1}{T} &= 0, \\ \frac{p_\alpha}{T_\alpha} - \frac{p}{T} &= 0, \\ s_\alpha - \frac{u_\alpha}{T} - \frac{pv_\alpha}{T} + \frac{\mu}{T} &= 0. \end{aligned} \quad (1.28)$$

上式中第一个条件为热平衡条件, 它指出在单元复相系中达到平衡时, 各相的温度相等; 第二个条件是力学平衡条件, 它指出平衡时各相的压强相等; 第三个条件是相平衡条件, 它指出相平衡时, 共存各相的化学势相等。以上平衡条件可以统一写成:

$$T_\alpha = T, \quad p_\alpha = p, \quad \mu_\alpha = \mu. \quad (1.29)$$

如果上述平衡条件不能满足, 那么体系就会发生一个 **真实的** 变动 (而不再是虚变动), 这个真实的变动的方向一定是使体系的总的熵的变化大于零。例如, 如果化学平衡条件没有满足 (而其他两个条件得到满足), 那么体系会向 $\sum_\alpha (-\frac{\mu_\alpha}{T_\alpha} + \frac{\mu}{T})\delta n_\alpha > 0$ 的方向变动。也就是说, 如果某个 $\mu_\alpha > \mu$, 与之相应的 $\delta n_\alpha < 0$, 即该相的摩尔数会减少。也就是说, 物质会从化学势高的相经过相变到化学势低的相, 这也恰恰就是化学势这个名称的由来。

¶ 系统的熵的一级变分为零仅仅保证了熵取 **极值**, 而不一定是按照熵判据的要求是一个 **极大值**。要确保这个极值是个极大值, 就必须研究熵的二级变分, 或者说考虑熵判据中的稳定性条件。现在我们假定前面讨论的熵判据的平衡条件 (1.29) 都已经满足了。我们来进一步讨论这个平衡是稳定的条件。利用变分的语言, 体系的熵为一个稳定的极大值的条件是: 在相应的约束条件下, 熵的二级变分小于零:

$$\delta^2 S - \frac{1}{T}\delta^2 U - \frac{p}{T}\delta^2 V + \frac{\mu}{T}\delta^2 n \leq 0. \quad (1.30)$$

熵的二级变分的表达式为:

$$\delta^2 S = \sum_\alpha (n_\alpha \delta^2 s_\alpha + 2\delta n_\alpha \delta s_\alpha + s_\alpha \delta^2 n_\alpha). \quad (1.31)$$

对于内能和体积我们也有类似的表达式。同时, 将公式 (1.26) 再取一次变分可以得到:

$$\delta^2 s_\alpha = \frac{\delta^2 u_\alpha}{T_\alpha} + \frac{p_\alpha}{T_\alpha} \delta^2 v_\alpha + \frac{1}{T_\alpha} (\delta p_\alpha \delta v_\alpha - \delta T_\alpha \delta s_\alpha). \quad (1.32)$$

将这些表达式带入公式 (1.30), 并且利用平衡条件直接进行计算, 我们可以得到:

$$\delta^2 S - \frac{1}{T}\delta^2 U - \frac{p}{T}\delta^2 V + \frac{\mu}{T}\delta^2 n = \sum_\alpha \frac{n_\alpha}{T_\alpha} (\delta p_\alpha \delta v_\alpha - \delta T_\alpha \delta s_\alpha). \quad (1.33)$$

平衡稳定性要求此式对于任意的 n_α 都要小或等于零，因此我们得到对于每一个均匀相（我们略去角标 α ）都有：

$$\delta p \delta v - \delta T \delta s \leq 0 . \quad (1.34)$$

这就是我们要求的、系统达到平衡的稳定性条件。

¶ 利用平衡的稳定性条件 (1.34)，我们可以得到一系列的 **热力学不等式**。例如，如果我们选取 T 和 v 为独立变量，我们可以把 δp 和 δs 表示为 T 和 v 的变分，再利用麦氏关系我们得到：

$$\frac{c_v}{T} (\delta T)^2 - \left(\frac{\partial p}{\partial v} \right)_T (\delta v)^2 > 0 . \quad (1.35)$$

因此我们看到，平衡稳定性要求：

$$c_v > 0 , \quad - \left(\frac{\partial p}{\partial v} \right)_T > 0 , \quad (1.36)$$

即定容比热和等温压缩系数都是正的，这些正好符合我们物理上的直觉。

4.2 单元复相系中的相变

¶ 考虑一个简单的 PVT 系统。当两相（我们把他们分别记为 α 相和 β 相）达到平衡时，根据平衡条件我们有：

$$T_\alpha = T_\beta = T , \quad p_\alpha = p_\beta = p , \quad \mu_\alpha(T, p) = \mu_\beta(T, p) . \quad (1.37)$$

原则上，上式确定了两相平衡时 (T, p) 平面上的一条曲线，它把 (T, p) 平面分成两个区域，分别对应于两相。在 (T, p) 平面上这样的图被称为 **相图**。我们 **日常** 见到的物质的相往往是比较有限的。它们大致可以分为 **气相**（或称 **非凝聚相**）和 **凝聚相**（一般包括 **液相** 和 **固相**，其中液相又可以称为 **软凝聚相**）。气相只可能有一个相，固相一般可以有多个相存在，它们往往对应于不同的晶格结构或对称性。液相 **通常** 情况下也只有一个相，但有些特殊的物质可以有多个液相（比如液氦、液晶）。我们在图 1.2 中画出了一个典型的相图，图中的三条曲线将 (T, p) 平面分为三个区域，分别对应于固相、液相和气相。

¶ 利用热力学理论，可以对相变曲线的斜率做出预言，这就是著名的 **克拉珀龙**（Clapeyron）**方程**。假定 (T, p) 和 $(T + dT, p + dp)$ 为相变曲线上临近的两个点，我们有：

$$d\mu_\alpha = d\mu_\beta . \quad (1.38)$$

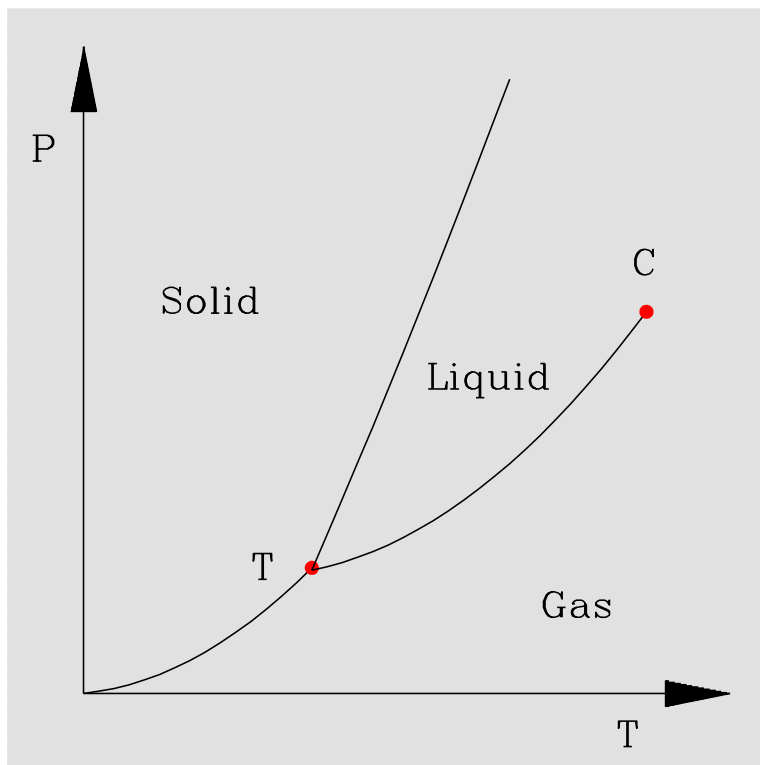


图 1.2: 一个典型的相图，由固相、液相和气相组成。

对于两相分别利用热力学关系:

$$d\mu = -s dT + v dp, \quad (1.39)$$

其中 s 和 v 分别为摩尔熵和摩尔体积，带入上式我们得到:

$$\frac{dp}{dT} = \frac{s_\beta - s_\alpha}{v_\beta - v_\alpha} = \frac{L}{T(v_\beta - v_\alpha)}. \quad (1.40)$$

这就是所谓的 **克拉珀龙方程**。它是克拉珀龙首先得到的，只是他当时用的是错误的热质说。克劳修斯首先运用正确的热力学理论导出了这个方程，因此也有人称之为克拉珀龙-克劳修斯方程。

¶ 在克拉珀龙方程中，我们引入了 **相变潜热**：

• • • • •

$$L \equiv T(s_\beta - s_\alpha) = h_\beta - h_\alpha . \quad (1.41)$$

由于两相的熵一般并不相等，所以一般说来相变潜热不等于零，也就是说，体系在发生相变时，它的熵和体积会有一个跃变。这样的相变被称为一级相变。注意到每摩尔的吉布斯函数（即化学势）在相变时是连续的（即相平衡条件），熵和体积是吉布斯函数的一阶导数，我们可以说：一级相变是吉布斯函数连续而它的一阶导数不连续的相变。在一级相变的相变曲线上，可能存在这样的点它们被称为临界点，当相变曲线上的点趋于这个点时，一级相变的熵和体积跃变会逐渐趋于零，相应的相变潜热也趋于零。于是系统将经过一个所谓的高级相变。按照 Ehrenfest 的分类，吉布斯函数及其一阶导数都连续，但吉布斯函数的二阶导数不连续的相变称为二级相变；吉布斯函数及其一、二阶导数都连续，但吉布斯函数的三阶导数不连续的相变称为三级相变，如此等等。目前，实验上已经发现许多二级相变，但实际上它们与 Ehrenfest 的定义并不是十分吻合。因为在多数已经发现的二级相变中，吉布斯函数的二阶导数实际上在相变点附近是发散的。关于二级相变的理论，我们会在第 5 章中作出简要的介绍。

¶ 二级相变的一个典型的例子是铁磁相变。当外磁场为零时，一般物质的磁极化强度 M 也为零，这样的物质称为顺磁物质。但是，还有一类物质，当外磁场为零时也具有不为零的磁极化强度。这就是所谓的铁磁物质。当温度变化时，物质可以发生在顺磁物质和铁磁物质之间转换，这就是铁磁相变。具体地说，存在一个临界温度，又称为居里温度 T_c ，当温度 $T > T_c$ 时，系统处在顺磁性的相；当温度 $T < T_c$ 时，系统处在铁磁性的相。从微观上讲，铁磁相相应于组成系统的微观粒子的磁矩，由于相互之间的交换相互作用都自发地沿着同一个方向，从而使得系统即使在外磁场为零时也具有非零的磁极化强度。关于二级相变在临界点附近的统计理论，我们会在第 5 章中讨论。

5 多元系的热力学

¶ 一个体系的热力学函数可以分为广延量和强度量两类。一个热力学函数，如果在其他条件不变时，将各个组元的摩尔数（或者说粒子数）变为原先的 λ 倍，该热力学量也变为原先值的 λ 倍，那么这个热力学函数就被称为广延量。象体系的内能、体积、熵以及各种热力学势（例如：焓、自由能、自由焓等等）都是广延量。另外一些热力学量，象温度、压强、密度、化学势等，它们则保持不变，这些热力学量称为强度量。

如果我们定义下列物理量:

$$\begin{cases} v_i = \left(\frac{\partial V}{\partial n_i} \right)_{T,p,n_{j \neq i}} , \\ u_i = \left(\frac{\partial U}{\partial n_i} \right)_{T,p,n_{j \neq i}} , \\ s_i = \left(\frac{\partial S}{\partial n_i} \right)_{T,p,n_{j \neq i}} . \end{cases} \quad (1.42)$$

其中偏导数 $\left(\frac{\partial V}{\partial n_i} \right)_{T,p,n_{j \neq i}}$ 中的下标 $n_{j \neq i}$ 代表除了 n_i 以外的所有其他的 n_j 不变。这些物理量分别称为第 i 个组元的 **偏摩尔体积**、**偏摩尔内能** 和 **偏摩尔熵**。于是, 利用广延量的性质, 体系的体积、内能和熵可以写成:

$$\begin{cases} V = \sum_{i=1}^k n_i v_i , \\ U = \sum_{i=1}^k n_i u_i , \\ S = \sum_{i=1}^k n_i s_i . \end{cases} \quad (1.43)$$

由于吉布斯函数也是广延量, 我们也有:

$$G = \sum_{i=1}^k n_i \mu_i , \quad (1.44)$$

其中 μ_i 是第 i 个组元的 **偏摩尔吉布斯函数**:

$$\mu_i = \left(\frac{\partial G}{\partial n_i} \right)_{T,p,n_j} . \quad (1.45)$$

它也称为第 i 个组元的 **化学势**。

¶ 对吉布斯函数 $G(T, p, n_1, n_2, \dots, n_k)$ 取全微分, 我们得到:

$$dG = -SdT + Vdp + \sum_{i=1}^k \mu_i dn_i . \quad (1.46)$$

类似地, 对于内能我们有:

$$dU = TdS - pdV + \sum_{i=1}^k \mu_i dn_i . \quad (1.47)$$

这就是多元均匀系的热力学基本微分方程。对公式 (1.44) 求微分并与式 (1.46) 比较, 我们得到:

$$SdT - Vdp + \sum_{i=1}^k n_i d\mu_i = 0, \quad (1.48)$$

即 $k+2$ 个强度量之间有一个关系, 这称为 **吉布斯关系**。

¶ 现在我们可以利用 4.1 节中引入的吉布斯函数判据来得到多元系的复相平衡条件。为了简单起见, 我们暂时假设多元系的 k 个组元之间不会发生化学反应, 并且每一个组元可以有两个相: α 相和 β 相。我们假设第 i 个组元在 α 相和 β 相中的摩尔数发生的虚变动分别为 δn_i^α 和 δn_i^β 。各个组元总的摩尔数不变要求:

$$\delta n_i^\alpha + \delta n_i^\beta = 0, \quad (i = 1, 2, \dots, k). \quad (1.49)$$

在温度和压强保持固定时, 两相的吉布斯函数的虚变化为:

$$\delta G^\alpha = \sum_{i=1}^k \mu_i^\alpha \delta n_i^\alpha, \quad \delta G^\beta = \sum_{i=1}^k \mu_i^\beta \delta n_i^\beta. \quad (1.50)$$

于是总的吉布斯函数的虚变动为 (考虑到约束条件 (1.49)):

$$\delta G = \delta G^\alpha + \delta G^\beta = \sum_{i=1}^k (\mu_i^\alpha - \mu_i^\beta) \delta n_i^\alpha. \quad (1.51)$$

吉布斯函数判据告诉我们, 平衡态的吉布斯函数最小。于是我们得到平衡条件:

$$\mu_i^\alpha = \mu_i^\beta, \quad (i = 1, 2, \dots, k), \quad (1.52)$$

也就是说, 两相中各个组元的化学势必须分别相等。

¶ 下面我们讨论体系中各个组元之间可以发生化学反应的情况。由于对于任何一个化学反应总可以把 **反应物** 移到 **生成物** 的一边并且改变化学反应式前系数的符号, 所以一个普遍的化学反应可以写成:

$$\sum_i \nu_i A_i = 0, \quad (1.53)$$

其中 ν_i 为化学反应方程式的系数。按照我们的约定: 生成物的系数为正, 反应物的系数为负。 A_i 是相应于生成物和反应物的化学符号。当化学反应进行时, 各个组元的摩尔数的变化都按比例进行, 这称为 **道尔顿定律**。即 $\Delta n_i = \nu_i \Delta n$ 。显然当 $\Delta n > 0$ 时, 化学反应向正向进行; 当 $\Delta n < 0$ 时, 化学反应向反向进行。

历史 1.1 道尔顿在历史上最大的贡献是在化学方面。他被认为是第一个从实验上证实了化学反应中各个物质总是按照一定的比例进行反应的。这实际上成为物质是由原子或分子组成的间接的证据，在物理、化学乃至整个科学发展史中都具有十分重要的地位。也许有些令人啼笑皆非的是，道尔顿的发现实际上具有某种虚假的成份。在道尔顿的年代，进行化学实验的仪器与设备还十分简陋。到了二十世纪，有多组对于科学史有兴趣的科学家尝试着按照道尔顿当年的记载，运用当时的仪器来重复道尔顿的实验。他们得到的结论是：以道尔顿当时的条件，他决不可能做出如此精确的实验。因此他们认为，几乎可以肯定的是，当时道尔顿实际上**人为地**改造了实验数据来为他的结论辩护，尽管他的结论仍然是具有划时代意义的。究竟是什么使得他坚信化学反应应当按照一定比例进行呢？我想这大约是他对于古希腊原子说和简单和谐比例的宗教般的信心吧。

¶ 一个化学反应的进行往往伴随着热量的吸收或放出。为了简化讨论，我们将主要研究在等温等压条件下进行的化学反应。我们知道，在等温等压条件下，一个过程中所吸收的热量等于过程前后焓的改变。在化学上，把 ν_i 摩尔数生成物产生后所吸收的热量称为（定压）**反应热**。按照定义，定压反应热 Q_p 为：

$$Q_p \equiv \Delta H = \sum_i \nu_i h_i, \quad (1.54)$$

其中 h_i 为组元 i 的偏摩尔焓。显然，按照我们的约定， $\Delta H > 0$ 对应于吸热反应； $\Delta H < 0$ 对应于放热反应。由于焓是态函数，因此 ΔH 只与反应的初态和末态有关而与中间的过程无关。所以，如果最初的反应物先经过一个化学反应到达一个中间态，再经过另一个化学反应到达终态；这个过程**的反应热一定是上述两个化学反应的反应热之和**。这在热化学上称为**赫斯**（Hess, 1840）**定律**。反应热是温度的函数，如果保持压强不变，对温度求偏微商：

$$\left(\frac{\partial \Delta H}{\partial T}\right)_p = \sum_i \nu_i c_{p,i}. \quad (1.55)$$

这个方程称为**基尔霍夫**（Kirchhoff）**方程**。它告诉我们，一旦知道了组元的热容量和某个温度时的反应热，就可以确定其他温度时的反应热。

¶ 当化学反应在等温等压条件下进行时，利用吉布斯函数判据，我们可以得出化学反应达到平衡的条件。为此，我们假设有一个虚变动 $\delta n_i = \nu_i \delta n$ ，那么在这个虚变动中，系统吉布斯函数的变化为：

$$\delta G = \sum_i \mu_i \delta n_i = \delta n \sum_i \nu_i \mu_i. \quad (1.56)$$

于是我们看到，化学反应达到平衡的**化学平衡条件**是：

$$\sum_i \nu_i \mu_i = 0 . \quad (1.57)$$

如果化学平衡条件 (1.57) 没有满足, 化学反应就会向使吉布斯函数减小的方向进行。具体地说, 按照我们对于各个 ν_i 符号的约定, 如果 $\sum_i \nu_i \mu_i < 0$, 化学反应将向正向进行; 如果 $\sum_i \nu_i \mu_i > 0$, 化学反应将向反向进行。因此, 只要求出多元系中各个组元的化学势 μ_i , 利用公式 (1.57) 就可以得到系统化学平衡时满足的条件。



相关的阅读

本章是热力学基本理论的一个简单的叙述。由于许多概念在普通物理热学中都有过讨论, 所以我们这里的论述是比较简略的。如果读者希望看到更为详尽的讨论, 我推荐王竹溪先生的《热力学》[2]。如果读者觉得我们这里的论述还是过于冗长, Plischke 和 Bergerson 的第一章可以做为一个不错的浓缩版 [11]。多数英文的教科书都是将热力学揉和在统计物理中讲述, 只有一些比较古老 (或者说经典) 的书中, 例如 [7], 是仅仅讲述热力学的内容的。如果读者对于偏导数的运算不十分清楚的话, 可以参考王竹溪的《热力学》[2] 中的有关总结。

第二章 统计系综

本章提要

- 统计系综的概念 (6)
- 量子统计与经典统计 (7)
- 微正则系综 (8)
- 正则系综与巨正则系综 (9)
- 热力学公式 (10)
- 涨落的准热力学理论 (11)
- 近独立子系的统计分布 (12) 与涨落 (13)

系

综理论的建立标志着分子运动论向统计物理的过渡。它是统计物理的普遍理论，不仅适用于无相互作用的体系——这类体系在统计物理中被称为近独立子系——而且可以用来研究普遍的、有相互作用的系统。在系综理论建立的初期，人们总是试图从纯力学的原理出发来导出统计物理的系综理论。后来人们逐渐认识到：这种纯力学的做法实际上是一种误导，统计物理的系综理论应当是以统计物理所特有的基本原理——而不是纯力学的基本原理——为基础的。系综理论首先是玻耳兹曼、吉布斯等人提出的。后来逐渐被完善并可以与量子力学兼容。我们这一章将主要讲述经典的（或半经典的）统计系综理论。

6 经典系综理论的基本概念

¶ 我们首先讨论一个经典的、多自由度体系的描述方法。在经典力学的范畴内，一个由 N 个 **全同粒子** 组成的、有相互作用的经典力学系统的自由度数目 $f = Nr$ ，其中 r 是每一个粒子的自由度数目。这样一个经典系统在任意时刻的运动状态可以由该时刻的 f 个 **广义坐标** q_1, \dots, q_f ，以及与之共轭的 **广义动量** p_1, \dots, p_f 来描述。以 $q_1, \dots, q_f; p_1, \dots, p_f$ 这 $2f$ 个变量为“坐标”可以构成一个 $2f$ 维的参数空间，我们称之为系统的 **相空间**¹ 或 **Γ 空间**。系统任意时刻的运动状态可以用相空间中的一个点来描述，这称为系统运动状态的 **代表点**。当系统运动状态随时间发生改变时，其代表点就在相空间中随时间变化从而划出一条轨道，我们称之为系统的 **相轨道**。

¶ 在经典力学中，系统的运动遵从经典的 **哈密顿 正则方程**：²

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \quad (i = 1, 2, \dots, f), \quad (2.1)$$

其中 $H(q, p)$ 称为系统的 **哈密顿量**。³ 在给定初始条件下，这些方程就确定了系统的相轨道。在运动过程中，系统的哈密顿量 $H(q, p)$ 是一个守恒量：

$$H(q, p) = E, \quad (2.2)$$

其中 E 为系统的总能量。⁴ 这个方程在系统的相空间中确定了一个曲面（这是一个 $(2f - 1)$ 维的超曲面），称为该力学体系相空间中的 **能量曲面**。如果能量守恒，系统的相轨道将始终处于能量曲面上。如果系统的能量不是严格守恒，而是在 E 到 $E + \Delta E$ 之间，那么系统的相轨道将处在：

$$E \leq H(q, p) \leq E + \Delta E, \quad (2.3)$$

所确定的 **能壳** 之内。一般说来，系统有可能还有其他的守恒量，例如平动动量、角动量等等。但是如果取特殊的参照系，比如取没有整体平移和转动的系统，同时，忽略掉出现偶然简并的可能性，那么唯一的守恒量就是系统的总能量。所以，系统的能量在统计物理中起着特别重要的作用。

¹有的书中称之为 **相宇**。

²如果你还没有修炼到哈密顿力学，你可以将这套运动方程想象成力学体系的拉格朗日方程；如果你连拉格朗日方程也没有修炼过，你可以把它们想象成体系的牛顿运动方程。

³为了使符号尽量简化，我们将用 q 和 p 来分别代替 q_1, \dots, q_f 和 p_1, \dots, p_f 。所以，系统的哈密顿量 $H(q_1, \dots, q_f; p_1, \dots, p_f)$ 就简记为 $H(q, p)$ 。

⁴从数学上讲，这个方程实际上给出了哈密顿方程 (2.1) 的一个初积分。

¶ 当我们测量一个宏观系统的某个宏观物理量时，这个测量一般会持续一段时间，比如说： $t_0 < t < t_0 + \tau$ ，其中 τ 是一个宏观短而微观长的时间间隔。⁵ 所谓“宏观短”，是指在这个时间间隔内，系统的宏观物理量还没有发生任何可观测的变化；所谓“微观长”，是说从微观角度看，在该时间间隔内，系统的微观运动状态实际上已经发生了很大的变化。从系统相空间的角度看，系统的代表点在相空间中的位置已经发生了明显的变化。如果我们要测量的宏观物理量的微观对应为 $B(q, p)$ ，麦克斯韦和玻耳兹曼认为：我们对于该宏观量的测量值实际上是其相应微观量在这一段时间中的平均值：

$$\bar{B}(t_0) = \frac{1}{\tau} \int_{t_0}^{t_0+\tau} dt B(q(t), p(t)) . \quad (2.4)$$

由于测量的时间间隔 τ 对于微观来讲是足够的长，为了数学的方便，在实际的考虑中往往取 $\tau \rightarrow \infty$ 。这个表达式虽然物理意义十分明确，但在实际应用中却很难给出任何有意义的结论。原因就在于我们实际上无法得到一个宏观系统的相轨道的明确表达式，所以上面的式子基本上只能停留在定义的层面上，不可能进行真正明确的计算。

¶ 至此，我们仅仅涉及了系统的“纯力学”性质。我们发现，由于宏观系统的复杂性，我们很难从纯力学上继续对它的宏观物理量进行任何具体的计算。⁶ 为了能够进一步得到系统的宏观统计性质，玻耳兹曼意识到必须引入统计物理的假设。这个假设有多种引入的方法。其中最为著名的就是所谓的“各态历经假设” (Ergodic Hypothesis)。这个假设是：在足够长的时间内，系统的代表点将会在系统的能量曲面上的各个区域停留相同的时间。很显然，如果各态历经假设成立，那么我们可以定义系统的代表点在系统能量曲面上各点出现的几率密度 $\rho(q, p, t)$ ，它的物理意义是： $\rho(q, p, t)dqdp$ 代表在时刻 t ，相空间中的点 (q, p) 附近的相体积元 $dqdp$ 内系统代表点出现的几率。那么，我们有：

$$\bar{B}(t_0) = \int dqdp \rho(q, p, t_0) B(q, p) , \quad (2.5)$$

¶ 象公式 (2.5) 这样的平均值被称为“系综平均”。需要注意的是，随着各态历经假设的引入，我们实际上已经转换了我们考察系统的视角。具体地说，我们已经从一个纯力学的角度变为一个具有某种“统计”意义的视角。因为，当我们在说系统的代表点在相空间

⁵这些概念是麦克斯韦和玻耳兹曼等人首先提出的。玻耳兹曼是统计物理的创始人之一。他的观点一直是从纯力学的概念出发来推导统计物理的结论，他还首先提出了各态历经假设和微正则系综。

⁶由于近年来计算机的发展，人们可以对于一个自由度数目相当大的系统（尽管仍然不是无限大）进行直接具体的计算，这种方法被称为“分子动力学” (Molecular Dynamics) 方法。在这类方法中，人们真正地用数值的方法去解系统的运动方程，然后进行平均来计算系统的各种物理量。这也许是最接近 Maxwell 和 Boltzmann 原始的想法的处理。关于这一点可以参考第 21.2 节中的讨论。

出现的几率时，我们实际上已经不是在考虑一个宏观系统了，而是在考虑大量的、具有同样宏观性质的系统的集合。这种具有相同宏观性质的系统的集合就称为系综，它也可以看成是一个系统在不同时刻的代表点的集合。用概率统计的语言来说，从这样大量的系统样本中取样，就可以确定系统代表点出现的几率。不难想象，如果各态历经假设成立，那么一个宏观系统物理量的时间平均值与其系综平均值是等价的。⁷

在统计物理发展的初期，许多的物理学家和数学家都致力于各态历经理论的研究，他们试图证明各态历经假设，⁸从而将统计物理放在纯力学的基础之上。但是研究的结果却与人们的初衷并不一致。人们发现了一些可以严格证明各态历经的系统，同时也发现了一些不各态历经的系统。最为糟糕的是，能够完成严格证明的往往是一些极为简化的系统，这与统计物理所要研究的复杂宏观系统形成了鲜明对照。另一方面，如果我们撇开统计物理的纯力学基础的问题，利用统计物理理论人们可以得到十分丰富的物理结果，这些结果可以被大量的物理实验所直接证实。逐渐地，另一种观念开始占了上风。这就是以吉布斯（W. Gibbs）为代表的，我们可以称之为“纯统计”的观点。这种观点认为：既然从纯力学的角度出发来建立统计物理是困难的，不引入假设甚至几乎是不可能的，我们不如承认：统计物理有它自身的特殊规律；这种规律就是独立于纯力学规律的统计规律。具体地说，吉布斯从一开始就假设系统的宏观物理量是相应的微观物理量的系综平均值。也就是说，吉布斯假设，统计物理的出发点就是公式(2.5)，而统计物理的主要任务就是确定系综的几率分布 $\rho(q, p, t)$ 的形式。至于说系综平均值是否与时间平均值相同，这是另一个问题。它应当是力学，而不是统计物理来回答的问题。这是一种完全不同于玻耳兹曼的观念，它完全摆脱了统计物理中形而上（机械的）的因素，或者说隐去了对于纯力学的不必要的依赖。统计物理的正确性将由大量的实验来直接验证，而不是通过纯力学的推导来间接验证。

当然，我们上面的讨论并不意味着研究各态历经理论是没有意义的。恰恰相反，这个领域近年来还相当活跃。它将有助于澄清统计物理的许多基础问题，这显然是有意义的，但这只是问题的一个方面。⁹另一方面，实际上近年对于各态历经理论的研究更多的是为了人们加深对于纯力学，特别是非线性力学和混沌性的认识，而不再单单是为统计物

⁷这一点可以严格证明，当然是在严格定义了各态历经之后。

⁸我们这里仅仅提出了著名的各态历经假设。实际上，各态历经假设是统计物理假设中出现比较早、也是讨论比较广泛的一个。还存在一些其他的相关假设，例如所谓的混和假设。这个假设是说：如果在能量曲面上的一个紧致的点集，经过运动方程的演化，该点集会弥撒于整个能量曲面的各个部分，但按照分析力学中的刘维定理，其总体积并不变化。形象的说，就像是在面团上撒上一点糖，然后运动方程就像揉面过程一样，经过一段时间，起初撒的糖便弥撒到整个面团了。从现代力学的观点看，这说明系统的运动是混沌的。混合在数学上又分为强、弱两种。可以证明强混合导致弱混合，而弱混合导致各态历经，但倒过来不一定正确。

⁹这方面的阶段性成果可以从 Birkhoff 和 von Neumann 的定理中得到概括，有兴趣的同学可以参考专著 [14] 的第一章。

理的基础进行铺垫了。原因也十分简单，统计物理的正确性已经被无数次的实验所证实。人们对它的正确性已经毫不怀疑了。因此，研究从力学规律如何过渡到统计规律的过程本身变得更为重要和吸引人，而不是最终的结果。做为统计物理的课程，我们的侧重点将放在统计物理本身的理论和应用上，我们将不会过多地讨论统计物理的力学基础问题，有兴趣的同学可以去参考相应的书籍和文章。

7 量子统计与经典统计

¶ 以上讨论的统计物理的概念都是源于纯经典的系统。在这一节中，我们简单介绍一下量子力学对于统计物理的影响。¹⁰我们会在这个课程中初步体会到，量子力学对于统计物理的影响是十分重要和深远的。从量子力学出发系统的处理统计物理问题的理论必须从量子力学的密度算符（又称为密度矩阵）表述出发，这超出了本课程的范围，将会在后续课程（例如量子统计）中讨论。本课程将仅仅涉及到量子化的基本概念，不会涉及到具体的量子动力学。

7.1 量子力学对系统的描述及其对统计的影响

从原则上讲，量子力学对于统计物理的影响主要体现在两个方面：一是物理量的量子化（量子性）；二是量子的全同性原理的作用（全同性）。下面我们分别从量子性和全同性这两个方面来考察量子力学对于系统统计性质的影响。

¶ 首先，在量子力学中系统的物理量，如能量、动量、角动量、磁矩（自旋）等等，原则上讲都只能取分立的数值。因此，物理量都由其相应的量子数来描写。虽然每一个物理量都有相应的量子数，当一个力学量取定一个量子数时，它也就具有确定的数值，但并不是所有的物理量都可以同时具有确定的数值。或者说，并不是所有物理量都可以同时测量。例如，地球人都知道的量子力学中著名的海森堡不确定性关系告诉我们：量子系统的任意一个广义坐标 q_i 与它的共轭动量 p_i 是不可能同时被测准的。如果我们对于广义坐标的测量的不确定度为 Δq_i ，对于相应的广义动量的测量的不确定度为 Δp_i ，那么一定有：

$$\Delta q_i \Delta p_i \simeq h, \quad (2.6)$$

其中 h 是普朗克（Planck）常数。这就是著名的海森堡不确定性关系。也就是说，在量子力学的描述中，一个力学体系的状态不可能完全指定它的所有物理量的量子数，而只能

¹⁰我们将不假设读者已经熟悉量子力学。我们所要利用的量子力学知识将是十分初步的，即能量、动量等物理量的量子化、波粒二象性（以及相应的海森堡不确定性关系）以及全同性原理等。当然，对于量子力学的更深入的了解将会使得本课程中的某些讨论更好理解。

够同时确定其中的一部分物理量的量子数。这最大可能的、可以同时测量的、独立的物理量所对应的量子数在量子力学中称为该系统的一套 **好量子数** (即与守恒量对应的量子数)。这套独立的量子数的数目与系统的自由度数相同。这就是量子力学对于一个力学体系的描述方法。

不确定性关系 (2.6) 实际上意味着: 一个由量子力学描写的系统在任意一对共轭变量的子相空间 (也就是 q_i 和 p_i 所张成的二维空间) 中的投影实际上不能够用一个几何的点来描述 (或者说不能同时确定), 而是一个“面积”大约是 h 的小马赛克。因此, 对于一个具有 $f = Nr$ 个自由度的量子力学体系, 它的一个量子状态在其相空间中对应于一个 (超) 体积为 h^{Nr} 的小 (超) 立方体。也就是说, 按照量子力学的观点,¹¹ 每一个量子化的轨道 (或者说 **量子态**) 对应于其经典相空间 ($2f = 2Nr$ 维的) 中大小为 $\Delta q \Delta p \sim h^{Nr}$ 的一个体积元。相应的, 力学系统的整个相空间实际上被分割为每个体积为 h^{Nr} 的 **相体积元**。因此, 按照量子力学的观点, 我们 **只能说** 系统的代表点存在于相空间的某个点附近的、体积为 h^{Nr} 的小体积元中, 而不可能更精确地确定系统在这个小体积元中的位置。或者说, 任何试图更精确地确定系统代表点在相空间中位置的企图都是量子力学的基本原理所不允许的。这就是物理量的量子性对于系统统计性质的影响。

与此相应, 上一节讨论的经典系统物理量的系综平均值 (2.5) 在量子力学情形下必须用相空间中 **分立的** 求和—而不是积分—来替代:

$$\bar{B}(t_0) = \sum_{p,q} \Delta q \Delta p \rho(q, p, t_0) B(q, p) . \quad (2.7)$$

¶ 量子力学对于系统统计性质的影响除了体现在物理量的量子化上以外, 还体现在量子力学的 **全同性原理** 上。这个原理指出: 微观全同粒子是不可分辨的, 一个由全同粒子组成的系统中, 如果将两个粒子所处的量子状态对换, 系统的微观运动状态 (或者说量子态) 并不发生改变。全同性原理还决定了构成系统的粒子的统计性质可以分为两大类: **玻色子** 和 **费米子**。用量子力学的语言来说, 一个全同玻色子系统的 **波函数**, 对于交换任意两个粒子的状态是 **对称** 的; 一个全同费米子系统的波函数, 对于交换任意两个粒子的状态是 **反对称** 的。这种微观全同粒子体系在粒子交换下的对称性质又与构成体系的微观粒子的内秉角动量—**自旋**—密切相关。

量子力学的角动量与经典力学的角动量的最大区别在于, 任何一个量子力学的角动量 \mathbf{J} 的三个分量: (J_x, J_y, J_z) 中的任意两个都不可能同时被确定, 正像一个坐标与相应的动量不能同时确定一样。但是, 我们可以同时确定一个角动量平方 \mathbf{J}^2 和角动量的任意一个分量 (比如说 J_z)。量子力学的分析指出, 它们的取值都是分立的:

¹¹确切地说, 这是量子力学中准经典近似 (又被称为 WKB 近似) 的结论。

$$\mathbf{J}^2 = j(j+1)\hbar^2, \quad J_z = m\hbar, \quad (2.8)$$

这里 $j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$ 称为这个角动量的（主）量子数； $m = -j, -j+1, \dots, j$ 称为角动量分量的 **磁量子数**。因此，对于一个给定的角动量量子数 j ，相应的磁量子数可以取从 $-j$ 到 $+j$ 一共 $(2j+1)$ 个不同的值。注意，量子力学的基本关系确定了任何角动量的量子数 j 只能是整数或者是半整数。

任何微观粒子，即使它没有任何的相对于原点的运动，也都具有确定的（在量子力学意义上确定的）角动量，这被称为该微观粒子的 **自旋角动量**，或简称为 **自旋**。任何微观粒子的自旋角动量 \mathbf{S} 一定也满足：

$$\mathbf{S}^2 = s(s+1)\hbar^2, \quad S_z = m_s\hbar, \quad (2.9)$$

这里 s 称为自旋角动量的量子数，通常又简称 s 为粒子的自旋，它也只能是整数或半整数；给定了一定种类的微观粒子，它的量子数 s 就完全确定了。¹² 粒子自旋的磁量子数， $m = -s, -s+1, \dots, s$ 仍然可以取 $(2s+1)$ 个不同的值。

在量子力学中，一个微观粒子的角动量是与它的 **磁矩** 密切联系在一起的。¹³ 一般来说，如果一个微观粒子同时具有轨道角动量 \mathbf{L} 和自旋角动量 \mathbf{S} ，那么它的磁矩可以写成：

$$\boldsymbol{\mu} = \mu_B(g_s\mathbf{S} + g_l\mathbf{L}), \quad (2.10)$$

其中 $\mu_B = |e|\hbar/(2mc)$ 是所谓的 **玻尔磁子**； g_s 和 g_l 分别称为该粒子自旋角动量和轨道角动量的 **回转因子**。¹⁴ 如果我们在空间加一个均匀磁场 \mathbf{B} ，一个磁矩与磁场的相互作用能量可以写为：

$$H_{Zeeman} = -\boldsymbol{\mu} \cdot \mathbf{B} = -\mu_B(g_s\mathbf{S} + g_l\mathbf{L}) \cdot \mathbf{B}. \quad (2.11)$$

这个能量被称为 **塞曼能量**（Zeeman energy）。

我们这个课程之所以要关注微观粒子的自旋是因为一个微观粒子的自旋量子数 s 与该微观粒子的统计性质有着一一对应的关系。这个关系被称为 **自旋—统计定理**。具体来

¹²例如，对于 $s = 1/2$ 的粒子，我们通常称该粒子具有 $1/2$ 的自旋。大家熟悉的电子就是具有 $1/2$ 的自旋。

¹³这一点实际上在经典电动力学中也是如此。

¹⁴这里我们采用了高斯单位制。对于象电子这样的基本粒子： $g_s = -2$ ， $g_l = -1$ 。

说，这个定理告诉我们：自然界中的微观粒子可以按照其自旋或统计性质分为两类。一类称为**费米子**，另一类称为**玻色子**。费米子的自旋量子数 s 一定是**半整数**，它们遵从所谓的**泡利不相容原理**，即：在同一个微观量子态上只允许有至多一个相同的费米子占据。我们称这种统计法为**费米-狄喇克 (Fermi-Dirac) 统计**。全同费米子系统的波函数对于两个粒子的交换是反对称的。玻色子的自旋量子数 s 一定是**整数**，它们不受 Pauli 不相容原理的限制，在同一个微观量子态上允许占据任意多个玻色子。人们称这种统计法为**玻色-爱因斯坦 (Bose-Einstein) 统计**。¹⁵ 全同玻色子系统的波函数对于两个粒子的交换是对称的。有关微观粒子的分类和基本性质，参见表 2.1。

表 2.1: 微观粒子统计性质的分类

分类	粒子的自旋量子数	泡利不相容原理	两粒子交换时波函数性质
玻色子	整数	不遵从	对称
费米子	半整数	遵从	反对称

自然界中已发现的**基本粒子**中电子、 μ 轻子、 τ 轻子、中微子、夸克都是费米子；光子、胶子、 W^\pm 中间玻色子、 Z^0 中间玻色子等都是玻色子。其它的微观粒子都是上述粒子的复合粒子。由偶数个费米子复合而成的粒子具有整数自旋量子数，¹⁶ 是玻色子，如： π 介子、 K 介子、 ^4He 原子等。由奇数个费米子复合而成的粒子具有半整数自旋量子数，是费米子，如：质子、中子、 ^3He 原子等。

¶ 在统计物理研究的全同粒子构成的系统中，将会遇到两类情况。在一类系统中，任意两个全同粒子所对应的“粒子云”（量子力学的波函数的一种形象化的描述）可以发生重叠（例如气体），这时我们必须考虑全同性原理的影响。这类系统我们称之为**非定域系**。在另一类系统中，粒子所对应的“粒子云”不会重叠。这时，我们可以用它们的位置来区分不同的粒子（例如固体中原子在其平衡位置附近的小振动），这时我们可以不考虑全同性原理的影响。这类系统我们称之为**定域系**。

从统计物理发展的历史来看，人们首先研究的是服从经典力学的、定域系的统计，这就是著名的**麦克斯韦-玻耳兹曼 (Maxwell-Boltzmann) 统计**。麦克斯韦和玻耳兹曼

¹⁵ 自旋统计定理的理论证明需要用到相对论性量子场论的知识。目前发现的基本粒子按照自旋 S 可分为三类： $S = 0, 1/2$ 和 1 。自旋 $1/2$ 的粒子由旋量场描述，其基本产生、湮灭算符满足反对易关系（自动保证了 Pauli 不相容原理），属于费米子；自旋是 0 或 1 的粒子分别由标量场和矢量场描述，其基本产生、湮灭算符满足对易关系，属于玻色子。

¹⁶ 有时我们还将“自旋量子数”简称为“自旋”。所以，也可以说玻色子具有整数自旋；费米子具有半整数自旋。

最初的统计法忽略了构成系统的微观粒子的[○]一切量子效应：既忽略了量子的动力学效应，即用经典力学描述粒子的运动（从而所有物理量都是连续的，而不是分立的），又忽略了量子力学全同性的统计效应（如费米子、玻色子的区别）。麦克斯韦-玻耳兹曼统计又被称为[○]经典统计。与之相对的是所谓[○]量子统计。量子统计包括：[○]费米-狄喇克（Fermi-Dirac）[○]统计和[○]玻色-爱因斯坦（Bose-Einstein）[○]统计。

这里我们特别提请读者注意所谓[○]经典统计这个词语的用法。如前所说，在麦克斯韦与玻耳兹曼的年代，他们的统计法是真正意义上的经典统计，它忽略了一切量子效应（物理量的分立性和微观粒子的全同性）。我们或许应当称之为[○]纯经典统计。事实上，在目前许多的统计物理书籍中，麦克斯韦-玻耳兹曼统计除了包括[○]纯经典统计以外，还包括了不忽略物理量的分立性但是忽略了微观粒子全同性的统计方法。这种统计方法我们或许可以称之为[○]量子的麦克斯韦-玻耳兹曼统计。按照这种理解，目前的所谓经典统计，也就是（广义的）麦克斯韦-玻耳兹曼统计与量子统计的区别就在于是否考虑微观粒子的全同性。

7.2 量子统计向经典统计的过渡

由于量子力学对于统计物理的影响主要体现在物理量的[○]量子性和微观粒子的[○]全同性两个方面，因此量子统计要过渡到经典统计也需要涉及这个两方面的条件。这些条件被统称为[○]经典极限条件：

- 如果系统典型的量子力学能级间隔 ΔE 比起与温度 T 相应的热激发能 $k_B T$ 要小得多时，系统能量的分立性变得不显著，如果其他的相关的物理量的分立性也同时可以忽略，这时系统的所有相关物理量可以看成是[○]准连续的。这时，象公式 (2.7) 所给出的量子力学的系综平均完全可以相当精确地用经典力学的系综平均 (2.5) 来替代。因此，这个条件又被称为[○]准连续条件。
- 如果组成系统的微观粒子的统计性质（即是玻色子还是费米子）对于系统而言并不重要，这时无论是玻色统计还是费米统计都过渡到麦克斯韦-玻耳兹曼统计，这个条件又被称为[○]非简并条件。

对于准连续条件，由于在相邻的相体积元中系统的物理量（特别是能量）的改变很小，可以看成是[○]准连续的。因此，前面提到的由于量子性带来的相空间中的求和可以很好地用相空间中的积分来近似，其中相空间中的体积元与相应的量子态数目（或者称为简并度）有下列简单的对应关系：

$$d\Omega \equiv \prod_{i=1}^{N_r} (dq_i dp_i) \longleftrightarrow \frac{d\Omega}{h^{N_r}} \text{ 个量子态} . \quad (2.12)$$

这是我们在后面处理准连续经典极限时经常要用到的关系。它实际上给出了量子态在相空间的密度，因此被称为 **态密度**。在以后的具体应用中，我们会具体分析准连续极限的适用条件。

我们强调：只有准连续条件和非简并条件这两个条件 **同时** 得到满足时，系统的统计性质才真正能够用 **纯经典统计** 来处理，这时我们就得到 **纯经典** 的统计物理的结果；还有相当多的系统是只满足上述两个条件中的一个，这时系统的统计性质应当说是具有量子影响的，不能用纯经典的麦克斯韦-玻耳兹曼统计来处理。具体来说，如果系统满足准连续条件而不满足非简并条件，那么系统必须用玻色统计或费米统计来处理，但其物理量仍然可以看成是准连续的，也就是说，在计算物理量的系综平均值时可以用积分来代替分立的求和；¹⁷ 同样，如果系统满足非简并条件而不满足准连续条件，那么系统可以用量子的麦克斯韦-玻耳兹曼统计来处理，其中的物理量必须看成是量子化的。¹⁸ 当然，也有的系统上述两个条件都不满足，这时必须完全用量子统计的方法来处理，既需要考虑微观粒子的全同性，又需要考虑到物理量的分立性。

8 微正则系综

¶ 考虑一个 **经典的** 孤立系所对应的系综，当系统达到平衡时一切宏观物理量都将不依赖于时间，所以 **几率分布密度** 一定也是与时间无关的。系统的能量是守恒的，所以系统的代表点一定在其能量曲面上。即便如此，系统所能够取的微观态数目还是大量的。要确定孤立系统的几率密度，实际上需要一个统计物理中的基本原理，这个原理被称为 **等几率原理**。这个原理认为：对于处于平衡态的孤立系，系统各个可能的微观态出现的几率相等。因此，对于能量严格守恒的孤立系而言，我们有：

$$\rho(q, p, t) = C \delta(H(q, p) - E) , \quad (2.13)$$

也就是说，几率密度在能量曲面以外是零；在能量曲面上是一个常数。有时为了方便，我们假设系统的能量并不是等于一个恒定的常数，而是可以处在一个小的能量区间内，这时几率分布密度为：

$$\rho(q, p) = C \quad E \leq H(q, p) \leq E + \Delta E , \quad (2.14)$$

¹⁷具体的例子就是第三章中所讨论的量子理想气体（光子气体、自由电子气体、声子气体等等）

¹⁸最典型的例子是第 18.2 小节中关于双原子分子气体热容量的讨论。

而在其他能量区域 $\rho = 0$ 。与这样的分布对应的系综称为 **微正则系综**。如果可以运用准经典近似，那么在能壳 $[E, E + \Delta E]$ 之内的系统的微观状态数为：

$$\Omega(E) = \frac{1}{N! h^{Nr}} \int_{E \leq H(q,p) \leq E + \Delta E} dq dp , \quad (2.15)$$

其中我们已经考虑了粒子的全同性，如果是定域系，那么上式中的因子 $N!$ 应当去掉。为了对于系统的微观态数目有个数量级的概念，我们来计算一下由 N 个（全同的）单原子分子组成的理想气体的微观态数目。¹⁹ 系统的哈密顿量可以写成：

$$H = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} , \quad (2.16)$$

因此，系统的微观态数目 $\Omega(N, E, V)$ 作为粒子数 N 、系统能量 E 和体积 V 的函数²⁰ 可以由下列积分算出：

$$\Omega(N, E, V) = \frac{V^N}{N! h^{Nr}} \int_{E \leq H \leq E + \Delta E} d^3 p_1 \cdots d^3 p_N , \quad (2.17)$$

这个积分正是 $3N$ 维空间中半径是 $\sqrt{2mE}$ 和半径是 $\sqrt{2m(E + \Delta E)}$ 的两个球面之间所夹的球壳的体积。所以我们得到：

$$\Omega(N, E, V) = \frac{3N}{2E} \left(\frac{V}{h^3} \right)^N \frac{(2\pi m E)^{3N/2}}{N! \left(\frac{3N}{2}\right)!} \Delta E , \quad (2.18)$$

由此可见，微观态数目随系统的能量的增加非常快速地增长（请记住 N 的量级是 10^{23} 左右）。

现在我们 **定义** 系统的 **熵** $S(N, E, V)$ 为：

$$S = k_B \log \Omega(N, E, V) , \quad (2.19)$$

这个定义实际上就是著名的 **玻耳兹曼关系**。于是，将单原子分子理想气体的微观态数目 (2.18) 代入，并且利用大宗量的斯特林（Stirling）公式，我们就得到了单原子分子理想气体的熵：

$$S = N k_B \log \left[\left(\frac{4\pi m E}{3N h^2} \right)^{3/2} \frac{V}{N} \right] + \frac{5}{2} N k_B , \quad (2.20)$$

¹⁹ 这也许是我们唯一能严格计算的例子了。

²⁰ 其实系统的微观态数目还依赖于我们所选取的能壳的宽度 ΔE （正比于 ΔE ）。但是，正如我们下面要看到的，我们感兴趣的是系统微观态数目的对数，而 $\log \Delta E$ 的贡献与其他部分比较是可以忽略的（见下面的熵的公式 (2.20) 中的讨论）。所以，我们在 Ω 中没有明显地写上对于 ΔE 的依赖。

其中我们已经忽略了诸如 $\log(\Delta E/E)$ 和 $\log(3N/2)$ 的项。由于粒子数 N 的数量级是非常巨大的 (大约是 10^{23} 的量级), 所以对于一般的 $(\Delta E/E)$, 其对数都远远小于其他正比于 N 的项; 同时, 正比于 $\log N$ 的项与正比于 N 的项相比, 也是可以忽略的。于是我们发现: 我们上面定义的微正则系综的熵是一个 **广延量**, 它正比于粒子数 N (在所有 **强度量** 不变的情况下); 同时, 系统的熵对于微正则系综的能壳宽度 ΔE 的依赖也趋于零。²¹ 细心的读者这时可能愿意将我们从微正则系综理论得到的单原子分子理想气体的熵的表达式 (2.20) 与以往热力学中的结果进行比较。如果我们代入单原子分子理想气体的内能表达式 $E = 3Nk_B T/2$, 我们发现两者是完全一致的。这从一个方面说明了我们的熵的定义 (2.19) — 至少对于单原子分子理想气体这样一个简单的系统 — 是与热力学理论体系中对熵的定义一致的。下面我们还要更为普遍地来阐明这一点。

为此我们考虑两个系统, 分别称为系统 1 和系统 2。原先它们各自是达到平衡的孤立系。现在, 我们让两个系统接触, 使它们之间可以交换能量,²² 但每一个系统的粒子数和体积仍然保持不变。显然, 这时两个系统各自都不再构成孤立系。但是如果我们将两个系统一起看成一个整个的系统, 这个总系统的能量 E_0 仍然是守恒的, 可以用微正则系综描写。也就是说, 总系统的能量: $E_0 = E_1 + E_2$ 是一个守恒的常数。而且, 这时总系统的微观态数目 Ω_0 就是两个子系统的微观态数目的乘积:

$$\Omega = \Omega_1(N_1, E_1, V_1)\Omega_2(N_2, E_2, V_2), \quad (2.21)$$

由于两个子系统可以交换能量, 所以在达到平衡以后, 两个系统的能量的平均值 (期望值) 将分别达到一个平衡的值: \bar{E}_1 和 \bar{E}_2 , 它们满足约束条件: $\bar{E}_1 + \bar{E}_2 = E_0$, 所以其中一个的取值就确定了另一个。按照孤立系的熵增加原理, 总系统达到平衡时 \bar{E}_1 的取值应当是使得 **总系统** 的熵取极大值。为此, 我们将上式取对数并对 E_1 求导数, 令其为零得到:

$$\left(\frac{\partial S_1(N_1, E_1, V_1)}{\partial E_1} \right)_{N_1, V_1} \Big|_{E_1 = \bar{E}_1} = \left(\frac{\partial S_2(N_2, E_2, V_2)}{\partial E_2} \right)_{N_2, V_2} \Big|_{E_2 = E_0 - \bar{E}_1}, \quad (2.22)$$

其中我们利用了能量的约束条件。这个式子告诉我们, 当两个任意系统通过能量交换达到平衡时, 它们具有相同的 $\left(\frac{\partial S(N, E, V)}{\partial E} \right)_{N, V}$ 值。与热力学第零定律和第二定律比较, 我们发现这个相同的值可以定义为绝对温度的倒数。所以我们有:

$$\frac{1}{T} = \left(\frac{\partial S(N, E, V)}{\partial E} \right)_{N, V}. \quad (2.23)$$

²¹ 如果我们在计算微观态数目 (2.15) 的时候没有放入 $(1/N!)$ 的因子, 那么我们由此得到的熵将不满足广延量的要求。这个问题仅仅出现在熵常数中。它对于通常的热力学过程不会造成影响, 但是在讨论气体的混合这样的不可逆过程时, 会造成所谓的吉布斯佯谬。具体的讨论参见第 19 节。

²² 原则上, 我们要求这两个系统的能量交换过程进行得无限缓慢。用热力学的语言, 就是这个交换过程是一个准静态过程。否则, 下面的等式 (2.21) 原则上会变成不等式 (熵增加)。

类似地，如果我们讨论两个系统可以交换粒子数，同时也可以发生体积改变（但是总粒子数和总体积保持不变），结合热力学中两个系统达到平衡的条件，我们就可以验证：

$$\frac{p}{T} = \left(\frac{\partial S(N, E, V)}{\partial V} \right)_{N, E}, \quad \frac{\mu}{T} = - \left(\frac{\partial S(N, E, V)}{\partial N} \right)_{E, V}. \quad (2.24)$$

其中 p 和 μ 分别是系统的压强和化学势。总结以上的讨论，我们实际上已经建立了一个孤立系的热力学基本微分方程：

$$dS = \frac{1}{T} dE + \frac{p}{T} dV - \frac{\mu}{T} dN. \quad (2.25)$$

因此，从微正则系综理论出发，我们说明了：如果我们利用玻耳兹曼关系 (2.19) 来定义系统的熵，那么我们可以得到与热力学理论完全一致的结果。注意我们这里的讨论是相当普遍的，没有任何涉及到具体系统的性质。所以我们可以说，玻耳兹曼关系是自然界的一个普遍关系。实际上不仅对平衡态，对于非平衡态，玻耳兹曼关系也是熵的唯一合理的统计定义。公式 (2.25) 告诉我们：微正则系综中通过系统微观态数目所确立的熵，实际上是以系统的内能 E 、体积 V 和粒子数 N 为独立变量的特性函数。或者说，对于一个固定内能、固定体积、固定量子数的系统，与之自然对应的系综是微正则系综；与之对应的特性函数是熵（或者等价地说是系统的微观态数目）。一旦求出了系统的熵（微观态数目），系统的所有热力学性质就可以很容易地得到了。读者不难验证，利用公式 (2.23) 和公式 (2.24)，我们得到了单原子分子经典理想气体的熟知的结果：

$$\begin{aligned} E &= \frac{3}{2} N k_B T, \quad pV = N k_B T, \\ \mu &= k_B T \log \left[\frac{p}{k_B T} \left(\frac{h^2}{2\pi m k_B T} \right)^{3/2} \right]. \end{aligned} \quad (2.26)$$

9 正则系综与巨正则系综

上一节中讨论的微正则系综是对应于能量 E 、体积 V 和粒子数 N 的系综。一旦系统微观态数目（或者说系统的熵）计算出来以后，我们就可以得到系统的所有热力学性质。但是微正则系综在实际运用上并不十分便利，在多数实际应用中，我们需要知道系统在固定温度和粒子数时的热力学性质。这就对应于我们现在要讨论的正则系综。

正则系综的分布函数可以通过让系统与一个大热源接触来导出。为此我们仍然考虑两个系统，一个是我们要研究的系统；另一个是具有恒定温度 T 的大热源。²³ 我们现在关心的是，当两者达到平衡时，我们所研究的系统处在一个指定的量子态 S 的几率 ρ_S 是多少。我们很快会发现，这个几率实际上只依赖于系统所处的量子态 S 的能量（能级）

²³即热容量趋于无穷大的系统。

E_S 。²⁴ 按照等几率原理，系统处在这个指定的量子态的几率一定是正比于这时总系统——也就是我们所考虑的系统加上大热源——的总微观状态数。由于所考虑的系统已经处在一个指定的量子态，所以总系统的微观状态数就是大热源具有能量 $E_0 - E_S$ 时热源的微观状态数目 $\Omega(E_0 - E_S)$ 。所以我们要求的几率为：

$$\rho(E_S) \propto e^{\log \Omega(E_0 - E_S)}, \quad (2.27)$$

同时，由于我们假设热源很大，所以我们一定有 $E_S \ll E_0$ 。所以我们可以做级数展开：

$$\log \Omega(E_0 - E_S) \sim \log \Omega(E_0) - E_S \left(\frac{\partial \log \Omega(E)}{\partial E} \right)_{E=E_0},$$

利用前一节的结果 (2.23)，上式中的偏微商正是大热源的绝对温度之倒数：

$$\left(\frac{\partial \log \Omega(E)}{\partial E} \right)_{E=E_0} = \beta = \frac{1}{k_B T}.$$

因此，我们所考虑的与大热源接触并达到平衡的系统处于一个指定量子态 S 的几率为：

$$\rho_S = \frac{1}{Z} e^{-\beta E_S}, \quad (2.28)$$

其中 E_S 就是我们所指定的量子态的能量。这就是著名的吉布斯 **正则分布**。具有这样几率分布的系综被称为 **正则系综**。上式中的归一化常数 Z 由下式给出：

$$Z = \sum_S e^{-\beta E_S}, \quad (2.29)$$

其中求和遍及所有系统可能的量子态 S ，它被称为正则系综的 **配分函数** (partition function)，它实际上是温度 T 、粒子数 N 、体积 V 的函数。²⁵ 我们下面将会看到：一旦得到了系统的配分函数，系统的所有的平衡态热力学性质就完全确定了。²⁶ 显然，正则系综是对应于固定温度、体积和粒子数的系综。

¶ 与正则系综的讨论类似，我们还可以讨论一个与大热源兼大粒子源接触的系统的统计性质。这时系统可以与源交换能量和粒子数，所以只有总能量 E_0 和总粒子数 N_0 是守

²⁴ 对于一个给定的量子态 S ，如何计算它的能量 E_S 是一个纯粹量子力学的问题，不是一个统计物理的问题。因此在我们统计物理的讨论中，我们都将假设对于任意一个给定的量子态 S ，它的能量 E_S 是已知的。

²⁵ 这里指的是最简单的 PVT 系统，否则的话，体积应当换为相应的广义位移。

²⁶ 配分函数的符号 Z 是首先由 Planck 使用的。玻耳兹曼用 *Zustandsumme* 来称呼配分函数，其实玻耳兹曼的名称倒是更贴切一些。所以 Planck 就以 Z 来代表配分函数。

恒的。现在我们希望求出我们所考虑的系统处在具有指定的粒子数 N 的一个指定的 N -粒子量子态 S 的几率 $\rho_{N,S}$ 。我们完全可以仿照正则系综的讨论得到：

$$\rho_{N,S} \propto e^{\log \Omega(E_0 - E_S^{(N)}, N_0 - N)}, \quad (2.30)$$

其中 S 是我们指定的、具有 N 个粒子的、系统可能的一个量子态， $E_S^{(N)}$ 是这个量子态的能量。我们仍然可以假定： $E_S^{(N)} \ll E_0$ 和 $N \ll N_0$ ，于是我们可以将微观态数目的对数（熵）展开，并利用关系 (2.23) 和 (2.24) 得到：

$$\rho_{N,S} = \frac{1}{\Xi} e^{-\beta E_S^{(N)} - \alpha N}, \quad (2.31)$$

其中 $\beta = 1/(k_B T)$ 为大热源的温度倒数，而 $\alpha = -\mu/(k_B T)$ 与源的温度和化学势有关。

上面公式中的归一化常数 Ξ 被称为系统的巨配分函数，它由归一化条件定出：

$$\Xi = \sum_N e^{-\alpha N} \sum_S e^{-\beta E_S^{(N)}}. \quad (2.32)$$

分布 (2.31) 被称为巨正则分布，满足巨正则分布的系综被称为巨正则系综。显然，巨正则系综是对应于固定温度、体积和化学势的系综。

¶ 上面巨正则分布的讨论中我们假定系统只包含一种化学组分。如果我们所考虑的系统中有 k 种化学物质，并且假定各个化学组分之间没有化学反应，同时我们假定要考虑的系统可以与热源兼粒子源交换所有组分的粒子。这时，公式 (2.31) 显然应推广为：

$$\rho_{N_1, N_2, \dots, N_k; S} = \frac{1}{\Xi} e^{-\beta E_S^{\{N_i\}} - \sum_{i=1}^k \alpha_i N_i}. \quad (2.33)$$

其中 $\alpha_i = -\mu_i/(k_B T)$ 与源的温度和第 i 种组元的化学势有关。这时，巨配分函数 $\Xi(T, V, \mu_1, \dots, \mu_k)$ 仍然由归一化条件定出：

$$\Xi = \sum_{S, \{N_i\}} e^{-\beta E_S^{\{N_i\}} - \sum_{i=1}^k N_i \alpha_i}. \quad (2.34)$$

这个分布我们会在讨论混合理想气体时用到（参见第 19 节）。

10 热力学公式

¶ 所谓热力学公式是指将处于平衡态的系统的热力学量与系统的统计配分函数联系起来的公式。我们首先以正则分布为例，来推导热力学公式。正则系综所对应的系统是一个具有固定粒子数 N 、体积 V （或一般的一个广义位移 y ）和温度 T 的宏观系统。利用系统的内能是系统微观总能量的系综平均值，我们首先得到系统内能的热力学公式：

$$U = \bar{E} = \frac{1}{Z} \sum_S E_S e^{-\beta E_S} = -\frac{\partial}{\partial \beta} \log Z . \quad (2.35)$$

¶ 类似的，对于广义力 Y ，如果它所对应的广义位移是 y ，系统的能级 E_S 会依赖于广义位移 y ，那么我们有：

$$Y = \frac{1}{Z} \sum_S \frac{\partial E_S}{\partial y} e^{-\beta E_S} = -\frac{1}{\beta} \frac{\partial}{\partial y} \log Z , \quad (2.36)$$

这个公式的一个重要特例就是关于简单 PVT 系统的压强公式：

$$p = \frac{1}{\beta} \frac{\partial}{\partial V} \log Z . \quad (2.37)$$

¶ 系统熵的统计表达式需要与热力学第二定律比较得到。首先注意到我们引入的系统正则配分函数是 β 和广义位移 y 的函数（假定粒子数恒定），所以一定有：

$$d \log Z = \frac{\partial \log Z}{\partial \beta} d\beta + \frac{\partial \log Z}{\partial y} dy , \quad (2.38)$$

因此，利用上面的系统内能和广义力的统计表达式，我们得到：

$$\beta(dU - Y dy) = d \left(\log Z - \beta \frac{\partial \log Z}{\partial \beta} \right) . \quad (2.39)$$

这个式子说明 β 是微分式 $dU - Y dy$ 的一个 **积分因子**，与我们熟知的热力学第二定律比较，我们得知熵的表达式为：

$$S = k_B \left(\log Z - \beta \frac{\partial \log Z}{\partial \beta} \right) . \quad (2.40)$$

因此，对于一个给定粒子数 N 、体积 V 、温度 T 的系统，只要求出它的正则配分函数 Z ，就可以完全确定它的所有热力学函数，系统的热力学性质也就完全确定了。这个特性与热力学中的亥姆霍兹自由能 $F = U - TS$ 是类似的。利用这里推导出来的热力学公式，很容易直接证明：

$$F = -k_B T \log Z . \quad (2.41)$$

也就是说，正则系综的配分函数与系统的亥姆霍兹自由能完全等价。

¶ 巨正则系综所对应的系统是一个具有固定化学势 μ 、体积 V （或一般的一个广义位移 y ）和温度 T 的宏观系统。它的热力学公式的推导与正则系综的情形完全类似，我们不再重复其推导，而只是列出最后得到热力学公式：

$$\bar{N} = -\frac{\partial}{\partial \alpha} \log \Xi, \quad (2.42)$$

$$U = -\frac{\partial}{\partial \beta} \log \Xi, \quad (2.43)$$

$$Y = -\frac{1}{\beta} \frac{\partial}{\partial y} \log \Xi, \quad (2.44)$$

$$S = k_B \left(\log \Xi - \alpha \frac{\partial \log \Xi}{\partial \alpha} - \beta \frac{\partial \log \Xi}{\partial \beta} \right), \quad (2.45)$$

其中 Ξ 是系统的巨配分函数，它被看成是 α , β , V （或任意给定广义位移 y ）的函数。因此，对于一个给定化学势 μ 、体积 V 、温度 T 的系统，只要求出它的巨正则配分函数 Ξ ，就可以完全确定它的热力学性质。这个特性与热力学中的巨势 $J = F - \mu N = -pV$ 是类似的。利用这里推导出来的热力学公式，很容易直接证明：

$$J = -k_B T \log \Xi. \quad (2.46)$$

¶ 服从正则分布的系统具有固定的温度，但是它可以与大热源发生能量的交换，所以系统的能量并不是固定的，而是可以在其平均值附近 **涨落**。这种涨落可以利用前面的热力学公式加以计算。经过简单的计算我们得到：

$$\langle (E - \bar{E})^2 \rangle = -\frac{\partial \bar{E}}{\partial \beta} = k_B T^2 C_V, \quad (2.47)$$

所以我们看到，系统的能量的涨落直接与温度以及系统的热容量有关。一个服从正则分布的系统的能量 **相对涨落** 的平方为：

$$\frac{\langle (E - \bar{E})^2 \rangle}{\bar{E}^2} = \frac{k_B T^2 C_V}{\bar{E}^2}. \quad (2.48)$$

由于 C_V 和 \bar{E} 都是广延量，它们与系统的总粒子数 N 成正比。因此上式实际上与系统的总粒子数 N 成反比。对于一个宏观的热力学系统， N 的数量级是 10^{23} ，因此能量的相对涨落（与 $1/\sqrt{N}$ 成正比）是很小的。

类似的，对于巨正则分布的系统，其粒子数也可以有涨落。简单的计算表明，系统粒子数相对涨落的平方为：

$$\frac{\langle (N - \bar{N})^2 \rangle}{\bar{N}^2} = \frac{k_B T}{\bar{N}^2} \left(\frac{\partial \bar{N}}{\partial \mu} \right)_{T, V}, \quad (2.49)$$

运用热力学的关系可以证明上式可以化为：

$$\frac{\langle (N - \bar{N})^2 \rangle}{\bar{N}^2} = \frac{k_B T}{V} \kappa_T, \quad (2.50)$$

其中 $\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_T$ 是系统的 **等温压缩系数**。由于 V 是广延量，我们再次看到：对于一个宏观热力学系统，只要等温压缩系数不是很大，系统粒子数的相对涨落仍然是非常小，与 $1/\sqrt{N}$ 成正比。只有在系统趋近于它的 **临界点** 时，等温压缩系数 κ_T 将趋于无穷，这时系统内的粒子数涨落也可以很强。有关系统在临界点附近的涨落问题，本课程的后面（第 27 节）还会更加详细地讨论。

¶ 由前面的讨论我们得知：当构成宏观系统的微观粒子数目很大时，系统的能量、粒子数等物理量的涨落都是非常小的。其相对涨落一般按照 $1/\sqrt{N}$ 的行为趋于零。如果忽略掉这些涨落效应，那么微正则系综、正则系综和巨正则系综都给出一个宏观系统 **等价的** 统计描述。也就是说，如果忽略涨落效应，用不同系综计算出来的宏观系统的热力学量将是一致的。选取不同的统计系综来处理，只不过相当于选取不同的热力学特性函数。当然，如果系统的涨落不能忽略（例如对于比较小的介观系统），那么不同的统计系综的结果一般是不等价的。

11 涨落的准热力学理论

¶ 在上一节中我们简单讨论了正则系综中的能量涨落以及巨正则系综中的粒子数涨落，实际上我们可以讨论一个任意热力学量的涨落。这个理论被称之为涨落的准热力学理论，是爱因斯坦首先提出的。我们考虑一个热力学系统，它与一个大的热源接触，热源具有确定的温度 T 和压强 p 。我们所考虑的系统与热源一起组成的复合系统是一个孤立系，所以当系统的能量和体积发生一个涨落 ΔE 和 ΔV 时，热源的能量与体积也相微观粒子数目应地发生一个变化 ΔE_r 和 ΔV_r ，它们满足：

$$\Delta E + \Delta E_r = 0, \quad \Delta V + \Delta V_r = 0, \quad (2.51)$$

当系统的能量与体积取其最可几值（同时也是平均值） \bar{E} 和 \bar{V} 时，复合系统的熵具有最大值 $\bar{S}^{(0)}$ ，它与复合系统的微观态数目之间的关系由玻耳兹曼关系描述。由于每个微观态都是等几率地出现（等几率原理），所以系统的能量和体积出现 ΔE 和 ΔV 的涨落的几率正比于：

$$W \propto \exp \left(\frac{S^{(0)} - \bar{S}^{(0)}}{k_B} \right), \quad (2.52)$$

由于 $\Delta S^{(0)} = \Delta S + \Delta S_r$ ，并且：

$$\Delta S_r = (\Delta E_r + p\Delta V_r)/T = -(\Delta E + p\Delta V)/T, \quad (2.53)$$

其中我们运用了约束条件 (2.51)。²⁷于是我们得到：

$$W \propto \exp \left(\frac{T\Delta S - \Delta E - p\Delta V}{k_B T} \right). \quad (2.54)$$

²⁷这里我们对于我们研究的系统和热源进行了不同的处理。对于热源，我们将它的热力学基本微分方程只展开到一阶；对于系统，我们则保留到二阶。这相当于忽略了热源（假定它足够的大，以至于其涨落可以忽略）的涨落。

现在我们需要将上面的式子化为用两个独立的物理量的涨落（例如 ΔE 和 ΔV ）来表达。这可以利用 ΔE 的表达式以及热力学第二定律得到：

$$\begin{aligned}\Delta E - T\Delta S + p\Delta V &= \frac{1}{2} \left(\frac{\partial^2 E}{\partial S^2} (\Delta S)^2 + 2 \frac{\partial^2 E}{\partial S \partial V} (\Delta S \Delta V) + \frac{\partial^2 E}{\partial V^2} (\Delta V)^2 \right) + \dots \\ &= \frac{1}{2} (\Delta S \Delta T - \Delta p \Delta V) .\end{aligned}\quad (2.55)$$

特别要注意的是，上面这个公式的最后一行中我们已经将整个系统的熵的变化表达为一般的形式。这个形式已经可以运用到更普遍的情形（例如，系统的温度和压强不一定是固定的情形）。所以我们最终得到一个普遍的表达式：

$$W \propto \exp \left(-\frac{\Delta S \Delta T - \Delta p \Delta V}{2k_B T} \right) . \quad (2.56)$$

¶ 公式 (2.56) 可以有多种用途。我们可以取任何我们感兴趣的两个独立变量来研究他们的涨落。例如，如果我们以体积 V 和温度 T 为独立变量，我们得到：

$$\begin{aligned}\Delta S &= \frac{C_V}{T} \Delta T + \left(\frac{\partial p}{\partial T} \right)_V \Delta V , \\ \Delta p &= \left(\frac{\partial p}{\partial T} \right)_V \Delta T + \left(\frac{\partial p}{\partial V} \right)_T \Delta V ,\end{aligned}\quad (2.57)$$

带入涨落的普遍公式 (2.56)，得到：

$$W \propto \exp \left(-\frac{C_V}{2k_B T^2} (\Delta T)^2 + \left(\frac{\partial p}{\partial V} \right)_T \frac{(\Delta V)^2}{2k_B T} \right) . \quad (2.58)$$

这个公式告诉我们温度的涨落和体积的涨落由两个独立的高斯分布描写，所以我们得到：

$$\overline{(\Delta T)^2} = T^2 k_B / C_V , \quad \overline{(\Delta V)^2} = V \kappa_T k_B T , \quad (2.59)$$

其中 $\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_T$ 为等温压缩系数。这个公式也说明了，一个热力学系统如果要对于温度和体积的涨落保持稳定，它的定容热容量和等温压缩系数必定是非负的。如果体系的粒子数保持不变，那么系统体积的涨落也可以表达成系统密度的涨落。通过这个公式我们还看到，在临界点附近，系统的等温压缩系数趋于无穷大，系统内部的密度涨落也发散，这时原先的两相的分界面就会消失。同时，如果有可见光入射到系统内部，会观察到所谓的 **临界乳光现象**。我们将在以后详细讨论临界点附近的涨落和关联现象（第 27 节中的讨论）。显然如果取其他的变量为独立变量，我们还可以得到其他物理量的涨落，这些都留作练习。

12 近独立子系的统计分布

¶ 本节中，我们将利用前面建立起来的统计系综理论来讨论最为简单的系统：由无相互作用的子系统构成的宏观系统。这类系统又被称为 **近独立子系**。²⁸ 一个近独立子系的总能量 E 可以表达为组成系统的各个子系统（我们姑且称之为粒子）的能量之和：

$$E = \sum_{i=1}^N \varepsilon_i, \quad (2.60)$$

其中 ε_i 是构成系统的第 i 个子系统（粒子）的能量。需要注意的是：近独立子系只要求构成系统的各个子系统之间没有相互作用，从而系统的总能量就是各个子系统能量之和，但是各个子系统可以和外界有相互作用。我们前面引入的统计系综理论是关于处于平衡态的宏观系统的普遍理论，当然也可以用来处理特殊的近独立子系。由于近独立子系的简单性，一个近独立子系的配分函数（无论是正则系综的正则配分函数还是巨正则系综的巨配分函数）都可以得到简化。本节的讨论分为两个部分：首先，我们利用正则系综来处理经典的、也就是不考虑量子全同性影响的近独立子系的统计问题，这将导致著名的麦克斯韦—玻耳兹曼分布；然后我们讨论量子的近独立子系（又称为量子理想气体）的统计问题，这将导致费米—狄拉克分布或玻色—爱因斯坦分布。

对于一个由 N 个粒子构成近独立子系，各个粒子是否可分辨，或者说是否考虑量子力学全同性原理，将直接影响我们对于这个系统状态的描述。如果粒子是 **可分辨** 的，那么整个系统的状态 S 就由每一个可以分辨的粒子所处的单粒子态所决定，即： $S = (s_1, s_2, \dots, s_N)$ 。换句话说，给定了每一个粒子所处的量子态，就唯一地确定了整个系统的一个状态，反之亦然。如果粒子是不可分辨的，量子力学的全同性原理告诉我们，我们只能以各个单粒子态上的粒子数： $\{a_s : s = 1, 2, \dots\}$ 来描述系统。也就是说，我们只能以每个单粒子态 s （它遍及所有可能的单粒子态）上有几个粒子，我们无法区别是哪几个粒子，因为它们在量子力学上是全同的。在表 2.2 中，我们总结了全同性原理对于系统状态描述的影响。

表 2.2: 微观粒子组成系统的状态描述以及粒子数按照量子态的平均分布

构成系统的粒子	全同性原理	整个系统状态描述	粒子数按照量子态的平均分布
可分辨	不起作用	$S = (s_1, s_2, \dots, s_N)$	M.B. 分布
不可分辨	起作用	$S = (a_1, a_2, \dots)$	F.D 或 B.E. 分布

¶ 如前所述，对于一个可分辨的粒子组成的近独立子系，**系统的**量子态 S 由每个子系统所处的单粒子态： (s_1, s_2, \dots, s_N) 给出。这里我们用符号 (s_1, s_2, \dots, s_N) 表示：第

²⁸之所以称为近独立子系而不是独立子系是由于完全独立的子系统构成的系统是无法达成热平衡的。

一个子系统处于单粒子态 s_1 ，第二个子系统处于单粒子态 s_2 ， \dots ，第 N 个系统处于单粒子态 s_N 。²⁹ 这时近独立子系的总能量显然就是： $E = \sum_{i=1}^N \varepsilon_{s_i}$ ，这里 ε_{s_i} 表示单粒子态 s_i 的能量。

如果我们利用 **正则系综**，我们发现近独立子系的正则配分函数可以写成：

$$\begin{aligned} Z &= \sum_S e^{-\beta E_S} = \sum_{s_1, s_2, \dots, s_N} e^{-\beta \varepsilon_{s_1}} e^{-\beta \varepsilon_{s_2}} \dots e^{-\beta \varepsilon_{s_N}} \\ &= \left(\sum_{s_1} e^{-\beta \varepsilon_{s_1}} \right) \left(\sum_{s_2} e^{-\beta \varepsilon_{s_2}} \right) \dots \left(\sum_{s_N} e^{-\beta \varepsilon_{s_N}} \right) \\ &= \left(\sum_s e^{-\beta \varepsilon_s} \right)^N \equiv z^N. \end{aligned} \quad (2.61)$$

我们这里再次指出，在推导公式 (2.61) 的时候实际上没有考虑粒子的全同性的影响。如果对于全同费米子，那么就要求单粒子态 s_1, s_2 等等不能有两个相同；对于全同的玻色子，如果两个单粒子态同时被占据，交换两个粒子并不产生新的系统的量子态。因此，我们得到如下重要的结论：对于不考虑全同性影响的近独立子系，其系统的正则配分函数 Z 完全由其 **子系配分函数** z 给出：

$$Z = z^N, \quad z = \sum_s e^{-\beta \varepsilon_s}. \quad (2.62)$$

类似地，我们很容易求出一个粒子处于某个单粒子态 s 的几率：

$$p_s = \frac{1}{z} e^{-\beta \varepsilon_s}. \quad (2.63)$$

因此，单粒子态 s 上的平均粒子数 \bar{a}_s 为：

$$\bar{a}_s = e^{-\alpha - \beta \varepsilon_s}, \quad e^{-\alpha} = N/z. \quad (2.64)$$

经典的（不考虑量子力学全同性影响的）近独立子系的粒子所满足的这个分布就是著名的 **麦克斯韦—玻耳兹曼分布**。满足这个分布的统计有时又被称为 **经典统计**。利用这些结论以及我们正则配分函数的热力学公式，我们可以将（不考虑粒子全同性的）近独立子系的热力学函数用系统的子系配分函数 z 表达出来。例如，对于系统的内能，我们有：

$$U = -N \frac{\partial \log z}{\partial \beta}. \quad (2.65)$$

²⁹ 显然，这样一个描述本身就依赖于粒子是可以分辨的事实，因为如果粒子是不可分辨的，我们就无法做上述的指定。我们仅仅能够说某个单粒子态上有几个子系统占据，而无法区分这几个子系统哪个是第一个或第二个。

¶ 如果我们必须考虑量子力学全同性的影响，那么上面公式 (2.61) 的推导不能成立。³⁰ 由于在近独立子系中，系统的总能量就是各个子系统（粒子）能量之和，所以一个由量子的全同粒子组成的近独立子系的统计性质将完全由各个单粒子量子态上的粒子数所确定。处在各个单粒子量子态上面的一组粒子数被称为粒子数对于单粒子态的一个分布。具体地说，如果我们用 s 来标记单粒子的量子态，处于量子态 s 上的粒子数记为 a_s ，那么一组完整的 a_s 的集合，即： $\{a_s : s = 1, 2, \dots\}$ 被称为系统粒子数按照单粒子量子态的一个分布，或简称分布。为了简化记号，我们有时也将它记为： $\{a_s\}$ 。对于任意的一个分布 $\{a_s\}$ ，它显然满足：

$$\sum_s a_s = N, \quad \sum_s \varepsilon_s a_s = E, \quad (2.66)$$

其中 N 和 E 分别是系统的总粒子数和总能量。显然系统各种分布出现的概率并不是相同的。我们感兴趣的是在某个给定单粒子量子态 s 上出现的粒子数 a_s 的平均值： \bar{a}_s 。这样的一组平均值： $\{\bar{a}_s\}$ 称为系统粒子数按照单粒子量子态的平均分布。

对于由全同粒子构成的近独立子系，计算系统的巨配分函数是比较方便的。按照巨正则系综的理论，系统的巨配分函数 Ξ 可以表达为（参见公式 (2.32)）：

$$\Xi = \sum_{S, N} e^{-\beta E_S^{(N)} - \alpha N}, \quad (2.67)$$

这里的求和首先是对于具有固定粒子数 N 的系统的量子态求和，然后再对于粒子数 N 求和。这就等价于对于系统所有可能的量子态 S — 无论粒子数为多少 — 求和。同时注意到：对于理想量子气体，系统的每一个量子态 S 是与粒子数按照单粒子量子态的分布 $\{a_s\}$ 一一对应的。所以，对于系统的所有量子态的求和就是对于不同单粒子量子态分布的求和利用其粒子数及总能量的表达式 (2.66)，我们得到：

$$\begin{aligned} \Xi &= \sum_{\{a_s\}} \prod_s e^{-\beta a_s \varepsilon_s - \alpha a_s}, \\ &= \prod_s \sum_{a_s} e^{-(\beta \varepsilon_s + \alpha) a_s}. \end{aligned} \quad (2.68)$$

读者可以发现，上面这个等式正好体现了全同性的效应。它与公式 (2.61) 的最大不同是：公式 (2.61) 中是对于单粒子态 s_i 求和；这里是对于单粒子态上的粒子数 a_s 求和。也就是说，由于全同性的影响，我们只在乎这个单粒子态上有几个粒子，而不在于是哪几个粒子。³¹ 需要注意的是，这个等式仅仅对于巨正则配分函数才是严格成立的，对于正则配分函数，由于其总粒子数必须守恒，因此全同性的影响会对各个 a_s 的求和加上一个约束条

³⁰ 这时候，即使是无相互作用的近独立子系，要严格求出其正则配分函数也不是十分容易的。但是，这是求出系统的巨配分函数却是比较容易的。

³¹ 事实上，量子力学的全同性不允许我们区分是哪几个粒子。

件；对于巨配分函数，我们反正要对所有可能的 N 求和，因此就没有了这个约束条件，这使得这个求和可以更容易地求出。

现在我们需要区分两种不同的情形：即构成系统的粒子是玻色子还是费米子。我们将分别称这两种系统为 **理想玻色气体** 和 **理想费米气体**。对于理想玻色气体，它不受泡利不相容原理的限制，所以上式中对于 a_s 的求和可以从 0 延伸到无穷大；对于理想费米气体，由于受到泡利不相容原理的限制，上式中对于 a_s 的求和只能取 0 和 1 两个可能的值。因此，我们就得到理想量子气体的巨配分函数的表达式：

$$\begin{aligned}\Xi &= \prod_s \left(1 \pm e^{-\alpha - \beta \varepsilon_s}\right)^{\pm 1}, \\ \log \Xi &= \pm \sum_s \log \left(1 \pm e^{-\alpha - \beta \varepsilon_s}\right),\end{aligned}\quad (2.69)$$

其中上（下）面一个符号对应于理想费米（玻色）气体。

¶ 为了求出某个指定单粒子量子态 s 上的平均粒子数，我们定义：

$$\zeta_s = \pm \log \left(1 \pm e^{-\alpha - \beta \varepsilon_s}\right), \quad (2.70)$$

那么我们可以轻易证明，某个单粒子量子态 s 上的平均粒子数 \bar{a}_s 由下式给出：

$$\bar{a}_s = -\frac{\partial \zeta_s}{\partial \alpha}. \quad (2.71)$$

利用前面的表达式，我们得到：

$$\bar{a}_s = \frac{1}{e^{\alpha + \beta \varepsilon_s} \pm 1}. \quad (2.72)$$

这就是著名的 **玻色分布** 和 **费米分布**。它给出了一个量子理想气体中粒子数按照单粒子量子态的平均分布。

公式 (2.64) 和 公式 (2.72) 告诉我们：无论对于经典的近独立子系还是对于量子的近独立子系，一个给定单粒子量子态上粒子数的平均值只和该量子态的能量 ε_s 有关，而与该量子态的其他量子数无关。因此，具有相同能量的单粒子态上面的粒子数的平均值是相等的。具体地说，如果某个单粒子能级 ε_l 的 **简并度** 为 ϖ_l ，那么这个单粒子 **能级** 上的平均粒子数为：

$$\bar{a}_l^{M.B.} = \varpi_l e^{-\alpha - \beta \varepsilon_l}, \quad (2.73)$$

$$\bar{a}_l^{F.D./B.E.} = \frac{\varpi_l}{e^{\alpha + \beta \varepsilon_s} \pm 1}. \quad (2.74)$$

如果费米或玻色分布中与化学势有关的参数 $\alpha = -\mu/(k_B T)$ 满足：

$$e^\alpha \gg 1, \quad (2.75)$$

这时，费米和玻色分布的区别将消失，两者都退化为公式 (2.64) 所给出的、经典的麦克斯韦—玻耳兹曼分布。这个条件被称为 **非简并条件**。上述非简并条件也可以等价地表述为，对于每个能级：

$$\varpi_l \gg \bar{a}_l. \quad (2.76)$$

这个公式的物理含义十分明确： ϖ_l 表示某个指定能级中可供占据的总的量子态的数目；如果在每个能级上平均的粒子数远远小于可供占据的数目，那么两个粒子同时占据到一个量子态的几率就是可以忽略的，而恰恰是两个粒子同时占据一个量子态时，量子力学的全同性原理才起作用。所以，如果非简并条件成立，实际上量子力学的全同性原理所起的作用是可以忽略的，这时当然量子统计（费米—狄拉克统计与玻色—爱因斯坦统计）都回到量子力学全同性原理不起作用的经典统计，即麦克斯韦—玻耳兹曼统计。

在这一节中，我们利用系综理论推导了近独立子系的统计分布。对于不考虑量子力学全同性影响的、经典近独立子系，我们得到了麦克斯韦—玻耳兹曼分布 (2.64)；对于量子力学全同性起作用的量子近独立子系（量子理想气体），我们得到了费米—狄拉克分布和玻色—爱因斯坦分布 (2.72)。在结束本节之前，我们对它们之间的关系和适用范围等作如下的几点评述：

- 我们这里的讨论只是从普遍的系综理论出发，抽象地导出了经典的近独立子系满足的麦克斯韦—玻耳兹曼分布以及量子的近独立子系的费米—狄拉克分布或玻色—爱因斯坦分布。这些分布所对应的具体实例将在后面两章中加以详细的介绍。关于经典的麦克斯韦—玻耳兹曼分布的例子可以参考本讲义的第 18 节和第 19 节；关于玻色—爱因斯坦分布的例子可以参见第 14 节、第 15 节和第 16 节；关于费米—狄拉克分布的例子可以参考第 17 节。
- 麦克斯韦—玻耳兹曼分布、费米—狄拉克分布、玻色—爱因斯坦分布都仅仅是对于近独立子系才严格成立。也就是说，对于有相互作用的系统，其粒子数按照单粒子态的分布实际上是十分复杂的。具体要计算这种分布也必须利用普遍的系综理论。上面这三种分布都只是在粒子之间的相互作用可以忽略时才是真实系统分布的一个好的近似。对于服从经典描述的、有相互作用的系统，我们将在第四章（第 20、21、22 节）和第五章中利用普遍的系综理论进行讨论。对于完全由量子力学描写的、有相互作用的系统，其统计性质的讨论完全超出了本课程的范围，将会在后续课程（量子统计课程）中进行讨论。这里顺便提一下某某 **统计** 与某某 **分布** 之间

的区别。统计一词实际上更加广泛。例如，对于任意的有相互作用的系统，处理的方法一定是经典统计或量子统计。当然，前者可以视为后者的经典极限。量子统计又可以分为费米—狄拉克统计和玻色—爱因斯坦统计。它们在经典极限下退化为麦克斯韦—玻耳兹曼统计。但是，如果系统存在相互作用（相互关联），那么任何一种情形下粒子的分布都不是上面提到的三种分布中的任何一种。只有对于近独立子系，这三种分布才严格成立。

- 我们这里利用系综理论求出的是某个单粒子态上占据的粒子数的系综平均值 \bar{a}_s 。另外一种导出这些分布的方法是去计算所谓的 **最可几分布**，又称为 **最概然分布**。它是指使得系统微观态数目取极大值的分布。按照等几率原理，最可几分布出现的几率也是最大的。由于可以证明最可几分布所对应的极大值对宏观系统而言实际上是一个十分尖锐的极大值，因此由此导出的最可几分布实际上与我们利用系综理论计算的平均分布几乎是相同的。利用拉格朗日乘子导出最可几分布的方法，我们列为本章后面的习题，供有兴趣的读者演练。当然，大家也可以参考汪志诚教材上面的推导 [1]。但是我们需要指出的是，从推导过程来看，利用系综理论的推导更加严格。
- 对于量子的近独立子系，如果 **非简并条件** 得到满足，那么两种量子分布的区别消失，同时它们都回到经典的麦克斯韦—玻耳兹曼分布。这个条件是独立于前面提到的 **准连续条件**。换句话说，非简并条件只是保证了量子力学的全同性原理不再起明显的作用，但并不一定保证系统的物理量是准连续的（参见第 7.2 小节的讨论）。

13 近独立子系中粒子数分布的涨落

¶ 作为前面推导出来的近独立子系三种统计分布的一个应用，我们来讨论粒子数按照量子态分布的涨落问题。

对于量子分布（玻色或费米），一个给定的单粒子量子态 s ，其上占据的平均粒子数为 \bar{a}_s ，那么容易验证这个平均值的涨落可以简单表达为：

$$\langle (a_s - \bar{a}_s)^2 \rangle = -\frac{\partial \bar{a}_s}{\partial \alpha} . \quad (2.77)$$

利用平均分布的表达式带入我们得到：

$$\langle (a_s - \bar{a}_s)^2 \rangle = \bar{a}_s (1 \pm \bar{a}_s) , \quad (2.78)$$

其中 + 号对应于玻色—爱因斯坦分布，- 号对应于费米—狄拉克分布。我们发现，如果温度比较低，费米—狄拉克分布指出，单粒子态的平均占据数一定是 1 或 0。因此，这时对于费米—狄拉克系统，其分布的涨落很小。对于玻色系统，由于没有 Pauli 不相容原理

的限制，它的相对涨落一般要大得多。这个公式可以应用在黑体辐射（光子气体）的能量涨落中。³²

¶ 讨论经典的麦克斯韦—玻耳兹曼分布的涨落反而相对要复杂一些。为此，我们现写出一个粒子占据一个单粒子态 s 的几率：

$$p_s = \frac{e^{-\beta\varepsilon_s}}{z}, \quad (2.79)$$

其中 $z = \sum_s e^{-\beta\varepsilon_s}$ 为子系配分函数。显然 $\sum_s p_s = 1$ 。于是，各个单粒子态 s 上分别有 a_s 个粒子的几率为一个多项式分布：³³

• • • • •

$$P_{\{a_s\}} = \frac{N!}{\prod_s a_s!} \prod_s (p_s)^{a_s}. \quad (2.80)$$

于是一个给定单粒子态上占据的粒子数的平均值可以由下式求出：

$$\bar{a}_s = \sum_{\{a_s\}} a_s P_{\{a_s\}} = p_s \frac{\partial}{\partial p_s} \left(\sum_{\{a_s\}} P_{\{a_s\}} \right),$$

但是上式的括号内的表达式正好就是：

$$\sum_{\{a_s\}} P_{\{a_s\}} = \left(\sum_s p_s \right)^N. \quad (2.81)$$

因此，我们自然得到：³⁴

$$\bar{a}_s = N p_s = \frac{N}{z} e^{-\beta\varepsilon_s}. \quad (2.82)$$

这正是麦克斯韦—玻耳兹曼分布。利用这种方法，我们可以计算出 a_s 的涨落：

$$\begin{aligned} \langle (a_s - \bar{a}_s)^2 \rangle &= \bar{a}_s^2 - \bar{a}_s^2 \\ &= \left(p_s \frac{\partial}{\partial p_s} \right)^2 \left(\sum_s p_s \right)^N - \left[p_s \frac{\partial}{\partial p_s} \left(\sum_s p_s \right)^N \right]^2, \end{aligned}$$

经过一些简单的计算我们得到：

$$\langle (a_s - \bar{a}_s)^2 \rangle = \bar{a}_s, \quad (2.83)$$

其中我们假定了 $N \gg 1$ 。这就是麦克斯韦—玻耳兹曼分布的涨落公式。


³²虽然光子气体的化学势为零，因此公式 (2.77) 对于光子气体不能适用，但是最终的结果 (2.78) 仍然是成立的。

³³不熟悉的同学可以参考本讲义后面的附录。


³⁴注意，在这类计算中我们必须首先计算对 p_s 的偏微商然后再利用 p_s 的归一条件。



相关的阅读



这一章我们讨论了经典的系综理论，对于量子系综理论只是介绍了一下它的概念，而对于其表述方式（密度算符的表述方式）则没有涉及。有一定量子力学基础的读者可以参考 Landau 的书 [4]。无论是最为经典的麦克斯韦-玻耳兹曼分布，还是量子的玻色分布和费米分布都可以从近独立子系的最可几分布得到。这种推导方法在一些普通物理热学的书中都有所讨论。如果读者对此并不了解，可以参考汪志诚先生的书 [1]。



第三章 量子理想气体

本章提要

- 理想玻色气体与玻色爱因斯坦凝聚 (14)
- 黑体辐射 (15)
- 固体热容量 (16)
- 理想费米气体 (17)

这

一章中，我们将利用前一章得到的关于量子近独立子系的一般公式来具体研究所谓的量子理想气体。我们这里所说的量子理想气体是指满足量子统计分布（费米分布或玻色分布）的近独立子系。它们的具体形态并不一定是我们日常生活中所说的气体形态。如果构成一个量子力学体系的多个粒子之间存在的相互作用，但在某些特定的条件下其相互作用可以忽略，这时这个体系仍然可以用量子近独立子系（量子理想气体）很好地近似。当粒子之间的相互作用不可忽略时，我们就必须从普遍的量子系综理论（密度算符描述）出发了。这样的系统有时又被称为量子液体，将在后续课程中讨论。本课程将只讨论量子理想气体。按照构成系统的粒子是玻色子还是费米子，量子理想气体可以分为理想玻色气体和理想费米气体两大类。

14 理想玻色气体

这一节中我们利用玻色-爱因斯坦分布来讨论一般的非相对论性的理想玻色气体。我们将分为弱简并和强简并两种情况讨论。在弱简并的情形下，体系的性质可以与经典的理想气体进行比较。事实上，体系的热力学量可以表达成一个级数展开式。这个展开式的首项就是经典理想气体所对应的物理量。后续的项可以看成是全同性原理所造成的量子修

正。在强简并的情形，体系则会出现所谓的玻色-爱因斯坦凝聚现象。这是一个纯粹的量子现象，也是近年来理论和实验十分关注的方向。

14.1 弱简并的非相对论性理想玻色气体

¶ 为了简单起见，这一小节中我们假定组成气体的玻色子的自旋是 0。¹ 同时，我们假定这些玻色子的能量不太高以至于可以利用非相对论来处理。所以，单粒子的能量为：

$$\varepsilon = \frac{\mathbf{p}^2}{2m}, \quad (3.1)$$

其中 m 是玻色子的质量， \mathbf{p} 是它的平动动量。我们假定温度不太低，或气体的体积足够大，从而粒子的平动动能可以看成是准连续的。因此按照第 12 节中的公式 (2.69)，系统巨配分函数对数中的求和可以用积分来代替。由于一个非相对论性的粒子的动能 - 无论是玻色子还是费米子 - 总是具有式 (3.1) 的形式，我们可以引入所谓的单粒子量子态的态密度 (density of states, DOS) 的概念。我们用 $g(\varepsilon)d\varepsilon$ 来表示系统中单粒子能量处在 $[\varepsilon, \varepsilon + d\varepsilon]$ 之中的单粒子量子态的数目。那么由于：

$$\frac{V d^3\mathbf{p}}{h^3} = \frac{4\pi V p^2 dp}{h^3} = \frac{2\pi V}{h^3} (2m)^{3/2} \varepsilon^{1/2} d\varepsilon,$$

其中第一步我们将 $d^3\mathbf{p}$ 写到动量空间的球坐标中 ($p = |\mathbf{p}|$ 表示动量的大小) 而第二步我们利用了 $\varepsilon = p^2/(2m)$ 换成用单粒子能量来表达。于是我们得到系统的态密度 $g(\varepsilon)$ ：

$$g(\varepsilon) = \frac{2\pi V}{h^3} (2m)^{3/2} \varepsilon^{1/2} \quad (3.2)$$

注意，这个态密度仅仅考虑了平动自由度相关的态密度，没有考虑粒子的其他自由度（例如自旋）。显然，对于自旋为 s 的粒子，这个态密度还需要乘以自旋的简并度 $(2s + 1)$ 。

利用系统的态密度，我们可以将体系的巨配分函数的对数的积分进行无量纲化（即令 $\beta\varepsilon = x$ ），得到的结果是，

$$\log \Xi = -\frac{2\pi V}{h^3} (2mk_B T)^{3/2} \int_0^\infty \log(1 - e^{-\alpha-x}) x^{1/2} dx. \quad (3.3)$$

值得注意的是，由于能态密度的式子中的因子 $\varepsilon^{1/2}$ ， $\varepsilon = 0$ 能级（即单粒子基态）上粒子的贡献实际上是被忽略了。这在温度不是很低时是可以的，因为这时一个能级上的粒子数与总粒子数比是个小量。² 如果温度很低，低于某个凝聚温度时，会出现所谓玻色

¹如果我们讨论的玻色子的自旋不是零，而是 s ，那么系统的巨配分函数的对数 (3.3)（以及其他由它导出的物理量）需要乘以一个因子： $(2s + 1)$ 。

²尽管平均来讲，占据这个单粒子态（基态）上的粒子数比占据任何一个其他单粒子态上的都多。

— 爱因斯坦凝聚，它的标志是：系统中会有宏观数目的粒子凝聚到 $\varepsilon = 0$ 的量子态上。这时我们就必须将这个态单独考虑。我们将在下一小节处理。现在，我们假定温度仍然高于这个凝聚温度，因此公式 (3.3) 就确定了一个非相对论理想玻色气体的巨配分函数。

由于玻色子的单粒子能量从零一直准连续地延伸到正无穷，因此，对于任意一个给定的正数，在其附近总有一个无限接近它的单粒子能级存在。这个事实实际上决定了理想玻色气体的化学势必定是非正的（也就是说它是负数或零）。因为如果不是这样，我们就可以找到一个与化学势无限接近的单粒子能级，按照第 12 节中玻色分布的公式 (2.72)，在这个能级上的平均粒子数将会任意地大，这将与所有能级上粒子数加起来是有限大矛盾。因此，我们可以假定 $\alpha > 0$ （或者说 $\mu < 0$ ）。这时，我们可以将上面积分中的对数函数展开：

$$\log(1 - e^{-\alpha-x}) = - \sum_{j=1}^{\infty} \frac{e^{-j(\alpha+x)}}{j} . \quad (3.4)$$

将此展开式带入巨配分函数的表达式 (3.3) 并逐项积分得到：

$$\log \Xi = \frac{V}{h^3} (2\pi m k_B T)^{3/2} \sum_{j=1}^{\infty} \frac{e^{-j\alpha}}{j^{5/2}} \equiv \frac{V}{\lambda_T^3} g_{5/2}(z) . \quad (3.5)$$

其中 $z = e^{-\alpha}$ 通常被称为体系的活度 (fugacity)，它直接与化学势相关。我们还定义了粒子的热波长 λ_T 为：

$$\lambda_T \equiv \frac{h}{(2\pi m k_B T)^{1/2}} . \quad (3.6)$$

同时我们引进了函数：

$$g_s(z) = \sum_{j=1}^{\infty} \frac{z^j}{j^s} , \quad s > 0 . \quad (3.7)$$

这个函数一般不能用初等函数来表达。当 $z = 1$ 时，它的函数值就是著名的黎曼 ζ -函数： $g_s(z = 1) = \zeta(s)$ 。由于 $\alpha > 0$ 即 $0 < z < 1$ ，上式中的级数是一致收敛的。因此，前面的逐项积分在数学上是许可的。现在，可以利用我们已经得到的热力学公式计算出所有热力学量：

$$\bar{N} = \frac{V}{\lambda_T^3} g_{3/2}(z) , \quad (3.8)$$

$$U = \frac{3}{2} k_B T \frac{V}{\lambda_T^3} g_{5/2}(z) , \quad (3.9)$$

$$pV = k_B T \frac{V}{\lambda_T^3} g_{5/2}(z) = \frac{2}{3} U , \quad (3.10)$$

$$S = k_B \left(\frac{5}{2} \log \Xi + \bar{N} \alpha \right) . \quad (3.11)$$

我们引入一个无量纲变量 y ，其定义为：

$$y = \frac{Nh^3}{V} (2\pi mk_B T)^{-3/2} = n\lambda_T^3, \quad (3.12)$$

当变量 $y \ll 1$ 时，我们称玻色子系统是弱简并的。这时系统中粒子的平均距离大于粒子本身热运动能量的德布罗意波长。因此，我们预料这时量子效应应当是比较小的。

下面我们主要讨论体系的物态方程。根据前面的公式 (3.8) 和 (3.10)，我们可以将体系的物态方程写为：

$$\frac{pV}{Nk_B T} = \frac{g_{5/2}(z)}{g_{3/2}(z)}. \quad (3.13)$$

但是这并不是通常的物态方程的形式。为了得到通常的物态方程，我们需要将上式等号右边的活度 z 用变量 y （温度、数密度）来表达。根据公式 (3.8)， y 和 z 之间的关系为：

$$y = g_{3/2}(z) \equiv \sum_{j=1}^{\infty} \frac{z^j}{j^{3/2}}. \quad (3.14)$$

当弱简并条件 $y \ll 1$ 成立时，我们可以反解出活度 z 作为 y 的级数：

$$z = y - \frac{1}{2^{3/2}} y^2 + \left(\frac{1}{4} - \frac{1}{3^{3/2}} \right) y^3 - \left(\frac{1}{8} + \frac{5}{8^{3/2}} - \frac{5}{6^{3/2}} \right) y^4 + \dots. \quad (3.15)$$

将这个结果代回到公式 (3.13) 中，我们就得到了弱简并的理想玻色气体的状态方程：

$$\frac{pV}{Nk_B T} = 1 - \frac{1}{2^{5/2}} y - \left(\frac{2}{3^{5/2}} - \frac{1}{8} \right) y^2 - \left(\frac{3}{32} + \frac{5}{2^{11/2}} + \frac{3}{6^{3/2}} \right) y^3 - \dots. \quad (3.16)$$

我们看到，弱简并理想玻色气体的状态方程与经典理想气体的状态方程相比有着偏差。上面的式子就是对于气体密度的系统展开。当弱简并条件成立时，这些量子修正是小的。我们还注意到，这种对于经典理想气体的偏差并不是由于动力学的相互作用引起的。我们所研究的构成玻色气体的微观粒子之间完全没有任何物理的相互作用。即便如此，玻色气体的压强也比具有同样温度和密度的纯经典理想气体的压强要小，仿佛玻色子之间有等效的相互吸引似的。造成这个效应的根源是量子力学的全同性原理，换句话说，是由于采取了不同于经典的统计法。因此，我们称这种关联为统计关联，正是这种统计关联使得理想玻色气体表现出一种等效的相互吸引。

14.2 玻色爱因斯坦凝聚

下面我们讨论极端强简并的情形，也就是 y 接近于 1 或更大时的情况。如果温度趋于零或密度很大时，按照 y 的定义式 (3.12)， y 的数值是可以任意大的；但是根据公

式 (3.14), y 的数值永远不会大于 $g_{3/2}(1) = \zeta(3/2)$ 。造成这个矛盾的根源是由于气体的密度 n 与化学势的关系 (3.8) 实际上在温度很低时是不正确的, 必须重新考虑。正如我们在推导公式 (3.3) 时所提到的, 这个公式中忽略了能量为零的单粒子态上的粒子对于系统物理量的贡献。这个近似在温度不太低时是可以的。但是, 随着温度的降低, 化学势也不断升高;³ 当温度降到某个临界温度 T_c 时, 化学势变为非常接近于零。如果温度进一步降低, 由于化学势不可能成为正的, 因此化学势实际上将不再变化而是维持为零。⁴ 这时, 随着温度的降低, 系统中就会有越来越大量的粒子凝聚到能量最低的单粒子态 (基态) 上。直到绝对零度时, 所有粒子都会凝聚到基态上。这种无相互作用系统中, 宏观数量的玻色子凝聚到能量最低的单粒子态上的现象就被称为 **玻色 — 爱因斯坦凝聚**。玻色 — 爱因斯坦凝聚可以看成是一种特殊的 **相变**。

按照前面的分析, 我们必须将单粒子的基态与激发态分开处理: 基态对于巨配分函数的贡献以求和的形式记入; 激发态的贡献仍然可以用积分来替代求和:

$$\log \Xi = -\frac{2\pi V}{h^3} (2mk_B T)^{3/2} \int_0^\infty \log(1 - e^{-\alpha-x}) x^{1/2} dx - \log(1 - e^{-\alpha}) . \quad (3.17)$$

相应的粒子数为:

$$\bar{N} = \frac{V}{\lambda_T^3} g_{3/2}(z) + \frac{z}{1-z} \equiv \bar{N}_{\varepsilon>0} + \bar{N}_{\varepsilon=0} , \quad (3.18)$$

其中 $z = e^{-\alpha}$ 为体系的活度, 而函数 $g_{3/2}(z)$ 的定义由公式 (3.7) 给出。容易验明, 上式右边的两项分别对应于位于单粒子激发态 (即 $\varepsilon > 0$ 的所有态) 和基态 (即 $\varepsilon = 0$ 的态) 上的粒子数。

公式 (3.18) 值得我们更细致的分析。首先, 由于活度 z 满足 $0 < z < 1$, 又由于对宏观系统 $\bar{N} \gg 1$, 因此我们发现, 只要 z 不是非常接近于 1, 公式 (3.18) 中的第二项 — 也就是单粒子基态上的粒子数 — 都是可以忽略的。如果我们忽略第二项的贡献, 我们就回到了我们上一小节得到的结果, 这时没有任何凝聚的发生。但是, 如果 $z \simeq 1 - O(1/\bar{N})$, 那么公式 (3.18) 中的第二项就变得不可忽略了。但是由于第一项中 $g_{3/2}(z)$ 对 z 的依赖在 $z \simeq 1$ 附近并不敏感, 我们完全可以将其中的 z 宗量取为 1。因此, 这时我们发现: 处于激发态的总粒子数目由:

$$\bar{N}_{\varepsilon>0} = \frac{V}{\lambda_T^3} \zeta(3/2) , \quad (3.19)$$

给出, 而基态上的粒子数则由: $\bar{N}_{\varepsilon=0} = \bar{N} - \bar{N}_{\varepsilon>0}$ 给出, 或者等价地写为:

³注意玻色气体的化学势永远都是非正的。所以, 我们说它升高, 是指其绝对值趋于零。

⁴事实上, 化学势仍然是负的, 只不过非常接近于零。它与零的差别对于所有考虑的物理量都不会造成实质的影响。

$$n_0(T) = n \left[1 - \left(\frac{T}{T_c} \right)^{3/2} \right]. \quad (3.20)$$

其中的相变临界温度可以由下列条件决定：

$$n\lambda_{T_c}^3 = \zeta(3/2) \simeq 2.612, \Rightarrow T_c = \frac{2\pi}{(2.612)^{2/3}} \frac{\hbar^2}{mk_B} n^{2/3}. \quad (3.21)$$

当温度 T 接近 T_c 时，基态数密度几乎是零；随温度降低，基态数密度也增加，直到零温时，所有粒子都凝聚到基态上从而形成玻色—爱因斯坦凝聚。这是符合物理直觉的，因为没有任何理由阻止多个玻色子同时占据一个量子态。这个现象首先是爱因斯坦经过理论分析后提出的，但实验上的验证直到近几年才真正实现。

¶ 发生玻色—爱因斯坦凝聚后的其他物理量也可以类似求出，所注意的只是将基态的贡献与其他态的贡献分开考虑。例如，利用关系： $pV = k_B T \log \Xi$ 我们可以得到理想玻色气体的压强 p ，另一方面，由于我们总是有： $pV = (2/3)U$ ，我们同时也得到了体系的内能。我们将它们统一写为：

$$p = \frac{2U}{3V} = \begin{cases} \frac{k_B T}{\lambda_T^3} g_{5/2}(z) = nk_B T \frac{g_{5/2}(z)}{g_{3/2}(z)} & T > T_c \\ \frac{k_B T}{\lambda_T^3} \zeta(5/2) = \frac{\zeta(5/2)}{\zeta(3/2)} nk_B T \left(\frac{T}{T_c} \right)^{3/2} & T < T_c \end{cases} \quad (3.22)$$

注意，对于 $T > T_c$ ，这里的结果与上一小节的结果完全一致。其中的活度 z 也必须通过公式 (3.8) 反解为体系数密度和温度的函数（见公式 (3.15) 处的讨论）。

从系统的内能出发并对温度求导数就可以求出系统的热容量：⁵

$$\frac{C_V}{Nk_B} = \begin{cases} \frac{15}{4} \frac{g_{5/2}(z)}{g_{3/2}(z)} - \frac{9}{4} \frac{g_{3/2}(z)}{g_{1/2}(z)}, & T > T_c \\ \frac{15}{4} \frac{\zeta(5/2)}{\zeta(3/2)} \left(\frac{T}{T_c} \right)^{3/2}, & T < T_c \end{cases} \quad (3.23)$$

也就是说，当温度在临界温度以下时，系统的热容量与 $T^{3/2}$ 成正比。实际上可以证明：在理想玻色气体的玻色—爱因斯坦凝聚相变中，系统的热容量是连续的。但是，热容量对于温度的导数在临界温度左右不相等，有一个跃变。因此按照 Ehrenfest 的定义，这是一个三级相变。

¶ 虽然在理论上，爱因斯坦早就提出了出现这样的相变的可能性，但在实验上要验证它却相当困难。首先，它需要相当低的温度。人们曾在液氦的超流相变处发现了类似的行为以致于怀疑液氦的超流相变可能是爱因斯坦提出的玻色—爱因斯坦凝聚相变。但后来发现并不是。液氦的超流性是由于氦原子之间的相互作用引起的。而且，其相变的特性是一个典型的二级相变而不是三级相变。随后，人们一直试图寻找一种真正玻色—爱因

⁵当 $T < T_c$ 时的计算十分简单；当 $T > T_c$ 时需要计算 $\partial z / \partial T$ 。这可以从公式 (3.8) 得到。

斯坦凝聚的例子。实验上的困难就在于必须寻找一个系统，它的相变的起因完全是（或者至少是主要是）由于统计关联⁶而不是力学上的相互作用引起的。由于真实的气体总是具有动力学的相互作用，因此要实现真正意义上的玻色-爱因斯坦凝聚，就必须使得气体足够稀薄，以至于其原子或分子间的相互作用可以忽略。这实际上要求极低的温度。因此，只有近年来低温技术（激光冷却和磁约束的技术）得到充分发展后，才可能在实验上实现。当然，与此同时也产生了许多新的诺贝尔奖获得者。

15 黑体辐射

¶ 利用纯粹热力学的方法就可以证明辐射场的能量密度与温度的四次方成正比：

$$u = aT^4 . \quad (3.24)$$

这被称为 Stefan-Boltzmann 定律。⁶ 这里我们将利用统计物理的方法来讨论平衡的辐射场问题。这个问题可以从两个不同的角度来处理，所得到的结果是完全一致的。一个方法是将空窖中的电磁场展开成简正振动⁷的叠加，这些简正模式正是平面波。每一个平面波由它的波矢 \mathbf{k} 以及偏振方向 \mathbf{s} 描述，波矢 \mathbf{k} 的方向代表了平面波的传播方向，而它的大小 k 与电磁波的圆频率 $\omega_{\mathbf{k}}$ 的关系为：

$$\omega_{\mathbf{k}} = ck , \quad (3.25)$$

其中 c 为真空中的光速。在空窖中，波矢 \mathbf{k} 的每一个分量都是量子化的。当然，如果空窖的尺度 L 很大，波矢实际上是准连续的。为简单起见，我们假定空窖是一个边长为 L 的立方体，运用周期边条件，⁷那么我们有：

$$\mathbf{k} = \frac{2\pi}{L} \mathbf{n} , \quad (3.26)$$

其中 \mathbf{n} 是一个分量取整数的矢量。对于一个给定的 \mathbf{k} ，电磁波可以有⁸两个独立的偏振方向。由于 \mathbf{k} 可以取无穷多个值，⁸所以空窖中的辐射场是一个具有无穷多自由度（每个自由度由一个固定的 \mathbf{k} 和确定的偏振方向 \mathbf{s} 描写）的力学体系。利用简正模式（或者说平面波展开）的好处是，这些简正模式在统计上是相互独立的，没有相互作用，每一个模式都是一个简谐振子。将这些谐振子量子化以后，电磁场的总能量可以写成：

$$E = \sum_{\mathbf{k},s} (n_{\mathbf{k},s} + 1/2) \hbar \omega_{\mathbf{k}} . \quad (3.27)$$

⁶关于这一点可以参考汪志诚先生的教科书 [1] 中的第 §2.6 节的讨论。

⁷实际上，如果运用更为复杂的边界，并不会影响我们得到的结论。

⁸也就是说，三维整数矢量 $\mathbf{n} \in \mathbb{Z}^3$ 可以取所有整数值。

因此，黑体辐射可以看成是由不同的 \mathbf{k} 和 s 的简正模构成的近独立子系。值得注意的是，对于一个由 (\mathbf{k}, s) 所描写的简正模，它的圆频率只与波矢 \mathbf{k} 的大小有关（公式 3.25），与波矢的方向、偏振 s 等无关。

按照这种逻辑，对于一个给定的波矢 \mathbf{k} 和偏振 s ，电磁场系统的正则配分函数 Z 可以写为：

$$Z = \sum_{\{n_{\mathbf{k},s}\}} e^{-\beta E} = \prod_{\mathbf{k},s} z_{\mathbf{k},s} ,$$

$$z_{\mathbf{k},s} = \sum_{n_{\mathbf{k},s}=0}^{\infty} e^{-\beta(n_{\mathbf{k},s}+1/2)\hbar\omega_s(\mathbf{k})} = \frac{e^{-\beta\hbar\omega_s(\mathbf{k})/2}}{1 - e^{-\beta\hbar\omega_s(\mathbf{k})}} , \quad (3.28)$$

因此，系统的总能量可以写成各个简正模式的贡献之和：

$$U = \sum_{\mathbf{k},s} -\frac{\partial}{\partial\beta} \log z_{\mathbf{k},s} . \quad (3.29)$$

对于偏振 s 的求和只是贡献一个因子 2，同时利用在空窖中，波矢的大小在 k 到 $k + dk$ 中的简正振动的个数： $8\pi V k^2 dk / (2\pi)^3 = \frac{V}{\pi^2 c^3} \omega^2 d\omega$ ，我们得到空窖中圆频率 $d\omega$ 范围内的辐射场能量 $U(\omega, T) d\omega$ 为：⁹

$$U(\omega, T) d\omega = \frac{V}{\pi^2 c^3} \frac{\hbar\omega^3}{e^{\frac{\hbar\omega}{k_B T}} - 1} d\omega , \quad (3.30)$$

这就是著名的普朗克黑体辐射公式。对圆频率积分就得到热力学中的能量密度的结果。

下面我们用另一种观点来讨论这个问题。我们不从简正模式出发，而是从简正模式的激发—光子—的角度出发来研究黑体辐射。辐射场的振动模式的变化可以看成是不断地吸收或发射光子的过程。所谓黑体辐射就是这些光子所组成的理想气体。由于光子可以自由地被产生和吸收，将不再有光子数守恒的条件。与此相应，光子气体的化学势为零。光子的能量 ε ，动量 \mathbf{p} 与辐射场的圆频率 ω 和波矢 \mathbf{k} 的关系是：

$$\mathbf{p} = \hbar\mathbf{k} , \quad \varepsilon = \hbar\omega = cp , \quad (3.31)$$

也就是说，光子是零质量的粒子，它永远以光速运动。光子是玻色子，而且光子之间的相互作用很弱，可以看成是理想玻色气体。¹⁰所以，在一个能级为 ε_l 上的平均光子数目为：

⁹实际上，还有一项对应于各个简正模式的零点能的贡献被扔掉了。这一项与温度无关，不参与任何物理量的贡献。它只是影响能量的零点的选取（这一项积分后实际上是发散的）。

¹⁰在自由电子和正电子十分稀少的空间，光子与光子之间的相互作用是非常弱的。按照量子电动力学的估计，最低级的贡献来自光子—光子散射过程，这个过程中交换虚的正负电子对。它对于散射截面的贡献正比于 α_{EM}^4 ，其中 $\alpha_{EM} \sim 1/137$ 为电磁作用的精细结构常数。这是一个非常小的量，因此将光子气看成理想气体是一个非常好的近似。

$$\bar{a}_l = \frac{\varpi_l}{e^{\beta\varepsilon_l} - 1}, \quad (3.32)$$

其中 ϖ_l 为该单粒子能级所对应的简并度。由于我们假定空腔是具有宏观尺度的，这时光子的能级是准连续的，我们只要求出光子的态密度就可以了。光子的自旋是 1，但是它的自旋沿其运动方向的分量只能取 ± 1 两个值。¹¹ 简单的计算发现，空腔中在 ω 到 $\omega + d\omega$ 之间的光子的量子态个数为： $\frac{V}{\pi^2 c^3} \omega^2 d\omega$ 。将其乘以一个光子的能量 $\hbar\omega$ ，我们马上得到在相应频率范围内的辐射场能量的普朗克公式：

$$U(\omega, T)d\omega = \frac{V}{\pi^2 c^3} \frac{\hbar\omega^3}{e^{\frac{\hbar\omega}{k_B T}} - 1} d\omega, \quad (3.33)$$

这就是普朗克 (M. Planck) 在 1900 年发表的关于辐射场能量密度随频率分布的内插公式。它在当时几乎完美地拟合了高频和低频的实验曲线。后来，爱因斯坦等人认真地考虑了普朗克公式中提出的能量量子化的方案。从此，人类开始走出经典物理的笼罩，步入一个全新的量子世界。

¶ 在高频极限下，普朗克公式变为：

$$U(\omega, T)d\omega = \frac{V}{\pi^2 c^3} \hbar\omega^3 e^{-\frac{\hbar\omega}{k_B T}} d\omega, \quad (3.34)$$

这正是维恩 (W. Wien, 1864 - 1928) 在 1896 年提出的经验公式 (后被称为 **维恩公式**)。它说明，在高频区域辐射场的能量随频率的分布指数地趋于零。这一特性从量子的观点也是很好理解的。在高频区域，高频光子的能量远远大于热激发的平均能量尺度 $k_B T$ ，因此产生这样高频光子的几率非常之小。所以，这类高能光子对于辐射场的能量几乎是没有贡献的，这正象气体分子的高振动能级被冻结，对热容量没有贡献一样。

在另一个极限，也就是低频极限下，我们得到：

$$U(\omega, T)d\omega = \frac{V}{\pi^2 c^3} \omega^2 k_B T d\omega, \quad (3.35)$$

所以辐射场的能量随频率的分布正比于 $k_B T$ ，而且比例系数正好是在这个频率范围内的振动模式数。也就是说，这恰恰是经典能均分定理的结果，它是瑞利 (Rayleigh) 和金斯 (Jeans) 首先提出的，因此被称为 **瑞利 - 金斯公式**。这个公式也被当时低频辐射的实验所证实。

¶ 我们可以将普朗克公式对频率积分，得到辐射场的总能量：

¹¹ 这一点与一般有质量的自旋是 1 的粒子不同。其原因是光子所对应的电磁场具有 **规范对称性**。这种规范对称性要求光子的自旋分量不能取零的值，而只能取 ± 1 。用简正模的语言，这对应于电磁波的左旋和右旋圆偏振。光子的自旋分量如果取零，对应于纵偏振的电磁波。电磁学的常识告诉我们：自然界中并不存在纵偏振的电磁波。

$$U = \frac{\pi^2 k_B^4}{15c^3 \hbar^3} VT^4 . \quad (3.36)$$

即辐射场的能量密度与温度的四次方呈正比，而且其比例系数可以用自然界的基本常数来表示，这是热力学理论无法预言的。著名的 Stefan-Boltzmann 定律告诉我们，黑体辐射的能流辐射强度为： $J = \frac{c}{4} \frac{U}{V} = \sigma T^4$ ，其中 Stefan 常数 σ 可以用自然界基本常数表达为：

$$\sigma = \frac{\pi^2 k_B^4}{60c^2 \hbar^3} . \quad (3.37)$$

具体的数值计算得到的 σ 的数值与实验上观测到的高度吻合。

显然，如果将能均分定理（或者说 Rayleigh-Jeans）的结果 (3.35) 对圆频率积分，将得到发散的结果。这一点被厄任菲斯特（P. Ehrenfest, 1880 – 1933）称为 **紫外灾难**¹²，它说明对于辐射场应用经典统计将得到多么荒谬的结果。其实瑞利、金斯包括普朗克等人并没有认为经典统计的结果可以应用于高频的辐射场。因此，从这个角度来讲，当时（1900 年左右）并没有什么紫外灾难之说。但不可否认的是，Ehrenfest 后来的论证¹³使得许多人相信：经典统计的确已经不能应用在辐射场这样的对象上了。经典的物理理论必须变革，代之以全新的量子理论。

¶ 黑体辐射的普朗克公式的最为完美的验证并不是来自地球上的任何一个实验室，而来自空间卫星关于宇宙微波背景辐射的观测。¹⁴ 宇宙微波背景辐射，或者简称为 CMB (Cosmic Microwave Background)，可能是迄今为止宇宙大爆炸的最直接的实验证据。大约在大爆炸发生后 30 万年左右，宇宙冷却到原子可以稳定的存在。这时，宇宙空间中原先存在的大量的自由电子被原子所俘获，组成了电中性的、稳定的原子。从这时起，原先与自由电子剧烈反应（康普顿散射）的光子这时几乎可以自由地穿过宇宙空间而不被散射。也就是说，仅仅从这时起，宇宙才“混沌初开”，开始变得“透明”起来。这个现象在宇宙学中称为光子的 **脱耦**。光子脱耦的时刻对于宇宙中的光子来说是一个十分重要的时刻。从这时起，宇宙中的光子气开始可以用自由光子气体（或者说黑体辐射）来很好地描述了。在微波背景辐射形成的时候，宇宙的温度（或者说微波背景辐射所对应的黑体辐射的温度）还是很高的。随着宇宙的继续膨胀，到现在这个温度已经降低到大约 2.7K 左右。来自卫星的辐射谱与普朗克的黑体辐射公式所预言的 **惊人地** 相符。其符合程度足以将微波背景辐射的温度确定到 4 ~ 6 位有效数字。这种符合程度不仅仅超过了当年普朗克的同事们验证普朗克公式的程度，而且远远超过了任何目前人工实验室中能达到的程度。因为，任何的人工实验装置都仅仅是理想黑体辐射的一个近似，人们无法将外界的各种干扰忽略。而宇宙的特点就是，它已经是一切，不存在宇宙之外的干扰。当然，CMB 提供给我们的物理信息远远不止验证黑体辐射这么简单，事实上，它可以看成是宇宙年轻时

¹²P. Ehrenfest, 玻耳兹曼的学生。一个伟大的物理学家，一个不幸的灵魂。

¹³这个词最早出现在 1911 年 P. Ehrenfest 的论述之中，尽管他可能这之前就已经有类似的讨论了。

¹⁴早期的观测卫星有 COBE 卫星，最近又有 WMAP, PLANCK 卫星的观测。

(只有 30 万年, 与目前宇宙 100 多亿年比较, 可以称为年轻) 留下的一张旧照片, 它可以为我们研究早期宇宙的各种物理性质提供重要的实验证据。

16 固体的热容量

¶ 在经典的麦克斯韦-玻耳兹曼统计理论中, 一个重要的结果就是能均分定理。这个定理应用于固体热容量, 就得到了著名的杜隆-珀替定律, 它告诉我们: 固体的摩尔热容量是一个常数 $3R$ 。实际固体的热容量在高温时都满足这个定律; 但在低温时, 几乎都不满足这个定律。造成这个定律被违反的原因就是量子效应。

历史 3.1 历史上, 爱因斯坦是第一个研究固体(低温)热容量量子理论的人。他在早期量子论方面的工作首先是在光量子方面(1905年), 他提出了光是量子化的观念, 解释了光电效应。随后一年, 也就是1906年, 爱因斯坦提出了他的固体热容量的量子理论, 第一次解释了固体热容量在低温时为什么会偏离经典的能均分定理(或杜隆-珀替定律)的预言。这个工作的伟大意义就在于, 它诞生于量子论的初期, 远早于量子力学的诞生(1925年), 甚至远早于玻尔的原子模型的诞生(1916年)。在1906年, 量子理论还处于它的萌芽期。多数人都对量子理论抱悲观和怀疑的态度。实际上, 就连提出量子论的普朗克当时都对于量子论抱悲观态度。普朗克非常欣赏爱因斯坦的狭义相对论, 并且是狭义相对论的积极支持者。但他对于爱因斯坦在量子论方面的尝试曾经发表了非常悲观的评论。直到1913年, 他还认为爱因斯坦发展量子理论是爱因斯坦的一个“失误”。由此可以看出量子论在萌芽期的发展是多么地艰难。这也从一个侧面反映了爱因斯坦的固体热容量的量子理论的伟大意义。可以毫不夸张地说, 爱因斯坦的固体热容量量子理论是当时人们所知道的寥寥无几的量子理论之一。

爱因斯坦假设固体中的原子的振动的 $3N$ 个独立简正模式的振动频率都相等, 将其圆频率记为 ω 。这些振子的能量为:¹⁵

$$\varepsilon_n = \left(n + \frac{1}{2} \right) \hbar\omega . \quad (3.38)$$

由于原子只是做小振动, 它们是定域的、可分辨的。用我们现在的的话来说, 爱因斯坦将整个固体系统看成是 $3N$ 个可分辨的振子组成的近独立子系。对应于一个振子的子系配分函数为:

$$z = \sum_{n=0}^{\infty} e^{-\beta\hbar\omega(n+1/2)} = \frac{e^{-\frac{\beta\hbar\omega}{2}}}{1 - e^{-\beta\hbar\omega}} . \quad (3.39)$$

整个系统的正则配分函数 $Z = z^{3N}$, 因此固体的内能为:

$$U = -3N \frac{\partial}{\partial \beta} \log z = 3N \frac{\hbar\omega}{2} + \frac{3N\hbar\omega}{e^{\beta\hbar\omega} - 1} . \quad (3.40)$$

¹⁵这个公式中的 $1/2$ 实际上在爱因斯坦的原先讨论中是没有的。不过这样一个常数仅仅影响能量的零点的选取, 对于热容量没有影响。

由此，我们计算出固体的热容量为：

$$C_V = \left(\frac{\partial U}{\partial T} \right)_V = 3Nk_B \left(\frac{\hbar\omega}{k_B T} \right)^2 \frac{e^{\beta\hbar\omega}}{(e^{\beta\hbar\omega} - 1)^2} . \quad (3.41)$$

于是，我们可以定义一个系统的特征温度：

$$\theta_E \equiv \frac{\hbar\omega}{k_B} , \quad (3.42)$$

它称为固体的 **爱因斯坦温度**。我们看到，当温度足够高时，即 $\theta_E \ll T$ 时，爱因斯坦理论预言的固体热容量趋于常数 $3Nk_B$ ，与杜隆-珀替定律刚好一致；反之，当温度很低时， $T \ll \theta_E$ ，爱因斯坦理论预言的固体热容量指数地趋于零。在图 3.1 中，我们示意性地画出了爱因斯坦理论中的固体热容量随温度的变化。实验上，固体的低温热容量的确随温度

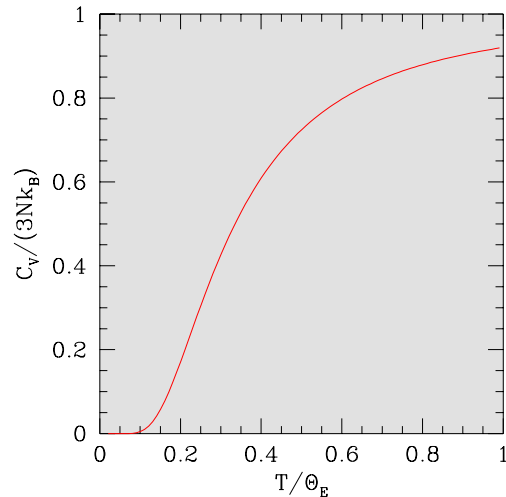


图 3.1: 爱因斯坦关于固体热容量的量子理论的预言。当温度很高时，它回到经典的杜隆-珀替定律；当温度很低时，它指数地趋于零。这是人类关于固体热容量的第一个量子理论。

趋于零而趋于零，但是并不象爱因斯坦理论所预言的这么快（指数地），而是按照温度的幂次趋于零。造成这种不一致的原因是因为爱因斯坦假设固体中的原子的 **所有** 振动模式的频率都是相同的。这是一个过于简化的假设。实际上固体振动的频率是有一个连续分布的。德拜 (P.J.W. Debye) 理论正是对于这一点进行了改进。但考虑到爱因斯坦理论是人类历史上早期的几个量子理论之一，它在物理学上的意义已经远远超出了仅仅解释固体的低温热容量这一范畴。因此，尽管爱因斯坦理论在定量地解释固体热容量上并没有德拜理论成功，但是它在物理学中的重要意义应当说远在德拜理论之上。

¶ 爱因斯坦引入量子化的概念从而定性地解释了固体热容量在低温时趋于零地事实，尽管定量上并不十分成功。真正的定量上比较成功的理论是由德拜 (P.J.W. Debye, 1884

— 1966) 完成的, 不过那又是爱因斯坦提出固体热容量理论六年以后 (1912 年) 的事情了。

历史 3.2 德拜 (Peter Joseph William Debye), 荷兰物理学家、化学家, 后来加入美国籍。1884 年生于荷兰的 Maastricht, 后来在 Aachen 工业大学学习深造并获得电机工程师。1906 年随导师 Sommerfeld 去慕尼黑大学并在那里获博士 (1908 年)。1911 年接替爱因斯坦任苏黎士大学理论物理教授。后来又在德国、荷兰等多所大学任教。1934 年, Debye 创立了 Max Planck 物理研究所。由于他不肯加入德国籍与纳粹不合, 1940 年去美国任 Cornell 大学化学系主任, 后入美国籍。1966 年病逝。德拜的贡献不仅仅在固体热容量方面, 他还对于 X 射线衍射学、低温物理 (他最先提出绝热去磁冷却的方法并获得 1K 以下的低温)、分子极化、溶液理论等方面都做出了卓越贡献。他因此获得了 1936 年的诺贝尔化学奖。我们在第 22 节中会介绍他关于稀薄等离子体的统计理论。这个理论可以应用在化学溶液中。

当时, 德拜刚刚接替爱因斯坦任苏黎士大学理论物理教授, 便开始研究如何改进爱因斯坦的固体热容量理论。他首先注意到, 固体中的振动实际上与固体中的 **声波** 相对应, 因此, 振动的频率与波长 (或者说波数) 是相互联系的。所以, 固体振动的频率也不是一个固定的常数, 而是随波长 (或者说波矢量的大小) 有一个分布。当我们将固体中的声波量子化以后, 我们可以讨论激发的量子, 它们被称为 **声子**; 就象在黑体辐射中我们可以讨论光子一样。于是, 振动对于热容量的贡献可以等价的变为讨论声子气体的热容量问题。声子也是玻色子, 它的数目也是不守恒的, 所以一切的讨论都和我们上一节中光子气体的讨论类似。唯一的不同在以下几个方面: 首先, 声子的能谱和偏振与光子不同。对应于一个给定的波矢 \mathbf{k} , 三维固体中的声波可以有三个不同的偏振方向: 两个横偏振和一个纵偏振。这两个不同偏振的声波在固体中的传播速度不同:

$$\omega = c_l k, \quad \omega = c_t k, \quad (3.43)$$

其中 c_l 和 c_t 分别是固体中的纵声波声速和横声波声速。¹⁶ 一般说来, $c_l > c_t$, 即纵波比横波要快。所以, 在圆频率 ω 到 $\omega + d\omega$ 之间声子 (包括横声子和纵声子) 的总量子态数目是:

$$g(\omega)d\omega = \frac{V}{2\pi^2} \left(\frac{1}{c_l^3} + \frac{2}{c_t^3} \right) \omega^2 d\omega, \quad (3.44)$$

这个数目也就是频率处在 ω 到 $\omega + d\omega$ 之间的固体的简谐振动的模式数。因此, 与黑体辐射不同的是, 与固体所对应的总振动自由度数目不是无穷大, 而应当等于三维固体的总的振动自由度的数目: $3N$, 其中 N 是构成固体的原子的数目。所以我们必须有:

$$\int_0^{\omega_D} g(\omega)d\omega = 3N, \quad (3.45)$$

¹⁶ 固体中的声速可以通过固体的弹性模量和密度计算出来。例如: $c_l = \sqrt{E/\rho}$, 其中 E 是固体的杨氏模量, ρ 是固体的密度。

其中 ω_D 是一个截止频率, 称为 **德拜截止频率**。于是, 固体的内能可以写成:

$$U = U_0 + \int_0^{\omega_D} g(\omega) \frac{\hbar\omega}{e^{k_B T} - 1} d\omega \quad (3.46)$$

现在我们引入无量纲变量:

$$y = \frac{\hbar\omega}{k_B T}, \quad x = \frac{\hbar\omega_D}{k_B T} \equiv \frac{\theta_D}{T}, \quad (3.47)$$

式中的 θ_D 称为固体的 **德拜温度**, 它的数值必须由理论预言与实验相拟合得到。再引

入 **德拜函数** $D(x)$:

$$D(x) = \frac{3}{x^3} \int_0^x \frac{y^3 dy}{e^y - 1}. \quad (3.48)$$

于是, 固体的内能可以十分简洁地写成:

$$U = U_0 + 3Nk_B T D(x). \quad (3.49)$$

当温度很高时, $T \gg \theta_D$ 或者说 $x \ll 1$ 时, 我们发现 $D(x) \sim 1$ 。所以, 当温度很高时, 固体的内能以及热容量就回到经典的杜隆-珀替定律的结果, 即:

$$U = U_0 + 3Nk_B T, \quad C_V = 3Nk_B. \quad (3.50)$$

在另一个极限下, 也就是说温度很低时, $x \gg 1$, 我们可以将德拜函数定义中的积分上限移到无穷大, 所以 $D(x) \sim (\pi^4/(5x^3))$ 。所以我们得到:

$$U = U_0 + 3Nk_B \frac{\pi^4 T^4}{5 \theta_D^3}, \quad C_V = 3Nk_B \frac{4\pi^4}{5} \left(\frac{T}{\theta_D}\right)^3, \quad (3.51)$$

这就是著名的德拜 T^3 律, 它告诉我们固体的低温热容量随温度的三次方趋于零。这个预言对于非金属固体非常好地与实验符合。金属固体的热容量在 $3K$ 以上时, 也符合得很好, 当温度进一步降低时, 由于传导电子的贡献变得重要 (参见第 17 节), 所以晶格振动的热容量不能完全解释固体的热容量, 必须要加上传导电子的贡献。

¶ 最后需要说明的是, 德拜理论仍然是对于固体中声子能谱的一种 **简化** 处理, 它在长波极限时很好地与实验吻合。但是对于波长较短的固体振动, 并不能很好地描述。这方面, 可以采用更为复杂的理论模型的计算。另外, 在实验上, 可以通过 **中子散射实验** 来测定固体中声子的态密度。这些属于固体物理中较为专门的课题, 有兴趣的读者可以参考固体物理学中的有关书籍。

¶ 象声子这样的物理对象与我们前一节讨论的光子有很多相似的地方, 但是它们也有着本质的区别。声子并不是所谓的基本粒子。它只是晶格振动的一种体现。它被称

为准粒子，或者称为元激发。这个概念在凝聚态物理的多个方面有着广泛的应用。许多系统的动力学自由度本身之间具有很强的相互作用，例如：固体的各个原子之间相互作用是很强的。但是从这样具有强相互作用的基态上激发出的准粒子（例如声子）却往往相互作用不强，在低温时它们往往可以看成是无相互作用的理想气体。因此，可以相当好地用理想气体模型来近似。这样的概念也可以用来讨论更为复杂的量子流体（例如液氦、费米液体等等）。

17 理想费米气体

这一节中我们讨论理想费米气体的性质。我们将首先讨论弱简并的理想费米气体，虽然这并没有什么实际的物理系统相对应，但它有助于揭示由于统计所带来的关联；然后我们讨论强简并的费米气体（Sommerfeld 理论），多数金属中的电子在常温下可以近似地看成是强简并的理想费米气体。

17.1 弱简并的理想费米气体

与讨论玻色气体的情形类似，我们假定我们研究的理想费米气体是非相对论性的。同时，我们首先考虑弱简并的情形，即： $e^{-\alpha} \ll 1$ 。由于粒子的平动动能可以看成是准连续的，体系的单粒子量子态的态密度仍然由公式 (3.2) 给出。我们将对于单粒子态的求和换成积分，理想费米气体的巨配分函数的对数为（参见第 12 节中的公式 (2.69)）：

$$\log \Xi = \frac{2\pi(2s+1)V}{h^3} (2mk_B T)^{3/2} \int_0^\infty \log(1 + e^{-\alpha-x}) x^{1/2} dx . \quad (3.52)$$

由于是弱简并，我们可以将对数函数展开成级数并逐项积分得到：

$$\log \Xi = \frac{(2s+1)V}{h^3} (2\pi mk_B T)^{3/2} \sum_{j=1}^{\infty} (-)^{j-1} \frac{e^{-j\alpha}}{j^{5/2}} , \quad (3.53)$$

$$\bar{N} = \frac{(2s+1)V}{h^3} (2\pi mk_B T)^{3/2} \sum_{j=1}^{\infty} (-)^{j-1} \frac{e^{-j\alpha}}{j^{3/2}} , \quad (3.54)$$

$$U = -\frac{\partial}{\partial \beta} \log \Xi = \frac{3}{2} k_B T \log \Xi , \quad (3.55)$$

$$p = \frac{1}{\beta} \frac{\partial}{\partial V} \log \Xi = \frac{2U}{3V} \quad (3.56)$$

$$S = k_B \left(\frac{5}{2} \log \Xi + \bar{N} \alpha \right) . \quad (3.57)$$

类似于理想玻色气体的情形，我们引进变量 y ：

$$y = \frac{\bar{N} h^3}{(2s+1)V(2\pi mk_B T)^{3/2}} , \quad (3.58)$$

于是, 当 y 很小时, 我们可以解出:

$$e^{-\alpha} = y + \frac{1}{2^{3/2}}y^2 + \left(\frac{1}{4} - \frac{1}{3^{3/2}}\right)y^3 + \left(\frac{1}{8} + \frac{5}{8^{3/2}} - \frac{5}{6^{3/2}}\right)y^4 + \dots, \quad (3.59)$$

并带入热力学公式 (3.56) 得到:

$$\begin{aligned} \frac{pV}{Nk_B T} &= \frac{2U}{3Nk_B T} = \frac{\log \Xi}{N} \\ &= 1 + \frac{1}{2^{5/2}}y - \left(\frac{2}{3^{5/2}} - \frac{1}{8}\right)y^2 + \left(\frac{3}{32} + \frac{5}{2^{11/2}} + \frac{3}{6^{3/2}}\right)y^3 - \dots \end{aligned} \quad (3.60)$$

注意, 这个公式与玻色子的相应公式 (3.16) 极为类似, 只要将 y 的符号换一下就可以从一个公式得到另一个。与玻色子类似, 我们再一次看到了所谓的 **统计关联** 现象: 理想费米气体的压强, 比具有同样温度和密度的经典理想气体的压强要大, 仿佛构成理想费米气体的粒子之间有一个等效的排斥作用, 这种等效排斥作用实际上源自泡利不相容原理。

真实世界中弱简并的理想费米气体的例子并不多, 也许在宇宙早期温度很高时的正负电子气体可以近似看成是弱简并的理想费米气体。我们日常见到的, 例如所有金属在常温时的情况, 实际上都可以近似地看成强简并的理想费米气体。

17.2 强简并的理想费米气体

¶ 作为强简并理想费米气体的一个例子, 我们在这一小节中讨论金属中的电子气体。将金属中的电子看成是 **自由的** 电子气体本身是一个近似。这种近似显然只能够用于金属中的 **巡游电子**。一般来说, 构成金属的原子的内层电子不会是巡游电子。它们被原子实紧紧束缚着, 显然不能够看成是自由电子。如果我们仅仅考虑巡游电子, 同时我们作为一个初步的近似, 忽略巡游电子与原子实之间的相互作用, 同时忽略巡游电子之间的库仑相互作用, 我们就得到了一个关于金属中巡游电子的理想费米气体模型, 它被称为 **索莫非 (Sommerfeld) 模型**。¹⁷ 索莫非模型的意义在于, 它可以看成是固体电子论的最初级近似。如果在它的基础上在考虑电子与晶格的相互作用, 我们就可以得到 **能带论**; 如果再考虑了电子-电子相互作用, 我们实际上就得到了一个强关联的多体理论 (一个非相对论性的量子场论)。

简单的数量级估计就可以发现, 在常温以下的多数金属中的巡游电子都是强简并的。例如, 如果我们利用上一节中的 y 的定义, 我们不难估计出: $y \sim 10^7/T^{3/2}$, 其中温度 T

¹⁷你可能会觉得这样的近似方法近乎“疯狂”。的确, 如果你对金属中电子与原子实, 或者是电子与电子之间的库仑相互作用做一个数量级的估计, 你会发现它们是相当大的。但是另一方面, 索莫非模型的确可以解释金属的许多性质。这件事一直困扰了物理学家相当多年。

以 K 为单位，并且我们利用了典型的金属的密度。于是，当温度大约是 $300K$ 时， y 的数值大约是 3400，远远大于 1。所以，在一般温度下，金属中的巡游电子 **必须** 看成是强简并的费米气体来处理。这个工作首先是索莫非完成的。他利用量子统计的费米分布，首先讨论了金属中的巡游电子的物理性质，相当完美地对金属的多种性质给出了合理的解释。

历史 3.3 Sommerfeld 主要澄清了当时的金属理论（称为 Drude 模型）与实验上的矛盾。Drude 模型是一个基于 **经典统计** 的自由电子气体模型。它与 Sommerfeld 模型的区别就在于不是利用费米分布，而是用经典的麦克斯韦 - 玻耳兹曼分布。按照这个模型，金属的电导率 σ 、热导率 κ 满足一个简单的关系： $\kappa/(\sigma T) = 2k_B^2/e^2$ ，其中 e 为电子电荷。这个定律被很多实验所证实，被称为 Wiedermann-Franz 定律。但唯一的遗憾是，实验上测出的比值要比 Drude 模型预言的大 50% 左右。Sommerfeld 首先运用费米分布研究了这个问题，他证明经典的 Drude 模型的预言是错误的，比例系数应当是： $\kappa/(\sigma T) = \pi^2 k_B^2/(3e^2)$ ，恰好比原先的结果大约大 50%。我们知道如果一个理想费米气体可以看成是弱简并的，那么利用经典的麦克斯韦 - 玻耳兹曼分布将是一个不错的近似。但是，正如我们上面估计的，通常温度下的金属中的巡游电子一定是强简并的。这就是为什么索莫非模型与 Drude 模型有着巨大的区别。

¶ 如果温度为零，那么费米分布随能量的依赖关系其实十分简单，就是一个阶梯函数。具体的说，当能量小于化学势时，所有的量子态上的粒子数为 1，而当能量大于化学

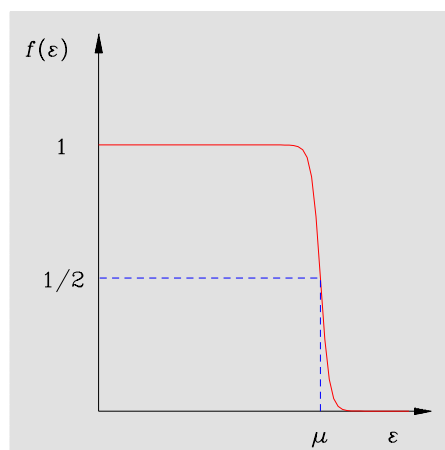


图 3.2: 低温时费米分布函数随能量的变化。在能量远离化学势时，分布函数仍然接近于阶梯函数；在化学势附近 $k_B T$ 的范围内，分布函数由 1 变为 0。

势时，所有的量子态上的粒子数为 0。当温度比较低时，也就是 $\mu/(k_B T) \gg 1$ ，费米分布的分布函数基本上仍然类似于一个阶梯函数。只是在 $|\varepsilon - \mu| \sim k_B T$ 的范围内分布函数从 1 快速地变为 0，参见图 3.2。

首先研究一下零温时的情形是很有帮助的。在零温时，所有能量低于化学势的系统的量子态都被填充。所以，零温时的化学势 μ_0 与系统的粒子数有简单的关系：

$$N = \int_0^{\mu_0} \frac{4\pi V}{h^3} (2m)^{3/2} \varepsilon^{1/2} d\varepsilon, \quad (3.61)$$

注意这里的态密度是公式 (3.2) 的两倍是因为我们考虑了电子的自旋， $(2s + 1) = 2$ 。将上式积分出来，我们就得到了零温时的化学势 μ_0 ：

$$\mu_0 \equiv \varepsilon_F = \frac{\hbar^2}{2m} (3\pi^2 n)^{2/3}, \quad (3.62)$$

其中，我们引进了符号 ε_F ，它称为费米能量，来表示零温时的化学势。式中 $n = N/V$ 代表自由电子的数密度。

费米能量是一个非常重要的概念。由于 Pauli 不相容原理的影响，零温时全同费米子系统的基本图像是，每个粒子都试图从最低可能的单粒子能级一直向上填充，直至某个能量 ε_F 。高于该能量的单粒子能级都是空的，而低于该能量的单粒子能级都是填满的。粒子所填充到的最高的单粒子能量就是费米能量，又称为费米面（类比于海平面）。对于自由的非相对论性的费米子来说，其动量空间（或者 \mathbf{k} 空间）的等能面为一球面。因此人们又称之为费米球。¹⁸ 费米球的“半径”称为费米动量：

$$p_F \equiv \hbar k_F = \hbar (3\pi^2 n)^{1/3}. \quad (3.63)$$

其中 k_F 称为费米波矢。类似定义的量还有所谓的费米速度： $v_F \equiv p_F/m$ 。因此费米能量可以等价地写为： $\varepsilon_F \equiv p_F^2/(2m) \equiv \hbar^2 k_F^2/(2m) \equiv (1/2)mv_F^2$ 。

与玻色子体系不同的是，零温时，理想费米气体的能量不为零，而是等于：

$$U = \int_0^{\mu_0} \frac{4\pi V}{h^3} (2m)^{3/2} \varepsilon^{3/2} d\varepsilon = \frac{3}{5} N \varepsilon_F. \quad (3.64)$$

如果我们将典型的金属中的自由电子的数密度、电子质量等常数带入，不难估计出零温时金属中电子的费米能量的数量级。我们再次发现：即使在常温时， $\varepsilon/(k_B T) \sim 300$ ，这说明，常温时的热激发能量远小于费米能量，所以这个理想费米子系统是强简并的。同样的，零温时费米气体的压强也不为零，事实上我们可以用

$$p = \frac{2U}{3V} = \frac{2}{5} n \varepsilon_F, \quad (3.65)$$

¹⁸如果考虑金属中电子与原子实构成的晶格之间的相互作用，那么电子的单粒子能级 $\varepsilon(\mathbf{k})$ 将不再是球面，而是波矢 \mathbf{k} 的一个复杂的函数。它称为该粒子的能带。这时费米面的概念仍然可以运用，只是它不再是球面。具体的例子可以参见固体物理。

得到体系的压强。这个压强一般称为费米气体的简并压。

下面我们讨论温度不为零的情形。如前所说，我们将讨论温度不太高的情形，也就是说仍然有： $\varepsilon_F \gg k_B T$ 。这时，我们可以做一个低温展开。系统的粒子数和能量可以写成：

$$N = \frac{4\pi V}{h^3} (2m)^{3/2} \int_0^\infty \frac{\varepsilon^{1/2} d\varepsilon}{e^{\frac{\varepsilon-\mu}{k_B T}} + 1}, \quad (3.66)$$

$$U = \frac{4\pi V}{h^3} (2m)^{3/2} \int_0^\infty \frac{\varepsilon^{3/2} d\varepsilon}{e^{\frac{\varepsilon-\mu}{k_B T}} + 1}, \quad (3.67)$$

我们发现，这些物理量都可以写成下述的积分形式：

$$I = \int_0^\infty \frac{\eta(\varepsilon) d\varepsilon}{e^{\frac{\varepsilon-\mu}{k_B T}} + 1}, \quad (3.68)$$

其中对于粒子数和能量，函数 $\eta(\varepsilon)$ 分别为 $c\varepsilon^{1/2}$ 以及 $c\varepsilon^{3/2}$ ，而 $c = \frac{4\pi V}{h^3} (2m)^{3/2}$ 是一个常数。

Sommerfeld 首先证明了，积分 (3.68) 在低温时，具有下列的展开式：¹⁹

$$I = \int_0^\mu \eta(\varepsilon) d\varepsilon + \frac{\pi^2}{6} (k_B T)^2 \eta'(\mu) + \frac{7\pi^4}{720} (k_B T)^4 \eta'''(\mu) + \dots, \quad (3.69)$$

这个展开式因此被称为索莫非展开。利用索莫非展开，我们就得到了有限温度时粒子数和能量的表达式：

$$N = \frac{2}{3} c \mu^{3/2} \left[1 + \frac{\pi^2}{8} \left(\frac{k_B T}{\mu} \right)^2 \right], \quad (3.70)$$

$$U = \frac{2}{5} c \mu^{5/2} \left[1 + \frac{5\pi^2}{8} \left(\frac{k_B T}{\mu} \right)^2 \right]. \quad (3.71)$$

上面的第一个式子给出了温度不为零时的化学势，将常数 c 的数值带入，我们得到：

$$\mu = \mu_0 \left[1 - \frac{\pi^2}{12} \left(\frac{k_B T}{\mu_0} \right)^2 \right]. \quad (3.72)$$

将此式带入能量的表达式 (3.71)，我们得到：

$$U = \frac{3}{5} N \mu_0 \left[1 + \frac{5\pi^2}{12} \left(\frac{k_B T}{\mu_0} \right)^2 \right]. \quad (3.73)$$

¹⁹也许有的同学注意到了，这个展开式的系数是所谓的伯努利数 (Bernoulli numbers)。如果你没听说过这个名称，就不必介意了，反正不影响什么。

这个式子的一个重要的结果就是巡游电子对于金属 **热容量** 的贡献：

$$C_V = \left(\frac{\partial U}{\partial T} \right)_V = Nk_B \frac{\pi^2}{2} \left(\frac{k_B T}{\mu_0} \right), \quad (3.74)$$

即巡游电子在低温时对金属热容量的贡献与绝对温度成正比。由于在常温下 $k_B T \ll \mu_0$ ，所以在 **常温** 下，传导电子的热容量的贡献比起离子振动的热容量要小很多，可以忽略。但是在低温时，按照德拜理论，离子振动的热容量随 T^3 趋于零，而巡游电子的热容量仅仅按照 T 趋于零。所以在 **足够低** 的温度时（对于多数金属这个温度大约是几个开尔文），巡游电子对金属热容量的贡献就开始变得重要了。公式 (3.74) 与经典统计结果的区别是十分显著的。按照经典统计的结果，传导电子也要参与能均分。所以，巡游电子对于金属热容量的贡献应当数量级为 Nk_B 。但在实验上金属的热容量在常温时很好地符合杜隆-珀替定律，根本看不到巡游电子的贡献，似乎金属热容量几乎完全由离子振动给出。这在经典的麦克斯韦-玻耳兹曼统计中是无法解释的。Sommerfeld 首先运用量子的费米分布，解释了这个现象的原因。按照费米分布，并不是所有的电子都可以被热激发的。如果温度不够高， $k_B T$ 远小于费米能量时，只有在费米能量附近的、 $k_B T/\mu_0$ 比例的电子可以被热激发从而参与能均分。其余的绝大部分电子都被深深地埋在费米能量以下。在那里，所有量子态都被占据，因此，泡利不相容原理阻止了这些电子进行小能量差的跃迁。所以，对于金属热容量有贡献的只是费密面附近的那“一小撮”电子，从而巡游电子对金属热容量的贡献的数量级应当是 $Nk_B(k_B T/\mu_0)$ ，这个估计与严格的公式 (3.74) 完全相符。

17.3 磁场中的强简并理想费米气体

¶ 本小节中我们讨论在外加磁场中的强简并理想费米气体的磁性质。我们讨论理想费米气体近似下的强简并电子气。这个系统具有十分重要的意义，它可以作为金属或半导体中磁性质的一个近似描述，虽然真实的金属或半导体中的磁性质的贡献不仅仅局限于电子气。为明确起见，本小节中我们假定外加的均匀磁场沿着 $+z$ 的方向且大小为 B_0 。

在一个弱的外磁场中的电子的磁性质会存在两方面的贡献：一方面，每个自由电子都具有正比于其自旋角动量的磁矩（参见第 7.1 小节的公式 (2.10)）。这些磁矩在外磁场的作用下会倾向于顺着磁场而排列，这就导致所谓的 **泡利顺磁性**。另一方面，我们知道在磁场中运动的经典的电子会按照其经典轨道做螺旋运动。这种螺旋运动经过量子化后的结果就是电子会占据一系列类似谐振子的能级，这些能级被称为 **朗道能级**。朗道能级上的电子实际上会产生抗磁性，这被称为 **朗道抗磁性**。

¶ 泡利顺磁性的理论很容易用我们前一小节给出的强简并理想费米气体的图像给出。在没有加外磁场时，按照前面的公式 (3.61)，自由电子的能态密度 $g(\varepsilon)d\varepsilon$ 为：

$$g(\varepsilon)d\varepsilon = \frac{4\pi V}{h^3} (2m)^{3/2} \varepsilon^{1/2} d\varepsilon$$

注意这个态密度是自旋向上和自旋向下电子态密度之和。换句话说，如果我们仅仅考虑一个自旋取向的自由电子态密度，我们有：

$$g_+(\varepsilon) = g_-(\varepsilon) = \frac{2\pi V}{h^3} (2m)^{3/2} \varepsilon^{1/2} \quad (3.75)$$

而总的态密度是两者之和： $g(\varepsilon) = g_+(\varepsilon) + g_-(\varepsilon) = 2g_+(\varepsilon) = 2g_-(\varepsilon)$ 。

当我们加上一个沿着 z 轴的外磁场时，这时自旋向上和向下的电子的能级会发生劈裂：

$$\varepsilon_{\pm} = \frac{\mathbf{p}^2}{2m} \pm \mu_B B_0. \quad (3.76)$$

其中 \pm 分别对应于自旋向上（顺着磁场）、向下（逆着磁场）的能量。²⁰ 换句话说，自旋向下的电子具有稍小的能量。由于 $\mu_B B_0$ 为常数，因此自旋向上或向下的电子的态密度仍然由 (3.75) 给出，只不过其宗量分别有一个向下、向上的平移罢了。作为零温时的电子气系统，它具有一个共同的化学势 μ 。因此我们有：

$$N_{\pm} = \int_{\pm\mu_B B_0}^{\mu} g_{\pm}(\varepsilon \mp \mu_B B_0) d\varepsilon. \quad (3.77)$$

令积分变量 $\varepsilon' = \varepsilon \mp \mu_B B_0$ 后可以将上式等价地写成：

$$N_{\pm} = \int_0^{\mu \mp \mu_B B_0} g_{\pm}(\varepsilon) d\varepsilon \quad (3.78)$$

化学势 μ 仍然由下式确定：

$$N \equiv N_+ + N_- = \frac{4\pi V}{3h^3} (2m)^{3/2} \left[(\mu + \mu_B B_0)^{3/2} + (\mu - \mu_B B_0)^{3/2} \right]. \quad (3.79)$$

在一般的应用中，磁场往往是比较小的： $\mu_B B_0 \ll \mu$ 。因此在零级近似下，化学势（费米能） μ 与粒子数密度的关系与没有外磁场时完全一样，由公式 (3.62) 给出。下一级的修正至少是 $(\mu_B B_0 / \varepsilon_F)^2$ 的量级。只要磁场不是特别强，这一般都是可以忽略的。²¹

虽然磁场对总的粒子数影响极小，但是它会影响到自旋向上、向下的粒子数只差。这个差值正好给出体系的净磁矩，而单位体积内的磁矩就是磁化强度：

$$M = \frac{1}{V} (\mu_B) (N_- - N_+) \quad (3.80)$$

$$= \frac{4\pi}{3h^3} (2m)^{3/2} \mu^{3/2} 2 \cdot \frac{3}{2} \mu_B \left(\frac{\mu_B B_0}{\mu} \right) \quad (3.81)$$

$$= \frac{3n\mu_B^2}{2\varepsilon_F} B_0. \quad (3.82)$$

由此我们得到电子气的磁化率：

²⁰注意，电子的磁矩的方向与它的自旋角动量是正好相反的，因为它带负电！参见第 7.1 小节的公式 (2.10)。

²¹一个有用的估计是记住玻尔磁子的大小： $\mu_B \sim 5.79 \times 10^{-5} \text{ eV/T}$ 。也就是说对于一个特斯拉的磁场， $\mu_B B_0$ 也只有大约 10^{-5} 电子伏。但是一般金属中电子气的费米能大概是几个电子伏。

$$\chi_{\text{para.}} = \frac{3n\mu_B^2}{2\varepsilon_F} \quad (3.83)$$

这就是所谓的 **泡利顺磁性** (Pauli paramagnetism) 的磁化率。

泡利顺磁性仅仅考虑了电子的自旋引起的磁矩，没有考虑到电子轨道的影响。事实上，一个经典的、非相对论性的电子（假定忽略其自旋磁矩与外场的相互作用）在一个外加静磁场中的哈密顿量可以写为：

$$H = \frac{1}{2m} \left(\mathbf{P} - \frac{e}{c} \mathbf{A} \right)^2 \quad (3.84)$$

其中 \mathbf{A} 是描述磁场的 **矢势**， \mathbf{P} 为粒子的 **正则动量**， $\mathbf{p} = \mathbf{P} - (e/c)\mathbf{A}$ 则为其机械动量。²²

对于一个经典的粒子在磁场中的运动情况我们已经知道，它的运动模式为沿着磁场方向的匀速平移运动，再叠加在垂直磁场平面内的匀速圆周运动。该匀速圆周运动的频率，称为带电粒子的 **回旋频率** (cyclotron frequency) 的表达式为：

$$\omega_B = \frac{eB_0}{mc} \quad (3.85)$$

在量子化之后，沿磁场方向的匀速平动继续保持，而在垂直平面内的回旋运动变为一个具有圆频率 ω_B 的谐振子。该粒子的能级可以写为：²³

$$\varepsilon = \frac{\mathbf{p}_z^2}{2m} + \left(n + \frac{1}{2} \right) \hbar\omega_B \quad (3.86)$$

其中 \mathbf{p}_z 为粒子沿着磁场方向的机械动量，它是一个守恒的量子数； n 则标志了谐振子的能级。这就是著名的 **朗道能级** (Landau level)。

现在我们简单分析一下，在加上外磁场后量子的电子气对磁场的响应。原则上我们需要计算体系的巨配分函数的对数 $\ln \Xi$ ，或者等价地说巨势 $J = -k_B T \ln \Xi$ 。正如我们前面分析的，假定磁场不是很强，因此体系的粒子数密度与化学势（费米能）之间的关系基本维持不变。我们需要考虑的是 J 中包含磁场的部分的修正。体系的磁化率可以表达为：

$$\chi = \frac{1}{V} \frac{\partial^2}{\partial B_0^2} (k_B T \ln \Xi) \quad (3.87)$$

详细的计算给出 (细节参见 [4]):

$$k_B T \ln \Xi = \frac{2}{5} N \varepsilon_F \left[1 - \frac{5}{32} \left(\frac{\hbar\omega_B}{\varepsilon_F} \right)^2 \right] \quad (3.88)$$

这个表达式的首项就是没有磁场时的结果： $k_B T \ln \Xi = (2/3)U = (2/5)N\varepsilon_F$ ；第二项恰恰体现了磁场的影响。将其带入磁化率的表达式后我们得到：

²²我们看到虽然动能表达式仍然是 $\mathbf{p}^2/(2m)$ ，但是它与正则动量的关系则不同了。注意到在量子化的过程中，在相空间中与粒子的坐标形成固定体积相格的是正则动量而不是机械动量。

²³参见任何一本量子力学教材，例如 [6]。

$$\chi_{\text{dia}} = -n \frac{\mu_B^2}{2\varepsilon_F}, \quad (3.89)$$

其中的负号说明体系是抗磁的，这就是所谓的朗道抗磁性。对于一个自由电子气系统，泡利顺磁性和朗道抗磁性同时存在。系统的真正磁响应是两者的叠加。由于泡利顺磁性更大，因此总体上自由电子气体现出顺磁性。比较公式 (3.83) 和公式 (3.89) 我们发现，两者有一个十分简单的关系：

$$\chi_{\text{dia}} = -\frac{1}{3}\chi_{\text{para}}. \quad (3.90)$$

这个关系仅对自由电子气是正确的。对于固体中的电子气来说，电子与晶格的相互作用可以很好地用电子的有效质量 m_e^* 来刻画。这时上面的关系可以写为：

$$\chi_{\text{dia}} = -\frac{1}{3} \left(\frac{m_e}{m_e^*} \right)^2 \chi_{\text{para}}. \quad (3.91)$$

这个公式可以很好地描写固体能带中的电子气的磁性质。由于有效质量 m_e^* 可以显著地不同于真空中电子的质量 m_e ，因此朗道抗磁性的贡献可以超过泡利顺磁性的贡献从而使系统整体上体现出抗磁性。²⁴



相关的阅读

我们这里对于量子统计的讨论仅仅限于近独立的系统，或者说没有相互作用的量子系统。近年来随着玻色-爱因斯坦凝聚的实验进展，对于理想玻色气体的讨论显然是十分有意义的。对于费米气体的讨论，特别是强简并费米气体的讨论，我们仅仅讨论了金属中的电子气体。其它的应用还有很多，例如天体物理中的白矮星、中子星的问题，有兴趣的读者可以参考 [12] 中的讨论。

²⁴ 固体的磁性质则更为复杂，还可以包括其原子实的贡献。

第四章 经典流体

本章提要

- 理想气体及其热容量 (18)
- 混合理想气体 (19)
- 非理想气体的状态方程 (20)
- 液体的热力学性质 (21)
- 稀薄等离子体的统计性质 (22)

这

一章中，我们将着重讨论经典流体，它实际上包含了十分丰富的物理对象，例如我们日常生活中遇到的气体和液体。我们称之为经典，主要是指我们将利用经典的统计而不是量子统计。正如我们多次指出的，物质世界本质上是量子的。即使对于我们下面要讨论的最为经典的对象，例如经典理想气体，我们也会发现，要合理地解释经典理想气体的热容量，也需要量子力学的知识。在本章中，我们将首先利用正则系综的方法讨论一下经典理想气体，特别是它的热容量；然后，我们利用巨正则系综来讨论混合理想气体的性质。对于经典非理想气体的讨论是本章中的重点、也是难点之一，我们将介绍经典的 Mayer 集团展开理论。随后，我们将简单介绍液体的统计物理处理方法。随着近年来软凝聚态物理的发展，对这些内容做一些基础的介绍还是有意义的。最后，我们将讨论德拜关于强电介质溶液的理论。

18 经典理想气体

18.1 单原子分子理想气体

¶ 我们首先考虑由单原子分子组成的经典理想气体。一般来说，惰性元素原子组成的稀薄气体在常温下通常可以看成是单原子分子的理想气体。这类系统的能量仅仅包含所有粒子的平动动能：

$$E = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m}, \quad (4.1)$$

其中 \mathbf{p}_i 是第 i 个分子的平动动量； m 为该分子的质量。由于系统的能量是各个分子能量之和，系统的正则配分函数 Z 就是各个子系统配分函数 z 的乘积：

$$Z = \sum_S e^{-\beta E_S} = \left(\sum_s e^{-\beta \epsilon_s} \right)^N = z^N. \quad (4.2)$$

我们考虑气体分子处在一个宏观大小的容器中。我们可以估计出气体中每个分子的典型能级间隔大小：

$$\Delta \epsilon \sim \frac{\hbar^2}{mL^2},$$

只要对于不太低的温度和宏观的尺寸 L ，我们发现都有： $\Delta \epsilon \ll k_B T$ ，所以粒子的能量可以看成是准连续的。于是，我们可以将单原子分子理想气体的正则配分函数写成：

$$Z = \frac{1}{N! h^{3N}} \int e^{-\frac{\beta}{2m} \sum_{i=1}^N \mathbf{p}_i^2} d^3 \mathbf{r}_i d^3 \mathbf{p}_i. \quad (4.3)$$

需要指出的是，上面式子中的因子 $1/N!$ 是考虑到气体分子的全同性所引入的。它可以说是量子力学的全同性原理在经典极限下的一个“遗迹”。¹对于坐标的积分可以积出得到体积 V ，剩下的动量部分的积分都是高斯积分。完成这些积分我们得到：

$$Z = \frac{1}{N!} z^N, \quad z = V \left(\frac{2\pi m}{\beta h^2} \right)^{3/2}. \quad (4.4)$$

其中 z 是与一个分子相对应的子系配分函数。

¶ 按照压强的统计力学公式 (2.37)，很容易求出单原子分子理想气体的压强：

$$p = \frac{1}{\beta} \frac{\partial}{\partial V} \log Z = \frac{N k_B T}{V}, \quad (4.5)$$

¹这个因子首先是吉布斯首先引入的。他引入这个因子是为了解决所谓的“吉布斯佯谬”。如果没有这个因子，那么系统的熵将不是一个广延量从而会产生“吉布斯佯谬”。参见第 19 节中关于混合熵的讨论。

也就是说，这种气体满足理想气体状态方程。按照内能的统计物理公式 (2.35)，我们有：

$$U = -\frac{\partial}{\partial \beta} \log Z = \frac{3}{2} N k_B T , \quad (4.6)$$

与经典能均分定理的结论一致。由此我们看到单原子分子气体的摩尔热容量是常数： $3R/2$ 。类似地，系统的熵可以按照熵的统计物理公式 (2.40) 求出：

$$S = \frac{3}{2} N k_B \log T + N k_B \log \frac{V}{N} + \frac{3}{2} N k_B \left[\frac{5}{3} + \log \left(\frac{2\pi m k_B}{h^2} \right) \right] . \quad (4.7)$$

它满足熵是广延量的要求。

18.2 双原子分子理想气体及其热容量

¶ 下面我们来看一下双原子分子理想气体。这是一个比单原子分子复杂得多的系统。这时，系统的能量由分子的平动动能 $\varepsilon^{(t)}$ 、转动动能 $\varepsilon^{(r)}$ 和振动能 $\varepsilon^{(v)}$ 分别贡献：

$$E = \sum_{i=1}^N \varepsilon_i^{(t)} + \varepsilon_i^{(r)} + \varepsilon_i^{(v)} . \quad (4.8)$$

由于总能量是各分子能量之和，所以系统的正则配分函数仍然可以写成：

$$Z = \frac{1}{N!} z^N , \quad z = z^{(t)} \cdot z^{(r)} \cdot z^{(v)} , \quad (4.9)$$

其中 z 是一个分子的子系配分函数， $z^{(t)}$ 、 $z^{(r)}$ 和 $z^{(v)}$ 分别是与一个分子的平动、转动和振动对应的子系配分函数。平动部分的子系配分函数 $z^{(t)}$ 与单原子分子的子系配分函数 (4.4) 是相同的。需要注意到的一点是，系统配分函数 Z 对于体积的依赖仅仅来源于平动部分。因此，不需要对转动和振动部分进行任何具体的计算，我们立刻发现：双原子分子理想气体的压强仍然由理想气体状态方程给出。

¶ 与气体的压强相比，双原子分子气体的内能和热容量则要复杂许多。这时，必须考虑气体分子的转动和振动对热容量的贡献。气体的热容量问题一直是统计物理历史发展中倍受关注的问题，它也是验证统计物理理论的试金石。在 1860 年麦克斯韦提出他的著名的速度分布律时就发现，理论预言的气体热容量并不能完全解释所有实验的结果。但不容否认的是，理论预言的确与大多数的实验结果吻合。这件事一直困扰着经典统计物理学家们很久。一直到量子力学诞生，人们才认识到，那些理论与实验的不吻合完全是因为不恰当地应用了经典力学的结果。如果运用量子力学的方法，则所有理论的预言与实验高度地一致。而那些纯经典统计的理论预言与实验吻合得好的例子，恰恰是量子效应不显

著，我们期待经典物理应当适用的例子。下面我们将从量子的玻耳兹曼分布出发来研究双原子分子理想气体的热容量，并讨论何时可以回到经典的能均分定理的结果。

¶ 根据热力学公式，双原子分子理想气体的内能也可以写成平动、转动和振动部分的贡献之和：

$$U = -N \frac{\partial}{\partial \beta} \log z = U^{(t)} + U^{(r)} + U^{(v)} . \quad (4.10)$$

相应的，双原子分子理想气体的热容量也可分为三部分之和：

$$C_V = C_V^{(t)} + C_V^{(r)} + C_V^{(v)} . \quad (4.11)$$

我们前面已经分析过，只要气体处在宏观的容器中并且温度不太低，气体分子的平动能级总是可以看成是准连续的，所以完全可以用积分代替求和。因此气体分子的质心平动部分对热容量的贡献正如经典统计中能均分定理所预言的，即： $C_V^{(t)} = \frac{3}{2} N k_B$ 。

¶ 下面我们考虑双原子分子的振动。如果我们将双原子分子的振动看成是简谐振动，²设其圆频率为 ω ，按照量子力学，分子振动的能量是量子化的，因此振动部分的子系配分函数应当用对于所有振动态的求和而不是积分来计算。简单的计算表明，振动部分对于内能的贡献是：

$$U^{(v)} = \frac{N \hbar \omega}{2} + \frac{N \hbar \omega}{e^{\beta \hbar \omega} - 1} . \quad (4.12)$$

我们看到，振动部分对于内能的贡献依赖于振动的特征能量 $\hbar \omega$ 与 $k_B T$ 之比。在通常温度（不太高）下，由于对于一般的双原子分子都有 $\theta_v \equiv \hbar \omega / k_B \gg T$ ，其中 θ_v 称为分子振动的振动特征温度，其量级一般为 $10^3 K$ ，³远大于通常的室温。于是我们得到分子的振动对于热容量的贡献为：

$$C_V^{(v)} = N k_B \left(\frac{\theta_v}{T} \right)^2 e^{-\frac{\theta_v}{T}} , \quad (4.13)$$

也就是说，振动部分对于气体热容量的贡献指数地趋于零。在通常温度下，气体分子的振动对于其热容量的贡献是非常小的。这实际上是能级量子化后的典型行为，我们在第 16 节中讨论固体热容量的爱因斯坦理论是也得到类似的行为。从物理上讲，这意味着气体分子的振动能级的间隔 $\hbar \omega$ 远远大于 $k_B T$ ，所以，分子被激发到高振动能级的几率几乎为零。也就是说，所有的振动模式都被冻结在基态上，从而对于系统热容量没有贡献。事

²分子的振动是一个比较复杂的量子力学问题。在简谐近似下，它可以按照分子的对称性分为多个简正模式，每一个模式具有一定的频率。这些频率可以通过分子的光谱线来测定。对于双原子分子来说，它的振动只有一个自由度，因此只有一个振动本征频率。

³下面的数量级估计有助于我们了解分子的振动特征温度。一般来说，分子振动的特征能量为 $10^{-1} eV$ ，相当于 $10^3 K$ 左右。这就决定了多数气体分子的振动特征温度是这个数量级。

实上，对于大多数的气体分子，由于它们的振动特征温度都相当大，以至于当温度不太高时，振动自由度根本不参加能量均分，从而对气体的热容量没有明显贡献；如果温升高到 $10^3 K$ 以上时，这些分子往往又已经变得不稳定，可能分解为原子。因此对于这类分子构成的气体，它们的振动自由度完全对热容量没有可观的贡献。当然，例外也是有的。如果气体分子的某个振动模式的固有频率格外的低，这时，与该频率对应的振动特征温度也特别低。于是这种振动模式可以对气体的热容量有可观的贡献。

¶ 最后我们讨论双原子分子的转动对热容量的贡献。与分子的振动不同，分子的转动所对应的特征能量一般是小于室温时的 $k_B T$ 的。我们首先以异核的双原子分子（例如：HCl, CO 等）为例，来说明转动自由度对于气体热容量的贡献。按照量子力学，一个双原子分子转动的能级为：

$$\varepsilon^{(r)} = \frac{j(j+1)\hbar^2}{2I}, \quad (4.14)$$

其中 $j = 0, 1, \dots$ 称为 **转动量子数**，它标志了双原子分子的角动量； I 是分子的转动惯量。这个能级的简并度是 $2j+1$ 。所以，分子转动部分的子系配分函数为：

$$z^{(r)} = \sum_j (2j+1) e^{-\frac{j(j+1)\hbar^2}{2Ik_B T}} = \sum_j (2j+1) e^{-j(j+1)\frac{\theta_r}{T}}, \quad (4.15)$$

其中我们引入了分子的 **转动特征温度** $\theta_r \equiv \hbar^2/(2Ik_B)$ ，它依赖于分子的转动惯量。实验上，可以通过分子的光谱线的测量来确定。对于大多数双原子分子来说，它们的转动特征温度一般是 $10^0 - 10^1 K$ 的数量级。因此，在通常室温情形下，我们有 $T \gg \theta_r$ 。这意味着，分子的转动能级可以看成是准连续的。也就是说，上式中的求和可以用积分来替代。如果令 $x = (\theta_r/T)j(j+1)$ ，那么， $dx = (\theta_r/T)(2j+1) dj$ （因为 $dj = 1$ ）。所以，我们得到：

$$z^{(r)} = \int_0^\infty \frac{T}{\theta_r} dx e^{-x} = \frac{2I}{\beta \hbar^2}. \quad (4.16)$$

由此，我们得到双原子分子转动自由度对于其内能的贡献为：

$$\begin{aligned} U^{(r)} &= -N \frac{\partial}{\partial \beta} \log z^{(r)} = N k_B T, \\ C_V^{(r)} &= N k_B. \end{aligned} \quad (4.17)$$

这正是经典统计中能均分定理的结果：双原子分子的两个转动自由度各贡献 $\frac{1}{2} k_B T$ 的能量，或者说，各贡献 $\frac{1}{2} k_B$ 的热容量，所以经典统计中，双原子分子的转动自由度对摩尔热容量的贡献为 $N_A k_B = R$ 。

如果气体的温度比较低，接近气体的转动特征温度，那么转动配分函数 (4.15) 中分子转动能级的分立性将变得显著。这时我们不能用积分来代替求和，于是分子转动对于热容

量的贡献也将明显地偏离能均分的结果。这种偏离完全是由于分子转动能级的量子效应造成的，因此是利用纯经典物理无法解释的。对于分子转动配分函数 (4.15) 的数值计算表明，考虑了量子力学效应以后的结果完全可以与实验的结果很好地吻合。这就回答了当年困扰麦克斯韦和玻耳兹曼的问题。

¶ 如果双原子分子的两个原子核是相同的，例如 H_2 , O_2 等，那么量子的全同性原理还将起作用。其结果是会影响到双原子分子的转动量子数 j 的可能取值，使得它与双原子分子的两个原子核的核自旋位形发生关联。这可以说是一种十分奇特的关联现象。

以 H_2 分子为例，它可以有两种核自旋的位形，一种是两个氢核的核自旋平行，这种氢被称为正氢 (Orthohydrogen)；另一种是两个氢核的核自旋反平行，这种氢被称为仲氢 (Parahydrogen)。两种氢的分子相对稳定，很难互相转换。在通常自然界的氢气中，正氢占 3/4；仲氢占 1/4。全同性原理要求，仲氢分子的转动量子数 j 只能取偶数；正氢分子的转动量子数 j 只能取奇数。⁴因此，如果温度接近分子的转动特征温度，这时分子转动能级的分立性是可观的，正氢和仲氢的转动部分的配分函数也应当分别来求。我们用 $z_o^{(r)}$ 和 $z_p^{(r)}$ 来分别表示正氢和仲氢的转动部分的子系配分函数，于是我们有：

$$\begin{aligned} z_o^{(r)} &= \sum_{j=1,3,\dots} (2j+1)e^{-j(j+1)\frac{\theta_r}{T}}, \\ z_p^{(r)} &= \sum_{j=0,2,\dots} (2j+1)e^{-j(j+1)\frac{\theta_r}{T}}, \end{aligned} \quad (4.18)$$

由于自然界中正氢和仲氢各占 3/4 和 1/4，所以，氢气的热容量中来自转动自由度的贡献为：

$$C_V^{(r)} = \frac{3}{4}C_{V_o}^{(r)} + \frac{1}{4}C_{V_p}^{(r)}, \quad (4.19)$$

其中 $C_{V_o}^{(r)}$ 和 $C_{V_p}^{(r)}$ 分别是正氢和仲氢的转动自由度对热容量的贡献。对于氢来说，实验上测定的转动特征温度大约是 85K，所以在通常室温下仍然有 $\theta_r \ll T$ ，也就是说，可以利用经典近似。于是我们发现，正氢和仲氢的子系配分函数的求和变成：

$$\sum_{j=1,3,\dots} (2j+1)e^{-j(j+1)\frac{\theta_r}{T}} \sim \sum_{j=0,2,\dots} (2j+1)e^{-j(j+1)\frac{\theta_r}{T}} \sim \frac{1}{2} \sum_{j=1}^{\infty} (2j+1)e^{-j(j+1)\frac{\theta_r}{T}} \quad (4.20)$$

⁴由于氢核是费米子，所以两个氢核组成的系统的波函数必须是反对称的。这个核系统的波函数可以写成自旋波函数与轨道波函数（也就是两个氢核构成的转子的波函数）的乘积。如果氢核的自旋相互平行，构成总自旋为 1 的“三重态”，那么系统的自旋波函数就是对称的，所以系统的轨道波函数必须是反对称的，才能使两者的乘积是反对称的，这就要求转动量子数 j 必须是奇数。类似地，如果氢核的自旋相互反平行，构成总自旋为 0 的“单态”，那么系统的自旋波函数就是反对称的，所以系统的轨道波函数必须是对称的，才能使两者的乘积是反对称的，这就要求转动量子数 j 必须是偶数。

也就是说，在较高温度（经典近似）下，正氢和仲氢的区别消失，我们仍然得到经典统计的结果：双原子分子转动自由度对于气体摩尔热容量的贡献为 R 。氢的转动特征温度是所有分子中最高的，所以，当温度降到 $100K$ 附近或以下时，经典近似就不能用了。这时，正氢和仲氢的区别也变得显著。要精确计算氢的热容量中转动自由度的贡献，必须用分立的求和而不能积分；同时，正氢和仲氢也要分开处理。这样得到的理论预言与实验很好地吻合。另一种需要类似处理的是氢的同位素组成的分子 D_2 ，它也具有较高的转动特征温度。但是需要注意的是，由于氘核是核自旋为 1 的玻色子，因此正氘与仲氘的丰度以及与之对应的转动量子数也不同于氢的情形。具体地说，自然界中有 $2/3$ 的氘的自旋波函数是对称的，其转动量子数 j 必须取偶数；另外有 $1/3$ 的氘的自旋波函数是反对称的，其转动量子数 j 必须取奇数。其他的讨论则完全与氢的情形类似。

¶ 最后，我们说明一下为什么我们一直没有考虑分子内电子的运动对于热容量的贡献。一般来说，气体分子内部的电子的运动是不会对气体的热容量有可观的贡献的。原因就在于内部电子的量子化的能级间隔很大，一般在 eV 的数量级。我们知道， $1eV \sim 10^4K$ ，所以，在通常温度下这些内部自由度都不可能激发，因此它们对于气体的热容量没有贡献。如果温度真的高到 10^4K ，气体分子往往会分解成原子。因此，对于分子构成的气体而言，其分子内部的电子运动对于热容量的贡献总是可以忽略的。

在分子的量子力学的能级中，其内部电子的不同能级之间的能级差 $\Delta\epsilon^{(e)}$ 是最大的，数量级是 $1 \sim 10eV$ ；其次是分子的振动能级的能级差 $\Delta\epsilon^{(v)}$ ，大约是 $10^{-1}eV$ ；能级差最小的是分子的转动能级，其能级差大约是 $10^{-4} \sim 10^{-3}eV$ 。⁵ 这就造成以上三种运动的特征温度分别为： 10^4K ， $10^2 \sim 10^3K$ 和 $10 \sim 10^2K$ 。当气体的温度远大于某一种运动的特征温度时，对于那种运动模式就可以运用经典统计的能均分定理；反之，如果温度接近其特征温度，那么对于该运动模式我们只能运用量子力学的结果；在另一个极端，如果温度远低于某种运动模式的特征温度，那么该运动模式就很难被激发，系统基本上被冻结在该模式的基态上，因此，这个模式就几乎不会对热容量有所贡献。总之，只要我们仔细区分量子与经典统计的适用范围，统计物理理论都得到了与实验高度吻合的结果。

⁵这一点在量子力学中是可以得到解释的。造成这种能级间隔的差异来源于电子质量与原子核质量之间所存在的巨大差异。因此，在研究分子的量子力学能级时，可以先认为原子核是固定不动的，电子在固定的核的库仑场中运动。由此可以解出电子运动的能级，称为电子项（electronic term），也就是我们所说的内部电子能级 $\epsilon^{(e)}$ 。当电子处在某个电子态时，它对原子核有个反馈作用，这种作用等效于为原子核提供了一个相互作用势能。对于双原子分子而言，两个原子核可以在这个等效势能的最小值附近振动，这就是分子的振动能级 $\epsilon^{(v)}$ 。最后，原子核体系整体的转动给出分子的转动能级 $\epsilon^{(r)}$ 。这就是所谓的 Born-Oppenheimer 近似。简单的数量级估计可以说明，这三种能级的间隔大致满足： $\Delta\epsilon^{(v)} = \sqrt{m/M}\Delta\epsilon^{(e)}$ ， $\Delta\epsilon^{(r)} = (m/M)\Delta\epsilon^{(e)}$ ，其中 m ， M 分别是电子和原子核的质量。

19 混合理想气体及其化学反应

¶ 前面讨论了单元理想气体的性质，下面我们来考虑由多种（具体地说，我们假定有 k 种）分子组成的混合理想气体的性质。我们用 N_i 其中 $i = 1, 2, \dots, k$ 来表示各个化学组分的分子数目；总的分子数目 $N = \sum_{i=1}^k N_i$ 。

我们将利用巨正则系综来研究这个问题。由于是理想气体，所以系统的总能量仍然是各个分子的能量之和。于是，系统的巨配分函数可以表示为：

$$\log \Xi = \sum_{i=1}^k e^{-\alpha_i} z_i, \quad (4.21)$$

其中 z_i 为第 i 个组元的分子的子系配分函数。类似于我们对于双原子分子理想气体的讨论，它包含了分子平动、转动、振动以及其他可能的内部自由度的积分或求和。一般来说，分子平动部分的能级总是可以看成是准连续的，因此对于平动能级的求和可以代为积分并分离出来。这部分包含了配分函数 z_i 对于体积的依赖：

$$z_i = z_i^{(t)} \cdot z_i'(T), \quad z_i^{(t)} = V \left(\frac{2\pi m_i}{\beta h^2} \right)^{3/2}, \quad (4.22)$$

其中 m_i 是第 i 个组元分子的质量。式中的 $z_i'(T)$ 包含了对于分子除平动以外其他运动量子态的求和或积分。一般来说，它是相当复杂的，但它仅仅是温度的函数，与体积无关。按照巨正则系综的热力学公式，第 i 个组元的分子数与其化学势的关系是：

$$N_i = - \frac{\partial \log \Xi}{\partial \alpha_i} = e^{-\alpha_i} z_i, \quad (4.23)$$

由于配分函数 z_i 对于体积的依赖仅仅来自平动部分，我们得到气体的压强：

$$p = \frac{1}{\beta} \frac{\partial \log \Xi}{\partial V} = \sum_{i=1}^k p_i, \quad p_i = \frac{N_i k_B T}{V}. \quad (4.24)$$

这个结果反映了所谓的 **道尔顿分压定律**。这是英国化学兼物理学家道尔顿（Dalton）首先发现的混合气体的规律：**混合理想气体的压强 p 等于各个化学组分所产生的分压 p_i 之和**。混合理想气体的内能也可以利用系综理论的普遍热力学公式 (2.43) 求出：

$$U = \sum_{i=1}^k \bar{N}_i \left(\frac{3}{2} k_B T - \frac{d \log z_i'(\beta)}{d\beta} \right) = \sum_{i=1}^k \bar{N}_i \bar{\varepsilon}_i \quad (4.25)$$

其中的 $\bar{\varepsilon}_i$ 代表了一个第 i 组元的分子的平均能量。

公式 (4.23) 可以利用来求出各个组元与 **一个分子** 相对应的化学势：

$$\mu_i = k_B T \log \left[\frac{p_i}{k_B T z'_i(T)} \left(\frac{h^2}{2\pi m_i k_B T} \right)^{3/2} \right]. \quad (4.26)$$

在热力学中，一般感兴趣一摩尔物质的化学势，它可以写成：

$$\begin{aligned} \mu_i &= RT [\phi_i(T) + \log(p_i)] , \\ \phi_i(T) &= \log \left[\frac{1}{k_B T z'_i(T)} \left(\frac{h^2}{2\pi m_i k_B T} \right)^{3/2} \right]. \end{aligned} \quad (4.27)$$

特别值得关注的是混合理想气体的熵。按照普遍热力学公式 (2.45) 直接计算我们得到：

$$S = k_B \sum_{i=1}^k \bar{N}_i \left(\frac{5}{2} + \alpha_i - \beta \frac{d \log z'_i(\beta)}{d\beta} \right),$$

将化学势的表达式 (4.26) 代入后得到：

$$S = k_B \sum_{i=1}^k \bar{N}_i (1 + \beta \varepsilon_i - \phi_i(T) - \log p) - k_B \sum_{i=1}^k \bar{N}_i \log x_i. \quad (4.28)$$

这个公式的第一项实际上就是各个组分的熵的简单相加，仿佛其他组分不存在一样。最后一项只与各个组员的粒子数和相对浓度有关，而与温度、压强等等无关。它被称为混合熵：

$$S_{mix} = -k_B \sum_{i=1}^k \bar{N}_i \log x_i. \quad (4.29)$$

由于各个 x_i 均小于 1，因此混合熵永远是正的，这表明气体混合是一个不可逆的熵增加过程。

为了说明这一点，考虑一个中间被隔板分隔成左右两部分的容器。我们将两种不同化学成分的气体（例如氧气和氢气）分别放入左边和右边。并且我们假设两种气体的温度和压强相等。这时，由于左右两边的气体都是化学纯的理想气体，因此系统的总的熵显然就是两种气体的熵之和：

$$S = k_B \sum_{i=1}^2 N_i \left(\frac{5}{2} - \phi_i(T) - \log p - \beta \frac{d \log z'_i(\beta)}{d\beta} \right). \quad (4.30)$$

现在，如果我们打开隔板允许两种气体进行混合。整个系统变成一个混合理想气体，其温度、压强保持不变，这时我们必须用公式 (4.28) 来计算它的熵。显然，系统的熵比混合之前多出来的部分就是前面给出的混合熵 4.29。

¶ 这里我们特别讨论一下与此相关的所谓吉布斯佯谬的问题。公式 (4.28) 仅仅适用于不同的气体之间的混合。如果错误地将它用于同种气体的“混合”，就会得到同种气体在等温、等压下混合也存在熵变的荒谬结论。这就是所谓的吉布斯佯谬。按照量子力学的全同性原理，同种气体的分子之间是全同的，但它与不同气体的分子之间是不同的，这有着“天壤之别”，两种关系之间绝不存在一个连续的过渡。⁶ 事实上，只要熵本身是广延量，就能够避免同种分子混合时的吉布斯佯谬。

¶ 上面求出的混合理想气体中各组元的化学势还可以用来研究混合理想气体的化学反应问题。按照我们在 5 节中的讨论，多元系化学平衡的条件为（见公式 (1.57) 及其后的讨论）： $\sum_i \nu_i \mu_i = 0$ 。于是，我们得到混合理想气体达到化学平衡时的条件：

$$\prod_i p_i^{\nu_i} = K_p(T) . \quad (4.31)$$

这个条件被称为质量作用定律。其中我们定义了定压平衡恒量 $K_p(T)$ ：

$$\log K_p(T) = - \sum_i \nu_i \phi_i(T) . \quad (4.32)$$

注意，定压平衡恒量只是温度的函数，与压强无关。如果把 $p_i = x_i p$ 带入质量作用定律，其中 $x_i = N_i/N = p_i/p$ 为第 i 种组元的摩尔百分比，我们有：

$$\prod_i x_i^{\nu_i} = K(p, T) , \quad K(p, T) = p^{-\sum_i \nu_i} K_p(T) , \quad (4.33)$$

其中 $K(p, T)$ 也称为平衡恒量，但它一般与压强有关。这是质量作用定律的另一种表达形式。

¶ 将定压平衡恒量对温度求导数，得到：

$$\frac{d}{dT} \log K_p(T) = \frac{\Delta H}{RT^2} , \quad (4.34)$$

其中 $\Delta H = \sum_i \nu_i h_i$ 为反应热。这个方程称为范霍夫（Van't Hoff, 1877）方程。它把平衡恒量随温度的变化与反应热联系起来。如果是一个吸热反应，即 $\Delta H > 0$ ，则随着

⁶一种吉布斯佯谬的变种就是考虑两种气体分子，假想一种是染成白色，另一种染成红色。这显然是不同的气体分子，因此等温、等压混合时有熵增。现在设想让红色气体分子的颜色“连续地”变浅，它就会连续地变为白色气体分子。这样就会产生吉布斯佯谬。但是，这种假想实验实际上是量子力学全同性原理所不允许的。

温度的增加, 平衡恒量也增加。所以如果原来化学反应达到平衡, 随温度增加, 平衡被破坏, 导致 $\prod p_i^{\nu_i} < K_p(T)$, 也就是 $\sum_i \nu_i \mu_i < 0$ 。所以会导致化学反应向正向进行, 即会吸收热量以对抗温度的增加。反之, 如果是放热反应, 则温度的增加会使反应向逆向进行, 即会吸收热量以抵消温度的增加。这正是所谓的 **勒沙特列** (Le Chatelier, 1884)

原理 的一个例子。勒沙特列原理是说: 把平衡的某一个因素加以改变以后, 将使平衡态向抵消原来因素改变的效果的方向转移。

20 实际气体的状态方程

理想气体是实际气体在十分稀薄情形下的一个极限情形, 真正的气体往往只是近似地符合理想气体的规律并与之有着系统的偏离。利用统计系综理论, 我们可以系统地计算实际气体 (假定它偏离理想状态不远) 对于理想气体的偏离, 这一节中我们就来讨论这个问题。

20.1 迈尔's 集团展开理论

处理经典的非理想气体的系统理论首先是迈尔 (J.E. Mayer, M.G. Mayer) 得到的。这就是所谓的 **集团展开理论**。我们首先以 **单原子分子** 组成的经典气体为例来阐述这个理论。⁷ 单原子分子气体的哈密顿量为:

$$H = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} + \sum_{i<j} \phi(r_{ij}), \quad (4.35)$$

其中 \mathbf{p}_i 是第 i 个分子的平动动量, m 是气体分子的质量; $\phi(r_{ij})$ 是第 i 个分子与第 j 个分子之间的相互作用的势能, 我们假定它只是两个分子之间距离 $r_{ij} \equiv |\mathbf{r}_i - \mathbf{r}_j|$ 的函数。

要计算系统的物态方程或其他任何热力学性质, 我们必须计算出系统的巨配分函数 Ξ :

$$\Xi = \sum_{N=0}^{\infty} \left(\frac{z}{\lambda_T^3} \right)^N Q_N(T, V), \quad (4.36)$$

其中 $z = e^{\beta\mu}$ 与化学势有关; $\lambda_T = h/\sqrt{2\pi mk_B T}$ 为分子的热波长; $Q_N(T, V)$ 是具有 N 个粒子的系统的正则配分函数中的 **位形积分**。:

$$Q_N(T, V) = \frac{1}{N!} \int \cdots \int \exp \left(-\beta \sum_{i<j} \phi(r_{ij}) \right) d^3\mathbf{r}_1 \cdots d^3\mathbf{r}_N, \quad (4.37)$$

⁷对于存在偶极相互作用的双原子分子气体, 有兴趣的同学可以参考本章后面的习题。

所以，实际气体的巨配分函数的计算完全依赖于系统位形积分的计算。

在集团展开方法中，我们首先定义一个双粒子坐标的函数 f_{ij} ：

$$f_{ij} = e^{-\beta\phi(r_{ij})} - 1 . \quad (4.38)$$

这一节中，我们将假定分子之间的相互作用是短程的，也就是说，势函数 $\phi(r_{ij})$ （从而函数 f_{ij} 也是如此）当分子之间距离 r_{ij} 超过其作用力程时迅速地趋于零。这个假设是有其物理根据的，多数中性分子组成的气体都可以满足这个条件。⁸要确定分子之间相互作用势的具体形式是个十分复杂的量子力学（原子分子物理）问题。我们这里并不准备深入分析。我们所需要的只是势函数的短程性，这个力程一般只是分子大小的几倍。所以，函数 f_{ij} 只是在空间中的很小的（大约只有几个分子尺度的大小）区域中才不为零。集团展开理论正是将 f_{ij} 作为微扰来处理的一种系统的展开方法。我们可以将位形积分写成：

$$\begin{aligned} Q &= \frac{1}{N!} \int \cdots \int \prod_{i < j} (1 + f_{ij}) d^3\mathbf{r}_1 \cdots d^3\mathbf{r}_N \\ &= \frac{1}{N!} \int \cdots \int \left(1 + \sum_{i < j} f_{ij} + \sum_{i < j, k < l} f_{ij} f_{kl} + \cdots \right) d^3\mathbf{r}_1 \cdots d^3\mathbf{r}_N . \quad (4.39) \end{aligned}$$

要进一步将上式化简，我们有必要引入几何（图论）的方法来代表展开式 (4.39) 中的各项。首先，我们将所有分子标号，从 1 到 N 。我们在每个标号旁用一个圆点来表示一个分子，因此，含有 N 个分子的系统将有 N 个具有标号的圆点来表示。如果展开式 (4.39) 中的某一项含有某个 f_{ij} ，我们就相应地将代表第 i 个分子与第 j 个分子之间连一条直线。这样一来，公式 (4.39) 中的每一项都可以与一个（具有标号的）图对应。这些图一定包含有 N 个具有标号的点，同时还包含一定数目的直线将某些点连接起来。每个图中所包含的直线的数目就是它所对应的项中所含的 f 因子的数目。例如，公式 (4.39) 中第一项所对应的图就是 N 个圆点，没有任何直线连接任何点；公式 (4.39) 中一次项（共有 $N(N-1)/2$ 项）都含有一个 f 因子，所以其中的每一项所对应的图的拓扑构形都是 $(N-2)$ 个孤立的点加上两个被直线连接起来的点，它们之间的区别只是分子的标号不同而已；公式 (4.39) 中二次项所对应的图则含有两条直线。具体来说，又可能具有两种不同的拓扑结构：一种是两条直线分别联结四个不同的点；另一种拓扑结构则是两条直线将三个点串起来。如果计算到更高阶，所对应的图形也更复杂，拓扑结构也更丰富。

现在我们注意到：从图的角度来说，每个有标号的图都由若干个连接的、有标号的子图组成。我们称这些几何上连接在一起的子图为集团。集团又按照它所包含的分子的个数（就是点的数目）分为单粒子集团、二粒子集团、三粒子集团，等等。现在注意到：

⁸因此，本节中讨论的方法不能直接应用于完全电离的等离子体。在等离子体中，粒子之间的相互作用是长程的库仑相互作用。这个情形我们将在第 22 中讨论。

公式 (4.39) 中在进行对于所有分子位置的积分时, 那些处在不同集团中的分子位置的积分实际上可以相互独立地进行。显然, 整个图所对应的积分可以分解为对于构成它的不同集团的积分之乘积。由于我们要对于所有分子的位置进行积分, 因此任意一个集团中分子的不同标号所产生的贡献在积分后是相同的。也就是说, 如果对于构成一个集团的点的标号重新排列次序, 并不改变积分以后的贡献。所以, 对于一个集团的积分值而言, 真正重要的是该集团所对应的图的拓扑构形。为此, 对于一个一般的集团 c , 如果我们用 n_c 表示这个集团所包含的点数, 我们定义 **集团积分**:

$$b_c(T) = \frac{1}{n_c!V} \sum_P \int d^3\mathbf{r}_1 d^3\mathbf{r}_2 \cdots d^3\mathbf{r}_{n_c} \prod_{l \in c} f_l, \quad (4.40)$$

其中连乘遍及该集团中所有不同的连线 l , 而积分号前面的求和代表对于构成该集团的点的标号: $(1, 2, \cdots, n_c)$ 的不同置换 P 求和。注意到展开式 (4.39) 是对于直线的数目 (也就是 f 因子的个数) 进行展开 (而不是所含点的数目), 所以最低级的集团积分就是单粒子集团积分, 它不含有任何 f 因子: $b_{\bullet}(T) = 1$; 下一阶是含有一个 f 因子的双粒子集团积分, 它只有一种拓扑结构 $\bullet-\bullet$:

$$b_{\bullet-\bullet}(T) = \frac{1}{2!V} \int d^3\mathbf{r}_1 d^3\mathbf{r}_2 f_{12}. \quad (4.41)$$

再高一阶的集团积分是具有两个 f 因子的三粒子集团积分, 该集团具有类似 $\bullet-\bullet-\bullet$ 的形状:

$$b_{\bullet-\bullet-\bullet}(T) = \frac{1}{3!V} \int d^3\mathbf{r}_1 d^3\mathbf{r}_2 d^3\mathbf{r}_3 (f_{12}f_{13} + f_{12}f_{23} + f_{13}f_{23}), \quad (4.42)$$

高阶的集团的拓扑构形可以更为复杂, 例如具有三个 f 因子的集团就可以有三种不同的拓扑结构。

¶ 现在我们来考虑整个系统的位形积分。如前所述, 系统的位形积分由所有可能的、具有标号的图之和给出; 每一个图的贡献又可以写成构成该图的各个集团积分的乘积。如果一个有标号图中, 某种拓扑结构的集团 c 出现 m_c 次, 而该集团具有 n_c 个点, 那么这类图对于系统位形积分的总贡献为:

$$\frac{1}{N!} \left(\frac{N!}{\prod_c m_c! (n_c!)^{m_c}} \prod_s [V n_s! b_s(T)]^{m_s} \right). \quad (4.43)$$

这里的组和数:

$$\frac{N!}{\prod_c m_c! (n_c!)^{m_c}}$$

正是将 $N = \sum_c m_c n_c$ 个分子组合成相应集团的不同组合方式数。

现在我们将这个结果带回到公式 (4.36) 中就得到系统的巨配分函数:

$$\Xi = \sum_{N=0}^{\infty} \left(\frac{z}{\lambda_T^3} \right)^N \sum_{\{m_c\}} \prod_c \frac{[V b_c(T)]^{m_c}}{m_c!},$$

其中求和 $\sum_{\{m_c\}}$ 代表对于所有可能的 m_c 求和，但这些 m_c 必须满足约束条件： $N = \sum_c m_c n_c$ 。由于有了前面对于 N 的求和，因此上式可以化为对于所有的、不加任何约束条件的 m_c 求和。这个求和正好是指数函数的展开式，于是我们得到：

$$\Xi = \prod_c \exp \left[\left(\frac{z}{\lambda_T^3} \right)^{n_c} V b_c(T) \right] . \quad (4.44)$$

利用热力学公式： $pV/(k_B T) = \log \Xi$ ，我们就得到系统的压强为：

$$\frac{p}{k_B T} = \sum_c \left[\left(\frac{z}{\lambda_T^3} \right)^{n_c} b_c(T) \right] . \quad (4.45)$$

这仍然不是系统状态方程的标准形式。我们仍然需要关于粒子数的展开式：

$$n \equiv \frac{N}{V} = \sum_c n_c \left[\left(\frac{z}{\lambda_T^3} \right)^{n_c} b_c(T) \right] . \quad (4.46)$$

这个式子可以利用来反解出 z 作为 n 的展开式。将此带回公式 (4.45)，我们就可以得到非理想气体的状态方程：

$$\frac{p}{k_B T} = n + B_2(T)n^2 + B_3(T)n^3 + \dots , \quad (4.47)$$

其中所谓的 **第二维里系数** 和 **第三维里系数** 的表达式为：

$$\begin{aligned} B_2(T) &= -\frac{1}{2} \int d^3 \mathbf{r} f(r) , \\ B_3(T) &= -\frac{1}{3V} \int d^3 \mathbf{r}_1 d^3 \mathbf{r}_2 d^3 \mathbf{r}_3 f_{12} f_{23} f_{13} . \end{aligned} \quad (4.48)$$

更高阶的维里系数也可以由类似的（更为复杂）表达式给出。因此，如果我们可以计算出各种集团积分，从而得到各级的维里系数，我们就可以完全确定系统的热力学性质。利用迈尔的集团展开理论，我们发现非理想气体的状态方程具有 **昂尼斯方程** 的形式。因此，统计物理的集团展开给了昂尼斯方程一个微观解释。

¶ 需要指出的是，维里展开仅仅对于那些偏离理想气体不太远的气体系统才是一个好的（收敛比较快的）展开。如果气体的密度很大，接近相变附近液体的密度，那么维里展开就不能很好的加以应用了。这时，我们会发现维里展开中的高阶项的贡献不可忽略。如果我们有能力计算出足够多的维里系数，那么利用一些 **重求和方法**（resummation methods）有可能适当扩大维里展开的适用范围，⁹ 但对于真正稠密的气体和液体，维里展开不能很好描述系统的性质。

⁹例如，可以利用所谓的 Padé 近似。有兴趣的读者可以参见 [10] 第十一章第三节中的讨论。

20.2 维里系数的计算

¶ 要具体的计算维里系数必须知道分子之间相互作用势的形式。一个非常常用的半经验公式是所谓的列纳德-琼斯 (Lenard-Jones) 势:

$$\phi(r) = \phi_0 \left[\left(\frac{r_0}{r} \right)^{12} - 2 \left(\frac{r_0}{r} \right)^6 \right], \quad (4.49)$$

也就是说, 在小的距离上是一个正比于 $1/r^{12}$ 的排斥势, 而在大的距离上是一个正比于 $1/r^6$ 的吸引势。这个势能有时又被称为 12-6 势, 它由两个参数描述: ϕ_0 标度势能的大小; r_0 标度相互作用的力程。¹⁰ 在图 4.1 中我们画出了列纳德-琼斯势的示意图。列纳德-琼斯势可以相当好地描述分子之间的相互作用, 但是利用它计算维里系数是比较复杂的。为了简化计算, 我们将采用更为简化的势模型:

$$\begin{aligned} \phi(\mathbf{r}) &= +\infty & r < r_0 \\ \phi(\mathbf{r}) &= -\phi_0 \left(\frac{r_0}{r} \right)^6 & r \geq r_0 \end{aligned} \quad (4.50)$$

这个势在大距离上与列纳德-琼斯势一致, 在小距离处我们用硬心势 (或者称钢球势) 来替代列纳德-琼斯势中的 $1/r^{12}$ 的排斥势。这样的相互作用又被称为 **硬心** 列纳德-琼斯势。利用球坐标进行计算, 我们立刻得到如下的第二维里系数表达式:

$$B_2(T) = b - \frac{a}{Nk_B T}, \quad (4.51)$$

其中常数 a 和 b 与势能参数 r_0 和 ϕ_0 之间的关系为:

$$a = \frac{2\pi}{3} N^2 \phi_0 r_0^3, \quad b = \frac{2\pi}{3} N r_0^3 \quad (4.52)$$

而气体的状态方程可以近似写为:

$$p = \frac{Nk_B T}{V - b} - \frac{a}{V^2}, \quad (4.53)$$

这正是范德瓦耳斯 (Van der Waals) 气体的状态方程。所以我们看到范氏气体可以看成是分子之间由一个短程排斥和长程吸引时的气体的状态方程。方程中的常数 b 体现了分子的固有体积 (排斥) 效应; 而常数 a 则体现了分子之间的相互吸引。

¹⁰这个势是一个半经验的公式。由它计算出来的实际气体的状态方程与实验符合的相当好。中性分子在大距离上的、正比于 $1/r^6$ 相互吸引势是有它的量子力学基础的。这种相互作用被称为 **范德瓦耳斯** 相互作用, 它实际上是电偶极相互作用的量子力学体现。参见相应的量子力学教材 [5]。

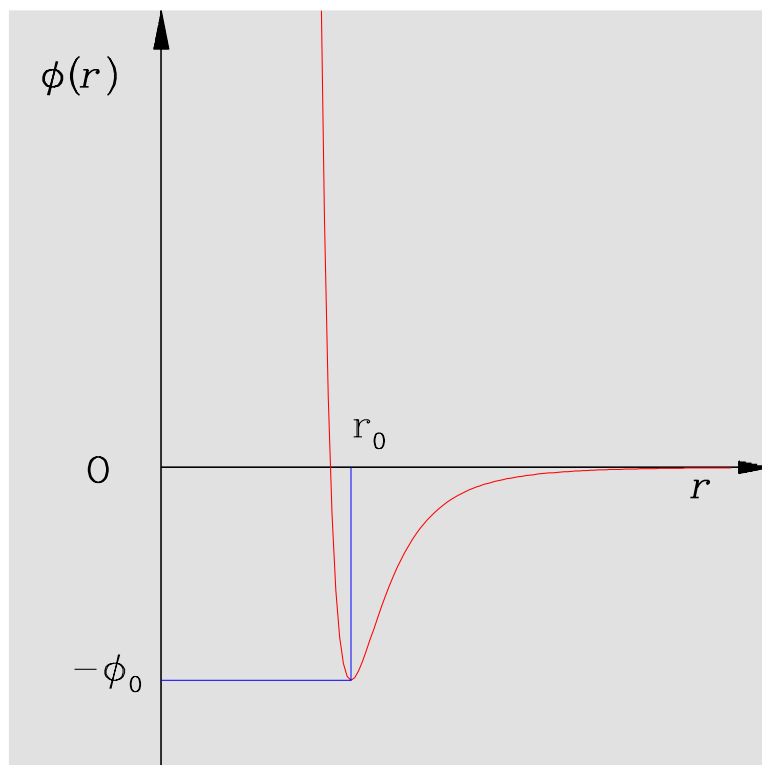


图 4.1: 列纳德 - 琼斯势的示意图。

21 液体的热力学性质

我们在上一节中提到，集团展开方法给出的结果一般不能够应用到密度很大的流体，特别是并不能很好地描述真实液体^①的热力学性质。这一节中，我们就对经典液体的统计理论作一个简单的介绍。经典的液体无疑与我们的日常生活密切相关（水是生命之源，呵呵）。但是，要对于液体给出一个完满的理论解释却是相当困难的。直至今日，仍然有许多没有解决的问题。应当说，液体是比固体和气体都要难于处理和难以捉摸的物理体系。

21.1 对分布函数

¶ 在处理稠密气体和液体的统计理论中，处于核心的物理量是各种分布函数^②。考虑函数：

$$P(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \frac{1}{N!Q_N(T, V)} \exp\left(-\beta \sum_{i < j} U(\mathbf{r}_i - \mathbf{r}_j)\right). \quad (4.54)$$

它代表了 N 个分子分别处在 $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N$ 的几率密度。现在，我们定义如下
 的单粒子分布函数 $n_1(\mathbf{x})$ 和 双粒子分布函数 $n_2(\mathbf{x}_1, \mathbf{x}_2)$ ：

$$n_1(\mathbf{x}) = \sum_{i=1}^N \langle \delta(\mathbf{x} - \mathbf{r}_i) \rangle, \quad (4.55)$$

$$n_2(\mathbf{x}_1, \mathbf{x}_2) = \sum_{i \neq j} \langle \delta(\mathbf{x}_1 - \mathbf{r}_i) \delta(\mathbf{x}_2 - \mathbf{r}_j) \rangle. \quad (4.56)$$

更为一般的 s 粒子分布函数也可以类似的定义。它们的物理意义是十分明显的。单粒子分布函数 $n_1(\mathbf{x})$ 实际上代表了在空间位置 \mathbf{x} 附近的一个单位小体积元内找到的平均粒子数，或者说是在单位体积内找到一个粒子的几率乘以总粒子数。双粒子分布函数则代表了在空间位置 \mathbf{x}_1 和 \mathbf{x}_2 分别找到一对粒子的几率乘以总的粒子对数目。

在一个均匀的流体中， $n_1(\mathbf{x})$ 实际上与位置无关并且就等于流体的分子数密度： $n_1(\mathbf{x}) = N/V$ 。对于双粒子分布函数，它一般只是 $|\mathbf{x}_1 - \mathbf{x}_2|$ 的函数。¹¹因此，我们可以定义所谓的对分布函数（pair distribution function） $g(|\mathbf{x}_1 - \mathbf{x}_2|)$ ：

$$n_2(|\mathbf{x}_1 - \mathbf{x}_2|) \equiv \left(\frac{N}{V}\right)^2 g(|\mathbf{x}_1 - \mathbf{x}_2|). \quad (4.57)$$

显然，当 $|\mathbf{x}_1 - \mathbf{x}_2| \rightarrow \infty$ ，我们有： $g(|\mathbf{x}_1 - \mathbf{x}_2|) \rightarrow 1$ 。通过对分布函数的傅立叶变换，我们可以定义所谓的结构因子 $S(q)$ ：¹²

$$S(q) - 1 = \frac{N}{V} \int d^3\mathbf{r} [g(\mathbf{r}) - 1] e^{-i\mathbf{q}\cdot\mathbf{r}}. \quad (4.58)$$

结构因子之所以重要，是因为它是可以通过电磁波或粒子束在液体中的散射实验而直接测量的物理量。所以，通过这些实验人们可以得到结构函数，从而确立流体中的双粒子分布函数。

¶ 流体的很多物理量都可以利用双粒子分布函数，（或者说流体的对分布函数）来表达。例如，流体中总势能（从而总内能也可以轻易得到）的平均值可以表达为：

¹¹这里我们假定我们讨论的流体是均匀的、各向同性的。如果流体不是各向同性的，比如液晶就是这种系统，原则上双粒子分布函数还依赖于方向。

¹²一般来讲，结构因子应当依赖于三维矢量 \mathbf{q} 。如果我们假定流体是均匀的、各向同性的，那么（正如 $g(\mathbf{r})$ 只是 r 的函数一样）它只是 $q = |\mathbf{q}|$ 的函数。

$$\begin{aligned}
 \langle U \rangle &= \frac{1}{2} \int d^3 \mathbf{x}_1 d^3 \mathbf{x}_2 U(\mathbf{x}_1 - \mathbf{x}_2) n_2(\mathbf{x}_1, \mathbf{x}_2) \\
 &= \frac{N^2}{2V} \int d^3 \mathbf{r} U(r) g(r) .
 \end{aligned} \tag{4.59}$$

¶ 另一个与对分布函数有密切关联的物理量是流体的（等温）压缩率： $\kappa_T = -(1/V)(\partial V/\partial p)_T$ 。为了找到这种联系，我们首先注意到：对于一个固定粒子数的系统（正则系综），我们有：

$$\int d^3 \mathbf{x}_1 n_2(|\mathbf{x}_1 - \mathbf{x}_2|) = N(N-1)/V .$$

如果我们允许系统的粒子数涨落（例如巨正则系综），那么上面式子的等号右边就应当换为： $\langle N(N-1) \rangle / V$ 。同时，对分布函数的定义也应当修改为：

$$g(|\mathbf{x}_1 - \mathbf{x}_2|) = \frac{V^2}{\langle N \rangle^2} n_2(|\mathbf{x}_1 - \mathbf{x}_2|) .$$

于是，我们发现：

$$\int d^3 \mathbf{r} [g(r) - 1] = -\frac{V}{\langle N \rangle} + V \frac{(\Delta N)^2}{\langle N \rangle^2} , \tag{4.60}$$

其中我们引入了粒子数涨落 $(\Delta N)^2$ ，它的定义是：

$$(\Delta N)^2 = \langle (N - \langle N \rangle)^2 \rangle .$$

按照第 10 节中的公式 (2.50)，粒子数的涨落与流体的等温压缩系数直接联系，所以我们得到：

$$n \int d^3 \mathbf{r} [g(r) - 1] = -1 + nk_B T \kappa_T . \tag{4.61}$$

这个重要的关系被称为 **涨落物态方程**，或压缩物态方程。在许多文献中，称：

• • • • •

$$h(r) \equiv g(r) - 1 , \tag{4.62}$$

为 **对关联函数**（pair correlation function），它实际上衡量了流体中不同位置的两个粒子之间的 **关联**。对于理想气体，由于 $g(r) = 1$ ，所以 $h(r) = 0$ ，即理想气体分子之间是无关的，这一点也可以从公式 (4.61) 中得到验证（这时，等号两边都等于零）。在通常情况下，液体的 $nk_B T \kappa_T$ 是比较小的。所以我们说在通常情况下液体 **几乎** 是不可压缩的。这时：

$$n \int d^3 \mathbf{r} [g(r) - 1] \approx -1 .$$

例外的情况发生在临界点附近。这时，流体的压缩率趋向于无穷大。涨落物态方程 (4.61) 告诉我们，这个发散起因是对关联函数 $h(r)$ 在大的 r 处衰减得不够快，从而方程 (4.61) 的左边的积分发散，这时我们称流体中具有长程关联。这种在临界点附近所特有的涨落和关联，我们在下一章讨论二阶相变的涨落和关联时还会更为详细地讨论（参见第 27 节中的讨论）。

¶ 利用维里定理，我们还可以建立另一个与流体中对分布函数密切相关的重要方程：

$$pV = Nk_B T \left[1 - \frac{n}{6k_B T} \int d^3\mathbf{r} [\mathbf{r} \cdot \nabla U(\mathbf{r})] g(\mathbf{r}) \right]. \quad (4.63)$$

这个方程被称为维里物态方程。需要指出的是，公式 (4.63) 和公式 (4.61) 都是严格成立的关系，所需的条件只是流体分子遵从经典力学，同时它们之间的相互作用只有二体相互作用（忽略多体作用）。但是，这两个物态方程都与流体中的对分布函数有关，由于利用各种理论计算对分布函数是很困难的事情，在这类计算中往往必须做出各种近似。这些近似往往会造成公式 (4.63) 和公式 (4.61) 所给出的物态方程不一致。因此，在某种近似下这两者之间的差别恰恰可以作为衡量一种近似是否可信的一个标准。

21.2 流体中对分布函数的计算

¶ 前面提到，从统计物理的基本原理出发计算流体中的对分布函数 $g(\mathbf{r})$ 是一个十分困难的问题，在理论上往往必须做出各种近似。最为直接而且著名的是所谓的 Ornstein-Zernike 方程的 Percus-Yevick 近似。这种近似给出一个对关联函数 $h(\mathbf{r})$ 所满足的积分方程（称为 Percus-Yevick 方程），这个方程可以利用解析或数值的方法进行求解。这里我们不准备深入地讨论这些问题，有兴趣的读者可以参考相关的文献 [10]。我们只是想指出：所有这些传统的理论近似方法往往可以对某些系统（比如对于刚球模型）给出不错的结果，但没有任何一个可以统一地给出不同模型的结果。

¶ 另一方面，随着近年来计算机的发展，另外一些纯数值的研究方法开始显示出它们的活力。其中最为典型的的就是蒙特卡罗（Monte Carlo）方法和分子动力学（Molecular Dynamics）方法。下面我们简单介绍一下分子动力学方法。

分子动力学方法的出发点是回到原始的玻耳兹曼和麦克斯韦的想法，将系统的热力学量看成是组成系统的分子相应微观量的时间平均。也就是说，“放弃”了系综平均的理念而回到按照粒子的运动轨道来进行时间平均。我们在介绍系综理论时曾经提到，要解析地完成这个步骤是不可能的，因为多粒子系统的运动方程太复杂了。但是利用计算机，这一点并不难做到，只要给定初始条件和各个分子之间的相互作用势能，计算机就可以将多粒子的牛顿运动方程进行数值积分，得到各个粒子随时间演化的轨道。¹³将相应的物理量对于这些轨道进行时间平均，我们就可以得到系统的热力学性质。

¹³这里我们所说的“不难做到”是指没有原则上的困难。当然，当你将这些想法输入计算机并进

目前，一般分子动力学计算中可以取相当多的粒子（比如说 10^5 左右），当然这与真实系统中粒子数的数量级仍然相差很多。但是这一点其实往往并不会对于最后的计算结果造成太多的影响。这里面的原因就在于，流体中的任何一个分子的运动一般来说只会受到它附近一些分子的影响；粗略地说，只有与它距离小或等于几个关联长度的那些分子会对它有影响，其他流体分子的运动几乎与它是独立的。只要流体不是非常地接近它的临界点，流体中的关联长度就是比较小的。因此我们进行分子动力学研究时，可以只去研究处于几个关联长度尺度内的分子。所以，即便是数值研究中的分子数仅仅是 10^3 ，往往也能得到不错的结果。当然，如果我们关心流体非常接近临界点附近的行为，这时我们分子动力学的样本也不可避免地要取得更大，否则结果就会明显地偏离真实的结果。

¶ 总之，稠密气体和液体的热力学性质是一个理论上十分重要同时又是相当困难的问题。这里，我们特别强调了近年来数值计算和模拟的重要性。事实上，目前这些数值方法得到的结果几乎已经成为标准的“实验结果”。任何相关的理论计算最终都要与这些“计算机实验”的结果进行比较来判别它们是否正确。当然，要真正体会这些数值方法的威力和魅力，与其在这里干巴巴地说教，不如真正地利用这些方法去解决一个（哪怕是非常简单的）问题，有兴趣的读者可以参考有关的书籍和文献并试图解决本章后面的练习题。

22 稀薄等离子体的统计性质

¶ 这一节中我们讨论稀薄的经典等离子体的统计性质。等离子体是由电离了的、带有正负电荷的粒子组成的流体，整个系统仍然呈现电中性。由于正负粒子之间的库仑相互作用是长程的，因此上一节中讨论的计算方法不能直接应用。我们这里首先研究稀薄的等离子体，它与经典的理想气体偏离不大。从物理上说，这意味着组成等离子系统的各个粒子之间的相互作用势能 $(ze)^2 n^{1/3}$ （其中 z 是粒子的电荷数）要比每个粒子的平均动能 $k_B T$ 小很多，即等离子体的粒子数密度 n 必须满足：

$$n \ll \left(\frac{k_B T}{z^2 e^2} \right)^3 . \quad (4.64)$$

我们称这样的等离子体为稀薄等离子体

¶ 我们假定第 i 种离子的平均数密度和电荷数分别为 n_{i0} 和 z_i （为正负整数），那么系统由于静电相互作用引起的系统的内能的变化 U_{int} 可以写成：

$$U_{int} = \frac{V}{2} \sum_i e z_i n_{i0} \phi_i , \quad (4.65)$$

其中 ϕ_i 是某一个第 i 种离子所感受到的、由于其他离子所产生的平均静电势。在平衡的均匀等离子体中，这个静电势是一个与位置和时间无关的常量。

行计算时，仍然会遇到相当多具体的困难：程序的编写、调试、检验、优化等等。

为了求出 ϕ_i ，我们考虑一个特定的离子周围的离子分布。如果我们用 n_i 表示距离我们选定的离子为 r 处的、第 i 种离子的数密度，那么：

$$n_i(r) = n_{i0} \exp\left(-\frac{z_i e \phi(r)}{k_B T}\right), \quad (4.66)$$

其中 $\phi(r)$ 是距离我们选定的离子为 r 处的静电势。另一方面，静电势 $\phi(r)$ 与产生它的电荷密度的关系是：

$$\nabla^2 \phi = -4\pi e \sum_i z_i n_i(r). \quad (4.67)$$

由于我们假定等离子体是稀薄的，于是公式 (4.66) 可以展开为：

$$n_i(r) = n_{i0} - \frac{n_{i0} e z_i}{k_B T} \phi(r), \quad (4.68)$$

将此式带入方程 (4.67) 并且利用电中性条件： $\sum_i z_i n_{i0} = 0$ ，我们得到：

$$\nabla^2 \phi - \kappa^2 \phi = 0, \quad \xi_D^{-2} \equiv \kappa^2 = \frac{4\pi e^2}{k_B T} \sum_i n_{i0} z_i^2. \quad (4.69)$$

这就是在一个指定离子周围的静电势所满足的方程。参数 $\xi_D \equiv (1/\kappa)$ 具有长度的量纲，被称为 **德拜屏蔽长度**。如果我们选定的离子的电荷数为 z_i ，这个方程的解具有如下形式：

$$\phi(r) = e z_i \frac{e^{-\kappa r}}{r}. \quad (4.70)$$

这个解形式保证了在 $r \rightarrow 0$ 时，静电势具有标准的库仑形式： $e z_i / r$ 。当距离增加时，由于其他离子的极化所产生的屏蔽作用，使得静电势变成一个屏蔽了的库仑势，而屏蔽的特征距离就是上面引入的德拜屏蔽长度 ξ_D 。如果我们将上式中的指数展开，我们有：

$$\phi(r) = \frac{e z_i}{r} - e z_i \kappa + \dots, \quad (4.71)$$

其中上面忽略掉的项在 $r \rightarrow 0$ 时没有贡献。第一项是我们指定的离子本身产生的库仑静电势；第二项是其他离子产生的、在我们所指定的离子位置处的静电势。这正是我们公式 (4.65) 中的 ϕ_i 。所以我们立刻得到等离子体系统的静电相互作用对系统内能的贡献：

$$U_{int} = -e^3 \sqrt{\frac{\pi}{V k_B T}} \left(\sum_i N_i z_i^2 \right)^{3/2}, \quad (4.72)$$

其中 $N_i = n_{i0} V$ 为第 i 种离子的总粒子数。

¶ 我们现在可以积分热力学公式 $U = -(\partial \log Z / \partial \beta)$ 并利用关系 $F = -k_B T \log Z$ 得到系统的自由能：

$$F = F_0 - \frac{2e^3}{3} \sqrt{\frac{\pi}{Vk_B T}} \left(\sum_i N_i z_i^2 \right)^{3/2}, \quad (4.73)$$

其中 F_0 为理想气体的自由能。将上式对于体积微分就得到系统的状态方程：

$$p = \frac{Nk_B T}{V} - \frac{e^3}{3V^{3/2}} \sqrt{\frac{\pi}{k_B T}} \left(\sum_i N_i z_i^2 \right)^{3/2}. \quad (4.74)$$

我们发现它不是维里展开的形式。原因就是粒子之间的相互作用是长程的库仑相互作用。



相关的阅读

这一章我们讨论了系综理论用于经典流体的例子。首先，我们从系综的角度讨论了理想气体。我们的侧重点放在双原子分子理想气体的热容量，特别是它的量子修正；混合理想气体及其化学反应。理想气体的其他内容则没有涉及。系综理论的一个经典应用是非理想流体的物态方程，它的完整理论是 Mayer 的集团展开理论，可以按照课时的情况讲述完整的理论或它的一个简化的版本。我们讨论了理想溶液的统计理论，得到了理想溶液的化学势。由此出发可以讨论许多热力学问题，可以按照课时多少具体安排。随后，我们讨论了经典的等离子体的统计性质，讲述了 Debye-Hückel 理论。这个理论可以直接应用到强电解质的溶液理论中去。

第五章

二级相变的平均场理论

本章提要

- 自旋模型的微观机制与处理方法 (23)
- Ising 模型的平均场近似 (24)
- 相变的朗道理论 (25)
- Bragg-Williams 近似 (26)
- 临界点附近的涨落与关联 (27)
- 易兴模型的高温展开 (28)
- 具有连续对称性的系统的相变 (29)
- 对称性破缺与普适性 (30)

二级相变又被称为连续相变，它在近代凝聚态物理和统计物理的发展中占有极为重要的地位，这主要是由于许多重要的物理发现都与二级相变有关联，例如：超流相变、超导相变、铁磁相变、反铁磁相变等等。这些相变现象不仅仅在纯物理学的研究中占据着重要的地位，对这些物理现象更深入的理解也为人们以后技术和应用的发展奠定了基础。

自旋模型 是人们在研究相变问题时所引入的一类特殊的统计物理模型。这类模型的哈密顿量一般只包含晶格中不同格点上的 **自旋变量**（一般地讲，是指晶格原子的总角动量）之间的相互作用。人们发现，铁磁相变、反铁磁相变等许多与磁性有关的二级相变都可以利用某种自旋模型来相当好地加以刻划。其实，研究自旋模型的意义还远不止如此。人们逐渐发现，其他二级相变（例如超导相变、超流相变等等）的临界性质也可以与

自旋模型的临界性质建立起某种对应关系，这就是所谓的 **普适性**（参见 30 节的讨论）。也就是说，自旋模型在其二级相变临界点附近的物理行为实际上具有超出这个模型之外的 **共性**。

在这一章中，我们准备以自旋模型为例来讨论二级相变的问题。我们首先简单说明一下自旋模型的由来以及处理自旋模型的几种常用的理论方法；然后我们以最为简单的自旋模型—易兴模型（Ising model）—为例来说明平均场近似的方法；由此我们进一步指出易兴模型在临界点附近的行为；接着我们从另一个角度考察平均场近似，说明它与朗道二级相变理论的关系；随后，我们利用对偶性来严格计算二维正方晶格上易兴模型的临界温度并由此引出超出平均场近似的其他处理方法；最后，我们将讨论一下其他自旋模型，特别是具有连续对称性的自旋模型。由于二级相变内容的丰富，显然不可能在一章中覆盖所有内容。我们的讨论将主要围绕平均场近似来展开，对于其他方法只是简单地介绍。有兴趣的同学可以参考有关的书籍 [11, 10, 16]。

23 自旋模型的微观机制与处理方法

自旋模型的动力学自由度是定义在某个晶格的格点上的自旋变量（原子的角动量）。自旋模型的哈密顿量只包含各个格点上的这些自旋变量之间的相互作用。这种相互作用的微观机制实际上是由量子力学的 **交换相互作用** 所引起的。所谓交换相互作用实际上是由于量子力学的 **全同性原理** 所引出的纯粹量子效应。下面我们就简要地说明一下交换相互作用的起因，更详细的讨论可以参阅相关的量子力学教材。

为了简单起见，我们考虑两个全同的、具有一个价电子的 **电中性** 原子（例如氢原子），它们之间的距离为 a 。我们会看到：由于量子力学全同性原理的要求，这个系统的能级会与两个价电子的自旋状态发生关联。如果我们 **暂时** 忽略两个原子之间的库仑相互作用，这实际上忽略了以下四种库仑相互作用：第一个电子与第二个原子实之间的相互作用；第二个电子与第一个原子实之间的相互作用；两个电子之间的库仑相互作用；两个原子实之间的相互作用。当上述相互作用被忽略时，系统就是两个完全 **独立的** 原子。因为电子是费米子，所以由两个价电子组成的量子力学系统的总波函数必定是 **反对称** 的。由于系统真实的相互作用哈密顿量并不明显依赖于自旋，因此这两个电子构成的量子力学系统的总波函数可以写成它们的 **空间波函数** 与 **自旋波函数** 的乘积。所以，如果两个价电子的自旋波函数是对称的，那么它们的空间波函数一定得是反对称的；反之，如果两个价电子的自旋波函数是反对称的，那么它们的空间波函数一定得是对称的。在忽略两个原子之间相互作用的时候，上述两种量子状态的能量是完全相同的，我们记为 K 。

现在，让我们开始考虑两个原子之间的相互作用。也就是说，我们记入上面忽略的四种库仑相互作用。这时候我们发现：上述两种量子状态所对应的能量发生了变化。具体来

说，自旋平行的量子态（三重态）的能量变为 $K - J$ 而对应于自旋反平行的状态（单态）的能量变为 $K + J$ 。其中 J 被称为 **交换积分**。它与上述四种库仑相互作用在两个原子波函数之间的重叠积分有关。而且它原则上可以是正的，也可以是负的。具体到氢原子，交换积分 $J < 0$ 。因此，两个氢原子处于单态（即两个电子的自旋反平行的态）时，系统的能量比较低。这就是两个氢原子形成共价键结合成氢分子的状态。但是，也有的原子之间的交换积分是正的。这时两个电子的自旋波函数反对称，即它们的自旋反平行时（即两个价电子构成的系统的总自旋量子数 $S = 0$ ）系统反而具有较高的能量；当两个电子自旋波函数对称，即它们的自旋平行时（即两个价电子构成的系统的总自旋量子数 $S = 1$ ）则具有较低的能量。因此，这两种情况之间的相对能量差可以用一个等效的哈密顿量写成：

$$H_{ex} = -JS_1 \cdot S_2, \quad (5.1)$$

其中常数 J 就是 **交换积分**；而 S_1 和 S_2 分别为两个价电子的自旋。这个表达式就是所谓的 **交换相互作用**。¹ 在量子力学中可以证明：交换积分 J 与两个电子的单电子波函数的重叠有关，所以它的数值一般随两个原子间的距离增加而 **指数地** 趋于零。因此，在考虑晶格上的原子自旋之间的相互作用时，往往可以假定只有 **近邻的** 原子之间才有交换相互作用。在多数情况下，交换相互作用的强度一般也是比较弱的，只有在铁磁体中交换相互作用才变得十分重要。

¶ 现在我们可以引入所谓的 **自旋模型**。一个最简单的、也是最著名的自旋模型就是 **易兴模型**（Ising model）。² 考虑一个晶格，我们用 i 来标记这个晶格上的格点。在每一个格点上有一个自旋变量 σ_i 。在易兴模型中，每个自旋变量只能取两个分立的值： $\sigma_i = \pm 1$ 。在没有外磁场存在的情况下，易兴模型的哈密顿量为：

$$H = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j, \quad (5.2)$$

其中 σ_i 是在第 i 个格点上的自旋变量， J 是交换能量，求和中的符号 $\sum_{\langle ij \rangle}$ 代表对晶格上所有的 **近邻对** 求和。由于前面提到的交换相互作用的短程性，易兴模型的哈密顿量中

¹这里需要注意的一点是：所谓交换相互作用并不是物质世界四种基本相互作用（引力、电磁、强、弱）之外的另外一种相互作用。它只是由于量子力学的全同性原理所造成的、已知的相互作用（这里就是电磁相互作用）的特殊体现形式而已。说它特殊，是因为这种形式是纯粹量子的，没有经典的对应。

²更为确切地讲，应当称之为 Lenz-Ising 模型。这个模型首先是 Lenz 在 1920 年提出的。他的学生 Ising 在 1925 年和他一起研究了一维的情形。当时 Lenz 是想让 Ising 研究一下铁磁相变的性质。Ising 首先严格的解了一维的模型并发现在任何非零的温度没有相变发生。

仅仅保留了最近邻的自旋变量之间的相互作用。显然，如果交换能量 $J > 0$ ，为了使得系统的能量较小，那么两个相邻格点上的自旋倾向于相互平行（即同时取 $+1$ 或同时取 -1 ）；反之，如果交换能量 $J < 0$ ，那么两个相邻格点上的自旋倾向于相互反平行。

易兴模型的引入是为了研究所谓的铁磁—顺磁相变。由于每个晶格格点上的磁矩的系综平均值与该格点上的自旋变量的平均值成正比：³ $\langle \mu_i \rangle \propto \langle \sigma_i \rangle$ ，因此，系统的磁化强度的大小 \mathcal{M} 也一定与每个晶格上自旋变量的系综平均成正比： $\mathcal{M} \propto \langle \sigma_i \rangle$ 。因此，在自旋模型的讨论中，人们往往就称每个晶格上的自旋变量的平均值 $\langle \sigma_i \rangle$ 为系统的磁化强度。如果在外磁场 \mathcal{H} 为零时磁化强度仍然不为零，我们就称系统具有铁磁性，或者说系统中有自发磁化。铁磁性不可能在很高的温度出现。高温时，系统一般总是呈现顺磁性。能够出现铁磁性的系统都存在一个铁磁相变的临界温度，它又被称为居里温度。铁磁性只在温度低于其居里温度时显现。因此，如果我们期望没有外磁场的易兴模型的哈密顿量 (5.2) 能够描写铁磁体，它必须能够在低温时存在一个有自发磁化的相（又称为铁磁相），即在低温下必须有： $\langle \sigma_i \rangle \neq 0$ 。这一点从上面关于能量的粗略分析可以看出一些苗头。例如，当 $J > 0$ 时，相邻的自旋倾向于平行，因此，如果在温度足够低时，系统中所有的自旋都相互平行，显然这时系统中的磁化强度（或者说 $\langle \sigma_i \rangle$ ）不为零。也就是说，有可能形成铁磁相。同样的，当 $J < 0$ 时，系统中有可能形成反铁磁相。注意，这里我们只是说有可能，不是说一定能。我们下面的分析会看到：系统中能否形成铁磁相不仅仅依赖于能量（哈密顿量）的分析，还必须分析统计涨落的影响。例如，我们可以严格的证明（参考本章后面的习题）：在一维晶格上，上述的易兴模型不可能在任何非零的温度形成铁磁相。其中的原因就是虽然从能量的角度分析，系统倾向于形成铁磁相，但是在一维统计涨落足以在任何非零的温度破坏这些铁磁相。总之，易兴模型中能否发生铁磁—顺磁相变，必须通过详细的统计物理计算和分析才能够最终确定。

铁磁性只是自然界存在的各种磁性中的一种。其他类型的磁性还包括：顺磁性、抗磁性、反铁磁性、亚铁磁性等。顺磁性其实是最为普通和常见的磁性。具有顺磁性的系统在外磁场为零时没有非零的自发磁化，或者说当外磁场 $\mathcal{H} = 0$ 时系统的序参量 $m = 0$ ；如果加上一个非常小的外磁场，那么系统所具有的磁化强度也正比于外磁场，其中的比例系数（对顺磁性而言，这个系数还是正的）就是磁化率。

公式 (5.2) 是没有外加磁场时的易兴模型的哈密顿量。如果我们存在一个外加的均匀磁场 \mathcal{H} ，那么易兴模型的哈密顿量中还要加上一项各个自旋（也就是磁矩）与外磁场的塞曼能量（参见第 7.1 小节的公式 (2.11)）：⁴

³按照晶格的平移不变性，每个晶格格点上的自旋变量的系综平均值与位置无关，即 $\langle \sigma_i \rangle$ 实际上与格点的位置 i 无关。

⁴注意，本章中所谓的“外磁场”实际上表示各个格点上的磁矩的大小 μ 与所加的外磁场 H_{ext} 的乘积： $\mathcal{H} = \mu H_{ext}$ 。因此，本章中所谓的“磁场” \mathcal{H} 实际上具有能量的量纲。

$$H = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - \sum_i \sigma_i \mathcal{H} , \quad (5.3)$$

这就是有外加均匀磁场时的易兴模型的哈密顿量。

¶ 如果我们对于格点上的自旋的取值加以改变，例如，不是限制它们取分立的值而是可以取一个单位圆上的连续值，那么我们就得到了所谓的（经典的）**XY - 模型**：

$$H = -J \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j , \quad (5.4)$$

其中 \mathbf{S}_i 是定义在第 i 个格点上的自旋，它有两个分量： $\mathbf{S}_i = (\mathbf{S}_i^x, \mathbf{S}_i^y)$ 并且满足： $(\mathbf{S}_i^x)^2 + (\mathbf{S}_i^y)^2 = 1$ 。如果我们允许格点上的自旋可以有三个分量，我们就得到了著名的（经典的）**海森堡模型**：⁵

$$H = -J \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j , \quad (5.5)$$

其中 \mathbf{S}_i 是在第 i 个格点上的自旋，它有三个分量： $\mathbf{S}_i = (\mathbf{S}_i^x, \mathbf{S}_i^y, \mathbf{S}_i^z)$ 并且满足： $(\mathbf{S}_i^x)^2 + (\mathbf{S}_i^y)^2 + (\mathbf{S}_i^z)^2 = 1$ 。⁶ 也就是说，海森堡模型中各格点上的自旋取值于一个单位球面。类似于易兴模型，如果存在外加磁场，那么 XY - 模型和海森堡模型的哈密顿量中也应当加上相应的塞曼能量项：

$$H = -J \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j - \sum_i \mathbf{S}_i^z \mathcal{H} , \quad (5.6)$$

其中，不失一般性，我们假定了外磁场沿着 z 方向。

¶ 由于上述自旋模型的各个自由度之间（也就是各个自旋变量之间）具有相互作用，它们显然不能看成是近独立子系来处理。要研究它们的统计物理性质必须从普遍的系综理论出发。以正则系综为例，易兴模型的正则配分函数可以写为：

$$Z = \sum_{\{\sigma_i\}} e^{-\beta H[\sigma_i]} , \quad (5.7)$$

其中 $H[\sigma_i]$ 为易兴模型的哈密顿量 (5.2) 或 (5.3)。上式中的求和是指对于一切自旋变量 $\{\sigma_i\}$ 的可能值求和，即：

$$\sum_{\{\sigma_i\}} \equiv \sum_{\sigma_1=\pm 1} \sum_{\sigma_2=\pm 1} \cdots \sum_{\sigma_N=\pm 1} ,$$

⁵一般地讲，如果每个格点上的自旋有 n 个分量，并且它取值于一个 $(n-1)$ 维的球面上，即满足： $\sum_{a=1}^n (\mathbf{S}_i^a)^2 = 1$ ，那么这样的模型称为非线性 $O(n) - \sigma$ 模型。因此 XY - 模型就是非线性 $O(2) - \sigma$ 模型；海森堡模型就是非线性 $O(3) - \sigma$ 模型。

⁶与这些经典统计模型对应的还有所谓的量子易兴模型、量子 XY - 模型、量子海森堡模型等。在量子模型中每个晶格上的自旋变量不再具有经典的取值而是量子力学的算符。

其中对于每一个自旋变量的求和都有两项，分别对应于该自旋变量取 $+1$ 和 -1 。因此，一个易兴模型系统的配分函数的求和总共包含 2^N 项，其中 N 是所考虑的系统总的格点数。

自旋模型的求解一直是近代统计物理发展的主要方向之一。遗憾的是，绝大多数模型很难进行严格求解。⁷作为一个可以严格求解的例子，同学们可以尝试求解一维的易兴模型（本章后面的习题）。对于多数的、无法解析求解自旋模型，人们往往需要依赖于其他近似的方法，这些近似方法主要有：

☞ **平均场近似**：主要思想是将其他自旋对于某一个给定自旋的相互作用用一个自治的、常数 **平均场** 来替代。这样一来，自旋模型中不同格点上的自旋变量变成了相互统计独立的变量，从而可以轻易求解。其优点是物理图象清晰，容易进行计算；缺点是对于一些问题（特别是低维中的问题）往往得到定量上（有时甚至是定性上）错误的结果。可以证明，平均场近似实际上在系统的空间维数趋于无穷时变成严格的理论。

☞ **高温展开** 或其他展开方法。其优点是可以相当精确地得到系统离临界点比较远时候的性质；缺点是在临界点附近展开式的收敛性越来越差。可以证明，在临界点处这个级数发散。另外，要得到准确的结果，往往需要计算级数展开中的许多项，计算量比较大。作为一个例子，我们将在后面（见第 28 节）简要介绍一下易兴模型的高温展开。

☞ **重整化群** 方法。原则上是研究靠近临界点（所谓临界区域）时系统性质的严格方法。但实际上必须与其他近似方法联合使用（例如 ϵ -展开等）。一般说来，其计算也相当的复杂，有时甚至必须进行数值计算（所谓的数值重整化群）。重整化群方法的介绍超出了本课程的范畴，只有留待后续课程（例如量子统计课程）处理了。

☞ **蒙特卡罗**（Monte Carlo）数值模拟方法。这是一个理论上严格的方法，但是需要进行计算机 **数值模拟**。对于计算机计算资源有一定的要求（取决于所研究的问题）。随着计算机能力的普遍提高，这种方法越来越成为这方面研究的主流方法之一。

在本章中，我们将主要讨论最为简单的平均场近似方法，对于易兴模型的高温展开方法也会稍做介绍。

⁷所谓严格求解是指可以 **解析地** 计算出系统的配分函数（从而也可以得到所有热力学量）。一个里程碑性的成果是昂萨格（Onsager）在 1944 年发表的关于二维易兴模型的严格解。这是人们得到的第一个非平庸的、有相互作用的统计模型的正则系综严格解。后来，又有一大类二维的模型找到了严格解。但是这些求解的方法无法推广到高维模型。二维模型具有严格的可解性是得益于二维的 **共形群**（Conformal Group）的无穷多的对称性造成的。

24 Ising 模型的平均场近似

¶ 考虑处在均匀外磁场 \mathcal{H} 中的易兴模型，它的哈密顿量由式 (5.3) 给出：

$$H = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - \sum_i \sigma_i \mathcal{H} . \quad (5.8)$$

对于一个给定的格点 i 上的自旋 σ_i ，哈密顿量中与它有关的项为： $-\sigma_i(\mathcal{H} + J \sum_{j, \langle ij \rangle} \sigma_j)$ 。也就是说，这个自旋除了感受到外磁场 \mathcal{H} 以外，还感受到与它相邻的格点上的自旋变量之和所产生的一个等效的磁场： $J \sum_{j, \langle ij \rangle} \sigma_j$ 。但是，这个所谓的等效磁场并不是一个常数磁场，而是与它近邻格点上的自旋变量有关，因此是一个在统计上涨落变化的磁场，这样的问题仍然无法严格处理。平均场近似的思路就是用这个等效磁场的系综平均值（从而是一个常数）来代替等效磁场本身。所以我们引入一个平均场：

$$\mathcal{H}_{eff} = \langle J \sum_{j, \langle ij \rangle} \sigma_j \rangle = qJ \langle \sigma_j \rangle , \quad (5.9)$$

其中 q 是晶格的配位数，即每个格点的近邻格点数目，它完全由晶格的几何结构所决定。

⁸ 注意，由于晶格的平移对称性，所有的 $\langle \sigma_i \rangle$ 都是相同的，并不依赖于格点位置 i ；同样的，等效平均场 \mathcal{H}_{eff} 也是一个常数磁场，与位置无关。于是，我们引入在平均场近似下的哈密顿量：⁹

$$H^{(MF)} = - \sum_i \sigma_i (\mathcal{H} + \mathcal{H}_{eff}) , \quad (5.10)$$

我们看到：在平均场近似下，易兴模型的求解问题化成了一个顺磁体的、近独立子系统统计问题。这种方法首先是由外斯（Weiss）提出的，它把周围的自旋对某个给定自旋所贡献的等效磁场 \mathcal{H}_{eff} 称为分子场，因此这个等效磁场又被称为外斯分子场。

平均场近似下的系统的配分函数可以很容易的计算出：

$$Z = \sum_{\{\sigma_i\}} e^{\beta(\mathcal{H} + \mathcal{H}_{eff}) \sum_i \sigma_i} = (2 \cosh[\beta(\mathcal{H} + \mathcal{H}_{eff})])^N . \quad (5.11)$$

由此，我们可以计算出每个自旋的系综平均值（也就是系统的磁化强度）：

$$\langle \sigma_i \rangle = \tanh[\beta(\mathcal{H} + \mathcal{H}_{eff})] , \quad (5.12)$$

但是按照自洽性条件 (5.9)，我们应当有：

⁸ 例如，对于一个三维的简单立方晶格，配位数 $q = 6$ ；对于一个二维的正方晶格，配位数 $q = 4$ ；对于一个二维的三角晶格，配位数 $q = 6$ ，等等。

⁹ 我们用 $H^{(MF)}$ 来表示进行了平均场近似以后的系统哈密顿量，以区别于原来的哈密顿量 H 。注意，两者并不相等！

$$\mathcal{H}_{eff} = qJ \tanh[\beta(\mathcal{H} + \mathcal{H}_{eff})] , \quad (5.13)$$

这个式子给出了 \mathcal{H}_{eff} 所必须满足的一个 **自洽方程**。

¶ 在自旋模型中，我们称每个格点上的自旋变量的系综平均值：

$$m = \langle \sigma_i \rangle \quad (5.14)$$

为系统的 **序参量** (order parameter)。我们前面提到过，这个序参量正比于系统的磁化强度。如果在外磁场 \mathcal{H} 为零时序参量仍然不为零，我们就称系统具有 **铁磁性**，也可以说系统中存在 **铁磁序**。系统中是否具有铁磁序与系统的各个参数有关。对于易兴模型，如果给定了其哈密顿量中的各个参数，那么系统是否具有自发磁化（或者说，铁磁序）就由温度决定。当温度降低到某个 **临界温度**（居里温度）时，系统会发生一个从顺磁相到铁磁相的相变。这个相变在没有外磁场的情形下是一个二级相变。当温度低于临界温度时，系统可以有非零的自发磁化。下面我们利用平均场近似的结果来讨论易兴模型的自发磁化问题。

¶ 根据平均场近似的结果，当外磁场 $\mathcal{H} = 0$ 时，分子场必须满足的自洽性条件 (5.13) 为：

$$\mathcal{H}_{eff} = qJ \tanh(\beta \mathcal{H}_{eff}) . \quad (5.15)$$

这个 **非线性方程** 的解可以分为两个情形来讨论。注意到双曲函数在小的 \mathcal{H}_{eff} 处的行为是： $\tanh(\beta \mathcal{H}_{eff}) \sim \beta \mathcal{H}_{eff}$ ，因此，如果 $qJ/k_B T < 1$ ，那么上面的方程将只有一个零解： $\mathcal{H}_{eff} = 0$ ；反之，如果 $qJ/k_B T > 1$ ，那么上面的方程除了必定存在的零解以外，还有两个非零的解，这两个非零解互为相反数。同时可以证明：这时两个非零解所对应的系统的自由能更小。因此，这时系统真正取得的应当是非零解。按照定义 (5.14) 和公式 (5.9)，非零的 \mathcal{H}_{eff} 就意味着非零的序参量 m ，即自发磁化。所以我们得到平均场近似下的易兴模型发生自发磁化的临界温度为：

$$T_c = \frac{qJ}{k_B} , \quad (5.16)$$

或 $k_B T_c / J = q$ 。注意，这个结果与晶格的维数无关，只与配位数 q 和交换能量 J 有关。

¹⁰利用双曲函数的泰勒展开式:

$$\tanh x \sim x - \frac{1}{3}x^3 + \dots, \quad (5.17)$$

我们可以求出在临界温度附近的分子场的大小, 进而求出平均自发磁化的数值:

$$m = \langle \sigma_i \rangle = \sqrt{3} \left(1 - \frac{T}{T_c} \right)^{1/2}, \quad (5.18)$$

这个公式适用的范围是温度非常接近于临界温度。因此, 在临界点附近系统序参量的大小与温度的关系为:

$$m = \langle \sigma_i \rangle \sim (T_c - T)^{1/2}. \quad (5.19)$$

上面这个公式的“变种”实际上在各种二级相变中都会出现。在任何一个二级相变中, 都可以定义一个“序参量” m 。当系统的温度 T 从高于临界温度 T_c 的一方过渡到低于临界温度一方时, 序参量会发生一个从零连续地变为非零的转变。这正是这类相变被称为连续相变的一个原因。在临界点附近, 实验和各种理论都已经证明, 系统的序参量与温度的关系可以普遍地写成:

$$m \sim \begin{cases} 0 & T > T_c \\ (T_c - T)^\beta & T < T_c \end{cases}, \quad (5.20)$$

其中 β 称为序参量临界指数。于是我们发现: 易兴模型的平均场近似给出的序参量临界指数为: $\beta = 1/2$ 。一般来讲, 真实铁磁系统中的序参量临界指数 β 并不严格等于 $1/2$ 。这也说明了平均场近似只是一个比较粗略的近似, 并不能完全定量地描写实际系统的相变性质。

¶ 公式(5.20)定义的只是系统众多的临界指数中的一个。实际上, 还有其他一些临界指数, 它们分别描写其他一些物理量在临界点附近的行为。例如, 如果我们考虑(等温)磁化率:

$$\chi_T = \left(\frac{\partial m}{\partial \mathcal{H}} \right)_{T, \mathcal{H}=0}. \quad (5.21)$$

利用平均场近似下的序参量表达式:

$$m = \tanh[\beta(\mathcal{H} + Jqm)] \sim \beta(\mathcal{H} + qJm), \quad (5.22)$$

¹⁰在平均场近似下, 一个二维三角晶格上的易兴模型与三维简单立方晶格上的易兴模型具有同样的临界温度(假定它们具有相同的 J 的话)。实际上两者的临界温度是很不相同的, 因为维数会对模型的临界行为起十分重要的作用。这一点是平均场近似下无法解释的。

其中我们假定外磁场很小，而且温度足够接近临界温度从而序参量 m 也是小量。我们由此求出：

$$m = \frac{\mathcal{H}}{k_B(T - T_c)} .$$

于是，我们可以求得等温磁化率：

$$\chi_T \sim (T - T_c)^{-1} , \quad \mathcal{H} = 0 , T \rightarrow T_c , \quad (5.23)$$

在临界点附近，一个任意系统的等温磁化率与温度的关系可以普遍地写成：

$$\chi_T \sim |T - T_c|^{-\gamma} , \quad (5.24)$$

其中 γ 被称为 **磁化率 临界指数**。所以我们得知：在平均场近似下的易兴模型的磁化率临界指数 $\gamma = 1$ 。当然，真实系统的铁磁相变中的临界指数 γ 并不严格等于 1。

¶ 如果我们研究在 $T = T_c$ 时序参量对于外磁场的依赖关系，我们可以定义另外一个临界指数：

$$\mathcal{H} \sim m^\delta \quad T = T_c , \quad (5.25)$$

其中 δ 称为 **序参量对磁场的 临界指数** 在平均场近似下的易兴模型，如果 $T = T_c$ ，那么：

$$m = \tanh(m + \beta\mathcal{H}) \sim (m + \beta\mathcal{H}) - \frac{1}{3}(m + \beta\mathcal{H})^3 + \dots ,$$

从而我们得到： $\delta = 3$ 。

¶ 另外一个重要的临界指数是反映（固定外磁场时的）热容量 $C_{\mathcal{H}}$ 在临界点附近的行为的：

$$C_{\mathcal{H}} \sim |T - T_c|^{-\alpha} . \quad (5.26)$$

临界指数 α 被称为 **热容量 临界指数** 在平均场近似下，易兴模型的热容量在临界点附近并不发散。它实际上有一个有限的跳跃（不连续）：

$$C_{\mathcal{H}} = \begin{cases} 3Nk_B/2 & T \rightarrow T_c^- \\ 0 & T \rightarrow T_c^+ \end{cases} . \quad (5.27)$$

因此，平均场近似下易兴模型的热容量临界指数 $\alpha = 0$ 。

25 相变的朗道理论

朗道 (L.D. Landau) 在 1936 年建立了一套关于 (二级) 相变的热力学唯象理论。所谓唯象理论是区别于从系统微观图象出发的统计物理理论, 它主要是指基于一些基本的物理假设, 然后运用热力学的基本原理得到的理论。

朗道理论主要的假设包括: 二级相变的两个相可以用一个序参量 m 来刻画。在一个相中序参量 $m \equiv 0$ 。在另一个相中, $m \neq 0$ 。在临界点序参量连续地从零变为非零。也就是说, 在临界点处一定有 $m = 0$ 。在临界点附近, 系统的自由能可以展开成系统的序参量 m (记住, 在临界点附近系统的序参量是一个小量) 的幂级数; 系统在平衡态所取的序参量的值对应用系统自由能的极小值。需要指出的是, 实验和理论都证明了, 朗道理论的这个基本假设实际上并不成立。尽管如此, 朗道理论仍然为我们理解相变提供了有意义的信息, 而且它可以作为更为正确的理论 (比如重正化群理论) 的一个出发点。下面我们简要介绍一下相变的朗道理论。我们将首先介绍一下朗道关于二级相变的理论, 然后我们再简单讨论一下朗道关于一级相变和三临界点的理论。

25.1 二级相变的朗道理论

¶ 这时我们假设系统具有 $m \leftrightarrow -m$ 的对称性, 那么在临界点附近, 系统的自由能可以展开为:

$$F(m) = F_0 + \frac{a(T)}{2}m^2 + \frac{b(T)}{4}m^4 + \dots \quad (5.28)$$

我们进一步假定, 在临界点附近 (即 $T \sim T_c$ 时), 上面展开式中的系数 $a(T)$ 会发生一个符号的变化, 即假设: $a(T) = a_0(T - T_c)$, 其中 a_0 为一正数。我们同时假定系数 $b(T)$ 在临界点附近永远是正的。

按照朗道理论的基本假设, 系统真实的平衡态的序参量的取值由自由能极小条件来确定:

$$\left(\frac{\partial F}{\partial m}\right)_T = 0 = a(T)m + b(T)m^3 + \dots \quad (5.29)$$

于是我们发现: 当 $T > T_c$ 时, 由于 $a(T) > 0$, 从而使自由能极小的解为: $m = 0$, 也就是说, 这时系统的序参量为零 (没有自发磁化); 而当 $T < T_c$ 时, 由于 $a(T) < 0$, 从而使自由能极小的解为:

$$m = \sqrt{-a(T)/b(T)} \sim (T_c - T)^{1/2} \quad T < T_c \quad (5.30)$$

也就是说, 这时系统的序参量不为零 (有自发磁化)。这个公式告诉我们: 朗道理论预言了系统序参量临界指数 $\beta = 1/2$ 。值得注意的是, 这个数值与我们上一节中得到的易兴模型在平均场近似下的序参量临界指数 β 相同, 都等于 $1/2$ 。

如果我们加上一个小的外磁场 \mathcal{H} ，那么可以研究系统对于外场的响应，这时朗道自由能应当写为：

$$f(m) = f_0 - m\mathcal{H} + \frac{a(T)}{2}m^2 + \frac{b(T)}{4}m^4 + \dots \quad (5.31)$$

而自由能极小的条件 (5.29) 也应当改为：

$$\mathcal{H} = a(T)m + b(T)m^3 + \dots \quad (5.32)$$

所以在临界点附近我们有： $m \sim \mathcal{H}/|T - T_c|$ ，这意味着系统的磁化率临界指数为： $\gamma = 1$ ，再一次与上一节中易兴模型的平均场近似的相应结果一致。如果我们考虑温度正好等于临界温度，那么系数 $a(T) = 0$ ，于是，我们发现： $\mathcal{H} \sim m^3$ ，这说明序参量对外磁场的临界指数 $\delta = 3$ ，也与上一节易兴模型平均场近似的相应结果一致。类似地，我们可以讨论系统的热容量在临界点附近的行为。可以证明：平均场近似的结果一样，朗道理论预言热容量在临界点两边有一个跃变（这一点，留作练习）。

综合以上的讨论我们发现：对于一个二级相变，朗道理论可以得到与平均场近似完全一致的结果。这种一致性并不是偶然的。我们在第 26 节中将说明，朗道理论从微观上讲就是一种平均场理论。当然，这两种方法（本质上是一致的）得到的结果都与实际的情况不符。这也说明了，朗道理论（或者说平均场近似）只是一个定性或半定量的唯象理论，不是一个精确的理论。

25.2 一级相变的朗道理论

现在我们考虑另外一种可能性。我们假定系统的自由能按照序参量 m 展开为：

$$f(m) = f_0 + \frac{a(T)}{2}m^2 + \frac{b(T)}{4}m^4 + \frac{c(T)}{6}m^6 + \dots \quad (5.33)$$

但是系数 $b(T)$ 并不是正的，而是负数。为了使得系统仍然有稳定的状态，我们假定系数 $c(T) > 0$ 永远成立。这时，系统中可以出现一个 **一级相变**。也就是说，在某个温度 $T_c^{(1)}$ 时，¹¹系统的序参量 m 可以发生一个 **有限的** 跃变。这个临界温度由下列两式联合确定：

$$\begin{aligned} \left. \frac{\partial f}{\partial m} \right|_{m_0} &= 0 = am_0 + bm_0^3 + cm_0^5, \\ f(m_0) - f(0) &= 0 = \frac{a}{2}m_0^2 + \frac{b}{4}m_0^4 + \frac{c}{6}m_0^6 \end{aligned} \quad (5.34)$$

¹¹这里我们用 $T_c^{(1)}$ 来表示一级相变的临界温度，以区别于前面讨论的二级相变的临界温度 $T_c^{(2)}$ 。

这个方程的解为：

$$m_0^2 = -\frac{3b(T_c^{(1)})}{4c(T_c^{(1)})}, \quad (5.35)$$

而系统一级相变的临界温度满足：

$$a(T_c^{(1)}) = \frac{3b(T_c^{(1)})^2}{16c(T_c^{(1)})}. \quad (5.36)$$

我们发现：在一级相变的临界温度处，我们仍然有： $a(T_c^{(1)}) > 0$ ，也就是说，一级相变的临界温度 **高于** 二阶相变的临界温度 $T_c^{(2)}$ 。由于我们假定系数 $b(T)$ 在温度 $T_c^{(1)}$ 附近是负的，所以系统在发生一级相变时，它的序参量会从 $m = 0$ 跃变到 $m = m_0 = \sqrt{-3b(T_c^{(1)})/4c(T_c^{(1)})}$ 。

25.3 三临界点的朗道理论

¶ 上一小节我们看到，系统发生一级相变时，一般来说系数 $a(T)$ 是正的。如果在一级相变发生时，系数 $a(T)$ 也恰好等于零，这时一级和二级相变将 **同时** 发生。一般来讲，事情不会这么巧，特别是如果系数 a, b, c 等等仅仅是温度的函数时。但是，如果这些系数还依赖于另外一个参量，¹² 比如说称为 Δ ，那么，只要适当调节参量 Δ ，就 **有可能** 使得在某个温度 T_t 和 Δ_t 时满足： $a(T_t, \Delta_t) = b(T_t, \Delta_t) = 0$ 。这时，我们称系统处于一个 **三临界点** (tricritical point)。在三临界点附近，系统的临界性质会表现出与一级或二级相变不同的性状。朗道理论也可以用来定性地讨论系统在三临界点附近的行为。

这时，一级相变的曲线满足下列方程：

$$a(\Delta, T) - \frac{3b^2(\Delta, T)}{16c(\Delta, T)} = 0, \quad (5.37)$$

而二级相变的临界点由 $a(\Delta, T) = 0$ 给出。这两条曲线在 $\Delta - T$ 平面的交点就是三临界点。两条曲线的斜率可以按照下面的方法求出。对于二级相变曲线，我们有：

$$\left. \frac{d\Delta}{dT} \right|_{(2)} = -\frac{\partial a / \partial T}{\partial a / \partial \Delta} \equiv -\frac{a_T}{a_\Delta}, \quad (5.38)$$

其中我们用简化的符号 a_T 和 a_Δ 来分别代表函数 $a(\Delta, T)$ 对于温度 T 和参量 Δ 的偏导数。而对于一级相变，我们得到：

$$\left. \frac{d\Delta}{dT} \right|_{(1)} = -\frac{a_T c + c_T a - \frac{3}{8} b b_T}{a_\Delta c + c_\Delta a - \frac{3}{8} b b_\Delta}. \quad (5.39)$$

¹² 这些另外的参量可以是外磁场、外电场、压强等等。

于是我们看到，在两根曲线相交的三临界点附近，由于 $a(\Delta, T) \rightarrow 0$ 并且 $b(\Delta, T) \rightarrow 0$ ，我们发现一、二级相变的相变曲线的斜率在三临界点附近也变得相同。这就意味着，在三临界点不仅两条曲线相交，而且它们是光滑联结在一起的（导数也连续）。

¶ 在 $T - \Delta$ 二维平面上我们选取三临界点为原点并令：

$$\mathbf{x} = \begin{pmatrix} T - T_t \\ \Delta - \Delta_t \end{pmatrix},$$

那么三临界点意味着 $\mathbf{x} \rightarrow \mathbf{0}$ ，系统的临界性质依赖于 \mathbf{x} 如何趋于原点。为了考察不同的可能性，我们假定 a, b 在原点附近都线性地依赖于 \mathbf{x} ：

$$a(T, \Delta) = \mathbf{a} \cdot \mathbf{x} + \dots, \quad b(T, \Delta) = \mathbf{b} \cdot \mathbf{x} + \dots \quad (5.40)$$

其中 \mathbf{a}, \mathbf{b} 为两个非零的二维矢量。另一个系数 $c(T, \Delta)$ 基本可以认为是常数。现在利用自由能的极值条件 $\partial f / \partial m = 0$ 我们发现序参量 m 满足方程 $a + bm^2 + cm^4 = 0$ ，即：

$$(\mathbf{a} \cdot \mathbf{x}) + (\mathbf{b} \cdot \mathbf{x})m^2 + cm^4 = 0. \quad (5.41)$$

在三临界点附近， $\mathbf{x} \rightarrow \mathbf{0}$ ，因此我们一定有： $m^2 \rightarrow 0$ ，但是其具体的行为将依赖于 \mathbf{x} 如何趋于 $\mathbf{0}$ 。前面提到， \mathbf{a}, \mathbf{b} 是两个一般的二维矢量（既不平行也不垂直）。那么对于 $\mathbf{x} \rightarrow \mathbf{0}$ 且 $\mathbf{x} \perp \mathbf{b}$ ，我们有 $m^4 \sim |\mathbf{x}|$ ，或者：

$$m(T, \Delta) \sim |\mathbf{x}|^{1/4}. \quad (5.42)$$

但在 $T - \Delta$ 平面上也存在使 $\mathbf{x} \perp \mathbf{a}$ 的情况下趋于零的可能，这种情况下我们就得到：

$$m(T, \Delta) \sim |\mathbf{x}|^{1/2}. \quad (5.43)$$

因此我们看到，在三临界点附近，按照不同的方式趋于三临界点，我们可以得到不同的临界指数。一般来讲，处于所谓三临界区域的点由关系 (5.42) 来表征，它们由一组临界指数描写，通常这一组临界指数用一个下标 t 来表示；另外一些处于临界区域的点由关系 (5.43) 来表征，它们由另外一组临界指数来描写，通常用脚标 u 来表示。因此我们发现，朗道理论预言： $\beta_t = 1/4$ ， $\beta_u = 1/2$ 。我们也可以求出其他相应的临界指数： $\gamma_t = 2$ ， $\gamma_u = 1$ ， $\alpha_t = 1/2$ ， $\alpha_u = -1$ 。

¶ 最后我们指出，从统计物理的微观角度来看，朗道理论做为一个热力学理论实际上等价于统计物理中的平均场理论，一般讲它只能对临界现象做出 **定性的** 或 **半定量** 的解释，它对临界指数的预言 **往往** 是不准确的。¹³真正的精确理论计算需要利用其他的方法，例如 **重整化群** 方法、Monte Carlo 模拟等等。

¹³一个令人吃惊的结果是，尽管朗道理论对于普通临界指数的预言与实际情况并不相符；朗道理论对于三临界点附近的临界指数（那些带有脚标 t 的临界指数）的预言却几乎是严格的。

26 Bragg-Williams 近似：再看平均场近似

¶ 我们在第 24 节中按照 Weiss 的思路讨论了易兴模型的平均场理论。事实上，平均场近似的结果可以从另外一个角度出发得到，这就是本节介绍的 Bragg-Williams 近似。从本质上说，Bragg-Williams 近似与平均场近似完全等价。

在 Bragg-Williams 近似中，我们假定各个格点上的自旋变量取 +1 和 -1 的几率与其他格点的自旋取值无关，是相互统计独立的。于是，每个格点上的自旋变量 σ_i 取 +1 和 -1 的几率可以统一由一个参数 m 描写：

$$p_{\pm} = (1 \pm m)/2, \quad (5.44)$$

其中 m 的物理意义就是每个格点上自旋的平均值，或者说，就是系统的序参量：

$$m = \langle \sigma_i \rangle = p_+ - p_- . \quad (5.45)$$

这时，整个系统构成了一个近独立子系。每一个格点上的自旋变量是统计独立的。或者用更物理一些的语言，在 Bragg-Williams 近似下，系统变成了一个顺磁系统。利用序参量 m ，我们可以写出系统的自由能：

$$F(T, \mathcal{H}) = \langle E - TS \rangle = -\frac{NqJ}{2}m^2 - N\mathcal{H}m + Nk_B T \left(\frac{1+m}{2} \log \frac{1+m}{2} + \frac{1-m}{2} \log \frac{1-m}{2} \right), \quad (5.46)$$

其中第一项是具有序参量 m 时的系统的相互作用能；第二项是各个自旋与外磁场的塞曼能量；第三项是系统熵对于自由能的贡献，它的计算可以利用玻耳兹曼对于熵的统计定义得到。¹⁴ 现在，我们要求参数 m 的取值必须使得自由能取极小值。这个条件给出：

$$-qJm - \mathcal{H} + \frac{k_B T}{2} \log \left(\frac{1+m}{1-m} \right) = 0,$$

或者等价的写成：

$$m = \tanh[\beta(qJm + \mathcal{H})]. \quad (5.47)$$

这个式子与第 24 节中我们得到的平均场近似下的自洽条件 (5.22) 完全一致。如果我们令上式中的外磁场等于零，我们马上就可以得到临界温度的结果： $T_c = qJ/k_B$ 。类似于前面第 24 节中的讨论，我们可以得到 Bragg-Williams 近似下的易兴系统的所有临界行为。当然，所有这些结果完全与第 24 节中平均场的结果相同。

¹⁴即： $S = k_B \log \Omega$ 。注意这时系统的微观状态数目就是： $\Omega = N!/(N_+!N_-!)$ ，其中 N_{\pm} 分别是自旋向上和向下的平均自旋数目，它们满足： $N_+ + N_- = N$ 。

¶ 从 Bragg-Williams 近似中的自由能 (5.46) 还可以发现一个重要的现象。如果我们令外磁场为零，同时假定温度接近于临界点，从而系统中的序参量 m 是一个小量，我们可以对自由能 $F(T, \mathcal{H} = 0)$ 对序参量做展开：¹⁵

$$\frac{F(T, \mathcal{H} = 0)}{N} = \frac{m^2}{2}(k_B T - qJ) + \frac{k_B T}{12} m^4 + \dots, \quad (5.48)$$

其中我们已经将与序参量无关的量略去了。于是我们发现，在平均场近似下（或者等价地说，在 Bragg-Williams 近似下），系统的自由能在临界点附近可以展开成序参量的幂级数，这正是朗道理论的最基本假设之一。同时，序参量的二次项的系数在临界点附近会变号，其他展开系数则都是恒正的。所有这些都与我们上一节的朗道理论中的假设一致。这种一致性再次告诉我们：平均场近似与唯象的朗道理论是等价的。这种等价性的另一个迹象就是我们前面已经提到的，两种方法给出的易兴模型的所有临界指数都相等。事实上，平均场近似与朗道理论的等价性可以在更为普遍的基础上加以证明（参考 [11, 16]）。

27 临界点附近的涨落与关联

¶ 前面我们仅仅讨论了系统在平衡时的序参量的确定，没有讨论序参量可能的涨落。这些涨落在通常情况下是很小的，但是如果系统非常接近临界点，那么涨落的影响就不能忽略不计了。

要讨论涨落，我们就必须讨论系统中非均匀的序参量的变化。我们可以借鉴朗道理论的思想来构造系统的自由能：

$$F[m(\mathbf{r})] = \int d\mathbf{r} \left[f_0(T) + \frac{a(T)}{2} m^2(\mathbf{r}) + \frac{d(T)}{2} [\nabla m(\mathbf{r})]^2 + \frac{b(T)}{4} m^4(\mathbf{r}) \right], \quad (5.49)$$

也就是说，系统的自由能是系统序参量 $m(\mathbf{r})$ 的泛函。这里我们略去了其他可能的高阶项。这种理论有时候又被称为朗道-金兹堡模型 (Landau-Ginzburg model)。¹⁶

关于这个自由能表达式我们需要稍微作一些说明。首先，我们这里采用了“连续的”符号。也就是说，空间坐标 \mathbf{r} 是连续的，而不是象前面所说的位于一些格点上。这实际上是基于一些假设的。我们实际上假设了系统可以分为一些宏观小、微观大的部分。每一个部分都可以用一个序参量 $m(\mathbf{r})$ 来描写，而且每一个部分内部系统是均匀的。当然，不同部分之间的序参量可以是不同的。具体来说，我们实际上假定 $m(\mathbf{r})$ 变化的特征长度比起

¹⁵ 更确切的展开应当用以温度和序参量为独立变量的热力学势，它与 $F(T, \mathcal{H})$ 相差一个勒让德变换。不过在外场为零时，两者数值是一样的。

¹⁶ 严格来说，Landau-Ginzburg 模型是描写超导体与外电磁场相互作用的一个模型。我们这里实际上只是借用其精神。

系统微观的格点间距大很多，以至于我们可以用“连续的”的语言来描写序参量。这一点在临界点附近是可以得到保证的，因为系统的关联长度趋于无穷大。

显然，对于均匀的（空间平移不变的）体系，使得自由能 (5.49) 极小的解 \bar{m} 也是均匀的并且满足：

$$\bar{m}^2 = \begin{cases} -a(T)/b(T) & T < T_c, \\ 0 & T > T_c. \end{cases} \quad (5.50)$$

这正前面讨论的朗道理论的结果。它们分别对应于系统的 **破缺相**（即 $\bar{m} \neq 0$ ）和 **对称相**（即 $\bar{m} = 0$ ）。现在考虑系统中存在的涨落，于是我们令

$$m(\mathbf{r}) = \bar{m} + \delta m(\mathbf{r}), \quad (5.51)$$

一个重要的物理量是序参量的 **关联函数**，它的最一般的定义是：

$$C(\mathbf{r}_1, \mathbf{r}_2) = \langle (m(\mathbf{r}_1) - \langle m(\mathbf{r}_1) \rangle)(m(\mathbf{r}_2) - \langle m(\mathbf{r}_2) \rangle) \rangle, \quad (5.52)$$

其中 $\mathbf{r}_1, \mathbf{r}_2$ 是两个任意的空间坐标。虽然对于最一般的系统它是两个坐标的函数，但是对于具有平移不变性的系统来说，它仅仅依赖于两者的相对坐标 $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ 。我们将仅仅考虑这类系统。对于这类系统我们有： $\langle m(\mathbf{r}_1) \rangle = \langle m(\mathbf{r}_2) \rangle = \bar{m}$ ，其中 \bar{m} 满足公式 (5.50)。关联函数是研究统计系统时十分重要的物理量，它包含了丰富的物理信息。我们这个课程无法“穷尽”它的所有内涵，我们将仅仅限于讨论它与临界涨落之间的关系。

对于平移不变的系统来说，将 $\delta m(\mathbf{r}) = m(\mathbf{r}) - \bar{m}$ 展开成傅立叶级数是方便的：

$$\delta m(\mathbf{r}) = m(\mathbf{r}) - \bar{m} = \frac{1}{V} \sum_{\mathbf{k}} \tilde{m}_{\mathbf{k}} e^{i\mathbf{k} \cdot \mathbf{r}}, \quad (5.53)$$

其中 V 是系统的体积， $\tilde{m}_{\mathbf{k}}$ 为序参量的傅立叶分量。它是一个 **复的** 变量。序参量 $\delta m(\mathbf{r})$ 的实性要求： $\tilde{m}_{\mathbf{k}} = \tilde{m}_{-\mathbf{k}}^*$ 。在一个有限的体积 V 中，所允许的 \mathbf{k} 的取值实际上是分立的。当然对于宏观的体积而言， \mathbf{k} 可以看成是准连续的。将公式两边乘以 $e^{-i\mathbf{p} \cdot \mathbf{r}}$ 并在体积 V 内对 \mathbf{r} 积分我们就可以得到傅里叶变换的逆变换：

$$\int d^3\mathbf{r} \delta m(\mathbf{r}) e^{-i\mathbf{p} \cdot \mathbf{r}} = \sum_{\mathbf{k}} \tilde{m}_{\mathbf{k}} \cdot \frac{1}{V} \int e^{i(\mathbf{k}-\mathbf{p}) \cdot \mathbf{r}} d^3\mathbf{r} = \tilde{m}_{\mathbf{p}}. \quad (5.54)$$

其中我们利用了： $(1/V) \int d^3\mathbf{r} \exp(i(\mathbf{k}-\mathbf{p}) \cdot \mathbf{r}) = \delta_{\mathbf{p},\mathbf{k}}$ 。因此， $\tilde{m}_{\mathbf{k}}$ 与 $\delta m(\mathbf{r})$ 是完全等价的。但是对于平移不变的体系而言，利用 $\tilde{m}_{\mathbf{k}}$ 来描写体系有它的方便之处。

我们下面希望用序参量涨落的傅里叶模式 $\tilde{m}_{\mathbf{k}}$ 来表达公式 (5.52) 中定义的关联函数。按照公式 (5.52) 中定义，我们有：

$$C(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{V^2} \sum_{\mathbf{k}_1, \mathbf{k}_2} \langle \tilde{m}_{\mathbf{k}_1} \tilde{m}_{\mathbf{k}_2} \rangle e^{i\mathbf{k}_1 \cdot \mathbf{r}_1 + i\mathbf{k}_2 \cdot \mathbf{r}_2}$$

现在我们将 \mathbf{r}_1 写为: $\mathbf{r}_1 \equiv \mathbf{r}_1 - \mathbf{r}_2 + \mathbf{r}_2 = \mathbf{r} + \mathbf{r}_2$, 其中 $\mathbf{r} \equiv \mathbf{r}_1 - \mathbf{r}_2$ 为两个点之间的相对坐标。正如我们前面指出过的, 对于平移不变的系统来说, 关联函数将只依赖于两点之间的相对坐标 \mathbf{r} , 不会明显依赖于 \mathbf{r}_2 。这样一来上式中的表达式变为:

$$C(\mathbf{r}_1, \mathbf{r}_2) \equiv C(\mathbf{r}) = \frac{1}{V^2} \sum_{\mathbf{k}_1, \mathbf{k}_2} \langle \tilde{m}_{\mathbf{k}_1} \tilde{m}_{\mathbf{k}_2} \rangle e^{i\mathbf{k}_1 \cdot \mathbf{r} + i(\mathbf{k}_2 + \mathbf{k}_1) \cdot \mathbf{r}_2}$$

因此我们可以将上式两边对于 \mathbf{r}_2 积分再除以体积 V 仍然得到类似的等式:

$$C(\mathbf{r}) = \frac{1}{V} \int d^3\mathbf{r}_2 C(\mathbf{r}) = \frac{1}{V^2} \sum_{\mathbf{k}_1, \mathbf{k}_2} \langle \tilde{m}_{\mathbf{k}_1} \tilde{m}_{\mathbf{k}_2} \rangle e^{i\mathbf{k}_1 \cdot \mathbf{r}} \delta_{\mathbf{k}_1 + \mathbf{k}_2, 0}$$

利用关联函数傅里叶模式之间的关系 $\tilde{m}_{\mathbf{k}} = \tilde{m}_{-\mathbf{k}}^*$, 我们发现关联函数可以表达为:

$$C(\mathbf{r}) = \frac{1}{V^2} \sum_{\mathbf{k}} \langle |\tilde{m}_{\mathbf{k}}|^2 \rangle e^{i\mathbf{k} \cdot \mathbf{r}} . \quad (5.55)$$

所以, 只要计算出序参量的傅立叶分量的模方的系综平均值 $\langle |\tilde{m}_{\mathbf{k}}|^2 \rangle$ 就可以得到系统的关联函数 $C(\mathbf{r})$ 。序参量的傅立叶分量反映了系统中序参量围绕其平均值的涨落情况。所以从这个式子也看出, 系统中的统计关联与涨落的强弱是联系在一起的。一般来讲, 系统内的涨落越强, 统计关联也越强。

为了计算 $\langle |\tilde{m}_{\mathbf{k}}|^2 \rangle$, 我们可以利用11节的关于涨落的准热力学理论。为此, 我们取温度和 $\tilde{m}_{\mathbf{k}}$ 为独立变量。因为温度的涨落与 $\tilde{m}_{\mathbf{k}}$ 的涨落相互独立, 我们可以进一步假定温度、体积恒定, 而仅仅研究 $\tilde{m}_{\mathbf{k}}$ 的涨落问题。这时涨落的普遍公式可以写成:

$$W \propto \exp\left(-\frac{\Delta F}{k_B T}\right) . \quad (5.56)$$

其中 ΔF 为系统存在涨落时, 自由能对于其平均值的偏离:

$$\Delta F = F - \bar{F} = \int \Delta f d\mathbf{r} , \quad (5.57)$$

按照朗道理论对于自由能的假设, 系统的自由能密度为:¹⁷

$$\Delta f = \frac{a(T)}{2} [\delta m(\mathbf{r})]^2 + \frac{d(T)}{2} [\nabla m(\mathbf{r})]^2 + \dots , \quad (5.58)$$

其中我们忽略了 $\delta m(\mathbf{r})$ 的高次项。平均场近似告诉我们, 在临界点附近有: $a(T) = a_0(T - T_c)$, 以及 $d(T) > 0$ 。所以, 我们可以得到:

$$\Delta F = \frac{1}{2V} \sum_{\mathbf{k}} (a(T) + d(T)\mathbf{k}^2) |\tilde{m}_{\mathbf{k}}|^2 . \quad (5.59)$$

¹⁷下面的计算是以对称相为例进行计算的, 即假定 $\bar{m} = 0$ 。所以这里的计算对应于 $T \rightarrow T_c^+$ 的情形, 即 $a(T) > 0$ 。如果要在破缺相 (其中 $a(T) < 0$) 中进行计算, 则需要考虑 $b(T)m^4$ 项中的贡献。结果是需要将下面的公式 (5.58) 中的 $a(T)$ 替换为 $2|a(T)|$ 。

这个公式充分体现了利用傅里叶分量 $\tilde{m}_{\mathbf{k}}$ 的优势。如果运用坐标空间的变量 $\delta m(\mathbf{r})$ 来表达, 体系自由能涨落 ΔF 关于 $\delta m(\mathbf{r})$ 是非对角的, 因为 ΔF 中涉及 $\nabla \delta m(\mathbf{r})$ 。但是如果利用傅里叶分量, 则 ΔF 关于 $\tilde{m}_{\mathbf{k}}$ 是完全对角的。将这样对角化的 ΔF 代入到涨落的概率 (5.56) 之中, 我们就得到完全独立的一系列高斯分布的乘积。于是我们得到:

$$\langle |\tilde{m}_{\mathbf{k}}|^2 \rangle = \frac{V k_B T}{a(T) + d(T) \mathbf{k}^2}, \quad (5.60)$$

将上式带入到关联函数的公式 (5.55) 中我们就得到:

$$C(\mathbf{r}) = \frac{k_B T}{V} \sum_{\mathbf{k}} \frac{e^{i\mathbf{k} \cdot \mathbf{r}}}{a(T) + d(T) \mathbf{k}^2}. \quad (5.61)$$

如果我们假定体积非常大 (宏观的), 用积分代替求和我们得到:

$$C(\mathbf{r}) = \frac{k_B T}{4\pi d(T)} \frac{e^{-r/\xi}}{r}, \quad (5.62)$$

其中我们定义了 **关联长度** ξ :

• • • •

$$\xi = \sqrt{\frac{d(T)}{a(T)}}, \quad (5.63)$$

关联长度的物理意义十分明确, 它代表了系统内两点之间的关联随距离衰减的特征长度。一般情况下, 关联长度是一个有限的正数。所以, 系统内部的两个点之间的关联随它们距离的增加而指数地衰减。但是, 如果系统处于临界点附近, 情况就不同了。由于在临界点附近 $a(T) \sim |T - T_c| \sim 0$, 我们发现在临界点附近关联长度 ξ 趋于无穷大:

$$\xi \sim |T - T_c|^{-\nu}, \quad \nu = \frac{1}{2}, \quad (5.64)$$

这个式子定义了另一个临界指数 ν 。在平均场近似下 (朗道理论), 临界指数 $\nu = 1/2$ 。在实际的系统中, 二阶相变临界点附近的关联长度也是发散的, 只不过临界指数 ν 并不等于平均场理论所预言的 $1/2$ 。在临界点上时, 系统内序参量的统计关联不再随距离而指数衰减, 而是表现出 **长程关联** (仅仅是按照距离的幂次衰减):

• • • •

$$C(\mathbf{r}) \sim r^{-D+2-\eta}, \quad (5.65)$$

其中 D 是系统的维数。上式实际上定义了另一个临界指数 η 。在平均场近似下, 我们有 $\eta = 0$ 。对于实际的系统, 实验测定的 $\eta \sim 0.1$ 。

¶ 通过这一节的讨论和分析我们看到, 关联与涨落是相互联系的。在临界点附近, 系统内的统计关联长度 ξ 发散, 系统内的涨落也趋于无穷。这是所有二阶相变 (临界现象) 最普遍、最明显的标志性特征, 同时也是造成普适性的主要原因。

28 易兴模型的严格解、高温展开和对偶性

¶ 前面几节中，我们侧重讨论了易兴模型的平均场近似。为了说明除了平均场近似以外还存在其他的方法，同时也为了与平均场近似进行比较，这一节中，我们介绍一下求解易兴模型的另外一些方法。我们将介绍一维易兴模型的严格解，说明它在有限温度并不存在相变；然后我们讨论针对二维正方晶格上的易兴模型的另一种方法——**高温展开方法**。同时，我们将顺便说明这时易兴模型所具有的一种特殊对称性——**对偶性**。

28.1 一维易兴模型的严格解

¶ 首先来看一维晶格上的 Ising 模型。为明确起见，我们假定系统具有 N 个格点并且我们加上**周期边条件**。我们将系统的哈密顿量写为：

$$H = -J \sum_{i=1}^N \sigma_i \sigma_{i+1} - \frac{\mathcal{H}}{2} \sum_{i=1}^N (\sigma_i + \sigma_{i+1}) \quad (5.66)$$

上式中我们将哈密顿量写为了更加对称的形式。注意，周期边条件意味着我们必须将第 $N+1$ 个格点与第一个格点等同起来，即： $\sigma_{N+1} \equiv \sigma_1$ 。于是，体系的正则系综配分函数为：

$$Z_N = \sum_{\{\sigma_i\}} \prod_{i=1}^N \exp \left[\beta \left(J \sigma_i \sigma_{i+1} + \frac{\mathcal{H}}{2} (\sigma_i + \sigma_{i+1}) \right) \right] . \quad (5.67)$$

现在我们引进所谓的**转移矩阵** (transfer matrix) \mathbb{T} 并且用 $+1, -1$ (而不是通常的 $1, 2$) 来标志它的行和列：

$$\mathbb{T} = \begin{pmatrix} \mathbb{T}_{1,1} & \mathbb{T}_{1,-1} \\ \mathbb{T}_{-1,1} & \mathbb{T}_{-1,-1} \end{pmatrix}, \quad (5.68)$$

其中的矩阵元分别等于：

$$\mathbb{T}_{1,1} = e^{\beta(J+\mathcal{H})}, \mathbb{T}_{-1,-1} = e^{\beta(J-\mathcal{H})}, \mathbb{T}_{-1,1} = \mathbb{T}_{1,-1} = e^{-\beta J} \quad (5.69)$$

注意这些矩阵元恰好是公式 (5.67) 中的每个因子在 σ_i 和 σ_{i+1} 分别取 ± 1 时的结果。于是系统的配分函数可以写为：

$$Z_N = \sum_{\{\sigma_i\}} \mathbb{T}_{\sigma_1 \sigma_2} \mathbb{T}_{\sigma_2 \sigma_3} \cdots \mathbb{T}_{\sigma_N \sigma_1} = \text{Tr} (\mathbb{T}^N) . \quad (5.70)$$

我们知道一个矩阵的迹在正交变换下是不变的。因此，我们需要做的就是将转移矩阵 \mathbb{T} 对角化，当然它的 N 次幂也相应地对角化了。令转移矩阵 (5.68) 的模较大的本征值为 λ_1 ，较小的为 λ_2 ，那么我们有：

$$Z_N = (\lambda_1)^N + (\lambda_2)^N \simeq (\lambda_1)^N, \quad (5.71)$$

其中我们忽略了 λ_2 的贡献是因为 N 是一个非常大的数 (10^{23} 的量级)。将矩阵元 (5.69) 代入, 我们解得:

$$\lambda_1 = e^{\beta J} \cosh(\beta \mathcal{H}) + \sqrt{e^{2\beta J} \sinh^2(\beta \mathcal{H}) + e^{-2\beta J}}$$

所以体系的自由能可以写为:

$$F = -Nk_B T \ln \left[e^{\beta J} \cosh(\beta \mathcal{H}) + \sqrt{e^{2\beta J} \sinh^2(\beta \mathcal{H}) + e^{-2\beta J}} \right]. \quad (5.72)$$

我们可以计算系统的序参量:

$$m = \langle \sigma_1 \rangle = -\frac{1}{N} \frac{\partial F}{\partial \mathcal{H}}$$

经过一些初等的运算后得到:

$$m = \frac{\sinh(\beta \mathcal{H})}{\sqrt{\sinh^2(\beta \mathcal{H}) + e^{-4\beta J}}}. \quad (5.73)$$

我们发现, 如何令外磁场为零 $\mathcal{H} = 0$, 那么序参量 $m = 0$, 这对于任何非零的温度都是如此。因此体系不会发生顺磁-铁磁相变。

¶ 我们还可以讨论一维 Ising 链的自旋关联函数。为此, 类比于第 27 节的定义式 (5.52), 这时的关联函数为: $C(i, j) = \langle \sigma_i \sigma_j \rangle$ 。由于体系具有一维平移不变性, 我们可以令

$$C(j) = \langle \sigma_1 \sigma_j \rangle. \quad (5.74)$$

注意两者之间的间隔是 $(j - 1)$ 倍的格距。

为了计算这个关联函数我们暂时将 σ_i 与 σ_{i+1} 之间的耦合参数写成 J_i :

$$H = - \sum_{i=1}^N J_i \sigma_i \sigma_{i+1}. \quad (5.75)$$

我们在最后会令 $J_1 = J_2 = \dots = J_N = J$, 但是在计算过程中保留它们的不同会给计算带来很大方便之处。于是, 自旋关联函数可以表达为:

$$C(j) = \frac{1}{Z_N} \sum_{\{\sigma_i\}} (\sigma_1 \sigma_j) e^{\sum_{i=1}^N \beta J_i \sigma_i \sigma_{i+1}}$$

其中 $Z_N = Z_N(J_1, \dots, J_N) = \sum_{\{\sigma_i\}} e^{\sum_{i=1}^N \beta J_i \sigma_i \sigma_{i+1}}$ 为体系的配分函数。为了计算这个配分函数, 我们再次利用前面引入的转移矩阵。对于具有不同耦合参数的一维 Ising 模型的哈密顿量, 它的配分函数可以写为:

$$Z_N = \text{Tr} \left(\prod_{i=1}^N \mathbb{T}^{(i)}(J_i) \right), \quad (5.76)$$

其中第 i 个转移矩阵为:

$$\mathbb{T}^{(i)}(J_i) = \begin{pmatrix} e^{\beta J_i} & e^{-\beta J_i} \\ e^{-\beta J_i} & e^{\beta J_i} \end{pmatrix} = e^{\beta J_i} \cdot \mathbb{1} + e^{-\beta J_i} \tau_1, \quad (5.77)$$

其中 τ_1 是转移矩阵空间的 Pauli 矩阵。为了操作方便我们定义投影算符:

$$P_{\pm} = \frac{1 \pm \tau_1}{2}. \quad (5.78)$$

这样一来上面的转移矩阵可以写为:

$$\mathbb{T}^{(i)}(J_i) = 2 [\cosh(\beta J_i) P_+ + \sinh(\beta J_i) P_-]. \quad (5.79)$$

现在我们可以利用投影算符的性质得到:

$$\prod_{i=1}^N \mathbb{T}^{(i)}(J_i) = 2^N \left[P_+ \cdot \prod_{i=1}^N \cosh(\beta J_i) + P_- \cdot \prod_{i=1}^N \sinh(\beta J_i) \right].$$

于是系统的配分函数可以表达为:

$$Z_N = 2^{N-1} \left(\prod_{i=1}^N \cosh(\beta J_i) + \prod_{i=1}^N \sinh(\beta J_i) \right). \quad (5.80)$$

现在注意到:

$$\sigma_1 \sigma_j = (\sigma_1 \sigma_2)(\sigma_2 \sigma_3) \cdots (\sigma_{j-1} \sigma_j),$$

所有插入的中间的因子都统统等于 $(\sigma_k)^2 \equiv +1$ 。但是这样一来, 上式中的每个因子可以通过对相应的 J_i 求偏导数获得。因此, 我们有:

$$C(j) = \frac{1}{\beta^{j-1} Z_N} \left. \frac{\partial^{j-1} Z_N(J_1, \dots, J_N)}{\partial J_1 \partial J_2 \cdots \partial J_{j-1}} \right|_{J_i=J} = \frac{\tanh^{j-1}(\beta J) + \tanh^{N-j+1} \tanh(\beta J)}{1 + \tanh^N(\beta J)} \quad (5.81)$$

由于我们总是有 $\tanh(\beta J) < 1$ 且 $(N-j+1) \gg 1$, 所以我们可以分母中略去 $\tanh^N(\beta J)$, 分子中略去 $\tanh^{N-j+1}(\beta J)$ 。因此我们得到:

$$C(j) \propto e^{-j/\xi}, \quad \xi = -\frac{1}{\ln(\tanh(\beta J))}. \quad (5.82)$$

其中我们定义了关联长度 ξ 。对于极低温, $\ln(\tanh(\beta J)) \simeq -2e^{-2\beta J}$, 因此我们得到:

$$\xi \simeq \frac{1}{2} e^{2\beta J}. \quad (5.83)$$

所以在低温时, 关联长度指数地增大, 但是它永远不会发散。

28.2 二维易兴模型的高温展开与对偶性

¶下面我们讨论二维易兴模型。为了简单和明确起见，我们考虑一个二维正方晶格上的、无外磁场的易兴模型：

$$H = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j , \quad (5.84)$$

其中所有的自旋变量 $\{\sigma\}$ 都定义于二维正方晶格的格点上并且只能取值 ± 1 。系统的正则配分函数为：

$$Z = \sum_{\{\sigma_i\}} \prod_{\langle ij \rangle} e^{\beta J \sigma_i \sigma_j} . \quad (5.85)$$

现在注意到易兴模型的特殊性：由于所有的 σ_i 只能取 ± 1 ，我们有：

$$e^{\beta \sigma_i \sigma_j} = \cosh(\beta J) + \sigma_i \sigma_j \sinh(\beta J) . \quad (5.86)$$

所以我们可以将配分函数写成：

$$Z = [\cosh(\beta J)]^{2N} \cdot \sum_{\{\sigma_i\}} \prod_{\langle ij \rangle} [1 + \sigma_i \sigma_j \tanh(\beta J)] . \quad (5.87)$$

这个式子特别适合来做所谓的高温展开。在温度很高时， $\tanh(\beta J) \ll 1$ ，因此上式中的连乘式中最大的一项贡献就是每一个因子都取 1。接下来的贡献就是尽可能的少取 $\tanh(\beta J)$ 的因子而多取 1 的因子。但是需要注意的一点是，每当我们取了一个 $\tanh(\beta J)$ 的因子时，与之一起必然同时出现一个 $\sigma_i \sigma_j$ 的因子。由于每一个格点上的自旋变量 σ_i 的求和取 ± 1 两个值，因此如果某个给定的 σ_i 的因子仅仅出现奇数次，那么对于这个自旋变量的求和以后必定给出零的贡献，因为： $(+1) + (-1) \equiv 0$ 。因此，所有的非零贡献中，必定每个自旋变量都出现偶数次（包括零次）。这个事实可以十分方便地用几何的方法来表述。从几何上说，如果我们将带来每一个 $\sigma_i \sigma_j \tanh(\beta J)$ 的因子的近邻对 $\langle ij \rangle$ 都作一个标记，例如将这些键都画成红色，那么所有这些被标记的近邻对必定在二维正方晶格上构成一个闭合回路，因为只有闭合的回路才能保证这个回路所连接的每一个格点上的 σ_i 因子都是偶数次幂。如果我们将一个孤立的（不在任何闭合回路中的）格点本身也看成是一种特殊的（长度为零的）闭合回路，那么我们发现：公式 (5.87) 中的求和的每一项都与一个几何上的图（diagram）一一对应。这些图中，每一个格点都必定在某个闭合回路（包括孤立的点，它被视为零长度的闭合回路）中。而每一个特定的图对于配分函数的贡献的规则也十分简单：每一个闭合回路中的紧邻联结贡献一个因子 $\tanh(\beta J)$ ；孤立的点不贡献。所以，易兴模型的配分函数可以写成：

$$Z = [\cosh(\beta J)]^{2N} \cdot \sum_{\Gamma} [\tanh(\beta J)]^{L(\Gamma)} , \quad (5.88)$$

其中 Γ 代表格点上所有不同的闭合回路， $L(\Gamma)$ 代表这个闭合回路的周长（以晶格格距为单位）。闭合的回路保证了每个晶格点上的自旋变量一定出现偶数次。我们看到，首项的贡献是 1，这对应于一个零长度的闭合回路。下一项的非零贡献是正比于 $\tanh^4(\beta J)$ ，这对应于晶格上的某个小方块回路（周长为 4）的贡献，等等。我们这里并不打算详细讨论这个高温展开，我们仅仅是指出这个展开到目前为止是严格的，没有进行任何的近似。在进行真正计算时，人们必须将展开式截断到某一阶。¹⁸ 由于这些图是与高温展开联系起来的，因此它们被称为高温展开图。利用几何的图形来表达物理的公式，这在我们这个课程中已经不是第一次了。前面（第 20 节）讨论的迈尔集团展开也是这种思想。统计物理中经常会利用图论的方法来计算某种展开。这种利用几何上的图来计算展开的方法同时也出现在其他分支中，例如量子场论的微扰展开（所谓的费曼图）。

公式 (5.88) 中对于格点上所有不同的闭合回路的求和也可以用更为明确的方法加以表述。为此，我们在格点的所有键上定义一个整数，对于那些组成闭合回路的键 l ，我们约定： $n_l = 1$ ；对于那些不属于闭合回路的键 l ，我们定义： $n_l = 0$ 。于是，易兴模型的配分函数的高温展开式也可以写成下面的等价的形式：

$$Z = [\cosh(\beta J)]^{2N} \cdot \sum_{\{n_l\}, \partial\{n_l\}=0} [\tanh(\beta J)]^{\sum_l n_l} , \quad (5.89)$$

其中符号 $\partial\{n_l\} = 0$ 表示所有取值为 $n_l = 1$ 的键构成一个闭合回路（这是数学上的一个符号，表示没有边界）；而 $\sum_l n_l$ 表示所有键上的 n_l 之和，也就是所有构成闭合回路的键的总数。只要从几何上数出格点上不同长度的拓扑不等价的闭合回路的数目，就可以给出易兴模型的配分函数一个很好的近似展开式。这种计数的工作对于周长比较小的闭合回路是直接了当的。但是，随着回路周长的增加，这个工作会越来越烦琐。事实上，到比较高的阶数，往往必须借助于计算机来完成。总之，经过适当的计算，我们就可以得到易兴模型配分函数的高温展开。可以证明，这种展开在二阶相变的临界点以上一定是收敛的。但是需要指出的是，在临界点处这个级数一定会发散。

¶ 现在我们要利用易兴模型的高温展开的形式表达式来引入所谓的对偶变换。由原来晶格的每个小方格的中心可以构成一个与原来晶格类似的二维正方晶格。这个晶格称为原来晶格的对偶晶格。于是，对于满足 $\partial\{n_l\} = 0$ 条件的每一组 $\{n_l\}$ ，我们可以在原来晶格的对偶晶格上重新定义一个自旋变量 $\tau_a = \pm 1$ ，这些新的自旋变量的取值按照下列法则来进行：如果原来晶格的某个键上 $n_l = 1$ ，这个键两侧的 τ 就取异号；如果原来晶格的某个键上 $n_l = 0$ ，这个键两侧的 τ 就取同号。这样一来，对于每一组满足 $\partial\{n_l\} = 0$ 条件的 $\{n_l\}$ ，我们正好有：

$$\sum_l n_l = \sum_{\langle ab \rangle} (1 - \tau_a \tau_b) / 2 , \quad (5.90)$$

¹⁸对于易兴模型，人们已经计算这个展开式到几十阶。当然，这里的困难主要在于数清楚不同的图的数目。

即我们将对于原来晶格上的所有键的一个求和（等式左边）与对偶晶格上的键的求和联系起来。稍微在纸上比画一下¹⁹就可以发现：条件 $\partial\{n_l\} = 0$ 刚好保证了这样定义的 τ_a 是自洽的，不会出现任何的矛盾。如果所有的非零的 n_l 不构成一个回路，就可能会有一个对偶晶格上的格点的自旋变量无法自恰定义（一方面要求取 +1，另一方面要求取 -1。）。现在我们定义一个对偶温度 $\tilde{\beta}$ ：

$$e^{-2\tilde{\beta}J} = \tanh(\beta J) , \quad (5.91)$$

这样的变换称为对偶变换，我们可以利用它将原先的易兴模型的配分函数表达成新的变量 τ_a 的配分函数，最为重要的一点是：新的的自旋变量的配分函数，仍然是一个易兴模型的配分函数（除去一些无关紧要的因子之外）：

$$Z(\beta) = [\sinh(2\beta J)]^N \frac{1}{2^{N+1}} \sum_{\{\tau_a\}} e^{\tilde{\beta}J \sum_{\langle ab \rangle} \tau_a \tau_b} . \quad (5.92)$$

由此我们得到了两个对偶的系统的自由能之间的关系：

$$\frac{F(\beta)}{N} + \frac{1}{\beta} \log \sinh(2\beta J) = \frac{F(\tilde{\beta})}{N} , \quad (5.93)$$

由于按照定义 (5.91) 容易证明：

$$\sinh(2\beta J) \sinh(2\tilde{\beta} J) = 1 , \quad (5.94)$$

至此我们发现：自由能和对偶温度的表达式都说明两个系统互为对方的对偶变换。

对偶变换 (5.94) 将系统的高温和低温联系在了一起。换句话说，易兴模型在极低温下的配分函数可以用它在极高温时的配分函数来表达，而后者可以利用高温展开很好的计算。另外，当易兴系统达到二级相变的临界点时，系统内的关联长度达到无穷。这时它的对偶格点上的易兴模型的关联长度一定也是无穷大。所以，我们知道临界温度一定满足： $\beta_c = \tilde{\beta}_c$ 。于是我们求出了二维正方晶格上易兴模型的严格临界温度满足：

$$\sinh(2\beta_c J) = 1 , \quad \beta_c J = \frac{1}{2} \log(1 + \sqrt{2}) = 0.440686794 \dots , \quad (5.95)$$

¹⁹强烈建议大家画一下，如果你真想搞明白的话。

这个对偶性的结论首先是 Kramers 和 Wannier 在 1941 年得到的。这个结果早于 Onsager 公布他的著名的二维易兴模型的严格解 (1944 年)。这个严格的结果也同时说明平均场近似的结果是不正确的。当然, 仅仅依靠对偶性虽然能够确定系统的严格的临界温度, 但仍然不足以解出系统的所有热力学性质。二维正方晶格上的易兴模型的完全解析解首先由 Onsager 得到。有兴趣的同学可以参考 [11, 16] 中的讨论。

¶ 对偶变换以及对偶性还被应用于理论物理的其他方面。对于每个晶格上的自旋模型都可以在其对偶晶格上定义它的对偶模型。但是二维正方晶格上易兴模型的特性是, 它的对偶模型仍然是二维正方晶格上的易兴模型, 或者说二维正方晶格上的易兴模型具有 **自对偶性**。正是这种自对偶性使得我们能够利用它来严格的确定临界温度。其他具有 $\bullet \bullet \bullet \bullet$ 自对偶性的模型还有正方晶格上的所谓的 Potts 模型。²⁰

29 具有连续对称性系统的相变

¶ 前面我们主要讨论了自旋模型中最为简单的一个模型 - 易兴模型 - 的统计性质。这一节中, 我们简要介绍一下其他稍微复杂一些的自旋模型, 也就是自旋变量可以连续取值的自旋模型。为了明确起见, 我们将讨论 **海森堡模型** (Heisenberg model)。我们的讨论将仅仅局限于平均场近似 (Bragg-Williams 近似) 的框架之内。类似的讨论可以轻易地推广到其他自旋模型 (例如 XY-模型)。我们随后讨论与 Heisenberg 模型的相变相对应的朗道-金兹堡模型。这里由于对称性的不同, 会出现在易兴模型中没有过的现象。

29.1 海森堡模型 Weiss 分子场理论

在海森堡模型中, 每个晶格格点上的自旋变量有三个分量, 它取值在一个单位球面上, 也就是说: $\mathbf{S}_i = (S_{i,x}, S_{i,y}, S_{i,z})$ 并且满足: $S_{i,x}^2 + S_{i,y}^2 + S_{i,z}^2 = 1$ 。海森堡模型的哈密顿量可以写为:

$$H = -J \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j - \sum_i \mathbf{S}_i \cdot \mathcal{H}, \quad (5.96)$$

其中 \mathcal{H} 也是一个三个分量的矢量, 它代表了外加的均匀磁场。不失普遍性, 我们假定它沿着 z 方向, 其大小为 \mathcal{H} 。

按照平均场近似的思想, 我们将某个给定的自旋变量紧邻的其他自旋变量用一个有效的平均场来替代:

$$\mathcal{H}_{eff} = qJ \langle S_{i,z} \rangle, \quad (5.97)$$

²⁰所谓的 q -态 Potts 模型的哈密顿量为: $H = -J \sum_{\langle ij \rangle} (2\delta_{\sigma_i \sigma_j} - 1)$, 其中每个晶格点上的自旋变量可以取 q 个不同的值。也就是说, 当一个近邻对上的自旋变量取值相同时对于哈密顿量有 $-J$ 的贡献; 当它们取值不同时对于哈密顿量有 $+J$ 的贡献。显然, 易兴模型对应于 $q = 2$ 的 Potts 模型。

并且引入平均场哈密顿量:

$$H^{(MF)} = - \sum_i \mathbf{S}_{i,z} (\mathcal{H}_{eff} + \mathcal{H}). \quad (5.98)$$

显然, 等效磁场 \mathcal{H}_{eff} 应当也沿着 z 方向 (均匀外加磁场的方向)。

在平均场近似下, 每个格点上的自旋变量的统计平均只有 z 方向不为零:

$$\mathbf{m}_z = \langle \mathbf{S}_{i,z} \rangle = \left(\coth[\beta(\mathcal{H}_{eff} + \mathcal{H})] - \frac{1}{\beta(\mathcal{H}_{eff} + \mathcal{H})} \right). \quad (5.99)$$

结合 (5.99) 和 (5.97) 给出 \mathcal{H}_{eff} 所满足的 **自洽条件**:

$$\mathcal{H}_{eff} = qJ \left(\coth[\beta(\mathcal{H}_{eff} + \mathcal{H})] - \frac{1}{\beta(\mathcal{H}_{eff} + \mathcal{H})} \right). \quad (5.100)$$

按照 $\coth(x)$ 在 $x \sim 0$ 附近的展开式: $\coth(x) \sim 1/x + x/3 + \dots$, 我们得到海森堡模型的临界温度满足:

$$T_c = \frac{qJ}{3k_B}. \quad (5.101)$$

我们同样可以在平均场近似下讨论各种临界指数。我们得到的结论是: 所有的临界指数都与易兴模型的临界指数相同。这个结论当然仅仅是由于采用了平均场近似, 而且实际上也是不正确的。

29.2 对称性的考虑与 Goldstone 模式

¶ 这一小节的讨论完全类比于第 27 节中关于 Ising 模型中涨落与关联的讨论。当然也有不同的地方。最为显著的不同之处在于 Ising 模型中我们的自由能具有所谓的 Z_2 对称性, 而这里 Heisenberg 模型的对称性变为了 $O(3)$ 。

我们在第 27 节中用坐标依赖的序参量 $m(\mathbf{r})$ 来描写 Ising 模型中的涨落与关联。注意到在体系的自由能 (5.49) 中仅仅保留了 $m(\mathbf{r})$ 的偶数幂次。这个源于 Ising 模型的哈密顿量具有所谓的 Z_2 对称性: 即如果令每个 $\sigma_i \rightarrow (-\sigma_i)$, 则哈密顿量不变。这意味着序参量 $m(\mathbf{r}) = \langle \sigma_{\mathbf{r}} \rangle$ 构成的自由能也应当具有 $m(\mathbf{r}) \rightarrow (-m(\mathbf{r}))$ 的对称性。这就是我们在自由能 (5.49) 中仅仅保留了 $m(\mathbf{r})$ 的偶数幂次的原因。

对于海森堡模型来说, 体系的哈密顿量具有所谓的 $O(3)$ 对称性。即如果我们将每个自旋变量 \mathbf{m}_i 做任意的三维转动:

$$\mathbf{m}_i \rightarrow \mathbf{m}'_i = R_{ij} \mathbf{m}_j \quad (5.102)$$

那么哈密顿量也是不变的。因此与之相应的序参量 - 我们记为 $\mathbf{m}(\mathbf{r})$ - 由它构成的体系自由能应当也具有转动不变性。需要注意的是, 现在的序参量 $\mathbf{m}(\mathbf{r}) = \langle \mathbf{S}_{\mathbf{r}} \rangle$ 在空间的每个点 \mathbf{r}

具有三个分量: $m^a(\mathbf{r})$, $a = 1, 2, 3$ 。于是类似于自由能 (5.49), 我们可以写出海森堡模型相应的 Landau-Ginzburg 自由能:

$$F[\mathbf{m}(\mathbf{r})] = \int d^3\mathbf{r} \left[\frac{a}{2} [m^a(\mathbf{r})m^a(\mathbf{r})] + \frac{d}{2} [\nabla m^a(\mathbf{r})] \cdot [\nabla m^a(\mathbf{r})] + \frac{b}{4} [m^a(\mathbf{r})m^a(\mathbf{r})]^2 \right] \quad (5.103)$$

其中我们略写了系数 a, b, d 对温度的依赖。

类似于 Ising 模型的讨论, 我们可以讨论使自由能取极小的均匀的解: $\bar{\mathbf{m}}$ 。这也是一个三分量的矢量。我们令:

$$\bar{\mathbf{m}} = \bar{m} \cdot \mathbf{n}, \quad (5.104)$$

其中 $\bar{m} = |\bar{\mathbf{m}}|$ 而 $\mathbf{n} = \bar{\mathbf{m}}/\bar{m}$ 为 $\bar{\mathbf{m}}$ 方向上的单位矢量。类似于公式 (5.50), 我们可以得到 \bar{m} 所满足的方程:

$$\bar{m}^2 = \begin{cases} -a(T)/b(T) & T < T_c, \\ 0 & T > T_c. \end{cases} \quad (5.105)$$

但是 $\bar{\mathbf{m}}$ 的方向 \mathbf{n} 则完全是任意的。这也并不奇怪。对空间均匀的分布来说, 自由能 (5.103) 仅仅依赖于 $\mathbf{m}(\mathbf{r})$ 的大小。

下面我们讨论系统围绕 $\bar{\mathbf{m}}$ 的涨落。我们的讨论将在所谓的 **破缺相** 中进行, 即 $T < T_c$ 。在第 27 节的 Ising 模型的讨论中, 我们的计算是在对称相中进行的。这里我们选择了破缺相, 因为只有在破缺相中我们才会发现 Ising 模型中没有的现象—Goldstone 模式的出现。²¹ 为此我们令:

$$\mathbf{m}(\mathbf{r}) = \bar{\mathbf{m}} + \phi(\mathbf{r}) = \bar{m}\mathbf{n} + \phi(\mathbf{r}). \quad (5.106)$$

其中 $\bar{m}^2 = -a/b$ 。下面我们需要做的就是将这个展开式带入到自由能 (5.103) 中并将其用 $\phi(\mathbf{r})$ 来表达。包含梯度的项 (方括号中的第二项) 最为简单, 因为我们有 $\nabla m^a(\mathbf{r}) \cdot \nabla m^a(\mathbf{r}) = \nabla \phi \cdot \nabla \phi$ 。另外两项中都包含 $m^a(\mathbf{r})m^a(\mathbf{r})$, 它的展开为

$$m^a(\mathbf{r})m^a(\mathbf{r}) = \bar{m}^2 + 2\bar{m}\mathbf{n} \cdot \phi + \phi \cdot \phi,$$

当我们把这个表达式带入到自由能 (5.103) 中的第一、第三项时, 按照不同的 ϕ 的幂次, 会出现下列各项:

- ϕ 的零次幂: 这就是一个常数项, 我们记为 F_0
- ϕ 的一次幂: 这项的系数一定为零, 因为 $\bar{m}\mathbf{n}$ 是 F 的一个极小值;
- ϕ 的二次幂: 这项可以表达为:

$$\int d^3\mathbf{r} \left[\frac{a}{2} \phi \cdot \phi + \frac{b\bar{m}^2}{4} (2\phi \cdot \phi + 4(\mathbf{n} \cdot \phi)^2) \right]$$

²¹如果在对称相中计算, 那么 Heisenberg 模型中的计算无非就是 Ising 模型计算的翻版而已。没有任何新的东西。

需要注意的是, 由于 $b\bar{m}^2 = -a$, 因此上式中第二项中的 $\phi \cdot \phi$ 的部分刚好与第一项相消。因此准确到 ϕ 的平方阶, 我们得到系统的自由能为:

$$F^{(2)}[\phi] = \int d^3\mathbf{r} \left[\frac{d}{2} \nabla\phi \cdot \nabla\phi + |a|(\mathbf{n} \cdot \phi)^2 \right]. \quad (5.107)$$

• ϕ 的更高幂次: 这部分代表了超出二阶涨落的部分, 我们称之为相互作用部分。

我们下面的讨论将仅仅考虑到涨落 ϕ 的平方阶。更改阶的相互作用的部分我们暂时不去考虑。

完全类比于第 27 节中关于 Ising 模型中涨落与关联的讨论, 现在我们从自由能的公式 (29.2) 出发可以计算相应的关联函数。将这里的自由能展开式与第 27 节中的相应表达式进行比较是有益的。对于只有分立对称性的 Ising 模型, 公式 (5.58) 给出了它到涨落的二阶的自由能表达式。它与 (29.2) 十分类似。注意到我们总可以将场 ϕ 进行分解:

$$\phi(\mathbf{r}) = \mathbf{n}\sigma(\mathbf{r}) + \phi_T(\mathbf{r}), \quad (5.108)$$

其中的场 $\phi_T(\mathbf{r})$ 是与磁化方向 \mathbf{n} 垂直的分量, 即:

$$\phi_T(\mathbf{r}) \cdot \mathbf{n} = 0 \quad (5.109)$$

换句话说, 我们将场 ϕ 按照系统的自发磁化的方向 \mathbf{n} 分为平行于 \mathbf{n} 的分量 σ 和垂直于 \mathbf{n} 的分量 ϕ_T 。那么我们有:

$$\nabla\phi \cdot \nabla\phi = \nabla\sigma \cdot \nabla\sigma + \nabla\phi_T \cdot \nabla\phi_T,$$

于是我们可以将前面的二阶自由能 () 改写为:

$$F^{(2)}[\sigma, \phi_T] = \int d^3\mathbf{r} \left[\frac{d}{2} \nabla\phi_T \cdot \nabla\phi_T + \frac{d}{2} (\nabla\sigma)^2 + |a|\sigma^2 \right]. \quad (5.110)$$

我们看到, 上式中包含 σ 场的部分与公式 5.58 完全类似, 只不过 $2|a|$ 替代了原先的 a ; 而包含 ϕ_T 的部分形式也十分类似, 只不过相当于令 $a \equiv 0$ 。

关联函数的具体计算与第 27 节完全类似。我们分别将涨落场的纵向和横向的分量分别作傅里叶变换, 就可以分别得到用相应的傅里叶模式 $-\tilde{\sigma}_{\mathbf{k}}$ 和 $\tilde{\phi}_{\mathbf{k}}$ 表达的自由能。它们都与公式 (5.59) 类似, 只不过对于纵向情形 $a(T)$ 变为 $2|a(T)|$ 而对横向分量则令 $a(T) \equiv 0$ 。由此我们看到, 对于具有连续对称性的自旋模型来说, 它围绕其序参量的涨落分为两种典型的类型:

1. 一种涨落模式是涨落的方向与系统自发磁化平行, 即所谓的纵向涨落。其行为也与单分量的 Ising 模型的情况十分类似。具体来说, 其关联函数一般随着距离指数衰减的:

$$C(r) = \frac{k_B T}{4\pi d(T)} \frac{e^{-r/\xi}}{r}, \quad (5.111)$$

其中的关联长度的表达式为：

$$\xi = \sqrt{\frac{d(T)}{2|a(T)|}}. \quad (5.112)$$

在临界点附近它会发散： $\xi \sim |T - T_c|^{-1/2}$ 。如果我们令 $M = 1/\xi$ ，那么关联函数又可以写为 $C(r) \sim e^{-Mr}/r$ 。一般来说在临界区域 $M \neq 0$ ，只有当 $T \rightarrow T_c$ 时， $M \rightarrow 0$ 。我们又称 M 为相应涨落的 **能隙**。因此，对于平行于自发磁化的涨落来说，它是有能隙的。这种涨落我们又称之为 **Higgs 模式**。

2. 另一种涨落的模式是与系统的自发磁化强度的方向垂直，即横向涨落。相应的关联函数相当于在上面的表达式中令 $a(T) \equiv 0$ ，或者等价地说， $M = 1/\xi = 0$ 。因此，这类涨落属于永远没有能隙的涨落，相应的关联函数也是随着距离幂次衰减的： $C(r) \sim 1/r$ 。这样的无能隙的涨落被称为 **Goldstone 模式**。²² 如果将这种涨落想象为准粒子，它们对应于无质量的粒子。因此也被称为 **Goldstone 粒子**，或者 **Goldstone 玻色子**。

前面曾经提到，如果是无能隙的 **Goldstone 模式** 的涨落，其关联函数衰减得要慢很多。因此，体系也体现出 **长程关联**。关联过于强一般就有可能破坏磁有序的形成。在讨论一维 **Ising 模型** 严格解时我们就指出过，一维 **Ising 模型** 之所以没有相变就是因为在一维系统的长程关联足以破坏业已形成的磁有序。因此该模型在一维没有自发磁化的相变。到了二维，**Ising 模型** 可以有顺磁-铁磁的二级相变。这是因为它的涨落和关联-如我们在第 27 节分析的那样-总是有能隙的，这并不是足够强。但是在二维，如果我们考虑具有 **连续对称性** 的自旋模型（例如 **XY-模型** 或者 **Heisenberg 模型**），那么由于上面提及的 **Goldstone 模式** 的存在，它们是不会发生顺磁-铁磁的二级相变的。换句话说，即使形成了某种磁有序，**Goldstone 模式** 的涨落是如此之强，完全可以破坏掉这样的磁有序。这个结论又被称为 **Mermin-Wagner 定理**。²³ 它的证明远远超出了本课程的内容，有兴趣的同学可以参考底下的参考文献。

30 对称性破缺与临界现象中的普适性

¶ 人们通过对于二级相变的研究发现：系统发生二级相变所对应的两个相（或者说系统在两个相所处的状态）实际上具有不同的 **对称性**。系统在较高温度时所处的相一

²²其实最先意识到这个现象的是南部阳一郎 (Yoichiro Nambu)。因此这里提及的所有带有 Goldstone 开头的词汇，也被称为 Nambu-Goldstone 加上后面相应的词汇。例如：Nambu-Goldstone 模式，Nambu-Goldstone 粒子，Nambu-Goldstone 定理等等。

²³又称为 Mermin-Wagner-Hohenberg 定理。在粒子物理和场论中往往又被称为 Coleman 定理。相关的文献为：N.D. Mermin and H. Wagner, Phys.Rev.Lett.**17**, 1133 (1966); P.C. Hohenberg Phys.Rev.**158**, 383 (1967); S. Coleman Commun.Math.Phys.**31**, 259 (1973)。

般具有较高的对称性，而在临界点以下的相一般具有较低的对称性。以我们讨论的易兴模型的铁磁—顺磁相变为例，在临界温度以上，系统处于顺磁相，这个相中的序参量 $m = \langle \sigma_i \rangle = 0$ ，同时系统所处的状态具有较高的对称性。具体来说，如果我们将每一个自旋变量改变符号，序参量一定也改变符号。而序参量 $m = 0$ 的顺磁相正好满足： $m = -m = 0$ 。也就是说，用序参量描写的系统的状态对于 $\sigma_i \rightarrow -\sigma_i$ 的变换是不变的。我们称系统的状态具有一个 Z_2 的分立对称性。²⁴ 实际上，这时系统所具有的对称性与系统的哈密顿量的对称性相同，因为对于易兴模型来说，哈密顿量也具有 Z_2 的分立对称性。²⁵

在居里温度以下，易兴系统处于铁磁相。这时系统有一个自发磁化。用序参量来描写就是： $m \neq 0$ 。因此系统的序参量对于自旋变量的符号改变不是不变的： $m' \equiv -m \neq m$ 。所以我们说铁磁相不具有 Z_2 的分立对称性。也就是说，铁磁—顺磁的二级相变过程可以看成是系统从一个具有 Z_2 对称性的相到一个不具有这个对称性的相的转变过程。或者用更为专业一些的语言来说：易兴模型的铁磁—顺磁相变（二级相变）是一个对称性破缺的过程。

上面关于对称性的讨论实际上不是易兴模型所特有的，而是所有二级相变的共同特性。例如，类似的讨论可以运用于无外场的海森堡模型。海森堡模型与易兴模型的最大区别是它的哈密顿量具有一个连续的对称性。也就是说，我们可以将其自旋变量的三个分量（它们构成一个三维矢量）利用一个空间转动矩阵加以变换，由于海森堡模型的哈密顿量只依赖于两个自旋变量的内积，因此它的哈密顿量在任意空间转动变换下是不变的。这种对称操作显然具有连续的无穷多种。因此我们称它具有一个连续对称性。在群论中，这个对称性被称为 $O(3)$ ，即三维空间转动对称性。在高温顺磁相中，海森堡模型的序参量 $\mathbf{m} = \langle \mathbf{S}_i \rangle = 0$ 。因此，在任何空间转动矩阵的作用下： $\mathbf{m}'_i = R_{ij} \mathbf{m}_j = 0 = \mathbf{m}$ 。也就是说，在高温相（顺磁相）中，海森堡模型所处的状态也具有与其哈密顿量相同的对称性：它们对于任意空间转动都是不变的。但是，在低温相（铁磁相）中，系统存在一个自发的磁化。因此海森堡系统的序参量 $\mathbf{m} \neq 0$ 。相应的，在一个任意的三维转动变换下： $\mathbf{m}'_i = R_{ij} \mathbf{m}_j \neq 0$ 一般来说并不等于 \mathbf{m} ；除非这个转动正好是以 \mathbf{m} 方向为轴的转动。所以，这时海森堡系统所处的状态仅仅是对于那些以 \mathbf{m} 为轴的二维转动是不变的；对于一般的三维转动，系统的状态不是不变的。也就是说，在低温相中，海森堡系统具有较低的对称性（用群论的符号就是 $O(2)$ 对称性）。我们再次看到，这个二阶相变的过程可以看成是一个对称性破缺的过程：系统从高温的、具有三维转动不变对称性的相（顺磁相）变化到低温的、仅仅具有二维转动不变对称性的相（铁磁相）。

在临界现象中，系统的对称性起着决定性的作用。在连续相变中，对称性的变化标志

²⁴也就是说，这种作用于自旋变量的对称变换只有有限多种。具体到易兴模型的 Z_2 对称性，作用于自旋变量 σ_i 的对称变换只有两种：一种是不变，一种是改变一个符号。

²⁵这里用到的符号 Z_2 以及下面讨论中的 $O(3)$ 和 $O(2)$ 都是群论中的符号。每一种对称性都可以用一个称为群的代数结构来描写。

了两个不同的相。一般来讲，在一个相中系统具有较高的、与系统的哈密顿量相同的对称性（一般是高温相），在另一个相中系统具有相对低的对称性。²⁶ 需要注意的是，系统的哈密顿量本身无论温度高低都是具有较高对称性的。但是如果温度低于临界温度时，系统所处的状态，或者更确切地说，系统能量最低的态（基态）却不具备这个对称性。因此，P. W. Anderson 称这种现象为 **对称性自发破缺**。事实上，所有的二级相变（或者说连续相变）都可以看成是对称性自发破缺的过程。

¶ 二级相变还有一个共同的特性，那就是二阶相变系统中的关联长度在临界点附近发散（参见公式 (5.64) 后相关的讨论）。这种发散性造就了二阶相变的另一个普遍特性，即 **标度性**。²⁷ 粗略地来说，由于系统中的关联长度是系统中唯一的物理尺度，而这个尺度在临界点附近发散，这就使得处于临界点附近的物理系统没有了标度。换句话说，我们可以用不同的尺子来测量系统，其结果都相似。进一步来讲，用不同的尺子测量相当于对长度作一个标度变换，即所有长度都乘以一个正的实数。但是由于关联长度 $\xi \rightarrow \infty$ ，因此，标度变换后它仍然发散。也就是说，系统在临界点呈现出 **自相似性**。这种自相似性使得系统的物理量具有所谓的 **标度律**。这种标度律正是利用 **重整化群方法** 研究临界现象的理论基础所在。

¶ 我们在自旋模型的讨论中主要是以易兴模型为例子的。实际上，人们还研究过许多其他的自旋模型。这些模型的物理背景千差万别，来自物理的各个领域而不仅仅局限于统计物理。人们惊奇地发现，许多原先完全无关的物理模型，它们的某些临界行为（例如临界指数）竟然表现出高度地一致性。在临界点附近，系统内的关联长度趋于无穷；物理量都有确定的临界行为，由相应的临界指数描写。不仅如此，人们发现这些临界指数具有某种普适性。例如，易兴模型的临界指数在气液相变的临界点附近也被发现。两个看似毫无关系的相变却拥有相同的临界指数。这种现象又在其他一些模型中得到验证。于是，人们总结出所谓的临界现象的 **普适性**（Universality）的概念：系统在连续相变的临界点附近的行为由一些具有一定普适性的临界指数描写。这些临界指数只与系统在两相中的对称性、系统的维数等有关；与系统的微观细节（例如晶格结构等等）无关。

注意，普适性并不是说系统的所有的性质都与系统的微观细节无关。它只是告诉我们临界指数（实际上还有其他一些物理量）是普适的。我们前面提到的系统的临界温度 T_c 就不是一个普适的量，它与系统的微观细节有关。应当指出的是，普适性只有在自旋之间的相互作用是 **短程** 时才能成立的。如果系统中的相互作用的力程不是短程的，那么普适性的结论有可能不成立。²⁸ 普适性的根本物理原因恰恰在于前面提到的事实，即当系统

²⁶ 用群论的语言来说，具有较低对称性的相所对应的对称群是具有较高对称性的相所对应的对称群的 **子群**。例如，对于海森堡模型，我们有： $O(2) \subset O(3)$ 。

²⁷ 实际上，更确切地说应当称之为“无标度性”。

²⁸ 一般来说，我们要求相互作用是局域的，也就是说，相互作用随着距离的增加必须指数地减小。

接近临界区域时，系统内的特征关联长度 ξ 趋于无穷大。因此，这时只有大尺度上的物理才变得重要而系统在微观尺度上的细节完全不起任何作用。具有不同微观细节的模型所体现出来的普适性的规律实际上是连续空间中的一个统计场论。



相关的阅读

这一章我们讨论了相变，特别是二级相变的统计理论。我们的讨论是以自旋模型为具体的例子进行的，但很多临界性质实际上是普适的。我们所用的方法主要是平均场理论，对于其他方法也做了简单的介绍。此外，我们还利用朗道理论讨论了一级相变和三临界点的性质。虽然所有的讨论主要是以最为简单的易兴模型为例子的，我们所讨论的平均场的方法也可以运用到其他模型中。我们对于具有连续的对称性的模型做了简单的介绍。

附录 A 概率与随机过程

我们这个课程中经常地使用的一个概念就是几率，或者用数学上更为通用的说法叫做概率。¹另外，随即过程也是与统计物理密切相关的数学分支。在这个课程中我们并不需要概率论和随即过程的详细知识，我们只是将概率论和随机过程的基本概念和结果做一个介绍而并不要求数学上的严格。对此有偏好的人可以去参照专门的数学书籍。

1 概率、随机变量与分布函数

从数学上讲，概率是定义在一个样品空间（sample space）上的一种归一的测度（measure）。要讨论清楚概率的数学基础，必须利用集合和测度的数学概念。我们这里并不想仔细的论述，有兴趣的读者可以参考有关的数学教材。所谓归一，就是说所有可能事件中必有至少一个发生（这个集合称为全集，一般记为 Ω ）的概率是一： $P(\Omega) \equiv 1$ 。

两个事件，以集合 A 和 B 来表示，那么集合 $A \cap B$ 则代表了 A 与 B 同时出现的概率。如果 $A \cap B = \emptyset$ ，其中 \emptyset 代表空集，也就是说 A 与 B 两个事件不可能同时出现，我们说 A 与 B 两个事件是互相排斥的。两个事件 A 与 B 的并： $A \cup B$ 代表这两个事件至少有一个出现的事件集合。如果这两个事件是互相排斥的，我们有： $P(A \cup B) = P(A) + P(B)$ 。更为一般地，我们有：

$$P(A \cup B) = P(A) + P(B) - P(A \cap B), \quad (\text{A.1})$$

如果两个事件 A 与 B 满足： $P(A \cap B) = P(A)P(B)$ ，我们称这两个事件是独立的。两个事件的条件概率 $P(B|A)$ 代表了出现事件 B 的同时，事件 A 出现的概率，它由下式定义：

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \quad (\text{A.2})$$

条件概率实际上是以 A 为全集时，事件 B 出现的概率。

¹我们在课程中总是用“几率”这个词，与此相关的词语还有：“几率密度”、“等几率”、“最可几”等等。如果用数学上更通用的“概率”一词，上述词语应当变为：“概率密度”、“等概率”、“最概然”等等。

一个 **随机变量** (stochastic variable) X 是从样本空间到实数空间的一个映射, 它的取值我们一般用 $\{x_i\}$ 来表示。随机变量可以分为离散随机变量和连续随机变量两类。所谓离散随机变量是指样本空间 Ω 上的随机变量 X 的可能取值的集合: $X(\Omega) = \{x_1, x_2, \dots\}$ 是一个分立的可数集合。我们可以由 $X(\Omega)$ 为样本空间构成一个概率空间, 它对于每一个 X 的可能取值 x_i 赋予一个概率 $f(x_i)$ 。显然, $f(x_i)$ 必须是正的, 并且满足归一条件:

$$\sum_i f(x_i) = 1, \quad (\text{A.3})$$

这样的概率函数 f 称为 (离散的) 随机变量 X 的分布函数。

一个连续的随机变量 X 的取值集合 $X(\Omega)$ 是一个不可数的集合, 这里我们以实数域 R 为例, 但也可以是更为复杂的集合。这时我们可以以 $X(\Omega) = R$ 为样本空间, 构造一个概率空间。如果存在一个分段连续的函数 $f(x)$, 使得 X 的取值落在区间 $[a, b]$ 中的概率等于:²

$$P(a \leq X \leq b) = \int_a^b f(x) dx, \quad (\text{A.4})$$

我们就称 $f(x)$ 为 (连续的) 随机变量 X 的分布函数或称为 **概率密度**。

2 常见的概率分布函数

2.1 二项式分布

如果我们进行 N 次统计独立的实验, 每次实验的结果只可能为两个: 比如说 $+1$ 和 -1 , 他们对应的概率分别为 p 和 $q = 1 - p$ 。那么在总共 N 次实验中, $+1$ 出现 n_1 次, -1 出现 n_2 次 (显然 $n_1 + n_2 = N$) 的组合概率为:

$$P_N(n_1) = \frac{N!}{n_1!n_2!} p^{n_1} (1-p)^{N-n_1}, \quad (\text{A.5})$$

这个分布就称为 **二项式分布**。

一个最为直接的物理的例子, 就是所谓的一维无规行走问题, 或者说, 这也是一维粒子的布朗运动的一个相当好的描述 (至少在长时间尺度上)。

2.2 Poisson 分布

如果在二项式分布中令 $N \gg 1$ 和 $p \ll 1$, 但是 $Np = \bar{n}$ 为有限, 那么我们得到:

$$P_N(n) = \frac{\bar{n}^n}{n!} e^{-\bar{n}}, \quad (\text{A.6})$$

这个分布是归一化的: $\sum_n P_N(n) = 1$, 并且 n 的平均值为 \bar{n} , 而其涨落为: $\overline{(\Delta n)^2} = Np(1-p) \approx Np = \bar{n}$ 。

²注意, 这里的积分是在 Lesbegue 意义下的积分, 而不是 Riemann 意义下的积分。

2.3 高斯正态分布

一个连续随机变量 x ，如果其分布函数为：

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right), \quad (\text{A.7})$$

那么我们称之为高斯（正态）分布，其中 \bar{x} 为变量 x 的平均值（期望值）， σ^2 是变量 x 的均方涨落。

3 随机过程

考虑一个具有随机过程的系统，该系统的性质由一个随机变量 ξ 来描述，我们用 $P_1(\xi_1, t_1)$ 来表示随机变量 ξ 在时刻 t_1 时取值为 ξ_1 的概率（密度）；用 $P_2(\xi_2, t_2; \xi_1, t_1)$ 来表示随机变量 ξ 在时刻 t_1 时取值为 ξ_1 、在时刻 t_2 时取值为 ξ_2 的联合概率。更为普遍的，我们用 $P_n(\xi_1, t_1; \xi_2, t_2; \cdots; \xi_n, t_n)$ 来表示随机变量 ξ 在 t_1, t_2, \cdots, t_n 时刻分别取值为 $\xi_1, \xi_2, \cdots, \xi_n$ 的联合概率。这些概率分布可以逐级地被约化为次一级的概率分布：

$$\int P_n(\xi_1, t_1; \xi_2, t_2; \cdots; \xi_n, t_n) d\xi_n = P_{n-1}(\xi_1, t_1; \xi_2, t_2; \cdots; \xi_{n-1}, t_{n-1}). \quad (\text{A.8})$$

最后，概率分布 $P_1(\xi_1, t_1)$ 是归一化的。

人们通常感兴趣随机变量 ξ 在不同时刻的关联函数。系统的一个 n 点关联函数被定义为：

$$\langle \xi_1(t_1)\xi_2(t_2)\cdots\xi_n(t_n) \rangle = \int \xi_1\xi_2\cdots\xi_n P_n(\xi_1, t_1; \xi_2, t_2; \cdots; \xi_n, t_n) d\xi_1 d\xi_2 \cdots d\xi_n. \quad (\text{A.9})$$

一个所谓 **平稳随机过程** 是指对于任意的 n 和 t_0 都有：

$$P_n(\xi_1, t_1 + t_0; \xi_2, t_2 + t_0; \cdots; \xi_n, t_n + t_0) = P_n(\xi_1, t_1; \xi_2, t_2; \cdots; \xi_n, t_n). \quad (\text{A.10})$$

也就是说，所有的概率分布具有时间平移不变性。物理系统在达到平衡或稳态时，其概率分布一定是平稳的。做为上面这个定义的一个特例，一个平稳随机过程的概率 $P_1(\xi_1, t_1) = P_1(\xi_1)$ 与时间无关。

我们引入（一到一）的条件概率 $P_{1|1}(\xi_1, t_1 | \xi_2, t_2)$ ，它的定义是：在时刻 t_1 取值为 ξ_1 的随机变量 ξ 在时刻 t_2 取值 ξ_2 的概率（密度）。它满足：

$$P_1(\xi_1, t_1) P_{1|1}(\xi_1, t_1 | \xi_2, t_2) = P_2(\xi_1, t_1; \xi_2, t_2), \quad (\text{A.11})$$

条件概率 $P_{1|1}(\xi_1, t_1 | \xi_2, t_2)$ 又被称为 **转移概率**（transition probability），它反映了一个随机过程在不同时刻、不同状态之间的转移的几率，我们有：

$$P_1(\xi_2, t_2) = \int P_1(\xi_1, t_1) P_{1|1}(\xi_1, t_1 | \xi_2, t_2) d\xi_1. \quad (\text{A.12})$$

类似于上面引入的一到一条件概率，一般来讲，我们还可以引入 p 到 q 的条件概率： $P_{p|q}(\xi_1, t_1; \cdots; \xi_p, t_p | \xi_{p+1}, t_{p+1} \cdots; \xi_{p+q}, t_{p+q})$ 。

4 Markov 过程

一个一般的随机过程中，刻划它的概率函数由一系列的密度 P_n 和条件概率 $P_{p|q}$ 来描述。如果一个随机过程对过去的记忆是短暂的（好幸福啊），具体的说，如果它只有最近的过去的记忆，那么这个过程从 $n-1$ 到 n 的条件概率必定满足：

$$P_{n-1|1}(\xi_1, t_1; \cdots; \xi_{n-1}, t_{n-1} | \xi_n, t_n) = P_{1|1}(\xi_{n-1}, t_{n-1} | \xi_n, t_n) . \quad (\text{A.13})$$

一个满足上式的随机过程称为一个 **马尔可夫过程** (Markov process)，一个马尔可夫过程完全由 $P_1(\xi, t)$ 和其转移概率 $P_{1|1}(\xi_1, t_1 | \xi_2, t_2)$ 所确定。例如，对于 P_3 ，我们有：

$$P_3(\xi_1, t_1; \xi_2, t_2; \xi_3, t_3) = P_1(\xi_1, t_1) P_{1|1}(\xi_1, t_1 | \xi_2, t_2) P_{1|1}(\xi_2, t_2 | \xi_3, t_3) . \quad (\text{A.14})$$

也就是说，在一个马尔可夫过程中，每一次的跃迁都是独立的，多次跃迁的几率就是单次跃迁几率的乘积。

下面我们研究一个马尔可夫过程中概率密度随时间的变化率。取一个无穷小时间间隔 τ 并取 $t_2 = t_1 + \tau$ ，我们有：

$$P_1(\xi_2, t_1 + \tau) = \int P_1(\xi_1, t_1) P_{1|1}(\xi_1, t_1 | \xi_2, t_1 + \tau) d\xi_1 . \quad (\text{A.15})$$

现在，我们将上式对小的 τ 做泰勒展开。注意到： $P_{1|1}(\xi_1, t_1 | \xi_2, t_1) = \delta(\xi_1 - \xi_2)$ ，同时，考虑到 $P_{1|1}$ 的归一化在 $O(\tau)$ 也得以保持，我们必定有：

$$P_{1|1}(\xi_1, t_1 | \xi_2, t_1 + \tau) = \delta(\xi_1 - \xi_2) \left(1 - \tau \int d\xi W_{t_1}(\xi_1, \xi) \right) + \tau W_{t_1}(\xi_1, \xi_2) . \quad (\text{A.16})$$

这里我们引入了单位时间中的转移概率 $W_t(\xi_1, \xi_2)$ ，于是，我们得到：

$$\frac{\partial P_1(\xi_2, t)}{\partial t} = \int d\xi_1 [W(\xi_1, \xi_2) P_1(\xi_1, t) - W(\xi_2, \xi_1) P_1(\xi_2, t)] , \quad (\text{A.17})$$

这就是著名的 **主方程**。注意，我们假定了单位时间的跃迁几率 $W(\xi_1, \xi_2)$ 并不明显地依赖于时间。

细心的读者一定会注意到：这里的主方程与我们在非平衡统计中推导的玻耳兹曼微分积分方程十分类似。实际上，玻耳兹曼微分积分方程可以作为主方程的一个特例来得到。

5 Fokker-Planck 方程

如果随机变量 ξ 的取值是连续的，那么它的变化只能连续地进行，这时主方程可以进一步简化为一个纯粹的微分方程，这就是著名的福克-普朗克方程 (Fokker-Planck

equation)。我们用 $y = \xi_2 - \xi_1$ 来表示随机变量的微小变化，将主方程对 y 展开，我们得到：

$$\frac{\partial P_1(\xi, t)}{\partial t} = -\frac{\partial}{\partial \xi} [\alpha_1(\xi) P_1(\xi, t)] + \frac{1}{2} \frac{\partial^2}{\partial \xi^2} [\alpha_2(\xi) P_1(\xi, t)] , \quad (\text{A.18})$$

其中函数 $\alpha_1(\xi)$ 和 $\alpha_2(\xi)$ 为跃迁的矩，其定义为：

$$\alpha_n(\xi) = \int dy y^n W(\xi, \xi + y) . \quad (\text{A.19})$$

物理学中许多方程都具有福克-普朗克方程的形式，典型的就是扩散方程。如果想象力丰富一些，你还会联想到薛定谔方程。

参 考 书

- [1] 汪志诚, 《热力学•统计物理》, 高等教育出版社, 2003年第三版。
- [2] 王竹溪, 《热力学》, 高等教育出版社, 1960年第二版。
- [3] 王竹溪, 《统计物理导论》, 人民教育出版社, 1956年第一版。
- [4] L. D. Landau and E. M. Lifschitz, *Statistical Physics*, Pergamon Press, 1994.
- [5] L. D. Landau and E. M. Lifschitz, *Electrodynamics of Continuous Media*, Pergamon Press, 1994.
- [6] L. D. Landau and E. M. Lifschitz, *Quantum Mechanics*, Pergamon Press, 1994.
- [7] M. W. Zemansky, *Heat and Thermodynamics*, McGraw-Hill, 1968.
- [8] K. Huang, *Statistical Mechanics*, John Wiley & Sons, 1967.
- [9] H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics*, John Wiley & Sons, 1985.
- [10] L. Reichl *A Modern Course in Statistical Physics*, Austin, University of Texas Press, 1980.
- [11] M. Plischke and B. Bergerson, *Equilibrium Statistical Physics*, Prentice Hall, 1989.
- [12] F. Mohling, *Statistical Mechanics: methods and applications*, Publishers Creative Services Inc., John Wiley and Sons, Inc. 1982.
- [13] M. Rasetti, *Modern Methods in Equilibrium Statistical Mechanics*, World Scientific, Singapore, 1986.
- [14] R. Jancel, *Foundations of Classical and Quantum Statistical Mechanics*, Pergamon Press, Oxford, 1969.
- [15] R. Balescu, *Equilibrium and Non-equilibrium Statistical Mechanics*, Wiley, New York, 1975.
- [16] C. Itzykson and J. M. Drouffe, *Statistical Field Theory*, Vol. I, Cambridge University Press, 1989.

- [17] J.D. Jackson, *Classical Electrodynamics*, John Wiley & Sons, Inc., 1975.
- [18] Stephen G. Brush, *Statistical Physics and Atomic Theory of Matter: from Boyle and Newton to Landau and Onsager*, Princeton University Press, 1983.
- [19] 《中国大百科全书》(物理卷), 中国大百科全书出版社。
- [20] Irving M. Klotz and Robert M. Rosenberg, *Chemical thermodynamics: basic theory and methods*, John Wiley and Sons, Inc. 2000.
- [21] W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed., Cambridge University Press, 1992.

索引

- XY 一模型, 95
- Born-Oppenheimer 近似, 75n
- Goldstone 模式, 120
- Higgs 模式, 120
- Mermin-Wagner 定理, 120
- Monte Carlo 数值模拟, 96
- Monte Carlo 模拟, 104
- Potts 模型, 116
- Wiedermann-Franz 定律, 62
- 爱因斯坦温度, 57
- 昂尼斯方程, 82
- 标度律, 122
- 标度性, 122
- 玻耳兹曼关系, 29
- 玻尔磁子, 25
- 玻色-爱因斯坦凝聚, 48, 50
- 玻色-爱因斯坦统计, 26, 27
- 玻色分布, 41
- 玻色子, 24, 26
- 波函数, 24
- 长程关联, 87, 109
- 重求和方法, 82
- 重整化群, 104
- 重整化群方法, 96, 122
- 磁化率, 94
- 磁矩, 25
- 磁量子数, 25
- 代表点, 20
- 单相系, 2
- 单元系, 2
- 道尔顿分压定律, 76
- 德拜 T^3 律, 59
- 德拜函数, 59
- 德拜截止频率, 59
- 德拜屏蔽长度, 89
- 德拜温度, 59
- 等几率原理, 28
- 等温压缩系数, 36
- 第二维里系数, 82
- 第三维里系数, 82
- 定压平衡恒量, 78
- 定域系, 26
- 杜隆-珀替定律, 56
- 对称相, 107
- 对称性破缺, 121
- 对称性自发破缺, 122
- 对分布函数, 85
- 对关联函数, 86
- 对偶变换, 114
- 对偶晶格, 114

- 对偶温度, 115
对偶性, 110
- 多项式分布, 44
多元系, 2
多元系的复相平衡条件, 16
- 二级相变, 14
- 反应热, 17
反应物, 16
范德瓦耳斯方程, 83
范德瓦耳斯相互作用, 83n
范霍夫方程, 78
- 非定域系, 26
非简并条件, 42
非理想气体, 79
非凝聚相, 12
非线性 $O(n) - \sigma$ 模型, 95
费米-狄拉克统计, 27
费米-狄喇克统计, 26
费米波矢, 63
费米动量, 63
费米分布, 41
费米面, 63
费米能量, 63
费米球, 63
费米速度, 63
费米子, 24, 26
- 分布, 40
分布函数, 84
 单粒子, 85
 双粒子, 85
分立对称性, 121
分子场, 97
分子动力学方法, 21n, 87
- 复相系, 2
- 高级相变, 14
高温展开, 96, 113
高温展开方法, 110
高温展开图, 114
- 各态历经假设, 21
- 固相, 12
- 关联长度, 109
关联函数, 107
- 广延量, 14, 30
广义动量, 20
广义力
 热力学, 3
广义坐标, 20
 热力学, 3
- 哈密顿方程, 20
哈密顿量, 20
- 海森堡模型, 95, 116
- 焓, 8
- 好量子数, 24
- 赫斯定律, 17
- 化学平衡条件, 10, 17
化学势, 9, 15
 名称的由来, 11
- 回转因子, 25
- 混合假设, 22n
- 活度, 48
- 基尔霍夫方程, 17
吉布斯分布, 32
吉布斯关系, 16
吉布斯函数, 8

- 吉布斯佯谬, 30n, 78
 集团, 80
 集团积分, 81
 集团展开理论, 79

 简并度, 41
 简并压, 64
 简正模式, 72

 交换积分, 93
 交换相互作用, 92, 93
 角动量量子数, 25

 结构因子, 85

 近独立子系, 38

 经典极限, 27
 经典流体, 69
 经典统计, 27

 居里温度, 94
 巨配分函数
 非理想气体的, 79
 巨正则分布, 33

 绝对热力学温标, 5
 绝对热力学温度, 5
 绝对温度, 5

 卡诺定理, 4
 卡诺循环
 理想气体的, 4

 克拉珀龙方程, 12
 克劳修斯不等式, 5

 空间波函数, 92

 朗道-金兹堡模型, 106
 朗道抗磁性, 65
 朗道能级, 65, 67

 勒让德变换, 8
 勒沙特列原理, 79

 理想玻色气体, 41, 46
 理想费米气体, 41, 46
 力学平衡条件, 10

 连续对称性, 121
 连续相变, 99

 量子理想气体, 46
 量子数, 23
 量子态, 24
 量子统计, 27
 量子性, 23

 列纳德-琼斯势, 83

 临界点, 36
 临界乳光现象, 37
 临界温度, 98
 临界指数, 99
 磁化率, 100
 热容量, 100
 序参量, 99
 序参量对磁场的, 100

 刘维定理, 22

 麦克斯韦-玻耳兹曼分布, 39
 麦克斯韦-玻耳兹曼统计, 26
 麦克斯韦关系, 8
 麦氏关系, 8

 蒙特卡罗方法, 87, 96

 密度矩阵, 23
 密度算符, 23

 内能, 3

 能带论, 61
 能均分定理, 56

- 能壳, 20
 能量守恒定律, 3
- 凝聚相, 12
- 泡利不相容原理, 26
 泡利顺磁性, 65, 67
- 配分函数
 巨正则系综的, 33
 正则系综的, 32
- 配位数, 97
- 偏摩尔吉布斯函数, 15
 偏摩尔内能, 15
 偏摩尔熵, 15
 偏摩尔体积, 15
- 平衡的稳定性条件, 12
 平衡态, 1
 平衡条件, 2, 10
 平均场, 96
 平均场近似, 96, 97
 平均分布, 40
 平稳随机过程, 126
- 破缺相, 107
- 普朗克公式, 53
 普适性, 92, 122
- 气体的热容量
 理想气体的, 72
- 气相, 12
- 强度量, 14, 30
- 全同性, 23
 全同性原理, 24, 92
- 热波长, 48
 热动平衡, 10
 热力学不等式, 12
- 热力学第二定律, 4
 Clausius 表述, 4
 Kelvin 表述, 4
- 热力学第零定律, 2
 热力学第一定律, 3
 热力学公式
 巨正则系综的, 34
 正则系综的, 33
- 热力学函数
 理想气体的, 70
- 热力学基本微分方程, 6
 PVT 系统的, 7
- 热力学势, 8
 热力学体系, 1
 状态方程, 2
 状态空间, 2
- 热平衡条件, 10
 热容量, 9
 定容, 9
- 软凝聚相, 12
- 瑞利-金斯公式, 54
- 弱简并, 49
- 塞曼能量, 25
- 三临界点, 103
- 熵, 6
 微正则系综的, 29
- 熵判据, 10
 熵增加原理, 6
- 声子, 58
 生成物, 16
- 实际气体, 79
 矢势, 67
- 顺磁性, 94

- 索莫非模型, 61
- 索莫非展开, 64
- 态变量, 1
- 态函数, 2
- 态密度, 28, 47
- 特性函数, 31
- 铁磁—顺磁相变, 94
- 铁磁相, 94
- 铁磁性, 94, 98
- 铁磁序, 98
- 统计分布的涨落, 43
- 统计关联, 49
- 图, 113
- 外斯分子场, 97
- 微观粒子分类, 24
- 微观态数目, 29
 - 理想气体的, 29
- 维恩公式, 54
- 维里定理, 87
- 维里物态方程, 87
- 维里系数, 82
- 位形积分, 79
- 温度, 2
- 稳定条件, 10
- 稳衡态, 2
- 稀薄等离子体, 88
- 系综, 22
 - 不同系综的等价性, 36
 - 巨正则, 31, 33
 - 平均值, 21
 - 微正则, 28
 - 正则, 31, 32
- 系综理论, 19
- 相, 2
- 相变潜热, 13
- 相轨道, 20
- 相空间, 20
 - 能量曲面, 20
 - 体积元与量子态数目的对应关系, 27
- 相平衡条件, 10
 - 单元复相系的, 11
- 相体积元, 24
- 相图, 12
- 相宇, 20n
- 效率
 - 热机的, 3
- 虚变动, 10
- 序参量, 98, 105
- 巡游电子, 61
- 液相, 12
- 一级相变, 14
 - 朗道理论中的, 102
- 易兴模型, 93
- 硬心列纳德—琼斯势, 83
- 元激发, 60
- 涨落, 37
 - 巨正则系综的粒子数, 35
 - 普遍公式, 37
 - 正则系综的能量, 35
- 涨落物态方程, 86
- 振动特征温度, 72
- 正氢, 74
- 正则动量, 67
- 正则方程, 20
- 正则分布, 32

- 正则系综, 31
- 质量作用定律, 78
- 中子散射实验, 59
- 仲氢, 74
- 周期边条件, 110
- 转动量子数, 73
- 转动特征温度, 73
- 转移矩阵, 110
- 状态方程, 2
- 准静态过程, 2
- 准粒子, 60
- 准连续, 27
- 准连续条件, 27
- 紫外灾难, 55
- 子系配分函数, 70
- 自对偶性, 116
- 自发磁化, 94
- 自恰方程, 98
- 自旋, 24
- 自旋—统计定理, 25
- 自旋波函数, 92
- 自旋角动量, 24
- 自旋模型, 91, 93
- 自旋统计定理, 26n
- 自由能, 8
- 最概然分布, 43
- 最可几分布, 43