

统计力学

马上庚

V1.3

2025 年 1 月 27 日

序

本书是供研究所教学和参考之用。统计力学是一门基础科学。近十几年来，这门科学扩充得很快，是个研究热门。本书的目的，不是在把这门学问作通盘介绍，而是把最基本的概念陈述清楚，再用一些较新的题材，来作为应用例子。概念和应用并重。

今天的统计力学，已分科太细，应用太广。读者不可能读一两本教科书而成专家。教本的目的，只是帮忙读者打稳基础，把一些基本概念弄清楚，学一两手简单的“招数”，以免以后自创门路时吃亏。目前市上教本甚多，不乏经典之作，但在观念上，题材上，已略嫌陈旧。新书多偏于专门。时人多急于研究，少致力于教本。再者新知累积太快，写一本长久性的教本，已是绝顶困难的事。此书是一本简短的书，供一时之需。希望对今天的读者有所帮助。

统计力学是理论物理的一部门。它最出色之处，是它的应用。应用范围包括物理、天文、化学、材料、以至于生物学。它是一个非常方便的工具，使我们能把物质的分子结构性质，和大尺度的热力、电磁等性质，连接起来。就难易程度来说，统计力学的应用大致可分成“初等”和“高等”。初等部分大致是“理想气体”，包括量子理想气体。凡是粒子间相互作用不重要的情形，如自由电子模型，都包括了。其应用已非常广泛。高等部分是对相互作用的讨论，应用在相互作用很重要的情况，如相变现象。高等部分自然是困难得多。除了少数特殊情形，近似法和数值计算是仅有的解法。简单、有效的近似法，往往是把问题简化成某一种理想气体模型。“平均场”近似法是一个重要的例子。

在我看来，高等统计力学是在固态物理兴起之后的产物。虽然在四十年前统计力学和相变现象的关系已建立，但大多数的应用、发展，却是近二十年的事，尤其是近十年来，更为蓬勃。大多数的应用，都是应固态物理之用。原因是固态物理现象繁多，实验方法突飞猛进，而所观测的物质，仍不似生物物体之复杂，因此理论的分析还大致可以配合。许多新的概念，都是从固态物理得来。

本书的大部分篇幅,用在例子的讨论.初等的例子用来澄清概念,高等例子用来讨论现象和方法.事实上,基本概念和应用是不能分开的.本书所有的例子,直接或间接,多少都是用来陈述一些基本概念.应用例子的选择,力求简明,所以都是些极其简单的模型.题材比较偏重于和固态物理有关的.我尽量选比较新的、有趣的、能激发读者思考的题材.这本书应该可使读者略窥统计力学的新貌.

我写这本书的动机之一,是藉此发表一点自己的意见.我对一些基本概念的看法,和传统的看法不大一样.我觉得统计力学是一门“头重脚轻”的学问.也就是说,有很多成功的应用,但对基本原理的了解,却是很幼稚.但一般教本,常把统计力学看成一门完美无缺的学问.

大致说来,统计力学是以一些计算法则为中点,而后向上讨论法则的应用,向下讨论它的根源.(这些计算法则可以说是玻尔兹曼的求熵公式,或可以说是某状态发生概率和 $e^{-H/T}$ 成正比, H 是该状态的能量, T 是温度).应用方面的材料,琳琅满目.但追溯根源的材料,相形之下太少,更少有适用于教本者.原因之一是多年来应用的成功,使人对这计算法则有了很强的信心.虽然对它的来源不清楚,只要会用它,也就不管清不清楚了.

还有一个原因是对根源的不了解.也就是说,至今仍不知道如何从力学观点来了解这些计算法则.这些法则至今仍是一假设.许多书都是以计算法则为起点,不问根源,把这门学问看成一门纯计算的学问.这种看法并无可厚非.糟的是,许多教本为了对学生有个交代,搞一些似通非通的花言巧语,企图把这些法则说成一个已成立的“定律”.学生不懂,反正没关系,只要到时会用法则代公式解题目就行.这是我以前的学习经验.

当然,许多成功的应用,并不需要深刻、正确的了解.这是科学发展的一大特色.不过,应用一推广,迟早会出纰漏,进一步的了解就成了急务.今天,统计力学的发展,已到了必须追溯根源的时候.

这统计力学的假设,事实上是一个“独立性”或“不规则”性的假设.如果能了解这假设,则有助于了解所有其他不规则现象,例如计算器的随机数程序,反之亦然.

我以为对科学,或对一切学问,最基本的必要态度是“知之为知之,不知为不知.”所以,这本书很注重统计力学的弱点,计算法则的限度,及可疑的地方.虽然本书对重要的问题没有解答,但尽量把问题说清楚.

从力学,电磁学的经验,学生们总习惯于明确的定律.在此我却会一直提

出各种靠不住的地方。因此会带来一种不安全感。我希望这种不安全感会使读者变得小心一点，养成怀疑的习惯。

本书分七篇，共三十章。这七篇是

- 第一篇 平衡 第一至四章
- 第二篇 假设 第五至九章
- 第三篇 概率 第十至十三章
- 第四篇 应用 第十四至十九章
- 第五篇 变动 第二十至二十二章
- 第六篇 溯源 第二十三至二十六章
- 第七篇 凝结 第二十七至三十章

第一篇可以说是一个“准备活动”，复习一下常用的概念和名词，将理想气体（包括量子气体）的种种温习了。

第二、三、六篇着重在基本概念的讨论。统计力学是建立在一个大胆的假设上，这假设的含义并不是很明确。本书从分子运动的观点来讨论这假设，避免抽象的观念，如“系综”。我的看法是，这系综概念是不必要而且不合事实。概率在此是统计数字，一个工具，不是一个原理。

第五篇讨论一些不平衡的现象，目的是要讨论平衡的维持。暂稳态、数值模拟包括在这篇。

第四、七篇着重在现象的分析，即应用。包括一些较新的题材，如冻结杂质，超流性，二度空间的电荷气体，等等。这些内容不是报道文献中的成果，而可说是把我的一点研习感想，用最直截的方式，写出来供读者参考。我对这些题目的看法，并不一定是最正确的，而且可能偏见颇多。读者善自斟酌。

本书的讨论，不大严密，也就是说，许多结论都没有严密的证明，而只用品简单的例子来示范一下。这并不是说严密的证明不重要，而是它超出了我的知识范围。本书的数学水平不高，用到的数学技巧也不多。可以说完全没有复杂的计算。复杂的计算非常重要，但我以为只有读者自己实地去计算才能学到计算方法。在书中讲一大堆不会有太多好处。

有些十分重要的材料嫌太专门，没有纳入本书，例如临界现象和“重整化群”，（虽然相变现象颇占了一些篇幅）。不平衡现象谈得很少。应用例子偏重于简单的固态物理问题。热力学也讲得很少。抽象的，形式上的东西尽量避免。有些材料，仅见于杂志，而尚未见于其它书本，如第十四章讲的位力系数和碰撞时间的关系，第二十四章讲的重聚现象，二十五章的从轨迹求熵。许多重点、

说法和一般教本不同,如强调观测时间和平衡的关系,暂稳态的重要,及超流态和“绕圈数”之守恒的密切关系,等.

阅读本书,并不需要很深的数学修养.大学量子力学所用的数学就够了.物理方面,读者需要有大学程度的力学、量子力学、热力学、电磁学的常识.本书对读者的要求,主要是谨慎,独立的思考.统计力学在基本物理学中,是最难的一门.本书不是容易念的.习题是内容中重要的一部分.不做习题,则不但应用技术学不到,基本概念也不会学到.

读者须牢记:这本书上说的,或其他书上说的,都不是您自己的见解.自己对统计力学的看法,是要自己去思考,去建立.这本书和其他的书,只是作您的参考,不是要牵着您走.

我写这本书还有一个动机.有鉴于今天中文科学书籍之极度缺乏,这本书虽不足以有很大的作用,但至少有一点点帮助.我深信基础科学,如不用本国文字,一定说不清.教本必须国人自写.近百年来科学人员多习惯于用外文写研究报告.请他用中文写书,就说“我不能”.事实上是“不为”,非不能也.我自己国文程度甚差.写此书,是一大胆尝试.开始觉得生硬,但很快就顺手了.到底科学不是文学.用本国语言写科学书,只要有内容,文字绝无问题.本书如不易读,主要是内容难懂.作者的国文差,是次要原因.

这本书的大部分,是在1977年及1981年,我在(台湾)清华大学执教时所写.许多位教授和同学,给了我指示,帮助及鼓励,特别是李怡严、单越、阎爱德、倪维斗、石育民、黄幸美、郎棣、颜晃彻、陈通等教授,及林其隆、杨芳铿、陈俊良、林达观等同学.在此向他们致谢.清华物理研究所的行政、文书同仁也给了我不少协助,特别是列位所长及张月梅小姐,我也一并致谢.

我更要谢谢我父、母亲.对我的工作,他们一向尽心尽力支持.这本书得以顺利完成,也靠他们的鼓励和支持.我对他们,却少有孝敬.今以此书献,以略慰慈心.

本书内容,待指正之处甚多.希望科学界人士不吝赐教.

1982年春,马上庚谨识

注意事项

1. 教学和阅读

本书提供一学年(九个月课程)的教材. 主要对象是物理、材料科学、化学的研究生.

大致上, 前三篇较浅易, 后四篇较深, 每一章都是由浅而深. 教材的取舍, 当然是由教师斟酌学生的程度和需要, 及教师自己的意见来决定. 以下是几点建议, 并指出一些教学、阅读该注意的事项.

(甲) 请特别注意讨论, 分析不同的看法, 不要用教条式的讲法. 头三篇的材料, 最好不削弱. 第六篇是总论性的, 亦不应略去.

(乙) 数学式子的推导, 不可忽略. 本书避免了任何复杂的计算, 目的是使学生可以把书中所有的式子推导做一遍. 推导的主要功效是了解各种近似, 假设的意义和效果.

(丙) 每一章后, 有讨论问题. 这些问题包括许多常见的练习题. 但也有很多供讨论用的, 没有“标准答案”, 我自己也不知道如何回答的问题. 此类问题, 不应忽略.

(丁) 本书所列参考资料, 并非必要教材, 而是供进一步研习参考之用, 或仅指明图表的来处. 我以为, 研究所教学, 仍应以训练学生自己的思考为主. 先自己想, 再去翻书刊, 以查出某一项特别资料. 自己概念越清楚, 参考数据越有用. 当然, 多看参考书, 仍是有用的. 见多则识广. 可惜的是, 今天大部分科学文献是外文书刊, 不易吸收. (不过, 如果基本概念清楚, 常识够, 外文科学书刊并不难读. 概念不清, 外文再好也没用). 因此, 教师的任务, 也要包括介绍一点其他资料, 以充实教材. 统计力学的中文教本和参考书籍, 为数亦不少. 教师和学生应多涉猎.

(戊) 学习时遭遇困难的原因, 大多是初级的东西没学好, 或忘了, 不是新资料太难, 也不是缺乏参考书. 常常复习初等的材料, 是十分重要的. “温

故而知新”，真是再正确不过。

2. 名词

今天的科学，有一特点，就是名词和术语的泛滥。（西方的拼音文字，很适合于大量制造新字，只可惜除了几个专家之外，没人看得懂，过一段时间，字又被改了，或忘了）。因此有人误以为学习科学，就是认些名词，说几句术语。这种误解，中外都常见。

对一门基础科学而言，名词术语并不重要。教本的目的，是把一套基本概念，分析方法说清楚。重点在解释，并不需要很多术语或名词，更不需要外国文字。

名词、术语不是科学的本身，纯是一种习惯，是一种表达工具。越少，越简单，越好，但不能没有。初等科学的名词术语，早已通用于中学教本及常识书刊，本书亦沿用之。至于高等科学，由于中文书刊少，还没有建立一套基本的名词术语习惯。本书大致沿用教育部审定者，（见“科学大辞典”，自然科学文化事业公司，1981年版），即照英文名词习惯直译者。但英文名词用法不合理之处甚多，习惯亦多有不合用于中文的。因此，本书自定义了少数名词。例如，本书不用“哈密顿函数”，而用“总能函数”，或“能量函数”。我以为，像能量这样常用的观念，不该沿用英文人名的用法。还有几个：不用“化学势”，用“人口压”。不用“玻色粒子”，“费米粒子”，而用“合羣粒子”，“不合羣粒子”*，等。阅读本书时，自然会了解这些名词的用意。既然中文科学名词还没有一套习惯用法，本书的用法，可算是一点建议。希望读者勿在名词术语上花功夫。教本的内容，在讨论和解释，不在名词和术语。更切忌把中文名词翻成英文，发现见过这英文字，就自以为懂了。认得英文名词和了解名词的意义，完全是两回事。

3. 单位和常数表

为求讨论的方便，数学式子的简短，本书把热常数 k （即玻尔兹曼常数）设为 1，常见的 kT 只写成 T 。温度 T 看成一能量，能量亦可用温度单位度量。

*为方便阅读，本简体本没有坚持原作者的用法，而改常用名词。

量子常数 \hbar (即普朗克常数除以 2π), 亦设为 1. 不过在有些常数的角色不明显时, 还是写出来. 以下单位和常数表可以作换算之用.

热常数 k = 每 1 K 的能量, 即

$$\begin{aligned} k \times 1 \text{ K} &= 1.38 \times 10^{-16} \text{ erg} \\ &= 8.62 \times 10^{-5} \text{ eV}. \end{aligned}$$

克分子数 N = 每克分子量物质中的分子数
 $= 6.02 \times 10^{23}$.

气体常数 $R = kN = 1.98$ 卡/克分子.

注意: $1 \text{ eV} \approx 10^4 \text{ K}$,

1卡/克分子 $\approx 0.5 \text{ K/分子}$.

量子常数 $\hbar = 1.05 \times 10^{-27} \text{ erg} \cdot \text{sec}$.

$$h = 2\pi\hbar.$$

$$1 \text{ eV}/h = 2.42 \times 10^{14} \text{ sec}^{-1}.$$

质子质量 = 1.67×10^{-24} 克.

电子质量 = $m = 9.11 \times 10^{-28}$ 克.

电子电荷 = $-e$,

$$e^2 = (4.80 \times 10^{-10})^2 \text{ erg} \cdot \text{cm}.$$

氢原子半径 (波耳半径)

$$= a = 0.529 \text{ \AA}.$$

氢原子电子束缚能量 = $e^2/2a$

$$= 13.6 \text{ eV}.$$

波耳磁矩 = $\mu_B = e\hbar/2mc = 9.27 \times 10^{-21} \text{ erg/G}$

$$\text{G} = 1 \text{ 高斯}.$$

4. 常用数学数据表

以下列出几个常用公式与数字, 这并不能取代数学手册, 但在手边没有手册时, 可能会有帮助.

(甲) δ 函数

$$\begin{aligned}\theta(x) &= 1, & x > 0, \\ &= \frac{1}{2}, & x = 0, \\ &= 0, & x < 0,\end{aligned}$$

$$\frac{d}{dx}\theta(x) = \delta(x)$$

$$\int_a^b \delta(x) dx = \theta(b) - \theta(a)$$

$$\delta(ax) = \frac{\delta(x)}{|a|}$$

$$\operatorname{sgn}x = \frac{x}{|x|}$$

$$\frac{d}{dx}\operatorname{sgn}x = 2\delta(x)$$

$$\delta(x-a) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik(x-a)}$$

$$\sum_{n=-\infty}^{\infty} \delta(x-a-nb) = \frac{1}{b} \sum_{m=-\infty}^{\infty} e^{i2\pi m(x-a)/b}$$

(乙) 常用数据和公式

$$\Gamma(n) = (n-1)!$$

$$\int_0^{\infty} dx x^n e^{-x} = \Gamma(n+1) = n!$$

$$\Gamma(1+n) = n\Gamma(n)$$

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} = 1.772$$

$$\Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\sqrt{\pi}$$

$$\ln\Gamma(1+x) = -Cx + O(x^2)$$

$$C = 0.5772 = -\int_0^{\infty} dx e^{-x} \ln x$$

$$N! = N^N e^{-N} (2\pi N)^{1/2} \left[1 + \frac{1}{12N} + \frac{1}{288N^2} + \cdots \right]$$

$$\int_0^{\infty} \frac{dx}{e^x + 1} = \ln 2$$

$$\int_0^{\infty} \frac{x dx}{e^x + 1} = \frac{\pi^2}{12}$$

$$\int_0^{\infty} \frac{x^2 dx}{e^x + 1} = \frac{3}{2} \zeta(3), \quad \zeta(3) = 1.202$$

$$\int_0^{\infty} \frac{x dx}{e^x - 1} = \frac{\pi^2}{6}$$

$$\int_0^{\infty} \frac{x^2 dx}{e^x - 1} = 2\zeta(3)$$

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

$$\int_{-\infty}^{\infty} dx e^{-x^2/(2a^2)} e^{-ikx} = \sqrt{2\pi a^2} e^{-1/2k^2 a^2}$$

$$\tanh^{-1} x = \frac{1}{2} \ln \frac{1+x}{1-x} = x + \frac{x^3}{3} + \cdots, \quad |x| < 1$$

$$\coth^{-1} x = \frac{1}{x} + \frac{1}{3x^3} + \cdots, \quad |x| > 1$$

(丙) 方便数字

$$\ln 2 = 0.693$$

$$\ln 10 = 2.30$$

$$\sqrt{\pi} = 1.77$$

$$6! = 720, \quad 9! = 362880, \quad 12! \approx 4.8 \times 10^8$$

$$2^{10} = 1024, \quad 2^{15} = 32768, \quad 2^{20} \approx 10^6$$

目录

1. 教学和阅读	7
2. 名词	8
3. 单位和常数表	8
4. 常用数学数据表	9
第一篇 平衡	21
第一章 平衡	25
1. 平衡和观测时间	25
2. 平衡和分子运动	26
3. 大尺度改变和局部平衡	27
4. 物质的结构	28
5. 补充、结论	29
第二章 热力学的基本概念	33
1. 平衡态和可逆过程	33
2. 功和热	33
3. 绝热过程和熵	35
4. 绝对温度和熵	40
5. 理想气体	42
6. 热力学第二定律	46
7. 热机的效率和速率	47
第三章 细致平衡定律	53
1. 状态的能量分布	54
2. 费米粒子的分布	57

3.	守恒定律, 化学势和熵的常数项	60
4.	波色粒子的分布, 热辐射	62
5.	化学反应, 浓度定律	64
第四章	金属电子	69
1.	高密度理想气体	69
2.	热容率	72
3.	周期性结构的影响	75
4.	电子引起的结构改变	79
第二篇	假设	85
第五章	分子运动幅度和熵, 基本假设	89
1.	运动幅度	89
2.	计算熵的假设	90
3.	费米粒子	92
4.	理想气体	94
5.	高维空间的一些特征	97
6.	尖峰积分和大数定律	99
7.	能量均分定律	100
8.	活动范围面面观	101
9.	轨迹和活动范围	104
第六章	基本假设的初步应用	109
1.	部分和总体 分类和综合	109
2.	振动	114
3.	德拜模型	116
4.	声子和第二声波	121
5.	一般结论, 平均值和正态分布	128
第七章	计算法则	135
1.	各种热位能	135
2.	各式计算法则	138

目录	15
第八章 示范例子	149
1. 时间尺度和活动范围	149
2. 氢分子气体	153
3. 一维空间模型	158
第九章 杨、李定理	167
1. 宏观极限 (即“热力学极限”)	168
2. 定理的推广	172
第三篇 概率	175
第十章 概率和统计	179
1. 概率的应用和滥用	179
2. 分布和平均值的统计	181
3. 概率的一般定义	188
第十一章 独立和纷乱	193
1. 独立现象、定义和后果	193
2. 独立的检验	197
3. 不规则数列	199
4. 散射实验和相干函数的测定	201
第十二章 许多独立变量之和	209
1. 正态分布	210
2. 中心极限定理, 简单说法	211
3. 高次平均值和连接值	214
4. 连接值展开定理	217
5. 中心极限定理, 广义说法	218
6. 重复试验和概率的测定	220
7. 不规则运动, 扩散	222
8. 宏观变数的起伏	224
9. 起伏和基本假设的延伸	226

第十三章 相干函数	233
1. 响应和起伏	233
2. 密度相干函数	236
3. 费米粒子气体	241
4. 量子力学中的响应和相干函数	243
第四篇 应用	249
第十四章 理想气体定律的修正	253
1. 分子间的作用力	253
2. 理想气体定律的修正	255
3. 碰撞的逗留时间	256
4. 量子力学结果	260
5. 束缚态, 勒文生定理	262
第十五章 多相共存	267
1. 气相和液相	267
2. 水珠的消长	271
3. 潜热和分子作用能	273
4. 熔解, 林德曼公式	274
5. μ_1, μ_2 的定义	275
第十六章 磁性	281
1. 顺磁性	281
2. 磁化率	285
3. 带电粒子的反磁性	286
4. 自旋的相互作用	293
第十七章 易形模型	299
1. 易形铁磁, 一维空间	299
2. 二维空间, 易形铁磁性的证明	302
3. 其它易形模型	307

目录	17
第十八章 杂质和运动	311
1. 溶液和渗透压	311
2. 有效作用能	312
3. 低密度情形	314
4. 走动杂质和固定杂质	315
5. 非晶态	320
第十九章 静电作用	325
1. 短距离和长距离都重要	325
2. 离子气体和溶液	326
3. 金属电子	328
4. 电子晶体	330
5. 二维空间的电荷气体模型	331
第五篇 变动	337
第二十章 气体运动方程式	341
1. 流动和碰撞	341
2. 没有碰撞的情形, 离子振动	343
3. 零点声波	345
4. 碰撞和扩散	347
5. 碰撞和声波, 粘性和热传导	350
6. H -定理	354
第二十一章 扩散方程式	357
1. 简单的应用例子	357
2. 暂稳态	361
3. 转换成波动方程式	363
4. 扩散方程式的推导	364
5. 二态小集团模型	368
第二十二章 数值模拟	375
1. 分子运动数值解	376

2. 随机数列	377
3. 随机数模拟法	378
4. 该注意的概念问题	380
第六篇 溯源	385
第二十三章 热力学定律	389
1. 绝热过程	389
2. 绝热过程和熵	392
3. 热力学第二定律	393
4. 热力学第三定律	396
5. 非晶态和第三定律	398
第二十四章 重聚现象	403
1. 粘液瓶示范	403
2. 示范的分析	405
3. 自旋重聚实验	407
4. 离子气体电荷重聚	409
第二十五章 从运动轨迹求熵	413
1. 巧合次数和区域大小	414
2. 物体各部的独立性	416
3. 相干时间	418
4. 计算过程	418
5. 暂稳态的熵	421
6. 热力学第三定律	421
第二十六章 基本假设的来源	427
1. 基本假设	427
2. 遍历性和系综	429
3. 混合性和独立性	431
4. 概率和实验	434
5. 轨迹的不稳性	436

目录	19
6. 一般独立性问题	439
7. 应用方面的困难	440
第七篇 凝结	443
第二十七章 平均场解	447
1. 平均场	447
2. 总磁矩	449
3. 热位能和多相共存	452
4. 易形格子气体	454
5. 范德瓦尔斯公式	457
6. 凝结现象的共同性质	461
第二十八章 分界面的起伏	467
1. 液面模型	468
2. 易形模型分界面	471
3. 晶体表面, 电荷气体模型	474
第二十九章 连续对称模型	481
1. 平面向量模型	481
2. 晶体的密度起伏	483
3. 量子向量模型	486
4. 连续对称和柔软波	489
5. 凝结的缺陷	491
第三十章 超流性原理	501
1. 量子格子气体	502
2. 基态和低温模型	504
3. 流动状态, 绕圈数	507
4. 超流的稳定性	509

第一篇

平衡

本篇分四章, 略述最基本的概念, 并讨论些简单例子. 第一、二章复习平衡的概念及热力学的定律. 特别强调平衡态非瞬时形态以及观测时间的重要性. 热力学的最突出的特点是“熵”的出现. 第三章讨论细致平衡定律, 强调分子间作用为平衡之本. 第四章讨论金属电子, 复习一些基本固态物理常识, 及统计力学的初步计算方法.

这一篇可以说是一个准备工作, 各种基本概念, 将在以后各篇重新分析.

第一章 平衡

本章讨论平衡这个最基本的概念，特别强调平衡和观测时间的关系。从分子运动的时间尺度来看，平衡是一长时间观测下的状况，而非一瞬时形象。讨论平衡，必须考虑观测时间的大致长度。

1. 平衡和观测时间

基本概念的来源，总是一些简单的现象。平衡是热力学和统计力学最基本的概念。现在用几个简单的例子，来说明平衡的意义。并且谈一谈它所牵涉到的一些问题。

平衡是指一种不变动和静止的状态。例如一张古画挂在墙上，即是一种不变动的状态。不过这种不变动的状态只是一种近似，而不是绝对。这张古画从它的产生到现在，已经过了不少的变动。例如，由于和空气的化学作用，有些颜色已经不是以前早先的颜色了，纸也变黄了。所以讲到平衡，我们必须牵涉到观测时间的问题。如果观测时间很短，譬如说“一秒钟”、“半秒钟”，或者“两三天”，或者“一两年”，这张画都可以看成是在一种平衡状态，因为在上述的观测时间内，它的改变是相当微不足道的。不过假如现在用一段很长的时间做观测时间的话，譬如一百年、两百年，那么这张画便不能被看成是在平衡状态了。因为它的改变已经是很显著，不能被忽略了，且不提挂画的钉子或墙壁的早已破损。多少是微不足道？多少是不能忽略？这当然是要看分析的目的来决定了。一般说来，讨论平衡，必须在一定观测时间范围内。这范围的规定，要看观测者的需要和兴趣。

现在看另外一个例子。从一个热水瓶里面倒一些开水在一个小茶杯内，几秒钟之后杯中的水便成为一种静止不变的状态。这状态就是平衡状态。杯中水的容积，可以从水的高度及杯子的口径量出。水的温度则可以用温度计量出。在一两秒钟之内，这些测量值大致都是不变的，也就是说，观测时间必

须小于一两秒钟. 如果时间太长, 则温度的改变已经显著, 水就不能被看成是在一平衡状态. 反之, 一个钟头之后, 水的温度已经和室温一样了. 如果室温在数小时内不变, 水的温度也因此不会改变, 所以这时的观测时间只要是在数小时范围之内, 这杯水都可以算成是在平衡状态了. 不过水分子因不断的从水面上蒸发出去, 虽然在数小时之内, 水的容积不会改变很多, 但如果时间太长, 比如两三天之后, 水就慢慢的干掉了, 所以观测时间长过两三天, 这杯水就不能算是在平衡状态中. 过了好几天后, 水完全干了, 这个干杯子又可以算做是在一种平衡状态中. 这个空杯子并不是在一种绝对不变的状态下, 因为杯子的分子也是在不断的蒸发中, 虽然在几年之内并没有显著的改变.

这些例子指出, 观测的时间不能太长, 否则平衡就没有意义. 世界上没有绝对不变的东西, 不过我们可以用各种不同的器具使平衡的时间范围延长. 譬如我们可以把古画放在镜框里, 把水放在热水瓶里. 不过绝对的封闭是不可能的, 总会有一些能量流入流出. 注意: 没有一种容器是可以永远不坏的, 任何在张力或剪切力下的物质, 都是在一种暂稳的状态之中, 而不是绝对的稳定. 只要观测时间不是很长, 物体没有什么明显变化, 这种状态, 我们就叫它做平衡, 这样的定义是相当模糊的, 比较明确的定义和进一步的了解, 要等我们分析过很多例子之后.

2. 平衡和分子运动

平衡这个现象, 是从一个大尺度的观点来看的一个现象. 大尺度就是比分子大小大很多的尺度, 也就是所谓“宏观”尺度. 以上讨论的时间也是一种宏观的时间. 也就是说, 比分子运动的时间尺度大很多的时间. 从“微观”的观点看, 也就是说从一个分子大小尺度的观点来看, 那张古画是由许许多多分子所构成的, 这些分子不断地振动, 前述那一杯水中的水分子也是不断做非常复杂的运动, 没有一个分子是静止不动的. 大尺度的现象和性质, 是这些非常复杂的分子的运动, 在观测时间内的平均. 从小尺度的观点来看, 平衡是一个非常特殊的运动状况. 这个运动状况的特点是: 在某观测时间范围内, 大尺度的性质不变. 统计力学的目的, 就是要了解分子运动和平衡的关系.

平衡既然是静止不变的状态, 则所有因时间而变的现象, 都可以称为是不平衡的现象. 热、电的传导, 声音的传播, 以至于刮风、下雨都是例子. 不平衡现象的了解, 也是须要从分子运动来着手. 要了解平衡必须分析维持平衡

的种种因素,也须要讨论产生变化的种种因素.所以,研究平衡,免不了要牵涉到不平衡的问题.

以上我们强调平衡和观测时间的密切关系,主要原因是使读者对统计力学的应用范围有较正确的印象.一般教科书多从“绝对平衡”的概念着手.所谓“绝对平衡”是理想化的,它是指绝对封闭的物系,在无限长的时间观测下的状态.或也有用更抽象的概念作为平衡的定义.这样的做法,并无可厚非,且开始时,使讨论比较简明,但是读者的思想会比较受限制.许多统计力学可以分析的问题,在这理想化的定义下,都成了“不平衡”的问题.我们从头开始就把时间放在平衡的定义里,这看来像是增加麻烦,但以后读者在应用统计力学时,会灵活得多.

3. 大尺度改变和局部平衡

以上我们强调,平衡是一大尺度的现象.从分子尺度看来,它不是一个瞬时性质,而是一个长时间观测下的性质.从大尺度看来,观测时间不能太长,否则难以维持不变.现在我们用流体力学的公式为例,来更进一步地强调大小尺度的重要性.

大尺度的气体的运动,多可以用以下方程式解释:

$$\begin{aligned}\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) &= 0, \\ \rho \frac{\partial \mathbf{v}}{\partial t} + \rho (\mathbf{v} \cdot \nabla) \mathbf{v} &= -\nabla p.\end{aligned}\quad (1)$$

式中 $\rho(\mathbf{r}, t)$ 是在位置 \mathbf{r} , 时刻 t 的质量密度, $\mathbf{v}(\mathbf{r}, t)$ 是气流速度, $p(\mathbf{r}, t)$ 是气压, $\rho \mathbf{v}$ 是气流量, 也是动量密度. 第一式指出质量守恒, 第二式表明牛顿定律, 即动量改变等于所受的力. 我们把气体的粘性略去了, 以求简明.

[(1) 的意义, 用积分式看得更清楚

$$\begin{aligned}\frac{\partial}{\partial t} \int \rho d^3 \mathbf{r} + \int d\sigma \cdot \mathbf{v} \rho &= 0, \\ \frac{\partial}{\partial t} \int \rho \mathbf{v} d^3 \mathbf{r} + \int (d\sigma \cdot \mathbf{v}) \rho \mathbf{v} &= - \int d\sigma p.\end{aligned}$$

$d^3 \mathbf{r}$ 的积分包括某区域内的点, $d\sigma$ 积分包括该区的表面.]

注意 (1) 中的时间, 空间尺度都是宏观尺度, 微分量 dt 是看成一个相当长的宏观的观测时间,

$$dt \gg \tau. \quad (2)$$

τ 是分子运动时间尺度, 在此大约是“平均自由时间”即一个分子在两次碰撞间的自由飞行时间. 空间的微分 $d^3\mathbf{r} = dx dy dz$ 是很大的空间, 比分子运动空间尺度大很多,

$$dx \gg \lambda. \quad (3)$$

λ 是“平均自由距离”, $\lambda \sim \tau v$, v 是分子平均速度. 在观测时间 dt 内, 容积 $d^3\mathbf{r}$ 内的气体大致是一平衡态, 是为“局部平衡”. 不过如时间比 dt 长很多, 空间比 $d^3\mathbf{r}$ 大很多, 则有变化而不平衡. 这变化可由 (1) 描述.

(2) 和 (3) 指出了 (1) 的应用范围, p 、 v 、 ρ 各量的改变必须相当平缓, τ 和 λ 的大小要看情形.

有一点该特别注意: (2) 和 (3) 并不是很明确, dt , dx 都是越大越好, 其最低限度是什么? 10τ , 10λ 还是 1000τ , 1000λ ? 原则上这些问题可以用微观的分析来回答. 也就是说, 原则上可以从分子运动观点导出大尺度运动方程式, 但事实上这些问题是极其困难的问题, 以后我们会讨论这些特殊情形, 以指出这困难的程度.

4. 物质的结构

在不同的环境下, 物质会有不同的平衡性质. 所谓环境是指温度、压强、磁场, 等因素. 这些环境因素在某些限度内, 可以在实验室内制造和控制. 所谓平衡性质, 是指物体的结构以及热力、电磁、弹性, 等性质. 统计力学面临的问题是如何从分子的运动来分析、了解这些性质. 现在简单地复习一下物质结构的一些特征. 一般物质都可以看做是一大堆核子、电子的合成物. 不同的合成、运动方式, 造成不同的结构和性质. 如果一个或数个核子吸住几个电子而结合成一个个体 (即原子或分子), 且各个个体间不结合, 各自运动, 则成流体. 如果核子不能东跑西跑, 则成固体. 如果作不规则排列, 则成“非晶体”或“玻璃体”, 如果作整齐排列, 则为“晶体”. 固体中若有自由来往的电子, 则为导体, 否则为“绝缘体”. 低温的氦液 (近于绝对零度) 是“超流体”, 一点粘性也没有. 许多金属、合金, 在低温时是“超导体”, 一点电阻也没有.

每个电子都具有自旋, 形成磁偶极矩, 或“磁针”. 在磁场内, 众偶极矩顺着磁场排列, 是为“顺磁性”. 如果这种排列不需要外加磁场, 则是“铁磁性”. 如果磁针的排列是规则的一上一下, 则为“反铁磁性”. 如果磁针排列是混乱

的, 则为“非晶磁性”或“玻璃磁性”。

以上只是略提几个不同的结构特点. 不同的分子, 不同的晶体, 不同的胶体, 何止几千万种. 今天材料科学兴盛, 人工合成的物质更是与日俱增.*

物质的结构, 和其他性质随温度、压强和其它环境因素而改变. 这些改变可说是种类繁多. 现在把一些较熟知的例子列在表 1. 从冰点 273 K(即绝对温度 273 度) 向上看: 水在 373 K, 一大气压下, 沸腾汽化, 铁的磁性在 1043 K 消失, 铁在 1808 K 熔为液态, 钨在 3653 K 熔解. 从冰点往下看, 氧气在 90 K 液化, 二氧化碳在 78.5 K 由气体结合成固体(干冰), 铅在 7.19 K 变成超导体, 氦在 2.17 K 变成超流体. 以上例子是一些随温度升降的突然的改变, 即所谓“相变”. 突变发生时的温度称为“相变温度”. 不在相变温度时, 小幅度的温度变更, 只会引起小幅度的物性改变.

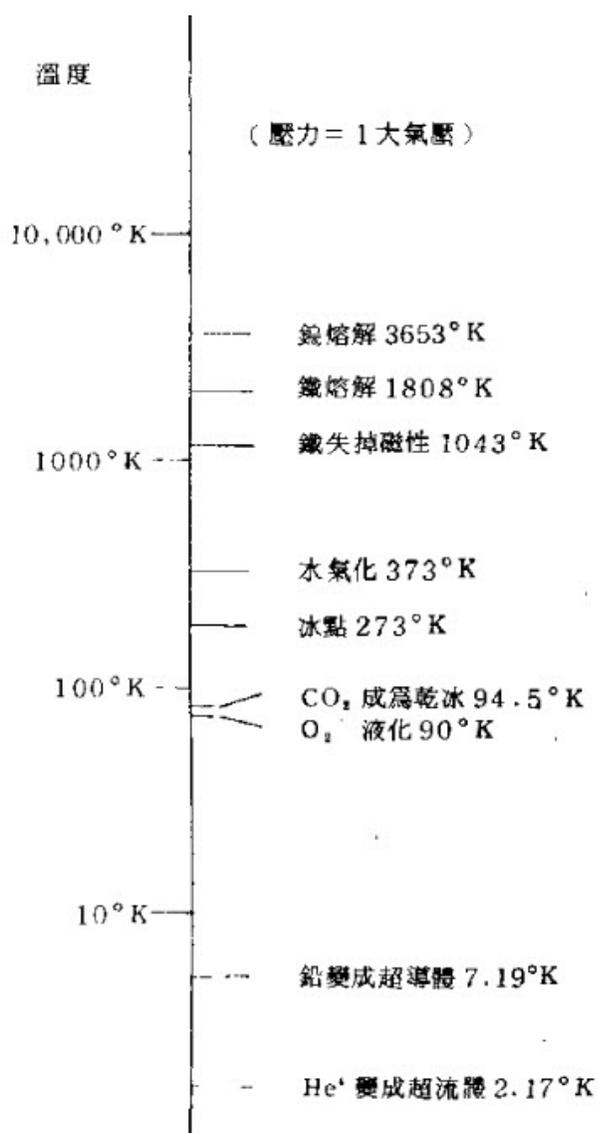
5. 补充、结论

现在总结一下, 并作几点补充:

- ① 从分子运动观点来看, “平衡”不是指某一时刻各分子的位置和速度, 而是指一段长时间的运动状况. 平衡的观测不是瞬时的, 而是长时间的.
- ② 这观测时间的长度, 是由观测者决定, 平衡性质是指各种量在这时间内的平均值. 例如晶体结构是由诸核子的位置所确立. 这结构是一平衡性质. 此地所指的“位置”, 是一核子位置在观测时间内的平均位置. 每个核子都在振动, 其瞬时位置一直在改变.
- ③ 平衡的特征是: 这些平均值不随观测时间的长度而异. 例如, 观测时间加倍或减半, 各平均值不变. 但是, 如果观测时间改变太多, 超出了某一范围, 则这些平均值就变了. 这物体的状态就是不平衡了. 这范围愈小平衡性质愈模糊.
- ④ 我们把所有的这些平均值都看作平衡性质, 不仅是几个大尺度的性质. 例如, 固体中每一分子的平均位置都是一平衡性质. 晶体中固定的杂质、裂缝都是平衡性质.

*见材料科学书籍, 如唐江涛 (1971).

表 1: 物性随温度改变的一些例子



- ⑤ 有些平衡性质和分子运动的关系不大明显. 例如温度, 和以后要仔细讨论的“熵”.
- ⑥ 以上对平衡的讨论当然不完全, 更明确的定义和更深的了解需要谈到许多别的特性, 例如“细致平衡”, 及物体各部运动的“独立性”等. 再者, 所谓大小尺度的分野, 有时不一定很清楚. 也就是说“宏观”和“微观”之间, 有时并没有明确的界限.

总而言之, 我们需要分析的, 不总是一些清清楚楚的题目, 而常是一些复杂而含糊的现象.

讨论问题一

1, 炭有两种常见的晶体结构, 即石墨和钻石. 在日常时间尺度下, 两者都有明确的平衡性质.

问: 在无限长的时间尺度下, 那个比较稳定? (答: 石墨比钻石稳定. 读者可参考固态物理数据, 找出两者的能量.)

2, 硫的晶体有许多种, 它的液体、气体的构造都十分有趣. 读者去找本化学书或有关硫的书去看看.

3, 图 1 示一杯子, 封在一低温盒子里, 杯中装了液态氦, 如果温度降低到 2.18 K 以下, 杯中的液氦就会流到杯外, 直到杯内外液面高度相同为止. 请阅参考书, 了解这现象.

4, 如果温度比 2.18 K 高些, 杯中的液氦是不是有办法跑出来?

5, 杯子里是水, 温度是室温, 情形又如何?



图 1

提示: 4、5 和 3 的主要不同处在时间尺度.

第二章 热力学的基本概念

热力学的讨论对象, 主要是能量的交换 (功和热) 及平衡性质. 它不考虑物质的结构, 只分析宏观的变化. 本章只是对基本的假设和概念作一复习, 以作为讨论统计力学的准备.

1. 平衡态和可逆过程

热力学的基本假设是: 在一定环境之下, 物体有一定的平衡性质. 环境由温度、压强、磁场等外力所确定, 一旦这些因素定了, 平衡形态也定了. 这假设在很多情况下, 颇与经验符合, 但例外情形也不少.

这假设有一个重要的结果, 即“可逆”过程的存在, 也就是说, 把环境因素改了, 再改回来, 则物体的形态也会回到原来的形态. 如果不回到原来形态, 则同一环境下就可以有不同的形态了. 例如, 把一瓶室温的水加到高温, 再冷回室温, 这瓶水又回到原来形态, 所以这些过程是可逆的. 但把一块铁加热到很高的温度, 再突然冷却下来, 这块铁的许多性质会改变. 不同的加热冷却速度可以导致不同的结构. 所以这些过程不是可逆的. 要了解这些过程和铁的性质, 必须要讨论详细结构的改变.

我们现在只讨论可逆过程. 一般说来, 平缓的改变过程多是可逆的, 而突然剧烈的改变常为不可逆. 当然, 平缓和突然要看是照什么时间标准. 大致说来, 改变过程经过的时间, 必须比分子运动的时间尺度要长很多, 这种过程才是平缓.

2. 功和热

一个物体的能量, 可由外界因素来改变, 如把物体拉长, 或压缩, 或加以磁场等. 这些外力都可以对这物体做功, 而增加物体的能量. “力” 乘上 “位

移”, (即坐标的改变), 即得所作之功. 例如, 以拉力 f 把一铁棒子拉长 dx , 便对它做功 dW ,

$$dW = f dx. \quad (1)$$

符号 d 上多加一横, 表示 dW 并非一函数 “ W ” 之微分, 而只是表示一小量而已. 如果在磁场 h 中, 总磁矩增加 dM , 则物体能量增加为

$$dW = h dM. \quad (2)$$

如果固定棒子一端, 而加转矩 η 于另一端, 扭转 $d\theta$ 角, 则能量增加

$$dW = \eta d\theta. \quad (3)$$

再举两个力和坐标的例子: 压强 p 是改变容积 V 的力, 它所做的功是

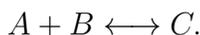
$$dW = -p dV. \quad (4)$$

式中 $-dV$ 是被压缩了的容积. 另外, 如 “化学势” μ 是改变粒子数 N 的力. 例如, 一碗水中溶有氧气, 其分子数为 N . 水面上空气的氧分子, 对水面下的氧分子来说, 是具有化学势的环境, 如果压进去 dN 个分子, 则 μ 做的功为

$$dW = \mu dN. \quad (5)$$

同样, 空气中的氮气, 水蒸气也对这碗水有化学势. μ' , μ'' 分别是能改变水中氮分子数 N' 和水分子数 N'' 的力, 每一种不同的粒子都具有化学势.

化学势大概不是个熟知的概念, 因此我们补几句. 化学势是增加粒子数的力量, 在讨论化学反应时, 十分重要. 各种分子的化学势决定反应进行的方向, 例如分子 A , B 合成 C , 或 C 分解为 A , B 的反应



分子 C 的存在, 对 A , B 粒子数加压, 反之, A , B 对 C 的粒子数加压, 在平衡时, 两边化学势抵消

$$\mu_A + \mu_B = \mu_C.$$

以后讲化学反应时, 会再讲这些.

此地化学势和社会科学中讲的人口压有相同意义, 早婚、多子、长寿都会增加人口压. 增加就业机会, 改善生活环境会减少人口压. 都市兴起, 就业机

会多, 人口压落到乡村之下, 于是人口涌向都市, 但人一多, 环境开始恶化, 职业减少, 都市人口压上升, 于是人口的流动就会慢下来。

长度 X , 磁矩 M , 扭转角 θ , 粒子数 N, N', N'' 都是计量物体的形象、状态的坐标。它们都可称作“总体坐标”。而拉力 f , 磁力 h , 转矩 η , 化学势 μ, μ', μ'' , 都是外力, 是改变坐标而做功的力量。

这些总体坐标是大尺度坐标, 或称为“热力学坐标”。至于计量详细分子结构的小尺度坐标, 则不在热力学内出现。当然, 小尺度的分子运动不但存在, 而且十分重要。能量可以从物体内流出或由物体外流入, 而以上各总体坐标可以不变。这就是所谓热流, 是小尺度运动的结果。小尺度运动是绝对不能忽略的。但我们并不需要把每个分子的坐标拿出来分析。热力学最突出的特点是: 它另加了一个新的总体坐标来总计这些小尺度运动的程度。这新的总体坐标就是“熵”, 我们用符号 S 代表, 增加这新坐标的“力”就是绝对温度, 它所加的能量是

$$dQ = TdS. \quad (6)$$

为什么“ T ”刚好是绝对温度? 这可以说是巧合, 也可以说是一种选择。现在来仔细讨论这个问题, 并复习一些名词和定义。

我们称 (1) 至 (5) 这些能量的改变的总和为“功能” dW , 是因为大尺度坐标改变而来, 是外力对物体所作之功。称 (6) 为“热能” dQ , 是因小尺度分子运动的改变而来, 是由外界流入的能量。总能量的改变为二者之和。

$$dE = dW + dQ. \quad (7)$$

这是热力学第一定律。这定律陈述能量守恒定律, 指出功和热两种不同的改变能量的方式。

—

3. 绝热过程和熵

如果 $dQ = 0$, 即在改变过程中没有热流出入, 则此过程称为“绝热过程”。绝对的绝热不大可能。照经验, 有些材料比较能绝热, 如石头、棉被。有些很差, 如金属。如果改变过程快, 则热量出入少。时间越长, 越不易绝热。以绝热过程为出发点, 我们可以一步步订出“熵”这个概念。为求简明, 我们先考虑某物体的一个坐标 X 。图 1 中的横轴是 X , 纵轴为该物体的总能量 E 。(X ,

E) 平面上每一点代表一平衡态, 从平衡态 A , 经由绝热过程, 可以把形态改变到 B , 或用其他绝热过程, 亦可达到同一条曲线上任一点. 同样, 我们可以从 A' 开始, 用绝热过程达到 B' , 或曲线 $A'B'$ 上的诸点. 从一条曲线到另一条曲线, 则必须加热或放热.

图 1 上的各曲线, 都是这样由绝热过程画出. 这群曲线, 可以用一函数 $g(X, E)$ 的等值线表示:

$$g(X, E) = \sigma. \quad (8)$$

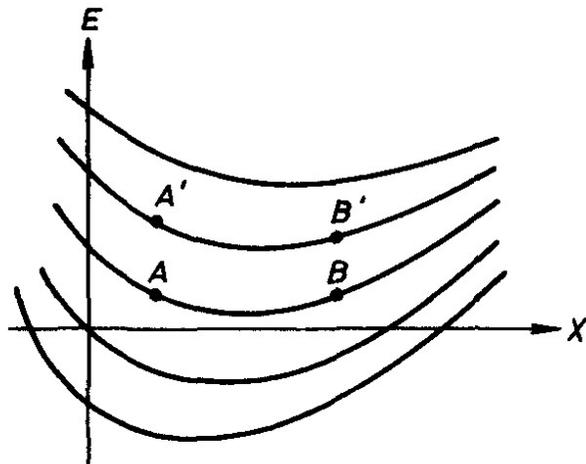


图 1

A

B

A'

B'

E

每一条曲线有一个常数 σ . 这情形和地图上的等高线一样, 同一高度的点, 可连成一条曲线. 反之, 平面上任一群曲线, 都可以想做某地势图的等高线. $g(X, E)$ 为 (X, E) 地点的“高度”. 当然, 我们可以用其他函数来表示这群曲线, 只要这函数的值在同一条曲线上不变, 例如取 $g(X, E)^3 = \sigma$ 也可以. 这些曲线当然都是微分方程 $dE = f dx$ 的解, σ 是积分常数, 每个 σ 确定一条曲线. 以上讨论的结论是: 我们有了一个新的量 σ , 称之曰“熵”, 不过它的定义仍有些选择余地.

从 (8) 把 E 解出, 得

$$E = E(X, \sigma), \quad (9)$$

也就是说, 物体的能量是坐标 X 和熵 σ 的函数. 如果熵为一定, (9) 就是一条绝热过程画出的曲线. 如果 X 不变, 则能量的改变等于流入的热量:

$$dQ = dE = \left(\frac{\partial E}{\partial \sigma} \right)_X d\sigma. \quad (10)$$

把这结果并入热力学第一定律, 得

$$\begin{aligned} dE &= f dx + \tau d\sigma, \\ f &= \left(\frac{\partial E}{\partial X} \right)_\sigma, \\ \tau &= \left(\frac{\partial E}{\partial \sigma} \right)_X, \\ d\sigma &= \frac{dQ}{\tau}. \end{aligned} \quad (11)$$

以上的分析, 可以稍加推广, 以包括其他各坐标的改变:

$$\begin{aligned} dE &= \sum_k f_k dX_k + \tau d\sigma \\ f_k &= \left(\frac{\partial E}{\partial X_k} \right)_\sigma \\ \tau &= \left(\frac{\partial E}{\partial \sigma} \right)_{X_1, X_2, \dots} \end{aligned} \quad (12)$$

其中 f_k, X_k , 是指 (1) 至 (5) 列举的力和坐标.

因此, 熵可以算作是一个新“坐标”和 X_k 并列, τ 是一个新的“力”和 f_k 并列.

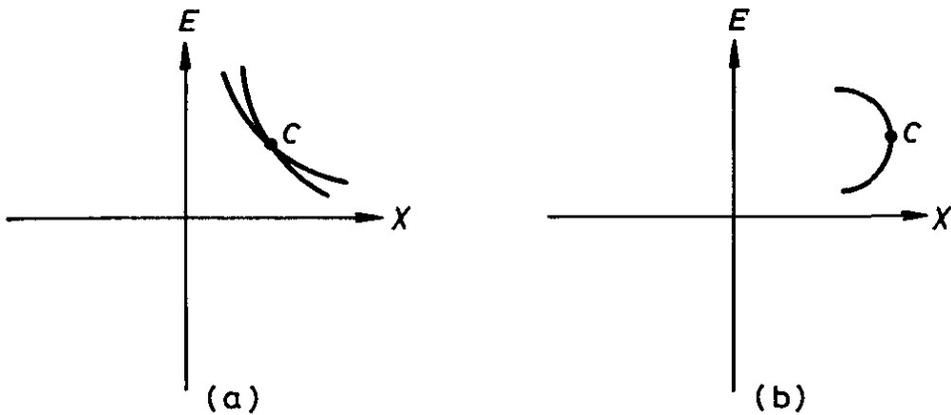


图 2

有两个问题还须要讨论:

- ① 图 1 上画的曲线, 都不相交. 事实上我们已假定交叉点不存在. 如果有交叉点, 如图 2(a) 中的 C 点, 则在 C 点 f 有两个不同的值. 因为 $f = dE/dx$ 是曲线的斜率, f 是平衡下的外力, 也是物体内部支撑住外力的力. 例如压强 p 是外加于物体的力, 也是物体内部的压强, 抵住外力而成平衡. 如果说同一平衡态可以产生两个不同的 f 值, 则实在是与经验不合. 如果每一平衡态的 f 只有一个值, 则图 2(a) 中的交点是不可能的. 打回头的曲线, 如图 2(b) 所示, 亦不可能, 因为在 C 点的 f 值将为 $+\infty$ 及 $-\infty$. 因此, 我们总可以把熵选作随 E 而增加的函数.

注意: 热力学的基本假设 (即平衡态可以由少数坐标完全确定) 至为重要. 如果有其他坐标未被考虑, 他们改变了而我们不知道, 则图 2(a) 和 (b) 的情形都属可能, 以上熵的定义步骤, 也不可靠了. 这基本假设是否正确, 要看所应用的问题. 热力学的隐忧是: 到底要考虑多少坐标才能确定一平衡态? 如何才能知道有没有漏掉一个?

- ② 以上熵的定义, 尚有些选择余地. 如果我们找到 $g(X, E)$, 且由 (8) 定出

了 σ , 则另一定义 $\bar{\sigma}$

$$\begin{aligned}\bar{\sigma} &= A(\sigma), \\ \bar{\tau} &= \frac{\tau}{\left(\frac{d\sigma}{d\bar{\sigma}}\right)}, \\ dQ &= \bar{\tau}d\bar{\sigma}\end{aligned}\quad (13)$$

也可以适用, $A(\sigma)$ 可以任意选择, 唯一的条件是它必须随 σ 增加. 这条件很宽, 所以我们颇可以选个合意的东西来做焓.

4. 绝对温度和熵

在上节, 由绝热过程画出的众曲线 (见图 1), 引入了一个新“坐标”和一个新的“力”. 即熵 σ 及 $\tau = \left(\frac{\partial E}{\partial \sigma}\right)_X$, 见 (9)~(12). 读者或许会问, 为什么不干脆用温度来做一坐标? 为什么要搞个熵出来? 原则上用温度作一新坐标并无不可, 我们可以用定温过程画曲线, 以代替图 1 的曲线. 但是如此“功”和“热”的分辨就比较不清楚. 再则, 如以温度为一“坐标”, 则熵免不了会以一个“力”的姿态出现. 现在我们把平衡的一些特性考虑进去, 调整一下熵的定义, 把 τ 改装成绝对温度 T , 把 σ 改装成一般用的熵. 把 T 当做“力”, 把熵当做“坐标”, 是一个方便的想法.

平衡中的物体有一个重要的特点: 它的每一部分也是在平衡中. 若把某部分看成一个物体, 则其它部分就成了它的环境. 因此所谓“环境因素”, 也必须是物体各部分的性质. 最重要的环境因素是温度, 因为不管物体的结构是多复杂, 多不均匀, 平衡中的物体各部分都会有同样的温度. (当然, 如果某部分是和其他各部分隔开的, 使热不能交流, 则温度可以不同). 这各部分等温的现象, 对 σ 和 τ 的性质, 有何指示? 请看下面.

现在把一个物体看成两部分, 1 和 2 的组合. 照上节的办法, 定出物体的 σ 和 τ , 同时也为这两部分分别定出 σ_1, τ_1 , 及 σ_2, τ_2 . 因为平衡态可由温度及各坐标确定, 所以 $\sigma, \tau, \sigma_1, \tau_1, \sigma_2, \tau_2$ 都是温度 θ 和各坐标的函数. 我们现在用 θ 来表示某一测量标准下的温度, 摄氏、华氏, 什么标准都可以. 令 X_{1k} , $k = 1, 2, 3, \dots$ 为第 1 部分的各坐标. 同样, 令 X_{2k} 为第 2 部分的各坐标. 如果在某过程中, 有热 dQ 流入此物体. 其中 dQ_1 流入第一部分, dQ_2 流入第二部分, 则 $dQ = dQ_1 + dQ_2$, 即

$$\tau d\sigma = \tau_1 d\sigma_1 + \tau_2 d\sigma_2. \quad (14)$$

这个微分式指出 σ 是 σ_1 和 σ_2 的函数, 也就是说, σ 虽是随 X_{1k} , X_{2k} 和 θ 而改变, 但其改变是通过 σ_1 和 σ_2 的改变来决定 (σ_1 和 σ_2 分别是 X_{1k} 和 X_{2k} 的函数, 也是 θ 的函数), 所以我们可以说

$$\sigma = \sigma(\sigma_1, \sigma_2). \quad (15)$$

括号内不再有其他变量了. 当然, σ 对 σ_1 , σ_2 的微分也只能是 σ_1 和 σ_2 的函数:

$$\begin{aligned} \left(\frac{\partial \sigma}{\partial \sigma_1} \right)_{\sigma_2} &= \frac{\tau_1}{\tau} = U_1(\sigma_1, \sigma_2), \\ \left(\frac{\partial \sigma}{\partial \sigma_2} \right)_{\sigma_1} &= \frac{\tau_2}{\tau} = U_2(\sigma_1, \sigma_2). \end{aligned} \quad (16)$$

τ_1 只是 X_{1k} 和 θ 的函数, 和 X_{2k} 无关. 所以 $\tau = \tau_1/U_1(\sigma_1, \sigma_2)$ 和 X_{2k} 的关系是透过 σ_2 . 同理由 (16) 的第二式, 我们知道 τ_2 和 X_{1k} 的关系是透过 σ_1 , 所以说

$$\tau = \tau(\theta, \sigma_1, \sigma_2). \quad (17)$$

把这结果代入 (16), 我们发现 τ_1 , τ_2 是 σ_1 , σ_2 和 θ 的函数. 所以

$$\begin{aligned} \tau_1 &= \tau_1(\sigma_1, \theta), \\ \tau_2 &= \tau_2(\sigma_2, \theta). \end{aligned} \quad (18)$$

τ_1 不能是 σ_2 的函数, 因为 τ_1 只和 X_{1k} 有关. (18) 也指出

$$\tau = \tau(\sigma, \theta). \quad (19)$$

因为每部分有的特质, 总体也有. (我们可以把差不多整个物体定为第 1 部分).

现在留意, (16) 中的比值 τ_1/τ , τ_2/τ , 并不以 θ 为变数, 而 τ_1 , τ_2 , τ 本身都和 θ 有直接关系. 惟一的可能性是

$$\begin{aligned} \tau_1 &= T(\theta)S'_1(\sigma_1), \\ \tau_2 &= T(\theta)S'_2(\sigma_2), \\ \tau &= T(\theta)S'(\sigma). \end{aligned} \quad (20)$$

也就是说他们和 θ 的关系是透过某一函数 $T(\theta)$. S'_1 , S'_2 和 S' 可以是不同的函数, 但是 τ_1 , τ_2 , τ 都有相同的 $T(\theta)$, 如此 τ_1/τ , τ_2/τ 这些比值就不会以 θ 为

变数了. 因为任何两个物体都可以放在一起成一个大物体, 所以 $T(\theta)$ 一旦测出, 就可以普遍通用于任何物体.

现在我们把熵的定义调整. 依照 (13) 定 $S(\sigma)$ 为熵: $\bar{\sigma} = S$, $dA/d\sigma = S'(\sigma)$,

$$S(\sigma) = \int S'(\sigma)d\sigma,$$

$$dQ = \tau d\sigma = T dS. \quad (21)$$

并定 $T(\theta)$ 为“绝对温度”. $T(\theta)$ 的单位仍可任意选择, 因为在 (20) 中, 若以常数 C 乘 $T(\theta)$, 再以 C 除 S'_1, S'_2, S' , 则原式不变, 我们可以定冰点温度为 $T = 273\text{K}$, 是为一般的“绝对温度”. 也可以用其他单位. 一旦 T 的单位定了, S 的单位也定了. 但 S 仍有一积分常数未定, 以后再讨论它.

用 T, S 代 (14) 中的 τ, σ 即得

$$dS = dS_1 + dS_2, \quad (22a)$$

两边的 T 已被约去.

这结果立即可以推广: 把物质分成 m 部分, 则

$$dS = \sum_{i=1}^m dS_i. \quad (22b)$$

每个 dS_i 积分可得 S_i , 但有一未定积分常数. 这积分常数该如何选择呢? 既然 dS 是个总和量, 即总体的 dS 是各部 dS_i 之和, 最好 S 也是个总和量, 即

$$S = \sum_{i=1}^m S_i. \quad (22c)$$

要使 S 成为总和量, 每个 dS_i 的积分常数就不能完全随意选择了. 我们现在把 (22c) 当作一个要求, 这个要求在计算上很容易满足, 但也有些细节需要特别注意, 下节会举例说明. 熵的总和性, 并不是理所当然或显而易见. 以上对熵的分析是十分抽象, 现在来看个例子.

5. 理想气体

现在我们把以上的分析, 应用到理想气体上. 目的不是要导出理想气体的性质, 只是要把以上抽象的讨论, 作一个简单的示范.

理想气体的压强 p 和总能量 E 和容积 V 之间, 有一个简单的关系, 即

$$p = \frac{2E}{3V}. \quad (23)$$

它的由来如下: 令 \mathbf{v} 为一气体分子的速度, v_x 为其垂直于容器某一墙壁的分量. 每次碰壁的动量改变 (冲量) 是 $2mv_x$, 压强是墙上每秒每单位面积所得的冲量:

$$\begin{aligned} p &= \frac{1}{2} \int d^3\mathbf{v} f(\mathbf{v}) v_x 2mv_x \\ &= \int d^3\mathbf{v} f(\mathbf{v}) mv_x^2 \\ &= \frac{2}{3} \int d^3\mathbf{v} f(\mathbf{v}) \frac{1}{2} mv^2 \\ &= \frac{2E}{3V}. \end{aligned} \quad (24)$$

式中 $d^3\mathbf{v} f(\mathbf{v})$ 是速度为 \mathbf{v} 的分子密度, $d^3\mathbf{v} f(\mathbf{v}) v_x$ 则为其在 X 方向的流量. 开头有 $\frac{1}{2}$, 是因为平均而言只有一半分子是朝这墙走, 其它离墙而去, 不产生压强. 速度方向分布是均匀的, 所以积分下的 v_x^2 可换成 $\frac{1}{3}(v_x^2 + v_y^2 + v_z^2) = \frac{1}{3}v^2$. (24) 之成立, 和 $f(\mathbf{v})$ 的详情无关.

假定分子数 N 不变, 则容积 V 是唯一的坐标 (还没有把熵算进去), 绝热过程画出的曲线是由

$$dE = -pdV \quad (25)$$

决定. 将 (24) 代入, 得

$$\frac{dE}{E} + \frac{2}{3} \frac{dV}{V} = 0. \quad (26)$$

解之得

$$EV^{2/3} = \sigma. \quad (27)$$

σ 是一积分常数, 每个 σ 的值, 确立一条在 (V, E) 平面上的曲线 (见 (8)), 我们可以令 σ 为熵, 则

$$\tau = \left(\frac{\partial E}{\partial \sigma} \right)_V = V^{-2/3}. \quad (28)$$

现在我们有坐标 V 和 σ , 它们的值确定这理想气体的平衡态. 气体的能量是 $E = \sigma V^{-2/3}$. 温度的概念尚未出现. 上节的讨论指出: 只要有了测定的温度 θ , 绝对温度 $T(\theta)$ 就定了. 不过, 我们得先有 θ 才行. 压强和能量却

是明确的力学概念, 所以我们能找出他们的关系 (24). 但温度却不然, 它的测定虽然容易, 但它和分子运动的关系, 却不清楚. 好在我们可以利用实验结果. 根据实验, 稀疏的气体的压强 p , 乘上容积 V , 再除以分子数目 N , 在一定的温度下不变, 即在不同的 p 、 V 、 N 情况下

$$\frac{pV}{N} = f(\theta) \quad (29)$$

只和温度有关, 温度 θ 用什么标准都可以. $f(\theta)$ 可以用实验测定.

上一节讨论指出, τ 必是绝对温度 $T(\theta)$ 乘上一个 σ 的函数 $S'(\sigma)$, (见 (20)). 所以, 从 (28) 的 τ , 和理想气体的实验结果 (29), 我们可以求出绝对温度来. 步骤如下: 从 (27) 和 (23) 求出 pV , 代入 (29), 得

$$f(\theta) = \frac{2\sigma}{3N} V^{-2/3}. \quad (30)$$

因此, (28) 可写成

$$\tau = f(\theta) \frac{3N}{2\sigma}. \quad (31)$$

现在定绝对温度 $f(\theta)/k = T(\theta)$, 即

$$\begin{aligned} kT &= \frac{pV}{N}, \\ S'(\sigma) &= \frac{3Nk}{2\sigma}. \end{aligned} \quad (32)$$

如此, 绝对温度可以由 p 、 V 和 N 来测定. 已有的温度标准和 T 的关系, 因此由 $f(\theta)$ 订出, k 是一常数, 由 T 的单位决定. $\frac{pV}{N}$ 的单位是能量单位, 如果 T 是一般的度数, 则

$$k = 1.38 \times 10^{-23} \text{ J/K} \quad (33)$$

是“玻尔兹曼”常数, 简称“热常数”. 如果干脆用能量的单位来量 T , 则

$$k = 1. \quad (34)$$

我们以后用 $k = 1$, 省很多麻烦. k 的选择只是单位的选择, 随时可以更改.

从 (32) 得

$$S = \int S'(\sigma) d\sigma = \frac{3}{2} N \ln \sigma + \text{常数}. \quad (35)$$

式中的常数, 可以是 N 的函数. 因为 (25) 没有把 dN 算进去, 所以 N 一直是个不变量. 把 (27) 代入 (35), 并把 (35) 的常数写成

$$N \ln \left[\frac{C^{3/2}}{N^{5/2}} \right],$$

得

$$S = N \ln \left[\frac{V}{N} \left(\frac{CE}{N} \right)^{3/2} \right], \quad (36)$$

是为理想气态的熵, $\frac{E}{N}$ 是每个粒子的平均能量, $\frac{N}{V}$ 是密度, C 是一个不随 N 变的常数. 因为 S 须是一总和量, 在温度和密度一定时, 必须和 N 成正比.

现在提一下所谓“混合熵”的问题. 两个同温, 同压的气体, 各有 N_1, N_2 个分子, 各占容积 V_1, V_2 . 现在把他们之间通条管子, 两气体就互相混合. 混合前后的总压强, 温度都一样. 其实我们看不出有任何变化, (假设气体有相同的颜色.) 令

$$V = V_1 + V_2, \quad N = N_1 + N_2, \quad E = E_1 + E_2,$$

混合前二气体的熵可由 (36) 算出

$$\begin{aligned} S_1 &= \frac{3}{2} N_1 \ln \varepsilon C + N_1 \ln \frac{V_1}{N_1}, \\ S_2 &= \frac{3}{2} N_2 \ln \varepsilon C + N_2 \ln \frac{V_2}{N_2}, \end{aligned} \quad (37)$$

$$\begin{aligned} \varepsilon &= \frac{E_1}{N_1} = \frac{E_2}{N_2} = \frac{E}{N}, \\ \frac{V_1}{N_1} &= \frac{V_2}{N_2} = \frac{V}{N}. \end{aligned}$$

二熵之和是

$$S = S_1 + S_2 = \frac{3}{2} N \ln \varepsilon C + N \ln \frac{V}{N}. \quad (38)$$

混合后, 气体能量不变, 但各气体一同占有两个容器, 容积是 V , 所以他们的熵是

$$\begin{aligned} S'_1 &= \frac{3}{2} N_1 \ln \varepsilon C + N_1 \ln \frac{V}{N_1}, \\ S'_2 &= \frac{3}{2} N_2 \ln \varepsilon C + N_2 \ln \frac{V}{N_2}. \end{aligned} \quad (39)$$

它们的和 $S' = S'_1 + S'_2$ 要比 S 大些:

$$S' - S = N_1 \ln \frac{N}{N_1} + N_2 \ln \frac{N}{N_2}, \quad (40)$$

是所谓“混合熵”, 由混合而来.

以前曾说过,任一物体可分成几个部分,我们可以分别计算熵,再相加,得物体之熵.这混合熵的例子,叫我们要特别小心,如果一气体是同一种分子组成,我们就不可把分子分成两组,用(39)分别计算熵再相加.如果只有一种分子,答案是(38),因为同一种分子是没有什么好混合的了.这一点在我们讨论物体各部分的“独立性时”会再考虑.

6. 热力学第二定律

这个定律和一般守恒定律,牛顿定律、电学定律都不一样,可以说是物理学中最特殊的定律.他说,没有一个机器可以不断地把物质的内能抽出来做功.举个例子:造一个轮船引擎,取海水分子的能量,来推动轮船.不用油,不用蒸汽,不用核能,那多好?但热力学第二定律说这办不到.也就是说,没有一个机器,其唯一作用是把内能改成功能.当然,一般的引擎是把内能改成功能,如蒸汽机,把水蒸气的能量来作功.但他们都有副作用,都要浪费一些能量,而且必须要把温度加高(比环境高),才能用.

这定律可以说是由经验建立.到现在为止,还没有被违反过.但也没有真正地被证明过.

从这个定律,立刻可以得到一些结论,例如:

把热能从低温处移到高温,必须做功.如不必做功,则可以把海水能量送到蒸汽机里,再由蒸汽机做功.如此则违反了第二定律.又例如:一个物体不会自己变成一半冷,一半热.

最重要的一个结论是:一个绝热的物体之熵,不会减少.如果会减少我们就可以加热,使熵增加,提高温度,再由绝热过程做功,把热能用掉,再让它的熵减少,如此可以不断地把热转换成功,违反了第二定律.以上每一个结论,都可以看成第二定律的一个说法.内容都一样.

因此,一个平衡的物体,不能对外做功,它的熵值必是极大.以后我们会再讨论这定律.

7. 热机的效率和速率

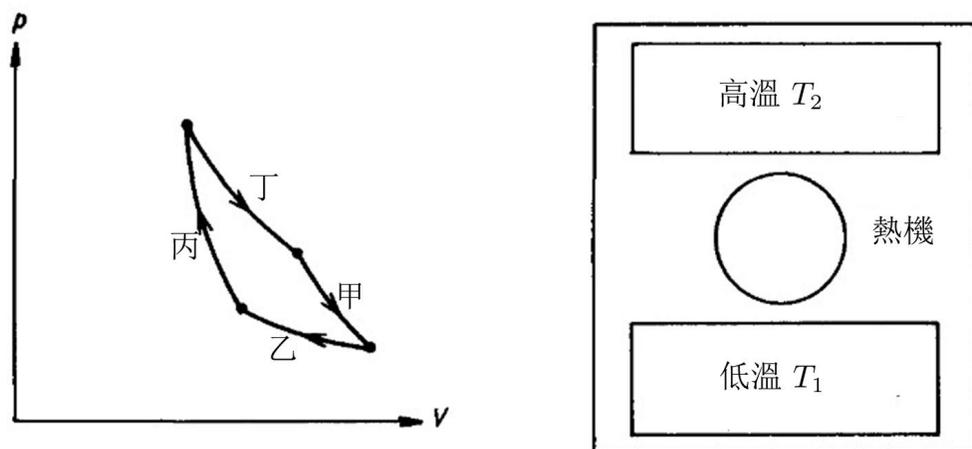


图 3

先复习一下“理想热机”即“卡诺热机”。假设一个气缸里装了理想气体。这热机有四个过程：见图 3。

(甲) 绝热膨胀, 做功。

(乙) 等温压缩, 放热, 并受功, (和低温环境接触, 温度 T_1)。

(丙) 绝热压缩, 受功。

(丁) 等温膨胀, 吸热, 并做功, (和高温环境接触, 温度 T_2)。

所谓“理想”, 是指吸热、放热过程十分缓慢, 机器和高温环境接触时, 机内外温度看成一样 T_2 , 热慢慢流入。和低温环境接触时, 内外温度是 T_1 , 热慢慢流出。

每次循环, 净做功为

$$W = Q_2 - Q_1, \quad (41)$$

即吸入热减去放出热。热机的效率的定义是

$$\eta = \frac{W}{Q_2} = 1 - \frac{Q_1}{Q_2}, \quad (42)$$

Q_1 是浪费掉的热能, 根据第二定律, 以上的循环过程必使环境之总熵增加。我们现在把高低温环境看作一个绝热的大物体 (即图 3 中的大框框)。总熵之增加为

$$\frac{Q_1}{T_1} - \frac{Q_2}{T_2} = \frac{Q_2 - W}{T_1} - \frac{Q_2}{T_2} \geq 0. \quad (43)$$

因此

$$\eta \leq 1 - \frac{T_1}{T_2}. \quad (44)$$

对理想热机而言, $Q_1 = T_1 \Delta S$, $Q_2 = T_2 \Delta S$, (见图 4). $\Delta S = S'' - S'$. 因此, $\eta = 1 - T_1/T_2$. 理想热机的效率, 是最高效率.

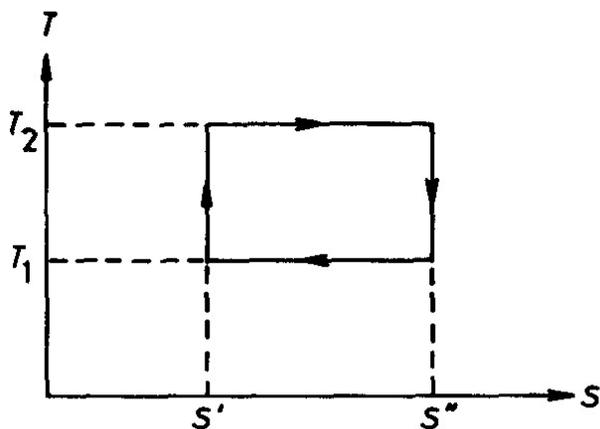


图 4

如果热机在吸热和放热时, 温度和环境相同, 则吸热和放热的速率必为零. 因此, 要达到最高效率, 热机就得是无限慢的机器. 从这一观点, 当然非常不理想. 为了增加速度, 必须牺牲一点效率. 如何才是最好的折衷办法? 这当然要看实用的情况. 下面我们粗略地估计一下速度和效率的关系.

令 T'_1 为热机和低温环境接触时的温度, T'_2 为它和高温环境接触时的温度. 当然 $T_2 > T'_2 > T'_1 > T_1$, 是必要条件. 令

$$\begin{aligned} \alpha_1 &\equiv \frac{T'_1}{T_1}, \\ \alpha_2 &\equiv \frac{T_2}{T'_2}. \end{aligned} \quad (45)$$

热机的效率是

$$\eta = 1 - \frac{T'_1}{T'_2} = 1 - \frac{T_1}{T_2} \alpha_1 \alpha_2. \quad (46)$$

每次循环所费的时间约是放热、吸热时间之和:

$$t \sim \frac{Q_2}{T_2 - T'_2} + \frac{Q_1}{T'_1 - T_1}, \quad (47)$$

以上假设热流量和温差成正比.

因为 $Q_2 \propto T_2'$, $Q_1 \propto T_1'$, 因此

$$t \sim \frac{T_2'}{T_2 - T_2'} + \frac{T_1'}{T_1' - T_1} \\ = \frac{\alpha_1 \alpha_2 - 1}{(\alpha_1 - 1)(\alpha_2 - 1)}. \quad (48)$$

图 5 显示 α^2 的函数 η 及 $1/t$, 即假设 $\alpha_1 = \alpha_2 = \alpha$ 时引擎的效率以及循环频率.

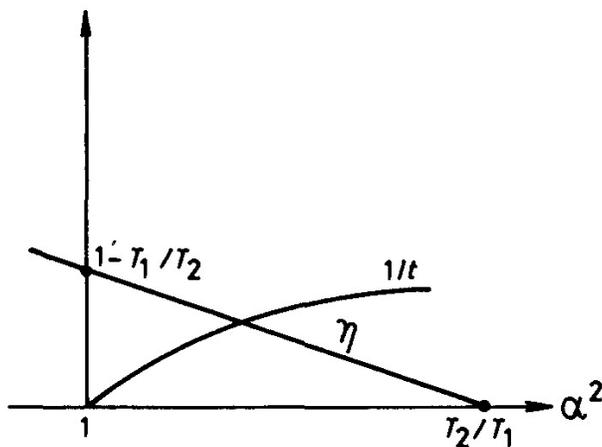


图 5

如果 T_1/T_2 很小, 即本来理想效率很高, 则 η/t 之极大值约在 $\alpha = (T_2/T_1)^{1/3}$, 在此值的效率是

$$\eta \approx 1 - \left(\frac{T_1}{T_2}\right)^{1/3}. \quad (49)$$

因此, 效率大受影响. 如果 T_1/T_2 接近 1, 即本来理想效率就不高, 则 η/t 之极大值约在 $\alpha \approx 1 + \frac{1}{4} \left(1 - \frac{T_1}{T_2}\right)$, η 之值为

$$\eta \approx \frac{1}{2} \left(1 - \frac{T_1}{T_2}\right), \quad (50)$$

即理想值之一半. 以上这些结果只是一粗略估计. η/t 之值并不一定是越大越合适. 但这些结果的确指出实用热机效率比理想热机效率差得很远.

以上各节只是把热力学的几个基本概念作了个简单的介绍, 指出各“坐标”和“力”的概念, 并把熵和温度归纳到这组概念中. 热力学的讨论, 当然不是到此为止. 进一步的讨论, 将分散在以后各章节, 和统计力学混在一起讲.

讨论问题二

1. 拉力即负的压强, 读者以前可能学到过压强必须为正的热力学定理. 在第一章已提过, 在长时间下, 拉力下的物体为不稳定, 但在短时间内, 拉力可以为正, 热力学性质仍很明确, 各弹性系数都是很有用的热力学性质.

读者应复习热力学讲稳定性的部分.

2. 第 3, 4 节的内容, 大致是所谓“Caratheodory 定理”及其后果. 吴大猷(1979), Wannier(1966) 书中都有类似讨论. 读者应把第 3, 4 节中所有数学步骤自己推演一次.
3. 两个绝热流体 1 和 2, 体积固定, 热容率和温度各为 C_1 、 T_1 和 C_2 、 T_2 ($T_1 > T_2$). 若以一卡诺热机, 从 1 取热、排热于 2, 直到 1, 2 均达相同的温度 T_0 .

(A) 试计算 T_0 和卡诺热机所作的功.

(B) 若热量可直接由 1 流至 2, 则达平衡时, 温度为何? 熵的变化若干?

(C) 若 C_1, C_2 均为正, 试证熵值增大.

以上假定 C_1, C_2 大致不随温度改变.

4. 导出 (48). 如果热流量不和温度成正比, 结果就不一样了. 令流量为 $A(\Delta T) + B(\Delta T)^2$, 其中 ΔT 为温差, A, B 为常数, 如 B 项较 A 项小很多, 求 (48)–(50) 的修正.
5. 物态方程式代表变量间的一个关系, 这些变量包括热力学中的力和坐标. 例如 $pV = NT$ 是理想气体的物态方程式. 物态方程式可以直接由实验测定.

某磁体的物态方程式是

$$M = \frac{Ah}{T},$$

M 为磁矩, h 为外加磁场, A 为一常数. 求熵和 h 的关系. (提示: 由 $dE = TdS + hdM$ 可得 $\frac{\partial S}{\partial h} = \frac{\partial m}{\partial T}$).

6. 热力学的“隐忧”是: 不知道要有多少坐标或力才能确定平衡态. 一定的环境下, 不一定有确定的平衡态. 这在第一节就谈过. 例如钢的性质

要看碳原子在铁中如何排列. 知道有多少碳, 多少铁, 不足以定平衡态. 热力学不是一门完整的学问, 应用时必须考虑很多其他细节.

有两个铜球, 一样重, 温度相同, 同在一大气压下, 但一个比另一个大 10 倍, 为什么?

答: 大的中间是空的.

这并非开玩笑的问题, 读者多想想, 坐标, 时间尺度, 等等.

第三章 细致平衡定律

分子运动可以看作一连串的反应, 反应就是改变. 分子碰撞, 使各个动量改变, 就是一种反应. 如果分子结构因反应而改变, 则为化学反应. 细致平衡的意思是: 在平衡中, 每个反应发生次数和反方向的反应次数相同. 也就是说, 不但有宏观尺度的平衡, 而且每一个微观尺度的反应都平衡. 细致平衡可以从时间正反对称的观点来了解, 如果时光倒流, 分子的运动定律是不变的. 平衡是一个不变的状态. 无论时光正流或倒流, 看起来是完全一样, 正, 反方向的反应次数也就必须一样.

细致平衡可以直接用来分析理想气体, 包括高密度的量子气体. 所谓量子气体, 是指因密度太高, 气体粒子的波动性必须考虑进去. 理想气体中的反应很容易分析. 加一些看来十分合理的假设, 即可由细致平衡求出粒子的能量分布, 并阐明温度、化学势和各反应守恒量的关系. 细致平衡还可以反过来用, 即从已知的粒子分布来决定反应的一些特性. 这些是本章的主要内容.

细致平衡可以用来做统计力学的出发点. 原则上它可以用来分析任何平衡物体. 现在通用的“蒙特卡罗模拟方法”就是以细致平衡为基础的计算法则, 这方法用计算机的数值计算来模拟各种反应. 我们会在第二十二章讲它.

细致平衡并不能取代第五章的基本假设, 因为我们对较复杂的反应尚无法作仔细分析. 即使理想气体的分析, 也必须假设各反应为独立, 在应用细致平衡时, 我们免不了会引入基本假设, 其出现方式不同而已.

1. 状态的能量分布

“状态”也就是“轨道”。例如，原子模型中的电子状态，是最熟知的。每个状态由一组量子数 (n, l, m, σ) 指定。状态的能量是：

$$-\frac{Z^2 R}{n^2}, \quad n = 1, 2, 3, \dots, \\ R = 13.6\text{eV}. \quad (1)$$

Z 是核子电荷，(以电子电荷的大小 e 为单位)。 l, m 的意义是角动量及其在某方向的投影， σ 是电子自旋在某方向的投影。

每个状态只能容纳一个电子，因为电子是“费米”粒子。自然界中粒子可分两类，一类为波色，一类为费米，后者必须遵守“独占定律”，即每个粒子不能和同样的粒子同占一状态。前者不受此限。以上的原子模型，是把 Z 个电子放在能量最低的 Z 个状态中，原子中的电子状态是由核子的静电势能决定。一般问题中的粒子状态要由势能决定，我们复习一下最简单的情形，并计算状态的能量分布。

(甲) 能量分布的统计

先看一个自由粒子的诸状态。这粒子是在一个容积为 L^3 的立方容器内，并有“循环边界”，也就是说 $x = 0$ 和 $x = L$ 是同一面， y, z 方向亦然。这容器因此实在是没有边界的，是个三度空间的“环式”容器。量子力学的分析结果是：粒子的形象或形态可以由三个整数 $(n_x, n_y, n_z) = \mathbf{n}$ 指定 (暂不考虑粒子自旋或其他内在形态)。 \mathbf{n} 的能量是

$$\varepsilon_{\mathbf{p}} = \frac{p^2}{2m}, \quad (2) \\ \mathbf{p} = \frac{2\pi}{L}(n_x, n_y, n_z), \\ n_x, n_y, n_z = 0, \pm 1, \pm 2, \dots$$

状态 \mathbf{n} 的“波函数”是一个平面波

$$\varphi_{\mathbf{n}} = \frac{1}{L^{3/2}} e^{i\mathbf{p}\cdot\mathbf{r}}, \quad (3)$$

(我们把 \hbar 当做 1，以简化式子，需要时再指明)。每形象是一状态。各状态的能量分布是

$$L^3 g(\varepsilon) = \sum_{n_x=-\infty}^{\infty} \sum_{n_y=-\infty}^{\infty} \sum_{n_z=-\infty}^{\infty} \delta \left[\varepsilon - \frac{1}{2m} \left(\frac{2\pi}{L} \right)^2 (n_x^2 + n_y^2 + n_z^2) \right]. \quad (4)$$

如果 L 很大, ε 不小, (4) 式收集的状态都有很大的 n_x, n_y, n_z , 则 n_x, n_y, n_z 可当做连续变量, \sum 可用积分代替

$$\sum_{n_x} \sum_{n_y} \sum_{n_z} = \int dn_x \int dn_y \int dn_z. \quad (5)$$

最后, 对 \mathbf{n} 的积分可改做对 \mathbf{p} 的 [见 (2)], 因此 (4) 变成了

$$\begin{aligned} L^3 g(\varepsilon) &= \frac{L^3}{(2\pi)^3} \int d^3 \mathbf{p} \delta\left(\varepsilon - \frac{p^2}{2m}\right) \\ &= \frac{L^3}{2\pi^2} m \sqrt{2m\varepsilon}. \end{aligned} \quad (6)$$

注意: 如果以上不用循环边界, 而用“硬壁边界”, 则波函数 $\varphi_n(\mathbf{r})$ 在边界须为零, 则 $\mathbf{n} = (n_x, n_y, n_z)$,

$$\begin{aligned} \varphi_n &= \left(\frac{2}{L}\right)^{3/2} \sin(xp_x) \sin(yp_y) \sin(zp_z), \\ \mathbf{p} &= \frac{\pi}{L}(n_x, n_y, n_z), \\ \varepsilon_{\mathbf{p}} &= \frac{p^2}{2m}, \quad n_x, n_y, n_z = 1, 2, 3, \dots \end{aligned} \quad (7)$$

(7) 和 (2) 不大一样.

读者试证: 用 (7) 算出的 $g(\varepsilon)$ 和 (6) 一样, 只要 L 很大, 边界条件对 $g(\varepsilon)$ 没有影响.

(乙) 短波近似下的能量统计

在短波近似法 (WKB 近似法) 中, 粒子的轨道可用古典力学求出, 但每个轨道必须适合一“量子条件”. 先看一个粒子, 以一维空间的情形为例, 这量子条件是:

$$\oint p(x) dx = nh, \quad (8)$$

$$\varepsilon_n = \frac{p(x)^2}{2m} + U(x). \quad (9)$$

式中 $U(x)$ 是粒子的势能, $p(x)$ 是动量, h 是量子常数, n 是一整数. (8) 的积分是沿着状态空间的轨道做, 见图 1.

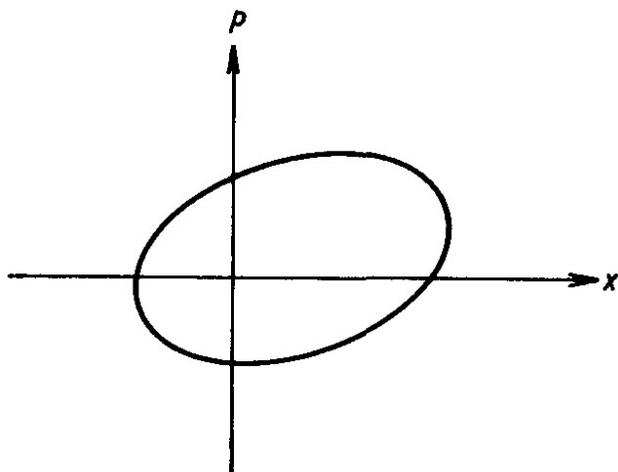


图 1

量子条件 (8) 决定此粒子的各能量 ε_n 和各轨道, 每一整数指明一轨道, 每一轨道是一状态. 这方法只合用于短波情形, 即在 $\frac{h}{p}$ 这段距离内, $U(x)$ 改变得很少. $\frac{h}{p}$ 是粒子的波长. (8) 的积分恰好是轨道包围的面积.

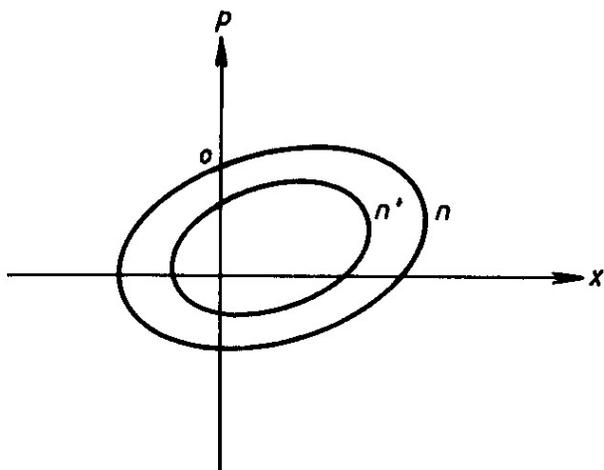


图 2

图 2 中的环形区域, 是由两轨道 n 和 n' 划定, 这两轨道之间有 $n - n'$ 个轨道在这区域内, 依照 (8)

$$n - n' = \frac{1}{h} \int dp dx. \quad (10)$$

也就是说, 区域面积除以 h 就得区域中的轨道数. 因此, 若要统计状态, 我们

只须求面积再除以 h , 同时我们也可以把诸状态想象成均匀地分布于状态空间的点, 其密度为 $1/h$. 所以

$$L^3 g(\varepsilon) = \frac{1}{h} \int dp dx \delta \left(\varepsilon - \frac{p^2}{2m} - U(x) \right). \quad (11)$$

(11) 这积分, 并不需要把轨道解出, 不必从 (8) 求出每个 ε_n , 只要知道 $U(x)$ 就行了, 所以短波近似法用起来十分方便, 但读者须牢记它的可用范围. $U(x)$ 必须在距离 h/p 之内几乎不变, 一般说来, 这近似法只合用于高能量状态.

如果是三度空间, 则 (11) 可以推广为

$$L^3 g(\varepsilon) = \frac{1}{h^3} \int d^3 p d^3 r \delta \left(\varepsilon - \frac{p^2}{2m} - U(\mathbf{r}) \right). \quad (12)$$

如果粒子自旋非 0, 则需要加一个自旋量子数来指定状态. 波色粒子的自旋为整数 (以 \hbar 为单位), 费米粒子的自旋为半整数 (整数加 $\frac{1}{2}$).

最常讨论的费米粒子是电子, 其自旋为 $\frac{1}{2}$, 波色粒子有光子 (自旋为 1)、 He^4 (自旋为 0 等).

有些固体中或液体中的小幅度运动可以用特别的“粒子”来描述. 例如声子 (描述晶格振动), 自旋波 (或称“磁子”, 描述磁矩振动) 等. 这些粒子多是波色性的. (注意, 因为这些“粒子”在真空中并不存在, 而是以物质为传播介质. 一般物质并非均匀, 或有晶格结构, 所以“自旋”这概念必须由物质结构的各种对称性取代).

2. 费米粒子的分布

现在用细致平衡定律来求得粒子的能量分布, 一方面说明这定律的内容, 一方面强调碰撞是理想气体的基本性质, 没有碰撞, 就没有热力学的特性.

理想气体之所以为“理想”, 是指碰撞时间 (即粒子接触时间) 很短, 接触半径小, (即粒子作用有效距离小).

每次碰撞就是一次反应, 我们假设各反应为独立的事件, 不考虑各反应之间的关联. 这是一个相当有效的假设, 使理论分析大为简化.

假想一费米粒子气体, 令 $f_{\mathbf{k}}$ 为 \mathbf{k} 状态的平均粒子数. \mathbf{k} 指粒子的动量和自旋. 由于碰撞, 各粒子不断地换状态. 例如

$$1 + 2 \longleftrightarrow 3 + 4, \quad (13)$$

即在 $\mathbf{1}, \mathbf{2}$ 的二粒子相互作用而跳到 $\mathbf{3}, \mathbf{4}$ 二状态. 这反应的速率是

$$f_1 f_2 (1 - f_3) (1 - f_4) R. \quad (14)$$

f_1, f_2 是状态 $\mathbf{1}, \mathbf{2}$ 都有粒子的概率, $(1 - f_3)(1 - f_4)$ 为 $\mathbf{3}, \mathbf{4}$ 状态都没有粒子的概率, 因为粒子是费米子, 如果 $\mathbf{3}, \mathbf{4}$ 有一个是满的, 则反应不能进行. 这速率必须和反方向的反应速率相同, 即

$$f_1 f_2 (1 - f_3) (1 - f_4) R = f_3 f_4 (1 - f_1) (1 - f_2) R'. \quad (15)$$

根据量子力学的计算, $R = R'$. 这是因为量子力学方程式在时间反演下不变,

$$\begin{aligned} R &= |M|^2, \\ R' &= |M^*|^2 = R. \end{aligned} \quad (16)$$

M 是散射算符在 $|\mathbf{1}, \mathbf{2}\rangle$ 和 $|\mathbf{3}, \mathbf{4}\rangle$ 形态间的作用幅度, 如果时间反演, $|\mathbf{1}, \mathbf{2}\rangle$ 和 $|\mathbf{3}, \mathbf{4}\rangle$ 对换, M 即成 M^* . 从 (15) 得

$$\frac{f_1}{1 - f_1} \times \frac{f_2}{1 - f_2} = \frac{f_3}{1 - f_3} \times \frac{f_4}{1 - f_4}. \quad (17)$$

如果有三个粒子作用

$$\mathbf{1} + \mathbf{2} + \mathbf{3} \longleftrightarrow \mathbf{4} + \mathbf{5} + \mathbf{6},$$

细致平衡的条件导致和 (17) 类似的结果

$$\frac{f_1}{1 - f_1} \times \frac{f_2}{1 - f_2} \times \frac{f_3}{1 - f_3} = \frac{f_4}{1 - f_4} \times \frac{f_5}{1 - f_5} \times \frac{f_6}{1 - f_6}. \quad (18)$$

(17) 和 (18) 是非常苛刻的条件, 所有的 f_k 必须适合. 我们可从 (17), (18) 找出 f_k 来. 当然, 如果所有的 f_k 都合乎

$$\frac{f_k}{1 - f_k} = \text{常数}, \quad (19)$$

则 (17), (18) 都适合了, 但这表示每个 f_k 都一样, 无论状态的能量多高. 这是不合理的. 要解 (17), (18), 我们还得考虑反应的守恒量. (17), (18) 的反应前后, 能量和粒子数是守恒的:

$$(17): \text{粒子数} = 2, \text{能量} = \varepsilon_1 + \varepsilon_2 = \varepsilon_3 + \varepsilon_4,$$

$$(18): \text{粒子数} = 3, \text{能量} = \varepsilon_1 + \varepsilon_2 + \varepsilon_3 = \varepsilon_4 + \varepsilon_5 + \varepsilon_6. \quad (20)$$

因此, (17), (18) 还有一解, 即

$$\frac{f_{\mathbf{k}}}{1 - f_{\mathbf{k}}} = e^{-\alpha - \beta \varepsilon_{\mathbf{k}}}, \quad (21)$$

α, β 为常数. 证明很简单, 将 (21) 代入 (17), (18), 并用 (20), 即得

$$\begin{aligned} e^{-2\alpha - \beta(\varepsilon_1 + \varepsilon_2)} &= e^{-2\alpha - \beta(\varepsilon_3 + \varepsilon_4)}, \\ e^{-3\alpha - \beta(\varepsilon_1 + \varepsilon_2 + \varepsilon_3)} &= e^{-3\alpha - \beta(\varepsilon_4 + \varepsilon_5 + \varepsilon_6)}. \end{aligned} \quad (22)$$

(21) 当然就是

$$f_{\mathbf{k}} = \frac{1}{e^{\alpha + \beta \varepsilon_{\mathbf{k}}} + 1} \quad (23)$$

这是适合细致平衡条件的粒子分布. α, β 为众粒子所共有, 应该和温度及化学势有关. 总能量 E 及总粒子数 N 和 $f_{\mathbf{k}}$ 的关系是

$$\begin{aligned} E &= \sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} f_{\mathbf{k}} = V \int d\varepsilon g(\varepsilon) \frac{\varepsilon}{e^{\alpha + \beta \varepsilon} + 1}, \\ N &= \sum_{\mathbf{k}} f_{\mathbf{k}} = V \int d\varepsilon g(\varepsilon) \frac{1}{e^{\alpha + \beta \varepsilon} + 1}. \end{aligned} \quad (24)$$

如果气体很稀, 即 $f_{\mathbf{k}} \ll 1$, 则

$$f_{\mathbf{k}} \approx e^{-\alpha - \beta \varepsilon_{\mathbf{k}}}. \quad (25)$$

这是熟知的理想气体结果, 即“玻尔兹曼分布”, 由状态能量可以立即算出. 因为 $g(\varepsilon) \propto \sqrt{\varepsilon}$ (见 (5)), 得

$$\frac{E}{N} = \frac{3}{2\beta}. \quad (26)$$

我们发现 $1/\beta$ 即是温度 T , 因为我们从热力学就知道 $E/N = \frac{3}{2}T$. α 可以由 (24) 求得. 用 (25) 作 f , 得

$$\begin{aligned} \frac{N}{V} &= e^{-\alpha} \int d\varepsilon g(\varepsilon) e^{-\beta \varepsilon} \\ &= e^{-\alpha} (2s + 1) (2\pi m T)^{3/2} / h^2, \end{aligned} \quad (27)$$

即

$$\alpha = \ln \left((2s + 1) (2\pi m T)^{3/2} \frac{V}{N h^2} \right). \quad (28)$$

s 为粒子自旋. 理想气体的化学势 μ 可以由热力学结果 (2.36) 得来

$$\begin{aligned} -\frac{\mu}{T} &= \left(\frac{\partial S}{\partial N} \right)_E = \ln \frac{V}{N} \left(\frac{CE}{N} \right)^{3/2} - \frac{5}{2} \\ &= \ln \frac{V}{N} + \ln \left(\frac{3}{2} TC \right)^{3/2} - \frac{5}{2}. \end{aligned} \quad (29)$$

(28) 和 (29) 全等, 除了常数 C 未定外. 我们可以定义

$$\frac{\mu}{T} = -\alpha, \quad (30)$$

因而决定 C 之值. (2.36) 变成

$$S = N \ln \left(\frac{V}{N} \left(\frac{4\pi m E}{3N h^3} \right)^{3/2} (2s+1) \right) + \frac{5}{2} N. \quad (31)$$

把 (30) 代入 (23) 则得常见的“费米”分布:

$$f_{\mathbf{k}} = \frac{1}{e^{(\varepsilon_{\mathbf{k}} - \mu)/T} + 1}. \quad (32)$$

注意, 以上我们用已知的理想气体性质来决定 α, β 的意义, 但 (32) 在任何温度都可以用. 理想气体可看作一个温度计及化学势计. 说得更明白些: 把气体容器的一角落的势能 $U(\mathbf{r})$ 加的很高使气体在这角落的密度变得非常低, 而使理想气体的公式可用. 但 μ, T 为气体所有部分共有, 因为各反应牵连到所有粒子. 因此, 一旦测得这角落的 T 和 μ , 则知全气体的 T 和 μ .

一般密度时的熵可以由热容率积分得来

$$S = \int^T \left(\frac{\partial E}{\partial T} \right)_V \frac{dT}{T}. \quad (33)$$

积分常数可以由 $T \rightarrow \infty$ 之值, 即理想气体 (31) 式决定.

因此, 从细致平衡, 及第二章热力学的结果, 我们决定了费米粒子气体的粒子分布及所有的平衡性质. 熵的定义中的常数 C 由 (30) 决定了, 这是一个方便的选择. 我们在下节末还要讨论这一问题. 最重要的费米粒子是电子, 我们在下一章讨论.

3. 守恒定律, 化学势和熵的常数项

以上的分析很明显地指出守恒定律的影响. (21) 式是由粒子数及能量守恒律得来, 因此 T 和 μ 是用以限定总能量和总粒子数的“力”.

如果还有别的不变量, 则 (17), (18) 还可以有其他的解. 例如, 动量守恒, 即

$$\begin{aligned} (17) \text{ 的 } \mathbf{p}_1 + \mathbf{p}_2 &= \mathbf{p}_3 + \mathbf{p}_4, \\ (18) \text{ 的 } \mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3 &= \mathbf{p}_4 + \mathbf{p}_5 + \mathbf{p}_6. \end{aligned} \quad (34)$$

这个守恒律容许

$$\frac{f_{\mathbf{k}}}{1 - f_{\mathbf{k}}} = e^{-\alpha - \beta \varepsilon_{\mathbf{k}} + \beta \mathbf{v} \cdot \mathbf{p}_{\mathbf{k}}} \quad (35)$$

为 (17), (18) 之解, 则

$$f_{\mathbf{k}} = \frac{1}{e^{(\varepsilon_{\mathbf{k}} - \mathbf{v} \cdot \mathbf{p}_{\mathbf{k}} - \mu)/T} + 1}. \quad (36)$$

\mathbf{v} 是整个气体的流速, 但如果物体的容器不动, 则和器壁碰撞会使粒子动量不守恒, \mathbf{v} 必须是 0.

如果碰撞前后总自旋不变, 则

$$f_{\mathbf{k}} = \frac{1}{e^{(\varepsilon_{\mathbf{k}} - \mu - \mathbf{s} \cdot \mathbf{h})/T} + 1}. \quad (37)$$

\mathbf{h} 可看作维持总自旋的外力, 如果 $s = \frac{1}{2}$, 则 $\mathbf{s} \cdot \mathbf{h} = \pm \frac{1}{2} h$.

总之, 每一个守恒律会引出一个“力”来. 以上的 $1/T$, $-\mu/T$, \mathbf{v}/T , \mathbf{h}/T 都是明显的例子. 反过来, 每一个守恒律被破坏, 则会使一个力为零, 例如容器壁使动量不守恒, 故 $\mathbf{v} = 0$. 如果粒子数不守恒, 即可因反应增减, 则 μ 必是 0, 比较复杂的例子是化学反应, 例如有三个分子 A, B, C 混合在一起而有反应



A, B, C 之粒子数不守恒, 但不是完全独立. 把细致平衡用到 (38), 则 (17) 之推广为

$$\frac{f_A}{1 - f_A} \times \frac{f_B}{1 - f_B} = \frac{f_C}{1 - f_C}. \quad (39)$$

(21) 式子仍可用, 只是各 α 须适合

$$\alpha_A + \alpha_B = \alpha_C, \quad (40)$$

由 (30), 得

$$\mu_A + \mu_B - \mu_C = 0. \quad (41)$$

三个本来为独立的化学势, 因 (38) 的反应, 只剩下两个, 其他一个由 (41) 决定了. 每一种化学反应会导致一个各化学势之间的关系.

(41) 这结果, 还可以从别的角度来了解. 在平衡时, 化学反应 (38) 不能做功, 即

$$\mu_A dN_A + \mu_B dN_B + \mu_C dN_C = 0. \quad (42)$$

但根据 (38), $dN_A = dN_B = -dN_C$, 即得 (41). 这是一个热力学的结果, 并不需要由 (39) 得来.

由此可见, (30) 实在是一个十分理想的选择, 如果我们不用 (30), 而用

$$\mu/T = -\alpha + \eta, \quad (43)$$

η 为一常数, 则由 (40)、(41) 可见 η 不能是任何常数而必须适合

$$\eta_A + \eta_B = \eta_C. \quad (44)$$

不仅如此, 分子 A, B, C 及所有和它们有化学反应的分子, 都必须同时考虑. 每一个化学反应都有类似 (41) 的一个关系, 每个 η 都要适合这些关系, 当然, 要选择这样的 η 只有从守恒律着手. 所有化学反应, 核子反应所共同遵守的是核粒子 (质子或中子) 的总数的守恒, 因此, η_A 必须选为

$$\eta_A = \text{常数} \times (A \text{ 分子的核粒子数}). \quad (45)$$

这“常数”必须为所有分子共有. 如果我们要把熵的定义扩充到核粒子数都不一定守恒的情况下, 唯一的选择是把所有的 η 全选为零.

因此, 从细致平衡, 及量子力学中的状态概念, 我们得到 α, β 及分布函数 (23). 由状态的统计, 量子常数 h 就跑进了 α , (见 (28)). 再由守恒律的考虑, h 自然地进入了熵, 完成熵的定义.

4. 波色粒子的分布, 热辐射

波色粒子不受独占定律的限制. 因此它的反应率也不同. (14) 中的 $(1 - f_3)(1 - f_4)$ 是独占定律的结果, 不适用于波色粒子. 根据量子力学的计算, (13) 的反应率该是

$$f_1 f_2 (1 + f_3)(1 + f_4) R, \quad (46)$$

也就是, 把 (14) 的负号改成正号便是. 这结果并不是很明显, 须要费些周章才能得出, 在此我们把它当作已知. (46) 表示, 若 **3**、**4** 状态中已经有了粒子, 别的粒子会更喜欢去.

现在, 我们只要把 (14)–(18) 的负号改成正号, 结果就是波色粒子的式子. (32) 现在变成

$$f_{\mathbf{k}} = \frac{1}{e^{(\varepsilon_{\mathbf{k}} - \mu)/T} - 1}. \quad (47)$$

无论是波色子还是费米子, 在低密度的情况下, $f_{\mathbf{k}}$ 都大约是“玻尔兹曼分布”

$$f_{\mathbf{k}} \approx e^{(\mu - \varepsilon_{\mathbf{k}})/T}. \quad (48)$$

这可以称为“古典分布”. (47) 和 (32) 都是“量子分布”. 上一节的所有结果, 都可以推广到波色粒子.

光子是波色粒子, 平衡态中的光子气体曰“热辐射”. 光子不断地被容器壁射入容器内, 也不断地被器壁吸收, 如果器内还有其他分子, 则这些分子也可以吸光放光. 光子数因此不守恒, 故化学势为零. 光子分布是

$$f_{\mathbf{k}} = \frac{1}{e^{\omega_{\mathbf{k}}/T} - 1}. \quad (49)$$

式中 $\omega_{\mathbf{k}}$ 是频率. 也就是光子能量, ($\hbar = 1$), $\varepsilon_{\mathbf{k}} = \hbar\omega_{\mathbf{k}}$. 光子状态的能量分布可以用第 1 节的方法求出, 只要把 (6) 的 $\varepsilon_{\mathbf{p}} = \frac{p^2}{2m}$ 改成 $\varepsilon_{\mathbf{p}} = cp$ 就是, c 是光速, 并计入 2 偏振方向:

$$g(\omega) = \frac{2}{(2\pi)^3} \int d^3p \delta(\omega - cp) = \frac{\omega^2}{\pi^2 c^3}. \quad (50)$$

因此热辐射能量的频率分布为

$$\varepsilon(\omega) = g(\omega) \frac{\omega}{e^{\omega/T} - 1} = \frac{\omega^3 / \pi^2 c^3}{e^{\omega/T} - 1}. \quad (51)$$

这就是“普朗克”分布, 这个分布的发现是量子力学诞生的前奏. 它比 (46) 这类反应率的计算要早. 事实上, 爱因斯坦曾用它和细致平衡来推断光的吸收率和放射率的一些特性. 我们稍提一下:

令某类分子 A 可以吸收频率 ω 的光而变成激发态 A^* . A^* 的能量要比 A 高出 ω , 在平衡中, A 的粒子数 N 和 A^* 的粒子数 N^* 之比为,

$$\frac{N^*}{N} = e^{-\omega/T}. \quad (52)$$

此式可由细致平衡导出 (见问题 10). 光的吸收率是和 N 及 $\varepsilon(\omega)$ 成正比, 放射率是和 N^* 成正比. 它们必须相等, 因此

$$RN\varepsilon(\omega) = R'N^* \quad (53)$$

R' 是一个 A^* 的放射率, R 是一个 A 和一个 ω 能量光子相互作用时的吸收率, 从 (52) 和 (53) 得

$$\begin{aligned} R' &= R \frac{N}{N^*} \varepsilon(\omega) = R \frac{\omega^3}{\pi^2 c^3} \frac{e^{\omega/T}}{e^{\omega/T} - 1} \\ &= R \frac{\omega^3}{\pi^2 c^3} [1 + f(\omega)] \\ &= R \frac{\omega^3}{\pi^2 c^3} + R\varepsilon(\omega). \end{aligned} \quad (54)$$

这 $1 + f$ 自然和 (46) 的 $1 + f$ 一样.

(54) 指出:

- ① 即使 $\varepsilon(\omega) = 0$, 即在无光子的情况下, A^* 会自动放射而成 A , 放射率为

$$r = R \frac{\omega^3}{\pi^2 c^3}, \quad (55)$$

也就是说 A^* 的“半衰期”为 $\frac{\ln 2}{r}$.

- ② $f \neq 0$ 使 A^* 放射的更快, 是为“被刺激”的放射. 这是波色粒子的特性.
- ③ 吸收和放射率的大小由一个数字 R 决定. R 决定于分子结构及其和光子作用细节.

这些结果都是有关光和分子作用的. 从细致平衡加上普朗克分布而得到这些结果, 可以说是出乎意料, 详细的量子力学计算必须把 A , A^* 及光子都用量子分析, 才能得到 (54) 的结果, (及 R 的值).

5. 化学反应, 浓度定律

在第 3 节讲守恒定律时, 曾提到化学反应, 每一种化学反应, 减少了一个守恒律, 而导致一个化学势间的关系, 如 (41). 在此我们利用 (41) 这类关系, 来求出各分子浓度间的关系.

我们就从反应 (38) 开始, 假定分子 A, B, C 混成一个理想气体, 粒子数分布都循古典分布 (48). 每个分子的能量除了动能之外还有内在运动的能量, 如转动, 振动:

$$\begin{aligned}\varepsilon_{\mathbf{k}} &= \frac{p_{\mathbf{k}}^2}{2m} + \omega_{\mathbf{k}}, \\ f_{\mathbf{k}} &= e^{(\mu - \varepsilon_{\mathbf{k}})/T}.\end{aligned}\quad (56)$$

A 分子的总数是

$$\begin{aligned}N_A &= \sum_{\mathbf{k}} f_{\mathbf{k}A} \\ &= V e^{\mu_A/T} \frac{1}{(2\pi\hbar)^3} \int d^3\mathbf{p} e^{-p^2/2m_A T} \sum_{\alpha} e^{-\omega_{A\alpha}/T} \\ &\equiv V e^{\mu_A/T} e^{-\phi_A/T}.\end{aligned}\quad (57)$$

ϕ_A 是温度的函数, 和分子的质量 m_A 以及各内形态能量 $\omega_{A\alpha}$ 有关.

令 $n_A \equiv N_A/V$ 为 A 分子的粒子数密度或“浓度”, 则

$$\begin{aligned}\mu_A &= \phi_A + T \ln n_A, \\ \mu_B &= \phi_B + T \ln n_B, \\ \mu_C &= \phi_C + T \ln n_C.\end{aligned}\quad (58)$$

由 (41) 得

$$0 = \phi_A + \phi_B - \phi_C + T \ln \frac{n_A n_B}{n_C}$$

或

$$\frac{n_A n_B}{n_C} = e^{-(\phi_A + \phi_B - \phi_C)/T}.\quad (59)$$

这就是平衡浓度定理. (59) 的右边是可以由分子特性计算的, 而且只是温度的函数. 如果知道分子结构, 就可以计算有关浓度的资料. 这对化学工程师来说, 是十分的有用. 若是各分子浓度可以测得而分子结构并不清楚, 则 (59) 可以用来推断分子结构, 读者可参阅化学书籍, 以求对这方面的更深入了解.

讨论问题三

1. 分布函数

“分布”是最常见的统计资料,如“年龄分布”,“所得分布”,学生“考试成绩”等等.一个分布,可用一条曲线表示,图3是国民年龄分布. $f(x)$ (以万为单位)

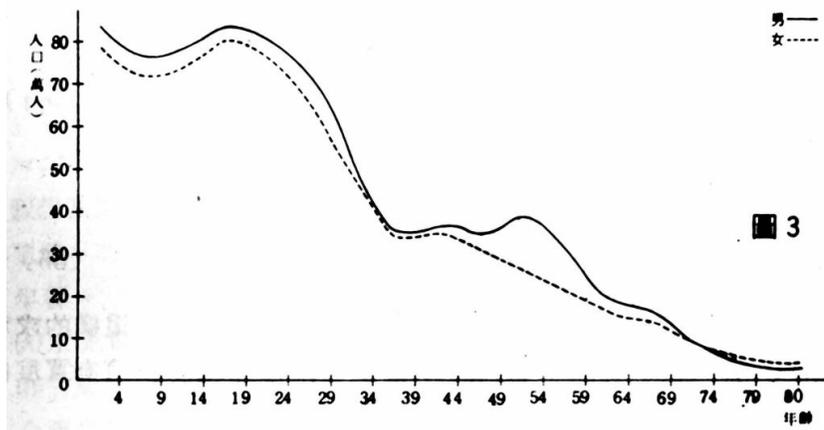


图 3

这曲线下的总面积是总人数:

$$\int_0^{\infty} f(x)dx = \text{总人数}. \quad (60)$$

年纪小于 x' 的人数 $N(x')$ 是 $x = x'$ 的左边曲线下的面积:

$$N(x') = \int_0^{x'} f(x)dx, \quad (61)$$

$$f(x) = \frac{dN(x)}{dx}. \quad (62)$$

这分布函数的意义是: $f(x)dx$ 是年龄在 x 和 $x + dx$ 之间的人数,要统计年纪小于 x' 的人数,我们可以去查全国名册,再计人数:

$$N(x') = \sum_s \theta(x' - x(s)). \quad (63)$$

s 是人名, $x(s)$ 是 s 先生或女士的年纪,他若比 x' 小, $\theta(x' - x(s))$ 就是 1, 否则就不计, (63) 表示这统计的过程, 年龄分布 $f(x')$ 形式上是 N 的微分 (见 (62)):

$$\begin{aligned} f(x') &= \frac{dN(x')}{dx'} \\ &= \sum_s \delta(x' - x(s)). \end{aligned} \quad (64)$$

在一小段年纪 $\Delta x'$ 内的人数是

$$f(x')\Delta x' = \sum_s \int_{x'}^{x'+\Delta x'} dx \delta(x - x(s)). \quad (65)$$

分布的定义需以 (65) 为准, 我们以后见到 (64) 这样的求和式, 就该立即想它的积分, 或者把 δ 函数想成一个稍有宽度的尖峰.

2. 试求自由粒子的状态的能量分布:

(A) 在 2 维空间.

(B) 在 1 维空间.

(C) 在 3 维空间的结果.

3. 用 (12) 式求状态能量分布, 假定

(A) $U(r) = \frac{1}{2}Kr^2,$

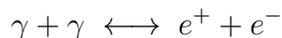
(B) $U(r) = -Ze^2/r.$

用 (B) 可以大致估计重原子 (大 Z) 的能量. 读者不妨试试.

4. 细致平衡来自时间方向的对称性. 这对称性是平衡态的一大特征. 有些状态, 虽然看来持久不变, 但没有对称性, 我们可称之为不平衡. 例如有热流通过的物体, 虽然热流维持不变, 但不算是平衡态. 但是作等速直线运动的物体, 或等速转动的物体, 是可以算作平衡. 读者仔细分析一下这些情况.

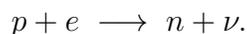
5. 第 2, 3 节指出熵和量子力学中的“状态”概念有密切的关系. 更指出各反应守恒量的重要. 读者试从 (2.36), (30) 求出 (31).

6. 某星球中有以下反应进行:



γ 为光子, e^\pm 为正电子及电子, 求平衡下的 e^+ , e^- 粒子数.

某高密度原子、电子、中子气体有以下反应进行



ν 为中微子, 产生后就跑掉了. 电子压 (即电子化学势) $\mu_e \gg mc^2$, $m =$ 电子质量, 求平衡的 p, e, n 粒子数 (中子量).

7. 假想某星球在时刻 $t = 0$ 时爆炸. 假设此星球为一质子与电子的混合气体. 有一探测器在距离 R 远处探测电子和质子的流量. $R \gg$ 星球半径.

(A) 求流量和时间的关系. 这关系和粒子的速度分布有关.

(B) 若爆炸后温度为 T , 求流量.

8. 波色粒子的理想气体可以用 (47) 分析. (A) 求 μ 和 N/V 的关系.

注意, 当温度低过某临界温度 T_c 时, $\mu = 0$, f_0 (即 $\mathbf{k} = 0$ 状态的粒子数) 变成和 V 成正比, 即有大量粒子占到最低状态去. 这现象一般称为“波色-爱因斯坦”凝聚现象.

(B) 求临界温度 T_c .

(C) 求 f_0 和温度的关系.

(D) 求热容率.

(E) 证明在一维、二维空间内, 此凝聚现象不存在.

9. (38) 式至 (45) 式的讨论并不限于费米粒子. 事实上 (38) 式的分子 C 必须是一波色粒子, (两个费米粒子只能合成一波色粒子).

试将第 3 节的讨论推广, 以包括波色粒子.

10. 从反应



导出 (52). 此结果和 A 的波色性无关.

第四章 金属电子

金属电子自然是固态物理中的一大部门. 在这一章我们挑出一点最简单的题目, 来作为上节理论的应用例子. 先把自由电子模型略为介绍, 再谈众所熟知的热容率. 对自由电子影响最大的是晶体的结构. 我们用一维空间的模型来略窥电子和晶体结构的相互影响. “派若氏的不稳定现象”(即一维空间的金属性为不稳定) 虽是专指一维空间而言, 但颇能指出一些重要的道理, 在此将予以仔细讨论.

1. 高密度理想气体

一般金属, 密度都很高, 原子排成晶体, 原子的内层电子都被核子紧紧吸住. 外围电子则游荡于各原子之间. 许多金属的特性都是由这些游荡的电子的运动而定. 现在我们把这些电子看作一理想费米气体. 是为“自由电子模型”. 这样简单的看法, 有意料不到的好成绩. 先看看一些这理想气体模型的结果:

我们用 \mathbf{k} 来指定各电子状态, $\varepsilon_{\mathbf{k}}$ 为 \mathbf{k} 状态的能量. 令 μ 为电子压, 即电子的化学势 (也就是日常所说的“电压”乘以电子电荷). 每个状态至多可以容一个电子. 令 $f_{\mathbf{k}}$ 为状态 \mathbf{k} 的平均粒子数, (见 (3.32)):

$$f_{\mathbf{k}} = f(\varepsilon_{\mathbf{k}} - \mu), \quad (1)$$

$$f(\varepsilon) \equiv \frac{1}{e^{\varepsilon/T} + 1}. \quad (2)$$

状态的能量分布 (态密度) 是个重要的数据. 令

$$g(\varepsilon) \equiv \frac{1}{V} \sum_{\mathbf{k}} \delta(\varepsilon - (\varepsilon_{\mathbf{k}} - \mu_0)),$$
$$\mu_0 \equiv \mu(T = 0). \quad (3)$$

$g(\varepsilon)$ 为单位容积的状态能量分布, 可简称为“能位密度”. ε 是以 μ_0 为准的能量尺度, $g(\varepsilon)$ 的计算, 在第三章第 1 节会有详细介绍.

$$\begin{aligned} g(\varepsilon) &= 2 \int \frac{d^3 \mathbf{p}}{(2\pi)^3} \delta\left(\varepsilon + \mu_0 - \frac{p^2}{2m}\right) \\ &= \frac{m}{\pi^2} [2m(\varepsilon + \mu_0)]^{1/2}. \end{aligned} \quad (4)$$

这模型的热力性质, 多可以用 f 与 g 表示. 例如电子密度 N/V 及能量密度 E/V 为

$$\begin{aligned} \frac{N}{V} &= \int d\varepsilon g(\varepsilon) f(\varepsilon - \mu), \\ \frac{E}{V} &= \int d\varepsilon g(\varepsilon) f(\varepsilon - \mu) (\varepsilon + \mu_0). \end{aligned} \quad (5)$$

如果温度 $T = 0$, 则能量较低之状态全被填满, 满到所有状态都有了电子. 这最后的一批状态是所谓“费米”面. 此层面以上的状态全空. 这费米面可以称作“顶层面”. 这层面的能量为 μ_0 , 即 $T = 0$ 时的化学势, 常以 ε_F 表示

$$\varepsilon_F \equiv \mu(T = 0) = \mu_0. \quad (6)$$

费米面可想作电子动量空间中的一个球面. 球半径为 p_F , 为顶层动量, 或“费米”动量.

$$\frac{p_F^2}{2m} = \varepsilon_F. \quad (7)$$

能量比 ε_F 小的状态共有 N 个

$$\frac{4\pi}{3} p_F^3 \times V \times \frac{2}{(2\pi)^3} = N, \quad (8)$$

即 (球体积) \times (体积) \times (自旋形态数 (= 2)) $(2\pi\hbar)^3$, $\hbar = 1$. 因此

$$p_F = (3\pi^2 n)^{1/3}. \quad (9)$$

$n \equiv \frac{N}{V}$, $a \equiv \hbar/p_F$ 约是电子间的平均距离,

$$\varepsilon_F \sim \frac{\hbar^2}{ma^2} \quad (10)$$

约为每电子的零点能量. 这零点能量是电子波动性的结果, 亦可说是“测不准原理”的必然结果. 要把电子塞到大小为 a 的空间, 则必有 \hbar/a 的动量.

表 1: 顶层电子的性质

	电子密度 $N/V, \text{cm}^{-3}$	动量 k_F, cm^{-1}	速度 $v_F, \text{cm/sec}$	能量 ε_F, eV	$T_F = \varepsilon_F/k$ K
Li	4.6×10^{22}	1.1×10^8	1.3×10^8	4.7	5.5×10^4
Na	2.5	0.90	1.1	3.1	3.7
K	1.34	0.73	0.85	2.1	2.4
Rb	1.08	0.68	0.79	1.8	2.1
Cs	0.86	0.63	0.73	1.5	1.8
Cu	8.50	1.35	1.56	7.0	8.2
Ag	5.76	1.19	1.38	5.5	6.4
Au	5.90	1.20	1.39	5.5	6.4

取自 Kittel(1966), p.208

一般金属电子间距离约为 10^{-8}cm , $\varepsilon_F \sim 1\text{eV} \sim 10^4\text{K}$, (见表 1). 因此在一般的温度下, $T \ll \varepsilon_F$, 即温度“很低”. 只有在费米面附近的电子有机会跳到外层, 大多数的电子都停留在层下, 不能做任何改变.

$$\begin{aligned}\varepsilon_F &= \frac{\hbar^2}{2m} (3\pi^2 N/V)^{2/3}, \\ k_F &= (3\pi^2 N/V)^{1/3} = p_F/\hbar, \\ v_F &= p_F/m, \\ T_F &= \varepsilon_F/k.\end{aligned}$$

注意, ε_F 高的原因是电子的质量很小. 密度高也是一原因. 这电子气体模型和 H_2 , O_2 之类的稀疏、常见分子气体是完全不同的. 在 $T \rightarrow 0$ 时, 这电子气体的总能量是

$$\begin{aligned}E(0) &= \frac{3}{5} N \varepsilon_F \\ &= \frac{3N}{10m} \left(\frac{3\pi^2 N}{V} \right)^{2/3} \propto V^{-2/3}.\end{aligned}\quad (11)$$

产生的压强, 是为“零点压强”:

$$p(0) = - \left(\frac{\partial E}{\partial V} \right)_N = \frac{2}{3} \frac{E}{V} = \frac{2}{5} \frac{N}{V} \varepsilon_F. \quad (12)$$

如果和理想气体的 $p = \frac{N}{V}T$ 比较, 我们发现这 $P(0)$ 相当于温度 $2\varepsilon_F/5$, 且 N/V 比一般气体密度大很多. 因此它是一个非常大的压强, 这压强把一块金属像吹气球似的往外胀. 全靠原子间的内聚力, 使它不炸开, 这内聚力的来源自然是电子的负电和离子的正电, 内聚力的计算颇是一门学问, 读者可参考固态物理的书.

2. 热容率

受到温度影响的只是一小部分电子. 因此空状态或“空子”的概念就应需要而生. 顶层下的状态几乎全满, 空状态少, 我们称一空状态为一“空子”. 现在这模型成了一个空子和外层电子的新混合气体. (费米面以上称为“外层”, 以下称为“内”.) 空子的动量小于 p_F , 外电子的大于 p_F . T 越低, 则此新气体越稀疏. 在 $T = 0$ 时, 这气体就消失了. 这是一非常方便的想法.

令 $T = 0$ 时的总能量为零, 即无空子亦无外电子的状态. 空子也是费米的, 因为每状态至多有一空子. 因此, 能量为 $-\varepsilon'$ 的状态, 可能产生能量为 ε' 的空子. 空子的平均粒子数为 (每状态):

$$1 - f(-\varepsilon' - \mu) = f(\varepsilon' + \mu). \quad (13)$$

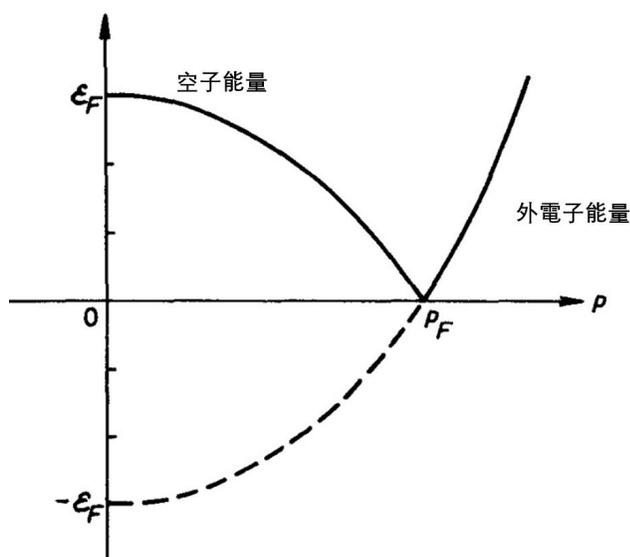


图 1

现在能量的原点是顶层面的能量, 即 $\mu_0 = 0$. 空子的能量是把一电子从内层提到顶层所需的能量 (见图 1). 既然 $T \ll \varepsilon_F$, 空子的能量和外电子的能量都不大可能太大于 T . 在这段能量范围内, $g(\varepsilon)$ 大致不变, 即

$$g(\varepsilon) \approx g(0) = mp_F/\pi^2. \quad (14)$$

如此, 空子、外电子的能量分布一样, 因此 $\mu \approx 0$, 各种计算也可以简化不少. 例如, 总能量是

$$\frac{1}{V}[E(T) - E(0)] = \int_0^\infty d\varepsilon 2g(0)f(\varepsilon)\varepsilon. \quad (15)$$

$2g(0)$ 是空子加外电子的能位密度. 空子的能量不会超过 ε_F . 但 $\varepsilon_F \gg T$, 所以 (15) 的积分上限可作 ∞ . 这积分很容易:

$$\int_0^\infty d\varepsilon f(\varepsilon)\varepsilon = T^2 \int_0^\infty dx \frac{x}{e^x + 1} = \frac{\pi^2}{12} T^2. \quad (16)$$

代入 (15), 微分一次得热容率:

$$\frac{1}{V}C = \frac{1}{V} \frac{\partial E}{\partial T} = \frac{\pi^2}{3} g(0)T. \quad (17)$$

这结果和稀疏的理想气体的 $C = \frac{3}{2}N$ 大不相同. 在那情形下, 每个分子都供给热容率 $\frac{3}{2}$. 在此情况下, 只有小部分电子参与运动, 参与的电子数约为 NT/ε_F ,

$$Vg(0) \int_0^\infty d\varepsilon f(\varepsilon) \sim NT/\varepsilon_F, \quad (18)$$

每参与电子供给的 C 约为 1, 故 $C \sim N(T/\varepsilon_F)$. 由 (17) 得

$$C = N \frac{\pi^2}{2} (T/\varepsilon_F). \quad (19)$$

研究金属的初期, 常把电子当成一稀疏理想气体, 但实验发现 C 比 $\frac{3}{2}N$ 小太多. 直到费米分布被发现, 这热容率问题才有满意的解释.

(19) 只是 T 的一次项. 下一项是什么? 如果 $g(\varepsilon) = g(0)$, 就不会有下一项. 因此, 我们把 $g(\varepsilon)$ 展开到 $O(\varepsilon)$, 来求下一项:

$$\begin{aligned} g_+(\varepsilon) &\approx g(0) + \varepsilon g'(0), \\ g_-(\varepsilon) &\approx g(0) - \varepsilon g'(0). \end{aligned} \quad (20)$$

g_+ 是外层电子的能位密度, g_- 是空子的. g_+ 既不等于 g_- , $T \neq 0$ 时的化学势亦非零, 外电子和空子的每状态粒子数为

$$\begin{aligned} & f(\varepsilon - \mu), \\ \text{及} \quad & 1 - f(-\varepsilon - \mu) = f(\varepsilon + \mu), \end{aligned} \quad (21)$$

μ 很容易求出, 因为空子数和外电子数相等:

$$\int_0^\infty d\varepsilon g_+(\varepsilon) f(\varepsilon - \mu) = \int_0^\infty d\varepsilon g_-(\varepsilon) f(\varepsilon + \mu). \quad (22)$$

μ 和 $g'(0)$ 均为小值. 将 (22) 展开, 解 μ 得

$$\mu = -\frac{T^2 \pi^2 g'(0)}{6g(0)}. \quad (23)$$

因为 $g'(0)/g(0) = 1/2\varepsilon_F$, 所以 $\mu \sim -T^2/\varepsilon_F$. 除了 μ 之外, $g'(0)$ 对能量 (15) 还可能有直接的修正, (15) 该是

$$\frac{1}{V}[E(T) - E(0)] = \int_0^\infty d\varepsilon g_+(\varepsilon) f(\varepsilon - \mu) \varepsilon + \int_0^\infty d\varepsilon g_-(\varepsilon) f(\varepsilon + \mu) \varepsilon.$$

如果展开, 则 $g'(0)$ 各项全消去了. 因此, (17) 或 (19) 的结果是不含 T^2 项的. 所以

$$C = N \frac{\pi^2 T}{2 \varepsilon_F} (1 + O(T^2/\varepsilon_F^2)), \quad (24)$$

修正项来自 $g''(0)$, 及 $g'(0)^2$.

除了电子热容率之外, 金属的原子振动也供给热容率. 在室温以上约为 $3N$, 大占优势. 金属的底拜温度 θ_D 多在 100 K 至 400 K 左右. 晶体振动的热容率在 $T > \theta_D$ 时约为 $3N$ (这些待第六章再详谈, 此地且先引用结果.) 在温度很低时, 即 $T \ll \theta_D$, 则为 $N(T/\theta_D)^3$. 因为 $\varepsilon_F \gg \theta_D$, 所以这 T^3 项比 (24) 的修正项要大很多, 因此, 在低温时, 总热容率为

$$C/N = \frac{\pi^2}{2} \left(\frac{T}{\varepsilon_F} \right) + \frac{12\pi^4}{5} \left(\frac{T}{\theta_D} \right)^3. \quad (25)$$

读者请注意: (20) 至 (24) 的分析非常重要. 它指出电子热容率中没有 T^2 这一项, 只有 T 和 T^3 . 否则 (25) 就会多一项 T^2 .

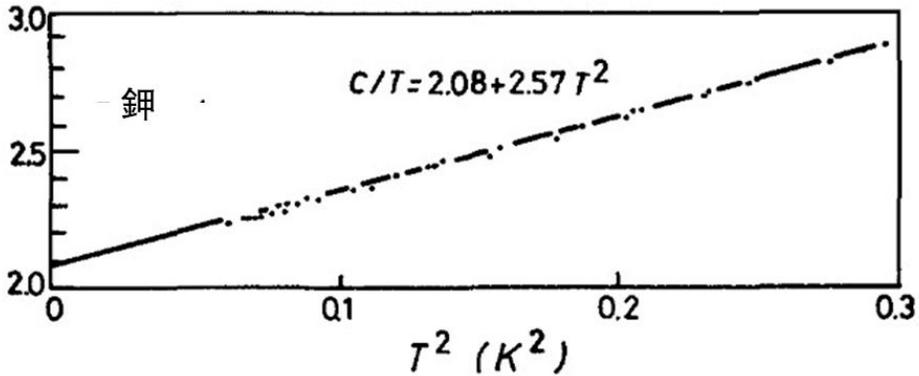


图 2. 钾的低温热容率 C/T (10^4 erg/ K^2 /克分子). 取自 W. H. Lien and N. E. Phillips, Phys. Rev. 133A 1370(1964).

图 2 是 $(C/T, T^2)$ 的图, 根据 (25) 该是一直线, 这实验是对自由电子模型的一次重要的考验. 从实验结果可以测定 (25) 中 T 和 T^3 的系数. T 的系数是

$$r \equiv \frac{\pi^2 m}{p_F^2} = \frac{\pi^2 k^2 m}{\hbar^2 (3\pi^2 n)^{2/3}}. \quad (26)$$

我们把 k, \hbar 放进去了. n 是电子密度. 实验求得的 r 比 (26) 的值要大些. 原因并不很清楚. 当然, 自由电子模型实在过分简略, 能做到这么多, 已是预料不到的. 电子和杂质, 和晶体振动的作用, 加上晶体的周期性结构, 都尚未考虑进去.

一般所谓的金属性, 就是类似自由电子的特性, 也可以说是费米面附近 $g(\varepsilon)$ 大致不变的后果. 如果 $g(\varepsilon)$ 受到影响, 则金属性亦受影响.

3. 周期性结构的影响

以上的讨论指出能位密度 $g(\varepsilon)$ 的重要性. 固体中有很多因素会影响 $g(\varepsilon)$. 晶体的结构自然是一主要因素. 以下用一个最简单的一维空间模型来谈一下周期性原子排列的影响.

电子状态的能量 ε 必须从波动方程式解出

$$-\frac{1}{2m} \frac{d^2}{dx^2} \psi(x) + U(x)\psi(x) = \varepsilon\psi(x). \quad (27)$$

$U(x)$ 是一周期函数

$$U(x) = U(x + a) \quad (a \text{ 为晶格常数}), \quad (28)$$

代表原子排列对电子的影响. 为求简单, 假设 U 十分微弱, 并假设

$$\int_0^L U(x) dx = 0, \quad (29)$$

L 为物体总长. 如果略去 U , 则 (27) 的解自然简单:

$$\begin{aligned} \psi_p &= \frac{1}{\sqrt{L}} e^{ipx}, \\ \varepsilon_p &= \frac{p^2}{2m}. \end{aligned} \quad (30)$$

现在来计算 U 对 ε_p 的修正, 量子力学的展开法告诉我们 ε_p 的一次修正为:

$$\langle \psi | U \psi \rangle = \frac{1}{L} \int_0^L U(x) dx = 0. \quad (31)$$

不过这答案不完全合用, 因为此式只合用于不接近重迭的能量. 在 (30), 有重迭发生, 即 ψ_p 和 ψ_{-p} 的能量重迭, 都是 $p^2/2m$. $U(x)$ 可写成

$$\begin{aligned} U(x) &= \sum_{n=-\infty}^{\infty} U_n e^{inqx}, \\ q &\equiv 2\pi/a. \end{aligned} \quad (32)$$

$U(x)$ 在 $|p\rangle$ 和 $|p + nq\rangle$ 之间有作用:

$$\langle p | U | p' \rangle = \sum_{n=-\infty}^{\infty} U_n \delta_{p, p+nq}, \quad (33)$$

因此, 如果

$$p \approx -p + nq,$$

即

$$p \approx \frac{n}{2}q, \quad n = \pm 1, \pm 2, \dots \quad (34)$$

则 (31) 的结果无效. 现在看看 $n = \pm 1$ 的情形, 令

$$\begin{aligned} p_1 &\equiv \frac{1}{2}q + k, \\ p_2 &\equiv -\frac{1}{2}q + k, \quad k \ll q. \end{aligned} \quad (35)$$

量子力学告诉我们要把这两个形态一同考虑, 才能求得正确的一次修正项. 能量算符在这两态的空间内是以方阵 H 表示:

$$H = \begin{pmatrix} \varepsilon_1 & H_{12}^* \\ H_{12} & \varepsilon_2 \end{pmatrix}, \quad (36)$$

$$H_{12} = U_1,$$

$$\begin{aligned} \varepsilon_1 &= \frac{1}{2m} \left(\frac{q^2}{4} + k^2 + qk \right), \\ \varepsilon_2 &= \frac{1}{2m} \left(\frac{q^2}{4} + k^2 - qk \right). \end{aligned} \quad (37)$$

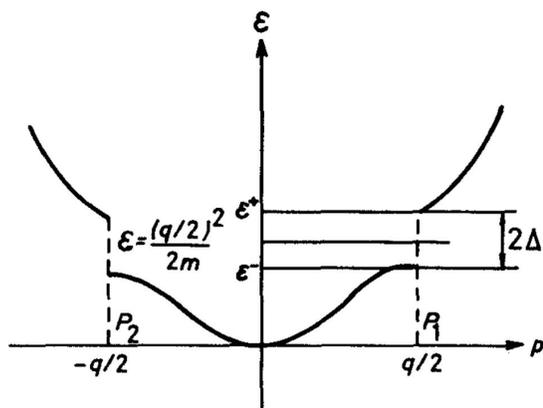


图 3

H 的本征值是

$$\begin{aligned} \varepsilon_{\pm} &= \frac{1}{2m} \left(\frac{q^2}{4} + k^2 \right) \pm \sqrt{\frac{q^2 k^2}{4m^2} + \Delta^2}, \\ \Delta &\equiv |U_1|. \end{aligned} \quad (38)$$

图 3 画出这结果, 在 $k = 0$, 即 $p = \pm q/2$ 时, $\varepsilon_+ - \varepsilon_- = 2\Delta$, $n = \pm 2, \pm 3, \dots$ 等的情况亦类似. 即能位密度在能量为

$$\frac{1}{2m} \left(\frac{n\pi}{a} \right)^2, \quad n = \pm 1, \pm 2, \pm 3, \dots, \quad (39)$$

处, 发生一小段空白 (见图 4). 这是周期性势能的最重要的效果.

令电子的总数为 N . 如果

$$N = L/a, \quad (40)$$

即平均每格一电子, 则 p_F 是什么? 因为每个 p 有两状态, (两自旋方向), 所以

$$\begin{aligned} N &= \frac{2L}{2\pi} \int_{-p_F}^{p_F} dp = \frac{2p_FL}{\pi}, \\ \therefore p_F &= \frac{\pi}{2a} = \frac{q}{4}, \end{aligned} \quad (41)$$

即“顶层面”是 $p = p_F$ 及 $p = -p_F$ 两点, 因为空白在 $p = \pm q/2$, 距 p_F 尚远, 所以只要 $T/\varepsilon_F \ll 1$, 上节的一些金属性质都会有.

如果每格子有两个电子, 则

$$p_F = \frac{\pi}{a} = \frac{q}{2}, \quad (42)$$

$\frac{q}{2}$ 和 $-\frac{q}{2}$ 之间的状态就全占满了, 电子如要激发, 必先跳上一层楼, 必须要有 2Δ 以上的能量. 我们可以略算一下 $g(\varepsilon)$ 及热容率.

我们把能量原点定在空白的正当中, 得

$$\begin{aligned} g(\varepsilon) &= \frac{2}{2\pi} \int dp \delta(\varepsilon - \varepsilon_F) \\ &\approx \left(\frac{2ma}{\pi^2} \right) \frac{|\varepsilon|}{\sqrt{\varepsilon^2 - \Delta^2}}, \quad \varepsilon^2 > \Delta^2. \end{aligned} \quad (43)$$

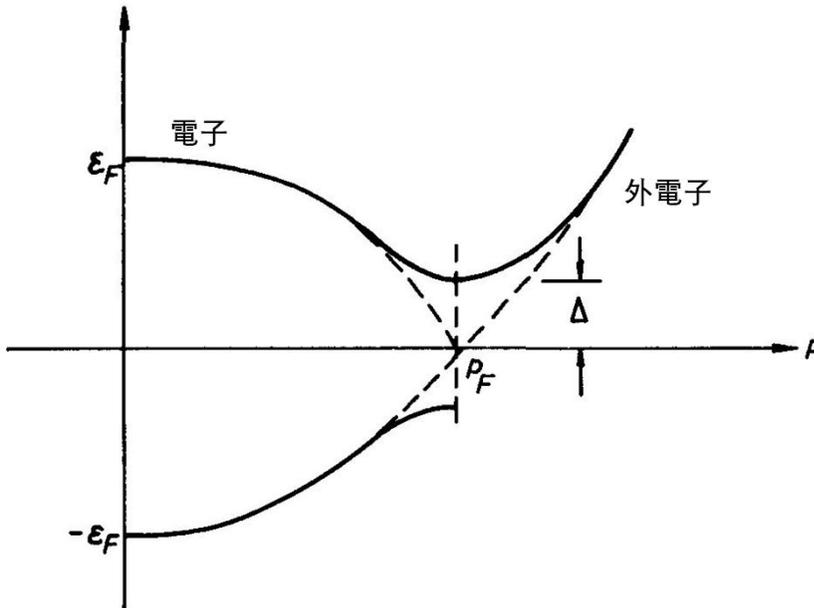


图 5

以上假设 $\varepsilon \ll \varepsilon_F$. 我们再用上节的空子想法来求出总能量, 空子为内层的空状态, 在空白之下. (见图 5) 空子和外电子数目相等且 (43) 的 $g(\varepsilon) \approx g(-\varepsilon)$. 所以 $\mu = 0$, 总能量和热容率是

$$\begin{aligned} \frac{1}{N}[E(T) - E(0)] &= \frac{L}{N} \int_0^{\infty} d\varepsilon 2g(\varepsilon) \frac{1}{e^{\varepsilon/T} + 1} \\ &\approx 4ma^2\pi^{-3/2}\Delta^{3/2}\sqrt{2T}e^{-\Delta/T}, \\ \frac{1}{N}C &\approx 4\sqrt{2}ma^2\pi^{-3/2}\Delta^{5/2}T^{-3/2}e^{-\Delta/T}. \end{aligned} \quad (44)$$

以上假设 $T \ll \Delta$.

这个结果自然和 (19) 的情形完全不同. 这能量空白使内层电子难以出去. $e^{-\Delta/T}$ 是个最重要的因子. 空子和外电子的粒子数和物体的导电性有直接的关系, 如果 Δ 不很大, 这粒子数可以用温度来作有效的控制. 这物体成了一种半导体. 如果 Δ 很大, 这粒子数很小, 即成了绝缘体.

4. 电子引起的结构改变

晶体结构, 造成了电子能量的一段空白. 上节例子中, 如 $N = \frac{2L}{a}$, 则这空白的的影响非常之大. 我们已看到 Δ 对 C 的影响, 现在来看看它对零点能量的

影响.

$$\begin{aligned} \frac{1}{N}[E(\Delta) - E(\Delta = 0)] &\approx a \int_{-\varepsilon_F}^{-\Delta} d\varepsilon (g(\varepsilon) - g_0(\varepsilon)) \varepsilon \\ &\approx -\pi^{-2} m a^2 \Delta^2 \ln(\varepsilon_F/\Delta) < 0, \\ \varepsilon_F &= \frac{p_F^2}{2m} = \frac{\pi^2}{2ma^2}. \end{aligned} \quad (45)$$

这结果是用近似式 (43) 来计算. 更精确的计算结果, 也差不多. 此地的主要结论是 $E(0)$ 因这空白而降低了.

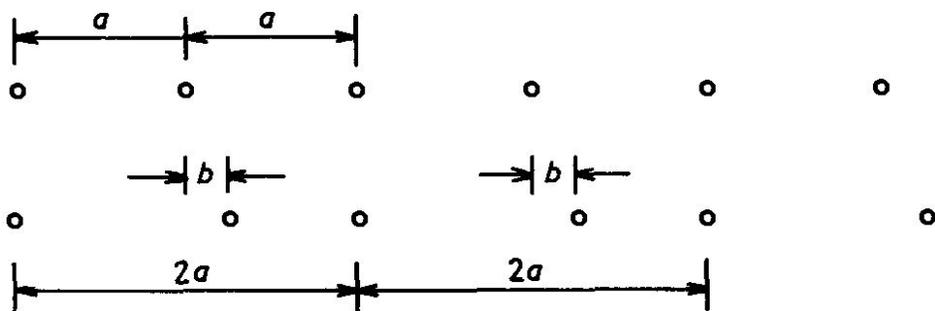


图 6

现在我们回到 $N = L/a$ 的情形, 即 $p_F = \frac{1}{4}q$. 在这情形, ε_F 距空白很远. 我们提过, 这情形下的结果和没有 $U(x)$ 的情形差不多, 但是, 晶体的原子是可以移动的, 如果每隔一个原子移动 b (如图 6), 则周期由 a 变成了 $2a$. 如果 b 不很大, 则晶体的能量的增加和 b^2 成正比, 因为 $b = 0$ 是平衡位置:

$$\frac{1}{N}[E(b) - E(b = 0)] = \frac{1}{2}Kb^2, \quad (46)$$

K 为弹性系数. 但移动后, 电子所见的 $U(x)$ 势能也改了. 新的势能的周期是 $2a$, 所以

$$\begin{aligned} U(x, b) &= U(x, b = 0) + b \sum_{n=-\infty}^{\infty} e^{inq'x} U'_n + O(b^2), \\ q' &= \frac{2\pi}{2a} = \frac{q}{2}. \end{aligned} \quad (47)$$

因此, 我们有了新的 bU'_1 ,

$$|bU'_1| = \Delta'. \quad (48)$$

这个 bU_1' 和以上的 U_1 的效果一样, 生出一段空白, 只是空白发生在

$$p = \pm \frac{q'}{2} = \pm \frac{q}{4}, \quad (49)$$

也就是说, 发生在现在的 $\pm p_F$. 这新的空白会把电子能量减低, 减低的量可以把 Δ' 代 (45) 的 Δ , $2a$ 代 a , 而求得

$$\begin{aligned} \frac{1}{N}[E(b) - E(b=0)] = & -4ma^2\pi^{-2}b^2|U_1'|^2 \ln\left(\frac{\varepsilon_F}{b|U_1'|}\right) \\ & + \frac{1}{2}Kb^2. \end{aligned} \quad (50)$$

最后一项为 (46) 的结果. 如果 b^2 很小, 则 $\ln(1/b)$ 就很大, 所以, 只要 b 够小, 这能量的改变就会是负的. $E(b)$ 极小值可以由 $\frac{\partial E(b)}{\partial b} = 0$ 求出, b 的平衡值为

$$b \sim \frac{\varepsilon_F}{|U_1'|} e^{-1/\lambda}, \quad \lambda = 8\pi^{-2}ma^2|U_1'|^2/K. \quad (51)$$

结论是, 周期为 a 的结构是不稳定的. 由于电子和晶体的作用, 周期为 $2a$ 的结构能量较低, 且有能量空白出现于 $p = \pm p_F$. 原先的金属性质因此失去, 这是派若氏指出的不稳定现象*.

以上的分析的最重要结论是: 晶体的结构和电子的运动是互为因果, 必须同时考虑.

这分析虽然十分简单, 但却明白地指出: 一维空间的晶体不可以有金属性. 也就是说, 在 $p = \pm p_F$ 一定会有能量空白出现. 注意: 以上的计算是用 $p_F = q/4, q/2$, 读者不难看出, p_F 可以是任何值, 即使 p_F 和 q 为不可通分的, 以上的结论都不变: 晶体的结构周期必是 $2\pi/(2p_F)$.

当然, 这模型没有考虑晶体的振动, 一维空间的振动有很大的振幅. (见第十三章末节). 再者, 电子间的作用亦未考虑, 因此, 以上的这个结论仍颇有商榷的余地. 实验室中的“一维空间”物体, 必是某物体的一部分. 例如某些晶体的电子特别容易在某方向 x 运动而不容易朝 y, z 方向运动, 这种物体可想成由一束一维空间晶体组成. 不过此类物体比以上模型复杂的很多, 以上的结果只可供参考, 不能立即直接应用, 无论在理论上, 或在实验上, 一维空间的电子运动的许多问题, 还尚待解决.

*Peierls (1955)

讨论问题四

1. 本章计算成分较大, 读者须导出每一结果. 尤其是最后两节, 如不自己做一遍, 就完全不能体会其中含义.

2. 试证低温的电子压强是

$$p(T) = \frac{2E}{3V} = \frac{2N}{5V} E_F \left[1 + \frac{5\pi^2}{12} \left(\frac{T}{\varepsilon_F} \right)^2 + \dots \right].$$

这是低温的物态方程式.

3. 高密度电子气体也在天文学中出现, 例如“白暗星”(或“白矮星”)的理論. 大致说来, 此类星为一团高温高压的氦核子及电子.

求电子压强和星球半径的关系. 电子速度接近光速, 须用相对论分析.

这题目在黄氏书中有详细讨论. (Huang (1963), P.230).

4. 令 ϕ 为金属电子化学势 μ 和金属外界的势能差, (见图 7). 假定温度 T 相当低, $T \ll \phi, \varepsilon_F$.

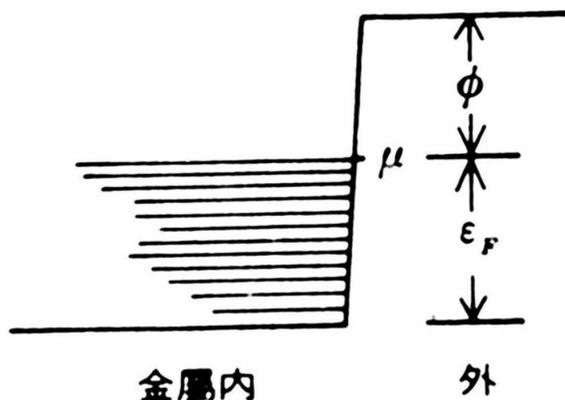


图 7

用光照射金属表面, 可以打出电子来, 是为“光电现象”. 令 I 为每光子产生的跳出电子.

(A) 在 $T = 0$ 时, I 和光的频率 ω 有何关系?

(B) 在 $T \neq 0$, 求 $I(\omega)$, 注意, 在 $T \neq 0$ 时, 即使没有光射入, 也会有电子跳出, 是为热电子辐射现象.

5. 假如金属中有条裂缝, 则缝中电子密度为何?

注意, 图 7 是个极其简化的金属表面模型, 和真正金属表面相去甚远.

6. 第 4 节谈的是一维空间模型. 二、三维空间的情形复杂的多, 但原则上是类似的. 晶体和电子的作用产生新的周期性的密度. 可称“密度波”, 亦称“电荷密度波”.

第二篇

假设

统计力学是建立在一个大胆的假设上, 即玻尔兹曼的求熵公式

$$S = \ln \Gamma,$$

Γ 是“活动范围”的体积. 这活动范围代表物体中分子整体状态的变化范围, Γ 度量各运动幅度之积.

本篇分五章. 先把这假设的动机和内容作一简介, 再用些例子说明它的使用法, (第五、六章). 第七、八章是进一步的推演及更多例子, 讨论重点在假设的含义, 及平衡观测时间, 反应速率和活动范围的关系. 第九章讨论“宏观极限”或“热力学极限”的数学问题. 我们用“杨、李定理”之一作为讨论中心.

本篇目的在于把统计力学的基本假设的功用及它的含糊之处同时指出, 使读者大致知道如何用它, 而不致盲目地被它牵着走.

所有的统计力学结果, 都可以从这假设及所谓“大数定律”导出, 不需要其他新概念. 读者须注意, 这假设并没有被证明过, 我们对它的了解, 并不完全.

第五章 分子运动幅度和熵, 基本假设

在热力学中, 熵是对小尺度分子运动的一种度量. 在上一章我们强调这一点, 并把熵看作一个“坐标”, 和体积、粒子数、磁矩等其他力学、电磁学坐标一样待遇. 到底熵是怎样度量分子运动? 统计力学对这问题提供了答案. 这是它最重要的成就. 这个答案并不是从什么基本原则一步步推导出来, 而是一个大胆的假设. 直到今天, 它的由来和定义仍不是完全清楚. 本章讨论这个假设所必须涉及的各种概念和计算问题, 包括假设的动机, 状态的统计, 大数定律, 高维空间的特性等.

1. 运动幅度

熵和分子运动的关系在第二章讨论理想气体时, 已经露了一点线索. 我们再看 (2.36)

$$S = N \ln \left(\left(\frac{CE}{N} \right)^{3/2} \frac{V}{N} \right), \quad (1)$$

N 是分子总数, $\frac{E}{N}$ 是每分子的平均能量, $\frac{V}{N}$ 是每分子平均所占的体积, 令

$$\begin{aligned} \frac{(\Delta p)^2}{2m} &\equiv \frac{E}{N}, \\ (\Delta x)^3 &\equiv \frac{V}{N}. \end{aligned} \quad (2)$$

m 是每分子的质量. 由于不断地碰撞, 各分子的动量和位置, 都不断地改变. 大致说来, Δp 是各分子动量改变的幅度, Δx 是各分子位置改变的幅度. 读者大概会问: 每个分子都可以到容器任一点, $(\Delta x)^3$ 是 V , 为什么用 $\frac{V}{N}$? 原因是: 每个分子大致跑不出 $\frac{V}{N}$ 这个范围: 跑出去了, 另一个跑进来, 就和没跑出去, 另一个没跑进来一样.

现在可以把 (1) 写成

$$S = N \ln \left(\frac{\Delta x \Delta p}{w} \right)^3 = \ln \Gamma, \quad (3)$$

$$\Gamma \equiv \left(\frac{\Delta x \Delta p}{w} \right)^{3N}.$$

$w = \frac{2m}{C}$ 是一常数. Γ 是一个 $6N$ 维空间区域的体积 (以 w^{3N} 为单位). 这区域有 $3N$ 个方向的长度为 Δx , $3N$ 个为 Δp .

这个区域可以说是代表这 N 个分子活动的程度, 它的每一方向的长度都代表一个运动变量的改变幅度, 共有 $6N$ 个变量, 即 $3N$ 个位置坐标及 $3N$ 个动量, $(x_k, y_k, z_k, p_{xk}, p_{yk}, p_{zk}, k = 1, 2, \dots, N)$.

所以, 熵和运动幅度的关系, 在这理想气体的例子中, 已是非常的明显了.

2. 计算熵的假设

当然, 理想气体是一个十分简单的特殊情形. 它的熵在上一节得来十分容易, 一般物体就不是这样简单了. 不过理想气体的结果, 已明确地指出, 了解熵的意义, 须从分子运动幅度方面着手. 现在我们把 (3) 推广.

在任一时刻, N 个分子的位置和动量可由上述 $6N$ 维空间的一点代表. 这一点叫做这群体在该时刻的“形象”, 或“状态”, 这空间叫做“状态空间”或“相空间”. 状态随时间的改变, 在这空间内划出一道“轨迹”来. 所以, 轨迹是真正出现的状态的集合. 在一段相当长的观测时间内, 这轨迹就成了一条很长的曲线, 蜿蜒曲折地散布在状态空间的一个区域里. 这个区域在 $6N$ 个方向的大小, 就是 $6N$ 个运动变量的改变幅度. 我们称这区域为“轨迹分布范围”或“状态活动范围”. 令 Γ 为此区域的体积, 并假设熵为 Γ 的对数

$$S = \ln \Gamma. \quad (4)$$

这是统计力学的基本假设*. 它不是从什么已知的定律推导出来. 在此我们把它看作 (3) 的建议. 到底 (4) 中的 S 是否就是热力学的熵, 则要由它在热力学上的应用来证实. 以上 Γ 的定义并不完全, 我们只取了些名字, 说了些话, 到底这活动范围怎么从分子运动求出? 这问题非常困难, 不过我们暂时避开它, 以后再谈. 现在只用以下假设来搪塞一下.

*这假设是玻尔兹曼先生 (Boltzmann 1844–1906) 的不朽之作.

基本假设的补充: 形象活动的范围, 等于观测时间内的不变量所允许的范围.

这确实是一个大胆的假设, 但并非难以捉摸. 我们看一个简单的例子, 就可以体会到这假设的含义. 一只在房间里飞来飞去的苍蝇有十分复杂的轨迹, 它的位置不停地改变. 如果观测时间比苍蝇从房间一边飞到另一边的时间长很多, 我们可以问: 苍蝇位置的改变幅度是什么? 苍蝇的位置是由三个坐标 (x, y, z) 代表. 我们用不着去分析轨迹, 就可以说 x, y, z 的改变幅度是由地板、天花板和四壁决定, 这些是“不变量”. 活动范围就是这房间, 房间的体积是各改变幅度的乘积.

我们现在要把 (x, y, z) 空间推广到 $6N$ 个运动变量所定义的状态空间, 这些墙壁、天花板的位置推广到各种不变量. 无论分子运动多复杂, 状态空间里的活动范围应该可以由这些不变量决定, 这就是这假设的含义.

有一点要特别强调: 必须要有运动, 有改变, 才能有活动范围, 这是“活动”二字的用意. 苍蝇如不飞, 则房间不能代表它的活动范围. 各分子必须运动, 不但运动, 还要不断地碰撞, 以改变其动量. 这假设的成立, 全仗着不断地变动.

对一个气体来说, 不变量只有总能量、体积、和分子总数. 固体的情形比较复杂, 除了能量, 总分子数之外, 可能还有各种结构上的不变量, 如冻结住了的杂质原子的位置等. 凡是不变量容许的状态, 一概包括在范围内. 这个补充假设非常了不起, 它把一个复杂的轨迹问题, 换成一个比较简单的范围计算问题. 当然这假设是否正确, 要由轨迹的分析结果决定. 这补充假设是看成基本假设的一部分.

现在澄清一下一个重要问题, 这个问题是: 该用什么变量来定义状态空间? 在刚才的讨论中, 我们以各分子的位置和动量为变量. 为什么不用位置的平方和能量为变量? 这是一个重要的问题, 因为用不同的变量, 就会有不同的体积. 因此, 变量不确定, Γ 就没有意义了. 现在举例说明这一点, 且先看一个粒子的动量 p 和能量 $E = p^2/2m$. 令某区域的定义为 $0 < E < E_0$, 或 $|p| < \sqrt{2mE_0}$. 如以 p 为变量, 此区域的体积是

$$\Gamma = 2\sqrt{2m}\sqrt{E_0}. \quad (5)$$

如果以 E 为变量, 则体积是

$$\Gamma' = E_0. \quad (6)$$

常数 $2\sqrt{2m}$ 和单位并不重要, 重要的是 $\Gamma \propto \sqrt{E_0}$ 而 $\Gamma' \propto E_0$. 在计算体积时, 能量高的部分在 Γ 中占的份量较小, 而在 Γ' 中占的较多. 所以, 这变量问题就是一个权衡范围中各部分的份量的问题, 也就是一个统计份量的问题. 我们现在把这假设加上一条硬性规定, 以避免这一问题. 这规定的制定如下:

凡是有连续变量, 统计份量的问题总是弄不清楚, 根本解决之道只有放弃连续变量, 退而改用分立值的变数. 量子力学供给了这一条退路, 状态空间该想作一群分立的点, 每一点代表一状态. 我们现在就规定:

Γ 是活动范围内状态的总数, 也就是说每状态的份量算是 1.

现在我们求熵的方法原则已大致齐备, 我们从理想气体的熵 (3) 开始, 指出它和运动幅度的关系, 再提出了活动范围的概念, 最后用一个假设, 把熵的计算问题换成一个统计问题, 即计算活动范围内状态数目. 我们必须记住, 这假设只是一个假设, 并没有经过证明, 它的含义和正确性, 有待进一步的分析.

现在, 我们先讨论一下技术方面的问题, 作为计算熵的准备.

3. 费米粒子

这基本假设的含义, 不是几句话或一两个例子所能说清楚. 其实这本书的主要目的就是要澄清这假设的含义. 我们用最简单的例子开始.

在第三章, 我们已经用细致平衡定律, 讨论了理想气体, 包括量子气体. 粒子间的碰撞或反应, 是分析的关键. 现在我们从基本假设着手, 来看同一问题.

令 L 个位子有相同的能量 ε . 其中有 N 个位子被费米粒子占住. 如果有某种原因, 使这些粒子不断地改换位子, 则熵是多少?

要用基本假设求熵, 必先求得 Γ , 即不变量容许的状态总数. 此地的不变量是 N, L, ε . 因此, Γ 是把 N 个粒子放在 L 个位子的放法的总数. 每位子至多放一个. 这总数是

$$\Gamma = \frac{L!}{N!(L-N)!} \quad (7)$$

根据基本假设, 熵是 $S = \ln\Gamma$.

但是 (7) 式中的一些阶乘, 如不乘出来, 是没什么用的. 事实上, 基本假设只合用于多粒子的群体, 即 $N, L \gg 1$. 熵本来就是群体的性质. 如果 L, N , 都很大, (7) 可以用大数定律简化. 这定律的推导在本章第 6 节. 现在我们用

它的结果, 暂不谈推导, 以求讨论的连贯. 如 $N \gg 1$, 则

$$N! \approx N^N e^{-N} (2\pi N)^{1/2} [1 + O(1/N)]. \quad (8)$$

用这结果代入 (7), 简化, 得

$$S = L \left(\frac{N}{L} \ln \frac{L}{N} + \frac{L-N}{L} \ln \frac{L}{L-N} \right) + \frac{1}{2} \ln \frac{L}{2\pi(L-N)N}. \quad (9)$$

假定 L, N 为同一数量级, 即 N/L 为 $O(1)$, 则 $S = O(L)$, (9) 的最后一项为 $O(\ln L)$, 可以略去. 因为每位子的能量都是 ε , 故总能量为

$$E = \varepsilon N. \quad (10)$$

我们现在把 (9) 的结果对 N 微分一下, 得

$$dS = \ln \frac{L-N}{N} dN. \quad (11)$$

根据热力学, 熵的微分和温度 T , 化学势 μ 有关, 即

$$\begin{aligned} dS &= \frac{1}{T} dE - \frac{\mu}{T} dN \\ &= \frac{\varepsilon - \mu}{T} dN. \end{aligned} \quad (12)$$

以上的 dE 即是 εdN , (见 (10)). (12) 和 (11) 全等, 因此

$$\ln \frac{L-N}{N} = \frac{\varepsilon - \mu}{T},$$

故

$$f \equiv \frac{N}{L} = \frac{1}{e^{(\varepsilon - \mu)/T} + 1}. \quad (13)$$

这结果当然就是第三章的结果, f 是平均每位子的粒子数.

读者会问: 这物体是不是太简化? 每位子能量都是 ε , 而真正物体内部每位子的能量都不一样. 答案是: 这物体可以看作某大物体的一部分, 即把能量和 ε 差不多的位子之集看成此例中的物体. 现在最重要的要求是, 大物体的每一部分都有相同的温度 T 和化学势 μ . 这个要求其实是基本假设的最重要的成果. 也就是说, 如果基本假定成立, 我们可以证明 T 和 μ 在物体各部都有同值. 这个证明, 留到下一章, 我们先细看一个有不同能量粒子的情形.

4. 理想气体

现在我们来考虑一个比较复杂的例子, 即把一理想气体的所有的粒子全考虑进去, 用基本假设求熵.

令理想气体有 N 个分子, 坐标和动量为 $\mathbf{r}_k, \mathbf{p}_k, k = 1, 2, \dots, N$. 在第三章, 第 1 节, 已经求得单粒子状态的统计公式:

$$\int d\Gamma = \frac{1}{h^3} \int d^3\mathbf{r} d^3\mathbf{p}, \quad (14)$$

即在 $d^3\mathbf{p}d^3\mathbf{r}$ 的体积中有 $d\Gamma$ 个状态. 现在将 (14) 推广到 N 粒子状态的统计:

$$\begin{aligned} \int d\Gamma &= \int d\Gamma_1 d\Gamma_2 \cdots d\Gamma_N \\ &= \frac{1}{h^{3N}} \int d^3\mathbf{r}_1 d^3\mathbf{r}_2 \cdots d^3\mathbf{r}_N d^3\mathbf{p}_1 \cdots d^3\mathbf{p}_N \end{aligned} \quad (15)$$

即在 $6N$ 维空间的体积 $d^3\mathbf{r}_1 \cdots d^3\mathbf{r}_N d^3\mathbf{p}_1 \cdots d^3\mathbf{p}_N$ 中的状态数.

要计算 Γ , 必须考虑各不变量的限制. 此地的限制是: ① 体积不变, 各分子的位置必须在体积内, 和 ② 总能量不变, 为一定值 E .

① 使 \mathbf{r}_k 积分限于

$$0 < x_k, y_k, z_k < L, \quad k = 1, 2, \dots, N. \quad (16)$$

容器的体积是 $V = L^3$.

② 可以用一个 δ 函数 $\delta(E - H)$ 来达成, H 是哈密顿量

$$H = \sum_{k=1}^N \frac{p_k^2}{2m}. \quad (17)$$

(16) 和 (17) 的限制还不够, 还要把统计结果除以 $N!$ 才行, 因为每粒子都一样, 交换任一对粒子的号码不会改变状态. (16), (17) 并未限制号码, 因此把状态多算了 $N!$ 倍. 这一点在本章开始时略谈过一下. 如果每粒子都不一样, 是不是就不除以 $N!$ 了? 这问题待本节末再谈.

现在用 (15) 来求 Γ , 加入 (16), (17) 的限制. 得

$$\begin{aligned} \Gamma(E) &= \frac{V^N}{N! h^{3N}} W(E), \\ W(E) &= \int d^3\mathbf{p}_1 d^3\mathbf{p}_2 \cdots d^3\mathbf{p}_N \delta(E - H). \end{aligned} \quad (18)$$

$W(E)$ 可以看成是一个 $3N$ 维空间里的球面积, 球的半径是由 $H = E$ 规定:

$$\left(\sum_{k=1}^N p_k^2 \right)^{1/2} = 2mE. \quad (19)$$

在这球面上的点合乎 (19). 计算球面积是一个有趣的数学问题, 我们把它放在下一节里, 结果见 (39), (31). 现在先把答案写下:

$$S = \ln \Gamma = N \left[\frac{3}{2} \ln \left(\frac{4\pi m E}{3N h^3} \right) + \frac{5}{2} + \ln \frac{V}{N} + O \left(\frac{1}{N} \ln N \right) \right] \quad (20)$$

有了这结果, 其它各量就可以用微分求出, 因为 $dE = -pdV + TdS$, 所以

$$\begin{aligned} \frac{1}{T} &= \left(\frac{\partial S}{\partial E} \right)_V = \frac{3N}{2E}, \\ p &= T \left(\frac{\partial S}{\partial V} \right)_E = T \frac{N}{V}. \end{aligned} \quad (21)$$

其他的性质, 如热容率, 膨胀系数等都可以计算出来.

我们曾强调过: 基本假设如要成立, 各变量必须要不断地变动. 这理想气体的各变量是如何变动呢? 位置变动是因为各分子有速度, 分子动量的改变则要靠碰撞, 或与容器壁, 或与其他分子. 如果分子和容器壁碰撞时, 常有能量的交换, 则即使各分子间不碰撞, 分子动量的大小、方向都常改变. 如果总是没有能量交换, 则动量的方向虽改变, 但大小不变, 这基本假设就不能成立. 如果分子间互相碰撞, 则动量改变就不缺乏了. 令 λ 为分子的“平均自由距离”, 即每分子在两次碰撞间的距离, 则

$$\tau = \frac{\lambda}{v} \quad (22)$$

为“平均自由时间”, v 为平均速度. 如果

$$L \gg \lambda, \quad (23)$$

即分子碰撞频繁, 则观测时间 t 必须合乎

$$t \gg \tau \quad (24)$$

的条件, 基本假设才可能成立. 如果 $L \ll \lambda$, 即分子间极少碰撞, 则

$$t \gg L/v \quad (25)$$

为必须条件, 空气分子在常温常压下的平均自由距离是 $\lambda \sim 10^{-5}\text{cm} = 1000\text{\AA}$, $v \sim 5 \times 10^4\text{cm/秒}$, $\tau \sim 10^{-9}\text{秒}$, 注意

$$\lambda \sim \frac{1}{n\sigma}$$

n 是密度, $\sigma \sim 10\text{\AA}^2$ 是分子碰撞有效面积. 空气分子间的平均距离是 $n^{1/3} \sim 30\text{\AA}$. 其它温度和压强下的 λ 和 τ 可以由 $v \propto \sqrt{T}$ 并以 T/p 代 n 求出.

当然, 分子间的碰撞是由相互作用引起, (17) 式并没有包括相互作用的能量. 因此, 我们一方面利用相互作用引起的碰撞, 来满足基本假设的应用条件, 另一方面在算熵时又忽略相互作用的能量. 这并不会引起矛盾, 因为分子和分子接触的时间非常短, 作用能量只在接触时发生. 碰撞的主要效果是使前后动量不同. 从活动范围的定义 (16), (17) 到各种性质的求出, 只是一连串的数学计算, 原则上不需要新的概念介入. 这基本假设一般的用法, 也大致如此, 只是计算细节会不一样, 而且复杂得多.

一旦活动范围定了, 接下来的数学计算往往是非常复杂的难题. 这类难题本身就是十分有趣, 有各种解决的方法. 我们将逐步讨论, 并引出新的概念, 以帮助我们了解.

有一点必须再三强调: 活动的范围的确定, 是第一步, 是最重要的一步, 并且常常是最不清楚, 最困难的一步.

以上的结果 (20) 是由基本假设算出来的熵, 它和第三章的 (3.31) 是一样的. 在第三章, 我们指出, 化学势 μ 的定义一旦确定, 熵也定了. 化学势的定义并没有什么选择的余地. 从守恒定律和化学反应的分析就把它定了. 如果不用第三章的定义, 则各公式中的 μ 必须改写为 $\mu + c$, c 是一个常数, 但每一种物质都将有一个不同的 c , 非常不便. 基本假设所定的熵, 并没有把任何物质的结构和性质列入定义步骤中, 因此, 由这假设算出的熵和化学势, 不对则已, 如对, 则 μ 中的常数 c 必是零. (20) 和 (3.31) 完全一样, 并非巧合.

我们再回过头来讨论 (15) 式中的 $N!$. 如果各分子都不一样, 是否把这 $N!$ 丢了就行? 当然, 这只是纸上谈兵, 要把 N 个不同种的分子放在一起, 而 $N \gg 1$, 恐非易事. 即使有这样多种分子存在, 放在一起大概也不是理想气体. 但讨论这问题是有原因的. 如果各分子不一样, 我们该问, 在观测时间内, 一个分子可以跑多远? 如果密度很稀, 容器很小, 则每个分子可以来往各壁之间多次, 因此, 活动范围仍以 (16), (17) 定义, 只要把 (15), (18) 的 $N!$ 拿掉就行了. 但如果密度大, 相互碰撞多, 在观测时间内一个分子走不了多远. 如此,

(16) 的限制不够, 还得加上时间有限所引起的限制. 如果这时间的限制不能有够明确的定义, 则基本假设就无法使用了. 我们以后会常讨论时间的限制和基本假设的应用问题.

当然, 如果 $N!$ 拿掉, 则 S 就不和 N 成正比, 即非广延量. 这是小容器, 分子不互撞的情形, 不和 N 成正比并没有不对. 各分子既不同且来往于四壁之间, 则无法分成有多个同样性质的部分. 如果把容器隔成两半, 则一边的分子就不能跑到另一边去, 和不分割的情况不一样了, 熵也就不一样了.

(15), (18) 不但合用于气体, 而且合用于稀溶液中的溶质分子, 只是下列条件要遵守: 溶质分子必须动得快, 即在观测时间内所跑的距离远大于溶质分子之间的平均距离. 这些式子不合用于散布在固体中的, 冻结住的杂质分子. 冻住的分子的位置是不变的. 这些将在讲杂质时 (第十八章) 详细分析. 注意, 如果分子是冻住的, 则各分子是否同类就不是问题. (15) 式中的每个 r_k 积分都会限制在一定的冻结位置附近. 不管分子是否同种, $N!$ 都不能要.

总之, 活动范围的划定, 必须由分子运动决定. 每一个因子都有其意义, 不能随便丢掉或凑入.

5. 高维空间的一些特征

在第 3 节的简例中, 状态空间共有 2^L 个状态, 也就是说, L 个位子, 每位子可以有两个状态. 活动范围是在这空间中的一部分, 由总粒子数决定.

在上一节, 状态空间是用一 $6N$ 维空间中的一些点来定义. 每一点是一状态. 点的密度是 $1/h^{3N}$, 平均分布在这 $6N$ 维空间内, 活动范围是由 $E = H$, 即 (18) 中的 $\delta(E - H)$ 决定. (N 已经定了). 活动范围可看成一球面. 在第 3 节的例子, 状态空间可看成 L 维空间中的立方体的 2^N 个顶点, 活动范围的几何意义就不大清楚了.

严格说来, 一群有限个分立的点, 是没有几何意义的. 上节的状态, 也可以用各粒子位子的粒子数表示, $6N$ 维空间的表示法, 并非必要. 不过, 点数既多, 几何的观念常常很有用, 上节用的 $6N$ 维空间, 是一个很好的例子. 群体的特点是运动变量繁多, 状态空间因此和高维数的空间性质有些关系.

在这一节, 我们介绍一些高维空间的集合性质, 并讨论“大数定律”.

令 $x = (x_1, x_2, \dots, x_N)$ 代表 N 维向量空间的点, x 的“长度”定义是

$$|x| = \left(\sum_{i=1}^N x_i^2 \right)^{1/2}. \quad (26)$$

假定 N 是一个很大的数字, 像 10^{20} 之类. 如果至少有一部分的 x_i 不是零, 则

$$|x| \sim N^{1/2}. \quad (27)$$

这是一个大数. 如果每个 x_i 都稍稍改变一些, $\Delta x_i \sim \varepsilon$, 则

$$|\Delta x| \sim \varepsilon N^{1/2}. \quad (28)$$

即使 ε 不大, $|\Delta x|$ 还可以很大, 倒过来说, 虽然在这空间走了一段长距离, 每个 x_i 的改变仍可以是非常的小.

最特别的是高维空间里的体积, 现在看一个球体的体积. 令球半径为 R , 体积为 $Q(R) = CR^N$, 球的表面积便为 $Q'(R) = NCR^{N-1}$. 常数 C 可以利用以下的积分求出. 令

$$\begin{aligned} I &= \int_{-\infty}^{\infty} dx_1 dx_2 \cdots dx_N e^{-\frac{1}{2}(x_1^2 + x_2^2 + \cdots + x_N^2)} \\ &= \left[\int_{-\infty}^{\infty} dx e^{-\frac{1}{2}x^2} \right]^N \\ &= (2\pi)^{N/2}. \end{aligned} \quad (29)$$

如果改用球形坐标, 得

$$\begin{aligned} I &= \int_0^{\infty} dR Q'(R) e^{-\frac{1}{2}R^2} \\ &= NC \int_0^{\infty} dR R^{N-1} e^{-\frac{1}{2}R^2} \\ &= 2^{N/2} C \left(\frac{N}{2} \right)!. \end{aligned} \quad (30)$$

比较 (29) 和 (30), 即得 C , 所以球的体积是

$$Q(R) = \pi^{N/2} R^N / \left(\frac{N}{2} \right)!. \quad (31)$$

如果半径由 R 增加到 $(1 + \varepsilon)R$, 则 $Q(R)$ 增加到

$$Q(R)(1 + \varepsilon)^N. \quad (32)$$

即使 ε 不大, $(1 + \varepsilon)^N$ 可以是个惊人大数. 因此一个厚度为 ε 的球壳的体积

$$Q(R)[(1 + \varepsilon)^N - 1] \quad (33)$$

和球体积 (32) 相差微不足道, 所以球的体积是集中在表层上.

大致说来, 高维空间的长度是大数, 但体积是“特大数”.

$$\begin{aligned} \text{长度} &\sim a\sqrt{N}, \\ \text{体积} &\sim a^N. \end{aligned} \quad (34)$$

a 是一个尺度标准. 以上这些奇特的性质是因为这空间有极多的方向.

6. 尖峰积分和大数定律

以下用大数定律的推导为例, 说明一点常用的数学技巧, 我们把 $N!$ 写成一积分

$$\begin{aligned} N! &= \int_0^{\infty} dX e^{-X} X^N \\ &= \int_0^{\infty} dX e^{S(X)}, \end{aligned} \quad (35)$$

(1)

其中

$$S(X) \equiv N \ln X - X.$$

如果 N 是大数, X^N 随 X 之增加而剧增, e^{-X} 则剧减. $e^{S(X)}$ 除了在其极大值附近外都很小, 极大值很容易求出:

$$\frac{\partial S}{\partial X} = \frac{N}{X} - 1 = 0, \quad (36)$$

其解为 $X_c = N$, 在 X_c 附近

$$\begin{aligned} S(X) &= S(X_c) + \frac{1}{2} \left(\frac{\partial^2 S}{\partial X^2} \right)_c (X - X_c)^2 + \dots \\ &\approx N \ln N - N - \frac{1}{2N} (X - X_c)^2. \end{aligned} \quad (37)$$

因此

$$\begin{aligned} N! &\approx e^{N \ln N - N} \int_{-N}^{\infty} dx' e^{-\frac{1}{2N} x'^2} \\ &= N^N e^{-N} (2\pi N)^{1/2} (1 + O(1/N)). \end{aligned} \quad (38)$$

大致上, $X^N e^{-N}$ 是一尖峰, 中心在 $X_c = N$, 峰宽是 $\sqrt{2\pi N}$, 这宽度比 N 要小很多. 所以 x' 的积分下限可以算是 $-\infty$. 式中的 $O(1/N)$ 项, 需要更仔细的计算才能确定. (38) 就是所谓“大数定律”. 以上的积分法, 合用于尖峰形的积分函数.

用 (38), 球体积 $Q(R)$ (31) 和球面积 $Q'(R)$ 可以简化成

$$\begin{aligned} Q(R) &\approx e^{\frac{N}{2} \left(\ln \frac{2\pi R^2}{N} + 1 - \frac{1}{N} \ln \pi N \right)}, \\ Q'(R) &\approx e^{\frac{N}{2} \left(\ln \frac{2\pi R^2}{N} + 1 - \frac{1}{N} \ln \pi N + \frac{2}{N} \ln \frac{N}{R} \right)}. \end{aligned} \quad (39)$$

因此, 当 $N \gg 1$ 时,

$$\begin{aligned} \frac{1}{N} \ln Q(R) &\approx \frac{1}{N} \ln Q'(R) \\ &\approx \frac{1}{2} \left(\ln \frac{2\pi R^2}{N} + 1 \right) + O\left(\frac{\ln N}{N}\right). \end{aligned} \quad (40)$$

7. 能量均分定律

我们再看一个示范例子, 即“能量均分定律”, 这定律的内容是: 如果哈密顿量是 M 个平方项之和

$$H = \sum_{i=1}^M A_i q_i^2 \quad (41)$$

q_i 是坐标或动量. A_i 为正值常数, 则总能量 E 和温度 T 的关系是

$$\frac{E}{M} = \frac{1}{2} T \quad (42)$$

也就是说, 平均每一变量 q_i 的运动, 分配到能量 $\frac{1}{2}T$.

现在回到理想气体的结果 (21), 即

$$\frac{E}{3N} = \frac{1}{2} T. \quad (43)$$

再看哈密顿量 H (见 (17)) 是 $3N$ 项之和 (每个 p_k^2 实是三项 $p_{kx}^2 + p_{ky}^2 + p_{kz}^2$). 因此, (41) 指出每一项的平均能量是 $\frac{1}{2}T$, 这就是能量均分定律的一例.

另一例是一群单摆, 即简谐振动单元, 其哈密顿量为

$$H = \sum_{i=1}^N \left(\frac{1}{2} p_i^2 + \frac{1}{2} \omega_i^2 x_i^2 \right). \quad (44)$$

此地每个 p_i 算一个变数. H 共有 $2N$ 项, 每项能量, 照以上定律, 为 $\frac{1}{2}T$, 故总能量值为

$$E = \left(\frac{1}{2} T \right) \times 2N = NT. \quad (45)$$

这定律的推导十分简单. $\Gamma(E)$ 的计算和理想气体的大同小异

$$\begin{aligned} \Gamma(E) &= \frac{1}{h^M} \int dq_1 dq_2 \cdots dq_M \delta(E - H) \\ &= \frac{1}{h^M} (A_1 A_2 \cdots A_M)^{-\frac{1}{2}} \int dx_1 dx_2 \cdots dx_M \delta(E - x_1^2 - x_2^2 \cdots - x_M^2), \end{aligned} \quad (46)$$

$$x_i \equiv \sqrt{A_i} q_i.$$

(46) 的积分又是一个球面积, 用 (40), 得

$$\begin{aligned} S(E) = \ln \Gamma(E) &= \frac{M}{2} \left(\ln \frac{2\pi E}{M h^2} + 1 \right) - \frac{1}{2} \sum_{i=1}^M \ln A_i, \\ \frac{\partial S}{\partial E} &= \frac{1}{T} = \frac{M}{2} \cdot \frac{1}{E}. \end{aligned} \quad (47)$$

T 只和 $(M/2)\ln E$ 这项有关.

8. 活动范围面面观

从这一章开始, 我们就强调: 活动范围的意义, 是代表各变数的改变幅度. 以上繁复的技术性的讨论, 很可能使读者忘了这意义. 在这一节我们从几何的投影观点来分析, 使这意义更为扎实.

我们先穿插一段有关“投影”的常识. “投影”在现实生活中是再熟悉不过了. 我们看到的东, 都是三维空间在视网膜上的投影, 工程师画的图, 也是三维空间物体投在二维空间的影, 也就是一种从各方向的透视. 同理, n 维

向量空间中的点, 都可投影到 n' 维的“底面空间”里, 底面空间是 n' 个轴撑开的空间 (当然 $n' < n$), 它是这 n 维空间的“一面”或“一方面”. 当然, 一个点在每一面的投影都是一点, 我们要看的是一群点在各方面的投影, 看这些点在这些投影里的分布.

投影里的分布密度的计算法, 可以用一个例子说明: 假设一 3 维空间的球面上

$$x^2 + y^2 + z^2 = R^2 \quad (48)$$

均匀地密布着很多点, 这些点在 XY 平面上的投影密度是什么? 答案是

$$\begin{aligned} \gamma(x, y) &= \frac{\Gamma}{4\pi R^2} \int dz \delta(\sqrt{x^2 + y^2 + z^2} - R) \\ &= \frac{\Gamma}{2\pi R} \frac{\theta(R^2 - x^2 - y^2)}{\sqrt{R^2 - x^2 - y^2}}. \end{aligned} \quad (49)$$

式中 Γ 是点的总数, 球面的投影是一圆盘子, 点的密度在盘子边附近很高, 读者可自己找出原因. 要得到在 X 轴上的投影密度, 只需再积分一次.

$$\begin{aligned} \gamma(x) &= \frac{\Gamma}{4\pi R^2} \int dz dy \delta(\sqrt{x^2 + y^2 + z^2} - R) \\ &= \frac{\Gamma}{2R} \theta(R - |x|). \end{aligned} \quad (50)$$

这些积分的意义, 是把垂直于投影面的各方向的点数加起来.

一般说来, 在 n 维空间的球面上的均匀分布, 投影在 n' 维空间的密度是

$$\begin{aligned} \gamma(x_1, x_2, \dots, x_{n'}) &\propto \theta(R - r) (R^2 - r^2)^{\frac{n-n'}{2}-1} \\ r^2 &= x_1^2 + x_2^2 + \dots + x_{n'}^2. \end{aligned} \quad (51)$$

如果 $n \gg n'$, 则密度集中在中心, 因为

$$(R^2 - r^2)^{\frac{M}{2}} \approx R^M e^{-\frac{Mr^2}{2R^2}}, \quad (52)$$

只要 M 是一大数, 这样的分布大致集中在

$$r \sim \frac{R}{\sqrt{M}}$$

之内, 这和 (49) 和 (50) 的情形大不相同. (52) 是一个“正态分布”或“高斯分布”. 这种分布, 在第 6 节讨论尖峰积分时出现过, 以后会再常常出现.

理想气体的状态活动范围, 是一个在 $6N$ 维空间里的球面, 各状态可由分立的点代表, 点的分布密度是均匀的, 这活动范围在各方向的大小就是各运动变量的改变幅度, 这些幅度该是怎样量出来呢? 最直接的办法是把这活动范围向各方面投影, 从投影上点的分布, 可以得到答案.

现在把活动范围内的点投影到 \mathbf{p}_1 空间里, 投影上的点的密度分布是

$$r(\mathbf{p}_1) = \frac{V^N}{N!h^{3N}} \int d^3\mathbf{p}_2 \cdots d^3\mathbf{p}_N \delta\left(E - \frac{p_1^2}{2m} - H'\right) \quad (53)$$

$$H' = \sum_{i=2}^N \frac{p_i^2}{2m}, \quad H = \frac{p_1^2}{2m} + H'. \quad (54)$$

(53) 中, \mathbf{p}_1 没有被积分, 而是固定的. $\gamma(\mathbf{p}_1)d^3p_1$ 是第一粒子动量在 d^3p_1 内的各状态的总数. H' 是除去了第一粒子后的能量. 所以依照 $\Gamma(N, E)$ 的定义 (18)

$$\begin{aligned} \gamma(\mathbf{p}) &= \frac{V}{Nh^3} \Gamma\left(N-1, E - \frac{p^2}{2m}\right) \\ &= \frac{V}{Nh^3} e^{S\left(N-1, E - \frac{p^2}{2m}\right)} \\ &= \frac{V}{Nh^3} e^{S(N, E) + \frac{\mu}{T} - \frac{1}{T}\left(\frac{p^2}{2m}\right) + O\left(\frac{1}{N}\right)} \propto e^{\frac{\mu}{T} - \frac{p^2}{2mT}}. \end{aligned} \quad (55)$$

这结果其实就是 (52), $2mE = R^2$, $p^2 = r^2$, $T = R^2/N$. (55) 式中各符号的意义该很清楚,

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_N, \quad \frac{\mu}{T} = -\left(\frac{\partial S}{\partial N}\right)_E. \quad (56)$$

化学势 μ 和 S 的关系, 可以从第二章的功、热讨论得出

$$dE = \mu dN + T dS. \quad (57)$$

(其他的微分未考虑), S 对 N 、 E 的二次微分都是 $\frac{1}{N}$ 级数量, 因此, 活动范围在 \mathbf{p}_1 方面, 或任一 \mathbf{p}_i 方面的投影, 是一个正态分布, 这分布范围的大小大致是

$$\Delta p = \sqrt{mT}. \quad (58)$$

这可算是分子动量改变幅度的一个估计.

我们也可以用大一点的空间作投影面, 例如 $(\mathbf{p}_1, \mathbf{p}_2)$ 空间, 或更大的空间, 以上的分析只须稍稍推广, 即得投影点密度分布:

$$\gamma(\mathbf{p}_1, \mathbf{p}_2 \cdots \mathbf{p}_n) \propto e^{-\frac{1}{T} \sum_{k=1}^n \frac{p_k^2}{2m}} e^{\frac{\mu n}{T}}. \quad (59)$$

只要 $n \ll N$, 这结果就成立.

一般说来, 所谓“投影”就是“分类收集”“分类统计”, 把同一类型的状态收集起来, 计其数目而已.

以上的投影分析告诉我们活动范围往各方面看去是什么样子, 有多大, 这些当然都是基本假设的后果. 从这里, 我们可以回到本章的第 1 节, 把本节的讨论和第 1 节的互相印证比较.

9. 轨迹和活动范围

现在我们对活动范围的几何性质, 略有了一点概念. 在第 2 节, 我们用苍蝇在一房间飞的例子, 来说明活动范围的意义. 苍蝇的轨迹分布在房间里.

不过, N 个分子的状态空间和苍蝇的房间不大一样, 因为 N 是大数, 不是 3. 状态空间的体积是 e^N 数量级的大数. 轨迹是在观测时间内发生的状态. 有多少状态发生? 大致是

$$n \sim N \frac{T}{\tau}. \quad (60)$$

τ 是平均自由时间, T/τ 是每分子在时间 T 内碰撞的次数. 每撞一次, 状态改一次. n 是一个 N 级大数. (当然, 不碰撞时, 分子位置亦变, 故 n 比 (60) 要大些, 但仍是 N 级大数). N 量级比 e^N 量级差得太远了. 也就是说轨迹在活动范围中是一条微乎其微的细丝. 活动范围, 必须由轨迹定义, 但这条细丝, 何以能撑起这广大的体积? 乍看起来, 这细丝只能留在一个极小的区域中. 这问题很难回答. 在第廿五章我们会作更详细的分析. 在此我们由几何观点, 来作一粗略分析.

令 $(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N) \equiv \mathbf{p}$ 为 N 个气体分子的动量. \mathbf{p} 的空间是一个 $3N$ 维空间. 为求简明, 我们不讨论位置空间.

如果两分子碰撞 (例如分子 1, 2), 则此二分子的动量由 $\mathbf{p}_1, \mathbf{p}_2$ 改成 $\mathbf{p}'_1, \mathbf{p}'_2$. 令

$$\mathbf{q}_{12} \equiv (\mathbf{p}'_1 - \mathbf{p}_1, \mathbf{p}'_2 - \mathbf{p}_2, 0, 0, 0, \dots, 0). \quad (61)$$

也就是说, 状态由 \mathbf{p} 移动到 $\mathbf{p} + \mathbf{q}_{12}$. 令下一次的碰撞为分子 i 和 j 的碰撞. 则状态移到 $\mathbf{p} + \mathbf{q}_{12} + \mathbf{q}_{ij}$, 在 n 次碰撞之后, 有 n 个 \mathbf{q} 向量. 各个 \mathbf{q} 的长度大小不一. 它们的平均就是各分子的动量改变的幅度, 即活动范围的各边长度.

现在注意, 如果 \mathbf{q}_{12} 和 \mathbf{q}_{ij} 的 12 和 ij 为不同的分子, 则 \mathbf{q}_{12} 和 \mathbf{q}_{ij} 为互

相垂直的向量, 即

$$\mathbf{q}_{12} \cdot \mathbf{q}_{ij} \equiv \sum_{\alpha=1}^N \mathbf{q}_{12\alpha} \cdot \mathbf{q}_{ij\alpha} = 0. \quad (62)$$

原因是 \mathbf{q}_{12} 只有 $\alpha = 1$ 和 2 分量非零, 而 \mathbf{q}_{ij} 只有 $\alpha = i$ 和 j 分量非零.

如果 n 次碰撞, $n \sim N$, 则 n 个 \mathbf{q} 向量几乎都是互相垂直的, 因为各部份的分子都在碰撞.

我们现在需要一点想象力. 如有几根棍子, 要撑出一个空间来, 该如何办? 答案是, 各棍子该互相垂直, 才能撑出个空间来. 造房子时, 栋梁都是互相垂直的. 现在我们在 $3N$ 维空间, 有 $n \sim 3N$ 个 \mathbf{q} 向量做的棍子. 要撑个空间, 当然要把各个 \mathbf{q} 放成互相垂直的方向, 才能撑出个 $3N$ 维空间的体积来. 如果各个棍子的方向分布在所有方向, 则撑不出来. 因为有 e^N 个方向, 需要 e^N 根棍子.

轨迹是各个 \mathbf{q} 向量连起来的. 就是因为各个 \mathbf{q} 互相垂直, 轨迹才可能撑出一个 $3N$ 维空间的体积. 轨迹中的状态虽然比起体积为数极微, 但可以像房屋的栋梁一样, 撑出一个体积来.

以上是对轨迹和活动范围的一个粗略的解释. 各 \mathbf{q} 向量的垂直性, 在以后会解释成“独立性”. 到第六篇时, 再分析.

讨论问题五

1. 把第 3 节的分析, 用到波色粒子.

(A) 求熵.

(B) 求粒子数和 T, μ 的关系.

2. 某顺磁体的总能量函数是

$$H = - \sum_{i=1}^N s_i h, \quad s_i \text{ 可以是 } +1 \text{ 或 } -1. \quad (63)$$

磁元 s_i 的改变是由某种运动法则决定, 在此不管. h 是一固定常数, 即磁场.

(A) 根据基本假设, 求熵.

(B) 求能量 E 和温度 T 的关系. 总磁矩 $M = \sum_{i=1}^N s_i$ 即 $-E/h$.

(C) 这模型和第 3 节的费米粒子颇有相似之处. 试述二者之间的关系.

3. 令 a 为一 N 维空间立方体的边长. 它的体积是 a^N . 试求它的外接球和内切球的体积, 并比较它们的大小. (N 是一大数)

4. 一均匀绳子, 质量为 M , 长度为 L , 两端固定, 张力为 F .

(A) 从下面方程式求出基本振动 (简正) 频率.

$$\frac{\partial^2 y}{\partial t^2} = c^2 \frac{\partial^2 y}{\partial x^2}, \quad c^2 = FL/M, \quad 0 < x < L. \quad (64)$$

(B) 求能量, 熵, 比热和温度的关系, 假设最高频率为 ω_D .

5. 假设上题的绳子是在温度 T 的气室中, 令 d 为绳子的直径, η 为空气的粘滞系数.

(A) 考虑空气阻力, 求方程式的修正项.

(B) 观测时间须甚长于起伏周期 τ , 平衡才有意义, 若考虑空气阻力, 试求 τ .

(C) 上题 (B) 的结果, 受不受 τ 的影响? 如果绳子是浸在水中, 则结果受影响吗?

6. 导出 (51), (52), 并澄清 (55) 和 (52) 的关系. (55) 当然就是第三章已谈过的玻尔兹曼分布, 在此是看成一个高度空间球体的投影.

7. 求第 2 题的活动范围在 s_1 方向的投影.

8. 立方体的投影.

n 维空间立方体的顶点可以用坐标 $(1, 1, \dots, 1)$, $(-1, 1, 1, \dots, 1)$, $(1, -1, 1, \dots)$...
 $1)$, 共 2^n 个向量代表. 称这些向量为 \mathbf{R}_i , $i = 1, 2, \dots, 2^n$, 任取两个垂直的向量 $\mathbf{e}_x, \mathbf{e}_y$,

$$\mathbf{e}_x \cdot \mathbf{e}_x = 1, \quad \mathbf{e}_y \cdot \mathbf{e}_y = 1, \quad \mathbf{e}_x \cdot \mathbf{e}_y = 0. \quad (66)$$

则 \mathbf{R}_i 在 xy 平面上的投影为

$$R_{ix} = \mathbf{R}_i \cdot \mathbf{e}_x. \quad \text{及} \quad R_{iy} = \mathbf{R}_i \cdot \mathbf{e}_y. \quad (67)$$

(A) 试画三维空间立方体的投影图, 即 $n = 3$. 避免重迭的顶点投影.

(B) 试画四度、五度立方体的投影图. 这些图应给读者一点高度空间的“印象”.

9. 第 9 节的分析指出, n 次碰撞, $n \geq 3N$, 轨迹就可以撑出一个 $3N$ 维动量空间的体积来. 因此, 所需的时间约是 $T \geq \tau$, (见 (60)).

若要撑出 $3N$ 维位置空间的体积, 要多少时间才够?

第六章 基本假设的初步应用

这一章我们从基本假设推导出一些平衡状态的特性, 即物体各部分的温度、压强、化学势等“力”都一样. 利用这些特性, 我们分析一些应用例子, 即晶体振动方面的问题, 包括德拜模型和第二声波.

1. 部分和总体 分类和综合

一个大物体可以看作它各部分的组合, 每一部分都是一个群体. 现在我们从基本假设出发, 看看总体的熵和各部分的熵有什么关系. 先把物体分为两部份, 1 和 2. 两部份的状态 s_1 和 s_2 指定之后, 总体的状态 s 也就定了. 因此,

$$s = (s_1, s_2). \quad (1)$$

总体能量 H 是两部分能量 H_1, H_2 之和加上一些交界上的作用能 H_{12}

$$H(s) = H_1(s_1) + H_2(s_2) + H_{12}(s_1, s_2). \quad (2)$$

现在把 H_{12} 略去, 理由是它只在两部份交界处有效, 为量甚小. 这并不是说它不重要, 两部份间能量的交易, 全靠它. 不过它的主要效果是使交界两边能量交流, 对总能量之值影响不大, 因此我们暂时将它略去, 于是总体的熵是

$$S(E) = \ln\Gamma(E), \quad (3)$$

$$\begin{aligned} \Gamma(E) &= \sum_{s_1} \sum_{s_2} \delta(E - H_1(s_1) - H_2(s_2)) \\ &= \int dE_1 dE_2 \delta(E - E_1 - E_2) \Gamma_1(E_1) \Gamma_2(E_2) \\ &= \int dE_1 dE_2 \delta(E - E_1 - E_2) e^{S_1(E_1) + S_2(E_2)}, \end{aligned} \quad (4)$$

其中

$$\begin{aligned}\Gamma_i(E_i) &= \sum_{s_i} \delta(E_i - H_i(s_i)) \\ &\equiv e^{S_i(E_i)}, \quad i = 1, 2.\end{aligned}\tag{4'}$$

我们再用尖峰积分法来做 (4), 在此地必须求 $S_1(E_1) + S_2(E_2)$ 的极大值. 不过 E_1, E_2 不是独立变数, 而是受 $E_1 + E_2 = E$ 所限制. 凡是在此类限制下求极大值的问题, 都可用“外力法 (拉格朗日乘子方法)*”解决. 现在提一下这方法的原理.

假设一粒子其位能是 $V(\mathbf{r})$, 但它的位置 \mathbf{r} 被限制在一曲面 $\phi(\mathbf{r}) = 0$ 上. 问: 平衡位置在那里? 即 $V(\mathbf{r})$ 在这曲面上的极小值在那里? 这粒子不但受力 $-\nabla V(\mathbf{r})$, 而且受曲面的力, 这曲面加于粒子之力是和曲面垂直 (假设没有摩擦力), 故和 $-\nabla\phi(\mathbf{r})$ 成正比, 可写成 $-\lambda\nabla\phi$. 在平衡处此力和 $-\nabla V(\mathbf{r})$ 抵消

$$-\nabla V - \lambda\nabla\phi = -\nabla(V + \lambda\phi) = 0, \quad \phi(\mathbf{r}) = 0,\tag{5}$$

λ 计量曲面施力之强度. (5) 指出:

我们必须求 $(V + \lambda\phi)$ 的极小值, 平衡点和 λ 的值可由 (5) 两式子联立解出. 对这粒子而言, $\phi(\mathbf{r}) = 0$. 这限制的效果是一个外加的力, λ 计其强度, 这外力把它托住在曲面上.

现在回到 (4). 要找 $S_1(E_1) + S_2(E_2)$ 的极大值, 我们用外力法, 照 (5) 写下

$$\begin{aligned}\frac{\partial}{\partial E_1} (S_1(E_1) + \beta(E - E_1 - E_2)) &= 0, \\ \frac{\partial}{\partial E_2} (S_2(E_2) + \beta(E - E_1 - E_2)) &= 0, \\ E - E_1 - E_2 &= 0,\end{aligned}\tag{6}$$

β 是外力强度. (6) 的意义十分明显, $\frac{\partial S_1}{\partial E_1} = T_1$, $\frac{\partial S_2}{\partial E_2} = T_2$ 是两部分温度, (6) 指出

$$T_1 = T_2 = \frac{1}{\beta} = T,\tag{7}$$

即二温度相等时, (4) 的积分函数为极大.

*即所谓 Lagrange multiplier 方法.

令 $E_1 = \bar{E}_1$, $E_2 = \bar{E}_2$ 为 (6) 之解, $S_1 + S_2$ 在这极大点附近的展开式是

$$\begin{aligned} & S_1(E_1) + S_2(E_2) \\ &= S_1(\bar{E}_1) + S_2(\bar{E}_2) + \frac{1}{T} (E_1 - \bar{E}_1 + E_2 - \bar{E}_2) \\ &\quad - \frac{1}{2} \left[\frac{1}{K_1} (E_1 - \bar{E}_1)^2 + \frac{1}{K_2} (E_2 - \bar{E}_2)^2 \right] \\ &\quad + \dots \end{aligned} \quad (8)$$

$$\begin{aligned} \frac{1}{K_i} &\equiv - \left(\frac{\partial^2 S_i}{\partial E_i^2} \right)_{E_i = \bar{E}_i} = (T^2 C_i)^{-1}, \quad i = 1, 2. \\ C_i &= \frac{\partial E_i}{\partial T_i}. \end{aligned} \quad (9)$$

(8) 中的一次项在 $E_1 + E_2 = \bar{E}_1 + \bar{E}_2 = E$ 的条件下是零, 将 (8) 代入 (4) 得

$$\begin{aligned} \Gamma(E) &\approx e^{(S_1(\bar{E}_1) + S_2(\bar{E}_2))} \int dE'_1 dE'_2 \delta(E'_1 + E'_2) \\ &\quad \times e^{-\frac{E_1'^2}{2K_1} - \frac{E_2'^2}{2K_2}}, \\ &= e^{(S_1(\bar{E}_1) + S_2(\bar{E}_2))} \sqrt{2\pi K}. \end{aligned} \quad (10)$$

$$K \equiv \frac{K_1 K_2}{K_1 + K_2}.$$

因此总体的熵是

$$S(E) = S_1(\bar{E}_1) + S_2(\bar{E}_2) + \ln(\sqrt{2\pi K}), \quad (11)$$

最后一项可以略去, 因为它只是 $\ln N$ 级的数量, 而 S_1, S_2 是 N 级的, N 是粒子总数.

以上的推演, 看来十分简单, 但其中不很明显的含义, 需要进一步说明. $\Gamma(E)$ 是 $\Gamma(E_1)\Gamma(E_2)$ 的积分 (见 (4)). $\Gamma(E_1)\Gamma(E_2)$ 是一个“分类统计”的结果, 这分类统计收集了一种特别类型的状态, 即第 1 部份能量为 E_1 , 且第 2 部份能量是 E_2 的状态. 这“类”字是指 (E_1, E_2) 而言, 每组 (E_1, E_2) 之值为一类. 当然 $E_1 + E_2 = E$, 因为这类状态是在活动范围内. (4) 式把这些分类统计的结果来一个“综合”, 把诸类状态数加起来, 得到活动范围总体积 $\Gamma(E)$. (4) 的积分是一个“综合统计”.

以上的计算指出, $\Gamma(E_1)\Gamma(E_2)$ 在 (\bar{E}_1, \bar{E}_2) 处为极大. 而且 (\bar{E}_1, \bar{E}_2) 附近的诸类状态的数目, 几乎是所有收集了的状态之总数. 所谓“附近”是指

$$\begin{aligned} E'_1 &= E_1 - \bar{E}_1 \sim \sqrt{K_1}, \\ E'_2 &= E_2 - \bar{E}_2 \sim \sqrt{K_2}. \end{aligned} \quad (12)$$

因为

$$\Gamma_1(E_1)\Gamma_2(E_2) \propto e^{-\frac{E_1^2}{2K_1} - \frac{E_2^2}{2K_2}}, \quad (12')$$

(见 (10)), 所以离 (\bar{E}_1, \bar{E}_2) 远的状态, 为数甚少, 也就是说 (4) 的积分是

$$\begin{aligned} \Gamma(E) &\approx \Gamma_1(\bar{E}_1)\Gamma_2(\bar{E}_2) \times \Delta E, \quad \Delta E = \sqrt{2\pi K}, \\ S(E) &= \ln\Gamma(E) \approx S_1(\bar{E}_1) + S_2(\bar{E}_2). \end{aligned} \quad (13)$$

(虽然 $\Delta E \sim \sqrt{N}$, 但 $S_1, S_2 \sim N$, 因此 $\ln\Delta E$ 可以略去).

也就是说, “综合”后的熵是“分类”熵的极大值. “综合”步骤事实上成了求极大值的步骤: 在分类统计结果中, 找出状态最多的一类. 这一类的熵, 就几乎是所有各类综合的熵, 也就是说:“整个活动范围可以说是被这一类的状态占满了”. 这个结果是够奇怪的了, 它可以是“超级大数”, 即 e^N 级大数的特性. 像 $\Gamma(E_1), \Gamma(E_2)$ 这种函数都是超级大函数, 变量一改, 这些超大函数的改变也是超大, 所以函数之各值真是大小悬殊, 其极大值之大比其他非极大值大得太多. 因此除了极大值之外, 其他都不重要.

这状态最多的一类, 正好是 $T_1 = T_2$ 的一类, 也就是两部份温度相同的一类. 其他各类对 $S(E)$ 的影响是微不足道的, 这个结论把平衡现象和统计力学的基本假设连起来了.

以上的讨论, 可以推广如下: 我们把物体分成 m 部份, 每部份的状态以 s_i 指名, 每一部份的熵为 $S_i(E_i)$, 能量函数为 $H_i(s_i)$, $S_i(E_i) = \ln\Gamma_i(E_i)$,

$$\Gamma_i(E_i) = \sum_{s_i} \delta(E - H_i(s_i)) \quad , \quad i = 1, 2, \dots, m. \quad (14)$$

物体的总熵是

$$\begin{aligned} S(E) &= \ln\Gamma(E) \\ \Gamma(E) &= \sum_{s_1} \sum_{s_2} \cdots \sum_{s_m} \delta(E - H_1(s_1) - H_2(s_2) - \cdots - H_m(s_m)), \end{aligned} \quad (15)$$

以上 (4) 至 (11) 的讨论, 可以用来得到

$$S(E) = \sum_{i=1}^m S_i(\bar{E}_i), \quad (16)$$

T 和 \bar{E}_i 是以下联立方程式之解

$$\begin{aligned} \frac{\partial S_i(E_i)}{\partial E_i} &= \frac{1}{T}, \quad i = 1, 2, \dots, m, \\ E &= E_1 + E_2 + \dots + E_m. \end{aligned} \quad (17)$$

读者可自己导出这结果.

当然, 分类统计不限于依能量分类, 以上的结果, 可以推广到任何运动变量, 只要这变量是 N 量级的大数. 现在讨论一个重要例子.

以上的物体各部份中的粒子数是看作一定的. 但如果各部份间有粒子的交流, 则各部份中的粒子数 $\hat{N}_1(s_1), \hat{N}_2(s_2), \dots, \hat{N}_m(s_m)$ 就成了运动变量. 假设其总数不变:

$$N = \sum_{i=1}^m N_i(s_i), \quad (18)$$

令

$$\begin{aligned} S_i(E_i, N_i) &= \ln \Gamma_i(E_i, N_i), \\ \Gamma_i(E_i, N_i) &= \sum_{s_i} \delta(E_i - H_i(s_i)) \delta(N_i - \hat{N}_i(s_i)). \end{aligned} \quad (19)$$

即不但按能量分类, 也同时按粒子数分类. 这是 (14) 的推广, 结论 (16), (17) 之推广为

$$S(E, N) = \sum_{i=1}^m S_i(\bar{E}_i, \bar{N}_i), \quad (20)$$

\bar{E}_i, \bar{N}_i 及 T, μ 是以下联立方程式之解

$$\begin{aligned} \frac{\partial S_i}{\partial E_i} &= \frac{1}{T}, \quad i = 1, 2, \dots, m. \\ \frac{\partial S_i}{\partial N_i} &= -\frac{\mu}{T}, \\ E &= E_1 + E_2 + \dots + E_m, \\ N &= N_1 + N_2 + \dots + N_m. \end{aligned} \quad (21)$$

因此, 各部份共有同一温度, 也共有同一化学势 μ , 进一步的讨论会指出各部份有相同的压力, 假如各部份的体积也是运动变量的话.

这些讨论指出: 从基本假设我们可以直截了当地导出了平衡状态的一些特性, 使我们对这假设有更多信心. 读者请注意一些不正确的口头禅, 如: 统计力学“证明熵是极大的”, “证明平衡的存在”. 事实上是假设了熵是极大 (分类统计比综合的小, 总熵是综合统计结果), 但这假设仍是个假设而已, 并没有证明.

现在来看些应用例子.

2. 振动

量子力学的讨论问题很多, 但有四样最简单的结果, 必须熟记: ① 自由粒子, ② 简谐振动, ③ 氢原子, ④ 自旋 $\frac{1}{2}$. 这些是最基本的常识, 自由粒子状态的统计在前面讲了不少, 以后会再碰到, 现在来看简谐振动.

一个固体的小幅度振动, 可以看作一群基本简谐振动单元的组合. 每个基本振动都有它的频率, 根据量子力学, 频率为 ω 的振动, 其能量为 ε_n

$$\varepsilon_n = \left(n + \frac{1}{2} \right) \omega, \quad (\hbar = 1), \quad n = 0, 1, 2, \dots \quad (22)$$

也就是说, 振动的形态可由整数 n 决定. 如果有 N 个基本振动单元, 则状态可由 N 个整数 $(n_1, n_2, \dots, n_N) \equiv s$ 来指名, 总能量是

$$H(s) = \sum_{i=1}^N \left(n_i + \frac{1}{2} \right) \omega_i, \quad (23)$$

现在来求熵

$$\begin{aligned} S(E) &= \ln \Gamma(E), \\ \Gamma(E) &= \sum_s \delta(E - H(s)), \\ &= \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \cdots \sum_{n_N=0}^{\infty} \delta \left(E - \sum_{i=1}^N \left(n_i + \frac{1}{2} \right) \omega_i \right). \end{aligned} \quad (24)$$

这些和不太好求. 我们现在把这 N 个振动单元分组, 每组的频率都差不多, 然后求每一组的熵, 再用上节的结果把各组的熵加起来便是.

令第 α 组有 N_α 个单元, 各单元频率都是 ω_α , 则这组的熵是

$$\begin{aligned} S_\alpha(E_\alpha) &= \ln \Gamma_\alpha(E_\alpha), \\ \Gamma_\alpha(E_\alpha) &= \sum_{n_1=0}^{\infty} \cdots \sum_{n_{N_\alpha}=0}^{\infty} \delta \left(E_\alpha - \sum_{i=1}^{N_\alpha} \left(n_i + \frac{1}{2} \right) \omega_\alpha \right). \end{aligned} \quad (25)$$

(25) 这些和比较容易求, 因为只有一个频率 ω_α , 令

$$M_\alpha \equiv \frac{E_\alpha}{\omega_\alpha} - \frac{1}{2}N_\alpha, \quad (26)$$

则 (25) 中的 δ -函数是[†]

$$\frac{1}{\omega_\alpha} \delta \left(M_\alpha - \sum_{i=1}^{N_\alpha} n_i \right). \quad (27)$$

假设有 N_α 个盒子, 如果把 n_i 想成在第 i 个盒子里放的球的数目, 球的总数必须是 M_α 才能被 (25) 统计. 因此, Γ_α 是把 M_α 个球放在 N_α 个盒子里的放法总数, (除以 ω_α):

$$\Gamma_\alpha(E_\alpha) = \frac{1}{\omega_\alpha} \frac{(M_\alpha + N_\alpha - 1)!}{M_\alpha! (N_\alpha - 1)!}. \quad (28)$$

读者可自己导出 (28). 这和上一章的计算球面积不大一样, 但目的和原则一样.

依照上节的结果, 总熵是各组熵之和的极大值, (见 (16), (17));

$$S(E) = \sum_\alpha S_\alpha(E_\alpha), \quad (29)$$

$$\frac{\partial S_\alpha}{\partial E_\alpha} = \frac{1}{T}, \quad E = \sum_\alpha \bar{E}_\alpha. \quad (30)$$

\bar{E}_α 和 T 须从 (30) 解出, 用大数定律, $\ln N! \approx N \ln N - N$, (28) 可以简化, 得:

$$\begin{aligned} S_\alpha(E_\alpha) &= \ln \Gamma_\alpha(E_\alpha) \\ &= (M_\alpha + N_\alpha) \ln(M_\alpha + N_\alpha) - M_\alpha \ln M_\alpha - N_\alpha \ln N_\alpha, \end{aligned} \quad (31)$$

$$\begin{aligned} \frac{\partial S_\alpha}{\partial E_\alpha} &= \frac{1}{\omega_\alpha} \frac{\partial S_\alpha}{\partial M_\alpha} \\ &= \frac{1}{\omega_\alpha} \ln \frac{M_\alpha + N_\alpha}{M_\alpha} \\ &= \frac{1}{T}. \end{aligned} \quad (32)$$

从 M_α 的定义, (26) 及 (32) 得

$$\bar{E}_\alpha = N_\alpha \left(\frac{1}{e^{\omega_\alpha/T} - 1} + \frac{1}{2} \right) \omega_\alpha. \quad (33)$$

[†] M_α 是看成一整数, (27) 式的 δ -函数的变量也是以整数看待, 我们定义 $\delta(M) = 1$, 如果 $M = 0$; $\delta(M) = 0$, 如果 $M \neq 0$, M 限为整数.

总能量和 T 的关系是由 (30) 决定, 即

$$E = \sum_{\alpha} N_{\alpha} \left(\frac{1}{e^{\omega_{\alpha}/T} - 1} + \frac{1}{2} \right) \omega_{\alpha}. \quad (34)$$

熵和温度的关系可由 (31)、(32) 和 (29) 得来:

$$S = \sum_{\alpha} N_{\alpha} \left(\frac{\omega_{\alpha}/T}{e^{\omega_{\alpha}/T} - 1} + \ln \left(\frac{1}{1 - e^{-\omega_{\alpha}/T}} \right) \right). \quad (35)$$

在许多应用上, 频率的分布很密, 以上二式的和可以用积分代替:

$$\begin{aligned} \sum_{\alpha} N_{\alpha} f(\omega_{\alpha}) &= \int d\omega g(\omega) f(\omega) V, \\ Vg(\omega) &= \sum_i \delta(\omega - \omega_i) = \sum_{\alpha} N_{\alpha} \delta(\omega - \omega_{\alpha}). \end{aligned} \quad (36)$$

$g(\omega)$ 是各振动单元的频率分布, 视物体而异, 在举例讨论 $g(\omega)$ 之先, 我们先指出零点能量的重要.

当 $T \rightarrow 0$ 时, 每个单元的能量是 $\frac{E_{\alpha}}{N_{\alpha}} = \frac{1}{2}\omega_{\alpha}$ (见 (33)), 这就是所谓“零点能量”, 这当然是十分重要的一项. 在高温时, 即 $\frac{\omega_{\alpha}}{T} \ll 1$ 时,

$$\begin{aligned} e^{\omega/T} - 1 &\approx \frac{\omega}{T} + \frac{1}{2} \left(\frac{\omega}{T} \right)^2 + O\left(\frac{\omega^3}{T^3} \right) \\ &= \frac{\omega}{T} \left(1 + \frac{\omega}{2T} \right) + O\left(\frac{\omega^3}{T^3} \right), \end{aligned} \quad (37)$$

代入 (33) 得

$$\begin{aligned} \frac{\bar{E}_{\alpha}}{N_{\alpha}} &= T - \frac{\omega_{\alpha}}{2} + O\left(\frac{\omega_{\alpha}^2}{T} \right) + \frac{\omega_{\alpha}}{2} \\ &= T + O\left(\frac{\omega_{\alpha}^2}{T} \right) \end{aligned} \quad (38)$$

这头一项 T 当然是上章讨论过的“能量均分”结果, 每单元的能量为 T . $O(\omega_{\alpha}^2/T)$ 这项在 T 很大时可略去. 注意零点能量仍是十分重要, 它把 $-\omega_{\alpha}/2$ 这项抵消了. 若非零点能量, 高温的能量将有一常数项出现.

3. 德拜模型

现在把以上的结果用到晶体振动上. 必要的工作是找出频率分布 $g(\omega)$. 一般晶体的构造十分复杂, 其振动单元的分析是专门题目. 在此我们只略提一下一个十分粗略的估计, 所谓“德拜模型”.

低频率的振动, 都是弹性波, 其频率和波长成反比. 令 k 为波数, 即 $2\pi/\text{波长}$, 则

$$\begin{aligned}\omega_{\parallel k} &= c_{\parallel} k, \\ \omega_{\perp k} &= c_{\perp} k.\end{aligned}\quad (39)$$

\parallel 指纵波, \perp 指横波. c_{\parallel}, c_{\perp} 为它们的传播速度. (39) 是弹性力学的结果, 只合用于波长很长 ($\frac{1}{k} \gg$ 晶格大小) 的振动.

$g(\omega)$ 的计算和第三章统计粒子位子的能量分布大同小异. 先解弹性波动方程式, 得平面波, (或驻立波, 视边界条件而定). 如果我们用循环边界条件, 则 \mathbf{k} 必须适合

$$\begin{aligned}\mathbf{k} &= \frac{2\pi}{L} (n_x, n_y, n_z), \\ n_x, n_y, n_z &= 0, \pm 1, \pm 2, \dots\end{aligned}\quad (40)$$

假设晶体是每边长 L 的立方体. 因此振动单元是由 3 个整数来指定. 频率分布是 (照 (36) 定义).

$$Vg(\omega) = \sum_{n_x} \sum_{n_y} \sum_{n_z} [\delta(\omega - \omega_{\parallel k}) + 2\delta(\omega - \omega_{\perp k})]. \quad (41)$$

横波有两种 (左旋, 右旋), 纵波只一种. 再把 (41) 的求和改为积分, 得

$$\begin{aligned}g(\omega) &= \frac{1}{(2\pi)^3} \int d^3k [\delta(\omega - c_{\parallel} k) + 2\delta(\omega - c_{\perp} k)] \\ &= \frac{1}{(2\pi)^3} 4\pi\omega^2 \left(\frac{1}{c_{\parallel}^3} + \frac{2}{c_{\perp}^3} \right).\end{aligned}\quad (42)$$

这只适用于低频率的振动, ω 很大时, (42) 显然没有对的理由. 不过我们如果只求粗略的估计, 则可以用 (42) 了事, 但还要注意:

$$\int g(\omega) d\omega = \frac{N}{V} \quad (43)$$

是个必要条件, (见 (36)). 如果晶体内有 N_A 个原子, 则 $N = 3N_A$, 因为每个原子有三个方向可以动, 为了适合 (43), 令 ω_D 为最高频率, 即

$$g(\omega) = 0, \quad \omega > \omega_D. \quad (44)$$

如果 $\omega < \omega_D$, 则用 (42) 为 $g(\omega)$, 由 (43) 得

$$\begin{aligned}\omega_D^3 &= \frac{2\pi^2 N}{V} c^3, \\ \frac{3}{c^3} &\equiv \frac{1}{c_{\parallel}^3} + \frac{2}{c_{\perp}^3}.\end{aligned}\quad (45)$$

因此, ω_D 可以由弹性波速测定, 这个 $g(\omega)$ 是为“德拜模型”. ω_D 叫做“德拜”频率, $g(\omega)$ 既定, 各种平衡性质可以从 (34)~(36) 及微分式求出, 最重要的例子是热容率 (见图 1)

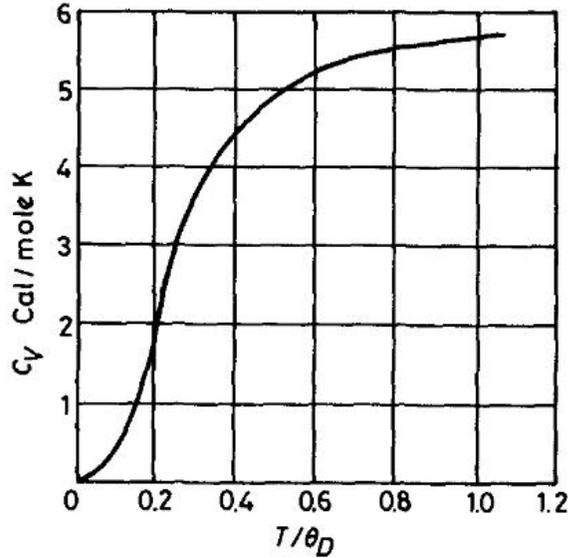


图 1, 德拜模型的热容率, 见 (46), (47)(图见 Kittel, (1966), P.179)

$$C_V = \frac{\partial E}{\partial T} = 3N_A f\left(\frac{\omega_D}{T}\right), \quad (46)$$

其中

$$f(y) = \frac{3}{y^3} \int_0^y dx \frac{x^4 e^x}{(e^x - 1)^2}, \quad (47)$$

读者可自导出. 注意我们用的单位是 $\hbar = k = 1$, 如果用一般单位, 则 $\frac{\omega_\alpha}{T}$, 该是 $\frac{\hbar\omega}{kT} \equiv \frac{\theta_D}{T}$. θ_D 是一般所谓“德拜温度”. 根据 (46), 所有固体的 $\frac{C_V}{N_A}$ 都是 $\frac{\theta_D}{T}$ 的同一函数. 这个结论和实验颇相符合. 符合程度由表 1 可见一斑. (注意: 各 θ_D 都和室温差不多).

表 1 德拜温度[‡]

固体	θ_D (从 C_V 得来) $^{\circ}\text{K}$	θ_D (从弹性得来) $^{\circ}\text{K}$
NaCl	308	320
KCl	230	246
Ag	225	216
Zn	308	305

表中右行是由弹性系数求出的 θ_D (见 (45)). 中行是由实验求 C_V , 再校正 θ_D , 尽量使 (46) 和实验结果相符合. 如果德拜模型是完全正确的话, 中行和右行该是一样, 事实上相差不远.

现在问题是: 这样马虎的模型, 为什么会如此成功? 主要原因是 ① $g(\omega)$ 的低频率部份并无误差. 因此当 $T \ll \theta_D$ 时, (46) 应该很正确. ② 在高温下, $T \gg \theta_D$, 只有 $g(\omega)$ 的积分, 即单元总数, 出现在 C_V 里, 因此在高温 (46) 也不坏. 高温低温都抓定了, 中间也差不到那里去.

图 2, 图 3 是一些比较正确的 $g(\omega)$ 计算结果.

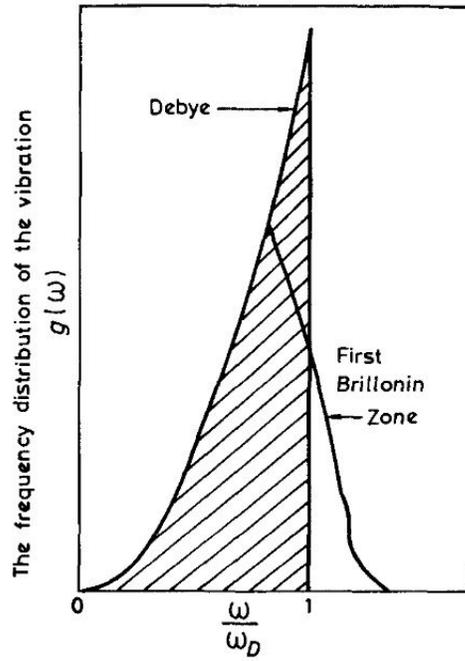


图 2. 振动频率分布. (数据仍然参考 Kittel (1966), P.177) 这图中的细线是由 (42) 计算, 但 k 的积分是取在第一布里渊区, 晶体为立方. (方敬尧图)

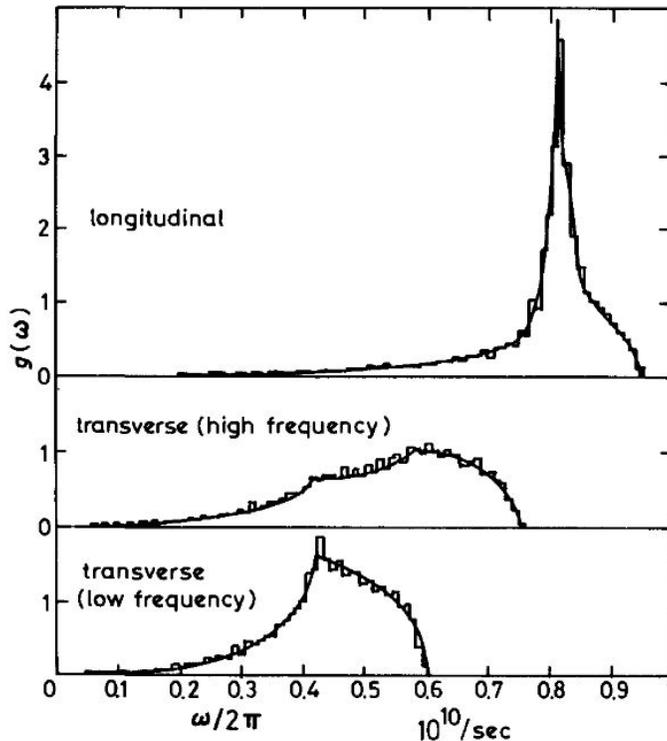


图 3, 铝的振动频率分布, 从 2791 个波向量统计得来. (见 Kittel(1966), P.177) 波向量空间的“第一布里渊区”的形状, 和晶体结构对称有关. 读者可参考固态物理的书籍.

4. 声子和第二声波

简谐振动的最大特点是: 其状态可以用一正整数 n_i 指定, 且此形态的能量是 $n_i\omega_i + \frac{1}{2}\omega_i$, ω_i 是此振动的频率. 因此, 把 n_i 想成某种“粒子”的数目, 是再自然不过了. 每个粒子有能量 ω_i . 上两节讲过, 一个固体的振动可以看成 $3N$ 个振动单元的运动. 因此它可以看作一群粒子的运动, 每个振动单元 i 是一个“位子”, n_i 是该位子中的粒子数. 这些粒子是合群的, 因为 n_i 可以大于 1, (见第三章对粒子位子及合群性的讨论). 如果固体是一晶体, 则振动单元为平面波, 平面波可以用波向量 \mathbf{k} 指定, 于是 $n_{\mathbf{k}}$ 是动量为 \mathbf{k} ($=\hbar\mathbf{k}$, $\hbar=1$) 的粒子数. 晶体的振动, 因此可以从一个波色粒子气体的角度来分析, 这些粒子叫做“声子”.

从这“声子气体”的分析,可以得出和振动有关的固体性质. 因为我们对气体比较熟悉, 所以如果一个群体, 可以用一个气体代表, 我们就很高兴, 立刻把有关气体的现象统统搬出来, 用它们来推断这群体的特性.

有一个很普遍的现象, 就是“声波”. “声子”是弹性波, 可以算是“声波”, 不过我们现在有一个声子气体, 这气体中, 也该可以有声波, 这个声波是靠这声子气体为传播介质, 和声子本身不一样, 我们叫它“第二声波”. 现在来看这有趣的问题.

普通气体中的声波, 可以用流体力学方程式 (见第一章 (1.1)) 来解释, 现在复习一下. 假定声波振幅很小, (1.1) 可以简化成

$$\frac{\partial \rho'}{\partial t} + \rho \nabla \cdot \mathbf{v} = 0, \quad (48a)$$

$$\rho \frac{\partial \mathbf{v}}{\partial t} = -\nabla p = -\frac{dp}{d\rho} \nabla \rho'. \quad (48b)$$

此地, ρ' 是指质量密度的改变, ρ 指没有声波时的密度, 从这二式, 立刻可得波动方程式

$$\frac{\partial^2 \rho'}{\partial t^2} = c^2 \nabla^2 \rho', \quad (49)$$

$$c^2 = \frac{dp}{d\rho}. \quad (50)$$

c 是传播速度. 因为热的扩散远比能量的周期性改变要慢, 所以声波中的密度, 压强改变是绝热过程. 假定振动周期远比分子的平均自由时间长, 波长远比平均自由距离大, (见 (1.1) 后局部平衡的讨论), c^2 可用平衡性质量出

$$c^2 = \left(\frac{\partial p}{\partial \rho} \right)_S. \quad (51)$$

理想气体的绝热过程中, (见 (2.24), (2.27))

$$p \propto V^{-5/3} \propto \rho^{5/3}.$$

故理想气体中的声速为

$$c = \sqrt{\frac{5p}{3\rho}} = \sqrt{\frac{5T}{3m}}. \quad (52)$$

因此, 我们该先求出像 (48a) (48b), 但合用于声子气体的方程式. 但是, 我们的声子无质量可言. 再者, 声子的总数并非常数, 而是一直在变. 别忘了, 声子数只是代表振动单元的形态. 当形态改变, 声子数也改了. 因此, 声子气体没

有 (48a) 这样的东西. 不过每个声子有一定的能量, 我们可以写下一个能量守恒公式:

$$\frac{\partial \varepsilon}{\partial t} + \nabla \cdot \mathbf{J} = 0. \quad (53)$$

$\varepsilon(\mathbf{r}, t)$ 是能量密度, $\mathbf{J}(\mathbf{r}, t)$ 是能流量. (48b) 式是牛顿第二定律, 即动量改变率等于所受之力. 每个声子都有动量. 令 $\mathbf{q}(\mathbf{r}, t)$ 为动量密度, 所以

$$\frac{\partial \mathbf{q}}{\partial t} = -\nabla p. \quad (54)$$

从局部平衡的观点来看, ε , \mathbf{q} , \mathbf{J} 都是温度 T 和流速 \mathbf{v} 的函数. 要找出这些函数, 我们还得先做点准备工作.

现在暂时离开本题, 考虑一个动量为 \mathbf{P} 的物体, 假设这物体在自由运动中, 以致于 \mathbf{P} 不变. 照我们的假设, 此物之熵为

$$\begin{aligned} S(E, \mathbf{P}) &= \ln \Gamma(E, \mathbf{P}), \\ \Gamma(E, \mathbf{P}) &= \sum_s \delta(E - H(s)) \delta(\mathbf{P} - \mathbf{P}(s)). \end{aligned} \quad (55)$$

$\mathbf{P}(s)$ 是状态 s 的总动量. 现在又依前例, 把这物体看作 m 个部份之组合. 各部份之动量为 $\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_m$, (16), (17) 及 (20), (21) 的推广为

$$S(E, \mathbf{P}) = \sum_{i=1}^m S_i(\bar{E}_i, \bar{\mathbf{P}}_i). \quad (56)$$

$\bar{E}_i, \bar{\mathbf{P}}_i$ 及 T, \mathbf{v} 是以下联立方程式之解

$$\begin{aligned} \frac{\partial S_i}{\partial \bar{E}_i} &= \frac{1}{T}, \\ \frac{\partial S_i}{\partial \bar{\mathbf{P}}_i} &= -\frac{\mathbf{v}}{T}, \quad i = 1, 2, \dots, m, \end{aligned} \quad (57)$$

$$\begin{aligned} E &= E_1 + E_2 + \dots + E_m, \\ \mathbf{P} &= \mathbf{P}_1 + \mathbf{P}_2 + \dots + \mathbf{P}_m. \end{aligned} \quad (58)$$

也就是说, 物体各部份除了有共同的温度外, 还共有同一速度 \mathbf{v} . \mathbf{v} 可以解释成物体的速度, 原因如下: 我们可以用物体的重心坐标系为观测标准. 用这个标准, 总能量是 $E = \frac{P^2}{2M}$, M 是总质量, 而总动量为零, 熵仍一样, (因为熵是一个统计值, 和坐标系无关)

$$S(E, \mathbf{P}) = S\left(E - \frac{\mathbf{P}^2}{2M}, 0\right). \quad (59)$$

因此

$$\frac{\partial S}{\partial \mathbf{P}} = \frac{\partial S}{\partial E} \left(-\frac{\mathbf{P}}{M} \right) = -\frac{\mathbf{v}}{T}, \quad (60)$$

\mathbf{P}/M 是物体的速度 \mathbf{v} .

在 (57), \mathbf{v}/T 是一个“力”, 和 $1/T$ 类似.

现在我们把这些结果用到声子气体上, 假定有一团声子气体, 其总动量为 \mathbf{P} , 我们把这些声子分组, 令第 α 组包括动量差不多是 \mathbf{k}_α 且振动方向一样的声子 (同一组的声子有相同的频率, 因为频率是由波向量和振动方向决定.) 这组的熵是

$$S_\alpha(\mathbf{P}_\alpha) = \ln \Gamma_\alpha(\mathbf{P}_\alpha),$$

$$\Gamma_\alpha(\mathbf{P}_\alpha) = \sum_{n_1=0}^{\infty} \cdots \sum_{n_{N_\alpha}=0}^{\infty} \delta \left(\mathbf{P}_\alpha - \sum_{i=1}^{N_\alpha} n_i \mathbf{k}_\alpha \right), \quad (61)$$

N_α 是这组声子位子的总数. n_i 是第 i 个位子中的声子数, \mathbf{P}_α 是这组声子的总动量. 注意, 因为每个声子动量都是 \mathbf{k}_α 且振动方向一致, \mathbf{k}_α 定了, ω_α 也定了, 所以一旦 \mathbf{P}_α 定了, 能量 E_α 也定了, 用不着再用一个 δ -函数

$$\delta \left(E_\alpha - \sum_{i=1}^{N_\alpha} \left(n_i + \frac{1}{2} \right) \omega_\alpha \right)$$

来限制能量, 现在令

$$\mathbf{P}_\alpha = M_\alpha \mathbf{k}_\alpha, \quad (62)$$

(61) 的 δ 函数即可以用

$$a \delta \left(M_\alpha - \sum_{i=1}^{N_\alpha} n_i \right), \quad (63)$$

代替[§], $S_\alpha(\mathbf{P}_\alpha)$ 就和 (31) 一样了:

$$S_\alpha(\mathbf{P}_\alpha) = (M_\alpha + N_\alpha) \ln(M_\alpha + N_\alpha) - M_\alpha \ln M_\alpha - N_\alpha \ln N_\alpha, \quad (64)$$

[§]注意 a 是一个不重要的常数, 取了对数后可以略去.

$$\delta \left(\mathbf{P}_\alpha - \sum_{i=1}^{N_\alpha} n_i \mathbf{k}_\alpha \right) = (k_{\alpha x}, k_{\alpha y}, k_{\alpha z})^{-1} \delta \left(M_\alpha - \sum_{i=1}^{N_\alpha} n_i \right),$$

$$\delta \left(E_\alpha - \sum_{i=1}^{N_\alpha} \left(n_i + \frac{1}{2} \right) \omega_\alpha \right) = \frac{1}{\omega_\alpha} \delta \left(M_\alpha - \sum_{i=1}^{N_\alpha} n_i \right).$$

$\ln a$ 可以略去, 因为 M_α, N_α 都是大数. M_α 当然就是在 α 组内的声子总数.

把这些分组结果加起来, 须要用到 (56) (57). 因为

$$\begin{aligned} E_\alpha &= \left(M_\alpha + \frac{1}{2} N_\alpha \right) \omega_\alpha, \\ \mathbf{P}_\alpha &= M_\alpha \mathbf{k}_\alpha, \end{aligned} \quad (65)$$

所以我们可以把 (57) 改写成

$$\begin{aligned} dS_\alpha &= \frac{1}{T} dE_\alpha - \frac{\mathbf{v}}{T} \cdot d\mathbf{P}_\alpha \\ &= \frac{1}{T} (\omega_\alpha - \mathbf{k}_\alpha \cdot \mathbf{v}) dM_\alpha, \end{aligned}$$

即

$$\begin{aligned} \frac{dS_\alpha}{dM_\alpha} &= \frac{1}{T} (\omega_\alpha - \mathbf{k}_\alpha \cdot \mathbf{v}) \\ &= \ln \frac{M_\alpha + N_\alpha}{M_\alpha}, \end{aligned} \quad (66)$$

这和 (32) 差不多, 只是 ω_α 改成了 $\omega_\alpha - \mathbf{k}_\alpha \cdot \mathbf{v}$, 令

$$\bar{n}(\omega) \equiv \frac{1}{e^{\omega/T} - 1}, \quad (67)$$

则从 (66) 解出 M_α 得

$$\frac{M_\alpha}{N_\alpha} = \bar{n}(\omega_\alpha - \mathbf{k}_\alpha \cdot \mathbf{v}), \quad (68)$$

(57) 的最后两式是

$$\begin{aligned} E &= \sum_\alpha N_\alpha \bar{n}(\omega_\alpha - \mathbf{k}_\alpha \cdot \mathbf{v}) \omega_\alpha + \sum_\alpha \frac{1}{2} N_\alpha \omega_\alpha, \\ \mathbf{P} &= \sum_\alpha N_\alpha \bar{n}(\omega_\alpha - \mathbf{k}_\alpha \cdot \mathbf{v}) \mathbf{k}_\alpha. \end{aligned} \quad (69)$$

能流量是能量的流量, 每个声子有能流量 $\frac{\omega_\alpha c \hat{\mathbf{k}}_\alpha}{V}$, $c \hat{\mathbf{k}}_\alpha$ 是流速, $\hat{\mathbf{k}}$ 是 \mathbf{k} 的方向, 因此总能流量是

$$\mathbf{J} = \sum_\alpha N_\alpha \bar{n}(\omega_\alpha - \mathbf{k}_\alpha \cdot \mathbf{v}) \omega_\alpha \frac{c \hat{\mathbf{k}}_\alpha}{V}. \quad (70)$$

现在以 $\frac{3V}{(2\pi)^3} \int d^3k$ 代替 $\sum_\alpha N_\alpha$, 从 (69) 和 (70) 得能量密度 E/V , 动量密度

P/V , 及能量流 \mathbf{J}

$$\begin{aligned}\varepsilon &= \frac{E}{V} = \frac{3}{(2\pi)^3} \int d^3k \bar{n}(\omega_k - \mathbf{k} \cdot \mathbf{v}) \omega_k, \\ \mathbf{q} &= \frac{P}{V} = \frac{3}{(2\pi)^3} \int d^3k \bar{n}(\omega_k - \mathbf{k} \cdot \mathbf{v}) \mathbf{k}, \\ \mathbf{J} &= \frac{3}{(2\pi)^3} \int d^3k \bar{n}(\omega_k - \mathbf{k} \cdot \mathbf{v}) \omega_k c \hat{\mathbf{k}}, \\ \bar{n}(\omega) &= \frac{1}{e^{\omega/T} - 1},\end{aligned}\tag{71}$$

各式中的 3 是指三个振动方向. 我们没有计入纵横波频率的差异, 是为了避免式子的繁琐. 零点能量也略去了, 因为它不会影响以下的讨论.

因为 \bar{n} 是 T 的函数, ε , \mathbf{J} , 和 \mathbf{q} 就成了 T 和 \mathbf{v} 的函数, 现在我们把 T 和 \mathbf{v} 想成位置 \mathbf{r} 和时刻 t 的函数, $T = T(\mathbf{r}, t)$, $\mathbf{v} = \mathbf{v}(\mathbf{r}, t)$. 因此 ε , \mathbf{q} , 和 \mathbf{J} 就成了 \mathbf{r} 和 t 的函数了. 我们把它们用到 (53) 和 (54) 来讨论声子气体的运动. 注意: 别忘了此地 \mathbf{r} 和 t 是宏观尺度, dt , $d\mathbf{v}$ 都是比声子运动尺度要大很多的量, 这些尺度还须要讨论的.

用 (71), 我们可以把 (53) 和 (54) 化成 T 和 \mathbf{v} 的方程式, 因为 \mathbf{v} 是小量, 所以

$$\bar{n}(\omega - \mathbf{k} \cdot \mathbf{v}) \approx \bar{n}(\omega) - \mathbf{k} \cdot \mathbf{v} \frac{\partial \bar{n}}{\partial \omega},\tag{72}$$

把它带入 (71) 得

$$\begin{aligned}\mathbf{J} &= \frac{3}{(2\pi)^3} \int d^3k \omega c \hat{\mathbf{k}} \left(-\mathbf{k} \cdot \mathbf{v} \frac{\partial \bar{n}}{\partial \omega} \right) \approx c^2 \rho \mathbf{v}, \\ \mathbf{q} &\approx \rho \mathbf{v}, \\ \rho &\equiv \frac{3}{(2\pi)^3} \int d^3k \frac{k^2}{3} \left(-\frac{\partial \bar{n}}{\partial \omega} \right).\end{aligned}\tag{73}$$

以上 ω 以 ck 代, ρ 可以解释成声子气体的“质量密度”, 还有

$$\begin{aligned}\frac{\partial \varepsilon}{\partial t} &= \frac{\partial \varepsilon}{\partial T} \frac{\partial T}{\partial t}, \\ \nabla p &= \frac{1}{3} \nabla \varepsilon = \frac{1}{3} \frac{\partial \varepsilon}{\partial T} \nabla T, \\ \frac{\partial \varepsilon}{\partial T} &= \frac{3}{(2\pi)^3} \int d^3k \frac{\omega^2}{T} \left(-\frac{\partial \bar{n}}{\partial \omega} \right) = \frac{3c^2}{T} \rho.\end{aligned}\tag{74}$$

以上用到 $p = \frac{1}{3}\varepsilon$, $\frac{\partial \bar{n}}{\partial \omega} = -\frac{\omega}{T} \frac{\partial n}{\partial \omega}$. 读者可自证实把 (73), (74) 代入 (53) 及 (54) 简化得

$$\begin{aligned}\frac{\partial T}{\partial t} &= -\frac{1}{3}T\nabla \cdot \mathbf{v}, \\ \frac{\partial \mathbf{v}}{\partial t} &= -c^2\nabla T.\end{aligned}\quad (75)$$

下一步自然是最简单不过了

$$\frac{\partial^2 T}{\partial t^2} = c_2^2 \nabla^2 T, \quad c_2 = \frac{c}{\sqrt{3}}. \quad (76)$$

我们终于找到了声波方程式, 是为“第二声波”, 其声速是声子速度的 $\frac{1}{\sqrt{3}}$. 但是, 这结果只是分析的简单部份. 困难部份还没开始, 即各种假设的可靠性的分析. 如果有不可靠的假设, 则 (76) 也不可靠了. 现在看看以上的讨论中假设了些什么:

- ① 各声子的动量, 必须不停地变动. 如果没有这些变动, 我们的基本假设就不能用. 变动的由来有
 - (a) 振动不完全是简谐的, 其结果是各振动单元间的相互作用, 使各个位子的声子数不断地改变.
 - (b) 晶体中的各种缺陷, 杂质原子也产生类似效果.

我们假设: 这些作用一方面够强, 可维持平衡现象, 一方面够弱, 不影响以上的统计步骤.

- ② 各声子动量虽变, 但总动量要守恒 (如果没有压强坡度的话). 从波动的观点来看, 动量守恒的原因是介质均匀. 真空是均匀的, 所以有动量守恒定律. 晶体不是均匀介质, 所以声子动量不守恒. 但是晶体中的原子排列是周期性, 周期为晶格大小. 因此, 对长波声子而言, 介质是均匀的, 这短距离的周期性变化不会发生什么作用. 因此, 以上的 (a) 项变化, 不会影响一群长波声子的总动量. 但是短波声子如果多了, 动量就不守恒了. (54) 就有了问题. 短波的是能量高的. 因此温度要很低

$$T \ll \theta_D \quad (77)$$

短波声子才会少, (54) 才可能正确. 不过 T 小, 则长波声子也少 (声子总数 $\propto T^3$). 气体太稀, 也无法有声波. 以上的 (b) 项变化, 即杂质缺陷,

引起的改变, 就完全不遵守动量守恒. 因此, 只有非常完美无瑕的晶体, 才能使 (54) 成立.

因此第二声波只能在温度不高也不低, 完美晶体内存在. 完美的晶体要做的大, 很不容易. 但声波的速度大 (10^5 cm/秒左右), 波长在 1cm, 即有频率每秒 10^5 次. 频率这样高, “局部平衡” 就须要快速声子反应才能维持. 所以反应速度也得小心考虑. 这个问题因此是相当的复杂, 超出了此地的讨论范围. 我们只提一下, 根据报导, 在 10°K 左右, 像 CdS 这样的晶体可以维持第二声波[¶].

因为这种第二声波不容易存在, 所以平时不受重视. 不过它的教育价值很高. 以上的讨论是从非常抽象的基本假设, 一直讲到声子气体内的波动, 确是一个有用的示范例子.

5. 一般结论, 平均值和正态分布

这章的讨论重点, 在总体和部份的关系. 平衡下的物体各部的各种“力”都一样, 即有相同的温度, 压强, 化学势. 整个物体若是在动, 则各部有相同的速度 (如果都在转动, 则有相同的转速, 读者可试试从基本假设引出). 这都是基本假设的后果, 这种力的均衡, 是平衡的特色.

我们注意到, 每个力都和一个不变量有关: 温度和能量, 压力和体积, 化学势和粒子数, 速度和动量. 一旦这些力不均匀, 则会有这些不变量的流动. 声波的分析是一个示范例子. 缓慢大尺度的改变, 多是和这些不变量有关的.

物体的各部份, 不一定要按照所占的空间区域来划分. 以上对振动的讨论指出: 我们可以依照振动单元的能量或动量来分组. 每一组算是物体的一部分. 不过, 分部而后个别计算熵的方法, 只有在各部大致独立时, 才有用. 各部独立的意思只是各部运动互不相关, 熵可以个别统计. 这独立性是个十分复杂的问题, 我们以后会仔细讨论.

从统计观点看来, 这一章谈的莫非是分类和综合步骤, 乃大数值函数的一些特色. 以上的结果, 不一定要和能量、粒子数之类的不变量有关. 我们可以按任何宏观量的值分类. 现在把这一点说明如下:

[¶]见 Prohofskey and Krumhansl, (1964) 及文中所列举资料.

令 $\hat{A}(s)$ 为一个 N 量级的运动变量, 令

$$\begin{aligned}\Gamma(E, A) &= \sum_s \delta(E - H(s))\delta(A - \hat{A}(s)) \\ &\equiv e^{S(E, A)},\end{aligned}\quad (78)$$

这是一个分类统计的结果. 每一个 A 值成一类. 在能量为 E 的诸状态中, 收集 \hat{A} 值为 A 者而计其数目. $\Gamma(E, A)$ 当然就是各状态的 (E, A) 分布. $\Gamma(E)$ 可以由 $\Gamma(E, A)$ 综合得来,

$$\begin{aligned}\Gamma(E) &= \int dA \Gamma(E, A) \\ &= \int dA e^{S(E, A)},\end{aligned}\quad (79)$$

这积分又可用尖峰法, 先求 $S(E, A)$ 之极大值:

$$\frac{\partial S(E, A)}{\partial A} = 0. \quad (80)$$

令其解为 $\bar{A}(E)$. 如果 \bar{A} 为极大值, 则

$$(A - \bar{A})^2 \left(\frac{\partial^2 S}{\partial A^2} \right)_E \equiv -\frac{(A - \bar{A})^2}{K_A} < 0 \quad (81)$$

为必要条件, 由 (15) 得

$$\begin{aligned}\Gamma(E) &\approx e^{S(E, \bar{A})} \int dA' e^{-A'^2/2K_A} \\ &= e^{S(E, \bar{A})} \sqrt{2\pi K_A},\end{aligned}\quad (82)$$

$$\begin{aligned}S(E) &= \ln \Gamma(E) \\ &= S(E, \bar{A}(E)).\end{aligned}\quad (83)$$

【例】 $\hat{A}(s) = H_1(s) - H_2(s)$

此地 H_1, H_2 是物体的第 1 部份和第 2 部分的能量 (见 (1), (2)), 在本章开始时分析过.

$$\begin{aligned}\Gamma(E, A) &= \sum_s \delta(E - H(s))\delta(A - \hat{A}(s)) \\ &= \sum_{s_1, s_2} \delta(E - H_1(s_1) - H_2(s_2)) \delta(A - H_1(s_1) + H_2(s_2)) \\ &= \sum_{s_1, s_2} \delta\left(\frac{E + A}{2} - H_1(s_1)\right) \delta\left(\frac{E - A}{2} - H_2(s_2)\right) \\ &= \Gamma_1\left(\frac{E + A}{2}\right) \Gamma_2\left(\frac{E - A}{2}\right),\end{aligned}\quad (84)$$

$\Gamma_1(E_1), \Gamma_2(E_2)$ 的定义见 (4'), 因此

$$\begin{aligned} S(E, A) &= \ln \Gamma(E, A) \\ &= S_1 \left(\frac{E+A}{2} \right) + S_2 \left(\frac{E-A}{2} \right), \end{aligned}$$

(80) 的结果是

$$\begin{aligned} \frac{\partial S}{\partial A} &= \frac{1}{2} \left(\frac{\partial S_1}{\partial E_1} - \frac{\partial S_2}{\partial E_2} \right) = 0, \\ E_1 &= \frac{1}{2}(E+A), \\ E_2 &= \frac{1}{2}(E-A). \end{aligned} \quad (85)$$

(85) 当然是和 (6), (7) 完全一样, $\bar{A} = \bar{E}_1 - \bar{E}_2$, 是 (85) 之解, \bar{E}_1, \bar{E}_2 是 (6), (7) 之解.

以上的分析说明: 从基本假设, 我们可以计算出任一个宏观变量 $\hat{A}(s)$ 的值 \bar{A} . 计算步骤是分类统计得熵, 而后求极大值 (见 (80)). 各状态的 A 分布总是一个正态分布,

$$\frac{\Gamma(E, A)}{\Gamma(E)} = \frac{1}{\sqrt{2\pi K_A}} e^{-(A-\bar{A})^2/2K_A}. \quad (86)$$

因此, \bar{A} 也可以写成

$$\begin{aligned} \bar{A} &= \int dA \frac{\Gamma(E, A)}{\Gamma(E)} A \\ &= \frac{1}{\Gamma(E)} \sum_s \delta(E - H(s)) \hat{A}(s), \end{aligned} \quad (87)$$

即 \bar{A} 是 $\hat{A}(s)$ 在活动范围中各状态的平均值. 原因是: 几乎整个活动范围中的 $\hat{A}(s)$ 值都是 $\bar{A}(1 + O(\frac{1}{\sqrt{K_A}}))$, 而 $\frac{1}{\sqrt{K_A}} \sim \frac{1}{\sqrt{N}}$, 是个很小的数目.

现在再回头来温习一下活动范围的意义, 它是状态空间里的一个区域, 运动轨迹散布其中. 它是一个由轨迹撑起来的区域. 以上的结论是: 活动范围中差不多每个地方的 $\hat{A}(s)$ 值都是 \bar{A} . 既然轨迹散布于这区域中, 我们不难想见, 在轨迹上差不多每一点, $\hat{A}(s)$ 值也都是 \bar{A} .

一般实验所测定的, 都是观测时间内一些宏观变量的时间平均值:

$$\langle \hat{A} \rangle = \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} dt \hat{A}(s(t)). \quad (88)$$

$s(t)$ 是状态空间中的轨迹, 代表详细的分子运动, 观测时间是从 $t = 0$ 到 $t = \mathcal{T}$. 既然 $\hat{A}(s(t))$ 差不多总是 \bar{A} , 我们的结论是

$$\langle \hat{A} \rangle = \bar{A}. \quad (89)$$

因此, 除了求熵之外, 基本假设对计算 $\langle \hat{A} \rangle$ 也有明显的指定. 只要知道了活动范围, \bar{A} 就可以算出了.

读者或会问: 既然几乎每一状态的 $\hat{A}(s)$ 值都是 \bar{A} , 何不随意取一状态, 算出 \bar{A} 便是? 何必去平均? 这问题牵涉甚多, 原则上是可以, 但在实际计算上, “随意” 取一状态是个不简单的问题, 平均值的计算仍属必要, 以后会详细讨论.

注意: 我们还未讨论如何从轨迹来决定活动范围. 这是个难题. 以后会讨论. 现在活动范围的划定是假设的一部分, 即: 凡是不变量容许的状态, 都是包括在范围内. 因此, 所有以上的结论, 都只是假设的延伸而已.

$\langle \hat{A} \rangle = \bar{A}$ 这句话也可以表达如下:

宏观运动变量的时间平均值, 等于其在活动范围内的平均值.

读者请务必注意: ① 这个结论只能合用于宏观变数, 即 N 量级变量, ② 这结论不是说轨迹在活动范围中每一点花同样的时间. 因为几乎整个活动范围中所有地方的 \hat{A} 值都是 \bar{A} , 轨迹只要不是 “太离谱”, 只要不到那些 “几乎整个” 之外的地方呆着, $\langle \hat{A} \rangle = \bar{A}$ 就可以成立. 因此 $\langle \hat{A} \rangle = \bar{A}$ 并不是一个令人惊异的结论. 它并没有对轨迹作苛刻的要求.

把活动范围分成 M 个大区域, M 不是很大的数字, 例如 $M \sim \sqrt{N}$. 令每区域体积一样. 如果要求轨迹在每个大区域内花的时间大致一样, $\langle \hat{A} \rangle = \bar{A}$ 就可以舒舒服服地成立了. 这个要求, 和要求轨迹通过每一点, 是完全不同的.

虽然这要求 “并不苛刻”, 但轨迹是否适合这要求, 仍是一个问题. 如果轨迹偏要缩到 $\hat{A}(s) \neq \bar{A}$ 的角落里去, 我们只有承认我们搞错了活动范围, 或基本假设不对.

从这一节和上一节, 我们可以看到, 基本假设对熵、对平衡现象、对宏观变量的平均值, 大致都有很圆满的解释. 这样简单的假设, 居然有此威力, 确是很令人难以理解.

最后, 我们略提一下正态分布的意义. 像 (86) 这样地分布一再地出现, 原因何在? 略看我们的推演, 就知道这是 A 为 N 级大数的结果. N 级大数都是由许多分子运动变量加起来的. 而这些分子变量大致是 “独立” 的, 因为物体可以分成许多大致独立的部份. 以后我们会讲 “中心极限定理”, 这定理大致

是说: 如果 A 是很多大致独立变量之和, 则 A 的分布是正态分布. 因此, 宏观变量的正态分布, 可以说是物体各部大致独立的结果.

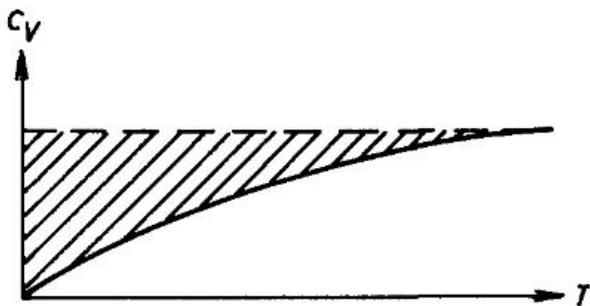
讨论问题六

1. 利用分组求熵和综合方法 (21), 求出 T, μ 和 N, E 的关系, 假设
 - (A) 物体是 N 个波色粒子,
 - (B) 是 N 个费米粒子.
2. 导出 (46), (47), 并证明在低温, 高温

$$\begin{aligned} C_V &\approx \frac{12\pi^4}{5} N_A \left(\frac{T}{\theta_D}\right)^3, & T \ll \theta_D, \\ C_V &\approx 3N_A, & T \gg \theta_D. \end{aligned} \quad (90)$$

即低温的 T^3 定律, 高温的“爱因斯坦定律”. 注意, 金属电子的热容率是和 T 成正比, 见第四章.

3. 图标某晶体的热容率



试证: 斜线部分面积是零点振动能量, 有 N 个原子组成一个一维空间的“晶体”, 求振动的频率分布 $g(\omega)$.

4. 热辐射, (即黑体辐射) 的理论和晶体振动理论颇有雷同之处. (见第三章, 第 4 节). 试作详细比较.
5. 热辐射是一光子气体, 在这气体内, 可不可以有声波的现象? 试作详细分析.

6. 第 4 节的计算稍为繁杂, 不过对基本概念和技巧的了解, 应颇有帮助. 读者应仔细推演一次, 注意每一细节. 第二声波先在氦液中出现, 读者可参考吴家玮 (1972), Wilks (1967).
7. 第 5 节的重点在数量级, 即 $S = O(N)$, $\Gamma = O(e^N)$. 若非这样的数量级, 基本假设就不能发生效力. 这数量级的要求, 并不简单. 如果有长程的相互作用, 则此种数量级就不一定成立.

试解下列模型:

$$H = -\frac{J}{2} \sum_{i,j} s_i s_j, \quad i, j = 1, 2, \dots, N \quad (91)$$

$$s_i = \pm 1.$$

此所谓“长程作用易形模型”. 每单元 s_i 和所有其他单元都有作用. “易形模型”的一般讨论在第十七章.

提示: 把 H 写成 $-\frac{1}{2}JM^2$, M 为所有 s_i 之和, 即 $\Gamma(E)$ 可立即算出.

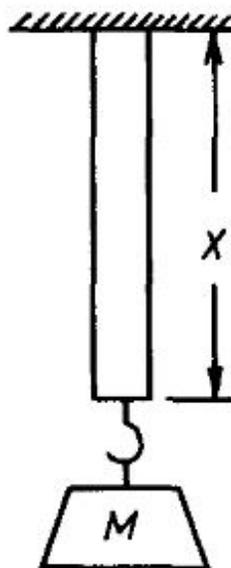
分析此模型各量的数量级. 将问题 (5.2) 和本题作一比较.

第七章 计算法则

基本假设就是一个计算熵的法则。这个法则，可以改写成许多不同的新计算法则。这些新计算法则的含义和后果，都和基本假设一样，只是计算方法有所改变。这些新法则，先计算某些“热位能”，再由热位能计算熵和其他平衡性质。在许多情形下，这样的计算比直接计算熵要简便。我们先讲各式热位能定义和意义，再从基本假设导出各式计算法则。

1. 各种热位能

在以上讨论中的“能量”，都是指物体内部分子运动，和相互作用的能量，或称“内能”。现在，我们可以把物体和外力作用有关的能量加入，而定出各式各样的“热位能”，或称“势能”或称“自由能”，及其他怪名字。把不同的外力加入，会得到不同的热位能。现在来看个例子。



【例一】

图 1 中的铁条子被一外力 f 拉长. 这外力用一锤子 (质量 M) 供给, $f = Mg$. g 是重力加速度. 令 X_0 为无拉力时铁条的长度.

铁条中的内能大约是

$$E(S, X) = E(S, X_0) + \frac{1}{2}K(S)(X - X_0)^2. \quad (1)$$

S 是熵, 其他的坐标都未标明. $K(S)$ 是这铁条的绝热弹性系数. 在平衡状态下, 外力 f 和内在的张力相等:

$$f = K(S)(X - X_0) = \left(\frac{\partial E}{\partial X}\right)_S. \quad (2)$$

现在我们把锤子的位能加上, 得

$$\begin{aligned} H(S, X) &= E(S, X) - MgX \\ &= E(S, X) - fX, \end{aligned} \quad (3)$$

也就是把坐标 X 和环境的作用能 $-fX$ 加入. 平衡状态下,

$$\left(\frac{\partial H}{\partial X}\right)_S = 0, \quad (4)$$

即 H 为极小值. 这当然和 (2) 完全一样. 在力学中, 我们习惯于把平衡想成能量的极小状态. 所以 (4) 比 (2) 好看些. 如果没挂锤子, 讨论其他性质时, $E(S, X)$ 比较方便, 如果挂了锤子, $H(S, X)$ 比较合用.

以上我们没考虑铁条子四周的空气, 这是另一环境因素. 如果空气的温度一定, 我们可以定义

$$F(S, X) = E(S, X) - TS, \quad (5)$$

T 是空气的温度, 它的作用和锤子重力 f 的作用相似. 此地的坐标 S 和 (3) 的 X 相似. 在平衡状况下, 铁条子的温度 $\frac{\partial E}{\partial S}$ 和空气温度 T 一样, 即

$$\begin{aligned} \left(\frac{\partial E}{\partial S}\right)_X &= T \quad \text{或} \\ \left(\frac{\partial F}{\partial S}\right)_X &= 0. \end{aligned} \quad (6)$$

F 是一极小值, 和 (4) 相似. 当然, 我们可以定出另一个热位能:

$$G(S, X) = E(S, X) - TS - fX \quad (7)$$

把两个外力 T, f 和坐标 S, X 的作用能都包括进去. 在一定 T 和 f 的环境下, G 为极小.

我们统称 H, F, G 为“热位能”. H 可以叫做“定拉力”热位能. F 是“定温”热位能. G 是“定拉力又定温”热位能.

【例二】

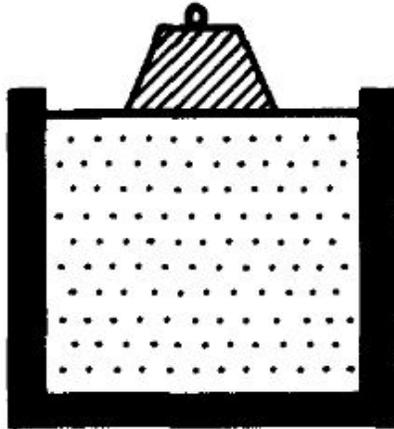


图 2

图 2 中的气体被压在一气缸中, 底面积为 A , 高为 y . 压强 p 由一活塞上的锤子供给.

$$p = Mg/A, \quad Mg \text{ 为锤子重量.}$$

循上例, 读者可自定出热位能:

$$\begin{aligned} H &= E + Mgy \\ &= E + pV. \end{aligned} \tag{8}$$

式中 $V = Ay$. 此外其他各热位能

$$\begin{aligned} F &= E - TS, \\ G &= E + pV - TS, \\ \Omega &= E - \mu N - TS, \end{aligned} \tag{9}$$

也可依同理定出.

这些热位能各有特别的微分式. 例如 (3) 的 H 的微分是

$$\begin{aligned} dH &= dE - fdx - Xdf \\ &= TdS - Xdf. \end{aligned} \quad (10)$$

因为 $dE = TdS + fdx$, (5) 的 F 的微分是

$$dF = -SdT + fdx. \quad (11)$$

因此, H 可看作 S 和 f 的函数, $H = H(S, f)$. F 可看作 T 及 X 的函数, $F = F(T, X)$. 所以如果一定要把 T 看成一个“坐标”, 则 $-S$ 以“力”的姿态出现. 把 f 看成“坐标”, $-X$ 看成“力”, 从数学式子看来, 不是很离谱的.

我们把 (9) 的各热位能的微分列于表 1

表 1 热位能的微分

	定 义	微 分
E	E	$dE = TdS - pdV + \mu dN$
H	$E + pV$	$dH = TdS + Vdp + \mu dN$
F	$E - TS$	$dF = -SdT - pdV + \mu dN$
G	$E - TS + pV$	$dG = -SdT + Vdp + \mu dN$
Ω	$E - TS - \mu N$	$d\Omega = -SdT - pdV - Nd\mu$

各种偏微分式可以自表中直接读出, 如

$$\begin{aligned} S &= -\left(\frac{\partial F}{\partial T}\right)_{V,N}, & N &= -\left(\frac{\partial \Omega}{\partial \mu}\right)_{T,V}. \\ V &= \left(\frac{\partial G}{\partial p}\right)_{T,N}, & T &= p\left(\frac{\partial V}{\partial S}\right)_{E,N}. \\ \mu &= -T\left(\frac{\partial S}{\partial N}\right)_{E,V}, & p &= -\left(\frac{\partial \Omega}{\partial V}\right)_{T,\mu}. \end{aligned} \quad (12)$$

这些微分式, 在计算时十分有用.

2. 各式计算法则

(甲): 定能法则和定温法则

基本假设中的“活动范围”是状态空间中一个有一定能量的区域, 域中的每一点的能量都是 E . 熵的计算可以缩写成

$$\begin{aligned}\Gamma(E) &= \sum_s \delta(E - H(s)), \\ S(E) &= \ln \Gamma(E),\end{aligned}\tag{13}$$

式中的 s 指各状态. $H(s)$ 是状态 s 的能量. 这求和的计算是一个统计的过程. 式中的 $\delta(E - H)$ 的目的是限制能量, 我们只收集能量为 E 的状态, 统计其总数. 这个统计步骤, 叫做“定能法则”. 在应用上, 这法则并不很方便. 我们要把它修改一下. 这一节讲的就是各种修改方法.

现在把 (13) 中的 $\delta(E - H)$ 改成

$$\Delta(E - H) \equiv e^{-\beta(H-E)},\tag{14}$$

令

$$\bar{\Gamma}(E) = \sum_s \Delta(E - H).\tag{15}$$

$\beta > 0$, 其值待稍后再定. 注意: 在 (13), 我们只收集能量为 E 的状态, 并予以相同份量. 在 (15), 所有状态都被收集, 予以不同份量. 每状态的份量和 $e^{-\beta H(s)}$ 成正比. 表面上看来, 整个活动范围的划定在 (15) 是十分含糊, 似乎没有意义.

这个 $\Delta(E - H)$ 当然和 $\delta(E - H)$ 不一样. 所以 $\bar{\Gamma}$ 和 Γ 也不一样, 它俩的关系是

$$\begin{aligned}\bar{\Gamma}(E) &= \sum_s \int dE' \Delta(E - E') \delta(E' - H(s)) \\ &= \int dE' \Gamma(E') e^{-\beta(E' - E)} \\ &= \int dE' e^{S(E') - \beta(E' - E)}\end{aligned}\tag{16}$$

假设 $\partial S(E')/\partial E' > 0$, 则 $\Gamma(E')$ 随 E' 剧增. E' 为大数, 故 $e^{-\beta E'}$ 随 E' 剧减. 因此 (16) 的积分函数为一尖峰, 其极大值可由

$$\frac{\partial}{\partial E'} [S(E') - \beta(E' - E)] = 0\tag{17}$$

决定, 即 $\partial S(E')/\partial E' = \beta$. 现在选定

$$\beta = \frac{1}{T} \equiv \frac{\partial S(E)}{\partial E}.\tag{18}$$

如此, 则 (16) 的积分函数的极大值在 $E' = E$. 既然 $S(E')$, E' 都是大数, 我们可用前一章讨论大数定律所用的方法来求这积分. (见 (3.56) 至 (3.59)): 把 $S(E')$ 展开, 只留到二次项:

$$\bar{\Gamma}(E) \approx e^{S(E)} \int dE'' e^{-(E'')^2/2K} \quad (19)$$

$$E'' \equiv E' - E,$$

$$\frac{1}{K} \equiv -\frac{\partial^2 S}{\partial E^2} = \left(T^2 \frac{\partial E}{\partial T} \right)^{-1}. \quad (20)$$

因为 S , E , E' 都和 N 成正比, 所以 K 也是, 并且 (19) 的积分函数是一尖峰, 其宽为 $\sqrt{2\pi K} \propto \sqrt{N}$. 所以 E'' 积分的结果是 $\sqrt{2\pi K}$, 因此

$$S(E) = \ln \bar{\Gamma}(E) - \frac{1}{2} \ln(2\pi K), \quad (21)$$

$$\ln \bar{\Gamma} \propto N, \quad \ln K \sim \ln N. \quad (22)$$

所以, 只要 N 很大, $\ln N$ 可以略去,

$$S(E) = \ln \bar{\Gamma}(E),$$

$$\bar{\Gamma}(E) = \sum_s e^{-(H(s)-E)/T}. \quad (23)$$

也就是说, $\Delta(E-H)$ 和 $\delta(E-H)$ 虽不同, 但只要 N 很大, 它们就有完全一样的结果. 原因是 $\Delta(E-H)$ 大致切去了高能的状态. 它虽不压抑低能量状态 (反而增加其份量), 但低能量的状态为数太少, 因此真正在 $\bar{\Gamma}(E)$ 里算数的是能量为 E 左右的各状态. 在 (23) 中 T 和 E 的关系还得透过 (18), 这不大方便. 因为温度 T 比较容易测定, 而能量 E 不容易, 所以最好把 E 消去, 把 T 留下. 令

$$F = -T \ln Z,$$

$$Z = \sum_s e^{-H(s)/T}. \quad (24)$$

则从 (23) 可得

$$Z \approx e^{S(E)} e^{-E/T},$$

$$F = E - TS. \quad (25)$$

这当然就是上节讲的热位能之一. S 可以直接由微分得来. (见 (12)). Z 叫作“定温配分函数”.

(24) 的计算中, T 是固定的. 所以, 这个求 F 再求 S 的方法叫做“定温法则”. 它和以上的“定能法则”有一样的效果. 在计算过程中, 它比定能法则方便. F 是所谓“定温热位能”.

(乙): 限能法则

既然可以把 (13) 中的 $\delta(E - H)$ 改成 $\Delta(E - H)$, 我们当然也可以把它改成 $\theta(E - H)$ 也就是不计能量大于 E 的状态. 证明很简单:

$$\begin{aligned} \sum_s \theta(E - H(s)) &= \int dE' \theta(E - E') \Gamma(E') \\ &= e^{S(E)} \int_0^E dE' e^{S(E') - S(E)} \\ &\approx e^{S(E)} \int_0^\infty dX e^{-\frac{X}{T} - \frac{X^2}{2K}}. \end{aligned} \quad (26)$$

式中 $X \equiv E - E'$, K 是由 (20) 定义. 所以, 只要 N 很大,

$$S(E) = \ln \sum_s \theta(E - H(s)). \quad (27)$$

我们叫这计算步骤为“限能法则”, 这法则在某些计算中, 十分有用. 在一般讨论中也相当有用.

(丙): 定压定温法则

以上的计算法则中, 我们已经假定物体的各状态都有一定的体积 V . 如果状态空间包括了不同体积的状态, 则体积 $\hat{V}(s)$ 和能量 $H(s)$ 一样, 也成了状态 s 的函数. 以上的各法则的名字, 应加上“定容”二字. 例如“定容定能法则”. 计算式子也该改写一下. 例如 (13) 该写做

$$\begin{aligned} \Gamma(E, V) &= \sum_s \delta(E - H(s)) \delta(V - \hat{V}(s)), \\ S(E, V) &= \ln \Gamma(E, V). \end{aligned} \quad (28)$$

我们可以依样画葫芦, 把 $\delta(V - \hat{V})$ 改成

$$\begin{aligned} \Delta(V - \hat{V}) &\equiv e^{-p(V - \hat{V})/T}, \\ \frac{p}{T} &= \frac{\partial S(E, V)}{\partial V}. \end{aligned} \quad (29)$$

此地 p 正好是物体所受的压强. T 是温度. 如此我们可以得到“定能定压法则”. 我们可以把 $\delta(E - H)$ 也同时换成 $\Delta(E - H)$, 则得“定温定压法则”. 这

两种法则的计算步骤和以上的各法则相似, 读者可以自己推出, 作为练习. 在这里只提一下“定温定压热位能”, G 的定义:

$$\begin{aligned} G &= -T \ln \bar{Z}, \\ \bar{Z} &= \sum_s e^{-(H(s)+pV(s))/T}. \end{aligned} \quad (30)$$

还是 (24) 的推广. \bar{Z} 为“定温定压配分函数”.

(25) 之推广为

$$\begin{aligned} G &= E - TS + pV, \\ S &= - \left(\frac{\partial G}{\partial T} \right)_p, \\ V &= \left(\frac{\partial G}{\partial p} \right)_T. \end{aligned} \quad (31)$$

(丁): 开放法则

在以上诸法则中, 每个状态的分子总数是假定为一样的. 现在, 如果不同分子数的状态也包括在状态空间内, 则分子数 $\hat{N}(s)$ 也成了状态的函数. 以上的法则都在收集状态时用了 $\delta(N - \hat{N}(s))$ 以确定分子数为 N . 我们又可以沿例把 $\delta(N - \hat{N}(s))$ 换成

$$\begin{aligned} \Delta(N - \hat{N}(s)) &= e^{(\hat{N}(s)-N)\mu/T}, \\ \frac{\mu}{T} &= \left(\frac{\partial S(E, V)}{\partial N} \right)_E. \end{aligned} \quad (32)$$

μ 正是分子的化学势. 我们又可以依据写出一大堆法则来. 在此只提一下最常用的“定温开放法则”, 熵可用以下式子求:

$$\begin{aligned} S(E, N) &= \ln \tilde{\Gamma}(E, N), \\ \tilde{\Gamma}(E, N) &= \sum_s e^{-(H(s)-\mu\hat{N}(s)-E+\mu N)/T}. \end{aligned} \quad (33)$$

这是 (23) 的推广. 这法则可以直接以 μ, T 为定值来计算一个“定温开放热位能” Ω :

$$\begin{aligned} \Omega &= -T \ln \xi, \\ \xi &= \sum_s e^{-(H(s)-\mu\hat{N}(s))/T}. \end{aligned} \quad (34)$$

ξ 为“定温开放配分函数”.

然后熵和分子数可由 Ω 的微分得来

$$\begin{aligned}\Omega &= E - TS - \mu N, \\ S &= - \left(\frac{\partial \Omega}{\partial T} \right)_{\mu}, \\ N &= - \left(\frac{\partial \Omega}{\partial \mu} \right)_T.\end{aligned}\quad (35)$$

以上是 (24), (25) 之推广.

“开放”一词的用意本是指分子不必按硬性限制在物体中. 但是“不开放”物系的熵, 仍可以用这法则计算.

(戊): 极大值法则.

不变量所容许的所有的形像, 都包括在活动范围里. 因此, 这范围的体积是一极大值 (如果不是极大, 则表示还可以放更多的状态到活动范围中). 因此, 基本假设就是一个熵为极大值的假设. 在许多计算中, 这极大的特性十分有用. 以上的诸法则也可以看作求极大值的结果. 我们用定温法则来说明这一点.

令 $\hat{A}(s)$ 为一大值变数. 用它的值, 我们可以把状态空间分成区域. 每区之 $A(s)$ 值为一定. 令

$$\Gamma(A) \equiv \sum_s \delta(A - \hat{A}(s)) \equiv e^{S(A)}, \quad (36)$$

这是某区域之体积. 用以上用惯了的尖峰积分法, 得

$$\begin{aligned}S(A) &= \ln \bar{\Gamma}(A), \\ \bar{\Gamma}(A) &= \sum_s e^{-\alpha(\hat{A}(s)-A)} \\ \alpha &= \frac{\partial S(A)}{\partial A}, \\ A &= \frac{\sum_s \hat{A}(s) e^{-\alpha \hat{A}(s)}}{\sum_s e^{-\alpha \hat{A}(s)}} = \langle \hat{A} \rangle.\end{aligned}\quad (37)$$

我们现在问: 该选用什么 $\alpha \hat{A}(s)$, 才能使 S 为极大而且 $\langle H \rangle = E$? $\langle H \rangle$ 的定义是

$$\langle H \rangle = \frac{\sum_s H(s) e^{-\alpha \hat{A}(s)}}{\sum_s e^{-\alpha \hat{A}(s)}}. \quad (38)$$

这 $\langle H \rangle$ 是能量在该区域的平均值. 因为 H 为大值函数, $H(s)$ 之值在区内几乎所有地方都是 $\langle H \rangle$. 选择 $\alpha \hat{A}(s)$ 的意思是选择每个 s 的 $\alpha \hat{A}(s)$ 值. 为了求 S 之极大值, 必须解

$$\frac{\partial}{\partial \hat{A}(s)} [S - \beta \langle H \rangle] = 0 \quad (39)$$

β 为“外力系数”, (见外力法的解释, 第五章, (5.5) 式上下). 微分的结果是 (稍加整理后):

$$\alpha(\hat{A}(s) - \langle \hat{A} \rangle) - \beta(H(s) - \langle H \rangle) = 0$$

即 $\alpha \hat{A}(s) = \beta H(s)$ 为最佳选择. 这答案当然就是定温法则. 以上的诸法则也可以用类似步骤得来.

这极大值法则有用的地方, 当然不是导出以上诸法则, 而是导出近似的计算法则. 例如, 当 $H(s)$ 太复杂, 以致于定温法则中 Z 的计算太难, 我们可以用一较简单的 $H'(s)$ 取代 $H(s)$. $H'(s)$ 中含有几个可以调整的参数 $\alpha_1, \alpha_2, \dots, \alpha_m$. 然后用

$$\frac{\partial}{\partial \alpha_i} [S - \beta \langle H' \rangle] = 0, \quad i = 1, 2, \dots, m. \quad (40)$$

来决定各 α_i 之值使 S 为极大, 同时限 $\langle H \rangle$ 为一定值. 以后, 我们常用的“平均场法则”就可用这方法求出. 注意, 如果令

$$\rho(s) \equiv \frac{e^{-\alpha \hat{A}(s)}}{\sum_s e^{-\alpha \hat{A}(s)}}, \quad (41)$$

则 (37) 的 S 可写作

$$S = - \sum_s \rho(s) \ln \rho(s), \quad (42)$$

这是一种常见的写法.

请务必注意以上的 $\hat{A}(s)$ 必须是一大值变量, 即 N 量级大数, 以上的各结果才有意义. 所有的法则都是从基本假设得来. 除了这假设之外, 没有其他假设, 不需要其他概念. 许多人喜欢用 (42) 为起点, 然后用 S 极大的条件来定 $\rho(s)$. 这当然就是此地讲的极大值法则. (42) 这个写法, 虽然普遍地被采用, 当作最基本式子, 但它的特别形式却稍有神秘感. 此地的讨论指出, 它和定能法则, 限能法则, 及以上所有法则都一样. 这形式只是一种特别的写法而已.

以上所有这些计算法则, 都是有相同的效果, 只是细节不同. 它们都是统计活动范围内状态的数目. 这活动范围的划定, 或用 $\delta(E-H)$, 或用 $e^{-(H-E)/T}$, 或用 $\delta(V-\hat{V})$, 或用 $e^{-(\hat{V}-V)p/T}$ 等等, 都没有关系, 结果都一样.

这些法则的导出,并不需要新概念的介入,它们只是把基本假设写成各种形式而已.各种热位能也只是各种总体坐标特别组合而已.当然这些组合在热力学中,有特别的意义,在第1节已讨论过一些.在下一章,我们用一些例子来说明这些法则的应用.

讨论问题七

1. 复习第六章第1节的“外力法”.将本章第1节的讨论和外力法作一比较.
2. 第2节的各式计算法则,和各量的数量(即 S, E, N, V 等为 $O(N)$, Γ 为 $O(e^N)$)分不开的.试证如数量级不对,则各计算法则就互相矛盾,都不管用了.
3. 本章的计算法则,和其他书上的“系综”有相同的计算效果.定温法则即所谓“正则系综”,定温开放法则即“巨正则系综”.本章的法则是基本假设 $S = \ln\Gamma$ 的后果.读者如果比较喜欢用系综的想法,当然也可以,只须注意系综的许多概念(如无限多个物体)都是抽象而不可当真,且须牢记时间尺度之不可忽略.
4. 将所有计算法则应用到理想气体上.
5. 某气体容器壁上有 N_A 个“吸附位子”,每个位子至多可以抓住一个气体分子,结合能是 B ,如果气体的压强是 p ,温度是 T ,求被吸住分子的总数.
6. 某弹性金属线的温度是 T ,因受拉力 f 而成折线形,如果 f 忽然除去,线的最后平衡温度将是什么?(见图3)

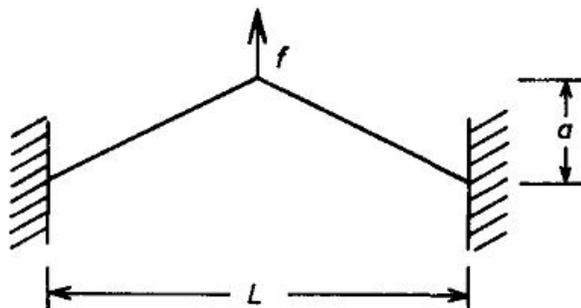


图 3

7. 用定温法则解一维空间的易形模型

$$H = -J \sum_{i=1}^{N-1} s_i s_{i+1}, \quad s_i = \pm 1.$$

8. 一群粒子, 各有半径 a , 质量 m , 散在水中. 水温为 T , 重力向下. 求粒子的高度分布. 答: $\rho(y) = \rho(0)e^{-mgy/T}$.

9. 求以下群体的熵, 物态公式, (即压强、体积、温度间的关系) 和热容率. 群体为 N 分子, 住在一维空间, 不能互相穿过, (所以次序是排定了) 每个分子的大小是 a , 不能压缩或膨胀, 容器的体积 (即容器长度) 为 L . 那一种法则最方便?

10. 如果上题各分子的相互作用能量是

$$\begin{aligned} U(x_i - x_{i+1}) &= \varepsilon, & |x_i - x_{i+1}| < a \\ &= 0, & |x_i - x_{i+1}| > a \end{aligned}$$

并限于左右邻, x_i 为第 i 个分子的位置, 求物态公式, 熵和热容率.

11. N 个带正电 q 和 N 个带负电 $-q$ 的电气离子, 限制在二维空间边长 L 的正方形内运动. 相互作用的位能是 $\zeta(|\mathbf{r}_i - \mathbf{r}_j|) = -q_i q_j \ln|\mathbf{r}_i - \mathbf{r}_j|$, 其中 $q_i, q_j = \pm q$.

试写出此系统的状态方程式并证明当 $T < \frac{q^2}{2}$ 时, 定温配分函数 Z 不存在.

这模型将在第十九章再出现.

12. 假设某三维空间的气体分子为圆饼子, 每个饼子的转动惯量为 I .

(A) 求熵、能量和温度的关系.

(B) 求热容率在 $T \ll \frac{1}{2I}$ 及 $T \gg \frac{1}{2I}$ 时之值.

13. 以上各章的讨论, 除了统计粒子位子之外, 并没有用到许多量子力学. 事实上基本假设已经合用于量子力学, 因为它只是统计物体的状态而已, 并不须要加入任何新概念. 在计算、分析时, 都要注意各变量的次序, 因为量子力学的运动变量是用算符代表.

量子力学的特色, 由自旋 $1/2$ 示范是最理想的. 一个自旋 $1/2$ 只有两个“独立”状态, 但也可以有此二者的“组合”状态. 组合状态可以代表自旋向量, 可指向任一方向. 统计状态时, 可以用任一方向为准, 考虑顺此方向或逆此方向的二状态.

读者应复习有关自旋 $1/2$ 的功课.

14. 令 $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_N$ 为 N 个自旋 $1/2$, 其哈密顿算符为

$$H = -J \sum_{i=1}^{N-1} \mathbf{s}_i \cdot \mathbf{s}_{i+1}.$$

(A) 求零点能量, 基态的简并度. (即有多少形态, 其能量为零点能量).

(B) 求能量最低的激发态.

第八章 示范例子

本章的目的是用简单的例子, 说明上一章导出的计算法则, 并指出观测时间长短的重要性. 首先我们用定能法则分析一个理想气体模型. 此气体的分子有两个内在形态. 再用定温法分析. 然后用定温法分析 H_2 气体的热容率, 主要是讨论 H_2 分子转动的效果. H_2 分子的核子自旋可以是 $s = 1$ 或 0 . 在普通观测时间内, 若没有催化剂帮忙, $s = 1$ 的分子不会因碰撞而变成 $s = 0$. 因此, $s = 1$ 的粒子数和 $s = 0$ 的粒子数各别都是不变量, 不受温度改变的影响. 如果有催化剂帮忙, 使 s 作快速改变, 则 $s = 1$ 和 $s = 0$ 的粒子数会依温度而异, 热容率也就和粒子数各别不变的情形大不相同了. 最后用一个一维空间的模型作为定温定压法的示范, 并说明拉力下的物体不能永久不断.

1. 时间尺度和活动范围

假想某分子, 有两个内在形态, 一为分子的“基态”或“零态”, 能量为 0 , 一为“激发态”, 其能量为 ε , 现在讨论此种分子组成的气体.

第一步是要决定这气体的状态空间和活动范围. 各分子除了跑来跑去, 互相碰撞之外, 它们的内在形态也可能会由于碰撞而改变. 令 n_i 为分子 i 的内形坐标. 分子在基态, 则 $n_i = 0$, 否则 $n_i = 1$. 现在有两种可能性:

① 内在形态的改变很快, 几乎每撞必改, 或观测时间很长, 即观测时间比改变周期要长得得多. (此地“周期”并不表示规则性的改变, 只是指各 n_i 不停的变动, 任一个 n_i 反复一次的大致时间.)

② 内在形态改变很慢, 在观测时间内几乎不变, 每个 n_i 都不变, 即碰撞不易引起改变, 而观测时间不长.

在第 ② 种情形下, 这气体可以说是两种分子的混合气体, 一种是基态分子, 一种是激发态分子, 二者各自为政. 我们等下再谈.

在第 ① 种情形, 我们必须考虑内在形态之改变, 除了位置和动量之外,

还要考虑各 n_i 的改变.

哈密顿量是

$$H = \sum_{i=1}^N \left(\frac{\mathbf{P}_i^2}{2m} + \varepsilon n_i \right). \quad (1)$$

活动范围可定为 $E = H$, 并且各分子位置须在容器内. 这和第五章的 (5.16)(5.17) 相同, 只是 H 复杂一些. 现在活动范围定了, 最主要的一步已完成.

下一步是把熵和其他性质算出来, 可以用上节讲的任一法则. 我们先用定能法则. 令

$$\Gamma_0(E_0) = e^{S_0(E_0)} \quad (2)$$

为 $n_1 = n_2 = \cdots = n_N = 0$ 情形下的活动范围体积. (即 (5.18) 给的 $\Gamma(E_0)$). 现在, 除了对各动量、位置积分外, 还必须对 n_1, \cdots, n_N 求和:

$$\begin{aligned} \Gamma(E) &= \sum_{n_1=0}^1 \sum_{n_2=0}^1 \cdots \sum_{n_N=0}^1 \frac{V^N}{N! h^{3N}} \int d^3 \mathbf{p}_1 \cdots d^3 \mathbf{p}_N \delta(E - H) \\ &= \sum_{n_1=0}^1 \cdots \sum_{n_N=0}^1 \Gamma_0(E - \varepsilon N_1), \end{aligned} \quad (3)$$

$$N_1 = n_1 + n_2 + \cdots + n_N,$$

N_1 是激发态分子总数. 上式可写做

$$\begin{aligned} \Gamma(E) &= \sum_{N_1=0}^N \frac{N!}{N_1! (N - N_1)!} \Gamma_0(E - \varepsilon N_1) \\ &\equiv \int dN_1 e^{S(E, N_1)}, \end{aligned} \quad (4)$$

$$S(E, N_1) \equiv S_0(E - \varepsilon N_1) + N_1 \ln \frac{N}{N_1} + (N - N_1) \ln \frac{N}{N - N_1}. \quad (5)$$

(4) 的组合数指出, 把 N 个分子分成 $N_1, N - N_1$ 两部份, 有多少种分法. 这也就是

$$\begin{aligned} \sum_{n_1=0}^1 \sum_{n_2=0}^1 \cdots \sum_{n_N=0}^1 \delta(n_1 + n_2 + \cdots + n_N - N_1) \\ = \frac{N!}{N_1! (N - N_1)!}. \end{aligned} \quad (6)$$

以上的 N_1 是看作大数. (5) 是用大数定律 (见 (5.38)) 简化 (4) 后的结果.

(4) 的积分, 又可以用第三章的尖峰积分法, 先求 S 之极大值:

$$\left(\frac{\partial S}{\partial N_1}\right)_E = 0,$$

即

$$-\varepsilon \left(\frac{\partial S}{\partial E_0}\right) + \ln \frac{N - N_1}{N_1} = 0. \quad (7)$$

式中 $E_0 \equiv E - \varepsilon N_1$ 是众分子的总动能. (7) 的解是

$$\bar{N}_1 = \frac{N}{e^{\varepsilon/T} + 1}, \quad (8)$$

$$\begin{aligned} \frac{1}{T} &= \frac{\partial S_0}{\partial E_0} = \frac{\partial S}{\partial E}, \\ \Gamma(E) &\approx e^{S(E, \bar{N}_1)} \int dN' e^{-N'^2/2K}, \\ N' &\equiv N_1 - \bar{N}_1, \\ \frac{1}{K} &\equiv \frac{2\varepsilon^2}{3NT^2} + \frac{N}{N_1(N - N_1)}. \end{aligned}$$

因此, 得出物体之熵是

$$\begin{aligned} S &= S(E, \bar{N}_1) \\ &= S_0(E_0) + \frac{N\varepsilon/T}{e^{\varepsilon/T} + 1} + N \ln(1 + e^{-\varepsilon/T}). \end{aligned} \quad (9)$$

S_0 是分子移动部份的熵, 其余是内在形态变动之熵. 各种平衡性质, 现在可用微分得出. 例如, 热容率是

$$\begin{aligned} C_V &= T \frac{\partial S}{\partial T} \\ &= C_{V0} + \frac{N(\varepsilon/T)^2 e^{-\varepsilon/T}}{(1 + e^{-\varepsilon/T})^2}. \end{aligned} \quad (10)$$

见图 1, 图中的曲线, 在 $T = \varepsilon/2$ 附近隆起. $C_{V0} = 3/2 N$ 是动能引起的热容率.

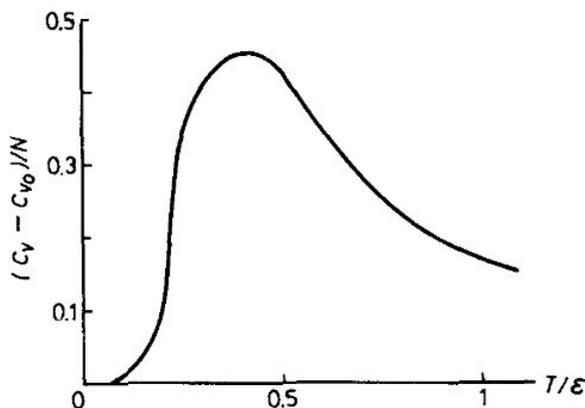


图 1

这是第 ① 种情形, 即分子内在形态改变得很快. 在第 ② 种情形下, 各 n_i 不变. N_1 因此不变. 这气体是两不同分子的混合体, 成分不变, 故其热容率为 $C_{V_0} = 3/2 N$, 没有 (10) 中的后一项, 即没有图 1 中的隆起.

从大尺度观点看, 分子内形态的快速改变, 促使 N_1 成为 $N_1(T)$. 在 $N_1 = \bar{N}_1(T)$ 时, 熵为极大值. 这种快速改变, 拓宽了活动范围.

以上的热容率结果也可以用 $C_V = \partial E / \partial T$ 来分析. 总能量是

$$E = \frac{3}{2} NT + \varepsilon N_1. \quad (11)$$

如果 N_1 为变量, 即第 ① 种情形, 则 $N_1 = \bar{N}_1(T)$,

$$C_V = \frac{3}{2} N + \varepsilon \frac{\partial \bar{N}_1}{\partial T}, \quad (12)$$

即得 (10), (\bar{N}_1 见 (8)). 如果 N_1 为不变量, 即第 ② 种情形, 则

$$C_V = \frac{3}{2} N.$$

也就是说, 如果分子内状态不断改变, 则基本假设决定 N_1 . 如果 N_1 不变, 则基本假设当然对 N_1 值没有话好讲. 图 1 曲线隆起部份是由 $\partial \bar{N}_1 / \partial T$ 而来.

现在我们把这问题用定温法则做. 要计算的是

$$\begin{aligned} F &= -T \ln Z, \\ Z &= \sum_s e^{-H/T} \\ &= \frac{V^N}{N! h^{3N}} \int d^3 \mathbf{p}_1 \cdots d^3 \mathbf{p}_N \sum_{n_1=0}^1 \cdots \sum_{n_N=0}^1 e^{-H/T} \end{aligned} \quad (13)$$

因为 H 的每一项只含有一个变量, 因此, 各积分及计和都可以分别做. 得

$$Z = \frac{V^N}{N! h^{3N}} \left[\int d^3 \mathbf{p} e^{-p^2/2mT} \right]^N \left[\sum_{n=0}^1 e^{-H/T} \right]^N, \quad (14)$$

$$\begin{aligned} F &= N(f_0 + f_1), \\ f_0 &= -T \left(\ln \frac{V}{N} + 1 + \frac{3}{2} \ln \frac{1}{\lambda} \right), \\ \lambda &\equiv h / \sqrt{2\pi m T}, \\ f_1 &= -T \ln (1 + e^{-\varepsilon/T}). \end{aligned} \quad (15)$$

用微分可得熵及其他量, 读者自己试. 以上结果是指第 ① 种情形, 即各 n_i 在观测时间内改变了很多次. 第 ② 种情形是各 n_i 不变, 则气体有 N_1 个激发态分子, $N - N_1$ 个基态分子, N_1 为不变量, (13) 的对 n_1, n_2, \dots, n_N 求和可略去, $N!$ 要改成 $N_1!(N - N_1)!$.

如此求得的热位能是

$$F = N f_0 - T \left[N_1 \ln \frac{N}{N_1} + (N - N_1) \ln \frac{N}{N - N_1} \right] + \varepsilon N_1. \quad (16)$$

f_0 见 (15), 右第二项为“混合熵”所致 (见 (2.40)). 从 (16) 求得的热容率, 就只是 $3/2 N$ 而已. 注意, 在一定温度下, 改变 N_1 使 F 为极小, 则得 (13) 的结果. 也就是说, n_i 的变化, 使 F 成极小值.

显然定温法则比定能法则要简便许多. 读者须记住, 在原则上, 结果上, 它和定能法则是一样的. 在分析过程中, 它以温度为能量尺度, 这比总能量 E 容易捉摸.

2. 氢分子气体

现在来看自然界中最简单的分子, 即由二原子组成者. 我们先复习一下量子力学的分析结果.

分子的整体和内部运动可分成几部份:

A, 电子运动

电子轻、核子重, 对电子而言, 核子大致不动. 电子的基态和第一个激发态有很大的能量差别 ($\sim 5\text{eV} \sim 5 \times 10^4\text{K}$), 只要温度不太高, 电子形态可以说只有基态而已.

B, 核子振动

核子各带正电而互斥, 但被电子拉住、包住. 振动能量为

$$\left(n + \frac{1}{2}\right)\omega, \quad (17)$$

$n = 0, 1, 2, 3, \dots$, 振动频率 ω 之值约为 6000K 左右.

C, 整体转动

因为振动幅度小, 分子可视为有一定大小的物体, 质量集中在两核子上. 令转动角动量为 J , 则能量为

$$\frac{1}{2I}J(J+1), \quad (18)$$

$J = 0, 1, 2, 3, \dots$, $1/(2I)$ 之值约为 80K.

D, 核子自旋

H_2 的两核子各有自旋 $1/2$, 相互作用甚小.

E, 电子自旋

H_2 的两个电子自旋在基态互相抵消, 不必考虑.

F, 整体移动

动能 $p^2/2m$.

不只是 H_2 , 一般双原子的分子的振动能量都较转动能量为高 (见表 1).

表 1 振动转动能量比较 (单位 K)

	H_2	N_2	O_2	NO	HCl
$\hbar\omega/k$	6100	3340	2230	2690	4140
$\hbar/2Ik$	85.4	2.9	2.1	2.4	15.2

因此, 在一大段温度中, 振动状态在基态 ($n = 0$) 而转动常被激发. 以下的讨论, 将振动略去, 分子的整体移动, 在讨论理想气体时讲够了, 所以也略去. 只剩下转动及核子自旋. 能量为

$$H = \frac{1}{2I}J(J+1), \quad (19)$$

核子自旋的能量太小, 不计. 现在来计算热位能, 照定温法则

$$\begin{aligned}
 F &= N(f_1 + f_2), \\
 f_1 &= -T \ln \sum_J (2J+1) e^{-J(J+1)\alpha}, \\
 f_2 &= -T \ln 4, \\
 \alpha &\equiv \frac{1}{2IT}.
 \end{aligned} \tag{20}$$

f_1 为转动热位能, f_2 为核自旋热位能, 核子自旋共有四态, 每核子两态. 这样的结果, 只合用于 HD 分子 (D 指重氢核子), 而不合用于 H_2 分子. H_2 的两核子为相同粒子. 不合群, 对换则波函数必须变号. 波函数是转动部份及自旋部份之积:

$$\varphi \propto Y_{JM}(\hat{\mathbf{R}}) \chi_{sm}, \tag{21}$$

$\hat{\mathbf{R}}$ 为核子之相对位置之方向. $M = -J, -J+1, \dots, J-1, J$. s 为核子总自转, $s=1$ 有三态

$$\begin{aligned}
 m &= 1, & 0, & -1, \\
 \chi_{1m} &= |\uparrow\uparrow\rangle, & \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle), & |\downarrow\downarrow\rangle.
 \end{aligned} \tag{22}$$

$$\chi_{00} = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle). \tag{23}$$

箭头指核子自旋方向, 核子互换, 则 $\hat{\mathbf{R}}$ 换成 $-\hat{\mathbf{R}}$, χ_{1m} 不变, χ_{00} 变号, 而

$$Y_{JM}(-\hat{\mathbf{R}}) = (-1)^J Y_{JM}(\hat{\mathbf{R}}). \tag{24}$$

所以为了要使 φ 变号, $s=1$ 只能配单数 J , $s=0$ 只能配双数 J . 这是量子力学的规矩. 因此, 虽然核子自旋和分子转动间并无相互作用能量, J 和 s 却非独立, f_1 和 f_2 不能分开统计, 只能合起来统计:

$$\begin{aligned}
 f_1 + f_2 &= \ln z, \\
 z &= 3 \sum_{\text{单}J} (2J+1) e^{-J(J+1)\alpha} \\
 &\quad + \sum_{\text{双}J} (2J+1) e^{-J(J+1)\alpha}.
 \end{aligned} \tag{25}$$

从这个结果可以计算熵和热容率

$$\begin{aligned}
 C_V &= T \frac{\partial S}{\partial T}, \\
 \frac{C_V}{N} &= T \frac{\partial^2}{\partial T^2} (T \ln z).
 \end{aligned} \tag{26}$$

图 2 的曲线是 (26) 计算结果, 图中的一串点是实验结果, 显然和理论不符, 原因何在?

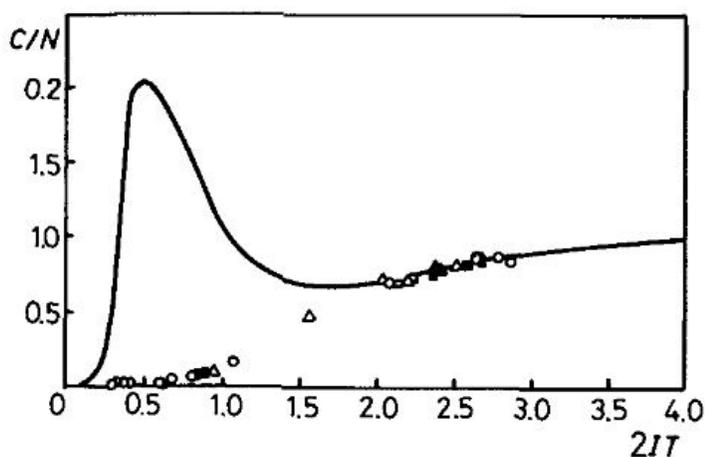


图 2. H_2 的转动热容率. 实线由 (25)(26) 计算, 点为实验结果. 计算与实验在低温不吻合.

原因很简单, H_2 的情形和上节讲的第 ② 种情形相似, 但 (25) 只合用于第 ① 种情形, 也就是说, H_2 的核子自旋, 在一般观测时间内是不变的, 但 (25) 背后的假设是不断地改变. 一个 H_2 核子自旋和其他 H_2 核子自旋作用太弱, 和电子作用也太弱. 且电子自旋抵消了, 不易和核子自旋作用. 核子被电子包着, 分子碰撞时, 核子自旋不会感觉到. 因此, 实验测定的结果是短时间的结果. (25) 不能用, (除非加入有效的催化剂, 使核子自旋经过催化作用而快速改变).

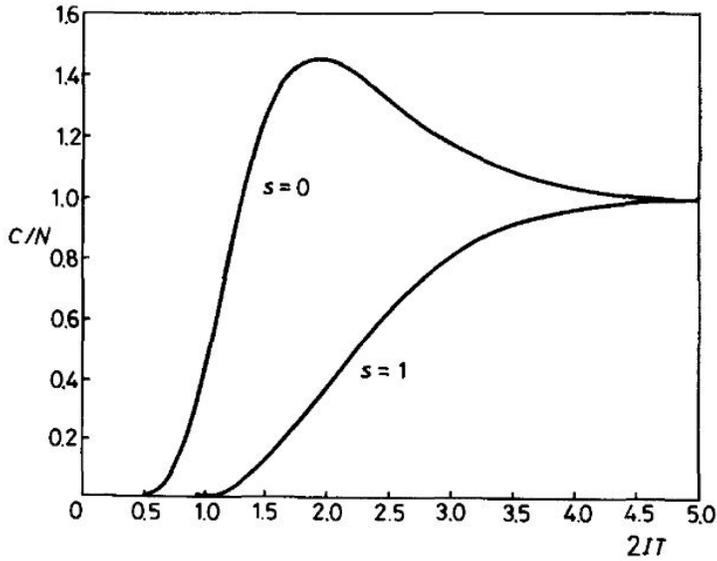


图 3. H₂ 核子自旋为 $s = 0$ 及 $s = 1$ 分别的热容率曲线

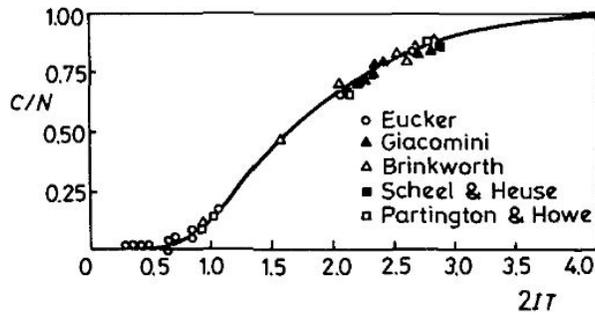


图 4. H₂ 转动热容率, 由 (26)(27) 计算, 与实验结果符合. 以上三图见 Wannier(1966)P.219, 220, 221

正确的计算, 必须把 $s = 1$ 的分子数 N_1 和 $s = 0$ 的分子数 N_0 看作不变量. 活动范围是要依照不变量定义, 这样求出来的热容率是

$$C_V = N_0 \frac{\partial \epsilon_0}{\partial T} + N_1 \frac{\partial \epsilon_1}{\partial T}, \quad (27)$$

$$\begin{aligned}\varepsilon_0(T) &= -\frac{1}{2I} \frac{\partial}{\partial \alpha} \sum_{\text{双}J} (2J+1) e^{-J(J+1)\alpha}, \\ \varepsilon_1(T) &= -\frac{1}{2I} \frac{\partial}{\partial \alpha} \sum_{\text{单}J} (2J+1) e^{-J(J+1)\alpha}.\end{aligned}\quad (28)$$

ε_0 和 ε_1 各为 $s=0$ 及 $s=1$ 的每分子能量, $\partial\varepsilon_0/\partial T$ 及 $\partial\varepsilon_1/\partial T$ 的计算结果见图 3. 图 4 的曲线是 (27) 的结果, 假设 $N_0/N = 1/4$, $N_1/N = 3/4$, 和实验相符.

注意, 图 2 曲线的隆起和图 1 的相似, 如果核子自旋不断地改变, 则 N_0 , N_1 之值都要由温度决定.

$$C_V = \frac{\partial}{\partial T} [\bar{N}_0(T)\varepsilon_0(T) + \bar{N}_1(T)\varepsilon_1(T)] \quad (29)$$

这就和 (26) 一样. \bar{N}_0 和 $\bar{N}_1 = N - \bar{N}_0$ 随 T 而变, 导致图 2 曲线的隆起. 注意, 在 T 很小时 ($T \ll 1/2I$), $\bar{N}_1(T)$ 很小, 因为 $J(J+1)$ 至少是 2:

$$\frac{\bar{N}_1(T)}{N} \approx e^{-2\alpha} = e^{-1/IT}$$

我们可以估计图 2 曲线的低温部分:

$$\frac{1}{I} \frac{\partial \bar{N}_1(T)}{\partial T} \propto \frac{e^{-1/IT}}{I^2 T^2}. \quad (30)$$

$\partial\bar{N}_1/\partial T$ 是一隆起曲线, 极大值在 $T = 1/2I$, 和 (10) 相似, 当然上一节例子的目的, 就是要使读者对本节的内容先有个准备.

3. 一维空间模型

以上的各例子, 是各种理想气体. 分子间的作用的效果, 除了使各变量不断地改变之外, 其他效果, 一概被略去. 在统计计算中, 相互作用能未被计入, 稀疏的气体 and 理想气体是相当近. 但一般的物体, 相互作用有许多重要的效果. 一旦要把相互作用计入, 数学上的问题就变得十分困难? 因此, 可以解的简单的模型, 就十分可贵. 虽然它是过分简略, 但常可以用来大致阐明一些相互作用的效果, 各种一维空间的模型最简单, 将被反复讨论.

现在看一群分子, 限在一直线上. 分子位置为 x_i , $i = 0, 1, 2 \cdots N$, 且 $x_0 < x_1 < x_2 < \cdots < x_N$. 我们把 x_0 固定在 $x_0 = 0$. 此物的“体积”或“长

度”是 x_N . 假设总能函数为

$$H = \sum_{i=1}^N \left[\frac{p_i^2}{2m} + U(x_i - x_{i-1}) \right]. \quad (31)$$

也就是说, 相邻的分子相互作用是近距离相斥, 远则相吸引, 如图 5 所示. 现在用定温法则分析, 计算热位能 G :

$$G = -T \ln \bar{Z},$$

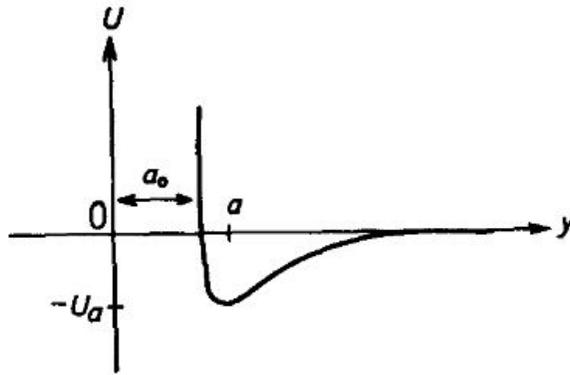


图 5

$$\begin{aligned} \bar{Z} = e^{-F_0/T} \int_0^\infty dx_N \int_0^{x_N} dx_{N-1} \cdots \int_0^{x_2} dx_1 \\ \cdot \exp \left\{ -\frac{1}{T} \left[\sum_{i=1}^N U(x_i - x_{i-1}) + px_N \right] \right\}, \end{aligned} \quad (32)$$

$$e^{-F_0/T} \equiv \frac{1}{h^N} \int d^3 \mathbf{p}_1 \cdots d^3 \mathbf{p}_N \exp \left\{ -\sum_{i=1}^N \frac{p_i^2}{2mT} \right\}, \quad (33)$$

$$F_0 = -\frac{1}{2} NT \ln \left(\frac{2\pi mT}{h} \right).$$

此地的 x_N 即 (7.30) 中的 $\hat{V}(s)$.

现在把积分变量改为

$$\begin{aligned} y_1 &= x_1, \\ y_2 &= x_2 - x_1, \\ y_3 &= x_3 - x_2, \\ &\dots\dots\dots, \\ y_N &= x_N - x_{N-1}. \end{aligned} \quad (34)$$

这些是相邻分子间的距离, x_N 是它们的和: $x_N = y_1 + y_2 + \dots + y_N$, \bar{Z} 的积分变得非常简单:

$$\bar{Z} = \zeta^N e^{-F_0/T}, \quad (35)$$

$$\zeta = \int_0^\infty dy e^{-[U(y)+py]/T}, \quad (36)$$

$$G = -\frac{1}{2}NT \ln \left(\frac{2\pi mT}{h} \right) - NT \ln \zeta.$$

ζ 是 p 和 T 的函数, 如果温度高, 即

$$\frac{U_a}{T} \ll 1, \quad (37)$$

则

$$\begin{aligned} \zeta &\approx \int_{a_0}^\infty dy e^{-py/T} = \frac{T}{p} e^{-pa_0/T}, \\ G &\approx -\frac{1}{2}NT \ln \left(\frac{2\pi mT}{h} \right) - NT \ln \left(\frac{T}{p} \right) + pa_0N. \end{aligned} \quad (38)$$

体积可以用微分求出

$$V = \left(\frac{\partial G}{\partial p} \right)_T = \frac{NT}{p} + Na_0,$$

即

$$p(V - a_0N) = NT. \quad (39)$$

因此, 在高温时这群体一如理想气体, 只是其“有效体积”为 $V - a_0N$. a_0 是两分子间最短的距离.

求熵也容易:

$$\begin{aligned} S &= - \left(\frac{\partial G}{\partial T} \right)_p \\ &= N \left[\frac{1}{2} \ln \left(\frac{2\pi mT}{h} \right) + \frac{3}{2} \right]. \end{aligned} \quad (40)$$

现在看看低温的情形, 即

$$\frac{U_a}{T} \gg 1,$$

则 ζ 的主要来源是 $U(y) + py$ 的极小值附近. 假设 p 不是太大, 极小值在 a 附近, 将 $U(y)$ 展开, 留到三次项, 得

$$U(y) \approx U_a + \frac{1}{2}\alpha(y-a)^2 - \frac{1}{6}\alpha'(y-a)^3. \quad (41)$$

这是 $U(y)$ 在 $y = a$ 附近的近似值, $U(y) + py$ 的极小值在 $\partial U/\partial y = -p$ 处, 即在

$$\bar{y} \approx a - \frac{p}{\alpha} + \frac{1}{2} \frac{\alpha'}{\alpha} \left(\frac{p}{\alpha}\right)^2. \quad (42)$$

因此

$$\zeta \approx e^{-[U(\bar{y})+p\bar{y}]/T} \left[\frac{2\pi T}{U''(\bar{y})} \right]^{1/2}, \quad (43)$$

$$U''(\bar{y}) \approx \alpha + \frac{\alpha'}{\alpha} p.$$

把这结果代入 (36), 稍加整理, 得

$$G = \frac{1}{2}NT \ln \left(\frac{h}{2\pi mT} \right) - \frac{1}{2}NT \ln \left(\frac{2\pi T}{\alpha} \right) + NT \frac{\alpha' p}{2\alpha^2} + N \left[U_a + pa - \frac{p^2}{2\alpha} - \frac{1}{6}\alpha' \left(\frac{p}{\alpha}\right)^3 \right], \quad (44)$$

$$V = \frac{\partial G}{\partial p} = N \left[a - \frac{p}{\alpha} - \frac{\alpha'}{2\alpha} \left(\frac{p}{\alpha}\right)^2 + \frac{T\alpha'}{2\alpha^2} \right]. \quad (45)$$

这些是低温的结果. 因为 $y \approx a$ 的意思是相邻分子距离约是 a , 所以这些结果描写晶体的特性. (38)、(39) 的高温结果是描写气体的特性.

如果 (41) 的 $U(y)$ 展开只留到第二次项, 则成了简谐运动的位能. 第六章讲晶体振动时即是用这种近似位能, 从 (45) 可见, 如果用简谐位能, 即略去 α' , 则 V 和温度的关系就失去了. 固体随温度的膨涨, 是由非简谐作用而来.

这模型虽然能显露一些晶体和气体的特性, 但它并没有“变态”的特性. 当温度由高而低, 它不会忽然在某温度自气态变成固态或液态. 因为 (35) 是 p 的连续, 可微分函数, 所以 p - V 等温曲线, 没有忽然改变斜率的地方. 一般说来, 在一维空间, 短程相互作用 (即每分子只和邻近的几个分子作用) 的群体不会有相变性质. 这是以后的题目.

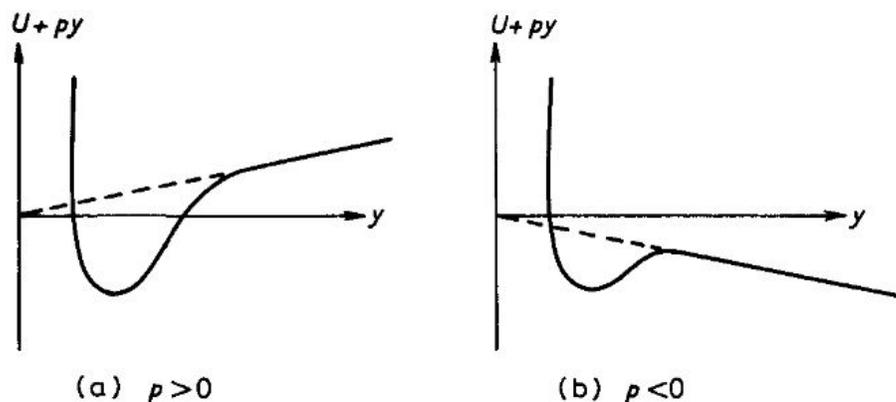


图 6

还有个非常重要的结果: 如果

$$p \leq 0, \quad (46)$$

即无压强, 或有拉力, 则 (35) 的积分就出了毛病. 这是因为 $U(y)$ 在 y 很大时必须消失 (即作用力为短程), 积分全靠 $e^{-yp/T}$ 来收敛. 所以这种情形下, 无论这短程作用是多强, 这结论都逃不掉. $p > 0$ 的情形很简单. 图 6(a) 告诉我们 y 该留在曲线的最低处左右. $p < 0$ 的情形是一个山谷, 但翻过一座屏障, 即是无止境的下坡. (当然, 如果拉力很强, 这山谷也没有了.) 这结果可以作以下解释: 如果每个 y , 都在谷底附近, 则这些分子是连成一链. 如果有一个以上的 y , 跳出谷外, 则链子就断了, 拉力也失. 这例子结论合用于任何物体, 即在拉力下, 任何东西终被拉断. (见图 6(b)).

当然, 这拉力的问题, 是一个时间的问题, 在短时间内, 拉力下的固体是处于平衡状态. 只是我们在选择活动范围作计算时, 必须略去拉断了情形的状态. 怎样略去? 这问题却还没有完美的答案. 不过在许多情形下, 并不难处理. 例如 (44) 和 (45) 是由展开得来. 即使 $p < 0$, 这结果仍可用, 只要我们假装 (35) 的积分没有问题.

讨论问题八

1. N 个原子整齐的排列形成一个完美的晶体, 若将 n 个其中的原子从晶格位置移到原子之间的空隙处, 这就形成 n 个缺陷的不完美晶体. 原子

之间空隙数目 N' 与原子数 N 是等量级的, 若 ω 为将原子从晶格移到空隙所需的能量, 试证

$$\frac{n^2}{(N-n)(N'-n)} = e^{-\omega/T}. \quad (47)$$

2. 上题可用细致平衡解. 某原子吸收一能量为 ω 的声子, 就跑到空隙里去, 留下一空位子. 由此反应, 可得上题结果.
3. 求以上缺陷引起的熵和热容率.
4. 以上晶体从高温 T 骤冷至低温 T' . 在 T' , 原子已不能从晶格位置转到空隙处. 以上结果仍可用吗?
5. 如果将一分子从某晶体中踢出所需能量为 ε . 求该晶体的蒸气压. 假定蒸气为一单原子理想气体, 压强很低, 并利用第六章有关晶体振动的结果.
6. 双原子分子的振动幅度很大时, 为“非简谐振动”. 位能曲线如图所示. 振动能阶可以用下式作近似,

$$\varepsilon_n = \left(n + \frac{1}{2}\right) \omega - a \left(n + \frac{1}{2}\right)^2 \omega, \quad n = 0, 1, 2, \dots \quad (48)$$

a 为代表非简谐性程度的参数. (见图 7)

计算由于非简谐性对比热的影响至 a 的第一次.

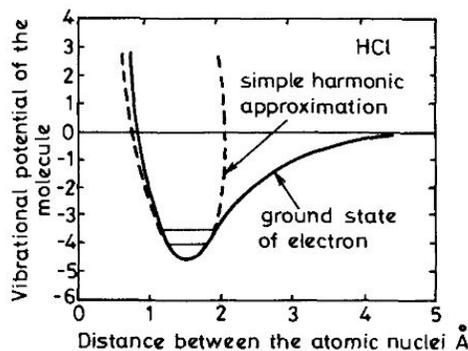


图 7

7. 试证; 如略去量子效应, 则带电粒子的热位能不受磁场影响.

提示: 总能函数为

$$H = \sum_{i=1}^N \frac{1}{2m_i} \left[\mathbf{p}_i - \frac{e_i}{c} \mathbf{A}(\mathbf{r}_i) \right]^2 + U(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N). \quad (49)$$

8. 某些分子具有两相近的电子运动能阶 (如氧、一氧化氮), 其能阶差距远小于 T 值, 足以使较高的能阶具有不少的粒子数, 而影响热力学的性质.

(A) 证明电子运动热容率为

$$C = wN e^{\varepsilon/T} \left[\frac{\varepsilon/T}{e^{\varepsilon/T} + w} \right]^2, \quad (50)$$

式中已假设低能阶自身能值为零. 且有 w_0 个位子, 高能阶能值为 ε , 有 w_1 个位子, $w = w_0/w_1$.

(B) 试分析 C 在高温和低温极限下的近似, 并证明 C 的极大值是在

$$\frac{\varepsilon}{RT} = \ln w + \ln \left(\frac{\varepsilon/T + 2}{\varepsilon/T - 2} \right). \quad (51)$$

(C) 当分子具有三个等距的能阶, 且 $w_i = 1$, 其热容率的形式为何?

(D) 取 $\frac{\varepsilon}{T}$ 值在 1.0~4.0 之间, 计算双能阶分子 ($w_0 = 1, w_1 = 3$) 的热容率极大值, 与 (C) 中结果比较.

(E) 一氧化氮有两个二重能阶, 能距异常的小, $\varepsilon = 174\text{K}$. 氧分子两能阶差 = 11300K, 在高温下 $w_0 = 3, w_1 = 2$. 估计此两分子热容率极大值时, 温度和热容率各为多少?

9. 一方盒高 h , 底面积 A . 内有分子 N 个, 成一气体, 温度为 T , 重力向下. 每分子质量为 m .

(A) 求顶和底面的气压.

(B) 求四壁的气压.

10. 一个圆柱形的容器, 半径 R , 长 L , 绕轴旋转, 角速度为 ω . 容器内有气体, 温度为 T , 每分子质量为 m . 求内壁压强.

注意: 在随器转动的坐标系内, 此气体是在平衡态, 有固定的边界. 读者必须先求出在此坐标系内的总能函数.

11. 如果考虑量子力学, 上题的解又是如何?

12. 某一维空间物体有 N 分子, 总能为

$$H = \sum_{i=1}^N \frac{p_i^2}{2m} + \sum_{i=1}^N U(x_i - x_{i=1})$$

$$x_0 < x_1 < x_2 < \cdots < x_N.$$

$$U(y) = \frac{1}{2}m\omega^2(y - a)^2, \quad y > 0.$$
(52)

$x_0 = 0$. 压强 p 加在第 N 个分子上.

(A) 求此物的熵、热容率和温度 T 的关系.

(B) 求 $\langle x_N \rangle$.

以上计算可以假设 $m\omega^2 a^2 \gg T$, 即略去

$$\alpha \equiv e^{-m\omega^2 a^2 / 2T}$$
(53)

(C) 求此物之热膨胀系数 (即长度和温度的关系). α 不可完全略去, 但可略去高次项.

13. 某物的哈密顿量为

$$H = - \sum_{i=1}^N h n_i s_i,$$

$n_i = 0$ 或 1 , $s_i = -1$ 或 1 .

n_i 可以想作在位子 i 的粒子数. 这些粒子可以跑来跑去. 如果粒子不在, 则 h 不发生作用. 粒子之总数为 N' .

(A) 求熵及热容率.

求出之后, 检查一下, 看 $\varepsilon N'$ 很小时, 或 $N' \approx N$ 时答案是否合理.

(B) 如果粒子不能跑来跑去, 但分布散乱, 求熵及热容率.

第九章 杨、李定理

从以上几章的分析和示范例子可见,用基本假设来解释热力学,是非常的理想.第七章所列的计算法则,和热力学的观点完全吻合,也就是说,在一定的环境之下,平衡性质一定.由于物体为宏观,即分子数 N 为大数,第七章的分析指出,用不同的统计份量做计算,可以得到一样的结果.但是,我们的分析,不够严谨.分析过程中并未考虑各式子的数学意义.各示范例子,都是过分简化,可以立即解出的模型.虽然这些例子使我们对这些计算法则和基本假设有了大致的认识和信心,但并不足以指出一般情形下可能遭遇的困难.在这一章,我们用比较严谨的数学步骤来讨论.如果分子粒子数 N 是有限,则各式热位能的计算是没有问题的.虽然我们需要的 N 非无限,但却是要 10^{20} 之类的大数,因此, $N \rightarrow \infty$ 的极限问题,就成了用基本假设解释热力学的重要问题,即一般所谓“宏观极限”问题.用严谨的数学分析来讨论此类极限问题,已成为统计力学的一部门.这一部门的创始工作,可以说是三十年前的杨、李定理*.杨、李的定理是为分析变态而作.此后的严谨数学分析也大多针对相变问题.原因是相变问题模型太不容易解,马虎一点的分析靠不住.但杨李定理的应用,却是十分的普遍.在这一章我们讨论杨、李定理之一,即热位能 Ω/V 在 $V \rightarrow \infty$ 极限之存在.本章目的不是讨论相变,(相变在以后诸章讨论),而是指出这极限存在的原因,指出宏观极限和分子作用能,及统计步骤的关系.从定理的证明,可以看到平衡中宏观物体的最重要特性,即它可以分成许多部分,每一部分都是在平衡状态中.此地所强调的倒不是严密的证明方法,而是对基本假设的更深一层的认识,并提醒读者数学步骤的严密,并不能保证结果有正确的物理意义,小心为要.

*Yang and Lee(1952), 黄克孙书中附录之一亦有相当详细的介绍 (Huang(1963), p.458).

1. 宏观极限 (即“热力学极限”)

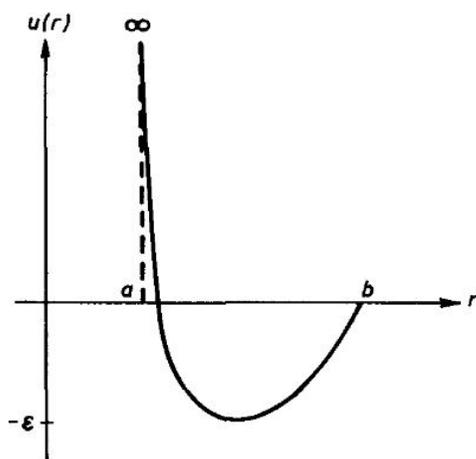


图 1

先订好一个模型. 令物体的体积为 V . 分子间的相互作用能如图 1. $u(r)$ 的必要性质是

$$\begin{aligned}
 \text{(甲)} \quad & u(r) = 0, \quad \text{如果 } r > b, \\
 \text{(乙)} \quad & u(r) = \infty, \quad \text{如果 } r < a, \\
 \text{(丙)} \quad & u(r) > -\varepsilon.
 \end{aligned} \tag{1}$$

也就是说, 分子间的距离不得小于 a , 因此体积 V 中可容纳的人口为有限:

$$N < N_m = V / \left(\frac{4\pi}{3} a^3 \right). \tag{2}$$

吸引力为有限, 即 ε 为有限. 相距太远无作用, 即 b 为有限, 只要 (1) 适合, $u(r)$ 改成别的样子无所谓.

现在我们用定温开放法则来分析, 热位能 Ω 的定义是

$$\xi \equiv e^{-\Omega/T} = \sum_s e^{-(H(s) - \mu N(s))/T}, \tag{3}$$

$$\begin{aligned}
 p &= -\frac{\Omega}{V}, \\
 H &= \sum_{i=1}^N \frac{p_i^2}{2m} + U, \\
 U &= \frac{1}{2} \sum_{i,j=1}^N u(r_i - r_j), \\
 \sum_s &\equiv \sum_{N=0}^{N_m} \frac{1}{N! h^{3N}} \int_V d^3\mathbf{r}_1 \cdots d^3\mathbf{r}_N d^3\mathbf{p}_1 \cdots d^3\mathbf{p}_N.
 \end{aligned} \tag{4}$$

先把各个 \mathbf{p}_i 积分做了, 得

$$\xi(z, V) = \sum_{N=0}^{N_m} z^N Z_N(z, V), \tag{5}$$

$$Z_N(z, N) \equiv \frac{1}{N!} \int_V d^{3N}\mathbf{r} e^{-U/T}, \tag{6}$$

$$z \equiv \frac{e^{\mu/T}}{h^3} \int d^3\mathbf{p} e^{-p^2/2mT} = \frac{e^{\mu/T}}{h^3} (2\pi mT)^{3/2}, \tag{7}$$

$$d^{3N}\mathbf{r} \equiv d^3\mathbf{r}_1 d^3\mathbf{r}_2 \cdots d^3\mathbf{r}_N.$$

ξ, Ω, Z 都是 T, V 和 z 的函数. 杨李的第一定理是:

只要 (1) 成立, $z > 0$, 则

$$\lim_{V \rightarrow \infty} \left(-\frac{\Omega}{TV} \right) \equiv \lim_{V \rightarrow \infty} \frac{1}{V} \ln \xi(z, V) \tag{8}$$

极限存在, 和容器的形状无关, 而且是一个 z 的连续、渐增的函数. 以上假设物体的表面积小于 $V^{2/3}$ 乘以一定的常数, 即表面没有作过份的曲折.

极限 (8) 存在的意思是: V 只要够大, Ω/V 就不随 V 变, 如果把 V 分成 n 部分, 只要每部份都很大, 则每一部份分开计算的 Ω/V 也一样. 说明的步骤也是从这观点开始.

我们把物体分成 n 部分, 并把各部分交接处划定一厚度为 b 的“分界区”(见图 2). 令分界区的总体积为 V' . 除去分界区, 剩下的体积是 nV_0 . 如果我们限定分界区内不许有分子, 则 ξ 变成

$$[\xi(z, V_0)]^n, \tag{9}$$

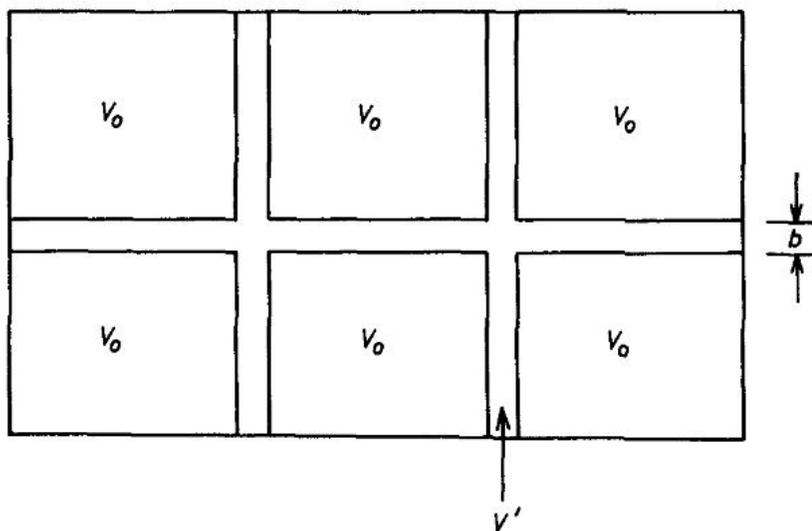


图 2

因为分子间作用有效距离为 b , 不能超越分界区而相互作用. 所以每个区域的积分可以分开做, 故得 (9).

从 (6) 可见, 如果把积分范围缩小, 则结果减少, 因为积分函数是正的. 因此, 如果把分界区切掉, 则 Z_N 只能减少, ξ 也只能减少, 即

$$[\xi(z, V_0)]^n \leq \xi(z, V). \quad (10)$$

所以, 把物体切成 n 块, 并丢掉一层分界区, 就把 ξ 弄小了一些.

下一步是把分界区放回去, 并简化其中作用而得 ξ 之上限. 我们把 (6) 改写成

$$Z_N(z, V) = \sum_{N'=0}^N \frac{1}{N'!(N-N')!} \int_{V'} d^{3N'} \mathbf{r} \int_{V-V'} d^{3(N-N')} \mathbf{r} e^{-U/T} \quad (11)$$

也就是把 $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N'}$ 限制在分界区, 其他的在 n 个 V_0 区内. 每个粒子至多和 $(b/a)^3$ 个其他粒子作用, (条件 (1)—(甲), (乙)). 每一对粒子间之作用能不低过 $-\varepsilon$, (条件 (1)—(丙)). 如果作用能 U 中和分界区有关的各项, 被 $-\varepsilon(b/a)^3 N'$ 取代, 则值变小, 即

$$U \geq \sum_{k,j=N'+1}^N u(\mathbf{r}_k - \mathbf{r}_j) - \varepsilon N'(b/a)^3. \quad (12)$$

右第一项合计即是除去分界区的作用能. 将 (12) 代入 (11), 积分, 得

$$z^N Z_N(z, V) \leq \sum_{N'=0}^N \frac{z^N e^{\varepsilon(\frac{b}{a})^3 \frac{N'}{T}} (V')^{N'}}{N'!} Z_{N-N'}^0. \quad (13)$$

$Z_{N-N'}^0$ 是把分界区切去后 $N - N'$ 个分子的统计分量. 把 (13) 代入 (5), 得

$$\xi(z, V) \leq \sum_{N'=0}^{\infty} \frac{z^{N'} e^{\varepsilon(\frac{b}{a})^3 \frac{N'}{T}} (V')^{N'}}{N'!} \sum_{N''=0}^N z^{N''} Z_{N''}^0. \quad (14)$$

以上我们把 $N - N'$ 写成 N'' , 并除去 $N' < N$ 的约束. 现在最后的一个合计即是 $(\xi(z, V_0))^n$, 第一个合计可以立刻得出, (14) 成了

$$\begin{aligned} [\xi(z, V_0)]^n \leq \xi(z, V) \leq e^{\sigma V'} [\xi(z, V_0)]^n, \\ \sigma \equiv z e^{(\varepsilon/T)(b/a)^3}. \end{aligned} \quad (15)$$

左边的下限是把 (10) 一并写入.

V' 是分界区的体积, 约是分界面面积乘以厚度 b , 即

$$V' = cV_0^{2/3}. \quad (16)$$

c 为一常数. 令

$$\phi(z, V_0) \equiv \frac{1}{(V/n)} \ln \xi(z, V_0) \quad (17)$$

则从 (15) 得

$$\begin{aligned} \phi(z, V_0) \leq \frac{1}{V} \ln \xi(z, V) \leq \phi(z, V_0) + \sigma c V_0^{-1/3}, \\ \left| \frac{1}{V} \ln \xi(z, V) - \phi(z, V_0) \right| \leq \sigma c V_0^{-1/3}. \end{aligned} \quad (18)$$

也就是说, 只要 V_0 够大, $\phi(z, V_0)$ 和 $V^{-1} \ln \xi(z, V)$ 间的差别就很小了. 即单位体积热位能可以用物体的一部分来计算, 不在乎用那一部分, 计算的误差是和该部分的大小成反比. 这已经把“热力学极限”在这模型内证明得很清楚了. 以上的分析中, V_0 不必是个立方体, 只要 (16) 成立就行. 每一部分也不必是一样大.

从物理学的观点, (18) 已是所要的结果. 若是要达到数学上的严格那么从 (8) 到 (18), 可能要多加几句数学术语. 我们在此不予计较.

因为

$$\frac{\partial p}{\partial \mu} = \frac{N}{V}, \quad z \propto e^{\mu/T} \quad (19)$$

所以

$$\frac{N}{V} = \frac{1}{T} z \frac{\partial p}{\partial z} = \frac{1}{V} \sum_{N=1}^{N_m} N z^N Z_N \bigg/ \sum_{N=0}^{N_m} z^N Z_N. \quad (20)$$

(20) 中每项都是正值, 故 p 一定是随 z 而增, 不管 V 是大是小. 既然 $V \rightarrow \infty$ 的极限存在, $p(z)$ 在此极限亦必是随 z 而增. 再者

$$\frac{N}{V} < \frac{N_m}{V} = \frac{1}{a^3}, \quad (21)$$

故 $\partial p / \partial z$ 为有限, 即 p 为连续. 杨、李的第一定理就此证明. 杨、李的第二定理在讨论相变时十分有用. 读者可自阅文献, 本书将不讨论.

2. 定理的推广

现在再回头看看 (1) 的三条件, 即有限的有效作用距离 (甲), 及近距离的互斥 (乙), 及吸引力不可太大 (丙). (甲) 不合则有远程作用, 就无法定“分界区”. (乙), (丙) 不合则可能会有不稳定的情形, 即所有的粒子可能都集合到一小区域去了. 有关宏观极限的问题中, 最重要的莫非是以上定理的推广. 即把条件放松, 看有什么结果. 我们略提一两个结果.

(甲) 可以略为放松, 只要 $u(r)$ 消失得够快, 即 $r > b$ 时,

$$u(r) \leq Ar^{-\lambda}, \quad \lambda > d, \quad (22)$$

就行, d 是物体所存在的空间度数, 一般 $d = 3$. (乙) 也可以放松一点, 在 $r < a$ 时

$$u(r) \geq Br^{-\lambda'}, \quad \lambda' > d, \quad (23)$$

就行. (丙) 不能通融.

但是最重要的作用力却都不合乎这些条件, 物质是由核子的正电和电子的负电作用而成. 还有磁偶极的作用. 这些电磁作用, 都是长程作用. 而且正负电荷间的作用能在 $r \rightarrow 0$ 时为 $-\infty$. 三个条件, 无一适合. 杨李定理因而不能直接用来讨论物质的稳定性和宏观极限. 以下略提电荷作用和量子性质的效果.

从电磁学的经验, 我们知道未抵消的电荷总是会跑到表面上来. 因此, 能量及其他量都会和物体的形状有关. 因此, 只有正负电相抵消的物体才有宏观极限可言.

量子力学也必须考虑, 否则正、负电子的粒子可以无限制地接近. 所有的粒子都集中到一点去了. 量子力学指出, 一个粒子被局限在很小的空间则其动能会加大. 这个定理使粒子不集中到一点去. 但是, 如果粒子都是合群的, 我们可以想见, 粒子虽不集中于一点, 但仍会集中, 宏观极限仍会有问题. 我们必须请独占定律来帮忙, 也就是说, 必须要有费米粒子. 已经有人证明: 如没有独占定律, 每粒子的零点能量会比 $-N^{2/5}$ 低, 即在 $N \rightarrow \infty$ 时, 每粒子的能量趋向 $-\infty$. 如果所有的负电粒子 (或正电粒子, 或全部) 为不合群, (粒子种类为有限, 否则独占定律仍不管用), 则每粒子的零点能量才为有限. 这些是相当不简单的证明. 宏观极限的证明, 现在也有相当的成绩[†]. 基本假设的地位已经是非常的稳固, 因为各种严密的分析结果, 都没有和事实相矛盾.

最后, 我们必须强调: 对此类严密的分析结果, 必须要采取一种谨慎的态度, 不能轻易接受它的结论, 必须先了解分析中的假设, 及各种极限的意义. 例如, 从杨、李的定理, 我们可以下定论: 在一定的温度、化学势下, 只能有一种平衡态, 不可能有暂稳态的存在. 事实上当然是有各种不同的暂稳晶体结构存在, 也有过饱和蒸汽存在. 这是因为这定理是以基本假设为出发点, 且没有考虑时间尺度, 只有无限长的时间观测之下, 这假设才可能完全正确. 所谓严密的证明, 是指在假设和结论间, 没有作任何近似. 而一般的假设, 都是为求简化而忽略了某些事实. 统计力学的基本假设, 就是一个过份简化的假设, 用途虽广, 但并非全能. 除了假设之外, 某些极限, 如体积 $V \rightarrow \infty$ 之类, 也要小心. 真实的物体, 是有限的. 因此, 极限的意义必须和事实不相矛盾. 常发生的情形是

$$\lim_{V \rightarrow \infty} \frac{1}{\ln V} = 0. \quad (24)$$

这个 0 的意义, 就不太清楚. 即是最大物体的体积, 它的对数也不会大到那里去. 虽然从数学的观点 (24) 是毫无疑问, 十分严密的, 但从物理的观点, (24) 是不够严密, 颇有疑问的. 因此, 要了解一个“严密”的结论, 必须对证明的细节, 也要特别小心. 一般说来, 结论的真义, 总是在证明步骤中出现的.

[†]见 Lebowitz and Lieb(1969), Dyson(1967).

有关严密证明的全盘介绍, 见 Griffiths(1972).

讨论问题九

1. 第 1 节的证明指出, 一旦哈密顿量定了, 一切热力性质全定了. 这是个严密的证明. 这证明也指出基本假设的含糊之处, 因为这证明否定了暂稳态的存在, 而我们把暂稳态看成平衡, 只要观测时间不太长.

以石墨和钻石为例, 试指出从碳原子群体模型及“严密的证明”, 可把钻石的存在否定了.

2. 推导出第 1 节所有的结果.
3. 如果模型是在 d 维空间, 证明细节有何变更? 读者以后可能在文献中遇到 $d \rightarrow \infty$ 的模型 (当然这只是理论上的讨论, 有时 $d > 3$ 的模型十分有用). 在 $d \rightarrow \infty$ 时, 第 1 节的证明就不合用了. 也就是说一般 $d \rightarrow \infty$ 的统计力学模型, 很可能靠不住, 因为它的宏观极限有问题.

第三篇

概率

概率是一个十分有用而十分容易被滥用的概念。本篇把统计和概率之不可分反复地指出,并讨论各种应用例子。概率在此是一个由数据的统计而得来的数字,而不是代表某人的猜测或无知程度。统计是一个整理数据,表达数据的步骤。它是一个工具,而不是一个原理。我们要整理的资料是分子运动的各种细节。

第十章讨论概率和统计的关系。第十一章讨论独立或纷乱的现象。这现象非常重要,但很不容易了解。此地的讨论,是对现象的陈述,而不究其来历。

第十二章讨论中心极限定理,对独立现象和概率的意义作更进一步的分析,更阐明一些基本假设的结果。这假设和独立现象是分不开的。第十三章的内容是一些有关相干函数的计算,一些起伏、响应等概念的应用例子。

第十章 概率和统计

在以上五章, 我们已经对统计力学的基本假设, 作了初步的介绍. 虽然用的示范例子是过份简化的模型, 但大致不难看出: ① 这基本假设可用来分析、解释所有的平衡现象和导出所有热力学的结果. ② 这假设和应用并不需要很多统计学的理论或概念, 以上几章谈到的, 就足够了. ③ 有关分子运动轨迹的资料, 也不需要多少. 分子运动扮演的角色, 只是使基本假设成立而已. 一旦活动范围定了, 分子运动就可置诸脑后.

但是若要对基本假设的含义作更深一层的了解, 我们需要对分子运动作更进一步的分析, 需要更多统计学的工具. 因此, 虽然基本假设的应用才刚开始讲, 我们暂停下来, 穿插一些统计学的材料. 本章的重点在“概率”的定义, 我们强调统计资料是概率定义的根本.

1. 概率的应用和滥用

统计学是一门整理资料的科学. 概率又称“或然率”或“可能率”, 是统计学的基本概念之一. 在数学理论中, 概率有很严格的定义, 它的种种性质, 也有明文规定. 在应用方面, 概率这个概念却常常被滥用. 因为它的定义常常没弄清楚. 现在我们来谈两种不同的用法. 在不同用法中, 概率有不同的意义. 如果不搞清楚这些不同的意义, 就会导致不正确的结论, 和带来不必要的烦恼. 特别是钻理论的学生, 对这两种不同的用法特别容易搞混.

在第一种用法中, 概率是两个统计数字之比. 例如, 某人丢骰子, 丢了六百次, 其中有 98 次是 1. 那么, 我们说丢骰子丢出 1 的概率, 或可能率, 是 98 除以 600, 差不多是 $1/6$. 这是一个实验的结果. 这个实验的结果是和骰子的构造, 和掷骰子的人的手运动有关. 照我们经验, 如果这个人, 或任何其他入, 再丢 600 次或者 1000 次, 10000 次, 总是会有 $1/6$ 的次数得 1. 这是一个事实, 是一个自然现象.

概率的第二种用处,是用来衡量一个人对某件事情怀疑的程度,或者不确定的程度.譬如说,一个押注的人看不到桌上骰子的点数.如果,你问他骰子上面几点.他说:“我不知道,因此每一种不同的点数都有 $1/6$ 的概率.”在这个用法概率不是从统计数字得来,只表示他全不知道.

以上这两种用法绝对不能混为一谈,举个例子:我们问某人,把骰子丢 600 次,会有多少“1 点”出现.某人说:“既然骰子只有六个可能性,既然我什么都不知道,所以 1 点出现 100 次”.这个结果刚好被他猜对.但是他用的理由却是不对,他用他的“不知道”来推断一个自然现象.这是不可以的.再举一个例子:假如我们问某人:“新竹每一年有几天下雨?”某人回答说“既然只有两种可能性,下雨或者不下雨,我对新竹一点都不知道,所以一年有六个月下雨,六个月不下雨.”这个回答所根据的理由,和上一个例子中的理由一样,是用他的无知来推测自然现象.新竹下雨过去的记录,可以从新竹县志或气象局查出.将来的下雨情形,可以由过去记录加上其他数据来推测,不能以“不知道”来推测.

以上强调概率两种用法的不同,并不是说第二种用法有什么不对.其实,每一个实验的结果,都须要注明误差.这误差并不表示实验做错了,只表示结果的正确程度.这是实验者的判断.如果他说误差是 0.1% ,这表示他不能确定 0.1% 内的数字.现在举个例子来说明.

某人用水银压力计,到某大街上去量气压.他的结果说:“某年、月、日、时、分、气压是 1.02 ± 0.05 大气压.”他的判断是,那个时候的气压应该是在 0.97 及 1.07 之间.到底在这范围内的那一点,他不知道.为了估计这个误差,他必须考虑到他的压力计的较正是多正确,他的眼睛是否被街上的废气弄花了,以致于看不清压力计上的尺码.

请注意,这误差和压力的“起伏”完全是两回事,绝对不能搞混了.气压随时间而变,是为起伏.一般水银压力计所量,是一种平均压力,在一秒钟或更长时间内的平均值.长时间的起伏可以用此压力计量出.长时间的起伏都是由于气候变化.短时间的快速起伏,是由于车子、路人和修理马路的机器等所引起.水银压力计量不出,但是可用一个录音机测定.这些噪音引起的压力起伏幅度很小,大约是 10^{-10} 大气压.这和以上水银计 5% 的误差是毫无关系的.

在这本书里,绝大多数情况下,概率只有一个用法,这就是第一个用法,除非特别注明我们将不考虑第二个用法.以后每次提到概率,我们都要讨论

它是怎样由统计资料求得.

在此我们还需要强调一点, 即概率的值必须不因数据多少而异, 才有意义. (当然, 数据不能太少.) 例如, 投骰子 1200 次, 统计 1 点出现次数, 我们可以用前 600 次, 求得 1 点的概率, 或用后 600 次, 或用整个 1200 次? 求得的概率大致都是 $1/6$. 如果用前 600 次得 $1/12$, 后 600 次得 $1/4$, 则虽然我们可以说 1200 次的统计结果是 $1/6$, 但这结果的意义很有问题. 也就是说, $1/6$ 这结果, 要经得起重复测定的考验, 才有意义.

2. 分布和平均值的统计

在前几章, 我们讨论过“分布”, 如国民年龄分布, 粒子形象的能量分布等. 分布是分类统计的结果. 例如, 把国民按年龄分类, 每一年龄算一类, 而后统计每一类的人数, 于是得到图 3.1 所示的年龄分布曲线. 这是一项有关众国民的资料. 它的统计过程, 可以用一个抽象的式子表示

$$f(X') = \sum_s \delta(X' - X(s)). \quad (1)$$

s 是各人的名字, $X(s)$ 是 s 先生 (女士) 的年龄, \sum_s 是统计人口数. δ -函数是分类, 即收集年龄为 X' 的人名. 年龄在 X_1 和 X_2 之间的人数是

$$N(X_1 < X < X_2) = \int_{X_1}^{X_2} dX' f(X'). \quad (2)$$

读者可复习一下第三章.

“分布”也常作“分配”. 我们可以说每段年龄分配到多少人口. 用那个名词都无所谓.

(1) 也就是一般分类统计的步骤. 分析的对象总是一群“样品”(每个样品都有他的名字或编号 s) 的某种特性 $X(s)$, 而后收集 $X(s)$ 相同的, 归之为为一类. 每类中的样品数叫做该类的“统计份量”. 把各类的统计份量列表或画成曲线, 就成了该特性的“分布”, 或样品的“分配”. 这是有关这些样品的一项资料. 这项资料, 有时仍嫌繁琐. 我们可以用些“平均值”来大致表示分布的情形. 例如, X 的平均值是

$$\begin{aligned} \langle X \rangle &\equiv \int dX' \rho(X') X', \\ &= \frac{1}{N} \sum_s X(s). \end{aligned} \quad (3)$$

$$\rho(X) \equiv \frac{f(X)}{\int dX' f(X')}. \quad (4)$$

N 是样品总数, $\rho(X)dX$ 是 $f(X)dX$ 所占的百分比. $(X - \langle X \rangle)^2$ 的平均值是

$$\begin{aligned} \langle (\Delta X)^2 \rangle &= \langle (X - \langle X \rangle)^2 \rangle \\ &= \int dX' (X' - \langle X \rangle)^2 \rho(X') \\ &= \frac{1}{N} \sum_s (X(s) - \langle X \rangle)^2, \end{aligned} \quad (5)$$

$\langle X \rangle$ 是分布的“重心”所在. 而 $\langle (\Delta X)^2 \rangle$ 是估计分布的分散的程度. $\rho(X)$ 叫做“概率的 X 分布”. 它只是分统计份量除以总份量. 这是定义. 至于 $\rho(X)$ 和“机会”、“可能”等概念有什么关系, 我们在此不必过问. 这要看 $\rho(X)$ 的应用.

任何 $X(s)$ 的函数 $\vartheta(X)$, 都有它的平均值. 例如 $\rho(X')$ 就是 $\delta(X' - X(s))$ 的平均值, 1 的平均值是 1:

$$\begin{aligned} \langle \vartheta(X) \rangle &= \int dX' \vartheta(X') \rho(X') \\ &= \frac{1}{N} \sum_s \vartheta(X(s)). \end{aligned} \quad (6)$$

$$\begin{aligned} \rho(X') &= \langle \delta(X' - X(s)) \rangle, \\ &= \frac{1}{N} \sum_s \delta(X' - X(s)), \end{aligned} \quad (7)$$

$$\langle 1 \rangle = \int \rho(X) dX = 1 = 100\%. \quad (7')$$

(7') 是所谓“标准化”或“归一化”条件. 总百分比必是 100%.

我们可以把分类统计的结果作部份的综合, 也就是把某些类归为大类. 例如年龄小于 30 岁的可归为“青年”类, 70 岁到 80 岁为“人生开始”类等等. 每一类的人数为其统计份量. 一般说来, 样品中的一组 R (即样品集中一子集) 的统计份量可以定义如下. 令 R 的“指定函数” $\theta_R(s)$ 为

$$\begin{aligned} \theta_R(s) &= 1, \quad \text{如果 } s \in R, \\ \theta_R(s) &= 0, \quad \text{如果 } s \notin R. \end{aligned} \quad (8)$$

因此

$$N_R \equiv \sum_s \theta_R(s) \quad (9)$$

为 R 组的统计份量. 令

$$P(R) \equiv \frac{N_R}{N} = \langle \theta_R \rangle \quad (10)$$

为 R 组的“概率”, 这是定义. 注意: $\delta(X' - X(s))dX'$ 是 R 的一个例子, 即 $X(s)$ 值在 X' 及 $X' + dX'$ 间的样品之集就是 R .

如果 R 是 $X(s) < X'$ 样品之集, 则

$$P(R) = P(X < X') = \langle \theta(X' - X) \rangle = \int_{-\infty}^{X'} dX \rho(X) \quad (11)$$

为“ $X < X'$ 之概率”.

这些被统计的“样品”可以是一群人, 一群物, 或一些事件的记录, 实验数据, 状态空间中的形象等. 现在再看些例子.

以下是新竹电话本 (民国 67 年) 第 205、206、214 页中姓马和姓梁的电话号码的个位数字:

$$\begin{aligned} &53631836117711836080666871376898 \\ &57815291573155577898809703583555 \\ &65652793069222509039677800596755 \end{aligned} \quad (12)$$

共有 $N = 96$ 户, 每户为一“样品”, 而以上的数字为每户的一个“特性”, $X(s)$, $s = 1, 2, \dots, 96$. 定概率分布为

$$\rho(X') = \frac{\sum_{s=1}^N \delta(X' - X(s))}{N} = \frac{N_{X'}}{N}. \quad (13)$$

$N_X = (X \text{ 出现的次数}), N = 96$.

此地的 δ -函数的变量只有整数, 故它的定义是

$$\begin{aligned} \delta(M) &= 1, \quad \text{如 } M = 0, \\ &= 0, \quad \text{如 } M \neq 0. \end{aligned} \quad (14)$$

$\rho(X)$ 可以用以上数据统计

X	0	1	2	3	4	5	6	7	8	9
$N(X)$	9	9	4	11	0	17	12	13	12	9
$\rho(X)\%$	9.37	9.37	4.16	11.46	0	17.71	12.50	13.54	12.50	9.37

(15)

从 $\rho(X)$, 各种平均值可以求得, 如

$$\begin{aligned}\langle X \rangle &= 4.95, \\ C_0 &\equiv \langle (\Delta X)^2 \rangle = \langle X^2 \rangle - \langle X \rangle^2 = 7.94.\end{aligned}\quad (16)$$

(15), (16) 是从 (12) 统计出来的结果, 至于这些结果有何意义, 还得要多花工夫才能了解. 例如, 为什么“4”这号码在 96 户中竟不出现一次? 同样有趣而且更重要的问题是为什么其他号码的 $\rho(X)$ 差不多而又有些差别? 第一问题的答案不难找出, 但第二问题却不是容易回答的.

再看个例子, 从 (12) 这些数字, 定义

$$Y(s) = X(s)X(s+1), \quad s = 1, 2, \dots, N'. \quad (17)$$

$N' = N - 1 = 95$, 这些是 (12) 数列中相邻数字的乘积. Y 的平均值是

$$\begin{aligned}\langle Y \rangle &= \frac{1}{N'} \sum_{s=1}^{N'} X(s)X(s+1) \\ &= 24.05.\end{aligned}\quad (18)$$

现在定义

$$\begin{aligned}C_1 &\equiv \langle Y \rangle - \langle X \rangle^2 = -0.45, \\ \frac{C_1}{C_0} &= \frac{C_1}{\langle (\Delta X)^2 \rangle} = -0.057.\end{aligned}\quad (19)$$

C_1 是所谓相邻数字的相干值, 同理, 我们可以定

$$\begin{aligned}C_k &\equiv \frac{1}{N''} \sum_{s=1}^{N''} X(s)X(s+k) - \langle X \rangle^2, \\ k &= 0, 1, 2, 3, \dots \\ N'' &= N - k.\end{aligned}\quad (20)$$

当然 $C_0 = \langle (\Delta X)^2 \rangle$. N'' 不是 N , 所以 k 必须比 N 小很多, 以求平均值的定义大致一样. C_k 是“相干函数”, 是“距离” k 的函数. 统计结果 C_1 比 $\langle (\Delta X)^2 \rangle$ 小很多, (见 (18)、(19)). 我们把这结果解释成 (12) 中相邻的号码大致“不相干”或互相“独立”. 照我们的“直觉”, 如果这些数字是“不规则”的, 则各 $X(s)$ 会有的高于 $\langle X \rangle$, 有的低于 $\langle X \rangle$, 毫无规则, C_1, C_2, \dots 都该是零. 在下一章, 我们将对这些概念作有系统的介绍.

注意, 这些相干函数当然是和号码排列有关, s 本是指定电话用户排列顺序, 并无统计意义. s 的数值在此是照电话本的排列顺序而定. 这顺序是我们新加的概念, 各 s 之间的“距离”也是照这顺序规定.

我们的目的是要把统计学用来整理分子运动资料, 许多实验都是在度量各种分子运动变量的平均值. 我们的统计步骤是要以分子运动为根据, 各种概率平均值都必须从分子运动数据来定义. 时间的顺序, 空间的位置, 是最基本的概念. 各种相干函数, 在讨论运动的规则性时, 是最重要的工具.

现在我们看一个比较不同的示范例子.

令一垂直平面为 xy 平面. 重力向下, y 轴向上, x 轴水平. 一个粒子在这个平面中 x 轴以上运动. x 轴是一硬地板, 粒子掉在地板上就反跳. 假设没有能量损失. 水平运动被限制在 $0 < x < L$, 为求简单起见, 把 $x = 0$ 和 $x = L$ 认定为同一点, 如此这个粒子的运动面成为一个圆柱面, 其周长为 L , 这是所谓周期性的边界限制或“循环边界”(见图 1)

这粒子的运动, 遵照力学的运动方程式. 这方程式的解, 是这粒子的轨迹. 轨迹非常简单, 水平运动一直不变绕着圆柱走. 垂直运动是上下不停. 任何性质都可由这个运动轨迹求出, 观测时间是 $0 < t < \mathcal{T}$.

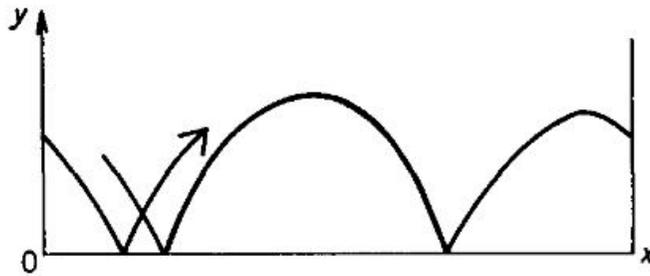


图 1

我们先问: 粒子的高度分布如何? 这问题并不清楚, 因为“分布”必须要有明确的统计定义, 而我们还没有下这定义. 一般有关运动变量的统计, 多是定时间为统计份量, 因为实验的结果多是对时间的平均值. 我们现在把粒子花在某段高度的时间, 定义为该段高度的统计份量. 令

$$\rho(y') = \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} dt \delta(y' - y(t)). \quad (21)$$

$y(t)$ 是在时刻 t 的高度, $\rho(y')$ 就是概率的高度分布. (21) 是 (7) 的一个例子.

此地的“样品集”是轨迹, 每一时刻 t 有一“样品”, 即在 t 时粒子的状态. $\rho(y')$ 很容易计算 ($\rho \propto |dt/dy|$, 即 ρ 和垂直速度成反比):

$$\begin{aligned}\rho(y) &= \frac{1}{2}[(h-y)h]^{-1/2}[1 + O(\tau/T)], \quad \text{如 } y < h. \\ &= 0, \quad \text{如 } y > h.\end{aligned}\quad (22)$$

h 是粒子最高可及的高度, 读者可自己导出 (22). τ 是上下一次的周期

$$\tau = \frac{2p_{y0}}{mg}, \quad (23)$$

p_{y0} 是在 $y = 0$ 时的垂直动量.

如果 $T \gg \tau$, 则 $O(\tau/T)$ 可略去. $\rho(y)$ 的积分

$$\begin{aligned}\int_{y_1}^{y_2} \rho(y) dy &= \mathcal{T}(y_1 < y < y_2) / T \\ &= P(y_1 < y < y_2)\end{aligned}\quad (24)$$

可定义为高度在 y_1, y_2 间的概率. $\mathcal{T}(y_1 < y < y_2)$ 就是粒子花在这段高度的时间.

高度的相关函数 (和 (20) 的 C_k 相类似)

$$\begin{aligned}C(t') &= \frac{1}{T} \int_0^T dt y'(t) y'(t+t'), \\ y'(t) &\equiv y(t) - \langle y \rangle, \quad t' \ll \tau.\end{aligned}\quad (25)$$

不难计算出来, (见图 2)

$$\begin{aligned}C(t') &= \frac{g\tau^2}{8} \left[\frac{4}{45} - \frac{2}{3}t''^2 + \frac{2}{3}t''^3 - \frac{1}{6}t''^4 \right], \\ t'' &= t'/2\tau, \quad \text{其中 } 0 < t' < \tau.\end{aligned}\quad (25')$$

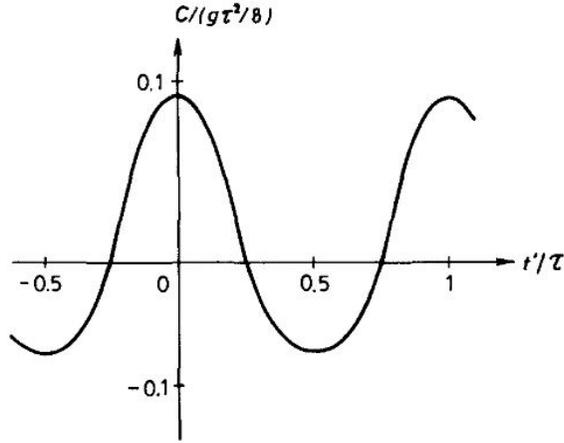


图 2: $C(t)$ 为一周期函数, 因为 $y(t)$ 是个周期函数.

再看在时间 $0 < t < \mathcal{T}$ 内, 地板所受的平均场. 粒子不着地时, 地板不受力. 着地时, 因为地板很硬, 受力很大. 每落地一次, 粒子动量的改变, 是粒子落地前动量的两倍, 即 $2p_y$. 如果每隔一段时间 τ 落地一次, 则第 n 次落地的时刻是 $n\tau$, 地板受力是

$$W(t) = \sum_{n=-\infty}^{\infty} 2p_{y_0} \delta(t - n\tau), \quad (26)$$

平均场是

$$\begin{aligned} \langle W \rangle &= \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} dt W(t) \\ &= 2p_{y_0} \frac{N}{\mathcal{T}} \\ &= \frac{2p_{y_0}}{\tau} \left(1 + O\left(\frac{\tau}{\mathcal{T}}\right) \right). \end{aligned} \quad (27)$$

式中 N 是在 $0 < t < \mathcal{T}$ 这段时间内落地的次数. $O(\tau/\mathcal{T})$ 这一项, 要看最后一次落地, 和 $t = \mathcal{T}$ 中间, 过了多久. 如果 $\mathcal{T} \gg \tau$, 则此项可略去,

$$\langle W \rangle = 2p_{y_0}/\tau = mg, \quad (28)$$

τ 之值见 (23).

以上的这些例子, 是统计步骤的例子. 概率、分布和平均值的概念, 只是用来整理、报导已有的资料. 它们并不涉及数据的解释, 并不涉及原理的探求. 这些例子还不足以说明概率的一般定义和用法. 在下节, 我们转向广泛的讨论.

3. 概率的一般定义

以上的例子? 指出概率是如何从统计得来. 现在我们把概率的定义用最广义的、通用的, 但可惜是比较抽象的字眼叙述如下. 然后再举些例子.

要定义概率, 需要三样东西: (甲) 样品集、(乙) 分类、(丙) 每类的统计份量. 这些东西要适合一些条件:

(甲) 样品集: 称之为 E , 必须要有样品, (至少必须要有一个样品).

(乙) 定“分类”为一群 E 之子集, 每子集为一类. 分类必须适合以下条件: 如第 R_1 为一类, R_2 为一类, 则 $R_1 \cup R_2$ 即二类之综合, 亦是一类. $R_1 \cap R_2$, 即二类之重合, 亦为一类. 还有, 如 R 为一类, 不属于 R 的样品, 亦为一类.

(丙) 每类须予以一统计份量 $P(R)$, 且

$$\begin{aligned} P(R) &\geq 0, \\ P(R_1 \cup R_2) &= P(R_1) + P(R_2) - P(R_1 \cap R_2), \\ P(E) &= 1. \end{aligned}$$

最后一个条件 $P(E) = 1$ 只是把统计份量“标准化”或“归一化”. 如果不适合, 我们只须要把每个 $P(R)$ 除以 $P(E)$ 就适合了. 这定义十分广泛. 凡适合以上各条件的, 就是概率. 这是数学家的概率定义. 也是现代概率理论的出发点*. 注意, 这广泛的定义并不提到统计, 虽然由统计定义的概率总会适合这广泛定义. 在许多概率问题中, 统计意义都略去不谈, 引为理所当然. 但有时会出毛病. 我们此地的重点是在统计. 因此在应用概率理论时, 我们还需要时时注意各种定义和结论的统计意义.

【例】掷骰子 (其一)

令样品集为 $E = \{1, 2, 3, 4, 5, 6\}$. 分类为所有 E 的子集. 概率 $P(R)$ 为 R 集内的样品数除以 6. 这概率的定义并未涉及统计, 不过我们知道它的统计意义. 注意, 此地的“样品”不是上节例中的实验记录, 或运动轨迹, 而是抽象的六个可能出现的点数. 我们可以想象骰子的 36 万次的记录, 按出现的点数分为六组. E 的六个样品为这六组.

【例】掷骰子 (其二)

*概率的数学分析, 见杨维哲“概率论”(1979).

用同样的 E , 我们改用另一分类 (R, R', E) :

$$\begin{aligned} R &= \{1, 2, 3\} & R' &= \{4, 5, 6\}, \\ P(R) &= 0.55, & P(R') &= 0.45. \end{aligned} \quad (29)$$

这样的概率定义, 也合乎以上 (甲)(乙)(丙) 条件, 这概率可用到某个歪骰子上. 本例说明, 分类不必分得很细, 只要适合条件 (乙) 就行. 本例的这群子集, 并不包括上例中的许多子集如 $\{1, 3\}$, $\{2, 5, 6\}$, $\{2\}$ 等.

【例】 在所有的整数中, 任取一个, 得偶数之概率是多少? 答案显然是 $1/2$. 得 10 之倍数之概率? 答: $1/10$.

这些答案是“理所当然”, 其统计意义是, 从 0 到 N (N 为大数) 约有 $1/2 N$ 个偶数, $N/10$ 个 10 之倍数.

【例】 在半径为 1 的圆上任画一弦, 弦长小于 1 的概率有多少?

这个问题, 看来简单但无意义, 因为“任画一弦”没有“理所当然”的统计意义. 如果画一万根弦, 该如何画? 是平行的, 而均匀的分布在圆内? 还是有均匀分布的角度? 还是有均匀分布的在圆上的端点? 每一种情形都会导致一个不同的概率. 一般说来, 连续的样品集都没有“理所当然”的统计份量.

现在来看两个特别重要的例子.

【例】 轨迹的时间分配.

令 E 为某物体状态空间中轨迹的分布范围, 即活动范围.

现在我们定范围中的各状态为样品, 把状态按某运动变量 $X(s)$, $s \in E$, 之值分类. 如果 R 为一类, 则定义 R 之概率为

$$P(R) = \mathcal{T}_R / \mathcal{T}, \quad (30)$$

R 是 E 之子集, 即活动范围中的一区域. \mathcal{T}_R 是轨迹花在 R 内的时间. \mathcal{T} 是总观测时间.

【例】 和上例一样, 只是定义概率为

$$P'(R) = \frac{\Gamma(R)}{\Gamma(E)}, \quad (30')$$

$\Gamma(R)$ 为 R 之容积, 即 R 内状态总数. $\Gamma(E)$ 为活动范围内状态之总数.

注意: (30) 是直接以运动轨迹为根据而定. (30') 并不提及轨迹, (但 E 应该由轨迹划定). 记得在第五章讲基本假设时, 我们借助于量子力学, 使活动范围成了一个分立而非连续的样品集. 如此 $\Gamma(E)$, $\Gamma(R)$ 才不会含糊. 当然, 无论状态是连续还是分立, \mathcal{T}_R 和 \mathcal{T} 是不含糊的.

从概率的广泛定义, 可以定出平均值. 令

$$\begin{aligned} X_0 < X_1 < X_2 < X_3 \cdots < X_n, \\ R_k = \{s; X_{k-1} < X(s) < X_k\}, \quad k = 1, 2, 3, \cdots, n. \end{aligned} \quad (31)$$

也就是说凡是 $X(s)$ 在 X_{k-1}, X_k 之间的样品, 都收集在 R_k 内. X_0 该比 $X(s)$ 的极小值小, X_n 比 $X(s)$ 的极大值大. 令

$$X_k - X_{k-1} = dX, \quad k = 1, 2, \cdots, n, \quad (32)$$

为相当小的数字, 则 $X(s)$ 之平均值 $\langle X \rangle$ 及分布函数 $\rho(X)$ 是

$$\begin{aligned} \langle X \rangle &= \sum_{k=1}^n X_k P(R_k) = \int dX \rho(X) X, \\ \rho(X_k) dX &\equiv P(R_k). \end{aligned} \quad (33)$$

【例】用 (30) 为概率定义. 令 $X(s)$ 为一宏观变数. (其值为 N 量级大数). 则

$$\rho(X) = \frac{1}{\sqrt{2\pi K}} e^{-(X-\bar{X})^2/2K} \quad (34)$$

X 和 K 的定义和第六章末节的 A, K_A 相同. (34) 即 (6.86). $A = X, K = K_A$.

在此要特别强调: “分类” 是样品集的一群子集, 但不必包括所有的子集, 只要条件 (乙) 适合就可以. (掷骰子第二例已说明这一点). 如果我们把活动范围用宏观变数之值分类, 则每一类 R 都是范围中的一大区域, 即 $\Gamma(R)$, $\Gamma(E)$ 为同数量级 (e^N 级) 之大数. ($\ln \Gamma(R), \ln \Gamma(E)$ 均为 N 量级). 如此, (30) 定义的 $P(R)$ 才可能有明确的平衡意义, 原因如下:

轨迹是在观测时间内真正出现的状态之集. 轨迹中状态之总数由运动决定. 如果大约每隔时间 τ , 每分子改变其状态一次, 则轨迹中状态总数约为

$$N\mathcal{T}/\tau \quad (35)$$

这和 $\Gamma(E) \sim e^N$ 比较来, 自然微不足道. 因此, 如果 R 是一个很小的子集, 轨迹很可能就不经过 R , \mathcal{T}_R 就是零. 如果碰巧在某时间经过 R , 则 $\mathcal{T}_R \neq 0$. 这样的概率, 对分析平衡性质, 没有用处. 但如果 R 的容积和 E 相当, 则轨迹花在 R 内的时间才可能和 \mathcal{T} 相当, $P(R)$ 才可能大致不随 \mathcal{T} 之增减而异, 而成为一平衡特性. 因此 (30) 所指的分类, 只该包括这种很大的区域.

讨论问题十

1. 一个路灯离地距离为 h . 求地上受光的强度的分布.
2. 某气体容器壁上有一小孔, 半径为 a , 壁厚为 b . 求漏出气体分子的速度方向的分布. 讨论 a, b 及平均自由距离大小和答案的关系.
3. 变数的变换.

已知 $\rho(x)$ 分布, 如果 $x' = f(x)$, x' 之分布为何?

答:

$$\begin{aligned}\rho(x') &= \langle \delta(x' - f(x)) \rangle \\ &= \int dx \rho(x) \delta(x' - f(x))\end{aligned}\quad (36)$$

注意, 每个 x' 可能有几个 x 适合 $x' = f(x)$. 每种情形都要统计进去.

令 $r = (x, y)$, r 的分布在方块 ($0 < x < 1, 0 < y < 1$) 内为均匀, 即 $\rho(x, y) = 1$. 求极坐标 (r, θ) 值之分布 $\rho(r, \theta), \rho(r), \rho(\theta)$.

4. 从电话本中取 100 个号码的最后 4 位数, 叫它们 $x_1, x_2, \dots, x_N, N = 100$
 - (A) 求 $\langle x \rangle, \langle x^2 \rangle$.
 - (B) 求 $C_k, k = 0, 1, 2$

$$C_k \equiv \frac{1}{N-k} \sum_{s=1}^{N-k} (x_s - \langle x \rangle)(x_{s+k} - \langle x \rangle).$$

(C) 讨论这列数字的不规则性.

5. 用袖珍计算器, 算出第二节最后例子的粒子落地点的坐标. 统计落地点的分布. 这分布自然和 L 及 l 的相对大小有关. $l \equiv v_x \tau, v_x$ 是水平速度, τ 是上下一次的周期. (见 (23)).

第十一章 独立和纷乱

独立就是现象与现象之间互不相干因此相互之间没有关连、看不出因果, 即日常所谓的纷乱、不规则. 统计学对这类现象有定量的分析方法. 独立是一个非常重要的现象, 它使许多复杂的问题简化, 但它的来源却不是很清楚. 本章叙述独立的定义, 并举例说明它的后果. 许多运动变量都有大致互相独立的起伏, 但独立性的检验? 却又不是很简单的事. 相干值是十分重要的概念. 它指出各变量间不独立的程度. 许多运动变量的相干值, 可以用散射实验测定. 本章最后一节讨论此类实验.

1. 独立现象、定义和后果

先复习两个定义:

R_1, R_2 的联合概率:

$$P(R_1 \cap R_2) \equiv P(R_1, R_2). \quad (1)$$

在已给条件 R_2 下, R_1 的概率

$$P(R_1|R_2) \equiv \frac{P(R_1 \cap R_2)}{P(R_2)}. \quad (2)$$

以 R_2 为条件的意思是把 R_2 当样品集, 然后在 R_2 内分类. 从这些定义, 可以定出变量 X, Y 的“联合分布” $\rho(X, Y)$ 及“条件分布” $\rho(X|Y)$. 定义步骤和 (10.31), (10.33) 相似. 我们以 (X, Y) 之值将样品集分类. (见 (10.31)、(10.33)).

令

$$R_k = \{s; X_{k-1} < X(s) < X_k\}, \quad (3)$$

$$Q_l = \{s; Y_{l-1} < Y(s) < Y_l\}.$$

定义

$$\rho(X_k, Y_l) dX dY \equiv P(R_k \cap Q_l), \quad (4)$$

$$\rho(X|Y) \equiv \rho(X, Y)/\rho(Y).$$

读者可试证

$$\rho(X) = \int dY \rho(X, Y). \quad (5)$$

同理, 我们可以订出更多变量的联合分布, 及条件下的分布.

独立的定义如下:

如果

$$P(R_1, R_2) = P(R_1)P(R_2),$$

或

$$P(R_1|R_2) = P(R_1). \quad (6)$$

则 R_1, R_2 为“互相独立”, 即 R_2 对 R_1 “没有影响”.

变量 X, Y 互相独立的定义是

$$\rho(X, Y) = \rho(X)\rho(Y), \quad (7)$$

或

$$\rho(X|Y) = \rho(X).$$

即 X 的统计资料和 Y 无关. 如果 X, Y 互相独立, 则任何函数 f, g 都合乎

$$\langle f(X)g(Y) \rangle = \langle f(X) \rangle \langle g(Y) \rangle. \quad (8)$$

令 X, Y 之“相干函数”, 或“相关函数”为

$$\begin{aligned} C &\equiv \langle XY \rangle - \langle X \rangle \langle Y \rangle \\ &= \langle (X - \langle X \rangle)(Y - \langle Y \rangle) \rangle, \end{aligned} \quad (9)$$

因此, 如果 X, Y 互相独立, 则 $C = 0$.

独立是一个现象, 看来简单, 但其由来不很明显. 如果没有这现象, 概率的应用价值就会大打折扣. 许多概率的应用问题, 都是从某些简单的独立变量, 做些排列组合的计算, 去求比较复杂的变数的分布或平均值. 独立现象就是混乱, 不规则, 纷散的现象. 相干函数是度量变量之间有多少“规则”存在. 我们特别强调, 独立是需要由统计资料验证. 它并非理所当然的. 要检验几个变量是否互相独立, 并非易事, 下节再说. 现在先看些独立现象的应用例子.

【例】 N 个气体分子在容积 V 内. 令 v 为某小区域的容积. n 为区内的分子数. 求 n 的分布 $\rho(n)$.

定义 n_i 如下: ($i = 1, 2, \dots, N$)

如果第 i 个分子在小区内, 则 $n_i = 1$, 否则 $n_i = 0$. 因此

$$n = \sum_{i=1}^N n_i. \quad (10)$$

现在假设 ① n_i 互相独立. ② $n_i = 1$ 的概率为

$$p = \langle n_i \rangle = \frac{v}{V}. \quad (11)$$

从这两个假设得

$$\rho(n) = \binom{N}{n} p^n (1-p)^{N-n}. \quad (12)$$

(此所谓“二项分布”).

如果 N 、 V 都很大, 密度为 $\frac{N}{V} = \alpha$, 则 (12) 可简化为

$$\rho(n) = \frac{(\alpha v)^n}{n!} e^{-\alpha v}. \quad (13)$$

(此所谓“卜阿松分布”).

【例】 在密度为 α 的气体中, 邻近分子间的距离 r 的分布为何?

答:

$$\rho(r) = 4\pi r^2 \alpha e^{-\frac{4\pi r^2 \alpha}{3}}. \quad (13')$$

$\alpha 4\pi r^2 dr$ 是在 r 和 $r + dr$ 间有一分子之概率.

$e^{-\frac{4\pi r^2 \alpha}{3}}$ 是在 r 之内没有其他分子的概率.

【例】 某长链分子 (在水中游动), 由 N 个键连成. 每键长为 l . 假设各键方向为独立变量, 每键方向分布均匀. 求 $\langle R^2 \rangle$, \mathbf{R} 为分子首尾相对位置.

$$\begin{aligned} \mathbf{R} &= \mathbf{r}_1 + \mathbf{r}_2 + \cdots + \mathbf{r}_N, \\ |\mathbf{r}_i| &= l, \\ \langle R^2 \rangle &= \sum_{i,j} \langle \mathbf{r}_i \cdot \mathbf{r}_j \rangle = \sum_i \langle \mathbf{r}_i^2 \rangle = Nl^2, \end{aligned} \quad (14)$$

$\sqrt{\langle R^2 \rangle} = \sqrt{N}l$ 可以看作这分子的大小.

现在看一个比较复杂的例子.

【例】 令 $\mathbf{u} + \mathbf{R}$ 为某晶体中某原子的位置, \mathbf{R} 为其平衡位置, 即 $\langle \mathbf{u} \rangle = 0$. 求 $\langle \mathbf{u}^2 \rangle$.

\mathbf{u} 为基本振动 ϕ_j 之组合

$$\mathbf{u} = \sum_j \eta_j \phi_j(\mathbf{R}), \quad (15)$$

$$\phi_j(\mathbf{R}) = \frac{1}{\sqrt{N}} \mathbf{e}_j e^{i\mathbf{k}_j \cdot \mathbf{R}_j}. \quad (16)$$

基本振动的幅度是平面波. \mathbf{e}_j 为振动方向. \mathbf{k}_j 为其波向量. N 是原子总数. η_j 是我们的运动变量. 现在假设各 η_j 互相独立. 则

$$\begin{aligned} \langle \mathbf{u}^2 \rangle &= \sum_j \langle \eta_j^2 \rangle |\phi_j|^2 = \frac{1}{N} \langle A \rangle, \\ A &\equiv \sum_i \eta_i^2, \end{aligned} \quad (17)$$

A 是一大数. 我们可以用第五、六章的方法去求. 振动总能量是

$$H = \sum_i \left(\frac{p_i^2}{2m} + \frac{m\omega_i^2 \eta_i^2}{2} \right) \quad (18)$$

m 是原子质量, ω_i 为 i 单元的频率. 因为每一频率都为许多单元所共有, 因此我们可以直接利用“能量均分”结果, (假定温度高, 略去量子效应), 得

$$\begin{aligned} \frac{1}{2} N_\alpha T &= \sum_{i=1}^{N_\alpha} \frac{1}{2} m \omega_\alpha^2 \eta_i^2, \\ A_\alpha &= \sum_{i=1}^{N_\alpha} \eta_i^2 = \frac{N_\alpha T}{m \omega_\alpha^2}. \end{aligned} \quad (19)$$

在此, 我们假设这 N_α 个单元共有大致相同的频率 ω_α . 然后再把不同的 α 组加起来, 得

$$\begin{aligned} A &= \sum_\alpha A_\alpha = \sum_\alpha \frac{N_\alpha T}{m \omega_\alpha^2} \\ &= \int d\omega g(\omega) \frac{T}{m \omega^2} V. \end{aligned} \quad (20)$$

$g(\omega)$ 为振动频率分布, 见第四章 (6.36). 代入 (17), 得

$$\langle u^2 \rangle = \frac{V}{N} \int d\omega g(\omega) \frac{T}{m \omega^2}. \quad (21)$$

读者或许会问: 为什么要加 (19) 这一步, 搞个 N_α 进来?

为什么不直接说: 每个单元

$$\left\langle \frac{1}{2} m \omega_i^2 \eta_i^2 \right\rangle = \frac{T}{2}, \text{ 能量均分,} \quad (22)$$

$$\therefore \langle \eta_i^2 \rangle = \frac{T}{m \omega_i^2},$$

不是更简单吗? 如此固然没有错, 但我们在此要对平均值 $\langle \dots \rangle$ 的意义特别小心留意. $\langle \mathbf{u}^2 \rangle$ 的意义是 \mathbf{u}^2 对时间的平均值. 在第五章讨论能量均分时, 只用了基本假设定熵, 而后求出 N 个振动单元的总能量, 结果是 NT . N 必须是大数. 这是计算的必要条件. 因此 (22) 中 $\langle \dots \rangle$ 符号的意义十分不清楚.

(20) 是应用基本假设的结果, 因为它是个大数. 所以 (17) 右边的取平均值 $\langle \dots \rangle$ 没有作用. 这一点若不清楚, 请复习第六章末节.

(21) 是个十分重要而有趣的结果. 它指出低频率振动对原子的位移颇有影响 (因为积分中的 $\frac{1}{\omega^2}$). 如 ω 小, 则 $g(\omega) \propto \omega^2$, (见 (6.42)), 正好把 $\frac{1}{\omega^2}$ 抵消了. 但如果晶体是在二度空间内 (例如附在一平面上的一层原子), 则 $g(\omega) \propto \omega$. (21) 这积分就出了毛病. 一般说来, d 度空间的晶体有 $g(\omega) \propto \omega^{d-1}$ 结果, (21) 大致是

$$\langle \mathbf{u}^2 \rangle \propto \frac{T}{\theta_D} \lambda^2 \int_0^1 \frac{dx}{x} x^{d-2}, \quad (23)$$

$$\lambda \equiv \frac{h}{\sqrt{2\pi m k \theta_D}}.$$

θ_D 是底拜温度, λ 大致是原子在此温度的波长, (23) 用的 $g(\omega)$ 是 (6.42) 的推广 ($x = \frac{\omega}{\omega_D}$). 当 $d \leq 2$ 时, $\langle \mathbf{u}^2 \rangle \rightarrow \infty$, 即低频率振动, 使原子位移扩大到不可收拾. 这结果以后会详谈, (见第廿九章). 除此以外, 以上结果指出, 原子位移不大, 因为 θ_D 大致是 300 K (见表 (6.1)), λ 大致是 0.5 Å 左右.

2. 独立的检验

上节的各例子, 明显地指出了独立的重要性. 互相独立的变数, 可以个别考虑. 这把许多复杂的问题, 简化不少. 现在我们来讨论独立现象的一些基本问题.

此地谈的“独立”或“不相干”并不是指分离或不相连接. 例如, 有两个单摆互不相连, 各自摆动. 令其位移各为 $X_1(t) = \cos \omega t$, $X_2(t) = \cos(\omega t + \phi)$. 这两单摆照日常的说法是独立不相干的. 但是, 如果它们有相同的频率, 则

(假定 $\langle X_1 \rangle = \langle X_2 \rangle = 0$)

$$\begin{aligned} C_{12} &= \langle X_1 X_2 \rangle = \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} dt X_1(t) X_2(t) \\ &= \frac{1}{2} \cos \phi + O\left(\frac{1}{\omega \mathcal{T}}\right). \end{aligned} \quad (24)$$

即 X_1, X_2 不是独立变量 (假设 $\phi \neq \pi/2$). 因此, 分离、不相连, 或无相互作用, 都不能保证独立性. 若把上例“时髦化”, 则是两个雷射光源, 频率相同. 两道光的干涉清晰可测.

注意 ① 我们一向把观测时间 \mathcal{T} 看作有限长的时间. \mathcal{T} 的长短, 视分析的对象而定. ② 在此我们重申统计为概率和平均值的资料来源. 不同的统计步骤, 决定不同的概率和平均值. 每种问题都不一样, 需要某种对它合用的统计步骤. 有许多实验是要作多次观测, 除了每次时间平均之外, 还得来一次总平均. 甚至整个过程还要在几个不同的物体上重复, 再来一次平均. 无论如何, 只要弄清楚统计步骤, 一切都没问题.

要适合上节所述独立的定义, 是不简单的事. 要检验一群变数是否互相独立, 非常困难. 事实上, 一般应用并不需要完全的独立性. 也就是说, 即使一群变量并不完全适合独立性的定义, 许多有用的结果也不会受太大影响.

我们先看个例子, 以说明独立性条件之不易适合.

【例】某样品集为 4 个球, 编号为 $S = 1, 2, 3, 4$. 分类为所有子集. 概率为子集中球数除以 4, 令 X, Y, Z 为三个变数, (三种样品性质). 变数各有二值, +1 或 -1, 我们指定: X 为颜色, +1 为黄, -1 为红. Y 为重量, +1 为重, -1 为轻. Z 为质料, +1 为硬, -1 为软.

资料列表如下:

特性 \ 球	球			
	1	2	3	4
色 X	黄 1	黄 1	红 -1	红 -1
量 Y	重 1	轻 -1	重 1	轻 -1
质 Z	硬 1	软 -1	软 -1	硬 1

显然, 黄红各半, 轻重各半, 硬软各半. 因此,

$$\rho(X) = \rho(Y) = \rho(Z) = \frac{1}{2}. \quad (25)$$

再者, 黄球中有一半是重的, 一半是轻的, 一半硬, 一半软, 红球中亦然. X, Y, Z 两两独立, 即

$$\begin{aligned}\rho(X|Y) &= \rho(X), \\ \rho(X|Z) &= \rho(X), \\ \rho(Y|Z) &= \rho(Y) \quad \text{等}.\end{aligned}\tag{26}$$

但是 X, Y, Z 却不是互相独立, 因为

$$\rho(X|Y, Z) \neq \rho(X)\tag{27}$$

又重又硬的只能是黄的? 没有红的, 轻软者也必是黄的.

并且, 黄轻必软, 红轻必硬, 等等. 因此, 虽然任两个变量都互相独立, 三个变量并不能互相独立.

由此例不难看出, 若有许多变数, 检验独立性质是一大难题. 如果证实了任两变量互相独立, 并不表示任三变量互相独立. 同理, 验证了三; 不足以知四.

我们要讨论的变量, 自然是有关各分子运动的变数. 由上节诸例子可见一斑. 变量太多, 要检验其独立性, 真是太困难了. 但是, 这种检验, 并不常需要. 例如, 上节中长链分子一例, 求 $\langle R^2 \rangle$. 只要

$$\sum_{i \neq j} \langle \mathbf{r}_i \cdot \mathbf{r}_j \rangle \ll Nl^2,\tag{28}$$

则结果 $\langle R^2 \rangle = Nl^2$ 就可以成立. 也就是说, 只要各两个 \mathbf{r}_i 独立就行了. 是否每三个独立, 每四个独立, 并没有关系. 在很多应用中, 每两个独立性检定了, 就够了. 当然, 也有许多情况要求得多些.

独立现象有一个十分重要的结果, 即中心极限定理, 我们在下一章讲它, 这定理可用来估计许多变数的独立性.

3. 不规则数列

上一章的第 2 节 (见 (10.12)) 曾讨论从电话本中抄下的一系列数字. 这列数字是“不规则数列”的一个例子. 这类数列, 又称“随机数列”, 列中之数曰“随机数”.

投骰子多次所现的点数,也是一列不规则数. 我们大致知道什么叫不规则,每次投骰子,都是“独立事件”. 这一次和上一次,下一次都“不相干”. 这样得出的数列是“不规则”的或“混乱的”.

我们常把“随机”和“随意”或“任意”或“不知”“不确定”混为一谈. 投骰子时是随意投,事先不知结果. 这种想法是不对的. 例如,我们“任意”写下的一列数字,大概都不够“不规则”. 如要一列够不规则的数列,我们必需要去查“随机数表”,或用计算器程序求出. 读者想想:查表,计算而得的数列,是再“确定”不过了. 任意写的,反而不合格.

投骰子求出的数列是够不规则了. 但请注意:我们需要骰子为工具. 骰子给出的数列,不得“任意”更改.

当然,我们需要对“不规则”有明确的定义. 用这定义来检验一数列,决定其“不规则”的程度. 我们需要以统计为定义之本. 数列必须很长. 定义很简单:令 $X_1, X_2, X_3 \dots$ 为某数列. 如果各数互相独立,则此数列为不规则. 互相独立的定义是由统计决定:

$$\text{令 } \langle f(X) \rangle \equiv \frac{1}{N} \sum_{i=k}^{N+k} f(X_i) \quad (29)$$

这平均值必须大致不随 N, k 改变. N 为一大数. 假设数列比 N 要长许多. 令

$$X_i(r) = X_{i+r}, \quad r = 0, 1, 2, \dots \quad (30)$$

如果

$$\begin{aligned} \frac{1}{N} \sum_{i=k}^{N+k} f(X_i) g(X_{i+r}) &\equiv \langle f(X(0))g(X(r)) \rangle \\ &= \langle f(X) \rangle \langle g(X) \rangle, \quad r = 1, 2, 3, \dots, \end{aligned} \quad (31)$$

则各 $X(r)$ 互相独立,即数列为不规则. f, g 为随便两个函数. 在许多应用,我们只须要考虑相干函数

$$C(r) \equiv \langle X(r)X(0) \rangle - \langle X \rangle^2. \quad (32)$$

$$\text{如果 } r \neq 0, \quad C(r) = 0, \quad (33)$$

则数列可视为不规则.

我们“任意”写一系列数字,大概不会通过 (31) 或 (33) 的考验. (10.12) 数列大致通过. (见 (10.19)). 和统计力学有关的数列,多是时间的函数即物体的

各运动变量. 虽不是一列分立数字, 但不规则性仍可定义. 例如 $n(t)$ 为气体中某小区域内的分子数. 我们可以定义它的时间相干函数:

$$\begin{aligned} C(t) &\equiv \langle n(t)n(0) \rangle - \langle n \rangle^2 \\ &\equiv \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} dt' [n(t'+t)n(t') - \langle n \rangle^2], \\ \langle n \rangle &\equiv \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} dt' n(t') \end{aligned} \quad (34)$$

观测时间 \mathcal{T} 须是相当长. 如果 $C(t)$ 只在 $t \leq \tau$ 时非零, 即

$$C(t) = 0, \quad t \gg \tau, \quad (35)$$

则我们称 τ 为相干时间. 在大于 τ 的时间尺度下, $n(t)$ 为不规则, 或先后独立.

根据许多实验的结果, 平衡中物体的各变量, 都有这先后独立的特性. (这特性的由来, 仍不清楚), 由于电子计算器的发明, 很长的随机数列 (可以百万计) 可以求得, 利用这些数列, 我们可以摹拟分子运动各变量, 以计算平衡性质, 这种计算, 将在第二十二章讨论.

4. 散射实验和相干函数的测定

要看某物体的一些运动细节, 我们可以射入一些粒子, (如光、中子、电子等), 测量被物体散射的粒子的动量、能量及自转的分布. 从测量的结果, 可以推测物体的一些底细. 散射实验直接测定的是某些相干值, 现在略略复习一点常识.

假定我们用中子照射某物体. 物体的众分子对中子的影响是一个对中子作用的位能 $U(\mathbf{x}, t)$. 假定这作用能是

$$U(\mathbf{x}, t) = \sum_i v(\mathbf{x} - \mathbf{r}_i(t)) \quad (36)$$

\mathbf{x} 是中子的位置, $\mathbf{r}_i(t)$ 是分子 i 在时刻 t 的位置. 中子和一个分子间的作用能为 $v(\mathbf{x} - \mathbf{r})$. $\mathbf{x} - \mathbf{r}$ 为中子和分子的相对位置. $U(\mathbf{x}, t)$ 为中子和所有分子的作用能. 假定各分子位置 $\mathbf{r}_i(t)$ 为已知. 略去中子自转, 中子的波动方程式是

$$i \frac{\partial}{\partial t} \psi(\mathbf{x}, t) = -\frac{1}{2m} \nabla^2 \psi(\mathbf{x}, t) + U(\mathbf{x}, t) \psi(\mathbf{x}, t), \quad (37)$$

m 为中子质量. 令 $a_{\mathbf{q}}(t)$ 为 ψ 的平面波展开系数:

$$\begin{aligned}\psi(\mathbf{x}, t) &= \sum_{\mathbf{q}'} a_{\mathbf{q}'}(t) e^{-i\epsilon' t} e^{i\mathbf{q}' \cdot \mathbf{x}}, \\ \epsilon' &\equiv \frac{\mathbf{q}'^2}{2m}.\end{aligned}\quad (38)$$

令入射中子的动量为 \mathbf{q} . 令 $t = 0$ 时

$$\psi(\mathbf{x}, 0) = e^{i\mathbf{q} \cdot \mathbf{x}}. \quad (39)$$

假设 U 为非常微弱的位能, 则 (37) 可作

$$i \frac{\partial}{\partial t} a_{\mathbf{q}'}(t) = \langle \mathbf{q}' | U | \mathbf{q} \rangle e^{i\omega t} + O(U^2), \quad (40)$$

$$\begin{aligned}\langle \mathbf{q}' | U | \mathbf{q} \rangle &= \int d^3x e^{-i\mathbf{k} \cdot \mathbf{x}} U(\mathbf{x}, t) \\ &\equiv U_{\mathbf{k}}(t),\end{aligned}\quad (41)$$

$$\begin{aligned}\mathbf{k} &\equiv \mathbf{q}' - \mathbf{q}, \\ \omega &\equiv \epsilon' - \epsilon = \frac{1}{2m} (\mathbf{q}'^2 - \mathbf{q}^2) \\ &= \frac{1}{2m} (\mathbf{q} + \mathbf{q}') \cdot \mathbf{k}.\end{aligned}\quad (42)$$

\mathbf{k} 是中子因散射而得的动量, ω 为所得的能量. 因此, 在时间 \mathcal{T} , 中子动量改变成 \mathbf{q}' 之幅度为 (40) 之解

$$a_{\mathbf{q}'}(\mathcal{T}) = -i \int_0^{\mathcal{T}} dt e^{i\omega t} U_{\mathbf{k}}(t), \quad (43)$$

它的平方除以 \mathcal{T} , 为单位时间的散射概率

$$\begin{aligned}P(\mathbf{q} \rightarrow \mathbf{q}') &= \frac{|a_{\mathbf{q}'}(\mathcal{T})|^2}{\mathcal{T}} \\ &= \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} dt_1 \int_0^{\mathcal{T}} dt_2 e^{i\omega(t_1 - t_2)} U_{\mathbf{k}}^*(t_2) U_{\mathbf{k}}(t_1) \\ &\equiv P(\mathbf{k}, \omega),\end{aligned}\quad (44)$$

$P(\mathbf{q} \rightarrow \mathbf{q}')$ 可以从实验求得. 不同的 \mathbf{q}, \mathbf{q}' 值导致不同的 \mathbf{k}, ω 值. 从 $P(\mathbf{k}, \omega)$, 我们可以求出一些相干函数. 我们先看对 ω 积分的结果.

(甲) 同时相干值

$$\text{因为 } \int_{-\infty}^{\infty} d\omega e^{i\omega t} = 2\pi\delta(t),$$

所以 (44) 的 ω 积分是

$$\begin{aligned} \int_{-\infty}^{\infty} d\omega P(\mathbf{k}, \omega) &= \frac{2\pi}{T} \int_0^T dt U_{\mathbf{k}}^*(t) U_{\mathbf{k}}(t) \\ &\equiv 2\pi \langle U_{\mathbf{k}}^* U_{\mathbf{k}} \rangle. \end{aligned} \quad (45)$$

注意: 除非特别注明, 尖括号 $\langle \dots \rangle$ 指时间平均值. $U_{\mathbf{k}}$ 和分子密度有直接关系, 由 (36) 得

$$\begin{aligned} U_{\mathbf{k}}(t) &= v_{\mathbf{k}} \rho_{\mathbf{k}}(t), \\ v_{\mathbf{k}} &= \int d^3\mathbf{x} e^{-i\mathbf{k}\cdot\mathbf{x}} v(\mathbf{x}), \\ \rho_{\mathbf{k}}(t) &= \int d^3\mathbf{x} e^{-i\mathbf{k}\cdot\mathbf{x}} \rho(\mathbf{x}, t), \\ \rho(\mathbf{x}, t) &= \sum_i \delta(\mathbf{x} - \mathbf{r}_i(t)). \end{aligned} \quad (46)$$

因此, 从 (45) 可得

$$A \equiv \frac{1}{2\pi |v_{\mathbf{k}}|^2} \int d\omega P(\mathbf{k}, \omega) = \langle \rho_{\mathbf{k}}^* \rho_{\mathbf{k}} \rangle. \quad (47)$$

令

$$\begin{aligned} \langle \rho_{\mathbf{k}}^* \rho_{\mathbf{k}} \rangle &\equiv |\langle \rho_{\mathbf{k}} \rangle|^2 + \langle \rho_{\mathbf{k}}^* \rho_{\mathbf{k}} \rangle_c, \\ \langle \rho(\mathbf{x}) \rho(\mathbf{x}') \rangle &\equiv \langle \rho(\mathbf{x}) \rangle \langle \rho(\mathbf{x}') \rangle + \langle \rho(\mathbf{x}) \rho(\mathbf{x}') \rangle_c, \end{aligned} \quad (48)$$

则

$$A = |\langle \rho_{\mathbf{k}} \rangle|^2 + \int d^3\mathbf{x} d^3\mathbf{x}' e^{-i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} \langle \rho(\mathbf{x}) \rho(\mathbf{x}') \rangle_c. \quad (49)$$

于是, 分子密度的起伏的相干函数可以从散射实验求得. 假定 $\rho(\mathbf{x})$ 和 $\rho(\mathbf{x}')$ 在 $|\mathbf{x} - \mathbf{x}'| \gg \xi$ (相干距离) 之外不相干. 再假定 ξ 为一微观尺度, 物体为均匀, 则 (49) 可以简化成

$$A = V \int d^3\mathbf{r} e^{-i\mathbf{k}\cdot\mathbf{r}} \langle \rho(\mathbf{r}) \rho(0) \rangle_c. \quad (50)$$

假定 $\mathbf{k} \neq 0$. V 是物体的容积.

(乙) 异时相干值

如果不对 ω 积分, 我们可以从 (44) 得到分子密度在不同时间的相干值. 分析如下:

令 $t = (t_1 + t_2) / 2, t' = t_1 - t_2$ 得

$$P(\mathbf{q} \rightarrow \mathbf{q}') = \frac{1}{\mathcal{T}} \int dt' e^{i\omega t'} \int dt U_{\mathbf{k}}^* \left(t - \frac{t'}{2} \right) U_{\mathbf{k}} \left(t + \frac{t'}{2} \right), \quad (51)$$

t, t' 的积分区域见图 1

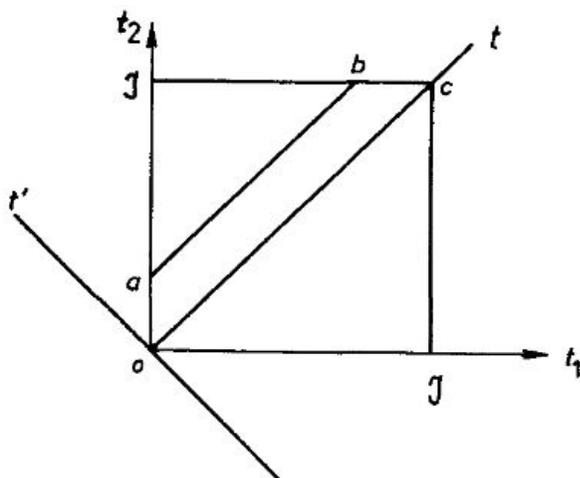


图 1

在一定的 t' , 求 dt 积分是沿着 45° 的斜线, 如图中的 a, b . 令

$$\begin{aligned} \frac{1}{t_b - t_a} \int_{t_a}^{t_b} dt U_{\mathbf{k}}^* \left(t - \frac{t'}{2} \right) U_{\mathbf{k}} \left(t + \frac{t'}{2} \right) \\ \equiv |\langle U_{\mathbf{k}} \rangle|^2 + \langle U_{\mathbf{k}}^* U_{\mathbf{k}} \rangle_c, \end{aligned} \quad (52)$$

平均值是在 t_a 到 t_b 这段时间内的平均值. (52) 的最后一项是 $U_{\mathbf{k}}$ 和它时间 t' 后之值的相干值. 现在假设: 如果 t' 比某相干时间 τ 长很多则这相干值趋于零. 并且, 假设 $\langle U_{\mathbf{k}} \rangle$ 约为一常数. 如果这些假设成立, 则 $t_b \approx \mathcal{T}, t_a \approx 0$. 将 (52) 代入 (51) 得

$$P(\mathbf{k}, \omega) = 2\pi\delta(\omega) |\langle U_{\mathbf{k}} \rangle|^2 + \int dt' e^{i\omega t'} \langle U_{\mathbf{k}}^* U_{\mathbf{k}}(t') \rangle_c. \quad (53)$$

用 (46), 这式子可化为

$$\frac{1}{|v_{\mathbf{k}}|^2} P(\mathbf{k}, \omega) = 2\pi\delta(\omega) |\langle \rho_{\mathbf{k}} \rangle|^2 + V \int dt d^3\mathbf{r} e^{i(\omega t - \mathbf{k} \cdot \mathbf{r})} \langle \rho(\mathbf{r}, t) \rho(0, 0) \rangle_c, \quad (54)$$

这是 (49) 的推广. (49) 是此式的 ω 积分. 这些结果把密度的起伏特性和散射率连起来. 相干值是对时间的平均值. 定义非常清楚. 这些结果并不涉及任何热力学概念, 更未涉及统计力学的基本假设.

讨论问题十一

1. 令 x_1, x_2, \dots, x_n 为独立变量. 求乘积

$$y = x_1 x_2 \dots x_n \quad (55)$$

之分布. 假定每个 x_i 的分布都是正态, 且 $\langle x_i^2 \rangle = 1, \langle x_i \rangle = 0$

2. 求

$$y = \left(\frac{1}{n} \sum_{i=1}^n x_i^2 \right)^{1/2} \quad (56)$$

之分布. x_i 如上题. (杨维哲, 104 页)

3. 令数列 x_1, x_2, x_3, \dots 由

$$x_{n+1} = 4\lambda x_n (1 - x_n), \quad n = 1, 2, 3, \dots \quad (57)$$

计算得来. x_1 和 λ 之值须先选定. 试用袖珍计算器求此数列. 取 $0 < x_1, \lambda < 1$. 在 λ 接近于 1 时, 此数列近乎不规则. (进一步分析, 见 Feigenbaum (1978)).

4. 令

$$x_n = \sin(2^n), \quad n = 1, 2, \dots \quad (58)$$

讨论此数列的不规则性.

5. 令随机变数 x 在 $0 < x < 1$ 中有均匀分布. 求

$$t \equiv \tau \ln \frac{1}{x} \quad (59)$$

的分布. τ 是一常数. 答: $\frac{1}{\tau} e^{-t/\tau}$, 即指数分布.

因此, 我们可以从平均分布的随机数列, 得到一指数分布的随机数列.

6. 正态分布将在第十二章详细讲, 在此先预习一下. 此分布该早已不陌生. 气体分子速度分布即为正态.

$$\rho(v_x) = \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-v_x^2/2\sigma^2} \quad (60)$$

$$\sigma^2 \equiv \langle v_x^2 \rangle = \frac{T}{m}.$$

求正态分布 $\rho(y)$ 的

(A) $\langle y^{2n} \rangle$

(B) $\langle e^{-iky} \rangle$.

7. 从平均分布的数列求正态分布数列.

这不像指数分布一般简单 (见问题 5), 但也不太难. 步骤如下:

(甲) 取平均分布的随机数 $0 < r < 1$, $0 < \theta < \pi$.

(乙) 令

$$Z = \left[2 \ln \frac{1}{r} \right]^{1/2}, \quad (61)$$

$$t = \sigma Z \cos \theta.$$

则 t 为正态分布之随机数, $\langle t^2 \rangle = \sigma^2$, $\langle t \rangle = 0$.

试导出以上步骤. 注意: $r = e^{-Z^2/2}$, 故 Z 的分布一如二度空间分子的速度分布. t 则是速度在一指定方向之投影. 读者可自创其他方法.

8. 如果 (36) 改成

$$U(\mathbf{x}, t) = \sum_i \delta(\mathbf{x} - \mathbf{r}_i(t)) \mathbf{s}_i(t) \cdot \boldsymbol{\sigma}(t), \quad (62)$$

即中子自转 $\boldsymbol{\sigma}$ 和分子自转 \mathbf{s}_i 有作用, 则散射率和分子的自转相干函数关系为何?

注意, 此地 U, σ 都是 2×2 方阵. 中子的波函数有二分量. 散射过程有自转不变的, 也有自转改变的. 分子自转可看成古典力学变数.

9. 试将第 4 节的内容全用量子力学分析.

10. 试证

$$\begin{aligned} & (A) \quad P(A_1, A_2, A_3, \cdots, A_n) \\ & \quad = P(A_1|A_2, A_3, \cdots, A_n) P(A_2|A_3, A_4, \cdots, A_n) \quad (63) \\ & \quad \cdots P(A_{n-2}|A_{n-1}, A_n) P(A_{n-1}|A_n) P(A_n). \end{aligned}$$

(B) 令 $A_1 \cup A_2 \cdots \cup A_n = E$. 且各类不相交, 即 $A_i \cap A_j = A_i \delta_{ij}$. 证明 (“贝叶斯定理”):

$$P(A_i|B) = \frac{P(A_i) P(B|A_i)}{\sum_j P(A_j) P(B|A_j)} \quad (64)$$

11. 某考生做答时有概率 $(1-p)$ 的情形是猜, p 的情形是知道如何答, 若是猜, 猜对的概率是 r , 今设此生答对了, 你想他是不是猜的?

12. 令 y_{ij} 为 $n \times n$ 方阵的各元素, $y_{ij} = +1$ 或 -1 , 各有概率 $1/2$, 各 y_{ij} 互相独立. 求这种方阵的行列式平方

$$\xi_n = [\text{Det}(y)]^2 \quad (65)$$

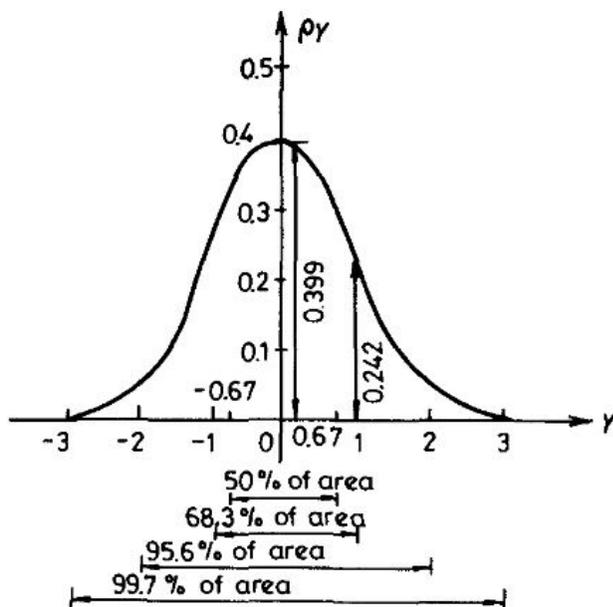
的平均值. 答: $\langle \xi_n \rangle = n!$.

(以上 3 题及解答, 见杨维哲 “概率论” 第 0 章).

第十二章 许多独立变量之和

独立性有一个十分重要的结果, 就是“中心极限定理”. 这定理指出: 如果 N 为一大数, 则 N 个大致独立的变量之和, 大致作正态分布. 分布之宽度与 \sqrt{N} 成正比, “大致”的意思, 当然会仔细讲. 我们用两种方法来推导这定理, 第一种直截了当, 第二种比较转弯抹角. 第一种见定理之大概, 第二种探其深处. 然后我们举例讨论这定理的应用. 我们穿插一点有关“连接值”的材料, 连接值在一般讨论展开计算法会十分有用. 在这一章我们用它来分析中心极限定理. 然后用这定理来讨论统计力学的基本假设, 并把这假设扩充, 使它在原则上和应用上都更趋完整有效.

1. 正态分布

图 1 正态分布 $\sigma = 1$

正态分布的定义是

$$\rho(Y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-Y^2/2\sigma^2}. \quad (1)$$

它的一些特性由图 1 可见.

它的特性也可以用各种平均值的表示:

$$\begin{aligned} \langle Y^2 \rangle &= \sigma^2, \\ \langle Y^{2n} \rangle &= \frac{(2n)!}{2^n n!} \sigma^{2n}, \\ \langle Y^{2n+1} \rangle &= 0, \quad n = 0, 1, 2, \dots \end{aligned} \quad (2)$$

这些结果可直接由 (1) 积分得来, 它们也可以用一个十分简单的式子表示:

$$\begin{aligned} f(k) &= \langle e^{-ikY} \rangle = e^{-k^2\sigma^2/2}, \\ \rho(Y) &= \frac{1}{2\pi} \int dk e^{ikY} f(k), \end{aligned} \quad (3)$$

$f(k)$ 是 $\rho(Y)$ 的“特性函数”. 把它展开成 k 的级数, 即得 (2).

2. 中心极限定理, 简单说法

令 X_1, X_2, \dots, X_N 为 N 个互相独立的变量. 假设 $\langle X_i \rangle = 0$. 令

$$Y \equiv \frac{1}{\sqrt{N}} (X_1 + X_2 + \dots + X_N). \quad (4)$$

现在来求 Y 的特性函数:

$$\begin{aligned} \langle e^{-ikY} \rangle &= \left\langle e^{-ik(X_1+X_2+\dots+X_N)/\sqrt{N}} \right\rangle \\ &= \left\langle \prod_{j=1}^N e^{-ikX_j/\sqrt{N}} \right\rangle \\ &= \prod_{j=1}^N \left\langle e^{-ikX_j/\sqrt{N}} \right\rangle \\ &= \exp \sum_{j=1}^N A_j(k/\sqrt{N}), \end{aligned} \quad (5)$$

$$A_j(k/\sqrt{N}) \equiv \ln \left\langle e^{-ikX_j/\sqrt{N}} \right\rangle. \quad (6)$$

假设 N 为大数, 把 $A_j(k/\sqrt{N})$ 展开, 得

$$A_j(k/\sqrt{N}) \approx A_j(0) + \frac{k^2}{2N} A_j''(0) + \frac{k^3}{N} O(N^{-1/2}). \quad (7)$$

注意: k 的一次项是 0, 因为它和 $\langle X_j \rangle = 0$ 成正比, 二次项是

$$A_j''(0) = -\langle X_j^2 \rangle. \quad (8)$$

这些可以直接由 (6) 得来. 将 (7) 代入 (5), 即得

$$\begin{aligned} \langle e^{-ikY} \rangle &= \exp \left\{ -\frac{1}{2} k^2 \sigma^2 + O\left(\frac{1}{\sqrt{N}}\right) \right\}, \\ \sigma^2 &\equiv \frac{1}{N} \sum_{j=1}^N \langle X_j^2 \rangle. \end{aligned} \quad (9)$$

这就是“中心极限定理”, 即 Y 的分布为正态, 因为 (9) 和 (3) 一样, 只要我们略去 $O(1/\sqrt{N})$. 如果 $\langle X_j \rangle \neq 0$ 则

$$\langle Y \rangle = \frac{1}{\sqrt{N}} \sum_{j=1}^N \langle X_j \rangle, \quad (10)$$

$Y - \langle Y \rangle$ 之分布为正态, 把 (9) 的 Y 改为 $Y - \langle Y \rangle$ 既可.

即然 Y 之分布为正态, 宽 σ , 所以 Y/\sqrt{N} 之分布亦然. 其宽为 σ/\sqrt{N} , 变数 Y/\sqrt{N}

$$\frac{Y}{\sqrt{N}} = \frac{1}{N} (X_1 + X_2 + \cdots + X_N) \quad (11)$$

可以说是 N 个 X_i 的“平均”, 如果 N 很大, 则 σ/\sqrt{N} 很小, 即“平均”和这“平均”的平均值

$$\left\langle \frac{1}{N} (X_1 + X_2 + \cdots + X_N) \right\rangle \quad (12)$$

相差甚微, 即起伏很小. 也就是说, 把 N 个独立变量加起来, 除以 N , 各 X_i 的起伏就大致抵消掉了. 当然 $\sqrt{N}Y = X_1 + X_2 + \cdots + X_N$ 的分布宽度是 $\sqrt{N}\sigma$. 把 N 个独立变量加起来, 起伏之增长是 \sqrt{N} , 不是 N .

注意: (9) 的 k 不可以太大, 如果 $k \geq \sqrt{N}$, 则 (9) 不成立. 因此, 所谓 Y 的正态分布, 是限于 Y 的低次方平均值之计算. 现在看些例子.

【例】二项分布

在第十一章 (见 (11.10) 左右的讨论) 我们曾讨论

$$n = \sum_{i=1}^N n_i, \quad (13)$$

$$n_i = 0, \text{ 或 } 1,$$

的分布, $n_i = 1$ 的概率为 p , 且各 n_i 互相独立, 结果 n 的分布为二项分布

$$\rho(n) = \binom{N}{n} p^n q^{N-n}, \quad (14)$$

$$q = 1 - p.$$

用中心极限定理, 得

$$\rho(n) \approx \frac{1}{\sqrt{2\pi N\sigma^2}} e^{-(n-\langle n \rangle)^2/2N\sigma^2} \quad (15)$$

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (\langle n_i^2 \rangle - \langle n_i \rangle^2)$$

$$= p(1-p) = pq. \quad (16)$$

表 1 示 $N = 2500$, $p = 0.02$ 的 $\rho(n)$, 大致告诉我们 (14) 和 (15) 有多接近. 注意

$$\begin{aligned}\langle n \rangle &= 50, \\ N\sigma^2 &= 49.\end{aligned}\tag{16}$$

表 1

n	(14)	(15)
25	0.0000	0.0001
30	0.0006	0.0010
35	0.0052	0.0057
40	0.0212	0.0205
45	0.0460	0.0442
50	0.0569	0.0570
55	0.0424	0.0442
60	0.0199	0.0205
65	0.0061	0.0057
70	0.0013	0.0010
75	0.0002	0.0001

【例】令变数 $X_i, i = 1, 2, \dots, N$ 之分布为

$$\rho(X_i) = \frac{a_i}{\pi(a_i^2 + X_i^2)},\tag{18}$$

a_i 为正值常数. 求各 X_i 之总和之分布.

令总和为

$$X = X_1 + X_2 + \dots + X_N.\tag{19}$$

因为 $\langle X_i^2 \rangle$ 为无穷大, 故中心极限定理不能用. 读者试证: X 之分布为

$$\begin{aligned}\frac{Na}{\pi(N^2a^2 + X^2)}, \\ a \equiv \frac{1}{N} \sum_{i=1}^N a_i.\end{aligned}\tag{20}$$

3. 高次平均值和连接值

我们先把相干值的定义推广. 先前所定

$$\langle X^2 \rangle = \langle X \rangle^2 + \langle X^2 \rangle_c,$$

即

$$\langle X^2 \rangle_c = C \equiv \langle X^2 \rangle - \langle X \rangle^2. \quad (21)$$

$\langle X^2 \rangle_c$ 可叫做二次“相干值”. 我们再给它多加一个名字, 叫“连接值”. 不同变数的二次相干值或连接值 $\langle X_i X_j \rangle_c$ 的定义是

$$\langle X_i X_j \rangle = \langle X_i \rangle \langle X_j \rangle + \langle X_i X_j \rangle_c. \quad (22)$$

如果 X_i, X_j 互相独立, 则其相干值为 0, 或连接值为 0. (21) 是 (22) 的特殊情形, $i = j$ 的情况. 再推广, 令

$$\begin{aligned} \langle X_i X_j X_k \rangle &= \langle X_i \rangle \langle X_j \rangle \langle X_k \rangle \\ &+ \langle X_i X_j \rangle_c \langle X_k \rangle + \langle X_i X_k \rangle_c \langle X_j \rangle \\ &+ \langle X_j X_k \rangle_c \langle X_i \rangle + \langle X_i X_j X_k \rangle_c. \end{aligned} \quad (23)$$

最后一项是“三次连接值”.

三次的情形, 比二次复杂多了. 三个中如果有两个不互相独立, 则三个就不是独立的了. X_i, X_j, X_k 三者之相干, 有来自两两相干者, 除此之外是来自“纯净”的三者相干, 是为“三次连接值” $\langle X_i X_j X_k \rangle_c$. 如果 X_i, X_j, X_k 互相独立, 则三者中任两者必互相独立, 即 $\langle X_i X_j \rangle_c = \langle X_j X_k \rangle_c = \langle X_k X_i \rangle_c = 0$. 不过, 如果任两变量独立, 并不保证三变量互相独立 (见上一章第 2 节). 因此, 讨论二次以上乘积的平均值, 我们可以把“相干值”看成一些“连接值”乘积的组合.

四次连接值 $\langle X_i X_j X_k X_l \rangle_c$ 可照样定义:

$$\begin{aligned} \langle X_i X_j X_k X_l \rangle &= \langle X_i \rangle \langle X_j \rangle \langle X_k \rangle \langle X_l \rangle \\ &+ \langle X_i X_j \rangle_c \langle X_k \rangle \langle X_l \rangle + \cdots \\ &+ \langle X_i X_j \rangle_c \langle X_k X_l \rangle_c + \cdots \\ &+ \langle X_i X_j X_k \rangle_c \langle X_l \rangle + \cdots \\ &+ \langle X_i X_j X_k X_l \rangle_c. \end{aligned} \quad (24)$$

上述的特殊情形, 如 (23) (24) 中各变量相等, 则

$$\begin{aligned} \langle X^3 \rangle &= \langle X \rangle^3 + 3 \langle X^2 \rangle_c \langle X \rangle + \langle X^3 \rangle_c \\ \langle X^4 \rangle &= \langle X \rangle^4 + 6 \langle X^2 \rangle_c \langle X \rangle^2 + 3 \langle X^2 \rangle_c^2 \\ &\quad + 4 \langle X^3 \rangle_c \langle X \rangle + \langle X^4 \rangle_c. \end{aligned} \tag{25}$$

(22)–(25) 叫做多次平均值的“连接值展开式”. 这些展开式可以用简图代表 (见图 2).

这些图有助于记忆和分析, 每个连接值都以一组连接起来的点代表, 分离的是各变量个别的平均值.

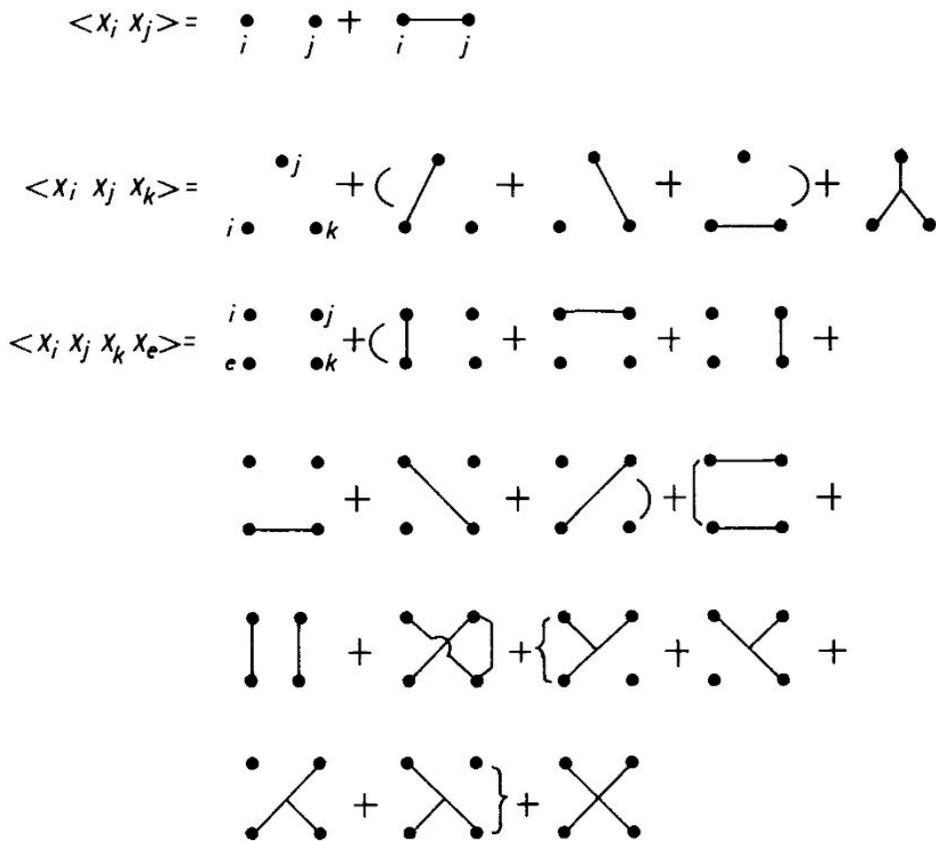


图 2 连接值展开式

更高次的平均值的展开式更长, 在下一节我们会讨论合用于所有次数的展开定理. 现在先看看 $\langle X^2 \rangle_c$ 在高次方平均值里的地位. 令 $\langle X_i \rangle = 0$ (如果 $\langle X_i \rangle \neq 0$, 我们可以把 $\langle X_i \rangle$ 从 X_i 中减去, 则新变数之平均值为 0. 因此, 要求

$\langle X_i \rangle = 0$, 算不了什么.) 在 $2n$ 次平均值的展开式中, 有 A_n 项只含二次连接值,

$$A_n = \frac{(2n)!}{2^n n!}. \quad (26)$$

A_n 这数目自然是把 $2n$ 变数配成对的配法总数. 因此, 如有一个变数 X , 则 $\langle X^{2n} \rangle$ 的展开式为

$$\langle X^{2n} \rangle = \frac{(2n)!}{2^n n!} \langle X^2 \rangle_c^n + R. \quad (27)$$

R 中各项, 每项含有至少一个至少三次或四次的连接值, 同理

$$\langle X^{2n+1} \rangle = \left(\frac{2n+1}{3} \right) A_{n-1} \langle X^3 \rangle_c \left[\langle X^2 \rangle_c^{n-1} + R' \right], \quad (28)$$

R' 中每项含有至少一个至少三次或四次的连接值.

读者请注意:

① 各连接值可以写成各平均值之展开式, 把 (21)–(25) 倒过来解就是. 例如:

$$\begin{aligned} \langle X_i X_j \rangle_c &= \langle X_i X_j \rangle - \langle X_i \rangle \langle X_j \rangle, \\ \langle X^3 \rangle_c &= \langle X^3 \rangle - 3 \langle X^2 \rangle \langle X \rangle + 2 \langle X \rangle^3. \end{aligned} \quad (29)$$

② 如果 Y_i 为 X_i 的线性组合, 即

$$Y_i = \sum_j \alpha_{ij} X_j, \quad (30a)$$

α_{ij} 为常数, 则

$$\langle Y_i Y_j \rangle_c = \sum_{k,l} \alpha_{ik} \alpha_{jl} \langle X_k X_l \rangle_c, \quad (30b)$$

更高次的亦然, 读者试证之.

如果 $Y = f(X)$ 不是 X 的线性组合, 则像 (30b) 这种结果就不成立. 例如:

$$\begin{aligned} \text{令 } \langle X \rangle &= 0; \quad Y = X^2. \\ \text{则 } \langle Y^2 \rangle_c &= \langle Y^2 \rangle - \langle Y \rangle^2 \\ &= \langle X^4 \rangle - \langle X^2 \rangle^2, \\ \text{而 } \langle X^4 \rangle_c &= \langle X^4 \rangle - 3 \langle X^2 \rangle^2. \\ \text{所以 } \langle Y^2 \rangle_c &\neq \langle X^4 \rangle_c. \end{aligned} \quad (31)$$

也就是说 $\langle \dots \rangle_c$ 下这个 c 字是要靠变量来定义的. 变量一改, 定义也改, 只有在 (30) 线性组合的情形下不改.

4. 连接值展开定理

这定理是

$$\begin{aligned} f_c(k) &\equiv \ln \langle e^{-ikX} \rangle \\ &= \sum_{n=1}^{\infty} \frac{(-ik)^n}{n!} \langle X^n \rangle_c \\ &= \langle (e^{-ikX} - 1) \rangle_c. \end{aligned} \quad (32)$$

$f_c(k)$ 可称作 $\rho(X)$ 的“连接特性函数”.

这定理证明如下:

先把特性函数展开

$$\begin{aligned} f(k) &\equiv \langle e^{-ikX} \rangle \\ &= \sum_{m=0}^{\infty} \frac{(-ik)^m}{m!} \langle X^m \rangle. \end{aligned} \quad (33)$$

再把 $\langle X^m \rangle$ 写成连接值之组合, 像图 1. 我们画 m 个点, 把这些点分组, 每组内的点连接起来, 各组间不连, 每一组相连的点为一连接值. 令 m_n 为含有 n 点的组数, 即有 m_1 组有一点, m_2 组有两点, 等等. 各组点数之和是:

$$m = \sum_n n m_n. \quad (34)$$

m_n 可以是零, 或正整数. 各 m_n 指定了, 仍有许多不同的分组法, 有多少? 有

$$\frac{m!}{m_1!(1!)^{m_1} m_2!(2!)^{m_2} \dots m_n!(n!)^{m_n} \dots}. \quad (35)$$

(注意: 如果 $m_i = 0$ 则 $0!$ 定为 1). 这结果很容易了解: m 点有 $m!$ 种排列. 每组中各点交换, 不产生新组, 故要除以各 $n!$. 同数各组交换亦不算新的分组, 故除以各 $m_n!$, 因此

$$\begin{aligned} \langle X^m \rangle &= m! \sum_{m_1, m_2, \dots} \frac{1}{m_1!} \left(\frac{\langle X \rangle}{1!} \right)^{m_1} \frac{1}{m_2!} \left(\frac{\langle X^2 \rangle_c}{2!} \right)^{m_2} \dots \frac{1}{m_n!} \left(\frac{\langle X^n \rangle_c}{n!} \right)^{m_n} \\ &\quad \dots \delta(m - m_1 - 2m_2 - 3m_3 - \dots - nm_n - \dots). \end{aligned} \quad (36)$$

对 m_1, m_2, \dots 的合计, 各项必须适合 (34). 将 (36) 代入 (33) 得

$$\begin{aligned}
 f(k) &= \langle e^{-ikX} \rangle \\
 &= \sum_m (-ik)^m \sum_{m_1, m_2, \dots} \delta(m - m_1 - \dots - nm_n - \dots) \\
 &\quad \cdot \frac{1}{m_1! m_2! \dots m_n! \dots} \left(\frac{\langle X \rangle}{1!} \right)^{m_1} \dots \left(\frac{\langle X^n \rangle_c}{n!} \right)^{m_n} \dots \\
 &= \prod_{n=0}^{\infty} \sum_{m_n=0}^{\infty} \frac{1}{m_n!} \left(\frac{\langle X^n \rangle_c}{n!} \right)^{m_n} (-ik)^{nm_n} \\
 &= \exp \left\{ \sum_{n=1}^{\infty} \frac{(-ik)^n}{n!} \langle X^n \rangle_c \right\} \\
 &= \exp \{ f_c(k) \}.
 \end{aligned} \tag{37}$$

(32) 就由此得证, 这是一个非常重要的结果.

以上这定理可以推广到多变量的展开:

$$\begin{aligned}
 \ln \left\langle \exp \left\{ -i \sum_{\alpha} k_{\alpha} X_{\alpha} \right\} \right\rangle \\
 = \left\langle \exp \left\{ -i \sum_{\alpha} k_{\alpha} X_{\alpha} \right\} - 1 \right\rangle_c.
 \end{aligned} \tag{38}$$

读者试证之.

5. 中心极限定理, 广义说法

现在来看正态分布的各连接值. 各次的平均值 (见 (2)) 可以和 (27), (28) 比较, 或者用 (2) 和 (32), 得

$$f_c(k) = -\frac{1}{2} k^2 \sigma^2, \tag{39}$$

即

$$\begin{aligned}
 \langle Y^2 \rangle_c &= \sigma^2, \\
 \langle Y^n \rangle_c &= 0, \quad n = 3, 4, 5, \dots
 \end{aligned} \tag{40}$$

也就是说, 除了二次外, 所有连接值都是 0, 这就是正态分布性质之总结.

现在来讲中心极限定理. 以下的推导是要把这定理的条件和伸缩性尽量说清楚. 令 X_1, X_2, \dots, X_N 为 N 个变量, N 相当大, 且

$$\begin{aligned}
 & \text{(甲)} \quad \langle X_i \rangle = 0, \\
 & \text{(乙)} \quad \frac{1}{N} \sum_{i,j} \langle X_i X_j \rangle_c \equiv \sigma^2 = O(1), \\
 & \text{(丙)} \quad \frac{1}{N} \sum_{i,j,k} \langle X_i X_j X_k \rangle_c \equiv \mu_3 \leq O(1), \\
 & \frac{1}{N} \sum_{j_1, \dots, j_n} \langle X_{j_1} X_{j_2} \cdots X_{j_n} \rangle_c \equiv \mu_n \leq O(N^{(n-3)/2}).
 \end{aligned} \tag{41}$$

$$\text{令 } Y = \frac{1}{\sqrt{N}} (X_1 + \cdots + X_n). \tag{42}$$

(定理): 如果 (甲)(乙)(丙) 成立, 则

$$\begin{aligned}
 & \text{(i)} \quad \langle Y \rangle = 0, \\
 & \text{(ii)} \quad \langle Y^2 \rangle_c = \sigma^2, \\
 & \text{(iii)} \quad \langle Y^n \rangle_c \leq O\left(\frac{1}{\sqrt{N}}\right), \quad n = 3, 4, \dots.
 \end{aligned} \tag{43}$$

这头两行 (i)(ii) 太明显了. (iii) 是这定理的本身, 我们先看 $\langle Y^3 \rangle_c$: 由 (42) 得

$$\langle Y^3 \rangle_c = \frac{1}{N^{3/2}} \sum_{i,j,k} \langle X_i X_j X_k \rangle_c = \frac{\mu_3}{\sqrt{N}}, \tag{44}$$

$n > 3$ 只是 (44) 的推广:

$$\begin{aligned}
 \langle Y^n \rangle_c &= \frac{1}{N^{n/2}} \sum_{j_1, \dots, j_n} \langle X_{j_1} X_{j_2} \cdots X_{j_n} \rangle_c \\
 &= \frac{\mu_n}{N^{n/2-1}}.
 \end{aligned} \tag{45}$$

因此, 只要 (41) 成立, (43) 就成立, 定理就这样证明了. 现在仔细看看它的含义:

① 结论 (43) 的含义是: 因为除了 $O(1/\sqrt{N})$, (40) 和 (43) 相同. 所以, Y 的各次平均值可以由正态分布来计算, 只要 $O(1/\sqrt{N})$ 可略去. 有一个明显的限制, 即高次的平均值不能用这结果. 因为 $\langle Y^n \rangle$ 的连接值展开会有很多项, 当项数有 \sqrt{N} 这么多, 这 $O(1/\sqrt{N})$ 就不能忽略. 因此, Y 的分布可视为“正态”, 只限于低次平均值之计算.

② 如果 X_1, \dots, X_N 为互相独立的变量, (41) 即成立. $\sigma^2, \mu_n, n = 3, 4, \dots$ 都是 $O(1)$.

③ 如果它们不互相独立, 但每个变量只和至多 m 个其他变量相干, 且 $m \ll N$, 则 (乙) 式计和中至多有 Nm^2 项. (丙) 式至多有 Nm^n 项, (41) 也成立, 定理也可以用.

④ 从一个实用的观点来看, 我们要算的常常只是 Y 的低次方的平均值而已. 因此, 只要 (41) 的 (甲)(乙) 及 (丙) 中的少数几个 n 成立, 定理就可以用了, 用不着考虑 X_1, \dots, X_N 是否真独立. 这是这定理所以有用的原因. 我们不苛求, 则它要求得也很少.

事实上, 要检验一变量是否独立, 是难以做到的事, 如果 N 很大, 则更不用说. 因此, 我们用中心极限定理时, 都用以上的实用观点, 这就是本章开始时提到的“大致”独立, 和“大致”正态分布的意义.

这个定理不但很有实用价值, 而且对一些基本概念, 也可作非常重要的指示. 现在来看些示范例子.

6. 重复试验和概率的测定

“重复实验”在许多概率的应用上很重要. 例如, $1/6$ 为得到骰子上 1 点的概率. 这 $1/6$ 的意思是: N 次重复实验的结果, “约有” $N/6$ 次出现 1 点. 这“约有”是什么意思? 照中心极限定理, “约有”是指 $(1/\sqrt{N})$ 的百分误差. 说的要仔细些: 投 N 次骰子, 令出现 1 点次数为 Y

$$Y = \sum_{i=1}^N n_i \quad (46)$$

$n_i = 1$, 如果第 i 次得 1 点,

否则 $n_i = 0$.

如果 N 很大, 则 Y/N 趋近于 $1/6$. 如何趋近? 我们把 Y 的试验重复 M 回, (每回投 N 次) 得 M 个 Y 值, Y_1, Y_2, \dots, Y_M . 根据中心极限定理, 这些值的分布该是正态, 且 Y/N 的分布宽度该是

$$\sigma = \frac{1}{\sqrt{N}} \left(\frac{5}{36} \right)^{1/2}. \quad (47)$$

这个值可以由这 M 回试验测定

$$\begin{aligned}\sigma^2 &\equiv \left\langle \left(\frac{Y}{N} - \frac{1}{6} \right)^2 \right\rangle \\ &= \frac{1}{M} \sum_{\alpha=1}^M \left(\frac{Y_{\alpha}}{N} - \frac{1}{6} \right)^2.\end{aligned}\quad (48)$$

也就是说, N 越大, Y/N 越和 $1/6$ 接近. 如果不是越接近, 则 $1/6$ 这数字的意义就成了问题. 每次投骰子, 都被看成一“独立事件”. 而一般讲概率, 多假设很多次重复独立实验, 某事件的“发生概率”的定义只能正确到 $O(1/\sqrt{N})$, N 是重复试验的次数. 如果不重复多次, 则这“发生概率”只是一个数学名词而已.

以上的讨论用到了中心极限定理. 由此看来, 概率本身的测定就需要这定理, 而这定理的导出先要有概率. 因此, 这定理和概率这概念本身是分不开的.

中心极限定理在某些情况下不能用, 那么在这些情况下, 概率本身的定义是否就成了问题? 我们可用一例子来讨论这问题. 令某变数 Y 的分布为

$$\rho(Y) = \frac{1}{\pi(Y^2 + 1)}.\quad (49)$$

这分布的 $\langle Y^2 \rangle$ 为无穷大, 如果做 N 次试验, 得 Y_1, Y_2, \dots, Y_N , 令

$$Z = \frac{1}{N} \sum_{i=1}^N Y_i.\quad (50)$$

根据第 2 节的例子, (见 (18)—(20)) Z 的分布和 (49) 一样, 即

$$\rho(Z) = \frac{1}{\pi(Z^2 + 1)},\quad (51)$$

和 N 无关, 也就是说, 无论做多少次试验, Y 的平均值都无法测准! 这和上面的 $O(1/\sqrt{N})$ 完全不同. 如果连 Y 的平均值都无法测定, 概率有何意义?

以上的结果 (51) 确是告诉我们不能用重复实验结果来计算 Y 的平均值. 不过, 用重复实验来定 $\rho(Y)$ 仍是可以的, 做 N 次试验, 统计 Y 发生在 a, b 之间的次数 $n(a, b)$:

$$\begin{aligned}n(a, b) &= \sum_{i=1}^N n_i \\ n_i &= 1, \text{ 如果 } a < Y_i < b \\ &\text{否则 } = 0.\end{aligned}\quad (52)$$

n 是 N 个独立变量之和, 且 $\langle n_i^2 \rangle < 1$ 为有限, 所以中心极限定理可用:

$$\frac{n(a, b)}{N} = \int_a^b dY \rho(Y) + O\left(\frac{1}{\sqrt{N}}\right). \quad (53)$$

从不同的 a, b 值, 及 (53), $\rho(Y)$ 可以测定. 虽然 $\rho(Y)$ 测定, 但都不是精确到可以求平均值

$$\int_{-\infty}^{\infty} dY \rho(Y) Y \quad (54)$$

的地步, 因为 $Y \geq N$ 的概率是 $O(1/N)$, 但对 (54) 的影响是 $NO(1/N) = O(1)$. 这例子是告诉我们, 如果某变量 X 的 $\langle X^2 \rangle_c \rightarrow \infty$, 或 $\langle X^3 \rangle_c \rightarrow \infty$, 则 $\langle X \rangle$ 的意义也就有了问题.

7. 不规则运动, 扩散

假想一群粒子, 如花粉, 重分子之类, 游于水中不断地被水分子碰. 令 \mathbf{p} 为某粒子的动量, 则

$$\frac{d\mathbf{p}}{dt} = -\alpha\mathbf{p} + \mathbf{f}(t) \quad (55)$$

是该粒子所受的力, 这力自然是相当复杂. (55) 是一假设, $-\alpha\mathbf{p}$ 是摩擦阻力. \mathbf{f} 则假设为不规则. (55) 之解为

$$\mathbf{p}(t) = \int_0^{\infty} dt' e^{-\alpha t'} \mathbf{f}(t - t'). \quad (56)$$

现在假设 \mathbf{f} 作快速起伏, $\langle \mathbf{f} \rangle = 0$, 有相干时间 τ , 且

$$\frac{1}{\alpha} \gg \tau. \quad (57)$$

也就是, 在时间 $1/\alpha$ 内, 起伏许多次. 相干时间的意义见上一章末节. 令

$$\begin{aligned} C(t) &= \frac{1}{3} \langle \mathbf{f}(t) \cdot \mathbf{f}(0) \rangle \\ &\equiv \frac{1}{3\mathcal{T}} \int_0^{\mathcal{T}} dt' \mathbf{f}(t' + t) \cdot \mathbf{f}(t'), \end{aligned} \quad (58)$$

\mathcal{T} 是一段很长的观测时间, $\mathcal{T} \gg 1/\alpha$. 如果 $|t| > \tau$, 则 $C(t) \approx 0$, \mathbf{f} 在时间 τ 之外, 是先后独立. 因为 $1/\alpha \gg \tau$; 所以 (56) 指出 $\mathbf{p}(t)$ 可以看成是许多独立

变量之和. 不同时间的 \mathbf{f} (相隔 τ 以上) 是独立的. 因此, 我们可以用中心极限定理得 p_x, p_y, p_z 的分布:

$$\begin{aligned}\rho(\mathbf{p}) &= \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^3 e^{-p_x^2/2\sigma^2} e^{-p_y^2/2\sigma^2} e^{-p_z^2/2\sigma^2} \\ &= \frac{1}{(2\pi\sigma^2)^{3/2}} e^{-p^2/2\sigma^2},\end{aligned}\quad (59)$$

$$\begin{aligned}\sigma^2 &= \langle p_x^2 \rangle = \langle p_y^2 \rangle = \langle p_z^2 \rangle \\ &= \int_0^\infty dt'_1 dt'_2 e^{-\alpha(t'_1+t'_2)} C(t'_1 - t'_2).\end{aligned}\quad (60)$$

上式由 (56) 的平方及 (58) 得来. 因为 $1/\alpha \gg \tau$, 上式积分是

$$\begin{aligned}\langle p_x^2 \rangle &= \int_0^\infty dt e^{-2\alpha t} \int_{-2t}^{2t} dt'' C(t'') \\ &\approx \frac{1}{2\alpha} \int_{-\infty}^\infty dt'' C(t'').\end{aligned}\quad (61)$$

根据能量均分定理, $\langle p_x^2 \rangle = mT$. 因此, 从 (61) 得

$$\alpha = \frac{1}{2Tm} \int_{-\infty}^\infty dt'' C(t'').\quad (62)$$

这是个十分重要的结果, 它把摩擦系数和液体分子运动造成的起伏连起来了, 是为“起伏消耗定理”, 以后再详谈.

用中心极限定理, 也可以求得粒子位移的分布. 令 $\mathbf{v} = \mathbf{p}/m$ 为某粒子的速度, 则它从时刻 t 到 $t+t'$ 的位移是

$$\begin{aligned}\Delta \mathbf{r} &\equiv \mathbf{r}(t+t') - \mathbf{r}(t) \\ &= \int_0^{t'} dt'' \mathbf{v}(t''+t).\end{aligned}\quad (62')$$

$\mathbf{v}(t'')$ 为一不规则的速度, 它的时间相干函数

$$\langle \mathbf{v}(t') \mathbf{v}(0) \rangle = \langle v^2 \rangle e^{-\alpha t'}\quad (63)$$

这结果是由 (55) 之解

$$\mathbf{v}(t'+t) = \mathbf{v}(t)e^{-\alpha t'} + \frac{1}{m} \int_0^{t'} dt'' e^{-\alpha t''} \mathbf{f}(t'+t-t'').$$

得来. 假设速度 $\mathbf{v}(t)$ 和 t 以后的 \mathbf{f} 不相干. 如果 (62) 的 t' 比 $1/\alpha$ 大很多, 则 $\Delta\mathbf{r}$ 是很多独立变量之和. 根据中心极限定理, $\Delta\mathbf{r}$ 的分布是正态:

$$\rho(\Delta\mathbf{r}) = \frac{1}{(2\pi\sigma^2)^{3/2}} e^{-(\Delta\mathbf{r})^2/2\sigma^2} \quad (64)$$

$$\sigma^2 = \frac{1}{3} \langle (\Delta\mathbf{r})^2 \rangle.$$

σ^2 可以由 (62) 及 (63) 得来.

$$\begin{aligned} \frac{1}{3} \langle (\Delta\mathbf{r})^2 \rangle &= \frac{1}{3} \int_0^{t'} dt_1 \int_0^{t''} dt_2 \langle \mathbf{v}(t_1) \cdot \mathbf{v}(t_2) \rangle \\ &= \frac{1}{3} \int_0^{t'} dt_1 \int_{-2t}^{2t} dt'' \langle v^2 \rangle e^{-\alpha|t''|} \\ &\approx \frac{2}{3} \langle v^2 \rangle \frac{t'}{\alpha}. \end{aligned} \quad (65)$$

这结果可以写成比较熟知的扩散公式:

$$\rho(\Delta\mathbf{r}, t') = \frac{1}{(4\pi Dt')^{3/2}} e^{-(\Delta\mathbf{r})^2/4Dt'}, \quad (66)$$

$$D \equiv \frac{\langle v^2 \rangle}{3\alpha} = \frac{T}{m\alpha}.$$

D 是扩散系数.

以上各量的平均值, 都是对时间的平均. 如果同时观测许多粒子, 则平均值看做是对这些粒子的平均也是一样. 如果在小区域 $d^3\mathbf{r}'$ 内的粒子数为 $n(\mathbf{r}')d^3\mathbf{r}'$, 则时间 t 后这些粒子的位置分布为

$$n(\mathbf{r}, t) = \int d^3r' \rho(\mathbf{r} - \mathbf{r}', t) n(\mathbf{r}') \quad (67)$$

这 $n(\mathbf{r}, t)$ 是扩散方程式

$$\frac{\partial}{\partial t} n(\mathbf{r}, t) = D\nabla^2 n(\mathbf{r}, t) \quad (68)$$

之解.

8. 宏观变数的起伏

在第六章, 最后一节, 我们提到, 大值变量分布总是正态分布. (见 (6.86)). 这分布是指活动范围各状态的统计结果, 这结论只适用于大值 (N 量级) 变量, 这样的正态分布, 是不是可以和中心极限定理扯上关系呢?

现在举例讨论这问题. 把某物体分成两半 1、2. 令

$$\hat{A}(s) = H_1(s_1) - H_2(s_2) \quad (69)$$

H_1, H_2 为两部的能量, 这个变量在第六章分析过, (见 (6.84), (6.85) 及讨论) 它的分布是

$$\frac{\Gamma(E, A)}{\Gamma(E)} = \frac{1}{\sqrt{2\pi K_A}} e^{-(A-\bar{A})^2/2K_A}. \quad (70)$$

\bar{A} 是 (6.85) 之解. 在此 $\bar{A} = 0$, 因为两部份一样大. K_A 的定义是

$$-\frac{1}{K_A} = \frac{\partial^2 S}{\partial A^2} = -\frac{1}{4} \left(\frac{\partial^2 S_1}{\partial E_1^2} + \frac{\partial^2 S_2}{\partial E_2^2} \right). \quad (71)$$

稍加整理, 即得

$$K_A = T^2 C_v. \quad (72)$$

$C_v = \partial E / \partial T$ 是物体的热容率.

现在我们直接从分子运动的观点来分析 \hat{A} 值的分布. 我们先把 \hat{A} 写成能量密度的积分:

$$\hat{A}(s(t)) = \int d^3r f(\mathbf{r}) \epsilon(\mathbf{r}, t) \quad (73)$$

$$\begin{aligned} f(\mathbf{r}) &= 1, \quad \text{如果 } \mathbf{r} \text{ 在第 1 部分} \\ &= -1, \quad \text{如果 } \mathbf{r} \text{ 在第 2 部分} \end{aligned} \quad (74)$$

$\epsilon(\mathbf{r}, t)$ 是能量密度. 例如, 这物体若是一群粒子, 则

$$\begin{aligned} s &\equiv (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, \mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N), \\ \epsilon(\mathbf{r}, t) &= \sum_i \delta(\mathbf{r} - \mathbf{r}_i(t)) \frac{p_i^2(t)}{2m}, \\ &+ \frac{1}{2} \sum_{i,j} U(\mathbf{r}_i(t) - \mathbf{r}_j(t)) \delta\left(\mathbf{r} - \frac{\mathbf{r}_i(t) + \mathbf{r}_j(t)}{2}\right). \end{aligned} \quad (75)$$

$p_i^2/2m$ 为粒子之动能. $U(\mathbf{r}_i - \mathbf{r}_j)$ 为 i, j 间之相互作用能. $\mathbf{r}_i(t), \mathbf{p}_i(t)$ 为在时刻 t 的位置及动量. 因此 (73) 事实上是一大群变量之组合. 组合中每一项都是动量和位置的函数, 每个变量都是随时刻 t 改变. 现在来求 $\langle \hat{A}^2 \rangle_c$. 由 (73) 得

$$\langle \hat{A}^2 \rangle_c = \int d^3r d^3r' f(\mathbf{r}) f(\mathbf{r}') \langle \epsilon(\mathbf{r}) \epsilon(\mathbf{r}') \rangle_c. \quad (75')$$

此地的平均值是观测时间内的平均值:

$$\begin{aligned}\langle \hat{A} \rangle &= \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} dt \hat{A}(s(t)), \\ \langle \hat{A}^2 \rangle_c &= \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} dt \hat{A}^2(s(t)) - \langle \hat{A} \rangle^2.\end{aligned}\quad (76)$$

现在假设相离很远的各点的能量密度不相干, 即

$$\langle \epsilon(\mathbf{r})\epsilon(\mathbf{r}') \rangle = 0, \quad \text{如果 } |\mathbf{r} - \mathbf{r}'| \gg \xi. \quad (77)$$

ξ 为相干距离, 为一微观尺度. 再假设物体为均匀, 则

$$\langle \epsilon(\mathbf{r})\epsilon(\mathbf{r}') \rangle_c = \langle \epsilon(\mathbf{r} - \mathbf{r}')\epsilon(0) \rangle_c. \quad (78)$$

原点 0 可取为物体中央. 由这些假设, 得

$$\langle \hat{A}^2 \rangle_c = V \left[\int d^3r'' \langle \epsilon(\mathbf{r}'')\epsilon(0) \rangle_c + O\left(\frac{\xi}{L}\right) \right], \quad (79)$$

$V = L^3$ 是物体的总容积. 最后一项是物体表面, 及 1, 2 部份交界面而来的误差. 因为 $\xi \ll L$, 所以 (73) 是一群大致独立的变量的组合.

现在引用中心极限定理, 得 \hat{A} 值之分布

$$\begin{aligned}\rho(\hat{A}) &= \frac{1}{\sqrt{2\pi N\sigma^2}} \exp\left\{-\frac{(A - \langle A \rangle)^2}{2N\sigma^2}\right\}, \\ \langle \hat{A} \rangle &= 0, \\ \sigma^2 &= \frac{V}{N} \int d^3r \langle \epsilon(\mathbf{r})\epsilon(0) \rangle_c.\end{aligned}\quad (80)$$

这个分布, 是从分子运动观点, 和独立的假设 $\xi \ll L$, 得来. 能量密度的相干函数是可以由实验测定的, 此类测定方法在上章末节说明过.

9. 起伏和基本假设的延伸

现在我们可以用以上的结果, 来对基本假设作更深一层的分析. 这假设是为求熵而设, 把熵和活动范围的体积连起来. 活动范围包括各不变量容许的状态. 有了熵, 热力学的各量就可以计算了. 在第六、七、八章, 我们是举了不少例子, 推出了一些计算法则, 来说明这基本假设应用.

简单、方便是这假设的特点. 有了它, 我们不需要详细的分子运动数据, 就可以计算热力性质.

我们三番五次地强调: 活动范围必须要由运动来决定, 它代表各变量起伏的幅度. 但基本假设并没有指出如何计算起伏的大小. 第五到七章的分析只指出宏观变量的起伏很小, 并没有明确的定量结论. 不错, 分布 (70) 是从活动范围中的状态统计得来. 它指出几乎整个活动范围内的 $\hat{A}(s)$ 值都是 \bar{A} , 百分误差为 $O(1/\sqrt{N})$. 不过正态分布函数 (70) 本身和分子运动的关系并不清楚 (见第六章末节).

在上章末节, 我们指出, 相干函数可以用散射实验测量. 上一节我们讨论了 \hat{A} 的起伏和相干函数的关系. 结论是, 起伏是可以由散射测定. 实验测定的各式平均值, 都是对时间的平均值. 变量值的分布, 是指时间的分配. (80) 这个分布的意义如下: 令

$$R = \{s; A_1 < A(s) < A_2\}, \quad (81)$$

即, 所有状态 s , 其 $\hat{A}(s)$ 值在 A_1, A_2 之间者, 都在 R 内. 令 \mathcal{T}_R 为轨迹花在 R 内的时间, 则 (80) 的意义是

$$\int_{A_1}^{A_2} dA \rho(A) = \frac{\mathcal{T}_R}{\mathcal{T}}. \quad (82)$$

\mathcal{T} 是总观测时间. 这可以看成是一个纯粹由轨迹的时间分配而定的量. 它并未涉及熵和基本假设.

(70) 这个分布是由活动范围中的区域大小而来. 令 $\Gamma(R)$ 为 R 的容积, 则

$$\int_{A_1}^{A_2} dA \frac{\Gamma(E, A)}{\Gamma(E)} = \frac{\Gamma(R)}{\Gamma(E)}. \quad (83)$$

这纯粹是区域容积的计算, 并未涉及轨迹和时间的概念.

(80) 的 $\rho(A)$ 和 (70) $\Gamma(E, A)/\Gamma(E)$ 显然意义不同, 但形式相同. 同为正态分布. 它们是不是相等呢? 相等的条件是 $K_A = N\sigma^2$, “见 (72), (80)”, 即

$$T^2 C_v = V \int d^3r \langle \epsilon(\mathbf{r}) \epsilon(0) \rangle_c. \quad (84)$$

这条件是否成立, 可以用热力实验测左边, 用散射实验测右边, 看等不等. 或者用基本假设计算左边, 用其他方法计算右边, 看看结果是否一样. 且不多谈这些实验和计算的细节. 它们的结果都是两边相等. 这些结果指出

$$\frac{\mathcal{T}_R}{\mathcal{T}} = \frac{\Gamma(R)}{\Gamma(E)}. \quad (85)$$

也就是说轨迹花在 R 内的时间, 是和 R 的大小成正比. 注意: R 是由 A 的分类而定, A 是个大值变数. 因此 R 是个大区域. 我们曾提过, R 必须是大区域, \mathcal{T}_R 才有意义. (见 (10.34)~(10.35) 讨论).

虽然以上的讨论是用一个特别的例子, 但一般的结论可想而知. R 可以是任何大区域, 式 (85) 都成立. 这结论可以讲成: “轨迹均匀地分布在活动范围内” 或 “活动范围内任一状态都有相同的出现概率”. 随便怎么讲都可以, 只是必须记住式 (85) 的 R 必须是大区域. R 若太小, 则 \mathcal{T}_R 无意义.

式 (85) 可以看作基本假设的延伸. 它大致告诉我们轨迹是如何分散盘桓在活动范围中. 现在, 我们不但可以用活动范围来计熵, 及其他热力学坐标 \hat{A} 之平均值, 我们还可以用它来计算 \hat{A} 的起伏. \hat{A} 之平均值为 $O(N)$. 而其起伏为 $O(\sqrt{N})$. 从 $O(N)$ 到 $O(\sqrt{N})$, 确是非常重要的步骤.

我们再次强调; 此地所讲 \hat{A} 的起伏, 是指 \hat{A} 随时间的改变, 不是对 \hat{A} 测量的不精确度. 一般说来这测不准度是从物体表面来的. 体外, 体内的界限不清, 造成 $O(N^{2/3})$ 的观测误差. 物体表面附近的分子数是 $O(N^{2/3})$. 这误差比起伏 $O(N^{1/2})$ 要大很多. 两者绝不可以混为一谈. \hat{A} 的平均值的测定和它的起伏的测定, 是用完全不同的方法. 前者的 $O(N^{2/3})$ 误差并不表示后者 $O(N^{1/2})$ 不能测定. 前者是观测时间内的平均值, 是 “不变部分” 而后者是起伏, 是 “改变” 的部分, 或频率不等于零的部分.

读者大概注意到, 以上用作例子的是 $\hat{A} = H_1 - H_2$ (见 (69)). 为什么不用 $H_1 + H_2 = H$ 为例? 原因是 H 为一守恒量. 它的起伏, 是由物体和外界的界面的作用决定. 如果界面是绝热的, H 的起伏就很小. 如果是开放的, 则起伏较大. 我们用 $H_1 - H_2$, 可以避开这问题. $H_1 - H_2$ 不是个守恒量, 能量密度的起伏, 用 $H_1 - H_2$ 来分析比较方便. 用散射实验分析的是波分量

$$\epsilon_{\mathbf{k}} = \int d^3r e^{-i\mathbf{k}\cdot\mathbf{r}} \epsilon(\mathbf{r}), \quad (86)$$

$H_1 - H_2$ 可以算是 $k \rightarrow 0$ 而非零的极限, 即 $k \sim 1/L$, L 为物体的大小. 如果 $k = 0$, 则 $\epsilon_0 = H$.

请注意, 起伏的现象, 告诉我们分子, 原子之存在. 也就是说, 物质不是连续的, 而是由分立的个体组成. 如果是连续的, 则 N 必是无穷大, 起伏 (百分比) $O(N^{1/2})/N = O(N^{-1/2})$ 将是零. N 为有限, 才有起伏. 因为我们用了热常数 $= k = 1$ 的单位, 所以容易忽略这热常数在各公式中的地位. kT 是一个微观量. 它大致是每一个分子的能量. 如果物质是连续的, k 必是零. k 之非零,

表示分立个体之存在. 热力学的式子, 是没有 k 的. 但是涉及起伏而又有温度出现的式子, 一定会有 k 出现. 例如 (62) 和 (84)

$$m\alpha = \frac{1}{2kT} \int dt' C(t'), \quad (87)$$

$$\frac{TC_v}{V} = \frac{1}{kT} \int d^3r \langle \epsilon(\mathbf{r})\epsilon(0) \rangle_c. \quad (88)$$

式中的 α , m 和相干函数都是和温度无关的概念. TC_v/V 是一个宏观量. 剩下的 T 一定要有 k 相陪, 否则单位就不对了, 相干函数当然是起伏的特性.

讨论问题十二

1. 求二项分布

$$\rho(n) = \binom{N}{n} p^n (1-p)^{N-n}$$

的连接特性函数

$$\ln \langle e^{-ikn} \rangle \quad (89)$$

答:

$$N \ln [1 - p(1 - e^{-ik})].$$

2. 在上题中, N , n 都很大的情形下, $\rho(n)$ 和正态分布有何不同?
 3. 求分布 (18) 的特性函数, 再推出 (20).
 4. 连接值展开定理, 对统计力学而言是个极其重要的定理. 令

$$H = H_0 + U$$

(A) 试证

$$F = F_0 - T \ln \langle e^{-U/T} \rangle_0,$$

式中 $\langle \rangle_0$ 的定义是对 $e^{-H_0/T}$ 的平均:

$$\langle A \rangle_0 = \sum A e^{-H_0/T} / Z_0, \quad (90)$$

F_0 是由 H_0 求得的热位能.

(B) 由连接值展开定理, 试证

$$F = F_0 - T \langle e^{-U/T} - 1 \rangle_{0c}. \quad (91)$$

即 U 对 F_0 的修正, 是 U/T 的各次连接值.

(C) 令 H_0 为某气体动能, U 为相互作用能, 并假设短程相互作用. 试证 F 对 U 的各展开项都和容积成正比. 这结果若非连接值展开定理, 就极难证明.

提示: 用第九章第 1 节的模型, 把 U 写成

$$U = \frac{1}{2} \int d^3r d^3r' \rho(\mathbf{r}) \rho(\mathbf{r}') u(\mathbf{r} - \mathbf{r}') \quad (92)$$

注意

$$\langle \rho(\mathbf{r}) \rho(\mathbf{r}') \rangle_0 = Z \delta(\mathbf{r} - \mathbf{r}') \quad (93)$$

理想气体的密度在不同地方是不相干的. 先试求 $\langle U \rangle_{0c}$, $\langle U^2 \rangle_{0c}$, 高次项的情形就大致知道了.

5. 有效能量

令物体的状态为 $s = (s_1, s_2)$, s_1 和 s_2 各为物体两部分的状态. 令总能量为

$$H(s) = H_1(s_1) + H_2(s_2) + H_{12}(s_1, s_2). \quad (94)$$

如果某变量 $A(s_1)$ 只和 s_1 有关, 试证其平均值可以用“有效能量” $H'_1(s_1)$ 计算, H' 的定义是

$$\begin{aligned} H'_1(s_1) &= H_1(s_1) - T \ln \langle e^{-H_{12}/T} \rangle_2 \\ \langle e^{-H_{12}/T} \rangle_2 &\equiv \frac{\sum_{s_2} e^{-H_{12}/T} e^{-H_2/T}}{\sum_{s_2} e^{-H_2/T}}. \end{aligned} \quad (95)$$

也就是说, H'_1 是 H_1 加上一修正项. 修正的来源是和第 2 部分的作用. 如果 H_{12} 和 H_1 的大小相若, 则此修正项不能忽略.

以上求有效能量的步骤, 又是一分类的例子, 也是一投影的例子. (复习第五章末节, 及第六章第 1 节).

6. 注意, 在上题, s_1 在状态空间 ($s = (s_1, s_2)$ 之集) 中, 代表一区域 $R(s_1)$ 即 $\{s_1$ 固定, 所有 s_2 之集}. 如果第 2 部份有 N_2 粒子, N_2 为大数, 则此区域为一大区域, 容积为 e^{N_2} 级大数:

$$\Gamma(R(s_1)) = O(e^{N_2}). \quad (96)$$

令

$$P(s_1) = \Gamma(R(s_1)) / \Gamma. \quad (97)$$

Γ 为整个物体的活动范围体积. 试讨论 $P(s_1)$ 的意义. 它是否可以代表 s_1 出现的概率? 令 N_1 为第 1 部分的粒子数, 如 N_1 不是大数, $P(s_1)$ 的数量级为何? 如 N_1 是大数, 又如何?

注意: $\Gamma = O(e^{N_1+N_2})$.

第十三章 相干函数

以上几章, 讨论独立和相干的概念在统计力学中的重要性. 本章的重点在相干函数的示范例子. 因此, 数学运算稍多一点. 从这些例子, 可以对相干的概念有更结实一点的认识. 首先我们把上一章讲的响应和起伏的关系再简单地归纳一下, 并定义响应函数. 然后是密度相干函数及响应函数在气体模型中的计算例子, 包括稀疏气体, 稠密气体 (一度空间模型), 及不合群量子气体. 量子力学带来的一些特别性质, 也略作分析.

1. 响应和起伏

借中心极限定理, 上一章的分析把变量 $\hat{A}(s)$ 在活动范围的分布, 及其起伏的相干值连了起来. 热容率和能量密度起伏的关系是一个重要的示范例子. 热容率是一种“响应效率”, 它度量物体对温度改变的一种反应, 现在我们用比较抽象一点的步骤, 导出一般响应和起伏间的关系. 以下的结果更进一步地指出, 相干值其实也可以看成一大值变量, 它在活动范围中几乎每个地方的值都一样.

假想一物体的总能函数为 H , 现在加一外力 λ , 使总能量变成

$$H - \lambda \hat{A}. \quad (1)$$

例如 λ 是磁场, \hat{A} 是总磁矩, 或 λ 为拉力, \hat{A} 为物长. 由于这外力, 引起物体内分子的反应, 而作若干改变. 这些反应, 可以由各变量的改变测定. 例如, 某变数 \hat{B} 之值可由平均值算出

$$B = \langle \hat{B} \rangle = \frac{\sum_s \hat{B} e^{-(H-\lambda\hat{A})/T}}{\sum_s e^{-(H-\lambda\hat{A})/T}}. \quad (2)$$

B 自然是 λ 的函数, 由上式微分得

$$\begin{aligned}\chi &\equiv \frac{\partial B}{\partial \lambda} = \frac{1}{T} [\langle \hat{B} \hat{A} \rangle - \langle \hat{B} \rangle \langle \hat{A} \rangle] \\ &\equiv \frac{1}{T} \langle \hat{B} \hat{A} \rangle_c.\end{aligned}\quad (3)$$

这是所谓“微分响应率”。

【例】令 $\hat{A} = \hat{B} = H$, 即总能量本身。

则

$$(H - \lambda A)/T \approx H/((1 + \lambda)T), \quad (4)$$

λ 即是温度的增加百分比 (假定 $\lambda \ll 1$)。由 (3) 得

$$T \frac{\partial E}{\partial T} = \frac{1}{T} \langle H^2 \rangle_c,$$

即

$$\begin{aligned}T^2 C_v &= \int d^3 r d^3 r' \langle \epsilon(\mathbf{r}) \epsilon(\mathbf{r}') \rangle_c \\ &= V \int d^3 r'' \langle \epsilon(\mathbf{r}'') \epsilon(0) \rangle_c.\end{aligned}\quad (5)$$

这就是 (12.84)。 $\epsilon(\mathbf{r})$ 为能量密度。

【例】令 λ 为磁场 h , $\hat{B} = \hat{A}$ 为总磁矩 \hat{M} , 则由 (3) 得

$$\begin{aligned}\chi &= \frac{\partial M}{\partial h} = \frac{1}{T} \langle \hat{M}^2 \rangle_c \\ &= \frac{V}{T} \int d^3 r'' \langle m(\mathbf{r}'') m(0) \rangle_c.\end{aligned}\quad (6)$$

$m(\mathbf{r})$ 为磁矩密度, χ 是磁感应率, 或磁化效率。

由这些结果可见, 相干函数可以看成一种大值变量, 因为它可以由 B 的微分而得, 而 B 是大值变数 \hat{B} 之值。

现在再倒退一步, 问: 外力 λ 对热位能的影响是什么? 热位能是

$$\begin{aligned}F &= -T \ln Z, \\ Z &= \sum_s e^{-(H - \lambda \hat{A})/T}.\end{aligned}\quad (7)$$

令 Z_0 为 Z 在 $\lambda = 0$ 时之值, 则

$$Z = Z_0 \left\langle e^{\lambda \hat{A}/T} \right\rangle_0. \quad (8)$$

平均值 $\langle \dots \rangle$ 是指 $\lambda = 0$ 时的平均值:

$$\langle \hat{B} \rangle_0 \equiv \sum_s \hat{B} e^{-H/T} / \sum_s e^{-H/T}. \quad (9)$$

现在用上一章的连接值展开定理, (见 12.32) 至 (12.38)), 得

$$\ln \langle e^{\lambda \hat{A}/T} \rangle_0 = \frac{\lambda}{T} \langle \hat{A} \rangle_0 + \frac{\lambda^2}{2T^2} \langle \hat{A}^2 \rangle_{0c} + \dots \quad (10)$$

将 (10) 代入 $\ln Z$, 得

$$\begin{aligned} F &= -T \ln Z_0 - T \ln \langle e^{\lambda \hat{A}/T} \rangle_0 \\ &= F_0 - \lambda \langle \hat{A} \rangle_0 - \frac{\lambda^2}{2T} \langle \hat{A}^2 \rangle_{0c} + \dots \end{aligned} \quad (11)$$

因此,

$$\begin{aligned} -\frac{\partial F}{\partial \lambda} &= \langle \hat{A} \rangle \\ -\frac{\partial^2 F}{\partial \lambda^2} &= \frac{\partial}{\partial \lambda} \langle \hat{A} \rangle = \frac{1}{T} \langle \hat{A}^2 \rangle_c. \end{aligned} \quad (12)$$

等等. (12) 式中未在 $\langle \dots \rangle$ 下加注 0, 因为不必, 以上展开可以用任何 λ 值为原点.

因此, 各种平均值, 相干值都可以和热位能的微分扯上关系. 从这观点看来, 既然基本假设指定了计算熵及热位能的法则, 它也就指定了它们的微分的计算法则. 因此, 基本假设也就指定了所有平均值和相干值的计算法则.

以上的讨论, 没有考虑到量子力学. 因此, 有些结论可能须要修改. 如果 $[\hat{A}, H] \neq 0$, 则 (8) 需修正, 因为

$$e^{-(H-\lambda\hat{A})/T} \neq e^{-H/T} e^{\lambda\hat{A}/T},$$

这在本章末再讨论. 现在暂不考虑量子力学.

【例】 响应函数和相干函数. 这些只是 (3) 的例子而已. 令

$$\begin{aligned} \hat{B} &= \int d^3r a(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} \equiv a_{\mathbf{k}} \\ \hat{A} &= \hat{B}^* \end{aligned} \quad (13)$$

代入 (3) 得

$$\begin{aligned} VG_{\mathbf{k}} &\equiv \frac{1}{T} \int d^3r d^3r' \langle a(\mathbf{r}) a(\mathbf{r}') \rangle_c e^{-i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} \\ &= \frac{V}{T} \int d^3r'' \langle a(\mathbf{r}'') a(0) \rangle_c e^{-i\mathbf{k}\cdot\mathbf{r}''} \\ &\equiv \frac{V}{T} C_{\mathbf{k}}. \end{aligned} \quad (14)$$

C_k 是能量起伏的相干函数的波分量. G_k 叫做“响应函数”的波分量, 它的傅里叶积分, 即“合波量”为

$$G(\mathbf{r} - \mathbf{r}') \equiv \frac{1}{(2\pi)^3} \int d^3k G(\mathbf{k}) e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}. \quad (15)$$

它的意义是: 如果 $\lambda(\mathbf{r})$ 代表一个外力分布, 则

$$\langle a(\mathbf{r}) \rangle = \int d^3r' G(\mathbf{r} - \mathbf{r}') \lambda(\mathbf{r}'). \quad (16)$$

以上假定 λ 很小, 且 $\langle a \rangle$ 在 $\lambda = 0$ 时为零. (这结果可由 $\langle a_k \rangle = \lambda_k G_k$ 得来). 响应函数指出因果的关系. $\lambda(\mathbf{r}')$ 是因, $\langle a(\mathbf{r}) \rangle$ 是果. $G(\mathbf{r} - \mathbf{r}')$ 指出由在 \mathbf{r}' 的单位前因在 \mathbf{r} 所产生的后果, 它也可称为“传播函数”.

【例】以上的结果自然可用在开放分布. 令 $\hat{A} = \hat{B} = \hat{N}$, 即分子数, 则 λ 为化学势 μ 之增加. (3) 的结果即成了

$$\begin{aligned} \frac{\partial N}{\partial \mu} &= \frac{1}{T} \langle \hat{N}^2 \rangle_c \\ &= \frac{V}{T} \int d^3r \langle \rho(\mathbf{r}) \rho(0) \rangle_c. \end{aligned} \quad (17)$$

此地 $\rho(\mathbf{r})$ 是分子密度. 注意

$$\frac{\partial N}{\partial \mu} = V \frac{\partial n}{\partial \mu} = V \frac{\partial n / \partial p}{\partial \mu / \partial p} = V n \frac{\partial n}{\partial p}, \quad (18)$$

以上用到 $n = \partial p / \partial \mu$, $n \equiv N/V = \langle \rho(\mathbf{r}) \rangle$. 因此

$$n \frac{\partial n}{\partial p} = \frac{1}{T} \int d^3r \langle \rho(\mathbf{r}) \rho(0) \rangle_c. \quad (19)$$

所以, 密度对应力的响应是和密度起伏有关.

2. 密度相干函数

现在来看密度相干函数的求法. 令

$$n(\mathbf{r}|0) \equiv \frac{1}{n} \langle \rho(\mathbf{r}) \rho(0) \rangle. \quad (20)$$

这是在 $\mathbf{r} = 0$ 放了一个分子时的密度分布, 注意

$$\begin{aligned} \langle \rho(\mathbf{r}) \rho(0) \rangle &= (\text{在 } \mathbf{r} = 0 \text{ 有分子之概率}) \times (\text{0 有分子条件下, 在 } \mathbf{r} \text{ 有分子之概率}) \\ &\equiv n \times n(\mathbf{r}|0). \end{aligned} \quad (21)$$

因此, (20) 可以看作一“条件概率”.

如果物体是一理想气体, 则

$$n(\mathbf{r}|0) = \delta(\mathbf{r}) + n. \quad (22)$$

右第一项是放在 0 的分子的密度, 第二项 n 是其他分子的平均分布.

我们可以把 (22) 应用到 (19) 来求压力和密度的关系. 由 (22), (20), 得

$$\langle \rho(\mathbf{r})\rho(0) \rangle_c = n\delta(\mathbf{r}). \quad (23)$$

代入 (19), 得

$$\frac{dn}{dp} = \frac{1}{T}. \quad (24)$$

这当然就是理想气体定律 $p = nT$.

如果分子间有相互作用能 $u(\mathbf{r}_i - \mathbf{r}_j)$, ($\mathbf{r}_i, \mathbf{r}_j$ 为分子 i, j 之位置), 则在 $\mathbf{r} = 0$ 放一分子后, 分子的分布不再是 (22), 而是

$$n(\mathbf{r}|0) = ne^{-u(\mathbf{r})/T} + \delta(\mathbf{r}). \quad (25)$$

此式在 r 很大时趋近于 n . (23) 现在成了

$$\langle \rho(\mathbf{r})\rho(0) \rangle_c = n [\delta(\mathbf{r}) + n(e^{-u(\mathbf{r})/T} - 1)]. \quad (26)$$

这式子只有在气体密度低时是正确的, 因为 (25) 是在位能 $u(\mathbf{r})$ 内理想气体的分布. 如果密度高, 则必须考虑这些分子的作用. 所以 (26) 的最后一项是低密度下的第一修正项.

把 (26) 代入 (19), 得

$$T \frac{\partial n}{\partial p} = 1 + n \int d^3r (e^{-u(\mathbf{r})/T} - 1), \quad (27)$$

积分后即得

$$p = nT + n^2T \int d^3r (1 - e^{-u(\mathbf{r})/T}). \quad (28)$$

这是理想气体公式加上一修正项, 这修正项将在下一章详细讨论. 此地只说明 (19) 的一个应用而已.

相干距离 ξ 的定义是: 相干函数在距离大于 ξ 时渐趋于零. 由 (23) 可见, 理想气体的相干距离是零. 由 (26) 可见, 稀疏的气体, 其分子作用 $u(r)$ 的有

效距离就是 ξ . 严格说来, 不同的相干函数可能会有不同的相干距离. 例如密度相干函数的相干距离就可能和自转相干函数的不一样, 但一般说来, 各相干距离大小都差不多.

以上例子是稀疏的气体, 下面看一个一度空间例子, 密度可高可低. 这就是第八章第 3 节的例子. 现在来求 $\langle \rho(\mathbf{x})\rho(0) \rangle$. 把一粒子固定在原点, 令右边的第一粒子的位置为 y_1 , 第二粒子位置为 $y_1 + y_2$, 第 n 粒子位置为 $y_1 + y_2 + \dots + y_n$. 因此,

$$n(x|0) = \delta(x) + \langle \delta(x - y_1) \rangle + \langle \delta(x - y_1 - y_2) \rangle \\ + \dots + \langle \delta(x - y_1 - y_2 - \dots - y_n) \rangle + \dots \quad (29)$$

$n(x|0)$ 的定义是 (20). 各个 y_i 是相邻分子间的距离, (29) 专用于 $x \geq 0$. 如果 $x \leq 0$, 相似的式子可照样写下. 从 (29) 得

$$C_k = \int dx e^{-ikx} \langle \rho(x)\rho(0) \rangle \\ = n \int dx e^{-ikx} n(x|0) \\ = n \left[1 + \sum_{m=1}^{N/2} (\phi_k^m + \phi_{-k}^m) \right], \quad (30)$$

$$\phi_k \equiv \langle e^{-iky} \rangle = \frac{1}{\zeta} \int_0^\infty dy e^{-(py+u(y))/T} e^{-iky}, \\ \zeta \equiv \int_0^\infty dy e^{-(py+u(y))/T}. \quad (31)$$

p 是压力. 这 ζ 已在第八章第 3 节讨论过 ((8.35)). $u(y)$ 是相邻分子间的作用能, 以上假设原点左、右两边都有 $N/2$ 个分子. (30) 中的 ϕ_{-k}^m 是由左边的分子供给, 以上并假定 $k \neq 0$.

(30) 中的和可以求出, 把 N 看成 ∞ , 得

$$C_k = n \left(\frac{1}{1 - \phi_k} + \frac{1}{1 - \phi_{-k}} - 1 \right) \\ = n \frac{1 - |\phi_k|^2}{|1 - \phi_k|^2}. \quad (32)$$

现在看一个特殊的作用能 $u(y)$, 即硬球作用

$$u(y) = \infty, \quad y < b, \\ = 0, \quad y > b. \quad (33)$$

b 是硬球的直径.

将 (33) 代入 (31), 得

$$\begin{aligned}\phi_k &= \frac{p/T}{(p/T + ik)} e^{-ika} \\ &= \frac{1}{1 + ik(a-b)} e^{-ika}, \\ a &\equiv 1/n.\end{aligned}\tag{34}$$

以上用到 $p(a-b) = T$. 这气体一如理想气体, 只是有效体积是 $L - Nb = N(a-b)$. 从 (34), (32) 得

$$C_k = \frac{nk^2(a-b)^2}{|1 - e^{-ikb} + ik(a-b)|^2}.\tag{35}$$

注意, 这结果和温度无关, 因为 $n(y)$ 并无能量尺度标准. 也就是说, 一群硬球在乱动, 无论温度是多少, 密度相干函数不会受影响. 如果 $a \gg b$, 即密度低时, (35) 约是

$$C_k \approx n \left(1 - \frac{2b}{a}\right).\tag{36}$$

此式合用于 $kb \ll 1$. 如果密度很高, 即 $a-b$ 很小, 则在 k 不很大时, (35) 的分母在 $k \approx 2\pi m/b$ 时很小, m 为整数

$$\begin{aligned}1 - e^{-ikb} + ik(a-b) &\approx \frac{1}{2} (k'b)^2 + i \left[2\pi m \left(\frac{a}{b} - 1\right) + k'a\right], \\ k &\equiv 2\pi m/b + k',\end{aligned}\tag{37}$$

以上 $k'a$, $1 - a/b$, 均视为小量. 因此, (35) 的分母在 $k = k_m$ 时为零

$$\begin{aligned}k_m &\approx \frac{2\pi m}{a} - i4\pi^2 m^2 (a-b)^2 / a^3 \\ m &= \pm 1, \pm 2, \pm 3, \dots\end{aligned}\tag{38}$$

所以 C_k 是一群尖峰, 坐落在 $2\pi m/a$, 峰的宽度是 $4\pi^2 m^2 a^{-3} (a-b)^2$. (见图 1), 如果 m 很大, 则 (38) 不合用. 这些尖峰的意义是:

$$C(x) \equiv \langle \rho(x)\rho(0) \rangle_c = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk C_k e^{ikx}\tag{39}$$

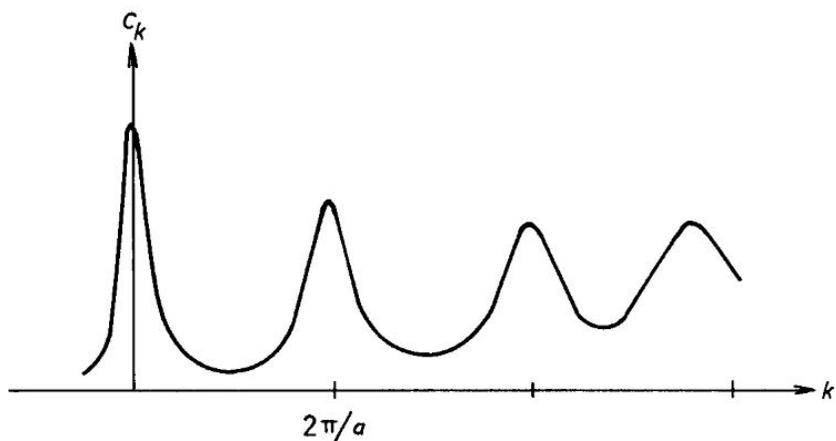


图 1

约是以 a 为周期的函数, 但逐渐随 $|x|$ 减小. (38) 的虚数部份为减小率. 因此, 相干距离约为

$$\xi = \frac{a^3}{4\pi^2(a-b)^2},$$

因为

$$C(x) \sim e^{-|x|/\xi}. \quad (40)$$

这个例子的目的是指出分子间类似硬球的强斥力, 会造成一段大致周期性的密度分布. 在一般液体中, 这现象早已为实验者所观测到. (见图'2).

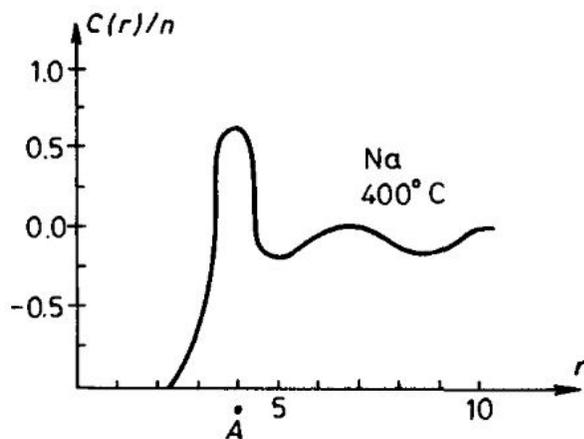


图 2 液态钠在 400°C 的密度相干函数

3. 费米粒子气体

现在来看量子气体中的密度相干情形, 先看费米粒子.

要计算密度相干函数, 最简单的方法是从它的定义着手: 它是一个散射率:

$$C_{\mathbf{k}} = \int d^3r e^{-i\mathbf{k}\cdot\mathbf{r}} \langle \rho(\mathbf{r}) \rho(0) \rangle_c \quad (41)$$

是和散射率成正比. 散射过程是: 一外来粒子射入, 把动量 \mathbf{k} 传给物体, 再走掉. 因此, 把一个粒子的动量由 \mathbf{p} 变成了 $\mathbf{p} + \mathbf{k}$. 我们知道这散射率必和 $f_{\mathbf{p}}(1 - f_{\mathbf{p}+\mathbf{k}})$ 成正比 (见第三章):

$$C_{\mathbf{k}} \propto \frac{2}{V} \sum_{\mathbf{p}} f_{\mathbf{p}}(1 - f_{\mathbf{p}+\mathbf{k}}), \quad (42)$$

右边的 2 是由自转 1/2 的两态而来. $f_{\mathbf{p}}$ 是 \mathbf{p} 位子的平均粒子数

$$f_{\mathbf{p}} = \frac{1}{e^{(\epsilon-\mu)/T} + 1}, \quad (43)$$

$1 - f_{\mathbf{p}+\mathbf{k}}$ 是 $\mathbf{p} + \mathbf{k}$ 位子空着的概率. (42) 的比例常数为 1, 因为当 f 很小时, (高温低密度时)(42) 就成了已知的理想气体结果 (23).

从 (42) 可以求出

$$\begin{aligned} C(\mathbf{r}) &= \langle \rho(\mathbf{r}) \rho(0) \rangle_c \\ &= \int \frac{d^3k}{(2\pi)^3} C_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} \\ &= n\delta(\mathbf{r}) - 2|f(\mathbf{r})|^2, \end{aligned} \quad (44)$$

$$f(\mathbf{r}) \equiv \frac{1}{V} \sum_{\mathbf{p}} f_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{r}} = \frac{I(\mathbf{r})}{2\pi^2 r} \quad (45)$$

$$I(\mathbf{r}) \equiv \int_0^\infty dp \frac{p \sin pr}{e^{(\epsilon-\mu)/T} + 1}.$$

这 $I(r)$ 积分已被人积出*:

$$\begin{aligned} I(\mathbf{r}) &= \frac{\partial}{\partial r} \left[\frac{1/2\xi}{\sin h(r/2\xi)} \sin(p_F r) \right], \\ 2\xi &\equiv \frac{p_F}{mT}, \\ \mu &= \frac{p_F^2}{2m}. \end{aligned} \quad (46)$$

*见 Landau and Lifshitz (1980) p.358

如果, $r \ll \xi$, 或 $T \rightarrow 0$, 则

$$\begin{aligned} I(\mathbf{r}) &\approx \frac{\partial}{\partial r} \left[\frac{1}{r} \sin p_F r \right], \\ C(\mathbf{r}) &\approx n\delta(\mathbf{r}) - \frac{p_F^4}{2\pi^4 r^2} \left[\frac{\cos p_F r}{p_F r} - \frac{\sin p_F r}{(p_F r)^2} \right]^2. \end{aligned} \quad (47)$$

这结果的特点是它带有周期性的改变, 周期为 $2\pi/2p_F$. 这是粒子分布在顶层突然切断所致, 也可以说是费米粒子的独占性使密度围着原点形成一层层的结构.

在 $T \rightarrow 0$, 没有相干距离的存在, 也就是说 $C(r)$ 不是以 $e^{-r/\xi}$ 的形式减弱. 不过 $C(r) \propto 1/r^4$, r 很大时仍然消失得相当快.

如果 $T > 0$, 就有了一个相干距离 ξ , (见 (46)) 在 r 很大

$$\begin{aligned} C(\mathbf{r}) &\approx \frac{p_F^2 e^{-r/\xi}}{2\pi^4 \xi^2 r^2} \sin^2 p_F r, \\ \xi &= \frac{1}{p_F} \left(\frac{\epsilon_F}{T} \right). \end{aligned} \quad (48)$$

$\epsilon_F = p_F^2/2m$. 在低温时, ξ 比粒子间距离大很多. 在导出 (48) 的过程中我们假设了 $\xi \gg 1/p_F$.

现在来看响应函数 $G(\mathbf{r})$, 它和 $C(\mathbf{r})$ 的关系. 由 (14) 得 $G = C/T$, 只在经典力学中成立, 在量子力学中不成立, 这一点稍后再讲, 现在直接看一下 G .

$G(\mathbf{r})$ 的意义是: 若加上一个位能 $-\lambda\delta(\mathbf{r})$, 则密度的改变为 $\lambda G(\mathbf{r}) + O(\lambda^2)$. 现在我们就加这样一个位能, 再计算密度. 这位能的结果是把粒子的波函数改了. 令气体的容积为一圆球, 原点在球心. 只有动量 $l = 0$ 的波函数被修改, 因 $\delta(\mathbf{r})$ 只影响了原点, 这波函数是

$$\begin{aligned} \varphi_0(\mathbf{r}) &= \frac{1}{\sqrt{V}} \frac{\sin(pr + \delta_0)}{pr} \\ &\approx \frac{1}{\sqrt{V}} \left[\frac{\sin pr}{pr} + \frac{\delta_0 \cos pr}{pr} \right], \\ \delta_0 &= -\frac{\lambda mp}{2\pi}. \end{aligned} \quad (49)$$

δ_0 读者自己导出.

密度的改变是

$$\lambda G(\mathbf{r}) = 2 \sum_{\mathbf{p}} f_{\mathbf{p}} \left[|\varphi_0(\mathbf{r})|^2 - |\varphi_0(\mathbf{r})|_{\lambda=0}^2 \right]. \quad (50)$$

右边的 2 是由两自转态而来. 因此, 由 (50), (49) 稍加整理得

$$G(\mathbf{r}) = \frac{-m}{2\pi^2 r^2} I(2r), \quad (51)$$

$I(2r)$ 正好是 (45), 但需将 r 改成 $2r$. 在 $r \ll \xi$, 或 $T \rightarrow 0$ 时, 由 (46) 得

$$G(\mathbf{r}) \approx -\frac{2mp_F^4}{\pi^2} \left[\frac{\cos 2p_F r}{(2p_F r)^3} - \frac{\sin 2p_F r}{(2p_F r)^4} \right]. \quad (52)$$

在 $T > 0$ 时, ξ 就很重要, $r \geq \xi$ 时,

$$G(\mathbf{r}) \approx -\frac{p_F m e^{-r/\xi}}{2\pi^3 r^2 \xi} \cos(2p_F r). \quad (53)$$

因此, $C(r)$ 和 $G(r)$ 虽然有许多共同性质, 如周期性的变化及相干距离, 但它们之间的关系却不很明显.

4. 量子力学中的响应和相干函数

现在提一下响应和起伏在量子力学中的关系. 我们再从 (2) 开始, \hat{B} 、 A 和 H 都是算符, 不一定可以对易, 因此, (3) 不一定对. 现在来求 (2) 对 λ 的展开的一次项, 令

$$U(\beta) \equiv e^{\beta H} e^{-\beta(H-\lambda\hat{A})} \quad (54)$$

则它的微分是

$$\begin{aligned} \frac{\partial U}{\partial \beta} &= e^{\beta H} (H - H + \lambda\hat{A}) e^{-\beta(H-\lambda\hat{A})} \\ &= \lambda\hat{A}(\beta)U(\beta), \end{aligned} \quad (55)$$

$$\hat{A}(\beta) \equiv e^{\beta H} \hat{A} e^{-\beta H}.$$

积分得

$$U(\beta) = 1 + \lambda \int_0^\beta d\tau \hat{A}(\tau) + O(\lambda^2). \quad (56)$$

(2) 式现在可写作

$$B = \frac{\text{Tr} e^{-H/T} U(1/T) \hat{B}}{\text{Tr} e^{-H/T} U(1/T)}, \quad (57)$$

状态空间的统计和 \sum_s 现在改写成本征和 Tr . 将 (56) 代入 (57) 得

$$\chi = \frac{\partial B}{\partial \lambda} = \int_0^{1/T} d\tau [\langle \hat{A}(\tau) \hat{B} \rangle - \langle \hat{A} \rangle \langle \hat{B} \rangle]. \quad (58)$$

当然, 如果 $[\hat{A}, H] = 0$, 也就是说 \hat{A} 如果是守恒量, (58) 和 (3) 就没有两样, (3) 以下的结论也就跟着成立了.

当然, $[\hat{A}, H]$ 在 $\hbar \rightarrow 0$ 时也趋于零, 即在古典力学中可略去.

令 $|s\rangle$ 为定态, 即 H 的本征态, 则

$$\int_0^{1/T} d\tau \langle A(\tau)B \rangle = - \int \frac{d\omega}{2\pi} \frac{S(\omega)}{\omega}, \quad (59)$$

$$\begin{aligned} S(\omega) &\equiv 2\pi (1 - e^{-\omega/T}) \frac{1}{Z} \sum_{s,s'} e^{-E_s/T} A_{ss'} B_{s's} \delta(\omega - E_{s'} + E_s) \\ &\equiv (1 - e^{-\omega/T}) S'(\omega), \\ Z &\equiv \sum_s e^{-E_s/T}. \end{aligned} \quad (60)$$

E_s 为能量本征值, $A_{ss'} \equiv \langle s|A|s'\rangle$. 读者可自证

$$\begin{aligned} \langle AB \rangle &= \int \frac{d\omega}{2\pi} S'(\omega) \\ &= \int \frac{d\omega}{2\pi} \frac{S(\omega)}{1 - e^{-\omega/T}}. \end{aligned} \quad (61)$$

以上的结果总结如下:

$S(\omega)$ 是响应函数的“频率谱”而 $S'(\omega)$ 是相干函数的频率谱, 只有在低频率时, 即 $\omega \ll T$ 时

$$\frac{1}{\omega} S(\omega) \approx \frac{1}{T} S'(\omega). \quad (62)$$

哈密顿算符 H 不但是代表能量, 而且兼管运动的一切细节. $S'(\omega)$ 事实上就是起伏的频率分布:

$$\begin{aligned} S'(\omega) &= \int dt e^{i\omega t} \langle \hat{A}(t) \hat{B} \rangle, \\ \hat{A}(t) &\equiv e^{iHt} \hat{A} e^{-iHt}. \end{aligned} \quad (63)$$

$\hat{A}(t)$ 代表 \hat{A} 之随时间变化. 因此, 从响应函数的分析, 可以同时分析运动的问题[†]. ω 一方面是能量差, 一方面也代表 \hat{A} 的运动时间尺度. 在古典力学中, $\hbar\omega$ 是被看成小得可以略去的, 即 $\hbar\omega/T \rightarrow 0$.

因此, 响应函数和相干函数之间的关系, 在量子力学中比较复杂. 同理, 热位能的微分也不能像 (12) 一样简单地写成起伏的相干值.

[†]这种分析技术, 由日本人 Matsubara(1955) 及一些俄国人发展出来. 见 Abrikosov 等的著作 (1963).

现在看个简例. 再考虑费米粒子气体的密度起伏. 令

$$\begin{aligned}\hat{A} &= \frac{1}{\sqrt{V}}\rho_{\mathbf{k}} \\ \hat{B} &= \frac{1}{\sqrt{V}}\rho_{-\mathbf{k}} = \frac{1}{\sqrt{V}}\rho_{\mathbf{k}}^*\end{aligned}\quad (64)$$

(60) 可以用来求 $S(\omega)$ 和 $S'(\omega)$. 它看来复杂, 其实用起来却不难.

$\rho_{\mathbf{k}}^*$ 的作用是把一粒子的动量增加 \mathbf{k} , 这在讨论散射实验时讲过, 写得仔细些:

$$\rho_{\mathbf{k}}^* = \sum_{\mathbf{p}, \sigma} a_{\mathbf{p}+\mathbf{k}, \sigma}^* a_{\mathbf{p}\sigma}, \quad (65)$$

$\sigma = \pm 1$ 是自转方向, $a_{\mathbf{p}\sigma}$ 是消去一个粒子 (动量 \mathbf{p} 自转 σ) 的算符, $a_{\mathbf{p}+\mathbf{k}}^*$ 生出一个粒子. (65) 就是说: 把在 (\mathbf{p}, σ) 位子中的粒子送到 $(\mathbf{p} + \mathbf{k}, \sigma)$ 位子去. 因此, (60) 中的 $|s'\rangle$ 和 $|s\rangle$ 不同的地方是 $|s\rangle$ 中的某粒子的动量增加了 \mathbf{k} . (60) 中的对 s 求和

$$\frac{1}{Z} \sum_{s'} e^{-E_s/T} |\langle s' | \rho_{\mathbf{k}}^* | s \rangle|^2 \delta(\omega - E_{s'} + E_s) \quad (66)$$

就是对各 $|s\rangle$ 求平均值. 现在考虑 (65) 中的一项 $a_{\mathbf{p}+\mathbf{k}\sigma}^* a_{\mathbf{p}\sigma}$. 如果 $|s\rangle$ 状态有粒子在 (\mathbf{p}, σ) 位子, 且无粒子在 $(\mathbf{p} + \mathbf{k}, \sigma)$ 位子, 则 $\langle s' | a_{\mathbf{p}+\mathbf{k}\sigma}^* a_{\mathbf{p}\sigma} | s \rangle$ 是 1, 否则是零. 其为 1 的概率是 $f_{\mathbf{p}}(1 - f_{\mathbf{p}+\mathbf{k}})$. (66) 中的和就是计算这概率. 因此, (66) 是

$$f_{\mathbf{p}}(1 - f_{\mathbf{p}+\mathbf{k}}) \delta(\omega - \epsilon_{\mathbf{p}+\mathbf{k}} - \epsilon_{\mathbf{p}}). \quad (67)$$

这只是 (65) 中一项的结果. 要计算 (66), 还得对 \mathbf{p}, σ 求和, 注意 $E_s - E_{s'} = \epsilon_{\mathbf{p}+\mathbf{k}} - \epsilon_{\mathbf{p}}$, 因为只有一个粒子的动量有改变.

代入 (60), 得

$$\begin{aligned}S'(\omega) &= \frac{2\pi}{V} \sum_{\mathbf{p}} 2f_{\mathbf{p}}(1 - f_{\mathbf{p}+\mathbf{k}}) \delta(\omega - \epsilon_{\mathbf{p}+\mathbf{k}} + \epsilon_{\mathbf{p}}), \\ S(\omega) &= \frac{2\pi}{V} \sum_{\mathbf{p}} 2(f_{\mathbf{p}} - f_{\mathbf{p}+\mathbf{k}}) \delta(\omega - \epsilon_{\mathbf{p}+\mathbf{k}} + \epsilon_{\mathbf{p}}).\end{aligned}\quad (68)$$

由于 δ -函数, (68) 里的粒子数有以下关系

$$f_{\mathbf{p}} - f_{\mathbf{p}+\mathbf{k}} = f_{\mathbf{p}}(1 - f_{\mathbf{p}+\mathbf{k}}) (1 - e^{-\omega/T}). \quad (69)$$

读者试证之.

将 (68) 代入 (61), (59), 即得 (42) 及响应函数

$$G_{\mathbf{k}} = 2 \int \frac{d^3 p}{(2\pi)^3} \frac{f_{\mathbf{p}} - f_{\mathbf{p}+\mathbf{k}}}{\epsilon_{\mathbf{p}} - \epsilon_{\mathbf{p}+\mathbf{k}}}. \quad (70)$$

这积分不容易做, 我们只看 k 很小的情形, 在这情形

$$\begin{aligned} f_{\mathbf{p}} - f_{\mathbf{p}+\mathbf{k}} &\approx -\mathbf{k} \cdot \mathbf{v} \frac{\partial f_{\mathbf{p}}}{\partial \epsilon_{\mathbf{p}}}, \\ \epsilon_{\mathbf{p}} - \epsilon_{\mathbf{p}+\mathbf{k}} &\approx -\mathbf{k} \cdot \mathbf{v}, \\ \mathbf{v} &\equiv \frac{\partial \epsilon_{\mathbf{p}}}{\partial \mathbf{p}}. \end{aligned} \quad (71)$$

$$G_{\mathbf{k}} \approx 2 \int \frac{d^3 p}{(2\pi)^3} \frac{\partial f_{\mathbf{p}}}{\partial \epsilon_{\mathbf{p}}} \approx \frac{m p_F}{\pi^2}.$$

这结果也可以由 (51) 积分得来

$$\lim_{\mathbf{k} \rightarrow 0} G_{\mathbf{k}} = \int d^3 r G(\mathbf{r}).$$

用 (45) 来求 $I(2r)$, 这积分很容易, 得 (71)

讨论问题十三

1. 在上两章, 物体各部的独立性, 中心极限定理, 散射实验, 相干函数等概念, 都扯在一起. 读者应复习一遍. 本章比较着重在计算方面的问题. 读者应导出每一式子.
2. 热位能的展开式 (11) 是一个连接值的级数. 这展开式在上一章问题 4. 也讨论过.

令该问题中的作用 u 为

$$u(r) = \frac{e^2}{r},$$

即粒子带电荷 e .

(A) 求这模型的展开式 (e^2 的级数) 的头两项.

(B) 以上 (A) 的答案是否合理? 即求出的热位能是否和容积、粒子数成正比? 如果“不合理”, 则原因何在?

注意, 在许多计算中, 常含有各种无限大的积分出现. 这些积分常被视为一种数学上的毛病, 而设法把积分修理或干脆丢掉. 但是这些无限大事实上都是有原因的. 或是模型不合理, 或是极限步骤有问题, 或是某种近似法所使然. 往往这些看来不合理的结果会带来重要的新发现. 总之, 当你面临一个不合理的结果, 你必须设法了解它. 无限大的积分不是零, 不可以丢掉.

3. 响应函数代表物体对外力的响应. 当然, 物体中的任一分子对其他分子来说都是一“外力”, 因此, 响应函数也可以用来分析物体本身的结构. 以下是一个例子.

假想一硬球气体, 令 $G_{0\mathbf{k}}$ 为密度响应函数. $C_{0\mathbf{k}} = G_{0\mathbf{k}}/T$, $C_{0\mathbf{k}}$ 即第 2 节谈的密度相干函数. 现在假想硬球之间又多加了相互吸引力, 作用能为 $-v(\mathbf{r})$, r 为二球间的距离, $v > 0$.

(A) 试证响应函数 $G_{\mathbf{k}}$ 大约是

$$G_{\mathbf{k}} = \frac{G_{0\mathbf{k}}}{(1 - v_{\mathbf{k}}G_{0\mathbf{k}})}. \quad (72)$$

$v_{\mathbf{k}}$ 是 $v(r)$ 的波分量. 这个近似结果和第二十七章的“平均力”近似解内容相似.

(B) 如果 $v_{\mathbf{k}}G_{0\mathbf{k}}$ 趋近于 1, 则 (72) 指出 $G_{\mathbf{k}}$ 变得很大. 注意 $G_{0\mathbf{k}} = C_{0\mathbf{k}}/T$, 而 $C_{0\mathbf{k}}$ 不随 T 变, (硬球的特性, 见第 2 节的例子). 因此, 当

$$T \rightarrow v_{\mathbf{k}}C_{0\mathbf{k}}$$

时, 此物对外力非常敏感, 密度起伏很大, 是一个不稳定的情形. 这可以算是一个“相变”现象的雏形.

令硬球的半径为 a , 令

$$-v(r) = -\epsilon\theta(b - r). \quad (73)$$

求 $T_0 \equiv v_0C_{00}$, 并讨论此温度的意义. (26) 式可用来计算 C_{00} , 这当然只是一粗略的近似计算.

4. 试证一度空间模型 (29)—(32) 不可能有不稳定的情形发生, 即 $G_{\mathbf{k}}$, $C_{\mathbf{k}}$ 都是有限的. 因此不能有相变现象发生.

5. 试将 (29)—(32) 的结果对密度展开, 用以讨论低密度近似结果 (22)—(28).
6. 最后一节的量子力学计算, 可以在固态物理书籍中找出, 不过读者应自己推导, 不要依赖书本.

第四篇

应用

本篇的六章, 讨论一些稍微复杂一点的应用例子. 这些例子, 是用来把一些熟知的自然现象, 作一粗略的分析. 以上各篇中的例子, 多是些理想气体. 本篇以下的例子, 大致以分子间的相互作用为重点, 并介绍磁性、杂质、多相共存现象. 相互作用的分析, 是近代统计力学的内容. 各种理想气体是初等统计力学的领域. 但是, 相互作用是一非常复杂的问题, 今天的数学技巧, 仍不足用. 所做到的是一些极其简化的模型和近似的计算法.

第十四章讨论理想气体定律的修正, 即把分子间相互作用看成一小量, 而求其展开式的第一项. 这章的讨论, 重点在了解相互作用和碰撞时间对热位能的影响. 第十五章简单地复习一下多相共存现象. 这现象可以说是相互作用的最奇妙的后果. 第十六章复习各式磁性现象. 第十七章介绍易形模型. 这是一个最简单的模型, 可用来分析许多不同的现象, 是一个十分重要的工具. 第十八章讨论杂质, 特别强调游动杂质和冻结杂质的不同. 第十九章讨论静电作用. 这是一个在长距离仍不能忽视的作用, 导致许多奇特的现象.

第十四章 理想气体定律的修正

我们讲理想气体时, 分子间的作用, 是看作碰撞的原因, 但碰撞时间长度, 和其他细节却略去了. 因此, 考虑到的效果, 只是使分子动量不断的改变而已. 事实上, 相互作用的效果不只如此, 在这一章, 我们讨论这些作用和它对理想气体定律 $p = TN/V$ 的影响. 我们只讨论低密度下的第一个修正项, 即 $T(N/V)^2$ 的系数 B , 所谓“第二位力系数”. 先复习一下分子间的作用能, 再指出它和 B 的关系. 然后引出碰撞逗留时间的概念, 把碰撞所扮演的角色作更深一层的描述, 对散射理论的一些基本概念, 也稍作讨论.

1. 分子间的作用力

分子间的作用力, 自然要看分子的结构而定. 复杂的分子之间的作用, 十分的复杂, 我们略为复习一下简单分子间的作用.

如果两分子间的距离近到电子层互相接触, 则发生强斥力. 这大致是由于电子为费米粒子, 互不兼容. 如果两分子相距很远, 则会有一点引力. 原因大致是电偶极作用. 原子单独时不会有电偶极, (即电子的重心必与原子核重心重合). 但这可以说是一种时间的平均性质, 电偶极是在高速运转, 平均是零. 如有二原子, 则电子仍会互相影响, (由不同原子组成的分子可以自备偶极矩, 不必其他分子帮忙.) 相互作用能是和 r^{-6} 成正比. r 是原子或分子间的距离, 这 r^{-6} 可以大致解释如下:

在原点放一个原子. 它有个高速运转的偶极矩 \mathbf{d} . 在 r 点的电场是 (略去细节)

$$\mathbf{E} \sim \frac{\mathbf{d}}{r^3}. \quad (1)$$

如在 r 放另一原子, 则 \mathbf{E} 使它的轨道偏向一方, 而产生偶极矩

$$\mathbf{d}' \sim \mathbf{E}. \quad (2)$$

因此二原子间产生作用能

$$U(r) = -\mathbf{d}' \cdot \mathbf{E} \sim -\frac{\mathbf{d}' \cdot \mathbf{d}}{r^3} \sim -\frac{d^2}{r^6}. \quad (3)$$

虽然 d, d' 的平均值均为零, 但 d^2 的平均非零.

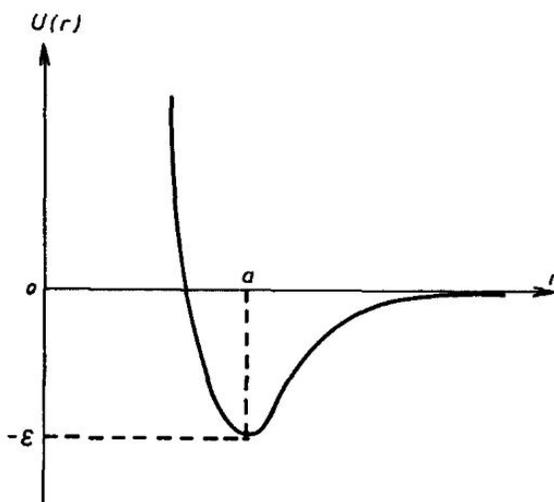


图 1

短距离的斥力和长距离的引力造成一个如图 1 所示的相互作用能. 其极小处之值为 $-\epsilon$, 在 $r = a$. 一般常用所谓“6—12 位能”*

$$U(r) = \epsilon \left(\left(\frac{a}{r} \right)^{12} - 2 \left(\frac{a}{r} \right)^6 \right) \quad (4)$$

来作为近似. 一些例子见表 1, 这样粗略的作用模型, 却有非常广泛的用途, 气体、固体、液体的许多性质, 都可以相当精确地由它来解释.

表 1 作用能极小值及距离

	$a(\text{\AA})$	$\epsilon(\text{焦耳})$
He	2.2	1×10^{-22}
H ₂	2.7	4
Ar	3.2	15
N ₂	3.7	13
CO ₂	4.5	40

Tabor, (1979) p.29

*或称 Lenard-Jones 位能.

2. 理想气体定律的修正

现在来看理想气体公式 $p = \frac{N}{V}T$ 的修正. 我们只讨论 $(N/V)^2$ 项. 我们可以从开放法则出发. 在第九章, 展开式已列出 (见 (9.3) 至 (9.7)). 在此, 取到 z^2 项:

$$\begin{aligned} p &= \frac{T}{V} \ln \sigma, \\ \frac{TN}{V} &= z \frac{\partial p}{\partial z}, \end{aligned} \quad (5)$$

$$\begin{aligned} \ln \sigma &= zZ_1 + z^2 \left(Z_2 - \frac{1}{2}Z_1^2 \right), \\ \frac{N}{V} &= \frac{1}{V} \left[zZ_1 + 2z^2 \left(Z_2 - \frac{1}{2}Z_1^2 \right) \right], \end{aligned} \quad (6)$$

$$z \equiv \frac{e^{\mu/T}}{h^3} (2\pi mT)^{3/2} = \text{理想气体密度},$$

$$Z_1 = V, \quad (7)$$

$$Z_2 = \frac{1}{2} \int d^3\mathbf{r}_1 d^3\mathbf{r}_2 e^{-u(|\mathbf{r}_1 - \mathbf{r}_2|)/T}. \quad (8)$$

把以上式子整理一下, 消去 z , 即得

$$p = \frac{N}{V}T \left(1 + B \frac{N}{V} \right), \quad (9)$$

$$\begin{aligned} B &\equiv -\frac{1}{V} \left(Z_2 - \frac{1}{2}Z_1^2 \right) \\ &= -\frac{1}{2V} \int d^3\mathbf{r}_1 d^3\mathbf{r}_2 \left[e^{-u(|\mathbf{r}_1 - \mathbf{r}_2|)/T} - 1 \right] \\ &= -\frac{1}{2} \int d^3\mathbf{r} \left[e^{-u(r)/T} - 1 \right], \end{aligned} \quad (10)$$

因此, 要求 B , 只需要分析一对粒子的作用. (10) 这积分的值大致是 b^3 , b 是 $u(\mathbf{r})$ 的有效距离. 由此可见, 以上的展开式是

$$zb^3 \sim \frac{N}{V}b^3 \quad (11)$$

的级数. 所谓“低密度”, 是指在相互作用的有效距离内, 不常有分子, 也就是说, 相互作用能不常生效.

因为 $u(r)$ 大致是由一个强斥力, 加上一个弱引力形成, (见图 1), 因此

$$\begin{aligned} B &\approx \frac{2\pi}{3}a^3 + \frac{1}{2T} \int_a^\infty u(r)d^3\mathbf{r} \\ &\equiv v_0 - \frac{y}{T}. \end{aligned} \quad (12)$$

代入 (9) 得

$$\begin{aligned} p &= \frac{N}{V}T \left(1 + \frac{Nv_0}{V} \right) - y \left(\frac{N}{V} \right)^2 \\ &\approx \frac{NT}{(V - Nv_0)} - y \left(\frac{N}{V} \right)^2. \end{aligned} \quad (13)$$

因此, 理想气体公式的修正大致可以看成 ① 由于斥力, 分子可以活动的体积从 V 减小至 $(V - Nv_0)$, ② 由于分子间的吸引, 压力减低. 当一个分子往墙上撞, 背后常有人拉一把, 因此撞得轻一些.

(13) 式就是有名的“范德瓦尔斯”方程式, 在讨论相变时会再出现.

B 是所谓“第二位力系数”. 密度展开常称作“位力展开”. (9) 式可直接由位力定理求得, 可作练习用. (见讨论问题)

从气体压力的测量结果、可以推测 B , 进一步利用公式 (10) 大致可定出图 1 的作用能.

3. 碰撞的逗留时间

以上的分析, 十分简单. 现在从一个不同的观点来看. 虽然比较复杂, 但它把分子运动的角色, 描述得清楚些.

早在讨论“细致平衡”时, 就强调了碰撞的重要性. 理想气体不是指碰撞不发生, 而是指碰撞发生的时间十分短. 因此, 理想气体公式的修正, 就是由碰撞时间引起. 也就是说, 每个分子的自由时间受到了影响. 以下的分析的结论是 (10) 式的 B 是和碰撞的“逗留时间”成正比.

我们先把 $z^2 Z_2$ 改写一下:

$$z^2 Z_2 = \frac{e^{-2\mu/T}}{2h^6} \int d^3\mathbf{p}_1 d^3\mathbf{p}_2 d^3\mathbf{r}_1 d^3\mathbf{r}_2 e^{-H/T}, \quad (14)$$

$$\begin{aligned} H &= \frac{p_1^2}{2m} + \frac{p_2^2}{2m} + u(|\mathbf{r}_1 - \mathbf{r}_2|) \\ &= \frac{P^2}{2M} + \frac{p^2}{2m'} + u(r), \end{aligned} \quad (15)$$

$$d^3\mathbf{p}_1 d^3\mathbf{r}_1 d^3\mathbf{p}_2 d^3\mathbf{r}_2 = d^3\mathbf{R} d^3\mathbf{P} d^3\mathbf{r} d^3\mathbf{p}.$$

式中的 \mathbf{P} 是二粒子的总动量, $M = 2m$ 是总质量. \mathbf{p} 是重心坐标系中的相对位置及动量, $m' = m/2$

$$\begin{aligned} \mathbf{P} &= \mathbf{p}_1 + \mathbf{p}_2, & \mathbf{p} &= \frac{1}{2}(\mathbf{p}_1 - \mathbf{p}_2), \\ \mathbf{R} &= \frac{1}{2}(\mathbf{r}_1 + \mathbf{r}_2), & \mathbf{r} &= \mathbf{r}_1 - \mathbf{r}_2. \end{aligned} \quad (16)$$

我们假设 $u(\mathbf{r})$ 和 \mathbf{r} 的方向无关. 将 (15) 代入 (14), d^3R 和 d^3P 的积分不涉及 $u(\mathbf{r})$, 可以立即做了, 得

$$\begin{aligned} z^2 Z_2 &= \frac{e^{2\mu/T}}{h^3} (2\pi MT)^{3/2} V \times V \int d\epsilon g(\epsilon) e^{-\epsilon/T}, \\ Vg(\epsilon) &\equiv \frac{1}{h^3} \int d^3\mathbf{r} d^3\mathbf{p} \delta\left(\epsilon - \frac{p^2}{2m'} - u(r)\right). \end{aligned} \quad (17)$$

$g(\epsilon)$ 自然是粒子轨道的能量分布, 它可以直接由轨道的统计得来. 各轨道可以由其守恒量加以分类. 守恒量中除了能量 ϵ 之外, 还有角动量

$$\mathbf{l} = (\mathbf{r} \times \mathbf{p})/\hbar. \quad (18)$$

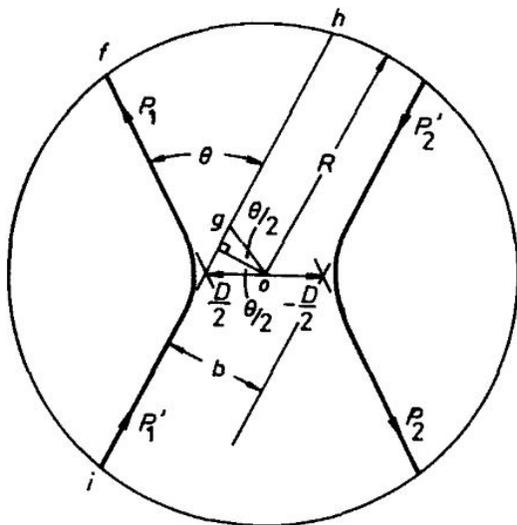


图 2 重心坐标系中的粒子轨道

(见图 2). (17) 可作

$$g(\epsilon) = \frac{1}{8\pi^3\hbar} \int 2\pi l dl \frac{dp_r}{d\epsilon} dr d\Omega. \quad (19)$$

式中的 p_r 是 \mathbf{p} 在 \mathbf{r} 方向的分量, $d\Omega$ 积分是对入射方向的积分. 因为 $d\epsilon/dp_r = v_r$ 是在 \mathbf{r} 方向的速度, 所以, $\mathbf{v}_r = d\mathbf{r}/dt$.

$$\frac{dp_r}{d\epsilon} dr = \frac{dt}{dr} dr = dt \quad (20)$$

是时间的微分, (19) 成为

$$Vg(\epsilon) = \frac{1}{8\pi^3\hbar} \int 2\pi l dl 4\pi \int dt. \quad (21)$$

如果 $u(r) = 0$, 则

$$z^2 Z_2 = \frac{1}{2} Z_1^2. \quad (22)$$

所以, 将 (21) 代入 (17), 得

$$\begin{aligned} z^2 \left(Z_2 - \frac{1}{2} Z_1^2 \right) &= \frac{e^{2\mu/T}}{h^3} (2\pi MT)^{3/2} V \\ &\times \frac{1}{\pi\hbar} \int_0^\infty l dl \int_0^\infty d\epsilon \cdot t'(\epsilon, l) e^{-\epsilon/T}, \end{aligned} \quad (23)$$

$$\begin{aligned} t'(\epsilon, l) &\equiv \int dt(\text{轨道}) - \int dt(u=0\text{轨道}) \\ &= \text{逗留时间}. \end{aligned} \quad (24)$$

这也就是图 2 中从 i 到 f 所费的时间减去从 i 到 h 的时间. 图中这轨道的逗留时间显然是负的, 即跑出来得比没有碰撞还要快. 像图 3 的情形 $t'(\epsilon, l)$ 就是正的.

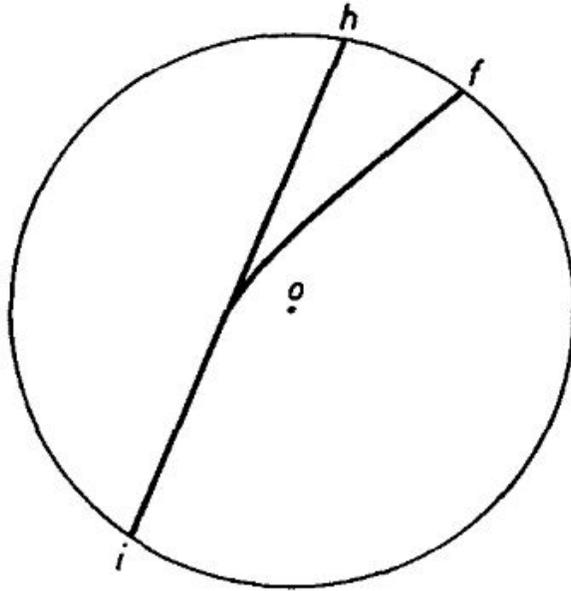


图 3 同图 2, 只是相互作用力为引力, 粒子 2 的轨道未画出.

以上分析, 总结一下, 是

$$B = \frac{2^{3/2}\lambda^3}{\pi\hbar} \int dl \int d\epsilon t'(\epsilon, l) e^{-\epsilon/T}, \quad (25)$$

$$\lambda \equiv h/(2\pi mT)^{1/2}.$$

这个结果的意义是: B 和轨道的关系, 只是逗留的时间而已. 这逗留时间是可以从轨道的自由部分定义. 在图 2 中, 只需测得粒子入圈时间和出圈时间. 至于轨道在圈内的详情, 则全不必知道, 也就是说, 要计算 B , 我们只需要知道入射前及射出后的数据, 不必知道作用时的数据.

从第 3 节到此的分析, 有一个重点, 即系数 B 并不必直接涉及分子作用的细节. 它可以直接以作用前后的各量表示, 虽然 B 是一个统计力学的概念, 而且由 (10) 看来, 它是和作用 $u(r)$ 有直接的关联, 但是以上分析的结论是, 统计并不必涉及短距离的细节, 只要 $u(r)$ 有效范围之外的数据就够了. 量子力学的结果更是明显, 波函数 $\psi(r)$ 在小 r 处的细节全不需要. 重要的是在 r 很大时, 自由区域内的 $\psi(r)$. 束缚态则可当作一自由粒子.

以上这些分析的进一步推广, 请参阅文献[†].

[†]Dashen, Ma, Bernstein (1969), Dashen and Ma (1971).

4. 量子力学结果

用量子力学来计算 B , 步骤和以上的差不多. 有两个要修正的地方. 一是粒子的统计性质 (波色和费米), 这不是很困难的事, 我们不讨论它. 二是 (17) 的 $g(\epsilon)$ 须要用量子力学计算, 而不用 (17) 这短波近似. 这是个有趣的问题, 讨论如下.

我们把 \mathbf{r} 限制在一个半径为 L 的球中. 只有在原点附近, $u(r)$ 才不是零. $g(\epsilon)$ 可以写成

$$g(\epsilon) = \sum_{l=0}^{\infty} (2l+1) \sum_n \delta(\epsilon - \epsilon_{nl}), \quad (26)$$

ϵ_{nl} 来自波动方程式

$$\left(-\frac{1}{2m'} \nabla^2 + u(r) \right) \psi_{nlm} = \epsilon_{nl} \psi_{nlm}, \quad (27)$$

$$\psi_{nlm} = Y_{lm}(\mathbf{r}) \phi_{ln}(r),$$

的解. 令在边界上的 ϕ_{ln} 值为零, 以定 ϵ_{nl} . 边界 $r = L$, 距原点很远. 在 $u(r)$ 的势力范围外,

$$\phi_{ln}(r) \sim \frac{1}{r} \sin \left(kr + \frac{l\pi}{2} + \delta_l \right). \quad (28)$$

k 的值, 决定于 $\phi(L) = 0$, 即

$$kL + \frac{l\pi}{2} + \delta_l = n\pi, \quad n = \text{整数}, \quad (29)$$

$$\epsilon_{nl} = \frac{k^2}{2m},$$

δ_l 是量子散射理论中的“相位移”. 它的意义简述如下.

我们学初等光学时, 学到“像”的概念. 反射镜, 透镜把来自某光源的光反射了, 或折射了等等. 最方便的分析方法是说反射镜或透镜造成了一个“像”. 光好像从“像”射出似的, 像位一旦求出, 问题就解决了. 散射理论中的“相位移”有相同的效用. 如果 $u(r) = 0$, 即粒子不受任何作用, 则 $\delta_l = 0$, $\phi_{ln}(r)$ 可写作入射波和出射波之和, 即 (28) 是

$$\phi_{ln}(r) \propto -(-1)^l e^{-ikr} + e^{ikr}, \quad (30)$$

末项为出射波. 右第一项为入射波, 前面的 $-(-1)^l$ 和三维空间的几何有关, 我们可以不要管. 如果 $u(r) \neq 0$, 则出射波受影响, (28) 成为

$$\phi_{ln}(r) \propto -(-1)^l e^{-ikr} + e^{2i\delta_l} e^{ikr}, \quad (31)$$

即 (30) 的第二项多了个 $e^{2i\delta_l}$ 乘数. 因此, 碰撞的结果, 是把出射波的振动相角加了 $2\delta_l$. 要看粒子如何射出, 我们把动量在 k 附近的波组合起来, 使成一厚度有限的球面波, 再看这波如何向外移动:

$$\psi(r, t) \propto \int dq A(q) e^{2i\delta_l(k+q)} e^{i(k+q)r} e^{-i\epsilon(k+q)t}. \quad (32)$$

以上只考虑出射波, 并把时间因子放进去,

$$\begin{aligned} A(q) &= \int dr e^{-iqr} a(r), \\ a(r) &= \frac{1}{2\pi} \int dq e^{iqr} A(q), \end{aligned} \quad (33)$$

$a(r)$ 是一个比 $1/k$ 宽很多的包络曲线, (见图 4)

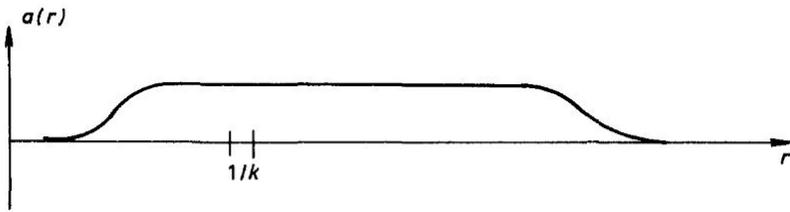


图 4

因为 $a(r)$ 很宽, 故 $A(q)$ 很窄, (32) 的指数可以展开, 得

$$\begin{aligned} \delta_l(\mathbf{k} + \mathbf{q}) &\approx \delta_l(\mathbf{k}) + \mathbf{q} \frac{\partial \delta_l}{\partial \mathbf{k}}, \\ \epsilon(\mathbf{k} + \mathbf{q}) &\approx \frac{1}{2m'} (\mathbf{k} + \mathbf{q})^2 \approx \epsilon(\mathbf{k}) + \frac{\mathbf{k}}{m'} \cdot \mathbf{q}. \end{aligned} \quad (34)$$

应用 (34) 的近似 (32) 可写成

$$\begin{aligned} \psi(r, t) &\propto e^{2i\delta_l(k)} e^{ikr - i\epsilon(k)t} \int dq A(q) e^{iq(r - vt - r_0)} \\ &\propto e^{2i\delta_l(k)} e^{ikr - i\epsilon(k)t} a(r - vt + r_0), \end{aligned} \quad (35)$$

$$\begin{aligned} v &\equiv k/m', \\ -r_0 &\equiv 2 \frac{\partial \delta_l}{\partial k}. \end{aligned} \quad (36)$$

(35) 的解释十分清楚: 出射波以速度 v 向外移动, 它的来源在 r_0 , 也就是说, “像” 的位置在 $r = r_0$, 如果 $\delta_l = 0$, 则来源在 origin. 作用能 $u(r)$ 的结果是在 r_0 成 “像”.

现在回到 (29), 及 (26). 我们要的是 $g(\epsilon) - g_0(\epsilon)$. $g_0(\epsilon)$ 是 $u(r) = 0$ 时的能量分布, 从 (29) 可得

$$\frac{dn}{d\epsilon} - \left(\frac{dn}{d\epsilon}\right)_0 = \frac{1}{\pi} \frac{d\delta_l}{d\epsilon}. \quad (37)$$

这 $dn/d\epsilon$ 是指在能量间隔 $d\epsilon$ 内有多少角动量为 l 的状态. 因此, 由 (29), (37) 得

$$g(\epsilon) - g_0(\epsilon) = \sum_{l=0}^{\infty} (2l+1) \frac{1}{\pi} \frac{d\delta_l}{d\epsilon}. \quad (38)$$

从这式子得[‡]

$$B = -2^{3/2} \lambda^3 \sum_{l=0}^{\infty} (2l+1) \frac{1}{\pi} \int d\epsilon e^{-\epsilon/T} \frac{d\delta_l}{d\epsilon}. \quad (39)$$

是为量子力学计算的第二位力系数. 此式尚未考虑粒子的统计性. 如果是波色粒子, 则 l 必为偶数. 如果为费米粒子, 则要看粒子的自旋. 如是 $\frac{1}{2}$, 二自旋平行, 则 l 必须是奇数, 整个式子须乘以 3, 如二自旋反向, 则 l 为偶数. 平行和反向的结果加起来才是总系数 B .

现在可以把 (39) 和 (25) 作一比较. 由 (36) 得

$$\begin{aligned} 2 \frac{d\delta_l}{d\epsilon} &= 2 \frac{d\delta_l}{dk} \bigg/ \frac{d\epsilon}{dk} \\ &= -r_0/v = t'(\epsilon, l). \end{aligned} \quad (40)$$

r_0 是像的位移, 除以速度, 正好就是逗留时间, 也就是说, 射出波看来是迟了 t' . 因此, 除了 l 该改成 $l + \frac{1}{2}$, 积分该改为求和之外, (25) 和 (39) 是完全一样的.

5. 束缚态, 勒文生定理

以上的分析, 没有包括两粒子的束缚态. 如果两粒子可结合而成一新分子, 则是一种化学反应



A_2 的化学势为 2μ , 质量为 $2m \equiv M$, 能量为动能减去结合能 W :

$$\frac{P^2}{2M} - W. \quad (42)$$

[‡]这结果见 Beth and Uhlenbeck (1937).

这 A_2 气体的压力是

$$p_2 = T \frac{e^{(2\mu+W)/T}}{\lambda_2^3} \approx 2^{3/2} T \lambda^3 \left(\frac{N}{V}\right)^2 e^{W/T}, \quad (43)$$

$$\lambda_2 \equiv \frac{h}{(2\pi MT)^{1/2}} = \frac{\lambda}{\sqrt{2}}.$$

要造一个 A_2 需要两个 A , 因此, (41) 的反应, 反使总压力减少 p_2 :

$$p = \frac{N}{V} T - p_2 \quad (44)$$

$$\approx \frac{N}{V} T \left(1 - 2^{3/2} \left(\frac{N}{V}\right) \lambda^3 e^{W/T}\right).$$

和 (9) 比较一下, 可见束缚态 A_2 的效果是

$$B = -2^{3/2} \lambda^3 e^{W/T}. \quad (45)$$

碰撞和束缚态的效果统统算在一起, 可以写成

$$g(\epsilon) - g_0(\epsilon) = \sum_{l=0}^{\infty} (2l+1) \left[\sum_{\alpha} \delta(\epsilon + W_{l\alpha}) + \frac{1}{\pi} \frac{d\delta_l}{d\epsilon} \right]. \quad (46)$$

也就是 (38) 式加上束缚态 $l\alpha$. 当然, 如果某个 l 没有束缚态, $\delta(\epsilon + W_{l\alpha})$ 就不存在. (25) 也没有包括束缚态的影响, 也该修正.

我们顺便提一下散射理论中的一个有趣的定理, 即“勒文生定理”[§]

$$n_l = \frac{1}{\pi} [\delta_l(0) - \delta_l(\infty)]. \quad (47)$$

式中 n_l 是角动量为 l 的束缚态数目. 在这低能和高能相位移之差. 这定理把散射实验的结果, 即 δ_l 和束缚态扯到一起, 看来十分玄妙. 但是, 从统计的观点来看, 这定理的意义十分明显. 说明如下.

令 $g_l(\epsilon)$ 为角动量为 l 的诸状态时能量分布, 则

$$g_l(\epsilon) - g_{0l}(\epsilon) = (2l+1) \left[\sum_{\alpha} \delta(\epsilon + W_{l\alpha}) + \frac{1}{\pi} \frac{d\delta_l}{d\epsilon} \right]. \quad (48)$$

是角动量为 l 的形像总数, 即整个角动量为 l 的状态空间的总体积, 应该和相互作用 $u(r)$ 无关. 作用的功效只是把各状态的能量改变一下, 把波函数重新组和, 但不会改变状态的总数. 因此

$$\int_{-\infty}^{\infty} d\epsilon [g_l(\epsilon) - g_{0l}(\epsilon)] = 0. \quad (49)$$

[§]Levinson 定理, 见量子力学书籍.

(48) 右边的积分就是

$$(2l+1) \left[n_l + \frac{1}{\pi} (\delta(\infty) - \delta(0)) \right]. \quad (50)$$

由 (49), (50) 即得 (47). (读者注意, 一般的情形下, 虽然 (49) 成立, 但若是把各 l 项加起来, 就不一定对了, 原因是级数不一定收敛.)

讨论问题十四

1. 试证“位力定理”, (从牛顿定律开始):

$$\sum_{i=1}^N \langle \mathbf{r}_i \cdot \mathbf{F}_i \rangle + 2K = 0. \quad (51)$$

\mathbf{r}_i 是第 i 个粒子的位置, \mathbf{F}_i 是它所受的力, K 是总动能, $\langle \dots \rangle$ 是长时间的平均值, 共有 N 个粒子.

2. 某气体的体积为 V , 总能量为

$$H = \sum_{i=1}^N \frac{p_i^2}{2m} + \frac{1}{2} \sum_{i,j} U(|\mathbf{r}_i - \mathbf{r}_j|). \quad (52)$$

(A) 用“位力定理”证明:

$$\text{如果 } U = 0, \quad \text{则 } pV = \frac{2}{3}K. \quad (53)$$

(B) 令 U 的有效距离为 a , 即

$$U(r) = 0, \quad r > a. \quad (54)$$

试证: 如果 $a \rightarrow 0$, 则即使 $U(r)$, $r < a$, 很大, pV 仍然大约是 $\frac{2}{3}K$.

(C) 如果 $\alpha \equiv Na^3/V \ll 1$, 求第二位力系数 B :

$$pV = NT \left(1 + B \left(\frac{N}{V} \right) + O(\alpha^2) \right). \quad (55)$$

从 (51) 开始. 再利用上章的相干函数得 $\sum_i \langle \mathbf{r}_i \cdot \mathbf{F}_i \rangle$.

3. 某金属内有固定的杂原子, 电子和杂原子的作用能为 $U(r)$, r 为电子和一杂原子间的距离. 令电子散射的相位移为 δ_l .

(A) 求此杂质引起的热容率改变. (提示: 热容率只和 $g(\epsilon)$ 有关), 假定杂原子的密度很低.

(B) 如果杂原子是硬球, 半径为 a , 试将 (A) 的结果算出. 考虑 $k_F a \ll 1$ 和 $k_F a \gg 1$ 两特殊情形. k_F 是顶层电子的动量.

在 $k_F a \gg 1$ 的情形下, 答案很容易求, 不必经过 δ_l , 只须考虑杂质所占的体积.

4. 高次位力系数, 比较复杂. 展开式须从第十二章的问题 4 做起. 有关这一方面的书籍很多, 例如 Brout and Carruthers (1963) 的第 1 章. 读者可自修.

第十五章 多相共存

一般说来, 在一定的环境下, 例如在一定的温度、压力、磁场下, 物系的性质, 如密度、熵、磁矩方向等, 都有一定的值. (微小的起伏不计). 但在某些特殊环境下, 有些性质却不一定, 而可以有两个或更多的值. 例如在 100°C , 一大气压下, 水可呈气相 (密度小) 或液相 (密度大). 再如一磁铁, 如无外加磁场, 则其磁矩不一定, 而可指任一方向. 在这些特殊环境下, 蒸汽和水可以共存, 磁铁的磁矩可以有一半朝东, 一半朝西. 这样的现象曰多相共存, 这是一极为广泛, 极其困难的题目, 也是近代统计力学中最引人注目的一部门. 这一章只将最基本的一些常识和概念作一简介.

1. 气相和液相

一般原子之间的作用力, 在短距离时为强斥力, 在较长距离有弱引力, (见图 1). 因此, 各原子大致是一些略相吸引的硬球. 温度很低时, $F = E - TS$ 的极小值决定于 E , 故呈晶体, 使 E 小. 高温时, F 决定于 $-TS$, 故呈气相, S 大也. 液体出现在中间温度, 一方面原子间距离和 a 相近 (见图 1), 一方面位置不定, 熵亦不小. 液体中原子的详细运动情形非常复杂, 有短时间晶状排列, 也有滑动、转动、振动等运动. 其详情视原子或分子的结构和其他性质而异, 将不在此地讨论.

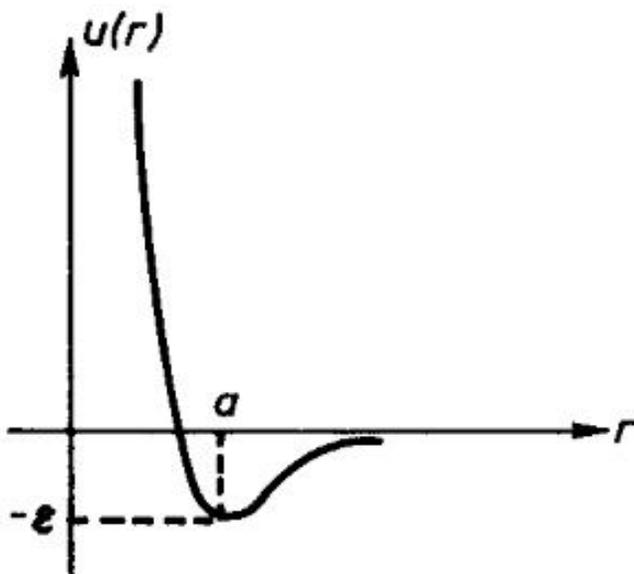


图 1

现在, 讲一点二相共存的简单常识. 当二相共存时, 例如一瓶水和水上的蒸汽, 这物体就不是一均匀的物体, 因为水和蒸汽的密度不同. 不过水这一部分是均匀的. 同样, 蒸汽的一部分也是均匀的. 这两部分可看作两个物体, 在定温定压的环境中. 既然分子可以自由地从一物系到另一物系, 故在平衡状态下, 分子的化学势在水中和在蒸汽中是一样的,

$$\mu_1(T, p) = \mu_2(T, p). \quad (1)$$

乍看起来, μ_1, μ_2 的意义很明显, 但细看却有问题, 要费些周章才能说清楚, 现在且就 (1) 谈些简单的结论. 把这些周章延到末节再说.

(1) 可以说是在 (T, p) 平面上一条曲线的定义. 只有在这曲线上的点才适合 (1) 这条件. 这曲线可称为相 1 和相 2 的分界线 (见图 2), 气相, 液相的分界线之终点称为“临界点”. 此点以上就没有气液之别了, 同理, 若是要有三相共存, 则 (T, p) 需合乎

$$\mu_1(T, p) = \mu_2(T, p) = \mu_3(T, p). \quad (2)$$

这 (T, p) 可看作 (1, 2), (2, 3), (3, 1) 三条分界线的交点, 即“三相点”.

气相部分, 曰蒸汽. 平衡的气压曰“饱和蒸汽压”. 现在复习几个名词, 及常识.

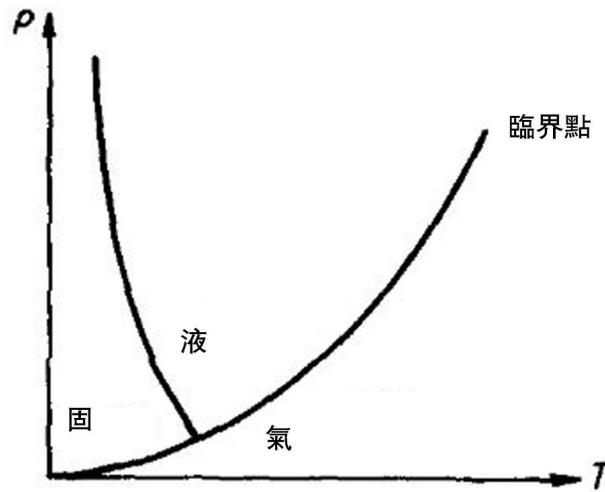


图 2

(甲) 潜热:

潜热 q 是把一分子从一相送到另一相所需的能量, 令

$$\begin{aligned} \epsilon_1 &= E_1/N_1, & \epsilon_2 &= E_2/N_2, \\ v_1 &= V_1/N_1, & v_2 &= V_2/N_2, \\ s_1 &= S_1/N_1, & s_2 &= S_2/N_2. \end{aligned} \quad (3)$$

要把一分子从相 2 送到相 1, 不但需要能量 $\epsilon_1 - \epsilon_2$, 还要对压力做功 $p(v_1 - v_2)$. 因此

$$q = (\epsilon_1 - \epsilon_2) + p(v_1 - v_2). \quad (4)$$

因为 $\mu_1 - \mu_2 = 0$, 且 $\mu_{1,2} = \epsilon_{1,2} - Ts_{1,2} + pv_{1,2}$, 故

$$(\epsilon_1 - \epsilon_2) - T(s_1 - s_2) + p(v_1 - v_2) = 0.$$

即

$$q = T(s_1 - s_2). \quad (5)$$

潜热是可以直接测量的, 因此, 二相的熵差可以直接测量.

(乙) 交界线方程式

沿着交界线 $\mu_1(T, p) = \mu_2(T, p)$ 取微分, 得 $d\mu_1 = d\mu_2$,

$$\text{即 } -s_1dT + v_1dp = -s_2dT + v_2dp, \quad (6)$$

因为 $d\mu = -sdT + vdp$. 整理之, 得

$$\frac{dp}{dT} = \frac{s_1 - s_2}{v_1 - v_2} = \frac{q}{T(v_1 - v_2)}. \quad (7)$$

q 是由 (5) 而来, 这就是 1, 2 两相的交界方程式. 此式至为重要, 因为 q , T , v_1 , v_2 , 都是可以直接测量的, 因此整个交界线可以用积分算出. 如二相之一是气体, 则交界线称为“蒸汽压曲线”. *

现在来用 (7) 看气相和固相在极低温度的交界线, 以此作为一简单的应用例子. 令 1 指气相, 2 指固相, 温度既低, 气压亦低, 故可将气体作理想气体看. 固体密度较气体大得多, 故

$$v_1 - v_2 \approx v_1. \quad (8)$$

因为 $pv_1 = T$, 故

$$\frac{dp}{dT} = \frac{qp}{T^2}. \quad (9)$$

现在

$$\begin{aligned} q &= \epsilon_1 - \epsilon_2 + p(v_1 - v_2) \\ &\approx \epsilon_1 - \epsilon_2 + T \approx \epsilon_1 - \epsilon_2. \end{aligned} \quad (10)$$

$\epsilon_1 - \epsilon_2$ 为将一分子踢出固体的能量. 固体外气压甚低, T 很小, 可以略去. 因此, 如 $\epsilon_1 - \epsilon_2$ 大致不随 T 变化, (9) 立即可以解出, 得固体的饱和蒸汽压.

$$p \propto e^{-(\epsilon_1 - \epsilon_2)/T}. \quad (11)$$

这自然是意料中事.

(丙) 蒸发率

每单位面积的液面, 每秒有多少分子蒸发? 现在来估计一下这蒸发率. 令液体和其蒸汽呈平衡状态. 则蒸发率等于气体分子进入液体的速率. 假设蒸汽很稀, 可当作一理想气体, 则每秒每平方厘米气体分子, 撞上液面的有

$$n\bar{v}_z = n \frac{\int_0^\infty dp_z (p_z/m) e^{-\beta p_z^2/2m}}{\int_{-\infty}^\infty dp_z e^{-\beta p_z^2/2m}} = \frac{nT}{\sqrt{2\pi mT}}. \quad (12)$$

式中 n 是气体密度, v_z 是垂直于液面的平均速度. 有些撞上液面后不留下, 却立即回升. 因此真正进入液体的只有 (12) 的一部分, $0 < \alpha < 1$, 这 α 可称作“胶着系数”. 既然进去的等于出来的, 故

$$\text{蒸发率} = \alpha \frac{p}{\sqrt{2\pi mT}}. \quad (13)$$

* (7) 式亦称 Clausius-Claypeyron 方程式.

这 α 的值要看详细运动情形, 液面结构等等. p 是饱和蒸汽压 $= nT$.

现在调回头来说, 即使蒸汽不饱和, (13) 仍然该是对的. 因为蒸发率和有多少分子进来不会有太大关系. 只要把 T 当作液体的温度就行了, 饱和气压可当作一温度的已知函数. 当然, 蒸汽不饱和, 液体蒸出的多, 进来的少, 过一会儿就干了.

蒸发和沸腾不一样. 如一气泡在液体中出现后, 能支持住不缩小, 这种温度就叫沸点, 气泡中的压力是饱和蒸汽压. 因此, 当饱和蒸汽压等于液体压力时, (即外加压力, 且常是一大气压) 再加热就造成沸腾.

2. 水珠的消长

以上的讨论着重于二相的共存, 对二相之间的分界面未作任何考虑. 如在分界面的分子, 比在液中、气中的少很多, 则可以略去不计. 但有时没有少很多, 则不能略去. 例如一个小水珠, 其表面上的分子对整个水珠的性质影响颇大. 水珠的成长是凝结现象的第一步, 这是一门大学问, 在讨论雨、露、霜、雾时不可少, 气象学者们在这方面下了不少工夫, 现在我们把水珠消长的最基本概念提一下.

水珠所处的环境, 是一个定温、定压, 而且是开放的环境. 我们定义热位能 Λ 为

$$\Lambda(p, T, N) = G(p, T, N) - \mu N. \quad (14)$$

N 为水珠内的分子数, μ 是环境的化学势, 即水蒸汽的化学势, G 为水珠的定温定压位能. 在平衡时, Λ 的极小在

$$\partial\Lambda/\partial N = 0 \text{ 处, 即 } \mu = \partial G/\partial N. \quad (15)$$

如果 N 是一大数, 则 $G(N)$ 该是

$$G(N) = Ng + N^{2/3}\eta + \dots. \quad (16)$$

略去的是比 $N^{2/3}$ 次数低的各项. 如果 $N \gg N^{2/3}$, 而略去 $N^{2/3}$, 即略去分界面, 则 G 必须是 $\mu'N$, $\mu'(T, p)$ 是液体的化学势. 因此, $g = \mu'$. 第二项的来源是水珠表面张力. 在表面的分子数和 $N^{2/3}$ 成正比, 只要表面够平. 这些分子有一边没有其他分子作邻居, 少了吸引力, 因此少了负的作用能, 故表面

积越大, 能量越高. η 因此和表面张力 α 成正比. (此 α 与 (13) 的 α 无关) $\eta N^{2/3} = \alpha 4\pi r^2$, 所以

$$\alpha = \frac{\eta}{4\pi} \left(\frac{4\pi}{3} n' \right)^{2/3}, \quad (17)$$

n' 是液体的密度, r 是水珠半径. 注意, 我们在此并没有对这水珠作确切的定义. 水珠的“表面”并非一确定的几何面, 那些分子在“水珠内”, 那些算是在“水珠外”, 并没有明确的定义. 所谓水珠的表面, 只是一大略的内外分界. 在 (16) 中, 略去的项并无太大意义, 因为 N 之值本身的意义有限. (大致说来, N 的意义只限于 $\pm\sqrt{N}$ 之外, 注意 $\sqrt{N} < N^{2/3}$).

将 (16) 代入 (14), 得

$$\Lambda(N) \approx (\mu' - \mu) N + N^{2/3} \eta. \quad (18)$$

图 3 示 $\Lambda(N)$. 最后一项总是随 N 而增的, $(\mu' - \mu)N$ 这项可要看 μ' 和 μ 那个大些, 如果 $\mu' > \mu$, 则 $\Lambda(N)$ 一直是随 N 而增, 如果 $\mu' < \mu$, 则先增而后落, μ 是水珠四周水蒸汽的化学势, μ' 是在同一温度, 压力下液体的化学势, μ 和 μ' 可以用蒸汽的压力 p 和饱和蒸汽压 p' 表示. 把蒸汽当作一理想气体, 则

$$p = e^{\beta\mu} \frac{z}{\lambda^3}, \quad (19)$$

$$\frac{z}{\lambda^3} \equiv \int \frac{d^3\mathbf{p}}{(2\pi)^3} e^{-\beta p^2/2m} \sum_{\alpha} e^{-\beta\epsilon_{\alpha}} \quad (20)$$

(20) 中的 α 是指水分子内部结构的状态, 从 (19) 得

$$\begin{aligned} \mu &= T (\ln p - \ln (z/\lambda^3)), \\ \mu' &= T (\ln p' - \ln (z/\lambda^3)), \end{aligned} \quad (21)$$

$$\mu - \mu' = T \ln (p/p') \equiv T \ln s.$$

式中的 $s = p/p'$ 称为“饱和度”, 注意 p' 是当水和蒸汽平衡时的蒸汽压, 即饱和气压. 如果水珠四周的气压 p 比 p' 小, 即 $s < 1$, 则曰“未饱和”. 如 $s > 1$, 则称“过饱和”, 过饱和则 $\mu' < \mu$. 当 N 够大时, $\Lambda(\mu)$ 会落下 (见图 3). N 趋向 $\Lambda(N)$ 低处, 则 N 将无限制增加, 即水珠越长越大. 如未饱和, 则 N 不会增大, 水珠将消失. 当然最重要的一个特点是在过饱和蒸汽中, 水珠的体积, 如是小於临界体积 N_c , (见图 3), 也不免趋于消失的方向. 因此, 若非先有些大于临界体积的水珠, 蒸汽就不会凝成水. N_c 可由 (18) 微分求出, 得

$$N_c^{1/3} = \frac{2\eta}{3T \ln s}. \quad (22)$$

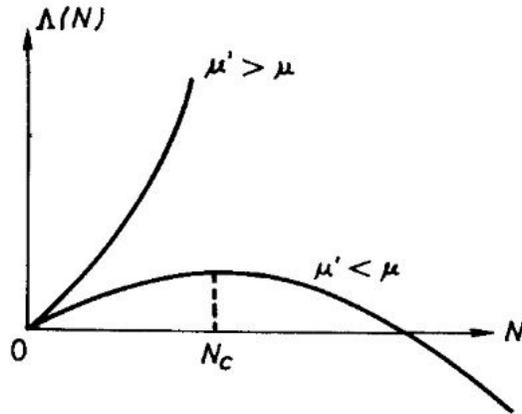


图 3

因为饱和度 s 是以其对数 $\ln s$ 方式出现, 所以它对 N 的影响不太大, s 加个几倍对 N_c 不会发生大作用. 在 $T \sim 273\text{K}$, $s \sim 250\%$ 时, $N_c \sim 300$, 相当于半径约 13 的水珠, 虽然这临界体积并不大, 但事实上在水蒸汽中要三百个分子靠碰撞而聚合却不容易. 据实验 $s \sim 500\%$ 的过饱和蒸汽都不会凝结. 这当然是指纯水蒸汽而言, 如果汽中有杂质, 如灰尘、水粒, 各种带电的粒子, 如离子等, 则凝结在 s 稍大于 1 就可以很快发生. 杂质成了凝结的中心, 水珠在各粒子上开始形成. 有关这一方面的常识, 读者可以从气象学书籍中得到, 许多十分有趣的问题, 如雾、露、霜、雪、雨点的运动, 人造雨的技术等, 都在气象学中.

3. 潜热和分子作用能

在第 1 节提到, 固相和气相间的潜热, 即“升华热”, 和把一原子踢出固体的能量相近. “见 (10)”. 这能量当然和图 1 这条曲线有关. 固体中, 原子间的距离大约是 a , 一对原子间的作用能约为 $-\epsilon$. 因此, 要踢出一个原子, 必须拆散 n 对原子, 假定每原子有 n 个邻居. 因此, 升华热是

$$q = \frac{1}{2}n\epsilon. \quad (23)$$

这是因为 N 个原子间只能有 $1/2Nn$ 对邻居. fcc(面心方) 及 hcp(六角密集) 的晶体, 都有 $n = 12$ 即 $q = 6\epsilon$, 表 1 是一些此类晶体的升华热及 ϵ 的数字. ϵ 可以由计算, 或气体的性质求得. “见第十五章”

表 1

晶体	升华热实验	$q(10^3\text{J/mole})$ 计算 (23)	$\epsilon(\text{焦耳})$
He	0.08	0.33	0.9×10^{-22}
Ne	1.3	1.7	4.8
Ar	7.3	5.9	16.5
Kr	10.0	9.0	25

Tabor(1979) p.152.

表 1 中的分子, 都是球形, 图 1 的作用能量可靠. 此表中所谓“实验”值, 其实是熔解热及蒸发热之和. 升华热不容易直接测定. 熔解热和蒸发热大致是由固相经液相而至于气相所需的能量, 应该和升华热相去不远. 表 1 中的数字, 只是大略的数字. 但确指出分子间的作用能和升华热的关系. 氦是一个例外, 原因是氦原子很轻, 有大幅的零点振动, 它的晶体原子间距离不是在图 1 的极小左右, 而分散到更大的值.

晶体密度和气体的差很多, 但和液体的差不多. 因此, 可以想见熔解热应比升华热小很多, 而蒸发热和升华热差不多. 事实上是如此, 一般的熔解热只有升华热的三十分之一至十分之一间.

4. 熔解, 林德曼公式

固体熔成液体, 是一个非常复杂的现象, 固体的原子作一定的排列. 温度升高, 则原子振动幅度加大, 而且会产生结构上的缺陷. 温度升到熔点, 则整个结构“瓦解”, 成了液体. 熔点温度和压力有关, 可用交界线方程式 (7) 决定. 如果固相体积小于液相, 则加压使熔点上升 (即比较不易熔). 如果固相体积较液相小, 则加压使熔点降低, 冰和水是最重要的例子.

因为固相和液相的密度相差很小, 图 1 的作用能曲线, 对于熔解的了解, 就无甚帮助, 因为熔解前后的 r 都是在 a 附近. 但是, 有人用十分简略的假设, 得到一个熔点和作用能的关系, 有意料不到的好结果. 这是所谓林德曼公式[†]. 我们略提一下.

假定: 晶体中分子的振幅大到某一程度, 晶体即熔解, 振幅和温度有关, 由能量均分定律得

$$\frac{1}{2}K \langle (r - a)^2 \rangle = \frac{1}{2}T. \quad (24)$$

[†]Lindemann 公式, 见固态物理书籍.

K 是作用能曲线在最低点的二次微分. 照假定, 振幅到 αa 时, 晶体熔解, 即

$$Ka^2\alpha^2 = T_0. \quad (25)$$

T_0 为熔点温度. K 是可以直接由弹性系数决定. 立方晶体的杨氏系数 E 和 K 的关系是 $K = Ea$. 因此,

$$T_0 = E \frac{\alpha^2}{nk}. \quad (26)$$

$n = 1/a^3$ 是晶体的密度, k 是热常数 (一向设为 $k = 1$ 者), E, n 都可以直接测定, 现在要凑一个 α 值, 使 (26) 尽量和测得的熔解温度符合. 表 2 是一些结果, (26) 的成绩可以说是惊人. 所取 $\alpha = 1/10$ 指出, 振幅不需要很大, 晶体就熔了.

表 2 林德曼公式计算的熔点 $\alpha^2 = 1/100$

	$T_0(26)$ 式	T_0 实验
铅	400 K	600 K
银	1100	1270
铁	1800	1800
钨	4200	3650
盐	1200	1070
水晶	1900	2000

取自 Tabor(1979) p.246.

因此, 要了解熔解的道理, 必须要了解为什么 (26) 如此成功.

5. μ_1, μ_2 的定义

以上的讨论主要是复习有关气液共存的一些现象, 现在来澄清一点基本概念上的问题, 也就是在 (1) 下提到的问题. 在 (1) 中, μ_1 和 μ_2 的定义并不清楚. 乍看起来, μ_2 是气体的化学势, μ_1 是液体的化学势, 但统计力学的基本假设只容许一个 μ , 也就是温度压强热位能 $G(T, p)$, 除以 N :

$$e^{-G(T,p)/T} = \sum_s e^{-(H(s)+pV(s))/T}, \quad (27)$$

$$\mu = \frac{G}{N}.$$

从 (27) 看来, G 的计算并未提到气相或液相或共存. 在一定的温度压力下, 只有一个 G , 也只有一个 μ , 那里来的两个 μ_1, μ_2 呢? 现在来把以上 μ_1, μ_2 的意义仔细检讨一下.

在一定压力下, (1) 的意思说我们可以找到两条曲线 $\mu_1(T)$ 和 $\mu_2(T)$ 的交点. 由这交点定气液共存的温度. 也就是假定 $\mu_2(T)$ 在 $T(p)$ 以下 (至少在 $T(p)$ 附近) 仍有意义, $\mu_1(T)$ 在 $T(p)$ 以上也仍有意义. 否则谈何交点? 不过 $\mu_1(T)$ 在 $T(p)$ 以上, 及 $\mu_2(T)$ 在 $T(p)$ 以下的意义是什么呢? 从统计力学的基本假设看来, 是无意义的, 因为 (27) 明白的指出, 只有一个 $\mu(T)$, 即一条曲线. 我们应该把 (1) 和图 4 看作图 5 的情形. 即只有一条曲线. 在 $T(p)$, 曲线的斜率为不连续. 图 4 中“不该有”的部分去掉, 即得图 5.

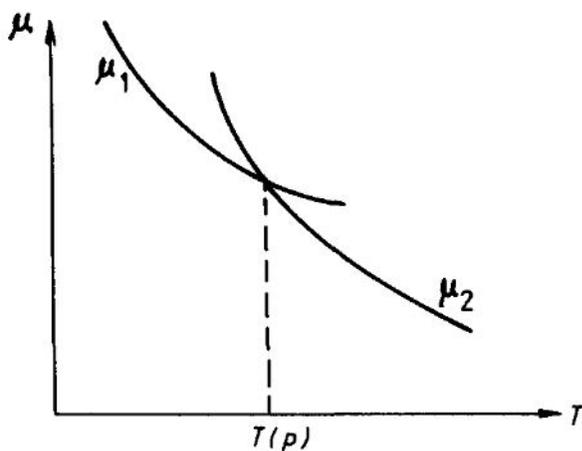


图 4

现在的问题是, (27) 看来是一个 T, p 的连续函数, 也看不出它会导致不连续的斜率, 求和中的每一项, 即每一统计份量

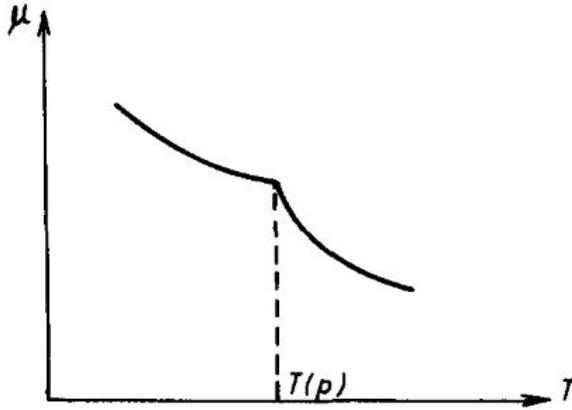


图 5

$$e^{-(H+pV)/T} \quad (28)$$

都是 T 和 p 的平滑, 可微分多次的函数, 不连续的斜率何从来起? 这个数学问题自然是极其重要. 如果 (27) 中的合计中的项数为有限, 则当然不会有任何不连续性发生, 但是分子数 N 为大数, 当此数趋向无限大, 连续性即成问题. 杨、李指出 $N \rightarrow \infty$ 时, 不连续性可能发生, 只要 N 够大, $\mu(T)$ 的曲线的斜率看来就是不连续的了. 此后这方面的文献甚多, 读者可自阅参考资料[‡].

虽然图 5 中的曲线和 $T(p)$ 的意义已是大致清楚, 但图 4 中的 μ_1, μ_2 的延伸部分, 也并非无意义. 在 $T(p)$ 之下, 仍可以有气体存在, 即过冷气体 (即过饱和蒸汽). 在 $T(p)$ 之上, 亦可有液体存在, 即过热液体, 只是这些过冷过热的东西, 是暂稳. 加以扰动, 则气化或凝结. 用来观测高能量带电粒子的“云室”和“气泡室”即是用过冷蒸汽和过热液体作成. 粒子过处, 蒸汽凝成小水珠, 液体呈现小气泡. 我们在第 2 节讨论水珠时, 并未遭遇任何困难, 只是时间长短, 成为一重要问题. 直接用基本假设 (27) 来做计算, 常会失去暂稳态的性质, 而暂稳态往往是代表性的特征.

最后, 我们提一下名词上的问题. 因为, 如图 5, 热位能的一次微分不连续, 所以像气液间的相变被称为“一级相变”. 如果某种相变中, 二次微分不连

[‡]Yang and Lee (1952), Griffiths (1972).

续, 如热容率

$$C = T \frac{\partial S}{\partial T} = -T \frac{\partial^2 G}{\partial T^2}, \quad (29)$$

则这种相变应叫做“二级”。事实上一般所指的二级相变是像由顺磁性到铁磁性, 由氦液到超流氦之类的改变, 这些改变都不该叫做“二级”, 因为在改变发生时的温度 (即临界温度), 热容率是无限大, 只有在某些近似计算下 (平均场计算), C 才是有限而不连续. 这“一级”, “二级”的名称, 不必重视.

讨论问题十五

1. 液体的蒸汽压曲线, 亦大致适合 (9). 实验指出在 99°C 和 101°C 的水蒸汽压是 733.7 及 788.0mm(汞柱高). 试用 (9) 求水的蒸发潜热.

(水的蒸发潜热在 100°C 是每克分子 9717 卡).

2. 实验结果指出, 许多液体的蒸发熵 (蒸发潜热除以温度) 在沸点时 (一大气压下) 都几乎相同. 见表 3, (取自 Zemansky (1957) p.322 Trouton's Rule). 这些结果, 该如何解释?

	T_b	L	ΔS
Ne	27.2	415	15.3
Ar	87.3	1560	17.9
Kr	120	2158	18.0
Xe	165	3021	18.3
Rn	211	3920	18.6
F ₂	85.0	1562	18.4
Cl ₂	239	4878	20.4
HCl	188	3860	20.5
HBr	206	4210	20.4
HI	238	4724	19.9
N ₂	77.3	1330	17.2
O ₂	90.2	1630	18.1
NO	121	3293	27.1
N ₂ O	373	9717	26.0
CS ₂	319	6400	20.0
SO ₂	263	5955	22.6
H ₂ S	213	4463	21.0
N ₂ O	185	3956	21.4
NH ₃	240	5581	23.3
CH ₄	112	1955	17.5
CF ₄	145	3010	20.7
CCl ₄	350	7140	20.4
CH ₂ O	254	5850	23.0
CF ₂ Cl ₂	243	4850	20.0
CHCl ₃	334	6970	20.8
C ₂ N ₂	252	5576	22.1
C ₂ H ₆	185	3517	19.1
C ₆ H ₆	353	7350	20.8
(C ₂ H ₅) ₂ O	307	6470	21.1

	T_b	L	ΔS
Li	1599	30800	19.3
Na	1156	21800	18.8
K	1030	18600	18.1
Rb	985	17100	17.4
Cs	986	15700	15.9
Hg	630	14100	22.4
Ga	2540	62000	24.4
In	2273	51600	22.7
Cd	1040	23900	22.9

T_b = 沸点温度(K), L = 蒸散热 (卡 / 克分子), $\Delta S = L/T$.

3. 由表 3 的结果及 (9), 试证

$$\left(\frac{dT}{T}\right) / \left(\frac{dp}{p}\right) \approx 0.1.$$

dp 为压改变, dT 为 dp 引起的沸点温度的改变.

4. 某液体的沸点在一山顶上为 95°C , 在山下为 105°C , 其潜热为 1000 卡 / 克分子, 求山高.

5. 讨论二维、和一维空间中水珠的消长. (即推广第 2 节的讨论).

6. 在气液共存的状态下, $\partial p / \partial V = 0$, 即总体积可变而压力不改.

第十三章的公式 (13.19) 在这情形下的意义如何? 试分析之.

7. 在临界温度 T_c 以上, 无气体和液体之分. 在 T_c 之上, $\partial p / \partial V < 0$. 在接近 T_c 的温度, $-\partial p / \partial V$ 很小. 试讨论在临界温度附近的中子或光子的散射实验.

第十六章 磁性

从我们的祖宗黄帝到现在, 磁性一直是一个研究和应用的对象. 在近代的统计力学和许多自然科学的领域中, 磁性占了极重要的地位. 磁性的作用, 大略可分两类: 一是外加磁场和个别电子或原子的作用, 二是各电子或原子之间的作用. 如果第二类作用较弱, 问题就比较简单. 一般的顺磁性和电子轨道引起的反磁性, 都属此类. 如果第二类作用很强, 则问题就十分复杂, 因为我们需要考虑许多单元的集体行为. 铁磁性, 反铁磁性等, 都是属于这第二类作用的问题. 当然, 第二类的问题也必须考虑外加磁场. 本章讨论顺磁性的问题和电子轨道的反磁性, 最后复习一下相互作用的成因.

1. 顺磁性

令 $s_1, s_2 \dots s_N$ 为 N 个自旋单元, 每单元来自一电子, 或一原子、一核子等. 令总能量为

$$\begin{aligned} H &= H_0 - \mathbf{h} \cdot \mathbf{M}, \\ \mathbf{M} &= g\mu_B \sum_i \mathbf{s}_i. \end{aligned} \quad (1)$$

$g\mu_B$ 是磁偶极和自旋之比, 如果自旋为电子自旋, 则

$$\begin{aligned} g &= 2, \\ \mu_B &= e/(2mc) = 9.27 \times 10^{-18} \text{ erg/kG} \\ &= 6.72 \times 10^{-2} \text{ K/kG}. \end{aligned} \quad (2)$$

m 是电子质量, $\text{kG} \equiv$ 一千“高斯”. (1) 式中的 H_0 是其他能量, 即所有不含 h 的能量. 现在假定 H_0 不含 s_i .

假定 $s = 1/2$. 每个单元有两个状态, 即顺着 \mathbf{h} 或与 \mathbf{h} 相反, 热位能十分

容易求:

$$\begin{aligned}
 e^{-F/T} &= \sum_s e^{-H/T} \\
 &= \left(e^{h'/T} + e^{-h'/T} \right)^N e^{-F_0/T}, \\
 F &= F_0 - NT \ln \left(2 \cosh \frac{h'}{T} \right).
 \end{aligned} \tag{3}$$

F_0 是 H_0 部份供给的热位能,

$$h' \equiv \frac{1}{2} g \mu_B h. \tag{4}$$

由微分可得熵及其他各量:

$$S = - \left(\frac{\partial F}{\partial T} \right)_h = S_0 + N \left[\ln \left(2 \cosh \frac{h'}{T} \right) - \frac{h'}{T} \tanh \frac{h'}{T} \right], \tag{5}$$

$$M = - \left(\frac{\partial F}{\partial h} \right)_T = \frac{1}{2} g \mu_B N \tanh \frac{h'}{T}, \tag{6}$$

$$E = E_0 - hM, \tag{7}$$

$$F' \equiv E_0 - TS \equiv F + hM. \tag{8}$$

式中 M 是沿着 h 方向的总磁矩, F' 的意义等下就讲. E_0, S_0 是由 H_0 计算而得的能量和熵.

注意此地的热位能 F 是在一定磁场下的热位能, 包括了各自转在磁场内的位能, 即 $-\mathbf{h} \cdot \mathbf{M}$. 磁矩 \mathbf{M} 是看作一变数, 而非不变量. 这一点我们用气体的情形比较一下可说得更清楚. 在第七章, 我们谈热位能定义时指出: “能量” E 是内能, 不包括外加压力引起的位能, 如果包括, 则能量和定温热位能 E 和 F 就改成定压量即

$$E \rightarrow A = E + pV, \tag{9}$$

$$F \rightarrow G = F + pV.$$

pV 是外加压力 p 产生的位能.

热力学第一定律中的“功”, 是专对内能而言, 即

$$\begin{aligned}
 dW &= -pdV, \\
 p &= - \left(\frac{\partial E}{\partial V} \right)_S = - \left(\frac{\partial F}{\partial V} \right)_T.
 \end{aligned} \tag{10}$$

但是在 A, G 中, p 是不变量, 而 V 是运动变量, 其平衡值是

$$V = \left(\frac{\partial A}{\partial p} \right)_S = \left(\frac{\partial G}{\partial p} \right)_T. \quad (11)$$

前面 (1)、(3) 和 (7), 都包括和 h 作用的位能. 因此, E_0 和 F' 相当于 (9) 中的 E, F , 不包括位能 $-hM$, ($-hM$ 相当于 pV):

$$\begin{aligned} dW &= h dM, \\ h &= \left(\frac{\partial E_0}{\partial M} \right)_S = \left(\frac{\partial F'}{\partial M} \right)_T, \end{aligned} \quad (12)$$

而 (10) 的 A, G 相当于 (7), (3) 的 E, F :

$$M = - \left(\frac{\partial E}{\partial h} \right)_S = - \left(\frac{\partial F}{\partial h} \right)_T. \quad (13)$$

F 是定磁场, 定温热位能, 而 F' 是定磁矩定温热位能.

一般说来, 任何磁性物体的热位能 F 可视为 M 和 T 的函数, 而 $F = F' - hM$ 视为 h 和 T 的函数:

$$\begin{aligned} dF' &= -SdT + h dM, \\ dF &= -SdT - Mdh. \end{aligned} \quad (14)$$

这并不限于顺磁性.

现在提一下“绝热去磁”冷却法, 或“核子冷却”法. 这是一个低温实验重要的成就. 所谓冷却, 就是要吸热. 低温实验中, 我们尽量把要观测的物体中的能量吸出来. 也就是要把它环境的温度降低. 所谓低温环境, 就是很会吸热的环境. 要吸热必须要很容易动. 在低温, 一般物质都成了固体, 振动在 10^{-2} K 以下可以说是停止了. 唯一还可以动的只有核子的自旋, 它们之间的相互作用能非常之小*. 在 1 K 时它们还可以自由翻身. “绝热去磁”法, 就是用这种运动来吸热, 这方法的理论如下. 图 1 是熵在不同磁场下的曲线 (5). (S_0 不计). 如果 $h = 0$, 则 $S/N = \ln 2$. 图中 $h_2 < h_1$, 磁场越高, S 越低, 因为磁场限制了自旋的活动范围.

冷却的步骤如下: 先加高磁场, 使自旋排列起来, (见图 1, 由 A 到 B), 然后逐渐把磁场减小, 同时保持绝热. 绝热过程中熵不变, 因此, 磁场减小, 温度也降低了, 是为绝热去磁的冷却法.

*一对核子自旋相距 10^{-8} cm 时的磁偶极作用能大约为 10^{-6} K.

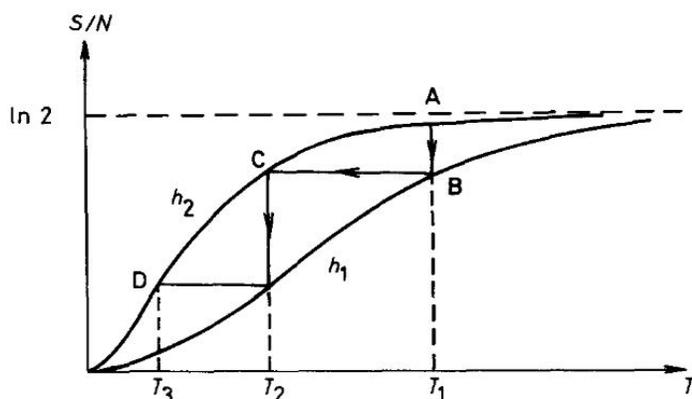


图 1

注意, 当磁场降低时, 总磁矩 M 并没有变. 熵只是 M 的函数, (5) 可以写成

$$S = \frac{1}{2}N \left[(1+m)\ln\frac{2}{1+m} + (1-m)\ln\frac{2}{1-m} \right], \quad (15)$$

$$m \equiv M/N.$$

熵没变, M 也没变. 磁场低, 则自旋容易翻身, 容易翻身的意思就是低温, 因为能量容易进来, 也就是容易吸取体外的热.

图 1 中的过程 $A \rightarrow B \rightarrow C$ 可以再往下重复, 到 D , 或再重复. 当然可以用其他不同大小的磁场.

以上的理论很简单, 做起实验却又是一回事. 平衡须要够长的时间来定义. 在观测时间内, 每个自旋必须有很多次的变动. 自旋的变动要靠它和其他自旋和环境的作用. 我们一方面要求这些作用强, 使自旋不断地改变, 以达成平衡, 缩短观测时间, 一方面又要这些作用弱, 使作用能不影响到自旋的自由. 真是“又要马儿好, 又要马儿不吃草”, 实验者面临的难题, 可以想见. 约在十五年前, 有人用金属中的核子自旋来做成这种实验, 达到低于 10^{-4} K 的温度[†]. 用金属的原因是自由电子在极低的温度下仍能活动. 在低温, 可以活动的电子粒子数是和 T 成正比, 即热容率和 T 成正比, 虽然这很小, 但比起晶体振动的 T^3 热容率要强多了. 核子自旋和电子有一点点作用. 靠电子的来往疏通, 核子自旋可达平衡. 观测时间往往要一星期.

[†]Osgood and Gookind(1967), Dundon and Gookind (1974).

2. 磁化率

“磁化率”或“磁感应率”的定义是

$$\chi = \frac{\partial M}{\partial h}, \quad (16)$$

一般都指 $h \rightarrow 0$ 时的情形, 即 $\chi = M/h$, 上一节的模型的磁化率由 (6) 可知:

$$\chi = \frac{1}{4} \frac{N (g\mu_B)^2}{T} \quad (17)$$

这个和温度成反比的结果是很普遍的现象, 一般称作“居里定律”[‡]. 在第十三章 (见 (13.6)), 我们指出

$$\chi = \frac{1}{T} \langle M^2 \rangle_c. \quad (18)$$

如果各自旋为独立, 则

$$\begin{aligned} \langle M^2 \rangle_c &= \frac{1}{3} \left\langle \left(\sum_i s_i \right)^2 \right\rangle_c (\mu_B g)^2 \\ &= \frac{1}{3} N s(s+1) (\mu_B g)^2. \end{aligned} \quad (19)$$

式中的 $1/3$ 是因为 M 的意思是沿 \mathbf{h} 方向的磁矩, 即 $\langle M^2 \rangle = \frac{1}{3} \langle \mathbf{M} \cdot \mathbf{M} \rangle$. 如果 $s = \frac{1}{2}$, 则得 (17). 注意, (18) 假设了 $[\mathbf{M}, H] = 0$, 即 \mathbf{M} 为守恒量, 或略去量子效应.

现在来看一个费米粒子群体的磁化率, 例如金属电子, 或 He^3 . 我们只考虑电子自旋, 不考虑其轨道引起的磁性.

在 $T = 0$ 时, 粒子位子在顶层能量 ϵ_F 以下客满, 层外全空. 每一动量有两位子, 自旋向上或向下. 因此, 在 $T = 0$ 时, 自旋全被锁住了, 一半向下, 一半向上. 在 $T > 0$ 时, 顶层外有些外层粒子, 内有些空子. 这些粒子, 空子的自旋可以活动, 可以造成磁感态. 这些粒子、空子的粒子数是和 T 成正比, 因此, χ 在低温时是一常数, 因为 (18) 分母也是 T . 要计算这常数很简单, 令 N_{\pm} 为自旋朝上朝下的粒子数, 则

$$\begin{aligned} M &= \frac{1}{2} g\mu_B (N_+ - N_-), \\ \langle M^2 \rangle_c &= \frac{1}{4} (g\mu_B)^2 [\langle N_+^2 \rangle_c + \langle N_-^2 \rangle_c], \end{aligned} \quad (20)$$

[‡]居里夫人之夫所发现.

$\langle N_+N_- \rangle_c = 0$, 因为 N_+ , N_- 不相干.

现在再用第十三章的关系 (见 (13.17)):

$$\langle N_{\pm}^2 \rangle_c = T \frac{\partial N_{\pm}}{\partial \mu} = \frac{1}{2} T \frac{\partial N}{\partial \mu} \quad (21)$$

在低温 $\partial N / \partial \mu$ 即是能位密度 (见 (4.14)):

$$Vg(0) = V \frac{mp_F}{\pi^2}, \quad (22)$$

也就是说, 顶层能量, 即 μ , 增加 $d\mu$, 占满的位子数就增加 $dN = Vg(0)d\mu$. 由 (18)、(20)、(21) 得

$$\chi = \frac{1}{4} (g\mu_B)^2 \frac{\partial N}{\partial \mu} = \frac{V}{4} (g\mu_B)^2 \frac{mp_F}{\pi^2}. \quad (23)$$

注意: 在高温, $\langle N_{\pm}^2 \rangle_c = N_{\pm}$, 则 χ 又回到 (17). (23) 是所谓泡利[§]磁化率.

实验测得 He^3 的磁化率远比 (23) 的大 (He^3 的 $g\mu_B$ 必须用 He^3 核子的值). 原因是原子间的作用. He^3 液体实在不是理想气体. 由于互斥, $\langle N_+N_- \rangle_c < 0$, 因而把 $\langle M^2 \rangle_c$ 加大不少.

3. 带电粒子的反磁性

带电粒子在磁场中会走圆圈轨道. 这圆圈会产生和外加磁场反向的磁矩, 造成反磁性. 但这样的看法并不可靠, 图 2 示一些轨道, 磁场向下 (假定电荷为正), 圆圈轨道的磁矩向上, 现在我们在气体中任取一区域 (图 2 的虚线内). 完全在区内的轨道造成向上的磁矩, 但靠边的轨道却造成反时针的电流, 形成向下的磁矩.

[§]发现独占定律的那位泡利, Pauli.

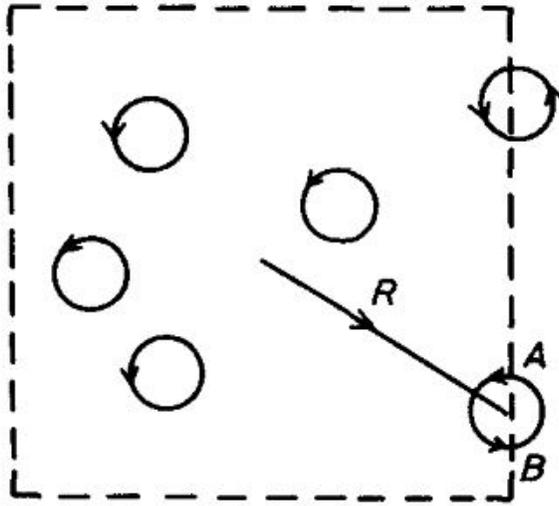


图 2

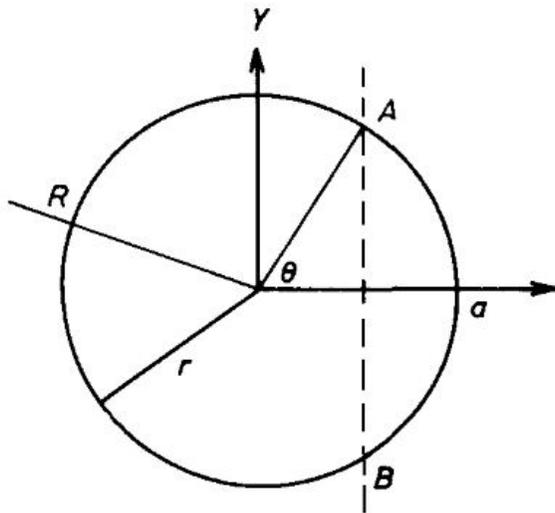


图 3

我们可以证明在区内的总磁矩为零, 如下: 一个轨道的磁矩是和

$$\sigma \equiv \frac{1}{2} \oint \mathbf{r} \times d\mathbf{r} \quad (24)$$

成正比. 积分是沿着轨道做, 如果轨道是一整圈, 则 σ 是圈内面积. 现在假定磁场指向 $-z$ 方向, 轨道都在 x, y 平面内, z 方向的运动略去. 如果一轨道在

图 2 区内, 则

$$\sigma = \pi a^2. \quad (25)$$

σ 方向指向 z , a 是圈圈的半径. 如果轨道有一部分在区外, 就比较麻烦, 令 \mathbf{R} 为圆心的位置, 则

$$\begin{aligned} \mathbf{r} &\equiv \mathbf{R} + \mathbf{r}', \\ \boldsymbol{\sigma} &= \frac{1}{2}\mathbf{R} \times \int d\mathbf{r}' + \frac{1}{2} \int \mathbf{r}' \times d\mathbf{r}'. \end{aligned} \quad (26)$$

图 3 是图 2 中右边的轨道的放大图. (26) 右第一项的积分是

$$\int d\mathbf{r}' = \text{向量 } AB = -2a \sin \theta \hat{\mathbf{y}}. \quad (27)$$

$\hat{\mathbf{y}}$ 为 Y 方向的单位向量. 凡是中心在距边界 $\pm a$ 之内的轨道, 都不全在区内, 这样的轨道在 dR 长的一段边界内有

$$n(2adR) \quad (28)$$

个. n 是粒子密度 (粒子数/面积), 因此, 这些靠边的轨道的总磁矩是, 由 (27), (28),

$$\begin{aligned} 2andR \frac{1}{2} \mathbf{R} \times (-2a \langle \sin \theta \rangle) \hat{\mathbf{y}} \\ = -\pi a^2 \frac{n}{2} \mathbf{R} \times d\mathbf{R}, \end{aligned} \quad (29)$$

$$\langle \sin \theta \rangle = \frac{1}{2} \int_0^\pi \sin \theta d(\cos \theta) = \frac{\pi}{4}. \quad (30)$$

式中的 $d\mathbf{R}$ 是沿着边界反时针方向. (26) 的最后一项略去了, 因为它最后供给的磁矩只是和边界长成正比, 而 (29) 是和面积成正比. 将 (29) 沿边界积分, 得

$$-\pi a^2 \frac{n}{2} \oint \mathbf{R} \times d\mathbf{R} = -\pi a^2 n \mathbf{A}. \quad (31)$$

\mathbf{A} 是区域面积方向朝 z . (31) 的大小正好与在区内完整轨道的总磁矩相同, 不过方向相反. 因此, 整个区内的磁矩为零.

以上的分析足以证明轨道不能产生总磁矩, 这是古典力学的结果. 其实, 磁场只能改变粒子运动方向, 不能改变其能量. 热位能只和能量有关, 因此, 不是磁场的函数. 既非磁场的函数, 微分必为零, 即总磁矩为零.

$$M = -\frac{\partial F}{\partial h} = 0. \quad (32)$$

如果把量子力学考虑进去, 则情形又稍有不同, 量子力学计算的结果是: 粒子位子的能量分布受到了磁场的影响, 因此, 热位能也受影响, 磁矩因而不是零. 这个结果是非常的重要[¶]. 我们只略提一下重点. 为求简明, 只考虑 xy 平面运动, 磁场在 z 方向.

由量子力学的波动方程式可解得单粒子的能量:

$$\begin{aligned}\epsilon_n &= \left(n + \frac{1}{2}\right)\omega, \quad n = 0, 1, 2, 3, \dots \\ \omega &\equiv \frac{eh}{mc},\end{aligned}\tag{33}$$

h 是磁场 ($\hbar = 1$), c 是光速, ω 是粒子的转动频率. 因此, 本来 $p^2/2m$ 的能量, 成了简谐振动的能量, 粒子位子的能量分布成了

$$L^2 g(\epsilon) = \sum_n W_n \delta(\epsilon - \epsilon_n)\tag{34}$$

L^2 是物体的面积, W_n 是能量 ϵ_n 的统计份量, 在没有磁场时 ϵ_n 和 ϵ_{n+1} 间的位子数是

$$L^2 \int_{\epsilon_n}^{\epsilon_{n+1}} d\epsilon g_0(\epsilon) \approx L^2 \omega g_0(\epsilon_n),\tag{35}$$

$g_0(\epsilon)$ 是在无磁场时的能位密度. ϵ 从 0 算起. 以上假设了 $g_0(\epsilon)$ 在这段能量中大致不变, 所以

$$g(\epsilon) \approx g_0(\epsilon) \omega \sum_{n=0}^{\infty} \delta\left(\epsilon - \left(n + \frac{1}{2}\right)\omega\right).\tag{36}$$

本来在 ϵ_n 和 ϵ_{n+1} 之间的位子, 全挤到同一能量了.

有了 $g(\epsilon)$, 热位能和其他平衡性质就可以计算了. 自由电子的热位能是

$$\frac{1}{L^2} \Omega = -T \int_0^{\infty} d\epsilon g(\epsilon) \ln(1 + e^{-(\epsilon - \mu)/T}).\tag{37}$$

注意, $g(\epsilon)$ 的积分和 $g_0(\epsilon)$ 的积分只差在: 一个是分立点的和, 一个是连续积分, 如果 ω 很小, 则所差有限. 图 4 指出积分与和的不同. 任一积分可写成

$$\int_0^{\infty} R(\epsilon) d\epsilon \approx \sum_{n=0}^{\infty} R\left(\left(n + \frac{1}{2}\right)\omega\right) - \frac{\omega^2}{24} \left(\frac{dR}{d\epsilon}\right)_{\epsilon=0} + O(\omega^3).\tag{38}$$

[¶]兰道 (Landau) 反磁性理论, 见任何固态物理教科书.

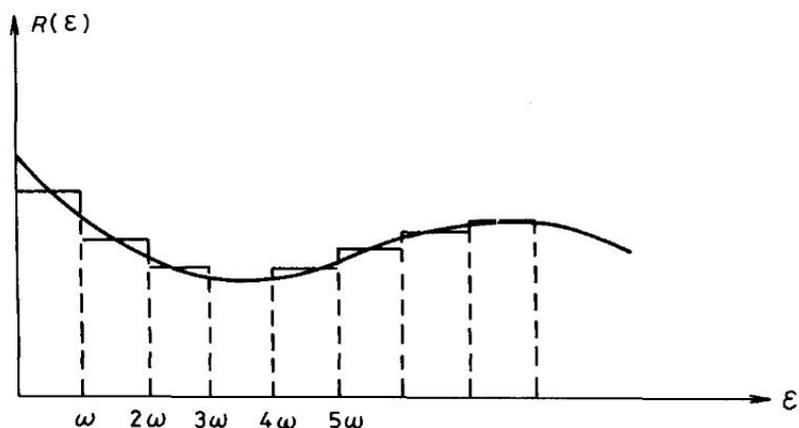


图 4

只要 $R(\epsilon)$ 相当平滑, 且积分为有限. (38) 的和是图 4 中诸方块的面积, 方块的高是 $R((n + \frac{1}{2})\omega)$, 读者可自导出 (38). 因此, 只有在端点 $\epsilon = 0$ 的值和 ω 有关. 所以, (37) 是

$$\begin{aligned} \frac{1}{L^2}\Omega - \frac{1}{L^2}\Omega_0 &= \frac{-\omega^2}{24} T \frac{d}{d\epsilon} [g_0(\epsilon) \ln(1 + e^{-(\epsilon-\mu)/T})]_{\epsilon=0} \\ &= \frac{\omega^2 m}{24\pi} \frac{e^{\mu/T}}{1 + e^{\mu/T}} \end{aligned} \quad (39)$$

注意, $g_0(\epsilon) = m/\pi$ 是个常数. Ω_0 是 $\omega = 0$ 时的热位能.

如果气体密度很低, 则 $e^{\mu/T}$ 很小, (39) 约是

$$\frac{1}{L^2}(\Omega - \Omega_0) \approx \frac{\omega^2 m}{24\pi} e^{\mu/T}. \quad (40)$$

理想气体的密度是

$$n = 2e^{\mu/T} \int \frac{d^2 p}{(2\pi)^2} e^{-p^2/2mT} = \frac{mT}{\pi} e^{\mu/T}. \quad (41)$$

由 (40), (41) 得总磁矩

$$\begin{aligned} M &= -\frac{\partial \Omega}{\partial h} = -N \frac{\mu_B^2 h}{3T} \\ \omega &= 2\mu_B h. \end{aligned} \quad (42)$$

因此, 粒子的轨道仍是有一点反磁性. 但这纯是量子效应. 这结果也合用于三维空间, 因为理想气体的 z 方向运动是独立的.

如果气体密度很高且 $\mu = \epsilon_F \gg T$, 则 (39) 成了

$$\frac{1}{L^2} (\Omega - \Omega_0) \approx \frac{(\mu_B h)^2 m}{6\pi},$$

$$M \approx -N \frac{\mu_B^2 h}{3\epsilon_F}. \quad (43)$$

这结果只合用于二维空间, 因为在三维空间, 高密度的费米气体的各方向运动不是独立的.

在高密度时, 如果 T 很小, 即 $\omega/T \gg 1$, 显然 (39), (43) 是一定不对的, 因为积分函数不能看成光滑曲线. 在 $\epsilon = \mu$ 附近, 在低温时总是有快速改变发生, 例如图 5 的 $R(\epsilon)$. 方块面积, 和曲线下的面积可以大不相同, 当 ω 增大时, 5ω 会向右移, (5, 6) 间的方块面积减少. 要待 (4, 5) 间的方块移过来才增大, 但再移又减少. 因而成了一种周期性的改变. 温度愈低, 这种改变越明显. 注意, 每当 $(n + \frac{1}{2})\omega$ 经过 μ , 方块面积就大改一次, 即在

$$n + \frac{1}{2} = \frac{\mu}{\omega}. \quad (44)$$

时. 因此, 变化是 $1/\omega$ 的周期性函数, 这是在固态物理中的一非常重要的研究题目, 在此不多谈. 注意, ω 的展开式, 显然不成立, 无论 ω 多小, 这种周期性的变化都存在. ω 的级数是无法形容这样的变化. 因此, 我们再回头来检讨一下 (38). 这是一个数学问题.

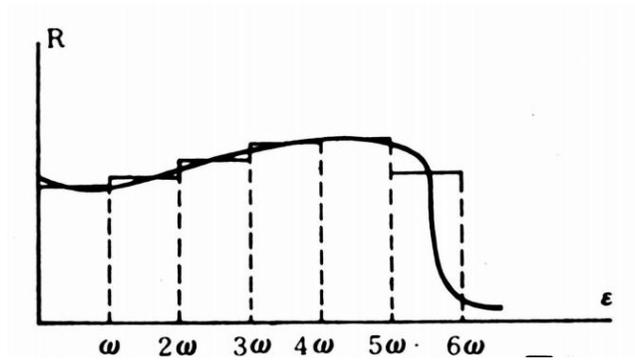


图 5

一个函数 $R(\epsilon)$ 的积分和图 4 中方块面积到底差多少?(38) 是直接了当的对 ω 展开, 但以上 (44) 的结果稍令人不安. 现在把 (36) 写成

$$\frac{g(\epsilon)}{g_0(\epsilon)} = \sum_{n=0}^{\infty} \delta\left(\frac{\epsilon}{\omega} - \frac{1}{2} - n\right) \equiv q\left(\frac{\epsilon}{\omega} - \frac{1}{2}\right), \quad (45)$$

$$q(x) \equiv \sum_{k=-\infty}^{\infty} e^{2\pi i k x}. \quad (46)$$

(46) 是周期为 1 的 δ -函数, 我们只管 $\epsilon > 0$. 因此, 任一函数之和可以写成

$$\begin{aligned} \omega \sum_{n=0}^{\infty} R\left(\left(n + \frac{1}{2}\right)\omega\right) &= \omega \int_0^{\infty} dx R(x\omega) q\left(x - \frac{1}{2}\right) \\ &= \sum_{k=-\infty}^{\infty} \omega \int_0^{\infty} dx R(x\omega) e^{2\pi i k\left(x - \frac{1}{2}\right)} \\ &= \sum_{k=-\infty}^{\infty} R_k (-1)^k, \end{aligned} \quad (47)$$

$$R_k \equiv \int_0^{\infty} d\epsilon R(\epsilon) e^{2\pi i k \epsilon / \omega}. \quad (48)$$

$R(\epsilon)$ 的积分即是 R_0 , 因此,

$$\begin{aligned} \omega \sum_{n=0}^{\infty} R\left(\left(n + \frac{1}{2}\right)\omega\right) - \int_0^{\infty} d\epsilon R(\epsilon) \\ = \sum_{k=1}^{\infty} (R_k + R_{-k}) (-1)^k \end{aligned} \quad (49)$$

此所谓“泊松求和公式”之一例. 如果这 k 级数收敛得快, 则非常有用. 看两个例子:

【例】 $R(\epsilon) = e^{-\epsilon/T}$, 则

$$R_k + R_{-k} = \frac{2\omega^2/T}{(\omega/T)^2 + (2\pi k)^2}, \quad (50)$$

无甚奇特之处, 级数收敛.

【例】 $R(\epsilon) = e^{-\epsilon^2/2\sigma^2}/\sqrt{2\pi}\sigma$, 则

$$R_k + R_{-k} = e^{-\frac{1}{2}(2\pi\mu\sigma)^2/\omega^2}. \quad (51)$$

这式子就很奇怪了, 它不能用 ω 的级数展开, 在 $\omega = 0$ 时, 每一次微分都是零, 如用 (38), 则得零. 但 (49) 的 k 级数却收敛得很快. 这例子给我们一个教训: 有些看来天衣无缝的展开式, 却是靠不住的.

注意, $R(\epsilon)$ 如果在 $\epsilon = \mu$ 有任何不连续性, 或快速改变, (48) 立即会有

$$e^{2\pi i k \mu / \omega}$$

出现, 即有 $1/\omega$ 的周期函数出现.

4. 自旋的相互作用

现在复习一下自旋间的相互作用, 最重要的是短距离的“交换作用”和长程的磁偶极作用. 在固体中, 由于晶体结构、杂质等因素, 各式作用名目繁多. 这些作用可以引起磁性的千变万化. 在此主要讨论交换作用.

(甲) 短距离的交换作用

这是电子自旋间最强的作用. 大致说来, 它是电子间的静电斥力, 和电子的不合群性所引起, 不合群性即是波函数的反对称性. 设想二电子, 其波函数可作

$$\phi(1, 2) = \phi(\mathbf{r}_1, \mathbf{r}_2) \chi(s_1, s_2). \quad (52)$$

$\mathbf{r}_1, \mathbf{r}_2$ 为电子位置, s_1, s_2 为自旋方向. 电子为不合群, 所以

$$\phi(1, 2) = -\phi(2, 1). \quad (53)$$

现在已知: 如自旋同向, 则 $\chi(s_1, s_2) = \chi(s_2, s_1)$, 如反向则 $\chi(s_1, s_2) = -\chi(s_2, s_1)$, 因此,

$$\begin{aligned} \text{同向则: } \quad \phi(\mathbf{r}_1, \mathbf{r}_2) &= -\phi(\mathbf{r}_2, \mathbf{r}_1), \\ \text{反向则: } \quad \phi(\mathbf{r}_1, \mathbf{r}_2) &= \phi(\mathbf{r}_2, \mathbf{r}_1). \end{aligned} \quad (54)$$

所以, 在 $\mathbf{r}_1 = \mathbf{r}_2$ 时, 同向的 ϕ 必须是 0, 而反向的 ϕ 不必. $\mathbf{r}_1 = \mathbf{r}_2$ 是指二电子在一起的情形, 因此, 自旋同向则二电子不常接近, 不常接近则静电斥力引起的位能小. 因此, 自旋同向时的能量较低. 自旋反向时, 电子比较相近, 故电能大, 以上这堆话的结论可以用下式表示

$$\text{有效作用能} = -J \mathbf{s}_1 \cdot \mathbf{s}_2, \quad (55)$$

J 是一个常数. 我们来做一个十分粗略的计算: 假定两个氢原子核固定在两点 O 和 \mathbf{R} , 令

$$\begin{aligned} \phi_1(\mathbf{r}) &= \phi(\mathbf{r}), \\ \phi_2(\mathbf{r}) &= \phi(\mathbf{r} - \mathbf{R}). \end{aligned} \quad (56)$$

$\phi(\mathbf{r})$ 为氢原子的 $1s$ 波函数, 略去电子间的静电作用, 则波函数为 (52)

$$\phi_{\pm}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} (\phi_1(\mathbf{r}_1) \phi_2(\mathbf{r}_2) \pm \phi_1(\mathbf{r}_2) \phi_2(\mathbf{r}_1)), \quad (57)$$

自旋同向则用 ϕ_- , 反向则用 ϕ_+ . 在这近似计算下, ϕ_+ , ϕ_- 的能量相同. 现在加上电子间的静电作用

$$u(\mathbf{r}_1 - \mathbf{r}_2) = \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}. \quad (58)$$

用一次近似计算, 得

$$E_{\pm} = \langle \phi_{\pm} | u | \phi_{\pm} \rangle = \int d^3\mathbf{r}_1 d^3\mathbf{r}_2 |\phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2)|^2 \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \pm \frac{J}{2}, \quad (59)$$

$$\frac{J}{2} \equiv \int d^3\mathbf{r}_1 d^3\mathbf{r}_2 \phi_1(\mathbf{r}_2)\phi_2(\mathbf{r}_1)\phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2) \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}.$$

ϕ_+ 和 ϕ_- 的能量之差为 $E_+ - E_- = J$, 也就是说自旋同向的能量较反向低了 J , (55) 的含义就是如此. 注意

$$\begin{aligned} \mathbf{s}_1 \cdot \mathbf{s}_2 &= \frac{1}{4} \quad \text{如果 } \mathbf{s}_1, \mathbf{s}_2 \text{ 同向,} \\ &= -\frac{3}{4} \quad \text{如果反向.} \end{aligned} \quad (60)$$

J 是一个短程作用, 因为 $\phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_1)$ 要 ϕ_1, ϕ_2 在 \mathbf{r}_1 同时不等于 0 才不等于 0. 也就是要靠 ϕ_1, ϕ_2 的重迭. 但 ϕ_1 中心在 $\mathbf{r} = 0$, ϕ_2 中心在 $\mathbf{r} = \mathbf{R}$

$$\begin{aligned} \phi_1(\mathbf{r}) &\sim \frac{1}{r} e^{-r/a}, \\ \phi_2(\mathbf{r}) &\sim \frac{1}{|\mathbf{r} - \mathbf{R}|} e^{-|\mathbf{r} - \mathbf{R}|/a}. \end{aligned} \quad (61)$$

a 是氢原子半径, 因此, 如果 $R \gg a$, 则 $J \rightarrow 0$. 如果 $R \leq a$ 左右, 则 J 很强.

以上粗略的分析, 大致指出 (55) 是怎么来的. (55) 的含意是两个电子状态的能差. 奇怪的是, J 和电子的磁性, 竟是毫无关系, 自旋的介入, 是由于波函数的交换对称性, 而能差的来源是电子间的静电作用. 整个物体对外加磁场的反应, 可以说是被 J 左右. 因此, 虽然每个电子是一个磁偶极, 但左右磁性的不是这个磁偶极的磁性.

自旋间有效作用能的计算, 当然不是如此简单, 连 (55) 的基本概念都不是完全清楚. 我们的目的是要把电子自旋的运动拿出来讨论, 把所有其他的运动的总效果放在一个有效作用强度 J 里面. 这 J 的责任可是太大了一点. 如果有三个自旋, 有效作用是否可以写成

$$-J_{12}\mathbf{s}_1 \cdot \mathbf{s}_2 - J_{23}\mathbf{s}_2 \cdot \mathbf{s}_3 - J_{31}\mathbf{s}_3 \cdot \mathbf{s}_1? \quad (62)$$

答案是否定的, 没这么好事^{||}. 唯一的希望是 (62) 是一个近似. 当然, 最保险的做法是把所有电子的运动直接解了, 自旋问题自然解决, 用不着谈有效作用能, 但这做法太困难. 如果我们只对自旋有兴趣, 不计较其他, 则由有效作用着手, 仍是比较方便的途径.

目前磁学中最常用的模型之一就是 (55) 为准, 把自旋总能量写成

$$H = -\frac{1}{2} \sum_{i,j} J_{ij} \mathbf{s}_i \cdot \mathbf{s}_j - \mathbf{h} \cdot \sum_i \mathbf{m}_i, \quad (63)$$

$$\mathbf{m}_i = g\mu_e \mathbf{s}_i.$$

(63) 中的 J_{ij} 是自旋 i, j 间的“交换作用”能量. (63) 即所谓“向量模型”, 或“海森堡”模型. 用量子力学分析, 各 \mathbf{s}_i 是算符, H 也是. 我们假定 (63) 为正确, 求得一些结果后, 再和实验比较, 而后定 J 的值以求和实验符合, 或者用计算求 J , 再用以推测实验. 总之, (63) 是一个分析的工具. 对某些情形 (63) 可能很适用. 对其他可能需要修正, 譬如说加一些别的项, 如

$$J_{ijk} \mathbf{s}_i \cdot \mathbf{s}_j \times \mathbf{s}_k \quad (64)$$

之类. \mathbf{s}_i 不一定是一个电子的自旋, 而可以是一个原子的总自旋. J_{ij} 也不一定是正值.

(乙) 长程磁偶极作用.

两个磁偶极 $\mathbf{m}_1, \mathbf{m}_2$ 相距 \mathbf{r} , 则作用能是

$$\frac{\mathbf{m}_1 \cdot \mathbf{m}_2}{r^3} - \frac{3(\mathbf{m}_1 \cdot \mathbf{r})(\mathbf{m}_2 \cdot \mathbf{r})}{r^5} \quad (65)$$

这个作用能, 虽然比交换作用要弱许多, 但它是长程作用. 如果诸自旋有集体排列现象, 则这作用就不可忽略了.

相互作用的后果, 是非常的不容易分析. 无论是自旋间的作用, 或是分子间的引力和斥力, 只要一牵涉到许多单元的相干问题, 我们的普通分析能力就不够用了. 我们尽量把模型简化, 但即使最简单的模型也是十分困难的题目, 下一章我们介绍易形模型. 目的之一是对磁性作一粗略的了解.

讨论问题十六

1. 体积为 V 的容器, 内有 N 个粒子的稀薄气体, 器壁上有 $N'(N' \ll N)$ 个单元, 每个单元有二种状态, 能量为 $\pm h$, h 正比于外加磁场, 起始温

^{||}有关 J 的理论, 可参考 Mattis(1965).

度为 T . 假设气体分子和单元之间有能量交流.

(A) 将磁场慢慢地关掉, 求温度的改变 ΔT , 假设没有热能损失.

(B) 若磁场突然关掉, 求 ΔT .

2. 压力冷却法

将液体的 He^3 绝热加压, 使成固体, 则温度下降. 这冷却法原理由彭末软恰 (Pomeranchuk) 指出, 故亦常以他名字称之. 以下为说明原理的步骤, 读者试讨论之.

(A) 求低温 He^3 的熵, 用一费米气体模型. 顶层能量约为 5 K.

(B) 固体的 He^3 的熵在低温几乎全来自 He^3 的自旋. 在 10^{-3} K 以下, He^3 呈反铁磁性, 各自旋有冻结现象.

把液相的熵和固相的熵画在同一图上, 同为 T 的函数.

(C) 由以上简图解释加压冷却的原理.

3. 磁化率和自旋的相干函数有关. (19) 是自旋互不相干的结果. 假想一模型, 有 N 个自旋 $1/2$, $\mathbf{s}_i, i = 1, 2, \dots, N$.

(A) 如果 $\mathbf{s}_i, \mathbf{s}_{i+1}, i = 1, 3, 5, \dots$ 总是平行的, 而 $\mathbf{s}_i, \mathbf{s}_{i+2}$ 不相干, 求磁化率.

(B) 如果 \mathbf{s}_i 和 $\mathbf{s}_{i+1}, i = 1, 3, 5, \dots$ 总是方向相反, 则磁化率如何?

讨论金属电子自旋的相干情形, 并和“反铁磁性”物体比较. 反铁磁物体的自旋是作周期性的排列. 例如在立方晶格上, 每自旋和它六个近邻自旋方向相反.

这一题和第 2 题当然是有关系的.

4. 将热位能 (37) 直接写成

$$\Omega = -T \sum_n W_n \ln (1 + e^{-(\epsilon_n - \mu)/T}). \quad (66)$$

然后对 h 微分, 得

$$\begin{aligned}
 M &= -\frac{\partial \Omega}{\partial h} \\
 &= \sum_n W_n \left(-\frac{\partial \epsilon_n}{\partial h} \right) f_n + T \sum_n \frac{\partial W_n}{\partial h} \ln (1 + e^{-(\epsilon_n - \mu)/T}) \quad (67) \\
 f_n &\equiv \frac{1}{e^{(\epsilon_n - \mu)/T} + 1}.
 \end{aligned}$$

(67) 右第一项可解释成各电子的圆形轨道为磁矩, (把每一个 n 看成一轨道). 第二项该如何解释? 先复习 (54) 至 (31) 的讨论.

5. 一个电子在磁场中的哈密顿函数可以写成

$$\begin{aligned}
 H &= \frac{1}{2m} \left(\mathbf{p} + \frac{e}{c} \mathbf{A} \right)^2 \\
 &= \frac{1}{2m} [p_x^2 + (p_y + m\omega x)^2]. \quad (68)
 \end{aligned}$$

此地的磁场 h 是和 (x, y) 平面垂直. 我们略去 Z 方向的运动. $A = hx$, 指向 y 方向, $\omega = eh/mc$.

(A) 由古典力学方程式, 试证电子轨道是一圆圈. 圆心在 (x_0, y_0) , 且

$$p_y = -m\omega x_0. \quad (69)$$

是一不变量.

(B) 令 $x - x_0 = x'$, 则 (67) 可写成

$$H = \frac{1}{2m} (p_x^2 + m^2\omega^2 x'^2). \quad (70)$$

从量子力学来看, $[H, p_y] = 0$, 故 p_y 为一不变量, (70) 是一简谐振动的哈密顿算符, 其本征值为 (33). 读者试求 $g(\epsilon)$. 本题大致是“兰道反磁性论”的主要内容.

第十七章 易形模型

最简单的变量有二值 (如果只有一值, 就不可能变), 这种变量可称为“易形”变数. 这二值可指定为 $+1$ 和 -1 , 一如易经的阴、阳. 最简单的群体模型是一群易形变量, 这样的模型我们统称之为易形模型. 这类模型, 在近代统计力学中用途非常之广. 虽然和真实物体相去甚远, 但它却能描述许多现象的大致原理. 古人作易经就是企图用最简单的模型来分析一些自然现象. 虽然我们在此不谈易经但精神和目的是一样; 看“万物”的现象如何用“阴阳动”来解释. 在这一章, 我们介绍最简单的易形磁铁模型, 一维空间的解, 及二维空间的铁磁性证明. 再提一下其他内容相同的易形模型. 本章提到的, 都是有关相变现象的模型. 易形模型在这一方面最受注目, 但其用途不止于相变.

1. 易形铁磁, 一维空间

近代物理中的易形模型是由磁铁模型开始. 令 $s_r (r = 1, 2, \dots, N)$ 为 N 个易形变量, 其值 ± 1 代表两个磁针方向. 每个 s_r 是一个运动单元. 它们的总能量是

$$H(s) = -J \sum_{r=1}^{N-1} s_r s_{r+1} - h \sum_{r=1}^N s_r. \quad (1)$$

这是一个一维空间的磁铁模型. 相邻的单元, (或称“自旋”), 有相互作用. h 是“外加磁场”^{*}.

这模型没有任何有关运动的内容. 也就是说, 它没有指出 s_r 如何因时而异. 不像气体模型和量子力学模型, 能量函数或算符可以决定运动方程式. 此地 s_r 的运动须要另外指定. 现在我们只假设各 s_r 不断地改变, 使基本假设成立.

这模型的平衡性质可由定温法则求出. 我们先看 $h = 0$ 时之解.

^{*}这是德国人 Ising 的论文题目. 其后凡是此模型的推广多被称为“Ising 模型”.

如果温度 $T = 0$, 则所有 s_r 都为 $+1$ 或都为 -1 , 使 H 为极小, 其值为

$$E(T = 0) = -J(N - 1). \quad (2)$$

如果 $T > 0$, 则有些 s_r 为 $+1$ 有些为 -1 . $+1$ 和 -1 分界处做“分界点”(见图 1). 在 $T = 0$ 时, 没有分界点. 在低温时, 分界点很少. 每个分界点的能量是 $2J$. 这模型成了一个分界点气体的模型. 分界点粒子数非守恒量, 故其化学势为 0. 每个 r 和 $r + 1$ 之间只能有一个分界点, 所以我们可以用费米粒子的结果:

$$\begin{aligned} N'/N &= \frac{1}{e^{2J/T} + 1}, \\ \frac{1}{N}[E(T) - E(0)] &= \frac{2J}{e^{2J/T} + 1}. \end{aligned} \quad (3)$$

N' 是分界点的粒子数. 以上假设 $N, N' \gg 1$. 读者可自计算其他量.

因为 $N'/N \neq 0$. (除非 $T = 0$), 所以 s_r 没有作单方向的排列, 总磁矩 M 因而是 0:

$$\frac{M}{N} = \frac{1}{N} \sum_{r=1}^N s_r = 0 \quad (4)$$

也就是说, 这模型没有铁磁性 (除非 $T = 0$), 这是个相当普遍的结果. 一般一维空间的模型不会有铁磁性. 这结论是指很长的物体, 即

$$N \gg \xi \equiv e^{2J/T}. \quad (5)$$

ξ 是两分界点间的平均距离. 如果 N 比 ξ 要小, 则各 s_r 都是一样了, 这物体可以看作磁铁. 在一定的温度下, 只要物体够长, 铁磁性会因分界点的出现而消失.

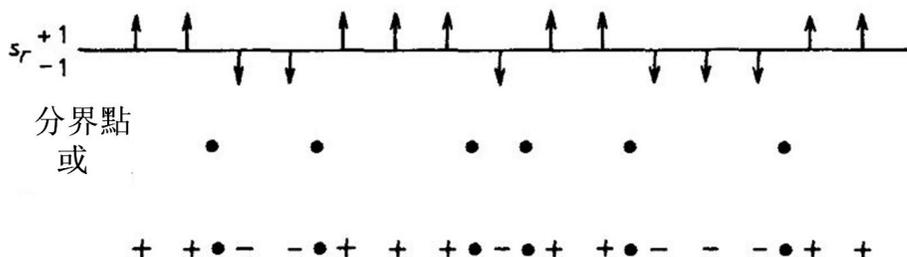


图 1

现在来看 $h \neq 0$ 的情形. 图 1 指出: 各分界点把物体分成许多段, 每段中的单元同值. 令 y_1, y_2, y_3, \dots 为各段长度, 即 y_i 为第 i 段中的单元数. 因此, 总能量 (1) 可写成

$$H(y_1, y_2, \dots) = 2JN' + h(y_1 - y_2 + y_3 - y_4 \cdots - y_{N'}) \quad (6)$$

这能量是由零点能量算起, 假定第一段中的 s_r 为 $+1$, 第二段的为 -1 , 等等, 并假定 N' 为偶数. $2JN'$ 为 N' 个分界点的能量. 总长为各段长之和

$$N = y_1 + y_2 + \cdots + y_{N'}, \quad (7)$$

这可以看作分界点气体的体积. 我们可以用定温定压法则来决定各平衡性质. 这和第十章最后一例相同. 令

$$\begin{aligned} G(p, T) &= -T \ln Z', \\ Z' &= \sum_{y_1, \dots, y_{N'}} \exp\{-(H + pN)/T\}. \end{aligned} \quad (8)$$

将 (6) 代入 (8), 得

$$Z' = e^{-2JN'/T} \left[\sum_{y=1}^{\infty} e^{-y(p+h)/T} \right]^{N'/2} \left[\sum_{y=1}^{\infty} e^{-y(p-h)/T} \right]^{N'/2}, \quad (9)$$

$$\begin{aligned} G(p, T) &= 2N'J + \frac{1}{2}N'T \left[\ln(e^{(p+h)/T} - 1) \right. \\ &\quad \left. + \ln(e^{(p-h)/T} - 1) \right]. \end{aligned} \quad (10)$$

压力 p 和分界点粒子数 N' 可以由

$$\begin{aligned} \frac{\partial G}{\partial p} &= N, \\ \frac{\partial G}{\partial N'} &= 0. \end{aligned} \quad (11)$$

决定. 结果是

$$\begin{aligned} N &= \frac{1}{2}N' \left[\frac{1}{1 - e^{-(p+h)/T}} + \frac{1}{1 - e^{-(p-h)/T}} \right], \\ e^{-4J/T} &= (e^{(p+h)/T} - 1)(e^{(p-h)/T} - 1). \end{aligned} \quad (12)$$

从这两式, 可以解出 N' 和 p . 注意, p 是分界点的“压力”, 就是使各个 y 减小之力. 温度越高, p 越大. 磁场 h 对单元为 $+1$ 各段是一张力, 使 y 加大, 对单元 -1 者为压力. 由 (12) 的第二式可得

$$e^{p/T} = \cosh \frac{h}{T} \pm \sqrt{\cosh^2 \frac{2h}{T} - 1 + e^{-4J/T}}. \quad (13)$$

根号前的符号必须为正, 方能使压力 p 有意义. 把 (13) 代入 (12) 的第一式即得 N' .

在 h 很小时, 总磁矩及磁化率是

$$\begin{aligned} M &= - \left(\frac{\partial G}{\partial h} \right)_{p,T,N} \\ &= \frac{hN'}{T} e^{-p/T} (1 - e^{-p/T})^{-2}, \end{aligned} \quad (14)$$

$$\chi = \frac{M}{h} = \frac{N}{T} e^{2J/T} = \frac{N}{T} \xi. \quad (15)$$

如果各单元为独立, 则 $\chi = N/T$. 现在多了一个 ξ . 这是各单元维持一致方向的平均距离.

2. 二维空间, 易形铁磁性的证明

一维空间的例子是最简单的. 二维空间以上的易形模型就不容易解了. 但是, 若要证明易形磁铁 (在二维以上空间) 有铁磁性, 却不很难. 证明铁磁性, 即是证明在没有磁场而且温度不必是零的情形下, 磁矩的平均值非零. 现在来作此一证明.

二维空间的模型定义如下. 假想一围棋盘格子. 每格点 i 上有一个单元 s_i . 总能量是

$$H(s) = -\frac{1}{2} \sum_{i,j} J_{ij} s_i s_j - h \sum_i s_i, \quad (16)$$

$$\begin{aligned} J_{ij} &= J \quad \text{如 } i, j \text{ 相邻,} \\ &= 0 \quad \text{如 } i, j \text{ 不相邻.} \end{aligned}$$

每一点和四点相邻. 我们先简单说明铁磁性为什么会发生, 再循理作严格证明.

上节讲的分界点概念, 可以推广到二维空间. 图 2 示一些单元之值, (+1 或 -1, 简作 +, -).

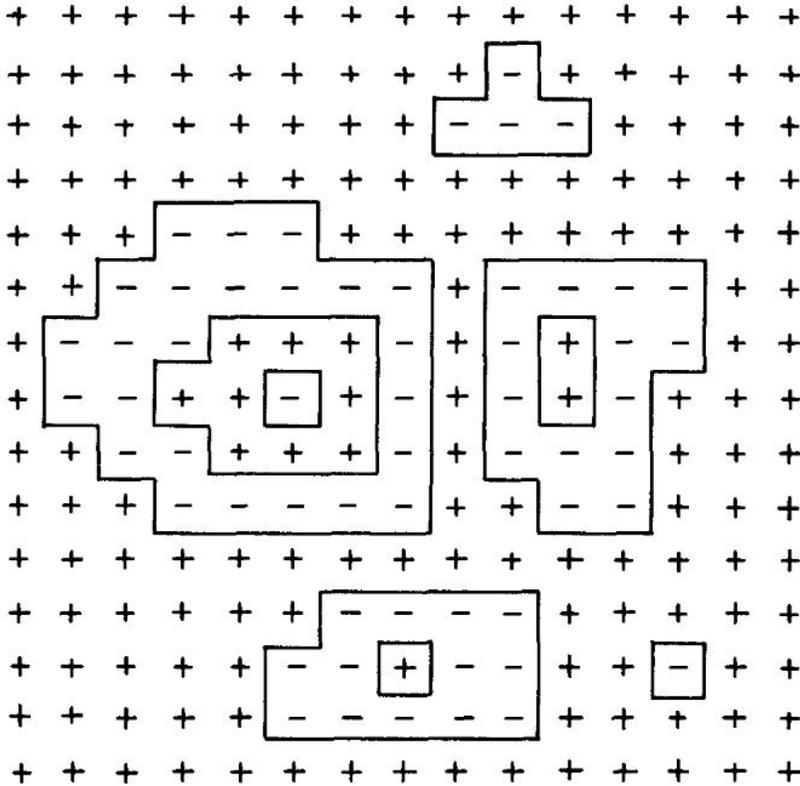


图 2

凡相邻而反向的单元之间, 都有一分界点. 这些点可以连成“分界线”, 划出区域来. 这分界线的长度, 可以由在线的点数计算. 总磁矩是正区面积减去负区面积. 要产生分界线, 必须要花能量. 在 $T = 0$ 时, 全体方向一致, 没有分界线. 假设方向为正. 把温度加到 T , 则负区会发生. 如果某负区的界线长为 L , 则能量为 $2JL$, 因为每分界点的能量为 $2J$. 因此, 这负区的产生概率约为

$$e^{-2JL/T}. \quad (17)$$

假设 $h = 0$, L 至少是 4. 因此, 只要温度够低, 大的负值不易发生, 因此多数单元为正, 总磁矩为正, 是为铁磁性.

这和一维空间的情形不一样. 在一维空间, 只要分界点的密度非零, 就有正区和负区, 各区大小可以不再花能量而改变. 因此平均起来正和负的单元一样多. 在二维空间, 虽然分界线的密度非零, 但温度低时负区不易扩张, 因

为扩张须要更长的分界线. 三维空间结果也相似, 分界线成了分界面. 在低温下, 除了些小区域之外, 所有的单元方向都一致.

现在把以上的话, 用严谨的数学步骤重说一遍[†].

任取一单元 s_1 . 令 s_1 为 $+1$, -1 的概率各为 p_+ , p_- . 则

$$\langle s_1 \rangle = p_+ - p_- \quad (18)$$

现在要证明, 只要 T 够小, 则 $\langle s_1 \rangle \neq 0$.

假设棋盘四周边上的单元都固定为 $+1$. 我们来证明 p_- 比 p_+ 小.

如果 s_1 是 -1 , 则它一定是被一圈, 或三圈, 或五圈, 分界线围绕着. 圈数必为单数. (如果围了两圈或四, 六, 圈, 则 s_1 必是 $+1$. 如果没有圈子围它, 自然是 $+1$.)

如果把 s_1 和最内圈里的单元改成 $+1$, 则少了一圈, 成了一个 $s_1 = +1$ 的状态. 因此, 任一个 $s_1 = -1$ 的状态的概率可以写成

$$e^{-2JL/T} \times e^{-2JL'/T} / Z \quad (19)$$

L 是最内圈的长度, L' 是所有其他分界线的总长, 包括其他各处的分界线. Z 是总份量.

(19) 只是一个状态的概率. 要求 p_- , 我们先计算在其他圈子固定下, 有多少状态其内圈长为 L :

$$g(L) = \text{内圈长为 } L \text{ 的状态数.} \quad (20)$$

这 $g(L)$ 自然是要看其他分界线的位置. 其他分界线占住的地方是不可以再画圈子的. 但是, 如果我们忽视其他分界线, 计算所有绕 s_1 一圈的状态数, 则我们就多算了. 因此, 在其他分界线固定的条件下, s_1 为 -1 的概率一定比

$$\left(\sum_L g(L) e^{-2JL/T} \right) e^{-2JL'/T} / Z \quad (21)$$

要小. 式中 $g(L)$ 的计算忽略其他分界线的存在. 现在我们把其他分界线的各状态加起来. (21) 括号外的就是一个 $s_1 = +1$ 的状态的概率, 因为最内圈去了之后 s_1 就成了 $+1$. 因此, 加起来就成了 p_+ . 结果是

$$p_- < \left(\sum_L g(L) e^{-2JL/T} \right) p_+ \quad (22)$$

[†]内容和派若氏 (Peierls, (1936).) 的证明相同, 细节稍有改变.

$g(L)$ 指绕着 s_1 一圈分界线有多少画法. 从任一点开始画, 每多画一单位有三种画法, 即除了上一步的方向, 其他三方向都可以走. 画 L 单位, 即走 L 步, 有 $4 \times 3^{L-1}$ 个画法, 4 是因为第一步有四个方向可走. 我们要求在 L 步后回到原处, 因此, 真正的总数比 $4 \times 3^{L-1}$ 小. 再者, 从圈上任一点开始都可以, 左右旋方向也一样, 因此, $4 \times 3^{L-1}$ 太夸大了些, 应除以 $2L$. 我们可以说

$$g(L) < \frac{4}{2L} 3^{L-1} \left(\frac{L}{4}\right)^2 \quad (23)$$

绝无问题. 式中 $(L/4)^2$ 是分界线所能包括的最大面积, 即以 $L/4$ 为边的正方形面积. 因为 s_1 可以在圈内任一点, 故 $g(L)$ 必须有面积这一因子. 将 (23) 代入 (22), 得

$$p_- < \sum_{L=4}^{\infty} \frac{L}{8} 3^{L-1} e^{-2JL/T}. \quad (24)$$

注意: 最短的一圈是 $L = 4$, 且 L 必为双数.

(22) 尾的 p_+ 略去了, 因为 $p_+ < 1$. (24) 的级数很容易计算, 把单数项也加入, 以求简单, 得

$$p_- < \frac{5}{24} \frac{e^{-4\alpha}}{(1 - e^{-\alpha})}, \quad (25)$$

$$\alpha \equiv \frac{2J}{T} - \ln 3,$$

$$\langle s_1 \rangle = 1 - 2p_- > 1 - \frac{5}{12} \frac{e^{-4\alpha}}{(1 - e^{-\alpha})}. \quad (26)$$

当 $T = J$ 时, $e^{-\alpha}$ 已小于 $1/2$, 因此, 如 $T < J$, 以上的结论是

$$\langle s_1 \rangle > \frac{91}{96}. \quad (27)$$

证明完毕.

在以上的证明, 我们假设在边界上的单元为 $+1$. 如果不作任何假设, 则 $\langle s_1 \rangle$ 必是零, 因为这模型在 $h = 0$ 时有 $+ -$ 对称性, 即

$$H(s) = H(-s). \quad (28)$$

铁磁性的意思是: 只要一点点的偏向力量就可以造成 $\langle s_1 \rangle \neq 0$. 边界只是总值的很小一部分. 以上的证明还可以改良, 使这 s_i 定为 $+1$ 的部分更缩小些. 用一很小的磁场来指定方向亦可.

以上的证明, 可以立刻推广到三维或三维以上的空间. 读者试证之.

二维空间的模型已被解出[‡], 只是数学过程太烦难, 不在此讨论. 以下是一些结果. (见图 3)

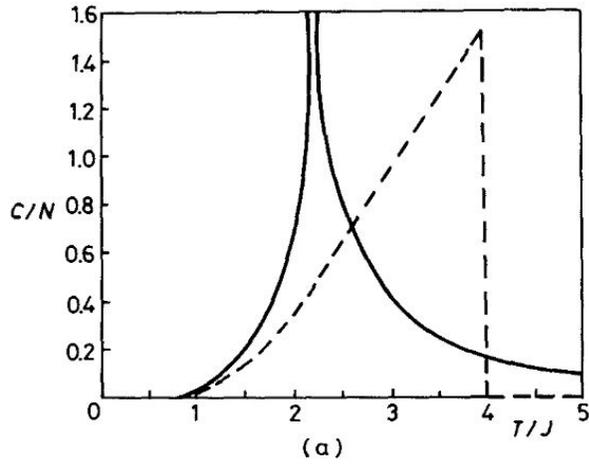


图 3a, 热容率

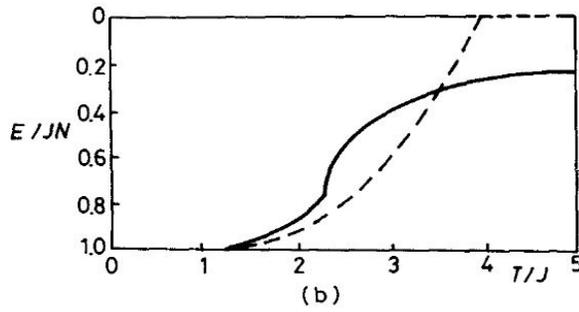


图 3b, 能量

[‡]Onsager(1944).

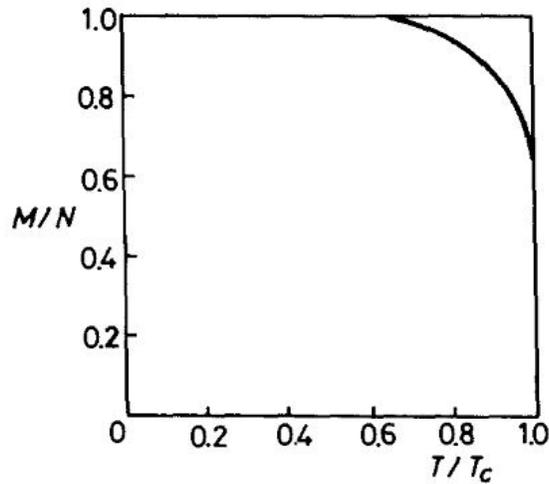


图 3c, 磁矩. 虚线为平均场解结果 (见第二十七章, 平均场解的 T_c 为 $4J$).

$$\sinh \frac{2J}{T_c} = 1, \quad (29)$$

$$T_c = 2.269J,$$

$h = 0$, $T < T_c$ 的总磁矩是[§]

$$M = N \frac{\cosh^2 2J/T}{\sinh^4 2J/T} [\sinh^2(2J/T) - 1]^{1/8}. \quad (30)$$

在低温, M/N 几乎就是 1. 在接近 T_c 时,

$$\frac{M}{N} \propto (T_c - T)^{1/8}. \quad (30')$$

3. 其它易形模型

以上的磁铁模型, 稍为改头换面一下, 就变成了别种物体的模型. 因此, 很多不同的现象, 可以用同一模型分析, 使我们了解这些不同现象的相同地方. 现在举几个例子.

(甲) 易形反铁磁体模型

[§]Yang(1952).

如果把 (16) 的 J 改成 $-J (J > 0)$, 则相邻单元相反时能量较低. 在低温时, 相邻单元反向, 这模型成了反铁磁模型.

当 $h = 0$ 时, 如果把每隔一格的 s_i 改成 $-s_i$, 则相邻单元和以前一样, 因为又多了一个负号, 把刚才改 $-J$ 的负号消了. 因此, 如果 $h = 0$, J 是正是负, 不会影响到热位能. 计算热位能时, 每个 s_i 的正负二值都要加进去的. 由以上两节的结果可知, 一维空间的易形反铁磁模型, 没有反铁磁性, 二维空间有.

如果 $h \neq 0$, 则铁磁模型和反铁磁模型就不同了. 需要重新分析.

(乙) 合金模型

令 $s_i = 1$ 代表在 i 点有一个原子 A , $s_i = -1$ 代表在 i 点有一个原子 B . 令 ϵ_{AA} , ϵ_{BB} , ϵ_{AB} 为相邻原子间的作用能. 这就是一个“合金”的模型. 这模型可以化成一个易形磁铁模型. 一对相邻原子 ij 的能量可写成

$$\begin{aligned} \epsilon_{ij} &= -J s_i s_j - \frac{1}{z} h (s_i + s_j) + K, \\ H &= \frac{1}{2} \sum_{i,j} \epsilon_{ij}. \end{aligned} \quad (31)$$

z 是每单元的邻居数.

如果二原子都是 A , 则 $\epsilon_{ij} = \epsilon_{AA}$, $s_i = s_j = 1$, 即

$$-J - \frac{2h}{z} + K = \epsilon_{AA}. \quad (32)$$

同理,

$$-J + \frac{2h}{z} + K = \epsilon_{BB}, \quad (33)$$

$$J + K = \epsilon_{AB}.$$

由 (32), (33) 解 J , h , K 得

$$\begin{aligned} J &= \frac{1}{2} \epsilon_{AB} - \frac{1}{4} (\epsilon_{AA} + \epsilon_{BB}), \\ K &= \frac{1}{2} \epsilon_{AB} + \frac{1}{4} (\epsilon_{AA} + \epsilon_{BB}), \\ h &= \frac{z}{4} (\epsilon_{BB} - \epsilon_{AA}). \end{aligned} \quad (34)$$

用 (34), 则 (16) 就成了一合金模型. 在温度够低时, A 原子聚到一起, B 和 B 聚到一起, 形成多相共存, (假定 $J > 0$). 如果 $J < 0$, 则低温时 A 、 B 原子会作相互排列, 和反铁磁相似.

(丙) 易形格子气体模型

令 $s_i = 1$ 代表在 i 处有一分子, $s_i = -1$ 代表在 i 处没有分子. 同一地方不能有两个或更多的分子. 令相邻分子之作用能为 $-\epsilon$, 再把相邻两点的能量写成 (31), 则

$$\begin{aligned} -J - \frac{2h}{z} + K &= -\epsilon, & s_i = s_j &= 1, \\ -J + \frac{2h}{z} + K &= 0, & s_i = s_j &= -1, \\ J + K &= 0, & s_i = 1, & s_j = -1. \end{aligned} \quad (35)$$

解之得

$$\begin{aligned} J &= \frac{1}{4}\epsilon, \\ h &= \frac{z}{4}\epsilon. \end{aligned} \quad (36)$$

因此, (16) 也是一个气体模型. $-\epsilon$ 代表气体分子间的引力. 格子的限制代表一种斥力, 使每格只能有一分子. 只要温度够低, 这模型可有两相. 高密度, 即 $\langle s_i \rangle > 0$, “液相”, 和低密度, $\langle s_i \rangle < 0$, “气相”.

以上的例子指出, 许多不同的现象可用同一模型分析.

以上只是三个最简单的例子. 还有数不清的易形模型. 不均匀的作用, 不规则的作用, 长程作用, 等等. 分析的现象包括相变、杂质、边界, 等等. 以后我们会讲一些.

虽然易形模型用途很广, 但有许多现象是不能用它来分析的. 例如有球形对称的问题, 自旋波, 等等. 自旋是一向量, 而易形变数没有角度的概念在里面. 因此, 在用易形模型分析某种现象时[¶], 必须先了解模型先天上受到的限制.

讨论问题十七

1. 一维空间的模型, 多可用所谓“转移矩阵”法, 只要相互作用限于近邻.

令

$$H = \sum_{i=1}^N K(s_i, s_{i+1}), \quad (37)$$

$$s_{N+1} = s_1.$$

[¶]易形模型颇有一些物理、化学、数学的人才问津. 文献太多. 最近 Fisher(1981) 的简介内列举了一些. 若要搜集资料, 可由此着手.

K 可以是 s_i, s_{i+1} 的任何函数.

(A) 试证:

$$Z = \sum_{s_1, s_2, \dots, s_N} e^{-H/T} = \text{Tr}(Q^N), \quad (38)$$

矩阵 Q 的定义是

$$Q_{\alpha\beta} \equiv e^{-K(\alpha,\beta)/T}, \quad (39)$$

是为“转移矩阵”.

(B) 试证

$$F = -T \ln Z = N\lambda. \quad (40)$$

λ 是 K 的最低本征值, 即 $e^{\lambda/T}$ 是 Q 的最大本征值. N 的大数.

(C) 用转移矩阵法, 解模型 (1), 先证明

$$Q = \begin{pmatrix} e^{(J+h)/T} & e^{-J/T} \\ e^{-J/T} & e^{(J-h)/T} \end{pmatrix}, \quad (41)$$

2. 以上的 s_i 不必限于二值. 如果 s_i 有 n 值, 则 Q 为 $n \times n$ 矩阵. 读者试解 $n = 3$ 的情形, 即 $s_i = 1, 0, -1$, 能量仍为 (1).
3. 令某量子力学模型有 n 个状态, 其总能算符为一个 $n \times n$ 的矩阵 K . 如果 $\psi(0)$ 为在时刻 $t = 0$ 时的状态向量, 则在时刻 t 的状态为

$$\psi(t) = e^{-iKt}\psi(0). \quad (42)$$

由此可见, 一维空间统计力学的问题和量子力学的求零点能量类似. (42) 的 K 有如 (39) 的 K , 它相当于 $-N/T$, e^{-iKt} 相当于 (38) 的 Q^N . 此地 N 是空间的长度, 而 t 是量子力学的时间. 因此, 易形模型 (1) 相当于量子力学的一个自旋 1/2 模型:

$$K = a + \mathbf{b} \cdot \boldsymbol{\sigma}. \quad (43)$$

b 是“磁场”, a 为一常数. σ 为自旋 1/2 算符. 试求 a, b 和 j, h 的关系.

4. 试将第 2 节的证明推广到三维空间易形模型.
5. 如果 2 维空间易形模型的温度很低, $T \ll T_c$, 且有一很小的正磁场 h , 则大多数单元都是 +1. 在某时刻忽然把磁场反过来, 成 $-h$. 试讨论以后的运动情况. (提示: 参考第十五章“水珠的消长”一节).

第十八章 杂质和运动

不同的分子混在一物体中，可曰“杂”。溶液、合金都是例子。就杂原子的运动情况来分，可分成两类，一为“走动”，如水中溶的氧。一为“固定”，如钢铁中的锰。这运动情况和物体的平衡性质大有关系。本章强调这两种情况的不同，统计力学的分析亦大不相同。溶液中的溶质分子，一如气体有渗透压，固体中的杂原子，固定不动，就和气体的行为完全不同了。走不动自然就谈不上渗透了。无论其位置分布是混乱或规则，不动和动就是不一样，不可混为一谈。混乱而凝固的物体曰“非晶体”或“玻璃体”。非晶体中的分子运动常有不同的时间尺度。无法把动和不动作明显的划分。本章举例略提一下非晶体的特征。

1. 溶液和渗透压

渗透压是溶质分子运动产生的压力。这压力可以用“半透膜”来观测，半透膜让溶剂分子透过，却不让溶质分子通过。

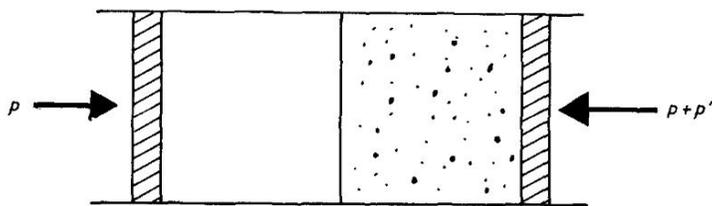


图 1

图 1 画的是一个半透膜，把液体分成两半。右半有溶质，左半没有。两边的压力用活塞维持住。右边的溶质分子去不了左边，而溶剂分子却可以来往左右，不受半透膜影响。右边比左边高出的压力叫做渗透压。图 1 的压力也可

以用重力来维持, 如图 2 所示, U 形管中间为半透膜. 左右两个活塞一样重. 右边就要多加一些重量才能保持平衡, 所加的重量就可以就用适量的溶液来代替, 即

$$p' = \rho gh. \quad (1)$$

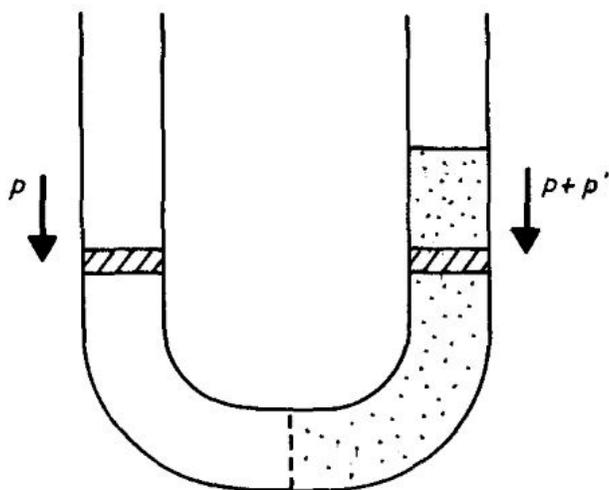


图 2

ρ 是溶液的质量密度. 其实把两边的活塞都拿掉也没关系, 因此, 左边的液面看来是被推上去了. 由 (1) 可见, 渗透压是很容易测定的.

如果溶液很稀, 则溶质分子就很像一个理想气体. 其压力是

$$p' = n'T, \quad (2)$$

n' 为溶质分子密度.

2. 有效作用能

我们现在来仔细分析溶质分子在溶剂 (以下简称“水”) 中的作用. 假定有 N 个水分子, N' 个溶质分子. 令哈密顿量为

$$H = H_0 + H_1. \quad (3)$$

H_0 包括水分子的动能, 水分子间的作用能及溶质分子的动能. H_1 是溶质分子之间及其与水分子的作用能:

$$H_1 = \sum_{k=1}^N \sum_{i=1}^{N'} u(\mathbf{r}_k - \mathbf{r}'_i) + \frac{1}{2} \sum_{i,j=1}^{N'} v(\mathbf{r}'_i - \mathbf{r}'_j). \quad (4)$$

\mathbf{r}'_i 为各溶质分子的位置, \mathbf{r}_k 为水分子位置. u, v 为一对分子间的作用能.

总热位能 $F(N, V, N', V', T)$ 可以用定温法则计算

$$e^{-F/T} = \sum_{s'} \sum_s e^{-[H_0(s) + H_1(s, s')]/T}, \quad (5)$$

$$s \equiv (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N), \quad s' \equiv (\mathbf{r}'_1, \mathbf{r}'_2, \dots, \mathbf{r}'_{N'}).$$

式中的 s, s' 指水分子及溶质分子的坐标. V, V' 为水分子, 溶质分子所占的体积, V' 为 V 的一部分. 分子的动能部分可以立即积分算出, 和 V, V' 无关, 我们不再管它. (5) 的和可以分两步做, 先做 s 再做 s'

$$e^{-F/T} = e^{-F_0/T} \sum_{s'} e^{-W(s')/T}, \quad (6)$$

$$e^{-F_0/T} \equiv \sum_s e^{-H_0(s)/T},$$

$$e^{-F_0/T - W(s', T)/T} \equiv \sum_s e^{-[H_0(s) + H_1(s, s')]/T}. \quad (7)$$

F_0 是略去 H_1 时的热位能. $W(s', T)$ 是溶质分子的“有效作用能”. 它是在各 s' 固定时的热位能. 这又是一个分类统计的例子, 是在一定的 s' 下, 统计水分子状态 s 而得. W 代表一种作用能, 是由水分子运动, 溶质自己之间的作用, 溶质和水之间的作用综合而得. 这有效作用能是一个非常重要的概念, 当我们要把物体的某一部分, (如此地的溶质分子), 拿出来单独考虑时, 这一部分和它的环境间的作用, 有时十分强, 必须考虑进去. 用这有效作用能的概念, 我们可以把环境看成一“介质”. 其功用是使这当事的一部分的作用能有适当的修正, 然后我们可以把介质忘了. 这种做法是十分常见的, 例如, 在物质中的静电作用. 外加电荷被物质分子影响的结果, 可以用一介电系数表示. 外加电荷除以介电系数就是其“有效电荷”. 以有效电荷为电荷, 即可把介质忘了. 这是一个十分方便的处理方法.

必须要特别留意的是: “有效作用能”是为某一种用途而定, 不能作别的用. 例如, 此地的 $W(s', T)$, 是用来分析平衡现象, 不能直接用来分析分子的

运动. 当然, 它可能对分析分子运动有所帮助. 但它本身不是一个哈密顿量, 即不能用它的微分求分子运动方程式.

当然, 我们也可以定义水分子之间的有效作用能. 由于溶质分子的介入, 水分子之间的作用受了影响, 可以用有效作用能代表.

(7) 也可以写成一个平均值:

$$\begin{aligned} e^{-W/T} &= \langle e^{-H_1/T} \rangle_0 \\ &\equiv \frac{\sum_s e^{-H_0/T} e^{-H_1/T}}{\sum_s e^{-H_0/T}}, \end{aligned} \quad (8)$$

也就是对水分子运动的平均值, 溶质分子不动. 注意, W 不是 H_1 的平均值, 而是一个相当复杂的计算过程.

$$W = -T \ln \langle e^{-H_1/T} \rangle_0. \quad (9)$$

3. 低密度情形

有效作用能的计算自然是不容易的. 但在特殊情形下, 近似的计算可以适用. 我们来看低密度的情形.

密度低则各溶质分子多半时间是分开的, 其间相互作用 v 可以略去, 但和水的作用 u 不能略去 (见 (4)). 但水分子之间的作用为短程. 如果溶质分子间的距离够大, 则各不相干, (8) 的计算可以简化如下:

$$\langle e^{-H_1/T} \rangle_0 \approx \prod_{i=1}^{N^*} \left\langle \exp \left(-\frac{1}{T} \sum_k u(\mathbf{r}_k - \mathbf{r}'_i) \right) \right\rangle_0, \quad (10)$$

即 (4) 的末项略去, 右第一项的各 \mathbf{r}_i 为独立.

再者, 水为均匀, (10) 中的平均值不因 \mathbf{r}'_i 而异. 所以在这近似下,

$$W \approx -TN' \ln \left\langle \exp \left(-\frac{1}{T} \sum_k u(\mathbf{r}_k) \right) \right\rangle_0 \equiv N' f'. \quad (11)$$

f' 和 \mathbf{r}' 无关. 将 (11) 代入 (6), 得

$$\begin{aligned} \sum_{s'} e^{-w/T} &= e^{-N' f'/T} \frac{V^{N'}}{N'^!} \\ F &= F_0 + N' f' - TN' \left(\ln \frac{V'}{N'} + 1 \right). \end{aligned} \quad (12)$$

这结果指出, 每溶质分子有热位能 f' , 除此之外, 它的平衡性质一如理想气体分子. 渗透压是

$$p' = -\frac{\partial F}{\partial V'} = \frac{TN'}{V'}, \quad (13)$$

因为 F_0 非 V' 的函数, f' 是 T 的函数. (13) 和 (2) 是同一回事. 我们的重要结论是: 无论水分子和溶质分子之间的作用有多强, 只要溶质分子密度够低, 渗透压就是 (13).

在低密度时, 溶质和水的作用效果, 全包括在 f' 里了. 现在用个例子说明 f' 的重要性.

【例】水面上溶质的气压为 p , 求水中溶质的密度 N'/V' .

溶质分子可以自由出入水面, 因此化学势 μ 在水面上下一样, 水面以上的气体可视为理想气体, 所以化学势是

$$\begin{aligned} \mu &= T \ln(\lambda^3 p/T), \\ \lambda &\equiv \frac{h}{\sqrt{2\pi mT}}. \end{aligned} \quad (14)$$

水面下的化学势可以由 (12) 求得

$$\mu = \frac{\partial F}{\partial N'} = T \ln\left(\lambda^3 \frac{N'}{V'}\right) + f'. \quad (15)$$

注意 (12) 的 F_0 中包括了溶质动能所供给的各项. 事实上在水中溶质除了能量 f' 之外和理想气体一样. 由 (14), (15) 得

$$\frac{N'}{V'} = \frac{p}{T} e^{-f'/T}. \quad (16)$$

因此, 如果知道 f' 和 p , 就可以算出在水中溶质的密度.

如果溶质密度不是很低, 则必须予 (10) 以修正. 结果是除了 f' 之外, 还有溶质分子之间的有效相互作用能. 这一点在下节会更清楚.

4. 走动杂质和固定杂质

以上讲的溶液中的溶质, 是游动杂质, 在观测时间内, 溶质分子不断地跑来跑去.

如果杂质分子不跑来跑去, 而固定在一位置, 则为固定杂质, 或“冻结”杂质. 在固体内的杂质多是固定不动的. 在粘性高的液体内, 杂质也可以不动.

动或不动, 是一个时间尺度的问题. 杂质的行动, 可用扩散来描述. 在时间 t 内扩散之距离约为

$$l \sim (Dt)^{1/2} \quad (17)$$

(见第十二章第九节), D 是扩散系数. 如果 n 是杂质密度, 则 $n^{-1/3}$ 为平均距离, 令 \mathcal{T} 为观测时间, 则杂质在

$$\begin{aligned} \mathcal{T} &\gg \frac{1}{Dn^{2/3}}, \text{ 为游动} \\ \mathcal{T} &\ll \frac{1}{Dn^{2/3}}, \text{ 为固定.} \end{aligned} \quad (18)$$

在固体中, 一般杂原子的扩散系数遵循以下实验结果

$$D = D_0 e^{-E/T} \quad (19)$$

图 3 示碳原子在铁中的扩散系数. $D_0 = 0.02\text{cm}^2/\text{sec}$, $E = 0.87\text{eV}$.

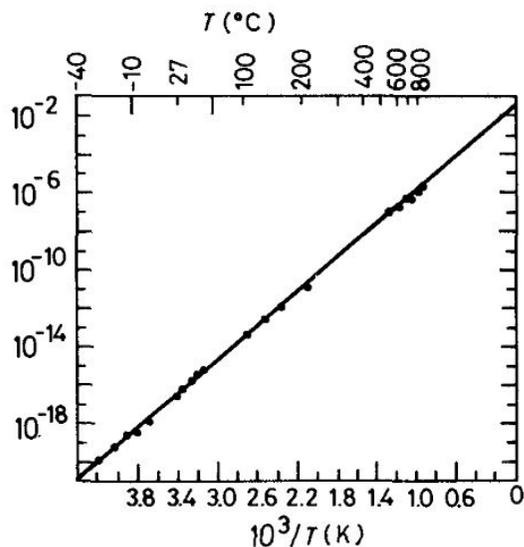


图 3. (取自 Kittel(1966)p.569)

(19) 式中的指数显示: 杂原子要走一步, 必须要越过一个能量为 E 的屏障. 它的四周被其他原子围住, 要挤一步很不容易. D 对温度极为敏感, 由 (19) 可见. 一般的 E 都在 1 至 2eV 左右, 即 10^4 K 左右. 在低温时, D 是微不足道, 杂质除了作小幅度振动外, 是不走动的. 在许多情形下, 只有在固体将要熔解的温度, 杂原子才能走动.

在应用基本假设时, 走动杂质和固定杂质是两个完全不同的情形. 前者一如以上各节的溶液, 杂原子的坐标不断地改变, 一如气体分子. 后者每个杂原子位置都是一个不变量, 一如固体原子. 无论杂原子的位置分布是如何的混乱, 或是如何的有规则, 它们每个的位置都是不变量. 这一点非常重要. 现在来看一个例子.

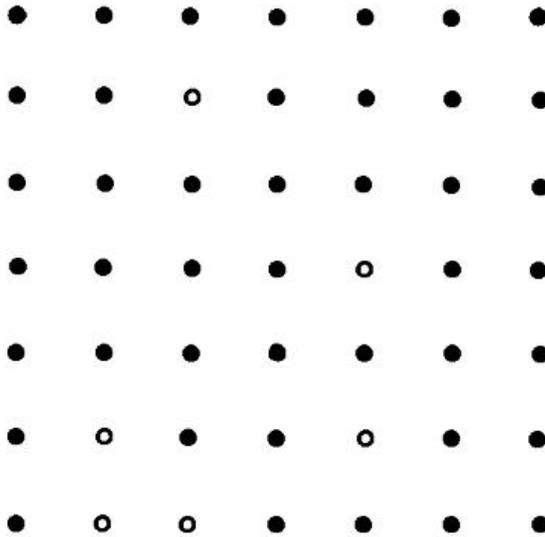


图 4

图 4 示一棋盘, 放满了两种原子, 正子 (黑) 及负子 (白). N_+ , N_- 为这两种原子的粒子数. 令同种原子间的作用能为 $-J$, 异种之间为 J , 作用限于相邻原子.

现在假定 $N_- \ll N_+$, 即负子很少, 可视为“杂原子”. 如果 $N_- = 0$, 即没有负子, 则总能量为 $-\frac{1}{2}NzJ$, 此地 $z = 4$, 是每原子的邻居数目. N 为总原子数, 即棋盘上位子数. 现在把一个正子换成负的, 则能量的增加为 $2Jz$. 因此, 如果负子密度非常小, 则总能量为

$$H_0 = -\frac{1}{2}NzJ + 2JzN_- \quad (20)$$

但如果有负子相邻, 则 (20) 须要修改. 每一对相邻的负子, 要多加 $-4J$ 的能

量, 因此, 总能量为

$$H = H_0 + H_1, \quad (21)$$

$$H_1 = \frac{1}{2} \sum_{i,j=1}^{N_-} u(|\mathbf{R}_i - \mathbf{R}_j|),$$

$\mathbf{R}_i, i = 1, 2, \dots, N_-$ 为负子的位置,

$$u(0) = \infty,$$

$$u(1) = -4J,$$

$$u(r) = 0, \quad \text{如 } r > 1.$$

相邻原子之距离为 1. 这当然就是第十六章尾的易形气体模型. 正子看成一个介质, 或“溶剂”, 负子为“溶质”. 不过我们现在要讨论的是这些原子是否走动的问题.

一个原子要走动, 必须和一个邻居交换位子. 交换位子需要被“推一把”. 在一般晶体中, 这是一个相当复杂的过程. 各原子是在不断地做小幅度振动. 如果刚好西邻向南移了一点, 自己向北移了一点, 而东邻正好向西推一把, 则挤向西邻位子的机会就大些 (见图 5). 这种过程本身就是一门研究专题.

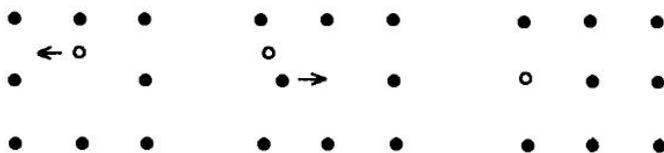


图 5

在此我们只假定此种过程的存在, 并假定过程本身很快. 即不发生则已, 一发生就很快结束. 我们只考虑发生前后状态, 过程中的状态略去了.

如果负子时常走动, 则这问题和溶液问题相同, 只是原子的位置被限制在分立的棋盘目上. 在这简单的模型里, (20) 的 $2Jz$ 即是 (12) 式中的 f' . 在上一节我们并未讨论杂质原子间的作用. 在此, (21) 式的 H_1 有十分重要的效果. 由上一章易形模型的讨论可知, 在低温时, 负子会聚到一起, 即和正子

分开, 以降低能量. 现在来看一点简单的计算, 令 F 为定温热位能, 则

$$F = -T \ln Z$$

$$Z \equiv \frac{1}{N_-!} \sum_{\mathbf{R}_1} \cdots \sum_{\mathbf{R}_{N_-}} e^{-H/T} \quad (23)$$

如果略去 J/T , 则结果是理想气体:

$$F \approx F_0 = H_0 - TS_0,$$

$$S_0 = N_- \ln \left(\frac{N}{N_-} + 1 \right). \quad (24)$$

如果负子浓度低, 则可用低密度的展开法 (见第十四章, 第 1 节). 我们只写下答案, 细节读者自己计算:

$$F = F_0 + F_1,$$

$$F_1 = -\frac{TN_-^2 v}{2N} (e^{3J/T} - 1) + \frac{TN_-^2}{2N}. \quad (25)$$

以上假设温度不太低, 负子不聚在一起. 任一负子, 它的某邻居为负子的概率是

$$p_-(T) = \frac{N_-}{N} e^{3J/T}. \quad (26)$$

(23)—(26) 是负子走动的结果.

如果负子不动, 则情形大不相同. 每个 \mathbf{R}_i 都是不变量, 无状态的改变. 因此热位能是

$$F'(\mathbf{R}_1, \cdots, \mathbf{R}_{N_-}) = H, \quad (27)$$

如此而已. H 是由 (21)、(20) 定义, F' 之值要看各负子的位置, 这些位置是事先定好的, 例如, 由高温冷却下来. 但是这些位置和冷却过程有关, 不同的冷法, 负子的位置分布就不一样了. 一旦各 \mathbf{R}_i 被冻住, 它们就不再受温度改变的影响 (只要不再热到又开始动的温度), 因此对熵, 对热容率都不影响. 注意, 无论负子的分布是如何混乱, 熵都不受影响, 因为没有运动. (27) 和 (24), (25) 可以说是非常不同. 杂原子对物性的影响, 和杂原子的走动与否, 关系实在太大.

在特殊情形下, (27) 的 F' 可以简化, 例如, 这物体是在某高温 T . 突然冷却下来, (是所谓“淬冷”), 而在 T_0 , 负子是可以走动. 如此, 负子的分布是反映

冷却之前 $T = T_0$ 时的分布. 在那时, 每负子的邻居中平均有 $zp_-(T_0)$ 个负子. (p_- 见 (26)). 因此, F' , 即 H 之值为

$$F'(\mathbf{R}_1, \dots, \mathbf{R}_{N_-}) = -\frac{1}{2}NzJ + 2JzN_- - \frac{1}{2}N_-z(3J)\frac{N_-}{N}e^{3J/T_0}. \quad (28)$$

注意, (28) 并非 F' 的“平均值”. F' 并无起伏, 没有什么好平均的, 平均每个负子有多少负邻居, 是一项统计数字, 由 N_- 个负子位置统计得来. 我们假设这平均值是 $zp_-(T_0)$.

简言之, 杂原子如不断地走动, 则状态空间中的轨迹就有个活动范围, \mathbf{R}_i 的改变幅度越大, 范围也越大. 如果不动, 则活动范围只是一点.

活动范围须由运动决定. 以上不是第一个例子. 在第八章讲氢分子时, 就已再三强调过.

以上讨论两种情形, “走动”和“固定”. 如果时间尺度是在这两者之间怎么办? 即走不太远又不完全固定. 这种情形就很麻烦了, 观测时间的长短, 及运动的细节都必须考虑进去. 也说是说, 问题不再是“平衡”的问题, 而是“不平衡”的问题了.

有许多物体中, 时间尺度没有明显的划分, 而是有长有短. 无论观测时间是多长, 都无法直接使用基本假设, 这种情形是最头痛的了, 在“非晶相”中常发生. 下节对非晶态作一简介.

5. 非晶态

上一节讨论的重点, 在“动”和“不动”的分别. 至于杂原子有多少, 倒是不重要. 在上个例子中, 如果正、负原子数目差不多, 结果会复杂一些, 但基本概念完全一样. 如果数目差不多, “杂质”这名词就不甚妥当, 必须叫这模型“混合体”, 如果各原子是固定的且分布混乱, 我们称之为“非晶体”, 或“非晶态”, 混乱而凝固也. 或称之为“不规则物体”.

“非晶”这个概念, 并不须要有不同种的分子相混. 同种的分子也可以有混乱而凝固的结构. 凡是有不规则的冻结结构, 都可称为“非晶”. 玻璃就是一个例子, 冰也是一个例子.

自然界的固体, 很少不是非晶体, 不是有冻结的杂质, 就是有不规则的结构. 在今天的材料学中, 非晶体是一主要部门. 用统计力学仔细分析非晶体, 还是最近十年的事, 理论尚在草创阶段.

以下用一个例子, 说明一些和非晶态有关的概念.

近年来“非晶磁体”是一相当受重视的研究题目. 锰铜合金是一个最常分析的非晶磁体. 它是在铜中熔入少许锰原子, 再冷到低温. 锰原子的最外围电子和铜的自由电子混到一起去了, 剩下锰离子就固定不能动了. 它带有自旋 $5/2$, 这自旋和金属电子的自旋作用, 微妙地影响了金属电子的运动, 是所谓“康斗”问题. 这问题我们不谈, 由于锰离子和金属电子的作用, 使两个锰离子自旋 (以下简称自旋) 发生了相互作用. 也就是说自旋 1 影响电子, 电子再影响自旋 2. 这相互作用大约是

$$J(\mathbf{R}_{12}) \mathbf{s}_1 \cdot \mathbf{s}_2, \quad (29)$$

$$J(R) = A \left[\frac{\cos 2p_F R}{(2p_F R)^3} - \frac{\sin 2p_F R}{(2p_F R)^4} \right].$$

这式子合用于锰原子密度低, 温度低的情况. 式中的 p_F 是铜的自由电子的顶层动量, R_{12} 是自旋 1, 2 间的距离. A 是一常数

$$A \sim \frac{K^2}{\epsilon_F} \left(\frac{p_F a}{2} \right)^6 \frac{1}{2\pi^3}, \quad (30)$$

K 是自旋和自由电子自旋的作用能. a 约是晶格的大小, ϵ_F 是自由电子顶层能量. 这相互作用 (29) 是所谓 RKKY 作用*, 这作用最重要的特点是随 R 变号. 这个特点来自电子自旋的相干函数, $J(R)$ 大致和这相干函数成正比, 因为 s_1, s_2 的作用, 全靠自由电子的来往于其间. 这以 π/p_F 为周期的起伏, 是由于自由电子分布, 在顶层的突然切断 (见第十三章).

因为 p_F 很大, 且锰原子的位置分布是混乱的, 所以各自自旋间的相互作用 $J(R_{ij})$ 有些正有些负, 成了一种不规则的作用. 注意, 各锰原子的位置是固定的, 因此各 $J(R_{ij})$ 也是固定的. 一般的实验在 $T \leq 20$ K, 锰的成分约在 10% 以下.

实验者的兴趣, 不是锰的混乱分布, 而是这分布引起的混乱自旋作用. 温度高时, 这些自旋形成一个顺磁体, 各自自旋可以自由改变方向. 但当温度降低过某温度 T_0 时, 有些自旋就开始冻结, 温度再降, 冻结得越多. 自旋的冻结方向是不规则的, 总磁矩是零. 这是自旋的“非晶态”或称“自旋玻璃相”.

这种非晶态有许多有趣的性质. 最重要的是它的“顽磁性”, 即用磁场将它磁化, 但除去磁场后, 磁矩 M 并不立即消失, 而作极缓慢的减退:

$$M(t) - M(0) \sim -aT \ln(t/\tau) \quad (31)$$

*Ruderman, Kittel, Kasuya, Yoshida(RKKY) 作用. 详细推导见 Mattis(1965), p.193-198.

a 是一常数, τ 是一微观时间尺度. 不但是磁矩, 其他量也是作此种对数式的改变, 此种改变, 显示着不同的运动时间尺度的存在, 即有些变量变得快, 有些太慢, 在观测时间内可算是冻结, 有些在中间. 这不但是非晶磁体的特性, 也是许多非晶体的特性. 要了解非晶体, 免不了要牵涉到不平衡的问题. 因此, 我们在讨论运动时再分析.

讨论问题十八

1. 某溶液中, 溶质的渗透压为 p' . 试证溶液的蒸气压 $p + \Delta p$ 比纯溶剂的蒸气压 p 要低,

$$\Delta p = -\frac{\rho_g}{\rho} p' \quad (32)$$

ρ 是溶液质量密度, ρ_g 是溶剂的蒸汽质量密度. 以上假设溶质的蒸气压非常低.

提示: 图 6 示一 O 形管, 有半透膜隔成两边. 由两边液面的压力差及液面高度差可得 (32). 溶质蒸气压太低, 没时间把溶质分子由上管送到左边去.

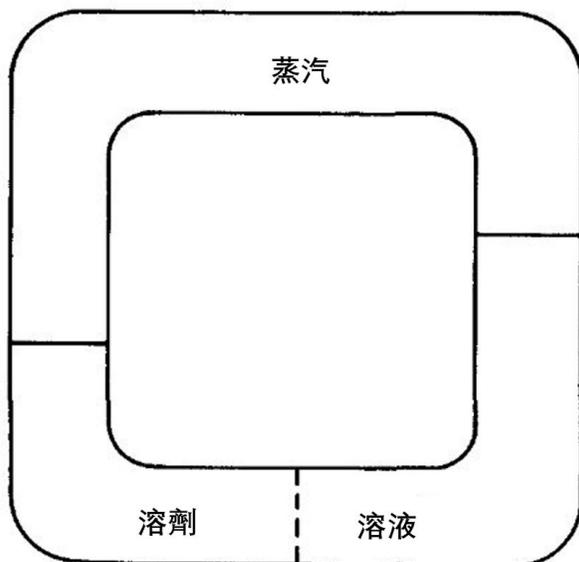


图 6

2. 计算海水中食盐的渗透压.

如果用半透膜, 可把海水升高多少? 读者会发现一相当大的数目.

3. 河口发电构想: 图 7 示河口筑两道水坝. 淡水在坝底发电, 再由半透膜入海. (从海洋学家 J. Issacs 听来). 试讨论此构想的可行性.

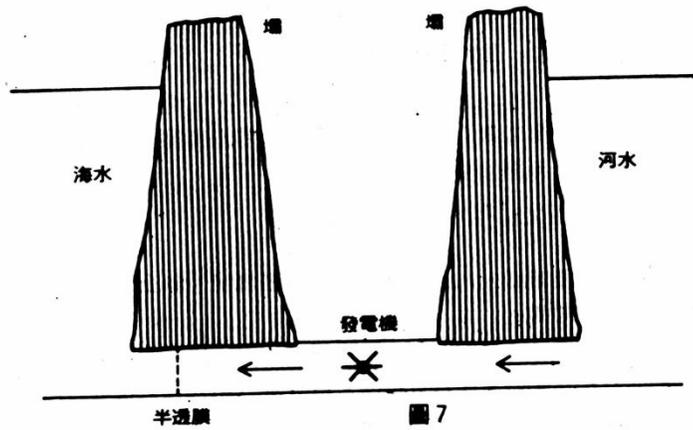


图 7

4. 复习讨论问题十二, 5, 并和本章第 2 节比较.

5. 一维空间的易形模型

$$H = -J \sum_{i=1}^{N-1} s_i s_{i+1} - h \sum_{i=1}^N s_i \quad (33)$$

可转换成另一相似模型, 步骤如下:

(甲) 令 $s_1, s_3, s_5 \dots$ 为物体的一部份 (1)

s_2, s_4, \dots 为另一部份 (2).

试写下求 H' 的式子. H' 是 (1) 部分的有效作用能.

(乙) 令

$$s'_1 = s_1, \quad s'_2 = s_3, \quad s'_3 = s_5 \dots \quad (34)$$

把 H'_1 写成

$$H'_1 = -J' \sum_{i=1}^{N'-1} s'_i s'_{i+1} - h' \sum_{i=1}^N s'_i + F_0, \quad (35)$$

$$N' = \frac{1}{2}N,$$

(假定 N 为偶数). 求 J' , h' 及 F_0 . 假设 $h = 0$.

答:

$$h' = 0$$

$$J' = \frac{1}{2}T \ln \left(\cosh \frac{2J}{T} \right), \quad (36)$$

$$F_0 = -\frac{1}{4}NT \ln \left(4 \cosh \frac{2J}{T} \right).$$

这模型十分简单, H' 和 H 有相同的结构.

6. 在上题, 若 $h \neq 0$, 求 h' , J' , F_0 .
7. 令 $N = 2^m$, 上题模型 H 在重复转换 m 次之后, 即成一常数, 即热位能 F . 求 F . 假定 $2J/T \leq 1$.
8. 假定某一维空间易形磁铁模型为

$$H = -J \sum_{i=1}^{N-1} s_i s_{i+1} (1 - \lambda n_i). \quad (37)$$

$n_i = 0$ 或 1 , λ 为一常数. n_i 可以看成一杂原子粒子数, 即如有一杂原子介于 s_i 和 s_{i+1} 之间, 则作用减弱. 令

$$N' = \sum_{i=1}^N n_i.$$

- (A) 如杂原子的位置不定, 跑来跑去, 求热位能, 熵, 热容率及磁感应.
- (B) 如杂原子不动, 且分布散乱, 求以上各量.

第十九章 静电作用

静电作用在短距离很强, 在长距离虽比较弱, 但同时可以和很多粒子作用, 因此它不能用简单的展开法来分析. 本章用一些例子来指出静电作用的特色, 重点在长距离的作用. 分析方法是由各粒子对某一定电荷的感应着手, 再讨论“屏蔽现象”. 讨论的例子包括离子气体和离子溶液, 金属电子, 并简单介绍二维空间电子晶体的实验. 最后, 我们讨论一抽象的二维空间电荷气体, 电荷间作用能是 $q^2 \ln r$, q 是电荷, r 是二电荷间距离. 这气体在某温度 T_0 以上为离子气体, 在 T_0 以下, 正负离子配成双而成电偶极分子, 这模型在近年统计力学中相当重要, 它的主要特性可以用非常简单的分析来解释.

1. 短距离和长距离都重要

静电作用能, 是和两作用电荷间的距离成反比. 它没有一个特别的距离尺度. 令 e_i 为粒子 i 的电荷, 则一群粒子的静电位能是

$$U = \frac{1}{2} \sum_{i,j} \frac{e_i e_j}{r_{ij}}. \quad (1)$$

因为没有一特别尺度, 我们可以把这物体缩小或放大, 即令 $r_{ij} = \lambda r'_{ij}$, 则 U 在新的尺度中就只多了个 $1/\lambda$ 乘数而已. 如果 U 是负的, 则各粒子会跑到一点去, $U \rightarrow -\infty$. 如果 U 是正的, 则电荷会四散到无穷远去, 以减低能量. 很显然, 其他效应必须存在, 单靠静电作用无法形成稳定的物体.

在第九章, 我们略略提到, 若要有“宏观极限”或“热力学极限”的成立, 必要条件有

(甲) 总电荷为零, 即 $\sum_i e_i = 0$.

(乙) 量子力学的动能.

(丙) 至少有一种电荷 (正或负皆可) 是费米粒子.

条件 (甲) 是指未抵消的电荷会散去 (到物体表面, 如果表面有法子留它). 条件 (乙) 可以避免所有粒子都集中到一点去. 量子论的要求是: 粒子的位置越受限制, 动能越高, 即

$$\text{动能} \sim \frac{\hbar^2}{mr^2} \quad (2)$$

$r =$ 位置的限制尺度. 因此 r 太小时, (2) 会比位能 e^2/r 大. 条件 (乙) 还不够, 还要有独占定律来帮忙, 这是条件 (丙)(见第九章第 2 节).

因此, 讨论静电作用时, 必须和其他效应一同讨论, 或用近似法避开上述的不稳定性. 在统计力学中, 和静电作用直接有关的问题很多. 在这些问题中, 静电作用的主要特色是长距离的作用. 由于 $1/r$ 在 r 很大时仍是可观, 一个粒子可以同时和许多粒子作用, 以下诸节的讨论的重点也是这长距离作用.

2. 离子气体和溶液

假定有 N 个粒子带正电 e , N 个带负电 $-e$, 位能为 (1), 现在不考虑量子力学. 假定粒子间的距离很近时, 含有某种斥力使它们不致于无限制的接近.

现在来看粒子密度的相干函数, 再考虑热位能, 我们照第十三章的方法去做 (见第十三章第 2 节). 先在原点放一正电 e , 看看别的粒子如何分布.

令 $n_{\pm}(\mathbf{r})$ 为正子 (带正电者) 及负子的密度. 令 $\phi(\mathbf{r})$ 为电位. 则

$$\begin{aligned} n_{\pm}(\mathbf{r}) &= ne^{\mp e\phi(\mathbf{r})/T} \\ &\approx \mp ne\phi(\mathbf{r})/T + n. \end{aligned} \quad (3)$$

n 是正子-负子的平均密度.

$\phi(\mathbf{r})$ 和电荷密度 $e(n_+ - n_-)$ 的关系是

$$-\nabla^2\phi(\mathbf{r}) = 4\pi e(n_+(\mathbf{r}) - n_-(\mathbf{r})). \quad (4)$$

将 (3) 代入 (4), 得

$$\begin{aligned} \nabla^2\phi &= \frac{\phi}{b^2}, \\ \frac{1}{b^2} &\equiv \frac{8\pi e^2 n}{T}, \end{aligned} \quad (5)$$

解:

$$\phi(\mathbf{r}) = \frac{e}{r} e^{-r/b}.$$

如果没有其他粒子, 单单一个电荷放在原点, 产生的电位是 e/r . 因此, 其他的粒子的效果是把远处的电位减弱. 现在多了一个尺度 b , 所谓“屏蔽层厚度”. 在原点四周, 有一层电荷, 密度是

$$\begin{aligned} e(n_+ - n_-) &= -\frac{\nabla^2 \phi}{4\pi} \\ &= -\frac{e}{4\pi b^2 r} e^{-r/b}. \end{aligned} \quad (6)$$

这是“屏蔽层”, (6) 的对 $4\pi r^2 dr$ 的积分是 $-e$. 所以, 放在原点的 $+e$ 被遮住了, 在距离 b 之外, 它的影响趋向零. 温度越高, b 越长, (5) 越正确. 注意, 由 (5) 得

$$\frac{8\pi e^2 n^{1/3}}{T} = \frac{1}{(b^3 n)^{2/3}} \quad (7)$$

左边约是作用能 e^2/r 和动能 T 之比, ($r \sim$ 粒子距离 $\sim n^{-1/3}$). 右边分母是在屏蔽层中的粒子数的 $2/3$ 次方. 因此, 以上的近似计算是在 $b^2 n$ 很大时有效, 但仍然是在低密度有效.

以上的 $n_{\pm}(\mathbf{r})$ 是一“条件分布”, 条件是在原点有一正电 e . 因此, 照第十三章的写法, 应该是 $n_{\pm}(\mathbf{r}|0+)$ (见 13.20). 从 $n_{\pm}(\mathbf{r})$, 密度相干函数即可求得:

$$\langle \rho_{\pm}(\mathbf{r}) \rho_{\pm}(0) \rangle_c = \mp \frac{e^2 n^2}{rT} e^{-r/b}. \quad (8)$$

由此可见, 这气体的相干距离为 b . 本来静电作用没有距离尺度, b 这尺度是由密度、温度和电荷组合而成 (见 (5)).

从 (8) 可以算出作用能 (1):

$$\begin{aligned} U &= \frac{V}{2} \int d^3 \mathbf{r} \frac{e^2}{r} \langle (\rho_+(\mathbf{r}) - \rho_-(\mathbf{r})) (\rho_+(0) - \rho_-(0)) \rangle \\ &= -V \frac{T}{8\pi b^3} = -\frac{e^2}{b} N \\ &= -\frac{V(8\pi)^{1/2} e^3 n^{3/2}}{\sqrt{T}}. \end{aligned} \quad (9)$$

从这结果可以求出热容率, 热位能等量. 注意

$$U \propto -\frac{N^{3/2}}{\sqrt{VT}}, \quad E = E_0 + U. \quad (10)$$

读者试证

$$\begin{aligned}\frac{\partial E}{\partial T} &= C_V = C_{0V} - \frac{1}{2} \frac{U}{T}, \\ S &= S_0 + \frac{1}{3} \frac{U}{T}, \\ F &= F_0 + \frac{2}{3} U, \\ p &= p_0 + \frac{1}{3} \frac{U}{V}, \\ \mu &= \mu_0 + \frac{U}{N}.\end{aligned}\tag{11}$$

带 0 的符号是指 $e = 0$ 时的量.

如果这些粒子是某溶液中的溶质, 如盐水中的 Na^+ , Cl^- , 以上结果可以直接应用. 唯一要修正的是电荷. 在水中, 由于水分子有电偶极, 离子的电荷会被遮掉一部分. 以上的 e^2 须除以介电系数 ϵ . 在水中, $\epsilon = 88$, (11) 的各量成了离子的性质. p 成了渗透压 (见第十八章).

在水中, 因为 $\epsilon = 88$ 相当大, 离子间的静电作用减弱了不少. 作用能和动能之比是 (7), 如果要 (7) 很小, 则

$$n \ll \frac{(kT)^3}{8\pi (e^2/\epsilon)^3}.\tag{12}$$

以上把热常数 k 写出来了. 把数字代入, $kT = 1.38 \times 10^{-16} \times 300 \text{ erg}$, $e^2 = (4.8 \times 10^{-10})^2 \text{ erg} \cdot \text{cm}$, $\epsilon = 88$, 得

$$n \ll 1.6 \times 10^{20} / \text{cm}^3.\tag{13}$$

这条件并不很苛刻, 一克分子量的溶质有 6.23×10^{23} 个分子.

这一节的分析本是为离子溶液而作, 是所谓“德拜—休克尔”理论.*

3. 金属电子

上节的分析步骤, 也可以用到金属电子上, 主要的不同点是

- ① 金属中的正电 (即核子) 大致是不动的.
- ② 电子的密度很高, $\epsilon_F \gg T$, 而且静电作用能 $e^2/r \sim e^2 p_F$ 比 T 也大很多. 因此讨论静电作用时, 温度可以先不管.

*Debye-Huckel 理论, 更详细的分析, 请阅物理化学书积.

作用能和动能之比是

$$\alpha \equiv \frac{e^2}{\epsilon_F r} = \frac{2me^2}{p_F} \sim \frac{n^{-1/3}}{a}, \quad (14)$$

$$a \equiv \frac{\hbar^2}{me^2}.$$

因此, 这个比值约是电子间的距离 $n^{-1/3}$ (n 是密度), 除以“波耳半径” a , (即氢原子半径 0.53\AA). 这比值不是小数目, 因此, 金属电子的静电作用不弱, 在以下的分析, 仍然假装作用很弱, 以求计算方便. 所得的结果因而不甚精确.

我们又在原点放一个电荷 e , 看看各电子如何应付 (核子不动, 不反应).

令 $\phi(\mathbf{r})$ 为电位. 静电公式是

$$-\nabla^2\phi = -4\pi en'(\mathbf{r}). \quad (15)$$

n' 是电子密度减掉 N/V , N 是总电子数, V 是体积. 我们现在把不动的正电看作均匀分布的正电, 其总值正好把电子总电荷抵消. $n'(\mathbf{r})$ 是电子在位能 $-e\phi(\mathbf{r})$ 内的密度分布 (减去 N/V). 令 ϵ_F 为顶层电子的动能, 则

$$\epsilon_F + (-e\phi(\mathbf{r})) = \mu, \quad (16)$$

即位能加动能得顶层能量, 在 $T = 0$ 时, 这就是化学势. 假设 $\phi(\mathbf{r})$ 是一个 \mathbf{r} 的平缓函数. 计算电子位子的能量分布, 即能位密度, 可以用短波近似: (见 (3.12)),

$$g(\epsilon) = \frac{1}{Vh^3} \int d^3\mathbf{r} d^3\mathbf{p} \delta\left(\epsilon - \frac{p^2}{2m} + e\phi(\mathbf{r})\right) \quad (17)$$

$$= \frac{1}{V} \int d^3\mathbf{r} g_0(\epsilon + e\phi(\mathbf{r})).$$

$g_0(\epsilon)$ 是 $\phi = 0$ 时的能位密度. $g_0(\epsilon + e\phi)$ 可以解释成在 \mathbf{r} 处的能位密度. 能量标准降了 $e\phi$, 这是因为位能 $-e\phi$ 的缘故.

现在假设 $e\phi(\mathbf{r})$ 很小, 在 \mathbf{r} 处的电子密度是

$$n'(\mathbf{r}) + n = \int_0^\mu d\epsilon g_0(\epsilon + e\phi)$$

$$\approx \int_0^\mu d\epsilon \left[g_0(\epsilon) + e\phi \frac{\partial g_0}{\partial \epsilon} \right], \quad (18)$$

$$n'(\mathbf{r}) = e\phi g_0(\mu) = e\phi \frac{mp_F}{\pi^2}.$$

$g_0(\mu)$ 即第四章的 $g(0)$. 将 (18) 代入 (15), 得

$$\begin{aligned}\nabla^2\phi &= \frac{\phi}{b^2}, \\ \frac{1}{b^2} &\equiv \frac{4e^2mp_F}{\pi} = \frac{2\alpha p_F^2}{\pi} \\ &= \frac{6\pi e^2n}{\epsilon_F}, \\ \phi(\mathbf{r}) &= \frac{e}{r}e^{-r/b}.\end{aligned}\tag{19}$$

这结果和 (5) 自然颇有类似之处. α 是 (14) 的比值.

以上的分析, 假设 α 很小, 但一般的金属的 α 都不小, 因此, 以上的结果只告诉我们静电屏蔽现象很强, 静电作用的有效距离很短.

4. 电子晶体

如果电子密度非常小, 以致于

$$\epsilon_F \ll T \ll e^2/r,\tag{20}$$

则量子力学可以不计, 这群体在低温时该成一晶体. 注意, 我们假设有正电作均匀的分布 (如果不是均匀的分布, 而是一般固体中的晶体排列, 则完全是另一问题). 只有电子可以动, 这均匀的正电分布全属虚构, 在三维空间中没有这样的东西. 但是, 在实验室内, 可以用其他效应来制造这均匀正电的效果. 在二维空间, 这均匀正电可以说是成功地造成了. 现在把有关的实验略提一下.

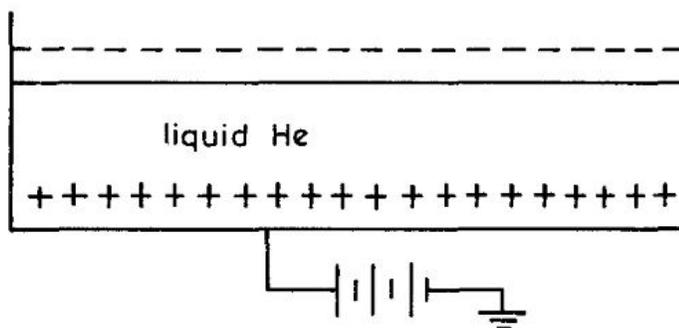


图 1

图一示实验大致情况. 一个平底锅中装了氦液, 温度非常低, (0.5 K 左右) 在液面上洒了一些电子, 这些电子是观测对象. 锅底加了一些正电, 一方面把电子往下拉, 一方面造成一个均匀正电的效果.

电子进不了液内, 因为被氦原子挡住. 液面对电子有点吸引力, 使电子靠在液面上. 电子很轻, 它的波动性使它略在液面之上[†]. 因为温度很底, 氦蒸气压可以略去, 液面上空可算是真空, 液面上的电子成了一个二维空间里的群体. 电子可以在液面上运动, 相互作用是它们之间的静电斥力. 没有杂质. 电子密度可以控制 (从 10^5 至 $10^8/\text{cm}^2$). 这真是和理论中的群体很相近了. 当然, 液面仍是会有振动, 会稍稍影响到电子运动.

如果温度够低, 这些电子就会排成一个晶体. 这晶体又该如何用实验证实呢? 答案是由晶体的振动. 晶体的振动频率和它的结构有关. 实验的步骤是: 在垂直方向多加一微弱的振动电场, 也就是周期性地把电子往液面上按. 这样按就会引起氦液的表面波动, 这表面波动就会把电子摇晃起来. 如果表面波动的频率和波长, 正好和晶体的某一基本振动的相同, 则会引起共振, 将振动电场的能量吸收了. 从共振的频率, 可以推断晶体的存在.

有关这一方面的理论和实验, 近来非常多, 读者有兴趣可自己去阅读[‡].

5. 二维空间的电荷气体模型

现在来讲另一二维空间的“电荷”群体, 或“漩涡”群体, 我们在学静电学时, 总会讲二维空间的公式

$$-\nabla^2\phi = 2\pi q\delta(\mathbf{r}). \quad (21)$$

ϕ 是一个电荷放在原点所造成的电位:

$$\phi(\mathbf{r}) = -q\ln(r/a). \quad (22)$$

a 是一个常数. 在流体力学中, 如有一强度为 q 的漩涡在原点, 则 $\phi(\mathbf{r})$ 是流速位, 即 $\nabla\phi$ 是流速.

[†]电子和液面的结合能是 $0.7 \times 10^{-3}\text{eV}$, 平均位置在液面上 100\AA .

[‡]可以从 Grimes and Adams(1979) 一文着手.

我们现在考虑一群这样的电荷, 它们之间的作用能定为

$$H = -\frac{1}{2} \sum_{i,j} q_i q_j \ln \left(\frac{r_{ij}}{a} \right),$$

$$\sum_i q_i = 0, \quad (23)$$

并令 r_{ij} 不得小于 a . r_{ij} 是 i 和 j 电荷之间的距离. $q_i = +q$, 或 $-q$. 这群体当然不是上节讲的电子群体. 上节的是 e^2/r 交互作用. 如果我们把 (21) 称作“静电”方程式, 则 (23) 也可以名正言顺地叫“电荷”群体了. (23) 这模型在近年来颇被重视, 因为它和一些其他的二维空间模型, 如磁性, 超流性、熔解等, 有密切的关系. 这许多模型的分析, 是由柯斯特立次、陶勒士开始, 现在一般以此二人名字称之[§].

这 $\ln(r/a)$ 的相互作用, 自然是长程作用, 不但是长程, 而且随 r 变大. 这样强的作用, 使第 2, 3 节的分析方法难以胜任. 柯、陶两人的贡献是将这模型的主要特性抓住了. 这特性是: 如果温度低, 正子 (载 q 者) 和负子 (载 $-q$ 者) 会配成对, 形成一个“分子”, 或“电偶极”. 如果温度高过某温度 T_c , 则这些分子就拆散了, 各走各的, 这特性可以大致了解如下.

用第 2 节的想法开始, 在原点放一正子. 这正子会产生电位 $\phi(\mathbf{r})$, 我们现在问: 有没有屏蔽现象?(屏蔽是指完全屏蔽, 即第 2, 3 节那种) 如果在 r 很大时,

$$\phi(\mathbf{r}) \longrightarrow -\frac{q}{\epsilon} \ln(r/a), \quad (24)$$

ϵ 为一常数, 即介电系数, 则我们说没有自由电荷, 只有电偶极, 即正负都配成了双. 如果在 $r \rightarrow \infty$ 时, $1/\epsilon$ 趋向 0, 则我们说有屏蔽现象 (q/ϵ 是没有遮掉的电荷). 图 2a 画了一个圆圈, 半径 r' . 圆心为原点, 放着 q . q/ϵ 是圆圈内的总电荷, 如果分子小, (即配偶间距离小) 则不会造成屏蔽, 因为二极同在圈内, 不会改变圈内电荷. 在圆圈边界的偶极或会跨在边界上, 但这些只造成一有限的 ϵ . 如果要造成屏蔽, 必须要有很多比 r 还大的分子, 这些大分子, 一端在圈内, 一端远在别处, 如这些大分子太多, 则圈内就如多一些独身的正子或负子, 它们的“另一半”离开太远了. 由第 2、3 节的经验, 这些独身的会造成屏蔽现象. 现在来估计一下这些大分子的粒子数.

[§]Kosterlitz and Thouless (1973), Kosterlitz (1973).

令分子的偶极间距离为 r' , 它的概率分布是

$$\begin{aligned} f(r') &\propto e^{+q\phi(r')/T} \sim e^{-(q^2/T)\ln(r'/a)} \\ &= \left(\frac{r'}{a}\right)^{-q^2/T}. \end{aligned} \quad (25)$$

因此, 比 r 大的分子的粒子数密度是

$$\int_r^\infty dr' 2\pi r' f(r'). \quad (26)$$

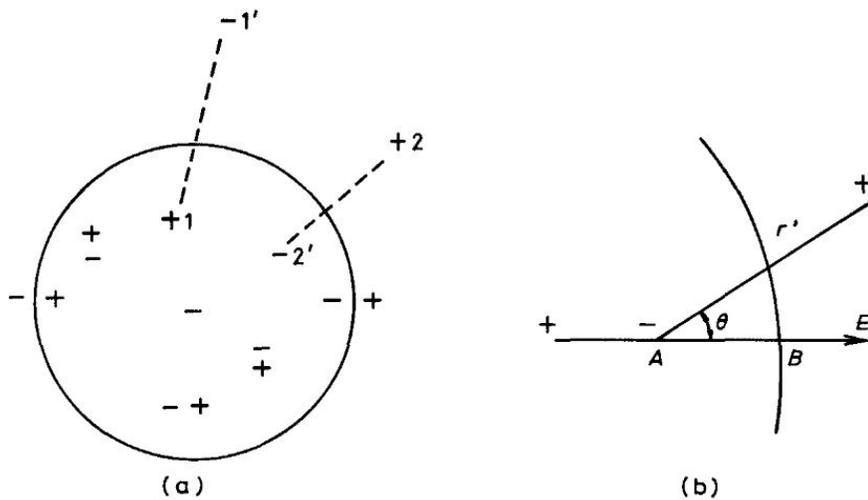


图 2. 小分子可视为电偶极, 但大分子, 如 $11'$, $22'$, 就和独行自由电荷差不多.

在圈内有一端的大分子人数约是

$$\begin{aligned} &\sim \pi r^2 \int_r^\infty dr' 2\pi r' \left(\frac{r'}{a}\right)^{-q^2/T} \\ &\sim r^{4-q^2/T}. \end{aligned} \quad (27)$$

如果这结果在 $r \rightarrow \infty$ 时趋于 0, 则没有屏蔽, 否则就有. 因此, 在

$$T < T_0 = \frac{q^2}{4}, \quad (28)$$

时, 没有屏蔽, 即正负子配成双, 以上的分析, 还需要修正一下. 在 (25), ϕ 用的是 $q\ln r/a$, 但由于偶极间可能有别的分子, 这 ϕ 必须修正, 不过这修正在温度不接近 T_0 时不会很重要.

现在来算一下 $T < T_0$ 时的介电系数 ϵ , 假定每分子都比 r 小很多. 图 2b 画了一分子, 负子在圈内, 正子在外. 我们要计算此类跨界分子放在圈内的总电荷. 假定 $r' \ll r$, 则角度 θ 的概率分布约为

$$\frac{e^{qEr' \cos \theta/T}}{\int_0^\pi d\theta e^{qEr' \cos \theta/T}} \approx \frac{1}{\pi} \left(1 + qEr' \cos \frac{\theta}{T} \right). \quad (29)$$

E 是在 r 处的电场, 并假定 E 很弱. 假定分子的位置分布为均匀, 密度为 n , 则图 2b 中的距离 AB , 平均分布在 0 和 $r' \cos(\theta)$ 之间, 即分布在沿界宽为 $r' \cos(\theta)$ 的环中, 此环的面积是 $2\pi r r' \cos(\theta)$. 所以在圈内的多余电荷为

$$n(-q) \times 2\pi r \times r' \cos \theta. \quad (30)$$

如果 $\cos \theta$ 是负的, 则表示正电在圈内, 负的在外. 对 r', θ 取平均值, 由 (29), 得圈内总电荷

$$-2\pi r E n \langle r'^2 \rangle \frac{q^2}{2T} + q = \frac{q}{\epsilon}. \quad (31)$$

由“高斯定理”, 知 $2\pi r E = 2\pi q/\epsilon$. 因此,

$$\epsilon = 1 + \frac{\pi n \langle r'^2 \rangle q^2}{T}. \quad (32)$$

r'^2 的平均值可以由 (25) 求得:

$$\begin{aligned} \langle r'^2 \rangle &= \frac{\int_0^\infty f(r') r'^3 dr'}{\int_0^\infty f(r') r' dr'} \\ &= a^2 \left[\frac{1 - (2T/q^2)}{1 - (4T/q^2)} \right] = a^2 \left[\frac{1 - (T/2T_0)}{1 - (T/T_0)} \right]. \end{aligned} \quad (33)$$

当 T 接近 T_0 时, 此式不合用.

如果温度高, 则第 1, 2 节的方法可以用. 读者试证, 在 r 很大时

$$\begin{aligned} \phi(\mathbf{r}) &\sim \frac{q}{\sqrt{r/b}} e^{-r/b} \\ \frac{1}{b^2} &\equiv \frac{4\pi n q^2}{T}. \end{aligned} \quad (34)$$

这和 (5) 很像. 这是一个离子气体的性质, 以上的结果 (32), (33), (34) 大致把这模型的主要特性澄清了. 这些都是非常简单的结果, 不需要任何复杂的分析.

如果要知道 T 在 T_0 附近的性质, 则需要比此处稍微复杂一点的计算.

读者可参考文献.[¶]

[¶]Young(1979), 此文比柯、陶二人的文章清楚些.

讨论问题十九

1. 原子的“统计模型”(即“托马斯—费米”模型).

重原子有许多电子, 其密度可以用 (15)–(17) 求出. 核子在原点, 电位 $\phi(\mathbf{r})$ 及电子密度 $n(\mathbf{r})$ 都有球面对称性.

(A) 求 $n(r)$ 和 $\mu, \phi(r)$ 的关系. ($n(r)$ 和 $(\mu - e\phi)^{3/2}$ 成正比).

(B) 由 (15) 得 ϕ 的微分方程式.

(C) 加入条件

$$Z = \int d^3r n(r). \quad (35)$$

以求 μ, Ze 是核子电荷.

以上的各方程式, 并不易解, 可用数值解.

2. 金属表面模型.

把金属中的核子想成平均分布的正电, 在 $x < 0$, 密度 $\frac{N}{V}$, 在 $x > 0$, 密度为 0, 试用上题的方法, 求 $n(x)$, 即电子密度分布. 在 $x \rightarrow -\infty$ 时, $n(x) \rightarrow N/V$. 在 $x = 0$ 左右, $n(x)$ 比较复杂, 这就是金属表面在 $x \rightarrow \infty, n(x) \rightarrow 0$.

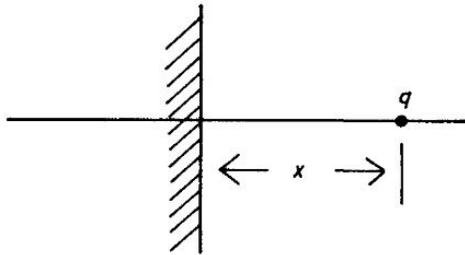


图 3

3. 电荷 q 位于一绝缘体之外 (见图 3), 求电荷所受的力. 绝缘体的介电系数为 ϵ .

提示: 假想绝缘体内有一虚像电荷.

$$q' = -q \frac{\epsilon - 1}{\epsilon + 1}. \quad (36)$$

4. 由上题的结果, 求一电子在 x 的位能 $U(x)$.
- (A) 写下在此位能中的波动方程式. 假定电子不能进入绝缘体.
- (B) 解之得电子的能量及波函数.
- (C) 将 He^4 液体的 ϵ 值代入, 求电子的零点能量, 及其平均位置.
5. 复习讨论问题 7.11, 此题指出在 $T < q^2/2$ 时, Z 不存在.
- 注意, $q^2/2$ 是 $2T_0$, (见 (28)), 不是 T_0 . 不过, 如果限定粒子间距离不得小于 a , 则 Z 在低温亦无问题 (见 (23)).
6. 上题的内容是, 在 $T < 2T_0$ 时, 正负电荷已是在一种束缚状态之下.
- (A) 令电荷 q 在原点, $-q$ 在 r . 令 $T < 2T_0$. 试证 r 为有限的概率为 1, 也就是说, 此二电荷是在束缚态.
- (B) 试证在 $T_0 < T < 2T_0$ 时, 介电系数 ϵ 不存在. 即虽然是在束缚态, 但 $\langle r^2 \rangle \rightarrow \infty$. 这一点由 (33) 看来, 十分明显. 因此, 在 T_0 至 $2T_0$ 这一段温度, 一对正负电荷, 分开得相当远. 因此别的电荷会介入其中. 因此以上 r 为有限的结论就靠不住了. 这一段温度的分析十分复杂.
7. 在 3 度空间, 电荷气体没有像 T_0 这种临界温度. 在低温, 正、负电荷会结合在一起, 在高温成离子气体. 从结合的分子气体到离子气体是一个连续的改变.
- (A) 令电荷 $q_i q_j$ 间的作用为

$$\begin{aligned}
 U(\mathbf{r}) &= \frac{q_i q_j}{r}, \quad r > a, \\
 &= \infty, \quad r \leq a, \\
 q_i &= \pm e.
 \end{aligned}
 \tag{37}$$

令粒子平均密度为 n . 假定 $na^3 \ll 1$. 什么样的情形可看作离子气体? 试作粗略分析.

(B) 令正电荷为核子, 负为电子. 把量子力学考虑进去. 试分析电离程度 (即有多少原子被电离), 和温度、压力的关系.

(C) 从以上结果, 讨论离子溶液, 及第 2 节的结论.

第五篇

变动

基本假设是一个十分有效的分析工具. 它并不直接讨论分子运动. 分子运动的功效只是使这假设成立而已. 但我们一直强调, 这假设在不同的情况下, 内容亦不同, 必须依照运动情况决定. 归根结底, 分子运动决定一切. 这一篇的讨论, 以运动为重点, 但目的是对平衡态作更进一步的认识. 玻尔兹曼的气体运动方程式 (第二十章) 是一个十分重要的题目. 我们在此用它来讨论几个小幅度变动的现象, 如声波及热传导. 第二十一章讨论扩散方程式, 重点在暂稳态. 这两章是一连串的例子. 这些讨论使我们从缓慢变化的现象体会平衡的意义. 第二十二章讨论直接分子运动来求平衡性质的方法. 这些方法是电子计算机问世以后的产物, 是近代物理的一大成就, 不但是计算工具, 而且使我们对统计力学的最基本意义有更深的了解.

第二十章 气体运动方程式

平衡性质, 在基本假设的分析下, 可以解释得十分圆满, 不平衡的性质, 却没有这样的有效的分析工具, 不平衡现象的理论, 可以说还在草创阶段. 气体运动方程式, 或称“玻尔兹曼”方程式, 是一个少有的有用工具. 在这一章中我们举例说明它的一些用法. 这个方程式, 很干净利落地把不规则碰撞的效果放在一边, 把规则的运动效果放在另一边. 规则的运动是平缓的, 碰撞是快速的. 玻尔兹曼方程式已经把这快速的, 不规则的碰撞平均了. 因此, 这方程式中的时间空间尺度都要比分子接触时间、分子大小要大很多. 本章的主要内容是一些例子, 用这些例子来说明运动方程式的含义和一些用法. 例子包括电离振动, 零点声波, 杂质散射, 声波的传播和消失. 最后提一下 H 定理.

1. 流动和碰撞

气体运动方程式, 一般称“玻尔兹曼”方程式, 是讨论理想气体运动的基本工具. 群体运动极其复杂, 即使理想气体都不例外. 玻尔兹曼方程式和他的求熵基本假设, 可以说是前无古人, 后无来者的杰作之一.

这方程式可以由直觉写下, 如果要从牛顿定律或别的力学定律求出, 却是千难万难. 以下就是一个简单的说法, 把这方程式凑出来.

令 $f(\mathbf{r}, \mathbf{p}, t)$ 为一群粒子在时刻 t 的分布函数. 这是在六度 (\mathbf{r}, \mathbf{p}) 空间内的分布. 假定分子数为守恒量, 则守恒方程式是

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial \mathbf{r}} \cdot (\mathbf{v}f) + \frac{\partial}{\partial \mathbf{p}} \cdot (m\mathbf{a}f) = 0. \quad (1)$$

\mathbf{v} 是 \mathbf{p}/m , 即流速, \mathbf{a} 是加速度. $(\mathbf{v}, m\mathbf{a})$ 是在 (\mathbf{r}, \mathbf{p}) 空间的流速, (1) 是一般三维空间粒子守恒公式

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{v}\rho = 0 \quad (2)$$

的推广. ρ 是粒子密度. (1), (2) 假设粒子不会忽然从一点跳到别处去, 只是平缓地流着. 如果把碰撞考虑进去, 则粒子的动量会突然作大幅改变, (1) 就不对了. (粒子位置仍无法作大幅跳动). 现在加上一修正项, 把 (1) 改成

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} + \mathbf{F} \cdot \frac{\partial f}{\partial \mathbf{p}} = \left(\frac{\partial f}{\partial t} \right)_c. \quad (3)$$

$\mathbf{F} = m\mathbf{a}$ 是每粒子所受的外力, 右边是由碰撞引起的修正项, 是每秒钟有多少粒子被撞进 (\mathbf{p}, \mathbf{r}) , 减去有多少从 (\mathbf{p}, \mathbf{r}) 碰到别处去:

$$\begin{aligned} \left(\frac{\partial f}{\partial t} \right)_c &= \int (f'' f''' - f f') R d^3 \mathbf{p}' d^3 \mathbf{p}'' d^3 \mathbf{p}''', \\ R &\propto \delta(\mathbf{p} + \mathbf{p}' - \mathbf{p}'' - \mathbf{p}''') \delta(\epsilon + \epsilon' - \epsilon'' - \epsilon'''), \end{aligned} \quad (4)$$

R 是二粒子反应 $\mathbf{p} + \mathbf{p}' \longleftrightarrow \mathbf{p}'' + \mathbf{p}'''$ 的反应率, 它含有动量、能量守恒的 δ 函数,

$$\begin{aligned} f &\equiv f(\mathbf{r}, \mathbf{p}, t), \\ f' &\equiv f(\mathbf{r}, \mathbf{p}', t), \\ f'' &\equiv f(\mathbf{r}, \mathbf{p}'', t), \\ f''' &\equiv f(\mathbf{r}, \mathbf{p}''', t). \end{aligned} \quad (5)$$

在第三章讨论细致平衡时曾指出正向反应和反向的 R 是一样的. 如果是费米粒子, 且密度大, 则 (4) 中的 $f'' f''' - f f'$ 要改成

$$f'' f''' (1 - f) (1 - f') - f f' (1 - f'') (1 - f'''). \quad (6)$$

这些如不清楚, 请复习第三章第 2 节. f 是每个粒子的粒子数. 在 $d^3 \mathbf{p} d^3 \mathbf{r}$ 中有 $d^3 \mathbf{p} d^3 \mathbf{r} / h^3$ 个位子, 即

$$\frac{f d^3 \mathbf{p} d^3 \mathbf{r}}{h^3}$$

是在 $d^3 \mathbf{p} d^3 \mathbf{r}$ 内的粒子数. 如果 $f \ll 1$, 即量子效应可以略去, 则可以把 $1/h^3$ 包括在 f 的定义中.

在平衡状况下, 碰撞不会改变 f , 因此 (4) 或 (6) 是零. 从此可得平衡粒子分布 (见第三章). 注意, 只要 f 是粒子能量 ϵ 的函数, $f(\mathbf{r}, \mathbf{p}) = f(\epsilon)$,

$$\epsilon \equiv \frac{p^2}{2m} + u(\mathbf{r}), \quad \left(\mathbf{F} = -\frac{\partial u}{\partial \mathbf{r}} \right). \quad (7)$$

即 $f(\mathbf{r}, \mathbf{p}) = f(\epsilon(\mathbf{r}, \mathbf{p}))$, 则 (3) 的左边就是零. 如要使右边消失, 则 $f(\epsilon)$ 必须是特别的函数, 如 $e^{-\alpha - \beta\epsilon}$ 使 (4) 消失, $1/(e^{\alpha + \beta\epsilon} + 1)$ 使 (6) 消失.

注意, $(\partial f/\partial t)_c$ 是和平衡性质无关的. 玻尔兹曼方程式可用来计算任何性质, 包括平衡性质, 但在平衡性质的答案中, R 不该出现. 如果 \mathbf{F} 纯是外力, 则平衡性质即理想气体性质. 不过我们可以把 \mathbf{F} 看成 f 的函数而把相互作用的效果考虑进去一部份.

(3) 式的左边形容缓慢的运动, 右边是不规则, 快速碰撞的效果. 在一体积 $d^3\mathbf{r}$, 时间 dt 内, $(\partial f/\partial t)_c dt d^3\mathbf{p} d^3\mathbf{r}$ 必须是一大数. 每个碰撞是看成一瞬时的事件, 发生于一点. 我们先看些 $(\partial f/\partial t)_c = 0$ 的例子, 再看 $(\partial f/\partial t)_c$ 不是零的例子.

2. 没有碰撞的情形, 离子振动

我们从最简单的例子开始, 假定 $\mathbf{F} = 0$, 且无碰撞, 则

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} = 0. \quad (8)$$

这方程的解是

$$f(\mathbf{r}, \mathbf{p}, t) = \phi(\mathbf{r} - \mathbf{v}t, \mathbf{p}), \quad (9)$$

$\phi(\mathbf{r}, \mathbf{p})$ 是在 $t = 0$ 时的分布. 一粒子在时刻 t 如位置是 \mathbf{r} , 则在 $t = 0$ 时是 $\mathbf{r} - \mathbf{v}t$.

【例】 某气体在 $t = 0$ 时集中在原点, 温度为 T , 求以后的密度分布 $n(\mathbf{r}, t)$.

$$\begin{aligned} n(\mathbf{r}, t) &= \int d^3\mathbf{p} f(\mathbf{r}, \mathbf{p}, t) \\ &= \int d^3\mathbf{p} \phi(\mathbf{r} - \mathbf{v}t, \mathbf{p}), \end{aligned} \quad (10)$$

$$\phi(\mathbf{r}, \mathbf{p}) = \frac{N}{(2\pi mT)^{3/2}} e^{-p^2/2mT} \delta(\mathbf{r}).$$

因此,

$$n(\mathbf{r}, t) = \left(\frac{m}{t}\right)^3 \frac{N}{(2\pi mT)^{3/2}} e^{-\frac{m\mathbf{r}^2}{2Tt^2}}. \quad (11)$$

气体散开, 因为各分子速度不同.

【例】 电离振动 (平面)

现在看一个比较复杂的例子, 即上一章讲的平面上的电子 (见上一章第 4 节), 只是现在假设电子成一很稀的气体. 令 $n'(\mathbf{r}, t)$ 为电子密度分布减去平均

密度 N/L^2 , L^2 为平面面积, 则

$$\begin{aligned} \mathbf{F}(\mathbf{r}, t) &= e\nabla\phi, \\ \frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f + \mathbf{F} \cdot \frac{\partial f}{\partial \mathbf{p}} &= 0, \\ \phi(\mathbf{r}, t) &= - \int d^2r' \frac{e}{|\mathbf{r} - \mathbf{r}'|} n'(\mathbf{r}', t). \end{aligned} \quad (12)$$

现在假定

$$\begin{aligned} f &= f_0 + f' e^{-i\omega t + i\mathbf{k} \cdot \mathbf{r}} \\ f_0 &\equiv \frac{n}{(2\pi m T)} e^{-p^2/2mT}, \\ f' &\ll f_0, \end{aligned} \quad (13)$$

其他各量 ϕ , n' 也是平面波, 并且是很小的量. 从 (12), (13) 得

$$-i(\omega - \mathbf{k} \cdot \mathbf{v})f' + i \frac{2\pi e^2 n' \mathbf{k} \cdot \mathbf{v} f_0}{k T} = 0. \quad (14)$$

注意

$$\begin{aligned} \int d^2\mathbf{r} \frac{1}{r} e^{-i\mathbf{k} \cdot \mathbf{r}} &= \frac{2\pi}{k}, \\ \frac{\partial f_0}{\partial \mathbf{p}} &= -\mathbf{v} \frac{f_0}{T}. \end{aligned} \quad (15)$$

由 (14) 得

$$\int d^2\mathbf{p} f' = n' = n' \frac{2\pi e^2}{k} \int d^2\mathbf{p} \frac{\mathbf{k} \cdot \mathbf{v} f_0 / T}{\omega - \mathbf{k} \cdot \mathbf{v}}. \quad (16)$$

因此, 振动的条件是

$$1 = \frac{2\pi e^2}{k} \int d^2\mathbf{p} \frac{\mathbf{k} \cdot \mathbf{v} f_0 / T}{\omega - \mathbf{k} \cdot \mathbf{v}} \quad (17)$$

在 $\frac{k}{\omega}$ 很小时, 将右边对 $\frac{\mathbf{k} \cdot \mathbf{v}}{\omega}$ 展开, 得

$$\begin{aligned} 1 &= \frac{2\pi e^2}{k} \int d^2\mathbf{p} \frac{(\mathbf{k} \cdot \mathbf{v})^2}{\omega^2} f_0 / T + \dots \\ &\approx \frac{2\pi e^2 n}{m} \cdot \frac{k}{\omega^2}, \end{aligned} \quad (18)$$

$$\omega^2 = \frac{2\pi e^2 n}{m} k. \quad (19)$$

因此, 振动频率和 \sqrt{k} 成正比, 还有 $O(k^{3/2})$ 的虚数项, 代表衰减率. 在 k 很小时 k/ω 的确很小, k/ω 的展开因此行得通.

(19) 这结果不只合用于平面上的电子气体, 也合用于平面上的电子晶体. 以上没有考虑碰撞, 但是即使有碰撞, (19) 仍然成立, (19) 可以直接由流体力学公式求出 (见讨论问题 1, 2, 3). (19) 的特点是: 它的来源是电子间的静电作用, 这作用产生很强的回复力量, 和固体的弹性力相似.

注意, (12) 式的力 \mathbf{F} 是一种“平均场”. 由于粒子的运动, 粒子受的力会作快速改变, 而 (12) 的 \mathbf{F} 没有包括这种改变, 它只包括了缓慢的 $n'(\mathbf{r}, t)$ 的改变, t 的尺度比碰撞的时间尺度要长很多. 当然, $(\partial f/\partial t)_c$ 的目的是将碰撞的效果考虑进去, 但并未把粒子密度的相干考虑进去.

3. 零点声波

用上例的方法, 可以解释费密粒子气体的“零点声波”现象. 这种声波相当于上例的电离振动, 不过作用是短程作用. 在实验室中, 不带电的、费密粒子物体, 只有在低温的 He^3 液体符合 $T/\epsilon_F \ll 1$ 的条件. (He^3 液体在常压下的 p_F 约是 10^{-8}cm^{-1} , ϵ_F 约为 5 K, 相互作用很强, 但其性质和自由粒子仍很相像. 这是俄人兰道的分析结果). 零点声波是在 T/ϵ_F 很小时才能发生的现象.

现在仍照 (12) 的做法, (12), (13) 现在改成

$$\begin{aligned} \mathbf{F} &= -\nabla\phi, \\ \frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f + \mathbf{F} \cdot \frac{\partial f}{\partial \mathbf{p}} &= 0, \\ \phi &= \int d^3\mathbf{r}' u(\mathbf{r} - \mathbf{r}') n'(\mathbf{r}', t), \\ f &= f_0 + f' e^{-i\omega t + i\mathbf{k} \cdot \mathbf{r}}, \\ f_0 &= \frac{1}{(e^{\epsilon_p/T} + 1)}, \\ \epsilon_p &\equiv \frac{p^2}{2m} - \epsilon_F. \end{aligned} \tag{20}$$

$u(\mathbf{r})$ 是二粒子间的相互作用能, \mathbf{r} 是二粒子相对位置. 现在假定 $T \rightarrow 0$, 则 $f_0 = \theta(-\epsilon_p)$, (14), (15) 成了

$$\begin{aligned} (\omega - \mathbf{k} \cdot \mathbf{v}) f' - n' u_{\mathbf{k}} \mathbf{k} \cdot \mathbf{v} \delta(\epsilon_p) &= 0, \\ \int d^3\mathbf{r} u(\mathbf{r}) e^{-i\mathbf{k} \cdot \mathbf{r}} &= u_{\mathbf{k}}, \end{aligned} \tag{21}$$

$$\frac{\partial f_0}{\partial \mathbf{p}} = -\mathbf{v} \delta(\epsilon_p). \quad (22)$$

(16), (17) 也是依样画葫芦, 得

$$\begin{aligned} n' &= 2 \int \frac{d^3 \mathbf{p}}{(2\pi)^3} f' \\ 1 &= 2u_{\mathbf{k}} \int \frac{d^3 \mathbf{p}}{(2\pi)^3} \frac{\mathbf{k} \cdot \mathbf{v}}{\omega - \mathbf{k} \cdot \mathbf{v}} \delta(\epsilon_p) \\ &= u_{\mathbf{k}} g(0) \int \frac{d\Omega}{4\pi} \frac{kv_F \cos \theta}{\omega - kv_F \cos \theta}, \\ d\Omega &\equiv 2\pi \sin \theta d\theta. \end{aligned} \quad (23)$$

$g(0) = mp_F/\pi^2$ 是在顶层的能位密度. (见第三、四章). (23) 的积分很简单, 得

$$\begin{aligned} 1 &= u_{\mathbf{k}} g(0) \left[\frac{1}{2} \ln \frac{s+1}{s-1} - 1 \right], \\ \omega &= v_F s k. \end{aligned} \quad (24)$$

振动频率可以由此式解出, 在作用不强时, 即 $u_{\mathbf{k}} g(0)$ 很小时, s 和 1 很近:

$$\begin{aligned} \frac{1}{u_{\mathbf{k}} g(0)} &\approx \frac{1}{2} \ln \frac{2}{s-1}, \\ s &\approx 1 + 2e^{-2/u_{\mathbf{k}} g(0)}. \end{aligned} \quad (25)$$

波速是 sv_F , 和 v_F 很近. 这个振动就是“零点声波”^{*}.

这零点声波和“普通的”声波有何不同? 固体的声波, 是籍原子间的作用力传播, 即是一种弹性波, 和这零点声波相似. 晶体中的音子 (见第六章), 可以叫做零点声波, 也是“普通”的声波.

如果我们用“普通”声波公式

$$c^2 = \frac{\partial p}{\partial n}. \quad (26)$$

来计算费密粒子气体在 $T \rightarrow 0$ 时的音速 c , 则得 (见第四章第 1 节的零点压力公式),

$$c = \frac{v_F}{\sqrt{3}}. \quad (27)$$

这结果当然和 (24), (25) 不同, 仔细分析一下, 就会发现这“普通声波”是外层电子和顶层下的空子气体里的声波, 也就是说, 在低温时, 费密粒子群体可以

^{*}更详细的讨论, 见 Lifshitz and Pitaevskii(1980).

看成一稀疏的空子加外电子的气体, 这气体中的声波有速度 $v_F/\sqrt{3}$. 这声波和第六章讲的“第二声波”很像. 第二声波是在晶体中的音子气体中的声波, 速度是

$$c_2 = \frac{c_0}{\sqrt{3}}. \quad (28)$$

c_0 是音子速度. 这和 (27) 来自同一原理, 即由碰撞来传播.

因此, 我们最好不要理会“第二”, “普通”, (一般称“普通”为“第一”), 等名词, 直接去了解振动原理. 这些名词的科学价值有限.

零点声波是由相互作用 $u(\mathbf{r})$ 促成, 即使没有碰撞, 这声波仍可以传播. 像 (27), (28) 所指的声波, 又是如何给碰撞传播呢? 靠碰撞传播的意思是: 如无碰撞, 则不能传播, 读者看了第 5 节后, 就知道这句话的含义.

4. 碰撞和扩散

现在来看碰撞的功效, 这是非常复杂的问题. 因此, 我们暂时不谈声波和粒子间的碰撞, 而先看一个比较简单的问题, 即粒子和固定的杂质间的碰撞.

假定有很多杂原子, 固定在某金属中, 自由电子的运动方程式是 (3), 碰撞项是

$$\begin{aligned} \left(\frac{\partial f}{\partial t}\right)_c &= \frac{1}{4\pi} \int d\Omega' \gamma (f' - f), \\ f' &\equiv f(\mathbf{r}, \mathbf{p}', t), \\ \mathbf{p}' &= p. \end{aligned} \quad (29)$$

\mathbf{p} 和 \mathbf{p}' 只是方向不同. $d\Omega'$ 是对 \mathbf{p}' 方向的积分, γ 是散射率, 和杂质密度成正比, 我们假设 γ 不是散射角的函数, 而只是一常数. 注意

$$f(1 - f') - f'(1 - f) = f - f'. \quad (30)$$

(29) 比 (4) 要简单不少.

假定 f 和平衡分布 f_0 差不多, 而且所差是一长波的平面波:

$$\begin{aligned} f &\equiv f_0 + f_1 = [1 + \psi(\mathbf{p}, t)e^{i\mathbf{k}\cdot\mathbf{r}}] f_0, \\ \psi &\ll 1. \end{aligned} \quad (31)$$

带入 (3), (29) 得

$$\begin{aligned} \frac{\partial \psi}{\partial t} &= -(K_1 + K) \psi, \\ K_1 &\equiv i\mathbf{k} \cdot \mathbf{v}, \quad \mathbf{v} \equiv \frac{\mathbf{p}}{m}, \end{aligned} \quad (32)$$

$$\begin{aligned}
 -K\psi &\equiv \frac{\gamma}{4\pi} \int d\Omega' (\psi' - \psi), \\
 \psi' &\equiv \psi(\mathbf{p}', t).
 \end{aligned}
 \tag{33}$$

因为我们假设粒子的能量不因碰撞而改变, 因此只有 \mathbf{p} 的方向是一变数, 其大小 p 是一常数. ψ 可以看作方向 $\hat{\mathbf{p}}$ 的函数.

现在用线性空间的技巧来解 (32). 集所有 $\hat{\mathbf{p}}$ 的函数为空间. 二函数 ϕ, ψ 的内积定为

$$\begin{aligned}
 (\psi, \phi) &\equiv \frac{1}{4\pi} \int d\Omega \psi(\hat{\mathbf{p}})\phi(\hat{\mathbf{p}}) \\
 &= (\phi, \psi).
 \end{aligned}
 \tag{34}$$

现在要求 $K_1 + K$ 的本征值 γ_α 及本征向量 ψ_α :

$$(K_1 + K)\psi_\alpha = \gamma_\alpha\psi_\alpha. \tag{35}$$

(32) 的解是

$$\begin{aligned}
 \psi &= \sum_{\alpha} a_{\alpha}(t)\psi_{\alpha}, \\
 a_{\alpha}(t) &= a_{\alpha}(0)e^{-\gamma_{\alpha}t}.
 \end{aligned}
 \tag{36}$$

因为我们假定 k 很小, K_1 可以用展开式考虑, 我们先求 K 的本征量. 很明显, 由 (33) 可见

$$\psi_{00} = 1, \quad \gamma_{00} = 0, \quad \text{即} \quad K\psi_{00} = 0, \tag{37}$$

是一解, 而且, 任何积分为 0 的 $\psi_{0\alpha}$ 都是 K 的本征向量, 本征值为 γ :

$$\begin{aligned}
 K\psi_{0\alpha} &= \gamma\psi_{0\alpha}, \\
 (\psi_{00}, \psi_{0\alpha}) &= \int \frac{d\Omega}{4\pi} \psi_{0\alpha} = 0, \\
 \psi_{0\alpha} &= \sqrt{4\pi}Y_{lm}, \quad \alpha = (l, m), \quad l \neq 0.
 \end{aligned}
 \tag{38}$$

也就是说, K 只有两个本征值, 0 和 γ . 0 的本征向量是 1. 而 γ 的有无限多个本征向量.

现在求对 K_1 的展开, 即对 k 的展开, 本征值的一次项是零, 因为 K_1 是和 $\mathbf{v} = \mathbf{p}/m$ 成正比, 而 $\psi_{0\alpha}^2$ 有 $\pm\mathbf{p}$ 的对称性. 所以

$$\gamma_{1\alpha} = (\psi_{0\alpha}, K_1\psi_{0\alpha}) = 0. \tag{39}$$

下一步是看二次项, 由初等量子力学的经验, γ_0 的二次修正项可以立即写下:

$$\gamma_{20} = \sum_{\alpha \neq 0} \frac{(\psi_{00}, K_1 \psi_{0\alpha})^2}{0 - \gamma_{\alpha 0}}. \quad (40)$$

这合计中唯一非零的项来自 $\phi_{0\alpha} = \sqrt{4\pi} Y_{10} = \sqrt{3} \cos \theta$. θ 是 \mathbf{k} 和 \mathbf{p} 的夹角,

$$\begin{aligned} (\psi_{00}, K_1 \sqrt{3} \cos \theta) &= \frac{i\sqrt{3}}{4\pi} \int d\Omega k v \cos^2 \theta \\ &= \frac{i}{\sqrt{3}} k v. \end{aligned} \quad (41)$$

代入 (40), 得

$$\begin{aligned} \gamma_0 = \gamma_{20} &= -\frac{k^2 v^2}{3\gamma} \equiv -D k^2, \\ \gamma_\alpha &= \gamma + O(k^2), \quad \alpha \neq 0. \end{aligned} \quad (42)$$

如果 \mathbf{k} 很小, γ_α , $\alpha \neq 0$, 几乎就是 γ , 我们不必管它的 $O(k^2)$ 项, 但是 $O(k^2)$ 是 γ_0 的全部, 非管不可. (36) 因此是

$$\begin{aligned} \psi &\approx a_0(0) e^{-Dk^2 t} + \sum_{\alpha} a_\alpha(0) e^{-\gamma t}, \\ D &= \frac{v^2}{3\gamma} \equiv \frac{v^2 \tau}{3}. \end{aligned} \quad (43)$$

$1/\gamma$ 是平均自由时间 τ . 因此, D 是和 τ 成正比. (43) 的最后一项消失得快. $\tau = 1/\gamma$ 是它的时间尺度. 但右边第一项, 在 k 很小时, 消失得十分慢. 这自然是粒子数守恒的结果. 在 $k=0$ 时, 这一项代表总粒子数的改变. 一旦改了, 就不会因碰撞而消失, 由 (31), (43) 可见, 如果 $t \gg \tau$, 则 f_1 适合扩散公式

$$\frac{\partial}{\partial t} f_1 = D \nabla^2 f_1. \quad (44)$$

此地的 f_1 代表一群能量相同的粒子. 因为我们假定杂质不动, 粒子碰撞前后的能量不会变, 因此, 每个粒子的能量都是不变量.

以上的分析指出扩散和碰撞的关系. 我们没有考虑杂质的振动、晶体的振动或电子间的碰撞. 这全是为了简化计算. 现在这分析步骤已大致说清楚. 我们可以用它来看比较复杂的情形.

5. 碰撞和声波, 粘性和热传导

现在用运动方程式 (3), (4) 来仔细分析碰撞和声波的关系, 这纯是理想气体的分析, 略去作用能, 只考虑碰撞.

现在把 f 写成

$$f = f_0 + f_1 = [1 + \psi(\mathbf{p}, t)e^{i\mathbf{k}\cdot\mathbf{r}}] f_0(\mathbf{p}). \quad (45)$$

f_0 见 (13). 假定 $\psi \ll 1$, 将 (45) 代入 (1), (4) 得

$$\begin{aligned} \frac{\partial \psi}{\partial t} &= -(K_1 + K)\psi, \\ K_1 &\equiv i\mathbf{k} \cdot \mathbf{v}, \end{aligned} \quad (46)$$

$$K\psi \equiv \int d^3\mathbf{p}' d^3\mathbf{p}'' d^3\mathbf{p}''' f_0' R(\psi + \psi' - \psi'' - \psi'''), \quad (47)$$

$$\begin{aligned} f_0' &\equiv f_0(\mathbf{p}'), \\ \psi' &\equiv \psi(\mathbf{p}', t) \\ \psi'' &\equiv \psi(\mathbf{p}'', t), \quad \text{等.} \end{aligned} \quad (48)$$

现在用上节的技巧来解 (46), 集所有 \mathbf{p} 的函数为空间. 二向量 ψ, ϕ 的内积定为

$$\begin{aligned} (\phi, \psi) &\equiv \int d^3\mathbf{p} f_0(\mathbf{p}) \phi(\mathbf{p}) \psi(\mathbf{p}) \\ &= (\psi, \phi). \end{aligned} \quad (49)$$

如能求得 $K_1 + K$ 的本征值 γ_α 及本征向量 ψ_α , 即

$$(K_1 + K)\psi_\alpha = \gamma_\alpha \psi_\alpha. \quad (50)$$

则 (46) 的解是

$$\begin{aligned} \psi &= \sum_{\alpha} a_{\alpha}(t) \psi_{\alpha}, \\ a_{\alpha}(t) &= a_{\alpha}(0) e^{-\gamma_{\alpha} t}. \end{aligned} \quad (51)$$

因此, 须要求 ψ_{α} 及 γ_{α} . K_1 很简单, 但 K 很复杂, K 的本征值大小可由 $(\psi, \mathbf{K}\psi)$ 估计:

$$(\psi, \mathbf{K}\psi) = \frac{1}{4} \int d^3\mathbf{p} d^3\mathbf{p}' d^3\mathbf{p}'' d^3\mathbf{p}''' f_0 f_0' R(\psi + \psi' - \psi'' - \psi''')^2. \quad (52)$$

这式子由 (47) 代 (49) 中的 ϕ 得来, 注意能量和动量的守恒 (见 (4) R 内的 δ 函数). 由 (52) 可见, K 的本征值没有负的. (读者试证 $(\psi, \mathbf{K}\phi) = (\phi, \mathbf{K}\psi)$). R 是碰撞率, 故 K 的本征值大致和 $1/\tau$ 相若, τ 是平均自由时间.

现在要考虑的是长波运动, 即 k 很小. 我们把 $\gamma_\alpha, \psi_\alpha$ 对 K_1 展开:

$$\begin{aligned}\gamma_\alpha &= \gamma_{0\alpha} + \gamma_{1\alpha} + \dots, \\ \psi_\alpha &= \psi_{0\alpha} + \psi_{1\alpha} + \dots,\end{aligned}\tag{53}$$

$$K\psi_{0\alpha} = \gamma_{0\alpha}\psi_{0\alpha},$$

$$\gamma_{1\alpha} = O(k),\tag{54}$$

$$\gamma_{2\alpha} = O(k^2), \quad \text{等等.}$$

如果每个 $\gamma_{0\alpha}$ 都和 $1/\tau$ 相若, 则不会有长时间的振动了, K 的最大特点是: 由于粒子数、能量、动量的守恒 (见 (4) 的 δ 函数), K 有五个本征向量, 其本征值为零:

$$\begin{aligned}\psi_{00} &= 1/\sqrt{A_0}, \\ \psi_{0i} &= p_i/\sqrt{A_i}, \quad i = 1, 2, 3,\end{aligned}\tag{55}$$

$$\psi_{04} = (\epsilon - \bar{\epsilon})/\sqrt{A_4}.$$

$$K\psi_{0\alpha} = 0,\tag{56}$$

$$(\psi_{0\alpha}, \psi_{0\beta}) = \delta_{\alpha\beta}, \quad \alpha, \beta = 0, 1, 2, 3, 4.$$

以上 $\sqrt{A_\alpha}, \bar{\epsilon}$ 是常数, 使 (56) 成立:

$$\begin{aligned}A_0 &= n, \\ A_{1,2,3} &= n \langle p_i^2 \rangle = nmT, \\ A_4 &= n \langle (\epsilon - \bar{\epsilon})^2 \rangle = \frac{3}{2}nT^2, \\ \bar{\epsilon} &\equiv \langle \epsilon \rangle = \frac{3}{2}T.\end{aligned}\tag{57}$$

平均值 $\langle \dots \rangle$ 是对 f_0 的平均值, 这五个 $\gamma_{0\alpha} = 0, \alpha = 0, 1, \dots, 4$, 是最低的本征值了. 现在要看 $\gamma_{1\alpha}, \alpha = 0, 1, \dots, 4$. 这 $\gamma_{1\alpha}$ 就是 5×5 矩阵

$$K_{1\alpha\beta} \equiv (\psi_{0\alpha}, K_1\psi_{0\beta})\tag{58}$$

的本征值. 令 \mathbf{k} 的方向为第 3 方向, 则

$$K_1 = i\mathbf{k} \cdot \mathbf{v} = \frac{ikp_3}{m}.\tag{59}$$

显然 $K_{1\alpha\beta} = 0$, $\alpha, \beta = 1, 2$, 因为 p_1, p_2 和 p_3 是垂直的方向. 所以此二方向的 γ_1 为零. 其他三个 $\psi_{0\alpha}$, $\alpha = 0, 3, 4$, 造成一 3×3 矩阵

$$K_1 = ik \begin{pmatrix} 0 & u_0 & 0 \\ u_0 & 0 & u_4 \\ 0 & u_4 & 0 \end{pmatrix},$$

$$u_0 = \frac{1}{m} \left(\frac{A_3}{A_0} \right)^{1/2} = \left(\frac{T}{m} \right)^{1/2},$$

$$u_4 = \frac{2}{3} \left(\frac{A_4}{A_3} \right)^{1/2} = \left(\frac{2T}{3m} \right)^{1/2}. \quad (60)$$

K_1 的本征值是

$$0, \quad \pm ikc,$$

$$c \equiv \sqrt{u_0^2 + u_4^2} = \left(\frac{5T}{3m} \right)^{1/2}. \quad (61)$$

因此, 除了 $\gamma_{1\pm} = \pm ikc$ 之外, 其他三个 $\gamma_{1\alpha}$ 都是零, 下一次项 $\gamma_{2\alpha}$ 就比较麻烦了,

$$\gamma_{2\alpha} = \sum_{\beta>4} \frac{K_{1\alpha\beta}^2}{-\gamma_{0\beta}}. \quad (62)$$

这计算需要知道 K 的其他本征值及向量. 因此必须用到 R 的细节. 因为 $\gamma_{0\beta} \sim 1/\tau$, $k_1 \sim ikv$, 因此, 大致

$$\gamma_{2\alpha} \sim k^2 v^2 \tau \sim k^2 \tau T/m. \quad (63)$$

我们在此不企图计算 (62), 且以 (63) 为粗略答案, 总结如下

$$\begin{aligned} \gamma_{\pm} &= \pm ick + D_3 k^2, \\ \gamma_{1,2} &= D_1 k^2, \\ \gamma_4 &= D_4 k^2, \\ \gamma_{\beta} &\sim 1/\tau, \quad \beta > 4, \\ D_{1,3,4} &\sim v^2 \tau \sim \tau T/m. \end{aligned} \quad (64)$$

ψ_{\pm} 造成周期性运动, 是向 $\pm \mathbf{k}$ 方向传播的声波

$$f_1 \propto e^{-\gamma_{\pm} t} e^{i\mathbf{k} \cdot \mathbf{r}} = e^{i\mathbf{k} \cdot \mathbf{r} \pm ickt - D_3 k^2 t},$$

则

$$\frac{\partial^2}{\partial t^2} f_1 = c^2 \nabla^2 f - D_3 \frac{\partial}{\partial t} \nabla^2 f_1. \quad (65)$$

$D_3 k^2$ 是消失率, 因为 $D_3 \propto \tau$, 故 τ 越大, 消失得越快. 如无碰撞, 则 $\tau \rightarrow \infty$, 则无声波. 这就是说声波靠碰撞来传播. 这结果相当明显, 如无碰撞, 则粒子走直线, 不会有波动产生. ψ_{\pm} 是 $\psi_{00}, \psi_{03}, \psi_{04}$ 的组和, 因此是粒子密度振动, 能量振动和动量 (沿 \mathbf{k} 方向) 振动的组和.

ψ_{12} 是粒子动量 (垂直于 \mathbf{k}) 的变化, 是一种“切动”或“扭动”. D_1 和粘性成正比. 如

$$f_1 \propto \psi_{1,2} e^{i\mathbf{k}\cdot\mathbf{r} - \gamma_{1,2} t} \sim e^{i\mathbf{k}\cdot\mathbf{r} - D_1 k^2 t},$$

则

$$\frac{\partial}{\partial t} f_1 = D_1 \nabla^2 f_1. \quad (66)$$

ψ_4 是能量的变化, 也带一点粒子密度变化, 如

$$f_1 \propto \psi_4 e^{i\mathbf{k}\cdot\mathbf{r} - \gamma_4 t} \sim e^{i\mathbf{k}\cdot\mathbf{r} - D_4 k^2 t},$$

则

$$\frac{\partial}{\partial t} f_1 = D_4 \nabla^2 f_1, \quad (67)$$

这是热扩散或传导, 以上 (65)—(67) 中写下的微分方程式, 都可以看作 (64) 的另一种表达方式. (64) 总结了理想气体的声波, 粘性和导热性质. 各系数原则上都可以用 (62) 计算.

以上分析指出, 像 (65)—(67) 这类的方程式, 其时间尺度必须比 τ 大很多, 且空间尺度比平均自由距离 l 大很多:

$$\begin{aligned} kv &\ll \frac{1}{\tau}, \\ \frac{1}{k} &\gg \tau v = l. \end{aligned} \quad (68)$$

更进一步的分析, 请阅文献[†].

现在把这节的结果和上两节的比较一下. 这一节的分析, 是把碰撞看成主要部分, 在 (46), K 看成大值, K_1 看成小量. 虽然结果是以 $K = 0$ 的空间为主, 但如没有 K , 本节的展开就无法成立. 这是一个以 $kv/(1/\tau) = kl$ 为小量的展开. 在 2, 3 两节没有考虑碰撞, 而考虑了由粒子密度改变引起的力.

[†]可从 Forster(1975) 着手.

这一节是讲的理想气体. 但同样的分析可用于高密度费米气体, 结论也相似. 但有一个很大的不同点, 在低温的费米气体, 空子和外层粒子的粒子数是和 T 成正比, 不是常数. 因此, 在低温时粒子间没有什么碰到的机会. 即使碰到了还要受独占定律的限制. 不是空的位子, 粒子不许去. 也就是说, 两个粒子相遇的机会是和 T 成正比, 它们的能量约是 T . 如果要碰撞, 撞后的总能量是 T , 各人的能量也必是 T , 不可以一个比 T 大很多一个比 T 小很多. 因此比 T 小很多的地方不许去. 所以撞后的状态也受限制. 因此,

$$\frac{1}{\tau} \propto T^2 \quad (69)$$

因此, 当 $T \rightarrow 0$ 时, 就无法用碰撞来传播声波. 只有零点声波可以传播了, 它用不着碰撞.

6. H -定理

我们略提一下玻尔兹曼的 H 定理, 这定理说, 根据他的方程式 (3), 且 $(\partial f/\partial t)_c \neq 0$, 则

$$H \equiv \int d^3\mathbf{p} d^3\mathbf{r} f \ln \left(\frac{1}{f} \right) \quad (70)$$

必随时间增加. 这定理的证明很简单. 其实在上一节已经证了八九分. 在上节, 我们指出算符 K 的本征值全是正的. (见 (52)). 这已经证明了 $f_1 \equiv f - f_0$ 会趋于零, 即 f 会趋向 f_0 , 而 H 在 $f = f_0$ 时为极大值. (这也很容易证明). 虽然这证明只限于很小的 f_1 , 但已经把要点指出了.

读者立即会指出 H 即是熵 (只差一常数). 我们在此提醒一句: 如果碰撞频繁, H 是熵, 否则不是.

注意, 即使没有碰撞, 也可能会渐趋向 f_0 . 一个开始时不均匀的分布因为粒子速度不同, 会渐渐变成均匀. 例如在 $t = 0$ 时, 如

$$f(\mathbf{r}, \mathbf{p}, 0) = f_0(\mathbf{p}) (1 + ae^{i\mathbf{k}\cdot\mathbf{r}}),$$

则

$$f(\mathbf{r}, \mathbf{p}, t) = f_0(\mathbf{p}) (1 + ae^{i\mathbf{k}\cdot\mathbf{r}} e^{-i\mathbf{k}\cdot\mathbf{v}t}). \quad (71)$$

在 t 很大时, 最后一项对 \mathbf{p} 的任何积分都趋向于零, 因为它的改变周期趋向于 ∞ , 说 f 趋向于 f_0 并没有错. 但我们不能说 $H(f)$ 趋向 $H(f_0)$, 即熵并未增

加. 照我们的看法, 若要用 H 作熵, 必须要有碰撞, 使粒子的动量不断地改变. 如无碰撞, H 就不是熵, 因为在动量空间里, 即 $(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N)$ 的 $3N$ 度空间里, 轨迹只是一点, 唯有碰撞, 才能使轨迹分散开来, 如无碰撞, 每个 \mathbf{p}_i 都是不变量, H 自然不是熵. 注意, 第 4 节讲的那种碰撞, 也不能使 H 成为熵, 因为它不能改变粒子的能量.

当然, 把 (71) 代入 (70) 再求 dH/dt , 我们会发现 $dH/dt = 0$ (如果没有碰撞). 这是因为 (71) 的最后一项虽然快速改变, 但仍能使 dH/dt 消失.

在第二十三章, 我们讲“重聚现象”时, 会再讨论碰撞和熵的问题. 这现象大略指出粒子分散开了, 熵不一定增加, 在第二十四章, 我们会仔细讨论如何用运动轨迹来计算熵. 这些都是十分重要的问题.

讨论问题二十

1. 电离振动在每本离子气体、固态物理的书中都会谈到.

假定正电荷不动, 电子作小幅振动.

(A) 试从流体力学公式导出

$$\begin{aligned} \frac{\partial n'}{\partial t} + n \nabla \cdot \mathbf{v} &= 0, \\ m \frac{\partial \mathbf{v}}{\partial t} &= -e \mathbf{E}, \\ \nabla \cdot \mathbf{E} &= -4\pi e n'. \end{aligned} \quad (72)$$

以上 n 是平均密度, n' 是电子密度的改变, \mathbf{v} 是流速. 讨论 (72) 的可靠限度, 注意碰撞的作用.

(B) 解 (72), 求振动频率, 得

$$\omega_p = \left(\frac{4\pi e^2 n}{m} \right)^{1/2}. \quad (73)$$

2. 上题只合用于 3 维空间. 试将 (72) 修改, 使能合用于平面的情形, 并导出 (19). 再用第 2 节的方法, 求 3 维空间的结果.
3. 如果在 (72) 的第 2 式加上一项 ∇p , 即压力梯度, (A) 试证

$$\frac{\partial^2 n'}{\partial t^2} + \omega_p^2 n' - \frac{1}{m} \frac{\partial p}{\partial n} \nabla^2 n' = 0, \quad (74)$$

并解之.

(B) 将 (17) 用到 3 维空间情形, 并对 k 展开, 留下头两项. 将所得结果和 (74) 的结果比较.

4. 试证离子气体的介电系数为

$$\epsilon = \frac{1}{(1 - \omega_p^2/\omega^2)}. \quad (75)$$

ϵ 的定义是: 如有一外加电荷分布 $qe^{ikr-i\omega t}$. 则物体中产生一电位分布

$$\frac{4\pi q}{k^2 \epsilon} e^{ik \cdot r - i\omega t} \quad (77)$$

以上假设 q 为小量.

5. 一高速带电粒子在一金属表面上空飞过, 速度和表面平行, 距离表面 x . 试讨论金属内电子的反应. 本题可以和上章第 3 题联想, 试证金属表面可有频率为 $\sqrt{2}\omega_p$ 的振动存在. (详细讨论见 Ruduick(1967)).

6. 读者须仔细导出第 4. 5. 节的计算, 以体会碰撞的意义, 并和第 3 节作一比较.

试将第 5 节的计算推广到低温费密粒子的情形.

7. 证实 (61) 的 c^2 即是理想气体的 $\frac{1}{m} \left(\frac{\partial p}{\partial n} \right)_s$, 即绝热压缩系数.

第二十一章 扩散方程式

长时间的改变, 是短时间改变的累积. 累积的结果, 往往和短时间运动的一些细节大致无关. 因此, 我们常可以把这些细节简化, 以避免繁杂的计算, 而有关长时间运动的结论并不受什么影响. 扩散方程及其各推广式就是这种简化了的分析工具. 这些方程式是非常粗略的模型, 效力相当有限. 但如果我们只要求粗略的答案, 这些方程式就非常有用. 本章的重点不在它们的理论基础, 而在一些简单的应用, 多数的例子和暂稳态有关. 屏障的翻越问题是暂稳态的分析重点.

这些方程式本身就是有趣的数学问题, 数学家已有群尽的分析, 巧妙的解法. 从物理的观点看来, 这些方程式到底多粗略? 如何去改进? 仍是未解决的问题. 也就是说, 我们仍无法从更基本的原理 (牛顿定律, 量子力学) 来了解这些方程式. 在特殊情形下, 可以有略深入的了解, 例如在上一章, 我们用玻尔兹曼方程式引出扩散. 但一般说来, 群体的运动的了解, 仍在起头的阶段. 此类方程式几乎是仅有的方便可用者, 知其可用, 而不知其所以然.

1. 简单的应用例子

在第十二章, 我们已经从不规则运动方程式推出了扩散方程式 (见 (12.55) 至 (12.68)). 扩散方程式可以看成中心极限定理的结果. 这一节谈这方程式的一些简单的性质及应用例子. 第 4 节再复习它和不规则运动的关系.

扩散方程式可以写成一粒子数守恒公式:

$$\begin{aligned}\frac{\partial \rho}{\partial t} &= -\nabla \cdot \mathbf{J}, \\ \mathbf{J} &= r\mathbf{F}\rho - D\nabla\rho.\end{aligned}\tag{1}$$

ρ 是一群粒子的粒子数密度分布, \mathbf{J} 为粒子数流量, \mathbf{F} 为每粒子所受的外力, D 为扩散系数. 流量有两项, 一是由扩散产生, 和密度的梯度成正比但方向相

反, 另一项和外力成正比, r 为比例常数. D 和 r 都看作已知的性质, 或可以测定, 或可以由其他理论计算. 它们都是温度的函数, 是粒子所处的环境的特性. r 的意义是一摩擦系数, 如不管 D , 则

$$\mathbf{J} = r\mathbf{F}\rho. \quad (2)$$

即流量和 \mathbf{F} 成正比, $r\mathbf{F}$ 是流速. 如果粒子是半径为 a 的小球而在一流体介质中, 则

$$r = \frac{1}{6\pi a\eta}. \quad (3)$$

η 是流体的粘度. r 和 D 之间有密切的关系. 因为, 在平衡时, $\mathbf{J} = 0$. 令 $\mathbf{F} = -\nabla U$, U 为外力位能, 则

$$\begin{aligned} -r(\nabla U)\rho - D\nabla\rho &= 0, \\ \rho &\propto e^{-(r/D)U}. \end{aligned} \quad (4)$$

在平衡时, ρ 应该和 $e^{-U/T}$ 成正比, 因此

$$\frac{r}{D} = \frac{1}{T}. \quad (5)$$

这是个十分重要的结果, 在第十二章已求出过, (见 (12.66)). 也就是第三章的细致平衡条件. 现在看些例子.

【例】某电子器材的效果可以用图 1 的 $U(x)$ 表示. 电子在 0 和 b 之间受力 $-F$, 在 b 点受很大的向右的力:

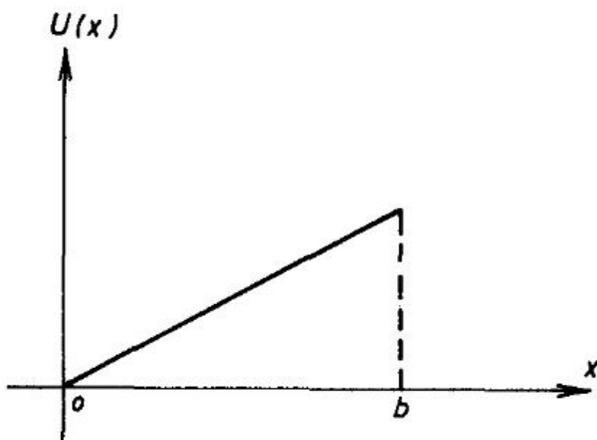


图 1

$$\begin{aligned}
 U(x) &= xF, \quad 0 < x < b, \\
 &= 0, \quad \text{其他 } x.
 \end{aligned}
 \tag{6}$$

这是一个简单的“二极管”模型. 电子向右流受力小, 向左则受力大.

在平衡时, 没有电流. 如果加一电场 E , 电流为何?

加了电场, 则 $U(x)$ 成了

$$U'(x) \equiv U(x) - exE. \tag{7}$$

由 (1), (5) 得

$$\begin{aligned}
 J &= D \left(-\frac{\partial U'}{\partial x} \frac{\rho}{T} - \frac{\partial \rho}{\partial x} \right) \\
 &= -De^{-U'/T} \frac{\partial}{\partial x} \left(\rho e^{U'/T} \right).
 \end{aligned}
 \tag{8}$$

电流稳定后, J, ρ 不再改变. J 不是 x 的函数. ρ 在 $(0, b)$ 之外亦为常数. 因此, 由 (8) 积分得

$$\begin{aligned}
 J &= \rho D (1 - e^{-eEb/T}) \Big/ \int_0^b dx e^{U'/T} \\
 &\approx \frac{\rho DF}{T} e^{-bF/T} (e^{eEb/T} - 1).
 \end{aligned}
 \tag{9}$$

以上假定 $eE \ll F$, Eb 是跨过 $(0, b)$ 的电位差. 虽然 $E = 0$ 时无电流, 但在 $eEb \gg T$ 时, E 向右电子可以畅行, E 向左则电流很小. (见图 2)

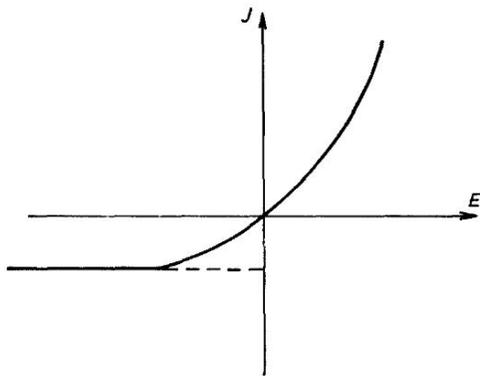


图 2

【例】 翻越屏障

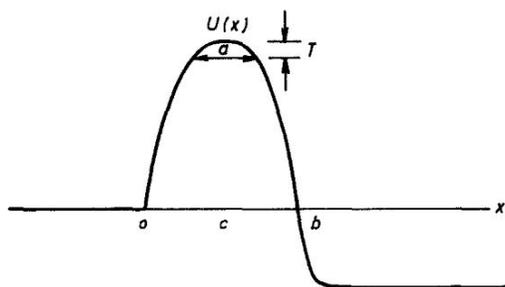


图 3

$U(x)$ 如图 3 所示, 起一“屏障”. 假定在屏障左边 $\rho \neq 0$, 而在屏障右边 $\rho = 0$, 并假定屏障相当高, $U_c \equiv U(c) \gg T$. 慢慢会有粒子翻过屏障, 掉到右边去. 这向右的流量是多少?

因为流量很小, 左边情况大致一定. 故 $J, \rho(0)$ 均大约不变, $\rho(b) \approx 0$. 用上例的步骤立即可得

$$J = \frac{\rho(0)D}{\int_0^b dx e^{U(x)/T}}. \quad (10)$$

将 $U(x)$ 在 $x = c$ 附近展开.

$$U(x) \approx U_c - \frac{1}{2}\alpha(x - c)^2, \quad (11)$$

代入 (10) 得

$$J \approx \rho(0)D \frac{e^{-U_c/T}}{\sqrt{2\pi T/\alpha}}. \quad (12)$$

屏障的高度 U_c 必是一决定性的因素. 注意, (9) 式中的 $e^{-bF/T}$ 也来自同一原因. $\sqrt{T/\alpha} \equiv a$ 是在 $U_c - T$ 处屏障的宽 (见图 3). 最难越过的是最高处, 其他地方相形之下都算不了什么.

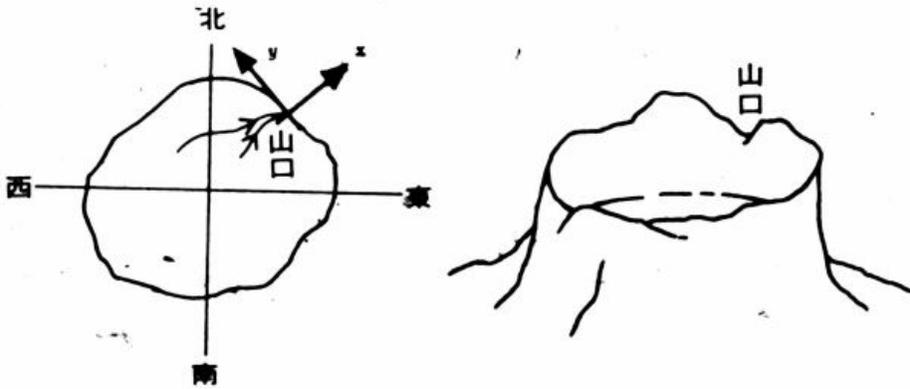


图 4

【例】图 4 示一二维空间的屏障，一个区域四周被很高的 $U(r)$ 包围。在东北方向有一略低之处，形成一山口，或“鞍点”。区内粒子密度为 ρ ，求流出区外的速率。

要走出去，必须翻过屏障。由上例的经验可知，最重要的因素是出去的路上最高点。显然由山口走出一条困难最少的路，因为走别的路须要翻过更高的山。把 x 轴穿过山口，则 (12) 可以用。粒子流出率约是

$$\begin{aligned}
 J &\approx \rho D \frac{\sqrt{\alpha\beta}}{(2\pi T)} e^{-U_c/T}, \\
 \alpha &\equiv - \left(\frac{\partial^2 U}{\partial x^2} \right)_c, \\
 \beta &\equiv \left(\frac{\partial^2 U}{\partial y^2} \right)_c.
 \end{aligned} \tag{13}$$

以上 c 是指山口处的值。 ρ 是区内密度。

2. 暂稳态

以上的例子是暂稳态的例子。在图 3 左边的粒子，或是在图 4 区内的粒子，都是在暂稳态。如果屏障高，它们可以很久不跑掉。这个时间尺度大致是

$$\tau = \tau_0 e^{\Lambda/T} \tag{14}$$

$\Lambda = U_c$ 是必须翻越的屏障高度。（此地用符号 Λ 因为它像个屏障）。 τ_0 是一个和 Λ 没有什么关系的时间尺度。我们可以就把它看成在屏障低时的时间尺

度. Λ/T 是最重要的因素. τ_0 相形之下就没有太大作用了.

(14) 可以用来估计一般暂稳态的寿命.

【例】直立的蛋可以维持多久?

在端午节, 或任何一天, 有耐心的人可以把蛋站起来. 站立的蛋是在一暂稳态, 现在来估计一下这暂稳态的屏障高度 Λ . 蛋之所以能直立, 是因为它的重量把它压在桌面上, 造成一个很小的接触面, 从重心向下的垂直线穿过此接触面. 要把这推倒, 重心必须先略为升高. 令 a 为接触面半径, h 为重心高度. θ 为蛋轴和垂直线的夹角, θ 很小时, 重心的升高为

$$\begin{aligned}\Delta h &\approx a\theta + h(\cos\theta - 1) \\ &\approx a\theta - \frac{1}{2}h\theta^2.\end{aligned}\quad (15)$$

这式子假设接触的是个坚硬的平面. $a\theta$ 是接触面中心的升高, $-\frac{1}{2}h\theta^2$ 是重心到接触面中心的高度的改变. 这假设并不太正确. 正确的分析必须考虑到蛋壳的弹性, 壳面的粗糙度, 蛋白蛋黄的流动等.

由 (15) 得蛋的位能 (图 5),

$$U(\theta) \approx mg\Delta h. \quad (16)$$

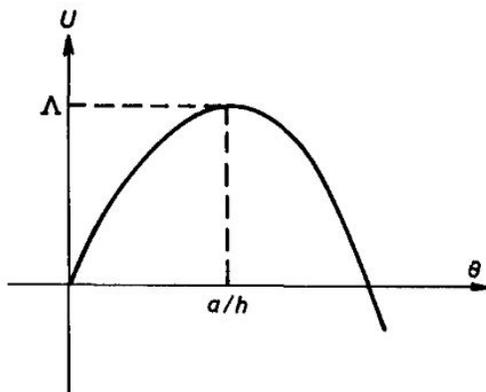


图 5

它的极大值在 $\theta = a/h$, 即

$$\Lambda = \frac{mga^2}{2h}. \quad (17)$$

一个蛋约重 30 克, 假定 $a \sim 0.03\text{cm}$, $h \sim 1.5\text{cm}$, 则 $\Lambda \sim 10\text{erg}$. 在室温

$$\frac{\Lambda}{T} \sim 10^{15}. \quad (18)$$

根据 (14), 不管 τ_0 多小, $e^{\Lambda/T}$ 使 τ 大得不成话. 也就是说, 单靠分子在平衡态的运动, 不足以使蛋跌倒. 在这种情形下, Λ 是一宏观量, 而 T 是一分子运动的能量. 要把蛋推倒, 某一分子的能量要有 10erg, 因此, 要分析蛋的站立时间, 必须考虑别的因素, 例如它的腐化, 室内的微风等, 这是以上分析的结论.

一般说来, (14) 的用途是在 Λ/T 不是太大但又不太小的情形. 如果 Λ/T 很大则表示更复杂的过程必须考虑.

3. 转换成波动方程式

以上例子的重点在流量 J , 现在来细看 (1) 的时间微分一面. 最方便的分析方法, 是把 (1) 改写成量子力学的波动方程式的样子, 再用我们量子力学的经验来分析. (1) 式的 $\nabla \cdot \mathbf{J}$ 中, 有 $\nabla \rho$ 项, 而波动方程式无此一项. 此一次微分不是一对称的算符, 颇不方便. 现在想法子把它去掉, 令

$$\rho(\mathbf{r}) \equiv \psi_0(\mathbf{r})\psi(\mathbf{r}), \quad (19)$$

$$\psi_0^2(\mathbf{r}) \equiv e^{-U(\mathbf{r})/T} \quad (20)$$

代入 (1), 稍加整理, 得

$$\begin{aligned} \frac{\partial \psi}{\partial t} &= -K\psi \\ K &= -D [\nabla^2 + W(\mathbf{r})], \\ W(\mathbf{r}) &\equiv -\frac{1}{\psi_0} \nabla^2 \psi_0. \end{aligned} \quad (21)$$

(21) 除了一些小节之外, 和量子力学公式一样. K 一如哈密顿算符. K 的“基态”为 ψ_0 , 能量为零:

$$K\psi_0 = 0. \quad (22)$$

量子力学的技巧, 可以立刻采用.

【例】, $U(x) = \alpha^2 x^2$, 解扩散方程式.

由 (20) 得

$$\psi_0(x) = \exp\left(-\frac{1}{2}\alpha^2 x^2 / T\right) \quad (23)$$

这当然就是简谐运动的基态波函数. K 的其他本征函数自然就是简谐运动的其他本征函数. 用不着多说, (1) 的解是

$$\begin{aligned}\rho(x, t) &= \psi_0(x)\psi(x, t), \\ \psi(x, t) &= \sum_{n=1}^{\infty} A_n \psi_n(x) e^{-r_n t} + A_0 \psi_0(x), \\ r_n &= \frac{2nD\alpha^2}{T}, \\ \psi_n(x) &= \sqrt{\frac{2}{T}} H_n(\alpha x) e^{-\frac{1}{2}\alpha^2 x^2 / T},\end{aligned}\tag{24}$$

A_n 是常数, 视 $t = 0$ 时的密度而定.

【例】解硬壁盒子内的扩散方程式

$$\begin{aligned}U(x) &= 0, \quad 0 < x < L \\ &= \infty, \quad \text{其他 } x\end{aligned}\tag{25}$$

此例的用意是提醒读者边界条件的重要, 扩散方程式在硬壁处的条件是

$$\mathbf{J}_{\perp} = 0,\tag{26}$$

\perp 指和壁垂直的方向. 这和 $\psi = 0$ 的条件大不相同, 因此, 此题的解是

$$\begin{aligned}\rho(x, t) &= \frac{N}{L} + \sum_{n=1}^{\infty} A_n \psi_n(x) e^{-r_n t}, \\ \psi_n(x) &= \cos \frac{n\pi}{L} x, \quad r_n = D \frac{n^2}{L^2}.\end{aligned}\tag{27}$$

N 是总粒子数.

4. 扩散方程式的推导

以上讲的 $\rho(\mathbf{r})$ 是粒子在 \mathbf{r} 空间的分布, 其实 \mathbf{r} 不必是位置, 任何变量都可以. $\rho(\mathbf{r})/N$ 是变数 \mathbf{r} 的概率分布. 概率的定义是由 N 个粒子的资料统计而得, 即样品集为 $(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$. N 也不必一定是粒子数. N 可以是重复实验的次数, \mathbf{r}_i 是第 i 次实验的结果.

我们复习一下第十二章的结果, 现在考虑由不规则速度 $v(t)$ 产生的位移, 并导出位移的分布.

从时刻 t' 开始, 在时间 Δt 内的位移是

$$\Delta x = \int_0^{\Delta t} dt'' v(t' + t''). \quad (28)$$

如果 v 的相干时间很短, 比 Δt 短很多, 则中心极限定理指出, Δx 的分布 $p(\Delta x)$ 为常态, 且

$$\begin{aligned} \langle \Delta x \rangle &= \langle v \rangle \Delta t, \\ \sigma^2 &\equiv \langle (\Delta x)^2 \rangle_c = \Delta t \int_{-\infty}^{\infty} dt'' \langle v(t'' + t') v(t') \rangle_c \end{aligned} \quad (29)$$

$x(t)$ 的分布可以由 $x(t')$ 的分布得来,

$$\begin{aligned} x(t) &= x(t') + \Delta x, \\ t &\equiv t' + \Delta t \\ \rho(x, t) &= \int d(\Delta x) p(\Delta x) \rho(x', t'), \end{aligned} \quad (30)$$

$$p(\Delta x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2\sigma^2} (\Delta x - \langle v \rangle \Delta t)^2 \right\}. \quad (31)$$

由直接微分, 可以证实 $p(\Delta x)$ 适合以下方程式

$$\frac{\partial p}{\partial \Delta t} = \langle v \rangle \frac{\partial p}{\partial (\Delta x)} + \frac{\sigma^2}{2\Delta t} \frac{\partial^2 p}{\partial (\Delta x)^2} \quad (32)$$

因为 $\Delta t = t - t'$, $\Delta x = x - x'$, 所以, 从 (30) 得

$$\begin{aligned} \frac{\partial \rho}{\partial t} &= \int dx' \frac{\partial p}{\partial (\Delta t)} \rho(x', t') \\ &= \int dx' \left[\langle v \rangle \frac{\partial p}{\partial (\Delta x)} + \frac{\sigma^2}{2\Delta t} \frac{\partial^2 p}{\partial (\Delta x)^2} \right] \rho(x', t') \\ &= \frac{\partial}{\partial x} \int dx' \langle v \rangle p(\Delta x) \rho(x', t') \\ &\quad + \frac{\partial^2}{\partial x^2} \int dx' \frac{\sigma^2}{2\Delta t} p(\Delta x) \rho(x', t'). \end{aligned} \quad (33)$$

现在假定 Δt 不很大, x, x' 相近, $\sigma^2, \langle v \rangle$ 可视为不随 x' 改变. 如此, (32) 成为

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x} \langle v \rangle \rho + \frac{1}{2} \frac{\partial^2}{\partial x^2} \frac{\sigma^2}{\Delta t} \rho. \quad (34)$$

以上的步骤和第十二章 ((12.55) 至 (12.68)) 的大致相同. 此地把 $\langle v \rangle \neq 0$ 的情形考虑进去, 并允许 $\langle v \rangle, \sigma^2/(\Delta t)$ 为 x, t 的平缓函数. 如果变量不止一个, 以上的推导可以推广, 并不困难. 令 $x_\alpha, \alpha = 1, 2, \dots$ 为变数, 则

$$\begin{aligned} \frac{\partial \rho}{\partial t} &= - \sum_{\alpha} \frac{\partial}{\partial x_{\alpha}} J_{\alpha}, \\ J_{\alpha} &= - \frac{\langle \Delta x_{\alpha} \rangle}{\Delta t} \rho - \frac{1}{2} \sum_{\beta} \frac{\partial}{\partial x_{\beta}} \left(\frac{\langle \Delta x_{\alpha} \Delta x_{\beta} \rangle_c}{\Delta t} \rho \right). \end{aligned} \quad (35)$$

(34), (35) 和 (1) 的关系已是十分清楚. 这些结果, 全靠中心极限定理得来. 如果这定理不合用, 则它们就没有什么意义了. (35) 即是所谓“福克—普朗克”方程式. 它也就是扩散方程式 (1), 写得稍微仔细一点, 并指出如何从不规则运动的细节来分析 D 和 $r\mathbf{F}$.

(35) 式中的 x_α 是有连续值的变量. 如果是分立值的变量, 则扩散方程式的面目需要改换一番. 举例说明如下.

令 $s_i = \pm 1, i = 1, 2, \dots, N$, 易形变量 s_i , 为最简单的变量, s_i 的改变必是突然的改变. 令 ρ_+, ρ_- 各为位子 +1 和 -1 的粒子数或概率, 则

$$\begin{aligned} \frac{d\rho_+}{dt} &= -r_{-+}\rho_+ + r_{+-}\rho_-, \\ \frac{d\rho_-}{dt} &= r_{-+}\rho_+ - r_{+-}\rho_-. \end{aligned} \quad (36)$$

dtr_{-+} 是在时间 dt 内 s_i 由 +1 变到 -1 的概率, dtr_{+-} 为由 -1 变成 +1 的概率. r_{+-}, r_{-+} 是反应率, 在平衡时, ρ_{\pm} 不变, 所以

$$\frac{r_{-+}}{r_{+-}} = \frac{\rho_-}{\rho_+}. \quad (37)$$

如果变量有许多值, 则 (36) 的推广是

$$\frac{d\rho_{\alpha}}{dt} = - \sum_{\beta} (r_{\beta\alpha}\rho_{\alpha} - r_{\alpha\beta}\rho_{\beta}). \quad (38)$$

α, β 为变数各值的编号, 即各位子的名字.

如果有外力使 α 位子的能量为 ϵ_{α} , 则平衡时

$$\rho_{\alpha} \propto e^{-\epsilon_{\alpha}/T}. \quad (39)$$

因此, 各反应率 $r_{\alpha\beta}$ 须适合

$$\frac{r_{\alpha\beta}}{r_{\beta\alpha}} = e^{-(\epsilon_{\alpha} - \epsilon_{\beta})/T}. \quad (40)$$

这也就是第三章讲的细致平衡条件.

(1), (35), (38) 这些方程式, 都是分子运动的一些粗略模型. 它们把运动的来源看成一不规则的“扰动”, 或不定时的突变. 这些扰动或突变必须适合某些条件, 最重要的是独立性, 及细致平衡条件. 此类模型虽然粗略, 但是非常可贵. 一则因为它们大致抓住了一些分子运动的特点, 再则更细致的分析往往过份复杂.

【例】 消去高能位

令 s_i 有三值 $+1, 0, -1$. 能量各为 (见图 6) $\epsilon_+, 0, -\epsilon_-$. 并令 $\epsilon_{\pm} \gg T$. s_i 不能从 $+1$ 直接变成 -1 , $r_{+-} = r_{-+} = 0$.

我们先得定义各反应率, 这些反应率必须适合细致平衡条件 (40). 我们用最简单的模型: $r_{\alpha\beta} = 1$, 如果 $\epsilon_{\alpha} < \epsilon_{\beta}$, 而 $r_{\alpha\beta} = e^{-(\epsilon_{\alpha}-\epsilon_{\beta})/T}$, 如果 $\epsilon_{\alpha} > \epsilon_{\beta}$, 即能量减少的反应率为 1, 增加者为 $e^{-\epsilon/T}$, ϵ 为反应所需的能量. 如此则

$$\begin{aligned} \frac{d\rho_{\pm}}{dt} &= -\lambda_{\pm}\rho_{\pm} + \rho_0. \\ \frac{d\rho_0}{dt} &= -2\rho_0 + \lambda_+\rho_+ + \lambda_-\rho_-, \\ \lambda_{\pm} &\equiv e^{-\epsilon_{\pm}/T}. \end{aligned} \quad (41)$$

0 是个高能位子, ± 1 为低能位子. 在 0 高能位的时间短而在低位者长. 因此, 我们可将 ρ_0 的方程式解了, 假装 ρ_{\pm} 不变:

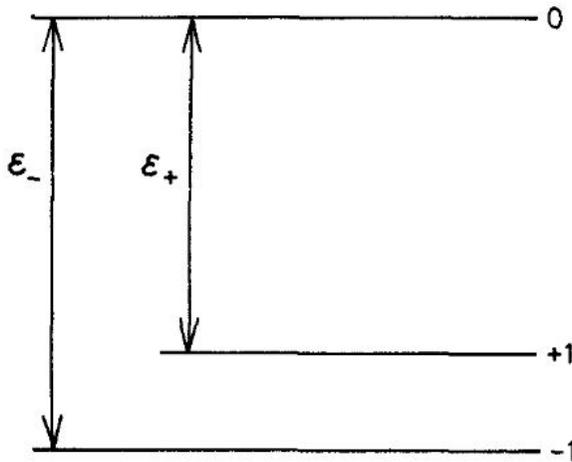


图 6

$$\begin{aligned}\rho_0(t) &= \int_{-\infty}^t dt' e^{-2(t-t')} [\lambda_+ \rho_+(t') + \lambda_- \rho_-(t')] \\ &\approx \frac{1}{2} \lambda_+ \rho_+(t) + \frac{1}{2} \lambda_- \rho_-(t).\end{aligned}\quad (42)$$

代入 (36) 第一式得

$$\frac{d\rho_{\pm}}{dt} = -\frac{1}{2} \lambda_{\pm} \rho_{\pm} + \frac{1}{2} \lambda_{\mp} \rho_{\mp}.\quad (43)$$

这就和 (36) 一样了. 高能位子 0 已消去, 也就是说, 在讨论慢动作时, + 和 - 两个位子的概率变化可由 (43) 决定, 不必考虑 0.

一般说来, 如果不止有三个位子, 但仍有两个特别低能的位子, 且被高能位子隔开, 则 (43) 也成立, 只是 λ_{\pm} 更复杂些.

注意, $\lambda_{\pm} = e^{-\epsilon_{\pm}/T}$ 正是和 (12) 一样的由屏障引起的因子. 凡是翻越屏障的反应, 其反应率都有这种因子.

5. 二态小集团模型

非晶物体, 像玻璃、岩石, 及第十八章提到的非晶磁体之类, 都有十分复杂的结构. 不过, 在某些特殊情况之下, 这些复杂的群体的低温性质, 可以用“二态集团模型”来解释. 这模型假设物体大致是冻结住的, 只有一小部分, 分散在物体各部的特别分子或自旋, 比较容易活动. 每个活动单元是一小集团, 由几个分子或自旋结成. 每一集团的最低能的两个形态, 在低温时最重要. 这二态模型只考虑这两态. 假定两态之间有屏障, 如此形态改变甚慢.

每个小集团的特性是由此二态的能差 ϵ 及二态间的屏障高度 Λ 决定. 图 7 的垂直方向是能量, 水平方向指小集团的状态空间. 从 + 态到 - 态必须翻过高度为 Λ 的屏障.

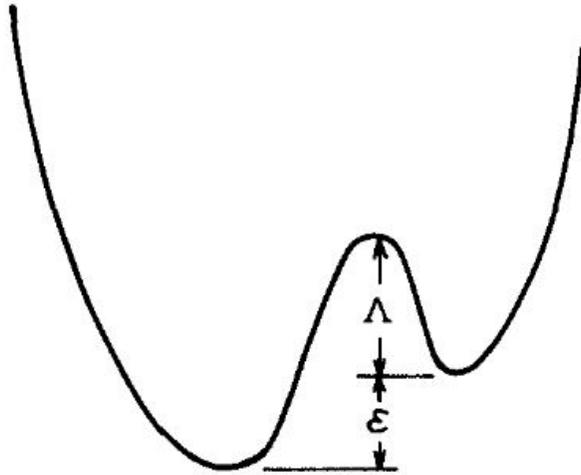


图 7

小集团很多, 各有不同的 Λ , ϵ . 假定集团数目够大, 我们可以统计 Λ , ϵ 的分布, 并用 (38) 之类的方程式讨论二态的概率 ρ_{\pm} . 假定

$$\begin{aligned} \frac{d\rho_+}{dt} &= -\frac{1}{\tau}\rho_+ + \frac{1}{\tau}e^{-\epsilon/T}\rho_-, \\ \frac{d\rho_-}{dt} &= \frac{1}{\tau}\rho_+ - \frac{1}{\tau}e^{-\epsilon/T}\rho_-; \end{aligned} \quad (44)$$

$$\frac{1}{\tau} \equiv \frac{1}{\tau_0}e^{-\Lambda/T}. \quad (45)$$

$1/\tau_0$ 是一常数, 约是高温时的反应率. 从这方程式的解和 ϵ , Λ 的分布, 可以推测各种性质.

读者可能有点不耐烦了. 左一个假设右一个假定, 到底这些集团是什么? Λ 和 ϵ 又是怎么分布? 现在用个简单模型来大致澄清一下.

令 H 为一易形磁体模型的哈密顿量

$$H = - \sum_{i,j} J_{ij} s_i s_j, \quad (46)$$

$s_i = \pm 1$, $i = 1, 2, \dots, N$. 作用能 J_{ij} 或正或负, 或大或小, 各不相同, 但不因时而异. 各 J_{ij} 的值成一分布, 令分布为常态, 且限于相邻的单元, 即各 s_i 占在一棋盘格子上, 不相邻的单元无相互作用. s_i 的运动是遵照以下规则: 如果 s_i 变号会使总能增加, 则反应率是 $e^{-\Delta E/T}$, ΔE 是增加能量. 如变号会使能量

减少, 则反应率为 1. 这也就是上节例子中用的规则, 也是下一章要仔细讨论的.

这模型已被仔细分析过*. 低温性质大致可用小集团模型解释. 在低温, 各个 s_i 大多冻结, 成了一个“不规则磁体”, 或“非晶磁体”, s_i 有的冻在 +1, 有的冻在 -1, 要改变 s_i , 需要能量 $2|h_i|$, h_i 是邻居在 s_i 处产生的磁场.

$$h_i = \sum_j J_{ij}s_j. \quad (47)$$

在低温时, h_i 小的单元会动. h_i 大就冻住了. 有的 h_i 虽不小, 但 s_i 和一相邻的单元者都翻个身时, 能量增加不多, 甚至减少. 如果 s_1, s_2 都变, 则能量增加为

$$\epsilon = 2|h_1| + 2|h_2| - 4|J_{12}|. \quad (48)$$

这可由 (47) 得来. 两个都变, 则 J_{12} 不生作用. 如果 J_{12} 很强, 则 h_1, h_2 虽大, 但 ϵ 可以很小, 甚至成了负的. 不过, 单元的运动是一个一个变, 不可以两个一同变. 因此, 能量必先增加 $2|h_1|$ 或 $2|h_2|$ 再下降, 一如图 7. 令 $|h_1| < |h_2|$, $\epsilon > 0$, 则先翻 s_1 再翻 s_2 的屏障较低:

$$\Lambda = 2|h_1| - \epsilon. \quad (49)$$

ϵ 由 (48) 定义. 这是一个“二元集团”. 二元集团和多元集团的分析也差不多. Λ, ϵ 的分布则要靠直接统计.

这个例子大致说明小集团的由来, 给读者一点“实在感”. 现在回到 (44), 看看这集团模型的结论.

(44) 的解非常的简单. 令 $\rho_+ + \rho_- = 1$, 则

$$\begin{aligned} \frac{d\rho_+}{dt} &= -\frac{1}{\tau}\rho_+ (1 + e^{-\epsilon/T}) + \frac{1}{\tau}e^{-\epsilon/T}, \\ \rho_+(t) &= f + e^{-rt}(\rho_+(0) - f), \end{aligned} \quad (50)$$

$$\begin{aligned} f &\equiv \frac{1}{(e^{\epsilon/T} + 1)} \\ r &\equiv \frac{1}{\tau_0}e^{-\Lambda/T} (1 + e^{-\epsilon/T}). \end{aligned} \quad (51)$$

$\rho_+(0)$ 为 $t=0$ 时的 + 位子被占的概率, 其值要看这物体是如何做成的. 现在估计一下能量的变化, 以说明这解的应用.

*详细分析见 Dasgupta, Ma, and Hu(1979).

总能量是

$$E(t) = \int d\epsilon d\Lambda N(\epsilon, \Lambda) \rho'_+(0) \epsilon e^{-rt} + \text{常数}, \quad (52)$$

$$\rho'_+(0) \equiv \rho_+(0) - f$$

$N(\epsilon, \Lambda)$ 是各集团的 ϵ, Λ 分布. r 由 (51) 定义, rt 可写成

$$rt = e^{-\frac{1}{T}(\Lambda-x)/T},$$

$$x \equiv T \ln(t/\tau'), \quad (53)$$

$$\tau' \equiv \frac{\tau_0}{(1 + e^{-\epsilon/T})} \approx \tau_0.$$

τ' 和 τ_0 差不多, 只要 $\epsilon > 0$. 显然 Λ 积分的结果是 x 的函数, τ' 只是以 $\ln \tau'$ 出现于 x , 因此把 τ' 看成 τ_0 亦可. 如果 $\epsilon < 0$, 且 $|\epsilon|/T$ 很大, 则 r 就不小了, e^{-rt} 可以略去. 我们现在把 $\epsilon < 0$ 情形略去, 由 (52) 得

$$E(t) = \int_0^\infty d\Lambda g(\Lambda) e^{-rt} + \text{常数}, \quad (54)$$

$$g(\Lambda) = \int_0^\infty d\epsilon N(\epsilon, \Lambda) \rho'_+(0) \epsilon.$$

e^{-rt} 这函数 (见图 8), 在 $\Lambda \gg x$ 时约为 1, 在 $\Lambda \ll x$ 约为零, 即

$$e^{-rt} = \exp\left(-e^{-\frac{1}{T}(\Lambda-x)}\right) \rightarrow 1, \quad \frac{(\Lambda-x)}{T} > 0, \quad (55)$$

$$\rightarrow 0, \quad \frac{(\Lambda-x)}{T} < 0.$$

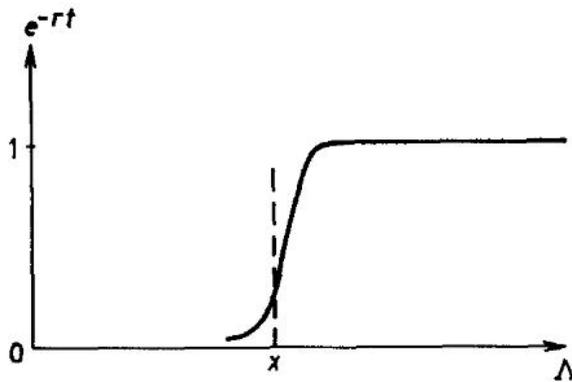


图 8

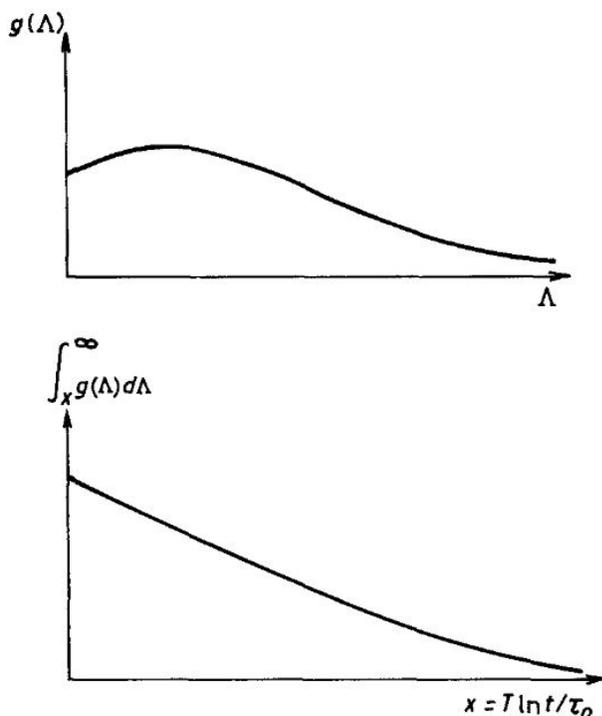


图 9

从 0 到 1, 转变区宽约为 T . 如 T 很小, 即 $g(\Lambda)$, 在 Λ 变到 $\Lambda + T$ 时, 改变不多, 则

$$e^{-rt} \approx \theta(\Lambda - x) \quad (56)$$

也就是说, 一个小集团的 Λ 如大于 $x = T \ln t / \tau_0$, 则小集团来不及改变, 冻住了. 如 $\Lambda < x$, 则它有足够时间改变, 解冻了. 注意, 时间是以 $\ln t$ 出现. 将 (56) 代入 (54) 得

$$E(t) \approx \int_x^\infty d\Lambda g(\Lambda) + \text{常数}. \quad (57)$$

只要 $g(\Lambda)$ 不是太怪异的函数, $E(t)$ 就会是一个 x 的平缓, 下降的函数 (见图 9). 也就是说, E 随 $T \ln t$ 逐渐下降, 这是非常慢的改变. 不只是能量, 其他性质亦然. 如磁矩, 一旦由磁场造成, 则不因磁场的消失而消失. 而慢慢地以 $T \ln t / \tau_0$ 方式减弱. 这种现象, 对研究古岩石磁性的人, 一向是非常熟悉的. 近来非晶磁体的研究, 亦得力于此类分析.

这小集团模型是极其粗略的. 做了不少的假设. 集团的重叠, 相互作用都未考虑. 集团如果很大, 则只看两态是不够的. 但以上的结论, 即对数式的

$T \ln t / \tau_0$ 改变, 是相当简单的结果, 对模型的很多细节不是很敏感. 说穿了, 这结果就是, $r = \frac{1}{\tau_0} e^{-\Lambda/T}$ 的化身. 只要有屏障高度 Λ 的分布, 就会有对数式的改变出现. 至于二态、三态之别, 甚至有没有小集团, 大概都无关紧要. 这集团模型, 只大致告诉我们 Λ 的分布的来源的一种.

类似的量子力学模型, 早在玻璃的理论中出现. 请自参阅文献[†].

讨论问题二十一

1. 设想有一面镜子, 直径远小于分子的平均自由径, 用一根细丝悬于温度 T , 压力为 p 的稀薄气体中. 试证镜子感受到的力矩为 $-\zeta\omega$, 且

$$\zeta = \frac{2m\bar{v}pI}{\sigma T}$$

$\bar{v} = \sqrt{8T/\pi m}$, I 为转动惯量, σ 为镜子的单位面积质量, m 为气体分子的质量, ω 为旋转角速度.

2. 有一表面光滑的球在温度为 T 的稀薄气体内作等速度运动. 球的半径大于气体分子的半径, 但小于气体分子的平均自由径. 假设球与气体分子之间为弹性碰撞, 求球所受到的阻力.
3. 试证在球面上的扩散方程式为

$$\frac{\partial \rho}{\partial t} = \frac{D}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \rho}{\partial \theta} \right) + \frac{D}{\sin^2 \theta} \frac{\partial^2 \rho}{\partial \phi^2}. \quad (58)$$

$\rho(\theta, \phi)$ 为在球面上的分布函数. (θ, ϕ) 为普通极坐标. 读者必须注意, (58) 的右手边不能写成

$$\frac{\partial^2 \rho}{\partial \theta^2} + \frac{\partial^2 \rho}{\partial \phi^2}$$

扩散的来源是不规则的动力. 在 (58), 我们假定此力的方向分布均匀, 且在球面上到处一样.

4. 试由 (35) 导出 (58).
5. 一圆管, 长 L , 截面积 A , 以角速度 ω 绕一端转, 转轴与管垂直. 管内有稀溶液, 有 N 溶质分子. 温度为 T , 转轴处的压力为 p . 假定每溶质分子的体积为 v , 质量为 m , 溶剂为水.

[†] 见 Anderson, Halperin and Varma(1972).

- (A) 求溶质密度分布 $n(r)$, r 为分子至轴的距离. 特别留意 m/v 值的影响.
- (B) 求熵及其他热力性质.
- (C) 如果在时刻 $t = 0$ 时, 所有的溶质都集中在转轴的一端, 分析密度分布 $n(\mathbf{r}, t)$ 在以后的改变情形, 估计趋向平衡分布所需的时间.
6. 某些大分子可看作小棍子, 长 l , 半径为 a , $l \gg a$, 这些分子溶在某溶剂中. 溶剂粘性系数为 η .
- 在时刻 $t = 0$ 时, 这些棍子都指向同一方向. 求此后的方向分布. (提示: 利用 (58)).
7. 某岩石中有许多小磁矩 m_i , $i = 1, 2, 3, \dots \dots N$, $m_i^2 = m^2 = \text{常数}$, 哈密顿量为

$$H = - \sum_i (h_i \cos \theta_i + A_i \cos^2 \theta_i). \quad (59)$$

各个 A_i , h_i 都不一样, 其分布为

$$\begin{aligned} p(h_i) &= \frac{1}{\sqrt{2\pi b^2}} e^{-h_i^2/2b^2}, \\ p(A_i) &= \frac{\theta(A_i)}{a} e^{-A_i/a}. \end{aligned} \quad (60)$$

假定 $a \gg b$. A_i , h_i 形容各磁矩所接触的环境. θ_i 为 m_i 和某一定方向的夹角. 各磁矩不断地被其环境中的扰动影响, 作不规则运动.

假定在时刻 $t = 0$ 时所有的 $\theta_i = 0$. 试求此段总磁矩的变化. 假定 $T \ll a$.

读者应先认清这模型中各量的意义, 将模型简化. 这是一个暂稳态的模型. A_i 项使 m_i 大致偏向两个方向, 即 $\theta_i = 0$ 或 π .

第二十二章 数值模拟

电子计算机的问世,对统计力学有极其深远的影响,不但使较复杂的计算问题得以解决,而且在基本概念上也有重大的贡献.本章略述两种计算方法,一是“分子运动数值解”,即解运动方程式,把所有分子的位置、速度统统算出来.二是“随机数模拟”,这是把运动方程式用随机数列取代,专用来求平衡性质,亦称“蒙特卡罗”法.本章主要讲第二种.

分子运动解的计算,是不折不扣的“数值实验”,把运动轨迹算出来,再作各种观测.随机模拟也模拟出轨迹来,只是用一个随机的运动模型,而不用牛顿定律.它也可以算是一种数值实验.这些数值实验有什么好处呢?第一,这些计算,不用近似.一旦模型定了,计算就定了,一切结果,你可以看成该模型的“实验结果”(当然,实验步骤必须正确,观测时间的长短,模型的大小等都必须考虑.).这比我们用纸用笔的理论分析要强,用不着再一个假设,再一个近似,东丢一项,西切半个积分.因此,计算的结果可以和真正实验比较,以决定模型的正确性,并可以和理论分析比较,以决定各近似,假设的正确性.第二,这些计算,算出了每个分子在每个时刻的一切运动细节.这些资料,在真正的实验中是无法测得的.当然,数值实验比较容易控制,用不着担心样品的纯化,低温系统漏热等问题.

由于这些好处,数值实验已成为十分普遍的工具.此类计算既是“实验”,则非自己动手不能领略其要义.各种程序的设计,全属经验,少见于书刊,也难解释.本章所谈,仅是一些原则上的问题,为读者略作心理准备而已.

这些计算,除了是很有用的工具之外,更能加深我们对统计力学的认识,使基本概念更清楚,这一方面的讨论,将占去本章许多篇幅.

1. 分子运动数值解

这个方法的原理十分简单, 把分子运动方程式 (牛顿定律) 交给计算机去解就是了.

在计算机发明之前, 要解三个分子以上的运动方程式, 是不可思议的事, 更别谈什么群体运动. 因此, 统计力学的学者有根本放弃分子运动者, 即不承认运动轨迹的物理意义, 因为不可能解得出来. 这种想法使统计力学走向抽象化, 纯数学化的路. 直到今天, 传统的统计力学仍以抽象的概率来作平衡态的定义, 而不愿涉及轨迹的讨论.

计算机问世之后, 对分子运动有兴趣的人立刻用它来解运动方程式, 看看平衡现象是否出现. 大师费米及门人乌拉姆解了一些振动方程式, 发现平衡不了, 因而引起了注目. 逐渐地, 更多人用更进步的机器来解运动方程式, 来对平衡作根本的了解. 费米、乌拉姆的困难是振动单元太少, 多放几个, 平衡现象就不难出现. 在今天, 用计算机解群体运动方程式已是统计力学的一部门. 每个计算考虑以千计的粒子, 作百万次的碰撞.

这种计算是一种实验, 由计算机把一群体在某时间内的轨迹算出来. 每分子在每一时刻的位置、速度都算出来了. 任何有关此群体在这一段时间内的数据都可以从这轨迹整理出来, 各平衡性质可以由这时间内的各平均值得来. 温度可以由每粒子的平均动能得来, 这种“数值实验”比真正的实验要容易控制, 而且可以作任何观测, 大尺度, 小尺度, 平衡不平衡, 都可以.

当然, 这种计算和真的实验仍有很大的区别:

(甲) 几千个粒子的群体, 比宏观物体小得太多. 百万碰撞的时间比宏观时间短得多.

(乙) 无法考虑量子力学, 解群体的量子波动方程式仍非今天的计算机所能办到.

这两个弱点看来是难以克服的. 但由于物体各部的独立性, (甲) 并不很严重, 对平衡性质的分析而言, 只要考虑比相干长度 ξ 大很多的模型, 观测时间比相干时间 τ 长很多就行. 如果 ξ 够小, 则用不着考虑很多粒子, τ 够短则不必有很长的时间. ξ, τ 的长短要看情形. 照现有的计算成绩看来, 一千个粒子的模型在一般情形下足够, 时间也足够. 在特殊情形下, 如临界点附近, ξ, τ 都很大, 就需要更大的模型, 往往超出今天计算机的能力. 弱点 (乙) 比较严重, 许多重要的问题都不能脱离量子力学. 例如金属电子, 超流体等. 这使计算的应用范围大受限制, 虽然, 这计算仍有很大的应用范围, 很多现象, 如熔

解、结晶, 及一些重原子、分子的群体性质, 大多和量子力学无关.

从基本概念的角度看来, 此类计算的成功有深远的意义. 它确定了力学方程式, 即牛顿定律的地位, 也就是说, 一切热力现象, 一切统计力学的概念, 归根结底都得以力学为本. 平衡是一种特别的运动状况, 必须遵守力学的定律. 运动的轨迹代表一切有关运动的数据, 任何概念, 任何量, 都必须能用这数据表示, 从这些数据计算. 有些热力学概念, 没有明显的力学意义. 这些是和熵有关的量. 轨迹和熵的关系, 是一个十分重要的题目. 在第二十五章, 我们将作详细讨论, 并指出一个由轨迹直接定熵的方法.

解运动方程式是一个复杂的计算, 有许多题目并不要求运动的真实性. 因此, 我们可以用比较简化的运动模型代替牛顿定律. 随机模拟就是一个简化的运动模型.

2. 随机数列

随机模拟是用随机数列决定运动, 不用牛顿定律, 随机数列的作用就是产生不规则的动力.

在第十, 至十二章, 我们曾谈到随机数列, 这是一连串的数目字, 适合一些条件 (见第十一章第 3 节). 现在的计算机, 多附有产生随机数列的程序, 最普通的是平均分布在 0 和 1 之间的随机数列, 以后提到随机数列, 就是指这种, 除非另外指明.

制造随机数列的程序, 就是一个快速的骰子, 此类程序, 是经验的结果. 目前并没有理论说如何的程序会算出较合标准的数列. 注意, 程序是个完全硬性规定的步骤, 一如牛顿定律. 造出来的数列为什么会不规则? 这问题和我们的基本问题很像: 为什么牛顿定律会造出不规则的分子运动? 答案都不知道.

随机数列可以说是概率的定义工具. 如果没有这种快速产生的数列, 则概率这概念就无法成为有用的计算工具. 今天的很多科学研究都依赖随机数列, 这岂是三、四十年前所能预料?

有了随机数程序, 我们可以放心去“任取一数”, 或“任取三者之一”等, 作“任意”的选择. 例如, 要在 1, 2, 3 中任取一数, 怎么取? 我们叫计算机程序算一随机数 x . 如果 x 小于 $1/3$ 则取 1, 在 $1/3, 2/3$ 之间, 则取 2, 大于 $2/3$ 则取 3. (我们已假定 x 的概率分布是在 0, 1 之间的均匀分布). 如要以概率 p_1

取 1, p_2 取 2, p_3 取 3, 如何取? 答案是, $x < p_1$, 则取 1, $p_1 < x < p_1 + p_2$, 则取 2, $p_1 + p_2 < x$, 则取 3.

注意: 这种“任取”如只行一两次, 则是毫无意义的, 必须要重复很多次, 概率才有意义.

3. 随机数模拟法

我们把随机模拟的计算法则, 看作一不规则运动模型. 计算的步骤, 视模型而定. 现在用易形模型为例, 来说明其原理.

令哈密顿量为

$$H(s) = -\frac{1}{2} \sum_{i,j} J_{ij} s_i s_j. \quad (1)$$

$s_i = \pm 1, i = 1, 2, \dots, N$. 模拟步骤如下:

- ① 先任取一组 s_i 的值.
- ② 任取一 s_i . 计算 h_i , 及 ΔE

$$\begin{aligned} h_i &\equiv \sum_j J_{ij} s_j, \\ \Delta E &\equiv 2s_i h_i \end{aligned} \quad (2)$$

③ 如果 $\Delta E \leq 0$, 则把 s_i 改成 $-s_i$, 如果 $\Delta E > 0$, 则以 $e^{-\Delta E/T}$ 的概率将 s_i 改成 $-s_i$, (即取随机数 x , 如 $x < e^{-\Delta E/T}$, 则改 s_i , 否则不改).

- ④ 再做 ②

这模拟步骤可谓简单. ① 是决定起始状态, 任何状态都可以, ② 是挑一个单元来翻. ③ 是要决定到底翻不翻. ΔE 是 s_i 翻后的能量改变, h_i 是邻居在 s_i 处造成的“磁场”. 如果能量减少, 则改 s_i 成 $-s_i$, 如能量增加, 则改成 $-s_i$ 的概率是 $e^{-\Delta E/T}$. 这个法则适合细致平衡的要求, 即 $s_i \rightarrow -s_i$ 的概率和 (同一 h_i 下的) $-s_i \rightarrow s_i$ 的概率之比是 $e^{2s_i h_i/T}$, 所以

$$\frac{P(s_i|h_i)}{P(-s_i|h_i)} = e^{2s_i h_i/T}. \quad (3)$$

$P(s_i|h_i)$ 是在 h_i 一定的条件概率. 每做一次 ② 和 ③, 时间就增加 1 单位. 做了许多次后, (3) 中的概率就有足够的样品来定义了, 一次又一次地重复, 就模拟出了一连串的状态, 即一条轨迹.

令 $s(t) \equiv [(s_1(t), s_2(t), \dots, s_N(t))]$ 为在时刻 t 的状态, 则任何量 $\hat{A}(s)$ 的平均值可以求得

$$\langle \hat{A} \rangle = \frac{1}{\mathcal{T}} \sum_{t=t_0}^{t_0+\mathcal{T}} A(s(t)). \quad (4)$$

\mathcal{T} 是一相当大的数目, t_0 之前的状态不用. 原因是, 开始的一些状态, 可能和“典型的”平衡态下看到的相差太远. $t=0$ 的形象是乱取的, 因此需要等一段时间, 以“达到平衡”. 这些问题就在下节再谈. 现在先看看另一个运动模型.

令 s_i 在时间 dt 内改变的概率为 $dt W_i$,

$$W_i = \frac{e^{-s_i h_i/T}}{\cosh(h_i/T)} \quad (5)$$

这也是一个运动模型. 注意:

$$\frac{W_i(s_i)}{W_i(-s_i)} = e^{-2s_i h_i/T}, \quad (6)$$

即这模型也适合细致平衡的要求. 用这模型模拟, 也很简单. 令

$$\Omega \equiv \sum_i W_i. \quad (7)$$

从时刻 t 算起, 过了时间 t' 后, 毫无动静的概率是 $e^{-\Omega t'}$. 因此在 t' 内不变而在其后 dt' 内 s' 改变的概率是

$$e^{-\Omega t'} W_i dt', \quad (8)$$

依照 (5), (7), (8) 模拟步骤是

- ① 从时刻 t 的状态, 计算 W_i 和 Ω .
- ② 取两个随机数, x, y . 令

$$t' = \frac{-\ln x}{\Omega}, \quad (9)$$

并令新时刻为 $t + t'$.

③ 把 $(0, 1)$ 这段区间分成 N 段, 每段长 W_i/Ω , $i = 1, 2, \dots, N$. 如果 y 落在第 j 段, 则将 s_j 改成 $-s_j$. 现在回到 ①.

这个模拟步骤比上一个复杂一些, 上个步骤每一步任取一单元, 再决定翻不翻它, 这个步骤每次决定静止的时间, 再决定翻那一个单元. 平均值的计算

亦稍有不同:

$$\langle \hat{A}(s) \rangle = \frac{1}{\mathcal{T}} \sum_{k=k_0}^{k_0+n} t'_k A(s(t_k)),$$

$$\mathcal{T} = \sum_{k=k_0}^{k_0+n} t'_k.$$
(10)

t'_k 是第 k 次改变前的静止时间长度, $s(t_k)$ 是在这时间内的状态. 在时间 \mathcal{T} 内, 共有 n 次改变.

4. 该注意的概念问题

上节的模拟步骤, 有十分明显的含义. 各单元按不规则的, 各处不相干的原动力震撼着. 各单元因此而运动. 但也受到其他单元的影响. 因此各单元可能不是互相独立的. 影响限于邻近的单元, 即相互作用非零者. 因此, 远方的单元大概是不相干的. 因此, 这步骤是十分合理的运动模型.

这步骤并不提到基本假设, 唯一和热力学扯上关系的地方是在改变概率中的 $e^{-\Delta E/T}$ 因子. 温度是这样跑进来的. 在第三章, 细致平衡的讨论限于理想气体. 在此各单元有相互作用, 因此情形比较复杂, 需要多讨论一下.

在此 s_i 的改变是在一定的 h_i 之下, 也就是说 s_i 的环境决定 s_i 的改变率. 细致平衡条件是在同一环境下, $s_i \longleftrightarrow -s_i$ 的正反方向的概率相等. 由此得 (3). (3) 可以稍加推广, 只要 s_j 的邻居 s_i (即 J_{ij} 非零者) 一定, h_i 就定了, 所以

$$\frac{P(s_i, s'_i, s''_i, \dots)}{P(-s_i, s'_i, s''_i, \dots)} = e^{2s_i h_i / T}. \quad (11)$$

s'_i, s''_i, \dots 包括所有和 s_i 有相互作用者, 也可以包括更多单元, 由 (11) 可得 (3). $P(s_i, s'_i, \dots)$ 是联立概率.

如果 s_i, s'_i, s''_i, \dots 包括了所有的单元, 则 (11) 的结论十分明显, 即

$$P(s_1, s_2, \dots, s_N) \propto e^{-H(s)/T}. \quad (12)$$

这样的说法时常被看成平衡的证明. 不过, 这说法是有漏洞的. 我们曾一再强调, 概率的定义, 需要有足够的样品作为统计资料的来源. 此地的样品, 就是轨迹中的状态, 即模拟出的状态. 以上的诸概率, 必须由这些状态来定义. 如果 N 不是个小数目, 则 $P(s_1 \dots s_N)$ 的定义需要许多状态, 即 $\sim 2^N$ 个. 如果 $N = 100$, 即使轨迹中有一千万个状态, 仍距 2^{100} 甚远. (12) 只是一个数学

结论, 即对无限长轨迹的结论. 不过这个“无限长”是没有物理意义的. (11) 只合用于较少数的单元, 如 $P(s_1, s_2, \dots, s_{10})$ 之类, 注意 (s_1, \dots, s_N) 代表状态空间中的一点, 而 (s_1, \dots, s_{10}) 代表此空间的一区域. 而这区域是由 $s_1 \dots s_{10}$ 的值划定, 域中共有 2^{N-10} 个点, 这是一个很大的区域. (s_1, s_2, s_3) 代表一个有 2^{N-3} 点的区域. 这种大区域的概率是有物理意义的, 因为轨迹在其中花的时间可以够长, 够用来计算概率. 这些在第十章末尾曾讨论过. 读者可复习一下. 基本假设分析下的平衡, 并不涉及 (12) 这单状态概率的概念, 只涉及大区域的体积. 而平衡性质是由大区域决定的. 在讨论基本假设及概率的诸章中, 我们已说了很多. 在做数值模拟计算时, 这些抽象的讨论结果, 就变得很明显. 读者最好能有机会做一点此类计算.

读者大概会问: 可不可以用模拟来计算热位能? 配分函数也可以写成一平均值:

$$\begin{aligned} e^{-F/T} = Z &= \sum_s e^{-H/T}, \\ \langle e^{H/T} \rangle &= \frac{1}{Z} \sum_s e^{-H/T} e^{H/T} = \frac{2^N}{Z}, \\ Z &= \frac{2^N}{\langle e^{H/T} \rangle}. \end{aligned} \quad (13)$$

把 $e^{H/T}$ 的时间平均值算出, 则 Z 和 F 就有了. 不过这办不到, 除非 N 很小. H 是 N 级大数. $e^{H/T}$ 是特大数, 其改变也是特大, 除非用无限长的时间, (13) 是没有用的. 这“无限长”的时间是指比 e^N 级特大数更长的时间, 如果要算 $F = E - TS$, 只有直接求 $E = \langle H \rangle$ 和熵 S , 从轨迹求熵是第二十五章的主题.

在第六章我们提过, 在活动范围中几乎所有的状态都有相同的宏观性质, 如 $A(s)$ 为一宏观变量, 则在活动范围中绝大多数的 s 的 $A(s)$ 值都一样:

$$A(s) = \langle A(s) \rangle \left(1 + O\left(\frac{1}{\sqrt{N}}\right) \right). \quad (14)$$

因此, 看来只要任取一状态 s 就知道 $\langle A \rangle$ 了, 何必做什么模拟?

事实上, 要在活动范围中任取一点颇不容易. 活动范围是很大, 体积是

$$\Gamma = e^S. \quad (15)$$

如果 $S \sim 0.5 \ln_2 N$, 则 $\Gamma \sim 2^{0.5N}$, 但这 Γ 比整个状态空间又小得太多.

$$\frac{\Gamma}{2^N} \sim 2^{-0.5N}. \quad (16)$$

如果 $N = 100$, 则这比值已小得不成话. 因此如果在状态空间中任取一点, 想要碰到活动范围中, 是毫无希望的. (因此, 用 (12) 作概率, 再用普通的随机积分法来计算平均值, 也是行不通的). 要进活动范围, 只有让这物体的运动, 逐渐把状态带进去.

在第 1 节提过, 计算中的 N , 或 100, 或 1000 仍不是 10^{20} , 我们要从这种不太大的 N 来推测热力性质, 岂非妄想? 情形不这么糟. 宏观物体有各部独立的性质, 只要模型比相干距离大很多就够了. 这模型可以看成一大物体中的一部分. 当然, 边界成了一个重要的考虑问题. N 不很大, 则 $N^{-1/3}$, $N^{-1/2}$ 都不是很小的数字. 一般的计算, 常用循环边界, 即无边界的模型. 这作法相当成功.

由于 N 非宏观大数, 许多结论 (如 (14) 式的结论) 都得重新考虑过. 模拟法至今仍是在累积经验的阶段, 还没有很稳固的理论基础.

模拟乃运动模型之解, 故应用不限于平衡. (随机模拟在长一点的时间尺度下, 也可以算是十分逼真的运动模型) 对暂稳态的分析, 模拟法是十分理想的工具. 暂稳态在短时间内是平衡态, 时间一长, 改变就出现. 除了暂稳态之外, 凡在统计力学的基本假设不太管用的情形下, 模拟法就是一个可贵甚至唯一可靠的工具. 许多不平衡的问题, 如结晶, 熔解等, 都已由模拟法分析, 获得十分有趣的结果*.

讨论问题二十二

1. 复习问题十一, 5, 7.
2. 数值模拟有许多实际困难. 主要是模型太小, 模拟时间太短. 但是许多有趣的问题, 偏偏是相干时间或相干距离长的问题. 因此, 读者在阅读文献时, 或自己做模拟计算时, 必须特别小心. 常常因为对结果没有正确的解释, 而导致不正确的结论. 试用易形模型为例, 讨论在低温时 $h = 0$ 和 $h \neq 0$ 的时间问题.
3. 试讨论在 N, V 不太大时, 第七章的各计算法则有何不同.
4. 试讨论模型的边界问题.

*阅读文献, 可由 Binder (1979), Berne (1977), Abraham (1979) 开始.

5. 第 3 节讲的是定温环境下的运动. 试设计定能随机模拟步骤, 及限能模拟步骤. (参考第 3 题).
6. 设计随机模拟步骤, 仿一气体的运动. 注意, 分子的动能可以不管, 每一步只把一分子挪一小距离, 再计算能量差. 假定有定温环境.
7. 如果要模仿定压环境, 该如何办? 定压则体积也成了运动变量. 如果用循环边界, 要改变体积, 就不很容易.

有一个法子, 是把分子的位置写成 $b\mathbf{r}_i, i = 1, 2, \dots, N$. 然后把 $(b, r_1, r_2, \dots, r_N)$ 空间看成此物体的状态空间, 用

$$H(b\mathbf{r}_1, \dots, b\mathbf{r}_N) + pb^3L^3$$

来作模拟. L^3 是固定的体积, p 是压力, \mathbf{r}_i 是限制在 L^3 的体积内. 试设计模拟程序, 并讨论此法的得失.

8. 第七章的每一法则, 都有一个随机模拟步骤. 读者自己去设计程序. 特别记住第 3 题的讨论.
9. 低温的随机模拟, 模拟出的是能量低的各状态. 如果温度很低, 则零点 (即基态), 就会常出现. 因此, 一面模拟, 一面温度逐渐降低, 就可以模拟出零点, 求出哈密顿函数 $H(s)$ 的极小值. 因此, 随机模拟法可以用来作一般求极小值的工具, 不只是做统计力学题目. $H(s)$ 可以是任何函数, s 是任何变量. 例如, s 代表某人环球旅行的一个路程, $H(s)$ 是该路程的费用. 路程必须经过某些地方. 如要决定最省钱的路程, 可用随机模拟法. 先想好一个改变 s 的步骤, 再模拟. 先用高温, 再逐渐降低温度.

(这种求极小的方法, 在工业方面已大有用处, 得力于物理学家 S. Kirkpatrick 等的研究).

这方法有许多困难, 视题目而异. 读者试讨论暂稳态的问题, 即模拟出的可能是一暂稳态, 而非基态.

随机模拟法还有许多不同的用途. 读者自己可多作设计.

第六篇

溯源

以上诸篇, 对统计力学的概念和应用, 大略作了一个介绍. 读者到此已非初学. 现在我们再回头分析一些基本概念, 用比较深入、有经验的眼光来看本书开始时的一些题目.

第二十三章讨论热力学的三大定律, 推敲它们的含义. 第二十四章例举几个“重聚现象”, 扫除一些对熵的误解. 这些现象本身都是极其有趣的. 第二十五章介绍直接从运动轨迹求熵的一个方法, 使统计力学的所有基本概念都和分子运动衔接上. 第二十六章略述有待解决的一些原则上的难题.

第二十三章 热力学定律

本章重新讨论第二章的热力学基本概念, 把时间尺度的角色作更进一步的说明, 把基本假设和热力学的关系作一总结, 包括热力学第三定律. 我们先讨论熵为什么在绝热过程前后不变, 再讨论熵的增加问题. 事实上熵之为极大是基本假设的一部分, 也就是说, 基本假设可以说是已经假设了第二定律. 最后我们讨论第三定律, 及可逆过程扮演的角色. 一般第三定律的“违反”是因为不可逆过程的介入.

1. 绝热过程

热力学第一定律, 一方面是重申能量守恒定律, 一方面也把能量分成两种, 一是功能, 一是热能. 在第二章已谈过一些. 功和热之分, 是要靠绝热过程来定义. 现在再来看绝热过程.

令某物体的哈密顿函数为 $H(s, L)$, s 指物体的状态, L 为一不变量, 例如, 某气体在一容器中, 容器壁对气体分子的影响, 可以用一个位能 U 代表.

$$H = \text{动能} + \text{分子间相互作用} + \sum_{i=1}^N U(\mathbf{r}_i). \quad (1)$$

$U(\mathbf{r})$ 的图形见图 1, 在 y, z 方向也类似. 接近器壁处, U 增高, 代表壁对分子的斥力. 在 x 靠近 L 处,

$$U(\mathbf{r}) = v(L - x), \quad (2)$$

即位能是分子位子和墙壁间距离的函数. 在此 L 看成一个不变量. 或一个“参数”.

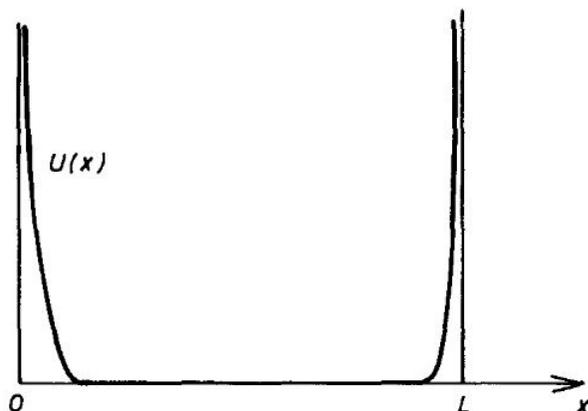


图 1

如果这气体的环境是绝热, 则总能量即是 (1), 没有和外界交换能量, U 为固定的函数. 在这情形下, $H(s, L)$ 为守恒量, 即

$$\frac{\partial}{\partial t} H(s(t), L) = 0. \quad (3)$$

现在我们将 L 改变, 即把一壁移动, 从 L 以等速移到 $L + \Delta L$, 假定费时 \mathcal{T} , 移动之后, 能量的改变是

$$\begin{aligned} \Delta E &= \int_0^{\mathcal{T}} dt \frac{dH}{dt} = \int_0^{\mathcal{T}} dt \frac{\partial H}{\partial L} \frac{dL}{dt} \\ &= \frac{\Delta L}{\mathcal{T}} \int_0^{\mathcal{T}} dt \frac{\partial H}{\partial L} \\ &= \Delta L \left\langle \frac{\partial H}{\partial L} \right\rangle, \end{aligned} \quad (4)$$

$$\left\langle \frac{\partial H}{\partial L} \right\rangle \equiv \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} dt \frac{\partial H}{\partial L}. \quad (5)$$

注意 $dL/dt = \Delta L/\mathcal{T}$, 以上用到 (3), 即若 L 不变, 能量亦不变. 能量的改变是由 L 的改变引起. (4), (5) 的平均值, 是过程时间内的平均值, 如果 ΔL 很小, 则 (5) 中 $H(s, L)$ 的 L 可视为不变. 结论是: 绝热过程能量的改变是

$$dE = \left\langle \frac{\partial H}{\partial L} \right\rangle dL. \quad (6)$$

$\langle \partial H / \partial L \rangle$ 为时间平均值. 以上纯是力学结果, 和热力学, 统计力学的基本假设无关.

由 (2) 式, 得

$$\begin{aligned}\frac{\partial H}{\partial L} &= \sum_i \frac{\partial U(\mathbf{r}_i)}{\partial L} = - \sum_i \frac{\partial v}{\partial x_i} \\ \left\langle \frac{\partial H}{\partial L} \right\rangle dL &= -pdV \\ p &\equiv \frac{1}{A} \left\langle - \sum_i \frac{\partial v}{\partial x_i} \right\rangle, \quad dV \equiv AdL.\end{aligned}\tag{7}$$

A 是容器在 y, z 方向的切面积, $-\partial v/\partial x_i$ 为器壁加在分子 i 上之力. (6), (7) 当然是第二章例举绝热过程之一.

如果过程时间 \mathcal{T} 比物体内分子运动相干时间长, 则 $\langle \partial H/\partial L \rangle$ 是一平衡性质 (假定在过程前物体是在平衡态). \mathcal{T} 不能太长, 太长则不能绝热. 因此, \mathcal{T} 一方面要比相干时间长, 但从宏观观点 \mathcal{T} 必须很短.

注意, 千万不要把这绝热过程和量子力学中的“平缓过程”混为一谈*. 我们在此澄清一下. 在量子力学中, 哈密顿算符 H 可以是某参数 λ 的函数. 因此, 各能量 $E_n(\lambda)$ 也是 λ 的函数, n 是定态的量子数. 如果将 λ 缓慢的改变, 则有如下结果:

(甲) 各能量保持原来的格局, 只要把 $\lambda(t)$ 代入即可, 即各能量是 $E_n(\lambda(t))$,

(乙) 如果开始时状态是 n , 则一直留在 n 态. 这是量子力学的“平缓过程定理”, 条件是过程时间 \mathcal{T} 必须很长

$$\mathcal{T} \gg \frac{\hbar}{\Delta E}\tag{8}$$

ΔE 是两相邻 E_n 之差. 这个定理, 合用于任何孤立物体, 当然也合用于此地讲的绝热物体. 从 (甲), (乙) 结果可以立刻得到 (6). 但是这定理的条件 (8) 是不合用于宏观的物体. 如果把 (8) 用到宏观物体, 则 ΔE 非常的小, \mathcal{T} 必须大于 $e^{10^{20}}$. (即轨迹走过活动范围中每一状态所需的总时间). 即

$$\begin{aligned}\Delta E &\sim \frac{E}{\Gamma(E)} \\ \mathcal{T} &\gg \frac{\hbar\Gamma(E)}{E} \sim e^N.\end{aligned}\tag{9}$$

这当然是个无理的要求. 因此, 量子力学的这条定理, 根本不能用来解释热力学的绝热过程. 绝热过程的要求是 \mathcal{T} 大于起伏的相干时间, 这是一个很小的时间, 和 (9) 扯不上关系.

*在英文书中, 此二者用同一字 adiabatic, 因而常引起误解, 用中文者, 不该犯此错误.

我们现在可以把功能定为由宏观时间尺度的变动而来的, 即改变时间为宏观者. 由微观运动而引起的能量改变定为热.

2. 绝热过程和熵

以上分析并没有涉及熵的概念. 现在我们来看熵的改变和绝热过程的关系, 从基本假设开始. 我们用第七章的“限能法则”计熵 (见 (7.27)):

$$S(E, L) = \ln \sum_s \theta(E - H(s, L)). \quad (10)$$

用这法则的原因纯粹是为了方便. 将 (10) 微分, 得

$$\begin{aligned} dS &= \left(\frac{\partial S}{\partial E} \right)_L dE + \left(\frac{\partial S}{\partial L} \right)_E dL, \\ \left(\frac{\partial S}{\partial L} \right)_E &= - \frac{\sum_s \delta(E - H) \frac{\partial H}{\partial L}}{\sum_s \theta(E - H)} \\ &= - \frac{\sum_s \delta(E - H) \frac{\partial H}{\partial L}}{\sum_s \delta(E - H)} \cdot \frac{\sum_s \delta(E - H)}{\sum_s \theta(E - H)}. \end{aligned} \quad (11)$$

最后一行的最后一个因子就是 $(\partial S / \partial E)_L$. 另一因子正是由定能法则计算的平均值

$$\left\langle \frac{\partial H}{\partial L} \right\rangle. \quad (12)$$

因此

$$dS = \left(\frac{\partial S}{\partial E} \right)_L \left[dE - \left\langle \frac{\partial H}{\partial L} \right\rangle dL \right], \quad (13)$$

$\partial H / \partial L$ 为宏观变数. (12) 这平均值, 应该和时间平均值 (5) 一样. 这是以上诸讨论的结果. 如果基本假设成立, 这结果就成立. 现在, 由 (6) 可见, 如果 dL, dE 是由一个绝热过程造成, 则 $dS = 0$. 这是个极其重要的结果, 由基本假设和力学结果 (6) 直接推出. 基本假设并不提过程, (6) 亦不涉及熵. 在第二章, 熵这个概念是由绝热过程推出. 由以上结果可见, 热力学和统计力学的基本假设, 可以说是吻合得十分完美.

经 (13) 可得

$$\left\langle \frac{\partial H}{\partial L} \right\rangle = \left(\frac{\partial E}{\partial L} \right)_S \quad (14)$$

即“力”是能量对位移的微分, 只是要把熵定住. (13) 可以写成更熟悉的方式:

$$\begin{aligned} dE &= TdS + \left\langle \frac{\partial H}{\partial L} \right\rangle dL, \\ \frac{1}{T} &\equiv \left(\frac{\partial S}{\partial E} \right)_L. \end{aligned} \quad (15)$$

3. 热力学第二定律

热力学第二定律, 是超出平衡性质之外的定律. 它指出熵的改变方向. 如果提到改变, 当然就不只是谈平衡态而已.

首先, 请务必注意: 熵的定义, 需要一段观测时间 \mathcal{T} . 熵是由状态空间中的轨迹在这时间中的分布范围, 即活动范围, 来定义. 熵没有瞬时的意义.

因此, 要谈“某时刻的”熵, 我们必须用一个比 \mathcal{T} 更长的时间尺度. 举个例子, 图 2 的气体分两部分, 中间的壁上有小洞. 开始时右边为真空, 气体慢慢漏到右边.

定义熵的观测时间 \mathcal{T} 必须符合条件

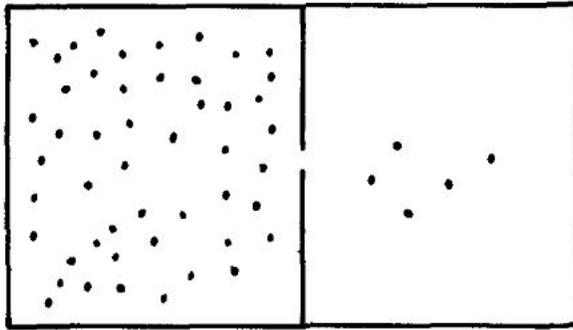


图 3

$$\mathcal{T} \gg \tau. \quad (16)$$

τ 是分子的平均自由时间. 要谈熵的改变率 dS/dt , 这 dt 必须比 \mathcal{T} 大. 也就是说漏气率必须小:

$$\left| \frac{1}{N_1} \frac{dN_1}{dt} \right| \ll \frac{1}{\mathcal{T}} \quad (17)$$

N_1 是左边分子数. 在时间 \mathcal{T} 内, N_1 可看成不变, 熵可计出. 如果漏得太快, 则熵在过程中是无法定义的.

当然, 我们可以讨论在漏之前 (洞未开) 的熵和漏到两边分子一样多之后的熵, 不谈漏气过程中之熵, 因为漏前、漏后, 都有足够的时间来定义熵.

根据第二定律, 绝热环境下的物体, 熵不会减少. 从基本假设来看, 这个定律的解释相当明显. 基本假设说: 熵是 Γ 的对数, Γ 是活动范围内的状态总数, 活动范围包括所有不变量所容许的状态. 也就是说, 熵已经被假设成极大了, 因为能包括的都包括了.

或许在短时间内, 活动范围因时间不足而受限制, 但时间一长, 所容许的范围就大了. 以上漏气的例子就是一示范. 由于洞小, 在短时间内分子被限在一边, 但在长时间内, 分子可以到另一边去, 因此范围大了, 熵也大了. 当然在两边粒子数相等时, 有没有洞都无所谓了 (见“分类综合”的讨论, 第六章). 因此, 熵的有增无减, 是基本假设的一部分. 上一节的讨论指出, 绝热过程前后熵不变. 因此, 对物体做功或对外做功不能使熵减少. 当然以上讲的绝热过程, 是慢动作的过程. 如果动作很快, 比分子相干时间更快, 这样的改变, 会不会使熵减少? 这问题是超出基本假设之外, 不能用它回答. 谁有如此快的动作, 能把高能的分子和低能的分开? 至少现在没人或机器能做到, 但要证明不可能做到却也困难. 文献中有证明这是不可能做到的. 这些证明事实上是假设连物体带机器一同都在平衡态, 即把基本假设用到物体加上机器. 这当然会得到熵不减的结果, 因为基本假设已假设了这结果. 我们看一个例子.

假想在图 2 的洞内装一对飞轮 (见图 3). 各轮上打一小洞. 飞轮以高速运转.

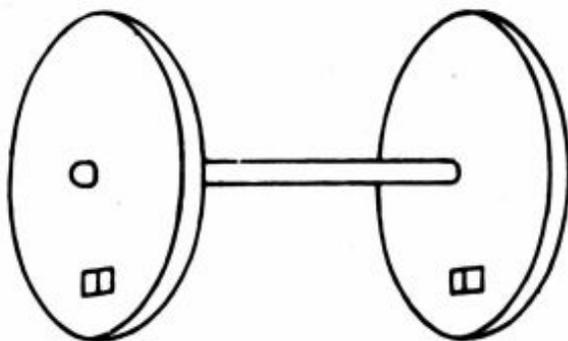


图 3a, 挖了小洞的飞轮

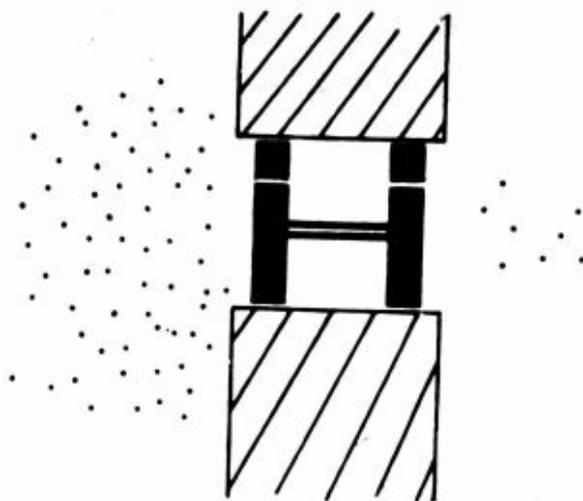


图 3b, 飞轮装在图 2 中间的隔板孔内

假定开始时右室是空的, 左室中高速分子比较容易通过轮上小洞. 如果速度 v 够大, 即

$$v > a\omega$$

a 是小洞宽, ω 是飞轮角速度, 则射入第一洞就可以穿过第二洞. 如 $v < a\omega$, 则通过第一洞就过不了第二洞. 当然, 碰壁多次之后仍有机会钻过.

乍看起来,这对飞轮可以使高速分子通行,低速分子被挡在一边.因此右边的温度就比左边高了,违反了第二定律,其实不然.

开始时,右边的分子确是速度高,但是飞轮不断地被分子撞击.飞轮之间也积了不少分子.因此,左右两边的能量可以由传导交流,时间一久,左右两室,连同飞轮的温度都一样.以上假定了基本假设的正确性,(即假定各处温度相同).一旦用了这基本假设,第二定律就不会被违反.无论是多精巧的机器,都不能例外.

因此,若证明了基本假设,就证明了第二定律.

4. 热力学第三定律

这定律的内容是:在温度 T 趋近于零时,熵也趋近于零[†].

从运动的概念来看,这定律指出在 $T \rightarrow 0$ 时,运动停止(量子力学的“零点运动”除外).说得更仔细些,在 $T \rightarrow 0$ 时,形象改变也趋于停止.因为改变停止,活动范围缩小,熵也趋于零.因此,从这运动观点来看,这定律是十分容易解释的.

从热力学的观点来看,这定律使熵有了绝对的定义.在热力学中,只有熵之差的定义.二平衡态 1, 2 的熵差为

$$S_2 - S_1 = \int_1^2 \left(\frac{dQ}{T} \right) = \int_1^2 dT \left(\frac{C}{T} \right) \quad (18)$$

$$C = \frac{dQ}{dT} = \text{热容率}$$

dQ 为所加的热,积分是沿着由 1 到 2 的可逆过程做的.因此,熵的定义,差了一个积分常数.这常数的选择,在第二至五章谈过.

因为我们已对统计力学十分熟悉,对这常数的决定不觉得稀奇.下面是个标准例子,说明第三定律的内容.

如果原子 A 和原子 B 可以结合成分子 AB .在极高的温度下, AB 的气体会成了 A, B 的混合气体,熵是

$$S_{AB} = S_A + S_B + N \ln 2. \quad (19)$$

[†]这是 Nernst 在 1905 年的假设.

$N \ln 2$ 是“混合熵”， N 是 A 的粒子数，也是 B 的粒子数。高温的熵和 $T = 0$ 的熵之差可由 (18) 测定：

$$\begin{aligned} S_A(T) &= \int_0^T dT' \frac{C_A(T')}{T'} + S_A(0), \\ S_B(T) &= \int_0^T dT' \frac{C_B(T')}{T'} + S_B(0), \\ S_{AB}(T) &= \int_0^T dT' \frac{C_{AB}(T')}{T'} + S_{AB}(0), \end{aligned} \quad (20)$$

式中 C_A, C_B, C_{AB} 为 A, B , 及 AB 的热容率。在相变过程中，必须另加潜热项。这三种热容率和各潜热都可以由实验测定。(19) 是热力学的结果 (见第二章)。因此，一旦选定了常数 $S_A(0)$ 和 $S_B(0)$ ，则 $S_{AB}(0)$ 就定了。同理，如可合成化合物 A_2B_3, AB_4 等，则它们的熵的未定常数也可由 $S_A(0), S_B(0)$ 决定。

显然，只要把每个元素物质的熵定了，所有的化合物的熵也定了。再进一步：每个元素都是由电子，核子构成，核子由中子质子构成，则如纯中子物体，质子物体，电子物体的熵中的常数定了，一切物体的熵的常数也定了。很明显：熵的定义中，该有一个基本常数，和物体结构无关的常数。第三定律把所有的 $S(0)$ 设为零。这个选择和第三章，第 3 节的选择实在是同一回事。这个选择并没有什么选择的余地。第三定律被发现时 (1905)，统计力学尚未诞生。

照以上的结论任何物体的熵，从 $T = 0$ 到高温，都可以由热容率的测定而求出，即

$$S(T) = \int_0^T dT' \frac{C}{T'} \quad (21)$$

加上相变时的潜热项。在高温气态时，熵的计算很简单，即理想气体的熵，加上一点修正。现在，我们可以用实验来鉴定 (21) 是否和高温的理论相符合。事实上在这些统计力学结果未出世之前，已有人用第三定律，及 (21) 求出了理想气体的熵[‡]，和以后的统计力学结果相符，并发现常数中有量子常数出现。当时的量子常数 h 才诞生不久。

现在再回到 (19), (20) 的讨论。把 $T = 0$ 时的熵设为零，则 (20) 的三个积分必须在 T 很大时适合 (19)。这是一个十分苛刻的结论。有很多实验可以做，注意， A, B 分开，或是化合成 AB ，是它们不同的存在形态。(20) 表示这些不同形态下的热容率间有关系存在。把这结论推广一点：如果一种分子，可以形成不同的结构，则在不同结构下的热容率之间也有关系存在。

[‡]在 1911 年 Sackur, 1912 年 Tetrode 做出。

许多元素的固态, 可以有不同的结构. 在一般的观测时间内, 这些结构都是平衡态 (如果时间是无限长, 则只有一种结构是稳定, 其他是暂稳态). 它们的热容率都可以测出. 一个最常提起的例子是锡. 金属锡冷到 $T_0 = 292\text{ K}$ 以下成暂稳, 称之为“白态”. 稳定的结构是一种半导体, “灰态”. 在温度 T_0 , 灰、白态可以共存, 一如水的气液共存. 灰态吸热可成白态. 在 T_0 之下, 白态是暂稳, 但在一般观测时间下改变得太少, 因此适合我们的平衡定义. 因此, 白态的熵在 T_0 是

$$S_1(T_0) = \int_0^{T_0} dT \frac{C_1(T)}{T}. \quad (22)$$

C_1 是白态的热容率. S_1 也可以由灰态热容率 C_2 积分而得:

$$S_1(T_0) = \int_0^{T_0} dT \frac{C_2(T)}{T} + \frac{L}{T_0}, \quad (23)$$

L 是由灰变白的潜热. (22), (23) 必须相等, 这就是 C_1, C_2 间的关系.

在以上的讨论, 我们对 $T \rightarrow 0$ 这极限没有小心定义. 在实验室制造低温是一门大学问. 要多低的温度下熵才是零? 这自然要看物体内运动的能量尺度. 当温度小于所有的运动能量尺度, 则可以算是零度. 在固体中, 能量最小的运动是核子自旋的改变. 如果温度比这能量高, 则核子自旋仍可以有熵. 如果核子自旋的运动和其他运动、物质结构无关, 则可以略去不管. 例如在以上锡的例子中, 核子自旋和灰态白态没有什么关系. 只要温度不是太低, 每个核子有熵 $\ln(2I + 1)$, I 是核子自旋. 这虽不是零, 但因核子数不变, 它不会影响以上的结果, (即 (22), (23) 各加 $N \ln(2I + 1)$, 不影响 C_1, C_2 间的关系). 如果温度低到各核子间相互作用能时, C_1, C_2 当然也会受到核子运动的影响. $T \rightarrow 0$ 的意义是温度比这相互作用能还低. 这时所有的运动都停止了.

注意, 以上所说的熵趋于零是指

$$\lim_{T \rightarrow 0} \lim_{N \rightarrow \infty} \frac{S(T)}{N} \rightarrow 0, \quad (24)$$

即每分子的熵趋于零. 能量尺度是指每分子的能量尺度.

5. 非晶态和第三定律

大约三十年前, 有些化学家, 用实验测定很多较复杂的固体的热容率, 例如冰, 及一些由分子 (即不是单原子固体) 构成的固体. 后来又有玻璃的热

容率实验. 这些实验结果都很明显地违反了热力学第三定律. 也就是说, 令 $S(T)$ 在高温气态时为已知, (气体的统计力学结果), 则实验结果

$$S(0) \equiv S(T) - \int_0^T dT \frac{C(T)}{T} \quad (25)$$

不等于零 (总是大于零). 有关这些实验, 见参考书[§].

其实, 这些实验, 并没有真正违反了第三定律, 因为熵的定义 (18) 要求可逆过程积分, 而这些实验中的过程免不了渗入了不可逆成分. 从这观点来看, 是实验不合定律的条件, 并非定律不对. 从运动的观点来看, 在 $T \rightarrow 0$, 运动停止, 熵也趋于零. 第三定律不会被违反. 这些实验的解释如下.

在低温, 冰, 玻璃一类的物质是“非晶体”(见第十七章), 有许多不同的时间尺度. 有些运动变量作快速改变, 有些作缓慢改变. 并且这些时间尺度对温度很敏感. 温度降低一些, 一些动得快的变数, 就冻结了. 温度升高时, 一些变量就“解冻”. 从高温冷下, 可以冻结成不同的结构. 虽然在大尺度下, 不同的结构看来都一样, 但在小尺度下, 却是不一样. 因此, 解冻之后再冻结, 很难回到原来的冻结结构, 所以过程是不可逆的.

当温度增加时, 本来可以动的变量增加了改变幅度. 除此之外, 本来在同一观测时间内不变的, 现在也动了. 因此不难想见

$$dS > \frac{dQ}{T}. \quad (26)$$

如果没有时间的限制, 则 $dS = dQ/T$. 根据 (26), (25) 式定义的 $S(0)$ 一定是大于零的.

注意, 如果某物体的零点, 或基态, 是重迭的, 即有许多最低能状态, 且重迭数的对数和 N 成正比, 则 $S(0)/N$ 就不是零. 这和以上讲的非晶性质毫无关系. 最简单的例子是 N 个没有相互作用的自旋 I , 它们的熵是 $N \ln(2I + 1)$, 不管温度是多低. 还有许多模型, 其基态的熵非零. 但自然界中的物体内很难有绝无作用者.

到此读者会问, 如果免不了要牵涉到不可逆过程, 那末熵就不能用热容率来测定. 熵到底该如何定义呢? (26) 中的 dS 又是什么? 根据基本假设, 熵是由状态空间中的活动范围体积所定. 但如果时间尺度不很清楚, 这范围可以定吗? 这些问题, 必须从运动轨迹的分析求解答. 见第二十五章.

[§]Wilks(1961).

讨论问题二十三

1. 如果 (11)—(13) 中的 $\langle \partial H / \partial L \rangle$ 和 (5), (6) 中的时间平均值相同, 则绝热过程下熵不变. 过程时间 \mathcal{T} 如比相干时间大很多, 则以上结论成立.

在气体中, 相干时间是看作平均自由时间, 但是, 第二十章的例子指出, 声波、扩散现象使某些相干时间加长不少. 例如, 长波振动可以迟迟不消失.

现在考虑第 1 节的过程 (见图 1), 把 L 以等速移到 $L + \Delta L$. 费时 \mathcal{T} .

- (A) 写下流体力学公式, 以讨论这过程引起的运动.
- (B) 这些公式是微分方程式, L 的改变是边界的改变, 也可看成一参数的改变. 现在这问题就很像量子力学中的缓慢过程问题. 微分方程式相当于单粒子波动方程式.
- (C) $\Delta L / \mathcal{T}$ 应该多小才不致于激起振动.
- (D) 如果分子的平均自由距离比 L 大, 以上的讨论和结论又如何?

2. 复习第八章第 2 节, 氢分子气体的分析. 只考虑分子转动.

(A) 用热容率积分求 $T \rightarrow 0$ 时的熵. 试证

$$S(0) = N_1 \ln 3. \quad (27)$$

N_1 是自旋为 1 的分子粒子数.

(B) H_2 分子的真正基态是自旋为零的. 因此由热容率积分求出的熵, 可能和真正的基态无关. 这该是非常明显. 读者自己举几个其他例子.

3. 某物体由 N 单元 $s_i, i = 1, 2, \dots, N$ 组成. 总能量为

$$H = - \sum_i s_i h_i + \sum_i (1 - s_i^2) \Lambda_i, \quad (28)$$

每单元有三个形态, $s_i = +1, 0, -1$. h_i, Λ_i 为固定, 但在每个单元有不同的值, 其分布为

$$p(h) = \frac{1}{\sqrt{2\pi b^2}} \exp(-h^2/2b^2),$$

$$p(\Lambda) = \frac{1}{a} e^{-\Lambda/a}. \quad (29)$$

假定 $a \gg h$. 这模型和问题二十一, 6 差不多.

假定有以下运动规则:

如果 $\Lambda_i > T \ln t$, 则 s_i 不能变.

如果 $\Lambda_i < T \ln t$, 则 s_i 可以变动.

t 为常数, 可以看成观测时间.

(A) 如果 $T \gg a$, 求总能量、热容率和熵.

(B) 将此物体逐渐冷却, 每一段时间 t , 温度下降 a/n . 令 $n = 100$, 令 $b = a/10$. 从 $T = 2a$ 开始, 试求出各量和 T 的关系并作图. 注意: 如果 s_i 不能动, 则其熵为零.

(C) 用 (B) 的热容率 dE/dT 积分来求熵, 并和 (B) 直接求出的熵作一比较. 就此讨论热力学第三定律.

第二十四章 重聚现象

熵的来源是不规则的运动。我们一直反复地强调运动的重要性。不规则的排列，散乱的分布，如不随时而变，则都和熵没有关系，一定要有运动才行。我们已举了不少的例子，指出活动范围和熵必须由运动来决定。在这一章，我们叙述三个颇富戏剧性的实验，即“自旋重聚”，“粘液瓶”实验，及“离子重聚”实验，前者是韩氏*的核子自旋共振的实验。他把自旋的相对进动方向倒转，使散乱了的自旋又重聚。“粘液瓶”是把分散在粘液中的颜色重聚起来，是一个非常简单的教室示范实验。韩氏用它来解释他的自旋重聚实验。（据他说，这示范不是他首创，但我以为至少他可居推广之功）。韩氏实验之后，有更多类似实验，包括“离子重聚”†。韩氏的目的，不是讨论熵。以后有人做类似的实验，也不是为了讨论熵。此类实验，在文献中称为“回响”现象。实验的技术和结果，都是极有趣，极有价值。我们的目的，是进一步阐明熵的意义。我们先仔细讲粘液瓶示范，这是教室示范，十分简单。再分析这示范的原理。然后再略提一下自旋重聚及离子重聚实验。这两个实验，都不是教室内可做，理论也复杂。在此地所述，不是原来文献中所述，而是简化过，为求阐明原理的假想实验，使读者可以大致了解实验的意义。

1. 粘液瓶示范

这个示范十分简单。任何学校都可以做。瓶子的构造是：两个同轴圆筒（透明塑料或玻璃制），外筒固定，内筒可用一把手推转（见图 1）。内外筒之间盛粘液，玉米浆即可。没有玉米浆，可以用太白粉加水搅拌煮成。粘液必须透明。越粘越好。内外筒之间之距离约一公分，外筒直径约 10 公分。大一点或小一点都没关系。

*Hahn(1950).

†Malmberg 等 (1968), O'Neil and Gould(1968).

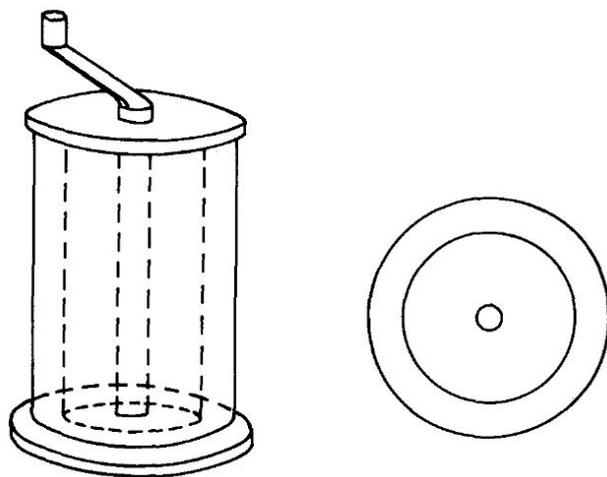


图 1

用一小烧杯, 将一些粘液和少许颜料调匀. (颜料可用食物颜料, 或原子笔心挤出的墨汁, 其他颜料大概也可以). 颜料分子均匀地分散在小烧杯的粘液里. 只需要一点点颜料即可, 看得出颜色就行了.

示范步骤如下:

(甲) 将小烧杯中的有色粘液用玻璃管吸一点, 再伸入瓶中. (瓶子盖上面必须有个洞, 使玻璃管可以出入). 再把有色粘液徐徐吹入已装好的无色粘液中, 边吹边将吹管上提, 使色液成一条线.

(乙) 将把手朝一方向推几圈, 贴靠内筒的分子被带动, 而靠外筒的不动. 近内筒的被推得较远, 近外筒的较不远. 因此, 这条色线里的颜料分子就被分散开了. 推上一两圈, 颜色已不见, 再推上十来圈, 颜料分子已分散到每一个角度. 在这时, 任何实验都可以鉴定颜料分子的混乱分布, 和理想气体相似, 只是不像气体分子一样的运动.

(丙) 将把手朝反方向推同样圈数, 那条色线即重新出现. 这时可见学生们脸上惊讶的表情. 颜料分子又重新聚合到原来的地方了. 若不是亲眼看到, 并自己推把手, 真是难以相信.

这个现象之所以令人惊讶, 当然是颜料分子的从分散而重聚. 一般对熵的解释是“混乱”的程度, 热力学第二定律是指“越来越混乱”的必然性. 此地示范的重聚现象, 显然和第二定律不符. 有人的解释是: 颜料分子重聚时, 熵从瓶子经过把手一直流到示范者的手臂里去了. 这类解释是由于对熵的误解, 是经不起考验的. 手臂推把手, 无论往那一方向推, 都是做功, 和瓶子里的

颜料分子的排列无关.

事实上, 这重聚现象和第二定律并无抵触之处.

2. 示范的分析

我们先叙述一个不正确的分析, 再讲正确的分析.

不正确的分析是: 当颜料在未放入瓶中时, 每个颜料分子的位置的“不确定”程度是色线的容积 V_0 , 概率分布为 $1/V_0$, 所以有熵

$$\begin{aligned} S_0 &= \ln \left(\frac{V_0^N}{N!} \right) + N\sigma \\ &= N \ln(V_0/N) + N + N\sigma, \end{aligned} \quad (1)$$

N 是颜料分子数, N/V_0 是小烧杯中颜料分子的密度. σ 是由于每分子小幅度振动引起的熵.

当把手被推了好些圈之后, 颜料分子位置的概率分布是 $1/V$, V 是瓶内粘液的容积, 因此, 熵是

$$\begin{aligned} S &= \ln \left(\frac{V^N}{N!} \right) + N\sigma \\ &= N \ln(V/N) + N + N\sigma, \end{aligned}$$

即

$$S - S_0 = N \ln(V/V_0). \quad (2)$$

这是分散后熵的增加. 看来这分析很合理. 但显然有毛病, 因为推回后颜料分子又回到 V_0 . 熵又减少了. 违反了第二定律. (如果熵流入了手臂, 这股热流不难感觉到). 以上的分析的毛病是在对熵的误解. 在定义熵时, 活动范围、概率分布, 是要由运动幅度来决定. V_0 , V 和颜料分子的运动幅度没有关系. 因此 (1)、(2) 都不对.

我们再次强调: 熵和观测的“不确定”扯不上关系, 更和观测者的“知识”或“无知”扯不上关系. 如果用这些概念来作熵的定义, 就自然地导致 (1), (2) 的错误结果. 现在来看正确的分析.

这重聚现象事实上毫不值得惊讶. 每个颜料分子, 除了作极小幅度的振动之外, 是静止的, 绝不像气体分子似的跑来跑去. 在数小时的观测时间内, 是不会有扩散现象发生, 示范中所有的时间尺度是我们眼睛的暂留时间, 约

1/10 秒左右. 在这种时间尺度下, 熵是小幅振动的熵, 和颜料分子在液中的分布无关. 因此, 将把手推几圈, 只稍微把整个物系稍稍弄热了一点. 由于摩擦, 无论朝那方向推, 粘液, 颜料分子振动都稍稍大一点, 熵稍稍增加一点. 这些和分子分布扯不上关系. 因此, 熵只是 $N\sigma$, 和 V_0, V 无关.

我们还可以这样看: 因为颜料分子很少, (在小烧杯内就十分稀了), 因此每个颜料分子的四周上下都是粘液分子. 因为它不动, 它的坐标和环境是不变的, (小幅度振动除外), 一如固体中被困住的杂分子. 因此, 熵只可能和颜料分子的总数有关, 不会和它们的位置有关. 这一点, 在讲冻结杂质时, 已经分析过.

因此, 在我们示范观测时间内, 熵在颜料分子重聚时和在分散时一样 (略去摩擦所生的热). 分散或相聚, 在这时间尺度下, 都是平衡态. 颜料分子的位置为不变量. “分散态”和“相聚态”有同样的熵. 这和第二定律毫无抵触之处.

数日之后, 颜料分子会散开一些, 不能再推回原处了. 这时, 熵增加了吗? 答案当然是没有. 刚才的结论是: 熵和颜料分子的位置分布无关. 数日后再看时所用的时间尺度仍是“看”的尺度, 即 1/10 秒到几分钟.

如果瓶中不是粘液, 而是粘性极低的流体, 则颜料分子将迅速地运动. (1), (2) 也就没错了. 重要的时间尺度是

$$\begin{aligned} \tau &\sim \frac{l^2}{D}, \\ l &= \left(\frac{V}{N}\right)^{1/3}. \end{aligned} \quad (3)$$

D 是颜料分子扩散系数, l 是颜料分子间的距离. 改变粘性就改变 D , 也就改变了时间尺度. 这个示范粘液的 D 非常小, 以至于 τ 可达十天左右, 超出平衡态的观测时间. 如果不用玉米浆而用水, 则 τ 比观测时间小很多, (1)、(2) 就合用了.

注意: 扩散不但和液体的粘性有关, 也和颜料分子的大小, 形状有关. 以上的示范, 该选用大分子颜料. 一般食物颜料好像效果都好. 当然, 最重要的是粘液本身必须够粘, 使粘液分子本身不扩散.

3. 自旋重聚实验

在此略提一下韩氏实验的原理. 读者可阅其论文以了解细节. 在固定的磁场内, 磁偶极 m 会作“进动”. 进动方程式是

$$\frac{dm}{dt} = \lambda h \times m \quad (4)$$

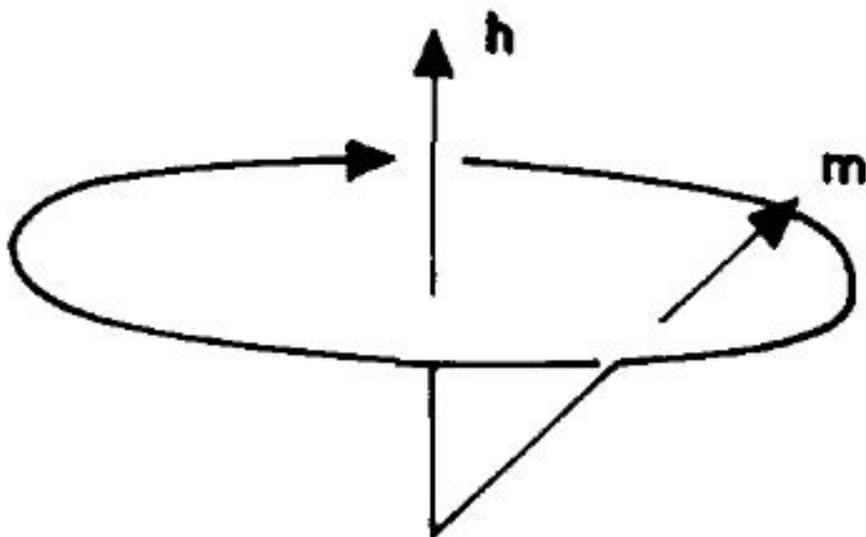


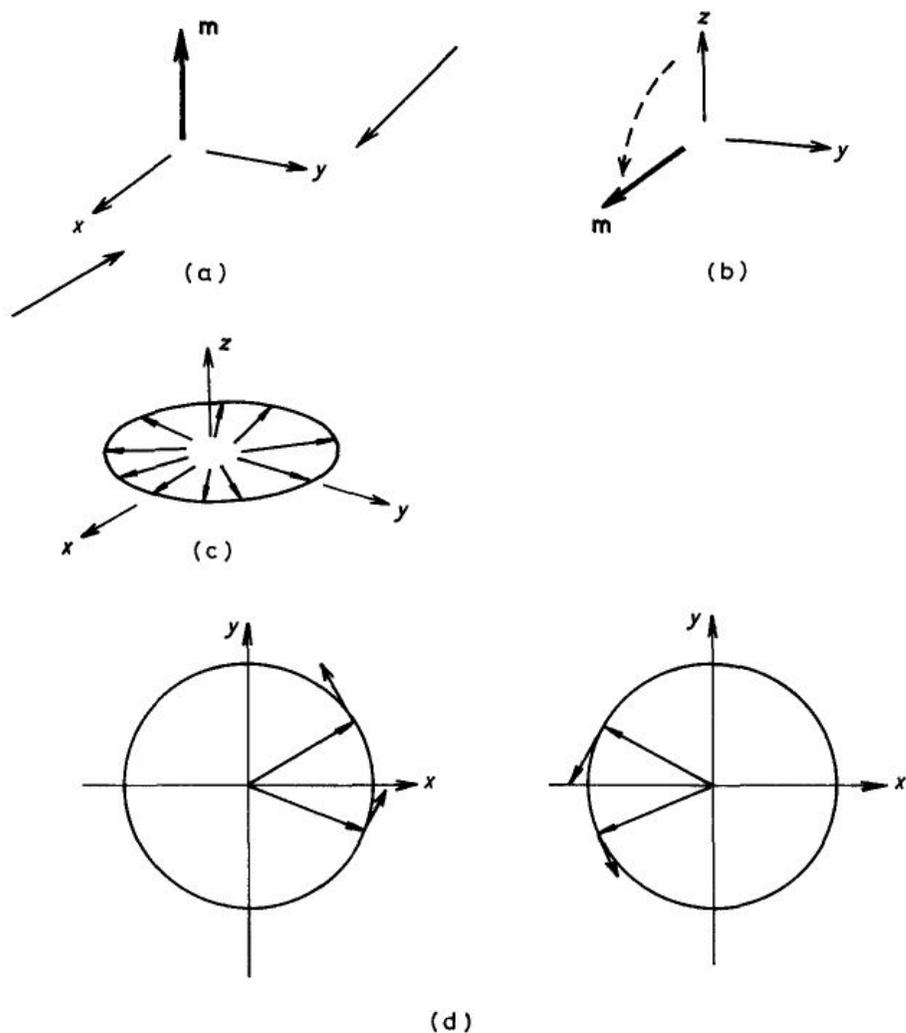
图 2

λ 为一常数, h 是磁场 (见图 2). λh 为进动角速度. 实验可用水中氢核子的自旋磁偶极.

如果我们用一个转动的坐标系, 转速为 λh , 则在这坐标系内, m 是不动的. 近代的技巧, 已经可以在这动坐标系内做许多实验. 韩氏的实验步骤大致是:

(甲) 用磁场 h (沿 z 方向) 将一试管水中的 H 核子磁偶极朝 h 方向排列 (见图 3(甲)).

磁场是稍稍有点不均匀. 物体各部分的磁场稍有不同. 令 h_0 为平均值.



(d)

图 3

(乙) 在转速为 λh_0 的坐标系中, 在 y 方向加一磁场使各 m 进动到 x 方向, 然后立即把磁场关掉. 这种闪电式的磁场, 可以说是一掌把 m 从 z 方向打到 x 方向, 转 90° (图 3(b)).

(丙) 因为各处磁场稍有不同, 在转动坐标系中, 各处自旋的进动速率也稍有不同. 过了相当时间 t 后, 各处自旋方向都不一样了, 成了混乱分布. 总磁矩为零, 量不出了. 图 3(c) 中的箭头, 代表各自旋方向, 平面图上画了其中两个自旋, 及进动速度, 快的在前, 慢的在后.

(丁) 再用 (乙) 的方法, 在 y 方向来个闪电式的磁场, 不过这一回磁场留

长一倍时间, 使所有的自旋都进动 180° . 如此整个图 3(c) 的一盘子箭头完全翻个身, 一如把烧饼翻面. 图 3(d) 指出 (c) 中的两箭头翻了身之后的模样. 注意, 翻身之后, 进动方向未改, 但以前慢的在后, 快的在前, 现在是快的在后, 慢的在前.

也就是说, 所有的相对角度全变了号. 如此, 过了时间 t 后, 各自又重聚了. 总磁矩又重现, 可以测得. 是为自旋重聚, 或“自旋回响”.

最后一步相当于粘液实验中的反向回推, 其他步骤也是异曲同工.

注意, 这实验中自旋方向分散的来源是不均匀的磁场, 是由磁铁产生. 虽然是不均匀, 但却不因时而变, 因此, 各自旋的进动速度虽稍有不同, 但都是一定的. 因此, 虽然各自旋方向有混乱的分布, 但没有混乱的运动, 因此没有熵的增加, 或减少.

4. 离子气体电荷重聚

和以上实验类似的有不少. 最有趣的大概是离子气体里的电荷重聚现象. 这现象比较复杂, 此地只略提一下原理.

在一个极稀疏的气体中, 平均自由时间很长, 观测时间若是比这时间短, 则碰撞就不会发生. 我们曾多次指出, 没有碰撞, 分子动量就不变, 也就不能供给熵了.

虽然没有碰撞, 但由于各粒子速度不同, 任何不均匀的密度分布, 在很短时间内会消失. 此所谓“兰道氏消失现象”. 例如, 在时刻 $t = 0$ 时, 若有分布

$$f(\mathbf{r}, \mathbf{v}, t = 0) = f_0 (1 + \alpha e^{i\mathbf{k}\cdot\mathbf{r}}) \quad (5)$$

$f_0(v)$ 为均匀平衡分布, α 为一小常数, 则以后的分布是

$$\begin{aligned} f(\mathbf{r}, \mathbf{v}, t) &= f(\mathbf{r} - \mathbf{v}t, \mathbf{v}, 0) \\ &= f_0 (1 + \alpha e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{v}t)}). \end{aligned} \quad (6)$$

在 t 很大时, $e^{-i\mathbf{k}\cdot\mathbf{v}t}$ 随 \mathbf{v} 改变得很快, 如果对 \mathbf{v} 积分, 就等于没有一样. 粒子的密度可由 (6) 积分, 得

$$\begin{aligned} n(\mathbf{r}, t) &= n_0 + \alpha e^{i\mathbf{k}\cdot\mathbf{r}} \int d^3\mathbf{v} f_0(\mathbf{v}) e^{-i\mathbf{k}\cdot\mathbf{v}t} \\ &= n_0 \left(1 + \alpha e^{i\mathbf{k}\cdot\mathbf{r}} e^{-\frac{1}{2}k^2 t^2 T/m} \right). \end{aligned} \quad (7)$$

即不均匀部分在时间 $1/k\bar{v}$ 后消失. $\bar{v} \sim \sqrt{T/m}$ 为平均速度.

这个离子重聚实验指出, 虽然不均匀分布消失, 但仍可以“重新出现”, 只是出现方式不甚直接而已. 也就是说, 即使 t 很大, $e^{-ik \cdot vt}$, 仍有其效果, 不能略去. 这个效果是利用离子气体的电荷测出. 大致步骤如下:

(甲) 在时刻 $t = 0$ 加一不均匀电场 $\propto e^{ik \cdot r}$ 于离子气体, 造成不均匀的粒子分布 (5). 关掉电场, 其后分布为 (6). 不均匀的电荷分布在时间 $1/k\bar{v}$ 后消失.

(乙) 在时刻 $t = t'$ 加另一不均匀电场 $\propto e^{-ik' \cdot r}$, 再关掉. 因此, 分布又多了一项

$$f_0(\mathbf{v})\alpha' e^{-ik' \cdot (\mathbf{r} - \mathbf{v}(t-t'))}. \quad (8)$$

电场关掉后 $1/k\bar{v}$, 密度又恢复了均匀. 但是这并非全解, 而只是和外加电场成正比的项, 即一次项. 还有高次项. 二次项中有和 $\alpha\alpha'$ 成正比者, (以 (6) 代 (8) 中的 $f_0(\mathbf{v})$ 即得):

$$f_0(\mathbf{v})\alpha\alpha' \exp [i(\mathbf{k} - \mathbf{k}') \cdot (\mathbf{r} - \mathbf{v}t) - ik' \cdot \mathbf{v}t']. \quad (9)$$

因此, 如果把 \mathbf{k}' 和 \mathbf{k} 定在同一方向, 则在时刻

$$t = \frac{k'}{k' - k} t' \quad (10)$$

(9) 式指数中含 \mathbf{v} 的各项相消, 得

$$f_0(\mathbf{v})\alpha\alpha' e^{i(\mathbf{k} - \mathbf{k}') \cdot \mathbf{r}}. \quad (11)$$

它的对 \mathbf{v} 积分不是零. 如果 \mathbf{k}' , \mathbf{k} 取得很相近, 则 (10) 式的 t 可以比 $1/k\bar{v}$ 大很多. 在这时候 (11) 造成一个不均匀的密度, 这个不均匀的电荷密度可以测得, 是为一种特别的重聚现象. 也就是说, 在 (6)、(8) 中的不均匀项看来都消失后, 它们的干涉效果却带来一个不均匀的密度分布, 重申它们的不可忽略.

因此, 不均匀密度的消失, 不能看作熵的增加, 它的重现, 不能解释成熵的减少. 在这实验的时间尺度下, 粒子之间无碰撞. 在动量空间里, 状态不变, 即粒子动量无起伏, 亦不供给熵. 第二定律也没有被违反.

讨论问题二十四

1. 试从状态空间观点, 分析粘液瓶示范. 令 $(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \equiv \mathbf{R}$ 为颜料分子位置.

- (A) 在开始时, R 的活动范围在什么地方?
 - (B) 分子被分散了之后, 活动范围的位置, 形状, 大小有何改变?
 - (C) 重聚之后又如何?
2. 离子气体重聚现象比较复杂, 但也可以从状态空间观点来分析. 读者试试.

第二十五章 从运动轨迹求熵

如果我们知道在观测时间内每一分子的详细运动经过, 则物体的任何性质都可以求出. 很多平衡的性质都是在观测时间内的平均值, 如压力、磁化强度、能量等. 虽然一般常见物体都太复杂, 但是简单的模型, 有上百上千粒子或自旋单元的, 已可以用计算机分析 (见第二十一章). 现在我们要讨论的问题是: 熵, 是不是也可以从这群体运动经过的数据中算出来? 这问题的回答应该是“可以”. 既然整个运动过程都知道了, 还有什么物理性质不能算的? 但是, 熵是和能量、压力, 这些量不一样, 它不是某一量的时间平均值. 在力学和电磁学中, 没有熵这个概念, 不过, 如果熵的测定须要超出整个运动数据之外, 那熵这个概念必须超出科学之外了. 现在, 要解决的问题是一个实际计算方法的问题, 要设计一套方法, 一步步地从运动经过的数据, 把熵算出来. 我们在此讨论一个特别简单的方法, 称为“巧合法”. 更好的方法一定存在, 只是有待发现而已.

读者或许会问, 为什么要讨论这个由运动数据求熵的问题? 讨论这问题有两个目的: 第一个目的是原则上的. 从第五章开始, 我们就一直反复说明熵和运动的关系. 许多例子的目的是指出熵的定义必须以运动为本. 但是这些例子都是十分简单的例子. 那些量是不变量, 那些量是不断地改变, 都非常清楚. 应用基本假设时, 只把改变的各运动变量当成变量, 不变的当成常数. 形象空间中的活动范围, 是各常数所容许的所有状态之集. 然而, 这种变和不变的划分, 不一定总是如此清楚. 尤其是分析“非晶”物体时, 在未把题目解出之前, 完全不知道那些量是不变, 那些量是改变. 无论如何, 活动范围的决定, 归根结底, 必须由轨迹定义.

第二个目的是实用, 一套从运动求熵的方法, 可以用来分析各式模型, 尤其是一些和暂稳状态有关的.

玻尔兹曼公式是

$$S = \ln \Gamma(E),$$

式中 S 是熵, E 是物体的总能量. $\Gamma(E)$ 是活动范围 Ω 的体积, 这区域中, 每个状态的能量都是 E . 本章的讨论仍是以此式为出发点. 我们的最基本立场是: $\Gamma(E)$ 须由轨迹本身来划定. “巧合法”是由轨迹计算 $\Gamma(E)$ 的一个程序.

我们先说明“巧合法”的原理, 再以易形模型为例, 说明计算步骤.

读者请注意, 本章的计算方法及技巧, 都是在草创阶段. 结果亦属初步. 把这些草创初步的东西拼入本书的原因, 主要是因为以轨迹定熵的想法十分重要而尚不普及*.

1. 巧合次数和区域大小

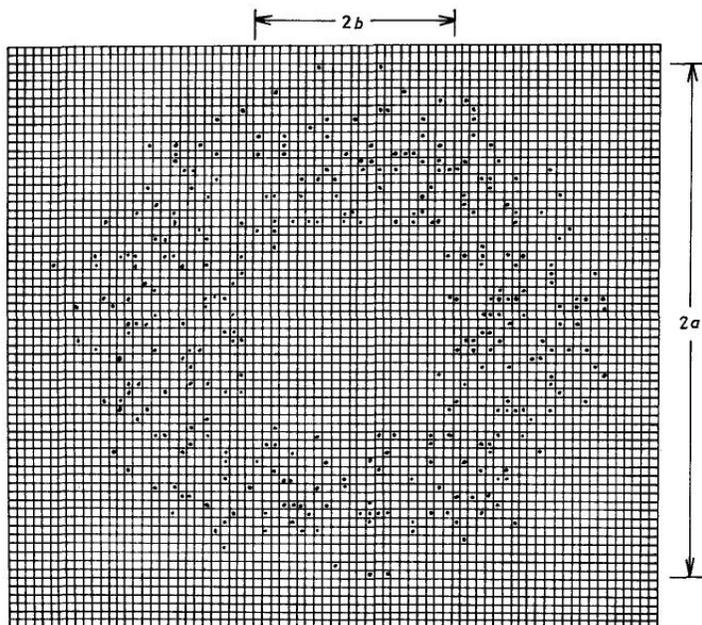


图 1

从巧合次数求熵的基本原理非常简单. 今借图 1 说明, 图示一群黑点零乱地散布在棋盘上的一区域. 散布的区域 Ω , 显然是个铜钱形, 这区域 Ω 的面积 Γ 是 $\pi(a^2 - b^2)$ 个格子, 在这 Γ 个格子中, 只有少数个格子有黑点, 大多数是空的. 如果我们没有这张图, 只知道棋盘上每个格子内的点数, 那这面积就不是这么容易求出了. 但是, 如果我们晓得这些黑点是在 Ω 中作完全不规则

*本章材料出自研究报告, 见马上庚 (1981), Ma(1981).

的, 完全散乱的分布, 则 Ω 的大小 Γ 可以估计. 方法如下:

令黑点总数为 n , Ω 中格子数为 Γ . 因为在 Ω 中分布是散乱的, 虽然 Γ 比 n 大很多, 但仍会有些格子中占着两个黑点以上. 任取一对黑点, 这两点落在同一格子内的概率是 $1/\Gamma$, 我叫它“巧合率” $R = 1/\Gamma$. 取 N_t 对黑点, 则“巧合次数” N_c 是 $N_t R$. N_t 最多可以有 $\frac{1}{2}n(n-1)$. 因此, 如果巧合次数测出了, 则 Γ 可以求出

$$\Gamma = \frac{1}{R} = \frac{N_t}{N_c}. \quad (1)$$

这个测 Γ 的方法, 可称作“巧合法”. 这法子很简单, 它用不着考虑分布区域的形状, 或其他几何性质. 这法子可以用来求熵, 每个棋盘格子可以想作一状态, 每个黑点可想作轨迹中的一状态. 用巧合法求熵, 就是从轨迹中挑出 N_t 对状态来, 每对相同的状态, 算一次巧合. 如有 N_c 次巧合, 则巧合率可照 (1) 算出, 熵可用

$$S = \ln(1/R) \quad (2)$$

算出, 这就是巧合计熵法的基本原理.

不过, 如要实际的应用, 还得先解决若干问题, 澄清一些细节. 我们先得复习一下轨迹和熵的重要性质, 再定解决方针.

以上我们假定分布在 Ω 中的点是完全散乱的, 也就是有平均的概率分布. 我们现在把这条件放松一些, 以包括不均匀分布的情形. 我们把轨迹中的诸状态分成组, 令每一组 λ 占去的时间为 \mathcal{T}_λ , 概率为

$$P_\lambda = \frac{\mathcal{T}_\lambda}{\mathcal{T}}. \quad (3)$$

\mathcal{T} 为总观测时间. 现在假定每一组 λ 中的各状态是平均地分散在一区域 Ω_λ 中, 令 Γ_λ 为 Ω_λ 的容积, R_λ 为此区域中求得的巧合率, 现在定熵为 $\ln(1/R)$ 的平均值

$$S = \sum_{\lambda} P_\lambda \ln(1/R_\lambda), \quad (4)$$

是为 (2) 的推广式. R_λ 为任选取 λ 组中一状态, 再任选一状态而得巧合的概率

$$R_\lambda = \frac{P_\lambda}{\Gamma_\lambda}, \quad (5)$$

P_λ 是得 λ 组中任一状态的概率, 而且恰巧选得先定之状态的概率为 $1/\Gamma_\lambda$. 令

$N_{t\lambda}$, $N_{c\lambda}$ 为在 λ 组内所取的比较次数及所得的巧合次数, 则

$$S = \sum_{\lambda} P_{\lambda} \ln(\Gamma_{\lambda}/P_{\lambda}), \quad (6)$$

$$\Gamma_{\lambda} = N_{t\lambda}/N_{c\lambda}.$$

至于如何分组, 则要看应用情况. 例如, 由于和环境的能量交易, 物体的总能量不保持一定, 则 λ 可用来指名不同的能量.

有两点要注意: ① 巧合法的用途在求 Γ_{λ} . 而概率 P_{λ} 在普通的数值状态计算中可直接由 \mathcal{T}_{λ} , (3) 求得, 或由其他方法获得. ② 分组不能分得太细, 细分的极端是每一状态成一组, (4) 变成

$$S = - \sum_s p_s \ln p_s \quad (7)$$

此地 s 是指名各状态. 但是如要测定 p_s , 我们必须测定轨迹在 s 状态的时间 \mathcal{T}_s . 这时间无法测定, 除非状态的总数很小使 \mathcal{T}_s 相当长. 常言统计力学和信息论相同, 这一点我们要澄清一下. 在信息论中, (7) 是一基本公式, s 指名符号, p_s 为其在使用中 (文字) 出现的概率, S 用来度量这些符号可以确定多少消息. 各符号相当于统计力学中物体的各状态, 符号在文字中出现相当于轨迹. 在信息论中, 符号的总数很小, 每个符号在使用中出现很多很多次, 所以每个 p_s 都可以正确地测出. 在统计力学中, 情形完全不同. s 的总数很大而轨迹短. 在一轨迹中, 只有一极小部分的状态出现. 因此, 一般情形下, 每一组 λ 必须包括很多 s 才能使 P_{λ} 为可测之量, 因此从直接测定的观点来看, (7) 在信息论中为一实用的定义, 而在统计力学中并无直接用途, 不合为熵的定义. 如果就 (7) 而以信息论来讨论统计力学, 则有失重点.

2. 物体各部的独立性

这巧合法可以用较少数的点来测定一个很大区域的容积, 即 $n \ll \Gamma$. 但是, 这法子显然是有其限度的. 刚才提过 N_t 最多不过 $\frac{1}{2}n(n-1)$, 所以

$$N_c \leq \frac{1}{2} \frac{n^2}{\Gamma}. \quad (8)$$

巧合次数 N_c 必须相当大. 即 $N_c \gg 1$, Γ 才是个有意义的数字, 因此 n 不能太小, 即

$$n \geq \sqrt{2\Gamma}. \quad (9)$$

因此, 轨迹必须很长, 因为须要测定的 Γ 总是非常大的. 由 N 个单元构成的群体, 它的 Γ 是以 N 的指数式增大

$$\begin{aligned}\Gamma &= e^S, \\ S &\propto N.\end{aligned}\tag{10}$$

但是轨迹的长度 n 是和 N 及观测时间 \mathcal{T} 成正比

$$n \sim N(\mathcal{T}/\tau)\tag{12}$$

即物体的状态在时间 \mathcal{T} 内改变了 n 次. τ 是每一单元改变一次所需的时间, 如果 N 很大, \mathcal{T} 在合理范围内, 则 (9) 是一个不能满足的要求. 这确是一个困难, 但这困难并不是太严重. Γ 的以 N 的指数式增大, 是由于物体各部大致互相独立 (见第十二章). 这是说, 间隔远各部的分子运动是互不相干的. 如果 $A(\mathbf{x}_1)$ 是在 \mathbf{x}_1 的某运动变量, $B(\mathbf{x}_2)$ 是在 \mathbf{x}_2 的一变数, 且 $|\mathbf{x}_1 - \mathbf{x}_2| \gg \xi$, 则

$$\langle A(\mathbf{x}_1) B(\mathbf{x}_2) \rangle - \langle A(\mathbf{x}_1) \rangle \langle B(\mathbf{x}_2) \rangle = 0\tag{13}$$

此地 $\langle \dots \rangle$ 是在观测时间内的平均值, ξ 是“相干距离”. 物体如比 ξ 大很多, 则是由许多大致互相独立的部分组成. 每一部分都比 ξ 大. 总熵是各部分熵之和. 各部分的总数是和 N 成正比, 所以 $S \propto N$ 而 Γ 随 N 的指数式增大, 因此, 只要 ξ 够小, 这“巧合法”可以用来求物体各部分的熵.

如果一物体有两个独立部分, A 和 B , 则巧合率 R 是 $R_A R_B$. 因为巧合在 A 内发生也同时在 B 内发生的概率是两分别发生概率之积, 因此

$$\begin{aligned}S &= \ln [1/(R_A R_B)] \\ &= S_A + S_B\end{aligned}\tag{14}$$

独立巧合率相乘和熵相加的特性是并行的, 如果把一物体分成大致独立的部份, 则

$$S = \sum_A S_A + \sum_{A,B} S_{AB}.\tag{15}$$

式中 S_A 是 A 部分的熵. 如果 A 和 B 相邻, 则不完全互相独立而需要 S_{AB} 作为修正. S_{AB} 可以从 S_A , S_B 和 S_{A+B} 求得

$$S_{AB} = S_{A+B} - S_A - S_B\tag{16}$$

式中 S_{A+B} 是由 A 和 B 合成的熵. 各部的独立性, 使总体的各性质可以由小部分 (比 ξ 大些) 的性质推测. 就因为这一点, 一般的数值模拟模型可以用来推断许多群体性质, 即使这些模型只用上几百或更少个运动单元. 这些模型虽小, 但只要比相关长度 ξ 大些就够了. 用更大的模型, 只把各部独立性显示得更清楚而已. 当然, ξ 的大小要看讨论的问题而定, 如果 ξ 很大 (例如在临界现象问题中) 则直接的模拟模型在今天的计算能力内是不管用的. 此地讨论的巧合法, 也只能实用在 ξ 较小的情形.

3. 相干时间

“巧合法”须要检验不相干的各状态. 虽然轨迹并非一系列完全散乱的点, 但相隔长时间发生的各状态是独立不相干的. 如果 A 为某运动变量在某时刻的值, $B(t)$ 为某变量在时刻 t 之后之值, 且 $t \gg \tau$, 则

$$\langle AB(t) \rangle - \langle A \rangle \langle B \rangle = 0. \quad (17)$$

τ 是“相干时间”, 如果观测时间比 τ 大很多, 我们可以选择相隔 τ 以上的状态来比较. (17) 指出相隔时间久的运动间的独立性, (13) 指出相隔距离远运动间的独立性. 这些特性是此地计算的必须条件, 在此地的讨论中, 我们只假定这些条件成立, 而不深究其来历.

基本原理的陈述到此为止. 当然还有许多原则上和实际运算的问题, 我们希望在以下讨论计算步骤时大致予以澄清.

4. 计算过程

现在, 我们把上述的巧合法用来计算易形模型的熵, 列出实际计算步骤, 当作一示范例子. 计算有两个阶段, 第一是演算出一段轨迹, 第二是求巧合次数. 第一段即是普通的模型计算, 其结果是第二段的输入数据. 我们主要的讨论是在第二段, 为求明确, 我们先把第一段交代清楚. (模型原理, 及细节, 见第二十章).

(甲) 易形运动模型

这模型的状态是由 N 个单元 s_1, \dots, s_N 的值确定. 每个单元的值是 $+1$ 或 -1 , 它的改变依照以下的运动规则: 从任一时刻 t 的状态, 我们可算出每

个单元的改变率 (单位时间内的改变概率):

$$W_i = \frac{e^{-s_i h_i / T}}{\cosh(h_i / T)}. \quad (18)$$

此处 T 是此物系的温度, h_i 是加在 s_i 上的引力

$$h_i = h + \sum_j J_{ij} s_j, \quad (19)$$

h 是外加力或“磁场”, 末后一项是各单元加在 i 单元的总力, J_{ij} 是单元 i 和 j 之间的作用强度, 是已知且不变. 现在运动规则已完全确定了.

轨迹是由以上概率, 及计算机内的随机数程序算出. 细节请看第二十二章, 第 3 节.

随机数的参与, 并不是必要. 在此用一随机模拟模型, 纯是为了方便. 由轨迹求熵的原则, 合用于任何轨迹, 唯一的条件是长时间、长距离外的独立性. 这种独立性也就是平衡的特性.

(乙) 计数程序

这程序是要把轨迹中的几个状态记录下来, 然后一对对抽出来比较, 看有多少次巧合.

以下所叙述的程序是一个简单而效率不高的程序, 目的是阐明“巧合法”的原理.

这程序不直接记录轨迹中的状态, 而只把一连串变值单元的名字 i_1, i_2, \dots, i_n 记下来. 要比较两个状态, 只须考虑这两形象发生时刻之间的改变. 例如, 要比较 i_2 前的状态 A 和 i_{10} 后的状态 B , 我们算一下 i_1, i_2, \dots, i_{10} 这名单中有几个出现单数次的单元. 一个单元如改值双数次, 就等于不改. 改值单数次的单元总数, 叫做 A, B 的“差别”, 如无差别, 则 A 和 B 为同一状态, 于是我们得到一巧合. 以下再谈一些细节.

(丙) 能量和形像的有效体积

这模型是一个定温物系, 能量有其起伏, 所以须要把状态按能量 E 分组, 计算 $\Gamma(E), P(E)$, 并用 (6) 求熵:

$$S = \sum_E P(E) \ln[\Gamma(E)/P(E)]. \quad (20)$$

E 组的概率可由模型程序供给,

$$P(E) = \frac{\mathcal{T}(E)}{\mathcal{T}}. \quad (21)$$

$\mathcal{T}(E)$ 是能量 E 出现的时间, \mathcal{T} 是整个轨迹的总时间, 因此出现的能量也必须记录下来.

为求增加巧合次数, 我们可以把巧合的定义放马虎一点. 比较一对状态时, 只要它们的差别不超过 m , 且能量相同, 就算是巧合. m 是一个小整数, 如 0, 1 或 2. 如 $m = 0$, 则巧合的定义和先前的一样. 令 V_s 为状态 s 的“有效体积”, 也就是和 s 差别不超过 m 的状态的总数. 也就是说, 如果 s' 和 s 之差别不超过 m , 并和 s 有相同能量, s' 就算在这体积内 (如 $m = 0$, 则 $V_s = 1$) 计算巧合率时, 比较次数必须照 V_s , 斟酌修正. V_s 必须在模型程序中算出, 并记录下来备用. 没有其他东西要记录了.

(丁) 程序的限度

以上所述的程序可以用来计算多大的熵? 一次分析多少单元? 答案自然要看是用在那个问题上. 不过我们现在可以粗略地估计一下一般的限度. 如果轨迹有 n 个状态, 分成 k 组, 每组中的 n/k 个状态有差不多一样的能量, 于是每组中的巧合次数约为

$$N_c(E) \sim \frac{1}{2} \left(\frac{n}{k}\right)^2 \frac{v}{\Gamma(E)} \quad (22)$$

式中 v 大约是一般 V_s 之值. N_c 越大, 求他的 $\Gamma(E)$ 越准, 如果要求 $N_c(E) \geq 25$ 则可以量出的熵最大限度是

$$S \sim \ln\left(\frac{v}{k^2}\right) + \ln\left(\frac{n^2}{50}\right) \quad (23)$$

为增大此限度, 可以把 v 加大, k 减小, 但实际上是否可以办到, 则要看是什么题目了. 如果 v 够大, (23) 右第一项不很重要, 则出现的状态的总数 n 可以决定 (23). 如果 $n \sim 10^6$, 则 $S \sim 23$, 即使 n 大很多, S 也不会大很多. 如 n 只有 1000, S 也不会小很多. 这当然是因为对数随 n 的变化很小. 如果每单元有熵 0.5 则这程序可以测量 $N \sim 40$ 个单元. 因此, 这程序的应用, 只限于很小的物体, 或者可以分成大致独立小部分的物体. 当然这要求限制了程序的应用范围, 但并非太严重, 可用范围仍相当广.

以上的程序, 已被用来计算“非晶磁体”模型的熵[†].

当然以上的程序可以改良的地方甚多, 最重要的是如何选择运动单元并分组, 使不相关的单元分在不同组里. 如此, 巧合计数法可以发挥最大的效用.

[†]Ma and Payne(1980).

5. 暂稳态的熵

许多暂稳物体, 都有大小悬殊的时间尺度, 例如在 H_2 气体中, H_2 的自旋 (1 或 0) 很少会因碰撞而改变, 所以其相关时间很长. 分子速度的相关时间, 即平均自由时间, 却很短. 因此只要观测时间长度是在这长短尺度之间, 平衡性质就十分明确. 长的不必管, 只要注意短的 (见第八章).

但许多物体中, 长的短的时间尺度没有明显的分野, 而是从长到短, 样样都有. 这种头痛的尺度分布, 常在“非晶”物质中出现. 状态空间被高低不一的屏障隔成许多区域, 低的易过, 高的难越. 因此, 轨迹分布的区域, 要看观测时间的长度了, 熵的值亦然. 我们可以大致估计一下熵如何随这时间长度而异. 越过一个高度 Λ 的屏障, 约需时

$$t \sim \tau_0 e^{\Lambda/T}. \quad (24)$$

式中 τ_0 是一短时间尺度, 即无屏障时的相关时间. 当 $t \gg \tau$, 高于

$$\Lambda_c \sim T \ln(t/\tau_0) \quad (25)$$

的屏障实在是过不了, 而低于 Λ_c 的可以轻易越过, 因此, 熵的值大致是以观测时间的对数而改变. 这种改变很小, 尤其是在低温. 熵的意义仍不含糊.

6. 热力学第三定律

在热力学中, 两平衡态, 1 和 2, 的熵之差是

$$S_2 - S_1 = \int_1^2 \frac{dQ}{T}. \quad (26)$$

1 和 2 必须由可逆过程相连. dQ/dT 是在此过程中, 一步步用实验测定的热容率. 除了一积分常数外, (26) 可作熵的实验定义. 这常数可以用高温气态的熵来决定 (用理想气体理论), 或令 $T = 0$ 时的熵为零 (热力学第三定律, 见第二十三章).

一般模型的热容率, 可以用数值模型算出. 再用 (26), 我们可以算出熵. 因此, 轨迹和熵可以连上. 不过, 这个计算是和以上讨论的完全不一样. 它必须要一连串其他平衡态的轨迹, 不只是观测下平衡态的轨迹, 所以不合乎由观测下轨迹求熵的要求.

在 (26) 式中, 1 和 2 之间的过程必须要是可逆的. 否则 (26) 的意义就不清楚了. 这个条件相当苛刻. 不可逆的情形是: 当温度加高再降回原来温度, 而轨迹从一区域转到了另一区域, 不回到原来的区域. 有时, 不可逆的现象不十分明显. 例如, 玻璃可以有許多不同的结构. 加温再冷却后, 结构就变了. 不过, 这不同的结构常有相同的大尺度性质, 很难察觉不可逆现象的发生.

当然, 不管可逆不可逆, (26) 可以用来作熵的定义. 这个定义, 是一般讨论玻璃所用的.

我们由高温的熵来定 (26) 所做的积分常数. 如果在高温和 $T = 0$ 之间有不可逆情形发生, 则积分所得的熵在 $T = 0$ 就可能不是零, 违反了热力学第三定律. 这是所谓“冻结熵”. 由轨迹求出的熵是不会违反第三定律的, 因为它直接度量轨迹在状态空间里的活动范围. 当 $T \rightarrow 0$, 运动停止, 这范围就缩小到一点了. 这运动停止的情况, 是第三定律的基本含义.

所以, 若有不可逆的情形发生, 则在某些温度下, 此地所定义的轨迹熵会与 (26) 定义的比热熵有差异. 这差异可用来作不可逆程度的一种度量. 注意, 如加热 dQ 则比热熵的改变 dS' 为

$$dS' = \frac{dQ}{T} \quad (27)$$

这即是 dS' 的定义 (26). 但是, 轨迹熵的改变 dS 就可能大些

$$dS \geq \frac{dQ}{T} \quad (28)$$

因为加热过程可能是不可逆.

注意, 在去热时, 即 $dS < 0$ 时, 熵的降低比 dQ/T 要多, 即 (28) 的反向过程. 这是因为温度降低, 时间尺度也改变, 轨迹被限制之故. 但这并未违反热力学第二定律. 如果去热时熵降低得多了些, 但加热时熵又增加得多了些. 因此无法用这多降之熵把热转成功能.

讨论问题二十五

1. 一群点在平面上, 或任何连续空间中的分布, 都可以用巧合法测量它的散乱程度. 由图 1 可见, Γ 是和格子数成正比 (见 (1)), 和一格子大小 v 成反比. 令 $S(v) = \ln\Gamma(v)$.

(A) 以 $S(v)$ 为 $\ln(1/v)$ 的函数作图. 如分布完全散乱, 该得何种图形?
(答: 直线, 斜率为 1).

(B) 用此法考验一随机数列. 从计算机取 $(0, 1)$ 平均分布数列.

(C) 用此法考验数列

$$X_m = \sin(2^m) \quad , \quad m = 1, 2, 3, \dots \quad (29)$$

注意, 如果有 n 个点, 有格子 $i = 1, 2, \dots, M$, 格子 i 中有 n_i 个点, 则 (1) 的 N_t 和 N_c 是

$$\begin{aligned} N_t &= \frac{1}{2}n(n-1) \\ N_c &= \frac{1}{2}\sum_{i=1}^M n_i(n_i-1). \end{aligned} \quad (30)$$

注意, 在 (29), 2^m 事实上只是 2^m 除以 2π 的余数, 这些余数, 照我们的直觉, 该是平均分布在 0 和 2π 之间, 因此, (29) 的各 x_m 不会作平均分布.

复习问题十一, 3, 4.

- 从纯粹牛顿力学观点来看, 状态空间是一连续性的空间. 因此, 如无量子力学, 熵必须由画格子来定义. 因此, 熵的值和格子大小有关. 但如格子取在适当范围之内, 使熵和格子大小成上题所说的直线关系, 则熵的定义不含糊, 试讨论此问题. 注意, 由此可见, 牛顿力学至少须在某一尺度内, 造成先后独立的轨迹, 使一群散乱的点得以取出.
- 第 4 节讲的一些细节, 只是用来示范, 而不是很有效率的方法. 计算机专家们有不少比较数目字的方法. 例如, 有一列数字 $A_i, i = 1, 2, \dots, n$. 有所谓“排列程序”, 把这些数字依大小排列起来. 相同的数字, 因而在排成后相邻. 此类程序, 所花时间约和 $n \ln n$ 成正比.

(A) 如何把物体的状态表成一数字?

(B) 如何用排列程序求巧合率?

试讨论各细节.

- 另有一种方法, 可以用从轨迹求熵. 其推导如下. 令 $s_i, i = 1, 2, \dots, N$ 为运动变量, $P(s_1, s_2, \dots, s_N)$ 为这些变数的联立概率. 把熵写成

$$S_N = - \sum_{s_1} \sum_{s_2} \dots \sum_{s_N} P \ln P. \quad (31)$$

(A) 令

$$P(s_1, s_2, \dots, s_N) = p(s_1 | s_2, \dots, s_N) P(s_2, \dots, s_N) \quad (32)$$

$P(s_1 | s_2, \dots, s_N)$ 为 (s_2, \dots, s_N) 一定下 s_1 的概率, $P(s_2, \dots, s_N)$ 为 (s_2, \dots, s_N) 的联立概率.

试证

$$S_N = - \sum_{s_1} \sum_{s_2} \cdots \sum_{s_N} P(s_1, s_2, \dots, s_N) \ln p(s_1 | s_2, \dots, s_N) + S_{N-1}. \quad (33)$$

S_{N-1} 是 (s_2, \dots, s_N) 的熵.

如果 s_1 只和邻近的少数单元 s_2, \dots, s_m 有关, 则

$$p(s_1 | s_2, \dots, s_N) = p(s_1 | s_2, \dots, s_m), \quad m \ll N. \quad (34)$$

(B) 令 $P(s_1, s_2, \dots, s_m)$ 为少数 m 单元的联立概率.

试证

$$S = - \sum_{s_1} \cdots \sum_{s_m} p(s_1, s_2, \dots, s_m) \ln p(s_1 | s_2, \dots, s_m) \\ - \sum_{s_2} \cdots \sum_{s_m} p(s_2, s_3, \dots, s'_m) \ln p(s_2 | s_3, \dots, s'_m) + S_{N-2}. \quad (35)$$

(C) 以上每一步分出一单元. 最后 S 中的每一项都只含条件概率, 只和邻近的少数单元有关. 如果这些条件概率可以从轨迹求出, 则 S 就可以算出.

试把 (A), (B) 步骤用到一度空间的易形模型.

(D) 试分析二度空间易形模型. 如果每个条件概率是四个邻居形象一定下的概率, 则此法就可以说是非常理想. 但可惜免不了少了几个邻居. 因此, 条件概率反而得包括距离较远的单元, 免不了涉及相干距离内的许多单元.

(E) 试讨论此法的优点与缺点. 这方法用途甚广, 原先并非用来分析数值模型的, 文献中的推导亦和以上的不大一样, 读者可参考 Kikuchi (1955), Alexandrowicz (1975), Meirovitch (1977).

5. 试将本章所定的熵, 用到上一章粘液瓶示范的分析上. 颜料分子的轨迹是什么? 如何定巧合率?

从上章问题 1 开始.

6. 物体各部份的大致独立, 为定义熵的必要条件. 不过这些“部份”并不一定要按照位置来分. 只要可以把 N 个变量分成 m 组, $m = O(N)$, 各组大致独立, 就行了.

(A) 讨论理想气体的情形. (提示: 按动量分).

(B) 讨论晶体振动情形.

第二十六章 基本假设的来源

统计力学的发展很奇怪. 在序中提过, 它是一门“头重脚轻”的学问. 从基本假设出发, 主要是向上发展. 物理、化学、材料等各行各的专家, 把它应用到各种模型上, 成功地解释了不少现象. 数学家把它美化、抽象化, 用严密的证明导出它的一些后果. 向下的发展, 就比较差了. 物理学家们对基本假设的根源, 虽有兴趣, 但不很热衷, 这问题太难了.

物理学家本来没有逃避难题的习惯, 大概感于应用的成就, 看轻了了解根源的重要. 于是这向下探源的工作, 虽本是物理学家的事情, 却交到一些大数学家手中. 只可惜数学家喜欢的是漂亮的数学, 有时会把物理的一些要求忽略了.

事实上, 基本假设的了解, 不只是一个物理问题. 也是对所有不规则现象的了解. 如何从规则的运动定律, 导致不规则的结果? 这是一个很普遍的问题.

原理方面是一难题, 应用方面, 也并非容易. 虽然应用相当成功, 但是大多计算都用不一定可靠的近似法. 照以上诸篇的经验, 一旦相互作用不能略去, 则问题变得非常困难. 这困难和对基本假设的难题, 大概是分不开的.

在这一章, 我们先总结一下本书对基本假设的看法, 也就是对统计力学的基础的看法. 再讨论“遍历性”、“系综”、“混合性”之类的概念. 然后对传统的一些看法作一检讨和批评. 轨迹的“不稳性”, 在本章也略作介绍.

这一章的讨论, 不能算是对统计力学现状的客观描述, 而是本作者的一些主观意见. 这些主见, 在以上诸章已反复说过多次. 此地再总结一下.

1. 基本假设

现在先复习一下基本假设在以上各篇的讲法. 我们把运动变数分成两类; 一为不变者, 即在观测时间内不变者, 二为快速变动者, 即在此时间内不断地

改变者. 然后我们定义“活动范围”为所有变动变量所包括的状态之集, 即合乎不变者条件下 (即不变者为定值), 所有状态之集. 以上诸篇再三强调; 不变者是在观测时间内的不变者, 视时间长短而异, 在许多情形下, 那些变, 那些不变, 并非显而易见. 活动范围的严格定义, 不免要用到详细运动轨迹.

令 Γ 为活动范围的体积, 即范围内状态的总数. 基本假设是: 熵为 Γ 的对数

$$S = \ln \Gamma. \quad (1)$$

这假设还要加上以下的条件才有用, 这条件是

$$S = O(N), \quad \text{即} \quad \Gamma = O(e^N). \quad (2)$$

N 是粒子数或变数的总数. 从 (1), (2) 整个热力学可以推出, 所有的平衡性质可以算出. 这是第二篇的结论, 也是第二十三章的讨论内容. 注意, 条件 (2) 极为重要, 没有它, 就没有热力学. 在第四篇以下, 我们把这条件看成各运动变量的大致独立性. 因此, (2) 就是独立性的假设. 在讨论中央极限定理时 (第十二章, 及在第六章末), 我们指出, 大值变量 ($O(N)$ 变量) 的时间平均值和它在活动范围中的平均值相同. 再者, 大值变数的起伏, 循常态分布, 为 $O(\sqrt{N})$. 这个结论也可以写成

$$P(R) \equiv \frac{\mathcal{T}(R)}{\mathcal{T}} = \frac{\Gamma(R)}{\Gamma}, \quad (3)$$

\mathcal{T} 是总观测时间, R 是一子集, 其容积为 $\Gamma(R)$ 且大小和 Γ 相若:

$$\Gamma(R) = O(e^N). \quad (4)$$

$\mathcal{T}(R)$ 为轨迹在 R 内的逗留时间. 这些也是 (2) 的结果 (见第十二章). $P(R)$ 可以解释成状态在 R 内的概率. 注意, \mathcal{T} 为有限时间, 绝不能看成 $O(e^N)$ 的大数. R 必须是大子集, 否则 (3) 无意义, (见第十二章第 9 节). 大值变数的起伏, 和热位能的微分有直接的关系 (见第十三章). 此地“大值变数”指物体各部变数之和, 如总磁矩是各部磁矩之和. 其平均值可能是零. 此类变量也包括了像

$$\int d^3\mathbf{r} \rho(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}}$$

这样的变量, 它是密度乘 $e^{-i\mathbf{k}\cdot\mathbf{r}}$ 之和. 总之, 上述的基本假设, 足以用来分析所有热力和散射实验.

总而言之, (1) 是一个了不起的大胆假设, 它把微观运动变量经 Γ 和热力学的熵连了起来. 从推理的观点来看, (2) 是这假设的根本. 要了解统计力学, 必须了解 (2).

现在我们先提一下“遍历性”和“系综”, 这是统计力学发展过程中两个重要的概念.

2. 遍历性和系综

上述基本假设, 并不是一个过分苛刻的要求, 因为我们只讨论大值变量, 及大区域. 读者须注意, 在文献中, 这基本假设以各种不同的方式出现, 内容亦常不同. 玻尔兹曼写下这假设时, 他的想法是: 轨迹会经过活动范围中的每一状态, 即有“遍历性”, 因此无限长时间平均等于范围平均. 这想法是错的, 因为这无限长的时间必须比 $O(e^N)$ 长很多, 而一般观测时间为 $O(1)$. 要使任何一变量的时间平均等于其范围平均, 是一苛刻的要求. 本书说法并不作此要求, 而只要求大值变量的时间平均等于其范围平均. 这要求不苛刻, 因为大值变量之值, 在活动范围中几乎到处一样, 轨迹不必经过每一状态, 即不必“遍历”.

轨迹的遍历性, 一直是力学中的一大问题*. 许多数学家致力于证明某些模型的遍历性. 只可惜成果对统计力学还没有很大的帮助.

读者请注意, 文献中名词的定义并不统一. “遍历”一词在不同的文章内可有不同的含义. 有些地方它的意义如上述. 有些地方, 它指轨迹“大致”经过每一点. 有些人则称基本假设成立的情形为“遍历”, 凡基本假设不合用为“不遍历”, 真是一团糟. 读者必须要了解当时意义, 以免受害.

玻尔兹曼的“遍历”说, 立刻被人攻击, 他的一整套理论, 包括 H 定理, 也都被人怀疑. 大概因此他忧愤而终.

玻尔兹曼以后的统计力学, 教吉布斯打扮了一下, 成为今天的一般说法, 也就是所谓“系综”说.

系综说是把一物体的平衡态想成一“系综”, 即无限多个结构相同的物体. 每物体的状态都是形象空间中的一点. 这些点分布在状态空间内, 可想成一“理想气体”(每一点为一“分子”, 需要至少 $\sim e^N$ 个“分子”来定义分布).

平衡态的性质, 是这系综的平均性质. 这种抽象的看法, 有很多好处, 因

*见 Arnold (1968), 杨维哲 (1978) 第六章, 及其他“Ergodic Theory”方面的书.

为我们对一团气体的流动、分布有相当强的直觉. 吉布斯因这想法, 推出了不少有用的公式 (相当于第七章的算法则), 和应用结果. 系综也成了传统的基本概念.

不过这系综只是一个抽象的东西. 事实上是只有一个物体, 没有无限多个. 这一点吉布斯本人在他书中也指出, 但今天却常被忽略. 系综这概念, 只是数学概念, 它并不能解决玻尔兹曼假设的问题. 如果玻尔兹曼的“遍历”被斥为不合事实, 那这系综更是不合事实了. 系综说的内容和玻尔兹曼假设相同, 它把系综分散在活动范围. 系综平均即活动范围平均.

吉布斯之后, 许多学者企图把这系综概念更抽象化, 或赋于更实在的意义. 今天, 最普遍的想法是以系综表示实验者对物体的知识[†], 或“不确定程度”. 这种不确定的程度是看成一种“事前概率”, 也就是说, 在观测之前, 物体状态有一个概率分布. 对这种看法, 在以上诸篇已批评了不少, (尤其是在第二十四章). 在下一节我们再把这种看法和本书的看法作一比较, 并作一检讨. 在此我们先提一下系综观点的一些常识.

系综即是状态空间中的一团云雾, 它的密度可用一分布函数 ρ 表之. 现在假想状态空间是一 $6N$ 维空间, 物体是 N 个粒子, 其运动遵守牛顿定律. 吉布斯讨论 ρ 的改变, 并把熵定义为

$$S_G = - \int \rho \ln \rho d\Gamma, \quad (5)$$

此地的 ρ 适合归一条件 $\int \rho d\Gamma = 1$. (5) 的积分范围是整个状态空间.

系综中的每一系统的运动都遵守牛顿定律, 走出一条轨迹来. 各系统走出的轨迹不会相交, (如相交, 则表示同一起始状态可导致不同的运动轨迹). 读者在学力学时, 必曾学过“刘维定理”, 即沿着轨迹, 密度 ρ 不变. 现在在状态空间中划定一区域 $R(0)$, 并令域中每点都照力学定律走动. 这些点在一段时间 t 后成了另一区域 $R(t)$. 根据刘维定理, $R(0)$ 和 $R(t)$ 的容积是一样的, 只是区域的形状会改变. 从这些结果可证明吉布斯的熵 (5) 是个不变量. 显然 (5) 不能用来讨论熵的增加现象. 许多人设法解释、修补这个“熵不变”的不合理结论. 读者不难想见, 如果 $R(0)$ 是个超球似的区域, 在长时间后, $R(t)$ 就会像变形虫一般地走了样, 再拉成了细长的面线状物, 蜿蜒曲折地纠成一堆. 如果在 $t = 0$ 时 ρ 只在 $R(0)$ 内非零, 则在时刻 t , ρ 只在这种面线中非零. (5) 是这面球或面线的体积, 并不代表这堆面线的混乱程度. 我们不讨论 (5) 的修

[†] 见 Tolman (1962).

补, 因为它是建立在一个不正确的概念上. 平衡态本来就不是一个瞬时概念, 熵自然是须要一段时间来定义. 这一点从本书一开始就强调. 系综概念企图用无限多个物系的瞬时性质来代表平衡, 是一个不合理的做法. 这一点在第 5 节再仔细讨论.

3. 混合性和独立性

近代研究统计力学根源的学者, 多致力于“混合性”的分析. 所谓混合性, 即指上节的“面线” $R(t)$ 均匀地分布在活动范围中, 如果 $t \rightarrow \infty$. 即

$$\lim_{t \rightarrow \infty} \frac{\Gamma(R(t) \cap R')}{\Gamma} = \frac{\Gamma(R(0))}{\Gamma} \cdot \frac{\Gamma(R')}{\Gamma} \quad (6)$$

R 是活动范围的任一子集, Γ 是活动范围的总体积, $\Gamma(R')$ 是 R' 的体积. 也就是说 $R(t)$ 在 R' 内的体积, 和 R' 的体积成正比.

混合性比起遍历性来, 要合理些. 它的要求是任一区域 $R(0)$ 都会演化成一团平均分布在活动范围里的细丝. 这混合性的分析是一复杂的数学问题, 本作者是门外汉. 有关这方面的书籍, 最近有一些[‡].

混合性的理论, 可以说是现在普遍被接受的想法. 它比玻尔兹曼、吉布斯时的思想进步, 但仍可以说是由吉布斯理论演化而来.

现在把本书的看法改写成一个和混合性有点相似的式子, 再作总检讨. 本书的看法是: 基本假设即是一独立性的假设, 也就是相干时间及相干距离的假设.

在第 1 节我们指出独立性 (2) 是基本假设的根本. 从上一章轨迹求熵的讨论可见, 活动范围必须由轨迹来定义, 来了解, 而独立性是用轨迹定义熵的必要性质. 这独立性是长时间外和长距离外的独立. 也就是说: 有相干时间 τ 和相干距离 ξ . 如果 $|t_1 - t_2| \gg \tau$ 或 $|\mathbf{r}_1 - \mathbf{r}_2| \gg \xi$, 则

$$\begin{aligned} & \langle A(\mathbf{r}_1, t_1) B(\mathbf{r}_2, t_2) \rangle \\ & \equiv \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} dt A(\mathbf{r}_1, t + t_1) B(\mathbf{r}_2, t + t_2) \\ & \longrightarrow \langle A(\mathbf{r}_1) \rangle \langle B(\mathbf{r}_2) \rangle, \end{aligned} \quad (7)$$

[‡]参考 Sinai (1977), Krylov (1979).

平均值是在观测时间 \mathcal{T} 内的时间平均, \mathcal{T} 可以有相当伸缩性, 即 \mathcal{T} 的改变不影响平均值. A, B 是在 $\mathbf{r}_1, \mathbf{r}_2$ 附近的“局部”运动变数. “附近”是指比 ξ 小的距离. “局部”即是只和少数运动变数有关.

如果能从力学定律推导出 (7), 则活动范围可定, (2) 亦成立, 整个基本假设就建立了. 以上并没有提供从 (7) 得 (2) 的数学步骤. 这些步骤, 恐怕不是太容易的, 或许需要些其他条件. 但主要的一步是要了解 (7), 即要了解独立性如何从力学产生.

此地所谓的独立性 (7) 和混合性 (6) 有何不同? 现在看一个简单的例子.

假想两个一度空间模型: (甲) 粒子被限在硬壁盒子中, 位置是 $0 < x < L$. (乙) 粒子被一弹簧紧住, 而作简谐运动, 令其频率为 ω . 现在看看此二模型的“混合性”如何.

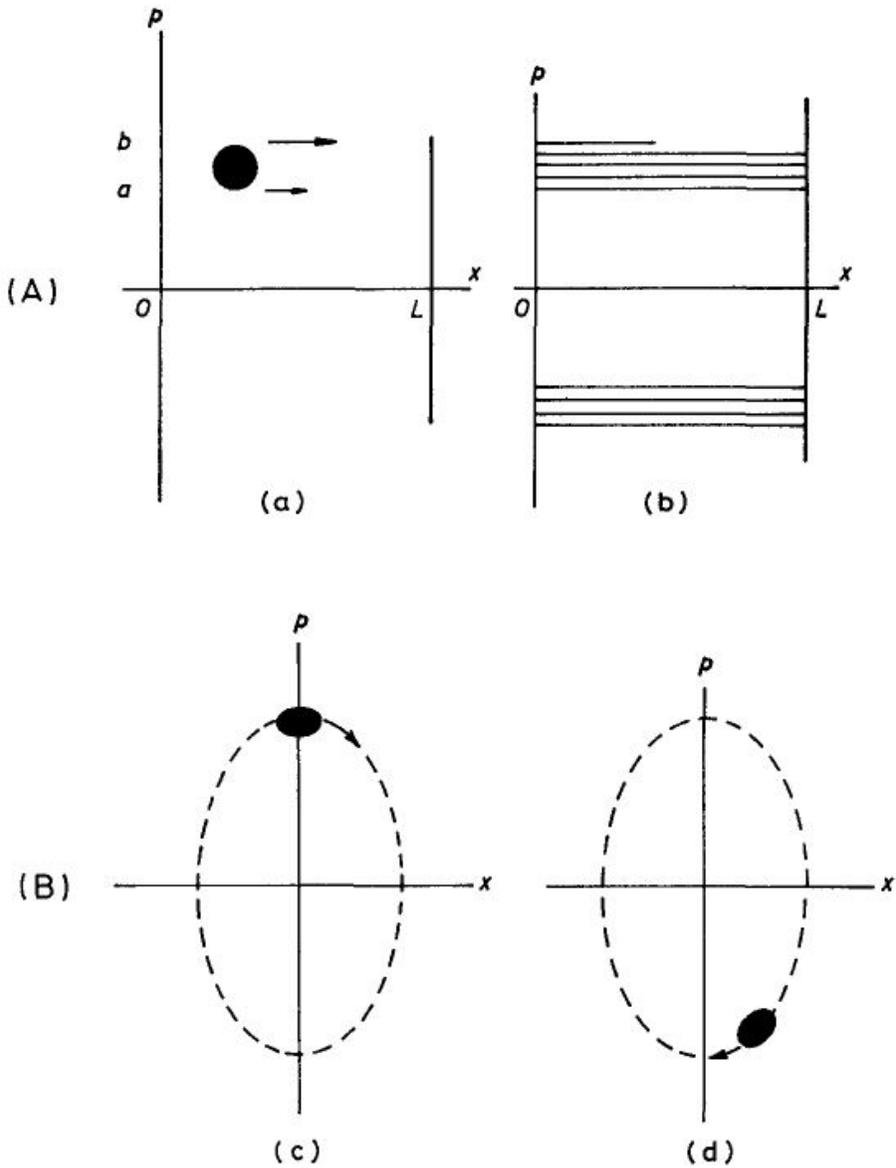


图 1

图 1(a) 中的黑饼区域, 经过一段长时间后, 变成图 1(b) 中的黑面条形区域. 时间越久, 这些面条越细越密. 黑区的总面积不变, 只是把这黑饼拉成细丝而已. 显然这些细丝均匀地分布在

$$a < |p| < b \quad , \quad 0 < x < L. \quad (8)$$

的区域中 $\Delta p = b - a$ 是黑饼的直径. 这均匀的分布在 $t \gg \tau$ 时发生

$$\tau \equiv \frac{\Delta p}{mL}, \quad (9)$$

m 是粒子的质量. 我们可以说这模型是有混合性的.

模型 (乙) 不大一样. 图 1(c) 中的黑饼会沿着椭圆轨道走, 形状会稍变, 但不会被抽成丝. 无论时间多长, 像图 1(b) 那样的情形不会发生 (见图 1(d)). 因此, 这模型没有混合性.

因此, 从混合性的分析来看, 这两模型完全不同.

但是, 从独立性的观点来看, 此二模型没有什么不同. 即二者均无独立性. 二者都是周期性的运动, 都没有纷乱现象发生.

本书的观点不把起始的状态看成一群点集, 而只考虑一个起始状态. 模型 (甲) 的运动周期由 $|\mathbf{p}|$ 决定, 是一个常数.

以上的例子自然过份简单. 且不论在一般复杂的模型中 (6) 和 (7) 是否大不相同, 至少这例子指出基本观点的不同.

4. 概率和实验

今天的文献中, “概率” 常指和吉布斯的“系综”一样的东西. 讨论平衡, 讨论熵的增加, 总不免以概率分布的演化为对象, 看看这概率如何从一起始的分布 (或“事前概率”), 逐渐分散开来, 布满整个活动范围. 相干时间的概念, 是指这概率分散开来所需要的时间.

本作者以为, 这“概率”的意义, 非常不清楚. 现在我们且不谈量子力学. 统计力学的概念应合用于古典力学. 如果不先把古典力学的情形弄清楚, 量子情形大概也弄不清.

一种说法是: 这概率是代表观测者的知识. 他对物体知道得越多, 则概率分布越集中. 这种说法显然不合适. 物体的运动, 和观测者的心理状态不该有关系. 照这说法, 不同的人看同样的物体时, 就必须说这些人的知识完全一样. 这说法难以接受.

另一说法是: 这概率代表实验的结果, 代表实验量出的知识. 这表面上看来比上一说法合理, “实验” 这名词给人一种客观的印象. 但这说法和上一说法大同小异. 问题是: 什么实验? 要做多少实验才够? 或曰: “宏观实验”、“可能做到的实验”. 这答案不对的. 从固定杂质的讨论, 及二十四章的重聚现象的分析可见, 有时每一杂原子的位置都得算是已知. 这是可能做到的实验吗?

统计力学文献中讨论的“实验”，往往和真的实验没有什么关系。例如，“正则系综”代表确定的温度，实验者可以确定能量到 $O(1/\sqrt{N})$ 的精确度。其他大值变量也可测准到 $O(1/\sqrt{N})$ 精确度。事实上， $N \sim 10^{22}$ ，故 $N^{-1/2} \sim 10^{-11}$ 。没有什么实验可做到如此高的精确度，且不论总能量该如何测定。再者，如果“概率”趋于零，则解释成“不会发生”，或“实验看不到”。这也不合事实。例如，任何暂稳态，其“概率”都趋近于零，包括钻石。今天的实验室中的物体，大半是在暂稳态中。

文献中的一大弱点，就是把活动范围，或平衡态的概率分布，看作理所当然的为已知。总能函数 H 一定，平衡概率就很简单地定了，即所有能量为一定值 E 的状态有相同的概率。这种看法，可以说是研究气体留下来的习惯。（本书已再次说明，活动范围的决定，不是这样简单，而要看真正的轨迹如何分布）。这习惯对分析统计力学的根源，是一大障碍。

本作者认为，吉布斯的系综及概率，和力学扯不上关系，对统计力学而言，是不必要且不合事实的概念。这些概念，虽可以用来帮助我们作抽象的分析，但对探讨统计力学的根源来说，是无关的。如要对这根源有正确的分析，必先摆脱这种系综概率想法的控制。

今天的计算机，已可以算出一些简单模型的轨迹，已可以从轨迹算出各热力学性质和一切性质，（见第二十二章，及第二十五章）。所以“实验”的问题，暂时不必介入根源的讨论。这些计算机的数据，代表完全客观的性质。当然，了解这些简单模型并不足以奠定统计力学的基础。但如果连这些模型的结果都不能了解，则真是“前途无亮”了。数值解只需要牛顿定律，根本不须要系综、概率的概念。

本书对概率的定义大不相同。本书把概率看成一个整理资料的工具，不是一个物理概念，不是一件必须品，纯是为了方便。必先有数据，有数据，才能统计、分类、综合、计算概率，平衡或不平衡，轨迹是所有的数据，一切性质必来自轨迹。

独立性 (7) 是轨迹的性质。活动范围须由轨迹定。平衡是轨迹的特性，因此不是一瞬时特性。在系综的看法中，平衡是一个瞬时的概率分布。读者到现在，应该对这两种看法的不同，分辨得相当清楚。

以上对系综和文献中的概率观，作了十分严格的批评，是为了澄清最基本的概念。但是，以上的批评，并不是对文献中成果的价值予以否定。以上提过，我们对空间，对一团云雾的流动、分散、有十分强的直觉。系综的看法，给我

们一个新的抽象工具, 而且颇具一种美感. 假想的数学概念自然不该避免, 但要认清楚, 并作有效的利用.

5. 轨迹的不稳性

要从力学定律了解混合性或独立性, 是很困难的事. (6) 和 (7) 看来都和轨迹的细节相去甚远. 现在来讨论一下“不稳定性”, 看一些细节. 不稳定性的意义大致如下. 令 $s(t)$ 为轨迹. 如果起始状态 $s(0)$ 稍改, 成 $s(0) + \delta s(0)$, 则轨迹的改变为 $\delta s(t)$. 如果

$$\delta s(t) \propto e^{\alpha t}, \quad (9)$$

$\alpha > 0$, 则轨迹为“不稳”. 如果有许多不同的 $\delta s(0)$, 成一小团云雾, 堆在 $s(0)$ 四周, 则这堆云雾就会散开. 这合乎混合性的要求. 这也就是所谓“失之毫厘, 差之千里”.

在第 3 节的例子, 模型 (甲) 中, 如果开始时 p 改为 $p + \delta p$, 则

$$\delta x(t) = \frac{\delta p}{m} t, \quad (10)$$

这也是不稳性, 只是不如 (9) 的严重.

现在看一较复杂的例子. 假想一粒子和一些固定的硬球碰撞, 如图 2. 现在只看一平面的运动模型. 此粒子的状态是由粒子的位置和它的速度方向指定. 速度的大小不变. 现在且只看速度的方向. 轨迹由原点开始. 如果方向改了一点点 $\delta\theta(0)$, 则轨迹就改了. 每碰撞一次 $\delta\theta$ 都要改变. 假定硬球密度很低, 粒子的平均自由距离 λ 比球半径 a 大许多. 图 2 示两次碰撞的情形. 由图可见

$$\delta\theta' = 2\delta\alpha = 2\frac{\lambda\delta\theta}{a \cos \alpha}, \quad (11)$$

因此, $|\delta\theta'| \geq (2\lambda/a)|\delta\theta|$. 在 n 次碰撞之后,

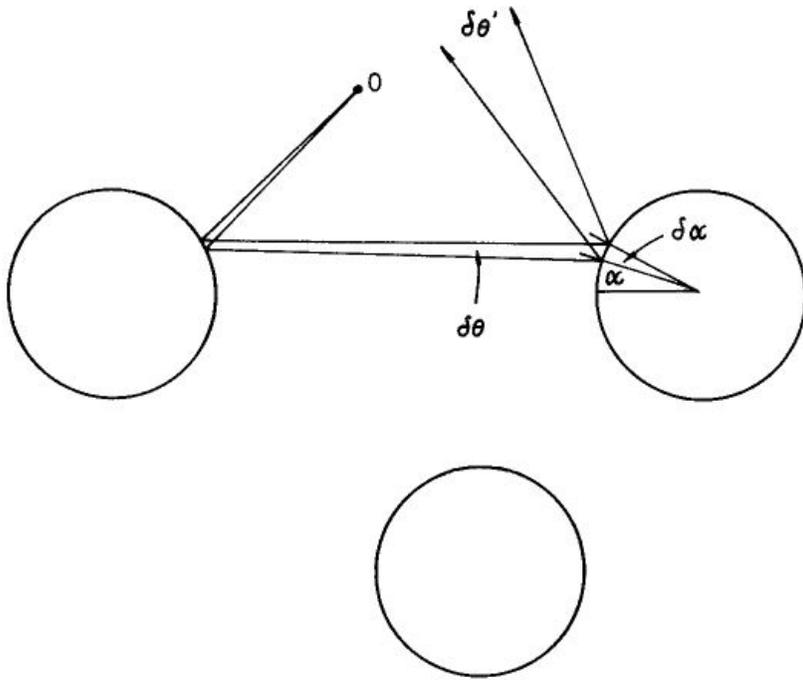


图 2

$$|\delta\theta(t)| \geq |\delta\theta(0)|e^{t/\tau},$$

$$\tau \equiv \lambda / \left(v \ln \frac{2\lambda}{a} \right). \quad (12)$$

v 是粒子速度. 这结果和 (9) 相同. 如时间够长, 即

$$t \geq \tau \ln(1/|\delta\theta(0)|)$$

$$= \frac{\lambda \ln|1/\delta\theta(0)|}{v \ln 2\lambda/a} \quad (13)$$

则 $|\delta\theta(t)| \sim 1$. 也就是说方向“乱了”. 以上的分析, 可以推广到气体模型, 结果差不多. 在此略提一下.

假想 N 个小硬球分子, 半径为 a . 假定分子密度很低. 令 $\mathbf{r}_i, i = 1, 2, 3, \dots, N$ 为各球位置.

令

$$\mathbf{r} \equiv (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N). \quad (14)$$

\mathbf{r} 空间为一 $3N$ 维空间. 这空间内有 $\frac{1}{2}N(N-1)$ 个“圆柱”:

$$|\mathbf{r}_i - \mathbf{r}_j| \leq a, \quad i, j = 1, 2, \dots, N. \quad (15)$$

\mathbf{r} 不能到柱内, 这是硬球使然. 于是, 这些分子的运动可由 $\mathbf{r}(t)$ 代表. $\mathbf{r}(t)$ 走的是直线, 但一碰到一个圆柱面就被反射. 这情形和上例相似. 图 2 中的圆饼子改想成 $3N$ 维空间的圆柱便成.

注意: $d\mathbf{r}/dt \equiv \mathbf{v}$ 是 \mathbf{r} 的速度, 且

$$v^2 = \sum_{i=1}^N \left(\frac{dr_i}{dt} \right)^2 \quad (16)$$

不变, 因为 v^2 和总能量成正比.

现在我们跟着 $\mathbf{r}(t)$ 走, 而得一条轨迹. 现在把 $\mathbf{v}(0)$ 的方向稍改, 成为 $\mathbf{v}(0) + \delta\mathbf{v}(0)$, 看看 $\delta\mathbf{v}(t)$ 如何改变, 一如上例, 只是 $\delta\mathbf{v}$ 比上例的 $\delta\theta$ 要复杂. $\delta\mathbf{v}$ 有 $3N$ 个方向. 好在每次碰撞, 只有两分子参与. 在两分子的重心坐标系内, 角度的分析和上例相似. 大致说来, 每次碰撞, $|\delta\mathbf{v}|$ 就增加约 λ/a 倍, λ 约是平均自由距离. 不过, 不同分子的碰撞, 使 $\delta\mathbf{v}$ 的方向改变. 不同分子碰撞所造成的 $\delta\mathbf{v}$ 的改变, 是互相垂直的. 例如, 分子 1, 2 相撞, 则只有 $\mathbf{v}_1, \mathbf{v}_2$ 改变, 故产生

$$\delta\mathbf{v}_{12} = (\delta\mathbf{v}_1, \delta\mathbf{v}_2, 0, 0, \dots, 0). \quad (17)$$

如果分子 7, 8 相撞, 则只有 $\delta\mathbf{v}_7, \delta\mathbf{v}_8$ 非零:

$$\delta\mathbf{v}_{78} = (0, 0, \dots, \delta\mathbf{v}_7, \delta\mathbf{v}_8, 0, 0, \dots, 0). \quad (18)$$

故

$$\delta\mathbf{v}_{12} \cdot \delta\mathbf{v}_{78} = 0. \quad (19)$$

这和第五章, 第 9 节的分析相似.

因此, 在 n 次碰撞之后, $\delta\mathbf{v}(t)$ 成为 n 个向量之和. 这 n 个向量大致是互相垂直的, 且

$$n \sim \frac{t}{\tau_0} N, \quad (20)$$

$$\tau_0 \equiv \frac{\lambda}{v}.$$

v = (分子平均速率). 每个向量的大小约为

$$\delta\mathbf{v}_i(t) \sim \left(\frac{\lambda}{a} \right)^{t/\tau_0} \delta\mathbf{v}_i(0). \quad (21)$$

注意, t/τ_0 是每分子的碰撞次数.

相干时间可以估计为 $\tau_0/\ln(\lambda/a)$, 和 (12) 相似. 既然各向量互相垂直, 则它们撑起一个体积

$$\prod_i \delta \mathbf{v}_i(t) \sim \left(\prod_i \delta \mathbf{v}_i(0) \right) \left(\frac{\lambda}{a} \right)^{3Nt/\tau_0}. \quad (22)$$

这个体积的增加是指数式的:

$$\left(\frac{\lambda}{a} \right)^{3Nt/\tau_0} = \exp [3N (t/\tau_0) \ln(\lambda/a)]. \quad (23)$$

\mathbf{v} 空间即是 $\mathbf{p} = (\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N)$ 动量空间, $\mathbf{p} = \mathbf{v}m$. 以上的分析, 大致指出 δv 如何膨胀起来. 注意: (23) 是一个 e^N 量级的大数. 以上只是个粗略的讨论[§].

以上的分析, 大致说明“不稳性”的意义. 如果轨迹有不稳性, 则很难“控制”, 或“非常敏感”. 因为起始的状态“差之毫厘”, 到后来的状态就“失之千里”. 如果有这样的不稳性, “混合性”就该没有什么问题了. 注意: 不稳性是一个力学的结果, 并未涉及概率的概念. 要用它来讨论混合性, 必先假想有“起始概率分布”.

在此类分析中, 很容易把混乱和独立的假设漫不经心的加入. 例如, 在 (13), 我们假定在 $|\delta\theta(t)|$ 大到 1 以后, 就和 $\delta\theta(0)$ “无关”了, 所以我们就可以不考虑 $\delta\theta(0)$ 了. 事实上, 以上的分析并没有说 $\delta\theta(t)$, 变大以后会有什么样的性质. (12) 只在 $\delta\theta(t)$ 很小时合用. $\delta\theta(t)$ 变大之后的情形, 我们不知道, 自然没有理由作结论. 这样的分析, 不能看作混合性或独立性的证明. (12) 或 (13) 只可以算是一个估计, 假设 $\delta\theta(t)$ 变大后就和 $t = 0$ 时的情形无关, 或独立. 不过这样的分析, 使我们对轨迹的特性, 认识得深一点.

本节的内容一部份, 来自阅读俄人克瑞洛夫的书[¶]的心得. 克氏的讨论十分深入. 但本作者的看法和他的不尽相同.

6. 一般独立性问题

总结一下: 从物理的观点, 了解统计力学, 必须了解

[§]更详细的分析, 见 Krylov (1979), p.193. 详细则已但未见严密.

[¶]Krylov(1979).

(甲) 相干时间 $\tau = O(1)$ 的存在. 也就是说, 不但相干时间必须为有限, 还须是 $O(1)$, 绝不是 $O(e^N)$ 的大数.

(乙) 相干距离 $\xi = O(1)$ 的存在. 也就是说, 物体可看成大致独立部份的组合. ξ 必须是 $O(1)$, 不能是 N 级大数.

这两点是独立性 (7) 的内容.

近年来, 力学的研究, 在计算不规则轨迹方面发展甚速. 许多简单的模型 (N 很小, 只有 3、4 以下个变量), 含有不规则 (即先后独立) 的运动. 这些研究指出, 不规则运动是牛顿定律的后果之一. 不过, 到现在为止, 所用的 N 仍太小, 不足以讨论物体各部的独立性, 即相干长度 ξ 的存在. 在 N 很小的情形下, 不规则运动的研究仍在数值模型阶段. 深入的了解尚未达到. 本作者认为研究方向应多着重于 N 值的影响, 尤其是 N 在 10 左右. 大 N 值即 $N \geq 10$ 造成的不规则运动, 可能和小 N 值的情形不大一样.

总之, 我们对基本假设的了解, 还幼稚得很.

虽然要从力学导出 (7) 看来是极其困难, 但从日常生活的经验, 这种独立性实在不稀奇. 独立性就是“不规则”和“纷乱”. 在第三篇讲了不少. 独立性既是如此普遍, 它该不只是牛顿定律的结果. 例如, 如果我们能了解随机数列, 就能对基本假设有所了解. 随机数列可由一计算程序算出, 应该较容易了解.

但事实上, 随机数列之所以不规则, 也是一个“知其然而不知其所以然”的情形.

从分子运动, 到掷骰子, 到随机数列, 以至于股票市场, 所有的不规则现象都还没有被了解. 纷乱的来源, 似乎和复杂的程度有关, 但不复杂的模型也可呈纷乱现象. 这些问题, 可能要等计算机的专家来帮助我们找答案.

7. 应用方面的困难

由以上诸篇的例子可见, 统计力学的应用, 是一个极端困难的数学问题. 如果我们不能把模型转换成一个类似理想气体的模型, 我们就无法解它. 也就是说, 我们必须把运动变量分成互相独立的部分, 再分别解每一部分. 在理想气体, 每一粒子为一个独立部份, 在晶体模型, 每一基本振动单元为一独立部分, 等等.

因此, 如果有 N 个变量, 我们必须寻求它们之间的某些关系, 把相干的部分“束成一捆”, 把独立性明朗化. 不只是本书中的各例, 无论多复杂的解,

(例如三度空间易形模型的严密解), 都必须把模型的独立性和相干特性分开, 才能解成. 在目前, 一般的模型都解不出, 只能用各式近似法去试, 或用数值解法.

本作者以为, 这解题的困难, 一部份来自我们对独立性的缺乏了解, 也就是本章讨论的, 对基本假设根源的缺乏了解. 交互作用促成独立性, 但也是相干的来源. 相干和独立搅在一起, 使题目难解. 从力学定律, 我们摸不到独立性. 而独立性使基本假设的应用简化.

目前对数值解的依赖渐深. 如第二十二章所说, 数值解法十分重要, 但从深入了解统计力学观点来看, 大概也不是长久之计.

讨论问题二十六

1. 第 5 节的分析指出, 气体的相干时间约是分子的平均自由时间. 这结论是有漏洞的. 试由第二十三章, 问题 1 指出这漏洞.

本题指出, 有些特别情况下, 相干时间很长. 这些特殊情况, 往往被忽略?

2. 试将第 5 节的分析方法, 用到晶体振动模型.

(A) 用一度空间模型, 见第八章第 3 节. 用 (8.41) 为相互作用. 先略去 α' 项. 求运动方程式.

(B) 加入 α' 项, 讨论其效果. 假设 α' 为小量.

(C) 求相干时间.

3. (A) 复习第十二章, 第 2 节. 此节指出, 要检验一群变量是否独立, 是颇不简单的事, 要一大群变量互相独立, 是一个很严格的要求.

(B) 复习第十二章, 第 5~8 节, 中央极限定理的要求, 及一般实验观测的要求, 并不是严格的独立性, 而只是“大致”独立性.

(C) 试讨论熵的定义和独立性的要求. 到底需要多少独立性才能给熵一个不含糊的定义? 这问题不易回答.

4. 有些近乎哲学的问题, 常在统计力学文献中出现. 例如时光的方向. 从力学定律, 看不出时光是倒流或是正流. 但由扩散现象, 可以察觉时光的方向.

事实上扩散仍不足以指定时光方向. 如果在 $t = 0$ 时有一些溶质分子在溶剂中聚作一堆. 过了一段时间, 它们就散开了. 但是, 我们问, 根据力学, 他们在 $t = 0$ 之前, 分布如何? 答案是: 也是扩散的. 因此, 时光的方向对称性仍在.

问题不是时光的方向, 而是何以分布随 $|t|$ 的增加而散开. 读者试就第二十四章的粘液瓶示范, 讨论 (推动把手时颜料分子散开、或聚合) 此问题.

5. 混乱和预测

混乱、独立, 都是可以测定的性质. 一系列随机数是否够乱, 可以检验. 只要由数列算出各相干值便是.

如果已经有了某随机数列中的数字, 如何从这 n 个数字计算下面 n 个数字? 这是预测问题.

照经验越乱越不容易预测.

当然, 如解出计算随机数列的程序, 则以下的数字自然可算出, 所以这是“解密码”问题.

乱的程度和难解的程度有何关系? 读者去想想.

注意, 有种看法是: “乱”的定义就是不能预测. 这种定义和此地的讨论无关.

第七篇

凝结

这一篇分四章, 讨论和凝结现象有关的题目. 凝结就是冻结, 如水的结冰, 顺磁体的变成铁磁体等等, 在第十五章讲的多相共存, 就是不同的凝结态的共存. 第十七章的易形模型是讨论凝结的主要模型之一. 本篇的重点在“序参量”的出现及其性质与后果. 多相共存可以想成物体在某些环境下, 序参量可以有不同的值, 如果物体的一部分的序参量和其他的部分的量不同, 则为多相共存. 如磁铁的一半有向东的磁矩, 另一半向西. (此地的磁矩为序参量). 如温度太高, 序参量就不见了.

第二十七章讲平均场解法. 这是一个近似算法. 它的好处是简单明了, 并且把各种凝结现象的共同特性表现得很清楚, 但嫌粗略. 它把序参量的起伏忽略了, 因而有时导致错误的结论. 第二十八章考虑多相共存的分界面, 分析它的起伏. 第二十九章讲有连续对称的情形. 这情形下, 序参量可以有连续不同的值, 如晶体中分子作规则排列, 但如把所有分子平移, 热位能不会改变. 平移使序参量改变, 且为连续性的改变. 后果之一是“柔软波”的存在, 即能量非常低的运动状态. 它可以造成序参量的大幅起伏. 第三十章讲超流性, 把它用凝结现象来解释. 这一章的重点在序参量的出现, 超流态的暂稳性, 及“缺陷”(漩涡) 所扮演的角色.

第二十七章 平均场解

平均场解法是一个极其重要的方法。在许多较复杂的问题中，这是唯一可以用来作分析的工具。无论它是多粗略，且答案常常完全离谱，它的简易明瞭使它在统计力学中占有极重要的地位。用易形模型来说明这方法是最理想不过。本章不但要说明计算方法，要特别指出此法容易被误解的地方。这法则就是把运动单元看成独立，即每一单元在一定的环境下运动，再由这些独立运动的平均值来决定环境。

本章的主要讨论模型是易形磁铁模型。我们用平均场解法来分析铁磁性，多相共存等性质。各种结论可以直接用到易形气体模型，以解释气液共存现象。范德瓦尔斯的气液共存理论也是一个平均场解。在此也有详细讨论。向量模型也可以用平均场解法。各种不同的凝结现象的共同点，在平均场解分析之下，变得很明显。

1. 平均场

在第十七章，我们讨论了一维空间易形磁铁模型的解。二维、三维空间的易形磁铁模型就不容易解了。近似解因而十分有用。平均场解是一个最简单的近似解，而且十分容易了解。令总能量为

$$H(s) = -\frac{1}{2} \sum_{i,j} s_i s_j J_{ij} - h \sum_i s_i, \quad (1)$$
$$J_{ij} = J \text{ 如果 } i, j \text{ 相邻}$$
$$= 0 \text{ 如不相邻.}$$

每个单元 s_i 的环境包括外加磁场 h 及邻近单元所达成的磁场。这总磁场是

$$h'_i = h + \sum_j J_{ij} s_j. \quad (2)$$

因为各单元不断地改变, 故 h'_i 也不断改变. 平均场法将 s_i 用其平均值代替, 得平均场

$$h' = h + Jv\langle s \rangle. \quad (3)$$

因为每个 s_i 的平均值都一样, 我们把编号 i 略去不写. (3) 的 v 是邻居的总数. 现在每个 s_i 都在一不变的磁场 h' 内. 其平均值是

$$\begin{aligned} m \equiv \langle s \rangle &= \frac{e^{h'/T} - e^{-h'/T}}{e^{h'/T} + e^{-h'/T}} \\ &= \tanh\left(\frac{h + Jvm}{T}\right), \end{aligned} \quad (4)$$

即

$$h = T \tanh^{-1} m - Jvm. \quad (5)$$

(h, m) 曲线见图 1.

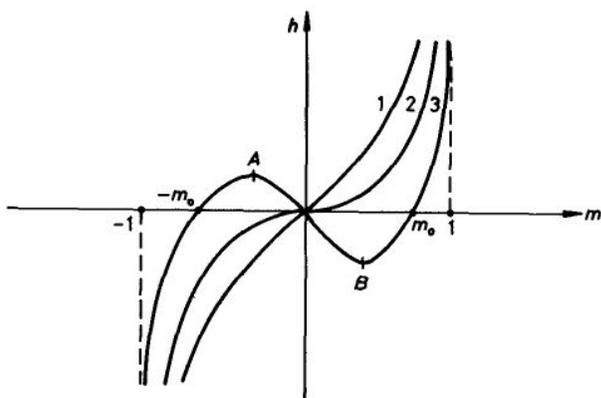


图 1

如果 T 够大, (5) 右第一项比第二项大, 则

$$\chi = \frac{\partial m}{\partial h} > 0 \quad (6)$$

显示顺磁性, 且在 $h = 0$ 时, $m = 0$ (曲线 1). 但如温度太低, 即

$$T < T_C \equiv Jv \quad (7)$$

时, 曲线就有一截违反了 (6), 且当 h 不太大时, m 可以有三值 (曲线 3). 如果 $T = T_C$ 则 χ 在 $h = 0$ 时为无限大, 曲线和横轴相切于原点 (曲线 2).

在 $T < T_C$, $h = 0$ 时, 有两个 m 非零的解 (见图 1, 2). 即

$$\begin{aligned} m &= \pm m_0(T), \\ m_0 &= \tanh(m_0 T_C / T). \end{aligned} \quad (8)$$

也有一个 $m = 0$ 的解. $m_0(T)$ 这条曲线在 T 接近于 0 时, 或近于 T_C 时为

$$\begin{aligned} m_0 &\approx 1 - 2e^{-2T_C/T}, \quad T < T_C, \\ m_0 &\approx \sqrt{3} \left(1 - \frac{T}{T_C}\right)^{1/2}, \quad T \approx T_C. \end{aligned} \quad (9)$$

现在我们要解决的问题是: 这些解有什么意义?

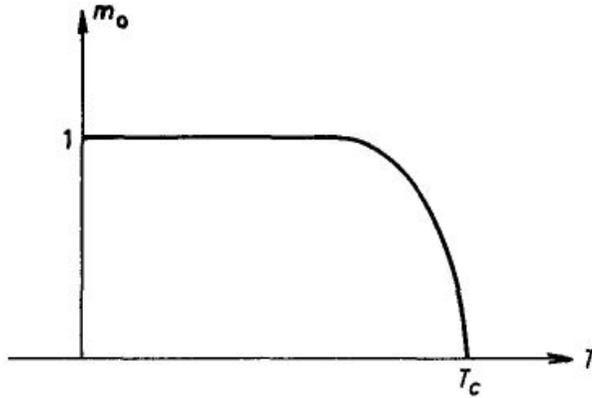


图 2

现在退一步, 看看在这平均场近似下, 热位能是什么. 从热位能的极小值, 我们可以对 $T < T_C$ 情形作一解释.

2. 总磁矩

平均场解法可以用“熵极大”法则求出. 这法则是: 用一简单的哈密顿函数 $H'(s)$ 取代 $H(s)$, 再用 H' 求熵及 $\langle H \rangle$, 再求 $E = \langle H \rangle$ 为定值下熵的极大值, 即在定温下热位能 $F = E - TS$ 的极小值, 以选定最好的 H' (见第七章末节). 现在我们用

$$H'(s) = - \sum_{i=1}^N s_i h' \quad (10)$$

来求熵, 视 h' 为待调整的参数. 这模型的熵非常容易求; 已求过多次:

$$\begin{aligned}
 S &= \ln \frac{N!}{N_+! N_-!} \\
 &= \frac{N}{2} \left[(m+1) \ln \frac{2}{1+m} + (1-m) \ln \frac{2}{1-m} \right], \\
 F_0(m) &= \langle H \rangle - TS, \\
 \langle H \rangle &= -\frac{1}{2} N v J m^2 - N m h', \\
 \langle s_i \rangle &= \tanh \frac{h'}{T} \equiv m.
 \end{aligned} \tag{11}$$

注意, m 和 h' 的关系为一单调函数, 毫无困难. 因此, 将 m 视为参数亦可. $F_0(m)$ 还不是热位能. 它的极小值才是热位能 (的近似).

现在调整 m 以求热位能的极小, 解

$$\frac{\partial F_0(m)}{\partial m} = \frac{\partial}{\partial m} [\langle H \rangle - TS] = 0, \tag{12}$$

结果就是 (5). 图 3 所示是 $F_0(m)$ 在 $T > T_C$ 和 $T < T_C$ 时的曲线. (5) 的解为 $F_0(m)$ 曲线水平点, 即图 2 曲线.

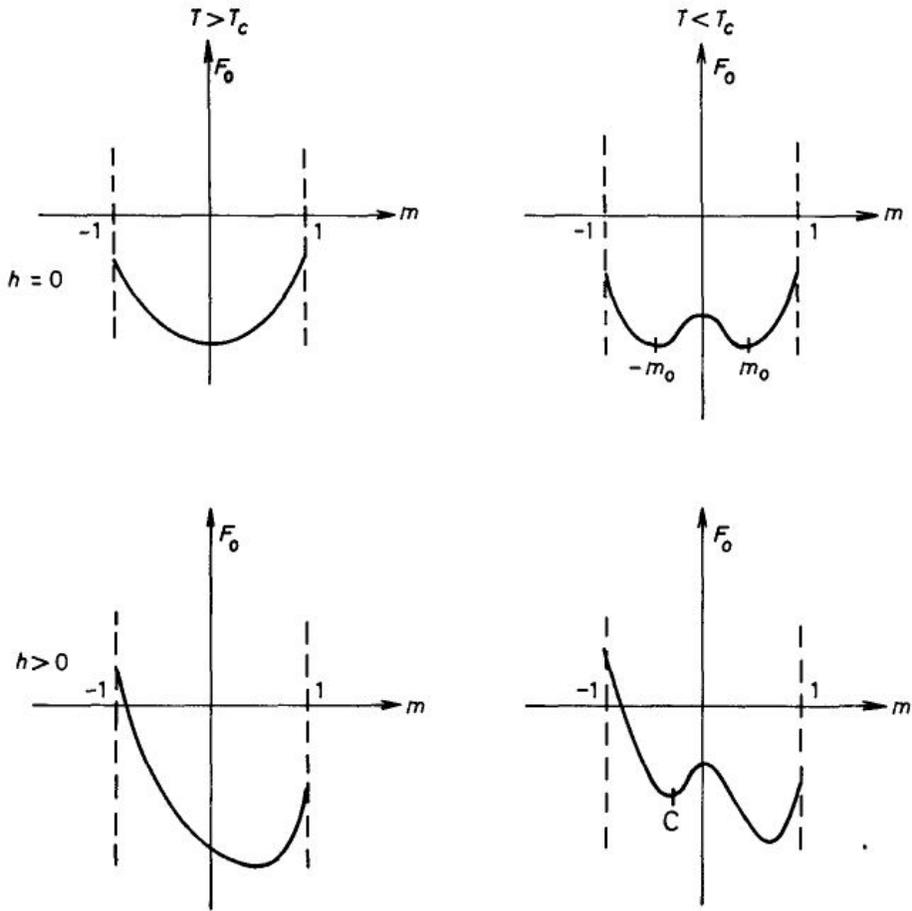


图 3

在 $T < T_C$, h 不太大时 (5) 有三解, 由图 3 可见. 不过有一解为 F_0 的极大, 此解为图 1, AB 之间者. 其他二解为 AB 之外者. 如 $h \neq 0$, 则只有一绝对极小, 即 m, h 同方向者. 如果 $h = 0$, 则有两极小, 即 $\pm m_0(T)$ (见 (8)). 图 3 中的 C 点为一极小, 但非绝对极小, m 和 h 反向. 这是一种暂稳态, 在章尾问题中会讨论.

现在, 只要 $h \neq 0$, $F_0(m)$ 就只有一个绝对极小值, 因此, $h \neq 0$ 的情形已解出, 不再有疑问. $h = 0, T < T_C$ 时, 有两个极小值. 这可以看作铁磁性的特征, 即没有磁场时, m 非零, 且方向不定. 现在我们把总磁矩 $M = Nm$ 的结果以图 4 表之.

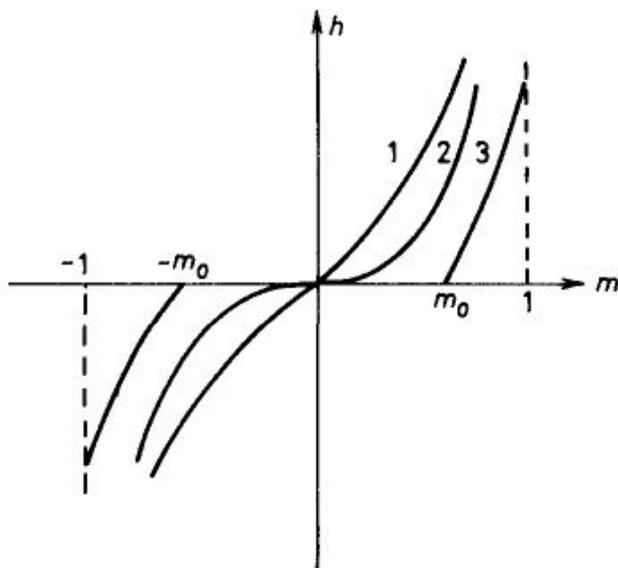


图 4

这图 4 就是图 1, 把不适合的部分去掉便得.

注意 $F_0(m)$, 即图 3 的曲线, 不是一定总磁矩下的热位能. (10) 是一个十分苛刻的限制, 不只是限制了总磁矩 Nm . 它把每一个 $\langle s_i \rangle$ 都限制成 m . 也就是说, (10) 至 (12) 的步骤是: 假定每个 $\langle s_i \rangle$ 都一样, 且每个 s_i 都独立, 则最接近真的热位能的选择是由 (12) 决定. 如果 $\langle s_i \rangle$ 不是到处一致, 则这步骤就必须修改才行.

3. 热位能和多相共存

我们方才强调图 3 的曲线 $F_0(m)$ 还不是热位能. 现在问, 热位能是什么? 答案很简单: 热位能即是

$$F(h, T) = F_0(m(h, T), h, T). \quad (13)$$

$F_0(m)$ 除了是 m 的函数外, 也是 h, T 的函数 (我们现在详细写出 h, T). $m = m(h, T)$ 即是图 4 的曲线, 由 $F_0(m)$ 的极小得来.

在 $h \neq 0$ 时 (13) 没问题. 在 $h = 0, T < T_C$ 时, $m(0, T) = \pm m_0(T)$, 虽有两解, 但 $F_0(m)$ 和 m 的符号无关. 故 (13) 亦无问题.

如果物体分成两部分, 其一有 $m = m_0$, 其二有 $m = -m_0$. 如果略去分界面的能量, 则 $F(h, T)$ 仍是一样. 分界面的能量是和其面积成正比, 而 F 是和

体积成正比. 因此分界面的能量只是一很小的修正项. 令第一部分占总体的百分比为 α , 则

$$\begin{aligned} m(0, T) &= \alpha m_0 - (1 - \alpha)m_0 \\ &= (2\alpha - 1)m_0, \quad 0 < \alpha < 1, \end{aligned} \quad (14)$$

即

$$-m_0 < m(0, T) < m_0.$$

当然, 把物体分成三部分, 或更多部分亦可. 只要分界面不是太大, 则热位能大致不变. 这可以看作“多相共存”的现象, 即 $m = m_0$ 和 $m = -m_0$ 的区域可以同时存在于平衡态中. 这些区域很像磁铁中的“磁畴”, 每块畴里的磁矩方向一定, 但各畴的磁矩不一. 但真的磁铁中有长距离作用, 比这易形模型复杂太多. 各区域间的分界面或移动, 或被杂质困住. 这是另一些问题. 我们现在的结论是, 在 $T < T_C$, $h = 0$, $m(0, T)$ 的值可在 $\pm m_0$ 之间任何值, 而 F 不变. 令

$$\begin{aligned} F'_0(m, T) &\equiv F_0(m, h, T) + hNm \\ &= -\frac{1}{2}NvJm^2 - TS(m), \end{aligned} \quad (15)$$

F'_0 只是 m 和 T 的函数. $-hM$ 一项没有包括. 令热位能 (13) 加上 hM 为

$$F + hM = F'(m(h, T), T). \quad (16)$$

这个热位能, 是看作总磁矩和温度的函数. 注意 F 和 F' 的微分为

$$\begin{aligned} dF &= -SdT - Mdh, \\ dF' &= -SdT + h dM. \end{aligned} \quad (17)$$

F' 是内能减 TS , 而 F 是内能减去 hM 再减 TS . 即 F 包括了外力 h 产生的位能 (见第十六章讨论).

根据以上的分析, 如果 $T < T_C$, 则

$$\begin{aligned} F'(m, T) &= F'_0(m, T), \quad \text{如 } |m| > m_0, \\ &= F'_0(m_0, T), \quad \text{如 } |m| < m_0. \end{aligned} \quad (18)$$

此地的 m 是 M/N 的平衡值, 即 $m(h, T)$.

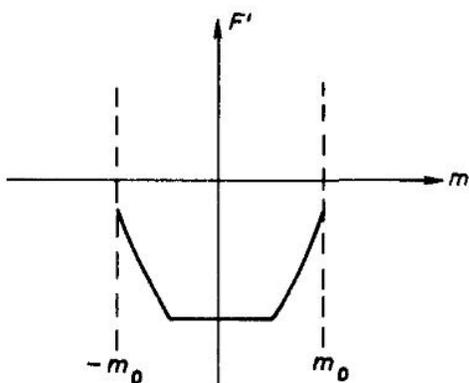


图 5

如果 $T > T_C$, 则 $F'(m, T) = F'_0(m, T)$.

图 5 就是图 3 右上曲线, 只是把底下作一割线. 以割线为 F' . 割线上的凸出部分不要了. 平衡时

$$\frac{\partial F'}{\partial M} = h. \quad (19)$$

如果 $h = 0$, 则 $F' = F$ 须为极小. 因为多相可以共存而 F 值不改, 所以在 m_0 、 $-m_0$ 间 $F' = F$ 必须是水平, 不会凸起来.

以上似乎啰里啰苏, 讲了半天, 结果只是把图 3 右上曲线的凸起削去而已. 这结果虽简单, 但 F , F' 和 F_0 之间的关系并不简单. 图 5 的意义和图 3 大不相同. 读者务必留意.

4. 易形格子气体

这模型是由哈密顿量 H 定义:

$$H - \mu N = -\frac{1}{2}\epsilon \sum_{i,j} n_i n_j - \mu \sum_i n_i, \quad (20)$$

$$N = \sum_i n_i.$$

此地的 n_i 是在格子 i 的粒子数, $n_i = 1$ 或 0 . (20) 的右第一项的和是限于相邻的 i, j . 因为每个 n_i 只有二值, 这模型是一易形模型. 令

$$\sigma_i = 2n_i - 1, \quad (21)$$

则 (20) 可化作易形磁铁模型:

$$\begin{aligned} H - \mu N &= H' - \left(\frac{1}{8}\epsilon v + \frac{\mu}{2} \right) V, \\ H' &= -\frac{1}{2}J \sum_{i,j} \sigma_i \sigma_j - h \sum_i \sigma_i, \\ J &= \frac{\epsilon}{4}, \\ h &= \frac{\mu}{2} + \frac{\epsilon v}{4}. \end{aligned} \quad (22)$$

注意, 现在用 V 指总格子数, 总单元数. 用 N 来指定气体粒子数. 以前的单元数 N 是现在的 V . 有了这模型和磁铁模型的关系, 我们可以把上几节的结果翻译成格子气体的性质. 热位能 Ω 的定义是

$$e^{-\Omega/T} = \sum_s e^{-(H-\mu N)/T}, \quad (23)$$

而磁铁模型的热位能 F 的定义是

$$e^{-F/T} = \sum_s e^{-H/T}. \quad (24)$$

因此, 由 (22) 可知

$$\Omega = F - \left(\frac{1}{2}\mu + \frac{1}{8}\epsilon v \right) V. \quad (25)$$

因为 $\Omega = -pV$, p 为压力, 所以 (注意磁铁模型的 N 是此地的 V).

$$\begin{aligned} p &= -\frac{F}{V} + \left(\frac{1}{2}\mu + \frac{1}{8}\epsilon v \right) \\ &= -\frac{F'}{V} + 2nh - \frac{1}{2}T_c, \end{aligned} \quad (26)$$

$$T_c = Jv = \frac{1}{4}\epsilon v,$$

$$n = \langle n_i \rangle = \frac{1}{2}(1 + m). \quad (27)$$

F' 是由 (16) 定义, 见 (16), (17), (18), (19), 及图 5. 我们要知道的是压力、密度 n 及温度 T 之间的关系. 因为 F' 已解出, (26) 即是答案, 只是不大明显.

先看 $T > T_c$. h 可以用 (5) 代, F' 由 (11) 代. 简化之, 得

$$p = T \ln \left(\frac{1}{1-n} \right) - 2n^2 T_c. \quad (28)$$

在低密度时, 即 $n \ll 1$ 时, $p \approx nT$, 即理想气体公式. 高密度时, 即 $n \rightarrow 1$, p 趋向 ∞ , 这是因为每格只能放一个粒子之故. (28) 可以改写成

$$p = (T - T_c) \ln \left(\frac{1}{1-n} \right) + T_c \left[\ln \left(\frac{1}{1-n} \right) - 2n^2 \right], \quad (29)$$

故

$$\frac{dp}{dn} = \frac{(T - T_c)}{(1-n)} + T_c \frac{(1-2n)^2}{(1-n)^2}. \quad (30)$$

可见在 $T = T_c$, $n = \frac{1}{2}$ 时, p 的一次和二次微分都是 0.

这即是“临界点”, 相当于 $h = 0$, $T = T_c$, $m = 0$. 临界压力是

$$p_c = T_c \left(\ln 2 - \frac{1}{2} \right) = 0.193T_c. \quad (31)$$

(p_c, T_c) 是平衡蒸汽压曲线的终点 (见图 6). 这曲线即是 $h = 0$, $T < T_c$ 画在 (p, T) 平面上. 将 (26) 的 h 设为 0, 将 (18) 的 $F'_0(m_0, T)$ 代 F' , 即得此曲线

$$\begin{aligned} p(T) &= -\frac{F'_0(m_0(T), T)}{V} - \frac{1}{2}T_c \\ &= -\frac{1}{2}T_c (1 + m_0^2(T)) + \frac{1}{2}T \ln \frac{4}{1 - m_0^2(T)}. \end{aligned} \quad (32)$$

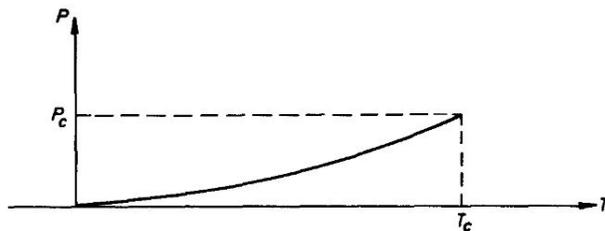


图 6

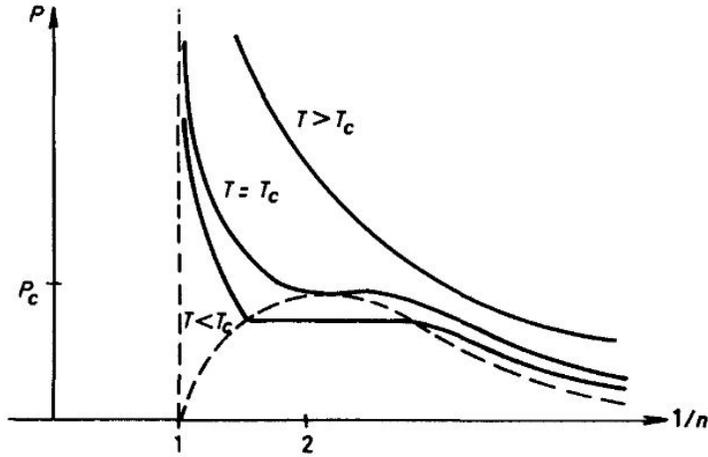


图 7

在 $T < T_c$, $h \neq 0$ 时, (28) 仍可用. 如果把 p 画成 $1/n = V/N$ 的函数, 则成图 7 的曲线. 水平的部分是气液共存的压力, 即饱和蒸汽压, 或平衡蒸汽压.

潜热 q 是每一分子由液态转入气态所需的热. 在这模型, $S(m_0)/V$ 和 $S(-m_0)/V$ 是一样的, 但密度 $n = N/V$ 由 $\frac{1}{2}(1 + m_0)$ 改成 $\frac{1}{2}(1 - m_0)$. 因此,

$$\begin{aligned} q &= T \left[\left(\frac{S}{N} \right)_{\text{气}} - \left(\frac{S}{N} \right)_{\text{液}} \right] \\ &= TS(m_0) \left(\frac{2}{1 - m_0} - \frac{2}{1 + m_0} \right) \frac{1}{V} \\ &= T \frac{2m_0}{1 - m_0^2} \left[(1 + m_0) \ln \left(\frac{2}{1 + m_0} \right) + (1 - m_0) \ln \left(\frac{2}{1 - m_0} \right) \right]. \end{aligned} \quad (33)$$

这结果也可以由 (32) 的微分得来, 即

$$\begin{aligned} \frac{dp}{dT} &= \frac{q}{T\Delta v} \\ \Delta v &= \left(\frac{1}{n} \right)_{\text{液}} - \left(\frac{1}{n} \right)_{\text{气}}. \end{aligned} \quad (34)$$

此式见第十五章.

5. 范德瓦尔斯公式

在上一节, 已经把平均场近似用于气液相变的问题. 范德瓦尔斯公式的内容也只是如此, 外表稍有不同而已. 本节的目的, 是多一次练习, 并介绍一

下“气液态万用公式”。

气体分子间的作用，有短程的斥力和稍远程的吸引力。现在我们把这气体用一理想气体模型取代，再调整模型中的参数使熵为极大，同时把总能量限制在定值。

令 S 为理想气体的熵。在定温定压下，能量是

$$H = \frac{3}{2}NT + pV - b\frac{N^2}{V}. \quad (35)$$

此式包括压力产生的位能 pV 及由分子间引力引起的作用能 $-bN^2/V$ 。现在求 S 的极大值而把 H 限为一定，即求

$$S(V - Nv_0) - \frac{H}{T} \equiv -\frac{1}{T}G_0(V) \quad (36)$$

的极大，或

$$G_0(V) = \frac{3}{2}NT + pV - \frac{bN^2}{V} - T \left[N \ln \left(\frac{V - Nv_0}{N} \right) + \frac{3}{2}N \ln T \right] \quad (37)$$

的极小值。现在 V 是一可调整的参数，一如第 2 节的 m 。在 (36)，熵所用的体积是 $V - Nv_0$ ，即这理想气体模型是有较小的容积。原因是分子间的斥力。 v_0 约是分子的体积。斥力使分子活动空间缩小。(36) 的极大值，即 (37) 的极小，可由

$$\frac{\partial G_0}{\partial V} = 0 \quad (38)$$

求得，也就是

$$p = \frac{NT}{V - Nv_0} - b\frac{N^2}{V}. \quad (39)$$

此式即范德瓦尔斯公式 (见图 8)。在高温时没问题。在温度太小时， $\partial p/\partial V$ 就可能是正的，即加压反而膨胀，不太合理。

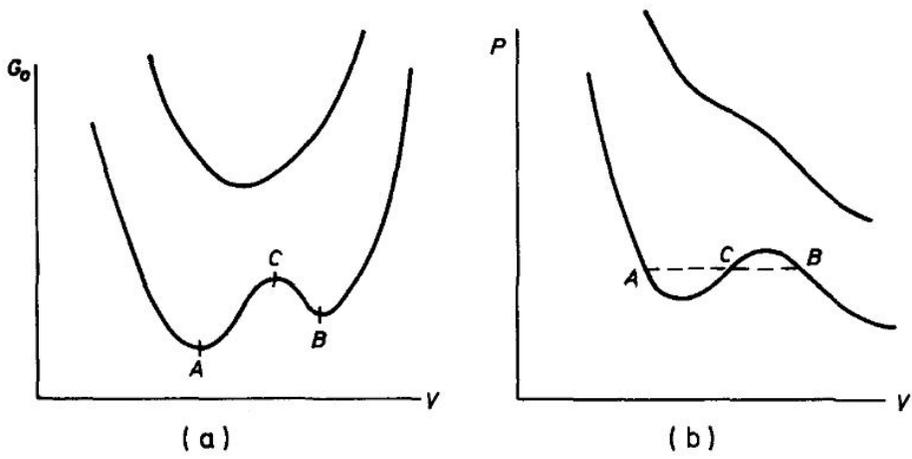


图 8

注意, (39) 是由在一定 p 下, G 的极小值求得. 图 8(a) 低温曲线有三处 $\partial G/\partial V = 0$. 这三个点在图 8 (b) 中也标出. 显然只有 A 处之值的 G_0 才是热位能, B, C 都得丢掉

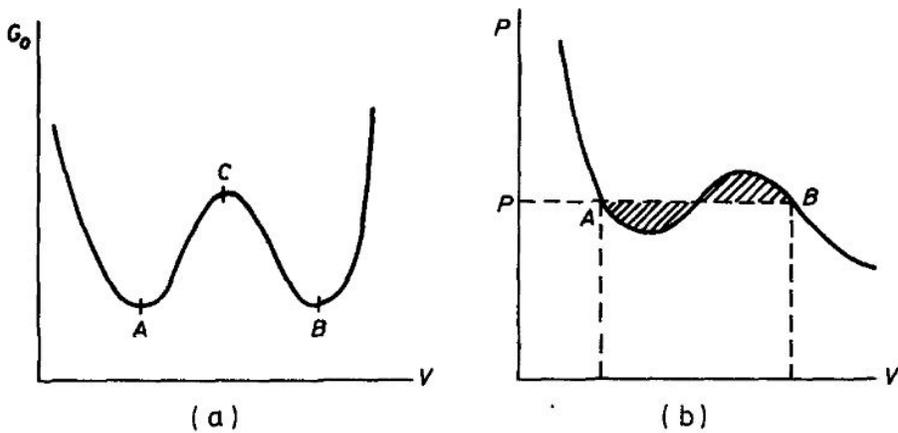


图 9

如果 G_0 在 A, B 两处同高, 则有二解. 在这情形下

$$\begin{aligned} 0 &= G_0(V_B) - G_0(V_A) \\ &= \int_{V_A}^{V_B} dV \frac{\partial G_0}{\partial V} = \int_{V_A}^{V_B} dV [p - p(V)] \\ &= (V_B - V_A)p - \int_{V_A}^{V_B} dV p(V). \end{aligned} \quad (40)$$

式中 $p(V)$ 是 (39) 的右手边, 即图 9(b) 的曲线. (40) 最后一项是曲线在 AB 下的面积. 此面积由 (40) 可见等于直线 AB 下的面积. 因此, 在二态可以共存时, 图 9(b) 的两块画满斜线部分有相同的面积. 所以, 若要求共存时的压力和体积, 只要在 $p = p(V)$ 曲线上画条直线, 令直线上曲线下的面积等于曲线上直线下方的面积, 用不着去计算热位能了. 此所谓“麦克斯韦法则”, 把一条曲线的不合理部分去掉. 每一个温度有一条曲线, 结果和图 7 大同小异, 虚曲线是所谓“共存曲线”. 共存曲线的最上一点为“临界点”. 此点以上就不会有二相共存了.

临界点是 $p = p(V)$ 曲线的一次, 二次微分都是零

$$\frac{\partial p}{\partial V} = 0, \quad \frac{\partial^2 p}{\partial V^2} = 0. \quad (41)$$

从范德瓦尔斯公式 (39) 至 (41), 可解出 V_c, T_c , 即临界体积及温度. 得

$$\begin{aligned} V_c &= 3v_0N, \\ T_c &= \frac{8b}{27v_0}, \\ p_c &= \frac{b}{27v_0^2}. \end{aligned} \quad (42)$$

现在用这些值为体积, 温度及压力的单位, 令

$$\begin{aligned} v' &\equiv \frac{V}{V_c}, \\ T' &\equiv \frac{T}{T_c}, \\ p' &\equiv \frac{p}{p_c}. \end{aligned} \quad (43)$$

则范德瓦尔斯公式成了

$$p' = \frac{8T'}{3v' - 1} - \frac{3}{v'^2}. \quad (44)$$

这个公式没有任何参数了. 因此, 如果范德瓦尔斯公式是对的, (44) 应该合用于所有气体. 只要测出临界点, 其他性质全可由 (44) 求得. 照实验结果, (44) 可以说是惊人的成功. 许多气体都适合 (44), 准确到 10% 以内. (44) 可称为“气液万用公式”*. 成功的原因, 是各种气体分子作用都差不多, 都是第十四章图 1 的样子.

注意: 平均场解并不能准确地解出 T_c . 刚才说的“准确到 10%”是指 (44). (43) 的 T_c, p_c, V_c 都是用的实验数值, 不是平均场解. 由第十七章的图 3 可见一斑.

6. 凝结现象的共同性质

以上我们用平均场方法讨论了液态气态, 和铁磁性的模型. 我们发现这些模型在表面上不同, 但内容一样. 因此, 它们讨论的不同现象, 应该有许多共同点. 现在总结整理如下.

(甲) 多相共存:

在一定的环境下, 可以有不同的平衡态. 在磁铁模型中, 多相指不同的磁矩方向. 在气液模型中, 多相是指不同的密度.

(乙) 临界点

在临界温度以上, 多相共存不再出现, 即只有一个平衡态. 在磁铁模型, 磁矩消失. 在气液模型, 无气液之分.

(丙) 序参量

因此, “序参量”的观念可以很自然地建立. 此量在临界温度以上为零. 在临界温度以下非零, 且可以有不同的值. 不同值代表不同的平衡态. 在磁铁中, 序参量可定义为平均磁矩. 在气液模型中, 可定义为密度减去临界点的密度. 在磁铁, 序参量是一个向量, 而在气液则是一标量. 注意: 以上各节的讨论, 事实上只限于标量, 磁铁模型的磁矩只限于两个方向. 向量是指可有连续性方向改变的量. 这是第二十九章的题材.

序参量是一个坐标. 和其他坐标相似. 每个坐标都可有一个“力”来控制它 (见第二章), 控制序参量的力可称“凝结力”. 磁场和压力 (或化学势) 是控制磁矩和密度的力, 是以上二例子的凝结力.

除了铁磁性和气液共存, 还有许多其他的凝结现象, 都有以上 (甲)、(乙)、

*英文文献称之为 Law of corresponding states.

(丙) 的特点. 表 1 列出一些常见的例子. 反铁磁性的固体中, 自旋方向作周期性的排列. 例如在立方晶格里, 任一个自旋的方向常和它的六邻自旋相反. He^4 超流性和超导电性的凝结量就很不明显了. 一个粒子位子可以是空的, 也可以是满的. 但量子力学还容许空态和满态的组和状态. 因此, 除了粒子数之外, 还有其他运动变量存在, 即“粒子幅度”. 读者若用自旋 $1/2$ 来想, 则自旋向上为“满”向下为空, 代表粒子数变数. 但除了上下之外, 自旋还可以朝左右南北各方向. 这些方向, 相当于“粒子幅度”. 这一点在第三十章会详谈. 超导体的情形也类似, 一对电子位子可空可满, 但也可以在不空又不满的组和态.

以上的共同点, 可以用简单的平均场解作一总结. 图 3 的各曲线在临界温度附近, 可以用以下简单式子代表. 即在 m 很小时

$$F_0(m) \approx \text{常数} + N(am^2 + bm^4 - hm) \quad (45)$$

$$a = a'(T - T_c).$$

现象	序参量	例如	临界温度 K
气液	密度	H_2O	647.05
铁磁性	磁矩	Fe	1044
反铁磁性	周期磁矩	FeF_2	78.26
超流性 He^4	粒子幅度	He^4	2.17
超导电性	配双电子幅度	Pb	7.19
混合液分离	两液体密度差	$\text{CCl}_4\text{-C}_7\text{F}_{14}$	301.8
合金分子排列	周期排列	CuZn	739

a', b 在 T_c 附近大致不变. m 是序参量. (45) 合用于任一凝结现象, 只是 a', b 的值各有不同. 设 $h = 0$, 由 $F_0(m)$ 的极小条件, 可知 m 在 $T > T_c$ 时为 0, 在 $T < T_c$ 时和 $(T_c - T)^{1/2}$ 成正比. $\partial m / \partial h$ 和 $|T - T_c|^{-1}$ 成正比. 热容率在 $T = T_c$ 处不连续. 这些特性在以上诸节大致都提到了, 读者可由 (45) 再推演一次. (45) 式即是所谓“兰道理论”. 它的内容和平均场解法一样, 但它的用意是把所有的凝结现象放在一起讨论. 这是极其重要的一步.

平均场解并不足以解释所有的凝结现象共同点. 更仔细的实验和理论, 发现在 T 和 T_c 很接近时, 平均场解的结论就不正确了, 例如

$$m \propto (T_c - T)^\beta,$$

$$\frac{\partial m}{\partial h} \propto |T - T_c|^{-\gamma}, \quad (46)$$

指数 β 不是 $1/2$, γ 亦非 1 . 而且 β, γ 的值和序参量的对称性有关. 易形模型有 $+m, -m$ 对称, 但向量模型 (即 \mathbf{m} 为一向量) 有连续性的对称, 各指数就不同. 同一对称性, 但如空间维数不同, 指数亦异. 如加了杂质, 可能又不一样. 临界点附近的性质, 统称“临界现象”. 虽然是很有趣的题目, 但略嫌太专门, 故不多谈了.

讨论问题二十七

1. 由图 3 可见, 在 $T < T_c$, h 不太大时, $F_0(m)$ 有一个绝对极小, 还有一局部极小.

绝对极小为平衡点.

如果把磁场 h 忽然转到相反方向, 则情形如何?

复习第十五章“水珠的消长”一节, 及第十七章问题 5. 试把图 3 的局部极小解释成一暂稳态.

2. 如果容许各点的平均场不一样, 则 (5) 的推广是

$$h_i = T \tanh^{-1} m_i - Jvm_i - J \sum_j (m_j - m_i), \quad (47)$$

$$m_i = \langle s_i \rangle.$$

式中的求和 j 限于 i 的近邻.

(A) 试导出 (47), 用第 1 节, 及第 2 节的方法.

(B) 如果 m_i 是 \mathbf{r}_i 的平缓函数试证

$$\sum_j (m_j - m_i) \approx a^2 \nabla^2 m(\mathbf{r}), \quad (48)$$

a 是近邻之间的距离. m 看成位置的连续函数. 因此,

$$h \approx T \tanh^{-1} m - Jvm - Ja^2 \nabla^2 m \quad (49)$$

外加磁场 h 也可以是 \mathbf{r} 的函数.

(C) 分界面

令 $h = 0, T < T_c$. 假定

$$\begin{aligned} m &= 0, & x &= 0, \\ m &\rightarrow m_0, & x &\gg a, \\ m &\rightarrow -m_0, & x &\ll -a, \end{aligned} \quad (50)$$

并假定 m 在 y, z 方向不改, $\pm m_0$ 为 $F_0(m)$ 的偶极小点 (见第 2, 3 节). 试求 $m(x)$, 并作图.

(D) 界面张力

试证由分界面引起的单位面积热位能是

$$\begin{aligned} \sigma &= \int_{-\infty}^{\infty} dx \left[\frac{1}{2a} J \left(\frac{dm}{dx} \right)^2 + f_0(m(x)) - f_0(m_0) \right], \\ f_0(m) &\equiv \frac{F_0(m)}{V}. \end{aligned} \quad (51)$$

σ 即界面张力. 将 (C) 结果代入即可求 σ .

(E) 界面位移

在 (50), 加了 $m(0) = 0$ 的要求. 如果把这要求改成 $m(b) = 0$, b 是任一常数, 则以上答案有何改变?

试讨论 b 的意义. 就此讨论分界面的起伏问题 (详细分析见下一章).

(F) 界面厚度

试证, 如 $x \gg a$,

$$\begin{aligned} m(x) - m_0 &\propto e^{-x/\xi}, \\ \xi &\equiv a \left[\frac{(T_c - T)}{J} \right]^{-1/2}. \end{aligned} \quad (52)$$

这结果只合用于 $\xi \gg a$ 的情形, 因为 (48), (49) 的假设是 m 在距离 a 之内改变得很少.

3. 试用平均场法解向量模型

$$H = -J \sum_{i,j} \mathbf{s}_i \cdot \mathbf{s}_j - \mathbf{h} \cdot \sum_i \mathbf{s}_i \quad (53)$$

\mathbf{s}_i 是长为 1 的向量, i, j 为立方晶格上的近邻. 求磁化率, 热容率.

4. 试用平均场法解以上向量模型, 只是 \mathbf{s}_i 为量子自旋 $1/2$ 算符.
5. 将 (44) 式仔细作图, 画出各 (p', v') 曲线, 共存曲线.
6. 从第 5 节的分析, 导出每分子蒸发熵

$$\Delta s = \frac{q}{T} = \ln \left(\frac{3v'_1 - 1}{3v'_2 - 1} \right), \quad (54)$$

v'_1, v'_2 各为气态、液态中的每分子体积除以每分子临界体积 v_c .

因此, 一旦 v'_1 测得, v'_2 和 Δs 即可算出. 注意, Δs 是一个对数, 因此对 v'_1, v'_2 的改变不很敏感. 既然万用公式 (44) 的应用范围很广, (54) 式也属“兼用”. 第十五章问题 2 也可以用此式大致解释.

7. 设有一类的物质, 它们有共同的作用位能形式, $\varphi(r) = \epsilon\phi(r/\sigma)$, ϕ 的形式固定, 但 ϵ, σ 随物质不同而异.

(A) 从定温法则证明若热力学变量取适当的单位, 则此类物质均有相同的状态方程式.

(B) 证明若此类物质有临界点, 则有相同的临界比值

$$\frac{p_c v_c}{T_c},$$

T_c, p_c, v_c 是在临界点的温度、压力和每分子体积. 这一点指出“万用公式”之所以为万用, 大致是因为各气体分子作用能有大致相同的形式.

8. 感应和相干

(A) 复习第十三章.

(B) 假定 $T > T_c$, m 很小, $h \propto \delta(\mathbf{r})$.

用 (49) 解 $m(\mathbf{r})$, 试证

$$\begin{aligned} G(\mathbf{r}) &\propto \frac{1}{r} e^{-r/\xi}, \\ \xi &\propto (T - T_c)^{-1/2}. \end{aligned} \quad (55)$$

$G(\mathbf{r})$ 是感应函数, 相干函数 $C(\mathbf{r})$ 可由 $G = C/T$ 求得. ξ 是相干长度.

9. 求 $T < T_c$ 时的感应和相干函数. 仍从 (49) 开始.

以上两题指出, 从平均场解法, 我们可以把起伏现象略作分析.

在临界点附近, 起伏大, 相干长度很长, 这简单的分析不大靠得住.

10. 在临界点附近, 散射率特别大, 因为散射率是和相干函数成正比. (见第十一章第 4 节). 试分析此一现象.

11. 反铁磁性, 易形模型

令总能量为

$$H = J \sum_{i,j} s_i s_j - h \sum_i s_i. \quad (56)$$

这和铁磁模型差不多, J 改成了 $-J$. ij 仍是立方晶格上的近邻. $J > 0$.

(A) 假定 $h = 0$, 解一维空间模型.

(B) 试证: 只要 $h = 0$, (56) 的热位能和铁磁模型一样, 即 J 改成 $-J$ 无所谓.

(C) 用平均场法解 (56), $h \neq 0$ 的情形. 求热容率及磁化率.

(D) 序参量在此模型该如何定义? 凝结力是什么? (不是 h).

第二十八章 分界面的起伏

上节讲的平均场解, 使我们对凝结现象有一个大致的了解, 告诉我们许多此类现象的共同特性, 例如相类似的多相共存曲线, 临界点等等. 不过我们还没有考虑起伏. 起伏对凝结有很大的影响, 起伏太大, 就不能凝结. 起伏的特性, 也和凝结有密切的关系.

起伏的分析, 往往可以分成三种情形: (1) 高温, (2) 低温, (3) 临界. (1) 是指比临界温度高不少的情形, 物体性质和凝结没有什么关系, 相干长度很短. 这种情形比较容易分析, 相互作用可以看成是一个不太重要的修正. (2) 是指比临界温度低了不少的温度. 依照平均场解, 凝结已成定局. 但起伏仍可以有十分重要的效果. (3) 是指在临界温度附近的情形. 这情形比较复杂. 在凝结的边缘, 一会儿凝一会儿不凝, 这一堆单元指向北, 另一堆又向南, 起伏幅度大, 相干长度大, 很不容易分析. 这是一个近年来颇受人瞩目的题目, 读者有兴趣可阅参考书*. 因为它稍嫌太专门一些, 我们不在此介绍.

起伏和空间的维数及序参量的对称性有十分密切的关系. 这一点, 平均场解没有能指出. 在第十一章, 我们曾计算晶体中原子的振幅, 发现在二维空间或一维空间, 振幅很大, 而在三维以上的空间就不大. 此类问题, 将在下一章作进一步的讨论.

在这一章, 我们讨论在低温时, 多相共存的分界问题. 这也是一个平均场解所忽略了的问题. 我们先用一简单的表面张力模型来看气、液分界的起伏, 再用易形模型及“电荷气体”模型来分析晶体表面. 最后一节的目的, 是指出看来不相干的模型竟可以有相同的内容.

*见 Ma (1976), Stanley (1971).

1. 液面模型

在第十七章, 我们解了一维空间的易形模型. 结论是没有铁磁性, 因此, 平均场解的结论不合用于这一维空间的模型. 原因很简单: 各分界点把物体分成“正区”和“负区”. 由于分界点可以移动而能量不改, 各区大小不断地作不规则的改变, 使平均磁矩消失. 这由分界点引起的起伏, 在平均场解中没有被考虑.

二、三维空间的易形模型就没有这种情形发生. “分界点”成了分界线, 或分界面. 一个在正区内的负区若要增大, 必须加大分界. 在低温时, 分界不易增大. 这是第十七章“派若氏证明”的内容.

但在讨论正态, 负态共存时, 仍要考虑到正、负区域的分界. 现在我们来谈二维空间分界线的起伏情形. 暂且不用易形模型, 而用一个更简化的模型. 图 1 示一分界线 $z(x)$. 线下为正区, 线上为负区. 我们把分界看成一连续曲线, 略去格子的分立性质.

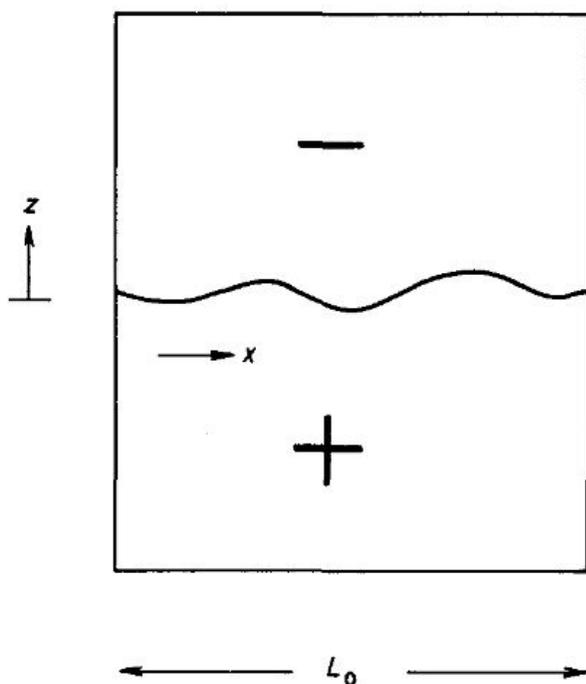


图 1

假定分界线的能量和其长度成正比:

$$\begin{aligned}
 H &= \alpha L \\
 L &= \int_0^{L_0} dx \left[1 + \left(\frac{dz}{dx} \right)^2 \right]^{1/2} \\
 &\approx L_0 + \frac{1}{2} \int_0^{L_0} dx \left(\frac{dz}{dx} \right)^2.
 \end{aligned} \tag{1}$$

α 是“表面张力”. 如果这分界是气, 液的分界, 则这模型该是相当正确 (如果没有重力的话). 以上假设坡度很小, 即 $|dz/dx| \ll 1$, 令

$$\begin{aligned}
 z(x) &= \frac{1}{\sqrt{L_0}} \sum_k \eta_k e^{ikx}, \\
 \eta_k &\equiv \frac{1}{\sqrt{L_0}} \int_0^{L_0} dx z(x) e^{-ikx}.
 \end{aligned} \tag{2}$$

代入 (1) 得

$$\begin{aligned}
 L &\approx L_0 + \frac{1}{2} \sum_k k^2 |\eta_k|^2, \\
 H &= \alpha L = \alpha L_0 + \frac{\alpha}{2} \sum_k k^2 |\eta_k|^2.
 \end{aligned} \tag{3}$$

由能量均分定律得

$$\langle \alpha k^2 |\eta_k|^2 \rangle = T. \tag{4}$$

令 Δz 为 x, x' 两点的高度差, 并求 $\langle (\Delta z)^2 \rangle$. 在 x, x' 相距很远时, Δz 的大小可以看成界线的“厚度”

$$\begin{aligned}
 \Delta z &\equiv z(x) - z(x') \\
 \langle (\Delta z)^2 \rangle &= 2 \langle z^2 \rangle - 2 \langle z(x)z(x') \rangle \\
 &= \frac{2}{L_0} \sum_k \frac{T}{\alpha k^2} [1 - \cos k(x - x')] \\
 &= \frac{2T}{\alpha} \int_{-\infty}^{\infty} \frac{dk}{2\pi k^2} [1 - \cos k(x - x')] \\
 &= \frac{2T}{\alpha} |x - x'|.
 \end{aligned} \tag{5}$$

如果 $|x - x'| \sim L_0$, 则 $\langle (\Delta z)^2 \rangle \sim (T/\alpha)L_0$. 这结果的由来是 (3) 式中的 k^2 因子. $k = 0$ 的移动是指所有的 $z(x)$ 都一样, 即均匀移动. 这种移动不费能量,

即把整个分界线往上或往下移不会改变能量. k 很小的振动, 即长波振动, 能量很小. 这些长波运动使分界面高度的起伏变得很大. 这分界的“厚度”因此和 \sqrt{L} 成正比. (即 $\langle(\Delta z)^2\rangle$ 的平方根). 这分界线因此是“粗糙”而非“光滑”.

现在把这简单的模型推广到三维空间中的分界面. 令界面的能量和其面积成正比.

$$\begin{aligned}
 H &= \alpha A \\
 A &= \int d^2\mathbf{r} [1 + (\nabla z)^2]^{1/2} \\
 &\approx L_0^2 + \frac{1}{2} \int d^2\mathbf{r} (\nabla z)^2.
 \end{aligned} \tag{7}$$

这是 (1) 的推广. α 是表面张力. 高度的起伏可再由 $\langle(\Delta z)^2\rangle$ 来表示. (5) 的推广是

$$\begin{aligned}
 \langle(\Delta z)^2\rangle &= \left(\frac{2T}{\alpha}\right) \int \frac{d^2\mathbf{k}}{(2\pi)^2 k^2} [1 - \cos \mathbf{k} \cdot (\mathbf{r} - \mathbf{r}')] \\
 &= \frac{T}{\pi\alpha} \int_0^{k_D} \frac{dk}{k} [1 - J_0(k|\mathbf{r} - \mathbf{r}'|)] \\
 &\approx \frac{T}{\pi\alpha} \ln \frac{|\mathbf{r} - \mathbf{r}'|}{a}, \\
 a &\approx \frac{1}{k_D}.
 \end{aligned} \tag{8}$$

我们把积分限在 k_D 之内. a 约是原子间的距离. (8) 只在 $|\mathbf{r} - \mathbf{r}'|$ 比 a 大很多时合用, J_0 是贝塞耳函数.

以上的结果 (8) 指出, 在 $|\mathbf{r} - \mathbf{r}'|$ 增大时, $\langle(\Delta z)^2\rangle$ 越来越大. 因此, 分界面是粗糙的, 不管温度多低, 界面的厚度是

$$\langle(\Delta z)^2\rangle^{1/2} \sim \left[\frac{T}{\alpha} \ln \frac{L_0}{a}\right]^{1/2}. \tag{9}$$

L_0^2 是界面在 (x, y) 面的面积. 注意, 即使 L_0 很大, (例如 1 公里), (9) 也不会大到那里去. 把 (9) 看成无限大是毫无意义的.

这模型把物体看成连续介质, 把分界看成弹性的线或面. 对一个气液分界来说这模型该是很正确, 只是未把重力考虑进去, 重力虽然弱, 但是液面之所以维持水平, 全靠重力, 现在估计一下重力的影响.

液面涨高 z , 则产生压力 $\rho g z$, ρ 是液体质量密度减去气体的. 故产生能量

(单位面积) $\frac{1}{2}\rho g z^2$, 因此, (7) 该加上一项, 而成为

$$H = \alpha A = \alpha L_0^2 + \frac{1}{2}\alpha \int d^2\mathbf{r} \left[(\nabla z)^2 + \frac{z^2}{\xi^2} \right], \quad (10)$$

$$\xi^2 \equiv \frac{\alpha}{\rho g}.$$

如此, (9) 中的 L_0 就该改成 ξ 了.

以上是气液分界的讨论. 如果我们要分析的分界是晶体的表面, 则晶体的结构就不能不考虑了. 易形模型的格子排列, 可以算是一种晶格的最粗略的写照. 如果把格子的分立性考虑进去, 或可以对晶体的表面有所认识.

二维空间中的分界线, 可以直接照易形模型 (17.16) 分析, 相当简单, 可以分析得很仔细. 结果也和以上的大同小异, 是一个很好的示范例子. 三维空间中的分界面就很复杂. 把格子考虑进去后, 低温下的分界面就不再是粗糙的了. 下节先看分界线.

2. 易形模型分界面

从不等式 (17.27) 可见, s_i 的平均值在 $T \leq J$ 时, 已是十分接近 1. 如果边界上的 s_i 取为 -1 , 则几乎所有的单元都接近于 -1 . 现在我们把棋盘上方边界取 -1 , 下方取 $+1$, 如此就一定有一条横贯分界线, 见图 2. 只要 T 不是很接近 T_c , 分界线上方的单元几乎全是 -1 , 下方的几乎全是 $+1$. 现在来大略计算分界线的热位能. 分界线的能量是它的总长乘以 $2J$. 总长是

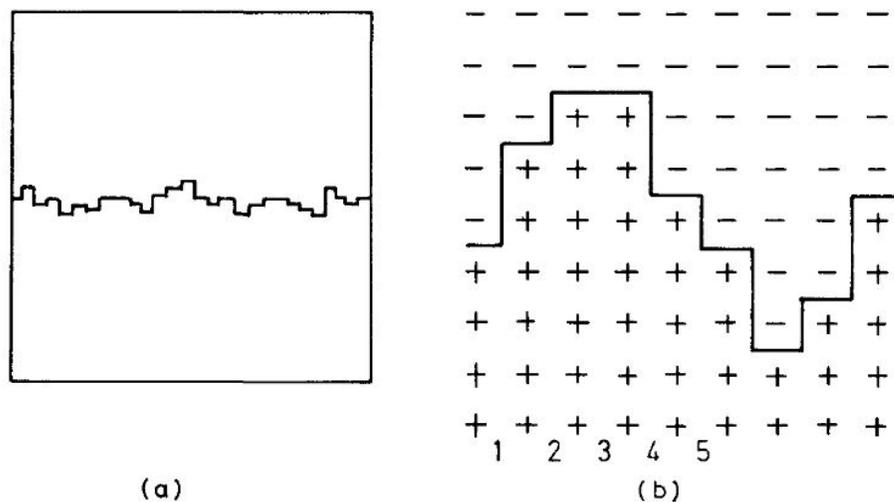


图 2

$$L = L_0 + \sum_{k=1}^{L_0} |y_k| \quad (11)$$

y_k 是, 从左算起, 第 k 个垂直分界线段长. 上图 (b) 中, $y_1 = 2$, $y_2 = 1$, $y_3 = 0$, $y_4 = -2$, $y_5 = -1$ 等. 因此分界线的热位能 F'' 是

$$F'' = -T \ln Z'',$$

$$Z'' = \sum_{y_1=-\infty}^{\infty} \sum_{y_2=-\infty}^{\infty} \dots \sum_{y_{L_0}=-\infty}^{\infty} e^{-2JL/T} \delta(y_1 + y_2 + \dots + y_{L_0}). \quad (12)$$

式中的 δ 函数迫使分界线的右端定在右边的中点. L_0 是个大数, 这 δ 函数的作用可以略去. (在 F'' 产生 $O(\ln L_0)$ 的误差). 略去 δ 函数后, Z'' 就很容易计算了:

$$Z'' = \zeta^{L_0}$$

$$\zeta = e^{-2J/T} \left(1 + 2 \sum_{y=1}^{\infty} e^{-2Jy/T} \right), \quad (13)$$

$$= e^{-2J/T} \tanh \left(\frac{J}{T} \right).$$

$$\bar{\alpha} \equiv \frac{F''}{L_0} = 2J + T \ln \tanh \left(\frac{J}{T} \right). \quad (14)$$

从大尺度观点来看, 分界线长是 L_0 , 因为不计上下的起伏. $\bar{\alpha}$ 是在温度 T 时的“表面张力”. (11) 式只在 T 很小时, 即 $T \leq J$ 时合用. 表面张力随 T 减, 但减得不多.

下一个问题是分界线的起伏幅度. 在第 k 个水平位置, 分界线高是

$$z_k = \sum_{j=1}^k y_j \quad (15)$$

因此, 从界线上一点到另一点的高差是

$$\Delta z \equiv z_{k+n} - z_k = \sum_{j=k+1}^{k+n} y_j. \quad (16)$$

只要 $n \ll L_0$, 各个 y_j 都是独立的变量. 用中心极限定理得 Δz 的分布

$$\begin{aligned} \rho(\Delta z) &\approx \frac{1}{\sqrt{2\pi}\sigma} e^{-(\Delta z)^2/2\sigma^2}, \\ \sigma^2 &= n \langle y^2 \rangle \\ &= \frac{2ne^{-2J/T}}{(1 - e^{-2J/T})^2} \\ &= \frac{n}{[2 \sinh^2(J/T)]}. \end{aligned} \quad (17)$$

因此, 这界线起伏得很厉害, 中间的起伏幅度约是

$$\frac{\sqrt{L_0}}{\sinh(J/T)}. \quad (18)$$

$\sqrt{L_0}$ 虽然比 L_0 小很多, 但不能算是一个微观尺度. 因此, 这分界线是一条曲折、粗糙的界线, 而不是一条光而平的直线.

以上的分析没有考虑图 3 所示的“悬空”型的区域, 或孤立的小区域. 不过, 即使包括了这些, 以上的结论, 即分界线为粗糙, 不受影响.

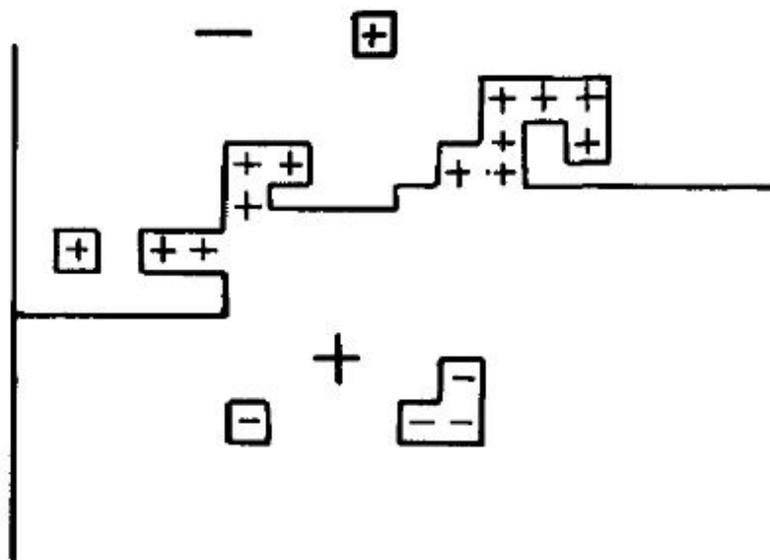


图 3

3. 晶体表面, 电荷气体模型

三十年前有人指出, 在低温, 晶体表面该是光滑的. 但当温度超过某温度则晶面会变粗糙. 如用易形模型分析, 此温度差不多该是二维空间易型磁铁的临界温度[†]. 这些人的说法大致如下

[†]Burton, Cabrera, and Frank (1951).

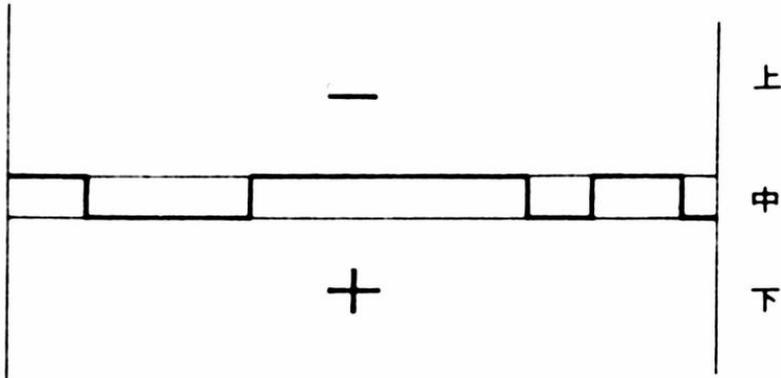


图 4

图 4 是一三维空间易形模型格子的截面图. 中层之上单元全为 $-$, 之下全为 $+$. 中层的单元则可正可负, 因为中层单元上邻为 $-$, 下邻为 $+$, 所以上下邻的影响抵消, 即中层的单元如改变, 其能量改变完全是由中层单元决定, 和上下都无关. 因此, 中层单元形成一个二维空间的易形模型. 如果温度低于 T_{c2} , 即低于二维空间易形模型的临界温度, 则中层单元就凝结成 $+$ 的, 或 $-$ 的. 当然, 仍会有一些起伏改变. 但这分界面除了这小小的起伏之外, 是光滑的. 图 4 中的粗线代表分界面, 当然, 分界面上、下的一层都有上述性质, 即不受其上邻, 下邻影响. 因此, 起伏不只限于中层. 不过, 以上的分析大致说来是非常合理的. 只要 $T \ll T_{c2}$, 分界面该是光滑的. 这分析也指出: 第 1 节的模型 (7) 之所以不能形成光滑的分界面, 是由于没有分立的层面. 层面之分立, 为光滑界面的必要条件.

用易形模型分析, 太复杂了一些. 我们现在把模型 (7) 加上一点分立层面的影响, 就可以大致看出光滑和粗糙的转变.

以下的分析是把模型 (7) 转换成一个电荷气体模型 (见第十九章第 5 节), 再用这电荷模型来分析. 怎么会和电荷扯上关系? 原因是 (7) 式中的 $(\nabla_z)^2$ 量和静电作用有密切的关系. 如果把 z 看成电位, 则 $-\nabla z$ 为电场 \mathbf{E} , (7) 正好是电场能量

$$H \propto \int d^2\mathbf{r} E^2, \quad (19)$$

而静电能量的极小值是由 $\nabla \cdot \mathbf{E} = 0$ 决定, $\nabla^2 z = 0$. 任何和 z 扯上关系的作

用, 都可以看成电荷的作用. 例如 H 中加一项

$$\int d^2\mathbf{r}q(r)z(r), \quad (20)$$

则 $q(r)$ 可看作电荷分布, 因为 $z(r)$ 是电位. 以下的步骤虽然比较繁杂, 但电荷观念的介入并非不自然. 现在的问题是如何使连续 z 和分立 z 的不同点暴露出来[‡].

现在把 (7) 改成

$$H = \alpha L_0^2 + \frac{\alpha}{2} \sum_{\mathbf{R}} \sum_{\mathbf{a}} [z(\mathbf{R}) - z(\mathbf{R} + \mathbf{a})]^2. \quad (21)$$

\mathbf{R} 是水平方向的格子位子, \mathbf{a} 是指向水平方向四邻的向量. 热位能是

$$F = -T \ln Z, \quad (22)$$

$$Z = \prod_k \sum_{z(\mathbf{R})=-\infty}^{\infty} e^{-H/T}.$$

$z(\mathbf{R})$ 现在是整数, 每层间的距离是 1. (21)、(22) 只是把 (7) 放在分立的格子上. 这不是易形模型, 而是一个新的模型. 如果把 (22) 的求和改为对 z 的积分, 则这模型和 (7) 就没有什么不同了. 水平方向的格子分立性, 没有太大的影响. 等下就明白了.

求积分与求和间的差, 可以用第十六章的“泊松求和公式”(见 (16.49)). 在此 z 的值可从 $-\infty$ 到 ∞ . 这公式可写成

$$\sum_{n=-\infty}^{\infty} f(n) = \sum_{q=-\infty}^{\infty} f_q = \int_{-\infty}^{\infty} dz f(z) + \sum_{q=1}^{\infty} (f_q + f_{-q}), \quad (23)$$

$$f_q \equiv \int_{-\infty}^{\infty} dz e^{-i2\pi qz} f(z).$$

详细的推导, 见第十六章.

(22) 式中的每个求和都用上式代入, 得

$$Z = Z_0 \prod_{\mathbf{R}} \sum_{q(\mathbf{R})=-\infty}^{\infty} \left\langle \exp \left\{ -i2\pi \sum_{\mathbf{R}} q(\mathbf{R})z(\mathbf{R}) \right\} \right\rangle_0. \quad (24)$$

[‡]以下的方法, 见 Chui and Weeks (1976), Weeks (1980).

Z_0 是把 $z(\mathbf{R})$ 看成连续变量所求得的 Z , $\langle \dots \rangle$. 是对这连续值模型求的平均值:

$$Z_0 \equiv \prod_{\mathbf{R}} \int dz(\mathbf{R}) e^{-H/T} \quad (25)$$

$$\langle A \rangle_0 \equiv \prod_{\mathbf{R}} \int dz(\mathbf{R}) A e^{-H/T} Z_0^{-1}.$$

如果 $z(\mathbf{R})$ 为连续变量, 则其分布为正态, 因为 (21) 是 z 的二项式. 既是正态, 则

$$\left\langle \exp \left\{ -i2\pi \sum_{\mathbf{R}} q(\mathbf{R}) z(\mathbf{R}) \right\} \right\rangle_0 \quad (26)$$

$$= \exp \left\{ -\frac{(2\pi)^2}{2} \sum_{\mathbf{R}} \sum_{\mathbf{R}'} q(\mathbf{R}) q(\mathbf{R}') \langle z(\mathbf{R}) z(\mathbf{R}') \rangle_{0c} \right\}.$$

这是正态分布的性质 (见 (12.32)、(12.38) 及第十二章的讨论). $\langle z(\mathbf{R}) z(\mathbf{R}') \rangle_{0c}$ 的计算和 (8) 差不多, 如果 $|\mathbf{R} - \mathbf{R}'|$ 很大, 则

$$\langle z(\mathbf{R}) z(\mathbf{R}') \rangle_{0c} \approx -\frac{T}{2\pi\alpha} \ln \frac{|\mathbf{R} - \mathbf{R}'|}{a} + \text{常数}. \quad (27)$$

代入 (26), (24) 得

$$Z = Z_0 \prod_{\mathbf{R}} \sum_{q(\mathbf{R})=-\infty}^{\infty} \exp \left\{ -\frac{1}{2T'} \sum_{\mathbf{R}, \mathbf{R}'} q(\mathbf{R}) q(\mathbf{R}') u(\mathbf{R} - \mathbf{R}') \right\}, \quad (28)$$

$$u(\mathbf{R} - \mathbf{R}') = -\ln \frac{|\mathbf{R} - \mathbf{R}'|}{a},$$

$$T' \equiv \frac{\alpha}{2\pi T}.$$

这个模型, 正好是第十九章第 5 节讲的二维空间电荷气体模型, 只是电荷可以是任何整数, 且“温度” T' 是和 T 成反比, 和“表面张力” α 成正比. $q(\mathbf{R})$ 是在 \mathbf{R} 处的电荷. 如果所有的 $q(\mathbf{R})$ 都是零, 则我们回到模型 (7). 因此, 分立层面的影响, 可以想成这些电荷的影响. 注意, 由于 $+q$ 和 $-q$ 的对称, 总电荷是零. 如要计算 $\langle (\Delta z)^2 \rangle$, 除了 (8) 之外, 我们还得考虑这些电荷的影响. $\langle z(\mathbf{R}) z(\mathbf{R}') \rangle_c$ 可以由以下步骤求出: 将 $q(\mathbf{R}), q(\mathbf{R}')$ 改成 $q(\mathbf{R}) + \lambda, q(\mathbf{R}') + \lambda'$, 令 λ, λ' 为很小的量. 如此, 热位能的增加可以由 (24) 计出, 得

$$\frac{\partial^2 F}{\partial \lambda \partial \lambda'} = -\frac{4\pi^2}{T} \langle z(\mathbf{R}) z(\mathbf{R}') \rangle_c, \quad (29)$$

$$\langle (\Delta z)^2 \rangle \equiv 2 \langle z^2 \rangle - 2 \langle z(\mathbf{R}) z(\mathbf{R}') \rangle_0.$$

λ, λ' 可以看成在 \mathbf{R}, \mathbf{R}' 加放的电荷, 因此, (29) 是这两个电荷间的作用能 (除以 $\lambda\lambda'$). 因此求 $z(\mathbf{R}), z(\mathbf{R}')$ 的相干值成了求二电荷间的作用能.

由第十九章的分析可知, 在“低温”, 即 T' 很小时, 正、负电荷会自己成双, 介电系数为有限, 而电荷间作用仍为对数式, 即

$$\langle z(\mathbf{R})z(\mathbf{R}') \rangle_c \propto -\ln \frac{|\mathbf{R} - \mathbf{R}'|}{a}, \quad (30)$$

$$T' < \frac{1}{4}, \quad \text{即} \quad T > \frac{2\alpha}{\pi}.$$

1/4 是由 (19.28) 得来, 最小的 q 是 1. 如果 $T' > 1/4$, 则有独行离子存在, 产生“屏蔽”现象, 因此

$$\langle z(\mathbf{R})z(\mathbf{R}') \rangle_c \propto \frac{e^{-|\mathbf{R}-\mathbf{R}'|/b}}{\sqrt{|\mathbf{R}-\mathbf{R}'|}}, \quad (31)$$

$$T' > \frac{1}{4}, \quad \text{即} \quad T < \frac{2\alpha}{\pi},$$

$$\frac{1}{b^2} = \frac{4\pi n}{T'}. \quad (32)$$

n 是电荷 +1 或 -1 的密度 (见 (19.34)). 大一些的, $|q| > 1$, 的电荷不发生什么作用.

因此, 在高温, $T > 2\alpha/\pi$, 第 1 节的粗糙界面结论不变. 在低温, $T < 2\alpha/\pi$ 时, 界面可视为光滑, 起伏的相干距离为 b . 电荷的密度 n 须要计算, 在此不谈.

以上的分析, 是用“泊松求和公式”把分立层面的影响写成一群电荷的影响. 再用静电作用的模型来求得解答. 这种把一个模型转换成另一模型再求解的做法, 是一个常见到的技巧. 这静电模型在下两章会再度出现.

晶体表面的粗糙或光滑, 对结晶速度很有影响. 这方面文献甚多[§]. 读者必须注意, 以上的模型并不足以分析固体表面的结构问题. 固体的表面, 是一门大学问, 必须从分子间的作用力着手.

讨论问题二十八

1. 复习第二十七章问题 2, (1) 式中的 y 和 (7) 式中的 z 相当于该题中的 b . 讨论该题的尺度 ξ 在本章第 1 节模型中的意义.

[§]见 Chui and Weeks (1980), Krumbhaar in Binder (1979).

2. (A) 求模型 (1) 的热位能 F . 可由 (3) 开始. 每一个 k 单元的热位能都很容易求.

(B) 令 $\bar{\alpha} = F/L_0$, 这可称为“表面张力”. 这是大尺度的表面张力, 和小尺度的表面张力 α (见 (1), (3)) 不一样, 试讨论 $\bar{\alpha}$, α 的意义, 并估计其差别. 用水作例子.

(C) 将这结果和第 2 节的结果比较.

3. 从易形分界模型导出连续界面模型.

(A) 由模型 (11) 开始, 取

$$k = 0, b, 2b, 3b, \dots, mb, \dots \quad (33)$$

诸点, 令 $b \gg 1$, 这些点如果用直线连起来, 就成了一个粗略化的界面. mb 和 $(m+1)b$ 间的详细变化都看不见了. 试证 (12) 的 Z'' 可写成,

$$Z'' = \prod_{m=1}^{L'} \sum_{\eta_m=-\infty}^{\infty} \zeta(\eta_m) \delta(\eta_1 + \eta_2 + \dots + \eta_{L'}), \quad (34)$$

$$L' = L_0/b$$

$$e^{2jT} \zeta(\eta) = \prod_{k=1}^m \sum_{y_k=-\infty}^{\infty} e^{-2J|y_k|T} \delta(\eta - y_1 - \dots - y_N). \quad (35)$$

(B) 利用中心极限定理, 试证

$$\ln \zeta(\eta) \approx b \ln \zeta - \frac{1}{2} \ln(2\pi b \langle y^2 \rangle) - \frac{\eta^2}{2b \langle y^2 \rangle}. \quad (36)$$

ζ 的定义见 (13), (14), $\langle y^2 \rangle$ 见 (17).

(C) 定义有效总能量为

$$H'(\eta_1, \eta_2, \dots, \eta_{L'}) = -T \sum_{m=1}^{L'} \ln \eta(\eta_m). \quad (37)$$

试证由 H 所得的热位能和由 (11) 的 $H = 2JL$ 所得相同. 这只是检验以上步骤的正确性.

(D) 在 (36), 只有最后一项和 η 有关, 其他都是常数, 试证

$$H' = \text{常数} + \frac{\alpha}{2} \int dx \left(\frac{dz}{dx} \right)^2$$

$$\alpha \equiv \frac{T}{\langle y^2 \rangle}. \quad (38)$$

z 看成 $x = mb$ 的平缓函数, dx 即是 b , $\eta/b = dz/dx$. 上式的积分只是对 m 求和的缩写.

因此, 第 2 节的模型和第 1 节的模型大致一样的, 只是尺度不同. (38) 或 (1) 的尺度是 b , 而 (11) 用的尺度是 1. 注意, 在此 $b \gg 1$, $dz \gg 1$, 并注意 (38) 的常数项不是 αL_0 .

(E) 注意 (38) 的 α 是个复杂的 T 的函数 (见 (14)). (38) 式的常数项非常重要. 如果要得到正确的 $\bar{\alpha}$, 这常数项一点都不能马虎. 读者必须了解 H' 的意义和来历. 务必多加思考. 试讨论 b 的重要性. 从 (38) 看来, 并无 b 的影子. 但如没有 b , (38) 的意义就有了问题.

4. 用 (10) 计算水的 ξ 值, 并用 (9) 计算室温下水面厚度.

5. 讨论一个水珠形状的起伏.

(A) 将 (7) 推广, 以讨论近于球面的表面.

(B) 用球面正交函数分析水珠形状.

第二十九章 连续对称模型

本章讨论三个例子, 平面向量模型 (即 XY 模型), 晶体模型以及量子向量模型. 用相当粗略的分析, 来说明序参量的起伏. 起伏幅度和空间的维数有关, 和对称性也有密切的关系. 这些例子指明这些关系, 最后我们讨论这些模型的共同点, 即连续性的对称以及凝结. 在这种冻结情形下, 会有低能量的长波“柔软波”出现. 在一、二维空间的模型中, 柔软波能使凝结软化而动摇. 除了平缓的柔软波之外“缺陷”也是一种重要的起伏. 我们用二维空间的 XY 模型为例, 略谈缺陷的来源及后果.

1. 平面向量模型

我们先看一个最简单的向量模型, 其哈密顿量是

$$H = -\frac{1}{2} \sum_{i,j} J_{ij} \mathbf{s}_i \cdot \mathbf{s}_j, \quad (1)$$
$$J_{ij} = J, \quad \text{如果 } i, j \text{ 相邻,}$$
$$= 0, \quad \text{如不相邻.}$$

\mathbf{s}_i 为长度为壹的平面向量, $\mathbf{s}_i = (s_{ix}, s_{iy})$, $s_{ix}^2 + s_{iy}^2 = 1$. 这就是所谓 XY “模型”. 和下一章的超流体模型相似. 我们现在来看向量方向的起伏. 先把 (1) 简化一点, 令

$$s_{ix} = \cos \theta_i,$$
$$s_{iy} = \sin \theta_i, \quad (2)$$

$$\text{则 } \mathbf{s}_i \cdot \mathbf{s}_j = \cos(\theta_i - \theta_j).$$

现在假定相邻的角度都差不多, 则

$$\cos(\theta_i - \theta_j) \approx 1 - \frac{1}{2}(\theta_i - \theta_j)^2.$$

令 \mathbf{R} 指定各向量的位置, (1) 可以化为

$$H = \frac{1}{4} J \sum_{\mathbf{R}} \sum_{\mathbf{a}} [\theta(\mathbf{R}) - \theta(\mathbf{R} + \mathbf{a})]^2 + \text{常数}. \quad (3)$$

\mathbf{a} 指向 \mathbf{R} 的各邻居. 最后, 令

$$\begin{aligned} \theta(\mathbf{R}) &= \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \theta_{\mathbf{k}} e^{i\mathbf{k} \cdot \mathbf{R}}, \\ \theta_{\mathbf{k}} &\equiv \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} \theta(\mathbf{R}) e^{-i\mathbf{k} \cdot \mathbf{R}}. \end{aligned} \quad (4)$$

代入 (3) 得

$$\begin{aligned} H &= \frac{1}{2} \sum_{\mathbf{k}} J_{\mathbf{k}} |\theta_{\mathbf{k}}|^2, \\ J_{\mathbf{k}} &\equiv J \sum_{\mathbf{a}} |1 - e^{-i\mathbf{k} \cdot \mathbf{a}}|^2 = 2J \sum_{\mathbf{a}} (1 - \cos \mathbf{k} \cdot \mathbf{a}). \end{aligned} \quad (5)$$

在 k 很小时, $J_{\mathbf{k}} \propto k^2$. 因此, 各 $\theta_{\mathbf{k}}$ 独立, 为正态分布, 一切平均值都容易计算. 能量均分定律指出

$$\langle |\theta_{\mathbf{k}}|^2 \rangle = \frac{T}{J_{\mathbf{k}}}. \quad (6)$$

由此可以计算

$$\begin{aligned} \langle (\Delta\theta)^2 \rangle &\equiv \langle (\theta(\mathbf{R}) - \theta(0))^2 \rangle \\ &= \frac{2}{N} \sum_{\mathbf{k}} \frac{T}{J_{\mathbf{k}}} (1 - \cos \mathbf{k} \cdot \mathbf{R}). \end{aligned} \quad (7)$$

如果 R 很长, 则小 k 诸项最重要. 把 $J_{\mathbf{k}}$ 写成小 k 的近似 $J_{\mathbf{k}} \approx \alpha k^2 a^2$, 则

$$\langle (\Delta\theta)^2 \rangle = \frac{2T}{N\alpha} \sum_{\mathbf{k}} \frac{1}{a^2 k^2} (1 - \cos \mathbf{k} \cdot \mathbf{R}). \quad (8)$$

这个计算在一维、二维空间都已在上一章做过 (见 (28.5), (28.8)),

$$\langle (\Delta\theta)^2 \rangle = \begin{cases} \frac{2TR}{\alpha a}, & d = 1, \\ \frac{T}{\pi\alpha} \ln \frac{R}{a}, & d = 2, \\ \frac{T}{\pi^2\alpha} \left(k_D a - \frac{\pi a}{4R} \right), & d = 3. \end{cases} \quad (9)$$

d 指空间的维数, $d = 3$ 的积分亦不难, 注意 $1/k^2$ 即是 $-\nabla^2$ 解, 即静电方程式

$$-\nabla^2 \phi = \delta(\mathbf{R}) \quad (10)$$

的解. 除了一些细节之外, (9) 即是 (10) 之解. $1/k_D$ 约是格子的大小. 由 (9) 可见, 角度的起伏在 $d \leq 2$ 时, 随 R 增大, 和上章第一节的液面包伏情形一样. 各个 \mathbf{s}_i 的相干函数可由 (9) 求得:

$$\begin{aligned} C(\mathbf{R}) &\equiv \langle \mathbf{s}(\mathbf{R}) \cdot \mathbf{s}(0) \rangle = \langle \cos(\theta(\mathbf{R}) - \theta(0)) \rangle \\ &= \operatorname{Re} \langle e^{i(\theta(\mathbf{R}) - \theta(0))} \rangle \\ &= e^{-((\Delta\theta)^2)/2}. \end{aligned} \quad (11)$$

这最后一步是利用 θ 的正态分布 (见连接值定理 (12.32), (12.38)), 由 (9) 得

$$C(\mathbf{R}) = \begin{cases} e^{-(T/\alpha)R/a}, & d = 1, \\ \left(\frac{R}{a}\right)^{-T/2\pi\alpha} & d = 2, \\ e^{-Tk_D a/\pi^2\alpha} \left[1 + \frac{\pi}{4k_D R}\right] & d = 3. \end{cases} \quad (12)$$

以上假定 $R \gg a$.

从这些结果可见, $d \leq 2$ 的情形下, 无论 T 多小, 只要 $T \neq 0$, $C(R)$ 都随 R 减小, 渐趋于零. 只有在 $d = 3$ 时 $C(R)$ 趋近于常数. 也就是说, $d \leq 2$ 时, 没有集体排列的现象, 起伏太大了. 注意在 $d = 2$ 时, $C(R)$ 消失得很慢, 在一般合理的尺度下, $C(R)$ 不能趋于零. 这一点在上一章也提过. (9) 中的 $\ln(R/a)$ 不会很大. 因此, 在 $d = 2$, 凝结现象不能说没有.

照 (12) 的结论, $d = 2, 3$ 在任何温度下都有排列现象. 也就是说, 在 (12) 中找不出临界温度的影子, 怎么回事? 原因是以上假设了 $\Delta\theta$ 为小量. 如果 $|\Delta\theta| \geq \pi$, 则以上的结果就没什么意义. π 和 $-\pi$ 是同一方向, θ 和 $\theta + 2n\pi$, $n = \pm 1, \pm 2, \dots$, 是一样的. 因此, (9), (12) 这些结论只合用于低温. 以上无排列的结晶也不是严密的证明, 不过严密的证明是存在, 第 4 节再提. (一维空间的模型很容易解, (12) 的结论不错, 读者试解之).

我们现在再看一个例子, 即晶体, 乍看起来和这 XY 模型大不相同, 但事实上十分相似.

2. 晶体的密度起伏

晶体中的原子作周期性排列, 因此, 密度不均匀. X 光的散射实验所观测的是散射强度, 它和

$$\langle |\rho_{\mathbf{k}}|^2 \rangle = |\langle \rho_{\mathbf{k}} \rangle|^2 + \langle |\rho_{\mathbf{k}}|^2 \rangle_c \quad (13)$$

成正比.

实验的分析一如第十一章的第 4 节, 由于晶体的周期性排列, 在某些特别的 \mathbf{k} 值下, 即 \mathbf{k} 等于所讲“倒逆向量”^{*}之一时, $\langle \rho_{\mathbf{k}} \rangle$ 就不等于零. 我们先来看 $\langle \rho_{\mathbf{k}} \rangle$ 和振动位移 \mathbf{u} 的关系:

$$\begin{aligned} \langle \rho_{\mathbf{k}} \rangle &\equiv \int d^3\mathbf{r} e^{-i\mathbf{k}\cdot\mathbf{r}} \langle \rho(\mathbf{r}) \rangle, \\ &= \sum_{\mathbf{R}} \langle e^{-i\mathbf{k}\cdot(\mathbf{R}+\mathbf{u}(\mathbf{R}))} \rangle. \end{aligned} \quad (14)$$

式中的 \mathbf{R} 指各原子的平均位置, $\mathbf{R} + \mathbf{u}(\mathbf{R})$ 为各原子坐标. $\mathbf{u}(\mathbf{R})$ 为运动变量, 现在略去量子力学, $\mathbf{u}(\mathbf{R})$ 是一群简谐运动坐标之和:

$$\mathbf{u}(\mathbf{R}) = \frac{1}{\sqrt{N}} \sum_{\lambda q} e^{i\mathbf{q}\cdot\mathbf{R}} \eta_{q\lambda} \hat{\mathbf{e}}_{\lambda q}. \quad (15)$$

每个 $\eta_{q\lambda}$ 的概率分布都是正态分布, 且

$$\langle \eta_{q\lambda} \rangle = 0, \quad \langle |\eta_{q\lambda}|^2 \rangle = \frac{T}{m\omega_{q\lambda}^2}. \quad (16)$$

正态分布的连接值只到二次为止. 因此, 用第十三章的连接值定理 (见 (12.32), (12.38)) 得

$$\begin{aligned} \langle e^{-i\mathbf{k}\cdot\mathbf{u}(\mathbf{R})} \rangle &= e^{-\frac{1}{2}\langle (\mathbf{k}\cdot\mathbf{u}(\mathbf{R}))^2 \rangle} \\ &\equiv e^{-W(\mathbf{k})}. \end{aligned} \quad (17)$$

W 在第十一章已算过, (见 (11.21)). 用德拜模型, 得

$$W(\mathbf{k}) = \frac{k^2 T \omega_D}{4\pi^2 n m c^3}, \quad d = 3. \quad (18)$$

n 是原子密度 N/V , m 为原子质量, c 为音速, ω_D 为德拜频率. (13) 的第一项 $|\langle \rho_{\mathbf{k}} \rangle|^2$ 为 X-光散射强度的主要项. 当 $\mathbf{k} = \mathbf{G}$ 刚好是“倒逆向量”之一时,

$$\begin{aligned} e^{i\mathbf{G}\cdot\mathbf{R}} &= 1, \\ |\langle \rho_{\mathbf{G}} \rangle|^2 &= N^2 e^{-2W(\mathbf{G})}. \end{aligned} \quad (19)$$

温度越高, 这强度越小. e^{-2W} 即所谓“德拜-瓦勒因子”, 指出因振动所引起的散射衰落.

^{*}倒逆向量 \mathbf{G} 代表周期性结构. $e^{i\mathbf{G}\cdot\mathbf{r}}$ 的周期和晶体的周期重合.

现在来计算 (13)

$$\begin{aligned}\langle |\rho_{\mathbf{k}}|^2 \rangle &= \sum_{\mathbf{R}} \sum_{\mathbf{R}'} e^{-i\mathbf{k} \cdot (\mathbf{R} - \mathbf{R}')} \left\langle e^{-i\mathbf{k} \cdot (\mathbf{u}(\mathbf{R}) - \mathbf{u}(\mathbf{R}'))} \right\rangle \\ &= N \sum_{\mathbf{R}} e^{-i\mathbf{k} \cdot \mathbf{R}} e^{-\frac{1}{2} \langle (\mathbf{k} \cdot \Delta \mathbf{u})^2 \rangle}, \\ \Delta \mathbf{u} &\equiv \mathbf{u}(\mathbf{R}) - \mathbf{u}(0).\end{aligned}\quad (20)$$

最后一步又用了连接值定理, 和 (17) 相似, $\langle (\mathbf{k} \cdot \Delta \mathbf{u})^2 \rangle$ 的计算, 和上一章第 1 节的 $\langle (\Delta z)^2 \rangle$, 及本章的 $\langle (\Delta \theta)^2 \rangle$ 都差不多, (见 (8), (9)):

$$\begin{aligned}\langle (\mathbf{k} \cdot \Delta \mathbf{u})^2 \rangle &= \frac{2k^2 T}{nmc^2} \int \frac{d^d \mathbf{q}}{(2\pi)^d} \frac{1}{q^2} (1 - \cos \mathbf{q} \cdot \mathbf{R}) \\ &= \frac{2T}{\alpha} R, \quad d = 1, \\ &= \frac{T}{\pi \alpha} \ln \left(\frac{R}{a} \right), \quad d = 2, \\ &= \frac{T}{\pi^2 \alpha} \left(\frac{\omega_D}{c} - \frac{\pi}{4R} \right), \quad d = 3,\end{aligned}\quad (21)$$

$$\alpha \equiv \frac{nmc^2}{k^2}.$$

这和 (9) 可以说是一样, 积分中的 $1/q^2$ 是从 (16) 式中的 $1/\omega_{\mathbf{q}\lambda}^2$ 来的. $\omega_{\mathbf{q}\lambda}^2 = c^2 q^2$. 这 q^2 是 ∇^2 的化身, 弹性体的能量是和位移的微分的平方成正比. 如果整个弹性体作一整体位移, 即 $q = 0$, 则不增加能量, q 很小时, 能量亦小. 由于小 q 振动, 即长波振动, 使振幅 \mathbf{u} 变大. 由 (18), (21) 可知, 在 $d \leq 2$ 时, 振幅已大得使晶体失去周期性的排列. 当然, 在 $d = 2$ 时, $\ln(R/a)$ 不会很大, 周期性并未全失.

现在把 (21) 代入 (20), 令 $\mathbf{k} = \mathbf{G}$, 得

$$\langle |\rho_{\mathbf{G}}|^2 \rangle = \begin{cases} \frac{2Nn\alpha}{T} & d = 1, \\ e^{-2W(\mathbf{G})} [N^2 + O(N^{5/3})] & d = 3 \end{cases}\quad (22)$$

$$\alpha \equiv \frac{nmc^2}{G^2}.$$

$d = 2$ 的情形比较特别,

$$\begin{aligned} \langle |\rho_{\mathbf{G}}|^2 \rangle &= \sum_{\mathbf{R}} \left(\frac{R}{a} \right)^{-T/2\pi\alpha} \\ &\approx \left(1 - \frac{T}{4\pi\alpha} \right) N^{2-(T/2\pi\alpha)}, \quad T < 4\pi\alpha, \\ &\approx \left(\frac{T}{4\pi\alpha} - 1 \right) N, \quad T > 4\pi\alpha. \end{aligned} \quad (23)$$

因此, 三维空间晶体的 X 光散射的强点, (即 $\mathbf{k} = \mathbf{G}$ 的散射), 有强度 $\sim N^2$, 加上 $N^{5/3}$ 及更小的修正量. e^{-2W} 随 T 减小 (见 (18), (19)). N^2 的存在代表晶体的周期性结构, 一维空间的晶体, 则无周期性结构, $\langle |\rho_{\mathbf{G}}|^2 \rangle$ 只和 N 成正比. 二维空间晶体在温度高的时候 $\langle |\rho_{\mathbf{G}}|^2 \rangle$ 不显出周期性结构, 而在 $T < 4\pi\alpha$ 时, 是介于 N^2 和 N 之间[†].

3. 量子向量模型

以上的各例, 都没有牵涉到量子力学, 现在看一个最简单的量子磁铁模型, 其哈密顿算符为

$$H = -\frac{1}{2}J \sum_{\mathbf{R}} \sum_{\mathbf{a}} \mathbf{s}(\mathbf{R}) \cdot \mathbf{s}(\mathbf{R} + \mathbf{a}). \quad (24)$$

\mathbf{R} 为格子位置, \mathbf{a} 指向 \mathbf{R} 的邻居. $\mathbf{s}(\mathbf{R})$ 为自旋 1/2 算符, 其三个分量是

$$s_x = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad s_y = \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad s_z = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (25)$$

另外有两个方便的组合, s_+ , s_-

$$\begin{aligned} s_{\pm} &\equiv s_x \pm i s_y \\ s_+ &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad s_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \end{aligned} \quad (26)$$

每一个 \mathbf{R} 都有这样一套算符.

[†]更仔细的分析, 见 Imry and Gunther (1971).

H 的基态很容易看出, 只要所有的自旋都指向一方, H 就是极小, 现在挑出一个特别的方向, $-Z$ 的方向, “向下” 的方向, 所有自旋都指向下. 称此态为 Φ_0

$$\begin{aligned} H\Phi_0 &= E_0\Phi_0, \\ E_0 &= -\frac{1}{4}JdN. \end{aligned} \quad (27)$$

d 是空间维数, $2d$ 是每格子的邻居数 (假定立方格子). N 是格子总数, 一如以上诸例. 注意, \mathbf{s} 有三个分量, 即 \mathbf{s} 是一三度空间的向量, 而格子的空间维数 d 可以有选择.

现在来看 H 的激发态, 最简单的是把一自旋倒过来, 得

$$\varphi(\mathbf{R}) \equiv s_+(\mathbf{R})\Phi_0. \quad (28)$$

一共有 N 个这样的的激发态, 但它们不是 H 的本征向量, 即不是定能态. 好在 $H\varphi$ 不会牵涉到其他种类的激发态, (只牵涉到反转一个自旋的). 因此, 由此类状态的组合, 可得定能态. 最简单的是

$$\begin{aligned} \sum_{\mathbf{R}} \varphi(\mathbf{R}) &= S_+\Phi_0, \\ S_+ &\equiv \sum_{\mathbf{R}} s_+(\mathbf{R}). \end{aligned} \quad (29)$$

这组和事实上是另一个基态, 而非激发态. 因为 H 是个标量, 不因坐标系的转移而变, 即 H 有转动对称性:

$$\begin{aligned} [H, \mathbf{S}] &= 0, \\ \mathbf{S} &\equiv \sum_{\mathbf{R}} \mathbf{s}(\mathbf{R}). \end{aligned} \quad (30)$$

\mathbf{S} 是总角动量. 因此, 将 S_{\pm} 施于 Φ_0 , 即得一同样能量的状态, 再施之, 又得一个, 一共有 $N+1$ 个基态.

现在令

$$\varphi_{\mathbf{k}} \equiv \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} \varphi(\mathbf{R}). \quad (31)$$

这状态是 H 的本征状态, 其能量 (高出 E_0) $\omega_{\mathbf{k}}$ 为

$$\begin{aligned} (H - E_0) \varphi_{\mathbf{k}} &= \omega_{\mathbf{k}} \varphi_{\mathbf{k}}, \\ \omega_{\mathbf{k}} &= \frac{1}{2}J \sum_{\mathbf{a}} (1 - \cos \mathbf{k} \cdot \mathbf{a}). \end{aligned} \quad (32)$$

这结果直接由 (31) 代入 (24) 求得. 注意, H 在格子的平移下不变 (假定循环边界, 即无边界), 故 H 的本征状态必为平移算符的本征状态, 即平面波. (31) 是一个“单磁子”的状态, 只有一个自旋倒过来. 这“磁子”可以看成一粒子, 能量为 $\omega_{\mathbf{k}}$, 动量为 \mathbf{k} . (28) 可以看成限在 \mathbf{R} 的磁子.

在低温时, 磁子粒子数密度小, 这模型又成了一理想气体模型. (我们把两个磁子, 或更多磁子重迭的情形略去).

磁子是波色粒子, 因为 $[\mathbf{s}(\mathbf{R}), \mathbf{s}(\mathbf{R}')] = 0$ (只要 $\mathbf{R} \neq \mathbf{R}'$, 即不重迭.) 磁子粒子数不是守恒量, 故化学势为零, 粒子数分布一如晶体振动的声子分布. 总能量为

$$\begin{aligned} E - E_0 &= \sum_{\mathbf{k}} \frac{\omega_{\mathbf{k}}}{e^{\omega_{\mathbf{k}}/T} - 1} \\ &= \frac{V}{(2\pi)^d} \int d^d \mathbf{k} \frac{\omega_{\mathbf{k}}}{e^{\omega_{\mathbf{k}}/T} - 1} \end{aligned} \quad (33)$$

每翻一自旋, Z 方向的总角动量就加 1, 磁子总数是

$$\begin{aligned} N' &= \sum_{\mathbf{k}} \frac{1}{e^{\omega_{\mathbf{k}}/T} - 1}, \\ |\langle \mathbf{S} \rangle| &= \frac{N}{2} \left(1 - \frac{2N'}{N} \right). \end{aligned} \quad (34)$$

N' 越大, 总磁矩越小. 从 (33), (34) 可计算热力性质.

如 k 很小, 则 (32) 成为

$$\begin{aligned} \omega_{\mathbf{k}} &\approx \lambda k^2, \\ \lambda &\equiv \frac{1}{4} J a^2. \end{aligned} \quad (35)$$

如果 $\omega_{\mathbf{k}} \ll T$, 则

$$\frac{1}{e^{\omega_{\mathbf{k}}/T} - 1} \approx \frac{T}{\omega_{\mathbf{k}}}. \quad (36)$$

我们可以估计一下 N' :

$$\begin{aligned} \frac{N'}{N} &\approx \frac{1}{n(2\pi)^d} \int d^d \mathbf{k} \left(\frac{T}{\lambda k^2} \right) \\ &\approx \left(\frac{T}{n\lambda} \right) L, \quad d=1, \\ &\approx \frac{T}{n\lambda} \ln \left(\frac{L}{a} \right), \quad d=2, \\ &\approx \frac{T k_D}{2\pi \lambda n}, \quad d=3. \end{aligned} \quad (37)$$

$k_D \sim 1/a$, 是 k 积分的上限, L 是物体的大小, $n \equiv N/V$. 上两节的故事, 现在又重演一次. 在 $d \leq 2$ 情形下, 无论 T 多小, 只要 L 够大, N'/N 可以超过 $1/2$. 如 N'/N 太大, 则总磁矩消失 (见 (34)). 在 $d = 2$ 时, $\ln(L/a)$ 不会太大, 因此, 只要 T 够小, 铁磁性仍可维持. 在 $d = 3$, 铁磁性在低温是没问题的. 在 $d = 1$, 铁磁性就不能存在了.

在这个模型里, 自旋的起伏是用磁子来形容, 没有用向量的方向. 其实不用量子力学, 可以很快地得到以上 (37) 的结论. 在 (36), 我们已经把量子力学丢了. 本节的目的之一就是要指出, (37) 这一类的结论, 和量子力学无关.

4. 连续对称和柔软波

从液面的起伏, 到自旋的起伏, 我们反复得到同一类型的结果. 显然这一连串的例子在反映某一种普遍的现象.

这些例子的模型, 有一共同点, 即哈密顿函数 H 有某种连续性的对称, 而在凝结态下, 这连续性的对称因冻结而遭受破坏.

所谓对称, 即是把某些变数改了, 而 H 不变. 例如把坐标方向改了, 向量模型的能量不变, 把坐标平移, 原子间作用能不变, 把液面整个上移、下移, 能量亦不变. 对称性透过统计力学的基本假设, 成了平衡态的均匀性. 旋转对称的结果是各方向的均匀, 不能有特别的方向出现. 平移对称的结果是均匀的空间分布, 不能有不均匀的密度. 这是因为状态空间中活动范围的划定, 只和总能量有关.

所以, 凝结现象违反了这均匀性质. 状态空间中的运动轨迹被限制在基本假设规定的活动范围的一部份, 没有时间跑出去. 平衡的状态下出现特别的方向, 如总磁矩, 不均匀的密度, 如固体, 如液面.

在此我们再提醒一次: 凝结现象不只限于规则的凝固态. 非晶磁体中冻住的自旋是混乱的. 照我们的定义, 这冻结状态仍指定了特别的方向. 如果把坐标系转了, 状态就不同了. 虽然所有自旋的向量和是零, 但每个自旋都有其平均值, 指向某个方向. 我们可以任取三个垂直的自旋的方向, 订为附在体内的方向指标. 这些指标是冻住不变的. 任何固体都违反了空间分布的均匀, 不只是晶体. 原子位置冻住了, 平移的对称就破坏了, 晶体也好, 不规则排列也好. 我们可以挑选四个不在同一平面上的原子位置, 来确定这固体的位置和方向.

何谓连续性的对称? 连续的反面是分立. 易形磁铁模型的对称是分立的, 把各 s_i 改成 $-s_i$, 能量不变. 这个变数的转换不是连续的. 向量模型的对称是连续的, 坐标轴方向可以连续地改变. 坐标的平移也是连续的改变.

连续性的对称被冻的结果之一, 是很多基态. 例如上例中的磁铁有 $N + 1$ 个基态, 也就是说磁矩凝结在任一方向均可. 固体可以停在任一位置, 有了这许多基态, 就不难想见许多能量非常小的激发态. 解释如下:

一个群体是由许多部份组成, 例如以上的磁铁模型, 可以看成许多磁铁放在一起. 在基态, 所有自旋指向一方, 现在把各部份的自旋稍转一点, 如所有自旋都转了同一角度, 则能量不变. 但如各部所转稍有不同, 转角平缓地由一部改到另一部, 则能量会稍高, 这就是一个低能的激发态. 这些话可以用量子力学式子写下, 把各自旋, 以 x 轴为轴, 转 θ 角的算符是

$$e^{-i\theta S_x} \approx 1 - i\theta S_x = 1 - i\theta \sum_{\mathbf{R}} s_x(\mathbf{R}), \quad (38)$$

假定 θ 很小, 如果每个自旋被转了不同的角度, 则算符是

$$U \equiv 1 - i \sum_{\mathbf{R}} \theta(\mathbf{R}) s_x(\mathbf{R}). \quad (39)$$

如果每个 θ 都相同, 则 $UHU^{-1} = H$, 即 $U\Phi_0$ 也是基态. H 是许多项之和, 每一项都是有对称性的, 每一项都只有少数几个自旋算符. 如 (24) 中的

$$H_{\mathbf{R}} \equiv -J\mathbf{s}(\mathbf{R}) \cdot \mathbf{s}(\mathbf{R} + \mathbf{a}). \quad (40)$$

如果这两个自旋被转了相同的角, 则此项不变, 但如果被转了不同的角, 即 $\theta(\mathbf{R}) \neq \theta(\mathbf{R} + \mathbf{a})$, 则此项要变. 变多少呢? 这就要看这两自旋本来的方向如何, 如果它们本来成直角, 则此项改变和 $\theta(\mathbf{R}), \theta(\mathbf{R} + \mathbf{a})$ 之差成正比. 如本来平行, 则和此二角差的平方成正比. 令 $\theta(\mathbf{R}) \propto e^{-i\mathbf{k} \cdot \mathbf{R}}$, 则

$$UH_{\mathbf{R}}U^{-1} - H_{\mathbf{R}} = O(ka) \quad \text{或} \quad O((ka)^2). \quad (41)$$

总之, 只要 H 不牵涉到相距太远的自旋, (41) 就成立, 不一定要 (40) 这个特别形式. 因此, 我们的结论是

$$\begin{aligned} \langle \Psi | (H - E_0) | \Psi \rangle &= O(ka) \quad \text{或} \quad O(k^2 a^2), \\ \Psi &= (U - 1)\Phi_0, \\ \theta(\mathbf{R}) &\propto e^{-i\mathbf{k} \cdot \mathbf{R}}. \end{aligned} \quad (42)$$

k 可以有很多不同的值, 都很小. 因此, 我们证明了低能激发态的存在. 以上分析不限于自旋模型, 不用量子力学也行得通. 不过, 只限于短程作用的模型. (41), (42) 特别把 ka 的 a 写出来. 如果 a 太大则结果就不保险, 如果是静电作用之类的模型, 则这结果就靠不住了.

以上这结果, 即连续对称被冻时必有低能激发态, (能量 $\rightarrow 0$, 当 $k \rightarrow 0$) 和量子场论中的“戈德斯通定理”[‡]有相同的含义.

晶体振动的声子, 上例中的磁子, 都是此类低能量激发态. 液面的表面张力波亦然. 我们统称之为“柔软波”, 其波长长而频率低也.

如果序参量的分布是混乱的, 如非晶磁体中的自旋, 则柔软波不一定是有一定波长的波. 它的形状要由自旋的分布, 相互作用能来决定. 不规则的固体中, 柔软波也不一定是平面波, 这一方面的数据还非常少.

因此, 无论是量子模型或古典模型, 凡是有连续性对称被冻, 则有柔软波, 这些柔软波如果幅度大, 则使凝结动摇. 在一维, 二维空间, 长波的份量较重, 柔软波就把序参量软化了. 以上的一连串例子, 都是反映此一现象.

这种凝结软化现象, 已有严密的分析. 结论和以上的诸例大致一样, 但说明了许多重要细节[§].

5. 凝结的缺陷

以上讨论的起伏, 是序参量的平缓起伏. 这是连续性对称的特征. 改变平缓, 则能量低. 起伏的幅度, 可以看成空间的平缓、连续的函数. XY 模型的 $\theta(\mathbf{R})$, 晶体模型的 $u(\mathbf{R})$ 都可看成 \mathbf{R} 的连续函数, 因为它们在短距离内 (格子大小) 的改变十分小.

但是格子虽小, 到底不是无限小. 不连续的情形仍可以发生, 不但发生, 而且至为重要. 现在就用二维空间的 XY 模型为例, 说明一下.

图 1 示两个平面上的向量分布. 令 $\alpha(\mathbf{R})$ 为 \mathbf{R} 和水平方向的夹角, \mathbf{R} 是以图中的 0 为原点来计量. 图中每一黑点是一个 \mathbf{R} , 图 1 向量分布是

[‡]所谓 Goldstone 定理.

[§]见 Hohenberg (1967), Mermin and Wagner(1966). 这些文章很容易读, 是很重要的文献.

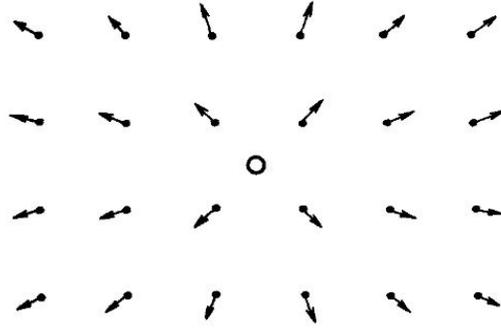


图 1a, 图 (a) 中 $\theta(\mathbf{R})$ 的“等值线”, 注意 π 和 $-\pi$ 的意义一样.

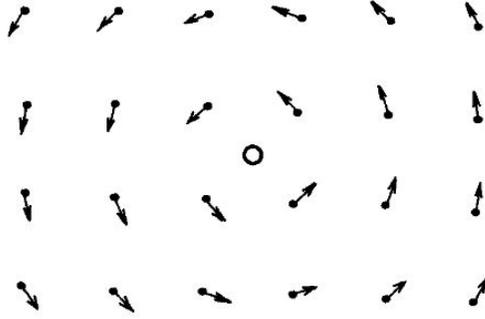


图 1b, 等值线同图 1a, 只是各值要加 $\pi/2$.

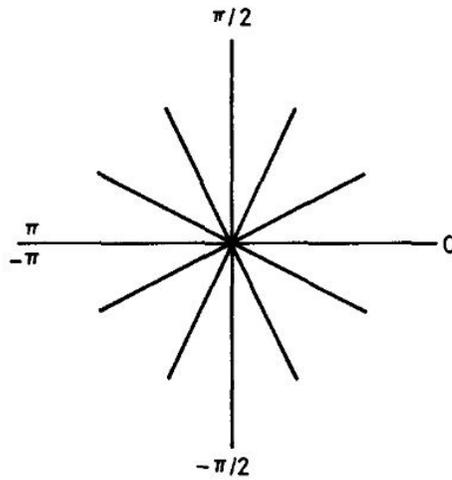


图 1c

$$\begin{aligned}\theta(\mathbf{R}) &= \alpha(\mathbf{R}), \quad \text{图 1(a), (c),} \\ \theta(\mathbf{R}) &= \alpha(\mathbf{R}) + \frac{\pi}{2}, \quad \text{图 1(b), (c).}\end{aligned}\tag{43}$$

在距 0 远的地方, $R \gg a$, $\theta(\mathbf{R})$ 的确是一个平缓的函数. 但是在 0 附近, 它就不是了, 在 0 周围的四个向量, 其邻居的 θ 角差为 $\pi/2$. 如果 $R \gg a$, 则

$$\begin{aligned}\frac{1}{2} \sum_{\mathbf{a}} (\theta(\mathbf{R}) - \theta(\mathbf{R} + \mathbf{a}))^2 \\ \approx a^2 (\nabla\theta)^2 = a^2 \frac{1}{R^2}.\end{aligned}\tag{44}$$

如果把 (44) 代入 (3), 即得

$$\begin{aligned}H' &= \frac{J}{2} a^2 \sum_{\mathbf{R}} (\nabla\theta)^2 \\ &= \frac{J}{2} \int d^2\mathbf{R} (\nabla\theta)^2 \approx \pi J \ln \left(\frac{L}{a'} \right).\end{aligned}\tag{45}$$

πL^2 是此物体的面积. 因为在 $R \approx a$ 时, (44) 不正确. 因此, a' 和 a 略有不同. 据仔细分析的结果, 如果用

$$\frac{a}{a'} = 2\sqrt{2}e^c, \quad c = 0.577,\tag{46}$$

则对正方格子来说, (45) 是个很好的近似[¶].

读者一定认得, (43) 是

$$\nabla^2\theta = 0\tag{47}$$

之解, 只要不太接近原点. 这也是说若非原点附近的“缺陷”, (43) 使能量

$$H \approx \frac{1}{2} J \int d^2\mathbf{r} (\nabla\theta)^2\tag{48}$$

成为一极小值. 因此, 除了原点左近的缺陷之外, (43) 是相当的“完美”.

注意, 绕此缺陷一周, (43) 的 $\theta(\mathbf{R})$ 改变了 2π

$$\oint \nabla\theta \cdot d\mathbf{R} = 2\pi.\tag{49}$$

这些是流体力学和静电学中常见的东西. 如果 $\theta(\mathbf{R})$ 看成流体力学中的“速度位”, 即 $\nabla\theta$ 为流速, 则图 1 的 θ 形象是一个“漩涡”. 二维空间的静电学中, 一

[¶]见 Kosterlitz (1974), P.1049.

点电荷产生的电场是和 $1/R$ 成正比. 因此, (44) 也可以看成一个电荷的结果. 因此, 这缺陷的作用很容易分析.

当然, (45) 指出, 一个漩涡或电荷的能量随 $\ln L$ 增大. 但如有一双相反的电荷载或漩涡, 则能量只和它们之间的距离有关. (见图 2)

图 2 中的 $\theta(\mathbf{R})$ 是

$$\theta(\mathbf{R}) = \alpha(\mathbf{R}) - \alpha(\mathbf{R} - \mathbf{R}'). \quad (50)$$

一对漩涡的能量是

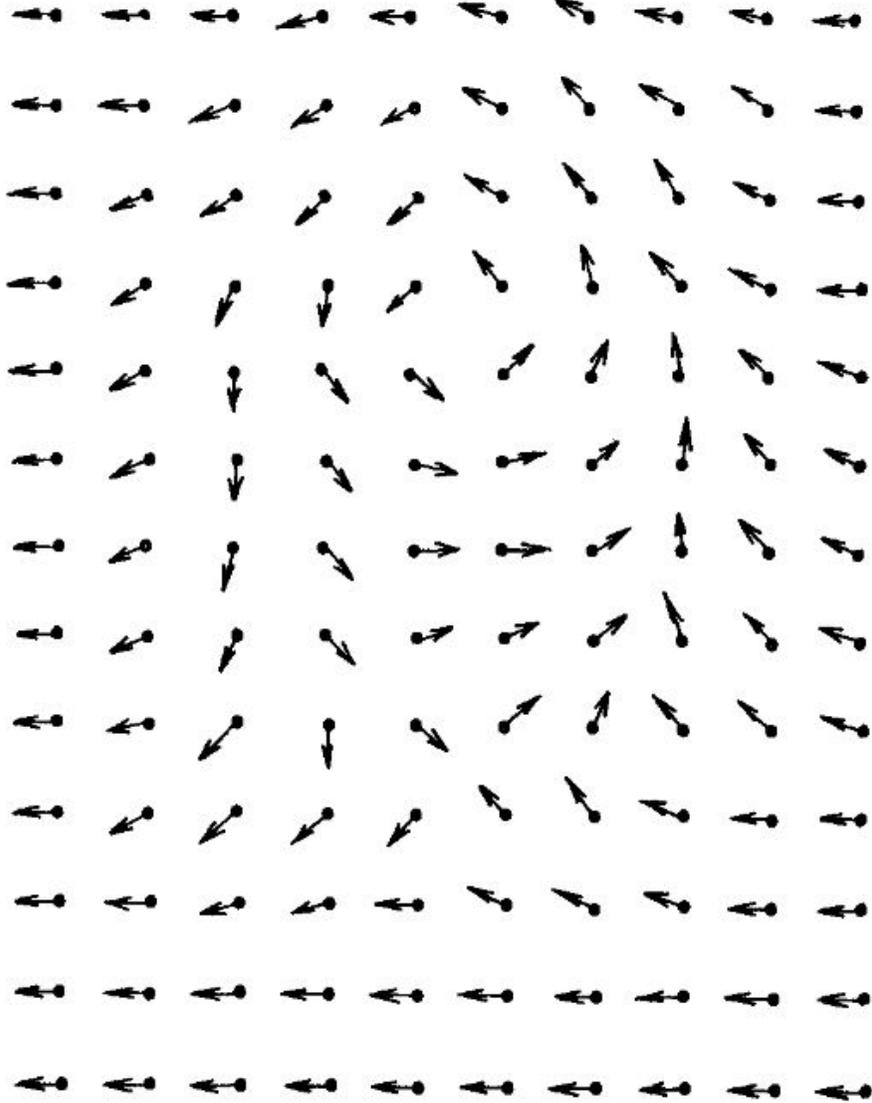


图 2

$$H' = 2\pi J \ln \left(\frac{R'}{a'} \right). \quad (51)$$

如果有许多漩涡, 则

$$H' = -\frac{1}{2} \sum_{i,j} q_i q_j \ln \left(\frac{|\mathbf{R}_i - \mathbf{R}_j|}{a'} \right). \quad (52)$$

\mathbf{R}_i 为漩涡眼的位置, 或电荷的位置.

$$\frac{1}{\sqrt{2\pi J}} q_i = \frac{1}{2\pi} \oint \nabla\theta \cdot d\mathbf{R}, \quad (53)$$

是绕着 \mathbf{R}_i 的积分, 是漩涡强度, 或电荷量. 现在把 (52) 的意义总结如下:

在每个 \mathbf{R}_i 的地方, 挖一个半径为 a' 的小洞, 在这挖了洞的平面上, 我们解 $\nabla^2\theta = 0$. 这个方程式的解, 是最平缓函数, 解出后代入 (48) 计算能量, 即得 (52). 每个小洞都有个条件, 即 (53) 积分必是一预定的整数. 还有一条件是电荷净值为零:

$$\sum_i q_i = 0. \quad (54)$$

(52) 是一群缺陷的能量, 这能量不是无限大, 因为 a' 不是零, 即格子的大小不是零.

$\theta(\mathbf{R})$ 既是电位, 而电荷对电位的影响在第十九章已讨论过. 在低温时, 正负电荷会配成双, 造成一个有限的介电系数. $\theta(\mathbf{R})$ 的起伏、相干, 大致仍遵照本章第 1 节的结果. 如果温度太高, 则电荷分散, $\theta(\mathbf{R})$ 的起伏的相干距离就成为有限, 凝结就瓦解了. 瓦解的温度, 照第十九章的粗略估计, 是 $T_0 \approx q^2/4 = \pi J/2$ †.

由此可见, 缺陷可以导致凝结的瓦解. 注意, 在温度高时, 相邻的向量就不再有相近的方向, 因此, “缺陷” 的观念, 就没有什么用处. 无一处 “完美”, 谈什么 “缺陷” 呢? 以上的电荷模型, 是假定缺陷很少, $\theta(\mathbf{R})$ 是一个大致完美的平缓函数, 只有少数地方挖了洞. 配双的电荷, 分离的电荷, 都建立在这大致完美的假设上. 不过, 电荷由配双而分离的温度 $T_0 \approx \pi J/2$, 已是相当高的温度, 即和 J 相若. 在这样高的温度下, 这大致完美的假设, 即缺陷密度小的假设, 是否成立? 如果用上一章的平均场解法去求临界温度, 答案是 $T_c = 2J$. (这非常简单, 读者自证之.) 这和 $T_0 = \pi J/2$ 差不了多少, 这真是令人担心了. 不过, 大概也是运气, 在 $T = T_0$ 时, 电荷的密度仍很低. 大致说来, 每个电荷的能量大概是 $= q^2 \ln l/a$, 平均密度是 $(a/l)^2$, l 是电荷间的平均距离. 因此,

$$\left(\frac{a}{l}\right)^2 \sim e^{-\epsilon/T} \sim \left(\frac{a}{l}\right)^{q^2/T}. \quad (55)$$

因此, 若要 $a/l \geq 1$, 必须要有 $T \geq q^2 = 4T_0$, 在 $T = T_0$ 附近, 电荷密度该不很大. 不过, 这粗略的估计并不足以完全消除以上提出的隐忧. 后来有人用数值演象, 才有更好一点的佐证.

†仔细的分析, 见 Kosterlitz (1974).

以上是二维空间 XY 模型的讨论, 若是三维空间, 则平面上的“小洞”就成了“细管子”了, 即缺陷是“漩涡线”. “电荷”不再合用, 必须改成线状电流, $\nabla\theta$ 指向磁场方向. 漩涡线可自绕成一圈, 成一漩涡环, 或者线的两端站在物体的边界上. 这样的缺陷不只牵涉到一两个格子, 而是一大串格子, 能量因此很高. 所以, 在平衡时不该出现. 但是, 这漩涡线是一暂稳的东西, 一旦形成, 就不易消失. 要除去一个漩涡, 须把很多单元重新排列 (二维空间的亦然.). 这些在下章讲超流性时会再提到.

晶体中的缺陷就复杂得多了, 最简单的是“位错”(见图 3), 图中的晶体上半多了一行, 其他缺陷种类繁多, 在此不谈. 读者可参考材料科学的书籍. 缺陷的理论, 亦见于其他模型. 文献在近几年来胜于雨后之春笋**. 读者自己去探吧!

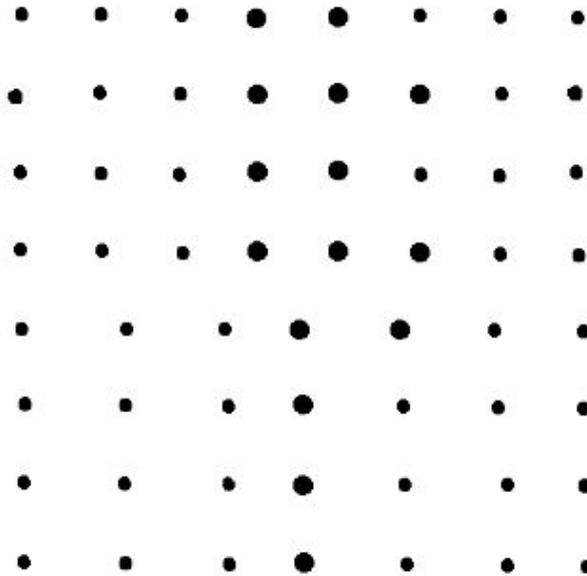


图 3

讨论问题二十九

1. 试解一维空间的 XY 模型, 直接用 (1), 不要用 (3).

(A) 计算热位能, 熵, 热容率, 并作图, 横轴为 T .

**例如 Young (1979), Nelson and Halperin (1980).

(B) 计算相干函数 $C(R)$, (定义见 (11)).

(C) 将以上结果和 (3) 之解作一比较.

(D) 如果相邻两向量之间的夹角大于直角, 我们就说此二单元之间有一“缺陷”. 且把这当成此模型中缺陷的定义. 试计算缺陷密度.

2. 上题中, 在定义熵时, 是否遭遇到困难? 如何解决?

注意, s_i 是连续变量. 凡是连续变量, 都需要用一点量子力学来帮忙定义熵.

3. 复习第二十七章. 试用平均场学解 XY 模型 (1). 求临界温度 T_c .

4. 将 (1) 加上一磁场项

$$- \sum_i \mathbf{h}_i \cdot \mathbf{s}_i. \quad (56)$$

将第二十七章问题 2, 8 的分析推广到 XY 模型来. 假定 h_i 很小.

(A) 推广 (27.47). 考虑 $T \ll T_c$, 及 T 接近 T_c 两特殊情形. 注意, 在 $T < T_c$ 时, 序参量指向一定方向.

(B) 求相干函数 $C(R)$, $T \ll T_c$, 或 T 接近 T_c . 和第 1 节的分析比较一下.

5. 将以上三题的分析用到 3 维向量模型, 即 \mathbf{s}_i 为三维空间向量. 位置空间的维数 d 另定. 此所谓“古典海森堡”模型.

6. 波色粒子气体, 规范变换.

(A) 复习讨论问题三, 第三章的讨论略去了作用能量. 现在加一项作用能

$$H' = \frac{1}{2} \sum_{i,j} u(\mathbf{r}_i - \mathbf{r}_j), \quad (57)$$

$\mathbf{r}_i, i = 1, 2, \dots, N$ 是粒子的位置.

(B) 令 $\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$ 为此物体的一个波函数. 定义“规范变换”为

$$\Psi' = e^{-i\alpha N} \Psi. \quad (58)$$

α 为“规范转角”.

试证哈密顿算符在规范变换下不变, 即

$$e^{i\alpha N} H e^{-i\alpha N} = H. \quad (59)$$

此地 N 是看成一算符. 此式子说粒子数是守恒的.

(C) 现在暂且略去 H' , 只考虑动能:

$$H_0 = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} N_{\mathbf{k}}, \quad \epsilon_{\mathbf{k}} = k^2/2m. \quad (60)$$

$N_{\mathbf{k}}$ 是在位子 \mathbf{k} 中的粒子数.

这可以看成一群简谐振动单元, 因为 $N_{\mathbf{k}}$ 是正整数, 因此 \mathbf{k} 位子可以看成是一个单摆, 能量可写成

$$\epsilon_{\mathbf{k}} N_{\mathbf{k}} = \frac{1}{2} \epsilon_{\mathbf{k}} (p^2 + q^2 - 1). \quad (61)$$

p, q 是单摆的“动量”和“坐标”. 状态由 (q, p) 平面上的位置代表.

试证规范变换量是 (q, p) 平面上的坐标转换:

$$e^{i\alpha N_{\mathbf{k}}}(q, p)e^{-i\alpha N_{\mathbf{k}}} = (q \cos \alpha + p \sin \alpha, p \cos \alpha - q \sin \alpha). \quad (62)$$

写得更简单些:

$$e^{i\alpha N_{\mathbf{k}}} a_{\mathbf{k}} e^{-i\alpha N_{\mathbf{k}}} = a_{\mathbf{k}} e^{-i\alpha} \quad (63)$$

$$a_{\mathbf{k}} \equiv \frac{1}{\sqrt{2}}(q + ip), \quad N_{\mathbf{k}} = a_{\mathbf{k}}^* a_{\mathbf{k}}. \quad (64)$$

$a_{\mathbf{k}}$ 是所谓“消去符”, $a_{\mathbf{k}}^*$ 是“产生符”.

(D) 问题三, 8 指出, 在低温, N_0 很大. 因此, $\mathbf{k} = 0$ 位子可以用古典力学来讨论. $\langle a_0 \rangle$ 是一个 $\sqrt{N_0}$ 量级的大数, 这可以看成一凝结现象, 即“波色-爱因斯坦”凝结现象.

试把第 4 节的柔软波分析, 用到这模型.

(E) 假定作用能 H' 不略去, 但凝结仍出现, 即 $N_0 = O(N)$, 再用第 4 节的分析讨论低能量的激发态.

波色粒子的问题, 在下一章还会再讨论.

7. 讨论量子反铁磁性模型, 即把 (24) 的 J 改成 $-J$. 这模型不像铁磁性一般简单, 读者可参考固态物理书籍.

8. (A) 计算 $W(k)$, $d = 1, 2$. (见 (17), (18)), 并和 (21) 比较.

(B) 求任一原子的位置分布

$$\langle \delta(\mathbf{r} - \mathbf{u}(\mathbf{R})) \rangle.$$

(C) 假定一铜膜 (只有一层铜原子) 的大小为 $L = 1$ 公里, 估计 $\langle u^2 \rangle$ 值.

9. 假想 N 个原子连成一维空间的长链中有 N' 个杂原子, 其在链中分布为不规则, $N \gg N'$. 哈密顿函数如第八章, 问题 12, 只是杂原子的质量 m' 比其他原子的 m 要大些.

(A) 试计算杂质对热力性质的影响.

(B) 试讨论杂质对各原子振幅的影响.

10. 某 d 维空间晶体中; 有 N' 个原子被杂原子代替. 杂原子分布为不规则. 上题即是 $d = 1$ 的情形. 讨论杂原子对第 2 节各项结论的影响.

11. 复习基本的平面流体力学和静电学. 特别留意复变量理论的应用.

(A) 令 $z = x + iy$, $f(z)$ 为 z 的函数. 如果在某处 $f(z)$ 为 z 的“解析函数”, 即 df/dz 存在, 试证在该处

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) f = 0. \quad (65)$$

令 $f = \phi + i\psi$. 则 ϕ, ψ 各适合 $\nabla^2 \phi = 0, \nabla^2 \psi = 0$.

(B) 试证 ψ 的等值线和 ϕ 的等值线相垂直, 即 $\nabla \phi \cdot \nabla \psi = 0$.

(C) 令 $f(z) = \ln z$. 讨论如何用此函数分析漩涡和电荷点.

第三十章 超流性原理

超流性是指无阻力、不休止地流动, 不需外力推动. 这种现象可在低温的 He^4 液体 ($T < 2.17\text{K}$) 发生, 也可在 He^3 液体 (T 低于 10^{-3}K 左右) 内发生. 许多金属电子在低温时也有此现象, 温度的限制视金属而定, 是为“超导电性”现象. He^3 的超流性极其复杂. 由于 He^3 和电子都是费米粒子, He^3 的理论和超导体略有相似处. He^4 是波色粒子, 情形简单些. 所有的超流现象, 都有一些基本特性. 本章并不准备把 He 液和超导体的资料作任何有系统的介绍. 本章的目的只求对超流性的最基本原理作一分析. 超流性是一个颇富神秘感的性质. 在此, 我们用最简单的模型去解释它.

超流性牵涉到两个现象, 第一是凝结, 第二是暂稳态. 前者为后者的必要条件. 流动的状态是一暂稳态, 在短时间内 (短时间可以长达许多年) 就是一个平衡态, 热位能为一极小值, 流动不会因扰动而停止 (扰动不能太大). 我们讨论暂稳态的目的是了解使流量减少的过程. 流动态的热位能是比静止态要高, 非绝对极小. 但如果要从流动态变到静止态, 热位能必先增加才行, 即必先越过一屏障 (见第二十一章对暂稳态的讨论). 我们必须了解这屏障是什么, 才算是对超流性有初步的了解.

如果没有凝结现象, 流动的稳定性就不会发生. 超流性的神秘, 就在这序参量. 到底是什么东西凝结了? 这一点, 我们用一个磁铁模型解释. 先建立一个格子气体模型. 这模型的内容和上章的自旋 $\frac{1}{2}$ 磁铁模型一样. 由于我们对磁铁模型有相当的认识, 在此可以用它来了解格子气体的性质, 一如在第二十七章用易形磁铁来分析气液共存. 此地的流体模型是量子力学中的格子气体, 它是量子力学中自旋 $\frac{1}{2}$ 模型的化身. 序参量即是磁矩. 没有什么神秘.

建立了序参量之后, 就可以讨论超流性. 我们先把模型简化成一个平面向量模型, 即上章的 XY 模型. 再讨论一环形体积中的序参量. “绕圈数”是一个重要的观念, 它代表凝结向量在沿环走一周后所改变的角度除以 2π , 即转了几圈. 这个绕圈数 n 是个守恒量, 如果不考虑凝结缺陷. 超流速是和 n 成

成正比. 缺陷即是漩涡. 本章后半部的内容主要是分析漩涡如何改变 n . 由此可以了解超流性的大致原理. 详细的分析主要以二维空间模型为主 (即 He^4 附在一光滑表面上的模型). 第十九章及上一章的经验, 在此都很有用. 理论上, 二维空间比较简单, 但实验比较困难. 实验和理论都是近几年的事, 文献非常多, 但本章的目的, 并不是对这些文献作任何介绍. 许多有趣的其他问题, 如各种声波, 低温的热力性质等, 都没有提到. 超导电性是一门极其有趣且重要的科学, 在此也不讨论. 它的基本原理和 He^4 颇有些相同的地方, 虽然电子是费米粒子, 且带电. 还有一个重要题目, 就是在临界点附近的超流性及其他性质. 我们也不讨论. 因此本章的内容可以说是不能再少了, 只讨论凝结和流动的暂稳性. 对这两个重点, 我们作比较仔细的分析.

1. 量子格子气体

现在来凑一个格子气体模型. 令 $n(\mathbf{R})$ 为在格子 \mathbf{R} 的粒子数, 其值为 1 或 0, 令*

$$n(\mathbf{R}) - \frac{1}{2} = s_z(\mathbf{R}). \quad (1)$$

s_z 即是一自旋 $\frac{1}{2}$ 的 z 分量. 格子气体模型因此和自旋 $\frac{1}{2}$ 模型有相同的内容.

粒子数为零的状态, 即真空, 即是所有自旋指向 $-z$ 的状态, 称之为 Φ_0 ,

$$\sum_{\mathbf{R}} n(\mathbf{R}) \Phi_0 = 0. \quad (2)$$

$s_+(\mathbf{R})\Phi_0$ 为单粒子状态, 粒子在 \mathbf{R} . 如粒子为一平面波, 则其波函数为

$$\psi_{\mathbf{k}} = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} s_+(\mathbf{R}) \Phi_0. \quad (3)$$

$s_+(\mathbf{R})$ 把 $s_z(\mathbf{R})$ 加 1, 即加一粒子于 \mathbf{R} . 因此, 上章第 3 节的“磁子”即此地的粒子. 现在要写出一个合理的全能算符 H . 先问动能是什么? 一个粒子的动能该是 $k^2/2m$. 不过现在粒子是限制在分立的格子上, 要找个完全合理的动能算符不大容易. 现在找个差不多的, 令

$$K \equiv \frac{1}{4ma^2} \sum_{\mathbf{R}} \sum_{\mathbf{a}} [s_+(\mathbf{R}) - s_+(\mathbf{R} + \mathbf{a})] [s_-(\mathbf{R}) - s_-(\mathbf{R} + \mathbf{a})]. \quad (4)$$

*这个 n 不是以后要讲的“绕圈数”.

如果 $ka \ll 1$, 则施 K 于 (3) 得

$$K\psi_{\mathbf{k}} \approx \frac{k^2}{2m}\psi_{\mathbf{k}}. \quad (5)$$

α 是指向格子邻居的向量, 假设格子为立方. (4) 式亦可以写成

$$\begin{aligned} K = & -\frac{1}{2}J \sum_{\mathbf{R}} \sum_{\mathbf{a}} \mathbf{s}(\mathbf{R}) \cdot \mathbf{s}(\mathbf{R} + \mathbf{a}) \\ & + \frac{1}{2}J \sum_{\mathbf{R}} \sum_{\mathbf{a}} s_z(\mathbf{R})s_z(\mathbf{R} + \mathbf{a}) + JN'd, \end{aligned} \quad (6)$$

$$J \equiv \frac{1}{ma^2},$$

N' 是粒子总数. 因此, K 很像磁铁模型中的交换作用. 我们只考虑粒子数不变的情形, 即 N' 为常数. 现在我们干脆就讨论一个模型

$$\begin{aligned} H &= K + U \\ &= -\frac{1}{2}J \sum_{\mathbf{R}} \sum_{\mathbf{a}} \mathbf{s}(\mathbf{R}) \cdot \mathbf{s}(\mathbf{R} + \mathbf{a}) + \text{常数}, \\ U &\equiv -\frac{1}{2}J \sum_{\mathbf{R}} \sum_{\mathbf{a}} n(\mathbf{R})n(\mathbf{R} + \mathbf{a}). \end{aligned} \quad (7)$$

也就是说, 这气体有动能 K , 位能 U , 位能是相邻粒子的吸引作用. 注意, $n = s_z + \frac{1}{2}$. 因此 U 就是 (6) 的右第二项加一常数 (N' 为常数). 因此, 这模型和铁磁模型完全一样了. 最后, 我们需要一个动量算符. 令

$$\mathbf{P} = -\frac{i}{2} \sum_{\mathbf{R}} \sum_{\mathbf{a}} \frac{1}{2} [s_+(\mathbf{R}) + s_+(\mathbf{R} + \mathbf{a})] \frac{1}{a} [s_-(\mathbf{R} + \mathbf{a}) - s_-(\mathbf{R})] \frac{\mathbf{a}}{a}. \quad (8)$$

这算符在 $ka \ll 1$ 时, 施于 (3) 的 $\psi_{\mathbf{k}}$ 得

$$\mathbf{P}\psi_{\mathbf{k}} \approx \mathbf{k}\psi_{\mathbf{k}}. \quad (9)$$

\mathbf{P} 大致可以算作动量算符. 在章尾的问题中, 会讨论一些 K 和 \mathbf{P} 的细节.

提出这个模型的目的之一, 是指出量子力学的几个重要特点.

易形磁铁模型不是一个自旋 $\frac{1}{2}$ 的向量模型, 易形模型的运动单元有二值, 自旋 $\frac{1}{2}$ 也有二值, 但两者大不相同. 自旋 $\frac{1}{2}$ 是一向量. 虽然每个自旋 $\frac{1}{2}$ 只有两个状态, 但它仍是一个向量. 从两个状态弄出一个三维空间里的向量, 是量

子力学的杰作. 没有量子力学是办不到的. 如果许多个自旋 $\frac{1}{2}$ 放在一起, 指向同一方向, 则这总体的向量性质, 可以用古典的观念去分析.

易形格子气体模型是易形磁铁模型的化身, 而量子格子气体模型 (7) 则是自旋 $\frac{1}{2}$ 模型的化身. 易形气体是古典气体模型, 其状态是用粒子分布来指定, 一如易形铁磁中各单元之值. 量子气体模型的状态也以粒子分布指定, 一如自旋 $\frac{1}{2}$ 模型中各 $s_z(\mathbf{R})$ 之值. 但除了 s_z 之外, 还有 s_x 和 s_y . 也就是说, 在量子力学中, 所有的算符, 都是运动变量, 如果有两个状态, $n = 1, 0$, 则除了 n 之外, 一定还有两个其他的变量, s_x, s_y .

自旋 $\frac{1}{2}$ 模型可以有凝结态, 总自旋 \mathbf{S} 可以凝结在任一方向. S_x, S_y, S_z 都可以不等于零. 同理, 量子气体模型也可以有凝结态, 也可以有不等于零的 S_x, S_y . ($S_z = N' - \frac{1}{2}N$, N' 为粒子总粒子数, N 为格子总数). 这不同于零的 S_x, S_y , 是古典气体中没有的东西. 但从磁铁模型来看, 是理所当然, 不足为奇的事. S_x, S_y 的凝结, 是超流性的必要条件.

模型 (7) 可以看成是一个 He^4 群体的锥形. 格子大小 a 可以看成原子的大小. 每格至多一原子的限制, 代表原子间的短距离斥力. U 是特别的吸引力, 其目的是凑成 H 的简单形式, 不足以写实, 亦无伤也.

注意, 这是一个波色粒子模型. 每格子至多一原子的限制, 不是指粒子不合群, 而是指粒子为“硬球”型. 粒子波函数是对称的.

2. 基态和低温模型

现在来分析模型 (7), 也就是一个自旋 $\frac{1}{2}$ 磁铁模型, 和上一章的 (24) 一样. 唯一新加的条件是

$$S_z = N' - \frac{N}{2} \quad (10)$$

是个不变量, 因为 N' 是总粒子数, 是不变量.

$$N = \frac{V}{a^3} \quad (11)$$

是格子数, V 是总体积, a^3 是格子体积. 先看 H 的基态, 我们早知道基态是所有自旋平行的状态, 自旋的方向受到 (10) 的限制, 限在一个环上, 见图 1. \mathbf{S} 的总长是 $\frac{1}{2}N$, 即 N 个自旋 $\frac{1}{2}$ 排起来. \mathbf{S} 和 $-z$ 方向成 θ 角. 这角度由 S_z 及 $|\mathbf{S}|$ 之值决定. ϕ 是沿环角, 任何值都可以.

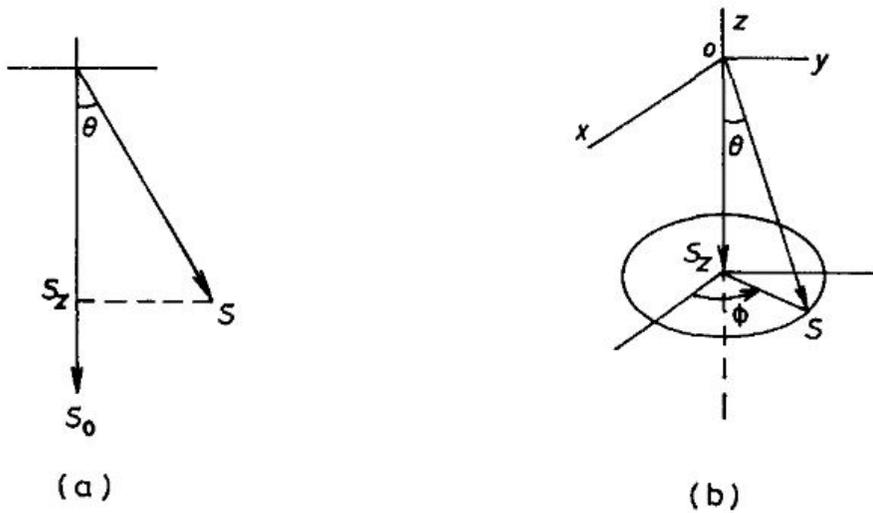


图 1

这基态的波函数十分容易求. 图 1(a) 中的 S_0 指的是 Φ_0 , 即无粒子态, 也就是所有自旋向下. 要得 N' 粒子的基态, 只要把 S_0 转个 θ 角就行了:

$$\Psi = e^{-i\phi S_z} e^{-i\theta S_y} \Phi_0, \quad (12)$$

即先以 y 方向为轴, 转 θ 角, 再以 z 为轴, 转 ϕ 角, 即得图 1(b) 中 S 的方向, Ψ 即是基态波函数. 这些转动可以立即做出, 因为每个自旋可以分别转, 读者试证:

$$\begin{aligned} \Psi &= \prod_{\mathbf{R}} (u + v s_+(\mathbf{R})) \Phi_0, \\ u &\equiv e^{i\phi/2} \cos \theta/2, \\ v &\equiv -e^{i\phi/2} \sin \theta/2, \end{aligned} \quad (13)$$

$|u|^2$ 是格子 \mathbf{R} 没有粒子的概率, $|v|^2$ 是有一粒子的概率. 是因为 u 和 v 的巧妙组和, 使 S_x, S_y 不等于零, 而角度 ϕ 因此有了意义. 从自旋模型的观点, 这种组和是不足为奇的. 自旋 $\frac{1}{2}$ 指向任一方向的状态, 都可以从“向上”和“向下”两个状态组和而得.

如果将每个自旋的 ϕ 角稍作改变, 且相邻自旋的 ϕ 改变差得少, 则得一低能的激发态 (见上章第 4 节的讨论), 令

$$\Psi' \equiv e^{-i \sum_{\mathbf{R}} \phi(\mathbf{R}) s_z(\mathbf{R})} \Psi. \quad (14)$$

这就是把 $\mathbf{s}(\mathbf{R})$ 转 $\phi(\mathbf{R})$ 角的结果. 如果每个 $\phi(\mathbf{R})$ 都一样, 则 Ψ 只是另一基态 (图 1(b) 环上的另一点).

Ψ' 的能量可以估计如下: 各 $\mathbf{s}(\mathbf{R})$ 被转了之后, $\mathbf{s}(\mathbf{R}) \cdot \mathbf{s}(\mathbf{R} + \mathbf{a})$ 改变多少? 本来各自旋是平行的, 这乘积是 $\frac{1}{4}$, 现在 $\mathbf{s}(\mathbf{R})$ 被转 $\phi(\mathbf{R})$, $\mathbf{s}(\mathbf{R} + \mathbf{a})$ 被转 $\phi(\mathbf{R} + \mathbf{a})$, 这乘积的改变是

$$\frac{1}{4} \sin^2 \theta \cos[\phi(\mathbf{R}) - \phi(\mathbf{R} + \mathbf{a})]. \quad (15)$$

因为转角是以 z 方向为轴, 只有 xy 平面方向的向量受转动影响. 所以 (15) 中有 $\sin^2 \theta$. 因此, 总能量的改变是

$$\begin{aligned} \langle \Psi' | H | \Psi' \rangle - \langle \Psi | H | \Psi \rangle &\equiv H', \\ H' &= -\frac{J'}{2} \sum_{\mathbf{R}} \sum_{\mathbf{a}} \cos[\phi(\mathbf{R}) - \phi(\mathbf{R} + \mathbf{a})], \\ J' &= \frac{1}{4} J \sin^2 \theta. \end{aligned} \quad (16)$$

θ 这个角是由 $S_z = N' - \frac{1}{2}N$ 和 $|\mathbf{S}| = N/2$ 决定:

$$\begin{aligned} \sin^2 \theta &= 1 - \left(\frac{2N'}{N} - 1 \right)^2, \\ &= 4n(1-n) \\ n &\equiv \frac{N'}{N}. \end{aligned} \quad (17)$$

因此

$$J' = Jn(1-n). \quad (18)$$

(16) 式恰好是上章讲的平面向量模型, 或“XY 模型”. ϕ 角的振动即是上章谈的一种“柔软波”. 各自旋的排列是一凝固现象, 这模型有以 Z 方向为轴的旋转对称性. 此模型的低温性质可以用 (16) 的 H' 作哈密顿函数分析. 注意, (16) 不再是一量子模型, 各变数 $\phi(\mathbf{R})$ 是看成普通的古典力学变数. 当然 H' 这模型是由量子模型 (7) 得来. 它是 (7) 的低温模型. 它只考虑了各自旋的 ϕ 角运动.

H' 也可以写成比较好看一点的形式. 因为 $J \equiv 1/ma^2$, 且

$$\begin{aligned} \sum_{\mathbf{R}} &\approx \frac{1}{a^3} \int d^3 \mathbf{R} \\ \cos[\phi(\mathbf{R}) - \phi(\mathbf{R} + \mathbf{a})] &\approx 1 - \frac{1}{2} (\mathbf{a} \cdot \nabla \phi)^2. \end{aligned} \quad (19)$$

因此,

$$H' \approx \frac{\rho}{2m} \int d^3\mathbf{R} (\nabla\phi)^2 + \text{常数}, \quad (20)$$

$$\rho \equiv \frac{1}{a^3} n(1-n) = \frac{N'}{V} \left(1 - \frac{N'}{V} a^3\right). \quad (21)$$

如格子很小, 则 ρ 是粒子密度. 读者别忘了 (18) 是一个磁铁模型的化身.

3. 流动状态, 绕圈数

现在假想一个长为 L 的脚踏车胎形圆环, z 坐标沿环, 截面积为 A , 我们的格子气体放在里面. $V = AL$, 能量是 H' , 用 (20) 比较简单.

如果 ϕ 为常数, 则 H' 为极小即基态. 现在来求 H' 的其他极小值. 令 H' 对 $\phi(\mathbf{R})$ 的微分为零, 得

$$\nabla^2\phi = 0. \quad (22)$$

$\phi = \text{常数}$ 是一个解. 但是,

$$\phi = \phi_n \equiv \frac{2\pi n}{L} z, \quad n = \pm 1, \pm 2, \dots, \quad (23)$$

也是 (22) 的解, 因为 $\phi + 2n\pi$ 的意义和 ϕ 完全一样, ϕ 是一个角度. z 坐标沿环, $z = 0$ 和 $z = L$ 是同一点. 故 $\phi(x, y, L) - \phi(x, y, 0)$ 必须是 2π 的整数倍. 把 (23) 代入 (20), 得

$$H' = \frac{\rho V}{2m} \left(\frac{2\pi n}{L}\right)^2 \equiv E_n. \quad (24)$$

E_n 是一个 H' 的极小值, 不是极大值, 因为如把 $\phi = \phi' + \phi_n$ 代入 (20) (ϕ' 很小) 即得 $E_n + \text{正数}$.

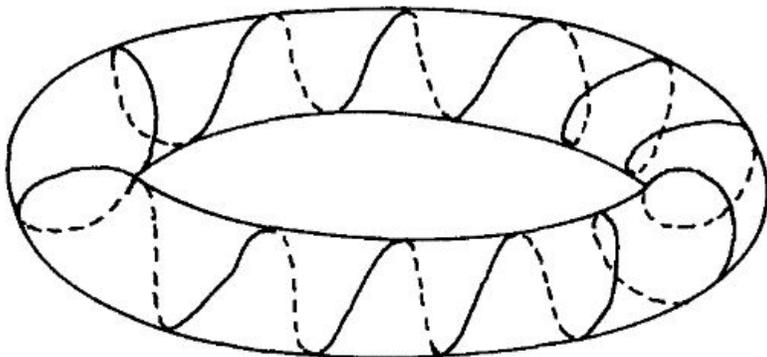


图 2

我们称 n 为“绕圈数”，这个名字的意义由图 2 可见。图 2 示一圆环，截面为半径为 $\frac{1}{2} \sin \theta = s$ 的圆， z 坐标沿环，环长为 L 。($s_{\perp} \cos \phi, s_{\perp} \sin \phi$) 就成了绕在环上的一条曲线了。($s_{\perp} \cos \phi, s_{\perp} \sin \phi$) 即 (s_x, s_y)。因此图 2 画的，是 (23) 代表的序参量，沿着环的分布，是这曲线所绕的圈数

$$\text{绕圈数} = \frac{1}{2\pi} \int_0^L dz \frac{\partial}{\partial z} \phi(x, y, z). \quad (25)$$

照现在的模型，这 ϕ_n 代表的状态是稳定的。绕圈数 n 是改不了的。由图 2 可见，一条绕在环上的线如不能切断再接起来，是无法改变绕圈数的。也就是说， n 是一个守恒量。不管柔软波如何振动，只要 ϕ 是连续函数，它的绕圈数变不了。虽然 $E_n > 0$ ， ϕ_n 代表一个稳定状态。在温度 T 非零时，热位能 F_n 可以由

$$H' = E_n + \frac{\rho}{2m} \int d^3 \mathbf{R} (\nabla \phi')^2, \quad (26)$$

$$\phi = \phi_n + \phi',$$

来算，答案显然是 $F_n = E_n + F_0$ 。这 F_0 是以 $n = 0$ 为基态的热位能。 ϕ_n 这状态，是流动的。我们需要用 (8) 来计算总动量。现在

$$s_{\pm}(\mathbf{R}) = \frac{1}{2} \sin \theta e^{\pm i\phi(\mathbf{R})} \quad (27)$$

(见图 1(b))，注意，(14) 的 Ψ' 可写成

$$\Psi' = \prod_{\mathbf{R}} [u(\mathbf{R}) + v(\mathbf{R})s_+(\mathbf{R})] \Phi_0, \quad (28)$$

其中 u, v 是把 (13) 的 u, v 中的 ϕ 以 $\phi(\mathbf{R})$ 代，(27) 是指 $s_{\pm}(\mathbf{R})$ 对 Ψ' 的平均值，即

$$s_+ = v^* u, \quad (29)$$

$$s_- = u^* v.$$

(8) 可以简化成

$$\mathbf{P} \approx \rho \int d^3 \mathbf{R} \nabla \phi, \quad (30)$$

以上用到

$$\frac{1}{4a^3} \sin^2 \theta = \rho, \quad (31)$$

见 (17), (19). 因此, 将 ϕ_n 代入 (30), 得沿环的动量

$$\begin{aligned} P_z &= \rho V \frac{2\pi n}{L} \equiv m\rho V v_s, \\ v_s &\equiv \frac{2\pi n}{L} \frac{1}{m}. \end{aligned} \quad (32)$$

这是一个稳定的总动量, 这就是超流性, v_s 是所谓“超流速度”. n 是不能变的, n 之所以不能变, 是因为 $\sin\theta$ 非零 (见图 1), 即有 $\mathbf{s}(\mathbf{R})$ 的凝结. 这是一个总体的, 热力学的结果. 加高一点温度, 加一点杂质, 不会影响这稳定的平衡态, 超流体不是流体力学中的粘性趋于零的情形. 粘性趋于零仍不是零, 加一点杂质, 就会影响流速, 使它慢下来. 这样的流动不是暂稳态, 而是不稳态, 超流状态是完全不同的.

以上绕圈数 n 之所以不变, 是由于我们只考虑了平缓的 $\phi(\mathbf{R})$, 视之为 \mathbf{R} 的连续函数. 如果我们把不连续的情形, 即上章第 5 节讲的“缺陷”也考虑进去, 则 n 就不能永久维持不变了. 这模型既是 XY 模型, 则缺陷是漩涡. 在下一节, 我们看漩涡如何使以上的稳定流动变成暂稳, 甚至成为不稳态.

4. 超流的稳定性

上节的结论是: 超流性是靠“绕圈数” n 的守恒来维持. 破坏 n 的守恒, 就破坏了超流性.

现在来看缺陷的影响. 先从一维空间模型说起. (以上的模型是三维空间的, 但要改成一维, 二维的情形很容易). n 之所以守恒, 是靠 $\phi(\mathbf{R})$ 的连续性, 即图 2 的那条曲线是连续的. 但是 \mathbf{R} 不是连续的, 而是一群分立的点. 在一维空间, 这些点排成一串, 像个项链似的, 图 2 中的这条曲线也是一串分立的点, 如果有相邻的 ϕ 值差接近了 π , 则这条曲线可以说是断了. 这切断的概率约是

$$e^{-\epsilon/T}, \quad \epsilon = \frac{\rho a}{2m} \left(\frac{\pi}{a}\right)^2. \quad (33)$$

即 $(\partial\phi/\partial z)^2 dz \sim (\pi/a)^2 a$. 切断的地方约有 N_π 处,

$$N_\pi = N e^{-\epsilon/T} = \frac{L}{a} e^{-\epsilon/T}, \quad (34)$$

假定 $N_\pi \gg 1$, 则切断的情形不停地发生, 当然切断之后又会接回去. 但是断了又接上, 前后角差可以是零, 也可以是 $\pm 2\pi$, 也就是说 n 可以不变, 也可以增加或减少, 即 n 不再是守恒量. 超流性也随之消失.

如果 N_π 很小, 即 T 很小, L/a 亦不大, 则超流态可视为暂稳态, n 的改变很慢. 切断的时间尺度是 τ ,

$$\tau \approx \frac{\tau_0}{N_\pi}, \quad (35)$$

$$\begin{aligned} \frac{dn}{dt} &\approx -\frac{1}{T} \left(\frac{\partial E_n}{\partial n} \right) \frac{1}{\tau} \\ &\approx \frac{n}{\tau_0} \frac{4\pi\rho}{maT} e^{-\epsilon/T}, \end{aligned} \quad (36)$$

$1/\tau_0$ 可以看成一高温时的反应率. 这是一个粗略的估计.

现在看二维空间的情形. 二维空间的缺陷是漩涡, 也可以看成电荷. $\nabla\phi(\mathbf{R})$ 可以看成电场. 从第十九章, 及上一章, 我们知道这些电荷在低温时会两个一对地配成电偶极, 其作用可以用一个介电系数 ϵ 代表. 电场在加了电荷后小了 ϵ 倍:

$$\text{电场} \propto \frac{1}{\epsilon} \left(\frac{\Delta\phi}{L} \right), \quad \Delta\phi = 2\pi n.$$

$$\text{因此, 电场能量} \propto \frac{1}{\epsilon} \left(\frac{\Delta\phi}{L} \right)^2. \quad (37)$$

有缺陷和没缺陷的分别在 ϵ 之值. 如无缺陷则 $\epsilon = 1$, 如有则 $\epsilon > 1$. (24) 可以改写成

$$E_n = \frac{\rho_s V}{2m} \left(\frac{2n\pi}{L} \right)^2, \quad \rho_s \equiv \frac{\rho}{\epsilon}. \quad (38)$$

因此, 在低温, 流量仍可以维持, 只是小了些. (30) 的 ρ 改成 ρ_s 即可.

ρ_s 度量序参量的坚韧程度, 即

$$\text{能量} \propto \rho_s \left(\frac{\Delta\phi}{L} \right)^2. \quad (39)$$

$\Delta\phi/L$ 是 s 被“扭转”或“歪曲”的程度. 我们可称 ρ_s 为“韧度”. 韧度不等于零, 才能维持一个扭转了的序参量.

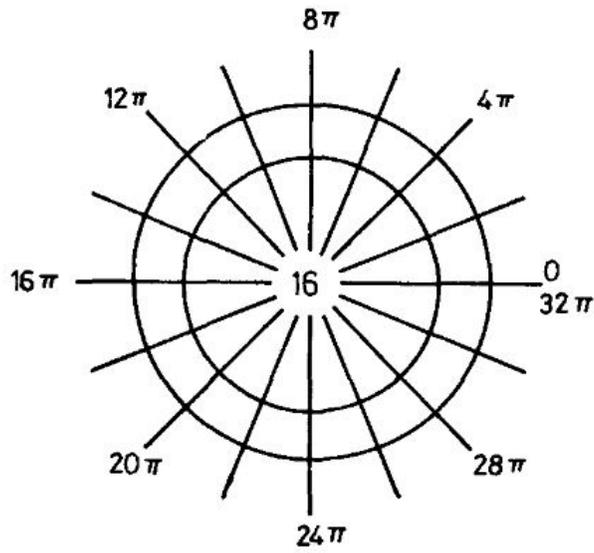


图 3

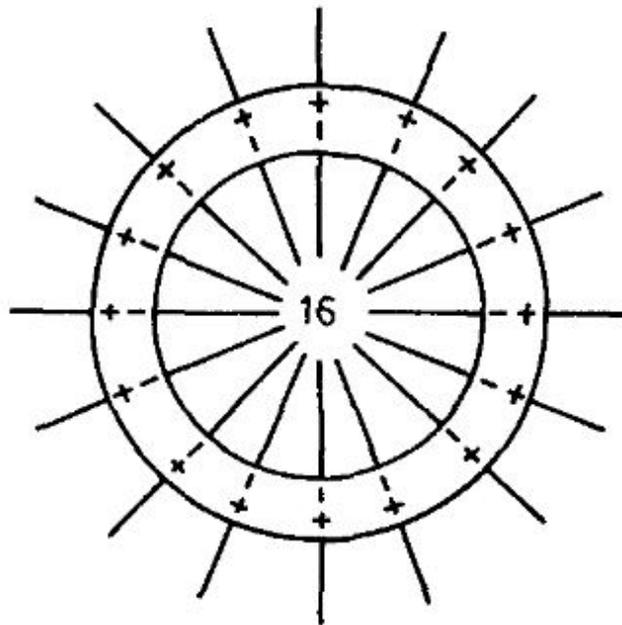


图 4

如果温度太高, $T > T_0$.

$$T_0 \approx \frac{q^2}{4} = \frac{\pi\rho}{2m}, \quad (40)$$

则各电偶极分散而成自由电荷, 则 $\epsilon \rightarrow \infty$, 因而 $\rho_s \rightarrow 0$, 超流态不再存在. 注意, 电荷单位 q 和 ρ/m 的关系是 $q = 2\pi\rho/m$. (见 (20), (29.48) 及 (29.53)).

以上的讨论, 可以借一些图来说明. 图 3 示一脚踏车轮子模样的东西. 轮胎是我们的二维空间模型, 周长为 L . 四射的直线是 ϕ 的定值线. 如果胎中没有漩涡, 则这图代表 ϕ_n , $n = 16$. 总动量由 (30) 可得. 当然, 我们可以把这图看成一个大漩涡, 中间是个大洞, 超流体在胎里流. 大漩涡的强度是 $16q$.

图 3 中四射的直线, 即 ϕ 的定值线, 亦可看成由中央电荷 nq 射出的电力线, 电场方向是由中心向外. 这也是一个很方便的看法. 电场强度是

$$\frac{nq}{(L/2\pi)} = \frac{2\pi nq}{L} = qmv_s, \quad (41)$$

$L/2\pi$ 是轮子的半径. 假定胎宽比 L 小很多.

如果有自由电荷出现, 即 $T > T_0$, 则自由电荷就把电力线屏蔽了, 如图 4. 胎内不再有电力线, 即不再有流量了.

如果 $T < T_0$, 电荷正负一对一对地配成了偶极, 电力线有 $16/\epsilon$ 根仍可透入.

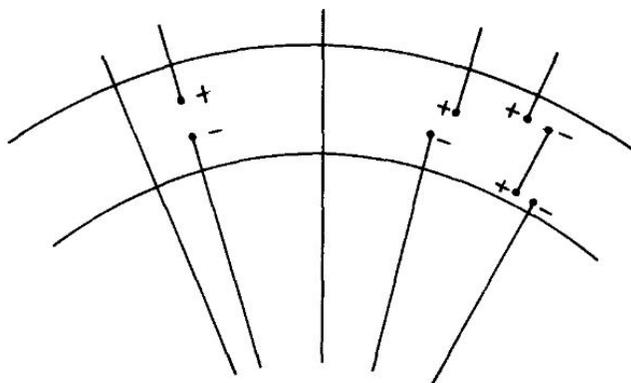


图 5

图 5 是把上二图的轮胎放大, 看一部分. 电力线有的穿过, 有的碰到偶极, 在偶极间的电场就小些. 平均起来, 电场小了 ϵ 倍 (见第十九章末节的讨论). 流量因此也小了 ϵ 倍.

现在来看在 $T < T_0$ 时, 流动态是否稳定. 这情形比一维空间的情形复杂多了. 在此只作一个十分粗略的分析. 我们将不考虑漩涡的详细运动情形, 只从能量减低过程的观点来看.

要改变流速, 必须要改变绕圈数 n . 从图 5 来看, 这就是要改变物体内的

电力线数目. 每减一根, n 即降低 1. 要减一根电力线, 只要拆散一对结合的正负电荷, 然后正电荷向上跑, 负电荷向下跑, 如此, 体内的 n 就减少了.

当然, 正负电荷间有引力, 否则不会结合. 要拆散它们, 不大容易, 要相当大的能量才行. 不过一旦拆开, 电力线所代表的“电场”就会把负电引向图 5 的内环, 正电推向外环, 以减低能量. 这又是一个翻越屏障的问题. 这电场是和流速成正比. 因此, 流速越大, 偶极越容易被拆散, n 下降得越快. 现在来看些细节.

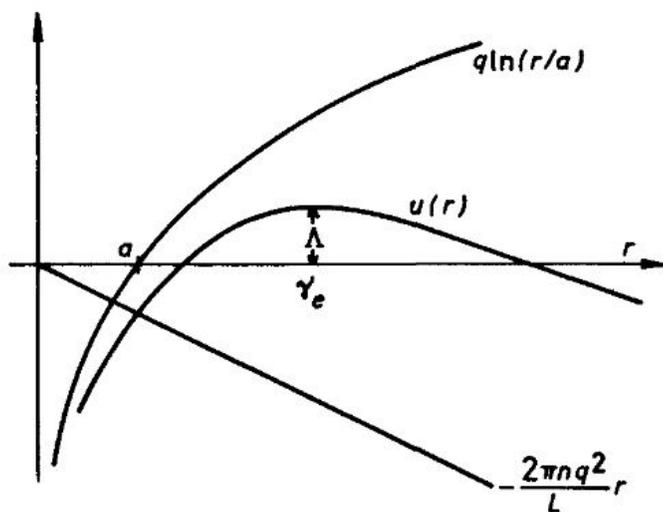


图 6

一个偶极的能量是 $+q$ 和 $-q$ 的作用能加上在外加电场

$$E = \frac{2\pi nq}{L\epsilon} = \frac{mv_s q}{\epsilon} \quad (42)$$

内的位能, 即

$$u(r) = \frac{1}{\epsilon} q^2 \ln \frac{r}{a} - qEr, \quad (43)$$

r 是 $+q, -q$ 间的距离. 图 6 示 $u(r)$, 其极大值 Λ 在 $r = r_c$,

$$r_c = \frac{L}{2\pi n} = \frac{1}{mv_s}. \quad (44)$$

因此, 流速越大, r_c 越小, Λ 之值是

$$\begin{aligned} \Lambda &= \frac{q^2}{\epsilon} \left(\ln \frac{r_c}{a} - mv_s r_c \right) \\ &= \frac{2\pi \rho_s}{m} \left(\ln \frac{1}{mv_s a} - 1 \right). \end{aligned} \quad (45)$$

如果 r 比 r_c 小很多, 则偶极的结合十分牢固. 要拉开一对, 十分困难, 必须越过屏障 Λ . 我们可以估计一下单位时间单位长度内有多少自由电荷流出. 我们用第二十一章的结果 (见 (21.12)). 流量是

$$J_\pi \approx \frac{DN_\pi}{V} \frac{e^{-\Lambda/T}}{\Delta r}, \quad (46)$$

$$(\Delta r)^2 \approx \frac{T}{(\partial^2 u / \partial r^2)_c} = \frac{\epsilon T}{q^2} r_c^2. \quad (47)$$

D 是电荷的扩散系数, N_π/V 是电偶极密度. V 是物体的总面积. Δr 是图 6 屏障峰顶的宽. 每一拆散的偶极使 n 减低 1, 故

$$\begin{aligned} \frac{dn}{dt} &= -LJ_\pi, \\ \frac{dv_s}{dt} &= -\frac{2\pi}{Lm} \frac{dn}{dt} = -\frac{2\pi}{m} J_\pi. \end{aligned} \quad (48)$$

将 (45) 代入 (46), (48), 得

$$\frac{dv_s}{dt} \approx -v_s \left(\frac{q^2}{\epsilon T} \right) \frac{D}{a^2} \left(\frac{1}{mv_s a} \right)^{-2\pi\rho_s/mT} \quad (49)$$

以上假定 $\frac{N_\pi}{V} \sim \frac{1}{a^2} e^{-q^2/T}$, $\frac{q^2}{\epsilon} = \frac{2\pi\rho_s}{m}$. 扩散系数必须另外计算. 它和这 He^4 薄膜所附的物体表面及漩涡的流体力学有关.

以上最重要的结论是, 二维空间的超流性是个暂稳态. 流速越大, 下降得也越快. 不过在低温, 低速的情形下, 它是相当的稳定. 这是由于 $e^{-\Lambda/T}$ 所引起的 (49) 的最后一因子所致.

(49) 式的解在很长的时间是, (假定 $mv_s a \ll 1$):

$$\begin{aligned} v_s &\sim \frac{1}{ma} \left(\frac{t}{\tau} \right)^{-\alpha T}, \\ \alpha &= \frac{m}{2\pi\rho_s}, \\ \frac{1}{\tau} &\sim \frac{D}{\alpha T a^2}. \end{aligned} \quad (50)$$

在低温时, v_s 消失得非常慢.

如果 mv_s 接近 $1/a$, 则图 6 的屏障就不见了, 也就没有超流性可言了.

三维空间的情形更加复杂. 缺陷是漩涡线. 由于它牵涉到一连串原子 (因为是一条线), 故能量很大. 因此它本身就是暂稳或稳定, 不会在平衡态中

忽然出现或忽然消失. 三维空间的超流状态非常稳定. 如果本来就有了一些漩涡线在边界上, 流速太大时就会把它拉长, 变成个大圈圈, 或移到管子另一边, 以减少绕圈数. 这些问题, 是非常有趣的流体力学问题, 早在十几年前, 即有很多数据, 比二维空间的工作要早很多[†]. 三维空间超流性的失去, 总是因为流速太高, 而最高的超流速和容器的细节大有关系. 暂稳态的屏障都非常高, 由于漩涡线的能量大, 平衡态中的分子能量不足以越过这样的屏障.

我们就讲到此地.

讨论问题三十

本章要求稍高, 读者需多下工夫?

1. 氦的种种特性, 在各参考书中叙述甚详, (例如吴家玮 (1972), Wilks (1967), Puterman (1974) 等), 读者可自阅读.
2. (A) 证明 (5). 注意

$$\begin{aligned} [s_+(\mathbf{R}), s_-(\mathbf{R}')] &= 2s_z(\mathbf{R})\delta_{\mathbf{R}\mathbf{R}'}, \\ [s_z(\mathbf{R}), s_{\pm}(\mathbf{R}')] &= \pm s_{\pm}(\mathbf{R})\delta_{\mathbf{R}\mathbf{R}'}. \end{aligned} \quad (51)$$

(B) 证明 (9).

3. 在 N'/N 很小时, (N' 为粒子数, 见 (10)), 试证这模型和理想波色子模型相近. 令

$$\psi(\mathbf{R}) \equiv a^{-3/2}s_-(\mathbf{R}), \quad \psi^*(\mathbf{R}) \equiv a^{-3/2}s_+(\mathbf{R}), \quad (52)$$

则

$$\begin{aligned} [\psi(\mathbf{R}), \psi^*(\mathbf{R}')] &= \delta(\mathbf{R} - \mathbf{R}'), \\ a^{-3}\delta_{\mathbf{R}\mathbf{R}'} &\approx \delta(\mathbf{R} - \mathbf{R}'). \end{aligned} \quad (53)$$

试证 $\psi^*(\mathbf{R})\psi(\mathbf{R})$ 大约是粒子密度算符, 注意, 这模型和理想波色气体仍不一样, 因为每格只能有一粒子, 故粒子间仍有作用.

4. 读者如对第 2 节的推理感到吃力, 则应复习量子力学讲自旋 $\frac{1}{2}$ 部分, 及有关坐标方向的转换.

[†] 见 Wilks (1967) 的书. Puterman (1974), Anderson (1966).

注意:

$$e^{-i\theta s_\alpha} = \cos \frac{\theta}{2} - 2is_\alpha \sin \frac{\theta}{2}, \quad \alpha = x, y, z. \quad (54)$$

(12) 中的 Φ_0 , 每自旋都向下, 故每自旋的状态还可用 $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ 代表.

5. 从 (14) 导出 (15), (16).

6. 从 (18) 求 (20) 的步骤如不清楚, 表示读者对泛函数微分还没经验. 对一函数 $\phi(x)$ 微分, 只须记住

$$\frac{\delta}{\delta\phi(x)}\phi(x') = \delta(x - x') \quad (55)$$

即可. 遇到困难时, 最保险的解决方法是把 x 分成小段, 把每一段的 $\phi(x)$ 看成一变数.

试求 $\delta F/\delta\phi(x)$:

$$(A) \quad F = \int dx dx' \phi^2(x) \phi^2(x') A(x, x'). \quad (56)$$

$$(B) \quad F = \int dx \phi^2(x) \left(\frac{d\phi}{dx} \right)^2. \quad (57)$$

$$(C) \quad F = \int dx \left[\phi(x)h(x) + (\nabla^2\phi(x))^2 \right] + \left(\int dx \phi^2(x) \right)^2. \quad (58)$$

7. 本章的讨论, 限于低温的情形. 若是 T 和临界温度接近, 则第 2 节的模型就不合用了.

(A) 用平均场法解模型 (7). 条件 (10) 仍必须符合. 求临界温度 T_c , 热容率, 及序参量 $\langle s_- \rangle$. 假定序参量到处相同.

(B) 假定序参量为不均匀:

$$\langle s_-(\mathbf{R}) \rangle = \psi(\mathbf{R}). \quad (59)$$

试证 $\psi(\mathbf{R})$ 约为

$$-\nabla^2\psi - A\psi + B\psi^3 = 0 \quad (60)$$

的解, (假定 $T_c - T$ 很小). 从 (7) 开始, 计算 A, B . 参考第二十七章问题 2, 第二十九章问题 3. 用第二十七章第 2 节的方法, 求出和 (27.11), (27.12) 相似的各式, 包括

$$F_0(\psi) = \langle H \rangle - TS \propto \int d^3 \mathbf{R} \left[\frac{1}{2} |\nabla \psi|^2 - \frac{1}{2} A |\psi|^2 + \frac{B}{4} |\psi|^4 \right], \quad (61)$$

(27.12) 式和 (60) 相当. 由 $\delta F / \delta \psi = 0$ 即得 (60).

(C) 用 (60) 取代 (21), 讨论第 3 节的各结论. 绕圈数的观念仍合用.

8. 上题的模型, 曾被用来分析细金属线的超导电性. 虽然我们没有讨论超导电性的理论, 但上题模型是可以用的. 本章的基本观念亦合用, 只是细节不同.

常数 A, B 须要由金属模型计算得来.

试讨论绕圈数 n 的改变. 注意: 这模型的 $|\psi|$ 是可以改变的.

即图 2 的曲线不必一定要在那手镯表面上. 不过, 如要 n 改变, 则曲线在某处必须切透手镯. 也就是说, 绕圈数由 n 减到 $n - 1$ 的过程中, $|\psi|$ 必须在某处减少至 0, 再增加. 这过程已有人仔细分析过, (见 Langer and Ambegaokar (1966)).

基本的想法是:

① 状态空间是 ψ 之集. ψ 是 z 的函数, z 是沿环坐标. 假定 ψ 不随 x, y 改变.

② (60) 决定 $F_0(\psi)$ 在此空间的极小值.

$$\psi_n \propto e^{i2\pi n z / L} \quad (62)$$

为 F_0 的极小处, ψ_0 处为绝对极小.

③ 从 ψ_n 走到 ψ_{n-1} , 必须越过 $F_0(\psi)$ 的高处, 即越过屏障. 屏障最低处, 第一山口, 或鞍点. 好在 (60) 不但可以用来求极小, 极大或山口点, 也是 (60) 的解. 令鞍点为 ψ' .

因此, 屏障高为

$$\Lambda = F_0(\psi') - F_0(\psi_n). \quad (63)$$

④ 由第二十一章的讨论, 可估计 n 至 $n-1$ 过程的反应率

$$\frac{1}{\tau} = \frac{1}{\tau_0} e^{-\Lambda/T}. \quad (64)$$

τ_0 不大容易计算, 但主要的因子 $e^{-\Lambda/T}$ 是由屏障得来.

读者自己先试试, 再去翻参考数据, 如果读者要作更广泛、详细的阅读, 可学 Putterman (1974).

9. 声子气体

本章的重点在超流性. 波色粒子的热力性质, 也很有趣, 但没有在本章讨论, 读者可自己导出一些结论如下.

(A) 复习上一章问题 6, 试证哈密顿函数可写成

$$H = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} a_{\mathbf{k}}^* a_{\mathbf{k}} + \frac{1}{2V} \sum_{\mathbf{k}, \mathbf{p}, \mathbf{q}} u_{\mathbf{q}} a_{\mathbf{k}+\mathbf{p}}^* a_{\mathbf{p}-\mathbf{q}}^* a_{\mathbf{p}} a_{\mathbf{q}},$$

$$u_{\mathbf{q}} \equiv \int d^3 \mathbf{r} e^{i\mathbf{q} \cdot \mathbf{r}} u(r). \quad (65)$$

(见 (29.57) 至 (29.64)).

(B) 假定作用能 u 很弱. 在低温时 $N_0 = a_0^* a_0$ 和 N 差不多, 其他 $a_{\mathbf{k}}$ 可看作小量. 因此, H 约是

$$H \approx \sum_{\mathbf{k}} (\epsilon_{\mathbf{k}} + 2nu_{\mathbf{k}}) a_{\mathbf{k}}^* a_{\mathbf{k}}$$

$$+ n \sum_{\mathbf{k}} u_{\mathbf{k}} (a_{-\mathbf{k}}^* a_{\mathbf{k}}^* + a_{-\mathbf{k}} a_{\mathbf{k}}) + \text{常数}, \quad (66)$$

$$n \equiv \frac{N}{V}.$$

(C) 注意, 位子 \mathbf{k} 只和 $-\mathbf{k}$ 相干. 因此, 模型 (66) 很容易解.

试求零点能量及波函数.

每一对 $(\mathbf{k}, -\mathbf{k})$ 位子的零点波函数是

$$\Psi_0 = \sum_m A_m (a_{-\mathbf{k}}^* a_{-\mathbf{k}}^*)^m \Phi_0. \quad (67)$$

Φ_0 是无粒子状态, 即真空. 求 A_m .

(D) 试证激发态可以看成一“声子理想气体”. 声子可由位子 $(\mathbf{k}, -\mathbf{k})$ 的激发态定义. 令

$$\alpha_{\mathbf{k}} \equiv u_{\mathbf{k}} a_{\mathbf{k}} - v_{\mathbf{k}} a_{-\mathbf{k}}^* \quad (68)$$

这是声子的“消去符”, $\alpha_{\mathbf{k}}^*$ 为“产生符”. $u_{\mathbf{k}} v_{\mathbf{k}}$ 是常数. 试由

$$\begin{aligned} [\alpha_{\mathbf{k}}, \alpha_{\mathbf{k}}^*] &= 1 \\ \alpha_{\mathbf{k}} \Psi_0 &= 0 \end{aligned} \quad (69)$$

求 $u_{\mathbf{k}}, v_{\mathbf{k}}$, 并证明

$$\begin{aligned} H &= \sum_{\mathbf{k}} \omega_{\mathbf{k}} \alpha_{\mathbf{k}}^* \alpha_{\mathbf{k}} + \text{常数}, \\ \omega_{\mathbf{k}}^2 &\equiv 2nu_{\mathbf{k}}\epsilon_{\mathbf{k}} + \epsilon_{\mathbf{k}}^2. \end{aligned} \quad (70)$$

如果 k 很小, 则

$$\begin{aligned} \omega_{\mathbf{k}} &\approx ck, \\ c^2 &\equiv nu_0/m. \end{aligned} \quad (71)$$

这是“声子”一词的动机.

以上是四十年前 Bogoliubov 的结果.

(E) 用 (70) 求热力性质.

注意: 在旧一点的教科书及文献中, 常把超流性和声子的能量连想在一起, 把 $c > 0$ 看成超流性的来源, 流速小于 c 则射不出声子因而不能损失能量. 这种看法是不对的 (固体也有声子, 却无超流性). 本章的讨论强调绕圈数的守恒及序参量韧度为超流性的根源, 读者请特别注意.

中文参考数据

李怡严 (1967)“大学物理学”(1967 年, 东华书局).

吴大猷 (1979)“热力学, 气体运动论, 统计力学”(1979 年联经出版事业).

吴家玮 (1972)“氦和多体物理”(1972 年中华书局).

唐江涛 (1971) 主编, “材料”(1971 年科学图书社).

马上庚 (1981)“物理”(中国物理学会) 第二期, 2.

杨锥哲 (1979)“概率论”(1979 年国立编译馆, 正中书局).

英文参考数据

- Abraham, F. F. (1979) *Phys. Rep.* 53, No.2.
- Alexandrowicz, Z. (1975) *J. Stat. Phys.* 13231.
- Abrikosov, A. A., L. P. Gorkov and I. E. Dzyaloshinski (1963) *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, New Jersey).
- Anderson, P. W. (1966) *Rev. Mod. Phys.* 38, 298.
- Anderson, P. W., B. I. Halperin and C. Varma (1972) *Philosophical Mag.* 25, 1.
- Arnold, V. I. and A. Avez (1968) "Ergodic Problems of Classical Mechanics" (W. A. Benjamin, New York).
- Berne, B. J., (1977), editor, "Statistical Mechanics" (Plenum Press, New York).
- Beth, E., and G. E. Uhlenbeck (1937) *Physica* 4, 915.
- Binder, K. (1979) *Monte Carlo Methods in Statistical Physics* (Springer-Verlag, Berlin).
- Brout, R., and P. Carruthers (1963) "Lectures in the Many Electron Problem" (Gordon and Breach, New York).
- Burton, W. K., N. Cabrera and F. C. Frank (1951) *Phil. Trans. Roy. Soc. London* 243A, 299.
- Chui, S. T. and J. D. Weeks (1976) *Phys. Rev. B* 14, 4978.
- Dasgupta, C., S. Ma and C. K. Hu (1979), *Phys. Rev. B* 20, 3837.
- Dashen, R., S. Ma and H. J. Bernstein (1969) *Phys. Rev.* 187, 345.
- Dashen, R. and S. Ma (1971) *Phys. Rev. A* 4, 700.
- Dundon, J. M. and J. M. Goodkind (1974) *Phys. Rev. Lett.* 32, 1343-6.
- Dyson, F. J. (1967) *J. Math. Phys.* 8, 1538.
- Feigenbaum, M. J. (1981) *J. Stat. Phys.* 1925.
- Fisher, M. E. (1981) *Physica* 106 A, 28.
- Forster, D. (1975) "Hydrodynamic Fluctuations, Broken Symmetry and Correlation Functions" (W. A. Benjamin, New York).
- Gibbs, J. W. (1960) "Elementary Principles of Statistical Mechanics" (Dover Publications, New York).
- Griffiths, R. B. (1972) in "Phase Transition and Critical Phenomena" Vol. 1, C. Domb and M. S. Green, eds. (Academic Press, New York).

- Grimes, C. C. and G. A. Adams (1979) *Phys. Rev. Lett.* 42, 795.
- Imry, Y. and L. Gunther (1971) *Phys. Rev.* B3, 3939.
- Hahn, E. L. (1950) *Phys. Rev.* 80, 580.
- Hohenberg, P. C. (1967) *Phys. Rev.* 158, 383.
- Huang, K. (1963) "Statistical Mechanics" (J. Wiley, New York).
- Kikuchi, F. (1951) *Phys. Rev.* 81, 988.
- Kittel, C. (1956), (1966) "Introduction to Solid State Physics", 3rd ed. (J. Wiley, New York).
- Kosterlitz, J. M. and D. J. Thouless, (1973) *J. Phys. C* 6, 1181.
- Kosterlitz, J. M. (1974) *J. Phys. C* 1, 1046.
- Krylov, N. S. (1979) "works on the Foundations of Statistical Physics" (Princeton University Press, Princeton).
- Langer, J. S., and V. Ambegaokar (1967) *Phys. Rev.* 164, 498.
- Landau, L. D. and E. M. Lifshitz (1980) "Statistical Physics" (Pergamon Press, London).
- Lebowitz, J. L. and E. H. Lieb (1969) *Phys. Rev. Lett.* 22, 631.
- Lifshitz, E. M. and L. P. Pitaevskii (1980) "Statistical Physics" (Pergamon Press, London).
- Ma, S. K. (1976) "Modern Theory of Critical Phenomena" (Benjamin, New York).
- Ma, S. K. (1981a) *J. Stat. Phys.* 26, 221.
- Ma, S. K. and M. Payne (1981) *Phys. Rev.* B24, 3984.
- Malmberg, J. H., C. B. Wharton, R. W. Gould, and T. M. O'Neils, (1968) *Phys. Fluids* 11, 1147.
- Mattis, D. C. (1965) "The Theory of Magnetism" (Harper and Row, New York).
- Matsubara, T. (1955) *Prog. Theor. Phys. (Kyoto)* 14, 351.
- Meirovitch, H., (1977) *Chem. Phys. Lett.* 45, 389.
- Mermin, N. D. and H. Wagner (1966) *Phys. Rev. Lett.* 17, 1133.
- Nelson, D. R. and B. I. Halperin (1980) *Phys. Rev.* B21, 5312.
- O'Neil, T. M. and R. W. Gould (1968) *Phys. Fluids* 11, 134.
- Onsager, L. (1944) *Phys. Rev.* 65, 117.

- Osgood, E. B. and J. M. Goodkind (1967) *Phys. Rev. Lett.* 18, 894.
- Peierls, R. (1936) *Camb. Phil. Soc.* 32, 477.
- Peierls, R. (1955) *Quantum Theory of Solids*. (Oxford Clarendon Press, London).
- Peierls, R. (1979) *Surprises in Theoretical Physics* (Princeton University Press, Princeton).
- Prohofsky, E. W. and J. A. Krumhansl (1964) *Phys. Rev.* A133, 1403.
- Putterman, S. J. (1974) "Super fluid Hydrodynamics" (North-Holland, Amsterdam).
- Rudnick, J. (1970) Ph.D. Thesis, University of California, San Diego.
- Sinai, Ya. G. (1977) "Introduction to Ergodic Theory" (Princeton University Press, Princeton).
- Stanley, H. E. (1971) "Introduction to Phase Transition and Critical Phenomena" (Oxford University Press, London).
- Tabor, D. (1979) *Gases, Liquids and Solids* 2nd ed. (Cambridge University Press, Cambridge).
- Tolman, R. C. (1962) *Principles of Statistical Mechanics* (Oxford University Press, London).
- Wannier, G. H. (1966) "Statistical Physics" (J.Wiley, New York).
- Weeks, J. D. (1980) in "Ordering in Strongly Fluctuating Condensed Matter Systems", T. Riste, ed. (Plenum, New York).
- Wilks, J. (1961) "Third Law of Thermodynamics" (Oxford University Press, London).
- Wilks, J. (1967) "Liquid and Solid Helium" (Oxford University Press, London).
- Yang, C. N. (1952) *Phys. Rev.* 85, 809.
- Yang, C. N. and T. D. Lee (1952) *Phys. Rev.* 87, 404.
- Young, A. P. (1979) *Phys. Rev.* B19, 1855.
- Zemansky, M. W., (1957) "Heat and Thermodynamics", 4th ed. (McGraw-Hill, New York).