普通高等教育“十五”国家级规划教材

理论物理基础教程

刘连寿　主编

高等教育出版社

ISBN 7-04-013099-8
9 787040 130997 >
定价 57.00 元

普通高等教育“十五”国家级规划教材

理论物理基础教程

刘连寿 主编

刘连寿　陈立华
汪德新　李家荣　编



高等教育出版社
HIGHER EDUCATION PRESS

图书在版编目（CIP）数据

理论物理基础教程／刘连寿主编．—北京：高等教育出版社，2003.10
ISBN 7-04-013099-8

Ⅰ.理… Ⅱ.刘… Ⅲ.理论物理学—高等学校—教材 Ⅳ.041

中国版本图书馆 CIP 数据核字(2003)第 069444 号

策划编辑 胡凯飞　**责任编辑** 钟金城　董洪光　**封面设计** 刘晓翔
责任绘图 朱　静　**责任印制** 宋克学

出版发行	高等教育出版社	**购书热线**	010-64054588
社　　址	北京市西城区德外大街 4 号	**免费咨询**	800-810　0598
邮政编码	100011	**网　　址**	http://www.hep.edu.cn
总　　机	010-82028899		http://www.hep.com.cn
经　　销	新华书店北京发行所		
印　　刷	北京中科印刷有限公司		
开　　本	787×960　1/16	**版　　次**	2003 年 10 月第 1 版
印　　张	50.25	**印　　次**	2003 年 10 月第 1 次印刷
字　　数	950 000	**定　　价**	57.00 元

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换。

版权所有　侵权必究

序

20世纪中叶以后，随着物理学的迅速发展，在普通高等学校本科教学中，综合性的“理论物理”课程被分成四门：《理论力学》，《电动力学》，《热力学与统计物理》，《量子力学》. 学生再也见不到作为一门统一学科的《理论物理》. “四大力学”之间的本质联系，不同“力学”的统一物理思想、处理问题的统一方法，不能充分反映出来.

为克服上述弊端，20世纪八十年代前后，国内外出现了一些统一的大学“理论物理”教材，如美国 E. C. Harris 著《Introduction to modern theoretical physics》，北京大学彭恒武，徐锡中著《理论物理基础》等. 这些教材的共同特点是“充分注意到理论物理的统一性”（彭恒武），使学生“对物理学的统一性获得更深的理解”（E. G. Harris）.

随着科学技术的发展，理论物理已成为不仅是理科而且是许多工科专业的基础课程. 1985年理科教学指导委员会《理论物理》教材编审小组兰州会议指出，为适应科学技术的发展和理论物理教学改革的趋势，迫切需要一本统一的理论物理教材，它应克服分别开设“四大力学”互不通气的毛病，具有物理思想一贯，系统风格一致的特点. 根据这一精神，我们于1986年编写出版了《理论物理简明教程》.

《简明教程》从1986年至1995年一直在华中师范大学物理系本科教学中使用，并被一些兄弟院校的物理专业本科及某些非物理专业，如“应用力学”、“应用数学和系统工程”、“电子工程”、“自动控制”、“材料科学”等专业的研究生教学中多次采用.

1996年经国家教委批准，在华中师范大学设立了“国家理科（物理学）基础科学研究和教学人才培养基地”. 为适应基地班的教学，我们对原《简明教程》进行了重大修改，更名为《理论物理基础教程》. 在基地班的教学中连续试用五届. 根据教学经验反复修订，重印三次，成为现在这本书.

我们的目标不是为非物理专业提供一个水平较低、时数较少的四门“力学”的合订本；而是为克服四大力学分别开设的弊端，用统一的物理思想、统一的演绎方法；注意前后呼应、相互渗透、综合运用；将理论物理在一个较高的水平上，作为一个整体，介绍给物理专业本科学生及某些需要理论物理知识较多的非物理专业研究生.

本教材的特点是，

（1）面向客观物理对象，按照从低速到高速、从宏观到微观、从单体到多体、从动力学到统计系综的线索组织教材. 各部分相互呼应、相互渗透，以便学生在所学

内容逐步深入的同时，受到综合运用理论物理规律的能力的训练.

(2)面向科学技术的高速发展，根据新一代青年接受知识的能力，在各部分内容的讲述中以基本规律为出发点，厚今薄古，删除陈旧落后、以及和普通物理重复的内容，精简学时、保证质量. 与此同时，在讲述上由浅入深，配合大量例题，以保证教材的可接受性.

(3)全书以守恒律和对称性为基础，以拉格朗日表述和哈密顿表述贯彻始终，提高了教学的起点，加深了学生对理论物理的理解. 使整个教材的水平更加适应科技发展对有创新能力人才的需要.

(4)在书中编入了数十道要求用计算机求解的习题，以培养学生用计算机研究物理问题的初步能力.

大体上说来，本书的第一篇是分析力学，第二、三、四篇是电动力学，五、六篇是量子力学，第七篇至第十篇是平衡和非平衡统计物理及统计热力学，全书是一个完整的整体，各部分之间相互渗透、相互呼应、相互引用，是一本完整的教材. 在教师恰当的处理下，也可以选用其中一部分. 书中一些打星号的内容，在教学时数不够的情况下可以略去，不致影响后续内容的学习.

在本书的编写修订过程中，得到武汉大学王治梁、田德诚、成都科技大学郑希特、华中科技大学吴大进、清华大学邝宇平、国防科技大学兰马群、华中师范大学杨约翰等同行专家的批评与鼓励. 作者在此表示深忱的感谢.

我校基地班同学在学习过程中，对本书(试用本)认真钻研、热情讨论，提出了许多宝贵的意见和建议，使书的质量得到不断提高，在此深表感谢.

作者在此还要感谢高燕敏在本书试用本的反复修改、重印的过程中，对书稿的输入付出的巨大劳动和提出的意见和建议.

理论物理是一个内容丰富，不断发展的学科，理论物理教学的改革是一个系统工程. 限于编者的水平，书中的错误和不足之处在所难免，真诚地希望得到批评和指正.

编　　者

2003 年 6 月 20 日

目　录

绪论 …… (1)

第一篇　低速宏观物体的运动 …… (3)

第一章　低速宏观运动的基本原理 …… (5)
§1.1.1　无约束质点的拉格朗日方程 …… (5)
§1.1.2　有约束情况下的拉格朗日方程 …… (10)
§1.1.3　最小作用量原理 …… (15)
§1.1.4　伽利略相对性原理　自由质点的拉格朗日函数 …… (19)
(一)惯性参考系　(二)伽利略变换　(三)伽利略相对性原理　(四)自由质点的拉格朗日函数
习题 …… (22)

第二章　守恒律 …… (25)
§1.2.1　动量和能量 …… (25)
(一)循环坐标与广义动量　(二)能量　(三)相互作用质点系的拉格朗日函数　(四)机械能
§1.2.2　守恒律与对称性的关系　角动量 …… (29)
(一)时间的均匀性与能量守恒　(二)空间的均匀性与动量守恒　(三)空间的各向同性与角动量守恒　(四)质点组的角动量
§1.2.3　质点组的动量、能量与角动量　质心 …… (32)
(一)质点组的动量定理　(二)质心　质心系　(三)动量和能量的变换　(四)角动量的变换
习题 …… (37)

第三章　有心力场中的运动 …… (39)
§1.3.1　二体问题　约化质量 …… (39)
§1.3.2　有心力场中运动的一般分析 …… (40)
(一)角动量守恒　等面积定律　(二)能量守恒　运动形式的分类　(三)运动方程的解　(四)比耐公式

§1.3.3 平方反比引力 开普勒问题 …… (44)
(一)运动形式的分类 (二)运动轨道 (三)行星的运动 开普勒问题
§1.3.4 碰撞与散射 …… (48)
(一)散射问题的提法 散射截面 (二)α粒子的散射 卢瑟福公式 (三)硬表面的散射 *(四)实验室系和质心系中的散射
习题 …… (55)

第四章 微振动 …… (57)
§1.4.1 无阻尼的微振动 …… (57)
(一)自由振动方程 (二)自由振动方程的解 (三)受迫振动方程的解 (四)受迫振动的能量转移
§1.4.2 阻尼振动 共振 …… (60)
(一)无阻尼的共振 (二)阻尼振动方程的解 (三)有阻尼情况下的共振 (四)通过共振时能量吸收率和相位的变化
§1.4.3 多自由度的耦合振动 …… (64)
(一)弱耦合的二振子系统 (二)对称矩阵的本征值与本征矢 (三)多自由度耦合振子的集体振动模式 简正坐标
习题 …… (75)

第五章 刚体的运动 …… (77)
§1.5.1 刚体的角速度、角动量与转动能量 …… (78)
(一)不同轴转动之间的关联 角速度矢量 (二)角动量矢量与转动惯量张量 惯量椭球(三)刚体的能量 (四)刚体的定点运动和定轴运动
§1.5.2 刚体的运动方程 …… (86)
(一)动量定理与角动量定理 (二)刚体的静平衡 (三)刚体的动平衡 (四)刚体的自由运动 *(五)刚体定点运动的运动学方程
§1.5.3 非惯性系中的运动 …… (92)
习题 …… (94)

第六章 低速宏观运动规律的正则形式 …… (97)
§1.6.1 哈密顿方程 …… (97)
(一)勒让德变换 (二)哈密顿函数 (三)哈密顿方程 (四)哈密顿方程的变分原理 (五)相空间
§1.6.2 守恒律 泊松括号 …… (101)
(一)力学量对时间的导数 (二)用泊松括号表出的运动方程 (三)能量守

恒与动量守恒　(四)泊松括号的定义和性质
§1.6.3　正则变换 …… (103)
(一)正则变换的定义　(二)正则变换的生成函数　(三)正则变换举例
(四)时间演化作为正则变换
* §1.6.4　作为坐标和时间函数的作用量 哈密顿-雅可比方程…… (107)
(一)作为坐标和时间函数的作用量　(二)哈密顿-雅可比方程　(三)利用哈密顿-雅可比方程求解动力学问题
习题 …… (110)

第二篇　高速运动的规律 …… (113)

第一章　狭义相对论的实验基础和基本原理 …… (115)
§2.1.1　迈克耳孙－莫雷实验 …… (115)
§2.1.2　相对论的基本原理 …… (117)
习题 …… (118)

第二章　相对论的时空理论 …… (119)
§2.2.1　间隔　相对论的时空结构 …… (119)
(一)事件和两事件的间隔　(二)间隔不变——光速不变原理的数学表示
(三)间隔的分类
§2.2.2　相对论的时空坐标变换公式 …… (123)
§2.2.3　相对论的速度变换公式 …… (125)
§2.2.4　相对论的时空性质 …… (127)
(一)同时的相对性　(二)运动时钟的延缓　(三)运动尺度的缩短
* §2.2.5　相对论时空理论的若干应用 …… (134)
(一)菲索流水实验　(二)光行差现象　(三)多普勒效应
§2.2.6　四维协变量及方程的协变性 …… (136)
习题 …… (141)

第三章　高速粒子的运动规律 …… (143)
§2.3.1　相对论力学的基本方程 …… (143)
§2.3.2　高速运动粒子的守恒定律 …… (148)
§2.3.3　电磁场中的带电粒子 …… (151)
(一)电磁场中带电粒子的作用量　(二)电磁场中带电粒子的运动方程
(三)电磁场张量

习题 …… (153)

第四章 电磁现象的基本规律 …… (155)
§2.4.1 电磁规律的三维表述 …… (155)
(一)电荷守恒律 麦克斯韦方程组和洛伦兹力公式 (二)介质的电磁性质方程
§2.4.2 电磁场的边值关系 …… (158)
(一)电荷电流分布的描述 (二)电磁场的边值关系
§2.4.3 电磁场的守恒定律 …… (161)
(一)电磁场与电荷的能量转化与守恒 (二)电磁场与电荷的动量转化与守恒
§2.4.4 电磁规律的四维表述 …… (166)
(一)四维电流密度矢量 电荷守恒定律的四维形式 (二)四维电势矢量和电磁场张量 麦克斯韦方程组的四维形式 (三)电磁场的变换公式 (四)洛伦兹力公式的四维形式
习题 …… (171)

第三篇 电磁场 …… (175)

第一章 静电场 …… (177)
§3.1.1 静电势 静电问题的唯一性定理 …… (177)
(一)静电势 (二)静电势满足的方程及边值关系 (三)静电问题的唯一性定理
§3.1.2 分离变量法 …… (182)
§3.1.3 电象法 …… (187)
§3.1.4 格林函数法 …… (191)
(一)泊松方程的格林函数 (二)常见的泊松方程第一类边值问题的格林函数 (三)格林函数的对称性 (四)用格林函数法求解泊松方程的第一类边值问题
§3.1.5 电多极矩法 …… (196)
§3.1.6 静电场的能量 电荷体系与外电场的相互作用 …… (202)
(一)静电场的能量 (二)电荷体系在外电场中的能量 (三)电多极子在外电场中所受的力和力矩
习题 …… (206)

第二章　静磁场 …… (211)
§3.2.1　矢势　静磁问题的唯一性定理 …… (211)
(一)矢势　(二)静磁场的唯一性定理　(三)矢势满足的方程与边值关系　*(四)用矢势法求解静磁场
§3.2.2　磁标势 …… (218)
(一)引入磁标势的条件　(二)磁标势的定义　磁标势满足的方程与边值关系　(三)用磁标势求解静磁场
§3.2.3　磁多极子 …… (222)
§3.2.4　静磁场的能量　电流分布与静磁场的相互作用 …… (224)
(一)静磁场的能量　(二)电流分布在外磁场中的能量　(三)磁偶极子与外磁场的相互作用能　(四)磁偶极子在外磁场中所受的力、力矩及其势能　(五)拉莫尔进动　(六)磁偶极子在外磁场中的势能与相互作用能的关系
§3.2.5　超导体的电磁性质 …… (229)
(一)超导体的基本电磁现象　(二)伦敦方程　(三)磁标势法
习题 …… (235)

第三章　电磁波的传播 …… (237)
§3.3.1　平面电磁波 …… (237)
(一)自由空间电磁场的波动方程　(二)平面电磁波　(三)单色电磁波的运动方程——亥姆霍兹方程　(四)平面电磁波的性质
§3.3.2　电磁波在介质分界面的反射与折射 …… (241)
(一)反射与折射定律　(二)费涅尔公式　(三)全反射　(四)光压
§3.3.3　电磁波在导体中的传播 …… (245)
(一)导体内自由电荷的分布　(二)良导体中单色电磁波的传播
§3.3.4　谐振腔 …… (248)
(一)腔内电磁场满足的方程和边界条件　(二)用分离变量法求解亥姆霍兹方程的边值问题　(三)谐振腔内的模式数　模密度
§3.3.5　波导管 …… (252)
(一)矩形波导管中的电磁波　(二)截止频率　(三)TE_{10}波的电磁场
习题 …… (256)

第四篇　电荷体系与电磁场的相互作用 …… (257)

第一章　宏观电荷体系辐射的电磁场 …… (259)
§4.1.1　电磁势的场方程　推迟势 …… (259)

(一)规范变换与规范不变性 (二)电磁势的场方程 (三)推迟势
§4.1.2 电偶极辐射 …… (262)
(一)推迟势的多极展开 (二)电偶极场及其性质 (三)电偶极辐射的角分布与总辐射功率
习题 …… (266)

第二章 运动带电粒子与电磁场的相互作用 …… (269)
§4.2.1 运动带电粒子的势和辐射的电磁场 …… (269)
(一)李纳-维谢尔势 (二)运动带电粒子的场
§4.2.2 韧致辐射与同步辐射 …… (273)
(一)加速带电粒子辐射的电磁场 (二)加速带电粒子的辐射功率和角分布 (三)韧致辐射 (四)同步辐射
§4.2.3 切连科夫辐射 …… (279)
(一)切连科夫辐射的理论解释 (二)在介质中运动的带电粒子激发的电磁场
§4.2.4 带电粒子的电磁场对粒子本身的反作用 …… (282)
(一)电磁质量及电子的经典半径 (二)辐射阻尼力 (三)谱线的自然宽度
§4.2.5 电磁波的散射、吸收和色散 …… (286)
(一)电磁波的散射 (二)电磁波的吸收 (三)电磁波在稀薄气体中的色散现象
习题 …… (291)

第五篇 微观粒子的运动规律 …… (293)

第一章 微观粒子状态的描述 …… (295)
§5.1.1 光的粒子性 …… (295)
(一)黑体辐射的研究 "能量子"的提出 (二)光电效应的研究 "光量子"的提出 (三)康普顿效应的研究 光的粒子性的进一步证实
§5.1.2 粒子的波动性 波函数 …… (299)
(一)原子光谱的分立性 玻尔量子化假设 (二)德布洛意波 (三)电子衍射实验 (四)微观粒子的波粒二象性 (五)波函数 (六)波函数的叠加原理
习题 …… (305)

第二章 力学量的算符 …… (307)

§5.2.1　测量结果的期望值　用算符表示力学量 …… (307)
(一)在给定状态里坐标的平均值　(二)在给定状态里动量的平均值　(三)用算符表示物理量　(四)线性厄米算符
§5.2.2　测量结果的概率分布　算符的本征值与本征函数 …… (312)
(一)力学量取值的谱　(二)本征值方程　(三)本征函数的正交性　(四)本征函数的归一化　(五)在给定状态里力学量取值的概率
§5.2.3　力学量的完全组　算符的对易关系 …… (325)
(一)力学量同时有确定值的条件　(二)对易关系　(三)算符乘积的厄米性　(四)力学量的完全组
§5.2.4　不确定关系 …… (332)
习题 …… (336)

第三章　运动方程和守恒律 …… (339)
§5.3.1　薛定谔方程 …… (339)
(一)薛定谔方程　(二)粒子数守恒与概率流密度　(三)解薛定谔方程的分离变量法　定态
§5.3.2　力学量对时间导数的算符　量子泊松括号　正则量子化……
…… (352)
(一)力学量对时间导数的算符　(二)正则量子化
§5.3.3　对称性与守恒律 …… (355)
(一)守恒量　(二)对称变换
习题 …… (361)

第四章　量子力学的表述形式 …… (363)
§5.4.1　希尔伯特空间 …… (363)
(一)几个不同表象的例子　(二)希尔伯特空间
§5.4.2　态矢量和算符 …… (367)
(一)态矢量　(二)态矢量的点积　(三)厄米算符　(四)厄米算符的本征值与本征矢　(五)力学量用厄米算符表示　(六)运动方程
§5.4.3　表象和表象变换 …… (372)
(一)态矢量在 F 表象中的分量形式　(二)算符在 F 表象中分量形式　(三)进入和退出表象　(四)用厄米矩阵表示力学量　(五)本征值方程与本征表象　(六)表象变换　(七)表象变换矩阵的么正性
§5.4.4　算符对易关系的确定 …… (380)
(一)物理量作为空间变换的生成元　奈特定理　(二)坐标与动量的对易关

系 (三)角动量分量之间的对易关系
§5.4.5 谐振子的占有数表象 …… (383)
§5.4.6 角动量的矩阵表示 …… (388)
(一)角动量的本征值 (二)角动量的矩阵元 (三)整数角动量 (四)电子的自旋角动量 泡利矩阵 *(五)角动量合成
§5.4.7 薛定谔绘景和海森伯绘景 …… (396)
(一)量子力学的不同绘景 (二)薛定谔绘景 (三)从薛定谔绘景到海森伯绘景
§5.4.8 用密度矩阵表示量子态 …… (399)
(一) 投影算符 (二)密度矩阵
习题 …… (402)

第五章 多粒子系统 全同性原理 …… (405)
§5.5.1 全同性原理 …… (405)
(一)全同性原理 (二)波函数的对称化和反对称化 (三)福克表象
§5.5.2 电子自旋波函数对粒子交换的对称性 …… (409)
习题 …… (413)

第六章 量子干涉现象 …… (415)
§5.6.1 路径积分量子化 …… (415)
* §5.6.2 双缝衍射 …… (418)
* §5.6.3 磁通量子 …… (421)
(一)阿喀洛诺夫——玻姆效应 (二)线圈内的磁通对线圈外电子能级的影响
习题 …… (423)

第七章 高速微观粒子的运动 …… (425)
§5.7.1 建立高速微观粒子运动方程遇到的困难 …… (425)
§5.7.2 自由粒子的狄拉克方程 …… (426)
(一) 方程的建立 (二) 概率密度与概率流密度 (三) 狄拉克海 (四) 狄拉克粒子的自旋 (五)自由狄拉克方程的解
* §5.7.3 电子的内禀磁矩 …… (431)
习题 …… (432)

第六篇 微观粒子在外场中的运动 …… (435)

第一章　外场中的定态问题 …… (437)
§6.1.1　坐标表象中定态方程的严格解 …… (437)
(一)坐标表象中的一维谐振子　(二)有心力场中运动的一般分析　(三)电子在库仑场中的运动
§6.1.2　定态微扰方法 …… (448)
(一)求解定态问题的近似方法　(二)微扰方法的基本方程　(三)无简并的未扰能级的微扰修正　(四)有简并的未扰能级的微扰修正
§6.1.3　氦原子 …… (458)
(一)微扰计算　(二)正氦与仲氦　(三)交换作用
§6.1.4　变分法 …… (461)
(一)变分原理　(二)作为近似方法的变分法
* §6.1.5　耦合振子系统　量子力学的正则变换 …… (464)
习题 …… (465)

第二章　量子跃迁 …… (467)
§6.2.1　研究时间演化的微扰方法　相互作用绘景 …… (467)
(一)相互作用绘景中的运动方程　(二)含时微扰　(三)恒定微扰的加入过程
§6.2.2　周期性外场引起的跃迁 …… (471)
(一)单色周期性的微扰　(二)非严格单色的微扰　(三)末态为连续谱的跃迁
§6.2.3　光的辐射和吸收 …… (474)
(一)微扰项的确定　(二)跃迁概率　(三)自发辐射　(四)谱线强度和选择定则
§6.2.4　不稳定系统的寿命和能级宽度　能量-时间不确定关系　(479)
(一)激发态原子的寿命　(二)不稳定系统的能级宽度　能量-时间不确定关系
习题 …… (481)

第三章　弹性散射 …… (483)
§6.3.1　散射问题的一般讨论 …… (483)
(一)散射截面　(二)散射问题的边界条件　散射振幅
§6.3.2　玻恩近似 …… (485)
§6.3.3　分波法 …… (490)

(一)自由粒子方程的球面波解 (二)中心场散射的分波展开 (三)散射振幅和散射截面
习题 …… (495)

第七篇 宏观系统统计理论的基本原理 …… (497)

第一章 统计理论基础 …… (499)
§7.1.1 宏观系统及其分类 …… (499)
§7.1.2 系统状态的宏观描述 平衡态 状态方程 …… (501)
(一)宏观量 宏观态 (二)平衡态 态参量和态函数 (三)状态方程
§7.1.3 系统状态的微观描述 统计规律性 …… (505)
(一)微观态 (二)微观态的量子力学描述 (三)微观态的经典力学描述 Γ空间 (四)统计规律性
§7.1.4 宏观量的统计性质 系综和系综平均 …… (516)
§7.1.5 密度函数和密度算符 经典和量子刘维尔定理 …… (519)
(一)密度函数 经典刘维尔定理及其推论 (二)密度算符 量子刘维尔定理及其推论
§7.1.6 微正则系综和微正则分布 …… (525)
习题 …… (526)

第二章 统计热力学基础 …… (529)
§7.2.1 热力学第一定律 功和热量的微观解释 …… (529)
(一)封闭系统中的热力学第一定律 (二)准静态过程及其功 (三)热量的计算 热容量 (四)热力学第一定律的微观解释
§7.2.2 热力学第二定律 熵 …… (534)
(一)支配自发过程方向和限度的物理量 (二)熵 (三)熵增加原理
§7.2.3 温度 化学势和压强 热力学基本方程 …… (540)
(一)热接触系统的平衡 温度 (二)处于扩散接触子系统的平衡 化学势 (三)处于力学接触子系统的平衡 压强 (四)热力学基本方程
§7.2.4 熵差的计算 …… (550)
(一)利用熵函数表达式求熵差 (二)利用热温比积分求熵差
§7.2.5 热力学势 麦克斯韦关系及它们的简单应用 …… (554)
(一)热力学势 (二)麦克斯韦关系 (三)特性函数 (四)普遍热力学关系的简单应用 (五)巨势
§7.2.6 气体的节流和绝热膨胀 磁冷却 …… (562)

(一)气体的节流膨胀　(二)气体的绝热膨胀　(三)磁冷却

§7.2.7　最大功原理　平衡的稳定性条件 …… (566)

(一)最大功原理　(二)平衡的稳定性判据和条件

§7.2.8　单元双相系的平衡　相图 …… (571)

(一)单元双相系的平衡条件和稳定性条件　(二)相平衡条件的讨论　相图

§7.2.9　一级相变和二级相变 …… (576)

(一)气-液等温转变的实验曲线　一级相变的特点　(二)范德瓦尔斯等温线的热力学分析　(三)正常导体和超导体之间的转变　(四)相变的分类　(五)二级相变的其他例子

§7.2.10　热力学第三定律 …… (588)

(一)热力学第三定律的能斯脱表述及其统计论证　(二)热力学第三定律的推论

习题 …… (590)

第三章　统计系综 …… (595)

§7.3.1　正则系综和正则分布 …… (595)

(一)正则系综和正则分布　(二)配分函数和热力学量的关系　(三)恒温系统的能量涨落

§7.3.2　固体的顺磁性模型和固体比热的爱因斯坦模型 …… (600)

(一)固体的顺磁性模型　(二)固体比热的爱因斯坦模型

§7.3.3　巨正则分布系综和巨正则分布 …… (606)

(一)巨正则系综和巨正则分布　(二)巨配分函数和热力学量的关系　(三)开放系统的粒子数涨落　(四)巨正则分布的简单应用

§7.3.4　理想气体的三种分布 …… (612)

(一)理想气体及其微观态的描述　单粒子态占有数集合　(二)用巨正则系综推导 F-D 分布和 B-E 分布　(三)由 F-D 分布和 B-E 分布过渡到 M-B 分布　(四)对非简并条件的讨论　(五)玻尔兹曼气体的配分函数

§7.3.5　三种系综的等价　其他系综 …… (620)

(一)三种系综的等价　(二)三种系综等价的解释　选用不同系综和进行勒让德变换的关系　*(三) T-p 系综

习题 …… (624)

第八篇　理想气体 …… (629)

第一章　玻尔兹曼气体 …… (631)

§8.1.1 由配分函数求热力学量 …… (631)
(一)分子配分函数的分解 (二)μ空间 分子配分函数在μ空间中的表示 (三)单原子理想气体的热力学量
§8.1.2 双原子分子气体的热容量 …… (636)
(一)转动和振动的配分函数和热容量 (二)对$C_V^{转}$和$C_V^{振}$的讨论及与实验的比较
§8.1.3 由玻尔兹曼分布求速度分布和空间分布 …… (640)
(一)麦克斯韦速度分布 (二)重力场中分子随高度的分布 (三)气体电介质的极化
§8.1.4 能量均分定理 …… (644)
§8.1.5 稀薄实际气体的状态方程 …… (646)
(一)Γ空间 系综配分函数在Γ空间中的表示 (二)非理想气体的配分函数 位形积分 (三)第二位力系数的计算 分子作用的刚球模型
习题 …… (652)

第二章 玻色气体和费米气体 …… (655)
§8.2.1 黑体幅射 光子统计 …… (655)
(一)黑体辐射的能谱曲线 斯特藩-玻尔兹曼公式 (二)光子气的分布函数和普朗克公式的推导 (三)维恩位移法则和斯特藩-玻尔兹曼定律 (四)黑体辐射的热力学
§8.2.2 晶格振动比热的德拜理论 声子统计 …… (659)
(一)声子概念的引入 (二)德拜频谱和晶格振动的热容量
* §8.2.3 液态^{4}He的性质和朗道的超流理论 …… (664)
(一)液态^{4}He的性质 (二)He Ⅱ的二流体模型 (三)朗道的超流理论 声子和旋子
§8.2.4 玻色-爱因斯坦凝聚 …… (669)
(一)凝聚温度T_C的计算 (二)基态粒子数N_0随温度的变化 (三)玻色理想气体的C_V (四)BEC的实验实现
§8.2.5 金属中的自由电子气 …… (673)
(一)0K下的自由电子气 (二)$T < T_F$时的自由电子气
习题 …… (678)

第九篇 涨落和临界现象的统计理论 …… (681)

第一章 涨落理论 …… (683)

§9.1.1　涨落的准热力学理论 …… (683)
(一)非孤立系统基本热力学量涨落的概率分布公式　(二)热力学量涨落的计算　(三)密度涨落和光的分子散射
§9.1.2　布朗运动 …… (689)
(一)布朗运动的朗之万理论　*(二)布朗粒子的概率分布　福克尔-普朗克方程
习题 …… (693)

第二章　临界现象 …… (695)
§9.2.1　临界指数和标度律 …… (695)
(一)流体系统的临界指数　(二)铁磁系统的临界指数
§9.2.2　二级相变的朗道－金兹堡理论 …… (698)
(一)序参量　(二)朗道－金兹堡理论
* §9.2.3　临界点邻域的涨落和关联 …… (702)
习题 …… (705)

第三章　临界现象的统计理论 …… (707)
§9.3.1　Ising 模型 …… (707)
(一)Ising 模型　(二)一维 Ising 模型的矩阵解法
§9.3.2　标度理论 …… (711)
(一)广义齐次函数　(二)Widom 的标度理论　(三)Kadanoff 的标度变换理论　(四)Fisher 和 Josephson 标度关系
* §9.3.3　重整化群方法简介 …… (717)
习题 …… (721)

第十篇　非平衡统计基础 …… (723)

第一章　稀薄气体中的输运过程 …… (725)
§10.1.1　玻尔兹曼方程 …… (725)
§10.1.2　气体粘滞系数和金属电导率的计算 …… (728)
(一)气体粘滞系数的计算　(二)金属电导率的计算
* §10.1.3　玻尔兹曼积分微分方程 …… (733)
(一)分子的弹性碰撞　(二)$(\partial f_1/\partial t)_{\perp}\,\mathrm{d}^3r\mathrm{d}^3v\delta t$　(三)两分子碰撞前后 $\mathrm{d}\boldsymbol{v}^3\mathrm{d}\boldsymbol{v}_1^3 = \mathrm{d}\boldsymbol{v}'^3\mathrm{d}\boldsymbol{v}_1'^3$ 的证明
* §10.1.4　H 定理 …… (737)

(一)H 定理 (二)细致平衡原理 (三)麦克斯韦速度分布 (四)H 函数与熵函数的关系 (五)H 定理的统计性质

习题 …… (741)

*第二章 BBGKY 序列 …… (743)

§10.2.1 约化概率密度 BBGKY 序列 …… (743)

§10.2.2 符拉索夫方程和玻尔兹曼方程的推导 …… (745)

(一)符拉索夫方程的推导 (二)玻尔兹曼积分微分方程的推导

附录 …… (749)

附录Ⅰ δ符号和δ函数 …… (751)

(一)克隆尼克δ符号 (二)三阶全反对称ε符号 (三)狄拉克δ函数

附录Ⅱ 矢量分析与张量计算初步 …… (755)

(一)梯度、散度、旋度的定义 (二)高阶导数与乘积的导数 (三)积分降维公式与格林定理 (四)正交曲线坐标中的矢量微分公式 (五)张量计算简介 (六)三维δ函数的一个表达式

附录Ⅲ 概率的基础知识 …… (764)

(一)随机事件和概率 (二)随机事件的概率分布 统计平均值 涨落 (三)二项式分布、泊松分布和高斯分布

附录Ⅳ 拉普拉斯方程的通解 …… (771)

(一)方程的分离变量 (二)勒让德多项式 (三)缔合勒让德函数 (四)球函数 (五)拉普拉斯方程的通解 (六)勒让德多项式的母函数

附录Ⅴ 一些积分公式 …… (776)

(一)高斯型积分 (二)高斯分布的矩 (三)玻色分布的积分 (四)费米分布的积分 (五)Γ函数 (六)雅可比行列式 (七)衍射积分

附录Ⅵ 电磁量的国际单位制与高斯单位制 …… (780)

(一)国际单位制和高斯单位制的基本量和基本单位 (二)国际单位制与高斯单位制之间的单位换算关系 (三)国际单位制与高斯单位制的重要常量及公式对照表

附录Ⅶ 物理常量与单位换算 …… (783)



绪 论

人类对自然界的认识是由感性到理性，从现象到本质，不断深入发展的过程. 自然科学的任务就是要在总结实验现象的基础上，深入进行理论分析，从感性认识发展到理性认识，揭示自然界的基本规律. 只有对自然界的基本规律有了清楚的了解，才能有力地改造自然，征服自然.

通过"普通物理"的学习，我们已经对物理实验现象有了比较全面的、定量的认识，并且对基本物理规律有了初步的了解. 在此基础上进一步学习"理论物理"就是要对基本物理规律得到更系统更完整的理解，并初步掌握用这些规律解决实际问题的方法.

物理规律和其它科学规律一样，不是在人们头脑中凭空产生，而是从大量实验事实中总结提高得到，并不断接受实验检验的. 然而，物理规律一旦建立起来以后，就要求能够从很少的一些假设（原理）出发，通过逻辑推理和数学推导，建立起整个物理学. 在学习理论物理的过程中，应当熟悉和掌握这种抽象的思维方法和处理问题方法.

人类对自然界的认识是永无止境的.在任何时候,人们所认识到的自然规律都只是客观规律的一种近似反映,有它一定的适用范围.在十九世纪末,经典力学和经典电动力学有了充分的发展,以致于有些物理学家以为已经建立起了最终的理论.然而,刚刚跨进二十世纪,经典物理和实验的严重矛盾就显露出来,推动着物理学进入以相对论和量子力学为代表的新时期.相对论和量子力学的发展对于社会生产起了巨大的推动作用,取得了辉煌的成就,然而它们的局限性在今天已经开始显现出来.未来的理论物理必将在它们的基础上有新的突破.

本书的目的是按照物质运动形式,从低速到高速、从宏观到微观、从单体到多体,对当前人们所认识到的理论物理的基本规律和基本方法作一个初步的系统介绍.



第一篇　低速宏观物体的运动

最简单的物质运动形态是机械运动，即物体各部分之间以及一个物体相对于其他物体的位置变化.

早在17世纪，牛顿就以"牛顿三定律"的形式总结了机械运动的基本规律，奠定了经典力学的基础. 进入20世纪以后，人们认识到牛顿力学只是客观物质运动规律的一种近似表述. 它只能描述速度远低于光速，特征尺度远大于原子尺度，即

$$v \ll c \approx 3 \times 10^{10}\text{cm/s}, \quad l \gg a \approx 10^{-8}\text{cm}$$

的低速宏观物体的运动. 然而，由于人们日常接触到的事物大量的是这种低速宏观物体，所以直到今天，经典力学仍然在人类的生产和生活中起着重要的指导作用. 另一方面，高速运动的规律（相对论）和微观运动的规律（量子力学）都是在经典力学的基础上建立起来的，并且以经典力学为它们的极限情况. 因此，深入地学习和掌握经典力学是十分必要的.

在这一篇里，我们讨论由经典力学描述的低速宏观物体的运动.



第一章　低速宏观运动的基本原理

经典力学的原理，除了用“牛顿三定律”表述以外，还有其他表述方式，例如：拉格朗日方式和哈密顿方式. 这些表述方式实质内容相同，只是表达形式有差别. 拉格朗日表述方式和哈密顿表述方式比牛顿表述方式站的角度更高，对客观规律的本质反映得更清楚，而且便于推广到高速和微观情况. 因此，在这一篇里我们将采用这两种方式表述经典力学的原理，并讨论它们的一些应用.

第一章讨论按拉格朗日方式表述的经典力学基本原理. 在第一节里，先就一种最简单的情况——单个质点在保守力场中的无约束运动，建立拉格朗日方程. 即使在这种简单情况下也可以看出，按拉格朗日方式建立运动方程比按牛顿方式建立运动方程有明显的优点. 在第二节再讨论更普遍的情况——有约束的情况. 在这种情况下，拉格朗日方式的优越性进一步表露出来. 在第三节里，将前两节的讨论提高到基本原理的高度，说明如何将经典力学原理表述成最小作用量原理的形式. 最小作用量原理是现代物理学原理的普遍形式. 除经典力学系统以外，相对论性力学系统和场论系统的基本原理都可以表述为最小作用量原理的形式. 这种形式的表述，深刻地反映了各种物理系统的本质，并为将它们量子化提供了有利的途径. 在这一章的第三节里，我们讨论经典力学的最小作用量原理，为今后的进一步学习打下基础. 最后，在第四节里，我们讨论惯性参考系和伽利略相对性原理，并以单个自由质点为例，说明如何从系统的基本物理性质（如：时空的均匀性和各向同性、运动的相对性等）出发，得到力学系统的拉格朗日函数. 相互作用质点系的拉格朗日函数将在 §1.2.1 中给出.

§1.1.1　无约束质点的拉格朗日方程

考虑一个不受约束的质点在保守力场中的运动. 质点的位置可以用笛卡儿坐标 (x,y,z) 表示：

$$\boldsymbol{r} = i x + j y + k z. \tag{1.1}$$

保守力场的势能是 $U(\boldsymbol{r})$. 在此力场中，质点受力

$$\boldsymbol{F} = -\frac{\partial U}{\partial \boldsymbol{r}}. \tag{1.2}$$

符号 $\partial U/\partial \boldsymbol{r}$ 表示 $U(\boldsymbol{r})$ 的梯度

$$\frac{\partial U}{\partial \boldsymbol{r}} = i\frac{\partial U}{\partial x} + j\frac{\partial U}{\partial y} + k\frac{\partial U}{\partial z}. \tag{1.3}$$

按照牛顿第二定律，这一质点的运动方程是

$$m\ddot{\boldsymbol{r}} = \boldsymbol{F} = -\partial U/\partial \boldsymbol{r}, \tag{1.4}$$

写成分量形式就是：

$$\left.\begin{aligned} m\ddot{x} &= F_x = -\partial U/\partial x, \\ m\ddot{y} &= F_y = -\partial U/\partial y, \\ m\ddot{z} &= F_z = -\partial U/\partial z. \end{aligned}\right\} \tag{1.5}$$

为了书写方便，以下将笛卡儿坐标 x,y,z 写成 x_1,x_2,x_3．(1.5)式写成

$$\left.\begin{aligned} m\ddot{x}_1 &= F_1 = -\partial U/\partial x_1, \\ m\ddot{x}_2 &= F_2 = -\partial U/\partial x_2, \\ m\ddot{x}_3 &= F_3 = -\partial U/\partial x_3. \end{aligned}\right\} \tag{1.6}$$

或合写为

$$m\ddot{x}_i = F_i = -\partial U/\partial x_i, \qquad i = 1,2,3. \tag{1.7}$$

在某些情况下，采用笛卡儿坐标不很方便．例如，当质点处于有心力场中时，势能只是质点距力心距离的函数，而与方向无关：

$$U = U(r). \tag{1.8}$$

在此情况下，更方便的是采用球坐标 (r,θ,φ)，它和笛卡儿坐标的关系是(见图1)：

$$\left.\begin{aligned} x_1 \equiv x &= r\sin\theta\cos\varphi, \\ x_2 \equiv y &= r\sin\theta\sin\varphi, \\ x_3 \equiv z &= r\cos\theta. \end{aligned}\right\} \tag{1.9}$$

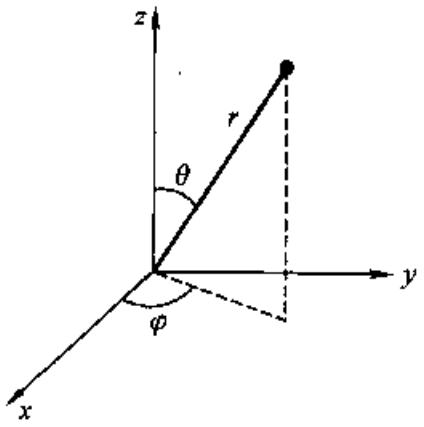


图1　球坐标

因此，有必要写出球坐标中的运动方程来代替笛卡儿坐标中的运动方程(1.7)式．

从牛顿运动方程(1.7)式出发，用(1.9)式进行变量代换，可以得到球坐标中的运动方程．但是，这样做手续较繁，而且显然是走了弯路．因为，在处理不同问题的时候，采用什么坐标系更方便，应该具体问题具体分析，而不应先入为主地把笛卡儿坐标放到一个特殊的地位．合理的做法显然不是对任何问题都先建立笛卡儿坐标中的运动方程，然后进行变量变换，而是根据具体问题的性质直接在合适的坐标系中建立运动方程．下面来讨论如何做到这一点．

为了普遍起见，先讨论一般情况，然后再举有心力场为例子．一个任意的无约束质点有三个自由度，所以，需要三个变量才能决定它的位置．这三个变量称为广义坐标，用 q_1,q_2,q_3 表示．它们和笛卡儿坐标之间有函数关系：

$$\left.\begin{aligned} x_1 &= x_1(q_1,q_2,q_3),\\ x_2 &= x_2(q_1,q_2,q_3),\\ x_3 &= x_3(q_1,q_2,q_3). \end{aligned}\right\} \tag{1.10}$$

或简写为(黑体的 $\boldsymbol{q}$ 代表 3 个 q_α,$\alpha = 1,2,3$):

$$x_i = x_i(\boldsymbol{q}), \qquad i = 1,2,3. \tag{1.11}$$

由于 x_i 是时间 t 的函数,所以 q_α 也是时间 t 的函数.

如果直接从矢量型的方程(1.4)出发,进行变量代换,就需要利用(1.11)式计算 $x_i(t)$ 对 t 的二阶导数.为了避免这一复杂性,将方程(1.4)左右两边点乘 $\mathrm{d}\boldsymbol{r}$,使它变成标量型的方程:

$$m\ddot{\boldsymbol{r}}\cdot\mathrm{d}\boldsymbol{r} = \boldsymbol{F}\cdot\mathrm{d}\boldsymbol{r} = -\frac{\partial U}{\partial \boldsymbol{r}}\cdot\mathrm{d}\boldsymbol{r}. \tag{1.12}$$

用分量写出是

$$\sum_{i=1}^{3} m\ddot{x}_i\mathrm{d}x_i = \sum_{i=1}^{3} F_i\mathrm{d}x_i = -\sum_{i=1}^{3}\frac{\partial U}{\partial x_i}\mathrm{d}x_i. \tag{1.13}$$

对(1.11)式求全微分

$$\mathrm{d}x_i = \frac{\partial x_i}{\partial q_1}\mathrm{d}q_1 + \frac{\partial x_i}{\partial q_2}\mathrm{d}q_2 + \frac{\partial x_i}{\partial q_3}\mathrm{d}q_3 = \sum_{\alpha=1}^{3}\frac{\partial x_i}{\partial q_\alpha}\mathrm{d}q_\alpha. \tag{1.14}$$

代入(1.13)式左边,得

$$\sum_{i=1}^{3} m\ddot{x}_i\mathrm{d}x_i = \sum_{i=1}^{3}\sum_{\alpha=1}^{3} m\ddot{x}_i\frac{\partial x_i}{\partial q_\alpha}\mathrm{d}q_\alpha = \sum_{\alpha=1}^{3}\left(\sum_{i=1}^{3} m\ddot{x}_i\frac{\partial x_i}{\partial q_\alpha}\mathrm{d}q_\alpha\right). \tag{1.15}$$

由于 i 和 α 是互不相关的指标,所以对它们的求和次序可以交换.

将此式和

$$\mathrm{d}U = \sum_{i=1}^{3}\frac{\partial U}{\partial x_i}\mathrm{d}x_i = \sum_{\alpha=1}^{3}\frac{\partial U}{\partial q_\alpha}\mathrm{d}q_\alpha \tag{1.16}$$

一道代入(1.13)式,得

$$\sum_{\alpha}\left[\sum_{i} m\ddot{x}_i\frac{\partial x_i}{\partial q_\alpha}\right]\mathrm{d}q_\alpha = -\sum_{\alpha}\frac{\partial U}{\partial q_\alpha}\mathrm{d}q_\alpha. \tag{1.17}$$

由于三个 $\mathrm{d}q_\alpha$ 相互独立,所以上式左右两边 $\mathrm{d}q_\alpha$ 的系数应该分别相等,于是得:

$$\sum_{i} m\ddot{x}_i\frac{\partial x_i}{\partial q_\alpha} = -\frac{\partial U}{\partial q_\alpha}, \qquad \alpha = 1,2,3. \tag{1.18}$$

下一步,要对 $\ddot{x}_i$ 进行变换.为此写出类似于分部积分的恒等式:

$$\sum_{i=1}^{3} m\ddot{x}_i\frac{\partial x_i}{\partial q_\alpha} = \frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=1}^{3}\left(m\dot{x}_i\frac{\partial x_i}{\partial q_\alpha}\right) - \sum_{i=1}^{3} m\dot{x}_i\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial x_i}{\partial q_\alpha}. \tag{1.19}$$

代入(1.18)式得:

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=1}^{3}\left(m\dot{x}_i\frac{\partial x_i}{\partial q_\alpha}\right)-\sum_{i=1}^{3}m\dot{x}_i\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial x_i}{\partial q_\alpha}=-\frac{\partial U}{\partial q_\alpha},\quad \alpha=1,2,3. \tag{1.20}$$

为了进一步化简上式,利用(1.11)式求 x_i 和 $\partial x_i/\partial q_\alpha$ 对 t 的导数:

$$\dot{x}_i=\frac{\mathrm{d}x_i}{\mathrm{d}t}=\sum_{\alpha=1}^{3}\frac{\partial x_i}{\partial q_\alpha}\cdot\frac{\mathrm{d}q_\alpha}{\mathrm{d}t}=\sum_{\alpha=1}^{3}\frac{\partial x_i}{\partial q_\alpha}\dot{q}_\alpha, \tag{1.21}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial x_i}{\partial q_\alpha}=\sum_{\beta=1}^{3}\frac{\partial^2 x_i}{\partial q_\alpha\partial q_\beta}\dot{q}_\beta=\frac{\partial}{\partial q_\alpha}\sum_{\beta=1}^{3}\frac{\partial x_i}{\partial q_\beta}\dot{q}_\beta=\frac{\partial}{\partial q_\alpha}\dot{x}_i. \tag{1.22}$$

(1.21)式右边 $\dot{q}_\alpha$ 的系数 $\partial x_i/\partial q_\alpha$ 只是 q_α 的函数,而不是 $\dot{q}_\alpha$ 的函数,所以将(1.21)式对 $\dot{q}_\alpha$ 求导,得到

$$\partial\dot{x}_i/\partial\dot{q}_\alpha=\partial x_i/\partial q_\alpha. \tag{1.23}$$

将(1.22)和(1.23)式一道代入(1.20)式,得到

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{i=1}^{3}m\dot{x}_i\frac{\partial\dot{x}_i}{\partial\dot{q}_\alpha}\right)-\sum_{i=1}^{3}m\dot{x}_i\frac{\partial\dot{x}_i}{\partial q_\alpha}=-\frac{\partial U}{\partial q_\alpha},\qquad \alpha=1,2,3. \tag{1.24}$$

质量为 m 的粒子的动能是

$$T=mv^2/2=\sum_{i=1}^{3}m\dot{x}_i^2/2. \tag{1.25}$$

将它对 q_α 和 $\dot{q}_\alpha$ 求偏导,

$$\frac{\partial T}{\partial q_\alpha}=\sum_{i=1}^{3}m\dot{x}_i\frac{\partial\dot{x}_i}{\partial q_\alpha},\qquad \frac{\partial T}{\partial\dot{q}_\alpha}=\sum_{i=1}^{3}m\dot{x}_i\frac{\partial\dot{x}_i}{\partial\dot{q}_\alpha}. \tag{1.26}$$

这正好就是(1.24)式左边的两个和式.因此,(1.24)式可以写为

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial\dot{q}_\alpha}-\frac{\partial T}{\partial q_\alpha}=-\frac{\partial U}{\partial q_\alpha},\qquad \alpha=1,2,3. \tag{1.27}$$

对于保守力场,势能 $U(r)$ 只是 q_α 的函数,而不是 $\dot{q}_\alpha$ 的函数,

$$\partial U/\partial\dot{q}_\alpha=0. \tag{1.28}$$

将(1.27)式的右边的项移到左边,利用(1.28)式,得到

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial\dot{q}_\alpha}-\frac{\partial L}{\partial q_\alpha}=0,\qquad \alpha=1,2,3. \tag{1.29}$$

其中

$$L=T-U \tag{1.30}$$

称为拉格朗日函数,而(1.29)式称为拉格朗日方程.

拉格朗日方程(1.29)式可以代替牛顿方程(1.7)式作为经典力学的基本运动方程.以上我们只是就无约束的单质点推出了拉格朗日方程,在下一节中再进行更普遍的推导.

由(1.29)式可见,拉格朗日方程是直接就任意广义坐标 q_α 写出的,排除了笛

卡儿坐标的特殊地位.这是它比牛顿形式的运动方程(1.7)式更为优越的地方.另一方面,从下一节的推导中可以看出,一切保守系的拉格朗日方程都有相同的形式,而不同系统的特点只表现在拉格朗日 L 作为 q_α 和 $\dot{q}_\alpha(\alpha = 1,2,\cdots,s)$ 的函数

$$L = L(q_1,\cdots,q_s,\dot{q}_1,\cdots,\dot{q}_s) \tag{1.31}$$

的函数形式上.(1.31)式是 $q_\alpha,\dot{q}_\alpha$ 的标量函数.用一个标量函数包含力学系统的全部性质,这是拉格朗日表述方式的最大优点,是牛顿方式所做不到的.

现在再来考虑有心力场(1.8)式,作为应用拉格朗日方程的一个简单例子.写拉格朗日方程的第一步是要得到拉格朗日函数(1.30),而为此先要写出球坐标中的动能表达式.考虑位矢为 $\boldsymbol{r}$ 的一点 P,它的球坐标是 (r,θ,φ).在 P 点,沿 r,θ,φ 增加的方向分别作三个单位矢量 $\boldsymbol{e}_r,\boldsymbol{e}_\theta,\boldsymbol{e}_\varphi$.这三个单位矢量相互垂直,形成位于 P 点的一个正交坐标系,称为局部标架.空间的不同点 P 有不同的局部标架,因而不能用它展开从原点 O 出发的位矢 $\boldsymbol{r}$,而只能用它展开从 P 点出发的位移矢量 $\mathrm{d}\boldsymbol{r}$.在现在讨论的球坐标情况下,位移 $\mathrm{d}\boldsymbol{r}$ 在局部标架中展开为

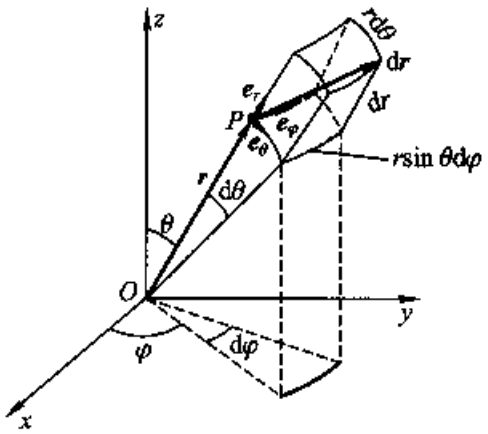


图2　球坐标中的位移

$$\mathrm{d}\boldsymbol{r} = \boldsymbol{e}_r\mathrm{d}r + \boldsymbol{e}_\theta r\mathrm{d}\theta + \boldsymbol{e}_\varphi r\sin\theta\mathrm{d}\varphi, \tag{1.32}$$

如图2.用 $\mathrm{d}t$ 除上式左右两边,得到

$$\dot{\boldsymbol{r}} = \boldsymbol{e}_r\dot{r} + \boldsymbol{e}_\theta r\dot{\theta} + \boldsymbol{e}_\varphi r\sin\theta\dot{\varphi}, \tag{1.33}$$

因而,动能是

$$T = \frac{1}{2}m\dot{\boldsymbol{r}}^2 = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\varphi}^2) \tag{1.34}$$

(由于 $\boldsymbol{e}_r,\boldsymbol{e}_\theta,\boldsymbol{e}_\varphi$ 相互垂直,所以它们的交叉乘积为零).将(1.34)式代入(1.30)式得到拉格朗日函数:

$$L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\varphi}^2) - U(r). \tag{1.35}$$

它的各个偏导数是

$$\frac{\partial L}{\partial\dot{r}} = m\dot{r},\quad \frac{\partial L}{\partial\dot{\theta}} = mr^2\dot{\theta},\quad \frac{\partial L}{\partial\dot{\varphi}} = mr^2\sin^2\theta\dot{\varphi}, \tag{1.36}$$

$$\frac{\partial L}{\partial r} = mr\dot{\theta}^2 + mr\sin^2\theta\dot{\varphi}^2 - \frac{\mathrm{d}U}{\mathrm{d}r}, \tag{1.37}$$

$$\frac{\partial L}{\partial\theta} = mr^2\sin\theta\cos\theta\dot{\varphi}^2, \tag{1.38}$$

$$\frac{\partial L}{\partial \varphi}=0. \tag{1.39}$$

代入拉格朗日方程(1.29),得到

$$\left.\begin{aligned} m\ddot{r} &= -\frac{\mathrm{d}U}{\mathrm{d}r}+mr\dot{\theta}^2+mr\sin^2\theta\dot{\varphi}^2,\\ \frac{\mathrm{d}}{\mathrm{d}t}(mr^2\dot{\theta}) &= mr^2\sin\theta\cos\theta\dot{\varphi}^2,\\ \frac{\mathrm{d}}{\mathrm{d}t}(mr^2\sin^2\theta\dot{\varphi}) &= 0. \end{aligned}\right\} \tag{1.40}$$

这样,我们就直接在球坐标系中写出了质点的运动方程.从牛顿方程(1.7)出发,进行变量代换(1.9),也可以得到这一方程,但推导过程要繁得多.方程(1.40)描述有心力场中质点的运动,这在以下第三章中将要详细讨论.

§1.1.2　有约束情况下的拉格朗日方程

上一节里,我们就不受约束的单个质点在保守力场中的运动,推导出了拉格朗日方程.现在来考虑更一般的情况,但是仍然只考虑保守系.关于有耗散的情况,将在后面(§1.4.2)结合具体例子讨论.

假定一个力学系统由 N 个质点组成,它们的质量和位矢分别是:

$$m_a,\boldsymbol{r}_a,\qquad a=1,2,\cdots,N. \tag{2.1}$$

牛顿运动方程是

$$m_a\ddot{\boldsymbol{r}}_a=\boldsymbol{F}_a,\qquad a=1,2,\cdots,N, \tag{2.2}$$

其中 $\boldsymbol{F}_a$ 是作用在第 a 个质点上的力.

用笛卡儿坐标的分量写出,(2.2)式共 $3N$ 个方程.然而,一般情况下这 $3N$ 个方程并不是独立的.例如,质量为 m 的质点用长为 l 的轻杆挂在一个固定点 O 上,形成一个单摆,则这一质点与 O 之间的距离保持为 l 不变.这就使这一质点的三个坐标只有两个独立.又如,质量为 m_1 和 m_2 的两个重块,安放在两个竖直光滑槽中,并用柔软轻绳通过一个滑轮相连,如图1.在这样的滑槽和软绳的限制下,这一系统(称为阿特伍德机)实际上只有一个自由度.

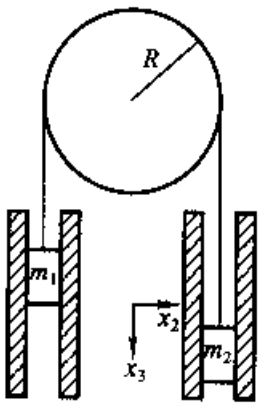


图1　阿特伍德机

以上两个例子中,自由度的数目之所以少于坐标的数目,是由于其中质点的运动受到了约束.在单摆情况下,质点的三个笛卡儿坐标需要满足一个等式

$$x_1^2 + x_2^2 + x_3^2 = l^2, \tag{2.3}$$

因而在三个变量 x_1, x_2, x_3 中，只有两个独立. 在阿特伍德机情况下，两个质点的六个笛卡儿坐标 $x_1^{(1)}, x_2^{(1)}, x_3^{(1)}, x_1^{(2)}, x_2^{(2)}, x_3^{(2)}$ 满足五个等式

$$\left.\begin{aligned} x_1^{(1)} &= x_1^{(2)} = 0, \\ -x_2^{(1)} &= x_2^{(2)} = R, \\ x_3^{(1)} + x_3^{(2)} &= \text{常数}, \end{aligned}\right\} \tag{2.4}$$

因而只有一个独立.

一般说来，如果由 N 个质点组成的力学体系的 $3N$ 个笛卡儿坐标

$$x_i^{(a)}, \qquad a = 1,2,\cdots,N; i = 1,2,3 \tag{2.5}$$

满足 $3N - s$ 个等式（不带指标的 x 代表(2.5)式中的全部 $x_i^{(a)}$）：

$$f_\alpha(x,t) = 0, \qquad \alpha = 1,2,\cdots,3N - s, \tag{2.6}$$

则这一力学体系只有 s 个独立坐标. 等式(2.6)称为约束方程.

如果约束方程中不含 t，则称为稳定约束. 以上所举的单摆和阿特伍德机都是稳定约束的例子. 如果单摆的悬点作简谐振动：

$$X(t) = A\cos \omega t,$$

则约束方程(2.3)式成为

$$(x_1 - A\cos \omega t)^2 + x_2^2 + x_3^2 = l^2. \tag{2.7}$$

这是不稳定约束的例子.

如果悬挂单摆的不是定长的轻杆，而是柔软的轻绳，则质点离悬点的距离不能大于 l，但可以小于 l，

$$x_1^2 + x_2^2 + x_3^2 \leqslant l^2. \tag{2.8}$$

这种由不等式表示的约束称为可解约束，它可以从一个方向上解除. 放在光滑曲面上的质点所受的约束也是可解约束.

约束总是通过一些外界物体（例如轻杆，滑槽，软绳）作用在所研究的系统中的质点上. 在运动过程中，系统中的质点对这些外界物体有作用力，同时受到这些物体的反作用力，称为约束反力. 正是约束反力的作用，迫使质点的坐标满足约束方程. 在运动过程中，约束反力是随时间变化的，不能预先知道，只能通过解运动方程求得. 但是，很多时候，我们并不关心约束反力，而只希望知道系统的运动情况. 因此，有必要从运动方程中消去约束反力，而只留下 s 个独立坐标所满足的方程.

对于一个有 s 个自由度的系统，可以由 $3N$ 个笛卡儿坐标组成 s 个函数 $q_\alpha, \alpha = 1,2,\cdots,s$ 作为描述系统运动的变量. 反过来 $3N$ 个笛卡儿坐标成为这 s 个变量的函数：

$$x_i^{(a)} = x_i^{(a)}(q_1, q_2, \cdots, q_s, t),$$

$$a = 1,2,\cdots,N; \qquad i = 1,2,3. \tag{2.9}$$

这一函数关系满足约束方程(2.6).这就是说,将(2.9)式代入约束方程(2.6),会使它成为恒等式.这样的 s 个变量 $q_\alpha(\alpha = 1,2,\cdots,s)$ 是决定系统中所有质点位置的独立变量,称为系统的广义坐标.

现在来推导广义坐标的运动方程,出发点是牛顿方程(2.2)式.

将作用在第 a 个质点上的力 $\boldsymbol{F}^{(a)}$ 分为主动力 $\boldsymbol{F}_{主}^{(a)}$ 和约束反力 $\boldsymbol{F}_{反}^{(a)}$ 两部分,即

$$\boldsymbol{F}^{(a)} = \boldsymbol{F}_{主}^{(a)} + \boldsymbol{F}_{反}^{(a)}. \tag{2.10}$$

运动方程(2.2)式成为

$$m_a\ddot{\boldsymbol{r}}^{(a)} = \boldsymbol{F}_{主}^{(a)} + \boldsymbol{F}_{反}^{(a)}. \tag{2.11}$$

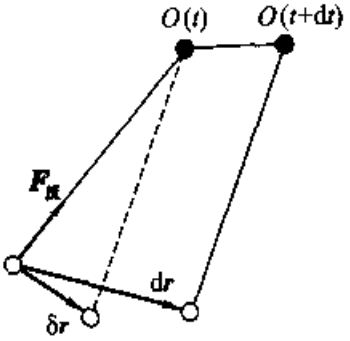


图 2 虚位移 δr 与实位移 dr

怎样能消掉约束反力呢?为了回答这一问题,先来分析几个例子.

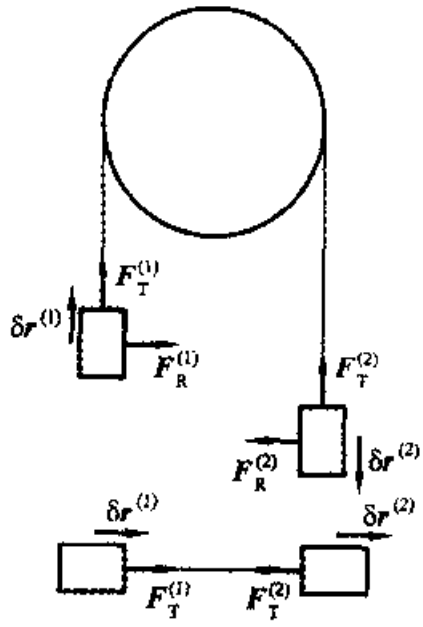


图 3　阿特伍德机的虚位移与约束反力

对于悬点 O 固定的单摆,质点只能在以 O 为心,以摆长 l 为半径的球面上运动.在任一时刻,约束所容许的位移称为虚位移,用 δr 表示.对单摆,虚位移在半径为 l 的球面的切面上,而约束反力沿此球面的半径,两者相互垂直,因而约束反力和虚位移的点积为零:

$$\boldsymbol{F}_{反} \cdot \delta\boldsymbol{r} = 0. \tag{2.12}$$

要注意,虚位移是某一时刻约束所容许的位移,它和实际发生的位移 $d\boldsymbol{r}$ 有区别.特别是在不稳定约束情况下,这一区别看得更明显.以悬点作谐振动的单摆为例,虚位移在以 t 时刻悬点所在位置 $O(t)$ 为心的球面上,而实际位移的始点和终点分别在以 $O(t)$ 和 $O(t + dt)$ 为心的两个球面上,如图 2.由图可见,约束反力 $\boldsymbol{F}_{反}$ 和虚位移 $\delta\boldsymbol{r}$ 垂直,而和实位移并不垂直.

一个力和虚位移的点积称为这个力所做的虚功.(2.12)式表明,作用在单摆上的约束反力所作的虚功为零.

再来看图 1 中的阿特伍德机.两个质点的虚位移 $\delta\boldsymbol{r}^{(1)}$ 和 $\delta\boldsymbol{r}^{(2)}$ 分别沿两个滑槽方向,而滑槽所施的约束反力 $\boldsymbol{F}_R^{(1)}$ 和 $\boldsymbol{F}_R^{(2)}$ 垂直于滑槽表面(不考虑摩擦力),因而 $\boldsymbol{F}_R^{(1)} \perp \delta\boldsymbol{r}^{(1)}$,$\boldsymbol{F}_R^{(2)} \perp \delta\boldsymbol{r}^{(2)}$,于是

$$\boldsymbol{F}_R^{(1)} \cdot \delta\boldsymbol{r}^{(1)} = 0, \qquad \boldsymbol{F}_R^{(2)} \cdot \delta\boldsymbol{r}^{(2)} = 0. \tag{2.13}$$

然而，软绳所施的约束反力 $\boldsymbol{F}_{\mathrm{T}}^{(1)}$，$\boldsymbol{F}_{\mathrm{T}}^{(2)}$ 分别和虚位移 $\delta\boldsymbol{r}^{(1)}$，$\delta\boldsymbol{r}^{(2)}$ 平行和反平行，而并不和它们垂直，如图 3，因而 $\boldsymbol{F}_{T}^{(1)}\cdot\delta\boldsymbol{r}^{(1)}\neq 0$，$\boldsymbol{F}_{T}^{(2)}\cdot\delta\boldsymbol{r}^{(2)}\neq 0$. 但是，它们两者之和为零. 为了看到这一点，注意到无摩擦的滑轮只起改变力和位移的方向的作用，所以图 3 的上图中阿特伍德机的受力情况，可以等效于图 3 的下图. 由此图可见，因为绳不能伸长，所以 $\delta\boldsymbol{r}^{(1)}=\delta\boldsymbol{r}^{(2)}$，又根据牛顿第三定律，$\boldsymbol{F}_{\mathrm{T}}^{(1)}=-\boldsymbol{F}_{\mathrm{T}}^{(2)}$，因此

$$\boldsymbol{F}_{\mathrm{T}}^{(1)}\cdot\delta\boldsymbol{r}^{(1)}+\boldsymbol{F}_{\mathrm{T}}^{(2)}\cdot\delta\boldsymbol{r}^{(2)}=0. \tag{2.14}$$

我们看到，在理想的无耗散力情况下，或者是约束反力本身所做的虚功为零，或者是各个质点所受到的约束反力所作的虚功之和为零. 显然，前一种情况也满足后一条件. 因此，可以一般地写出：

$$\sum_{a=1}^{N}\boldsymbol{F}_{\text{反}}^{(a)}\cdot\delta\boldsymbol{r}^{(a)}=0. \tag{2.15}$$

凡是满足条件(2.15)的约束称为理想约束.

结合(2.11)式和(2.15)式可以得到

$$\sum_{a=1}^{N}(\boldsymbol{F}_{\text{主}}^{(a)}-m_a\ddot{\boldsymbol{r}}^{(a)})\cdot\delta\boldsymbol{r}^{(a)}=0. \tag{2.16}$$

这就是说，在理想约束情况下，作用在各个质点上的主动力和惯性力所作的虚功之和为零. 这称为推广的虚功原理. 原来意义下的"虚功原理"是就平衡系统而言的，因而没有惯性力 $-m\ddot{\boldsymbol{r}}^{(a)}$ 这一项. (2.16)式将虚功原理推广到非平衡情况，将惯性力也包括进去.

如果是保守系，则有势能 U：

$$U=U(\boldsymbol{r}^{(1)},\boldsymbol{r}^{(2)},\cdots,\boldsymbol{r}^{(N)}), \tag{2.17}$$

而作用在第 a 个质点上的主动力是 U 对 $\boldsymbol{r}^{(a)}$ 的梯度：

$$\boldsymbol{F}_{\text{主}}^{(a)}=-\partial U/\partial\boldsymbol{r}^{(a)}. \tag{2.18}$$

(2.16)式，(2.18)式和上节的(1.12)式相似，只是在(2.16)式中实位移 $\mathrm{d}\boldsymbol{r}$ 换成了虚位移 $\delta\boldsymbol{r}$，而且还要对所有的质点求和. 变量变换(2.9)式也和上节的(1.11)式相似. 因此，可以进行和上节同样的推导. 但是，要注意(2.9)式和上节(1.11)式的差别：在 $x_i^{(a)}$ 的变量中有时间 t，即考虑了不稳定约束的情况. 因此，在对(2.9)式求微分时，比上节的(1.14)式多出了一个对 t 求导项：

$$\mathrm{d}x_i^{(a)}=\sum_{\alpha=1}^{s}\frac{\partial x_i^{(a)}}{\partial q_\alpha}\mathrm{d}q_\alpha+\frac{\partial x_i^{(a)}}{\partial t}\mathrm{d}t. \tag{2.19}$$

但是，出现在(2.16)式中的是虚位移 $\delta\boldsymbol{r}^{(a)}$ 而不是微分 $\mathrm{d}\boldsymbol{r}^{(a)}$. 根据虚位移的定义，在进行虚位移时，约束固定在某一时刻 t 不变，见图 2. 因此，在求虚位移时，(2.9)式中的 t 应视为常数，$\mathrm{d}t=0$，因而

$$\delta x_i^{(a)} = \sum_{\alpha=1}^{s} \frac{\partial x_i^{(a)}}{\partial q_\alpha} \delta q_\alpha. \tag{2.20}$$

和上节的(1.14)式相同. 这样一来,上节中对拉格朗日方程的推导完全适用于现在的情况,因而,类似于上节的(1.29)式和(1.30)式,得到拉格朗日方程:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_\alpha} - \frac{\partial L}{\partial q_\alpha} = 0, \qquad \alpha = 1,2,\cdots,s, \tag{2.21}$$

$$L = T - U. \tag{2.22}$$

有时简称拉格朗方程为拉氏方程,称拉格朗日 L 为拉氏量.

利用拉格朗日方程写出单摆和阿特伍德机的运动方程很容易. 我们来考虑一个略为复杂一点的例子.

【例】 质量为 M 的重块系在弹簧上,可以在 x 方向振动. 它的下面挂一长为 l,质量为 m 的单摆,如图4. 求这一系统的运动方程.

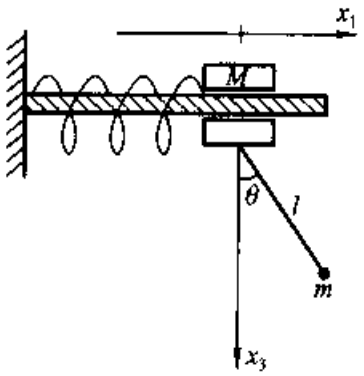


图4　振子与摆复合系统

解　这一系统只有三个自由度,可以选广义坐标为:质点 M 的坐标 X 和质点 m 相对于质点 M 的球面坐标 θ,φ. 质点 m 的笛卡儿坐标是

$$\left.\begin{aligned} x_1 &= l\sin\theta\cos\varphi + X, \\ x_2 &= l\sin\theta\sin\varphi, \\ x_3 &= l\cos\theta. \end{aligned}\right\} \tag{2.23}$$

系统的动能是

$$\begin{aligned} T &= \frac{1}{2}M\dot{X}^2 + \frac{1}{2}m(\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2) \\ &= \frac{1}{2}M\dot{X}^2 + \frac{1}{2}m\dot{X}^2 + \frac{1}{2}ml^2(\dot{\theta}^2 + \sin^2\theta\dot{\varphi}^2) \\ &\quad + ml\dot{X}(\cos\theta\cos\varphi\dot{\theta} - \sin\theta\sin\varphi\dot{\varphi}). \end{aligned} \tag{2.24}$$

势能 U 是弹性势能和重力势能之和

$$U = \frac{1}{2}kX^2 + mgl(1 - \cos\theta). \tag{2.25}$$

拉格朗日函数 $L = T - U$,而拉格朗日方程是

$$(M+m)\frac{\mathrm{d}\dot{X}}{\mathrm{d}t} + ml\frac{\mathrm{d}(\cos\theta\cos\varphi\dot{\theta} - \sin\theta\sin\varphi\dot{\varphi})}{\mathrm{d}t} = -kX, \tag{2.26}$$

$$ml\frac{\mathrm{d}(l\dot{\theta} + \dot{X}\cos\theta\cos\varphi)}{\mathrm{d}t} = ml^2\sin\theta\cos\theta\dot{\varphi}^2 - ml\dot{X}(\sin\theta\cos\varphi\dot{\theta} + \cos\theta\sin\varphi\dot{\varphi}) - mgl\sin\theta, \tag{2.27}$$

$$ml\frac{\mathrm{d}(l\sin^2\theta\dot{\varphi}-\dot{X}\sin\theta\sin\varphi)}{\mathrm{d}t}=-ml\dot{X}(\cos\theta\sin\varphi\dot{\theta}+\sin\theta\cos\varphi\dot{\varphi}).\quad(2.28)$$

这一组运动方程也可以直接从牛顿方程出发推导出来，但比以上利用拉格朗日方程的推导要复杂得多.

最后，让我们指出，在得到拉格朗日方程(2.21)式时，用了(2.17)式，即假定了是保守系，势能 U 不随时间变化. 所得结果可以推广到外场随时间变化的情况，

$$U=U(\boldsymbol{r}^{(1)},\boldsymbol{r}^{(2)},\cdots,\boldsymbol{r}^{(N)},t).\quad(2.29)$$

由于对虚位移而言，$\mathrm{d}t=0$，所以虚位移引起的 $U(\boldsymbol{r}^{(1)},\cdots,\boldsymbol{r}^{(N)},t)$ 的变化仍然和上节的(1.16)式类似：

$$\delta U=\sum_{\alpha=1}^{s}\frac{\partial U}{\partial q_\alpha}\delta q_\alpha.\quad(2.30)$$

因而以上的推导可以完全重复，得到拉格朗日方程(2.21)式. 当外场随时间变化时，系统是非保守系，但只要不存在耗散力，在每一时刻可以写出那一时刻的势能，如(2.29)式，拉格朗日方程(2.21)式就仍然成立. 当存在耗散力时，拉格朗日方程(2.21)式不成立. 在§1.4.2中将看到一个如何处理这类问题的例子.

§1.1.3　最小作用量原理

在上两节里，我们从牛顿方程出发推导了拉格朗日方程. 拉格朗日方程能够代替牛顿方程，作为经典力学的基本运动方程. 因此，不应该把它建立在牛顿方程的基础上，而应该有更根本的物理原理作为它的基础.

归根结底，建立运动方程的目的是通过对它求解来得知系统的真实运动状态. 我们现在的目标是要寻找一个比运动方程更为基本的物理原理，利用它可以不通过求解运动方程而直接回答同样的问题：什么是系统的真实运动状态？

对于由 N 个质点组成的系统而言，它的运动状态由这些质点的坐标作为 t 的函数

$$\boldsymbol{r}^{(\alpha)}=\boldsymbol{r}^{(\alpha)}(t),\qquad\alpha=1,2,\cdots,N\quad(3.1)$$

来描述. 如果有约束，使系统的自由度降为 s，则系统的运动状态由 s 个广义坐标作为 t 的函数

$$q^{(\alpha)}=q^{(\alpha)}(t),\qquad\alpha=1,2,\cdots,s\quad(3.2)$$

来描述. 任意的一组 s 个函数 $q^{(\alpha)}(t)(\alpha=1,2,\cdots,s)$ 描述系统的一个满足约束条件的运动状态. 但是，系统的运动还必须满足初条件. 初条件可以是给定某一时刻 t_1 的 $q^{(\alpha)}$ 和 $\dot{q}^{(\alpha)}$ 的值：

$$q^{(\alpha)}(t_1)=A_\alpha,\quad\dot{q}^{(\alpha)}(t_1)=B_\alpha,\qquad\alpha=1,\cdots,s,\quad(3.3)$$

也可以是给定某两个时刻 t_1 和 t_2 的 $q^{(\alpha)}$ 的值：

$$q^{(\alpha)}(t_1) = q_1^{(\alpha)}, \quad q^{(\alpha)}(t_2) = q_2^{(\alpha)}, \qquad \alpha = 1,\cdots,s. \tag{3.4}$$

下面采用后一种“初条件”.满足这一条件的任意一组函数 $q^{(\alpha)}(t)(\alpha = 1,\cdots,s)$ 都是系统的一个可能的运动状态.我们的任务就是要从所有可能的函数组 $q^{(\alpha)}(t)$ 中挑选出描述真实运动的那一组；也就是说，要从所有满足条件(3.4)的运动状态中挑选出实际发生的真实的运动状态.

问题的这一提法具有超出经典力学范围的普遍意义.考虑一个任意的物理系统，它不一定是经典力学系统，也可以是相对论性力学系统、电磁系统、以及其他更为复杂的系统.这些不同系统的运动状态有不同的具体描述方式.例如：力学系统的运动状态用广义坐标作为时间的函数 $q^{(\alpha)} = q^{(\alpha)}(t)(\alpha = 1,2,\cdots,s)$ 描述；电磁系统的运动状态用电场强度和磁场强度作为坐标和时间的函数：$\boldsymbol{E} = \boldsymbol{E}(\boldsymbol{r},t)$，$\boldsymbol{H} = \boldsymbol{H}(\boldsymbol{r},t)$ 描述.相应地，不同物理系统的运动规律的具体形式很不相同.对于经典力学系统是牛顿方程；对于电磁系统是麦克斯韦方程.但是，对于所有这些系统，物理学的根本任务都是：从一切可能的运动状态中挑选出真实的运动状态.

按照这一具有高度普遍性的提法，应该能够不通过解运动方程而直接给出对基本问题“什么是真实运动状态”的回答.为此，只能采用比较的方法，即：将所有可能的运动状态加以比较，从中找出真实的运动状态.但是，运动状态[对于力学系统就是满足条件(3.4)的 $q^{(\alpha)}(t)$]本身是不可比的.为了比较它们，最简单的方法是让它们每一个对应一个数，然后来比较这些数.

举一个粗浅的例子.好比各地方同时进行选举.不同地方的候选人各不相同，但都有一个共同的挑选方式，即让每个候选人对应一个数(得票数)，然后在各个地方都挑选出这个数字最大的候选者.又比如，托一个朋友上商场去买日用品：一支牙膏、一块肥皂、一个练习本，等等.但是商场里的货物品种繁多，怎样用一句话表明请他买什么？这一句话要能适用于所有不同种类的物品.显然，指明厂牌是行不通的，因为一种厂牌不能同时适用于不同的物品.为了能适用于挑选所有各种不同物品，只能对他说：让每一件物品对应一个数(标价)，然后对每种物品挑选所对应的数最小的那一个.

与此类似，对所有各种不同物理系统都适用的“从一切可能的运动状态中挑选出真实运动状态”的办法是：让每一个可能的运动状态对应一个数，真实的运动状态是所对应的数最小的那一个运动状态.

和每个运动状态所对应的这个数称为作用量.在物理学中，描述任意系统运动规律的基本原理是：一个物理系统实际发生的真实运动状态是所对应的作用量具有最小值的那个状态.这称为最小作用量原理.它具有非常广泛的适用性.尽管描述不同系统运动状态的函数各不相同，相应的运动方程千差万别，但是，它们都服

从最小作用量原理,它们的运动方程都可以通过最小作用量原理确定下来.

现在回过头来考虑经典力学系统.一个力学系统的运动状态用广义坐标作为时间的函数 $\boldsymbol{q}=\boldsymbol{q}(t)$ 描述.这里,为了书写简单,用黑体字母 $\boldsymbol{q}(t)$ 表示 s 个广义坐标 $q_\alpha(t)(\alpha=1,2,\cdots,s)$.每个满足初条件(3.4)的函数 $\boldsymbol{q}(t)$ 描述一个可能的运动状态.它对应于一个数——作用量,用 S 表示.要注意的是,和状态 $\boldsymbol{q}=\boldsymbol{q}(t)$ 对应的作用量 S 不是依赖于某一时刻 t 的 $\boldsymbol{q}$ 值,而是依赖于整个 $\boldsymbol{q}(t)$ 的函数形式.这在数学上可以表述为:在 S 对 $\boldsymbol{q}(t)$ 的依赖关系中,变量 t 已被积分掉.

具体说来,可以将作用量 S 表示成积分:

$$S=\int_{t_1}^{t_2}L(\boldsymbol{q}(t),\dot{\boldsymbol{q}}(t),t)\,\mathrm{d}t. \tag{3.5}$$

这里,$L(\boldsymbol{q}(t),\dot{\boldsymbol{q}}(t),t)$ 是广义坐标 $\boldsymbol{q}(t)$,广义速度 $\dot{\boldsymbol{q}}(t)$ 和时间 t 的一个函数,它决定了系统的全部力学性质,称为系统的拉格朗日函数(或称为系统的拉格朗日量,简称拉氏量).在拉格朗日函数中既有广义坐标 $\boldsymbol{q}(t)$ 又有它的一阶导数(广义速度)$\dot{\boldsymbol{q}}(t)$,而没有更高阶的导数是由于,只要给定初始时刻的位置和速度(而不需要加速度)就决定了系统的运动.

特别要注意的是拉格朗日函数 L 和作用量 S 在性质上的区别.它们两者都依赖于状态函数 $\boldsymbol{q}(t)$,但拉格朗日函数 L 是通过 $\boldsymbol{q}(t)$ 和它的导数 $\dot{\boldsymbol{q}}(t)$ 而依赖于 t,是 t 的一个复合函数;而在作用量 S 中,变量 t 已被积分掉,因而它不再依赖于 t,而只依赖于整个 $\boldsymbol{q}(t)$ 的函数形式.

我们知道,函数 $y=f(x)$ 可以看成由实数 x 到实数 y 的映射.与此类似,作用量可以看成是由函数 $\boldsymbol{q}(t)$ 到实数 S 的一个映射.所不同的是,在现在情况下"自变量"不是一个数,而是一个(一组)函数.这种由函数到数的映射称为泛函,常写为:$S=S[\boldsymbol{q}(t)]$,其中的"自变量" $\boldsymbol{q}(t)$ 用方括号括起来,以区别于在通常的函数 $y=f(x)$ 中用圆括号括起来的自变量 x.

现在来看,如何利用最小作用量原理从一切可能的运动状态中挑选出真实的运动状态.

满足条件(3.4)的任意一组(s 个)函数 $\boldsymbol{q}(t)$ 描述了系统的一个可能的运动状态.其中使作用量(3.5)式有极小值的那一组,就是系统实际发生的真实运动状态.因此,问题归结为在条件(3.4)下求泛函(3.5)的极值.

正像通常的函数 $y=f(x)$ 的极值条件是它的微分为零 $\mathrm{d}y=0$ 一样,泛函 $S=S[\boldsymbol{q}(t)]$ 的极值条件是它的变分 $\delta S=0$.

为了说明什么是泛函的变分,让我们回忆一下函数的微分的定义.函数 $y=f(x)$ 的微分 $\mathrm{d}f(x)$ 是由自变量 x 的一个微小改变 $\mathrm{d}x$ 所引起的函数 $f(x)$ 的准确到 $\mathrm{d}x$ 一次的变化:

$$f(x+\mathrm{d}x)-f(x)=\mathrm{d}f(x)+o(\mathrm{d}x).$$

与此类似,泛函 $S=S[\boldsymbol{q}(t)]$ 的变分 $\delta S[\boldsymbol{q}(t)]$ 是由它的"自变量" $\boldsymbol{q}(t)$ 的一个微小改变 $\delta\boldsymbol{q}(t)$ 所引起的准确到 $\delta\boldsymbol{q}(t)$ 一次的变化,

$$S[\boldsymbol{q}(t)+\delta\boldsymbol{q}(t)]-S[\boldsymbol{q}(t)]=\delta S[\boldsymbol{q}(t)]+o[\delta\boldsymbol{q}(t)].$$

这样,我们所感兴趣的,"从一切可能的运动状态中挑选出真实运动状态"的问题就归结为:在满足条件(3.4)的一切函数中挑选出满足条件:

$$\delta S=\delta\int_{t_1}^{t_2}L(\boldsymbol{q}(t),\dot{\boldsymbol{q}}(t),t)\mathrm{d}t=0 \tag{3.6}$$

的函数 $\boldsymbol{q}(t)$.

我们来说明,如何利用条件(3.6)推导代表实际运动状态的 $q^{(\alpha)}(t)(\alpha=1,\cdots,s)$ 所满足的微分方程(即力学系统的运动方程).

为了使讨论形象化,作示意图如图 1. 图中将 s 个广义坐标 $q^{\alpha}(\alpha=1,\cdots,s)$ 用一根纵坐标轴表示,而横坐标所表示的是时间 t. 满足条件(3.4)的一组 $q^{(\alpha)}(t)$ 由端点在 $A(t_1,\boldsymbol{q}_1)$ 和 $B(t_2,\boldsymbol{q}_2)$ 的一条曲线表示. 它按(3.5)式决定作用量 S 的一个值. 当 $q^{(\alpha)}(t)$ 在端点 A,B 固定的条件下发生改变 $\delta q^{(\alpha)}(t)$ [见图(1b)]时,曲线发生变化,成为 $q^{(\alpha)}(t)+\delta q^{(\alpha)}(t)$,如图 1(a),使得作用量 S 发生改变. 根据(3.5)式,S 的改变是由于被积函数改变. 因此,

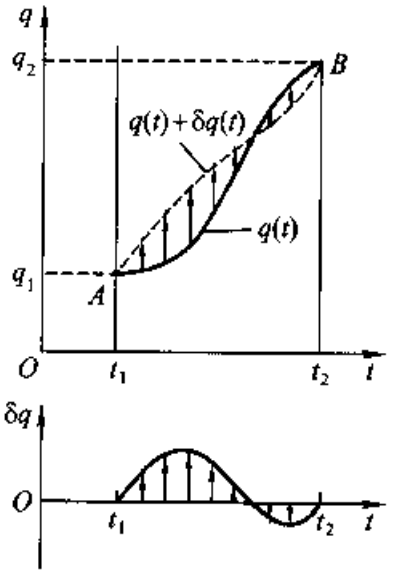


图 1　作用量变分的示意图

$$\delta S=\delta\int_{t_1}^{t_2}L(\boldsymbol{q}(t),\dot{\boldsymbol{q}}(t),t)\mathrm{d}t=\int_{t_1}^{t_2}\sum_{\alpha=1}^{s}\left(\frac{\partial L}{\partial q^{(\alpha)}}\delta q^{(\alpha)}+\frac{\partial L}{\partial\dot{q}^{(\alpha)}}\delta\dot{q}^{(\alpha)}\right)\mathrm{d}t. \tag{3.7}$$

对 $q^{(\alpha)}(t)$ 的变分 $\delta q^{(\alpha)}(t)$ 是在每一时刻 t 分别作出的[参看图(1b)],因此,它和对 t 求导的运算可以交换次序,即

$$\delta\dot{q}^{(\alpha)}=\delta\frac{\mathrm{d}}{\mathrm{d}t}q^{(\alpha)}=\frac{\mathrm{d}}{\mathrm{d}t}\delta q^{(\alpha)}. \tag{3.8}$$

代入(3.7)式右边第二项中,并对这一项进行分部积分,得到

$$\delta S=\sum_{\alpha=1}^{s}\frac{\partial L}{\partial\dot{q}^{(\alpha)}}\delta q^{(\alpha)}\Bigg|_{t_1}^{t_2}+\int_{t_1}^{t_2}\sum_{\alpha=1}^{s}\left(\frac{\partial L}{\partial q^{(\alpha)}}-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial\dot{q}^{(\alpha)}}\right)\delta q^{(\alpha)}\mathrm{d}t. \tag{3.9}$$

但是,我们已假定 $q^{(\alpha)}(t)$ 是在端点 A 和 B 固定的条件下改变,因而在时刻 t_1 和 t_2,$\delta q^{(\alpha)}$ 等于零,参看图 1(b),即

$$\delta q^{(\alpha)}(t_1)=\delta q^{(\alpha)}(t_2)=0. \tag{3.10}$$

因此,(3.9)式右方第一项为零. 于是有

$$\delta S = \int_{t_1}^{t_2} \sum_{\alpha=1}^{s} \left(\frac{\partial L}{\partial q^{(\alpha)}} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}^{(\alpha)}} \right) \delta q^{(\alpha)} \mathrm{d}t. \tag{3.11}$$

如果 $q^{(\alpha)}(t)$ 代表真实运动，则它应使 S 取极值，即对于任意的 $\delta q^{(\alpha)}$ 都有 $\delta S = 0$. 而这只在(3.11)式右方的被积函数中各个 $\delta q^{(\alpha)}$ 的系数都等于零时才可能，亦即应有

$$\frac{\partial L}{\partial q^{(\alpha)}} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}^{(\alpha)}} = 0, \qquad \alpha = 1,2,\cdots,s. \tag{3.12}$$

这称为拉格朗日方程，它就是实际运动的 $q^{(\alpha)}(t)$ 所应满足的微分方程.

比较(3.12)式和上一节的(2.21)式可见，两者完全一致. 这样，我们就用最小作用量原理作为基本原理推导出了经典力学的运动方程——拉格朗日方程.

最小作用量原理是适用于一切物理系统(其中包括一切力学系统)的基本原理. 它把不同系统的特点集中在一个标量函数——拉格朗日函数上.

需要注意的是，给定一个拉格朗日函数就完全决定了和它对应的力学系统的性质. 反过来，给定一个力学系统却不能完全决定对应的拉格朗日函数，在它上面还可以加上一个广义坐标和时间的任意函数对时间的全导数. 也就是说，$L = L(\boldsymbol{q},\dot{\boldsymbol{q}},t)$ 和

$$L' = L(\boldsymbol{q},\dot{\boldsymbol{q}},t) + \frac{\mathrm{d}}{\mathrm{d}t} f(\boldsymbol{q},t) \tag{3.13}$$

描述同一个力学系统. 其原因是，由 L 和 L' 得到的作用量 S 和 S' 的变分相等. 证明如下：

$$\delta S' = \delta \int_{t_1}^{t_2} \left(L(\boldsymbol{q},\dot{\boldsymbol{q}},t) + \frac{\mathrm{d}}{\mathrm{d}t} f(\boldsymbol{q},t) \right) \mathrm{d}t = \delta S + \delta f(\boldsymbol{q}(t_2),t_2) - \delta f(\boldsymbol{q}(t_1),t_1).$$

由于 $\boldsymbol{q}(t)$ 满足条件(3.10)，$\delta f(\boldsymbol{q}(t_i),t_i) = 0 (i = 1,2)$，因而 $\delta S' = \delta S$. **【证毕】**

这样，对于满足(3.13)式的两个拉格朗日函数 L 和 L'，由最小作用量原理得到的“真实运动”相同，因而它们描述同一个力学系统.

§1.1.4　伽利略相对性原理　自由质点的拉格朗日函数

(一)惯性参考系

为了研究物体的运动，必须确定它的位置，而为此又必须选定某物体作为参考. 特殊选定的，用来确定其他物体位置的物体称为参考系. 在不同的参考系中所观察到的物体运动会很不相同. 例如，放在行进中的火车车厢中的物体，当以地面为参考系时是运动的，而当以车厢为参考系时却是静止的. 因此，在研究机械运动的规律时，必须指明参考系. 在以上几节中，我们回避了这一问题，现在来作补充说明.

假定在所选用的参考系中，拉格朗日方程成立. 在这一参考系中作笛卡儿坐标，则质点的拉格朗日方程[§1.1.2(2.21)式]就是牛顿方程

$$m\ddot{\boldsymbol{r}} = m\dot{\boldsymbol{v}} = \boldsymbol{F}. \tag{4.1}$$

如果质点不受力作用，$\boldsymbol{F} = 0$，则

$$\dot{\boldsymbol{v}} = 0, \quad \boldsymbol{v} = \text{常矢量}. \tag{4.2}$$

即，不受外力作用的质点或者静止，或者作匀速直线运动. 凡是满足这一条件的参考系称为惯性系. 在以上几节中都隐含着采用惯性系作参考系的假定. 拉格朗日方程在惯性系中成立.

(二)伽利略变换

考虑一个任意的参考系 K. 质点在其中的位矢是 $\boldsymbol{r}$. 参考系 K′相对于 K 以速度 $\boldsymbol{V}$ 运动. 在 K′中同一质点的位置位矢是 $\boldsymbol{r}'$. 假定在 $t = 0$ 时刻，在 K 系和 K′系中所作的笛卡儿坐标原点重合，则在 t 时刻，$\boldsymbol{r}'$ 与 $\boldsymbol{r}$ 的关系是

$$\boldsymbol{r}' = \boldsymbol{r} - \boldsymbol{V}t, \tag{4.3}$$

如图 1. 另一方面，在 K′系和 K 系中，时间的计算相同，即有

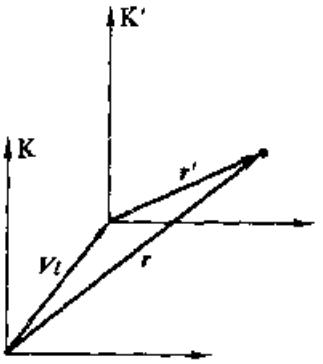


图 1　惯性参考系之间的变换

$$t' = t. \tag{4.4}$$

(4.3)式和(4.4)式是相互作匀速运动的两个参考系的坐标与时间的变换规律，称为伽利略变换.

根据(4.4)式，对(4.3)式两边取导数，得

$$\frac{\mathrm{d}\boldsymbol{r}'}{\mathrm{d}t'} = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} - \boldsymbol{V},$$

或

$$\boldsymbol{v}' = \boldsymbol{v} - \boldsymbol{V}. \tag{4.5}$$

这是相对作匀速运动的参考系之间的速度变换规律.

(三)伽利略相对性原理

由(4.5)式可见，如果 K 系是惯性系(不受力作用的自由质点在 K 系中或者静止，或者作匀速直线运动)，则相对 K 系作匀速运动的 K′系也是惯性系(自由质点在 K′系中也是或者静止，或者作匀速直线运动).

进一步对(4.5)式求导，得到

$$\frac{\mathrm{d}^2\boldsymbol{r}'}{\mathrm{d}t'^2} = \frac{\mathrm{d}^2\boldsymbol{r}}{\mathrm{d}t^2}.$$

在运动方程(4.1)式中，只有 $\ddot{\boldsymbol{r}}$ 而没有 $\dot{\boldsymbol{r}}$，因此，在不同惯性系中，运动方程的形式相同. 我们知道，正是运动方程反映了机械运动的规律. 所以，以上的讨论表明，在

不同惯性系中，机械运动的规律有相同的形式. 不可能通过机械运动规律的差别来区分不同的惯性系. 所有的惯性系对研究机械运动规律都是等效的. 这称为伽利略相对性原理.

(四)自由质点的拉格朗日函数

根据最小作用量原理，力学系统的全部性质集中在它的拉格朗日函数上. 因此，如何确定力学系统的拉格朗日函数是一个重要问题❶. 一般说来，一个力学系统的拉格朗日函数是根据它的一些对称性质(即在时空变换和内部变换下的不变性质)和一些普遍原理(如伽利略相对性原理)确定的. 下面我们以单个自由质点为例，说明如何根据这些根本性的原理和性质决定系统的拉格朗日函数.

对于单个自由质点，广义坐标可选为笛卡儿坐标 $\boldsymbol{r}$，$L = L(\boldsymbol{r},\dot{\boldsymbol{r}},t)$. 它在时间平移，空间平移和空间转动下应有不变性(时空均匀和各向同性).

时间平移变换是：

$$t \to t' = t + \tau. \tag{4.6}$$

要在这一变换下不变，拉格朗日函数 L 不能显含 t，$L = L(\boldsymbol{r},\dot{\boldsymbol{r}})$.

空间平移变换是：

$$\boldsymbol{r} \to \boldsymbol{r}' = \boldsymbol{r} + \boldsymbol{a}. \tag{4.7}$$

要在这一变换下不变，拉格朗日函数 L 不能显含坐标 $\boldsymbol{r}$，而只能是速度 $\dot{\boldsymbol{r}}$ 的函数，$L = L(\dot{\boldsymbol{r}})$.

空间转动改变矢量的指向. 要在这一变换下不变，拉格朗日函数 L 不能依赖于速度 $\dot{\boldsymbol{r}}$ 的方向，而只能依赖于它的大小，$L = L(v^2)$.

现在来考虑相对于 K 系以无穷小速度 $\boldsymbol{\varepsilon}$ 运动的另一惯性系 K′. 根据伽利略相对性原理，力学规律在 K 和 K′系中应该相同. 因此，在这两个参考系中的拉格朗日函数最多只能相差一个坐标的任意函数的时间全导数，即

$$L' = L(v'^2) = L(v^2) + \mathrm{d}f(\boldsymbol{r},t)/\mathrm{d}t. \tag{4.8}$$

按照伽利略变换，$\boldsymbol{v}' = \boldsymbol{v} + \boldsymbol{\varepsilon}$. 将 $L(v'^2) = L((\boldsymbol{v}+\boldsymbol{\varepsilon})^2)$ 对小量 $\boldsymbol{\varepsilon}$ 展开，保留到 $\boldsymbol{\varepsilon}$ 的一次项：

$$L(v'^2) = L(v^2) + \frac{\mathrm{d}L}{\mathrm{d}v^2}2\boldsymbol{v}\cdot\boldsymbol{\varepsilon}. \tag{4.9}$$

比较(4.8)、(4.9)两式可见

$$\frac{\mathrm{d}L}{\mathrm{d}v^2}2\boldsymbol{\varepsilon}\cdot\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \frac{\partial f}{\partial \boldsymbol{r}}\cdot\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} + \frac{\partial f}{\partial t}$$

❶ 在 §1.1.2 中得到拉格朗日的表达式 $L = T - U$，见该节(2.22)式，是从牛顿方程出发得到的. 在 §1.1.3 中由最小作用量原理推导拉格朗日方程时，没有得到拉格朗日的表达式. 现在要讨论的是如何从普遍原理出发，不依赖于牛顿方程，得到拉格朗日的表达式.

由于 $\boldsymbol{v} = \mathrm{d}\boldsymbol{r}/\mathrm{d}t$ 是任意矢量，上式左右两边 $\mathrm{d}\boldsymbol{r}/\mathrm{d}t$ 的同次项的系数应相等，即：右边第二项为零，而且左右两边 $\mathrm{d}\boldsymbol{r}/\mathrm{d}t$ 的系数相等，

$$\frac{\mathrm{d}L}{\mathrm{d}v^2}2\boldsymbol{\varepsilon} = \frac{\partial f}{\partial \boldsymbol{r}}.$$

上式左边不含 $\boldsymbol{r}$，右边不含 $\boldsymbol{v}$，因而只能是常数. 因此，$\mathrm{d}L/\mathrm{d}v^2$ 是一个不依赖于速度 $\boldsymbol{v}$ 的常数. 将这一常数写为 $m/2$，$\mathrm{d}L/\mathrm{d}v^2 = m/2$，积分得

$$L = mv^2/2. \tag{4.10}$$

以上是根据“有无穷小相对速度的参考系之间的变换满足伽利略相对性原理”的要求，得到了自由质点的拉格朗日函数(4.10)式. 可以证明，这一拉格朗日函数也满足有限相对速度变换下伽利略相对性原理的要求，见习题 11.

习　题

1. 设质点在势能场 $U(r)$ 中运动. 在笛卡儿坐标系中写出其拉格朗日方程.

2. 已知柱坐标 (ρ,φ,z) 与笛卡儿坐标的关系是

$x = \rho\cos\varphi$，　$y = \rho\sin\varphi$，　$z = z$，

如图 1. 设质点在轴对称势能场 $U(\rho)$ 中运动，写出其拉格朗日方程.

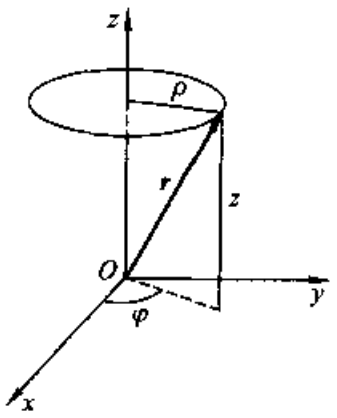


图 1　柱坐标

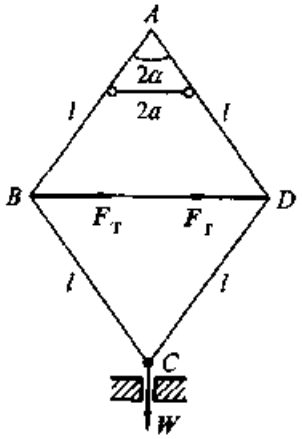


图 2　四根轻棒联结图

3. 长为 l 的绳系一小球，悬点按 $X = A\sin\omega(t - t_0)$ 方式运动，如 §1.1.2 图 4. 小球被限制在 (x,z) 平面内. $t = t_0$ 时，悬线铅直向下.

(a)求悬线和铅直线偏 $\delta\theta$ 所对应的虚位移 $\delta\boldsymbol{r}$.

(b)已知在这一时刻的角速度为 $\dot{\theta}$，求经过 $\mathrm{d}t$ 时间后的位移 $\mathrm{d}\boldsymbol{r}$. 问：当 $\mathrm{d}t\to 0$ 时，$\mathrm{d}\boldsymbol{r}$ 与 $\delta\boldsymbol{r}$ 有何差别？(答案用图表出)

4. 长度同为 l 的轻棒四根，相互联接成为一个可以无摩擦地改变顶角的菱形 $ABCD$. AB 和 AD 两棒无摩擦地支于同一水平线上相距 $2a$ 的两根钉上，BD 之间用一根轻棒联接. 在联接点(B 和 D)处，各棒之间可以无摩擦地转动. C 点上系一重物 W. C 点和重物受约束，只能上下运动. 设 A 点上的顶角为 2α，试用虚功原理求平衡时联接棒 BD 中的张力 F_T(参看图 2). 讨论 F_T 的方向与 α 的大小的关系. 问：什么情况下有 $F_T = 0$？说明其意义.

5. 气缸中气体的压强为 P，大气压强为 p，活塞半径为 R，曲柄半径为 r，在曲柄和连杆的联结点 A 上悬挂一个重物 Q 达到平衡(忽略摩擦). 试用虚功原理求重力 Q 和曲柄长度 l 以及曲柄相对于活塞轴的角度 ψ 之间的关系(参看图 3).

6. 从 §1.1.2(2.16)、(2.17)、(2.18)、(2.20)式出发，推导拉格朗日方程 §1.1.2(2.21)式.

7. 离心节速器的示意图如图 4. 质点 m_2 可以沿竖直方向运动，整个系统以匀角速度 ω 绕竖直轴转动，求拉格朗日函数.

8. 在质量为 m_1，棒长为 l_1 的单摆下方再悬一个质量为 m_2，棒长为 l_2 的摆. m_1 和 m_2 都在 xy 平面内运动（如图 5），忽略棒的质量，求此系统的拉格朗日函数.

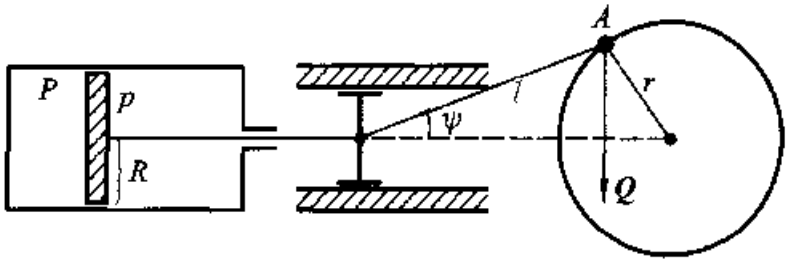


图 3　曲柄连杆机构

9. 质量为 M 的斜面可以无摩擦地在水平桌面上滑动. 斜面上无摩擦地放一滑块 m，如图 6. 写出拉格朗日方程，并求斜面的加速度 $\ddot{X}$ 和滑块相对于斜面的加速度 $\ddot{x}$.

10. 直接用拉格朗日方程［§1.1.2 (2.21)式］证明，由相差一广义坐标和时间的函数的时间全导数的两个拉格朗日函数 L' 和 L［§1.1.3(3.13)式］得到的运动方程相同.

11. 证明一维运动自由质点的拉格朗日函数［§1.1.4(4.10)式］满足有限相对速度变换下伽利略相对性原理的要求.

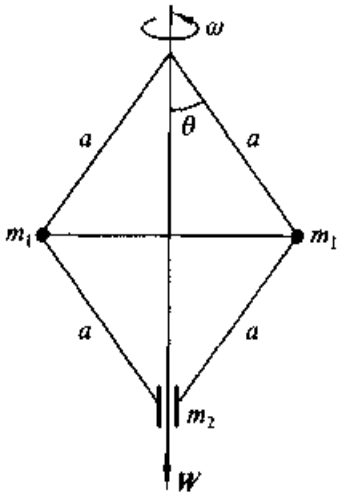


图 4　离心节速器

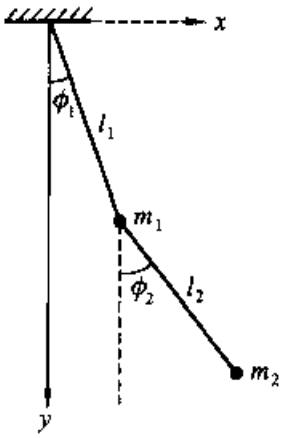


图 5　平面双摆

12. 已知一维运动自由质点的拉格朗日是 $L = m\dot{x}^2/2$.

(a) 证明：当按真实运动方式运动时，作用量是

$$S_0 = \frac{m(x_2 - x_1)^2}{2(t_2 - t_1)}.$$

(b) 设 $x(t_1) = a, x(t_2) = b$，求 S_0；并任意假定一种非真实的运动方式，计算相应的作用量 S_1，验证 $S_1 > S_0$.

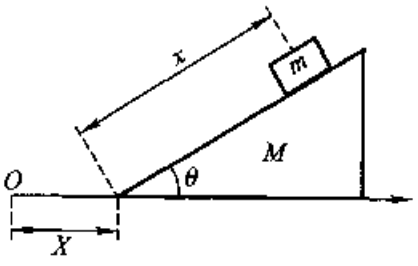


图 6　滑动斜面

第二章　守恒律

具有 s 个自由度的系统的运动方程是 s 个广义坐标 $q_\alpha(t)(\alpha = 1,2,\cdots,s)$ 的二阶常微分方程组. 在以下几章里，我们将就一些具体情况求这一组微分方程的解. 这一章暂不求解运动方程，而先来讨论另一个重要问题——守恒律.

广义坐标 $q_\alpha(t)$ 和它的导数（广义速度）$\dot{q}_\alpha(t)$ 都是随时间变化的，但是可以由它们组成一些不随时间变化的量，称为守恒量. 利用这些守恒量，可以帮助我们求运动方程的解，或者分析解的性质. 这一章讨论一些重要的守恒量.

§1.2.1　动量和能量

（一）循环坐标与广义动量

将拉格朗日方程 §1.1.2(2.21)式改写为

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_\alpha} = \frac{\partial L}{\partial q_\alpha}, \qquad \alpha = 1,\cdots,s. \tag{1.1}$$

这和笛卡儿坐标中的牛顿方程

$$\frac{\mathrm{d}}{\mathrm{d}t}p_i = f_i, \qquad i = 1,2,3 \tag{1.2}$$

相似. 因此，称

$$p_\alpha = \partial L/\partial \dot{q}_\alpha, \qquad \alpha = 1,\cdots,s \tag{1.3}$$

为广义动量，

$$f_\alpha = \partial L/\partial q_\alpha, \qquad \alpha = 1,\cdots,s \tag{1.4}$$

为广义力.

我们来举几个简单例子. 第一个例子是笛卡儿坐标中的单个粒子，此时有

$$L = T - U = \sum_{i=1}^{3}\frac{1}{2}m\dot{x}_i^2 - U(x_1,x_2,x_3). \tag{1.5}$$

在此情况下，广义动量

$$p_i = \partial L/\partial \dot{x}_i = m\dot{x}_i, \qquad i = 1,2,3 \tag{1.6}$$

就是通常的动量，而广义力

$$f_i = \partial L/\partial x_i = -\partial U/\partial x_i, \qquad i = 1,2,3 \tag{1.7}$$

就是通常的力. 第二个例子是球坐标 (r,θ,φ) 中的单个粒子. 拉格朗日函数是 §1.1.1(1.35)式：

$$L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\varphi}^2) - U(r,\theta,\varphi). \tag{1.8}$$

对应于三个广义坐标的广义动量和广义力分别是：

$$\left.\begin{array}{lll} r & p_r = \dfrac{\partial L}{\partial \dot r} = m\dot r & f_r = \dfrac{\partial L}{\partial r} = mr\dot\theta^2 + mr\sin^2\theta\dot\varphi^2 - \dfrac{\partial U}{\partial r}, \\ \theta & p_\theta = \dfrac{\partial L}{\partial \dot\theta} = mr^2\dot\theta & f_\theta = \dfrac{\partial L}{\partial \theta} = mr^2\sin\theta\cos\theta\dot\varphi^2 - \dfrac{\partial U}{\partial \theta}, \\ \varphi & p_\varphi = \dfrac{\partial L}{\partial \dot\varphi} = mr^2\sin^2\theta\dot\varphi & f_\varphi = \dfrac{\partial L}{\partial \varphi} = -\dfrac{\partial U}{\partial \varphi}. \end{array}\right\} \tag{1.9}$$

特别值得注意的是，和广义坐标 φ 对应的广义力简单地等于 $-\partial U/\partial\varphi$，不包含其他附加项. 因此，只要势能 U 不含 φ（例如绕 Oz 轴旋转对称的力场），就有 $f_\varphi = 0$. 而由(1.1)式就有

$$\mathrm{d}p_\varphi/\mathrm{d}t = 0, \tag{1.10}$$

即：绕 Oz 轴旋转对称的力场中，和广义坐标 φ 对应的广义动量[即绕 Oz 轴的角动量，见 §1.2.2(2.16)式]守恒.

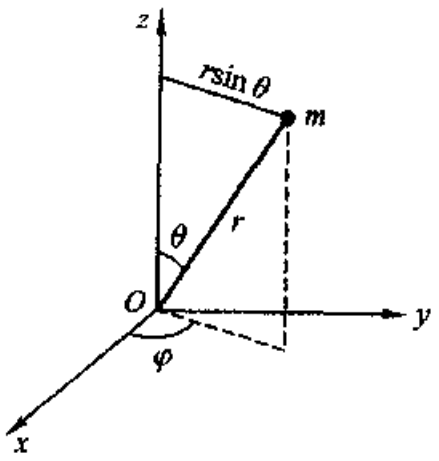


图1　有心力场中的质点

在上面第二个例子中，我们得到了广义动量 p_φ 守恒，其原因是在拉格朗日(1.8)式中不包含对应的广义坐标 φ. 这是一个有普遍意义的结论. 一般说来，如果在拉格朗日函数中不包含某一个广义坐标 q_α，则称这一广义坐标为循环坐标. 对于循环坐标 q_α 有 $\partial L/\partial q_\alpha = 0$，因而由拉格朗日方程得到

$$\frac{\mathrm{d}}{\mathrm{d}t}p_\alpha = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot q_\alpha} = \frac{\partial L}{\partial q_\alpha} = 0.$$

即：和循环坐标对应的广义动量守恒.

(二)能量

拉格朗日 L 是 q_α 和 $\dot q_\alpha$ 的函数. 它通过 $q_\alpha(t)$ 和 $\dot q_\alpha(t)$ 而依赖于时间. 但有时，L 还明显地含 t. 例如，假定系统处于随时间变化的外力场中，

$$L = T - U(q,t), \tag{1.11}$$

则 L 就明显地与 t 有关. 这种系统与产生外力场的“源”有能量交换，因而不是保守系.

现在我们考虑保守系，L 不明显含 t，而只是通过 q_α 和 $\dot q_\alpha$ 依赖于 t. 此时，L 对时间的导数是

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \sum_{\alpha=1}^{s}\left(\frac{\partial L}{\partial q_\alpha}\dot q_\alpha + \frac{\partial L}{\partial \dot q_\alpha}\ddot q_\alpha\right). \tag{1.12}$$

将拉格朗日方程 §1.1.2(2.21)式代入上式，得到

$$\frac{\mathrm{d}L}{\mathrm{d}t}=\sum_{\alpha=1}^{s}\left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial\dot{q}_\alpha}\right)\dot{q}_\alpha+\frac{\partial L}{\partial\dot{q}_\alpha}\frac{\mathrm{d}}{\mathrm{d}t}\dot{q}_\alpha\right]=\sum_{\alpha=1}^{s}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial\dot{q}_\alpha}\dot{q}_\alpha\right).\tag{1.13}$$

将上式的左边移到右边,得到

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{\alpha=1}^{s}\left(\frac{\partial L}{\partial\dot{q}_\alpha}\dot{q}_\alpha\right)-L\right]=0.\tag{1.14}$$

我们看到,对于保守系,上式方括号内的表达式是一个守恒量.定义它为机械能量,简称能量,用符号 E 表示

$$E=\sum_{\alpha=1}^{s}\left(\frac{\partial L}{\partial\dot{q}_\alpha}\dot{q}_\alpha\right)-L.\tag{1.15}$$

(1.14)、(1.15)式表明,保守系的能量守恒.

(三)相互作用质点系的拉格朗日函数

由(1.15)式给出的能量表达式中含有拉格朗日函数 L.在§1.1.4中讨论了不依赖于牛顿方程而从普遍原理确定拉格朗日函数的问题,得到了单个自由质点的拉格朗日函数[见§1.1.4(4.10)式].现在来讨论如何从普遍原理确定相互作用质点系的拉格朗日函数.为了简单起见,假定不存在约束.

先来看质点之间无相互作用的情况.此时,第 a 个质点的拉格朗日函数仍由§1.1.4(4.10)式

$$L_a=mv_a^2/2\tag{1.16}$$

给出.总系统的拉格朗日是它们的和

$$L=\sum_{a=1}^{N}L_a\qquad(\text{质点间无相互作用}).\tag{1.17}$$

由这一拉格朗日函数对于坐标 $\boldsymbol{r}_a$ 得到的拉格朗日方程就是无相互作用的单个质点 a 的拉格朗日方程.

如果质点之间有相互作用,总拉格朗日函数中必然有同时包含多个质点坐标的项❶.将这一项写为 $G(\boldsymbol{r}_1,\boldsymbol{r}_2,\cdots,\boldsymbol{r}_N,t)$,则相互作用质点系的拉格朗日函数就是

$$L=\sum_{a=1}^{N}\frac{1}{2}m_a\dot{\boldsymbol{r}}_a^2+G(\boldsymbol{r}_1,\boldsymbol{r}_2,\cdots,\boldsymbol{r}_N,t).\tag{1.18}$$

如果相互作用与时间无关,G 不显含时间 t,则 L 不显含时间 t,由前面的推导知能量守恒,且能量的表达式(1.15)成为

$$E=\sum_{a=1}^{N}m_a\dot{\boldsymbol{r}}_a^2-\left(\sum_{a=1}^{N}\frac{1}{2}m_a\dot{\boldsymbol{r}}_a^2+G(\boldsymbol{r}_1,\boldsymbol{r}_2,\cdots,\boldsymbol{r}_N,t)\right)$$

---

❶ 通常情况下,相互作用拉格朗日只包含质点的坐标,而不包含它们的速度.只在某些特殊情况下,例如存在磁力的时候,相互作用拉格朗日才包含速度.我们现在不考虑这种情况.

$$= \sum_{\alpha=1}^{N} \frac{1}{2} m_\alpha \dot{\boldsymbol{r}}_\alpha^2 - G(\boldsymbol{r}_1, \boldsymbol{r}_2, \cdots, \boldsymbol{r}_N, t).$$

右边第一个和式是 N 个质点之间无相互作用时的能量，称为质点系的总动能. 第二项是相互作用引起的附加能量，称为质点系的势能. 通常用符号 U 表示势能，则有 $U = -G$. 代入(1.18)式得到，相互作用质点系的拉格朗日函数是 $L = T - U$. [比较 §1.1.2(2.22)式. 注意，那一节的(2.22)式是从牛顿方程出发得到的，而这里的推导不依赖于牛顿方程.]

推广到有约束的情况，坐标 $\boldsymbol{r}_\alpha (\alpha = 1, \cdots, N)$ 要用广义坐标 $q_\alpha (\alpha = 1, \cdots, s)$ 代替. 拉格朗日函数是

$$L(\boldsymbol{q}, \dot{\boldsymbol{q}}) = T(\boldsymbol{q}, \dot{\boldsymbol{q}}) - U(\boldsymbol{q}), \qquad \boldsymbol{q} \equiv q_1, \cdots, q_s. \tag{1.19}$$

(四)机械能

注意，在笛卡儿坐标中，动能只是速度 $\dot{x}_i$ 的函数，不是坐标 x_i 的函数；而在广义坐标中，动能的表达式里既包含广义速度 $\dot{q}_\alpha$ 又包含广义坐标 q_α，例如见(1.8)式. 另一方面，从(1.8)式这个例子中也可以看到，和笛卡儿坐标中一样，动能是广义速度 $\dot{q}_\alpha$ 的二次齐次式. 这是一个有普遍意义的结论，它对于一切不显含时间的广义坐标变换

$$x_i^{(\alpha)} = x_i^{(\alpha)}(q_1, q_2, \cdots, q_s), \qquad \alpha = 1, 2, \cdots, N; \quad i = 1, 2, 3 \tag{1.20}$$

都成立. 证明如下：

将(1.20)式代入笛卡儿坐标中动能的表达式：

$$T = \frac{1}{2} \sum_{\alpha=1}^{N} \sum_{i=1}^{3} m_\alpha (\dot{x}_i^{(\alpha)})^2 = \frac{1}{2} \sum_{\alpha=1}^{N} \sum_{i=1}^{3} m_\alpha \left(\sum_{\alpha=1}^{s} \frac{\partial x_i^{(\alpha)}}{\partial q_\alpha} \dot{q}_\alpha \cdot \sum_{\beta=1}^{s} \frac{\partial x_i^{(\alpha)}}{\partial q_\beta} \dot{q}_\beta \right)$$

交换对 i 和对 α, β 求和的次序，定义

$$\sum_{\alpha=1}^{N} \sum_{i=1}^{3} m_\alpha \left(\frac{\partial x_i^{(\alpha)}}{\partial q_\alpha} \cdot \frac{\partial x_i^{(\alpha)}}{\partial q_\beta} \right) = m_{\alpha\beta}(\boldsymbol{q}), \tag{1.21}$$

得到

$$T = \frac{1}{2} \sum_{\alpha=1}^{s} \sum_{\beta=1}^{s} m_{\alpha\beta}(\boldsymbol{q}) \dot{q}_\alpha \dot{q}_\beta. \tag{1.22}$$

这表明，用广义坐标表出的动能是广义速度的二次齐次式，其中的系数矩阵一般情况下是广义坐标的函数. 【证毕】

根据齐次函数的欧拉定理，如果 $f(u_1, u_2, \cdots, u_s)$ 是 s 个变量 u_α 的 n 次齐次式，则

$$\sum_{\alpha=1}^{s} \frac{\partial f}{\partial u_\alpha} u_\alpha = nf. \tag{1.23}$$

现在，动能 T 是 $\dot{q}_\alpha$ 的二次齐次函数，因而，由(1.19)式得

$$\sum_{\alpha=1}^{s}\frac{\partial L}{\partial \dot{q}_\alpha}\dot{q}_\alpha = \sum_{\alpha=1}^{s}\frac{\partial T}{\partial \dot{q}_\alpha}\dot{q}_\alpha = 2T.$$

和(1.19)式一道代入(1.15)式,得到

$$E = T + U. \tag{1.24}$$

即,机械能等于动能与势能之和.在运动过程中,动能和势能都会变化,但它们之和,即总机械能保持不变.

§1.2.2　守恒律与对称性的关系　角动量

能量(指机械能,下同)守恒和动量守恒是机械运动的两条基本守恒律.它们的产生有深刻的物理原因.下面我们来证明,产生能量守恒和动量守恒的根源在于时间和空间的均匀性.

从现在开始到本章结束,我们限于考虑笛卡儿坐标系.

(一)时间的均匀性与能量守恒

假定系统处于变化的外场中,如§1.1.2(2.29)式,则时间不是均匀的.这意思是说,用不同的时刻作为计算时间的起点,所得到的运动方程不相同;以不同的时刻作为时间轴的原点对于研究系统的运动不等效.

但是,如果系统处于不变的外场中,或者更特殊些说,如果系统不处于外场中,则时间是均匀的,以不同时刻作为计算时间的起点,对于研究运动是等效的.在此情况下,拉格朗日不显含时间

$$\partial L/\partial t = 0, \tag{2.1}$$

而根据§1.2.1的分析知道,这导致能量守恒.由此可见,能量守恒是由时间的均匀性产生的.

(二)空间的均匀性与动量守恒

在没有外场的情况下,不但时间是均匀的,空间也是均匀的.这就是说,在没有外场的情况下,无论选用空间哪一点作为坐标原点来研究系统的运动都是等效的.我们来看看,空间的这一性质(均匀性)产生什么结果.

假定在某一坐标系中,系统的拉格朗日为 L.将坐标系平移一个常矢量 $\boldsymbol{\varepsilon}$,使拉格朗日改变 δL.如果空间是均匀的,则平移坐标系时拉格朗日应该不变,即

$$\delta L = 0 \quad (\text{坐标系平移 } \boldsymbol{\varepsilon}). \tag{2.2}$$

但是,坐标系平移 $\boldsymbol{\varepsilon}$ 使得每个粒子的坐标都改变 $\boldsymbol{\varepsilon}$

$$\delta \boldsymbol{r}_a = \boldsymbol{\varepsilon}, \quad a = 1,2,\cdots,N, \tag{2.3}$$

而速度 $\dot{\boldsymbol{r}}_a$ 不变(因为 $\boldsymbol{\varepsilon}$ 是与时间无关的常矢量):

$$\delta \dot{\boldsymbol{r}}_a = 0, \quad a = 1,2,\cdots,N. \tag{2.4}$$

因此，拉格朗日的变化是

$$\delta L = \sum_{a=1}^{N}\left(\frac{\partial L}{\partial \boldsymbol{r}_a}\cdot\delta\boldsymbol{r}_a + \frac{\partial L}{\partial \dot{\boldsymbol{r}}_a}\cdot\delta\dot{\boldsymbol{r}}_a\right) = \boldsymbol{\varepsilon}\cdot\sum_{a=1}^{N}\frac{\partial L}{\partial \boldsymbol{r}_a}. \tag{2.5}$$

如果空间均匀，则按(2.2)式，对任意矢量 $\boldsymbol{\varepsilon}$ 都有 $\delta L = 0$，因而由(2.5)式有❶

$$\sum_{a=1}^{N}\frac{\partial L}{\partial \boldsymbol{r}_a} = 0. \tag{2.6}$$

将拉格朗日方程代入，得到

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{a=1}^{N}\frac{\partial L}{\partial \dot{\boldsymbol{r}}_a}\right) = 0.$$

利用 §1.2.1(1.6)式可将上式写为

$$\mathrm{d}\left(\sum_{a=1}^{N}\boldsymbol{p}_a\right)\Big/\mathrm{d}t = \mathrm{d}\boldsymbol{P}/\mathrm{d}t = 0, \tag{2.7}$$

其中，$\boldsymbol{P} = \sum_a \boldsymbol{p}_a$ 是系统的总动量. 即：当空间均匀时，系统的总动量守恒. 由此可见，动量守恒是由空间的均匀性产生的.

如果系统只是沿某一方向(例如 x 方向)均匀，则只在 $\boldsymbol{\varepsilon} = \varepsilon\boldsymbol{e}_x$ 时才有 $\delta L = 0$. 此时，由 ε 的任意性得到 $P_x = \boldsymbol{e}_x\cdot\boldsymbol{P}$ 守恒. 即：当空间沿某一方向均匀时，这一方向上的动量守恒.

(三)空间的各向同性与角动量守恒

时间和空间的均匀性具体表示出来就是拉格朗日 L 在时间轴平移和空间坐标平移下的不变性. 由于拉格朗日 L 反映了系统的全部力学性质，所以，拉格朗日的不变性反映了所研究的系统的不变性. 系统对某种变换的不变性称为系统对这种变换的对称性. 因此可以说：系统具有时间平移对称性导致能量守恒，系统具有空间平移对称性导致动量守恒.

在有些情况下，例如，在无外场情况下或者在有心力场中，空间的各个方向是等效的，或者说：空间是各向同性的. 在此情况下，将系统转动，拉格朗日不变；系统具有转动对称性. 我们来看，这一对称性导致什么量守恒.

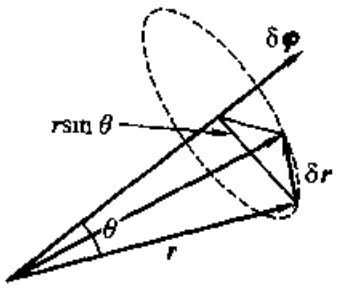


图 1　转动引起的位矢变化

令系统绕空间某一轴转动一个角度 $\delta\varphi$. 沿这一转动轴按右手螺旋方向作一个长为 $\delta\varphi$ 的矢量 $\delta\boldsymbol{\varphi}$，它代表了这一转动，如图 1. 用 θ 表示 $\boldsymbol{r}$ 和 $\delta\boldsymbol{\varphi}$ 之间的夹

❶ 当两个矢量的点积等于零时，由其中之一不等于零不能得到另一矢量必等于零的结论. 只有其中一个矢量是任意矢量，才能推得另一矢量为零.

角，则位矢 $\boldsymbol{r}$ 的端点绕 $\delta\boldsymbol{\varphi}$ 轴转动的半径是 $r\sin\theta$，因此 $\delta\boldsymbol{r}$ 的长度是

$$\delta r = \delta\varphi(r\sin\theta).$$

另一方面，当 $\delta\boldsymbol{\varphi}$ 小时，$\delta\boldsymbol{r}$ 垂直于 $\delta\boldsymbol{\varphi}$ 和 $\boldsymbol{r}$，因此 $\delta\boldsymbol{r}$ 等于 $\delta\boldsymbol{\varphi}$ 和 $\boldsymbol{r}$ 的叉乘积：

$$\delta\boldsymbol{r} = \delta\boldsymbol{\varphi}\times\boldsymbol{r}. \tag{2.8}$$

注意，此式成立的条件是位矢 $\boldsymbol{r}$ 的起点（坐标原点）在转动 $\delta\boldsymbol{\varphi}$ 的转轴上.

将(2.8)式对 t 求导，得到

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta\boldsymbol{r} = \delta\boldsymbol{\varphi}\times\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}.$$

但是，δ 和 $\mathrm{d}/\mathrm{d}t$ 是相互独立的，可以交换次序，因而上式可以写成

$$\delta\dot{\boldsymbol{r}} = \delta\boldsymbol{\varphi}\times\dot{\boldsymbol{r}}. \tag{2.9}$$

由此产生的拉格朗日的变化是

$$\delta L = \sum_{a=1}^{N}\left(\frac{\partial L}{\partial\boldsymbol{r}_a}\cdot\delta\boldsymbol{r}_a + \frac{\partial L}{\partial\dot{\boldsymbol{r}}_a}\cdot\delta\dot{\boldsymbol{r}}_a\right) = \sum_{a=1}^{N}\left(\frac{\partial L}{\partial\boldsymbol{r}_a}\cdot[\delta\boldsymbol{\varphi}\times\boldsymbol{r}_a] + \frac{\partial L}{\partial\dot{\boldsymbol{r}}_a}\cdot[\delta\boldsymbol{\varphi}\times\dot{\boldsymbol{r}}_a]\right).$$

利用矢量 $\boldsymbol{a}\cdot[\boldsymbol{b}\times\boldsymbol{c}] = \boldsymbol{b}\cdot[\boldsymbol{c}\times\boldsymbol{a}]$，可将上式写为

$$\delta L = \delta\boldsymbol{\varphi}\cdot\sum_{a=1}^{N}\left(\boldsymbol{r}_a\times\frac{\partial L}{\partial\boldsymbol{r}_a} + \dot{\boldsymbol{r}}_a\times\frac{\partial L}{\partial\dot{\boldsymbol{r}}_a}\right).$$

将拉格朗日方程代入，得到

$$\begin{aligned}\delta L &= \delta\boldsymbol{\varphi}\cdot\sum_{a=1}^{N}\left[\boldsymbol{r}_a\times\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial\dot{\boldsymbol{r}}_a}\right) + \frac{\mathrm{d}\boldsymbol{r}_a}{\mathrm{d}t}\times\frac{\partial L}{\partial\dot{\boldsymbol{r}}_a}\right]\\ &= \delta\boldsymbol{\varphi}\cdot\frac{\mathrm{d}}{\mathrm{d}t}\sum_{a=1}^{N}\left[\boldsymbol{r}_a\times\frac{\partial L}{\partial\dot{\boldsymbol{r}}_a}\right]\\ &= \delta\boldsymbol{\varphi}\cdot\frac{\mathrm{d}}{\mathrm{d}t}\sum_{a=1}^{N}\boldsymbol{r}_a\times\boldsymbol{p}_a.\end{aligned} \tag{2.10}$$

这里用了 §1.2.1(1.6)式.

假定系统具有转动对称性（无外力场或者是中心对称场），则在系统转动下拉格朗日应该不变，对任意 $\delta\boldsymbol{\varphi}$ 都有 $\delta L = 0$. 由(2.10)式知，对于这种系统

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{a=1}^{N}\boldsymbol{r}_a\times\boldsymbol{p}_a = 0. \tag{2.11}$$

这表明，在此情况下，矢量

$$\boldsymbol{L} = \sum_{a=1}^{N}\boldsymbol{r}_a\times\boldsymbol{p}_a \tag{2.12}$$

守恒. $\boldsymbol{L}$ 称为系统的角动量，我们看到，系统的转动对称性导致角动量守恒.

在有些情况下，系统并不具有完全的转动不变性，但是有绕某个轴，例如绕 z 轴的转动不变性. 此时，只在 $\delta\boldsymbol{\varphi} = \delta\varphi\boldsymbol{e}_z$ 时，才有 $\delta L = 0$. 由 $\delta\boldsymbol{\varphi}$ 的任意性得到角动

量的 z 分量，

$$L_z = \boldsymbol{e}_z \cdot \boldsymbol{L} = \sum_{a=1}^{N} [\boldsymbol{r}_a \times \boldsymbol{p}_a]_z = \sum_{a=1}^{N} (x_a p_{ya} - y_a p_{xa}) \tag{2.13}$$

守恒.

角动量的 z 分量可以用另一种方式表述. 根据 §1.1.1(1.9)式

$$x_a = r_a \sin\theta_a \cos\varphi_a, \qquad y_a = r_a \sin\theta_a \sin\varphi_a, \tag{2.14}$$

$$p_{xa} = m_a \dot{x}_a, \qquad p_{ya} = m_a \dot{y}_a. \tag{2.15}$$

将(2.14)式代入(2.15)式，再一道代入(2.13)式，注意绕 z 轴转动时 φ 改变而 θ 不变，得到

$$L_z = \sum_{a=1}^{N} m_a r_a^2 \sin^2\theta_a \dot{\varphi}_a = \sum_{a=1}^{N} p_{\varphi a}. \tag{2.16}$$

这里，$p_{\varphi a}$ 是质点 a 对应于广义坐标 φ 的广义动量，见 §1.2.1(1.9)式. 注意到 φ 是绕 z 轴转动的广义坐标，可以看出(2.16)的物理意义是：绕 z 轴转动的角动量就是和绕 z 轴转动的广义坐标 φ 对应的广义动量.

在结束本节之前，我们来证明一个比(2.8)式更普遍的公式.

令 $\boldsymbol{A}$ 为一个任意矢量，它和系统的转动轴的夹角为 θ，如图 2. 由图可见，当系统转动 $\delta\boldsymbol{\varphi}$ 时，矢量 $\boldsymbol{A}$ 的改变是

$$\delta\boldsymbol{A} = \delta\boldsymbol{\varphi} \times \boldsymbol{A}. \tag{2.17}$$

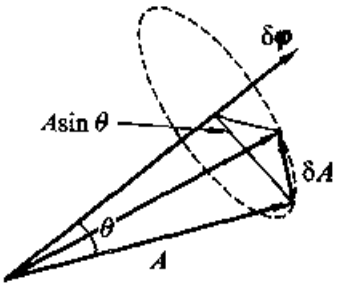


图 2　转动引起的矢量 $\boldsymbol{A}$ 变化

这是(2.8)式的推广. 它表明：当系统转动 $\delta\boldsymbol{\varphi}$ 时，矢量 $\boldsymbol{A}$ 的改变是 $\delta\boldsymbol{\varphi}$ 叉乘这一矢量. 此式成立的条件是矢量 $\boldsymbol{A}$ 的起点在转动 $\delta\boldsymbol{\varphi}$ 的转轴上. (2.8)式和(2.9)式都是这一普遍公式的特例.

§1.2.3　质点组的动量、能量与角动量　质心

本节讨论质点组的运动. 一个重要问题是设法从中分离出整个质点组作为一个整体的运动和质点在质点组内部的运动.

(一)质点组的动量定理

假定系统由 N 个质点组成，第 a 个质点和第 b 个质点 $(a \neq b)$ 之间的相互作用可以用相互作用势能 U_{ab} 表示，它是这两个质点之间的距离

$$r_{ab} = |\boldsymbol{r}_a - \boldsymbol{r}_b| = |\boldsymbol{r}_{ab}| = (\boldsymbol{r}_{ab} \cdot \boldsymbol{r}_{ab})^{1/2} \tag{3.1}$$

的函数. 根据复合函数求导法则

$$\frac{\partial U_{ab}}{\partial \boldsymbol{r}_a} = \frac{\mathrm{d}U_{ab}}{\mathrm{d}r_{ab}}\frac{\partial r_{ab}}{\partial \boldsymbol{r}_a}, \qquad \frac{\partial U_{ab}}{\partial \boldsymbol{r}_b} = \frac{\mathrm{d}U_{ab}}{\mathrm{d}r_{ab}}\frac{\partial r_{ab}}{\partial \boldsymbol{r}_b}$$

$$\frac{\partial r_{ab}}{\partial \boldsymbol{r}_a} = \frac{(\boldsymbol{r}_{ab}\cdot\boldsymbol{r}_{ab})^{-1/2}}{2}\sum_{i=1}^{3}\boldsymbol{e}_i\frac{\partial\left(\sum_{j=1}^{3}(r_{ab})_j(r_{ab})_j\right)}{\partial r_{ai}}$$

$$= \frac{1}{2r_{ab}}\sum_{i=1}^{3}\boldsymbol{e}_i\left(\frac{\partial(r_{ab})_j}{\partial r_{ai}}(r_{ab})_j + (r_{ab})_j\frac{\partial(r_{ab})_j}{\partial r_{ai}}\right)$$

注意到

$$\frac{\partial(r_{ab})_j}{\partial r_{ai}} = \frac{\partial(r_{aj}-r_{bj})}{\partial r_{ai}} = \delta_{ij}$$

(其中,克罗内克 δ 符号 δ_{ij} 的定义见附录Ⅰ),就得到

$$\partial r_{ab}/\partial \boldsymbol{r}_a = \boldsymbol{r}_{ab}/r_{ab}. \tag{3.2a}$$

同理有

$$\partial r_{ab}/\partial \boldsymbol{r}_b = -\boldsymbol{r}_{ab}/r_{ab}. \tag{3.2b}$$

因此,$\partial U_{ab}/\partial \boldsymbol{r}_a = -\partial U_{ab}/\partial \boldsymbol{r}_b$. 用

$$\boldsymbol{F}_{ab} = -\partial U_{ab}/\partial \boldsymbol{r}_a; \qquad \boldsymbol{F}_{ba} = -\partial U_{ab}/\partial \boldsymbol{r}_b \tag{3.3}$$

分别表示质点 b 作用在质点 a 上的力和质点 a 作用在质点 b 上的力,得到

$$\boldsymbol{F}_{ab} = -\boldsymbol{F}_{ba}, \tag{3.4}$$

这就是牛顿第三定律:质点 b 对质点 a 的作用力和质点 a 对质点 b 的反作用力大小相等,方向相反. 它成立的条件是两个质点之间的相互作用能只依赖于它们之间的距离:$U_{ab} = U_{ab}(r_{ab})$.

整个系统的势能等于系统内部质点之间的相互作用势能和外场中的势能两部分之和:

$$U = \sum_{a<b}^{N} U_{ab}(r_{ab}) + U_{外}(\boldsymbol{r}_1,\boldsymbol{r}_2,\cdots,\boldsymbol{r}_N). \tag{3.5}$$

注意,上式右边第一项是对 a 和 b 都由 1 到 N 求和,但要求 $a<b$. 这样一来,一方面保证了 $a\neq b$,另一方面又避免了对同一对质点 a 和 b 的相互作用势能重复计算两次. 将 U 对 $\boldsymbol{r}_a$ 求导再改号就得到作用在质点 a 上的力:

$$\boldsymbol{F}_a = -\frac{\partial U}{\partial \boldsymbol{r}_a} = \sum_{b=1}^{N}{}'\left(-\frac{\partial U_{ab}}{\partial \boldsymbol{r}_a}\right) + \left(-\frac{\partial U_{外}}{\partial \boldsymbol{r}_a}\right), \tag{3.6}$$

其中,求和号上打撇表示在对 b 的求和中去掉 $b=a$ 的项. 上式右边第一项是系统中其他质点作用在质点 a 上的力的合力,称为内力

$$\boldsymbol{F}_{a内} = \sum_{b=1}^{N}{}'\left(-\frac{\partial U_{ab}}{\partial \boldsymbol{r}_a}\right) = \sum_{b=1}^{N}{}'\boldsymbol{F}_{ab}, \tag{3.7}$$

而第二项是质点 a 在外场中所受的力，称为外力，

$$\boldsymbol{F}_{a外} = -\partial U_{外}/\partial \boldsymbol{r}_a. \tag{3.8}$$

于是(3.6)式成为

$$\boldsymbol{F}_a = \boldsymbol{F}_{a内} + \boldsymbol{F}_{a外}. \tag{3.9}$$

对于上述 N 个质点的质点组，在笛卡儿坐标中写出拉格朗日方程，得到

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}_i^{(a)}} = \frac{\partial L}{\partial x_i^{(a)}}, \qquad a = 1,\cdots,N; i = 1,2,3.$$

将 §1.2.1 的(1.6)式和(1.7)式代入并利用本节的(3.9)式，得到

$$\mathrm{d}\boldsymbol{p}_a/\mathrm{d}t = \boldsymbol{F}_a = \boldsymbol{F}_{a内} + \boldsymbol{F}_{a外}. \tag{3.10}$$

这里的 $a = 1,2,\cdots,N$ 共 N 个方程，将它们全部求和，得到

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{a=1}^{N}\boldsymbol{p}_a\right) = \sum_{a=1}^{N}\boldsymbol{F}_{a内} + \sum_{a=1}^{N}\boldsymbol{F}_{a外}. \tag{3.11}$$

此式左边的和式是质点组中所有质点的动量的矢量和，称为质点组的总动量

$$\boldsymbol{P} = \sum_{a=1}^{N}\boldsymbol{p}_a, \tag{3.12}$$

而(3.11)式右边第一项是作用在质点组内所有质点上的内力的矢量和. 对于任意两个粒子 a 和 b，在上述矢量和中，既包含 b 对 a 的作用力，又包含 a 对 b 的反作用力，因而由牛顿第三定律(3.4)式知

$$\sum_{a=1}^{N}\boldsymbol{F}_{a内} = \sum_{a<b}^{N}(\boldsymbol{F}_{ab} + \boldsymbol{F}_{ba}) = 0. \tag{3.13}$$

即：作用在质点组内所有质点上的内力的矢量和为零. 将(3.12)，(3.13)式代入(3.11)式，得到

$$\frac{\mathrm{d}\boldsymbol{P}}{\mathrm{d}t} = \sum_{a=1}^{N}\boldsymbol{F}_{a外}. \tag{3.14}$$

即：质点组的总动量对时间的导数等于作用在质点组内所有质点上的外力的矢量和. 这称为质点组的动量定理.

如果不存在外场，则

$$U_{外} = 0, \qquad \boldsymbol{F}_{a外} = 0, \qquad a = 1,2,\cdots,N.$$

此时，由(3.14)式知

$$\dot{\boldsymbol{P}} = 0, \qquad (\text{无外场时}), \tag{3.15}$$

即：无外场时(不受外力作用时)质点组的总动量守恒.

(二)质心　质心系

(3.12)式中定义的 $\boldsymbol{P}$ 是质点组作为一个整体的总动量，相应地可以定义质点组作为一个整体的质量 M 和速度 $\boldsymbol{V}_0$：

$$M = \sum_{a=1}^{N} m_a, \tag{3.16}$$

$$\boldsymbol{P} = M\boldsymbol{V}_0. \tag{3.17}$$

将(3.12)式代入(3.17)式得到

$$\boldsymbol{V}_0 = \frac{1}{M}\sum_{a=1}^{N} m_a\boldsymbol{v}_a = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{M}\sum_{a=1}^{N} m_a\boldsymbol{r}_a\right). \tag{3.18}$$

上式右边圆括号中的表达式决定了一个点的坐标

$$\boldsymbol{R}_C = \frac{1}{M}\sum_{a=1}^{N} m_a\boldsymbol{r}_a. \tag{3.19}$$

这个点称为质点组的质心. 由以上两式可见,质点组整体的速度就是质心运动的速度:

$$\boldsymbol{V}_0 = \mathrm{d}\boldsymbol{R}_C/\mathrm{d}t. \tag{3.20}$$

随质心一道运动的参考系称为质心系. 由(3.19)式知,质心系中 $\sum_{a=1}^{N} m_a\boldsymbol{r}_a =$ 常矢量. 对时间求导为零,即

$$\sum_{a=1}^{N} m_a\dot{\boldsymbol{r}}_a = 0 \qquad (\text{在质心系中}). \tag{3.21}$$

如果质点组不处于外场中,则由(3.15)、(3.17)、(3.20)式, $\dot{\boldsymbol{V}}_0 = 0$, $\boldsymbol{V}_0 = \dot{\boldsymbol{R}}_C =$ 常矢量. 即:无外场时,质心以恒速运动,质心系是惯性系.

(三)动量和能量的变换

动量守恒和能量守恒的成立不依赖于参考系的选取. 但是,在不同参考系中,动量和能量所取的值不同.

假定参考系 K′相对于参考系 K 以速度 $\boldsymbol{V}$ 运动,用 $\boldsymbol{v}_a$ 和 $\boldsymbol{v}'_a$ 分别表示第 a 个质点相对于 K 系和 K′系的速度, $\boldsymbol{P}$ 和 $\boldsymbol{P}'$ 分别表示质点组在 K 系和 K′系中的总动量,则由 §1.1.4(4.5)式

$$\boldsymbol{v}_a = \boldsymbol{v}'_a + \boldsymbol{V},$$

因而

$$\boldsymbol{P} = \sum_{a=1}^{N} m_a\boldsymbol{v}_a = \sum_{a=1}^{N} m_a\boldsymbol{v}'_a + \left(\sum_{a=1}^{N} m_a\right)\boldsymbol{V},$$

即

$$\boldsymbol{P} = \boldsymbol{P}' + \left(\sum_{a=1}^{N} m_a\right)\boldsymbol{V} = \boldsymbol{P}' + M\boldsymbol{V}, \tag{3.22}$$

这就是质点组的总动量在相互作匀速运动的坐标系之间的变换规则.

假定质点组不处于外场中. 用 E 和 E' 分别表示这一质点组在 K 系和 K′系中的能量,按 §1.2.1(1.24)式和 §1.1.4(4.5)式,得

$$E = \sum_{a=1}^{N} \frac{1}{2} m_a v_a^2 + U = \sum_{a=1}^{N} \frac{1}{2} m_a (\boldsymbol{v}'_a + \boldsymbol{V})^2 + U$$

$$= \left(\sum_{a=1}^{N} \frac{1}{2} m_a v'^2_a + U \right) + \frac{1}{2} M V^2 + \left(\sum_{a=1}^{N} m_a \boldsymbol{v}'_a \right) \cdot \boldsymbol{V},$$

其中 U 是质点组内各质点间的相互作用能，它只依赖于质点间的距离，因而和参考系无关.

再利用 §1.2.1(1.24)式和本节(3.12)式，得到

$$E = E' + \frac{1}{2} M V^2 + \boldsymbol{P}' \cdot \boldsymbol{V}. \tag{3.23}$$

这就是在以相对速度 $\boldsymbol{V}$ 运动的两个惯性系之间的能量变换规律.

如果 K′是质心系，则 $\boldsymbol{P}' = 0$，上式简化为

$$E = E' + \frac{1}{2} M V^2 \qquad (\text{K}' \text{为质心系}). \tag{3.24}$$

这表明，质点组在参考系 K 中的能量等于质点组在质心系中的能量加上将质点组的总质量附在质心上时质心在 K 系中的动能. 可以将质心系中的能量

$$E' = \sum_{a=1}^{N} \frac{1}{2} m_a \boldsymbol{v}'^2_a + U \tag{3.25}$$

看成质点组的内能. (3.24)式的意义是：质点组在参考系 K 中的能量等于质点组的内能加上质点组整体在 K 系中的动能. 这样，利用质心系就成功地将质点组在 K 系中的总能量分解成：质点组的全部质量集中在质心，在 K 系中的整体动能，和质点在质心系内部运动的能量，两部分之和.

(四)角动量的变换

和讨论动量与能量的变换时一样，假定参考系 K′相对于参考系 K 以速度 $\boldsymbol{V}$ 运动. 按照伽利略变换 §1.1.4(4.3)(4.5)式，$\boldsymbol{r}_a = \boldsymbol{r}'_a + \boldsymbol{V}t$，$\boldsymbol{v}_a = \boldsymbol{v}'_a + \boldsymbol{V}$.

K 系和 K′系的坐标原点有相对运动. 用 $\boldsymbol{L}$ 和 $\boldsymbol{L}'$ 分别表示在 K 系和 K′系中质点组相对于各自的坐标原点的总角动量

$$\boldsymbol{L} = \sum_{a=1}^{N} \boldsymbol{r}_a \times m_a \boldsymbol{v}_a = \sum_{a=1}^{N} \boldsymbol{r}_a \times m_a \boldsymbol{v}'_a + \sum_{a=1}^{N} m_a \boldsymbol{r}_a \times \boldsymbol{V}$$

$$= \sum_{a=1}^{N} \boldsymbol{r}'_a \times m_a \boldsymbol{v}'_a + \boldsymbol{V}t \times \sum_{a=1}^{N} m_a \boldsymbol{v}'_a + \sum_{a=1}^{N} m_a \boldsymbol{r}_a \times \boldsymbol{V}. \tag{3.26}$$

右边第一项是 K′系中质点组的总角动量 $\boldsymbol{L}'$. 第三项中的和式可以用质心的位矢表示为 $M\boldsymbol{R}_C$ [见(3.19)式]，因而当上式的第二项为零时就有

$$\boldsymbol{L} = \boldsymbol{L}' + \boldsymbol{R}_C \times M\boldsymbol{V} = \boldsymbol{L}' + \boldsymbol{R}_C \times \boldsymbol{P}. \tag{3.27}$$

(3.26)式的第二项为零有两种情况：一是 $t = 0$；另一是 K′系为质心系，因为由(3.

21)式知,在质心系中 $\sum_{a=1}^{N} m_a \boldsymbol{v}'_a = 0$. 这表明,当 K′系为质心系时(3.27)式在任意时刻都成立,即在任意惯性系 K 中,质点组的角动量 $\boldsymbol{L}$ 等于质点组在质心系 K′中的角动量 $\boldsymbol{L}'$ 加上质心集中了质点组的总质量以后在这一惯性系中的角动量 $\boldsymbol{R}_C \times \boldsymbol{P}$.

结合(3.24)式和(3.27)式得到重要结论:在任意惯性系中,质点组的能量和角动量分别等于质点组在质心系中的能量和角动量加上质心集中了质点组的总质量以后在这一惯性系中的能量和角动量. 常称质点组在质心系中的能量和角动量为质点组的内能和内禀角动量. 我们得到结论:质点组在任意惯性系中的能量和角动量分别等于质点组的内能和内禀角动量加上质点组整体质量集中在质心上的能量和角动量.

习　题

1. 设质点系统的拉格朗日 $L = T - U$,其中 $U = U(\boldsymbol{r}, \dot{\boldsymbol{r}})$ 是坐标和速度的函数.

(a)证明:和整个系统绕 z 轴转动的角度 φ 对应的广义动量不再是

$$L_z = \sum m_a r_a^2 \sin^2\theta_a \dot{\varphi}_a,$$

如 §1.2.2(2.16)式,而是

$$p_\varphi = L_z - \sum_{a=1}^{N} \boldsymbol{e}_z \cdot [\boldsymbol{r}_a \times \partial U / \partial \dot{\boldsymbol{r}}_a].$$

(b)已知在电磁场中电荷为 e 的粒子 $U = e\varphi - e\boldsymbol{v} \cdot \boldsymbol{A}$,其中 φ 和 $\boldsymbol{A}$ 是电磁场的标势和矢势(见 §2.3.3),求广义动量 p_φ.

2. 质量为 M 半径为 a 的半球形碗,放在光滑的水平桌面上,如图 1. 有一质量为 m 的滑块沿碗的内壁无摩擦地滑下. 用 α 表示滑块位置与球心的联线和竖直向上的直线的夹角. 整个系统起始时静止且 $\alpha = \alpha_0$. 求此滑块滑到 $\alpha = \alpha_1$ 时 $\dot{\alpha}$ 的值.

3. 质量为 m 的质点在三维空间中运动,势能是

$$U = \begin{cases} 0 & z > 0, \\ -U_0 & z \leqslant 0 \end{cases}$$

证明这一质点由 $z > 0$ 区域经过分界面进入 $z < 0$ 区域的运动轨迹等同于光线从空气入射到折射率为 $n = \sqrt{1 + U_0/E}$ 的介质所受到的折射. 其中,$E = mv_1^2/2$ 是质点在 $z > 0$ 区域中的动能.

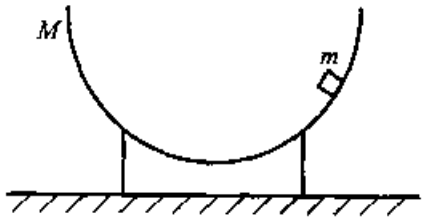


图 1　光滑水平桌面上的半球形碗

4. 求半径为 a,圆心角为 2θ 的均匀扇形薄片的质心.

5. 对于第一章第 9 题,分别在桌面坐标系 K 和沿 X 方向以速度 V 相对于桌面运动的坐标系 K′中写出滑块和斜面的能量以及整个系统的能量在 K 系和 K′系之间的变换方程.

6. 写出角动量的笛卡儿分量 L_x, L_y, L_z 和它的平方 L^2 用球坐标 (r, θ, φ) 表出的表达式.

7. 在下列场中运动的系统,动量 $\boldsymbol{p}$ 的什么分量守恒?角动量 $\boldsymbol{L}$ 的什么分量守恒?

(a)无穷大均匀平面所产生的场;

(b)无限长均匀柱所产生的场;

(c)两个点源所产生的场;

(d)均匀圆环所产生的场;

(e)均匀圆球所产生的场.

提示 在每种情况下,首先考虑什么样的空间变换会使得场源不变(因而所产生的场也不变).

第三章　有心力场中的运动

对于 N 个无约束质点系统，当 $N \geqslant 3$ 时求解运动方程是困难的，只是在 $N = 2$ 的情况下，才易于求解. 解 $N = 2$ 的质点组运动方程的问题称为二体问题. 本章讨论这类问题.

如果两个质点之间有吸引力，则在一定的初始条件之下，它们可能形成一个束缚体系，在有限空间范围内运动；而在另一些初始条件之下，它们也可能相互飞开，无限远离. 如果两个质点先相互飞近，然后再飞开，就称为碰撞. 如果两个质点之间是排斥力，则它们不可能形成束缚体系，而只能发生碰撞. 平行飞来的粒子束和“靶”中的粒子碰撞而散开到不同方向上称为散射.

在本章头两节里，我们对二体问题作一些一般性讨论；第三节讨论平方反比引力；第四节讨论散射问题.

§1.3.1　二体问题　约化质量

用 m_1, m_2 表示所考虑的两个质点的质量，$\boldsymbol{r}_1, \boldsymbol{r}_2$ 表示它们的位矢. 这个二粒子系统的运动可以分解为两个质点作为一个整体的运动和它们之间的相对运动两部分. 整体运动由它们的质心坐标

$$\boldsymbol{R} = \frac{m_1\boldsymbol{r}_1 + m_2\boldsymbol{r}_2}{m_1 + m_2} \tag{1.1}$$

描述；相对运动由它们之间的相对坐标

$$\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2 \tag{1.2}$$

描述，如图 1. 式(1.1)、(1.2)是由 $(\boldsymbol{r}_1, \boldsymbol{r}_2)$ 到 $(\boldsymbol{R}, \boldsymbol{r})$ 的变量变换. 由它们可解出

$$\boldsymbol{r}_1 = \boldsymbol{R} + \frac{m_2}{m_1 + m_2}\boldsymbol{r}; \quad \boldsymbol{r}_2 = \boldsymbol{R} - \frac{m_1}{m_1 + m_2}\boldsymbol{r}. \tag{1.3}$$

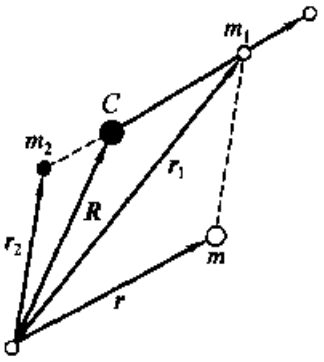


图 1　质心坐标与相对坐标

这一系统的拉格朗日是

$$L = \frac{1}{2}m_1\dot{\boldsymbol{r}}_1^2 + \frac{1}{2}m_2\dot{\boldsymbol{r}}_2^2 - U(r). \tag{1.4}$$

这里假定两质点系统不处于外力场中，并考虑到了两质点之间的相互作用势能是它们之间的距离 $r = |\boldsymbol{r}|$ 的函数. 将(1.3)式代入(1.4)式，得到

$$L = \frac{1}{2}m\dot{\boldsymbol{r}}^2 - U(r) + \frac{1}{2}M\dot{\boldsymbol{R}}^2. \tag{1.5}$$

其中

$$m = \frac{m_1 m_2}{m_1 + m_2} \tag{1.6}$$

称为两质点系统的约化质量，$M = m_1 + m_2$ 是系统的总质量. 由于系统不受外力作用，$\dot{\boldsymbol{R}} = \boldsymbol{V}$ 是常矢量，即：当两质点系统不处于外场中时，它们的质心作匀速直线运动. 这一运动是我们不感兴趣的. 为了消去它，可以转到质心系.

在质心系中 $\dot{\boldsymbol{R}} = \boldsymbol{V} = 0$，拉格朗日成为

$$L = \frac{1}{2}m\dot{r}^2 - U(r). \tag{1.7}$$

在(1.7)式中，不再有 $\boldsymbol{r}_1$ 和 $\boldsymbol{r}_2$，而只有 $\boldsymbol{r}$ 和它的长度 r. (1.7)式中的 L 是一个质量为 m，位矢为 $\boldsymbol{r}$ 的质点(如图 1 中的小圆)在有心力场 $U(r)$ 中的拉格朗日函数. 这样，求解二体问题[(1.4)式]被转化成为求解在有心力场中运动的单体问题[(1.7)式]. 在得到解以后，再由(1.3)式求得所需要的两个质点的位矢. 但应注意，这个在有心力场中运动的“单个质点”的“质量”是约化质量 m，如(1.6)式. 在以下几节中，我们就来求解单个质点 m 在有心力场 $U(r)$ 中运动的问题.

最后，应该指出，如果两个质点轻重悬殊，$m_1 \ll m_2$，则由(1.6)式和(1.3)式可见，

$$\begin{aligned} &\text{当 } m_1 \ll m_2 \text{ 时，} \quad m = \frac{m_1}{1 + m_1/m_2} \approx m_1, \\ &\boldsymbol{r}_1 = \frac{\boldsymbol{r}}{1 + m_1/m_2} \approx \boldsymbol{r}, \quad \boldsymbol{r}_2 = \boldsymbol{R} - \frac{m_1}{m_2}\boldsymbol{r}_1 \approx \boldsymbol{R}. \end{aligned} \tag{1.8}$$

这表明，在此情况下可以近似地认为，重质点 m_2 固定在质心位置上不动，成为“力心”，而轻质点 m_1 在它所产生的有心力场中运动.

§1.3.2　有心力场中运动的一般分析

(一)角动量守恒　等面积定律

在有心力场中，角动量 $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$ 守恒. 因此，在运动过程中，位矢 $\boldsymbol{r}$ 始终垂直于常矢量 $\boldsymbol{L}$，质点的位置始终在一个垂直于 $\boldsymbol{L}$ 的平面上. 这表明，有心力场中的运动是平面运动. 选择坐标系，使这一平面为 xz 面(也就是使 y 轴和矢量 $\boldsymbol{L}$ 平行)，则当采用球坐标系时，$\varphi = 0$ (参看 §1.1.1 图 1).

当这样选择坐标系时，拉格朗日函数 §1.1.1(1.35)式成为

$$L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - U(r). \tag{2.1}$$

在 L 中不包含变量 θ，说明 θ 是循环变量，相应的广义动量（即角动量）

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta}. \tag{2.2}$$

守恒.

对于一个任意的平面运动，角动量 p_θ 有明确的几何意义. 在 t 时刻作位矢 $r(t)$，经过 Δt 以后，它到达 $r(t+\Delta t)$，如图 1. 这一位矢扫过的面积 ΔA 可以用图中所画的三角形来近似地表示. 这一三角形可以近似地看成直角三角形，其两直角边的长度分别是 r 和 $r\Delta\theta$，因而

$$\Delta A = r^2\Delta\theta/2.$$

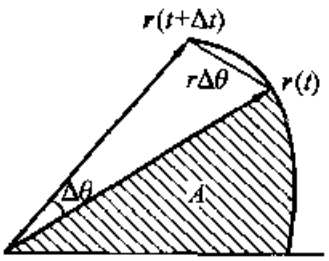


图 1　位矢扫过的面积

用 Δt 除，再取 $\Delta t \to 0$ 的极限，得到

$$\frac{dA}{dt} = r^2\frac{1}{2}\dot{\theta} = \frac{p_\theta}{2m}. \tag{2.3}$$

这表明，当粒子作平面运动时，角动量 p_θ 除以 $2m$ 等于位矢 r 在单位时间扫过的面积. 对于有心力场中的运动，角动量 p_θ 守恒，这一面积是常数

$$\frac{dA}{dt} = 常数 \qquad (有心力场). \tag{2.4}$$

这就证明了，在有心力场中运动的质点的位矢在相等时间内扫过相等的面积. 这称为等面积定律.

（二）能量守恒　运动形式的分类

另一个守恒量是能量

$$E = \frac{\partial L}{\partial \dot{r}}\dot{r} + \frac{\partial L}{\partial \dot{\theta}}\dot{\theta} - L = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2 + U(r). \tag{2.5}$$

将(2.2)式代入，得到

$$E = \frac{1}{2}m\dot{r}^2 + U(r) + \frac{p_\theta^2}{2mr^2}. \tag{2.6}$$

能量的这一表达式中只包含一个变量 r. 它在形式上类似于一维运动的能量. (2.6)式中的第一项 $m\dot{r}^2/2$ 依赖于“速度” $\dot{r}$，是等效的一维运动的“动能”；后两项依赖于“坐标” r，是等效的一维运动的“势能”. 将这一等效势能写为

$$U_{等效} = U(r) + p_\theta^2/(2mr^2), \tag{2.7}$$

其中的第二项称为离心能量：

$$U_{离心} = p_\theta^2/(2mr^2). \tag{2.8}$$

将 $U(r)$ 对 r 求导以后改号，得到作用在质点上的力 $F(r) = -\partial U/\partial r$. 将 $U_{等效}(r)$ 对 r 求导以后改号，得到一个等效力

$$F_{等效} = -\partial U_{等效}/\partial r = F(r) + p_\theta^2/(mr^3), \tag{2.9}$$

其中第二项是惯性离心力：

$$F_{离心} = p_\theta^2/(mr^3) = mr\dot{\theta}^2. \tag{2.10}$$

在 $F_{等效}$ 中出现惯性离心力的原因是由于上述等效一维运动实质上是在以角速度 $\dot{\theta}$ 转动的坐标系中观察质点的运动，而转动坐标是非惯性系. 在后面 §1.5.3 中，将要详细讨论非惯性系中所出现的惯性力.

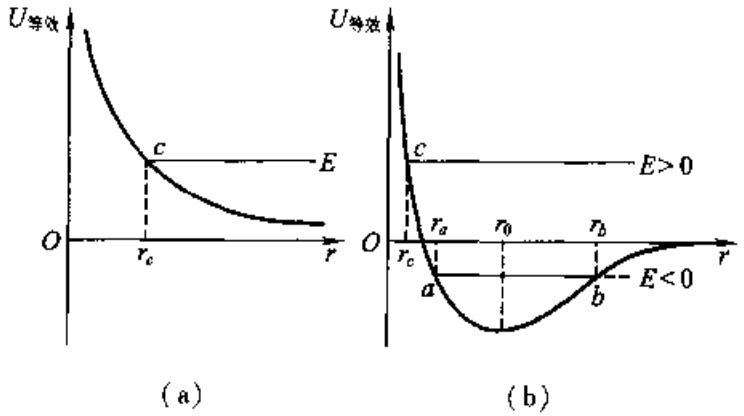


图 2　等效势能曲线

我们来对有心力场中可能发生的运动形式作一般性讨论.

首先，要确定在 $r\to\infty$ 和 $r\to 0$ 两个极端情况下，$U(r)$ 的行为. $r\to\infty$ 时两质点无限远离，它们之间的相互作用消失. 我们令此时的势能为零：

$$\text{当 } r\to\infty, \quad U(r)\to 0. \tag{2.11}$$

结合(2.7)式得到

$$\text{当 } r\to\infty, \quad U_{等效}\to 0. \tag{2.12}$$

另一方面，当 $r\to 0$ 时，$U(r)$ 可能成为正无穷(如果两质点相斥)，也有可能成为负无穷(如果两质点相吸). 无论在哪种情况下，离心能都趋于正无穷，见(2.8)式. 我们假定，即使两质点相吸，使得 $U(r)$ 趋于负无穷，它也不足以和离心能相抵；也就是假定吸力不是太强，因而当 $r\to 0$ 时，$U(r)$ 的绝对值仍小于离心势能. 这样由(2.7)式得到

$$\text{当 } r\to 0 \text{ 时}, \quad U_{等效}\to\infty. \tag{2.13}$$

由(2.12)、(2.13)式可见，等效势能 $U_{等效}$ 随 r 的变化有两种可能的情况：一种情况是，$U_{等效}$ 随 r 的增加而单调下降，如图 2(a)；另一种情况是，$U_{等效}$ 随 r 的增加从正无穷降到小于零，达到极小值 $U_{等效}^{\min}$ 后再上升，从下方趋于横轴，如图 2(b). 在前一情况下，能量 E 一定大于零. 在后一情况下，E 可能大于零，也可能小于零. 但由于(2.6)式中右边第一项不能为负，结合(2.7)式知 E 不能小于 $U_{等效}$，即

$$E\geqslant U_{等效}, \tag{2.14}$$

这一条件限制了质点的运动区域.

由图 2(b)可见，当 $E<0$ 时，质点被限制在有限区域 $r_a\leqslant r\leqslant r_b$ 中运动. 这是束缚运动. 当 $E>0$ 时，运动区域是 $r_c\leqslant r\leqslant\infty$，这是在无限空间中的运动. $E=0$ 是一个过渡情况，在此情况下质点也能运动到无穷远. 也就是说，

$$\left.\begin{array}{ll} E<0 & \text{束缚运动}, \\ E\geqslant 0 & \text{无限运动}. \end{array}\right\} \tag{2.15}$$

在 $E\geqslant 0$ 情况下，质点的运动是 r 由无穷大减小到 r_c，然后又增加到无穷大；与此同时，θ 也将发生相应的变化. 因此，这是从无穷远飞来的质点受到“力心”的作用而改变运动方向飞到无穷远的“散射”问题. 束缚运动和散射是两类不同类型的力学问题. 在本章第四节中将详细讨论散射问题.

(三)运动方程的解

对于有心力场中的运动，求解运动方程最方便的方法是利用守恒律.

由(2.6)式解出 t，得到

$$\mathrm{d}r/\mathrm{d}t=\sqrt{(2/m)(E-U(r))-(p_\theta/(mr))^2}. \tag{2.16}$$

这是一个一阶微分方程，可以分离变量，即

$$\mathrm{d}t=\frac{\mathrm{d}r}{\sqrt{(2/m)(E-U(r))-(p_\theta/(mr))^2}}, \tag{2.17}$$

积分得

$$t=\int\frac{\mathrm{d}r}{\sqrt{(2/m)(E-U(r))-(p_\theta/(mr))^2}}+\text{常数}. \tag{2.18}$$

再由(2.2)式得 $\mathrm{d}t=mr^2\mathrm{d}\theta/p_\theta$，令它和(2.17)式相等，得到

$$\mathrm{d}\theta=\frac{p_\theta\mathrm{d}r/r^2}{\sqrt{2m(E-U)-(p_\theta/r)^2}}, \tag{2.19}$$

积分得

$$\theta=\int\frac{p_\theta\mathrm{d}r/r^2}{\sqrt{2m(E-U)-(p_\theta/r)^2}}+\text{常数}. \tag{2.20}$$

对于具体的有心力场 $U(r)$，算出上式中的积分，就得到质点的运动轨道 $\theta=\theta(r)$.

由(2.18)式算出积分，得到 r 作为 t 的隐函数. 结合 $\theta=\theta(r)$ 又得到 θ 作为 t 的隐函数. 这样，就求得了运动方程的解.

(四)比耐公式

(2.19)式是决定运动轨道 $r=r(\theta)$ 的微分方程. 将它改写为

$$(p_\theta/r^2)(\mathrm{d}r/\mathrm{d}\theta)=\sqrt{2m(E-U)-(p_\theta/r)^2}, \tag{2.21}$$

令 $u=1/r$，则 $\mathrm{d}u/\mathrm{d}\theta=-(1/r^2)\mathrm{d}r/\mathrm{d}\theta$，代入(2.21)式得

$$p_\theta\mathrm{d}u/\mathrm{d}\theta=-\sqrt{2m(E-U)-(p_\theta u)^2}. \tag{2.22}$$

将此式对 θ 求导，得到

$$p_\theta \frac{d^2u}{d\theta^2} = -\frac{-2m(dU/d\theta) - 2p_\theta^2 u(du/d\theta)}{2\sqrt{2m(E-U) - (p_\theta u)^2}} = \frac{m dU/d\theta}{p_\theta du/d\theta} - p_\theta u. \tag{2.23}$$

势能 U 是 r 的函数，因而也是 $u = 1/r$ 的函数，并通过 u 而依赖于 θ. 根据复合函数求导法则

$$\frac{dU}{d\theta} = \frac{dU}{dr}\frac{dr}{d\theta} = Fr^2\frac{du}{d\theta} = = F\frac{1}{u^2}\frac{du}{d\theta}. \tag{2.24}$$

这里用了 $F = -dU/dr$ 和 $dr/d\theta = -r^2 du/d\theta$. 代入(2.23)式，得到

$$p_\theta^2 u^2\left(\frac{d^2u}{d\theta^2} + u\right) = -mF. \tag{2.25}$$

这是决定运动轨道 $u = u(\theta)$ 或 $r = 1/u(\theta)$ 的微分方程的一种更方便的形式，称为比耐公式. 这一公式可以代替(2.19)式，作为求运动轨道的出发点；另一方面，如果已知运动轨道，则由比耐公式可以得到作用力 F.

§1.3.3　平方反比引力　开普勒问题

一种重要的有心力是和 r^2 成反比的引力. 万有引力就是这样一种力. 带异号电荷的质点之间的库仑力也是这种力. 在这一节，我们就来讨论由平方反比引力所形成的力场中的运动.

可以近似地认为，行星只受到太阳引力的作用，而忽略行星之间的相互作用. 由于太阳远比行星重，所以可以采用§1.3.1(1.8)式的近似. 在这一近似下，行星的运动就是在平方反比引力作用下围绕力心(太阳)的运动. 但是，在历史上开始研究行星的运动时，并不知道这种运动所受到的是怎样的力. 从观察到的行星运动规律(开普勒定律)出发，反过来推断行星所受到的力，称为开普勒问题. 在本节的第三小节中将要讨论这个问题.

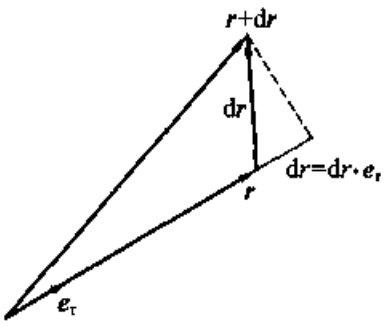


图1　dr 与 e_r 的点积

(一)运动形式的分类

设质点受到平方反比引力的作用. 写作用力为

$$\boldsymbol{F} = -\alpha \boldsymbol{e}_r/r^2, \tag{3.1}$$

其中 $\alpha > 0$ 是表征作用力强度的常数，$\boldsymbol{e}_r$ 是沿位矢 $\boldsymbol{r}$ 方向的单位矢量 $\boldsymbol{e}_r = \boldsymbol{r}/r$. 当质点移动 $d\boldsymbol{r}$ 时，力 $\boldsymbol{F}$ 作功

$$dW = \boldsymbol{F}\cdot d\boldsymbol{r} = -\alpha \boldsymbol{e}_r\cdot d\boldsymbol{r}/r^2 = -\alpha dr/r^2. \tag{3.2}$$

$\boldsymbol{e}_r\cdot d\boldsymbol{r}$ 是 $d\boldsymbol{r}$ 在 $\boldsymbol{r}$ 方向上的分量，因而等于 dr，如图1. 对 dW 积分，得到势能

$$U(r) = -\int \mathrm{d}W = -\frac{\alpha}{r} + 常数. \tag{3.3}$$

加上条件[见§1.3.2(2.11)式]

$$r \to \infty \text{ 时}, \quad U(r) \to 0 \tag{3.4}$$

得到

$$U(r) = -\alpha/r. \tag{3.5}$$

将上式代入§1.3.2(2.7)式,得到等效的一维运动势能

$$U_{等效} = \frac{p_\theta^2}{2mr^2} - \frac{\alpha}{r}. \tag{3.6}$$

容易看到,当 $r \to 0$ 时,上式右边第一项(离心能)比第二项 $U(r)$ 的绝对值增加得更快,因而§1.3.2(2.13)式成立.这样,运动就按 $E < 0$ 和 $E \geqslant 0$ 而区分为束缚运动和无限运动,见§1.3.2(2.15)式.

(二)运动轨道

为了得到平方反比引力情况下的运动轨道,最方便的办法是利用比耐公式.将 $F = -\alpha/r^2 = -\alpha u^2$ 代入§1.3.2(2.25)式,得到

$$\mathrm{d}^2u/\mathrm{d}\theta^2 + u = \alpha m/p_\theta^2. \tag{3.7}$$

作代换

$$\xi = u - \alpha m/p_\theta^2, \tag{3.8}$$

得到

$$\frac{\mathrm{d}^2\xi}{\mathrm{d}\theta^2} + \xi = 0.$$

这是谐振动方程,它的解是 $\xi = A\cos(\theta - \theta_0)$,因而,由(3.8)式有 $u = A\cos(\theta - \theta_0) + \alpha m/p_\theta^2$,或

$$r = \frac{1}{u} = \frac{p_\theta^2/(\alpha m)}{1 + (Ap_\theta^2/(\alpha m))\cos(\theta - \theta_0)}. \tag{3.9}$$

这里的 θ_0 和 A 是两个积分常数.由于所讨论的问题具有球对称性,可以通过转动坐标轴使 $\theta_0 = 0$,再令 $e = Ap_\theta^2/\alpha m$,得到

$$r = \frac{p}{1 + e\cos\theta}, \tag{3.10}$$

其中

$$p = p_\theta^2/(\alpha m). \tag{3.11}$$

(3.10)式是以坐标原点(力心)为焦点的圆锥曲线方程,其中的 e 称为离心率,p 称为焦点参数.

从§1.3.2(2.19)式出发,也可以得到同样的结果,而且可以知道参数 e 与能

量 E 的关系. 这里,我们不去计算该式中的积分,而是将已经得到的 $r = r(\theta)$ 关系[本节的(3.10)式]代入 §1.3.2(2.19)式:

$$d\theta = \frac{p_\theta \cdot (1+e\cos\theta)^{-2} \cdot e\sin\theta \cdot \left(\frac{\alpha m}{p_\theta^2}\right) \cdot (1+e\cos\theta)^2}{\sqrt{2mE + \frac{2m^2\alpha^2}{p_\theta^2}(1+e\cos\theta) - \frac{\alpha^2 m^2}{p_\theta^2}(1+e\cos\theta)^2}} d\theta,$$

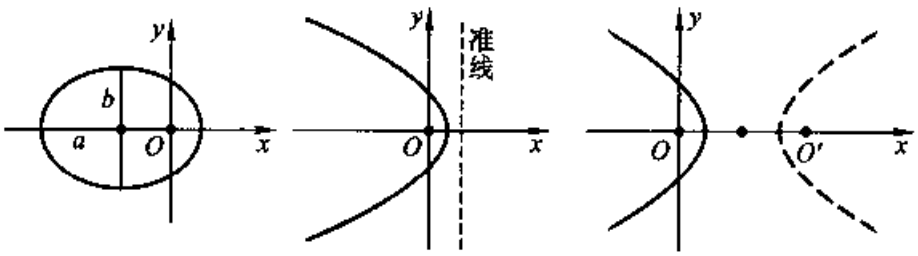
由此得到 e 和 E 的关系

$$e = \sqrt{1 + (2E/m)(p_\theta/\alpha)^2}. \tag{3.12}$$

由解析几何中圆锥曲线的理论知道,(3.10)式按 e 小于、等于和大于 1 而成为椭圆、抛物线和双曲线. 利用(3.12)式就有

$$\left.\begin{array}{llll} E<0 & 则 & e<1 & 轨道为椭圆; \\ E=0 & 则 & e=1 & 轨道为抛物线; \\ E>0 & 则 & e>1 & 轨道为双曲线, \end{array}\right\} \tag{3.13}$$

如图 2. 这和 §1.3.2(2.15)式中按等效势能曲线得到的分类相符. 注意,有心力场的力心 O 在焦点上,而不在椭圆和双曲线的对称中心上. 另外,对于椭圆和抛物线只存在平方反比引力情况. 对于双曲线,图中画为实线的一支对应于平方反比引力,画为虚线的一支对应于平方反比斥力.



(a)椭圆　　(b)抛物线　　(c)双曲线

图 2　圆锥曲线

(三)行星的运动　开普勒问题

由万有引力定律,行星和太阳之间的引力是

$$F = -GMm/r^2, \tag{3.14}$$

其中 m 和 M 分别是行星和太阳的质量,r 是它们之间的距离,G 是万有引力常数,见附录Ⅶ. 和(3.1)式比较可见,在此情况下,

$$\alpha = GMm. \tag{3.15}$$

如果近似地把行星和太阳作为二体问题处理,则前面得到的结果(3.10)~(3.12)式可以直接应用.

但是,历史上开普勒在总结出行星运动的三条经验规律的时候(1609~1619年),并不知道行星和太阳之间有平方反比引力. 只是在 50 年以后(1666 年),牛顿

才根据开普勒三定律推出了万有引力定律.下面就来说明,如何进行这一推论.

开普勒根据布拉赫观测行星位置的数据发现的行星运动三定律是:

第一定律　行星绕太阳作椭圆运动,并以太阳为椭圆的一个焦点;

第二定律　从太阳引向行星的位矢在相等时间内扫过相等的面积;

第三定律　行星公转周期的平方和轨道长半轴的立方成正比.

开普勒第二定律就是§1.3.2中讨论过的等面积定律.那一节的(2.3)式告诉我们

$$\mathrm{d}A/\mathrm{d}t = p_\theta/(2m), \tag{3.16}$$

即,位矢在单位时间扫过的面积(瞬间面积速度)正比于角动量.不论 $\mathrm{d}A/\mathrm{d}t$ 是否为常数,这一定律都成立.现在开普勒第二定律告诉我们, $\mathrm{d}A/\mathrm{d}t$ 是常数,这说明 p_θ 是常数——角动量守恒.但是,角动量守恒表明空间各向同性(见§1.2.2),因此,行星所受到的力只能是有心力.

再根据开普勒第一定律,行星的轨道是以太阳为一个焦点的椭圆.以原点为一个焦点的椭圆方程是

$$r = p/(1 + e\cos\theta), \qquad e < 1. \tag{3.17}$$

代入比耐公式§1.3.2(2.25)式,得到

$$F = -\frac{p_\theta^2}{m}\left(\frac{1 + e\cos\theta}{p}\right)^2\left(-\frac{e}{p}\cos\theta + \frac{1 + e\cos\theta}{p}\right) = -\frac{p_\theta^2/(mp)}{r^2} = -\frac{\alpha}{r^2}. \tag{3.18}$$

这表明,行星受到太阳的平方反比引力.

(3.18)式中的比例系数

$$\alpha = p_\theta^2/(mp) \tag{3.19}$$

除了依赖于行星质量 m 之外,还依赖于 p_θ 和 p 两个参数.为了确定这些参数,利用开普勒第三定律.

将(3.16)式积分,得到从 t_0 到 t 位矢扫过的面积 $A = p_\theta(t - t_0)/2m$.经过一个周期 T ,位矢扫过整个椭圆的面积 πab ,其中 a 和 b 分别是椭圆的长半轴和短半轴,即

$$[p_\theta/(2m)]T = \pi ab. \tag{3.20}$$

由此得到

$$T^2/a^3 = (2\pi m/p_\theta)^2 b^2/a. \tag{3.21}$$

根据解析几何的公式,椭圆的长半轴 a ,短半轴 b 和焦点参数 p 之间有关系 $p = b^2/a$,代入(3.21)式得到

$$T^2/a^3 = 4\pi^2 m(mp/p_\theta^2) = 4\pi^2 m/\alpha. \tag{3.22}$$

这里利用了(3.19)式.根据开普勒第三定律,上述比值是一个与行星无关的常数,

它只可能与太阳的性质(亦即太阳的质量 M)有关.令它等于 $4\pi^2/f(M)$,代入(3.18)式得 $F=-mf(M)/r^2$.按照假定,这一引力不仅存在于太阳和行星之间,而且存在于任何两个物体 m_1 和 m_2 之间,从而 $F_{m_1m_2}=-m_1f(m_2)/r_{12}^2$.显然, $F_{m_1m_2}$ 对 m_1 和 m_2 的交换应该有对称性,所以 $f(m)\propto m$,即 $F_{m_1m_2}\propto -m_1m_2/r_{12}^2$.这就是万有引力定律(3.14)式.这一定律是牛顿根据开普勒的行星运动定律分析归纳得到的.以后,开文迪许在 1798 年用扭秤法证明了万有引力不仅存在于星体之间,也存在于其他一切物体之间.

§1.3.4　碰撞与散射

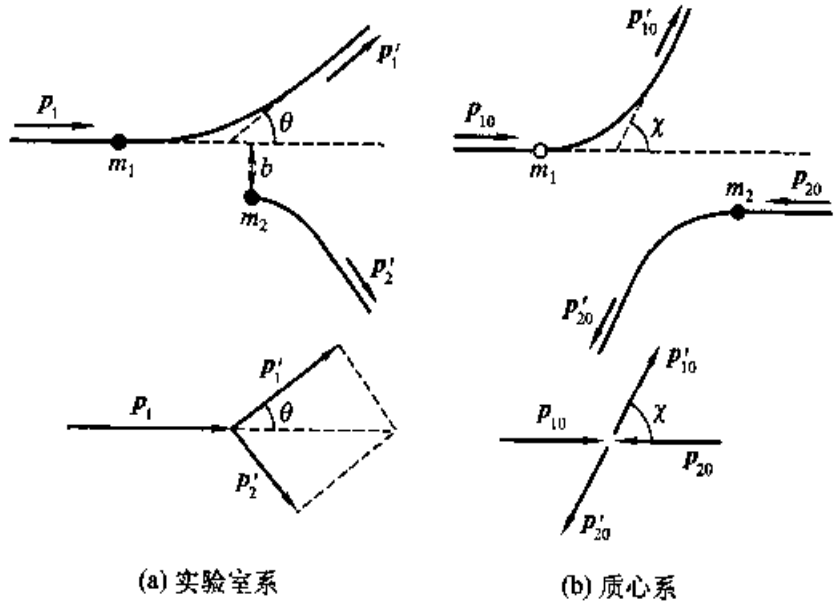
(一)散射问题的提法　散射截面

在§1.3.2 中看到,有心力场中的运动按能量 $E<0$ 和 $E\geqslant 0$ 分为束缚运动和无限运动两类.更一般地说,任一力学系统的运动可以按粒子的位矢在运动的全过程中始终保持有限和可以趋于无穷而区分为束缚运动和无限运动两大类.这两类问题的研究目标和研究方法很不相同.

在这一节里我们讨论后一类运动——无限运动.

这一类运动又有两种情况.一种是两个物体(把它们看成质点)之间的碰撞,例如两个台球的碰撞或彗星向地球飞来又飞走.另一种是一束粒子穿过靶的散射,例如卢瑟福的 α 粒子散射实验.

考虑两个质点之间的碰撞,假定在初始时刻,其中一个质点静止,称为靶粒子.而另一个质点运动,称为射弹粒子.这种参考系称为实验室系(Lab 系).给定的已知条件是:射弹粒子的初速度 v 和靶粒子到射弹粒子沿初速度的延长线之间的距离 b (称为瞄准距离),参看图 1.有了这些条件,能够完全确定运动路径,



(a) 实验室系　　(b) 质心系

图 1　碰撞过程示意图

而我们关心的正是运动路径.例如,对于彗星和地球的"碰撞",我们最关心的不是在无穷远处的彗星,而是它飞过地球时的路径.

散射问题的情况不同.此时,入射的是一束射弹粒子,而靶中也分布有许多靶粒子.不存在唯一确定的瞄准距离 b,因而也不存在唯一确定的运动路径.另一方面,射弹粒子经过靶粒子附近时的运动情况是一个微观过程,无法用仪器探测,不是我们关心的对象.探测器是放在和靶有宏观距离,对于射弹粒子和靶粒子碰撞而言可以看成是无穷远的地方.我们关心的不是单个射弹粒子在单个靶粒子附近的运动,而是从无穷远入射来的一束射弹粒子被靶散射到无穷远处的不同探测器中去的概率.

从粒子源发出的不是单个粒子,而是一束以相同速度 $\boldsymbol{v}$ 运动的粒子流.垂直于束流运动方向作一个平面.单位时间通过这一平面上单位面积的粒子数称为入射流强,用符号 J 表示.

考虑入射束流中的一个粒子 m_1,称它为射弹.取靶中的一个粒子 m_2 为散射中心.在远离 m_2 的地方,m_1 以匀速 $\boldsymbol{v}_1$ 运动,动量是 $\boldsymbol{p}_1$,如图 1a.经过碰撞以后,m_1 的动量成为 $\boldsymbol{p}'_1$,而 m_2 也受到反冲,得到动量 $\boldsymbol{p}'_2$.m_1 在碰撞前和碰撞后的运动方向之间的夹角,也就是 $\boldsymbol{p}'_1$,和 $\boldsymbol{p}_1$ 的夹角,称为散射角,用 θ 表示,见图 1a.

通常情况下,瞄准距离 b 越小,发生的相互作用越强,相应的散射角 θ 越大,b 是 θ 的函数 $b = b(\theta)$(也有例外,如硬锥体上的散射,见习题 7).

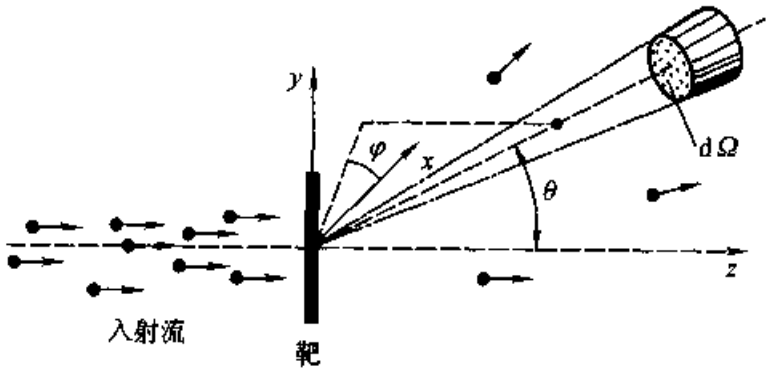


图 2　散射

对于一个固定的靶粒子,束流中的不同粒子有不同的瞄准距离,会被散射到不同的方向.因此,散射问题的提法不是问:射弹粒子被散射到什么方向?而是问:射弹粒子被散射到不同方向的概率是多大?

以散射中心为原点,$\boldsymbol{p}_1$ 方向为极轴作球坐标,如图 2.单位时间里,被散射到 θ,φ 方向的元立体角 $\mathrm{d}\Omega$ 中去的粒子数 $\mathrm{d}N$ 显然正比于 $\mathrm{d}\Omega$,也正比于入射流强 J.$\mathrm{d}N$ 与 $J\mathrm{d}\Omega$ 的比值:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{J}\frac{\mathrm{d}N}{\mathrm{d}\Omega} \tag{4.1}$$

表征粒子被散射到 θ,φ 方向的概率.注意,J 的因次是 $\dim J = \mathrm{L}^{-2}\mathrm{T}^{-1}$,而 $\mathrm{d}N/\mathrm{d}\Omega$ 的因次是 $\dim(\mathrm{d}N/\mathrm{d}\Omega) = \mathrm{T}^{-1}$,由此可见,$\mathrm{d}\sigma/\mathrm{d}\Omega$ 具有面积的因次,称为微分散射

截面.将它对整个立体角积分,得到总散射截面:

$$\sigma = \int (\mathrm{d}\sigma/\mathrm{d}\Omega)\,\mathrm{d}\Omega. \tag{4.2}$$

它也具有面积的因次,表征了入射粒子受到散射的总概率.形象化地,可以将作为散射中心的靶粒子看成等效于面积为 σ 的圆盘,凡是射到这一圆盘上的粒子都受到散射.

对于散射截面,常用的单位是靶恩,毫靶恩和微靶恩(b,mb,μb)

$$1\mathrm{b} = 10^{3}\mathrm{mb} = 10^{6}\mu\mathrm{b} = 10^{-24}\ \mathrm{cm}^{2}. \tag{4.3}$$

让我们来进一步讨论一下微分散射截面的物理意义.

由图 2 可见,所讨论的散射问题具有绕 z 轴的轴对称,散射的微分截面只依赖于 θ,而与 φ 无关.在以靶粒子为心的球面上作一个极角在 θ 到 $\theta + \mathrm{d}\theta$ 之间的环带,如图 3.这一环带所张的立体角是 $\mathrm{d}\Omega = 2\pi\sin\theta\mathrm{d}\theta$.再在远离散射中心(即入射粒子尚未受到散射中心作用处)垂直于入射方向作一个平面,以散射中心到这一平面的垂线和平面的交点为圆心,作一个内外半径分别是 b 和 $b + \mathrm{d}b$ 的圆环,其中 $b = b(\theta)$ 是和散射角 θ 对应的描准距离.凡是射到这个环上的粒子,都被散射到 $\mathrm{d}\Omega$ 内.单位时间射到这个圆环上的粒子数目是 $\mathrm{d}N = J\cdot|2\pi b\mathrm{d}b|$.代入(4.1)式得

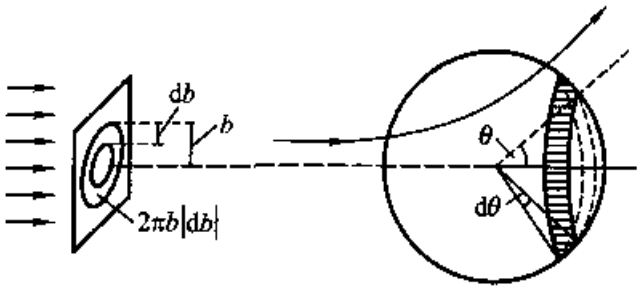


图 3　微分散射截面

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left|\frac{2\pi b\mathrm{d}b}{\mathrm{d}\Omega}\right| = \frac{1}{\sin\theta}\left|b\frac{\mathrm{d}b}{\mathrm{d}\theta}\right|. \tag{4.4}$$

我们看到,散射截面正比于垂直于入射方向的平面上一个圆环的面积 $|2\pi b\mathrm{d}b|$,参看图 3,因而有面积的因次.[这一讨论对某些特殊情况,如本章习题 8 不适用.在那种情况下,只能以散射截面的基本定义(4.1)式和(4.2)式作为计算散射截面的出发点.]

(二)α 粒子的散射　卢瑟福公式

散射实验是用来研究物质微观结构的有力工具.最早的散射实验是盖革和马斯登在 1909 年所做的 α 粒子束透过金属薄膜的散射.正是在这一实验观测的基础上,卢瑟福提出了原子的有核模型,奠定了现代原子物理学和核物理学的基础.

α 粒子由两个质子和两个中子组成,带电荷 $+2e$,其中 e 是电子所带电荷的绝对值(也就是质子所带的电荷).金属的原子核带正电荷 $+Ze$;Z 是原子序数,它和 α 粒子之间有平方反比的库仑斥力.采用高斯单位制,就有

$$F = 2Ze^2 e_r / r^2. \tag{4.5}$$

通常，金属的原子比 α 粒子重得多，可以认为它在散射过程中静止不动，忽略它的反冲. 这样，实验室系和质心系的区别消失，α 粒子是在固定力心的平方反比斥力(4.5)作用下被散射.

(4.5)式和 §1.3.3(3.1)式的差别只在于参数 α 的符号，在 §1.3.3(3.1)式中令

$$\alpha = -2Ze^2, \tag{4.6}$$

就得到本节的(4.5)式. 因此，对 §1.3.3 中的公式作代换(4.6)以后就能用到现在的问题中来. 特别是，对 §1.3.3 的(3.10) ~ (3.12)式作这一代换，立即得到运动轨道，如 §1.3.3 图 2c. 在现在所讨论的平方反比斥力情况下，运动轨道如该图的虚线所示. 为了描述粒子从无穷远沿水平方向入射，要将该图转动 θ_0 角，如图 4. 另外，为了和电子的电荷 e 相区别，将离心率用希腊字母 ε 表示. 这样得到

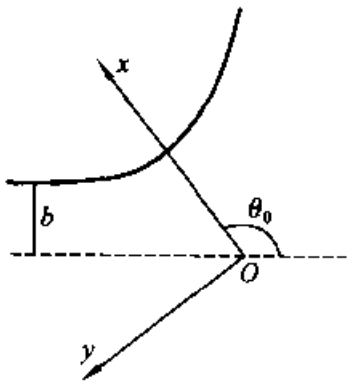


图 4　将 §1.3.3 图 2c 转动 θ_0 角

$$r = \frac{p}{1 + \varepsilon\cos(\theta - \theta_0)}, \tag{4.7}$$

$$p = -\frac{p_\theta^2}{2Ze^2 m}, \qquad \varepsilon = \sqrt{1 + \frac{2E}{m}\left(\frac{p_\theta}{2Ze^2}\right)^2}. \tag{4.8}$$

由于取了靶静止的近似，E 就是入射粒子的能量，而 p_θ 是入射粒子的角动量，由图 1 可见

$$E = \frac{1}{2}mv_\infty^2, \qquad p_\theta = mv_\infty b, \tag{4.9}$$

其中 v_∞ 是射弹粒子远离力心时的速度.

计算散射截面的第一步是计算瞄准距离 b 和散射角 θ 之间的关系. 为此可以利用圆锥曲线的性质. 由图 4 可见，散射角 θ 是双曲线的两条渐进线的夹角. 因而有

$$\cot\frac{\theta}{2} = \frac{b}{a} = \frac{\sqrt{c^2 - a^2}}{a} = \sqrt{\varepsilon^2 - 1}. \tag{4.10}$$

将(4.8)式和(4.9)式代入，得到 $\cot(\theta/2) = mv_\infty^2 b/(2Ze^2)$. 由此得

$$b = \frac{2Ze^2}{mv_\infty^2}\cot\frac{\theta}{2}, \qquad \frac{db}{d\theta} = -\frac{Ze^2}{mv_\infty{}^2\sin^2(\theta/2)}. \tag{4.11}$$

一道代入(4.4)式得到

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}=\frac{Z^2e^4}{m^2v_\infty^4\sin^4(\theta/2)}. \tag{4.12}$$

利用(4.6)式和(4.9)式,可改写上式为

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}=\left(\frac{\alpha}{4E}\right)^2\frac{1}{\sin^4(\theta/2)}. \tag{4.13}$$

这称为卢瑟福公式.

为了更直观,我们来具体讨论一下散射过程中射弹粒子的极坐标(r,ϑ)的变化.[为了避免和最终的散射角θ相混淆,将作为粒子运动过程中的坐标的极角写成花体的ϑ,并将(4.7)式中的θ_0写成ϑ_0].首先,将(4.7)式改写为

$$r=\frac{p_\theta^2/(2Ze^2m)}{\varepsilon\cos(\vartheta-\vartheta_0-\pi)-1}=\frac{p_\theta^2/(2Ze^2m)}{\varepsilon\cos(\vartheta-\theta_0)-1}. \tag{4.14}$$

其中$\theta_0=\pi+\vartheta_0$.

在散射过程中,粒子由无穷远飞来,又被散射到无穷远去,r由∞减小到r_{min},再增加到∞,在此过程中,ϑ单调地减小.相应地$\cos(\vartheta-\theta_0)$由$1/\varepsilon$增大到1,又减小到$1/\varepsilon$.由于ϑ单调减小,所以$\vartheta-\theta_0$由$+\arccos(1/\varepsilon)$减小到0,再减小到$-\arccos(1/\varepsilon)$,参看图5.即:

$$\left.\begin{array}{rccccc} r: & \text{由}\ \infty & \to & r_{min} & \to & \infty, \\ \cos(\vartheta-\theta_0): & \text{由}(1/\varepsilon) & \to & 1 & \to & (1/\varepsilon) \\ \vartheta-\theta_0: & \text{由}\ \arccos(1/\varepsilon) & \to & 0 & \to & -\arccos(1/\varepsilon) \\ \vartheta: & \text{由}\ \pi & \to & \theta_0 & \to & \theta \end{array}\right\}. \tag{4.15}$$

将最后两行相减,得到

$$\pi-\theta_0=\arccos(1/\varepsilon),\qquad \theta-\theta_0=-\arccos(1/\varepsilon),$$

消去θ_0,得

$$2\arccos(1/\varepsilon)=\pi-\theta. \tag{4.16}$$

代入(4.8)、(4.9)式,就得到卢瑟福公式(4.13).

(三)硬表面的散射

以上假定势能$U(r)$是r的连续函数.现在来考虑不满足这一条件的情况,例如在半径为a的硬球上的散

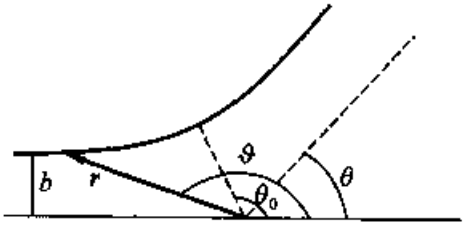


图5　散射角与瞄准距离的关系

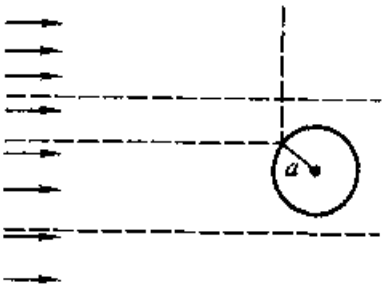


图6　硬球散射

射.此时,势能

$$U(r) = \begin{cases} 0 & \text{当 } r > a, \\ \infty & \text{当 } r \leqslant a. \end{cases} \tag{4.17}$$

由于硬球不吸收能量,射到球面上的粒子,反弹开来,动能不变,动量大小不变,切向动量分量也不变,因而法向动量分量大小不变而反号.这表明,从硬球表面弹开的粒子遵从"入射角等于反射角"的规律.

容易看到,对于(4.17)式的势能,只有射到球上的粒子才发生散射,而从球的四周射过的粒子不发生散射,因而总散射截面就等于球的横截面(参看图6)

$$\sigma = \pi a^2. \tag{4.18}$$

以上两条规律——总散射截面等于散射体的横截面;入射角等于反射角,对一切具有"硬表面"的物体上的散射都适用,而不限于硬球.在习题中有几个例子.

***〔四〕实验室系和质心系中的散射**

在α粒子散射情况下,靶粒子质量和射弹粒子质量相差很大(>100/4倍)可以不区分实验室系和质心系;但是,在一般情况下,不能忽略这一差别.下面讨论这两种参考系之间的变换.

用 m_1 和 m_2 表示射弹粒子和靶粒子的质量.在质心系中,m_1 和 m_2 在碰撞前的动量分别是 $\boldsymbol{p}_{10}$ 和 $\boldsymbol{p}_{20}$,满足

$$\boldsymbol{p}_{10} + \boldsymbol{p}_{20} = 0. \tag{4.19}$$

碰撞以后它们的动量 $\boldsymbol{p}'_{10}$ 和 $\boldsymbol{p}'_{20}$ 仍然满足

$$\boldsymbol{p}'_{10} + \boldsymbol{p}'_{20} = 0. \tag{4.20}$$

$\boldsymbol{p}'_{10}$ 相对于 $\boldsymbol{p}_{10}$ 转过的角度(也就是 $\boldsymbol{p}'_{20}$ 相对于 $\boldsymbol{p}_{20}$ 转过的角度)是质心系中的散射角,用 χ 表示,见图1b.

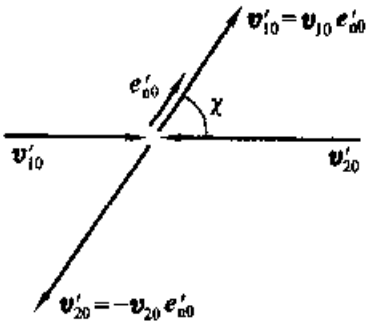


图7　质心系中的速度

(4.19)式和(4.20)式表达了质心系中碰撞过程的动量守恒.除此以外,还有能量守恒

$$(m_1 v_{10}^2/2) + (m_2 v_{20}^2/2) = (m_1 v'^2_{10}/2) + (m_2 v'^2_{10}/2), \tag{4.21}$$

或

$$(p_{10}^2/(2m_1)) + (p_{20}^2/(2m_2)) = (p'^2_{10}/(2m_1)) + p'^2_{20}/(2m_2)), \tag{4.22}$$

由(4.19)、(4.20)和(4.22)式知,碰撞前后的各个动量的大小都相等,

$$|\boldsymbol{p}_{10}| = |\boldsymbol{p}_{20}| = |\boldsymbol{p}'_{10}| = |\boldsymbol{p}'_{20}|, \tag{4.23}$$

能够变化的只是碰撞过程中动量转过的角度 χ.

在实验室系中,碰撞前 m_2 静止不动,m_1 以匀速 v_1 动量 $\boldsymbol{p}_1$ 运动.经过碰撞以后,m_1 的动量成为 $\boldsymbol{p}'_1$,而 m_2 也受到反冲,得到动量 $\boldsymbol{p}'_2$.m_1 在碰撞前和碰撞后的运动方向之间的夹角 θ 是实验室系中的散射角,见图1a.

在这两个参考系中,碰撞过程的动量守恒可写为:

质心系:　$(\boldsymbol{p}_{10} = m_1\boldsymbol{v}_{10}) + (\boldsymbol{p}_{20} = m_2\boldsymbol{v}_{20}) = (\boldsymbol{p}'_{10} = m_1\boldsymbol{v}'_{10}) + (\boldsymbol{p}'_{20} = m_2\boldsymbol{v}'_{20}) = 0$;

实验室系　$(\boldsymbol{p}_1 = m_1\boldsymbol{v}_1) + (\boldsymbol{p}_2 = m_2\boldsymbol{v}_2 = 0) = (\boldsymbol{p}'_1 = m_1\boldsymbol{v}'_1) + (\boldsymbol{p}'_2 = m_2\boldsymbol{v}'_2)$.

在质心系中,四个动量的大小都相等,如(4.23)式,因而

$$v'_{10} = v_{10}, \qquad v'_{20} = v_{20}. \tag{4.24}$$

即：在碰撞过程中，两个粒子的速度大小分别都不变，只是改变了方向. 用 e_{n0} 表示第一个粒子碰撞后速度方向的单位矢量，则

$$\boldsymbol{v}'_{10} = v_{10}\boldsymbol{e}'_{n0}, \qquad \boldsymbol{v}_{20} = -v_{20}\boldsymbol{e}'_{n0}, \tag{4.25}$$

如图 7，图中假定了 $m_1 > m_2$.

按照 §1.2.3(3.18)式，质心系相对于实验室系的速度是（注意 $\boldsymbol{v}_2 = 0$）

$$\boldsymbol{V} = \frac{m_1\boldsymbol{v}_1 + m_2\boldsymbol{v}_2}{m_1 + m_2} = \frac{m_1\boldsymbol{v}_1}{m_1 + m_2}. \tag{4.26}$$

利用 §1.1.4(4.5)式，得到

$$\left.\begin{aligned} \boldsymbol{v}_{10} &= \boldsymbol{v}_1 - \boldsymbol{V} = m_2\boldsymbol{v}_1/(m_1 + m_2), \\ \boldsymbol{v}_{20} &= \boldsymbol{v}_2 - \boldsymbol{V} = -m_1\boldsymbol{v}_1/(m_1 + m_2). \end{aligned}\right\} \tag{4.27}$$

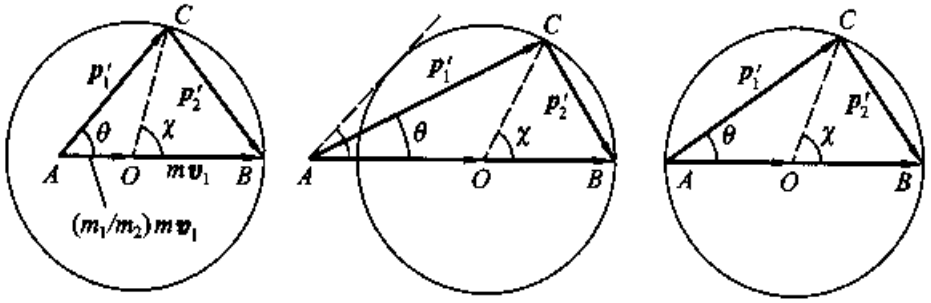
代入(4.25)式，再用 §1.1.4(4.5)式，得到碰撞后粒子在实验室系中的速度：

$$\left.\begin{aligned} \boldsymbol{v}'_1 &= \boldsymbol{v}'_{10} + \boldsymbol{V} = \frac{m_2 v_1 \boldsymbol{e}'_{n0}}{m_1 + m_2} + \frac{m_1\boldsymbol{v}_1}{m_1 + m_2}, \\ \boldsymbol{v}'_2 &= \boldsymbol{v}'_{20} + \boldsymbol{V} = \frac{-m_1 v_1 \boldsymbol{e}'_{n0}}{m_1 + m_2} + \frac{m_1\boldsymbol{v}_1}{m_1 + m_2}. \end{aligned}\right\} \tag{4.28}$$

将此二式分别乘以 m_1 和 m_2 得到

$$\left.\begin{aligned} \boldsymbol{p}'_1 &= mv_1\boldsymbol{e}'_{n0} + (m_1/m_2)m\boldsymbol{v}_1, \\ \boldsymbol{p}'_2 &= -mv_1\boldsymbol{e}'_{n0} + m\boldsymbol{v}_1. \end{aligned}\right\} \tag{4.29}$$

其中 $m = m_1m_2/(m_1 + m_2)$ 是约化质量.



(a) $m_1 < m_2$ 的情况　　(b) $m_1 > m_2$ 的情况　　(c) $m_1 = m_2$ 的情况

图 8　碰撞前后动量的变化

(4.29)式可以用图形表示. 为此，以 O 为圆心，

$$R = mv_1 \tag{4.30}$$

为半径作圆，如图 8. 图中以水平方向为入射方向，即 $\boldsymbol{p}_1$ 的方向，矢量

$$\boldsymbol{AO} = (m_1/m_2)m\boldsymbol{v}_1, \qquad \boldsymbol{OB} = m\boldsymbol{v}_1. \tag{4.31}$$

由(4.30)式知 B 点在圆周上，A 点在圆内或圆外由 $m_1 < m_2$ 或 $m_1 > m_2$ 而定，如图 3a 和 3b. 沿碰撞后第 1 粒子在质心系中的运动方向（$\boldsymbol{e}'_{n0}$ 方向）画一条半径 OC，则 $\angle COB = \chi$ 是质心系中的散射角，而矢量 $\boldsymbol{OC} = mv_1\boldsymbol{e}'_{n0}$ 为(4.29)式右边第一项. 因此，按(4.29)式有：$\boldsymbol{AC} = \boldsymbol{p}'_1$，$\boldsymbol{CB} = \boldsymbol{p}'_2$，$\angle CAB = \theta$ 是实验室系中的散射角. 由图 8 可见

$$\tan\theta = R\sin\chi/(\overline{AO} + R\cos\chi),$$

利用(4.30)、(4.31)式得到

$$\tan\theta = m_2\sin\chi/(m_1 + m_2\cos\chi). \tag{4.32}$$

这就是实验室系中的散射角 θ 和质心系中的散射角 χ 的关系.

由图 8 可见,当 $m_1 < m_2$ 时,碰撞后的粒子可沿任意方向运动(即不受运动学的限制,实际运动方向由动力学决定);当 $m_1 > m_2$ 时,碰撞后的粒子在实验室系中的运动受限制,最大散射角为 θ_{max}:

$$\sin\theta_{max} = R/\overline{AO} = m_2/m_1. \tag{4.33}$$

用 $E_1 = m_1v_1^2/2$ 表示实验室系中射弹粒子的入射能量.靶粒子在碰撞前的能量为零,碰撞后获得能量 $E'_2 = m_2v'^2_2/2$,这是射弹粒子转移给靶粒子的能量.由图 8 可见,最大能量转移发生在质心系中的散射角 $\chi = \pi$ 时,即瞄准距离 $b = 0$ 时.此时,由图 8 知

$$v'_{2对撞} = p_2/m_2 = 2R/m_2 = 2m_1v_1/(m_1 + m_2), \tag{4.34}$$

因而,射弹粒子向靶粒子转移的能量最大只能是

$$E'_{2对撞} = \frac{m_2}{2}v'^2_{2对撞} = \frac{4mE_1}{m_1 + m_2} = \left[1 - \left(\frac{m_1 - m_2}{m_1 + m_2}\right)^2\right]E_1. \tag{4.35}$$

由此可见,$m_1 - m_2$ 愈小(射弹粒子与靶粒子质量相差愈小),能转移的最大能量愈大.这就是核反应堆中采用含轻核多的物质作中子减速剂的原因.

最后,讨论 $m_1 \approx m_2$ 的情况,如图 3c.此时 θ 和 χ 分别是圆周角和圆心角,因而

$$\theta = \chi/2 \qquad (m_1 = m_2). \tag{4.36}$$

而由(4.35)式,

$$E'_{2对撞} = E_1 \qquad (m_1 = m_2). \tag{4.37}$$

即,当 $m_1 = m_2$ 的两个粒子对碰时,在实验室系中射弹粒子停止下来,全部能量转移给靶粒子.

习　题

1.矢量 r,R 和 r_1,r_2 之间有 §1.3.1(1.3)式的关系,试证明:$d^3rd^3R = d^3r_1d^3r_2$.

2.质点受到的有心力为

$$F = -\frac{1}{m}\left(\frac{\alpha}{r^2} + \frac{\beta}{r^3}\right),$$

其中 $\beta < p_\theta^2$.试证明其轨道方程为

$$r = \frac{a}{1 + e\cos k\theta}.$$

其中,$k^2 = 1 - \beta/p_\theta^2$,$a = k^2p_\theta^2/\alpha$,$e = Ak^2P_\theta^2/\alpha$,$A$ 为积分常数.

3.一质点在有心吸力作用下沿圆形轨道运动,力心在此圆的圆周上.求证这一有心力与距离的五次方成反比.

4.在一个顶角为 2α 的圆锥形光滑杯中放置一个质量为 m 的质点.圆锥的轴沿竖直方向,杯口朝上.求证当 $E > 0$ 时,质点在两个水平圆环之间的杯壁上运动,并写出决定这两个圆环半径

的方程.

5. 由椭圆的焦点 F 引出的一线段 FM，以均匀的角速度 ω 绕 F 点转动. 求证此线段与椭圆的交点 M 的速度是 $v = r\omega\sqrt{r(2a-r)}/b$，其中 a 和 b 是椭圆的半长轴和半短轴.

6. (a)有心力势能为 $U = -\alpha/r^n$. 分别对于 $n < 2$，$n = 2$ 和 $n > 2$ 画出有效势能 $U_{有效}$ 的曲线，并分别讨论这三种情况下的各种可能运动方式.

(b)证明，只有当 $n \geqslant 2$ 时，粒子才能落到力心上. 说明其物理原因，并对 $n = 2$ 计算粒子落到力心上的截面.

7. 半径为 a 的硬球势场是

$$U = \begin{cases} 0 & r > a, \\ \infty & r < a. \end{cases}$$

求粒子受这势场散射的有效截面.

8. 设有顶角为 θ，底半径为 a 的硬质圆锥体，锥内势能为 $U = \infty$，锥外势能为 $U = 0$. 粒子平行于锥轴入射，如图 1. 求散射的有效截面.

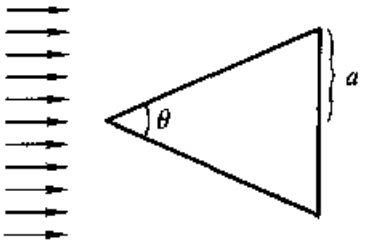


图 1　硬质圆锥体上的散射

9. 一个平面(取为 x,y 面)将空间分为两半. 在这两半区间中的势能分别为 U_1 和 U_2.

$$U = \begin{cases} U_1 & z > 0, \\ U_2 & z < 0, \end{cases}$$

$$u = U_2 - U_1 > 0.$$

质量为 m 的粒子以速度 v 从 $z > 0$ 区域飞到坐标原点处(取 x 轴垂直于速度 v 如图 2)，问：在什么情况下粒子射进 $z < 0$ 区域？在什么情况下粒子反射回 $z > 0$ 区域？求射进 $z < 0$ 区域或反射回 $z > 0$ 区域后粒子的速度 v'.

10. 球形势场

$$U = \begin{cases} 0 & r > a, \\ U_0 > 0 & r < a. \end{cases}$$

求粒子受这势场散射的有效截面.

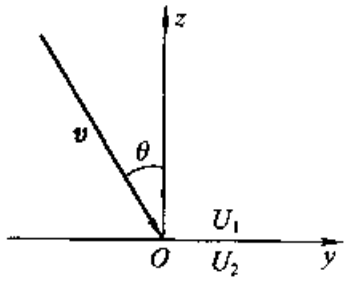


图 2　粒子入射到两区域的分界面上

11. 从 §1.3.4(4.8)，(4.9)，(4.16)式出发，推导卢瑟福公式.

12. 证明在实验室系中靶粒子相对于射弹粒子入射方向的反冲角 $\varphi = (\pi - \chi)/2$，其中 χ 是射弹粒子在质心系中的散射角.

13. 设质量为 m_1 和 m_2 的两个质点之间的相互作用势是半径为 a 的硬球势. 它们在质心系中的散射等效于约化质量为 $m = m_1m_2/(m_1 + m_2)$ 的质点在静止的硬球势场中的散射，如习题 7. 分别就 $m_1 < m_2$，$m_1 > m_2$ 和 $m_1 = m_2$ 三种情况求实验室系中的散射截面.

14. 在 $m_1 > m_2$ 情况下，实验室系中的散射角有极大值 θ_{max}，用 χ_0 表示和 θ_{max} 对应的质心系散射角. 当 $\theta < \theta_{max}$ 时，圆图[§1.3.4 图 8(b)]中的直线 AC 和圆交于两点，说明其意义.

提示　按 §1.3.4 图 1 的方式，画出和这两个交点对应的碰撞过程示意图，并加以讨论.

15. 试用计算机作出以不同瞄准距离 b 入射的 α 粒子受到平方反比斥力散射的动画.

第四章　微　振　动

系统在平衡位置附近的微小振动是一种很普遍的运动形式. 只要力学系统存在平衡位形(即势能有极小值的位形),就会发生这种运动. 本章就来讨论微振动.

在第一节里先讨论一种最简单的情况:单个自由度的无阻尼微振动. 第二节讨论阻尼振动和共振. 第三节讨论多自由度的耦合振动.

§1.4.1　无阻尼的微振动

(一)自由振动方程

我们考虑一种最简单的情况—— 一维系统,忽略阻尼. 用 q 表示这一力学系统的广义坐标,拉格朗日是

$$L = T - U = \frac{1}{2}a(q)\dot{q}^2 - U(q). \tag{1.1}$$

设 q_0 是系统的平衡位置,则在 q_0 处势能 U 有极小值

$$U'(q_0) = 0, \qquad U''(q_0) > 0. \tag{1.2}$$

在 q_0 附近作泰勒展开,保留到二次项

$$U(q) = U(q_0) + \frac{1}{2}U''(q_0)(q - q_0)^2. \tag{1.3}$$

代入(1.1)式,并注意到 $\dot{q}^2 = [\mathrm{d}(q - q_0)/\mathrm{d}t]^2$ 本身已经是二级小量,因此在它的系数 $a(q)$ 中可以令 $q = q_0$,得到

$$L = \frac{a(q_0)}{2}\left[\frac{\mathrm{d}}{\mathrm{d}t}(q - q_0)\right]^2 - \frac{U''(q_0)}{2}(q - q_0)^2 - U(q_0). \tag{1.4}$$

略去常数项 $-U(q_0)$,令

$$q - q_0 = x, a(q_0) = m, U''(q_0) = k, \tag{1.5}$$

得到

$$L = \frac{m}{2}\dot{x}^2 - \frac{1}{2}kx^2, \qquad k > 0. \tag{1.6}$$

这里的 m 是振动体系的等效"质量",k 是等效"劲度系数". 由此得到运动方程

$$\ddot{x} + \omega^2 x = 0, \tag{1.7}$$

其中 $\omega^2 = k/m$.

(二)自由振动方程的解

方程(1.7)的解可以写成

$$x = a\cos\omega t + b\sin\omega t = A\cos(\omega t + \varphi). \tag{1.8}$$

其中 A 和 φ 是两个积分常数，分别称为振幅和初相位. ω 是角频率，或简称为频率. 有时定义 $\nu = \omega/2\pi$ 为频率，但以下我们将不采用它.

利用欧拉公式 $e^{i\theta} = \cos\theta + i\sin\theta$ 可以将(1.8)式写成

$$x(t) = \mathrm{Re}X(t), \qquad X(t) = Ce^{i\omega t}, \tag{1.9}$$

其中符号 Re 表示取复量的实部，$C = A\exp(i\varphi)$ 是复振幅.

将解写成复数形式，运算比较方便. 如果是进行线性运算，可以先用复数形式运算，运算完毕再取实部：设 $x_1 = \mathrm{Re}X_1(t)$，$x_2 = \mathrm{Re}X_2(t)$，则

$$ax_1 + bx_2 = \mathrm{Re}(aX_1 + bX_2), \qquad \dot{x}_1 = \mathrm{Re}\dot{X}_1,$$

其中 a 和 b 是两个实数.

但是，这种方法只适用于线性运算. 对于非线性量(例如能量)就不能先用复数形式运算再取实部.

自由振动系统的能量对 x 和 $\dot{x}$ 都是二次，即

$$E = \frac{m}{2}\dot{x}^2 + \frac{1}{2}kx^2, \tag{1.10}$$

将(1.8)式代入，得到

$$E = m\omega^2A^2/2. \tag{1.11}$$

能量 E 正比于频率的平方和振幅的平方，与时间 t 无关. 这表明自由振动系统是保守系，能量守恒.

能量(1.11)式是二次式，不能用复数形式运算再取实部. 但是，能量也能用 $X(t)$ 表示出来. 由(1.9)式有 $\dot{X}(t) = iA\omega e^{i(\omega t+\varphi)}$，和(1.11)式比较可见

$$E = \frac{m}{2}\omega^2|X|^2 = \frac{m}{2}|\dot{X}|^2 = \frac{m}{2}\dot{X}\dot{X}^*. \tag{1.12}$$

上标 * 表示取复共轭.

(三)受迫振动方程的解

设振子受到一个随时间变化的外力场 $U_e(x,t)$ 的作用，则拉格朗日成为

$$L = \frac{m}{2}\dot{x}^2 - \frac{1}{2}kx^2 - U_e(x,t). \tag{1.13}$$

将 $U_e(x,t)$ 在平衡位置 $x = 0$ 附近展开，

$$U_e(x,t) = U_e(0,t) + U'_e(0,t)x,$$

其中头一项只是时间 t 的函数，对于运动方程没有贡献，可以略去. 令 $U'_e(0,t) = -F(t)$，得到

$$L = \frac{m}{2}\dot{x}^2 - \frac{1}{2}kx^2 + xF(t). \tag{1.14}$$

运动方程成为

$$\ddot{x} + \omega^2 x = F(t)/m. \tag{1.15}$$

我们采用降阶的方法来解这个二阶微分方程. 写

$$\begin{aligned}\ddot{x} + \omega^2 x &= \ddot{x} + i\omega\dot{x} - i\omega\dot{x} + \omega^2 x \\ &= \frac{d}{dt}(\dot{x} + i\omega x) - i\omega(\dot{x} + i\omega x).\end{aligned}$$

因此,令

$$X = (\dot{x} + i\omega x)/(i\omega) = x - i\dot{x}/\omega, \tag{1.16}$$

则(1.15)式成为

$$\dot{X} - i\omega X = F(t)/(i\omega m). \tag{1.17}$$

这是复函数 X 的一阶微分方程. 由(1.16)式可见,求出它的解以后取实部,就得到 $x(t)$.

非齐次一阶线性微分方程(1.17)可以用变动系数法求解. 当 $F(t) = 0$ 时,方程(1.17)的解 $X = C\exp(i\omega t)$;当 $F(t) \neq 0$ 时,可以假定解仍然有这一形式,只是 $C = C(t)$ 是 t 是函数: $X(t) = C(t)\exp(i\omega t)$. 代入(1.17)式,得到决定 $C(t)$ 的方程

$$dC(t)/dt = F(t)e^{-i\omega t}/(i\omega m). \tag{1.18}$$

求积分,得到

$$X(t) = e^{i\omega t}\left\{\int_{t_0}^{t}\frac{F(t)}{i\omega m}e^{-i\omega t}dt + X_0\right\}. \tag{1.19}$$

它可以写成

$$X(t) = X_0 e^{i\omega t} + e^{i\omega t}\int_{t_0}^{t}\frac{F(t)}{i\omega m}e^{-i\omega t}dt. \tag{1.20}$$

右边第一项取实部是自由振动方程(1.7)式的通解;第二项取实部,是强迫振动方程(1.15)式的一个特解. 这表达了一个熟知的事实:非齐次方程的通解等于齐次方程的通解加上非齐次方程的一个特解.

【例 1】 设在初始时刻,振子停在平衡位置不动. 在 $t = 0$ 到 τ 的时间内施以恒力 f,求 $t > \tau$ 以后的振动情况.

解 由所给的初始条件, $X_0 = 0$,将

$$F(t) = \begin{cases} 0 & \text{当 } t < 0 \text{ 和 } t > \tau \\ f & \text{当 } 0 \leqslant t \leqslant \tau \end{cases} \tag{1.21}$$

代入(1.20)式,得到

$$X(t) = e^{i\omega t}\int_0^{\tau}\frac{f}{i\omega m}e^{-i\omega t}dt = \frac{f}{m\omega^2}[e^{i\omega(t-\tau)} - e^{i\omega t}]. \tag{1.22}$$

取实部得到

$$x(t) = A\cos(\omega t + \varphi);$$

$$A = f\sqrt{(1-\cos\omega\tau)^2+\sin^2\omega\tau}/m\omega^2;$$
$$\varphi = \arctan\left[\sin\omega\tau/(1-\cos\omega\tau)\right]. \tag{1.23}$$

(四)受迫振动的能量转移

受迫振动的拉格朗日(1.14)式明显含 t,因而能量不守恒.施加强迫力的系统和振子系统之间有能量转移.

在 t 时刻,振子的能量是 $E = m(\dot{x}^2+\omega^2x^2)/2 = m\omega^2|X|^2/2$,见(1.12)式.将(1.20)式代入,可以得到 E 作为 t 的函数:

$$E(t) = \frac{1}{2m}\left|\mathrm{i}\omega mA\mathrm{e}^{\mathrm{i}\varphi}+\int_{t_0}^{t}F(t)\mathrm{e}^{-\mathrm{i}\omega t}\mathrm{d}t\right|^2. \tag{1.24}$$

假定在初始时刻振子不振动,然后在一个有限的时间间隔内施加外力 $F(t)$,则可以在(1.24)式中令 $A=0$,$t_0=-\infty$,$t=\infty$,得到外力作用完毕后转移给振子的总能量

$$E(\infty) = \frac{1}{2m}\left|\int_{-\infty}^{\infty}F(t)\mathrm{e}^{-\mathrm{i}\omega t}\mathrm{d}t\right|^2 = \frac{|\tilde{F}(\omega)|^2}{2m}. \tag{1.25}$$

符号 $\tilde{F}(\omega)$ 表示 $F(t)$ 的傅里叶变换.

【例2】 求按本节(1.21)式所给出的外力作用下转移给振子的总能量.

解 对(1.21)式作傅里叶变换:

$$\tilde{F}(\omega) = \int_{-\infty}^{\infty}F(t)\mathrm{e}^{-\mathrm{i}\omega t}\mathrm{d}t = \int_0^{\tau}f\mathrm{e}^{-\mathrm{i}\omega t}\mathrm{d}t = \frac{\mathrm{i}f}{\omega}(\mathrm{e}^{-\mathrm{i}\omega\tau}-1).$$

代入(1.25)式得到总的能量转移

$$E(\infty) = [f^2/(2m\omega^2)]\left[(\cos\omega\tau-1)^2+\sin^2\omega\tau\right]. \tag{1.26}$$

将(1.23)式代入(1.11)式也会得到同一结果.

§1.4.2 阻尼振动 共振

上一节讨论了无阻尼的受迫振动,一个重要的特殊情况是周期性外场

$$F(t) = f\cos(\gamma t+\beta). \tag{2.1}$$

代入上节(1.20)式再取实部,

$$x = \mathrm{Re}\left[X_0\mathrm{e}^{\mathrm{i}\omega t}+\mathrm{e}^{\mathrm{i}\omega t}\int_{t_0}^{t}\frac{F(t)}{\mathrm{i}\omega m}\mathrm{e}^{-\mathrm{i}\omega t}\mathrm{d}t\right] = \frac{f}{m(\omega^2-\gamma^2)}\cos(\gamma t+\beta)+\mathrm{Re}(X_1\mathrm{e}^{\mathrm{i}\omega t}).$$

积分下限的贡献已经归并到第二项中.令 $X_1 = A\mathrm{e}^{\mathrm{i}\varphi}$,得到

$$x = A\cos(\omega t+\varphi)+\frac{f}{m(\omega^2-\gamma^2)}\cos(\gamma t+\beta). \tag{2.2}$$

我们看到,质点的运动是按本征频率 ω 的振动和按强迫力频率 γ 的振动这样两部

分的叠加.

当强迫力的频率 γ 等于本征频率 ω 时,(2.2)式第二项的分母成为零,出现振幅趋于无穷大的非物理情况.其原因是由于忽略了阻尼.实际系统总是有阻尼.当振幅不大因而速度不大时,忽略阻尼是一个可以接受的近似.但是,当 $\gamma=\omega$ 时,如果忽略阻尼,振幅将不停地增长,最终会破坏容许忽略阻尼的条件,无阻尼振动的公式(2.2)式将不再成立.此时必须考虑阻尼.当考虑了阻尼以后,即使外力频率等于振子的本征频率,振幅也不会趋于无穷,而只会变得很大.当外力频率等于振子的本征频率时振幅变大的现象称为共振.在这一节里,我们来讨论这一重要现象.在第(一)小节先讨论无阻尼的共振,第(二)小节讨论阻尼振动方程的解,第(三)小节讨论有阻尼的共振,第(四)小节讨论在共振频率附近,振幅、相位和能量转移随频率的变化.

(一)无阻尼的共振

我们从(2.2)式出发.当 $\gamma=\omega$ 时,右边第二项的分母为零,因而此式不能直接应用.应注意到,在 $\gamma=\omega$ 情况下,(2.2)式右边的两项有相同的频率,而第一项的振幅 A 是待定常数.不能从第二项的振幅趋于无穷就认为整个振动的振幅趋于无穷.正确的做法是:先假定 $\gamma\neq\omega$,并在右边第一项中分出一部分,放到第二项中,得到

$$x=A\cos(\omega t+\varphi)+\frac{f}{m(\omega^2-\gamma^2)}[\cos(\gamma t+\beta)-\cos(\omega t+\beta)]. \quad (2.3)$$

注意,这里的 A 和 φ 是另外两个常数,和(2.2)式中的 A 和 φ 不同.

(2.3)式右边第二项当 $\gamma\to\omega$ 时成为 0/0 型不定式,利用洛必达法则得到

$$\lim_{\gamma\to\omega}\frac{f}{m(\omega^2-\gamma^2)}[\cos(\gamma t+\beta)-\cos(\omega t+\beta)]=\frac{ft}{2m\omega}\sin(\omega t+\beta).$$

代入(2.3)式,得到

$$x=A\cos(\omega t+\varphi)+ft\sin(\omega t+\beta)/2m\omega. \quad (2.4)$$

右边第二项的振幅和时间 t 成正比.这表明,在共振情况下(外力频率 γ 等于本征频率 ω 时),振动的振幅将随时间的流逝而无限增大.这显然是不合理的.当振幅增长到一定程度时,微振动的假设已不再成立.

然而,实际上不会发生振幅无限增长的情况.因为实际的运动总是有阻尼的.在共振情况下,随着 t 的增加,x 和 $\dot{x}$ 增大,以致即使是近似地也不能忽略阻尼.下面就来讨论阻尼振动问题,然后再考虑有阻尼情况下的共振.

(二)阻尼振动方程的解

在以上的全部讨论中都忽略了周围介质对运动物体的阻力,包括流体介质的粘滞阻力和固体介质之间的摩擦力.介质的阻力使机械运动的能量耗散,转化为热

能,并最终使机械运动停下来(如果没有外力补充能量).

对于介质的阻力,不存在势能函数,不能纳入拉格朗日中.解决办法是在运动方程中加进阻力项.

当速度不很大时,可以近似地认为阻力和速度成正比,即

$$f_{\text{阻力}} = -\alpha\dot{x},$$

负号表示阻力永远和运动方向相反.将它加到运动方程中,得到

$$m\ddot{x} = -kx - \alpha\dot{x}. \tag{2.5}$$

令

$$\omega_0^2 = k/m, \qquad 2\lambda = \alpha/m, \tag{2.6}$$

运动方程成为

$$\ddot{x} + 2\lambda\dot{x} + \omega_0^2 x = 0. \tag{2.7}$$

这是一个线性方程,可以令 $x = \mathrm{Re}X$,其中 X 是复数,满足

$$\ddot{X} + 2\lambda\dot{X} + \omega_0^2 X = 0.$$

我们求下述形式的解:$X = C\exp(\beta t)$.代入(2.7)式,得到特征方程

$$\beta^2 + 2\lambda\beta + \omega_0^2 = 0. \tag{2.8}$$

它的两个解是

$$\beta_1 = -\lambda + \mathrm{i}\omega, \qquad \beta_2 = -\lambda - \mathrm{i}\omega, \tag{2.9}$$

$$\omega = \sqrt{\omega_0^2 - \lambda^2}. \tag{2.10}$$

假定阻尼足够小,使得 $\lambda < \omega_0$.由此得到两个解,通解是它们的叠加:

$$x = \mathrm{Re}\{\mathrm{e}^{-\lambda t}[c_1\mathrm{e}^{\mathrm{i}\omega t} + c_2\mathrm{e}^{-\mathrm{i}\omega t}]\}. \tag{2.11}$$

这是频率为 ω 而振幅按指数衰减的振动.

(三)有阻尼情况下的共振

考虑有阻尼情况下的强迫振动,在(2.7)式右方加上周期性的强迫力,得到

$$\ddot{x} + 2\lambda\dot{x} + \omega_0^2 x = f\cos\gamma t/m. \tag{2.12}$$

复数形式的方程是

$$\ddot{X} + 2\lambda\dot{X} + \omega_0^2 X = f\mathrm{e}^{\mathrm{i}\gamma t}/m. \tag{2.13}$$

由(2.2)式知道,受迫振动中包含以本征频率 ω 振动和以强迫力频率 γ 振动的两项.我们关心后一项,令 $X = C\exp(\mathrm{i}\gamma t)$,代入(2.13)式得到

$$C = \frac{f}{m(\omega_0^2 - \gamma^2 + 2\mathrm{i}\lambda\gamma)} = \frac{f}{m}\frac{(\omega_0^2 - \gamma^2) - 2\mathrm{i}\lambda\gamma}{(\omega_0^2 - \gamma^2)^2 + 4\lambda^2\gamma^2} = c\mathrm{e}^{\mathrm{i}\delta}, \tag{2.14}$$

其中

$$c = \frac{f}{m\sqrt{(\omega_0^2 - \gamma^2)^2 + 4\lambda^2\gamma^2}}, \qquad \tan\delta = \frac{2\lambda\gamma}{\gamma^2 - \omega_0^2}, \qquad \delta < 0. \tag{2.15}$$

$\delta < 0$ 是由于 $\sin\delta < 0$，见(2.14)式. 由此得到方程(2.12)的一个特解

$$x = \mathrm{Re}[ce^{i(\gamma t+\delta)}] = c\cos(\gamma t + \delta). \tag{2.16}$$

非齐次方程(2.12)的通解是它的特解(2.16)加上齐次方程的通解(2.11)

$$x = Ae^{-\lambda t}\cos(\omega t + \varphi) + c\cos(\gamma t + \delta). \tag{2.17}$$

其中 c 和 δ 由(2.15)式给出，A 和 φ 由初始条件决定.

(2.17)式的第一项含有随时间指数衰减的因子 $\exp(-\lambda t)$，经过足够长时间以后，只剩下第二项，即(2.16)式. 我们看到，有阻尼的受迫振子，经过足够长的时间后，完全按强迫力的频率振动. 振动的位相落后于强迫力的位相[$\delta < 0$，见(2.15)式].

当外力频率等于系统的本征频率 $\gamma = \omega_0$ 时，发生共振●. 此时，按(2.15)式，振幅 c 很大，但并不随时间的增长而无限增大.

(四)通过共振时能量吸收率和相位的变化

我们考虑接近共振的情况，令 $\gamma = \omega_0 + \varepsilon$，$|\varepsilon| \ll \omega_0$. 此时有(注意，按假设 λ 也是小量)：

$$\gamma^2 - \omega_0^2 = (\gamma + \omega_0)(\gamma - \omega_0) = (2\omega_0 + \varepsilon)\varepsilon \approx 2\omega_0\varepsilon,$$

$$\lambda\gamma \approx \lambda\omega_0,$$

代入(2.15)式得到

$$c = \frac{f}{2m\omega_0\sqrt{\varepsilon^2 + \lambda^2}}, \qquad \tan\delta = \frac{\lambda}{\varepsilon}. \tag{2.18}$$

在共振处，$\varepsilon = 0$，$\tan\delta = \infty$，$\delta = -\pi/2$[注意按(2.15)式，$\delta < 0$]. 在远离共振的地方 $|\varepsilon| \gg \lambda$，但仍假定有 $|\varepsilon| \ll \omega_0$，$\tan\delta \to 0$，因而有 $\delta \to 0$(当 $\varepsilon < 0$)或 $\delta \to -\pi$(当 $\varepsilon > 0$). 由此可见，当外力的频率由低到高通过共振频率时，振动的相位改变 $-\pi$，在共振点处相位等于 $-\pi/2$，如图 1(a).

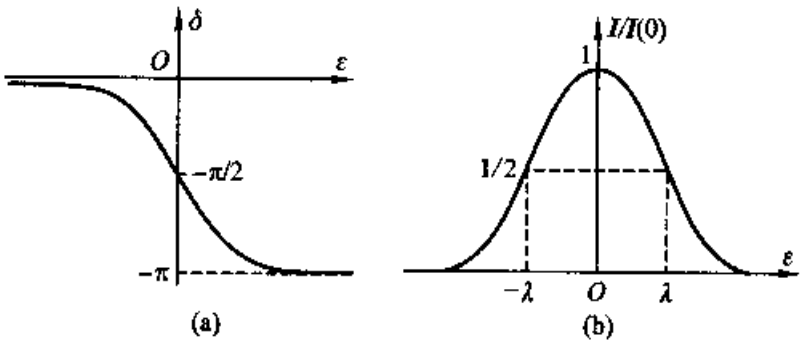


图 1 强迫力频率通过共振频率时相位和能量吸收率的变化
(a)相位的变化 (b)能量吸收率的变化

当振动达到稳定如(2.16)式时，振子的能量不再变化，克服阻尼所消耗的能

● 共振点的标志是平均能量吸收率 $\bar{I}$ 有极大值，相位 δ 通过 $-\pi/2$，参看图 1 和式(2.21).

量通过从外力源吸收能量来补充. 单位时间从外力源吸收的能量 I 等于克服阻力 $-\alpha\dot{x} = -2\lambda m\dot{x}$ 在单位时间内作的功：

$$I = 2\lambda m\dot{x}\mathrm{d}x/\mathrm{d}t = 2\lambda m\dot{x}^2. \tag{2.19}$$

将(2.16)式代入，得到

$$I = 2\lambda mc^2\gamma^2\sin^2(\gamma t + \delta). \tag{2.20}$$

对时间平均得到

$$\bar{I} = \lambda mc^2\gamma^2 = \frac{f^2}{4m}\frac{\lambda}{\varepsilon^2 + \lambda^2}. \tag{2.21}$$

这里用了(2.18)式，并取 $\gamma \approx \omega_0$. $\bar{I}$ 称为平均能量吸收率. 它作为强迫力频率的函数画在图1(b)上. 它呈现为一个峰，在强迫力频率通过共振频率（$\varepsilon = 0$）时达到极大值. 这称为共振吸收.

(2.21)式型的分布称为布雷特-维格纳分布，是描述共振曲线的一种普遍的分布. 容易看到，它在 $\varepsilon = 0$ 时有极大值；而在 $\varepsilon = \lambda$ 时降到极大值的一半. 用 S 表示和这里的能量吸收率对应的某一物理量，它依赖于外来频率 ω. 设在 $\omega = \omega_0$ 时达到共振. 则布雷特-维格纳分布可写成

$$S(\omega) = S_0\frac{\Gamma^2/4}{(\omega - \omega_0)^2 + \Gamma^2/4}, \tag{2.22}$$

其中 ω_0 是共振峰的位置，S_0 是共振峰的强度，Γ 是共振峰的宽度，即 S 降到 $S_0/2$ 时峰的宽度，见图2.

上两节讨论了单个振子. 在实际问题中常常会遇到多个相互耦合的振子. 下一节就来讨论多自由度的耦合振动. 首先，以简单的弱耦合二振子系统为例，介绍“集体振动模式”的概念，然后再推广到一般情况. 为了进行这一推广，在下一节里还将讨论一个有普遍意义的数学问题——对称矩阵的本征值与本征矢. 它不仅对于这里讨论的宏观系统的的振动问题，而且对于以后将要讨论的刚体转动问题（见§1.5.1）以及微观粒子运动问题（见第六篇），都有重要的应用.

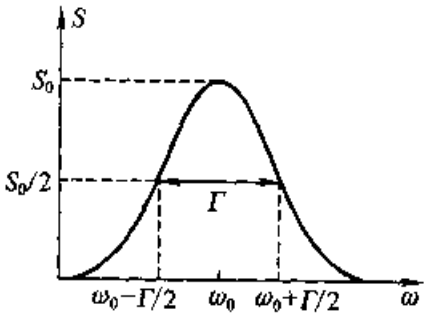


图2　布雷特-维格纳分布

§1.4.3　多自由度的耦合振动

在实际中常常遇到多个振子耦合的系统. 例如，晶体中的原子全部位于格点上时，势能有极小值. 每个原子偏离自己所处的格点的运动都是微振动. 因此，晶体可

以看成一个多振子系统.但是,这些振子不是独立的.任意一个原子偏离格点,会改变它周围的原子的势能场,使得这些原子势能的极小不再在它们原来的格点上.因此,一个原子偏离格点,会带动它周围的原子发生振动,而这些原子又会带动它们周围的原子振动.这样,一个原子偏离平衡位置会引起整个晶体中原子的协调振动.整个晶体中原子的协调振动称为晶体的集体振动模式.在本节里,我们就来研究多自由度耦合振子的集体振动模式.

用 $x_i(i=1,\cdots,s)$ 表示第 i 个振子对自己平衡位置的偏离.系统的拉格朗日可写为(用 $\boldsymbol{x}$ 表示 s 个 $x_i, i=1,\cdots,s$ 的集合)

$$L(\boldsymbol{x},\dot{\boldsymbol{x}}) = \sum_{i=1}^{s}\left(\frac{1}{2}m_i\dot{x}_i^2 - \frac{1}{2}k_i x_i^2\right) + \frac{1}{2}\sum_{i\neq j=1}^{s}\chi_{ij}x_i x_j. \tag{3.1}$$

其中第一项是 s 个自由振子的拉格朗日,第二项是振子之间的耦合.我们的目的是要进行坐标变换

$$x_i = x_i(Q_1,Q_2,\cdots,Q_s), \tag{3.2}$$

使得在新坐标 $Q_\alpha(\alpha=1,\cdots,s)$ 中,拉格朗日退耦合,即使得在拉格朗日中不存在 $Q_\alpha Q_\beta(\alpha\neq\beta)$ 乘积项.用 $\boldsymbol{Q}$ 表示 s 个 $Q_\alpha,\alpha=1,\cdots,s$ 的集合,则退耦合的拉格朗日

$$L(\boldsymbol{Q},\dot{\boldsymbol{Q}}) = \sum_{\alpha=1}^{s}L(Q_\alpha,\dot{Q}_\alpha) = \sum_{\alpha=1}^{s}\left(\frac{1}{2}m_\alpha\dot{Q}_\alpha^2 - \frac{1}{2}k_\alpha Q_\alpha^2\right) \tag{3.3}$$

成为 s 个独立的振子.新的广义坐标 Q_α 就描述系统的集体振动模式,而 $\omega_\alpha = k_\alpha/m_\alpha$ 是集体振动模式的频率.

下面先以一个最简单的情况——弱耦合二振子为例,说明如何从两个单振子的坐标转到集体振动坐标,使振动退耦合.然后再考虑多自由度的一般情况.

(一)弱耦合的二振子系统

设有两个质量为 m 的滑块分别固连在两个劲度系数为 k 的弹簧上,形成两个振子.两者之间用一根软弹簧相联,其劲度系数 $\chi \ll k$,如图 1.

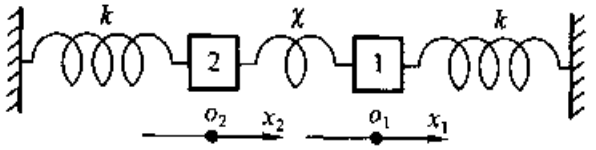


图 1 弱耦合的二振子

以滑块 1 和 2 的平衡位置为原点分别作两根坐标轴 o_1x_1 和 o_2x_2,则势能为

$$U(x_1,x_2) = \frac{1}{2}kx_1^2 + \frac{1}{2}kx_2^2 + \frac{1}{2}\chi(x_1-x_2)^2. \tag{3.4}$$

拉格朗日是

$$L = T - U = \frac{1}{2}m\dot{x}_1^2 + \frac{1}{2}m\dot{x}_2^2 - \frac{1}{2}kx_1^2 - \frac{1}{2}kx_2^2 - \frac{1}{2}\chi(x_1-x_2)^2$$

$$= \frac{1}{2}m\dot{x}_1^2 + \frac{1}{2}m\dot{x}_2^2 - \frac{1}{2}k_0x_1^2 - \frac{1}{2}k_0x_2^2 + \chi(x_1x_2)^2, \tag{3.5}$$

其中 $k_0 = k + \chi$. 由此得运动方程

$$\left.\begin{aligned} \ddot{x}_1 + \omega_0^2 x_1 &= \eta x_2, \\ \ddot{x}_2 + \omega_0^2 x_2 &= \eta x_1. \end{aligned}\right\} \tag{3.6}$$

其中,

$$\omega_0^2 = k_0/m = (k+\chi)/m, \qquad \eta = \chi/m. \tag{3.7}$$

我们来求方程(3.6)的如下形式的解:

$$x_1 = \mathrm{Re}[C_1 e^{i\omega t}], \qquad x_2 = \mathrm{Re}[C_2 e^{i\omega t}]. \tag{3.8}$$

它表示两个滑块以同一频率振动,形成一种集体振动模式. 代入(3.6)式得到

$$\left.\begin{aligned} (\omega_0^2 - \omega^2)C_1 - \eta C_2 &= 0, \\ -\eta C_1 + (\omega_0^2 - \omega^2)C_2 &= 0. \end{aligned}\right\} \tag{3.9}$$

这是两个未知数 C_1, C_2 的一个线性齐次联立方程,它有非零解的条件是系数行列式为零,即

$$\begin{vmatrix} \omega_0^2 - \omega^2 & -\eta \\ -\eta & \omega_0^2 - \omega^2 \end{vmatrix} = 0. \tag{3.10}$$

或 $\omega_0^2 - \omega^2 = \pm\eta$. 它有两个大于零的解:

$$\omega_1 = \sqrt{\omega_0^2 - \eta}, \qquad \omega_2 = \sqrt{\omega_0^2 + \eta}. \tag{3.11}$$

代入(3.9)式,得到 C_1, C_2 的两组解:

$C_1^{(1)} = C_2^{(1)}$, $\quad C_1^{(2)} = -C_2^{(2)}$,

如图2. 将这两组解写成矩阵形式:

$$u^{(1)} = \begin{pmatrix} C_1^{(1)} \\ C_2^{(1)} \end{pmatrix} = C_1^{(1)}\begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

$$u^{(2)} = \begin{pmatrix} C_1^{(2)} \\ C_2^{(2)} \end{pmatrix} = C_1^{(2)}\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

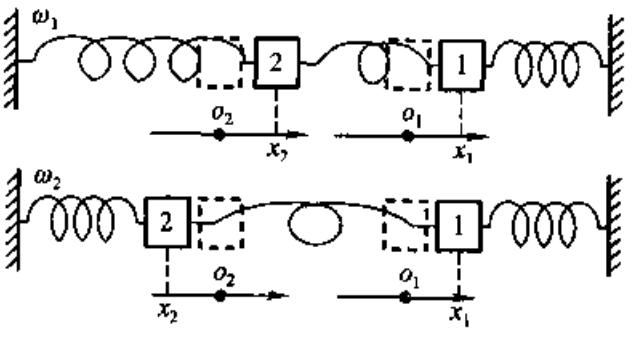


图 2　耦合振子的本征振动模式

容易看到它们相互正交,即

$$\sum_{i=1}^{2} u_i^{(1)} u_i^{(2)} = 0.$$

为了将它们归一化,可以令 $C_1^{(1)} = C_1^{(2)} = 1/\sqrt{2}$. 于是

$$u^{(1)} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad u^{(2)} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{3.12}$$

满足正交归一条件 $\sum_{i=1}^{2} u_i^{(a)} u_i^{(b)} = \delta_{ab}$.

由此可见,这一耦合振子系统有两个本征振动频率 ω_1, ω_2,如(3.11)式.和这两个频率对应有两种确定的集体振动模式

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mathrm{Re}\left[\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{i\omega_1 t}\right], \qquad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mathrm{Re}\left[\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{i\omega_2 t}\right].$$

分别画在图的上下二图上.一般情况下的振动是这两种振动模式的叠加,即

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \mathrm{Re}[A e^{i\omega_1 t}] + \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \mathrm{Re}[B e^{i\omega_2 t}].$$

上式方括号中的两项分别以频率 ω_1 和 ω_2 振动,代表了两种独立的集体振动模式.用 Q_1 和 Q_2 分别表示这两种独立的集体振动模式.由上式得到坐标变换关系:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} Q_1 + \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} Q_2. \tag{3.13}$$

$$x_1 = \frac{1}{\sqrt{2}}(Q_1 + Q_2), \qquad x_2 = \frac{1}{\sqrt{2}}(Q_1 - Q_2). \tag{3.14}$$

将(3.14)式代入(3.5)式,得到新坐标下的拉格朗日

$$L = \frac{1}{2}m\dot{Q}_1^2 - \frac{1}{2}kQ_1^2 + \frac{1}{2}m\dot{Q}_2^2 - \frac{1}{2}(k + 2\chi)Q_2^2. \tag{3.15}$$

这样,两种独立集体振动模式 Q_1 和 Q_2 退耦合.这两种独立集体振动模式的频率分别是 $\omega_1 = \sqrt{k/m}, \omega_2 = \sqrt{(k + 2\chi)/m}$,和(3.7)、(3.11)式一致.

(二)对称矩阵的本征值与本征矢

为了便于将以上讨论的二耦合振子系统推广到任意 s 个耦合振于的系统,将(3.9)式改写成矩阵形式:

$$\begin{pmatrix} \omega_0^2 & -\eta \\ -\eta & \omega_0^2 \end{pmatrix}\begin{pmatrix} C_1 \\ C_2 \end{pmatrix} = \omega^2\begin{pmatrix} C_1 \\ C_2 \end{pmatrix}. \tag{3.16}$$

令

$$S = \begin{pmatrix} \omega_0^2 & -\eta \\ -\eta & \omega_0^2 \end{pmatrix}, \qquad U = \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}, \qquad \lambda = \omega^2. \tag{3.17}$$

(3.16)式成为

$$SU = \lambda U. \tag{3.16'}$$

一列二行矩阵 U 可以看成一个二维空间中的矢量.一般说,2×2 对称矩阵 S 作用在一个任意二维空间矢量上,会改变它的大小和方向,即 SU 和 U 一般不平

行.但是,满足(3.16)式的矢量 U 受到 S 的作用,不改变方向,而只是乘上一个常数 λ.这样的矢量 U 称为矩阵 S 的本征矢,常数 λ 称为和本征矢 U 对应的本征值.方程(3.16)称为对称矩阵 S 的本征值方程.对比(3.9)式和(3.16)式可见,求耦合二振子系统的集体振动模式,归结为求解矩阵 S 的本征值方程(3.16)式.

推广到三维空间,将矢量 $\boldsymbol{u} = \boldsymbol{e}_1 u_1 + \boldsymbol{e}_2 u_2 + \boldsymbol{e}_3 u_3$ 写成一列三行矩阵

$$U \equiv \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}.$$

3×3 的矩阵 S 的本征值方程是

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \lambda \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}. \tag{3.18}$$

或写为

$$SU = \lambda U. \tag{3.18'}$$

如果 $\sigma_{ij} = \sigma_{ji}(i,j = 1,2,3)$,则称矩阵 S 为对称矩阵.对于对称矩阵有如下定理.

定理一 3×3 的对称矩阵 S 有 3 个独立的本征矢.对应的本征值为实数.

证 (3.18)式可写为:

$$SU - \lambda U = (S - \lambda \mathbf{1})U = 0, \tag{3.19}$$

其中 $\mathbf{1}$ 为单位矩阵

$$\mathbf{1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

(3.19)式写成明显形式是

$$\begin{pmatrix} \sigma_{11} - \lambda & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} - \lambda & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} - \lambda \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = 0. \tag{3.19'}$$

这是 3 个未知数 u_1, u_2, u_3 的一个齐次线性方程组.它有非零解的条件是由系数组成的行列式为零

$$\begin{vmatrix} \sigma_{11} - \lambda & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} - \lambda & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} - \lambda \end{vmatrix} = 0. \tag{3.20}$$

这是 λ 的 3 次代数方程,有 3 个根 $\lambda = \lambda_a (a = 1,2,3)$.将它们中的任一个代入(3.19)式,可以得到一个解 $u_1^{(a)}, u_2^{(a)}, u_3^{(a)}$.这就是和本征值 λ_a 对应的本征矢.

但应注意,三次代数方程(3.20)可能有三个实根,也可能有一个实根和一对相互共轭的复根.当 λ 是复数时,代入(3.19)式得到的解也是复数,因而不是有物理意义的本征值和本征矢.然而,这一情况在 S 是对称矩阵时不会出现.这就是定理的第二部分:对称矩阵的本征值是实数.下面来加以证明.

按假定,S 是实对称矩阵

$$\sigma_{ij}^{*} = \sigma_{ji}^{*} = \sigma_{ij} = \sigma_{ji}. \tag{3.21}$$

先假定 S 的本征值 λ 不是实数.将(3.18)式写成

$$\sum_{j=1}^{3} \sigma_{ij} u_j = \lambda u_i, \qquad i = 1,2,3. \tag{3.22}$$

取复共轭

$$\sum_{j=1}^{3} \sigma_{ij}^{*} u_j^{*} = \lambda^{*} u_i^{*}, \qquad i = 1,2,3.$$

改写 $i \rightleftharpoons j$,利用(3.21)式得到

$$\sum_{i=1}^{3} u_i^{*} \sigma_{ij} = \lambda^{*} u_j^{*}, \qquad j = 1,2,3.$$

左右同乘 u_j 并对 j 求和

$$\sum_{i=1}^{3} \sum_{j=1}^{3} u_i^{*} \sigma_{ij} u_j = \lambda^{*} \sum_{j=1}^{3} u_j^{*} u_j, \qquad i = 1,2,3.$$

再将(3.22)式左右同乘 u_i^{*} 对 i 求和

$$\sum_{i=1}^{3} \sum_{j=1}^{3} u_i^{*} \sigma_{ij} u_j = \lambda \sum_{i=1}^{3} u_i^{*} u_i, \qquad i = 1,2,3.$$

上两式的左边相同,右边的和式相等,因而 $\lambda^{*} = \lambda$,即 λ 是实数. **【证毕】**

当 λ 是实数时,由方程(3.19)得到的解 u_1, u_2, u_3 也是实数,组成有物理意义的矢量.因此,3×3 对称矩阵有 3 个实本征值,对应地有三个独立的本征矢.

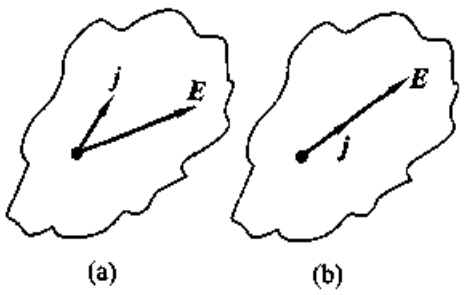


图 3 各向异性介质中的电流密度 j 和场强 E:(a) E 不沿 σ_{ij} 的本征矢方向 (b) E 沿 σ_{ij} 的本征矢方向

这一定理有广泛的应用.为了更形象化,这里再举一个例子.在一个各向同性介质中加电场 E 会产生电流.电流密度 $j = \sigma E$,其中 σ 是电导率.这就是微分形式的欧姆定律[见 §2.4.1(1.23)式].$j \parallel E$ 是由于介质各向同性.如果是各向异性介质,电流 j 一般不平行于 E,如图 3(a).它是 E 的矢量函数 $j = f(E_1, E_2, E_3)$.当电场不太强时,可以作泰勒展开到线性项

$$j_i = \sum_{j=1}^{3} (\partial f_i/\partial E_j) E_j = \sum_{j=1}^{3} \sigma_{ij} E_j.$$

可以证明，σ_{ij} 是对称矩阵. 按定理一，它有三个独立本征矢. 当电场 $\boldsymbol{E}$ 沿 σ_{ij} 的本征矢方向时，$j_i = \sum_{j=1}^{3} \sigma_{ij} E_j = \sum_{j=1}^{3} \lambda \delta_{ij} E_j = \lambda E_i$，因而 $\boldsymbol{j}$ 和 $\boldsymbol{E}$ 平行，如图 3(b).

注意，方程(3.18)是齐次方程，它的解可以乘上任意常数. 因此，和本征值对应的只是本征矢的方向，而相应的本征矢的长度不确定. 可以将本征矢“归一化”成为单位长度，即通过乘上一个常数使得 $u_i (i = 1,2,3)$ 满足

$$\sum_{i=1}^{3} u_i^2 = 1. \tag{3.23}$$

它可以用矩阵形式写为

$$\tilde{U} U = \mathbf{1}. \tag{3.23'}$$

其中 $\tilde{U}$ 是 U 的转置，是一行三列矩阵.

下一个定理讨论不同本征值的本征矢之间的关系.

定理二　对称矩阵对应于不相等本征值的本征矢相互正交.

证　写出和 $\lambda_a \neq \lambda_b$ 对应的本征值方程(3.22)式

$$\sum_{j=1}^{3} \sigma_{ij} u_j^{(a)} = \lambda_a u_i^{(a)}, \qquad \sum_{j=1}^{3} \sigma_{ij} u_j^{(b)} = \lambda_b u_i^{(b)}.$$

将两式分别乘上 $u_i^{(b)}$ 和 $u_i^{(a)}$ 并对 i 求和

$$\sum_{i=1}^{3}\sum_{j=1}^{3} u_i^{(b)} \sigma_{ij} u_j^{(a)} = \lambda_a \sum_{i=1}^{3} u_i^{(b)} u_i^{(a)}, \qquad \sum_{i=1}^{3}\sum_{j=1}^{3} u_i^{(a)} \sigma_{ij} u_j^{(b)} = \lambda_b \sum_{i=1}^{3} u_i^{(a)} u_i^{(b)}.$$

在第一式的左边交换求和指标 $i \rightleftharpoons j$，利用(3.21)式可见，它和第二式的左边相等，因而将两式的右边相减得到

$$(\lambda_a - \lambda_b) \sum_{i=1}^{3} u_i^{(a)} u_i^{(b)} = 0.$$

由于 $\lambda_a - \lambda_b \neq 0$，所以 $\sum_{i=1}^{3} u_i^{(a)} u_i^{(b)} = 0$，即：$\boldsymbol{u}^{(a)} \perp \boldsymbol{u}^{(b)}$.　　【证毕】

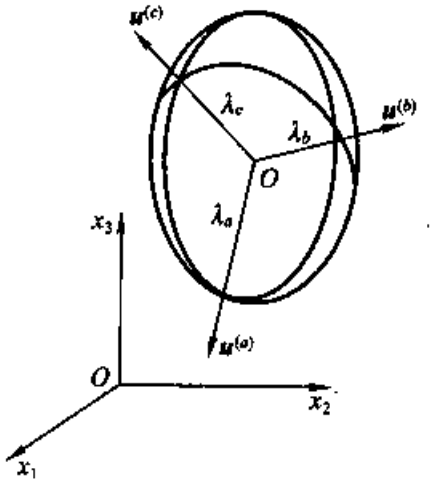


图 4　对称矩阵的本征椭球

从这两个定理可知，3×3 的对称矩阵有三个独立本征矢，对应于三个本征值. 如果这三个本征值互不相等，则对应的三个本征矢相互垂直. 可以为它提供一个直观的几何图像：画出三个本征矢，取其长度分别为对应的本征值. 用它们为主轴作一个椭球，如图 4. 这一椭球就是对称矩阵的几何表示. 只要给出这一椭球就知道了对称矩阵的本征值和本征矢，也

就知道了这一对称矩阵的全部性质.因而这一椭球包含了对称矩阵的全部信息,称为对称矩阵的本征椭球.

用本征椭球的三个主轴(即对称矩阵的三个本征矢)作为标架基矢作一个笛卡儿坐标系.在这一坐标系中,对称矩阵有对角形式

$$S_0 = \begin{pmatrix} \lambda_a & 0 & 0 \\ 0 & \lambda_b & 0 \\ 0 & 0 & \lambda_c \end{pmatrix}. \tag{3.24}$$

下标 0 表明这一矩阵是在以它的三个本征矢为基矢的坐标系中给出的.

当坐标系转动时,矢量 U 变成 U',它的三个分量 $u_{i'}(i' = 1,2,3)$ 是 U 的三个分量 $u_i(i = 1,2,3)$ 的线性组合

$$u_{i'} = \sum_{j=1}^{3} a_{i'i} u_i, \tag{3.25}$$

或用矩阵写为

$$U' = AU. \tag{3.25'}$$

转动矩阵 A 要满足一定的条件,即应能保证归一化的矢量在转动以后仍然归一化.按(3.23′)式就有

$$1 = \tilde{U}'U' = \tilde{U}\tilde{A}AU = \tilde{U}U.$$

对任意矢量 U,上式均应成立,因而 A 应满足条件

$$\tilde{A}A = 1, \qquad \tilde{A} = A^{-1}. \tag{3.26}$$

满足这一条件的矩阵 A 称为正交矩阵.注意,代表物理量的矩阵 S 是对称矩阵,$\tilde{S} = S$,见(3.21)式;而坐标转动矩阵 A 则是正交矩阵,$\tilde{A} = A^{-1}$,如(3.26)式.

坐标转动使代表物理量的矩阵 S 引起的变化可以通过以下考虑得到,即:在原坐标中 S 作用到矢量 U 上得到另一矢量 V:$SU = V$;坐标转动后,这一关系仍应成立,即应有:$S'U' = V'$,或 $S'AU = AV = ASU$.由于 U 是任意矢量,故 $S'A = AS$,利用(3.26)式就有

$$S' = AS\tilde{A}. \tag{3.27}$$

(3.25′)和(3.27)式就是在坐标转动下,矢量和矩阵的变换规律.容易证明,在坐标转动下,代表物理量的矩阵 S 的本征值和本征矢不变,见习题 9.

注意,在变换到另一坐标系时,矩阵 S_0 的各个分量都要发生变化,矩阵不再是对角的,但是物理量的本征值和本征矢不因坐标系的变换而变,因而相应的椭球在空间中的位置和形状不变.这正好象代表物理量(如力)的矢量的三个分量会因坐标系的变换而变,但矢量的方向和长度不变.同样地,代表物理量(如电导率)的矩阵的分量会因坐标系的变换而变,但是由它的本征值和本征矢决定的椭球固定在空间中不动,不依赖于所选用的坐标系.

最后来讨论有相等本征值的情况❶. 如果 $\lambda_a = \lambda_b$ 则和它们对应的本征矢 $\boldsymbol{u}_a$ 和 $\boldsymbol{u}_b$ 不一定正交. 此时有

定理三 如果对称矩阵 S 的两个本征值相等 $\lambda_a = \lambda_b = \lambda$，则和它们对应的本征矢 $\boldsymbol{u}_a$ 和 $\boldsymbol{u}_b$ 的线性组合 $\alpha\boldsymbol{u}_a + \beta\boldsymbol{u}_b$ 也是 S 的对应于同一本征值 λ 的本征矢.

证 写出本征值方程(3.22)式

$$\sum_{j=1}^{3} \sigma_{ij} u_j^{(a)} = \lambda u_i^{(a)},$$

$$\sum_{j=1}^{3} \sigma_{ij} u_j^{(b)} = \lambda u_i^{(b)}.$$

将两式分别乘上 α 和 β 并相加，得

$$\sum_{j=1}^{3} \sigma_{ij} (\alpha u_j^{(a)} + \beta u_j^{(b)}) = \lambda (\alpha u_i^{(a)} + \beta u_i^{(b)}).$$

这就证明了 $\alpha\boldsymbol{u}_a + \beta\boldsymbol{u}_b$ 也是 S 的对应于同一本征值 λ 的本征矢. 【证毕】

两个独立矢量 $\boldsymbol{u}_a$ 和 $\boldsymbol{u}_b$ 的线性组合形成一个平面. 因此，定理三表明，和两个相等本征值对应的不是两个特定的本征矢量，而是一个平面，在这一平面中的任意矢量都是和这一本征值对应的本征矢. 在此情况下，对应的椭球有两个主轴长度相同，是一个旋转椭球. 沿这两个主轴作的椭球的截面是一个圆. 这一截面上的任意矢量都可以看成椭球的主轴. 可以从中任选两个相互垂直的矢量作为椭球的主轴. 这样就得到了在本小节开始作出的论断：对于任意一个 3×3 对称矩阵 S（例如电导率矩阵）总可以找到三个相互垂直的方向，当矢量 $\boldsymbol{u}$ 沿这三个方向时，S 作用到矢量 $\boldsymbol{u}$ 上不改变它的方向.

很容易将以上对 3 维空间中矢量和矩阵的讨论推广到任意 s 维. 除了不能给出直观的三维空间图像（例如三维椭球）之外，其他结论对任意 s 维都成立.

先给出 s 维矢量的定义.

定义一 一组 s 个实数 $u_i (i = 1, \cdots, s)$ 称为 s 维空间中的矢量. 每个 u_i 称为这一矢量的分量.

定义二 两个矢量的对应分量相乘求和 $\sum_{i=1}^{s} u_i v_i$ 称为它们的标积.

定义三 如果两个矢量的对应分量成正比 $u_i = cv_i$ 就称它们相互平行.

定义四 如果两个矢量的标积等于零 $\sum_{i=1}^{s} u_i v_i = 0$ 就称它们相互正交.

矩阵的本征值方程(3.18)式成为

---

❶ 这对应于特征方程(3.20)有重根的情况. 此时(3.19′)的三个方程中，只有一个独立. 由这一个方程可以得到两个线性独立的解，即和 λ_a, λ_b 对应的两个独立本征矢 $\boldsymbol{u}_a$ 和 $\boldsymbol{u}_b$.

$$\begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1s} \\ \vdots & \ddots & \vdots \\ \sigma_{s1} & \cdots & \sigma_{ss} \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_s \end{pmatrix} = \lambda \begin{pmatrix} u_1 \\ \vdots \\ u_s \end{pmatrix}. \tag{3.28}$$

或写为[参看(3.22)式]

$$\sum_{j=1}^{s} \sigma_{ij} u_j = \lambda u_i, \qquad i = 1, \cdots, s. \tag{3.28'}$$

有关对称矩阵本征值和本征矢的三个定理都成立,只要在定理一的表述和三个定理的证明中将维数 3 改为 s,所有的证明和结论都有效.

总起来说,归结为如下定理:

定理四 $s \times s$ 的对称矩阵 S 有 s 个独立的本征矢.对应的本征值为实数.当这 s 个本征值各不相等时,对应的 s 个本征矢相互正交.可以将它们归一化成为一组 s 个正交归一的 s 维矢量 $u_i^{(a)}, i = 1, \cdots, s$ $(a = 1, \cdots, s)$.当在 s 个本征值中有 m 个本征值相等时,对应的 m 个独立本征矢的线性组合形成一个 m 维线性子空间(m 维"平面"),其中的任意矢量都是对称矩阵 S 对应于这一本征值的本征矢.可以从中选出 m 相互正交的矢量并加以归一化,成为这个 m 维线性子空间中的正交归一完备基.对于所有有相等本征值的本征矢都这样处理以后,得到一组 s 个矢量 $u_i^{(a)}, i = 1, \cdots, s$ $(a = 1, \cdots, s)$,满足正交归一条件

$$\sum_{i=1}^{s} u_i^{(a)} u_i^{(b)} = \delta_{ab}. \tag{3.29}$$

(三)多自由度耦合振子的集体振动模式 简正坐标

现在回过头来讨论 s 维耦合振子.首先注意,(3.4)式可以改写为

$$U(x_1, x_2) = \frac{1}{2}(k + \chi)x_1^2 + \frac{1}{2}(k + \chi)x_2^2 - \chi x_1 x_2 = \frac{1}{2}\sum_{ij=1}^{2} k_{ij} x_i x_j, \tag{3.30}$$

$$k_{11} = k_{22} = k + \chi, \qquad k_{12} = k_{21} = -\chi.$$

推广到 s 个自由度的一般情况就有

$$U = \frac{1}{2}\sum_{ij=1}^{s} k_{ij} x_i x_j. \tag{3.31}$$

于是有拉格朗日

$$L = \frac{1}{2}\sum_{i=1}^{s} m_i \dot{x}_i^2 - \frac{1}{2}\sum_{ij=1}^{s} k_{ij} x_i x_j. \tag{3.32}$$

其中的 k_{ij} 对指标 i 和 j 对称❶:

$$k_{ij} = k_{ji}. \tag{3.33}$$

---

❶ 由于 $x_i x_j = x_j x_i$,而 k_{ij} 是和它相乘并对 i, j 求和,所以即使 k_{ij} 中含有反对称分量,这种分量对(3.32)式也没有贡献,见附录(Ⅱ.15)式.

另外，由于 $x_i = 0, i = 1, \cdots, s$ 是所研究的系统的平衡位形，系统偏离平衡 ($x_i \neq 0$) 使势能增加，因而势能 U [(3.31)式]是一个恒正二次型

$$\sum_{ij=1}^{s} k_{ij} x_i x_j \geqslant 0. \tag{3.34}$$

等式只在所有的 s 个 x_i 都为零时才成立. 拉格朗日方程是

$$m_i \ddot{x}_i + \sum_{j=1}^{s} k_{ij} x_j = 0, \qquad i = 1, \cdots, s. \tag{3.35}$$

证　拉格朗日方程是

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{x}_i} - \frac{\partial L}{\partial x_i} = 0, \qquad i = 1, \cdots, s.$$

在将(3.32)式右边第一项(动能项)对 $\dot{x}_i$ 求偏导时，要注意自由角标和求和角标(哑角标)不能重复. 因此在求偏导之前应先将动能项中的求和角标换成另一符号，例如 j

$$\frac{\partial L}{\partial \dot{x}_i} = \frac{1}{2} \frac{\partial}{\partial \dot{x}_i} \sum_{j=1}^{s} m_j \dot{x}_j^2$$

利用 $\partial \dot{x}_j / \partial \dot{x}_i = \delta_{ij}$，就得到 $\partial L / \partial \dot{x}_i = m_i \dot{x}_i$. 对势能项也用类似的方法求导就得到(3.35)式.　　**【证毕】**

为了得到 s 个质点的集体振动模式，我们来求所有的 $x_i, i = 1, \cdots, s$ 以同一频率振动 $x_i = \mathrm{Re}[C_i \exp(\mathrm{i}\omega t)]$ 的解. 代入(3.35)式得

$$-m_i \omega^2 C_i + \sum_{j=1}^{s} k_{ij} C_j = 0, \qquad i = 1, \cdots, s. \tag{3.36}$$

(3.36)式可以写成标准的对称矩阵本征值方程的形式，将它乘以 $1/\sqrt{m_i}$，并令

$$v_i = \sqrt{m_i} C_i, \qquad k'_{ij} = \frac{1}{\sqrt{m_i m_j}} k_{ij}. \tag{3.37}$$

代入(3.36)式得到

$$\sum_{j=1}^{s} k'_{ij} v_j = \omega^2 v_i, \qquad i = 1, \cdots, s. \tag{3.38}$$

对照(3.22)式可见，这是对称矩阵 K' 的本征值方程. 它有 s 个本征值 $\omega_\alpha^2, \alpha = 1, \cdots, s$. 这些本征值不但是实的而且是正的. 为了看到这后一点，用 v_i 乘上式并对 i 求和，得到

$$\omega^2 = \sum_{ij=1}^{s} v_i k'_{ij} v_j \Big/ \sum_{i=1}^{s} v_i^2. \tag{3.39}$$

由(3.34)、(3.37)式就可以看到 $\omega^2 > 0$.

这样，我们就得到了 s 个(实)频率 $\omega_\alpha, \alpha = 1, \cdots, s$. 相应的本征矢 $V_i^{(\alpha)}$ ($v_i^{(\alpha)}, i$

$=1,\cdots,s$)通过(3.37)式决定了 x_i 的振动 $x_i^{(a)}=\mathrm{Re}[(v_i^{(a)}/\sqrt{m_i})\exp(i\omega_a t)]$. 这是所有 s 个自由度 $x_i, i=1,\cdots,s$ 以同一频率 ω_a 按特定的方式(由 $v_i^{(a)}/\sqrt{m_i}$ 决定)的协调振动,是一个集体振动模式,称为简正振动. 总共有 s 个($a=1,\cdots,s$)这样的集体振动模式. 实际的振动是这 s 个集体振动模式(简正振动)的线性叠加:

$$x_i=\sum_{a=1}^{s}\mathrm{Re}\left[A_a\frac{v_i^{(a)}}{\sqrt{m_i}}e^{i\omega_a t}\right], \tag{3.40}$$

A_a 是线性叠加的(复)系数.

用 $Q^{(a)}=A_a e^{i\omega_a t}$ 表示第 a 个集体振动模式的广义坐标,称为简正坐标. 坐标变换式是

$$x_i=\sum_{a=1}^{s}\frac{v_i^{(a)}}{\sqrt{m_i}}Q^{(a)}. \tag{3.41}$$

不同的本征矢 $V^{(a)}$ 有正交性,令 $\mu_a=\sum_{i=1}^{s}v_i^{(a)}v_i^{(a)}$,就有

$$\sum_{i=1}^{s}v_i^{(a)}v_i^{(b)}=\mu_a\cdot\delta_{ab}. \tag{3.42}$$

将 $v_i^{(a)}$ 乘以 $1/\sqrt{\mu_a}$ 可以使它归一化

$$u_i^{(a)}=v_i^{(a)}/\sqrt{\mu_a}, \tag{3.43}$$

$$\sum_{i=1}^{s}u_i^{(a)}u_i^{(b)}=\delta_{ab}. \tag{3.44}$$

由(3.38)式并利用(3.42)式可以得到:

$$\sum_{ij=1}^{s}v_i^{(a)}k'_{ij}v_j^{(b)}=\mu_a\omega_a^2\cdot\delta_{ab}. \tag{3.45}$$

再利用(3.43)式得到

$$\sum_{ij=1}^{s}u_i^{(a)}k'_{ij}u_j^{(b)}=\omega_a^2\delta_{ab}. \tag{3.46}$$

可以证明,在新的广义坐标 $Q^{(a)}$ 下拉格朗日退耦合成为 s 个独立振动

$$L=\frac{1}{2}\sum_{a=1}^{s}\mu_a\dot{Q}^{(a)2}-\frac{1}{2}\sum_{a=1}^{s}\mu_a\omega_a^2Q^{(a)2}. \tag{3.47}$$

见习题 10.

习 题

1. 由质量为 m_1 和 m_2 的两个原子结合成的分子,振动频率为 ω;由它们的质量为 m'_1 和 m'_2 的同位素结合成的分子,振动频率为 ω'. 证明

$$\left(\frac{\omega'}{\omega}\right)^2 = \frac{m_1 m_2 (m'_1 + m'_2)}{m'_1 m'_2 (m_1 + m_2)}.$$

2. 质量为 m 的穿孔小球可以自由地在一光滑直杆上滑动，如图 1. 将一根弹簧系在此球上，弹簧的另一端固定在离杆 l 的钉子上. 在平衡时弹簧的拉力为 F. 求此球在平衡位置附近作微小振动的频率.

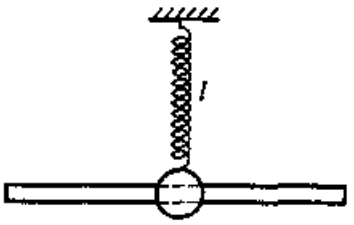


图 1 习题 2 图

3. 频率为 ω 的振子受到外力 F 的作用. 在 $t < 0$ 时没有振动，求 $t > T$ 时的振幅. 假定当 $t < 0$ 和 $t > T$ 时 $F = 0$. 而在 $0 < t < T$ 时，

(a) $F = F_0$，当 $0 < t < T$；

(b) $F = F_0 \sin \omega t$，当 $0 < t < T$.

4. 在强迫力频率通过共振频率时的相位变化曲线[§1.4.2 图 1(a)]的上方画出相位的正切 $\tan\delta$ 的变化曲线和受迫振动振幅的变化曲线.

5. 在阻力比较大，使得 $\lambda \geq \omega_0$ 的情况下，解阻力振动方程 §1.4.2(2.5)式.

6. 质量为 m 的两个原子位于质量为 M 的原子的两侧. 三者在一直线上，平衡时两端的两个原子和中心原子的距离都是 b. 求三者沿联线振动的简正频率和简正坐标.

7. 求第 1 章第 8 题中的平面双摆的简正频率.

8. 坐标转动矩阵 A 按 §1.4.3(3.25′)式定义. 试写出绕 z 轴转 θ 角的转动矩阵 A.

9. 证明坐标转动不改变对称矩阵的本征值和本征矢.

10. 证明正交矩阵 A（即能保证归一化的矢量在变换后仍然归一化的变换矩阵）的矩阵元 a_{ij} 满足条件

$$\sum_i a_{ik} a_{ij} = \delta_{kj}$$

[这也就是 §1.4.3(3.26)式的分量形式.]

11. 证明按 §1.4.3(3.41)式定义的简正坐标将耦合振子退耦合，即 §1.4.3(3.47)式.

第五章　刚体的运动

形状和大小都不变化的物体叫做刚体. 刚体可以看成是一个质点组,其中任何两个质点之间的距离都是常数. 这当然是一个抽象的模型,但是,对于通常的固体,在只考虑它们的整体运动时,常常可以近似地把它们看成刚体.

刚体只有整体运动的自由度. 为了确定刚体在空间的位置,可以作一个固连在刚体中的坐标系 $ox_1x_2x_3$. 它的原点 o 在静止坐标系 $Ox_0y_0z_0$ 中的位矢是 $\boldsymbol{R}_0$,如图 1. 过 o 点作 $oxyz$ 和 $ox_0y_0z_0$ 平行,然后用以下步骤将 $oxyz$ 转到 $ox_1x_2x_3$ 位置:

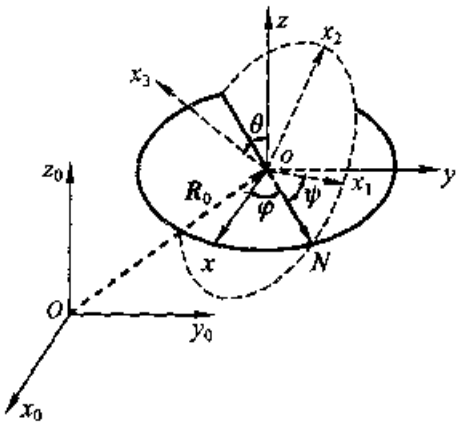


图 1　刚体位置的确定

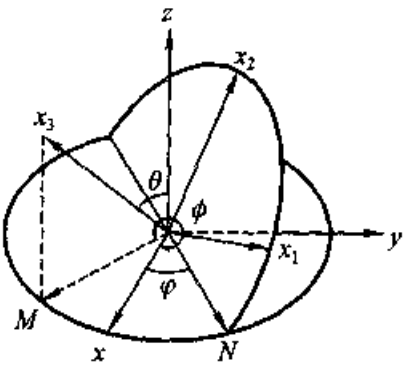


图 2　转动 oz 用的转轴 oN

首先要设法将 oz 转到 ox_3. 用通常决定球坐标中角度的方法,将 ox_3 投影到 xy 平面得到 oM. 令 ox_3 和 oz 的夹角为 θ,oM 和 ox 的夹角为 ϕ,如图 2. 显然,为了将 oz 转到 ox_3 需要转动 θ 角,问题是绕什么轴转? 这个转轴应该垂直于 zx_3 平面,而 oM 是在这一平面上而不是垂直于它. 因此,所需要的转轴 oN 应该垂直于 oM. 它和 ox 的夹角 $\varphi = \phi - 3\pi/2$,如图 2. 这样就得到将 $oxyz$ 转到 $ox_1x_2x_3$ 的步骤如下,参看图 1:

1. 将 $oxyz$ 绕 z 轴转 φ 角,使 ox 转到垂直于 zx_3 平面的位置 oN;

2. 绕 oN 转 θ 角,使 z 轴转到 ox_3 位置;

3. 再绕 ox_3 转 ψ 角,使 oN 轴转到 ox_1 位置.

这样三个角度 φ,θ,ψ(称为欧拉角)和 $\boldsymbol{R}_0$ 的三个分量共同决定了刚体的位置. 因此刚体有六个自由度.

当欧拉角 φ,θ,ψ 不变而 $\boldsymbol{R}_0$ 变化时,称为平动,此时刚体内部任何两点的联线在运动过程中始终保持平行. 如果 $\boldsymbol{R}_0$ 不变而欧拉角 φ,θ,ψ 变化,则称刚体绕固定点 o 转动.

刚体的基本运动形式是平动和转动. 如果将 o 点选为刚体的质心,则刚体的平动和整个刚体的质量集中在质心上的单质点运动规律相同. 体现出刚体特点的主

要运动方式是转动，因此我们将着重讨论刚体的转动.

在本章第一节里，讨论刚体转动的角速度、角动量和转动能量；第二节讨论刚体的运动方程以及它的一些应用.

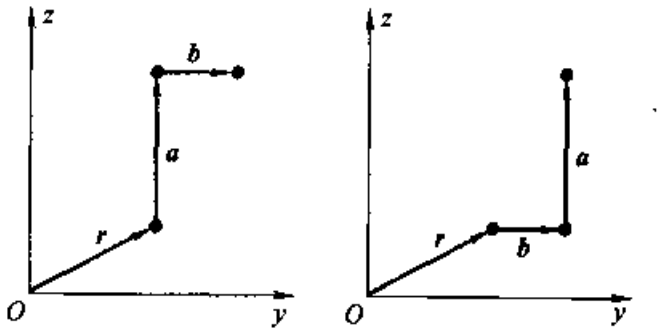
当把参考系固连在有加速运动（特别是有转动）的刚体上时，这种参考系是非惯性系. 本章的最后一节讨论质点在非惯性系中的运动.

§1.5.1　刚体的角速度、角动量与转动能量

刚体转动的最大特点是：绕不同轴的转动之间有相互关联. 这一特点在涉及转动的所有问题中都体现出来，因此我们首先对它作仔细的分析.

（一）不同轴转动之间的关联　角速度矢量

我们知道，沿不同方向的线位移是彼此独立的. 沿某一方向（例如 y 方向）的位移和沿另一方向（例如 z 方向）的位移是在它之前还是在它之后进行无关，如图 1. 这称为平行四边形法则：$\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a}$. 用 T_a 和 T_b 表示将位矢为 $\boldsymbol{r}$ 的点平移 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 的算符，则这一法则可写为：$T_bT_a\boldsymbol{r}=T_aT_b\boldsymbol{r}$，如图 1.



(a) 沿y位移之前先沿z位移　(b) 沿y位移之后再沿z位移

图 1　不同方向位移的独立性

可是，如果是转动，情况就不一样，绕某一轴（例如 y 轴）的转动和绕另一轴（例如 z 轴）的转动是在它之先进行还是在它之后进行有关，如图 2. 这表明，绕不同轴的转动，彼此之间有相互关联. 两个有限转动（即转角不是无穷小的转动）的合成不服从平行四边形法则，不能看成矢量.

但是可以证明，相继进行两个无穷小转动，和进行的次序无关，即：无穷小转动的合成服从平行四边形法则. 因而，刚体转动的角速度

$$\boldsymbol{\omega}=\frac{\mathrm{d}\boldsymbol{\varphi}}{\mathrm{d}t}=\lim_{\Delta t\to 0}\frac{\Delta\boldsymbol{\varphi}}{\Delta t} \tag{1.1}$$

的叠加满足平行四边形法则. 证明如下：

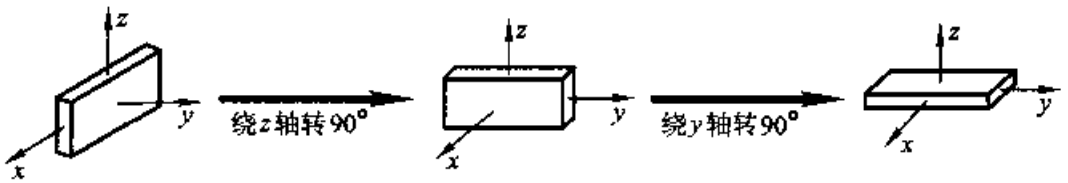
设转角 $\Delta\varphi\to 0$. 沿转动轴线，按右手螺旋方向作一个有向线段 $\Delta\boldsymbol{\varphi}$，其长度等于转动角 $\Delta\varphi$. 将坐标原点放在转动 $\Delta\boldsymbol{\varphi}$ 的转轴上，则在这一转动之下，任意位矢 $\boldsymbol{r}$ 的变化是

$$\Delta\boldsymbol{r}=\Delta\boldsymbol{\varphi}\times\boldsymbol{r}, \tag{1.2}$$

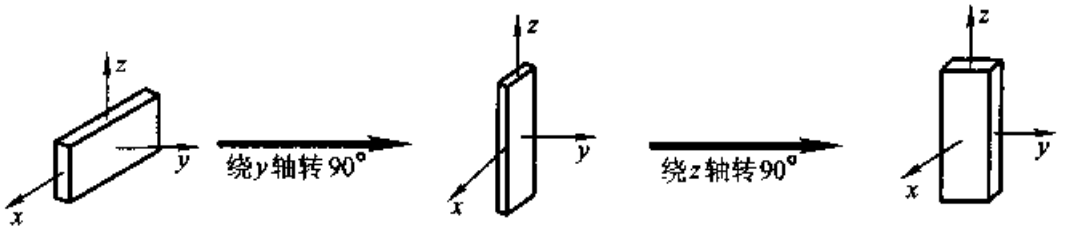
见 §1.2.2(2.8)式. 因此，无穷小转动可以看成是一种空间变换，它把空间的任一

点 r 变到另一点 $r + \Delta\varphi \times r$：

$$r \xrightarrow[\text{转动 } \Delta\varphi]{} r + \Delta\varphi \times r = (1 + \Delta\varphi \times)r. \tag{1.3}$$



(a) 绕y轴转之前先绕z轴转



(b) 绕y轴转之后再绕z轴转

图 2　绕不同轴转动相互关联

上式括号中的 $(1 + \Delta\varphi \times)$ 是一种运算符号，它作用在 r 上的规则和通常的乘法分配律一样，因而可以将 r 放进括号内，得到 $r + \Delta\varphi \times r$. 这一运算符号称为转动变换算符，用 $T_{\Delta\varphi}$ 表示：

$$T_{\Delta\varphi} = 1 + \Delta\varphi \times. \tag{1.4}$$

用 $\Delta\varphi_1$ 和 $\Delta\varphi_2$ 表示绕两个不同轴的转动. 先转 $\Delta\varphi_1$ 后转 $\Delta\varphi_2$ 使 r 变为

$$r \xrightarrow[\text{先转 } \Delta\varphi_1 \quad \text{再转 } \Delta\varphi_2]{} T_{\Delta\varphi_2}T_{\Delta\varphi_1}r = (1 + \Delta\varphi_2 \times)(1 + \Delta\varphi_1 \times)r$$

$$= r + \Delta\varphi_2 \times r + \Delta\varphi_1 \times r + o(\Delta\varphi). \tag{1.5}$$

先转 $\Delta\varphi_2$ 后转 $\Delta\varphi_1$ 使 r 变为

$$r \xrightarrow[\text{先转 } \Delta\varphi_2 \quad \text{再转 } \Delta\varphi_1]{} T_{\Delta\varphi_1}T_{\Delta\varphi_2}r = (1 + \Delta\varphi_1 \times)(1 + \Delta\varphi_2 \times)r$$

$$= r + \Delta\varphi_1 \times r + \Delta\varphi_2 \times r + o(\Delta\varphi). \tag{1.6}$$

由此可见

$$T_{\Delta\varphi_2}T_{\Delta\varphi_1}r = T_{\Delta\varphi_1}T_{\Delta\varphi_2}r + o(\Delta\varphi). \tag{1.7}$$

当 $\Delta\varphi \to 0$ 时，二级小量 $o(\Delta\varphi)$ 可以略去. 因而

$$\text{当 } \Delta\varphi \to 0 \text{ 时} \qquad T_{\Delta\varphi_2}T_{\Delta\varphi_1}r = T_{\Delta\varphi_1}T_{\Delta\varphi_2}r. \tag{1.8}$$

这就是说，两个无穷小转动与次序无关. 由角速度和无穷小转动的关系(1.1)式可见，两个角速度相加服从平行四边形法则. 【证毕】

进一步还可以证明，刚体转动的角速度和转轴的位置无关.

设在某一瞬间 t，刚体中的某一点 o 相对于惯性参考系 K_0 有速度 V_o. 考虑一

个相对于 K_0 以速度 $\boldsymbol{V}_o$ 运动的另一惯性参考系 K. 在时刻 t，o 点相对于 K 系静止，而整个刚体绕 o 点转动. 用 $\boldsymbol{\omega}$ 表示转动的瞬时角速度.

刚体中的任一点 P 相对于 o 点的坐标位矢 $\boldsymbol{r}_{oP}$ 在转动 $\mathrm{d}\boldsymbol{\varphi}$ 下的变化是 $\mathrm{d}\boldsymbol{r}_{oP} = \mathrm{d}\boldsymbol{\varphi} \times \boldsymbol{r}_{oP}$. 将 P 点相对于 K_0 系的位矢 $\boldsymbol{R}_P = \boldsymbol{R}_o + \boldsymbol{r}_{oP}$ 对 t 求导，得到

$$\boldsymbol{V}_P = \boldsymbol{V}_o + \boldsymbol{\omega} \times \boldsymbol{r}_{oP}. \tag{1.9}$$

同理，用刚体中的另一点 o' 代替 o 得到

$$\boldsymbol{V}_P = \boldsymbol{V}_{o'} + \boldsymbol{\omega}' \times \boldsymbol{r}_{o'P}. \tag{1.10}$$

另一方面，也可以将 o' 看成 P 点，按(1.9)式又有

$$\boldsymbol{V}_{o'} = \boldsymbol{V}_o + \boldsymbol{\omega} \times \boldsymbol{r}_{oo'}. \tag{1.11}$$

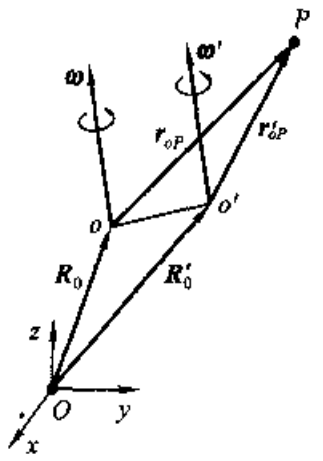


图 3　转动中心的变换

将 $\boldsymbol{r}_{oP} = \boldsymbol{r}_{oo'} + \boldsymbol{r}_{o'P}$ 代入(1.9)式并将(1.11)式代入(1.10)式，就有

$$\boldsymbol{V}_P = \boldsymbol{V}_o + \boldsymbol{\omega} \times \boldsymbol{r}_{oo'} + \boldsymbol{\omega} \times \boldsymbol{r}_{o'P} = \boldsymbol{V}_o + \boldsymbol{\omega} \times \boldsymbol{r}_{oo'} + \boldsymbol{\omega}' \times \boldsymbol{r}_{o'P}.$$

从而得到

$$\boldsymbol{\omega}' = \boldsymbol{\omega}. \tag{1.12}$$

即，刚体的角速度 $\boldsymbol{\omega}$ 与转动中心的选择无关，可以任意平行移动. **【证毕】**

综合以上两点可见，刚体的角速度 $\boldsymbol{\omega}$ 是一个按平行四边形法则叠加，可以任意平行移动的矢量.

最后，让我们来证明一个有用的公式.

(1.4)式中定义的 $T_{\Delta\varphi}$ 是转动变换算符. 它作用在位矢 $\boldsymbol{r}$ 上使 $\boldsymbol{r}$ 转动 $\Delta\boldsymbol{\varphi}$. 往往更关心的不是转动以后的位矢 $T_{\Delta\varphi}\boldsymbol{r}$，而是转动引起的位矢的变化

$$T_{\Delta\varphi}\boldsymbol{r} - \boldsymbol{r} = (T_{\Delta\varphi} - 1)\boldsymbol{r}. \tag{1.13}$$

我们写“转动引起的变化”的算符 $(T_{\Delta\varphi} - 1)$ 为 $R_{\Delta\varphi}$. 由(1.4)式知

$$R_{\Delta\varphi} = (T_{\Delta\varphi} - 1) = \Delta\boldsymbol{\varphi} \times. \tag{1.14}$$

我们来看两个这样的算符相继作用与作用次序的关系.

由(1.5)、(1.6)式容易看到

$$\begin{aligned}(R_{\Delta\varphi_1}R_{\Delta\varphi_2} - R_{\Delta\varphi_2}R_{\Delta\varphi_1})\boldsymbol{r} &= \Delta\boldsymbol{\varphi}_1 \times (\Delta\boldsymbol{\varphi}_2 \times \boldsymbol{r}) - \Delta\boldsymbol{\varphi}_2 \times (\Delta\boldsymbol{\varphi}_1 \times \boldsymbol{r}) \\ &= \Delta\boldsymbol{\varphi}_2(\Delta\boldsymbol{\varphi}_1 \cdot \boldsymbol{r}) - \Delta\boldsymbol{\varphi}_1(\Delta\boldsymbol{\varphi}_2 \cdot \boldsymbol{r}) \\ &= \boldsymbol{r} \times (\Delta\boldsymbol{\varphi}_2 \times \Delta\boldsymbol{\varphi}_1) \\ &= (\Delta\boldsymbol{\varphi}_1 \times \Delta\boldsymbol{\varphi}_2) \times \boldsymbol{r}. \end{aligned} \tag{1.15}$$

再用(1.14)式，可将上式写为

$$R_{\Delta\varphi_1}R_{\Delta\varphi_2} - R_{\Delta\varphi_2}R_{\Delta\varphi_1} = R_{\Delta\varphi_1\times\Delta\varphi_2}. \tag{1.16}$$

这表明，两个无穷小转动引起的变化的算符 $R_{\Delta\varphi_1}$ 和 $R_{\Delta\varphi_2}$ 相继进行的结果和次序

有关，不可对易. 注意，上式左边的每一项都是两个一级小量相乘，是二级小量. 右边也是二级小量. 小量的级次相配，因而右方不能略去.

我们看到，无穷小转动 T 可对易，见(1.8)式；而无穷小转动引起的变化 R 不可对易，见(1.16)式. 为了对其原因有一个形象化的理解，在图4上举了一个例子. 图4a上画出了位矢 r 先绕 x 轴作无穷小转动 $\Delta\varphi_1$ 成为 $T_x r$（简写 $T_{\Delta\varphi_1}$ 为 T_x），然后再绕 y 轴作无穷小转动 $\Delta\varphi_2$ 成为 $T_y T_x r$. 图4b是交换转动次序得到的 $T_x T_y r$. 由图可见，两者近似相等. 同样两个图上也分别画出了转动引起的改变 $(T_x - 1)r = R_x r$ 和 $(T_y - 1)R_x r = R_y R_x r$，以及 $R_y r$ 和 $R_x R_y r$. 由图显然可见，$R_y R_x r$ 和 $R_x R_y r$ 完全不相等.

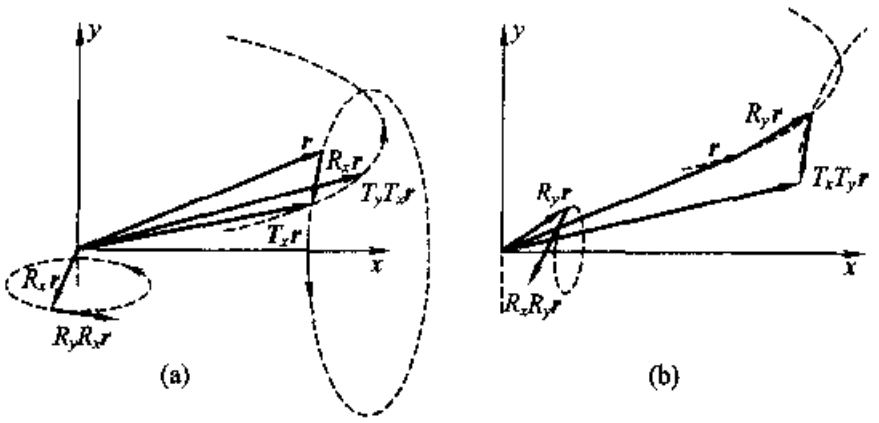


图4　两个转动算符 T_1, T_2 和它们引起的变化的算符 R_1, R_2 相继作用的结果的比较

(二)角动量矢量与转动惯量张量　惯量椭球

在本章引言中指出，刚体的位置由任一点 o 在固定坐标系 K_0 中的位矢和绕 o 点的欧拉角决定. 因此，刚体的运动可以归结为 o 点的运动和绕 o 点的转动. 现在来讨论刚体绕 o 点的转动.

一般情况下，刚体绕 o 点转动的轴不是固定的. 此时，它对 o 点的总角动量是§1.2.2(2.12)式.

$$\boldsymbol{L} = \sum_{a=1}^{N} \boldsymbol{r}_a \times \boldsymbol{p}_a = \sum_{a=1}^{N} m_a \boldsymbol{r}_a \times \boldsymbol{v}_a. \tag{1.17}$$

利用(1.9)式，并注意由于选了 o 点为坐标原点，$\boldsymbol{R}_o = 0$，有 $\boldsymbol{v}_a = \boldsymbol{\omega} \times \boldsymbol{r}_a$. 代入(1.17)式，得到

$$\boldsymbol{L} = \sum_{a=1}^{N} m_a \boldsymbol{r}_a \times (\boldsymbol{\omega} \times \boldsymbol{r}_a) = \sum_{a=1}^{N} m_a [r_a^2 \boldsymbol{\omega} - \boldsymbol{r}_a (\boldsymbol{r}_a \cdot \boldsymbol{\omega})]. \tag{1.18}$$

为了将此式写成类似于§1.2.3(3.17)式的形式，需要将 $\boldsymbol{\omega}$ 提到括号之外. 为此，利用

$$\boldsymbol{\omega} = \sum_{i=1}^{3} \boldsymbol{e}_i \omega_i = \sum_{i=1}^{3} \boldsymbol{e}_i (\boldsymbol{e}_i \cdot \boldsymbol{\omega}) = \sum_{i=1}^{3} [\boldsymbol{e}_i \boldsymbol{e}_i] \cdot \boldsymbol{\omega}. \tag{1.19}$$

右边方括号中的 $\boldsymbol{e}_i\boldsymbol{e}_i$ 是并矢(见附录Ⅱ). 代入(1.18)式,得到

$$\boldsymbol{L} = \sum_{a=1}^{N}\left[(m_a\boldsymbol{r}_a^2)\sum_{i=1}^{3}\boldsymbol{e}_i\boldsymbol{e}_i - m_a\boldsymbol{r}_a\boldsymbol{r}_a\right]\cdot\boldsymbol{\omega}. \tag{1.20}$$

或写为

$$\boldsymbol{L} = \vec{\boldsymbol{I}}\cdot\boldsymbol{\omega}. \tag{1.21}$$

其中,

$$\vec{\boldsymbol{I}} = \sum_{a=1}^{N}\left[(m_a\boldsymbol{r}_a^2)\sum_{i=1}^{3}\boldsymbol{e}_i\boldsymbol{e}_i - m_a\boldsymbol{r}_a\boldsymbol{r}_a\right]. \tag{1.22}$$

是一个并矢,称为转动惯量并矢,或转动惯量张量. 由(1.21)式可见,角动量 $\boldsymbol{L}$ 是转动惯量并矢 $\vec{\boldsymbol{I}}$ 与角速度 $\boldsymbol{\omega}$ 的点积.

(1.21)式和§1.2.3(3.17)式有类似的结构,即:刚体的动量、角动量分别和它的速度、角速度有线性关系. 这一线性关系的系数是反映刚体性质的参量——质量、转动惯量张量.

转动惯量并矢 $\vec{\boldsymbol{I}}$ 有九个分量,它们不难从(1.22)式求出. 例如:

$$\left.\begin{aligned} I_{xx} &= \boldsymbol{e}_1\cdot\vec{\boldsymbol{I}}\cdot\boldsymbol{e}_1 = \sum_{a=1}^{N}m_a(\boldsymbol{r}_a^2 - x_a^2) = \sum_{a=1}^{N}m_a(y_a^2 + z_a^2),\\ I_{xy} &= \boldsymbol{e}_1\cdot\vec{\boldsymbol{I}}\cdot\boldsymbol{e}_2 = \sum_{a=1}^{N}(-m_a x_a y_a) = I_{yx},\quad \cdots \end{aligned}\right\} \tag{1.23}$$

将它们排成矩阵就是:

$$I = \begin{pmatrix} \sum m(y^2+z^2) & -\sum mxy & -\sum mxz \\ -\sum myx & \sum m(z^2+x^2) & -\sum myz \\ -\sum mzx & -\sum mzy & \sum m(x^2+y^2) \end{pmatrix}. \tag{1.24}$$

这里为了书写简单,省去了表示不同质点的下标 a.

实际上,刚体常常可以看成质量连续分布. 用 $\mathrm{d}m$ 表示在体积元 $\mathrm{d}\tau = \mathrm{d}z\mathrm{d}y\mathrm{d}z$ 中的质量,则

$$\mathrm{d}m = \rho(x,y,z)\mathrm{d}\tau, \tag{1.25}$$

其中 $\rho(x,y,z)$ 是密度. 此时,(1.22)、(1.23)式中的求和应改成积分:

$$\vec{\boldsymbol{I}} = \int\left(r^2\sum_{i=1}^{3}\boldsymbol{e}_i\boldsymbol{e}_i - \boldsymbol{r}\boldsymbol{r}\right)\rho\mathrm{d}\tau,\qquad I_{xx} = \int(y^2+z^2)\rho\mathrm{d}\tau,\qquad I_{xy} = -\int xy\rho\mathrm{d}\tau,\quad \cdots. \tag{1.26}$$

通常称 I_{xx}, I_{yy}, I_{zz} 为刚体对三个坐标轴 x,y,z 的主转动惯量,而 I_{xy}, I_{yz}, I_{zx} 为惯量积.

比较§1.2.3(3.17)式和本节(1.21)式可见,刚体的平动和转动之间有如下

对应关系：

	平动		转动
速度	$\boldsymbol{v}$	角速度	$\boldsymbol{\omega}$
动量	$\boldsymbol{p}$	角动量	$\boldsymbol{L}$
质量	M	转动惯量	$\ddot{\boldsymbol{I}}$

可见，和质量对应的是转动惯量，但前者是标量，后者是并矢（矩阵或张量）.

将角速度 $\boldsymbol{\omega}$ 的三个分量排成 3 行 1 列的矩阵，利用转动惯量矩阵(1.24)式可将(1.21)式写成矩阵形式

$$L = I\omega, \qquad L_i = \sum_{j=1}^{3} I_{ij}\,\omega_j, \qquad i = 1,2,3. \tag{1.27}$$

或写成明显的分量形式

$$\left.\begin{aligned} L_x &= I_{xx}\omega_x + I_{xy}\omega_y + I_{xz}\omega_z \\ L_y &= I_{yx}\omega_x + I_{yy}\omega_y + I_{yz}\omega_z \\ L_z &= I_{zx}\omega_x + I_{zy}\omega_y + I_{zz}\omega_z \end{aligned}\right\}. \tag{1.28}$$

由于转动惯量 I 是对称矩阵，见(1.23)、(1.26)式，所以可以直接应用 §1.4.3(二)的结果：

(1)一般情况下(只要 I 不是正比于单位矩阵)，角动量 $\boldsymbol{L}$ 一般不平行于角速度 $\boldsymbol{\omega}$. 这可以从(1.28)式明显看到.

(2)存在三个特殊的方向，称为刚体转动的本征方向. 当角速度沿这三个方向时，角动量和角速度平行.

(3)刚体绕本征方向 $a(a = 1,2,3)$ 转动的转动惯量 I_a 由特征方程[见 §1.4.3(3.20)式]

$$| I_{ij} - I_a\delta_{ij} | = 0 \tag{1.29}$$

决定. 将所得到的 I_a 代入本征值方程[即 §1.4.3 (3.18)或(3.22)式]

$$\sum_{j=1}^{3} I_{ij}\,\omega_j^{(a)} = I_a\omega_i^{(a)} \qquad (i = 1,2,3), \tag{1.30}$$

可解出三个相互垂直的本征方向 $\boldsymbol{\omega}^{(a)}$ ($a = 1,2,3$)，称为惯量主轴.

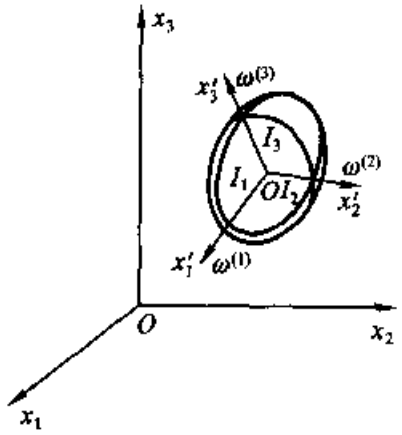


图 5　刚体的惯量椭球

(4)在三个相互垂直的惯量主轴 $\boldsymbol{\omega}^{(a)}$ ($a = 1,2,3$)上，以三个本征值 I_a 为半轴作出的椭球，称为刚体的惯量椭球，参看图 5 及 §1.4.3 图 4. 它就是

转动惯量矩阵的本征椭球.

如果选相互垂直的三个惯量主轴作坐标轴建立坐标系，如图 5 中的 $ox'_1x'_2x'_3$，则在这一坐标系中，所有的惯量积都为零，转动惯量张量具有对角形式：

$$I = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix}. \tag{1.31}$$

在这一坐标系中，(1.27)式成为

$$L_1 = I_1\omega_1, \quad L_2 = I_2\omega_2, \quad L_3 = I_3\omega_3. \tag{1.32}$$

证　设 $\boldsymbol{\omega}_\alpha$ 为 $\vec{\boldsymbol{I}}$ 的本征矢，$\vec{\boldsymbol{I}}\boldsymbol{\omega}_\alpha = I_\alpha\boldsymbol{\omega}_\alpha(\alpha = 1,2,3)$. 在 $\boldsymbol{\omega}_\alpha$ 方向上的单位矢量 $\boldsymbol{e}_{n\alpha} = \boldsymbol{\omega}_\alpha / |\boldsymbol{\omega}_\alpha|$ 满足 $\vec{\boldsymbol{I}}\boldsymbol{e}_{n\alpha} = I_\alpha\boldsymbol{e}_{n\alpha}, \boldsymbol{e}_{n\alpha}\cdot\boldsymbol{e}_{n\beta} = \delta_{\alpha\beta}(\alpha,\beta = 1,2,3)$.

设 $\boldsymbol{e}_i(i = 1,2,3)$ 为所用的坐标系的三个基矢，它们满足正交归一条件 $\boldsymbol{e}_i\cdot\boldsymbol{e}_j = \delta_{ij}$. 在此坐标系中，$\vec{\boldsymbol{I}}$ 的分量 $I_{ij} = \boldsymbol{e}_i\cdot\vec{\boldsymbol{I}}\cdot\boldsymbol{e}_j$ 有 9 个分量，如(1.24)式.

由于 $\boldsymbol{e}_{n\alpha}(\alpha = 1,2,3)$ 也满足正交归一条件 $\boldsymbol{e}_{n\alpha}\cdot\boldsymbol{e}_{n\beta} = \delta_{\alpha\beta}$，所以也可以用来作为一个坐标系的基矢. 在这一坐标系中 $\vec{\boldsymbol{I}}$ 的分量为

$$I_{\alpha\beta} = \boldsymbol{e}_{n\alpha}\cdot\vec{\boldsymbol{I}}\cdot\boldsymbol{e}_{n\beta} = \boldsymbol{e}_{n\alpha}\cdot I_\beta\boldsymbol{e}_{n\beta} = I_\beta\boldsymbol{e}_{n\alpha}\cdot\boldsymbol{e}_{n\beta} = I_\beta\delta_{\alpha\beta}.$$

即：$\vec{\boldsymbol{I}}$ 在以自己的本征矢为基矢的坐标系中只有 3 个对角分量，如(1.31)式.

【证毕】

我们来讨论由以 $\boldsymbol{e}_i$ 为基矢的坐标向以 $\boldsymbol{e}_{n\alpha}$ 为基矢的坐标的坐标变换，

$$\boldsymbol{e}_{n\alpha} = \sum_{i=1}^{3} a_{\alpha i}\boldsymbol{e}_i, \qquad \boldsymbol{e}_i = \sum_{\alpha=1}^{3} a_{i\alpha}^{-1}\boldsymbol{e}_{n\alpha}, \qquad \sum_{i=1}^{3} a_{\alpha i}a_{i\beta}^{-1} = \delta_{\alpha\beta}.$$

其中，$a_{\alpha i}$ 为坐标变换矩阵，$a_{i\alpha}^{-1}$ 是它的逆矩阵. 由于 $\boldsymbol{e}_i(i = 1,2,3)$ 和 $\boldsymbol{e}_{n\alpha}(\alpha = 1,2,3)$ 都是正交归一矢量组，它们之间的变换矩阵 A 是正交矩阵，满足条件 $\sum_{i=1}^{3} a_{\alpha i}a_{\beta i} = \delta_{\alpha\beta}$，证明见习题 4[参看 §1.4.3(3.26)式]. 因而有

$$I_{\alpha\beta} = \boldsymbol{e}_{n\alpha}\cdot\vec{\boldsymbol{I}}\cdot\boldsymbol{e}_{n\beta} = \sum_{ij=1}^{3} a_{\alpha i}\boldsymbol{e}_i\cdot\vec{\boldsymbol{I}}\cdot a_{\beta j}\boldsymbol{e}_j = \sum_{ij=1}^{3} a_{\alpha i}I_{ij}a_{i\beta}^{-1}.$$

这就是 $\vec{\boldsymbol{I}}$ 在两个坐标系之间的变换规律.

求惯量主轴的一般方法是解方程(1.29)、(1.30). 但是在刚体具有对称性的情况下常常可以直接看出惯量主轴. 例如，如果刚体是一个边长为 a,b,c 的质量均匀分布的长方体，则通过长方体中心的惯量主轴方向就是边长 a,b,c 的方向. 如果 $b = c \neq a$，则在半轴 b,c 平面内的任意方向都是惯量主轴，参看 §1.4.3 定理三.

再如，半径可以忽略的细棒，绕棒轴的转动惯量为零，而通过棒的中心垂直于棒的任意方向的转动惯量都相等. 这种细棒称为转子，它的特点是只有两个而不是三个转动自由度，因为当棒半径为零时，绕棒轴的转动无意义.

刚体的角动量一般不平行于角速度的原因可以如下理解. 对于一个不完全对称的刚体，(1.32)式中的三个 I_α 不全相等. 假定 $I_2 \neq I_3$. 当刚体绕(2,3)平面，即 (y,z) 平面上某一轴有角速度 $\boldsymbol{\omega}$ 时，可将它分别投影到 y 轴和 z 轴上，得到角速度分量 ω_2, ω_3. 由此有角动量分量 $L_2 = I_2\omega_2, L_3 = I_3\omega_3$. 由于 $I_2 \neq I_3$，由 L_2 和 L_3 合成的角动量 $\boldsymbol{L}$ 和 $\boldsymbol{\omega}$ 不平行，如图 6.

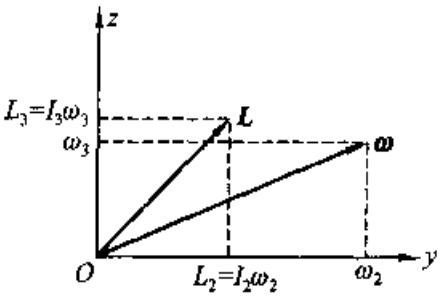


图 6　角动量不平行于角速度

(三) 刚体的能量

通常感兴趣的是刚体的自由运动和在重力场中的运动. 刚体在重力场中的势能等于全部质量集中在质心时的质点的势能. 需要讨论的是刚体的动能 $T = \sum mv^2/2$，将(1.9)式代入得到

$$T = \sum_\alpha \frac{m_\alpha}{2}(\boldsymbol{V} + \boldsymbol{\omega} \times \boldsymbol{r}_\alpha)^2$$

$$= \sum_\alpha \frac{m_\alpha V^2}{2} + \sum_\alpha m_\alpha \boldsymbol{V} \cdot [\boldsymbol{\omega} \times \boldsymbol{r}_\alpha] + \sum_\alpha \frac{m_\alpha}{2}(\boldsymbol{\omega} \times \boldsymbol{r}_\alpha)^2. \qquad (1.33)$$

右边第一项等于 $MV^2/2$，是全部质量集中在质心上时的平动动能. 第二项是

$$\sum m_\alpha \boldsymbol{V} \cdot [\boldsymbol{\omega} \times \boldsymbol{r}_\alpha] = (\sum m_\alpha \boldsymbol{r}_\alpha) \cdot [\boldsymbol{V} \times \boldsymbol{\omega}] = 0,$$

因为我们已经选质心为运动坐标系的原点. 在第三项中展开

$$(\boldsymbol{\omega} \times \boldsymbol{r}_\alpha) \cdot (\boldsymbol{\omega} \times \boldsymbol{r}_\alpha) = \boldsymbol{r}_\alpha \cdot [(\boldsymbol{\omega} \times \boldsymbol{r}_\alpha) \times \boldsymbol{\omega}] = r_\alpha^2\omega^2 - (\boldsymbol{r}_\alpha \cdot \boldsymbol{\omega})^2.$$

代入(1.33)式得到

$$\left.\begin{aligned} T &= T_{\text{平动}} + T_{\text{转动}} \\ T_{\text{平动}} &= MV^2/2 \\ T_{\text{转动}} &= \sum_\alpha (m_\alpha/2)[\omega^2 r_\alpha^2 - (\boldsymbol{\omega} \cdot \boldsymbol{r}_\alpha)^2] \end{aligned}\right\}. \qquad (1.34)$$

注意，(1.33)式右边第二项为零，是由于我们选取了质心作为固连在刚体上的运动坐标系的原点. 如果不是以质心作运动坐标系的原点，则刚体的动能不能按(1.34)式那样写成平动动能与转动动能之和. 也就是说，只有用质心作为运动坐标系的原点，刚体的平动和转动才能相互分离，成为独立自由度. 用其他点作运动坐标系的原点时，不能作这样的分离(参看 §1.2.3 中的讨论).

由(1.22)式，

$$\boldsymbol{\omega}\cdot\vec{\boldsymbol{I}}\cdot\boldsymbol{\omega}=\sum_{a=1}^{N}m_a\Big[r_a^2\sum_{i=1}^{3}\omega_i\omega_i-(\boldsymbol{\omega}\cdot\boldsymbol{r}_a)^2\Big].$$

因而可以写

$$T_{\text{转动}}=\frac{1}{2}\boldsymbol{\omega}\cdot\vec{\boldsymbol{I}}\cdot\boldsymbol{\omega}=\frac{1}{2}\sum_{ij=1}^{3}I_{ij}\omega_i\omega_j. \tag{1.35}$$

如果将平动能量也改写为

$$T_{\text{平动}}=\frac{1}{2}\sum_{i=1}^{3}MV_iV_i=\frac{1}{2}\sum_{ij=1}^{3}M\delta_{ij}V_iV_j, \tag{1.36}$$

则又一次看到,在转动情况下和质量 M 对应的是转动惯量张量 I_{ij}.

(四)刚体的定点运动和定轴运动

在结束本节之前,让我们指出,同一刚体相对于不同点有不同的转动惯量.这可以从转动惯量的定义(1.22)式明显看出.对于一个不受约束的刚体,在不加说明时,都是指相对于其质心的转动惯量.其原因是,不受约束的刚体的运动可以看成是其质量集中在质心的平动加上绕质心的转动.

如果刚体受到约束使某一点固定,例如陀螺的尖端固定在一个小坑中.此时的转动惯量就是绕这一定点的转动惯量.(1.21)(1.22)式仍然适用,只是要以固定点作为位矢 $\boldsymbol{r}_a$ 的起点.绕这一定点的转动惯量矩阵有自己的惯量主轴和相应的本征值,因而也可以作出自己的惯量椭球.

如果刚体受到约束使一个轴固定,如飞轮的轴受到轴承的约束,则刚体的角速度和角动量都沿这一固定轴的方向,两者平行.其比值 I 是一个标量,

$$\boldsymbol{L}=I\boldsymbol{\omega}. \tag{1.37}$$

§1.2.2(2.16)式属于这一情况.其中选固定的转轴为 z 轴,因而

$$L_z=I\omega_z,\qquad L_x=L_y=\omega_x=\omega_y=0,$$

$$I=\sum_{a=1}^{N}m_ar_a^2\sin^2\theta_a,\qquad \omega_z=\dot{\varphi}.$$

(1.37)式和平动情况下的§1.2.3(3.17)式相对应.

§1.5.2 刚体的运动方程

刚体有六个自由度,但如果对它加上进一步的约束,自由度数目还会减少.例如,令刚体内的一点固定,则只有绕这一点转动的三个自由度;令刚体中的一根轴固定,则只剩下绕此轴转动的一个自由度,又如靠在墙上的梯子有实际意义的是向下滑动的一个自由度.在这种有约束的情况下,我们所关心的常常不仅是刚体本身的运动,而且也包括约束反力.例如,常常需要知道轴承作用在高速转子上的力,地

面作用在梯子上的摩擦力，等等. 用拉格朗日方程不容易得到约束反力，所以我们回到牛顿表述方式，将刚体看成 N 个质点的质点组来建立运动方程.

(一)动量定理与角动量定理

在 §1.2.3 中已经得到了质点组的动量定理. 它当然也适用于刚体这种特殊的质点组. 由 §1.2.3 的(3.14)式和(3.17)式，

$$M\ddot{\boldsymbol{R}}_0 = \sum \boldsymbol{F}_{a外} = \boldsymbol{F}, \tag{2.1}$$

其中，M 是刚体的总质量，$\boldsymbol{R}_0$ 是质心在静止参考系中的位矢，$\boldsymbol{F}$ 是作用在刚体上的合外力.

再将 §1.2.3 的(3.10)式改写为

$$m_a\ddot{\boldsymbol{R}}_a = \boldsymbol{F}_{a内} + \boldsymbol{F}_{a外}, \tag{2.2}$$

其中，$\boldsymbol{R}_a$ 是第 a 个质点在静止坐标系中的位矢，

$$\boldsymbol{R}_a = \boldsymbol{R}_0 + \boldsymbol{r}_a, \tag{2.3}$$

$\boldsymbol{r}_a$ 是质点 a 在以质心为原点的运动坐标系中的位矢.

用 $\boldsymbol{R}_a$ 叉乘(2.2)式，再对 a 求和. 注意

$$\boldsymbol{R}_a \times \ddot{\boldsymbol{R}}_a = \frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{R}_a \times \dot{\boldsymbol{R}}_a),$$

而由 §1.2.3(3.2)式知道，质点 b 对质点 a 的作用力 $\boldsymbol{F}_{ab}$ 和联接 ab 的矢量 $\boldsymbol{R}_a - \boldsymbol{R}_b$ 平行，因而 $(\boldsymbol{R}_a - \boldsymbol{R}_b) \times \boldsymbol{F}_{ab} = 0$. 由此易证，当 $\boldsymbol{R}_a \times \boldsymbol{F}_{a内}$ 对 a 求和时，最后结果为零

$$\sum_{a=1}^{N} \boldsymbol{R}_a \times \boldsymbol{F}_{a内} = 0. \tag{2.4}$$

这样就得到

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{a=1}^{N}(\boldsymbol{R}_a \times m_a\dot{\boldsymbol{R}}_a) = \sum_{a=1}^{N}\boldsymbol{R}_a \times \boldsymbol{F}_{a外}.$$

此式左边的和式是质点组对固定点 O 的总角动量 $\boldsymbol{L}$，而右边是对同一点的总力矩 $\boldsymbol{M}$，因而

$$\frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t} = \boldsymbol{M}. \tag{2.5}$$

但是，相对于质心的转动惯量表达式比较简单，因此，以质心为中心来讨论刚体的转动往往更方便. 为此，将(2.3)式代入(2.2)式，得到

$$m_a\ddot{\boldsymbol{R}}_0 + m_a\ddot{\boldsymbol{r}}_a = \boldsymbol{F}_{a内} + \boldsymbol{F}_{a外},$$

用 $\boldsymbol{r}_a$ 叉乘上式再对 a 求和. 注意到按照质心的定义，$\sum m_a\boldsymbol{r}_a = 0$，就类似于(2.4)式和(2.5)式得到

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{a=1}^{N}(\boldsymbol{r}_a \times m_a\dot{\boldsymbol{r}}_a) = \sum_{a=1}^{N}\boldsymbol{r}_a \times \boldsymbol{F}_{a外},$$

即

$$\frac{\mathrm{d}\boldsymbol{L}^{(o)}}{\mathrm{d}t} = \boldsymbol{M}^{(o)}, \tag{2.6}$$

其中，上标 (o) 表示是相对于质心的角动量和力矩.

(2.5)式和(2.6)式分别是对固定点和对质心的角动量定理.(注意，除对质心以外，对其他随刚体运动的点，类似的式子不成立).它们都是三个(分量)方程，加上动量定理(2.1)式的三个方程，总共六个方程构成刚体的运动方程组，它们完全决定了刚体的六个自由度的运动.

(二)刚体的静平衡

在特殊情况下，如果刚体在外力作用下处于平衡，则由(2.1)和(2.5)式，有

$$\boldsymbol{F} = 0, \qquad \boldsymbol{M} = 0. \tag{2.7}$$

这是刚体的平衡方程.其中，$\boldsymbol{F}$ 是作用在刚体上的外力的矢量和，$\boldsymbol{M}$ 是对任意固定点的力矩的矢量和.

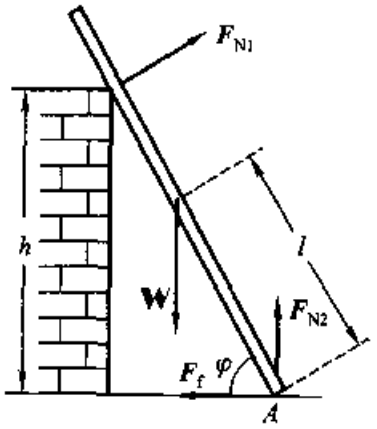


图 1　梯子的平衡

【例 1】　重为 W，长为 $2l$ 的均匀梯子，靠在高 h 的墙上.墙对梯子的支持是光滑的，而地面粗糙，梯子与地而的摩擦系数为 μ，求能保持平衡的最小角度 φ_0.

解　作用在梯子上的力如图 1，由 x,y 两方向上的合力为零得

$$F_{N_1}\cos(\pi/2-\varphi_0) - F_f = 0, \qquad F_{N_1}\sin(\pi/2-\varphi_0) + F_{N_2} - W = 0;$$

由对 A 点的力矩为零得：

$$Wl\cos\varphi_0 - F_{N_1}h/\sin\varphi_0 = 0.$$

联立求解，得到

$$F_f = Wl\sin\varphi_0\sin 2\varphi_0/2h, \qquad F_{N_2} = W(1 - l\cos\varphi_0\sin 2\varphi_0/2h).$$

F_f 最大只能等于 μF_{N_2}，因而得

$$l\sin\varphi_0\sin 2\varphi_0 = \mu(2h - l\cos\varphi_0\sin 2\varphi_0).$$

这就是决定 φ_0 的方程.

(三)刚体的动平衡

刚体(例如飞轮)绕固定轴转动时会受到轴承的约束反力，反过来刚体对轴承施加压力.现代机器的转动部分转速很高，对轴承的压力会产生破坏性作用，因此，减少或消除这种压力是很重要的.如果绕固定轴转动的刚体，对轴承不产生附加压

力，我们就说刚体处于动平衡.

【例 2】　刚体在相距 l 的两个轴承 A、B 的约束下转动，求动平衡条件.

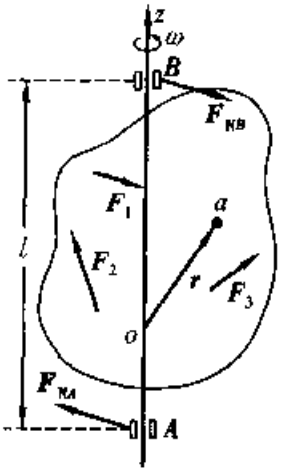


图 2　转动刚体对轴承的附加压力

解　选旋转轴 $\overline{AB}$ 为 z 轴，如图 2，o 为 z 轴上任一点. $\boldsymbol{F}_1, \boldsymbol{F}_2, \cdots, \boldsymbol{F}_n$ 为作用在刚体上的外力，合外力为 $\boldsymbol{F} = \sum \boldsymbol{F}_i$. 轴承对刚体的约束反力是 $\boldsymbol{F}_{NA}$ 和 $\boldsymbol{F}_{NB}$，在 $\boldsymbol{F} + \boldsymbol{F}_{NA} + \boldsymbol{F}_{NB}$ 共同作用下，刚体只绕轴转动，而轴静止不动，因而位于轴上的 o 点静止. 刚体中不在转轴上的任意点 a 有速度 $\boldsymbol{v}_a$. 按 §1.5.1 的(1.9)式(注意到 $\boldsymbol{\omega}$ 沿 z 方向)，

$$\boldsymbol{v}_a = \boldsymbol{\omega} \times \boldsymbol{r}_a \quad v_{ax} = -\omega y_a, \quad v_{ay} = \omega x_a. \tag{2.8}$$

对于现在讨论的定轴转动问题，将运动分为质心的平动和绕质心的转动并不方便. 因此，直接从 §1.2.3(3.11)(3.13)式和相对于固定点的角动量定理[本节(2.5)式]出发. 由 §1.2.3(2.11)(3.13)式有

$$\sum m_a \dot{\boldsymbol{v}}_a = \boldsymbol{F}_{NA} + \boldsymbol{F}_{NB} + \boldsymbol{F}. \tag{2.9}$$

利用(2.8)式

$$\begin{aligned}\dot{\boldsymbol{v}}_a &= \boldsymbol{\omega} \times \dot{\boldsymbol{r}}_a + \dot{\boldsymbol{\omega}} \times \boldsymbol{r}_a = \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}_a) + \dot{\boldsymbol{\omega}} \times \boldsymbol{r}_a \\ &= \boldsymbol{\omega}(\boldsymbol{\omega} \cdot \boldsymbol{r}_a) - \omega^2 \boldsymbol{r}_a + \dot{\boldsymbol{\omega}} \times \boldsymbol{r}_a\end{aligned}$$

代入(2.9)式得到

$$M[\boldsymbol{\omega}(\boldsymbol{\omega} \cdot \boldsymbol{r}_C) - (\omega^2)\boldsymbol{r}_C + \dot{\boldsymbol{\omega}} \times \boldsymbol{r}_C] = \boldsymbol{F}_{NA} + \boldsymbol{F}_{NB} + \boldsymbol{F}. \tag{2.10}$$

其中 $\boldsymbol{r}_C = \sum m_a \boldsymbol{r}_a / M$ 是刚体质心相对于 o 点的位矢，见 §1.2.3(3.19)式.

由固定点的角动量定理(2.5)式

$$\mathrm{d}\boldsymbol{L}/\mathrm{d}t = \boldsymbol{M}_N + \boldsymbol{M}_F, \tag{2.11}$$

其中，$\boldsymbol{M}_N$ 和 $\boldsymbol{M}_F$ 分别是约束反力 $\boldsymbol{N}_A, \boldsymbol{N}_B$ 和外力 $\boldsymbol{F}$ 所产生的力矩，(2.10)式和(2.11)式就是所考虑问题的运动方程组.

(2.10)式的分量形式是

$$\begin{cases} -M\omega^2 x_C - M\dot{\omega} y_C = F_{NAx} + F_{NBx} + F_x, & (2.12) \\ -M\omega^2 y_C + M\dot{\omega} x_C = F_{NAy} + F_{NBy} + F_y, & (2.13) \\ 0 = F_{NA_z} + F_{NB_z} + F_z. & (2.14) \end{cases}$$

为了将(2.11)式写成分量形式，利用 §1.5.1(1.21)式和(1.24)式，并注意到 $\omega_x = \omega_y = 0, \dot{z}_a = 0$，得到

$$\left.\begin{aligned} M_{Nx}+M_{Fx} &= I_{xz}\dot{\omega}+\dot{I}_{xz}\omega = I_{xz}\dot{\omega}-\sum m_a\dot{x}_a z_a\omega,\\ M_{Ny}+M_{Fy} &= I_{yz}\dot{\omega}+\dot{I}_{yz}\omega = I_{yz}\dot{\omega}-\sum m_a\dot{y}_a z_a\omega,\\ M_{Fz} &= I_{zz}\dot{\omega}. \end{aligned}\right\} \tag{2.15}$$

在写第三式时已注意到 $\boldsymbol{F}_{NA}$ 和 $\boldsymbol{F}_{NB}$ 的作用点在 z 轴上，故 $M_{Nz}=0$．再将(2.8)式代入前两式，利用 §1.5.1(1.23)式，得到

$$I_{xz}\dot{\omega}-I_{yz}\omega^2 = M_{Nx}+M_{Fx}, \tag{2.16}$$

$$I_{yz}\dot{\omega}+I_{xz}\omega^2 = M_{Ny}+M_{Fy}, \tag{2.17}$$

$$I_{zz}\dot{\omega} = M_{Fz}. \tag{2.18}$$

(2.12)～(2.14)，(2.16)～(2.18)等六个方程就是运动方程(2.10)式和(2.11)式的分量形式．

(2.18)式是定轴转动的运动方程，(2.14)式是沿 z 轴方向的平衡方程．它们都是我们所熟悉的．现在有兴趣的是(2.12)，(2.13)，(2.16)，(2.17)四个方程．按照动平衡条件要求，在外力为零时，对轴承的压力为零．它归结为：

$$\left.\begin{aligned} \omega^2 x_C+\dot{\omega}y_C &= 0, \qquad \omega^2 y_C-\dot{\omega}x_C = 0,\\ \omega^2 I_{yz}-\dot{\omega}I_{xz} &= 0, \qquad \omega^2 I_{xz}+\dot{\omega}I_{yz} = 0. \end{aligned}\right\}$$

这是两个变量 ω 和 $\dot{\omega}$ 的四个方程．要它们有非零解，只有

$$x_C = y_C = 0, \qquad I_{yz} = I_{xz} = 0. \tag{2.19}$$

这就是说，只有当转轴过质心，方向沿惯量主轴时，才能达到动平衡．

(四)刚体的自由运动

现在来考虑不受外力作用也不受任何约束的刚体的自由运动．在运动方程(2.1)式和(2.6)式中，令 $\boldsymbol{F}=0$，$\boldsymbol{M}^{(o)}=0$，得到

$$\ddot{\boldsymbol{R}}_0 = 0, \qquad \dot{\boldsymbol{L}}^{(o)} = 0. \tag{2.20}$$

前一式说明，质心以匀速度运动．选用质心系就消去了这一运动，只剩下绕质心的转动．由后一式知，这一转动的角动量 $\boldsymbol{L}$ 守恒(省去上标 o)．

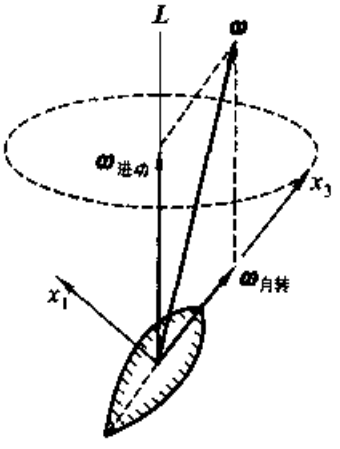


图 3　对称陀螺的规则进动

用 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3$ 表示三个惯量主轴方向，I_1,I_2,I_3 表示沿这三个主轴的转动惯量．

首先考虑 $I_1=I_2=I_3=I$ 的情况，这称为球对称陀螺．此时有 $\boldsymbol{L}=I\boldsymbol{\omega}$．这说明角速度 $\boldsymbol{\omega}$ 守恒——球对称陀螺的自由运动是绕空间任一固定轴以匀角速转动．

沿某一个主轴的转动惯量为零的情况(称为转子)也与此类似．此时 $I_1=I_2=$

$I_1, I_3 = 0$，角动量 $\boldsymbol{L} = I\boldsymbol{\omega}$ 垂直于 x_3，在 x_1x_2 平面内. 自由转子绕 x_1x_2 平面内任一固定轴匀速旋转.

$I_1 = I_2 \neq I_3$，称为对称陀螺(以区别于 $I_1 \neq I_2 \neq I_3$ 的不对称陀螺). 我们考虑一般情况，假定 $\boldsymbol{L}$ 不在主轴方向上. 由于 x_1x_2 平面内的任一轴都是主轴，所以在任一瞬时都可以选 x_2 垂直于 $\boldsymbol{L}$. 这样 $L_2 = 0$. 面由 §1.5.1(1.32)式就有 $\omega_2 = 0$. 因此在这一瞬时，$\boldsymbol{L}, \boldsymbol{\omega}, x_3$ 在同一平面上. 但是这一瞬时是任意选取的. 所以在刚体转动过程中，$\boldsymbol{L}, \boldsymbol{\omega}, x_3$ 始终在同一平面上，如图 3. 将 $\boldsymbol{\omega}$ 分解到 x_3 和 $\boldsymbol{L}$ 的方向上，分别称为 $\boldsymbol{\omega}_{自转}$ 和 $\boldsymbol{\omega}_{进动}$. 前者是刚体绕 x_3 轴转动的角速度，而后者是 x_3 轴绕 $\boldsymbol{L}$ 进动的角速度. 用 θ 表示 $\boldsymbol{L}$ 和 x_3 的夹角，则

$$\omega_{进动}\sin\theta = \omega_1 = L_1/I_1 = L\sin\theta/I_1.$$

因此，

$$\omega_{进动} = L/I_1. \tag{2.21}$$

由此可见，对称陀螺自由转动的运动情况是：一方面绕 x_3 轴自转，另一方面 x_3 轴绕空间一固定轴($\boldsymbol{L}$ 轴)进动. 在此过程中，x_3 轴与 $\boldsymbol{L}$ 之间的夹角 θ 保持不变. 这称为规则进动.

$I_1 \neq I_2 \neq I_3$ 的不对称陀螺运动情况十分复杂，这里不再讨论.

*(五)刚体定点运动的运动学方程

假定在运动过程中刚体的某一点(不一定是质心)固定不动. 我们来讨论它的运动情况. 由于可能受有外力，所以一般没有角动量守恒. 我们不求解动力学方程，而只限于讨论定点运动的运动学.

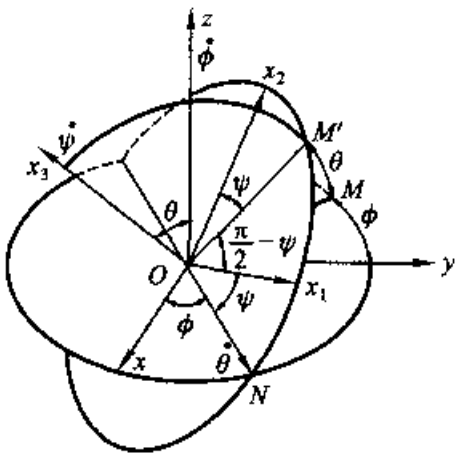


图 4　欧拉角

在有一点固定的情况下，刚体只有三个自由度，最方便的是用三个欧拉角来表示，如本章引言图 1. 按定义，欧拉角 φ 是刚体绕固定轴 oz 转过的角度，因而称为进动角. ψ 是绕轴 ox_3 转过的角度，称为自转角；θ 是 ox_3 和 oz 的夹角，叫章动角. 相应地，$\dot{\varphi}, \dot{\psi}, \dot{\theta}$ 就是进动角速度，自转角速度和章动角速度. 它们分别指向 oz, ox_3, oN 方向，如图 4. 我们来把它们投影到三个惯量主轴 x_1，x_2, x_3 的方向上来.

由图 4 可见，oN 在 x_1x_2 平面上，和 ox_1 夹角 ψ，因而 $\dot{\boldsymbol{\theta}}$ 在 ox_1, ox_2 上的分量分别是 $\dot{\theta}\cos\psi$ 和 $-\dot{\theta}\sin\psi$. 另一方面，想象以 $ox_1x_2x_3$ 为基础作球坐标系，则 oZ 在这一坐标系中的方位角是 θ 和 $\pi/2-\psi$，因而 $\dot{\boldsymbol{\varphi}}$ 在 ox_1, ox_2, ox_3 方向的分量分别是 $(\dot{\varphi}\sin\theta\sin\psi, \dot{\varphi}\sin\theta\cos\psi, \dot{\varphi}\cos\theta)$. $\dot{\boldsymbol{\psi}}$ 本身在 ox_3 方向上. 因此，ox_1, ox_2, ox_3 方向上的角速度是

$$\left.\begin{aligned}\omega_1 &= \dot{\varphi}\sin\theta\sin\psi + \dot{\theta}\cos\psi,\\ \omega_2 &= \dot{\varphi}\sin\theta\cos\psi - \dot{\theta}\sin\psi,\\ \omega_3 &= \dot{\varphi}\cos\theta + \dot{\psi}.\end{aligned}\right\} \tag{2.22}$$

这称为欧拉运动学方程.

如果已知 $\omega_1,\omega_2,\omega_3$，由欧拉运动学方程可以解出 $\dot{\psi},\dot{\varphi},\dot{\theta}$，从而得到刚体的自转，进动和章动情况. 对于对称陀螺，这一方程可以大大简化. 在此情况下，两个主轴可以在 x_1x_2 平面上任意选，因此可以让 x_1 轴在所讨论的时刻沿 oN 方向，使 $\psi = 0$，于是(2.22)式化为

$$\omega_1 = \dot{\theta}, \quad \omega_2 = \dot{\varphi}\sin\theta, \quad \omega_3 = \dot{\varphi}\cos\theta + \dot{\psi}. \tag{2.23}$$

§1.5.3 非惯性系中的运动

如果将参考系固连在一个刚体上，则只要刚体不是相对于惯性系静止或作匀速平动，这一参考系就是非惯性系. 在这一节里，我们讨论在非惯性系中质点的运动.

用 K_0 表示某一惯性参考系. 设 K′系相对于 K_0 系以速度 $\boldsymbol{V}(t)$ 平动，而 K 系相对于 K′系以角速度 $\boldsymbol{\omega}(t)$ 转动，如图 1，则 K 系是一个既有平动加速度，设 $\dot{\boldsymbol{V}} \neq 0$，又有转动，设 $\boldsymbol{\omega} \neq 0$，的非惯性系.

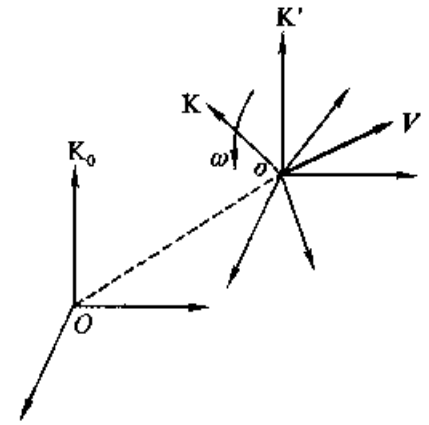


图 I K_0，K′，K 三个参考系的关系

在着手建立 K 系中的运动方程之前，先证明一个引理：

引理 设 K 系相对于 K′系以角速度 $\boldsymbol{\omega}$ 转动，则任意矢量 $\boldsymbol{A}$ 在 K′系中的时间导数 $\mathrm{d}'\boldsymbol{A}/\mathrm{d}t$ 和它在 K 系中的时间导数 $\mathrm{d}\boldsymbol{A}/\mathrm{d}t$ 之间有关系：

$$\mathrm{d}'\boldsymbol{A}/\mathrm{d}t = \mathrm{d}\boldsymbol{A}/\mathrm{d}t + \boldsymbol{\omega} \times \boldsymbol{A}. \tag{3.1}$$

证 由于 K 系转动，所以它的坐标基矢随时间变化. 用 $\boldsymbol{e}_i(t)$ 和 $\boldsymbol{e}'_i(i = 1,2,3)$ 分别表示 K 系和 K′系的坐标基矢. 矢量 $\boldsymbol{A}(t)$ 用它们展开为

$$\boldsymbol{A}(t) = \sum_{i=1}^{3} \boldsymbol{e}_i(t)A_i(t) = \sum_{i=1}^{3} \boldsymbol{e}'_iA'_i(t). \tag{3.2}$$

对 t 求导得到

$$\sum_{i=1}^{3} \boldsymbol{e}_i(t)\,\mathrm{d}A_i/\mathrm{d}t + \sum_{i=1}^{3} A_i\mathrm{d}\boldsymbol{e}_i(t)/\mathrm{d}t = \sum_{i=1}^{3} \boldsymbol{e}'_i\mathrm{d}A'_i(t)/\mathrm{d}t. \tag{3.3}$$

$\boldsymbol{e}_i(t)$ 是单位矢量，长度不变. 它随 t 变化是由于 K 系以角速度 $\boldsymbol{\omega}(t)$ 转动. 按照 §1.2.2(2.17)式，$\mathrm{d}\boldsymbol{e}_i(t)/\mathrm{d}t = \boldsymbol{\omega} \times \boldsymbol{e}_i(t)$. 因而左边第二项是

$$\sum_{i=1}^{3} A_i\mathrm{d}\boldsymbol{e}_i(t)/\mathrm{d}t = \sum_{i=1}^{3} \boldsymbol{\omega} \times \boldsymbol{e}_i(t)A_i(t) = \boldsymbol{\omega} \times \boldsymbol{A}. \tag{3.4}$$

另一方面，(3.3)式左边第一项和右边分别为

$$\sum_{i=1}^{3} \boldsymbol{e}_i(t)\,\mathrm{d}A_i/\mathrm{d}t = \mathrm{d}\boldsymbol{A}/\mathrm{d}t, \qquad \sum_{i=1}^{3} \boldsymbol{e}'_i\,\mathrm{d}A'_i(t)/\mathrm{d}t = \mathrm{d}\boldsymbol{A}'/\mathrm{d}t. \tag{3.5}$$

它们分别是矢量 $\boldsymbol{A}$ 在 K 系和 K′系中的时间导数. 将(3.4)、(3.5)式代入(3.3)式，就得到(3.1)式.　【证毕】

在应用(3.1)式的时候要特别注意，矢量 $\boldsymbol{A}$ 本身并不依赖于所用的坐标系是 K 系还是 K′系，如(3.2)式. 只有矢量 $\boldsymbol{A}$ 的时间导数才依赖于所用的坐标系是相互之间有角速度 $\boldsymbol{\omega}$ 的两个坐标系 K 和 K′之中的哪一个，如(3.1)式.

用 $\boldsymbol{v}_0$，$\boldsymbol{v}'$ 和 $\boldsymbol{v}$ 分别表示质点在 K_0，K′和 K 系中的速度，则根据上述引理和 §1.1.4(4.5)式有

$$\boldsymbol{v}_0 = \boldsymbol{v}' + \boldsymbol{V}, \qquad \boldsymbol{v}' = \mathrm{d}'\boldsymbol{r}/\mathrm{d}t = \mathrm{d}\boldsymbol{r}/\mathrm{d}t + \boldsymbol{\omega} \times \boldsymbol{r},$$

$\mathrm{d}\boldsymbol{r}/\mathrm{d}t$ 是在 K 系中观察的质点速度，等于 $\boldsymbol{v}$，因而

$$\boldsymbol{v}_0 = \boldsymbol{V} + \boldsymbol{v} + \boldsymbol{\omega} \times \boldsymbol{r}. \tag{3.6}$$

从 K_0 系(惯性系)中看，运动方程是

$$m\frac{\mathrm{d}\boldsymbol{v}_0}{\mathrm{d}t} = \boldsymbol{f}, \tag{3.7}$$

其中 $\boldsymbol{f}$ 是质点所受到的力. 但是，在将(3.6)式代入(3.7)式求导时，要注意各项的时间导数的含义.

(3.6)式右边第一项 $\boldsymbol{V}$ 是 K 系和 K′系的共同坐标原点 o 在 K_0 系中的速度. 它(在 K_0 系中)的时间导数即 o 点在 K_0 系中的加速度

$$\boldsymbol{W} = \dot{\boldsymbol{V}}. \tag{3.8}$$

(3.6)式右边第二项 $\boldsymbol{v}$ 是质点在 K 系中的速度，它在 K 系和 K′系中的时间导数是不同的. 现在我们要的是在 K_0 系中的时间导数，由于 K′系相对于 K_0 系只有平动，所以在 K_0 系中的时间导数和在 K′系中的时间导数是相同的. 对于(3.6)式右边第三项也同样如此. 因而(3.6)式的时间导数可写为

$$\frac{\mathrm{d}\boldsymbol{v}_0}{\mathrm{d}t} = \frac{\mathrm{d}\boldsymbol{V}}{\mathrm{d}t} + \frac{\mathrm{d}'\boldsymbol{v}}{\mathrm{d}t} + \frac{\mathrm{d}'(\boldsymbol{\omega} \times \boldsymbol{r})}{\mathrm{d}t}. \tag{3.9}$$

利用(3.1)式得到上式右边第二项

$$\mathrm{d}'\boldsymbol{v}/\mathrm{d}t = \mathrm{d}\boldsymbol{v}/\mathrm{d}t + \boldsymbol{\omega} \times \boldsymbol{v} = \boldsymbol{a} + \boldsymbol{\omega} \times \boldsymbol{v}, \tag{3.10}$$

其中 $\boldsymbol{a} = \mathrm{d}\boldsymbol{v}/\mathrm{d}t$ 是在 K 系中看质点的加速度. (3.9)式右边第三项

$$\begin{aligned}\mathrm{d}'(\boldsymbol{\omega} \times \boldsymbol{r})/\mathrm{d}t &= \mathrm{d}'\boldsymbol{\omega}/\mathrm{d}t \times \boldsymbol{r} + \boldsymbol{\omega} \times (\mathrm{d}'\boldsymbol{r}/\mathrm{d}t)\\ &= (\mathrm{d}\boldsymbol{\omega}/\mathrm{d}t) \times \boldsymbol{r} + \boldsymbol{\omega} \times (\mathrm{d}\boldsymbol{r}/\mathrm{d}t + \boldsymbol{\omega} \times \boldsymbol{r}).\end{aligned}$$

这里头一行是乘积求导公式，第二行用了引理(3.1)式，并注意了 $\boldsymbol{\omega} \times \boldsymbol{\omega} = 0$. 写 $\mathrm{d}\boldsymbol{\omega}/\mathrm{d}t = \dot{\boldsymbol{\omega}}$，$\mathrm{d}\boldsymbol{r}/\mathrm{d}t = \boldsymbol{v}$，就有

$$\mathrm{d}'(\boldsymbol{\omega} \times \boldsymbol{r})/\mathrm{d}t = \dot{\boldsymbol{\omega}} \times \boldsymbol{r} + \boldsymbol{\omega} \times \boldsymbol{v} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}). \tag{3.11}$$

将(3.8)、(3.10)、(3.11)式代入(3.9)式得到

$$d\boldsymbol{v}_0/dt = \boldsymbol{W} + \boldsymbol{a} + 2\boldsymbol{\omega} \times \boldsymbol{v} + \dot{\boldsymbol{\omega}} \times \boldsymbol{r} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}). \tag{3.12}$$

代入(3.7)式并移项,得到

$$m\boldsymbol{a} = \boldsymbol{f} - m\boldsymbol{W} + m\boldsymbol{r} \times \dot{\boldsymbol{\omega}} + 2m\boldsymbol{v} \times \boldsymbol{\omega} + m\boldsymbol{\omega} \times (\boldsymbol{r} \times \boldsymbol{\omega}). \tag{3.13}$$

这就是在非惯性系 K 中的运动方程.

方程(3.13)的左边是质量乘上非惯性系中观察到的加速度 $\boldsymbol{a}$,所以方程右边是在非惯性系中观察到的质点所受到的各种力.第一项 $\boldsymbol{f}$ 是外力;第二项 $-m\boldsymbol{W}$ 是由于非惯性系的平动加速度产生的惯性力;第三项 $m\boldsymbol{r} \times \dot{\boldsymbol{\omega}}$ 是由角加速度产生的惯性力;第四项 $2m\boldsymbol{v} \times \boldsymbol{\omega}$ 线性地依赖于速度 $\boldsymbol{v}$,称为科里奥利力;第五项 $m\boldsymbol{\omega} \times (\boldsymbol{r} \times \boldsymbol{\omega})$ 线性地依赖于位矢 $\boldsymbol{r}$,称为惯性离心力.

习　题

1.决定下列物体相对于质心的惯量主轴及绕主轴转动的转动惯量.设物体总质量为 m,且分布均匀.

(a)长为 l 的细棒;(b)半径为 R 的球;(c)半径为 R 高为 h 的圆柱;(d)边长为 a,b,c 的长方体.

用计算机分别画出这四个物体及它们的惯量椭球.

2.三个质量为 m 的质点放置在 $(a,0,0)$,$(0,a,2a)$ 和 $(0,2a,a)$,相互之间用质量可忽略的轻杆刚性连接,求相对于坐标原点的惯量主轴和绕主轴转动的转动惯量.用计算机画出这个三质点系统及这一系统绕坐标原点的惯量椭球.

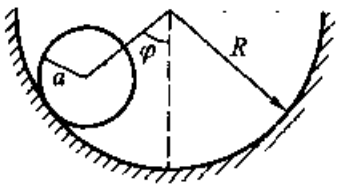


图 1　习题 5 图

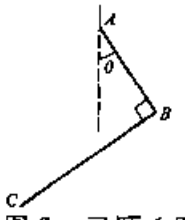


图 2　习题 6 图

3.求习题 2 中的三质点系统的质心.求这一系统相对于质心的惯量主轴,和绕各个主轴的转动惯量.用计算机画出相应的惯量椭球.

4.将正交归一矢量组 $\boldsymbol{e}_i(i = 1,2,3)$ 变换到另一正交归一矢量组 $\boldsymbol{e}_\alpha(\alpha = 1,2,3)$ 的矩阵 A,

$$\boldsymbol{e}_\alpha = \sum_{i=1}^{3} a_{\alpha i}\boldsymbol{e}_i \qquad (\alpha = 1,2,3)$$

称为正交矩阵.试证明正交矩阵的矩阵元满足

$$\sum_{i=1}^{3} a_{\alpha i}a_{\beta i} = \delta_{\alpha\beta}$$

5.半径为 a 的均匀圆柱沿半径为 R 的圆柱形表面滚动($R > 2a$)如图 1,求其动能.

6.两根长为 a 和 b 的均质棒 AB,BC 在 B 点形成直角刚性连接,A 点用绳悬挂,如图 2.求证在平衡时,AB 棒和竖直线的夹角 θ 满足 $\tan\theta$

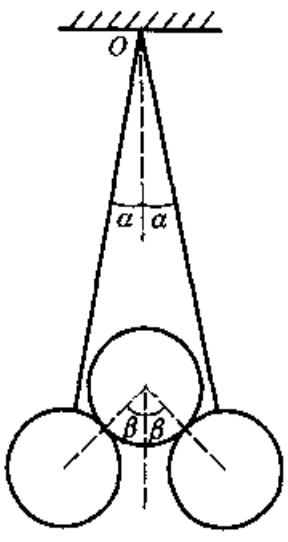


图 3　习题 7 图

$= b^2/(a^2 + 2ab)$.

7. 两个相同的均质光滑球用轻杆悬在定点 O 上，此两球又支持一个等重的均质球，如图 3，求角 α 与 β 的关系.

8. 一个均质圆盘，其质量为 20kg，半径 0.2m，由距离 1m 的两轴承支持，重心在转轴的中点，但转轴和盘面法线偏离 1°，设盘以 12000 转/min 匀角速度转动，求轴上的压力（忽略重力）.

9. 证明 §1.5.2(2.4)式 $\sum_{a=1}^{N} \boldsymbol{R}_a \times \boldsymbol{F}_{a内} = 0$.

10. 为什么自行车骑得快就不倒，骑慢了或停下来就会倒？手不握把手骑自行车时，想要右转弯该怎么办？

提示　用一根直线段联接两个长方形，代表自行车和前后轮. 画出自行车前进时车轮转动的角动量矢量，以及人的身体向右（或向左）偏时，人的重心产生的力矩在 Δt 时间内引起的角动量增量，讨论由角动量的变化导致的自行车偏转.

11. 荡秋千的人在经过竖直位置时站起来，而在经过最高点时蹲下去，使秋千越荡越高，说明其原因. 在竖直位置上站起来时，角速度突然变大，说明其原因.

12. 抛物线形金属丝，以匀角速 ω 绕竖直轴转动，如图 4. 丝上套一质量为 m 的小环，无摩擦地沿金属丝滑动，已知抛物线方程为 $x^2 = 4ay$，求小环运动的微分方程.

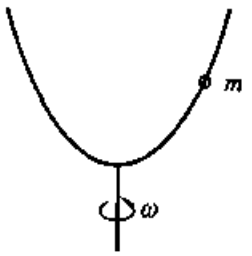


图 4　习题 12 图

13. 试讨论 §1.5.3(3.1)式和 §1.2.2(2.17)式之间的关系.

14. 根据 §1.5.3(3.13)式右边第四项，讨论科里奥利力对地球上河流的影响.

15. 根据 §1.5.3(3.13)式右边第五项，讨论惯性离心力的意义，并作图.

第六章　低速宏观运动规律的正则形式

在第一章里，我们用拉格朗日表述方式给出了低速宏观运动的规律. 在这种表述方式中，力学系统的状态通过给定(广义)坐标 q_α 和(广义)速度 $\dot{q}_\alpha$ 来确定. 拉格朗日 L 是 $q_\alpha, \dot{q}_\alpha$ (可能还有 t)的函数. 在以后几章里，给出了拉格朗日方程的一些应用. 现在来讨论低速宏观运动规律的另一种表述方式——正则表述方式.

从拉格朗日方程 §1.1.2(2.21)式可以看出，q_α 和 $\dot{q}_\alpha$ 的地位不平等. 这是它的一个缺点. 如果改用(广义)坐标 q_α 和(广义)动量 p_α 来决定系统的状态，则可以得到更为对称的方程，称为正则运动方程. 在本章第一节里就来建立正则运动方程. 第二节讨论守恒律，然后在第三节里讨论与正则运动方程有关的变换. 最后，在第四节讨论作用量的意义，并推导作用量作为 q, t 的函数所满足的哈密顿-雅可比方程. 它是经典力学运动方程的另一形式.

§1.6.1　哈密顿方程

(一)勒让德变换

为了从以 $q_\alpha, \dot{q}_\alpha$ 为变量转换到以 q_α, p_α 为变量，可以采用勒让德变换.

首先说明一下什么是勒让德变换.

设有两个变量 x, y 的两个函数 $P(x,y), Q(x,y)$ 满足

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \tag{1.1}$$

则 $P\mathrm{d}x + Q\mathrm{d}y$ 是一个全微分式

$$\mathrm{d}f = P(x,y)\mathrm{d}x + Q(x,y)\mathrm{d}y. \tag{1.2}$$

在四个变量 x, y, P, Q 中可以任选 $(x,y), (x,Q), (P,y), (P,Q)$ 中的一组为独立变量. 例如，为了将独立变量由 x, y 换为 x, Q，可以根据乘积微分的公式写出

$$\mathrm{d}(Qy) = y\mathrm{d}Q + Q\mathrm{d}y. \tag{1.3}$$

(1.2)式减(1.3)式得到

$$\mathrm{d}(f - Qy) = P\mathrm{d}x - y\mathrm{d}Q. \tag{1.4}$$

这样，我们就得到了两个变量 x 和 Q 的全微分式，代替原来的 x 和 y 的全微分式. 由(1.2)式得(1.4)式称为勒让德变换.

经过勒让德变换以后的函数是

$$g = f - Qy. \tag{1.5}$$

应该特别注意的是，在变换之前，f 是 x 和 y 的函数，如(1.2)式；而在变换以后，g

是 x 和 Q 的函数：$g = g(x,Q)$. 只要从 $Q = Q(x,y)$ 解出 $y = y(x,Q)$ 代入(1.5)式中，就使 g 成为 x 和 Q 的函数.

(二)哈密顿函数

设系统的拉格朗日函数是 $L = L(\boldsymbol{q},\dot{\boldsymbol{q}},t)$，其中 $\boldsymbol{q}$ 和 $\dot{\boldsymbol{q}}$ 分别代表 s 个 $q_\alpha,\dot{q}_\alpha(\alpha = 1,2,\cdots,s)$. 将 t 作为固定参量，求 L 在固定时刻 t 对 $\boldsymbol{q}$ 和 $\dot{\boldsymbol{q}}$ 的全微分(t 是固定参量，因而 $\mathrm{d}t \approx 0$，参看 §1.1.2)

$$\mathrm{d}L = \sum_{\alpha=1}^{s}\frac{\partial L}{\partial q_\alpha}\mathrm{d}q_\alpha + \sum_{\alpha=1}^{s}\frac{\partial L}{\partial \dot{q}_\alpha}\mathrm{d}\dot{q}_\alpha. \tag{1.6}$$

利用 §1.2.1 的(1.3)和(1.1)式，

$$\frac{\partial L}{\partial \dot{q}_\alpha} = p_\alpha, \qquad \frac{\partial L}{\partial q_\alpha} = \dot{p}_\alpha. \tag{1.7}$$

代入(1.6)式，得到

$$\mathrm{d}L = \sum_{\alpha=1}^{s}\dot{p}_\alpha\mathrm{d}q_\alpha + \sum_{\alpha=1}^{s}p_\alpha\mathrm{d}\dot{q}_\alpha. \tag{1.8}$$

为了转到以 q_α,p_α 为变量，进行勒让德变换. 写

$$\mathrm{d}\Big(\sum_{\alpha=1}^{s}p_\alpha\dot{q}_\alpha\Big) = \sum_{\alpha=1}^{s}\dot{q}_\alpha\mathrm{d}p_\alpha + \sum_{\alpha=1}^{s}p_\alpha\mathrm{d}\dot{q}_\alpha. \tag{1.9}$$

用(1.9)式减(1.8)式，得到

$$\mathrm{d}\Big(\sum_{\alpha=1}^{s}p_\alpha\dot{q}_\alpha - L\Big) = \sum_{\alpha=1}^{s}\dot{q}_\alpha\mathrm{d}p_\alpha - \sum_{\alpha=1}^{s}\dot{p}_\alpha\mathrm{d}q_\alpha. \tag{1.10}$$

这样就完成了从变量 $q_\alpha,\dot{q}_\alpha$ 到变量 q_α,p_α 的勒让德变换.

变换后的函数

$$H = \sum_{\alpha=1}^{s}p_\alpha\dot{q}_\alpha - L. \tag{1.11}$$

作为 q_α 和 $p_\alpha(\alpha = 1,2,\cdots,s)$ 的函数称为哈密顿函数. 它也可能显含 t(如果拉格朗日显含 t)，即

$$H = H(\boldsymbol{q},\boldsymbol{p},t). \tag{1.12}$$

上式中，$\boldsymbol{q},\boldsymbol{p}$ 代表 s 个 $q_\alpha,p_\alpha(\alpha = 1,2,\cdots,s)$. 将(1.11)式代入(1.10)式，得到

$$\mathrm{d}H = \sum_{\alpha=1}^{s}\dot{q}_\alpha\mathrm{d}p_\alpha - \sum_{\alpha=1}^{s}\dot{p}_\alpha\mathrm{d}q_\alpha. \tag{1.13}$$

将(1.11)式和 §1.2.1(1.15)式比较，并注意(1.7)式可见，当能量 E 用坐标和动量 $\boldsymbol{q},\boldsymbol{p}$ 表出时，其表达式就是哈密顿函数 $H(\boldsymbol{q},\boldsymbol{p})$. 但应注意，两者的意义不同. 哈密顿函数 H 的意义在于它作为 $\boldsymbol{q},\boldsymbol{p}$ 的函数的函数形式，见下面(1.15)式；而能量 E 是一个物理量，它的意义在于它所取的具体数值.

(三)哈密顿方程

写出 H 的全微分

$$dH = \sum_{\alpha=1}^{s} \frac{\partial H}{\partial p_\alpha} dp_\alpha + \sum_{\alpha=1}^{s} \frac{\partial H}{\partial q_\alpha} dq_\alpha. \tag{1.14}$$

和(1.13)式比较可见

$$\frac{dq_\alpha}{dt} = \frac{\partial H}{\partial p_\alpha}, \quad \frac{dp_\alpha}{dt} = -\frac{\partial H}{\partial q_\alpha}, \qquad \alpha = 1,2,\cdots,s. \tag{1.15}$$

这就是用哈密顿函数写出的力学系统的运动方程,称为哈密顿方程,又叫正则方程.

拉格朗日方程[§1.1.2(2.21)式]是 s 个函数 $q_\alpha(t)(\alpha = 1,2,\cdots,s)$ 的二阶微分方程;哈密顿方程(1.15)式是 $2s$ 个函数 $q_\alpha(t), p_\alpha(t)$ 的一阶微分方程.两者可以等效地作为系统的运动方程.

在哈密顿方程(1.15)式中,q_α, p_α 的地位平等,这是哈密顿方程的特点.正因为如此,所以称它为正则方程.q_α 和 p_α 称为相互共轭的正则变量.

(四)哈密顿方程的变分原理

哈密顿方程也可以从变分原理得到.为此,由(1.11)式解出

$$L = \sum_{\alpha=1}^{s} p_\alpha \dot{q}_\alpha - H. \tag{1.16}$$

代入§1.1.3(3.5)式,得到作用量

$$S = \int_A^B \left(\sum_{\alpha=1}^{s} p_\alpha dq_\alpha - H dt \right), \tag{1.17}$$

其中的积分限为

$$t_A = t_1, \quad t_B = t_2, \quad q_{\alpha A} = q_\alpha(t_1), \quad q_{\alpha B} = q_\alpha(t_2).$$

下面来证明,(1.17)式中的 S 的极值条件导致正则方程(1.15)式.

在证明之前先强调指出,按定义,拉格朗日函数是(广义)坐标 q_α,(广义)速度 $\dot{q}_\alpha$ 和时间 t 的函数.如果用 $H(\boldsymbol{q},\boldsymbol{p},t)$ 写出,如(1.16)式,就不再是原来意义下的"拉格朗日函数".我们只是利用它将作用量写成(1.17)式.因此,应该直接以(1.17)式作为以下讨论的出发点.

证　我们有

$$\delta S = \sum_{\alpha=1}^{s} \int_A^B \left[\delta p_\alpha dq_\alpha + p_\alpha d\delta q_\alpha - \frac{\partial H}{\partial q_\alpha} \delta q_\alpha dt - \frac{\partial H}{\partial p_\alpha} \delta p_\alpha dt \right],$$

对第二项分部积分

$$\int_A^B p_\alpha d\delta q_\alpha = p_\alpha \delta q_\alpha \Big|_A^B - \int_A^B \delta q_\alpha dp_\alpha.$$

按§1.1.3(3.10)式

$$\delta q_{\alpha A} = \delta q_{\alpha B} = 0,$$

故上式中积分出的项为零，于是得到

$$\delta S = \sum_{\alpha=1}^{s}\int_A^B \left\{\delta p_\alpha\left(\mathrm{d}q_\alpha - \frac{\partial H}{\partial p_\alpha}\mathrm{d}t\right) - \delta q_\alpha\left(\mathrm{d}p_\alpha + \frac{\partial H}{\partial q_\alpha}\mathrm{d}t\right)\right\}, \tag{1.18}$$

极值条件是 $\delta S = 0$，而由于各个 $\delta q_\alpha, \delta p_\alpha$ 相互独立，它们的系数应分别为零. 这样就得到哈密顿方程(1.15)式. 【证毕】

我们知道，在拉氏方程的变分原理中，拉氏量可以相差一个坐标和时间的任意函数对时间的全导数，见 §1.1.3(3.13)式. 在哈密顿方程中，$p_\alpha, q_\alpha(\alpha = 1,2,\cdots, s)$ 是独立变量，因而在哈密顿方程的变分原理中，作用量的被积函数可以相差一个坐标、动量和时间的任意函数 $f(\boldsymbol{q},\boldsymbol{p},t)$ 对时间的全微分，即

$$S' = \int_A^B \left(\sum_{\alpha=1}^{s} p_\alpha \mathrm{d}q_\alpha - H\mathrm{d}t + \mathrm{d}f\right) \tag{1.19}$$

和 S 描述同一个力学系统. 证明如下：

证 S' 和 S 之差

$$S' - S = f(\boldsymbol{q}_B,\boldsymbol{p}_B,t_B) - f(\boldsymbol{q}_A,\boldsymbol{p}_A,t_A)$$

是常数，即只依赖于端点 A,B 的 $\boldsymbol{q}_A,\boldsymbol{p}_A,\boldsymbol{q}_B,\boldsymbol{p}_B$ 的值，而不依赖于 $\boldsymbol{q}(t),\boldsymbol{p}(t)$. 因而 $\delta S' = \delta S$，由 S' 和 S 得到相同的运动方程. 【证毕】

(五)相空间

由共轭变量 q_α, p_α 形成的空间称为相空间，以别于单由 q_α 形成的位形空间. 为了说明相空间的意义，我们讨论笛卡儿坐标中的单个质点. 它的位形空间是由 3 个坐标分量 $x_i(i = 1,2,3)$ 形成的 3 维欧氏空间.

我们要强调指出的是，在任一时刻 t，给定位形空间中的一点 $\boldsymbol{r}(t)$，不能完全决定质点的运动. 为了决定质点的运动，还必须知道这一时刻位矢的导数 $\dot{\boldsymbol{r}}(t)$. 而这意味着，需要知道相邻时刻的 $\boldsymbol{r}(t)$.

为了使得给定空间中的一点能完全决定质点的运动，我们将 3 个坐标分量 $x_i(i = 1,2,3)$ 和 3 个动量分量 $p_i(i = 1,2,3)$ 合在一起，形成一个 6 维欧氏空间，称为这一质点的相空间. 在任一时刻 t，给定相空间中的一点，就完全决定了质点的运动.

我们说相空间是一个欧氏空间的原因是，其中任意两点 P 和 Q 的距离的平方定义为

$$(\overline{PQ})^2 = \sum_{i=1}^{3}(x_{Pi} - x_{Qi})^2 + \sum_{i=1}^{3}(p_{Pi} - p_{Qi})^2. \tag{1.20}$$

这和位形空间中任意两点 P 和 Q 的距离的平方

$$(\overline{PQ})^2 = \sum_{i=1}^{3}(x_{Pi} - x_{Qi})^2 \tag{1.21}$$

有相同的形式，因而都称为欧氏空间.

质点在相空间中的代表点随时间 t 的变化所描出的曲线称为质点的相轨迹. 对于周期运动，相轨迹是闭合曲线.

§1.6.2　守恒律　泊松括号

(一)力学量对时间的导数

在哈密顿形式下，力学体系的基本变量是 q_α, p_α. 任意力学量 f 是 q_α 和 p_α ($\alpha = 1, \cdots, s$) 的函数，也可能显含 t: $f = f(\boldsymbol{q}, \boldsymbol{p}, t)$，其中 $\boldsymbol{q}, \boldsymbol{p}$ 代表 s 个 q_α, p_α ($\alpha = 1, \cdots, s$). f 除了有可能显含 t 以外，还通过 $q_\alpha(t), p_\alpha(t)$ 而依赖于 t，因而 f 对 t 的导数是

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \sum_{\alpha=1}^{s}\left(\frac{\partial f}{\partial q_\alpha}\dot{q}_\alpha + \frac{\partial f}{\partial p_\alpha}\dot{p}_\alpha\right).$$

将哈密顿方程[§1.6.1(1.15)式]代入上式，得到

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \sum_{\alpha=1}^{s}\left(\frac{\partial f}{\partial q_\alpha}\frac{\partial H}{\partial p_\alpha} - \frac{\partial f}{\partial p_\alpha}\frac{\partial H}{\partial q_\alpha}\right). \tag{2.1}$$

引进符号

$$[H, f]_{\mathrm{PB}} \equiv \sum_{\alpha=1}^{s}\left(\frac{\partial H}{\partial p_\alpha}\frac{\partial f}{\partial q_\alpha} - \frac{\partial H}{\partial q_\alpha}\frac{\partial f}{\partial p_\alpha}\right), \tag{2.2}$$

称之为 H 和 f 的泊松括号[泊松括号的普遍定义见第(四)小节]，则有

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + [H, f]_{\mathrm{PB}}. \tag{2.3}$$

由此可见，物理量 f 对 t 的导数包含两项，一项是 f 对 t 的偏导数，另一项是哈密顿函数 H 和 f 的泊松括号.

(二)用泊松括号表出的运动方程

在(2.3)式中令 $f = q_\alpha$ 和 p_α，则有

$$\mathrm{d}q_\alpha/\mathrm{d}t = [H, q_\alpha]_{\mathrm{PB}}, \qquad \mathrm{d}p_\alpha/\mathrm{d}t = [H, p_\alpha]_{\mathrm{PB}}. \tag{2.4}$$

这是用泊松括号表出的运动方程. (2.4)式和上节(1.15)式一致. 其原因是

$$[H, q_\alpha]_{\mathrm{PB}} = \partial H/\partial p_\alpha, \qquad [H, p_\alpha]_{\mathrm{PB}} = -\partial H/\partial q_\alpha. \tag{2.5}$$

证　在(2.2)式中令 $f = q_\alpha$ 得到[将求和指标(哑标)改写为 β，以免指标混淆]

$$[H, q_\alpha]_{\mathrm{PB}} = \sum_{\beta=1}^{s}\left(\frac{\partial H}{\partial p_\beta}\frac{\partial q_\alpha}{\partial q_\beta} - \frac{\partial H}{\partial q_\beta}\frac{\partial q_\alpha}{\partial p_\beta}\right). \tag{2.6}$$

但是，q_α 和 p_β 是独立变量，因而

$$\partial q_\alpha / \partial p_\beta = 0, \tag{2.7}$$

而由附录(I.4)式,

$$\partial q_\alpha / \partial q_\beta = \delta_{\alpha\beta}. \tag{2.8}$$

将(2.7)式和(2.8)式代入(2.6)式,利用 δ 符号的性质就得到(2.5)式中的第一式.(2.5)中的第二式也可以类似地证明. **【证毕】**

(三)能量守恒与动量守恒

由(2.3)式知,如果 f 不显含时间 t,则

$$\frac{\mathrm{d}f}{\mathrm{d}t} = [H, f]_{PB} \quad (\text{如果 } f \text{ 不显含时间 } t). \tag{2.9}$$

因此,对于不显含 t 的 f,如果 $[H, f]_{PB} = 0$,则 $\mathrm{d}f/\mathrm{d}t = 0$ ——f 守恒.即:不显含 t 的力学量守恒的充分必要条件是它和 H 的泊松括号等于零.

由§1.6.1(1.11)、(1.7)式和§1.2.1(1.15)式知,在 H 不显含时间情况下,$E = H$.利用泊松括号的定义(2.2)式可见,$[H,E]_{PB} = 0$.因此,如果 H 不显含 t,则 E 是守恒量,即能量守恒.

再设哈密顿函数 H 中不包含某一个广义坐标,$\partial H/\partial q_\alpha = 0$,则由(2.5)式知 $[H,p_\alpha]_{PB} = 0$,因而 p_α 守恒.这一结论也可以直接从哈密顿方程[上节(1.15)式]得出.这样的 q_α 称为循环坐标.

(四)泊松括号的定义与性质

对于 p_α 和 q_α 的任意两个函数 f 和 g,泊松括号的定义为

$$[f,g]_{PB} \equiv \sum_{\alpha=1}^{s}\left(\frac{\partial f}{\partial p_\alpha}\frac{\partial g}{\partial q_\alpha} - \frac{\partial f}{\partial q_\alpha}\frac{\partial g}{\partial p_\alpha}\right). \tag{2.10}$$

从这一定义显然可见,$[f,g]_{PB}$ 在 f 和 g 互换时改号

$$[f,g]_{PB} = -[g,f]_{PB}, \tag{2.11}$$

因而又有

$$[f,f]_{PB} = 0. \tag{2.12}$$

如果 f 和 g 中有一个是常数 c,则由(2.10)式知

$$[f,c]_{PB} = 0. \tag{2.13}$$

另外,由(2.10)式容易证明

$$[f,c_1g_1 + c_2g_2]_{PB} = c_1[f,g_1]_{PB} + c_2[f,g_2]_{PB}, \tag{2.14}$$

$$[f,g_1g_2]_{PB} = g_1[f,g_2]_{PB} + [f,g_1]_{PB}\,g_2. \tag{2.15}$$

(2.15)式类似于对两函数的乘积求导的公式.因此,如果将 $[f,\cdots]_{PB}$ 看成是作用在位于省略号处的函数上的一种运算,则它是具有和求导运算类似的性质[(2.15)式]的线性运算[(2.14)式].

一种重要的特殊情况是 g 为某一个正则变量 q_α 或 p_α.此时,由(2.10)式,

$$[f,q_\alpha]_{PB}=\sum_{\beta=1}^{s}\left(\frac{\partial f}{\partial p_\beta}\frac{\partial q_\alpha}{\partial q_\beta}-\frac{\partial f}{\partial q_\beta}\frac{\partial q_\alpha}{\partial p_\beta}\right)=\sum_{\beta=1}^{s}\left(\frac{\partial f}{\partial p_\beta}\right)\delta_{\alpha\beta}=\frac{\partial f}{\partial p_\alpha},$$

$$[f,p_\alpha]_{PB}=\sum_{\beta=1}^{s}\left(\frac{\partial f}{\partial p_\beta}\frac{\partial p_\alpha}{\partial q_\beta}-\frac{\partial f}{\partial q_\beta}\frac{\partial p_\alpha}{\partial p_\beta}\right)=-\sum_{\beta=1}^{s}\left(\frac{\partial f}{\partial q_\beta}\right)\delta_{\alpha\beta}=-\frac{\partial f}{\partial q_\alpha},$$

即

$$[f,q_\alpha]_{PB}=\partial f/\partial p_\alpha,\qquad [f,p_\alpha]_{PB}=-\partial f/\partial q_\alpha,\tag{2.16}$$

[(2.5)式是此式的特例]. 由此立即得到

$$[q_\alpha,q_\beta]_{PB}=0,\quad [p_\alpha,p_\beta]_{PB}=0,\quad [p_\alpha,q_\beta]_{PB}=\delta_{\alpha\beta}.\tag{2.17}$$

最后,让我们来证明一个重要的恒等式——雅可比恒等式

$$[f,[g,h]]_{PB}+[g,[h,f]]_{PB}+[h,[f,g]]_{PB}=0.\tag{2.18}$$

证　由定义(2.10)式,得到(省去求和符号 $\sum_\alpha\sum_\beta$)

$$\begin{aligned}[f,[g,h]]_{PB}=&\frac{\partial f}{\partial p_\alpha}\frac{\partial g}{\partial p_\beta}\frac{\partial^2 h}{\partial q_\alpha\partial q_\beta}+\frac{\partial f}{\partial p_\alpha}\frac{\partial h}{\partial q_\beta}\frac{\partial^2 g}{\partial q_\alpha\partial p_\beta}-\\&\frac{\partial f}{\partial p_\alpha}\frac{\partial h}{\partial p_\beta}\frac{\partial^2 g}{\partial q_\alpha\partial q_\beta}-\frac{\partial f}{\partial p_\alpha}\frac{\partial g}{\partial q_\beta}\frac{\partial^2 h}{\partial q_\alpha\partial p_\beta}+\\&\frac{\partial f}{\partial q_\alpha}\frac{\partial g}{\partial q_\beta}\frac{\partial^2 h}{\partial p_\alpha\partial p_\beta}+\frac{\partial f}{\partial q_\alpha}\frac{\partial h}{\partial p_\beta}\frac{\partial^2 g}{\partial p_\alpha\partial q_\beta}-\\&\frac{\partial f}{\partial q_\alpha}\frac{\partial h}{\partial q_\beta}\frac{\partial^2 g}{\partial p_\alpha\partial p_\beta}-\frac{\partial f}{\partial q_\alpha}\frac{\partial g}{\partial p_\beta}\frac{\partial^2 h}{\partial p_\alpha\partial q_\beta}.\end{aligned}\tag{2.19}$$

为了醒目,采用符号

$$((fg))\equiv\frac{\partial f}{\partial p_\alpha}\frac{\partial g}{\partial p_\beta}\frac{\partial^2 h}{\partial q_\alpha\partial q_\beta}+\frac{\partial f}{\partial q_\alpha}\frac{\partial g}{\partial q_\beta}\frac{\partial^2 h}{\partial p_\alpha\partial p_\beta}-\frac{\partial f}{\partial q_\alpha}\frac{\partial g}{\partial p_\beta}\frac{\partial^2 h}{\partial p_\alpha\partial q_\beta}-\frac{\partial f}{\partial p_\alpha}\frac{\partial g}{\partial q_\beta}\frac{\partial^2 h}{\partial q_\alpha\partial p_\beta}.$$

显然,它对 f,g 对称:$((fg))=((gf))$. 于是可以写(2.19)式为

$$[f,[g,h]]_{PB}=((fg))-((fh)).$$

将 f,g,h 轮换,得到

$$[g,[h,f]]_{PB}=((gh))-((gf)),\qquad [h,[f,g]]_{PB}=((hf))-((hg)).$$

将此三式相加,右边各项相消得零. 从而证明了(2.18)式.　**【证毕】**

§1.6.3　正则变换

(一)正则变换的定义

在§1.1.2(2.9)式中定义广义坐标 $q_\alpha(\alpha=1,2,\cdots,s)$ 为:能够决定系统中所有质点位置的独立变量. 显然,如果

$$Q_\alpha=Q_\alpha(q_1,q_2,\cdots,q_s,t),\qquad \alpha=1,2,\cdots,s\tag{3.1}$$

为 $q_\alpha(\alpha = 1,2,\cdots,s)$ 的单值可逆函数，则由 $Q_\alpha(\alpha = 1,2,\cdots,s)$ 可以决定 $q_\alpha(\alpha = 1,2,\cdots,s)$，从而也就决定了系统中所有质点的位置. 因此，$Q_\alpha(\alpha = 1,2,\cdots,s)$ 也是系统的广义坐标. 举例来说，对于自由质点，笛卡儿坐标和球坐标之间或球坐标与柱坐标之间的变换都是(3.1)式型的变换.

(3.1)式型的变换表示广义坐标的选取不唯一. 这一论断对于拉格朗日表述方式和哈密顿表述方式都是对的.

但是，对于哈密顿表述方式来说，q_α 和 p_α 的地位平等，因而能够允许有一种比(3.1)式更广泛的变换，即同时包含 q_α 和 $p_\alpha(\alpha = 1,2,\cdots,s)$ 的变换：

$$\left.\begin{aligned} Q_\alpha &= Q_\alpha(\boldsymbol{q},\boldsymbol{p},t) \\ P_\alpha &= P_\alpha(\boldsymbol{q},\boldsymbol{p},t) \end{aligned}\right\}, \qquad \alpha = 1,2,\cdots,s, \tag{3.2}$$

其中的宗量 $(\boldsymbol{q},\boldsymbol{p})$ 表示 $(q_1,q_2,\cdots,q_s;p_1,p_2,\cdots,p_s)$. 如果经过这样变换以后，能够找到一个函数 $H'(\boldsymbol{Q},\boldsymbol{P},t)$，[$\boldsymbol{Q},\boldsymbol{P}$ 表示 $(Q_1,\cdots,Q_s;P_1,\cdots P_s)$]，使得运动方程[§1.6.1(1.15)式]

$$\dot{q}_\alpha = \partial H/\partial p_\alpha, \quad \dot{p}_\alpha = -\partial H/\partial q_\alpha, \qquad \alpha = 1,2,\cdots,s, \tag{3.3}$$

在转到新变量后，也具有同一正则形式：

$$\dot{Q}_\alpha = \partial H'/\partial P_\alpha, \quad \dot{P}_\alpha = -\partial H'/\partial Q_\alpha, \qquad \alpha = 1,2,\cdots,s, \tag{3.4}$$

则称变换(3.2)为正则变换. 函数 $H'(\boldsymbol{Q},\boldsymbol{P},t)$ 对于正则变换以后的坐标和动量起哈密顿函数的作用. 正则变换后的运动方程也可以类似于§1.6.2(2.4)式用 H' 和 Q_α,P_α 的泊松括号写出，而且变换以后的 Q_α,P_α 的泊松括号也满足§1.6.2(2.17)式.

(二)正则变换的生成函数

由§1.6.1(1.17)式知，正则运动方程(3.3)式可以由变分原理

$$\delta\int_A^B\left(\sum_{\alpha=1}^{s}P_\alpha\dot{q}_\alpha - H\right)\mathrm{d}t = 0 \tag{3.5}$$

导出. 类似地，方程(3.4)式可以由变分原理

$$\delta\int_A^B\left(\sum_{\alpha=1}^{s}P_\alpha\dot{Q}_\alpha - H'\right)\mathrm{d}t = 0 \tag{3.6}$$

导出.

(3.5)式和(3.6)式同时成立，并不意味着这两个积分的被积函数一定相等，而是意味着：两个被积函数之间，最多只能差一个坐标、动量的任意函数 F 对时间的全微分，如§1.6.1(1.19)式. 因此，由(3.5)式和(3.6)式得到

$$\sum_{\alpha=1}^{s}P_\alpha\dot{q}_\alpha - H = \sum_{\alpha=1}^{s}P_\alpha\dot{Q}_\alpha - H' + \frac{\mathrm{d}F}{\mathrm{d}t}. \tag{3.7}$$

(3.7)式中的 F 称为正则变换(3.2)式的生成函数. 它是 $\boldsymbol{p},\boldsymbol{q},\boldsymbol{P},\boldsymbol{Q},t$ 的函数，

总共 $4s+1$ 个变量. 由于有 $2s$ 个等式(3.2)式,所以除 t 以外,独立的变量只有 $2s$ 个. 这 $2s$ 个独立变量可以有四种不同取法[参看 §1.6.1(1.2)式后面的讨论]:

$$F_1(\boldsymbol{q},\boldsymbol{Q},t),F_2(\boldsymbol{q},\boldsymbol{P},t),F_3(\boldsymbol{p},\boldsymbol{Q},t),F_4(\boldsymbol{p},\boldsymbol{P},t). \tag{3.8}$$

如果选用 F_1 则可以展开 $\mathrm{d}F/\mathrm{d}t$ 为

$$\frac{\mathrm{d}F_1}{\mathrm{d}t}=\sum_{\alpha=1}^{s}\frac{\partial F_1}{\partial q_\alpha}\dot{q}_\alpha+\sum_{\alpha=1}^{s}\frac{\partial F_1}{\partial Q_\alpha}\dot{Q}_\alpha+\frac{\partial F_1}{\partial t}. \tag{3.9}$$

代入(3.7)式,经过移项整理,得到

$$\sum_{\alpha=1}^{s}\left(p_\alpha-\frac{\partial F_1}{\partial q_\alpha}\right)\dot{q}_\alpha=\sum_{\alpha=1}^{s}\left(P_\alpha+\frac{\partial F_1}{\partial Q_\alpha}\right)\dot{Q}_\alpha+\left(H-H'+\frac{\partial F_1}{\partial t}\right).$$

由于 q_α,Q_α 是独立变量,$\dot{q}_\alpha,\dot{Q}_\alpha$ 相互独立,因而上式中的 $\dot{q}_\alpha$ 和 $\dot{Q}_\alpha$ 的 1 次方和零次方的系数必须分别等于零,从而有

$$p_\alpha=\frac{\partial F_1}{\partial q_\alpha},\quad P_\alpha=-\frac{\partial F_1}{\partial Q_\alpha},\quad H'=H+\frac{\partial F_1}{\partial t}. \tag{3.10}$$

这样,只要给定了生成函数 $F_1(\boldsymbol{q},\boldsymbol{Q},t)$,就可以由前面 $2s$ 个方程解出 Q_α,P_α 作为 $\boldsymbol{q},\boldsymbol{p},t$ 的函数,即正则变换(3.2)式;同时,由(3.10)的最后一式可以得到变换后的哈密顿函数 H'. 显然,如果生成函数不显含时间 t 则变换前后的哈密顿相等,$H'=H$,只是独立变量不同(对于 H 是 $\boldsymbol{p},\boldsymbol{q}$;对于 H' 是 $\boldsymbol{P},\boldsymbol{Q}$).

将(3.10)式代入(3.9)式得到

$$\frac{\mathrm{d}F_1}{\mathrm{d}t}=\sum_{\alpha=1}^{s}p_\alpha\dot{q}_\alpha-\sum_{\alpha=1}^{s}P_\alpha\dot{Q}_\alpha+(H'-H). \tag{3.11}$$

为了得到生成函数 F_2 决定的正则变换,需要将(3.11)式右边的 $P_\alpha\dot{Q}_\alpha$ 换成 $Q_\alpha\dot{P}_\alpha$. 这可以通过勒让德变换来实现. 具体说来,就是在(3.11)式的左右两边加上一个恒等式

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{\alpha=1}^{s}P_\alpha Q_\alpha\right)=\sum_{\alpha=1}^{s}P_\alpha\dot{Q}_\alpha+\sum_{\alpha=1}^{s}Q_\alpha\dot{P}_\alpha,$$

并令

$$F_2=F_1+\sum_{\alpha=1}^{s}P_\alpha Q_\alpha, \tag{3.12}$$

结果得到

$$\frac{\mathrm{d}F_2}{\mathrm{d}t}=\sum_{\alpha=1}^{s}p_\alpha\dot{q}_\alpha+\sum_{\alpha=1}^{s}Q_\alpha\dot{P}_\alpha+(H'-H). \tag{3.13}$$

由此得到

$$p_\alpha=\frac{\partial F_2}{\partial q_\alpha},\quad Q_\alpha=\frac{\partial F_2}{\partial P_\alpha},\quad H'=H+\frac{\partial F_2}{\partial t}. \tag{3.14}$$

由此可以解出生成函数 F_2 所决定的正则变换.

由生成函数 F_3, F_4 决定的正则变换可以类似地用勒让德变换的方法得到.

(三)正则变换举例

首先考虑由 $F_2(\boldsymbol{q}, \boldsymbol{P}, t)$ 生成的变换 .

最简单的情况是

$$F_2 = \sum_\alpha q_\alpha P_\alpha. \tag{3.15}$$

此时,由(3.14)式知

$$p_\alpha = P_\alpha, \qquad q_\alpha = Q_\alpha, \qquad H' = H.$$

这是恒等变换,也就是不变换.

略为复杂一点的是

$$F_2 = \sum_{\alpha=1}^{s} f_\alpha(\boldsymbol{q}, t) P_\alpha. \tag{3.16}$$

这里 $f_\alpha(\alpha = 1, 2, \cdots, s)$ 是 $(q_1, q_2, \cdots, q_s, t)$ 的给定函数. 由(3.14)式有

$$Q_\alpha = f_\alpha(q_1, q_2, \cdots, q_s, t). \tag{3.17}$$

这就是(3.1)式型的变换.

再来看由 $F_1(\boldsymbol{q}, \boldsymbol{Q}, t)$ 生成的变换 .

最简单的情况是

$$F_1 = \sum_\alpha q_\alpha Q_\alpha. \tag{3.18}$$

由(3.10)式得到

$$p_\alpha = Q_\alpha, \qquad P_\alpha = -q_\alpha, \qquad H' = H. \tag{3.19}$$

这样,老动量成了新坐标,老坐标成了新动量(改号). 这就清楚地表明,由于在哈密顿表述方式中,坐标和动量处于平等地位,所以坐标和动量的区分失去了绝对的意义. 正因此称 q_α, p_α 为相互共轭的正则变量.

最后一个例子是一维线性谐振子. 由拉格朗日 §1.4.1(1.6)式,写 x 为 q 可以得到哈密顿

$$H = p^2/(2m) + m\omega^2 q^2/2. \tag{3.20}$$

取生成函数为

$$F_1 = \frac{1}{2} m\omega q^2 \cot Q, \tag{3.21}$$

则由(3.10)式得到

$$p = m\omega q \cot Q, \qquad P = m\omega q^2/(2\sin^2 Q). \tag{3.22}$$

由此解出

$$q = \sqrt{2P/(m\omega)} \sin Q, \qquad p = \sqrt{2m\omega P} \cos Q. \tag{3.23}$$

由于 F_1 不显含时间,所以变换后的哈密顿仍然是(3.20),只要用(3.23)式代其中

的 q 和 p .结果得到

$$H' = \omega P\cos^2 Q + \omega P\sin^2 Q = \omega P. \tag{3.24}$$

这样,变换以后的哈密顿不包含 Q ,表明 Q 是循环坐标,因而对应的正则变量 P 是守恒量.

将(3.24)式代入(3.4)式,得到运动方程

$$\dot{Q} = \frac{\partial H'}{\partial P} = \omega. \tag{3.25}$$

它的解是 $Q = \omega t + \varphi$,其中 φ 是一个积分常数.代入(3.23)式得到

$$q = A\sin(\omega t + \varphi). \tag{3.26}$$

这就是通常的谐振子解.振幅 A 是一个常数,守恒量 P 是 $m\omega A^2/2 = E/\omega$ [见 §1.4.1(1.11)式].

在这个例子中,正则变换使得 Q 成为循环坐标,从而求解运动方程的问题简化成一个简单的积分[(3.25)式].这个例子本身是平庸的,但是它表明,只要能找到合适的正则变换,在很多情况下能够使求解复杂运动方程的问题大为简化.

(四)时间演化作为正则变换

为了简单,下面的讨论限于一维.所得结果不难推广到任意 s 维,只是书写上较为复杂.用 $q(t)$,$p(t)$ 表示 t 时刻的广义坐标和广义动量;用

$$Q(t) = q(t + \mathrm{d}t), \qquad P(t) = p(t + \mathrm{d}t) \tag{3.27}$$

表示 $t + \mathrm{d}t$ 时刻的广义坐标和广义动量.显然它们都满足哈密顿方程:

$$\frac{\mathrm{d}q}{\mathrm{d}t} = \left(\frac{\partial H}{\partial p}\right)_q, \quad \frac{\mathrm{d}p}{\mathrm{d}t} = -\left(\frac{\partial H}{\partial q}\right)_p. \tag{3.28}$$

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = \left(\frac{\partial H'}{\partial P}\right)_Q, \quad \frac{\mathrm{d}P}{\mathrm{d}t} = -\left(\frac{\partial H'}{\partial Q}\right)_P. \tag{3.29}$$

其中 $H'(Q,P)$ 是用 Q,P 表出的哈密顿函数.为了避免混淆,在(3.28)(3.29)式中用下标表明求偏导时保持不变的量.由(3.27)式可见,从 $q(t)$,$p(t)$ 到 $Q(t) = q(t + \mathrm{d}t)$,$P(t) = p(t + \mathrm{d}t)$ 是力学系统的时间演化,而由(3.28)(3.29)式可见,可以将它看成是一个正则变换.

* §1.6.4　作为坐标和时间函数的作用量　哈密顿-雅可比方程

(一)作为坐标和时间函数的作用量

在 §1.1.3 中定义作用量为

$$S = \int_{t_1}^{t_2} L(\boldsymbol{q}(t), \dot{\boldsymbol{q}}(t), t)\,\mathrm{d}t. \tag{4.1}$$

当固定 t_1 ,t_2 两个时刻的坐标

$$q_\alpha(t_1) = q_1^{(\alpha)}, \quad q_\alpha(t_2) = q_2^{(\alpha)}, \quad \alpha = 1,2,\cdots,s \tag{4.2}$$

时，作用量是 $q_\alpha(t)$ 的泛函，

$$S = S[q(t)]. \tag{4.3}$$

这一泛函的极值条件，即拉格朗日方程，决定了实际运动的路径 $q_\alpha(t) = \bar{q}_\alpha(t)$．我们称它为“经典路径”．和经典路径对应的作用量称为经典作用量 $S_{经典} = S[\bar{q}(t)]$．在前面的章节中，我们关心的是粒子的运动情况，它由 $\bar{q}_\alpha(t)$ 描述，在利用作用量的极值条件得到 $q_\alpha(t)$ 以后，就不再关心作用量所取的值 $S_{经典}$（在第一章习题12中举过一个计算 $S_{经典}$ 所取的值的例子）．然而，以后将看到，不仅是使作用量取极值的 $q_\alpha(t)$，而且是作用量的极值 $S_{经典}$ 本身，都有重要的物理意义．

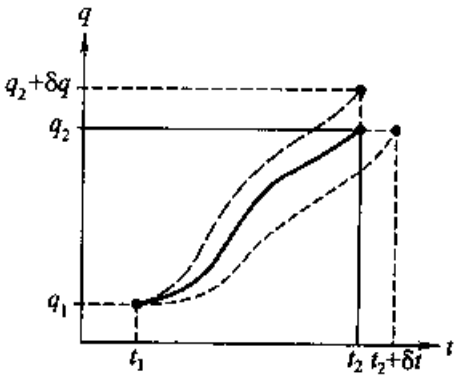


图1　经典路径随端点的变化

本节讨论 $S_{经典}$ 作为端点的函数的意义．只考虑上端点 (t_2,q_2)．固定 t_1 时刻的 $q_1^{(\alpha)}$，改变 t_2 时刻的 $q_2^{(\alpha)}$．对于每一个 $q_2^{(\alpha)}$ 的值都有一个 $S_{经典}$ 的值，如图1中的长虚线．类似地，固定 $q_2^{(\alpha)}$，改变 t_2，也会得到不同的 $S_{经典}$，如图1中的短虚线．因此，$S_{经典}$ 是 $q_2^{(\alpha)}$ 和 t_2 的函数．略去下标“经典”和“2”，就可以写：

$$S = S(q,t). \tag{4.4}$$

注意(4.4)式右边的圆括号和(4.3)式右边的方括号的区别．(4.4)式中的圆括号表明，S 是上端点 $(q_\alpha,t;\alpha = 1,\cdots,s)$ 的多元函数，此时默认 S 为经典作用量．(4.3)式中的方括号表明，S 是 $q_\alpha(t)$ 的泛函，此时默认 $q_\alpha(t)$ 满足端点条件(4.2)式．

当 $q_\alpha(t)$ 改变 δq_α 时，S 的变化如§1.1.3(3.9)式

$$\delta S = \sum_{\alpha=1}^{s} \frac{\partial L}{\partial \dot{q}_\alpha}\delta q_\alpha \Big|_{t_1}^{t_2} + \int_{t_1}^{t_2} \sum_{\alpha=1}^{s} \left(\frac{\partial L}{\partial q_\alpha} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_\alpha}\right)\delta q_\alpha \mathrm{d}t. \tag{4.5}$$

右边第一项是由于端点的变化引起的 S 的变化，如本节图1；第二项是在固定端点条件下由于 $q_\alpha(t)$ 的变化引起的 S 的变化，如§1.1.3图1．当 $q_\alpha(t) = \bar{q}_\alpha(t)$ 为经典路径时，第二项为零．其原因是，经典路径 $\bar{q}_\alpha(t)$ 满足拉格朗日方程；或者换句话说，是因为满足端点条件(4.2)式的经典路径 $\bar{q}_\alpha(t)$ 使 S 取极值，相应的 S 的一级变分[即(4.5)式右边第二项]为零．因此，和经典路径对应的经典作用量的变化是

$$\delta S = \sum_{\alpha=1}^{s} \frac{\partial L}{\partial \dot{q}_\alpha}\delta q_\alpha \Big|_{t_1}^{t_2}. \tag{4.6}$$

按假定，下端点是固定的，因而 $\delta q_\alpha(t_1) = 0$，再由§1.2.1(1.3)式 $p_\alpha = \partial L/\partial \dot{q}_\alpha$ 就有

$$\delta S = \sum_{\alpha=1}^{s} p_\alpha \delta q_\alpha. \tag{4.7}$$

由此知，S 对广义坐标 q_α 的偏导等于相应的广义动量，

$$\partial S/\partial q_\alpha = p_\alpha. \tag{4.8}$$

（二）哈密顿-雅可比方程

我们来推导 $S(q,t)$ 所满足的方程．由定义(4.1)式知，$S(q,t)$ 对 t 的全导数等于拉格朗日

函数

$$dS/dt = L. \tag{4.9}$$

另一方面，根据多元函数求导的法则，

$$\frac{dS}{dt} = \frac{\partial S}{\partial t} + \sum_{\alpha=1}^{s} \frac{\partial S}{\partial q_\alpha}\dot{q}_\alpha.$$

将(4.8)、(4.9)式代入，得到

$$L = \frac{\partial S}{\partial t} + \sum_{\alpha=1}^{s} p_\alpha \dot{q}_\alpha.$$

于是有

$$\frac{\partial S}{\partial t} = L - \sum_{\alpha=1}^{s} p_\alpha \dot{q}_\alpha = -H(\boldsymbol{q},\boldsymbol{p},t). \tag{4.10}$$

其中 H 是哈密顿函数，见 §1.6.1(1.11)式. 将(4.8)式代入，得到

$$\frac{\partial S}{\partial t} + H\left(\boldsymbol{q},\frac{\partial S}{\partial \boldsymbol{q}},t\right) = 0. \tag{4.11}$$

这就是 S 作为 $q_\alpha, t(\alpha = 1,\cdots,s)$ 的函数所满足的偏微分方程，称为哈密顿-雅可比方程.

哈密顿-雅可比方程是 $s+1$ 个变量的一阶偏微分方程. 它的完全积分有 $s+1$ 个任意常数. 由于在方程中只包含未知函数 S 的偏导数，而不包含 S 本身，见(4.11)式，因而，如果 $S = f$ 是方程的解，则 $S = f + C$[C 为任意常数也必是方程的解. 这表明，在 $s+1$ 个任意常数中有一个是相加常数. 因此，哈密顿-雅可比方程的完全积分可写成

$$S = S(t,q_1,\cdots,q_s,c_1,\cdots,c_s) + C. \tag{4.12}$$

除 C 以外，其他 s 个常数 $c_1,\cdots,c_s$ 没有一个是相加性常数.

(三)利用哈密顿-雅可比方程求解动力学问题

下面来讨论，在求得哈密顿-雅可比方程的完全积分(4.12)式以后，如何得到描述力学系统运动的 $q_\alpha(t), p_\alpha(t)(\alpha = 1,\cdots,s)$.

哈密顿-雅可比方程的完全积分 $S(t,q_1,\cdots,q_s,c_1,\cdots,c_s)$ 本身不直接描述系统的运动，但可以将它作为正则变换的生成函数来描述正则变量 $q_\alpha, p_\alpha(\alpha = 1,\cdots,s)$ 随时间的变化. 在 §1.6.3 中已说明，利用生成函数可以将正则变量 $q_\alpha, p_\alpha(\alpha = 1,\cdots,s)$ 变换到一组新的正则变量 Q_α，$P_\alpha(\alpha = 1,\cdots,s)$. 作为一种特殊情况，令 Q_α, P_α 为一个特定时刻(例如初始时刻) t_0 的 q_α, p_α

$$Q_\alpha(t) = q_\alpha(t_0) = A_\alpha, \qquad P_\alpha(t) = p_\alpha(t_0) = c_\alpha. \tag{4.13}$$

由于 Q_α, P_α 是初始时刻 t_0 的正则变量，不依赖于 t，对 t 而言它们是常数. 这样的正则变换将任意时刻 t 的正则变量和初始时刻 t_0 的正则变量(即初条件)联系起来，从而描述了系统的运动.

将(4.8)式和 §1.6.3(3.10)，(3.14)式比较可见，S 适合于作第一类或第二类生成函数. 我们令 S 为第二类生成函数，选取任意常数 c_α 为

$$c_\alpha = P_\alpha, \qquad \alpha = 1,\cdots,s. \tag{4.14}$$

这样就有

$$S(t,q_1,\cdots,q_s,c_1,\cdots,c_s) = F_2(t,q_1,\cdots,q_s,P_1,\cdots,P_s). \tag{4.15}$$

由 §1.6.3(3.14)的第二式

$$\partial S/\partial P_\alpha = Q_\alpha. \qquad \alpha = 1,\cdots,s. \tag{4.16}$$

这是一组 s 个变量 $q_\alpha(\alpha = 1,\cdots,s)$ 的 s 个方程,从中可以解出 $q_\alpha(\alpha = 1,\cdots,s)$ 作为 $P_\alpha,Q_\alpha,t(\alpha = 1,\cdots,s)$ 的函数

$$q_\alpha = q_\alpha(t,P,Q). \tag{4.17}$$

再利用(4.8)式得到

$$p_\alpha = p_\alpha(t,P,Q). \tag{4.18}$$

这就是所需要的,由初条件(4.13)式决定的正则变量随时间 t 的变化.

例 一维谐振子的哈密顿是

$$H = \frac{1}{2m}(p^2 + m^2\omega^2q^2), \tag{4.19}$$

因而有哈密顿-雅可比方程

$$\frac{1}{2m}\left[\left(\frac{\partial S}{\partial q}\right)^2 + m^2\omega^2q^2\right] + \frac{\partial S}{\partial t} = 0. \tag{4.20}$$

在此方程中,对 t 的偏导数单独出现,因而可以将解写成

$$S(q,c,t) = W(q,c) - ct. \tag{4.21}$$

其中,$W(q,c)$ 满足方程

$$\frac{1}{2m}\left[\left(\frac{\partial W}{\partial q}\right)^2 + m^2\omega^2q^2\right] = c. \tag{4.22}$$

积分得 $W = \sqrt{2mc}\int dq\ \sqrt{1 - m\omega^2q^2/(2c)} + C$. 由此得到哈密顿-雅可比方程的解

$$S = \sqrt{2mc}\int dq\ \sqrt{1 - m\omega^2q^2/(2c)} - ct + C.$$

相加性常数 C 对结果没有贡献,可以略去.按(4.14)式写 $c = P$ 有

$$S = \sqrt{2mP}\int dq\ \sqrt{1 - m\omega^2q^2/(2P)} - Pt. \tag{4.23}$$

代入(4.16)式

$$\frac{\partial S}{\partial P} = Q = \sqrt{\frac{2m}{P}}\int \frac{dq}{\sqrt{1 - m\omega^2q^2/(2P)}} - t = \frac{1}{\omega}\sin^{-1}q\ \sqrt{m\omega^2/(2P)} - t.$$

由此得

$$q = A\sin(\omega t + \varphi), \tag{4.24}$$

其中已改写 $\sqrt{2P/(m\omega^2)} = A,\omega Q = \varphi$.

习 题

1.分别写出单个质点在笛卡儿坐标,柱坐标和球坐标中的哈密顿函数.

2.对于有心力场中的质点,建立其哈密顿方程,并消去其循环坐标.

3.画出质量为 m,角频率为 ω 的一维谐振子以不同振幅 A 振动时的相轨迹.求各条相轨迹所包围的面积.

4.求质点的动量 $\boldsymbol{p}$ 的笛卡儿分量和角动量 $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$ 的笛卡儿分量之间的泊松括号.

5.求角动量 $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$ 的各个笛卡儿分量之间的泊松括号.

6. 设 f 为 $\boldsymbol{r}$ 和 $\boldsymbol{p}$ 的具有球对称性的函数，证明 f 和 L_z 的泊松括号为零：$[f,L_z]_{PB}=0$.

提示　由于球对称性，f 只能是 $\boldsymbol{r}^2,\boldsymbol{p}^2$，和 $\boldsymbol{r}\cdot\boldsymbol{p}$ 的函数，$f=f(\boldsymbol{r}^2,\boldsymbol{p}^2,\boldsymbol{r}\cdot\boldsymbol{p})$.

7. 作出由 $F_3(\boldsymbol{p},\boldsymbol{Q},t)$ 和 $F_4(\boldsymbol{p},\boldsymbol{P},t)$ [见 §1.6.3(3.8)式]决定的正则变换.

8. 根据 §1.6.4(4.23)式求谐振子的动量 $p(t)$.

9. 在 §1.6.4(4.10)式中作函数代换 $\psi=e^{(i/\hbar)S}$，其中 $\hbar$ 是普朗克常数，见附录Ⅶ. 求 ψ 所满足的方程.



第二篇　高速运动的规律

在第一篇里阐述了低速宏观物体的运动规律. 实验表明，这些规律对于高速运动不再适用. 例如，一个通过 10MV（兆伏）电势差加速的电子，其速度为 0.9988c（其中 c 为真空中的光速）. 如果让电子通过 40MV 电势差，按照低速运动的动能公式 $T = mv^2/2$，其速度应增加一倍，但实验结果表明，电子的速度仅由 0.9988c 提高到 0.9999c. 大量的事实证明，如果将低速运动的规律应用到速度接近于光速的范围，就会得出不正确的结果.

1905 年爱因斯坦提出狭义相对论，建立了高速运动的基本规律，成功地解决了上述困难. 相对论是对旧时间空间观念的根本变革，对近代物理学的发展起了重大作用.

在这一篇里，我们就来讨论以狭义相对论为基础的高速运动规律. 第一章介绍狭义相对论的实验基础和基本原理；第二章讨论由狭义相对论引起的时空结构的变化；第三、四章讨论建立在狭义相对论基上的高速运动粒子和电磁场的运动规律.

第一章　狭义相对论的实验基础和基本原理

狭义相对论的基本原理是从高速运动现象，特别是电磁波的传播现象中揭示出来的.

在19世纪中叶，随着生产的发展，人们发现了一系列电磁现象的实验定律. 1864年麦克斯韦把这些实验定律总结为麦克斯韦方程组，并预言了电磁波的存在. 在19世纪后期，随着电磁波的发现及无线电技术的发展，人们逐渐认识到电磁场是一种客观存在. 但是，由于受到机械论的影响，仍把电磁场看作某种充满整个宇宙的特殊介质“以太”的运动形态，把电磁波看成是以太介质的某种振动的传播形式，这就是所谓的“以太理论”. 另一方面，由麦克斯韦方程组能导出波动方程，得到电磁波的传播速度为 c 的结论(见 §3.3.1). 按照以太理论，这一结论只对“以太参考系”成立，因为根据伽利略的旧时空观，见 §1.1.4(4.3)、(4.4)式，对于其他惯性参考系，光的速度不可能为 c. 由此出发，会得到结论，利用电磁运动可以确定一个特殊的参考系，从而使相对于这一参考系的运动成为“绝对运动”. 于是，经典力学中“一切惯性系等价”的相对性原理(见 §1.1.4)在电磁现象中不再成立.

但是，在19世纪末对光速的测量证实了“在真空中光速相对任何惯性系都等于 c”的结论，从而否定了特殊参考系的存在，并为普遍适用于机械运动和电磁运动的相对性原理的建立奠定了基础.

本章先介绍具有代表性的迈克耳孙－莫雷实验(1887)年，然后介绍相对论的基本原理.

§2.1.1　迈克耳孙-莫雷实验

以太理论认为，只有在相对以太静止的特殊参考系里，真空中沿各个方向的光速才是 c.

因此，如果存在以太，而地球相对以太运动的速度为 v，则地球上的观察者观察到的光速应为 $c-v$，因而光速的大小依赖于 c 与 v 的夹角. 只要我们测出各个方向光速的差异，就有可能探知地球相对以太的运动，这就是迈克耳孙-莫雷实验的基本设想.

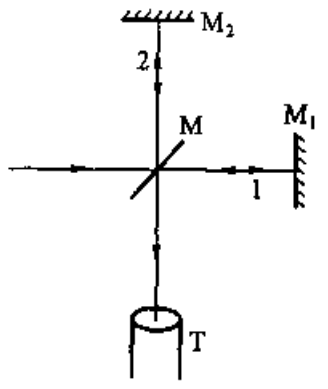


图1　迈克耳孙干涉仪示意图

迈克耳孙-莫雷实验的装置如图1所示. 迈克耳孙干涉仪是固定在地球上的.

地球绕太阳公转运动的速度 $v \approx 30\ \mathrm{km \cdot s^{-1}}$. 测量在不同季节的白天和晚上进行，因而可以估计测量过程中，地球相对以太参考系的运动速度至少和 v 有同一数量级.

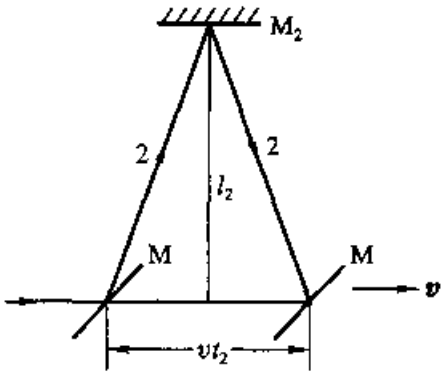


图 2　波束 2 的光程

设由光源 S 发出的一束平行光线被半镀银镜 M 分成两束：透过 M 的光束 1 被镜 M_1 反射回 M，再被 M 反射至望远镜 T；被 M 反射的光束 2 经 M_2 反射后回到 M，然后透过 M 到达望远镜 T. 两束光在望远镜的像平面上形成干涉条纹. 由于装置中 M 与光束方向的夹角为 45°，M_1 与 M_2 近似互相垂直，故在望远镜中将会观察到由空气尖劈所形成的等厚度干涉条纹.

设 $MM_1 = l_1$，$MM_2 = l_2$，又设地球相对以太参考系的速度 v 沿 MM_1 方向，则可求得两束光的相位差. 根据伽利略的速度合成定理[§1.1.4(4.5)式]，光束 1 从 M 到 M_1，再从 M_1 到 M 所经历的时间为

$$t_1 = \frac{l_1}{c - v} + \frac{l_1}{c + v} = \frac{2l_1}{c} \cdot \frac{1}{1 - (v/c)^2}. \tag{1.1}$$

光束 2 从 M 到 M_2，再从 M_2 到 M 所经历的时间 t_2 应满足(见图 2)

$$2\sqrt{l_2^2 + (vt_2/2)^2} = ct_2,$$

即

$$t_2 = \frac{2l_2}{\sqrt{c^2 - v^2}} = \frac{2l_2}{c} \cdot \frac{1}{\sqrt{1 - (v/c)^2}}. \tag{1.2}$$

由此可见，两束光的相位差为

$$\Delta\varphi = \frac{2\pi}{T}(t_2 - t_1) = \frac{2\pi}{T}\frac{2}{c}\left[\frac{l_2}{\sqrt{1 - (v/c)^2}} - \frac{l_1}{1 - (v/c)^2}\right]. \tag{1.3}$$

如果将仪器旋转 90°，则两束光位置互换，两束光的相位差为

$$\Delta\varphi' = \frac{2\pi}{T}(t'_2 - t'_1) = \frac{2\pi}{T}\frac{2}{c}\left[\frac{l_2}{1 - (v/c)^2} - \frac{l_1}{\sqrt{1 - (v/c)^2}}\right]. \tag{1.4}$$

因此，两束光的相位差的改变为

$$\Delta\varphi' - \Delta\varphi = \frac{2\pi}{\lambda}(l_1 + l_2)\frac{v^2}{c^2}. \tag{1.5}$$

在得到上式时已将分母按泰勒级数展开，取到 v^2/c^2 项.

干涉条纹移动的条数为：

$$\Delta N = \frac{\Delta\varphi' - \Delta\varphi}{2\pi} = \frac{(l_1 + l_2)}{\lambda}\frac{v^2}{c^2}. \tag{1.6}$$

迈克耳孙-莫雷实验的 $l_1 + l_2 \approx 22$ m，而 $v/c = 10^{-4}$，如取 $\lambda = 5.5 \times 10^{-7}$ m，则 $\Delta N = 0.4$．因此，干涉条纹将移动 0.4 个条纹的距离，根据当时的实验水平，这样的条纹移动，肯定能被探测到．但是，在一年四季的白天和夜晚进行实验，都没有观察到条纹的移动．

解释迈克耳孙-莫雷实验结果有两种可能．一是认为任何惯性系中所有方向测得的真空中的光速均为 c．当仪器转动 90°后，两束光的相位差没有变化，因而 $\Delta N = 0$．但是，在任何惯性系测得的光速均为 c，将明显与伽利略变换矛盾，这就要摒弃旧的时空观念．第二种考虑是仍然保留以太参考系的概念，而设法提出其他假设解释实验结果．

1892 年前后，洛伦兹和斐兹杰惹提出了“收缩假说”，他们认为所有物体在相对以太运动的方向上有一按因子 $\sqrt{1-(v/c)^2}$ 的收缩．这个假设能解释迈克耳孙-莫雷实验的结果．但是，如果洛伦兹－斐兹杰惹的收缩理论是正确的话，相对以太运动的透明物质将可发现双折射现象．但是 1902 年瑞利作的实验和 1904 年布拉斯作的实验都得到否定的结果．

另一个假设是“以太牵引说”．它假定以太会粘附于各种有质量的物体上，于是在地球周围的以太被地球所牵引，跟地球一起运动．这样，在地面附近所作的实验就不存在相对以太的运动．迈克耳孙-莫雷实验的结果也就得到解释，但是这又与菲索的流水实验及光行差现象（见 §2.2.5）矛盾．

此外还有各种假设，均能解释部分现象而不能解释另一部分现象，在此不再赘述．

§2.1.2　相对论的基本原理

1905 年，爱因斯坦总结了前人的一系列实验事实，在“论运动物体的电动力学”一文中提出了两条相对论的基本假设：

1. **相对性原理**❶：包括机械运动规律和电磁运动规律在内的一切物理规律在所有惯性系中是相同的，不存在特殊的惯性系．

2. **光速不变原理**：在所有惯性系内，真空中的光速沿任何方向均为 c，并与光源运动无关．

由这两个假设可以直接导出整个狭义相对论．这一理论不仅成功地解释了当时已有的全部实验事实，而且预言了新的效应，为以后的实验所证实．例如，1964

❶ 为了区别于伽利略相对性原理（§1.1.4），有时称这里的相对性原理为“爱因斯坦相对性原理”．

年和1966年柯尔法格等人让 π^0 介子以速度 $v = 0.9975c$ 运动，在运动中衰变为两个光子．实验测得沿 π^0 介子运动方向放出的光子速度与静止光源测得的光速一致．1972年，西欧核子研究中心的贝利和皮加索等人观测沿圆形轨道飞行的 μ 子束，对飞行 μ 子的寿命进行测量，证实了时钟延缓效应（见§2.2.4）．1975年斯耐德等人测定的横向多普勒效应也与理论一致．半个世纪以来，大量的关于质能关系和运动学的实验，无一不与相对论的预言相符合．这一切都证明了，建立在相对性原理与光速不变原理上的狭义相对论的巨大成功．

容易看出，爱因斯坦的相对性原理与旧时空观存在着深刻的矛盾．我们用下面两个例子来说明．

首先，伽利略的速度合成定理与光速不变原理相矛盾．设K′参考系以 $\boldsymbol{v} = v\boldsymbol{e}_x$ 相对K参考系运动．如果光源在K′系静止，光源在K′系发出的光沿 $\pm \boldsymbol{e}_x$ 方向的光速均为 c．根据伽利略的速度合成定理，在K系沿 $\pm \boldsymbol{e}_x$ 方向的光速应分别为 $c \pm v$．这与光速不变原理存在着明显的矛盾．

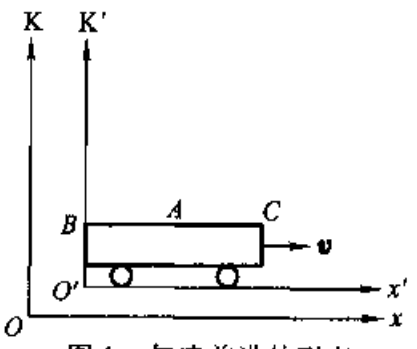


图1　匀速前进的列车

其次，在伽利略变换中，时间是绝对的，对子不同惯性系，$t = t'$．因此，如果两个事件在K′系中是同时发生的，$\Delta t' = 0$，则对任一惯性系K来说，这两个事件也必然是同时发生的，$\Delta t = 0$．但是从光速不变原理来看，在K′系中同时发生的两个事件，在K系看来就不一定是同时的了．例如，火车（K′系）以 $\boldsymbol{v} = v\boldsymbol{e}_x$ 相对地面（K系）作匀速直线运动（图1）．设在 $t' = 0$ 时有一个雷电打在车厢中点 A 处．我们把闪光传到车厢末端的 B 点称为事件1，把闪光传到车厢前端的 C 点称为事件2，已知 $BA = AC$．这样，由于光速在K′系沿各个方向均为 c，故两闪光同时传至 B 点和 C 点，即K′系观察者测得事件1和事件2是同时发生的．但在K系观察者看来，雷打在 A 处后，当光向四面八方传播时，B 点迎着闪光移动，而 C 点则背着闪光移动．考虑到光速在K系沿各个方向也为 c，所以闪光应先到达 B 点，后到达 C 点，即事件1和事件2在K系中看来，就不是同时发生的了．

由此可见，根据低速运动所形成的时空观（牛顿-伽利略时空观）仅在一定范围内是正确的．当研究高速运动特别是电磁波传播的现象时，就会发现旧时空观与实验事实的深刻矛盾．因此，必须从具有更广泛实验基础的爱因斯坦相对性原理出发，建立新的相对论时空观．

习　题

1．间隔不变性：$x'^2 + y'^2 + z'^2 - c^2t'^2 = x^2 + y^2 + z^2 - c^2t^2$ 是相对论时空观的一个基本关系．试证明伽利略变换不能保持此关系成立．

第二章　相对论的时空理论

本章由相对论的基本原理出发建立相对论的时空理论. 首先给出相对论时空理论的一个重要概念——间隔，并导出光速不变的数学表示——间隔不变. 然后，由间隔不变及惯性系的时空均匀性导出洛伦兹变换式，它给出同一事件在不同惯性系之间的时空坐标变换关系. 最后由洛伦兹变换出发得出速度变换公式及同时的相对性、运动时钟延缓、运动尺度缩短等一系列重要推论.

本篇除特别声明外，所提到的参考系均为惯性系.

§2.2.1　间隔　相对论的时空结构

(一)事件和两事件的间隔

物质运动是由许许多多事件构成的，每一个事件可由其发生的地点和时间来表示，记作 (x,y,z,t).

为了方便起见，引入一个虚构的四维空间，并以 x，y,z,t 为它的四个轴. 这样就可以用四维空间的一点表示一个事件. 这种点称为世界点. 一个质点在不同时刻处于不同的空间位置，因而可以用一系列的世界点来描述它的运动. 这些世界点连成一条世界线. 一维运动可以用二维空间的世界线来描述. 如果一个质点一直停留在 x_0 点，它的世界线就是一条平行于 t 轴的直线，如图 1 中的虚线；如果一个质点以速度 v 沿 x 轴作匀速直线运动，它的世界线就是一条斜率为 $\tan\theta = 1/v$ 的直线，如图 1 中的实线.

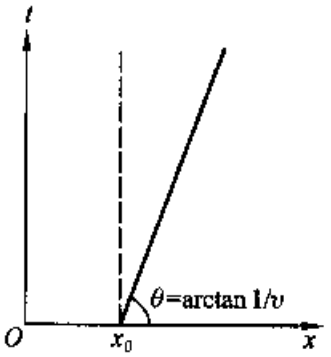


图 1　质点的世界线

如果在某一参考系中，事件 1 用 (x_1,y_1,z_1,t_1) 表示，事件 2 用 (x_2,y_2,z_2,t_2) 表示. 那么 $\Delta r = \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}$ 是这两个事件的空间间隔，$\Delta t = t_2 - t_1$ 是这两个事件的时间间隔，而

$$\Delta s = \sqrt{c^2(\Delta t)^2 - (\Delta r)^2} \tag{1.1}$$

称为这两个事件的间隔. 无穷小的间隔通常用 ds 表示.

(二)间隔不变 —— 光速不变原理的数学表示

现在来证明，光速不变原理的数学表示就是间隔不变，即：两个事件的间隔不因参考系的变换而改变，

$$\Delta s = \Delta s', \tag{1.2}$$

其中 Δs 是在 K 系中观察的事件 1 和事件 2 的间隔，如(1.1)式；$\Delta s'$ 是在 K′系中观察的同样两个事件的间隔：

$$\Delta s' = \sqrt{c^2(\Delta t')^2 - (\Delta r')^2}, \tag{1.3}$$

其中，$\Delta t' = t'_2 - t'_1$，$\Delta r' = \sqrt{(\Delta x')^2 + (\Delta y')^2 + (\Delta z')^2}$.

为了证明(1.2)式，我们先看一个简单例子. 设事件 1 是 t_1 时刻在 x_1, y_1, z_1 处闪光，事件 2 是 t_2 时刻在 x_2, y_2, z_2 处接收到这个闪光. 由光速不变原理可知：

$$\Delta s = \sqrt{c^2(\Delta t)^2 - (\Delta r)^2} = 0, \qquad \Delta s' = \sqrt{c^2(\Delta t')^2 - (\Delta r')^2} = 0.$$

因而 $\Delta s = \Delta s' = 0$. 这就对于这一特例证明了(1.2)式.

一般情况下，任意两个事件，可能是上述由光讯号联系的两事件，可能是用其他讯号联系的两事件，也有可能是没有因果关系的两事件. 在后两种情况下，$\Delta s \neq 0$，$\Delta s' \neq 0$，可以证明它们两者仍然相等. 先证一个引理.

引理 从 (x, y, z, t) 到 (x', y', z', t') 的变换是线性变换.

证 根据惯性系中时间和空间的均匀性. 所有空间上的点和时间上的时刻都是等价的. 因此，两个特定事件的空间间隔和时间间隔的测量结果，与这一测量在参考系的何时何地进行无关.

但是，如果从 (x, y, z, t) 到 (x', y', z', t') 的变换不是线性的，则与上述均匀性的要求相矛盾. 这可以从一个简单的反例看出. 假设 $x' = Ax^2$，A 为常数，则 $x'_2 - x'_1 = A(x_2^2 - x_1^2)$，取一根 1 米长的直尺，两端分别放在 $x_1 = 1$ 和 $x_2 = 2$，则 $x'_2 - x'_1 = A(2^2 - 1^2) = 3A$，若将同一尺子的两端分别放在 $x_1 = 2$ 和 $x_2 = 3$，则 $x'_2 - x'_1 = A(3^2 - 2^2) = 5A$，于是，在 K′系中测量这根直尺的结果，与测量的空间位置有关，违反了空间的均匀性. 关于时间均匀性的要求也有类似的情况. 这就证明了，惯性系的时空均匀性要求惯性系之间的时空变换是线性变换. **【证毕】**

根据这一引理就知道，$(\Delta s')^2 = c^2(\Delta t')^2 - (\Delta r')^2$ 可以写成关于 x, y, z, t 的二次式. 既然 $(\Delta s)^2$ 与 $(\Delta s')^2$ 都是 x, y, z, t 的二次式，Δs 和 $\Delta s'$ 必有线性关系：$\Delta s = \alpha \Delta s' + \beta$.

如果两事件是由光讯号联系，则 Δs 与 $\Delta s'$ 均为零，如上述. 如果两事件是用非常接近光速的讯号相联系，则 Δs 与 $\Delta s'$ 将接近于零. 这两种情况都是 Δs 与 $\Delta s'$ 普遍关系的特例，因此必有 $\beta = 0$；$\Delta s = \alpha \Delta s'$.

由时间空间的均匀性和空间的各向同性可知，上式中的比例系数不能依赖于时空点的坐标，也不能依赖于两参考系相对速度的方向，而只可能依赖于两参考系相对速度的大小 v，因此

$$\Delta s = \alpha(v) \Delta s'.$$

又由相对性原理，两个参考系是等价的，所以也应有

$$\Delta s' = \alpha(v)\Delta s.$$

两式联立，即有 $\alpha^2 = 1$，$\alpha = \pm 1$．这样，α 实际上与 v 也无关．

最后，我们用反证法证明 $\alpha \neq -1$．

证　设有三个惯性系 K，K′，K″，如果 $\alpha = -1$，则有

$$\Delta s = -\Delta s', \qquad \Delta s' = -\Delta s'', \qquad \Delta s'' = -\Delta s.$$

由此得到 $\Delta s = -\Delta s' = \Delta s'' = -\Delta s$，即 $\Delta s = 0$．但是，对于不是由光讯号相联系的事件，$\Delta s \neq 0$，这就证明了 $\alpha \neq -1$．由此知，$\alpha = 1$，$\Delta s = \Delta s'$．**【证毕】**

这样就证明了(1.2)式，即两个事件的间隔在所有惯性系中均相等．这是光速不变原理的数学表述．

(三)间隔的分类

从间隔的定义(1.1)式可见，间隔有三种取值的可能：$\Delta s = 0$，Δs 为实数，Δs 为虚数；因而可以把间隔分成三类．由间隔不变性可知，间隔的划分是绝对的，不因参考系的变换而改变．

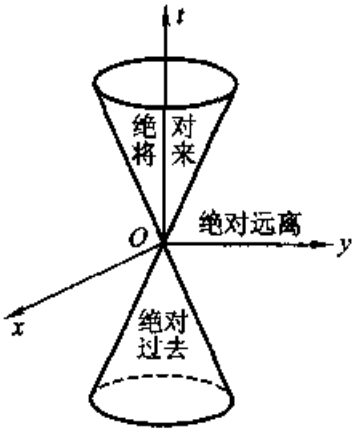


图 2　光锥

为了确定起见，假设在 $t = t' = 0$ 时刻，不同参考系的四维坐标原点重合，并取这一点 O 为第一事件的世界点．设第二事件 P 的时空坐标为 (x,y,z,t)，这样两个事件的间隔 s 是：

$$s^2 = c^2t^2 - (x^2 + y^2 + z^2) = c^2t^2 - r^2.$$

为了直观起见，只画出二维空间和一维时间的时空图，如图 2．事件 P 的世界点在 xy 面上的投影表示事件的发生地点，在 t 轴的投影表示事件发生的时间．于是事件 P 与事件 O 的间隔存在如下三类：

1．$s = 0$，称为零间隔．

如果事件 P 与事件 O 可以用光讯号联系，则 $r = ct$，因而 $s = 0$．

这种事件 P 位于时空图以 O 为顶点的锥面上，锥面母线的斜率 $\tan\theta = 1/c$．由于锥面上的点均可用光讯号与点 O 联系，故称这个锥面为光锥．

2．s 为实数，称为类时间隔．

若 s 为实数，$s^2 = c^2t^2 - r^2 > 0$，则 $r < ct$，即两事件可用低于光速的讯号联系．

例如事件 O 是“开枪”，事件 P 是“鸟落”．讯号传播的速度是子弹的平均速度 $\bar{v}$．于是，OP 线的斜率 $\tan\theta = 1/\bar{v}$ 将大于光锥母线的斜率，所以事件 P 的世界点在光锥之内．

对于在光锥之内的事件，可以找到一个参考系 K′，使事件 P 和事件 O 发生在 K′系的同一点．也就是说，必定存在一个参考系 K′，在其中两事件的空间间隔 $r' =$

0. 这是因为,由 $s'^2 = s^2 > 0$ 有

$$c^2t'^2 - r'^2 = c^2t^2 - r^2 > 0 \tag{1.4}$$

将 $r' = 0$ 代入,得 $c^2t'^2 > 0$,这个不等式是能够成立的. 实际上,选取由开枪地点以子弹的平均速度向鸟落地点运动的参考系 K′,则"开枪"与"鸟落"这两个事件均发生在 K′系的同一点.

但是,不可能找到这样一个参考系,在这个参考系中,这两个事件是同时发生的. 也就是说,不存在参考系 K′,在其中两事件的时间间隔 $t' = 0$. 因为将 $t' = 0$ 代入(1.4)式,得 $-r'^2 > 0$,这个不等式不可能成立.

同样也不可能找到一个参考系,其 $t' < 0$. 也就是说,不可能找到一个参考系,先看见"鸟落",然后才"开枪":"开枪"和"鸟落"这种有因果关系的事件,其因果关系是绝对的,不依赖于参考系而转移. 这一论断不能简单地从(1.4)式得出. 在§2.2.4 中将利用洛伦兹变换给出证明. 因此,位于下半光锥的事件 P 相对子事件 O 称为绝对过去;位于上半光锥的事件 P 相对子事件 O 称为绝对将来.

当然,说光锥内的事件 P 与 O 有可能存在因果关系,并不一定表示它们实际上存在因果关系. 例如"某甲凌晨 0:00 在寝室关灯"是事件 O,"凌晨 0:05 宿舍旁边树上的苹果掉下来"是事件 P. 这两个事件的空间间隔与时间间隔存在 $r < ct$ 的关系(因而 $s^2 > 0$),因此是有可能存在因果关系的,但实际上它们确实又不存在因果关系. 我们这里强调的是"有可能",而不是"一定有".

3. s 为虚数,称为类空间隔.

若 s 为虚数,$s^2 = c^2t^2 - r^2 < 0$,则 $r > ct$,即两事件不可能用光讯号联系. 按照相对论,真空中的光速是物质运动的最大速度,也是一切相互作用的最大传播速度,因而这两事件不可能存在因果关系. 这样的两个事件的时序在不同参考系中可以颠倒.

例如,事件 O 是"某甲凌晨 0:00 在寝室关灯",事件 P 是"凌晨 0:05 有一宇宙飞船经过火星上空". 这两个事件的空间间隔与时间间隔存在 $r > ct$ 的关系(因而 $s^2 < 0$),因此是不可能存在因果关系的. 在四维空间中,如果我们把世界点 O 与 P 连一根直线,既然 $r > ct$,直线的斜率 $\tan\theta$ 将小于 $1/c$,所以这样的事件 P 位于光锥之外.

对于位于光锥之外的事件 P,可以找到一个参考系 K′,使事件 P 与 O 同时发生. 因为 $s'^2 = s^2 < 0$,即

$$c^2t'^2 - r'^2 = c^2t^2 - r^2 < 0, \tag{1.5}$$

将 $t' = 0$ 代入,得到 $-r'^2 < 0$,这个不等式是能够成立的.

但是,我们不能找到一个参考系 K′,使这样的两个事件在同一点发生. 也就是不存在这样的 K′系,在其中两事件的空间间隔 $r' = 0$. 将 $r' = 0$ 代入(1.5)式得到

$c^2t'^2 < 0$，这个不等式不可能成立. 在任何参考系中，这样的两个事件都不可能在空间同一点发生，因而在光锥外的事件 P 相对于事件 O 称为绝对远离.

以上就是零间隔、类时间隔、类空间隔这三类间隔的定义和特征.

§2.2.2　相对论的时空坐标变换公式

这一节从惯性系的时空均匀性和间隔不变性出发，导出同一事件的时空坐标在不同惯性参考系之间的变换关系，即洛伦兹变换.

为确定起见，仍然假设 K 系与 K′系的坐标轴互相平行，K′系以 $\boldsymbol{V} = V\boldsymbol{e}_x$ 相对 K 系作匀速直线运动，在 $t = t' = 0$ 时，K 系和 K′系的坐标原点重合.

为了方便起见，引入新的变量来表示 x, y, z, t：

$$x_1 = x, \quad x_2 = y, \quad x_3 = z, \quad x_4 = ict. \tag{2.1}$$

这样的四维空间的第四个坐标是虚数，称为四维复欧氏空间. 这样，两事件的间隔可表示为

$$\begin{aligned}(\Delta s)^2 &= -c^2(i\Delta t)^2 - [(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2] \\ &= -\sum_{\mu=1}^{4} \Delta x_\mu \Delta x_\mu = -\sum_{\mu=1}^{4} (\Delta x_\mu)^2.\end{aligned}$$

为了书写方便，今后除特别声明外，在每一项中凡有重复指标时，即表示要对重复指标求和，省去求和符号. 如为拉丁字母指标，就表示由 1 到 3 求和；如为希腊字母指标，就表示由 1 到 4 求和. 这称为爱因斯坦惯例. 这样，上式可简写为：

$$(\Delta s)^2 = -(\Delta x_\mu)^2.$$

由此可见，两事件间隔平方的负值 $-(\Delta s)^2$ 可理解为四维复欧氏空间中和这两个事件对应的世界点的距离平方. 于是间隔不变可表示为四维空间距离不变：

$$(\Delta x_\mu)^2 = (\Delta x'_\mu)^2. \tag{2.2}$$

根据惯性系时空均匀性的要求，时空坐标的变换必须是线性的；根据间隔不变性的要求，时空坐标的变换必须保持四维空间的距离不变. 保持距离不变的线性变换就是正交变换. 正交变换包括坐标系统的平移和旋转. 而坐标系统的平移仅相当于同一惯性系时空原点的改变. 因而由惯性系 K 到惯性系 K′的变换在数学上可以表示为四维复欧氏空间中的旋转.

既然 K′系相对 K 系的运动沿 x 方向，在时空坐标变换过程中 $y = y', z = z'$，因此四维复欧氏空间中的旋转又可归结为 x_1x_4 二维复欧氏平面中的旋转.

设转角为 θ，则变换关系为

$$\begin{cases} x'_1 = x_1\cos\theta + x_4\sin\theta, \quad x'_2 = x_2, \quad x'_3 = x_3, \\ x'_4 = -x_1\sin\theta + x_4\cos\theta, \end{cases} \tag{2.3}$$

或者

$$\begin{cases} x_1 = x'_1\cos\theta - x'_4\sin\theta, \quad x_2 = x'_2, \quad x_3 = x'_3, \\ x_4 = x'_1\sin\theta + x'_4\cos\theta. \end{cases} \tag{2.4}$$

这两个式子在形式上与普通欧氏空间的转动相同,但由于目前第 4 轴的坐标取虚值,所以 θ 就不能是实数,而只能是虚数,以使得 $\sin\theta$ 是虚数, $\cos\theta$ 大于 1.❶

纯虚宗量的三角函数就是实宗量的双曲函数. 令 $\theta = i\eta$,则 $\sin\theta = \sin(i\eta) = i\mathrm{sh}\eta$, $\cos\theta = \cos(i\eta) = \mathrm{ch}\eta$,其中,双曲函数 $\mathrm{sh}\eta$, $\mathrm{ch}\eta$ 定义为:

$$\mathrm{sh}\eta = (e^{\eta} - e^{-\eta})/2, \qquad \mathrm{ch}\eta = (e^{\eta} + e^{-\eta})/2.$$

K′系原点的坐标为

$$x'_1 = 0, \qquad x_1 = Vt \tag{2.5}$$

将第一式代入(2.4)式,就有

$$x_1 = -x'_4\sin\theta, \qquad x_4 = x'_4\cos\theta, \tag{2.6}$$

两式相除,利用(2.5)的第二式即可求得 θ 与 V 的关系:

$$\tan\theta = -x_1/x_4 = -Vt/ict = iV/c, \tag{2.7}$$

因而

$$\left.\begin{aligned} &\sin\theta = \tan\theta/\sqrt{1+(\tan\theta)^2} = i\beta/\sqrt{1-\beta^2} = i\beta\gamma, \\ &\cos\theta = 1/\sqrt{1+(\tan\theta)^2} = 1/\sqrt{1-\beta^2} = \gamma. \end{aligned}\right\} \tag{2.8}$$

其中

$$\beta = V/c, \qquad \gamma = 1/\sqrt{1-(V/c)^2} = 1/\sqrt{1-\beta^2}. \tag{2.9}$$

将(2.8)式代入(2.3)式即有:

$$\begin{cases} x'_1 = \gamma(x_1 + i\beta x_4), \quad x'_2 = x_2, \quad x'_3 = x_3, \\ x'_4 = \gamma(-i\beta x_1 + x_4). \end{cases} \tag{2.10}$$

将(2.1)式代入(2.10)式,即得时空坐标的变换式:

$$x' = \gamma(x - Vt), \quad y' = y, \quad z' = z, \quad t' = \gamma\left(t - \frac{V}{c^2}x\right). \tag{2.11}$$

用 x', y', z', t' 表出 x, y, z, t 的反变换式,可以通过(2.8)式代入(2.4)式得到,也可以利用相对性原理求得. 因为 K 系与 K′系是等价的,它们之间的变换应有相同的形式,只不过 K 系相对 K′系的速度 $\boldsymbol{V} = -V\boldsymbol{e}_x$ 而已. 因此,只要在(2.11)式中将 x', y', z', t' 与 x, y, z, t 互换的同时,将 V 换为 $-V$ 即得:

$$x = \gamma(x' + Vt'), \quad y = y', \quad z = z', \quad t = \gamma\left(t' + \frac{V}{c^2}x'\right). \tag{2.12}$$

---

❶ 正弦和余弦可用欧勒公式表示: $\sin\theta = (e^{i\theta} - e^{-i\theta})/2i$, $\cos\theta = (e^{i\theta} + e^{-i\theta})/2$. 由此可见,当 θ 为纯虚数时, $\sin\theta$ 是虚数,而 $\cos\theta > 1$.

(2.11)式和(2.12)式统称为洛伦兹变换式,它给出同一事件在 K 系和 K′系的时空坐标变换关系,也就是两系的观察者在事件发生处对时空坐标的测量结果.

当我们提到某一参考系的"观察者"时,不妨作如下形象的理解.对于每一个参考系,设想在空间的每一点都有它的一个观察者,因而在整个空间中就有无穷多个相对参考系静止的观察者,他们手里各自拿着本参考系已调整到同步的时钟.如果发生了某一事件,K 系和 K′系站在事件发生处的观察者将分别记下该事件相应的时空坐标.这两组测量值即洛伦兹变换式中的数值.这两组测量值是两个参考系的观察者在空间同一点在该事件发生时测量的,所以他们互相公认:"你的读数确实是那样,我的读数也确实是这样,尽管我们的读数不相同".

从洛伦兹变换直接得到如下结论:

(1)在相对论中,光速 c 具有极限速度的特征.因为由(2.9)式可见,当 $V > c$ 时,γ 变为虚数,时空坐标的变换失去意义.这一点与到目前为止的实验事实相符合.

(2)惯性参考系之间的相对速度不单是不能大于 c,也不能等于 c,否则 $\gamma \to \infty$,洛伦兹变换也失去意义.在第三章将会看到,只有静质量为零的粒子才能以光速运动,一个参考系必须有携带着时钟的观察者,静质量不可能为零,所以它实际上是不可能以光速运动的.

(3)当 $V \ll c$ 时,$\gamma \approx 1$,洛伦兹变换(2.11)式过渡到伽利略变换,即 §1.1.4 (4.3)、(4.4)式:$x' = x - Vt, y' = y, z' = z, t' = t$.

§2.2.3　相对论的速度变换公式

在上一节得到了洛伦兹变换式,使我们能从一个事件在某一参考系的时空坐标求得同一事件在其他参考系的时空坐标.本节将从一个物体相对某一参考系的速度求得同一物体相对其他参考系的速度.

设物体在 K 系中的速度为:

$$v_x = \mathrm{d}x/\mathrm{d}t, \quad v_y = \mathrm{d}y/\mathrm{d}t, \quad v_z = \mathrm{d}z/\mathrm{d}t. \tag{3.1}$$

若 K′系以 $\boldsymbol{V} = V\boldsymbol{e}_x$ 相对 K 系运动,则由洛伦兹变换式

$$\begin{cases} x' = \gamma(x - Vt), \quad y' = y, \quad z' = z, \\ t' = \gamma[t - (V/c^2)x]. \end{cases} \tag{3.2}$$

两边取微分,即有

$$\mathrm{d}x' = \gamma(\mathrm{d}x - V\mathrm{d}t), \quad \mathrm{d}y' = \mathrm{d}y, \quad \mathrm{d}z' = \mathrm{d}z, \quad \mathrm{d}t' = \gamma[\mathrm{d}t - (V/c^2)\mathrm{d}x].$$

用第四个方程除前面三个方程,可得

$$\begin{cases} \dfrac{dx'}{dt'} = \dfrac{dx - Vdt}{dt - (V/c^2)dx}, \\ \dfrac{dy'}{dt'} = \dfrac{dy\sqrt{1-(V/c)^2}}{dt - (V/c^2)dx}, \\ \dfrac{dz'}{dt'} = \dfrac{dz\sqrt{1-(V/c)^2}}{dt - (V/c^2)dx}. \end{cases}$$

用 dt 除上述方程右边的分子和分母，即有

$$\begin{cases} v'_x = \dfrac{dx'}{dt'} = \dfrac{v_x - V}{1 - Vv_x/c^2}, \\ v'_y = \dfrac{dy'}{dt'} = \dfrac{v_y\sqrt{1-(V/c)^2}}{1 - Vv_x/c^2}, \\ v'_z = \dfrac{dz'}{dt'} = \dfrac{v_z\sqrt{1-(V/c)^2}}{1 - Vv_x/c^2}. \end{cases} \tag{3.3}$$

这就是相对论的速度变换公式.

像以前对洛伦兹变换作过的那样，在进行 v_x, v_y, v_z 与 v'_x, v'_y, v'_z 互换的同时，将 V 换为 $-V$ 即得反变换式：

$$\begin{cases} v_x = \dfrac{v'_x + V}{1 + Vv'_x/c^2}, \\ v_y = \dfrac{v'_y\sqrt{1-(V/c)^2}}{1 + Vv'_x/c^2}, \\ v_z = \dfrac{v'_z\sqrt{1-(V/c)^2}}{1 + Vv'_x/c^2}. \end{cases} \tag{3.4}$$

在非相对论极限下（$V \ll c, v \ll c$），(3.4)式即过渡到经典的速度变换公式 §1.1.4(4.5)式

$$v'_x \approx v_x - V, \qquad v'_y \approx v_y, \qquad v'_z \approx v_z. \tag{3.5}$$

【例1】 若在 K 系 $v = c$，试证明对任一参考系 K′有 $v' = c$.

证 这里只就粒子在 K 系中的速度 $\boldsymbol{v}$ 和 K′系相对于 K 系的速度 $\boldsymbol{V}$ 平行（或反平行）的情况给出证明. 一般情况下的证明留作习题. 设在 K 系有 $v_x = c$，若 K′系相对 K 系以 $\boldsymbol{V} = V\boldsymbol{e}_x$ 运动，则有：

$$v'_x = \frac{v_x - V}{1 - Vv_x/c^2} = \frac{c - V}{1 - Vc/c^2} = c.$$

如果 K′系相对 K 系以 $\boldsymbol{V} = -V\boldsymbol{e}_x$ 运动，则有：

$$v'_x = \frac{v_x - (-V)}{1 - (-V)v_x/c^2} = \frac{c + V}{1 + Vc/c^2} = c.$$

这表明，无论 K′系相对 K 系沿 x 轴正向或负向运动，v'_x 均为 c. 【证毕】

光速 c 在相对论中的地位就仿佛 ∞ 速度在牛顿力学中的地位一样，因为 ∞ 的速度通过伽利略变换变到任何参考系仍为 ∞ 的速度. 由此可见，c 具有极限速度的特征.

【例 2】 若粒子在 K 系的速度 $v < c$，验证对任一参考系 K′有 $v' < c$.

证 设在 K 系中，粒子在 $\mathrm{d}t$ 时间内移动的距离为 $\mathrm{d}r = \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2}$，将 $\mathrm{d}r = v\mathrm{d}t$ 代入，即有：

$$v^2\mathrm{d}t^2 = \mathrm{d}r^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2.$$

根据间隔不变性：

$$c^2\mathrm{d}t^2 - (\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2) = c^2\mathrm{d}t'^2 - (\mathrm{d}x'^2 + \mathrm{d}y'^2 + \mathrm{d}z'^2),$$

可得

$$(c^2 - v^2)\mathrm{d}t^2 = (c^2 - v'^2)\mathrm{d}t'^2.$$

由 $\mathrm{d}t^2 > 0, \mathrm{d}t'^2 > 0, c^2 - v^2 > 0$，即有 $c^2 - v'^2 > 0$，因而，$v' < c$. 【证毕】

【例 3】 两束电子迎面相对运动，每束电子相对于实验室的速度 $v = 0.9c$，试求：

(a)实验室中观察者观察到的两束电子之间的相对速度.

(b)相对一束电子静止的观察者观察另一束电子的速度.

解 (a)实验室观察者观察到两束电子之间的相对速度为

$$0.9c - (-0.9c) = 1.8c.$$

(b)设实验室参考系为 K′系，沿 x 轴负向运动的电子束的固有参考系为 K 系，则 K′系相对 K 系的速度为 $\boldsymbol{V} = 0.9c\boldsymbol{e}_x$.

沿 x 轴正向运动的电子束相对 K′系的速度 $v'_x = 0.9c$，相对另一束电子束的速度，即相对 K 系的速度为：

$$v_x = \frac{v'_x + V}{1 + v'_xV/c^2} = \frac{0.9c + 0.9c}{1 + (0.9c \times 0.9c)/c^2} = \frac{1.8c}{1 + (0.9)^2} \approx 0.994c.$$

由此可见，由第三者观察两物体之间的相对速度可以超过真空中的光速 c，而两物体本身之间的相对速度则以 c 为极限.

§2.2.4　相对论的时空性质

上两节从时空均匀性和光速不变性原理出发，得到了高速运动情况下的时空坐标变换公式和速度变换公式. 这些公式和低速情况下的相应公式——伽利略变

换公式很不相同.因此,在高速情况下,我们所习惯的关于时空性质的许多观念都需要改变.这一节首先讨论同时的相对性,然后讨论高速运动的时钟和直尺读数的变化.

(一)同时的相对性

在§2.1.2中曾利用光速不变原理证明,当一个雷电击中列车车厢中的A点时,列车参考系测得闪光将同时到达车厢两端的B点和C点,但地面参考系却测得这两个事件不同时,这是同时相对性的一个特例.

一般说来,对于任意两个事件,可以按照§2.2.1的方法来判断它们是否有可能是同时事件.如果事件1和2有类时间隔,则事件2相对于事件1说来,或者是绝对过去,或者是绝对将来(参看§2.2.1图2),在任何参考系中,这两个事件的时序都不能颠倒,也不可能同时发生.但是,如果事件1和2有类空间隔,则当选择不同参考系时,它们可能有任意时序;也有可能选择一个特殊参考系使它们是同时事件.下面利用洛伦兹变换具体证明这一论断.

证 设事件1和事件2有类空间隔

$$\Delta s^2 = c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 < 0. \tag{4.1}$$

按照洛伦兹变换,在K′系中这两个事件发生的时间是

$$t'_1 = \gamma[t_1 - (V/c^2)x_1], \qquad t'_2 = \gamma[t_2 - (V/c^2)x_2]. \tag{4.2}$$

两式相减,得到

$$t'_2 - t'_1 = \gamma[(t_2 - t_1) - (V/c^2)\cdot(x_2 - x_1)]. \tag{4.3}$$

假定在K系中,$t_2 > t_1$,我们来证明,存在参考系K′使两事件同时发生($t'_2 = t'_1$),或按相反时序发生($t'_2 < t'_1$).

将$t'_2 = t'_1$代入(4.3)式,得到

$$V = c\cdot c(t_2 - t_1)/(x_2 - x_1). \tag{4.4}$$

将(4.1)式代入(4.4)式可知$|V| < c$,因而只要令K′系以可实现的速度$\boldsymbol{V} = V\boldsymbol{e}_x$相对于K系运动,就有$t'_2 = t'_1$.注意,按(4.4)式,$V$可能大于零(如$x_2 > x_1$),也可能小于零($x_2 < x_1$),相应地,$V$可能沿$x$轴的正方向或负方向.

再看能否选择K′系使$t'_2 < t'_1$.由(4.1)式得$c(t_2 - t_1) < |x_2 - x_1|$,因而可以选$V$,使

$$c(t_2 - t_1)/|x_2 - x_1| < |V|/c < 1. \tag{4.5}$$

让V和$x_2 - x_1$有相同符号,则无论它们二者是大于零,还是小于零,都有

$$V\cdot(x_2 - x_1) = |V|\cdot|x_2 - x_1| > c^2(t_2 - t_1),$$

代入(4.3)式便可得到$t'_2 < t'_1$.

因此,只要按$x_2 > x_1$或$x_2 < x_1$选择沿x轴正方向或负方向运动的K′系,使其速度的大小满足(4.5)式,则在K′系中,两事件发生的时序改变.比较(4.5)式和

(4.4)式知,使两事件时序改变的参考系,比两事件同时发生的参考系,相对于K系说来,运动速度更快.

注意,如果事件1和2有类时间隔,则(4.1)式中的不等号反转,以上办法得到的 $|V|>c$,因而这样的参考系K′不存在.这就证实了:类时间隔事件的时序有绝对意义,不因参考系而变.　**【证毕】**

(二)运动时钟的延缓

现在利用洛伦兹变换来讨论,在不同参考系测量同一物理过程所经历的时间间隔有什么关系.例如,有一个物体以 $\boldsymbol{V}=V\boldsymbol{e}_x$ 作匀速运动,在 t_1 时它位于 x_1 处,t_2 时位于 x_2 处,因此在实验室参考系K测得这一物理过程所经历的时间间隔为 $\Delta t=t_2-t_1$.如果我们以K′系与这个物体相固联,那么在K′系测得这同一物理过程所经历的时间间隔 $\Delta t'=t'_2-t'_1$ 为若干呢?

在相对论中,通常将与物体相对静止的参考系称为固有参考系(或本征参考系),从而把与物体固联的时钟测出的时间间隔称为固有时间间隔(或本征时间间隔),简称为固有时;而相对物体运动的参考系测出的时间间隔则称为非固有时间间隔.显然,固有时可以认为是在固有参考系同一地点发生的两个事件的时间间隔,也就是同一个时钟测量出来的时间间隔,这个时钟一直与被观察物体固联在一起;而非固有时间间隔是相对物体运动的参考系中位于两个不同地方的两个时钟测得的时间间隔.

利用洛伦兹变换

$$t_1=\gamma[t'_1+(V/c^2)x'_1],\qquad t_2=\gamma[t'_2+(V/c^2)x'_2].$$

两式相减,并注意到K′系与物体相固联,$x'_1=x'_2$,即有 $\Delta t=\gamma\Delta t'$.

通常我们用 $\Delta\tau$ 表示固有时,上式可写为:

$$\Delta t=\gamma\Delta\tau>\Delta\tau. \tag{4.6}$$

因为 $\gamma=1/\sqrt{1-(V/c)^2}>1$.

这说明:一个物理过程所经历的时间间隔是相对的,在物体的固有参考系测得的固有时 $\Delta\tau$ 是最短的.从其他参考系看来,物体运动的速度 v 越大,则 γ 越大,进行同一物理过程所需要的时间 Δt 越长.这种其他参考系认为固有参考系时钟走慢了的效应称为时钟延缓效应.

若时钟不是作匀速直线运动,则在不同时刻应引入不同的与时钟有相同瞬时速度的惯性系,这时运动时钟的读数 $\mathrm{d}\tau=\sqrt{1-v^2/c^2}\,\mathrm{d}t$ 仍成立.当实验室参考系的时钟经历 t_2-t_1 的时间,则运动时钟经历的时间(固有时)为

$$\Delta\tau=\int_{t_1}^{t_2}\sqrt{1-(v^2/c^2)}\,\mathrm{d}t. \tag{4.7}$$

我们曾经提到,1972年西欧核子研究中心观测沿圆形轨道飞行的 μ 子束,对飞行 μ

子寿命进行测量. 已知 μ 子静止时的平均寿命为 $\Delta\tau = 2.197134 \times 10^{-6}$ s，在实验中让 μ 子的速度达到 $v = 0.998c$，测得运动 μ 子的平均寿命为 26.37×10^{-6} s.

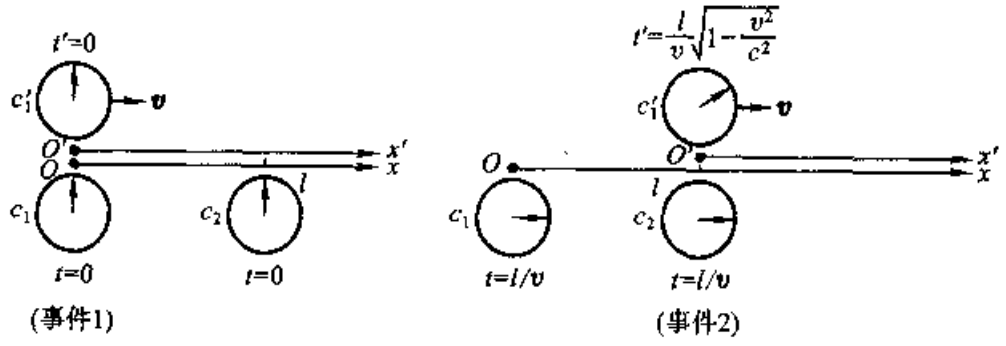
若按运动时钟延缓的公式(4.6)计算，在实验室参考系中 μ 子的平均寿命应为

$$\Delta t = \gamma \Delta\tau = \frac{2.197134 \times 10^{-6}}{\sqrt{1-(0.998)^2}} = 26.69 \times 10^{-6}\text{s},$$

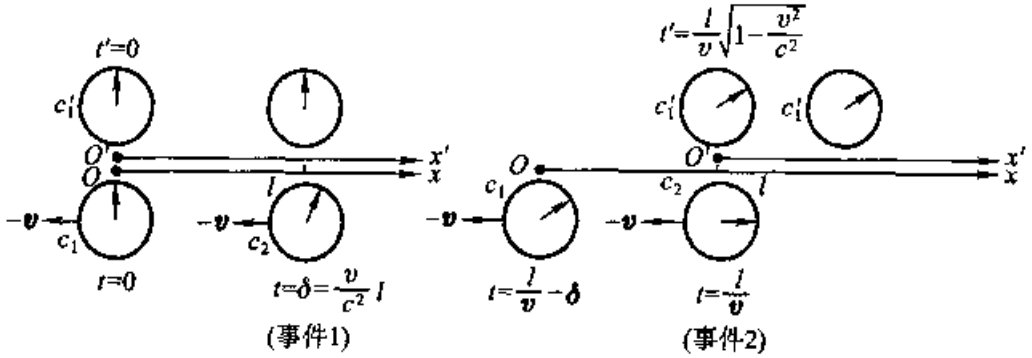
和实验测量结果在 1.1% 精度内符合.

由于运动和静止是相对的，因而时钟延缓效应也是相对的. 下面通过 K 系与 K′系两个互相联系的对钟过程来证明这一结论.

证　对于同一参考系的两个时钟 A 和 B 的校准，可以采用如下的方法：把一个作为信号的光源放在 AB 连线的中点，并要求位于 A，B 处的两个观察者在接到光源发出的信号时将他们的钟同时校准为相同读数. 按照这个方法，即可校准同一参考系的所有时钟.



(a) K系观察者认为K′系时钟延缓



(b) K′系观察者认为K系时钟延缓

图 1　相对运动的两参考系中的对钟过程

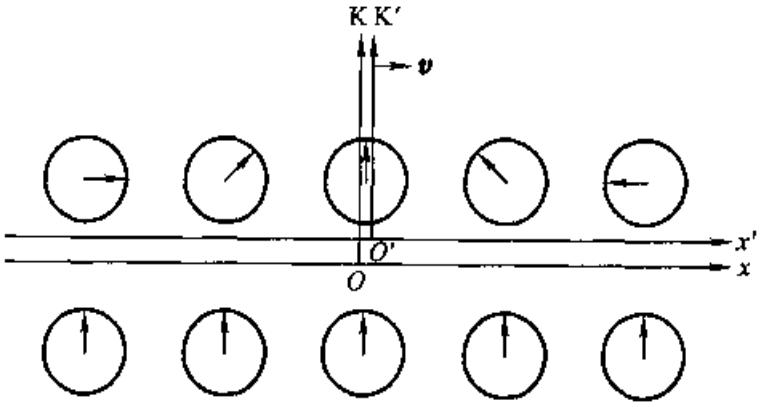
现在，让 K′系的一个钟与 K 系的两个已校准好的钟先后相遇(见图 1)以进行对比. 设事件 1 是 K′系的 c'_1 钟与 K 系的 c_1 钟相遇，K 系与 K′系测得事件 1 的时空坐标都是(0,0)，如图 1(a)左图所示；事件 2 是 K′系的 c'_1 钟与 K 系的 c_2 钟相遇，K 系与 K′系测得事件 2 的时空坐标分别为 (l,t) 及 $(0,t')$，如图 1(a)的右图所示.

由此可见，这两个事件都发生在K′系的钟 c'_1，即发生在K′系的同一点（$x' = 0$）上.

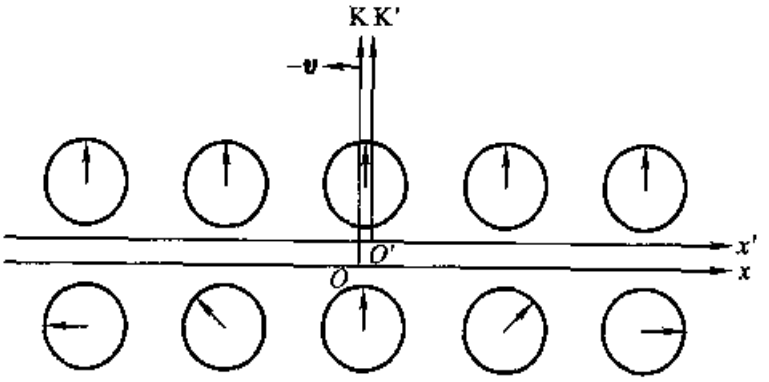
既然K′系相对K系的速度为 $\boldsymbol{v} = v\boldsymbol{e}_x$，故K系测得 c'_1 在 t 时间内走过的距离为 $l = vt$. 由洛伦兹变换即可求得事件2发生时 c'_1 钟的读数为：

$$t' = \gamma[t - (v/c^2)x] = \gamma t(1 - v^2/c^2)$$
$$= (l/v)\sqrt{1 - (v/c)^2} = l/(v\gamma) < l/v. \tag{4.8}$$

K系观察者认为K′系中 c'_1 钟的读数 $t' = l/(v\gamma)$ 小于K系中的时钟读数 $t = l/v$，所以K系观察者认为K′系的时钟变慢.



(a) K系观察者看K′系不同地点的时钟



(b) K′系观察者看K系不同地点的时钟

图2　运动参考系不同地点时钟的读数

但是，K′系的观察者却看到K系的时钟延缓. 这个结论初看起来似乎是不可能的，因为在事件2发生时[见图1(a)右图]，c'_1 钟的读数显然比 c_2 钟的读数小. 这些读数又是两个参考系的观察者所公认的. 但是，必须注意到，站在K′系观察者的立场上来观察同一过程时，必须让K系的一个钟与K′系的两个已校准好的钟先后相遇[见图1(b)]来进行比较. 设事件1是K系的 c_2 钟与K′系的 c'_2 钟相遇，K系和K′系测得事件1的时空坐标分别是 (l,δ) 和 $(l', 0)$，其中 δ 是这一事件发生时 c_2 钟的读数，如图1(b)的左图所示. 事件2是K系的 c_2 钟与K′系的 c'_1 钟相遇，K系和K′系测得事件2的时空坐标分别是 (l,t) 和 $(0,t')$，如图1(a)和1(b)的右图所示. 这里，我们把两个对钟过程的事件2选成相同的，以便说明两者为什么没有矛盾. 在这个对钟过程中，事件1和事件2都发生在K系的钟 c_2 上，即发生在K系的同一点 $(x = l)$ 上.

问题的关键在于，从K′系的观察者看来，K′系的钟已校准好，但由于同时的相对性，K′系观察者认为：K系的时钟沿着 x 的正方向一个比一个超前. 沿着 x 的负

方向一个比一个落后[见图2(b)及习题1]. 因此,在事件1(c'_2 与 c_2 相遇)发生时,c_2 钟的读数并不是零,而是 δ. 至于 δ 为若干,可由洛伦兹变换求出. 将 $t' = 0$ 及 $x = l$ 代入即可求出 c_2 与 c'_2 相遇时 c_2 的读数 δ. 由

$$t' = \gamma[t - (v/c^2)x] = \gamma[\delta - (v/c^2)l] = 0,$$

得 $\delta = vl/c^2$. 而当事件2(c'_1 与 c_2 相遇)发生时,这两个钟的读数如图1a及b的右图,这两个读数同样是由洛伦兹变换确定的. 也是两参考系观察者所公认的.

于是,从K′系观察者看来,从事件1到事件2这个过程中,两系的时钟记录的时间间隔分别是:

$$K' \text{ 系 } c'_2 \text{ 钟:} \qquad \Delta t' = \frac{l}{v}\sqrt{1 - \frac{v^2}{c^2}} - 0 = \frac{l}{v}\sqrt{1 - \frac{v^2}{c^2}}$$

$$K \text{ 系 } c_2 \text{ 钟:} \qquad \Delta t = \frac{l}{v} - \frac{v}{c^2}l = \frac{l}{v}\left(1 - \frac{v^2}{c^2}\right).$$

两式相比,可得

$$\frac{\Delta t'}{\Delta t} = \frac{1}{\sqrt{1 - (v^2/c^2)}} = \gamma > 1. \tag{4.9}$$

即 $\Delta t' > \Delta t$,K' 系观察者看到K系时钟的 Δt 较小,故认为K系的时钟延缓. 这就证明了,对于两个惯性参考系而言,时间延缓的相对性.【证毕】

其实,时钟延缓效应的相对性与"固有时短"的论断是一致的. 因为当K系观察者比较时钟时,是用K′系的一个时钟 c'_1 与K系的两个时钟比较,两事件发生在K′系的同一点上($x' = 0$),故K′系的钟 c'_1 记录的 $\Delta t'$ 是固有时. 由于固有时短,故 $\Delta t' < \Delta t$,因此K系认为K′系的时钟延缓. 而K′系比较时钟时,是用K系的一个时钟 c_2 与K′系的两个时钟比较,两事件发生在K系的同一点上($x = l$),故K系的时钟 c_2 所记录的 Δt 也是固有时. 由于固有时短,$\Delta t < \Delta t'$,因此K′系认为K系的时钟延缓.

(三)运动尺度的缩短

下面利用洛伦兹变换讨论运动物体的长度与该物体的固有长度之间的关系. 所谓固有长度就是在相对物体静止的固有参考系所测得的物体长度.

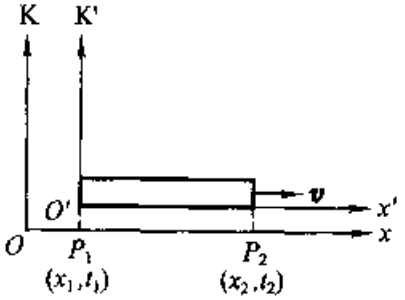


图3　运动物体的长度和固有长度

设物体在K′系静止,K′系以 $\boldsymbol{V} = V\boldsymbol{e}_x$ 相对K系运动. 于是,物体的固有长度为

$$l_0 = x'_2 - x'_1. \tag{4.10}$$

在相对物体运动的K系测得的物体长度,是物体在K系同一时刻两端的距离,见图3. 我们称物体的后端经过 P_1 点为第1事件,记为 (x_1, t_1);物体前端经过

P_2 点为第 2 事件，记为 (x_2, t_2)．由 $t_1 = t_2$，可得运动物体的长度为：

$$l = x_2 - x_1, \quad (\text{当 } t_1 = t_2). \tag{4.11}$$

利用洛伦兹变换：

$$x'_1 = \gamma(x_1 - vt_1), \qquad x'_2 = \gamma(x_2 - vt_2).$$

两式相减，利用 $t_1 = t_2$，即有

$$l_0 = x'_2 - x'_1 = \gamma(x_2 - x_1) = \gamma l. \tag{4.12}$$

因为 $\gamma > 1$，可见物体的固有长度 l_0 最长．在物体相对它以 v 运动的参考系内，物体的长度为

$$l = l_0\sqrt{1 - (v^2/c^2)} = l_0/\gamma. \tag{4.13}$$

这称为运动尺度缩短．

运动尺度缩短效应也是相对的．即 K 系观察到固定在 K′系的物体长度缩短了；同样，K′系观察到固定在 K 系的物体长度也按同一比例缩短了．

证　设物体在 K 系静止，即 K 系是这个物体的固有参考系，物体的固有长度 $l_0 = x_2 - x_1$．在 K′系测得物体的长度为 $l' = x'_2 - x'_1$，当然 x'_1 及 x'_2 是在 $t'_1 = t'_2$ 的条件下测定的．由洛伦兹变换：

$$x_1 = \gamma(x'_1 + vt'_1), \qquad x_2 = \gamma(x'_2 + vt'_2).$$

两式相减，将 $t'_1 = t'_2$ 代入，即有

$$l_0 = x_2 - x_1 = \gamma(x'_2 - x'_1) = \gamma l'.$$

或

$$l' = l_0\sqrt{1 - (v^2/c^2)} = l_0/\gamma. \tag{4.14}$$

即 K′系观察到固定在 K 系的物体长度缩短了．　**【证毕】**

注意，(4.13)式给出 $x_2 - x_1 = (x'_2 - x'_1)/\gamma$，而(4.14)式给出 $x'_2 - x'_1 = (x_2 - x_1)/\gamma$，两式并无矛盾．因为前式是在 K 系同时发生的两事件 $(t_1 = t_2)$，后式是在 K′系同时发生的两事件 $(t'_1 = t'_2)$．

最后，还应指出，运动时钟延缓效应与运动尺度缩短效应是有密切联系的．从下面的例子可以看出这一点．

【例 1】　设在 6000m 高层大气中产生了一个寿命为 $\tau_0 = 2 \times 10^{-6}$ s 的 μ 子．在产生时，该 μ 子在朝地球的方向具有 $v = 0.998c$ 的速度，问它能否越过大气层？试分别按地球参考系 K 及 μ 子的固有参考系 K′进行讨论．

解　(a)在 K 系中测定，由子运动时钟延缓效应，测得 μ 子的寿命

$$\tau = \gamma\tau_0 = \frac{2 \times 10^{-6}}{\sqrt{1 - (0.998)^2}} = 31.6 \times 10^{-6}\text{s}.$$

在 K 系测得 μ 子能飞越的距离为：

$$s = v\tau = 0.998 \times 3 \times 10^8 \times 31.6 \times 10^{-6} = 9470\text{m},$$

而大气层的厚度为 $l = 6000$ m，故 $s > l$，因而 μ 子能穿越大气层.

(b)在 K′系中测定，μ 子能飞越的距离为

$$s' = v\tau' = 0.998 \times 3 \times 10^8 \times 2 \times 10^{-6} = 598.8\text{m},$$

由于运动尺度缩短，K′系观察者测得大气层的厚度为

$$l' = l\sqrt{1-(v/c)^2} = 6000\sqrt{1-(0.998)^2} = 378\text{m},$$

由于 $s' > l'$，故 μ 子能穿越大气层.

由此可见，μ 子能穿越大气层这一客观事实，在 K 系描述为 μ 子的寿命延长；在 K′系描述为大气层厚度变薄，两者得出的结论一致.

* §2.2.5　相对论时空理论的若干应用

本节应用洛伦兹变换和速度变换讨论几个历史上有名的实验.

(一)菲索流水实验

实验装置如图 1. 由光源 S 发出的光被半透明反射镜 M 分为两束. 第一束光以顺时针方向经 $MM_1M_2M_3MT$ 进入望远镜 T；第二束光以逆时针方向经 $MM_3M_2M_1MT$ 进入望远镜 T. 图中管道充满水. 问当水静止时及流动时两束光产生的干涉条纹有何变化？已知水流速度为 V，$AB = CD = l$，水的折射率为 n.

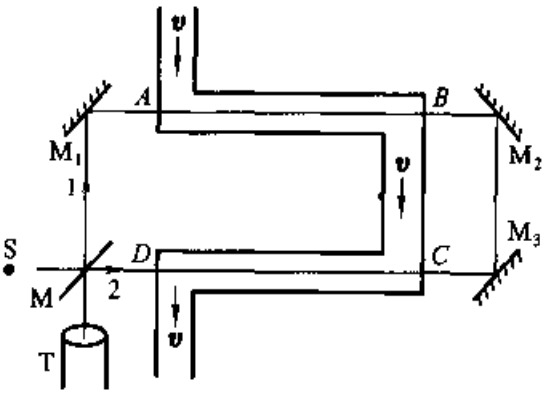


图 1　菲索流水实验示意图

为简单起见，我们仅讨论 AB 段. 因为在光的传播方向与水流方向的关系上，CD 段与 AB 段相同，所以关于 AB 段的结论也适用于 CD 段. 设实验室参考系为 K，水流的固有参考系为 K'(指 AB 段). 故 K'系相对 K 系以 $\boldsymbol{V} = V\boldsymbol{e}_x$ 运动. 在 AB 段，光束 1 以 $\boldsymbol{v} = \boldsymbol{e}_x c/n$ 相对 K'系传播，由速度变换公式可知，光束 1 相对 K 系的传播速度为

$$v_x = \frac{v'_x + V}{1 + Vv'_x/c^2} = \frac{(c/n) + V}{1 + (V/nc)} \approx \left(\frac{c}{n} + V\right)\left(1 - \frac{V}{nc}\right) \approx \frac{c}{n} + V\left(1 - \frac{1}{n^2}\right),$$

即顺水的光束 1 在 AB 段和 CD 段相对 K 系的光速大小均为

$$v^{(1)} = \frac{c}{n} + V\left(1 - \frac{1}{n^2}\right). \tag{5.1}$$

在 AB 段，光束 2 以 $\boldsymbol{v}' = -\boldsymbol{e}_x c/n$ 相对 K'系传播，由速度变换公式可知，光束相对 K 系的传播速度为

$$v_x = \frac{v'_x + V}{1 + Vv'_x/c^2} = \frac{-(c/n) + V}{1 - (V/nc)} \approx -\frac{c}{n} + V\left(1 - \frac{1}{n^2}\right),$$

即逆水的光束 2 在 AB 段和 CD 段相对 K 系光速的大小均为(对 v 取绝对值)：

$$v^{(2)} = \frac{c}{n} - V\left(1 - \frac{1}{n^2}\right). \tag{5.2}$$

进入望远镜的两束光的相位差 $\Delta\varphi$ 是由于两束光通过 AB 段和 CD 段所需时间不同而引起的. 因此，当管道中的水从静止到以 V 流动所引起的干涉条纹的移动数 ΔN 为

$$\Delta N = \frac{\Delta\varphi}{2\pi} = \frac{\Delta t}{T} = \frac{1}{T}\left(\frac{2l}{v^{(2)}} - \frac{2l}{v^{(1)}}\right) \approx \frac{4ln^2V(1 - 1/n^2)}{c\lambda}. \tag{5.3}$$

计算结果与实验结果相符.

(二)光行差现象

18 世纪初，天文学家布莱德雷发现，恒星在一年的不同时间里，其表观位置有变化，这称为光行差现象. 下面对这一现象给出解释.

光行差现象 是由于地球绕太阳公转运动具有 $V = 30\ \mathrm{km \cdot s^{-1}}$ 的速度引起的.

设太阳参考系为 K，与地球瞬时保持相对静止的惯性参考系为 K′. 当地球在公转轨道上的不同位置时，相应的 K′系是不同的. 现在我们仅讨论地球在某一特定位置时的情形. 选地球相对 K 系的速度方向为 x 轴方向，远处恒星射来的星光可看作平行光，取星光所在平面为 xoy 平面、设 θ 角和 θ' 角分别为星光对 K 系和 K′系的 x 轴的夹角，因而

$$\begin{cases} v_x = c\cos\theta \\ v_y = c\sin\theta \end{cases} \qquad \begin{cases} v'_x = c\cos\theta' \\ v'_y = c\sin\theta' \end{cases}.$$

利用相对论的速度变换公式即可求得星光相对太阳的方向 θ(称"真实方向")与星光相对地球的方向 θ'(称"表观方向")的关系. 由

$$v_x = \frac{v'_x + V}{1 + Vv'_x/c^2}, \qquad v_y = \frac{v'_y\sqrt{1 - (V/c)^2}}{1 + Vv'_x/c^2}.$$

即可得

$$\tan\theta = \frac{v_y}{v_x} = \frac{v'_y\sqrt{1 - (V/c)^2}}{v'_x + V} = \frac{\sin\theta'\sqrt{1 - (V/c)^2}}{\cos\theta' + V/c} \approx \frac{\sin\theta'}{\cos\theta' + V/c}. \tag{5.4}$$

上面利用了 $V \ll c$，$\sqrt{1 - (V/c)^2} \approx 1$ 的近似. 按照三角学公式，由上式得到

$$\sin\theta = \frac{\tan\theta}{\sqrt{1 + \tan^2\theta}} \approx \frac{\sin\theta'}{1 + V\cos\theta'/c}, \tag{5.5}$$

$$\cos\theta = \frac{1}{\sqrt{1 + \tan^2\theta}} \approx \frac{\cos\theta' + V/c}{1 + V\cos\theta'/c}, \tag{5.6}$$

故

$$\Delta\theta = \theta' - \theta \approx \sin(\theta' - \theta) \approx \sin\theta'\cos\theta - \cos\theta'\sin\theta \approx V\sin\theta'/c.$$

由于 θ' 是星光与地球公转运动方向(K′系的 x' 轴方向)的夹角，在地球公转过程中，地球公转运动的方向不断变化，恒星的表观位置也就不断变化，因而星光的表观方向不断变化. 其变化周期恰好是一年. 上述理论与光行差现象的观测结果完全一致.

(三)多普勒效应

声音的多普勒效应是常见的. 迎面而来的火车拉响汽笛，我们感觉汽笛声音变尖了，即频率

升高了. 这是因为当汽笛接近我们的时候,它在 1s 内发出的声波数将在不到 1s 的时间内进入我们的耳朵. 既然在比较短的时间内传来比较多的波数,声音就变尖了. 相反,当火车在我们背后离去时,汽笛声变得低沉了. 这是由于在相同时间内传到我们耳内的声音的波数减少了的缘故.

利用洛伦兹变换很容易证明光也有多普勒效应.

设在实验室参考系 K 的坐标原点 o 发出光沿 x 方向传播. 光振动的相位为

$$\varphi = \omega t - kx = 2\pi\nu(t - x/c). \tag{5.7}$$

如果观察者(K′系)以 $V = Ve_x$ 相对 K 系运动,将洛伦兹变换式代入:

$$\varphi' = 2\pi\nu\left[\gamma\left(t' + \frac{v}{c^2}x'\right) - \frac{1}{c}\gamma(x' + vt')\right]$$

$$= 2\pi\nu\gamma\left(1 - \frac{v}{c}\right)\left(t' - \frac{x'}{c}\right) = 2\pi\nu\sqrt{\frac{1-\beta}{1+\beta}}\left(t' - \frac{x'}{c}\right). \tag{5.8}$$

将上式与 K′系的相位表示式 $\varphi' = 2\pi\nu'(t' - x'/c)$ 相比较,即有:

$$\nu' = \sqrt{(1-\beta)/(1+\beta)}\,\nu. \tag{5.9}$$

可见,当观察者远离光源运动时,他所接收到的光频率变低.

同样,如果观察者(K′系)以 $V = -Ve_x$ 迎着光源运动时,可以将上式中的 v 换为 $-v$,β 换为 $-\beta$,即有

$$\nu' = \sqrt{(1+\beta)/(1-\beta)}\,\nu. \tag{5.10}$$

即观察者接受到的光频率增高. 这就是光的多普勒效应.

§2.2.6 四维协变量及方程的协变性

在 §2.2.2 中曾把惯性系之间的时空坐标变换形式上看作四维空间的转动,由此导出了洛伦兹变换式:

$$x' = \gamma(x - Vt), \quad y' = y, \quad z' = z, \quad t' = \gamma\left(t - \frac{V}{c^2}x\right), \tag{6.1}$$

$$x = \gamma(x' + Vt'), \quad y = y', \quad z = z', \quad t = \gamma\left(t' + \frac{V}{c^2}x'\right). \tag{6.2}$$

后来,利用洛伦兹变换及由此得到的相对论速度变换公式解释了高速运动以及电磁波传播中的一系列现象.

本节将首先引入四维位矢 $x_\mu(\mu = 1,2,3,4)$ 这样一个统一的物理量来表示一个事件的时空坐标,并利用变换矩阵 $A = (a_{\mu\nu})$ 将洛伦兹变换写成更为简洁的形式. 在讨论变换矩阵 A 的性质之后,根据物理量在四维空间转动下的变换性质,将物理量分成四维各阶张量,统称为四维协变量.

四维协变量的引入,不仅使同一物理量在不同惯性系的变换关系更为简洁;而且,以下将发现,除了时间与空间坐标之外,频率与波矢、能量与动量、标势与矢势、电荷密度与电流密度、电场强度与磁感强度等等均各自组成一个统一的四维协变

量，从而使我们更深刻地认识到它们之间的内在联系；最后，使得有可能把所有物理规律写成符合相对性原理要求的形式，即本节末尾将要介绍的方程的协变性.

首先，引入变换矩阵

$$A = (a_{\mu\nu}) = \begin{pmatrix} \gamma & 0 & 0 & i\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -i\beta\gamma & 0 & 0 & \gamma \end{pmatrix}. \tag{6.3}$$

将洛伦兹变换(6.1)写成矩阵形式：

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \\ x'_4 \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & i\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -i\beta\gamma & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}. \tag{6.4}$$

用矩阵元写出就是

$$x'_\mu = a_{\mu\nu} x_\nu, \tag{6.5}$$

这就是洛伦兹变换的四维形式.

现在讨论变换矩阵元 $a_{\mu\nu}$ 及变换矩阵 A 的性质.

由间隔不变性可知：$x'_\mu x'_\mu = x_\mu x_\mu$，将(6.5)式代入，即有：

$$a_{\mu\rho} x_\rho a_{\mu\tau} x_\tau = x_\mu x_\mu. \tag{6.6}$$

利用克罗内克 δ 符号，可将上式写成 $a_{\mu\rho} a_{\mu\tau} x_\rho x_\tau = \delta_{\rho\tau} x_\rho x_\tau$，故

$$a_{\mu\rho} a_{\mu\tau} = \delta_{\rho\tau}. \tag{6.7}$$

定义 $\tilde{a}_{\mu\nu}$ 为转置矩阵 $\tilde{A}$ 的矩阵元：

$$\tilde{a}_{\mu\nu} = a_{\nu\mu}$$

可将(6.7)式写成：$\tilde{a}_{\rho\mu} a_{\mu\tau} = \delta_{\rho\tau}$，用矩阵符号写出就是：

$$\tilde{A}A = \mathbf{1}. \tag{6.8}$$

其中 $\tilde{A}$ 是变换矩阵 A 的转置矩阵，$\mathbf{1}$ 为单位矩阵：

$$\mathbf{1} = (\delta_{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{6.9}$$

(6.7)、(6.8)式称为正交变换条件，即保持距离不变的线性变换(四维空间中的转动)所满足的充分必要条件. 参看 s 维空间中保持矢量长度不变的正交变换所满足的条件 §1.4.3(3.26)式及该章习题 10.

现在再回到变换式(6.5). 我们知道，x_1, x_2, x_3, x_4 这四个量作为一个事件的时空坐标，可以看成是四维空间的一个矢量的四个分量，这个矢量就称为四维位矢.

(6.5)式表明,洛伦兹变换可以看成是四维空间中的转动[参看 §2.2.2(2.3)、(2.4)式].

同样,如果有四个量 B_1,B_2,B_3,B_4 在洛伦兹变换(即四维空间转动)时,其变换规则与 x_μ 的变换规则相同

$$B'_\mu = a_{\mu\nu}B_\nu, \tag{6.10}$$

则 B_μ 也称为四维矢量.

具有两个指标的量 $A_{\mu\tau}$,如果在四维空间转动(洛伦兹变换)时,它的两个指标分别遵守和四维位矢相同的变换关系

$$A'_{\mu\nu} = a_{\mu\rho}a_{\nu\tau}A_{\rho\tau}, \tag{6.11}$$

则称它为四维二阶张量.同样,遵守变换关系

$$A'_{\mu\nu\lambda} = a_{\mu\rho}a_{\nu\tau}a_{\lambda\sigma}A_{\rho\tau\sigma}, \tag{6.12}$$

的量 $A_{\mu\nu\lambda}$ 称为四维三阶张量.高阶张量的定义可以类推.

这样,我们就可以按照物理量在四维空间转动下的变换性质将物理量分为标量(不变量或零阶张量),矢量(一阶张量),二阶张量及其他高阶张量.由于这些物理量在四维空间转动时,每个指标都具有和四维位矢相同的变换性质,所以称为四维协变量.

不难证明,两个四维矢量 B_μ 与 D_μ 的标积是标量(不变量),因为:

$$B'_\mu D'_\mu = a_{\mu\rho}B_\rho a_{\mu\tau}D_\tau = \delta_{\rho\tau}B_\rho D_\tau = B_\rho D_\rho.$$

这种将两个矢量(更一般说是张量)各有一个指标取成相同,并对它求和的运算称为指标的缩并.

根据一个式子含有多少个自由指标,就可确定它是多少阶张量❶,如 $B_\mu B_\mu$ 只有哑指标(求和指标),没有自由指标,故为零阶张量,即标量.又如 $A_{\mu\nu}B_\nu$ 有一个自由指标 μ,因而是一阶张量,即矢量,余类推.

现在我们列举一些四维协变量.

间隔　$ds^2 = c^2dt^2 - (dx^2 + dy^2 + dz^2) = -dx_\mu dx_\mu$ 没有自由指标,是标量.

固有时　$d\tau = ds/c$ 也是标量,因为 c 是常数,而 ds 是标量.

四维速度矢量　定义为:

$$u_\mu = dx_\mu/d\tau, \tag{6.13}$$

由于 dx_μ 是四维矢量,$d\tau$ 是标量,所以 u_μ 是四维矢量.由定义出发,不难求得 u_μ 的各分量与三维矢量的关系.

设物体相对 K 系的速度为 $v = dx/dt$,则

$$d\tau = ds/c = \sqrt{c^2dt^2 - dx_i^2}/c = dt\sqrt{1 - v^2/c^2}, \tag{6.14}$$

---

❶ 这是因为我们所讨论的物理量都是张量,所讨论的运算都是张量运算(求和,缩并等等).

代入(6.13)式,并采用如下记号

$$u_\mu = \frac{1}{\sqrt{1-(v/c)^2}}\left(\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}, ic\right) = \gamma_v(\boldsymbol{v}, ic). \tag{6.15}$$

上式圆括号中逗号前的三维矢量 $\gamma_v\boldsymbol{v}$ 是 u_μ 的空间分量;逗号后的分量 $\gamma_v ic$ 为 u_μ 的时间分量.今后对于四维矢量都将采用这样的简单记号.注意,(6.15)式中的 $\gamma_v = 1/\sqrt{1-(v/c)^2}$ 不同于 $\gamma = 1/\sqrt{1-(V/c)^2}$.前者的 v 是物体相对 K 系的速度的大小;后者的 V 是 K′系相对 K 系的速度的大小.

从(6.15)式可以看出,如果物体的速度远小于光速,$\gamma_v \approx 1$,则物体四维速度的空间分量即为物体的三维速度 $\boldsymbol{v}$,其时间分量的模是真空中的光速.

四维速度矢量在不同惯性系之间的变换关系为:

$$u'_\mu = a_{\mu\nu}u_\nu, \tag{6.16}$$

由此可以导出三维速度的相对论变换式,见习题 9.

四维波矢量的定义可根据电磁波的相位是不变量的性质引入.关于电磁波的相位是不变量的定量证明,将在 §2.4.4 中讨论电磁场变换时给出.这里我们只给一个定性的描述.

证　在 K 系和 K′系研究一系列事件,设事件 0 是单色平面电磁波

$$\begin{cases}\boldsymbol{E} = \boldsymbol{E}_0\cos(\omega t - \boldsymbol{k}\cdot\boldsymbol{x}),\\ \boldsymbol{B} = \boldsymbol{B}_0\cos(\omega t - \boldsymbol{k}\cdot\boldsymbol{x}).\end{cases} \tag{6.17}$$

在 $t = 0$ 时 $\boldsymbol{x} = 0$ 处出现极大值,其相位 $\varphi = \omega t - \boldsymbol{k}\cdot\boldsymbol{x} = 0$,而事件 1,2,…,$n$,… 分别是 $t_1 = T, t_2 = 2T, \cdots, t_n = nT, \cdots$ 时在 $\boldsymbol{x} = 0$ 处 $\boldsymbol{E}$ 和 $\boldsymbol{B}$ 出现极大值,其位相分别为 $\varphi = 2\pi, 4\pi, \cdots, n\pi, \cdots$.显然,在事件 0,1,2,… 发生时,K′系在事件发生处的观察者也观察到 $\boldsymbol{E}$ 及 $\boldsymbol{B}$ 达到极大值,即相应的 $\varphi' = 2\pi, 4\pi, \cdots, n\pi, \cdots$.尽管 K′系测得的极大值,$\boldsymbol{E}'_0$、$\boldsymbol{B}'_0$ 和 $\boldsymbol{E}_0$、$\boldsymbol{B}_0$ 不同,这一系列事件的时空坐标 x', t' 与 x, t 也不同,但是,在这一系列事件发生时的位相 $\varphi' = \varphi = 2\pi, 4\pi, \cdots, n\pi, \cdots$ 却是一致的.对于 $\boldsymbol{E}, \boldsymbol{B}$ 变化过程中任意的 φ 值也可以类似地讨论.　　**【证毕】**

既然 $\varphi = \varphi'$(即 $\omega t - \boldsymbol{k}\cdot\boldsymbol{x} = \omega' t' - \boldsymbol{k}'\cdot\boldsymbol{x}'$)是洛伦兹变换(即四维空间转动变换)的不变量,而四维位矢 $(\boldsymbol{x}, ict)$ 是四维空间的矢量,可见 $(\boldsymbol{k}, i\omega/c)$ 也组成一个四维矢量,称为四维波矢量

$$k_\mu = (\boldsymbol{k}, i\omega/c). \tag{6.18}$$

四维波矢和四维位矢的标积 $k_\mu x_\mu = k'_\mu x'_\mu$ 在四维空间转动下不变,也就是相位在洛伦兹变换下不变:$\omega t - \boldsymbol{k}\cdot\boldsymbol{x} = \omega' t' - \boldsymbol{k}'\cdot\boldsymbol{x}'$.四维波矢量的空间分量就是电磁波的三维波矢,时间分量实际上决定于电磁波的角频率 ω,因为 $1/c$ 是普适常数.

四维波矢量的变换为:

$$k'_\mu = a_{\mu\nu}k_\nu, \tag{6.19}$$

写成矩阵形式,即为

$$\begin{pmatrix} k'_1 \\ k'_2 \\ k'_3 \\ \mathrm{i}\omega'/c \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & \mathrm{i}\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\mathrm{i}\beta\gamma & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{pmatrix} = \begin{pmatrix} \gamma(k_1 + \mathrm{i}\beta \cdot \mathrm{i}\omega/c) \\ k_2 \\ k_3 \\ \gamma(-\mathrm{i}\beta k_1 + \mathrm{i}\omega/c) \end{pmatrix}.$$

由相应的矩阵元相等可得

$$\begin{aligned} &k'_1 = \gamma[k_1 - (V/c^2)\omega], \quad k'_2 = k_2, \quad k'_3 = k_3, \\ &\omega' = \gamma(\omega - Vk_1). \end{aligned} \tag{6.20}$$

逆变换式为:

$$\begin{aligned} &k_1 = \gamma[k'_1 + (V/c^2)\omega'], \quad k_2 = k'_2, \quad k_3 = k'_3, \\ &\omega = \gamma(\omega' + Vk'_1). \end{aligned} \tag{6.21}$$

利用四维波矢量的变换可以解释多普勒效应及光行差现象.

首先讨论多普勒效应. 设实验室参考系为 K,光源的固有参考系 K′以 $\boldsymbol{V} = V\boldsymbol{e}_x$ 相对 K 系运动,因而 $\omega' = \omega_0$ 是光源的固有角频率. 设 θ 和 θ' 分别为 $\boldsymbol{k}$ 和 $\boldsymbol{k}'$ 与 $\boldsymbol{v}$ 的夹角,亦即与 $\boldsymbol{e}_x$ 的夹角,因此

$$\begin{cases} k_1 = k\cos\theta = = (\omega/c)\cos\theta, \\ k'_1 = k'\cos\theta' = (\omega_0/c)\cos\theta', \end{cases} \tag{6.22}$$

代入(6.20)式的第四式,即有 $\omega_0 = \gamma[\omega - (v\omega/c)\cos\theta]$,或

$$\omega = \frac{\omega_0}{\gamma(1 - \beta\cos\theta)} = \frac{\omega_0\sqrt{1-\beta^2}}{1 - \beta\cos\theta}, \tag{6.23}$$

我们分别讨论 $\theta = 0, \pi/2, \pi$ 这几种情况:

(a) $\boldsymbol{k}$ 与 $\boldsymbol{v}$ 同向,$\theta = 0$;$\omega = \omega_0\sqrt{(1+\beta)/(1-\beta)} > \omega_0$,即当观察者迎着光源的运动方向观察时,测得光波的角频率大于其固有角频率 ω_0.

(b) $\boldsymbol{k}$ 与 $\boldsymbol{v}$ 反向,$\theta = \pi$;$\omega = \omega_0\sqrt{(1-\beta)/(1+\beta)} < \omega_0$,即当观察者背着光源的运动方向观察时,测得光波的角频率小于其固有角频率 ω_0. 这两种情况都和 §2.2.5 的结果一致,而推导更简洁.

(c) $\boldsymbol{k}$ 与 $\boldsymbol{v}$ 垂直,$\theta = \pi/2$;$\omega = \omega_0\sqrt{1-\beta^2} < \omega_0$,即当观察者垂直于光源的运动方向观察时,测得光波的角频率小于其固有角频率. 这称为横向多普勒效应,也与实验相符.

利用四维波矢量的变换同样可以解释光行差现象. 仍然取太阳参考系为 K,瞬时与地球相对静止的惯性参考系为 K′. 在该瞬时,K′系相对 K 系的速度方向为 x

轴方向，星光选在 xoy 平面，θ 及 θ' 分别为星光相对 K 系 x 轴及 K′系 x' 轴的夹角.

由(6.22)，(6.21)式

$$\cos\theta = \frac{c}{\omega}k_1 = \frac{c\gamma[k'_1 + (v/c^2)\omega']}{\gamma(\omega' + vk'_1)} = \frac{\cos\theta' + \beta}{1 + \beta\cos\theta'},$$

略去 β^2 项，即可求得

$$\sin\theta = \sqrt{1 - \cos^2\theta} = \sin\theta'/(1 + \beta\cos\theta').$$

于是 $\Delta\theta = \theta' - \theta \approx \sin(\theta' - \theta) = \beta\sin\theta'$，与上节利用相对论速度变换公式所得结果一样.

最后，介绍一下方程协变性的含义.

相对性原理指出，所有惯性系是等价的，物理规律在所有惯性系中是相同的. 因此，衡量一个方程是否满足相对性原理的要求，就要看这个方程在不同惯性系能否表示成相同形式，而这决定于方程的每一项是否属于同类协变量. 也就是说，如果方程的每一项都是标量，或者都是四维矢量，或者都是同阶的四维张量，那么在洛伦兹变换下，各项均按同一规律变化，因而就能保证方程的形式不变. 例如，方程

$$A_\mu = B_\mu + C_\mu,$$

当由 K 系变到 K′系时，

$$A'_\mu = a_{\mu\nu}A_\nu = a_{\mu\nu}(B_\nu + C_\nu) = a_{\mu\nu}B_\nu + a_{\mu\nu}C_\nu = B'_\mu + C'_\mu.$$

可见，在 K′系的方程 $A'_\mu = B'_\mu + C'_\mu$ 与在 K 系的方程具有相同的形式. 这种在参考系变换时方程保持形式不变的性质称为方程的协变性.

物理规律应满足相对论的要求，因而所有的方程都应具有洛伦兹变换下的协变性. 由于这一原因，所有表达物理规律的方程都应能写成张量之间的等式的形式.

习　题

1. 设 K′系相对 K 系有速度 $V = Ve_x$，证明从 K 系看 K′系不同位置的时钟，沿 x 轴正方向一个比一个落后，沿 x 轴负方向一个比一个超前；反之，从 K′系看 K 系不同位置的时钟，沿 x 轴正方向一个比一个超前，沿 x 轴负方向一个比一个落后，如 §2.2.4 图 2.

2. 设有两根互相平行的尺，其固有长度均为 l_0. 它们以相同的速率 v 相对于某一参考系运动，但运动方向相反，且平行于尺子. 求站在一根尺上测量另一尺的长度.

3. 固有长度为 l_0 的车厢，以速度 v 相对地面运行，车厢的后壁以速度 v_0 向前推出一个小球，求地面观察者看到小球从后壁到前壁的运动时间.

4. 一辆以速度 v 运动的列车(K′系)在经过某一高大建筑物时，建筑物的避雷针跳起一脉冲电火花，电光迅速传播，照亮了铁路线的两铁塔，求 K′系观察者测得两铁塔被电光照亮的时间差. 已知两铁塔分别位于建筑物的前方和后方，三者在一直线上，与列车前进方向一致. 铁塔至

建筑物的地面距离已知都是 l_0.

5. 一列固有长度为 l_0 的列车,以恒定速度 v 驶过固有长度也是 l_0 并与之平行的站台.列车观察者已将车头 b 和车尾 a 上两个时钟校准好;站台观察者已将站台起点 A 和终点 B 上两个时钟校准好.当车头的时钟 b 与站台起点的时钟 A 相遇时,设 b 和 A 两处的观察者测得这两个时钟的读数均为 t_0 ,问站台的观察者看来:

(a)在该时刻是否所有的时钟均指示 t_0 ? (b)当时钟 a,A 相遇时,各时钟的读数是多少? (c)当时钟 b,B 相遇时,各时钟的读数是多少? (d)当时钟 a,B 相遇时,各时钟的读数是多少?

列车上的观察者看来,结果是否相同?

6. 列车 CD (K′系)沿铁路以匀速 v 相对地面(K 系)运动,地面观察者测得列车 CD 与隧道 AB 等长,当列车进入隧道时:

(a)若 K 系观察者测得 CA 相遇时,两个雷电同时打在隧道出口 A 与入口 B 处,问 K′系观察者看到什么现象?

(b)若 K′系观察者测得 CA 相遇时,两个雷电同时打在隧道出口 A 与入口 B 处,问 K 系和 K′系分别观察到什么现象?

7. 有一光源 S 与接收器 R 相对静止,距离为 l_0,S-R 装置浸在无限的液体介质(静止折射率为 n)中,试对下列三种情况计算光源发出讯号到接收器接收到讯号所经历的时间.

(a)液体介质相对 S-R 装置静止.

(b)液体沿着 S-R 连线方向以速度 v 流动.

(c)液体垂直于 S-R 连线方向以速度 v 流动.

8. 一把直尺相对 K 系静止,直尺与 x 轴交角为 θ ,今有一观察者以速度 v 沿 x 轴运动,问他测得直尺 与 x 轴交角 θ' 有何变化?

9. 具体写出四维速度变换公式 §2.2.6(6.16)式的各分量式.导出三维速度的变换公式.

提示 第四分量式 $u_4 = \alpha_{4\nu}u_\nu$ 可化为 $\gamma_{v'} = \gamma\gamma_v(1 - Vv_1/c^2)$.

10. 粒子在 K 系中的速度大小为 $v = c$,方向任意.若 K′系相对 K 系以速度 $\boldsymbol{V} = v\boldsymbol{e}_x$ 运动,求证在 K′系中也有 $v' = c$.

11. 考虑沿某一方向的一维运动.用下式定义"快度" y :

$$\beta = v/c = \tanh y.$$

其中, $\tanh y = (e^y - e^{-y})/(e^y + e^{-y})$ 是双曲正切函数.

(a)证明 $y = \dfrac{1}{2}\ln\dfrac{1+\beta}{1-\beta}$.

(b)设 K′系相对于 K 系以速度 $v = c\cdot\tanh y_1$ 运动,K″系相对于 K′系以速度 $v' = c\cdot\tanh y$ 运动.试证明 K″系相对于 K 系的快度是 $y_2 = y_1 + y$.

12. 一个位于 yoz 平面的平面镜以 $\boldsymbol{V} = V\boldsymbol{e}_x$ 相对实验室 K 系运动,一束频率为 ω_0 与水平面成 θ_0 夹角的平面光波自右向左入射到镜面上,求反射光波的频率 ω 及反射角 θ.

第三章　高速粒子的运动规律

很容易证明，牛顿定律不符合相对论协变性的要求. 经典的牛顿定律只适用于低速运动的质点和质点系，对于高速运动的粒于不成立. 例如，一个初速为零，质量为 m 的粒子在恒力的作用下，按照经典的牛顿定律，这个粒子将有恒定的加速度. 这样，当物体受力的时间足够长，物体的速度必将超过真空中的光速 c，这就与相对论关于“物体运动速度以光速 c 为极限”的论断相矛盾，也是与实验事实相违背的. 因此，必须建立符合相对论基本原理的动力学方程.

本章首先建立相对论力学基本方程，它满足相对论协变性的要求，并且当物体的速度 $v \ll c$ 时，过渡到经典的牛顿定律. 其次，将讨论相对论的质量－能量守恒定律和动量守恒定律，并运用它们讨论粒子分裂和结合的问题.

本章第一、二节讨论高速自由粒于，第三节讨论在电磁场中运动的高速带电粒于.

§2.3.1　相对论力学的基本方程

我们知道，在经典力学里，动量是物体的质量和速度的乘积. 那么，在相对论中怎样定义动量呢？从 §2.2.6 的讨论知道，为了符合相对论的要求，需要把动量扩展成四维协变量. 由于在 $v \ll c$ 时，相对论力学基本方程过渡到经典的牛顿定律，因而四维协变物理量也应当能合理地过渡到相应的经典力学量. 正如我们已介绍过的，四维速度

$$u_\mu = \gamma_v(\boldsymbol{v}, ic) = \left(\frac{\boldsymbol{v}}{\sqrt{1-(v/c)^2}}, \frac{ic}{\sqrt{1-(v/c)^2}}\right), \tag{1.1}$$

在 $v \ll c$ 时，它的空间分量便过渡到物体的经典速度 $\boldsymbol{v}$. 因此，可以这样定义四维动量：

$$p_\mu = m_0 u_\mu. \tag{1.2}$$

其中 m_0 是物体的静质量. m_0 是一个标量，而标量与四维矢量的乘积仍为四维矢量，这样便保证了按(1.2)式定义的 p_μ 是四维矢量.

p_μ 可以写成

$$p_\mu = \left(\frac{m_0\boldsymbol{v}}{\sqrt{1-(v/c)^2}}, \frac{i}{c}\frac{m_0c^2}{\sqrt{1-(v/c)^2}}\right) = \left(\boldsymbol{p}, \frac{i}{c}E\right), \tag{1.3}$$

它的空间分量

$$\boldsymbol{p} = \frac{m_0\boldsymbol{v}}{\sqrt{1-(v/c)^2}} = m\boldsymbol{v}, \tag{1.4}$$

称为物体的相对论动量，它在 $v \ll c$ 时过渡到经典动量 $m_0\boldsymbol{v}$. (1.4)式右方的

$$m = m_0/\sqrt{1-(v/c)^2}, \tag{1.5}$$

称为物体的相对论质量，又称为动质量. 显然，在不同惯性系测得物体的速度 $\boldsymbol{v}$ 不同，因而物体的动质量也就不同. 在 $v \ll c$ 时，m 过渡到经典质量 m_0. 在本章中，将简称动质量为“质量”.

p_μ 的时间分量中的

$$E = \frac{m_0c^2}{\sqrt{1-(v/c)^2}} = mc^2 \tag{1.6}$$

具有能量的量纲，它的物理意义在建立了相对论力学基本方程以后再回过头来讨论.

现在我们引入四维力矢量

$$K_\mu = (\boldsymbol{K}, K_4), \tag{1.7}$$

将牛顿定律 $\boldsymbol{F} = \mathrm{d}\boldsymbol{p}/\mathrm{d}t$ 推广到四维形式：

$$K_\mu = \mathrm{d}p_\mu/\mathrm{d}\tau, \tag{1.8}$$

这就是相对论力学基本方程. 因为 $\mathrm{d}\tau$ 是标量，所以 $\mathrm{d}p_\mu/\mathrm{d}\tau$ 是四维矢量，于是方程(1.8)式两边均为四维矢量，从而保证了这一方程的相对论协变性.

下面来看，要使(1.8)式的前三个方程在 $v \ll c$ 时能过渡到经典的牛顿定律，四维力 K_μ 应该是怎样的.

利用(1.3)式和 §2.2.6(6.14)式可将四维方程(1.8)式写成

$$K_\mu = \left(\gamma_v\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t}, \mathrm{i}\frac{\gamma_v}{c}\frac{\mathrm{d}E}{\mathrm{d}t}\right). \tag{1.9}$$

我们要求它的前三个方程具有 $\boldsymbol{F} = \mathrm{d}\boldsymbol{p}/\mathrm{d}t$ 的形式，其中 $\boldsymbol{F}$ 为三维力. 为此，令 K_μ 的空间分量 $\boldsymbol{K}$ 与三维力 $\boldsymbol{F}$ 有如下关系

$$\boldsymbol{K} = \gamma_v\boldsymbol{F} = \boldsymbol{F}/\sqrt{1-(v/c)^2}. \tag{1.10}$$

当 $v \ll c$ 时，$\boldsymbol{p}$ 过渡到 $m_0\boldsymbol{v}$，方程(1.8)的前三个方程过渡到经典的牛顿定律. 也就是说，只要 K_μ 是空间分量由(1.10)式规定的一个四维矢量，方程(1.8)就既满足协变性的要求，也满足经典过渡的要求.

现在，我们来讨论 p_μ 和 K_μ 的时间分量以及(1.8)式的第四个方程的物理意义.

为了了解 p_4 的含义，取(1.6)式对 t 的导数：

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{m_0}{[1-(v/c)^2]^{\frac{3}{2}}}\boldsymbol{v}\cdot\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}$$

$$= \frac{m_0}{\sqrt{1-(v/c)^2}}\left[1+\left(\frac{v}{c}\right)^2\frac{1}{1-(v/c)^2}\right]\boldsymbol{v}\cdot\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}$$

$$= \boldsymbol{v}\cdot\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{m_0\boldsymbol{v}}{\sqrt{1-(v/c)^2}}\right] = \boldsymbol{v}\cdot\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = \boldsymbol{F}\cdot\boldsymbol{v}. \tag{1.11}$$

因为 $\boldsymbol{F}\cdot\boldsymbol{v}$ 是外力对物体作功的功率，它应等于物体的总能量的增加率. 这样，(1.11)式表明 E 就是物体的总能. 由(1.3)式可见，p_μ 的空间分量是物体的相对论动量 $\boldsymbol{p}$，时间分量是 i/c 乘以物体的总能量 E. 因此，四维动量 p_μ 又称为能量－动量四维矢量.

将(1.11)式代入(1.9)式，即得 K_μ 的时间分量为

$$K_4 = (\mathrm{i}/c)\gamma_v\boldsymbol{F}\cdot\boldsymbol{v}, \tag{1.12}$$

将(1.10)式和(1.12)式代入(1.7)式，可将四维力 K_μ 写成

$$K_\mu = \gamma_v(\boldsymbol{F}, \mathrm{i}\boldsymbol{F}\cdot\boldsymbol{v}/c), \tag{1.13}$$

即四维力 K_μ 的空间分量是 γ_v 与力 $\boldsymbol{F}$ 的乘积，时间分量是 $\mathrm{i}\gamma_v/c$ 与外力对物体所作功率 $\boldsymbol{F}\cdot\boldsymbol{v}$ 的乘积. 方程(1.8)的第四个方程实际上就是动能定理(1.11)式.

在静质量守恒的情况下，(1.8)式还可以写成下述形式❶

$$K_\mu = \frac{\mathrm{d}(m_0u_\mu)}{\mathrm{d}\tau} = m_0\frac{\mathrm{d}u_\mu}{\mathrm{d}\tau} = m_0w_\mu, \tag{1.14}$$

式中

$$w_\mu = \frac{\mathrm{d}u_\mu}{\mathrm{d}\tau} = \left(\frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\boldsymbol{v}}{\sqrt{1-(v/c)^2}}, \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\mathrm{i}c}{\sqrt{1-(v/c)^2}}\right) \tag{1.15}$$

称为四维加速度. 它的平方 $w_\mu w_\mu$ 没有自由指标，所以

$$w_\mu w_\mu = \text{不变量}. \tag{1.16}$$

【例】 在实验室参考系 K 中有一个粒子在 $t=0$ 时从坐标原点出发沿 x 轴作初速为零的变速直线运动. 在粒子瞬时静止的惯性系 K_0 测得粒子的加速度 $\boldsymbol{w}_0 = w_0\boldsymbol{e}_x$ 为恒定值.

(a)问在实验室参考系 K 测得粒子的 $v(t)=?, x(t)=?$

(b)问当 $t\to\infty$ 时，K 系测得粒子的 $v=?$

解 (a)在 K_0 系粒子的四维加速度为 $w_\mu^{(0)} = (w_0,0,0,0)$，由此得 $w_\mu^{(0)2} = w_0^2$. 在 K 系中

$$w_1 = \frac{1}{\sqrt{1-(v/c)^2}}\frac{\mathrm{d}}{\mathrm{d}t}\frac{v}{\sqrt{1-(v/c)^2}} = \frac{1}{(1-v^2/c^2)^2}\frac{\mathrm{d}v}{\mathrm{d}t},$$

❶ 在静质量守恒的情况下，m_0 才能提出微商号外. 注意，“m_0 是不变量”是指在参考系变换时它保持不变，而“m_0 是守恒量”是指 m_0 不随时间变化. 我们在下节将看到，m_0 不一定是守恒的.

$$w_2 = w_3 = 0.$$

$$w_4 = \frac{ic}{\sqrt{1-(v/c)^2}}\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{\sqrt{1-(v/c)^2}} = \frac{iv}{c(1-v^2/c^2)^2}\frac{\mathrm{d}v}{\mathrm{d}t}.$$

因而 $w_\mu^2 = w_1^2 + w_4^2$. 根据(1.16)式 $w_\mu^{(0)2} = w_\mu^2$, 因而

$$w_0 = \sqrt{w_\mu^2} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{v}{\sqrt{1-(v/c)^2}}.$$

由此得到 $v = w_0 t\sqrt{1-(v/c)^2}$. 解之, 得

$$v = w_0 t/\sqrt{1+(w_0 t/c)^2}. \tag{1.17}$$

再积分一次, 得到

$$x = \int v\mathrm{d}t = \int \frac{w_0 t}{\sqrt{1+(w_0 t/c)^2}}\mathrm{d}t = \frac{c^2}{w_0}\sqrt{1+\left(\frac{w_0 t}{c}\right)^2} + 常数.$$

由 $t = 0$ 时 $x = 0$ 可知, 积分常数 $= -c^2/w_0$, 故

$$x = c^2(\sqrt{1+(w_0 t/c)^2} - 1)/w_0.$$

(b) 当 $t \to \infty$ 时, 用 t 除(1.17)式右边的分子分母, 可得

$$\lim_{t\to\infty} v(t) = \lim_{t\to\infty}\frac{w_0}{\sqrt{(1/t)^2+(w_0/c)^2}} = c.$$

可见, 即使在恒力的作用下, 无论受力时间有多长, 物体的速度仍以光速 c 为其极限.

至此, 已阐明了 p_μ, K_μ 的各分量的物理意义. 相对论中的动量 $\boldsymbol{p}$, 能量 E, 力 $\boldsymbol{F}$, 功率 $\boldsymbol{F}\cdot\boldsymbol{v}$ 在不同参考系的变换关系可以由 $p'_\mu = a_{\mu\nu}p_\nu$ 及 $K'_\mu = a_{\mu\nu}K_\nu$ 求得. 在这里, 用得着 $\gamma = 1/\sqrt{1-(V/c)^2}$, $\gamma_v = 1/\sqrt{1-(v/c)^2}$ 和 $\gamma_{v'} = 1/\sqrt{1-(v'/c)^2}$ 之间的关系:

$$\gamma_{v'} = \gamma\gamma_v(1 - Vv_x/c^2). \tag{1.18}$$

现在来考察物体的总能量 E 与静能 E_0 和动能 T 的关系. 由物体的总能公式(1.6)可见. 当物体的速度 $v = 0$ 时, 物体的总能变为

$$E_0 = m_0 c^2. \tag{1.19}$$

这称为物体的静能. 对于微观粒子, 例如电子, 静能 E_0 (或和它成正比的静止质量 m_0)是它的固有性质. 而对于一个宏观尺度的物体内部, 存在着分子无规则运动的动能, 分子间的势能, 原子内部乃至核内的势能等等, 它们共同组成物体的静能.

物体的动能就是物体的总能与静能之差

$$T = E - E_0 = mc^2 - m_0 c^2 = \frac{m_0 c^2}{\sqrt{1-(v/c)^2}} - m_0 c^2, \tag{1.20}$$

在低速情况下，$v \ll c$，展开上式的根式，并保留到二阶小量，即有：

$$T = m_0c^2\left(1 + \frac{1}{2}\frac{v^2}{c^2} + \cdots\right) - m_0c^2 \approx \frac{1}{2}m_0v^2.$$

这正是经典力学中物体动能的公式.

最后，我们建立质能关系式. 因为 $p_\mu p_\mu$ 没有自由指标，是不变量，利用和物体有相同瞬时速度的惯性系 K_0 及任意参考系 K 的 $p_\mu p_\mu$ 值相等，即有：

$$(\mathrm{i}E_0/c)^2 = \boldsymbol{p}\cdot\boldsymbol{p} + (\mathrm{i}E/c)^2.$$

这里利用了物体在 K_0 系中的动量为零. 再将 $E_0 = m_0c^2$ 代入，即可求得物体在 K 系的总能 E，动量 $\boldsymbol{p}$ 与物体的静质量 m_0 的关系：

$$E^2 = \boldsymbol{p}^2c^2 + m_0^2c^4, \tag{1.21}$$

这就是著名的质能关系式. 它也可以改写为

$$p_\mu p_\mu = (\mathrm{i}E/c)^2 + \boldsymbol{p}^2 = -m_0^2c^2. \tag{1.21a}$$

这一等式在 4 维复欧氏空间中是一个半径为 $\mathrm{i}m_0c$ 的“球壳”，因而常称为质壳条件.

顺便指出上面几个公式的关系：

①公式 $E_0 = m_0c^2$ 是公式 $E = \sqrt{\boldsymbol{p}^2c^2 + m_0^2c^4}$ 的特例，因为物体在其静止参考系的动量为零.

②上面利用 $p_\mu p_\mu$ = 不变量 导出(1.21)式，实际上，(1.21)式还可直接由(1.6)式求得.

③将 $\boldsymbol{p} = m\boldsymbol{v}$ 和 $E = mc^2$ 相除可得

$$\boldsymbol{p} = E\boldsymbol{v}/c^2. \tag{1.22}$$

在以上的讨论中都假定粒子的静质量 $m_0 \neq 0$. 前面已经举了很多例子说明这种粒子的运动速度一定小于光速，其原因还可以从(1.6)式看出. 因为，如果 $m_0 \neq 0$ 而 $v = c$，则将有 $E = \infty$，而这是不可能的. 但是，对子光子或其他以光速运动的粒子(例如中微子)，速度 $v = c$. 此时由(1.6)式知，为使 E 不为无穷大，只有 $m_0 = 0$；反之，如果 $m_0 = 0$，为使 $E \neq 0$ 必须 $v = c$. 因此，

$$v = c \quad \Leftrightarrow \quad m_0 = 0. \tag{1.23}$$

对于这种粒子，由(1.22)式得到

$$p = E/c \qquad (\text{当 } m_0 = 0). \tag{1.24}$$

在本节里，用比较直观的方法，从四维速度得到四维动量(1.3)式，进而得到运动方程(1.8)式. 更为严格的推导，应该以最小作用量原理为出发点. 首先根据作用量是洛伦兹变换的不变量这一要求，写出自由粒子的作用量

$$S = -m_0c\int_a^b \mathrm{d}s, \tag{1.25}$$

其中，ds 是间隔，m_0 是一个反映粒子性质的常数. 然后，利用最小作用量原理得到自由粒子的能量、动量和运动方程，见习题 3. 在外场中运动的粒子的作用量将在下一篇中讨论.

§2.3.2 高速运动粒子的守恒定律

设封闭系在 $t = 0$ 时由 N 个粒子组成，E_{a0}，$\boldsymbol{p}_{a0}$ 及 E_a，$\boldsymbol{p}_a$ 分别表示第 a 个粒子在 $t = 0$ 及 t 时的能量和动量. 封闭系的能量守恒定律与动量守恒定律可表示为

$$\sum_{a=1}^{N} E_{a0} = \sum_{a=1}^{N'} E_a, \qquad \sum_{a=1}^{N} \boldsymbol{p}_{a0} = \sum_{a=1}^{N'} \boldsymbol{p}_a. \tag{2.1}$$

注意符号 E_{a0} 和 E_{0a} 的区别，前者是 $t = 0$ 时第 a 个粒子的能量，而后者是第 a 个粒子的静能. 另外，$t = 0$ 及 t 时粒子数 N 和 N' 可以不同.

下面运用能量守恒定律讨论结合能问题. 为简单起见，我们讨论一个静质量为 m_0，速度为 $\boldsymbol{v}$ 的物体，由于内部的原因分裂为两个静质量分别为 m_{01} 和 m_{02}，速度为 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的物体. 在这个分裂过程中，物体的动质量、静质量、静能是否守恒呢？

在这个问题中，能量守恒可写为

$$\frac{m_0 c^2}{\sqrt{1 - (v/c)^2}} = \sum_{a=1}^{2} \frac{m_{0a} c^2}{\sqrt{1 - (v_a/c)^2}}. \tag{2.2}$$

两边除以 c^2，由 §2.3.1 的(1.5)式可见动质量守恒：

$$m = \sum_{a=1}^{2} m_a. \tag{2.3}$$

这说明，在相对论中能量守恒与质量守恒是一致的. (2.3)式与(2.2)式统称为质量－能量守恒定律.

从(2.2)式还可以看到：由于 $v = v_1 = v_2$ 一般不成立，所以静质量一般不守恒. 即一般说来

$$m_0 \neq m_{01} + m_{02}. \tag{2.4}$$

两边乘以 c^2，可见静能一般也不守恒

$$m_0 c^2 \neq m_{01} c^2 + m_{02} c^2. \tag{2.5}$$

为了更直观地理解物体分裂过程中的能量转化过程，将(2.1)式的总能 E 写成静能 E_0 和动能 T 之和：

$$m_0 c^2 + T = (m_{01} + m_{02}) c^2 + (T_1 + T_2). \tag{2.6}$$

如果 $m_0 c^2 > (m_{01} + m_{02}) c^2$，则 $(T_1 + T_2) > T$. 这表示当物体分裂后，静能较原来减少了，而动能增加了. 也就是说，物体 m 的一部分静能释放出来转化为 m_1，m_2 的动能.

物体的分裂是否会释放能量决定于 m_0c^2 与 $(m_{01}+m_{02})c^2$ 的大小，我们称

$$\Delta m_0 = m_0 - (m_{01}+m_{02}) \tag{2.7}$$

为质量亏损，它是当物体分裂时减少的静质量，而称

$$\Delta E_0 = (m_{01}+m_{02})c^2 - m_0c^2 = -\Delta m_0c^2 \tag{2.8}$$

为结合能，它是 m_1 与 m_2 结合为 m 时放出的能量.

如前所述，如果 $m_0c^2 > (m_{01}+m_{02})c^2$，则 $\Delta m_0 > 0$，$\Delta E_0 < 0$，质量亏损为正，物体分裂后静能减少，一部分静能被释放出来转变为两个分裂物的动能. 这意味着物体能自动分裂.

如果 $(m_{01}+m_{02})c^2 > m_0c^2$，则 $\Delta m_0 < 0$，$\Delta E_0 > 0$，即质量亏损为负. 这意味着分裂后系统的静能增加了. 因此，必须从外界供给物体以足够的能量(至少要等于结合能)才有可能使物体分裂. 结合的过程刚好与分裂的过程相反，所以，此时结合过程将会释放能量.

相对论的这一结论预言了利用原子能的可能性. 例如可以通过原于弹的裂变和氢弹的聚变获得能量. 实验表明，重核(如铀核)的质子和中子结合得不像中等原于核(原子序数为 100 左右)那么紧，因此重核的每个核子的平均结合能要比中等原子核小. 这样，当核子结合成重核时释放出来的能量 $\sum m_{0a}c^2 - m_0c^2$，要小于这些核子结合成两个中等原子核 x 和 y 时释放出来的能量 $\sum m_{0a}c^2 - (m_{0x}+m_{0y})c^2$，这里 m_{0a} 为第 a 个核子的静质量. m_0 为重核的静质量，m_{0x} 和 m_{0y} 分别为两个中等原子核的静质量. 也就是说，存在着不等式：

$$0 < \sum m_{0a}c^2 - m_0c^2 < \sum m_{0a}c^2 - (m_{0x}+m_{0y})c^2$$

即

$$\sum m_{0a}c^2 > m_0c^2 > (m_{0x}+m_{0y})c^2.$$

由此可见，当铀核裂变为两个中等原子核 x 和 y 时，将要释放出 $m_0c^2 - (m_{0x}+m_{0y})c^2$ 的能量. 聚变的情况刚好相反. 聚变是由结合能较小的氘核和氚核通过聚合过程结合为结合能比较大的氦核而释放出能量.

【例 1】氦核是由两个质于和两个中子组成. 已知：

氢原子的静质量 $=4.002603$amu，

氢原子的静质量 $=1.007825$amu，

中子的静质量 $=1.008665$amu.

求两个质子和两个中子结合为氦核时放出的结合能. (amu 为原子质量单位，见附录Ⅶ).

解　质量亏损为

$$\Delta m_0 = 4.002603 - 2 \times (1.007825 + 1.008665) = -0.030377\text{amu}.$$

当两个质子与两个中子结合为氦核时放出的能量即为(参看附录Ⅶ)

$$\Delta E_0 = -\Delta m_0 c^2 = 0.030377 \times 931.5\text{MeV} = 28.30\text{MeV}.$$

【例2】　已知 π 介子的静质量 $m_{0\pi} = 139.57\text{MeV}/c^2$，$\mu$ 子的静质量 $m_{01} = 105.66\text{MeV}/c^2$，中微子的静质量 $m_{02} = 0$．当一个带电的 π 介子衰变为 μ 子和中微子时，求 π 介子质心系中 μ 子的动量、能量和速度．

解　在 π 介子质心系 K 中，衰变前 π 介子的动量和能量分别为：

$$\boldsymbol{p}_\pi = 0, \qquad E_\pi = E_{0\pi} = m_{0\pi}c^2.$$

衰变后 μ 子和中微子的的动量和能量分别为：

$$\mu\text{子：} \qquad \boldsymbol{p}_1, \qquad E_1 = \sqrt{p_1^2c^2 + m_{01}^2c^4},$$

$$\text{中微子：} \qquad \boldsymbol{p}_2, \qquad E_2 = \sqrt{p_2^2c^2 + 0} = p_2c$$

由动量守恒定律：$\boldsymbol{p}_1 + \boldsymbol{p}_2 = \boldsymbol{p}_\pi = 0$，故 $|\boldsymbol{p}_1| = |\boldsymbol{p}_2| = p$．

由能量守恒定律，$m_{0\pi}c^2 = \sqrt{p^2c^2 + m_{01}^2c^4} + pc$，由此解出：

$$p = c(m_{0\pi}^2 - m_{01}^2)/2m_{0\pi} = 29.79\text{MeV}/c,$$

$$E_1 = E_\pi - E_2 = m_{0\pi}c^2 - pc = c^2(m_{0\pi}^2 + m_{01}^2)/2m_{0\pi} = 109.78\text{MeV}.$$

将 §2.3.1(1.22)式应用于 μ 子，即有 $p_1/E_1 = v_1/c^2$，由此得 μ 子的速度：

$$v_1 = p_1c^2/E_1 = 29.29c/109.78 = 0.271c.$$

【例3】　一个封闭的开水瓶中盛有 1kg 开水，求当水从 100℃ 冷却到 0℃ 时静质量的变化．

解　水由于热量的散失引起的总能变化为：

$$\Delta E_0 = -1 \times 10^3 \times 100 = -10^5\text{cal} = -4.2 \times 10^5\text{J},$$

由此引起水的静质量的变化为：

$$\Delta m_0 = \Delta E_0/c^2 = -4.2 \times 10^{12}/9 \times 10^{20} = -4.67 \times 10^{-12}\text{kg}.$$

这个例子说明，水的静质量变化是原来的静质量的 4.67×10^{-12} 倍，所以在日常生活中是不可能觉察到的．

值得注意的是，在经典力学中，质量守恒定律和能量守恒定律是互相独立的两个定律．从经典的观点看，封闭开水瓶中的水的质量守恒，而能量不守恒．相对论则指出，质量与能量存在着 $E = mc^2$ 的关系．质量和能量的变化是同步的．曾经有人认为，水温从 100℃ 下降到 0℃，水的质量减少了，而外界物体获得了能量，就认为是水的质量转化为外界物体的能量，从而得出“物质可以转化为能量”的错误结论．其实，在水的质量和能量都减少的同时，外界物体的能量和质量都增加了．这只不过是质量和能量从一个物体转移到另一个物体，或由一种形式转化为另一种形式而已．又如，由电子和正电子组成的“电子对”在一定条件下可以转化为一对高

能光子. 这同样不能认为电子的质量转化为光子的能量，而应该看到一方面电子对的动能和静能全部转化为光子对的能量，同时电子对的质量也全部转化为光子对的质量. 总之，一个封闭系统遵守统一的质量-能量守恒定律.

§2.3.3　电磁场中的带电粒子

(一)电磁场中带电粒子的作用量

在上两节中，我们讨论了高速运动的自由粒子. 现在来讨论在外场中的高速运动粒子.

首先要注意的是，当自由粒子以高速运动时，能量(乘上 i/c)是一个四维矢量的第四分量，见 §2.3.1(1.3)式. 类似地，高速粒子在外场中的势能也是一个四维矢量的第四分量. 具体考虑电荷为 e 的带电粒子，它在外场中的势能可写为 $U = e\phi$，其中 ϕ 是电势. 在带电粒子以高速运动的情况下，电势应扩展成一个四维矢量 A_μ：

$$A_\mu = (\boldsymbol{A}, i\phi/c). \tag{3.1}$$

其中，$\boldsymbol{A}$ 称为电磁场的三维矢势，而 A_μ 则是四维矢势. 在这里，我们是根据相对论的基本原理，由电势 ϕ 是四维矢量的第四分量，就知道一定存在一个相应的三维矢势 $\boldsymbol{A}$. 它的物理意义将在(3.10)式、(3.12)式和 §2.4.4、§3.2.1 中讨论.

A_μ 和带电粒子的四维位矢 $dx_\mu = (d\boldsymbol{r}, icdt)$ 的点积 $A_\mu dx_\mu$ 是标量. 利用它可以将带电粒子在电磁场中的作用量写成

$$S = \int_a^b (-m_0 c ds + eA_\mu dx_\mu), \tag{3.2}$$

其中，第一项是自由粒子的作用量 §2.3.1(1.25)式；第二项反映场与带电粒子的相互作用. A_μ 描述电磁场的特征，edx_μ 描述带电粒子的特征.

将 $ds = cdt\sqrt{1 - v^2/c^2}$ [§2.2.6(6.14)式]和(3.1)式一道代入(3.2)式得到

$$S = \int_a^b (-m_0 c^2\sqrt{1 - v^2/c^2} + e\boldsymbol{A}\cdot\boldsymbol{v} - e\phi)dt. \tag{3.3}$$

于是，带电粒子在电磁场中的拉格朗日量是

$$L = -m_0 c^2\sqrt{1 - v^2/c^2} + e\boldsymbol{A}\cdot\boldsymbol{v} - e\phi. \tag{3.4}$$

按 §1.2.1(1.3)和(1.15)式，带电粒子在电磁场中的广义动量和能量是

$$\boldsymbol{p}_{粒场} = \frac{\partial L}{\partial \boldsymbol{v}} = \frac{m_0\boldsymbol{v}}{\sqrt{1 - v^2/c^2}} + e\boldsymbol{A} = \boldsymbol{p} + e\boldsymbol{A}. \tag{3.5}$$

$$E_{粒场} = \frac{\partial L}{\partial \boldsymbol{v}}\cdot\boldsymbol{v} - L = \frac{m_0 c^2}{\sqrt{1 - v^2/c^2}} + e\phi = E + e\phi. \tag{3.6}$$

其中 $\boldsymbol{p}$ 和 E 分别是无外场时带电粒子的动量和能量[见 §2.3.1(1.4)(1.6)式]. (3.5)式和(3.6)式表明,只要在自由带电粒子的表达式中作代换

$$\begin{aligned} &\boldsymbol{p} \to \boldsymbol{p} + e\boldsymbol{A}, \qquad E \to E + e\phi, \quad \text{国际单位制} \\ &\boldsymbol{p} \to \boldsymbol{p} + (e/c)\boldsymbol{A}, \quad E \to E + e\phi, \quad \text{高斯单位制} \end{aligned} \tag{3.7}$$

就得到和电磁场相互作用的带电为 e 的粒子的相应表达式. 有时称这为最小电磁相互作用. 关于国际单位制和高斯单位制,见附录Ⅵ.

(二)电磁场中带电粒子的运动方程

运动方程由最小作用量原理得到[§1.1.3(3.12)式]:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \boldsymbol{v}} = \frac{\partial L}{\partial \boldsymbol{r}}. \tag{3.8}$$

其中拉格朗日量 L 由(3.4)式给出.

(3.8)式左边的偏导 $\partial L/\partial \boldsymbol{v}$ 就是带电粒子在电磁场中的广义动量 $\boldsymbol{p}_{\text{粒场}}$,如(3.5)式. 为求(3.8)式右边的偏导 $\partial L/\partial \boldsymbol{r} \equiv \operatorname{grad} \boldsymbol{r}$,需要用到一个矢量公式[见附录(Ⅱ.22)式]:

$$\operatorname{grad}(\boldsymbol{a}\cdot\boldsymbol{b}) = (\boldsymbol{a}\cdot\nabla)\boldsymbol{b} + (\boldsymbol{b}\cdot\nabla)\boldsymbol{a} + \boldsymbol{b}\times\operatorname{rot}\boldsymbol{a} + \boldsymbol{a}\times\operatorname{rot}\boldsymbol{b}.$$

由此得到

$$\partial L/\partial \boldsymbol{r} = \nabla L = e(\boldsymbol{v}\cdot\nabla)\boldsymbol{A} + e\boldsymbol{v}\times\operatorname{rot}\boldsymbol{A} - e\operatorname{grad}\phi.$$

代入拉格朗日方程(3.8)式,得到

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{p} + e\boldsymbol{A}) = e(\boldsymbol{v}\cdot\nabla)\boldsymbol{A} + e\boldsymbol{v}\times\operatorname{rot}\boldsymbol{A} - e\operatorname{grad}\phi. \tag{3.9}$$

需要特别注意的是,矢势 $\boldsymbol{A} = \boldsymbol{A}(\boldsymbol{r},t)$ 是位置 $\boldsymbol{r}$ 和时间 t 的函数. 它对时间的全导数 $\mathrm{d}\boldsymbol{A}/\mathrm{d}t$ 包含对时间的偏导 $\partial\boldsymbol{A}/\partial t$ 和由于粒子运动使 $\boldsymbol{r}$ 变化而引起的改变两部分:$\mathrm{d}\boldsymbol{A}/\mathrm{d}t = \partial\boldsymbol{A}/\partial t + (\boldsymbol{v}\cdot\nabla)\boldsymbol{A}$. 代入(3.9)式得到带电粒子在电磁场中的运动方程

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = -e\frac{\partial\boldsymbol{A}}{\partial t} - e\operatorname{grad}\phi + e\boldsymbol{v}\times\operatorname{rot}\boldsymbol{A}. \tag{3.10}$$

(3.10)式右边是带电粒子在电磁场中所受的力. 它分为两种类型. 第一类(第一、二项)与粒子的速度无关;而第二类(第三项)线性地依赖于粒子的速度,其方向和速度垂直. 单位电荷所受的力称为场强. 和前一类型力(和速度无关的力)对应的场强称为电场强度,用 $\boldsymbol{E}$ 表示:

$$\boldsymbol{E} = -\frac{\partial\boldsymbol{A}}{\partial t} - \operatorname{grad}\phi. \tag{3.11}$$

和后一类力(线性地依赖于速度的力)对应的场强称为磁感应强度,用 $\boldsymbol{B}$ 表示:

$$\boldsymbol{B} = \operatorname{rot}\boldsymbol{A}. \tag{3.12}$$

将(3.11),(3.12)式代入(3.10)式得到用场强表出的运动方程

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = e\boldsymbol{E} + e\boldsymbol{v} \times \boldsymbol{B}. \tag{3.13}$$

(三)电磁场张量

电场强度 $\boldsymbol{E}$ (3.11)式和磁感应强度 $\boldsymbol{B}$ (3.12)式是三维矢势 $\boldsymbol{A}$ 和标势 ϕ 对三维坐标 $\boldsymbol{r}$ 和时间 t 的导数.四维形式下的场强应表示为四维矢势 A_μ 对四维坐标 x_μ 的导数.

定义电磁场张量为

$$F_{\mu\nu} = \frac{\partial A_\nu}{\partial x_\mu} - \frac{\partial A_\mu}{\partial x_\nu}. \tag{3.14}$$

不难看到它的分量就是三维电场强度和磁感应强度(见习题 10):

$$(F_{\mu\nu}) = \begin{bmatrix} 0 & B_z & -B_y & -\mathrm{i}E_x/c \\ -B_z & 0 & B_x & -\mathrm{i}E_y/c \\ B_y & -B_x & 0 & -\mathrm{i}E_z/c \\ \mathrm{i}E_x/c & \mathrm{i}E_y/c & \mathrm{i}E_z/c & 0 \end{bmatrix}. \tag{3.15}$$

习　题

1.根据经典物理,要把一个电子加速到光速,应使用多大的电势差?应用这个电势差,根据相对论,电子所获得的速度将为多少?在这个速度下,它的质量和动能为多少?

2.当速度比较大但尚未接近光速时,可以在经典力学上加相对论修正.试将 §2.3.1(1.20)式展开到 β^4 项,求动能的相对论修正.

3.已知静质量为 m_0 的相对论性自由粒子的作用量是四维标量 $S = -m_0 c\int_a^b \mathrm{d}s$ 其中 $\mathrm{d}s = \sqrt{-\mathrm{d}x_\mu^2}$ 是间隔,如 §2.3.1(1.25)式.

(a)求相对论性自由粒子的拉格朗日函数 L;

(b)利用 §1.2.1(1.3)、(1.15)式求相对论性自由粒子的动量 $\boldsymbol{p}$ 和能量 E;

(c)利用最小作用量原理推导相对论性自由粒子的运动方程.

4.静质量为 m_0 的两个高速粒子相互对碰.一个粒子的原有速度为 $0.8c$,另一个为 $0.6c$,假设碰撞是完全非弹性的,使两个粒子粘在一起.用 m_0 及 c 回答下列问题:(a)碰撞后的动量是多少?(b)牛顿力学所预言的动量是多少?(c)碰撞后总能量是多少?(d)碰撞后总的静质量是多少?(e)碰撞后的动能是多少?

5.一个静质量为 M_0 的物体,当静止时分裂成静止质量为 M_{01} 和 M_{02} 的两部分,求这两部分的能量 E_1 和 E_2 及速度 v_1 和 v_2 各为若干?

6.频率为 ω 的光子(能量为 $\hbar\omega$,动量 $\hbar k$)碰在自由电子上,试证明:

(a)电子不可能吸收这个光子,否则能量和动量守恒律不能满足;

(b)电子可以散射这个光子,散射后光子频率 ω' 比散射前光子频率小(这和经典理论中散射光频率不变的结论不同,称为康普顿效应),

$$\omega - \omega' = [2\hbar/(m_0c^2)]\omega\omega'\sin^2(\theta/2),$$

亦即散射光波长

$$\lambda' = \lambda + [4\pi\hbar/(m_0c)]\sin^2(\theta/2),$$

λ 为散射前光子的波长，m_0 为电子的静止质量，θ 为散射光子与入射方向的夹角.

7. 一个总质量为 m_0 的激发原子，对所选定的坐标系静止. 它在跃迁到能量比之低 ΔE_0 的基态时，发射一个光子(能量为 $\hbar\omega$，动量 $\hbar k$)，同时受到光子的反冲，因此光子的频率不能正好是 $\omega = \Delta E_0/\hbar$，而要略小一些. 证明这个频率 $\omega = (\Delta E_0/\hbar)[1 - \Delta E_0/(2m_0c^2)]$.

8. 不稳定的粒子在运动中衰变为运动方向夹角 θ，静止质量分别为 m_{01} 和 m_{02}，动量分别为 $\boldsymbol{p}_1$ 和 $\boldsymbol{p}_2$ 的两个粒子，求衰变前母体粒子的运动速度及其静止质量.

9. 频率为 ω 的光子被匀速运动的自由电子散射，试求散射光子的频率 ω' 与其运动方向的关系. 设散射前自由电子的动量为 $\boldsymbol{p}_e$，它与入射光子运动方向的夹角为 θ，散射后光子的运动方向与其入射方向的夹角(散射角)为 θ'.

10. 根据四维矢势 A_μ 及四维坐标 x_μ 的定义，写出电磁场张量

$$F_{\mu\nu} = \frac{\partial A_\nu}{\partial x_\mu} - \frac{\partial A_\mu}{\partial x_\nu},$$

并将它的各个分量用电场强度 $\boldsymbol{E}$ 和磁感应强度 $\boldsymbol{B}$ 表出[即 §2.3.3(3.15)式].

11. 根据四维二阶张量 $F_{\mu\nu}$ 在参考系 K 和相对于 K 以速度 $\boldsymbol{V} = V\boldsymbol{e}_x$ 运动的参考系 K′之间的变换规则 §2.2.6(6.11)式，写出电场强度 $\boldsymbol{E}$ 和磁感应强度 $\boldsymbol{B}$ 在这两个参考系之间的变换规则.

第四章　电磁现象的基本规律

本章要阐明电磁运动形态的基本规律. 电磁场是物质存在的一种形态，它具有能量、动量等物质运动的基本属性. 电磁场的状态可以用电场强度 $\boldsymbol{E}(\boldsymbol{x},t)$ 和磁感应强度 $\boldsymbol{B}(\boldsymbol{x},t)$（或者等效地用四维形式的电磁场张量 $F_{\mu\nu}$）来描述，而电磁场的变化规律由电磁场的运动方程决定❶. 本章先在三维形式下介绍自然界普遍规律之一的电荷守恒律、作为电磁场运动方程的麦克斯韦方程组以及描述电磁场对荷电物体作用力的洛伦兹公式；接着导出两种介质分界面上的场方程——电磁场的边值关系；然后讨论场与电荷体系的能量转化和守恒定律以及动量转化和守恒定律；最后给出电磁规律的四维表述.

电磁量有许多种单位制. 在具体问题中，为了方便，习惯上往往采用各自不同的单位制. 大体上说来，对宏观电磁问题，常常采用国际单位制；而在讨论微观问题时，习惯上常采用高斯单位制. 在本章和下两篇中将采用国际单位制；而在第五篇到第十篇中将采用高斯单位制. 在附录Ⅵ中对这两种单位制及它们之间的换算有简单的介绍.

§2.4.1　电磁规律的三维表述

本节首先归纳一下由实际中总结出的电磁现象的基本规律，然后给出介质电磁性质的实验定律.

（一）电荷守恒律　麦克斯韦方程组和洛伦兹力公式

18 世纪中叶以来，人们总结了一系列电磁运动形态的实验规律，其中包括电荷守恒律：

$$\oint \boldsymbol{J}_f \cdot \mathrm{d}\boldsymbol{\sigma} + \frac{\mathrm{d}}{\mathrm{d}t}\int \rho_f \mathrm{d}V = 0, \tag{1.1}$$

以及库仑定律、毕奥－萨伐尔定律和法拉第电磁感应定律. (1.1) 式中的 $\boldsymbol{J}_f$ 为传导电流密度，ρ_f 为自由电荷密度，$\mathrm{d}\boldsymbol{\sigma}$ 和 $\mathrm{d}V$ 分别为进行积分的面积元矢量和体积元.

1863 年，麦克斯韦在总结这些定律的基础上，提出了位移电流假设，建立了真空中的麦克斯韦方程组：

---

❶ 在本章和第三、四篇里，常用 $\boldsymbol{x}$ 表示场点的三维坐标，用 $\boldsymbol{x}'$ 表示源点（电荷所在位置）的三维坐标，而用 $\boldsymbol{r}$ 表示从源点 $\boldsymbol{x}'$ 指向场点 $\boldsymbol{x}$ 的矢量：$\boldsymbol{r} = \boldsymbol{x} - \boldsymbol{x}'$.

$$\begin{cases} \oint \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{\sigma} = Q_f/\varepsilon_0, & (1.2) \\ \oint \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{l} = -\int \dfrac{\partial \boldsymbol{B}}{\partial t} \cdot \mathrm{d}\boldsymbol{\sigma}, & (1.3) \\ \oint \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{\sigma} = 0, & (1.4) \\ \oint \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{l} = \int \mu_0 \left(\boldsymbol{J}_f + \varepsilon_0 \dfrac{\partial \boldsymbol{E}}{\partial t} \right) \cdot \mathrm{d}\boldsymbol{\sigma}. & (1.5) \end{cases}$$

这里 $\boldsymbol{E}$ 是电场强度，$\boldsymbol{B}$ 是磁感应强度，ε_0 和 μ_0 是真空中的介电常数和磁导率.

为了由真空中的麦克斯韦方程组导出介质中的麦克斯韦方程组，只要注意到，介质是许多电子和原子核分布在真空中的一个电荷体系. 在电磁场的作用下，介质将发生极化与磁化，在介质内出现宏观的电荷、电流分布. 因此，只要把自由电荷、传导电流以及介质中由于极化和磁化出现的宏观电荷、电流同时代入真空中的麦克斯韦方程组，即可得到介质中的麦克斯韦方程组：

$$\begin{cases} \oint_S \boldsymbol{D} \cdot \mathrm{d}\boldsymbol{\sigma} = Q_f, & (1.6) \\ \oint_L \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{l} = -\dfrac{\mathrm{d}}{\mathrm{d}t} \int_S \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{\sigma}, & (1.7) \\ \oint_S \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{\sigma} = 0, & (1.8) \\ \oint_L \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{l} = I_f + \dfrac{\mathrm{d}}{\mathrm{d}t} \int_S \boldsymbol{D} \cdot \mathrm{d}\boldsymbol{\sigma}. & (1.9) \end{cases}$$

这里引入了两个辅助量

$$\boldsymbol{D} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P}, \quad (1.10)$$

$$\boldsymbol{H} = \boldsymbol{B}/\mu_0 - \boldsymbol{M}, \quad (1.11)$$

其中，$\boldsymbol{D}$ 称为电位移矢量，$\boldsymbol{H}$ 称为磁场强度，$\boldsymbol{P}$ 为极化强度，$\boldsymbol{M}$ 为磁化强度. 这些方程的建立及其物理意义在一般《电磁学》教科书中均有论述，在此不再重复.

利用矢量分析中的高斯定理和斯托克斯定理[见附录(Ⅱ.12)，(Ⅱ.14)式]，即可求得介质中麦克斯韦方程组的微分形式为：

$$\begin{cases} \nabla \cdot \boldsymbol{D} = \rho_f, & (1.12) \\ \nabla \times \boldsymbol{E} = -\dfrac{\partial \boldsymbol{B}}{\partial t}, & (1.13) \\ \nabla \cdot \boldsymbol{B} = 0, & (1.14) \\ \nabla \times \boldsymbol{H} = \boldsymbol{J}_f + \dfrac{\partial \boldsymbol{D}}{\partial t}. & (1.15) \end{cases}$$

例如，利用高斯定理可将(1.6)式变为

$$\int_V \nabla \cdot \boldsymbol{D} \mathrm{d}V = \int_V \rho_f \mathrm{d}V.$$

由于区域 V 任意，有 $\nabla \cdot \boldsymbol{D} = \rho_f$. 类似地，利用斯托克斯定理，可将(1.7)式变为：

$$\int_S \nabla \times \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{\sigma} = -\int_S \frac{\partial \boldsymbol{B}}{\partial t} \cdot \mathrm{d}\boldsymbol{\sigma},$$

由于积分区域 S 是任意的，故有 $\nabla \times \boldsymbol{E} = -\partial \boldsymbol{B}/\partial t$.

同理，电荷守恒定律(1.1)的微分形式为：

$$\nabla \cdot \boldsymbol{J}_f + \partial \rho_f / \partial t = 0. \tag{1.16}$$

将静止电荷在静电场中的受力公式与稳定电流在静磁场中的受力公式推广到普遍情形，得到洛伦兹力密度公式：

$$\boldsymbol{f} = \rho_f \boldsymbol{E} + \boldsymbol{J}_f \times \boldsymbol{B}. \tag{1.17}$$

一个电荷为 e，速度为 $\boldsymbol{v}$ 的粒子所受到的电磁力为

$$\boldsymbol{F} = e\boldsymbol{E} + e\boldsymbol{v} \times \boldsymbol{B}. \tag{1.18}$$

这称为洛伦兹力公式. 这两个公式的正确性均为其后的大量实验所证实.

麦克斯韦方程组、洛伦兹力公式和电荷守恒定律正确地反映了电磁场的运动规律以及场与带电物质相互作用的规律，是电动力学的理论基础.

在求解实际问题时，还需要加上介质的电磁性质方程.

(二)介质的电磁性质方程

对于各向同性的线性介质，其电磁性质方程为

$$\boldsymbol{P} = \chi_e \varepsilon_0 \boldsymbol{E}, \tag{1.19}$$

$$\boldsymbol{M} = \chi_m \boldsymbol{H}, \tag{1.20}$$

其中，χ_e 称为极化率，χ_m 称为磁化率. 将上两式代入(1.10)式与(1.11)式，即有：

$$\boldsymbol{D} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P} = \varepsilon_0 (1 + \chi_e) \boldsymbol{E} = \varepsilon_0 \varepsilon_r \boldsymbol{E} = \varepsilon \boldsymbol{E}, \tag{1.21}$$

$$\boldsymbol{B} = \mu_0 (\boldsymbol{H} + \boldsymbol{M}) = \mu_0 (1 + \chi_m) \boldsymbol{H} = \mu_0 \mu_r \boldsymbol{H} = \mu \boldsymbol{H}, \tag{1.22}$$

其中，ε 称为介质的介电常数，$\varepsilon_r = (\varepsilon/\varepsilon_0) = 1 + \chi_e$ 称为相对介电常数；μ 称为磁介质的磁导率，$\mu_r = (\mu/\mu_0) = 1 + \chi_m$ 称为相对磁导率.

对于导电介质来说，还有欧姆定律

$$\boldsymbol{J}_f = \sigma \boldsymbol{E}. \tag{1.23}$$

其中 σ 称为介质的电导率.

公式(1.21)，(1.22)，(1.23)统称介质的电磁性质方程，它们反映了介质的宏观性质.

对于各向异性介质，其介电常数和电导率是张量 ε_{ij} 和 σ_{ij}，它们决定了 $\boldsymbol{D}$、$\boldsymbol{J}$ 与 $\boldsymbol{E}$ 的关系

$$D_i = \varepsilon_{ij} E_j, \qquad (J_f)_i = \sigma_{ij} E_j. \tag{1.24}$$

对于非线性介质，$\boldsymbol{D}$、$\boldsymbol{J}_f$ 和 $\boldsymbol{E}$ 的一次式、二次式、三次式等等有关，这是非线性光学研究的问题. 铁磁性物质的 $\boldsymbol{B}$ 与 $\boldsymbol{H}$ 也存在非线性关系，常用磁化曲线和磁滞回线来描述. 这是铁磁学研究的问题. 在这里我们不作详细讨论.

§2.4.2 电磁场的边值关系

上节给出了各向同性均匀介质的麦克斯韦方程组. 但是，在介质分界面上，由于可能出现面电流和面电荷的分布，场量发生不连续的跃变，微分形式的麦克斯韦方程组将不能应用. 本节要建立介质分界面两侧场量的关系，称为电磁场的边值关系. 它实质上是介质分界面上的场方程.

由于边值关系牵涉到电荷电流分布的描述，为此先加以讨论：

(一)电荷电流分布的描述

电磁学中已遇到点电荷、面电荷和体电荷的概念. 实际上电荷都有一定的大小，只不过当观察点到电荷的距离远大于电荷的线度时，可以把电荷看作点电荷；当观察点到电荷的距离远大于电荷薄层的厚度时，可以把电荷看作面电荷；当观察点到电荷的距离远大于电荷的横截面时，可以把电荷看作线电荷.

使用这些概念时相应地引入了反映电荷密度的量，如电荷体密度 $\rho = \Delta Q/\Delta V$，电荷面密度 $\sigma = \Delta Q/\Delta\sigma$，电荷线密度 $\eta = \Delta Q/\Delta l$. 其中 ΔQ 分别是在体积 ΔV 内，面积 $\Delta\sigma$ 的面内，长度 Δl 的线上的电量.

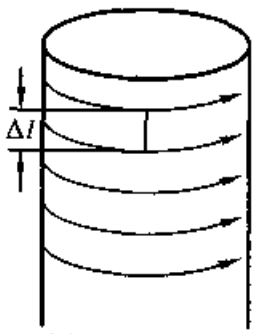


图 1 面电流

值得注意的是，当电荷作面分布或线分布时，有可能对应于很大的体密度. 因为当电荷作面分布时，$\rho = \Delta Q/(\Delta\sigma \cdot d) = \sigma/d$，其中 d 是带电薄层的厚度. 由于 d 很小，ρ 有可能很大. 同样，当电荷作线分布时，它的 $\rho = \Delta Q/(\Delta l \cdot A) = \eta/A$，其中 A 是带电细线的横截面积. 由于 A 很小，ρ 有可能很大. 利用 δ 函数(见附录Ⅰ)，位于 x_0 点的点电荷 Q 的电荷密度可表示为：$\rho(x) = Q\delta^3(x - x_0)$.

再讨论电流分布. 当电流通过导线时，用电流强度 $I = \Delta Q/\Delta t$ 描述，它表示单位时间通过导体横截面的电量. 当有体电流通过大块导电介质时，用电流密度 $\boldsymbol{J}$ 描述，$\boldsymbol{J}$ 的方向沿电流方向，其大小 $J = \Delta I/\Delta\sigma$，它表示垂直通过单位横截面的电流，可以理解为体电流的面密度. 当有面电流流过某一曲面时(图 1)，则用面电流密度 $\boldsymbol{\alpha}$ 描述，$\boldsymbol{\alpha}$ 的方向沿电流方向，其大小 $\alpha = \Delta I/\Delta l$，它表示垂直通过单位横截线的电流，可以理解为面电流的线密度.

(二)电磁场的边值关系

现在由麦克斯韦方程组的积分形式出发，推导电磁场的边值关系.

1. 界面法向分量的边值关系.

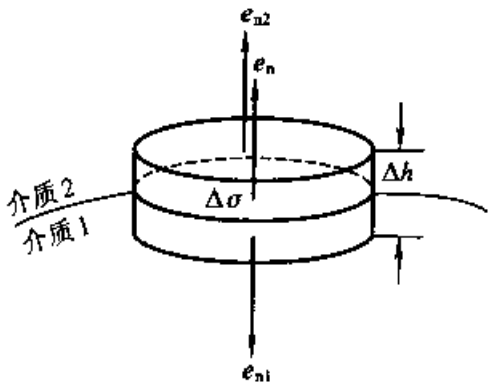


图2　求场量法向分量边值关系的扁平盒

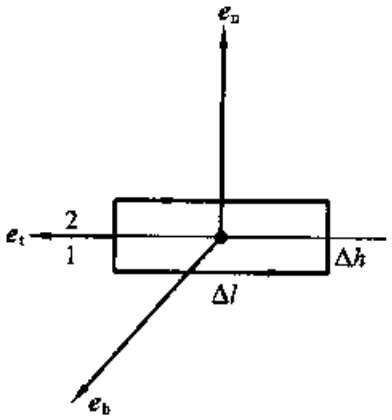


图3　求场量切向分量边值关系的矩形回路

电场强度法向分量的边值关系可以由

$$\oint \boldsymbol{E}\cdot \mathrm{d}\boldsymbol{\sigma} = (Q_f + Q_P)/\varepsilon_0$$

导出. 为此,在边界面上作一个扁平盒(如图2),盒的上、下底平行于介质1、2分界面的小面元 $\Delta\sigma$,盒的高为 Δh,且 Δh 相对于 $\Delta\sigma$ 的线度是高级小量. 因此,计算通过盒表面的电通量时可以略去通过侧面的通量. 同时,考虑到扁平盒上、下底的外法线方向与面元 $\Delta\sigma$ 法线方向的关系为 $\boldsymbol{e}_{n2} = \boldsymbol{e}_n$ 及 $\boldsymbol{e}_{n1} = -\boldsymbol{e}_n$,则有

$$\varepsilon_0(E_{2n}\Delta\sigma - E_{1n}\Delta\sigma) = (\sigma_f + \sigma_P)\Delta\sigma,$$

其中 σ_f 和 σ_P 分别是分界面上的自由电荷面密度和极化电荷面密度. 即

$$\varepsilon_0(E_{2n} - E_{1n}) = \sigma_f + \sigma_P, \tag{2.1}$$

同理,由 $\oint \boldsymbol{P}\cdot \mathrm{d}\boldsymbol{\sigma} = -Q_P$ 可得极化强度法向分量的边值关系

$$P_{2n} - P_{1n} = -\sigma_P, \tag{2.2}$$

由 $\oint \boldsymbol{D}\cdot \mathrm{d}\boldsymbol{\sigma} = Q_f$ 可得电位移矢量法向分量的边值关系

$$D_{2n} - D_{1n} = \sigma_f. \tag{2.3}$$

当然(2.3)式也可由 $\boldsymbol{D} = \varepsilon_0\boldsymbol{E} + \boldsymbol{P}$ 与(2.1)、(2.2)式联立求得.

根据同样的道理,由 $\oint \boldsymbol{B}\cdot \mathrm{d}\boldsymbol{\sigma} = 0$ 可得磁感应强度法向分量的边值关系:

$$B_{2n} - B_{1n} = 0, \tag{2.4}$$

由 $\boldsymbol{B} = \mu\boldsymbol{H}$ 可得磁场强度法向分量的边值关系

$$\mu_2 H_{2n} = \mu_1 H_{1n}. \tag{2.5}$$

由(2.1)(2.2)(2.3)式可见,面电荷的存在是使 E_n, P_n, D_n 在介质分界面上发生跃变的原因.

2. 界面切向分量的边值关系

磁场强度切向分量的边值关系可以由 $\oint \boldsymbol{H}\cdot \mathrm{d}\boldsymbol{l}=I_{\mathrm{f}}+\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_S \boldsymbol{D}\cdot \mathrm{d}\boldsymbol{\sigma}\right)$ 导出. 为确定起见,取 $\boldsymbol{e}_{\mathrm{n}}$ 为界面法向单位矢量,$\boldsymbol{e}_{\mathrm{t}}$ 为某一切向单位矢量,$\boldsymbol{e}_{\mathrm{b}}=\boldsymbol{e}_{\mathrm{n}}\times \mathrm{e}_{\mathrm{t}}$,如图 3 所示. 在垂直于界面的平面上作一个狭窄的矩形回路,其中,边长为 Δl 的长边与界面平行,短边的边长 $\Delta h \ll \Delta l$. 矩形的面积元矢量用 $\Delta\boldsymbol{\sigma}=\Delta l\Delta h\boldsymbol{e}_{\mathrm{b}}$ 表示. 既然 Δh 相对 Δl 是高级小量,计算 $\boldsymbol{H}$ 的环流时略去在 Δh 边上的贡献,即有:

$$\boldsymbol{H}_2\cdot\boldsymbol{e}_{\mathrm{t}}\Delta l-\boldsymbol{H}_1\cdot\boldsymbol{e}_{\mathrm{t}}\Delta l=\boldsymbol{\alpha}_{\mathrm{f}}\cdot\boldsymbol{e}_{\mathrm{b}}\Delta l+(\partial\boldsymbol{D}/\partial t)\cdot\boldsymbol{e}_{\mathrm{b}}\Delta l\Delta h,$$

其中 $\boldsymbol{\alpha}_{\mathrm{f}}$ 是传导电流的线密度. 既然 Δh 为高级小量,上式右方第二项亦可略去,由此得

$$H_{2\mathrm{t}}-H_{1\mathrm{t}}=\boldsymbol{\alpha}_{\mathrm{f}}\cdot\boldsymbol{e}_{\mathrm{b}}. \tag{2.6}$$

若 $\boldsymbol{e}_{\mathrm{b}}$ 刚好与面电流方向相同,则有

$$H_{2\mathrm{t}}-H_{1\mathrm{t}}=\alpha_{\mathrm{f}}\ . \tag{2.7}$$

若 $\boldsymbol{e}_{\mathrm{b}}$ 为任意方向,(2.6)式可写成矢量形式:

$$\boldsymbol{\alpha}_{\mathrm{f}}\cdot\boldsymbol{e}_{\mathrm{b}}=(\boldsymbol{H}_2-\boldsymbol{H}_1)\cdot\boldsymbol{e}_{\mathrm{t}}=(\boldsymbol{H}_2-\boldsymbol{H}_1)\cdot(\boldsymbol{e}_{\mathrm{b}}\times\boldsymbol{e}_{\mathrm{n}})=[\boldsymbol{e}_{\mathrm{n}}\times(\boldsymbol{H}_2-\boldsymbol{H}_1)]\cdot\boldsymbol{e}_{\mathrm{b}}.$$

由 $\boldsymbol{e}_{\mathrm{b}}$ 的任意性,可得

$$\boldsymbol{e}_{\mathrm{n}}\times(\boldsymbol{H}_2-\boldsymbol{H}_1)=\boldsymbol{\alpha}_{\mathrm{f}}\ . \tag{2.8}$$

(2.6)式与(2.8)式是等价的,(2.7)式是特殊情况. 它们均指出,面传导电流 $\boldsymbol{\alpha}_f$ 的存在是磁场强度切向分量跃变的原因.

由 $\boldsymbol{B}=\mu\boldsymbol{H}$ 可得磁感应强度切向分量的边值关系

$$\boldsymbol{e}_{\mathrm{n}}\times\left(\frac{\boldsymbol{B}_2}{\mu_2}-\frac{\boldsymbol{B}_1}{\mu_1}\right)=\boldsymbol{\alpha}_{\mathrm{f}}\ . \tag{2.9}$$

同理,由 $\oint\boldsymbol{E}\cdot\mathrm{d}\boldsymbol{l}=-\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_S\boldsymbol{B}\cdot\mathrm{d}\boldsymbol{\sigma}\right)$ 可得电场强度切向分量的边值关系

$$E_{2\mathrm{t}}-E_{1\mathrm{t}}=0. \tag{2.10}$$

或

$$\boldsymbol{e}_{\mathrm{n}}\times(\boldsymbol{E}_2-\boldsymbol{E}_1)=0. \tag{2.11}$$

同理,由 $\oint\boldsymbol{M}\cdot\mathrm{d}\boldsymbol{l}=I_{\mathrm{M}}$ 可得磁化强度切向分量的边值关系

$$\boldsymbol{e}_{\mathrm{n}}\times(\boldsymbol{M}_2-\boldsymbol{M}_1)=\boldsymbol{\alpha}_{\mathrm{M}}, \tag{2.12}$$

其中 $\boldsymbol{\alpha}_{\mathrm{M}}$ 是磁化电流线密度.

其他物理量的边值关系均可按上述方法求得,不再一一赘述. 在下一篇里,将会看到边值关系在求解实际问题中的重要作用.

§2.4.3　电磁场的守恒定律

(一)电磁场与电荷的能量转化与守恒

物理学的发展史告诉我们,人们对于一种新的运动形态的认识,往往是通过它与已知运动形态的能量转化和守恒而得来.例如,对势能的认识是通过势能与动能的相互转化得来;对作为分子热运动的热能的认识是通过热能和机械能的转化与守恒定律得来.下面,我们从电磁场能量与带电体机械能的转化与守恒定律出发,探讨电磁场能量密度和能流密度的表达式.

考虑一个空间区域 V,它的界面为 σ.能量转化与守恒定律为

$$\begin{pmatrix}\text{单位时间通过界面}\\ \sigma\text{ 流入 }V\text{ 内的能量}\end{pmatrix}=\begin{pmatrix}\text{单位时间场对 }V\\ \text{内电荷所作的功}\end{pmatrix}+\begin{pmatrix}\text{单位时间 }V\text{ 内电}\\ \text{磁场能量的增量}\end{pmatrix}. \tag{3.1}$$

为了将它写成定量的表达式,需要引入新的物理量:

电磁场的能量密度 $w(\boldsymbol{x},t)$.它表示 t 时 $\boldsymbol{x}$ 点附近单位体积的场能.于是体积 V 内电磁场的能量为 $\int_V w\mathrm{d}V$.而 V 内电磁场能量单位时间的增加为 $\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_V w\mathrm{d}V\right)$.

电磁场的能流密度 $\boldsymbol{S}(\boldsymbol{x},t)$,其方向表示 t 时 $\boldsymbol{x}$ 处电磁场能量流动的方向,其大小表示 t 时 $\boldsymbol{x}$ 处单位时间内垂直流过单位横截面的电磁场能量.因此,单位时间内通过封闭面 S 流入 V 内的电磁场能量为 $-\oint_S \boldsymbol{S}\cdot\mathrm{d}\boldsymbol{\sigma}$.

电磁场对电荷的作用力密度 $\boldsymbol{f}$.以 $\boldsymbol{v}$ 表示电荷运动速度,则单位时间场对 V 内电荷所作的功为 $\int_V \boldsymbol{f}\cdot\boldsymbol{v}\mathrm{d}V$.

将这些量代入(3.1)式,即得能量转化与守恒定律

$$-\oint_S \boldsymbol{S}\cdot\mathrm{d}\boldsymbol{\sigma}=\int_V \boldsymbol{f}\cdot\boldsymbol{v}\mathrm{d}V+\frac{\mathrm{d}}{\mathrm{d}t}\int_V w\mathrm{d}V. \tag{3.2}$$

用高斯定理将左边换成体积分,然后令被积函数相等,得到相应的微分形式为

$$\nabla\cdot\boldsymbol{S}+\frac{\partial w}{\partial t}=-\boldsymbol{f}\cdot\boldsymbol{v}. \tag{3.3}$$

如果 V 包括全空间,则通过无限远界面的能流为零,面积分项无贡献,(3.2)式变为

$$\int_\infty \boldsymbol{f}\cdot\boldsymbol{v}\mathrm{d}V=-\frac{\mathrm{d}}{\mathrm{d}t}\int_\infty w\mathrm{d}V. \tag{3.4}$$

即电磁场对全空间电荷所作的总功率等于场的总能量的减少率,亦即场与电荷的总能量守恒.

利用麦克斯韦方程组及洛伦兹力公式与(3.3)式联立，即可求得电磁能流密度 S 与电磁能量密度 w 的表达式. 因为

$$\begin{aligned} -f\cdot v &= -(\rho_f E+\rho_f v\times B)\cdot v = -E\cdot\rho_f v = -E\cdot J \\ &= -E\cdot[\nabla\times H-\partial D/\partial t] \\ &= \nabla\cdot(E\times H)-H\cdot(\nabla\times E)+E\cdot\partial D/\partial t \\ &= \nabla\cdot(E\times H)+H\cdot\partial B/\partial t+E\cdot\partial D/\partial t. \end{aligned} \tag{3.5}$$

比较(3.3)式与(3.5)式可得

$$S = E\times H, \tag{3.6}$$

$$\frac{\partial w}{\partial t} = E\cdot\frac{\partial D}{\partial t}+H\cdot\frac{\partial B}{\partial t}. \tag{3.7}$$

严格说来，(3.6)式与(3.7)式仅是符合(3.3)式的一种可能的表达式，而不是唯一的表达式. 只是由于运用这个表达式得到的一切推论均与实验结果符合才证实了它的正确性.

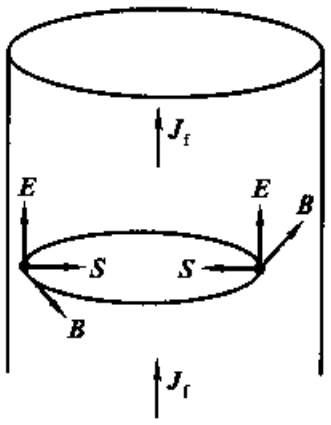


图 1　通电直导线中的电磁能流

在真空情况下，$D=\varepsilon_0 E$，$H=B/\mu_0$，(3.6)式和(3.7)式可以简化为

$$S = \frac{1}{\mu_0}E\times B, \tag{3.8}$$

$$w = \frac{1}{2}\left(\varepsilon_0 E^2+\frac{1}{\mu_0}B^2\right). \tag{3.9}$$

对于各向同性的线性介质，$D=\varepsilon E$，$B=\mu H$，(3.6)式和(3.7)式可表示为

$$S = \frac{1}{\mu}E\times B, \tag{3.10}$$

$$w = \frac{1}{2}(E\cdot D+H\cdot B). \tag{3.11}$$

下面我们讨论关于能流密度 S 的例子.

【例 1】 试以电池供电的电路为例，简要说明能量传输的物理图像.

解 我们先考察一段长度为 l，横截面为 $A=\pi a^2$，电流为 I_f 的圆柱形直导线，讨论能流密度 S 的方向以及单位时间内从侧面流入该段导体的电磁能流.

设导线中的电流方向如图 1 所示. 由欧姆定律 $J_f=\sigma E$ 可知，E 与 J_f 同向；而由右手螺旋法则可知，B 沿逆时针方向. 由此可见，$S=E\times H$ 的方向沿着圆柱的径向向内.

将电场 $E=J_f/\sigma$ 及圆柱导线表面处的磁场 $H=I_f/2\pi a=J_f\pi a^2/2\pi a=aJ_f/2$ 代入即有：$S=(J_f/\sigma)\cdot(aJ_f/2)=(a/2\sigma)J_f^2$. 于是，单位时间从导体表面流入导

体内部的电磁能量为：

$$S \cdot 2\pi al = \frac{a}{2\sigma} J_f^2 2\pi al = (J_f \pi a^2)^2 \frac{l}{\sigma \pi a^2} = I_f^2 R.$$

这说明，单位时间从侧面流入该段导体的电磁能量，等于该段导体在单位时间内所消耗的焦耳热 $I_f^2 R$. 由此可见，电磁能量是从导线表面流向导线内，而不是沿着导线流入该段导体.

整个电池供电线路的能流方向如图 2 所示. 在电池内部，由于化学力的作用，$\boldsymbol{J}_f$ 与 $\boldsymbol{E}$ 反向，电磁场对电流作负功（功率密度 $\boldsymbol{f} \cdot \boldsymbol{v} = \boldsymbol{J}_f \cdot \boldsymbol{E} < 0$），此处能流密度 $\boldsymbol{S}$ 的方向由电池内指向电池外. 这说明电池提供的能量是通过电磁场传到导线各个部分的. 值得注意的是，在导线表面上存在面电荷分布（特别是在导线拐弯的地方），因面在导线外部 $\boldsymbol{S}$ 的方向不一定垂直于导体表面.

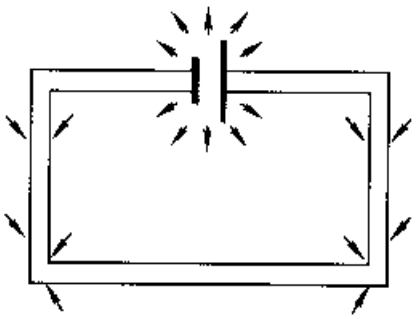


图 2　电池供电电路的能流方向

（二）电磁场与电荷的动量转化与守恒

现在通过体积 V 内电荷的机械动量与电磁场动量的转化与守恒定律得出电磁场动量密度和电磁场动量流密度的表达式.

考虑空间的某一区域 V，它的界面为 σ. 区域 V 内的电荷和电流的分布由电荷密度 $\rho_f(\boldsymbol{x},t)$ 及电流密度 $\boldsymbol{J}_f(\boldsymbol{x},t)$ 描述. 电磁场与电荷的动量转化与守恒定律为

$$\begin{pmatrix}\text{单位时间通过界面}\\ \sigma\text{ 流入 }V\text{ 内的动量}\end{pmatrix} = \begin{pmatrix}\text{单位时间 }V\text{ 内}\\ \text{电荷动量的增量}\end{pmatrix} + \begin{pmatrix}\text{单位时间 }V\text{ 内电}\\ \text{磁场动量的增量}\end{pmatrix}. \quad (3.12)$$

为了将它写成定量的表达式，同样需要引入新的物理量.

回忆在上一段中，在考虑能量守恒与转化的时候，曾引入电磁场的能量密度 $w(\boldsymbol{x},t)$，它表示 t 时 $\boldsymbol{x}$ 点附近，单位体积的场能. 类似地，有电磁场的动量密度 $\boldsymbol{g}(\boldsymbol{x},t)$，它表示 t 时 $\boldsymbol{x}$ 点附近，单位体积电磁场的动量. 由于动量是矢量，所以电磁场的动量密度 $\boldsymbol{g}(\boldsymbol{x},t)$ 是坐标 $\boldsymbol{x},t$ 的矢量函数. 单位时间中，体积 V 内电磁场动量的增量是 $\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_V \boldsymbol{g}\mathrm{d}V\right)$.

同样，在考虑能量守恒与转化的时候，曾引入电磁场的能流密度 $\boldsymbol{S}(\boldsymbol{x},t)$，它是一个三维矢量，其方向表示能量流动的方向，大小表示单位时间垂直流过单位横截面的电磁场能量. 现在，由于动量本身是矢量，所以电磁场的“动量流密度”不是简单的矢量，而是两个矢量的“并矢”（即二阶张量）. 动量流密度常用 $\overleftrightarrow{\boldsymbol{T}} = \sum T_{ij}\boldsymbol{e}_i\boldsymbol{e}_j$ 表示. 它和坐标基矢 $\boldsymbol{e}_i$ 的点积 $\boldsymbol{e}_i \cdot \overleftrightarrow{\boldsymbol{T}} = \sum T_{ij}\boldsymbol{e}_j$ 是一个矢量，它等于单位

时间流过垂直于 e_i 的单位横截面的电磁场动量. 因而 $\mathrm{d}\boldsymbol{\sigma}\cdot\vec{\vec{T}}$ 就是单位时间内电磁场由 $\mathrm{d}\boldsymbol{\sigma}$ 后方流向 $\mathrm{d}\boldsymbol{\sigma}$ 前方的动量. 所以，$-\oint_S \mathrm{d}\boldsymbol{\sigma}\cdot\vec{\vec{T}}$ 就是单位时间内通过封闭面 S 流入 V 内的电磁场动量.

我们知道，电磁场对电荷的作用力密度为 $\boldsymbol{f}$，它使 V 内电荷的动量发生变化. 因此，单位时间体积 V 内电荷动量的增量即为 $\int_V \boldsymbol{f}\mathrm{d}V$.

将这些量代入(3.12)式，即有

$$-\oint_S \mathrm{d}\boldsymbol{\sigma}\cdot\vec{\vec{T}} = \int_V \boldsymbol{f}\mathrm{d}V + \frac{\mathrm{d}}{\mathrm{d}t}\int_t \boldsymbol{g}\mathrm{d}V. \tag{3.13}$$

利用面积分化体积分的公式[见附录(Ⅱ.23)式] $\oint_S \cdots \mathrm{d}\sigma = \int_V \nabla \cdots \mathrm{d}V$，将面积分换成体积分，即有：$\int_V \boldsymbol{f}\mathrm{d}V = -\int_V \mathrm{d}V\ \nabla\cdot\vec{\vec{T}} - \frac{\mathrm{d}}{\mathrm{d}t}\int_V \boldsymbol{g}\mathrm{d}V$. 由于积分体积 V 的任意性，得到与积分形式的(3.13)式相对应的微分形式：

$$\boldsymbol{f} = -\nabla\cdot\vec{\vec{T}} - \partial\boldsymbol{g}/\partial t. \tag{3.14}$$

如果 V 包括全空间，则通过无限远界面的动量为零，因而面积分项无贡献. (3.13)式变为

$$\int_\infty \boldsymbol{f}\mathrm{d}V = -\frac{\mathrm{d}}{\mathrm{d}t}\int_\infty \boldsymbol{g}\mathrm{d}V. \tag{3.15}$$

即电磁场对全空间电荷的作用力等于场的总动量的减少率. 亦即场与电荷的总动量守恒.

利用麦克斯韦方程组及洛伦兹力公式与(3.14)式联立，即可求得电磁场动量流密度 $\vec{\vec{T}}$ 与电磁场动量密度 $\boldsymbol{g}$ 的表达式.

作 $\nabla\times\boldsymbol{E} = -\partial\boldsymbol{B}/\partial t$ 与 $\varepsilon_0\boldsymbol{E}$ 的矢积，得

$$\varepsilon_0(\nabla\times\boldsymbol{E})\times\boldsymbol{E} - \varepsilon_0\boldsymbol{E}\times\partial\boldsymbol{B}/\partial t = 0. \tag{3.16}$$

由 $\nabla\times\boldsymbol{H} = \boldsymbol{J}_f + \partial\boldsymbol{D}/\partial t$ 可求得 $\boldsymbol{J}_f$ 为：

$$\boldsymbol{J}_f = \frac{1}{\mu_0}\nabla\times\boldsymbol{B} - \varepsilon_0\frac{\partial\boldsymbol{E}}{\partial t}. \tag{3.17}$$

由 $\nabla\cdot\boldsymbol{D} = \rho_f$ 可求得 ρ_f 为：

$$\rho_f = \varepsilon_0\nabla\cdot\boldsymbol{E}. \tag{3.18}$$

将 $\nabla\cdot\boldsymbol{B} = 0$ 两边乘以 $\boldsymbol{B}/\mu_0$ 可得

$$(\nabla\cdot\boldsymbol{B})\boldsymbol{B}/\mu_0 = 0. \tag{3.19}$$

此外，注意到

$$-\varepsilon_0\frac{\partial}{\partial t}(\boldsymbol{E}\times\boldsymbol{B}) = -\varepsilon_0\left[\frac{\partial\boldsymbol{E}}{\partial t}\times\boldsymbol{B} + \boldsymbol{E}\times\frac{\partial\boldsymbol{B}}{\partial t}\right]. \tag{3.20}$$

将(3.16)～(3.20)式代入洛伦兹力密度公式,即有

$$\begin{aligned}\boldsymbol{f} &= \rho_t \boldsymbol{E} + \boldsymbol{J}_f \times \boldsymbol{B} \\ &= \varepsilon_0 (\nabla \cdot \boldsymbol{E})\boldsymbol{E} + \left(\frac{1}{\mu_0}\nabla \times \boldsymbol{B} - \varepsilon_0 \frac{\partial \boldsymbol{E}}{\partial t}\right) \times \boldsymbol{B} \\ &= \varepsilon_0 (\nabla \cdot \boldsymbol{E})\boldsymbol{E} + \frac{1}{\mu_0}(\nabla \cdot \boldsymbol{B})\boldsymbol{B} + \frac{1}{\mu_0}(\nabla \times \boldsymbol{B}) \times \boldsymbol{B} \\ &\quad + \varepsilon_0 (\nabla \times \boldsymbol{E}) \times \boldsymbol{E} - \varepsilon_0 \frac{\partial}{\partial t}(\boldsymbol{E} \times \boldsymbol{B}).\end{aligned}$$

再利用习题的结果:

$$(\nabla \cdot \boldsymbol{E})\boldsymbol{E} + (\nabla \times \boldsymbol{E}) \times \boldsymbol{E} = \nabla \cdot \left(\boldsymbol{EE} - \frac{1}{2}\overleftrightarrow{e} E^2\right),$$

$$(\nabla \cdot \boldsymbol{B})\boldsymbol{B} + (\nabla \times \boldsymbol{B}) \times \boldsymbol{B} = \nabla \cdot \left(\boldsymbol{BB} - \frac{1}{2}\overleftrightarrow{e} B^2\right),$$

其中 $\overleftrightarrow{e}$ 为单位张量,即有

$$\boldsymbol{f} = \nabla \cdot \left[\varepsilon_0 \boldsymbol{EE} + \frac{1}{\mu_0}\boldsymbol{BB} - \frac{1}{2}\overleftrightarrow{e}\left(\varepsilon_0 E^2 + \frac{1}{\mu_0}B^2\right)\right] - \varepsilon_0 \frac{\partial}{\partial t}(\boldsymbol{E} \times \boldsymbol{B}).$$

和(3.14)式相比较,即有

$$\overleftrightarrow{T} = -\varepsilon_0 \boldsymbol{EE} - \frac{1}{\mu_0}\boldsymbol{BB} + \frac{1}{2}\overleftrightarrow{e}\left(\varepsilon_0 E^2 + \frac{1}{\mu_0}B^2\right), \tag{3.21}$$

$$\boldsymbol{g} = \varepsilon_0 \boldsymbol{E} \times \boldsymbol{B}. \tag{3.22}$$

利用(3.10)、(3.11)和(3.22)式还可求得电磁场动量密度 $\boldsymbol{g}$ 与电磁场能流密度 $\boldsymbol{S}$ 的关系

$$\boldsymbol{g} = \varepsilon_0 \boldsymbol{E} \times \boldsymbol{B} = \varepsilon_0 \mu_0 \boldsymbol{E} \times \boldsymbol{H} = \boldsymbol{S}/c^2. \tag{3.23}$$

其中 $\varepsilon_0 \mu_0 = 1/c^2$.

最后,让我们指出,$-\overleftrightarrow{T}$ 又称为电磁场应力张量,特别是在稳恒的情况下,(3.13)式变为

$$\boldsymbol{F} = \int_V \boldsymbol{f} \mathrm{d}V = -\oint_S \mathrm{d}\boldsymbol{\sigma} \cdot \overleftrightarrow{T}. \tag{3.24}$$

这就是作用于 V 内电荷系统的合力. 利用(3.24)式可以计算场对电荷的作用力以及电磁波对物体的辐射压力.

【例2】　在真空中 $(0,0,-a)$ 及 $(0,0,+a)$ 处分别放置点电荷 $+q$、$-q$,求 $-q$ 所受到的作用力.

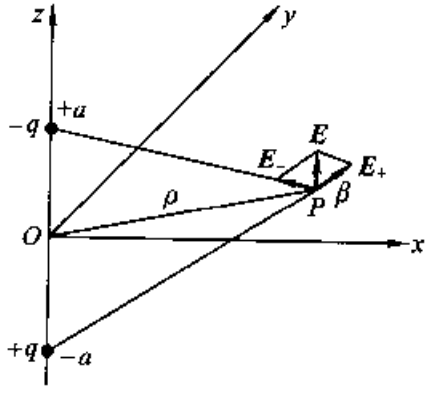


图3　电荷 $-q$ 的受力分析

解 利用(3.24)式计算 $-q$ 所受到的作用力

$$\boldsymbol{F} = -\oint \mathrm{d}\boldsymbol{\sigma}\cdot\overleftrightarrow{\boldsymbol{T}} = \varepsilon_0\oint \mathrm{d}\boldsymbol{\sigma}\cdot\left(\boldsymbol{EE}-\frac{1}{2}E^2\overleftrightarrow{\boldsymbol{e}}\right).$$

首先选择一个便于计算的包围 $-q$ 的封闭面. 例如可取由 $z=0$ 的平面 S_0，及球心在原点半径 $R\to\infty$ 的半球面 S_R 组成的封闭面. 于是

$$\boldsymbol{F} = \varepsilon_0\int_{S_0}\mathrm{d}\boldsymbol{\sigma}\cdot\left(\boldsymbol{EE}-\frac{1}{2}E^2\overleftrightarrow{\boldsymbol{e}}\right)+\varepsilon_0\int_{S_R}\mathrm{d}\boldsymbol{\sigma}\cdot\left(\boldsymbol{EE}-\frac{1}{2}E^2\overleftrightarrow{\boldsymbol{e}}\right).$$

因为在 $R\to\infty$ 的半球面上，$E\approx R^{-2}$，$E^2\approx R^{-4}$，$\mathrm{d}S = R^2\mathrm{d}\Omega\approx R^2$，故第二项积分趋于零而略去，因而只要计算在 $z=0$ 平面上的积分. 令 $\rho^2 = x^2+y^2$（见图3），则在 $z=0$ 平面上，电场强度 $\boldsymbol{E}$ 为

$$\boldsymbol{E} = 2E_+\cos\beta\boldsymbol{e}_z = 2\frac{q}{4\pi\varepsilon_0(a^2+\rho^2)}\frac{a}{\sqrt{a^2+\rho^2}}\boldsymbol{e}_z = \frac{qa}{2\pi\varepsilon_0}(a^2+\rho^2)^{-\frac{3}{2}}\boldsymbol{e}_z.$$

因而力

$$\boldsymbol{F} = \varepsilon_0\int_{S_0}(-\mathrm{d}\sigma\boldsymbol{e}_z)\cdot\frac{q^2a^2}{4\pi^2\varepsilon_0^2}\frac{1}{(a^2+\rho^2)^3}\left(\boldsymbol{e}_z\boldsymbol{e}_z-\frac{1}{2}\overleftrightarrow{\boldsymbol{e}}\right)$$

$$= -\frac{q^2a^2\boldsymbol{e}_z}{8\pi^2\varepsilon_0}\int_0^{2\pi}\mathrm{d}\varphi\int_0^{\infty}\frac{\rho\mathrm{d}\rho}{(a^2+\rho^2)^3} = -\frac{q^2\boldsymbol{e}_z}{4\pi\varepsilon_0(2a)^2}.$$

可见它与库仑定律的结果一致. 它清楚地表明，“静电力”不是电荷之间直接进行的“超距作用力”，而是通过电磁场动量流传递的.

§2.4.4 电磁规律的四维表述

相对性原理指出，物理规律在所有惯性系中是相同的，电磁规律也不例外. 为了明显地表现出电磁规律的协变性，最好的办法是将它们写成四维形式，使得方程中的每一项都是同类协变量. 为此，我们利用四维电流密度矢量 J_μ、四维电势矢量 A_μ 和四维二阶反对称电磁场张量 $F_{\mu\nu}$，将电荷守恒律、麦克斯韦方程及洛伦兹力公式写成协变形式. 最后，利用四维协变量的变换性质导出电磁场的变换公式. 为了简单起见，本节只讨论真空中的电磁场.

(一)四维电流密度矢量 电荷守恒定律的四维形式

为了确定四维电流密度 J_μ 的各个分量，首先找出与它有关的不变量. 由电荷守恒定律可知，带电粒子的总电量不因其运动的速度而变化. 所以 Q 是不变量. 带电粒子的总电量是

$$Q = \int\rho\mathrm{d}V, \tag{4.1}$$

当带电粒子以速度 v 运动时，它占有的体积由于运动长度缩短而收缩，收缩的因子

为 $\gamma_v^{-1} = \sqrt{1-(v/c)^2}$. 与此同时，由于 Q 是不变量，电荷密度 ρ 相应要增加 $\gamma_v = 1/\sqrt{1-(v/c)^2}$ 倍. 因此

$$\rho = \rho_0/\sqrt{1-(v/c)^2} = \gamma_v \rho_0, \tag{4.2}$$

其中 ρ_0 是在相对粒子静止的参考系 K_0 测得的固有电荷密度，而 ρ 是在粒子相对它以 v 运动的参考系 K 测得的电荷密度. 显然，ρ_0 是不变量，而对于不同的参考系，测得的 ρ 不同. 由(4.2)式可见，在参考系 K 中，有电流密度

$$\boldsymbol{J} = \rho \boldsymbol{v} = \gamma_v \rho_0 \boldsymbol{v}. \tag{4.3}$$

与四维速度矢量的定义 $u_\mu = \gamma_v(\boldsymbol{v}, ic)$ 相比较，定义四维电流密度矢量为

$$J_\mu = \rho(\boldsymbol{v}, ic) = \rho_0 \gamma_v(\boldsymbol{v}, ic) = \rho_0 u_\mu. \tag{4.4}$$

当 $v \ll c$ 时，$\gamma_v \sim 1$，J_μ 的空间分量 $\rho_0 \boldsymbol{v}$ 就是经典的电流密度 $\boldsymbol{J}$，所以(4.4)式的定义是合理的.

电流密度 $\boldsymbol{J}$ 与电荷密度 ρ 合为四维矢量，显示了这两个物理量的统一性. 如果在惯性系 K_0 中粒子静止，只有电荷密度 ρ_0；从相对于 K_0 作匀速运动的惯性系 K 中看，粒子运动，出现了电流密度 $\boldsymbol{J} = \rho_0 \gamma_v \boldsymbol{v}$，与此同时，电荷密度变为 $\rho = \gamma_v \rho_0$. 因此 ρ 和 $\boldsymbol{J}$ 是一个统一物理量的不同方面，当参考系变换时，它按

$$J'_\mu = a_{\mu\nu} J_\nu \tag{4.5}$$

变换.

利用四维电流密度矢量 J_μ 可将电荷守恒定律写成协变形式. 采用简化符号 $\partial_\mu = \partial/\partial x_\mu$，即有

$$\partial_\mu J_\mu = \frac{\partial J_\mu}{\partial x_\mu} = \frac{\partial J_i}{\partial x_i} + \frac{\partial(ic\rho)}{\partial(ict)} = \nabla \cdot \boldsymbol{J} + \frac{\partial \rho}{\partial t} = 0. \tag{4.6}$$

$\partial_\mu J_\mu$ 没有自由指标，是四维标量(即洛伦兹变换下的不变量)，在惯性系变换时其值不变，故上式对任意惯性系成立.

(二)四维电势矢量和电磁场张量　麦克斯韦方程组的四维形式

在 §2.3.3 中已经引入了四维势矢量

$$A_\mu = (\boldsymbol{A}, i\varphi/c), \tag{4.7}$$

和四维二阶反对称电磁场张量

$$F_{\mu\nu} = \frac{\partial A_\nu}{\partial x_\mu} - \frac{\partial A_\mu}{\partial x_\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{4.8}$$

利用电磁场张量可以将麦克斯韦方程组写成下列形式：

$$\begin{cases} \partial_\nu F_{\mu\nu} = \mu_0 J_\mu, & (4.9) \\ \partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0. & (4.10) \end{cases}$$

这两个式子写成三维形式就是 §2.4.1 的(1.12)-(1.15)式，见习题 13.

(三)电磁场的变换公式

为了求得电磁场在不同惯性系之间的变换有两种方法.

一是利用四维势矢量 A_μ 的变换关系,即

$$A'_\mu = a_{\mu\nu}A_\nu, \tag{4.11}$$

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}, \qquad \boldsymbol{E} = -\nabla\varphi - \frac{\partial \boldsymbol{A}}{\partial t}, \tag{4.12}$$

$$\boldsymbol{B}' = \nabla' \times \boldsymbol{A}', \qquad \boldsymbol{E}' = -\nabla'\varphi' - \frac{\partial \boldsymbol{A}'}{\partial t'}, \tag{4.13}$$

其中 $\nabla' = \boldsymbol{e}'_i \partial/\partial x'_i$.

二是利用电磁场张量 $F_{\mu\nu}$ 的变换关系,即

$$F'_{\mu\nu} = a_{\mu\lambda}a_{\nu\tau}F_{\lambda\tau}, \tag{4.14}$$

这时可利用 §2.3.3(3.15)式逐个求出 $\boldsymbol{E}$ 和 $\boldsymbol{B}$ 的分量. 例如要求 E'_2,可由,

$$\begin{aligned} iE'_2/c = F'_{42} &= a_{4\lambda}a_{2\tau}F_{\lambda\tau} = a_{41}a_{2\tau}F_{1\tau} + a_{44}a_{2\tau}F_{4\tau} \\ &= a_{41}a_{22}F_{12} + a_{44}a_{22}F_{42} = i\gamma(E_2 - VB_3)/c. \end{aligned}$$

在推导过程中,注意到 $a_{\mu\nu}$ 与 $F_{\mu\nu}$ 都有一些矩阵元为零. 先对 $\lambda = 1,2,3,4$ 迭加,仅存两项;然后再对 $\tau = 1,2,3,4$ 迭加,仍存两项.

此外,也可以把(4.14)式改写为矩阵乘积的形式来求. 由

$$F'_{\mu\nu} = a_{\mu\lambda}F_{\lambda\tau}\tilde{a}_{\tau\nu},$$

写成矩阵相乘的形式,即为

$$F' = AF\tilde{A}, \tag{4.15}$$

其中 $F = (F_{\mu\nu})$, $A = (a_{\mu\nu})$,而 $\tilde{A}$ 为 A 的转置矩阵. 因此,只要将右边三个矩阵相乘之后,由(4.15)式两边相应矩阵元相等,即可求得 $\boldsymbol{E}$ 和 $\boldsymbol{B}$ 的各分量的变换关系.

用以上任一种方法得到的电磁场的变换公式可归纳为:

$$\begin{cases} E'_1 = E_1, \quad E'_2 = \gamma(E_2 - VB_3), \quad E'_3 = \gamma(E_3 + VB_2), & (4.16) \\ B'_1 = B_1, \quad B'_2 = \gamma\left(B_2 + \dfrac{V}{c^2}E_3\right), \quad B'_3 = \gamma\left(B_3 - \dfrac{V}{c^2}E_2\right). & (4.17) \end{cases}$$

不难发现,电磁场平行于两参考系之间的相对速度 $\boldsymbol{V}$ 的分量和垂直于 $\boldsymbol{V}$ 的分量的变换公式分别为

$$\begin{cases} \boldsymbol{E}'_\parallel = \boldsymbol{E}_\parallel, \qquad \boldsymbol{E}'_\perp = \gamma(\boldsymbol{E} + \boldsymbol{V} \times \boldsymbol{B})_\perp, & (4.18) \\ \boldsymbol{B}'_\parallel = \boldsymbol{B}_\parallel, \qquad \boldsymbol{B}'_\perp = \gamma\left(\boldsymbol{B} - \dfrac{\boldsymbol{V}}{c^2} \times \boldsymbol{E}\right)_\perp. & (4.19) \end{cases}$$

当 $V \ll c$ 时,即过渡到非相对论的电磁场变换式,

$$\begin{cases} \boldsymbol{E}' = \boldsymbol{E} + \boldsymbol{V} \times \boldsymbol{B}, & (4.20) \\ \boldsymbol{B}' = \boldsymbol{B} - \boldsymbol{V} \times \boldsymbol{E}/c^2. & (4.21) \end{cases}$$

矢势 $\boldsymbol{A}$ 与标势 φ 统一为四维势矢量 A_μ，以及磁场 $\boldsymbol{B}$ 与电场 $\boldsymbol{E}$ 统一为四维电磁场张量 $F_{\mu\nu}$，反映了电磁场的统一性和相对性. 在给定的参考系中，电磁场表现出不同的性质. 如电场对场中的电荷有作用力，磁场对场中的电流有作用力，两者性质不同. 但是，电场与磁场的变换关系表明，电场与磁场在一定条件下是可以相互转化的. 例如，在参考系 K 中有一个静止的电荷，即只有 ρ_f 存在，故在 K 系中只观察到电场；但在相对 K 系以 $\boldsymbol{V}$ 运动的 K′系中，电荷就是运动的，故既存在有 ρ'_f，也存在 $\boldsymbol{J}'_f$，既观察到电场也观察到磁场. 可见电场与磁场可以相互转化.

【例 1】　求以匀速 $\boldsymbol{V} = \boldsymbol{v} = ve_x$ 运动的带电粒子 e 的电磁场.

解　设实验室参考系为 K，取参考系 K′的原点固定在 e 上，即 K′系相对 K 系以 $\boldsymbol{v} = ve_x$ 运动. 仍然设 $t = 0$ 时两系的原点重合.

由于 e 相对 K′系静止，故 K′系观察到电磁场为：

$$\begin{cases} \boldsymbol{E}' = e\boldsymbol{r}'/(4\pi\varepsilon_0 r'^3), & (4.22) \\ \boldsymbol{B}' = 0. & (4.23) \end{cases}$$

利用电磁场的反变换式，即将公式(4.16)，(4.17)中 $\boldsymbol{E}$，$\boldsymbol{B}$ 与 $\boldsymbol{E}'$，$\boldsymbol{B}'$ 交换的同时将 V 换为 $-\boldsymbol{V} = -\boldsymbol{v}$，可得

$$\begin{cases} E_x = \dfrac{ex'}{4\pi\varepsilon_0 r'^3}, \quad E_y = \dfrac{\gamma ey'}{4\pi\varepsilon_0 r'^3}, \quad E_z = \dfrac{\gamma ez'}{4\pi\varepsilon_0 r'^3}, & (4.24) \\ B_x = 0, \quad B_y = -\dfrac{v}{c^2}E_z, \quad B_z = \dfrac{v}{c^2}E_y. & (4.25) \end{cases}$$

现在的任务是用 K 系测得的 x, y, z, r，取代 K′系测得的 x', y', z', r'. 为简单起见，只求 $t = 0$ 时的电磁场，因为对于匀速运动的带电粒子，电磁场的性质不随时间而变.

将洛伦兹变换 $x' = \gamma x, y' = y, z' = z$ 代入(4.24)式，可得：

$$\boldsymbol{E} = \frac{\gamma e\boldsymbol{r}}{4\pi\varepsilon_0[(\gamma x)^2 + y^2 + z^2]^{3/2}} = \frac{e\boldsymbol{r}}{4\pi\varepsilon_0\gamma^2\left[\dfrac{1}{\gamma^2}(x^2 + y^2 + z^2) + \left(1 - \dfrac{1}{\gamma^2}\right)x^2\right]^{3/2}}.$$

利用

$$\left(1 - \frac{1}{\gamma^2}\right)x^2 = \left[1 - \left(1 - \frac{v^2}{c^2}\right)\right]x^2 = \frac{v^2x^2}{c^2} = \left(\frac{\boldsymbol{v}\cdot\boldsymbol{r}}{c}\right)^2.$$

可得

$$\begin{cases} \boldsymbol{E} = \dfrac{e\boldsymbol{r}}{4\pi\varepsilon_0\gamma^2[r^2/\gamma^2 + (\boldsymbol{v}\cdot\boldsymbol{r}/c)^2]^{3/2}}, & (4.26) \\ \boldsymbol{B} = -\dfrac{v}{c^2}E_z\boldsymbol{e}_y + \dfrac{v}{c^2}E_y\boldsymbol{e}_z = \dfrac{v}{c^2}(\boldsymbol{e}_x \times \boldsymbol{E}) = \dfrac{\boldsymbol{v}}{c^2} \times \boldsymbol{E}. & (4.27) \end{cases}$$

下面讨论当 $v \ll c$ 及 $v \approx c$ 时电磁场的分布.

当 $v \ll c$，将 $\gamma \approx 1$ 代入(4.26)、(4.27)式可得

$$\begin{cases} \boldsymbol{E} = \dfrac{e\boldsymbol{r}}{4\pi\varepsilon_0 r^3} = \boldsymbol{E}_0, & (4.28) \\ \boldsymbol{B} = \dfrac{\boldsymbol{v}}{c^2} \times \boldsymbol{E} = \dfrac{\boldsymbol{v}}{c^2} \times \dfrac{e\boldsymbol{r}}{4\pi\varepsilon_0 r^3} = \dfrac{\mu_0 e\boldsymbol{v} \times \boldsymbol{r}}{4\pi r^3}. & (4.29) \end{cases}$$

因为带电粒子在 $t = 0$ 时在原点处，故 $\boldsymbol{r}$ 即为从源点指向场点的矢量. (4.28)式及(4.29)式表明，库仑定理与毕奥－沙伐尔定律仍近似正确.

当 $v \approx c$ 时，γ 很大，而 $\gamma^{-2} = (1 - v^2/c^2)$ 很小. 在平行于 $\boldsymbol{v}$ 的方向上，分母的 $r^2/\gamma^2 \ll (\boldsymbol{v} \cdot \boldsymbol{r}/c)^2$，故保留 $(\boldsymbol{v} \cdot \boldsymbol{r}/c)^2$ 项；而在垂直于 $\boldsymbol{v}$ 的方向上，$\boldsymbol{v} \cdot \boldsymbol{r} = 0$，故仅 r^2/γ^2 起作用. 此外，由于 $x' = 0$，$r = x - x' = x$，故

$$\begin{cases} \boldsymbol{E}_{\parallel} = \left(1 - \dfrac{v^2}{c^2}\right) \dfrac{e\boldsymbol{r}}{4\pi\varepsilon_0 (\boldsymbol{v} \cdot \boldsymbol{x}/c)^3} = \left(1 - \dfrac{v^2}{c^2}\right) \dfrac{e\boldsymbol{x}}{4\pi\varepsilon_0 r^3} \ll \boldsymbol{E}_0, \\ \boldsymbol{E}_{\perp} = \gamma \dfrac{e\boldsymbol{x}}{4\pi\varepsilon_0 r^3} \gg \boldsymbol{E}_0. \end{cases}$$

这说明高速运动的带电粒子的电场集中在与 $\boldsymbol{v}$ 垂直的方向上，特别是当 $v \approx c$ 时，整个电场的能量几乎集中在通过粒子且垂直于 $\boldsymbol{v}$ 的平面附近，其电场线分布在平行于 $\boldsymbol{v}$ 的一个平面上的投影如图 1 中的直线. 磁场线则是圆心在粒子运动轨迹上的圆圈. 图中用 · 与 × 表示 $\boldsymbol{B}$ 垂直于纸面，方向向上(·)与向下(×). 电磁场的能流密度 $\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H}$ 是以粒子为心的圆圈，并且 $\boldsymbol{S}$ 的方向系由粒子的后方经圆圈指向粒子的前方. 这说明作匀速运动的带电粒子是没有辐射的，整个电磁场跟随着粒子向前运动. 这一点也可定量证明如下：因为 $\boldsymbol{E}$ 与 $\boldsymbol{e}_R$ 同向，所以 $\boldsymbol{S} \cdot \boldsymbol{e}_R = (\boldsymbol{E} \times \boldsymbol{H}) \cdot \boldsymbol{e}_R = (\boldsymbol{e}_R \times \boldsymbol{E}) \cdot \boldsymbol{H} = 0$. 将此图绕粒子运动轨迹旋转可见，电场线形成一个个圆锥面，而能流密度则形成一些同心球面.

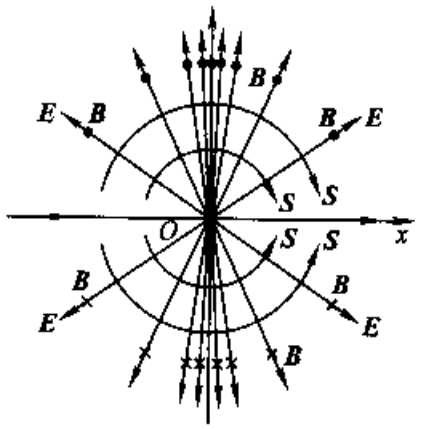


图 1 匀速运动带电粒子激发的电磁场

【例 2】　设 K 系电磁波由

$$\boldsymbol{E} = \boldsymbol{E}_0 \cos\varphi, \quad \boldsymbol{B} = \boldsymbol{B}_0 \cos\varphi \quad (\text{其中 } \varphi = \omega t - \boldsymbol{k} \cdot \boldsymbol{x})$$

描述，K′系电磁波由

$$\boldsymbol{E}' = \boldsymbol{E}'_0 \cos\varphi', \quad \boldsymbol{B}' = \boldsymbol{B}'_0 \cos\varphi' \quad (\text{其中 } \varphi' = \omega' t' - \boldsymbol{k}' \cdot \boldsymbol{x}')$$

描述. 试证明 $\varphi' = \varphi$.

证　由电磁场变换公式(4.16)，有 $E'_{\parallel} = E_{\parallel}$，即有

$$E'_{01}\cos\varphi' = E_{01}\cos\varphi.$$

由 t 与 x 的任意性，即有：$E'_{01} = E_{01}$，$\varphi' = \varphi$．　【证毕】

最后，让我们指出，由电磁场张量 $F_{\mu\nu}$ 可以构成两个独立的不变量（标量）❶：

$$\begin{cases} F_{\mu\nu}^2/2 = \boldsymbol{B}^2 - \boldsymbol{E}^2/c^2, & (4.30) \\ (\mathrm{i}c/8)\varepsilon_{\alpha\beta\gamma\delta}F_{\alpha\beta}F_{\gamma\delta} = \boldsymbol{E}\cdot\boldsymbol{B}, & (4.31) \end{cases}$$

其中 $\varepsilon_{\alpha\beta\gamma\delta}$ 是四阶完全反对称单位张量（见附录Ⅰ）．

利用(4.30)式和(4.31)式可以由一个参考系电磁场的性质推导出其他参考系电磁场的某些性质．例如，若在 K 系 $\boldsymbol{E}\perp\boldsymbol{B}$，则由(4.30)式可知：$\boldsymbol{E}'\cdot\boldsymbol{B}' = \boldsymbol{E}\cdot\boldsymbol{B} = 0$，故在任何参考系均有 $\boldsymbol{E}'\perp\boldsymbol{B}'$，或者 $\boldsymbol{E}'$ 与 $\boldsymbol{B}'$ 之一为零．

如果在 K′系中 $\boldsymbol{E}' = 0$，由 $\boldsymbol{B}'^2 - \boldsymbol{E}'^2/c^2 = \boldsymbol{B}^2 - \boldsymbol{E}^2/c^2$ 可得

$$B' = \sqrt{B^2 - (E/c)^2}. \qquad (4.32)$$

（四）洛伦兹力公式的四维形式

利用四维力矢量、速度矢量和电磁场张量，可将洛伦兹力公式写成四维形式：

$$K_\mu = eF_{\mu\nu}u_\nu. \qquad (4.33)$$

不难证明，它的前三个分量式就是洛伦兹力公式：

$$\boldsymbol{F} = \boldsymbol{K}/\gamma_v = e(\boldsymbol{E} + \boldsymbol{v}\times\boldsymbol{B}). \qquad (4.34)$$

它的第四个分量式为功率公式：

$$\mathrm{d}E/\mathrm{d}t = e\boldsymbol{E}\cdot\boldsymbol{v}. \qquad (4.35)$$

习　题

1．设 $\boldsymbol{r} = (x-x')\boldsymbol{e}_x + (y-y')\boldsymbol{e}_y + (z-z')\boldsymbol{e}_z$，且 ∇ 是对场点 $\boldsymbol{x}$ 求微商的算符，∇' 是对源点 $\boldsymbol{x}'$ 求微商的算符．试求

(a) ∇r；(b) $\nabla' r$；(c) $\nabla\frac{1}{r}$；(d) $\nabla'\frac{1}{r}$；(e) $\nabla\times\frac{\boldsymbol{r}}{r^3}$；(f) $\nabla\cdot\frac{\boldsymbol{r}}{r^3}$；

(g) $\nabla\cdot\boldsymbol{r}$；(h) $\nabla\times\boldsymbol{r}$；(i) $(\boldsymbol{a}\cdot\nabla)\boldsymbol{r}$；(j) $\nabla(\boldsymbol{a}\cdot\boldsymbol{r})$；(k) $\nabla\cdot[\boldsymbol{E}_0\sin(\boldsymbol{k}\cdot\boldsymbol{r})]$；

(l) $\nabla\times[\boldsymbol{E}_0\sin(\boldsymbol{k}\cdot\boldsymbol{r})]$；(m) $\nabla\mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}$；(n) $\nabla\cdot[\boldsymbol{E}_0\mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}]$；

(o) $\nabla\times[\boldsymbol{E}_0\mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}]$，其中 $\boldsymbol{a}$，$\boldsymbol{k}$，ω 是常量．

2．若 $\boldsymbol{M}$ 是常矢量，证明除 $R=0$ 点外，矢量 $\boldsymbol{A}^{(1)} = \mu_0\boldsymbol{M}\times\boldsymbol{R}/(4\pi R^3)$ 的旋度等于标量 $\varphi_m^{(1)} = \boldsymbol{M}\cdot\boldsymbol{R}/(4\pi R^3)$ 的梯度的负值，即 $\nabla\times\boldsymbol{A}^{(1)} = -\mu_0\nabla\varphi_m^{(1)}(R\neq 0)$，其中 R 为坐标原点到场点的距离，$\boldsymbol{R}$ 的方向由原点指向场点．

3．设 $\boldsymbol{x}'$ 为源点坐标，在小区域 V 内变动．坐标原点在 V 内，而 $\boldsymbol{x}$ 为场点坐标．试证明：

(a) $f(\boldsymbol{x}-\boldsymbol{x}') = f(\boldsymbol{x}) - (\boldsymbol{x}'\cdot\nabla)f(\boldsymbol{x}) + \frac{1}{2!}(\boldsymbol{x}'\cdot\nabla)^2 f(\boldsymbol{x}) + \cdots$

---

❶ 严格说来，$\boldsymbol{E}\cdot\boldsymbol{B}$ 是赝标量，它在空间反演时改号．

(b) $\frac{1}{r}=\frac{1}{R}-(\boldsymbol{x}'\cdot\nabla)\frac{1}{R}+\frac{1}{2!}(\boldsymbol{x}'\cdot\nabla)^2\frac{1}{R}+\cdots$.

提示 令 $F(t)=f(\boldsymbol{x}-\boldsymbol{x}'t)=f(x-x't,y-y't,z-z't)$ 其中 $t\ll 1$,然后将 $F(t)$ 展开为麦克劳林级数.

4. 已知 $\nabla'\cdot\boldsymbol{J}_t(\boldsymbol{x}')=0$,试证明(为书写方便,将 $\boldsymbol{J}_t$ 的 i 分量记为 J_i)

$$\nabla'\cdot[\boldsymbol{J}_t(\boldsymbol{x}')x'x'_i]=\boldsymbol{J}_t(\boldsymbol{x}')x'_i+J_i(\boldsymbol{x}')\boldsymbol{x}',$$

其中 x'_i 及 $J_i(\boldsymbol{x}')$ 分别为 $\boldsymbol{x}'$ 及 $\boldsymbol{J}_t(\boldsymbol{x}')$ 的 i 分量.

提示 利用张量散度公式 $\nabla'\cdot(f\boldsymbol{g}_i)=\boldsymbol{g}_i(\nabla'\cdot f)+(f\cdot\nabla')\boldsymbol{g}_i$,并令 $f=\boldsymbol{J}_t(\boldsymbol{x}')$,$\boldsymbol{g}_i=\boldsymbol{x}'x'_i$.

5. 试证明 $(\nabla\cdot\boldsymbol{E})\boldsymbol{E}+(\nabla\times\boldsymbol{E})\times\boldsymbol{E}=\nabla\cdot(\boldsymbol{EE}-\overleftrightarrow{e}E^2/2)$,其中 $\overleftrightarrow{e}=\boldsymbol{e}_x\boldsymbol{e}_x+\boldsymbol{e}_y\boldsymbol{e}_y+\boldsymbol{e}_z\boldsymbol{e}_z$ 为单位张量.

6. 试利用几何关系写出:

(a)球坐标系 $\boldsymbol{e}_R,\boldsymbol{e}_\theta,\boldsymbol{e}_\phi$ 与 $\boldsymbol{e}_x,\boldsymbol{e}_y,\boldsymbol{e}_z$ 的关系,以及 $\boldsymbol{e}_R,\boldsymbol{e}_\theta,\boldsymbol{e}_\phi$ 分别对 R,θ,ϕ 的偏导数;

(b)柱坐标系 $\boldsymbol{e}_\rho,\boldsymbol{e}_\phi,\boldsymbol{e}_z$ 与 $\boldsymbol{e}_x,\boldsymbol{e}_y,\boldsymbol{e}_z$ 的关系,以及 $\boldsymbol{e}_\rho,\boldsymbol{e}_\phi,\boldsymbol{e}_z$ 分别对 ρ,ϕ,z 的偏导数.

7. 已知一个电荷体系的偶极矩定义为

$$\boldsymbol{p}_e(t)=\int_V\rho_t(\boldsymbol{x}',t)\boldsymbol{x}'\mathrm{d}V'.$$

试利用电荷守恒律 $\nabla\cdot\boldsymbol{J}_t+\partial\rho_t/\partial t=0$ 证明 $\boldsymbol{p}_e$ 的变化率为:

$$\mathrm{d}\boldsymbol{p}_e(t)/\mathrm{d}t=\int_t\boldsymbol{J}_t(\boldsymbol{x}',t)\mathrm{d}V'.$$

8. 有一内外半径分别为 R_1、R_2 的空心介质球,介质的介电常数为 ε,介质均匀带电且其电荷密度为 ρ_t.求:

(a)空间各点的电场;

(b)极化体电荷和极化面电荷分布.

9. 内外半径分别为 R_1、R_2 的无穷长中空导体圆柱,沿轴向流有稳恒均匀传导电流 I_t,导体的磁导率为 μ.求磁感应强度和磁化电流.

10. 试用边值关系证明:在绝缘介质与导体分界面上,在静电情况下,导体外的电场线总是垂直于导体表面;在稳定电流情况下,导体内的电场线总是平行于导体表面.

11. 在一平行板电容器的两板上加上 $v=v_0\cos(\omega t)$ 的电压.若平板为圆形,半径为 a,板间距离为 d.试求:

(a)两板间的位移电流;

(b)电容器离轴为 R 处的磁场强度;

(c)电容器内的能量密度和能流密度的平均值.

12. 利用 $\boldsymbol{E}$ 和 $\boldsymbol{B}$ 的变换规律证明 $\boldsymbol{E}\cdot\boldsymbol{B}$ 和 $E^2-c^2B^2$ 都是洛伦兹变换下的不变量.

13. 证明麦克斯韦方程组的四维形式 §2.4.4(4.9)、(4.10)式写成三维形式就是 §2.4.1 的(1.12)~(1.15)式.

14. 设参考系 K 中有一均匀恒定且不垂直的电场 $\boldsymbol{E}$ 和磁场 $\boldsymbol{B}$.

(a)问对于怎样的参考系 K',在其中的电场与磁场互相平行,即指出满足上述要求的 K'系

相对 K 系的速度 $\boldsymbol{V}$ 的大小和方向.

(b)满足上述要求的参考系是否唯一的,为什么?

15. 在实验室参考系中有两个电量相等,距离为 y_0 的正电荷 q_1 和 q_2. 设它们以 $\boldsymbol{v}=v\boldsymbol{e}_x$ 运动,求 q_1 对 q_2 的作用力 $\boldsymbol{F}_{21}$. 问:这一作用力是引力还是斥力?



第三篇　电　磁　场

在上一篇里，介绍了高速运动的规律，给出了电荷守恒定律，麦克斯韦方程组和洛伦兹力公式的三维和四维形式，它们是经典电磁场的普遍规律.

本篇从这些普遍规律出发，讨论三种情形下的场. 第一章和第二章讨论静电场和静磁场，第三章讨论电磁波的传播. 所用的方法是：引入标势和矢势来表示场，导出标势和矢势所遵守的方程和边值关系，运用多种数学方法求出势，从而获得场的分布，以及相关的电荷、电流分布.

第一章　静　电　场

本章讨论电磁场的一种特殊情形，即电荷静止或只有稳恒电流的情形，在这两种情形下，空间的电荷分布均不随时间而变，因而场量也不随时间而变.

静电场的基本问题是给定电荷的分布 $\rho_f(\boldsymbol{x})$ 以及周围空间中介质和导体的分布，求解静电场. 本章首先引入静电场的标势，求得静电势所满足的微分方程及边值关系，进而讨论静电问题的唯一性定理，它指出给定哪些条件，静电场的解才唯一确定. 随后介绍几种最常用的求解静电边值问题的方法——分离变量法、电象法、格林函数法和多极矩法. 最后讨论静电场的能量，以及电荷体系与外电场的相互作用.

§3.1.1　静电势　静电问题的唯一性定理

(一)静电势

在静电情形下，起作用的麦克斯韦方程及介质的电磁性质方程为：

$$\begin{cases} \nabla \times \boldsymbol{E} = 0, & (1.1) \\ \nabla \cdot \boldsymbol{D} = \rho_f, & (1.2) \\ \boldsymbol{D} = \varepsilon \boldsymbol{E}. & (1.3) \end{cases}$$

这三个方程是解决静电问题的基础.

由(1.1)式及斯托克斯定理可知，电荷 q 在静电场中沿封闭路线运动一周，静电场对 q 所作的功为

$$\oint \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{l} = q\oint \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{l} = q\int (\nabla \times \boldsymbol{E}) \cdot \mathrm{d}\boldsymbol{\sigma} = 0.$$

也就是说，静电场对电荷所作的功与路径无关而只与起点和终点有关，因而可以引入位置的单值函数 φ(P)来表示静电场对单位正电荷所作的功

$$\int_{P_1}^{P_2} \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{l} = \varphi(P_1) - \varphi(P_2), \qquad (1.4)$$

其中 $\varphi(P_1) - \varphi(P_2)$ 称为 P_1 点与 P_2 点之间的电势差，而 φ(P)称为电势. 这里利用电场对单位正电荷作的功来定义电势差，但电势的值还没有完全确定. 必须确定电势的零参考点，空间各点的电势才可能有确定值. 通常当电荷分布在有限区域时，可令 $\varphi(\infty) = 0$；而当电荷分布一直延伸到无限远处时，则只能把电势的零参考点取在有限距离处. 因为这时若仍令 $\varphi(\infty) = 0$，则在有限距离上的 φ 将出现无穷大. 此外，在实际问题中如果有导体接地，我们将取大地的电势为零.

利用(1.4)式可以求得相距为 dl 的两点间的电势差. 因为

$$\int_{P_1}^{P_2} d\varphi = \varphi(P_2) - \varphi(P_1) = -\int_{P_1}^{P_2} \boldsymbol{E} \cdot d\boldsymbol{l}.$$

由 P_1 点和 P_2 点的任意性,即有:

$$d\varphi = -\boldsymbol{E} \cdot d\boldsymbol{l}. \tag{1.5}$$

利用 $d\varphi = \nabla\varphi \cdot d\boldsymbol{l}$ 可得 $\boldsymbol{E} \cdot d\boldsymbol{l} = -d\varphi = -\nabla\varphi \cdot d\boldsymbol{l}$,由 $d\boldsymbol{l}$ 的任意性得到:

$$\boldsymbol{E} = -\nabla\varphi. \tag{1.6}$$

即:电场强度 $\boldsymbol{E}$ 等于电势 φ 的负梯度. 梯度 $\nabla\varphi(\boldsymbol{x}) = \text{grad}\ \varphi(\boldsymbol{x})$ 的方向指向 $\boldsymbol{x}$ 点处 $\varphi(\boldsymbol{x})$ 增加得最快的方向,于是(1.6)式说明 $\boldsymbol{x}$ 点处的 $\boldsymbol{E}$ 指向 $\varphi(\boldsymbol{x})$ 减小得最快的方向. (1.6)式就是静电情况下的 §2.3.3(3.11)式.

在电磁学中介绍过求静电势的方法,若已知电荷分布,则可利用点电荷的电势公式与电势叠加原理来求. 例如,点电荷体系的电势为

$$\varphi(\boldsymbol{x}) = \sum_a Q_a/(4\pi\varepsilon_0 r_a),$$

其中 r_a 是从点电荷 Q_a 的位置 $\boldsymbol{x}_a$ 到场点 $\boldsymbol{x}$ 的距离. 当电荷连续分布时,若电荷密度 $\rho_f(\boldsymbol{x})$ 已知,则 $\boldsymbol{x}$ 点的电势为

$$\varphi(\boldsymbol{x}) = \int \frac{\rho_f(\boldsymbol{x}')\,dV'}{4\pi\varepsilon_0 r},$$

其中 r 是从源点位置 $\boldsymbol{x}'$ 到场点 $\boldsymbol{x}$ 的距离.

当然也可以根据电荷分布求电场强度 $\boldsymbol{E}$,再按(1.4)式求电势,其中取 P_1 点的电势为零,取 P_2 为场点.

总之,电磁学所介绍的方法都要先给出电荷分布,这就存在很大的局限性. 例如,在一个导体球附近放一个点电荷 Q,求空间各点的电势. 这时,电磁学所介绍的方法就无能为力了,因为导体球上感应电荷的分布并未给出. 这时,只能求解 φ 的边值问题. 为此就要知道:1)φ 满足的方程;2)φ 在导体或介质分界面上的边值关系;3)φ 在求解区域的边界 S 上的边界条件.

通常边界条件要预先给定,因此,为了求出 φ,先要寻找 φ 满足的方程及边值关系.

(二)静电势满足的方程及边值关系

对于均匀的各向同性线性介质内部,将(1.6)式代入(1.3)式后再代入(1.2)式,可得 $\rho_f = \nabla \cdot \boldsymbol{D} = \varepsilon \nabla \cdot \boldsymbol{E} = -\varepsilon \nabla^2\varphi$,

$$\nabla^2\varphi = -\rho_f/\varepsilon. \tag{1.7}$$

这就是静电势 φ 所满足的方程,称为泊松方程.

在两种介质的分界面上,φ 的边值关系为

$$\varphi_1 = \varphi_2, \tag{1.8}$$

$$\varepsilon_2 \partial\varphi_2/\partial n - \varepsilon_1 \partial\varphi_1/\partial n = -\sigma_f. \tag{1.9}$$

证 设 P_1 与 P_2 分别是介质 1 及介质 2 边界两侧无限靠近的两点，$\overline{P_1P_2} = \Delta l \to 0$. 由(1.5)式得

$$\varphi_2 - \varphi_1 = \varphi(P_2) - \varphi(P_1) = -\boldsymbol{E} \cdot \Delta \boldsymbol{l} = 0.$$

最后一个等号是考虑到即使在界面附近 $\boldsymbol{E}$ 也是有限量，因而当 $\Delta l \to 0$ 时其点积为零，故有 $\varphi_1 = \varphi_2$，这就是(1.8)式.

将 $\boldsymbol{D} = \varepsilon \boldsymbol{E}$ 与 $\boldsymbol{E} = -\nabla\varphi$ 代入 $\boldsymbol{D}$ 的边值关系 $\boldsymbol{e}_n \cdot (\boldsymbol{D}_2 - \boldsymbol{D}_1) = \sigma_f$ 可得：

$$\sigma_f = -\boldsymbol{e}_n \cdot (\varepsilon_2 \nabla \varphi_2 - \varepsilon_1 \nabla \varphi_1) = -(\varepsilon_2 \partial\varphi_2/\partial n - \varepsilon_1 \partial\varphi_1/\partial n),$$

其中 $\boldsymbol{e}_n$ 为由介质 1 指向介质 2 的法线(见图 1). 将负号移至 σ_f 一侧，即为(1.9)式. 【证毕】

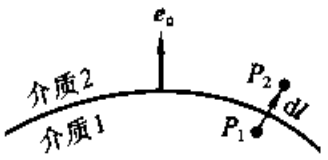


图 1 两种介质的分界面

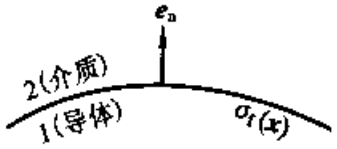


图 2 导体与介质的分界面

在导体与介质分界面上 φ 的边值关系为(参看图 2)

$$\varphi_1 = \varphi_2 = 常数, \tag{1.10}$$

$$\varepsilon_2 \partial\varphi_2/\partial n = -\sigma_f, \tag{1.11}$$

$$\oint \varepsilon_2 (\partial\varphi_2/\partial n)\, d\sigma = -Q_f. \tag{1.12}$$

证 上面关于(1.8)式的证明在这里仍然有效. (1.10)式中 $\varphi_1 =$ 常数 是由导体的特殊性质决定的. 因为，如果导体不是等势体，导体中的自由电子就会在电场作用下运动，使导体内电荷重新分布，直到达到以下情况：导体内部不带电，电荷只分布在导体表面上；导体内部电场为零，导体表面上电场沿法线方向；导体表面为等势面，整个导体为等势体. 因此，(1.10)式是静电情形下导体表面必须满足的条件.

将(1.10)式代入(1.9)式可得(1.11)式. 设介质 1 为导体，于是(1.9)式左边的第二项 $\partial\varphi_1/\partial n = 0$. (1.11)式中的 $\varphi_2(\boldsymbol{x})$ 是介质 2 的电势，但是由(1.10)式可知，介质无限靠近导体表面处的电势实际上等于导体的电势.

(1.12)式是对(1.11)式取面积分而得. 若封闭面取为导体表面，则(1.12)式右方的 Q_f 就是该导体所带的总电量. 【证毕】

在分离变量法中要利用已知的 Q_f 来确定 $\varphi(\boldsymbol{x})$ 中的待定系数，而在求出 $\varphi(\boldsymbol{x})$ 后往往可以利用(1.11)式来确定导体表面的电荷分布 $\sigma_f(\boldsymbol{x})$.

在两种导电介质的分界面上，φ 的边值关系为

$$\begin{cases}\varphi_1 = \varphi_2, & (1.13)\\ \sigma_1 \partial\varphi_1/\partial n = \sigma_2 \partial\varphi_2/\partial n. & (1.14)\end{cases}$$

证 (1.8)式的证明在这里仍然有效,故有(1.13)式.

(1.14)式中的 σ_1 及 σ_2 分别为两种导电介质的电导率. (1.14)式可以利用欧姆定律 $\boldsymbol{J}_f = \sigma\boldsymbol{E}$,电荷守恒律 $\nabla \cdot \boldsymbol{J}_f + \partial\rho_f/\partial t = 0$ 及稳恒条件 $\partial\rho_f/\partial t = 0$ 来证明. 我们有

$$0 = -\int \partial\rho_f/\partial t \mathrm{d}V = \int \nabla \cdot \boldsymbol{J}_f \mathrm{d}V = \oint \boldsymbol{J}_f \cdot \mathrm{d}\boldsymbol{\sigma} = \oint \sigma\boldsymbol{E} \cdot \mathrm{d}\boldsymbol{\sigma}.$$

仿照 §2.4.2 的图 2 在介质界面上取一扁平盒作为封闭面,利用上式可导出(1.14)式. **【证毕】**

(三)静电问题的唯一性定理

设区域 V 内自由电荷分布 $\rho_f(\boldsymbol{x})$ 给定,电介质分布也给定,且 $\boldsymbol{D} = \varepsilon\boldsymbol{E}$ 成立;在 V 的边界 S 上电势 $\varphi(\boldsymbol{x})|_S$ 或电势的法向导数 $\partial\varphi(\boldsymbol{x})/\partial n|_S$ 给定,则 V 内电场唯一确定. φ 和 $\partial\varphi(\boldsymbol{x})/\partial n$ 右边的竖线和下标 S 表明取表面 S 上的值.

证 为简单起见,设 V 内电介质的介电常数为 ε. 设两个解 $\varphi'(\boldsymbol{x})$ 及 $\varphi''(\boldsymbol{x})$ 满足唯一性定理的条件,作函数

$$\boldsymbol{K}(\boldsymbol{x}) = (\varphi' - \varphi'')(\boldsymbol{D}' - \boldsymbol{D}''),$$

其中,$\boldsymbol{D}' = \varepsilon\boldsymbol{E}' = -\varepsilon\nabla\varphi'$,$\boldsymbol{D}'' = \varepsilon\boldsymbol{E}'' = -\varepsilon\nabla\varphi''$. 求 $\boldsymbol{K}(\boldsymbol{x})$ 的散度,并利用 $\nabla \cdot \boldsymbol{D}' = \nabla \cdot \boldsymbol{D}'' = \rho_f$,即有:

$$\begin{aligned}\nabla \cdot \boldsymbol{K}(\boldsymbol{x}) &= (\varphi' - \varphi'')(\nabla \cdot \boldsymbol{D}' - \nabla \cdot \boldsymbol{D}'') + (\nabla\varphi' - \nabla\varphi'') \cdot (\boldsymbol{D}' - \boldsymbol{D}'')\\ &= -(\boldsymbol{E}' - \boldsymbol{E}'') \cdot (\boldsymbol{D}' - \boldsymbol{D}'') = -\varepsilon(\boldsymbol{E}' - \boldsymbol{E}'')^2. \qquad (1.15)\end{aligned}$$

另一方面,由高斯定理及 $\boldsymbol{D} \cdot \mathrm{d}\boldsymbol{\sigma} = -\varepsilon(\nabla\varphi) \cdot \boldsymbol{e}_n \mathrm{d}\sigma = -\varepsilon(\partial\varphi/\partial n)\mathrm{d}\sigma$,即有

$$\begin{aligned}\int_V \nabla \cdot \boldsymbol{K}(\boldsymbol{x})\mathrm{d}V &= \oint_S \boldsymbol{K}(\boldsymbol{x}) \cdot \mathrm{d}\boldsymbol{\sigma} = \oint_S (\varphi' - \varphi'')(\boldsymbol{D}' - \boldsymbol{D}'') \cdot \mathrm{d}\boldsymbol{\sigma}\\ &= -\oint_S \varepsilon(\varphi' - \varphi'')(\partial\varphi'/\partial n - \partial\varphi''/\partial n)\mathrm{d}\sigma = 0. \qquad (1.16)\end{aligned}$$

面积分在 V 的边界 S 上进行,若在 S 上给定的是电势 $\varphi(\boldsymbol{x})|_S$,则被积函数的 $(\varphi' - \varphi'')$ 因子为零;若在 S 上给定的是电势的法向导数 $\partial\varphi'/\partial n|_S$,则被积函数的 $(\partial\varphi'/\partial n - \partial\varphi''/\partial n)$ 因子为零. 总之,(1.16)式应为零.

对(1.15)式进行体积分并将(1.16)式代入,即有

$$-\int_V \varepsilon(\boldsymbol{E}' - \boldsymbol{E}'')^2 \mathrm{d}V = 0.$$

由于 $\varepsilon > 0$,且 $(\boldsymbol{E}' - \boldsymbol{E}'')^2 \geqslant 0$,由积分为零必然有

$$\boldsymbol{E}' = \boldsymbol{E}''$$

即:静电场是唯一的. 由 $\boldsymbol{E} = -\nabla\varphi$ 可见,电势 φ' 与 φ'' 至多只能相差一个常数,这显

然是无关紧要的. 这样,唯一性定理得证. 【证毕】

如果 V 内含有导体,并且给定每个导体上的电势,则只要把所有导体表面 S_a $(a=1,2,\cdots)$ 看作边界面,并将上面的定理应用于 S 及 $S_a(a=1,2,\cdots)$ 之间的区域 V',由上述定理可知,这个区域内的电场将被唯一确定.

如果不是给定每个导体上的电势,而是给定每个导体上的总电荷 $Q_a(a=1,2,\cdots)$,也可在 V' 区域内应用上面的方法证明. 但应注意,(1.16)式中应包含对各个导体表面进行积分. 对于这种积分,由于导体静电平衡条件,尽管 $(\varphi'-\varphi'')$ 没有给定,但它们必为常数,因而可以提出积分号外,于是积分变为

$$-\int_V \varepsilon(\boldsymbol{E}'-\boldsymbol{E}'')^2 \mathrm{d}V = (\varphi'-\varphi'')\oint_{S_a}(\boldsymbol{D}'-\boldsymbol{D}'')\cdot \mathrm{d}\boldsymbol{\sigma}$$

$$= -(\varphi'-\varphi'')\oint_{S_a}(\partial\varphi'/\partial n - \partial\varphi''/\partial n)\mathrm{d}\sigma = 0.$$

这里对 $\partial\varphi'/\partial n$ 与 $\partial\varphi''/\partial n$ 分别应用了(1.12)式.

唯一性定理不仅告诉我们,需要知道哪些条件,静电场的解才能被唯一确定,而且还告诉我们在具体求解静电场的问题时,可以根据已知条件对问题提出尝试解. 只要尝试解能满足唯一性定理所要求的条件,则这个尝试解就是唯一正确的解.

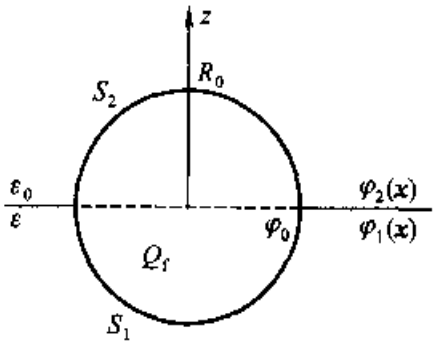


图 3 一半浸没在液体中的导体球

【例 1】 半径为 R_0 的导体球,带自由电荷总量为 Q_f. 今使之一半浸在介电常数为 ε 的液体中,另一半露在真空中. 求静电势、静电场强、束缚电荷分布和自由电荷分布.

解 设导体的电势为 φ_0,液体的电势为 $\varphi_1(x)$,真空的电势为 $\varphi_2(x)$. 采用图 3 所示的球坐标系,定解条件为:

$$\text{方程} \qquad \nabla^2\varphi = 0 \quad (R > R_0), \tag{1.17}$$

$$\text{边界条件} \qquad \lim_{R\to\infty}\varphi = 0, \tag{1.18}$$

$$\text{边值关系} \qquad \varphi_1|_{R=R_0} = \varphi_2|_{R=R_0} = \varphi_0, \tag{1.19}$$

$$-\int_{S_1}\varepsilon(\partial\varphi_1/\partial R)\mathrm{d}\sigma - \int_{S_2}\varepsilon_0(\partial\varphi_2/\partial R)\mathrm{d}\sigma = Q_f, \tag{1.20}$$

$$\varphi_1|_{\theta=\pi/2} = \varphi_2|_{\theta=\pi/2} \quad (R > R_0), \tag{1.21}$$

$$\varepsilon\,\partial\varphi_1/\partial\theta|_{\theta=\pi/2} = \varepsilon_0\,\partial\varphi_2/\partial\theta|_{\theta=\pi/2} \quad (R > R_0). \tag{1.22}$$

由此提出尝试解(C 为待定常数)

$$\left.\begin{aligned}\varphi_1 &= \varphi_2 = C/R \quad (R > R_0),\\ \varphi_0 &= C/R_0 \quad (R < R_0).\end{aligned}\right\} \tag{1.23}$$

不难证明,尝试解满足(1.17)至(1.22)各式.将(1.23)式代入(1.20)式确定 C:

$$C = Q_f/[2\pi(\varepsilon_0 + \varepsilon)].$$

将 C 代回(1.23)式,有

$$\varphi_1 = \varphi_2 = Q_f/[2\pi(\varepsilon_0 + \varepsilon)R] \quad (R > R_0),$$

$$\varphi_0 = Q_f/[2\pi(\varepsilon_0 + \varepsilon)R_0] \quad (R < R_0).$$

由唯一性定理可知,这个解是正确的.

电场强度 $\boldsymbol{E}$ 为

$$\boldsymbol{E} = -\nabla\varphi = Q_f\boldsymbol{R}/[2\pi(\varepsilon_0 + \varepsilon)R^3] \quad (R > R_0),$$

$$\boldsymbol{E} = 0 \quad (R < R_0).$$

自由电荷分布由 $\sigma_f = \boldsymbol{e}_n \cdot (\boldsymbol{D}_2 - \boldsymbol{D}_1)$ 确定

下半球面 $\sigma_f = \boldsymbol{e}_R \cdot (\boldsymbol{D}_1 - 0)|_{R=R_0} = \varepsilon Q_f/[2\pi(\varepsilon_0 + \varepsilon)R_0^2]$,

上半球面 $\sigma_f = \boldsymbol{e}_R \cdot (\boldsymbol{D}_2 - 0)|_{R=R_0} = \varepsilon_0 Q_f/[2\pi(\varepsilon_0 + \varepsilon)R_0^2]$,

液体平面 $\sigma_f = \boldsymbol{e}_\theta \cdot (\boldsymbol{D}_1 - \boldsymbol{D}_2)|_{z=0} = 0$ (因 $\boldsymbol{e}_\theta \cdot \boldsymbol{R} = 0$).

束缚电荷分布由 $\sigma_P = -\boldsymbol{e}_n \cdot (\boldsymbol{P}_2 - \boldsymbol{P}_1)$ 确定:

下半球面 $\sigma_P = -\boldsymbol{e}_R \cdot (\boldsymbol{P}_1 - 0)|_{R=R_0} = (\varepsilon_0 - \varepsilon)Q_f/[2\pi(\varepsilon_0 + \varepsilon)R_0^2]$,

上半球面 $\sigma_P = -\boldsymbol{e}_R \cdot (\boldsymbol{P}_2 - 0)|_{R=R_0} = 0$ (因 $\boldsymbol{P}_2 = 0$),

液体平面 $\sigma_P = -\boldsymbol{e}_\theta \cdot (\boldsymbol{P}_1 - \boldsymbol{P}_2)|_{z=0} = 0$ (因 $\boldsymbol{e}_\theta \cdot \boldsymbol{R} = 0$).

由本题的讨论可见,导体球上、下两半球上自由电荷及束缚电荷分布都不是球对称的,但总的电荷分布($\sigma_f + \sigma_P$)是球对称的.因此,$\boldsymbol{D}$ 与 $\boldsymbol{P}$ 不是球对称的,但 $\boldsymbol{E}$ 是球对称的.由 $\boldsymbol{E} = -\nabla\varphi$ 可见,$\boldsymbol{E}$ 的球对称分布保证导体球表面是等势面.

§3.1.2 分离变量法

现在讨论用分离变量法求解泊松方程的边值问题

$$\begin{cases} \nabla^2\varphi = -\rho_f/\varepsilon, & (2.1) \\ \varphi|_S = f_1(x) \quad 或 \quad \partial\varphi/\partial n|_S = f_2(x) & (2.2) \end{cases}$$

用分离变量法求静电势的解题步骤如下:

1.根据界面的形状选择适当的坐标系,并确定零电势参考点.本节着重讨论用球坐标系处理问题,用柱坐标系处理泊松方程的边值问题留到第二章中介绍.

2.将求解泊松方程的边值问题化为拉普拉斯方程的边值问题,即将定解问题的(2.1)式和(2.2)式化为

$$\begin{cases} \nabla^2\varphi = 0, & (2.3) \\ \varphi|_S = f_1(x) \quad 或 \quad \partial\varphi/\partial n|_S = f_2(x). & (2.4) \end{cases}$$

能够这样做的原因是因为实际问题中的自由电荷往往只分布在导体表面而不在空间中,因此可以将导体从求解区域排除出去,并将导体表面作为求解区域 V 的边界. 这样,在导体以外的区域中的电势将满足拉普拉斯方程. 导体上电荷的作用由边值关系反映,而导体的电势也在求得导体周围介质的电势后利用边值关系 $\varphi_1=\varphi_2$ 求得.

如果作了上述处理后,我们所讨论的区域 V 内仍有自由电荷分布,则可利用电势叠加原理 $\varphi=\varphi_f+\varphi'$ 来处理. 其中 φ_f 是与 V 中自由电荷分布有关的电势(它可以包括由于这些自由电荷分布引起的某些束缚电荷所激发的电势,见例 2);而 φ' 仍满足拉普拉斯方程,它是束缚电荷或 φ' 所在区域 V 外电荷分布所产生的电势. φ_f 通常可以根据经验写出,于是求解 φ 的问题就转化为求解 φ' 的问题,即把求解泊松方程的边值问题转化为求解拉普拉斯方程的边值问题.

3. 根据问题的对称性选择通解的形式

根据本节的讨论,我们可以利用下表选择通解的形式,其中 l,m 为球谐函数的参数,见附录Ⅳ:

对称性	解与坐标的关系	l,m 值	通 解 的 形 式
无对称性	与 R,θ,ϕ 有关	$l>m$	见附录(Ⅳ.35)式
轴对称性	与 ϕ 无关	$m=0$	$\varphi=\sum[a_lR^l+b_lR^{-(l+1)}]P_l(\cos\theta)$
球对称性	与 θ,ϕ 无关	$l=m=0$	$\varphi=a+b/R$

4. 由边界条件与边值关系定系数,求电势.

边界条件与边值关系务必写齐全,系数才能全部定出来.

边界条件是指讨论区域的边界给定的 $\varphi|_S$ 或 $\partial\varphi/\partial n|_S$. 如果讨论的是全空间,往往要利用 $R\to0$ 及 $R\to\infty$ 时 φ 有限的条件. 因为,除了点电荷所在点之外,静电势应处处有限. 一个例外是全空间均匀的外场,当静电势的零参考点取在有限距离处时,在 $R\to\infty$ 时电势自然不可能保持有限,这时 $R\to\infty$ 处的边界条件见例 1.

边值关系根据不同情况采用 §3.1.1(1.8) ~ (1.14)式中适当的公式.

系数求出后,代回通解表达式,即可求得电势.

5. 由电势求其他物理量.

下面举几个例子说明解题方法.

【例 1】 在均匀外电场中置入半径为 R_0 的导体球,试求下列两种情况下的电势:1. 导体球接在电池上使球与地保持电势差 Φ_0;2. 导体球所带总电荷为 Q_0.

解 1. **导体球接电池**

(a)坐标系与电势零参考点. 选坐标原点在球心,z 轴沿 $\boldsymbol{E}_0$ 方向的球坐标系,

如图 1. 设导体球置入前原点的电势为 0.

(*b*)方程. 除球面上有自由电荷分布外,球内电势 φ_1 及球外电势 φ_2 均满足拉普拉斯方程 $\nabla^2\varphi=0$.

(*c*)对称性及通解形式. 本问题有轴对称性,通解为:

$$\text{球内}\quad \varphi_1 = \Phi_0, \tag{2.5}$$

$$\text{球外}\quad \varphi_2 = \sum_l [a_l R^l + b_l R^{-(l+1)}]\mathrm{P}_l(\cos\theta). \tag{2.6}$$

(*d*)边界条件与边值关系.

①无穷远处电势不变.

$$\varphi_2 \mid_{R\to\infty} = -E_0 R\cos\theta. \tag{2.7}$$

因为在无穷远处,导体上电荷对电势的贡献可以忽略,即 $R\to\infty$ 处的电势与导体置入前的电势

$$\varphi_2(R) = -\int_0^R \boldsymbol{E}_0\cdot \mathrm{d}\boldsymbol{l} = -E_0 R\cos\theta$$

相同. 只要 $R\to\infty$,此公式对导体置入前后均正确.

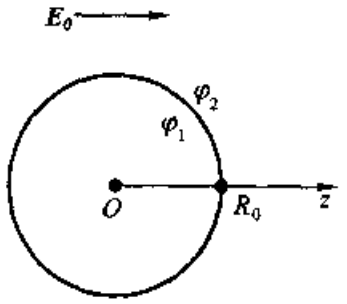


图 1 均匀外场中的导电球

②界面上的电势连续.

$$\varphi_2(R_0) = \varphi_1(R_0) = \Phi_0. \tag{2.8}$$

(*e*)定系数,得电势. 将(2.7)式代入(2.6)式,得

$$\sum_l a_l R^l \mathrm{P}_l(\cos\theta) = -E_0 R\cos\theta,$$

因为 $\mathrm{P}_1(\cos\theta)=\cos\theta$,上式可写为

$$\sum_l a_l R^l \mathrm{P}_l(\cos\theta) = -E_0 R\mathrm{P}_1(\cos\theta) \tag{2.9}$$

分别用 $\mathrm{P}_1(\cos\theta)$ 及 $\mathrm{P}_k(\cos\theta)$ $(k\neq 1)$ 乘上式,然后从 -1 到 $+1$ 对 $x=\cos\theta$ 积分,利用勒让德多项式的正交归一性,附录(Ⅳ.27)式,可得

$$l=1\text{ 时},\quad a_1 R = -E_0 R,\quad \text{即}\quad a_1 = -E_0;$$

$$l\neq 1\text{ 时},\quad a_l R^l = 0,\quad \text{即}\quad a_l = 0\quad (\text{当 } l\neq 1).$$

实际上,勒让德多项式的正交性使得(2.9)式两边 $\mathrm{P}_l(\cos\theta)$ 的系数相等.

将 a_l 代入(2.8)式,即有

$$-E_0 R_0\cos\theta + \sum_l b_l R_0^{-(l+1)}\mathrm{P}_l(\cos\theta) = \Phi_0.$$

同理,利用 $\mathrm{P}_l(\cos\theta)$ 的正交性可得:$b_0=\Phi_0 R_0$,$b_1=E_0 R_0^3$,$b_l=0(l\neq 0,1)$. 将 a_l 及 b_l 代入(2.6)式,可得:

$$\varphi_1 = \Phi_0,$$

$$\varphi_2 = -E_0 R\cos\theta + \Phi_0 R_0/R + E_0 R_0^3\cos\theta/R^2$$

其中,

$-E_0R\cos\theta$ 是导体置入前 R 点的电势；

Φ_0R_0/R 是导体及电池置入后，电池给导体充以电量 $Q=4\pi\varepsilon_0\Phi_0R_0$，带电导体球在球外距球心 R 处激发的附加电势；

$E_0R_0^3\cos\theta/R^2=\boldsymbol{R}_0^3\boldsymbol{E}_0\cdot\boldsymbol{R}/R^3$ 是导体置入后，在电场作用下使导体产生了感应电荷，这感应电荷的电偶极矩在球外各点所激发的附加电势.

2. 导体球所带的总电荷为 Q_0

其通解形式与 1. 相同，仅将边界条件与边值关系归纳如下：

$$\varphi_2(\boldsymbol{R})\mid_{R\to\infty}=-E_0R\cos\theta, \tag{2.10}$$

$$\varphi_2(R_0)=\varphi_1(R_0)=\Phi_0(\text{未知的待定常数}), \tag{2.11}$$

$$-\oint(\partial\varphi_2/\partial n)\,\mathrm{d}\sigma=Q_0/\varepsilon_0. \tag{2.12}$$

仿前可得

$$\varphi_2=-E_0R\cos\theta+\Phi_0R_0/R+E_0R_0^3\cos\theta/R^2.$$

由(2.12)式可确定 Φ_0，利用

$$\partial\varphi_2/\partial n=\partial\varphi_2/\partial R=-E_0\cos\theta-\Phi_0R_0/R^2-2E_0R_0^3\cos\theta/R^3.$$

代入(2.12)式，因面积分在 $R=R_0$ 上进行，故有

$$-\int_0^{2\pi}\mathrm{d}\phi\int_0^{\pi}\left[-3E_0\cos\theta-\frac{\Phi_0}{R_0}\right]R_0^2\sin\theta\,\mathrm{d}\theta=\frac{Q_0}{\varepsilon}.$$

积分后可得 $\Phi_0=Q_0/(4\pi\varepsilon_0R_0)$，故

$$\varphi_1=Q_0/(4\pi\varepsilon_0R_0),$$

$$\varphi_2=-E_0R\cos\theta+Q_0/(4\pi\varepsilon_0R)+E_0R_0^3\cos\theta/R^2.$$

各项的物理意义可仿前进行讨论.

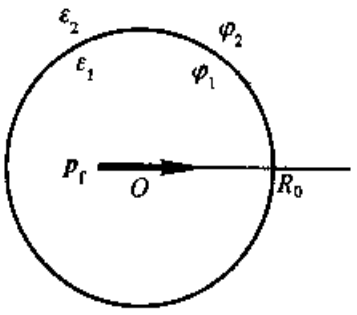


图 2 自由电偶极子位于介质球中心

*【例 2】 介电常数为 ε_1 的均匀介质球的球心置一自由电偶极子 $\boldsymbol{p}_f$，如图 2，球外充满介电常数为 ε_2 的介质，求空间各点的电势和极化电荷分布.

解 (a)坐标系与电势零参考点. 取坐标原点在球心，z 轴沿 $\boldsymbol{p}_f$ 方向的球坐标系，电势零点取在 ∞.

(b)方程. 根据电势叠加原理，球内外各点的电势可看成电偶极子激发的电势 $\varphi_f=\boldsymbol{p}_f\cdot\boldsymbol{R}/(4\pi\varepsilon_1R^3)$ 与球面上面极化电荷 σ_P 激发的电势 φ' 之和：

$$\varphi=\varphi'+\varphi_f=\varphi'+\boldsymbol{p}_f\cdot\boldsymbol{R}/(4\pi\varepsilon_1R^3). \tag{2.13}$$

显然，面极化电荷 σ_P 在介质球内外激发的电势满足拉普拉斯方程

$$\nabla^2\varphi'=0.$$

(c)对称性与通解的形式. 本题有轴对称性，故通解形式为：

$$\text{球内：}\quad \varphi_1=\sum_l[a_lR^l+b_lR^{-(l+1)}]P_l(\cos\theta)+\boldsymbol{p}_f\cdot\boldsymbol{R}/(4\pi\varepsilon_1R^3), \tag{2.14}$$

$$\text{球外：}\quad \varphi_2 = \sum_l [c_l R^l + d_l R^{-(l+1)}] P_l(\cos\theta) + (\boldsymbol{p}_l \cdot \boldsymbol{R})/(4\pi\varepsilon_1 R^3). \tag{2.15}$$

值得注意的是，球外电势本来就满足拉普拉斯方程（因为球外没有自由电荷），所以(2.15)式是否要第二项是无关紧要的. 有这一项在定系数时较方便. 但应注意，如果不加第二项，算出的 d_l 会不相同.

(*d*)边界条件与边值关系.

①电荷分布在有限区域内，故取

$$\varphi_2 |_{R\to\infty} = 0. \tag{2.16}$$

②由坐标原点的电势 $\varphi_1|_{R=0}$ 有限，而中心位于坐标原点的电偶极子对坐标原点的电势没有贡献，故束缚电荷对坐标原点的电势的贡献：

$$\varphi_1 |_{R\to 0} = \sum_l [a_l R^l + b_l R^{-(l+1)}] P_l(\cos\theta)|_{R\to 0} \text{ 有限}, \tag{2.17}$$

③在球面上电势连续：

$$\varphi_1 |_{R=R_0} = \varphi_2 |_{R=R_0}, \tag{2.18}$$

④在球面上没有自由电荷分布，故

$$\varepsilon_1 \partial\varphi_1/\partial R |_{R=R_0} = \varepsilon_2 \partial\varphi_2/\partial R |_{R=R_0}, \tag{2.19}$$

(*e*)定系数，求电势. 将(2.15)式代入(2.16)式，有

$$\varphi_2 |_{R\to\infty} = \sum_l c_l R^l P_l(\cos\theta)|_{R\to\infty} = 0.$$

由勒让德多项式的正交性可得 $c_l = 0$，故

$$\varphi_2 = \sum_l d_l R^{-(l+1)} P_l(\cos\theta) + \varphi_f. \tag{2.20}$$

由(2.17)式，当 $R\to 0$ 时 $a_l R^l \to 0$，故

$$\sum_l b_l R^{-(l+1)} P_l(\cos\theta)|_{R\to 0} \text{ 有限},$$

可见 $b_l = 0$，代入(2.14)式可得

$$\varphi_1 = \sum_l a_l R^l P_l(\cos\theta) + \varphi_f, \tag{2.21}$$

将(2.20)式和(2.21)式代入(2.18)式，并由 $P_l(\cos\theta)$ 的正交性得

$$a_l = d_l/R_0^{2l+1}. \tag{2.22}$$

将(2.20)式及(2.21)式代入(2.19)式得

$$\varepsilon_1\left[\sum_l l a_l R_0^{l-1} P_l(\cos\theta) - \frac{2p_l\cos\theta}{4\pi\varepsilon_1 R_0^3}\right] = \varepsilon_2\left[-\sum_l \frac{(l+1)d_l P_l(\cos\theta)}{R_0^{l+2}} - \frac{2p_l\cos\theta}{4\pi\varepsilon_1 R_0^3}\right].$$

由 $P_1(\cos\theta) = \cos\theta$ 及 $P_l(\cos\theta)$ 的正交性即有：

$$\varepsilon_1 a_1 - \frac{p_l}{2\pi R_0^3} = -\frac{2\varepsilon_2 d_1}{R_0^3} - \frac{\varepsilon_2 p_l}{2\pi\varepsilon_1 R_0^3}, \tag{2.23}$$

$$l\varepsilon_1 a_l R_0^{l-1} = -(l+1)\varepsilon_2 d_l/R_0^{l+2} \quad (l \neq 1). \tag{2.24}$$

由(2.22)式及(2.24)式得

$$a_l = d_l = 0 \quad (l \neq 1). \tag{2.25}$$

由(2.22)式及(2.23)式得

$$a_1 = d_f/R_0^3 = (\varepsilon_1 - \varepsilon_2)p_f/[2\pi\varepsilon_1(\varepsilon_1 + 2\varepsilon_2)R_0^3]. \tag{2.26}$$

将系数代入通解,即有

$$\text{球内:}\quad \varphi_1 = \frac{(\varepsilon_1 - \varepsilon_2)\boldsymbol{p}_f \cdot \boldsymbol{R}}{2\pi\varepsilon_1(\varepsilon_1 + 2\varepsilon_2)R_0^3} + \frac{\boldsymbol{p}_f \cdot \boldsymbol{R}}{4\pi\varepsilon_1 R^3},$$

$$\text{球外:}\quad \varphi_2 = \frac{3\boldsymbol{p}_f \cdot \boldsymbol{R}}{4\pi(\varepsilon_1 + 2\varepsilon_2)R^3}.$$

(*f*)求极化电荷面密度. 现在再来计算在 $R = R_0$ 球面上的极化电荷面密度. 对 $\nabla \cdot (\varepsilon_0 \boldsymbol{E}) = \nabla \cdot (\boldsymbol{D} - \boldsymbol{P}) = \rho_f + \rho_P$ 进行积分,可得

$$\varepsilon_0 \oint \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{\sigma} = Q_f + Q_P.$$

在球面上取一足够小的表面积 $\Delta\sigma$,作一个上下底与 $\Delta\sigma$ 平行的扁平盒作为高斯面,可得

$$\varepsilon_0 \boldsymbol{e}_n \cdot (\boldsymbol{E}_2 - \boldsymbol{E}_1) = \sigma_f + \sigma_P.$$

在球面上没有自由电荷分布,即 $\sigma_f = 0$. 再将 $\boldsymbol{e}_n \cdot \mathrm{E} = -\partial\varphi/\partial R$ 代入,即有

$$\begin{aligned}\sigma_P &= \varepsilon_0[-\partial\varphi_2/\partial R + \partial\varphi_1/\partial R]\big|_{R=R_0} \\ &= \varepsilon_0[-\partial\varphi'_2/\partial R + \partial\varphi'_1/\partial R]\big|_{R=R_0} \\ &= (2\varepsilon_0 d_1 \cos\theta/R_0^3) + \varepsilon_0 a_1 \cos\theta.\end{aligned}$$

将(2.26)式代入即有

$$\sigma_P = 3\varepsilon_0 a_1 \cos\theta = 3\varepsilon_0(\varepsilon_1 - \varepsilon_2)p_f \cos\theta/[2\pi\varepsilon_1(\varepsilon_1 + 2\varepsilon_2)R_0^3].$$

(*g*)求极化电荷体密度 ρ_P. 中心位于坐标原点的电偶极子 $\boldsymbol{p}_f = q\boldsymbol{l}$ 的电荷密度为

$$\rho_f = -q\delta^3[\boldsymbol{x} + \boldsymbol{e}_z l/2] + q\delta^3[\boldsymbol{x} - \boldsymbol{e}_z l/2],$$

因面极化电荷密度为

$$\rho_P = \left(\frac{\varepsilon_0}{\varepsilon_1} - 1\right)\rho_f = -\left(\frac{\varepsilon_0}{\varepsilon_1} - 1\right)q\delta^3\left[\boldsymbol{x} + \boldsymbol{e}_z \frac{l}{2}\right] + \left(\frac{\varepsilon_0}{\varepsilon_1} - 1\right)q\delta^3\left[\boldsymbol{x} - \boldsymbol{e}_z \frac{l}{2}\right].$$

它相当于中心位于坐标原点的电偶极矩为 $\boldsymbol{p}_P = -(1 - \varepsilon_0/\varepsilon_1)\boldsymbol{p}_f$ 的电偶极子.

容易证明,我们原先给出的 $\varphi_f = \boldsymbol{p}_f \cdot \boldsymbol{R}/(4\pi\varepsilon_1 R^3)$ 正是自由电荷电偶极子 $\boldsymbol{p}_f$ 与束缚电荷电偶极子 $\boldsymbol{p}_P$ 的共同贡献. 我们仍将介质看成真空中的一个电荷体系,则 $\boldsymbol{p}_f$ 与 $\boldsymbol{p}_P$ 共同激发的电势即为

$$\varphi_f = \frac{(\boldsymbol{p}_f + \boldsymbol{p}_P) \cdot \boldsymbol{R}}{4\pi\varepsilon_0 R^3} = \frac{[1 - (1 - \varepsilon_0/\varepsilon_1)]\boldsymbol{p}_f \cdot \boldsymbol{R}}{4\pi\varepsilon_0 R^3} = \frac{\boldsymbol{p}_f \cdot \boldsymbol{R}}{4\pi\varepsilon_1 R^3}.$$

§3.1.3 电 象 法

电象法是在某些特殊情形下求解泊松方程边值问题的一个比较简便的方法. 它常应用于区域内仅有一个或几个点电荷,区域的边界是导体或介质并具有一定的对称性的情形. 下面分析几个例子,这些例子颇具代表性,不少较为复杂的问题均可看成是这几个例子的组合.

【例 1】 与无穷大接地导体平板距离为 a 处的 A 点有一个点电荷 $+Q$，如图 1. 求空间电势分布.

解 先分析一下本题的电荷分布及电场分布. 点电荷 $+Q$ 的存在使导体上表面出现异号感应电荷，整个电场就是由 $+Q$ 及这些负的感应电荷所共同产生的. 同理，如果不是导体平板而是介质板，则整个电场就是由电荷 $+Q$ 与介质板表面的束缚电荷所产生的.

电象法的基本思想就是通过引进一些假想的电荷（称为象电荷）来取代这些感应电荷或束缚电荷的作用，从而把求解区域的电场看作是原电荷的电场与象电荷电场的叠加.

现在通过上面的例子说明如何引入象电荷. 显然，正确引入象电荷是成功地运用电象法的关键.

首先，选坐标原点在 A 对导体板的垂足处，z 轴沿 OA 方向，即 Q 的坐标为 $(0,0,a)$. 因导体板接地，设其电势为零.

其次，分析电场分布并引入象电荷. 由于 $+Q$ 的作用，在导体上表面出现了负的感应电荷，即电场线将由 $+Q$ 发出而终止于导体表面（及无穷远处）. 而由导体的静电平衡条件可知，导体表面是等势面，电场线应垂直于等势面，也就是垂直于导体板表面；而在导体板下方，由于导体已接地，没有电场. 于是可以估计电场线的分布大体上如图 1 的实线所示.

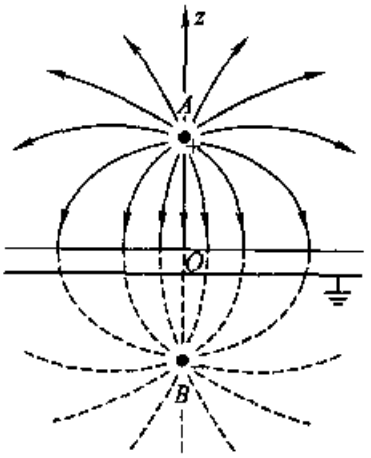


图 1 无穷大接地导体平板

将图 1 实线的电场分布与拿走导体板，但在 A 点仍置电荷 $+Q$，在 B 点（它是 A 关于导体板上表面的镜象位置）置电荷 $-Q$ 的电场分布相比较，可以看出，在 $z>0$ 区域两者的电场很有可能相同（在 $z<0$ 区域当然不同）.

因此，在求解 $z>0$ 区域的电场时，拿走导体板而用 B 点的象电荷 $-Q$ 取代导体上的感应电荷的作用，提出尝试解

$$\varphi = \frac{Q}{4\pi\varepsilon_0 r_+} + \frac{-Q}{4\pi\varepsilon_0 r_-} \quad (z>0), \tag{3.1}$$

其中，由 $+Q$ 到场点 (x,y,z) 的距离为

$$r_+ = \sqrt{x^2+y^2+(z-a)^2}; \tag{3.2}$$

由 $-Q$ 到场点 (x,y,z) 的距离为

$$r_- = \sqrt{x^2+y^2+(z+a)^2}. \tag{3.3}$$

而对 $z<0$ 区域，不必引入象电荷，直接由导体的静电屏蔽作用可知

$$\varphi = 0 \quad (z<0). \tag{3.4}$$

由(3.1)式与(3.4)式给出的尝试解是否正确,要用静电问题的唯一性定理来回答.如果(3.1)式和(3.4)式满足泊松方程、边界条件和边值关系,由唯一性定理可知,它就是正确的解.

为此,我们先分区写出定解条件.对于 $z>0$ 区域,定解条件为:

$$\begin{cases}\nabla^2\varphi = -(Q/\varepsilon_0)\delta^3(x,y,z-a), & (3.5)\\ \varphi|_{z=0} = 0, & (3.6)\\ \varphi|_{R\to\infty} = 0. & (3.7)\end{cases}$$

为了验证尝试解(3.1)确实满足(3.5)、(3.6)、(3.7)式,需要用到 δ 函数的公式[见附录(Ⅰ.23)式]

$$\nabla^2(1/r) = -4\pi\delta^3(\boldsymbol{r}).$$

公式表明,若 $r\neq 0$,则 $\nabla^2(1/r)=0$.因此用拉普拉斯算符作用于(3.1)时,第二项 $\nabla^2(1/r_-)=0$,故

$$\nabla^2\varphi = \frac{Q}{4\pi\varepsilon_0}\left(\nabla^2\frac{1}{r_+}-\nabla^2\frac{1}{r_-}\right) = -\frac{Q}{\varepsilon_0}\delta^3(r_+) = -\frac{Q}{\varepsilon_0}\delta^3(x,y,z-a).$$

对于 $z<0$ 区域,定解条件为

$$\begin{cases}\nabla^2\varphi = 0, & (3.8)\\ \varphi|_{z=0} = 0, & (3.9)\\ \varphi|_{R\to\infty} = 0. & (3.10)\end{cases}$$

显然,尝试解(3.4)式必能满足这些条件,这样就证明了尝试解(3.1)、(3.4)式是正确的解.

这种求解静电场的方法就称为电象法,因为所引入的象电荷是在原电荷关于导体表面的镜象位置引入的,而且这样一对位于镜象位置的电荷的符号是相反的.应用电象法的关键是正确地引入象电荷,这样才可能提出正确的尝试解.从上面的例子可见,引入象电荷要遵守两个原则:一是象电荷必须在求解的区域之外引入.如此例求解区域为 $z>0$ 区域,其象电荷只能在 $z<0$ 区域引入.否则必将改变求解区域的电荷分布 $\rho_f(\boldsymbol{x})$,违反泊松方程 $\nabla^2\varphi=-\rho_f/\varepsilon_0$.二是象电荷的位置与电量必须满足边界条件与边值关系.此例将 $-Q$ 引入到 A 点关于导体上表面的镜象位置 B 处,并且象电荷与原电荷的电量大小相等、符号相反,因而 $+Q$ 和 $-Q$ 到 $z=0$ 平面上任一点的距离相等,从面保证了 $\varphi|_{z=0}=0$ 的要求;此外,既然电场是由两个点电荷激发的,$\varphi|_{R\to\infty}=0$ 的要求自然也满足.由此可见,如果象电荷电量的大小、符号及位置不适当,会违背边界条件与边值关系的要求.上例比较简单,可以一目了然地看出象电荷应有的大小、符号和位置;如果问题比较复杂,往往就要先引入一些反映象电荷电量与位置的参数,然后通过边界条件和边值关系来确定这些参数.

【例2】 真空中有一半径为 R_0 的接地导体球,距球心为 $a(a>R_0)$ 处有一个

点电荷 Q,求空间各点的电势.

解 (*a*)坐标系与电势零参考点.取坐标原点在球心,z 轴沿 OQ 方向的球坐标系.设导体球电势为零.

(*b*)定解问题.

球内区域($R<R_0$):

$$\begin{cases} \nabla^2\varphi_1 = 0, & (3.11)\\ \varphi_1\big|_{R=R_0} = 0. & (3.12) \end{cases}$$

球外区域($R>R_0$):

$$\begin{cases} \nabla^2\varphi_2 = -(Q/\varepsilon_0)\delta^3(x,y,z-a), & (3.13)\\ \varphi_2\big|_{R=R_0} = 0, & (3.14)\\ \varphi_2\big|_{R\to\infty} = 0, & (3.15)\\ \varepsilon_0\partial\varphi_2/\partial n\big|_{R=R_0} = -\sigma_f. & (3.16) \end{cases}$$

(*c*)引入象电荷,提出尝试解.

在球内区域,由于导体的静电屏蔽作用可知,$\varphi_1=0$ 满足定解条件(3.11)、(3.12)式.由唯一性定理,这是正确的解.

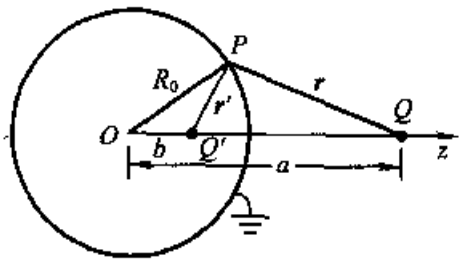


图2 接地导体球

为了求球外区域的电势,需要引入象电荷.这里电场线是由 $+Q$ 发出而终止于导体球面上感应产生的负电荷(及无穷远)处.由电场线的形状及其趋势可以判断,在球内引入一个负电荷 Q' 是适当的.让我们再次强调,在求解球外区域电势时,象电荷只能在球内区域引入.由对称性可以判断,负电荷位于 z 轴上是合理的.但具体 $Q'=?$ 以及 Q' 距球心的距离 $b=?$ 则难以判断,要根据边界条件来确定.由此可提出尝试解

$$\varphi_2 = Q/(4\pi\varepsilon_0 r) + Q'/(4\pi\varepsilon_0 r'), \qquad (3.17)$$

其中

$$\text{电荷 } Q \text{ 到场点距离:}\quad r=\sqrt{x^2+y^2+(z-a)^2}, \qquad (3.18)$$

$$\text{电荷 } Q' \text{ 到场点距离:}\quad r'=\sqrt{x^2+y^2+(z-b)^2}. \qquad (3.19)$$

(*d*)将尝试解代入定解条件,确定待定参数而得解.

将(3.17)式代入(3.14)式,即可求得 Q' 及 b.

$$\varphi_2\big|_{R=R_0} = \left[Q/(4\pi\varepsilon_0 r) + Q'/(4\pi\varepsilon_0 r')\right]\Big|_{R=R_0} = 0.$$

即对球面上任一点 P,应有

$$\frac{Q}{r}+\frac{Q'}{r'}=0,\quad \text{即}\quad \frac{\overline{PQ'}}{\overline{PQ}}=\frac{r'}{r}=-\frac{Q'}{Q}=\text{常数}. \qquad (3.20)$$

显然，若令 $\triangle OQ'P \backsim \triangle OPQ$，则 $\overline{Q'P}/\overline{PQ} = \overline{OQ'}/\overline{OP} = \overline{OP}/\overline{OQ}$，即

$$r'/r = b/R_0 = R_0/a, \tag{3.21}$$

$$b = R_0^2/a. \tag{3.22}$$

将(3.21)式代入(3.20)式得：

$$Q' = -R_0 Q/a. \tag{3.23}$$

由于 R_0，a，Q 已给定，故象电荷的位置 b 及电量 Q' 均已确定. 将(3.22)、(3.23)式代入(3.17)式，即有(注意，观察点 P 可在球面上或球面外)

$$\varphi_2 = \frac{Q}{4\pi\varepsilon_0\sqrt{x^2+y^2+(z-a)^2}} + \frac{-R_0 Q/a}{4\pi\varepsilon_0\sqrt{x^2+y^2+(z-R_0^2/a)^2}}. \tag{3.24}$$

不难证明，由(3.14)式导出的(3.24)式满足(3.13)式与(3.15)式，所以是正确的解. 将(3.24)式代入(3.16)式，即可求得导体表面的电荷分布.

本题求解的四步是电象法解题的一般步骤.

运用电象法解题的关键是要掌握前面提到的引入象电荷的两个原则. 本节讨论了无穷大导体平板与导体球附近一个点电荷所产生的电场. 这两类问题比较简单，但是它是讨论许多复杂问题的基础. 有的问题需要引入许多个甚至无穷多个象电荷才能完全满足定解条件. 例如，在两块互相平行的无穷大接地导体板之间的任意位置上置一点电荷，就要引入无穷多个象电荷. 这时，电势将被表示为一个无穷级数. 可以证明，这样得到的无穷级数是收敛的，并且通常只要取它的前三、四项就足够精确.

§3.1.4 格林函数法

上节利用电象法求出了一个点电荷激发的、在一定边界条件下的静电场. 但电象法有局限性，即当场源是一个或若干个点电荷时才能奏效. 如果场源为连续分布的电荷则无能为力. 由于连续分布的电荷可以看成点电荷的叠加，利用叠加原理就可以由一个点电荷的场求得同样边界条件下任意电荷分布的场. 这种求解静电场的方法称为格林函数法，点电荷(单位正电荷)所激发的电势称为格林函数.

(一)泊松方程的格林函数

泊松方程的边值问题可表示为：

$$\begin{cases} \nabla^2\varphi(\boldsymbol{x}) = -\rho_f(\boldsymbol{x})/\varepsilon_0, & (4.1) \\ [\alpha\partial\varphi(\boldsymbol{x})/\partial n + \beta\varphi(\boldsymbol{x})]\big|_S = f(\boldsymbol{x}). & (4.2) \end{cases}$$

当 $\alpha=0$ 时为第一类边界条件；$\beta=0$ 为第二类边界条件；α、β 均不为零为第三类边界条件. 当 $f(\boldsymbol{x})=0$ 时为齐次边界条件.

格林函数 $G(x,x')$是位于 x'的单位正电荷所激发的、满足齐次边界条件的电势.

泊松方程的三类边值问题对应三类格林函数,它们的定解问题为

$$\begin{cases} \nabla^2 G(x,x') = -\delta^3(x-x')/\varepsilon_0, & (4.3) \\ [\alpha \partial G(x,x')/\partial n + \beta G(x,x')]\big|_S = 0. & (4.4) \end{cases}$$

但是,应该指出,对于第二类齐次边界条件,当研究区域为有限时,这个定解问题没有解[1][2]. 实际上,方程(4.3)表明 V 内有单位正电荷存在. 第二类齐次边界条件 $\partial G(x,x')/\partial n\big|_S = 0$,表示点源产生的场在边界 S 上电场强度的法向分量 $E_n = -\partial G(x,x')/\partial n$ 处处为零,因而方程与边界条件是不相容的. 解决办法是引入广义格林函数,它满足的方程及边界条件为

$$\begin{cases} \nabla^2 G(x,x') = -\dfrac{1}{\varepsilon_0}\delta^3(x-x') + \dfrac{1}{\varepsilon_0 V}, & (4.5) \\ \left.\dfrac{\partial G(x,x')}{\partial n}\right|_S = 0. & (4.6) \end{cases}$$

总之,通过求解(4.3)、(4.4)式或(4.5)、(4.6)式即可求得在给定区域 V 的泊松方程的各类边值问题的格林函数.

(二)常见的泊松方程第一类边值问题的格林函数

1. 无界空间的格林函数.

定解问题为

$$\begin{cases} \nabla^2 G(x,x') = -\delta^3(x-x')/\varepsilon_0, & (4.7) \\ G(x,x')\big|_{R\to\infty} = 0. & (4.8) \end{cases}$$

这个定解问题是下述物理问题的数学表示:单位正电荷 $Q=1$ 置于 $x' = x'e_x + y'e_y + z'e_z$ 处,求空间 $x = xe_x + ye_y + ze_z$ 处的电势 $G(x,x') = ?$ 这个问题的解就是库仑定律:

$$\begin{aligned} G(x,x') &= \frac{1}{4\pi\varepsilon_0 |x-x'|} \\ &= \frac{1}{4\pi\varepsilon_0 \sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2}}. \end{aligned} \quad (4.9)$$

这个无界区域的格林函数又称为泊松方程的基本解.

2. 上半空间的格林函数.

定解问题为

---

[1] 郭敦仁. 数学物理方法. 人民教育出版社,1965,343~346.

[2] [美]J. D. 杰克逊. 经典电动力学. 朱培豫译,人民教育出版社,1980,47~50.

$$\begin{cases} \nabla^2 G(\boldsymbol{x},\boldsymbol{x}') = -\delta^3(\boldsymbol{x}-\boldsymbol{x}')/\varepsilon_0, & z>0; \quad (4.10) \\ G(\boldsymbol{x},\boldsymbol{x}')\big|_{z=0} = 0, \quad G(\boldsymbol{x},\boldsymbol{x}')\big|_{R\to\infty} = 0. & (4.11) \end{cases}$$

这个定解问题是下述物理问题的数学表示：在 $z=0$ 处有一无限大的接地金属板，在 $\boldsymbol{x}'$ 处有一单位正电荷 Q，求金属板上方任一点 $\boldsymbol{x}$ 处的电势 $G(\boldsymbol{x},\boldsymbol{x}')$. 上节已利用电象法求解了这个问题. 引入象电荷 $Q'=-1$（如图 1 所示），可得

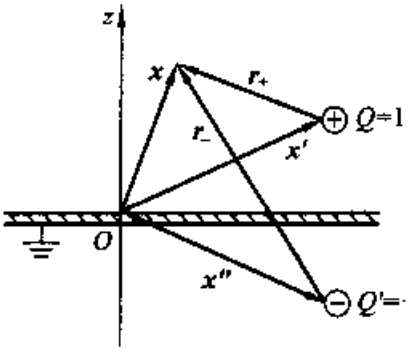


图 1 无穷大接地金属板

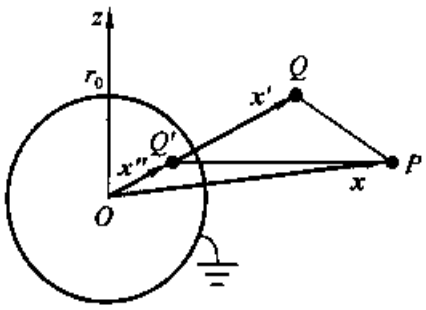


图 2 接地金属球

$$\begin{aligned} G(\boldsymbol{x},\boldsymbol{x}') &= \frac{1}{4\pi\varepsilon_0 r_+} + \frac{-1}{4\pi\varepsilon_0 r_-} \\ &= \frac{1}{4\pi\varepsilon_0}\left[\frac{1}{\sqrt{(x-x')^2+(y-y')^2+(z-z')^2}} - \right. \\ &\quad \left. \frac{1}{\sqrt{(x-x')^2+(y-y')^2+(z+z')^2}}\right]. \end{aligned} \quad (4.12)$$

式中 r_+ 为电荷 Q 到场点 $\boldsymbol{x}$ 的距离，r_- 为象电荷 Q' 到场点 $\boldsymbol{x}$ 的距离.

3. 球外空间的格林函数.

定解问题为

$$\begin{cases} \nabla^2 G(\boldsymbol{x},\boldsymbol{x}') = -\delta^3(\boldsymbol{x}-\boldsymbol{x}')/\varepsilon_0, & R>R_0; \quad (4.13) \\ G(\boldsymbol{x},\boldsymbol{x}')\big|_{R=R_0} = 0, \quad G(\boldsymbol{x},\boldsymbol{x}')\big|_{R\to\infty} = 0. & (4.14) \end{cases}$$

这个定解问题可看作下述物理问题的数学表示：在一个接地金属球外 $\boldsymbol{x}'$ 处有一单位正电荷 Q，求球外空间任一点 $\boldsymbol{x}$ 处的电势 $G(\boldsymbol{x},\boldsymbol{x}')=?$，这个问题在上节例 2 中用电象法中已原则上解决，差别仅在于电荷不在 Z 轴上，场点及源点用球坐标表示如下（见图 2）：

场点位矢 $\boldsymbol{x}=(R,\theta,\phi)$；

单位正电荷 Q 位矢 $\boldsymbol{x}'=(R',\theta',\phi')$；

象电荷 $Q'=-QR_0/R'$ 位矢 $\boldsymbol{x}''=(R'',\theta'',\phi'')=(R_0^2/R',\theta',\phi')$.

令 $\boldsymbol{r}=\boldsymbol{x}-\boldsymbol{x}'$，$\boldsymbol{r}'=\boldsymbol{x}-\boldsymbol{x}''$，则

$$r=\sqrt{R^2+R'^2-2RR'\cos\alpha},$$

$$r'=\sqrt{R^2+R''^2-2RR''\cos\alpha}=\sqrt{R^2+(R_0^2/R')^2-2R(R_0^2/R')\cos\alpha}.$$

格林函数为

$$G(\boldsymbol{x},\boldsymbol{x}') = \frac{1}{4\pi\varepsilon_0}\left(\frac{Q}{r}+\frac{Q'}{r'}\right) = \frac{1}{4\pi\varepsilon_0}\left(\frac{1}{r}-\frac{R_0}{r'R'}\right)$$

$$= \frac{1}{4\pi\varepsilon_0}\left[\frac{1}{\sqrt{R^2+R'^2-2RR'\cos\alpha}} - \frac{1}{\sqrt{(RR'/R_0)^2+R_0^2-2RR'\cos\alpha}}\right], \tag{4.15}$$

式中 $\cos\alpha$ 由下式求得，

$$\begin{aligned} RR'\cos\alpha &= \boldsymbol{x}\cdot\boldsymbol{x}' = xx'+yy'+zz' \\ &= RR'(\sin\theta\cos\phi\sin\theta'\cos\phi' + \sin\theta\sin\phi\sin\theta'\sin\phi' + \cos\theta\cos\theta') \\ &= RR'[\sin\theta\sin\theta'\cos(\phi'-\phi)+\cos\theta\cos\theta'], \end{aligned}$$

约去 RR'，即得

$$\cos\alpha = \sin\theta\sin\theta'\cos(\phi'-\phi)+\cos\theta\cos\theta'. \tag{4.16}$$

(三)格林函数的对称性

$$G(\boldsymbol{x},\boldsymbol{x}') = G(\boldsymbol{x}',\boldsymbol{x}) \tag{4.17}$$

证 设 $\boldsymbol{x}'$ 为源点，$\boldsymbol{x}$ 为场点，定解问题为

$$\begin{cases} \nabla^2 G(\boldsymbol{x},\boldsymbol{x}') = -\delta^3(\boldsymbol{x}-\boldsymbol{x}')/\varepsilon_0, & (4.18) \\ [\alpha\partial G(\boldsymbol{x},\boldsymbol{x}')/\partial n + \beta G(\boldsymbol{x},\boldsymbol{x}')]|_S = 0. & (4.19) \end{cases}$$

设 $\boldsymbol{x}''$ 为源点，$\boldsymbol{x}$ 为场点，定解问题为(对第二类办界条件，仅讨论 $V\to\infty$ 的问题)

$$\begin{cases} \nabla^2 G(\boldsymbol{x},\boldsymbol{x}'') = -\delta^3(\boldsymbol{x}-\boldsymbol{x}'')/\varepsilon_0, & (4.20) \\ [\alpha\partial G(\boldsymbol{x},\boldsymbol{x}'')/\partial n + \beta G(\boldsymbol{x},\boldsymbol{x}'')]|_S = 0. & (4.21) \end{cases}$$

用 $G(\boldsymbol{x},\boldsymbol{x}'')$ 乘(4.18)式后减去 $G(\boldsymbol{x},\boldsymbol{x}')$ 乘(4.20)式，再对 V 积分，并利用 δ 函数的性质，可得

$$\begin{aligned} &\int_V [G(\boldsymbol{x},\boldsymbol{x}'')\nabla^2 G(\boldsymbol{x},\boldsymbol{x}') - G(\boldsymbol{x},\boldsymbol{x}')\nabla^2 G(\boldsymbol{x},\boldsymbol{x}'')]\mathrm{d}^3\boldsymbol{x} \\ &= -\frac{1}{\varepsilon_0}\int_V [G(\boldsymbol{x},\boldsymbol{x}'')\delta^3(\boldsymbol{x}-\boldsymbol{x}') - G(\boldsymbol{x},\boldsymbol{x}')\delta^3(\boldsymbol{x}-\boldsymbol{x}'')]\mathrm{d}^3\boldsymbol{x} \\ &= -\frac{1}{\varepsilon_0}[G(\boldsymbol{x}',\boldsymbol{x}'') - G(\boldsymbol{x}'',\boldsymbol{x}')]. \end{aligned} \tag{4.22}$$

利用体积分化面积分的公式 $\int_V \nabla\cdots\mathrm{d}V = \oint_S \cdots \mathrm{d}\boldsymbol{\sigma}$ [见附录(Ⅱ.23)式]

$$\begin{aligned} \int_V(\varphi\nabla^2\psi - \psi\nabla^2\varphi)\mathrm{d}V &= \int_V \nabla\cdot(\varphi\nabla\psi - \psi\nabla\varphi)\mathrm{d}V \\ &= \oint_S(\varphi\nabla\psi - \psi\nabla\varphi)\cdot\mathrm{d}\boldsymbol{\sigma}, \end{aligned}$$

这称为格林公式. 再利用 $\nabla\psi\cdot\mathrm{d}\boldsymbol{\sigma} = \nabla\psi\cdot\boldsymbol{e}_n\mathrm{d}\sigma = (\partial\psi/\partial n)\mathrm{d}\sigma$，得到

$$\int_V(\varphi\nabla^2\psi - \psi\nabla^2\varphi)\mathrm{d}V = \oint_S\left(\varphi\frac{\partial\psi}{\partial n} - \psi\frac{\partial\varphi}{\partial n}\right)\mathrm{d}\sigma. \tag{4.23}$$

利用(4.23)式可将(4.22)式写为

$$-\frac{1}{\varepsilon_0}[G(x',x'')-G(x'',x')]$$

$$=\oint_S\left[G(x,x'')\frac{\partial G(x,x')}{\partial n}-G(x,x')\frac{\partial G(x,x'')}{\partial n}\right]\mathrm{d}\sigma. \tag{4.24}$$

另一方面,用 $G(x,x'')$ 乘(4.19)式后减去 $G(x,x')$ 乘(4.21)式得

$$\left[G(x,x'')\frac{\partial G(x,x')}{\partial n}-G(x,x')\frac{\partial G(x,x'')}{\partial n}\right]\Big|_S=0,$$

代入(4.24)式便证明 $G(x',x'')=G(x'',x')$,将 x'' 换为 x 就是(4.17)式.【证毕】

格林函数的对称性有明显的物理意义:在相同的齐次边界条件下,位于 x' 的点源在 x 点激发的场等于位于 x 的点源在 x' 点激发的场.

(四)用格林函数法求解泊松方程的第一类边值问题

泊松方程的第一类边值问题为

$$\begin{cases}\nabla^2\varphi(x)=-\rho_f(x)/\varepsilon_0, & (4.25)\\ \varphi(x)|_S=f(x). & (4.26)\end{cases}$$

1. 写出与 $\varphi(x)$ 有相同边界且属于同类边值问题的格林函数所满足的方程与边界条件

$$\begin{cases}\nabla^2G(x,x')=-\delta^3(x-x')/\varepsilon_0, & (4.27)\\ G(x,x')|_S=0. & (4.28)\end{cases}$$

2. 写出自变量为 x' 的格林公式

$$\int_V[\psi(x')\nabla'^2\varphi(x')-\varphi(x')\nabla'^2\psi(x')]\mathrm{d}V'$$

$$=\oint_S\left[\psi(x')\frac{\partial\varphi(x')}{\partial n'}-\varphi(x')\frac{\partial\psi(x')}{\partial n'}\right]\mathrm{d}\sigma'.$$

3. 令 $\varphi(x')$ 为待求电势,$\psi(x')=G(x',x)$,便有

$$\int_V[G(x',x)\nabla'^2\varphi(x')-\varphi(x')\nabla'^2G(x',x)]\mathrm{d}V'$$

$$=\oint_S\left[G(x',x)\frac{\partial\varphi(x')}{\partial n'}-\varphi(x')\frac{\partial G(x',x)}{\partial n'}\right]\mathrm{d}\sigma'.$$

将(4.25)、(4.26)、(4.27)、(4.28)式代入上式,便有

$$-\frac{1}{\varepsilon_0}\int_V[G(x',x)\rho_f(x')-\varphi(x')\delta^3(x'-x)]\mathrm{d}V'=-\oint_S f(x')\frac{\partial G(x',x)}{\partial n'}\mathrm{d}\sigma'.$$

利用 δ 函数的性质及格林函数的对称性,得到

$$\varphi(x)=\int_V G(x,x')\rho_f(x')\mathrm{d}V'-\varepsilon_0\oint_S f(x')\frac{\partial G(x,x')}{\partial n'}\mathrm{d}\sigma', \tag{4.29}$$

上式右边第一项反映 V 内电荷分布激发的势，第二项反映 V 外电荷分布激发的势.

【例 1】 在接地无穷大金属平板上方有电荷分布 $\rho_f(x,y,z)$，求平板上方的电势分布.

解 取金属平板为 $z=0$ 平面，其定解问题为

$$\begin{cases}\nabla^2\varphi(\boldsymbol{x}) = -\rho_f(x,y,z)/\varepsilon_0, & z>0;\\ \varphi(\boldsymbol{x})\big|_{z=0}=0, \quad \varphi(\boldsymbol{x})\big|_{R\to\infty}=0.\end{cases}$$

1. 求与 $\varphi(\boldsymbol{x})$ 有相同边界且属于同类边值问题的格林函数 $G(\boldsymbol{x}',\boldsymbol{x})$，如(4.12)式所示.

2. 将 $G(\boldsymbol{x}',\boldsymbol{x})$，$\rho_f(\boldsymbol{x}')$ 及 $\varphi(\boldsymbol{x}')$ 代入(4.29)式，即得

$$\begin{aligned}\varphi(\boldsymbol{x}) &= \int_V G(\boldsymbol{x},\boldsymbol{x}')\rho_f(\boldsymbol{x}')\,\mathrm{d}V'\\ &= \frac{1}{4\pi\varepsilon_0}\int_{-\infty}^{\infty}\mathrm{d}x'\int_{-\infty}^{\infty}\mathrm{d}y'\int_0^{\infty}\mathrm{d}z'\rho_f(x',y',z')\times\\ &\quad\left[\frac{1}{\sqrt{(x-x')^2+(y-y')^2+(z-z')^2}}-\right.\\ &\quad\left.\frac{1}{\sqrt{(x-x')^2+(y-y')^2+(z+z')^2}}\right].\end{aligned}$$

§3.1.5 电多极矩法

在电磁学中已介绍过点电荷、电偶极子和电四极子所激发的势，它们分别与 R，R^2 及 R^3 成反比. 如果我们能够把分布在坐标原点附近一个很小的区域 V 内的电荷体系在远处激发的势，看成是位于原点的电荷 Q，以及中心在原点的电偶极子 $\boldsymbol{p}_e$，电四极子 $\overleftrightarrow{D}$，…所激发的势的叠加，就可以根据问题所要求的精度，适当地利用前几项之和近似地表示上述电荷体系的势. 这种方法称为多极矩法.

一个电荷分布为何能看成在原点的点电荷以及中心在原点的各级多极于的叠加呢？我们先从物理思想上讨论，然后再利用多元函数的泰勒公式作出定量证明.

为简单起见，设讨论的电荷分布是在原点附近 $\boldsymbol{x}'$ 处的一个点电荷 $+q$，我们来说明，这一点电荷可以看成是位于原点的点电荷 $+q$ 以及中心在原点的各级多极子的叠加(见图 1).

作为零级近似，可把这一电荷简单地看作位于原点的一个点电荷 $+q$.

精确一点，在图 1(a)的基础上，在原点 o 上同时放置一个 $+q$ 和一个 $-q$，显然图 1(b)的电荷体系与图 1(a)是等价的，因为位于同一点的 $\pm q$ 互相抵消了. 但图

1(b)却可以看成是位于原点的电荷 $+q$ 以及中心位于 $\boldsymbol{x}'/2$ 的电偶极子 $\boldsymbol{p}_e=q\boldsymbol{x}'$ 的叠加，如果我们把 $\boldsymbol{p}_e$ 的中心近似认为就在坐标原点，这就是一级近似.

如果要求更精确一点，可以在图 1(b)的基础上再在 $-\boldsymbol{x}'/2$ 和 $\boldsymbol{x}'/2$ 处同时置一对 $\pm q$，如图 1(c). 这两对电荷也互相抵消了，所以这个电荷体系也与图 1(a)完全等价. 但现在电荷体系可以看成是位于坐标原点的电荷 q，中心位于原点的电偶极子 $\boldsymbol{p}_e$，以及中心位于 $\boldsymbol{x}'/4$ 的电四极子 $\vec{\boldsymbol{d}}$ 的叠加. 如果我们把 $\vec{\boldsymbol{d}}$ 的中心近似地认为也在坐标原点，这就是二级近似.

仿此，可以一直作下去. 显然，如果引入的电多极子的级数越高，把它的中心近似认为在原点所带来的偏离也就越小.

既然一个在原点附近的点电荷可以这样分解，而一个复杂的电荷分布也无非是许许多多点电荷的叠加，把每一个点电荷都按此分解，那么整个电荷体系也就可以看成是位于原点的电荷 Q，中心在原点的电偶极子 $\boldsymbol{p}_e$、电四极子 $\vec{\boldsymbol{D}}$、…的叠加了. 子是，给定一个电荷体系，求出它的 $Q,\boldsymbol{p}_e,\vec{\boldsymbol{D}}$……即可求得这个电荷体系在远处产生的势. 下一节将利用图 1 的思想进一步说明电荷体系在外电场中的能量也可看成是描述该电荷体系的一系列多极子在外电场中的能量之和.

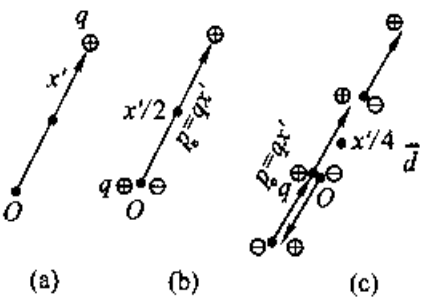


图 1 一个点荷可以看作各级电多极子的叠加

现在，我们在数学上给出严格的证明. 讨论问题的出发点是任意电荷分布的电势公式

$$\varphi(\boldsymbol{x})=\int_V \mathrm{d}V'\rho_f(\boldsymbol{x}')/(4\pi\varepsilon_0 r), \tag{5.1}$$

其中 $\boldsymbol{x}$ 为场点坐标，$\boldsymbol{x}'$为源点坐标，坐标原点取在电荷分布的区域 V 内. 令 $R=|\boldsymbol{x}|$ 及 $r=|\boldsymbol{x}-\boldsymbol{x}'|$.

由多元函数的泰勒公式

$$f(\boldsymbol{x}-\boldsymbol{x}')=f(\boldsymbol{x})-\boldsymbol{x}'\cdot\nabla f(\boldsymbol{x})+\frac{1}{2}(\boldsymbol{x}'\cdot\nabla)^2 f(\boldsymbol{x})+\cdots, \tag{5.2}$$

取

$$f(\boldsymbol{x}-\boldsymbol{x}')=1/|\boldsymbol{x}-\boldsymbol{x}'|=1/r, \tag{5.3}$$

则

$$f(\boldsymbol{x})=1/|\boldsymbol{x}|=1/R, \tag{5.4}$$

将(5.3)、(5.4)式代入(5.2)式，即有

$$\frac{1}{r}=\frac{1}{R}-\boldsymbol{x}'\cdot\nabla\frac{1}{R}+\frac{1}{2!}(\boldsymbol{x}'\cdot\nabla)^2\frac{1}{R}+\cdots, \tag{5.5}$$

将(5.5)式代入(5.1)式,并利用$(\boldsymbol{x}'\cdot\nabla)^2=\boldsymbol{x}'\boldsymbol{x}':\nabla\nabla$,便有

$$\begin{aligned}\varphi(x)&=\frac{1}{4\pi\varepsilon_0}\int_V \mathrm{d}V'\rho_f(\boldsymbol{x}')\left[\frac{1}{R}-\boldsymbol{x}'\cdot\nabla\frac{1}{R}+\frac{1}{2}(\boldsymbol{x}'\cdot\nabla)^2\frac{1}{R}+\cdots\right]\\&=\frac{1}{4\pi\varepsilon_0}\left[\frac{1}{R}\int_V\rho_f(\boldsymbol{x}')\mathrm{d}V'-\int_V\mathrm{d}V'\rho_f(\boldsymbol{x}')\boldsymbol{x}'\cdot\nabla\frac{1}{R}+\right.\\&\left.\frac{1}{6}\int_V\mathrm{d}V'3\rho_f(\boldsymbol{x}')\boldsymbol{x}'\boldsymbol{x}':\nabla\nabla\frac{1}{R}+\cdots\right]\\&=\frac{1}{4\pi\varepsilon_0}\left[\frac{Q}{R}-\boldsymbol{p}_e\cdot\nabla\frac{1}{R}+\frac{1}{6}\overleftrightarrow{D}:\nabla\nabla\frac{1}{R}+\cdots\right],\end{aligned}\tag{5.6}$$

其中,

$$Q=\int_V\rho_f(\boldsymbol{x}')\mathrm{d}V' \qquad \text{为体系总电荷;}\tag{5.7}$$

$$\boldsymbol{p}_e=\int_V\rho_f(\boldsymbol{x}')\boldsymbol{x}'\mathrm{d}V' \qquad \text{为体系电偶极矩;}\tag{5.8}$$

$$\overleftrightarrow{D}=\int_V 3\rho_f(\boldsymbol{x}')\boldsymbol{x}'\boldsymbol{x}'\mathrm{d}V' \qquad \text{为体系电四极矩;}\tag{5.9}$$

并且,

$$\varphi^{(0)}=Q/(4\pi\varepsilon_0R) \qquad \text{是 } Q \text{ 激发的电势;}\tag{5.10}$$

$$\varphi^{(1)}=-\boldsymbol{p}_e\cdot\nabla(1/R)/(4\pi\varepsilon_0) \qquad \text{是 } \boldsymbol{p}_e \text{ 激发的电势;}\tag{5.11}$$

$$\varphi^{(2)}=\overleftrightarrow{D}:\nabla\nabla(1/R)/(24\pi\varepsilon_0) \qquad \text{是 } \overleftrightarrow{D} \text{ 激发的电势.}\tag{5.12}$$

这样,利用(5.6)~(5.9)式即可求得真空中的电荷体系在远处激发的电势.这种方法就是多极矩法.通常讨论原子核物理中的问题,需要精确到电四极矩项,而电八极矩及更高级的多极矩,实际上很少用到.

现在说明电荷分布的对称性与体系的电多极矩是否为零的关系.如果电荷体系的正电荷与负电荷的电量大小相等,则电荷体系的 Q 为零;如果电荷分布对原点对称,则由(5.8)式可证电荷体系的 $\boldsymbol{p}_e$ 为零;如果电荷分布具有球对称性,则电荷体系的 $\overleftrightarrow{D}$ 为零.后者可用高斯定理来证明:作一个包围整个电荷体系的球面为高斯面,由于电荷分布的球对称性,因而球面上的 $\boldsymbol{E}$ 大小相等,方向只能与径向相同或相反.由此可得 $E=Q/(4\pi\varepsilon_0R^2)$,即 $\varphi=Q/(4\pi\varepsilon_0R)$,与(5.6)式相比较,即得 $\boldsymbol{p}_e=0$ 及 $\overleftrightarrow{D}=0$.这里电荷分布具有球对称就必然对原点对称,故 $\boldsymbol{p}_e=0$ 是很自然的.



图2 电四极子

下面,我们着重讨论电四极矩的性质.

①体系的电四极矩 $\overleftrightarrow{D}$ 是一个张量，它的九个分量是

$$D_{ij} = \boldsymbol{e}_i \cdot \overleftrightarrow{D} \cdot \boldsymbol{e}_j = \int_V 3\rho_f(x')(\boldsymbol{e}_i \cdot \boldsymbol{x}')(\boldsymbol{x}' \cdot \boldsymbol{e}_j)\mathrm{d}V' = \int_V 3\rho_f(\boldsymbol{x}')x'_i x'_j \mathrm{d}V'. \tag{5.13}$$

为了对体系的电四极矩有更加具体的了解，我们计算下面的例子.

【例 1】 已知电荷分布如图 2 所示，求这个电荷体系的电四极矩 $\overleftrightarrow{D}$.

解 这个电荷体系的电荷密度可表示为

$$\rho_f(\boldsymbol{x}) = q\delta^3(\boldsymbol{x}' - b\boldsymbol{e}_z) - q\delta^3(\boldsymbol{x}' - a\boldsymbol{e}_z) - q\delta^3(\boldsymbol{x}' + a\boldsymbol{e}_z) + q\delta^3(\boldsymbol{x}' + b\boldsymbol{e}_z).$$

代入(5.9)式，利用 $\int \boldsymbol{f}(\boldsymbol{x}')\delta^3(\boldsymbol{x}' - \boldsymbol{x})\mathrm{d}V' = f(\boldsymbol{x})$ 可得

$$\begin{aligned}
\overleftrightarrow{D} &= \int_V 3\rho_f(\boldsymbol{x}')\boldsymbol{x}'\boldsymbol{x}'\mathrm{d}V' \\
&= \int_V 3[q\delta^3(\boldsymbol{x}' - b\boldsymbol{e}_z) - q\delta^3(\boldsymbol{x}' - a\boldsymbol{e}_z) - q\delta^3(\boldsymbol{x}' + a\boldsymbol{e}_z) + q\delta^3(\boldsymbol{x}' + b\boldsymbol{e}_z)]\boldsymbol{x}'\boldsymbol{x}'\mathrm{d}V' \\
&= 3q[bb\boldsymbol{e}_z\boldsymbol{e}_z - aa\boldsymbol{e}_z\boldsymbol{e}_z - (-a)(-a)\boldsymbol{e}_z\boldsymbol{e}_z + (-b)(-b)\boldsymbol{e}_z\boldsymbol{e}_z] \\
&= 6q(b^2 - a^2)\boldsymbol{e}_z\boldsymbol{e}_z = 6q(b-a)(b+a)\boldsymbol{e}_z\boldsymbol{e}_z = 6p_f l\boldsymbol{e}_z\boldsymbol{e}_z
\end{aligned} \tag{5.14}$$

其中应用了 $p_f = q(b-a)$ 及 $l = (b+a)$.

②$\overleftrightarrow{D}$ 是一个对称张量. 因为

$$D_{ij} = \int_V 3\rho_f(\boldsymbol{x}')x'_i x'_j \mathrm{d}V' = \int_V 3\rho_f(\boldsymbol{x}')x'_j x'_i \mathrm{d}V' = D_{ji}. \tag{5.15}$$

由此可见，$\overleftrightarrow{D}$ 至多只有六个独立分量，即三个对角元 D_{11}, D_{22}, D_{33} 和三个非对角元 $D_{12} = D_{21}, D_{23} = D_{32}, D_{31} = D_{13}$.

③电四极矩实际上只有五个独立分量.

为了证明这一论断，我们重新定义一个张量[1]：

$$\overleftrightarrow{\mathscr{D}} = \int_V [3\boldsymbol{x}'\boldsymbol{x}' - R'^2\overleftrightarrow{e}]\rho_f(\boldsymbol{x}')\mathrm{d}V', \tag{5.16}$$

称为约化电四极矩张量. 式中的 $\overleftrightarrow{e}$ 是单位张量，显然 $\overleftrightarrow{\mathscr{D}} \neq \overleftrightarrow{D}$，但是可以证明 $\overleftrightarrow{\mathscr{D}}$ 与 $\overleftrightarrow{D}$ 所激发的势是相同的，因而在(5.6)式中用 $\overleftrightarrow{\mathscr{D}}$ 取代 $\overleftrightarrow{D}$ 对 $\varphi(\boldsymbol{x})$ 没有影响.

证 由(5.12)式出发：

$$\begin{aligned}
\varphi^{(2)}(\boldsymbol{x}) &= \frac{1}{24\pi\varepsilon_0}\overleftrightarrow{D} : \nabla\nabla(1/R) \\
&= \frac{1}{24\pi\varepsilon_0}\left[\overleftrightarrow{\mathscr{D}} + \int_V R'^2\overleftrightarrow{e}\rho_f(\boldsymbol{x}')\mathrm{d}V'\right] : \nabla\nabla(1/R)
\end{aligned}$$

[1] 将(5.13)、(5.16)式和转动惯量张量的定义 §1.5.1(1.26)式比较，可以看到它们的类似性. 差别只在于转动惯量张量的迹不为零，因而其六个分量独立.

$$=\frac{1}{24\pi\varepsilon_0}\vec{\mathscr{D}}:\nabla\nabla(1/R)+\frac{1}{24\pi\varepsilon_0}\int_V R'^2\rho_f(x')\nabla^2\frac{1}{R}\mathrm{d}V'$$

$$=\frac{1}{24\pi\varepsilon_0}\vec{\mathscr{D}}:\nabla\nabla(1/R). \tag{5.17}$$

因为场点 x 是在远离电荷体系的地方，而坐标原点选在电荷体系内部，因此场点到坐标原点的距离 $R=|x|\neq 0$，故 $\nabla^2(1/R)=0$，最后一个积分式的被积函数为零. 这就证明了可以用约化电四极矩张量 $\vec{\mathscr{D}}$ 取代电四极矩张量 $\vec{D}$ 来描述电荷体系的性质. 【证毕】

很容易证明，约化电四极矩的迹为零，即

$$\sum_i \mathscr{D}_{ii}=\int_V[(3x'^2_1-R'^2)+(3x'^2_2-R'^2)+(3x'^2_3-R'^2)]\rho_f(x')\mathrm{d}V'$$

$$=\int_V 3(x'^2_1+x'^2_2+x'^2_3-R'^2)\rho_f(x')\mathrm{d}V'=0, \tag{5.18}$$

这说明，由任意两个对角元可以求得第三个对角元，可见 $\vec{\mathscr{D}}$ 的独立分量为五个. 【证毕】

下面举例说明如何用多极矩法求电荷体系激发的电势.

【例 2】 半轴为 a,b,c 的椭球体内均匀带电，总电量为 Q，求它的电偶极矩、电四极矩以及准确到电四极矩的在远处的电势，并进一步讨论旋转椭球（$a=b$）和圆球（$a=b=c$）的特殊情况.

解 电荷体系在远处激发的电势（准确到电四极矩）为

$$\varphi(x)=\frac{1}{4\pi\varepsilon_0}\left[\frac{Q}{R}-\boldsymbol{p}_e\cdot\nabla\frac{1}{R}+\frac{1}{6}\vec{\mathscr{D}}:\nabla\nabla\frac{1}{R}\right].$$

在本题的情况下，电荷 Q 已给定，由于电荷分布对原点的对称性可知 $\boldsymbol{p}_e=0$，因此关键是求得电荷体系的 $\vec{\mathscr{D}}$. 为此，

(a) 引入广义球坐标. 计算 $\vec{\mathscr{D}}$ 的困难在于椭球面的方程为

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1, \tag{5.19}$$

因此，利用(5.16)式计算 $\vec{\mathscr{D}}$ 时积分限比较复杂. 为简化计算，作坐标变换：

$$X=x/a,\quad Y=y/b,\quad Z=z/c, \tag{5.20}$$

则椭球面方程(5.19)式变为半径为 1 的球面方程

$$X^2+Y^2+Z^2=1, \tag{5.21}$$

这样，积分限就容易确定. 再在 XYZ 坐标系中引入球坐标，即有

$$x=aX=a\sin\theta\cos\phi, \tag{5.22}$$

$$y=bY=b\sin\theta\sin\phi, \tag{5.23}$$

$$z = cZ = cr\cos\theta. \tag{5.24}$$

这就是广义球坐标. 其体积元由雅可比行列式求得为

$$\mathrm{d}x\mathrm{d}y\mathrm{d}z = abcr^2\sin\theta\mathrm{d}r\mathrm{d}\theta\mathrm{d}\phi. \tag{5.25}$$

(b)计算约化电四极矩张量 $\ddot{\mathscr{D}}$. 为书写方便,本题全部略去源点坐标的撇号. 由(5.16)式得

$$\mathscr{D}_{ij} = \boldsymbol{e}_i \cdot \ddot{\boldsymbol{\mathscr{D}}} \cdot \boldsymbol{e}_j = \rho_f\int_V (3x_ix_j - R^2\delta_{ij})\mathrm{d}V. \tag{5.26}$$

当 $i\neq j$ 时,由于电荷分布对原点的对称性可证

$$\mathscr{D}_{ij} = 3\rho_f\int_V x_ix_j\mathrm{d}V = 0,$$

当 $i=j$ 时,

$$\mathscr{D}_{ii} = \rho_f\int_V (3x_i^2 - R^2)\mathrm{d}V.$$

现分别计算如下:

$$3\rho_f\int_V x_1^2\mathrm{d}V = 3\rho_f a^3bc\int_0^1 r^4\mathrm{d}r\int_0^{2\pi}\cos^2\phi\mathrm{d}\phi\int_0^{\pi}\sin^3\theta\mathrm{d}\theta = \frac{4}{5}\pi\rho_f a^3bc,$$

$$3\rho_f\int_V x_2^2\mathrm{d}V = 3\rho_f ab^3c\int_0^1 r^4\mathrm{d}r\int_0^{2\pi}\sin^2\phi\mathrm{d}\phi\int_0^{\pi}\sin^3\theta\mathrm{d}\theta = \frac{4}{5}\pi\rho_f ab^3c,$$

$$3\rho_f\int_V x_3^2\mathrm{d}V = 3\rho_f abc^3\int_0^1 r^4\mathrm{d}r\int_0^{2\pi}\mathrm{d}\phi\int_0^{\pi}\cos^2\theta\sin\theta\mathrm{d}\theta = \frac{4}{5}\pi\rho_f abc^3.$$

将三式相加,并利用 $R^2 = x_1^2 + x_2^2 + x_3^3$,即有

$$\rho_f\int_V R^2\mathrm{d}V = 4\pi\rho_f abc(a^2 + b^2 + c^2)/15 = Q(a^2 + b^2 + c^2)/5.$$

再代入 $\mathscr{D}_{ij}$,得

$$\begin{aligned}
\mathscr{D}_{11} &= Q(2a^2 - b^2 - c^2)/5,\\
\mathscr{D}_{22} &= Q(-a^2 + 2b^2 - c^2)/5,\\
\mathscr{D}_{33} &= Q(-a^2 - b^2 + 2c^2)/5,\\
\ddot{\boldsymbol{\mathscr{D}}} &= Q[(2a^2 - b^2 - c^2)\boldsymbol{e}_x\boldsymbol{e}_x + (-a^2 + 2b^2 - c^2)\boldsymbol{e}_y\boldsymbol{e}_y\\
&\quad + (-a^2 - b^2 + 2c^2)\boldsymbol{e}_z\boldsymbol{e}_z]/5.
\end{aligned} \tag{5.27}$$

(c)准确到电四极矩的在远处的电势.

在所要求的精度下,体系在远处的电势为

$$\varphi(\boldsymbol{x}) = \varphi^{(0)} + \varphi^{(1)} + \varphi^{(2)}, \tag{5.28}$$

其中,

$$\varphi^{(0)} = Q/(4\pi\varepsilon_0 R), \tag{5.29}$$

$$\varphi^{(1)} = -\boldsymbol{p}_e \cdot \nabla (1/R)/(4\pi\varepsilon_0) = 0, \tag{5.30}$$

$$\varphi^{(2)} = \frac{1}{24\pi\varepsilon_0}\overleftrightarrow{\mathscr{D}} : \nabla\nabla\frac{1}{R} = \frac{1}{24\pi\varepsilon_0}\sum_{i=1}^{3}\mathscr{D}_{ii}\frac{\partial^2}{\partial X_i^2}\frac{1}{R}, \tag{5.31}$$

而 $\partial^2(1/R)/\partial x_i^2 = (3x_i^2 - R^2)/R^5$，因此：

$$\begin{aligned}\varphi^{(2)} &= \frac{Q}{120\pi\varepsilon_0}\left[\frac{3x^2 - R^2}{R^5}(2a^2 - b^2 - c^2) + \frac{3y^2 - R^2}{R^5}(-a^2 + 2b^2 - c^2)\right.\\ &\quad \left. + \frac{3z^2 - R^2}{R^5}(-a^2 - b^2 + 2c^2)\right]\\ &= Q[a^2(3x^2 - R^2) + b^2(3y^2 - R^2) + c^2(3z^2 - R^2)]/40\pi\varepsilon_0 R^5,\\ \varphi &= \frac{1}{4\pi\varepsilon_0}\left\{\frac{Q}{R} + \frac{Q[a^2(3x^2 - R^2) + b^2(3y^2 - R^2) + c^2(3z^2 - R^2)]}{10R^5}\right\}.\end{aligned} \tag{5.32}$$

(d)旋转椭球的情形. 将 $a = b$ 代入上式后化简为：

$$\varphi = \frac{1}{4\pi\varepsilon_0}\left[\frac{Q}{R} + \frac{Q(a^2 - c^2)(1 - 3\cos^2\theta)}{10R^3}\right]. \tag{5.33}$$

(e)圆球的情形. 将 $a = b = c$ 代入(5.32)式即有 $\varphi = Q/(4\pi\varepsilon_0 R)$，即均匀带电球体在远处的电势等于球体的电量全部集中在球心的点电荷在该处的电势. 实际上，当电荷分布具有球对称时，体系的 $\boldsymbol{p}_e = 0$，$\overleftrightarrow{\mathscr{D}} = 0$. 所以 $\varphi = \varphi^{(0)}$.

§3.1.6 静电场的能量 电荷体系与外电场的相互作用

(一)静电场的能量

在§2.4.3中已导出各向同性的线性介质中电磁场的能量密度为

$$w = \frac{1}{2}(\boldsymbol{E}\cdot\boldsymbol{D} + \boldsymbol{H}\cdot\boldsymbol{B}), \tag{6.1}$$

见该节的(3.11)式，因而电场的能量密度为

$$w = \frac{1}{2}\boldsymbol{E}\cdot\boldsymbol{D}, \tag{6.2}$$

全空间的电场能量为

$$E = \int_\infty w\mathrm{d}V = \frac{1}{2}\int_\infty \boldsymbol{E}\cdot\boldsymbol{D}\mathrm{d}V. \tag{6.3}$$

它可以是变化电磁场中的电场能量，也可以用来表示静电场的能量.

但是，对于静电场可以引入静电势 φ，而 φ 是标量，计算起来比 $\boldsymbol{E}$ 和 $\boldsymbol{D}$ 方便，故常用 φ 来表示静电场的能量. 由 $\boldsymbol{E} = -\nabla\varphi$ 及 $\nabla\cdot\boldsymbol{D} = \rho_f$ 可得：

$$\boldsymbol{E}\cdot\boldsymbol{D} = -(\nabla\varphi)\cdot\boldsymbol{D} = -[\nabla\cdot(\varphi\boldsymbol{D}) - \varphi\nabla\cdot\boldsymbol{D}] = -\nabla\cdot(\varphi\boldsymbol{D}) + \varphi\rho_f. \tag{6.4}$$

将(6.4)式代入(6.3)式，有 $E = -\frac{1}{2}\int_\infty \nabla\cdot(\varphi\boldsymbol{D})\mathrm{d}V + \frac{1}{2}\int_\infty \varphi\rho_f\mathrm{d}V$. 利用高斯定理，

可得 $\int_\infty \nabla\cdot(\varphi\boldsymbol{D})\mathrm{d}V=\oint_\infty \varphi\boldsymbol{D}\cdot\mathrm{d}\boldsymbol{\sigma}=0$. 这是因为在封闭面上,$\varphi\propto 1/R, D\propto 1/R^2$, $\mathrm{d}\sigma\propto R^2$,故当 $R\to\infty$ 时,面积分为零. 于是

$$E=\frac{1}{2}\int_\infty \varphi\rho_f\mathrm{d}V. \tag{6.5}$$

注意,(6.5)式只能应用于静电场,因为推导过程中应用了 $\boldsymbol{E}=-\nabla\varphi$. 其次,(6.5)式不是对任意积分区域成立,因而不能将其中的被积函数 $\varphi\rho_f/2$ 等同于能量密度 w. 例如,一个均匀带电的介质球,它的电场分布在整个空间,因而电场的能量也分布在整个空间. 但如果认为 $w(\boldsymbol{x})=\varphi(\boldsymbol{x})\rho_f/2$,则只有在 $\rho_f(\boldsymbol{x})$ 不为零的地方电场的能量密度才不为零,这显然是不正确的.

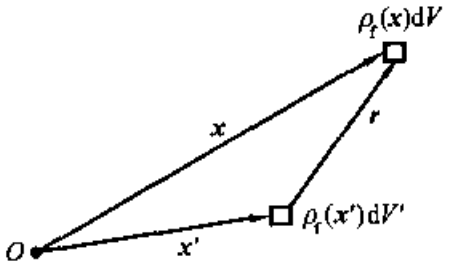


图 1 $\rho_f(\boldsymbol{x})$ 与 $\rho_f(\boldsymbol{x}')$ 描述同一电荷分布

最后,我们将 $\varphi(\boldsymbol{x})=\int \mathrm{d}V'\rho_f(\boldsymbol{x}')/(4\pi\varepsilon r)$ 代入(6.5)式,即得电场能量为

$$E=\frac{1}{2}\int\left[\int\frac{\rho_f(\boldsymbol{x}')\mathrm{d}V'}{4\pi\varepsilon r}\right]\rho_f(\boldsymbol{x})\mathrm{d}V=\frac{1}{8\pi\varepsilon}\int\mathrm{d}V\int\mathrm{d}V'\frac{\rho_f(\boldsymbol{x})\rho_f(\boldsymbol{x}')}{r}, \tag{6.6}$$

其中 $r=|\boldsymbol{x}-\boldsymbol{x}'|$.

注意,这里的 $\rho_f(\boldsymbol{x})$ 与 $\rho_f(\boldsymbol{x}')$ 描述的是同一个电荷分布,参看图 1. 这个电荷分布在 $\boldsymbol{x}$ 点的电势是

$$\varphi(\boldsymbol{x})=\int \mathrm{d}V'\rho_f(\boldsymbol{x}')/(4\pi\varepsilon r),$$

这样的电势分布又使得所给定的这个电荷分布 $\rho_f(\boldsymbol{x})$ 具有按公式 $E=\int_\infty \mathrm{d}V\varphi(\boldsymbol{x})\rho_f(\boldsymbol{x})/2$ 计算的能量. 因为这里包含了两个积分过程,所以用不同的积分变量来区别它们. 此外,公式(6.6)应用的范围除了受(6.5)式的限制外,还要求介质均匀充满,即所有电场 $\boldsymbol{E}\neq 0$ 的地方都均匀充满介电常数为 ε 的各向同性的线性介质. 这是公式 $\varphi(\boldsymbol{x})=\int \mathrm{d}V'\rho_f(\boldsymbol{x}')/(4\pi\varepsilon r)$ 成立的条件. 由此可见,静电场能量的三个表达式(6.3),(6.5),(6.6)各自的应用范围是不同的.

(二)电荷体系在外电场中的能量

电荷体系在外场中的能量即电荷体系与电磁场的相互作用能.

设外场由电荷分布 $\rho_e(\boldsymbol{x})$ 产生,外场的电势分布为 $\varphi_e(\boldsymbol{x})$;电荷体系的电荷分布为 $\rho_f(\boldsymbol{x})$,它激发的电势为 $\varphi(\boldsymbol{x})$.

将两个电荷分布 $\rho_e(\boldsymbol{x})$ 和 $\rho_f(\boldsymbol{x})$ 合起来看成一个大的电荷体系,其电场能量由(6.5)式给出:

$$E = \frac{1}{2}\int_{\infty} \mathrm{d}V(\varphi_e + \varphi)(\rho_e + \rho_f). \tag{6.7}$$

但当$\rho_e(\boldsymbol{x})$与$\rho_f(\boldsymbol{x})$单独存在时,它们各自的电场能分别为

$$E_e = \frac{1}{2}\int_{\infty} \mathrm{d}V\varphi_e\rho_e, \tag{6.8}$$

$$E_f = \frac{1}{2}\int_{\infty} \mathrm{d}V\varphi\rho_f, \tag{6.9}$$

因而电荷体系$\rho_f(\boldsymbol{x})$在外场$\varphi_e(\boldsymbol{x})$中的能量(电荷体系与外电场的相互作用能)为

$$E_{\text{相互}} = E - (E_e + E_f) = \frac{1}{2}\int \mathrm{d}V[(\varphi_e + \varphi)(\rho_e + \rho_f) - (\varphi_e\rho_e + \varphi\rho_f)]$$

$$= \frac{1}{2}\int \mathrm{d}V(\varphi_e\rho_f + \varphi\rho_e).$$

将上式右边第二项的积分次序交换后再交换积分变量$\boldsymbol{x}$与$\boldsymbol{x}'$的符号,可以证明上式的两项是相等的:

$$\int \varphi(\boldsymbol{x})\rho_e(\boldsymbol{x})\mathrm{d}V = \int \mathrm{d}V\rho_e(\boldsymbol{x})\int \mathrm{d}V'\frac{\rho_f(\boldsymbol{x}')}{4\pi\varepsilon_0 r} = \int \mathrm{d}V'\rho_f(\boldsymbol{x}')\int \mathrm{d}V\frac{\rho_e(\boldsymbol{x})}{4\pi\varepsilon_0 r}$$

$$= \int \mathrm{d}V\rho_f(\boldsymbol{x})\int \mathrm{d}V'\frac{\rho_e(\boldsymbol{x}')}{4\pi\varepsilon_0 r} = \int \mathrm{d}V\rho_f(\boldsymbol{x})\varphi_e(\boldsymbol{x}).$$

由此得

$$E_{\text{相互}} = \int \varphi_e\rho_f \mathrm{d}V. \tag{6.10}$$

这就是电荷分布为$\rho_f(\boldsymbol{x})$的体系在电势为$\varphi_e(\boldsymbol{x})$的外电场中所具有的能量,亦即电荷体系与外电场的相互作用能.

现在,如果电荷体系分布在一个小区域V内(坐标原点仍取在V内),将这个小区域的电势$\varphi_e(\boldsymbol{x})$在坐标原点展开,则电荷体系在外电场中的能量可以看成描述该电荷体系的一系列电多极子在外电场中的能量之和.

证 将$\varphi_e(\boldsymbol{x})$在$\boldsymbol{x}=0$点展开成泰勒级数

$$\varphi_e(\boldsymbol{x}) = \varphi_e(0) + \boldsymbol{x}\cdot\nabla\varphi_e(0) + \frac{1}{2!}(\boldsymbol{x}\cdot\nabla)^2\varphi_e(0) + \cdots, \tag{6.11}$$

代入(6.10)式,便有:

$$E_{\text{相互}} = \int_V \rho_f(\boldsymbol{x})[\varphi(0) + \boldsymbol{x}\cdot\nabla\varphi_e(0) + \frac{1}{2!}\boldsymbol{x}\boldsymbol{x}:\nabla\nabla\varphi_e(0) + \cdots]\mathrm{d}V$$

$$= [\int_V \rho_f(\boldsymbol{x})\mathrm{d}V]\varphi_e(0) + [\int_V \rho_f(\boldsymbol{x})\boldsymbol{x}\mathrm{d}V]\cdot\nabla\varphi_e(0) +$$

$$\frac{1}{6}\left\{\int_V [3\rho_f(\boldsymbol{x})\boldsymbol{x}\boldsymbol{x} - R^2\overleftrightarrow{e}]\mathrm{d}V\right\}:\nabla\nabla\varphi_e(0) + \cdots$$

$$= Q\varphi_e(0) + \boldsymbol{p}_e \cdot \nabla \varphi_e(0) + \frac{1}{6}\overleftrightarrow{\mathscr{D}} : \nabla\nabla \varphi_e(0) + \cdots \qquad (6.12)$$

上述结论得证. 【证毕】

我们看到,一方面,电荷体系所激发的电势可以看成描述该电荷体系的一系列电多极子所激发的电势之和,如上节(5.6)式;另一方面,电荷体系在外场中的能量也可以看成描述该电荷体系的一系列电多极子在外场中的能量之和,如本节(6.12)式.

(三)电多极子在外电场中所受的力和力矩

利用第一篇第一章介绍过的虚功原理,并注意到电多极子在外电场中的势能 U 等于电多极子与外电场的相互作用能 $E_{\text{相互}}$,很容易得到电多极子在外场中所受的力和力矩.

设处于一定位形的电荷体系在外电场中的势能为 U. 当它的位形发生微小变化时,例如发生平移 $\delta\boldsymbol{l}$ 或转动 $\delta\boldsymbol{\theta}$ 时,电场力作功

$$\delta A = \boldsymbol{F} \cdot \delta\boldsymbol{l} \quad 或 \quad \delta A = M_{力}\,\delta\theta. \qquad (6.13)$$

电场力作功将使势能 U 减少,即

$$\delta A = -\delta U. \qquad (6.14)$$

将(6.14)式代入(6.13)的第一式,有:

$$\boldsymbol{F} \cdot \delta\boldsymbol{l} = -\delta U \quad 或 \quad F_l \delta l = -(\partial U/\partial l)\delta l.$$

由于 δl 的任意性,可得

$$F_l = -\partial U/\partial l.$$

如果将 δl 分别取 $\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z$ 的方向,则有

$$\boldsymbol{F} = \sum_{k=1}^{3} F_k \boldsymbol{e}_k = -\sum_{k=1}^{3} (\partial U/\partial x_k)\boldsymbol{e}_k = -\nabla U, \qquad (6.15)$$

即电荷体系所受的力等于电荷体系在外场的势能的负梯度.

将(6.14)式代入(6.13)式的第二式,得

$$M_{力}\,\delta\theta = -\delta U = -(\partial U/\partial\theta)\delta\theta,$$

因而

$$M_{力} = -\partial U/\partial\theta, \qquad (6.16)$$

即电荷体系所受的力矩等于 U 对 θ 的偏导数的负值.

现在利用(6.15)、(6.16)式计算外电场对电荷 Q、电偶极子 $\boldsymbol{p}_e$、电四极子 $\overleftrightarrow{\mathscr{D}}$ 的作用力和力矩.

位于坐标原点的电荷 Q 在外电场中的势能等于(6.12)式的第一项

$$U^{(0)} = Q\varphi_e(0), \qquad (6.17)$$

因而外电场对电荷 Q 的作用力与力矩分别为

$$F^{(0)} = -\nabla U^{(0)} = -Q\nabla\varphi_e(0) = QE_e(0), \tag{6.18}$$

$$M_{力}^{(0)} = -\partial Q\varphi_e(0)/\partial\theta = 0. \tag{6.19}$$

中心位于坐标原点的电偶极子 $\boldsymbol{p}_e$ 在外电场中的势能等于(6.12)式的第二项

$$U^{(1)} = \boldsymbol{p}_e\cdot\nabla\varphi_e(0) = -\boldsymbol{p}_e\cdot\boldsymbol{E}_e(0), \tag{6.20}$$

因而外电场对电偶极子 $\boldsymbol{p}_e$ 的作用力与力矩分别为

$$\begin{aligned} F^{(1)} &= -\nabla U^{(1)} = \nabla[\boldsymbol{p}_e\cdot\boldsymbol{E}_e(0)] \\ &= \boldsymbol{p}_e\times[\nabla\times\boldsymbol{E}_e(0)] + [\boldsymbol{p}_e\cdot\nabla]\boldsymbol{E}_e(0) \\ &= \boldsymbol{p}_e\cdot\nabla\boldsymbol{E}_e(0), \end{aligned} \tag{6.21}$$

$$M_{力}^{(1)} = -\partial U^{(1)}/\partial\theta = \partial[pE_e(0)\cos\theta]/\partial\theta = -pE_e(0)\sin\theta,$$

即

$$\boldsymbol{M}_{力}^{(1)} = \boldsymbol{p}_e\times\boldsymbol{E}_e(0). \tag{6.22}$$

(6.21)式说明,电偶极子只有在不均匀电场中才受到合力不为零的电场力的作用. 例如,用皮毛摩擦橡胶棒能吸引通草球,是因为通草球在静电感应下成为电偶极子,而电偶极子在带电橡胶棒的不均匀电场中受力被吸引向橡胶棒. (6.22)式表示的力矩与分别考虑电偶极子的正负电荷受力而得到的力矩是相同的.

中心位于坐标原点的电四极子 $\overleftrightarrow{\mathscr{D}}$ 在外电场中的势能等于(6.12)式的第三项

$$U^{(2)} = \frac{1}{6}\overleftrightarrow{\mathscr{D}}:\nabla\nabla\varphi_e(0) = -\frac{1}{6}\overleftrightarrow{\mathscr{D}}:\nabla\boldsymbol{E}_e(0), \tag{6.23}$$

因而电四极子只是在不均匀电场中其势能才不为零. 它在电场中所受的作用力和力矩分别为:

$$F^{(2)} = \frac{1}{6}\nabla[\overleftrightarrow{\mathscr{D}}:\nabla\boldsymbol{E}_e(0)], \tag{6.24}$$

$$M_{力}^{(2)} = \frac{1}{6}\frac{\partial[\overleftrightarrow{\mathscr{D}}:\nabla\boldsymbol{E}_e(0)]}{\partial\theta}. \tag{6.25}$$

由此可见,只有在不均匀电场中,电场对电四极子的作用力与力矩才不为零.

习 题

1. 一块极化介质的极化矢量为 $\boldsymbol{P}(x')$,根据偶极子静电势公式,极化介质所产生的静电势为

$$\varphi = \int_V\frac{\boldsymbol{P}(x')\cdot\boldsymbol{r}}{4\pi\varepsilon_0 r^3}\mathrm{d}V'.$$

另外,根据极化电荷公式 $\rho_P = -\nabla'\cdot\boldsymbol{P}(x')$ 及 $\sigma_P = \boldsymbol{e}_n\cdot\boldsymbol{P}$,极化介质所产生的电势又可表示为

$$\varphi = -\int_V\frac{\nabla'\cdot\boldsymbol{P}(x')}{4\pi\varepsilon_0 r}\mathrm{d}V' + \oint_S\frac{\boldsymbol{P}(x')\cdot\mathrm{d}\boldsymbol{\sigma}'}{4\pi\varepsilon_0 r}.$$

试证明以上两表达式是等价的.

2. 试证明：

①在面电荷两侧，电势法向微商有跃变，而电势是连续的；

②面电荷密度分别为 $\pm\sigma_f$，而相互间的距离 $l\to 0$ 的两个面称为偶极层，$P=\lim\limits_{\substack{\sigma_f\to\infty\\ l\to 0}}\sigma_f l$ 称为面偶极密度，见图 1. 证明在偶极层两侧电势有跃变

$$\varphi_2-\varphi_1=\frac{1}{\varepsilon_0}e_n\cdot P,$$

而电势的法向微商是连续的.

提示 计算偶极层两侧的势

$$\varphi(x)=\int_{S'}\frac{\sigma_f(x')\,ds'}{4\pi\varepsilon_0|x-x'|}+\int_{S''}\frac{\sigma_f(x'')\,ds''}{4\pi\varepsilon_0|x-x''|}$$

时，应利用三元函数的泰勒公式展开（l 为小量）：

$$\frac{1}{|x-x''|}=\frac{1}{|x-x'+l|}=\frac{1}{|x-x'|}+l\cdot\nabla\frac{1}{|x-x'|}$$

3. 证明均值定理：在无电荷空间任一点的静电势之值，等于以该点为球心的任一球面上势的平均值.

4. 试给出导电介质中稳定电场的唯一性定理及其证明，已知导电介质的电导率为 $\sigma(x)$.

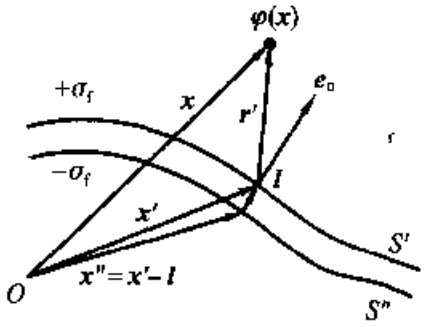


图 1 偶极层示意图

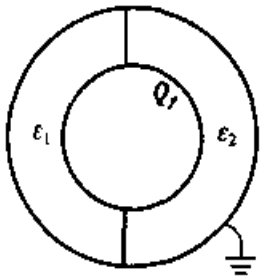


图 2 习题 5 图

5. 在两同心导体球壳之间充以两种介质，左半部介电常数为 ε_1，右半部介电常数为 ε_2，设内球壳带总电荷 Q_f，外球壳接地，见图 2. 求电场和球壳上的电荷分布.

6. 在均匀外电场 E_0 中置入一个带均匀自由电荷 ρ_f 的绝缘介质球（介电常数 ε），求空间各点的电势.

7. 均匀介质球的中心置一点电荷 Q_f，球的介电常数为 ε，球外为真空，试用分离变量法求空间电势，并将结果与高斯定理所得结果比较.

8. 一半径为 R_0 的球面，在球坐标 $0\leqslant\theta<\pi/2$ 的半球面上电势为 φ_0，在 $\pi/2<\theta\leqslant\pi$ 的半球面上电势为 $-\varphi_0$，求球外空间各点的电势.

9. 空心导体球壳的内、外半径分别为 R_1 和 R_2，球中心置一偶极子 p_f，球壳上带电荷 Q_f，求空间各点的电势和电荷分布.

10. 在一很大的电解槽中充满电导率为 σ_2 的液体，使其中流着均匀的电流 J_{f0}. 今在液体中置入一个电导率为 σ_1 的小球，求稳恒时电流分布和面电荷分布. 讨论 $\sigma_1\gg\sigma_2$ 及 $\sigma_2\gg\sigma_1$ 两种情况下电流分布的特点.

11. 有一个半径为 R_0 的绝缘金属球，位于半径为 R'_0 的空心接地金属球内. 这两个金属球的球心距离为 d，而且 $d\ll R_0$，$d\ll R'_0$. 内球的总电荷为 Q_f. 求在内球面上的面电荷分布和作用在内球上的力 F. 所求的解要准确到 d 的一次项.

提示 选两球心连线为 z 轴，提出具有轴对称性的尝试解

$$\varphi = \sum [a_l R^l + b_l R^{-(l+1)}] d^l \mathrm{P}_l(\cos\theta),$$

按题意仅保留到 $l=1$ 项，并注意到内、外球面方程分别为 $R = R_0$ 及 $R' = R'_0 + c\cos\theta$.

12. 在接地的导体平面上有一半径为 a 的半球凸部，半球的球心在导体平面上，点电荷 Q 位于系统的对称轴上，并与平面相距为 $b(b>a)$，试用电象法求空间电势.

13. 有一点电荷 Q_f 位于两个互相垂直的接地导体平面所围成的直角空间内. 它到两平面的距离分别为 a 和 b，求空间电势.

14. 半径为 R_0 的导体球外充满均匀绝缘介质 ε. 导体球接地. 离球心为 a 处 $(a>R_0)$ 置一点电荷 Q_f. 试用分离变量法求空间各点的电势. 证明所得结果与电象法结果相同.

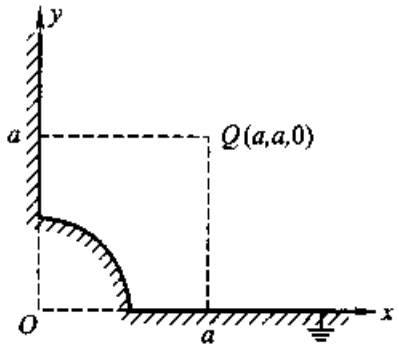


图 3 第 17 题图

15. 接地空心导体球的内外半径分别为 R_1 和 R_2. 在球内离球心为 $a(a<R_1)$ 处放置一个点电荷 Q_0. 用电象法求电势. 导体球上的感应电荷有多少？分布在内表面还是外表面？

16. 设两种介电常数分别为 ε_1 和 ε_2 的介质分别充满左右两个半无限空间，两者的分界面为 $x=0$ 平面. 有自由点电荷 Q_f 位于介质 1 的 $(x_0,0,0)$ 处，求介质 1 与介质 2 的电场分布.

17. 已知两块互相垂直的、接地半无穷大金属平板与球心在 O 点，半径为 R_0 的部分金属球面把空间分成两个区域，见图 3. 有一电量为 Q_f 的点电荷置于 $(a,a,0)$ 点，求电荷所在空间区域的电势.

18. 有一半径为 R_0，球心位于坐标原点的带电导体球的电势为 φ_0，另一无穷大接地导体平板位于 $x=a$ 处，其中 $a \gg R_0$，见图 4. 求 $x<a, R>R_0$ 区域内的电势. 要求精确到 R_0/a 的二级小量.

19. 已知在 $x=0$ 及 $x=l$ 处有两块互相平行的无穷大接地金属平板. 在两平行板间的 c 点处有一点电荷 Q_f，求空间电势.

20. 设有由两个半无穷大平面围成的直角无穷容器，其内充满电导率为 σ 的液体. 取该两平面为 xz 面和 yz 面. 在 (x_0,y_0,z_0) 和 $(x_0,y_0,-z_0)$ 两点分别置正负电极并通以电流 I_f，求导电液中的电势.

21. 在无穷大导体平面上有半径为 a 的圆，圆内和圆外用极狭窄的绝缘环绝缘. 设圆内电势为 V_0，导体板其余部分电势为 0，试用格林函数法求上半空间的电势.

22. 在半径为 a 的中空接地的导体球壳内沿 z 轴放一长为 $2a$ 的均匀荷电细导线（两端与球壳绝缘），总电荷为 Q_f，试用格林函数法求壳内电势.

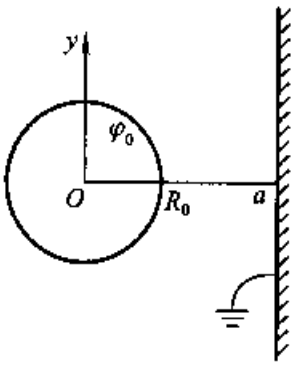


图 4 第 18 题图

提示 球内的电荷密度

$$\rho(r,\theta,\phi) = C[\delta(\theta) + \delta(\theta-\pi)]Q_f/(ar^2\sin\theta).$$

C 为待定系数.

23. 设有两个偶极子，它们的偶极矩大小相等，并都指向 z 轴方向. 其一位置在原点，另一位置在：(a) $\theta=\pi/2$，距原点为 R；(b) $\theta=0$，距原点为 R. 求两种情况下的相互作用能和相互作用力.

24. 有一个电矩为 p_e 的偶极子，位于距无穷大导体平面 a 处. 求导体平面对该偶极子的吸引力.

第二章　静　磁　场

本章讨论电磁场的另一种特殊情形，即只有稳恒电流和永久磁铁的情形. 此时场量也不随时间而变.

静磁场的基本问题是给定电流分布 $\boldsymbol{J}_f(\boldsymbol{x})$ 和磁介质分布，求解磁场.

求解静磁场常用的方法有：①如果电流分布具有轴对称性，可直接用安培回路定律 $\oint \boldsymbol{H}\cdot \mathrm{d}\boldsymbol{l}=I_f$ 求解；②如果仅涉及均匀介质中的磁场，则可以用毕奥-萨伐尔定律 $\boldsymbol{B}(\boldsymbol{x})=\mu\int \mathrm{d}V'\boldsymbol{J}_f(\boldsymbol{x}')\times\boldsymbol{r}/(4\pi r^3)$ 求解；③对于具有一定对称性的问题，可以根据经验提出尝试解，再利用静磁问题的唯一性定理进行验证；④对于较为复杂的问题，则要引入磁矢势 $\boldsymbol{A}$ 或磁标势 φ_m，求解 $\boldsymbol{A}$ 或 φ_m 微分方程的边值问题. 前面两种方法在电磁学中已讨论过，这一章着重讨论后两种方法.

§3.2.1　矢势　静磁问题的唯一性定理

静磁场是由稳恒电流或永磁体激发的. 但是，在稳恒电流激发的磁场问题中，实际上要维持稳恒电流，就要有电源和导线，在电源的极板和导线的表面上就一定会存在电荷分布. 因而我们面临的问题实际上是静电与静磁的问题，这时的麦克斯韦方程组为：

$$\begin{cases} \nabla\times\boldsymbol{E}=0, & (1.1)\\ \nabla\cdot\boldsymbol{D}=\rho_f, & (1.2)\\ \nabla\cdot\boldsymbol{B}=0, & (1.3)\\ \nabla\times\boldsymbol{H}=\boldsymbol{J}_f. & (1.4) \end{cases}$$

不难发现，在稳恒情形下，电场和磁场没有直接联系，可以将电场问题与磁场问题分开解决. 而方程(1.3)、(1.4)即为求解静磁问题的基本方程❶.

本节首先根据(1.3)式引入磁场的矢势. 接着利用矢势给出静磁问题的唯一性定理的证明和应用. 然后导出矢势满足的方程及边值关系. 最后举例说明利用矢势求解静磁问题的方法.

(一)矢势

由(1.3)式 $\nabla\cdot\boldsymbol{B}=0$ 可见磁场是无源场，因为无源场必可表为另一矢量场的旋度，由此得矢势 $\boldsymbol{A}$ 的定义为[即 §2.3.3(3.12)式]

$$\boldsymbol{B}=\nabla\times\boldsymbol{A}. \qquad (1.5)$$

---

❶ 在本节中为了避免和电荷密度 ρ 相混，将柱坐标变量记作 R,ϕ,z.

矢势 $\boldsymbol{A}$ 的物量意义是：$\boldsymbol{A}$ 沿任一闭合回路的环流 $\oint_l \boldsymbol{A}\cdot \mathrm{d}\boldsymbol{l}$ 等于以该回路为界的任一曲面的磁通量 $\int_S \boldsymbol{B}\cdot \mathrm{d}\boldsymbol{\sigma}$，这可以直接利用斯托克斯定理及矢势的定义(1.5)式得证：

$$\oint_l \boldsymbol{A}\cdot \mathrm{d}\boldsymbol{l} = \int_S \nabla \times \boldsymbol{A}\cdot \mathrm{d}\boldsymbol{\sigma} = \int_S \boldsymbol{B}\cdot \mathrm{d}\boldsymbol{\sigma}. \tag{1.6}$$

从矢势的定义可见，若给定 $\boldsymbol{A}$，则 $\boldsymbol{B}$ 就唯一确定. 但对于一个给定的磁场 $\boldsymbol{B}$，相应的矢势 $\boldsymbol{A}$ 却不是唯一的. 这是由于任何函数 ψ 的梯度的旋度恒为零：$\nabla \times (\nabla \psi) = 0$. 因此，如 $\boldsymbol{A}$ 通过 $\nabla \times \boldsymbol{A}$ 给出 $\boldsymbol{B}$，则 $\boldsymbol{A}' = \boldsymbol{A} + \nabla \psi$ 通过 $\nabla \times \boldsymbol{A}' = \nabla \times (\boldsymbol{A} + \nabla \psi) = \nabla \times \boldsymbol{A}$ 也给出相同的 $\boldsymbol{B}$. 由于 $\boldsymbol{A}$ 的这种任意性，可以给 $\boldsymbol{A}$ 加上一定的辅助条件(称为规范条件). 常用的辅助条件是(由这一条件决定的规范称为库仑规范)：

$$\nabla \cdot \boldsymbol{A} = 0. \tag{1.7}$$

这能够给计算带来方便. 但是，对于给定的磁场 $\boldsymbol{B}$，能否找到同时满足(1.5)式与(1.7)式的矢势呢？设 $\nabla \times \boldsymbol{A} = \boldsymbol{B}$，但 $\nabla \cdot \boldsymbol{A} = u(\boldsymbol{x})$，我们来证明，总可以找到一个 $\boldsymbol{A}'$，使 $\nabla \cdot \boldsymbol{A}' = 0$ 与 $\nabla \times \boldsymbol{A}' = \boldsymbol{B}$ 同时成立. 令

$$\boldsymbol{A}' = \boldsymbol{A} + \nabla \psi, \tag{1.8}$$

则

$$\nabla \cdot \boldsymbol{A}' = \nabla \cdot (\boldsymbol{A} + \nabla \psi) = u + \nabla^2 \psi = 0.$$

因此，只要取泊松方程

$$\nabla^2 \psi = -u.$$

的一个解代入(1.8)式，所得的 $\boldsymbol{A}'$ 必同时满足 $\nabla \cdot \boldsymbol{A}' = 0$ 及 $\nabla \times \boldsymbol{A}' = \boldsymbol{B}$ 的要求. 本章后面所取的 $\boldsymbol{A}$ 均假定已满足库仑规范条件 $\nabla \cdot \boldsymbol{A} = 0$.

(二)静磁场的唯一性定理

设区域 V 内电流分布 $\boldsymbol{J}_f(\boldsymbol{x})$ 及磁介质分布给定，并且磁介质内 $\boldsymbol{B} = \mu \boldsymbol{H}$ 成立，在 V 的边界 S 上 $\boldsymbol{A}$ 或 $\boldsymbol{H}$ 的切向分量给定，则 V 内磁场唯一确定.

证 设有两个解 $\boldsymbol{B}'$ 和 $\boldsymbol{B}''$ 满足唯一性定理的条件，即：

$$\boldsymbol{B}' = \nabla \times \boldsymbol{A}', \quad \boldsymbol{B}'' = \nabla \times \boldsymbol{A}''; \tag{1.9}$$

$$\boldsymbol{B}' = \mu \boldsymbol{H}', \quad \boldsymbol{B}'' = \mu \boldsymbol{H}''; \tag{1.10}$$

$$\nabla \times \boldsymbol{H}' = \nabla \times \boldsymbol{H}'' = \boldsymbol{J}_f. \tag{1.11}$$

令

$$\boldsymbol{b} = \boldsymbol{B}' - \boldsymbol{B}'', \quad \boldsymbol{h} = \boldsymbol{H}' - \boldsymbol{H}'', \quad \boldsymbol{a} = \boldsymbol{A}' - \boldsymbol{A}'', \tag{1.12}$$

则

$$\begin{aligned} \nabla \times \boldsymbol{h} &= \nabla \times (\boldsymbol{H}' - \boldsymbol{H}'') = \boldsymbol{J}_f - \boldsymbol{J}_f = 0 \\ \boldsymbol{b} \cdot \boldsymbol{h} &= (\nabla \times \boldsymbol{a}) \cdot \boldsymbol{h} = [\nabla \cdot (\boldsymbol{a} \times \boldsymbol{h}) - \boldsymbol{a} \cdot (\nabla \times \boldsymbol{h})] \\ &= \nabla \cdot (\boldsymbol{a} \times \boldsymbol{h}). \end{aligned} \tag{1.13}$$

将上式两边进行体积分，并利用高斯定理，即有：

$$\int_V \boldsymbol{b}\cdot\boldsymbol{h}\mathrm{d}V=\int_V \nabla\cdot(\boldsymbol{a}\times\boldsymbol{h})\mathrm{d}V=\oint_S(\boldsymbol{a}\times\boldsymbol{h})\cdot\mathrm{d}\boldsymbol{\sigma}$$

$$=\oint_S(\mathrm{d}\boldsymbol{\sigma}\times\boldsymbol{a})\cdot\boldsymbol{h}=\oint_S\mathrm{d}\sigma(\boldsymbol{e}_n\times\boldsymbol{a})\cdot\boldsymbol{h},$$

或

$$\int_V \boldsymbol{b}\cdot\boldsymbol{h}\mathrm{d}V=-\oint_S\mathrm{d}\sigma(\boldsymbol{e}_n\times\boldsymbol{h})\cdot\boldsymbol{a}. \tag{1.14}$$

将(1.12)式代入(1.13)式及(1.14)式后分别得到

$$\int_t(\boldsymbol{B}'-\boldsymbol{B}'')\cdot(\boldsymbol{H}'-\boldsymbol{H}'')\mathrm{d}V=\oint_S[\boldsymbol{e}_n\times(\boldsymbol{A}'-\boldsymbol{A}'')]\cdot(\boldsymbol{H}'-\boldsymbol{H}'')\mathrm{d}\sigma, \tag{1.15}$$

$$\int_V(\boldsymbol{B}'-\boldsymbol{B}'')\cdot(\boldsymbol{H}'-\boldsymbol{H}'')\mathrm{d}V=-\oint_S[\boldsymbol{e}_n\times(\boldsymbol{H}'-\boldsymbol{H}'')]\cdot(\boldsymbol{A}'-\boldsymbol{A}'')\mathrm{d}\sigma, \tag{1.16}$$

若给定 $\boldsymbol{A}$ 在边界 S 上的切向分量，则 $\boldsymbol{e}_n\times\boldsymbol{A}'=\boldsymbol{e}_n\times\boldsymbol{A}''$；若给定 $\boldsymbol{H}$ 在边界 S 上的切向分量，则 $\boldsymbol{e}_n\times\boldsymbol{H}'=\boldsymbol{e}_n\times\boldsymbol{H}''$. 由(1.15)、(1.16)式可见，无论给定边界 S 上 $\boldsymbol{A}$ 的切向分量或 $\boldsymbol{H}$ 的切向分量，$\int_V(\boldsymbol{B}'-\boldsymbol{B}'')\cdot(\boldsymbol{H}'-\boldsymbol{H}'')\mathrm{d}V$ 均为零. 利用(1.10)式，即有

$$\int_V\mathrm{d}V(\boldsymbol{B}'-\boldsymbol{B}'')^2/\mu=\int_V\mathrm{d}V(\boldsymbol{B}'-\boldsymbol{B}'')\cdot(\boldsymbol{H}'-\boldsymbol{H}'')=0. \tag{1.17}$$

因为 $\mu>0$，$(\boldsymbol{B}'-\boldsymbol{B}'')^2\geqslant 0$，当积分为零时必然有 $\boldsymbol{B}'=\boldsymbol{B}''$，即静磁场是唯一的.

【证毕】

下面举例说明如何直接根据静磁问题的唯一性定理解题.

【例 1】 设有无穷长的线电流 I_f 沿 z 轴流动，在 $z<0$ 空间充满磁导率为 μ 的均匀介质，$z>0$ 区域为真空，试用唯一性定理求磁感应强度 $\boldsymbol{B}$，然后求出磁化电流分布.

解 (a)坐标系. 使用柱坐标系，其轴沿线电流，坐标原点位于线电流与介质分界面的交点. 设 $z<0$ 及 $z>0$ 区域的磁感应强度分别为 $\boldsymbol{B}_1$ 和 $\boldsymbol{B}_2$.

(b)方程. 为便于提出尝试解，方程取积分形式：

$z<0$ 区域：
$$\oint_l\boldsymbol{B}_1\cdot\mathrm{d}\boldsymbol{l}=\mu I_f \tag{1.18}$$

$z>0$ 区域：
$$\oint_l\boldsymbol{B}_2\cdot\mathrm{d}\boldsymbol{l}=\mu_0 I_f \tag{1.19}$$

任意封闭面：
$$\oint_S\boldsymbol{B}\cdot\mathrm{d}\boldsymbol{\sigma}=0. \tag{1.20}$$

(c)边界条件与边值关系.

① $\lim\limits_{R\to\infty}\boldsymbol{B}=0$, (1.21)

② 由 $\boldsymbol{e}_n\times(\boldsymbol{H}_2-\boldsymbol{H}_1)|_{z=0}=\boldsymbol{\alpha}_f$ 得 $\boldsymbol{e}_z\times\left(\frac{\boldsymbol{B}_2}{\mu_0}-\frac{\boldsymbol{B}_1}{\mu}\right)\Big|_{z=0}=0$, (1.22)

③ 由 $\boldsymbol{e}_n\cdot(\boldsymbol{B}_2-\boldsymbol{B}_1)|_{z=0}=0$ 得 $\boldsymbol{e}_z\cdot(\boldsymbol{B}_2-\boldsymbol{B}_1)|_{z=0}=0$. (1.23)

(d)提出尝试解.

$$\begin{cases} z<0 \text{ 区域:} & B_1=\mu I_f\boldsymbol{e}_\phi/(2\pi R),\\ z>0 \text{ 区域:} & B_2=\mu_0 I_f\boldsymbol{e}_\phi/(2\pi R). \end{cases} \tag{1.24}$$

(e)验证:不难证明,尝试解满足方程、边界条件和边值关系,由唯一性定理可知,这是正确的解.

(f)求磁化电流分布.先求出磁化强度 $\boldsymbol{M}$,然后由它得磁化电流密度 $\boldsymbol{J}_M(\boldsymbol{x})$、磁化电流 I_M 以及两介质分界面上的磁化电流线密度 $\boldsymbol{\alpha}_M$.

①磁化强度 $\boldsymbol{M}=\boldsymbol{B}/\mu_0-\boldsymbol{H}=(1/\mu_0-1/\mu)\boldsymbol{B}$.

②磁化电流密度 $\boldsymbol{J}_M=\nabla\times\boldsymbol{M}=(1/\mu_0-1/\mu)\nabla\times\boldsymbol{B}$

$$=(1/\mu_0-1/\mu)\mu\,\boldsymbol{J}_f.$$

③磁化电流 $I_M=\int_S\boldsymbol{J}_M\mathrm{d}\boldsymbol{\sigma}=(1/\mu_0-1/\mu)\mu\int_S\boldsymbol{J}_f\cdot d\boldsymbol{\sigma}$,

因为在 $z<0$ 区域,$\mu=\mu$;在 $z>0$ 区域,$\mu=\mu_0$,所以,

$$I_M=\begin{cases}(1/\mu_0-1/\mu)\mu\,I_f, & z<0,R=0;\\ 0, & z>0.\end{cases}$$

由于 $\boldsymbol{J}_M(\boldsymbol{x})=(1/\mu_0-1/\mu)\mu\,\boldsymbol{J}_f(\boldsymbol{x})$ 仅仅存在于 $z<0$ 及 $R=0$ 处,故 I_M 仅存在于 $z<0$ 及 $R=0$ 处.

④磁化电流线密度 $\boldsymbol{\alpha}_M=\boldsymbol{e}_n\times(\boldsymbol{M}_2-\boldsymbol{M}_1)$.

因为 $z>0$ 区域为真空,故 $\boldsymbol{M}_2=0$,在两介质交界面处,$\boldsymbol{e}_n=\boldsymbol{e}_z$,所以,

$$\boldsymbol{\alpha}_M=-\boldsymbol{e}_z\times\boldsymbol{M}_1=-\boldsymbol{e}_z\times(1/\mu_0-1/\mu)\boldsymbol{B}_1$$

$$=-\left(\frac{1}{\mu_0}-\frac{1}{\mu}\right)\frac{\mu I_f}{2\pi R}\boldsymbol{e}_z\times\boldsymbol{e}_\varphi=\left(\frac{\mu}{\mu_0}-1\right)\frac{I_f}{2\pi R}\boldsymbol{e}_R.$$

由此可见,若 $z<0$ 区域为顺磁介质($\mu>\mu_0$),则 I_M 与 I_f 同向,I_M 流至界面 $z=0$ 处成为面电流,以线密度 $\boldsymbol{\alpha}_M$ 沿 $\boldsymbol{e}_R$ 流向无穷远处;若 $z<0$ 区域为抗磁介质($\mu<\mu_0$),则 I_M 与 I_f 反向,在界面 $z=0$ 处以线密度 $\boldsymbol{\alpha}_M$ 沿 $-\boldsymbol{e}_R$ 方向从无穷远处会聚到 $R=0$ 处成为电流 I_M 再沿 $-\boldsymbol{e}_z$ 方向流向无穷远处.

(三)矢势满足的方程与边值关系

在均匀的各向同性的线性介质内部,将 $\boldsymbol{B}=\nabla\times\boldsymbol{A}$,$\nabla\cdot\boldsymbol{A}=0$ 及 $\boldsymbol{B}=\mu\boldsymbol{H}$ 代入 $\nabla\times\boldsymbol{H}=\boldsymbol{J}_f$,即有:

$$\boldsymbol{J}_f=\nabla\times\frac{\boldsymbol{B}}{\mu}=\frac{\nabla\times(\nabla\times\boldsymbol{A})}{\mu}=\frac{\nabla(\nabla\cdot\boldsymbol{A})-\nabla^2\boldsymbol{A}}{\mu}=-\frac{\nabla^2\boldsymbol{A}}{\mu},$$

即

$$\nabla^2 \boldsymbol{A} = -\mu \boldsymbol{J}_f. \tag{1.25}$$

可见 $\boldsymbol{A}$ 满足泊松方程.

值得注意的是,使用直角坐标时 $\boldsymbol{A}$ 的各个分量仍满足泊松方程

$$\nabla^2 A_i = -\mu J_i. \tag{1.26}$$

(为书写方便,暂将 $\boldsymbol{J}_f$ 的 i 分量记作 J_i.)这是由于 $\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z$ 与 x 无关,因此,若用 $\boldsymbol{e}_x$ 点乘(1.25)式两边,可将 $\boldsymbol{e}_x$ 移入拉普拉斯算符之内,即

$$\nabla^2(\boldsymbol{e}_x \cdot \boldsymbol{A}) = \boldsymbol{e}_x \cdot \nabla^2 \boldsymbol{A} = \boldsymbol{e}_x \cdot (-\mu \boldsymbol{J}_f) = -\mu(\boldsymbol{e}_x \cdot \boldsymbol{J}_f)$$

$$\nabla^2 A_x = -\mu J_x.$$

同理 A_y 与 A_z 也满足类似的方程.

但是,曲线坐标系的单位矢量,如球坐标系的 $\boldsymbol{e}_R, \boldsymbol{e}_\theta, \boldsymbol{e}_\varphi$ 及柱坐标系的 $\boldsymbol{e}_\rho, \boldsymbol{e}_\phi$,与 x 有关,不能从拉普拉斯算符的外面移到里面,因面 $\nabla^2 A_\rho \neq \mu(J_f)_\rho$,…等等.

当介质分区均匀时,将 $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ 代入§2.4.2的(2.4)式和(2.9)式,

$$\boldsymbol{e}_n \cdot (\boldsymbol{B}_2 - \boldsymbol{B}_1) = 0, \qquad \boldsymbol{e}_n \times [(\boldsymbol{B}_2/\mu_2) - (\boldsymbol{B}_1/\mu_1)] = \boldsymbol{\alpha}_f,$$

即得矢势的边值关系

$$\boldsymbol{e}_n \cdot [(\nabla \times \boldsymbol{A}_2) - (\nabla \times \boldsymbol{A}_1)] = 0, \tag{1.27}$$

$$\boldsymbol{e}_n \times \left[\frac{1}{\mu_2}(\nabla \times \boldsymbol{A}_2) - \frac{1}{\mu_1}(\nabla \times \boldsymbol{A}_1)\right] = \boldsymbol{\alpha}_f, \tag{1.28}$$

矢势的边值关系还有其他等价的形式.如(1.27)式也可写成更简单的形式.在 $\nabla \times \boldsymbol{A} = \boldsymbol{B}$ 两边点乘 $\mathrm{d}\boldsymbol{\sigma}$ 后作面积分,利用斯托克斯定理后可得

$$\int_s \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{\sigma} = \int_s \nabla \times \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{\sigma} = \oint_l \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{l}. \tag{1.29}$$

将(1.29)式应用于介质分界面上的一个矩形回路,仿照§2.4.2的作法,略去作为高级小量的面积分项(参看该节图3),即有

$$\boldsymbol{e}_t \cdot (\boldsymbol{A}_2 - \boldsymbol{A}_1) = 0. \tag{1.30}$$

式中 $\boldsymbol{e}_t$ 为界面上沿任意方向的切向单位矢量.(1.30)式指出,$\boldsymbol{A}$ 的切向分量连续:$A_{2t} = A_{1t}$.

讨论 当采用库仑规范时,在区域1,2均满足 $\nabla \cdot \boldsymbol{A} = 0$ 的条件,这时应该也可以由

$$\oint_S \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{\sigma} = \int_V \nabla \cdot \boldsymbol{A} \mathrm{d}V = 0, \tag{1.31}$$

仿照§2.4.2的作法,将(1.31)式应用于介质分界面上的一个扁平圆柱盒,得出

$$\boldsymbol{e}_n \cdot (\boldsymbol{A}_2 - \boldsymbol{A}_1) = 0. \tag{1.32}$$

式中 $\boldsymbol{e}_n$ 为界面上法向单位矢量.(1.32)式指出,$\boldsymbol{A}$ 的法向分量连续 $A_{2n} = A_{1n}$.综合

(1.30)式和(1.32)式,即有

$$\boldsymbol{A}_1 = \boldsymbol{A}_2. \tag{1.33}$$

上述论证是否成立的关键在于介质分界面上的$\nabla \cdot \boldsymbol{A}$是否存在.如果$\nabla \cdot \boldsymbol{A}$不存在,则(1.31)式、(1.32)式及(1.33)式均不成立;若$\nabla \cdot \boldsymbol{A}$存在,$\boldsymbol{A}$必连续.在例1中,下述矢势给出(1.24)式表示的磁感应强度(注意,柱坐标的$R = \sqrt{x^2 + y^2}$)

$$\boldsymbol{A}_1 = -\mathbf{e}_z \frac{\mu I_f}{2\pi} \ln R, \qquad \boldsymbol{A}_2 = -\mathbf{e}_z \frac{\mu_0 I_f}{2\pi} \ln R. \tag{1.34}$$

此时在$z=0$平面上$\boldsymbol{A}_1|_{z=0} \neq \boldsymbol{A}_2|_{z=0}$.

但是由于给定$\boldsymbol{B}$之后,$\boldsymbol{A}$不是唯一的,前面已经证明,$\boldsymbol{A}' = \boldsymbol{A} + \nabla\psi$(式中$\psi$满足$\nabla^2\psi = 0$)也给出同样的$\boldsymbol{B}$,而且仍然满足库仑规范的条件.因此总可以找到合适的ψ,使$\boldsymbol{A}'$满足(1.33)式.在上例中若取

$$\psi_1 = \frac{(\mu - \mu_0)}{2\pi} I_f z \ln R, \qquad \psi_2 = 0. \tag{1.35}$$

利用$\boldsymbol{e}_R = \cos\phi \boldsymbol{e}_x + \sin\phi \boldsymbol{e}_y$,可得

$$\boldsymbol{A}'_1 = \boldsymbol{A}_1 + \nabla\psi_1 = -\boldsymbol{e}_z \frac{\mu_0 I_f}{2\pi} \ln R + \boldsymbol{e}_R (\mu - \mu_0) \frac{I_f z}{2\pi R}, \tag{1.36}$$

$$\boldsymbol{A}'_2 = \boldsymbol{A}_2 + \nabla\psi_2 = -\boldsymbol{e}_z \frac{\mu_0 I_f}{2\pi} \ln R. \tag{1.37}$$

利用$\boldsymbol{e}_\phi = -\sin\phi \boldsymbol{e}_x + \cos\phi \boldsymbol{e}_y$即可证明(留给读者):$\boldsymbol{A}'_1$及$\boldsymbol{A}'_2$给出(1.24)式的$\boldsymbol{B}$;$\boldsymbol{A}'_1$与$\boldsymbol{A}'_2$均满足库仑规范的条件;并且$\boldsymbol{A}'_1|_{z=0} = \boldsymbol{A}'_2|_{z=0}$.因此,条件(1.33)也是可以采用的.

*(四)用矢势法求解静磁场

利用矢势法求静磁场一般比较复杂,我们用两个较简单而有意义的例子说明解题的方法.

【例2】 设在全空间充满磁导率为μ的均匀介质,且电流分布$\boldsymbol{J}_f(x)$给定,求矢势.

解 这个题目可以利用与静电问题的类比法求解.我们选用直角坐标系,并为书写方便,暂将$\boldsymbol{J}_f$的i分量记作J_i.定解问题是:

$$\nabla^2 A_i = -\mu J_i, \tag{1.38}$$

$$\lim_{R\to\infty} A_i = 0. \tag{1.39}$$

这个定解问题与静电场中下述问题很相似:在全空间充满介电常数ε的均匀介质,且电荷分布$\rho_f(x)$给定,求静电势.其定解问题是:

$$\nabla^2 \varphi = -\rho_f/\varepsilon, \tag{1.40}$$

$$\lim_{R\to\infty} \varphi = 0. \tag{1.41}$$

可见,两者都满足泊松方程,并且边界条件也相同.因此,只要作替换

$$\varphi \to A_i, \tag{1.42}$$

$$\rho_f/\varepsilon \to \mu J_i. \tag{1.43}$$

即可将(1.40)、(1.41)的解变为(1.38),(1.39)的解.φ 的解为:

$$\varphi = \int dV' \rho_f(x')/(4\pi\varepsilon r), \tag{1.44}$$

因此 A_i 的解为

$$A_i = \int dV' \mu J_i(x')/(4\pi r), \tag{1.45}$$

矢势 A 即为

$$A = \sum_{i=1}^{3} e_i A_i = \sum_{i=1}^{3} e_i \int \frac{\mu J_i(x')\,dV'}{4\pi r} = \int \frac{\mu J_i(x')\,dV'}{4\pi r}. \tag{1.46}$$

*【例 3】 半径为 a 的圆环通有电流 I_f,求矢势与磁感应强度.

解 (a)坐标系.选原点在圆环中心的球坐标系.场点坐标为(R,θ,ϕ),源点坐标为(R',θ',ϕ').

(b)利用对称性,把求 A 的问题简化为求 A_y 的问题.空间的矢势由(46)式给出,将 $\mu=\mu_0$,$J_f dV' = I_f dl'$代入(1.46)式,即有

$$A = \frac{\mu_0}{4\pi}\int \frac{J_f(x')\,dV'}{r} = \frac{\mu_0}{4\pi}\oint \frac{I_f dl'}{r}. \tag{1.47}$$

对于上式中沿圆环的回路积分可作如下分析.从场点(图 1 中 P 点)到圆环所在平面作垂线.从这一垂线和平面的交点(图 1 中 M 点)穿过圆环心作直线 MN,将圆环分成两半.在这两半中处于对称位置的两个无穷小矢量 dl'_1 和 dl'_2 之和沿 e_ϕ 的方向,见图 1.因此,(1.47)式中的回路积分结果,得到一个沿 e_ϕ 方向的矢量.

由此可见,A 只有沿 e_ϕ 方向的分量 A_ϕ,并且积分号前均为给定的常数,因而 A 的大小只与 $r=|x-x'|$有关.既然源点 x'分布在一个圆环上,由对称性可见,A 的大小只与场点的 R 与 θ 坐标有关,而与场点的 ϕ 坐标无关.

既然 A 的大小决定于 A_ϕ,而 A_ϕ 又与 ϕ 无关,因此,选 OM 为 x 轴,求出在 xoz 平面上的 A_ϕ 便可求得 A,而在 xoz 平面上的 $A_\phi = A_y$.于是,求 A 的问题归结为求解 A_y.

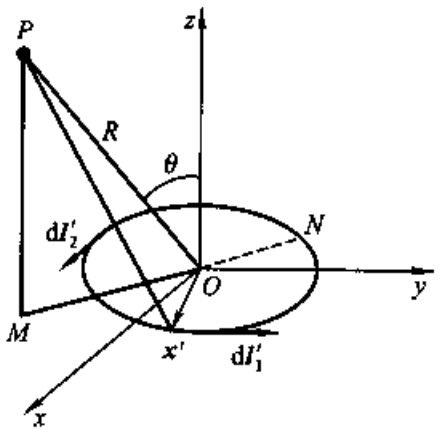


图 1 通电圆环的磁场

(c)求 $A = e_y A_y$.将矢势

$$A = \frac{\mu_0}{4\pi}\oint \frac{I_f dl'}{r},$$

与 e_y 点乘可得

$$A_y = \frac{\mu_0}{4\pi}\oint \frac{I_f dl'_y}{r}. \tag{1.48}$$

但

$$dl'_y = dl'\cos\phi' = a\cos\phi' d\phi', \tag{1.49}$$

$$r = |x-x'| = \sqrt{x^2 + x'^2 - 2x\cdot x'}, \tag{1.50}$$

式中 $x\cdot x'$可以这样计算:因为场点 x 在 xOz 平面上,故 $y=0$;而源点在 xOy 平面上,故 $z'=0$.同

时注意到$|x|=R$,$|x'|=a$,即有

$$x = R\sin\theta e_x + R\cos\theta e_z,$$

$$x' = a\cos\phi' e_x + a\sin\phi' e_y.$$

求两式的点积,并利用单位矢量的正交性,即有

$$x \cdot x' = aR\sin\theta\cos\phi'. \tag{1.51}$$

将(1.49)~(1.51)式代入(1.48)式,可得

$$A_\phi = A_y = \frac{\mu_0}{4\pi}\int_0^{2\pi} \frac{I_f a\cos\phi' d\phi'}{\sqrt{R^2 + a^2 - 2Ra\sin\theta\cos\phi'}}. \tag{1.52}$$

这是一个椭圆积分,通过查表可查出积分值.但对于远场($R \gg a$)及近轴场($R \ll a$ 或 $\sin\theta \ll 1$),因为$2Ra\sin\theta\cos\phi'/(R^2+a^2) \ll 1$,故上式可写成

$$A_\phi = \frac{\mu_0 I_f a}{4\pi}\int_0^{2\pi} \frac{1}{\sqrt{R^2+a^2}} \frac{\cos\phi' d\phi'}{\sqrt{1 - 2Ra\sin\theta\cos\phi'/(R^2+a^2)}}.$$

然后将根式按泰勒级数展开,并取至第二项,即有

$$A = e_\phi A_\phi = \frac{\mu_0 I_f a e_\phi}{4\pi}\int_0^{2\pi} \frac{1}{\sqrt{R^2+a^2}}\left(1 + \frac{Ra\sin\theta}{R^2+a^2}\cos\phi'\right)\cos\phi' d\phi'$$

$$= \frac{\mu_0 I_f a^2 R\sin\theta}{4(R^2+a^2)^{3/2}} e_\phi. \tag{1.53}$$

(d)由 $B = \nabla \times A$ 得 B.利用球坐标的公式,不难求得:

$$B_R = \frac{1}{R\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta A_\phi) \approx \frac{\mu_0 I_f a^2\cos\theta}{2(R^2+a^2)^{3/2}},$$

$$B_\theta = -\frac{1}{R}\frac{\partial}{\partial R}(RA_\phi) \approx \frac{a_0 I_f a^2 (R^2 - 2a^2)}{4(R^2+a^2)^{5/2}}\sin\theta,$$

$$B_\phi = 0.$$

这些公式的适用范围是远场和近轴场.

§3.2.2 磁标势

上节我们引入了矢势 A 求解静磁场.由于 A 是矢量,求 A 比较困难.特别是在曲线坐标中 A 的各分量并不满足泊松方程.球坐标系的 A_R, A_θ, A_ϕ 一般不能分离开来.因此,希望能引入磁标势 φ_m 来求解静磁场.

本节首先讨论引入磁标势的条件,然后导出磁标势满足的方程及边值关系,最后举例说明利用磁标势求解静磁场的方法.

(一)引入磁标势的条件

我们知道,静电场之所以能引入标势 φ 是因为静电场 E 是势场,其数学表示是 $\nabla \times E = 0$,或 $\oint E \cdot dl = 0$.故可令 $E = -\nabla\varphi$.

对于静磁场来说,$\nabla \times H = J_f \neq 0$,因此一般说不能引入磁标势.而且,即使是

把 $J_f \neq 0$ 的区域从求解的区域中挖掉，则余下的地方虽然有 $\nabla \times H = 0$，仍然还不能引入磁标势. 例如图 1(a)中一个通电圆线圈，如果仅仅把圆线圈所在的空间挖去，余下的空间中的磁场仍然不是势场. 例如图 1(a)的回路 L 与电流 I_f 相链环，由安培环路定律可知 $\oint H \cdot dl = I_f$. 这说明 H 的线积分与路径有关，不能定出一个仅依赖于位置的势函数. 由此可见，能够引入磁标势的条件是：在求解的区域内满足 $\oint H \cdot dl = 0$.

那么，在实际问题中如何实现这个条件呢？以上面的电流圈为例，为了保证 $\oint H \cdot dl = 0$，就要求在求解的区域 V 内任何封闭回路都不与电流 I_f 相链环. 用数学语言就是：在求解区域内任何封闭回路都能通过连续的变形而缩成一点. 为了达到这一要求，就要挖去以电流圈为边界的一个曲面，若讨论的是距通电流圈远处的静磁场，则可将以线圈为边界的内部曲面挖掉(见图 1(b)的阴影线部分)；若讨论的是通电流线圈轴线附近的静磁场，则可将以线圈为边界的外部曲面挖掉(见图 1(c)的阴影部分).

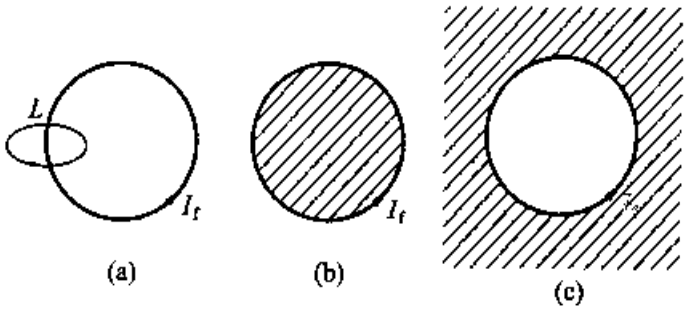


图 1 由复连通区变成单连通区

如果一个区域内任何封闭回路都可以通过连续变形缩成一点，则这一区域就称为单连通区域. 因此，引入磁标势的条件可以归纳为：没有传导电流的单连通区域.

(二)磁标势的定义 磁标势满足的方程与边值关系

磁标势 φ_m 的定义为：

$$H = -\nabla \varphi_m. \tag{2.1}$$

为了导出磁标势 φ_m 满足的方程，先引入假想的束缚磁荷密度的概念. 为此，将 $B = \mu_0(H + M)$ 代入 $\nabla \cdot B = 0$，即有：

$$\nabla \cdot B = \mu_0(\nabla \cdot H + \nabla \cdot M) = 0,$$

或

$$\nabla \cdot H = -\nabla \cdot M. \tag{2.2}$$

在讨论静电场时也遇到过类似的方程：

$$\nabla \cdot E = \rho_f/\varepsilon = (\rho_f + \rho_p)/\varepsilon_0. \tag{2.3}$$

通过类比，引入假想的束缚磁荷密度 ρ_m，则

$$\nabla \cdot H = \rho_m/\mu_0. \tag{2.4}$$

当然，电学量与磁学量有区别. 首先，这里没有和自由电荷密度 ρ_f 相应的自由磁荷

密度. 这是由于不存在"磁单极"的缘故. 其次,对于均匀磁介质,磁导率 μ 为常数,因而可以提出∇号之外,即

$$\rho_m/\mu_0 = \nabla \cdot \boldsymbol{H} = \nabla \cdot (\boldsymbol{B}/\mu) = (1/\mu)\nabla \cdot \boldsymbol{B} = 0. \tag{2.5}$$

这说明,在均匀磁介质内部肯定没有体束缚磁荷密度 ρ_m 的分布. 例如,一根铁棒被均匀磁化,在铁棒内部没有束缚磁荷的体分布,而只有面分布. 但是对于电介质,在介质内部有自由电荷分布的地方,就有束缚电荷的体分布. 以均匀电介质为例(ε 为常数)

$$\rho_P = -\nabla \cdot \boldsymbol{P} = -\nabla \cdot (\boldsymbol{D} - \varepsilon_0 \boldsymbol{E}) = -\left(1 - \frac{\varepsilon_0}{\varepsilon}\right)\nabla \cdot \boldsymbol{D} = \left(\frac{\varepsilon_0}{\varepsilon} - 1\right)\rho_f. \tag{2.6}$$

可见这个差异的根源也在于没有自由磁荷密度.

现在来推导磁标势 φ_m 所满足的方程. 将磁标势的定义(2.1)式代入束缚磁荷密度的定义(2.4)式,即得

$$\nabla^2 \varphi_m = -\rho_m/\mu_0. \tag{2.7}$$

可见磁标势 φ_m 满足泊松方程.

在均匀介质内部,$\rho_m = 0$,φ_m 满足拉普拉斯方程

$$\nabla^2 \varphi_m = 0. \tag{2.8}$$

与静电场比较,差别仍在于没有"自由磁荷".

磁标势的边值关系:当界面上没有传导电流时,

$$\varphi_{m1} = \varphi_{m2}, \tag{2.9}$$

当界面上有传导电流 $\boldsymbol{\alpha}_f$ 时,

$$\boldsymbol{e}_n \times (\nabla \varphi_{m1} - \nabla \varphi_{m2}) = \boldsymbol{\alpha}_f, \tag{2.10}$$

线性介质:

$$\mu_1 \frac{\partial \varphi_{m1}}{\partial n} = \mu_2 \frac{\partial \varphi_{m2}}{\partial n}, \tag{2.11}$$

任意介质:

$$\mu_0\left(\frac{\partial \varphi_{m2}}{\partial n} - \frac{\partial \varphi_{m1}}{\partial n}\right) = \mu_0 \boldsymbol{e}_n \cdot (\boldsymbol{M}_2 - \boldsymbol{M}_1) = -\sigma_m. \tag{2.12}$$

其中 $\sigma_m = -\mu_0 \boldsymbol{e}_n \cdot (\boldsymbol{M}_2 - \boldsymbol{M}_1)$ 称为面束缚磁荷密度. 它是与静电场中面束缚电荷密度 $\sigma_P = -\boldsymbol{e}_n \cdot (\boldsymbol{P}_2 - \boldsymbol{P}_1)$ 相类比而得名.

(2.9)式的证明与静电场 $\varphi_1 = \varphi_2$ 的证明相仿. 由磁场边值关系 $\boldsymbol{e}_n \times (\boldsymbol{H}_2 - \boldsymbol{H}_1) = \boldsymbol{\alpha}_f$,可证明(2.10)式. (2.12)式的证明可将 $\boldsymbol{H} = -\nabla \varphi_m$ 代入 $\boldsymbol{B} = \mu \boldsymbol{H}$,得

$$B_n = \boldsymbol{e}_n \cdot \boldsymbol{B} = -\mu \boldsymbol{e}_n \cdot \nabla \varphi_m = -\mu \partial \varphi_m / \partial n .$$

代入 $B_{1n} = B_{2n}$ 即为(2.11)式. (2.12)式的证明由 $\boldsymbol{B} = \mu_0(\boldsymbol{H} + \boldsymbol{M})$ 出发,这里 $\boldsymbol{B}$ 与 $\boldsymbol{M}$ 以及 $\boldsymbol{B}$ 与 $\boldsymbol{H}$ 不一定存在线性关系. 用 $\boldsymbol{e}_n$ 点乘之后移项,即有

$$\boldsymbol{e}_n \cdot (\boldsymbol{H}_1 - \boldsymbol{H}_2) = \boldsymbol{e}_n \cdot (\boldsymbol{M}_2 - \boldsymbol{M}_1).$$

再利用 $\boldsymbol{e}_n \cdot \boldsymbol{H} = -\partial \varphi_m / \partial n$ 及面束缚磁荷密度 σ_m 的定义代入上式,即可得(2.12)

式. 显然,(2.12)式为任意磁介质的边值关系. 只要将线性介质的条件 $\boldsymbol{B}=\mu\boldsymbol{H}$ 代入 $\boldsymbol{M}$,便有

$$\boldsymbol{M}=\left(\frac{\boldsymbol{B}}{\mu_0}-\boldsymbol{H}\right)=\left(\frac{\mu}{\mu_0}-1\right)\boldsymbol{H}=-\left(\frac{\mu}{\mu_0}-1\right)\nabla\varphi_m.$$

再将此式代入(2.12)式,便可得到(2.11)式.

(三)用磁标势求解静磁场

现在举例说明用磁标势解题的步骤.

【例 1】 求磁化矢量为 $\boldsymbol{M}_0$ 的均匀磁化铁球产生的磁场.

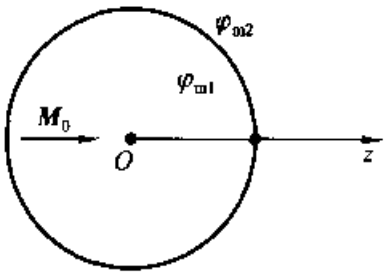


图 2 均匀磁化铁球的磁场

解 利用分离变量法解题.

(a)参考系与磁标势零参考点. 选坐标原点在球心,z 轴沿 $\boldsymbol{M}_0$ 的球坐标系,如图 2. 磁标势的零参考点取在无穷远点.

(b)方程. 因真空及均匀介质中的 $\rho_m=0$,故

$$\nabla^2\varphi_{m1}=0,\qquad \nabla^2\varphi_{m2}=0. \tag{2.13}$$

(c)对称性及通解的形式. 本题有轴对称性,故

$$\varphi_{m1}=\sum a_l R^l \mathrm{P}_l(\cos\theta),\qquad \varphi_{m2}=\sum b_l R^{-(l+1)}\mathrm{P}_l(\cos\theta). \tag{2.14}$$

这个形式已考虑到 $\varphi_{m1}|_{R=0}$ 有限及 $\varphi_{m2}|_{R\to\infty}=0$.

(d)边值关系.

① $$\varphi_{m1}|_{R=R_0}=\varphi_{m2}|_{R=R_0}, \tag{2.15}$$

② $$[\partial\varphi_{m2}/\partial R-\partial\varphi_{m1}/\partial R]|_{R=R_0}=-\boldsymbol{e}_R\cdot\boldsymbol{M}_0. \tag{2.16}$$

后式是将 $\boldsymbol{e}_n=\boldsymbol{e}_R$,$\boldsymbol{M}_1=\boldsymbol{M}_0$,$\boldsymbol{M}_2=0$ 代入(2.12)式得到的.

(e)定系数,得磁标势. 不难求得:

$$\varphi_{m1}=\frac{\boldsymbol{M}_0\cdot\boldsymbol{R}}{3},\qquad \varphi_{m2}=\frac{R_0^3\boldsymbol{M}_0\cdot\boldsymbol{R}}{3R^3}. \tag{2.17}$$

(f)球内外的磁场.

下节的(3.12)式将给出磁偶极子的磁标势为 $\varphi_m=\boldsymbol{m}\cdot\boldsymbol{R}/(4\pi R^3)$,与(2.17)式比较可见,均匀磁化铁球外的磁场是磁偶极子产生的场,磁矩为

$$\boldsymbol{m}=4\pi R_0^3\boldsymbol{M}_0/3=V\boldsymbol{M}_0. \tag{2.18}$$

其中 V 是铁球的体积,$\boldsymbol{M}_0$ 是单位体积内的磁偶极矩. (2.18)式的意义是显然的.

球内的磁场可由(2.17)式求得:

$$\begin{aligned}\boldsymbol{H}_1&=-\nabla\varphi_{m1}=-\boldsymbol{M}_0/3,\\ \boldsymbol{B}_1&=\mu_0(\boldsymbol{H}+\boldsymbol{M}_0)=2\mu_0\boldsymbol{M}_0/3.\end{aligned} \tag{2.19}$$

§3.2.3 磁多极子

本节讨论在小区域 V 内的稳恒电流分布 $\boldsymbol{J}_f(\boldsymbol{x}')$ 在远处所产生的磁场. 我们证明,这个磁场的矢势也可以看成描述所讨论的电流分布的各级磁多极子在远处产生的矢势之和.

电流分布 $\boldsymbol{J}_f(\boldsymbol{x}')$ 激发的磁矢势为

$$\boldsymbol{A}(\boldsymbol{x}) = \mu_0\int_V \mathrm{d}V' \frac{\boldsymbol{J}_f(\boldsymbol{x}')}{4\pi r}. \tag{3.1}$$

其中源点到场点的距离 $r = |\boldsymbol{x} - \boldsymbol{x}'|$,将 $1/r$ 展开代入上式,得

$$\boldsymbol{A}(\boldsymbol{x}) = \frac{\mu_0}{4\pi}\int \mathrm{d}V' \boldsymbol{J}_f(\boldsymbol{x}')\left[\frac{1}{R} - \boldsymbol{x}'\cdot\nabla\frac{1}{R} + \frac{1}{2!}(\boldsymbol{x}'\cdot\nabla)^2\frac{1}{R} + \cdots\right], \tag{3.2}$$

由于数量级的考虑,通常略去磁四极矩及更高级的磁多极矩,只讨论展开式的前两项.

第一项为 $(\mu_0/4\pi R)\int_V \mathrm{d}V'\boldsymbol{J}_f(\boldsymbol{x}')$,$R$ 与积分变量 $\boldsymbol{x}'$ 无关,已提出积分号外. 容易证明积分 $\int_V \mathrm{d}V'\boldsymbol{J}_f(\boldsymbol{x}')$ 为零. 由于稳恒电流的连续性,$\boldsymbol{J}_f(\boldsymbol{x}')$ 线各自形成闭合曲线. 如果将一束 $\boldsymbol{J}_f(\boldsymbol{x}')$ 线在空间所构成的管道称为流管,则对每一小流管说来,线元 $\mathrm{d}\boldsymbol{l}$ 的方向与 $\boldsymbol{J}_f(\boldsymbol{x}')$ 的方向一致,故

$$\boldsymbol{J}_f(\boldsymbol{x}')\mathrm{d}V' = \boldsymbol{J}_f(\boldsymbol{x}')\mathrm{d}s\mathrm{d}\boldsymbol{l} = I_f\mathrm{d}\boldsymbol{l}, \tag{3.3}$$

这样,对每一小流管均有

$$\int \boldsymbol{J}_f(\boldsymbol{x}')\mathrm{d}V' = \oint I_f\mathrm{d}\boldsymbol{l} = I_f\oint\mathrm{d}\boldsymbol{l} = 0.$$

因为全空间中所有流管的结果相同,故(3.2)式的第一项对磁矢势的贡献为零. 实际上,这是不存在磁单极的数学表示. $\boldsymbol{A}^{(0)}(\boldsymbol{x}) = 0$ 意味着体系的总磁荷项对磁矢势无贡献.

(3.2)式的第二项 $\boldsymbol{A}^{(1)}(\boldsymbol{x})$ 是体系磁偶极矩的贡献. 利用

$$\boldsymbol{x}'\cdot\nabla\frac{1}{R} = \sum_{i=1}^{3} x'_i\nabla_i\frac{1}{R},$$

可将它改写为

$$\boldsymbol{A}^{(1)}(\boldsymbol{x}) = -\frac{\mu_0}{4\pi}\sum_{i=1}^{3}\nabla_i\frac{1}{R}\int \boldsymbol{J}_f(\boldsymbol{x}')x'_i\mathrm{d}V'. \tag{3.4}$$

为书写方便,暂将 $\boldsymbol{J}_f(\boldsymbol{x})$ 的 i 分量记作 J_i,再在积分内加减 $J_i(\boldsymbol{x}')$,即有

$$\int \boldsymbol{J}_f(\boldsymbol{x}')x'_i\mathrm{d}V' = \frac{1}{2}\int[(J_fx'_i + J_ix') + (J_fx'_i - J_ix')]\mathrm{d}V' \tag{3.5}$$

利用第二篇第四章习题 4 的结果

$$\nabla' \cdot (\boldsymbol{J}_f x' x'_i) = \boldsymbol{J}_f x'_i + J_i x', \tag{3.6}$$

可将(3.5)式右边的第一项改写为

$$\int (\boldsymbol{J}_f x'_i + J_i x') \mathrm{d}V' = \int \nabla' \cdot (\boldsymbol{J}_f x' x'_i) \mathrm{d}V' = \oint \mathrm{d}\sigma' \cdot (\boldsymbol{J}_f x' x'_i) = 0.$$

这是因为包围整个电流体系的封闭面上 $\boldsymbol{J}_f(x') = 0$. 于是,(3.5)式可写为

$$\int \boldsymbol{J}_f(x') x'_i \mathrm{d}V' = \frac{1}{2} \int (\boldsymbol{J}_f x'_i - \boldsymbol{J}_i x') \mathrm{d}V'. \tag{3.7}$$

代入(3.4)式,即有

$$\begin{aligned} \boldsymbol{A}^{(1)}(x) &= -\frac{\mu_0}{8\pi} \int \sum_{i=1}^{3} (\boldsymbol{J}_f x'_i - x' J_i) \nabla_i \frac{1}{R} \mathrm{d}V' \\ &= -\frac{\mu_0}{8\pi} \int \left[\boldsymbol{J}_f \left(x' \cdot \nabla \frac{1}{R} \right) - x' \left(\boldsymbol{J}_f \cdot \nabla \frac{1}{R} \right) \right] \mathrm{d}V' \\ &= -\frac{\mu_0}{4\pi} \left[\frac{1}{2} \int x' \times \boldsymbol{J}_f \mathrm{d}V' \right] \times \nabla \frac{1}{R}. \end{aligned} \tag{3.8}$$

这里利用了三矢量矢积的公式以及 $\nabla(1/R)$ 与积分变量无关,可置入或提出积分号的性质.

定义电流分布的磁偶极矩为

$$\boldsymbol{m} = \frac{1}{2} \int_V x' \times \boldsymbol{J}_f(x') \mathrm{d}V'. \tag{3.9}$$

代入(3.8)式可得

$$\boldsymbol{A}^{(1)}(x) = -\frac{\mu_0}{4\pi} \boldsymbol{m} \times \nabla \frac{1}{R} = \frac{\mu_0 \boldsymbol{m} \times \boldsymbol{R}}{4\pi R^3}, \tag{3.10}$$

这就是电流体系的磁偶极矩在远处所激发的磁矢势. 与 $\boldsymbol{A}^{(1)}(x)$ 相应的磁场为

$$\begin{aligned} \boldsymbol{B}^{(1)} &= \nabla \times \boldsymbol{A}^{(1)} = \nabla \times \left[-\frac{\mu_0}{4\pi} \boldsymbol{m} \times \nabla \frac{1}{R} \right] \\ &= -\frac{\mu_0}{4\pi} \boldsymbol{m} \nabla^2 \frac{1}{R} + \frac{\mu_0}{4\pi} (\boldsymbol{m} \cdot \nabla) \nabla \frac{1}{R} = \frac{\mu_0}{4\pi} (\boldsymbol{m} \cdot \nabla) \nabla \frac{1}{R}. \end{aligned} \tag{3.11}$$

由于这里讨论的是远离区域 V 的场,$R \neq 0$,故 $\nabla^2(1/R) = 0$. 若取

$$\varphi_m^{(1)} = \boldsymbol{m} \cdot \boldsymbol{R} / (4\pi R^3), \tag{3.12}$$

利用第二篇第四章习题 2 的结果,$\boldsymbol{B}^{(1)}$ 可表示为:

$$\boldsymbol{B}^{(1)} = \nabla \times \boldsymbol{A}^{(1)} = \nabla \times \frac{\mu_0 \boldsymbol{m} \times \boldsymbol{R}}{4\pi R^3} = -\mu_0 \nabla \varphi_m^{(1)}, \tag{3.13}$$

这与磁标势理论得到的结果 §3.2.2(2.1)式一致.

最后,对于电流为 I_f 面积为 a 的平面通电线圈,可以利用(3.7)式证明它的磁矩为 $m = I_f \boldsymbol{a}$,其中面积元矢量 $\boldsymbol{a}$ 的方向与电流方向有右手螺旋关系.(留作练习).

§3.2.4 静磁场的能量 电流分布与静磁场的相互作用

(一)静磁场的能量

在§2.4.3中给出各向同性的线性介质中电磁场的能量密度为

$$w = (\boldsymbol{E}\cdot\boldsymbol{D}+\boldsymbol{H}\cdot\boldsymbol{B})/2, \tag{4.1}$$

见该节的(3.11)式,因而磁场能量密度为:

$$w = \boldsymbol{H}\cdot\boldsymbol{B}/2. \tag{4.2}$$

将稳恒电流磁场的麦克斯韦方程§3.2.1(1.4)式$\nabla\times\boldsymbol{H}=\boldsymbol{J}_f$代入可得静磁场能量为

$$\begin{aligned} E &= \frac{1}{2}\int_\infty \boldsymbol{B}\cdot\boldsymbol{H}\mathrm{d}V = \frac{1}{2}\int_\infty(\nabla\times\boldsymbol{A})\cdot\boldsymbol{H}\mathrm{d}V \\ &= \frac{1}{2}\int_\infty \nabla\cdot(\boldsymbol{A}\times\boldsymbol{H})\mathrm{d}V + \frac{1}{2}\int_\infty \boldsymbol{A}\cdot(\nabla\times\boldsymbol{H})\mathrm{d}V \\ &= \frac{1}{2}\oint_\infty \mathrm{d}\boldsymbol{\sigma}\cdot(\boldsymbol{A}\times\boldsymbol{H}) + \frac{1}{2}\int\boldsymbol{A}\cdot\boldsymbol{J}_f\,\mathrm{d}V \\ &= \frac{1}{2}\int_\infty \boldsymbol{A}\cdot\boldsymbol{J}_f\,\mathrm{d}V. \end{aligned} \tag{4.3}$$

其中,由于$\boldsymbol{A}^{(0)}=0;\boldsymbol{A}^{(1)}\propto 1/R^2,\boldsymbol{H}^{(1)}\propto 1/R^3;\boldsymbol{A}^{(2)}\propto 1/R^3,\boldsymbol{H}^{(2)}\propto 1/R^4,\cdots,\mathrm{d}\sigma\propto R^2$,故当$R\to\infty$时面积分项为零.

(二)电流分布在外磁场中的能量

电流分布在外磁场中的能量即电流分布与外磁场的相互作用能.

设外磁场的矢势$\boldsymbol{A}_e(\boldsymbol{x})$由$\boldsymbol{J}_e(\boldsymbol{x})$激发,电流分布$\boldsymbol{J}_f(\boldsymbol{x})$激发的矢势为$\boldsymbol{A}(\boldsymbol{x})$.

若把两个电流分布$\boldsymbol{J}_e(\boldsymbol{x})$与$\boldsymbol{J}_f(\boldsymbol{x})$合起来看作一个大的电流分布,其磁场能由(4.3)式给出:

$$E = \frac{1}{2}\int(\boldsymbol{A}+\boldsymbol{A}_e)\cdot(\boldsymbol{J}_f+\boldsymbol{J}_e)\mathrm{d}V. \tag{4.4}$$

但是当$\boldsymbol{J}_e(\boldsymbol{x})$和$\boldsymbol{J}_f(\boldsymbol{x})$单独存在时,它们各自的磁场能量为

$$E_e = \frac{1}{2}\int\boldsymbol{A}_e\cdot\boldsymbol{J}_e\,\mathrm{d}V, \tag{4.5}$$

$$E_f = \frac{1}{2}\int\boldsymbol{A}\cdot\boldsymbol{J}_f\,\mathrm{d}V, \tag{4.6}$$

因而电流分布$\boldsymbol{J}_f(\boldsymbol{x})$在外磁场$\boldsymbol{A}_e(\boldsymbol{x})$中的能量,亦即电流分布与外磁场的相互作用能就是

$$E_{相互} = E - E_e - E_f = \int\boldsymbol{A}_e\cdot\boldsymbol{J}_f\,\mathrm{d}V, \tag{4.7}$$

其中应用了

$$\int \boldsymbol{A}(x)\cdot\boldsymbol{J}_e(x)\,\mathrm{d}V=\int\left[\frac{\mu_0}{4\pi}\int\frac{\boldsymbol{J}_f(x')\,\mathrm{d}V'}{r}\right]\cdot\boldsymbol{J}_e(x)\,\mathrm{d}V$$

$$=\int\left[\frac{\mu_0}{4\pi}\int\frac{\boldsymbol{J}_e(x)\,\mathrm{d}V}{r}\right]\cdot\boldsymbol{J}_f(x')\,\mathrm{d}V'$$

$$=\int\boldsymbol{A}_e(x')\cdot\boldsymbol{J}_f(x')\,\mathrm{d}V'.$$

将最后一式的积分变量 x' 改为 x，就得到(4.7)式.

(三)磁偶极子与外磁场的相互作用能

现在讨论一个磁矩为 $\boldsymbol{m}=I_f\boldsymbol{a}$ 的通电平面线圈与外磁场的相互作用能. 将 $\boldsymbol{J}_f\mathrm{d}V=I_f\mathrm{d}\boldsymbol{l}$ 及 $\oint\boldsymbol{A}_e\cdot\mathrm{d}\boldsymbol{l}=\int_S\boldsymbol{B}_e\cdot\mathrm{d}\boldsymbol{\sigma}$ 代入(4.7)式，就有

$$E^{(1)}_{\text{相互}}=\int\boldsymbol{A}_e\cdot\boldsymbol{J}_f\,\mathrm{d}V=I_f\oint\boldsymbol{A}_e\cdot\mathrm{d}\boldsymbol{l}=I_f\int_S\nabla\times\boldsymbol{A}_e\cdot\mathrm{d}\boldsymbol{\sigma}=I_f\int_S\boldsymbol{B}_e\cdot\mathrm{d}\boldsymbol{\sigma}.\quad(4.8)$$

取坐标原点在线圈平面内，将 $\boldsymbol{B}_e(x)$ 在原点展开：

$$\boldsymbol{B}_e(x)=\boldsymbol{B}_e(0)+x\cdot\nabla\boldsymbol{B}_e(0)+\frac{1}{2}(x\cdot\nabla)^2\boldsymbol{B}_e(0)+\cdots.$$

将第一项(零级近似)$\boldsymbol{B}_e(0)$ 代入(4.8)式，并将它提出积分号外(因为它是与积分变量无关的常量)，利用线圈平面的面积元矢量 $\boldsymbol{a}=\int_S\mathrm{d}\boldsymbol{\sigma}$，就有

$$E^{(1)}_{\text{相互}}=I_f\int_S\boldsymbol{B}_e(0)\cdot\mathrm{d}\boldsymbol{\sigma}=\boldsymbol{B}_e(0)\cdot I_f\boldsymbol{a}=\boldsymbol{m}\cdot\boldsymbol{B}_e(0).\quad(4.9)$$

(4.9)式实际上可推广到小区域内的电流分布. 因为一般电流分布可以分成许多闭合流管，对每一流管(4.9)式成立. 将所有流管的贡献叠加起来，就对整个电流分布成立.

值得注意的是，在形式上(4.9)式和电偶极子与外电场的相互作用能 §3.1.6(6.20)式 $E^{(1)}=-\boldsymbol{p}\cdot\boldsymbol{E}_e(0)$ 差一负号，其物理根源留在本节最后讨论.

(四)磁偶极子在外磁场中所受的力、力矩及其势能

由洛伦兹力密度公式 $\boldsymbol{f}=\rho_f\boldsymbol{E}+\boldsymbol{J}_f\times\boldsymbol{B}$ 可见，分布在小区域 V 内的电流 $\boldsymbol{J}_f(x)$ 受外磁场 $\boldsymbol{B}_e(x)$ 的作用力为

$$\boldsymbol{F}=\int_V\boldsymbol{J}_f(x')\times\boldsymbol{B}_e(x')\,\mathrm{d}V',\quad(4.10)$$

坐标原点仍取在电流分布区域 V 内. 在坐标原点展开 $\boldsymbol{B}_e(x')$，并代入(4.10)式，得[注意 $\nabla\boldsymbol{B}_e(0)$ 是二阶张量]

$$\boldsymbol{F}=\int_V\boldsymbol{J}_f(x')\times\boldsymbol{B}_e(0)\,\mathrm{d}V'+\int_V\boldsymbol{J}_f(x')\times[(x'\cdot\nabla)\boldsymbol{B}_e(0)]\,\mathrm{d}V'+\cdots,\quad(4.11)$$

第一项由于 $\boldsymbol{B}_e(0)$ 与积分变量无关而提出积分号外，再利用稳恒电流 $\boldsymbol{J}_f(x')$ 的连续性可得 $\int_V\boldsymbol{J}_f(x')\mathrm{d}V'=0$，于是

$$\boldsymbol{F}^{(0)} = -\boldsymbol{B}_e(0) \times \int_V \boldsymbol{J}_f(x')\mathrm{d}V' = 0. \tag{4.12}$$

第二项是磁偶极矩项. 利用坐标原点在 V 内而外磁场电流分布 $\boldsymbol{J}_e(x)$ 在 V 外可得 $\boldsymbol{J}_e(0)=0$,因此 $\nabla \times \boldsymbol{B}_e(0)=\mu_0 \boldsymbol{J}_e(0)=0$. 于是:

$$x' \times [\nabla \times \boldsymbol{B}_e(0)] = \nabla [x' \cdot \boldsymbol{B}_e(0)] - (x' \cdot \nabla)\boldsymbol{B}_e(0) = 0.$$

这里 $\nabla \boldsymbol{B}_e(0) = \nabla \boldsymbol{B}_e(x)|_{x=0}$. 将 $x' \cdot \nabla \boldsymbol{B}_e(0) = \nabla[x' \cdot \boldsymbol{B}_e(0)]$ 代入(4.11)式第二项的被积函数后,再利用 $\nabla \times \boldsymbol{J}_f(x')=0$[因为 $\boldsymbol{J}_f(x')$ 与 x 无关],加上等于零的 $[x' \cdot \boldsymbol{B}_e(0)]\nabla \times \boldsymbol{J}_f(x')$ 项,即有

$$\begin{aligned}\boldsymbol{J}_f(x') \times [x' \cdot \nabla \boldsymbol{B}_e(0)] &= \boldsymbol{J}_f(x') \times \nabla [x' \cdot \boldsymbol{B}_e(0)] \\ &= -\{\nabla [x' \cdot \boldsymbol{B}_e(0)] \times \boldsymbol{J}_f(x') + [x' \cdot \boldsymbol{B}_e(0)] \nabla \times \boldsymbol{J}_f(x')\} \\ &= -\nabla \times \{\boldsymbol{J}_f(x')[x' \cdot \boldsymbol{B}_e(0)]\},\end{aligned}$$

代入(4.11)式第二项后,因为 ∇ 与积分变量 x' 无关而将 ∇ 提到积分号外,即得

$$\begin{aligned}\boldsymbol{F}^{(1)} &= -\nabla \times \int \boldsymbol{J}_f(x')[x' \cdot \boldsymbol{B}_e(0)]\mathrm{d}V' \\ &= -\nabla \times \sum_{i=1}^{3} B_i(0) \int \boldsymbol{J}_f(x') x'_i \mathrm{d}V' \Big].\end{aligned} \tag{4.13}$$

利用 §3.2.3(3.6)式后,再利用两矢量的点积及三矢量的矢积公式,可得❶:

$$\begin{aligned}\boldsymbol{F}^{(1)} &= -\nabla \times \Big[\sum_{i=1}^{3} B_i(0) \int \frac{1}{2}(\boldsymbol{J}_f x'_i - x' J_i)\mathrm{d}V'\Big] \\ &= -\nabla \times \int \frac{1}{2}\{\boldsymbol{J}_f[x' \cdot \boldsymbol{B}_e(0)] - x'[\boldsymbol{J}_f \cdot \boldsymbol{B}_e(0)]\}\mathrm{d}V' \\ &= -\nabla \times \Big\{\Big[\frac{1}{2}\int x' \times \boldsymbol{J}_f(x')\mathrm{d}V'\Big] \times \boldsymbol{B}_e(0)\Big\} \\ &= -\nabla \times [\boldsymbol{m} \times \boldsymbol{B}_e(0)].\end{aligned} \tag{4.14}$$

这里,已经利用了电流分布的磁矩 $\boldsymbol{m}$ 的定义式 §3.2.3(3.9)式. 由于 $\boldsymbol{m}$ 为常矢量,$\boldsymbol{B}_e(0)$ 的散度为零,上式可写为

$$\boldsymbol{F}^{(1)} = (\boldsymbol{m} \cdot \nabla)\boldsymbol{B}_e(0) = \nabla [\boldsymbol{m} \cdot \boldsymbol{B}_e(0)]. \tag{4.15}$$

这与电偶极子 $\boldsymbol{p}_e$ 在静电场中受力的公式 §3.1.6(6.21)式 $\boldsymbol{F}^{(1)} = \nabla(\boldsymbol{p}_e \cdot \boldsymbol{E}_e)$ 相似. (4.15)式指出一个很有趣的事实:虽然磁场不是保守场($\nabla \times \boldsymbol{B} \neq 0$),但磁偶极子在磁场中所受到的作用力 $\boldsymbol{F}^{(1)}$ 却是一个保守力,它等于势函数 $-\boldsymbol{m} \cdot \boldsymbol{B}_e(0)$ 的负梯度. 我们称

$$U^{(1)} = -\boldsymbol{m} \cdot \boldsymbol{B}_e(0) \tag{4.16}$$

---

❶ 为书写简单起见,$\boldsymbol{B}_e$ 及 $\boldsymbol{J}_f$ 的分量分别简记为 B_i、J_i.

为磁偶极子在磁场中的势能，它准确到一个与磁偶极子位置无关的常数. (4.16)式说明，磁场与磁偶极子的相互作用有使磁偶极子与磁场平行的倾向. 因为在这种情况下，它的势能最低.

电流分布 $J_f(x')$ 在外磁场中所受的力矩有两种计算方法.

方法一：将 $U^{(1)} = -m \cdot B_e(0)$ 代入 $M_{力}^{(1)} = -\partial U^{(1)}/\partial\theta$ 可得

$$M_{力}^{(1)} = -\frac{\partial}{\partial\theta}[-m \cdot B_e(0)] = \frac{\partial}{\partial\theta}[mB_e(0)\cos\theta] = -mB_e(0)\sin\theta,$$

式中负号表示力矩使 θ 角减小. 写成矢量公式即为

$$M_{力}^{(1)} = m \times B_e(0).$$

方法二：

$$M_{力} = \int_V x' \times f(x')\mathrm{d}V' = \int_V x' \times [J_f(x') \times B_e(x')]\mathrm{d}V'. \tag{4.17}$$

将 $B(x')$ 在坐标原点展开，仅取展开式第一项 $B_e(0)$ 代入，即得磁偶极子在外磁场所受力矩为

$$\begin{aligned} M_{力}^{(1)} &= \int x' \times [J_f(x') \times B_e(0)]\mathrm{d}V' \\ &= \int \{J_f(x')[x' \cdot B_e(0)] - [x' \cdot J_f(x')]B_e(0)\}\mathrm{d}V'. \end{aligned} \tag{4.18}$$

现证明上式右边的第二项积分为零. 利用稳恒电流的条件 $\nabla' \cdot J_f(x') = 0$，可得

$$\nabla' \cdot [x'^2 J_f(x')] = (\nabla' x'^2) \cdot J_f(x') + x'^2 \nabla' \cdot J_f(x') = 2x' \cdot J_f(x').$$

由此 $M_{力}^{(1)}$ 的第二项可化为等子零的面积分：

$$\begin{aligned} B_e(0)\int x' \cdot J_f(x')\mathrm{d}V' &= \frac{1}{2}B_e(0)\int \nabla' \cdot [x'^2 J_f(x')]\mathrm{d}V' \\ &= \frac{1}{2}B_e(0)\oint x'^2 J_f(x') \cdot \mathrm{d}\sigma' = 0. \end{aligned}$$

因此(4.18)式右边只剩下第一项. 重复(4.13)式到(4.14)式的证明可得

$$M_{力}^{(1)} = \int J_f(x')[x' \cdot B_e(0)]\mathrm{d}V' = m \times B_e(0). \tag{4.19}$$

显然，这与电偶极子 p_e 在静电场中所受的力矩 §3.1.6(6.22)式形式上相似.

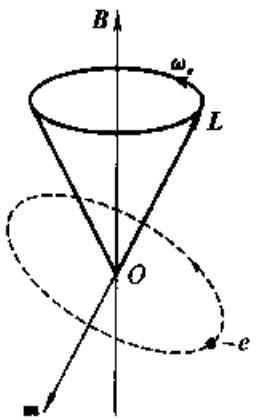


图 1 拉莫尔进动

(五)拉莫尔进动

现在讨论当原子置于均匀磁场 B 中时，磁场对原子中电子运动的影响. 首先，原子中所有的电子构成了一个具有相同荷质比 q/m 并环绕共同力心运动的电荷体系. 我们用 $\rho(x)$ 表

示这个电荷体系的质量密度,则其电流密度 $\boldsymbol{J}_f$ 为:

$$\boldsymbol{J}_f = \rho_f \boldsymbol{v} = (q/m)\rho\, \boldsymbol{v}. \tag{4.20}$$

由此可将电荷体系的磁偶极矩 § 3.2.3(3.9)式表示为:

$$\boldsymbol{m} = \frac{1}{2}\int_V (\boldsymbol{x}' \times \boldsymbol{J}_f)\,\mathrm{d}V' = \frac{q}{2m}\int_V (\boldsymbol{x}' \times \rho\boldsymbol{v})\,\mathrm{d}V' = \frac{q}{2m}\boldsymbol{L}. \tag{4.21}$$

其中 $\boldsymbol{L} = \int (\boldsymbol{x}' \times \rho\boldsymbol{v})\,\mathrm{d}V'$ 为电荷体系的角动量. 由角动量定理 § 1.5.2(2.5)式

$$\frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t} = \boldsymbol{M}_{力} = \boldsymbol{m} \times \boldsymbol{B} = \frac{q}{2m}\boldsymbol{L} \times \boldsymbol{B} = \left(-\frac{q}{2m}\boldsymbol{B}\right) \times \boldsymbol{L}. \tag{4.22}$$

将此式和 § 1.5.3(3.1)式比较可见,电荷体系的角动量矢量 $\boldsymbol{L}$ 绕磁场 $\boldsymbol{B}$ 以角速度

$$\omega_e = -\frac{q}{2m}B \tag{4.23}$$

转动(如图 1 所示). 这个现象与陀螺转动轴的进动现象相似,称为拉莫尔进动,进动的角频率 ω_e 称为拉莫尔频率. 由于电子带负电,$q = -e$,所以 $\boldsymbol{\omega}_e$ 与 $\boldsymbol{B}$ 的方向一致,而 $\boldsymbol{m}$ 与 $\boldsymbol{L}$ 方向相反.

(六)磁偶极子在外磁场中的势能与相互作用能的关系

电偶极子与外电场的相互作用能和它在外电场中的势能相等. 但磁偶极子 $\boldsymbol{m}$ 与外磁场的相互作用能 $E^{(1)}_{相互} = \boldsymbol{m} \cdot \boldsymbol{B}_e$ 和它在外磁场中的势能 $U^{(1)} = -\boldsymbol{m} \cdot \boldsymbol{B}_e$ 差一负号. 现在讨论产生这一差异的物理原因. 不失一般性,假设磁矩为 $\boldsymbol{m} = I_f \boldsymbol{a}$ 的磁偶极子是由电流为 I_f,面积元矢量为 $\boldsymbol{a}$ 的小圆电流圈 f 产生,而外磁场是由电流为 I_e 的闭合回路 e 所产生.

磁偶极子 $\boldsymbol{m}$ 在场点 $\boldsymbol{x}$ 处的势能 $U^{(1)}$ 等于将磁偶极子从磁场外移动到场点 $\boldsymbol{x}$ 处克服磁场力所作的功. 但是,当把一个作为磁偶极子的载流线圈 f 移入磁场中时,穿过线圈的磁通量将发生变化,而在线圈 f 中引起感生电动势 $\mathscr{E}'_f$. 为了维持线圈中的电流不变,必须在线圈中再接上一个附加的电源 $\mathscr{E}_f$,这个电源的电动势可以随时间变化而始终与线圈 f 上的感生电动势 $\mathscr{E}'_f$ 相抵消,即

$$\mathscr{E}_f = -\mathscr{E}'_f = \mathrm{d}\Phi_f/\mathrm{d}t. \tag{4.24}$$

这里 Φ_f 是外磁场穿过线圈 f 的磁通. 因此,在 $\mathrm{d}t$ 时间内,附加电源所作的功为

$$\mathrm{d}W_f = \mathscr{E}_f\,\mathrm{d}q = \mathscr{E}_f I_f\,\mathrm{d}t = I_f\,\mathrm{d}\Phi_f. \tag{4.25}$$

因此,把磁偶极子从磁场外移动到场点 $\boldsymbol{x}$ 处,为保持线圈 f 的电流始终不变,附加电源要作功

$$W_f = \int \mathrm{d}W_f = \int I_f\,\mathrm{d}\Phi_f = I_f\,\Phi_f = I_f(\boldsymbol{a} \cdot \boldsymbol{B}_e) = \boldsymbol{m} \cdot \boldsymbol{B}_e. \tag{4.26}$$

另一方面,当磁偶极子移入磁场时,穿过闭合回路 e 的磁通也发生变化,从而在回路 e 引起感生电动势 $\mathscr{E}'_e$. 为了保持外磁场 $\boldsymbol{B}_e$ 不变,就要保持回路 e 的电流 I_e 不变,因此在回路 e 上同样要再接一个附加电源 $\mathscr{E}'_e$,它始终抵消回路 e 上的感

生电动势. 现在计算，当磁偶极子由磁场外移至场点 $\boldsymbol{x}$ 处时，这个电源所作的功 W_e.

设线圈 f 与线圈 e 的互感为 $M(t)$，由于要保证 I_f 与 I_e 不变，所以在线圈 f 与线圈 e 中，外接的附加电源电动势应分别为

$$\mathscr{E}_f = \mathrm{d}\Phi_f/\mathrm{d}t = \mathrm{d}(MI_e)/\mathrm{d}t = I_e\,\mathrm{d}M/\mathrm{d}t, \tag{4.27}$$

$$\mathscr{E}_e = \mathrm{d}\Phi_e/\mathrm{d}t = \mathrm{d}(MI_f)/\mathrm{d}t = I_f\,\mathrm{d}M/\mathrm{d}t. \tag{4.28}$$

两附加电源所提供的功率分别为

$$\mathrm{d}W_f/\mathrm{d}t = I_f\mathscr{E}_f = I_f I_e\,\mathrm{d}M/\mathrm{d}t, \tag{4.29}$$

$$\mathrm{d}W_e/\mathrm{d}t = I_e\mathscr{E}_e = I_e I_f\,\mathrm{d}M/\mathrm{d}t, \tag{4.30}$$

比较两式可见 $W_f = W_e$，再将(4.26)式代入，即有

$$W_e = W_f = \boldsymbol{m}\cdot\boldsymbol{B}_e. \tag{4.31}$$

因此，将磁偶极子 $\boldsymbol{m}$ 从磁场外移至场点 $\boldsymbol{x}$ 处时，它与外磁场的相互作用能 $E^{(1)}_{相互}$ 应等于外力克服磁场所作的功 $U^{(1)}$ 以及外接附加电源所作的功之和，即

$$E^{(1)}_{相互} = U^{(1)} + W_f + W_e = -\boldsymbol{m}\cdot\boldsymbol{B}_e + \boldsymbol{m}\cdot\boldsymbol{B}_e + \boldsymbol{m}\cdot\boldsymbol{B}_e = \boldsymbol{m}\cdot\boldsymbol{B}_e. \tag{4.32}$$

由此可见，造成 $E^{(1)}_{相互}$ 与 $U^{(1)}$ 不相等的根源在于，在没有外接附加电源的情况下，将磁偶极子由磁场外移动到 $\boldsymbol{x}$ 点时不能保持磁偶极子 $\boldsymbol{m}$ 与外磁场 $\boldsymbol{B}_e$ 不变；如果要保持 $\boldsymbol{m}$ 与 $\boldsymbol{B}_e$ 不变，必须由外接附加电源供给 $2\boldsymbol{m}\cdot\boldsymbol{B}_e$ 的能量，从而使相互作用能 $E^{(1)}_{相互}$ 比势能 $U^{(1)}$ 大 $2\boldsymbol{m}\cdot\boldsymbol{B}_e = -2U^{(0)}$，因而 $E^{(1)}_{相互} = -U^{(0)}$.

§3.2.5 超导体的电磁性质

本节介绍超导体的四种基本电磁现象，即零电阻、临界磁场与临界电流、完全抗磁性和同位素效应. 在此基础上，介绍处理超导问题的两种方法：一种是从超导体的零电阻性出发，把超导体看作具有特殊性质的导体，把超导电流看作传导电流的一部分，由二流体模型导出超导体的电磁性质方程——伦敦方程，然后利用麦克斯韦方程和伦敦方程讨论超导问题；另一种方法是从超导体的完全抗磁性出发，把超导体看作一种特殊的磁介质，把超导电流看作磁化电流(略去通常很小的分子磁化电流)，超导电流使超导体带有磁矩 $\boldsymbol{M}$，然后用磁标势法讨论超导问题.

(一)超导体的基本电磁现象

(1)零电阻性

1911 年昂纳斯发现水银的电阻率在 $T = 4.2\mathrm{K}$ 附近突然下降到零，这种电阻率为零的性质称为超导电性. 刚出现超导电性的温度称为临界温度，并记作 T_C. 在 T_C 以上，超导材料处于正常态；在 T_C 以下，超导材料处于超导态. 不同的物质材料，有不同的临界温度.

(2)临界磁场与临界电流

1914 年,昂纳斯又发现,将超导体置于磁场内,当磁场增大到某一临界值 H_C 时,超导体从超导态转变为正常态. 临界磁场 H_C 与温度 T 的经验关系式为

$$H_C(T) = H_C(0)[1 - (T/T_C)^2]. \tag{5.1}$$

$H_C(T)$ 的函数曲线如图 1 所示,曲线以下的区域为超导态,曲线以上的区域为正常态. 超导态与正常态是两种不同的相,它们之间的转变是二级相变的典型例子,在 §7.2.9 中将仔细讨论.

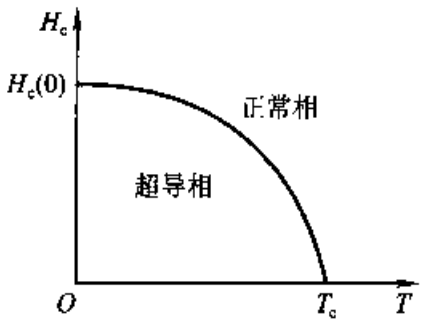


图 1 超导体的临界磁场 $H_c(T)$

当超导体内部的电流 $I > I_C$,因而电流 I 所产生的磁场 $H > H_c(T)$ 时,超导材料的超导性也会被破坏. 可以证明,$I_c(T)$ 也有与 $H_C(T)$ 类似的函数关系

$$I_c(T) = I_c(0)[1 - (T/T_C)^2]. \tag{5.2}$$

(3)完全抗磁性——迈斯纳效应

1933 年迈斯纳发现,只要 $T \leqslant T_C$,$H \leqslant H_C$,就有超导体内部的磁感应强度 $\boldsymbol{B} = 0$,和降温与加磁场这两个过程的先后次序无关. 也就是说,若超导材料已处于超导态,当加入外磁场时,只要磁场强度 $H \leqslant H_C$,则 B 不能进入超导体内部;若将处于正常态的超导材料置于 $H \leqslant H_C$ 的磁场内,当温度下降使之转变为超导态时,B 将从超导体内部被排出.

由零电阻性可知,超导体内部的电场 $\boldsymbol{E} = \boldsymbol{J}_f/\sigma \to 0$;又由麦克斯韦方程 $\partial \boldsymbol{B}/\partial t = -\nabla \times \boldsymbol{E} \to 0$ 可知,超导体内的 $\boldsymbol{B}$ 应为与时间无关的量. 但不能肯定 $\boldsymbol{B} = 0$,因此,完全抗磁性不可能由零电阻性导出,它是超导体的另一重要性质.

(4)同位素效应

1950 年,美国物理学家麦克斯韦和雷诺兹同时分别发现,超导体的临界温度 T_C 反比于构成该超导体的同位素质量的平方根(不同超导材料略有不同)

$$T_C \propto 1/\sqrt{M}. \tag{5.3}$$

由于同位素的结构完全相同,若超导体的性质决定于电子间的相互作用,则其 T_c 应相同;而同位素的差异是原子核的质量不同,故同位素效应表明,电子与晶格的相互作用是超导相变的关键所在.

(二)伦敦方程

为了解释超导体的电磁性质,戈尔特和卡西米尔子 1933 年提出了超导性的唯象理论——二流体模型. 即把超导体看作 $\mu \approx \mu_0$,$\varepsilon \approx \varepsilon_0$ 的普通导体的同时,假设超导体存在两种自由电子:正常电子和超导电子,它们具有不同的性质:

(1)正常电子与晶格碰撞而被散射,改变其能量和动量,产生电阻. 正常电子

的流动产生正常电流 J_n，它遵守欧姆定律

$$J_n = \sigma E. \tag{5.4}$$

(2)超导电子不会被晶格散射，运动不受阻尼，超导电子的流动产生超导电流 J_S.

这两种电子就像两种流体一样在超导体中流动. 在正常态中，超导体只有正常电子，它的行为就和正常导体一样，存在电阻；当温度降低到临界温度以下时，超导体进入超导态，这时超导体中就出现了超导电子，它们可以不受任何阻碍地在超导体中流动. 温度越低，超导电子的比例越大；当温度无限接近 0K 时，超导体中就只有超导电子存在了. 二流体模型解释了超导体的许多性质，但没有指出超导电子的起源. 1957 年巴丁、库珀(Cooper)和施列菲(Schrieffer)提出了超导的微观理论(简称 BCS 理论)，成功地解释了当时已知的各种超导现象. BCS 理论的基本物理图像是：所谓超导电子实际上是束缚在一起的一对电子，这种束缚来自电子与晶格的相互作用. 当一个电子在晶格中运动时，由于它与组成晶格的正离子的库仑吸引作用，会对原子整齐排列的晶格点阵带来一个扰动，使点阵局部发生畸变. 这种扰动会像波动一样传播而吸引另一个电子，于是这两个电子就会通过点阵的振动而耦合起来. 如果电子与晶格的作用足够强，使电子间通过晶格的间接吸引作用大于电子间的库仑排斥作用，则在两电子间产生一种净吸引力，导致两电子形成一个束缚态，就像两个氢原子结合成氢分子一样，结合成一个电子对，称为“库珀电子对”，简称库珀对. 库珀对的空间宽度为 10^{-4}cm，比氢分子中两氢原子的距离 10^{-8}cm 要大得多. 库珀还证明：当超导体没有电流时，库珀对的总动量为零(即两电子的动量大小相等，方向相反，并且两电子的自旋也相反)；而当有电流存在时，所有的库珀对都具有相同的总动量 p[在以 $v=p/(2m)$ 运动的观察者看来，与前述总动量为零的情况是一样的]. 数以亿万计的这种总动量相同的库相对形成有序的运动，产生无阻尼的超导电流. 计算表明，每个库珀对有结合能 Δ(数量级为 10^{-4}eV)，称为能隙. 当温度满足 $kT > \Delta$ 时，热运动将库珀对拆散，晶格(严格说来是晶格的振动)与电子的相互作用使导体产生电阻；当温度满足 $kT < \Delta$ 时，热运动不会破坏库珀对，晶格与电子的相互作用体现为使电子处于有序状态，形成超导电流. $T_c \approx \Delta/k$ 是超导体的临界温度.

现在由二流体模型导出超导体的电磁性质方程——伦敦方程.

(1)伦敦第一方程

在有电阻的正常导体中，电场的作用是维持电流；在无电阻的超导体中，电场使超导电子作加速运动. 设库珀对(含两个超导电子，其质量为 $2m$，电量为 $-2e$)的速度为 v，则由牛顿定律(约去 2)可得

$$m\dot{v} = -eE. \tag{5.5}$$

若单位体积内有 n_s 个超导电子,则超导电流密度

$$\boldsymbol{J}_S = -n_s e\boldsymbol{v}. \tag{5.6}$$

将(5.6)式对 t 求导后代入(5.5)式,并令 $\alpha = n_s e^2/m$,即有

$$\frac{\partial \boldsymbol{J}_S}{\partial t} = -n_s e\dot{\boldsymbol{v}} = \frac{n_s e^2}{m}\boldsymbol{E} = \alpha \boldsymbol{E}, \tag{5.7}$$

这就是伦敦第一方程,它给出超导电流和场的关系.

和正常导体中电流密度正比于场强[参看(5.4)式]不一样,超导体中是超导电流的时间导数正比于场强,如(5.7)式. 其原因是,超导电子的运动不受阻力. 如果在(5.5)式中加上与 $\boldsymbol{v}$ 成正比的阻力项,达到稳定时($\dot{\boldsymbol{v}}=0$)就有 $\boldsymbol{v}$ 与 $\boldsymbol{E}$ 成正比. 由(5.6)式就有电流密度 $\boldsymbol{J}$ 与 $\boldsymbol{E}$ 成正比. 这就是正常导体的情况.

利用伦敦第一方程可得两个重要结论. 一是当超导电流恒定时($\partial \boldsymbol{J}_S/\partial t = 0$),由(5.7)式得 $\boldsymbol{E}=0$,代入(5.4)式即有 $\boldsymbol{J}_n = 0$,故在电流恒定的情况下,导体内没有正常电流,只有无损耗的超导电流. 二是当超导电流交变时,例如 $J_S = J_S(0)e^{i\omega t}$,由(5.7)式可见 $E = \dot{J}_S/\alpha \neq 0$,因而 $\boldsymbol{J}_n = \sigma \boldsymbol{E} \neq 0$,即在交变情况下超导体是有电阻损耗的. 设电流频率为 ω,其交流损耗的百分比可估算如下:

$$\frac{J_n}{J_S} = \frac{\sigma}{\alpha/\omega} = \frac{m\sigma\omega}{n_s e^2} \approx 10^{-12}\omega,$$

由此可见,对于低频交流电,其交流损耗甚微.

(2)伦敦第二方程

将伦敦第一方程 $\boldsymbol{J}_S = \alpha \boldsymbol{E}$ 代入麦克斯韦方程,可得 $\nabla \times \dot{\boldsymbol{J}}_S \alpha \nabla \times \boldsymbol{E} = -\alpha \dot{\boldsymbol{B}}$. 交换上式左边对时间和空间导数的次序,移项后得

$$\frac{\partial}{\partial t}(\nabla \times \boldsymbol{J}_S + \alpha \boldsymbol{B}) = 0,$$

由此得

$$\nabla \times \boldsymbol{J}_S + \alpha \boldsymbol{B} = \text{常矢量} \quad (\text{与 } t \text{ 无关}).$$

设常矢量为零(其正确性由实验判断),即有

$$\nabla \times \boldsymbol{J}_S = -\alpha \boldsymbol{B} = -(n_s e^2/m)\boldsymbol{B}. \tag{5.8}$$

这就是伦敦第二方程,它给出超导电流与磁场的关系. (5.8)式说明超导电流是靠磁场来维持的.

由麦克斯韦方程及伦敦方程可以解释迈斯纳效应. 将恒定电流时 $\boldsymbol{J}_t = \boldsymbol{J}_S$ 代入麦克斯韦方程 $\nabla \times \boldsymbol{B} = \mu \boldsymbol{J}_t$ 可得

$$\nabla \times \boldsymbol{B} = \mu \boldsymbol{J}_S. \tag{5.9}$$

将上式两边取旋度 $\nabla \times (\nabla \times \boldsymbol{B}) = \mu \nabla \times \boldsymbol{J}_S$,并利用伦敦第二方程及 $\nabla \cdot \boldsymbol{B} = 0$,便

有 $-\nabla^2\boldsymbol{B}=-\mu\alpha\boldsymbol{B}$,

$$\nabla^2\boldsymbol{B}=(1/\lambda_L^2)\boldsymbol{B}, \tag{5.10}$$

其中 $\lambda_L=1/\sqrt{\mu\alpha}$.

设超导体占满 $z>0$ 的空间以及 $\boldsymbol{B}=B(z)\boldsymbol{e}_x$,则(1.10)式可简化为

$$\frac{\mathrm{d}^2B(z)}{\mathrm{d}z^2}=\frac{1}{\lambda_L^2}B(z),$$

其解为

$$\boldsymbol{B}=B(z)\boldsymbol{e}_x=B(0)\mathrm{e}^{-z/\lambda_L}\boldsymbol{e}_x, \tag{5.11}$$

即磁场随 z 的增加而指数地衰减,故磁场只能透入超导体的表面薄层. 通常称磁场 $B(z)$ 减小到 $B(0)$ 的 e^{-1} 时的距离为穿透深度. 由(5.11)式可见,穿透深度为

$$\lambda_L=1/\sqrt{\mu\alpha}=\sqrt{m/(\mu n_S e^2)}. \tag{5.12}$$

对一般超导体 $\lambda_L\sim10^{-7}\mathrm{m}$. 这说明在超导体内,磁场只在极薄的表层中存在. 在超导体的内部没有磁场,这就是迈斯纳效应.

同样可以证明,超导电流也只存在于超导体表面数量级为 λ_L 的薄层中. 在上例中,磁场只有 x 分量,将(5.11)式代入(5.9)式,即有

$$\boldsymbol{J}_S=\frac{1}{\mu}\nabla\times\boldsymbol{B}=\frac{1}{\mu}\frac{\partial B_x}{\partial z}\boldsymbol{e}_y=-\frac{B(0)}{\mu\lambda_L}\mathrm{e}^{-z/\lambda_L}\boldsymbol{e}_y. \tag{5.13}$$

超导电流所产生的磁场在超导体内部与外场反向,结果将外场屏蔽掉,导致超导体内部 $\boldsymbol{B}=0$.

(三)磁标势法

现在把超导体看成特殊的磁介质,把超导电流看作磁化电流,它使超导体带有磁矩 $\boldsymbol{M}$. 由于磁化电流与磁矩有下述关系:

$$\int_S\boldsymbol{J}_M\cdot\mathrm{d}\boldsymbol{\sigma}=\oint_l n\,i\boldsymbol{a}\cdot\mathrm{d}\boldsymbol{l}=\oint nm\cdot\mathrm{d}\boldsymbol{l}=\oint_L\boldsymbol{M}\cdot\mathrm{d}\boldsymbol{l}=\oint_S\nabla\times\boldsymbol{M}\cdot\mathrm{d}\boldsymbol{\sigma},$$

由 S 的任意性,即有

$$\nabla\times\boldsymbol{M}=\boldsymbol{J}_M.$$

现在把超导电流看作磁化电流,应将 $\boldsymbol{J}_M$ 改写为 $\boldsymbol{J}_S$,因此,

$$\nabla\times\boldsymbol{M}=\boldsymbol{J}_S. \tag{5.14}$$

设介质 1 为超导体,介质 2 为空气,用 $\boldsymbol{\alpha}_S$ 表示超导电流面密度,仿照 §2.4.2 (2.8)式及(2.12)式的推导,可得

$$\boldsymbol{e}_n\times\boldsymbol{M}=-\boldsymbol{\alpha}_S. \tag{5.15}$$

由于超导体内部的磁感应强度 $\boldsymbol{B}=0$,便有 $\mu_0(\boldsymbol{H}+\boldsymbol{M})=\boldsymbol{B}=0$,

$$\boldsymbol{M}=-\boldsymbol{H}. \tag{5.16}$$

又由 §2.4.1 的(1.20)式 $\boldsymbol{M}=\chi_m\boldsymbol{H}$ 可得超导体的磁化率

$$\chi_m = -1, \tag{5.17}$$

因而超导体的磁导率为零

$$\mu = \mu_0(1+\chi_m) = 0, \tag{5.18}$$

即超导体是完全抗磁体.

现在用磁标势法计算一个具体问题.

【例 1】 半径为 R_0 的超导球置于均匀外场 $\boldsymbol{H}_0$ 中,已知 $H_0 < H_C$, $T < T_C$,求超导球内外的磁场及超导面电流分布.

【解】 在此情况下没有传导电流,故在超导体内外均可用磁标势描述磁场.

(a)坐标系.取坐标原点在球心,z 轴沿 $\boldsymbol{H}_0$ 方向的球坐标系.

(b)方程.因真空及均匀介质中的束缚磁荷密度为零,故磁标势满足拉普拉斯方程

$$\nabla^2\varphi_1 = 0, \qquad \nabla^2\varphi_2 = 0. \tag{5.19}$$

(c)对称性及通解形式.本题有轴对称性,故

球内:　$\varphi_1 = \sum a_l R^l P_l(\cos\theta)$, (5.20)

球外:　$\varphi_2 = \sum [c_l R^l + d_l R^{-(l+1)}] P_l(\cos\theta)$, (5.21)

在(1.20)式中已考虑到 $\varphi_1|_{R=0}$ 有限的要求.

(d)边界条件与边值关系.

①均匀外场 H_0 在无限远处激发的磁标势为

$$\varphi_2|_{R=\infty} = -\boldsymbol{H}_0 \cdot \boldsymbol{x} = -H_0 R\cos\theta, \tag{5.22}$$

②球面上磁标势连续

$$\varphi_1|_{R=R_0} = \varphi_2|_{R=R_0}, \tag{5.23}$$

③球面上 $B_{2n} = B_{1n} = 0$,由此得

$$\partial\varphi_2/\partial R|_{R=R_0} = 0. \tag{5.24}$$

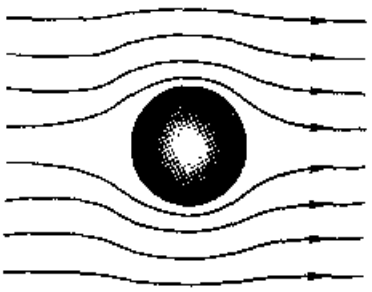


图 2　超导球内外的磁场

(e)定系数得磁标势.

$$\varphi_1 = -(2/3)H_0 R\cos\theta, \tag{5.25}$$

$$\varphi_2 = -H_0 R\cos\theta - \frac{H_0 R_0^3}{2R^2}\cos\theta = -H_0 R\cos\theta - \frac{R_0^3 \boldsymbol{H}_0 \cdot \boldsymbol{R}}{2R^3}. \tag{5.26}$$

(f)求磁场强度.利用 §3.2.2(2.1)式和本节(5.25)式求磁标势的负梯度即得球内的磁场强度

$$\boldsymbol{H}_1 = -\nabla\varphi_1 = (3/2)\nabla(H_0 R\cos\theta) = (3/2)\boldsymbol{H}_0. \tag{5.27}$$

由超导体的完全抗磁性可得 $0 = \boldsymbol{B}_1 = \mu_0(\boldsymbol{M} + \boldsymbol{H}_1)$,故磁化强度

$$\boldsymbol{M} = -\boldsymbol{H}_1 = -(3/2)\boldsymbol{H}_0. \tag{5.28}$$

球外的磁场强度 $\boldsymbol{H}_2 = -\nabla\varphi_2$.由(5.26)式可见,第一项的负梯度给出外磁场

$\boldsymbol{H}_0$;考虑到磁矩为 $\boldsymbol{m}$ 的磁偶极子激发的磁势为 $\varphi = \boldsymbol{m} \cdot \boldsymbol{R}/(4\pi R^3)$,与 φ_2 的第二项比较可以得到 $\boldsymbol{m} = -2\pi R_0^3 \boldsymbol{H}_0 = -(4\pi R_0^3/3)(3/2)\boldsymbol{H}_0 = \boldsymbol{MV}$. 故球外磁场是外磁场 $\boldsymbol{H}_0$ 加上磁矩 $\boldsymbol{m} = \boldsymbol{MV}$ 激发的场.

(g)球面上的超导电流密度.

$$\boldsymbol{\alpha}_2 = -\boldsymbol{e}_n \times \boldsymbol{M} = \boldsymbol{e}_n \times \boldsymbol{H}_1 = (3/2)\boldsymbol{e}_R \times \boldsymbol{H}_0 = -(3/2)H_0 \sin\theta \boldsymbol{e}_\phi.$$

磁场线的分布如图 2 所示.

习 题

1. 设 $x<0$ 半空间充满磁导率为 μ 的均匀介质,$x>0$ 半空间为真空,今有线电流 I_f 沿 z 轴流动,求磁感应强度与磁化电流分布.

2. 均匀无限长直圆柱形螺线管,每单位长度线圈匝数为 n,电流强度为 I. 试用唯一性定理求管内、外磁感应强度 $\boldsymbol{B}$.

3. 有一均匀带电的导体球壳,其半径为 R_0,总电荷为 Q_f,今使球壳绕自身某一直径以角速度 ω 旋转,试分别按矢势法和磁标势法求球内、外的磁场 $\boldsymbol{B}$.

4. 证明 $\mu \to \infty$ 的磁性物质表面为等磁势面.

5. 有一内外半径分别为 R_1 和 R_2 的空心球,位于均匀外磁场 H_0 内,球的磁导率为 μ,求空腔内的磁场 $\boldsymbol{B}$. 讨论 $\mu \gg \mu_0$ 时的磁屏蔽作用.

6. 设理想铁磁体的磁化规律为 $\boldsymbol{B} = \mu \boldsymbol{H} + \mu_0 \boldsymbol{M}_0$,$\boldsymbol{M}_0$ 是恒定的与 $\boldsymbol{H}$ 无关的量. 今将一个理想铁磁体做成的均匀磁化球($\boldsymbol{M}_0$ 为常量)浸入磁导率为 μ' 的无限介质中,求磁感应强度与磁化电流分布.

7. 电荷按体均匀分布的刚性小球,其总电荷为 Q,半径为 R_0. 它以角速度 ω 绕自身某一直径转动,求(a)它的磁矩,(b)它的磁矩与自转角动量之比(设质量 M_0 是均匀分布的).

8. 有一块磁矩为 $\boldsymbol{m}$ 的小永磁体,位于一块磁导率非常大的实物的平坦界面附近的真空中,求作用在小磁体上的力 $\boldsymbol{F}$.

第三章　电磁波的传播

前两章讨论的是稳恒场，本章讨论随时间变化的场. 变化的电磁场通过自身的互相激发而运动，形成电磁波. 本章先讨论无界空间中平面电磁波的主要性质. 然后利用电磁波的边值关系讨论电磁波在两种介质界面上的反射和折射现象，通过导出的反射定律和折射定律给出反射波和折射波的传播方向，而导出的菲涅尔公式则给出反射波和折射波的波幅. 接着讨论电磁波在导体中传播的规律，解释电磁波在导体中的趋肤效应等现象. 最后以谐振腔和波导管为例讨论电磁波在有界空间中的传播规律，它在微波技术和激光技术上有重要应用.

§3.3.1　平面电磁波

(一)自由空间电磁场的波动方程

所谓自由空间是指没有自由电荷($\rho_f=0$)和传导电流($\boldsymbol{J}_f=0$)的区域. 因为麦克斯韦方程组是电磁场的普遍规律，因而自由空间电磁场的波动方程可将 $\rho_f=0$，$\boldsymbol{J}_f=0$ 代入麦克斯韦方程求得. 这时，

$$\begin{cases} \nabla\times\boldsymbol{E}=-\partial\boldsymbol{B}/\partial t, & (1.1)\\ \nabla\times\boldsymbol{H}=\partial\boldsymbol{D}/\partial t, & (1.2)\\ \nabla\cdot\boldsymbol{D}=0, & (1.3)\\ \nabla\cdot\boldsymbol{B}=0. & (1.4) \end{cases}$$

在真空情况下，$\boldsymbol{D}=\varepsilon_0\boldsymbol{E}$，$\boldsymbol{B}=\mu_0\boldsymbol{H}$，利用它求(1.1)式的旋度，并利用(1.2)式即有

$$\nabla\times(\nabla\times\boldsymbol{E})=\nabla\times\left(-\frac{\partial\boldsymbol{B}}{\partial t}\right)=-\frac{\partial}{\partial t}(\mu_0\nabla\times\boldsymbol{H})=-\mu_0\varepsilon_0\frac{\partial^2\boldsymbol{E}}{\partial t^2}. \quad (1.5)$$

由∇算符公式及$\nabla\cdot\boldsymbol{E}=\nabla\cdot\boldsymbol{D}/\varepsilon_0=0$，即有

$$\nabla\times(\nabla\times\boldsymbol{E})=\nabla(\nabla\cdot\boldsymbol{E})-\nabla^2\boldsymbol{E}=-\nabla^2\boldsymbol{E}. \quad (1.6)$$

将(1.5)式与(1.6)式联立，并令$\sqrt{\mu_0\varepsilon_0}=1/c$，便有

$$\nabla^2\boldsymbol{E}-\frac{1}{c^2}\frac{\partial^2\boldsymbol{E}}{\partial t^2}=0. \quad (1.7)$$

同理

$$\nabla^2\boldsymbol{B}-\frac{1}{c^2}\frac{\partial^2\boldsymbol{B}}{\partial t^2}=0. \quad (1.8)$$

这就是真空中电磁场的波动方程.

(二)平面电磁波

电磁场波动方程仅依赖于一个坐标(例如 x)的解就是沿这个坐标轴正向或反向传播的平面波. 设 f 是 $\boldsymbol{E}$ 或 $\boldsymbol{B}$ 的任一直角分量,由(1.7)及(1.8)式可知 f 遵从运动方程

$$\left(\frac{\partial^2}{\partial x^2}-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)f=\left(\frac{\partial}{\partial x}-\frac{1}{c}\frac{\partial}{\partial t}\right)\left(\frac{\partial}{\partial x}+\frac{1}{c}\frac{\partial}{\partial t}\right)f=0. \tag{1.9}$$

令

$$\xi=x-ct,\qquad \eta=x+ct, \tag{1.10}$$

即

$$x=(\xi+\eta)/2,\quad t=(\eta-\xi)/(2c). \tag{1.11}$$

得到:

$$\begin{cases}\dfrac{\partial}{\partial\xi}=\dfrac{\partial x}{\partial\xi}\dfrac{\partial}{\partial x}+\dfrac{\partial t}{\partial\xi}\dfrac{\partial}{\partial t}=\dfrac{1}{2}\left(\dfrac{\partial}{\partial x}-\dfrac{1}{c}\dfrac{\partial}{\partial t}\right), & (1.12)\\[2ex] \dfrac{\partial}{\partial\eta}=\dfrac{\partial x}{\partial\eta}\dfrac{\partial}{\partial x}+\dfrac{\partial t}{\partial\eta}\dfrac{\partial}{\partial t}=\dfrac{1}{2}\left(\dfrac{\partial}{\partial x}+\dfrac{1}{c}\dfrac{\partial}{\partial t}\right). & (1.13)\end{cases}$$

将(1.12)、(1.13)式代入(1.9)式,就有

$$\partial^2 f/\partial\eta\partial\xi=0, \tag{1.14}$$

对 ξ 积分,得:

$$\partial f/\partial\eta=F(\eta). \tag{1.15}$$

这里的"常数项"是与积分变量 ξ 无关的"常数",但它可以是 η 的函数. 再对 η 积分,也要加上与积分变量 η 无关的"常数项" $f_2(\xi)$,即

$$f=\int F(\eta)\mathrm{d}\eta+f_2(\xi)=f_1(\eta)+f_2(\xi). \tag{1.16}$$

将(1.10)式代入(1.16)式,就有

$$f=f_1(x-ct)+f_2(x+ct). \tag{1.17}$$

我们推导(1.17)式的出发点是场量只与一个坐标 x 有关. 现有的解正如所期:在任何给定的时刻 t,在 $x=$常数的平面上各点,$\boldsymbol{E}$ 与 $\boldsymbol{B}$ 的各直角分量相等. 这就是"平面电磁波"这个名称的由来. 这个场值相同的面称为等相面. 等相面的传播速度由 $x\pm ct=$常数求得为

$$\mathrm{d}x/\mathrm{d}t=\pm c. \tag{1.18}$$

这表示电磁波在真空中以速度 $c=1/\sqrt{\varepsilon_0\mu_0}$ 沿 x 轴的正向或反向传播.

将真空中的"介电常数" ε_0 和"磁导率" μ_0 代入,得到 c 的值正好等于光在真空中的传播速度. 这就给光是电磁波提供了一个有力的证据.

(三)单色电磁波的运动方程——亥姆霍兹方程

当角频率为 ω 的电磁波在介质中传播时,介质内的束缚电荷在电磁场的作用下以角频率 ω 作强迫振动.通过求解束缚电子的运动方程(见 §4.2.5),可以证明介质的介电常数与电磁波的频率有关:$\varepsilon=\varepsilon(\omega)$.类似地,可以证明 $\mu=\mu(\omega)$,因此,线性介质的电磁性质方程为

$$\boldsymbol{D}_\omega(t)=\varepsilon(\omega)\boldsymbol{E}_\omega(t),\quad \boldsymbol{B}_\omega(t)=\mu(\omega)\boldsymbol{H}_\omega(t). \tag{1.19}$$

这种 ε,μ 随电磁波的频率而变的现象称为电磁波在介质中的色散.这样,当一个非单色的电磁波射入介质中时,为了找到它的传播规律,必须首先进行傅里叶分解,然后求得给定频率的单色电磁波的运动方程及其解,最后再进行叠加.

现在推导单色电磁波的运动方程.为了方便起见,采用复数形式表示具有确定频率的电磁场:

$$\begin{cases}\boldsymbol{E}(\boldsymbol{x},t)=\boldsymbol{E}(\boldsymbol{x})\cos\omega t=\mathrm{Re}[\boldsymbol{E}(\boldsymbol{x})e^{-i\omega t}]\to\boldsymbol{E}(\boldsymbol{x})e^{-i\omega t},\\ \boldsymbol{B}(\boldsymbol{x},t)=\boldsymbol{B}(\boldsymbol{x})\cos\omega t=\mathrm{Re}[\boldsymbol{B}(\boldsymbol{x})e^{-i\omega t}]\to\boldsymbol{B}(\boldsymbol{x})e^{-i\omega t}.\end{cases} \tag{1.20}$$

这里已略去 $\boldsymbol{E}_\omega$ 和 $\boldsymbol{B}_\omega$ 的脚标.由于方程线性,可以直接用指数进行运算,计算完毕后再取实部.现在,将 $\boldsymbol{D}=\varepsilon\boldsymbol{E}$ 及 $\boldsymbol{B}=\mu\boldsymbol{H}$ 代入麦克韦方程组,就有

$$\begin{cases}\nabla\times\boldsymbol{E}=-\partial\boldsymbol{B}/\partial t=i\omega\mu\boldsymbol{H}, & (1.21)\\ \nabla\times\boldsymbol{H}=\partial\boldsymbol{D}/\partial t=-i\omega\varepsilon\boldsymbol{E}, & (1.22)\\ \nabla\cdot\boldsymbol{E}=0, & (1.23)\\ \nabla\cdot\boldsymbol{H}=0. & (1.24)\end{cases}$$

对(1.21)式取旋度并将(1.22)式代入,得到

$$\nabla\times(\nabla\times\boldsymbol{E})=\omega^2\mu\varepsilon\boldsymbol{E}. \tag{1.25}$$

展开(1.25)式左边,并将(1.23)式代入,即可求得单色电磁波的运动方程——亥姆霍兹方程:

$$\nabla^2\boldsymbol{E}+k^2\boldsymbol{E}=0,\qquad \text{其中 } k=\omega\sqrt{\mu\varepsilon}. \tag{1.26}$$

值得注意的是,由一阶方程(1.21),(1.22),(1.23)式推出二阶方程(1.26)式时,解比原来增加了.也就是说满足方程(1.26)的电磁场不一定满足(1.23)式,因此必须将(1.26)式与(1.23)式联立.

将求得的 $\boldsymbol{E}$ 代入(1.21)式,即可得 $\boldsymbol{B}$

$$\boldsymbol{B}=\mu\boldsymbol{H}=-i\omega^{-1}\nabla\times\boldsymbol{E}. \tag{1.27}$$

当然也可以先求 $\boldsymbol{B}$,后求 $\boldsymbol{E}$,相应的方程为

$$\nabla^2\boldsymbol{B}+k^2\boldsymbol{B}=0,\quad \nabla\cdot\boldsymbol{B}=0,\quad \boldsymbol{E}=i(\omega\mu\varepsilon)^{-1}\nabla\times\boldsymbol{B}. \tag{1.28}$$

(四)平面电磁波的性质

亥姆霍兹方程最简单的解是平面波解

$$\boldsymbol{E}(\boldsymbol{x},t)=\boldsymbol{E}_0e^{i(k\cdot x-\omega t)}. \tag{1.29}$$

① 将(1.29)式分别代入(1.23)式和(1.27)式可得：

$$\nabla \cdot \boldsymbol{E} = \boldsymbol{E}_0 \cdot \nabla e^{i(k\cdot x-\omega t)} = i\boldsymbol{k}\cdot\boldsymbol{E} = 0, \tag{1.30}$$

$$\boldsymbol{B} = -\frac{i}{\omega}\nabla\times\boldsymbol{E} = -\frac{i}{\omega}[\nabla e^{i(k\cdot x-\omega t)}]\times\boldsymbol{E}_0 = \frac{1}{\omega}\boldsymbol{k}\times\boldsymbol{E}. \tag{1.31}$$

求 $\boldsymbol{k}$ 与 $\boldsymbol{B}$ 的点积，即有

$$\boldsymbol{k}\cdot\boldsymbol{B} = \omega^{-1}\boldsymbol{k}\cdot(\boldsymbol{k}\times\boldsymbol{E}) = \omega^{-1}(\boldsymbol{k}\times\boldsymbol{k})\cdot\boldsymbol{E} = 0. \tag{1.32}$$

(1.30)和(1.32)式说明电磁波是横波. 以上三式说明 $\boldsymbol{E},\boldsymbol{B},\boldsymbol{k}$ 组成右手螺旋系.

② 将(1.29)式代入(1.31)式，并引入波传播方向的单位矢量 $\boldsymbol{e}_n = \boldsymbol{k}/k$，则有：

$$\boldsymbol{B} = \sqrt{\mu\varepsilon}\,\boldsymbol{e}_n\times\boldsymbol{E}_0 e^{i(k\cdot x-\omega t)}. \tag{1.33}$$

由此可见，$\boldsymbol{B}(t)$ 与 $\boldsymbol{E}(t)$ 同位相，且其振幅之比为

$$|\boldsymbol{E}|/|\boldsymbol{B}| = 1/\sqrt{\mu\varepsilon} = v. \tag{1.34}$$

v 是电磁波在介质中的相速度.

③ 由(1.34)式可求得平面电磁波的能量密度为[§2.4.3(3.11)式]

$$w = \frac{1}{2}(\boldsymbol{E}\cdot\boldsymbol{D}+\boldsymbol{H}\cdot\boldsymbol{B}) = \frac{1}{2}(\varepsilon E^2 + B^2/\mu) = \varepsilon E^2 = B^2/\mu. \tag{1.35}$$

④ 电磁波的能流密度为[§2.4.3(3.10)式]

$$\boldsymbol{S} = \boldsymbol{E}\times\boldsymbol{H} = \sqrt{\varepsilon/\mu}\boldsymbol{E}\times(\boldsymbol{e}_n\times\boldsymbol{E}) = \sqrt{\varepsilon/\mu}E^2\boldsymbol{e}_n = \sqrt{1/\varepsilon\mu}\varepsilon E^2\boldsymbol{e}_n = vw\boldsymbol{e}_n. \tag{1.36}$$

这表明，单色平面电磁波的能流密度 $\boldsymbol{S}$ 与波的传播方向 $\boldsymbol{e}_n$ 一致，能量传播速度与波在介质中的相速相等.

⑤ 电磁波具有动量，其电磁场动量流密度张量为[参看 §2.4.3(3.21)式]：

$$\overleftrightarrow{T} = -\varepsilon\boldsymbol{E}\boldsymbol{E} - \frac{1}{\mu}\boldsymbol{B}\boldsymbol{B} + \overleftrightarrow{e}\frac{1}{2}(\varepsilon E^2 + \frac{1}{\mu}B^2).$$

相应地，电磁场的应力张量为 $-\overleftrightarrow{T}$.

设 $\boldsymbol{E} = E\boldsymbol{e}_x$，$\boldsymbol{B} = B\boldsymbol{e}_y$，$\boldsymbol{e}_n = e_z$，将(1.35)式代入即有：

$$\overleftrightarrow{T} = w\{-\boldsymbol{e}_x\boldsymbol{e}_x - \boldsymbol{e}_y\boldsymbol{e}_y + (\boldsymbol{e}_x\boldsymbol{e}_x + \boldsymbol{e}_y\boldsymbol{e}_y + \boldsymbol{e}_n\boldsymbol{e}_n)\} = w\boldsymbol{e}_n\boldsymbol{e}_n. \tag{1.37}$$

这说明，电磁波通过任意面的动量 $\overleftrightarrow{T}\cdot d\boldsymbol{\sigma}$ 是沿着波的传播方向，而动量流密度的大小等于能量密度.

既然电磁波具有动量，当电磁波投射到物体表面上时将对物体施加一定的压力，称为辐射压力. 一个有趣的天文现象是，当太阳或恒星经过巴纳德星云时，之所以不会把构成星云的小尘粒吸引过去，就是由于辐射压力的作用. 辐射压力在宇宙现象与原子现象中往往值得重视. 辐射压力的定量计算留在下一节进行.

§3.3.2　电磁波在介质分界面的反射与折射

本节从电磁场的边值关系出发讨论电磁波射到介质的分界面上时，反射波与折射波的传播方向及振幅大小，最后介绍全反射现象和光压.

(一)反射与折射定律

绝缘介质分界面($z=0$)上的电磁场边值关系为

$$\begin{cases} \boldsymbol{e}_z \times (\boldsymbol{E}_2 - \boldsymbol{E}_1) = 0, & (2.1) \\ \boldsymbol{e}_z \times (\boldsymbol{H}_2 - \boldsymbol{H}_1) = \boldsymbol{\alpha}_f = 0, & (2.2) \\ \boldsymbol{e}_z \cdot (\boldsymbol{D}_2 - \boldsymbol{D}_1) = \sigma_f = 0, & (2.3) \\ \boldsymbol{e}_z \cdot (\boldsymbol{B}_2 - \boldsymbol{B}_1) = 0. & (2.4) \end{cases}$$

这四个方程只有两个是独立的，故利用前两个来讨论即可.

按假设，分界面两边的介质都是均匀的，因而介质内电磁场满足波动方程. 设入射波、反射波和折射波分别为

$$\begin{cases} \boldsymbol{E} = \boldsymbol{E}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}, & (2.5) \\ \boldsymbol{E}' = \boldsymbol{E}'_0 e^{i(\boldsymbol{k}'\cdot\boldsymbol{r}-\omega' t)}, & (2.6) \\ \boldsymbol{E}'' = \boldsymbol{E}''_0 e^{i(\boldsymbol{k}''\cdot\boldsymbol{r}-\omega'' t)}. & (2.7) \end{cases}$$

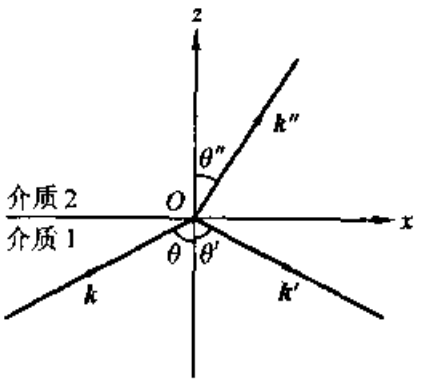


图1　电磁波在介质分界面上的反射和折射

这里假设当入射波为平面波时，反射波和折射波也均为平面波. 这个假设是否正确由它们能否满足电磁场边值关系来检验.

在界面 $z=0$ 处，因介质1的一侧有 $\boldsymbol{E}$ 及 $\boldsymbol{E}'$，介质2的一侧有 $\boldsymbol{E}''$(见图1)，代入(2.1)式可得

$$\boldsymbol{e}_z \times \boldsymbol{E} + \boldsymbol{e}_z \times \boldsymbol{E}' = \boldsymbol{e}_z \times \boldsymbol{E}'', \tag{2.8}$$

设 $\boldsymbol{e}_t$ 与 $\boldsymbol{e}_b$ 是在 $z=0$ 面上与 $\boldsymbol{e}_z$ 构成右手正交系的两个单位矢量，$\boldsymbol{e}_t \times \boldsymbol{e}_b = \boldsymbol{e}_z$. 因而有，$\boldsymbol{e}_b \cdot (\boldsymbol{e}_z \times \boldsymbol{E}) = (\boldsymbol{e}_b \times \boldsymbol{e}_z) \cdot \boldsymbol{E} = \boldsymbol{e}_t \cdot \boldsymbol{E} = \boldsymbol{E}_t$，用 $\boldsymbol{e}_b$ 点乘(2.8)式，并将(2.5)、(2.6)、(2.7)式代入，可得

$$\boldsymbol{E}_{0t} e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} + \boldsymbol{E}'_{0t} e^{i(\boldsymbol{k}'\cdot\boldsymbol{r}-\omega' t)} = \boldsymbol{E}''_{0t} e^{i(\boldsymbol{k}''\cdot\boldsymbol{r}-\omega'' t)}. \tag{2.9}$$

此式在界面 $z=0$ 成立，因此 x 和 y 可取任意值.

现在我们先证明 $\omega'' = \omega' = \omega$，然后再导出反射和折射定律.

在(2.9)式中令 $x=y=0$，得

$$\boldsymbol{E}_{0t} e^{-i\omega t} + \boldsymbol{E}'_{0t} e^{-i\omega' t} = \boldsymbol{E}''_{0t} e^{-i\omega'' t}. \tag{2.10}$$

对 t 求导，整理后得

$$\omega E_{0t}e^{-i\omega t} + \omega' E'_{0t}e^{-i\omega' t} = \omega'' E''_{0t}e^{-i\omega'' t}. \tag{2.11}$$

将(2.10)式与(2.11)式联立,消去 E''_{0t} 可得

$$(\omega'' - \omega)E_{0t}e^{-i\omega t} = (\omega' - \omega'')E'_{0t}e^{-i\omega' t}. \tag{2.12}$$

由 t 的任意性可得 $\omega = \omega'$. 代回(2.12)式,注意 $-E_{0t} \neq E'_{0t}$,即可求得 $\omega'' = \omega' = \omega$. 这是自然的,因为,当角频率为 ω 的电磁波射入介质后,将引起介质中束缚电子的振动而辐射. 这些辐射与原入射波一起形成反射波及折射波. 由于电子受迫振动的频率与入射波频率是相同的,见 §1.4.2(2.17)式后面的讨论,故反射波与折射波的频率也与入射波频率相同.

在(2.9)式中令 $y=0, t=0$,可得:

$$E_{0t}e^{ik_x x} + E'_{0t}e^{ik'_x x} = E''_{0t}e^{ik''_x x}. \tag{2.13}$$

此式在形式上与(2.10)式相同,用类似方法可证

$$k''_x = k'_x = k_x. \tag{2.14}$$

在(2.9)式中令 $x=0, t=0$,同理可证 $k_y'' = k_y' = k_y$. 因此,可以适当选择坐标系,使 $k_y'' = k_y' = 0$,即让 $\boldsymbol{k}, \boldsymbol{k}', \boldsymbol{k}''$ 均在 xoz 平面上. 由

$$k = k' = \omega\sqrt{\mu_1\varepsilon_1}, \quad k'' = \omega\sqrt{\mu_2\varepsilon_2}, \tag{2.15}$$

可得

$$k_x = \omega\sqrt{\mu_1\varepsilon_1}\sin\theta,\ k'_x = \omega\sqrt{\mu_1\varepsilon_1}\sin\theta',\ k''_x = \omega\sqrt{\mu_2\varepsilon_2}\sin\theta''. \tag{2.16}$$

将(2.16)式代入 $k_x = k'_x$,即得反射定律:

$$\theta = \theta'. \tag{2.17}$$

将(2.16)式代入 $k_x = k''_x$,即得折射定律:

$$\sin\theta/\sin\theta'' = \sqrt{\mu_2\varepsilon_2}/\sqrt{\mu_1\varepsilon_1} = n_2/n_1 = n_{21}. \tag{2.18}$$

(二)菲涅尔公式

现在讨论入射波、反射波与折射波的振幅关系. 因为对于每一波矢 $\boldsymbol{k}$ 有两个独立的偏振波,我们分别讨论 $\boldsymbol{E}$ 垂直于入射面(图2)和 $\boldsymbol{E}$ 平行于入射面(图3)这两种情形.

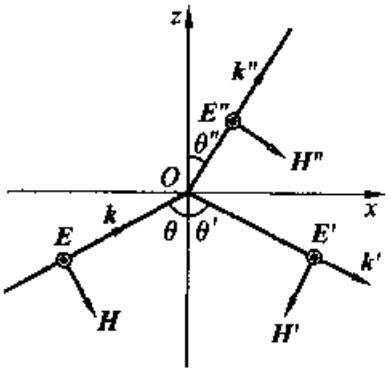


图2 E 垂直于入射面的情形

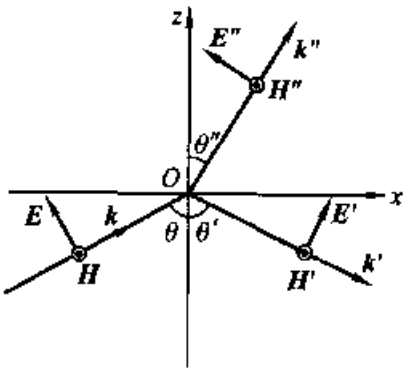


图3 E 平行于入射面的情形

在这两个图中,利用 $\boldsymbol{E}, \boldsymbol{H}$ 与 $\boldsymbol{k}$ 构成右手螺旋关系来标出电场与磁场的正方向. 如果将来证明 $E'/E < 0$,则说明当 $\boldsymbol{E}$ 取图上的方向时,$\boldsymbol{E}'$ 取与图上相反的方向,

即 $\boldsymbol{E}'$ 与 $\boldsymbol{H}'$ 要绕着 $\boldsymbol{k}'$ 轴旋转 180°才是实际方向.

对于 $\boldsymbol{E}$ 垂直于入射面的情形,用 $\boldsymbol{e}_x$ 点乘(2.8)式可得

$$E_y + E'_y = E''_y \qquad \text{即} \qquad E_\perp + E'_\perp = E''_\perp. \tag{2.19}$$

由(2.2)式可知与 $E_\perp$ 相联系的 $\boldsymbol{H}_{/\!/}$ 有

$$\boldsymbol{e}_z \times \boldsymbol{H}_{/\!/} + \boldsymbol{e}_z \times \boldsymbol{H}'_{/\!/} = \boldsymbol{e}_z \times \boldsymbol{H}''_{/\!/}. \tag{2.20}$$

由图 2 可见,$\boldsymbol{e}_z \times \boldsymbol{H}_{/\!/}$ 及 $\boldsymbol{e}_z \times \boldsymbol{H}''_{/\!/}$ 沿 $\boldsymbol{e}_y$ 方向,$\boldsymbol{e}_z \times \boldsymbol{H}'_{/\!/}$ 沿 $-\boldsymbol{e}_y$ 方向. 考虑到 $\boldsymbol{e}_z$ 与 $\boldsymbol{H}_{/\!/}$, $\boldsymbol{H}'_{/\!/}$, $\boldsymbol{H}''_{/\!/}$ 叉乘后分别沿 $\pm\boldsymbol{e}_y$ 方向,可见 $\boldsymbol{e}_z$ 与它们的夹角分别为 $\theta + (\pi/2)$, $\theta' + (\pi/2)$, $\theta'' + (\pi/2)$,故有

$$H_{/\!/} \cos\theta - H'_{/\!/} \cos\theta' = H''_{/\!/} \cos\theta''. \tag{2.21}$$

将 $H_{/\!/} = \sqrt{\varepsilon_1/\mu_1}E_\perp$, $H'_{/\!/} = \sqrt{\varepsilon_1/\mu_1}E'_\perp$, $H''_{/\!/} = \sqrt{\varepsilon_2/\mu_2}E''_\perp$, $\theta = \theta'$ 以及非铁磁质 $\mu_1 \approx \mu_2 \approx \mu_0$ 代入(2.21)式,可得

$$\sqrt{\varepsilon_1}E_\perp \cos\theta - \sqrt{\varepsilon_1}E'_\perp \cos\theta' = \sqrt{\varepsilon_2}E''_\perp \cos\theta''. \tag{2.22}$$

将(2.19)式与(2.22)式联立并利用 $\sin\theta/\sin\theta'' = \sqrt{\varepsilon_2/\varepsilon_1}$,可得

$$\begin{cases} E'_\perp/E_\perp = -\sin(\theta - \theta'')/\sin(\theta + \theta''), & (2.23) \\ E''_\perp/E_\perp = 2\cos\theta\sin\theta''/\sin(\theta + \theta''). & (2.24) \end{cases}$$

对于 $\boldsymbol{E}$ 平行于入射面的情形,类似可证

$$\begin{cases} E'_{/\!/}/E_{/\!/} = \tan(\theta - \theta'')/\tan(\theta + \theta''), & (2.25) \\ E''_{/\!/}/E_{/\!/} = 2\cos\theta\sin\theta''/[\sin(\theta + \theta'')\cos(\theta - \theta'')]. & (2.26) \end{cases}$$

(2.23) ~ (2.26)式称为菲涅尔公式,它的物理意义在光学课中已作了充分的讨论. 我们仅指出如下三点:①界面对振动方向不同的电磁波的反射系数不同;②布儒斯特定律,当 $\theta + \theta'' = 90°$ 时,$E'_{/\!/}/E_{/\!/} = \tan(\theta - \theta'')/\tan 90° = 0$,这时 $\boldsymbol{E}$ 平行于入射面的电磁波没有反射,反射光为全偏振光,θ 角称为布儒斯特角;③半波损失,若 $\varepsilon_2 > \varepsilon_1$ ($n_2 > n_1$),由(2.23)和(2.25)式可知,对于接近正入射($\theta = 0$)和掠入射($\theta = 90°$)的情形,$\boldsymbol{E}$ 与 $\boldsymbol{E}'$ 反向,反射波与入射波的相位有 180°的跃变.

(三)全反射

实验表明,当光从光密介质入射到光疏介质的分界面,如果入射角 θ 大于临界角 $\theta_{临}$ ($= \arcsin n_{21}$)时,入射波的能量将全部返回光密介质. 这种现象称为全反射现象. 现在我们通过计算折射波的 $\boldsymbol{E}''$ 及折射波的能流密度 $\boldsymbol{S}''$,讨论电磁波在上述条件下入射时所发生的物理过程.

首先,将 $k''_y = 0$, $k''/k = \sqrt{\mu_2\varepsilon_2}/\sqrt{\mu_1\varepsilon_1} = n_{21}$ 以及(2.14)式 $k''_x = k_x = k\sin\theta$ 代入 $k''^2 = (k''^2_x + k''^2_y + k''^2_z)$,可得

$$k''_z = \sqrt{k''^2 - k''^2_x} = \sqrt{k^2 n_{21}^2 - k^2\sin^2\theta} = ik\sqrt{\sin^2\theta - n_{21}^2} = i\kappa. \tag{2.27}$$

在 $\sin\theta > n_{21}$ 的条件下,$\kappa = k\sqrt{\sin^2\theta - n_{21}^2}$ 为实数,将(2.27)式代入(2.7)式,即有

$$\boldsymbol{E}'' = \boldsymbol{E}''_0 e^{i(k''\cdot x-\omega t)} = \boldsymbol{E}''_0 e^{-\kappa z} e^{i(k''_x x-\omega t)}. \tag{2.28}$$

这说明在介质 2 中,电磁波沿界面切向(e_x 方向)传播,而振幅 $E''_0 e^{-\kappa z}$ 沿 e_z 方向衰减. 透入介质 2 的厚度为:

$$\delta = 1/\kappa = 1/(k\sqrt{\sin^2\theta - n_{21}^2}) = \lambda_1/(2\pi\sqrt{\sin^2\theta - n_{21}^2}). \tag{2.29}$$

这里 λ_1 是电磁波在介质 1 中的波长. 若使用波长为 3cm 的微波,利用检波器可发现介质 2 中存在透射波. 理论和实验表明,电磁波是可以透入介质 2 的. 但问题在于,电磁波并不是不断地将能量输送到介质 2. 这可以通过计算折射波的能流密度 $\boldsymbol{S}''$得以证明.

为了简单起见,只讨论 $\boldsymbol{E}$ 垂直于入射面的分量. 将(2.28)式及(2.27)式代入 $\boldsymbol{B}'' = \sqrt{\mu_2\varepsilon_2}\, \boldsymbol{e}_n'' \times \boldsymbol{E}''$,得:

$$\boldsymbol{H}'' = \frac{\sqrt{\mu_2\varepsilon_2}}{\mu_2}\frac{\boldsymbol{k}''}{k''} \times \boldsymbol{E}'' = \sqrt{\frac{\varepsilon_2}{\mu_2}}\frac{k''_x \boldsymbol{e}_z + k''_z \boldsymbol{e}_x}{k''} \times \boldsymbol{e}_y E''$$

$$= \sqrt{\frac{\varepsilon_2}{\mu_2}}\frac{k''_x \boldsymbol{e}_z - i\kappa \boldsymbol{e}_x}{k''}\boldsymbol{E}''_0 e^{-\kappa z} e^{i(k''_x x-\omega t)}. \tag{2.30}$$

由于能流 $\boldsymbol{S}$ 是场强的二次式,不能够用复数形式计算,因而将 $\boldsymbol{E}''$及 $\boldsymbol{H}''$的实部代入 $\boldsymbol{S}'' = (\mathrm{Re}\boldsymbol{E}'') \times (\mathrm{Re}\boldsymbol{H}'')$,并利用(2.16)、(2.18)、(2.27)、(2.28)式进行整理(见习题 3),可得:

$$\begin{cases} S''_z = (k/2\mu_2\omega)E''^2_0 e^{-2\kappa z}\sin\theta\{1 + \cos[2(kx\sin\theta - \omega t)]\}, \\ S''_x = -(\kappa/2\mu_2\omega)E''^2_0 e^{-2\kappa z}\sin[2(kx\sin\theta - \omega t)]. \end{cases} \tag{2.31}$$

由此式可见,在界面法线 z 方向,能流密度 $\boldsymbol{S}''$的分量 S''_z 以 2ω 的角频率振动,在半个周期中能量由介质 1 进入介质 2,另半个周期中返回,平均而言没有能量进入介质 2. 而在界面切线 x 方向,则有能量定向流动,对时间平均 $S''_x \neq 0$. 这与实验结果完全符合,进一步计算反射波与入射波电场振幅之比,可得:

$$|E'/E| = |e^{-i\phi}| = 1, \quad \text{其中 } \phi = \arctan[\sqrt{\sin^2\theta - n_{21}^2}/\cos\theta]. \tag{2.32}$$

这说明,入射波与反射波电场振幅的大小相等,即入射波的能量将全部返回介质 1,这正是全反射现象. 另一方面,反射波与入射波的相位不同,这说明折射波参与了能量交换. 这些都是由电磁场边值关系所规定,而为实验结果所证实的.

(四)光压

电磁波对物体单位表面积上的法向辐射压力称为光压:

$$p_{光} = F_N/A = \boldsymbol{F}\cdot\boldsymbol{e}_N/A. \tag{2.33}$$

其中 $\boldsymbol{e}_N$ 为物体表面的法向单位矢量,A 为物体表面的面积,$\boldsymbol{A} = A\boldsymbol{e}_N$ 为其面积矢量.

现在利用电磁场应力张量计算电磁波对物质的辐射压力，即光压. 设平面电磁波以 θ 角入射到物质平坦的表面. 入射波与反射波沿传播方向的单位矢量分别为 e_n 及 e_n'（图 4），则入射波与反射波的电磁场应力张量分别为

$$-\overleftrightarrow{T} = -we_ne_n, \quad -\overleftrightarrow{T}' = -w'e_n'e_n'. \tag{2.34}$$

因此，物体所受的辐射压力（通常指对时间的平均值）为

$$\begin{aligned} \boldsymbol{F} &= -(\overline{\overleftrightarrow{T}} + \overline{\overleftrightarrow{T}'}) \cdot \boldsymbol{A} \\ &= -[\bar{w}Ae_n(e_n \cdot e_N) + \bar{w}'Ae_n'(e_n' \cdot e_N)]. \end{aligned} \tag{2.35}$$

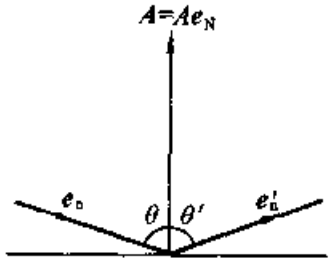


图 4　光压的计算

由图可知，$\angle(e_n, e_N) = \pi - \theta$，$\angle(e_n', e_N) = \theta$，再将反射系数 $R = |E'/E|^2 = \overline{w'}/\bar{w}$ 及 (2.35) 式代入 (2.33) 式，即得光压

$$p_{光} = \boldsymbol{F} \cdot e_N/A = \bar{w}(e_n \cdot e_N)^2 + \bar{w}'(e_n' \cdot e_N)^2 = \bar{w}(1+R)\cos^2\theta. \tag{2.36}$$

若入射光被物体完全反射（如理想导体的情形），则 $\bar{w}' = \bar{w}$，$R = 1$，这时光压为

$$p_{光} = 2\bar{w}\cos^2\theta. \tag{2.37}$$

若入射光被物体完全吸收（如绝对黑体的情形）则 $\bar{w}' = 0$，$R = 0$，这时光压为

$$p_{光} = \bar{w}\cos^2\theta. \tag{2.38}$$

若平面电磁波从一切可能方向入射到绝对黑体平面，则要在 (2.38) 式中对 θ 求平均，因为入射光来自 2π 立体角的范围，所以，

$$p_{光} = \frac{1}{2\pi}\int_0^{2\pi} d\phi \int_0^{\frac{\pi}{2}} \bar{w}\cos^2\theta \sin\theta d\theta = \frac{\bar{w}}{3} \equiv \frac{u}{3}, \tag{2.39}$$

其中用 u 表示平均能量密度 $\bar{w}$.

§3.3.3　电磁波在导体中的传播

我们知道，稳恒电流和低频交流电都是利用导体来传输的. 导体中的电磁场要引起自由电荷的重新分布和电流的流动；另一方面，电荷分布与电流又将激发电磁场. 导体内电磁波的传播过程实际上就是电磁场与自由电子相互作用的过程. 它决定了电磁波在导体中传播的性质. 本节首先从麦克斯韦方程组及电荷守恒律出发，讨论导体内自由电荷的分布规律，由此求得导体中单色电磁波的运动方程，进而讨论导体中单色平面电磁波的性质.

（一）导体内自由电荷的分布

将 $\boldsymbol{J}_f = \sigma \boldsymbol{E}$ 和 $\nabla \cdot \boldsymbol{D} = \rho_f$ 代入 $\nabla \cdot \boldsymbol{J}_f + \partial\rho_f/\partial t = 0$，可得

$$\partial\rho_f(\boldsymbol{x}, t)/\partial t = -(\sigma/\varepsilon)\rho_f(\boldsymbol{x}, t), \tag{3.1}$$

因而

$$\rho_f(\boldsymbol{x},t) = \rho_f(\boldsymbol{x},0)\mathrm{e}^{-(\sigma/\varepsilon)t}, \tag{3.2}$$

其中$\rho_f(\boldsymbol{x},0)$是$\boldsymbol{x}$点$t=0$时的电荷密度. (3.2)式说明导体内的电荷密度将按指数规律衰减,当$\tau=\varepsilon/\sigma$时$\rho_f(\boldsymbol{x},t)$减少到原来的$1/\mathrm{e}$,通常用它来表征衰减的寿命. 衰减寿命$\tau=\varepsilon/\sigma$表明$\rho_f(\boldsymbol{x},t)$的衰减与外场无关,只与导体的电导率σ与介电常数ε有关. σ大说明导体导电性能好,$\rho_f(\boldsymbol{x},t)$衰减得快; ε大说明环绕自由电荷周围的异号束缚电荷多,即束缚电荷的屏蔽作用大,导致$\boldsymbol{E}$与$\boldsymbol{J}_f$的散度减小,从而使$\rho_f(\boldsymbol{x},t)$衰减得慢. 对于良导体,$\tau=\varepsilon/\sigma$极短$\approx10^{-17}$s,这说明良导体的内部不能堆积电荷,电荷只能分布在导体表面上. 良导体内部可作$\rho_f=0$处理.

(二)良导体中单色电磁波的传播

将$\boldsymbol{J}_f=\sigma\boldsymbol{E}$, $\rho_f=0$及单色电磁波的表达式$\boldsymbol{E}(\boldsymbol{x},t)=\boldsymbol{E}(\boldsymbol{x})\mathrm{e}^{-\mathrm{i}\omega t}$和$\boldsymbol{B}(\boldsymbol{x},t)=\boldsymbol{B}(\boldsymbol{x})\mathrm{e}^{-\mathrm{i}\omega t}$代入麦克斯韦方程组,可得

$$\begin{cases} \nabla\times\boldsymbol{E}=\mathrm{i}\omega\mu\boldsymbol{H}, & (3.3)\\ \nabla\times\boldsymbol{H}=-\mathrm{i}\omega\varepsilon\boldsymbol{E}+\sigma\boldsymbol{E}=-\mathrm{i}\omega\varepsilon'\boldsymbol{E}, & (3.4)\\ \nabla\cdot\boldsymbol{E}=0, & (3.5)\\ \nabla\cdot\boldsymbol{H}=0. & (3.6) \end{cases}$$

其中

$$\varepsilon'=\varepsilon+\mathrm{i}\sigma/\omega \tag{3.7}$$

称为复介电常数. 与绝缘介质中的麦克斯韦方程组相比较,除用ε'替代ε外,别无差异. 因此在导体内的单色平面电磁波与介质中的解形式上相同,只是相应的

$$k=\omega\sqrt{\mu\varepsilon'} \tag{3.8}$$

为复数. 为此引入复波矢量

$$\boldsymbol{k}=\boldsymbol{\beta}+\mathrm{i}\boldsymbol{\alpha}, \tag{3.9}$$

其中$\boldsymbol{\alpha}$与$\boldsymbol{\beta}$为实矢量,则导体中的单色平面电磁波解即为

$$\boldsymbol{E}(\boldsymbol{x},t)=\boldsymbol{E}_0\mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}=\boldsymbol{E}_0\mathrm{e}^{-\boldsymbol{\alpha}\cdot\boldsymbol{x}}\mathrm{e}^{\mathrm{i}(\boldsymbol{\beta}\cdot\boldsymbol{x}-\omega t)}. \tag{3.10}$$

由(3.10)式可见,电磁波的振幅为$\boldsymbol{E}_0\mathrm{e}^{-\boldsymbol{\alpha}\cdot\boldsymbol{x}}$,衰减因子为$\mathrm{e}^{-\boldsymbol{\alpha}\cdot\boldsymbol{x}}$,故波矢量$\boldsymbol{k}$的虚部$\boldsymbol{\alpha}$的大小称为衰减常数. $\boldsymbol{k}$之所以出现虚部是由于介质的电导率$\sigma\neq0$造成的. $\sigma\neq0$就意味着电场对传导电流的功率密度$\boldsymbol{E}\cdot\boldsymbol{J}_f=\sigma E^2$的存在. 这正是导致导体中电磁波衰减的物理原因. 由(3.10)式可见,导体中单色平面电磁波的等幅面就是$\boldsymbol{\alpha}\cdot\boldsymbol{x}=$常数的面. 在垂直于$\boldsymbol{\alpha}$的平面上的任一点$\boldsymbol{x}$,均有$\boldsymbol{\alpha}\cdot\boldsymbol{x}=\alpha|x|\cos\theta=\alpha\cdot\overline{ON}=$常数,见图1. 这说明,$\boldsymbol{\alpha}$的方向就是单色平面电磁波等幅面的法线方向.

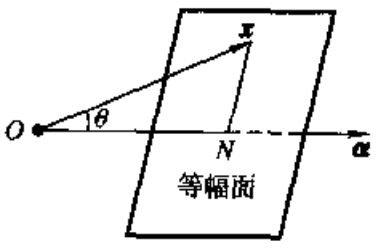


图1 $\boldsymbol{\alpha}$与等辐面垂直

(3.10)式表明,单色平面电磁波的相位是($\boldsymbol{\beta}\cdot\boldsymbol{x}-\omega t$),故波矢量 $\boldsymbol{k}$ 的实部 $\boldsymbol{\beta}$ 的大小称为相位常数.$\boldsymbol{\beta}$ 的方向就是导体中单色平面电磁波等相面的传播方向.

为了求得 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$,将 $\boldsymbol{k}=\boldsymbol{\beta}+\mathrm{i}\boldsymbol{\alpha}$ 和 $\varepsilon'=\varepsilon+\mathrm{i}\sigma/\omega$ 代入 $k=\omega\sqrt{\mu\varepsilon'}$,即有

$$\boldsymbol{k}^2=\beta^2-\alpha^2+2\mathrm{i}\boldsymbol{\alpha}\cdot\boldsymbol{\beta}=\omega^2\mu(\varepsilon+\mathrm{i}\sigma/\omega),\tag{3.11}$$

由(3.11)式两边实部与虚部相等,即有

$$\begin{cases}\beta^2-\alpha^2=\omega^2\mu\varepsilon,&(3.12)\\ \boldsymbol{\alpha}\cdot\boldsymbol{\beta}=\omega\mu\sigma/2.&(3.13)\end{cases}$$

将(3.12)(3.13)式与边值关系联立才能完全确定 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$.在普遍情形下,$\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的方向不相同.为简单起见,只讨论 $\boldsymbol{\alpha}/\!/\boldsymbol{\beta}$ 的情形(当电磁波垂直入射到导体表面时,将有这种情形发生).这时(3.13)式简化为 $\alpha\beta=\omega\mu\sigma/2$,与(3.12)式联立可得

$$\begin{cases}\alpha=\omega\sqrt{\mu\varepsilon/2}\left[\sqrt{1+(\sigma/\varepsilon\omega)^2}-1\right]^{1/2},&(3.14)\\ \beta=\omega\sqrt{\mu\varepsilon/2}\left[\sqrt{1+(\sigma/\varepsilon\omega)^2}+1\right]^{1/2}.&(3.15)\end{cases}$$

对于良导体 $\sigma\gg\varepsilon\omega$,则

$$\alpha=\beta\approx\sqrt{\omega\mu\sigma/2}.\tag{3.16}$$

穿透深度 δ 定义为波幅降至原值 $1/\mathrm{e}$ 的传播距离,由(3.10)式及(3.16)式可见:

$$\delta=1/\alpha=\sqrt{2/(\omega\mu\sigma)}.\tag{3.17}$$

对于铜来说,当频率为 50Hz 时 $\delta\sim0.9\mathrm{cm}$;当频率为 100MHz 时 $\delta\sim0.7\times10^{-3}\mathrm{cm}$.由此可见,在高频的情况下,电磁场及其在导体中激发的高频电流只能集中在导体表面的一个薄层内,这种现象称为趋肤效应.利用这一效应可做成电磁屏蔽装置.

若介质是不良导体,即 $\sigma\ll\varepsilon\omega$,展开(3.14)、(3.15)式的根式并保留前两项,可得

$$\begin{cases}\alpha=(\sigma/2)\sqrt{\mu/\varepsilon},&(3.18)\\ \beta=\omega\sqrt{\varepsilon\mu}.&(3.19)\end{cases}$$

可见 $\beta\gg\alpha$,衰减是很小的.

值得注意的是,是否是良导体应由 $\sigma\gg\varepsilon\omega$ 是否成立来衡量.例如,对于 $\nu=10^{20}\mathrm{Hz}$ 这样短波长的 X 射线来说,铜的 $\sigma/(\varepsilon\omega)\sim10^{-2}$,这样,铜就不是良导体了.这种 X 射线射到铜板上不仅不存在趋肤效应,面且能穿透铜板.

最后要指出,在良导体中,电磁波的能量主要是磁场能量,由(3.3)式及(3.10)式可得:

$$\boldsymbol{H}=\frac{1}{i\omega\mu}\nabla\times\boldsymbol{E}=\frac{1}{\mathrm{i}\omega\mu}\mathrm{i}\boldsymbol{k}\times\boldsymbol{E}=\frac{k}{\omega\mu}\boldsymbol{e}_n\times\boldsymbol{E}.\tag{3.20}$$

当 $\boldsymbol{\alpha}/\!/\boldsymbol{\beta}$ 时,将(3.16)式代入(3.9)式并利用 $\mathrm{e}^{\mathrm{i}\pi/4}=(1+\mathrm{i})/\sqrt{2}$ 可得:

$$k = \beta + \mathrm{i}\alpha = (1 + \mathrm{i})\sqrt{\omega\mu\sigma/2} = \sqrt{\omega\mu\sigma}\,\mathrm{e}^{\mathrm{i}\pi/4},$$

因而

$$\boldsymbol{H} = \sqrt{\sigma/\omega\mu}\,\mathrm{e}^{\mathrm{i}\pi/4}\boldsymbol{n} \times \boldsymbol{E}. \tag{3.21}$$

这说明,在良导体中,磁场的相位滞后电场 $\pi/4$,并且磁场能量密度与电场能量密度之比为

$$\frac{\mu H^2}{\varepsilon E^2} = \frac{\mu}{\varepsilon}\left|\frac{\boldsymbol{H}}{\boldsymbol{E}}\right|^2 = \frac{\mu}{\varepsilon}\frac{\sigma}{\omega\mu} = \frac{\sigma}{\varepsilon\omega} \gg 1. \tag{3.22}$$

§3.3.4 谐 振 腔

我们知道,低频无线电波是利用 LC 回路振荡器来激发的,其振荡频率为 $\omega_0 = 1/\sqrt{LC}$. 但是,对于高频电磁波,利用 LC 回路遇到了严重困难. 一是随着频率的提高,辐射损耗不断增加(见第四篇);二是高频情况下趋肤效应显著,导体单位表面积的功率损耗以及电容器内的介质损耗也要增大;三是为了得到高频,就要减小 L 和 C,但减小 L 和 C 受到一定的限制.

解决困难的办法是用谐振腔来代替 LC 回路. 它可以看成是在 LC 回路中减小 L 和 C 达到极限的情况. 例如,为了减小 C,可以把电容器作为两块面积很小而距离很大的平板,为了减小 L 则把电感线圈干脆用一根导线来代替,直接连接电容器的两板. 如果电感还不够小,可以把电感并联,例如用密密麻麻的一大排导线把电容器的上下两极板连接起来. 这样,在极限情形下就成为用导体板作成的一个长盒子,这就是谐振腔❶.

利用谐振腔,上述困难都得到解决,一是由于趋肤效应,电磁场全部被局限在金属空腔内,没有辐射损失. 二是作为高频电流通路的空腔内壁的表面积扩大了,故导体上单位表面积的损耗减小;又因为没有电容器上的支撑介质,也就没有介质损耗. 三是谐振频率实际上不再受 $\omega_0 = 1/\sqrt{LC}$ 的限制,可以获得很高的频率.

本节首先给出谐振腔内的电磁场满足的方程和边界条件,然后求出腔内的电磁波解,最后讨论谐振腔的频率由哪些参数确定.

(一)腔内电磁场满足的方程和边界条件

设谐振腔壁为理想导体(电导率 $\sigma \to \infty$),电磁波的穿透深度 $\delta = \sqrt{2/\omega\mu\sigma} \to 0$,即导体内的 $\boldsymbol{E} = 0, \boldsymbol{H} = 0$,取 $\boldsymbol{e}_n$ 为谐振腔壁指向腔内的法向单位矢量,则腔内电磁场遵守的方程及边界条件为:

---

❶ 在§5.1.1和§8.2.1中提到的黑体空腔可视为谐振腔.

$$\nabla^2 \boldsymbol{E} + k^2 \boldsymbol{E} = 0, \tag{4.1}$$

$$\nabla \cdot \boldsymbol{E} = 0, \tag{4.2}$$

$$\boldsymbol{B} = -(\mathrm{i}/\omega)\nabla \times \boldsymbol{E}, \tag{4.3}$$

$$\boldsymbol{e}_n \times \boldsymbol{E} = 0 \text{ 或 } E_t = 0, \tag{4.4}$$

$$\boldsymbol{e}_n \cdot \boldsymbol{B} = 0, \tag{4.5}$$

$$\boldsymbol{e}_n \times \boldsymbol{H} = \boldsymbol{\alpha}_f, \tag{4.6}$$

$$\boldsymbol{e}_n \cdot \boldsymbol{D} = \sigma_f. \tag{4.7}$$

(4.4)～(4.7)式中的 $\boldsymbol{E},\boldsymbol{B},\boldsymbol{D},\boldsymbol{H}$ 是边界上空腔一侧的场量. 为了方便,将 $E_t=0$ 代入 $\nabla \cdot \boldsymbol{E}=0$ 可得边界上对 E 的要求:

$$\partial E_n/\partial n = 0. \tag{4.8}$$

这样,利用(4.1)、(4.2)、(4.4)、(4.8)式可以求出 $\boldsymbol{E}$. 将 $\boldsymbol{E}$ 代入(4.3)式可求出 $\boldsymbol{B}$,将 $\boldsymbol{B}$ 代入(4.6)式可求出谐振腔内壁的面电流分布 $\boldsymbol{\alpha}_f$. 将 $\boldsymbol{E}$ 代入(4.7)式可求出谐振腔内壁的面电荷分布 σ_f.

(二)用分离变量法求解亥姆霍兹方程的边值问题

选直角坐标系如图 1 所示,由方程(4.1)可见,$\boldsymbol{E}$ 的任一直角分量 E_t,均满足亥姆霍兹方程

$$\nabla^2 E_t + k^2 E_t = 0, \tag{4.9}$$

将 $E_t(x,y,z)$ 分离变量,令,

$$E_t(x,y,z) = X(x)Y(y)Z(z), \tag{4.10}$$

代入(4.9)式后除以 XYZ,即有

$$\frac{1}{X}\frac{\partial^2 X}{\partial x^2} + \frac{1}{Y}\frac{\partial^2 Y}{\partial y^2} + \frac{1}{Z}\frac{\partial^2 Z}{\partial z^2} + k^2 = 0. \tag{4.11}$$

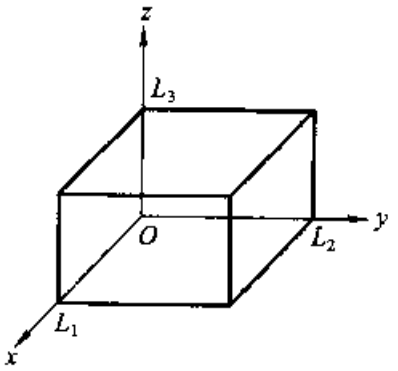


图 1　谐振腔

由于 $X(x),Y(y),Z(z)$ 相互独立,要(4.11)式成立,则每一项应为常数,分别令它们为 $-k_x^2,-k_y^2,-k_z^2$,便有

$$\mathrm{d}^2X/\mathrm{d}x^2 + k_x^2 X = 0, \tag{4.12}$$

$$\mathrm{d}^2Y/\mathrm{d}y^2 + k_y^2 Y = 0, \tag{4.13}$$

$$\mathrm{d}^2Z/\mathrm{d}z^2 + k_z^2 Z = 0, \tag{4.14}$$

$$k_x^2 + k_y^2 + k_z^2 = k^2 = \omega^2\mu_0\varepsilon_0. \tag{4.15}$$

这几个方程的解为

$$\begin{cases} X(x) = C_1\cos(k_x x) + D_1\sin(k_x x), & (4.16) \\ Y(y) = C_2\cos(k_y y) + D_2\sin(k_y y), & (4.17) \\ Z(z) = C_3\cos(k_z z) + D_3\sin(k_z z). & (4.18) \end{cases}$$

因而 E_x, E_y, E_z 的解都具有如下形式：

$$E_1(x,y,z) = [C_1\cos(k_x x) + D_1\sin(k_x x)] \cdot [C_2\cos(k_y y) + D_2\sin(k_y y)] \cdot [C_3\cos(k_z z) + D_3\sin(k_z z)]. \tag{4.19}$$

现在利用 $\nabla \cdot E = 0$ 及 $E_t = 0$ 来确定 E_x, E_y, E_z 式中的待定系数. 这里仅以 E_x 为例.

在 $x = 0$ 面上，电场的法向分量沿法向偏导为零：$\partial E_x/\partial x = 0$ 得 $D_1 = 0$；

在 $y = 0$ 面上，电场的切向分量为零：$E_x|_{y=0} = 0$ 得 $C_2 = 0$；

在 $z = 0$ 面上，同理 $E_x|_{z=0} = 0$ 得 $C_3 = 0$.

因而有

$$E_x = A_1\cos(k_x x)\sin(k_y y)\sin(k_z z). \tag{4.20}$$

这里已将三个待定常数的乘积 $C_1D_2D_3$ 用一个待定常数 A_1 表示.

在 $x = L_1$ 面，由 $\partial E_x/\partial x|_{x=L_1} = 0$ 得

$$A_1 k_x \sin(k_x L_1) \cdot \sin(k_y y) \cdot \sin(k_z z) = 0. \tag{4.21}$$

因 y 与 z 是变数，后两个因子不恒为零，故不平凡解为 $\sin k_x L_1 = 0$ 即 $k_x L_1 = m\pi$，由此得

$$k_x = m\pi/L_1 \qquad (m = 0,1,2\cdots). \tag{4.22}$$

在 $y = L_2$ 而上由 $E_x|_{y=L_2} = 0$，同理可得

$$k_y = n\pi/L_2 \qquad (n = 0,1,2\cdots). \tag{4.23}$$

在 $z = L_3$ 面上由 $E_x|_{z=L_3} = 0$，可得

$$k_z = p\pi/L_3 \qquad (p = 0,1,2\cdots). \tag{4.24}$$

对于 E_y 及 E_z 分量，同理可得

$$E_y = A_2\sin(k_x x)\cos(k_y y)\sin(k_z z), \tag{4.25}$$

$$E_z = A_3\sin(k_x x)\sin(k_y y)\cos(k_z z). \tag{4.26}$$

(4.22) ~ (4.24)式表示 k 的三个分量只能取一系列分立值，即在谐振腔里不是 k 沿任意方向的电磁波都能存在. 具体说来，可以将(4.22)式写成

$$m = \frac{k_x L_1}{\pi} = \frac{k\cos\theta_1 L_1}{\pi} = \frac{L_1\cos\theta_1}{\lambda/2}. \tag{4.27}$$

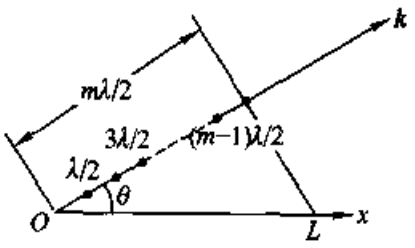


图 2 L_1 在 k 方向投影含半波长的整数倍

对 n 和 p 也有类似的结果. 这表明，只有谐振腔的三条边在 k 方向的投影都包含半波长的整数倍的那些电磁波才能在谐振腔中存在，见图 2. 这样的波称为驻波. 振幅最大的地方称为波腹，振幅最小的地方称为波节(参看 §5.2.2 图 1).

上面我们仅考虑到 $\nabla \cdot E = 0$ 对边界上的 E 的要求，对于谐振腔内的 E 仍要受 $\nabla \cdot E = 0$ 的制约，将(4.20)、(4.25)、(4.26)式代入，可得

$$\nabla \cdot E = -(A_1k_x + A_2k_y + A_3k_z)\sin(k_x x)\sin(k_y y)\sin(k_z z) = 0.$$

因为空腔内 x,y,z 是变量，后面三个因子不恒为零，故

$$A_1k_x + A_2k_y + A_3k_z = 0. \tag{4.28}$$

于是，A_1,A_2,A_3 这三个常数仅有两个是独立的. 因此，对于给定的每一组(m,n,p)值，只有两个独立的偏振波型. 即如果给定两个满足上述条件的线性无关的 $\boldsymbol{E}'$ 与 $\boldsymbol{E}''$，则任何满足上述条件的 $\boldsymbol{E}$ 均可由 $\boldsymbol{E}',\boldsymbol{E}''$ 线性地表示出来.

最后，我们导出谐振腔的谐振频率，因为谐振腔的三边长 L_1,L_2,L_3 是确定的，当(m,n,p)给定后，该电磁波的频率为

$$\omega_{mnp} = \frac{k}{\sqrt{\mu\varepsilon}} = \frac{\sqrt{k_x^2 + k_y^2 + k_z^2}}{\sqrt{\mu\varepsilon}} = \frac{\pi}{\sqrt{\mu\varepsilon}}\sqrt{\left(\frac{m}{L_1}\right)^2 + \left(\frac{n}{L_2}\right)^2 + \left(\frac{p}{L_3}\right)^2}. \tag{4.29}$$

这称为谐振腔的本征频率. 若 $L_1 \geqslant L_2 \geqslant L_3$，则谐振腔的最低频率为

$$\omega_{110} = (\pi/\sqrt{\mu\varepsilon})\sqrt{(1/L_1)^2 + (1/L_2)^2}. \tag{4.30}$$

(三)谐振控内的模式数　模密度

上面的(4.22)～(4.24)式已给出能在谐振腔内存在的电磁波所必须满足的条件为

$$k_x = m\pi/L_1, \quad k_y = n\pi/L_2, \quad k_z = p\pi/L_3. \tag{4.31}$$

其中，

$$m,n,p = 0,1,2\cdots. \tag{4.32}$$

这说明 $\boldsymbol{k}$ 只能取满足(4.31)和(4.32)式所规定的值. 一定的 m,n,p 对应着一个可能的 $\boldsymbol{k} = \sum k_i\boldsymbol{e}_i$. 那么，在谐振腔的体积 $V = L_1L_2L_3$ 中，在 $k_x \to k_x + \Delta k_x, k_y \to k_y + \Delta k_y, k_z \to k_z + \Delta k_z$ 范围内，$\boldsymbol{k}$ 可取值的数目是多少呢？为了回答这个问题，先求出在 Δk_x 范围内，k_x 可取多少种数值. 因为 m 每改变 1，k_x 就由一个可能取的 k_x 值变到另一个可能的 k_x 值. 故在 Δk_x 值范围内，k_x 可能取值的数目为

$$\Delta m = L_1\Delta k_x/\pi.$$

同理，在 Δk_y 和 Δk_z 范围内，k_y 和 k_z 可能取值的数目分别为

$$\Delta n = L_2\Delta k_y/\pi, \qquad \Delta p = L_3\Delta k_z/\pi.$$

因此，在上述给定的波矢范围内，$\boldsymbol{k}$ 可取值的数目为

$$\Delta N = \Delta m\Delta n\Delta p = L_1L_2L_3\Delta k_x\Delta k_y\Delta k_z/\pi^3 = V\Delta k_x\Delta k_y\Delta k_z/\pi^3. \tag{4.33}$$

现在我们作三个轴分别为 k_x,k_y,k_z 的直角坐标系，构成波矢空间，将波矢空间的直角坐标系换为波矢空间的球坐标系. 两坐标系间的体积元变换关系为 $\Delta k_x\Delta k_y\Delta k_z = k^2\Delta k\Delta\Omega$，故(4.33)式可写成

$$\Delta N = Vk^2\Delta k\Delta\Omega/\pi^3. \tag{4.34}$$

因为 m,n,p 均取正整数，故 $\boldsymbol{k}$ 只能在三维波矢空间的第一象限内，所对应的立体角范围是 $\Delta\Omega = 4\pi/8 = \pi/2$. 代入(4.34)式即得

$$\Delta N = \frac{Vk^2\Delta k}{\pi^3}\cdot\frac{\pi}{2} = \frac{Vk^2\Delta k}{2\pi^2}. \tag{4.35}$$

这就是在给定的 k 和 Δk 条件下，谐振腔内电磁场可能取的波矢数.

如果考虑到每一个波矢可以有两个独立的偏振，对应着两个独立的电磁场状态(模式)，则谐振腔在 $k\to k+\Delta k$ 范围内的模式数为 $Vk^2\Delta k/\pi^2$. 谐振腔单位体积内的模式数称为模密度 ρ_λ. 谐振腔内在 $k\to k+\Delta k$ 范围内的模密度为

$$\rho_\lambda = k^2\Delta k/\pi^2. \tag{4.36}$$

若用频率表示，由 $k=2\pi\nu/c$，$\Delta k=2\pi\Delta\nu/c$，则谐振腔在 $\nu\to\nu+\Delta\nu$ 范围内的模密度为

$$\rho_\nu = 8\pi\nu^2\Delta\nu/c^3. \tag{4.37}$$

§3.3.5 波导管

低频电磁波是用平行导线传输的，它的电磁场仅分布在传输线的周围，到了分米波段($\lambda\approx 10\text{cm}$)，为了避免辐射，将双线改为同轴线(由一根金属管和一根沿管轴的导线组成)，金属管起着屏蔽作用以防止电磁能向外辐射. 到了厘米波段($\lambda\approx 1\text{cm}$)，一方面由于趋肤效应，电流通过导体的有效截面减不而使焦耳热损耗随频率的升高而增加；另一方面，用来支撑金属导线的介质由于反复极化的损耗也随之增加. 这时，同轴线也不合适了，必须把金属导线抽掉而仅剩下一个金属管，这就是波导管. 它起着导引电磁场沿管的轴线方向在管内传播的作用. 常用的波导管截面是圆形或矩形，分别称为圆柱形波导管或矩形波导管，本节以矩形波导管为例介绍电磁波在管中的传播.

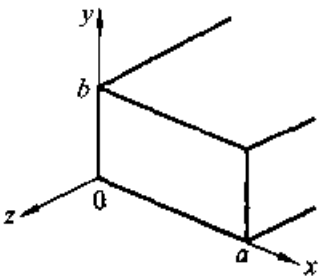


图 1 矩形波导管

(一)矩形波导管中的电磁波

矩形波导管的几何参数如图 1 所示，设电磁波沿 z 轴方向传播，定态电磁波满足的方程为

$$\nabla^2\boldsymbol{E} + k^2\boldsymbol{E} = 0, \tag{5.1}$$

$$\nabla\cdot\boldsymbol{E} = 0, \tag{5.2}$$

$$\boldsymbol{B} = -(\mathrm{i}/\omega)\nabla\times\boldsymbol{E}. \tag{5.3}$$

在管壁上还应满足边界条件：

法向分量

$$\partial E_z/\partial x\big|_{x=0} = \partial E_z/\partial x\big|_{x=a} = 0, \tag{5.4}$$

$$\partial E_y/\partial y\big|_{y=0} = \partial E_y/\partial y\big|_{y=b} = 0, \tag{5.5}$$

切向分量

$$E_y|_{x=0} = E_y|_{x=a} = 0, \quad E_z|_{x=0} = E_z|_{x=a} = 0, \tag{5.6}$$

$$E_x|_{y=0} = E_x|_{y=b} = 0, \quad E_z|_{y=0} = E_z|_{y=b} = 0. \tag{5.7}$$

由于电磁波沿 z 轴方向传播，它的解具有如下形式：

$$\boldsymbol{E}(\boldsymbol{x},t) = \boldsymbol{E}(x,y)e^{i(k_z z-\omega t)}. \tag{5.8}$$

这样，只要求出 $\boldsymbol{E}(x,y) = \sum_{j=1}^{3} E_j(x,y)\boldsymbol{e}_j$ 即可. 将(5.8)式代入亥姆霍兹方程(5.1)便有

$$0 = \nabla^2 \boldsymbol{E} + k^2 \boldsymbol{E} = (\nabla^2 + k^2)\boldsymbol{E}(x,y)e^{i(k_z z-\omega t)}$$

$$= \left[\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\boldsymbol{E}(x,y) + (k^2 - k_z^2)\boldsymbol{E}(x,y)\right]e^{i(k_z z-\omega t)}$$

由此得

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\boldsymbol{E}(x,y) + (k^2 - k_z^2)\boldsymbol{E}(x,y) = 0,$$

用 $\boldsymbol{e}_j$ 点乘便有

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)E_j(x,y) + (k^2 - k_z^2)E_j(x,y) = 0. \tag{5.9}$$

现在用分离变量法求 $E_j(x,y)$，令

$$E_j(x,y) = X(x)Y(y), \tag{5.10}$$

代入(3.9)式后可分离出两个常微分方程：

$$X''(x) + k_x^2 X(x) = 0,$$

$$Y''(y) + k_y^2 Y(y) = 0,$$

$$k^2 = k_x^2 + k_y^2 + k_z^2. \tag{5.11}$$

求出 $X(x)$ 及 $Y(y)$ 后，代入(5.10)式可得(5.8)式的 j 分量

$$E_j(\boldsymbol{x},t) = [C_1\cos(k_x x) + D_1\sin(k_x x)][C_2\cos(k_y y) + D_2\sin(k_y y)]e^{i(k_z z-\omega t)}. \tag{5.12}$$

下面利用边界条件定系数.

先求 $E_z(\boldsymbol{x},t)$. 由(5.4)式 $\partial E_z/\partial x|_{x=0} = 0$ 得 $D_1 = 0$，故

$$E_z(\boldsymbol{x},t) = C_1\cos(k_x x)[C_2\cos(k_y y) + D_2\sin(k_y y)]e^{i(k_z z-\omega t)}.$$

由(5.7)式 $E_z|_{y=0} = 0$ 得 $C_2 = 0$，故

$$E_z(\boldsymbol{x},t) = A_1\cos(k_x x)\sin(k_y y)\, e^{i(k_z z-\omega t)}. \tag{5.13}$$

同理可得

$$E_y(\boldsymbol{x},t) = A_2\sin(k_x x)\cos(k_y y)e^{i(k_z z-\omega t)}, \tag{5.14}$$

$$E_x(\boldsymbol{x},t) = A_3\sin(k_x x)\sin(k_y y)e^{i(k_z z-\omega t)}. \tag{5.15}$$

利用(5.4)式的第二个等式可确定 k_x

$$0 = \partial E_x/\partial x\big|_{x=a} = -A_1 k_x \sin(k_x a)\sin(k_y y)e^{i(k_z z-\omega t)},$$

故有

$$\sin k_x a = 0, \quad \text{即} \quad k_x = m\pi/a \quad (m = 0,1,2,\cdots). \tag{5.16}$$

利用(5.7)式可确定 k_y

$$0 = E_z\big|_{y=b} = -A_1\cos(k_x x)\sin(k_y b)e^{i(k_z z-\omega t)},$$

故有

$$\sin k_y b = 0, \quad \text{即} \quad k_y = n\pi/b \quad (n = 0,1,2,\cdots). \tag{5.17}$$

对于每一组(m,n)值,仅有两个独立的波模. 由管内$\nabla \cdot \boldsymbol{E} = 0$可得

$$0 = \nabla \cdot \boldsymbol{E} = (-A_1 k_x - A_2 k_y + A_3 i k_z)\sin(k_x x)\sin(k_y y)e^{i(k_z z-\omega t)},$$

故有

$$A_1 k_x + A_2 k_y - iA_3 k_z = 0. \tag{5.18}$$

上式表明,在A_1, A_2, A_3之中仅有两个独立. 由(5.13)、(5.14)、(5.15)式可见,当$\{k_x, k_y, k_z\}$给定时,$\boldsymbol{E}$就决定于$\{A_1, A_2, A_3\}$. 如果$\boldsymbol{E}'$决定于$\{A'_1, A'_2, A'_3\}$,$\boldsymbol{E}''$决定于$\{A''_1, A''_2, A''_3\}$,两者又是独立的波模,则任意的波模$\boldsymbol{E}$均可由$\boldsymbol{E}'$和$\boldsymbol{E}''$的线性组合得到. 也就是说,对每一组给定的(m,n),只有两种独立的波模.

在§3.3.1中已指出,在无界空间中的平面电磁波都是横波,$\boldsymbol{E}, \boldsymbol{B}, \boldsymbol{k}$互相垂直. 但在波导管中,电磁波就不一定是横波了,$\boldsymbol{E}$和$\boldsymbol{B}$都可能有沿$\boldsymbol{k}$方向的分量. 但是,应用最广的还是横电波(T E 波)和横磁波(T M 波),前者的$\boldsymbol{E} \perp \boldsymbol{k}$,后者的$\boldsymbol{B} \perp \boldsymbol{k}$.

(二)截止频率

波导管中(m,n)型电磁波的截止频率由下列各式确定:

$$k = \omega\sqrt{\mu\varepsilon}, \tag{5.19}$$

$$k^2 = k_x^2 + k_y^2 + k_z^2, \tag{5.20}$$

$$k_x = m\pi/a, \quad k_y = n\pi/b. \tag{5.21}$$

由(5.20)式可见,若k_x和k_y很大,以致$k_z^2 = k^2 - (k_x^2 + k_y^2) < 0$,则$k_z = i\kappa$为纯虚数. 这导致

$$\boldsymbol{E}(\boldsymbol{x},t) = \boldsymbol{E}(x,y)e^{i(k_z z-\omega t)} = \boldsymbol{E}(x,y)e^{i(i\kappa z-\omega t)} = \boldsymbol{E}(x,y)e^{-\kappa z - i\omega t}.$$

这说明,电磁场沿z方向衰减而不能沿波导管传播. 这样的电磁波称为消失波. 因此,能沿z轴方向传播的波应满足$k_z^2 > 0$. 通常称与$k_z^2 = 0$对应的频率为该波模的截止频率. 对于(m,n)型波,其截止角频率为

$$\omega_{c,mn} = \frac{k}{\sqrt{\mu\varepsilon}} = \frac{\sqrt{k_x^2 + k_y^2}}{\sqrt{\mu\varepsilon}} = \frac{\pi}{\sqrt{\mu\varepsilon}}\sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2}. \tag{5.22}$$

对于 TE_{10} 波(即 $m=1,n=0$ 的横电波),若 $a>b$,它的截止频率为

$$\nu_{c,10}=\frac{\omega_{c,10}}{2\pi}=\frac{1}{2\pi}\frac{\pi}{\sqrt{\mu\varepsilon}}\sqrt{\left(\frac{1}{a}\right)^2+\left(\frac{0}{b}\right)^2}=\frac{1}{2a\sqrt{\mu\varepsilon}}\tag{5.23}$$

若波导管内真空,相应的截止波长为

$$\lambda_{c,10}=\frac{c}{\nu_{c,10}}=2ac\sqrt{\mu_0\varepsilon_0}=2a.\tag{5.24}$$

可见截止波长与波导管尺寸同数量级.

通过上述讨论可以看出,波导管可以起滤波器的作用.对于给定的角频率 ω,可以适当选择波导管的尺寸 a 和 b,使波导管只让 TE_{10} 波通过,而不让其他波模通过.方法是只让 TE_{10} 波的 $k_z^2=k^2-k_x^2-k_y^2>0$,即要求

$$\left(\frac{\omega}{c}\right)^2>\left(\frac{m\pi}{a}\right)^2+\left(\frac{n\pi}{b}\right)^2,\tag{5.25}$$

选择 $a>b$,且令

$$(\omega/c)^2>(\pi/a)^2,\quad(\omega/c)^2<(\pi/b)^2,\quad(\omega/c)^2<(2\pi/a)^2,\tag{5.26}$$

则 TE_{10} 波可通过,TE_{01},TE_{20} 以及其他波模均不能通过.

(三)TE_{10} 波的电磁场

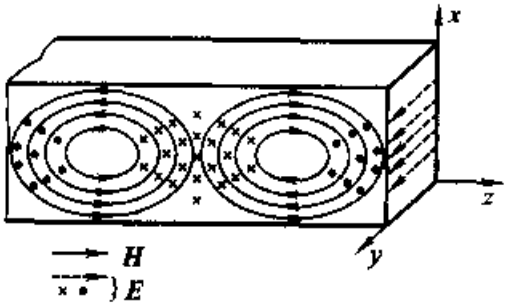


图 2　TE_{10} 波的电磁场

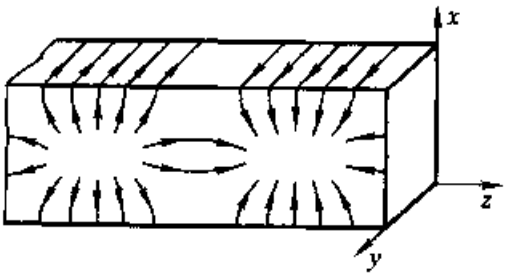


图 3　TE_{10} 波的管壁电流

TE_{10} 波的 $m=1,n=0$;由此得 $k_x=m\pi/a=\pi/a$,$k_y=n\pi/b=0$.作为横电波,其 $E_z=0$,故 $A_3=0$,将 $k_y=0$ 及 $A_3=0$ 代入(5.18)式即得 $A_1=0$.由此可见,$\boldsymbol{E}$ 仅有 E_y 分量

$$E_y=A_2\sin(k_x x)e^{i(k_z z-\omega t)}.\tag{5.27}$$

将(5.27)式代入(5.3)式,可得

$$\boldsymbol{H}=\frac{\boldsymbol{B}}{\mu}=-\frac{i}{\mu\omega}\nabla\times\boldsymbol{E}=-\frac{i}{\mu\omega}\left(-\boldsymbol{e}_x\frac{\partial E_y}{\partial z}+\boldsymbol{e}_z\frac{\partial E_y}{\partial x}\right)$$

$$=\frac{iA_2}{\mu\omega}\left[\boldsymbol{e}_x ik_z\sin\frac{\pi x}{a}-\boldsymbol{e}_z\frac{\pi}{a}\cos\frac{\pi x}{a}\right]e^{i(k_z z-\omega t)}.\tag{5.28}$$

通常采用待定常数 H_0 代替 A_2,即取 $A_2=i\mu\omega aH_0/\pi$,不为零的电磁场诸分量即为

$$H_z=H_0\cos(\pi x/a)e^{i(k_z z-\omega t)},\tag{5.29}$$

$$E_y = (i\mu\omega a/\pi)H_0\sin(\pi x/a)e^{i(k_z z-\omega t)}, \tag{5.30}$$

$$H_x = -(ik_z a/\pi)H_0\sin(\pi x/a)e^{i(k_z z-\omega t)}. \tag{5.31}$$

H_0 的大小由激发功率确定.

波导管的管壁电流由边值关系给出:

$$\boldsymbol{\alpha}_f = \boldsymbol{e}_n \times \boldsymbol{H}. \tag{5.32}$$

TE_{10}波的电磁场与管壁电流如图 2 及图 3 所示.

习 题

1. 平面电磁波垂直入射到金属表面,试证明透入金属内部的电磁波能量全部变为焦尔热.

2. 试证明:对于一定频率的电磁场,电磁场的边值关系不完全独立,即可以由其中的两式 $\boldsymbol{e}_n \times (\boldsymbol{E}_2 - \boldsymbol{E}_1) = 0$, $\boldsymbol{e}_n \times (\boldsymbol{H}_2 - \boldsymbol{H}_1) = \boldsymbol{\alpha}_f$,导出 $\boldsymbol{e}_n \cdot (\boldsymbol{D}_2 - \boldsymbol{D}_1) = \sigma_f$, $\boldsymbol{e}_n \cdot (\boldsymbol{B}_2 - \boldsymbol{B}_1) = 0$.

3. 试证明在全反射现象中,折射波的能流密度 S'' 的分量为 §3.3.2(2.31)式所示.

4. 大多数微波炉的工作频率为 2.5GHz,在此频率下,牛肉的 $\varepsilon = 49\varepsilon_0$, $\sigma = 2\Omega^{-1}\mathrm{m}^{-1}$,试估计穿透深度 $\delta = ?$ (Ω 是电阻的单位欧姆).

5. 平面电磁波由真空倾斜入射到导电介质表面上,入射角为 θ_i,求导电介质中电磁波的相速度和衰减长度. 若导电介质为金属,结果如何?

6. 一对无穷大的平行理想导体板,相距为 b,电磁波沿平行于板面的 z 方向传播. 设波在 x 方向是均匀的,求可能传播的波模.

7. 无限长的矩形波导管,在 $z=0$ 处被一块垂直地插入的理想导体平板完全封闭,求在 $z=-\infty$ 到 $z=0$ 这段管内可能存在的波模.

8. 电磁波 $\boldsymbol{E}(x,y,z,t) = \boldsymbol{E}(x,y)e^{i(k_z z-\omega t)}$ 在波导管中沿 z 方向传播,试使用 $\nabla \times \boldsymbol{E} = i\omega\mu_0 \boldsymbol{H}$ 及 $\nabla \times \boldsymbol{H} = -i\omega\varepsilon_0 \boldsymbol{E}$ 证明电磁场所有分量都可用 $E_z(x,y)$ 及 $H_z(x,y)$ 这两个分量表示.

9. 写出矩形波导管内磁场 $\boldsymbol{H}$ 满足的方程及边界条件.

10. 试证明矩形波导管内不存在 TM_{m0} 或 TM_{0n} 波.

11. 已知矩形波导管宽为 a,高为 b,在其中传播的 TE_{10} 波如 §3.3.5(5.29)、(5.30)、(5.31)式所示,求 TE_{10} 波的传输功率 P,管壁上的电荷密度 σ_f 及线电流密度 $\boldsymbol{\alpha}_f$.

12. 频率为 30×10^9Hz 的微波,在 0.7cm × 0.4cm 的矩形波导管中能以什么波模传播? 在 0.7cm × 0.6cm 的矩形波导管中能以什么波模传播?



第四篇　电荷体系与电磁场的相互作用

上一篇讨论了稳恒的电场、磁场，以及电磁波的传播. 在这一篇里，我们来讨论电荷体系与电磁场的相互作用.

本篇第一章讨论宏观电荷体系辐射的电磁场. 第二章讨论作任意运动的带电粒子激发的电磁场，特别是介绍在电子加速器中作加速运动的电子所产生的韧致辐射（直线加速运动）和同步辐射（圆周运动），以及在介质中高速运动（高于介质中的光速）的电子产生的切连科夫辐射. 在此基础上，利用经典理论讨论运动带电粒子激发的自有场和辐射场对粒子本身的反作用. 它们分别表现为粒子的电磁质量和辐射阻尼力.

最后，在第三章里，讨论外来电磁波通过介质时与介质中带电粒子的相互作用，和由此导致的电磁波的散射、吸收和色散现象.

第一章　宏观电荷体系辐射的电磁场

本章讨论宏观电荷体系辐射的电磁场. 第一节首先说明在普遍情形下描述电磁场的势不是唯一的,接着给出在不同规范下电磁势的场方程,并求出洛伦兹规范下电磁势的场方程的解——推迟势. 第二节讨论小区域电荷电流分布的辐射问题,通过对推迟势的多极展开着重讨论最基本且最重要的电偶极辐射,分析电偶极辐射场及其性质.

§4.1.1　电磁势的场方程　推迟势

(一)规范变换与规范不变性

在§2.3.3已引入标势 φ 和矢势 $\boldsymbol{A}$ 描述普遍的电磁场:

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}, \quad \boldsymbol{E} = -\nabla\varphi - \partial\boldsymbol{A}/\partial t. \tag{1.1}$$

现在来证明,给定的 $\boldsymbol{E}$ 和 $\boldsymbol{B}$ 并不对应于唯一的$(\boldsymbol{A},\varphi)$. 例如,对于任意的可微时空函数 $\psi(x,t)$,作变换:

$$\boldsymbol{A}' = \boldsymbol{A} + \nabla\psi, \quad \varphi' = \varphi - \partial\psi/\partial t \tag{1.2}$$

相应于$(\boldsymbol{A}',\varphi')$的 $\boldsymbol{B}'$和 $\boldsymbol{E}'$为:

$$\begin{aligned}\boldsymbol{E}' &= -\nabla\varphi' - \partial\boldsymbol{A}'/\partial t = -\nabla(\varphi - \partial\psi/\partial t) - \partial(\boldsymbol{A} + \nabla\psi)/\partial t\\ &= -\nabla\varphi - \partial\boldsymbol{A}/\partial t = \boldsymbol{E},\\ \boldsymbol{B}' &= \nabla\times\boldsymbol{A}' = \nabla\times(\boldsymbol{A} + \nabla\psi) = \nabla\times\boldsymbol{A} = \boldsymbol{B}.\end{aligned}$$

这说明$(\boldsymbol{A}',\varphi')$与$(\boldsymbol{A},\varphi)$描述同一电磁场. 变换式(1.2)称为势的规范变换,当势作规范变换时 $\boldsymbol{E}$ 和 $\boldsymbol{B}$ 保持不变的性质称为规范不变性. 势之所以有这种选择的自由是因为,确定一个矢量场必须同时给定它的旋度和散度. 势的定义式(1.1)只规定了$\nabla\times\boldsymbol{A}$ 而没有规定$\nabla\cdot\boldsymbol{A}$,因而可以任意选择$\nabla\cdot\boldsymbol{A}$ 的值. 为计算方便,人们通常选取$\nabla\cdot\boldsymbol{A}=0$ 或$\nabla\cdot\boldsymbol{A} = -c^{-2}\partial\varphi/\partial t$. 这两个辅助条件分别称为库仑规范和洛伦兹规范. 下面分别讨论在这两种规范下电磁势的场方程.

(二)电磁势的场方程

$\boldsymbol{A}$ 与 φ 的方程可以由麦克斯韦方程组出发推导出. 将 $\boldsymbol{H} = \boldsymbol{B}/\mu_0, \boldsymbol{D} = \varepsilon_0\boldsymbol{E}, \boldsymbol{B} = \nabla\times\boldsymbol{A}, \boldsymbol{E} = -\nabla\varphi - \partial\boldsymbol{A}/\partial t$ 代入$\nabla\times\boldsymbol{H} = \boldsymbol{J}_f + \partial\boldsymbol{D}/\partial t$ 并展开$\nabla\times(\nabla\times\boldsymbol{A})$,整理后得

$$\nabla^2\boldsymbol{A} - \frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2} - \nabla\left(\nabla\cdot\boldsymbol{A} + \frac{1}{c^2}\frac{\partial\varphi}{\partial t}\right) = -\mu_0\boldsymbol{J}_f, \tag{1.3}$$

再将 $\boldsymbol{D} = \varepsilon_0\boldsymbol{E}$ 及 $\boldsymbol{E} = -\nabla\varphi - \partial\boldsymbol{A}/\partial t$ 代入$\nabla\cdot\boldsymbol{D} = \rho_f$ 得

$$\varepsilon_0\nabla\cdot(-\nabla\varphi - \partial\boldsymbol{A}/\partial t) = \rho_f. \tag{1.4}$$

1. 采用库仑规范,其辅助条件为

$$\nabla \cdot \boldsymbol{A} = 0. \tag{1.5}$$

将(1.5)式代入(1.3)式与(1.4)式,可得电磁势的场方程为

$$\begin{cases} \nabla^2 \boldsymbol{A} - \dfrac{1}{c^2}\dfrac{\partial^2 \boldsymbol{A}}{\partial t^2} - \dfrac{1}{c^2}\dfrac{\partial \nabla \varphi}{\partial t} = -\mu_0 \boldsymbol{J}_t, & (1.6) \\ \nabla^2 \varphi = -\rho_t / \varepsilon_0. & (1.7) \end{cases}$$

这说明,采用库仑规范时,φ 满足泊松方程. 将求得的 φ 代入(1.6)式便可解出 $\boldsymbol{A}$.

2. 采用洛伦兹规范,其辅助条件为

$$\nabla \cdot \boldsymbol{A} + \frac{1}{c^2}\frac{\partial \varphi}{\partial t} = 0. \tag{1.8}$$

将(1.8)式代入(1.3)、(1.4)式,得到更对称的方程

$$\begin{cases} \nabla^2 \boldsymbol{A} - \dfrac{1}{c^2}\dfrac{\partial^2 \boldsymbol{A}}{\partial t^2} = -\mu_0 \boldsymbol{J}_t, & (1.9) \\ \nabla^2 \varphi - \dfrac{1}{c^2}\dfrac{\partial^2 \varphi}{\partial t^2} = -\dfrac{\rho_t}{\varepsilon_0}. & (1.10) \end{cases}$$

称为达朗伯方程. 由(1.8)~(1.10)式解出 $\boldsymbol{A}$、φ 后代入(1.1)式即可求得电磁场 $\boldsymbol{E}$,$\boldsymbol{B}$. 值得注意的是,不是所有满足达朗伯方程的解都满足麦克斯韦方程组,只有同时满足洛伦兹条件(1.8)式的解才是电磁场的解.

下面,我们在洛伦兹规范下求解达朗伯方程.

(三)推迟势

首先,由于矢势的直角分量 A_i 满足的方程与标势 φ 的方程形式上一致,因此可先求 φ,然后推广到 A_i 而得 $\boldsymbol{A}$. 其次,由于场的叠加性,可以先求 t 时刻位于坐标原点的电荷 $\mathrm{d}Q(t)$ 所激发的标势,再推广到任意电荷分布所激发的标势❶. 在这种情况下,电荷密度为 $\rho_t(\boldsymbol{x},t) = \mathrm{d}Q(t)\delta^3(\boldsymbol{x})$,标势的达朗伯方程为

$$\nabla^2 \varphi - \frac{1}{c^2}\frac{\partial^2 \varphi}{\partial t^2} = -\frac{1}{\varepsilon_0}\mathrm{d}Q(t)\delta^3(\boldsymbol{x}). \tag{1.11}$$

除坐标原点 $\boldsymbol{x}=0$ 以外的整个空间中,方程(1.11)变为

$$\nabla^2 \varphi - \frac{1}{c^2}\frac{\partial^2 \varphi}{\partial t^2} = 0. \tag{1.12}$$

因为问题具有球对称性,$\varphi(\boldsymbol{x},t)$ 与 θ,ϕ 无关,在球坐标系中(1.11)式可表示为:

$$\frac{1}{R^2}\frac{\partial}{\partial R}\left(R^2 \frac{\partial \varphi}{\partial R}\right) - \frac{1}{c^2}\frac{\partial^2 \varphi}{\partial t^2} = 0. \tag{1.13}$$

---

❶ 这实质上就是格林函数方法,参看 §3.1.4.

为了简化方程，作变换

$$\varphi(R,t) = u(R,t)/R, \tag{1.14}$$

代入方程(1.13)，得到 $u(R,t)$ 满足的方程

$$\frac{\partial^2 u}{\partial R^2} - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = 0. \tag{1.15}$$

这方程的解已在 §3.3.1 中求得为

$$u(R,t) = f_1(t - R/c) + f_2(t + R/c). \tag{1.16}$$

因此，除去 $x=0$ 点外，$\varphi(R,t)$ 的解为

$$\varphi(R,t) = \frac{f_1(t - R/c)}{R} + \frac{f_2(t + R/c)}{R}. \tag{1.17}$$

显然，第一项表示向外发射的球面波，第二项表示向内收缩的球面波. 在讨论辐射问题时，我们只需要第一项，故简单地取

$$\varphi(R,t) = \frac{f_1(t - R/c)}{R}. \tag{1.18}$$

为了确定在包含 $x=0$ 的全空间中满足方程(1.11)式的 $f_1(t-R/c)$ 的形式，根据与静电情形的对比提出一个合理的尝试解，因为静电情形应为(1.18)式的一个特例. 一个位于坐标原点的静止的点电荷 $\mathrm{d}Q$ 所激发的标势为

$$\varphi(R) = \mathrm{d}Q/(4\pi\varepsilon_0 R). \tag{1.19}$$

比较(1.18)与(1.19)式可得 $f_1(t-R/c) = \mathrm{d}Q(t-R/c)/(4\pi\varepsilon_0)$，从而得到尝试解

$$\varphi(R,t) = \mathrm{d}Q(t - R/c)/(4\pi\varepsilon_0 R). \tag{1.20}$$

这个解的正确性可以通过将(1.20)式代入(1.11)式得到验证，见习题 3.

如果电荷不是位于坐标原点的点电荷 $\mathrm{d}Q$，而是位于 x' 点的点电荷 $\mathrm{d}Q$，则在(1.20)式中源点到场点的距离 R 应该用 $r = |x - x'|$ 取代，即方程(1.11)的解应为

$$\varphi(x,t) = \mathrm{d}Q(x', t - r/c)/(4\pi\varepsilon_0 r). \tag{1.21}$$

对于空间中任意的电荷分布 $\rho_f(x,t)$，由场的叠加性可知，它激发的标势为：

$$\varphi(x,t) = \int \frac{\rho_f(x', t - r/c)\,\mathrm{d}V'}{4\pi\varepsilon_0 r}. \tag{1.22}$$

由于 A_i 满足的方程与 φ 满足的方程有相同形式，故在用 A_i 取代 φ 的同时，用 $\mu_0 J_i$ 取代 ρ_f/ε_0 即可得 A_i 的解❶，于是，

$$\boldsymbol{A} = \sum_i \boldsymbol{e}_i A_i = \int \frac{\sum_i \boldsymbol{e}_i \mu_0 J_i(x', t - r/c)}{4\pi r}\mathrm{d}V' = \frac{\mu_0}{4\pi}\int \frac{\boldsymbol{J}_f(x', t - r/c)}{r}\mathrm{d}V'. \tag{1.23}$$

不难证明，由(1.22)、(1.23)式确定的 φ 和 $\boldsymbol{A}$ 满足洛伦兹条件，见习题 4. 因

❶ 和前面一样，暂以 J_i 表示 $\boldsymbol{J}_f$ 在 i 方向的分量.

而，以上两式就是电磁场的解. 这两个公式说明，x 点 t 时刻的势是在 $(t-r/c)=t-|x-x'|/c$ 时刻在 x' 处的电荷电流分布所激发的，如图 1 所示. 这反映了电磁场具有一定传播速度这一事实. 由于电荷电流在空间某一点所产生的物理作用要经过一段时间才能传到场点，所以(1.22)，(1.23)式称为推迟势.

将推迟势代入 $\boldsymbol{B}=\nabla\times\boldsymbol{A}$ 及 $\boldsymbol{E}=-\nabla\varphi-\partial\boldsymbol{A}/\partial t$，即可求得随时间变化的电荷电流分布所激发的电磁场.

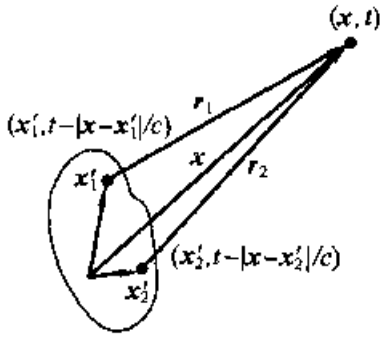


图 1　场点与场源

§4.1.2　电偶极辐射

现在讨论宏观电荷体系的辐射问题，所考虑的电荷体系的线度 l 远小于其辐射的波长，即 $l\ll\lambda$. 我们首先将推迟势作多极展开，并着重讨论电荷体系的电偶极矩 $\boldsymbol{p}_e=\int\rho_f(\boldsymbol{x}',t)\boldsymbol{x}'\mathrm{d}V'$ 所激发的 $\varphi^{(1)}$ 和 $\boldsymbol{A}^{(1)}$. 由此可求得电偶极场 $\boldsymbol{E}^{(1)}$ 和 $\boldsymbol{B}^{(1)}$，最后给出电偶极辐射的角分布及总辐射功率.

(一)推迟势的多极展开

上节已给出随时间变化的电荷体系激发的推迟势(坐标原点取在电荷体系内部)：

$$\varphi(\boldsymbol{x},t)=\frac{1}{4\pi\varepsilon_0}\int\frac{\rho_f(\boldsymbol{x}',t-r/c)}{r}\mathrm{d}V', \tag{2.1}$$

$$\boldsymbol{A}(\boldsymbol{x},t)=\frac{\mu_0}{4\pi}\int\frac{\boldsymbol{J}_f(\boldsymbol{x}',t-r/c)}{r}\mathrm{d}V'. \tag{2.2}$$

对于源点坐标 $\boldsymbol{x}'$ 来说，有 $|\boldsymbol{x}'|\leqslant l$；而对于辐射场的场点坐标 $\boldsymbol{x}$ 来说，有 $|\boldsymbol{x}|\geqslant l$. 由 $\boldsymbol{r}=\boldsymbol{x}-\boldsymbol{x}'=\boldsymbol{R}-\boldsymbol{x}'$ 知 $|\boldsymbol{x}'|=|\boldsymbol{R}-\boldsymbol{r}|$ 相对 R 与 r 是小量. 将 $f(r)=\rho_f(\boldsymbol{x}',t-r/c)/r$ 在 $r=R$ 处作泰勒展开：

$$\frac{\rho_f(\boldsymbol{x}',t-r/c)}{r}=\frac{\rho_f(\boldsymbol{x}',t-R/c)}{R}-\boldsymbol{x}'\cdot\nabla\frac{\rho_f(\boldsymbol{x}',t-R/c)}{R}+\frac{1}{2}(\boldsymbol{x}'\cdot\nabla)^2\frac{\rho_f(\boldsymbol{x}',t-R/c)}{R}+\cdots, \tag{2.3}$$

因而

$$\varphi(\boldsymbol{x},t)=\frac{1}{4\pi\varepsilon_0}\int\frac{\rho_f(\boldsymbol{x}',t-R/c)}{R}\mathrm{d}V'-\frac{1}{4\pi\varepsilon_0}\int\boldsymbol{x}'\cdot\nabla\frac{\rho_f(\boldsymbol{x}',t-R/c)}{R}\mathrm{d}V'+\cdots. \tag{2.4}$$

同理，将 $\boldsymbol{J}_f(\boldsymbol{x}',t-r/c)/r$ 在 $r=R$ 展开，再代回(2.2)式，即有

$$\boldsymbol{A}(\boldsymbol{x},t)=\frac{\mu_0}{4\pi}\int\frac{\boldsymbol{J}_f(\boldsymbol{x}',t-R/c)}{R}\mathrm{d}V'-\frac{\mu_0}{4\pi}\int\boldsymbol{x}'\cdot\nabla\frac{\boldsymbol{J}_f(\boldsymbol{x}',t-R/c)}{R}\mathrm{d}V'+\cdots. \tag{2.5}$$

(2.4)式的第一项是电荷体系总电量的贡献 $\varphi^{(0)}(\boldsymbol{x},t)=Q/(4\pi\varepsilon_0R)$. 在第二项中利用 R 及 ∇ 与积分变量 $\boldsymbol{x}'$ 无关而提出积分号外，可得

$$\varphi^{(1)}(\boldsymbol{x},t)=-\nabla\cdot\frac{1}{4\pi\varepsilon_0R}\int\boldsymbol{x}'\rho_f(\boldsymbol{x}',t-R/c)\mathrm{d}V'=-\nabla\cdot\frac{\boldsymbol{p}_e(t-R/c)}{4\pi\varepsilon_0R}. \tag{2.6}$$

这就是电荷体系随时间变化的电偶极矩所激发的标势.

在(2.5)式第一项中同样将 R 提出积分号外，利用第二篇第四章习题 7 的结果 $\dot{\boldsymbol{p}}_e(t)=\int_V\boldsymbol{J}_f(\boldsymbol{x}',t)\mathrm{d}V'$ 可得

$$\boldsymbol{A}^{(1)}(\boldsymbol{x},t)=\frac{\mu_0}{4\pi R}\int\boldsymbol{J}_f(\boldsymbol{x}',t-R/c)\mathrm{d}V'=\frac{\mu_0\dot{\boldsymbol{p}}_e(t-R/c)}{4\pi R}. \tag{2.7}$$

这就是电荷体系随时间变化的电偶极矩所激发的矢势. 注意，在稳恒情况下积分 $\int_V\boldsymbol{J}_f(\boldsymbol{x}')\mathrm{d}V'=0$，但在随时间变化的电磁场中，传导电流不处处连续，因而不能形成闭合的流管，积分式 $\int_V\boldsymbol{J}_f(\boldsymbol{x}',t-R/c)\mathrm{d}V'$ 不为零.

(二)电偶极场及其性质

将(2.6)式与(2.7)式代入 $\boldsymbol{B}=\nabla\times\boldsymbol{A}$ 及 $\boldsymbol{E}=-\nabla\varphi-\partial\boldsymbol{A}/\partial t$，即有

$$\boldsymbol{B}^{(1)}=\frac{\mu_0}{4\pi}\nabla\times\frac{\dot{\boldsymbol{p}}_e(t-R/c)}{R}, \tag{2.8}$$

$$\boldsymbol{E}^{(1)}=\frac{1}{4\pi\varepsilon_0}\left[\nabla\nabla\cdot\frac{\boldsymbol{p}_e(t-R/c)}{R}-\frac{1}{c^2}\frac{\ddot{\boldsymbol{p}}_e(t-R/c)}{R}\right]. \tag{2.9}$$

在电荷体系区域之外，电场满足波动方程 $\nabla^2\boldsymbol{E}=c^{-2}\partial^2\boldsymbol{E}/\partial t^2$，因此(2.9)式括号中的第二项可化为

$$\frac{1}{c^2}\frac{\ddot{\boldsymbol{p}}_e(t-R/c)}{R}=\nabla^2\left[\frac{\boldsymbol{p}_e(t-R/c)}{R}\right].$$

代入(2.9)式，则 $\boldsymbol{E}^{(1)}$ 变为

$$\begin{aligned}\boldsymbol{E}^{(1)}&=\frac{1}{4\pi\varepsilon_0}\left[\nabla\nabla\cdot\frac{\boldsymbol{p}_e(t-R/c)}{R}-\nabla^2\frac{\boldsymbol{p}_e(t-R/c)}{R}\right]\\&=\frac{1}{4\pi\varepsilon_0}\nabla\times\nabla\times\frac{\boldsymbol{p}_e(t-R/c)}{R}.\end{aligned} \tag{2.10}$$

下面，我们分别讨论近区($R\ll\lambda$)和远区($R\gg\lambda$)中电偶极场的性质. 为了简单起见，设电荷体系的电偶极矩按正弦规律变化，用复数表示即为

$$\boldsymbol{p}_e(t-R/c)=\boldsymbol{p}_0\exp[\mathrm{i}\omega(t-R/c)]. \tag{2.11}$$

不难证明，它对时间及空间的导数分别为

$$\dot{\boldsymbol{p}}_e(t-R/c) = \mathrm{i}\omega \boldsymbol{p}_e(t-R/c), \tag{2.12}$$

$$\ddot{\boldsymbol{p}}_e(t-R/c) = -\omega^2 \boldsymbol{p}_e(t-R/c), \tag{2.13}$$

$$\nabla \times \boldsymbol{p}_e(t-R/c) = -\mathrm{i}(\omega/c)\boldsymbol{e}_R \times \boldsymbol{p}_e(t-R/c), \tag{2.14}$$

$$\nabla \cdot \boldsymbol{p}_e(t-R/c) = -\mathrm{i}(\omega/c)\boldsymbol{e}_R \cdot \boldsymbol{p}_e(t-R/c). \tag{2.15}$$

由此可见，在计算 $\boldsymbol{B}^{(1)}$ 的(2.8)式中，对分母求梯度将产生一个 $1/R$ 的因子，而对分子求旋度将产生一个 $\omega/c = 2\pi/\lambda$ 因子. 因此，对子近区($R \ll \lambda$)，分子的旋度是小项，只要保留分母的梯度项；对于远区($R \gg \lambda$)，情况刚好相反，只要保留分子的旋度项；对于过渡区($R \sim \lambda$)，则两项均应保留.

1. 近区场.

$$\boldsymbol{B}^{(1)} = \frac{\mu_0}{4\pi}\left(\nabla \frac{1}{R}\right) \times \dot{\boldsymbol{p}}_e\left(t-\frac{R}{c}\right) = -\frac{\mu_0}{4\pi R^2}\boldsymbol{e}_R \times \dot{\boldsymbol{p}}_e\left(t-\frac{R}{c}\right). \tag{2.16}$$

按(2.10)式计算近区场时，仿照上述只保留 $1/R$ 的空间导数，而略去 $\boldsymbol{p}_e(t-R/c)$ 的空间导数，按∇算符运算公式可得

$$\boldsymbol{E}^{(1)} = \nabla \times \nabla \times \frac{\boldsymbol{p}_e(t-R/c)}{4\pi\varepsilon_0 R} = \frac{1}{4\pi\varepsilon_0}\nabla \times \left[\left(\nabla \frac{1}{R}\right) \times \boldsymbol{p}_e\left(t-\frac{R}{c}\right)\right]$$

$$= \frac{1}{4\pi\varepsilon_0}\left[\boldsymbol{p}_e\left(t-\frac{R}{c}\right)\cdot \nabla\left(\nabla \frac{1}{R}\right) - \nabla \cdot \left(\nabla \frac{1}{R}\right)\boldsymbol{p}_e\left(t-\frac{R}{c}\right)\right].$$

利用 $\boldsymbol{R} \neq 0$ 处 $\nabla^2(1/R) = \delta^3(\boldsymbol{R}) = 0$ 及 $\nabla\nabla(1/R) = (3\boldsymbol{e}_R\boldsymbol{e}_R - \overleftrightarrow{e})/R^3$，可得

$$\boldsymbol{E}^{(1)} = \frac{1}{4\pi\varepsilon_0 R^3}\left\{3\boldsymbol{e}_R\left[\boldsymbol{e}_R \cdot \boldsymbol{p}_e\left(t-\frac{R}{c}\right)\right] - \boldsymbol{p}_e\left(t-\frac{R}{c}\right)\right\}. \tag{2.17}$$

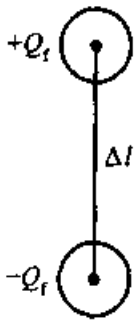


图1　振荡的电偶极系统

为了更直观地了解近区场的性质，考察图1所示的电偶极子系统. 它由两个相距为 $\Delta\boldsymbol{l}$ 的导体球组成，两导体球之间由细导线相连. 当导线上有交变电流 I_f 时，两导体球上的电荷将交替地变化，形成一个振荡的电偶极子. 这个电荷体系的电偶极矩即为 $\boldsymbol{p}_e = Q_f\Delta\boldsymbol{l}$. 因为 Δl 很短，同一时刻 Δl 上各点的 I_f 可以看作相同，故 $\dot{\boldsymbol{p}}_e = I_f\Delta\boldsymbol{l}$. 代入(2.16)式，便有

$$\boldsymbol{B}^{(1)} = \frac{\mu_0 I_f(t-R/c)\,\Delta\boldsymbol{l} \times \boldsymbol{R}}{4\pi R^3}. \tag{2.18}$$

这和静磁学中的毕奥-萨伐尔定律形式上一致. 而(2.17)式也与静电学中电偶极子的电场公式形式上一致. (2.17)与(2.18)式与静场公式的差异仅在子电偶极矩是随时间变化的，公式中的 $\boldsymbol{p}_e$ 与 $\boldsymbol{I}_f$ 均应取 $t-R/c$ 时的值，即应考虑推迟效应. 由子近区场与静场公式的相似性，所以近区也称为准静区.

2. 远区场. 采用上述近似及(2.13)、(2.14)式可得

$$\boldsymbol{B}^{(1)} = \frac{\mu_0 \nabla \times \dot{\boldsymbol{p}}_e(t - R/c)}{4\pi R} = -\frac{\mathrm{i}\mu_0\omega}{4\pi cR}\boldsymbol{e}_R \times \dot{\boldsymbol{p}}_e(t - R/c)$$

$$= \frac{-\mu_0}{4\pi cR}\boldsymbol{e}_R \times \ddot{\boldsymbol{p}}_e(t - R/c), \tag{2.19}$$

$$\boldsymbol{E}^{(1)} = \frac{\nabla \times \nabla \times \boldsymbol{p}_e(t - R/c)}{4\pi\varepsilon_0 R} = \frac{\boldsymbol{e}_R \times [\boldsymbol{e}_R \times \ddot{\boldsymbol{p}}_e(t - R/c)]}{4\pi\varepsilon_0 c^2 R}$$

$$= c\boldsymbol{B}^{(1)} \times \boldsymbol{e}_R. \tag{2.20}$$

由此可见,在远区的 $\boldsymbol{E}^{(1)}$ 与 $\boldsymbol{B}^{(1)}$ 均与 R 成反比,能流与 R^2 成反比,对球面积分后其总功率与球半径无关,因此在没有吸收的情况下,电磁能量可以传播到任意远处.

(三)电偶极辐射的角分布与总辐射功率

电偶极辐射场 $\boldsymbol{B}^{(1)}$,$\boldsymbol{E}^{(1)}$ 已由(2.19)、(2.20)式给出. 由此可求得能流密度 $\boldsymbol{S} = \boldsymbol{E}^{(1)} \times \boldsymbol{H}^{(1)}$,而 S 与 θ,ϕ 的关系即给出辐射的角分布. 再对一封闭球面积分即得总辐射功率.

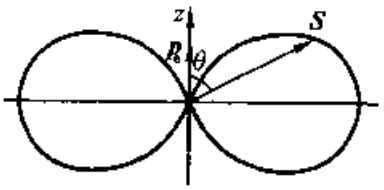


图 2　电偶极辐射的角分布

采用球坐标系计算. 坐标原点取在电荷体系电偶极矩的中心,z 轴沿 $\boldsymbol{p}_e$ 的方向. 利用 $-\boldsymbol{e}_R \times \boldsymbol{e}_z = \sin\theta\boldsymbol{e}_\phi$ 及 $\boldsymbol{e}_\phi \times \boldsymbol{e}_R = \boldsymbol{e}_\theta$,则远处的辐射场可表示为

$$\boldsymbol{B}^{(1)} = \frac{\mu_0 \ddot{p}_e(t - R/c)}{4\pi cR}\sin\theta\boldsymbol{e}_\phi, \tag{2.21}$$

$$\boldsymbol{E}^{(1)} = c\boldsymbol{B}^{(1)} \times \boldsymbol{e}_R = \frac{\ddot{p}_e(t - R/c)}{4\pi\varepsilon_0 c^2 R}\sin\theta\boldsymbol{e}_\theta, \tag{2.22}$$

可见 $\boldsymbol{E}^{(1)}$,$\boldsymbol{B}^{(1)}$ 与 $\boldsymbol{e}_R$ 三者构成右手螺旋系统.

远区的电磁能流密度为

$$\boldsymbol{S}^{(1)} = \boldsymbol{E}^{(1)} \times \boldsymbol{H}^{(1)} = c(\boldsymbol{B}^{(1)} \times \boldsymbol{e}_R) \times \boldsymbol{B}^{(1)}/\mu_0 = \frac{[\ddot{p}_{ee}(t - R/c)]^2}{16\pi^2\varepsilon_0 c^3 R^2}\sin^2\theta\boldsymbol{e}_R, \tag{2.23}$$

这说明 $\boldsymbol{S}$ 与 $\mathbf{e}_R$ 方向一致,即电磁能量是沿位矢方向向外传播的.

在(2.21)及(2.22)式中取 $\ddot{p}_e(t - R/c)$ 的实部代入(2.23)式,得到

$$\boldsymbol{S}^{(1)} = \frac{\omega^4 p_0^2 \sin^2\theta}{16\pi^2\varepsilon_0 c^3 R^2}\cos^2[\omega(t - R/c)]\boldsymbol{e}_R. \tag{2.24}$$

取一周期内的平均值即得平均能流密度

$$\overline{\boldsymbol{S}^{(1)}} = \frac{1}{T}\int_0^T \boldsymbol{S}\mathrm{d}t = \frac{\omega^4 p_0^2}{32\pi^2\varepsilon_0 c^3 R^2}\sin^2\theta\mathbf{e}_R. \tag{2.25}$$

这说明平均能流密度 $S \propto \sin^2\theta$，即在 $\theta=0,\pi$ 方向（沿偶极子方向）没有辐射，而在 $\theta=90°$ 的平面上（垂直于偶极子方向）辐射最强，如图 2.

将 $\overline{S^{(1)}}$ 对整个球面积分即得总辐射功率

$$\begin{aligned} P &= \oint_{\Gamma} \overline{S^{(1)}} R^2 \mathrm{d}\Omega \\ &= \int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \frac{\omega^4 p_0^2 \sin^2\theta}{32\pi^2 \varepsilon_0 c^3 R^2} R^2 \sin\theta \mathrm{d}\theta \\ &= \frac{\omega^4 p_0^2}{12\pi\varepsilon_0 c^3}. \end{aligned} \tag{2.26}$$

由此可见，若电偶极矩振幅 p_0 保持不变，则总辐射功率正比于频率的四次方，说明频率越高时总辐射功率越大.

习　题

1. 试写出规范变换的洛伦兹条件与达朗伯方程的四维形式.

2. 试分别求出采用洛伦兹规范和库仑规范时单色平面电磁波的矢势和标势，并指出这时 $\boldsymbol{A}$ 与 φ 是否唯一的，为什么？

提示　电磁场只依赖于 $\boldsymbol{A}$ 的横向分量. 采用洛伦兹规范时 $\boldsymbol{B}=\mathrm{i}\boldsymbol{k}\times\boldsymbol{A}$ 及 $\boldsymbol{E}=-c\boldsymbol{e}_n\times\boldsymbol{B}$；采用库仑规范时取 $\varphi=0$，可证 $\boldsymbol{A}=\boldsymbol{A}_0 e^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}$，再利用库仑条件 $\nabla\cdot\boldsymbol{A}=0$.

3. 证明 §4.1.1(1.20)式是该节方程(1.11)的解.

提示　要证明：①当 $x\neq 0$ 时，(1.20)式满足方程(1.11)；②因 $R=0$ 是(1.20)式的奇点，为了考察方程(1.11)在奇点的性质，只要将(1.20)式代入(1.11)式后，两边对以原点为心，半径 $\varepsilon\to 0$ 的球体 V_ε 进行积分，利用 δ 函数的定义，如果等式成立即得证.

4. 利用电荷守恒定律，验证推迟势 $\boldsymbol{A}$ 和 φ 满足洛伦兹条件 §4.1.1(1.8)式.

提示　应注意：①在计算 $\nabla'\cdot\boldsymbol{J}_t(x',t')$ 时，$\boldsymbol{J}_t(x',t')$ 不仅直接依赖于 x'，而且要通过中间变量 $t'=t-r/c=t-|x-x'|/c$ 依赖于 x'，故

$$\begin{aligned} \nabla'\cdot\boldsymbol{J}_t(x',t') &= \sum_{i=1}^{3} \frac{J_{ti}(x',t')}{\partial x'_i} \\ &= \sum_{i=1}^{3} \left[\frac{\partial J_{ti}(x',t')}{\partial x'_i}\bigg|_{t'=\text{常数}} + \frac{\partial J_{ti}(x',t')}{\partial t'}\cdot\frac{\partial t'}{\partial x'_i} \right] \\ &= \nabla'\cdot\boldsymbol{J}_t(x',t')\big|_{t'=\text{常数}} + \frac{\partial \boldsymbol{J}_t(x',t')}{\partial t'}\cdot\nabla t'. \end{aligned}$$

②电荷守恒律为 $\nabla'\cdot\boldsymbol{J}_t(x',t')\big|_{t'=\text{常数}} + \partial\rho(x',t')/\partial t'=0$，因为电荷守恒律是说明 t' 时刻 x' 处 $\boldsymbol{J}_t$ 的散度等于同一时刻同一点处 ρ_t 随时间的减小率，这与源点 x' 到场点 x 的距离 r 无关，故这里计算 $\boldsymbol{J}_t(x',t')$ 的散度时应把 t' 看作常数，而不需要通过中间变量 t' 求对 x' 的偏导数.

5. 试求氢原子中电子以不变的角频率 ω 沿半径为 a 的圆轨道绕核旋转时的辐射电磁场及平均能流.

6. 设有线偏振平面波 $\boldsymbol{E}=\boldsymbol{E}_0 e^{i(kz-\omega t)}$ 照射到一个绝缘介质球上（$\boldsymbol{E}_0$ 在 z 方向），引起介质极化. 极化矢量 $\boldsymbol{P}$ 是随时间变化的，因而产生辐射. 设平面波的波长 $2\pi/k$ 远大于球半径 R_0，求介质球所产生的辐射场和能流.

7. 设有一振辐为 p_0，频率为 ω 的电偶极子位于距离理想导体平面为 $a/2$ 处，p_0 平行于导体平面. 设 $a\ll\lambda$，求在 $R\gg\lambda$ 处的电磁场及辐射能流.

第二章　运动带电粒子与电磁场的相互作用

上一章讨论了宏观电荷体系辐射的电磁场，下面考虑运动带电粒子与电磁场的相互作用. 本章第一节给出运动带电粒子所激发的电磁场，其中包括自有场和辐射场. 第二节和第三节讨论几种特殊情况下带电粒子的辐射——韧致辐射、同步辐射和切连科夫辐射. 第四节讨论带电粒子的辐射场对粒子本身的反作用. 第五节讨论由于和带电粒子的相互作用引起的电磁波的散射、吸收和色散.

§4.2.1　运动带电粒子的势和辐射的电磁场

对子一个作任意运动的带电粒子激发的势和辐射的电磁场，通常有两种讨论方法：第一种方法是给出粒子电荷分布的假设，例如，把粒子看作一个均匀的带电球，然后利用推迟势公式进行计算：

$$\varphi(\boldsymbol{x},t) = \frac{1}{4\pi\varepsilon_0}\int\frac{\rho_f(\boldsymbol{x}',t-r/c)}{r}\mathrm{d}V', \tag{1.1}$$

$$\boldsymbol{A}(\boldsymbol{x},t) = \frac{\mu_0}{4\pi}\int\frac{\boldsymbol{J}_f(\boldsymbol{x}',t-r/c)}{r}\mathrm{d}V'. \tag{1.2}$$

第二种方法是利用势的洛伦兹变换. 如果能求得在相对粒子静止的参考系 K_0 中的势，通过洛伦兹变换就容易求得实验室参考系 K 中的势. 当然，这里有一个问题是，当粒子作任意运动时，固连在粒子上的参考系 K_0 是非惯性系，势的洛伦兹变换不能成立. 那么，能不能在每一时刻引入一个与粒子瞬间静止的惯性系 $\tilde{\mathrm{K}}$ 呢？这里又出现一个问题：粒子在 $\tilde{\mathrm{K}}$ 系中仅瞬时静止，它在 $\tilde{\mathrm{K}}$ 系激发的势是否与静电势相同呢？这个问题的答案可由推迟势公式(1.1)，(1.2)及运动带电粒子的 $\boldsymbol{J}_f=\rho_f\boldsymbol{v}$ 看出. 粒子产生的势只与粒子的速度有关，而与粒子的加速度无关，所以粒子在 $\tilde{\mathrm{K}}$ 系激发的势就是静电势. 再利用势的洛伦兹变换即可求得在实验室参考系的势. 由于每一时刻要引入一个 $\tilde{\mathrm{K}}$ 系，所以在一段观察时间内就有无穷多个 $\tilde{\mathrm{K}}$ 系，它们分别在不同瞬时与粒子相对静止，随后即以粒子在那一瞬时的速度继续作匀速直线运动而与粒子分道扬镳. 下面我们就按这种方法来求势.

(一)李纳-维谢尔势

设在实验室系 K 中，研究 t 时刻场点 $\boldsymbol{x}$ 的势，运动带电粒子的位矢为 $\boldsymbol{x}_e(t)$，称为现在位置. t 时刻 P 点的势是 $t'=t-r/c$ 时刻粒子在 $\boldsymbol{x}_e(t')$ 激发的(如图 1)，称 $\boldsymbol{x}_e(t')$ 为推迟位置，并令 $\boldsymbol{r}=\boldsymbol{x}-\boldsymbol{x}_e(t')$. 在 t' 时刻与粒子瞬时静止的惯性参考系 $\tilde{\mathrm{K}}$ 中，带电粒子激发的标势和矢势分别为：

$$\begin{cases}\tilde{\varphi} = e/(4\pi\varepsilon_0\tilde{r}), & \text{其中 } \tilde{r} = c(\tilde{t}-\tilde{t}'), \quad (1.3)\\ \tilde{\boldsymbol{A}} = 0. & (1.4)\end{cases}$$

式中 $\tilde{r}$ 是 $\tilde{\mathrm{K}}$ 系测得的粒子推迟位置与场点的距离.

利用 K 系与 $\tilde{\mathrm{K}}$ 系之间势的洛伦兹变换式可得:

$$\begin{cases}A_x = \gamma[\tilde{A}_x + v\tilde{\varphi}/c^2] = \gamma v\tilde{\varphi}/c^2, & (1.5)\\ A_y = \tilde{A}_y = 0, & (1.6)\\ A_z = \tilde{A}_z = 0, & (1.7)\\ \varphi = \gamma[\tilde{\varphi} + v\tilde{A}_x] = \gamma\tilde{\varphi}. & (1.8)\end{cases}$$

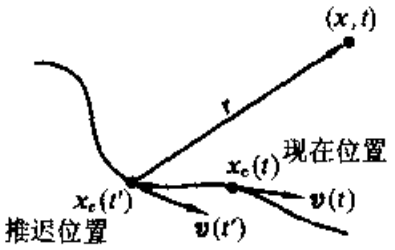


图 1　粒子的现在位置、推迟位置与场点

将矢势写成矢量形式,并将 $\tilde{\varphi}$ 代入(1.5)式和(1.8)式得到

$$\boldsymbol{A} = \sum_i A_i\boldsymbol{e}_i = \frac{\gamma v\tilde{\varphi}}{c^2}\boldsymbol{e}_x = \frac{\gamma e\boldsymbol{v}}{4\pi\varepsilon_0 c^2\tilde{r}}, \quad (1.9)$$

$$\varphi = \gamma\tilde{\varphi} = \frac{\gamma e}{4\pi\varepsilon_0\tilde{r}}. \quad (1.10)$$

现在将 $\tilde{r}$ 用 K 系的距离 r 表示出来,有

$$\tilde{r} = |\tilde{\boldsymbol{x}} - \tilde{\boldsymbol{x}}_e(t')| = c(\tilde{t}-\tilde{t}'), \quad (1.11)$$

及洛伦兹变换式

$$\begin{cases}\tilde{t} = \gamma(t - vx/c^2), & (1.12)\\ \tilde{t}' = \gamma(t' - vx_e'/c^2). & (1.13)\end{cases}$$

注意,我们观察的是 t 时刻 $\boldsymbol{x}$ 点的场,(1.12)式是这一事件的洛伦兹变换式;(1.13)式是相应的推迟时刻 t' 和推迟位置 $\boldsymbol{x}_e' = \boldsymbol{x}_e(t')$ 的洛伦兹变换式;$\boldsymbol{v}$ 是 t' 时刻 $\tilde{\mathrm{K}}$ 系相对 K 系的速度,亦即粒子在推迟位置的速度 $\boldsymbol{v}(t')$. 将(1.12)和(1.13)式代入(1.11)式,即有:

$$\tilde{r} = c\gamma[(t-t') - v(x-x_e')/c^2] = \gamma(r - \boldsymbol{v}\cdot\boldsymbol{r}/c). \quad (1.14)$$

将(1.14)式代入(1.9)、(1.10)式,即有

$$\boldsymbol{A} = \frac{e\boldsymbol{v}}{4\pi\varepsilon_0 c^2(r - \boldsymbol{v}\cdot\boldsymbol{r}/c)}, \quad (1.15)$$

$$\varphi = \frac{e}{4\pi\varepsilon_0(r - \boldsymbol{v}\cdot\boldsymbol{r}/c)}. \quad (1.16)$$

(1.15)与(1.16)式称为李纳 - 维谢尔势.

(二)运动带电粒子的场

将(1.15)、(1.16)式代入 $\boldsymbol{B} = \nabla\times\boldsymbol{A}$ 及 $\boldsymbol{E} = -\nabla\varphi - \partial\boldsymbol{A}/\partial t$ 即可求得实验室参考系 K 中的电磁场. 困难在于 $\boldsymbol{A}$ 与 φ 中的 r 及 $\boldsymbol{v}$ 均为 t' 的函数,而

$$t' = t - r/c = t - |\boldsymbol{x} - \boldsymbol{x}_e(t')|/c$$

是 x 的函数. 因此,在计算 E 与 B 之前,要先算出 $\partial t'/\partial t$ 及 $\nabla t'$. 由

$$\frac{\partial t'}{\partial t} = \frac{\partial}{\partial t}\left(t - \frac{r}{c}\right) = 1 - \frac{1}{c}\frac{\partial r(t')}{\partial t'}\frac{\partial t'}{\partial t}, \tag{1.17}$$

在 $r^2(t') = [x - x_e(t')]^2$ 两边对 t' 求导数可得

$$2r\frac{\partial r(t')}{\partial t'} = 2[x - x_e(t')] \cdot \left[-\frac{\mathrm{d}x_e(t')}{\mathrm{d}t'}\right] = -2r \cdot v(t').$$

两边除以 $2r$ 可得

$$\partial r(t')/\partial t' = -e_n \cdot v(t'). \tag{1.18}$$

其中 $e_n = r/r$ 是沿粒子到场点方向的单位矢量. 将(1.18)式代入(1.17)式后整理可得:

$$\frac{\partial t'}{\partial t} = \frac{1}{1 - e_n \cdot v/c}. \tag{1.19}$$

为计算 $\nabla t'$ 可以利用复合函数的偏导数公式及(1.18)式,得到

$$\begin{aligned}
\nabla t' &= \nabla\left(t - \frac{r}{c}\right) = -\frac{\nabla r}{c} = -\frac{1}{c}\sum_{i=1}^{3} e_i \frac{\partial r}{\partial x_i} \\
&= -\frac{1}{c}\sum e_i\left(\left.\frac{\partial r}{\partial x_i}\right|_{t'\text{不变}} + \frac{\partial t'}{\partial x_i}\frac{\partial r}{\partial t'}\right) \\
&= -\frac{1}{c}\left(\left.\nabla r\right|_{t'\text{不变}} + \nabla t'\frac{\partial r}{\partial t'}\right) \\
&= -\frac{e_n}{c} + \frac{e_n \cdot v}{c}\nabla t',
\end{aligned}$$

因而,

$$\nabla t' = \frac{-e_n}{(c - e_n \cdot v)}. \tag{1.20}$$

当粒子的速度 $v \ll c$ 时,(1.19)与(1.20)式简化为

$$\partial t'/\partial t = 1, \tag{1.21}$$

$$\nabla t' = -e_n/c. \tag{1.22}$$

现在计算 $v \ll c$ 条件下的辐射场,此时矢势 A 可简化为

$$A = \frac{ev(t')}{4\pi\varepsilon_0 c^2 r}, \qquad v \ll c. \tag{1.23}$$

仿照(1.20)式的推导方法可得

$$\begin{aligned}
B &= \nabla \times A = \left.\nabla \times A\right|_{t'\text{不变}} + \nabla t' \times (\partial A/\partial t') \\
&= \left.\nabla \times \frac{ev(t')}{4\pi\varepsilon_0 c^2 r}\right|_{t'\text{不变}} - \frac{e_n}{c} \times \frac{\partial}{\partial t'}\frac{ev(t')}{4\pi\varepsilon_0 c^2 r}.
\end{aligned}$$

$$=\frac{e}{4\pi\varepsilon_0 c^2}\left(\nabla\frac{1}{r}\right)\times\boldsymbol{v}-\frac{e\boldsymbol{e}_n}{4\pi\varepsilon_0 c^3}\times\left(\frac{\dot{\boldsymbol{v}}}{r}-\frac{\dot{r}\boldsymbol{v}}{r^2}\right)$$

$$\approx\frac{e\boldsymbol{v}\times\boldsymbol{r}}{4\pi\varepsilon_0 c^2 r^3}+\frac{e\dot{\boldsymbol{v}}\times\boldsymbol{r}}{4\pi\varepsilon_0 c^3 r^2},\qquad v\ll c.\ . \tag{1.24}$$

剩下一项含 $\dot{r}\boldsymbol{v}/c^2$，当 $v\ll c$ 时，可略去.

采用类似的方法，不难证明：

$$\boldsymbol{E}=\frac{e\boldsymbol{r}}{4\pi\varepsilon_0 r^3}+\frac{e}{4\pi\varepsilon_0 c^2 r^3}[\boldsymbol{r}\times(\boldsymbol{r}\times\dot{\boldsymbol{v}})],\qquad v\ll c. \tag{1.25}$$

(1.24)式与(1.25)式表明，$\boldsymbol{B}$ 与 $\boldsymbol{E}$ 的第一项与 r^2 成反比，仅存在于粒子附近，r 大处它很小而可略去，故这个场称为自有场；而 $\boldsymbol{B}$ 与 $\boldsymbol{E}$ 的第二项均与 r 成反比，即其能流与 r 无关，可见这部分场可以脱离电荷电流而独立存在，其能量可辐射到任意远处，故称为辐射场.

带电粒子的电偶极矩可以表示为 $\boldsymbol{p}_e\approx e\boldsymbol{x}_e$（参看 §3.1.5），因此 $\dot{\boldsymbol{p}}_e=e\boldsymbol{v}$，$\ddot{\boldsymbol{p}}_e=e\dot{\boldsymbol{v}}$. 这里电偶极矩本身与坐标原点选择有关，但决定电磁场的 $\ddot{\boldsymbol{p}}_e=e\dot{\boldsymbol{v}}$ 则与坐标原点的选择无关. 将 $\ddot{\boldsymbol{p}}_e$ 代入(1.24)与(1.25)式，得辐射场部分为

$$\boldsymbol{B}=\frac{\ddot{\boldsymbol{p}}_e\times\boldsymbol{e}_n}{4\pi\varepsilon_0 c^3 r}, \tag{1.26}$$

$$\boldsymbol{E}=\frac{\boldsymbol{e}_n\times(\boldsymbol{e}_n\times\ddot{\boldsymbol{p}}_e)}{4\pi\varepsilon_0 c^2 r}=c(\boldsymbol{B}\times\boldsymbol{e}_n). \tag{1.27}$$

可见，这与 §4.1.2 的(2.19)和(2.20)式给出的电偶极辐射公式一致. 表面的差异是用一个运动带电粒子的 $\ddot{\boldsymbol{p}}_e$ 替代整个电荷体系的 $\ddot{\boldsymbol{p}}_e$ 以及用 r 代替 R. 后者是因为电荷体系的 $\boldsymbol{p}_e$ 位于坐标原点，故 $\boldsymbol{p}_e$ 到场点的距离 $r=|\boldsymbol{x}-0|=R$. 这说明具有加速度的低速运动粒子（$\dot{\boldsymbol{v}}\neq0$，$v\ll c$）所辐射的电磁场就是电偶极辐射场，故它的辐射称为电偶极辐射，而 $v\ll c$ 称为电偶极辐射条件.

不难证明，辐射能流为

$$\boldsymbol{S}=\boldsymbol{E}\times\boldsymbol{H}=c(\boldsymbol{B}\times\boldsymbol{e}_n)\times\boldsymbol{B}/\mu_0=cB^2\boldsymbol{e}_n/\mu_0=\frac{e^2\dot{\boldsymbol{v}}^2\sin^2\angle(\boldsymbol{r},\dot{\boldsymbol{v}})}{16\pi^2\varepsilon_0 c^3 r^2}\boldsymbol{e}_n. \tag{1.28}$$

它说明，在与 $\dot{\boldsymbol{v}}$ 的垂直方向上，因为 $\sin^2\angle(\boldsymbol{r},\dot{\boldsymbol{v}})=1$，其辐射最强.

计算总辐射功率时可取瞬时加速度 $\dot{\boldsymbol{v}}$ 的方向为极轴，则 $\boldsymbol{r}$，$\dot{\boldsymbol{v}}$ 的夹角等于 θ，不难证明

$$P=\oint\boldsymbol{S}\cdot\mathrm{d}\boldsymbol{\sigma}=\frac{e^2\dot{\boldsymbol{v}}^2}{6\pi\varepsilon_0 c^3}. \tag{1.29}$$

这说明，总辐射功率正比于粒子加速度的平方，若粒子在真空中作匀速直线运动，无论速度有多大，均无辐射.

§4.2.2　韧致辐射与同步辐射

在电子加速器中，电子的速度接近光速，作加速运动的电子将会发生辐射. 在直线加速器中，电子沿直线运动，其 $\dot{\boldsymbol{v}}/\!/\boldsymbol{v}$，所产生的辐射称为韧致辐射；在同步加速器中，电子作匀速圆周运动，其 $\dot{\boldsymbol{v}}\perp\boldsymbol{v}$，所产生的辐射称为同步辐射. 本节先导出加速带电粒子辐射出的电磁场，给出辐射功率与角分布，然后分别讨论韧致辐射与同步辐射.

(一)加速带电粒子辐射的电磁场

下面由李纳－维谢尔势出发计算运动带电粒子辐射的电磁场. 为方便起见，引入

$$s = r - \boldsymbol{v}\cdot\boldsymbol{r}/c, \tag{2.1}$$

并利用 s 表示上节的(1.19)、(1.20)、(1.15)和(1.16)式如下：

$$\partial t'/\partial t = r/s, \tag{2.2}$$

$$\nabla t' = -\boldsymbol{r}/(cs), \tag{2.3}$$

$$\boldsymbol{A} = e\boldsymbol{v}/(4\pi\varepsilon_0 c^2 s), \tag{2.4}$$

$$\varphi = e/(4\pi\varepsilon_0 s). \tag{2.5}$$

同时，利用上节的(1.18)式可得

$$\frac{\partial s}{\partial t'} = \frac{\partial}{\partial t'}\left(r - \frac{\boldsymbol{v}\cdot\boldsymbol{r}}{c}\right) = \dot{r} - \frac{\dot{\boldsymbol{v}}\cdot\boldsymbol{r}}{c} - \frac{\boldsymbol{v}\cdot\dot{\boldsymbol{r}}}{c} = -\frac{\boldsymbol{r}\cdot\boldsymbol{v}}{r} - \frac{\dot{\boldsymbol{v}}\cdot\boldsymbol{r}}{c} + \frac{v^2}{c}, \tag{2.6}$$

$$\nabla\frac{1}{s}\bigg|_{t'\text{不变}} = -\frac{1}{s^2}\nabla s\bigg|_{t'\text{不变}} = -\frac{1}{s^2}\nabla\left(r - \frac{\boldsymbol{v}\cdot\boldsymbol{r}}{c}\right) = -\frac{1}{s^2}\left(\frac{\boldsymbol{r}}{r} - \frac{\boldsymbol{v}}{c}\right). \tag{2.7}$$

由于辐射场与 $1/r$ 成正比，且当 $v\approx c$ 时有 $s\approx r$，在计算加速带电粒子辐射的电磁场时仅保留 $1/r$ 及 $1/s$ 的项，而略去它们的高次幂项.

现分别计算 $\boldsymbol{E} = -\nabla\varphi - \partial\boldsymbol{A}/\partial t$ 的第一、二项.

$$\begin{aligned}\nabla\varphi &= \nabla\varphi\big|_{t'\text{不变}} + \nabla t'(\partial\varphi/\partial t')\\ &= \frac{e}{4\pi\varepsilon_0}\nabla\frac{1}{s}\bigg|_{t'\text{不变}} - \frac{\boldsymbol{r}}{cs}\frac{e}{4\pi\varepsilon_0}\frac{\partial}{\partial t'}\left(\frac{1}{s}\right)\\ &= -\frac{e}{4\pi\varepsilon_0 s^2}\left(\frac{\boldsymbol{r}}{r} - \frac{\boldsymbol{v}}{c}\right) - \frac{e\boldsymbol{r}}{4\pi\varepsilon_0 cs^3}\left(\frac{\boldsymbol{r}\boldsymbol{v}}{r} + \frac{\dot{\boldsymbol{v}}\cdot\boldsymbol{r}}{c} - \frac{v^2}{c}\right),\end{aligned}$$

上式除第四项外均含 $1/r$ 的高次幂，故仅考虑第四项，即

$$\nabla\varphi = -\frac{e}{4\pi\varepsilon_0 c^2 s^3}(\dot{\boldsymbol{v}}\cdot\boldsymbol{r})\boldsymbol{r}. \tag{2.8}$$

同理

$$\frac{\partial \boldsymbol{A}}{\partial t}=\frac{\partial \boldsymbol{A}}{\partial t'}\frac{\partial t'}{\partial t}=\frac{e\boldsymbol{r}}{4\pi\varepsilon_0 c^2 s}\frac{\partial}{\partial t'}\left(\frac{\boldsymbol{v}}{s}\right)=\frac{e\boldsymbol{r}}{4\pi\varepsilon_0 c^2 s}\left(\frac{\dot{\boldsymbol{v}}}{s}-\frac{\dot{s}}{s^2}\boldsymbol{v}\right)$$

$$=\frac{e\boldsymbol{r}}{4\pi\varepsilon_0 c^2 s^3}\left[s\dot{\boldsymbol{v}}+\left(\frac{r\boldsymbol{v}}{r}+\frac{\dot{\boldsymbol{v}}\cdot\boldsymbol{r}}{c}-\frac{v^2}{c}\right)\boldsymbol{v}\right],$$

上式除第一、三项外均含 $1/r$ 的高次幂，故仅考虑第一、三项，即

$$\frac{\partial \boldsymbol{A}}{\partial t}=\frac{e\boldsymbol{r}}{4\pi\varepsilon_0 c^2 s^3}\left[s\dot{\boldsymbol{v}}+\frac{\dot{\boldsymbol{v}}\cdot\boldsymbol{r}}{c}\boldsymbol{v}\right]. \tag{2.9}$$

将(2.8)、(2.9)式代入 $\boldsymbol{E}=-\nabla\varphi-\partial\boldsymbol{A}/\partial t$ 即得加速带电粒子的辐射场为

$$\boldsymbol{E}=\frac{e}{4\pi\varepsilon_0 c^2 s^3}\left[(\dot{\boldsymbol{v}}\cdot\boldsymbol{r})\boldsymbol{r}-r\left(r-\frac{\boldsymbol{v}\cdot\boldsymbol{r}}{c}\right)\dot{\boldsymbol{v}}-\frac{r(\dot{\boldsymbol{v}}\cdot\boldsymbol{r})}{c}\boldsymbol{v}\right]$$

$$=\frac{e}{4\pi\varepsilon_0 c^2 s^3}\boldsymbol{r}\times\left[\left(\boldsymbol{r}-\frac{r\boldsymbol{v}}{c}\right)\times\dot{\boldsymbol{v}}\right]$$

$$=\frac{e}{4\pi\varepsilon_0 c^2 r}\frac{\boldsymbol{e}_n\times\left[\left(\boldsymbol{e}_n-\frac{\boldsymbol{v}}{c}\right)\times\dot{\boldsymbol{v}}\right]}{\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{e}_n}{c}\right)^3}. \tag{2.10}$$

这里已将 $\boldsymbol{e}_n=\boldsymbol{r}/r$ 以及 $s=r-\boldsymbol{v}\cdot\boldsymbol{r}/c=r(1-\boldsymbol{v}\cdot\boldsymbol{e}_n/c)$ 代入. 同理可得

$$\boldsymbol{B}=\nabla\times\boldsymbol{A}=\nabla\times\boldsymbol{A}\big|_{t'\text{不变}}+\nabla t'\times\frac{\partial \boldsymbol{A}}{\partial t'}, \tag{2.11}$$

仿上可证第一项含 $1/r$ 的高次幂，第二项利用

$$\frac{\partial \boldsymbol{A}}{\partial t}=\frac{\partial \boldsymbol{A}}{\partial t'}\frac{\partial t'}{\partial t}=\frac{r}{s}\frac{\partial \boldsymbol{A}}{\partial t'},$$

得：

$$\frac{\partial \boldsymbol{A}}{\partial t'}=\frac{s}{r}\frac{\partial \boldsymbol{A}}{\partial t},$$

因而(2.11)式可写成

$$\boldsymbol{B}=\nabla t'\times\frac{\partial \boldsymbol{A}}{\partial t'}=-\frac{\boldsymbol{r}}{cs}\times\frac{s}{r}\frac{\partial \boldsymbol{A}}{\partial t}=-\frac{\boldsymbol{r}}{cr}\times\frac{\partial \boldsymbol{A}}{\partial t}. \tag{2.12}$$

由(2.8)式可见 $\boldsymbol{r}\times\nabla\varphi=0$，故可加在(2.12)式上，即有

$$\boldsymbol{B}=\frac{\boldsymbol{r}}{cr}\times\left(-\nabla\varphi-\frac{\partial \boldsymbol{A}}{\partial t}\right)=\frac{\boldsymbol{r}}{cr}\times\boldsymbol{E}=\frac{1}{c}\boldsymbol{e}_n\times\boldsymbol{E}. \tag{2.13}$$

其中 $\boldsymbol{e}_n=\boldsymbol{r}/r$ 为沿 $\boldsymbol{r}$ 方向的单位矢量. 由(2.10)和(2.13)式可见，加速带电粒子的辐射场有如下两个特点：一是辐射场的 $\boldsymbol{E}$ 和 $\boldsymbol{B}$ 均与加速度 $\dot{\boldsymbol{v}}$ 成正比(与 r 成反比)；二是加速带电粒子辐射的电磁波是横向的，$\boldsymbol{E}$、$\boldsymbol{B}$、$\boldsymbol{e}_n$ 互相垂直，并成右手螺旋关系.

(二)加速带电粒子的辐射功率和角分布

为了计算加速带电粒子的辐射功率,设粒子在 t' 时刻位于 O_1 点,在 $t'+\mathrm{d}t'$ 时刻位于 O_2 点,如图 1. 粒子在初时刻 t' 辐射的场在时刻 $t=t'+R/c$ 将到达以 O_1 为球心 R 为半径的球面 S_1 上;而在末时刻 $t'+\mathrm{d}t'$ 辐射的场在同一时刻 t 将到达以 O_2 为球心 R' 为半径的球面 S' 上. 由 $t=(t'+\mathrm{d}t')+R'/c=t'+R/c$ 得 $R'=R-c\mathrm{d}t'$. 由此可见,粒子在 $t'\sim t'+\mathrm{d}t'$ 的时间内辐射的能量在 t 时刻位于 S_1 和 S_2 这两个球面之间的区域.

现在计算在 $t\sim t+\mathrm{d}t$ 时间内通过 S_1 球面面元 $\mathrm{d}\sigma$ 的电磁能量($\mathrm{d}\sigma$ 在纸面的横截线为 CG,见图 1)

$$\mathrm{d}E = S\,\mathrm{d}t\,\mathrm{d}\sigma. \tag{2.14}$$

式中 S 为能流密度矢量,其大小为单位时间内垂直通过单位横截面的电磁能量. 将能流密度与能量密度的关系 $S=wc$[见 §3.3.1(1.36)式]代入,即有

$$\mathrm{d}E = wc\,\mathrm{d}t\,\mathrm{d}\sigma = w\mathrm{d}r\,\mathrm{d}\sigma, \tag{2.15}$$

其中 $\mathrm{d}r=c\,\mathrm{d}t$ 沿球面 S_1 的径向,见图 1. 利用 $\boldsymbol{e}_n=\boldsymbol{r}/r$,

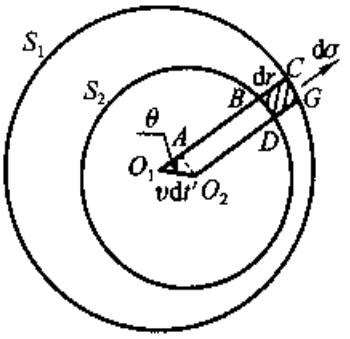


图 1　加速带电粒子的辐射

$\overline{O_1O_2}=\boldsymbol{v}\mathrm{d}t'$,$\overline{O_1C}=R$ 和 $\overline{O_2D}=R-c\mathrm{d}t'$,可得

$$\begin{aligned}\mathrm{d}r &= \overline{BC} = \overline{O_1C}-\overline{O_1A}-\overline{AB} = \overline{O_1C}-\overline{O_1A}-\overline{O_2D}\\ &= R-\boldsymbol{e}_n\cdot\boldsymbol{v}\mathrm{d}t'-(R-c\mathrm{d}t') = c(1-\boldsymbol{v}\cdot\boldsymbol{e}_n/c)\mathrm{d}t'.\end{aligned} \tag{2.16}$$

上式对任意方向的 $\mathrm{d}r$ 成立,在 $\boldsymbol{e}_n$ 与 $\boldsymbol{v}$ 同向处,$\mathrm{d}r$ 较小;而在 $\boldsymbol{e}_n$ 与 $\boldsymbol{v}$ 反向处,$\mathrm{d}r$ 较大. 将上式代入(2.15)式,便有

$$\mathrm{d}E = wc(1-\boldsymbol{v}\cdot\boldsymbol{e}_n/c)\mathrm{d}\sigma\,\mathrm{d}t'. \tag{2.17}$$

由此得粒子运动单位时间内辐射到 $\mathrm{d}\sigma$ 面的电磁能量为

$$\frac{\mathrm{d}E}{\mathrm{d}t'} = wc\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{e}_n}{c}\right)\mathrm{d}\sigma = S\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{e}_n}{c}\right)\mathrm{d}\sigma. \tag{2.18}$$

将上式对球面 S_1 积分即得粒子在其运动的单位时间内辐射的电磁能量,即辐射功率 $P(t')$

$$P(t') = \oint_{S_1} S\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{e}_n}{c}\right)\mathrm{d}\sigma. \tag{2.19}$$

式中的 S 为能流密度 $\boldsymbol{S}$ 的模,由(2.10)和(2.13)式求得

$$\begin{aligned}\boldsymbol{S} &= \boldsymbol{E}\times\boldsymbol{H} = \frac{1}{\mu_0}\boldsymbol{E}\times\boldsymbol{B} = \frac{1}{\mu_0 c}\boldsymbol{E}\times(\boldsymbol{e}_n\times\boldsymbol{E}) = \varepsilon_0 c\boldsymbol{E}^2\boldsymbol{e}_n\\ &= \frac{e^2\boldsymbol{e}_n}{16\pi^2\varepsilon_0 c^3 r^2}\left|\boldsymbol{e}_n\times\left[\left(\boldsymbol{e}_n-\frac{\boldsymbol{v}}{c}\right)\times\dot{\boldsymbol{v}}\right]\right|^2\Big/\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{e}_n}{c}\right)^6.\end{aligned} \tag{2.20}$$

将上式代入(2.19)式,并利用 $\mathrm{d}\Omega=\mathrm{d}\sigma/r^2$,便有

$$P(t') = \frac{e^2}{16\pi^2\varepsilon_0 c^3}\oint \left|\boldsymbol{e}_n \times \left[\left(\boldsymbol{e}_n - \frac{\boldsymbol{v}}{c}\right)\times \dot{\boldsymbol{v}}\right]\right|^2 \Big/ \left(1 - \frac{\boldsymbol{v}\cdot\boldsymbol{e}_n}{c}\right)^5 \mathrm{d}\Omega. \quad (2.21)$$

辐射功率的角分布是指粒子单位时间沿某个方向单位立体角所辐射的能量. 由上式可得

$$\frac{\mathrm{d}P(t')}{\mathrm{d}\Omega} = \frac{e^2}{16\pi^2\varepsilon_0 c^3}\left|\boldsymbol{e}_n \times \left[\left(\boldsymbol{e}_n - \frac{\boldsymbol{v}}{c}\right)\times \dot{\boldsymbol{v}}\right]\right|^2 \Big/ \left(1 - \frac{\boldsymbol{v}\cdot\boldsymbol{e}_n}{c}\right)^5. \quad (2.22)$$

和(2.20)式比较,辐射功率的角分布也可用能流密度 $\boldsymbol{S}$ 表示为

$$\frac{\mathrm{d}P(t')}{\mathrm{d}\Omega} = \boldsymbol{e}_n \cdot \boldsymbol{S}\, r^2\left(1 - \frac{\boldsymbol{v}\cdot\boldsymbol{e}_n}{c}\right). \quad (2.23)$$

取粒子位置为原点,z 轴沿 $\boldsymbol{v}$ 的方向,则 $\boldsymbol{e}_n$ 与 $\boldsymbol{v}$ 的夹角为 θ. 可以证明,当 $v \approx c$ 时,无论加速度沿那个方向,辐射能量将集中在以 $\boldsymbol{v}$ 方向为轴的下述锥角范围内:

$$\Delta\theta \approx \sqrt{1 - v^2/c^2}. \quad (2.24)$$

这个结论可以从(2.22)式的分母看出:当 $v \approx c$ 时,$(1 - \boldsymbol{v}\cdot\boldsymbol{e}_n/c)^{-5} = (1 - v\cos\theta/c)^{-5}$ 在 $\theta \approx 0$ 时很大,故仅需保留不高于二阶的小量,利用 $v \approx c$ 时,

$$\left(1 - \frac{v}{c}\right) \approx \left(1 - \frac{v}{c}\right)\frac{1}{2}\left(1 + \frac{v}{c}\right) = \frac{1}{2}\left(1 - \frac{v^2}{c^2}\right),$$

即有

$$\left(1 - \frac{\boldsymbol{v}\cdot\boldsymbol{e}_n}{c}\right) \approx 1 - \frac{v}{c}\cos\theta \approx 1 - \frac{v}{c}\left(1 - \frac{\theta^2}{2}\right) = 1 - \frac{v}{c} + \frac{\theta^2}{2} \approx \frac{1}{2}\left(1 - \frac{v^2}{c^2}\right) + \frac{\theta^2}{2}.$$

显然,当第二项不大于第一项时,分母作为上式的五次方比较小,分式的值比较大;而当第二项明显大于第一项时,分母将以上式的五次方的速率迅速增加,分式的值迅速减小. 由此可见,辐射能量将集中于 $\theta^2/2 \approx (1 - v^2/c^2)/2$ 的范围中,即(2.24)式所示.

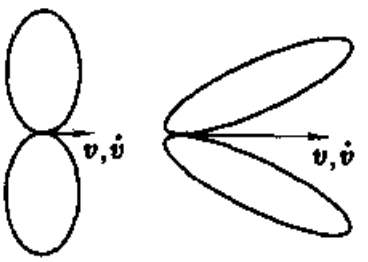
(三)韧致辐射

韧致辐射是指 $\dot{\boldsymbol{v}} /\!/ \boldsymbol{v}$ 时的辐射,最常见的例子是在直线加速器中,高速运动的电子轰击金属靶因减速而辐射出 X 射线.

现将 $\dot{\boldsymbol{v}} /\!/ \boldsymbol{v}$ 的条件代入(2.10)式,利用 $\boldsymbol{v}\times\dot{\boldsymbol{v}} = 0$,即得电子辐射的电磁场

$$\boldsymbol{E} = \frac{e}{4\pi\varepsilon_0 c^2 r}\frac{\boldsymbol{e}_n \times (\boldsymbol{e}_n \times \dot{\boldsymbol{v}})}{(1 - \boldsymbol{v}\cdot\boldsymbol{e}_n/c)^3}, \quad (2.25)$$

$$\boldsymbol{B} = \frac{1}{c}\boldsymbol{e}_n \times \boldsymbol{E} = -\frac{e}{4\pi\varepsilon_0 c^3 r}\boldsymbol{e}_n \times \dot{\boldsymbol{v}}. \quad (2.26)$$



(a) $v \ll c$　　(b) $v \approx c$

图 2　韧致辐射的角分布

辐射能流为

$$\boldsymbol{S}=\boldsymbol{E}\times\boldsymbol{H}=\frac{e^2\dot{v}^2}{16\pi^2\varepsilon_0c^3r^2}\frac{\sin^2\theta}{(1-v\cos\theta/c)^6}\boldsymbol{e}_n. \tag{2.27}$$

辐射功率的角分布可将上式代入(2.23)式即得

$$\frac{\mathrm{d}P(t')}{\mathrm{d}\Omega}=\boldsymbol{e}_n\cdot\mathrm{S}\,r^2\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{e}_n}{c}\right)=\frac{e^2\dot{v}^2}{16\pi^2\varepsilon_0c^3}\frac{\sin^2\theta}{(1-v\cos\theta/c)^5}. \tag{2.28}$$

辐射角分布如图2所示,其中(a)为$v\ll c$的辐射角分布;(b)为$v\approx c$时的辐射角分布.后者有两个明显的特点:一是保持与φ角无关,即分布和$v\ll c$情况一样以v为对称轴;二是出现$(1-v\cos\theta/c)^{-5}$因子,使得角分布曲线向前倾,并且随v的增大而加剧.

辐射功率可由(2.28)式对立体角积分而得

$$P=\oint\frac{\mathrm{d}P(t')}{\mathrm{d}\Omega}\mathrm{d}\Omega=\frac{e^2}{6\pi\varepsilon_0c^3}\frac{\dot{v}^2}{(1-v^2/c^2)^3}. \tag{2.29}$$

与低速运动的情况相比较[见§4.2.1的(1.29)式],多了一个$(1-v^2/c^2)^{-3}$因子,可见它比低速运动时的辐射功率强,且随v的增大而急剧增大.

为了用粒子所受的力表示辐射功率,可将$\dot{\boldsymbol{v}}/\!/\boldsymbol{v}$代入相对论力学基本方程

$$\boldsymbol{F}=\frac{\mathrm{d}}{\mathrm{d}t}\frac{m_0\boldsymbol{v}}{\sqrt{1-v^2/c^2}}=\frac{m_0\dot{\boldsymbol{v}}}{(1-v^2/c^2)^{3/2}}=\gamma^3m_0\dot{\boldsymbol{v}}. \tag{2.30}$$

由于此时$\dot{\boldsymbol{v}}$与$\boldsymbol{v}$方向相同,上式已将$\boldsymbol{v}\dot{v}=v\dot{\boldsymbol{v}}$代入.再将(2.30)式代入(2.29)式,即得

$$P(t')=\frac{e^2F^2}{6\pi\varepsilon_0m_0^2c^3}. \tag{2.31}$$

它表明,当粒子所受的作用力不变时,作直线运动的粒子的辐射功率与粒子的能量无关.

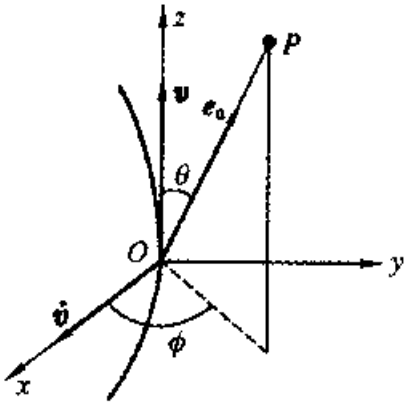


图3　同步辐射

(四)同步辐射

同步辐射是指$\dot{\boldsymbol{v}}\perp\boldsymbol{v}$时的辐射,最常见的例子是在电子同步加速器中作匀速圆周运动时发出的电磁辐射.这一现象于1948年被发现.

现取z轴沿粒子瞬时速度$\boldsymbol{v}$的方向,x轴沿粒子瞬时加速度$\dot{\boldsymbol{v}}$的方向,$\boldsymbol{e}_n=\mathrm{r}/r$为粒子指向场点$P$的单位矢量,如图3所示.由图可见:

$$\begin{aligned}\boldsymbol{v}\cdot\dot{\boldsymbol{v}}&=0,\\ \boldsymbol{v}\cdot\boldsymbol{e}_n&=v\cos\theta,\\ \dot{\boldsymbol{v}}\cdot\boldsymbol{e}_n&=|\dot{\boldsymbol{v}}|\sin\theta\cos\phi.\end{aligned} \tag{2.32}$$

由以上各式可得

$$\left|\boldsymbol{e}_n\times\left[\left(\boldsymbol{e}_n-\frac{\boldsymbol{v}}{c}\right)\times\dot{\boldsymbol{v}}\right]\right|^2=\left|(\boldsymbol{e}_n\cdot\dot{\boldsymbol{v}})\left(\boldsymbol{e}_n-\frac{\boldsymbol{v}}{c}\right)-\left(1-\frac{\boldsymbol{e}_n\cdot\boldsymbol{v}}{c}\right)\dot{\boldsymbol{v}}\right|^2$$

$$=\left|\dot{\boldsymbol{v}}\right|\sin\theta\cos\phi\left(\boldsymbol{e}_n-\frac{\boldsymbol{v}}{c}\right)-\left(1-\frac{v}{c}\cos\theta\right)\dot{\boldsymbol{v}}\Bigg|^2$$

$$=\dot{v}^2\left[\sin^2\theta\cos^2\phi\left(\boldsymbol{e}_n-\frac{\boldsymbol{v}}{c}\right)^2+\left(1-\frac{v}{c}\cos\theta\right)^2\right.$$

$$\left.-2\sin\theta\cos\phi\left(1-\frac{v}{c}\cos\theta\right)(\boldsymbol{e}_n\cdot\dot{\boldsymbol{v}})\right]$$

$$=\dot{v}^2\left[\left(1-\frac{v}{c}\cos\theta\right)^2-\left(1-\frac{v^2}{c^2}\right)\sin^2\theta\cos^2\phi\right]. \quad (2.33)$$

将此式代入(2.22)式即得辐射角分布

$$\frac{\mathrm{d}P(t')}{\mathrm{d}\Omega}=\frac{e^2\dot{v}^2}{16\pi^2\varepsilon_0c^3}\frac{(1-v\cos\theta/c)^2-(1-v^2/c^2)\sin^2\theta\cos^2\phi}{(1-v\cos\theta/c)^5}. \quad (2.34)$$

上式表明,角分布与 θ,ϕ 均有关,在粒子运动轨道平面上(即 $\boldsymbol{v}$ 与 $\dot{\boldsymbol{v}}$ 所确定的平面 xoz),因 $\phi=0$ 或 π,故有

$$\frac{\mathrm{d}P(t')}{\mathrm{d}\Omega}=\frac{e^2\dot{v}^2}{16\pi^2\varepsilon_0c^3}\frac{(1-v\cos\theta/c)^2-\sin^2\theta(1-v^2/c^2)}{(1-v\cos\theta/c)^5}. \quad (2.35)$$

由(2.35)式可见,在 $\theta=0$ 方向,辐射最强;而在 $\cos\theta_0=v/c$ 方向,$\mathrm{d}P(t')/\mathrm{d}\Omega=0$. 这表明:同步辐射集中在以粒子速度方向为中心的、张角为 $2\theta_0$ 的范围内,且 v 越大,θ_0 越小,辐射能量越集中,如图 4 所示.

将上式对 $\mathrm{d}\Omega$ 积分便得辐射总功率

$$P(t')=\frac{e^2\dot{v}^2}{6\pi\varepsilon_0c^3(1-v^2/c^2)^2}, \quad (2.36)$$

与偶极辐射比较[见 §4.2.1 的(1.29)式)],多出一个因子 $(1-v^2/c^2)^{-2}$.

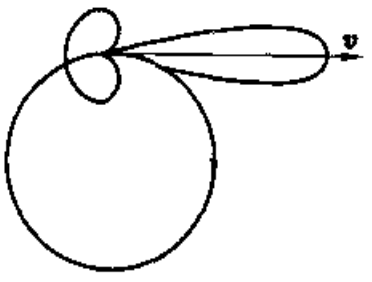


图 4 同步辐射的角分布

下面进一步用粒子所受的力来表示辐射功率. 当 $\dot{\boldsymbol{v}}\perp\boldsymbol{v}$(即 $\dot{\boldsymbol{v}}$ 没有沿 $\boldsymbol{v}$ 的分量)时,其速率 v 没有变化,将 $\dot{v}=0$ 代入相对论力学基本方程,即有

$$\boldsymbol{F}=\frac{\mathrm{d}}{\mathrm{d}t}\frac{m_0\boldsymbol{v}}{\sqrt{1-v^2/c^2}}=\frac{m_0\dot{\boldsymbol{v}}}{\sqrt{1-v^2/c^2}}=\gamma m_0\dot{\boldsymbol{v}}. \quad (2.37)$$

代入(2.36)式得粒子的辐射功率

$$P(t')=\frac{e^2F^2}{6\pi\varepsilon_0m_0^2c^3}\gamma^2. \quad (2.38)$$

再将带电粒子的能量

$$E = \frac{m_0c^2}{\sqrt{1-v^2/c^2}} = m_0c^2\gamma \tag{2.39}$$

即 $\gamma = E/(m_0c^2)$ 代入上式，即有

$$P(t') = \frac{e^2F^2}{6\pi\varepsilon_0 m_0^2c^3}\left(\frac{E}{m_0c^2}\right)^2 = \frac{e^2F^2}{6\pi\varepsilon_0 m_0^4c^7}E^2. \tag{2.40}$$

由此可见，在一定作用力下，同步辐射功率与粒子的能量平方成正比. 电子能量越高，辐射损耗就越大. 当加速器提供的功率与辐射损耗相等时，电子的能量便不能再增加，这是同步加速器的缺点. 为了获得较高能量的电子，需要采用直线加速器.

§4.2.3　切连科夫辐射

上节讨论了在真空中作加速运动的带电粒子所产生的辐射. 如果粒子的加速度 $\dot{v}=0$，则无辐射. 但是，切连科夫在 1934 年发现：带电粒子在介质中作匀速直线运动时，如果它的运动速度大于介质中的光速 $v_p = c/n$，则将产生电磁辐射，称为切连科夫辐射. 本节首先给出切连科夫辐射的理论解释，接着用傅里叶变换求出粒子辐射的电磁场的频率为 ω 的分量.

(一) 切连科夫辐射的理论解释

1937 年，塔姆和弗朗克提出了切连科夫辐射的理论解释. 当带电粒子通过介质时，介质内产生束缚电荷和诱导电流，它们将激发以相速度为 $v_p = c/n$ 的球面次波向外辐射. 若带电粒子的运动速度 $v < v_p$，则这些球面次波不可能相遇，就不可能产生相增长的干涉而加强. 它们将成为一个个孤立的球面波而逐渐衰减，不能形成辐射，如图 1(a)所示. 若 $v > v_p$，这些次波将由于干涉而加强，形成一个以粒子为顶点，以粒子运动方向为轴的锥形波阵面，如图 1(b)所示. 锥形波阵面的法线方向即为切连科夫辐射的方向. 辐射方向与粒子运动方向的夹角 θ_c 由下式确定：

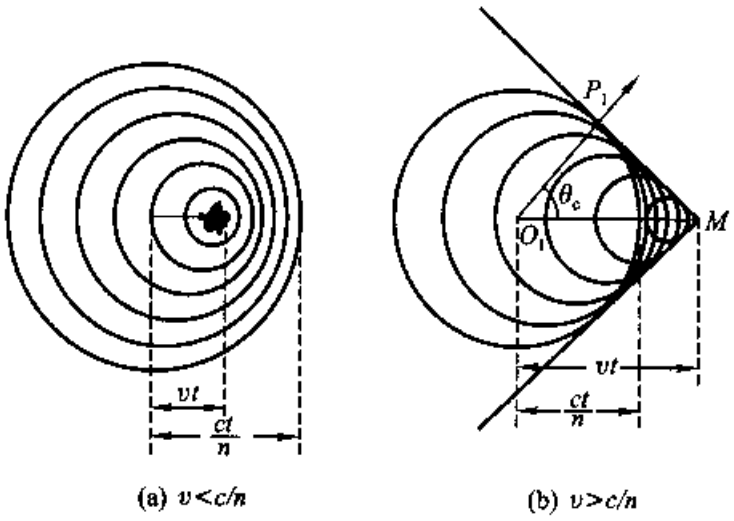


图 1　束缚电荷和诱导电流激发的球面次波

$$\cos\theta_c = \overline{O_1P_1}/\overline{O_1M} = (ct/n)/(vt) = c/(nv). \tag{3.1}$$

公式(3.1)与实验结果完全一致.

切连科夫辐射的性质可以用来测定其速度超过介质中光速的快速粒子的速度. 通过测量辐射角 θ_C,即可由公式(3.1)确定粒子的速度.

(二)在介质中运动的带电粒子激发的电磁场

设电量为 e 的粒子以 $\boldsymbol{v}=v\boldsymbol{e}_z$ 作匀速直线运动(略去粒子由于辐射损耗而带来的减速效应),其位矢为 $\boldsymbol{x}_e(t)=\boldsymbol{v}t$,则其电荷密度与电流密度分别为:

$$\rho_f(\boldsymbol{x},t)=e\delta^3[\boldsymbol{x}-\boldsymbol{x}_e(t)], \tag{3.2}$$

$$\boldsymbol{J}_f(\boldsymbol{x},t)=e\boldsymbol{v}\,\delta^3[\boldsymbol{x}-\boldsymbol{x}_e(t)]. \tag{3.3}$$

带电粒子在介质中运动所激发的束缚电荷与诱导电流所产生的宏观效应可以用介电常数 ε 和磁导率 μ 描述. 因此,运动带电粒子在介质中激发的电磁场可由推迟势[见 §4.1.1 的(1.23)式]

$$\boldsymbol{A}(\boldsymbol{x},t)=\frac{\mu_0}{4\pi}\int\frac{\boldsymbol{J}_f(\boldsymbol{x}',t-r/c)}{r}\mathrm{d}V'. \tag{3.4}$$

作适当的替换而得到. 另一方面,考虑到介质的 $\mu\approx\mu_0$(非铁磁介质的 $\mu_r\approx1$),而 $\varepsilon=\varepsilon_r\varepsilon_0$ 及 $n=\sqrt{\varepsilon_r\mu_r}\approx\sqrt{\varepsilon_r}$ 均与电磁辐射的角频率 ω 有关,在作替换之前应将电流密度 $\boldsymbol{J}_f(\boldsymbol{x},t)$ 展开为傅里叶积分

$$\boldsymbol{J}_f(\boldsymbol{x},t)=\int_{-\infty}^{\infty}\boldsymbol{J}_\omega(\boldsymbol{x})\mathrm{e}^{-\mathrm{i}\omega t}\mathrm{d}\omega. \tag{3.5}$$

由此得 $\boldsymbol{J}_f(\boldsymbol{x},t)$ 的频率为 ω 的分量

$$\boldsymbol{J}_\omega(\boldsymbol{x})=\frac{1}{2\pi}\int_{-\infty}^{\infty}\boldsymbol{J}_f(\boldsymbol{x},t)\mathrm{e}^{\mathrm{i}\omega t}\mathrm{d}t. \tag{3.6}$$

现在将(3.4)式中的 c 替换为 $v=c/n$ 即得在介质中运动的带电粒子的辐射矢势

$$\boldsymbol{A}(\boldsymbol{x},t)=\frac{\mu_0}{4\pi}\int\frac{\boldsymbol{J}_f(\boldsymbol{x}',t-nr/c)}{r}\mathrm{d}V'. \tag{3.7}$$

将(3.5)式代入上式即得

$$\boldsymbol{A}(\boldsymbol{x},t)=\frac{\mu_0}{4\pi}\int\frac{\mathrm{d}V'}{r}\int_{-\infty}^{\infty}\boldsymbol{J}_\omega(\boldsymbol{x}')\mathrm{e}^{-\mathrm{i}\omega\left(t-\frac{nr}{c}\right)}\mathrm{d}\omega. \tag{3.8}$$

与 $\boldsymbol{A}(\boldsymbol{x},t)$ 的傅里叶积分

$$\boldsymbol{A}(\boldsymbol{x},t)=\int_{-\infty}^{\infty}\boldsymbol{A}_\omega(\boldsymbol{x})\mathrm{e}^{-\mathrm{i}\omega t}\mathrm{d}\omega \tag{3.9}$$

相比较,即得带电粒子辐射的矢势的 ω 分量

$$\boldsymbol{A}_\omega(\boldsymbol{x})=\frac{\mu_0}{4\pi}\int\frac{\boldsymbol{J}_\omega(\boldsymbol{x}')\mathrm{e}^{\mathrm{i}\omega nr/c}}{r}\mathrm{d}V'. \tag{3.10}$$

将(3.6)式的 t 换为 t' 代入上式,便有

$$\boldsymbol{A}_\omega(\boldsymbol{x}) = \frac{\mu_0}{8\pi^2}\int_{-\infty}^{\infty} \mathrm{d}t' \mathrm{e}^{\mathrm{i}\omega\left(t'+\frac{nr}{c}\right)} \int \frac{\boldsymbol{J}_f(\boldsymbol{x}', t')}{r} \mathrm{d}V'.$$

再将运动带电粒子的电流密度(3.3)式代入上式,可得

$$\begin{aligned}\boldsymbol{A}_\omega(\boldsymbol{x}) &= \frac{\mu_0}{8\pi^2}\int_{-\infty}^{\infty} \mathrm{d}t' \mathrm{e}^{\mathrm{i}\omega\left(t'+\frac{nr}{c}\right)} \int \frac{e\boldsymbol{v}(t')\delta^3[\boldsymbol{x}' - \boldsymbol{x}_e(t')]}{r} \mathrm{d}V' \\ &= \frac{\mu_0 e}{8\pi^2}\int_{-\infty}^{\infty} \frac{\boldsymbol{v}(t')}{r} \mathrm{e}^{\mathrm{i}\omega(t'+nr/c)} \mathrm{d}t', \end{aligned} \tag{3.11}$$

式中 $r = |\boldsymbol{x} - \boldsymbol{x}_e(t')|$ 为带电粒子到场点的距离.

现在计算带电粒子在远处的辐射场. 若匀速直线运动的带电粒子在有限的区域内运动,取 z 轴沿 $\boldsymbol{v}$ 的方向. 设 $\boldsymbol{e}_R = \boldsymbol{x}/R$ 为由坐标原点指向场点的单位矢量,由于远处的 $|\boldsymbol{x}| \gg |\boldsymbol{x}_e(t')|$,(3.11)式中的 r 按下述方式取近似值:对于分母中的 r,取 $r = |\boldsymbol{x} - \boldsymbol{x}_e(t')| \approx |\boldsymbol{x}| = R$;对子相因子中的 r,略去 $|\boldsymbol{x}_e|$ 的二阶小量,利用余弦定理及近似式 $\sqrt{1-\varepsilon} \approx 1 - \varepsilon/2$(当 $\varepsilon \ll 1$). 即有

$$\begin{aligned} r &= |\boldsymbol{x} - \boldsymbol{x}_e(t')| = \sqrt{|\boldsymbol{x}|^2 - 2\boldsymbol{x}\cdot\boldsymbol{x}_e(t') + |\boldsymbol{x}_e(t')|^2} \\ &\approx R\sqrt{1 - 2\boldsymbol{e}_R\cdot\boldsymbol{x}_e/R} \approx R - z_e\cos\theta. \end{aligned} \tag{3.12}$$

因 $\boldsymbol{x}_e(t')$ 沿 z 轴,故 $\boldsymbol{e}_R$ 与 $\boldsymbol{x}_e$ 的夹角即 θ. 将(3.12)式代入(3.11)式,利用 $\boldsymbol{v}(t')\mathrm{d}t' = \boldsymbol{e}_z \mathrm{d}z_e$(其中 z_e 表示粒子在 z 轴上的位置),即得粒子辐射的矢势的 ω 分量:

$$\begin{aligned}\boldsymbol{A}_\omega(\boldsymbol{x}) &= \frac{\mu_0 e}{8\pi^2}\int_{-\infty}^{\infty} \frac{1}{R} \mathrm{e}^{\mathrm{i}\omega\left[\frac{z_e}{v}+\frac{n}{c}(R - z_e\cos\theta)\right]} \boldsymbol{e}_z \mathrm{d}z_e \\ &= \frac{e}{8\pi^2\varepsilon_0 c^2}\frac{\mathrm{e}^{\mathrm{i}kR}}{R}\boldsymbol{e}_z \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}\omega\left(\frac{1}{v}-\frac{n}{c}\cos\theta\right)z_e} \mathrm{d}z_e, \end{aligned} \tag{3.13}$$

式中 $k = n\omega/c$ 是介质中的波数,并已利用了熟知的公式 $\mu_0\varepsilon_0 c^2 = 1$.

利用 δ 函数的傅里叶积分表示[见附录(Ⅰ.14)式]

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}\xi z_e} \mathrm{d}z_e = \delta(\xi), \tag{3.14}$$

可将(3.13)式写成

$$\boldsymbol{A}_\omega(\boldsymbol{x}) = \frac{e}{4\pi\varepsilon_0 c^2}\frac{\mathrm{e}^{\mathrm{i}kR}}{R}\boldsymbol{e}_z \delta^3\left(\frac{\omega}{v} - \frac{n\omega}{c}\cos\theta\right), \tag{3.15}$$

仅当 δ 函数的宗量为零,即 $\omega/v - (n\omega/c)\cos\theta = 0$ 时,$\boldsymbol{A}_\omega(\boldsymbol{x})$ 才不为零. 因而

$$\cos\theta = c/nv. \tag{3.16}$$

与(3.1)式相比较,可见 θ 即 θ_C,即仅在 θ_C 方向才可能出现辐射电磁场,而在 $\cos\theta \neq c/nv$ 的方向没有辐射. 但考虑到 n 是 ω 的函数,特别是当 ω 很大时,介质的折射率 $n \approx 1$(见 §4.2.5),$v > c/n$ 得不到满足,因而辐射频谱在高频下截断. 对于低频部分,不同的 ω 对应于不同的 n,因而对应于不同的 θ_C,所以辐射将分布在有

一定宽度的辐射角内.

利用电磁场与矢势的关系(见第四篇第一章习题2的提示),并用 c/n 替代式中的 c,便有

$$\boldsymbol{B} = \mathrm{i}\boldsymbol{k} \times \boldsymbol{A} = \frac{\mathrm{i}n\omega}{c}\boldsymbol{e}_n \times \boldsymbol{A}, \tag{3.17}$$

$$\boldsymbol{E} = \frac{c}{n}\boldsymbol{B} \times \boldsymbol{e}_n. \tag{3.18}$$

由此得电磁场的 ω 分量与矢势的 ω 分量的关系

$$\boldsymbol{B}_\omega = \frac{\mathrm{i}n\omega}{c}\boldsymbol{e}_n \times \boldsymbol{A}_\omega, \tag{3.19}$$

$$\boldsymbol{E}_\omega = \frac{c}{n}\boldsymbol{B}_\omega \times \boldsymbol{e}_n. \tag{3.20}$$

(3.16)式表明 $\boldsymbol{k}$ 与 $\boldsymbol{A}_\omega$ 的夹角,即 $\boldsymbol{e}_n$ 与 $\boldsymbol{e}_z$ 的夹角为 θ. 利用 $\boldsymbol{e}_n \times \boldsymbol{e}_z = -\sin\theta\,\boldsymbol{e}_\phi$,可得

$$\boldsymbol{B}_\omega = -\frac{\mathrm{i}n\omega e}{4\pi\varepsilon_0 c^3}\frac{\mathrm{e}^{\mathrm{i}kR}}{R}\sin\theta\,\delta^3\left(\frac{\omega}{v} - \frac{n\omega}{c}\cos\theta\right)\boldsymbol{e}_\phi. \tag{3.21}$$

将上式代入(3.18)式,并利用 $\boldsymbol{e}_\phi \times \boldsymbol{e}_n = \boldsymbol{e}_\theta$,便有

$$\boldsymbol{E}_\omega = \frac{\mathrm{i}\omega e}{4\pi\varepsilon_0 c^2}\frac{\mathrm{e}^{\mathrm{i}kR}}{R}\sin\theta\delta^3\left(\frac{\omega}{v} - \frac{n\omega}{c}\cos\theta\right)\boldsymbol{e}_\theta. \tag{3.22}$$

根据(3.11)及(3.12)式还可计算出粒子通过单位距离单位频率间隔的辐射能量为

$$\frac{\mathrm{d}E(\omega)}{\mathrm{d}L} = \frac{e^2}{4\pi\varepsilon_0 c^2}\left(1 - \frac{c^2}{n^2v^2}\right)\omega. \tag{3.23}$$

具体的计算留作习题.

§4.2.4 带电粒子的电磁场对粒子本身的反作用

在§4.2.1中我们给出了低速运动带电粒子所激发的电磁场[见该节的式(1.24)、(1.25)]:

$$\boldsymbol{E} = \frac{e\boldsymbol{r}}{4\pi\varepsilon_0 r^3} + \frac{e}{4\pi\varepsilon_0 c^2 r^2}[\boldsymbol{r} \times (\boldsymbol{r} \times \dot{\boldsymbol{v}})], \tag{4.1}$$

$$\boldsymbol{B} = \frac{\mu_0 e\boldsymbol{v} \times \boldsymbol{r}}{4\pi r^3} + \frac{e\dot{\boldsymbol{v}} \times \boldsymbol{r}}{4\pi\varepsilon_0 c^3 r^2}, \tag{4.2}$$

并且指出第一项为自有场,第二项为辐射场. 本节分别讨论自有场与辐射场对粒子的反作用,并由此得到电子的经典半径和原子自发辐射的谱线宽度.

(一)电磁质量及电子的经典半径

首先讨论自有场对粒子的反作用. 由 §2.4.4 图 1 可以看出,匀速运动带电粒子所激发的电磁场是随着粒子向前运动的. 特别是,当 $v \ll c$ 时,这个电磁场过渡到 §2.4.4 的(4.28)、(4.29)式,这正是(4.1)、(4.2)式的第一项. 既然自有场与粒子本身不可分离,当测量带电粒子的能量时,自然就把不包含自有场的"裸"粒子的能量 E_m 和自有场的能量 E_{em} 同时测量进去了. 由相对论质能关系 $E = mc^2$ 可见,和 E_m 与 E_{em} 相对应,其质量也分为两部分:

$$m = m_m + m_{em}, \tag{4.3}$$

其中 m_m 是裸粒子的质量,而 m_{em} 是粒子的自有场的质量.

那么,对于电子来说,它的电磁质量有多大呢? 由于目前对于电子的内部结构还不清楚,只能作数量级的估计. 为此,假设电子的电荷分布在半径为 r_e 的球面上,这样求出电子的库仑能为

$$E = \frac{1}{2}\int_\infty \varepsilon_0 E^2 \mathrm{d}V = \frac{\varepsilon_0}{2}\int_{r_e}^{\infty}\left(\frac{e}{4\pi\varepsilon_0 R^2}\right)^2 4\pi R^2 \mathrm{d}R = \frac{e^2}{8\pi\varepsilon_0 r_e}.$$

因此电子的电磁质量为

$$m_{em} = \frac{E}{c^2} = \frac{e^2}{8\pi\varepsilon_0 c^2 r_e}. \tag{4.4}$$

由于电子的结构不清楚,r_e 不知道,m_{em} 也就不知道. 如果进一步假设

$$m_m \approx m_{em} \tag{4.5}$$

则可以利用实验测得的电子质量来估计电子的半径 r_e. 这样得到的 r_e 称为电子的经典半径. 将

$$m = m_m + m_{em} \approx 2m_{em}$$

代入(4.4)式,即得电子经典半径为:

$$r_e = \frac{1}{4\pi\varepsilon_0}\frac{e^2}{mc^2} = 2.818 \times 10^{-13}\mathrm{cm}. \tag{4.6}$$

这是根据电子的经典模型计算出来的结果. 实际上,在 $R \approx 10^{-10}$cm 的线度上量子效应已经重要,经典理论失效. 因此,r_e 仅仅在某些特殊情况下才显示电子半径的特征(例如见后面 §4.2.5 自由电子对电磁波的散射).

(二)辐射阻尼力

辐射场对粒子的反作用表现为场对粒子的反作用力,称为辐射阻尼力 $\boldsymbol{F}_{阻}$. 现在根据能量守恒定律寻找 $\boldsymbol{F}_{阻}$ 的形式. 在一般情况下计算 $\boldsymbol{F}_{阻}$ 是很困难的,因为当粒子的运动状态不同时,它的自有场也会发生变化. 这里,我们仅讨论一个非常重要的特殊情形,即电子作匀速圆周运动的情形. 这时,当粒子运动一周,回到它原来位置时,可以认为附近的自有场没有变化,于是电子的能量变化仅仅由于它的辐

射. 在低速情况下,总辐射功率已由上节求得为

$$P = e^2\dot{\boldsymbol{v}}^2/(6\pi\varepsilon_0 c^3). \tag{4.7}$$

因为辐射阻尼力所作的负功等于辐射出去的能量,所以

$$\int_{t_0}^{t_0+T}\boldsymbol{F}_{\text{阻}}\cdot\boldsymbol{v}\,\mathrm{d}t = -\int_{t_0}^{t_0+T}\frac{e^2\dot{\boldsymbol{v}}\dot{\boldsymbol{v}}}{6\pi\varepsilon_0 c^3}\,\mathrm{d}t = -\frac{e^2}{6\pi\varepsilon_0 c^3}\int_{t_0}^{t_0+T}\dot{\boldsymbol{v}}\cdot\mathrm{d}\boldsymbol{v}$$

$$= -\frac{e^2}{6\pi\varepsilon_0 c^3}\left[\dot{\boldsymbol{v}}\boldsymbol{v}\Big|_{t_0}^{t_0+T} - \int_{t_0}^{t_0+T}\boldsymbol{v}\cdot\mathrm{d}\dot{\boldsymbol{v}}\right] = \int_{t_0}^{t_0+T}\frac{e^2\ddot{\boldsymbol{v}}\cdot\boldsymbol{v}}{6\pi\varepsilon_0 c^3}\,\mathrm{d}t, \tag{4.8}$$

其中$\dot{\boldsymbol{v}}\cdot\boldsymbol{v}\Big|_{t_0}^{t_0+T}=0$是因为电子运动一周后回到原处时,电子的运动状态可以认为没有变化. 这样,(4.8)式说明,作为一个周期的平均效应,辐射阻尼力$\boldsymbol{F}_{\text{阻}}$为:

$$\boldsymbol{F}_{\text{阻}} = e^2\ddot{\boldsymbol{v}}/(6\pi\varepsilon_0 c^3). \tag{4.9}$$

显然,它不是一个瞬时力,只能代表一种平均效应.

值得注意的是,只有当辐射阻尼力远小子作用于粒子上的外力时,才可以利用辐射阻尼力的概念,否则将会出现电子自我加速的荒谬结果. 例如,当作用于电子的外力为零时,应用(4.9)式于牛顿第二定律,即有

$$m\dot{\boldsymbol{v}} = e^2\ddot{\boldsymbol{v}}/(6\pi\varepsilon_0 c^3). \tag{4.10}$$

这方程除了平凡解$\boldsymbol{v}=$常量外,尚有一"奔离"解:

$$\dot{\boldsymbol{v}}(t) = \dot{\boldsymbol{v}}(0)\,\mathrm{e}^{6\pi\varepsilon_0 mc^3 t/e^2}. \tag{4.11}$$

它表示电子将无限制地"自我加速". 这说明应用辐射阻尼力$\boldsymbol{F}_{\text{阻}}$于运动方程是有限制的,否则将导致不正确的结果.

(三)谱线的自然宽度

原子自发地由高能态过渡到低能态辐射电磁波,称为原子的自发辐射. 从经典电磁理论的角度来看,辐射阻尼力对原子的自发辐射过程起着决定性的作用.

我们采用经典谐振子作为原子的模型. 为确定起见,设振子在x轴上运动,弹性恢复力为$-kx$. 在原子自发辐射的情形中,辐射阻尼力远小于弹性恢复力(后面将作出验证). 因此,可以将$F_{\text{阻}}$作为瞬时力引人振于运动方程

$$m\ddot{x} = -kx + F_{\text{阻}}. \tag{4.12}$$

令$k/m=\omega_0^2$,并将(4.9)式代人,即有

$$\ddot{x} + \omega_0^2 x = e^2\dddot{x}/(6\pi\varepsilon_0 mc^3). \tag{4.13}$$

现在利用逐次近似法求解方程(4.13),首先略去$F_{\text{阻}}$,得到近似解$x=x_0\mathrm{e}^{-\mathrm{i}\omega_0 t}$,然后用这个解求$\dddot{x}$得$\dddot{x}=-\omega_0^2\dot{x}$,子是方程成为

$$\ddot{x} + \gamma\dot{x} + \omega_0^2 x = 0. \tag{4.14}$$

其中

$$\gamma = e^2\omega_0^2/(6\pi\varepsilon_0 mc^3). \tag{4.15}$$

(4.14)式就是§1.4.2中已经解过的阻尼振动方程，它的解是[见§1.4.2(2.11)式]

$$x = x_0 e^{-\gamma t/2} e^{-i\omega_0 t}, \tag{4.16}$$

这说明原子是一个振幅不断衰减的振子，其能量衰减到初始值的1/e的时间称为振子的寿命. 因为振子的能量正比于振子振幅的平方，故振子的寿命为：

$$\tau = 1/\gamma. \tag{4.17}$$

不难验证，辐射阻尼力 $F_{阻}$ 远小于弹性恢复力，因此把 $F_{阻}$ 作为瞬时力引入运动方程是允许的. 实际上，它们两者之比为

$$\left|\frac{-m\gamma\dot{x}}{-kx}\right| \approx \left|\frac{m\gamma i\omega_0 x}{m\omega_0^2 x}\right| = \frac{\gamma}{\omega_0} = \frac{e^2\omega_0^2}{6\pi\varepsilon_0 mc^3\omega_0} = \frac{2r_e\omega_0}{3c} = \frac{2r_e}{3}\frac{2\pi}{\lambda} \sim \frac{10^{-13}}{10^{-5}} \ll 1.$$

这里应用了电子的经典半径 $r_e \approx 10^{-13}$cm，并假设原子辐射的波长 $\lambda \approx 10^{-5}$cm.

现在来讨论辐射的频谱分布. 设在 $t=0$ 时辐射开始到达观察点，在该点辐射的场强为

$$\boldsymbol{E} = \int_{-\infty}^{+\infty} \boldsymbol{E}_\omega e^{-i\omega t} d\omega = \begin{cases} \boldsymbol{E}_0 e^{-\gamma t/2} e^{-i\omega_0 t}, & t \geqslant 0, \\ 0, & t < 0. \end{cases} \tag{4.18}$$

由此得 $\boldsymbol{E}$ 的傅里叶分量

$$\boldsymbol{E}_\omega = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \boldsymbol{E}(t) e^{i\omega t} dt = \frac{1}{2\pi}\int_0^{\infty} \boldsymbol{E}_0 e^{-\gamma t/2} e^{i(\omega-\omega_0)t} dt$$

$$= \frac{1}{2\pi}\frac{\boldsymbol{E}_0}{i(\omega-\omega_0) - \gamma/2}. \tag{4.19}$$

振子在给定方向上的辐射强度 I_ω 就是在单位时间单位频率间隔内辐射到该方向附近单位立体角内的辐射能量. 它正比于 $|\boldsymbol{E}_\omega|^2$. 由此得

$$I_\omega \propto \frac{1}{(\omega-\omega_0)^2 + (\gamma/2)^2}. \tag{4.20}$$

设总辐射强度为

$$I_0 = \int_{-\infty}^{+\infty} I_\omega d\omega, \tag{4.21}$$

将(4.20)式代入(4.21)式就可以求得比例系数为 $I_0\gamma/2\pi$，故

$$I_\omega = \frac{I_0}{2\pi}\frac{\gamma}{(\omega-\omega_0)^2 + (\gamma/2)^2}. \tag{4.22}$$

这表明，辐射强度遵从布雷特-维格纳共振分布，见§1.4.2(2.22)式.

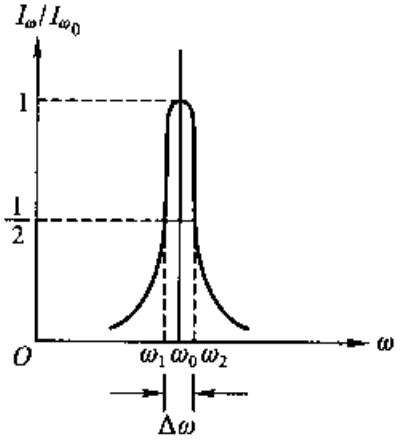


图1　谱线的自然宽度

图 1 给出 I_ω 随 ω 变化的曲线(比较 §1.4.2 图 1b). 当 $\omega=\omega_0$ 时,I_ω 有最大值:

$$I_{\omega_0}=\frac{I_0}{2\pi}\frac{\gamma}{(\gamma/2)^2}. \tag{4.23}$$

当 I_ω 下降到最大值的一半时,$I_{\omega_1}=I_{\omega_2}=I_{\omega_0}/2$,相应的 ω 值之差 $\omega_2-\omega_1=\Delta\omega$ 称为谱线宽度. 原子自发辐射所形成的谱线宽度称为谱线的自然宽度:

$$\Delta\omega=\omega_2-\omega_1=\gamma. \tag{4.24}$$

以上是利用经典理论对原子及其辐射的处理. 显然,它是很不完备的,只能粗略地解释若干特殊问题,如谱线的自然宽度问题. 实际上,如果按照经典理论,电子围绕原子核作圆周运动,由于 $\dot{\boldsymbol{v}}\neq 0$,必然不断地向外辐射电磁波而耗散其能量(参看第四篇第一章习题 5),最终要掉到原子核上而不可能存在稳定的原子. 这一结论与实验事实相违背. 完善地解决原子中电子的运动及其辐射有赖于量子理论,见第六篇.

§4.2.5 电磁波的散射、吸收和色散

前两节讨论了带电粒子激发的电磁场和这一电磁场对粒子本身的反作用. 本节研究外来电磁场与带电粒子的相互作用.

(一)电磁波的散射

先讨论自由电子对电磁波的散射. 设电子在外来电磁场作用下,其速度 $v\ll c$,则电子运动的幅度 $\sim vT\ll cT=\lambda$(λ 为入射波的波长). 因此,可用某定点的场强代替作用于电子的场强. 同时,由平面电磁波的性质 $|\boldsymbol{B}|/|\boldsymbol{E}|=1/c$ 可以知道,电磁波的磁场作用于电子的力与电场作用于电子的力之比 $|e\boldsymbol{v}\times\boldsymbol{B}|/|e\boldsymbol{E}|\approx v/c\ll 1$,因面只须考虑入射电磁波的电场对电子的作用力. 设入射波电场为

$$\boldsymbol{E}=\boldsymbol{E}_0 e^{-i\omega t},$$

将洛伦兹力及辐射阻尼力代入电子运动方程得

$$\ddot{\boldsymbol{x}}=\frac{e}{m}\boldsymbol{E}_0 e^{-i\omega t}+\frac{e^2}{6\pi\varepsilon_0 mc^3}\dddot{\boldsymbol{x}}. \tag{5.1}$$

这个方程的稳态解为

$$\boldsymbol{x}=\boldsymbol{x}_0 e^{-i\omega t}. \tag{5.2}$$

由此得 $\ddot{\boldsymbol{x}}=-\omega^2\dot{\boldsymbol{x}}$,再令 $\gamma=e^2\omega^2/(6\pi\varepsilon_0 mc^3)$,方程(5.1)变为

$$\ddot{\boldsymbol{x}}+\gamma\dot{\boldsymbol{x}}=(e/m)\boldsymbol{E}_0 e^{-i\omega t}. \tag{5.3}$$

将(5.2)式代入(5.3)式即可求得

$$\boldsymbol{x}_0=-e\boldsymbol{E}_0/[m(\omega^2+i\omega\gamma)]. \tag{5.4}$$

只要入射波的波长 $\lambda \gg r_e \approx 10^{-13}$cm，上式右边分母的实部与虚部之比为 $\omega^2/(\omega\gamma) = 3\lambda/4\pi r_e \gg 1$，故虚部可略去. (5.4)式变为

$$\boldsymbol{x}_0 = -e\boldsymbol{E}_0/(m\omega^2). \tag{5.5}$$

代入(5.2)式，即有

$$\boldsymbol{x} = -[e\boldsymbol{E}_0/(m\omega^2)]\mathrm{e}^{-\mathrm{i}\omega t}. \tag{5.6}$$

由此得 $\dot{\boldsymbol{v}}$ 的实部为

$$\dot{\boldsymbol{v}} = \ddot{\boldsymbol{x}} = (e/m)\boldsymbol{E}_0\cos\omega t. \tag{5.7}$$

为确定起见，取坐标系如图 1 所示. z 轴沿入射波传播方向，其电场在 xoy 平面上且与 x 轴夹角为 ϕ. 观察点 P 在 xoz 平面上，$\boldsymbol{r}$ 与 z 轴的夹角为 θ，$\boldsymbol{r}$ 与 $\boldsymbol{E}_0$ 的夹角为 α. 于是，

$$\boldsymbol{E}_0 \cdot \boldsymbol{r} = E_0 r\cos\alpha, \tag{5.8}$$

$$\boldsymbol{E}_0 = E_0(\cos\phi\boldsymbol{e}_x + \sin\phi\boldsymbol{e}_y), \tag{5.9}$$

$$\boldsymbol{r} = r(\sin\theta\boldsymbol{e}_x + \cos\theta\boldsymbol{e}_z). \tag{5.10}$$

由(5.9)，(5.10)式可得 $\boldsymbol{E}_0$ 与 $\boldsymbol{r}$ 的点积为：

$$\boldsymbol{E}_0 \cdot \boldsymbol{r} = E_0 r(\cos\phi\boldsymbol{e}_x + \sin\phi\boldsymbol{e}_y)\cdot(\sin\theta\boldsymbol{e}_x + \cos\theta\boldsymbol{e}_z) = E_0 r\sin\theta\cos\phi,$$

与(5.8)式比较，可得

$$\cos\alpha = \sin\theta\cos\phi. \tag{5.11}$$

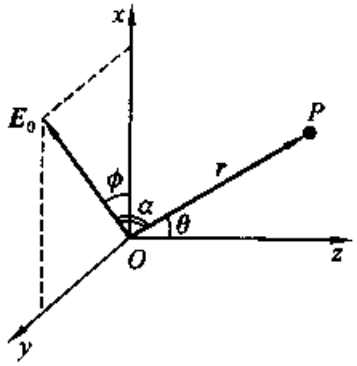


图 1　散射波角分布的计算

由于电子的运动速度 $v \ll c$，故电子的辐射是电偶极辐射. 利用(5.7)，(5.8)式及 §4.2.1 的(1.28)式可得散射波能流密度为

$$\boldsymbol{S} = \frac{e^2\sin^2\alpha}{16\pi^2\varepsilon_0 c^3 r^2}\frac{e^2}{m^2}E_0^2\cos^2\omega t\boldsymbol{e}_n. \tag{5.12}$$

因为 $\cos^2\omega t$ 对时间的平均值为 1/2，由(5.12)式和(5.11)式可得平均散射能流的大小为

$$\bar{S} = \frac{e^4E_0^2}{32\pi^2\varepsilon_0 c^3 m^2 r^2}\sin^2\alpha = \frac{r_e^2}{r^2}I_0(1-\sin^2\theta\cos^2\phi). \tag{5.13}$$

这里引入了入射波强度 I_0，其定义为平均入射能流

$$I_0 = \bar{S}_0 = c\varepsilon_0 E_0^2/2. \tag{5.14}$$

将(5.13)式对球面积分得散射波总平均功率. 再将(5.14)式代入，并引入具有面积量纲的量 $\sigma = 8\pi r_e^2/3$，即有

$$P = \oint \bar{S}r^2\mathrm{d}\Omega = r_e^2 I_0\int_0^{2\pi}\mathrm{d}\phi\int_0^{\pi}(1-\sin^2\theta\cos^2\phi)\mathrm{d}\theta = (8\pi r_e^2/3)I_0 = \sigma I_0. \tag{5.15}$$

这称为汤姆孙散射公式. 由于 I_0 是每秒种垂直入射到单位截面上的能量，上式说

明被散射的能量相当于射到面积为 $\sigma = 8\pi r_e^2/3$ 的截面上的能量，故面积 σ 称为自由电子对电磁波的汤姆孙散射截面. 当入射光子的能量远小于电子静止能量时，上述结论与实验结果相符. 所以，在一定条件下，电子的表现好像一个半径为 r_e 的球一样.

当入射波是非偏振波时，将(5.13)式对 ϕ 求平均，得到非偏振的入射波平均散射能流

$$\bar{S} = \frac{r_e^2 I_0}{r^2} \cdot \frac{1}{2\pi}\int_0^{2\pi} (1 - \sin^2\theta\cos^2\phi)\,\mathrm{d}\phi = \frac{r_e^2}{r^2}\frac{1+\cos^2\theta}{2} I_0. \tag{5.16}$$

在给定的方向 θ 附近，单位立体角的散射功率 $\bar{S}r^2$ 与入射波强度 I_0 之比，称为微分散射截面. 将(5.16)式代入即得汤姆孙微分散射截面：

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{\bar{S}r^2}{I_0} = \frac{r_e^2}{2}(1 + \cos^2\theta). \tag{5.17}$$

它给出了散射的角分布. 当光子能量 $\hbar\omega$ 远小于电子静止能量 m_0c^2 时，与实验结果相符.

现在讨论束缚电子对电磁波的散射. 我们以具有固有角频率 ω_0 的谐振子作为原子内束缚电子的模型. 在入射波电场 $\boldsymbol{E} = \boldsymbol{E}_0 \mathrm{e}^{-\mathrm{i}\omega t}$ 的作用下，电子的运动方程为

$$\ddot{\boldsymbol{x}} + \gamma\dot{\boldsymbol{x}} + \omega_0^2\boldsymbol{x} = (e/m)\boldsymbol{E}_0\mathrm{e}^{-\mathrm{i}\omega t}. \tag{5.18}$$

它与方程(5.3)的区别仅在于束缚电子要加上弹性恢复力 $-\omega_0^2\boldsymbol{x}$. 以(5.2)式代入可求得 $\boldsymbol{x}_0$，从而得到稳态解

$$\boldsymbol{x} = \frac{e}{m}\frac{1}{\omega_0^2 - \omega^2 - \mathrm{i}\omega\gamma}\boldsymbol{E}_0\mathrm{e}^{-\mathrm{i}\omega t} = \frac{e}{m}\frac{1}{\sqrt{(\omega_0^2-\omega^2)^2+\omega^2\gamma^2}}\boldsymbol{E}_0\mathrm{e}^{-\mathrm{i}(\omega t-\delta)}, \tag{5.19}$$

其中 $\tan\delta = \omega\gamma/(\omega_0^2 - \omega^2)$. 由此可求得 $\dot{\boldsymbol{v}} = \ddot{\boldsymbol{x}}$ 的实部为

$$\dot{\boldsymbol{v}} = -\frac{e}{m}\frac{\omega^2}{\sqrt{(\omega_0^2-\omega^2)^2+\omega^2\gamma^2}}\boldsymbol{E}_0\cos(\omega t - \delta). \tag{5.20}$$

类似(5.12)－(5.15)式，可求得平均散射能流、散射功率和散射截面分别为

$$\begin{cases} \bar{S} = \dfrac{e^4E_0^2}{32\pi^2\varepsilon_0c^3m^2r^2}\dfrac{\omega^4}{(\omega_0^2-\omega^2)^2+\omega^2\gamma^2}\sin^2\alpha, & (5.21)\\[2ex] P = \dfrac{8\pi}{3}r_e^2\dfrac{\omega^4}{(\omega_0^2-\omega^2)^2+\omega^2\gamma^2}I_0, & (5.22)\\[2ex] \sigma = \dfrac{8\pi}{3}r_e^2\dfrac{\omega^4}{(\omega_0^2-\omega^2)^2+\omega^2\gamma^2}. & (5.23) \end{cases}$$

当入射波的频率不同时，散射截面的取值分列如下：

(a) 当 $\omega \ll \omega_0$ 时，$\sigma = (8\pi r_e^2/3)(\omega/\omega_0)^4$，即低频时散射截面与 ω^4 成正比，这称为瑞利散射. 在分子散射的情况下，蓝光的强度远远超过红光的强度就是最明显

的例子. 天空之所以呈蓝色就是大气分子散射所造成的.

(b)当 $\omega \gg \omega_0$ 时,$\sigma = 8\pi r_e^2/3$,与自由电子散射的 σ 相同.

(c)当 $\omega = \omega_0$ 时,$\sigma = (8\pi r_e^2/3)(\omega/\gamma)^2$. 由于 $\omega/\gamma = 3\lambda/(4\pi r_e) \gg 1$,因此 $\omega = \omega_0$ 时散射截面远远超出汤姆孙散射截面. 这种当入射电磁波的角频率等于电子固有角频率时散射功率(因而散射截面)出现尖锐的极大值的现象称为共振散射(参看 § 1.4.2 图 1 和 § 4.2.4 图 1).

(二)电磁波的吸收

当具有连续谱的电磁波射到电子上时,只有 $\omega = \omega_0$ 的电磁波被强烈地吸收,从而形成一条吸收谱线.

现在计算电子单位时间内从入射波吸收的能量. 由能量守恒定律可知,当振子振幅稳定后,振子从入射波吸收的能量等于振子辐射的能量. 考虑到电磁波的辐射与 ω 有关,可将(5.22)式改写为

$$P(\omega) = \frac{8\pi}{3} r_e^2 \frac{\omega^4}{(\omega_0^2 - \omega^2)^2 + \omega^2\gamma^2} I_0(\omega), \tag{5.24}$$

其中 $I_0(\omega)$ 为单位频率范围内入射波在单位时间内射到单位面积上的能量. 将(5.24)式对 ω 积分,即可求得单位时间内振子幅射的能量为

$$\int_0^\infty P(\omega)\,\mathrm{d}\omega = \frac{8\pi}{3} r_e^2 \int_0^\infty \frac{\omega^4}{(\omega_0^2 - \omega^2)^2 + \omega^2\gamma^2} I_0(\omega)\,\mathrm{d}\omega. \tag{5.25}$$

由于积分的主要贡献来自 $\omega \approx \omega_0$ 处,故在被积函数中除 $(\omega_0^2 - \omega^2)$ 项外,其余 ω 均换作 ω_0. 于是,积分成为

$$\int_0^\infty \frac{\omega^4}{(\omega_0^2 - \omega^2)^2 + \omega^2\gamma^2}\mathrm{d}\omega = \int_0^\infty \frac{\omega^2\,\mathrm{d}\omega}{(\omega_0 - \omega)^2(\omega_0 + \omega)^2/\omega^2 + \gamma^2}$$

$$\approx \frac{\omega_0^2}{4}\int_0^\infty \frac{\mathrm{d}\omega}{(\omega_0 - \omega)^2 + (\gamma/2)^2} = \frac{\omega_0^2}{4}\int_{-\omega_0}^\infty \frac{\mathrm{d}u}{u^2 + (\gamma/2)^2}$$

$$\approx \frac{\omega_0^2}{4}\int_{-\infty}^{+\infty} \frac{\mathrm{d}u}{u^2 + (\gamma/2)^2} = \frac{\omega_0^2}{2\gamma}\arctan\frac{2u}{\gamma}\Bigg|_{-\infty}^{+\infty} = \frac{\pi\omega_0^2}{2\gamma} = \frac{3\pi c}{4r_e}.$$

上面作了变换 $u = \omega - \omega_0$,并利用 $\omega_0 \gg \gamma$,把下限近似取为 $-\infty$. 代入(5.25)式,即得单位时间内振子辐射的能量(亦即电子单位时间内从入射波吸收的能量)为 $2\pi^2 r_e c I_0(\omega_0)$.

(三)电磁波在稀薄气体中的色散现象

电磁波穿入介质时,传播方向和波长发生变化,称为折射,见 §3.3.2. 从微观的角度看,介质中的电磁波的形成过程是:当电磁波入射到介质中时,引起电子的强迫振动,由于电子作加速运动而辐射出次波,次波的互相叠加便形成了在介质中传播的电磁波. 不同的电磁波在介质中产生不同的次波,使得不同频率的波在介质

中有不同的折射率,从而形成色散现象.

以稀薄气体为例讨论电磁波在介质中的色散现象. 为此就要研究介质的复折射率 $\tilde{n}=c\sqrt{\mu\varepsilon}$ 与电磁波频率 ω 的关系. 对于非铁磁介质 $\mu\approx\mu_0$,而介电常数 ε 通过

$$\varepsilon\boldsymbol{E}=\boldsymbol{D}=\varepsilon_0\boldsymbol{E}+\boldsymbol{P}=\varepsilon_0\boldsymbol{E}+N\boldsymbol{e}\boldsymbol{x}$$

与电子在外电场作用下的位矢 $\boldsymbol{x}$ 相联系(其中 N 为介质中单位体积的电子数). 因此,利用束缚电子运动方程的解 $\boldsymbol{x}$ 即可求得 $\boldsymbol{P}$、ε 及 $\tilde{n}$.

由(5.19)式,可得在入射波电场作用下介质的电极化强度

$$\boldsymbol{P}=N\boldsymbol{e}\boldsymbol{x}=\frac{Ne^2}{m}\frac{1}{(\omega_0^2-\omega^2)-\mathrm{i}\omega\gamma}\boldsymbol{E}. \tag{5.26}$$

代入 $\varepsilon_r\varepsilon_0\boldsymbol{E}=\varepsilon_0\boldsymbol{E}+\boldsymbol{P}$ 可得介质的相对介电常数 ε_r 为

$$\varepsilon_r=1+\frac{Ne^2}{\varepsilon_0 m}\frac{1}{(\omega_0^2-\omega^2)-\mathrm{i}\omega\gamma}. \tag{5.27}$$

由于 ε_r 为复数,故可令 $\varepsilon_r=\varepsilon'_r+\mathrm{i}\,\varepsilon''_r$,由此得

$$\varepsilon'_r=1+\frac{Ne^2}{\varepsilon_0 m}\frac{\omega_0^2-\omega^2}{(\omega_0^2-\omega^2)^2+\omega^2\gamma^2}, \tag{5.28}$$

$$\varepsilon''_r=\frac{Ne^2}{\varepsilon_0 m}\frac{\omega\gamma}{(\omega_0^2-\omega^2)^2+\omega^2\gamma^2}. \tag{5.29}$$

ε'_r 及 ε''_r 对 ω 的依赖关系如图 2 所示.

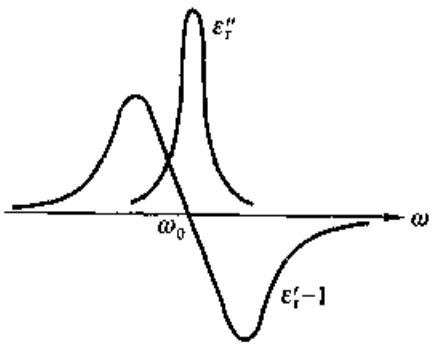


图 2　介质相对介电常数的实部与虚部

复折射率 $\tilde{n}$ 也可以表示为 $\tilde{n}=n+\mathrm{i}\eta$,

$$n+\mathrm{i}\eta=\tilde{n}=c\sqrt{\mu_0\varepsilon_0\varepsilon_r}=\sqrt{\varepsilon_r}. \tag{5.30}$$

将它两边平方,即有

$$n^2-\eta^2+\mathrm{i}2n\eta=\varepsilon_r=\varepsilon'_r+\mathrm{i}\varepsilon''_r. \tag{5.31}$$

令(5.31)式两边实部与虚部分别相等,然后将(5.28)、(5.29)式代入,考虑到原子中的电子可能有多个固有频率 $\omega_i(i=1,2,\cdots)$,其谱线宽度为 γ_i,在单位体积中和固有频率 ω_i 相应的电子数为 Nf_i,便有

$$n^2-\eta^2=1+\sum_i\frac{Ne^2}{\varepsilon_0 m}\frac{f_i(\omega_i^2-\omega^2)}{(\omega_i^2-\omega^2)^2+\omega^2\gamma_i^2}, \tag{5.32}$$

$$n\eta=\sum_i\frac{Ne^2}{2\varepsilon_0 m}\frac{f_i\omega\gamma_i}{(\omega_i^2-\omega^2)^2+\omega^2\gamma_i^2}. \tag{5.33}$$

由此可解出 n 和 η.

对于介质中的电磁波，

$$k = \frac{2\pi}{\lambda} = \frac{\omega}{v} = \bar{n}\frac{\omega}{c} = n\frac{\omega}{c} + \mathrm{i}\frac{\eta\omega}{c}. \tag{5.34}$$

于是，介质中电磁波的电场为

$$\boldsymbol{E} = \boldsymbol{E}_0 \mathrm{e}^{\mathrm{i}(\omega t - kx)} = \boldsymbol{E}_0 \mathrm{e}^{\mathrm{i}\left[\omega t - \frac{n\omega x}{c}\right] - \eta\frac{\omega}{c}x}.$$

其实部为

$$\boldsymbol{E} = \boldsymbol{E}_0 \mathrm{e}^{-\eta\omega x/c}\cos(\omega t - n\omega x/c). \tag{5.35}$$

这说明 $\bar{n}$（或 ε_r）的实部 n 引起色散，而虚部 η 引起吸收.

习　题

1. 点电荷 e 以速度 $\boldsymbol{v} = v\boldsymbol{e}_x$ 运动，在 $t=0$ 时经过原点，问 t 时刻到达 x 点的电磁信号是该电荷在何时发出的？此时点电荷离场点的距离有多远？

2. 已知在带电粒子瞬时静止的参考系 $\tilde{\mathrm{K}}$ 中，带电粒子激发的标势和矢势分别为 §4.2.1(1.3)(1.4)式

$$\tilde{\varphi} = e/[4\pi\varepsilon_0 c(\tilde{t} - \tilde{t}')], \qquad \tilde{\boldsymbol{A}} = 0. \tag{1}$$

试将它扩展成 4 维协变等式，由此得到任意参考系中的矢势和标势 §4.2.1(1.15)、(1.16)式.

提示　首先将 φ 在 $\tilde{\mathrm{K}}$ 系的表达式中的分母 $c(\tilde{t} - \tilde{t}')$ 扩展成 4 维矢量 $r_\mu = (\boldsymbol{x} - \boldsymbol{x}', \mathrm{i}c(t - t'))$. 然后注意 4 维速度矢量 §2.2.6(6.15)式 $u_\mu = \gamma_v(\boldsymbol{v}, \mathrm{i}c)$，在粒子瞬时静止的参考系中，$u_\mu$ 的空间分量为 0，因而可以写

$$A_\mu = -\frac{e}{4\pi\varepsilon_0 c}\cdot\frac{u_\mu}{r_\nu u_\nu}, \tag{2}$$

它在 $\tilde{\mathrm{K}}$ 系中就是(1)式. 在任意参考系中，利用 §2.2.2(2.1)和 §2.2.6(6.15)式，即得 §4.2.1(1,15)、(1,16)式.

3. 有一带电粒子沿 z 轴作简谐振动 $z = z_0 \mathrm{e}^{-\mathrm{i}\omega t}$，设 $z_0\omega \ll c$，

(a)求它的辐射场和能流；

(b)求它的自有场；并比较两者的不同.

4. 带电荷 e 的粒子在 xy 平面上绕 z 轴作匀速圆周运动，角频率为 ω，半径为 R_0. 设 $\omega R_0 \ll c$，试计算辐射场的频率和能流密度，讨论 $\theta = 0, \pi/4, \pi/2$ 及 π 处电磁场的偏振.

5. 设有一各向同性的带电谐振子(无外场时粒子受弹性恢复力 $-m\omega_0^2 r$ 作用)，处于均匀恒定外磁场 $\boldsymbol{B}$ 中，假设粒子速度 $v \ll c$，辐射阻尼力可以忽略，

(a)求振子运动的通解；

(b)讨论沿磁场方向和垂直于磁场方向上辐射场的频率和偏振.

6. 设电子在均匀外场 $\boldsymbol{B}$ 中运动，取磁场 $\boldsymbol{B}$ 的方向为 z 轴方向，已知 $t=0$ 时 $x = R_0, y = z = 0, \dot{x} = \dot{z} = 0, \dot{y} = v_0$. 设满足非相对论条件，求

(a)考虑辐射阻尼力的电子运动轨道；

(b)电子单位时间内辐射的能量.

7. 等离子体由带电粒子组成，第 i 种粒子的数密度为 N_i，带电量为 Z_ie，求其相对介电常数 ε 及电磁波能通过等离子体而不被吸收的极限频率.



第五篇　微观粒子的运动规律

以上几篇讨论了宏观世界的物理规律，如果企图用这些规律研究微观世界的物理现象，就会发现它们的局限性. 在这一篇和下一篇里，我们讨论微观粒子的运动.

微观粒子是指电子、质子、中子、介子等这类粒子. 自从发现电子以来，已发现了数百种微观粒子. 研究表明，微观粒子在 $10^{-8}\sim10^{-13}$ cm 范围内，即在分子、原子、原子核范围内的运动服从量子力学.

量子力学是研究分子、原子、原子核以及固体、液体的理论基础. 凡是从微观的观点来研究自然的任何一种学科，都要应用量子力学，如量子光学、量子化学、量子生物学等.

在这一篇里，我们建立量子力学的基本规律；下一篇讨论微观粒子在外场中的运动.

第一章　微观粒子状态的描述

微观粒子之所以有不同于宏观物体的运动规律，归根到底是由微观粒子本身的性质决定的. 历史上正是通过揭示微观粒子的波-粒二象性而逐步建立起量子力学. 在这一章，将介绍 19 世纪末和 20 世纪初经典物理所遇到的困难，以及人们在克服这些困难的过程中如何认识到微观粒子有不同于宏观物体的性质. 然后，再根据微观粒子的性质，说明应该用怎样的数学工具描述微观粒子的运动状态.

§5.1.1　光的粒子性

人们认识微观粒子的二象性是从揭示光的二象性开始的. 历史上对光的本性曾经有过粒子说和波动说的长期争论. 麦克斯韦创立光的电磁理论，证明光是电磁波，使人们对光的波动说几乎坚信不移. 但是，当一系列有关光的辐射和吸收的新实验事实被发现以后，研究深入到需要用微观机制来解释这些结果，就暴露了光的波动理论的局限性，并导致了人们对光的性质的进一步认识. 这大致经历了三个阶段.

(一)黑体辐射的研究　"能量子"的提出

一个能将投射在它上面的全部电磁波吸收的物体叫绝对黑体. 自然界没有真正的黑体. 实验室中的黑体是用不透明材料制成的，开有小孔的空腔. 因为通过小孔进入空腔内的电磁波经过多次反射才有可能重新从小孔射出，而在多次反射过程中，电磁波几乎已被全部吸收，故这一空腔的小孔是能将电磁波完全吸收的黑体.

当空腔中的辐射场与周围物体处于平衡态时，从小孔发出的辐射称为黑体辐射，其能量密度按频率的分布 $\rho_\nu \mathrm{d}\nu$ 的实验结果如图 1 中的实线.

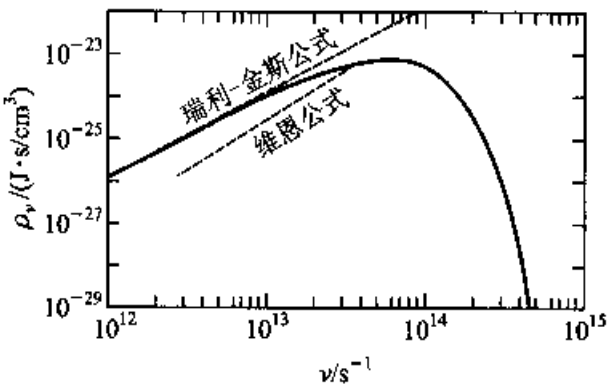


图 1　黑体辐射能量按频率的分布，实线是实验结果，虚线是维恩公式和瑞利-金斯公式的计算结果.

我们现在感兴趣的不是讨论黑体辐射的细节，而是把注意力集中在：用经典物理来解释这些结果会遇到什么困难，从而会揭示出什么新的物理概念.

历史上，维恩和瑞利－金斯曾经用经典物理理论推出了黑体辐射能量分布的

两个公式：

$$\text{维恩公式：}\qquad \rho_\nu \mathrm{d}\nu = c_1 \nu^3 \mathrm{e}^{-c_2\nu/T}\mathrm{d}\nu; \tag{1.1}$$

$$\text{瑞利－金斯公式：}\qquad \rho_\nu \mathrm{d}\nu = (8\pi\nu^2/c^3) k_B T \mathrm{d}\nu. \tag{1.2}$$

其中 k_B 是玻尔兹曼常量，T 是温度. 这两个公式分别只在辐射频率的较高和较低部分与实验符合，见图 1❶，但都不能解释中间频段的实验结果，这暴露了经典物理的困难.

摆脱这种困境的道路是：先找到圆满符合实验结果的经验公式，比较它和经典理论导出的公式的差异，分析这种差异所包含的物理实质，进而看出经典物理困难的症结，提出新的物理概念. 在这条道路上，普朗克于 1900 年符合出了正确的经验公式（即图 1 中的实线）：

$$\rho_\nu \mathrm{d}\nu = \frac{8\pi\nu^2}{c^3} \cdot \frac{h\nu}{\mathrm{e}^{h\nu/(k_B T)} - 1}\mathrm{d}\nu. \tag{1.3}$$

式中 h 为普朗克常量，见附录Ⅶ.

为了比较普朗克公式与经典物理导出的公式在物理实质上的差异，我们来分析瑞利－金斯公式的推导过程. 这个推导有关键性的两步，一是将空腔内的电磁场分解为各种频率的简正振动，求得空腔单位体积内频率在 $\nu \sim \nu + \mathrm{d}\nu$ 之间的振动方式的数目为 $(8\pi\nu^2/c^3)\mathrm{d}\nu$［见 §3.3.4(4.37)式］；二是根据经典统计物理的能量均分定理知，每一振动方式的平均能量为 $k_B T$，见 §8.1.4. 这样，密度分布就是(1.2)式. 将它和普朗克公式比较，明显看出两者的差异在于对每个振子平均能量的认识. 普朗克的公式表明，每个振子的平均能量不是 $k_B T$，正是从这里开始了对经典物理的突破.

每个振子的平均能量为 $k_B T$ 这个结论的得出，是基于经典物理中一个基本的概念，即：能量可以连续取值. 这一基本概念在证明能量均分定理时要用到. 在普通物理中用分子运动论证明过，气体分子运动的每个自由度分得平均能量 $k_B T/2$. 普遍的能量均分定理［见 §8.1.4(4.3)式］指出，在能量表示式中的每一个关于正则变量的平方项，对能量的平均值贡献 $k_B T/2$. 对于振子，$E = p^2/(2m) - m\omega^2 q^2/2$ 包含两个平方项，故其平均值为 $k_B T$. 特别要注意的是，证明能量均分定理的关键步骤［见 §8.1.4(4.8)式］是将每一个平方项都当成 p 和 q 的连续函数来进行积分. 因而，如果振子的平均能量不是 $k_B T$，就要动摇经典物理中关于能量总是连续变化，这样一个最基本的观念. 在 19 世纪，经典物理被认为是很完美的，要作出这

---

❶ 维恩公式(1.1)中的常数 c_1、c_2 不能由经典理论导出，只能通过符合实验来确定. 在图 1 中用的是 $c_1 = 8\pi h/c^3$，$c_2 = h/k_B$.

种突破很不容易.就是普朗克本人,也觉得是"很痛苦的".他知道,要得到正确的结果,必须有这个突破.但经典物理紧紧束缚着他,使他称自己作出的这个突破是一个"绝望的行动".他甚至在得到正确结果后,又花费了十年时间,试图仍将其结果纳入经典物理的框架.当然这是徒劳无益的,最终不得不相信必须有这个突破.

为得出振子平均能量的正确结果,普朗克提出了以下基本假定:频率为 ν 的谐振子,其能量是不连续变化的,只取一些分立的数值;这些分立值都是能量最小单元 $h\nu$ 的整数倍.能量的这种最小单元称为能量子.

在普朗克的假定中有两个要点:一个是谐振子能量的分立性,另一个是存在最小能量单元.从普朗克的假定出发可以证明[参看 §8.2.1(1.10)式],谐振子的平均能量是 $E = h\nu/(e^{h\nu/(k_BT)} - 1)$,从而就自然地得到了普朗克公式(1.3)式.

普朗克假定的成功,揭示出光的能量具有分立性和对于频率为 ν 的光存在有能量最小单元,即能量子 $h\nu$.“量子”这个词正是从此问世的.普朗克的上述假定在物理学发展史上是一个新的里程碑,从它开始了量子论的创立.

(二)光电效应的研究　“光量子”的提出

光电效应是指光照射到金属表面时有电子从金属中逸出的现象.光电效应的主要实验结论是:

①对一定物质而言,只有入射光的频率高于某个阈值时,才有光电子产生,并且光电子几乎是在受光照射后立即产生的(滞后时间约为 10^{-9}s);

②光电子的最大动能与引起它发射的光的频率成正比,与入射光的强度无关;

③光电子的密度与入射光的强度成正比.

经典物理无法解释光电效应.按照经典电磁理论,构成光波的电磁场使物质中的电子受迫振动,电子得到的能量取决于入射光的强度.入射光愈强,电子获得的能量愈大.获得足够大能量的电子能支付物质的逸出功而逸出成为光电子.即使入射光很弱,只要经过一段时间积累,电子也能获得足够大的能量,而仍会有光电子产生.经典物理的这些结论与实验不符.

光电效应是光被物质吸收产生的物理效应.在普朗克研究光辐射时提出的假定——物质辐射光的能量不连续,存在最小能量单元,的启示下,可以设想光被物质吸收时,其能量也具有分立性和存在能量最小单元.子是就有了在 1905 年由爱因斯坦提出的著名假设:光是粒子流;和频率为 ν 的光对应的每个粒子的能量为 $h\nu$,它是光的结构单元,称为光量子,或光子.

按照光量子假设,光电效应可以得到圆满解释:当光照射到物质上时,能量为 $h\nu$ 的光子被电子吸收,电子吸收的这份能量一部分消耗子逸出功 W,其余的成为光电子的动能:

$$mv^2/2 = h\nu - W, \tag{1.4}$$

这称为爱因斯坦公式.从这一公式可以直接看出:光电子的初动能由入射光的频率决定;对于给定的物质,存在一个产生光电效应的频率阈值,大小为 $W/\hbar$;光子的能量是一次被电子吸收,只要入射光的频率足够大,光电子就会立即产生.

对光电效应解释的成功,使人们对光的本性的认识前进了一步,认识到光的粒子性,即:光被物质发射和吸收时,或一般地说,光与物质相互作用时,表现为由光子构成的粒子流.

光子的速度即光速,因而按照相对论,它是静质量为零的粒子.根据相对论还知道,一个静质量为零,能量为 E 的粒子,其动量为 $p = E/c$ [见 §2.3.1(1.24)式],因而光子的动量为 $p = h\nu/c = h/\lambda$.

关系式

$$E = h\nu, \qquad p = h/\lambda \tag{1.5}$$

把具有确定频率和波长的波与具有确定能量和动量的粒子联系起来,反映了波动性和粒子性的联系,称为爱因斯坦关系式.

引入圆频率 $\omega = 2\pi\nu$,波矢量 $\boldsymbol{k} = \boldsymbol{e}_n(2\pi/\lambda)$,其中 $\boldsymbol{e}_n$ 是单位矢量,并令 $\hbar \equiv h/2\pi$,则爱因斯坦关系式可以写成:

$$E = \hbar\omega, \qquad \boldsymbol{p} = (h/\lambda)\boldsymbol{e}_n = \hbar\boldsymbol{k}. \tag{1.6}$$

(三)康普顿效应的研究 光的粒子性的进一步证实

光的粒子性在 1923 年研究康普顿效应时以非常直观的图像得到了进一步的证实.当波长为 λ 的 X 射线被石墨靶中的电子散射时,康普顿发现,在散射的 X 射线中,除波长不变的成分外,还有波长大于入射波长的成分 λ',在不同散射方向观察到的波长偏移 $\Delta\lambda = \lambda' - \lambda$ 不相同.

经典物理不能解释波长偏移.因为按照经典电磁理论,频率为 ν 的入射波的电场强迫散射体中的电子作相同频率的振动,受迫振动的电子只能辐射出相同频率的电磁波.这样,散射波的波长只能和入射波的波长相同.

如果把入射的 X 射线看成粒子流,X 射线被电子散射的过程看成是光子和电子这两种粒子的碰撞过程,则问题可以得到正确的解释(参看第一篇第三章习题 6).因为在碰撞过程中,入射光子把它的一部分能量传递给电子,散射后一定有能量比入射光子能量低的散射光子.按照爱因斯坦关系式,这意味着这些光子的波长比入射光子的波长要长.

考虑一个光子和一个电子的碰撞过程.假定电子在碰撞前处于静止状态.由于 X 射线的波长短,频率高,其光子的能量和电子的静能可以相比,碰撞后反冲,电子的速度和光速相比不能忽略,因而对于电子的能量和动量应该用相对论公式.碰撞前光子的频率为 ω,碰撞后变为 ω',由能量守恒:

$$\hbar\omega + mc^2 = \hbar\omega' + mc^2/\sqrt{1-(v/c)^2},$$

m 为电子的静质量. 由动量守恒：

$$\begin{cases}(\hbar\omega/c) = (\hbar\omega'/c)\cos\theta + (mv/\sqrt{1-(v/c)^2})\cos\theta', \\ 0 = (\hbar\omega'/c)\sin\theta - (mv/\sqrt{1-(v/c)^2})\sin\theta'. \end{cases}$$

消去 v 和 θ'，解得 $\omega - \omega' = [2\hbar/(mc^2)]\omega\omega'\sin^2(\theta/2)$，或

$$\Delta\lambda = \lambda' - \lambda = [4\pi\hbar/(mc)]\sin^2(\theta/2). \tag{1.7}$$

这一公式圆满地解释了实验观察到的 $\Delta\lambda$ 和 θ 的关系. 这样，就以粒子碰撞这种直观的粒子相互作用机制显示出了光和物质相互作用时呈现出的粒子性.

光在空间传播表现为电磁波，服从波动方程. 这种波动性的现象性特征是光的干涉、衍射. 光的波动性很好地描述了光的传播，但不能解释光与物质的相互作用. 光与物质相互作用时表现出粒子性，服从粒子间的相互作用机制. 这就是光的波粒二象性.

光的二象性被揭示，不仅使人们对光的性质的认识前进了一步，而且在人类物理思想的发展上起了巨大的启示作用. 在 20 世纪初，当经典物理在讨论原子问题遇到困难时，正是在光的量子论的启发下，使人们认识到粒子的波动性，并最终导致了量子力学的诞生.

§5.1.2　粒子的波动性　波函数

(一)原子光谱的分立性　玻尔量子化假设

在历史上人们早就认识到原子光谱的分立性. 以氢光谱为例，每条谱线的波数 $\tilde{\nu}$（它和频率的关系为 $\nu = c\tilde{\nu}$）可以统一表示为

$$\tilde{\nu} = R_H(m^{-2} - n^{-2}), \tag{2.1}$$

式中 $m = 1, 2, \cdots; n = m+1, m+2, \cdots; R_H$ 为里德伯常量，见附录Ⅶ. 由于 m 和 n 的取值都是分立的，所以该式清楚地反映出氢光谱的分立结构. 原子光谱的分立性也是经典物理所不能解释的. 例如，假定电子绕核作圆周运动，则其辐射频率就是 $\nu = v/(2\pi r)$. 可见 ν 随 r 连续变化，而且电子会因为不断辐射电磁波，消耗能量，最终掉到原子核上（参看 §4.2.2 及第四篇第一章习题 5）.

经典物理遇到的上述困难表明，不能简单地把经典物理应用到原子这样的微观领域内. 对于微观领域发生的过程必须从实验事实出发，提出新概念，建立新理论.

注意到，按照普朗克的量子概念，频率为 ν 的光，其能量子 $h\nu$；若把式(2.1)两边乘以 hc，则有 $h\nu = R_H hc(m^{-2} - n^{-2})$. 该式左边 $h\nu$ 是光子的能量，是原子辐射出来的，因而右边就应是原子在辐射前的能量 E_i 和辐射后的能量 E_f 之差：

$$h\nu = E_{\mathrm{i}} - E_{\mathrm{f}}. \tag{2.2}$$

这是将普朗克量子论用来分析原子光谱的分立性时自然得到的结果. 它揭示出原子光谱的不连续性是原子能量状态不连续性的反映. 正是从这种分析出发,玻尔提出了如下假定:

①电子并不是沿经典力学所允许的任一轨道运动,而只能沿满足量子化条件的轨道运动. 量子化条件为

$$\oint p_\alpha \mathrm{d}q_\alpha = n_\alpha h, \tag{2.3}$$

n_α 为正整数,q_α 和 p_α 为粒子的广义坐标和广义动量,α 标记粒子的空间自由度. 在满足(2.3)式的轨道上,电子的能量有确定值,处于稳定状态,称为定态. 注意,(2.1)式右边的积分是相空间中和相应轨道对应的椭园面积,见第一篇第六章习题 3.

②当电子从一个定态过渡到另一个定态时,会辐射或吸收光,其频率为

$$\nu = (E_n - E_m)/h. \tag{2.4}$$

玻尔的假定在物理学史上曾获得很大的成功. 它能很好地解释氢原子光谱的规律和多电子原子光谱的一些定性特征. 但是,玻尔理论有其局限性,它不能计算谱线的强度;对于多电子原子,不能给出光谱的定量结果.

玻尔理论存在困难是可以想像的,因为这个理论本身不是自洽的,缺乏逻辑的连贯性. 玻尔理论的结构是"经典力学加量子化条件". 一方面它引入了和经典物理不相容的量子概念,即量子化条件;另一方面它又保留了经典力学的基本观念,即粒子作轨道运动. 尽管如此,玻尔理论在量子论的发展史上仍然起了承前启后的作用. 既然它的成功之处在于它沿用了光的量子论思想,提出了量子化条件,那么它的困难之处,很可能在于它不是一贯的量子理论,而保留了粒子作轨道运动的经典概念.

(二)德布罗意波

怎样才能突破粒子作轨道运动的概念呢? 联想到人们对光的二象性的认识过程,德布罗意于 1923 年这样提出了问题,他说:"在光学上,比起波动的研究方法来,是过于忽略了粒子的方法;在物质理论上,是否发生了相反的错误呢? 是不是我们把粒子图像想得太多,而过于忽略了波的图像?"于是他提出粒子具有波动性的假设. 他推广描述光的二象性的爱因斯坦关系式,认为具有一定能量和动量的粒子,其运动可以用一定频率和波长的波来描述,它们之间的联系为

$$E = h\nu = \hbar\omega, \qquad \boldsymbol{p} = (h/\lambda)\boldsymbol{e}_n = \hbar\boldsymbol{k}, \tag{2.5}$$

这称为德布罗意关系式,λ 称为德布罗意波长. 应该注意的是,爱因斯坦关系式[§5.1.1(1.6)式]是描述以光速运动的光子的二象性,而经德布罗意推广后得到的德布罗意关系(2.5)式,对以任何速度运动的粒子都适用. 也就是说,关系式 E

$= \hbar\omega, p = \hbar k$ 是微观粒子的波动性和粒子性之间联系的反映，它无论对相对论性的粒子或非相对论性的粒子都成立.

(三)电子衍射实验

一个假设的正确与否，不在于它是否和原有的理论相协调，也不只在于它提出时所基子的根据，最重要的在于它能否被实验所检验. 干涉和衍射是波动性的特征，如果德布罗意的假设是正确的，当粒子流通过线度和德布罗意波长可以相比较的障碍物时，就应该呈现出干涉、衍射条纹.

现在来估算一下在通常情况下粒子的德布罗意波长，以便选择适当的"障碍物". 由(2.5)式得

$$\lambda = h/p. \tag{2.6}$$

对子非相对论的自由粒子 $E = p^2/(2m)$，因而有

$$\lambda = h/\sqrt{2mE}. \tag{2.7}$$

如果是通过电势差 V 加速电子，使其获得能量 $E = eV$，则电子的德布罗意波长为

$$\lambda = h/\sqrt{2meV} \approx \sqrt{150/V}\text{Å}. \tag{2.8}$$

可见，当 $V = 150$ V(伏)时，$\lambda = 1$ Å(埃)；当 V 更高时，λ 更短. 即知在通常情况下，电子的波长有 Å 的量级，它可以和晶体中格点的间距相比较，因而可以选用晶片作"光栅"，观察电子德布罗意波的衍射. 戴维逊和革末首先按照这种设想成功地观察到电子的衍射现象.

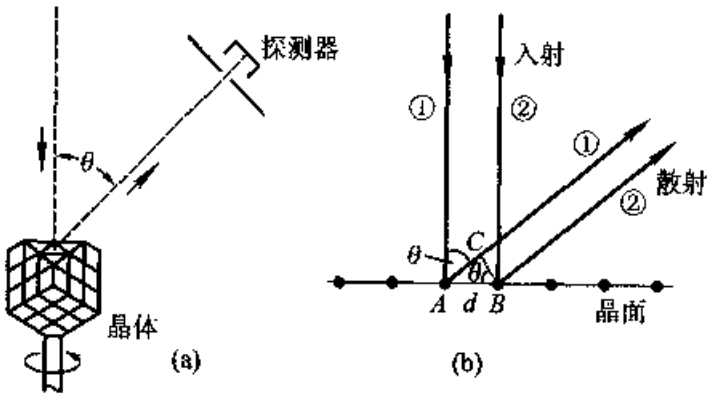


图 1　电子在晶面上的衍射

图 1(b)是电子在晶面上衍射的示意图. 设晶格原子的间距为 d，电子垂直入射到晶面上，并在 θ 方向进行观察. 电子被相邻原子 A、B 散射后，散射波的程差为 $AC = d\sin\theta$. 这些散射波是相干波，因而在 θ 方向可以观察到衍射现象. 根据干涉原理，当程差为波长 λ 的整数倍，即在

$$d\sin\theta = n\lambda \tag{2.9}$$

时，应出现衍射强度的极大值.

图 1(a)是戴维逊－革末实验的示意图. 转动探测器和晶体可以观察各个方向的衍射结果. 实验表明，衍射强度随 θ 的变化和波动理论给出的(2.9)式一致，从而证实了电子的波动性.

如果衍射晶体不是一整块单晶体，而是由许多微小晶体组成的多晶体，则由于

微小晶体的取向是混乱的，电子透过多晶薄膜后将沿入射方向轴对称地分布．这样，采用照相底板记录衍射结果，得到的是衍射环．用类似的方法，人们还观察到了其他粒子（例如中子）的衍射现象，从而从实验上充分证实了德布罗意波假设的正确性．

（四）微观粒子的波粒二象性

为了进一步理解微观粒子的波粒二象性，我们来分析电子衍射实验的过程．假定电子流很弱，以致电子是一个一个地通过“光栅”，则在底板上看到的将是一个一个不相重叠的斑点．这说明，当电子与底板上的原子相互作用时，体现出粒子性，作为一个完整的粒子进行能量、动量、电荷的交换．但它不是经典力学描述的粒子，即不按经典力学轨道运动．按照经典力学，在同样实验条件下的电子，由于其受力情况以及初始条件和边界条件都相同，应该有同样的轨道，因而底板上的斑点就应重合．然而实际上，电子在底板上的斑点并不重合，而是在底板的各处都出现．如果入射粒子流很弱，在开始时底板上会呈现零散的斑点，似乎没有一定的规律；但只要延长曝光时间，在底板上就会看到斑点有规律的分布，呈现为衍射图样．这种衍射图样表明，电子在传播过程中呈现出干涉、衍射这样一种“波”的特征行为．因此，从这一实验可以看到，电子（以及其他微观粒子）和光子一样，具有波粒二象性——在和物质相互作用时呈现出粒子性，服从粒子间的相互作用机制；而在空间中传播时，呈现出波动性，会发生干涉、衍射．

电子和其他微观粒子既然是“粒子”，怎么会有波动性呢？这初看起来似乎很难理解．仔细想想并不奇怪．我们可以反过来提一个问题，你又怎么知道电子是“粒子”呢？回顾历史，最早有迹象表明存在电子这样一种“粒子”的根据，是法拉第电解定律证实有一个最小的不可再分的电荷量存在．以后，密立根油滴实验证实，悬浮在电容器的电场中的油滴所改变的电荷，永远是这同一个最小电荷量的整数倍．这就表明，存在一种带有这个最小电荷量的粒子——电子．以后，利用质谱仪成功地测出了电子的质量，进一步证实了带有确定的质量和电荷的粒子——电子的存在．仔细考察这些实验，可以看出，没有任何一个实验看到了电子像一个经典小球那样在空间飞行．这些实验只不过是发现了电子在和物质相互作用时呈现出粒子性——一份份地交换能量、动量和电荷．另一方面，电子衍射实验表明，电子在传播过程中呈现出波动性．因此，电子作为一种微观客体有它自己的本性——简单说来就是波粒二象性．它并不是经典意义下的粒子，只不过是在和物质相互作用时呈现出粒子性；它也不是经典意义下的波，只不过是在传播过程中具有干涉、衍射这类波的性质．

（五）波函数

既然波粒二象性是微观粒子的本性，描述微观粒子状态的数学工具就应该能

反映这种性质.

我们知道,任何一种波都是用 r 和 t 的某种函数来描述. 例如,电磁波用电场和磁场作为 r、t 的函数 $\boldsymbol{E}=\boldsymbol{E}(r,t)$,$\boldsymbol{H}=\boldsymbol{H}(r,t)$,或四维矢势作为 r、t 的函数 $A_\mu=A_\mu(r,t)$ 来描述;声波以空气密度作为 r、t 的函数 $\rho=\rho(r,t)$ 来描述. 描述不同波的波函数当然不同,但它们有一个共同性质:满足叠加原理——两个波在空间中传播时,它们在空间每一点的波函数可以叠加. 这是波会发生干涉、衍射的根本原因.

既然微观粒子具有波动性,会发生干涉、衍射,它的状态就应该用一种满足叠加原理的波函数来描述. 我们用 $\Psi(r,t)$ 表示描述微观粒子状态的波函数. 也称为态函数.

然而,微观粒子并不是经典波. 当它和物质(包括测量仪器)相互作用时,不是像经典波那样弥散在全空间,而是呈现为一份份集中的能量、动量和电荷. 因此,描述微观粒子状态的波函数 $\Psi(r,t)$ 本身不是可观测量. 它的作用只在于用它可以对微观粒子的各种力学量(坐标、动量、能量、角动量等等)的观测结果作出预言.

为了说明利用波函数 $\Psi(r,t)$ 能对微观粒子的力学量的观测结果作出怎样的预言,让我们回到前面讨论过的电子衍射实验. 前面已经指出,如果入射流强不大,在开始时,板上出现的是一个个杂乱分布的斑点. 这表明,对于电子将在板的什么位置上出现,不能作出唯一的回答. 这一点并不令人奇怪,因为正是它体现了电子作为一种微观粒子,具有波粒二象性的特点. 如果对于电子将在板的什么位置上出现能作出唯一的回答,那将意味着电子沿确定的轨道运动,和经典小球没有什么区别.

前面还曾指出,经过足够长的时间,板上会出现衍射图样. 在这种衍射图样中的深条纹处的斑点密,表明电子打到板的这种位置的概率大;浅条纹处的斑点稀,表明电子打到这种位置的概率小. 整个衍射图样表示电子在板的不同位置出现的概率分布. 正是这种不同位置的概率分布是波函数 $\Psi(r,t)$ 所应该描述而且能够描述的.

任何一种波的强度都是正比于相应的波函数振幅的平方,因此,微观粒子出现在空间不同位置的概率分布应该用波函数模的平方 $|\Psi(r,t)|^2=\Psi^*(r,t)\Psi(r,t)$ 来描述❶. 具体说来,微观粒子在 t 时刻,在 r 处的体积元 d^3r 中出现的概率是

$$\mathrm{d}P(r,t)=|\Psi(r,t)|^2\mathrm{d}^3r=\Psi^*(r,t)\Psi(r,t)\mathrm{d}^3r. \tag{2.10}$$

❶ 由于 $\Psi(r,t)$ 本身不代表可观测量,因而它一般说是取复数值的函数,没有理由限制它只能取实数值. 要注意,$\Psi(r,t)$ 取复值和描述经典波的函数写成复数形式[例如§3.3.1(1.29)式]有根本区别. 对于经典波,写成复数形式只是为了使于运算,在得到最后结果时还要取实部或虚部,从而回到实函数. 而微观粒子的波函数 $\Psi(r,t)$ 本质上就是一个取复值的函数.

根据相斥事件概率相加的原理，在一个体积 V 中找到粒子的概率等于在 V 所包含的各个体积元 $d^3\boldsymbol{r}$ 中找到粒子的概率之和，因而可以由(2.10)式积分得到在 V 中找到粒子的概率：

$$P_V(t) = \int_V |\Psi(\boldsymbol{r},t)|^2 d^3\boldsymbol{r}. \tag{2.11}$$

如果将 V 取为全空间，则找到粒子是必然事件，相应的概率为 1，因此

$$\int_{全空间} |\Psi(\boldsymbol{r},t)|^2 d^3\boldsymbol{r} = 1. \tag{2.12}$$

这是波函数所应满足的归一化条件. 如果给定的 $\Psi(\boldsymbol{r},t)$，例如通过解运动方程得到的 $\Psi(\boldsymbol{r},t)$，不满足这一条件，可以在它上面乘一个常数 A（我们取 A 为实常数），使 $A\Psi$ 满足(2.12)式，

$$\int (A\Psi)^* (A\Psi) d^3\boldsymbol{r} = A^2 \int \Psi^* \Psi d^3\boldsymbol{r} = 1. \tag{2.13}$$

这样的一个常数因子称为归一化常数. 将它乘到 $\Psi(\boldsymbol{r},t)$ 上，不改变在空间不同位置找到粒子的相对概率，因而是可以允许的.

即使在 $\Psi(\boldsymbol{r},t)$ 已经归一化以后，仍然可以在它上面乘一个模为 1 的复常数 $e^{i\delta}$，因为 $|e^{i\delta}\Psi|^2 = |\Psi|^2$，所以两者描述同一物理状态：

$$\Psi(\boldsymbol{r},t) \quad 和 \quad e^{i\delta}\Psi(\boldsymbol{r},t) \quad 描述同一状态, \tag{2.14}$$

其中 δ 是一个任意的常数相位.

【例 1】 写出自由运动粒子的波函数.

解 质量为 m 的自由粒子有确定的能量 E 和动量 $\boldsymbol{p}$

$$E = \boldsymbol{p}^2/(2m). \tag{2.15}$$

根据德布罗意关系(2.5)式，它有确定的频率 ω 和波矢 $\boldsymbol{k}$

$$\omega = E/\hbar, \qquad \boldsymbol{k} = \boldsymbol{p}/\hbar. \tag{2.16}$$

这样的波是单色（ω 为常数）平面（$\boldsymbol{k}$ 为常矢量）波，单色平面波的波函数应是 $\boldsymbol{k}\cdot\boldsymbol{r} - \omega t = (\boldsymbol{p}\cdot\boldsymbol{r} - Et)/\hbar$ 的周期函数. 考虑到 Ψ 取复值就有

$$\Psi_p(\boldsymbol{r},t) = Ae^{i(\boldsymbol{p}\cdot\boldsymbol{r}-Et)/\hbar}. \tag{2.17}$$

这就是描述自由粒子的波函数，称为德布罗意波.

德布罗意波不能按(2.12)式归一化. 因为由(2.17)式，$|\Psi|^2 = A^2$ 是常数，表明在空间任何位置上找到粒子的概率都相同. 这正是“单色平面波”应有的性质. 但是，如果将这样的 $|\Psi|^2$ 对全空间积分，会得到无穷大. 如果想给它找一个满足(2.13)式的归一化常数 A，就会得到 A = 0. 出现这一矛盾的原因是，实际的粒子不可能在全空间中自由运动，而只能在某一个空间范围内运动. 用 V 表示这一空间范围的体积，则 $A = 1/\sqrt{V}$.

例如，可以选 V 为长宽高都是 L 的一个立方体，限制粒子在

$$0 \leqslant x \leqslant L,\quad 0 \leqslant y \leqslant L,\quad 0 \leqslant z \leqslant L \tag{2.18}$$

中运动，并取周期性边界条件❶

$$\begin{cases} \Psi_p(0,y,z,t) = \Psi_p(L,y,z,t)\ , \\ \Psi_p(x,0,z,t) = \Psi_p(x,L,z,t)\ , \\ \Psi_p(x,y,0,t) = \Psi_p(x,y,L,t)\ . \end{cases} \tag{2.19}$$

代入(2.17)式可见，这一边界条件限制了 p 只能取以下分立值：

$$\left.\begin{aligned} p_x = n_x 2\pi\hbar/L,\quad p_y = n_y 2\pi\hbar/L,\quad p_z = n_z 2\pi\hbar/L, \\ n_x, n_y, n_z = 0, \pm 1, \pm 2, \cdots. \end{aligned}\right\} \tag{2.20}$$

此时，归一化条件(2.13)式成为 $A^2\int_0^L \mathrm{e}^{-\mathrm{i}(\boldsymbol{p}\cdot\boldsymbol{r}-Et)/\hbar}\mathrm{e}^{\mathrm{i}(\boldsymbol{p}\cdot\boldsymbol{r}-Et)/\hbar}\mathrm{d}^3\boldsymbol{r} = A^2L^3 = 1$，因而，$A = L^{-3/2} = 1/\sqrt{V}$. 代入(2.17)式

$$\Psi_p(\boldsymbol{r},t) = L^{-3/2}\mathrm{e}^{\mathrm{i}(\boldsymbol{p}\cdot\boldsymbol{r}-Et)/\hbar}. \tag{2.21}$$

按这种方式的归一化称为箱归一化.

在量子力学中也常常考虑在无限空间中运动的自由粒子. 实际上，当体积 V 有宏观大小时，对于微观粒子就可以看成是无限空间. 此时，波函数要用一种和(2.12)、(2.13)式不同的方式归一化，这将在后面介绍.

(六)波函数的叠加原理

"波"和"粒子"的不同在于波弥散分布在空间中，但是并不"占据"这一空间区域. 两个或更多的波可以同时存在于同一空间区域. 当两个波传播到同一空间区域时，会发生叠加，相互加强或削弱. 这就是体现出波的特性的干涉现象. 由于微观粒子有波动性，其波函数具有叠加性，即：如果 $\Psi_1(\boldsymbol{r},t)$ 和 $\Psi_2(\boldsymbol{r},t)$ 是描述系统两个可能状态的波函数，则

$$\Psi(\boldsymbol{r},t) = c_1\Psi_1(\boldsymbol{r},t) + c_2\Psi_2(\boldsymbol{r},t) \tag{2.22}$$

也是一个描述系统状态的波函数. 其中 c_1 和 c_2 是两个(复)常数. 这称为波函数的叠加原理.

习　题

1. 计算下列各种频率的谐振子的能量子：

(a) $\nu = 50\mathrm{Hz}$ 的带电谐振子；

(b) $\nu = 10^{10}\mathrm{Hz}$ 的微波；

(c) $\nu = 10^{15}\mathrm{Hz}$ 的光波，

❶ 边界条件有不同取法，例如可以令边界上 $\Psi = 0$. 但是，周期性边界条件(2.19)式更便于计算.

进而指出为什么普通振子的能量不显分立性.

2. 计算下列粒子的德布罗意波长:

(a)能量为 100eV 的自由电子;

(b)被 10^{15} V 电势差加速的质子;

(c)以速度 $1m \cdot s^{-1}$ 运动的质量为 1kg 的质点;

(d) $T = 300$ K 时,以麦克斯韦分布[见 §8.1.3(3.7)式]的最概然速率运动的中子.

进而指出为什么在普通条件下观察不到粒子的波动性.

第二章　力学量的算符

上一章，通过分析微观粒子的波粒二象性，说明了微观粒子的状态用波函数 $\Psi(r,t)$ 描述. 利用 $\Psi(r,t)$ 可以知道测量粒子位置 r 得到不同结果的概率分布［§5.1.2(2.10)式］. 这一章将讨论反映微观粒子动力学性质的力学量应该用怎样的数学工具描述，从而使理论能够预言在给定状态里各种不同力学量的实验观测结果.

在经典力学里，在给定状态对质点的力学量进行观测，能得到确定的数值. 力学量的数值和状态之间有确定的对应关系. 描述状态是用正则变量，即：坐标和动量，而实际观测的力学量都是它们的函数，如能量 $E = p^2/(2m) + U(r)$，从而使理论能够对子在给定状态里力学量的观测结果，给出唯一确定的预言.

对于微观粒子，由于波粒二象性，在给定状态里测量力学量，通常不是得到唯一的结果，而是有一定概率分布的一系列可能值，因而在量子力学里，力学量和描述状态的波函数之间不是简单的对应关系. 这一章就来讨论，如何为微观粒子力学量的描述找到适当的数学工具. 通过它对波函数的运算能够预言在给定状态里力学量的观测结果.

§5.2.1　测量结果的期望值　用算符表示力学量

对于一个微观粒子系统，在给定的状态里，力学量以不同概率取不同值. 因此，对力学量的完整描述，首先要知道它所能取的值的谱，即各种可能观测到的值；还要能预言它在给定状态里取不同值的概率. 我们把对这一问题的全面回答留到下一节. 在这一节里，先从一个比较简单些的问题开始，即：如何决定在给定状态里测量力学量的期望值，或平均值. 正如在附录Ⅲ中指出的，数学期望值可以集中反映随机变量的统计特征，因此，掌握在给定状态里测量物理量的期望值的计算方法，对于量子系统的动力学是很重要的.

(一)在给定状态里坐标的平均值

为了简单起见，讨论一维情况，所得结果不难推广到三维.

根据§5.1.2(2.10)式，测量坐标 x 得值在 $x \sim x + \mathrm{d}x$ 之间的概率是

$$\mathrm{d}P(x,t) = |\Psi(x,t)|^2 \mathrm{d}x = \Psi^*(x,t)\Psi(x,t)\mathrm{d}x. \tag{1.1}$$

利用它可以计算 x 的平均值

$$\bar{x} = \int x \mathrm{d}P(x,t) = \int \Psi^*(x,t) x \Psi(x,t) \mathrm{d}x. \tag{1.2}$$

我们来讨论(1.2)式的含义.首先要明确,所提出的问题是:给定描述状态的波函数 $\Psi(x,t)$,问:测量坐标 x 会得到什么结果?我们知道,由于微观粒子的波粒二象性,对上述问题通常不能给出确定的答案,而只能给出概率性的答案.(1.2)式所给出的就是一个这种性质的答案.它告诉我们:在状态 $\Psi(x,t)$ 中测量坐标 x 所得结果的期望值是用 x 乘上 Ψ 再乘以 Ψ^* 并对全空间积分.

(二)在给定状态里动量的平均值

再来考虑动量 p_x 的平均值.这一问题的回答对于一种特殊情况非常简单,那就是对于德布罗意波[§5.1.2(2.21)式]

$$\Psi_{p_x}(x,t) = (1/\sqrt{L})\mathrm{e}^{\mathrm{i}(p_x x - Et)/\hbar} \tag{1.3}$$

所描述的状态.在这一状态里,由于动量 $p_x = \hbar k_x$ 有唯一确定的值,因而 $\bar{p}_x = p_x$.

利用(1.3)式所满足的归一化条件

$$\int_0^L \Psi^*_{p_x}(x,t)\Psi_{p_x}(x,t)\mathrm{d}x = 1, \tag{1.4}$$

可以将 $p_x = p_x$ 写为

$$p_x = p_x\int_0^L \Psi^*_{p_x}(x,t)\Psi_{p_x}(x,t)\mathrm{d}x = \frac{1}{L}\int_0^L \mathrm{e}^{-\mathrm{i}(p_x x - Et)/\hbar}p_x\mathrm{e}^{\mathrm{i}(p_x x - Et)/\hbar}\mathrm{d}x.$$

利用显然的等式

$$-\mathrm{i}\hbar\frac{\partial}{\partial x}\mathrm{e}^{\mathrm{i}p_x x/\hbar} = p_x\mathrm{e}^{\mathrm{i}p_x x/\hbar}, \tag{1.5}$$

可将上式进一步改写为

$$\begin{aligned}\bar{p}_x &= \frac{1}{L}\int_0^L \mathrm{e}^{-\mathrm{i}(p_x x - Et)/\hbar}\left(-\mathrm{i}\hbar\frac{\partial}{\partial x}\right)\mathrm{e}^{\mathrm{i}(p_x x - Et)/\hbar}\mathrm{d}x \\ &= \int_0^L \Psi^*_{p_x}(x,t)\left(-\mathrm{i}\hbar\frac{\partial}{\partial x}\right)\Psi_{p_x}(x,t)\mathrm{d}x.\end{aligned} \tag{1.6}$$

该式和求坐标 x 的平均值的(1.2)式有类似的结构,两者分别是用变量 x 乘在波函数 Ψ_{p_x} 上,和用微分算符 $-\mathrm{i}\hbar\partial/\partial x$ 作用在波函数 Ψ_{p_x} 上,再乘 $\Psi^*_{p_x}$ 并积分.

算符 $-\mathrm{i}\hbar\partial/\partial x$ 在计算动量平均值时起重要作用,我们称它为动量算符,用符号 $\hat{p}_x$ 表示:

$$\hat{p}_x = -\mathrm{i}\hbar\partial/\partial x. \tag{1.7}$$

将(1.6)式推广到计算任意状态 Ψ 中的动量平均值,就得到如下法则:当粒子处子状态 $\Psi(x,t)$ 时,粒子动量的平均值 $\bar{p}$ 可以通过对 $\Psi(x,t)$ 作如下运算而得到:将算符 $\hat{p} = -\mathrm{i}\hbar\partial/\partial x$ 作用于 $\Psi(x,t)$ 然后与 $\Psi^*(x,t)$ 相乘,再对粒子运动的整个空间积分:

$$\bar{p}_x = \int\Psi^*(x,t)\hat{p}_x\Psi(x,t)\mathrm{d}x. \tag{1.8}$$

(三)用算符表示物理量

算符 $\hat{p}_x = -\mathrm{i}\hbar\partial/\partial x$ 是在量子力学中表示动量的数学量,通过它对态函数 $\Psi(x,t)$ 的运算可以得到动量在 $\Psi(x,t)$ 所描述的状态中的平均值,如(1.8)式.

孤立的算符本身只是一种运算符号,没有直接的物理意义,它的物理意义体现在它对波函数的运算上.正是在这个意义上,人们说动量 p_x 用算符 $\hat{p}_x = -\mathrm{i}\hbar\partial/\partial x$ 表示.对于(1.2)式也可以作同样的理解,只是这时坐标算符 $\hat{x}$ 就是坐标 x 本身:

$$\hat{x} = x. \tag{1.9}$$

上面的结果推广到三维情况,就是

$$\bar{\boldsymbol{r}} = \int \Psi^*(\boldsymbol{r},t)\hat{\boldsymbol{r}}\Psi(\boldsymbol{r},t)\mathrm{d}\tau, \tag{1.10}$$

$$\bar{\boldsymbol{p}} = \int \Psi^*(\boldsymbol{r},t)\hat{\boldsymbol{p}}\Psi(\boldsymbol{r},t)\mathrm{d}\tau, \tag{1.11}$$

其中 $\mathrm{d}\tau = \mathrm{d}x\mathrm{d}y\mathrm{d}z \equiv \mathrm{d}^3\boldsymbol{r}$ 是空间体积元,而

$$\hat{\boldsymbol{p}} = -\mathrm{i}\hbar\nabla, \qquad \hat{\boldsymbol{r}} = \boldsymbol{r} \tag{1.12}$$

是三维动量和坐标算符.

推广到坐标和动量的任意函数 $F(\boldsymbol{r},\boldsymbol{p})$ 的平均值,有

$$\overline{F}(\boldsymbol{r},\boldsymbol{p}) = \int \Psi^*(\boldsymbol{r},t)\hat{F}\Psi(\boldsymbol{r},t)\mathrm{d}\tau, \tag{1.13}$$

其中

$$\hat{F} = F(\hat{\boldsymbol{r}},\hat{\boldsymbol{p}}). \tag{1.14}$$

它是把 $F(\boldsymbol{r},\boldsymbol{p})$ 中的坐标和动量换成算符而得到的 F 的算符.例如,粒子的角动量的经典力学表示为 $\boldsymbol{L} = \boldsymbol{r}\times\boldsymbol{p}$,在量子力学中它的算符则为

$$\hat{\boldsymbol{L}} = \hat{\boldsymbol{r}}\times\hat{\boldsymbol{p}} = -\mathrm{i}\hbar\boldsymbol{r}\times\nabla, \tag{1.15}$$

写成分量式为:

$$\left.\begin{aligned}
\hat{L}_x &= y\hat{p}_z - z\hat{p}_y = -\mathrm{i}\hbar(y\partial/\partial z - z\partial/\partial y),\\
\hat{L}_y &= z\hat{p}_x - x\hat{p}_z = -\mathrm{i}\hbar(z\partial/\partial x - x\partial/\partial z),\\
\hat{L}_z &= x\hat{p}_y - y\hat{p}_x = -\mathrm{i}\hbar(x\partial/\partial y - y\partial/\partial x).
\end{aligned}\right\} \tag{1.16}$$

又如,经典力学中的能量表示式为 $E = \boldsymbol{p}^2/(2m) + U(\boldsymbol{r})$,在量子力学中能量算符则为

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + U(\boldsymbol{r}). \tag{1.17}$$

这里假定了粒子是在保守力场 $U(\boldsymbol{r})$ 中运动,哈密顿函数 H 等于能量 E,见 §1.6.1,因面能量算符就是哈密顿算符 $\hat{H}$.

总结上而的讨论可以得到结论:

①在状态 Ψ 中力学量 F 的期望值可以通过算符 $\hat{F}$ 按(1.13)式对 Ψ 的运算而

得到.

②对于和经典力学中的力学量对应的量子力学中的力学量,其算符可以这样得到,即:将经典力学表示式中的 $\boldsymbol{r}$ 和 $\boldsymbol{p}$ 换成相应的算符 $\hat{\boldsymbol{r}}$ 和 $\hat{\boldsymbol{p}}$.

由于在一定极限下量子力学可以过渡到经典力学,因而一切经典力学中的力学量在量子力学中都对应存在.当然量子力学作为一个更广泛的理论,也包含有不存在经典对应的力学量,如以后要讨论的宇称和自旋.这种力学量的算符要根据它们自己的性质决定,以后将分别介绍.

(四)线性厄米算符

我们来讨论一下,用来表示力学量的算符的一般性质.

在数学中,一个算符 $\hat{L}$ 是表示一种运算符号,它的意义表现在:它对一个函数的运算结果会得到另一个函数,即

$$\hat{L}u = v. \tag{1.18}$$

这里 $\hat{L}$ 可以理解为微商,例如 $\hat{L} = \partial/\partial x$,或开方 $\hat{L} = \sqrt{\ }$,或一个数和 u 相乘,在后一情况下 $\hat{L}$ 就是这个数本身.不同算符体现在它们对 u 运算的结果所得到的 v 不同.

并不是任意的数学算符都可以用来在量子力学中表示力学量.能用来表示力学量的算符受到物理条件的限制.这首先是:算符的作用不应破坏波函数的叠加原理;同时,用算符计算出来的力学量平均值必须是实数.注意,由于波函数是复函数,上述第二个条件对力学量的算符加了很强的限制.

为了满足以上两个物理条件,在量子力学中表示力学量的算符必须是线性厄米自共轭算符,简称为线性厄米算符,或进一步简称为厄米算符.

算符 $\hat{L}$ 是线性的,是指它满足:

$$\hat{L}(c_1u_1 + c_2u_2) = c_1\hat{L}u_1 + c_2\hat{L}u_2, \tag{1.19}$$

其中 u_1 和 u_2 是两个任意函数,c_1 和 c_2 是两个任意(复)常数.显然 $\hat{L} = \partial/\partial x$ 是线性算符,而 $\hat{L} = \sqrt{\ }$ 不是.

表示力学量的算符之所以要是线性的,是由于叠加原理的限制,只有线性算符对波函数的运算,才不破坏叠加原理 §5.1.2(2.22)式.

对于算符 $\hat{L}$,可以按条件

$$\int u\tilde{\hat{L}}v\mathrm{d}\tau = \int v\hat{L}u\mathrm{d}\tau \tag{1.20}$$

定义它的转置算符 $\tilde{\hat{L}}$.

【例 1】 证明梯度算符的转置等于它自己乘以 -1:

$$\tilde{\nabla} = -\nabla. \tag{1.21}$$

证　先考虑一维情况. 根据定义

$$\int_a^b u\frac{\bar{\mathrm{d}}}{\mathrm{d}x}v\mathrm{d}x = \int_a^b v\frac{\mathrm{d}}{\mathrm{d}x}u\mathrm{d}x = vu\Big|_a^b - \int_a^b u\frac{\mathrm{d}}{\mathrm{d}x}v\mathrm{d}x.$$

如果 $a=-\infty$, $b=\infty$, 则由归一化条件 §5.1.2(2.12)式, 当 $x\to\pm\infty$ 时 $u,v\to 0$, 因而积分出的项 $vu\Big|_a^b$ 等于零. 如果 a,b 取有限值, 令 u,v 满足周期性条件, 积分出的项也等于0. 从而得到

$$\int_a^b u\frac{\bar{\mathrm{d}}}{\mathrm{d}x}v\mathrm{d}x = -\int_a^b u\frac{\mathrm{d}}{\mathrm{d}x}v\mathrm{d}x.$$

在三维情况下,

$$\int_V u\nabla v\mathrm{d}\tau = \int_V \nabla(uv)\mathrm{d}\tau - \int_V v\nabla u\mathrm{d}\tau.$$

利用附录(Ⅱ.26)式, 可将左边第一个积分改写成

$$\int_V \nabla(uv)\mathrm{d}\tau = \oint(uv)\mathrm{d}\boldsymbol{\sigma}.$$

如果是无穷远封闭面, $u,v\to 0$; 如果运动局限在有限区域, 在区域的边界上 u,v 或者等于零, 或者指数地趋于零[参看 §5.3.1 中的讨论(第333页)], 因而

$$\int_V u\nabla v\mathrm{d}\tau = -\int_V v\nabla u\mathrm{d}\tau.$$ 【证毕】

若把算符 $\hat{L}$ 中的虚因子 i 换成 $-\mathrm{i}$, 则所得的算符为 $\hat{L}$ 的复共轭算符 $\hat{L}^*$. 算符 $\hat{L}$ 的转置复共轭算符 $\tilde{\hat{L}}^*$ 称为它的厄米共轭算符

$$\hat{L}^\dagger = \tilde{\hat{L}}^*. \tag{1.22}$$

容易证明, $\hat{L}^\dagger$ 是满足下列条件的算符:

$$\int u^*\hat{L}^\dagger v\mathrm{d}\tau = \int v(\hat{L}u)^*\mathrm{d}\tau. \tag{1.23}$$

证　改写(1.20)式中的 v 为 v^*: $\int u\tilde{\hat{L}}v^*\mathrm{d}\tau = \int v^*\hat{L}u\mathrm{d}\tau$. 等式左右两边取复共轭, 利用(1.22)式就得到(1.23)式. 【证毕】

如果算符 $\hat{L}$ 等于它自己的厄米共轭算符 $\hat{L}^\dagger$:

$$\hat{L} = \hat{L}^\dagger, \tag{1.24}$$

则称 $\hat{L}$ 为厄米自共轭算符, 简称厄米算符. 按照(1.23)式, 厄米(自共轭)条件(1.24)式可写成:

$$\int u^*\hat{L}v\mathrm{d}\tau = \int v(\hat{L}u)^*\mathrm{d}\tau. \tag{1.25}$$

力学量的算符之所以必须是厄米的, 是由于力学量的平均值应为实数, 即应有 $\bar{F}=\bar{F}^*$:

$$\int \Psi^* \hat{F}\Psi \mathrm{d}\tau = \left(\int \Psi^* \hat{F}\Psi \mathrm{d}\tau\right)^* = \int \Psi (\hat{F}\Psi)^* \mathrm{d}\tau = \int \Psi^* \hat{F}^\dagger \Psi \mathrm{d}\tau. \quad (1.26)$$

最后一步用了(1.23)式. 如果 $\hat{F}$ 满足厄米条件(1.24)式,则(1.26)式成立.

总结前面的讨论,可以提出量子力学关于力学量的基本假定:力学量用线性厄米算符表示. 以下将线性厄米算符简称为厄米算符.

下面来证明坐标和动量算符是厄米算符.

证　为了简单起见,只考虑一维情况,三维情况下的证明是类似的.

对于坐标算符 $\hat{x} = x$,显然有 $\int u^* xv \mathrm{d}x = \int v(xu)^* \mathrm{d}x$.

将(1.21)式两边乘 $-\mathrm{i}\hbar$,取复共轭,利用(1.22)式,就可以看到动量算符 $\hat{p} = -\mathrm{i}\hbar\partial/\partial x$ 是厄米算符. **【证毕】**

§5.2.2　测量结果的概率分布　算符的本征值与本征函数

上一节的开始已经指出,在给定的状态里测量微观粒子的力学量,通常不是得到唯一确定的值,而是按一定概率分布的一系列值. 通过上一节的讨论看到,量子力学的一个基本假定是,力学量用线性厄米算符表示. 利用力学量的算符可以预言在给定状态里测量这一力学量所得结果的期望值——平均值. 这一节进一步讨论如何利用力学量的算符决定这一力学量所可能取的值,以及如何预言在一个给定状态里测量这一力学量得到各种不同值的概率.

力学量可能取的值由系统本身的性质和所处的外部条件决定,与运动状态无关. 例如,限制在体积为 L^3 的空间范围里运动的系统,其动量所能取的值是由 §5.1.2(2.20)式决定的那些分立值. 不管系统处于什么状态,测量动量都只能得到这样一些值. 但是,在不同状态里进行测量,得到这些值的概率分布不同. 由此可见,力学量的可能值的谱仅仅由这一力学量的算符和系统所处的条件(边界条件)决定;而在给定状态里,力学量取不同值的概率分布则既依赖于力学量的算符,也依赖于状态的波函数,由它们两者共同决定.

(一)力学量取值的谱

为了回答力学量所能取的值的谱,我们先来寻找一种特殊的状态,在这种状态中,力学量 F 有确定的数值 f●. 设 Ψ 是该状态的波函数,显然在这个状态中,力学量 F 的平均值 $\bar{f}$ 就等于这个确定值 f,所以在这个态中,个别测量结果对这个值的

● 为了避免混淆,在这一节里,我们用大写的 F 表示一个力学量,小写的 f 表示它所取的值,而 $\hat{f}$ 表示力学量 F 的算符.

偏差的平方平均值，即均方差 $\overline{(\Delta f)^2} = \overline{(\hat{f}-f)^2}$ 应为零. 定义 Δf 的算符为 $\Delta\hat{f} = \hat{f}-f$. 按 §5.2.1(1.13)式给出的平均值公式，均方差为零可表示成：

$$\overline{(\Delta f)^2} = \int \Psi^* (\Delta\hat{f})^2 \Psi \mathrm{d}\tau = 0 \tag{2.1}$$

由 $\hat{f}$ 的厄米性知 $\Delta\hat{f}$ 也是厄米的，于是上式可写为

$$\overline{(\Delta f)^2} = \int \Psi^* \Delta\hat{f}(\Delta\hat{f}\Psi)\mathrm{d}\tau = \int (\Delta\hat{f}\Psi)(\Delta\hat{f}\Psi)^* \mathrm{d}\tau = \int |\Delta\hat{f}\Psi|^2 \mathrm{d}\tau = 0.$$

由于 $|\Delta\hat{f}\Psi|^2$ 不能为负数，故要上式为零，只有 $\Delta\hat{f}\Psi = (\hat{f}-f)\Psi = 0$，即

$$\hat{f}\Psi = f\Psi. \tag{2.2}$$

满足这个条件的波函数 Ψ 描述力学量 F 有确定值 f 的状态.

加上一定的定解条件的方程 $\hat{f}\Psi = f\Psi$，称为算符 $\hat{f}$ 的本征值方程，f 称为 $\hat{f}$ 的本征值，相应的函数 Ψ 称为和本征值 f 对应的本征函数. $\hat{f}$ 的本征函数描述的状态称为 $\hat{f}$ 的本征态. 由上面的讨论可知，在算符 $\hat{f}$ 的本征函数所描述的状态中测量力学量 F，会得到确定的值，这个值就是和这个本征函数对应的本征值 f. 这样就把力学量算符的本征值和力学量所能取的值联系起来了. 于是有量子力学中关于力学量所能取的值的基本假定：在一个量子力学系统中，某一力学量所能取的值是这一力学量的算符在由系统性质决定的定解条件下的本征值.

如果系统处于物理量的算符 $\hat{f}$ 的本征态 Ψ 中，则这一物理量的平均值 $\bar{f}$ 就等于和这个本征态对应的本征值：

$$\bar{f} = \int \Psi^* \hat{f}\Psi \mathrm{d}\tau = f\int \Psi^*\Psi \mathrm{d}\tau = f.$$

这里用了本征值方程(2.2)式.

已知厄米算符的平均值为实数，而厄米算符的本征值等于它在本征态中的平均值，所以厄米算符的本征值是实数.

在一定的定解条件下解本征值方程，会得到一系列本征值，相应地可以求得一系列的本征函数. 这一系列本征值的集合称为本征值谱. 由于定解条件不同，本征值可能是分立的，也可能是连续的.

当本征值谱取分立值时，可以将本征值写成 $f_1, f_2, \cdots, f_i, \cdots$；相应的本征函数是 $\Psi_1, \Psi_2, \cdots, \Psi_i, \cdots$；而本征值方程写成：

$$\hat{f}\Psi_i = f_i\Psi_i, \qquad i = 1,2,\cdots, \tag{2.3}$$

取整数值的数 i 用来标志本征值和本征函数，称为量子数.

如果本征值连续地取值，则称这一力学量有连续谱. 此时本征函数是不可数的，通常就用相应的本征值来标记本征函数. 和本征值 f 对应的本征函数写成 Ψ_f，而本征值方程写成

$$\hat{f}\Psi_f = f\Psi_f \ . \tag{2.4}$$

注意，Ψ_f 的下标 f 是本征函数的标记，而写在 Ψ_f 前面的 f 是和这一本征函数对应的本征值.

在具体问题中，本征值方程的定解条件可以应用波函数所要满足的一般性条件及系统所处的外部条件得到. 波函数的一般性条件是：波函数在它的变量变化的全部区域内是单值、连续、有限的. 这些条件有时又称为波函数的标准条件. 单值性和有限性是波函数的物理意义所要求的，而在下一章中将看到，波函数所满足的运动方程包含有二阶导数，因而要求它是连续的，并至少有二阶导数存在.

在很多情况下，算符 $\hat{f}$ 都不含时间 t，例如 $\boldsymbol{r}$，$\boldsymbol{p}$ 以及它们的函数 $E(\boldsymbol{r},\boldsymbol{p})$ 都是这样. 在此情况下，由本征值方程(2.2)解出的本征函数 Ψ 也不含 t. (这里说的是薛定谔绘景，在其他绘景中不是这样，见 §5.4.7)

(二)本征值方程

下面讨论几个求解本征值方程的例子，以具体了解定解条件的作用和本征值谱的特征.

【例 1】 求动量算符的本征值和本征函数.

解 先看一维情况. $\hat{p}_x$ 的本征值方程是

$$-\mathrm{i}\hbar \mathrm{d}\psi_{p_x}/\mathrm{d}x = p_x\psi_{p_x}. \tag{2.5}$$

直接积分得：

$$\psi_{p_x} = A\mathrm{e}^{(\mathrm{i}/\hbar)p_x x}. \tag{2.6}$$

如果粒子在无限空间范围 $-\infty < x < \infty$ 中运动，则只要 p_x 是实数，它取任何值都可保证波函数满足标准条件，故知 p_x 有连续谱，其取值范围是 $-\infty < p_x < \infty$. 而如果粒子在有限空间范围 $0 \leqslant x \leqslant L$ 中运动，则 p_x 有分立谱，如 § 5.1.2(2.20)式

在三维情况下，动量算符的本征值方程为

$$-\mathrm{i}\hbar\left(\boldsymbol{e}_x\frac{\partial}{\partial x} + \boldsymbol{e}_y\frac{\partial}{\partial y} + \boldsymbol{e}_z\frac{\partial}{\partial z}\right)\psi_{\boldsymbol{p}} = (\boldsymbol{e}_x p_x + \boldsymbol{e}_y p_y + \boldsymbol{e}_z p_z)\psi_{\boldsymbol{p}} \tag{2.7}$$

由于各个方向的平移运动相互独立(见 §1.5.1 图 1)，可以写

$$\psi_{\boldsymbol{p}}(\boldsymbol{r}) = \psi_{p_x}(x)\psi_{p_y}(y)\psi_{p_z}(z). \tag{2.8}$$

代回上式，得到三个已求解过的一维方程(2.5)，因而

$$\psi_{\boldsymbol{p}}(\boldsymbol{r}) = A\mathrm{e}^{(\mathrm{i}/\hbar)(p_x x + p_y y + p_z z)} = A\mathrm{e}^{(\mathrm{i}/\hbar)\boldsymbol{p}\cdot\boldsymbol{r}} \tag{2.9}$$

不难看到，§5.1.2(2.17)式，在去掉时间因子 $\exp(-\mathrm{i}Et/\hbar)$ 以后，就是这里的(2.9)式. 和一维情况一样，如果粒子在无限空间中运动，则 $\boldsymbol{p}$ 有连续谱：

$$-\infty < p_x, p_y, p_z < \infty\ , \tag{2.10}$$

而如果粒子被限制在体积为 L^3 的方箱中，则 $\boldsymbol{p}$ 有分立谱，如 §5.1.2(2.20)式.

【例 2】 设质量为 m 的粒子做一维自由运动，求其能量算符的本征值和本征

函数.

解　这时粒子的能量算符是

$$\hat{H} = \hat{p}_x^2/(2m) = -[\hbar^2/(2m)]\mathrm{d}^2/\mathrm{d}x^2, \tag{2.11}$$

其本征值方程为

$$-[\hbar^2/(2m)]\mathrm{d}^2\psi(x)/\mathrm{d}x^2 = E\psi(x).$$

它可以写成

$$\mathrm{d}^2\psi(x)/\mathrm{d}x^2 + k^2\psi(x) = 0, \qquad k^2 = 2mE/\hbar^2. \tag{2.12}$$

这是二阶常系数线性常微分方程,它可以有两个基本解 $\mathrm{e}^{\pm ikx}$,通解是这两个基本解的叠加:

$$\psi_k = A\mathrm{e}^{ikx} + B\mathrm{e}^{-ikx}. \tag{2.13}$$

可以看出,k 取任何实数值,解都满足标准条件,故 $E = \hbar^2k^2/(2m)$ 有连续谱:

$$0 \leqslant E < \infty. \tag{2.14}$$

如果在(2.13)式中取 $B = 0$ 或 $A = 0$,就得到

$$\psi_k^{(+)} = A\mathrm{e}^{ikx}, \qquad \psi_k^{(-)} = B\mathrm{e}^{-ikx}. \tag{2.15}$$

和(2.6)式比较可见,它们都是动量 $\hat{p}_x$ 的本征函数,本征值分别为

$$p_x = \hbar k = \sqrt{2mE} \quad 和 \quad p_x = -\hbar k = -\sqrt{2mE}. \tag{2.16}$$

这表明:①对于自由粒子,能量算符和动量算符可以有共同的本征函数;②和同一个能量本征值 $E = \hbar^2k^2/(2m)$(除 $E = 0$ 以外)对应,有两个独立的本征函数,如(2.15)式. 这种和两个独立本征函数对应的本征值被称为是二重简并的.

一般说来,如果相应于力学量的一个本征值,有 r 个线性独立的本征函数,则称力学量的这个本征值是简并的,简并度为 r. 一维自由粒子的能量本征值是二重简并的,原因是,自由粒子的能量只依赖于动量的大小,而其本征函数却还依赖于动量的方向. 能量在动量的两个方向上是对称的,导致二重简并. 以后还会看到,简并常常和系统的某种对称性相联系.

【例3】　求角动量 $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$ 的本征值和本征函数.

解　为了计算方便,首先把角动量算符变换到球坐标系,在这种坐标系中有[见 §1.1.1(1.9)式]:

$$x = r\sin\theta\cos\varphi, \quad y = r\sin\theta\sin\varphi, \quad z = r\cos\theta. \tag{2.17}$$

以及

$$r^2 = x^2 + y^2 + z^2, \quad \cos\theta = \frac{z}{\sqrt{x^2 + y^2 + z^2}}, \quad \tan\varphi = \frac{y}{x}. \tag{2.18}$$

由它们可得

$$\left.\begin{aligned}
&\frac{\partial r}{\partial x}=\sin\theta\cos\varphi, \quad \frac{\partial r}{\partial y}=\sin\theta\sin\varphi, \quad \frac{\partial r}{\partial z}=\cos\theta,\\
&\frac{\partial\theta}{\partial x}=\frac{\cos\theta\cos\varphi}{r}, \quad \frac{\partial\theta}{\partial y}=\frac{\cos\theta\sin\varphi}{r}, \quad \frac{\partial\theta}{\partial z}=-\frac{\sin\theta}{r},\\
&\frac{\partial\varphi}{\partial x}=-\frac{\sin\varphi}{r\sin\theta}, \quad \frac{\partial\varphi}{\partial y}=\frac{\cos\varphi}{r\sin\theta}, \quad \frac{\partial\varphi}{\partial z}=0.
\end{aligned}\right\} \tag{2.19}$$

这样可以求得球坐标系中的角动量算符：

$$\begin{aligned}
\hat{L}_x &= -\mathrm{i}\hbar\left(x\frac{\partial}{\partial y}-y\frac{\partial}{\partial x}\right)\\
&= -\mathrm{i}\hbar\left[r\sin\theta\cos\varphi\left(\frac{\partial r}{\partial y}\frac{\partial}{\partial r}+\frac{\partial\theta}{\partial y}\frac{\partial}{\partial\theta}+\frac{\partial\varphi}{\partial y}\frac{\partial}{\partial\varphi}\right)-\right.\\
&\quad \left.r\sin\theta\sin\varphi\left(\frac{\partial r}{\partial x}\frac{\partial}{\partial r}+\frac{\partial\theta}{\partial x}\frac{\partial}{\partial\theta}+\frac{\partial\varphi}{\partial x}\frac{\partial}{\partial\varphi}\right)\right].
\end{aligned}$$

经过化简得到●

$$L_z = -\mathrm{i}\hbar\partial/\partial\varphi. \tag{2.20}$$

类似可得

$$\hat{L}_x = \mathrm{i}\hbar\left(\sin\varphi\frac{\partial}{\partial\theta}+\cot\theta\cos\varphi\frac{\partial}{\partial\varphi}\right), \tag{2.21}$$

$$\hat{L}_y = -\mathrm{i}\hbar\left(\cos\varphi\frac{\partial}{\partial\theta}-\cot\theta\sin\varphi\frac{\partial}{\partial\varphi}\right), \tag{2.22}$$

$$\hat{L}^2=\hat{L}_x^2+\hat{L}_y^2+\hat{L}_z^2=-\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right]. \tag{2.23}$$

现在来求解 $\hat{L}_z$ 的本征值方程

$$-\mathrm{i}\hbar\partial\psi(\varphi)/\partial\varphi = L_z\psi(\varphi). \tag{2.24}$$

直接积分可得

$$\psi(\varphi) = A\mathrm{e}^{(\mathrm{i}/\hbar)L_z\varphi}. \tag{2.25}$$

按照波函数的标准条件，$\psi(\varphi)$ 应是单值函数，因而有

$$\psi(\varphi) = \psi(\varphi+2\pi). \tag{2.26}$$

显然只有当 $L_z/\hbar = m$，且 $m = 0, \pm1, \pm2, \cdots$ 时，函数(2.25)才满足上述条件，因而得到 $\hat{L}_z$ 的本征值谱为

$$L_z = m\hbar, \qquad m = 0, \pm1, \pm2, \cdots. \tag{2.27}$$

m 称为磁量子数，它决定了角动量在 z 轴方向的投影. 相应的本征函数为

● 仿照§5.2.1(1.7)式定义一个和广义坐标 φ 对应的广义动量算符，即 $\hat{p}_\varphi = -\mathrm{i}\hbar\partial/\partial\varphi$，则(2.20)式成为 $\hat{L}_z = \hat{p}_\varphi$. 这就是§1.2.2(2.16)式的算符形式.

$$\psi_m(\varphi) = (1/\sqrt{2\pi})e^{im\varphi}, \tag{2.28}$$

这里已利用附录(Ⅰ.22)式由归一化条件 $\int_0^{2\pi}\psi_m^*\psi_m d\varphi = 1$ 求出了(2.25)式中的归一化常数.

总角动量的可能值通过解角动量平方算符的本征值方程 $\hat{L}^2\psi = L^2\psi$ 而得到. 注意 $\hat{L}^2$ 算符(2.23)是对于角变量 θ 和 φ 的微分算符,因而其本征函数是 θ 和 φ 的函数,可写为 $\psi(\theta,\varphi)$,即有:

$$\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}+\frac{L^2}{\hbar^2}\right]\psi(\theta,\varphi) = 0. \tag{2.29}$$

将(2.29)式和附录(Ⅳ.7)式比较可见,它们是同一个方程. 因此,(2.29)式的解(即 $\hat{L}^2$ 的本征函数)就是球谐函数. 在附录Ⅳ中指出,为使解处处单值、连续、有限,必须

$$L^2 = l(l+1)\hbar^2, \qquad l = 0,1,2,\cdots. \tag{2.30}$$

这就是 $\hat{L}^2$ 的本征值,l 称为角量子数. 由附录(Ⅳ.31)式还知道,对于每一个 l,有由不同 m 值

$$m = 0, \pm 1, \pm 2, \cdots, \pm l \tag{2.31}$$

相区别的 $2l+1$ 个解 $\psi(\theta,\varphi) = Y_{lm}(\theta,\varphi)$:

$$Y_{lm}(\theta,\varphi) = \sqrt{\frac{(l-|m|)!(2l+1)}{(l+|m|)!4\pi}}P_l^{|m|}(\cos\theta)e^{im\varphi}. \tag{2.32}$$

将(2.32)式和(2.28)式比较可以看到,$\hat{L}^2$ 的本征函数 $Y_{lm}(\theta,\varphi)$ 同时也是 $\hat{L}_z$ 的本征函数. $\hat{L}_z$ 的量子数(磁量子数) m 和 $\hat{L}^2$ 的量子数(角量子数) l 之间的关系是:$|m| \leqslant l$,如(2.31)式. 它有明显的物理意义,即:角动量的一个分量,不能比角动量矢量的长度更大. 但要注意,角动量矢量的长度 $\sqrt{L^2}$ 精确说来不是 $l\hbar$ 而是 $\sqrt{l(l+1)}\hbar$.

给定一个角量子数的值 l,就决定了 $\hat{L}^2$ 的一个本征值,而相应地却有用磁量子数 m 相区别的 $2l+1$ 个本征函数 $Y_{lm}(\theta,\varphi)$,故 $\hat{L}^2$ 的本征值有 $2l+1$ 重简并.

值得注意的是,$Y_{lm}(\theta,\varphi)$ 是 $\hat{L}^2$ 和 $\hat{L}_z$ 的共同本征态,然而却不是 $\hat{L}_x$ 和 $\hat{L}_y$ 的本征态(除了 $l = m = 0$ 的情况以外). 利用(2.21)、(2.22)式很容易验证这一结论. 这一重要事实表明,角动量的不同分量不能有共同的本征函数系. 其本质原因在于,不同方向的转动之间相互关联,见§1.5.1 图2.

【例4】　粒子在一维无限深势阱(势箱)

$$U(x) = \begin{cases} 0, & \text{当 } 0 < x < a, \\ \infty, & \text{当 } x = 0 \text{ 和 } x = a \end{cases} \tag{2.33}$$

中运动,求能量的本征值和本征函数.

解　粒子的能量本征值方程为

$$-[\hbar^2/(2m)]\mathrm{d}^2\psi(x)/\mathrm{d}x^2 + U(x)\psi(x) = E\psi(x).$$

为了得到在 $x = 0$ 和 $x = a$ 处方程的解应满足的条件，把方程写成：

$$\psi''(x)/\psi(x) = -(2m/\hbar^2)(E - U(x)). \tag{2.34}$$

当 $x \to 0$ 和 $x \to a$ 时，$U(x) \to \infty$，这只有在这两点 $\psi = 0$ 才有可能，即

$$\psi(0) = \psi(a) = 0. \tag{2.35}$$

在区域 $0 < x < a$，方程(2.34)成为

$$\psi''(x) + (2mE/\hbar^2)\psi = 0. \tag{2.36}$$

这是一个二阶常系数线性微分方程，它的一般解可取为：

$$\psi(x) = A\sin\left(\sqrt{2mE/\hbar^2}x\right) + B\cos\left(\sqrt{2mE/\hbar^2}x\right).$$

由边界条件 $\psi(0) = 0$ 得 $B = 0$，再由 $\psi(a) = 0$ 得 $A\sin\left(\sqrt{2mE/\hbar^2}a\right) = 0$．但是 $A \neq 0$，否则将有 $\psi(x) = 0$，不符合要求，因此，

$$\sin\left(\sqrt{2mE/\hbar^2}a\right) = 0. \tag{2.37}$$

这要求

$$\sqrt{2mE/\hbar^2}a = n\pi, \quad n = 1,2,\cdots. \tag{2.38}$$

注意 $n \neq 0$，否则也将有 $\psi(x) = 0$．

由(2.38)式知能量取分立值

$$E_n = n^2\pi^2\hbar^2/(2ma^2) = n^2h^2/(8ma^2), \quad n = 1,2,\cdots. \tag{2.39}$$

相应的本征函数为

$$\psi_n = A\sin(n\pi x/a). \tag{2.40}$$

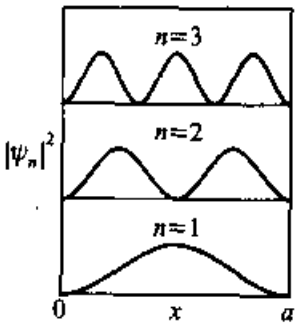


图 1　势箱中粒子的位置概率密度

利用归一化条件 $A^2\int_0^l \sin^2(n\pi x/a)\mathrm{d}x = 1$，得 $A = \sqrt{2/a}$．

归一化的本征函数是

$$\psi_n = \sqrt{2/a}\sin(n\pi x/a). \tag{2.41}$$

值得注意的是，由于 $n \neq 0$，所以由(2.39)式可知，被限制在无限深势阱中的粒子的最低能量(以势能极小值为计算起点)是

$$E_{\min} = \pi^2\hbar^2/(2ma^2) > 0. \tag{2.42}$$

在一个给定势能场中的最低能量状态称为给定势能场中的基态．基态能量称为零点能．由(2.42)式知，无限深势阱中粒子的基态能量(零点能)大于零，而不是等于零．这是一个有普遍意义的重要结果，后面将看到，它和微观粒子的波粒二象性有直接联系．如果是经典粒子，能量最低的状态是停在势箱中不动，其能量为零．

图 1 上画出了在 $n = 1,2,3$ 三个能量本征态中，粒子的位置概率密度 $|\psi_n|^2$．

由图可见，边界条件(2.35)式使概率分布成为驻波型的分布，$|\psi_n|^2=0$ 的点称为波节. 由图可见，基态($n=1$)在粒子运动区间内没有波节，第一激发态 $n=2$ 有一个波节，第二激发态 $n=3$ 有两个波节，…….

比较本节(2.40)式和§3.3.4(4.22)～(4.26)式，以及本节图1和§3.3.4图2，很容易看到势箱中的微观粒子和谐振腔中的电磁波之间的类似性. 其原因是，两种情况下，稳定的运动状态都是驻波，要求空间尺度等于半波长的整数倍. 正是这一要求导致谐振腔中存在各种电磁波模式和势箱中出现微观粒子能量的量子化. 由此可以清楚地看到，能量之所以量子化是由于微观粒子具有波动性.

(三)本征函数的正交性

厄米算符的本征函数的一个重要性质是正交性. “正交性”这个词是从矢量空间中借用来的. 在证明本征函数的正交性之前，先说明如何将函数看成矢量以及如何定义函数的标积.

如果由 n 个复数($a_1,a_2,\cdots,a_n$)形成 n 维复空间的矢量 $\boldsymbol{A}$，则两个矢量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的标积定义为

$$(\boldsymbol{A},\boldsymbol{B})=\sum_{i=1}^{n}a_i^*b_i. \tag{2.43}$$

这是三维空间矢量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的标积 $\boldsymbol{A}\cdot\boldsymbol{B}=\sum_{i=1}^{3}A_iB_i$ 的自然推广. 如果 $(\boldsymbol{A},\boldsymbol{B})=0$，则称矢量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 正交.

(2.43)式中的下标 i 是标志矢量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的不同分量的标号，它取分立值. 如果作一个形式的推广，让标志矢量分量的标号连续取值，也就是说让 $i\to x$，则 $a_i\to f(x)$，$b_i\to g(x)$. $f(x)$ 和 $g(x)$ 称为函数空间的矢量，用 F 和 G 表示. 它们的标积由(2.43)式中将对 i 的求和改成对 x 的积分而得到：

$$(F,G)=\int f^*(x)g(x)\mathrm{d}x. \tag{2.44}$$

如果 $(F,G)=0$，就说 $f(x)$ 和 $g(x)$ 正交. 这一定义可以不限于一维情况，因而有定义

$$\text{如果}\quad\int f^*g\mathrm{d}\tau=0,\quad\text{则函数 } f \text{ 和 } g \text{ 正交.} \tag{2.45}$$

现在来证明一个定理：厄米算符对应于不同本征值的本征函数相互正交. 为了确定起见，在证明中假定本征值有分立谱，对于连续谱可以同样地证明.

证　设 ψ_k 和 ψ_l 是力学量算符 $\hat{F}$ 的任意两个本征函数，相应的本征值分别是 f_k 和 f_l，且 $f_k\neq f_l$. 由 $\hat{F}$ 的厄米条件，有

$$\int\psi_k^*\hat{F}\psi_l\mathrm{d}\tau=\int\psi_l(\hat{F}\psi_k)^*\mathrm{d}\tau.$$

利用厄米算符的本征值为实数的性质，注意到 $\hat{F}\psi_l = f_l\psi_l$，$\hat{F}\psi_k = f_k\psi_k$，则上式成为

$$(f_l - f_k)\int \psi_k^* \psi_l \mathrm{d}\tau = 0.$$

而按假定 $f_l - f_k \neq 0$，故为使上式成立，要求 $\int \psi_k^* \psi_l \mathrm{d}\tau = 0$. 【证毕】

如果本征值 f_k 有 s 重简并，即对应于 f_k 有用 α 相区别的 s 个线性独立的本征函数 $\psi_{k\alpha}$

$$\hat{F}\psi_{k\alpha} = f_k\psi_{k\alpha}, \quad \alpha = 1,2,\cdots,s, \tag{2.46}$$

则 $\psi_{k\alpha}$ 和 $\psi_{k\beta}(\alpha \neq \beta)$ 不一定相互正交. 此时，可以由 s 个 $\psi_{k\alpha}$ 组合得到另外 s 个线性独立的函数：

$$\psi_{k\rho} = \sum_{\alpha=1}^{s} a_{\rho\alpha}\psi_{k\alpha}, \quad \rho = 1,2,\cdots,s. \tag{2.47}$$

用 $\hat{F}$ 作用到(2.47)式上，利用(2.46)式可见，这 s 个 $\psi_{k\rho}$ 也是本征值为 f_k 的本征函数 $\hat{F}\psi_{k\rho} = f_k\psi_{k\rho}$. 可以选系数 $a_{\rho\alpha}$ 使 s 个 $\psi_{k\rho}$ 相互正交，参看习题9.

将上述定理和§1.5.1中证明的：三个惯量主轴相互垂直，或者可以选成垂直(如果 $I_\alpha = I_\beta$)的定理相对比，可见两者非常类似，包括定理的证明方法和步骤[参看§1.4.3第(二)小点]也类似. 这一类似不是偶然的. 实际上，实三维空间中的对称矩阵[例如转动惯量矩阵§1.5.1(1.24)式]推广到复空间就是厄米矩阵，再推广到函数空间就是厄米算符. 对称矩阵的主值和主轴推广到函数空间就是厄米算符的本征值和本征函数[比较转动惯量的主值方程§1.5.1(1.30)式和厄米算符的本征值方程本节(2.2)式]. 利用这一对比，能加深对本征函数正交性定理的直观理解.

正像对于转动惯量可以选定三个相互正交的主轴一样，对于厄米算符 $\hat{F}$ 可以选定一组相互正交的本征函数系. 此时，属于同一简并本征值的线性独立本征函数已经按(2.47)式的办法正交化. 用一个统一的脚标 i 来标志这一正交函数系中的函数，则有

$$\int \psi_i^* \psi_j \mathrm{d}\tau = 0, \quad \text{当} \quad i \neq j. \tag{2.48}$$

(四)本征函数的归一化

我们在写出(2.48)式的时候，假定了 $\hat{F}$ 的本征值有分立谱. 在此情况下，ψ_i 可以按§5.1.2(2.12)式归一化

$$\int \psi_i^* \psi_i \mathrm{d}\tau = 1. \tag{2.49}$$

利用克罗内克 δ 符号[见附录(Ⅰ.1)式]可以将(2.48)、(2.49)式合并写为

$$\int \psi_i^* \psi_j \mathrm{d}\tau = \delta_{ij}. \tag{2.50}$$

这是分立谱本征函数的正交归一条件.

如果 F 的本征值有连续谱,则它的本征函数不能按 §5.1.2(2.12)式归一化. 但是,对应不同本征值的本征函数的正交性定理仍然成立,即

$$\int \psi_f^* \psi_{f'} \mathrm{d}\tau = 0, \qquad 当 \qquad f \neq f'. \tag{2.51}$$

这对应于分立谱情况下的(2.48)式. 和分立情况下的(2.50)式对应,将 δ 符号 δ_{ij} 换成 δ 函数 $\delta(f-f')$ (见附录Ⅰ),就有

$$\int \psi_f^* \psi_{f'} \mathrm{d}\tau = \delta(f-f'). \tag{2.52}$$

这是连续谱本征函数的正交归一条件. 根据 δ 函数的定义[见附录(Ⅰ.11)式],当 $f \neq f'$ 时,(2.52)式过渡到(2.51)式,而当 $f = f'$ 时,(2.52)式中的积分成为 ∞ ,这就是说 ψ_f 不能按 §5.1.2(2.12)式归一化. 我们用(2.52)式来规定连续谱本征函数 ψ_f 的"归一化",见下例.

【例 5】　将无限空间中动量的本征函数归一化.

解　先看一维情况,由本节(2.6)式

$$\int_{-\infty}^{\infty} \psi_{p'_x}^*(x) \psi_{p_x}(x) \mathrm{d}x = A^2 \int_{-\infty}^{\infty} \mathrm{e}^{(i/\hbar)(p_x - p'_x)\cdot x} \mathrm{d}x.$$

根据附录(Ⅰ.14)式,有

$$\delta(p_x - p'_x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{(i/\hbar)(p_x - p'_x)x} \mathrm{d}(x/\hbar).$$

将以上二式代入本节(2.52)式,得到 $A = 1/\sqrt{2\pi\hbar}$,因而归一化的动量本征函数是

$$\psi_{p_x}(x) = (1/\sqrt{2\pi\hbar})\mathrm{e}^{(i/\hbar)p_x x}. \tag{2.53}$$

三维情况下,动量的归一化本征函数是三个(2.53)式相乘:

$$\psi_p(r) = (2\pi\hbar)^{-3/2} \mathrm{e}^{(i/\hbar)p\cdot r}. \tag{2.54}$$

它满足归一化条件

$$\int \psi_{p'}^*(r) \psi_p(r) \mathrm{d}^3 r = \delta^3(p - p'). \tag{2.55}$$

我们来证明,按连续谱归一化条件(2.55)式归一化得到的(2.54)式和按箱归一化得到的 §5.1.2(2.21)式的一致性.

证　考虑一维情况.

①当 $L \to \infty$ 时,箱归一化的相邻动量值之差 $\Delta p = 2\pi\hbar/L \to 0$ [§5.1.2(2.20)式],动量由分立谱变成连续谱;

②此时，按§5.1.2(2.12)式归一化得到的§5.1.2(2.21)式 $\to 0$．为了克服这一困难，我们将归一化条件§5.1.2(2.12)式(略去时间因子)写为

$$\int \psi_p^{箱归一*}(x)\psi_{p'}^{箱归一}(x)\mathrm{d}x = \begin{cases} 0, & p \neq p', \\ 1, & p = p', \end{cases} \tag{2.56}$$

由此得归一化的波函数[§5.1.2(2.21)式]

$$\psi_p^{箱归一}(x) = L^{-1/2}\mathrm{e}^{(i/\hbar)px}.$$

③(2.56)式可写为

$$\sum_p \int \psi_p^{箱归一*}(x)\psi_{p'}^{箱归一}(x)\mathrm{d}x = 1.$$

或

$$\frac{1}{\Delta p}\sum_p \left[\int \psi_p^{箱归一*}(x)\,\psi_{p'}^{箱归一}(x)\mathrm{d}x\right]\Delta p = 1.$$

当 $L \to \infty$ 时，成为

$$\frac{1}{\Delta p}\int\left[\int \psi_p^{箱归一*}(x)\,\psi_{p'}^{箱归一}(x)\mathrm{d}x\right]\mathrm{d}p = 1.$$

和连续谱归一化条件(2.55)式

$$\int\left[\int \psi_p^{连续谱}(x)\psi_{p'}^{连续谱}(x)\mathrm{d}x\right]\mathrm{d}p = 1$$

比较可见，

$$\psi_p^{连续谱} = (\Delta p)^{-1/2}\psi_p^{箱归一} = \sqrt{L/2\pi\hbar}\,\psi_p^{箱归一} = (2\pi\hbar)^{-1/2}\mathrm{e}^{(i/\hbar)px}. \quad \textbf{【证毕】}$$

(五)在给定状态里力学量取值的概率

现在来回答本节一开始提出的问题：在给定状态里测量某一力学量得到不同值的概率是多少．

为此，再来看戴维逊－革末的电子衍射实验，如§5.1.2图1．经过衍射以后，在各个方向上都有可能发现电子，因此，电子处于没有确定动量的状态．用 $\Psi(\boldsymbol{r},t)$ 表示这一状态的波函数．在进行测量时，探测器放在一定的角度 θ，如§5.1.2图1(a)．所探测到的是动量沿 θ 方向的电子，因而是 $\boldsymbol{p}$ 有确定值的电子，这种电子用波函数 $\psi_{\boldsymbol{p}}(\boldsymbol{r})$ 描述．电子原本处于由 $\Psi(\boldsymbol{r},t)$ 描述的状态，而探测到的却是处于 $\psi_{\boldsymbol{p}}(\boldsymbol{r})$ 态的电子．可以认为 $\psi_{\boldsymbol{p}}(\boldsymbol{r})$ 态本来就包含在 $\Psi(\boldsymbol{r},t)$ 之中．探测器放在不同角度都能探测到电子，说明 $\Psi(\boldsymbol{r},t)$ 中不是只包含一个动量为 $\boldsymbol{p}$ 的 $\psi_{\boldsymbol{p}}(\boldsymbol{r})$，而是包含各种不同动量 $\boldsymbol{p}$ 的 $\psi_{\boldsymbol{p}}(\boldsymbol{r})$．换句话说，$\Psi(\boldsymbol{r},t)$ 是各种不同动量 $\boldsymbol{p}$ 的 $\psi_{\boldsymbol{p}}(\boldsymbol{r})$ 的叠加：

$$\Psi(\boldsymbol{r},t) = \sum_{\boldsymbol{p}} c_{\boldsymbol{p}}(t)\psi_{\boldsymbol{p}}(\boldsymbol{r}).$$

这里采用了箱归一化，因而 $\boldsymbol{p}$ 取分立值．在这一展开式中，$\psi_{\boldsymbol{p}}(\boldsymbol{r})$ 的系数的绝对值

平方 $|c_p(t)|^2$ 是在 $\psi(\boldsymbol{r},t)$ 中所包含的 $\psi_p(\boldsymbol{r})$ 的份额，它决定了在状态 $\Psi(\boldsymbol{r},t)$ 中动量取值 $\boldsymbol{p}$ 的概率.

回到一般情况，问在状态 $\Psi(\boldsymbol{r},t)$ 中测力学量 F，得到不同值的概率是多少？为此，先解 $\hat{f}$ 的本征值方程 $\hat{f}\psi_i(\boldsymbol{r}) = f_i\psi_i(\boldsymbol{r})$，然后将 $\Psi(\boldsymbol{r},t)$ 用本征函数 $\psi_i(\boldsymbol{r})$ 展开：

$$\Psi(\boldsymbol{r},t) = \sum_i c_i(t)\psi_i(\boldsymbol{r}). \tag{2.57}$$

为求展开式系数 c_i，将上式中的求和指标（哑标）i 换成 j，然后用 ψ_i^* 乘上式，并对 $d^3\boldsymbol{r}$ 积分：

$$\begin{aligned}\int\psi_i^*(\boldsymbol{r})\Psi(\boldsymbol{r},t)\mathrm{d}^3\boldsymbol{r} &= \sum_j c_j(t)\int\psi_i^*(\boldsymbol{r})\psi_j(\boldsymbol{r})\mathrm{d}^3\boldsymbol{r}\\ &= \sum_j c_j(t)\delta_{ij} = c_i(t).\end{aligned} \tag{2.58}$$

这里用了本征函数的正交归一条件(2.50)式.

我们来证明，展开式系数的绝对值平方 $|c_i(t)|^2$ 决定在 $\Psi(\boldsymbol{r},t)$ 中 F 取值 f_i 的概率 P_i. 为此，计算 F 的平均值. 按概率统计中平均值的定义[见附录(Ⅲ.9)式和(Ⅲ.6)式]，

$$\bar{f} = \sum_i P_i f_i, \qquad P_i \geqslant 0, \qquad \sum_i P_i = 1. \tag{2.59}$$

另一方面，将展开式(2.57)代入§5.2.1(1.13)式，有

$$\begin{aligned}\bar{f} &= \int\Psi^*\hat{f}\Psi\mathrm{d}^3\boldsymbol{r} = \sum_{ij} c_i^* c_j\int\psi_i^*\hat{f}\psi_j\mathrm{d}^3\boldsymbol{r} = \sum_{ij} c_i^* c_j f_j\int\psi_i^*\psi_j\mathrm{d}^3\boldsymbol{r}\\ &= \sum_{ij} c_i^* c_j f_j\delta_{ij} = \sum_i |c_i|^2 f_i.\end{aligned}$$

另外又有：

$$1 = \int\Psi^*\Psi\mathrm{d}^3\boldsymbol{r} = \sum_{ij} c_i^* c_j\int\psi_i^*\psi_j\mathrm{d}^3\boldsymbol{r} = \sum_{ij} c_i^* c_j\delta_{ij} = \sum_i |c_i|^2.$$

将这两个式子和(2.59)式比较，可见

$$P_i = |c_i|^2. \tag{2.60}$$

即：在状态 $\Psi(\boldsymbol{r},t)$ 中测力学量 F 得值 f_i 的概率，是 $\Psi(\boldsymbol{r},t)$ 用 $\hat{f}$ 的本征函数系展开的展开式中，ψ_i 的系数的绝对值平方.❶ 【证毕】

对于连续谱可以类似地讨论. 此时，展开式(2.57)成为

---

❶ 严格说来，$|c_i|^2$ 是在状态 $\Psi(\boldsymbol{r},t)$ 中的粒子处于 $\hat{f}$ 的本征态 ψ_i 的概率. 如果本征值 f_i 不简并，这也就是测得 f_i 的概率. 但是如果 f_i 有 s 重简并，对应有 s 个正交归一的本征态 $\psi_{i\alpha}$，$\alpha = 1,2,\cdots,s$，则测得 f_i 的概率是 $P_i = \sum_{\alpha=1}^{s} |c_{i\alpha}|^2$.

$$\Psi(\boldsymbol{r},t) = \int c_f(t)\psi_f(\boldsymbol{r})\,\mathrm{d}f. \tag{2.61}$$

利用 ψ_f 的正交归一条件(2.52)式可以求出展开系数

$$c_f(t) = \int \psi_f^*(\boldsymbol{r})\Psi(\boldsymbol{r},t)\,\mathrm{d}^3\boldsymbol{r}. \tag{2.62}$$

F 取值 f 的概率密度是

$$\rho_f = |c_f|^2, \tag{2.63}$$

它归一化为

$$\int \rho_f \mathrm{d}f = \int |c_f|^2 \mathrm{d}f = 1. \tag{2.64}$$

这样,我们就回答了,在给定状态里测量某一力学量得到不同值的概率分布如何计算的问题.

在以上的论证中,一个重要条件是:用本征函数 $\psi_f(\boldsymbol{r})$ 能够展开任意态函数 $\Psi(\boldsymbol{r},t)$,即:(2.57)式或(2.61)式能成立,这称为本征函数系的完备性.从数学上证明本征函数系的完备性比较复杂.从物理考虑可以推断,如果 $\Psi(\boldsymbol{r},t)$ 是微观粒子的一个实际运动状态,且在这一状态中能对力学量 F 进行测量,则一定能进行(2.57)式或(2.61)式的展开.

【例6】 粒子在如(2.33)式所示的宽度为 a 的一维无限深势阱(势箱)中运动,所处状态的波函数为 $\psi(x) = Ax(a-x)$,求在该状态粒子能量可取的数值,以及取这些数值的概率.

解 一维无限深势阱(势箱)中粒子的能量本征值和本征函数已在例4中求出,见(2.39)、(2.41)式.按(2.57)式将波函数 $\psi(x)$ 用本征函数展开:

$$Ax(a-x) = \sum_n c_n \sqrt{2/a}\sin(n\pi x/a),$$

其中

$$\begin{aligned} c_n &= \int_0^a \psi_n^*(x)\psi(x)\,\mathrm{d}x \\ &= \int_0^a Ax(a-x)\sqrt{2/a}\sin(n\pi x/a)\,\mathrm{d}x \\ &= A\sqrt{2/a}\,2a^3[1-(-1)^n]/(n\pi)^3. \end{aligned}$$

当 n 为偶数时,c_n 为0,所以粒子能量所取的值为

$$E_n = n^2\pi^2\hbar^2/(2ma^2), \qquad n = 1,3,5,\cdots.$$

其概率为 $P_n = |c_n|^2 = A^2 32a^5/(n^6\pi^6)$.由 $\psi(x)$ 的归一化条件得 $A^2 = 30/a^5$,因而

$$P_n = 960/(n\pi)^6, \qquad n = 1,3,5,\cdots.$$

现在来小结一下量子力学对微观粒子力学量的描述.对于微观粒子,在给定的

状态中力学量一般不具有确定值，而是以不同的概率取不同的值. 为了反映这一特征，量子力学建立了关于力学量的原理：力学量用线性厄米算符表示；力学量取值的谱是其算符的本征值谱；在给定状态中力学量取某一本征值的概率，是将状态波函数用本征函数系展开后，在展开式中的相应系数的绝对值平方. 根据这一原理，可以预言在给定状态测量力学量的结果.

§5.2.3　力学量的完全组　算符的对易关系

在第一章已经指出，量子力学中，系统的状态由波函数描述，波函数决定了给定状态里系统的所有性质.

在很多情况下，我们感兴趣的是某些力学量有确定值的状态. 例如，对于自由粒子，我们常常感兴趣的是动量有确定值的状态；对于中心力场中的粒子，如原子中的电子，我们感兴趣的是角动量有确定值的状态. 在这种情况下，系统的状态就是这些力学量的本征态，因而可以用这些力学量所取的值（即它们的一组本征值）的量子数，来标志系统的状态.

需要注意的是，在量子力学里，有些力学量不能同时有确定值，因而不能同时被测定. 例如，德布罗意波函数[§5.1.2(2.17)式]是动量 p_x 的本征函数，因而在它所描述的状态里，动量 p_x 有确定值，但是，坐标 x 却没有确定值. 在这一个状态中，$|\psi(x,t)|^2=|A|^2$ 是常数，说明粒子的坐标 x 可以以相同的概率取任何值.

人们自然不能期望用不能同时被测定的力学量来确定系统的状态. 因此，在量子力学里，用来标志系统状态的力学量必须是一些能同时有确定值，因而可以同时被测定的力学量. 我们知道，量子力学中力学量用线性厄米算符表示，因而就提出了这样的问题，如果已知几个力学量的算符，如何判断它们是否能同时被测定呢？这一节就来讨论力学量同时有确定值的条件.

（一）力学量同时有确定值的条件

由 §5.2.2 知道，如果系统处于力学量算符 $\hat{f}$ 的本征函数所描述的状态，则力学量 F 有确定值，就是相应的本征值. 这些本征函数形成正交完备函数系，见 §5.2.2(五)中的讨论. 显然，如果这个函数系中的函数也是其他几个力学量算符的本征函数，则在相应的状态中，这些力学量就都有确定值，分别取它们相应的本征值. 例如，对于自由粒子，德布罗意波函数 $\psi_{p_x}(x)=Ae^{ip_x x/\hbar}$ 既是粒子动量算符的本征函数，也是粒子能量算符的本征函数，所以动量和能量在此状态能同时有确定值，即它们各自的本征值 p_x 和 $p_x^2/(2m)$. 由不同 p_x 标志的 $\psi_{p_x}(x)$ 形成完备函数系. 因此，判断几个力学量是否同时可测的问题，实际上就是判断它们的算符是否有共

同本征函数系的问题●.

应该注意,力学量同时可测,要求它们的算符有共同的本征函数系,而不只是有个别的共同本征函数.例如,在§5.2.2例3中曾看到,$\hat{L}_x,\hat{L}_y,\hat{L}_z$有一个共同本征函数$Y_{00}$,然而它们没有共同本征函数系,因而不能同时用来作为标志系统状态的变量.

在回答算符有共同本征函数系的条件这一问题之前,先来简单讨论一下算符的代数运算.算符不是普通的数,但是可以对它们定义代数运算.

算符加法的定义是:若算符$\hat{A}$和$\hat{B}$分别作用到任一函数u上所得结果的和,等效于算符$\hat{C}$的作用,则称$\hat{C}$为$\hat{A}$与$\hat{B}$的和.即:

$$若\ \hat{C}u = \hat{A}u + \hat{B}u\quad(u\ 为任意函数),\quad 则\ \hat{C} = \hat{A} + \hat{B}. \tag{3.1}$$

算符乘法的定义是:若算符$\hat{A}$和$\hat{B}$相继作用到任一函数u上所得结果,等效于算符$\hat{C}$的作用,则称$\hat{C}$为$\hat{A}$和$\hat{B}$的积.即:

$$若\ \hat{C}u = \hat{A}(\hat{B}u)\quad(u\ 为任意函数),\quad 则\ \hat{C} = \hat{A}\hat{B}. \tag{3.2}$$

注意:

①在上述定义中的u必须是任意函数.如果只是对于某些函数有$\hat{C}u = \hat{A}u + \hat{B}u$或$\hat{C}u = \hat{A}(\hat{B}u)$,还不能说$\hat{C} = \hat{A} + \hat{B}$或$\hat{C} = \hat{A}\hat{B}$.

②算符的加法满足通常的代数法则,算符的乘法也满足通常的乘法结合律和分配律.但需特别强调的是,算符的乘法一般不满足交换律,即,一般说来,算符的乘积与次序有关,不可对易:$\hat{A}\hat{B} \neq \hat{B}\hat{A}$,其原因是,一般说来,$\hat{A}(\hat{B}u) \neq \hat{B}(\hat{A}u)$.

例如,很容易证明:

$$\left.\begin{aligned}&\hat{x}(\hat{p}_x\psi) = x(-i\hbar\partial/\partial x)\psi = -i\hbar x\partial\psi/\partial x,\\&\hat{p}_x(\hat{x}\psi) = -i\hbar\partial(x\psi)/\partial x = -i\hbar x\partial\psi/\partial x - i\hbar\psi \neq \hat{x}(\hat{p}_x\psi).\end{aligned}\right\} \tag{3.3}$$

式中ψ是一个任意函数.由上面的计算可以看出,$\hat{x}$和$\hat{p}_x$不可对易:$\hat{x}\hat{p}_x \neq \hat{p}_x\hat{x}$,即:坐标算符和动量算符的相同分量不可对易.又如,利用§5.2.2的(2.20)~(2.22)式很容易看到,对一个任意函数ψ,$L_x(L_y\psi) \neq L_y(L_x\psi)$,…,因而,$L_xL_y \neq L_yL_x$,…,即:角动量算符的不同分量之间不可对易.

算符的对易性与它们是否有共同本征函数系有密切关系.我们有以下定理:

定理一　若力学量算符$\hat{F}$和$\hat{G}$有共同本征函数系,则$\hat{F}$和$\hat{G}$相互对易:$\hat{F}\hat{G} = \hat{G}\hat{F}$.

证　设ψ_n是$\hat{F}$和$\hat{G}$的共同本征函数系中的任一函数,即有

● 为了行文简洁,以下常将两个力学量的算符没有共同本征函数系,简称为它们"不能同时有确定值",或"非同时可测";而将两个力学量的算符有共同本征函数系,简称为它们"能同时有确定值",或"同时可测".

$$\hat{F}\psi_n = f_n\psi_n, \qquad \hat{G}\psi_n = g_n\psi_n,$$

用 $\hat{G}$ 左乘第一式的两边，用 $\hat{F}$ 左乘第二式的两边，并将所得结果相减，则有

$$(\hat{F}\hat{G} - \hat{G}\hat{F})\psi_n = (f_n g_n - f_n g_n)\psi_n = 0.$$

由于 ψ_n 不是任意的函数，因而还不能由上式得出 $\hat{F}$ 和 $\hat{G}$ 可对易的结论. 但是，注意到任意函数都可用本征函数系展开，即 $\psi = \sum_n c_n\psi_n$，则从上式可以导致下面的结果：

$$(\hat{F}\hat{G} - \hat{G}\hat{F})\psi = \sum_n c_n(\hat{F}\hat{G} - \hat{G}\hat{F})\psi_n = 0.$$

这就意味着 $\hat{F}$ 和 $\hat{G}$ 相互对易，$\hat{F}\hat{G} - \hat{G}\hat{F} = 0$. 【证毕】

定理二　若 $\hat{F}$ 和 $\hat{G}$ 可对易，则它们有共同本征函数系.

证　如果 $\hat{F}$ 和 $\hat{G}$ 的本征值都没有简并，定理的证明比较简单. 设 ψ_n 是 $\hat{F}$ 的任一个本征函数，即有 $\hat{F}\psi_n = f_n\psi_n$. 用 $\hat{G}$ 从左边作用于上式两端，注意到 $\hat{F}\hat{G} = \hat{G}\hat{F}$，则有

$$\hat{F}(\hat{G}\psi_n) = \hat{G}\hat{F}\psi_n = f_n(\hat{G}\psi_n).$$

上式表明，$\hat{G}\psi_n$ 也是 $\hat{F}$ 的本征函数，且对应的本征值也是 f_n. 由于已假定 f_n 不简并，故 $\hat{G}\psi_n$ 和 ψ_n 描述同一个本征态，因而 $\hat{G}\psi_n$ 和 ψ_n 只差一个相乘的常数. 令这个常数为 g_n，则有 $\hat{G}\psi_n = g_n\psi_n$. 这样就证明了 ψ_n 也是 $\hat{G}$ 的本征函数. 在以上证明中，假定了 ψ_n 是 $\hat{F}$ 的任一个本征函数，所以也就证明了：若 $\hat{F}\hat{G} = \hat{G}\hat{F}$，则 $\hat{F}$ 和 $\hat{G}$ 有共同本征函数系. 【证毕】

在算符本征值有简并的情况下，经过稍复杂点的证明，也可以得到同样的结论. 留作练习.

以上两个定理表明，几个力学量同时可测的充分必要条件是它们的算符相互对易.

(二)对易关系

微观粒子的力学量一般说来不具有确定值，两个力学量一般说来不能同时有确定值. 这是量子力学与经典力学相区别的特点，是微观粒子波粒二象性的体现. 上面我们看到，这与力学量的算符不可对易有直接关系. 因此，不可对易性是量子力学中的力学量(算符)区别于经典力学中的力学量(普通的数)的根本性质.

通常称乘法服从交换律，因而可对易的数为 c 数；而称乘法不服从交换律，因而不可对易的数为 q 数. 经典力学中的力学量用 c 数表示；量子力学中的力学量用 q 数表示. 算符就是一种 q 数. 我们知道，矩阵的乘法一般不服从交换律，所以矩阵也是 q 数.

利用对易关系可以方便地研究两个算符的对易性质. 设 $\hat{A}, \hat{B}$ 为两个算符，则

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} \tag{3.4}$$

称为它们的对易关系.

【例2】 求坐标算符与动量算符的对易关系.

解 在计算对易关系的时候,特别要注意的是:算符的意义在于它对函数的作用,而算符的乘法是对函数的相继作用.因此,要计算算符的对易关系,必须把它们作用到任意函数上来计算.对于坐标算符和动量算符,

$$[\hat{x},\hat{p}_x] = \hat{x}\hat{p}_x - \hat{p}_x\hat{x} = x(-i\hbar\partial/\partial x) - (-i\hbar\partial/\partial x)x,$$

但是决不能误认为它等于零.应该把它作用到函数 ψ 上:

$$[\hat{x},\hat{p}_x]\psi = x(-i\hbar\partial/\partial x)\psi - (-i\hbar\partial/\partial x)x\psi.$$

右边的两项已在(3.3)式中分别算出,将它们相减得到 $[\hat{x},\hat{p}_x]\psi = i\hbar\psi$,由于 ψ 是任意函数,故有

$$[\hat{x},\hat{p}_x] = i\hbar. \tag{3.5}$$

类似地有

$$[\hat{y},\hat{p}_y] = i\hbar, \qquad [\hat{z},\hat{p}_z] = i\hbar. \tag{3.6}$$

另一方面,容易看到,坐标算符和动量算符的不同分量之间相互可对易:

$$[\hat{x},\hat{p}_y] = [\hat{y},\hat{p}_x] = \cdots = 0. \tag{3.7}$$

并且坐标算符的不同分量之间可对易,动量算符的不同分量之间也可对易.将(3.5)~(3.7)式写在一起就是

$$[\hat{x}_i,\hat{x}_j] = 0, \qquad [\hat{p}_i,\hat{p}_j] = 0, \qquad [\hat{p}_i,\hat{x}_j] = -i\hbar\delta_{ij}. \tag{3.8}$$

这就是坐标和动量算符各分量之间的全部对易关系.

将(3.8)式和§1.6.2(2.17)式比较可见,两者非常类似.只要将经典泊松符号按以下方式过渡为算符对易关系:

$$[A,B]_{PB} \quad \Rightarrow \quad (i/\hbar)[\hat{A},\hat{B}], \tag{3.9}$$

就可以由§1.6.2(2.17)式得到本节的(3.8)式.

为了便于计算不同力学量算符的对易关系,我们写出对易关系运算的代数法则[注意和§1.6.2(2.11)~(2.15)式及(2.18)式对比]:

$$[\hat{A},\hat{B}] = -[\hat{B},\hat{A}], \tag{3.10}$$

$$[\hat{A},\hat{A}] = 0, \tag{3.11}$$

$$[\alpha\hat{A},\beta\hat{B}] = \alpha\beta[\hat{A},\hat{B}], \tag{3.12}$$

$$[\hat{A},\hat{B}+\hat{C}] = [\hat{A},\hat{B}] + [\hat{A},\hat{C}], \tag{3.13}$$

$$[\hat{A}+\hat{B},\hat{C}] = [\hat{A},\hat{C}] + [\hat{B},\hat{C}], \tag{3.14}$$

$$[\hat{A},\hat{B}\hat{C}] = [\hat{A},\hat{B}]\hat{C} + \hat{B}[\hat{A},\hat{C}], \tag{3.15}$$

$$[\hat{A}\hat{B},\hat{C}] = [\hat{A},\hat{C}]\hat{B} + \hat{A}[\hat{B},\hat{C}], \tag{3.16}$$

$$[\hat{A},[\hat{B},\hat{C}]] + [\hat{B},[\hat{C},\hat{A}]] + [\hat{C},[\hat{A},\hat{B}]] = 0, \tag{3.17}$$

其中，α,β 为任意复数. 这些式子都不难从对易关系的定义(3.4)式直接证明. 特别值得注意的是(3.15)、(3.16)式类似于乘积求导法则. 例如，(3.15)式中算符 $\hat{A}$ 和另两个算符的乘积 $\hat{B}\hat{C}$ 的对易关系等于 $\hat{A}$ 和这一乘积中的两个因子分别取对易关系，乘上另一因子再相加. 只是在这样做的时候，乘积中两个因子的次序($\hat{B}$ 在 $\hat{C}$ 的左边)不能颠倒. 对于(3.16)式也可同样分析.

【例 3】 证明坐标、动量算符和角动量算符有如下对易关系(为书写简单，略去表示算符的尖角).

$$[L_x,y]=\mathrm{i}\hbar z,\quad [L_y,z]=\mathrm{i}\hbar x,\quad [L_z,x]=\mathrm{i}\hbar y, \tag{3.18}$$

$$[L_x,p_y]=\mathrm{i}\hbar p_z,\quad [L_y,p_z]=\mathrm{i}\hbar p_x,\quad [L_z,p_x]=\mathrm{i}\hbar p_y. \tag{3.19}$$

证　首先利用三阶完全反对称张量 ε_{ijk} 写出叉乘积(见附录Ⅰ)：

$$L_i=(\boldsymbol{r}\times\boldsymbol{p})_i=\sum_{lm}\varepsilon_{ilm}x_lp_m. \tag{3.20}$$

利用(3.16)式有

$$\begin{aligned}[L_i,x_j]&=\sum_{lm}\varepsilon_{ilm}[x_lp_m,x_j]=\sum_{l}\varepsilon_{ilm}([x_l,x_j]p_m+x_l[p_m,x_j])\\&=\sum_{lm}\varepsilon_{ilm}x_l(-\mathrm{i}\hbar\delta_{mj})=\mathrm{i}\hbar\sum_{l}\varepsilon_{ijl}x_l.\end{aligned} \tag{3.21}$$

这里用了(3.8)式以及 ε_{ilm} 对脚标的反对易性 $\varepsilon_{ilm}=-\varepsilon_{iml}$. 将(3.21)式明显写出就是(3.18)式. 类似地可以证明

$$[L_i,p_j]=\mathrm{i}\hbar\sum_{l}\varepsilon_{ijl}p_l. \tag{3.22}$$

明显写出就是(3.19)式. **【证毕】**

【例 4】 证明角动量算符不同分量之间，以及它们和角动量平方之间有如下对易关系：

$$[L_x,L_y]=\mathrm{i}\hbar L_z,\quad [L_y,L_z]=\mathrm{i}\hbar L_x,\quad [L_z,L_x]=\mathrm{i}\hbar L_y, \tag{3.23}$$

$$[\boldsymbol{L}^2,L_x]=[\boldsymbol{L}^2,L_y]=[\boldsymbol{L}^2,L_z]=0. \tag{3.24}$$

这些等式不难利用(3.15)式和(3.21)、(3.22)式证明，留作练习. (3.23)式可合写为

$$[L_i,L_j]=\mathrm{i}\hbar\sum_{l}\varepsilon_{ijl}L_l. \tag{3.25}$$

有时也写成

$$\boldsymbol{L}\times\boldsymbol{L}=\mathrm{i}\hbar\boldsymbol{L}. \tag{3.26}$$

将此式展开，可以看到它与(3.23)式等效. 注意，经典矢量的自叉乘为零. 这里的 $\boldsymbol{L}$ 为算符，各个分量不可对易，因而自叉乘不为零.

注意，在例 3 的证明中，没有将对易关系按定义(3.4)式展开，而是运用了对易关系的运算法则(3.16)式. 这使得运算非常简洁. 在进行对易关系的运算时，例

如在证明例4中的几个公式时，一定要注意充分利用对易关系的运算法则(3.10)～(3.16)式.

(三)算符乘积的厄米性

我们知道，量子力学中的物理量必须用厄米算符表示，其原因是厄米算符的本征值和平均值是实数. 但是，两个厄米算符的乘积不一定是厄米算符，因此需要加以研究.

定理三 两个算符乘积的厄米共轭等于它们各自的厄米共轭颠倒次序相乘.

$$(FG)^{\dagger} = G^{\dagger}F^{\dagger} \tag{3.27}$$

证 根据算符的厄米共轭条件 §5.2.1(1.23)式有

$$\int \psi^{*}(FG)^{\dagger}\psi \mathrm{d}\tau = \int \psi (FG)^{*}\psi^{*}\mathrm{d}\tau = \int \psi F^{*}(G\psi)^{*}\mathrm{d}\tau$$

$$= \int (G\psi)^{*}F^{\dagger}\psi \mathrm{d}\tau = \int (F^{\dagger}\psi)G^{*}\psi^{*}\mathrm{d}\tau = \int \psi^{*}G^{\dagger}F^{\dagger}\psi \mathrm{d}\tau.$$

由于 ψ 是任意函数，即得(3.27)式. **【证毕】**

定理四 两个厄米算符的乘积仍为厄米算符的充分必要条件是它们可对易.

证 设 $F^{\dagger} = F, G^{\dagger} = G$. 则由(3.27)式 $(FG)^{\dagger} = G^{\dagger}F^{\dagger} = GF$. 如果 F, G 对易，则由此得 $(FG)^{\dagger} = FG$，即 FG 为厄米算符. 反过来，如果 F 和 G 不可对易，则 $(FG)^{\dagger} \neq FG$，故 FG 不厄米. **【证毕】**

定理五 在 $\hat{F}$ 和 $\hat{G}$ 不可对易的情况下，它们的对易关系乘上虚因子 i 是厄米算符，即：如果 $\hat{F}$，$\hat{G}$ 厄米，而

$$[\hat{F}, \hat{G}] = i\hat{C}, \tag{3.28}$$

则 $\hat{C}$ 为厄米算符，证明见习题. 特殊情况下，$\hat{C}$ 是一个实常数. 考查前面的(3.5)式和例3、例4可见，它们都符合(3.28)式. **【证毕】**

(四)力学量的完全组

按照本节开始的分析，对于一个给定的系统，能同时可测的一组独立力学量的值可以用来标志系统的状态. 现在知道，这组力学量的算符是相互对易的. 我们称这样一组能完全标志系统状态的独立力学量为力学量的完全组. 完全组中力学量的数目等于系统的自由度数❶. 一个量子力学系统的自由度数是由实验确定的. 在一些简单情况下，量子系统的自由度数目等于相应的经典力学系统的自由度. 例如，无自旋的粒子，和经典力学中的粒子一样，自由度数为3.

---

❶ 在这一点上，量子力学与经典力学不同. 在经典力学中，为了确定系统的状态，所需的力学量的数目是系统自由度的两倍. 例如，为确定一个自由运动粒子的状态需给定初始时刻的 $x_i, p_i (i = 1,2,3)$，而在量子力学中，确定系统状态所需力学量的数目等于系统的自由度. 其原因是，经典力学运动方程对时间是二阶微分方程，而量子力学运动方程对时间是一阶的，见 §5.3.1(1.5)式.

【例 5】　氢原子中的电子在核的库仑场中运动，略去电子的自旋，试给出它的一个力学量完全组.

解　库仑场是有心力场，应该用角动量而不是动量来确定它的状态. 由例 4 知道，角动量的三个分量 $\hat{L}_x,\hat{L}_y,\hat{L}_z$ 不可对易，因而不能有共同本征函数系，不能取作力学量完全组. 由同一例题又知道 $\hat{\boldsymbol{L}}^2$ 和 $\hat{L}_z$ 有共同本征函数系，因此，只需要再找一个和它们对易的算符. 容易证明，能量算符 $\hat{H}$ 符合这一条件. 在有心力场中的能量是

$$E = \boldsymbol{p}^2/(2m) + U(r). \tag{3.29}$$

因而能量算符是

$$\hat{H} = [\hat{\boldsymbol{p}}^2/(2m)] + U(r) = -[\hbar^2/(2m)]\nabla^2 + U(r). \tag{3.30}$$

将球坐标中的拉普拉斯算符代入，得到

$$\hat{H} = -\frac{\hbar^2}{2mr^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{\hbar^2}{2mr^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right] + U(r).$$

再用 §5.2.2(2.23)式，得到

$$\hat{H} = -\frac{\hbar^2}{2mr^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\hat{\boldsymbol{L}}^2}{2mr^2} + U(r). \tag{3.31}$$

和 §1.3.2(2.6)式相比，可以看到它们之间的对应关系. 由于没有取 $\varphi = 0$ 的特殊坐标系，所以 §1.3.2(2.6)式中的 p_θ^2 应写为 $\boldsymbol{L}^2$，而同一式子右边的第一项可改写为 $p_r^2/(2m)$，过渡到算符

$$\boldsymbol{L}^2 \to \hat{\boldsymbol{L}}^2, \qquad p_r^2 \to -\frac{\hbar^2}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) \tag{3.32}$$

就成为(3.31)式. 注意到 $\hat{\boldsymbol{L}}^2$ 和 $\hat{L}_z$ 是作用在角变量 θ、φ 上的算符[见 §5.2.2(2.20)、(2.23)式]，容易看到 $\hat{H}$ 和 $\hat{\boldsymbol{L}}^2$ 及 $\hat{L}_z$ 相互对易. 因此，可以选 $\hat{H}$、$\hat{\boldsymbol{L}}^2$ 和 $\hat{L}_z$ 作为有心力场(例如氢原子的库仑场)中粒子的力学量完全组. 原子中电子的状态用 $\hat{H}$，$\hat{\boldsymbol{L}}^2$ 和 $\hat{L}_z$ 的量子数，主量子数 n，角量子数 l 和磁量子数 m 来标定：$\Psi = \Psi_{nlm}(\boldsymbol{r},t)$. 这是原子物理中熟知的事实.

如果已知确定系统状态的力学量完全组的一组本征值，描述系统这个状态的波函数也就知道了，就是这组力学量算符的，相应于这一组本征值的共同本征函数. 因而，用力学量完全组标志状态和用波函数描述状态是一致的. 当然，在量子力学中除了力学量有确定值的本征态外，还有力学量不具有确定值的状态，这种状态不能借助于力学量完全组的一组本征值来确定. 这种状态的波函数是力学量完全组的本征函数的叠加.

§5.2.4 不确定关系

在上一节里证明了,如果两个力学量的算符不可对易,则它们不能同时有确定值.这一节对这样两个物理量不能同时确定的程度作定量的研究.

设某一力学量 F 在状态 ψ 中的平均值为 $\bar{f}$. F 在该状态中取值的不确定程度用它对 $\bar{f}$ 的偏差的平方平均值 $\overline{(\Delta f)^2} = \overline{(\hat{f} - \bar{f})^2}$ 来量度.采用符号 $\Delta f = \sqrt{\overline{(\Delta f)^2}}$ 来表示 f 的偏差.如果两个力学量 F 和 G 不能同时确定,则它们在任意状态中的偏差 Δf 和 Δg 不能同时为零.现在要讨论的是, Δf 和 Δg 之间的定量关系.我们以坐标和动量为例来进行这一讨论.

为了简单起见,考虑一维情况,坐标和动量的算符不可对易,它们的对易关系是§5.2.3(3.5)式(略去动量算符的脚标 x):

$$[\hat{x},\hat{p}] = \mathrm{i}\hbar. \tag{4.1}$$

根据§5.2.3 定理一就知道,坐标和动量不能同时有确定值.

举几个具体的波函数为例.动量有确定值的状态是德布罗意波,即:§5.2.2(2.53)式:

$$\psi_p(x) = (2\pi\hbar)^{-1/2}\mathrm{e}^{(\mathrm{i}/\hbar)px}. \tag{4.2}$$

在这一个状态中, $p = p, \Delta p = 0$,动量 p 有确定值;而与此同时,坐标 x 的概率密度 $|\psi_p(x)|^2$ 是常数,即:坐标取任何值的概率都相等,完全没有确定值, $\Delta x = \infty$.

另一种极端情况是坐标有确定值 x_0 的状态.我们知道,在 x_0 点有一个点电荷的电荷密度可以用 δ 函数 $\delta(x - x_0)$ 表示(见附录Ⅰ).类似地,坐标有确定值 x_0 的量子力学状态的波函数也可以用 δ 函数表示为

$$\psi_{x_0}(x) = \delta(x - x_0). \tag{4.3}$$

在这一状态中, $x = x_0, \Delta x = 0$.为了知道在这一状态中动量取值的概率,将它用动量本征函数展开,并利用§5.2.2(2.62)式求展开系数:

$$c_p(t) = \int\psi_p^*(x)\delta(x - x_0)\mathrm{d}x = \psi_p^*(x_0) = (2\pi\hbar)^{-1/2}\mathrm{e}^{-(\mathrm{i}/\hbar)px_0}. \tag{4.4}$$

动量的概率密度 $|c_p|^2$ 是常数,表明动量取任意值的概率都相等,完全没有确定值, $\Delta p = \infty$.

以上两个例子表明,如果在某个状态中 $\Delta p = 0$,则 $\Delta x = \infty$;如果在某个状态中 $\Delta x = 0$,则 $\Delta p = \infty$.这是两种极端情况.在更一般的情况下, Δx 取一个非零的有限值($\Delta x = a > 0$),此时, Δp 不能是零,也不为无穷大,我们来建立 Δx 和 Δp 之间的定量关系.

考虑 Δx 和 Δp 都不为零的情况. 这样的状态 $\psi(x)$ 可以写成一系列动量为 p 的单色平面波——德布罗意波的叠加[这实质上是函数 $\psi(x)$ 的傅里叶展开]:

$$\psi(x) = (2\pi\hbar)^{-1/2}\int c_p e^{(i/\hbar)px}dp. \tag{4.5}$$

其中 c_p 是只在平均值 $\bar{p} = p_0$ 周围宽为 Δp 的区间内显著不为零的函数. 由单色平面波这样叠加而成的状态称为波包. 为了简单起见,假定 c_p 在 p_0 周围对称地分布.

由于 c_p 只在 p_0 周围的一个小区间内显著不为零,我们从 $\psi(x)$ 中抽出一个动量为 p_0 的单色平面波因子,写成

$$\psi(x) = g(x)e^{(i/\hbar)p_0x}. \tag{4.6}$$

归一化条件是

$$\int_{-\infty}^{\infty} |\psi|^2 dx = \int_{-\infty}^{\infty} |g(x)|^2 dx = 1. \tag{4.7}$$

我们来求波包的展开系数 c_p. 按 §5.2.2(2.62)式

$$c_p = (2\pi\hbar)^{-1/2}\int \psi(x)e^{-(i/\hbar)px}dx = (2\pi\hbar)^{-1/2}\int g(x)e^{(i/\hbar)(p_0-p)x}dx. \tag{4.8}$$

按假设 c_p 在 $p = p_0$ 周围对称地分布,因而 $g(x)$ 为偶函数. 对于不同的 $g(x)$,有不同的 c_p,下面举两个例子:

【例 1】 设 $g(x)$ 局限在 $|x| < a/2$ 之中,而在此区间中它是常数:

$$g(x) = \begin{cases} 0, & \text{当 } |x| > a/2, \\ 1/\sqrt{a}, & \text{当 } |x| \leqslant a/2, \end{cases} \tag{4.9}$$

如图 1(a),求 c_p 和 Δp.

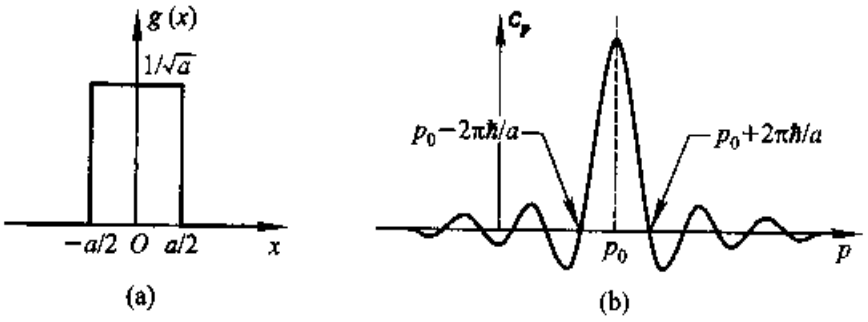


图 1　矩形波包

解　这一 $g(x)$ 显然是满足(4.7)式的偶函数. 由(4.8)式

$$c_p = \frac{1}{\sqrt{2\pi a\hbar}}\int_{-a/2}^{a/2} e^{-(i/\hbar)(p-p_0)x}dx = \sqrt{\frac{a}{2\pi\hbar}}\frac{\sin[(p-p_0)a/2\hbar]}{(p-p_0)a/2\hbar}. \tag{4.10}$$

这是一个以 $p = p_0$ 为中心,对称分布的振荡函数. 在 $p = p_0$ 处的主峰最高,以后的峰愈来愈低,如图 1(b). 由此可见,当 x 局限在 $(-a/2, a/2)$ 区间中时,p 的概率分布也集中在 p_0 的周围.

实际的波函数要满足连续性条件，在超出 $(-a/2, a/2)$ 区间时应连续地趋于零. (4.9)式是它的简化描述. 由于(4.9)式中有两个不连续点，按定义直接用(4.10)式计算 Δp 比较困难. 但可以利用图 1 作出估计. 由图 1(b)可见，$\Delta p \sim \pi\hbar/a$，因而

$$\Delta x \Delta p \sim \pi\hbar. \tag{4.11}$$

【例 2】 设 $|g(x)|^2$ 为 $\Delta x = a$ 的高斯分布：

$$g(x) = \frac{1}{(2\pi)^{1/4}\sqrt{a}} e^{-x^2/(2a)^2}, \tag{4.12}$$

[见附录(Ⅲ.33)式]，求 c_p 和 Δp.

解 (4.12)式中的 $g(x)$ 是对称的，利用积分公式(见附录Ⅴ)：

$$\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\pi/a}, \tag{4.13}$$

可以证明它满足归一化条件(4.7)式. 由(4.8)式

$$c_p = \frac{1}{\sqrt{2\pi a\hbar}} \cdot \frac{1}{(2\pi)^{1/4}} \cdot \int_{-\infty}^{\infty} \exp\left[-\frac{x^2}{4a^2} + \frac{i(p_0 - p)x}{\hbar}\right] dx \tag{4.14}$$

作代换 $\xi = x + 2a^2 i(p_0 - p)/\hbar$，上式成为❶

$$c_p = \frac{1}{\sqrt{2\pi a\hbar}} \cdot \frac{1}{(2\pi)^{1/4}} \cdot e^{-a^2(p_0-p)^2/\hbar^2} \cdot \int_{-\infty}^{\infty} e^{-\xi^2/4a^2} d\xi.$$

再利用(4.13)式得到

$$c_p = \sqrt{2a/\hbar}(2\pi)^{-1/4} e^{-a^2(p-p_0)^2/\hbar^2}. \tag{4.15}$$

可见，c_p 也是一个高斯分布，其中心点在 p_0，如图 2.

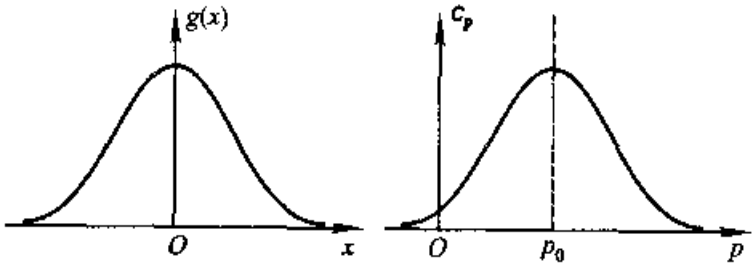


图 2 高斯型波包

利用 c_p 可将 p 的方均偏差写成

$$\overline{(\Delta p)^2} = \int_{-\infty}^{\infty} (p - p_0)^2 c_p^2 dp.$$

用积分公式(见附录Ⅴ)：

$$\int_{-\infty}^{\infty} x^2 e^{-ax^2} dx = \frac{\sqrt{\pi}}{2} a^{-3/2}. \tag{4.16}$$

得到 $\overline{(\Delta p)^2} = \hbar^2/(4a^2) = \hbar^2/(4\overline{(\Delta x)^2})$，因而

$$\Delta x \Delta p = \hbar/2. \tag{4.17}$$

❶ 这一代换相当于将积分路径在复平面上沿虚轴平移. 由于被积函数无奇点，利用柯西定理可以证明积分结果不变.

从这两个例子可以看到，对于不同的波包，Δx 和 Δp 的乘积略有不同. 但都 $\sim \hbar$. 可以证明，高斯型波包的 $\Delta x \Delta p$ 乘积最小（称为最小不确定波包），因而有

$$\Delta x \Delta p \geqslant \hbar/2. \tag{4.18}$$

这称为坐标动量的不确定关系.

不确定关系是微观粒子二象性的体现. 正是由于二象性，使坐标和动量不能同时确定，而不确定关系(4.18)式定量地表示了它们的不确定程度之间的关联.

坐标和动量不能同时确定，导致了其他一些力学量不能同时确定. 例如由坐标和动量的对易关系可以证明，角动量三个分量的算符不相互对易，见 §5.2.3 例 4. 这就是说，角动量的三个分量不能同时有确定值，不能期望像普通矢量那样，通过确定它的三个分量的数值来确定角动量矢量. 再如，由坐标和动量的对易关系可以证明，粒子的动能算符 $\boldsymbol{p}^2/(2m)$ 和势能 $U(\boldsymbol{r})$ 不可对易，不能同时有确定值. 因而不能期望像在经典力学中那样，通过分别确定动能和势能的大小来确定总能量的数值. 总能量的数值只能通过直接求解能量算符的本征值方程来决定，例如 §5.2.2 例 4.

坐标和动量的不确定关系常用来对一些物理问题作定性分析. 当知道系统的空间限度时，可以用它判断动量的取值限度；反之也是一样. 在近代高能碰撞实验中，常常用加速某种粒子作射弹去轰击靶粒子，探测靶粒子的结构. 为此要求射弹粒子有准单色性，即近似地有确定的动量，动量的相对偏差 $(\Delta p)/\bar{p}$ 不能大. 但是根据不确定关系，Δp 小意味着 Δx 大. 这就限制了对靶粒子微观结构的精细探测. 这和波长长的光分辨本领低，在微小粒子上会发生衍射，不能看清微粒的轮廓是同样的道理. 解决办法是提高射弹粒子的动量 p、对子同样的相对偏差，$\bar{p}$ 越大，可以容许的 Δp 就越大，从而 Δx 越小. 这就是为什么要用高能粒子流探测粒子微观结构的原因.

【例 3】 用动量为 4MeV/c 的 α 粒子探测原子的结构；用动量为 100GeV/c 的电子探测核子的结构. 粗略估计这两种情况下探测到的空间范围的尺度.

解 已知动量的单位（见附录Ⅶ）

$$1\text{eV}/c = 10^{-6}\text{MeV}/c = 10^{-9}\text{GeV}/c = 5.34 \times 10^{-23}\text{g} \cdot \text{cm} \cdot \text{s}^{-1},$$

因此，4MeV/c 的 α 粒子和 100GeV/c 的电子的动量分别是

$$p_\alpha = 4 \times 10^6\ \text{eV}/c = 2.1 \times 10^{-16}\text{g} \cdot \text{cm} \cdot \text{s}^{-1},$$

$$p_e = 10^{10}\text{eV}/c = 5.3 \times 10^{-13}\text{g} \cdot \text{cm} \cdot \text{s}^{-1}.$$

粗略估计，Δp_α 和 Δp_e 分别同 p_α 和 p_e 有相同量级，根据不确定关系，这两种情况下能探测到的空间范围的尺度分别是

$$\Delta x_\alpha \sim \frac{\hbar}{\Delta p_\alpha} \sim \frac{1.05 \times 10^{-27}}{4.2 \times 10^{-16}} \sim 2 \times 10^{-12}\text{cm},$$

$$\Delta x_e \sim \frac{\hbar}{\Delta p_e} \sim \frac{1.05 \times 10^{-27}}{1.06 \times 10^{-12}} \sim 10^{-15}\text{cm}.$$

【例 4】 粒子在宽度为 a 的一维无限深势阱中运动，试用不确定关系估计粒子的基态能.

解 由于粒子是关闭在宽度为 a 的势阱中，粒子坐标的不确定度不会超过 a，即 $\Delta x \leqslant a$. 按不确定关系(4.18)式，动量的不确定度 Δp 为 $\Delta p \geqslant \hbar/2a$. 个别测量中得到的粒子动量不会小于其不确定度，故得 $p \geqslant \hbar/(2a)$，进一步有

$$E = p^2/(2m) \geqslant \hbar^2/(8ma^2).$$

由此可知，粒子的基态能有 $\hbar^2/(8ma^2)$ 的量级，这和 §5.2.2 例 4 中通过精确计算得到的基态能 $\pi^2\hbar^2/(2ma^2)$ 在数量级上一致.

习 题

1. 证明角动量算符 $\hat{L} = -\mathrm{i}\hbar r \times \nabla$ 和哈密顿算符 $\hat{H} = \hat{p}^2/(2m) + U(r)$ 是厄米算符.

2. 若 $\hat{F}$ 不是厄米的，证明 $\hat{F} + \hat{F}^\dagger$ 和 $\mathrm{i}(\hat{F} - \hat{F}^\dagger)$ 是厄米的，并且 $\hat{F}$ 可写成

$$\hat{F} = (\hat{F} + \hat{F}^\dagger)/2 + \mathrm{i}(\hat{F} - \hat{F}^\dagger)/2\mathrm{i}$$

3. 利用在宽为 a 的势箱中形成驻波的条件——势箱宽度等于半波长的整数倍，求势箱中粒子的能谱.

4. 玻尔原子模型假定，氢原子中的电子在绕核的圆形轨道上运动. 将电子看成在这圆形轨道上传播的波. 利用波的稳定性条件(即驻波条件)求氢原子的基态能量.

5. 证明在力学量 F 有连续谱的情况下，在状态 $\Psi(r,t)$ 中测 F 得值 f 的概率密度 ρ_f 的公式—— §5.2.2(2.62)、(2.63)、(2.64)式.

6. 一维运动粒子的状态是

$$\psi(x) = \begin{cases} Axe^{-\lambda x}, & x \geqslant 0, \\ 0 & x < 0. \end{cases}$$

其中 $\lambda > 0$，求

(a)粒子动量的概率分布函数；

(b)粒子的平均动量.

7. 求绕固定轴转动的刚性转子的能量本征值和本征函数.

提示 平面转子的能量算符为 $\hat{H} = \hat{L}_z^2/(2I) = -[\hbar^2/(2I)](\partial^2/\partial\varphi^2)$，其中 I 为绕 z 轴的转动惯量.

8. 粒子在一维无限深势箱中运动，状态波函数为 $\psi(x) = A\sin(\pi x/a)\cos^2(\pi x/a)$，求粒子能量的可能值和相应的概率.

9. 若 $\hat{f}$ 的本征值 f 是 4 重简并的，四个独立的本征函数为 $\varphi_1, \varphi_2, \varphi_3, \varphi_4$，它们不相互正交. 试求对应这个本征值的一组四个相互正交的本征函数.

提示 这一问题的一般解可以通过 §5.2.2(2.47)式求得. 但实际上可以用更简便的方法.

先规定一个待求的函数为 $\phi_1 = \varphi_1$，再将 ϕ_1 和 φ_2 线性叠加成 ϕ_2，通过要求 ϕ_2 和 ϕ_1 正交来确定叠加系数. 以此类推.

10. 若 $\hat{F}$ 和 $\hat{G}$ 可对易，在它们的本征值有简并的情况下，证明它们有共同的本征函数系.

11. 证明角动量算符 $\hat{\boldsymbol{L}}$ 的各个笛卡儿分量之间，以及它们和 $\hat{\boldsymbol{L}}^2$ 之间的对易关系——§5.2.3(3.23)、(3.24)式.

12. 证明 $\psi(r,\theta,\varphi) = f(r)\sin^3\theta e^{im\varphi}$ 是 $\hat{\boldsymbol{L}}^2$ 和 $\hat{L}_z$ 的共同本征函数，并求相应的本征值，证明在此状态 $\hat{L}_x$，$\hat{L}_y$ 没有确定值.

13. 已知 $\hat{K} = \hat{A}\hat{B}$，且 $\hat{A}\hat{B} - \hat{B}\hat{A} = 1$. 证明：若 φ 为 $\hat{K}$ 的本征函数，对应的本征值为 λ，则 $u = \hat{A}\varphi$ 也是 $\hat{K}$ 的本征函数，对应的本征值为 $\lambda - 1$，而且 $v = \hat{B}\varphi$ 也是 $\hat{K}$ 的本征函数，对应的本征值为 $\lambda + 1$.

14. 用不确定关系估计氢原子的基态能.

第三章　运动方程和守恒律

在前两章里，分别讨论了微观粒子状态的描述和力学量的描述. 这一章讨论状态随时间的演化，建立量子力学的运动方程.

正像对于经典力学一样，守恒律对于量子力学也有十分重要的意义. 在这一章里，在建立了量子力学的运动方程以后，将接着讨论各种守恒律，并分析它们和对称性的关系.

§5.3.1　薛定谔方程

(一)薛定谔方程

微观粒子所服从的运动方程是量子力学的又一基本原理. 由于微观粒子的状态由波函数完全地描述，所以，按照因果律，运动方程应该回答的问题是：如何由系统初始时刻的波函数决定以后任一时刻的波函数.

作为描述状态随时间变化的方程，它自然应包含有对波函数 Ψ 的时间导数，而且应该是一阶的，否则仅由初始时刻的波函数不足以完全确定以后时刻的波函数. 另外，根据叠加原理，运动方程的任意两个解的线性组合也是方程的解. 这就意味着，运动方程必须是线性微分方程. 根据上面的分析，量子力学的运动方程可以一般地写成：

$$\frac{\partial \Psi}{\partial t} = \hat{L}\Psi, \tag{1.1}$$

其中 $\hat{L} = \hat{L}(r,t)$ 是一个作用在坐标 r 上的线性算符，在一般情况下可能依赖于时间 t.

为了知道 $\hat{L}$ 是怎样的算符，我们先讨论粒子作自由运动这一简单情况，以便从中得到关于 $\hat{L}$ 的形式的启示. 已经知道，粒子的自由运动状态用德布罗意平面波描述，即

$$\Psi(r,t) = Ae^{(i/\hbar)(p\cdot r - Et)}. \tag{1.2}$$

在此状态中，粒子的能量 $E = p^2/(2m)$. 由(1.2)式得到

$$-i\hbar \nabla \Psi = p\Psi,$$

$$\frac{\partial \Psi}{\partial t} = -\frac{i}{\hbar}E\Psi = -\frac{i}{\hbar}\frac{p^2}{2m}\Psi = \frac{i\hbar}{2m}\nabla^2\Psi.$$

这样就可以看出，对于自由粒子，(1.1)式中的算符 $\hat{L} = i\hbar \nabla^2/(2m)$. 为了进一步推广到在势能场中运动的一般情况，让我们注意到，自由粒子的哈密顿算符是

$\hat{H} = -\hbar^2 \nabla^2/(2m)$，因此可以将 $\hat{L}$ 写成 $\hat{H}/(\mathrm{i}\hbar)$．于是(1.1)式成为：

$$\mathrm{i}\hbar\partial\Psi/\partial t = \hat{H}\Psi. \tag{1.3}$$

这就是推广到一般情况下决定 Ψ 随时间演化的运动方程，称为薛定谔方程，其中 $\hat{H}$ 是系统的哈密顿算符．

例如，对于一个在保守场中运动的粒子

$$\hat{H} = -[\hbar^2/(2m)]\nabla^2 + U(r), \tag{1.4}$$

$U(r)$ 是粒子在保守场中的势能．在这种情况下，$\hat{H}$ 不依赖时间．对于在非保守场中运动的粒子，

$$\hat{H} = [-\hbar^2/(2m)]\nabla^2 + U(r,t), \tag{1.4'}$$

这时 $\hat{H}$ 显含时间．在这两种情况下，薛定谔方程都是

$$\mathrm{i}\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi + U\Psi. \tag{1.5}$$

薛定谔方程中的哈密顿算符是非相对论的，因此这一方程只适合于描述非相对论性的粒子．薛定谔方程有一个明显的特点，即含有虚因子 i，这使得波函数一般是实变量的复函数．

以上只是以一种看来比较易于接受的方式建立了薛定谔方程，而不能认为是对它的推导．象任何一种基本运动形态的方程一样，薛定谔方程是不能推导的，它的正确性，从理论上在于它和波函数 $\Psi(r,t)$ 的物理意义（$|\Psi|^2$ 是位置概率密度）相自洽，即：能保证粒子数守恒——概率守恒，见以下第（二）小节及 §5.7.1；从实践上，在于由它得到的结论能被实验检验．

薛定谔方程在形式上也可以这样得到，即在经典力学的能量表示式[例如 $E = p^2/(2m) + U(r)$]中作如下的替换：

$$\left.\begin{aligned} &E \to \mathrm{i}\hbar\partial/\partial t, \\ &p \to \hat{p} = -\mathrm{i}\hbar\nabla, \qquad r \to \hat{r} = r. \end{aligned}\right\} \tag{1.6}$$

再将等式两边作用到波函数 Ψ 上．

（二）粒子数守恒与概率流密度

薛定谔方程决定了波函数随时间的演化，但波函数的归一化，§5.1.2(2.12)式

$$\int_{\text{全空间}} \Psi^*(r,t)\Psi(r,t)\mathrm{d}^3r = 1 \tag{1.7}$$

却应该保持不变，其中的积分在所研究的系统的整个空间中进行．(1.7)式的意义是：系统中的粒子数始终是 1——粒子数守恒．我们来证明，薛定谔方程(1.4)式的确满足粒子数守恒的要求．

将体积元 d^3r 中找到粒子的概率 $\mathrm{d}P \approx |\Psi|^2\mathrm{d}^3r$ 用体积元 d^3r 除，得到在 r 处

单位体积内找到粒子的概率

$$\rho(\boldsymbol{r},t) = |\Psi(\boldsymbol{r},t)|^2 = \Psi^*(\boldsymbol{r},t)\Psi(\boldsymbol{r},t). \tag{1.8}$$

$\rho(\boldsymbol{r},t)$ 是粒子位于 $\boldsymbol{r}$ 处的概率密度. (1.7)式写成

$$\int_{全空间}\rho(\boldsymbol{r},t)\mathrm{d}^3\boldsymbol{r} = 1, \tag{1.9}$$

这表示概率守恒. 因为 $\rho(\boldsymbol{r},t) \geqslant 0$ [见(1.8)式],所以由(1.9)式知道,如果某些地方概率密度增加,必有另外一些地方概率密度减少,使总的概率守恒.

由于上述原因,可以形象化地把微观粒子的位置概率看成是按密度(1.8)式分布在空间中的一种"流体",它可以在空间中流动,但总量不能增加也不能减少.

用 $\boldsymbol{j}$ 表示概率流密度,它的方向指向概率流动的方向,而大小等于单位时间通过垂直于 $\boldsymbol{j}$ 的单位面积流过去的概率. 概率守恒(即粒子数守恒)的数学表述是

$$\oint_S \boldsymbol{j}\cdot\mathrm{d}\boldsymbol{\sigma} = -\frac{\mathrm{d}}{\mathrm{d}t}\int_V \rho\cdot\mathrm{d}\tau = -\int_V \frac{\partial\rho}{\partial t}\mathrm{d}\tau. \tag{1.10}$$

其中 S 是包围体积 V 的封闭面. 利用高斯定理可以将上式化成微分形式

$$\mathrm{div}\,\boldsymbol{j} = -\partial\rho/\partial t. \tag{1.11}$$

这和电荷守恒律的微分形式,§2.4.1(1.16)式,在形式上相同.

为了证明确实存在满足粒子数守恒要求[满足(1.11)式]的概率流密度,并找到它的表达式,取薛定谔方程(1.5)式的复共轭:

$$-\mathrm{i}\hbar\frac{\partial\Psi^*}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi^* + U\Psi^*, \tag{1.12}$$

用 Ψ^* 乘(1.5)式,用 Ψ 乘(1.12)式,相减得到

$$\mathrm{i}\hbar(\Psi^*\partial\Psi/\partial t + \Psi\partial\Psi^*/\partial t) = -[\hbar^2/(2m)](\Psi^*\nabla^2\Psi - \Psi\nabla^2\Psi^*).$$

它可以写成

$$\frac{\partial(\Psi^*\Psi)}{\partial t} = \frac{\mathrm{i}\hbar}{2m}\nabla\cdot(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*). \tag{1.13}$$

将它和(1.11)式对比,并注意到(1.8)式就看到,矢量

$$\boldsymbol{j} = -\frac{\mathrm{i}\hbar}{2m}(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*). \tag{1.14}$$

就是能满足粒子数守恒[(1.11)式]要求的概率流密度.

将(1.10)式中的积分体积扩大到整个空间,并注意到在无穷远的界面上 Ψ 趋于零,得到

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\infty \rho\mathrm{d}\tau = \frac{\mathrm{d}}{\mathrm{d}t}\int_\infty \Psi^*\Psi\mathrm{d}\tau = 0,$$

即:在整个空间找到粒子的概率与时间无关. 这就意味着粒子数守恒. 同时它也证明了波函数的归一化不随时间变化.

由于薛定谔方程是非相对论的运动方程,因而由它导出的粒子数守恒只适用于非相对论的量子力学.

【例 1】 求德布罗意波的概率流密度.

解 将德布罗意波函数 $\Psi(\boldsymbol{r},t) = A\mathrm{e}^{(\mathrm{i}/\hbar)(\boldsymbol{p}\cdot\boldsymbol{r}-Et)}$ 代入(1.14)式得到

$$\boldsymbol{j} = |A|^2\boldsymbol{p}/m = |A|^2\boldsymbol{v}. \tag{1.15}$$

在本例所讨论的状态(德布罗意波)中,动量 $\boldsymbol{p}$ 有确定值,$\Delta\boldsymbol{p} = 0$,因而速度 $\boldsymbol{v} = \boldsymbol{p}/m$ 也有确定值,$\Delta\boldsymbol{v} = 0$.根据不确定关系,$\Delta\boldsymbol{r} = \infty$,坐标 $\boldsymbol{r}$ 的值完全不确定.因此,不能象经典力学那样简单地将速度 $\boldsymbol{v}$ 看成坐标 $\boldsymbol{r}$ 的测量值的时间导数.在下一节中,将仔细讨论力学量的时间导数问题.

【例 2】 求球面波

$$\Psi(r,t) = Af(\theta,\varphi)\mathrm{e}^{\mathrm{i}kr}/r \tag{1.16}$$

在 $r\to\infty$ 时的概率流密度,以及通过半径为 r 的球面面积元 $S = r^2\mathrm{d}\Omega$ 的概率流.

解 球坐标中的梯度算符见附录(Ⅱ.43)式

$$\nabla = \boldsymbol{e}_r\partial/\partial r + \boldsymbol{e}_\theta(1/r)\partial/\partial\theta + \boldsymbol{e}_\varphi[1/(r\sin\theta)]\partial/\partial\varphi.$$

和(1.16)式一道代入(1.14)式,略去 r 大时 r^{-1} 的高级小量,得到

$$\boldsymbol{j} = \boldsymbol{e}_r|A|^2v|f(\theta,\varphi)|^2/r^2.$$

通过半径为 r 的球面面积元 $S = r^2\mathrm{d}\Omega$ 的概率流是

$$j_{\mathrm{d}\Omega} = |A|^2v|f(\theta,\varphi)|^2\mathrm{d}\Omega. \tag{1.17}$$

这表明,在 r 大时球面波(1.16)通过立体角元 $\mathrm{d}\Omega$ 的概率流只依赖于 θ,φ 而与 r 无关.

(三)解薛定谔方程的分离变量法 定态

当哈密顿算符与时间无关时,薛定谔方程存在具有重要意义的解.此时,薛定谔方程是

$$\mathrm{i}\hbar\partial\Psi/\partial t = -[\hbar^2/(2m)]\nabla^2\Psi + U(\boldsymbol{r})\Psi. \tag{1.18}$$

在这一方程中,作用在空间变量上的算符和作用在时间变量上的算符是独立的,因而可以用分离变量法求解.令

$$\Psi(\boldsymbol{r},t) = \psi(\boldsymbol{r})f(t).$$

代入上式,并用 $\psi(\boldsymbol{r})f(t)$ 除等式两边可得

$$\frac{\mathrm{i}\hbar}{f(t)}\frac{\mathrm{d}f(t)}{\mathrm{d}t} = \frac{1}{\psi(\boldsymbol{r})}\left[-\frac{\hbar^2}{2m}\nabla^2\psi(\boldsymbol{r}) + U(\boldsymbol{r})\psi(\boldsymbol{r})\right].$$

此式的左边只依赖于变量 t,右边只依赖于变量 $\boldsymbol{r}$,两边要相等,只能都等于一个共同的常数 E

$$\begin{aligned} &\mathrm{i}\hbar\mathrm{d}f(t)/\mathrm{d}t = Ef(t), \\ &-[\hbar^2/(2m)]\nabla^2\psi(\boldsymbol{r}) + U(\boldsymbol{r})\psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}). \end{aligned} \tag{1.19}$$

第一个方程的解显然是：

$$f(t) = ce^{-(i/\hbar)Et}. \tag{1.20}$$

将方程(1.19)及(1.4)式和§5.2.2(2.2)式对比，可见(1.19)式就是能量算符的本征值方程，因而 E 就是能量，$\psi(\boldsymbol{r})$ 就是能量的本征函数，用 $\psi_E(\boldsymbol{r})$ 表示. 这样，薛定谔方程的解可以写成

$$\Psi_E(\boldsymbol{r},t) = \psi_E(\boldsymbol{r})e^{-(i/\hbar)Et}. \tag{1.21}$$

它描述了能量有确定值的状态. 这种状态称为定态. 波函数(1.21)称为定态波函数，方程(1.19)称为定态薛定谔方程.

定态是粒子在保守场中运动时很常见的重要状态. 这种状态，除能量有确定值外，还有其他一些特点. 由(1.21)式可以看出，定态波函数对时间有谐和的依赖关系，这一依赖关系仅由定态的能量决定. 将(1.21)式代入(1.8)式和(1.14)式可见，时间因子消去，因而在定态，粒子的位置概率密度和概率流密度都不随时间改变；将(1.21)式代入§5.2.2(2.58)、(2.60)式和§5.2.1(1.13)式，时间因子也消去，因而在定态，力学量的概率分布和平均值也与时间无关. 总之，定态是一种力学性质稳定的状态.

由于定态是哈密顿算符的本征态，所以标志定态的力学量完全组中的各个算符必须全都和哈密顿算符 $\hat{H}$ 可对易.

由于薛定谔方程(1.18)式是一个线性微分方程，所以它的一般解应是定态解的线性组合. 例如，对于能量取分立谱的情况有

$$\Psi(\boldsymbol{r},t) = \sum_n c_n\psi_n(\boldsymbol{r})e^{-(i/\hbar)E_nt}. \tag{1.22}$$

这种一般解当然不再是能量有确定值的定态.

展开式的系数 c_n 可以通过初始条件

$$\Psi(\boldsymbol{r},0) = \psi(\boldsymbol{r}) \tag{1.23}$$

确定. 将(1.22)式代入(1.23)式得到

$$\psi(\boldsymbol{r}) = \sum_n c_n\psi_n(\boldsymbol{r}).$$

由此得

$$c_n = \int\psi_n^*(\boldsymbol{r})\psi(\boldsymbol{r})\mathrm{d}^3\boldsymbol{r}. \tag{1.24}$$

(1.22)和(1.24)式回答了本节一开始提出的问题：如何由系统初始时刻的波函数 $\Psi(\boldsymbol{r},0) = \psi(\boldsymbol{r})$ 决定以后任一时刻的波函数 $\Psi(\boldsymbol{r},t)$. 我们看到，回答这一问题的关键一步是求系统的定态波函数 $\psi_n(\boldsymbol{r})$ 和能谱 E_n. 这可以归结为求解定态薛定谔方程. 下面举一个例子. 另一个有重要意义的例子见本章习题3.

【例3】　转轴固定并限定在平面上作圆周运动的刚性转子称为平面转子，参

看第二章习题7.设其转动惯量为 I.在转子上作一根垂直于转轴的直线,它和垂直于转轴的固定坐标轴的夹角为 φ.若在 $t=0$ 时,平面转子的状态用波函数 $\psi=A\sin^2\varphi$ 描写,问它的状态如何随时间变化.

解 这归结为求满足初始条件的薛定谔方程的解.以下首先求方程的一般解,再用初始条件定解.

平面转子的定态问题已经在第二章习题7中解过.现复述如下.根据§1.5.1(1.35)、(1.37)式和§5.2.2(2.20)式,平面转子的哈密顿算符为 $\hat{H}=\hat{L}_z^2/2I=-(\hbar^2/2I)(\partial^2/\partial\varphi^2)$.按照前面的讨论,首先求定态薛定谔方程

$$[-\hbar^2/(2I)](\mathrm{d}^2/\mathrm{d}\varphi^2)\psi_m(\varphi)=E_m\psi_m(\varphi) \tag{1.25}$$

的解,再乘以时间因子得定态波函数:$\Psi_m(\varphi,t)=\psi_m(\varphi)\mathrm{e}^{-(\mathrm{i}/\hbar)E_m t}$.注意到波函数的单值条件

$$\psi_m(\varphi)=\psi_m(\varphi+2\pi), \tag{1.26}$$

不难得到定态方程(1.25)在 $(0,2\pi)$ 区间中归一化的解为

$$\psi_m(\varphi)=\sqrt{1/2\pi}\mathrm{e}^{\mathrm{i}m\varphi},\qquad m=0,\pm1,\pm2,\cdots \tag{1.27}$$

相应的定态能量为

$$E_m=m^2\hbar^2/(2I). \tag{1.28}$$

所求的薛定谔方程的一般解为

$$\Psi(\varphi,t)=\sum_m c_m\Psi_m(\varphi,t)=\sum_m c_m\psi_m(\varphi)\mathrm{e}^{-(\mathrm{i}/\hbar)E_m t}. \tag{1.29}$$

展开式系数 c_m 由初始条件决定

$$\Psi(\varphi,0)=\sum_m c_m\sqrt{1/2\pi}\mathrm{e}^{\mathrm{i}m\varphi}=A\sin^2\varphi. \tag{1.30}$$

和(1.27)式一道代入(1.24)式可以得到 c_m.在本例中也可以如下求出.注意到

$$A\sin^2\varphi=A\frac{1-\cos(2\varphi)}{2}=\frac{A}{2}-A\frac{\mathrm{e}^{\mathrm{i}2\varphi}+\mathrm{e}^{-\mathrm{i}2\varphi}}{4},$$

和(1.30)式比较可以定出系数 c_m.代入一般解(1.29),就得到满足初始条件的任意时刻的波函数:

$$\Psi(\varphi,t)=\frac{A}{2}-\frac{A}{4}(\mathrm{e}^{\mathrm{i}2\varphi}+\mathrm{e}^{-\mathrm{i}2\varphi})\mathrm{e}^{-(\mathrm{i}/\hbar)E_2 t}=\frac{A}{2}[1-\cos(2\varphi)\cdot\mathrm{e}^{-\mathrm{i}2\hbar t/I}].$$

再由归一化条件可求得 $A=2/\sqrt{3\pi}$.

(四)一维方形势

一维势能 $U(x)$ 随 x 的变化一般是连续函数.作为一种粗略的近似,我们把这种变化简化为在一些分离的点上发生.在这些点上 $U(x)$ 有不连续的跳跃;而在被这些点隔开的不同区域里,$U(x)$ 分别等于不同的常数.这样简化后的 $U(x)$ 称为一维方形势.这种一维方形势中的运动很容易求解,而由此得到的一些定性结论有

普遍意义.

先来看两个典型例子.

【例 4】　一维对称方形势阱. 势能是

$$U(x) = \begin{cases} 0, & \text{当 } |x| > a, \\ -U_0, & \text{当 } |x| < a, \end{cases} \tag{1.31}$$

如图 1. 试分析能量 E 满足 $-U_0 < E < 0$ 的粒子在这一势能场中的运动.

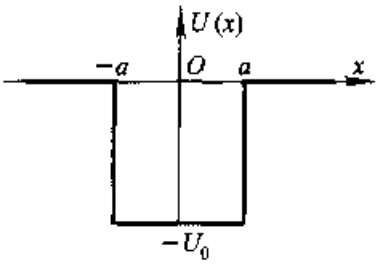


图 1　一维方势阱

解　定态方程 $-[\hbar^2/(2m)]\psi'' + U(x)\psi = E\psi$ 在图 1 所示的三个区域里分别成为

$$\left.\begin{aligned} &\psi''_{\mathrm{I}} + (2mE/\hbar^2)\psi_{\mathrm{I}} = 0, && x < -a, \\ &\psi''_{\mathrm{II}} + (2m/\hbar^2)(E+U_0)\psi_{\mathrm{II}} = 0, && |x| < a, \\ &\psi''_{\mathrm{III}} + (2mE/\hbar^2)\psi_{\mathrm{III}} = 0, && x > a. \end{aligned}\right\} \tag{1.32}$$

在 $U_0 < E < 0$ 情况下,方程的解可写为

$$\begin{cases} \psi_{\mathrm{I}} = Ae^{\lambda x} + A'e^{-\lambda x}, & (1.33) \\ \psi_{\mathrm{II}} = B\cos kx + B'\sin kx, & (1.34) \\ \psi_{\mathrm{III}} = Ce^{\lambda x} + C'e^{-\lambda x}. & (1.35) \end{cases}$$

$$\lambda = \sqrt{-2mE/\hbar^2} = \sqrt{2m|E|/\hbar^2}, \qquad k = \sqrt{2m(E+U_0)/\hbar^2}. \tag{1.36}$$

根据波函数的标准条件,为保证 $x \to \infty$ 时,ψ_{III} 有限,要求 $C = 0$;为保证 $x \to -\infty$ 时,ψ_{I} 有限,要求 $A' = 0$. 这样有 $\psi_{\mathrm{I}} = Ae^{\lambda x}$,$\psi_{\mathrm{III}} = C'e^{-\lambda x}$,$\psi_{\mathrm{II}}$ 仍由(1.34)式决定. 不难看到,当 $x \to \pm\infty$ 时,$\psi \to 0$,表明粒子束缚在有限空间内运动. 这种解是束缚态解.

可以进一步利用波函数的标准条件确定解的系数. 在 $x = \pm a$ 处,波函数及其一阶导数的连续性条件是:

$$\left.\begin{aligned} x = -a: \quad & Ae^{-\lambda a} = B\cos ka - B'\sin ka, \\ & \lambda Ae^{-\lambda a} = kB\sin ka + kB'\cos ka; \\ x = a: \quad & C'e^{-\lambda a} = B\cos ka + B'\sin ka, \\ & -\lambda C'e^{-\lambda a} = -kB\sin ka + kB'\cos ka \end{aligned}\right\} \tag{1.37}$$

这是待定系数 A, B, B', C' 的线性齐次方程组,其中的变量数 4 等于方程数 4. 它有非零解的条件是系数行列式为零,即:

$$\begin{vmatrix} e^{-\lambda a} & -\cos ka & \sin ka & 0 \\ \lambda e^{-\lambda a} & -k\sin ka & -k\cos ka & 0 \\ 0 & \cos ka & \sin ka & -e^{-\lambda a} \\ 0 & -k\sin ka & k\cos ka & \lambda e^{-\lambda a} \end{vmatrix} = 0. \tag{1.38}$$

应用关于行列式的拉普拉斯展开定理，以上行列式可分解为

$$\begin{vmatrix} e^{-\lambda a} & -\cos ka \\ \lambda e^{-\lambda a} & -k\sin ka \end{vmatrix} \cdot \begin{vmatrix} \sin ka & -e^{-\lambda a} \\ k\cos ka & \lambda e^{-\lambda a} \end{vmatrix} - \begin{vmatrix} e^{-\lambda a} & \sin ka \\ \lambda e^{-\lambda a} & -k\cos ka \end{vmatrix} \cdot \begin{vmatrix} \cos ka & -e^{-\lambda a} \\ -k\sin ka & \lambda e^{-\lambda a} \end{vmatrix}$$

$$= 2e^{-\lambda a}(k\cos ka + \lambda \sin ka)(\lambda \cos ka - k\sin ka). \tag{1.39}$$

但是 $e^{-\lambda a} \neq 0$，因而必有

$$k\tan ka = \lambda \quad 或 \quad k\cot ka = -\lambda. \tag{1.40}$$

这就是方程(1.38)有非零解的条件，它决定了方形势阱中的束缚态能量只能取分立值. 这是因为按(1.36)式，k 和 λ 是与粒子能量有关的，因而(1.40)式表明，能量只能取某些特定值. 微观粒子处于束缚态时，能量只能取分立值，称为能量量子化. 从以上例子看到，在量子力学中，能量的量子化是通过标准化的波函数满足外部约束条件而自然地得到. 本来，波函数的标准化条件只要求 ψ 有限，但由于势阱的约束，使得 $x \to \pm \infty$ 时 $\psi \to 0$. 正是这一约束使能量只能取分立值. (参看 $U_0 = \infty$ 的无限深势阱 §5.2.2 例 4.)

由(1.40)式决定的能量所能取的分立值可以用图解法求出. 令 $\xi = ka$，$\eta = \lambda a$，则(1.40)式成为

$$\eta = \xi\tan \xi \quad 或 \quad \eta = -\xi\cot \xi. \tag{1.41}$$

而按(1.36)式又有

$$\xi^2 + \eta^2 = (2m/\hbar^2)U_0 a^2. \tag{1.42}$$

对于一个确定的势阱，U_0 和 a 取确定的值，(1.42)式是一个有确定半径的圆. 这个圆和(1.41)式中的两组曲线的交点的坐标值，确定了粒子的能量值，交点的数目就是能级的个数，见图 2. 由图可以清楚地看出能量取值的不连续性.

将有非零解的条件(1.40)式代回方程(1.37)式，得到相应的解为：

$$\left.\begin{aligned} &当\ k\tan ka = \lambda\ 时，\quad B' = 0，\quad C' = A = Be^{-\lambda a}\cos ka，\\ &当\ k\cot ka = -\lambda\ 时，\quad B = 0，\quad C' = -A = B'e^{-\lambda a}\sin ka. \end{aligned}\right\} \tag{1.43}$$

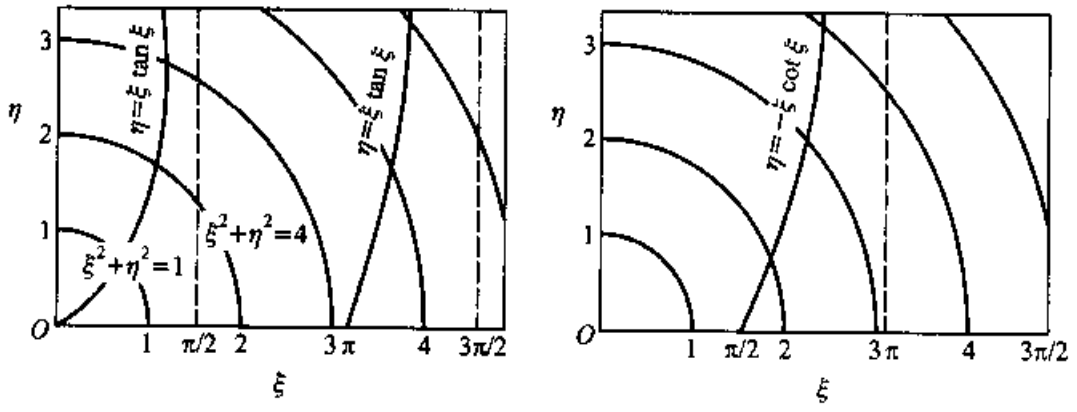


图 2　决定势阱中分立能级的图解法

从而求出相应的束缚态波函数为

$$\psi = \begin{cases} Ae^{\lambda x}, & x \leqslant -a \\ (Ae^{-\lambda a}\sec ka)\cos kx, & -a \leqslant x \leqslant a \\ Ae^{-\lambda x}. & a \leqslant x \end{cases} \tag{1.44}$$

或

$$\psi = \begin{cases} Ae^{\lambda x}, & x \leqslant -a \\ -(Ae^{-\lambda a}\csc ka)\sin kx, & -a \leqslant x \leqslant a \\ -Ae^{-\lambda x}. & a \leqslant x \end{cases} \tag{1.45}$$

它们分别和(1.40)式中的两个方程所决定的能级(图2中的两组交点)相对应.注意,(1.44)式[以及(1.45)式]中的波函数在三个区域中有不同表达式,但这并不是三个函数,而是同一个函数的三个部分,它们共同组成一个统一的波函数.

【例5】　势垒的穿透.一维方形势垒的势能是

$$U(x) = \begin{cases} 0, & \text{当} \mid x \mid > a, \\ U_0, & \text{当} \mid x \mid < a, \end{cases} \tag{1.46}$$

如图3,分析 $0 < E < U_0$ 的粒子在这一势能场中的运动.

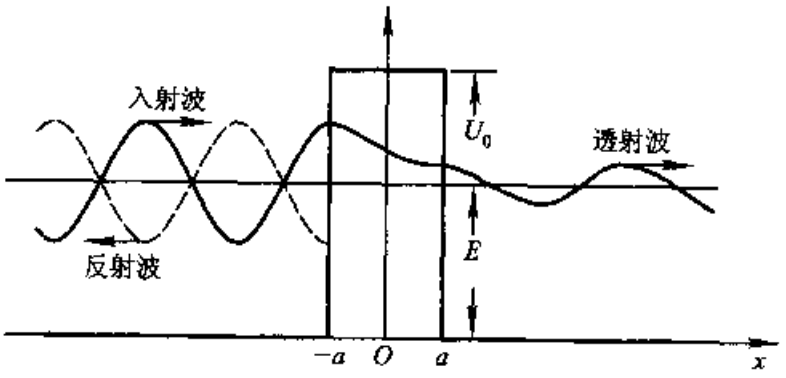


图3　一维方形势垒

解　运动方程和(1.32)式一样,只要将 U_0 前的正号改为负号.

由于 $0 < E < U_0$,势垒高度超过总能,区域Ⅱ是 $E < U$ 的区域,经典粒子不能进入这种区域,因而从左边射来的粒子,会全部被势垒挡住,不能到达势垒右边.然而微观粒子,却可以穿透势垒,到达势垒的右边.这是三维散射问题(见第六篇第三章)在一维情况下的简化.下面来求解这一问题.

定义 k 和 λ 为

$$k = \sqrt{2mE/\hbar^2}, \qquad \lambda = \sqrt{2m(U_0 - E)/\hbar^2}. \tag{1.47}$$

注意到 $0 < E < U_0$,可见 k 和 λ 都是实数.方程的解为

$$\begin{cases} \psi_{\text{Ⅰ}} = Ae^{ikx} + A'e^{-ikx}, \\ \psi_{\text{Ⅱ}} = Be^{-\lambda x} + B'e^{\lambda x}, \\ \psi_{\text{Ⅲ}} = Ce^{ikx} + C'e^{-ikx}. \end{cases} \tag{1.48}$$

和§5.2.2(2.6)式比较可见，$\psi_{\text{Ⅰ}}$ 和 $\psi_{\text{Ⅲ}}$ 右边的第一项是以动量 $p=\hbar k$ 向右行进的波，第二项是以动量 $p=-\hbar k$ 向左行进的波. 假定入射粒子从左边飞来，在势垒的作用下一部分被反射，飞到 $x=-\infty$；另一部分穿透过去，飞到 $x=+\infty$. 这样，在 $x>a$ 的区域里，只有向右行进的波而没有向左行进的波：

$$x\to+\infty,\qquad \psi(x)\sim e^{ikx} \tag{1.49}$$

这就是一维"散射"问题的边界条件.

利用边界条件(1.49)可见，(1.48)式中的 $C'=0$. 在 $\psi_{\text{Ⅲ}}$ 中只剩下 Ce^{ikx}，它代表透射波，其概率流密度是

$$j_{\text{透}}=|C|^2\hbar k/m. \tag{1.50}$$

$\psi_{\text{Ⅰ}}$ 中的两项分别代表入射波和反射波，概率流密度分别是

$$j_{\text{入}}=|A|^2\hbar k/m,\qquad j_{\text{反}}=-|A'|^2\hbar k/m. \tag{1.51}$$

穿透系数 D 和反射系数 R 分别是

$$\begin{aligned} D&=|j_{\text{透}}|/|j_{\text{入}}|=|C|^2/|A|^2,\\ R&=|j_{\text{反}}|/|j_{\text{入}}|=|A'|^2/|A|^2. \end{aligned} \tag{1.52}$$

因此，问题归结为，利用 $x=\pm a$ 的连续性条件决定 A'、C 和 A 的关系.

将 A 当作已知，则连续性条件成为四个未知数 B、C、A'、B' 的四个非齐次联立方程组(包含 A 的项是"非齐次项"). 这样的方程组有唯一解，从而定出 B、C、A'、B' 和 A 的函数关系，而按(1.52)式就决定了穿透系数 D 和反射系数 R. 在这一问题中，能量 E 的取值不受限制，因而能谱连续.

$x=-a$ 处的连续性条件 $\psi_{\text{Ⅰ}}(-a)=\psi_{\text{Ⅱ}}(-a)$，$\psi'_{\text{Ⅰ}}(-a)=\psi'_{\text{Ⅱ}}(-a)$ 是

$$\left.\begin{aligned} Be^{\lambda a}+B'e^{-\lambda a}&=Ae^{-ika}+A'e^{ika},\\ -\lambda(Be^{\lambda a}-B'e^{-\lambda a})&=ik(Ae^{-ika}-A'e^{ika}). \end{aligned}\right\} \tag{1.53}$$

$x=a$ 处的连续性条件 $\psi_{\text{Ⅱ}}(a)=\psi_{\text{Ⅲ}}(a)$，$\psi'_{\text{Ⅱ}}(a)=\psi'_{\text{Ⅲ}}(a)$ 是

$$\left.\begin{aligned} Be^{-\lambda a}+B'e^{\lambda a}&=Ce^{ika},\\ -\lambda(Be^{-\lambda a}-B'e^{\lambda a})&=ikCe^{ika}. \end{aligned}\right\} \tag{1.54}$$

将(1.53)的第二式乘以 $-i/k$ 再和第一式相加减，可以解出 A 和 A'：

$$\left.\begin{aligned} A&=\frac{1}{2}e^{ika}\left(\left(1+\frac{i\lambda}{k}\right)Be^{\lambda a}+\left(1-\frac{i\lambda}{k}\right)B'e^{-\lambda a}\right),\\ A'&=\frac{1}{2}e^{-ika}\left(\left(1-\frac{i\lambda}{k}\right)Be^{\lambda a}+\left(1+\frac{i\lambda}{k}\right)B'e^{-\lambda a}\right). \end{aligned}\right\} \tag{1.55}$$

它们可以合写成矩阵形式

$$\begin{pmatrix}A\\A'\end{pmatrix}=\frac{1}{2}\begin{pmatrix}(1+i\lambda/k)e^{(\lambda+ik)a} & (1-i\lambda/k)e^{(ik-\lambda)a}\\ (1-i\lambda/k)e^{(\lambda-ik)a} & (1+i\lambda/k)e^{-(\lambda+ik)a}\end{pmatrix}\cdot\begin{pmatrix}B\\B'\end{pmatrix}. \tag{1.56}$$

类似地由(1.54)式得到

$$\begin{pmatrix} B \\ B' \end{pmatrix} = \frac{1}{2}\begin{pmatrix} (1 + k/\mathrm{i}\lambda)\mathrm{e}^{(\mathrm{i}k+\lambda)a} \\ (1 - k/\mathrm{i}\lambda)\mathrm{e}^{(\mathrm{i}k-\lambda)a} \end{pmatrix} C. \tag{1.57}$$

将(1.57)式代入(1.56)式得到

$$\begin{pmatrix} A \\ A' \end{pmatrix} = \frac{1}{4}\begin{pmatrix} (1 + \mathrm{i}\lambda/k)\mathrm{e}^{(\lambda+\mathrm{i}k)a} & (1 - \mathrm{i}\lambda/k)\mathrm{e}^{(\mathrm{i}k-\lambda)a} \\ (1 - \mathrm{i}\lambda/k)\mathrm{e}^{(\lambda-\mathrm{i}k)a} & (1 + \mathrm{i}\lambda/k)\mathrm{e}^{-(\lambda+\mathrm{i}k)a} \end{pmatrix} \cdot$$
$$\begin{pmatrix} [1 + k/(\mathrm{i}\lambda)]\mathrm{e}^{(\mathrm{i}k+\lambda)a} \\ [1 - k/(\mathrm{i}\lambda)]\mathrm{e}^{(\mathrm{i}k-\lambda)a} \end{pmatrix} C = \begin{pmatrix} (\mathrm{ch}2\lambda a + \mathrm{i}\varepsilon\mathrm{sh}2\lambda a/2)\mathrm{e}^{2\mathrm{i}k'a} \\ -(\mathrm{i}\eta/2)\mathrm{sh}2\lambda a \end{pmatrix} C, \tag{1.58}$$

其中，

$$\varepsilon = \frac{\lambda}{k} - \frac{k}{\lambda}, \qquad \eta = \frac{\lambda}{k} + \frac{k}{\lambda}. \tag{1.59}$$

由此得到

$$\frac{C}{A} = \frac{\mathrm{e}^{-2\mathrm{i}ka}}{\mathrm{ch}\,2\lambda a + \mathrm{i}(\varepsilon/2)\mathrm{sh}\,2\lambda a}, \tag{1.60}$$

$$\frac{A'}{A} = \frac{A'}{C}\frac{C}{A} = \frac{-\mathrm{i}(\eta/2)\mathrm{sh}\,2\lambda a\mathrm{e}^{-2\mathrm{i}ka}}{\mathrm{ch}\,2\lambda a + \mathrm{i}(\varepsilon/2)\mathrm{sh}\,2\lambda a}. \tag{1.61}$$

代入(1.52)式得到透射系数 D 和反射系数 R：

$$D = \left|\frac{C}{A}\right|^2 = \frac{1}{\mathrm{ch}^2\,2\lambda a + (\varepsilon^2/4)\mathrm{sh}^2\,2\lambda a}, \tag{1.62}$$

$$R = \left|\frac{A'}{A}\right|^2 = \frac{(\eta^2/4)\mathrm{sh}^2\,2\lambda a}{\mathrm{ch}^2\,2\lambda a + (\varepsilon^2/4)\mathrm{sh}^2\,2\lambda a}. \tag{1.63}$$

由定义式(1.59)可见，$1 + \varepsilon^2/4 = \eta^2/4$．再利用双曲函数的平方差公式 $\mathrm{ch}^2x - \mathrm{sh}^2x = 1$，容易证明

$$D + R = 1. \tag{1.64}$$

这正是概率守恒所要求的．

在本例中，总能量比势垒的高度低（$E < U_0$），经典粒子会被势垒挡住；微观粒子却有一定的概率穿透势垒．这是波动性的体现．但有时形象地说成是从势垒里打穿了一个“隧道”，因而称它为隧道效应．隧道效应在原子核物理和固体物理中有广泛的应用．

如果势垒高而且厚 $\lambda a \gg 1$，则穿透概率 $D \ll 1$．此时 $\mathrm{ch}(2\lambda a) \approx \mathrm{e}^{2\lambda a}/2$．于是，(1.62)式成为

$$D = (16/\eta^2)\mathrm{e}^{-4\lambda a} = D_0\mathrm{e}^{-8a\sqrt{U_0-E}/\hbar}. \tag{1.65}$$

$D_0 = 16/\eta^2$ 是一个与 a 无关的系数．由此可见，隧道效应的概率随着势垒高度 U_0

和宽度 a 的增加而指数地减小.

如果势垒 $U(x)$ 不是方形的,则可以把它分成许多厚为 Δx 的方形势垒,如图4.每个这样的方形势垒的穿透系数都如(1.65)式,整个势垒的穿透系数是这些穿透系数的连乘积,它可以化为指数上的求和(积分)

$$D = D_0 \exp\left(-\frac{4}{\hbar}\int_a^b \sqrt{2m(U-E)}\,\mathrm{d}x\right). \tag{1.66}$$

以上两个例题虽然简单,却很典型.从中可以得到一些带有普遍意义的结论:

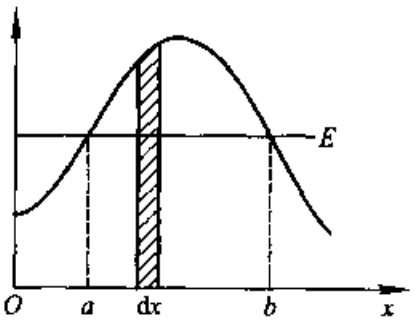


图4 任意形状的势垒

①当粒子被势能场约束在有限空间中运动,形成束缚态时,能谱分立;当粒子在无限空间中运动时,能谱连续.

②如果取势能的最低点为计算能量的起点,则分立谱(束缚态)的最低能量状态,称为基态,的能量不为零,称为零点能.在第二章中已经分析过,零点能大于0是波粒二象性的结果,可以用不确定关系估算.从图2上也可以看到零点能的存在.

③在例4中,$E<0$,在 $|x|>a$ 的区域Ⅰ和Ⅱ,总能 E 小于势能.这样的区域是经典粒子不能进入的区域.但是,微观粒子却能有一定的概率进入到 $E<U$ 的区域,进入后指数衰减.这也是微观粒子具有波动性的表现.为了帮助理解,可以对比在§3.3.3中讨论过的电磁波在导体表面上反射时能透入一定深度,但进入导体后指数衰减[参看§3.3.3(3.10)式].

需要指出,虽然微观粒子能够进入总能 E 小于势能 U 的区域,却不能在全空间里总能都小于势能.如果用 $U_{\min}$ 表示全空间中势能的最小值,则不存在 $E<U_{\min}$ 的状态.为了证明这一论断,首先注意到,经典物理量之间的关系在量子力学中以算符关系的形式成立,而如果对给定的状态 ψ 取这种算符关系的平均值就得到平均值之间的关系.因此,经典力学量之间的关系在量子力学中既以算符的形式成立,也以平均值的形式成立.例如经典力学中 $E=\boldsymbol{p}^2/(2m)+U=T+U$,在量子力学中对应为 $\hat{H}=\hat{T}+\hat{U}$ 和 $\bar{H}=\bar{T}+\bar{U}$,但 $\bar{T}\geqslant 0$,$\bar{U}\geqslant U_{\min}$,故 $\bar{E}\geqslant U_{\min}$.对定态就有 $E\geqslant U_{\min}$.

④在例4中,微观粒子能进入 $E<U$ 的"经典不能达到"区域,但是进入以后指数衰减.在例5情况下,在 $E<U_0$ 的"经典不能达到"区域里,波函数既有指数衰减的分量,也有指数增长的分量,即(1.48)式中的 B 和 B' 都不为零.其原因是,势垒的厚度有限,指数增长项不会使波函数成为无穷大,所以没有理由将它排除.而且必须要同时有衰减项和增长项才能满足边界上的连续性条件,参看图3.

最后，我们来讨论上两例中的另外两个区域，完整地结束这两个例子.

【例 6】　一维方形势垒如(1.46)式. 分析 $E > U_0$ 粒子的运动，参看图 5(a).

解　此时，按(1.47)式定义的 λ 是虚数，因而重新定义

$$k = \sqrt{2mE/\hbar^2}, \qquad k' = \sqrt{2m(E - U_0)/\hbar^2}. \tag{1.67}$$

只要在例 5 的结果中作代换 $\lambda \to \mathrm{i}k'$ 就可以得到本例所要求的结果. 在这一代换下，(1.59)式中的 $\varepsilon \to \mathrm{i}\varepsilon', \eta \to \mathrm{i}\eta'$，

$$\varepsilon' = \frac{k'}{k} + \frac{k}{k'}, \qquad \eta' = \frac{k'}{k} - \frac{k}{k'}. \tag{1.68}$$

同时，$\mathrm{ch}(2\lambda a) \to \cos(2k'a), \mathrm{sh}(2\lambda a) \to \mathrm{i}\sin(2k'a)$. 代入(1.62)、(1.63)式得到透射系数 D 和反射系数 R

$$D = \frac{1}{\cos^2(2k'a) + (\varepsilon'^2/4)\sin^2(2k'a)}, \tag{1.69}$$

$$R = \frac{(\eta'^2/4)\sin^2(2k'a)}{\cos^2(2k'a) + (\varepsilon'^2/4)\sin^2(2k'a)}. \tag{1.70}$$

【例 7】　一维方形势阱如(1.31)式，分析 $E > 0$ 时粒子的运动，参看图 5b.

解　此时，按(1.36)式定义的 λ 是虚数，因而重新定义

$$k = \sqrt{2mE/\hbar^2}, \qquad k' = \sqrt{2m(E + U_0)/\hbar^2}. \tag{1.71}$$

它们和(1.67)式中定义的 k 和 k' 一样，都是实数，因而(1.68)～(1.70)式不需要改变. 特别是，透射系数 D 和反射系数 R 仍然由(1.69)、(1.70)式给出，只不过其中的 k' 按(1.71)式定义.

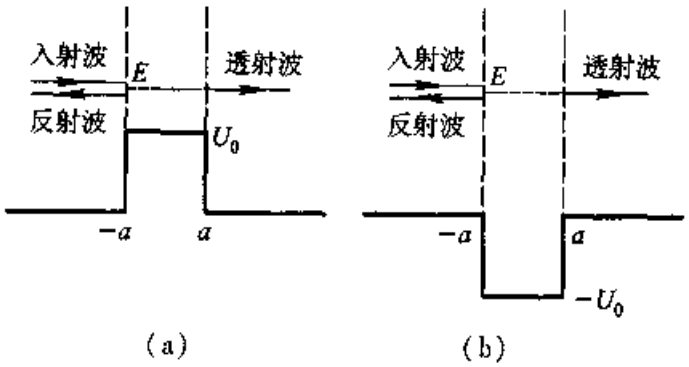


图 5　势垒(a)和势阱(b)的穿透和反射

讨论　这两个例子中，粒子运动的整个区域里都有 $E > U$，见图 5. 如果是经典粒子，会一直向前运动，在经过势垒或势阱时，速度发生变化，但不会受到反射. 在量子情况下则不同，只要势能有变化，尽管总能始终大于势能，也会引起反射.

以例 7 中所讨论的一维方形势阱为例，势能的变化只出现在 $x = -a$ 和 $x = a$ 两点[见图 5(b)]. 粒子(概率波)从左边飞来，在 $x = -a$ 处部分反射部分穿过. 穿过的波到达 $x = a$ 又会部分反射部分穿过. 因此，在最终形成反射波和透射波之前，有一个在 $x = -a$ 到 $x = a$ 的区域中来回反射的过程. 为了看到这种来回反射过程的影响，我们来讨论穿透系数 D 随能量的变化. 由(1.69)式可见，在

$$2k'a = n\pi, \qquad n = 1,2,\cdots \tag{1.72}$$

时，穿透系数 D 达到它的最大可能值 $D = 1$．将(1.69)式对 E 求导也可以看到(1.72)式是穿透系数 D 的极大值位置(见习题)．从(1.67)、(1.69)式又可看到，$E = 0$ 时 $D = 0$．于是得到 D 随 E 变化的图形如图6．穿透系数(推广到三维就是散射截面)在能量取某些值时达到极大的现象称为共振穿透(共振散射)．这里看到的是这种现象的一个最简单的例子．

容易看到，决定穿透系数 D 的极大值位置的条件(1.72)式正好就是决定势箱中能级的条件§5.2.2(2.38)式(注意，那里的势箱宽 a，而这里的势垒、势阱宽 $2a$)．这不是偶然的．实际上，两者都是粒子的概率波在势箱(或势垒、势阱)的区域内来回反射，相位一致，形成驻波的条件．

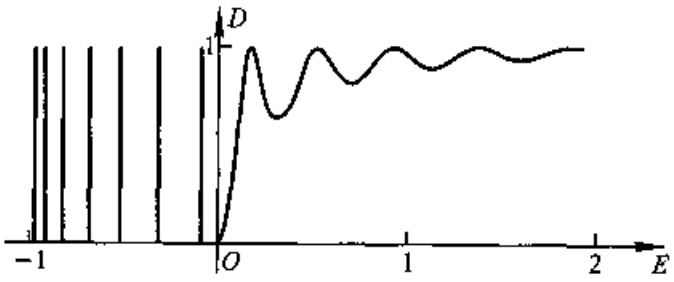


图6　穿透系数随能量的变化
左边的直线是势阱中能级的位置

§5.3.2　力学量对时间导数的算符　量子泊松括号　正则量子化

(一)力学量对时间导数的算符

在§5.3.1中讨论德布罗意波的概率流密度时，曾经将动量与质量之比 p/m 写成速度 v，并指出它不能简单地理解为坐标的测量值对时间的导数．实际上，坐标的测量值对时间的导数是不存在的．如果在某一时刻坐标 r 有确定值 $r(t)$，在这一时刻，动量的值就完全不确定 $\Delta p = \infty$，因而过了无限短的时间以后，坐标的值就完全不确定．

在量子力学中，“速度是坐标的时间导数”这一论断应该在算符的意义上，或测量结果的平均值的意义上理解．

定义　力学量 F 对时间的导数 $\mathrm{d}F/\mathrm{d}t$ 是另一个力学量，它的平均值等于力学量 F 的平均值对时间的导数：

$$\overline{\frac{\mathrm{d}F}{\mathrm{d}t}} = \frac{\mathrm{d}\bar{F}}{\mathrm{d}t}. \tag{2.1}$$

按§5.2.1(1.13)式写出在状态 Ψ 中 $\mathrm{d}F/\mathrm{d}t$ 和 F 的平均值，代入(2.1)式就有

$$\int \Psi^* \frac{\widehat{\mathrm{d}F}}{\mathrm{d}t}\Psi \mathrm{d}\tau = \frac{\mathrm{d}}{\mathrm{d}t}\int \Psi^* \hat{F}\Psi \mathrm{d}\tau. \tag{2.2}$$

这是力学量 F 对时间的导数的算符 $\frac{\mathrm{d}\hat{F}}{\mathrm{d}t}$ 的定义.

(2.1)式和(2.2)式的右边是力学量 F 的平均值对时间的导数.有两个原因导致力学量的平均值随时间变化:①波函数是时间的函数;② $\hat{F}$ 可能显含时间.这样:

$$\frac{\mathrm{d}\bar{F}}{\mathrm{d}t}=\int\Psi^{*}\frac{\partial\hat{F}}{\partial t}\Psi\mathrm{d}\tau+\int\frac{\partial\Psi^{*}}{\partial t}\hat{F}\Psi\mathrm{d}\tau+\int\Psi^{*}\hat{F}\frac{\partial\Psi}{\partial t}\mathrm{d}\tau \tag{2.3}$$

将薛定谔方程 $\partial\Psi/\partial t=\hat{H}\Psi/(\mathrm{i}\hbar)$ 及其复共轭 $\partial\Psi^{*}/\partial t=-(\hat{H}\Psi)^{*}/(\mathrm{i}\hbar)$ 代入得

$$\frac{\mathrm{d}\bar{F}}{\mathrm{d}t}=\int\Psi^{*}\frac{\partial\hat{F}}{\partial t}\Psi\mathrm{d}\tau+\frac{1}{\mathrm{i}\hbar}\int\Psi^{*}\hat{F}\hat{H}\Psi\mathrm{d}\tau-\frac{1}{\mathrm{i}\hbar}\int(\hat{H}\Psi)^{*}\hat{F}\Psi\,\mathrm{d}\tau.$$

利用 $\hat{H}$ 的厄米条件 $\int(\hat{H}\Psi)^{*}\hat{F}\Psi\,\mathrm{d}\tau=\int\Psi^{*}\hat{H}\hat{F}\Psi\mathrm{d}\tau$ 得到

$$\frac{\mathrm{d}\bar{F}}{\mathrm{d}t}=\int\Psi^{*}\left\{\frac{\partial\hat{F}}{\partial t}+\frac{\mathrm{i}}{\hbar}[\hat{H},\hat{F}]\right\}\Psi\mathrm{d}\tau. \tag{2.4}$$

式中 $[\hat{H},\hat{F}]=\hat{H}\hat{F}-\hat{F}\hat{H}$ 是 $\hat{H}$ 和 $\hat{F}$ 的对易关系.(2.4)式决定了在状态 Ψ 中力学量 F 的平均值如何随时间变化.

由(2.4)式可见,力学量 F 在任意状态中的平均值的时间导数,等于花括号中的算符在这一状态中的平均值.按(2.1)式就有

$$\frac{\mathrm{d}\hat{F}}{\mathrm{d}t}=\frac{\partial\hat{F}}{\partial t}+\frac{\mathrm{i}}{\hbar}[\hat{H},\hat{F}]. \tag{2.5}$$

特别要注意区分力学量对时间导数的算符 $\mathrm{d}\hat{F}/\mathrm{d}t$ 和力学量算符对时间的导数 $\partial\hat{F}/\partial t$ 的差别.例如,坐标算符 $\boldsymbol{r}$ 不含时间,$\partial\boldsymbol{r}/\partial t=0$,然而,坐标对时间的导数,即速度的算符是:

$$\hat{\boldsymbol{v}}=\frac{\mathrm{d}\hat{\boldsymbol{r}}}{\mathrm{d}t}=(\mathrm{i}/\hbar)[\hat{H},\boldsymbol{r}]. \tag{2.6}$$

在经典力学中,力学量对时间的导数可按 §1.6.2(2.3)式写成

$$\mathrm{d}f/\mathrm{d}t=\partial f/\partial t+[H,f]_{\mathrm{PB}}, \tag{2.7}$$

其中 $[H,f]_{\mathrm{PB}}$ 是(经典)泊松括号.(2.5)式是经典方程(2.7)式的量子对应,因而称 $(\mathrm{i}/\hbar)[\hat{H},\hat{F}]$ 为量子泊松括号,见 §5.2.3(3.9)式.

利用(2.5)式,并注意到坐标和动量算符都不明显依赖时间,可以得到:

$$\frac{\mathrm{d}\hat{\boldsymbol{r}}}{\mathrm{d}t}=(\mathrm{i}/\hbar)[\hat{H},\boldsymbol{r}],\qquad\frac{\mathrm{d}\hat{\boldsymbol{p}}}{\mathrm{d}t}=(\mathrm{i}/\hbar)[\hat{H},\hat{\boldsymbol{p}}]. \tag{2.8}$$

这是经典运动方程 §1.6.2(2.4)式的量子对应. 而且可以证明❶,当 $\hbar \to 0$ 时,量子泊松括号过渡到经典泊松括号,这样(2.8)式就过渡到经典运动方程, §1.6.2(2.4)式. 量子力学在 $\hbar \to 0$ 时过渡到经典力学的原因是很清楚的. 因为微观粒子的二象性体现在不确定关系中,但当 $\hbar \to 0$ 时, $\Delta x \cdot \Delta p_x \to 0$,即坐标 x 和动量 p_x 可以同时有确定值,回到经典力学. 还应指出的是,当一个量子系统接近于经典系统时,可以将该系统的运动方程按 $\hbar$ 的幂次展开而近似求解到前几级,这种方法称为半经典近似.

【例 1】 粒子在势能场 $U(\boldsymbol{r})$ 中运动,求坐标和动量对时间导数的算符.

解 哈密顿算符是 $\hat{H} = \hat{\boldsymbol{p}}^2/(2m) + U(\boldsymbol{r})$,代入(2.8)式

$$\begin{aligned}\frac{\mathrm{d}\hat{x}_i}{\mathrm{d}t} &= (\mathrm{i}/\hbar)[\hat{H}, x_i] = \mathrm{i}/(2m\hbar)\sum_j [\hat{p}_j^2, x_i] \\ &= \mathrm{i}/(2m\hbar)\sum_j (\hat{p}_j[\hat{p}_j, x_i] + [\hat{p}_j, x_i]\hat{p}_j) \\ &= \mathrm{i}/(m\hbar)\sum_j \hat{p}_j(-\mathrm{i}\delta_{ji}) = \hat{p}_i/m, \end{aligned} \tag{2.9}$$

$$\begin{aligned}\frac{\mathrm{d}\hat{p}_i}{\mathrm{d}t} &= (\mathrm{i}/\hbar)[\hat{H}, p_i] = (\mathrm{i}/\hbar)[U(\boldsymbol{r}), (-\mathrm{i}\hbar\partial/\partial x_i)] \\ &= [U(\boldsymbol{r}), \partial/\partial x_i]. \end{aligned} \tag{2.10}$$

将(2.10)式右方的对易关系作用到任意状态 $\Psi(\boldsymbol{r})$ 上,可以得到

$$[U(\boldsymbol{r}), \partial/\partial x_i] = -\partial U/\partial x_i, \tag{2.11}$$

因而(2.10)式成为 $\mathrm{d}\hat{p}_i/\mathrm{d}t = -\partial U/\partial x_i$,它和(2.9)式一道写成矢量形式就是

$$\frac{\mathrm{d}\hat{\boldsymbol{r}}}{\mathrm{d}t} = \frac{\hat{\boldsymbol{p}}}{m}, \qquad \frac{\mathrm{d}\hat{\boldsymbol{p}}}{\mathrm{d}t} = -\nabla U. \tag{2.12}$$

这再次表明,经典力学的运动方程在量子力学中以算符的形式成立.

(二)正则量子化

前面已经指出,(2.8)式和经典力学的正则运动方程 §1.6.2(2.4)式对应. 由此可见,如果采用正则变量(坐标和动量)来描述粒子,则从经典力学到量子力学的过渡可以通过将经典正则方程中的泊松括号换成量子泊松括号来实现,如 §5.2.3(3.9)式. 这种量子化方案称为正则量子化. 正则量子化在量子理论的发展和应用中具有重要的作用,因为按照这种量子化方案,只要知道一个物理系统的经典力学正则方程,就可以很方便地过渡到该系统的量子理论描述.

---

❶ 参看:曾谨言,量子力学,下册. 北京:科学出版社,1989.

§5.3.3　对称性与守恒律

(一)守恒量

在§5.3.1中指出,在定态中力学量的平均值不随时间变化.如果是非定态,则力学量的平均值一般要随时间变化.但是,有些力学量在任意状态中的平均值都不随时间变化:

$$\mathrm{d}F/\mathrm{d}t = 0 \quad (\text{对任意状态}), \tag{3.1}$$

这种力学量称为守恒量.由§5.3.2(2.1)式就知道,对于守恒量有

$$\overline{\frac{\mathrm{d}F}{\mathrm{d}t}} = 0 \quad (\text{对任意状态}), \tag{3.2}$$

因而必有 $\mathrm{d}\hat{F}/\mathrm{d}t = 0$.按§5.3.2(2.5)式就有

$$\frac{\mathrm{d}\hat{F}}{\mathrm{d}t} = \frac{\partial \hat{F}}{\partial t} + \frac{\mathrm{i}}{\hbar}[\hat{H},\hat{F}] = 0. \tag{3.3}$$

这是力学量 F 守恒的条件.

力学量的算符通常不明显地与时间有关:$\partial\hat{F}/\partial t = 0$.在此情况下,守恒量的条件是

$$[\hat{H},\hat{F}] = 0. \tag{3.4}$$

即:若不显含时间的力学量算符 $\hat{F}$ 和 $\hat{H}$ 对易,则 F 是守恒量.

注意,F 是守恒量并不意味着它在任意状态 ψ 中都取确定的数值.因为一个任意的状态 ψ 一般说来不是 $\hat{F}$ 的本征态,在状态 ψ 中,F 没有确定值.F 守恒只是它在 ψ 中的平均值(期望值)以及它在 ψ 中取不同值的概率分布不随时间变化.

我们来证明后一论断:如果 F 守恒,则它在任意状态 ψ 中取值的概率分布不随时间变化.

证　由(3.4)式,如果 F 守恒,则 $\hat{F}$ 和 $\hat{H}$ 对易,因而有共同本征函数系 $\varphi_n(n = 1,2,\cdots)$.令 $\psi = \sum_n c_n\varphi_n, c_n = \int\varphi_n^*\psi\mathrm{d}\tau$,则 F 在 ψ 中取值的概率分布是 $|c_n|^2$.利用薛定谔方程得到:

$$\begin{aligned}\mathrm{d}|c_n|^2/\mathrm{d}t &= (\mathrm{d}c_n^*/\mathrm{d}t)c_n + c_n^*\mathrm{d}c_n/\mathrm{d}t\\ &= \left(\int\varphi_n\frac{\partial\psi^*}{\partial t}\mathrm{d}\tau\right)c_n + c_n^*\int\varphi_n^*\frac{\partial\psi}{\partial t}\mathrm{d}\tau\\ &= -\frac{\mathrm{i}}{\hbar}\left[\left(\int\varphi_n(-\hat{H}^*\psi^*)\mathrm{d}\tau\right)c_n + c_n^*\int\varphi_n^*\hat{H}\psi\mathrm{d}\tau\right]\end{aligned}$$

$$= \frac{\mathrm{i}}{\hbar}\left(c_n\int\psi^*\hat{H}\varphi_n\mathrm{d}\tau - c_n^*\int\psi(\hat{H}\varphi_n)^*\mathrm{d}\tau\right)$$

$$= \frac{\mathrm{i}E_n}{\hbar}\left(c_n\int\psi^*\varphi_n\mathrm{d}\tau - c_n^*\int\psi\varphi_n^*\mathrm{d}\tau\right)$$

$$= (\mathrm{i}E_n/\hbar)(c_nc_n^* - c_n^*c_n) = 0,$$

这里利用了 $\hat{H}$ 的厄米性以及 $\hat{H}$ 的本征值 E_n 为实数的性质.　　**【证毕】**

【例 1】 证明自由粒子的动量守恒,中心场中粒子的角动量守恒.

解 自由粒子 $\hat{H} = \hat{\boldsymbol{p}}^2/(2m)$,因而:

$$[\hat{H},\hat{\boldsymbol{p}}] = [\hat{\boldsymbol{p}}^2/(2m),\hat{\boldsymbol{p}}] = 0.$$

粒子在中心场中运动时,

$$\hat{H} = -\frac{\hbar^2}{2mr^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\hat{\boldsymbol{L}}^2}{2mr^2} + U(r),$$

见 §5.2.3(3.31)式,而角动量算符 $\hat{\boldsymbol{L}}^2$、$\hat{L}_x$、$\hat{L}_y$、$\hat{L}_z$ 只与变量 θ、φ 有关,与 r 无关,它们与作用在 r 上的算符可对易,而且 $\hat{\boldsymbol{L}}^2$ 和 $\hat{\boldsymbol{L}}$ 的三个分量也可对易,因而有

$$[\hat{\boldsymbol{L}}^2,\hat{H}] = [\hat{L}_x,\hat{H}] = [\hat{L}_y,\hat{H}] = [\hat{L}_z,\hat{H}] = 0.$$

即知,粒子在中心场中运动时,角动量的平方以及角动量的三个分量都是守恒量.

(二)对称变换

在 §1.2.2 中已经看到,能量、动量和角动量的守恒与时空对称性有密切关系.我们来证明这一结论同样适用于量子力学系统.

一个力学系统的对称性就是它的运动规律的不变性.在经典力学里,运动规律由拉格朗日函数决定,因而时空对称性表现为拉格朗日函数在时空变换下的不变性(见 §1.2.2).在量子力学里,运动规律是薛定谔方程,它决定于系统的哈密顿算符 $\hat{H}$,因此,量子力学系统的对称性表现为哈密顿算符 $\hat{H}$ 的不变性.

用 S 表示某一空间变换,在这一变换下,任意波函数 Ψ 变为 $\Psi^{(S)}$

$$\Psi \to \Psi^{(S)} = \hat{S}\Psi. \tag{3.5}$$

设 $\hat{O}$ 为一个任意算符,它作用在 Ψ 上得到 Φ:

$$\hat{O}\Psi = \Phi. \tag{3.6}$$

经过变换 S 以后,应有

$$\hat{O}^{(S)}\Psi^{(S)} = \Phi^{(S)}. \tag{3.7}$$

用 $\hat{S}$ 作用在(3.6)式上,得到 $\hat{S}\hat{O}\Psi = \hat{S}\Phi = \Phi^{(S)}$.再定义逆变换 $\hat{S}^{(-1)}$ 为

$$\hat{S}^{(-1)}\hat{S} = \hat{S}\hat{S}^{(-1)} = \mathbf{1}, \tag{3.8}$$

$\mathbf{1}$ 表示恒等变换,即不变换.(3.8)式的含义是,变换 $\hat{S}$ 和逆变换 $\hat{S}^{(-1)}$ 相继作用的结果,相互抵消,等于不变换.因而可以写

$$\Phi^{(S)} = \hat{S}\hat{O}\mathbf{1}\Psi = \hat{S}\hat{O}\hat{S}^{-1}\hat{S}\Psi = \hat{S}\hat{O}\hat{S}^{-1}\Psi^{(S)}.$$

和(3.7)式比较可见：

$$\hat{O} \to \hat{O}^{(S)} = \hat{S}\hat{O}\hat{S}^{-1}. \tag{3.9}$$

(3.5)式和(3.9)式分别是在空间变换 S 之下，波函数 Ψ 和算符 $\hat{O}$ 的变换规律.

空间变换算符 $\hat{S}$ 有一个非常重要的性质——幺正性. 它来源于一个基本要求：空间变换应保持概率守恒. 根据这一要求，如果波函数在变换以前是归一化的(全空间中的总概率为1)，则在变换以后波函数 $\Psi^{(S)}$ 仍应保持为归一化. 即应有：

$$\int \Psi^{(S)*}\Psi^{(S)}\mathrm{d}\tau = \int \Psi^{*}\Psi\mathrm{d}\tau = 1. \tag{3.10}$$

将(3.5)式代入

$$\int (\hat{S}\Psi)^{*}(\hat{S}\Psi)\mathrm{d}\tau = \int \Psi^{*}\hat{S}^{\dagger}\hat{S}\Psi\mathrm{d}\tau = \int \Psi^{*}\Psi\mathrm{d}\tau.$$

由于 Ψ 是一个任意波函数，所以上式成立表明 $\hat{S}$ 应满足

$$\hat{S}^{\dagger}\hat{S} = 1, \qquad \hat{S}^{\dagger} = \hat{S}^{-1}. \tag{3.11}$$

这称为幺正性条件. 满足这一条件的算符称为幺正算符. 以上证明了，概率守恒要求空间变换算符是幺正算符. 将(3.11)式代入(3.9)式得到

$$\hat{O}^{(S)} = \hat{S}\hat{O}\hat{S}^{\dagger}. \tag{3.9'}$$

上述讨论适用于任意空间变换. 现在假定 S 是系统的一个对称变换，即系统的哈密顿算符在这一变换下不变：$\hat{H}^{(S)} = \hat{H}$. 由(3.9)式就有 $\hat{S}\hat{H}\hat{S}^{-1} = \hat{H}$. 从右边乘以 $\hat{S}$，利用(3.8)式得到 $\hat{S}\hat{H} = \hat{H}\hat{S}$，即：$\hat{H}$ 和 $\hat{S}$ 可对易：

$$\text{若 } \hat{S} \text{ 为系统的一个对称变换，则 } \quad [\hat{H},\hat{S}] = 0. \tag{3.12}$$

但应注意，$\hat{S}$ 是幺正算符，一般不是厄米算符[幺正算符的条件是(3.11)式 $\hat{S}^{\dagger} = \hat{S}^{-1}$ 而厄米算符的条件是§5.2.1(1.24)式 $\hat{L}^{\dagger} = \hat{L}$]，所以它本身一般不是守恒量算符，但它可以决定一个守恒量算符. 下面举几个例子.

1. 空间平移不变性与动量守恒

在空间平移 $\delta\boldsymbol{r}$ 时，波函数变化为❶

$$\psi(\boldsymbol{r}) \to \psi(\boldsymbol{r} - \delta\boldsymbol{r}) = (1 - \delta\boldsymbol{r}\cdot\nabla)\psi(\boldsymbol{r}). \tag{3.13}$$

这里假定了 $\delta\boldsymbol{r}$ 是小量，略去了它的高级项. 和(3.5)式比较可见，平移变换算符是

$$\hat{S}_{\delta\boldsymbol{r}} = 1 - \delta\boldsymbol{r}\cdot\nabla. \tag{3.14}$$

由(3.12)式知，如果系统有平移不变性(例

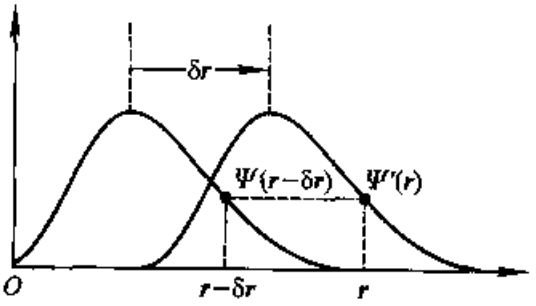


图1　波函数的平移

❶ 这里的"空间平移"系指函数平移. 由此得到的新波函数在 r 点的值等于原波函数在 $r - \delta r$ 处的值，故有(3.13)式，见图1. 也可以将"空间平移"理解为波函数不动而坐标平移. 这两种情况的变换公式中有一个符号差，但最终得到的守恒量算符相同.

如,无外场),则 $\hat{S}_{\delta r}$ 与 $\hat{H}$ 可对易,亦即 ∇ 与 $\hat{H}$ 可对易.但是,容易证明 ∇ 不是厄米算符,而 $\mathrm{i}\nabla$ 是厄米算符[见 §5.2.1(1.21)式].因此,$\mathrm{i}\nabla$ 是和平移不变性相联系的守恒量算符.乘以常数 $-\hbar$ 得到动量 $\hat{\boldsymbol{p}} = -\mathrm{i}\hbar\nabla$.因此,具有平移不变性的系统动量守恒.

2. 空间转动不变性与角动量守恒

当空间转动 $\delta\boldsymbol{\varphi}$ 时,按 §1.2.2(2.8)式有 $\delta\boldsymbol{r} = \delta\boldsymbol{\varphi} \times \boldsymbol{r}$,故波函数变化为

$$\begin{aligned}\psi(\boldsymbol{r}) \to \psi(\boldsymbol{r} - \delta\boldsymbol{\varphi} \times \boldsymbol{r}) &= \psi(\boldsymbol{r}) - (\delta\boldsymbol{\varphi} \times \boldsymbol{r}) \cdot \nabla\psi \\ &= \psi(\boldsymbol{r}) - \delta\boldsymbol{\varphi} \cdot (\boldsymbol{r} \times \nabla)\psi \\ &= [1 - \delta\boldsymbol{\varphi} \cdot (\boldsymbol{r} \times \nabla)]\psi.\end{aligned}$$

和(3.5)式比较可见,转动变换算符是

$$\hat{S}_{\delta\varphi} = 1 - \delta\boldsymbol{\varphi} \cdot (\boldsymbol{r} \times \nabla). \tag{3.15}$$

如果系统有转动不变性(例如,无外场或外场中心对称),则 $\hat{S}_{\delta\varphi}$ 与 $\hat{H}$ 可对易.这样 $\hat{\boldsymbol{L}} = -\mathrm{i}\hbar\boldsymbol{r} \times \nabla$ 与 $\hat{H}$ 可对易,因而有转动不变性的系统角动量 $\hat{\boldsymbol{L}}$ 守恒.

3. 空间反演对称性与宇称守恒

我们曾经指出,在量子力学中除了有和经典力学量对应的量以外,还有一些力学量没有经典力学量和它们对应.这类力学量的算符不能通过把经典表示式中的坐标和动量换成算符的方式得到,只能借助于对称性的分析来确定.

宇称就是没有经典对应的力学量,它是系统具有空间反演对称时的守恒量.空间反演变换是指粒子的位矢由 $\boldsymbol{r}$ 到 $-\boldsymbol{r}$ 的突变.这是一种不连续的空间变换,因而是经典力学中不会有的.

用算符 $\hat{I}$ 表示作用在波函数上的空间反演算符,其定义是:

$$\hat{I}\psi(\boldsymbol{r}) = \psi(-\boldsymbol{r}). \tag{3.16}$$

它把波函数的变量 $\boldsymbol{r}$ 变换为 $-\boldsymbol{r}$.用 $\hat{I}$ 再次作用到上式两边,得到

$$\hat{I}^2\psi(\boldsymbol{r}) = \hat{I}\psi(-\boldsymbol{r}) = \psi(\boldsymbol{r}).$$

即

$$\hat{I}^2 = 1. \tag{3.17}$$

是恒等变换.

(3.17)式表明 $\hat{I} = \hat{I}^{-1}$.另一方面,作为一个空间变换算符,$\hat{I}$ 应满足幺正性(3.11)式,$\hat{I}^{\dagger} = \hat{I}^{-1}$.故 $\hat{I}^{\dagger} = \hat{I}$.即 $\hat{I}$ 既是幺正算符,又是厄米算符.

如果系统具有空间反演对称,那就意味着 $[\hat{H}, \hat{I}] = 0$.由于 $\hat{I}$ 是厄米算符,所以它本身就是一个守恒量的算符.这种由系统空间反演对称决定的守恒量称为宇称.$\hat{I}$ 称为宇称算符,它没有明显的表示式,它的算符作用体现在它的定义式(3.16)中.

现在来求 $\hat{I}$ 的本征值.它的本征值方程为

$$\hat{I}\psi(\boldsymbol{r}) = I\psi(\boldsymbol{r}). \tag{3.18}$$

再用 $\hat{I}$ 作用一次，并利用(3.17)式，得

$$\hat{I}^2\psi(\boldsymbol{r}) = \hat{I}\,I\psi(\boldsymbol{r}) = I\,\hat{I}\psi(\boldsymbol{r}) = I^2\psi(\boldsymbol{r}) = \psi(\boldsymbol{r}).$$

由此得到 $I^2 = 1$，故 $\hat{I}$ 的本征值为

$$I = \pm 1. \tag{3.19}$$

相应地有两个本征态：

$$\hat{I}\psi_s(\boldsymbol{r}) = \psi_s(\boldsymbol{r}), \qquad \hat{I}\psi_a(\boldsymbol{r}) = -\psi_a(\boldsymbol{r}). \tag{3.20}$$

由(3.16)，(3.20)式可见 $\psi_s(-\boldsymbol{r}) = \psi_s(\boldsymbol{r})$，$\psi_a(-\boldsymbol{r}) = -\psi_a(\boldsymbol{r})$．这表明，宇称算符 $\hat{I}$ 的两个本征态 ψ_s，ψ_a 分别是偶函数和奇函数，它们分别对应宇称的两个本征值 $I = +1$，和 $I = -1$．由此可见，宇称的具体意义是表明波函数的奇偶性．

如果一个系统具有空间反演对称，即 $\hat{I}\hat{H} = \hat{H}\hat{I}$，则 $\hat{H}$ 的本征函数系也是 $\hat{I}$ 的本征函数系．这种系统的定态波函数有确定的宇称，即有确定的奇偶性——或为奇函数，或为偶函数．宇称守恒要求，状态波函数的奇偶性不随时间变化．

在 1956 年以前，人们一直认为自然界的各种基本相互作用过程都遵从宇称守恒．但是，李政道、杨振宁和吴健雄证实了，在弱相互作用过程中宇称不守恒，从而使人类对自然界的对称性有了新的认识．

【例 2】 利用宇称守恒重解 §5.3.1 例 4 中的对称方形势阱．

解 势能 $U(x)$ 如 §5.3.1(1.31)式(见该节图 1)，它满足 $U(-x) = U(x)$，故 $\hat{H} = -[\hbar^2/(2m)]\mathrm{d}^2/\mathrm{d}x^2 + U(x)$ 满足 $\hat{H}(-x) = \hat{H}(x)$，具有空间反演不变性．因而这一系统有宇称守恒．定态($\hat{H}$ 的本征态)是宇称的本征态，具有一定的奇偶性，$\psi(-x) = \pm\psi(x)$．

定态方程的解 §5.3.1(1.33) ~ (1.35)式加上无穷远边界条件以后，成为

$$\psi(x) = \begin{cases} A\mathrm{e}^{\lambda x} & \text{当 } x < -a, \\ B\cos kx + B'\sin kx & \text{当 } |x| < a, \\ C\mathrm{e}^{-\lambda x} & \text{当 } x > a. \end{cases} \tag{3.21}$$

对它分别加上条件 $\psi(-x) = \psi(x)$ 和 $\psi(-x) = -\psi(x)$，得到两类解：

$$\psi_s = \begin{cases} A\mathrm{e}^{\lambda x} & \text{当 } x < -a, \\ B\cos kx & \text{当 } |x| < a, \\ A\mathrm{e}^{-\lambda x} & \text{当 } x < a. \end{cases}$$

$$\psi_a = \begin{cases} A\mathrm{e}^{\lambda x} & \text{当 } x < -a, \\ B'\sin kx & \text{当 } |x| < a, \\ -A\mathrm{e}^{-\lambda x} & \text{当 } x > a. \end{cases} \tag{3.22}$$

这表明，宇称守恒条件使原来的 4 个未知系数 A，B，B'，C 简化为只剩下两个．

在 $x=-a$ 处的连续性条件 $\psi_{\mathrm{I}}(-a)=\psi_{\mathrm{II}}(-a)$，$\psi'_{\mathrm{I}}(-a)=\psi'_{\mathrm{II}}(-a)$，可以合写成 $\psi'_{\mathrm{I}}(x)/\psi_{\mathrm{I}}(x)\big|_{x=-a}=\psi'_{\mathrm{II}}(x)/\psi_{\mathrm{II}}(x)\big|_{x=-a}$，或

$$\mathrm{d}\ln\psi_{\mathrm{I}}(x)/\mathrm{d}x\big|_{x=-a}=\mathrm{d}\ln\psi_{\mathrm{II}}(x)/\mathrm{d}x\big|_{x=-a}, \tag{3.23}$$

这称为对数导数连续条件. 对 ψ_s 和 ψ_a 分别得到

$$\lambda=k\tan ka, \qquad \lambda=-k\cot ka. \tag{3.24}$$

这就是决定势阱中能谱的条件 §5.3.1(1.40)式. 很容易看到 $x=a$ 处的对数导数连续导致同一条件. 这并不奇怪，是势能的反演对称性决定的. 我们正是利用了由这一对称性导致的宇称守恒，使求能谱的工作大为简化[比较 §5.3.1 中从(1.37)式到(1.40)式的推导过程].

【例 3】 证明角动量和宇称有共同的本征函数系，并求角动量本征函数(球谐函数)的宇称.

解 宇称算符对 $\boldsymbol{r}$ 的作用是

$$\hat{I}\boldsymbol{r}\hat{I}^{-1}=-\boldsymbol{r}, \tag{3.25}$$

因此，

$$\hat{I}\hat{\boldsymbol{p}}\hat{I}^{-1}=\hat{I}(-\mathrm{i}\hbar\partial/\partial\boldsymbol{r})\hat{I}^{-1}=\mathrm{i}\hbar\partial/\partial=-\hat{\boldsymbol{p}}. \tag{3.26}$$

这表明，动量和坐标一样，在空间反演下改号. 这样的矢量称为真矢量，或极矢量.

对于角动量：

$$\hat{I}\hat{\boldsymbol{L}}\hat{I}^{-1}=\hat{I}(\boldsymbol{r}\times\hat{\boldsymbol{p}})\hat{I}^{-1}=\hat{I}\boldsymbol{r}\hat{I}^{-1}\times\hat{I}\hat{\boldsymbol{p}}\hat{I}^{-1}=-\boldsymbol{r}\times(-\hat{\boldsymbol{p}})=\boldsymbol{r}\times\hat{\boldsymbol{p}}=\hat{\boldsymbol{L}}. \tag{3.27}$$

这里，在 $\boldsymbol{r}\times\hat{\boldsymbol{p}}$ 中插入了 $\mathbf{1}=\hat{I}^{-1}\hat{I}$，并利用了叉乘和 $\hat{I}$ 可对易的性质. (3.27)式表明，角动量在空间反演下不改号. 这样的矢量在空间反演下和位矢 $\boldsymbol{r}$ 有相反的变换性质，因而称为赝矢量，或轴矢量.

由(3.27)式可见，$[\hat{I},\hat{\boldsymbol{L}}]=0$，$\hat{I}$ 和 $\hat{\boldsymbol{L}}$ 可对易，因而角动量和宇称有共同本征函数系.

在球坐标中，空间反演 $\boldsymbol{r}\to-\boldsymbol{r}$ 使得

$$r\to r, \quad \theta\to\pi-\theta, \quad \varphi\to\pi+\varphi \tag{3.28}$$

(参看 §1.1.1 图 1). 按附录(Ⅳ.31)式

$$Y_{lm}(\theta,\varphi)\propto P_l^{|m|}(\cos\theta)\mathrm{e}^{\mathrm{i}m\varphi}.$$

缔合勒让德多项式可写为[见附录(Ⅳ.29)式]

$$P_l^{|m|}(\cos\theta)=(1-x^2)^{|m|/2}\mathrm{d}^{|m|}P_l(x)/\mathrm{d}x^{|m|}\big|_{x=\cos\theta}.$$

右边的第一个因子是 $\cos\theta$ 的偶函数，而第二个因子是 $\cos\theta$ 的 $l-|m|$ 次多项式. 因而在 $\theta\to\pi-\theta$ 时，$\cos\theta\to-\cos\theta$，$P_l^{|m|}(\cos\theta)$ 乘因子 $(-1)^{l-|m|}=(-1)^{l-m}$；而 $\mathrm{e}^{\mathrm{i}m\varphi}$ 在 $\varphi\to\pi+\varphi$ 时，乘因子 $\mathrm{e}^{\mathrm{i}m\pi}=(-1)^m$. 因而在空间反演时

$$Y_{lm}(\theta,\varphi)\to(-1)^{l-m}(-1)^mY_{lm}(\theta,\varphi)=(-1)^lY_{lm}(\theta,\varphi).$$

因此，$Y_{lm}(\theta,\varphi)$ 的宇称是 $P = (-1)^l$.

习　题

1. 求在状态 $\psi_1(r) = R(r)Y_{10}(\theta,\varphi)$，$\psi_2(r) = R(r)Y_{11}(\theta,\varphi)$，$\psi_3(r) = R(r)Y_{1-1}(\theta,\varphi)$ 中的粒子的概率流密度，其中 Y_{10}，$Y_{1\pm1}$ 分别为 $l = 1$ 的球谐函数，见附录(Ⅳ.33)式.

2. 对相对论能量与动量的关系[§2.3.1(1.21)式] $E^2 = c^2p^2 + m^2c^4$ 作 §5.3.1(1.6)式的代换得到的方程称为克莱因-高登方程. 试写出这一方程，并证明对这一方程，满足概率守恒条件 $\partial\rho/\partial t + \nabla \cdot \boldsymbol{J} = 0$ 的概率密度 ρ 不是正定的.

3. 自由运动粒子初始时刻处于坐标有确定值 $\boldsymbol{r} = \boldsymbol{r}_0$ 的状态，求任意时刻的波函数 $\Psi(\boldsymbol{r},t)$.

4. 证明在定态中，力学量的概率分布不随时间变化.

5. 证明在具有分立能量本征值的定态中，动量的平均值为零.

6. 求在一维无限深势阱中运动的粒子的能量和波函数.

7. 质量为 m 的粒子在一维势阱

$$U(x) = \begin{cases} \infty, & \text{当 } x \leqslant 0, \\ -U_0, & \text{当 } 0 < x < a, \\ 0, & \text{当 } x \geqslant a \end{cases}$$

中运动，求粒子的能谱.

8. 质量为 m 的粒子在一维势能

$$U(x) = \begin{cases} U_1, & \text{当 } x \leqslant a, \\ 0, & \text{当 } a < x < b, \\ U_2, & \text{当 } x \geqslant b \end{cases}$$

的场中运动 $(0 < U_1 < U_2)$，证明当粒子能量 $E < U_1$ 时，能谱由关系式

$$k(b-a) = n\pi - \arcsin\left(\hbar k/\sqrt{2mU_1}\right) - \arcsin\left(\hbar k/\sqrt{2mU_2}\right)$$

决定，其中 $k = \sqrt{2mE/\hbar^2}$.

9. 证明 $E > 0$ 的粒子受到一维对称方形势阱[§5.3.1(1.31)式]散射的穿透系数 D [§5.3.1(1.69)式]在 $2k'a = n\pi(n = 1,2,\cdots)$ 处有极大值，在 $2k'a$ 近似等于 $(n + 1/2)\pi$ 处有极小值. 适当选取参数，用计算机画出 D 作为 E 的函数的图形.

10. 粒子在势能为 $U(x)$ 的保守场中运动，证明

$$\frac{\mathrm{d}\bar{x}}{\mathrm{d}t} = \frac{\bar{p}_x}{m}, \qquad \frac{\mathrm{d}\bar{p}}{\mathrm{d}t} = -\overline{\frac{\partial U}{\partial x}}.$$

11. 证明氢原子宇称守恒.

第四章　量子力学的表述形式

前面三章分别讨论了量子力学关于系统状态的假定，关于力学量的假定和关于运动方程的假定.除关于多粒子系统的全同性原理留到下一章讨论外，已基本学完了量子力学的原理.

任何一种运动形态的规律都有不同的表述形式.量子力学的表述形式也是多样的.在这一章，我们将建立希尔伯特空间的概念，并讨论由于希尔伯特空间中基底选择的不同，使量子力学原理有不同的表象；以及由于对时间演化的处理方法不同，使量子力学有不同的绘景.

量子力学理论也会由于所采用的量子化方案不同而有不同的形式.本教程主要采用正则量子化方案，即通过引入对易关系使力学量由 c 数变成 q 数(算符)而实现量子化，见§5.3.2.1948 年费曼提出了另一种量子化方案，即路径积分量子化，它在现代量子理论中有广泛的应用.在本篇第六章中将对路径积分量子化作简单介绍.

§5.4.1　希尔伯特空间

(一)几个不同表象的例子

我们先通过一个简单例子说明，量子力学理论能够有不同的表述形式.

在§5.2.2 中讨论了用算符 $\hat{F}$ 的本征函数系展开任意函数 $\Psi(r,t)$ 的问题，如那一节的(2.61)、(2.62)式.为了具体起见，令 f 为动量 p，并且只考虑一维情况.上述两个式子成为

$$\Psi(x,t) = \int c(p,t)\psi_p(x)\mathrm{d}p, \tag{1.1}$$

$$c(p,t) = \int \psi_p^*(x)\Psi(x,t)\mathrm{d}x. \tag{1.2}$$

将归一化的本征函数[§5.2.2(2.53)式]代入，得到

$$\Psi(x,t) = (1/\sqrt{2\pi\hbar})\int c(p,t)\mathrm{e}^{(\mathrm{i}/\hbar)px}\mathrm{d}p, \tag{1.3}$$

$$c(p,t) = (1/\sqrt{2\pi\hbar})\int \Psi(x,t)\mathrm{e}^{-(\mathrm{i}/\hbar)px}\mathrm{d}x. \tag{1.4}$$

熟悉傅里叶变换的读者一定会看出，这正是傅里叶变换和反变换的公式.这两个公式表明，由 $\Psi(x,t)$ 唯一决定了 $c(p,t)$；由 $c(p,t)$ 也唯一决定了 $\Psi(x,t)$.$c(p,t)$ 称为 $\Psi(x,t)$ 的傅里叶变换象函数，而 $\Psi(x,t)$ 称为 $c(p,t)$ 的傅里叶变换原函数.

因此,完全可以用 $c(p,t)$ 代替 $\Psi(x,t)$ 作为描述量子态的态函数. 由于 $c(p,t)$ 以 p 为变量,所以称为动量表象的态函数. 相应地称 $\Psi(x,t)$ 为坐标表象的态函数.

下面举一个例子来说明,在动量表象中和在坐标表象中一样,可以解决量子力学的问题.

【例 1】 粒子在均匀力场 $U(x) = -Fx$ 中运动,求定态波函数.

解 定态方程是

$$-[\hbar^2/(2m)](\mathrm{d}^2/\mathrm{d}x^2)\psi(x) - Fx\psi(x) = E\psi(x). \tag{1.5}$$

这是一个二阶变系数常微分方程,可以用级数解法求解. 我们不用这种办法,而用傅里叶变换法,令

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}}\int c(p)\mathrm{e}^{(\mathrm{i}/\hbar)px}\mathrm{d}p, \quad c(p) = \frac{1}{\sqrt{2\pi\hbar}}\int \psi(x)\mathrm{e}^{-(\mathrm{i}/\hbar)px}\mathrm{d}x. \tag{1.6}$$

我们有

$$\frac{\mathrm{d}\psi}{\mathrm{d}x} = \frac{\mathrm{i}/\hbar}{\sqrt{2\pi\hbar}}\int pc(p)\mathrm{e}^{(\mathrm{i}/\hbar)px}\mathrm{d}p, \qquad \frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = \frac{-1}{\hbar^2\sqrt{2\pi\hbar}}\int p^2c(p)\mathrm{e}^{(\mathrm{i}/\hbar)px}\mathrm{d}p;$$

$$\frac{\mathrm{d}c(p)}{\mathrm{d}p} = \frac{-\mathrm{i}/\hbar}{\sqrt{2\pi\hbar}}\int x\psi(x)\mathrm{e}^{-(\mathrm{i}/\hbar)px}\mathrm{d}x, \quad -\frac{\mathrm{i}}{\hbar}x\psi(x) = \frac{1}{\sqrt{2\pi\hbar}}\int \frac{\mathrm{d}c(p)}{\mathrm{d}p}\mathrm{e}^{(\mathrm{i}/\hbar)px}\mathrm{d}p.$$

将这几个式子和(1.6)式一道,代入方程(1.5)式. 由于函数系 $\mathrm{e}^{(\mathrm{i}/\hbar)px}$ 的正交完备性,被积函数应相等,得到

$$p^2/(2m)c(p) - \mathrm{i}\hbar F\mathrm{d}c(p)/\mathrm{d}p = Ec(p). \tag{1.7}$$

这是 $c(p)$ 的一阶微分方程,可以直接积分,得到

$$c_E(p) = A\exp\left[\frac{\mathrm{i}}{\hbar F}\left(\frac{p^3}{6m} - Ep\right)\right]. \tag{1.8}$$

这就是能量为 E 的定态在动量表象中的态函数.

讨论:

①由于在解方程时能量 E 未受到限制,所以能谱连续.

②动量的概率分布 $|c_E(p)|^2 = |A|^2$ 是常数,说明动量以均匀的概率分布.

③坐标的概率分布可以通过将(1.8)式代入(1.6)的第一式而得到

$$\psi(x) = \frac{A}{\sqrt{2\pi\hbar}}\int \exp\left\{-\frac{\mathrm{i}}{\hbar F}\left(\frac{p^3}{6m} - (E+Fx)p\right)\right\}\mathrm{d}p. \tag{1.9}$$

这就是用积分形式写出的坐标表象波函数,而位置概率密度是 $|\psi(x)|^2$.

我们看到,量子力学的状态既可以用坐标作为变量来描述(坐标表象),也可以用动量作为变量来描述(动量表象);更一般地,也应该可以用其他力学量,例如 F 作为变量来描述. 为了具体起见,假定 F 有分立谱,如 §5.2.2(2.57),(2.58)式. 为了书写简单,仍然只考虑一维,

$$\Psi(x,t) = \sum_i c_i(t)\psi_i(x), \tag{1.10}$$

$$c_i(t) = \int \psi_i^*(x)\Psi(x,t)\mathrm{d}x, \qquad i = 1,2,\cdots, \tag{1.11}$$

其中，$\psi_i(x)$ 是力学量 F 的本征函数，$\hat{F}\psi_i(x) = f_i\psi_i(x)(i = 1,2,\cdots)$.

注意，$\Psi(x,t)$ 能用 $\psi_i(x)$ 展开成(1.10)式是由于本征函数系 $\psi_i(x)$ 的完备性，见§5.2.2(五)中的讨论；而展开系数 $c_i(t)$ 能写成(1.11)式是由于本征函数系 $\psi_i(x)$ 的正交归一性.

由(1.10)、(1.11)式可见，$\Psi(x,t)$ 和 $c_i(t)$ 相互单值决定，只要给定 $\Psi(x,t)$ 就知道 $c_i(t)(i = 1,2,\cdots)$；而只要给定 $c_i(t)(i = 1,2,\cdots)$ 也就知道 $\Psi(x,t)$. 因此，完全可以用 $c_i(t)$ $(i = 1,2,\cdots)$ 来代替 $\Psi(x,t)$ 作为描述系统状态的态函数. 这样的描述方式称为 F 表象.

(二)希尔伯特空间

从上面的例子看到，量子力学理论既可以用坐标作自变量表述，也可以用动量作自变量表述；更一般地说，可以用任一力学量完全组作自变量表述(上例是一维运动，只要一个变量 x 或 p 或 F 就形成自变量完全组). 这许多不同的表述，相互之间完全等效. 这样就产生一个问题：为什么同一理论有这么多不同的表述方式呢？这些不同表述所反映的共同本质是什么呢？

为了有助于回答这一问题，让我们回忆一下三维空间中的经典力学. 我们知道，经典力学的运动状态既可以在坐标系 K 中用 $x(t),y(t),z(t)$ 描述；也可以在坐标系 K′中用 $x'(t),y'(t),z'(t)$ 描述. 为什么同一个力学系统的运动状态可以用不同方式描述呢？其原因是：在这许多不同描述方式的背后有一个共同的客观物理对象——质点的位置矢量 $\boldsymbol{r}(t)$. 力学运动状态的不同描述方式是位置矢量用不同坐标基矢展开的不同分量形式.

与此类似，可以预期，量子力学的运动状态之所以有不同描述方式(不同表象)的原因，是这许多不同描述方式背后有一个共同的客观物理对象——状态矢量；量子力学运动状态的不同描述方式是状态矢量用不同基矢展开的不同分量形式.

在§5.2.2中讨论厄米算符本征函数的正交性时曾经指出，本征函数，或者更一般地说任意状态的波函数，可以看成是一个无穷维复矢量空间中的矢量——状态矢量. 这一复矢量空间是通常的三维矢量空间的推广，如下式❶

---

❶ 表示态矢量的符号最方便的是下一节中将要引入的狄拉克符号. 目前，暂时用不带宗量的 ψ、φ 表示态矢量.

矢量	$\boldsymbol{a}$	$\to \psi$
自由度编号	i （取分立值）	$\to x$ （取连续值）
分量	a_i	$\to \psi(x)$
点积	$\boldsymbol{a}\cdot\boldsymbol{b}=\sum a_i^* b_i$	$\to (\psi,\varphi)=\int\psi^*(x)\varphi(x)\mathrm{d}x$
正交性	$\sum a_i^* b_i=0$，则 $\boldsymbol{a}\perp\boldsymbol{b}$	$\to \int\psi^*(x)\varphi(x)\mathrm{d}x=0$，则 $\psi\perp\varphi$

(1.12)

2.18 在同一节中，讨论厄米算符的本征函数形成正交完备函数系时，还曾经将它和转动惯量的惯量主轴形成三维空间的正交坐标系相对比，指出厄米算符可以看成是矢量空间中的对称矩阵的推广. 正如转动惯量的主轴可以作为三维空间中的正交坐标系来展开任意矢量一样，厄米算符的本征函数系，是一个正交完备函数系，也可以用来作为坐标基矢展开任意状态矢量.

仔细分析起来，这一推广过程有以下几步：

①在实 N 维矢量空间中，矢量的分量只能取实数值，推广成复 N 维矢量空间，矢量的分量可以取复数值. 此时，在矢量点积的定义中，左边一个矢量的分量应取复共轭. 这样才能保证任意矢量 $\boldsymbol{a}$ 和它自身的点积恒正，即矢量模的平方

$$|\boldsymbol{a}|^2=\boldsymbol{a}\cdot\boldsymbol{a}=\sum_{i=1}^{N}a_i^* a_i\geqslant 0. \tag{1.13}$$

按照这样的定义，两个不同矢量按相反次序的点积，即：$\boldsymbol{a}\cdot\boldsymbol{b}=\sum a_i^* b_i$ 和 $\boldsymbol{b}\cdot\boldsymbol{a}=\sum b_i^* a_i$，并不相等，而是相互复共轭：

$$\boldsymbol{a}\cdot\boldsymbol{b}=(\boldsymbol{b}\cdot\boldsymbol{a})^*. \tag{1.14}$$

②将空间维数 N 由有限推广到无穷. 当 N 有限时，矢量模的平方(1.13)式取有限值；当 $N\to\infty$ 时，(1.13)式右方成为无穷积数，可能发散. 因此，对于这种无穷维复空间的矢量应加上模有限的条件：

$$\sum_{i=1}^{\infty}a_i^* a_i \quad \text{收敛}. \tag{1.15}$$

③空间维数的标号由取分立值的 i 推广成取连续值的 x，$a_i\to f(x)$. 并将 a_i 所应满足的条件(1.15)式推广成为加在 $f(x)$ 上的条件

$$\int_{-\infty}^{\infty}\psi^*(x)\psi(x)\mathrm{d}x \quad \text{收敛}. \tag{1.16}$$

这相当于要求 $\psi(x)$ 能按 §5.1.2(2.12)式归一化●.

● 在 §5.1.2 的末尾曾指出："无限运动"的波函数不满足(1.16)式. 然而，真正的无限运动并不存在. 所谓"无限运动"只不过是运动范围达到了宏观的尺度. 所以，对(1.16)式的违反实际上只是形式上的.

④将三维空间中的对称矩阵（$A = \tilde{A}$）推广到无穷维空间，成为厄米算符 $F = F^{\dagger} = \tilde{F}^{*}$.

⑤三维空间中的对称矩阵（例如由转动惯量张量的分量组成的矩阵）有三个正交的主轴. 这三个主轴方向的单位矢量称为对称矩阵的本征矢量，由它们组成的坐标系可以展开三维空间中的任意矢量. 这表明，对称矩阵的本征矢量形成三维空间中的完备矢量组. 由此推广，就要求厄米算符的本征函数［满足 $\hat{F}\varphi_i(x) = f_i\varphi_i(x)$ 的 $\varphi_i(x)$］形成完备函数系，用它可以展开任意函数，即：展开式

$$\psi(x) = \sum_{i=1}^{\infty} c_i\varphi_i(x) \tag{1.17}$$

右边的函数项无穷级数收敛到 $\psi(x)$.

按照以上方式得到的模有限的无穷维复线性空间称为希尔伯特空间.

量子力学理论是建立在希尔伯特空间的基础上——量子系统的状态由希尔伯特空间中的矢量描述，而力学量由希尔伯特空间中的厄米算符描述.

我们知道，三维空间中的矢量 a 在不同坐标系中有不同的分量形式，而各不同的分量形式代表同一个客观对象，因而都可以等效地用来作计算. 与此类似，希尔伯特空间中的状态矢量也可以用不同的“坐标系”展开成 $\psi(x), c(p), \cdots$. 这就是量子力学理论能够有不同表象的本质原因.

在三维空间里对矢量进行具体计算，总是要选定坐标系，用这一坐标系中的分量形式运算；然而在推导公式和理解公式的物理意义时，采用抽象的矢量形式更为方便. 与此类似，将量子力学理论用希尔伯特空间中的抽象矢量形式表述出来，也能使表述更简练，意义更清楚. 有了这种抽象的表述，在进行实际计算时，可以根据不同问题的需要进入不同的表象. 在以下几节里，我们就来给出量子力学理论的这种抽象表述，并说明从这种抽象表述进入具体表象的方法.

§5.4.2　态矢量和算符

这一节在希尔伯特空间里给出量子力学原理的表述.

（一）态矢量

状态用希尔伯特空间的矢量描述，称为态矢量.

希尔伯特空间的矢量，可以用狄拉克所创立的符号表示. 这种符号是一个尖角指向右边的括号，并将标志不同状态的参量写在括号内. 例如，状态 a 的态矢量记为 $|a\rangle$，物理量 F 的第 i 个本征态矢量记为 $|i\rangle$，等等. 通常称这种矢量为狄拉克右矢，简称为右矢.

态矢量满足叠加原理：如果 $|a\rangle$、$|b\rangle$ 是两个态矢量，则

$$|c\rangle = \alpha|a\rangle + \beta|b\rangle \tag{2.1}$$

也是一个态矢量,其中 α 和 β 是两个任意复数.

态矢量的加法满足交换律和结合律:

$$|a\rangle + |b\rangle = |b\rangle + |a\rangle, \tag{2.2}$$

$$|a\rangle + (|b\rangle + |c\rangle) = (|a\rangle + |b\rangle) + |c\rangle. \tag{2.3}$$

数和态矢量的乘法也满足结合律:

$$\alpha(\beta|a\rangle) = (\alpha\beta)|a\rangle. \tag{2.4}$$

如果 n 个矢量,$|a_1\rangle, |a_2\rangle, \cdots, |a_n\rangle$ 存在以下关系:

$$\alpha_1|a_1\rangle + \alpha_2|a_2\rangle + \cdots + \alpha_n|a_n\rangle = 0, \tag{2.5}$$

其中 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 不全为零,则称这 n 个矢量线性相关. 如果找不到 n 个不全为零的数使(2.5)式成立,则称这 n 个矢量线性独立. 特别是如果 $|a\rangle = \alpha|b\rangle, \alpha \neq 0$,则称 $|a\rangle$ 和 $|b\rangle$ 线性相关,否则就称 $|a\rangle$ 和 $|b\rangle$ 线性独立.

(二)态矢量的点积

和每一个右矢对应有一个左矢,用尖角指向左边的括号表示,如 $\langle a|, \langle i|$. 右矢和左矢相互共轭:

$$|a\rangle^\dagger = \langle a|, \quad \langle a|^\dagger = |a\rangle. \tag{2.6}$$

这里的"共轭"是希尔伯特空间里一种抽象意义下的共轭,它的具体含义要在定义了矢量点积以后才明了.

状态 a 的态矢量可以等效地用右矢 $|a\rangle$ 或左矢 $\langle a|$ 表示. 两个状态 a 和 b 的态矢的点积是一个复数,记为:

$$\langle a|b\rangle \quad \text{是一个复数.} \tag{2.7}$$

注意,状态 a, b 的态矢的点积是取 a 的左矢形式和 b 的右矢形式相乘.

两个矢量点积的共轭,等于它们各自的共轭按相反的次序取点积:

$$\langle a|b\rangle^\dagger = (|b\rangle^\dagger)(\langle a|^\dagger) = \langle b|a\rangle. \tag{2.8}$$

由于 $\langle a|b\rangle$ 是一个普通的复数,所以它的共轭就是复共轭,因此,

$$\langle a|b\rangle^* = \langle b|a\rangle. \tag{2.9}$$

这和 §5.4.1(1.14)式一致.

注意,态矢的点积必须是左矢形式和右矢形式相乘,而不能用右矢和右矢或者左矢和左矢相乘,即不能写 $|a\rangle|b\rangle$ 或 $\langle a|\langle b|$. 形象化地说,点积一定要两头尖,形成完整的括号,如(2.7)式.

(三)厄米算符

希尔伯特空间中的算符 F 是一种对应规则,它把希尔伯特空间的任意矢量对应成另一矢量:

$$F|a\rangle = |c\rangle. \tag{2.10}$$

由于右矢本身是抽象矢量，所以算符 F 不能写成具体形式，而只能用右矢之间的对应关系(2.10)式来定义. 注意，希尔伯特空间的抽象算符，简单地用大写字母表示，而不在上面加尖角，以区别于具体表象(例如 x 表象)中的算符.

如果算符 F 满足以下条件

$$F(\alpha|a\rangle+\beta|b\rangle)=\alpha F|a\rangle+\beta F|b\rangle, \tag{2.11}$$

则称 F 为线性算符.

前面已经定义了复数 α 和矢量 $|b\rangle$ 的乘积 $\alpha|b\rangle$，从这个意义上可以把复数也看成一个线性算符. 按定义，

$$\langle a|\alpha|b\rangle=\alpha\langle a|b\rangle. \tag{2.12}$$

线性算符 F 的厄米共轭算符定义为

$$(F|a\rangle)^{\dagger}=(|a\rangle)^{\dagger}F^{\dagger}=\langle a|F^{\dagger}. \tag{2.13}$$

注意，(2.13)式和(2.8)式一样，都是根据一条普遍的规则：两个对象(矢量和矢量，或算符和矢量，或算符和算符)相乘的厄米共轭等于它们各自的厄米共轭按相反次序相乘. (对于算符和态矢，“相乘”是指算符对态矢的作用). 这一规则下面常要用到. 利用它还可以写：

$$\langle a|F|b\rangle^{*}=\langle a|F|b\rangle^{\dagger}=(|b\rangle)^{\dagger}(F)^{\dagger}(\langle a|)^{\dagger}=\langle b|F^{\dagger}|a\rangle. \tag{2.14}$$

显然有：

$$(F^{\dagger})^{\dagger}=F. \tag{2.15}$$

如果线性算符 F 和它自身厄米共轭

$$F^{\dagger}=F, \tag{2.16}$$

则称 F 为厄米自共轭算符，简称厄米算符. 由(2.14)式知：

$$\langle a|F|b\rangle^{*}=\langle b|F|a\rangle \quad (\text{如果 } F \text{ 厄米}). \tag{2.17}$$

这是算符厄米性的条件.

线性算符的一个例子是右矢和左矢的并矢：$|a\rangle\langle b|$. 为了看到它的确是一个算符，将它作用到任意矢量 $|c\rangle$ 上(注意：狄拉克矢量的连乘满足结合律)：

$$(|a\rangle\langle b|)|c\rangle=|a\rangle(\langle b|c\rangle).$$

但是，$\langle b|c\rangle$ 是一个数 $\langle b|c\rangle=\alpha$，而 α 乘矢量 $|a\rangle$ 得另一矢量 $|d\rangle$，因而有：

$$(|a\rangle\langle b|)|c\rangle=|d\rangle,$$

这就表明 $|a\rangle\langle b|$ 是一个算符. 我们看到，矢量 $|a\rangle$ 和 $|b\rangle$ 有两种类型的乘积：两头尖 $\langle a|b\rangle$ 是点积，得到一个数；两头平 $|a\rangle\langle b|$ 是并矢，得到一个算符.

(四)厄米算符的本征值与本征矢

算符 F 的本征值方程是

$$F|i\rangle=f_i|i\rangle \quad (\text{分立谱}), \qquad F|f\rangle=f|f\rangle \quad (\text{连续谱}). \tag{2.18}$$

f_i 或 f 称为算符 F 的本征值，$|i\rangle$ 或 $|f\rangle$ 称为和本征值 f_i 或 f 对应的本征矢. 关于厄米算符的本征值和本征矢有以下定理. 证明时假定本征值谱分立. 在连续谱情况下，证明是类似的.（不难看到以下关于希尔伯特空间厄米算符的定理和 §1.4.3 第(二)小点中讨论的实欧氏空间中对称矩阵的相应定理是对应的. 两者的证明方法也完全类似.）

定理一 厄米算符的本征值是实数.

证 设 $F^\dagger = F, F|i\rangle = f_i|i\rangle$. 用本征矢 $\langle i|$ 点乘上式，得到

$$\langle i|F|i\rangle = \langle i|f_i|i\rangle = f_i\langle i|i\rangle. \tag{2.19}$$

按(2.9)式取上式的共轭，得到

$$\langle i|F|i\rangle^* = (f_i\langle i|i\rangle)^* = f_i^*\langle i|i\rangle.$$

再利用(2.17)式知上两式左边相等，因而 $f_i^* = f_i$，f_i 为实数. 【证毕】

定理二 厄米算符对应于不同本征值的本征矢相互正交.

证 设 $F^\dagger = F, F|i\rangle = f_i|i\rangle, F|j\rangle = f_j|j\rangle$. 用本征矢 $\langle j|$ 和 $\langle i|$ 分别点乘以上二式：

$$\langle j|F|i\rangle = f_i\langle j|i\rangle, \qquad \langle i|F|j\rangle = f_j\langle i|j\rangle. \tag{2.19'}$$

取前一式的共轭，并利用 F 厄米条件(2.17)式，得到 $\langle i|F|j\rangle = \langle j|F|i\rangle^* = f_i^*\langle j|i\rangle^* = f_i\langle i|j\rangle$，这里用了定理一，$f_i^* = f_i$. 和以下第二式相减得到

$$(f_i - f_j)\langle i|j\rangle = 0.$$

按假设 $f_i \neq f_j$，故 $\langle i|j\rangle = 0$. 【证毕】

如果和厄米算符 F 的某一个本征值 f_i 相对应，有一个以上的线性独立本征矢：$|ia\rangle$，$|ib\rangle$，…，则称本征值 f_i 有简并.

定理三 如果 $|ia\rangle$，$|ib\rangle$ 是对应于厄米算符 F 的本征值 f_i 的线性独立本征矢，则它们的线性叠加 $\alpha|ia\rangle + \beta|ib\rangle$ 也是对应于同一本征值的本征矢，即：

$$\left.\begin{aligned} &\text{设 } F|ia\rangle = f_i|ia\rangle, \quad F|ib\rangle = f_i|ib\rangle, \\ &\text{则 } F(\alpha|ia\rangle + \beta|ib\rangle) = f_i(\alpha|ia\rangle + \beta|ib\rangle). \end{aligned}\right\} \tag{2.20}$$

证 按(2.11)式

$$\begin{aligned} F(\alpha|ia\rangle + \beta|ib\rangle) &= \alpha F|ia\rangle + \beta F|ib\rangle = \alpha f_i|ia\rangle + \beta f_i|ib\rangle \\ &= f_i(\alpha|ia\rangle + \beta|ib\rangle). \end{aligned}$$

【证毕】

由此可见，如果本征值 f_i 有简并，则和这一本征值对应有无穷多本征矢. 这些本征矢形成希尔伯特空间的一个线性矢量子空间. 用 s 表示这一线性空间中独立矢量的最大数目（即这一线性空间的维数），则称本征值 f_i 有 s 重简并. 为了对这一论断增加形象的理解，回忆转动惯量张量有两个主值相等的情况. 此时，惯量椭球成为旋转椭球，在这两个相等主值对应的两个主轴所在平面上的截面是一个圆. 这一平面上的任意矢量，都是对应同一主值的惯量主轴. 这种矢量有无穷多个，它们

形成一个二维线性空间(平面).

总结上述,厄米算符 F 对应于不同本征值的本征矢相互正交.当本征值有 s 重简并时,和它对应的本征矢不一定正交,但是可以按定理三从这些本征矢所形成的线性空间中选出 s 个正交矢量(参看第五篇第二章习题9).这样做了以后,就得到了 F 的一个正交完备矢量组.用 i 给这一矢量组中的矢量重新编号,成为

$$|i\rangle, \quad i=1,2,\cdots. \tag{2.21}$$

这一组矢量满足正交归一化条件:

$$\langle i|j\rangle=\delta_{ij}. \tag{2.22}$$

同时,它们又具有完备性,即:希尔伯特空间的任意矢量 $|a\rangle$ 可以用矢量组(2.21)展开:

$$|a\rangle=\sum_{i=1}^{\infty}a_i|i\rangle, \tag{2.23}$$

此时,右边的矢量项无穷级数收敛到矢量 $|a\rangle$ [比较§5.4.1的(1.17)式].

为了求(2.23)式中的展开系数 a_i,可以先将此式右边的哑标 i 改写为 j,然后用 $\langle i|$ 点乘等式两边:

$$\langle i|a\rangle=\sum_{j=1}^{\infty}a_j\langle i|j\rangle=\sum_{j=1}^{\infty}a_j\delta_{ij}=a_i, \qquad a_i=\langle i|a\rangle. \tag{2.24}$$

如果矢量组(2.21)式是完备的,则展开式(2.23)成立.将(2.24)式代入(2.23)式,得到

$$|a\rangle=\sum_{i=1}^{\infty}|i\rangle a_i=\sum_{i=1}^{\infty}|i\rangle\langle i|a\rangle. \tag{2.25}$$

由此可见,

$$\sum_{i=1}^{\infty}|i\rangle\langle i|=\mathbf{1}. \tag{2.26}$$

是恒等算符,它作用在任意矢量上仍得到这一矢量本身.(2.26)式就是矢量组(2.21)的完备性条件.

以上假定 F 的本征值有分立谱,对于连续谱情况,定理一至定理三仍然成立,只是(2.22)-(2.26)式改动如下:

正交归一条件:
$$\langle f|f'\rangle=\delta(f-f'), \tag{2.27}$$

完备性条件:
$$\int|f\rangle\langle f|\,\mathrm{d}f=\mathbf{1}, \tag{2.28}$$

矢量展开式:
$$|a\rangle=\int a_f|f\rangle\,\mathrm{d}f, \tag{2.29}$$

展开式系数:
$$a_f=\langle f|a\rangle. \tag{2.30}$$

(五)力学量用厄米算符表示

力学量用希尔伯特空间中的线性厄米算符表示.具体含义如下:

力学量 F 所能够取的值是它的算符的本征值. 考虑用矢量 $|a\rangle$ 描述的状态，将 $|a\rangle$ 用 F 的本征矢量组展开，如(2.23)(2.24)式

$$|a\rangle = \sum_{i=1}^{\infty} a_i |i\rangle, \qquad a_i = \langle i|a\rangle, \tag{2.31}$$

则在状态 $|a\rangle$ 中，物理量 F 取值 f_i 的概率是

$$P_i = |a_i|^2 = |\langle i|a\rangle|^2. \tag{2.32}$$

关于力学量同时有确定值的条件 §5.2.3 定理一和定理二，以及 §5.2.3 中定义的力学量完全组，也可以用希尔伯特空间中的抽象矢量和抽象算符表述出来，这里不再赘述.

(六)运动方程

态矢量随时间的演化由希尔伯特空间的一个线性厄米算符 H 决定：

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}|\Psi,t\rangle = H|\Psi,t\rangle. \tag{2.33}$$

H 称为哈密顿算符. 注意，在 $|\Psi,t\rangle$ 中只有一个变量 t，因而对 t 的求导是全导数，而不是偏导数.

§5.4.3 表象和表象变换

(一)态矢量在 F 表象中的分量形式

在上一节里，我们用希尔伯特空间中的抽象矢量和抽象算符表述了量子力学的基本原理. 在进行具体计算时，常常需要选用具体的“坐标系”，即进入具体的表象.

正如三维空间中的对称矩阵(例如转动惯量张量)的三个本征矢量可以作为三维空间中的坐标系的基矢一样，希尔伯特空间中的厄米算符的本征矢量组也可以作为希尔伯特空间中的“坐标系”的基矢. 这一对比可以类似于 §5.4.1(1.12)式写成：

对称张量	$\vec{I}$	→ 厄米算符 F
本征值方程	$\vec{I}e^{(\alpha)} = I_\alpha e^{(\alpha)}$	→ $F\|i\rangle = f_i\|i\rangle$ 或 $F\|f\rangle = f\|f\rangle$
坐标基矢	$e^{(\alpha)}(\alpha = 1,2,3)$	→ 表象基矢 $\|i\rangle, i = 1,2,\cdots$；或 $\|f\rangle, f \in [-\infty,\infty]$
展开任意矢量	$a = \sum_\alpha a_\alpha e^{(\alpha)}$	→ $\|a\rangle = \sum_i a_i\|i\rangle$ 或 $\|a\rangle = \int a_f\|f\rangle \mathrm{d}f$
矢量分量	$a_\alpha = e^{(\alpha)} \cdot a$	→ $a_i = \langle i\|a\rangle$ 或 $a_f = \langle f\|a\rangle$

(3.1)

分量 $a_i = \langle i|a\rangle (i = 1,2,\cdots)$ 或 $a_f = \langle f|a\rangle (-\infty < f < \infty)$ 是一组无穷

多个数，它们就是状态 $|a\rangle$ 在力学量 F 的表象中的分量形式. 例如，如果 F 是坐标 x 或动量 p，则

$$\Psi(x,t\rangle = \langle x|\Psi,t\rangle, \qquad c(p,t\rangle = \langle p|\Psi,t\rangle \tag{3.2}$$

就是坐标表象和动量表象中的态函数(波函数).

(二)算符在 F 表象中的分量形式

设有算符 G，它按 §5.4.2(2.10)式定义：

$$G|a\rangle = |c\rangle. \tag{3.3}$$

为了进入物理量 F 的表象，用 F 的本征矢 $\langle i|$ 左乘上式两边：

$$\langle i|G|a\rangle = \langle i|c\rangle = c_i. \tag{3.4}$$

等式右边已经是 F 表象的分量形式，而左边还不是. 为了将左边也进入 F 表象，采用"插 1 法"，即利用完备性条件 §5.4.2(2.26)式，写出

$$\langle i|G|a\rangle = \langle i|G\mathbf{1}|a\rangle = \sum_j \langle i|G|j\rangle\langle j|a\rangle = \sum_j \langle i|G|j\rangle a_j.$$

代入(3.4)式得到

$$\sum_j \langle i|G|j\rangle a_j = c_i. \tag{3.5}$$

此式中的 $\langle i|G|j\rangle$ 是有两个指标 i 和 j 的一组数，它们可以排成矩阵(白体大写字母 G 表示希尔伯特空间中的算符，黑体大写 $\boldsymbol{G}$ 表示矩阵，两者不应混淆)：

$$\boldsymbol{G} = \begin{pmatrix} G_{11} & G_{12} & \cdots \\ G_{21} & G_{22} & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix}, \qquad G_{ij} = \langle i|G|j\rangle. \tag{3.6}$$

将态矢量的分量 a_j, c_i 排成一列矩阵：

$$\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix}, \qquad \boldsymbol{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \end{pmatrix}. \tag{3.7}$$

这样，就可以将(3.5)式写成矩阵乘法：

$$\sum_j G_{ij}a_j = c_i \rightarrow \boldsymbol{G}\boldsymbol{a} = \boldsymbol{c}, \tag{3.8}$$

明显写出就是

$$\begin{pmatrix} G_{11} & G_{12} & \cdots \\ G_{21} & G_{22} & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \end{pmatrix}. \tag{3.9}$$

(三)进入和退出表象

总结上述，得到将抽象希尔伯特空间中的态矢量 $|a\rangle$ 和算符 G 进入到一个具体表象，例如力学量 F 的表象，的方法. 首先假定 F 有分立谱，$F|i\rangle = f_i|i\rangle$.

用 F 的本征矢 $\langle i|$ 左乘态矢 $|a\rangle$，就得到 F 表象中的态矢：$a_i = \langle i|a\rangle$；而用 F 的本征矢 $\langle i|$ 和 $|j\rangle$ 分别从左右两边乘算符 G，就得到 G 矩阵的矩阵元：$G_{ij} = \langle i|G|j\rangle$.

反之，将 F 表象中的态矢 a_i 和力学量 G 在 F 表象中的矩阵元 G_{ij} 用狄拉克矢量写出：$a_i = \langle i|a\rangle$，$G_{ij} = \langle i|G|j\rangle$，分别去掉左边的本征矢 $\langle i|$ 和左右两边的本征矢 $\langle i|$ 和 $|j\rangle$，就还原到抽象希尔伯特空间中的态矢 $|a\rangle$ 和算符 G.

进入连续谱表象，例如坐标 X 的表象，的方法是类似的：

用 X 的本征矢 $\langle x|$ 左乘态矢 $|a\rangle$，就得到 X 表象中的态矢（波函数）$a_x \equiv \psi_a(x) = \langle x|a\rangle$；用 X 的本征矢 $\langle x|$ 和 $|x'\rangle$ 分别从左右两边乘算符 G，就得到矩阵元 $G_{xx'} = \langle x|G|x'\rangle$.

通常，连续谱中的算符用微分算符，而不用矩阵表示. 能够这样做的原因是由于我们讨论的力学量对态矢的作用是局域的而不是延展的. 考虑 G 对态矢 $|a\rangle$ 的作用 $G|a\rangle = |c\rangle$，如(3.3)式. 为了进入 X 表象，左乘 $\langle x|$，并插入一个完备性条件 $\int |x'\rangle\langle x'|\,\mathrm{d}x' = \mathbb{1}$，得到

$$\int \langle x|G|x'\rangle\langle x'|a\rangle \mathrm{d}x' = \langle x|c\rangle. \tag{3.10}$$

通常的力学量 G，例如势能、速度、动量，只将 x 点的 $\langle x|c\rangle$ 和 x 点附近一个无穷小邻域中的 $\langle x'|a\rangle$ 联系起来. 因而可以写

$$\langle x|G|x'\rangle = \hat{G}_x\delta(x-x'), \tag{3.11}$$

其中，$\hat{G}$ 的下标 x 表示它是作用在变量 x 上的算符. 代入(3.10)式，得到 $\int \hat{G}_x\delta(x-x')\langle x'|a\rangle \mathrm{d}x' = \langle x|c\rangle$，或 $\hat{G}_x\langle x|a\rangle = \langle x|c\rangle$，即

$$\hat{G}_x\psi_a(x) = \psi_c(x). \tag{3.12}$$

（四）用厄米矩阵表示力学量

表示力学量的算符是厄米算符，满足 §5.4.2(2.17)式，因而

$$\langle j|G|i\rangle^* = \langle i|G|j\rangle, \qquad \text{或 } G_{ij} = G_{ji}^*. \tag{3.13}$$

或写为

$$\boldsymbol{G} = \tilde{\boldsymbol{G}}^* = \boldsymbol{G}^\dagger, \tag{3.14}$$

这里波线 ~ 表示转置；† 表示厄米共轭，即转置复共轭. 满足条件(3.13)、(3.14)式的矩阵称为厄米自共轭矩阵，简称厄米矩阵.

因此，在 F 表象中，力学量用厄米矩阵表示，它按矩阵乘法规则(3.8)、(3.9)式作用在态矢量上.

对于连续谱表象，如 X 表象，表示力学量的厄米矩阵 $\langle x|G|x'\rangle$ 对态矢的作用如(3.10)式所示. 此时，厄米矩阵 $\langle x|G|x'\rangle$ 通过(3.11)式和厄米算符 $\hat{G}$ 相联系.

(五)本征值方程与本征表象

在 F 表象中,力学量 G 的本征值方程是

$$\boldsymbol{G}a^{(n)} = g_n a^{(n)}, \tag{3.15}$$

其中 $n = 1,2,\cdots$ 是标志不同本征值的标号,按(3.8)式写成分量形式就是

$$\sum_j G_{ij} a_j^{(n)} = g_n a_i^{(n)}. \tag{3.16}$$

注意区分指标 n 和 i,j;前者是 G 的本征值的标号,在本征值方程(3.16)中只取一个特定值;而后者是表象分量的记号,取 1 到∞的所有值.

现在假定 F 就是 G,即:考虑力学量 G 在它自身表象中的矩阵.在希尔伯特空间中写出 G 的本征方程 $G\mid m\rangle = g_m \mid m\rangle$,于是按(3.6)式有

$$G_{nm} = \langle n \mid G \mid m\rangle = \langle n \mid g_m \mid m\rangle = g_m\langle n \mid m\rangle = g_m\delta_{nm}, \tag{3.17}$$

排成矩阵就是

$$\boldsymbol{G} = \begin{pmatrix} g_1 & & 0 \\ & g_2 & \\ 0 & & \ddots \end{pmatrix}. \tag{3.18}$$

即:力学量在它自身表象中是对角矩阵,其对角元就是力学量的本征值.

(六)表象变换

假定已知状态 $\mid a\rangle$ 和力学量 Q 在 F 表象中的分量,要求它们在 G 表象中的分量.这称为表象变换.

设 $F\mid i\rangle = f_i \mid i\rangle$,$G\mid n\rangle = g_n \mid n\rangle$.按假定,已知 $a_i = \langle i\mid a\rangle$ 和 $Q_{ij} = \langle i\mid Q \mid j\rangle$,我们的问题是要求 a_n 和 Q_{nm}.

为此,利用 F 的本征矢的完备性:

$$\sum_i \mid i\rangle\langle i \mid = \mathbf{1}. \tag{3.19}$$

把它插进 a_n 和 Q_{nm} 的定义中:

$$\begin{aligned} a_n &= \langle n \mid a\rangle = \langle n \mid \mathbf{1} \mid a\rangle \\ &= \sum_i \langle n \mid i\rangle\langle i \mid a\rangle = \sum_i \langle n \mid i\rangle a_i, \end{aligned} \tag{3.20}$$

$$\begin{aligned} Q_{nm} &= \langle n \mid Q \mid m\rangle = \langle n \mid \mathbf{1}\, Q\, \mathbf{1} \mid m\rangle \\ &= \sum_{ij} \langle n \mid i\rangle\langle i \mid Q \mid j\rangle\langle j \mid m\rangle \\ &= \sum_{ij} \langle n \mid i\rangle Q_{ij}\langle j \mid m\rangle. \end{aligned} \tag{3.21}$$

这就是 $a_i \to a_n$,$Q_{ij} \to Q_{nm}$ 的变换公式,也就是从 F 表象到 G 表象的变换.我们看到,在这一变换中起关键作用的是 $\langle n \mid i\rangle = \langle i \mid n\rangle^*$.给 n 一个定值,则 $\langle i \mid n\rangle$ 是在 F 表象中 G 的本征矢 $\mid n\rangle$ 的分量形式;让 n 取不同值,得到 F 表象中 G 的不同

本征矢. 只要在 F 表象中解 G 的本征值方程,就可以得到它们.

将 $\langle n \mid i\rangle (n = 1,2,\cdots; i = 1,2,\cdots)$ 写成

$$S_{ni} = \langle n \mid i\rangle. \tag{3.22}$$

将它们排成矩阵,用符号 S 表示[为了书写方便,以下直接用白体而不是黑体大写字母表示矩阵]:

$$S = [S_{ni}] = \begin{pmatrix} S_{11} & S_{12} & \cdots \\ S_{21} & S_{22} & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix} = \begin{pmatrix} \langle n_1 \mid i_1\rangle & \langle n_1 \mid i_2\rangle & \cdots \\ \langle n_2 \mid i_1\rangle & \langle n_2 \mid i_2\rangle & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix}. \tag{3.23}$$

应注意,S_{ni} 的第一个脚标,即矩阵 S 的行号,是算符 G 本征矢的编号;而 S_{ni} 的第二个脚标,即矩阵 S 的列号,是算符 F 本征矢的编号,两者不能混淆. 将 S_{ni} 明显写出就是:$\langle n = 1 \mid i = 1\rangle, \langle n = 1 \mid i = 2\rangle, \cdots$. 为了书写简单,在(3.23)式右边采用了简化符号 $\langle n_1 \mid i_1\rangle$ 来表示 $\langle n = 1 \mid i = 1\rangle, \cdots$.

在(3.21)式中出现的 $\langle i \mid n\rangle$ 是 S 的厄米共轭(即转置复共轭)矩阵 $S^\dagger = \tilde{S}^*$ 的矩阵元

$$(S^\dagger)_{in} = \langle i \mid n\rangle. \tag{3.24}$$

证明如下.

取(3.23)式的转置

$$\tilde{S} = \begin{pmatrix} \langle n_1 \mid i_1\rangle & \langle n_2 \mid i_1\rangle & \cdots \\ \langle n_1 \mid i_2\rangle & \langle n_2 \mid i_2\rangle & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix}.$$

再取复共轭,就得到 S 的厄米共轭

$$S^\dagger = \tilde{S}^* = \begin{pmatrix} \langle i_1 \mid n_1\rangle & \langle i_1 \mid n_2\rangle & \cdots \\ \langle i_2 \mid n_1\rangle & \langle i_2 \mid n_2\rangle & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix}.$$

这就是(3.24)式. 【证毕】

这样,(3.20)、(3.21)式成为

$$a_n = \sum_i S_{ni} a_i, \qquad Q_{nm} = \sum_{ij} S_{ni} Q_{ij} (S^\dagger)_{jm}. \tag{3.25}$$

a_i 和 a_n 分别是态矢 $\mid a\rangle$ 在 F 和 G 表象中的分量形式,可以把它们排成两个一列矩阵,用 $a^{(F)}$、$a^{(G)}$ 表示:

$$a^{(F)}=\begin{pmatrix}a_1\\a_2\\\vdots\\a_i\\\vdots\end{pmatrix},\qquad a^{(G)}=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\\\vdots\end{pmatrix}. \tag{3.26}$$

类似地将 Q_{ij} 和 Q_{nm} 排成矩阵，用 $Q^{(F)}$、$Q^{(G)}$ 表示，则由 $F\to G$ 的表象变换公式(3.25)式可以合写成：

$$a^{(G)}=Sa^{(F)},\qquad Q^{(G)}=SQ^{(F)}S^{\dagger}. \tag{3.27}$$

需要再次强调的是，由 F 表象到 G 表象的变换矩阵 S 实际上是由 G 在 F 表象中的本征矢所生成. 由(3.22)(3.23)式可以看出，S 矩阵的第 n 行，就是 G 的第 n 个本征矢，所以在实际计算中，只需要在 F 表象中求解 $\hat{G}$ 的本征值方程，并将所得的本征矢按行排列，就得到变换矩阵 S.

【例1】 写出态矢量 $|\Psi,t\rangle$ 在坐标表象和动量表象之间的变换公式.

解 态矢量 $|\Psi,t\rangle$ 在坐标表象和动量表象之间的分量形式分别是

$$\Psi(x,t)=\langle x|\Psi,t\rangle\quad 和\quad c(p,t)=\langle p|\Psi,t\rangle, \tag{3.28}$$

利用完备性条件

$$\int|p\rangle\langle p|\,\mathrm{d}p=1,\qquad 和\qquad \int|x\rangle\langle x|\,\mathrm{d}x=1, \tag{3.29}$$

分别插入(3.28)的两式中，得到

$$\left.\begin{aligned}\Psi(x,t)&=\int\langle x|p\rangle\langle p|\Psi,t\rangle\,\mathrm{d}p=\int\langle x|p\rangle c(p,t)\,\mathrm{d}p,\\ c(p,t)&=\int\langle p|x\rangle\langle x|\Psi,t\rangle\,\mathrm{d}x=\int\langle p|x\rangle\Psi(x,t)\,\mathrm{d}x.\end{aligned}\right\} \tag{3.30}$$

这里的 $\langle x|p\rangle$ 是动量在 x 表象中的本征函数，已知为

$$\langle x|p\rangle=\langle p|x\rangle^{*}=(1/\sqrt{2\pi\hbar})\mathrm{e}^{(i/\hbar)px}, \tag{3.31}$$

代入(3.30)式得到

$$\left.\begin{aligned}\Psi(x,t)&=(1/\sqrt{2\pi\hbar})\int c(p,t)\mathrm{e}^{(i/\hbar)px}\,\mathrm{d}p,\\ c(p,t)&=(1/\sqrt{2\pi\hbar})\int\Psi(x,t)\mathrm{e}^{-(i/\hbar)px}\,\mathrm{d}x.\end{aligned}\right\} \tag{3.30'}$$

这就是 §5.4.1 的(1.3)、(1.4)式.

【例2】 求动量表象中的坐标算符和动量算符.

解 我们用大写的 X 和 P 表示希尔伯特空间中的算符. 利用 X 的本征值方程 $X|x''\rangle=x''|x''\rangle$，以及 $\langle x'|x''\rangle=\delta(x'-x'')$，$x''\delta(x'-x'')=x'\delta(x'-x'')$，得到

$$x_{x'x''}=\langle x'|X|x''\rangle=x'\delta(x'-x''). \tag{3.32}$$

和(3.11)式一致. 注意,上式中的 X 是希尔伯特空间中的抽象算符,而 x' 和 x'' 是它的本征值,是普通的数. 利用完备性条件(3.29)式,写出坐标在动量表象中的矩阵元:

$$\langle p \mid X \mid p' \rangle = \int \langle p \mid x' \rangle \langle x' \mid X \mid x'' \rangle \langle x'' \mid p' \rangle \mathrm{d}x' \mathrm{d}x''.$$

将(3.32)式代入,得到

$$\begin{aligned}\langle p \mid X \mid p' \rangle &= \int \langle p \mid x' \rangle x' \delta(x' - x'') \langle x'' \mid p' \rangle \mathrm{d}x' \mathrm{d}x'' \\ &= \frac{1}{2\pi\hbar} \int x' \mathrm{e}^{(-\mathrm{i}/\hbar)(p-p')x'} \mathrm{d}x' = \frac{1}{2\pi\hbar} \left(-\frac{\hbar}{\mathrm{i}} \right) \int \frac{\partial}{\partial p} \mathrm{e}^{(-\mathrm{i}/\hbar)(p-p')x'} \mathrm{d}x' \\ &= \frac{\mathrm{i}\hbar}{2\pi} \frac{\partial}{\partial p} \int \mathrm{e}^{-\mathrm{i}(p-p')(x'/\hbar)} \mathrm{d}(x'/\hbar) = \mathrm{i}\hbar \frac{\partial}{\partial p} \delta(p - p').\end{aligned}$$

这里用了 δ 函数的表达式[附录(Ⅰ.14)式].

动量算符在动量表象中的矩阵元可以利用本征值方程 $P \mid p' \rangle = p' \mid p \rangle$ 写出:

$$\langle p \mid P \mid p' \rangle = p\delta(p - p').$$

将以上二式和(3.11)式比较,得到动量表象中的坐标和动量算符:

$$\hat{x} = \mathrm{i}\hbar \partial/\partial p, \qquad \hat{p} = p \quad (\text{动量表象}). \tag{3.33}$$

利用上式可以直接在动量表象中写出 §5.4.1 例 1 中的薛定谔方程,得到那一节的(1.7)式.

【例 3】 设算符 G 有分立谱,已知它在 x 表象中的本征函数 $\varphi_n(x)$,并设已知 x 表象中的任意态函数 $\Psi(x,t)$ 和算符 $\hat{F}_{(x)}$,求 G 表象中的态矢量 $\Psi_n(t)$ 和矩阵 F_{nm}.

解 已知 $\langle x \mid \Psi, t \rangle = \Psi(x,t)$, $\langle x \mid F \mid x' \rangle = \hat{F}_{(x)} \delta(x - x')$. 由此得

$$\begin{aligned}\Psi_n(t) &= \langle n \mid \Psi, t \rangle = \int \langle n \mid x \rangle \langle x \mid \Psi, t \rangle \mathrm{d}x \\ &= \int \langle n \mid x \rangle \Psi(x,t) \mathrm{d}x, \\ F_{nm} &= \langle n \mid F \mid m \rangle = \int \langle n \mid x \rangle \langle x \mid F \mid x' \rangle \langle x' \mid m \rangle \mathrm{d}x \mathrm{d}x' \\ &= \int \langle n \mid x \rangle \hat{F}_{(x)} \delta(x - x') \langle x' \mid m \rangle \mathrm{d}x \mathrm{d}x' \\ &= \int \langle n \mid x \rangle \hat{F}_{(x)} \langle x \mid m \rangle \mathrm{d}x.\end{aligned}$$

再将

$$\langle x \mid n \rangle = \langle n \mid x \rangle^* = \varphi_n(x)$$

代入以上二式,得

$$\left.\begin{aligned} \Psi_n(t) &= \int \varphi_n^*(x)\Psi(x,t)\mathrm{d}x, \\ F_{nm} &= \int \varphi_n^*(x)\ \hat{F}_{(x)}\varphi_m(x)\mathrm{d}x. \end{aligned}\right\} \tag{3.34}$$

这就是由 x 表象到 G 表象的变换公式.

(七)表象变换矩阵的幺正性

在§5.3.3中已经证明,为使空间变换不改变态矢的归一化(概率守恒),空间变换算符应为幺正算符,见那一节的(3.11)式.根据同样的理由,作为希尔伯特空间坐标变换的表象变换也不应该改变态矢的归一化,因而表象变换矩阵应该是幺正矩阵.

$$S^\dagger = S^{-1}, \qquad S^\dagger S = SS^\dagger = 1. \tag{3.35}$$

明显写出就是

$$\sum_i S_{ni}S_{im}^\dagger = \delta_{nm}, \qquad \sum_n S_{in}^\dagger S_{nj} = \delta_{ij}. \tag{3.36}$$

或写为

$$\sum_i S_{ni}S_{mi}^* = \delta_{nm}, \qquad \sum_n S_{ni}^* S_{nj} = \delta_{ij}. \tag{3.37}$$

利用(3.22)(3.24)式很容易证明这一结论.

我们来证明,由于表象变换是幺正变换,在表象变换下量子力学的基本公式不变,物理观测结果不变.

证　上述定理的证明分三步完成.

(1)表象变换不改变算符间的对易关系

设在 F 表象有 $[A^F,B^F] = \mathrm{i}C^F$,分别左乘以 S 右乘以 $S^\dagger$,就有

$$[SA^FS^\dagger, SB^FS^\dagger] = S[A^F,B^F]S^\dagger = S\mathrm{i}C^FS^\dagger,$$

即

$$[A^F,B^F] = \mathrm{i}C^F \quad \Rightarrow \quad [A^G,B^G] = \mathrm{i}C^G.$$

(2)表象变换不改变力学量的本征值

设在 F 表象力学量 L 的本征值方程为 $L^Fa^F = l^Fa^F$,在 G 表象为 $L^Ga^G = l^Ga^G$,现证明:$l^F = l^G$.由 $L^G = SL^FS^\dagger, a^G = Sa^F$,知 $SL^FS^\dagger Sa^F = l^GSa^F$.上式左边 $= SL^Fa^F = Sl^Fa^F = l^FSa^F$,和右边 l^GSa^F 比较可见 $l^G = l^F$.

(3)表象变换不改变矢量的内积

$\langle a \mid b\rangle$ 的 F 表象为 $\sum_i a_i^* b_i$,G 表象为 $\sum_n a_n^* b_n$.现证明它们是相等的.

$$\sum_n a_n^* b_n = \sum_n \Big(\sum_i S_{ni}a_i\Big)^* \Big(\sum_j S_{nj}b_j\Big) = \sum_n \sum_{ij} a_i^* S_{ni}^* S_{nj}b_j$$

$$= \sum_{ij} a_i^* \delta_{ij} b_j = \sum_{ij} a_i^* b_i.$$

这就证明了,由于表象变换是幺正变换,不改变算符的对易关系以及算符的本征值和矢量的内积.因此,测量力学量所能取的值和在一个状态中测量力学量的平均值及得到不同值的概率均不因表象而变. 【证毕】

§5.4.4 算符对易关系的确定

在§5.4.2中看到,量子力学的基本原理可以不用具体表象,直接在抽象的希尔伯特空间中表出.在抽象的希尔伯特空间中,力学量的算符没有具体表达式,它们的根本性质体现在它们之间的对易关系上.那么,怎样得到算符的对易关系呢?在§5.2.3(3.5)式中得到坐标和动量算符之间的对易关系是利用了它们在坐标表象中的具体形式.然而,算符的对易关系不应该依赖于具体表象.于是出现一个问题:如何不依赖于具体表象而直接确定算符的对易关系?本节就来讨论这一问题.

(一)物理量作为空间变换的生成元 奈特定理

我们将具体讨论坐标-动量算符的对易关系和角动量算符各个分量之间的对易关系.为此,先要回答一个问题:什么是坐标、动量和角动量?显然,坐标是描述空间位置的变量.那么,什么是动量呢?简单地回答:动量是质量乘速度$\boldsymbol{p}=m\boldsymbol{v}$是不对的,因为它只适用于低速运动的经典粒子.高速运动粒子的动量$\boldsymbol{p}=m\boldsymbol{v}/\sqrt{1-(v/c)^2}$[§2.3.1(1.4)式]和体积$V$中电磁场的动量$\boldsymbol{p}=\int_V \boldsymbol{g}\mathrm{d}V=\int_V \varepsilon_0 \boldsymbol{E}\times\boldsymbol{B}\mathrm{d}V$[§2.4.3(3.22)式]各有不同的表达式.因此,动量不能用一个具体表达式来定义,而应该从它的物理意义来定义.动量的基本物理意义在于它的守恒,即:它是空间平移不变时的守恒量(见§1.2.3和§5.3.3).例如,设一个由相互作用的带电粒子和电磁场组成的系统具有空间平移不变性,则尽管带电粒子和电磁场的动量的表达式很不相同,它们的和,即系统的总动量,保持守恒,见§2.4.3(3.12)式.因此,动量的基本物理意义可以表述为:动量是和空间平移变换对应的力学量,具有空间平移不变性的系统动量守恒.类似地,角动量的基本物理意义是:角动量是和空间转动变换对应的力学量,具有空间转动不变性的系统角动量守恒.

在§5.3.3中曾经利用坐标表象讨论过对称性与守恒律的关系.现在再在抽象希尔伯特空间中作普遍的讨论.

平移和转动都依赖于一个连续参量δ.对于平移,这个参量是移动的距离;而对于转动,它是转动的角度.在平移或转动δ下,态矢$|a\rangle$变为

$$|a\rangle\rightarrow|a'\rangle=O_\delta|a\rangle, \tag{4.1}$$

其中O_δ是和这一空间变换对应的希尔伯特空间算符.相应地,希尔伯特空间的任

意算符 A 变为

$$A \to A' = O_\delta A\, O_\delta^{-1}. \tag{4.2}$$

由于空间变换不改变态矢的归一化，所以 O_δ 是幺正算符，$O_\delta^{-1} = O_\delta^\dagger$. 当 $\boldsymbol{\delta}$ 很小时，可以将 O_δ 对它作泰勒展开到一级. 我们来证明，如果将这一展开写成（加进普朗克常数 $\hbar$ 是为了以后选用单位的方便）：

$$O_\delta = 1 - \frac{\mathrm{i}}{\hbar}\boldsymbol{F}\cdot\boldsymbol{\delta} = 1 - \frac{\mathrm{i}}{\hbar}\sum_{j=1}^{3} F_j\delta_j, \tag{4.3}$$

则其中 $\boldsymbol{F}$ 的三个分量 F_j 都是希尔伯特空间的厄米算符. 证明如下：

证

$$O_\delta = 1 - (\mathrm{i}/\hbar)\boldsymbol{F}\cdot\boldsymbol{\delta}, \qquad O_\delta^\dagger = 1 + (\mathrm{i}/\hbar)\boldsymbol{F}^\dagger\cdot\boldsymbol{\delta},$$

$$O_\delta O_\delta^\dagger = [1 - (\mathrm{i}/\hbar)\boldsymbol{F}\cdot\boldsymbol{\delta}][1 + (\mathrm{i}/\hbar)\boldsymbol{F}^\dagger\cdot\boldsymbol{\delta}]$$

$$= 1 + (\mathrm{i}/\hbar)(\boldsymbol{F}^\dagger - \boldsymbol{F})\cdot\boldsymbol{\delta} + o(\boldsymbol{\delta}),$$

其中 $o(\boldsymbol{\delta})$ 表示比 $(\boldsymbol{\delta})$ 更高级的小量. 由于 O_δ 是幺正算符，上式的左边等于 1. 略去二级小量后，为使上式右边等于 1 需要 $\boldsymbol{F} = \boldsymbol{F}^\dagger$. **【证毕】**

我们看到，每一个依赖于连续参量的空间变换 O_δ 都对应一个厄米算符 $\boldsymbol{F}$［见(4.3)式］. 这一个厄米算符 $\boldsymbol{F}$ 称为空间变换 O_δ 的生成元. 它就是和这一空间变换对应的力学量的算符. 例如，和平移变换对应的力学量是动量，因而平移变换的生成元就是动量算符；和转动变换对应的力学量是角动量，因而转动变换的生成元就是角动量算符.

如果系统具有某一空间变换 O 的不变性，则哈密顿算符在这一空间变换下不变. 根据(4.2)式知 $O_\delta H O_\delta^\dagger = H$. 右乘以 O_δ 得到

$$O_\delta H = H O_\delta, \qquad [H, O_\delta] = 0. \tag{4.4}$$

即：哈密顿算符和这一空间变换的算符可对易. 由此根据(4.3)式又知：哈密顿算符和这一空间变换的生成元可对易：$[H, \boldsymbol{F}] = 0$. 由于空间变换的生成元不显含时间，它和哈密顿算符可对易就表明它所代表的力学量守恒. 于是得到：

奈特定理　如果系统在某一空间变换下不变，则这一空间变换的生成元所对应的力学量守恒.

我们熟知的平移不变动量守恒和转动不变角动量守恒，都是奈特定理的特例.

（二）坐标与动量的对易关系

我们来根据动量是空间平移的生成元这一性质，推导坐标和动量的对易关系. 为了书写简单，只考虑一维. 不难将结果推广到三维.

用 $\hat{x}$ 表示坐标算符，$|x\rangle$ 是它的本征矢：

$$\hat{x}|x\rangle = x|x\rangle. \tag{4.5}$$

用 D_δ 表示使系统平移 δ 的算符❶:

$$D_\delta | x\rangle = | x+\delta\rangle. \tag{4.6}$$

将 D_δ 和 $\hat{x}$ 分别作用到(4.5)和(4.6)式上,

$$D_\delta\hat{x}|x\rangle = xD_\delta|x\rangle = x|x+\delta\rangle,$$
$$\hat{x}D_\delta|x\rangle = \hat{x}|x+\delta\rangle = (x+\delta)|x+\delta\rangle.$$

相减得到

$$(\hat{x}D_\delta - D_\delta\hat{x})|x\rangle = [\hat{x},D_\delta]|x\rangle = \delta\cdot|x+\delta\rangle.$$

将 D_δ 和 $|x+\delta\rangle$ 都对 δ 展开并在等式左右两边保留到一级,利用(4.3)式以及动量是空间平移的生成元的性质,得到

$$[\hat{x},1-(\mathrm{i}/\hbar)\hat{p}\delta]|x\rangle = \delta|x\rangle, \qquad -(\mathrm{i}\delta/\hbar)[\hat{x},\hat{p}]|x\rangle = \delta|x\rangle.$$

由于坐标本征矢 $|x\rangle$ 的完备性,算符作用到它上面相等就表明算符本身相等,因而由上式得

$$[\hat{x},\hat{p}] = \mathrm{i}\hbar. \tag{4.7}$$

这就是坐标算符和动量算符的对易关系.它是从"动量是空间平移的生成元"这一根本性质出发得到的希尔伯特空间中的对易关系.根据坐标表象中的算符得到的对易关系 §5.2.3(3.5)式是它的一个具体实现.

(三)角动量分量之间的对易关系

用 $T_{\delta\varphi}$ 表示和转动 $\delta\boldsymbol{\varphi}$ 对应的希尔伯特空间的算符.角动量 $\boldsymbol{J}$ 是它的生成元.按(4.3)式

$$T_{\delta\varphi} = 1 - \frac{\mathrm{i}}{\hbar}\boldsymbol{J}\cdot\delta\boldsymbol{\varphi} = 1 - \frac{\mathrm{i}}{\hbar}\sum_{i=1}^{3}J_i(\delta\varphi)_i. \tag{4.8}$$

转动使得位矢产生的变化所对应的算符是

$$R_{\delta\varphi} = T_{\delta\varphi} - 1 = -\frac{i}{\hbar}\boldsymbol{J}\cdot\delta\boldsymbol{\varphi} = -\frac{i}{\hbar}\sum_{i=1}^{3}J_i(\delta\varphi)_i. \tag{4.9}$$

在 §1.5.1 中证明了,由于绕不同轴的转动不可对易,绕两个不同轴转动 $\delta\boldsymbol{\varphi}_1$ 和 $\delta\boldsymbol{\varphi}_2$ 所引起的变化的算符 R 之间有对易关系 §1.5.1(1.16)式

$$[R_{\delta\varphi_1},R_{\delta\varphi_2}] = R_{\delta\varphi_1\times\delta\varphi_2}. \tag{4.10}$$

将(4.9)式代入,得到

$$(-\mathrm{i}/\hbar)^2[\boldsymbol{J}\cdot\delta\boldsymbol{\varphi}_1,\boldsymbol{J}\cdot\delta\boldsymbol{\varphi}_2] = -(\mathrm{i}/\hbar)\boldsymbol{J}\cdot(\delta\boldsymbol{\varphi}_1\times\delta\boldsymbol{\varphi}_2).$$

令 $(\delta\boldsymbol{\varphi}_1,\delta\boldsymbol{\varphi}_2)$ 等于 $(\boldsymbol{e}_x,\boldsymbol{e}_y),(\boldsymbol{e}_y,\boldsymbol{e}_z),(\boldsymbol{e}_z,\boldsymbol{e}_x)$,得到

---

❶ 将(4.6)式和 §5.3.3(3.13)式比较可见,这里是 $x+\delta$,而那里是 $r-\delta r$,两者相差符号.原因是,这里的 $|x\rangle$ 是(坐标算符的)本征矢,而那里的 r 是(坐标表象的)表象变量.参看 §5.3.3 图1.当该图中的波峰的宽度为零时,就是坐标的本征矢 $|x\rangle$,波峰的位置就是 x.波峰向前平移使 $x\to x+\delta$.

$$[J_x,J_y]=\mathrm{i}\hbar J_z,\quad [J_y,J_z]=\mathrm{i}\hbar J_x,\quad [J_z,J_x]=\mathrm{i}\hbar J_y. \tag{4.11}$$

这就是角动量的各个分量之间的对易关系.

在§5.2.2例3中用$\boldsymbol{L}=\boldsymbol{r}\times\boldsymbol{p}$定义的"轨道角动量"是普遍角动量$\boldsymbol{J}$的一个特例,因而也有(4.11)式型的对易关系,见§5.2.3(3.23)式.特别要注意的是,普遍意义下的角动量(转动变换的生成元)不仅包含轨道角动量$\boldsymbol{L}=\boldsymbol{r}\times\boldsymbol{p}$,而且包含另一种角动量——自旋角动量$\boldsymbol{s}$,详见§5.4.6.自旋角动量不能象轨道角动量那样写成微分算符,而只能写成矩阵.只有对易关系(4.11)式才是一切角动量都具有的普遍性质.

§5.4.5　谐振子的占有数表象

本章前几节说明了量子力学理论可以在希尔伯特空间中以抽象的形式建立起来,然后再根据所研究的问题的特点,选用适当的表象来求解.在前几章和下一篇里,我们讨论的大多数问题都采用坐标表象.由于坐标表象有连续谱,所以这些问题都归结为求解微分方程.在这一节和下一节里,我们举两个分立谱表象的例子.在此情况下,问题归结为求矩阵的本征值和矩阵元.这两个例子不仅有方法论上的价值,而且所得结果也有重要的实际意义.

第一个例子是线性谐振子,它的经典哈密顿函数是§1.6.3(3.20)式:

$$H=p^2/(2m)+m\omega^2x^2/2. \tag{5.1}$$

在量子力学里,谐振子的哈密顿算符也具有同一形式,只要将x和p由c数换成q数,就将经典谐振子量子化为量子谐振子.这一过渡可以通过按照§5.4.4(4.7)式规定x和p的对易关系而实现(在本节里,我们直接用x,p表示希尔伯特空间中的算符):

$$[x,p]=\mathrm{i}\hbar. \tag{5.2}$$

注意,我们没有采用具体表象.因而x和p都是希尔伯特空间中的抽象算符.它们不能写成具体形式,所能够知道的只是对易关系(5.2)式.我们就是要从(5.2)式和(5.1)式出发,求解谐振子的定态问题.

首先设法在(5.2)式中作变量变换,使它成为更易于求解的形式.用下式定义一个算符a

$$a=\alpha x+\mathrm{i}\beta p, \tag{5.3}$$

其中α和β是两个待定的实系数.由于在定义式中包含了虚因子i,所以a不是厄米算符,它的厄米共轭是

$$a^\dagger=\alpha x^\dagger+\mathrm{i}^*\beta p^\dagger=\alpha x-\mathrm{i}\beta p. \tag{5.4}$$

这里利用了算符x和p的厄米性:$x^\dagger=x,p^\dagger=p$.算符a和$a^\dagger$的对易关系是:

$$[a,a^\dagger] = [(\alpha x + i\beta p),(\alpha x - i\beta p)] = i\beta\alpha[p,x] - i\alpha\beta[x,p] = 2\alpha\beta\hbar. \quad (5.5)$$

可以通过调整系数 α 和 β 使得

$$[a,a^\dagger] = 1. \quad (5.6)$$

为此只需要

$$\alpha\beta = 1/(2\hbar). \quad (5.7)$$

这一等式没有完全决定 α 和 β，还可以对它们加上另一个条件，这一条件将在后面再规定.

现在来利用对易关系(5.6)式求哈密顿算符(5.1)式的本征值，即求谐振子的定态能量. 这通过以下几步来完成：

①首先注意 a 和 $a^\dagger$ 都不是厄米的，但是由于它们相互厄米共轭，所以它们的乘积

$$N = a^\dagger a \quad (5.8)$$

是厄米的.

证

$$N^\dagger = (a^\dagger a)^\dagger = a^\dagger (a^\dagger)^\dagger = a^\dagger a = N.$$

这里利用了算符(或矩阵)乘积的厄米共轭等于它们各自的厄米共轭颠倒次序相乘. **【证毕】**

②由于 N 是厄米算符，所以它的本征值是实数. 现在来进一步证明 N 的本征值 n 恒正：$n \geqslant 0$.

证 设 $N|n\rangle = n|n\rangle$ 为 N 的本征值方程. 用 $\langle n|$ 左乘上式，得到

$$\langle n|N|n\rangle = \langle n|a^\dagger a|n\rangle = \langle n|n|n\rangle = n\langle n|n\rangle = n. \quad (5.9)$$

算符 a 作用到矢量 $|n\rangle$ 上，得到另一矢量，用 $|b\rangle$ 表示它：

$$a|n\rangle = |b\rangle. \quad (5.10)$$

取厄米共轭，得 $\langle b| = \langle n|a^\dagger$，因此，(5.9)式可以写成

$$n = \langle b|b\rangle. \quad (5.11)$$

这是矢量 $|b\rangle$ 的模的平方，恒正，所以 $n \geqslant 0$. **【证毕】**

③在以上两步中没有用到对易关系，所以是对于任意算符 a 和它的共轭算符 $a^\dagger$ 都适用. 现在来进行最关键的一步，利用对易关系(5.6)式证明：如果 $|n\rangle$ 是 N 的本征矢，对应的本征值是 n，且 $a|n\rangle \neq 0$，则 $a|n\rangle$ 也是 N 的本征矢，对应的本征值为 $n-1$.

证 将 N 作用到 $a|n\rangle$ 上，并利用对易关系(5.6)式：

$$\begin{aligned} Na|n\rangle &= a^\dagger aa|n\rangle = (aa^\dagger - 1)a|n\rangle \\ &= aa^\dagger a|n\rangle - a|n\rangle = aN|n\rangle - a|n\rangle \\ &= an|n\rangle - a|n\rangle = (n-1)a|n\rangle. \end{aligned}$$

上式表明，只要 $a|n\rangle \neq 0$，则 $a|n\rangle$ 为 N 的本征矢，对应的本征值为 $n-1$.

【证毕】

用 $|n-1\rangle$ 表示 N 的本征值为 $n-1$ 的归一化本征矢，则

$$\text{当 } a|n\rangle \neq 0 \text{ 时，} \quad a|n\rangle = c|n-1\rangle, \tag{5.12}$$

c 为归一化常数.

④利用对易关系(5.6)式还可以证明：只要 $a^\dagger|n\rangle \neq 0$，$a^\dagger|n\rangle$ 也是 N 的本征矢，对应本征值为 $n+1$.

证

$$\begin{aligned} Na^\dagger|n\rangle &= a^\dagger a a^\dagger|n\rangle = a^\dagger(a^\dagger a+1)|n\rangle \\ &= a^\dagger N|n\rangle + a^\dagger|n\rangle = (n+1)a^\dagger|n\rangle. \end{aligned} \tag{5.13}$$

【证毕】

⑤由以上③、④两点知道，N 有一系列本征值，每两个相邻本征值之间相差 1，因而 N 的本征值谱可写为

$$\cdots, n+2, n+1, n, n-1, n-2, \cdots. \tag{5.14}$$

但是，如果这一个序列可以无限地向右方延伸下去，就会和第②条论断："N 的本征值恒正"相矛盾. 因此，序列(5.14)必须在右方某处，例如在 $n-n'$ 处截断. 为此就要求，对于本征矢 $|n-n'\rangle$，第③点中所述定理的条件 $a|n-n'\rangle \neq 0$ 不再成立，即

$$a|n-n'\rangle = 0 \quad (n' \text{ 为某一正整数或零}). \tag{5.15}$$

取共轭 $\langle n-n'|a^\dagger = 0$，和(5.15)式相乘

$$\begin{aligned} &\langle n-n'|a^\dagger a|n-n'\rangle = \langle n-n'|N|n-n'\rangle \\ &= (n-n')\langle n-n'|n-n'\rangle = n-n' = 0. \end{aligned}$$

因而，$n = n'$，其中 n' 为正整数或零：

$$n = 0, 1, 2, \cdots. \tag{5.16}$$

这就是 N 的本征值谱.

⑥通过这样几个步骤，就完成了求解对易关系(5.6)式的问题，求出了 $a^\dagger a$ 的本征值谱. 剩下的问题是将谐振子的哈密顿(5.1)式用算符 $a^\dagger, a$ 表出.

由(5.3)、(5.4)式解出

$$x = (a+a^\dagger)/(2\alpha), \quad p = (a-a^\dagger)/(2i\beta). \tag{5.17}$$

代入(5.1)式并化简. 在计算时要注意 a 和 $a^\dagger$ 是 q 数，不可对易，因而

$$\begin{aligned} (a+a^\dagger)^2 &= a^2 + a^{\dagger 2} + aa^\dagger + a^\dagger a, \\ (a-a^\dagger)^2 &= a^2 + a^{\dagger 2} - aa^\dagger - a^\dagger a. \end{aligned}$$

这样就得到

$$H = \left(\frac{m\omega^2}{8\alpha^2} - \frac{1}{8m\beta^2}\right)(a^2 + a^{\dagger 2}) + \left(\frac{m\omega^2}{8\alpha^2} + \frac{1}{8m\beta^2}\right)(aa^\dagger + a^\dagger a). \tag{5.18}$$

前面，对于 α 和 β 已经规定了一个条件(5.7)式，这里再补充一个条件使上式中非厄米的项 a^2 和 $a^{\dagger 2}$ 的系数为零. 这一条件是

$$m\omega^2/(8\alpha^2) = 1/(8m\beta^2). \tag{5.19}$$

和(5.7)式一道，解出

$$\alpha = \sqrt{m\omega/2\hbar}, \qquad \beta = 1/\sqrt{2m\hbar\omega}. \tag{5.20}$$

于是(5.18)式成为

$$H = (aa^\dagger + a^\dagger a)\hbar\omega/2. \tag{5.21}$$

利用对易关系(5.6)式，又可写上式为

$$H = (a^\dagger a + 1/2)\hbar\omega. \tag{5.22}$$

利用(5.16)式得到 H 的本征值：

$$E_n = (n + 1/2)\hbar\omega \quad (n = 0,1,2,\cdots). \tag{5.23}$$

这就是谐振子的能谱. 我们是利用对易关系(5.2)式，在抽象希尔伯特空间中，不用任何表象，求出了这一能谱.

(5.23)式是一个非常重要的结果. 它表明：

①谐振子的能级是等间距的，相邻能级之间相差 $\hbar\omega$.

②谐振子的基态能量，即零点振动能量是 $\hbar\omega/2$. 正如前面多次指出的，零点能的存在是波粒二象性的体现，也可以说是不确定关系的后果.

由(5.23)式知道，如果以基态能量作为计算能量的起点，则谐振子的定态能量是 $\hbar\omega$ 的整数倍. 如果谐振子和外界交换能量，它只能由一个定态跃迁到另一个定态，因而所交换的能量只能是 $\hbar\omega$ 的整数倍. 在第一章中曾经指出，交换能量时有最小单元 $\hbar\omega$ 存在，可以理解为是存在一种基元粒子，它的能量是

$$\varepsilon = \hbar\omega. \tag{5.24}$$

于是，谐振子处于第 n 个激发态，即能量为 $E_n - E_0 = n\hbar\omega$ 的状态，可以认为是存在 n 个这样的元粒子，占据在 $\varepsilon = \hbar\omega$ 的能级上，如图 1. 因此，n 被称为占有数，或粒子数. $N = a^\dagger a$ 是占有数算符，或粒子数算符.

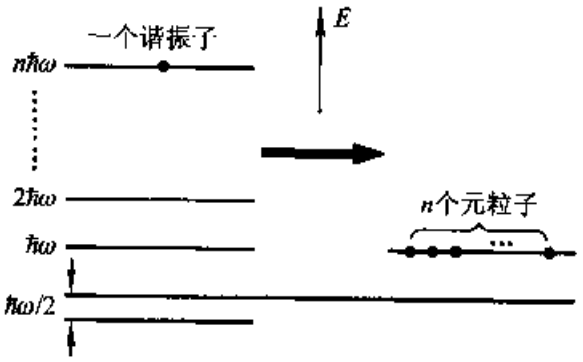


图 1　谐振子的激发态与占有数

由(5.12)、(5.13)式知，a 和 $a^\dagger$ 的作用分别使粒子数减少 1 和增加 1，因而被称为湮灭算符和产生算符.

最后，让我们写出 x 和 p 在粒子数表象（即谐振子的能量表象）中的矩阵元. 为此，先求 a 和 $a^\dagger$ 在粒子数表象中的矩阵元.

取(5.12)式左右两边的厄米共轭，再和(5.12)式相乘. 左右两边分别是

$$\langle n \mid a^\dagger a \mid n\rangle = \langle n \mid N \mid n\rangle = n,$$

和

$$\langle n-1 \mid c^* c \mid n-1\rangle = |c|^2 \langle n-1 \mid n-1\rangle = |c|^2.$$

因而有 $|c|^2 = n$. 适当选择相因子可以使 c 为实数，因而 $c = \sqrt{n}$，

$$\text{当 } a \mid n\rangle \neq 0 \text{ 时} \qquad a \mid n\rangle = \sqrt{n} \mid n-1\rangle. \tag{5.25}$$

为了得到 N 表象中 a 和 $a^\dagger$ 的矩阵元，用 $\langle n' \mid$ 乘(5.25)式：

$$\langle n' \mid a \mid n\rangle = a_{n'n} = \sqrt{n}\,\langle n' \mid n-1\rangle = \sqrt{n}\,\delta_{n',n-1}. \tag{5.26}$$

这表明，矩阵 a 的不为零的矩阵元是

$$a_{n-1,n} = \sqrt{n}. \tag{5.27}$$

再取(5.26)式的厄米共轭：

$$\langle n \mid a^\dagger \mid n'\rangle = \sqrt{n}\delta_{n',n-1},$$

在此式中将符号 n 和 n' 对换 $n \rightleftharpoons n'$，注意 $\delta_{n,n'-1}$ 表示 $n = n'-1, n' = n+1$，因而 $\delta_{n,n'-1} = \delta_{n',n+1}$

$$\sqrt{n'}\delta_{n,n'-1} = \sqrt{n+1}\delta_{n',n+1},$$

就得到

$$a^\dagger_{n'n} = \langle n' \mid a^\dagger \mid n\rangle = \sqrt{n+1}\delta_{n',n+1}, \tag{5.28}$$

因而，$a^\dagger$ 的不为零的矩阵元是

$$a^\dagger_{n+1,n} = \sqrt{n+1}. \tag{5.29}$$

利用(5.27)，(5.29)式，可以将 a 和 $a^\dagger$ 排成矩阵：

$$a = \begin{pmatrix} 0 & \sqrt{1} & \cdots & \cdots \\ 0 & 0 & \sqrt{2} & \cdots \\ \cdots & 0 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}, \quad a^\dagger = \begin{pmatrix} 0 & 0 & \cdots & \cdots \\ \sqrt{1} & 0 & 0 & \cdots \\ \cdots & \sqrt{2} & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}. \tag{5.30}$$

将(5.20)式代入(5.17)式，得到

$$x = \sqrt{\hbar/(2m\omega)}(a^\dagger + a), \qquad p = \mathrm{i}\sqrt{m\hbar\omega/2}(a^\dagger - a). \tag{5.31}$$

由(5.30)(5.31)式可以写出 x 和 p 在粒子数表象中的矩阵：

$$x = \sqrt{\frac{\hbar}{2m\omega}}\begin{pmatrix} 0 & \sqrt{1} & 0 & \cdots \\ \sqrt{1} & 0 & \sqrt{2} & \cdots \\ 0 & \sqrt{2} & 0 & \cdots \\ 0 & \cdots & \cdots & \cdots \end{pmatrix}, \tag{5.32}$$

$$p = \mathrm{i}\sqrt{\frac{m\hbar\omega}{2}}\begin{pmatrix} 0 & -\sqrt{1} & 0 & \cdots \\ \sqrt{1} & 0 & -\sqrt{2} & \cdots \\ 0 & \sqrt{2} & 0 & \cdots \\ 0 & \cdots & \cdots & \cdots \end{pmatrix}. \tag{5.33}$$

通过这一例子看到,虽然在希尔伯特空间里的算符不能写出明显形式,但是,只要有了对易关系,就可以求解本征值方程.这表明,正是对易关系反映了算符的根本性质.

§5.4.6 角动量的矩阵表示

作为在抽象希尔伯特空间中求解本征值方程的又一个例子,在这一节里从角动量的对易关系出发求解角动量的本征值方程.

应该指出的是,这里所说的角动量是指普遍意义下的角动量,即作为转动变换生成元的角动量,而不限于在§5.2.2中讨论过的轨道角动量.因而作为出发点的不是角动量的算符形式,而是它们的对易关系.若用 J 表示一般的角动量,则基本的对易关系是§5.4.4(4.11)式:

$$[J_x, J_y] = \mathrm{i}\hbar J_z, \quad [J_y, J_z] = \mathrm{i}\hbar J_x, \quad [J_z, J_x] = \mathrm{i}\hbar J_y. \tag{6.1}$$

(一)角动量的本征值

在§5.2.2例3中求出了轨道角动量 $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$ 的本征值和本征函数.现在讨论的是更普遍意义下的角动量 $\boldsymbol{J}$,因此只能从对易关系(6.1)式出发来求解它的本征值问题.

首先,将 J_x, J_y 线性组合,得到

$$J_+ = J_x + \mathrm{i}J_y, \qquad J_- = J_x - \mathrm{i}J_y. \tag{6.2}$$

这两个算符都不是厄米的,而是相互厄米共轭:

$$J_- = (J_+)^\dagger. \tag{6.3}$$

容易证明(见习题7)

$$[J_+, J_-] = 2\hbar J_z, \quad [J_z, J_+] = \hbar J_+, \quad [J_z, J_-] = -\hbar J_-. \tag{6.4}$$

可以用它们将 $\boldsymbol{J}^2$ 写成

$$\boldsymbol{J}^2 = J_- J_+ + J_z^2 + \hbar J_z. \tag{6.5}$$

证

$$\begin{aligned} J_- J_+ &= (J_x - \mathrm{i}J_y)(J_x + \mathrm{i}J_y) = J_x^2 + J_y^2 + \mathrm{i}(J_x J_y - J_y J_x) \\ &= J_x^2 + J_y^2 + J_z^2 - J_z^2 + \mathrm{i}(\mathrm{i}\hbar J_z) = \boldsymbol{J}^2 - J_z^2 - \hbar J_z. \end{aligned}$$

【证毕】

另外，我们还知道(见习题7)，$[\boldsymbol{J}^2, J_z] = 0$，即 $\boldsymbol{J}^2$ 和 J_z 有共同本征矢量组. 用 $|\lambda, m\rangle$ 表示这一矢量组中的一个矢量. 待求解的本征值方程是(写进常数 $\hbar$ 和 $\hbar^2$ 是为了单位制的方便)

$$\boldsymbol{J}^2 |\lambda, m\rangle = \lambda\hbar^2 |\lambda, m\rangle, \tag{6.6}$$

$$J_z |\lambda, m\rangle = m\hbar |\lambda, m\rangle. \tag{6.7}$$

首先在 $\boldsymbol{J}^2$ 有定值的条件下确定 J_z 的本征值. 这归结为在 λ 有定值的条件下确定 m 的取值. 解法的基本思路和上一节中解谐振子问题类似，即首先定出本征值的上、下限，再定出相邻两本征值的差，从而确定整个本征值谱. 步骤如下：

①注意到 z 轴的两个方向在物理上是等价的，所以对应于 $m = +|m|$，一定存在 $m = -|m|$. 另一方面，由于是在 $\boldsymbol{J}^2$ 有定值的条件下求 J_z 的本征值，因而 $|m|$ 必有最大值，若 J 是 $|m|$ 的最大值，则 m 的取值范围是

$$-J \leqslant m \leqslant J. \tag{6.8}$$

注意，到此为止 J 还只是 $|m|$ 的最大值，而不是 $\boldsymbol{J}^2$ 的量子数.

②利用对易关系(6.4)式可以得到

$$\begin{aligned} J_z J_\pm |\lambda, m\rangle &= J_\pm J_z |\lambda, m\rangle \pm \hbar J_\pm |\lambda, m\rangle \\ &= (m \pm 1)\hbar J_\pm |\lambda, m\rangle. \end{aligned} \tag{6.9}$$

这表明，若 $|\lambda, m\rangle$ 是 J_z 的本征矢，相应的本征值为 $m\hbar$，则只要 $J_\pm |\lambda, m\rangle$ 不等于零，它也是 J_z 的本征矢，相应的本征值为 $(m \pm 1)\hbar$. 因而可写

$$\text{当 } J_+ |\lambda, m\rangle \neq 0 \text{ 时}, \quad J_+ |\lambda, m\rangle = c |\lambda, m+1\rangle, \tag{6.10}$$

$$\text{当 } J_- |\lambda, m\rangle \neq 0 \text{ 时}, \quad J_- |\lambda, m\rangle = c' |\lambda, m-1\rangle. \tag{6.11}$$

这样，我们看到，从 J_z 的任一本征矢 $|\lambda, m\rangle$ 出发，经过 $J_\pm$ 的作用，可以生成一系列的 J_z 的本征矢，相邻两个本征矢对应的本征值相差 $\hbar$. 因而 J_z 的本征值谱为 $m\hbar$

$$m = J, J-1, J-2, \cdots -J, \tag{6.12}$$

共 $2J+1$ 个值.

③再求 $\boldsymbol{J}^2$ 的本征值. 用 $|\lambda, J\rangle$ 表示 m 取最大值 J 时的本征矢. 由于不存在 $m > J$ 的态，故

$$J_+ |\lambda, J\rangle = 0.$$

将 J_- 作用于上式，并利用(6.5)式得到

$$J_- J_+ |\lambda, J\rangle = (\boldsymbol{J}^2 - J_z^2 - \hbar J_z) |\lambda, J\rangle = 0. \tag{6.13}$$

注意 $|\lambda, J\rangle$ 是 $\boldsymbol{J}^2$ 和 J_z 的共同本征矢，相应的 $\boldsymbol{J}^2$ 本征值为 $\lambda\hbar^2$，J_z 本征值为 $J\hbar$，见(6.6)(6.7)式. 即

$$\boldsymbol{J}^2 |\lambda, J\rangle = \lambda\hbar^2 |\lambda, J\rangle, \qquad J_z |\lambda, J\rangle = J\hbar |\lambda, J\rangle.$$

代入(6.13)式得

$$(\boldsymbol{J}^2 - J_z^2 - \hbar J_z)\mid \lambda, J\rangle = (\lambda - J^2 - J)\hbar^2 \mid \lambda, J\rangle = 0.$$

而按假设 $\mid \lambda, J\rangle \neq 0$，故 $\lambda - J^2 - J = 0, \lambda = J(J+1)$，即，$\boldsymbol{J}^2$ 的本征值为 $J(J+1)\hbar^2$. 由此可见，J 实际上是 $\boldsymbol{J}^2$ 的量子数，因而以下将相应的本征矢 $\mid \lambda, m\rangle$ 改写为 $\mid J, m\rangle$.

④到此为止，我们还没有确定 J 应取什么值. 由 m 的取值序列(6.11)式：$J, J-1, J-2, \cdots -J$，知 $J = -J + N$，故 $J = N/2, N = 0, 1, 2, , \cdots$. 于是得到

$$\boldsymbol{J}^2 \text{ 的本征值} = J(J+1)\hbar^2, \qquad J = 0, 1/2, 1, 3/2, \cdots. \tag{6.14}$$

即 J 可取的值为整数，半整数或零.

以上从角动量的对易关系出发，得到了角动量的本征值. 所得结果在 $J = 0$ 和正整数的情况下和在坐标表象中解 $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$ 的本征值方程得到的结果一致，然而这里多出了 J 取半整数的可能性. 这一情况表明，微观粒子除了轨道角动量 $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$ 以外还有另一种角动量，称为自旋角动量. 下面还要进一步讨论这个问题.

【例 1】 求球对称空间转子的能级.

解 绕固定轴转动的“平面转子”的能级已在 §5.3.1 例 3 中求得. 现在讨论的是绕任意轴转动的空间转子. 为了简单，只考虑球对称转子，即转动惯量的三个主值都相等的情况. 用 I 表示转动惯量的主值. 哈密顿量为

$$H = \boldsymbol{J}^2/(2I)$$

由于不存在自旋的贡献，因而 $\boldsymbol{J}^2$ 的本征值 $J(J+1)\hbar^2$ 的量子数 J 只取零和正整数值. 将它写为 L，得到能级

$$E = \hbar^2 L(L+1)/(2I), \qquad L = 0, 1, 2, \cdots.$$

能级简并度是 $2L+1$. 和这 $2L+1$ 个简并度对应的量子数 $m = -L, -L+1, \cdots, L$.

(二)角动量的矩阵元

由于 $\boldsymbol{J}^2$ 和 J_x, J_y, J_z 以及 J_+ 均可对易，因而可以固定 $\boldsymbol{J}^2$ 的本征值量子数为 J. 这一本征值(除 $J = 0$ 外)是简并的. 力学量在自身表象中是对角矩阵，若选用 $\boldsymbol{J}^2$ 和 J_z 的共同表象，则 $\boldsymbol{J}^2$ 和 J_z 的矩阵可以由其本征值给出：

$$\langle J, m' \mid \boldsymbol{J}^2 \mid J, m\rangle = J(J+1)\hbar^2 \delta_{mm'}, \tag{6.15}$$

$$\langle J, m' \mid J_z \mid J, m\rangle = m\hbar\delta_{mm'}. \tag{6.16}$$

由(6.10)、(6.11)式可见，在此表象中，J_+ 和 J_- 的矩阵元由下式决定：

$$\langle J, m' \mid J_+ \mid J, m\rangle = c\langle J, m' \mid J, m+1\rangle = c\delta_{m', m+1}, \tag{6.17}$$

$$\langle J, m' \mid J_- \mid J, m\rangle = c'\langle J, m' \mid J, m-1\rangle = c'\delta_{m', m-1}, \tag{6.18}$$

即知非零矩阵元是 $\langle J, m+1 \mid J_+ \mid J, m\rangle$ 和 $\langle J, m-1 \mid J_- \mid J, m\rangle$，它们分别等于 c 和 c'.

为了确定 c，取(6.10)式的共轭和它自身相乘，并利用(6.3)式：

$$\langle J,m\mid J_-J_+\mid J,m\rangle = |c|^2\langle J,m+1\mid J,m+1\rangle = |c|^2,$$

利用(6.5)式，可进一步得到

$$|c|^2 = \langle J,m\mid J_-J_+\mid J,m\rangle = \langle J,m\mid \boldsymbol{J}^2 - J_z^2 - \hbar J_z\mid J,m\rangle$$
$$= J(J+1)\hbar^2 - m(m+1)\hbar^2 = (J-m)(J+m+1)\hbar^2.$$

适当选择位相因子，可以使 c 为实数，故

$$c = \sqrt{(J-m)(J+m+1)}\hbar.$$

这样求得 J_+ 的非零矩阵元为

$$\langle J,m+1\mid J_+\mid J,m\rangle = \sqrt{(J-m)(J+m+1)}\hbar. \tag{6.19}$$

由(6.3)式得 J_- 的非零矩阵元为

$$\langle J,m\mid J_-\mid J,m+1\rangle = \sqrt{(J-m)(J+m+1)}\hbar. \tag{6.20}$$

代入(6.2)式，可求得 J_x 和 J_y 的矩阵元为

$$\langle J,m\mid J_x\mid J,m+1\rangle = \langle J,m+1\mid J_x\mid J,m\rangle$$
$$= (\hbar/2)\sqrt{(J-m)(J+m+1)}. \tag{6.21}$$
$$\langle J,m\mid J_y\mid J,m+1\rangle = -\langle J,m+1\mid J_y\mid J,m\rangle$$
$$= (\mathrm{i}\hbar/2)\sqrt{(J-m)(J+m+1)}. \tag{6.22}$$

(6.15)、(6.16)、(6.21)、(6.22)式就是角动量算符在 $\boldsymbol{J}^2$ 和 J_z 共同表象中的矩阵表示.

(三)整数角动量

按照(6.14)式，角动量的量子数 J 可以取整数值(包括0)或半整数值. 由 $\boldsymbol{L} = \boldsymbol{r}\times\boldsymbol{p}$ 定义的轨道角动量是整数角动量的例子. 作为角动量，它也满足(6.1)式型的对易关系

$$[L_x,L_y] = \mathrm{i}\hbar L_z,\quad [L_y,L_z] = \mathrm{i}\hbar L_x,\quad [L_z,L_x] = \mathrm{i}\hbar L_y, \tag{6.23}$$

只不过 $\boldsymbol{L}^2$ [本征值为 $l(l+1)\hbar^2$]的量子数 l 只能取整数值或零而不能为半整数，

$$l = 0,1,\cdots. \tag{6.24}$$

对于轨道角动量，常用的是坐标表象. §5.2.2 例3 已在坐标表象中讨论过轨道角动量，得到它的本征函数(球谐函数)

$$\langle \boldsymbol{r}\mid l,m\rangle = \langle r,\theta,\varphi\mid l,m\rangle = Y_{lm}(\theta,\varphi)$$
$$= \sqrt{(l-|m|)!(2l+1)/[4\pi(l+|m|)!]}\,P_l^{|m|}(\cos\theta)\mathrm{e}^{\mathrm{i}m\varphi}. \tag{6.25}$$

如果采用 $\boldsymbol{L}^2, L_z$ 的共同表象，则 $\boldsymbol{L}$ 成为矩阵，其矩阵元也是(6.21)、(6.22)式(将 J 换成 l).

【例1】 写出 $l=1$ 时 $\boldsymbol{L}$ 的三个笛卡儿分量在 $\boldsymbol{L}^2, L_z$ 表象中的矩阵.

解　由(6.16)、(6.21)、(6.22)式得

$$L_x=\frac{\hbar}{\sqrt{2}}\begin{pmatrix}0&1&0\\1&0&1\\0&1&0\end{pmatrix},\quad L_y=\frac{\hbar}{\sqrt{2}}\begin{pmatrix}0&-\mathrm{i}&0\\\mathrm{i}&0&-\mathrm{i}\\0&\mathrm{i}&0\end{pmatrix},\quad L_z=\hbar\begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix}. \tag{6.26}$$

【例2】　将 $l=1$ 时角动量在坐标表象中的本征函数(球谐函数)用位矢 $\boldsymbol{r}$ 的笛卡儿分量 x、y、z 表出.

解　根据(6.25)式和附录(Ⅳ.33)式有

$$\langle \boldsymbol{r}\mid 1,\pm 1\rangle=Y_{1,\pm 1}(\theta,\varphi)=\sqrt{\frac{3}{8\pi}}\sin\theta e^{\pm i\varphi}=\sqrt{\frac{3}{8\pi}}\frac{x\pm iy}{r}, \tag{6.27a}$$

$$\langle \boldsymbol{r}\mid 1,0\rangle=Y_{10}(\theta,\varphi)=\sqrt{\frac{3}{4\pi}}\cos\theta=\sqrt{\frac{3}{4\pi}}\frac{z}{r}. \tag{6.27b}$$

由此可见,由位矢的分量组成的函数 $(x\pm iy)/r$,z/r 可以作为波函数描述角动量量子数 $l=1$ 的状态.

(四)电子的自旋角动量　泡利矩阵

很多实验表明,电子和其他一些粒子具有固有的角动量,称为自旋角动量,或简称为自旋.自旋是一种纯量子效应,没有经典对应.

用 s 表示自旋算符,s_x,s_y,s_z 为它的三个分量.作为一种角动量算符,它们满足对易关系:

$$[s_x,s_y]=i\hbar s_z,\quad [s_y,s_z]=i\hbar s_x,\quad [s_z,s_x]=i\hbar s_y. \tag{6.28}$$

实验证明,电子的自旋角动量在任何方向的投影都只取两个值,按(6.12)式,这两个值之间相差 $\hbar$,因而 $s_z=\pm\hbar/2$,由此知电子的自旋角动量量子数 $s=1/2$.另一些粒子,如 ρ 介子,的自旋角动量在任意方向的投影可以取三个值,$s_z=0,\pm\hbar$.这种粒子的自旋角动量量子数 $s=1$.

自旋没有经典对应,不可能用坐标和动量的算符把它表示出来.但是,可以在 s^2 和 s_z 的共同表象里,得到它的矩阵表示.电子的自旋 $s=1/2$,按(6.16)、(6.21)、(6.22)式可以求得电子的自旋角动量在 s^2,s_z 表象中的矩阵为

$$s_x=(\hbar/2)\sigma_x,\qquad s_y=(\hbar/2)\sigma_y,\qquad s_z=(\hbar/2)\sigma_z, \tag{6.29}$$

其中

$$\sigma_x=\begin{pmatrix}0&1\\1&0\end{pmatrix},\qquad \sigma_y=\begin{pmatrix}0&-\mathrm{i}\\\mathrm{i}&0\end{pmatrix},\qquad \sigma_z=\begin{pmatrix}1&0\\0&-1\end{pmatrix}. \tag{6.30}$$

称为泡利矩阵.

电子的自旋角动量的大小是固定不变的,$s=1/2$,因而它在任一方向都有 $(2s+1)=2$ 个投影,即两个指向.自旋指向不同,电子的自旋状态也不同.描述电子自旋状态的态函数在 s_z 表象中写出是二行一列矩阵,称为自旋波函数或者旋量.

由 s_z 矩阵，不难得知它有两个本征值，即 $\pm\hbar/2$，相应的本征波函数是：

$$\chi_{1/2}=\begin{pmatrix}1\\0\end{pmatrix},\qquad \chi_{-1/2}=\begin{pmatrix}0\\1\end{pmatrix}. \tag{6.31}$$

一般的自旋波函数（旋量）是它们的叠加：

$$\chi=c_+\chi_{1/2}+c_-\chi_{-1/2}=\begin{pmatrix}c_+\\c_-\end{pmatrix}\quad(|c_+|^2+|c_-|^2=1). \tag{6.32}$$

它描述 $s_z=\hbar/2$ 的概率为 $|c_+|^2$，$s_z=-\hbar/2$ 的概率为 $|c_-|^2$ 的状态.

通常习惯上相对于自旋角动量，将 $\boldsymbol{L}=\boldsymbol{r}\times\boldsymbol{p}$ 称为轨道角动量；相对于自旋波函数，将以 $\boldsymbol{r}$ 为自变量的波函数称为轨道波函数，并将由轨道波函数描述的运动称为轨道运动. 当然，这只是一种习惯上用的词，并不意味着电子真的沿轨道运动.

在考虑到电子有自旋的情况下，描述电子运动的力学量就包含 4 个，而不是 3 个独立的变量. 其中 3 个是“轨道”变量，例如 x,y,z，或 r,θ,φ；另一个是自旋变量 s_z.（注意 s^2 有唯一确定值，不是变量）. 常用 σ 表示 s_z 的量子数. 于是，电子的任意波函数可写成

$$\Psi(x,y,z,\sigma)\qquad 或\qquad \Psi(r,\theta,\varphi,\sigma). \tag{6.33}$$

变量 x,y,z（或 r,θ,φ）取连续值，而变量 σ 取两个分立值：$\pm1/2$.

一般情况下，电子的自旋和轨道运动的耦合很弱，此时可以把电子的总波函数写成轨道波函数和自旋波函数的乘积.

$$\Psi=\Psi(x,y,z)\,\chi(\sigma)\quad 或\quad \Psi(r,\theta,\varphi)\,\chi(\sigma), \tag{6.34}$$

其中 $\chi(\sigma)$ 如(6.32)式.

ρ 介子的自旋 $s=1$，它的自旋算符在 s^2 和 s_z 共同表象中是 3×3 矩阵，如(6.26)式

$$s_x=\frac{\hbar}{\sqrt{2}}\begin{pmatrix}0&1&0\\1&0&1\\0&1&0\end{pmatrix},\quad s_y=\frac{\hbar}{\sqrt{2}}\begin{pmatrix}0&-\mathrm{i}&0\\\mathrm{i}&0&-\mathrm{i}\\0&\mathrm{i}&0\end{pmatrix},\quad s_z=\hbar\begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix}. \tag{6.35}$$

*（五）角动量合成

我们来考虑系统中的两个角动量 j_1 和 j_2 合成为角动量 $\boldsymbol{J}$

$$\boldsymbol{J}=\boldsymbol{j}_1+\boldsymbol{j}_2$$

的问题. 这样的例子有：电子的轨道角动量 $\boldsymbol{L}$ 和自旋角动量 $\boldsymbol{S}$ 合成为总角动量 $\boldsymbol{J}=\boldsymbol{L}+\boldsymbol{S}$，或两个电子的自旋角动量 $\boldsymbol{s}_1$ 和 $\boldsymbol{s}_2$ 合成为总自旋 $\boldsymbol{S}=\boldsymbol{s}_1+\boldsymbol{s}_2$，等等.

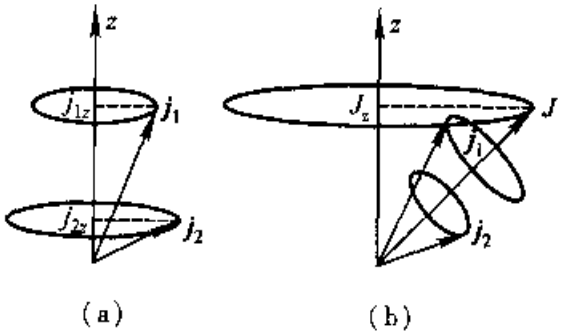


图 1　角动量合成的示意图

在用量子力学推导角动量合成之前，让我

们先用初等量子论的观点，作一个简单的形象化讨论.

用两根带箭头的线段表示 j_1 和 j_2，如图 1 a. 假定它们的 z 分量 j_{1z}、j_{2z} 有确定值. 此时，j_{1x}、j_{1y}、j_{2x}、j_{2y} 都没有确定值. 形象地，可以认为 j_1 和 j_2 绕 z 轴旋转，如图 1 a.

当 j_1 和 j_2 合成为 $\boldsymbol{J}$ 时 j_1 和 j_2 绕 $\boldsymbol{J}$ 旋转. 此时，j_{1z} 和 j_{2z} 不再有确定值，而只有 J_z 有确定值，如图 1(b).

因此，合成之前的状态是 j_1、j_2、j_{1z}、j_{2z} 有定值的状态；合成以后的状态是 j_1、j_2、J、J_z 有定值的状态.

当 j_1 和 j_2 平行时，J 有最大值 j_1+j_2；当 j_1 和 j_2 反平行时，J 有最小值 $|j_1-j_2|$. 由于量子化，相邻 J 值之间相差 1.

下面就来用量子力学对以上的形象化讨论作出严格的证明.

在量子力学中，描述系统的力学量完全集由一组可对易的算符组成. 从以上的讨论可以预期，由 j_1 和 j_2 组成的系统有两个力学量完全集——j_1^2，j_2^2，j_{1z}，j_{2z} 和 j_1^2，j_2^2，J^2，J_z. 下面通过研究这几个力学量算符之间的对易关系，来证明这一论断.

假定 j_1 和 j_2 的耦合很弱，因而相互间可对易

$$[j_1\ j_2]=0. \tag{6.36}$$

由于

$$\boldsymbol{J}^2=j_1^2+j_2^2+2\boldsymbol{j}_1\cdot\boldsymbol{j}_2,\qquad J_z=j_{1z}+j_{2z}, \tag{6.37}$$

所以

$$[j_1^2,\boldsymbol{J}^2]=[j_2^2,\boldsymbol{J}^2]=[j_1^2,J_z]=[j_2^2,J_z]=0. \tag{6.38}$$

但是，由于 $\boldsymbol{j}_1$（$\boldsymbol{j}_2$）的各个分量之间不对易，因而 $\boldsymbol{j}_1$（$\boldsymbol{j}_2$）和 $\boldsymbol{J}$ 不对易；同时，由于 $\boldsymbol{J}^2$ 中包含 $\boldsymbol{j}_1\cdot\boldsymbol{j}_2=j_{1x}j_{2x}+j_{1y}j_{2y}+j_{1z}j_{2z}$，因而 $\boldsymbol{j}_1$（$\boldsymbol{j}_2$）和 $\boldsymbol{J}^2$ 也不对易.

给定 j_1^2，j_2^2 的量子数 j_1，j_2，则 j_{1z}，j_{2z} 的量子数所能取的值是

$$\begin{aligned}m_1&=-j_1,-j_1+1,\cdots,j_1;\\ m_2&=-j_2,-j_2+1,\cdots,j_2;\end{aligned} \tag{6.39}$$

分别有 $2j_1+1$ 和 $2j_2+1$ 个值.

我们来求 $\boldsymbol{J}^2$，J_z 的量子数 J，M 所取的值.

这一系统有两组完备本征矢. 一组是 j_1^2，j_2^2，j_{1z}，j_{2z} 的共同本征矢 $|j_1,j_2,m_1,m_2\rangle$，另一组是 j_1^2，j_2^2，$\boldsymbol{J}^2$，J_z 的共同本征矢 $|j_1,j_2,J,M\rangle$. 两组本征矢的数目应该同样多. 下面根据这一考虑来确定量子数 J 和 M 所能取的值. 为了确定起见，假定 $j_2\leqslant j_1$.

由(6.37)式 $J_z=j_{1z}+j_{2z}$ 可知 $M=m_1+m_2$. 由于 m_1 和 m_2 的最大可能值是 $m_1=j_1$ 和 $m_2=j_2$，因此 M 的最大可能值是 $M=j_1+j_2$，它对应于 $J=j_1+j_2$；对于 M 的下一个值 j_1+j_2-1 有两种可能性：$m_1=j_1$，$m_2=j_2-1$ 和 $m_1=j_1-1$，$m_2=j_2$，它们对应于能够得到 $M=j_1+j_2-1$ 的两个 J 值：$J=j_1+j_2$ 和 $J=j_1+j_2-1$；见图 2. 我们看到，M 的值每减小 1，对应的状态数目就增加 1，相应地 J 的可能值也增加 1. 这样继续下去，直到 $M=j_1-j_2$. 当 M 再减小时，对应的状态数不再增加，保持等于 $2j_2+1$. 其原因是，m_2 已取遍了它所能取的全部值：j_2，j_2-1，$\cdots$，$-j_2$，不可能再取其他值. 当 M 继续减小时，状态数保持等于 $2j_2+1$，直到 $M=-(j_1-j_2)$. 在那以后，M 继续减小，对应的状态数随之减少. 最后，当 M 减小到 $M=-(j_1+j_2)$ 时，对应的状态数等于

1，即 $J = j_1 + j_2$ 的一个态. 这样，J 所能取的值就是：$j_1 + j_2, j_1 + j_2 - 1, \cdots, j_1 - j_2$.

以上是假定 $j_2 \leqslant j_1$. 如果反过来，$j_1 \leqslant j_2$，J 所能取的值就是：$j_1 + j_2, j_1 + j_2 - 1, \cdots, j_2 - j_1$. 合起来就有

$$J = j_1 + j_2, j_1 + j_2 - 1, \cdots, |j_1 - j_2|, \tag{6.40}$$

$$M = J, J-1, \cdots, -J. \tag{6.41}$$

现在来讨论角动量合成时态矢的变换.

如果两个角动量 j_1 和 j_2 完全没有相互作用，系统的完备本征矢是 $j_1^2, j_2^2, j_{1z}, j_{2z}$ 的共同本征矢 $|j_1, j_2, m_1, m_2\rangle$. 当 j_1 和 j_2 有微弱相互作用时，它们合成为 J，使 j_1 和 j_2 的分量不再是守恒量，但是 j_1^2, j_2^2 仍保持守恒，见(6.38)式. 这可以形象化地说成是：j_1 和 j_2 保持长度和夹角不变而作为一个整体绕 J 转动. 在此情况下，系统的完备本征矢成为 j_1^2, j_2^2, J^2, J_z 的共同本征矢 $|j_1, j_2, J, M\rangle$. 这两组本征矢 $|j_1, j_2, m_1, m_2\rangle$ 和 $|j_1, j_2, J, M\rangle$ 之间应该有相互叠加的关系：

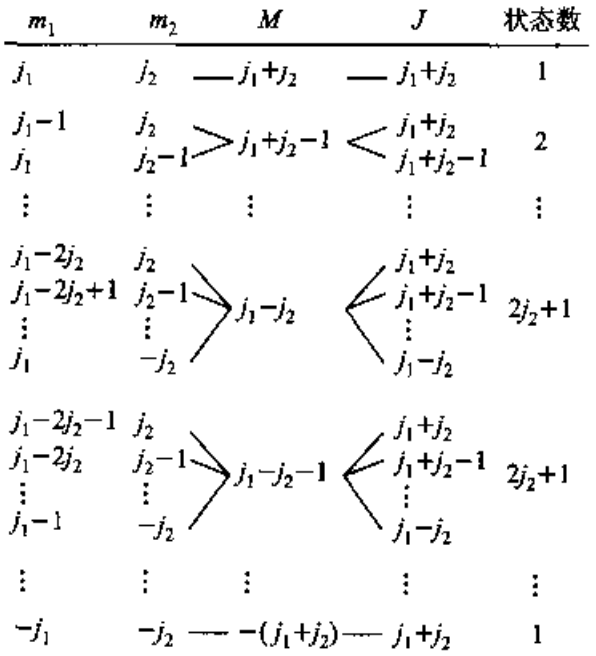


图 2　角动量合成的量子数相加规则

$$|j_1, j_2, m_1, m_2\rangle = \sum_{JM} |j_1, j_2, J, M\rangle \langle j_1, j_2, J, M | j_1, j_2, m_1, m_2\rangle, \tag{6.42}$$

$$|j_1, j_2, J, M\rangle = \sum_{m_1 m_2} |j_1, j_2, m_1, m_2\rangle \langle j_1, j_2, m_1, m_2 | j_1, j_2, J, M\rangle. \tag{6.43}$$

指标求和的范围由(6.39)、(6.40)、(6.41)式决定.

(6.42)、(6.43)式中的系数称为克莱布斯－高登系数，简称为 C－G 系数

$$C^{JM}_{j_1 m_1 j_2 m_2} = \langle j_1, j_2, J, M | j_1, j_2, m_1, m_2\rangle = \langle j_1, j_2, m_1, m_2 | j_1, j_2, J, M\rangle^*, \tag{6.44}$$

通常选相因子使 C－G 系数为实数，这时上式中的复共轭号 * 可以去掉.

C－G 系数的值可以查表. 在表Ⅰ和表Ⅱ中列出了 $j_2 = 1/2$ 和 $j_2 = 1$ 两种情况下的 C－G 系数.

表Ⅰ　$C^{J,(m+\mu)}_{j_1 m, 1/2, \mu}$

J \ μ	1/2	−1/2
$j + 1/2$	$\sqrt{(j+m+1)/(2j+1)}$	$\sqrt{(j-m+1)/(2j+1)}$
$j - 1/2$	$-\sqrt{(j-m)/(2j+1)}$	$\sqrt{(j+m)/(2j+1)}$

表Ⅱ $C^{j,(m+\mu)}_{j_1,m,(,\mu}$

J \ μ	1	0	-1
$j+1$	$\sqrt{\frac{(j+m+1)(j+m+2)}{(2j+1)(2j+2)}}$	$\sqrt{\frac{(j-m+1)(j+m+1)}{(2j+1)(j+1)}}$	$\sqrt{\frac{(j-m+1)(j-m+2)}{(2j+1)(2j+2)}}$
j	$-\sqrt{\frac{(j+m+1)(j-m)}{2j(j+1)}}$	$\frac{m}{\sqrt{j(j+1)}}$	$\sqrt{\frac{(j-m+1)(j+m)}{2j(j+1)}}$
$j-1$	$\sqrt{\frac{(j-m-1)(j-m)}{2j(2j+1)}}$	$-\sqrt{\frac{(j-m)(j+m)}{j(2j+1)}}$	$\sqrt{\frac{(j+m)(j+m-1)}{2j(2j+1)}}$

【例3】 写出轨道角动量(l)和自旋角动量($s=1/2$)合成为总角动量(JM)的波函数(假定 $l\neq 0$).

解 各个量子数的可能取值是 $J=l\pm 1/2,\mu=\pm 1/2,m=M-\mu$,由表Ⅰ得到

$$\psi_{J=l\pm 1/2,M}=\sum_{\mu}C^{(l\pm 1/2)M}_{l,M-\mu,1/2,\mu}Y_{l,M-\mu}\chi_{\mu}$$

$$=\frac{1}{\sqrt{2l+1}}\left(\pm\sqrt{l\pm M+1/2}Y_{l,M-1/2}\chi_{1/2}+\sqrt{l\mp M+1/2}Y_{l,M+1/2}\chi_{-1/2}\right). \tag{6.45}$$

其中,$\chi_{\pm 1/2}$ 是 $s=1/2$ 自旋波函数,见(6.31)式.将(6.31)式代入(6.45)式得到总角动量的本征波函数 $\psi_l^{J,M}$

$$\psi_l^{(l\pm 1/2),M}=\frac{1}{\sqrt{2l+1}}\begin{pmatrix}\pm\sqrt{l\pm M+1/2}Y_l^{M-(1/2)}\\ \sqrt{l\mp M+1/2}Y_l^{M+(1/2)}\end{pmatrix}. \tag{6.46}$$

§5.4.7 薛定谔绘景和海森伯绘景

(一)量子力学的不同绘景

通过前面的讨论看到,量子力学的原理包括关于状态的原理和关子力学量的原理两部分.但是,不论是描述状态的态矢量 $|a\rangle$,还是描述力学量的算符 F,都不是直接可观测量,直接可观测的是力学量 F 在状态 $|a\rangle$ 中取不同值的概率分布

$$P_n=|\langle n|a\rangle|^2. \tag{7.1}$$

其中 $|n\rangle$ 是算符 F 的本征矢,是用来展开态矢 $|a\rangle$ 的基底.

概率分布 P_n 会随时间变化,$P_n=P_n(t)$,这一变化是可以观测的,是量子理论应该能够预言的.由(7.1)式可见,P_n 随时间的变化既可以归因于态矢

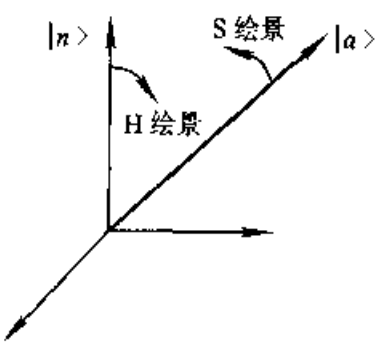


图1 S 绘景和 H 绘景的等效性

$|a\rangle$ 随时间变化，也可以归于基底 $|n\rangle$ 随时间变化. 这两种观点对物理观测结果 $P_n(t)$ 可以作出相同的预言，因而是完全等效的.

为了形象化地理解这两种不同的观点，在图 1 上画出一个示意图. 图中作为坐标基矢的三个矢量表示希尔伯特空间中力学量的本征矢 $|n\rangle(n=1,2,\cdots)$. 概率 $P_n(n=1,2,\cdots)$ 是标架基矢 $|n\rangle(n=1,2,\cdots)$ 和态矢 $|a\rangle$ 的点积的绝对值平方. $|a\rangle$ 和 $|n\rangle$ 的这一点积决定于它们之间的相对位置. 因此，概率 P_n 的时间演化既可以表示为基矢 $|n\rangle$ 不动，态矢 $|a\rangle$ 转动；也可以等效地表示为态矢 $|a\rangle$ 不动，而标架基矢 $|n\rangle$ 沿相反方向转动，如图 1. 直到现在，我们都假定力学量的算符不含时间，因而力学量本征矢不随时间变化，即认为基底是不动的. 因而力学量的概率分布随时间的变化被看成仅由态矢随时间变化而引起. 通常将这种描绘状态随时间变化的图像称为薛定谔绘景. 也可以这样看力学量的概率随时间的变化，即认为态矢不动，而基矢随时间变化，从而使概率 P_n 依赖时间. 这样描绘状态随时间变化的图像称为海森伯绘景. 注意："绘景"和"表象"是两个不同的概念. 表象指的是选定基底使态矢量和算符有具体表示；绘景是指描绘状态随时间变化的图像. 在同一绘景里，还可以有不同表象.

以下具体讨论薛定谔绘景和海森伯绘景，以及它们之间的联系.

(二)薛定谔绘景

在薛定谔绘景，如果力学量 F 不显含时间，则它不随时间变化. 根据本征值方程 $F|n\rangle = f_n|n\rangle$ 知，其本征矢也不随时间变化. 这样有

$$\text{薛定谔绘景(当 } F \text{ 不显含时间时)}\quad \mathrm{d}F/\mathrm{d}t = 0,\quad \mathrm{d}|n\rangle/\mathrm{d}t = 0. \tag{7.2}$$

在这种绘景里，状态随时间变化体现于态矢量随时间变化，可以用薛定谔方程描述：

$$\text{薛定谔绘景}\qquad i\hbar\mathrm{d}|t\rangle/\mathrm{d}t = H|t\rangle. \tag{7.3}$$

薛定谔方程告诉我们，若已知 $t=0$ 时系统的态矢 $|t=0\rangle$，则由方程可以决定以后任何时刻的态矢 $|t\rangle$. 薛定谔方程反映的这种因果关系也可以用下面定义的时间演化算符 $S(t)$ 来描述：

$$\text{薛定谔绘景}\qquad |t\rangle = S(t)|t=0\rangle, \tag{7.4}$$

$S(t)$ 作用在初始时刻的态矢 $|t=0\rangle$ 上，得到以后时刻的态矢 $|t\rangle$，显然，应有

$$S(0) = \mathbf{1}. \tag{7.5}$$

为保证波函数在任何时刻都是归一化的，$S(t)$ 应是幺正算符，因为，当且仅当

$$S^{\dagger}S = \mathbf{1}. \tag{7.6}$$

时，才有

$$\langle t|t\rangle = \langle 0|S^{\dagger}(t)S(t)|0\rangle = \langle 0|0\rangle = \mathbf{1}. \tag{7.7}$$

为确定 $S(t)$ 的形式，将(7.4)式代入薛定谔方程，得到

$$i\hbar dS(t)/dt = HS(t). \tag{7.8}$$

若 H 不明显依赖时间，上式有如下的形式解：

$$S(t) = e^{-(i/\hbar)Ht}. \tag{7.9}$$

这样，态矢量对时间的依赖关系(7.4)式可写成：

$$|t\rangle = e^{-(i/\hbar)Ht}|t=0\rangle. \tag{7.10}$$

这里，我们遇到了指数形式的算符，它按所展开的级数来定义.(7.10)式的意义是

$$|t\rangle = \sum_{n=0}^{\infty}\frac{1}{n!}\left(-\frac{i}{\hbar}Ht\right)^n|t=0\rangle. \tag{7.11}$$

不难验证，$S(t)$ 的逆算符是

$$S^{-1}(t) = e^{(i/\hbar)Ht}. \tag{7.12}$$

(三)从薛定谔绘景到海森伯绘景

在海森伯绘景里，态矢量不随时间变化

$$\text{海森伯绘景} \qquad d|\rangle^H/dt = 0. \tag{7.13}$$

也就是说，任何时刻 t 的态矢都和 $t=0$ 时的态矢一样. 令 $t=0$ 时刻薛定谔绘景的态矢和海森伯绘景的态矢重合：

$$|\rangle^H = |t=0\rangle^S. \tag{7.14}$$

则由(7.4)式知

$$|\rangle^H = S^{-1}|t\rangle^S = e^{(i/\hbar)Ht}|t\rangle^S. \tag{7.15}$$

这里用指标 H 和 S 分别标明海森伯绘景和薛定谔绘景中的量. 上式给出了从薛定谔绘景到海森伯绘景的幺正变换，其中，指数上的 H 是薛定谔绘景中的哈密顿算符. 以下将看到，它也等于海森伯绘景中的哈密顿算符.

现在来讨论两种绘景中的算符之间的变换，设在薛定谔绘景中力学量算符的定义式为

$$F^S|a\rangle^S = |b\rangle^S.$$

两边同时经受幺正变换(7.15)式：

$$e^{(i/\hbar)Ht}F^S|a\rangle^S = e^{(i/\hbar)Ht}|b\rangle^S = |b\rangle^H = e^{(i/\hbar)Ht}F^Se^{-(i/\hbar)Ht}e^{(i/\hbar)Ht}|a\rangle^S$$
$$= e^{(i/\hbar)Ht}F^Se^{-(i/\hbar)Ht}|a\rangle^H.$$

它应该具有 $F^H|a\rangle^H = |b\rangle^H$ 的形式，因而

$$F^H = e^{(i/\hbar)Ht}F^Se^{-(i/\hbar)Ht}. \tag{7.16}$$

这就是力学量的算符从薛定谔绘景到海森伯绘景的变换.

由(7.16)式可以看到，如果 F^S 和 $H^S = H$ 可对易，则 $F^H = F^S$. 这是因为，当 $[F^S, H] = 0$ 时，可以把 F^S 移到 $e^{(i/\hbar)Ht}$ 的左边，从而使(7.16)式中的两个指数因子消去. 特别是，由此就知道：

$$H^H = H^S. \tag{7.17}$$

即:海森伯和薛定谔绘景里的哈密顿算符相同.这就是在前面的各表示式中对哈密顿算符不标绘景标号的原因.

从(7.16)式可以明显看到,即使在薛定谔绘景中算符 F 不明显和时间有关,只要它和 H 不可对易,在海森伯绘景中它也会依赖时间.从(7.16)式直接求 F^{H} 对时间的微商可得

海森伯绘景(当 F 不显含时间时)　$\mathrm{d}F^{H}/\mathrm{d}t = (\mathrm{i}/\hbar)[H,F^{H}]$.　(7.18)

这就是海森伯绘景中的运动方程.本节开头已经指出,在海森伯绘景中,力学量概率分布随时间的变化体现在基矢随时间变化.方程(7.18)决定了力学量算符对时间的依赖,从而也就决定了基矢对时间的依赖,见本章习题17.

如果力学量显含时间,例如随时间变化的势能 $U(\boldsymbol{r},t)$,则(7.2)式和(7.18)式应修改为

$$\mathrm{d}F^{(S)}/\mathrm{d}t = \partial F^{(S)}/\partial t, \qquad \mathrm{d}F^{(H)}/\mathrm{d}t = \partial F^{(H)}/\partial t + (\mathrm{i}/\hbar)[H,F^{(H)}]. \tag{7.19}$$

§5.4.8　用密度矩阵表示量子态

(一)投影算符

在§5.4.2中已经指出,一个任意的态矢 $|\,a\rangle$ 可以用一组正交完备矢量组[§5.4.2(2.21)式]

$$|\,i\rangle, \qquad i = 1,2,\cdots. \tag{8.1}$$

展开为§5.4.2(2.25)式

$$|\,a\rangle = \sum_{i=1}^{\infty} |\,i\rangle\langle i\,|\,a\rangle. \tag{8.2}$$

此式的右边可以看成是算符 $|\,i\rangle\langle i|$ 作用在态矢 $|\,a\rangle$ 上再对 i 求和.这一算符称为投影算符,用 P_i 表示

$$P_i = |\,i\rangle\langle i\,|. \tag{8.3}$$

它作用到态矢 $|\,a\rangle$ 上得到的 $P_i|\,a\rangle = |\,i\rangle\langle i\,|\,a\rangle$ 是将态矢 $|\,a\rangle$ 投影到基矢 $|\,i\rangle$ 方向上得到的矢量.将它们相加,又回到原态矢 $|\,a\rangle$:

$$\sum_{i=1}^{\infty} P_i|\,a\rangle = |\,a\rangle, \tag{8.4}$$

见(8.2)式.其原因是由于基矢组(8.1)的完备性.

为了形象化地理解上述结果,考虑一个三维欧氏空间的矢量 $\boldsymbol{a}$.它可以用一组正交完备基矢 $\boldsymbol{e}_i(i = 1,2,3)$ 所形成的笛卡儿坐标系展开为

$$\boldsymbol{a} = \sum_{i=1}^{3} \boldsymbol{e}_i\boldsymbol{e}_i \cdot \boldsymbol{a}.$$

右边和式中的每一项 $e_ie_i\cdot a$ 是将矢量 a 投影到坐标轴上得到的矢量，见图1. 将它们相加就回到原矢量 a. 因此，并矢 e_ie_i 是投影到坐标轴 i 上的投影算符.

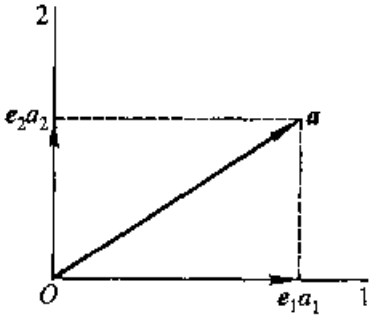


图1 矢量 a 在坐标轴上的投影

由§5.4.2(2.22)、(2.26)式易证，投影算符满足正交"归一"性

$$P_iP_j = \delta_{ij}P_i, \tag{8.5}$$

和完备性

$$\sum_{i=1}^{\infty} P_i = 1. \tag{8.6}$$

由(8.5)式知，

$$P_i^n = P_i \quad (n\text{ 为任意正整数}). \tag{8.7}$$

投影算符的这一性质称为幂等性.

(二)密度矩阵

我们知道，量子态 a 可以用态矢 $|a\rangle$ 描述. 它归一化为

$$\langle a|a\rangle = 1. \tag{8.8}$$

用它可以得到力学量 F 在态 a 中的平均值

$$F^{(a)} = \langle a|F|a\rangle, \tag{8.9}$$

以及在态 a 中测量力学量 F 得值 f_n 的概率幅

$$c_n = \langle n|a\rangle, \tag{8.10}$$

其中，$|n\rangle$ 是和本征值 f_n 对应的本征矢. 由此又得到在态 a 中测量力学量 F 得值 f_n 的概率

$$P_n^{(a)} = |c_n|^2 = |\langle n|a\rangle|^2. \tag{8.11}$$

下面来证明，量子态 a 可以等效地用由 $|a\rangle$ 和它的左矢 $\langle a|$ 组成的算符，

$$\rho_a = |a\rangle\langle a| \tag{8.12}$$

描述. ρ_a 称为态 a 的密度算符. 进入具体表象以后，它成为矩阵，称为密度矩阵. 它的归一化是

$$\mathrm{tr}\,\rho_a = 1. \tag{8.13}$$

用它可以得到力学量 F 在态 a 中的平均值

$$\overline{F}^{(a)} = \mathrm{tr}\,F\rho_a = \mathrm{tr}\,\rho_aF. \tag{8.14}$$

以及在态 a 中测量力学量 F 得值 f_n 的概率

$$P_n^{(a)} = \mathrm{tr}\,P_n\rho_a = \mathrm{tr}\,\rho_aP_n, \tag{8.15}$$

其中 P_n 是本征态 n 的投影算符

$$P_n = |n\rangle\langle n|. \tag{8.16}$$

证 (8.8)式的左边是一个 c 数，它可以看成一个 1×1 矩阵的迹

$$1 = \langle a \mid a \rangle = \mathrm{tr}\langle a \mid a \rangle = \mathrm{tr} \mid a \rangle\langle a \mid = \mathrm{tr}\,\rho_a.$$

在写第三个等式时用了"矩阵乘积求迹时可以将相乘的矩阵循环"的规则. 同理有

$$\bar{F}^{(a)} = \langle a \mid F \mid a \rangle = \mathrm{tr}\langle a \mid F \mid a \rangle = \mathrm{tr} \mid a \rangle\langle a \mid F = \mathrm{tr}\,\rho_a F = \mathrm{tr}\,F\rho_a.$$

$$P_n^{(a)} = \mid \langle n \mid a \rangle \mid^2 = \langle n \mid a \rangle\langle a \mid n \rangle = \mathrm{tr}\langle n \mid a \rangle\langle a \mid n \rangle$$

$$= \mathrm{tr} \mid n \rangle\langle n \mid a \rangle\langle a \mid = \mathrm{tr}\,P_n\rho_a = \mathrm{tr}\,\rho_a P_n. \qquad \text{【证毕】}$$

力学量 F 在态 a 中的平均值和在态 a 中测量力学量 F 得值 f_n 的概率是实际可测量. 利用密度矩阵 ρ_a 和态矢 $\mid a\rangle$ 都可以得到这些可测量，两者在这个意义上是等效的. 但是用密度矩阵得不到在态 a 中测量力学量 F 得值 f_n 的概率幅 c_n 的相位角. 这一相位角本身不直接可测，但是，在量子态的叠加中它起着关键性作用. 正是量子态之间的相位差在它们叠加的过程中产生了干涉和衍射这一类典型的波动现象. 例如，设态 c 是态 a 和 b 的叠加. 用态矢写出就是

$$\mid c\rangle = \alpha \mid a\rangle + \beta \mid b\rangle. \tag{8.17}$$

物理量 F 在态 c 中的平均值是

$$\bar{F}^{(c)} = \langle c \mid F \mid c \rangle = (\alpha^* \langle a \mid + \beta^* \langle b \mid) F(\alpha \mid a\rangle + \beta \mid b\rangle)$$

$$= P_1 F^{(a)} + P_2 \bar{F}^{(b)} + \alpha^*\beta\langle a \mid F \mid b\rangle + \beta^*\alpha\langle b \mid F \mid a\rangle, \tag{8.18}$$

其中 $P_1 = \mid \alpha \mid^2$，$P_2 = \mid \beta \mid^2$ 分别是态 a 和 b 在态 c 中所占的比例(态 a 和 b 在态 c 中的概率). (8.18)式的后两项是交叉项，它们体现了量子干涉. 用密度矩阵不能描述量子态间的干涉.

密度矩阵的一个重要应用是用于描述宏观系统的状态. 由于宏观系统的自由度数非常巨大(有阿伏伽德罗常数 $\sim 6 \times 10^{23}$ 的量级)，而且和周围环境有极为复杂的相互作用，不可能对它作完整的量子描述. 用 $\mid n\rangle (n = 1,2,\cdots)$ 表示系统的一组量子态. 力学量 F 在态 $\mid n\rangle$ 中的量子平均值是 $\bar{F}^{(n)} = \langle n \mid F \mid n\rangle$. 但是对于宏观系统，我们不知道它确切地处于那一个量子态，而只能知道它是以概率 p_n 处于态 $\mid n\rangle (n = 1,2,\cdots)$. 因而实际观测到的力学量 F 的平均值是

$$F = \sum_{n=1}^{\infty} P_n F^{(n)} = \sum_{n=1}^{\infty} P_n \langle n \mid F \mid n\rangle. \tag{8.19}$$

这称为系综平均，见 §7.1.4. 它包含分别由 $\langle n \mid \cdots \mid n\rangle$ 和 $\sum_{n=1}^{\infty} P_n \cdots$ 表示的量子和统计两重平均. 将(8.14)式代入，得到

$$F = \mathrm{tr}\,\rho F, \tag{8.20}$$

其中

$$\rho = \sum_{n=1}^{\infty} P_n \mid n\rangle\langle n \mid = \sum_{n=1}^{\infty} P_n \rho_n \tag{8.21}$$

是所研究的宏观态的密度矩阵. 这种量子和统计两重平均的应用见 §7.1.5(5.

12)(5.13)式.

(8.17)式那样的态矢相加称为量子态的叠加.由此得到的态 c 和态 a 态 b 一样都是纯量子态,简称纯态.(8.21)式那样的密度矩阵相加称为态的混合.由混合得到的不是纯量子态,而是混合态.混合态中丢失了态之间的量子干涉,因而它不是纯的量子对象.

习 题

1. 求一维无限深势阱(势箱)中粒子的坐标和动量在能量表象中的矩阵元.

2. 设 A 为一个厄米矩阵,本征值为 a_n,试证 e^{iA} 为幺正矩阵,本征值为 e^{ia_n}.

3. 在一组正交归一的基矢 $|1\rangle,|2\rangle,|3\rangle$ 所张成的矢量空间里,定义了

$$H=\hbar\omega\begin{pmatrix}1&0&0\\0&-1&0\\0&0&-1\end{pmatrix},\qquad B=b\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix},$$

ω,b 为实数,证明 H 和 B 对易,并求它们的共同本征矢.

4. 设厄米算符 $\hat{A}$,$\hat{B}$ 满足 $\hat{A}^2=\hat{B}^2=1$,且 $\hat{A}\hat{B}+\hat{B}\hat{A}=1$,求:

(a)在 A 表象中,$\hat{A}$ 和 $\hat{B}$ 的矩阵表示;

(b)在 B 表象中,$\hat{A}$ 和 $\hat{B}$ 的矩阵表示;

(c)在 A 表象中,$\hat{B}$ 的本征值和本征矢;

(d)在 B 表象中,$\hat{A}$ 的本征值和本征矢;

(e)由 A 表象到 B 表象的幺正变换矩阵 S.

5. 利用不确定关系估计线性谐振子的零点能.

6. 从动量算符在动量表象中的矩阵元出发,求坐标表象中的动量算符.

7. 从角动量的对易关系 §5.4.6(6.1)式出发,证明 §5.4.6(6.4)式.

$$[J_+,J_-]=2\hbar J_z,\qquad [J_z,J_+]=\hbar J_+,\qquad [J_z,J_-]=-\hbar J_-.$$

以及

$$[\boldsymbol{J}^2,J_x]=[\boldsymbol{J}^2,J_y]=[\boldsymbol{J}^2,J_z]=[\boldsymbol{J}^2,J_z]=0.$$

8. 已知 ρ 介子的自旋 $s=1$(确切地说是 ρ 介子自旋角动量的量子数 $s=1$).(a)写出 ρ 介子自旋 s_x、s_y 在 s^2,s_z 表象中的矩阵和本征矢量;(b)将 s_x、s_y 对角化.

9. 证明泡利矩阵有如下性质:

(a) $\sigma_i^2=1$;$\sigma_i\sigma_j=\mathrm{i}\sum_k\varepsilon_{ijk}\sigma_k$(当 $i\neq j$);或合写为:$\sigma_i\sigma_j=\delta_{ij}+\mathrm{i}\sum_k\varepsilon_{ijk}\sigma_k$.

(b) $\sigma_i\sigma_j+\sigma_j\sigma_i=2\delta_{ij}$;$\sigma_i\sigma_j-\sigma_j\sigma_i=2\mathrm{i}\sum_k\varepsilon_{ijk}\sigma_k$.

(c) $\mathrm{tr}\,\sigma_i=0$.

10. 设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为两个任意可对易矢量,求 $(\boldsymbol{\sigma}\cdot\boldsymbol{A})(\boldsymbol{\sigma}\cdot\boldsymbol{B})$.

11. 利用第 10 题的结果,证明 $e^{-i\theta\boldsymbol{e}_n\cdot\boldsymbol{\sigma}}=\cos\theta-i\boldsymbol{e}_n\cdot\boldsymbol{\sigma}\sin\theta$,其中 $\boldsymbol{e}_n$ 为一个任意单位矢量.

12. 设两个角动量 $\boldsymbol{j}_1$,$\boldsymbol{j}_2$ 合成为 $\boldsymbol{J}$,$\boldsymbol{j}_1+\boldsymbol{j}_2=\boldsymbol{J}$.证明:在 j_1,j_2 固定的情况下,两组基矢 $|j_1,j_2,$

$m_1, m_2\rangle$ 和 $|j_1, j_2, J, M\rangle$ 有相同的状态数.

13. 证明 C－G 系数的正交性

$$\sum_{m_1, m_2} C^{JM}_{j_1, m_1; j_2, m_2} C^{J'M'}_{j_1, m_1; j_2, m_2} = \delta_{JJ'}\delta_{MM'},$$

$$\sum_{m_1, m_2} C^{JM}_{j_1, m_1; j_2, m_2} C^{JM}_{j_1, m'_1; j_2, m'_2} = \delta_{m_1 m'_1}\delta_{m_2 m'_2}.$$

14. 两个自旋 $s_1 = s_2 = 1/2$ 的自旋波函数分别是：$|s_1, s_{1z}\rangle = \chi^{(1)}_{1/2}, \chi^{(1)}_{-1/2}$ 和 $|s_2, s_{2z}\rangle = \chi^{(2)}_{1/2}, \chi^{(2)}_{-1/2}$ [见 §5.4.6(6.31)式]. 利用 C－G 系数（§6.4.6 表 I）求它们合成为 $\boldsymbol{S} = \boldsymbol{s}_1 + \boldsymbol{s}_2$ 的波函数 $|S, S_z\rangle$.

15. 证明由 §5.4.6(6.46)式给出的 $\mathscr{Y}_l^{l\pm 1/2, M}$ 是 $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$ 的平方及 z 分量：$\boldsymbol{J}^2, J_z$ 的本征函数，并求相应的本征值.

15. 已知在薛定谔绘景中线性谐振子的哈密顿

$$\hat{H} = \hat{p}^2/(2m) + m\omega^2 x^2/2,$$

试求海森伯绘景中谐振子的坐标和动量算符.

17. 将薛定谔绘景中的本征值方程 $F|n\rangle = f_n|n\rangle$ 转到海森伯绘景，求海森伯绘景中的本征矢 $|n, t\rangle^{\mathrm{H}}$，并推导它随时间演化的方程. 将所得结果和薛定谔绘景中态矢随时间演化的薛定谔方程比较，说明其意义.

18. 写出自旋为 1/2 的粒子的自旋波函数 $|\chi\rangle$ 和密度矩阵 ρ_χ 的 s_z 表象. 利用它们计算 s_x 和 s_y 在状态 χ 中的平均值.

第五章　多粒子系统　全同性原理

这一章讨论多个相同粒子组成的系统.

在经典力学中,与处理单粒子系统相比较,处理相同粒子构成的多粒子系统时,新增加的需要考虑的因素就是系统的自由度增加,使具体计算增加困难,而在理论框架里并不需要添加新的原则.在量子力学中则不然,在处理多个微观粒子组成的系统时,粒子的全同性要求在理论框架里增加新的原理.本章就来介绍量子力学的这个基本原理——全同性原理.

§5.5.1　全同性原理

(一)全同性原理

在研究由许多相同的微观粒子组成的体系时,必须考虑到量子力学的又一基本原理——全同性原理.

相同的微观粒子,例如电子,具有完全相同的物理性质.当它们能够用经典力学描述时,具体说来就是,如果它们的德布罗意波长相对于运动的空间范围很短,描述它们的波函数在空间中的重叠很小,以致于可以认为它们各自沿自己的轨道运动时,它们仍不失去自己的"个性".在此情况下,可以给组成系统的每个粒子标一个号码.只要在某一瞬时这样标了号,则在以后可以追随着它们的运动,在任一时刻重新识别每个粒子.

但是,如果组成系统的各个粒子的波函数在空间中有显著的重叠,使得经典近似失效,情况就变了.由于不确定关系,关于粒子"轨道"的概念失去了意义.即使在某一瞬时,给粒子编了号,在另一瞬时仍然无法辨认它们.因此,在量子力学中,相同的粒子完全失去了自己的"个性",无法加以区分,这称为全同性原理.这一原理对由相同的微观粒子组成的多粒子系统的波函数加了很强的限制.

如果粒子有自旋,则描述粒子状态的力学量完全组除了空间变量 $\boldsymbol{r}$ 以外,还有自旋变量 σ.为了书写简单,用一个符号 q 统一表示它们:$q=(\boldsymbol{r},\sigma)$.为了具体起见,考虑由两个相同粒子组成的系统,用 q_1 和 q_2 表示描述这两个粒子状态的变量,用 $\psi(q_1,q_2)$ 表示描述粒子系统状态的波函数.这里,q_i 的脚标 $i=1,2$ 只不过表明系统中有两个粒子,由于全同性原理,无法确认那一个粒子是第 1 粒子,哪一个粒子是第 2 粒子.因此,将 q_1 和 q_2 交换得到的 $\psi(q_2,q_1)$ 和 $\psi(q_1,q_2)$ 描述的是同一状态.

在量子力学中,描述同一状态的波函数可以差一个相因子,而且只能差一个相

因子.因此

$$\psi(q_1,q_2) = e^{i\alpha}\psi(q_2,q_1). \tag{1.1}$$

其中 α 为某一实数,由粒子本身的性质决定.如果再在 $\psi(q_2,q_1)$ 中交换 q_1 和 q_2,也应有 $\psi(q_2,q_1) = e^{i\alpha}\psi(q_1,q_2)$.代入(1.1)式,得到

$$\psi(q_1,q_2) = e^{i\alpha}\psi(q_2,q_1) = e^{2i\alpha}\psi(q_1,q_2).$$

由此可见,$e^{2i\alpha} = 1$,$e^{i\alpha} = \pm 1$.代入(1.1)式,得到

$$\psi(q_1,q_2) = \pm\psi(q_2,q_1). \tag{1.2}$$

即:描述全同粒子系统的波函数,在交换两个粒子的全部变量时(简单说就是:在交换两个粒子时),只能是或者对称——交换后波函数不变,或者反对称——交换后波函数改号.

将全同性原理和叠加原理结合,可以得到一个重要结论:全同粒子系统的波函数,对于两个粒子的交换是对称还是反对称,完全决定于粒子本身的性质,与所处的状态以及系统中粒子的数目无关.

证 用反证法.设 $\psi_{\mathrm{I}}(q_1,q_2)$ 和 $\psi_{\mathrm{II}}(q_1,q_2)$ 分别对粒子交换对称和反对称:$\psi_{\mathrm{I}}(q_1,q_2) = \psi_{\mathrm{I}}(q_2,q_1)$,$\psi_{\mathrm{II}}(q_1,q_2) = -\psi_{\mathrm{II}}(q_2,q_1)$,则由它们叠加成的波函数 $\psi(q_1,q_2) = c_{\mathrm{I}}\psi_{\mathrm{I}} + c_{\mathrm{II}}\psi_{\mathrm{II}}$ 将是既不对称又不反对称:$\psi(q_1,q_2) = c_{\mathrm{I}}\psi_{\mathrm{I}}(q_2,q_1) - c_{\mathrm{II}}\psi_{\mathrm{II}}(q_2,q_1) \neq \pm\psi(q_2,q_1)$,违反了全同性原理.这表明,$\psi_{\mathrm{I}}$ 和 ψ_{II} 对于两粒子的交换必须有相同对称性.又设 $\psi(q_1,q_2,q_3)$ 对头两个粒子的交换对称,对后两个粒子的交换反对称,则有:

$$\begin{aligned}\psi(q_1,q_2,q_3) &= \psi(q_2,q_1,q_3) = -\psi(q_2,q_3,q_1) = -\psi(q_3,q_2,q_1)\\ &= \psi(q_3,q_1,q_2) = \psi(q_1,q_3,q_2) = -\psi(q_1,q_2,q_3).\end{aligned}$$

这样就产生了矛盾,从而表明,$\psi(q_1,q_2,q_3)$ 只能是对于任意两个粒子的交换都对称或者对于任意两个粒子的交换都反对称. **【证毕】**

如果某种粒子所组成的系统的波函数对于粒子的交换对称,就称这种粒子为玻色子;如果某种粒子所组成的系统的波函数对于粒子的交换反对称,就称这种粒子为费米子.采用这样的名称的原因是由子这两种粒子分别服从玻色-爱因斯坦统计和费米-狄拉克统计,见第七篇.

实验表明,自旋为整数或零的粒子,例如光子($s = 1$),ρ 介子($s = 1$),π 介子($s = 0$),都是玻色子;自旋为半整数的粒子,例如:电子、质子、中子($s = 1/2$)或 Δ 粒子($s = 3/2$),都是费米子.复合粒子(例如原子核)是费米子还是玻色子,决定于其中所含费米子的数目——含奇数个费米子的复合粒子是费米子,含偶数个费米子的复合粒子是玻色子.

(二)波函数的对称化和反对称化

考虑由 N 个粒子组成的系统.假定粒子之间的相互作用很微弱,在初级近似

下可以忽略，则可以定义单粒子态：$\psi_1,\psi_2,\cdots$，并假定这些单粒子态已经归一化. 整个系统的态函数是 N 个单粒子态函数的乘积：

$$\psi_{\nu_1}(q_1)\psi_{\nu_2}(q_2)\cdots\psi_{\nu_N}(q_N). \tag{1.3}$$

但是这一乘积对于任意两个粒子的交换 $q_i\rightleftharpoons q_j$，不具有一定的对称性，不满足全同性原理的要求. 为了得到满足全同性原理要求的波函数，需要将(1.3)式对称化(如果所讨论的是玻色子)，或反对称化(如果所考虑的是费米子).

以两个玻色子组成的系统为例，设它们处于以 ν_1,ν_2 为量子数的两个单粒子态. 当 $\nu_1=\nu_2$ 时，$\psi_{\nu_1}(q_1)\psi_{\nu_2}(q_2)$ 对 $q_1\rightleftharpoons q_2$ 的交换对称；因而是符合全同性原理要求的波函数. 当 $\nu_1\neq\nu_2$ 时，作对称组合：

$$\psi_S(q_1,q_2)=\frac{1}{\sqrt{2}}[\psi_{\nu_1}(q_1)\psi_{\nu_2}(q_2)+\psi_{\nu_1}(q_2)\psi_{\nu_2}(q_1)], \tag{1.4}$$

就得到对粒子交换对称的波函数，其中的因子 $1/\sqrt{2}$ 是为了归一化.

如果考虑的是费米子，则应进行反对称化：

$$\psi_A(q_1,q_2)=\frac{1}{\sqrt{2}}[\psi_{\nu_1}(q_1)\psi_{\nu_2}(q_2)-\psi_{\nu_1}(q_2)\psi_{\nu_2}(q_1)], \tag{1.5}$$

这就是对粒子交换反对称的波函数. 显然，当 $\nu_1=\nu_2$ 时，不可能组成反对称波函数，即两个费米子不能处于同一个量子态(泡利不相容原理).

(1.5)式可以写成行列式的形式：

$$\psi_A(q_1,q_2)=\frac{1}{\sqrt{2}}\begin{vmatrix}\psi_{\nu_1}(q_1) & \psi_{\nu_1}(q_2)\\ \psi_{\nu_2}(q_1) & \psi_{\nu_2}(q_2)\end{vmatrix}. \tag{1.6}$$

推广到 N 个费米子系统，其波函数可以写成 N 阶行列式：

$$\psi_A=\frac{1}{\sqrt{N!}}\begin{vmatrix}\psi_{\nu_1}(q_1) & \psi_{\nu_1}(q_2) & \cdots & \psi_{\nu_1}(q_N)\\ \psi_{\nu_2}(q_1) & \psi_{\nu_2}(q_2) & \cdots & \psi_{\nu_2}(q_N)\\ \cdots & \cdots & \cdots & \cdots\\ \psi_{\nu_N}(q_1) & \psi_{\nu_N}(q_2) & \cdots & \psi_{\nu_N}(q_N)\end{vmatrix}. \tag{1.7}$$

两个粒子的交换相当于这个行列式两列的交换，使行列式改号. 所以这一波函数对于任意两个粒子的交换有反对称性.

由(1.7)式可见，如果有两个粒子处子相同的量子态：$\nu_i=\nu_j$，则行列式的第 i 和第 j 行相同，使得行列式为零. 因此，由费米子组成的系统中，不可能有两个或两个以上的粒子处于同一量子态. 这就是泡利不相容原理. 因此泡利不相容原理是全同性原理的一个推论.

对于 N 个全同玻色子构成的系统，对称化的结果为

$$\psi_S(q_1,q_2,\cdots,q_N) = c\sum_P P\psi_{\nu_1}(q_1)\psi_{\nu_2}(q_2)\cdots\psi_{\nu_N}(q_N), \tag{1.8}$$

P 表示 N 个单粒子波函数的某一种排列，$\sum_P$ 表示对所有可能的排列求和. c 是归一化常数.

(三)福克表象

对称化和反对称化的波函数(1.8)式和(1.7)式不便于应用，因而对全同多粒子系统常采用另一种表象.

在(1.8)式和(1.7)式中，是先标明第 i 个粒子所处的状态 $\nu_i(i=1,2,\cdots,N)$，然后再进行对称化或反对称化. 为了避免这一复杂性，用另一办法来描述系统的状态. 那就是：将单粒子态编号：$1,2,\cdots,\nu,\cdots$，指明在各个态中的粒子数：

$$\begin{array}{cccccc} \text{单粒子态} & 1 & 2 & \cdots & \nu & \cdots \\ \text{粒子数} & n_1 & n_2 & \cdots & n_\nu & \cdots \end{array} \tag{1.9}$$

这样就描述了全同多粒子系统的状态. 显然，

$$\sum_{\nu=1}^{\infty} n_\nu = N. \tag{1.10}$$

这种描述方法从一开始就不给粒子编号，因而自动满足全同性原理的要求.

描述全同多粒子系统状态的这一方法所用的态矢是

$$|\, n_1,n_2,\cdots,n_\nu,\cdots\rangle. \tag{1.11}$$

相应的表象称为福克表象.

态矢(1.11)既适用于玻色子又适用于费米子. 差别在于，对玻色子 $n_\nu = 0,1,2,\cdots$ 没有限制；对费米子则只能有 $n_\nu = 0,1$. 这一差别通过作用在态矢(1.11)上的算符体现出来.

算符对态矢(1.11)作用的结果，改变了其中的 $n_\nu(\nu=1,2,\cdots)$，使它们减少或增加. 基元算符是使 n_ν 减少 1 的算符 a_ν 称为湮灭算符，和使 n_ν 增加 1 的算符 $a_\nu^\dagger$ 称为产生算符：

$$a_\nu |\, n_1,n_2,\cdots,n_\nu,\cdots\rangle = c_\nu |\, n_1,n_2,\cdots,n_\nu-1,\cdots\rangle, \tag{1.12}$$

$$a_\nu^\dagger |\, n_1,n_2,\cdots,n_\nu,\cdots\rangle = c'_\nu |\, n_1,n_2,\cdots,n_\nu+1,\cdots\rangle \tag{1.13}$$

其中 c_ν 和 c'_ν 是两个常数. 和 §5.4.5(5.12)及(5.6)式比较可见，只要有对易关系

$$[a_\mu,a_\nu] = [a_\mu^\dagger,a_\nu^\dagger] = 0, \qquad [a_\mu,a_\nu^\dagger] = \delta_{\mu\nu}. \tag{1.14}$$

(1.13)式就能满足，而且单粒子态 ν 中的粒子数算符

$$N_\nu = a_\nu^\dagger a_\nu \tag{1.15}$$

的本征值 $n_\nu = 0,1,2,\cdots$ [见 §5.4.5(5.16)式]，符合玻色子的要求.

对于费米子，n_ν 只能取 0 和 1 两个值. 用 b_ν 和 $b_\nu^\dagger$ 表示相应的消灭和产生算符. 我们有

$$b_\nu^\dagger \mid 0,0,\cdots,n_\nu = 0,\cdots\rangle = c_\nu \mid 0,0,\cdots,n_\nu = 1,\cdots\rangle. \tag{1.16}$$

其中 $\mid 0,0,\cdots,n_\nu = 0,\cdots\rangle$ 是没有粒子的状态，称为真空态，c_ν 是另一常数. 如果再用 $b_\nu^\dagger$ 作用一次，将会使 n_ν 变为 2，而这是不可能的. 因而必须有 $(b_\nu^\dagger)^2 = 0$. 取共轭得 $(b_\nu)^2 = 0$.

当 $\mu \neq \nu$ 时，$b_\mu^\dagger b_\nu^\dagger$ 作用到真空态 $\mid 0,0,\cdots,n_\mu = 0,\cdots,n_\nu = 0,\cdots\rangle$ 上会得到 $n_\mu = n_\nu = 1$ 的态，而 $b_\nu^\dagger b_\mu^\dagger$ 作用到真空态上也会得到同一状态，但是两个粒子交换了. 根据费米子的性质，所得结果应该相差符号. 因而有

$$b_\mu^\dagger b_\nu^\dagger = - b_\nu^\dagger b_\mu^\dagger, \qquad b_\mu^\dagger b_\nu^\dagger + b_\nu^\dagger b_\mu^\dagger = 0. \tag{1.17}$$

取共轭得

$$b_\mu b_\nu = - b_\nu b_\mu, \qquad b_\mu b_\nu + b_\nu b_\mu = 0. \tag{1.17'}$$

当 $\mu = \nu$ 时回到前面讨论过的 $(b_\nu^\dagger)^2 = 0, (b_\nu)^2 = 0$.

由(1.17)式可见，$b_\mu^\dagger b_\nu^\dagger$(或 $b_\mu b_\nu$) 交换后改号. 这种交换后改号的算符称为反对易算符，以区别于交换后不变的可对易算符. 定义对易子和反对易子为

$$[A,B] \equiv AB - BA; \qquad \{A,B\} \equiv AB + BA, \tag{1.18}$$

可以将(1.17)式写成类似于(1.14)式的头两个等式的形式

$$\{b_\mu, b_\nu\} = \{b_\mu^\dagger, b_\nu^\dagger\} = 0. \tag{1.19}$$

和(1.14)式的后一等式对应，有

$$\{b_\mu, b_\nu^\dagger\} = \delta_{\mu\nu}. \tag{1.19'}$$

容易证明，相应的粒子数算符 $N_\nu = b_\nu^\dagger b_\nu$ 的本征值为 0 和 1.

证　取粒子数算符的平方 $N_\nu^2 = b_\nu^\dagger b_\nu b_\nu^\dagger b_\nu$. 由(1.19')式有 $b_\nu b_\nu^\dagger = 1 - b_\nu^\dagger b_\nu$ 代入上式，利用 $(b_\nu^\dagger)^2 = b_\nu^2 = 0$，得到 $N_\nu^2 = N_\nu$. 其本征值满足 $n_\nu^2 = n_\nu$，因而有 $n_\nu = 0,1$. **【证毕】**

我们看到，满足对易关系(1.14)式的算符 $a_\nu, a_\nu^\dagger$ 是玻色子的湮灭算符和产生算符；满足反对易关系(1.19)(1.19')式的算符 $b_\nu, b_\nu^\dagger$ 是费米子的湮灭算符和产生算符. 全同多粒子系统中一切力学量的算符都可以用它们表示出来. 对于全同多粒子系统的这样一种描述方式称为二次量子化.

在本节的最后，让我们再次强调：以上说到的“两个粒子交换”是指描述这两个粒子的全部变量的交换，其中包括坐标变量和自旋变量. 坐标变量的函数称为“轨道”波函数，前面已经讨论得很多，下面来讨论电子的自旋波函数.

§5.5.2　电子自旋波函数对粒子交换的对称性

在 §5.4.6 中已经指出，电子的自旋 $s = 1/2$，因而 $s_z = \pm \hbar/2$，在 $\hat{s}^2$ 和 $\hat{s}_z$ 表象

中，自旋算符是

$$\hat{s} = \frac{\hbar}{2}\boldsymbol{\sigma},$$
$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.1}$$

如§5.4.6(6.29)、(6.30)式. 在这一表象中，自旋波函数是二行一列矩阵，如§5.4.6(6.32)式. 我们把它写成：

$$\chi = \begin{pmatrix} \chi(1/2) \\ \chi(-1/2) \end{pmatrix} = \chi(1/2)\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \chi(-1/2)\begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.2}$$

其中，$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ 和 $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ 分别是 $\hat{s}_z$ 的本征值为 $\pm\hbar/2$ 的本征函数，见§5.4.6(6.31)式. $P_\sigma = |\chi(\sigma)|^2, \sigma = \pm 1/2$，分别是在状态 χ 中 s_z 取值 $\pm\hbar/2$ 的概率. 由此可见，自旋波函数 $\chi(\sigma)$ 中的变量 σ 就是 s_z 所取的值（以 $\hbar$ 为单位）. 这和轨道波函数 $\psi(x)$ 中的变量 x 是坐标所取的值一样，只不过坐标变量能取无穷多个连续值，而自旋变量只能取两个分立值.

描述电子的全部变量 q 包括坐标变量 $\boldsymbol{r}$ 和自旋变量 σ，因而§5.5.1(1.2)式应写成（注意电子是自旋为半整数的费米子）

$$\psi(\boldsymbol{r}_1,\sigma_1;\boldsymbol{r}_2,\sigma_2) = -\psi(\boldsymbol{r}_2,\sigma_2;\boldsymbol{r}_1,\sigma_1). \tag{2.3}$$

在通常情况下，自旋和“轨道”运动的相互作用（称为旋轨作用）很微弱，在初级近似中可以忽略，而将电子的总波函数写成轨道波函数和自旋波函数的乘积：

$$\psi(\boldsymbol{r}_1,\sigma_1;\boldsymbol{r}_2,\sigma_2) = \psi(\boldsymbol{r}_1,\boldsymbol{r}_2)\,\chi(\sigma_1,\sigma_2). \tag{2.4}$$

这里对总波函数和轨道波函数采用同一符号 ψ，靠里面的变量来区分，因而不会混淆. 在此情况下，(2.3)式成为：

$$\psi(\boldsymbol{r}_1,\boldsymbol{r}_2)\,\chi(\sigma_1,\sigma_2) = -\psi(\boldsymbol{r}_2,\boldsymbol{r}_1)\,\chi(\sigma_2,\sigma_1). \tag{2.5}$$

为了满足此式，有两种可能：

$$\left.\begin{aligned} &\psi(\boldsymbol{r}_1,\boldsymbol{r}_2) = -\psi(\boldsymbol{r}_2,\boldsymbol{r}_1), \quad && \chi(\sigma_1,\sigma_2) = \chi(\sigma_2,\sigma_1); \\ &\psi(\boldsymbol{r}_1,\boldsymbol{r}_2) = \psi(\boldsymbol{r}_2,\boldsymbol{r}_1), \quad && \chi(\sigma_1,\sigma_2) = -\chi(\sigma_2,\sigma_1). \end{aligned}\right\} \tag{2.6}$$

换句话说，二电子系统的波函数如果对两个电子坐标的交换反对称，则对自旋变量的交换对称；如果对坐标的交换对称，则对自旋变量的交换反对称. 注意，尽管电子是费米子，如果只看轨道波函数或只看自旋波函数，则对其中变量的交换都有对称和反对称两种可能；通过轨道波函数和自旋波函数有相反的对称性，保证了总波函数满足全同性原理对反对称性的要求.

如果两个电子的自旋相互作用可以忽略，则可以把 $\chi(\sigma_1,\sigma_2)$ 写成两个单电子自旋波函数的乘积. 对于单电子的自旋波函数可以用 s_z 的本征函数

$$\chi_{1/2}(\sigma)=\delta_{\sigma,1/2},\qquad \chi_{-1/2}(\sigma)=\delta_{\sigma,-1/2} \tag{2.7}$$

作为正交归一的基底.[将(2.7)式按(2.2)式排成矩阵,就是§5.4.6的(6.31)式].这样,二电子系统的基底波函数有四个:

$$\begin{aligned}&\chi_{1/2}(\sigma_1)\chi_{1/2}(\sigma_2),\qquad \chi_{1/2}(\sigma_1)\chi_{-1/2}(\sigma_2),\\&\chi_{-1/2}(\sigma_1)\chi_{1/2}(\sigma_2),\qquad \chi_{-1/2}(\sigma_1)\chi_{-1/2}(\sigma_2).\end{aligned} \tag{2.8}$$

其中第一个和第四个是对称波函数,中间两个需要对称化和反对称化.这样就得到三个对称波函数和一个反对称波函数:

$$\left.\begin{aligned}\chi_1^{(s)}(\sigma_1,\sigma_2)&=\chi_{1/2}(\sigma_1)\chi_{1/2}(\sigma_2),\\\chi_2^{(s)}(\sigma_1,\sigma_2)&=\frac{1}{\sqrt{2}}[\chi_{1/2}(\sigma_1)\chi_{-1/2}(\sigma_2)+\chi_{-1/2}(\sigma_1)\chi_{1/2}(\sigma_2)],\\\chi_3^{(s)}(\sigma_1,\sigma_2)&=\chi_{-1/2}(\sigma_1)\chi_{-1/2}(\sigma_2),\\\chi^{(a)}(\sigma_1,\sigma_2)&=\frac{1}{\sqrt{2}}[\chi_{1/2}(\sigma_1)\chi_{-1/2}(\sigma_2)-\chi_{-1/2}(\sigma_1)\chi_{1/2}(\sigma_2)].\end{aligned}\right\} \tag{2.9}$$

定义二电子系统的总自旋算符为

$$\hat{s}=\hat{s}_1+\hat{s}_2, \tag{2.10}$$

其中 $\hat{s}_1$ 和 $\hat{s}_2$ 分别是两个电子的自旋算符,它们都具有(2.1)式的形式,但是分别作用在变量 σ_1 和 σ_2 上.我们来证明(2.9)式中的四个波函数都是 $\hat{S}^2$ 和 $\hat{S}_z$ 的共同本征函数,并求相应的本征值.

证　首先计算

$$\begin{aligned}\hat{S}_z\chi_1^{(s)}&=\hat{s}_{z_1}\chi_{1/2}(\sigma_1)\chi_{1/2}(\sigma_2)+\chi_{1/2}(\sigma_1)\hat{s}_{z_2}\chi_{1/2}(\sigma_2)\\&=(\hbar/2)\chi_{1/2}(\sigma_1)\chi_{1/2}(\sigma_2)+\chi_{1/2}(\sigma_1)[(\hbar/2)\chi_{1/2}(\sigma_2)]=\hbar\chi_1^{(s)}.\end{aligned} \tag{2.11}$$

因此 $\chi_1^{(s)}$ 是 $\hat{S}_z$ 的本征态,本征值为 $\hbar$.同样可以证明 $\chi_2^{(s)}$,$\chi_3^{(s)}$ 和 $\chi^{(a)}$ 都是 $\hat{S}_z$ 的本征态,对应的本征值分别为 0,$-\hbar$ 和 0.

在计算 $\hat{S}^2$ 的本征值时,应注意由于两个电子的自旋无相互作用(在零级近似下),$\hat{s}_1$ 的各个分量和 $\hat{s}_2$ 的各个分量之间均可对易,因而

$$\hat{S}^2=\hat{s}_1^2+2\hat{s}_1\cdot\hat{s}_2+\hat{s}_2^2=\hat{s}_1^2+\hat{s}_2^2+2\hat{s}_{1z}\hat{s}_{2z}+2(\hat{s}_{1x}\hat{s}_{2x}+\hat{s}_{1y}\hat{s}_{2y}). \tag{2.12}$$

容易看到

$$\hat{s}_1^2\chi_1^{(s)}=\hat{s}_1^2\chi_{\frac{1}{2}}(\sigma_1)\chi_{\frac{1}{2}}(\sigma_2)=(1/2)(3/2)\hbar^2\chi_{\frac{1}{2}}(\sigma_1)\chi_{\frac{1}{2}}(\sigma_2)=(3/4)\hbar^2\chi_1^{(s)}.$$

$\hat{s}_2^2\chi_1^{(s)}$ 也有同样结果.而

$$2\hat{s}_{1z}\hat{s}_{2z}\chi_1^{(s)}=2\hat{s}_{1z}\chi_{\frac{1}{2}}(\sigma_1)\hat{s}_{2z}\chi_{\frac{1}{2}}(\sigma_2)=2\frac{\hbar}{2}\chi_{\frac{1}{2}}(\sigma_1)\frac{\hbar}{2}\chi_{\frac{1}{2}}(\sigma_2)=\frac{\hbar^2}{2}\chi_1^{(s)}.$$

在以上的计算中,处处都用到了单粒子自旋波函数(2.7)式是单粒子自旋算符 $\hat{S}^2$ 和 $\hat{S}_z$ 的本征态这一事实,使计算简化.

在计算 $2\hat{s}_{1x}\hat{s}_{2x}$ 和 $2\hat{s}_{1y}\hat{s}_{2y}$ 的作用时,必须用矩阵运算,我们有:

$$\left.\begin{aligned}
\hat{s}_x\chi_{1/2} &= \frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\cdot\begin{pmatrix}1\\0\end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix}0\\1\end{pmatrix} = \frac{\hbar}{2}\chi_{-1/2},\\
\hat{s}_x\chi_{-1/2} &= \frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\cdot\begin{pmatrix}0\\1\end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix}1\\0\end{pmatrix} = \frac{\hbar}{2}\chi_{1/2},\\
\hat{s}_y\chi_{1/2} &= \frac{\hbar}{2}\begin{pmatrix}0 & -\mathrm{i}\\ \mathrm{i} & 0\end{pmatrix}\cdot\begin{pmatrix}1\\0\end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix}0\\\mathrm{i}\end{pmatrix} = \frac{\hbar}{2}\mathrm{i}\chi_{-1/2},\\
\hat{s}_y\chi_{-1/2} &= \frac{\hbar}{2}\begin{pmatrix}0 & -\mathrm{i}\\ \mathrm{i} & 0\end{pmatrix}\cdot\begin{pmatrix}0\\1\end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix}-\mathrm{i}\\0\end{pmatrix} = -\frac{\hbar}{2}\mathrm{i}\chi_{1/2}.
\end{aligned}\right\} \tag{2.13}$$

利用这几个公式,不难得到

$$2\hat{s}_{1x}\hat{s}_{2x}\chi_1^{(s)} = 2\hat{s}_{1x}\chi_{1/2}(\sigma_1)\hat{s}_{2x}\chi_{1/2}(\sigma_2) = \frac{\hbar^2}{2}\chi_{-1/2}(\sigma_1)\chi_{-1/2}(\sigma_2),$$

$$2\hat{s}_{1y}\hat{s}_{2y}\chi_1^{(s)} = 2\hat{s}_{1y}\chi_{1/2}(\sigma_1)\hat{s}_{2y}\chi_{1/2}(\sigma_2) = \frac{\hbar^2}{2}\mathrm{i}\chi_{-1/2}(\sigma_1)\mathrm{i}\chi_{-1/2}(\sigma_2).$$

两者符号正好相反,相加以后消去.集合以上结果得到:

$$\hat{S}^2\chi_1^{(s)} = \left(\frac{3}{2}\hbar^2 + \frac{\hbar^2}{2}\right)\chi_1^{(s)} = = 2\hbar^2\chi_1^{(s)} = 1\cdot(1+1)\hbar^2\chi_1^{(s)}.$$

我们看到,$\chi_1^{(s)}$ 是 $\hat{S}^2$,$\hat{S}_z$ 的共同本征态,本征值分别是 $1\cdot(1+1)\hbar^2$ 和 $1\hbar$,即 $S = 1, S_z = 1$.

类似地可以证明(2.9)式中的另外两个波函数 $\chi_2^{(s)}$ 和 $\chi_3^{(s)}$ 分别是 $\hat{S}^2$ 和 $\hat{S}_z$ 的共同本征态,对应 $S = 1$ 和 $S_z = 0, -1$;而(2.9)式中的 $\chi^{(a)}$ 是 $\hat{S}^2$ 和 $\hat{S}_z$ 的本征态,对应 $S = 0$ 和 $S_z = 0$(见习题2),因而可以写

$$\chi_1^{(s)} = \chi_{11}, \quad \chi_2^{(s)} = \chi_{10}, \quad \chi_3^{(s)} = \chi_{1-1}, \quad \chi^{(a)} = \chi_{00}. \tag{2.14}$$

此式右边各个 χ 的下标分别表示 S 和 S_z.

这一结果有明显的几何解释:具有交换对称的 $\chi^{(s)}$ 描述两个电子自旋平行的态,因而 $S = 1$,它有三种取向:$S_z = \pm 1, 0$;具有交换反对称的 $\chi^{(a)}$ 描述两个电子自旋反平行的态,因而 $S = 0, S_z = 0$.

由(2.6)式:

$$\left.\begin{aligned}
&\text{如果 } S = 1, \quad \text{则} \quad \chi = \chi^{(s)} \Rightarrow \psi = \psi^{(a)};\\
&\text{如果 } S = 0, \quad \text{则} \quad \chi = \chi^{(a)} \Rightarrow \psi = \psi^{(s)}.
\end{aligned}\right\} \tag{2.15}$$

这就是说:如果二电子系统的总自旋 $S = 1$,则轨道波函数对于坐标交换反对称;如果二电子系统的总自旋 $S = 0$,则轨道波函数对于坐标交换对称.一般情况下,

$\psi^{(a)}(\boldsymbol{r}_1,\boldsymbol{r}_2)$ 和 $\psi^{(s)}(\boldsymbol{r}_1,\boldsymbol{r}_2)$ 有不同能量. 因此,虽然自旋和轨道没有直接的相互作用,但由于全同性原理,二电子系统的自旋状态却影响了轨道运动的能量,这称为交换相互作用.

习　题

1. 两个粒子在同一势场中作一维简谐振动,分析这一系统的基态和第一激发态的能量简并情况. 若粒子间有和相对距离成正比的斥力作用,求上述两个能级能量的变化,若这两个粒子是自旋为 1/2 的全同粒子,求在各个能级系统的总自旋.

2. 证明 §5.5.2(2.9)式中的二电子自旋波函数是 S^2 和 S_z 的共同本征态,并求相应的本征值.

3. 将第四章习题 14 所得结果和 §5.5.2(2.9)式比较,并加以讨论.

4. 二电子系统的一般自旋波函数是 §5.5.2(2,8)式中四个函数的叠加

$$\chi = c_1\begin{pmatrix}1\\0\end{pmatrix}^{(1)}\cdot\begin{pmatrix}1\\0\end{pmatrix}^{(2)} + c_2\begin{pmatrix}1\\0\end{pmatrix}^{(1)}\cdot\begin{pmatrix}0\\1\end{pmatrix}^{(2)} + c_3\begin{pmatrix}0\\1\end{pmatrix}^{(1)}\cdot\begin{pmatrix}1\\0\end{pmatrix}^{(2)} + c_4\begin{pmatrix}0\\1\end{pmatrix}^{(1)}\cdot\begin{pmatrix}0\\1\end{pmatrix}^{(2)}$$

试解: $\hat{S}_z = \hat{S}_{z1} + \hat{S}_{z2}$ 的本征值方程 $\hat{S}_z\chi = S_z\chi$,求本征值和本征函数,并分析所得本征函数在两个电子自旋变量交换时的对称性质.

5. 总结量子力学的基本原理. 要求语言简洁内容充实,清晰地描绘出刻画在你大脑里的量子力学理论框架.

第六章　量子干涉现象

正如§5.1.2中指出的，微观粒子的波动性主要表现在其传播过程中有干涉现象，即：在波函数的叠加过程中，由于相位差而产生的加强或减弱现象. 在以上各章节中一般都没有涉及波的干涉，因而看不出波函数的相位的作用，这就是为什么常常可以简单地将相位取为零，例如§5.4.5(5.25)式、§5.4.6(6.19)～(6.22)式. 然而，在涉及干涉现象时，波函数的相位（更准确地说是不同波函数之间的相位差）起决定性作用. 本节就来讨论波函数叠加时的干涉现象.

§5.6.1　路径积分量子化

我们知道，由于微观粒子的波动性，不存在经典粒子那样确定的位置 $\boldsymbol{r}$，但是仍然可以将经典粒子关于"位置"的概念借用过来，给它赋与概率的意义，认为波函数 $\Psi(\boldsymbol{r},t)$ 的绝对值平方是粒子处于 $\boldsymbol{r}$ 处的概率密度.

由于不确定关系，如果在某一时刻 t 粒子处于位置 $\boldsymbol{r}$，则其动量（速度）的大小和方向完全不确定，因而对于微观粒子，不存在"运动路径"的概念. 然而，正像为了描述微观粒子，虽然它不存在确定的位置，仍然可以在概率的意义上将经典粒子关于"位置"的概念借用过来一样；对于微观粒子也可以借用经典粒子关于"路径"的概念，只要给它赋与概率的意义. 具体说来，如果粒子在 t_A 时刻位于 $\boldsymbol{r}_A$，在 t_B 时刻位于 $\boldsymbol{r}_B$，则对于经典粒子，路径 $\bar{\boldsymbol{r}}(t)$ 是完全确定的[这里用 $\boldsymbol{r}(t)$ 上加横线表示经典路径]；而如果是微观粒子，则各种路径都是可能的，每条路径各有自己的概率幅.

费曼假定，由 A 点 $(\boldsymbol{r}_A,t_A)$ 到 B 点 $(\boldsymbol{r}_B,t_B)$ 之间的一条路径 $\boldsymbol{r}(t)$ 的概率幅是

$$\phi[\boldsymbol{r}(t)] = A^{-1}\mathrm{e}^{(\mathrm{i}/\hbar)S[\boldsymbol{r}(t)]}, \tag{1.1}$$

其中

$$S[\boldsymbol{r}(t)] = \int_{t_A}^{t_B} L(\boldsymbol{r},\dot{\boldsymbol{r}},t)\,\mathrm{d}t \tag{1.2}$$

是满足条件

$$\boldsymbol{r}(t_A) = \boldsymbol{r}_A, \qquad \boldsymbol{r}(t_B) = \boldsymbol{r}_B \tag{1.3}$$

的路径 $\boldsymbol{r}(t)$ 的作用量，$L(\boldsymbol{r},\dot{\boldsymbol{r}},t)$ 是拉格朗日函数，参看§1.1.3图1及该节(3.5)式. (1.1)式中的 A^{-1} 是比例系数.

由(1.1)式显然可见，为了由 A 点到 B 点，一切符合条件(1.3)式的路径都是等概率的，$|\phi[\boldsymbol{r}(t)]|^2$ = 常数，只是相位不同. 初看起来，这似乎是不可能的. 因

为当系统很大,以致普朗克常量 $\hbar$ 可以忽略不计时,量子力学应该过渡到经典力学,那时只有一条路径[经典路径 $r(t)$]有 100% 的概率,其余路径的概率均为零.

产生这一表观上的矛盾的原因是由于,满足条件(1.3)的路径是连续分布的. 正像由于位矢 r 连续,所以 $|\Psi(r,t)|^2$ 是粒子位于 r 的概率密度,而不是概率一样;$|\phi[r(t)]|^2$ 也是路径 $r(t)$ 的概率密度,而不是概率. 同样,正像我们不能问粒子位于 r 的概率,而只能问粒子位于 $r \sim r + dr$ 之间的概率一样;我们也不能问 $r(t)$ 这一条路径的概率,而只能问 $r(t) \sim r(t) \pm \Delta r(t)$ 之间的所有路径的总概率,参看图 1. 这一概率是

$$P[r(t)] = A^{-1} \sum_{\Delta r(t)} e^{(i/\hbar)S[r(t)+\delta r(t)]}. \tag{1.4}$$

符号 $\sum_{\Delta r(t)}$ 表示对于从路径 $r(t)$ 到限制在 $r(t) \pm \Delta r(t)$ 之间的所有路径 $r(t) + \delta r(t)$ 求和❶.

当系统很大,以至作用量 S 远大于 $\hbar$ 时,由 $\delta r(t)$ 引起的 S 的改变 δS 一般说来也远大于 $\hbar$. 此时,在(1.4)式的求和中,不同路径 $r(t) + \delta r(t)$ 的相位差很大,相互抵消,使 $P[r(t)] = 0$. 只有当 $r(t)$ 使作用量 S 有极值

$$\delta S = 0 \tag{1.5}$$

时,(1.4)式的求和中,不同路径才在一级近似下有相同的相位,而相互加强. 因此,当系统很大时,在(1.4)式的求和中只有满足条件(1.5)式的路径的概率是 100%,而其他路径的概率都是零.

比较(1.5)式和 §1.1.3(3.6)式可见,满足此式的路径正是由"最小作用量原理"决定的经典路径. 由此可见,费曼的假定(1.1)式能自然地过渡到经典力学,而且在这个意义上,经典力学的"最小作用量原理"不再是独立的"原理",而是费曼原理的一个推论.

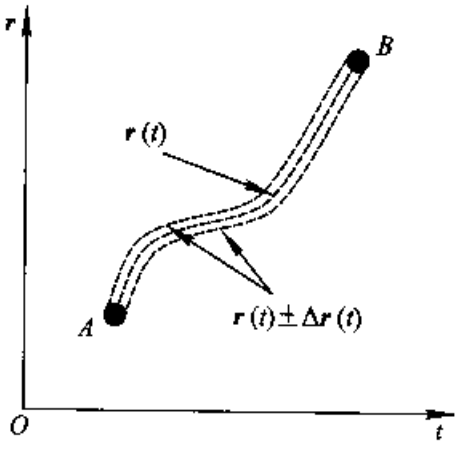


图 1　路径 $r(t)$ 和偏离它在 $\Delta r(t)$ 之内的路径组

我们看到,当系统很大时,费曼原理自然地过渡到经典力学. 还需要进一步考查的是,当系统很小时,费曼原理是否和由正则量子化得到的量子力学——薛定谔方程一致. 下面就来讨论这一问题. 为了书写简单,只考虑一维情况.

首先,让我们来考虑 t_A 时刻粒子在 x_A,而 t_B 时刻到达 x_B 的概率幅 $K(x_B, t_B;$

❶ 由于 $r(t)$ 连续分布,这一"求和"实际上是对路径 $r(t) + \delta r(t)$ 的积分,参看下节的(2.3)式.

$x_A,t_A)$. 它称为由 (x_A,t_A) 到 (x_B,t_B) 的传播函数，是满足(1.3)式的一切路径的概率幅 $\phi[x(t)] = A^{-1}e^{(i/\hbar)S[x(t)]}$ 之和：

$$K(x_B,t_B;x_A,t_A) = \sum_{\text{满足(1.3)的一切}x(t)} \phi[x(t)]. \tag{1.6}$$

让我们注意，波函数 $\Psi(x,t)$ 的物理意义是粒子在 t 时刻位于 x 的概率幅，而传播函数 $K(x,t;x_A,t_A)$ 的意义也是 t 时刻位于 x 的概率幅. 因此，传播函数 $K(x,t;x_A,t_A)$ 也是一个波函数，只不过它带有初条件——t_A 时刻有确定的位置 x_A（参看第五篇第三章习题3）. 一般的波函数 $\Psi(x,t)$ 应该有任意初条件，这种任意初条件可以用“t_A 时刻位于 x_A 的概率幅是 $\Psi(x_A,t_A)$”来表示. 将 t_A 时刻有确定位置 x_A 的“波函数” $K(x,t;x_A,t_A)$ 乘上 t_A 时刻位于 x_A 的概率幅为 $\Psi(x_A,t_A)$ 的初条件，再对 x_A 积分就得到具有任意初条件的波函数

$$\Psi(x,t) = \int_{-\infty}^{\infty} K(x,t;x_A,t_A)\Psi(x_A,t_A)\,dx_A. \tag{1.7}$$

为了和微分形式的量子力学运动方程——薛定锷方程比较，考虑经过无穷小时间间隔 ε 的传播. 此时传播函数 $K(x,t_A+\varepsilon;x_A,t_A)$ 可以近似地用 $\phi[x(t)] = A^{-1}e^{(i/\hbar)S[x(t)]}$ 代替：

$$K(x,t_A+\varepsilon;x_A,t_A) \approx A^{-1}e^{(i/\hbar)\int_{t_A}^{t_A+\varepsilon}L(x,\dot{x},t)\,dt}. \tag{1.8}$$

由于 ε 很小，可以按中值定理将被积函数 $L(x,\dot{x},t)$ 中的变量 x 取为积分区间中的平均值 $(x+x_A)/2$，而 $\dot{x}$ 取为 $(x-x_A)/\varepsilon$，

$$K(x,t_A+\varepsilon;x_A,t_A) \approx A^{-1}e^{(i\varepsilon/\hbar)L((x+x_A)/2,(x-x_A)/\varepsilon,t_A+\varepsilon/2)}. \tag{1.9}$$

代入(1.7)式并改写 t_A 为 t，得到

$$\Psi(x,t+\varepsilon) = A^{-1}\int_{-\infty}^{\infty} e^{(i\varepsilon/\hbar)L((x+x_A)/2,(x-x_A)/\varepsilon,t+\varepsilon/2)}\Psi(x_A,t)\,dx_A. \tag{1.10}$$

考虑在势能场 $U(x)$ 中运动的粒子，$L = m\dot{x}^2/2 - U(x)$，代入(1.10)式，得到

$$\Psi(x,t+\varepsilon) = A^{-1}\int_{-\infty}^{\infty} e^{(i\varepsilon/\hbar)[m(x-x_A)^2/\varepsilon^2 - U((x+x_A)/2)]}\Psi(x_A,t)\,dx_A. \tag{1.11}$$

由于 ε 很小，除了在 $(x-x_A)^2 \approx \varepsilon$ 的一个小区间中以外，$m(x-x_A)^2/\varepsilon$ 都很大，在对 dx_A 积分时，虚指数函数 $e^{(i/\hbar)m(x-x_A)^2/\varepsilon}$ 振荡很快，使积分为零. 因此，对积分有贡献的只是 $(x-x_A)^2 = \eta^2$ 是 ε 的同级小量的区间. 写 $x_A = x+\eta$，得到

$$\Psi(x,t+\varepsilon) = A^{-1}\int_{-\infty}^{\infty} e^{im\eta^2/2\hbar\varepsilon}e^{-(i\varepsilon/\hbar)U(x+\eta/2)}\Psi(x+\eta,t)\,d\eta. \tag{1.12}$$

由第一个虚指数因子可见，η^2 和 ε 是同级小量. 可以将(1.12)式的左右两边展开到 ε 的一级和 η 的二级，得到

$$\Psi(x,t) + \varepsilon\frac{\partial\Psi}{\partial t} = A^{-1}\int_{-\infty}^{\infty} e^{im\eta^2/2\hbar\varepsilon}\left[1 - \frac{i\varepsilon}{\hbar}U(x,t)\right]\times$$

$$\left[\Psi(x,t) + \eta\frac{\partial\Psi}{\partial x} + \frac{1}{2}\eta^2\frac{\partial^2\Psi}{\partial x^2}\right]\mathrm{d}\eta. \tag{1.13}$$

由附录(V.3)~(V.6)式有

$$\int_{-\infty}^{\infty} e^{im\eta^2/2\hbar\varepsilon}\mathrm{d}\eta = \left(\frac{2\pi i\hbar\varepsilon}{m}\right)^{1/2}; \qquad \int_{-\infty}^{\infty} e^{im\eta^2/2\hbar\varepsilon}\eta\mathrm{d}\eta = 0;$$

$$\int_{-\infty}^{\infty} e^{im\eta^2/2\hbar\varepsilon}\eta^2\mathrm{d}\eta = \frac{i\hbar\varepsilon}{m}\left(\frac{2\pi i\hbar\varepsilon}{m}\right)^{1/2}. \tag{1.14}$$

代入(1.13)式，ε 的零级项是

$$\Psi(x,t) = A^{-1}\left(\frac{2\pi i\hbar\varepsilon}{m}\right)^{1/2}\Psi(x,t)$$

由此可见，

$$A = \left(\frac{2\pi i\hbar\varepsilon}{m}\right)^{1/2}. \tag{1.15}$$

而由(1.13)式中 ε 的一级项得到

$$i\hbar\partial\Psi/\partial t = -[\hbar^2/(2m)]\partial\Psi^2/\partial x^2 + U(x,t)\Psi.$$

这正是薛定谔方程. 由此可见，由费曼关于路径概率幅的假定(1.1)式实现的量子化——称为“路径积分量子化”，和由正则变量的对易关系实现的量子化——“正则量子化”，是等效的.

* §5.6.2 双缝衍射

我们来应用费曼的路径积分量子化研究双缝衍射实验. 这一理想实验的示意图如图 1. 电子由 A 点发出，经过双缝 B_1 和 B_2 到达屏 C 上，发生干涉.

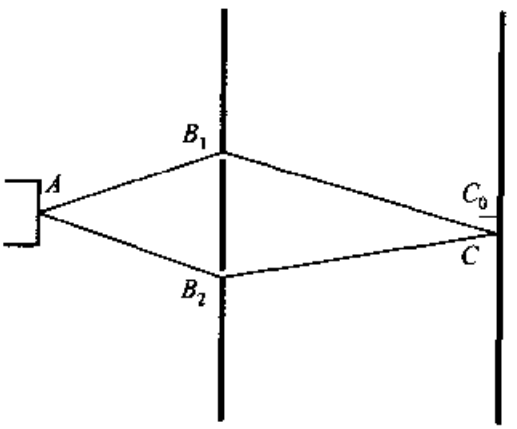


图 1 双缝衍射的示意图

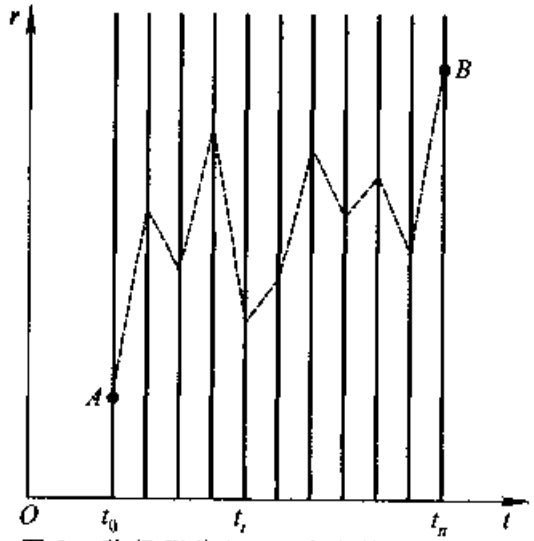


图 2 路径积分(2.8)式中的一条路径

首先讨论电子在 t_A 时刻由 A 出发，在 t_B 时刻到达 B（B 表示 B_1 或 B_2）

$$r(t_A) = r_A, \qquad r(t_B) = r_B \tag{2.1}$$

的概率幅，即由 (A,t_A) 到 (B,t_B) 的传播函数 §5.6.1(1.6)式. 我们将它写成 $\langle B \mid A\rangle$. 利用 §5.6.1(1.1)，(1.6)和(1.15)式，得到

$$\langle B \mid A\rangle = \sum_{\text{满足(2.1)的一切}r(t)} (2\pi i\hbar\varepsilon/m)^{-1/2} e^{(i/\hbar)S[r(t)]}. \tag{2.2}$$

由于路径 $r(t)$ 连续分布，所以这一"求和"实际上是对路径的"积分"，称为路径积分. 通常将它写成

$$\langle B \mid A\rangle = (2\pi i\hbar\varepsilon/m)^{-1/2}\int_{t_A}^{t_B} e^{(i/\hbar)S[r(t)]}\,\mathscr{D}r(t). \tag{2.3}$$

为了计算(2.3)式中的路径积分，将时间间隔 $T = t_B - t_A$ 均分为 n 份，$t_0 = t_A, t_1, t_2, \cdots, t_n = t_B$，其中每一份的宽度 $t_{i+1} - t_i = \varepsilon = T/n$. 在(2.3)式左边插进 $n-1$ 个完备性条件 $\int |r_i\rangle\langle r_i|\,dr_i, i = 1,2,\cdots,n-1$ ❶，

$$\langle B \mid A\rangle = \int\langle B \mid r_{n-1}\rangle\langle r_{n-1} \mid r_{n-2}\rangle\langle r_{n-2} \mid \cdots \mid r_1\rangle\langle r_1 \mid A\rangle\,dr_1\,dr_2\cdots dr_{n-1}. \tag{2.4}$$

容易看到，(2.4)式右边的被积函数中的尖括号 $\langle r_{i+1} \mid r_i\rangle$ 是由 r_i 到 r_{i+1} 的概率幅. 整个等式右边是由 A 经由 $r_1, r_2, \cdots, r_{n-1}$ 到 B 的曲折路径的概率幅，见图 2. 为了便于作图，在图 2 中用一根纵轴代表三维变量 r.

当 $n\to\infty$，$\varepsilon\to 0$ 时，可以将从 r_i 到 r_{i+1} 的路径用直线路径代. 这一直线路径就是经典自由粒子的运动路径 $r(t)$，其作用量已经在第一篇第一章习题 12 中算出为

$$S[r(t)] = \frac{m(r_{i+1} - r_i)^2}{2(t_{i+1} - t_i)}. \tag{2.5}$$

利用 §5.6.1(1.1)式就可以将(2.4)式写为

$$\langle B \mid A\rangle = \left(\frac{2\pi i\hbar\varepsilon}{m}\right)^{-1/2}\int e^{\frac{i}{\hbar}\frac{m(r_B - r_{n-1})^2}{2(t_B - t_{n-1})}}dr_{n-1}\cdot\cdots\cdot\left(\frac{2\pi i\hbar\varepsilon}{m}\right)^{-1/2}\int e^{\frac{i}{\hbar}\frac{m(r_3 - r_2)^2}{2(t_3 - t_2)}}dr_2$$
$$\left(\frac{2\pi i\hbar\varepsilon}{m}\right)^{-1}\int e^{\frac{i}{\hbar}\frac{m(r_2 - r_1)^2}{2(t_2 - t_1)}}e^{\frac{i}{\hbar}\frac{m(r_1 - r_A)^2}{2(t_1 - t_A)}}dr_1. \tag{2.6}$$

注意 $(r_{i+1} - r_i)^2 = (x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2 + (z_{i+1} - z_i)^2$，$dr_i = dx_i dy_i dz_i$，就看到(2.6)式右边可以写成三个积分之积：

$$\langle B_1 \mid A\rangle = I_x \cdot I_y \cdot I_z, \tag{2.7}$$

$$I_x = \left(\frac{2\pi i\hbar\varepsilon}{m}\right)^{-1/2}\int e^{\frac{i}{\hbar}\frac{m(x_B - x_{n-1})^2}{2(t_B - t_{n-1})}}dx_{n-1}\cdot\cdots\cdot\left(\frac{2\pi i\hbar\varepsilon}{m}\right)^{-1/2}\int e^{\frac{i}{\hbar}\frac{m(x_3 - x_2)^2}{2(t_3 - t_2)}}dx_2\cdot$$
$$\cdot\left(\frac{2\pi i\hbar\varepsilon}{m}\right)^{-1}\int e^{\frac{i}{\hbar}\frac{m(x_2 - x_1)^2}{2(t_2 - t_1)}}e^{\frac{i}{\hbar}\frac{m(x_1 - x_A)^2}{2(t_1 - t_A)}}dx_1. \tag{2.8}$$

---

❶ 上一节已经证明了路径积分量子化与正则量子化的一致性，所以这里用了正则量子化中的完备性条件. 实际上，(2.4)式也可以完全用路径积分量子化得到，即将图 2 中的折线看成一系列相继发生的事件，应用相继发生的事件概率幅相乘得到(2.4)式.

对于 I_y, I_z 也有类似的式子.

首先计算(2.8)式最右边的一个积分

$$\left(\frac{2\pi i\hbar\varepsilon}{m}\right)^{-1}\int e^{\frac{i}{\hbar}\frac{m(x_2-x_1)^2}{2(t_2-t_1)}}e^{\frac{i}{\hbar}\frac{m(x_1-x_A)^2}{2(t_1-t_A)}}dx_1 = \left(\frac{2\pi i\hbar\varepsilon}{m}\right)^{-1}\int e^{\frac{im}{2\hbar\varepsilon}[(x_2-x_1)^2+(x_1-x_A)^2]}\cdot dx_1.$$

写

$$(x_2-x_1)^2+(x_1-x_A)^2 = (1/2)(x_2-x_A)^2+2\left[x_1-(1/2)(x_2+x_A)\right]^2$$

并利用 §5.6.1(1.14)式,可以算出(2.8)式最右边的一个积分

$$\left(\frac{2\pi i\hbar\varepsilon}{m}\right)^{-1/2}\int\left(\frac{2\pi i\hbar\varepsilon}{m}\right)^{-1/2}e^{\frac{im}{2\hbar\varepsilon}[(x_2-x_1)^2+(x_1-x_A)^2]}\cdot dx_1 = \left(\frac{2\pi i\hbar 2\varepsilon}{m}\right)^{-1/2}e^{\frac{im}{2\hbar 2\varepsilon}(x_2-x_A)^2}. \tag{2.9}$$

代回(2.8)式,接下来算下一个积分,用同样的办法得到

$$\left(\frac{2\pi i\hbar\varepsilon}{m}\right)^{-1/2}\int\left(\frac{2\pi i\hbar 2\varepsilon}{m}\right)^{-1/2}e^{\frac{im}{2\hbar 2\varepsilon}[(x_3-x_2)^2+(x_2-x_A)^2]}dx_2 = \left(\frac{2\pi i\hbar 3\varepsilon}{m}\right)^{-1/2}e^{\frac{im}{2\hbar 3\varepsilon}(x_3-x_A)^2}. \tag{2.10}$$

这样一步步积分下去,到最后第 $n-1$ 步积分是

$$\left(\frac{2\pi i\hbar\varepsilon}{m}\right)^{-1/2}\int\left(\frac{2\pi i\hbar(n-1)\varepsilon}{m}\right)^{-1/2}e^{\frac{im}{2\hbar(n-1)\varepsilon}[(x_n-x_{n-1})^2+(x_{n-1}-x_A)^2]}dx_{n-1}$$

$$= \left(\frac{2\pi i\hbar n\varepsilon}{m}\right)^{-1/2}e^{\frac{im}{2\hbar n\varepsilon}(x_n-x_A)^2}. \tag{2.11}$$

注意到 $x_n = x_B, n\varepsilon = t_B - t_A$,就得到

$$I_x = \left(\frac{2\pi i\hbar(t_B-t_A)}{m}\right)^{-1/2}e^{\frac{im}{2\hbar(t_B-t_A)}(x_B-x_A)^2}. \tag{2.12}$$

代回(2.7)式最后得到由 A 到 B 的传播函数

$$\langle B\mid A\rangle = \left(\frac{2\pi i\hbar(t_B-t_A)}{m}\right)^{-3/2}e^{\frac{im}{2\hbar(t_B-t_A)}(r_B-r_A)^2}. \tag{2.13}$$

现在来应用所得结果讨论图 1 中的双缝衍射.为了写书简单,令 $t_A = 0, t_{B_1} = t_1, t_{B_2} = t_2, r_{B_1} = r_1, r_{B_2} = r_2, t = 0$ 时由 A 出发,在时刻 t 到 C 的概率幅是由 A 通过上缝 B_1 到 C 和由 A 经过下缝 B_2 到 C 的概率幅之和:

$$\langle C,t\mid A,0\rangle = \int_0^t\langle C,t\mid B_1,t_1\rangle\langle B_1,t_1\mid A,0\rangle dt_1 +$$

$$\int_0^t\langle C,t\mid B_2,t_2\rangle\langle B_2,t_2\mid A,0\rangle dt_2. \tag{2.14}$$

我们只关心粒子由 A 出发的时间和到达 C 的时间,而不关心它通过缝的时间.因此要对 t_1 和 t_2 积分.

(2.14)式右边的各个传播函数都是类似于(2.7)(2.8)式那样的路径积分,其结果都类似于(2.13)式.因而有

$$\langle C,t\mid A,0\rangle = \left(\frac{m}{2\pi i\hbar}\right)^3\int_0^t t_1^{-3/2}e^{\frac{im a^2}{2\hbar t_1}}\cdot(t-t_1)^{-3/2}e^{\frac{imb^2}{2\hbar(t-t_1)}}dt_1 +$$

$$\left(\frac{m}{2\pi i\hbar}\right)^3\int_0^t t_2^{-3/2}e^{\frac{imc^2}{2\hbar t_2}}\cdot(t-t_2)^{-3/2}e^{\frac{imd^2}{2\hbar(t-t_2)}}dt_2. \tag{2.15}$$

利用附录(V.29)式可以算出这几个积分,结果得到

$$\langle C,t\mid A,0\rangle = \sqrt{\frac{2\pi^2}{im}}\left(\frac{m}{2\pi i\hbar}\right)^{3} t^{-3/2}\cdot$$
$$\left[\left(\frac{a+b}{ab}\right)^{2} e^{\cdots}+\left(\frac{c+d}{cd}\right)^{2} e^{\cdots}\right]. \tag{2.16}$$

假定 A 点在 $\overline{B_1B_2}$ 的中垂线上，因而 $a = c$．令 C_0 为屏幕上位于 $B_1\overline{B_2}$ 的中垂线上的点，当 $C = C_0$ 时 $c = d$．我们只讨论在 C_0 点附近一个小区间中的衍射花样，因而近似地令 $(a+b)/(ab) \approx (c+d)/(cd)$．

写 $a + b = l_1$，$c + d = l_2$，(2.16)式成为

$$|\langle C,t\mid A,0\rangle|^2 = 常数 \times \left[1+\cos\left\{\frac{m}{2\hbar t}(l_1^{\,2}-l_2^{\,2})\right\}\right], \tag{2.17}$$

再令 $l_1/t \approx l_2/t = v = p/m$，可以写 $(l_1^{\,2}-l_2^{\,2})/t = 2v(l_1-l_2) = 2p(l_1-l_2)/m$．代入(2.17)式，并写 $p = 2\pi\hbar/\lambda$ 最后得到

$$|\langle C,t\mid A,0\rangle|^2 = 常数 \times \left[1+\cos\left(2\pi\frac{l_1-l_2}{\lambda}\right)\right]. \tag{2.18}$$

当 $(l_1-l_2)/\lambda = n$ 为整数时有峰　　(加强的干涉)，
当 $(l_1-l_2)/\lambda = n/2$ 为半整数时有谷　(减弱的干涉)． (2.19)

这样就得到了双缝衍射图样．

* §5.6.3　磁通量子

在本节里，我们来讨论量子力学波函数的相位导致的一种特殊效应——在电子运动区域以外的磁通对电子运动的影响．

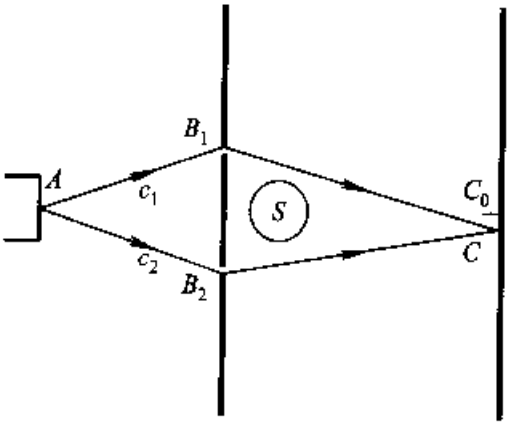


图 1　阿哈洛诺夫-玻姆效应示意图

(一)阿哈洛诺夫－玻姆效应

在双缝实验中加进一个圆柱线圈，如图 1，使圆柱中心产生磁场，并将线圈屏蔽，使电子不能穿透．用 A 表示矢势，则带电荷 $-e$ 的电子有势能 $(e/c)\boldsymbol{A}\cdot\dot{\boldsymbol{r}}$，见 §2.3.3(3.3)式(注意那里用的是国际单位制，而这里是高斯单位制)．拉格朗日是

$$L = (m/2)\dot{\boldsymbol{r}}^2 - (e/c)\boldsymbol{A}\cdot\dot{\boldsymbol{r}}. \tag{3.1}$$

作用量是

$$S = S_0 - \frac{e}{c}\int_A^r \boldsymbol{A}\cdot d\boldsymbol{r}, \tag{3.2}$$

其中 S_0 是无磁场 $(A = 0)$ 时自由粒子的作用量．

将(3.2)式代入 §5.6.2(2.14)式，计算 $\langle C,t\mid A,0\rangle$，结果是使 §5.6.2(2.16)式修改为

$$\langle C,t\mid A,0\rangle = 常数 \times [e^{\frac{im}{2\hbar t}l_1^{\,2}-\frac{ie}{\hbar c}\int c_1 A\cdot dr} + e^{\frac{im}{2\hbar t}l_2^{\,2}-\frac{ie}{\hbar c}\int c_2 A\cdot dr}]. \tag{3.3}$$

其中，c_1 和 c_2 分别是由 A 通过缝 B_1 和 B_2 到达 C 的路径，如图 2．在得到(3.3)式时已利用了条件 $(a+b)/ab \approx (c+d)/cd$．于是代替 §5.6.2(2.18)式有

$$|(C,t|A,0\rangle|^2 = 常数 \times \left\{1+\cos\left[2\pi\frac{l_1-l_2}{\lambda}-\frac{e}{\hbar c}\Phi\right]\right\}. \tag{3.4}$$

其中

$$\Phi = \left(\int_{c_1}-\int_{c_2}\right)\boldsymbol{A}\cdot d\boldsymbol{r} = \oint\boldsymbol{A}\cdot d\boldsymbol{r} = \int\nabla\times\boldsymbol{A}d\boldsymbol{S} = \int\boldsymbol{B}\cdot d\boldsymbol{S} \tag{3.5}$$

是穿过由 c_1 和 $-c_2$ 组成的闭合回路的磁通量，亦即穿过线圈 S 的磁通量.

由(3.4)式可见，代替§5.6.2(2.19)式有

$$\left.\begin{aligned}&当\frac{l_1-l_2}{\lambda}-\frac{e\Phi}{\hbar c}=n\text{ 为整数时有峰}\quad\text{（加强的干涉），}\\&当\frac{l_1-l_2}{\lambda}-\frac{e\Phi}{\hbar c}=\frac{n}{2}\text{ 为半整数时有谷}\quad\text{（减弱的干涉）.}\end{aligned}\right\} \tag{3.6}$$

阿哈洛诺夫-玻姆于1959年预言的这一现象第二年就被实验证实，称为阿哈洛诺夫-玻姆效应. 它的特点是，不在电子运动路径之内的磁场，也会影响电子运动的观测结果. 这是一个纯量子现象，是经典情况下不会发生的.

注意，当 $e\Phi/\hbar c = n$ 或 $\Phi = n(\hbar c/e)$ 时，(3.6)式和§5.6.2(2.19)式有相同的结果. 其中，

$$\Phi_0 = \hbar c/e = 4\times10^{-7}\text{G}\cdot\text{cm}^2\text{（高斯·厘米}^2\text{）} \tag{3.7}$$

称为磁通量子. 当线圈 S 中的磁通等于整数个磁通量子时，阿哈洛诺夫－玻姆效应消失.

（二）线圈内的磁通对线圈外电子能级的影响

在半径为 a 的圆柱线圈外套一个内径为 b 的同心圆柱，如图2. 电子被束缚在两个圆柱之间的区域中运动. 为了简单，不考虑平行于圆柱轴（z）和沿径向（r）的运动，而只考虑绕圆柱的 ϕ 角方向的运动.

按照“最小相互作用”§2.3.3(3.7)式得到哈密顿

$$\hat{H} = \frac{1}{2m}\left(\hat{p}_\phi-\frac{e}{c}A\right)^2. \tag{3.8}$$

其中，$\hat{p}_\phi = -(i\hbar/a)(\partial/\partial\phi)$ 是沿 ϕ 方向的动量.

能量的本征值方程是

$$\frac{1}{2m}\left(\frac{\hbar}{ia}\frac{\partial}{\partial\phi}-\frac{e}{c}A\right)^2\psi(\phi) = E\psi(\phi). \tag{3.9}$$

这一方程满足标准条件的解是

$$\psi_n(\phi) = e^{in\phi},\quad n=0,\pm1,\cdots. \tag{3.10}$$

对应的能量本征值是

$$E_n = \frac{1}{2m}\left(\frac{n\hbar}{a}-\frac{e}{c}A\right)^2 = \frac{\hbar^2}{2ma^2}\left(n-\frac{\Phi}{\Phi_0}\right)^2. \tag{3.11}$$

其中，

$$\Phi = 2\pi aA = \oint\boldsymbol{A}\cdot d\boldsymbol{l} = \int\boldsymbol{B}\cdot d\boldsymbol{S}$$

是穿过半径为 a 的线圈的磁通，Φ_0 是由(3.7)式定义的磁通量子.

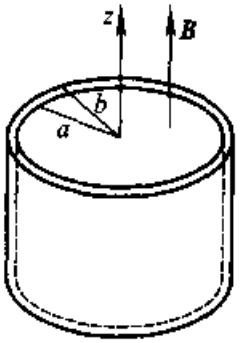


图1　磁通柱对柱外电子的影响

我们看到，在电子运动区间（半径为 a 和 b 的两个圆柱之间的区域）之外的磁场影响了电子的能级. 这一效应在线圈中的磁通 Φ 等于磁通量子 Φ_0 的整数倍时消失.

习　题

1. 计算 §5.6.2(2.9)式、(2.10)式和(2.11)式的积分.

2. 用计算机画出 §5.6.2(2.18)式给出的双缝衍射图样.

第七章　高速微观粒子的运动

§5.7.1　建立高速微观粒子运动方程遇到的困难

在§5.3.1中指出，薛定谔方程可以在经典能量表达式

$$E = p^2/(2m) + U(\boldsymbol{r}) \tag{1.1}$$

中作代换

$$E \to \mathrm{i}\hbar\partial/\partial t; \qquad \boldsymbol{p} \to -\mathrm{i}\hbar\nabla \tag{1.2}$$

而得到，见该节(1.6)式.因而很显然，薛定谔方程只适用于低速运动的微观粒子.在本章里，我们来讨论高速微观粒子的运动.

在试图建立高速微观粒子的运动方程之前，让我们先讨论一下，建立一个新的运动方程，或者更一般地说，建立一个新理论，需要遵守那些条件.当然，物理学是实验的科学.一个新理论必然是建立在实验的基础上，而且要受到实验的检验.然而，还应注意到另一个方面，即一个正确的理论必须内部自洽.

以量子力学的薛定谔方程为例，在写出方程

$$\mathrm{i}\hbar\partial\Psi/\partial t = -[\hbar^2/(2m)]\nabla^2\Psi + U\Psi \tag{1.3}$$

以后，就要考查它是否自洽.由于微观粒子的二象性，应该能够用波函数 $\Psi(\boldsymbol{r},t)$ 构成粒子的位置概率密度 $\rho(\boldsymbol{r},t)$.作为概率密度，它应该是正定的

$$\rho(\boldsymbol{r},t) \geqslant 0, \tag{1.4}$$

同时，还应满足概率守恒

$$\partial\rho/\partial t = -\nabla\cdot\boldsymbol{j}. \tag{1.5}$$

在§5.3.1中证明了，薛定谔方程(1.1)的确能满足这一要求，即能够找到正定的概率密度 $\rho(\boldsymbol{r},t)$ 及和它相对应的[满足概率守恒(1.3)式的]概率流密度 $\boldsymbol{j}(\boldsymbol{r},t)$.具体做法是：用 $\Psi^*(\boldsymbol{r},t)$ 乘方程(1.1)式，并用 $\Psi(\boldsymbol{r},t)$ 乘复共轭方程，两者相减得到

$$\mathrm{i}\hbar(\Psi^*\partial\Psi/\partial t + \Psi\partial\Psi^*/\partial t) = -[\hbar^2/(2m)][\Psi^*\nabla^2\Psi - \Psi\nabla^2\Psi^*].$$

或

$$\partial|\Psi|^2/\partial t = -\nabla\cdot\{[\hbar/(2m\mathrm{i})](\Psi^*\nabla\Psi - \Psi\nabla\Psi^*)\}. \tag{1.6}$$

只要令

$$\rho = |\Psi|^2, \quad \boldsymbol{j} = [\hbar/(2m\mathrm{i})](\Psi^*\nabla\Psi - \Psi\nabla\Psi^*), \tag{1.7}$$

(1.6)式就有概率守恒方程(1.5)式的形式，而且概率密度 $\rho = |\Psi|^2$ 满足正定条件(1.2)式，见§5.3.1(1.14)式.

现在来按照这一思路尝试建立高速微观粒子的运动方程.很自然的办法是在相对论的能量－动量关系式§2.3.1(1.21)式

$$E^2 = c^2p^2 + m^2c^4 \tag{1.8}$$

中,按(1.2)式作代换,得到

$$-\hbar^2(\partial^2/\partial t^2)\Psi = [c^2(-i\hbar\nabla)^2 + m^2c^4]\Psi.$$

或写为

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\Psi = \left(\nabla^2 - \frac{m^2c^2}{\hbar^2}\right)\Psi. \tag{1.9}$$

这称为克莱因－高登方程.

为了给克莱因－高登方程(1.9)式找到守恒的概率密度和概率流密度,用 Ψ^* 乘方程,并用 $\Psi(r,t)$ 乘复共轭方程,两者相减得到

$$\Psi^*(\partial^2/\partial t^2)\Psi - \Psi(\partial^2/\partial t^2)\Psi^* = c^2(\Psi^*\nabla^2\Psi - \Psi\nabla^2\Psi^*).$$

或

$$\frac{\partial}{\partial t}(\Psi^*(\partial/\partial t)\Psi - \Psi(\partial/\partial t)\Psi^*) = \nabla[c^2(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*)]. \tag{1.10}$$

为使这一等式具有概率守恒(1.5)式的形式,只要令

$$\rho = A[\Psi^*(\partial/\partial t)\Psi - \Psi(\partial/\partial t)\Psi^*], \tag{1.11}$$

$$j = -A[c^2(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*)], \tag{1.12}$$

A 为任意常数.然而,这样定义的"概率密度" ρ 不满足正定条件(1.2)式.

由于克莱因－高登方程找不到满足概率守恒的正定概率密度,因此它不适合于作为高速微观粒子的运动方程.

§5.7.2 自由粒子的狄拉克方程

(一)方程的建立

上一节中看到,在相对论的能量－动量关系式[上节(1.6)式]中作上节(1.2)式的代换得到的方程,没有正定的概率密度.狄拉克找到了克服这一困难的办法.

容易看到,其所以薛定谔方程[上节(1.3)式]有正定的概率密度,而克莱因－高登方程[上节(1.9)式]没有,是因为前者对时间的偏导是一阶而后者对时间的偏导是二阶.为了建立有正定概率密度的方程,狄拉克将上节(1.8)式开方

$$E = \sqrt{c^2p^2 + m^2c^4} \tag{2.1}$$

再按上节(1.2)式作代换.然而,开方 $\sqrt{\quad}$ 不是线性算符,不满足叠加原理的要求.狄拉克的办法是将它写成线性展开式

$$E/c = \boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta mc, \tag{2.2}$$

其中 $\alpha_i(i=1,2,3)$ 和 β 是四个待定“系数”.

为了确定 $\alpha_i(i=1,2,3)$ 和 β，将(2.2)式乘平方

$$(E/c)^2 = \sum_{i=1}^{3}\alpha_i^2 p_i^2 + \beta^2(mc)^2 + \sum_{i<j}(\alpha_i\alpha_j + \alpha_j\alpha_i) + \sum_{i=1}^{3}(\alpha_i\beta + \beta\alpha_i). \tag{2.3}$$

为了能回到相对论的能量－动量关系式[上节(1.8)式]，应有

$$\alpha_i^2 = 1(i=1,2,3), \qquad \beta^2 = 1, \tag{2.4}$$

$$\left.\begin{aligned}&\alpha_i\alpha_j + \alpha_j\alpha_i = 0 \quad (i\neq j),\\ &\alpha_i\beta + \beta\alpha_i = 0 \quad (i=1,2,3).\end{aligned}\right\} \tag{2.5}$$

显然，任何实数或复数都不能满足(2.5)式. 因此，狄拉克让 $\alpha_i(i=1,2,3)$ 和 β 为四个满足(2.5)式的矩阵. 例如，可以令它们是

$$\alpha_i = \begin{pmatrix}0 & \sigma_i\\ \sigma_i & 0\end{pmatrix} \quad (i=1,2,3), \qquad \beta = \begin{pmatrix}\mathbf{1} & \mathbf{0}\\ \mathbf{0} & -\mathbf{1}\end{pmatrix}, \tag{2.6}$$

其中，$\sigma_i(i=1,2,3)$ 是泡利矩阵，见§5.4.6(6.30)式，而 $\mathbf{1}$ 是 2×2 的单位矩阵. 将(2.6)式代入(2.2)式，并按上节(1.2)式作代换，得到

$$\mathrm{i}\hbar\partial\Psi/\partial t = [-\mathrm{i}\hbar c\boldsymbol{\alpha}\cdot\nabla + \beta mc^2]\Psi, \tag{2.7}$$

称为狄拉克方程. 它也可以写为

$$\mathrm{i}\hbar\partial\Psi/\partial t = [c\boldsymbol{\alpha}\cdot\hat{\boldsymbol{p}} + \beta mc^2]\Psi, \tag{2.7'}$$

由于 $\boldsymbol{\alpha}$ 和 β 是 4×4 的矩阵，狄拉克方程中的波函数 Ψ 是 4 行 1 列矩阵，而共轭波函数 $\Psi^\dagger$ 是 1 行 4 列矩阵

$$\Psi(\boldsymbol{r},t) = \begin{pmatrix}\psi_1(\boldsymbol{r},t)\\ \psi_2(\boldsymbol{r},t)\\ \psi_3(\boldsymbol{r},t)\\ \psi_4(\boldsymbol{r},t)\end{pmatrix},$$

$$\Psi^\dagger(\boldsymbol{r},t) = (\psi_1^*(\boldsymbol{r},t)\quad \psi_2^*(\boldsymbol{r},t)\quad \psi_3^*(\boldsymbol{r},t)\quad \psi_4^*(\boldsymbol{r},t)). \tag{2.8}$$

它归一化为

$$\int\Psi^\dagger\Psi\mathrm{d}^3\boldsymbol{r} = \sum_{i=1}^{4}\int\psi_i^*\psi_i\mathrm{d}^3\boldsymbol{r} = 1. \tag{2.9}$$

(二)概率密度与概率流密度

我们来讨论狄拉克方程的概率密度与概率流密度，看看是否能得到正定的守恒概率密度.

为此用 $\Psi^\dagger$ 乘狄拉克方程(2.7)式，并用 $\Psi(\boldsymbol{r},t)$ 乘厄米共轭方程，两者相减得到

$$\partial \Psi^{\dagger}\Psi/\partial t = -\nabla \cdot c\Psi^{\dagger}\boldsymbol{\alpha}\Psi.$$

和§5.7.1(1.5)式比较可见,狄拉克方程的概率密度和概率流密度分别是

$$\rho = \Psi^{\dagger}\Psi = \sum_{i=1}^{4} |\psi_i|^2, \tag{2.10}$$

$$\boldsymbol{j} = c\Psi^{\dagger}\boldsymbol{\alpha}\Psi. \tag{2.10'}$$

符合物理要求§5.7.1(1.4)、(1.5)式,概率密度是正定的.

(三)狄拉克海

我们看到,狄拉克通过将§5.7.1(1.8)式开方,克服了克莱因-高登方程的概率不正定问题.然而这样开方得到的能量有正负两个值,从而又出现了"负能问题".

注意,在非相对论情况下,计算能量的起点可以任意选取,能量的值可正可负;而在相对论情况下,计算能量的起点不能改变,只有正的能量才有物理意义.

实际上,§5.7.1(1.8)式是相对论力学的普遍公式,见§2.3.1(1.21)式.由它得到的能量本来就有正负两个解.对于宏观高速粒子(非量子情况),这不构成问题.因为在正能和负能之间有一个宽为$2mc^2$的能隙,而在非量子情况下,粒子的能量只能连续变化,因而当粒子有正能量时,其能量最多只能降低到"静能"mc^2,如图1.

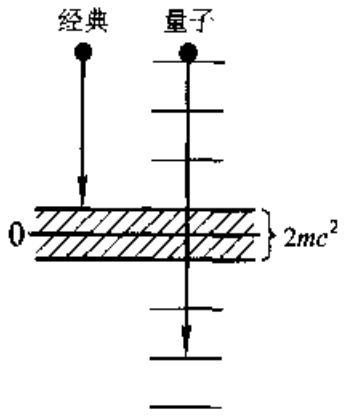


图1 相对论的负能量在经典和量子情况下产生的不同结果

但是,高速微观粒子的能量可以不连续地改变——量子跃迁.因而,即使电子原来处于正能态,也会自发地向负能态跃迁,如图1.这就使得"负能"成为一个严重问题.

为了解决这一问题,狄拉克提出一个假设.他假定,人们观察到的"真空"并不是没有电子,而是有无穷多个电子,充满了全部负能态.这种全部负能态充满了电子的状态称为狄拉克海.由于这个"负能电子海"总是存在于我们周围,如果它的状态不发生变化,我们就感觉不到它的存在,而认为它是"真空".在这种"真空态"中加进一个正能电子,则由于电子有半整数自旋,服从泡利不相容原理,这个正能态的电子不会向负能态跃迁,它的运动规律就和只有一个电子时的运动规律一样.

狄拉克海的假设不但克服了负难困难,而且有新的预言.当光子的能量$\hbar\omega$大于$2mc^2$时,真空(狄拉克海)中的负能电子有可能会被激发到正能态来,同时在拉克狄海中留下一个空穴,如图2.(注意,电子有两个自旋态,因而每个能态可以容纳两个电子.)当狄拉克海中充满电子时,我们认为是"真空",感觉到它的电荷为零.当拉克狄海中少了一个电子(有一个空穴)时,我们就会感觉到存在有一个带

正电荷的粒子.这种带有正电荷,而其他性质(质量、自旋等)和电子完全相同的粒子称为正电子.高频光将拉克狄海中的一个负能电子激发到正能态,形成一个空穴和一个(正能)电子,从人的感觉说来,就是光子转化成了一对电子和正电子,如图2.这一现象的确在实验中被观察到.

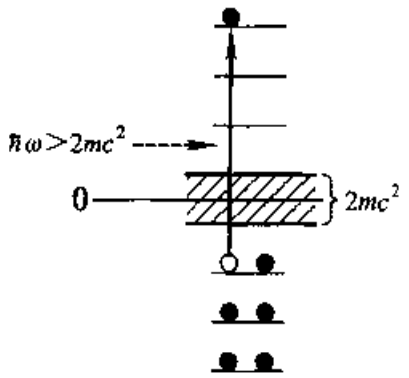


图2　高频光子产生电子—空穴(正电子)对

(四)狄拉克粒子的自旋

我们知道,无外场或在有心力场中的系统有空间转动不变性(对称性),由于这一对称性,这种系统的角动量守恒.其哈密顿算符和总角动量算符可对易.例如,在低速运动情况下,薛定谔方程的自由粒子哈密顿算符 $\hat{H}=-[\hbar^2/(2m)]\nabla^2$ 和轨道角动量 $\hat{\boldsymbol{L}}=-\mathrm{i}\hbar\boldsymbol{r}\times\nabla=\boldsymbol{r}\times\hat{\boldsymbol{p}}$ 可对易,见§5.3.3,因而由薛定谔方程描述的自由粒子的轨道角动量守恒.

然而,狄拉克方程(2.7)式的哈密顿算符

$$\hat{H}=-\mathrm{i}\hbar c\boldsymbol{\alpha}\cdot\nabla+\boldsymbol{\beta}mc^2=c\boldsymbol{\alpha}\cdot\hat{\boldsymbol{p}}+\boldsymbol{\beta}mc^2 \tag{2.11}$$

和 $\hat{\boldsymbol{L}}=-\mathrm{i}\hbar\boldsymbol{r}\times\nabla=\boldsymbol{r}\times\hat{\boldsymbol{p}}$ 不可对易:

$$\begin{aligned}[\hat{L},\hat{H}]&=[\hat{\boldsymbol{L}},c\boldsymbol{\alpha}\cdot\hat{\boldsymbol{p}}]=ce_i\alpha_j[\hat{L}_i,\hat{p}_j]\\&=\mathrm{i}\hbar ce_i\alpha_j\varepsilon_{ijk}\hat{p}_k=\mathrm{i}\hbar c\boldsymbol{\alpha}\times\hat{\boldsymbol{p}}\neq 0.\end{aligned} \tag{2.12}$$

在后一行中用了§5.2.3(3.22)式.这一结果表明,轨道角动量 $\boldsymbol{L}$ 不是狄拉克粒子的全部角动量,狄拉克粒子还有另一种角动量——自旋角动量 $\boldsymbol{S}$,它和轨道角动量 $\boldsymbol{L}$ 一道构成狄拉克粒子的总角动量:$\boldsymbol{J}=\boldsymbol{L}+\boldsymbol{S}$. $\hat{L}$ 和 $\boldsymbol{S}$ 分别和 $\hat{H}$ 都不对易,只有它们的和 $\hat{\boldsymbol{J}}$ 才和 $\hat{H}$ 对易

$$\hat{\boldsymbol{J}}=\hat{L}+S,\qquad[\hat{J},\hat{H}]=0. \tag{2.13}$$

由(2.12)、(2.13)式知,应该有 $[\boldsymbol{S},\hat{H}]=-\mathrm{i}\hbar c\boldsymbol{\alpha}\times\hat{\boldsymbol{p}}$. 我们就根据这一条件来求 $\boldsymbol{S}$.

为此,定义一个矩阵

$$\boldsymbol{\Sigma}=\begin{pmatrix}\boldsymbol{\sigma}&0\\0&\boldsymbol{\sigma}\end{pmatrix}, \tag{2.14}$$

它和 $\beta,\boldsymbol{\alpha}$ 矩阵之间的对易关系是

$$[\boldsymbol{\Sigma},\beta]=0,\qquad[\Sigma_i,\alpha_j]=2\mathrm{i}\,\varepsilon_{ijk}\alpha_k. \tag{2.15}$$

证明留作习题.由此易见

$$[\boldsymbol{\Sigma},\hat{H}]=[\boldsymbol{\Sigma},c\boldsymbol{\alpha}\cdot\hat{\boldsymbol{p}}]=ce_i[\Sigma_i,\alpha_j]\hat{p}_j=2\mathrm{i}c\hat{\boldsymbol{p}}\times\boldsymbol{\alpha}. \tag{2.16}$$

因此,只要令

$$\boldsymbol{S}=(\hbar/2)\boldsymbol{\Sigma}, \tag{2.17}$$

就有

$$[\boldsymbol{S},\hat{H}]=\mathrm{i}\hbar c\hat{\boldsymbol{p}}\times\boldsymbol{\alpha}. \tag{2.18}$$

结合(2.12)式就有(2.13)式.这就证明了,按(2.14)(2.17)式定义的 S 是狄拉克粒子的自旋算符.

可以证明(见习题4),S_z 的本征值是 $\pm \hbar/2$,而 S^2 的本征值是 $(3/4)\hbar^2$.因此,狄拉克方程描述的粒子的自旋为1/2.

我们看到,相对论情况和非相对论情况下对自旋的描述很不相同.在非相对论情况下,自旋和轨道运动是相互独立的.描述轨道运动的薛定谔方程既可以描述无自旋($S=0$)的粒子,也可以描述有任意自旋的粒子.总波函数是轨道波函数(薛定谔方程的解)和自旋波函数两个相互独立部分的乘积.而在相对论情况下则不然.自旋和轨道运动相互关联,因而一个运动方程只能描述一种自旋的粒子,例如,狄拉克方程描述的就只是自旋为1/2的粒子.

虽然不同自旋的粒子由不同的方程描述,然而对它们有一个共同要求,即:所有自旋不为零的粒子的多分量波函数中,每个分量都应满足克莱因-高登方程.这是由于它们都应满足相对论的能量-动量关系式.而不同自旋粒子波函数的各个分量之间的关系,则由它们各自的方程规定.

只有一个分量的克莱因-高登方程本应描述自旋为零的粒子,然而在§5.7.1中看到,这种描述方式的概率密度不正定.彻底解决这一问题的办法,是将克莱因-高登方程看成经典波动方程,然后将它量子化.即:克莱因-高登方程不能作为自旋为零的粒子的量子力学方程,但可以在量子场论中作为描述自旋为零的粒子的经典场方程.

(五)自由狄拉克方程的解

用分离变量法,将狄拉克方程的解(2.8)式写成

$$\Psi = \begin{pmatrix}\varphi\\ \chi\end{pmatrix} e^{-(i/\hbar)Et}, \qquad \varphi = \begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}, \qquad \chi = \begin{pmatrix}\psi_3\\ \psi_4\end{pmatrix}. \tag{2.19}$$

和(2.6)式一道代入狄拉克方程(2.7′),得到

$$E\begin{pmatrix}\varphi\\ \chi\end{pmatrix} = c\boldsymbol{\sigma}\cdot\boldsymbol{p}\begin{pmatrix}\chi\\ \varphi\end{pmatrix} + mc^2\begin{pmatrix}\varphi\\ -\chi\end{pmatrix}. \tag{2.20}$$

或写成两个联立方程

$$(E - mc^2)\varphi = c\boldsymbol{\sigma}\cdot\boldsymbol{p}\chi. \tag{2.21a}$$

$$(E + mc^2)\chi = c\boldsymbol{\sigma}\cdot\boldsymbol{p}\varphi. \tag{2.21b}$$

由此得

$$\varphi = \frac{c\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E - mc^2}\chi, \qquad \chi = \frac{c\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E + mc^2}\varphi. \tag{2.22}$$

狄拉克方程有正能解和负能解.我们来讨论在接近非相对论的情况下,这两种解之间的差别.对于非相对论的正能解(低速运动的电子),$E = mc^2 + \varepsilon, \varepsilon \ll mc^2$,

由(2.22)式知 $|\chi| \ll |\varphi|$，波函数的上分量 φ 是大分量，而下分量 χ 是小分量；对于非相对论的负能解(低速运动的正电子)，$E = -(mc^2 + \varepsilon)$，$\varepsilon \ll mc^2$，由(2.22)式知 $|\varphi| \ll |\chi|$，波函数的下分量 χ 是大分量，而上分量 φ 是小分量.

回过头来求相对论情况下的解. 将(2.22)式的第二式代入(2.21a)，将(2.22)的第一式代入(2.21b)，分别得到两个恒等式 $\varphi = \varphi$ 和 $\chi = \chi$，因此，采用(2.22)的第二式时，φ 可以任意取；采用(2.21)的第一式时，χ 可以任意取. 例如，采用(2.22)的第二式，可以取 φ 为 $\boldsymbol{\sigma}\cdot\hat{\boldsymbol{p}}$ 的本征矢. 在 $\boldsymbol{\sigma}^2$ 和 $\boldsymbol{\sigma}\cdot\hat{\boldsymbol{p}}$ 的共同表象中，$\boldsymbol{\sigma}\cdot\hat{\boldsymbol{p}}$ 有两个本征矢，分别代表右旋(R)和左旋(L)态

$$\boldsymbol{\sigma}\cdot\hat{\boldsymbol{p}}\varphi^{(\mathrm{R})} = +\varphi^{(\mathrm{R})}, \qquad \boldsymbol{\sigma}\cdot\hat{\boldsymbol{p}}\varphi^{(\mathrm{L})} = -\varphi^{(\mathrm{L})},$$

$$\varphi^{(\mathrm{R})} = \begin{pmatrix}1\\0\end{pmatrix}, \qquad \varphi^{(\mathrm{L})} = \begin{pmatrix}0\\1\end{pmatrix} \qquad (\boldsymbol{\sigma}^2, \boldsymbol{\sigma}\cdot\boldsymbol{p}\ \text{表象}). \tag{2.23}$$

再利用(2.22)的第二式，得到两个独立的正能解

$$u^{(\mathrm{R})} \propto \begin{pmatrix} 1 \\ 0 \\ \dfrac{cp}{E_p + mc^2} \\ 0 \end{pmatrix}, \qquad u^{(\mathrm{L})} \propto \begin{pmatrix} 0 \\ 1 \\ 0 \\ \dfrac{-cp}{E_p + mc^2} \end{pmatrix}. \tag{2.24a}$$

类似地，利用(2.22)的第一式得到两个独立的负能解

$$v^{(\mathrm{R})} \propto \pm \begin{pmatrix} \dfrac{-cp}{E_p + mc^2} \\ 0 \\ 1 \\ 0 \end{pmatrix}, \qquad v^{(\mathrm{L})} \propto \pm \begin{pmatrix} 0 \\ \dfrac{cp}{E_p + mc^2} \\ 0 \\ 1 \end{pmatrix}. \tag{2.24b}$$

* §5.7.3　电子的内禀磁矩

电磁场中电子的狄拉克方程可以在自由电子的狄拉克方程§5.7.2(2.7′)中作最小相互作用代换§2.3.3(3.7)式

$$\boldsymbol{p} \to \boldsymbol{p} - (e/c)\boldsymbol{A}, \qquad E = E - e\Phi, \tag{3.1}$$

而得到(注意，①电子的电荷为 $-e$；②本节所用的是高斯单位制)

$$\mathrm{i}\hbar\frac{\partial\Psi}{\partial t} = \left[c\boldsymbol{\alpha}\cdot\left(\boldsymbol{p} - \frac{e}{c}\boldsymbol{A}\right) + \beta mc^2 + e\Phi\right]\Psi. \tag{3.2}$$

这就是电磁场中的狄拉克方程.

类似于§5.7.2(2.19)式写

$$\Psi(\boldsymbol{r},t)=\begin{pmatrix}\varphi(\boldsymbol{r},t)\\ \chi(\boldsymbol{r},t)\end{pmatrix},\qquad \varphi(\boldsymbol{r},t)=\begin{pmatrix}\psi_1(\boldsymbol{r},t)\\ \psi_2(\boldsymbol{r},t)\end{pmatrix},\qquad \chi(\boldsymbol{r},t)=\begin{pmatrix}\psi_3(\boldsymbol{r},t)\\ \psi_4(\boldsymbol{r},t)\end{pmatrix}. \tag{3.3}$$

类似于§5.7.2(2.21)式得到

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\begin{pmatrix}\varphi\\ \chi\end{pmatrix}=c\boldsymbol{\sigma}\cdot\boldsymbol{\pi}\begin{pmatrix}\chi\\ \varphi\end{pmatrix}+e\Phi\begin{pmatrix}\varphi\\ \chi\end{pmatrix}+mc^2\begin{pmatrix}\varphi\\ -\chi\end{pmatrix}. \tag{3.4}$$

其中引进了符号

$$\boldsymbol{\pi}=\boldsymbol{p}-(e/c)\boldsymbol{A}. \tag{3.5}$$

定态方程是

$$E\begin{pmatrix}\varphi\\ \chi\end{pmatrix}=c\boldsymbol{\sigma}\cdot\boldsymbol{\pi}\begin{pmatrix}\chi\\ \varphi\end{pmatrix}+e\Phi\begin{pmatrix}\varphi\\ \chi\end{pmatrix}+mc^2\begin{pmatrix}\varphi\\ -\chi\end{pmatrix}. \tag{3.6}$$

在电子速度不很高的情况下，可以用非相对论近似，有 $E\approx mc^2$，相对于 mc^2 可以忽略 $e\Phi$，从而由(3.6)的第二式解出

$$\chi=[\boldsymbol{\sigma}\cdot\boldsymbol{\pi}/(2mc)]\varphi. \tag{3.7}$$

代入(3.4)的第一式，得到

$$\left(\mathrm{i}\hbar\frac{\partial}{\partial t}-mc^2\right)\varphi=\left(\frac{(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})}{2m}+e\Phi\right)\varphi. \tag{3.8}$$

注意到，在相对论的总能中减去静能 mc^2 才是非相对论能量，作函数代换

$$\psi=\mathrm{e}^{(-\mathrm{i}/\hbar)mc^2t}\varphi. \tag{3.9}$$

它使得 φ 的时间因子 $\exp[(-\mathrm{i}/\hbar)Et]$ 中的 E 是非相对论能量. 将(3.9)式代入(3.8)式并利用习题7的结果，得到

$$\mathrm{i}\hbar\frac{\partial\psi}{\partial t}=\left[\frac{(\boldsymbol{p}-(e/c)\boldsymbol{A})^2}{2m}-\frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot\boldsymbol{B}+e\Phi\right]\psi. \tag{3.10}$$

这就是自旋为1/2的微观粒子(电子)的非相对论运动方程，称为泡利方程.

注意，在泡利方程中除了由薛定谔方程作代换(3.1)式得到的项之外，还多出一项

$$-[e\hbar/(2mc)]\boldsymbol{\sigma}\cdot\boldsymbol{B}=-[e/(mc)]s\cdot\boldsymbol{B}. \tag{7.11}$$

其中，$s=(\hbar/2)\boldsymbol{\sigma}$ 是电子的自旋. 这表明，狄拉克方程描述的粒子(电子)有内禀磁矩

$$\boldsymbol{m}=-[e/(mc)]\boldsymbol{s}. \tag{3.12}$$

其磁矩与自旋之比为

$$|\boldsymbol{m}|/|\boldsymbol{s}|=e/(mc). \tag{3.13}$$

习 题

1. 利用§2.2.2(2.1)式定义的4维时空坐标 $x_\mu(\mu=1,2,3,4)$，将§5.7.1中的克莱因-高登方程(1.9)式写成4维协变形式.

2. 证明按§5.7.2(2.6)式定义的 $\alpha_i(i=1,2,3)$，β 的平方等于1[§5.7.2(2.4)式]，且满足反对易性§5.7.2(2.5)式.

3. 证明§5.7.2(2.15)式.

4. 证明§5.7.2(2,14)(2,17)式定义的 S_z 的本征值是 $\pm\hbar/2$，而 S^2 的本征值是 $(3/4)\hbar^2$.

5. 证明狄拉克方程的波函数 Ψ 的 4 个分量 ψ_1、ψ_2、ψ_3、ψ_4 分别满足克莱因－高登方程.

6. 证明 § 5.7.2(2,21)式的两个式子不是独立的. 将其中一个代入另一个得到恒等式.

7. (a)证明将 §5.7.2(2.22)的两个式子代入该节(2.21a)和(2.21b)得到两个恒等式. (b)证明 §5.7.2(2.24)式,并说明称 $\varphi^{(R)}$ 和 $\varphi^{(L)}$ 为右旋和左旋态的原因.

提示　先写出 $\boldsymbol{\sigma}\cdot\boldsymbol{p}$ 表象中的 $\boldsymbol{\sigma}\cdot\boldsymbol{p}$ 矩阵.

8. 利用波函数的归一化条件 §5.7.2(2.9)式,求(2.24a)(2.24b)式中的比例系数(即波函数的归一化系数).

9. 证明 $(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})(\boldsymbol{\sigma}\cdot\boldsymbol{\pi}) = \boldsymbol{\pi}^2 - (e\hbar/c)\boldsymbol{\sigma}\cdot\boldsymbol{B}$,$\boldsymbol{\pi}$ 的定义见 §5.7.3(3.5)式.



第六篇　微观粒子在外场中的运动

本篇讨论微观粒子在外场中的运动.

微观粒子在外场中的运动,大体上可以按系统的哈密顿是否依赖时间而分成两种情况:一是外场不依赖时间,问题归结为求解定态方程;另一是外场依赖时间,要解决的问题是在定态之间的跃迁.

本篇的第一章讨论各种定态问题,第二章讨论在与时间有关的外场影响下,从一个状态到另一个状态的跃迁.

粒子在外场中的运动又可按粒子运动的空间限度分为束缚态问题和散射问题,本篇第一章讨论束缚态问题,散射问题在第三章讨论.

第一章　外场中的定态问题

一般说来，如果外场与时间无关，量子力学的基本问题是定态问题，也就是通过解定态方程确定系统的定态能量和定态波函数的问题. 例如，原子的能谱、分子的能谱和固体的能谱都是定态问题. 但是，对于绝大多数实际物理系统而言，定态方程不能严格求解，因而需要建立相应的近似方法. 这一章首先讨论有严格解的两个基本物理模型；然后讨论定态方程最常用的两种近似解法——微扰法和变分法. 此外，还简单介绍正则变换在量子力学中的应用.

§6.1.1　坐标表象中定态方程的严格解

这一节在坐标表象中求解束缚态的定态方程. 采用坐标表象的好处是：①能够对能量量子化——能谱分立的原因有更直观的了解，直接看到正是由于微观粒子具有波动性，波函数的标准条件使得几何空间里的运动受到约束，从而导致能量量子化；②能直接得到几何空间中的位置概率分布，即所谓“概率云”，从而对微观粒子的运动有形象化的概念.

正如本章引言所指出，定态方程能严格求解的情况不多. 这一节讨论两个基本的可解模型——谐振子和氢原子.

(一)坐标表象中的一维谐振子

在§1.4.1中已经指出，任何有平衡位置的一维系统，将势能展开，保留到二级项，就是一个简谐振子. 现在来讨论简谐振子，而把高级项引起的修正留到下一节再讨论.

谐振子的哈密顿算符是 $\hat{H}=\hat{p}^2/(2m)+m\omega^2\hat{x}^2/2$. §5.4.5在抽象希尔伯特空间中求解了它的定态方程，得到了能谱以及坐标、动量在粒子数表象中的矩阵元. 现在在坐标表象中重新求解这一问题.

在坐标表象中，定态薛定谔方程是

$$-[\hbar^2/(2m)]\mathrm{d}^2\psi/\mathrm{d}x^2+(m\omega^2/2)x^2\psi=E\psi \tag{1.1}$$

作代换

$$\xi=ax,\qquad a=\sqrt{m\omega/\hbar},\qquad \lambda=2E/(\hbar\omega). \tag{1.2}$$

使方程(1.1)化为

$$\mathrm{d}^2\psi/\mathrm{d}\xi^2+(\lambda-\xi^2)\psi(\xi)=0. \tag{1.3}$$

这是一个变系数的二阶常微分方程，需要用级数解法.

1. 微分方程的级数解法

对于能量量子化起决定作用的是波函数在 $x \to \pm\infty$ ($\zeta \to \pm\infty$) 时的渐近性质,因此,首先讨论 $|\xi| \to \infty$ 时解的渐近行为.

当 $|\xi| \to \infty$ 时,λ 可以略去,方程成为

$$\frac{d^2\psi}{d\xi^2} - \xi^2\psi(\xi) = 0. \tag{1.4}$$

我们来求指数形式的渐近解.令 $\psi = \exp[f(\xi)]$,其中 $f(\xi)$ 是 ξ 的一个多项式.于是有

$$\psi = e^f, \qquad \psi' = f'e^f, \qquad \psi'' = (f'^2 + f'')e^f \approx f'^2 e^f.$$

这里,考虑到 $f(\xi)$ 是 ξ 的多项式,每求一次导数要降低一个方次,所以 f'' 所含 ξ 的方次低于 f'^2,当 $|\xi| \to \infty$ 时可以略去.代入(1.4)式,得到

$$f'^2 = \xi^2, \qquad f = \pm\xi^2/2.$$

这样,方程(1.3)的渐近解是:

$$\psi(\xi) \sim Ae^{\pm\xi^2/2}, \qquad \text{当 } |\xi| \to \infty. \tag{1.5}$$

我们需要的是 $|\xi| \to \infty$ 时保持有限的解,因而上式的指数中应取负号.

为了将波函数在 $x \to \pm\infty$ ($\zeta \to \pm\infty$) 时的渐近行为分离出来,作函数代换:

$$\psi(\xi) = e^{-\xi^2/2}H(\xi), \tag{1.6}$$

方程(1.3)化为

$$d^2H/d\xi^2 - 2\xi dH/d\xi + (\lambda - 1)H = 0. \tag{1.7}$$

这一方程称为厄米方程.

为了采用级数解法,首先要分析解的奇异性,而解的奇异性决定于微分方程系数的奇异性.这里,"方程的系数"指的是,最高阶导数项(对于二阶方程就是二阶导数项)的系数为 1 的情况下,方程中其余各项的系数.由(1.7)式的系数可以看出,方程在 $\xi = 0$ 处是解析的,因而方程的解 $H(\xi)$ 可以在这一点的邻域内展开成泰勒级数:

$$H(\xi) = \sum_{\nu=0}^{\infty} a_\nu \xi^\nu. \tag{1.8}$$

为决定展开式系数,将它代回(1.7)式,并将结果按 ξ 的幂次整理,得到:

$$2a_2 + 6a_3\xi + \cdots + (\nu+2)(\nu+1)a_{\nu+2}\xi^\nu + \cdots$$
$$= (1-\lambda)a_0 + \cdots + (2\nu - \lambda + 1)a_\nu\xi^\nu + \cdots.$$

令等式两边 ξ 的同次幂项相等,得到

$$a_{\nu+2} = \frac{(2\nu - \lambda + 1)}{(\nu+1)(\nu+2)}a_\nu. \tag{1.9}$$

这是展开式系数的递推公式.由 a_0 可以逐步求出所有 ν 为偶数的系数,由 a_1 可以逐步求出所有 ν 为奇数的系数.这样

$$\left.\begin{aligned} H(\xi) &= H_s(\xi) + H_a(\xi) \\ H_s(\xi) &= a_0 + a_2\xi^2 + \cdots, \\ H_a(\xi) &= a_1\xi + a_3\xi^3 + \cdots. \end{aligned}\right\} \tag{1.10}$$

H_s 和 H_a 分别是偶函数和奇函数，它们是宇称算符 $\hat{P}$ 的本征值为 $P = \pm 1$ 的本征函数. 由(1.1)式可见，谐振子的哈密顿算符 $\hat{H}$ 在空间反演变换 $(x \to -x)$ 之下不变，因而 $\hat{H}$ 和宇称算符 $\hat{P}$ 有共同本征函数系. 因此，定态应该是有确定宇称的态. 这可以在(1.10)式中令 $a_0 = 0$ 或 $a_1 = 0$ 而得到.

2. 利用标准条件定解　能量量子化

根据解的标准条件，要求

$$|\xi| \to \infty, \qquad \psi(\xi) \text{ 有限.} \tag{1.11}$$

由(1.6)式知，这归结为 $|\xi| \to \infty$ 时，$H(\xi)$ 趋于无穷不能比 $e^{\xi^2/2}$ 更快. 利用递推公式(1.9)式，可以得到 H_s 或 H_a 的相邻系数之比：$a_{\nu+2}/a_\nu$. 我们关心的是 $|\xi| \to \infty$ 时的渐近行为，此时，级数中 ξ 的幂次高的项（ν 大的项）占支配地位. 当 ν 大时，$a_{\nu+2}/a_\nu \sim 2/\nu$. 这和无穷级数

$$e^{\xi^2} = 1 + \frac{\xi^2}{1!} + \frac{\xi^4}{2!} + \cdots + \frac{\xi^\nu}{(\nu/2)!} + \cdots$$

的相邻系数比：$(\nu/2 - 1)!/(\nu/2)! = 2/\nu$ 相同. 这表明，

$$H(\xi) \sim e^{\xi^2}, \qquad \text{当 } |\xi| \to \infty, \tag{1.12}$$

这比 $e^{\xi^2/2}$ 趋于无穷更快，违反了标准条件(1.11)式.

由此可见，为使解满足标准条件，$H(\xi)$ 不能是无穷级数，而只能在某一项中断，成为一个多项式. 显然，ξ 的任何有限多项式都比 $e^{\xi^2/2}$ 趋于无穷要慢，因而能使按(1.6)式定义的 $\psi(\xi)$ 满足标准条件的要求.

为使 $H(\xi)$ 成为多项式，要求(1.8)式中的无穷级数从某一项 $\nu = n$ 起中断. 由递推公式(1.9)式知，为此只需

$$\lambda = 2n + 1, \qquad n = 0, 1, 2, \cdots. \tag{1.13}$$

如果 n 为偶数，则这一条件能使 H_s 成为多项式；如果 n 为奇数，则这一条件能使 H_a 成为多项式. 因此当(1.13)式中的 n 为偶数，应取偶宇称解 H_s；当 n 为奇数，应取奇宇称解 H_a.

3. 厄米多项式

方程(1.7)的多项式解称为厄米多项式，用符号 $H_n(\xi)$ 表示. 脚标 n 是多项式的次数. 将(1.13)式代入(1.7)式得到 $H_n(\xi)$ 所满足的方程

$$d^2H_n/d\xi^2 - 2\xi dH_n/d\xi + 2nH_n = 0. \tag{1.14}$$

容易检验

$$H_n(\xi) = (-1)^n e^{\xi^2} \frac{d^n}{d\xi^n}(e^{-\xi^2}). \tag{1.15}$$

是 ξ 的一个 n 次多项式，且满足方程(1.14).

证 用数学归纳法，令 $F_n = e^{\xi^2}(e^{-\xi^2})^{(n)}$，其中上标 (n) 表示求 n 阶导数. 假定 F_{n-1} 满足 $n-1$ 时的方程(1.14)：

$$F''_{n-1} - 2\xi F'_{n-1} + 2(n-1)F_{n-1} = 0, \tag{1.16}$$

要证明 F_n 满足(1.14). 为此取 $F_{n-1} = e^{\xi^2}(e^{-\xi^2})^{(n-1)}$ 的导数，得到 $F'_{n-1} = 2\xi F_{n-1} + F_n$，因而

$$\left.\begin{aligned} F_n &= F'_{n-1} - 2\xi F_{n-1} \\ F'_n &= F''_{n-1} - 2\xi F'_{n-1} - 2F_{n-1}, \\ F''_n &= F'''_{n-1} - 2\xi F''_{n-1} - 4F'_{n-1}. \end{aligned}\right\} \tag{1.17}$$

因而

$$\begin{aligned} F''_n - 2\xi F'_n + 2nF_n = {} & [F'''_{n-1} - 2\xi F''_{n-1} + (2n-4)F'_{n-1}] \\ & - 2\xi[F''_{n-1} - 2\xi F'_{n-1} + 2(n-1)F_{n-1}]. \end{aligned}$$

此式右方第二个方括号正是(1.16)式的右方，而第一个方括号是(1.16)式右方的导数，因而此式右方等于零，即 F_n 满足方程(1.14).

显然，$F_0 = 1$ 满足 $n = 0$ 的方程(1.14)；故 n 为任何正整数时，F_n 均满足方程(1.14). 由此知 $H_n = (-1)^n F_n$ 是方程(1.14)的解. **【证毕】**

利用(1.15)式得到头几个厄米多项式：

$$\begin{aligned} H_0(\xi) &= 1, \\ H_1(\xi) &= 2\xi, \\ H_2(\xi) &= 4\xi^2 - 2, \\ H_3(\xi) &= 8\xi^3 - 12\xi, \\ H_4(\xi) &= 10\xi^4 - 48\xi^2 + 12, \\ H_5(\xi) &= 32\xi^5 - 160\xi^3 + 120\xi. \end{aligned} \tag{1.18}$$

由(1.16)、(1.17)式知 $H_n(\xi)$ 满足递推关系：

$$H'_n(\xi) = 2nH_{n-1}(\xi), \tag{1.19}$$

$$H_{n+1}(\xi) - 2\xi H_n(\xi) + 2nH_{n-1}(\xi) = 0. \tag{1.20}$$

4. 能谱与定态波函数

将(1.13)式代入(1.2)式得到能谱：

$$E_n = (n + 1/2)\hbar\omega, \tag{1.21}$$

和 §5.4.5 中的结果一致. 在那里，是将坐标和动量由 c 数变为 q 数，满足一定的

对易关系，从而使能谱分立；而这里则明显看出，是描述粒子波动性的函数 $\psi(x)$ 受到势能场的约束，使能谱分立.

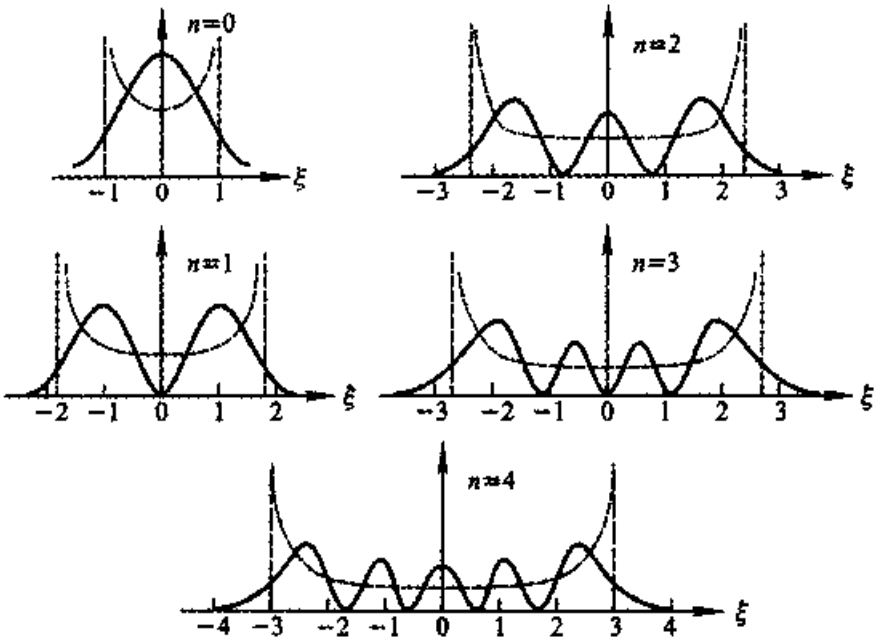


图 1　谐振子的概率密度

将厄米多项式代入(1.6)式，并利用(1.2)式，得到和定态能量(1.21)式对应的定态波函数：

$$\psi_n(x) = N_n e^{-\alpha^2 x^2/2} H_n(\alpha x). \tag{1.22}$$

其中

$$N_n = \sqrt{\alpha/(2^n n!\sqrt{\pi})}. \tag{1.23}$$

是归一化常数.

图 1 中画出了 $n = 0,1,2,3,4$ 的概率密度 $|\psi_n|^2$（图中的实线）. 图中的虚曲线是经典谐振子的平均位置密度. 按经典力学，谐振子的 $x = A\sin(\omega t + \varphi)$. 对时间求导，得 $v = -\omega\sqrt{A^2 - x^2}$，因而找到粒子的概率 $P(x) \sim v^{-1} \sim 1/\sqrt{A^2 - x^2}$，如图 1 中的虚曲线. 注意，$x = A$ 是经典转折点，对应于总能等子势能：$E = U$，由(1.3)式和(1.13)式可见，在这一点 $\xi = \sqrt{\lambda} = \sqrt{2n + 1}$，如图 1 中的虚直线. 由图 1 可见：①经典振子不能进入 $|x| > A$ 的区域，而量子振子则可以进入这一区域，进入以后概率指数衰减；②当 n 增大时，$\Delta E/E_n$ 减小，量子振子向经典振子过渡——量子振子的概率密度成为无限密集的波，其波峰的分布趋近于经典振子的概率分布（图 1 中的虚线）.

5. 两种表象的比较

前面已经提到，粒子数表象和坐标表象强调的方面不同，面得到的能谱相同.

同时,利用(1.22)式的 ψ_n 计算 x 和 p 的矩阵元:

$$x_{nm} = \int \psi_n^* x \psi_m \mathrm{d}x, \qquad p_{nm} = \int \psi_n^* (-i\hbar \mathrm{d}/\mathrm{d}x) x \psi_m \mathrm{d}x,$$

也可以得到和 §5.4.5(5.32)(5.33)式相同的结果.

反过来,利用 §5.4.5(5.32)式解矩阵 x_{nm} 的本征值方程:

$$\begin{pmatrix} x_{11} & x_{12} & \cdots \\ x_{21} & x_{22} & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} = x \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix}, \tag{1.24}$$

也可以求得 $a_n = \langle n \mid x \rangle = \langle x \mid n \rangle^* = \psi_n^*(x)$,从而得到定态波函数 $\psi_n(x)$,结果和(1.22)式一致●.

由此可见,两种表象完全等效.但是,为了解决不同的问题,采用不同的表象更为方便.

(二)有心力场中运动的一般分析

在第一篇第三章中讨论了经典粒子在有心力场中的运动,现在来考虑微观粒子在有心力场中的运动.

按定义,有心力场的势能只是 r 的函数,与 θ、φ 无关,如 §1.3.1(1.7)式:$U = U(r)$.这样的场具有中心对称性.由于场的这种对称性,所以在球坐标系讨论问题比较方便.在球坐标系,有心力场中粒子的哈密顿算符为 §5.2.3(3.31)式

$$\hat{H} = -\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\right) + \frac{\hat{L}^2}{2\mu r^2} + U(r), \tag{1.25}$$

式中 $\hat{L}^2$ 是角动量平方算符,见 §5.2.2(2.23)式,而 $\mu = m_1 m_2/(m_1 + m_2)$ 是二体问题的约化质量,如 §1.3.1(1.6)式.

在求解定态方程之前,我们先说明一个具有一般意义的问题,即:如果系统具有某种对称性,使得系统的哈密顿算符和其他若干相应的力学量算符对易,则我们欲求的定态解,实际上是这些力学量的共同本征函数系.也就是说,要求解的是和系统的对称性相应的一组相互对易的力学量完全组的共同本征函数系.利用力学量完全组不仅可以方便求解定态方程,而且能使我们更具体地了解到定态波函数描述了怎样的物理状态.

由于场具有中心对称性,$\hat{H}$,$\hat{L}^2$ 和 $\hat{L}_z$ 是一组相互对易的力学量完全组(在本节里不考虑粒子的自旋).我们所要求的定态解就是它们的共同本征函数系.在 §5.2.2 例3 中已经证明,$\hat{L}^2$ 和 $\hat{L}_z$ 的共同本征函数为球谐函数 $Y_{lm}(\theta,\varphi)$.又从 $\hat{H}$ 的表示式(1.25)可以看出,径向变量 r 和角向变量 θ,φ 是独立的,因而可令 $\hat{H}$,$\hat{L}^2$ 和 $\hat{L}_z$

● 参看:E. Merzbacher. Quantum Mechanics. John Wiley&Sons. 1961:361,362.

的共同本征函数为

$$\psi(r,\theta,\varphi) = R(r)Y_{lm}(\theta,\varphi), \tag{1.26}$$

问题就归结为求径向波函数 $R(r)$.

将(1.25),(1.26)式代入定态方程,得到:

$$\left[-\frac{\hbar^2}{2\mu}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)+\frac{\hat{L}^2}{2\mu r^2}+U(r)\right]R(r)Y_{lm}(\theta,\varphi) = ER(r)Y_{lm}(\theta,\varphi).$$

利用

$$\hat{L}^2Y_{lm}(\theta,\varphi) = l(l+1)\hbar^2Y_{lm}(\theta,\varphi).$$

得到径向波函数的方程

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right)-\frac{l(l+1)}{r^2}R+\frac{2\mu}{\hbar^2}[E-U(r)]R = 0. \tag{1.27}$$

由此可见,粒子在中心场中运动的能量完全由径向方程决定.

作函数替换

$$R(r) = u(r)/r, \tag{1.28}$$

代回方程(1.27),并注意

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right) = \frac{1}{r}\frac{\mathrm{d}^2u}{\mathrm{d}r^2}, \tag{1.29}$$

得到关于 $u(r)$ 的方程:

$$\frac{\mathrm{d}^2u}{\mathrm{d}r^2}+\left[\frac{2\mu}{\hbar^2}(E-U(r))-\frac{l(l+1)}{r^2}\right]u = 0. \tag{1.30}$$

为保证 $r=0$ 处 ψ 有限,要求 $r=0$ 时, $rR=0$,即

$$u(0) = 0. \tag{1.31}$$

这是方程(1.30)的解应满足的一个定解条件.

方程(1.30)类似于在势场

$$U_{等效}(r) = U(r)+l(l+1)\hbar^2/(2\mu r^2) \tag{1.32}$$

中作一维运动的定态方程. u 的归一化条件可写成

$$\int_0^\infty |u(r)|^2\mathrm{d}r = \int_0^\infty |R|^2r^2\mathrm{d}r = 1. \tag{1.33}$$

这也是径向波函数 $R(r)$ 的归一化条件. 显然,(1.32)式和 §1.3.2(2.7)式对应. 根据有效势(1.32)式可以知道, $E<0$ 为束缚态问题; $E>0$ 为无限运动,如 §1.3.2(2.15)式. 这一节只讨论 $E<0$ 的束缚态问题.

(三)电子在库仑场中的运动

1. 径向方程

现在讨论电子在原子核的库仑场中的运动. 在此情况下可以认为原子核不动,约化质量等于电子的质量 $\mu \approx m_e$. 设 Z 是核电荷数,则电子在原子核库仑场中的

势能为

$$U(r) = -Ze^2/r. \tag{1.34}$$

径向方程(1.30)成为

$$\frac{d^2u}{dr^2} + \left[\frac{2m_eE}{\hbar^2} + \frac{2m_eZe^2}{\hbar^2 r} - \frac{l(l+1)}{r^2}\right]u = 0. \tag{1.35}$$

为了便于求解,引入量纲为1变量

$$\rho = r/a_0, \qquad \alpha = \sqrt{-2E/E_1} \tag{1.36}$$

来代替 r 和 E. 这里 $a_0 = \hbar^2/(m_e e^2)$ 是玻尔轨道半径(见附录Ⅷ), E_1 是在这个轨道上电子的势能

$$E_1 = e^2/a_0 = m_e e^4/\hbar^2 = 27.21\text{eV}. \tag{1.37}$$

这样,方程(1.35)化为

$$d^2u(\rho)/d\rho^2 + [-\alpha^2 + 2Z/\rho - l(l+1)/\rho^2]u = 0. \tag{1.38}$$

2. 方程的级数解

和解厄米函数方程(1.3)式的方法一样,首先分析 $\rho \to \infty$ 时解的渐近行为. 在 $\rho \to \infty$ 时,方程(1.38)简化为

$$(d^2/d\rho^2 - \alpha^2)u(\rho) = 0, \tag{1.39}$$

它的一般解为 $u(\rho) = Ae^{-\alpha\rho} + Be^{\alpha\rho}$,为保证 $\rho \to \infty$ 解有限,要求 $B = 0$,这样

$$u(\rho) = Ae^{-\alpha\rho}, \quad \text{当} \quad \rho \to \infty. \tag{1.40}$$

这就是 $\rho \to \infty$ 时,解应具有的形式. 具有这一形式的方程(1.38)的解可设为

$$u(\rho) = e^{-\alpha\rho}f(\rho). \tag{1.41}$$

代回(1.38)式,得到 $f(\rho)$ 的方程

$$d^2f/d\rho^2 - 2\alpha df/d\rho + [2Z/\rho - l(l+1)/\rho^2]f = 0. \tag{1.42}$$

这一方程中的二阶导数项的系数为1. 由其余两项的系数可见,$\rho = 0$ 是方程的奇点,因而需要进一步分析这一奇点附近解的形式. 在方程(1.42)中只保留 $\rho \to 0$ 时最大的两项: $d^2f/d\rho^2 - l(l+1)/\rho^2 = 0$. 令 $f = \rho^t$,得到

$$t(t-1) - l(l+1) = 0, \qquad t = l+1 \quad \text{或} \quad -l. \tag{1.43}$$

为保证 $\rho \to 0$ 时解有限,只能取 $t = l+1$. 这样,满足定解条件的解为:

$$u(\rho) = e^{-\alpha\rho}f(\rho) = e^{-\alpha\rho}\rho^{l+1}\sum_{\nu=0}^{\infty} a_\nu \rho^\nu. \tag{1.44}$$

下一步要确定级数的系数. 将上式代回方程(1.42)并按 ρ 的幂次整理得到

$$\sum_{\nu=0}^{\infty}\{a_{\nu+1}[(\nu+l+2)(\nu+l+1) - l(l+1)] +$$

$$a_\nu[2Z - 2\alpha(\nu+l+1)]\}\rho^{\nu+l} = 0.$$

为使在 ρ 取任何值时上式都成立,要求 ρ 的同幂次项的系数为零,这给出:

$$a_{\nu+1} = \frac{2\alpha(\nu + l + 1) - 2Z}{(\nu + l + 2)(\nu + l + 1) - l(l+1)} a_\nu . \tag{1.45}$$

只要给定 a_0，就可按此递推公式依次得到所有系数.

3. 用标准条件定解　能量量子化

由(1.45)式可见，当 ν 大时，$a_{\nu+1}/a_\nu \approx 2\alpha/\nu$，这和 $e^{2\alpha\rho}$ 的相邻系数比相同. 因此，$u(\rho)$ 在 ρ 大时的渐近行为是

$$u(\rho) = e^{-\alpha\rho} f(\rho) \sim e^{-\alpha\rho} \cdot e^{2\alpha\rho} = e^{\alpha\rho} \to \infty , \tag{1.46}$$

违反了标准条件. 为了使解满足标准条件，无穷级数(1.43)必须从某一项(例如 $\nu = n_r$)起中断，成为多项式. 由递推公式(1.45)可见，这归结为要求

$$\alpha(n_r + l + 1) = Z, \qquad n_r = 0,1,2,\cdots . \tag{1.47}$$

写 $n_r + l + 1 = n$，则由于 n_r 和 l 都等于 $0,1,2,\cdots$，可见 n 只能为正整数，而且 $l < n$：

$$n = 1,2,\cdots; \qquad l = 0,1,2,\cdots,n-1, \tag{1.48}$$

而 $\alpha = Z/n$. 由于 α 和 E 通过(1.36)、(1.37)式联系，使得能量取分立值：

$$E_n = -m_e Z^2 e^4/(2\hbar^2 n^2). \tag{1.49}$$

由此可见：①基态能量(零点能) $E_0 = -m_e Z^2 e^4/2\hbar^2 \neq 0$（比较第五篇第二章习题 14）. ②基态能量不简并. ③激发态能量简并. 能量对量子数 m 简并的原因是由于中心对称，能量与角动量的取向无关. 除此以外，库仑场中的能量还有对量子数 l 的附加简并，总简并度为

$$\sum_{l=0}^{n-1} (2l+1) = n^2. \tag{1.50}$$

4. 定态波函数

按(1.47)式：$\alpha = Z/n$，代入(1.45)式得

$$a_{\nu+1} = -\frac{2Z}{n} \frac{n - (l + \nu + 1)}{(\nu + 1)(2l + \nu + 2)} a_\nu .$$

由 a_0 依次按上式计算级数的系数，得到

$$f(\rho) = a_0 \rho^{l+1} \left\{ 1 - \frac{n-l-1}{1!(2l+2)} \left(\frac{2Z\rho}{n}\right) + \frac{(n-l-1)(n-l-2)}{2!!(2l+2)(2L+3)} \left(\frac{2Z\rho}{n}\right)^2 + \cdots + (-1)^{n_r} \frac{(n-l-1)(n-l-2)\cdots 1}{(n-l-1)!(2l+2)\cdots(2l+n_r+1)} \left(\frac{2Z\rho}{n}\right)^{n_r} \right\}. \tag{1.51}$$

引入变量 ξ：

$$\xi = 2Z\rho/n = 2Zr/(na_0). \tag{1.52}$$

将(1.51)式代入(1.44)、(1.28)式，最终得到所求的径向波函数为

$$R_{nl}(\xi) = N_{nl} e^{-\xi/2} \xi^l L_{n+l}^{2l+1}(\xi). \tag{1.53}$$

式中 $L_{n+l}^{2l+1}(\xi)$ 称为缔合拉盖尔多项式，它等于(1.51)式中大括号内的多项式乘以 $[(n+l)!]^2/[(n-l-1)!(2l+1)!]$，$N_{nl}$ 是归一化常数：

$$N_{nl} = \left\{\left(\frac{2Z}{na_0}\right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3}\right\}^{1/2}, \tag{1.54}$$

下面列出了几个径向波函数 $R_{nl}(r)$

$$\begin{aligned}
R_{1,0}(r) &= (Z/a_0)^{3/2}2\exp(-Zr/a_0),\\
R_{2,0}(r) &= \left(\frac{Z}{2a_0}\right)^{\frac{3}{2}}\left(2-\frac{Zr}{a_0}\right)\exp\left(-\frac{Zr}{2a_0}\right),\\
R_{2,1}(r) &= \left(\frac{Z}{2a_0}\right)^{\frac{3}{2}}\frac{Zr}{a_0\sqrt{3}}\exp\left(-\frac{Zr}{2a_0}\right),\\
R_{3,0}(r) &= \left(\frac{Z}{3a_0}\right)^{\frac{3}{2}}\left[2-\frac{4Zr}{3a_0}+\frac{4}{27}\left(\frac{Zr}{a_0}\right)^2\right]\exp\left(-\frac{Zr}{3a_0}\right),\\
R_{3,1}(r) &= \left(\frac{2Z}{a_0}\right)^{\frac{3}{2}}\left(\frac{2}{27\sqrt{3}}-\frac{Zr}{81a_0\sqrt{3}}\right)\frac{Zr}{a_0}\exp\left(-\frac{Zr}{3a_0}\right),\\
R_{3,2}(r) &= \left(\frac{2Z}{a_0}\right)^{\frac{3}{2}}\frac{1}{81\sqrt{15}}\left(\frac{Zr}{a_0}\right)^2\exp\left(-\frac{Zr}{3a_0}\right).
\end{aligned} \tag{1.55}$$

在确定了径向波函数后，按前节的讨论知道，电子在原子核的库仑场中运动时，其定态方程的解就是

$$\left.\begin{aligned}
&\psi_{nlm}(r,\theta,\varphi) = R_{nl}(r)Y_{lm}(\theta,\varphi),\\
&n=1,2,\cdots; \qquad l=0,1,\cdots,n-1;\\
&m=-l,-l+1,\cdots,l.
\end{aligned}\right\} \tag{1.56}$$

这是能量 E，角动量平方 $\hat{\boldsymbol{L}}^2$ 和角动量分量 $\hat{L}_z$ 的共同本征函数，同时也是宇称的本征函数(见 §5.3.3 例 3)．相应的本征值是：

$$E_n = -m_e Z^2 e^4/(2\hbar^2 n^2), \quad \boldsymbol{L}^2 = l(l+1)\hbar^2, \quad L_z = m\hbar^2, \quad P = (-1)^l. \tag{1.57}$$

5. 氢原子中电子的概率分布

现在我们应用上面给出的电子在库仑场中的定态波函数来分析氢原子中电子在空间各点的概率分布．当状态波函数是 $\psi_{nlm}(r,\theta,\varphi)$ 时，在 (r,θ,φ) 点附近的体积元 $\mathrm{d}\tau$ 中找到电子的概率为

$$|\psi_{nlm}(r,\theta,\varphi)|^2\mathrm{d}\tau = |R_{nl}(r)|^2|Y_{lm}(\theta,\varphi)|^2 r^2\sin\theta\mathrm{d}r\mathrm{d}\theta\mathrm{d}\varphi. \tag{1.58}$$

可以看出，由于波函数是径向部分和角向部分的乘积，所以电子的位置概率也是两部分的乘积．

径向概率分布

$$P_{nl}(r)\mathrm{d}r = |R_{nl}(r)|^2 r^2\mathrm{d}r \tag{1.59}$$

表示的是在 $r \sim r+\mathrm{d}r$ 的球壳中发现电子的概率. 径向概率分布给出的图像和玻尔理论的图像是不同的. 按照玻尔理论，只能在 r 取一定值的轨道上才能找到电子；而按照量子力学，在 $r=0$ 到 $r=\infty$ 的空间中找到电子的概率一般都不为零. 但两者也有符合之处，如在 $n=1, l=0$ 时，按玻尔理论，电子的轨道半径为 a_0，而按量子力学，这时电子的径向概率是在 $r=a_0$ 处取极大值，如图 2. 当电子处于 2p，3d，4f 态时，径向概率依次在 $4a_0, 6a_0, 16a_0$ 处取极大值，而这些地方正是玻尔理论中相应的圆轨道半径.

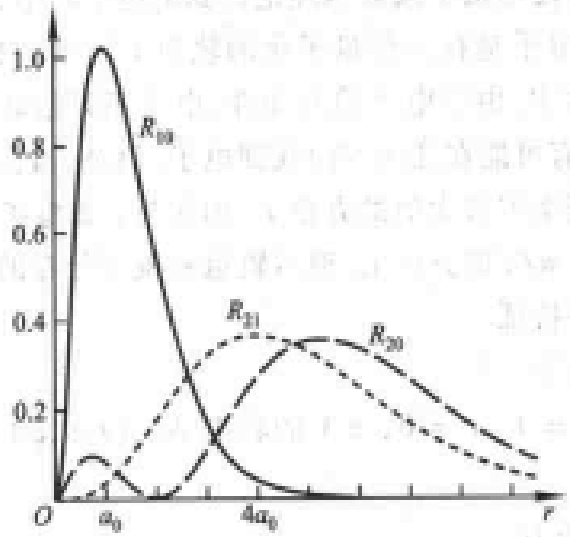


图 2　库仑场中 1s，2s，2p 态的径向概率分布

再来讨论角向概率分布，即在方向 (θ,φ) 附近的立体角中找到电子的概率

$$P_{lm}(\theta,\varphi)\mathrm{d}\Omega = |Y_{lm}(\theta,\varphi)|^2\mathrm{d}\Omega = N_{lm}^2[P_l^{|m|}(\cos\theta)]^2 d\Omega. \quad (1.60)$$

附录(Ⅳ.34)式[又见 §5.2.2(2.32)式]中给出了球谐函数 $Y_{lm}(\theta,\varphi)$

$$Y_{lm}(\theta,\varphi) = \sqrt{(l-|m|)!(2l+1)/[4\pi(l+|m|)!]}P_l^{|m|}(\cos\theta)\mathrm{e}^{\mathrm{i}m\varphi}, \quad (1.61)$$

$$Y_{00} = \sqrt{1/(4\pi)},$$

$$Y_{10} = \sqrt{3/(4\pi)}\cos\theta,$$

$$Y_{1\pm1} = \sqrt{3/(8\pi)}\sin\theta\mathrm{e}^{\pm\mathrm{i}\varphi},$$

$$Y_{20} = \sqrt{5/(16\pi)}(3\cos^2\theta-1),$$

$$Y_{2\pm1} = \sqrt{15/(8\pi)}\cos\theta\sin\theta\mathrm{e}^{\pm\mathrm{i}\varphi},$$

$$Y_{2\pm2} = \sqrt{15/(32\pi)}\sin^2\theta\mathrm{e}^{\pm2\mathrm{i}\varphi}.$$

由此得到角向概率分布如图 3. 注意 $|Y_{lm}(\theta,\varphi)|^2$ 不再与 φ 有关，所以这些图形绕 z 轴旋转对称. 例如，对 $l=1$ 的 p 态电子，若取 $m=\pm1$，有 $P_{1,\pm1}(\theta)=3\sin^2\theta/8\pi$. 它在 $\theta=\pi/2$ 处有最大值，在 $\theta=0$ 方向的值为零. 由于它绕 z 轴旋转对称，所以应将图 3 中的图形绕 z 轴旋转，由此得到的角向概率分布类似于平放

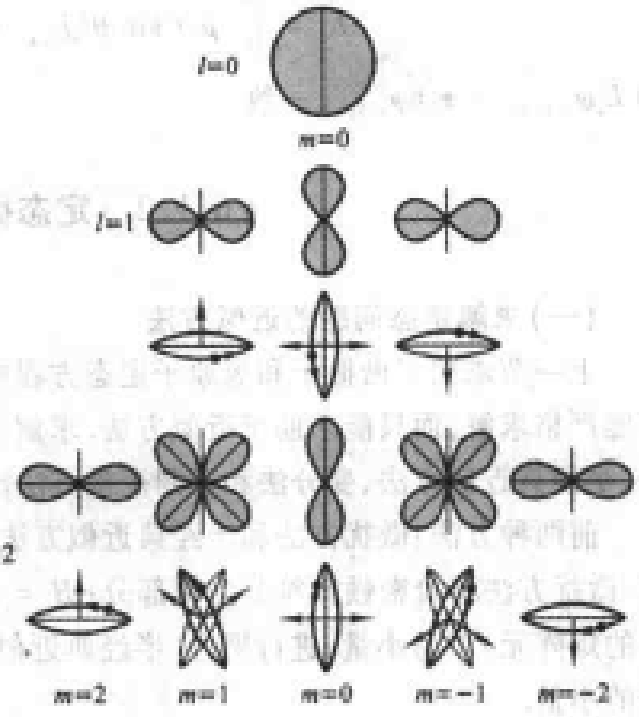


图 3　中心场中粒子的角向概率分布

着的充了气的车胎.

在图 3 中,在角向概率分布图的下面还给出了按玻尔理论得到的电子轨道,可以直观地比较两者的图像.按玻尔理论,电子是在一些量子化的轨道上运动,只在这些轨道上才能找到电子.而在量子力学中,由于电子的波动性,电子不是沿轨道运动,而是用波函数描述它的状态,因而有可能在任何方向找到电子.但是,值得注意的是,玻尔的轨道平面正好是在角向概率取极大值的方位上.例如对 p 态电子 ($l = 1$),当 $m = \pm 1$ 时,轨道平面是在 $\theta = \pi/2$ 的方位上.玻尔轨道和概率分布的这种对应说明,玻尔理论能给出原子的定性特征.

6. 氢原子中电子的概率流

在第五篇第三章习题 1 中已求得 $l = 1, m = 0, \pm 1$ 的状态 $\psi_{nlm}(r,\theta,\varphi) = R_{nl}(r)Y_{lm}(\theta,\varphi)$ 中的概率流密度:

$$j_{10} = 0,$$

$$j_{1,\pm 1} = \pm |R_{n1}|^2 \frac{\hbar}{\mu}\frac{3}{8\pi r}\sin\theta e_\varphi. \tag{1.62}$$

这表明,在球对称态 $\psi_{n,1,0}$ 中,电子的概率云是静止的,没有流动.在 $m = \pm 1$ 的状态 $\psi_{n,1,\pm 1}$ 中,电子的概率云分别绕 z 轴右旋和左旋转动.

电子概率云的转动产生角动量

$$L_z = \int_0^{2\pi} \mu r \sin\theta (j_{1,\pm 1}\cdot e_\varphi)\mathrm{d}\varphi = \pm\hbar, \tag{1.63}$$

和 $\hat{L}_z\psi_{n,1,\pm 1} = \pm\hbar\psi_{n,1,\pm 1}$ 一致.

§6.1.2 定态微扰方法

(一)求解定态问题的近似方法

上一节求出了谐振子和氢原子定态方程的严格解.在大多数情况下,定态方程不能严格求解,而只能求助于近似方法.求解定态问题常用的近似方法有:微扰方法、半经典近似方法、变分法和正则变换方法.

前两种方法:微扰方法和半经典近似方法,是采用小量展开来解定态薛定谔方程.微扰方法将哈密顿算符分为两部分:$H = H_0 + V$,将 V(实际上是 V 在 H_0 表象中的矩阵元)作为小量,进行展开.半经典近似方法则将普朗克常量 $\hbar$ 作为进行展开的小量.

后两种方法,变分法和正则变换方法,的出发点不再是定态薛定谔方程,而分别是两个和定态薛定谔方程等效的原理:变分原理和正则变换原理.按这两个原理求解定态问题,本来都是严格方法,但是在求解过程中采用近似就成了近似方法.

在用变分法时，是对希尔伯特空间作近似，即：只在能反映主要物理的一部分希尔伯特空间中作变分；在用正则变换方法时，是对哈密顿算符作近似，即：只保留能反映主要物理的哈密顿算符. 因此，运用这两种方法需要更多的物理考虑和技巧.

本节和下一节讨论微扰方法及其应用举例；随后两节简单介绍变分法和正则变换.

(二)微扰方法的基本方程

微扰方法的基本物理思想是：首先将真实系统加以简化，使简化系统的方程是可解的，然后再把真实系统哈密顿算符中没有考虑的因素考虑进来，作为对简化系统的“微扰”，在简化系统的解的基础上，得到真实系统的近似解.

举例说来，在外电场 $\boldsymbol{\varepsilon}$ 中的氢原子的哈密顿是

$$\hat{H} = -[\hbar^2/(2m_e)]\nabla^2 - (e^2/r) + e\boldsymbol{\varepsilon}\cdot\boldsymbol{r}. \tag{2.1}$$

相应的薛定谔方程无法严格求解. 因此，将它分为两部分

$$\hat{H} = \hat{H}_0 + \hat{V}, \tag{2.1'}$$

其中 $\hat{H}_0 = -[\hbar^2/(2m_e)]\nabla^2 - (e^2/r)$ 是不处于外场中的氢原子的哈密顿，它的定态薛定谔方程 $\hat{H}_0\psi_n^{(0)} = E_n^{(0)}\psi_n^{(0)}$ 已在上节中严格解出. 在此基础上来求由 $\hat{V} = e\boldsymbol{\varepsilon}\cdot\boldsymbol{r}$ 产生的对 $E_n^{(0)}$ 和 $\psi_n^{(0)}$ 的修正，就称为微扰方法.

微扰方法可以采用任意表象. 为了更普遍也为了简洁，我们将在抽象希尔伯特空间中推导微扰方法的公式，在应用这些公式进行计算时，再进入具体表象.

设 H 为系统的哈密顿算符. 将它分为两部分：$H = H_0 + H_1$，其中 H_0 称为未扰哈密顿. 按假定，H_0 的本征值方程

$$H_0|n^0\rangle = E_n^{(0)}|n^0\rangle \tag{2.2}$$

已经解出，其本征态 $|n^0\rangle$ 称为未扰态. H_1 称为微扰，假定它比 H_0 小[这一假定的确切含义见以下(2.26)式].

利用已经解出的未扰态能谱 $E_n^{(0)}$ 和态矢 $|n^0\rangle$，求定态方程

$$H|n\rangle = (H_0 + H_1)|n\rangle = E_n|n\rangle \tag{2.3}$$

的近似解，称为微扰方法.

为了清楚地看到 H_1 是小量，写 $H_1 = \lambda V$，λ 是用来标记小量幂次的参数，在计算完毕以后再令 $\lambda = 1$. 将 E_n 和 $|n\rangle$ 都按 λ 的方次展开：

$$\left.\begin{aligned} E_n &= E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots \\ |n\rangle &= |n^0\rangle + \lambda|n^1\rangle + \lambda^2|n^2\rangle + \cdots \end{aligned}\right\} \tag{2.4}$$

参数 λ 最后要等于 1，但它标志了小量的级别：和 λ^1 相乘的 $E_n^{(1)}$，$|n^1\rangle$ 是一级小量；和 λ^2 相乘的 $E_n^{(2)}$，$|n^2\rangle$ 是二级小量；….

将(2.4)式代入方程(2.3)式

$$(H_0+\lambda V)(|n^0\rangle+\lambda|n^1\rangle+\lambda^2|n^2\rangle+\cdots)$$
$$=(E_n^{(0)}+\lambda E_n^{(1)}+\lambda^2E_n^{(2)}+\cdots)(|n^0\rangle+\lambda|n^1\rangle+\lambda^2|n^2\rangle+\cdots). \tag{2.5}$$

将左右两边乘开,并令 λ 的同次幂项相等,得到

零级 $$H_0|n^0\rangle=E_n^{(0)}|n^0\rangle; \tag{2.6}$$

一级 $$H_0|n^1\rangle+V|n^0\rangle=E_n^{(0)}|n^1\rangle+E_n^{(1)}|n^0\rangle; \tag{2.7}$$

二级 $$H_0|n^2\rangle+V|n^1\rangle=E_n^{(0)}|n^2\rangle+E_n^{(1)}|n^1\rangle+E_n^{(2)}|n^0\rangle. \tag{2.8}$$

零级方程(2.6)即已经解出的方程(2.2).以下各级方程是:

一级方程 $$(H_0-E_n^{(0)})|n^1\rangle=-(V-E_n^{(1)})|n^0\rangle; \tag{2.9}$$

二级方程 $$(H_0-E_n^{(0)})|n^2\rangle=-(V-E_n^{(1)})|n^1\rangle+E_n^{(2)}|n^0\rangle. \tag{2.10}$$

下面分别按未扰态能级 $E_n^{(0)}$ 无简并和有简并两种情况求解.

(三)无简并的未扰能级的微扰修正

先解一级方程(2.9)式.用 $\langle{}^0n|$ 乘两边,得到

$$\langle{}^0n|(H_0-E_n^{(0)})|n^1\rangle=-\langle{}^0n|(V-E_n^{(1)})|n^0\rangle. \tag{2.11}$$

利用 H_0 的厄米性和未扰方程(2.6),有

$$\langle{}^0n|(H_0-E_n^{(0)})=[(H_0-E_n^{(0)})|n^0\rangle]^\dagger=0. \tag{2.12}$$

再根据未扰态的归一化

$$\langle{}^0n|E_n^{(1)}|n^0\rangle=E_n^{(1)}\langle{}^0n|n^0\rangle=E_n^{(1)},$$

由(2.11)式得到

$$E_n^{(1)}=\langle{}^0n|V|n^0\rangle. \tag{2.13}$$

这是一个重要结果,它表明:能量的一级修正等于微扰算符在未扰态中的平均值.

为了求态矢的修正,用 $\langle{}^0k|$ $(k\neq n)$ 乘一级方程(2.9)式:

$$\langle{}^0k|(H_0-E_n^{(0)})|n^1\rangle=-\langle{}^0k|(V-E_n^{(1)})|n^0\rangle$$
$$=-V_{kn}+E_n^{(1)}\langle{}^0k|n^0\rangle=-V_{kn}, \tag{2.14}$$

这里 V_{kn} 是微扰算符在未扰态之间的矩阵元

$$V_{kn}=\langle{}^0k|V|n^0\rangle. \tag{2.15}$$

对(2.14)式的左边,可以再一次应用(2.12)式中用过的方法,写

$$\langle{}^0k|H_0=(H_0|k^0\rangle)^\dagger=(E_k^{(0)}|k^0\rangle)^\dagger=E_k^{(0)}\langle{}^0k|. \tag{2.16}$$

能够这样写是由于 H_0 是厄米算符,$E_k^{(0)}$ 是实数.代入(2.14)式得到

$$(E_k^{(0)}-E_n^{(0)})\langle{}^0k|n^1\rangle=-V_{kn},\qquad k\neq n. \tag{2.17}$$

上式左边的 $\langle{}^0k|n^1\rangle$ 是态矢一级修正 $|n^1\rangle$ 在未扰态 $|k^0\rangle$ 表象中的分量.我们将它写成 $a_k^{(1)}$.这样就有

$$|\,n^1\rangle = \sum_k a_k^{(1)}\,|\,k^0\rangle, \qquad a_k^{(1)} = \langle {}^0k\,|\,n^1\rangle. \tag{2.18}$$

由于 $|\,k^0\rangle$ 已知，所以只要知道了所有的 $a_k^{(1)}$，也就知道了 $|\,n^1\rangle$. 由于假定了 $E_n^{(0)}$ 不简并，所以当 $k \neq n$ 时，$E_k^{(0)} \neq E_n^{(0)}$，因而由(2.17)式得

$$a_k^{(1)} = -V_{kn}/(E_k^{(0)} - E_n^{(0)}), \qquad k \neq n. \tag{2.19}$$

唯一尚未决定的 $a_n^{(1)}$ 可以根据归一化决定. 按(2.4)式取到 λ 的一级，有：

$$\begin{aligned} 1 &= (\langle {}^0n\,| + \lambda\langle {}^1n\,|)(|\,n^0\rangle + \lambda\,|\,n^1\rangle) \\ &= \langle {}^0n\,|\,n^0\rangle + \lambda 2\mathrm{Re}\langle {}^0n\,|\,n^1\rangle + \lambda^2\langle {}^1n\,|\,n^1\rangle, \end{aligned} \tag{2.20}$$

略去含 λ^2 的二级项. 由于 $|\,n^0\rangle$ 已归一化，$\langle {}^0n\,|\,n^0\rangle = 1$，因而，为使上式成立，只需

$$a_n^{(1)} = \langle {}^0n\,|\,n^1\rangle = 0. \tag{2.21}$$

这样，态矢的一级修正是

$$|\,n^1\rangle = \sum_k{}' \frac{V_{kn}}{E_n^{(0)} - E_k^{(0)}}\,|\,k^0\rangle. \tag{2.22}$$

其中，求和号上打撇表示对 $k \neq n$ 的所有 k 值求和.

以上在一级近似下求出了定态方程(2.3)式的解. 为求二级近似解，用 $\langle {}^0n\,|$ 乘二级方程(2.10)式

$$\langle {}^0n\,|\,(H_0 - E_n^{(0)})\,|\,n^2\rangle = -\langle {}^0n\,|\,(V - E_n^{(1)})\,|\,n^1\rangle + \langle {}^0n\,|\,E_n^{(2)}\,|\,n^0\rangle. \tag{2.23}$$

由(2.12)式知左边为零. 再用(2.21)式和(2.22)式得到

$$\begin{aligned} E_n^{(2)} &= \langle {}^0n\,|\,V\,|\,n^1\rangle = \langle {}^0n\,|\,V\Big(\sum_k{}' \frac{V_{kn}}{E_n^{(0)} - E_k^{(0)}}\,|\,k^0\rangle\Big) \\ &= \sum_k{}' \frac{V_{nk}V_{kn}}{E_n^{(0)} - E_k^{(0)}}. \end{aligned} \tag{2.24}$$

根据微扰算符的厄米性，$V_{kn} = V_{nk}^*$，有

$$E_n^{(2)} = \sum_k{}' \frac{|\,V_{nk}\,|^2}{E_n^{(0)} - E_k^{(0)}}. \tag{2.25}$$

这就是定态能量的二级修正. 再用 $\langle {}^0k\,|$ $(k \neq n)$ 乘二级方程(2.10)式，可以得到态矢的二级修正.

从以上的计算可以看到，在各级近似的微扰计算中，总是先求能级的修正，再求态矢的修正. 在求某一级近似的能级修正时，只需要用到上一级近似的态矢. 这是逐级迭代求解的特点.

另外，在以上计算中用到的(2.12)式和(2.16)式有普遍意义. 它们表明，厄米算符对狄拉克矢量的作用既可以向右作用到右矢上，也可以向左作用到左矢上. 例

如,(2.16)式表明,厄米算符 H_0 对左本征矢 $\langle {}^0k|$ 作用得到本征值 $E_k^{(0)}$ 乘本征矢;这和它对右本征矢 $|k^0\rangle$ 作用,得到本征值乘本征矢一样.这一规则在计算中常要用到.

为使微扰方法能够应用,要求微扰级数(2.4)式收敛.这是一个复杂问题,但很容易得到一个必要条件,即各级修正应该愈来愈小.由(2.22)式可见,这意味着要求

$$|V_{kn}| \ll |E_n^{(0)} - E_k^{(0)}|, \tag{2.26}$$

即:微扰算符在未扰态之间的矩阵元远小于未扰能级的间距,这就是前面说,微扰算符 H_1(即 V)比 H_0 小的含义.

在计算具体例子之前,我们再对(2.22)和(2.25)式作一点形式上的改写.

将(2.15)式代入(2.22)式,得到

$$|n^1\rangle = \sum_k{}' |k^0\rangle \frac{\langle {}^0k|V|n^0\rangle}{E_n^{(0)} - E_k^{(0)}}.$$

它可以写成

$$|n^1\rangle = G_n V|n^0\rangle, \qquad G_n = \sum_k{}' \frac{|k^0\rangle\langle {}^0k|}{E_n^{(0)} - E_k^{(0)}}. \tag{2.27}$$

再代入(2.24)式得到

$$E_n^{(2)} = \langle {}^0n|V|n^1\rangle = \langle {}^0n|VG_nV|n^0\rangle. \tag{2.28}$$

比较(2.13)式可见,能量的一级和二级修正 $E_n^{(1)}$ 和 $E_n^{(2)}$ 分别是算符 V 和 VG_nV 在未扰态 $|n^0\rangle$ 中的平均值.

【例 1】 求非谐振子的能级.

解 由 §1.4.1 的讨论知道,当系统有平衡位置时,将势能在平衡位置附近展开,只保留不为零的第一项,就是谐振子.如果进一步考虑展开式中的高级项,则得到非谐振子.将势能展开到四级项:

$$U(x) = m\omega^2x^2/2 + \alpha x^3 + \beta x^4, \tag{2.29}$$

则哈密顿算符可写为 $H = H_0 + V$

$$H_0 = -[\hbar^2/(2m)]\mathrm{d}^2/\mathrm{d}x^2 + (m\omega^2/2)x^2, \qquad V = \alpha x^3 + \beta x^4. \tag{2.30}$$

下面将看到,V 中的 x^3 项和 x^4 项所引起的能量修正有相同量级,因而应该同时保留.

在未扰态(即谐振子态),x 的矩阵元已在 §5.4.5 中算出,见那一节的(5.32)式.不为零的矩阵元是

$$x_{n-1,n} = x_{n,n-1} = \sqrt{n\hbar/(2m\omega)}. \tag{2.31}$$

利用矩阵乘法可以得到 x^2, x^3 的不为零的矩阵元

$$\left.\begin{aligned} x^2_{n,n} &= x_{n,n-1}x_{n-1,n} + x_{n,n+1}x_{n+1,n} = \frac{\hbar}{2m\omega}(2n+1), \\ x^2_{n,n-2} &= x^2_{n-2,n} = x_{n,n-1}x_{n-1,n-2} = \frac{\hbar}{2m\omega}\sqrt{(n^2-n)}, \\ x^2_{n,n+2} &= x^2_{n+2,n} = x_{n,n+1}x_{n+1,n+2} = \frac{\hbar}{2m\omega}\sqrt{(n+2)(n+1)}, \end{aligned}\right\} \tag{2.32}$$

$$\left.\begin{aligned} x^3_{n-1,n} &= x^3_{n,n-1} = x^2_{n,n-2}x_{n-2,n-1} + x^2_{n,n}x_{n,n-1} = \left(\frac{\hbar}{2m\omega}\right)^{3/2} 3\sqrt{n^3}, \\ x^3_{n,n-3} &= x^3_{n-3,n} = x^2_{n,n-2}x_{n-2,n-3} = \left(\frac{\hbar}{2m\omega}\right)^{3/2}\sqrt{n(n-1)(n-2)}. \end{aligned}\right\} \tag{2.33}$$

此外，$x^3_{n+1,n} = x^3_{n,n+1}$ 和 $x^3_{n+3,n} = x^3_{n,n-3}$ 可以分别由(2.33)中的两式作代换 $n \to n+1$ 和 $n \to n+3$ 而得到. 由此可见，x^3 没有对角元：$x^3_{n,n} = 0$，因此，αx^3 引起的一级能量修正为零. 不为零的是二级能量修正：

$$\begin{aligned} E^{(2)}_n(\alpha x^3) &= \alpha^2\left\{\frac{|x^3_{n,n-1}|^2}{E^{(0)}_n - E^{(0)}_{n-1}} + \frac{|x^3_{n,n-3}|^2}{E^{(0)}_n - E^{(0)}_{n-3}} + \frac{|x^3_{n,n+1}|^2}{E^{(0)}_n - E^{(0)}_{n+1}} + \frac{|x^3_{n,n+3}|^2}{E^{(0)}_n - E^{(0)}_{n+3}}\right\} \\ &= -\frac{15}{4}\frac{\alpha^2}{\hbar\omega}\left(\frac{\hbar}{m\omega}\right)^3\left(n^2+n+\frac{11}{30}\right), \end{aligned} \tag{2.34}$$

这里用了 $E^{(0)}_n = (n+1/2)\hbar\omega$.

由(2.31)～(2.33)式知，x^4 的对角元不为零：

$$x^4_{n,n} = x^2_{n,n-2}x^2_{n-2,n} + x^2_{n,n+2}x^2_{n+2,n} + (x^2_{n,n})^2 = [\hbar/(2m\omega)]^2 3(2n^2+2n+1), \tag{2.35}$$

由此引起的能量一级修正是

$$E^{(1)}_n(\beta x^4) = (3\beta/2)[\hbar/(m\omega)]^2(n^2+n+1/2). \tag{2.36}$$

由(2.29)式知 β 相对于 α 是高级小量，但和 α^2 同级. 因此，(2.34)式和(2.36)式属于同级小量. 准确到这一级小量的非谐振子能级是

$$E_n = E^{(0)}_n + E^{(1)}_n(\beta x^4) + E^{(2)}_n(\alpha x^3). \tag{2.37}$$

(四)有简并的未扰能级的微扰修正

如果未扰能级 $E^{(0)}_n$ 有简并，则当 $k \neq n$ 时也会因简并态所对应的能量相等使条件(2.26)式不满足. 现在来讨论在这种情况下如何计算微扰修正.

给态矢重新编号，将属于简并能级 $E^{(0)}_n$ 的态矢都标以量子数 n，并补充另一个量子数来区分它们. 如果 $E^{(0)}_n$ 有 s 重简并，则和它对应有 s 个正交归一的本征矢量 $|ni^0\rangle (i = 1,2,\cdots,s)$. 注意，其他未扰能级，不论是否有简并，仍用一个脚标 $k(\neq n)$ 标志. 如果从这一简并能级的某一本征矢，例如 $|nj^0\rangle$ 出发，按上面的方法做微扰，则(2.22)式中的求和既包括属于不同能级 $E^{(0)}_k \neq E^{(0)}_n$ 的态矢 $|k^0\rangle$，也包括属

于同一简并能级的不同态矢 $|ni^0\rangle (i \neq j)$：

$$|nj^1\rangle = \sum_k{}' \frac{|V_{kn}|^2}{(E_n^{(0)} - E_k^{(0)})} |k^0\rangle + \sum_i{}' \frac{|V_{ni,nj}|^2}{(E_{nj}^{(0)} - E_{ni}^{(0)})} |ni^0\rangle. \quad (2.38)$$

后一个和式各项的分母 $E_{nj}^{(0)} - E_{ni}^{(0)} = 0$，而分子 $V_{ni,nj}$ 一般说来不为零，微扰方法失效.

为了克服这一困难，注意 $E_n^{(0)}$ 的本征矢 $|ni^0\rangle (i = 1,2,\cdots,s)$ 形成一个 s 维子空间，其中的任一矢量都是和 $E_n^{(0)}$ 对应的 H_0 的本征矢. 可以从中任选 s 个正交归一矢量

$$|n\alpha^0\rangle = \sum_{i=1}^{s} c_{i\alpha} |ni^0\rangle, \quad c_{i\alpha} = \langle {}^0ni | n\alpha^0\rangle, \qquad \alpha = 1,2,\cdots,s \quad (2.39)$$

为基：

$$H_0 |n\alpha^0\rangle = E_n^{(0)} |n\alpha^0\rangle. \quad (2.40)$$

以 $|n\alpha^0\rangle$ 作为零级近似，一级近似方程(2.9)式成为

$$H_0 |n\alpha^1\rangle + V |n\alpha^0\rangle = E_n^{(0)} |n\alpha^1\rangle + E_{n\alpha}^{(1)} |n\alpha^0\rangle, \quad (2.41)$$

将(2.39)式中的 i 改写为 j 代入，再用 $\langle {}^0ni|$ 左乘，令 $\langle {}^0ni | V | nj^0\rangle = V_{ij}$，得

$$\sum_{j=1}^{s} V_{ij} c_{j\alpha} = E_{n\alpha}^{(1)} c_{i\alpha}. \quad (2.42)$$

这一线性齐次联立方程组有非零解的条件

$$|V_{ij} - E_{n\alpha}^{(1)} \delta_{ij}| = 0 \quad (2.43)$$

称为久期方程. 它是 $E_{n\alpha}^{(1)}$ 的一个 s 次代数方程. 由此解出 s 个根 $E_{n\alpha}^{(1)} (\alpha = 1,2,\cdots, s)$ 就是一级能量修正.

将 $E_{n\alpha}^{(1)}$ 代入(2.42)式，可解出 $c_{i\alpha}, i = 1,\cdots,s; \alpha = 1,\cdots,s$. 再代入(2.39)式，就得到 $|n\alpha^0\rangle, \alpha = 1,\cdots,s$.

将(2.41)式中的 α 改写为 β，用 $\langle n\alpha^0|$ 左乘，利用 $\langle {}^0n\alpha | H_0 = \langle {}^0n\alpha | E_n^{(0)}$，得到

$$E_{n\alpha}^{(1)} \delta_{\alpha\beta} = \langle {}^0n\alpha | V | n\beta^0\rangle. \quad (2.44)$$

特别是，当 $\beta = \alpha$ 时有

$$E_{n\alpha}^{(1)} = \langle {}^0n\alpha | V | n\alpha^0\rangle. \quad (2.44')$$

这表明，在通过解方程(2.43)、(2.42)得到 $|n\alpha^0\rangle$ 以后，就可以直接在 $|n\alpha^0\rangle$ 态中求微扰算符的平均值得到能量的一级修正 $E_{n\alpha}^{(1)}$，和未扰态没有简并的情况一样[比较(2.44′)式和(2.13)式]. 因此，称 $|n\alpha^0\rangle$ 为“正确的未扰态”.

由(2.44)式可见，当采用“正确的未扰态” $|n\alpha^0\rangle$ 时，微扰算符 V 在和 $E_n^{(0)}$ 对应的 s 维简并子空间中是对角的. 因此，在态矢的一级修正 $|n\alpha^1\rangle$ 用零级态矢展开的展开式中不包含 $|n\beta^0\rangle (\beta \neq \alpha)$ 的项：$\langle {}^0n\beta | n\alpha^1\rangle = 0 \ (\beta \neq \alpha)$.

用$\langle {}^0k \mid (k \neq n)$左乘(2.41)式得

$$\langle {}^0k \mid n\alpha^1\rangle = V_{k\alpha}/(E_n^{(0)} - E_k^{(0)}) \qquad (k \neq n). \tag{2.45}$$

这是一级态矢$\mid n\alpha^1\rangle$用零级态矢展开式中$\mid k^0\rangle (k \neq n)$的系数.类似于(2.20)式可以证明这一展开式中$\mid n\alpha^0\rangle$的系数为零:$\langle {}^0 n\alpha \mid n\alpha^1\rangle = 0$. 于是就有

$$\mid n\alpha^1\rangle = \sum_k{}' \frac{V_{k\alpha}}{E_n^{(0)} - E_k^{(0)}} \mid k^0\rangle. \tag{2.46}$$

其中撇号表示对$k \neq n$求和.

以下分几种情况讨论.

情况Ⅰ　久期方程的s个根都不相等

此时,在一级近似下微扰已将简并完全拉开.如果要进一步作二级修正,将(2.46)式代入二级方程(2.10)式

$$H_0 \mid n\alpha^2\rangle + V \mid n\alpha^1\rangle = E_n^{(0)} \mid n\alpha^2\rangle + E_{n\alpha}^{(1)} \mid n\alpha^1\rangle + E_{n\alpha}^{(2)} \mid n\alpha^0\rangle.$$

用$\langle {}^0n\alpha \mid$左乘上式两边.根据(2.12)式,等式两边第一项相消;又由于$\langle {}^0n\alpha \mid n\alpha^1\rangle = 0$,右边第二项为零;再将(2.46)式代入得到

$$E_{n\alpha}^{(2)} = \sum_k{}' \mid V_{k\alpha} \mid^2/(E_n^{(0)} - E_k^{(0)}). \tag{2.47}$$

这和未扰能级无简并时的(2.25)式相同.它表明,在选择了正确的未扰态以后,如果久期方程无重根,就可以直接应用无简并的微扰公式.

情况Ⅱ　久期方程的s个根中有$r \leqslant s$个相等

简并只是部分地被拉开,还有r重简并继续保持.往下再分成两种情况讨论:

情况Ⅱa　久期方程的r个相等的根$E_{n\alpha}^{(1)} (\alpha = 1, \cdots, r)$是一级小量

此时,一级方程(2.9)和二级方程(2.10)仍然可以分别考虑.在此情况下,可以和情况Ⅰ同样地计算二级能量修正$E_{n\alpha}^{(2)}$,如(2.47)式.如果这些$E_{n\alpha}^{(2)} (\alpha = 1, \cdots, r)$不全相等,简并就在二级近似下被拉开.

情况Ⅱb　久期方程的r个相等的根全为0,或者和二级小量同级

此时,$E_{n\alpha}^{(1)} (\alpha = 1, \cdots, r)$并不象它的上标"(1)"所标志的那样是一级小量,而是等于零或者是二级小量,"一级方程"(2.9)式或(2.41)式的右边第二项是二级小量,因而应将一级和二级方程[即(2.9)、(2.10)式]一道考虑.将(2.9)式和(2.10)式相加,略去三级小量$E_{n\alpha}^{(1)} \mid n\alpha^1\rangle$,并将二级小量$E_{n\alpha}^{(1)} + E_{n\alpha}^{(2)}$合写为$E_{n\alpha}^{(2)}$得到

$$\begin{aligned} &H_0(\mid n\alpha^1\rangle + \mid n\alpha^2\rangle) + V(\mid n\alpha^0\rangle + \mid n\alpha^1\rangle) \\ &\quad = E_n^{(0)}(\mid n\alpha^1\rangle + \mid n\alpha^2\rangle) + E_n^{(2)} \mid n\alpha^0\rangle. \end{aligned} \tag{2.48}$$

用$\langle {}^0ni \mid$左乘,利用(2.39)式和(2.46)式得到

$$\sum_{j=1}^{r}\left(V_{ij}+\sum_{k}{}'\frac{V_{ik}V_{kj}}{E_n^{(0)}-E_k^{(0)}}\right)c_{j\alpha}=E_n^{(2)}c_{i\alpha}. \tag{2.49}$$

和(2.42)式比较可见,差别只在于用圆括号中的表达式替换了那里的 V_{ij}. 这使久期方程(2.43)式变为

$$\left|\left(V_{ij}+\sum_{k}{}'\frac{V_{ik}V_{kj}}{E_n^{(0)}-E_k^{(0)}}\right)-E_n^{(2)}\delta_{ij}\right|=0,$$

$$i,j=1,2,\cdots,r. \tag{2.50}$$

由此解出 r 个 $E_n^{(2)}$ 如果不全相等,简并就被进一步拉开了.

我们来讨论一下能级简并及其被微扰拉开的物理原因. 正如前面多次指出的,能量(在这里是指未扰态能量)的简并常常和某种对称性相联系. 微扰破坏了对称性,使简并解除或部分解除. 有些情况下,可以通过这样的对称性分析直接写出正确的未扰态,而不必解久期方程(2.43)式和方程(2.42)式. 有了正确的未扰态,就可以直接应用未扰态能量无简并的微扰公式.

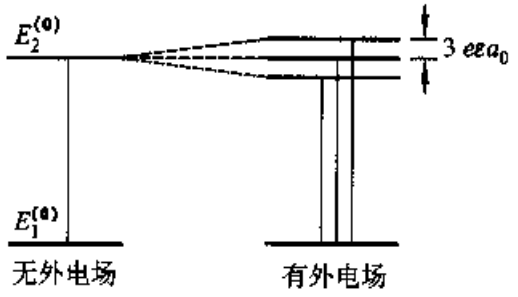


图 1 氢原子能级在外电场中的分裂

【例 2】 设氢原子受到外电场的作用,哈密顿如(2.1)式. 势能为 $e\varepsilon r\cos\theta$. 这里 ε 是场强大小,其方向为 z 轴的方向,r 是氢原于中电于离开原子核的距离,θ 为 $\boldsymbol{r}$ 与电场方向的夹角. 将外电场看成微扰,试计算在它的作用下氢原子的前两个能级的分裂情况和相应的谱线分裂情况.

解 在电场中氢原于的哈密顿算符为:

$$\hat{H}=\hat{H}_0+e\varepsilon r\cos\theta, \tag{2.51}$$

其中第一项 $\hat{H}_0$ 是无外场时氢原子的哈密顿算符,第二项是微扰 V:

$$\left.\begin{aligned}\hat{H}_0&=-[\hbar^2/(2m_e)]\nabla^2-(e^2/r),\\ V&=e\varepsilon r\cos\theta.\end{aligned}\right\} \tag{2.52}$$

H_0 的定态方程的解已在 §6.1.1 给出,并且知道 $E_n^{(0)}$ 有 n^2 重简并. 在微扰作用下,能级会发生分裂,简并将会有所消除. 现在要具体计算的是 $n=1$ 和 $n=2$ 时能级的分裂情况.

当 $n=1$ 时,$E_1^{(0)}$ 没有简并,因而它在外场作用下没有分裂现象.

当 $n=2$ 时,$E_2^{(0)}$ 有四重简并,相应的四个波函数为

$$\left.\begin{aligned}
\Phi_1 &= \psi_{200}^{(0)} = \frac{1}{4\sqrt{2\pi}}\left(\frac{1}{a_0}\right)^{\frac{3}{2}}\left(2-\frac{r}{a_0}\right)e^{-r/2a_0}\\
\Phi_2 &= \psi_{210}^{(0)} = \frac{1}{4\sqrt{2\pi}}\left(\frac{1}{a_0}\right)^{\frac{3}{2}}\left(\frac{r}{a_0}\right)e^{-r/2a_0}\cos\theta\\
\Phi_3 &= \psi_{211}^{(0)} = \frac{1}{8\sqrt{2\pi}}\left(\frac{1}{a_0}\right)^{\frac{3}{2}}\left(\frac{r}{a_0}\right)e^{-r/2a_0}\sin\theta e^{i\varphi}\\
\Phi_4 &= \psi_{21-1}^{(0)} = \frac{1}{8\sqrt{2\pi}}\left(\frac{1}{a_0}\right)^{\frac{3}{2}}\left(\frac{r}{a_0}\right)e^{-r/2a_0}\sin\theta e^{-i\varphi}
\end{aligned}\right\} \tag{2.53}$$

为计算 $E_2^{(0)}$ 的一级修正，需要知道微扰算符 $\hat{V}$ 在这四个简并波函数中的矩阵元. 注意到正交性条件[见附录(Ⅰ.22)式]

$$\int_0^{2\pi} e^{i(m-m')\varphi}d\varphi = 2\pi\,\delta_{mm'}, \tag{2.54}$$

可知 $\hat{V}$ 在有不同磁量子数的波函数之间的矩阵元为零. 而在(2.53)式的四个波函数中，只有 Φ_1 和 Φ_2 有相同磁量子数，因而不为零的非对角元只有 V_{12} 和 V_{21}. 另外，当用这几个波函数计算 $\hat{V}$ 的对角元素

$$V_{nn} = \int \Phi_n^* e\varepsilon r\cos\theta\Phi_n d\tau = \int |\Phi_n|^2 e\varepsilon r\cos\theta r^2 dr d\cos\theta d\varphi$$

时，由于 $|\Phi_n|^2\cos\theta$ 是 $\cos\theta$ 的奇函数[见(2.53)式]，因而上式积分为零. 这样，仅有的不为零的矩阵元是 V_{12} 和 V_{21}. 以下完成对它们的计算：

$$V_{12} = V_{21} = \frac{1}{32\pi}\left(\frac{1}{a_0}\right)^3\int\left(2-\frac{r}{a_0}\right)\frac{r}{a_0}e^{-r/a_0}\times\cos\theta e\varepsilon r\cos\theta r^2\sin\theta dr d\theta d\varphi = 3e\varepsilon a_0.$$

将这些结果代入久期方程(2.43)中，得

$$\begin{vmatrix}
-E_2^{(1)} & 3e\varepsilon a_0 & 0 & 0\\
3e\varepsilon a_0 & -E_2^{(1)} & 0 & 0\\
0 & 0 & -E_2^{(1)} & 0\\
0 & 0 & 0 & -E_2^{(1)}
\end{vmatrix} = 0,$$

即 $(E_2^{(1)})^2[(E_2^{(1)})^2-(3e\varepsilon a_0)^2] = 0$. 由它得到 $E_2^{(1)}$ 的四个根：

$$E_2^{(1)} = 3e\varepsilon a_0,\ -3e\varepsilon a_0,\ 0,\ 0. \tag{2.55}$$

这就是说，受微扰作用后 $E_2^{(0)}$ 分裂为三个能级，简并得到部分解除. 图 1 画出了 $E_1^{(0)}$ 和 $E_2^{(0)}$ 的分裂情况.

由于 $E_2^{(0)}$ 的能级分裂为三条能级，因而相应于从 $E_2^{(0)}$ 到 $E_1^{(0)}$ 辐射跃迁的一条谱线，在外电场中分裂成了三条. 谱线在外电场中的分裂现象叫做斯塔克效应.

§6.1.3 氦原子

除了氢原子外,最简单的原子是氦原子.它只有两个电子,其哈密顿算符为

$$\hat{H} = -\frac{\hbar^2}{2\mu}\nabla_a^2 - \frac{\hbar^2}{2\mu}\nabla_b^2 - \frac{2e^2}{r_a} - \frac{2e^2}{r_b} + \frac{e^2}{r_{ab}}, \tag{3.1}$$

其中 μ 为电子和氦原子核的约化质量,$\boldsymbol{r}_a$ 和 $\boldsymbol{r}_b$ 分别为二电子的位矢,r_{ab} 为二电子间的距离.下面首先用微扰方法来近似地计算能级,然后用全同性原理作一般性讨论.

(一)微扰计算

将(3.1)式的最后一项视为微扰,写

$$\hat{H} = \hat{H}_0 + \hat{V},$$

$$\hat{H}_0 = \hat{H}_{a0} + \hat{H}_{b0} = -[\hbar^2/(2\mu)]\nabla_a^2 - 2e^2/r_a - [\hbar^2/(2\mu)]\nabla_b^2 - 2e^2/r_b, \tag{3.2}$$

$$\hat{V} = e^2/r_{ab}. \tag{3.3}$$

未扰哈密顿算符 $\hat{H}_0$ 分成了仅含 $\boldsymbol{r}_a$ 与仅含 $\boldsymbol{r}_b$ 的两项之和,因此未扰态可用分离变量法写成仅含 $\boldsymbol{r}_a$ 与仅含 $\boldsymbol{r}_b$ 的二函数之积:

$$\psi_1(\boldsymbol{r}_a, \boldsymbol{r}_b) = \varphi_n(\boldsymbol{r}_a)\varphi_m(\boldsymbol{r}_b), \tag{3.4}$$

其中 $\varphi_n(\boldsymbol{r}_a)$ 与 $\varphi_m(\boldsymbol{r}_b)$ 是方程

$$[-\hbar^2\nabla^2/(2\mu) - 2e^2/r]\varphi_k(\boldsymbol{r}) = E_k^{(0)}\varphi_k(\boldsymbol{r}) \tag{3.5}$$

的解.因为 $\hat{H}_{a0}$ 与 $\hat{H}_{b0}$ 的形式完全相同,所以 $\varphi_n(\boldsymbol{r}_a)$ 与 $\varphi_m(\boldsymbol{r}_b)$ 是同一形式的方程(3.5)(即 $z = 2$ 的库仑场方程)的解.它已经在 §6.1.1 的(1.56)式中求出.

由(3.4)式知:

$$\begin{aligned}\hat{H}_0\psi_1 &= \varphi_m(\boldsymbol{r}_b)\hat{H}_{a0}\varphi_n(\boldsymbol{r}_a) + \varphi_n(\boldsymbol{r}_a)\hat{H}_{b0}\varphi_m(\boldsymbol{r}_b)\\ &= (E_n^{(0)} + E_m^{(0)})\psi_1 = E^{(0)}\psi_1,\end{aligned}$$

很明显,函数

$$\psi_2(\boldsymbol{r}_a, \boldsymbol{r}_b) = \varphi_m(\boldsymbol{r}_a)\varphi_n(\boldsymbol{r}_b) \tag{3.6}$$

同样也是 $\hat{H}_0$ 的本征态,而且也对应于本征值 $E_n^{(0)} + E_m^{(0)}$.未扰态中同一能量 $E_n^{(0)} + E_m^{(0)}$ 对应于两个态 ψ_1 和 ψ_2,所以它有二重简并.因为由 ψ_1 交换 $\boldsymbol{r}_a, \boldsymbol{r}_b$ 可得 ψ_2,所以这种简并叫做交换简并.但应注意,这种"简并"之所以产生,是由于我们在哈密顿中人为地分出了"微扰" V,而实际上电子不可能只与核作用而自己之间无相互作用.因此这种"交换简并"并不是真实具有的简并.

根据有简并情况下的一般微扰方法,令正确的零级波函数为

$$\psi = c_1\psi_1 + c_2\psi_2, \tag{3.7}$$

代入方程 $\hat{H}\psi = E\psi$，得到

$$V\psi = c_1 V\psi_1 + c_2 V\psi_2 = (E - E^{(0)})(c_1\psi_1 + c_2\psi_2).$$

分别乘上 $\psi_1^* \mathrm{d}\tau$ 和 $\psi_2^* \mathrm{d}\tau$ 积分得二方程：

$$V_{11}c_1 + V_{12}c_2 = (E - E^{(0)})c_1,$$
$$V_{21}c_1 + V_{22}c_2 = (E - E^{(0)})c_2.$$

其中

$$\begin{aligned} V_{11} &= \int |\psi_1|^2 V \mathrm{d}^3 r_a \mathrm{d}^3 r_b \\ &= e^2 \int \frac{|\varphi_n(\boldsymbol{r}_a)|^2 |\varphi_m(\boldsymbol{r}_b)|^2}{r_{ab}} \mathrm{d}^3\boldsymbol{r}_a \mathrm{d}^3\boldsymbol{r}_b; \\ V_{22} &= \int |\psi_2|^2 V \mathrm{d}^3 r_a \mathrm{d}^3 r_b = V_{11} = K; \end{aligned} \tag{3.8}$$

$$\begin{aligned} V_{12} &= \int \psi_1^* V \psi_2 \mathrm{d}^3\boldsymbol{r}_a \mathrm{d}^3\boldsymbol{r}_b \\ &= e^2 \int \frac{\varphi_n^*(\boldsymbol{r}_a)\varphi_m^*(\boldsymbol{r}_b)\varphi_m(\boldsymbol{r}_a)\varphi_n(\boldsymbol{r}_b)}{r_{ab}} \mathrm{d}^3\boldsymbol{r}_a \mathrm{d}^3\boldsymbol{r}_b, \\ V_{21} &= \int \psi_2^* V \psi_1 \mathrm{d}^3\boldsymbol{r}_a \mathrm{d}^3\boldsymbol{r}_b = V_{12} = A. \end{aligned} \tag{3.9}$$

$V_{11} = V_{22}$ 及 $V_{12} = V_{21}$ 是由于这些积分都是对于 $\boldsymbol{r}_a, \boldsymbol{r}_b$ 的遍及整个空间的积分，交换脚标 a, b 不改变积分的值.

令 $E = E^{(0)} + E^{(1)}$，上二式就成为

$$\left.\begin{aligned} (K - E^{(1)})c_1 + Ac_2 &= 0, \\ Ac_1 + (K - E^{(1)})c_2 &= 0. \end{aligned}\right\} \tag{3.10}$$

这就是 §6.1.2(2.42)式. 为使它对 c_1, c_2 有非零解，需要系数行列式等于零

$$\begin{vmatrix} K - E^{(1)} & A \\ A & K - E^{(1)} \end{vmatrix} = 0.$$

此即久期方程. 解之得 $E^{(1)} = K \pm A$，故得能级

$$E = E^{(0)} + K \pm A = E_m^{(0)} + E_n^{(0)} + K \pm A. \tag{3.11}$$

由此可见，交换简并被解除.

将 $E^{(1)} = K \pm A$ 代入(3.10)式得 $c_1 = \pm c_2$，故得二态：

$$\psi_s = (1/\sqrt{2})(\psi_1 + \psi_2), \qquad \psi_a = (1/\sqrt{2})(\psi_1 - \psi_2). \tag{3.12}$$

前者对于两粒子交换是对称态，后者是反对称态.

正确的未扰态(3.12)式是对称态和反对称态并不是偶然的. 它是由于全同性原理. 根据这一原理，波函数对两个电子的全部变量——空间变量 $\boldsymbol{r}$ 和自旋变量 σ

的交换反对称，因而空间波函数对 r_1 和 r_2 的交换或者对称或者反对称，见 §5.5.2 (2.6)式. 因此，利用全同性原理可以不必解久期方程而直接写出正确的未扰态 (3.12)式.

(二)正氦与仲氦

根据全同性原理，轨道波函数和自旋波函数应有相反的对称性，见 §5.5.2 (2.6)式. 用脚标 s 与 a 表示对称与反对称函数则有

$$\Psi_1 = \psi_s \chi_a, \qquad \Psi_2 = \psi_a \chi_s. \tag{3.13}$$

在氦的状态 Ψ_1 中，轨道波函数是对称函数，与它对应的自旋波函数只能是 χ_a [见 §5.5.2(2.9)第4式]，其自旋 $S = 0$，这种状态的氦称为仲氦.

在氦的状态 Ψ_2 中，轨道波函数是反对称函数，与它对应的自旋波函数有三个：χ_{s1}、χ_{s2}、χ_{s3} [见 §5.5.2(2.9)前三式]，其自旋 z 分量分别为 1，0，−1. 这种状态的氦称为正氦，它具有总自旋 $S = 1$.

在氦的基态中，两个电子都处于最低能态，此时轨道波函数

$$\psi = \varphi_1(\boldsymbol{r}_a)\varphi_1(\boldsymbol{r}_b)$$

是对称函数，故知氦的基态是仲氦.

(三)交换作用

我们来进一步讨论前面用微扰方法求得的能量(3.11)式的物理意义. 在(3.8)、(3.9)式中引用符号：

$$\rho_{nn}(\boldsymbol{r}) = -e|\varphi_n(\boldsymbol{r})|^2, \qquad \rho_{mm}(\boldsymbol{r}) = -e|\varphi_m(\boldsymbol{r})|^2, \tag{3.14}$$

$$\rho_{mn}(\boldsymbol{r}) = -e\varphi_m^*(\boldsymbol{r})\varphi_n(\boldsymbol{r}), \tag{3.15}$$

则得

$$K = V_{11} = \int \frac{\rho_{nn}(\boldsymbol{r}_a)\rho_{mm}(\boldsymbol{r}_b)}{r_{ab}} \mathrm{d}^3 r_a \mathrm{d}^3 r_b, \tag{3.16}$$

$$A = V_{12} = \int \frac{\rho_{mn}(\boldsymbol{r}_a)\rho_{nm}^*(\boldsymbol{r}_b)}{r_{ab}} \mathrm{d}^3 r_a \mathrm{d}^3 r_b, \tag{3.17}$$

由(3.14)式知 ρ_{nn} 可视为第 n 态的电子在空间的电荷密度分布，因此可以将 K 看作分别处于 n 态和 m 态的两个电子的电子云的相互作用能. 附加的一项 A 叫做交换作用能，它是由于全同性原理使系统的轨道运动状态必须对于粒子的交换或者对称或者反对称而产生的，参看 §5.5.2 末尾的讨论.

实际上“交换作用”并不是一种什么新的力的作用，它本质上仍然是库仑作用 e^2/r_{ab}，是库仑作用加上全同性原理而产生的.

最后，让我们再次强调，前面进行微扰计算是通过解久期方程的途径. 实际上，由全同性原理，不必解久期方程也可以知道正确的未扰态波函数一定是或者对称或者反对称. 因此，可以直接写出：$\psi = \psi_s$ 或 ψ_a，如(3.12)式，然后通过计算微扰

算符 V 在这两个态中的平均值[§6.1.2(2.44′)式]求 $E^{(1)}$:

$$E_s^{(1)} = \overline{\left(\frac{e^2}{r_{ab}}\right)} = \int |\psi_s|^2 \frac{e^2}{r_{ab}} \mathrm{d}^3 r_a \mathrm{d}^3 r_b$$

$$= \frac{e^2}{2} \int \frac{|\psi_1|^2 + |\psi_2|^2 + \psi_1^* \psi_2 + \psi_2^* \psi_1}{r_{ab}} \mathrm{d}^3 r_a \mathrm{d}^3 r_b$$

$$= K + A. \tag{3.18}$$

同样可求得

$$E_a^{(1)} = \int |\psi_a|^2 \frac{e^2}{r_{ab}} \mathrm{d}^3 r_a \mathrm{d}^3 r_b = K - A. \tag{3.19}$$

§6.1.4　变分法

在实际计算中,有些系统不适于用微扰方法求解定态方程,因而在量子力学中建立了其他求解定态问题的近似方法. 变分法是常用来计算系统的基态和前几个分立激发态的近似方法. 它是基于下述变分原理.

(一)变分原理

设系统的哈密顿算符为 $\hat{H}$,在任意态函数 ψ 所描述的状态中,能量的平均值为

$$\overline{H} = \int \psi^* \hat{H} \psi \mathrm{d}\tau \Big/ \int \psi^* \psi \mathrm{d}\tau. \tag{4.1}$$

这里,为了一般起见,不假定 ψ 已归一化.

变分原理指出:用任意波函数计算 $\overline{H}$ 所得结果中的最小值就是基态能 E_0 :

$$E_0 = \min(\overline{H}), \tag{4.2}$$

和这一最小值对应的波函数 ψ 就是基态波函数 φ_0 .

为了求第一激发态,对波函数 ψ 加上与 φ_0 正交的条件

$$\int \varphi_0^* \psi \mathrm{d}\tau = 0. \tag{4.3}$$

第一激发态能量 E_1 是一切符合条件(4.3)式的波函数所算出的 $\overline{H}$ 中的最小值

$$\text{对一切满足(1.3)式的 } \psi, \quad E_1 = \min(\overline{H}). \tag{4.4}$$

和这一极小值对应的波函数 ψ 就是 φ_1 .

对于更高的激发态,可以类推.

证　令 $\varphi_n (n = 1, 2, \cdots)$ 为哈密顿的正交归一本征函数, $\hat{H}\varphi_n = E_n \varphi_n$. 将任意波函数 ψ 用 φ_n 展开

$$\psi = \sum_n c_n \varphi_n. \tag{4.5}$$

代入(4.1)式,得到

$$\overline{H} = \sum_n E_n |c_n|^2 / \sum_n |c_n|^2. \tag{4.6}$$

用 E_0 表示基态能,则对一切 $n \neq 0$,均有 $E_n > E_0$. 因而,在(4.6)式中将全部 E_n 换成 E_0 只会使分子中的和式减小,而有

$$\overline{H} = \frac{\sum_n E_n |c_n|^2}{\sum_n |c_n|^2} \geqslant \frac{\sum_n E_0 |c_n|^2}{\sum_n |c_n|^2} = E_0. \tag{4.7}$$

只有当 $\psi = \varphi_0$ 时,$c_n = \int \varphi_n^* \psi = \delta_{n0}$,上式才取等号. 这就证明了,$E_0$ 是用任意波函数 ψ 算出的 $\overline{H}$ 中的最小值,如(4.2)式,而且使 $\overline{H}$ 取这个值的 ψ 就是 φ_0.

如果用满足(4.3)式的 ψ 计算 $\overline{H}$,则 $c_0 = \int \varphi_0^* \psi = 0$,展开式(4.6)中不包含 E_0 项,其中的一切 E_n 以 E_1 为最小. 因而类似于(4.2)式得到(4.4)式.

【证毕】

(二)作为近似方法的变分法

在以上的讨论中,是用一切可能的波函数来求 $\overline{H}$ 的极值,这样原则上能得到精确的基态和激发态的能量和波函数. 这在实际上很难做到,但提示了一种近似方法——变分法.

用变分法求基态能量和波函数的步骤是:

首先,根据实际问题的物理分析,选择含有待定参量 $\alpha,\beta,\cdots$ 的尝试波函数 $\psi(r;\alpha,\beta,\cdots)$,然后计算积分

$$J(\alpha,\beta,\cdots) = \frac{\int \psi^*(r;\alpha,\beta,\cdots)\hat{H}\psi(r;\alpha,\beta,\cdots)\,d\tau}{\int \psi^*(r;\alpha,\beta,\cdots)\psi(r;\alpha,\beta,\cdots)\,d\tau} \tag{4.8}$$

的极值

$$\partial J/\partial\alpha = \partial J/\partial\beta = \cdots = 0, \tag{4.9}$$

得到使积分取最小值的参量 $\alpha_0,\beta_0,\cdots$. 用它们按(4.8)式计算的结果就是基态能量的近似值,即近似地有

$$E_0 \approx J(\alpha_0,\beta_0,\cdots), \tag{4.10}$$

$$\varphi_0(r) \approx \psi(r;\alpha_0,\beta_0,\cdots). \tag{4.11}$$

进一步,给尝试波函数加上条件

$$\int \varphi_0^*(r)\psi(r;\alpha,\beta,\cdots)\,d\tau = 0. \tag{4.12}$$

在这一条件下求(4.8)式的极小值,就得到第一激发态能量的近似值.

这一方法的近似性在于，在应用变分原理(4.2)式和(4.4)式时，不是在由任意波函数 ψ 组成的全希尔伯特空间中求 $\bar{H}$ 的极小值，而只在由尝试波函数 $\psi(r;\alpha,\beta,\cdots)$ 组成的部分希尔伯特空间中求 $\bar{H}$ 的极小值.

在使用变分法求解定态问题时，关键性的工作是选择尝试波函数. 注意到波函数的标准条件和系统所具有的对称性，会帮助我们有效地完成这一工作.

【例】 已知氢原子的哈密顿算符为

$$\hat{H}_0 = -[\hbar^2/(2m_e)]\nabla^2 - e^2/r, \tag{4.13}$$

用变分法确定氢原子的基态能量和波函数.

解 首先讨论怎样选择尝试波函数. 从场的中心对称性知道，在定态中，电子的角动量有确定值. 对于基态，角动量为零，因而尝试波函数只依赖于 r 而与角度无关. 另外，作为描述束缚态的波函数，还要求 $r\to\infty$ 时，它趋于零. 因而我们可以选择尝试波函数为以下形式：

$$\psi = Ae^{-\beta r}, \tag{4.14}$$

β 是待定参量，A 为归一化常数，可以求得 $A^2 = \beta^3/\pi$.

将(4.13)式和(4.14)式代入(4.8)式，有

$$J(\beta) = -\frac{2\beta^3\hbar^2}{m_e}\int_0^\infty e^{-\beta r}\nabla^2 e^{-\beta r}r^2 dr - 4\beta^3 e^2\int_0^\infty e^{-2\beta r}r dr. \tag{4.15}$$

为了计算右边的第一个积分，利用球坐标中的公式[附录(Ⅱ.50)式]

$$\nabla^2 e^{-\beta r} = \frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}e^{-\beta r}\right),$$

代入(4.15)式右边第一项，进行分部积分，得到

$$\int_0^\infty e^{-\beta r}\nabla^2 e^{-\beta r}r^2 dr = -\int_0^\infty\left(\frac{d}{dr}e^{-\beta r}\right)^2 r^2 dr = -\frac{1}{4\beta}. \tag{4.16}$$

而(4.15)式右边的第二个积分为

$$\int_0^\infty e^{-2\beta r}r dr = \frac{1}{(2\beta)^2}. \tag{4.17}$$

这样，(4.15)式成为

$$J(\beta) = \beta^2\hbar^2/(2m_e) - \beta e^2. \tag{4.18}$$

按照极值条件(4.9)式，可得相应于基态，参量 β 的值为 $\beta_0 = 1/a_0$，其中 $a_0 = \hbar^2/m_e e^2$. 将 β_0 代入(4.18)式和(4.14)式就得到氢原子的基态能量和波函数：

$$E_0 = J(\beta_0) = -m_e e^4/(2\hbar^2), \qquad \psi_0 = (\pi a_0^3)^{-1/2}e^{-r/a_0}. \tag{4.19}$$

在这一例子中，由于通过对称性分析，选择了正确形式的尝试波函数(4.14)式，所以用变分法得到的基态能量和波函数同精确解一致，见 §6.1.1(1.49)式和(1.55)式的第一式.

* §6.1.5 耦合振子系统 量子力学的正则变换

在§1.6.3中曾经看到,让系统的哈密顿函数的变量经受适当的正则变换,可以得到更便于处理物理问题的哈密顿函数形式.这一节讨论如何将这一方法用到量子力学中.

在量子力学中,借助于正则变换求系统定态能量的基本思想是:让系统的哈密顿经受一个适当的正则变换,使得变换后的哈密顿是对角化的,其对角元素就是所求的能量.在多数情况下,将系统哈密顿严格对角化是困难的,通常是把哈密顿简化,保留它的本质部分;再通过正则变换把简化后的哈密顿对角化,得到定态能量的近似值.

在经典力学中,正则变换是在保持哈密顿方程形式不变条件下的变量变换.由§5.3.2知道,量子力学中哈密顿形式的运动方程是用量子泊松括号表示,即用哈密顿和坐标及动量的对易关系表示.考虑到幺正变换不改变算符间的对易关系,可见在量子力学中,正则变换是正则变量的幺正变换.

作为例子,讨论由彼此间存在相互作用的两个谐振子构成的系统.在§1.4.3中用经典力学研究过这种系统,证明了可以通过正则变换转到简正坐标,成为两个独立的谐振子.下面用量子力学讨论这同一问题.为了物理图像清晰起见,选用占有数表象.由§5.4.5知道,在占有数表象中,谐振子的哈密顿算符在扣除了零点能后可以写成

$$\hat{H} = \hbar\omega\hat{A}^{\dagger}\hat{A}. \tag{5.1}$$

$\hat{A}^{\dagger}$ 和 $\hat{A}$ 分别是粒子的产生和湮灭算符,它们之间的对易关系是

$$[\hat{A},\hat{A}^{\dagger}] = 1. \tag{5.2}$$

在占有数表象,正则变换就是从一种产生算符和湮灭算符到另一种产生算符和湮灭算符间的幺正变换,在这种变换下,产生算符和湮灭算符之间的对易关系的形式保持不变.

设系统的哈密顿算符为:

$$\hat{H} = \sum_{\alpha=1}^{2}\varepsilon_{\alpha}^{(0)}\hat{B}_{\alpha}^{\dagger}\hat{B}_{\alpha} + \sum_{\alpha\neq\beta}^{2}\varepsilon_{\alpha\beta}\hat{B}_{\alpha}^{\dagger}\hat{B}_{\beta}, \tag{5.3}$$

式中第一项是两个独立谐振子的哈密顿算符之和,第二项表示它们之间的相互作用势能算符,参数 $\varepsilon_{\alpha}^{(0)}$,$\varepsilon_{\alpha\beta}$ 是实数.产生算符和湮灭算符之间的对易关系是:

$$[\hat{B}_{\alpha},\hat{B}_{\beta}^{\dagger}] = \delta_{\alpha\beta}, \qquad [\hat{B}_{\alpha}^{\dagger},\hat{B}_{\beta}^{\dagger}] = [\hat{B}_{\alpha},\hat{B}_{\beta}] = 0. \tag{5.4}$$

我们用正则变换的方法将 $\hat{H}$ 对角化.设算符 $\hat{B}_{\alpha}$ 通过以下的正则变换而和新算符 $\hat{A}_{\mu}$ 相联系:

$$\hat{B}_1 = u_{11}\hat{A}_1 + u_{12}\hat{A}_2, \qquad \hat{B}_2 = u_{21}\hat{A}_1 + u_{22}\hat{A}_2. \tag{5.5}$$

变换矩阵 u 的幺正性条件

$$\sum_{\alpha}u_{\alpha\mu}u_{\alpha\nu}^{*} = \delta_{\mu\nu}, \qquad \sum_{\mu}u_{\alpha\mu}u_{\beta\mu}^{*} = \delta_{\alpha\beta}, \tag{5.6}$$

使得新算符仍有以下形式的对易关系

$$[\hat{A}_{\mu},\hat{A}_{\nu}^{\dagger}] = \delta_{\mu\nu}, \qquad [\hat{A}_{\mu}^{\dagger},\hat{A}_{\nu}^{\dagger}] = [\hat{A}_{\mu},\hat{A}_{\nu}] = 0. \tag{5.7}$$

将(5.5)式代入(5.3)式,写 $\varepsilon_{\alpha}^{(0)} = \varepsilon_{\alpha\alpha}$ 得变换后的哈密顿算符为:

$$\hat{H} = \sum_{\alpha=1}^{2} \varepsilon_{\alpha\alpha} \hat{B}_\alpha^\dagger \hat{B}_\alpha + \sum_{\alpha \neq \beta}^{2} \varepsilon_{\alpha\beta} \hat{B}_\alpha^\dagger \hat{B}_\beta = \sum_{\alpha,\beta} \varepsilon_{\alpha\beta} \hat{B}_\alpha^\dagger \hat{B}_\beta$$

$$= \sum_{\alpha,\beta} \sum_{\mu,\nu} \varepsilon_{\alpha\beta} u_{\alpha\nu}^* u_{\beta\mu} \hat{A}_\nu^\dagger \hat{A}_\mu , \tag{5.8}$$

注意到在粒子数表象中 $\hat{H}$ 的对角形式应是

$$\hat{H} = \sum_{\mu=1}^{2} E_\mu \hat{A}_\mu^\dagger \hat{A}_\mu , \tag{5.9}$$

可见，为了使 $\hat{H}$ 对角化，必须有

$$\sum_{\alpha,\beta} \sum_{\mu,\nu} \varepsilon_{\alpha\beta} u_{\alpha\nu}^* u_{\beta\mu} \hat{A}_\nu^\dagger \hat{A}_\mu = \sum_{\mu=1}^{2} E_\mu \hat{A}_\mu^\dagger \hat{A}_\mu , \tag{5.10}$$

由此可以得到幺正矩阵 u 所满足的方程. 事实上为使上面的等式成立，必须

$$\sum_{\alpha,\beta} \varepsilon_{\alpha\beta} u_{\alpha\nu}^* u_{\beta\mu} = E_\mu \delta_{\mu\nu} , \tag{5.11}$$

将上式左边的哑标 α 改写为 ρ，再用 $u_{\alpha\nu}$ 乘左右两边，对 ν 求和，利用幺正条件(5.6)式得到

$$\sum_{\beta} (\varepsilon_{\alpha\beta} - E_\mu \delta_{\alpha\beta}) u_{\beta\mu} = 0. \tag{5.12}$$

这是关于矩阵元 $u_{\beta\mu}$ 的齐次方程组，它有非零解的条件是

$$| \varepsilon_{\alpha\beta} - E_\mu \delta_{\alpha\beta} | = 0. \tag{5.13}$$

由这个久期方程可得 E 的两个根 $E_\mu (\mu = 1,2)$，

$$E_\mu = (\varepsilon_{11} + \varepsilon_{22} \pm \sqrt{(\varepsilon_{11} - \varepsilon_{22})^2 - 4\varepsilon_{12}\varepsilon_{21}}) / 2 . \tag{5.14}$$

对每一个根 E_μ，可以从方程(5.12)式和幺正条件(5.6)式一道决定相应的幺正矩阵 u 的矩阵元 $u_{\beta\mu}$，进而得到新的湮灭算符和产生算符 $\hat{A}_\mu$、$\hat{A}_\mu^\dagger$.

值得注意的是，由 $\hat{A}_\mu$、$\hat{A}_\mu^\dagger$ 描述的具有能量 E_μ 的谐振子不是物理真实的谐振子，真实的谐振子是哈密顿算符原始形式所描述的彼此有相互作用的谐振子. 因而通常把经过正则变换得到的谐振子称为赝谐振子. 注意到新的产生算符和湮灭算符是原来产生算符和湮灭算符的线性组合，这就表明，每一个赝谐振子描述的都不是哪一个真实谐振子的个别行为，而是描述有相互作用的两个谐振子的集体行为. 这和经典力学中简正坐标描述相互作用谐振子的集体行为相对应，参看 §1.4.3.

这种把有相互作用的多粒子系统，通过正则变换变成一个没有相互作用的赝粒子(准粒子)系统的方法，在物理理论中是屡见不鲜的，例如见第八篇第二章.

习　题

1. 两个耦合谐振子的哈密顿为：

$$\hat{H}_0 = -(\hat{p}_1^2 + \hat{p}_2^2)/(2m) + (m\omega_0^2/2)(x_1^2 + x_2^2) - \lambda x_1 x_2 ,$$

试变换到简正坐标(见 §1.4.3)化为两个独立谐振子后，求耦合振子的能级.

2. 单价原子中价电子在原子实的电场中的势能为：

$$U(r) = -e^2/r - \lambda e^2 a_0 / r^2 , \qquad 0 < \lambda \ll 1$$

a_0 为第一玻尔轨道半径，求价电子的能级.

3. 求氢原子中的电子处于 $\psi_{n,2,0}$、$\psi_{n,2,\pm1}$、$\psi_{n,2,\pm2}$ 状态中的概率流密度及电子云绕 z 轴的角动量.

4. 分别在直角坐标和球坐标中分离变量，求三维各向同性谐振子的能级.（对球坐标中的径向方程，用二阶变系数常微分方程的级数解法求解.）

5. 试用变量替换和函数替换方法，把三维各向同性谐振子的径向方程化为氢原子的径向方程，以求解该谐振子的能级.

提示　令 $r^2 = \rho, u(r) = \rho^{-1/4} v(\rho)$.

6. 一个质量为 m，电荷为 q 的粒子处于势能场 $V(x) = k^2 x^2/2$ 中，而且又受到平行于 x 轴的电场 ε 的作用.（a）直接由定态方程求粒子的能级.（b）将电场看作微扰，求粒子的能级.

7. 哈密顿在 H_0 表象中的矩阵为

$$\begin{pmatrix} E_1^{(0)} + a & b \\ b & E_2^{(0)} + a \end{pmatrix}$$

其中 a, b 为实数.（a）直接由定态方程求能量.（b）用微扰方法求能量至二级近似.

8. 在粒子数表象中哈密顿算符为：$\hat{H} = \alpha \hat{A}^{\dagger} \hat{A} + \lambda\alpha(\hat{A}^{\dagger} + \hat{A})$.（a）把 $\hat{H}$ 化成标准型的谐振子哈密顿算符，求能量本征值.（b）把 $\lambda\alpha(\hat{A}^{\dagger} + \hat{A})$ 看成微扰，求能量.

9. 质量为 μ，带电 $+Q$ 和 $-Q$ 的两个小球由长为 a 的刚性轻质杆联结，绕通过杆中点、垂直于杆的轴转动. 令转轴为 z 轴. 在 x 方向有均匀电场 E，z 方向有均匀磁场 B. 用微扰方法求转子的能谱.

10. 推导未扰态无简并情况下态矢的二级修正和能量的三级修正.

11. 设系统的哈密顿在未扰态能量表象中的矩阵为

$$\begin{pmatrix} E_1^{(0)} & 0 & \lambda_a \\ 0 & E_1^{(0)} & \lambda_b \\ \lambda_a^* & \lambda_b^* & E_2^{(0)} \end{pmatrix}, \qquad E_1^{(0)} > E_2^{(0)}, \qquad \lambda_a, \lambda_b \ll 1.$$

（a）用微扰方法求能量到二级修正；（b）求能量的严格解，并和（a）比较.

12. 质量为 m 的粒子在势能场 $V(r) = -V_0 e^{-r/2a}, V_0 > 0$ 中运动.（a）以 $\alpha e^{\beta r/a}$ 作尝试波函数，用变分法求基态能；（b）通过变量变换，令 $\rho = e^{-r/2a}$，解定态方程确定基态能量.

13. 用 $(a + bx)e^{-\beta x^2}$ 作尝试波函数，求谐振子的基态和第一激发态的能量和波函数.

第二章　量子跃迁

系统在外场作用下，从一个定态跃迁到另一个定态是一类常见的实际问题. 原子的受激跃迁就是这类问题的代表.

跃迁问题不是定态问题，它要讨论的是根据已知的系统的初始状态，来确定受到与时间有关的外场作用后，系统将变化到怎样的状态. 由于外场是与时间有关的，因而这类问题需要通过求解含时间的薛定谔方程来讨论. 在这一章，我们首先讨论求解含时薛定谔方程的微扰方法，然后再讨论在一些特殊类型的外场作用下引起的跃迁，并应用所得结果讨论光的发射与吸收以及弹性散射问题.

§6.2.1　研究时间演化的微扰方法　相互作用绘景

(一)相互作用绘景中的运动方程

迄今为止，我们采用的都是薛定谔绘景. 在这一绘景中，系统的时间演化由薛定谔方程描述，见§5.3.1 和§5.4.7. 若已知系统在初始时刻的状态波函数，则通过薛定谔方程可以决定系统在以后时刻的状态. 这是薛定谔方程的初值问题.

在§5.3.1 中讨论过系统哈密顿 $\hat{H}$ 不显含时间情况下的初值问题. 这时薛定谔方程的一般解为

$$\sum_m a_m \varphi_m(\boldsymbol{r}) \mathrm{e}^{-(\mathrm{i}/\hbar)E_m t}. \tag{1.1}$$

其中 φ_m 是 $\hat{H}$ 的定态解，它是包括 $\hat{H}$ 在内的一组力学量完全组的共同本征态. 而 a_m 由系统的初始状态波函数决定[见§5.3.1(1.24)式].

如果系统的哈密顿显含时间，则薛定谔方程为

$$\mathrm{i}\hbar\partial\psi(\boldsymbol{r},t)/\partial t = \hat{H}(t)\psi(\boldsymbol{r},t). \tag{1.2}$$

此时不能再象§5.3.1 那样用分离变量法解薛定谔方程得到如(1.1)式的解. 但如果有

$$\hat{H} = \hat{H}_0 + \hat{V}(t), \tag{1.3}$$

其中 $\hat{H}_0$ 不显含时间，存在定态解：

$$\hat{H}_0\varphi_m = E_m\varphi_m, \tag{1.4}$$

则仍然可以把解写成(1.1)式的形式，只不过其中的系数 a_m 是时间 t 的函数：

$$\psi(\boldsymbol{r},t) = \sum_m a_m(t)\varphi_m(\boldsymbol{r})\mathrm{e}^{-(\mathrm{i}/\hbar)E_m t}, \quad a_m(t) = \int \varphi_m^*(\boldsymbol{r})\psi(\boldsymbol{r},t)\mathrm{e}^{(\mathrm{i}/\hbar)E_m t}\mathrm{d}\tau. \tag{1.5}$$

这是解微分方程的一种常用的方法，称为变动常数法. 将(1.5)式代入(1.2)式，可以得到决定 $a_m(t)$ 的方程：

$$\mathrm{i}\hbar\sum_m \dot{a}_m\varphi_m \mathrm{e}^{-(\mathrm{i}/\hbar)E_m t} + \sum_m a_m E_m \varphi_m \mathrm{e}^{-(\mathrm{i}/\hbar)E_m t} = \sum_m a_m \hat{H}_0\varphi_m \mathrm{e}^{-(\mathrm{i}/\hbar)E_m t} + \sum_m a_m \hat{V}\varphi_m \mathrm{e}^{-(\mathrm{i}/\hbar)E_m t}. \tag{1.6}$$

由于 φ_m 是 $\hat{H}_0$ 的本征态，所以左边第二项和右边第一项消去. 将剩下的两个和式中的哑标 m 改为 k，乘上 φ_m^*，积分得到 $\mathrm{i}\hbar\dot{a}_m \mathrm{e}^{-(\mathrm{i}/\hbar)E_m t} = \sum_k a_k V_{mk}\mathrm{e}^{-(\mathrm{i}/\hbar)E_k t}$，其中

$$V_{mk} = \int \varphi_m^* \hat{V}\varphi_k \mathrm{d}\tau. \tag{1.7}$$

是算符 $\hat{V}$ 在 H_0 表象中的矩阵元. 令

$$\omega_{mk} = (E_m - E_k)/\hbar, \tag{1.8}$$

得到 $a_m(t)$ 的方程

$$\mathrm{i}\hbar\dot{a}_m(t) = \sum_k V_{mk}a_k(t)\mathrm{e}^{\mathrm{i}\omega_{mk}t}. \tag{1.9}$$

(1.9)式和薛定谔方程(1.2)式等效，是描述系统时间演化的方程. 值得特别注意的是，(1.9)式右边出现的是微扰算符 V，而不象(1.2)式那样是总哈密顿 H. 其原因是，在推导(1.9)式的过程中进行了一个绘景变换.

在 §5.4.7 中曾经指出，量子力学可以有不同的绘景. 不同绘景描述时间演化的方式不同. 例如，在薛定谔绘景中，算符不随时间变，而态矢随时间的变化由总哈密顿 H 决定；在海森伯绘景中，态矢不随时间变，而算符随时间的变化由总哈密顿 H 决定. 现在，$a_m(t)$ 是（H_0 表象中的）态矢，(1.9)式表明它随时间变化，但其变化规律不是由总哈密顿 H，而是由相互作用哈密顿 V 决定. 因此，(1.9)式是介于薛定谔绘景和海森伯绘景之间的一种新的绘景. 为了更清楚地看到在(1.9)式的推导过程中包含有绘景的变化，将(1.5)的第二式改写为

$$a_m(t) = \int \varphi_m^*(\boldsymbol{r})\mathrm{e}^{(\mathrm{i}/\hbar)E_m t}\psi(\boldsymbol{r},t)\mathrm{d}\tau = \int \langle m \mid \boldsymbol{r}\rangle \mathrm{e}^{(\mathrm{i}/\hbar)E_m t}\langle \boldsymbol{r} \mid t\rangle^{\mathrm{S}}\mathrm{d}^3\boldsymbol{r}$$

$$= \int \langle m \mid \mathrm{e}^{(\mathrm{i}/\hbar)H_0 t} \mid \boldsymbol{r}\rangle\langle \boldsymbol{r} \mid t\rangle^{\mathrm{S}}\mathrm{d}^3\boldsymbol{r} = \langle m \mid \mathrm{e}^{(\mathrm{i}/\hbar)H_0 t} \mid t\rangle^{\mathrm{S}}. \tag{1.10}$$

这里利用了厄米算符既可以向右也可以向左作用到态矢上[见 §6.1.2(2.16)式]的法则.

写 $a_m(t) = \langle m \mid t\rangle^{\mathrm{I}}$，则(1.10)式成为 $\langle m \mid t\rangle^{\mathrm{I}} = \langle m \mid \mathrm{e}^{(\mathrm{i}/\hbar)H_0 t} \mid t\rangle^{\mathrm{S}}$. 去掉左矢 $\langle m \mid$ 以后，可以明显看到，算符 $\mathrm{e}^{(\mathrm{i}/\hbar)H_0 t}$ 将态矢从薛定谔绘景变换到了一个新的绘景. 将这一算符和由薛定谔绘景到海森伯绘景的变换[见 §5.4.7(7.15)(7.16)式]

$$\mid\rangle^{\mathrm{H}} = \mathrm{e}^{(\mathrm{i}/\hbar)Ht} \mid t\rangle^{\mathrm{S}}, \tag{1.11a}$$

$$F^H = e^{(i/\hbar)Ht}F^S e^{-(i/\hbar)Ht}, \tag{1.11b}$$

所用的算符 $e^{(i/\hbar)Ht}$ 比较，可见两者形式上一致，只是用 H_0 代替了 H．由它实现的绘景称为相互作用绘景，用 I 表示

$$|t\rangle^I = e^{(i/\hbar)H_0t}|t\rangle^S, \tag{1.12a}$$

$$F^{(I)} = e^{(i/\hbar)H_0t}F^{(S)}e^{-(i/\hbar)H_0t}. \tag{1.12b}$$

三种绘景对时间演化的描述方式归纳于表 I．

表 I　三种绘景对时间演化的描述方式

	薛定谔绘景	海森伯绘景	相互作用绘景
态　矢	随 t 变 规律由 $\hat{H}$ 决定	不随 t 变	随 t 变 规律由 $\hat{V}$ 决定
算　符	不随 t 变	随 t 变 规律由 $\hat{H}$ 决定	随 t 变 规律由 $\hat{H}_0$ 决定
本征矢	同上	同上	同上

三种绘景在物理观测结果上等效，而在应用上各有方便之处．相互作用绘景特别对处理跃迁和散射问题比较方便，原因是在这一绘景中，态矢随 t 变化的规律仅由相互作用 V 决定，见(1.9)式．这一特点对于用微扰方法解含时方程十分有利．

(二)含时微扰

设在 $t<0$ 时，系统的哈密顿 $\hat{H}_0$ 不显含时间，有定态解．在 $t=0$ 时刻，系统处于某一定态 φ_n，它是包括 $\hat{H}_0$ 在内的力学量完全组 F 的一个共同本征态．从 $t=0$ 开始，加入外场．外场 V 一般与时间有关．这时，原力学量完全组不再是守恒量，系统因而不再处于原来的状态．状态的变化由运动方程(1.9)式决定．初始条件为

$$a_k(0) = \delta_{kn}. \tag{1.13}$$

它表示在 $t=0$ 时，系统处于状态 φ_n．

微扰方法的基本思想是将相互作用 V 看成微扰(写成 λV)，并将解展开为

$$a_m(t) = a_m^{(0)}(t) + \lambda a_m^{(1)}(t) + \lambda^2 a_m^{(2)}(t) + \cdots. \tag{1.14}$$

代回方程(1.9)，并令方程两边 λ 的同次幂系数相等，得到

$$\dot{a}_m^{(0)} = 0, \tag{1.15}$$

$$i\hbar\dot{a}_m^{(1)}(t) = \sum_k a_k^{(0)}(t)V_{mk}e^{-i\omega_{mk}t}, \tag{1.16}$$

$$i\hbar\dot{a}_m^{(2)}(t) = \sum_k a_k^{(1)}(t)V_{mk}e^{-i\omega_{mk}t}, \tag{1.17}$$

…

从上面这些方程可以看到，只要知道了初始时刻的零级波函数，就可以用迭代法逐级求出各级修正.

由初始条件(1.13)式，按(1.15)式可得

$$a_k^{(0)}(t) = \delta_{kn},$$

再把它代入(1.16)式，得到一级修正所满足的方程

$$i\hbar\dot{a}_m^{(1)}(t) = V_{mn}e^{i\omega_{mn}t}. \tag{1.18}$$

积分得到

$$a_m^{(1)}(t) = \frac{1}{i\hbar}\int_0^t V_{mn}e^{i\omega_{mn}t}dt, \qquad m \neq n. \tag{1.19}$$

(三)恒定微扰的加入过程

现在来应用以上得到的含时微扰方法，讨论在系统中加入恒定微扰的问题. 假定微扰是在 $t = 0$ 时刻加入，加入以后保持为常数，不再随 t 变. 对(1.19)式进行分部积分，得到

$$a_m^{(1)}(t) = \frac{V_{mn}e^{i\omega_{mn}t}}{E_n - E_m} + \int_0^t \frac{e^{i\omega_{mn}t}}{\hbar\omega_{mn}}\frac{dV_{mn}}{dt}dt, \qquad m \neq n. \tag{1.20}$$

我们从上式出发分析两种情况：

1. 缓慢加入的微扰

如果微扰是很缓慢地加入的，则称为绝热加入. 此时，(1.20)式第二项中的 dV_{mn}/dt 很小，可以略去. 将剩下的第一项和 §6.1.2(2.22)式比较可见，它代表微扰加入以后引起的定态的变化. 比那里多出时间因子 $e^{i\omega_{mn}t}$ 是由于这里 a_m 的定义(1.5)式中有时间因子.

这一结果说明，如果微扰是绝热变化的，则它不会导致系统发生从一个定态到另一定态的跃迁，而只是按 §6.1.2 中的定态微扰论引起定态的微小变化. 通常实验室中加入外磁场、外电场的过程，所经历的时间有秒的量级，比原子内部过程的特征时间 $\omega_{mn}^{-1} \sim 10^{-15}$ s 长得多，满足绝热加入的条件，因而加入过程不引起跃迁，而只是使能级移动.

2. 微扰突然加入时引起的跃迁

考虑另一种极端情况，假定微扰是突然加入的. 在(1.20)式中，对积分的贡献来自微扰突然加入的那个瞬间. 这个瞬间是极短暂的，它远小于系统从 E_n 态跃迁到 E_m 态所需的时间 $1/\omega_{mn}$，即：相对而言，在对积分有贡献的这个极短的时间内，$\omega_{mn}t$ 是很小的. 因而在(1.20)式中的因子 $e^{i\omega_{mn}t}$ 可以从积分号内移出.

这样，(1.20)式成为

$$a_m^{(1)}(t) = \frac{V_{mn}e^{i\omega_{mn}t}}{E_n - E_m} + \frac{e^{i\omega_{mn}t}}{E_m - E_n}\int_0^t \frac{dV_{mn}}{dt}dt = 0, \qquad m \neq n.$$

这表明，加入微扰以后，系统仍然停留在 φ_n 态，没有变化. 其原因是，微扰加入太快，在很短的时间 τ 内就完成了由 H_0 到 H 的变化. 这一时间远小于系统状态变化的特征时间：$\tau \ll 1/\omega_{mn}$，因而系统的状态来不及改变. 但是，在哈密顿由 H_0 变为 H 以后，$\varphi_n(r) = \langle r \mid n^0 \rangle$ 不再是定态，因而应将它用新哈密顿 H 的本征态（新的定态）展开. 在一级近似下，H 的本征态是 §6.1.2(2.22)式

$$\mid m\rangle = \mid m^0\rangle + \sum_k{}' \frac{V_{km}}{E_m^{(0)} - E_k^{(0)}} \mid k^0\rangle. \tag{1.21}$$

态矢 $\mid n^0\rangle$ 用 $\mid m\rangle$ 展开的系数 $\langle m \mid n^0\rangle$ 的绝对值平方，就是在 $\mid n^0\rangle$ 态中发现 $\mid m\rangle$ 态的概率

$$P_{mn} = \mid V_{mn} \mid^2 / (\hbar^2 \omega_{mn}^2). \tag{1.22}$$

这就是突然加入的微扰使系统由（H_0 的）n 态跃迁到（H 的）m 态的概率.

§6.2.2　周期性外场引起的跃迁

在很多实际问题中，微扰是以振荡方式随时间变化的. 现在我们来讨论这类微扰引起的跃迁. 这是下一节要讨论的光的发射与吸收的基础.

（一）单色周期性的微扰

对于单色周期性的微扰，例如单色光波，其算符一般地可以写成

$$\hat{V}(t) = \hat{A}\cos \omega t = \hat{F}(e^{i\omega t} + e^{-i\omega t}), \tag{2.1}$$

其中 $\hat{F}$ 是与时间无关的厄米算符. 按 §6.2.1(1.19)式，在此微扰作用下，系统由初态，即 $\hat{H}_0$ 的定态 φ_n 跃迁到 $\hat{H}_0$ 的另一个定态 φ_m 的跃迁概率振幅为

$$\begin{aligned} a_m(t) &= \frac{F_{mn}}{i\hbar}\int_0^t [e^{i(\omega_{mn}+\omega)t'} + e^{i(\omega_{mn}-\omega)t'}]\,dt' \\ &= -\frac{F_{mn}}{\hbar}\left[\frac{e^{i(\omega_{mn}+\omega)t} - 1}{\omega_{mn} + \omega} + \frac{e^{i(\omega_{mn}-\omega)t} - 1}{\omega_{mn} - \omega}\right]. \end{aligned} \tag{2.2}$$

从(2.2)式可以看出，跃迁概率振幅依赖于外场频率 ω 和系统的玻尔频率 ω_{mn} 之间的关系. 当 $\omega \neq \pm \omega_{mn}$ 时，它不随时间增长，而只是发生振荡性的变化. 当 $\omega = \omega_{mn}$ 时，第二项的分子分母都为零. 利用洛比达法则可以得知该项与 t 成正比，即随时间线性地增大，因而相对于它，第一项可以略去不计. 类似地，当 $\omega = -\omega_{mn}$ 时，第二项可以略去不计. 这样我们看到，在一般情况下，当 $\omega \neq \pm \omega_{mn}$ 时，外场只微弱地影响到系统的状态，不论外场作用时间多长，跃迁概率振幅都不会明显地增大，因而系统基本停留在原始的状态. 但是，当 $\omega = \pm \omega_{mn}$ 时，跃迁概率振幅随时间而增大，只要外场作用时间够长，系统一定会有足够大的概率从初态 φ_n 跃迁到状态 φ_m. 这类似于经典力学中的受迫振动：当外力频率不等于振于固有频率时，振幅始终不

大.但当外力频率等于振子的固有频率时,振幅会变得很大.因而通常把跃迁中的这种现象称为共振跃迁.

下面计算 $\omega = \pm \omega_{mn}$ 时的跃迁概率.

根据上面的分析,当 $\omega = \omega_{mn}$ 时,在(2.2)式中只有第二项起主要作用,因面这时的跃迁概率为

$$P_{nm}(t) = |a_m(t)|^2 = \frac{2}{\hbar^2}|F_{mn}|^2 \frac{[1-\cos(\omega_{mn}-\omega)t]}{(\omega_{mn}-\omega)^2}$$

$$= \frac{4}{\hbar^2}|F_{mn}|^2 \frac{\sin^2[(\omega_{mn}-\omega)t/2]}{(\omega_{mn}-\omega)^2}. \tag{2.3}$$

在附录(Ⅰ.16)式中证明了

$$\sin^2 xu/x^2 \underset{u\to\infty}{\longrightarrow} \pi u\delta(x). \tag{2.4}$$

令 $u = t/2$ 就看出,只要外场作用时间够长,(2.3)式可以写成

$$P_{nm}(t) = (2\pi t/\hbar^2)|F_{mn}|^2\delta(\omega_{mn}-\omega). \tag{2.5}$$

这就是当 $\omega = \omega_{mn}$ 时的跃迁概率.

现在来分析一下这种跃迁是怎样的物理过程.由于外场频率 ω 总是正的,$\omega = \omega_{mn}$ 意味着 $E_m - E_n > 0$,即末态能量大于初态能量.因而(2.5)式描述的是在外场作用下,系统吸收能量 $\hbar\omega_{mn} = E_m - E_n$ 而从低能态跃迁到高能态的概率.

类似地,当 $\omega = -\omega_{mn}$ 时,在(2.2)式中只有第一项才有贡献.重复上面类似的讨论可得这时的跃迁概率为:

$$P_{nm}(t) = (2\pi t/\hbar^2)|F_{mn}|^2\delta(\omega_{mn}+\omega). \tag{2.6}$$

由 $\hbar\omega = -\hbar\omega_{mn} = E_n - E_m > 0$ 知道,(2.6)式描述的是系统在外场作用下,从高能态跃迁到低能态的跃迁概率.在这个过程中系统会辐射能量 $\hbar\omega_{mn}$.

由(2.5)式和(2.6)式可见,在同一外场作用下,系统吸收能量从低能态 l 跃迁到高能态 k 的概率,和辐射能量从高能态 k 跃迁到低能态 l 的概率相等.这表明了微观过程的可逆性.

在(2.5)式和(2.6)式中,跃迁概率都与微扰作用的时间成正比.有实际意义的是单位时间的跃迁概率.从(2.5)式和(2.6)式得到相应的单位时间由 n 态到 m 态的跃迁概率为:

$$W_{nm} = \mathrm{d}P_{nm}(t)/\mathrm{d}t = (2\pi/\hbar^2)|F_{mn}|^2\delta(\omega_{mn}\pm\omega). \tag{2.7}$$

或者写成

$$W_{nm} = (2\pi/\hbar)|F_{mn}|^2\delta(E_m - E_n \pm \hbar\omega). \tag{2.8}$$

对于微扰作用下的辐射过程,上面两式的 δ 函数中取“+”号;对于吸收过程取“-”号.

(2.7)式和(2.8)式中都含有 δ 函数,我们来讨论它的物理意义.在(2.7)式中含有 $\delta(\omega_{mn} \pm \omega)$ 表明,只有当 $\omega = \pm \omega_{mn}$ 时,W_{nm} 才不为零,这是量子跃迁共振特性的反映;而在(2.8)式中所含的 $\delta(E_m - E_n \pm \hbar\omega)$,则反映了在跃迁过程中能量守恒.另一方面,我们又知道,当 $x = x'$ 时,$\delta(x - x') = \infty$,$\int f(x) \cdot \delta(x - x')\mathrm{d}x = f(x')$.按照 δ 函数的这个意义,我们可以得到结论:(2.7)式和(2.8)式只有出现在积分号内,而且被积函数含有 ω 或 E_m 时,才有实际意义.这正是我们下面要进一步讨论的两种情况.

(二)非严格单色的微扰

实际上,对系统发生微扰作用的不可能是严格的单色周期场,而必定是若干频率相近的单色场的叠加,它所导致的跃迁概率应是各种单色场导致的跃迁概率之和.在这种实际的微扰中,不同频率的单色场的振幅不同,这就是说,(2.1)式中的 $\hat{F}$ 是依赖于 ω 的,因而在(2.7)式中 F_{mn} 和频率有关,应写成 $F_{mn}(\omega)$.这样,分布在 $\Delta\omega$ 内的各种单色场引起的单位时间跃迁概率为:

$$W_{nm} = \frac{2\pi}{\hbar^2}\int_{\Delta\omega} |F_{mn}(\omega)|^2 \delta(\omega_{mn} \pm \omega)\mathrm{d}\omega. \tag{2.9}$$

如果在 $\Delta\omega$ 内包含有 $\omega = |\omega_{mn}|$ 的场,则上式给出

$$W_{nm} = (2\pi/\hbar^2)|F_{mn}(\omega_{mn})|^2; \tag{2.10}$$

如果在 $\Delta\omega$ 内没有包含 $\omega = |\omega_{mn}|$ 的场,则 $W_{nm} = 0$.

从(2.10)式可以看出,由非单色性微扰引起的从 n 到 m 态的跃迁概率,是由微扰中频率为 $|\omega_{mn}|$ 的那一部分单色频率的场决定的,这也正是量子跃迁的共振性质.

(三)末态为连续谱的跃迁

在周期性微扰的作用下,除了会引起分立能态之间的跃迁外,还会引起从某一初态到连续谱中的末态之间的跃迁.例如,原子中发生的光电效应就是在入射光的作用下,电子从原子中的束缚态到电离后的自由状态的跃迁.

以上得到的(2.8)式是分离能级之间的跃迁概率.为了用它讨论末态为连续谱的问题,将系统放在一个长宽高均为 L 的一个大立方盒子中,使能谱分立.然后再让 $L \to \infty$.为了使能谱分立,可以在盒子边界上加势能无穷的条件,如一维情况下的 §5.2.2 例 4.但是,更方便的是用周期性条件,如 §5.1.2(2.19)式:

$$\psi(x + L) = \psi(x), \quad \psi(y + L) = \psi(y), \quad \psi(z + L) = \psi(z). \tag{2.11}$$

在这样的大盒子中的运动是"准自由运动" $Ae^{\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}$ 其中的动量 $\boldsymbol{p}$ 取分立值,如 §5.1.2(2.20)式:

$$p_x = 2\pi\hbar n_x/L,\quad p_y = 2\pi\hbar n_y/L,\quad p_z = 2\pi\hbar n_z/L,$$
$$n_x, n_y, n_z = 0, 1, 2, \cdots. \tag{2.12}$$

用 ν 表示标志这种"准自由运动"状态的量子数[例如(2.12)式中的 n_x, n_y, n_z]. 我们关心的是在外场作用下跃迁到状态 ν 附近的能量间隔 dE_ν 中的概率. 用 $\rho(E_\nu)$ 表示状态 ν 附近的态密度,即单位能量间隔中的状态数,则末态能量在 E_ν 到 $E_\nu + dE_\nu$ 范围内的状态数为 $\rho(E_\nu)dE_\nu$. 系统在周期性微扰的作用下从初态 φ_n 单位时间跃迁到这些末态的概率为:

$$dW_{n\nu} = (2\pi/\hbar)\,|\,F_{\nu n}\,|^2\delta(E_\nu - E_n \pm \hbar\omega)\rho(E_\nu)dE_\nu; \tag{2.13}$$

跃迁到连续谱中任一个末态的概率则为

$$W_{n\nu} = \int(2\pi/\hbar)\,|\,F_{\nu n}\,|^2\delta(E_\nu - E_n \pm \hbar\omega)\rho(E_\nu)dE_\nu. \tag{2.14}$$

积分对所有可能的末态能量 E_ν 进行.

【例 1】 求长宽高均为 L 的一个大立方盒子中"准自由运动"状态的态密度[用周期性边条件(2.12)式].

解 由(2.12)式,在动量值 $p_x \sim p_x + \Delta p_x, p_y \sim p_y + \Delta p_y, p_z \sim p_z + \Delta p_z$ 之间的状态数是

$$\Delta N = [L/(2\pi\hbar)]^3\Delta p_x\Delta p_y\Delta p_z. \tag{2.15}$$

利用空间的各向同性,采用球坐标

$$\Delta N = [L/(2\pi\hbar)]^3 4\pi p^2\Delta p. \tag{2.16}$$

由 $p^2 = 2mE$ 有 $p\Delta p = m\Delta E$,代入(2.16)式得

$$\rho(E) = \frac{\Delta N}{\Delta E} = \frac{m^{3/2}\sqrt{E}}{\sqrt{2}\pi^2}\left(\frac{L}{\hbar}\right)^3. \tag{2.17}$$

§6.2.3 光的辐射和吸收

上节一般地讨论了在周期性微扰作用下导致的系统定态间的跃迁. 现在应用上节的结果,研究在入射光的作用下原子定态间的跃迁. 在入射光的作用下,处子低能态的原子可以通过吸收光而跃迁到相应的高能态,叫做光的吸收. 与此同时,在入射光的作用下,处子高能态的原子也会通过辐射光而跃迁到低能态,叫做原子的受激辐射. 光的吸收和受激辐射都是原子在外场作用下发生的跃迁,称为感应跃迁. 除了感应跃迁外,原子也会自发地从高能态跃迁到低能态而辐射,称为自发辐射. 这一节主要讨论感应跃迁,最后再对自发辐射作些说明.

根据微观过程的可逆性,在原子的两个给定的定态之间,吸收光和受激辐射的概率是相同的[见上节的(2.5)式和(2.6)式],因而我们只需任选一种作具体计

算. 以下计算光的吸收过程的概率.

(一)微扰项的确定

设频率为 ω 的入射光沿 z 轴方向传播,它的电场沿 x 轴线偏振,电场强度为

$$E_x = E_\omega^0 \cos(kz - \omega t), \tag{3.1}$$

E_ω^0 是这种单色场的振辐. 由于在入射光中有各种频率的单色场,不同单色场的振辐是不一样的,所以我们特别在振幅下面用了脚标 ω.

在入射光中还包含各种频率的磁场,但是将原子中电子受到的入射光的电场力和磁场力作一比较,可以发现 $|e(\boldsymbol{v}\times\boldsymbol{B})|/|e\boldsymbol{E}| \sim v/c \ll 1$ ❶,因而磁场的作用可以忽略.

现在关心的是电场对原子内电子的作用. 若以原子中心为坐标原点,则在(3.1)式中 z 的变化范围是原子的线度,量级为 10^{-8} cm,面对于光学光谱,波长的量级是 $10^{-4} \sim 10^{-5}$ cm,因而可以认为 $kz = 2\pi z/\lambda \ll 1$. 在(3.1)式中可以将 kz 忽略不计,即认为电场在原子范围内是均匀的,而有

$$E_x = E_\omega^0 \cos(\omega t). \tag{3.2}$$

注意到电场是在 x 方向,因而电子与电场的相互作用为:

$$\hat{V} = -exE_x = -exE_\omega^0 \cos(\omega t) = -exE_\omega^0 (e^{i\omega t} + e^{-i\omega t})/2. \tag{3.3}$$

上式表示的是一个电偶极矩为 $-ex$ 的偶极子与电场的相互作用. 这是忽略了磁场作用并把电场简化成(3.2)式后得到的. 通常称这种近似为电偶极近似.

将(3.3)式和 §6.2.2(2.1)式比较,可知

$$\hat{F} = -eE_\omega^0 x/2. \tag{3.4}$$

(二)跃迁概率

按照 §6.2.2(2.10)式,在入射光中只有 $\omega \approx \omega_{mn}$ 的单色光才会导致原子从 φ_n 态到 φ_m 态的跃迁. 将(3.4)式代入 §6.2.2(2.10)式,得到单位时间的跃迁概率

$$W_{nm} = [\pi e^2/(2\hbar^2)][E^0(\omega_{mn})]^2 |x_{mn}|^2. \tag{3.5}$$

$E^0(\omega_{mn})$ 是频率为 ω_{mn} 的光的振幅,它的平方和光的能量密度成正比. 我们取电磁场能量密度的平均值来描述相应光波的能量密度 $\rho(\omega)$:

$$\rho(\omega) = (\overline{E_\omega^2} + \overline{B_\omega^2})/(8\pi) = \overline{E_\omega^2}/(4\pi) = (E_\omega^0)^2/8\pi, \tag{3.6}$$

这里利用了

$$\overline{E_\omega^2} = (E_\omega^0)^2 \frac{1}{T}\int_0^T \cos^2\omega t\,dt = \frac{1}{2}(E_\omega^0)^2.$$

将(3.6)式代入(3.5)式,可将单位时间跃迁概率写成

---

❶ 由 §2.4.4(4.30)式知 $|\boldsymbol{B}|$ 和 $|\boldsymbol{E}|/c$ 同量级.

$$W_{nm} = (4\pi^2 e^2/\hbar^2)\rho(\omega_{mn}) | x_{mn} |^2. \tag{3.7}$$

以上的计算假定了入射光是在 x 方向偏振，实际上入射光的偏振是混乱的，各向均等的，因而原子从 n 态到 m 态的实际跃迁概率是对各个偏振方向的贡献取平均，即

$$\begin{aligned} W_{nm} &= [4\pi^2 e^2/(3\hbar^2)]\rho(\omega_{mn})[| x_{mn} |^2 + | y_{mn} |^2 + | z_{mn} |^2] \\ &= [4\pi^2 e^2/(3\hbar^2)]\rho(\omega_{mn}) | \boldsymbol{r}_{mn} |^2. \end{aligned} \tag{3.8}$$

这就是在入射光的作用下，原子单位时间从低能态 n 跃迁到高能态 m 的概率，它描述了原子吸收频率为 ω_{mn} 的光后引起的跃迁. 按照微观过程的可逆性，在入射光的作用下，单位时间原子从高能态 m 通过辐射光而跃迁到低能态 n 的概率也是：

$$W_{mn} = [4\pi^2 e^2/(3\hbar^2)]\rho(\omega_{mn}) | \boldsymbol{r}_{mn} |^2. \tag{3.9}$$

这就是受激辐射的跃迁概率，是研究激光的重要基础.

注意到 $-e\boldsymbol{r}$ 是电子的电偶极矩，从(3.8)式和(3.9)式可以看到，在偶极近似下，感应跃迁的概率是由原子的电偶极矩在相应的始态和末态中的矩阵元决定的.

(三)自发辐射

前面讨论的都是在入射光的作用下，原子所发生的跃迁. 而且对于入射光，没有像爱因斯坦的光量子假设那样把它看成光子流，而是看作经典的电磁场，因而这样的原子发光与吸收光的理论不是一贯的量子理论. 正是由于这个原因，这样的理论不能全面解释原子发光问题，其中最典型的是不能解释原子的自发辐射问题. 按照量子力学，处于某个定态的原子，如无外来的作用改变它的能量，它将永远处于那个定态.

解释原子发光的彻底量子理论是量子电动力学. 按照量子电动力学，我们平常所说的"没有"电磁场，实际上是没有光子存在而电磁场处于基态的情况，所以原子的自发辐射是原子受到处于基态的场的扰动而从高能态向低能态的跃迁.

当局限在量子力学中时，对于自发辐射，可以按照爱因斯坦的办法，通过要求原子和辐射场达到热平衡来处理.

在能量密度为 $\rho(\omega_{mn})$ 的光场激发下单位时间从 m 态跃迁到 n 态的受激辐射概率为 $W_{mn} = B_{mn}\rho(\omega_{mn})$，从 n 态跃迁到 m 态而吸收光的概率为 $W_{nm} = W_{mn}$ [见(3.8)(3.9)式]. 自发地辐射光面从 m 态跃迁到 n 态的概率为 A_{mn}. 设在高能态 m 中有 N_m 个原子，在低能态 n 中有 N_n 个原子. 达到平衡时，单位时间从 m 态跃迁到 n 态和从 n 态跃迁到 m 态的原子数应相等. 因而有

$$[A_{mn} + B_{mn}\rho(\omega)]N_m = B_{mn}\rho(\omega)N_n. \tag{3.10}$$

另一方面，由统计物理知道[见 §7.3.1(1.8)式]，达到平衡时原子数 N_m 和 N_n 之比应为

$$N_m/N_n = Ce^{-(E_m-E_n)/(k_BT)} = Ce^{-\hbar\omega/(k_BT)}. \tag{3.11}$$

其中 k_B 为玻尔兹曼常量. 代入(3.10)式得到

$$\rho(\omega) = \frac{A_{mn}/B_{mn}}{e^{\hbar\omega/(k_BT)}-1}. \tag{3.12}$$

由此可见,为能达到平衡,自发辐射的概率 A_{mn} 不能为零. 将(3.12)式和平衡辐射(称为"黑体辐射")的概率[即普朗克分布 §8.2.1(1.13)式]比较可见,自发辐射概率为:

$$A_{mn} = \frac{\hbar\omega_{mn}^3}{c^3\pi^2}B_{mn} = \frac{4e^2\omega_{mn}^3}{3\hbar c^3}|\boldsymbol{r}_{nm}|^2. \tag{3.13}$$

其中 B_{mn} 是受激辐射概率.

(四)谱线强度和选择定则

知道了辐射过程的跃迁概率,就可以进一步算出辐射的能量. 注意到 A_{mn} 是单位时间内由 m 态自发跃迁到低能态 n 的概率. N_m 是处于激发态 m 的原子数,则在单位时间内由 m 态跃迁到 n 态的原子数为 $A_{mn}N_m$. 由于每个原子在跃迁过程中辐射的能量为 $\hbar\omega_{mn}$,因而单位时间内原子辐射出的频率为 ω_{mn} 的光的能量为

$$I_{mn} = \hbar\omega_{mn}A_{mn}N_m = N_m[4e^2\omega_{mn}^4/(3c^3)]|\boldsymbol{r}_{nm}|^2. \tag{3.14}$$

这就是频率为 ω_{mn} 的谱线的强度.

利用跃迁概率还可以确定系统的定态之间可否发生跃迁的规则,即选择定则. 只有满足这些规则的定态之间才允许发生跃迁,否则是禁戒的.

我们现在来确定,在偶极近似下原子光谱的选择定则. 从前面的跃迁概率公式可以看出,原子能否从 m 态跃迁到 n 态,决定于电子的位矢在这两个态之间的矩阵元 $|\boldsymbol{r}_{mn}|$ 是否为零. 只有 $|\boldsymbol{r}_{mn}|\neq 0$ 时,这两个定态之间才能发生跃迁. 为此我们来计算 $|\boldsymbol{r}_{mn}|$ 的三个分量 x_{mn}、y_{mn}、z_{mn}.

假定电子是在有心力场中运动,如氢原子和碱金属原子中的电子那样. 这时电子的波函数为

$$\psi_{nlm} = R_{nl}(r)Y_{lm}(\theta,\varphi) = R_{nl}(r)P_l^{|m|}(\cos\theta)e^{im\varphi}. \tag{3.15}$$

考虑到波函数的形式,用球坐标计算较为方便. 为了便于利用正交性条件,将 $\sin\varphi$ 和 $\cos\varphi$ 写成指数形式,这样有:

$$\begin{aligned} x_{nlm,n'l'm'} = &\int_0^\infty r^3R_{nl}^*(r)R_{n'l'}(r)\mathrm{d}r\times \\ &\int_{-1}^1 P_l^{|m|}(\cos\theta)\sin\theta P_{l'}^{|m'|}(\cos\theta)\mathrm{d}\cos\theta\times \\ &\frac{1}{2}\int_0^{2\pi}(e^{i(m'-m+1)\varphi}+e^{i(m'-m-1)\varphi})\mathrm{d}\varphi, \end{aligned} \tag{3.16}$$

$$y_{nlm,n'l'm'} = \int_0^\infty r^3 R_{nl}^*(r) R_{n'l'}(r)\,\mathrm{d}r \times$$
$$\int_{-1}^{1} P_l^{|m|}(\cos\theta)\sin\theta P_{l'}^{|m'|}(\cos\theta)\,\mathrm{d}\cos\theta \times$$
$$\frac{1}{2\mathrm{i}}\int_0^{2\pi}(\mathrm{e}^{\mathrm{i}(m'-m+1)\varphi} - \mathrm{e}^{\mathrm{i}(m'-m-1)\varphi})\,\mathrm{d}\varphi, \tag{3.17}$$

$$z_{nlm,n'l'm'} = \int_0^\infty r^3 R_{nl}^*(r) R_{n'l'}^*(r)\,\mathrm{d}r \times$$
$$\int_{-1}^{1} P_l^{|m|}(\cos\theta)\cos\theta P_{l'}^{|m'|}(\cos\theta)\,\mathrm{d}\cos\theta \times$$
$$\int_0^{2\pi} \mathrm{e}^{\mathrm{i}(m'-m)\varphi}\,\mathrm{d}\varphi. \tag{3.18}$$

首先求 m 的选择定则，利用正交归一化条件[见附录(Ⅰ.22)式]：

$$\int_0^{2\pi} \mathrm{e}^{\mathrm{i}(m'-m)\varphi}\mathrm{d}\varphi = 2\pi\,\delta_{m'm}, \tag{3.19}$$

从(3.18)式看出，z 的矩阵元不为零的条件是：

$$\Delta m = m' - m = 0, \tag{3.20}$$

而类似地由(3.16)式和(3.17)式知，x 和 y 的矩阵元不为零的条件是

$$\Delta m = m' - m = \pm 1, \tag{3.21}$$

再来推导 l 的选择定则. 我们考虑 z 的矩阵元(3.18)式，并令 $m' = m$，而只注意对 θ 的积分. 利用缔合勒让德函数满足的关系式

$$\cos\theta \mathrm{P}_l^{|m|}(\cos\theta) = \frac{l+|m|}{2l+1}\mathrm{P}_{l-1}^{|m|}(\cos\theta) + \frac{l-|m|+1}{2l+1}P_{l+1}^{|m|}(\cos\theta), \tag{3.22}$$

和缔合勒让德函数的正交性可以看出，z 的矩阵元不为零的条件是：

$$\Delta l = l' - l = \pm 1, \tag{3.23}$$

对 x 和 y 的矩阵元作类似的讨论也可得到同样结论.

这样，总结上面的讨论，得到选择定则：

$$\Delta l = \pm 1, \qquad \Delta m = 0, \pm 1. \tag{3.24}$$

由于在各个矩阵元积分中，对 r 的积分在 n 和 n' 取任何值时都不为零，所以对 n 没有选择定则.

我们来说明选择定则的物理意义. 决定偶极辐射强度的矩阵元是

$$\boldsymbol{r}_{nlm,n'l'm'} = \langle nlm \mid \boldsymbol{r} \mid n'l'm' \rangle. \tag{3.25}$$

由 §5.4.6 例 2 知道，位矢 $\boldsymbol{r}$ 的分量可以组成角动量 $l'' = 1$ 的波函数. (3.25)式中的 $\boldsymbol{r} \mid n'l'm'\rangle$ 可以看成是角动量为 l' 和 l'' 的两个波函数的乘积. 根据 §5.4.6(6.40)式知，由 l' 和 $l''(=1)$ 合成的角动量 $L = l'+1, l', l'-1$. 根据角动量波函数的

正交性，由(3.25)式可知 $r_{nlm,n'l'm'} \neq 0$ 的条件是 $l = l' + 1, l', l' - 1$. 同理可知 $m = m' + 1, m', m' - 1$. 另外一方面，由 §5.3.3 例 3 又知道，角动量为 l 的态的宇称是 $(-1)^l$，而位矢 $\boldsymbol{r}$ 的宇称为 -1，因而宇称守恒禁戒了 $l = l'$. 这样就有选择定则(3.21)式. 由此可见，选择定则，从本质上说，是来源于空间对称性(空间各向同性和空间反演对称性)导致的(角动量和宇称)守恒律.

§6.2.4　不稳定系统的寿命和能级宽度　能量－时间不确定关系

(一)激发态原子的寿命

由于处于激发态的原子有一定的概率自发跃迁到低能态去，因而原子停留在激发态的时间是有限的. 现在来讨论处于激发态原子的寿命.

设 $t = 0$ 时处于激发态 m 的原子数为 N_m^0，而在任意时刻 t 处于激发态 m 的原子数为 $N_m(t)$. 注意到单位时间内自发跃迁到低能态 n 的概率为 A_{mn}，则知在 t 到 $t + dt$ 时间内自发跃迁到 n 态的原子数为

$$\mathrm{d}N_m = -A_{mn}N_m\mathrm{d}t. \tag{4.1}$$

积分得到 m 态原子数减少的规律为：

$$N_m = N_m^0 \mathrm{e}^{-A_{mn}t}. \tag{4.2}$$

由于向 n 态自发跃迁使得原子的激发态 m 不稳定，它的平均寿命是

$$\tau = \frac{1}{N_m^0}\left(-\int_0^\infty t\mathrm{d}N_m\right) = \frac{1}{A_{mn}}. \tag{4.3}$$

激发态原子是不稳定系统的一个例子. 在微观物理学中，这类不稳定系统很多. 例如，进行 α 衰变的原子核，π-N 散射的 3-3 共振(Δ 粒子)等等. 下面我们来证明，所有这些不稳定系统都不具有唯一确定的能量，它们的能级具有一定的宽度.

(二)不稳定系统的能级宽度　能量－时间不确定关系

假定有一种不稳定的粒子，平均寿命为 τ. 在 $t = 0$ 时，系统中的粒子数为 N_0，则由于蜕变，在 t 时刻，粒子数成为 $N(t) = N_0\mathrm{e}^{-t/\tau}$.

如果将还未蜕变的的整体看作一个系统(例如将 N 个处于激发态中的原子看成一个系统)，则这个系统的波函数中，依赖于时间的部分应该是

$$\psi \approx \mathrm{e}^{-t/(2\tau)}\mathrm{e}^{-(\mathrm{i}/\hbar)E_0 t}, \tag{4.4}$$

其中第二个因子是通常的时间因子，而第一个因子是考虑到了不稳定体系数目随时间而减少的规律，因为概率密度 $\approx |\psi|^2 \propto \mathrm{e}^{-t/\tau}$.

波函数(4.4)不是具有一定能量的定态，因为具有一定能量的定态的时间因子是 $\mathrm{e}^{-(\mathrm{i}/\hbar)E_0 t}$. 很明显，(4.4)式代表了随时间而衰减的波. 它可以看成是许多个具

有一定能量的状态的叠加. 这实际上就是将 ψ 展开成傅里叶积分:

$$\psi = \int C_E e^{-(i/\hbar)Et} dE. \tag{4.5}$$

我们来考虑,相互叠加而组成 ψ 的这些定态 $C_E e^{-(i/\hbar)Et}$ 位于多大的能量间隔之中. 也就是说看看系数 C_E 在多大的能量间隔范围中不等于零.

为了得到这一能量间隔,可以不必进行具体计算,而只需要与 §5.2.4 中讨论过的波包对比一下. 波包是不具有一定动量 p 的波,它可以看成是一些具有一定动量的波 $c_p e^{-(i/\hbar)px}$ 的迭加. 在 §5.2.4 中计算过,如果波包在空间范围 Δx 中不为零,则它由动量相差 Δp 的波组成,其中 $\Delta p \Delta x \geq \hbar$.

由以上对比可知,如果不稳定系统经过时间 Δt 后蜕变($\Delta t = \tau$),则它应该由能量间隔范围为 ΔE 的定态组成,而

$$\Delta E \Delta t \geq \hbar. \tag{4.6}$$

(4.6)式是适用于任意不稳定体系的普遍关系式,它称为能量 - 时间不确定关系. 但这个名词应该说是不够确切的,不要把它和坐标 - 动量不确定关系的意义相混淆. 我们知道,坐标 - 动量不确定关系的意义是说,不能用坐标与动量这一类表征粒子性质的物理量确切地描述微观粒子,微观粒子不能同时具有确定的坐标与动量. 显然不能照样地把(4.6)式理解成:"不能用能量与时间确切地描述微观粒子".

能量 - 时间不确定关系的物理意义,是说明了不稳定体系能量的准确度与其寿命的关系. 如不稳定体系的寿命是 $\Delta t = \tau$;则它不是具有确定能量的定态,但可看作是能量处子范围在 $\Delta E \geq \hbar/\tau$ 中的许多定态的迭加. 当寿命比较长时, ΔE 可以很小,此时仍可认为不稳定体系处于某种准定态能级中,只不过这一能级有一定的宽度 ΔE . 由于这一原因我们定义:

$$\Gamma = \hbar/\tau. \tag{4.7}$$

为能级的宽度.

例如,激发态原子的寿命一般说来是在 $10^{-7} \sim 10^{-8}$ s 之间. 因此,激发态能级的宽度为(参看附录Ⅶ)

$$\Gamma = \hbar/\tau \approx 10^{-19}\text{erg} \approx 10^{-7}\text{eV}. \tag{4.8}$$

我们知道,原子能级间距的数量级为 1eV,因而能级的宽度比能级的间距小很多. 由这样的能级之间的跃迁,得到的光谱是具有一定宽度的线. 对光谱的可见光部分,谱线的宽度一般数量级为 10^{-3} Å.

极限情况下,寿命 $\tau \to \infty$,就成了稳定体系,能级宽度 $\Gamma \to 0$,即稳定体系有确定的能量. 这是自然的. 对极不稳定的体系 $\tau \approx 0$,此时能级宽度 $\Gamma \to \infty$,这种体系根本不能当作近似孤立的体系.

习　题

1. 已知从薛定谔绘景到相互作用绘景的幺正变换为：$U = e^{(i/\hbar)H_0 t}$.

(a)证明在相互作用绘景中态矢的演化方程为

$$i\hbar\partial\psi_I/\partial t = V^I\psi_I,$$

算符的演化方程为

$$dF_I/dt = (-i/\hbar)[F^I, H^I].$$

(b)给出态矢演化方程的 H_0 表象.

2. 用微扰方法求跃迁概率振幅 a_n 至二级近似.

3. 电荷为 e 的谐振子在时间 $t = 0$ 时处于基态，$t > 0$ 以后在 $\varepsilon = \varepsilon_0 e^{-\gamma t}$ 的电场中. 求谐振子处于激发态的概率.

4. 设 $t = 0$ 时，氢原子处于基态，以后受到周期性的均匀电场 $\varepsilon = \sin\omega t$ 的作用而电离，电离后电子为自由状态. 求电场应有的最小角频率和在 t 时刻跃迁到电离状态的概率.

5. 计算氢原子处于 $n = 2, l = 1, m = 0$ 态的平均寿命.

第三章　弹性散射

散射是另一类基本动力学问题，在§1.3.4中用经典力学研究过这类问题．这一节用量子力学研究微观粒子间的散射．散射过程可以用粒子的状态是否因碰撞而发生改变区分为两种类型：一种碰撞的结果，粒子间只有动量交换，而粒子的内部状态不变．这类散射称为弹性散射．另一类碰撞使粒子的内部状态发生改变，如粒子被激发，甚至碎裂．这类散射称为非弹性散射．

碰撞的具体情况与粒子本身的结构以及它们之间的相互作用性质密切相关．通过对散射结果的分析，可以探知粒子的结构以及粒子间相互作用的规律，推动基础理论的发展．人们之所以能从原子到夸克这样一个层次又一个层次地深入认识物质结构，在很大程度上是依赖于对散射的研究．例如，卢瑟福的 α 粒子散射和美国斯坦福直线加速器中心 SLAC 的电子深度非弹性散射，就是在认识原子的有核结构和核子的部分子结构中起了关键性作用的散射实验．

§6.3.1　散射问题的一般讨论

(一)散射截面

在§1.3.4图2上画出了散射过程的示意图．设入射方向为 z 方向．入射粒子和靶粒子的相互作用由相互作用势能描述，并近似地认为相互作用范围是有限的．在入射粒子没有进入相互作用范围内时，可以用平面波描述；进入相互作用范围后，因受势场作用而散射到各个可能的方向．入射粒子被散射后的运动方向与入射方向之间的夹角 θ 称为散射角．粒子探测器装置在远离散射中心的地方，探测到的是被散射后又离开了相互作用范围的粒子．探测器的窗口对散射中心张成一个立体角，凡通过这个立体角的粒子都能被观测到．被散射到某一个立体角中的粒子数是基本的实验观测量．在理论上计算粒子被散射到 θ,φ 方向上单位立体角中的概率，是研究散射问题的中心课题．在§1.3.4(4.1)式中已经指出，这一概率可以用散射微分截面

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}=\frac{1}{J}\frac{\mathrm{d}N}{\mathrm{d}\Omega} \tag{1.1}$$

表征，其中 J 是入射粒子流密度，$\mathrm{d}N$ 是单位时间内散射到 $\mathrm{d}\Omega$ 中去的粒子数．

按照量子力学，粒子的位置概率分布由波函数决定，因而求粒子被散射后的空间分布，就归结为通过薛定谔方程求散射后的波函数．

设两个粒子间的相互作用势能为 $U(|\boldsymbol{r}_1-\boldsymbol{r}_2|)$，则求解系统的薛定谔方程可

归结为求解定态方程：

$$\left[-\left[\hbar^2/(2m_1)\right]\nabla_1^2-\left[\hbar^2/(2m_2)\right]\nabla_2^2\right]\Psi+U(|\boldsymbol{r}_1-\boldsymbol{r}_2|)\Psi=E_{总}\Psi. \tag{1.2}$$

这是个二体运动方程，它可以化为两个运动方程，一个描述质心的运动，另一个描述粒子之间的相对运动. 若选用质心系，则只需讨论粒子之间的相对运动方程：

$$-\left[\hbar^2/(2\mu)\right]\nabla^2\psi+U(r)\psi=E\psi, \tag{1.3}$$

其中 μ 是约化质量[见 §1.3.1(1.6)式]，r 是散射粒子相对靶粒子的位置矢量的长度，E 是相对运动的能量. 在实验上 E 是通过加速人射粒子而确定的，在计算中作为已知条件.

让我们进一步强调指出，在用(1.1)式计算散射截面时，隐含一个假定：在 r 大时，粒子不再受散射势场的作用，而自由运动. 只有在这一情况下，单位时间通过立体角 $\mathrm{d}\Omega$ 的粒子数 $\mathrm{d}N$ 才与 r 无关，(1.1)式才有意义. 这一假定意味着，势能 $U(r)$ 在 $r\to\infty$ 时下降足够快，例如：是短程作用，有力程 a，当 $r>a$ 时 $U=0$；或者，至少在 $r\to\infty$ 时，$U\to 0$ 比 r^{-1} 更快.

按照这一假定，在 $r>a$ 或 $r\to\infty$ 的区域里，粒子作自由运动，哈密顿为

$$\hat{H}\approx\hat{H}_0=\hat{\boldsymbol{p}}^2/(2\mu).$$

注意，$\hat{H}_0$ 和动量 $\hat{\boldsymbol{p}}$ 及角动量 $\hat{\boldsymbol{L}}^2,\hat{L}_z$ 均可对易：

$$[\hat{H}_0,\hat{\boldsymbol{p}}]=0,\qquad[\hat{H}_0,\hat{\boldsymbol{L}}^2]=0,\qquad[\hat{H}_0,\hat{L}_z]=0.$$

因而，自由粒子可以有平面波(德布罗意波)和球面波两种解.

根据散射问题的物理条件，在 $r>a$ 或 $r\to\infty$ 的区域里，波函数应该由人射平面波和散射球面波两部分组成. 这一要求就成为散射问题的波函数在 $r\to\infty$ 时所应满足的边界条件.

(二)散射问题的边界条件　散射振幅

由于场有中心对称性，所以和 §6.1.1(1.26)～(1.31)式一样，可以写

$$\psi(r,\theta,\varphi)=R(r)Y_{lm}(\theta\varphi), \tag{1.4}$$

$$R(r)=u(r)/r, \tag{1.5}$$

$$\mathrm{d}^2u/\mathrm{d}r^2+\left[(2\mu/\hbar^2)(E-U(r))-l(l+1)/r^2\right]u=0, \tag{1.6}$$

$$u(0)=0. \tag{1.7}$$

1. $r\to\infty$ 时波函数的渐近形式

上面已经指出，远离散射中心观察，粒子作自由运动. 这样，在 $r\to\infty$ 时，波函数应由两部分组成，一部分是描述沿 z 方向运动的人射粒子的平面波 $\psi_1=Ae^{ikz}$；另一部分是描述粒子被散射后从散射中心向外发散的球面出射波 $\psi_2=B(\theta,$

$\varphi)e^{ikr}/r$ ❶.

对于弹性散射，散射后粒子的能量不变，而且散射前后都假定了粒子是在自由状态，因而散射前后，波矢 k 的大小不变，球面波的振幅 B 与 θ,φ 有关，是因为粒子被散射到不同方向的概率不一样. 这样，$r\to\infty$ 时，波函数的一般渐近形式为：$r\to\infty$，$\psi\to Ae^{ikz}+B(\theta,\varphi)e^{ikr}/r$. 或者写 $B(\theta,\varphi)=Af(\theta,\varphi)$，而有

$$\psi\underset{r\to\infty}{\longrightarrow}A[e^{ikz}+f(\theta,\varphi)e^{ikr}/r]. \tag{1.8}$$

把它代回方程(1.3)后可以看出，只要相互作用势能 $U(r)$ 随 $r\to\infty$ 而趋于零的速度比 r^{-1} 快，(1.8)式就能满足方程(1.3)，而不论 $U(r)$ 的具体形式如何. 这就是说，(1.8)式的确是在广泛情况下都能成立的定态波函数的渐近形式. 它是在散射问题中求解定态方程的边界条件[比较一维“散射”问题的边界条件 §5.3.1(1.49)式. 注意，在一维情况下，球面散射波成为向右的透射波和向左的反射波.]. 在这一边界条件下解方程(1.3)式，可以求得受到具体势能 $U(r)$ 散射的球面波的 $f(\theta,\varphi)$.

2. $f(\theta,\varphi)$ 的物理意义

(1.8)式第二项是球面波，见 §5.3.1(1.16)式. 相应的概率流密度已在那一节中算出为 $j_{d\Omega}=|A|^2v|f(\theta,\varphi)|^2d\Omega$，见该节(1.17)式. 在现在的问题中，我们将波函数归一化为粒子数，因而当 r 大时，单位时间内，射进 (θ,φ) 方向附近的立体角 $d\Omega$ 中的粒子数为

$$dN=v|A|^2|f(\theta,\varphi)|^2d\Omega, \tag{1.9}$$

式中 v 为粒子速度. 又按 §5.3.1(1.15)式，由于入射粒子是平面波 ψ_i，入射粒子的概率流密度为 $J=|A|^2v$，因而散射微分截面为

$$d\sigma/d\Omega=dN/(Jd\Omega)=|f(\theta,\varphi)|^2. \tag{1.10}$$

由此可见，微分截面由球面散射波的振幅 $f(\theta,\varphi)$ 决定，通常称 $f(\theta,\varphi)$ 为散射振幅.

§6.3.2　玻恩近似

令

$$k^2=2\mu E/\hbar^2,\qquad V(r)=2\mu U(r)/\hbar^2, \tag{2.1}$$

则 §6.3.1 的方程(1.3)改写成

$$\nabla^2\psi+k^2\psi=V(r)\psi. \tag{2.2}$$

---

❶ 球面波的这一函数形式使得通过立体角元 $d\Omega$ 的概率流沿径向向外，在 r 大时强度与 r 无关，只依赖于 θ、φ，见 §5.3.1(1.17)式.

我们在一种最简单的情况下讨论方程的近似解法，这就是假定入射粒子的能量很大，以致可以把它与靶粒子的相互作用势能看成微扰.

未扰态的方程 $\nabla^2\psi^{(0)}+k^2\psi^{(0)}=0$ 是自由运动方程. 用微扰处理有两种办法. 一种办法是将微扰势能所引起的跃迁过程看成是从沿 z 方向入射的平面波到沿 θ, φ 方向出射的平面波之间的跃迁. 另一种方法是将这一跃迁过程看成是从沿 z 方向入射的平面波到球面出射波之间的跃迁，见图 1.［注意，自由粒子运动方程有平面波和球面波两种解，见下一节的(3.2)式和(3.12)式］. 按前一种办法计算散射截面留到习题 3 中再考虑. 这里采用第二种办法，将出射波看成球面波.

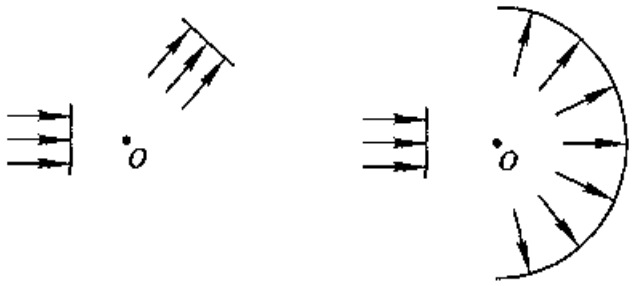


图 1　从平面波到平面波和从平面波到球面波的跃迁

按照这一考虑，将方程(2.2)式的解写成

$$\psi=\psi^{(0)}+u. \tag{2.3}$$

$\psi^{(0)}$ 是微扰未加入时粒子的波函数，即入射粒子平面波，u 是出射球面波，被认为是一级小量. 将上式代入(2.2)式，并略去二级小量，得到一级修正 $u(\boldsymbol{r})$ 所满足的方程：

$$\nabla^2u+k^2u=V(\boldsymbol{r})\psi^{(0)}, \tag{2.4}$$

按照上面的分析，我们要求的是方程在 $r\to\infty$ 时的渐近解，从而得到散射振幅. 为此，采用格林函数方法.

在 §3.1.4 中用格林函数方法求解过静电场. 在那里是将格林函数定义为点电荷激发的电势. 现在从更普遍的意义上说明一下格林函数方法的基本思想.

将方程(2.4)写为

$$Au=v, \tag{2.5}$$

其中 $A=\nabla^2+k^2$ 是一个微分算符，$v=V(\boldsymbol{r})\psi^{(0)}$ 是一个已知函数. 用 A 的逆算符 A^{-1} 乘方程的左右两边，就得到解 $u=A^{-1}v$. 式中的 A^{-1} 称为方程(2.5)的格林函数，常用符号 G 表示，它满足方程

$$AG=\mathbf{1}. \tag{2.6}$$

解这一方程，得到格林函数 G，就可以写出原方程的解

$$u=Gv. \tag{2.7}$$

用抽象符号写出的(2.5)～(2.7)式表述了格林函数方法的基本思想. 为了加深对格林函数方法基本思想的理解，再举 §3.1.4 的方程(4.25)为例. 这一方程可以写成本节(2.5)式的形式，其中 $A=-\varepsilon_0\nabla^2$. 本节(2.6)式即那里的(4.27)式；而本节(2.7)式就是那一节(4.29)式的第一项. §3.1.4(4.29)式还有第二项，是由于

边界条件 §3.1.4(4.26)式.

回到本节所讨论的问题.(2.4)式采用的是坐标表象.相应地,格林函数方程(2.6)式成为

$$(\nabla^2 + k^2)G(\boldsymbol{r} - \boldsymbol{r}') = \delta^3(\boldsymbol{r} - \boldsymbol{r}'). \tag{2.8}$$

而利用格林函数得到的解(2.7)式就是

$$u(\boldsymbol{r}) = \int G(\boldsymbol{r} - \boldsymbol{r}')V(\boldsymbol{r}')\psi(\boldsymbol{r}')\mathrm{d}^3\boldsymbol{r}'. \tag{2.9}$$

我们用傅里叶变换方法来解格林函数的方程(2.8)式.为了书写简单,暂时写 $\boldsymbol{r} - \boldsymbol{r}'$ 为 $\boldsymbol{r}$.格林函数 $G(\boldsymbol{r})$ 的傅里叶展开是

$$G(\boldsymbol{r}) = (2\pi)^{-3}\int g(\boldsymbol{k}')\mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{r}}\mathrm{d}^3\boldsymbol{k}'. \tag{2.10}$$

而由附录(Ⅰ.20)式有 $\delta^3(\boldsymbol{r}) = (2\pi)^{-3}\int \mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{r}}\mathrm{d}^3\boldsymbol{k}'$.和(2.10)式一道代入方程(2.8)式,利用 $\nabla^2\mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{r}} = -k'^2\mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{r}}$ 得到 $g(\boldsymbol{k}') = 1/(k^2 - k'^2)$.代入(2.10)式得到

$$\begin{aligned} G(\boldsymbol{r}) &= \frac{1}{(2\pi)^3}\int \frac{\mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{r}}}{k^2 - k'^2}\mathrm{d}^3\boldsymbol{k}' \\ &= \frac{1}{(2\pi)^3}\int_0^\infty k'^2\mathrm{d}k'\int_0^{2\pi}\mathrm{d}\varphi\int_0^\pi \frac{\mathrm{e}^{\mathrm{i}k'r\cos\theta}}{k^2 - k'^2}\sin\theta\mathrm{d}\theta \\ &= \frac{1}{4\pi^2 r}\left(\int_0^\infty (\mathrm{i}k')\frac{\mathrm{e}^{\mathrm{i}k'r}}{k'^2 - k^2}\mathrm{d}k' + \int_0^\infty (-\mathrm{i}k')\frac{\mathrm{e}^{-\mathrm{i}k'r}}{k'^2 - k^2}\mathrm{d}k'\right). \end{aligned}$$

在后一积分中作代换 $k' \to -k'$ 再和前一积分合并,得到

$$G(\boldsymbol{r}) = \frac{1}{4\pi^2 r}\int_{-\infty}^\infty \frac{\mathrm{i}k'\mathrm{e}^{\mathrm{i}k'r}}{k'^2 - k^2}\mathrm{d}k' = \frac{1}{4\pi^2 r}\frac{\mathrm{d}}{\mathrm{d}r}\int_{-\infty}^\infty \frac{\mathrm{e}^{\mathrm{i}k'r}}{k'^2 - k^2}\mathrm{d}k'.$$

但是,上式的被积函数在 $k' = \pm k$ 处有极点,因而积分发散.为了克服这一困难得到有意义的结果,在被积函数的分母中加一个无穷小的虚部 $-\mathrm{i}\eta$($\eta > 0$),在完成积分后再让 $\eta \to 0$:

$$G(\boldsymbol{r}) = \frac{1}{4\pi^2 r}\frac{\mathrm{d}}{\mathrm{d}r}\lim_{\eta\to 0}\int_{-\infty}^\infty \frac{\mathrm{e}^{\mathrm{i}k'r}}{k'^2 - (k^2 + \mathrm{i}\eta)}\mathrm{d}k'. \tag{2.11}$$

这相当于将极点从实轴移开.新的极点位于 $k' = \pm\sqrt{k^2 + \mathrm{i}\eta} \approx \pm(k + \mathrm{i}\varepsilon)$ 处,其中 $\varepsilon = \eta/2k$ 是另一无穷小正数,如图 2(a).上式中的积分是沿 k' 实轴的无限积分.由于被积函数的分子为 $\mathrm{e}^{\mathrm{i}k'r}$ 因而可以用上半平面半径为无穷大的半圆周来完成闭合回路,见图 2(a).在这一闭合回路中,被积函数只有一个极点 $k + \mathrm{i}\varepsilon$,利用留数定理得到

$$\lim_{\eta\to 0}\int_{-\infty}^\infty \frac{\mathrm{e}^{\mathrm{i}k'r}}{k'^2 - (k^2 + \mathrm{i}\eta)}\mathrm{d}k' = \lim_{\varepsilon\to 0} 2\pi i\mathrm{Res}(k + \mathrm{i}\varepsilon) = i\pi\frac{\mathrm{e}^{\mathrm{i}kr}}{k}.$$

代入(2.11)式得到格林函数：

$$G(\boldsymbol{r}) = -\frac{1}{4\pi}\frac{e^{ikr}}{r}. \tag{2.12}$$

利用这一格林函数，由(2.9)式得到方程(2.5)的解

$$u(\boldsymbol{r}) = -\frac{1}{4\pi}\int \frac{V(\boldsymbol{r}')\psi^{(0)}(\boldsymbol{r}')e^{ik|\boldsymbol{r}'-\boldsymbol{r}|}}{|\boldsymbol{r}'-\boldsymbol{r}|}d\tau'. \tag{2.13}$$

积分遍及势能存在的整个区域. 我们一开始已假定这个范围是有限的.

让我们指出，(2.11)式在被积函数的分母中加上了一个小虚部使得极点离开积分路径. 而当 $\eta \to 0$ 时，被积函数的极点回到实轴，在积分路径上“挤”出了两个半径为无穷小的圆弧，如图2(b). 因此，(2.11)式也可以理解为，被积函数的极点仍在实轴上，但积分路径 C 是沿实轴进行，在经过极点时分别用两个半径为无穷小的圆弧绕过去，如图2(b)中的实线，

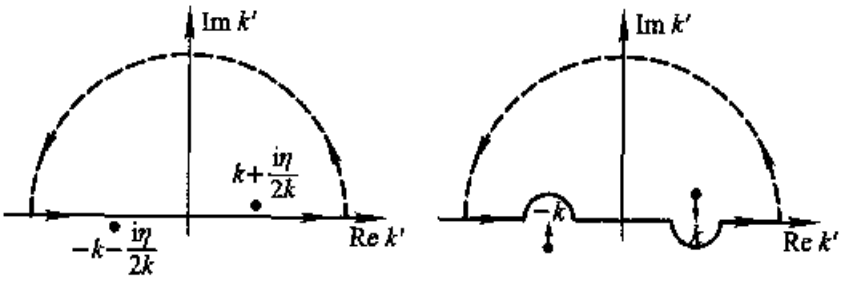


图2　在复 k' 平面上完成(2.11)式积分的回路

$$G(\boldsymbol{r}) = \frac{1}{4\pi^2 r}\frac{d}{dr}\int_C \frac{e^{ik'r}}{k'^2-k^2}dk'. \tag{2.11'}$$

(2.13)式给出了任一点 $\boldsymbol{r}$ 处的解. 为求得在 $r\to\infty$ 时的渐近解，将 $|\boldsymbol{r}'-\boldsymbol{r}|$ 的形式作一点改变：

$$|\boldsymbol{r}'-\boldsymbol{r}| = [(\boldsymbol{r}'-\boldsymbol{r})\cdot(\boldsymbol{r}'-\boldsymbol{r})]^{1/2} = [r^2+r'^2-2r(\boldsymbol{e}_r\cdot\boldsymbol{r}')]^{1/2}, \tag{2.14}$$

式中 $\boldsymbol{e}_r$ 是 $\boldsymbol{r}$ 方向的单位矢量. 注意到 r' 是有限的，在 $r\to\infty$ 时 r'/r 是小量，把上式两边同除以 r，按小量展开保留到一级，则有 $|\boldsymbol{r}'-\boldsymbol{r}| \approx r-\boldsymbol{e}_r\cdot\boldsymbol{r}'$. 代入(2.13)式，并在分母中略去比 r 小的量 $\boldsymbol{e}_r\cdot\boldsymbol{r}'$，就得到在 r 大时 u 的渐近形式

$$r\to\infty \qquad u(\boldsymbol{r}) = -\frac{1}{4\pi}\frac{e^{ikr}}{r}\int e^{-ik\boldsymbol{e}_r\cdot\boldsymbol{r}'}V(\boldsymbol{r}')\psi^{(0)}(\boldsymbol{r}')d\tau'. \tag{2.15}$$

将入射粒子平面波 $\psi^{(0)}$ 代入上式，并注意到 $z' = \boldsymbol{r}'\cdot\boldsymbol{e}_{z0}$，而 $\boldsymbol{e}_{z0}$ 是 z 方向的单位矢量，就可以将 $u(\boldsymbol{r})$ 的渐近形式写成

$$u(\boldsymbol{r}) = -\frac{1}{4\pi}\frac{e^{ikr}}{r}\int e^{ik(\boldsymbol{e}_{z0}-\boldsymbol{e}_r)\cdot\boldsymbol{r}'}V(\boldsymbol{r}')d\tau'. \tag{2.16}$$

将它和散射球面波的一般形式 §6.3.1(1.8)式第二项作一比较，得到散射振幅为

$$f(\theta,\varphi) = -\frac{1}{4\pi}\int e^{ik(\boldsymbol{e}_{z0}-\boldsymbol{e}_r)\cdot\boldsymbol{r}'}V(\boldsymbol{r}')d\tau'. \tag{2.17}$$

用矢量 $\boldsymbol{K}$ 表示波矢的改变：

$$\boldsymbol{K} = k(\boldsymbol{e}_{z0} - \boldsymbol{e}_r), \tag{2.18}$$

它的大小为

$$K = 2k\sin\theta/2. \tag{2.19}$$

这样，(2.17)式成为

$$f(\theta) = -\frac{1}{4\pi}\int e^{i\boldsymbol{K}\cdot\boldsymbol{r}'}V(\boldsymbol{r}')\,d\tau' = -\frac{\mu}{2\pi\hbar^2}\int e^{i\boldsymbol{K}\cdot\boldsymbol{r}'}U(\boldsymbol{r}')\,d\tau', \tag{2.20}$$

这里已用到了关系式(2.1).

如果场有中心对称，$U(r)$ 只依赖于 r 的大小，就可以进一步完成上式中的角向积分. 取极轴与 $\boldsymbol{K}$ 平行的极坐标系. 在这个坐标系中 $\boldsymbol{K}\cdot\boldsymbol{r}' = Kr'\cos\theta'$，于是有

$$\begin{aligned} f(\theta) &= -\frac{\mu}{2\pi\hbar^2}\int_0^\infty U(r')r'^2dr'\int_0^\pi e^{iKr'\cos\theta'}\sin\theta'd\theta'\int_0^{2\pi}d\varphi' \\ &= -\frac{2\mu}{K\hbar^2}\int_0^\infty rU(r)\sin(Kr)\,dr. \end{aligned} \tag{2.21}$$

这是利用微扰方法得到的一级近似结果，称为一级玻恩近似. 以上在得到这一近似时是把整个相互作用势能当作微扰，这意味着要求入射粒子的动能很大. 在这种条件下，按(2.21)式，只要知道中心对称散射势场 $U(r)$，就可以计算微分散射截面.

【例】 计算高速带电粒子被中性原子散射的微分散射截面. 设入射粒子的电荷为 $Z'e$，中性原子的原子序数为 Z.

解 首先分析散射势场的形式，原子核的库仑场与入射带电粒子的相互作用势能为 $ZZ'e^2/r$，但由于在原子周围存在带负电的电子云，它屏蔽库仑场对带电粒子的作用，使得在同一 r 处，相互作用效应减弱. 将这种屏蔽库仑场与带电粒子的相互作用势能写为

$$U(r) = (ZZ'e^2/r)e^{-r/a}, \tag{2.22}$$

a 为一表征原子半径的参量. (2.22)式的意义是用因子 $e^{-r/a}$ 描述电子云的屏蔽效应. 可以看出 $U(r)$ 保持了中心对称性，并随着 $r\to\infty$，比 r^{-1} 更快地趋于零.

将(2.22)式代入(2.21)式和§6.3.1(1.10)式，得到带电粒子被散射的微分截面为

$$\frac{d\sigma}{d\Omega} = \frac{4\mu^2Z^2Z'^2e^4}{K^2\hbar^4}\left|\int_0^\infty \sin(Kr)e^{-r/a}dr\right|^2 = \frac{4\mu^2Z^2Z'^2e^4}{\hbar^4}\frac{1}{(K^2+1/a^2)^2}. \tag{2.23}$$

当 $a\to\infty$ 时，(2.22)式成为无屏蔽的库仑场，而(2.23)式成为库仑场中散射的卢瑟福公式，见§1.3.4(4.13)式.

§6.3.3 分波法

在讨论粒子被中心场弹性散射的问题时，分波法是一种常用的求解散射振幅的方法.

由于势能场的中心对称性，因而在散射过程中角动量守恒. 这是分波法的基础. 描述角动量有确定值的波函数称为分波. 如 $l = 0$ 和 $l = 1$ 的波函数分别称为 s 波和 p 波等等. 为了应用角动量守恒分析散射过程，可以把受势能场作用前后的定态按分波展开. 即认为粒子受势能场作用前后都分别处于各种分波的叠加态中，而讨论势能场对各分波的效应. 这样做的结果会发现，势能场的效应在于改变分波的相位.

各种分波都是球面波. 分波法是将受势场作用前后的自由运动状态都作为球面波. 下面先给出自由粒子的球面波解.

(一)自由粒子方程的球面波解

自由粒子方程[$U(r) = 0$ 的方程]的平面波解是已知的. 我们来求它的球面波解.

令 $k = \sqrt{2\mu E}/\hbar$，自由粒子的径向方程 §6.3.1(1.6)式可写为

$$\mathrm{d}^2u_{kl}/\mathrm{d}r^2 + [k^2 - l(l+1)/r^2]u_{kl} = 0. \tag{3.1}$$

当 $l = 0$ 时其解为

$$u_{k0} = Ae^{ikr} + Be^{-ikr}. \tag{3.2}$$

如果方程(3.1)在全空间适用，它的解就应满足 $r = 0$ 的边条件 §6.3.1(1.7)式. 这样的解可以在(3.2)式中令 $B = -A$ 得到，或写为

$$u_{k0} = A\sin kr. \tag{3.3}$$

我们设法找一个递推关系，利用它从(3.3)式出发得到任意 l 的解. 为此，令

$$\chi_{kl} = R_{kl}/r^l = u_{kl}/r^{l+1}. \tag{3.4}$$

(当 $l = 0$ 时 $\chi_{k0} = R_{k0} = u_{k0}/r$). 它满足方程

$$\chi''_{kl} + 2(l+1)\chi'_{kl}/r + k^2\chi_{kl} = 0. \tag{3.5}$$

将它对 r 求导得到

$$\chi'''_{kl} + 2(l+1)\chi''_{kl}/r + [k^2 - 2(l+1)/r^2]\chi'_{kl} = 0. \tag{3.6}$$

另一方面，由 $\chi_{k,l+1}$ 所满足的方程

$$\chi''_{k,l+1} + 2(l+2)\chi'_{k,l+1}/r + k^2\chi_{k,l+1} = 0,$$

可知 $r\chi_{k,l+1}'' + 2(l+2)\chi_{k,l+1}' + k^2r\chi_{k,l+1} = 0$. 将 $(r\chi_{k,l+1})' = r\chi_{k,l+1}' + \chi_{k,l+1}$ 以及 $(r\chi_{k,l+1})'' = r\chi_{k,l+1}'' + 2\chi_{k,l+1}'$，代入，得到

$$(r\chi_{k,l+1})'' + 2(l+1)(r\chi_{k,l+1})'/r + [k^2 - 2(l+1)/r^2](r\chi_{k,l+1}) = 0.$$

显然，$(r\chi_{k,l+1})$ 所满足的这一方程和 χ'_{kl} 所满足的方程(3.6)式完全相同. 因而得到递推关系

$$\chi_{k,l+1} = \chi'_{kl}/r. \tag{3.7}$$

由此得 $\chi_{kl} = (1/r)(\mathrm{d}/\mathrm{d}r)\chi_{k,l-1} = [(1/r)(\mathrm{d}/\mathrm{d}r)]^2\chi_{k,l-2} = \cdots$，

$$\chi_{kl} = \left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)^l \chi_{k0} = \left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)^l R_{k0}. \tag{3.8}$$

将(3.3)(3.4)式代入，得到

$$R_{kl} = Ar^l\left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)^l \frac{\sin kr}{r}. \tag{3.9}$$

定义球贝塞尔函数

$$j_l(x) = (-x)^l\left(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\right)^l \frac{\sin x}{x}. \tag{3.10}$$

前几个球贝塞尔函数是：

$$j_0(x) = \frac{\sin x}{x},$$

$$j_1(x) = \frac{\sin x}{x^2} - \frac{\cos x}{x},$$

$$j_2(x) = \left(\frac{3}{x^3} - \frac{1}{x}\right)\sin x - \frac{3\cos x}{x^2}.$$

可以将 R_{kl} 写为

$$R_{kl} = Aj_l(kr), \tag{3.11}$$

其中 A 是另一个归一化系数. 代入§6.3.1(1.4)式就得到自由粒子方程的球面波解，

$$\psi_{klm}(r,\theta,\varphi) = Aj_l(kr)\mathrm{Y}_{lm}(\theta,\varphi). \tag{3.12}$$

其中 $k = \sqrt{2\mu E}/\hbar$. 这一解适用于全空间，包括 $r = 0$，因而是描述自由粒子运动的波函数.

既然平面波(德布罗意波) $\mathrm{e}^{\mathrm{i}k\cdot r}$ 和球面波(3.12)式都是描述自由粒子运动的波函数，它们之间必有相互叠加的关系. 令 k 指向 z 轴. 由于轴对称，在球面波展开式中只有 $m = 0$ 的项. 展开式系数由勒襄德多项式的性质决定，结果是

$$\mathrm{e}^{\mathrm{i}kz} = \sum_{l=0}^{\infty}(2l+1)\mathrm{i}^l j_l(kr)\mathrm{P}_l(\cos\theta). \tag{3.13}$$

利用(3.10)式很容易得到 x 大时球贝塞尔函数 $j_l(x)$ 的渐近表达式. 注意，对分母中的 x 求导会增大 x 的负幂次，这种项当 $x \to \infty$ 时可以略去. 因而只需对分子中的 $\sin x$ 求导 l 次. 每次求导运算都使宗量增加 $-\pi/2$：$-\mathrm{d}(\sin x)/\mathrm{d}x = -\cos x$

$= \sin(x - \pi/2)$. 因而,球贝塞尔函数的渐近表达式是

$$j_l(x) \sim \frac{1}{x}\sin\left(x - \frac{l\pi}{2}\right) \qquad x \to \infty. \tag{3.14}$$

将(3.14)式代入(3.13)式得到平面波的渐近表达式

$$e^{ikz} \sim \sum_{l=0}^{\infty}(2l+1)(i^l/kr)\sin(kr - l\pi/2)P_l(\cos\theta) \qquad r \to \infty. \tag{3.15}$$

由(3.3)式递推得到的解(3.11)式是向外幅射和向内收缩的波叠加而成的驻波. 它满足波函数在 $r = 0$ 的有限性条件 §6.3.1(1.7)式,因而和德布罗意波 $e^{ik\cdot r}$ 一样,是在全空间中自由运动的粒子的一种可能的状态波函数. 但是,在散射问题中,粒子并不是在全空间自由运动,而只是在 $r > a$ 的区域中服从自由运动方程,其中 a 是散射势能 $U(r)$ 的作用半径. 显然,$r = 0$ 不在这一区域中. 因此,散射问题在 $r > a$ 的区域中的波函数不需要满足 $r = 0$ 的有限性条件 §6.3.1(1.7)式. 解除了这一条件的限制以后,自由运动方程的球面波解可以分别由(3.2)式中的两项递推得到.

定义第 1 类和第 2 类球汉克尔函数为(1 类取上号,2 类取下号)

$$h_l^{(1,2)}(x) = (-x)^l\left(\frac{1}{x}\frac{d}{dx}\right)^l \frac{\mp ie^{\pm ix}}{x}. \tag{3.16}$$

它们和球贝塞尔函数的关系是

$$j_l(x) = \frac{1}{2}[h_l^{(1)}(x) + h_l^{(2)}(x)]. \tag{3.17}$$

前几个球汉克尔函数是:

$$h_0^{(1,2)}(x) = \mp i\frac{e^{\pm ix}}{x},$$

$$h_1^{(1,2)}(x) = \left(\mp\frac{i}{x^2} - \frac{1}{x}\right)e^{\pm ix},$$

$$h_2^{(1,2)}(x) = \left(\mp\frac{3i}{x^3} - \frac{3}{x^2} \pm \frac{i}{x}\right)e^{\pm ix}.$$

球汉克尔函数的渐近表达式是

$$h_l^{(1,2)}(x) \sim \frac{1}{x}e^{\pm i\left(x - \frac{(l+1)\pi}{2}\right)}, \qquad x \to \infty. \tag{3.18}$$

球贝塞尔函数是两类球汉克尔函数的线性组合,见(3.18)式. 和它独立的另一种线性组合是球诺依曼函数

$$n_l(x) = \frac{1}{2i}[h_l^{(1)}(x) - h_l^{(2)}(x)]. \tag{3.19}$$

头几个球诺依曼函数是

$$n_0(x) = -\frac{\cos x}{x},$$

$$n_1(x) = -\frac{\cos x}{x^2} - \frac{\sin x}{x},$$

$$n_2(x) = -\left(\frac{3}{x^3} - \frac{1}{x}\right)\cos x - \frac{3\sin x}{x^2}.$$

球诺依曼函数的渐近表达式是

$$n_l(x) \sim -\frac{1}{x}\cos\left(x - \frac{l\pi}{2}\right), \qquad x \to \infty. \tag{3.20}$$

让我们再次强调，真正的自由粒子（在全空间中自由运动的粒子）的球面波函数只能是满足条件§6.3.1(1.7)式的球贝塞尔函数；而另一方面，只在 $r > a$ 区域中服从自由粒子运动方程的波函数，则可以是球贝塞尔函数和球诺依曼函数的线性组合（或两类球汉克尔函数的线性组合）.

（二）中心场散射的分波展开

现在回过头来考虑在中心对称场 $U(r)$ 中散射的问题. 假定在 $r \to \infty$ 时 $U(r) \to 0$ 足够快（比 r^{-1} 趋于零更快）. 用 a 表示场的作用半径，即假定在 $r > a$ 已经可以认为 $U(r) = 0$. 这样，在 $r > a$ 时径向方程§6.3.1(1.6)式成为自由运动方程[本节(3.1)式]. 它的一般解可以写成球贝塞尔函数和球诺依曼函数的组合：

$$R_l(kr) = A_l j_l(kr) + B_l n_l(kr), \qquad r > a. \tag{3.21}$$

注意，球诺依曼函数 $n_l(kr)$ 在 $r = 0$ 时发散. 但是(3.21)式只适用于 $r > a$ 的区域，其中不包括 $r = 0$，因而系数 A_l, B_l 都不为零. 它们可以通过解 r 有限时的方程§6.3.1(1.6)式，求所得解在 r 大时的行为而得到.

将球贝塞尔函数和球诺依曼函数的渐近表达式(3.14)、(3.20)代入(3.21)式得到径向波函数的渐近表达式

$$R_l(kr) \sim (kr)^{-1}[A_l\sin(kr - l\pi/2) - B_l\cos(kr - l\pi/2)], \qquad r \to \infty, \tag{3.22}$$

抽出一个因子 $A_l' = \sqrt{A_l^2 + B_l^2}$，就使得上式方括号内的表达式成为一个三角函数：

$$A_l\sin(kr - l\pi/2) - B_l\cos(kr - l\pi/2) = A_l'\sin(kr - l\pi/2 + \delta_l),$$

$$\cos\delta_l = A_l/A_l', \qquad \sin\delta_l = -B_l/A_l', \qquad \tan\delta_l = -B_l/A_l.$$

这样就有

$$R_l(kr) \underset{r\to\infty}{\to} (A'_l/kr)\sin(kr - l\pi/2 + \delta_l). \tag{3.23}$$

于是散射波函数的渐近形式为：

$$\psi(r,\theta) \underset{r\to\infty}{\to} \sum_l \frac{A'_l}{kr}\sin\left(kr - l\pi/2 + \delta_l\right)\mathrm{P}_l(\cos\theta). \tag{3.24}$$

在(3.23)、(3.24)式中引进了一个重要的量 δ_l. 将上式和入射波的渐近形式(3.15)式作比较可以看出，δ_l 是第 l 个分波受到势场作用后产生的相移，称为分波相移. 分波有了相移是可以理解的. 因为角动量守恒，处于角动量为 l 的状态的粒子，在受到势能场的作用后仍处于角动量为 l 的状态，但相应的分波的波函数可以相差一个位相因子.

(三)散射振幅和散射截面

将(3.15)式和(3.24)式代入散射问题的边界条件 §6.3.1(1.8)式可得

$$\begin{aligned} f(\theta,\varphi)\mathrm{e}^{\mathrm{i}kr}/r &= (\psi - \mathrm{e}^{\mathrm{i}kz})\big|_{r\to\infty} \\ &= \frac{1}{kr}\sum_l \left[A'_l\sin\left(kr - l\pi/2 + \delta_l\right) - (2l+1)\mathrm{i}^l\sin\left(kr - l\pi/2\right)\right]\mathrm{P}_l(\cos\theta) \\ &= \left[\mathrm{e}^{\mathrm{i}kr}/(2\mathrm{i}kr)\right]\sum_l \left[A'_l\mathrm{e}^{\mathrm{i}\left(-l\pi/2+\delta_l\right)} - (2l+1)\mathrm{i}^l\mathrm{e}^{-\mathrm{i}l\pi/2}\right]\mathrm{P}_l(\cos\theta) - \\ &\quad \left[\mathrm{e}^{-\mathrm{i}kr}/(2\mathrm{i}kr)\right]\sum_l \left[-A'_l\mathrm{e}^{-\mathrm{i}\left(-l\pi/2+\delta_l\right)} + (2l+1)\mathrm{i}^l\mathrm{e}^{\mathrm{i}l\pi/2}\right]\mathrm{P}_l(\cos\theta). \end{aligned} \tag{3.25}$$

上式右边的两项分别是向外发散和向内收缩的波. 由于所关心的是散射波，因而第二项应为零. 注意到各个 P_l 的线性无关性，第二项为零的条件是：

$$-A'_l\mathrm{e}^{-\mathrm{i}\left(-\frac{1}{2}l\pi+\delta_l\right)} + (2l+1)\mathrm{i}^l\mathrm{e}^{\mathrm{i}\frac{1}{2}l\pi} = 0.$$

这导致

$$A'_l = (2l+1)\mathrm{i}^l\mathrm{e}^{\mathrm{i}\delta_l}. \tag{3.26}$$

取 $\mathrm{i}^l = \mathrm{e}^{\mathrm{i}l\pi/2}$，将 A'_l 代回(3.25)式，就得到散射振幅：

$$f(\theta) = \frac{1}{k}\sum_l (2l+1)\mathrm{e}^{\mathrm{i}\delta_l}\sin\delta_l\mathrm{P}_l(\cos\theta). \tag{3.27}$$

进一步可得散射微分截面：

$$\sigma(\theta) = |f(\theta)|^2 = \frac{1}{k^2}\Big|\sum_l (2l+1)\mathrm{e}^{\mathrm{i}\delta_l}\sin\delta_l\mathrm{P}_l(\cos\theta)\Big|^2. \tag{3.28}$$

和总截面：

$$\sigma_t = \int |f(\theta)|^2\mathrm{d}\Omega = \sum_l \frac{4\pi}{k^2}(2l+1)\sin^2\delta_l \equiv \sum_l \sigma_l. \tag{3.29}$$

$\sigma_l = (4\pi/k^2)(2l+1)\sin^2\delta_l$ 是第 l 个分波的散射截面. 上式表明，总截面等于各个分波的散射截面之和.

(3.28)式和(3.29)式给出了散射截面和分波相移的关系. 对不同势能场，分波相移是不同的. 为了计算 δ_l，应首先在 $r < a$ 的区域中求解具体势能场中的径向方程，得到 R_l，然后再在 $r > a$ 的区域和解的一般渐近形式作比较，从而得到具体

的 δ_l .

习　题

1. 证明相互作用的二粒子薛定谔方程 §6.3.1(1.2)式可以化为描述质心运动和相对运动的两个方程，其中后一方程即 §6.3.1(1.3)式.

2. 将 §6.3.2(2.11)式中的 $+i\eta$ 改为 $-i\eta$ 可以得到该节方程(2.4)的另一个格林函数. 求这一格林函数并说明它的物理意义.

3. (a)将常微扰 $V(r)$ 看成频率 $\omega = 0$ 的周期性微扰，求跃迁概率.

(b)在入射粒子动能很大时，将粒子在势场 $V(r)$ 中的散射过程看成自由粒子平面波在微扰 $V(r)$ 作用下的跃迁过程，求微分散射截面.

4. 粒子受到势能为

$$U(r) = \begin{cases} -U_0, & r \leqslant a \\ 0, & r > a \end{cases}$$

的场散射，$U_0 > 0$，用玻恩近似计算微分散射截面.

5. 粒子受到势能 $U(r) = \alpha/r^2$ 的散射，用玻恩近似计算微分散射截面.

6. 设散射场的势能仍为上题中的 $U(r)$，试用分波法求相移 δ_l.

7. 证明：弹性散射的总截面等于 $(4\pi/k)$ 乘上向前散射振幅的虚部(光学定理)：

$$\sigma = (4\pi/k)\mathrm{Im}f(\theta = 0).$$



第七篇　宏观系统统计理论的基本原理

到目前为止，我们研究的对象是单个或少数粒子的运动及其与场的相互作用，被研究的系统只有少量自由度. 从本篇开始，我们将研究由大量粒子（如分子、原子、离子、电子、光子等）和准粒子（如声子、极化子、激子、旋子等）组成的，具有大量随机变化自由度的宏观系统. 由于系统的自由度数目非常大（通常有 10^{23} 的数量级）和自由度的随机性，即使我们透彻地掌握了单个粒子的运动规律和粒子间相互作用的规律，也不可能写出全部运动方程，更无法准确知道并利用全部初始条件求解运动方程. 要强调指出的是，不能用纯粹力学方法研究具有大量随机自由度的宏观系统，不仅是由于技术上的困难，更主要的是，由于大量随机自由度的存在，导致性质上的全新规律. 因此，研究这类系统的方法必须有本质上的改变.

研究具有大量随机自由度的宏观系统的方法有两种.其一是宏观理论,即热力学;其二是微观理论,即统计物理学.热力学以三个由观察和实验总结出的基本定律为基础,运用严密的数学推导,结合由实验测得的物质的状态方程和热容量,来研究系统宏观特性之间的关系,而不涉及物质的微观结构和组成系统的微观粒子的运动.热力学的优点是,它的结论和所依据的定律一样,具有普适性和可靠性.然而热力学也有明显的局限性,主要表现在,它不能揭示热力学基本定律及其结论的微观本质和不能解释涨落现象.统计物理从物质的微观结构和粒子所遵从的力学规律出发,运用概率统计的方法来研究宏观系统的性质和规律,包括涨落现象.统计物理的优点是,它可以深入问题的本质,使我们对热力学定律及其结论获得更深刻的认识.但统计物理中对物质微观结构所提出的模型只是实际情况的近似代表,因而理论预言和实验观测不可能完全一致,必须不断修正.统计物理和热力学是两种相辅相成的方法,在研究宏观系统时应该综合运用.从理论本身来说,它们也是密切相关的.

统计物理和热力学都包括平衡和非平衡两部分.第七篇至第九篇是平衡理论,第十篇是非平衡理论.第七篇介绍平衡理论的基本概念、基本思想、基本原理和基本方法,第八和第九篇是平衡理论的应用.

第一章　统计理论基础

对于平衡系统有多种统计方法.其中,系综法(见§7.1.4)不仅适用于粒子间相互作用可以忽略的简单系统,也适用于粒子间相互作用不能忽略的复杂系统.换句话说,系综法是一种普遍的统计理论.因此,我们在多种统计方法中选择了系综理论.

本篇第一章介绍系综理论的基本概念和基本假设;第二章在系综基本假设的基础上引入热力学的基本概念和基本规律,并说明它们的微观本质;第三章介绍有广泛应用的正则系综和巨正则系综,讨论系综的等价性并对应用系综理论的基本步骤和方法给出总结.

§7.1.1　宏观系统及其分类

由大量粒子或准粒子组成的系统称为宏观系统,与系统发生相互作用的环境称为外界.宏观系统中粒子数 N 的数量级为 10^{23}.为了能用数学工具对这种含巨大数目粒子的系统进行分析,通常采用热力学极限的概念,即认为:$N \to \infty$,$V \to \infty$,但 N/V 有限.

在热力学中,往往根据系统和外界的关系将系统分为孤立系统、封闭系统和开放系统.根据系统的化学组分分为单元系和多元系,根据系统的物理性质分为单相系和复相系.在统计物理中,除以上分类外,还进一步根据系统内部粒子间动力学相互作用的情况分为近独立子系和非近独立子系;根据粒子运动范围分为定域粒子系和非定域粒子系.

孤立系、封闭系及开放系:和外界既无物质交换又无能量交换的系统称为孤立系.能量交换包括力的(机械力、电力、磁力)和非力的(热的)两种方式.处于不变外场(如重力场)中的系统,虽然受力的作用,但能量不发生改变,也是孤立系.根据定义,我们无法对孤立系进行观测,因为任何观测必然和系统发生作用.因此,孤立系是一种理想情况.但它在热力学和统计物理中却是一个重要概念.和外界有能量交换但无物质交换的系统称为封闭系.如气缸中的气体,变化外场中的介质.和外界既有能量交换又有物质交换的系统称为开放系.如一个由水和水蒸气组成的系统,若把水和蒸气单独作为研究对象,它们可以交换能量和粒子,因而都是开放系.

单元系和多元系:"元"是指独立的化学成分.如水是单元系,盐水溶液和黄铜是二元系.

单相系和复相系：物理性质均匀的部分叫“相”. 物理性质包括力学、电学、光学和热学等各种性质. 单相系又称为均匀系. 盐水溶液是二元单相系. 水-汽是单元双相系. 石墨和金刚石都由碳原子组成，但具有完全不同的物理性质，是固态碳的不同相. 在 5.2K 以下，^{4}He 气体可以液化，并有正常流体和超流体两种不同的相. 许多金属及合金具有超导性，当温度降低到某一特定临界温度时，金属（或合金）会从正常导体变为超导体，它们是同一种金属固态的两种不同相.

近独立子系和非近独立子系：相互之间的动力学作用可以忽略不计的子系称为近独立，相应的系统称为近独立子系；否则称为非近独立，相应的系统称为非近独立子系. 近独立子系的总能量可表为 $U=\sum_{i=1}^{N}\varepsilon_i$，其中 ε_i 代表第 i 个子系的能量，包括它的动能和在外场中的势能. 相互作用可以忽略不计的条件是，相互作用能的总和远小于子系本身的能量总和.

定域粒子系和非定域粒子系：定域粒子系形成有规则的晶格结构，定域粒子只能在某格点附近的小范围内运动. 非定域粒子系不存在规则的晶格结构，非定域粒子可以在整个系统内运动. 从统计物理角度考虑，定域粒子和非定域粒子的区别在于，根据量子力学的全同性原理，属性（质量、电荷、自旋）相同的非定域粒子是无法识别的，而属性相同的定域粒子却可以根据其所处格点位置而加以识别. 当然，若非定域粒子服从经典力学，即使是全同粒子，也可以从粒子运动的不同轨道来识别.

高温下的气体、金属中的自由电子、黑体空腔中的光子都可近似为近独立、非定域粒子系. 低激发态对固体的低温性质起主要作用，研究表明，固体的低激发态可以看成是一些近独立、非定域的准粒子（声子）的集合. 近独立、非定域粒子或准粒子组成的系统统称为理想气体. 在 §5.5.1 中已经说明，为了满足微观粒子全同性的要求，近独立、非定域粒子系统的波函数必须具有对称或反对称性，即交换系统中任意一对粒子时，系统的波函数不变或仅改变代数符号. 据此，可进一步将理想气体分为两类：波函数具有对称性的理想气体称为玻色气体，如光子气和声子气；波函数具有反对称性的理想气体称为费米气体，如金属中的自由电子气.

晶体中的原子（或分子、离子）是定域粒子. 研究固体的顺磁性质时，往往忽略磁原子间的相互作用，即采用近独立、定域粒子系模型.

研究铁磁-顺磁相变时，则必须考虑磁原子间的相互作用，所采用的模型（Ising 模型，见 §9.3.1）是非近独立、定域粒子系模型.

§7.1.2　系统状态的宏观描述　平衡态　状态方程

(一)宏观量　宏观态

作为例子,以发动机气缸中汽油蒸气和空气的混合物作为研究的对象——系统.点火前后,系统的化学组分不同.活塞往复运动时,系统的体积、系统作用在活塞上的压强及系统的温度都在变化.为了掌握发动机的工作情况,我们不需要了解混合气体中单个分子的运动,但必须测量组分、体积、压强和温度.显然,化学组分、体积、压强及温度描述的不是系统中单个分子的运动状态,而是系统总体的物理性质.描述宏观系统总体物理性质的量称为宏观量或热力学量.宏观量的例子还有能量、密度、磁场强度、电场强度、磁矩、电极矩等.

从不同的角度考虑,往往把宏观量分为内参量和外参量,广延量和强度量.内参量表示系统内部的状态,由组成系统的粒子的热运动情况决定,如气体的压强、温度,磁介质的磁化强度等.外参量表示系统周围环境的状态,或者说,表示加于系统的外界条件,如体积、外电场、外磁场等.由于系统和外界的划分是人为的,所以内参量和外参量的划分也不是绝对的.广延量和强度量则是按宏观量与系统质量的关系来划分的.广延量与系统的质量成正比,而强度量则与系统的质量无关.如,粒子数 N、体积 V、内能 U、总磁矩 $\boldsymbol{M}$ 等是广延量;而粒子数密度 $n = N/V$,比容 $v = V/N$,比内能 $u = U/V$,磁化强度 $\boldsymbol{m} = \boldsymbol{M}/V$,压强 p,温度 T 等则是强度量.一个强度量经常和一个广延量同时出现,这种成对的或共轭的强度量和广延量有:T 和 S(熵),μ(化学势)和 N,p 和 V,$\boldsymbol{H}$ 和 $\boldsymbol{M}$,f(张力)和 L(长度)等.

由宏观量描述的系统的状态叫做宏观态或热力学状态.

在平衡态热力学和统计物理中,我们讨论系统的一种特殊的宏观态,即平衡态.

(二)平衡态　态参量和态函数

经验表明,一个孤立系统,若初始宏观量在系统各部分具有不同的值(如气体的密度、压强和温度不均匀),则系统的宏观态会随时间变化,经过足够长的时间后,将会达到一个宏观性质不再随时间变化的状态.孤立系达到的这种不再随时间变化的宏观态叫做平衡态,否则叫做非平衡态.

在理解平衡态的概念时,应当注意:

①一旦孤立系达到平衡态,则描述系统总体性质的宏观量不再随时间变化,系统内部自然也不再有宏观物理过程,如热传导、对流、扩散及化学反应等.需要特别注意的是,不要把平衡态和稳定态混为一谈.设想将铜棒一端置于沸水中,另一端置于冰水混合物中,经过一段时间后,铜棒各处温度分布不再随时间变化.我们称

铜棒处于稳定态，但不是平衡态. 因为铜棒不是孤立系，铜棒内部存在着热传导过程.

②在平衡态下，虽然系统的宏观性质不再随时间变化，但组成系统的大量粒子却仍处在不停的无规则运动之中，只是这些粒子运动的平均效果不变而已. 因此，平衡态是一种宏观上的动态平衡，常称为热动平衡. 其次，在平衡态下，宏观量的数值仍会发生或大或小的涨落，在适当的条件下可以观察到这种涨落. 不过对于宏观系统，在一般情况下，涨落极其微小而可以忽略. 在热力学中不考虑涨落，认为在平衡态下宏观量有确定值.

③孤立系从开始的非平衡态到达平衡态所需的时间称为弛豫时间. 弛豫时间的长短不仅依赖于系统的大小，还和弛豫过程的性质或机制有关. 例如，同一系统中压强的均匀化比浓度均匀化要快，而电、磁极化的均匀化极快. 一般应取系统中不同过程的最长弛豫时间作为系统的弛豫时间 τ. 当对系统的观察所经历的时间 Δt 远大于系统的弛豫时间 τ 时，有理由相信系统已达到平衡态.

④平衡态的概念不限于孤立系. 封闭系、开放系可以看成是孤立系的一部分，另一部分是外界. 当孤立系处于平衡态时，作为其中一部分的封闭系或开放系当然也处于平衡态.

⑤当系统达到平衡态后，不仅描述其总体性质的各种宏观量不再随时间变化，而且这些宏观量之间存在一定的联系. 从数学上说，即平衡态的各个宏观量之间存在一定的函数关系，因而它们并非都能独立改变. 换句话说，平衡态具有非常简单的性质，只需少数必要且充分的独立宏观量就可以完全确定平衡态的性质. 我们把被选定的独立变化宏观量叫做态参量，其他宏观量则可表示为态参量的函数叫做态函数.

一个系统究竟需要几个态参量（即所需态参量的数目）必须根据所研究的问题由经验决定；选择哪几个宏观量作为态参量则是任意的，可以根据研究的方便决定. 这一思想在热力学方法中有重要的指导意义. 如对于单相封闭流体系，在和外界只交换膨胀功的情况下，态参量数为 2，可选（T,V）也可选（T,p）或（p,V）为态参量. 当研究固体的压电效应或磁致伸缩现象时，除膨胀功外系统和外界还有极化功或磁化功交换，态参量数为 3. 除（T,V）或（T,p）外，还须引入电场强度或极化强度，磁场强度或磁化强度为态参量. 对于可以发生化学反应的多元系，还必须选各组元的物质的量为态参量.

（三）状态方程

除温度 T 以外的态参量和 T 的函数关系叫做状态方程. 用 $x_1,x_2,\cdots,x_n$ 表示态参量，则状态方程的一般形式为

$$\mathscr{F}(T,x_1,x_2,\cdots,x_n)=0. \tag{2.1}$$

对于复相系，各相有自己的状态方程. 下面是一些常见的状态方程实例.

1. 气体

考虑 1mol 气体，则有理想气体的状态方程：

$$pv = RT. \tag{2.2}$$

考虑分子有一定大小及分子间有相互作用后，对(2.2)式的一种修正形式是：范德瓦耳斯方程：

$$(p + a/v^2)(v - b) = RT. \tag{2.3}$$

a 和 b 是两个常数，随气体性质面异. 在实验中，测出若干条 $T > T_C$（临界温度）下的状态方程线，就可以反解出 a,b 的值. 实际上，(2.2)式和(2.3)式可以看成是 p 按 v^{-1} 的幂级数展开式. 非理想气体状态方程更一般的形式是位力展开，即卡末林·昂内斯方程：

$$pv = RT(1 + B_2v^{-1} + B_3v^{-2} + \cdots), \tag{2.4}$$

$B_i(T)$ 叫第 i 级位力系数. 它是温度的函数，并与气体的性质有关.

2. 拉紧的弦

拉紧的弦的实验往往在大气压下进行，且体积变化可以忽略，因此可取张力 f 和长度 L 为态参量. 在弹性限度内胡克定律成立，故状态方程为：

$$f = k(t)(L - L_0). \tag{2.5}$$

L_0 是 $f = 0$ 时的长度，$k(t)$ 是与温度有关的劲度系数.

3. 表面膜

表面膜的研究是固体物理和化学物理中的重要分支. 表面是指两相分界处很薄的一层，由于层内分子与体内分子所处的环境不同而具有特殊的性质. 通常把表面层理想化为一几何面，取表面张力 σ 和表面面积 A 为态参量. 实验证明，纯液体与其蒸气达到平衡时，表面张力与表面面积的大小无关，仅仅是温度的函数. 状态方程是

$$\sigma = \sigma_0(1 - t/t')^n. \tag{2.6}$$

σ_0 是 0℃时的表面张力，t' 和 n 均为实验常数，前者与临界温度相差几度，后者的数值在 1 和 2 之间.

4. 顺磁固体

大多数顺磁固体的实验在大气压下进行，且只有很小的体积变化，因而可取磁场强度 $\boldsymbol{H}$ 和总磁矩 $\boldsymbol{M}$ 为态参量. 实验证明，在高温弱场时，顺磁固体的状态方程是

$$\boldsymbol{M} = c\boldsymbol{H}/T, \tag{2.7}$$

c 称为居里常数. 统计物理对顺磁质有特别兴趣，因为它是一个最便于处理的系统.

要指出的是，如果系统的状态不仅与态参量的瞬时值，而且还与其历史有关，

则不能给出状态方程.例如铁磁材料(图1)和发生范性形变的固体(图2).前者的磁化强度不是磁场强度的单值函数,后者的应力不是应变的单值函数.这种情况统称为滞后效应.

在热力学中,状态方程只能由实验确定,而不能由热力学理论导出.推导状态方程是统计物理的重要课题之一.在应用热力学解决实际问题时,经常要计算物质的某些热力学量,如熵 S 和内能 U 等;或研究某些热力学量的关系,如表示内能和体积关系的偏导数 $(\partial U/\partial V)_T$,这些量往往不便于直接测量.后而将看到,利用热力学规律可以把它们与状态方程及易于测量的热容量联系起来.因此状态方程和热容量是基本实验数据,在热力学中是很重要的.

下而介绍与状态方程密切联系的几个系数.知道了状态方程可以求这几个系数;反之,由实验测得这几个系数可以导出状态方程.对于液体和固体,常常用测量这几个系数来代替测状态方程.这些系数分别是等压膨胀系数 α,等容压强系数 β 和等温压缩系数 κ_T:

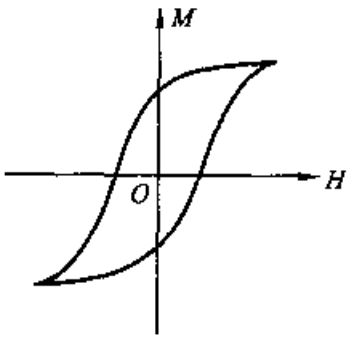


图1　磁滞回线

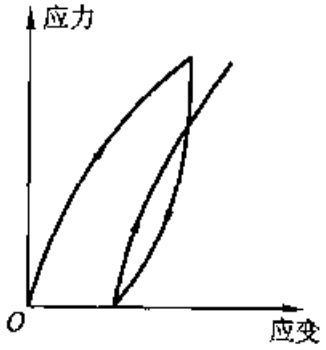


图2　范性形变

$$\alpha \equiv \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_p, \quad \beta \equiv \frac{1}{p}\left(\frac{\partial p}{\partial T}\right)_V, \quad \kappa_T \equiv -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T. \tag{2.8}$$

这三个系数并非彼此独立,测出其中两个就可算出第三个.把 V 作为 p,T 的函数,$V = V(p,T)$,可写出 $\mathrm{d}V = (\partial V/\partial T)_p \mathrm{d}T + (\partial V/\partial p)_T \mathrm{d}p$,令 $\mathrm{d}V = 0$,得

$$\left(\frac{\partial V}{\partial p}\right)_T\left(\frac{\partial p}{\partial T}\right)_V\left(\frac{\partial T}{\partial V}\right)_p = -1. \tag{2.9}$$

将(2.8)式代入得到

$$\alpha = p\beta\kappa_T. \tag{2.10}$$

对于固体和液体,升高温度时要维持体积不变相当困难,利用(2.10)式可通过测量 α 和 κ_T 来求得 β.以上方法不难推广到和外界交换其他功(如磁化功)的系统,并得到和 α,β,κ_T 相对应的系数.

【例1】　实验发现,对一橡皮带有

$$\left(\frac{\partial f}{\partial L}\right)_T = \frac{aT}{L_0}\left[1 + 2\left(\frac{L_0}{L}\right)^3\right], \qquad \left(\frac{\partial f}{\partial T}\right)_L = \frac{aL}{L_0}\left[1 - \left(\frac{L_0}{L}\right)^3\right]. \tag{2.11}$$

式中 f 是张力,a 是一常数,L_0 是无张力时的长度.(a)求 $(\partial L/\partial T)_f$;(b)求状态方

程.

解　(a)由状态方程 $\mathscr{F}(f,L,T)=0$ 得 $\left(\frac{\partial f}{\partial L}\right)_T\left(\frac{\partial L}{\partial T}\right)_f\left(\frac{\partial T}{\partial f}\right)_L=-1$，故

$$\left(\frac{\partial L}{\partial T}\right)_f=-\frac{(\partial f/\partial T)_L}{(\partial f/\partial L)_T}=-\frac{(aL/L_0)[1-(L_0/L)^3]}{(aT/L_0)[1+2(L_0/L)^3]},$$

$$\alpha\equiv\frac{1}{L}\left(\frac{\partial L}{\partial T}\right)_f=-\frac{1}{T}\frac{1-(L_0/L)^3}{1+2(L_0/L)^3}. \tag{2.12}$$

α 是线膨胀系数.

(b)

$$f(L,T)=\int_{L_0}^{L}\left(\frac{\partial f}{\partial L}\right)_T\mathrm{d}L+c(T)=aT\frac{L}{L_0}(1-L_0^3/L^3)+c(T),$$

$$(\partial f/\partial T)_L=a(L/L_0)(1-L_0^3/L^3)+c'(T).$$

与已知条件(2.11)比较,可知 $c'(T)=0$ 或 $c=$ 常数. 于是我们有:

$$f=aT(L/L_0)(1-L_0^3/L^3)+c.$$

因为 $L=L_0$ 时,$f=0$,所以 $c=0$,故橡皮带的状态方程为

$$f=aT(L/L_0)(1-L_0^3/L^3). \tag{2.13}$$

§7.1.3　系统状态的微观描述　统计规律性

统计物理的根本任务是,从宏观系统的微观结构模型和粒子遵从的力学规律出发,经过演绎,从理论上推导出宏观系统的性质和规律. 例如,从气体的微观模型和气体分子服从的力学规律导出气体的状态方程和气体热容量对温度的依赖关系. 因此,任何统计方法最先遇到的问题都是:回答如何在分子水平上从力学观点来描述系统;研究系统的宏观性质和微观粒子的力学行为有怎样的关系. 本节回答前一个问题,而后一个问题:系统宏观性质和粒子的力学行为的关系,则留到下一节讨论.

(一)微观态

通过考查系统中每个粒子的力学状态而确定的宏观系统的状态叫做系统的动力学状态,又叫做系统的微观态.

从本质上说,微观粒子遵从量子力学规律,因而宏观系统的微观态应该用量子力学方法描述. 但在某些情况下(如粒子质量较大,数密度较稀和温度较高的非定域粒子系),也可近似认为微观粒子遵从经典力学规律,从而可采用经典力学方法描述系统的微观态.

(二)微观态的量子力学描述

按照量子力学,和单粒子一样,宏观系统的动力学状态(或微观态)由特定表象如坐标表象或动量表象中的波函数描述;或更一般地说,由希尔伯特空间中的态矢量 $|\psi(t)\rangle$ 描述. 态矢量 $|\psi(t)\rangle$ 随时间的变化遵从薛定谔方程:

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle. \tag{3.1}$$

$\hat{H}$ 是宏观系统的哈密顿算符. 对于哈密顿不随时间变化的保守系统(统计物理研究的对象正是这种系统),可以对(3.1)式分离变量得到

$$\hat{H}|\psi_\nu\rangle = E_\nu|\psi_\nu\rangle. \tag{3.2}$$

给定 $\hat{H}$ 和边界条件,通过解方程(3.2)可以得到一系列能量本征值 E_ν 和能量本征态 $|\psi_\nu\rangle$. 每一个 $|\psi_\nu\rangle$ 就是宏观系统的一个量子动力学状态或微观态.

和单粒子不同的是,为确定宏观系统的一个量子动力学状态 $|\psi_\nu\rangle$,需要知道数量级为 $N\approx10^{23}$ 个量子数,因而宏观系统微观态的描述十分复杂. 值得庆幸的是,在统计物理中,我们并不需要知道微观态的详细知识. 归根结底,关心的仅仅是和给定宏观态(平衡态)对应的可达微观态数.

为了说明微观态以及微观态和宏观态的关系,考虑三个例子. 前两个虽不实际,但简单且对于了解统计物理方法有意义.

【例 1】 近独立定域谐振子系的微观态.

系统甲由 $N=3$ 个近独立,位于晶格格点上,具有相同圆频率 ω 的量子谐振子组成(在本篇第三章中将看到,这实际上是简化了的关于固体比热的爱因斯坦模型). 假定系统是孤立的,而且总能量有确定值 $E=9\hbar\omega/2$. 这一孤立系统的平衡态由 $(N=3, E=9\hbar\omega/2, V)$ 描述.

我们来考查如何描述这个系统的微观态. 由于忽略了谐振子间的相互作用,(3.2)式中的 $\hat{H}$ 可表为

$$\hat{H}=\sum_{k=1}^{3}\hat{H}(k), \tag{3.3}$$

$\hat{H}(k)$ 代表第 k 个量子谐振子的哈密顿算符. 在坐标表象中有

$$\hat{H}(k)=-[\hbar^2/(2\mu)]\mathrm{d}^2/\mathrm{d}x_k^2+(\mu\omega^2/2)x_k^2. \tag{3.4}$$

μ 和 ω 分别代表谐振子的约化质量和圆频率. 当把固体看成由近独立量子谐振子组成时,可以证明 ω 是固体体积 V 的函数:$\omega=f(V)$.

通过解单个谐振子的定态薛定谔方程

$$\hat{H}(k)\varphi_n(k)=\varepsilon_n(k)\varphi_n(k). \tag{3.5}$$

得到谐振子的能级公式[见 §6.1.1(1.21)式]

$$\varepsilon_n(k)=(n+1/2)\hbar\omega,\qquad n=0,1,2,\cdots, \tag{3.6}$$

以及和前几个能级对应的定态波函数 $\varphi_n(k)$，见 §6.1.1(1.18)、(1.22)、(1.23)式：

$$\begin{aligned} &n = 0 \quad \varphi_0(k) \approx (\alpha/\pi)^{1/4} e^{-\alpha x^2/2}, \\ &n = 1 \quad \varphi_1(k) = (4\alpha^3/\pi)^{1/4} x e^{-\alpha x^2/2}, \\ &n = 2 \quad \varphi_2(k) \approx (\alpha/4\pi)^{1/4} (2\alpha x^2 - 1) e^{-\alpha x^2/2}, \\ &\cdots \quad \cdots. \end{aligned} \tag{3.7}$$

其中，$\alpha = 2\pi\mu\omega/h$，见 §6.1.1(1.2)式. 注意，这里的 α 是那里的 α^2.

根据(3.3)式，系统 A 的能量值为

$$E = \varepsilon_{n_1}(1) + \varepsilon_{n_2}(2) + \varepsilon_{n_3}(3). \tag{3.8}$$

由于谐振子是定域粒子而可以识别，系统甲的波函数等于每个谐振子定态波函数之积

$$\psi_\nu = \varphi_{n_1}(1)\varphi_{n_2}(2)\varphi_{n_3}(3). \tag{3.9}$$

换句话说，系统的波函数 ψ_ν 或微观态 ν 由一套量子数 $\{n_1, n_2, n_3\}$ 决定.

表 I　近独立定域谐振子系统的微观态

波函数	能　级	能量分配方式 X		
		a	b	c
φ_4	$\varepsilon_4 = (9/2)\hbar\omega$			
φ_3	$\varepsilon_3 = (7/2)\hbar\omega$		③	
φ_2	$\varepsilon_2 = (5/2)\hbar\omega$			③
φ_1	$\varepsilon_1 = (3/2)\hbar\omega$	① ② ③		②
φ_0	$\varepsilon_0 = (1/2)\hbar\omega$		① ②	①
$\psi_\nu = \prod_k \varphi_{n_k}(k)$	$E = \sum_k \varepsilon_{n_k}(k)$	$\varphi_1(1)\varphi_1(2)\varphi_1(3)$	$\varphi_0(1)\varphi_0(2)\varphi_3(3)$ $\varphi_0(1)\varphi_3(2)\varphi_0(3)$ $\varphi_3(1)\varphi_0(2)\varphi_0(3)$	$\varphi_0(1)\varphi_1(2)\varphi_2(3)$ $\varphi_1(1)\varphi_0(2)\varphi_2(3)$ ……
微观态数 t_X		1	3	6
$\Omega = \sum_X t_X$		10		

表 I 中 X 代表能量分配方式，t_X 代表和每一种能量分配方式对应的波函数的个数，因而 $\Omega(N,E,V) = \sum_X t_X$ 代表和宏观态($N = 3, E = 9\hbar\omega/2, V$)相对应的总微观态数.

在粒子数 $N = 3$ 和总能量 $E = 9\hbar\omega/2$ 不变的条件下，总能量在三个粒子中有

三种不同的分配方式. a 方式：三个粒子各有 $3\hbar\omega/2$ 的能量；b 方式：两个粒子有 $\hbar\omega/2$ 的能量，另一粒子有 $7\hbar\omega/2$ 的能量. 究竟哪个粒子有 $7\hbar\omega/2$ 又有三种可能；c 方式：一个粒子有 $5\hbar\omega/2$ 的能量，一个粒子有 $3\hbar\omega/2$ 的能量和一个粒子有 $\hbar\omega/2$ 的能量. 究竟哪个粒子有 $5\hbar\omega/2$ 的能量，哪个粒子有 $3\hbar\omega/2$ 的能量，哪个粒子有 $\hbar\omega/2$ 的能量又有六种可能. 由以上分析可知，系统甲可以有表 I 所示的十个不同的波函数 ψ_ν，每一个 ψ_ν 代表系统甲的一个量子动力学状态或微观态，但它们都对应着系统甲的同一个宏观态（$N = 3, E = 9\hbar\omega/2, V$）.

【例 2】 近独立定域自旋粒子系的微观态.

系统乙由四个位于晶格格点，自旋为 1/2，磁矩为 μ_0 的磁偶极子组成，且处于不变的外磁场 $\boldsymbol{B}$ 中. 我们将不考虑磁偶极子在格点的振动并忽略所有磁偶极子间的相互作用，只考虑磁偶极子和外场的作用. 由于自旋为 1/2，每个磁偶极子的磁矩只能平行或反平行于外场，具有 $\mp\mu_0 B$ 的能量.

表 II 近独立定域自旋粒子系的微观态

l	n_1	n_2	n_3	n_4	M	E
1	+	+	+	+	$4\mu_0$	$-4\mu_0 B$
2	+	+	+	−	$2\mu_0$	$-2\mu_0 B$
3	+	+	−	+	$2\mu_0$	$-2\mu_0 B$
4	+	−	+	+	$2\mu_0$	$-2\mu_0 B$
5	−	+	+	+	$2\mu_0$	$-2\mu_0 B$
6	+	+	−	−	0	0
7	+	−	+	−	0	0
8	+	−	−	+	0	0
9	−	+	+	−	0	0
10	−	+	−	+	0	0
11	−	−	+	+	0	0
12	+	−	−	−	$-2\mu_0$	$2\mu_0 B$
13	−	+	−	−	$-2\mu_0$	$2\mu_0 B$
14	−	−	+	−	$-2\mu_0$	$2\mu_0 B$
15	−	−	−	+	$-2\mu_0$	$2\mu_0 B$
16	−	−	−	−	$-4\mu_0$	$4\mu_0 B$

表Ⅱ列举了这一系统的各种微观态. 表中，$n_1 \sim n_4$ 分别代表四个磁偶极子的自旋量子数. 为简单起见，用"+"表示 $n=1$，即磁矩与外场平行，用"-"代表 $n=-1$，即磁矩与外场反平行. E 和 M 代表系统的总能量和总磁矩. 系统乙的微观态由一套量子数 $\{n_1,n_2,n_3,n_4\}$ 确定，l 代表微观态的编号.

实际上，这是一个简化的顺磁固体模型. 由于这一模型可以明显地精确处理，因此在后面我们将多次利用它来引入统计物理中的概念和方法.

【例 3】 近独立非定域粒子系的微观态.

系统丙由 N 个全同的无相互作用的粒子组成，粒子在边长为 L 的立方盒中自由运动，假设系统是孤立的，其平衡态由 (N,E,V) 描述，实际上这是理想气体模型.

为讨论方便，我们先考虑一个自由粒子，$\varepsilon=(p_x^2+p_y^2+p_z^2)/(2m)$. 在周期性边界条件下，动量的本征值为[见 §5.1.2(2.20)式]

$$p_x=hn_x/L,\ p_y=hn_y/L,\ p_z=hn_z/L,\ n_x,n_y,n_z=0,\pm1,\cdots. \tag{3.10}$$

则

$$\varepsilon=[h^2/(2mL^2)](n_x^2+n_y^2+n_z^2). \tag{3.11}$$

以量子数 n_x,n_y 和 n_z 为三个垂直坐标轴构成"数空间"，这三个量子数的许可值用图表示就落在边长为 1 的各个立方体的中心，如图 1 所示.

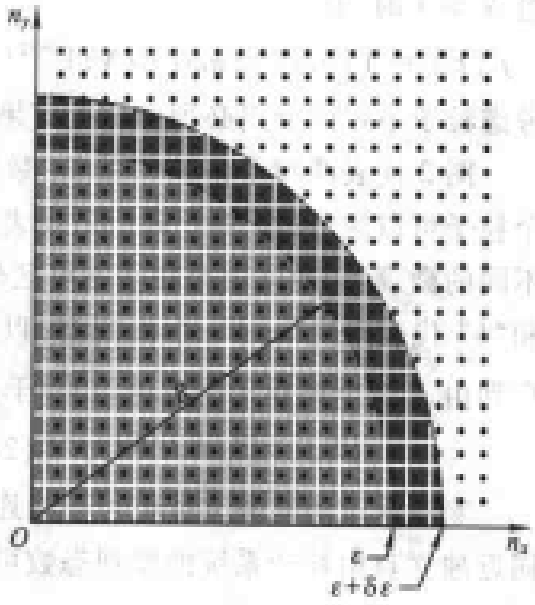


图 1 图上的点表示确定三维单粒子态的量子数的各种可能值 $n_x,n_y,n_z=\pm1,\pm2,\pm3,\pm4,\cdots$（$n_z$ 轴离开纸面朝外）. 相应于能量 ε 和 $\varepsilon+\delta\varepsilon$ 的 n_x,n_y,n_z 值位于被指明的球面上. 小圆弧内的区域包含粒子能量小于 ε 的所有 n 值. 大小圆弧之间的区域包含能量处于 ε 和 $\varepsilon+\delta\varepsilon$ 之间的所有 n 值.

当 L 有宏观大小，而 ε 取有意义的值（如 10^{-16}erg）时，量子数很大. 由(3.11)式 $n_x^2+n_y^2+n_z^2=(L/h)^2 2m\varepsilon=(R')^2$. 对于给定的 ε 值，满足这一方程的所有的点 (n_x,n_y,n_z) 落在半径为 R' 的球面上. 当粒子的能量小于 ε 时，粒子的力学状态由图中小圆弧内的"点"表示，力学状态数则等于半径为 $R=\sqrt{2m\varepsilon}$ 的球体积被 $(h/L)^3$ 除，即

$$f(\varepsilon)=(4\pi/3)(2m\varepsilon)^{3/2}(V/h^3). \tag{3.12}$$

粒子的能量在 $\varepsilon\sim\varepsilon+\delta\varepsilon$ 之间时，对应的单粒子力学状态则由图中大小圆弧之间的

区域的“点”代表.

再研究由这样的 N 个自由粒子组成的理想气体,理想气体的能量为

$$E = [1/(2m)]\sum_{i=1}^{N}(p_{xi}^2 + p_{yi}^2 + p_{zi}^2) = [h^2/(2mL^2)]\sum_{i=1}^{N}(n_{xi}^2 + n_{yi}^2 + n_{zi}^2).$$

我们暂时假设这 N 个粒子是可区分的,于是在由 $(n_{x1}^2, n_{y1}^2, n_{z1}^2, \cdots, n_{xN}^2, n_{yN}^2, n_{zN}^2)$ 这 $3N$ 个量子数为垂直坐标轴构成的 $3N$ 维“数空间”中,一个“整数点”代表理想气体系统的一个微观态. 在此 $3N$ 维空间中,半径为 $R = \sqrt{2mE}$ 的“球”体积为(本节稍后将给出证明)

$$\frac{\pi^{3N/2}}{\Gamma(3N/2+1)}R^{3N} = \frac{\pi^{3N/2}}{\Gamma(3N/2+1)}(2mE)^{3N/2}.$$

$\Gamma(\alpha)$ 是 Γ 函数(见附录 V). 系统的能量小于 E 时对应的微观态数为[对照(3.12)式]

$$j'(E) = \frac{\pi^{3N/2}}{\Gamma(3N/2+1)}\frac{(2mE)^{3N/2}}{(h/L)^{3N}}.$$

当 $N \gg 1$ 时,有

$$j'(E) \approx [(2e/3)2\pi m(E/N)]^{3N/2}(V^N/h^{3N}).$$

考虑粒子不可区分则必须对以上结果进行修正.

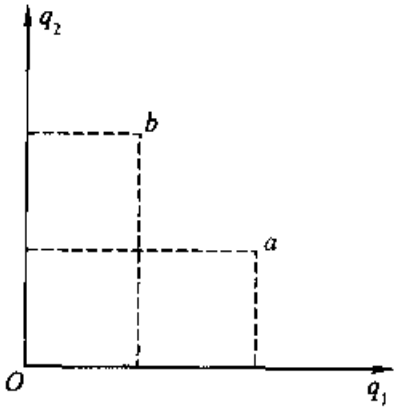


图 2　二粒子系统的状态

图 2 代表两粒子系统的量子“数空间”. 当两个粒子可以区分时,“a”和“b”代表系统的两个不同的微观态. 而当两个粒子不可区分时,则“a”和“b”代表系统的同一个微观态. 以上分析可推广到由 N 个全同粒子组成的系统. 于是有

$$j(E) = j'(E)/N! \approx [(2e/3)2\pi m(E/N)]^{3N/2}(V^N/N!h^{3N}). \qquad (3.13)$$

宏观系统的能量本征值可视为连续变量,在能量间隔 $E \sim E + \delta E$ 间,N 个全同近独立自由粒子系统的微观态数可表为

$$\Omega(E)\delta E = \frac{\partial j(E)}{\partial E}\delta E$$

$$= \frac{1}{N!h^{3N}}\left(\frac{2e}{3}2\pi m\right)^{3N/2}\frac{3}{2}\left(\frac{E}{N}\right)^{(3N/2-1)}V^N\delta E \propto E^{3N/2}\delta E. \qquad (3.14)$$

从以上例子可以清楚地看出以下几点:

①用量子力学方法描述时,系统的微观态由一套量子数确定,量子数的数目和宏观系统中的粒子数 N 有相同数量级.

②微观态和宏观态的关系是:一个微观态只能对应于一个宏观态,而一个宏观态往往对应着许多可能达到的微观态. 例如孤立系统甲和宏观态($N = 3, E =$

$9\hbar\omega/2, V$) 对应有如表 I 所示的 10 个可能达到的微观态. 若理想气体系统丙是孤立的,其宏观态用 $(N, V, E \sim E + \delta E)$ 描述,对应的可达微观态数如(3.14)式. 若系统乙是孤立的,且能量为 $2\mu_0 B$,其宏观态用($N = 4, E = 2\mu_0 B, M = -2\mu_0$) 表示,对应的可达微观态数是 4. 若系统乙和温度为 T 的大热源接触且处于平衡,其总能量和总磁矩可取表 II 中的任意值,这时系统乙封闭而非孤立,其宏观态用 ($N = 3, T, B$) 描述,对应的可达微观态数是 16.

③通常用 $\Omega(N, E, V)\delta E$ 表示系统的能量值介于 E 和 $E + \delta E$ 之间的可达微观态数,叫做宏观态 (N, E, V) 的统计权重. 我们把能量值无上限的系统叫做正常系统,如例 1 中的系统甲和例 3 中的系统丙. 从(3.14)式可以看到,对于正常系统有

$$\Omega(N, V, E)\delta E \propto E^N \delta E. \tag{3.15}$$

即粒子数 N 愈多,能量值 E 愈高,则统计权重愈大. $\Omega(N, V, E)$ 代表在 E 附近,单位能量间隔中的可达微观态数,叫做态密度. 因为 $N \approx 10^{23}$,所以 $\Omega(E)$ 是一个随 E 增大而迅速增大的天文数值.

④实际上没有完全不和外界发生相互作用的孤立系,也不存在粒子间完全不发生相互作用的独立粒子,因此,已知初始时刻处于某一微观态的孤立系,绝不会永远停留在这个初始微观态. 随着时间的推移,在外界作用和粒子间相互作用下,孤立系将会跃迁到宏观条件 (N, V, E 一定)所允许的其他微观态. 例如,若例 2 中的孤立系乙开始处于微观态 $\{+, +, -, -\}$,在以后的某一时刻,它会以不等于零的概率出现于其他微观态,如 $\{+, -, +, -\}$.

当然,非孤立系更会在宏观条件(例如 N, T, V 一定)允许下的微观态之间不断转移.

⑤系统的能量是外参量的函数,能量值会因外参量的改变而改变. 例如,系统乙的能量和外场强度 B 有关. 气体的能量和气体的体积 V 有关. 系统甲的能量和谐振子的圆频率 ω 有关,而 ω 和系统的体积 V 有关. V 和 B 都是外参量,它们表示加子系统的外界条件,改变 V 和 B 会使系统甲、乙和丙的能量发生改变. 系统的能量因外参量变化而变化的过程叫做作功.

(三)微观态的经典力学描述　Γ 空间

按照经典力学,粒子的运动状态可用其广义坐标矢量 $\boldsymbol{q}$ 和共轭广义动量矢量 $\boldsymbol{p}$ 描述. 用 r 表示单个粒子的自由度数,则 $\boldsymbol{q}$ 和 $\boldsymbol{p}$ 都是 r 维矢量. 于是,指明了宏观系统中每个粒子的广义坐标矢量和共轭广义动量矢量,就完全和唯一地确定了宏观系统的一个经典动力学状态. 换句话说,用经典力学描述时,宏观系统的微观态由以下一组量决定:

$$\{\boldsymbol{q}_1, \boldsymbol{p}_1; \boldsymbol{q}_2, \boldsymbol{p}_2; \cdots; \boldsymbol{q}_N, \boldsymbol{p}_N\}. \tag{3.16}$$

这 $2rN$ 个变量构成描述宏观系统的广义坐标和正则共轭广义动量. 由第一篇知道,

当用广义坐标和广义动量描述系统的状态时，力学系统的运动遵从正则方程. 对于保守力学系统，正则方程中的哈密顿函数就是系统的能量. 统计物理中研究的宏观系统由大量处于保守力场中的粒子组成，系统的哈密顿可表为

$$H = \sum_{i=1}^{N} H_i + U(q_1, q_2, \cdots, q_N). \tag{3.17}$$

其中 H_i 是第 i 个粒子在外场中的哈密顿，而 $U(q_1, q_2, \cdots, q_N)$ 是粒子间的相互作用能. 由(3.17)式可得

$$\partial H/\partial p_i = \dot{q}_i, \qquad \partial H/\partial q_i = -\dot{p}_i, \qquad i = 1, 2, \cdots, N. \tag{3.18}$$

(3.18)式是经典宏观系统的微观态随时间变化的正则方程. 它是一个庞大的方程组. 给定初始条件，积分正则方程组可得 $\{q_1(t), p_1(t); q_2(t), p_2(t); \cdots; q_N(t), p_N(t)\}$，即系统的微观态将随时间连续变化.

为了形象化地描述经典宏观系统所有可能的微观态，及微观态随时间的变化，想象中用(3.16)式中的总共 $S = 2rN$ 个广义坐标和共轭广义动量支起一个 S 维的笛卡儿坐标系，叫做经典宏观系统的相空间或 Γ 空间，r 是单个粒子的自由度数. Γ 空间中的每一个点确定了全部粒子的广义坐标和广义动量，代表了经典宏观系统的一个微观态. 反之，经典宏观系统的任何一个微观态也都能在 Γ 空间中找到对应的点，叫做系统的微观态代表点，如图 3 所示.

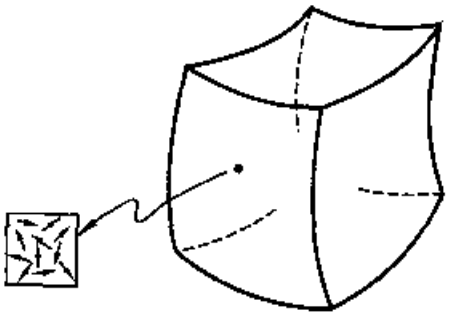


图 3　Γ 空间及微观态代表点

随着时间的流逝，经典宏观系统的微观态连续变化，其微观态代表点在 Γ 空间中画出一条线，称为相轨迹，参看 §1.6.1. 下面说明相轨迹的一些性质. 首先，系统从不同的初态出发，其代表点将画出不同的相轨迹，这些相轨迹互不相交. 这是因为哈密顿是广义坐标和广义动量的单值函数，所以正则方程的解是唯一的. 从一定的初态出发，代表点只能画出一条相轨迹. 如果两条轨迹相交于一点，则从该点出发，将有两条相轨迹，这是不可能的. 其次，相轨迹必需保持在 Γ 空间中的一个有限区域内，这是因为系统的体积直接限制了广义坐标 q 的取值，而有限的能量则限制了 q 和 p 的取值. 特别是，如果系统的能量具有精确值 E，则相轨迹将被限制在一个"多维等能面"上：

$$H(q_1, p_1; q_2, p_2; \cdots; q_N, p_N) = E. \tag{3.19}$$

若总能量处于能量间隔 $E \sim E + \delta E$ 范围内，则相轨迹被约束在一个能量"壳层"中.

在一定条件下，可以近似认为宏观系统中的微观粒子遵循经典力学，但由此得

出的结论不得与用量子力学得到的结论发生原则上的抵触.根据这一考虑,必须对 Γ 空间的某些细节进行修改.

按照量子力学,任何一个宏观系统,尽管其微观状态数 Ω 是一个天文大数,但只要 (N,E,V) 给定,则 Ω 终归会有一个有限的明确值.然而,若用 Γ 空间来描述孤立系统的微观态,则在相应的能量"壳层"中,微观态代表点的数目无穷多,即 $\Omega(E)\delta E \to \infty$.为了不和量子力学的结论冲突,使微观态数成为可数的,可以将 Γ 空间细分成许多大小相同的"相格",其"体积"为 $\delta q_1 \delta p_1 \cdots \delta q_N \delta p_N = h_0^r$,$h_0$ 是一常数.近似地认为一个相格代表经典宏观系统的一个微观态.显然"相格"选得愈小,微观态的确定愈精确.按照经典力学,常数 h_0 可以选得任意小,而量子力学则对 h_0 的大小施加了限制.根据不确定关系[见 §5.2.4(4.18)式],每个粒子的广义坐标和广义动量不能同时精确测定.为了不违背不确定关系,可以要求广义坐标和广义动量的不确定度满足

$$\Delta q_i \Delta p_i \geqslant h'.$$

即:把 Γ 空间分成体积为 h^{Nr} 的相格❶.

以上分析表明,为使经典力学所得结论合理,对 Γ 空间要作如下修改:将 Γ 空间分为大小相等的相格(大小为 h^{Nr}),每个相格代表经典宏观系统的一个微观态.

考虑 Γ 空间中的一个小体积元 $\mathrm{d}\Gamma = \mathrm{d}q_1 \mathrm{d}p_1 \cdots \mathrm{d}q_N \mathrm{d}p_N$,则其中包含的微观态数为 $\mathrm{d}\Gamma/(N!h^{Nr})$,其中 $N!$ 是考虑到量子力学中全同非定域粒子不可识别性质所作的修正.

经典宏观系统能量介于 $E \sim E + \delta E$ 时的微观态数为:

$$\Omega(N,E,V)\delta E = \frac{1}{N!h^{Nr}} \int\cdots\cdots\int_{E \leqslant H \leqslant E+\mathrm{d}E} \mathrm{d}q_1 \cdots \mathrm{d}q_N dp_1 \cdots dp_N. \tag{3.20}$$

【例 4】　再次考虑例 3 中的系统丙,但用经典力学处理.由同一种物质的 N 个单原子分子组成的理想气体,被封闭在边长为 L 的容器中,每个分子有三个广义坐标 (x_i,y_i,z_i) 和三个广义动量 $(p_{x_i},p_{y_i},p_{z_i})$.因此系统的 Γ 空间是 $6N$ 维的.系统的哈密顿量为

$$H = \frac{1}{2m}\sum_{i=1}^{N}(p_{x_i}^2 + p_{y_i}^2 + p_{z_i}^2).$$

这是 Γ 空间中某一等能面的方程.试按照经典力学的描述方法计算系统的能量介于 $E \sim E + \delta E$ 时,和宏观态 (N,E,V) 对应的微观态数.

---

❶ 这相当于将每个自由度的相空间分成大小为 h 的相格,和玻尔量子化条件 §5.1.2(2.3)式一致.另一方面,1911、1912 年 Sackur 和 Tetrode 用 Γ 空间计算单原子理想气体的熵时,发现将相格大小取为 h^{Nr} 后,所得理论值和实验值一致.

解　能量为 E 的等能面所包围的相体积可表示为

$$\Gamma(E) = \int\cdots\int\prod_i \mathrm{d}x_i\mathrm{d}y_i\mathrm{d}z_i \int_{H\leqslant E}\cdots\int\prod_i \mathrm{d}p_{x_i}\mathrm{d}p_{y_i}\mathrm{d}p_{z_i}.$$

坐标的积分区间为

$$0 \leqslant x_i, y_i, z_i \leqslant L, \qquad i = 1,2,\cdots,N.$$

因而有

$$\int\cdots\int\prod_i \mathrm{d}x_i\mathrm{d}y_i\mathrm{d}z_i = (L^3)^N = V^N.$$

为计算动量的积分，进行变量变换. 令

$$p_{x_i} = \sqrt{2mE}\varphi_{x_i}, \quad p_{y_i} = \sqrt{2mE}\varphi_{y_i}, \quad p_{z_i} = \sqrt{2mE}\varphi_{z_i}, \quad i = 1,2,\cdots,N.$$

得到

$$\int\cdots\cdots\int_{H\leqslant E}\prod_i \mathrm{d}p_{x_i}\mathrm{d}p_{y_i}\mathrm{d}p_{z_i} = C_{3N}(2mE)^{3N/2},$$

其中

$$C_{3N} = \int\cdots\cdots\int_{\sum_i(\varphi_{xi}^2+\varphi_{yi}^2+\varphi_{zi}^2)\leqslant 1}\prod_i \mathrm{d}\varphi_{x_i}\mathrm{d}\varphi_{y_i}\mathrm{d}\varphi_{z_i} = \frac{\pi^{3N/2}}{\Gamma(3N/2+1)}.$$

它是以无量纲变量 $\varphi_{x_i}, \varphi_{y_i}, \varphi_{z_i}(i = 1,2,\cdots,N)$ 为轴的 $3N$ 维空间中半径为 1 的球面所包围的球体积，证明见习题 5. 式中的 Γ 函数可在附录 V 中查到.

当 $N \gg 1$ 时，可采用斯特林公式（见附录 V）得 $C_{3N} = [2\mathrm{e}\,\pi/(3N)]^{3N/2}$. 因而理想气体在 Γ 空间中的某一等能面所包围的相体积为

$$\Gamma(E) = [2\pi m(2\mathrm{e}/3)]^{3N/2}(E/N)^{3N/2}V^N.$$

在能量为 $E \sim E + \delta E$ 间的相体积为

$$\begin{aligned}\Gamma(E+\delta E) - \Gamma(E) &= (\mathrm{d}\Gamma(E)/\mathrm{d}E)\delta E \\ &= (3/2)[2\pi m(2\mathrm{e}/3)]^{3N/2}(E/N)^{(3N/2)-1}V^N\delta E.\end{aligned}$$

态密度为

$$\begin{aligned}\Omega(E) &= [(\Gamma(E+\delta E) - \Gamma(E))/(N!h^{3N})] \\ &= \frac{1}{N!h^{3N}}\left(\frac{2\mathrm{e}}{3}2\pi m\right)^{3N/2}\frac{3}{2}\left(\frac{E}{N}\right)^{(3N/2)-1}V^N. \qquad (3.21)\end{aligned}$$

和(3.14)式一致. 这表明，当容器的体积 V 足够大，系统能谱连续时，理想气体微观态数的量子力学计算和经典力学计算相同.

（四）统计规律性

宏观系统动力学状态（微观态）的描述告诉我们，不管粒子遵从量子力学还是经典力学，为了考查宏观系统可达的微观态以及微观态随时间的变化，都必须写出巨大数目的运动方程，并利用初始条件对它们积分. 由于宏观系统中粒子数的巨大

和粒子相互作用的随机性，即使我们耗费得起纸张和时间，也无法掌握系统的初始条件．换句话说，我们根本无法知道，在某时刻系统所处的确切微观态和系统微观量所具有的确切值．乍一看来，似乎可以由此得出结论，随着粒子数的增多，宏观系统的性质变得不可想象的紊乱，以致找不到任何规律，也无法进行研究．但事实并非如此．我们将看到，虽然我们研究的是力学体系，系统中的每个粒子遵从力学规律，但由于大量粒子间的随机相互作用，使得系统整体具有了不能归结为单个粒子行为简单叠加的新性质和新规律，即统计性质和统计规律．这给出了用统计方法解决问题的可能性．

为使读者对统计规律的特点及其和力学规律的区别有初步了解，让我们来看一个演示实验——伽尔顿板实验（并参考习题8和9的解）．在一块竖直木板的上部，规则地钉上许多铁钉，木板的下部用竖直的隔板隔成许多等宽的狭槽，参看图4．从木板顶端漏斗形的入口处可以投入小球，板前复盖玻璃，以使小球留在狭槽内．

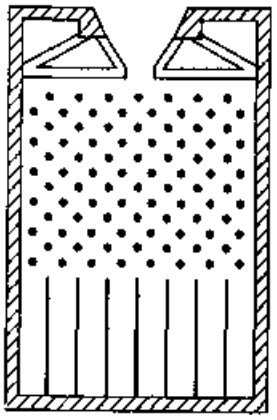


图4　演示统计规律的伽尔顿板实验

从入口处投入一个小球，小球在下落过程中先后和许多铁钉碰撞，最后落入某一狭槽．重复实验会发现，小球落入哪个狭槽是完全偶然的．

同时投入大量小球，则可以看到最后落入各狭槽中小球的数目不相等．在中央槽内分布的小球最多，离中央愈远的狭槽内小球数愈少．可以把小球按狭槽分布的情况用一条连续曲线表示．

重复实验则可发现，在投入少量小球的情况下，每次所得的分布曲线有显著的差别．当投入小球数很多时，每次所得分布曲线彼此近似重合．落入某指定狭槽的小球数和总小球数之比有稳定值，它代表小球落入该狭槽的概率．

改变伽尔顿板放置的情况，如将其倾斜，发生的情况和上述相似，只是小球按狭槽的分布曲线的形状会发生变化，即小球落入某狭槽的概率会发生改变．

总之，实验结果表明，尽管单个小球落入哪个狭槽是偶然的，少量小球按狭槽的分布也带有明显的偶然性，但大量小球按狭槽的分布情况则是稳定的．换句话说，在大量小球的情况下，小球落入某狭槽的概率服从一定的统计规律．改变实验条件，小球落入某狭槽的概率随之改变．

由伽尔顿板实验可以看出，统计规律具有以下特点：

（1）统计规律是对大量随机事件整体起作用的规律，把它用于少量粒子组成的系统时便失去意义．与此相反，动力学规律则是对单个粒子、少量粒子和大量粒子组成的系统皆有效的规律．

（2）在一定的实验条件下，一次事件的结果是偶然的，但事件处于某一可能结

果的概率是确定的. 改变实验条件，概率随之改变. 反映在统计物理中，则可表述为：在一定的宏观条件（如孤立、封闭或开放）下，某一时刻系统处于哪个微观态是偶然的，但处于某一微观态的概率是确定的. 改变宏观条件，不仅可达微观态发生变化，系统处于某一微观态的概率也随之改变.

简言之，统计规律是非确定论性的，它和确定论性的动力学规律："在一定的初始条件下，某一时刻系统必然处于一确定状态"有原则的区别.

(3)统计规律永远伴随着涨落现象. 在伽尔顿板的实验中，即使小球数很大，每次小球按狭槽的分布曲线也只是近似重合而非绝对一致，即存在着涨落. 这反映在统计物理中则表现为，总存在着偏离平均值的涨落. 在有些情况下涨落还会有非常大的作用.

(4)将牛顿第二定律 $F = m(\mathrm{d}^2x/\mathrm{d}t^2)$ 中的 t 换为 $-t$，方程的形式保持不变，这表示动力学定律对时间反演是对称的❶. 也就是说，过去和将来等价，运动是可逆的. 然而经验告诉我们，宏观系统状态的演化是不可逆的，过去和将来不等价，统计规律对时间反演是不对称的. 例如，热传导的傅里叶定律 $\mathrm{d}T/\mathrm{d}t = \lambda \nabla^2 T$ 和扩散的裴克定律 $\mathrm{d}n/\mathrm{d}t = D \nabla^2 n$（$T$ 和 n 分别为温度和浓度，λ 和 D 分别为导热系数和扩散系数）. 如果将 t 换为 $-t$，会得到 $\mathrm{d}T/\mathrm{d}t = -\lambda \nabla^2 T$ 和 $\mathrm{d}n/\mathrm{d}t = -D \nabla^2 n$. 这将意味着，热量会自发地由低温传送到高温，物质会自发地由低浓度区域输运到高浓度区域，而这是我们从来没有观察到的. 在 §7.2.2 中将讨论产生宏观系统不可逆性的原因.

§7.1.4　宏观量的统计性质　系综和系综平均

本节研究系统的宏观性质和微观粒子的力学行为之间的关系.

已知，描述宏观系统整体性质的量叫宏观量. 与此对应，人们往往把宏观系统处于某一微观态时具有的动力学量叫做微观量. 从微观粒子的力学行为来解释系统的宏观性质时，会遇到两种情况：有些宏观量有直接对应的微观量，如气体的压强和分子碰撞器壁的动量变化相对应；而有些宏观量则没有直接对应的微观量，如温度、熵和化学势.

对于有直接对应微观量的宏观量，系综理论的一个基本观点（也有人称之为系综理论的第一个基本假设）是：宏观量是微观量的时间平均，并等价于微观量的系综平均［见以下(4.1)式和(4.4)式］. 两种平均等价的依据是统计物理中一个基

❶ 量子力学中时间反演变换定义为，在 $t \to -t$ 的同时取复共轭［即在 §5.3.1(1.12)式中 $t \to -t$］. 在这样的变换下，波函数 $\psi(r,t) \to \psi^*(r,t)$，而薛定谔方程的形式保持不变.

本的，但迄今仍未得到严格证明的各态历经假说.

各态历经假说可表述如下：只要等待足够长的时间，宏观系统必将多次经历和宏观约束相适应的所有可达微观态.

为具体起见，我们先以气体的密度 $\rho = m\langle n\rangle_{时}$ 为例来说明宏观量是微观量的时间平均. m 是分子的质量，n 是单位体积中的分子数，符号"〈〉"代表求平均的运算. 考虑气体中的小体积元 dV，其中的分子数为 ndV. 为了能精确地测出某一时刻的密度 ρ，应把 dV 取得足够小，如 $dV = 10^{-9}cm^3$，并把一次观测的时间取得足够短，如 $10^{-6}s$. 但是，在这一很短的时间内，气体分子却发生了大量碰撞，例如在 1atm（大气压，见附录Ⅶ）和 0℃下，$dV = 10^{-9}cm^3$ 的体积元中，分子在 $10^{-6}s$ 内的碰撞次数约为 10^{14} 次. 也就是说，实际观测总是在一段时间中进行，而任何宏观短的时间，从微观看来却是足够长. 在这一宏观短微观长的时间中，不断有分子从 dV 出来，也不断有分子进入 dV. dV 中的分子数涨落不定. 我们观测到的气体密度实际是分子质量 m 和这一段时间中的平均粒子数 $\langle n\rangle_{时}$ 的乘积.

再回过头来讨论一般情况. 取一次实际观测时间为 T，系统的某微观量为 A. 由于系统所处的微观态不断变化，微观量 A 的值也随时间不断变化，即有 $A = A(t)$. 于是，宏观性质的一次观测值是相应微观量在这一段时间 T 中的平均，可表为

$$A_{观} = \frac{1}{T}\int_0^T A(t)\,dt = \langle A(t)\rangle_{时}. \tag{4.1}$$

事实上，统计物理所要解释的宏观量应该是在一定宏观条件下，很多次实际观测的平均结果. 从微观来看，这很多次实际观测的总时间很长很长，于是可把(4.1)式改写为

$$\langle A(t)\rangle_{时} = \lim_{T\to\infty}\frac{1}{T}\int_0^T A(t)\,dt. \tag{4.2}$$

要想根据(4.2)式从理论上计算出 $\langle A(t)\rangle_{时}$，必须解宏观系统的薛定谔方程或正则方程组，从而跟踪微观态和微观量随时间的变化 $A(t)$. 在上一节中已经说明，对于具有大量随机自由度的宏观系统，这是不可能的.

现在，我们换一个角度来考查上述的测量问题. 想象在一次观测时间 T 中进行了 $\mathcal{N}$ 次独立测量，每次独立测量时间非常短促，以致可以认为系统仅仅处在某一微观态（注意：这里想象中的一次独立测量不是实际上的一次观测），于是 $A_{观}$ 又可以看成是对这想象中的 $\mathcal{N}$ 次独立测量结果的平均，即(4.1)式可改写为

$$A_{观} = \frac{1}{\mathcal{N}}\sum_{l=1}^{\mathcal{N}} A^{(l)}. \tag{4.3}$$

$A^{(l)}$ 是第 l 次测量中微观量 A 的值，求和 $\sum_{l=1}^{\mathcal{N}}$ 是对在这一段时间 T 内进行的 $\mathcal{N}$ 次想

象的独立测量中系统经历到的微观态进行.

考虑到事实上宏观量是很多次实际观测结果的平均和各态历经假说,则在很多次的实际观测过程中,系统会多次遍历和宏观条件相适应的一切可能的微观态.用 $\mathscr{N}_\alpha$ 表示系统经历某微观态 α 的次数,相应的微观量用 A_α 表示.则由(4.3)式可以得到宏观量(多次实际观测的平均)的表达式为

$$A_{观} = \lim_{\mathscr{N}\to\infty}\frac{1}{\mathscr{N}}\sum_\alpha \mathscr{N}_\alpha A_\alpha = \sum_\alpha P_\alpha A_\alpha. \tag{4.4}$$

其中

$$P_\alpha = \lim_{\mathscr{N}\to\infty}(\mathscr{N}_\alpha/\mathscr{N}) = \lim_{T\to\infty}(\Delta t_\alpha/T). \tag{4.5}$$

(4.4)式中的求和 $\sum_\alpha$ 遍及宏观条件允许下的一切可能的微观态. P_α 代表在足够长的观测时间中,系统在微观态 α 逗留的概率.(4.4)式表明,只要知道系统处于每一微观态时微观量之值 A_α,和处于每一微观态 α 的概率 P_α,就可以计算宏观量 $A_{观}$.

为了形象化地表达(4.4)式所给出的统计平均,并避免长时间跟踪系统来确定 P_α,吉布斯提出了系综的概念.

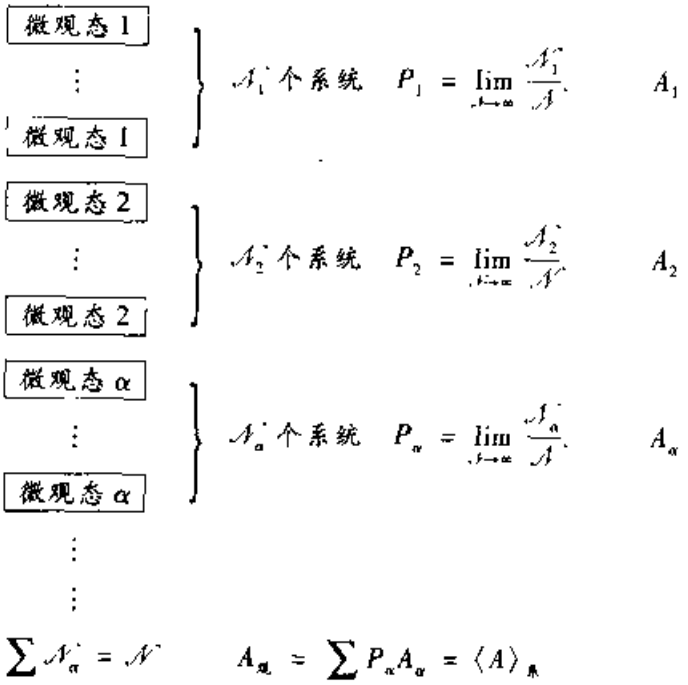


图 1　由 $\mathscr{N}$ 个系统组成的系综.

系综是大量($\mathscr{N}$个)系统的集合.系综中的每个系统和被研究系统具有完全相同的结构,受到完全相同的宏观约束,但可能处在不同的微观态.而且系综中的 $\mathscr{N}$ 个系统按(4.5)式的概率 P_α 分布在给定宏观约束下的所有可能微观态.简言之,系综是大量被研究系统复制品的集合.系综的概念可用图 1 表示.

引入系综的概念后,(4.5)式中的 P_α 又可解释为系综中处于微观态 α 的系统

数 $\mathcal{N}_\alpha$ 和系统总数 $\mathcal{N}$ 的比值，或从系综中任取一个系统时发现它处于微观态 α 的概率. 这叫做系综概率. 而(4.4)式则又可解释为：宏观量是微观量的系综平均 $\langle A\rangle_{系}$. 按照概率的定义(4.5)式，系综平均和对一段时间中进行的多次观测的平均完全等效.

为了帮助理解，让我们考虑一种熟悉的游戏. 抛出一颗骰子，问出现的平均点数是多少？解答这一问题可采用两种方法：一种是将一颗骰子抛出大量次数 $\mathcal{N}$，将每次出现的点数相加再除以 $\mathcal{N}$. 另一种方法是，将 $\mathcal{N}$ 颗相同的骰子交给众人，让众人以相同方式同时抛出，计算所出现的点数的平均值. 经验告诉我们，只要 $\mathcal{N}\to\infty$，两种方式会得出相同的结果. 第一种方法是对单个系统进行长时间的多次观测，求得的是对单个系统的时间平均. 第二种方法是在同一时刻对大量系统或系综进行考查，求得的是系综平均. 可见，在日常生活中我们早已在不知不觉中使用了系综的概念.

最后再强调几点：

①系综是统计物理的一个想象中的工具，而不是实际的客体. 实际的客体是组成系综的单元——系统.

②系综理论的基本观点是，宏观量是相应微观量的时间平均，而时间平均等价于系综平均. 只要求得系综的概率就可以回避跟踪系统微观态随时间演化这一无法解决的难题. 这正是引入系综的主要理由.

③微观量的时间平均等价于系综平均的依据是迄今还未得到严格证明的各态历经假说，因此也可以把时间平均等价于系综平均直接作为假设，交由实践去检验.

④对处于平衡态的宏观系统，其宏观量不随时间改变，则由(4.4)式可知，相应的系综概率必不依赖于时间. 这样的系综叫做稳定系综，它是平衡统计物理的研究对象.

⑤不难想象，系综概率和组成系综的系统所受的外界约束有关. 通常考虑三种最重要的宏观约束：孤立系(N,E,V 给定)；封闭恒温系(N,T,V 给定)和开放系(μ,T,V 给定，μ 为化学势). 相应的系综分别叫微正则、正则和巨正则系综. 平衡系综理论的首要任务是：研究系综概率必须满足的一般条件，和确定三类系综的概率. 下一节研究系综概率满足的一般条件. 本章最后一节介绍微正则系综及其概率，即微正则分布. 正则和巨正则系综及其分布将在本篇第三章中讨论.

§7.1.5　密度函数和密度算符　经典和量子刘维尔定理

上节(4.4)式表明，只要知道系综概率就可以计算出和任何微观量对应的宏

观量.因此,确定系综的概率是系综理论的关键.如果认为系统中的粒子遵循经典力学,则相应的系综叫经典系综;如果认为系统中的粒子遵循量子力学,则相应的系综叫量子系综.由于在经典系综和量子系综中,对微观态(宏观系统的动力学状态)的描述方法,及对微观量(宏观系统的动力学量)的表示方法不同,以致系综概率有不同的表示形式,分别叫密度函数和密度算符.在这一节,我们分别给出密度函数和密度算符的定义及它们随时间的演化方程——刘维尔定理,进而得到平衡统计中密度函数和密度算符应满足的一般要求.在第十篇中我们还将看到刘维尔定理不仅是平衡统计也是非平衡统计的基础.

(一)密度函数　经典刘维尔定理及其推论

已经知道,按照经典力学,系统的微观态由 Γ 空间中的点 $(\boldsymbol{q},\boldsymbol{p}) \equiv (q_1,\cdots,q_N;p_1,\cdots,p_N)$ 代表.在同一时刻,系综中每个系统在 Γ 空间中有各自微观态的代表点.用 $\mathrm{d}\Gamma = \mathrm{d}\boldsymbol{q}\mathrm{d}\boldsymbol{p}$ 表示 Γ 空间中的一个小相体积元.用 $\mathrm{d}\mathscr{N}$ 表示在某时刻 t,在 Γ 空间中某点 $(\boldsymbol{q},\boldsymbol{p})$ 附近相体积元 $\mathrm{d}\Gamma = \mathrm{d}\boldsymbol{q}\mathrm{d}\boldsymbol{p}$ 中的系统代表点数;$\mathscr{N}$ 是系综中的系统总数.显然,$\mathrm{d}\mathscr{N}$ 不仅应和 $\mathscr{N}$ 及 $\mathrm{d}\Gamma$ 成正比,还应依赖于 $\mathrm{d}\Gamma$ 所处的相点位置 $(\boldsymbol{q},\boldsymbol{p})$ 及时间 t,即有

$$\mathrm{d}\mathscr{N} = \mathscr{N}\rho(\boldsymbol{q},\boldsymbol{p},t)\mathrm{d}\boldsymbol{q}\mathrm{d}\boldsymbol{p}, \tag{5.1}$$

或

$$\rho(\boldsymbol{q},\boldsymbol{p},t) = \mathrm{d}\mathscr{N}/(\mathscr{N}\mathrm{d}\boldsymbol{q}\mathrm{d}\boldsymbol{p}). \tag{5.2}$$

$\rho(\boldsymbol{q},\boldsymbol{p},t)$ 叫做概率密度函数或简称密度函数.它表示在时刻 t,从系综中任取一个系统,发现它的微观态代表点处于 $(\boldsymbol{q},\boldsymbol{p})$ 附近的单位相体积元中的概率.根据概率的定义,密度函数应具有正定性和满足归一化条件:

$$\rho(\boldsymbol{q},\boldsymbol{p},t) \geqslant 0, \tag{5.3}$$

$$\int\rho(\boldsymbol{q},\boldsymbol{p},t)\mathrm{d}\boldsymbol{q}\mathrm{d}\boldsymbol{p} = 1. \tag{5.4}$$

(5.4)式中的积分在 Γ 空间中宏观条件允许的范围内进行.

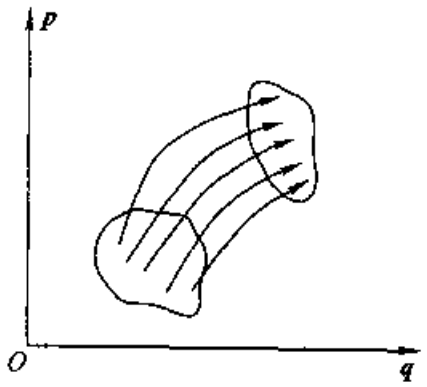


图 1　系综中系统微观态代表点在 Γ 空间中的运动

按照经典力学,系统的任何动力学量(即微观量)可表示为系统中所有粒子的广义坐标及广义动量的函数,即有 $A = A(\boldsymbol{q},\boldsymbol{p})$.于是在经典系综中,宏观量是微观量的系综平均,可由 §7.1.4(4.4)式进一步写为:

$$\langle A\rangle = \int A(\boldsymbol{q},\boldsymbol{p})\rho(\boldsymbol{q},\boldsymbol{p},t)\mathrm{d}\boldsymbol{q}\mathrm{d}\boldsymbol{p}. \tag{5.5}$$

下面来推导密度函数 $\rho(\boldsymbol{q},\boldsymbol{p},t)$ 随时间的演化方程. 随着时间的流逝,系综中 $\mathcal{N}$ 个系统的微观态都在变化,它们在 Γ 空间中的 $\mathcal{N}$ 个代表点将沿各自的相轨迹运动. 可以把它们想象为在 $S=2r\mathcal{N}$ 维空间中具有密度 $D(\boldsymbol{q},\boldsymbol{p},t)=\mathcal{N}\rho(\boldsymbol{q},\boldsymbol{p},t)$ 的流体的流动,如图 1 所示. $D(\boldsymbol{q},\boldsymbol{p},t)$ 表示在 $(\boldsymbol{q},\boldsymbol{p})$ 附近,单位相体积元中的代表点数,叫做代表点密度. 在普通三维空间中,流体运动时物质是守恒的. 类似于电动力学中的电荷守恒 §2.4.1(1.16)式和量子力学中的概率守恒 §5.3.1 (1.11) 式的守恒律以连续性方程表示:

$$\frac{\partial\rho_m}{\partial t}+\operatorname{div}(\rho_m\boldsymbol{v})=0.$$

其中 $\rho_m(x,y,z,t)$ 为流体密度, $\boldsymbol{v}(x,y,z,t)$ 是 t 时刻 (x,y,z) 处的流速, $\operatorname{div}(\rho_m\boldsymbol{v})$ 的分量形式为:

$$\operatorname{div}(\rho_m\boldsymbol{v})=\frac{\partial}{\partial x}(\rho_m v_x)+\frac{\partial}{\partial y}(\rho_m v_y)+\frac{\partial}{\partial z}(\rho_m v_z).$$

现在 $\mathcal{N}$ 个系统的微观态代表点在 Γ 空间中运动,代表点数既不会增多,也不会减少,是守恒的. 因而系综中的系统微观态代表点在 Γ 空间中运动时,代表点密度 $D(\boldsymbol{q}_1,\cdots,\boldsymbol{q}_N;\boldsymbol{p}_1,\cdots,\boldsymbol{p}_N;t)$ 服从类似的方程

$$\frac{\partial D}{\partial t}+\sum_{i=1}^{N}\left[\frac{\partial}{\partial\boldsymbol{q}_i}(D\dot{\boldsymbol{q}}_i)+\frac{\partial}{\partial\boldsymbol{p}_i}(D\dot{\boldsymbol{p}}_i)\right]=0. \tag{5.6}$$

这里 N 代表系统中的粒子数. 考虑复合函数的微商,上式中第二项可分解,并注意 $D=\mathcal{N}\rho$,于是对概率密度函数有

$$\frac{\partial\rho}{\partial t}+\sum_{i=1}^{N}\left\{\rho\left(\frac{\partial\dot{\boldsymbol{q}}_i}{\partial\boldsymbol{q}_i}+\frac{\partial\dot{\boldsymbol{p}}_i}{\partial\boldsymbol{p}_i}\right)+\left(\dot{\boldsymbol{q}}_i\frac{\partial\rho}{\partial\boldsymbol{q}_i}+\dot{\boldsymbol{p}}_i\frac{\partial\rho}{\partial\boldsymbol{p}_i}\right)\right\}=0.$$

利用保守力系统服从的正则方程组[§7.1.3(3.18)式],上式求和中的第一项为零,得

$$\frac{\partial\rho}{\partial t}+\sum_{i=1}^{N}\left\{\dot{\boldsymbol{q}}_i\frac{\partial\rho}{\partial\boldsymbol{q}_i}+\dot{\boldsymbol{p}}_i\frac{\partial\rho}{\partial\boldsymbol{p}_i}\right\}=\frac{\partial\rho}{\partial t}+\sum_{i=1}^{N}\left\{\frac{\partial H}{\partial\boldsymbol{p}_i}\frac{\partial\rho}{\partial\boldsymbol{q}_i}-\frac{\partial H}{\partial\boldsymbol{q}_i}\frac{\partial\rho}{\partial\boldsymbol{p}_i}\right\}=0.$$

利用泊松括号的定义[见 §1.6.2(2.2)式]

$$[H,\rho]_{PB}=\sum_{i=1}^{N}\left\{\frac{\partial H}{\partial\boldsymbol{p}_i}\frac{\partial\rho}{\partial\boldsymbol{q}_i}-\frac{\partial H}{\partial\boldsymbol{q}_i}\frac{\partial\rho}{\partial\boldsymbol{p}_i}\right\}.$$

最后得

$$\frac{\partial\rho}{\partial t}+[H,\rho]_{PB}=0. \tag{5.7}$$

考虑到 $\rho=\rho(\boldsymbol{q}_1,\cdots,\boldsymbol{q}_N;\boldsymbol{p}_1,\cdots,\boldsymbol{p}_N;t)$,由上式得 ρ 对 t 的全微分为零

$$\frac{\mathrm{d}\rho}{\mathrm{d}t}=\frac{\partial\rho}{\partial t}+[H,\rho]_{PB}=0. \tag{5.8}$$

(5.8)式是经典系综中密度函数随时间变化遵从的方程，叫做经典刘维尔定理. 它说明经典系综中系统微观态代表点在 Γ 空间中的运动，就象不可压缩流体在三维空间中流动一样，代表点密度或密度函数沿一条相轨迹随时间始终不变. 代表点没有集中或分散的趋向.

注意，连续性方程(5.6)式反映了相空间中的代表点数目守恒，它和一切满足物质守恒的流体(包括可压缩和不可压缩流体)的物质守恒律相对应；而刘维尔定理(5.8)式则表明，代表点沿相轨迹运动时，其密度始终不变，它对应于三维空间中不可压缩流体的不可压缩性.

如果密度函数不明显地依赖于时间，即有

$$\frac{\partial \rho}{\partial t} = 0. \tag{5.9}$$

则称相应的系综为稳定系综. 显然，对于稳定系综，根据(5.5)式，任何微观量 $A(\boldsymbol{q},\boldsymbol{p})$ 的系综平均将与时间无关. 因此，稳定系综代表处于平衡态的系统.

由刘维尔定理(5.8)可知，(5.9)式或稳定系综成立的条件是

$$[H,\rho]_{PB} = \sum_{i=1}^{N}\left\{\frac{\partial H}{\partial \boldsymbol{p}_i}\frac{\partial \rho}{\partial \boldsymbol{q}_i} - \frac{\partial H}{\partial \boldsymbol{q}_i}\frac{\partial \rho}{\partial \boldsymbol{p}_i}\right\} = 0. \tag{5.10}$$

要使上式成立有两种可能的方式：第一，$\rho(\boldsymbol{q},\boldsymbol{p}) =$ 常数，即稳定系综中系统微观态代表点均匀分布在宏观条件允许下的整个 Γ 空间，或平衡系统以相等的概率处于宏观条件允许下一切可达微观态. 这种均匀系综没有什么意义且不能保证满足归一化，可不予考虑. 第二，$\rho = \rho[H(\boldsymbol{q},\boldsymbol{p})]$，即稳定经典系综的密度函数仅仅通过哈密顿而依赖于所有粒子的广义坐标和广义动量. 这就是由平衡系统组成的稳定系综的密度函数必须满足的一般要求，而不管它是哪类系综. 例如在本篇第三章中我们将看到，对于正则系综有 $\rho_{正则} \propto e^{-H(q,p)/(kT)}$.

(二)密度算符　量子刘维尔定理及其推论

考虑由 $\mathcal{N} \gg 1$ 个结构相同，处于相同宏观条件下的系统组成的系综，假定其中的系统遵循量子力学. 虽然宏观条件一定，但系统仍存在许多可达的微观态. 系综中的系统处于第 α 个微观态的概率为 P_α. 由于系统遵循量子力学，系统的动力学状态或微观态由希尔伯特空间中的态矢量 $|\alpha(t)\rangle$ 或者由相应的密度矩阵 $\rho_\alpha(t) = |\alpha(t)\rangle\langle\alpha(t)|$ 描述(见§5.4.8)；系统的动力学量或微观量由线性厄米算符 $\hat{A}$ 表示；当系统处于由 $|\alpha(t)\rangle$ 或 $\rho_\alpha(t)$ 描述的第 α 个微观态时，动力学量(微观量)一般没有确定值，而只有确定的量子力学平均值：$\bar{A}_\alpha = \langle\alpha(t)|\hat{A}|\alpha(t)\rangle = \mathrm{tr}\hat{A}\rho_\alpha(t)$ [见§5.4.8(8.3)式]. 于是§7.1.4中关于宏观量是微观量的系综平均的(4.4)式应写为

$$A_{宏} = \langle\hat{A}\rangle_{系} = \lim_{\mathcal{N}\to\infty}\frac{1}{\mathcal{N}}\sum_\alpha \mathcal{N}_\alpha\bar{A}_\alpha = \sum_\alpha P_\alpha A_\alpha = \sum_\alpha P_\alpha \mathrm{tr}(\hat{A}\rho_\alpha). \tag{5.11}$$

(5.11)式表明,和经典系综不同,微观量的量子系综平均包含有两重平均:其一是量子力学平均 $\mathrm{tr}\hat{A}\rho_\alpha$,它是由微观粒子的波粒二象性所引起;其二是统计平均 $\sum_\alpha P_\alpha(\cdots)_\alpha$,它是由系综中的系统以不同的概率处于不同的微观态所引起.

注意,系综平均 $\sum_\alpha P_\alpha A_\alpha$ 是线性运算.它可以移到求迹号 tr 之内,即

$$\langle\hat{A}\rangle = \mathrm{tr}\left(\hat{A}\sum_\alpha P_\alpha\rho_\alpha\right) = \mathrm{tr}(\hat{A}\hat{\rho}). \tag{5.12}$$

在这里,我们引入了系综的密度算符 $\hat{\rho}$:

$$\hat{\rho} = \sum_\alpha P_\alpha\rho_\alpha = \sum_\alpha |\alpha(t)\rangle P_\alpha\langle\alpha(t)|. \tag{5.13}$$

其中 ρ_α 代表宏观约束下系统的一个可达纯微观态的密度矩阵,P_α 是系综中的系统处于该纯微观态的概率.和 §5.4.8(8.21)式比较可见,系综是一个混合态.$\hat{\rho}$ 是描述这一混合态的密度算符.利用它求平均,如(5.12)式,就使得量子平均和统计平均一次完成.

(5.12)式表明,微观量 $\hat{A}$ 的量子系综平均等于这一微观量的算符和系综的密度算符乘积的迹.知道了密度算符可以求得任何微观量的系综平均.由此可见,和经典系综类似,确定密度算符是量子系综理论的关键.

由于算符的迹和表象无关,故按(5.12)式取任意表象所得宏观量相同,这正是系统性质客观性的反映.为了更好地理解密度算符的物理意义,我们来写出它在某一表象中的矩阵元.

引入任意算符的本征值方程:$\hat{G}|l\rangle = g_l|l\rangle$.将 $|\alpha(t)\rangle$ 用 $|l\rangle$ 展开

$$|\alpha(t)\rangle = \sum_l c_l^\alpha(t)|l\rangle.$$

其中

$$c_l^\alpha(t) = \langle l|\alpha(t)\rangle.$$

$|c_l^\alpha(t)|^2$ 代表系统处于 $|\alpha(t)\rangle$ 时动力学量 $\hat{G}$ 取本征值 g_l 的概率.写出密度算符 $\hat{\rho}$ 在 $\hat{G}$ 表象中的矩阵元

$$\rho_{lm} = \langle l|\hat{\rho}|m\rangle = \sum_\alpha\langle l|\alpha(t)\rangle P_\alpha\langle\alpha(t)|m\rangle = \sum_\alpha c_l^\alpha(t)c_m^{\alpha*}(t)P_\alpha. \tag{5.14}$$

其对角元为

$$\rho_{ll} \equiv \rho_l = \sum_\alpha |c_l^\alpha(t)|^2 P_\alpha = \langle|c_l(t)|^2\rangle_{系综}. \tag{5.15}$$

由(5.15)式可见,密度算符的对角元是概率 $|c_l(t)|^2$ 的系综平均,因而恒大于或等于零.

密度算符有以下主要性质:

(1)满足归一化.　在(5.12)式中令 $\hat{A} = 1$,得

$$\mathrm{tr}\,\hat{\rho} = 1. \tag{5.16}$$

(2)密度算符是厄米算符. 由(5.14)式很容易得到

$$\rho_{lm} = \rho_{ml}^{*}. \tag{5.17}$$

即 $\hat{\rho}$ 是厄米算符. 由厄米算符的性质可知,存在一组正交完备基使密度算符对角化,以及密度算符的本征值是实数. 由归一化条件(5.16)式知,$\hat{\rho}$ 的对角元是小于或等于1的正实数.

下面来推导密度算符 $\hat{\rho}$ 随时间演化的方程. 将 $\hat{\rho}$ 的定义(5.13)式两边同乘以 $i\hbar$,并对时间求微商

$$i\hbar\frac{\partial\hat{\rho}}{\partial t} = \sum_{\alpha}\left\{P_{\alpha}\left(i\hbar\frac{\partial}{\partial t}\mid\alpha(t)\right)\langle\alpha(t)\mid + P_{\alpha}\mid\alpha(t)\rangle\left(i\hbar\frac{\partial}{\partial t}\langle\alpha(t)\mid\right)\right\}.$$

利用 $i\hbar\dfrac{\partial}{\partial t}\mid\alpha(t)\rangle = \hat{H}\mid\alpha(t)\rangle$, $-i\hbar\dfrac{\partial}{\partial t}\langle\alpha(t)\mid = \langle\alpha(t)\mid\hat{H}$ 立即得到

$$\partial\hat{\rho}/\partial t + (i/\hbar)[\hat{H},\hat{\rho}] = 0, \tag{5.18}$$

其中,$[\hat{H},\hat{\rho}] = \hat{H}\hat{\rho} - \hat{\rho}\hat{H}$ 是 $\hat{H}$ 和 $\hat{\rho}$ 的对易关系,而

$$(i/\hbar)[\hat{H},\hat{\rho}] = (i/\hbar)(\hat{H}\hat{\rho} - \hat{\rho}\hat{H}) \tag{5.19}$$

是量子泊松括号. (5.18)式和(5.7)式对应,叫做量子刘维尔定理.

系统处于平衡态时,相应的系综是稳定系综,即有 $\partial\hat{\rho}/\partial t = 0$. 于是由量子刘维尔定理可知,必有 $[\hat{H},\hat{\rho}] = 0$,即哈密顿算符 $\hat{H}$ 和密度算符 $\hat{\rho}$ 是可对易算符. 进而可得出两点推论:其一,稳定系综的密度算符 $\hat{\rho}$ 是哈密顿算符 $\hat{H}$ 的函数,且 $\hat{H}$ 不依赖于时间,即有 $\hat{\rho} = \hat{\rho}(\hat{H})$;其二,从量子力学已经知道,对易算符有共同的本征函数,而且算符在自身表象中为对角矩阵. 因此,稳定系综的密度算符在能量表象中是对角矩阵,且对角元是能量本征值的函数,即在能量表象中有

$$\rho_{mn} = \rho_{n}(E_{n})\delta_{mn}. \tag{5.20}$$

不难看出,密度算符 $\hat{\rho}$ 的具体形式和系综中系统所受宏观约束有关,即微正则、正则和巨正则系综的 $\hat{\rho}$ 有不同形式. 但量子刘维尔定理告诉我们,它们的 $\hat{\rho}$ 都必须满足以上一般要求. 由于密度算符 $\hat{\rho}$ 在能量表象中具有对角矩阵的简单特点,在利用(5.12)式求微观量的系综平均时,往往在能量表象中进行. 能量和能量表象在统计物理中有特别重要的作用. 以下求宏观量时无一例外均在能量表象中进行.

我们看到,在经典系综理论和量子系综理论中有完全对应的概念和公式. 将(5.5)式和(5.12)式比较可以看到,只要将经典系综公式中的 $A(\boldsymbol{q},\boldsymbol{p})$ 换为 $\hat{A}$,$\rho(\boldsymbol{q},\boldsymbol{p},t)$ 换为 $\hat{\rho}(t)$,对 Γ 空间的积分 $\int\cdots\int\mathrm{d}\boldsymbol{q}\mathrm{d}\boldsymbol{p}$ 换为对矩阵的对角元求和(即求迹),立刻得到相应的量子系综公式,反之亦然. 因此,为了节约篇幅,在以后的讨论中往往只写出一种情形下的公式.

§7.1.6　微正则系综和微正则分布

在§7.1.4的最后曾经指出，系综理论主要研究处于三种不同宏观条件下的平衡系统组成的三种稳定系综，即由孤立系统组成的微正则系综、由恒温封闭系统组成的正则系综和由开放系统组成的巨正则系综. 刘维尔定理的定态解只能告诉我们稳定系综的密度函数(或密度算符)必须满足的一般条件，而不能给出三种系综密度函数(或密度算符)的具体形式. 在系综理论中，微正则系综的密度函数(或密度算符)是作为一个基本假设提出的，而正则系综和巨正则系综的密度函数(或密度算符)则由这个基本假设推导出来. 本节介绍系综理论的这个基本假设.

由大量处于平衡态的孤立系统组成的系综叫做微正则系综. 孤立系统的能量介于 $E \sim E + \delta E$ 之间，且 $\delta E \ll E$，其平衡态由 (N,E,V) 描述.

显然，孤立系不可能到达这个能量范围之外的微观态. 但在 $E \sim E + \delta E$ 的能量范围内，孤立系可能达到的微观态是大量的. 这些微观态都满足给定的宏观条件，它们应当是平权的. 因此，在满足刘维尔定理要求的条件下，一个合理的假设是：平衡孤立系的一切可达微观态出现的概率相等. 这叫做先验等概率假设. 由这一假设得到的系综密度函数称为微正则分布.

微正则分布的经典表达式为

$$\rho(\boldsymbol{q},\boldsymbol{p}) = \begin{cases} \text{常数 } c, & \text{当 } E < H(\boldsymbol{q},\boldsymbol{p}) < E + \delta E, \\ 0 & \text{其他.} \end{cases} \tag{6.1}$$

$$c = \lim_{\delta E \to 0} \left(\int\cdots\int_{\delta E} \mathrm{d}\boldsymbol{q}\mathrm{d}\boldsymbol{p} \right)^{-1}. \tag{6.2}$$

这里用 $H(\boldsymbol{q},\boldsymbol{p})$ 表示系统所处微观态的能量. (6.1)式的意义是，经典微正则系综中所有系统的微观态代表点，稳定而且均匀地分布在 Γ 空间的能量壳层 $E \sim E + \delta E$ 内. $\lim\limits_{\delta E \to 0} \left(\int\cdots\int_{\delta E} \mathrm{d}\boldsymbol{q}\mathrm{d}\boldsymbol{p} \right)$ 是能量壳层的相体积.

微正则分布的量子表达式是

$$\rho_{mn} = \rho_n \delta_{mn}, \tag{6.3}$$

$$\rho_n = \begin{cases} 1/\Omega(N,E,V), & \text{当 } E < H(\boldsymbol{q},\boldsymbol{p}) < E + \delta E, \\ 0 & \text{其他.} \end{cases} \tag{6.4}$$

$\Omega(N,E,V)$ 是和宏观态 (N,E,V) 对应，孤立系可达的微观态数. 由量子刘维尔定理知道，在能量表象中，密度算符是对角矩阵，而(6.3)和(6.4)式表明，微正则系综的密度矩阵的对角元等于可达微观态数的倒数. 由§7.1.3(3.20)式，如果把经典统计理解为量子统计的经典极限，对于由 N 个全同非定域粒子组成的系统，在 Γ

空间的能量壳层中的微观态数应为

$$\Omega = \frac{1}{N!h^{3r}}\int\cdots\int_{\delta E}\mathrm{d}\boldsymbol{q}\mathrm{d}\boldsymbol{p}. \tag{6.5}$$

最后强调两点：第一，微正则分布是一个假设，它的正确性只能通过将它的推论和实验比较而加以验证. 它是一个基本假设，是整个统计物理理论的基石. 第二，由于计算 $\Omega(N,E,V)$ 的困难，直接运用微正则分布往往并不方便. 由微正则分布可以推得正则分布和巨正则分布. 用正则分布和巨正则分布研究具体系统的宏观性质要方便得多. 因此，我们只介绍微正则分布并用它推导正则分布和巨正则分布，而不过多地举例说明如何直接用微正则分布计算宏观量.

习 题

1. 试求理想气体的定压膨胀系数和等温压缩系数.

2. 某气体的定压膨胀系数和等温压缩系数分别为 $\alpha = nR/(pv)$，$\kappa_T = (1/p) + (a/V)$，其中 n,R,a 都是常数. 试求此气体的状态方程.

3. 一弹性棒的热力学状态可用它的长度 L，应力 f 描述. L,f 和温度 T 的关系即为其状态方程. 今设此弹性棒发生一微小变化，从一平衡态变到另一平衡态. 试证明：

$$\mathrm{d}f = -\alpha AY\mathrm{d}T + AY\mathrm{d}L/L.$$

其中 A 为棒的横截面积，α 为线胀系数，Y 为杨氏模量.

4. 对一气体的膨胀系数和压缩系数进行测量的结果，得出以下方程：

$$\left(\frac{\partial v}{\partial T}\right)_p = \frac{R}{p} + \frac{a}{T^2}, \qquad \left(\frac{\partial v}{\partial p}\right)_T = -Tf(p).$$

其中 a 是常数，$f(p)$ 只是 p 的函数. 证明：(a) $f(p) = R/p^2$，(b)状态方程为 $pv = RT - ap/T$.

5. 试给出半径为 1 的 n 维球体积 $V_n(1) = \pi^{n/2}/\Gamma(n/2+1)$ 的证明.

提示 半径为 R 的 n 维球体积和球面积可分别表为

$$V_n(R) = C_nR^n, \qquad S_n(R) = nC_nR^{n-1}, \tag{1}$$

由附录(V.3)式知积分：

$$I_n = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\mathrm{e}^{-\alpha(x_1^2+x_2^2+\cdots+x_n^2)}\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n = \left[\int_{-\infty}^{\infty}\mathrm{e}^{-\alpha x_1^2}\mathrm{d}x_1\right]^n = \left(\frac{\pi}{\alpha}\right)^{n/2}, \tag{2}$$

利用(1)式计算 I_n，并将结果和(2)式比较可得 C_n.

6. 利用附录(V.20)给出的斯特林公式：$\ln N! \approx N\ln N - N(N>10)$ 证明上题中的系数 C 满足下式：

$$C_{3N}(1) = \frac{\pi^{3N/2}}{\Gamma(3N/2+1)} \approx \left(\frac{2e\pi}{3N}\right)^{3N/2}. \tag{3}$$

7. N 个近独立频率为 ω 的三维谐振子组成的系统共有 $3N$ 个自由度，确定系统的微观态需

要$3N$个坐标x_i,y_i,z_i和$3N$个动量$p_{x_i},p_{y_i},p_{z_i},(i=1,2,\cdots,N)$，因此系统的$\Gamma$空间是$6N$维的.系统的哈密顿量为

$$H=\sum_{i=1}^{N}\left[\frac{1}{2m}(p_{x_i}^2+p_{y_i}^2+p_{z_i}^2)+\frac{m\omega^2}{2}(x_i^2+y_i^2+z_i^2)\right].$$

试求某一等能面所包围的相体积$\Gamma(E)$和态密度$\Omega(E)$.

8. 已知用计算机产生高斯随机数的方法是

$$y=X_1+X_2+\cdots+X_{12}-6 \tag{4}$$

$X_i(i=1,2,\cdots,12)$是由伪随机数产生器产生的在0和1之间均匀分布的随机数.

(a)试用计算机先后产生1000个和10000个这样的高斯随机数，并绘出其直方图.

提示　在直方图中，区间宽度可取为0.2，并在-2.9至2.9间绘出直方图.

(b)由上面结果说明统计规律的特点.

(c)用计算机计算证明，由(4)式所产生的高斯随机数是标准高斯随机数，即有：

$$\langle y\rangle=0\qquad\sigma_y^2=1 \tag{5}$$

9. 已知概率服从下式的正整数随机数列为泊松随机数列

$$P(n_i)=\mathrm{e}^{-m}m^{n_i}/n_i!\qquad n_i=0,1,2,\cdots\quad m=\langle n_i\rangle=\sigma_{n_i}^2. \tag{6}$$

用计算机产生泊松随机数的方法如下：由伪随机数产生器产生一个介于0～1之间的均匀分布随机数y，若满足下式，

$$\sum_{n_i=1}^{k}\mathrm{e}^{-m}m^{n_i}/n_i!\leqslant y\leqslant\sum_{n_i=1}^{k+1}\mathrm{e}^{-m}m^{n_i}/n_i! \tag{7}$$

则对应的k即为满足(6)式的泊松随机数.

(a)试编程先后产生10000个$m=1$和$m=2$的泊松随机数，并分别绘出其直方图.

(b)分别计算两组泊松随机数的平均值和方差函数.

(c)由解题结果说明统计规律的特点.

第二章　统计热力学基础

从发展历史看，热力学是以从实验中总结出的三个基本定律为基础，用逻辑推理方法建立起来的一门唯象科学. 热力学方法的特点是，在研究宏观系统的性质时，完全不考虑系统的微观结构和运动，而结果却十分可靠. 但科学发展至今，从统计物理的基本原理已经可以导出这些基本定律，人们对热力学中基本概念的微观本质也有了深刻认识. 所以本章将不循历史发展轨迹，而是在统计物理的基础上引入热力学的基本定律和基本概念，然后在对基本定律、基本概念的微观本质有深刻认识的基础上，进行热力学方法的训练. 这样做的好处是：从一开始就使读者对热力学获得更直观、更深入的理解.

本章第一节介绍热力学第一定律，着重对准静态过程中的功和热量进行微观解释. 在第二节中，从统计物理的基本假设——微正则分布出发，导出热力学第二定律. 第三节用第一、第二定律研究孤立系统中的各种平衡，从微观角度引入几个描述趋向平衡的方向和达到平衡的条件的强度量——温度，化学势和压强，进而得到热力学基本方程. 在第五至第七节中，从热力学基本方程出发，利用勒让德变换，引入几个辅助热力学函数（又称热力学势），导出一套普遍热力学关系及平衡的稳定性判据，并举例说明它们的应用. 掌握普遍热力学关系及稳定性判据的应用是热力学方法的核心. 第八、第九节讨论相变热力学. 相变的统计理论留到第九篇讨论. 在本章的最后介绍热力学第三定律及其统计解释.

§7.2.1　热力学第一定律　功和热量的微观解释

热力学第一定律是能量守恒与转换定律在一切涉及热现象的宏观过程中的具体表现. 本节先考虑它在封闭系统中的表达形式，并给予微观解释.

（一）封闭系统中的热力学第一定律

当封闭系从一个平衡态变化到另一个平衡态时，根据能量守恒定律有

$$\Delta U = \Delta Q + \Delta W. \tag{1.1}$$

U 叫做系统的内能，它是和系统的内部自由度（如分子平动的能量，分子转动和振动的能量，磁性原子在外场中的能量及电子的能量等）相关的能量的系综平均，即

$$U = \langle E \rangle = \sum_{\alpha} E_{\alpha} P_{\alpha}. \tag{1.2}$$

α 是微观态的指标，E_{α} 是系统处于微观态 α 时的能量，P_{α} 是系统处于微观态 α 的概率，$\sum_{\alpha}$ 表示对宏观条件下一切可能达到的微观态求和.

ΔW 代表由于外参量(如体积、磁场强度等)变化而引起的内能变化,叫做外界对系统所作的功. ΔQ 则代表当外参量不变时系统内能的变化,叫做系统从外界吸收的热量.

如果系统的状态发生一无穷小的变化,则(1.1)式化为

$$dU = đQ + đW. \tag{1.3}$$

因为 U 是态函数,故 dU 由初、末平衡态决定而与过程无关;而 dQ 和 $đW$ 则不仅与初、末平衡态有关,还依赖于变化过程.它们只是微小量而不是全微分.(1.1)式和(1.3)式是封闭系中热力学第一定律的表达式.下面讨论 $đQ$ 和 $đW$ 的计算.

(二)准静态过程及其功

一般说来,外界对系统所作的功要由实验测定.只有在无摩擦、无滞后效应的准静态过程中,外界对系统的功 $đW$ 才能用系统的态参量表示,从而可进行理论计算.下面首先说明什么是准静态过程.

若过程进行的每一中间态都是平衡态,则此过程称为准静态过程.显然,只有当过程进行得无限缓慢,以致系统有足够的时间恢复平衡时,准静态过程才能实现.因此,准静态过程是一个理想化的概念.但许多实际过程可以相当好地近似为准静态过程.只要一个物理量发生微小(宏观上微小)变化的时间比该物理量的弛豫时间长得多,就可以认为该过程足够缓慢,而近似作为准静态过程处理.如气体压强均匀化的速度大约为每秒几百米(分子热运动的平均速度),而气缸活塞的运动速度不超过每秒几米,因此认为气缸中的气体经历准静态过程是没有问题的.极化(不管是电的或磁的)过程的弛豫时间都很短,所以在不涉及变化极快的过程(如高频交变电场中的介质)时,总可以把极化过程看成是准静态的.

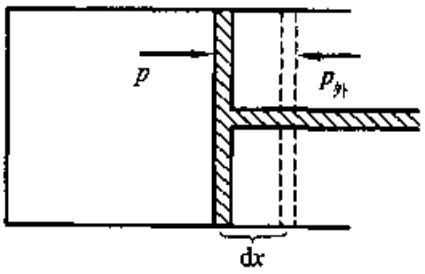


图 1　准静态过程的功

为了说明只有在无摩擦的准静态过程中,外界对系统的功才能用系统的态参量表示,我们考虑一个例子.如图 1,当圆柱形筒内的气体被活塞压缩 dV 时,外界对系统的功为 $đW = -p_{外}\,dV$.如果发生的不是准静态过程,则活塞迅速移动使气体的平衡受到破坏,它的压强变得不均匀,当然不能和外压强相等,因而,外界对气体的功不能用气体的态参量表示.即使过程是准静态地进行,如果活塞和筒壁的摩擦不能忽略,虽然气体在任何时刻都处于平衡态,从而有均匀的压强,但由于有摩擦力,压缩气体时,作用于活塞上的外压强大于气体内部的压强 $p_{外} > p$;反之,气体膨胀时,外压强小于气体压强 $p_{外} < p$.不难看出,两种情况下均有 $đW > -p\,dV$.即:虽然过程是准静态的,但因为有摩擦,外界对系统的功仍然不能用系统的态参量表示.只有在无摩擦的准静态过程中,不仅在任何时刻系统处于平衡态,有均匀

压强，而且外压强始终和气体压强相等 $p_{外}=p$，才有

$$\mathrm{d}W=-p_{外}\,\mathrm{d}V=-p\mathrm{d}V. \tag{1.3'}$$

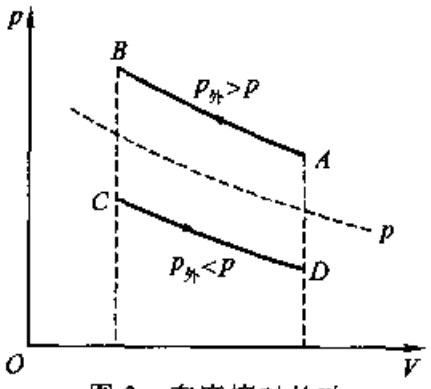


图 2　有摩擦时的功

(1.3′)式用描述系统平衡态的态参量把无摩擦准静态过程中的功定量地表示出来. 这样，利用状态方程就可以具体计算无摩擦准静态过程中的功.

以上情况可以用图 2 说明. 虚线代表准静态过程中气体的压强，实线代表有摩擦时的外压强. 在有摩擦的情况下如果先压缩气体，再使气体膨胀，而完成一个循环 $ABCDA$. 此路程与 §7.1.2 图 1 相似，具有滞后效应的特征. 在一切具有滞后效应的过程中，功都不能用态参量表示. 我们今后将不考虑这种情况，凡是提到准静态过程，都是指无摩擦、无滞后效应的准静态过程.

下面再介绍几种简单系统在准静态过程中功的表达式.

1. 液体表面膜

设有液体表面薄膜张于金属框上，如图 3. 长为 l 的金属丝 AA' 可以自由移动，液体膜的表面张力为 σ，所以作用在 AA' 上的外力等于 $2\sigma l$. 当 AA' 准静态地移动 $\mathrm{d}x$ 时，外界对液体表面膜的功是

$$\mathrm{d}W=2\sigma l\mathrm{d}x=\sigma\mathrm{d}A. \tag{1.4}$$

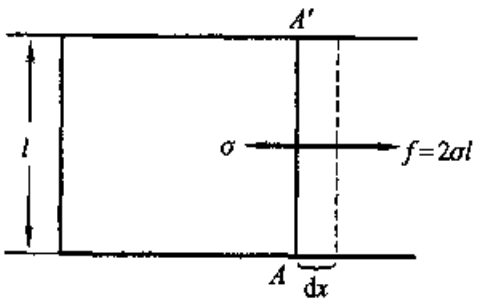


图 3　表面膜系统的功

2. 磁介质

将磁介质磁化可采用多种方法. 下面考虑将磁介质置于螺线管中的方法来推导磁化功的表达式，见图 4.

取一长为 l，截面积为 S，绕有 N 匝线圈的螺线管，设螺线管的长度比管的直径大得多，因此可以认为当线圈中通以电流 i 时，螺线管中的磁场是均匀的，且据安培环路定理有

$$H=\frac{4\pi}{c}\frac{N}{l}i=(4\pi/c)ni. \tag{1.5}$$

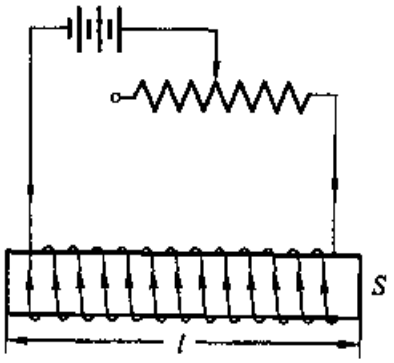


图 4　磁介质系统的功

再将磁介质置于螺线管中，亦即置于由(1.5)式表示的外场(或激化场)中，介质中的辅助量磁场强度 H 仍由(1.5)式表出. 当改变电流时，螺线管中将产生感应电动势 $\varepsilon=-(N/c)\mathrm{d}\varphi/\mathrm{d}t$，其中 $\mathrm{d}\varphi$ 是磁通量的变化：$\mathrm{d}\varphi=S\mathrm{d}B$. 因而

$\varepsilon = -(1/c)NS(\mathrm{d}B/\mathrm{d}t)$. 为了维持电流,电源必须反抗感应电动势作功

$$\mathrm{d}W = -\varepsilon i\mathrm{d}t = NiS\mathrm{d}B/c.$$

将其中的 i 用(1.5)式代入

$$\mathrm{d}W = lSH\mathrm{d}B/4\pi = VH\cdot\mathrm{d}B/4\pi$$

而 $\boldsymbol{B} = \boldsymbol{H} + 4\pi\boldsymbol{m}$,最后得

$$\mathrm{d}W = (V/4\pi)\boldsymbol{H}\cdot\mathrm{d}\boldsymbol{H} + V\boldsymbol{H}\cdot\mathrm{d}\boldsymbol{m} = (V/4\pi)\mathrm{d}(\boldsymbol{H}^2/2) + \boldsymbol{H}\cdot\mathrm{d}\boldsymbol{M}, \quad (1.6)$$

这是以磁场和介质一起作为系统时,外界对系统所作的功. 第一项与介质的性质无关,且 $\boldsymbol{H}^2/2$ 正是真空中磁场强度为 H 时,单位体积中场能的表达式,因此第一项表示使磁场强度从 $\boldsymbol{H}$ 增加到 $\boldsymbol{H}+\mathrm{d}\boldsymbol{H}$ 时,外界所作的功,或储存在真空螺线管中磁能的改变量;第二项则表示使介质的磁矩从 $\boldsymbol{M}$ 增加到 $\boldsymbol{M}+\mathrm{d}\boldsymbol{M}$ 时外界所作的功. 若仅以介质作为系统,则外界对系统所作的磁化功为

$$\mathrm{d}W = \boldsymbol{H}\cdot\mathrm{d}\boldsymbol{M}. \quad (1.7)$$

其中 $\boldsymbol{H}$ 表示外场或激化场. 在以上推导中,我们采用的是高斯单位制,若采用国际单位制,(1.7)式将由 $\mathrm{d}W = \mu_0\boldsymbol{H}\cdot\mathrm{d}\boldsymbol{M}$ 代替.

严格说来,在磁化过程中介质的体积也会发生变化,因而外界对介质所作的功包括两部分:

$$\mathrm{d}W = -p\mathrm{d}V + \boldsymbol{H}\cdot\mathrm{d}\boldsymbol{M}. \quad (1.7')$$

由以上几个例子可以看到,在准静态过程中外界对系统所作的功由一强度量和其共轭广延量的变化相乘得到. 引入广义力 Y_i 和广义坐标 x_i,并考虑外界和系统可能有多种功交换,则准静态过程的功一般可写为

$$\mathrm{d}W = \sum Y_i\mathrm{d}x_i. \quad (1.8)$$

最后,我们要指出,无摩擦无滞后效应准静态过程具有可逆性. 一个系统沿某一过程从状态 I 变到状态 II 后,如果能在逆方向沿同一过程从状态 II 回到状态 I,而不引起外界变化,则这个过程叫做可逆过程. 从图 2 可以看到,在无摩擦的准静态过程中,实线与虚线重合,这表示膨胀时系统所经过的一系列中间态与压缩时完全相同;而且由于过程的每一步沿正、反方向进行时,外界对系统所作的功大小相等、符号相反,根据热力学第一定律,系统与外界交换的热量也必然大小相等符号相反. 所以,当系统沿原来的路径回到原状态后,外界的变化被完全消除,即:无摩擦的准静态过程是可逆可程. 与此相反,在有摩擦的情况下,当外界条件逆转时,系统沿一条新的路径回到原来状态,于是外界对系统作了净功,由图 2 中的循环面积 $ABCDA$ 代表. 这部分功最终转变成热传给外界. 其结果,虽然系统回到了初态,而外界发生了变化. 因此,有摩擦热耗散或有滞后效应的准静态过程是不可逆过程.

(三)热量的计算　热容量

实验证明,在没有相变及化学反应时,系统吸入的热量与升高的温度成正比,

故热量可表为

$$\Delta Q = C\Delta T. \tag{1.9}$$

比例系数 C，即物质温度升高 1℃所需的热量，称为热容量，其定义为

$$C = \lim_{\Delta T\to 0}\frac{\Delta Q}{\Delta T} = \frac{\mathrm{d}Q}{\mathrm{d}T}. \tag{1.10}$$

单位质量物质的热容量是表征该物质性质的量，称为该物质的比热，常用小写字母 c 表示. 特别是，一摩尔物质的热容量称为摩尔比热，或摩尔热容量.

实验证明，在不同过程中使系统升高同样的温度所需热量不同，故热容量的测定是在特定条件下进行的. 对于流体，常见的是定体热容量和定压热容量，其定义如下：

$$C_V = \lim_{\Delta T\to 0}\left(\frac{\Delta Q}{\Delta T}\right)_V = \left(\frac{\mathrm{d}Q}{\mathrm{d}T}\right)_V, \tag{1.11}$$

$$C_p = \lim_{\Delta T\to 0}\left(\frac{\Delta Q}{\Delta T}\right)_p = \left(\frac{\mathrm{d}Q}{\mathrm{d}T}\right)_p, \tag{1.12}$$

括号的下角 V(或 p) 表示在进行这一过程时保持体积(或压强)不变. 对于单位物质，相应地就有定体比热和定压比热. 由以上诸式可得出热量的计算公式：

$$\mathrm{d}Q = C\mathrm{d}T, \quad (\mathrm{d}Q)_V = C_V\mathrm{d}T, \quad (\mathrm{d}Q)_p = C_p\mathrm{d}T. \tag{1.13}$$

由于热量是过程量，所以在 C_V 和 C_p 的定义(1.12)式的右方不是态函数的微商，不便于理论运算. 但利用热力学第一定律很容易将 C_V 和 C_p 表为态函数的微商.

根据热力学第一定律，在等容过程中有 $\Delta Q = \Delta U$，于是

$$C_V = \lim_{\Delta T\to 0}\left(\frac{\Delta Q}{\Delta T}\right)_V = \lim_{\Delta T\to 0}\left(\frac{\Delta U}{\Delta T}\right)_V = \left(\frac{\partial U}{\partial T}\right)_V. \tag{1.14}$$

对于一个简单系统，U 是 T,V 的函数，所以 C_V 也是 T,V 的函数.

对于准静态等压过程，热力学第一定律可写为：

$$\Delta Q = \Delta U + p\Delta V = \Delta(U + pV), \quad (\text{等压过程}). \tag{1.15}$$

用 U,p 和 V 组成一个新的态函数：

$$H = U + pV, \tag{1.16}$$

叫做焓，则

$$C_p = \lim_{\Delta T\to 0}\left(\frac{\Delta Q}{\Delta T}\right)_p = \left(\frac{\partial(U+pV)}{\partial T}\right)_p = \left(\frac{\partial H}{\partial T}\right)_p. \tag{1.17}$$

对于一个简单系统，C_p 和焓 H 一样，是 T,p 的函数. 由(1.15)式可见，焓的物理意义是：它等于等压过程中系统吸收的热量. 在下一章我们还要对焓的意义和作用作进一步的讨论.

将压强 p 换为其他广义力，体积 V 换为相应的共轭广义坐标，就可以得到其他

的热容量. 如对磁介质作代换：$P \to -H, V \to M$，可写出磁介质的热容量为 $C_M = (\partial U/\partial T)_M$ 和 $C_H = [\partial(U - \boldsymbol{H}\cdot\boldsymbol{M})/\partial T]_H$（注意区分磁场 $\boldsymbol{H}$ 和焓 H）.

（四）热力学第一定律的微观解释

设想系统的状态无限缓慢地经历一个微小变化，系统始终近似处于平衡态. 根据(1.2)式，对此准静态元过程有

$$\mathrm{d}U = \sum_\alpha E_\alpha \mathrm{d}p_\alpha + \sum_\alpha p_\alpha \mathrm{d}E_\alpha. \tag{1.18}$$

从数学上看，右方第一项代表各微观态的能量不变，而由于各微观态出现的概率发生变化所引起的内能改变；第二项则代表各微观态出现的概率不变，而由于各微观态的能量改变所引起的内能变化.

下面进一步分析这两项的物理意义. 在§7.1.3中已经指出，系统处于某一微观态时的能量，除和系统本身的性质有关外，只是外参量 x 的函数，即 $E_\alpha = E_\alpha(x)$. 外参量 x 变化时，微观态的能量发生变化. 反之，微观态的能量不变，意味着外参量没有变化. 由此可知，(1.18)式中第一项代表外参量不变时，系统内能的变化，或代表系统从外界吸收的热量. 即有

$$\sum_\alpha E_\alpha \mathrm{d}p_\alpha = \mathrm{d}Q. \tag{1.19}$$

将(1.19)式代入(1.18)式并和(1.3)式比较，可知(1.18)中的第二项代表由于外参量改变而引起的内能变化，即代表外界对系统所作的功. 于是有

$$\sum_\alpha p_\alpha \mathrm{d}E_\alpha = \sum_\alpha p_\alpha (\partial E_\alpha/\partial x)\mathrm{d}x = Y\mathrm{d}x = \mathrm{d}W. \tag{1.20}$$

其中 $Y = \sum_\alpha p_\alpha(\partial E_\alpha/\partial x) = \langle(\partial E_\alpha/\partial x)\rangle_{系}$. 在写出(1.20)式时，已经默认改变外参量 x 的唯一结果，只是改变微观态的能量. 根据§6.2.2中关于微扰绝热加入的讨论知道，这只有当外参量的变化无限缓慢（或系统经历准静态过程）时才能成立. (1.19)和(1.20)式说明了准静态过程中热量和功的微观本质.

§7.2.2 热力学第二定律 熵

根据热力学第一定律，孤立系统的能量必须保持不变或守恒，违反能量守恒的过程不可能发生. 但大量事实表明，孤立系统中不违反能量守恒的过程却并非都能发生. 例如，气体可以向真空自由膨胀，直到密度均匀为止，却不能自动收缩到一部分空间中使另一部分成为真空. 热量可以从高温传到低温，直到温度均匀为止，却不能自动从低温传到高温；运动物体由于摩擦损失了动能最终会停止，动能变成物体及周围环境的内能，这些内能却不会重新聚集并变成动能使物体再运动起来. 这一切都说明，除热力学第一定律以外，自然界还存在一个支配能量传递、转换的方

向和限度的规律.这就是热力学第二定律.

孤立系统中发生的过程往往叫做自发过程.以上例子表明,自发过程具有单向性,它们不仅不能直接反向进行,而且无论用什么曲折复杂的方法都不可能把它们的效果完全消除而不产生其他影响.这样的过程叫做不可逆过程.因此又可以说,热力学第二定律的实质在于:它指出了一切与热现象有关的自发过程是不可逆过程.在热学课程中已经证明,这些表面上完全不同的不可逆过程彼此等价,即从一个自发过程的不可逆性可以推断其他自发过程的不可逆性.因此,任何一个不可逆过程都可以作为第二定律的表述,如开尔文表述:*不可能从单一热源取热使之完全变为有用的功而不产生其他影响*,和克劳修斯表述:*不可能把热从低温物体传到高温物体,而不产生其他影响*.既然一切不可逆过程存在着内在联系,那就应该有支配一切不可逆过程发生的共同原因,有判断一切不可逆过程的进行方向和限度的共同标准.这个共同的原因是什么?共同的标准或表征这一共同标准的物理量又是什么?它的微观本质如何?它满足怎样的数学表达式?这些就是热力学第二定律的内容,也就是本节要回答的问题.

(一)支配自发过程方向和限度的物理量

为具体起见,我们以气体的自由膨胀为例,从微观角度来分析孤立系统中自发过程发生的原因,和支配自发过程的进行方向和限度的物理量.分析的出发点是系综理论的第二个基本假设——微正则分布.

如图1所示,设想将一定量的某种气体储存在一个由刚性绝热材料做成的盒子中.开始用一隔板将气体限制在盒子的左部 V_i 中.抽出隔板后,系统处于非平衡态,经过一段弛豫时间,气体充满整个盒子 V_f,最终达到平衡态.

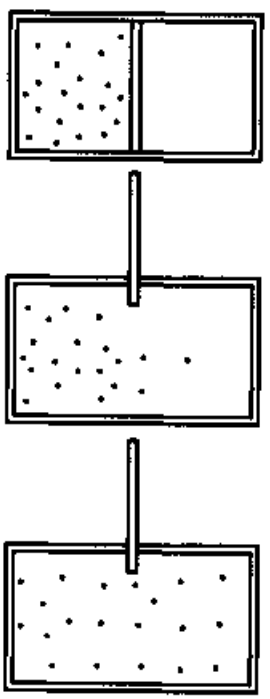


图1　气体的自由膨胀

这是一个典型的孤立系统中的自发过程——自由膨胀过程,如图1.在气体最终达到的平衡态,气体分子的数密度是各处均匀,且不随时间改变的.盒子的左、右半部各包含一半的分子数.当然,不是精确的一半,涨落总是存在的,不过和总分子数相比涨落一般很小.大的涨落,例如所有分子集中到盒子的左半部(即自动收缩),原则上说不是不可能,但事实上从未被观察到.

因为各处密度均匀,因此孤立系的平衡态完全由 (N,E,V) 描述.而对于非平衡宏观态,则除 (N,E,V) 外,还须指出各处的密度 $\rho(\boldsymbol{r},t)$.一般说来,非平衡宏观态应由 (N,E,V,α) 描述,α 代表附加变量,如 $\rho(\boldsymbol{r},t)$.用 $\Omega(N,E,V,\alpha)$ 代表和非平衡宏观态 (N,E,V,α) 对应的可达微观态数,用 $\Omega(N,E,V)$ 代表

和平衡宏观态 (N,E,V) 对应的可达微观态数. 下面我们仅以抽出隔板的瞬间和最终平衡态比较,证明存在着

$$\Omega(N,E,V,\alpha) \leqslant \Omega(N,E,V), \tag{2.1}$$

即孤立系处于平衡宏观态时,对应的可达微观态数最多.

抽出隔板后的瞬间,全部分子仍留在体积 V_i 中,根据 §7.1.3(3.21)式,系统的可达微观态数为

$$\Omega_i = cE^{3N/2}V_i^N\delta E.$$

达到最终的平衡态时,系统的能量 E 和粒子数 N 不变,但分子的活动空间扩大到 V_f,这时的可达微观态数为

$$\Omega_f = cE^{3N/2}V_f^N\delta E, \qquad \Omega_f = (V_f/V_i)^N\Omega_i.$$

即有

$$\Omega_f > \Omega_i.$$

在 §7.1.6 中介绍了微正则分布:平衡孤立系的一切可达微观态出现的概率相等. 现在把这一假设推广到一般情况,即:对于孤立系,和给定约束(N,E,V 一定)相适应的每一微观态出现的概率相等. 既然每个微观态是等概率的,那么各个宏观态就不可能是等概率的. 哪一个宏观态包含的微观态数目多,这个宏观态出现的机会就大. 换句话说,根据微正则分布,孤立系处于非平衡宏观态 (N,E,V,α) 的概率和 $\Omega(N,E,V,\alpha)$ 成比例;而处于平衡宏观态 (N,E,V) 的概率和 $\Omega(N,E,V)$ 成比例,具有最大值. 若通过特殊制备使孤立系最初偏离平衡态(如图 1 所示的方法),则孤立系统会从概率小、微观态数少的宏观态,向概率大、微观态数多的宏观态过渡,直到达到概率最大、微观态数最多的平衡态. 简言之,孤立系统中自发过程具有单向性的原因是,系统要从概率小的宏观态向概率大的宏观态变化,而支配一切不可逆过程的方向和限度的物理量则是和宏观态对应的微观态数 Ω.

(二)熵

代替微观态数 Ω,我们引入熵. 对处于宏观态 (N,E,V,α) 的孤立系统,熵的定义是

$$S(N,E,V,\alpha) = k_B\ln\Omega(N,E,V,\alpha). \tag{2.2}$$

k_B 是一个任意常数. 通过和实验比较(见下节的例 1)可知,它就是玻尔兹曼常数,即

$$k_B = 1.380\times10^{-16}\text{erg}\cdot\text{K}^{-1}$$
$$= 1.380\times10^{-23}\text{J}\cdot\text{K}^{-1}. \tag{2.3}$$

显然,(2.2)式对孤立系统的平衡态也适合,因为平衡态只不过对应于 α 的特殊值而已.

为什么要代替微观态数 Ω 面采用熵 S 来工作呢? 首先,$\ln\Omega$ 是 Ω 的单调函

数. 其次,在§7.1.3中已经指出Ω取巨大的天文数值,而$\ln\Omega$则小得多. 而最根本的原因是,正如我们马上要证明的那样,微观态数Ω是一个相乘量,而$\ln\Omega$则是一个相加量或广延量. 显然,为了描述系统的宏观热力学性质应该采用相加量(广延量)而不是相乘量.

(2.2)式表明,熵的意义和性质是:

1. 熵是宏观系统无序度的量度. 在给定的宏观约束(N,E,V一定)下,若$\Omega(N,E,V,\alpha)=1$则$S(N,E,V,\alpha)=0$,此时系统只有唯一可达微观态,系统的行为完全可以预言,系统完全有序. $\Omega(N,E,V,\alpha)$或$S(N,E,V,\alpha)$愈大,表示宏观系统可供选择的微观态愈多,于是系统行为的可预言性愈小,系统的无序度愈大. 因此,熵是宏观系统无序度的量度.

2. 熵是一个相加量或广延量. 考虑由两个独立子系统A和B组成的复合孤立系$A+B$. 设子系统的微观态数分别为Ω_A和Ω_B. 因为子系统A的每一个微观态可以和子系统B的Ω_B个微观态配合. 因此复合孤立系的微观态数$\Omega_{A+B}=\Omega_A\Omega_B$. 于是复合孤立系的熵为:$S_{A+B}=k_B\ln\Omega_{A+B}=k_B\ln\Omega_A+k_B\ln\Omega_B=S_A+S_B$,即熵是一个相加量或广延量.

3. 熵是态函数. 当孤立系处于平衡态时,其宏观态完全由(N,E,V)描述,则熵$S(N,E,V)$有完全确定的唯一值. 即熵是态函数.

4. 熵和玻尔兹曼常数有相同的量纲和单位. 其原因是Ω没有量纲.

熵是统计物理和热力学中特有的宏观量. (2.2)式是统计物理中一个非常重要的公式,文献上往往称为熵的玻尔兹曼公式,并被刻在玻尔兹曼的墓碑上,以示永远的纪念.

最后,介绍熵的一般定义.

先定义熵算符$\hat{\eta}$

$$\hat{\eta}=-\ln\hat{\rho}.$$

$\hat{\rho}$是密度算符. 根据§7.1.4中的公式,立刻可以写出熵算符$\hat{\eta}$的系综平均,乘上k_B后得到和熵算符$\hat{\eta}$对应的宏观量

$$S=k_B\langle\hat{\eta}\rangle=-k_B\mathrm{tr}(\hat{\rho}\ln\hat{\rho})=-k_B\sum_n(\hat{\rho}\ln\hat{\rho})_{nn}. \tag{2.4}$$

将微正则分布[§7.1.6(6.3)(6.4)式]代入上式,得到孤立系处于平衡时的熵

$$S=-k_B\sum_{\Omega}(1/\Omega)\ln(1/\Omega)=k_B\ln\Omega.$$

这和孤立系中熵的玻尔兹曼公式完全一致. 但(2.4)式有更普遍的意义,不仅适用于孤立系.

(三)熵增加原理

利用孤立系统熵的玻尔兹曼公式(2.2),由(2.1)式可知,当孤立系统由不平

衡趋向平衡时，

$$S(N,E,V,\alpha) \leqslant S(N,E,V). \tag{2.5}$$

或

$$\Delta S_{孤} = S(N,E,V) - S(N,E,V,\alpha) \geqslant 0. \tag{2.6}$$

等式不等式(2.5)或(2.6)就是热力学第二定律的数学表达式.其意义可表述如下:孤立系统的熵永不减少.在真实过程中,孤立系统的熵总是增加,当达到平衡态时,熵增加到最大值,并且不再改变.故往往又把热力学第二定律叫熵增加原理.熵增加原理的微观本质则是,孤立系总是从概率小(微观态数少)的宏观态,向概率大(微观态数多)的宏观态改变,直到达到概率最大(微观态数最多)的平衡态.即:孤立系总是从有序趋向无序,直到达到最无序的平衡态.

以上是从分析气体自由膨胀这一特殊的不可逆过程导出熵增加原理的.在本章第四节中,我们将通过具体计算证明,热量从低温自动传到高温,热机把从热源吸收的热量全部用来做功而不产生其他影响,都是违背熵增加原理的.换言之,熵增加原理包括了热力学第二定律的各种特殊表述,如克氏和开氏表述,为判断孤立系中一切不可逆过程进行的方向和限度,提供了一个共同的、深刻的判据.

最后,对热力学第一、第二定律作几点比较说明:①第一、第二定律都是关于能量传递和转换的规律,但第一定律涉及的是能量的数量,而第二定律涉及的是方向和限度.②第一定律指出了各种形式能量的共性,而第二定律则指出了内能(宏观系统内的无序运动能量)的特殊性.③能量守恒定律是一个确定论性的规律,对单个或少量粒子系统和大量粒子组成的宏观系统均成立,而第二定律则是一个统计性的规律.首先,它只对由大量粒子组成的宏观系统有效;其次,因为存在着涨落,孤立系统的熵并非绝对不能减少,只是减少的概率小到实际观察不到而已.

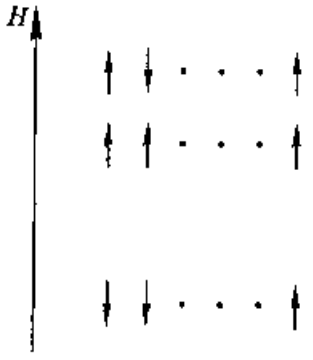


图2　晶格点阵自旋系统

【例1】　理想自旋系统的熵.系统由 N 个自旋为1/2,固有磁矩为 μ_0 的粒子组成.粒子位于晶格点阵的 N 个格点上,如图2,忽略粒子磁矩间的相互作用.系统置于外磁场 H 中,每个粒子的磁矩有两种可能取向:平行于外场(能量为 $-\mu_0 H$)和

反平行于外场(能量为$\mu_0 H$).求当有 n 个粒子的磁矩平行于外场时,系统的能量和熵.

解　当有 n 个粒子的磁矩平行于外场时,系统的能量为

$$E(n) = (N-n)\mu_0 H - n\mu_0 H = (N-2n)\mu_0 H.$$

应当注意,磁矩平行于外场的 n 个粒子在 N 个格点上有许多不同的排列方法.每一种不同的排列方法对应于具有相同能量 $E(n)$ 的不同微观态.

利用排列组合知识容易证明 n 个粒子在 N 个格点上作不同的排列方法的总数,即和 $E(n)$ 对应的微观态数为

$$\Omega(n,N,V) = N!/n!(N-n)!, \tag{2.7}$$

将(2.7)式代入熵的玻尔兹曼公式,并考虑斯特林近似公式[见附录(V.20)式]:

$$\ln N! \approx N\ln N - N, \quad (N > 10), \tag{2.8}$$

可得熵 S 对 n 的依赖,

$$S(n) = k_B[N\ln N - n\ln n - (N-n)\ln(N-n)]. \tag{2.9}$$

对于以下几个特殊情况,由(2.7)式和(2.9)式可得:

$$\begin{aligned} &n = N, && E = -N\mu_0 H, && \Omega = 1, && S = 0. \\ &n = N/2, && E = 0 && \Omega = 2^N, && S = Nk_B\ln 2. \\ &n = 0, && E = N\mu_0 H, && \Omega = 1, && S = 0. \end{aligned} \tag{2.10}$$

(2.10)式表明,当 N 个粒子的磁矩全部平行或反平行于外场时,系统均只有一个微观态,系统完全有序,熵为零;当各有一半粒子的磁矩平行和反平行于外场时,对应的微观态数最多,系统最无序,熵有极大值.以上分析可用图 3 更直观地表示出来.

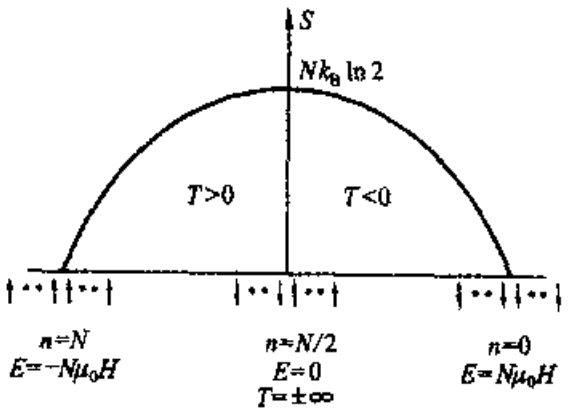


图 3　理想自旋系统的熵

为后面讨论的需要,对(2.7)式作进一步的近似计算.

引入 $m = [n-(N-n)]/2$,或 $n = (N/2)+m$ 和 $N-n = (N/2)-m$,(2.7)式和(2.9)式化为

$$\begin{aligned} &\Omega = N!/(N/2+m)!(N/2-m)!, \\ &S \approx k_B\ln\Omega = k_B\{\ln N! - [\ln(N/2+m)! + \ln(N/2-m)!]\}. \end{aligned} \tag{2.11}$$

因为

$$(N/2+m)! = (N/2)!(N/2+1)\cdots(N/2+m),$$

$$\ln(N/2+m)! = \ln(N/2)! + \sum_{s=1}^{m}\ln(N/2+s).$$

类似有

$$\ln(N/2-m)! = \ln(N/2)! - \sum_{s=1}^{m}\ln(N/2-s+1),$$

$$\ln(N/2+m)! + \ln(N/2-m)! \approx 2\ln(N/2)! + \sum_{s=1}^{m}\ln\frac{1+2s/N}{1-2s/N}. \quad (2.12)$$

这里我们已经近似用了 $(N/2-s+1)\approx N/2-s$.

当 $x \ll 1$ 时有

$$\ln(1\pm x)\approx \pm x, \qquad \ln[(1+x)/(1-x)]\approx 2x,$$

于是在 $1 \ll m \ll N$ 的条件下,有

$$\ln[(1+2s/N)/(1-2s/N)]\approx 4s/N, \quad (2.13)$$

$$\sum_{s=1}^{m}\ln\frac{1+2s/N}{1-2s/N}\approx\frac{4}{N}\sum_{s=1}^{m}s\approx\frac{4}{N}\frac{1}{2}m(m+1)\approx\frac{2m^2}{N}, \quad (2.14)$$

代入(2.11)式,得

$$S = k_B\{\ln N! - 2\ln(N/2)! - 2m^2/N\}, \quad (2.15)$$

再注意到

$$E = (N-2n)\mu_0 H = -2m\mu_0 H, \qquad 或 \qquad m = -E/2\mu_0 H,$$

又可把 S 表为 E 的函数

$$S = k_B\{\ln N! - 2\ln(N/2)! - [E^2/(2\mu_0^2 H^2 N)]\}. \quad (2.16)$$

§7.2.3 温度 化学势和压强 热力学基本方程

热力学第一、第二定律是热力学的基本定律.在本节,我们首先利用第一、第二定律研究孤立、均匀(单相)系统中各部分达到平衡的条件,从微观角度引入温度、化学势和压强,进而得到热力学最重要的方程——热力学基本方程.最后,通过对热力学基本方程的讨论,还将得到热力学第二定律的另一种常用的数学表达式.

(一)热接触系统的平衡 温度

由经验知道,两个可以交换能量(无功交换)而处于平衡的系统具有相同的温度.让我们来研究这种平衡是如何建立的,以及从统计物理的观点看,温度意味着什么.

为简单具体起见,设想被研究的孤立、均匀系统是储在由刚性绝热材料制成的盒中的某种气体,并被一位置固定的导热隔板分为两部分,如图 1 所示.两个子系统的体积和粒子数分别为 V_1, N_1 和 V_2, N_2,并均保持不变,而两个子系统的能量

E_1, E_2 则可以改变. 因根据热力学第一定律有 $dE_1 = -dE_2$，所以此孤立系统只有一个独立变量，可取为 E_1. 孤立系统的熵可表为 $S = S(E_1)$. 对每个子系统来说，能量的交换是在粒子数和体积不变下发生的，往往称这两个子系统处于热接触.

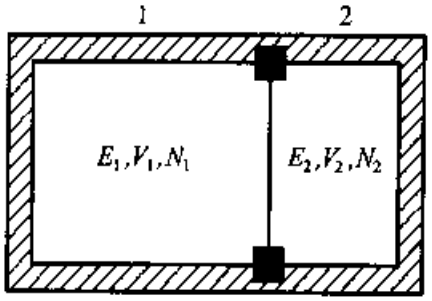


图 1　导热隔板分孤立系为两部分

根据热力学第二定律，此两子系统达到平衡的条件是复合孤立系的熵有最大值. 我们现在的问题是，复合孤立系的总能量如何分配，即：当子系统 1 的能量 E_1 取什么特殊值 $\bar{E}_1$ 时，复合孤立系的熵达到最大？为此，利用极大值的必要条件

$$(\partial S/\partial E_1)_{E_1=\bar{E}_1;N_1,V_1;N_2,V_2} = 0.$$

考虑到熵是广延量，复合孤立系的熵是子系统熵之和，即 $S = S_1 + S_2$，上式化为

$$(\partial S_1/\partial E_1)_{E_1=\bar{E}_1;N_1,V_1} + (\partial S_2/\partial E_2)_{E_2=\bar{E}_2;N_2,V_2}(dE_2/dE_1) = 0.$$

由于 $dE_1 = -dE_2$，最后得

$$(\partial S_1/\partial E_1)_{E_1=\bar{E}_1;N_1,V_1} = (\partial S_2/\partial E_2)_{E_2=\bar{E}_2;N_2,V_2}.$$

$\bar{E}_i$ 是子系统 i 处于平衡时的能量，即 E_i 的系综平均 $\langle E_i\rangle$，在热力学中用 U_i 表示，并称为系统的内能. 故上式又可简洁地表为

$$(\partial S_1/\partial U_1)_{N_1,V_1} = (\partial S_2/\partial U_2)_{N_2,V_2}. \tag{3.1}$$

(3.1)式是两个处于热接触的子系统之间达到平衡，没有净能量传递的条件，叫做热平衡条件. 当(3.1)式不满足时，复合孤立系统的熵将随时间增加

$$\frac{dS}{dt} = \frac{dS_1}{dt} + \frac{dS_2}{dt} = \left(\frac{\partial S_1}{\partial U_1} - \frac{\partial S_2}{\partial U_2}\right)\frac{dU_1}{dt} > 0.$$

若 $\dfrac{\partial S_1}{\partial U_1} > \dfrac{\partial S_2}{\partial U_2}$，则必有 $\dfrac{dU_1}{dt} > 0$ 和 $\dfrac{dU_2}{dt} < 0$.

即：当热接触的子系统未达到平衡时，会有净能量从偏导数 $(\partial S/\partial U)$ 低的子系统转移到偏导数 $(\partial S/\partial U)$ 高的子系统. 根据经验，只有当热接触子系统的温度相等时，复合孤立系才达到平衡；当子系统的温度不相等时，会有能量从高温传到低温. 理论分析和经验比较的自然结论是：偏导数 $(\partial S/\partial U)$ 和系统的温度成反比. 据此引入绝对温度来表示此偏导数

$$1/T = (\partial S/\partial U)_{N,V} = k_B(\partial \ln \Omega/\partial U)_{N,V}. \tag{3.2}$$

于是热平衡条件(3.1)式化为

$$T_1 = T_2. \tag{3.3}$$

对于由(3.2)式定义的绝对温度作几点说明：第一，(3.2)式定义的绝对温度 T，不仅和日常生活中关于“冷”“热”的经验一致，更重要的是，将(3.2)式用于理想气体时，将得到理想气体温度. 换句话说，(3.2)式引入的 T 和科学技术中早已

使用的理想气体温标是一致的. 在随后的例 1 中我们将看到这点. 第二,因为熵 S 和内能 U 都是广延量,因此它们的变化之比或温度是强度量. 第三,温度和熵一样,是一个纯统计性质的宏观量,没有和温度直接对应的动力学量或微观量. 不能通过直接求系综平均而得到温度. 换句话说,温度和熵一样,是统计物理和热力学中特有的物理量. 第四,(3.2)式说明,温度的微观意义是,它反映了熵或微观态数对内能的依赖关系. 通常情况下,微观态数 Ω 是内能 U 的单调增加函数,因而系统具有正的绝对温度. 若存在着某种特殊系统,在一定条件下,随着内能的增加,微观态数 Ω 减少,即:内能愈高,系统愈有序,则该系统将有负绝对温度. 如上节例题中的自旋系统,当处于高能级的自旋粒子数大于处于低能级的粒子数时,系统有负绝对温度. (3.2)式纠正了人们长期以来认为绝对温度只能是正数的观念. 负绝对温度概念的提出,对激光产生机制的解释有重要意义.

(二)处于扩散接触子系统的平衡　化学势

将图 1 中的隔板换为位置依然固定,但不仅允许导热还允许粒子穿过的隔板,即对两个子系统有 N_1, E_1 和 N_2, E_2 均可改变. 由于 $\mathrm{d}N_1 = -\mathrm{d}N_2$ 和 $\mathrm{d}E_1 = -\mathrm{d}E_2$,复合孤立系有两个独立变量,可取为 N_1 和 E_1. 我们称可以自由交换能量和粒子的子系统处于热接触和扩散接触. 现在的问题是,复合孤立系的能量和粒子数怎样分配时,或子系统 1 的能量 E_1 和粒子数 N_1 取什么特殊值 $\tilde{E}_1$ 和 $\tilde{N}_1$ 时,复合孤立系达到平衡,使熵有极大值? 仿照前面的推导,根据极大值的必要条件有

$$\left(\frac{\partial S_1}{\partial E_1}\right)_{E_1=\tilde{E}_1;N_1,V_1}+\left(\frac{\partial S_2}{\partial E_2}\right)_{E_2=\tilde{E}_2;N_2,V_2}(\mathrm{d}E_2/\mathrm{d}E_1)$$

$$+\left(\frac{\partial S_1}{\partial N_1}\right)_{N_1=\tilde{N}_1;E_1,V_1}+\left(\frac{\partial S_2}{\partial N_2}\right)_{N_2=\tilde{N}_2;E_2,V_2}(\mathrm{d}N_2/\mathrm{d}N_1)=0.$$

由于 $\mathrm{d}E_2 = -\mathrm{d}E_1, \mathrm{d}N_2 = -\mathrm{d}N_1$ 且 E_1, N_1 为独立变量,上式成立的条件是

$$\left(\frac{\partial S_1}{\partial E_1}\right)_{E_1=\tilde{E}_1;N_1,V_1}=\left(\frac{\partial S_2}{\partial E_2}\right)_{E_2=\tilde{E}_2;N_2,V_2};\qquad \left(\frac{\partial S_1}{\partial N_1}\right)_{N_1=\tilde{N}_1;E_1,V_1}=\left(\frac{\partial S_2}{\partial N_2}\right)_{N_2=\tilde{N}_2;E_2,V_2}.$$

和前面 $\tilde{E}_i = \langle E_i\rangle = U_i$ 一样,$\tilde{N}_i$ 是子系统 i 处于平衡态时 N_i 的系综平均 $\langle N_i\rangle$,在热力学中就用 N_i 表示,故上式可简洁地表为

$$\left(\frac{\partial S_1}{\partial U_1}\right)_{N_1,V_1}=\left(\frac{\partial S_2}{\partial U_2}\right)_{N_2,V_2};\qquad \left(\frac{\partial S_1}{\partial N_1}\right)_{U_1,V_1}=\left(\frac{\partial S_2}{\partial N_2}\right)_{U_2,V_2}.\tag{3.4}$$

(3.4)式是两个处于热接触和扩散接触的子系统之间达到平衡,没有净能量和净粒子传递的条件. 显然左边的等式表示两个子系统的温度相等. 但是,在这一条件满足时,如果右边的等式不满足,仍然会有扩散. 为了研究在温度均匀的情况下扩散平衡的条件,人们引入化学势的概念

$$\mu_i = -T_i\left(\frac{\partial S_i}{\partial N_i}\right)_{U_i,V_i} = -k_B T_i\left(\frac{\partial \ln \Omega_i}{\partial N_i}\right)_{U_i,V_i}, \qquad i = 1,2. \tag{3.5}$$

于是两个可以交换能量和粒子的子系统达到平衡的条件是

$$T_1 = T_2, \qquad \mu_1 = \mu_2. \tag{3.6}$$

当两个子系统具有相同温度，但尚未到达扩散平衡时，复合孤立系的熵仍将随时间增加，即有：

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \left(\frac{\partial S_1}{\partial N_1} - \frac{\partial S_2}{\partial N_2}\right)\frac{\mathrm{d}N_1}{\mathrm{d}t} = \left(-\frac{\mu_1}{T} + \frac{\mu_2}{T}\right)\frac{\mathrm{d}N_1}{\mathrm{d}t} > 0.$$

若 $\mu_2 > \mu_1$，且系统是正常系统（$T > 0$），则有 $\mathrm{d}N_1/\mathrm{d}t > 0$ 和 $\mathrm{d}N_2/\mathrm{d}t < 0$．即在温度相同的情况下，平均来说，粒子会从化学势 μ 高的地方向化学势 μ 低的地方转移，就象水从高压流向低压，热量从高温传到低温一样．因此可以把化学势解释为粒子转移能力的量度．

考虑由不同原子或分子组成的多元系统，显然以上讨论可以对不同化学组分分别进行．于是可以引入多元系统中，组分 r 的化学势的概念：

$$\mu_r = -T(\partial S/\partial N_r)_{U,V,N_{1\neq r}}. \tag{3.7}$$

脚标 $N_{1\neq r}$ 表示除组分 r 以外的其他组分的粒子数不变．

化学势是和温度同等重要的宏观量，特作以下几点说明：

①化学势和温度一样，没有对应的动力学量或微观量，不能直接通过求微观量的系综平均来得到．换句话说，化学势和熵、温度一样，是热力学和统计物理中特有的、纯统计性质的物理量．

②从(3.5)式可知，因为广延量 S 和广延量 N 变化之比是强度量，因此化学势和温度一样也是强度量，且具有能量的量纲．

③由(3.5)式知道，化学势的微观意义是，它反映了微观态数对粒子数的依赖关系．如果微观态数随着粒子数的增加而增多，则对正常系统（$T > 0$），化学势将具有负值．从推导过程还可看出，对粒子数不守恒的系统，化学势为零．

④从后面的例 1 和例 2 将看到，化学势往往来自两类贡献：粒子的势能和粒子的浓度．

⑤粒子从化学势高的地方向化学势低的地方转移的过程，可能引起物质流，也可能引起原子、分子的重新排列．通过这些过程，系统的不同密度之间，不同相之间及不同的化学组分之间重新建立起平衡．

和温度相比，人们往往对化学势比较生疏．实际上，对带电粒子，它可以用伏特计测量；对中性粒子，可通过测量渗透压的设备来测定，这些和温度计一样都是实验室的常用设备．

(三)处于力学接触于系统的平衡 压强

最后,将图1中的隔板换成可移动的导热隔板.这时,两个子系统的粒子数 N_1 和 N_2 固定,而它们的能量和体积 E_1, V_1, E_2, V_2 均可以改变.但由于 $\mathrm{d}E_1 = -\mathrm{d}E_2$, $\mathrm{d}V_1 = -\mathrm{d}V_2$,复合孤立系有两个独立变量,可取为 E_1 和 V_1.我们称子系统处于热接触和力学接触.

仿照前面的推导,对于两个处于热接触和力学接触于系组成的复合孤立系,熵有极大值的必要条件是:

$$\left(\frac{\partial S_1}{\partial E_1}\right)_{E_1=\bar{E}_1;N_1,V_1} + \left(\frac{\partial S_2}{\partial E_2}\right)_{E_2=\bar{E}_2;N_2,V_2} (\mathrm{d}E_2/\mathrm{d}E_1)$$

$$+\left(\frac{\partial S_1}{\partial V_1}\right)_{V_1=\bar{V}_1;E_1,N_1} + \left(\frac{\partial S_2}{\partial V_2}\right)_{V_2=\bar{V}_2;E_2,N_2} (\mathrm{d}V_2/\mathrm{d}V_1) = 0.$$

因为 $\mathrm{d}E_1 = -\mathrm{d}E_2$, $\mathrm{d}V_1 = -\mathrm{d}V_2$,且 E_1 和 V_1 是独立变量,为使上式成立,必须有

$$\left(\frac{\partial S_1}{\partial U_1}\right)_{N_1,V_1} = \left(\frac{\partial S_2}{\partial U_2}\right)_{N_2,V_2}; \quad \left(\frac{\partial S_1}{\partial V_1}\right)_{U_1,N_1} = \left(\frac{\partial S_2}{\partial V_2}\right)_{U_2,N_2}. \tag{3.8}$$

和前面的讨论类似,在(3.8)式中已考虑到 $\bar{V}_i = \langle V_i \rangle$,而在热力学中就用 V_i 表示.为了研究在温度相等的情况下力学平衡的条件,我们引入物理量 p_i

$$p_i = T_i\left(\frac{\partial S_i}{\partial V_i}\right)_{U_i,N_i} = k_B T_i\left(\frac{\partial \ln \Omega_i}{\partial V_i}\right)_{U_i,N_i}. \tag{3.9}$$

于是,两个热接触和力学接触的子系统达到平衡的条件(3.8)式化为:

$$T_1 = T_2, \qquad p_1 = p_2. \tag{3.10}$$

当两个子系统具有相同温度,但尚未达到力学平衡时,根据熵增加原理,复合孤立系统的熵仍将随时间增加,即有

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \left(\frac{\partial S_1}{\partial V_1} - \frac{\partial S_2}{\partial V_2}\right)\frac{\mathrm{d}V_1}{\mathrm{d}t} = \left(\frac{p_1}{T} - \frac{p_2}{T}\right)\frac{\mathrm{d}V_1}{\mathrm{d}t} > 0.$$

若 $p_1 > p_2$,则有 $\mathrm{d}V_1 > 0$ 和 $\mathrm{d}V_2 < 0$,即在温度相同的情况下,物理量 p 大的子系统的体积将膨胀,而物理量 p 小的子系统的体积会缩小.和日常生活经验比较,显然(3.9)式引进的 p 就是我们熟悉的压强.

虽然(3.9)式和(3.2)、(3.5)式类似,但压强和温度及化学势却十分不同.温度和化学势没有对应的微观量,而压强却有对应的微观量.按照统计物理的基本观点,压强是对应微观量的系综平均.下面按照这一基本观点来推导压强的统计表达式,并证明它和(3.9)式一致.

为具体起见,研究气缸中的气体,并假设气体经历这样一个特殊元过程:当系统的体积从 V 准静态地增加到 $V + \mathrm{d}V$ 时,系统保持在同一个微观态 α(即微扰是绝热变化的,见§6.2.6,于是系统的能量从 $E_\alpha(V)$ 变化到 $E_\alpha(V + \mathrm{d}V)$.根据量子

力学，在立方盒 V 中运动的单粒子处于量子态(n_x,n_y,n_z)时，其能量为

$$\varepsilon_n = [\hbar^2\pi^2/(2mV^{2/3})](n_x^2+n_y^2+n_z^2),$$

见 §5.2.2(2.39)式. 量子态不变而外参量 V 改变时，粒子的能量 ε_n 将改变. 因为系综中每个系统保持在它们的初始微观态，系综中的微观态数没有改变，所以这是一个等熵过程.

系统处在微观态 α 时，作用在活塞上的力为 $F_\alpha = -(\partial E_\alpha/\partial x)_{S,N}$，其中，$x$ 是活塞移动的距离. 用 A 表示活塞的面积，则有 $F_\alpha/A = -(\partial E_\alpha/\partial V)_{S,N}$. 将上式两边同时求系综平均，即同乘以系统处于微观态 α 的概率 P_α，然后对所有可达微观态求和，则得到

$$p = \langle F\rangle/A = \frac{1}{A}\sum_\alpha F_\alpha P_\alpha = -\sum_\alpha\left(\frac{\partial E_\alpha}{\partial V}\right)_{S,N}P_\alpha \tag{3.11}$$

$$\sum_\alpha\left(\frac{\partial E_\alpha}{\partial V}\right)_{S,N}P_\alpha = -\left(\frac{\partial}{\partial V}\sum_\alpha E_\alpha P_\alpha\right)_{S,N} = -\left(\frac{\partial U}{\partial V}\right)_{S,N}. \tag{3.12}$$

即：宏观量压强是微观量 $\partial E_\alpha/\partial V$ 的系综平均(反号)，而这一平均等于系统内能对体积的偏导数.

当系统的粒子数不变时，系统的熵是内能和体积的函数 $S = S(U,V)$. 熵的全微分是 $\mathrm{d}S = (\partial S/\partial U)_V\mathrm{d}U + (\partial S/\partial V)_U\mathrm{d}V$. 考虑等熵过程 $\mathrm{d}S = 0$，并用 $\mathrm{d}V$ 除上式得

$$\left(\frac{\partial S}{\partial U}\right)_V\left(\frac{\partial U}{\partial V}\right)_S + \left(\frac{\partial S}{\partial V}\right)_U = 0.$$

利用(3.2)式和(3.12)式，上式化为 $p/T = (\partial S/\partial V)_{U,N}$. 可见，从压强是微观量 $\partial E_\alpha/\partial V$ 的系综平均出发，重新得到了(3.9)式.

总结以上讨论可知，决定自发过程方向和限度的是系统微观态数对能量、粒子数和体积(或外参量)的依赖，而表示这种依赖关系的强度量是温度、化学势和压强.

【例 1】　理想气体的熵、温度、压强和化学势

解　理论和实验表明，在高温甚至常温下，都可以把单原子理想气体视为经典理想气体，可用 Γ 空间描述单原子分子气体的微观态. 在 §7.1.3 中已经求得经典理想气体的态密度为[见该节的(3.21)式]：

$$\Omega(N,E,V) = \frac{1}{N!h^{3N}}\left(\frac{2\mathrm{e}}{3}2\pi m\right)^{3N/2}\frac{3}{2}\left(\frac{E}{N}\right)^{3N/2-1}V^N.$$

在 $E \sim E+\delta E$ 间的微观态数为

$$\Omega(N,E,V)\delta E = \frac{1}{N!h^{3N}}\left(\frac{2\mathrm{e}}{3}2\pi m\right)^{3N/2}N\frac{3}{2}\left(\frac{E}{N}\right)^{3N/2}V^N\frac{\delta E}{E}. \tag{3.13}$$

将(3.13)式代入§7.2.2(2.2)式,得到单原子理想气体的熵

$$S(N,E,V) = \frac{3N}{2}k_B\ln\left(\frac{mEV^{2/3}}{3\pi\hbar^2}\right) - \frac{5}{2}k_BN\ln N + \frac{5}{2}Nk_B. \tag{3.14}$$

将(3.13)式代入(3.2)式,可得理想气体的温度

$$\frac{1}{T} = k_B\left(\frac{\partial\ln\Omega}{\partial U}\right)_{N,V} = k_B\left\{\frac{3}{2}N\frac{\partial\ln E}{\partial E} + \frac{\partial\ln(\delta E/E)}{\partial E}\right\}_{E=\bar{E}=U}.$$

因为 $\delta E \ll E$,忽略能量间隔的影响,最后得

$$U = (3/2)Nk_BT. \tag{3.15}$$

将(3.13)式代入(3.9)式,得到理想气体的压强

$$p = k_BT\left(\frac{\partial\ln\Omega}{\partial V}\right)_{U,N} = k_BT\left(\frac{\partial\ln V^N}{\partial V}\right)_{U,N} = Nk_BT\frac{1}{V},$$

$$pV = Nk_BT. \tag{3.16}$$

(3.15)和(3.16)式都是熟知的结果,并由此可以看出,由(3.2)式引入的绝对温度正是通常使用的理想气体温度.因此,熵的玻尔兹曼公式中的常数就是玻尔兹曼常数.

将(3.15)式代入(3.14)式,可以把理想气体的熵表为温度的函数

$$S(N,T,V) = k_BN\ln\left[\left(\frac{mk_BT}{2\pi\hbar^2}\right)^{3/2}\left(\frac{V}{N}\right)\right] + \frac{5}{2}Nk_B. \tag{3.17}$$

因为 V/N 是强度量,所以(3.17)式表明熵和粒子数 N 成正比,即熵是广延量.不难看出,若在态总数 $\Omega(N,E,V)$ 的表达式中,没有因考虑粒子全同性而引入的因子 $N!$,则所得到的熵将不满足广延性的要求.事实上早在量子力学发展之前,吉布斯正是为了使熵具有广延性,而硬塞进一个 $N!$ 来克服历史上被称为吉布斯佯谬的困难.

将(3.13)式代入(3.5)式,可求得经典理想气体的化学势.在将(3.13)式代入之前,注意利用(3.15)式将 $(E/N)^{3N/2}$ 换为 $(3k_BT/2)^{3N/2}$,并利用斯特林公式[附录(V.20)式],最后可得

$$\mu(N,T,V) = k_BT\ln\frac{N}{V} + \frac{3}{2}k_BT\ln\left(\frac{2\pi\hbar^2}{mk_BT}\right). \tag{3.18}$$

(3.18)式表明,理想气体的化学势除依赖于温度外,还和浓度 N/V 的对数成正比.当气体各部分温度均匀时,浓度高的区域化学势高.粒子从化学势高的区域向化学势低的区域转移的过程实际上是粒子从高浓度区向低浓度区扩散的过程.

【例2】 理想自旋系统的温度、磁化率和化学势

在§7.2.2的例题中,已经得到理想自旋系统熵的近似表达式,见该节(2.16)式.在 $1 \ll m \ll N$ 的条件下

$$S = k_B\left\{\ln N! - 2\ln\left(\frac{N}{2}\right)! - \frac{E^2}{2\mu_0^2H^2N}\right\} = k_B\left\{N\ln 2 - \frac{E^2}{2\mu_0^2H^2N}\right\}. \quad (3.19)$$

将(3.19)式代入(3.2)式,得到

$$1/T = (\partial S/\partial U)_{N,H} = -k_BU/(\mu_0^2H^2N). \quad (3.20)$$

当自旋系统在温度 T 达到平衡时,其内能和熵为

$$U = -\frac{N\mu_0^2H^2}{k_BT}, \quad S = k_BN\ln 2 - \frac{N\mu_0^2H^2}{2k_B^2T^2} = S_0 - \frac{N\mu_0^2H^2}{2k_B^2T^2}. \quad (3.21)$$

根据 m 的定义(见§7.2.2 例1)可知,自旋系统的能量又可表为 $E = -2m\mu_0H$. 温度为 T 时,系统的内能为

$$U = \langle E\rangle_{系综} = -2\langle m\rangle\mu_0H.$$

和(3.21)式比较

$$\langle m\rangle = N\mu_0H/(2k_BT). \quad (3.22)$$

系统的总平均磁矩为

$$M = 2\langle m\rangle\mu_0 = N\mu_0^2H/(k_BT).$$

磁化率为

$$\chi = (1/V)(\partial M/\partial H) = N\mu_0^2/(Vk_BT). \quad (3.23)$$

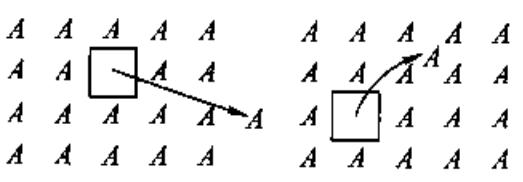
即理想自旋系统的磁化率和外场 H 无关,和温度成反比. 将(3.19)式代入(3.5)式可得 $\mu = -T(\partial S/\partial N)_{U,V} = -k_BT[\ln 2 + U^2/(2\mu_0^2H^2N^2)]$. 利用(3.21)式可进一步把 μ 表为可测参量 T 和 H 的函数

$$\mu(T,H) = -k_BT\ln 2 - \mu_0^2H^2/(2k_BT). \quad (3.24)$$

(3.24)式表明,系统的化学势除依赖于温度外,还依赖于外场或粒子在外场中的势能.

【例3】　肖特基缺陷的密度和温度的关系

真实晶体和理想晶体完整周期结构的偏差叫做缺陷. 没有杂质的理想配比晶体中的缺陷叫做结构缺陷. 点缺陷是结构缺陷中的一种,它会引起晶体体积、比热、扩散和导电等一系列变化. 由于不可避免的热扰动,格点粒子会迁移到表面,而在晶体内部留下空穴(图2),这种点缺陷叫做肖特基缺陷或空穴式缺陷.



(a) 空位式缺陷（肖特基缺陷）　(b) 填隙式缺陷（弗仑克尔缺陷）

图2　肖特基缺陷

设总粒子数为 N 的晶体中形成一个肖特基缺陷(把一个粒子从格点移到表面)所需要的能量为 ε_0. 则当晶体中有 n 个空穴时,系统的内能为 $U = n\varepsilon_0$. 由于这 n 个空穴可以在 $(N+n)$ 个格点上任意分布,其分布方式数或微观态数为 $\Omega(N,$

$U = n\varepsilon_0, V)$ $= (N+n)!/(N!n!)$. 和肖特基缺陷联系的熵为

$$S(N, U = n\varepsilon_0, V) = k_B \ln \Omega(N, U = n\varepsilon_0, V) = k_B \ln \frac{(N+n)!}{N!n!}.$$

利用斯特林公式[附录(V.20)式],得到

$$S(N, U = n\varepsilon_0, V) = k_B\{-N\ln N - n\ln n + (N+n)\ln(N+n)\}. \quad (3.25)$$

当晶体达到平衡时,利用(3.2)式有

$$1/T = (\partial S/\partial U)_{N,V} = [\mathrm{d}S(n)/\mathrm{d}n](\mathrm{d}n/\mathrm{d}U).$$

将 $U = n\varepsilon_0$ 和(3.25)式代入,最后得

$$\frac{1}{T} = \frac{k_B}{\varepsilon_0}\ln\frac{N+n}{n}, \quad \text{或} \quad \frac{n}{N} = \frac{1}{\mathrm{e}^{\varepsilon_0/(k_BT)} - 1}.$$

当 $\varepsilon_0 \gg k_BT$ 或 $n \ll N$ 时,上式简化为

$$n \approx N\mathrm{e}^{-\varepsilon_0/(k_BT)}. \quad (3.26)$$

(3.26)式表示,在平衡状态下,肖特基缺陷数随温度升高而指数增加.

(四)热力学基本方程

考虑一个多元均匀系,用 N_k 表示第 k 个组元的分子数. 设想系统经历任意可逆或不可逆元过程,但过程前后系统处于平衡态,则由态函数熵 $S = S(\{N_k\}, U, V)$,可得

$$\mathrm{d}S = \left(\frac{\partial S}{\partial U}\right)_{\{N_k\},V}\mathrm{d}U + \left(\frac{\partial S}{\partial V}\right)_{\{N_k\},U}\mathrm{d}V + \sum_k\left(\frac{\partial S}{\partial N_k}\right)_{U,V,N_{k'\neq k}}\mathrm{d}N_k.$$

利用(3.2),(3.7)和(3.9)式,上式化为

$$\mathrm{d}S = (1/T)\mathrm{d}U + (p/T)\mathrm{d}V - \sum_k(\mu_k/T)\mathrm{d}N_k.$$

或

$$\mathrm{d}U = T\mathrm{d}S - p\mathrm{d}V + \sum_k \mu_k \mathrm{d}N_k, \quad \text{(可逆或不可逆元过程)}. \quad (3.27)$$

对于和外界没有粒子交换的封闭系统,上式简化为

$$\mathrm{d}U = T\mathrm{d}S - p\mathrm{d}V, \quad \text{(可逆或不可逆元过程)}. \quad (3.28)$$

(3.27)和(3.28)式分别叫做开放均系和封闭均系的热力学基本方程. 因为在引入(3.2),(3.7)及(3.9)式时,同时应用了热力学第一、第二定律,因此热力学基本方程是第一、第二定律联合的结果,和热力学第一、第二定律一样,是热力学中最基本、最重要的关系,是本章以下各节讨论的出发点.

要特别强调的是,热力学基本方程对初、末态为平衡态的任意元过程皆成立,不管过程是否可逆. 但在可逆和不可逆过程中,方程右方各项的意义不同. 仅以封闭均匀系热力学基本方程为例来说明这一点.

我们知道,封闭系统经历任意可逆或不可逆过程时,均遵循热力学第一定律:

$$\mathrm{d}U = \bar{\mathrm{d}}Q + \bar{\mathrm{d}}W \quad \text{（可逆或不可逆元过程）.} \tag{3.29}$$

$\mathrm{d}U$ 代表系统内能的变化，$\bar{\mathrm{d}}Q$ 代表系统从外界吸收的热量，$\mathrm{d}W$ 代表外界对系统所作的功.

当系统是流体时，外界对系统作压缩功. 在可逆元过程中，由于过程无限缓慢且无摩擦，外压强等于系统的压强，外界对系统的功可以用系统的态参量表示，即有[见 §7.2.1(1.3′)式]

$$\mathrm{d}W = -p_{外}\,\mathrm{d}V = -p\mathrm{d}V \quad \text{（可逆元过程）.} \tag{3.30}$$

负号的引入是为了保证压缩时（$\mathrm{d}V < 0$），外界对系统作正功. 将(3.30)式代入(3.29)式，再和热力学基本方程(3.28)式比较，可得

$$\bar{\mathrm{d}}Q = T\mathrm{d}S \quad \text{（可逆元过程）.} \tag{3.31}$$

在不可逆过程中，由于存在摩擦和过程非无限缓慢，外压强不等于系统的压强，外界对系统的功不能用系统的态参量表示，且有(见 §7.2.1)

$$\bar{\mathrm{d}}W > -p\mathrm{d}V \quad \text{（不可逆元过程）.} \tag{3.32}$$

将(3.32)式代入(3.29)式，再和热力学基本方程(3.28)式比较，则得

$$\bar{\mathrm{d}}Q < T\mathrm{d}S \quad \text{（不可逆元过程）.} \tag{3.33}$$

即：虽然封闭系热力学基本方程对可逆和不可逆元过程皆成立，但在两种情况下，方程右方各项意义不同. 在可逆元过程中，$-p\mathrm{d}V$ 等于外界对系统的功，$T\mathrm{d}S$ 等于系统从外界吸收的热量；但在不可逆元过程中，$-p\mathrm{d}V$ 小于外界对系统的功，$T\mathrm{d}S$ 大于系统从外界吸收的热量. 例如，气体自由膨胀是一个不可逆过程，在自由膨胀中，虽然有 $\mathrm{d}U = \bar{\mathrm{d}}Q = \mathrm{d}W = 0$，但显然有 $-p\mathrm{d}V < 0$ 和 $T\mathrm{d}S > 0$.

设除 $-p\mathrm{d}V$ 外，外界对系统还有其他功. 用 $\sum_i Y_i\mathrm{d}x_i$ 表示在可逆元过程中，外界对系统所作的各种元功之和[见 §7.2.1(1.8)式]，$Y_i = \langle \partial E_\alpha/\partial x_i \rangle = T(\partial S/\partial x_i)_U$ 叫广义力，x_i 是和 Y_i 共轭的广义坐标，则封闭系的热力学基本方程可写为更一般的形式

$$\mathrm{d}U = T\mathrm{d}S + \sum_i Y_i\mathrm{d}x_i. \tag{3.34}$$

在固体物理和低温物理中，最重要的是外磁场对介质的磁化功 $\boldsymbol{H}\cdot\mathrm{d}\boldsymbol{M}$ 和电场对电介质的极化功 $\boldsymbol{E}\cdot\mathrm{d}\boldsymbol{P}$. $\boldsymbol{H}$ 和 $\boldsymbol{E}$ 分别代表外场的磁场强度和电场强度，$\boldsymbol{M}$ 和 $\boldsymbol{P}$ 分别代表系统的总磁矩和总电极矩. 因为磁化和极化过程的弛豫时间很短，所以一般都可当成可逆过程处理.

若用

$$\boldsymbol{H}\cdot\mathrm{d}\boldsymbol{M} = \mathrm{d}(\boldsymbol{H}\cdot\boldsymbol{M}) - \boldsymbol{M}\cdot\mathrm{d}\boldsymbol{H}$$

代入(3.34)式，并忽略介质体积的变化，则磁介质系统的热力学基本方程可化为

$$\mathrm{d}U^* = \mathrm{d}(U - \boldsymbol{H}\cdot\boldsymbol{M}) = T\mathrm{d}S - \boldsymbol{M}\cdot\mathrm{d}\boldsymbol{H}. \tag{3.35}$$

因为 $-\boldsymbol{H}\cdot\boldsymbol{M}$ 是介质在外场中的势能，故(3.35)式是将介质在外场中的势能做为系统内能 U^* 的一部分时，热力学基本方程的表达式. 对电介质有类似情况. 以上分析说明，一般说来，系统热力学基本方程的表达式不是唯一的，它依赖于系统内能的定义.

如果外界对系统有多种功，且系统含有多种组元，则开放均匀系统的热力学基本方程具有最一般的形式：

$$\mathrm{d}U = T\mathrm{d}S + \sum_i Y_i \mathrm{d}x_i + \sum_k \mu_k \mathrm{d}N_k. \tag{3.36}$$

$\sum_k \mu_k \mathrm{d}N_k$ 代表各组元粒子数改变所引起的内能变化，往往又称为化学功.

从以上讨论中，还可以得到热力学第二定律的另一种常用的数学表达式. 将(3.31)和(3.33)式合并得

$$\mathrm{d}S \geqslant \frac{\mathrm{d}Q}{T} \quad \begin{pmatrix} \text{等号对应于可逆元过程} \\ \text{大于号对应于不可逆元过程} \end{pmatrix}. \tag{3.37}$$

令 $\mathrm{d}Q = 0$，则有

$$\mathrm{d}S \geqslant 0 \quad \begin{pmatrix} \text{等号对应于可逆元过程} \\ \text{大于号对应于不可逆元过程} \end{pmatrix}. \tag{3.38}$$

(3.38)式表示：当经历可逆绝热过程时，系统的熵不变；经历不可逆绝热过程时，系统的熵增加. 孤立系统中的过程一定是不可逆绝热过程，因而孤立系统的熵总是增加.

§7.2.4　熵差的计算

熵是统计物理和热力学中特有的一个非常重要的物理量. 熵的玻尔兹曼公式 $S = k_B \ln \Omega(N,E,V)$ 说明了宏观量熵的微观本质，也给出了求熵值的方法. 但对于大多数实际系统，$\Omega(N,E,V)$ 的计算十分困难，因此，从统计的角度出发用玻尔兹曼公式求熵不是一种普遍易行的方法. 不过在许多实际问题中，我们往往感兴趣的不是系统在某一平衡态时熵的绝对数值，而是末、始平衡态的熵差. 根据熵是态函数和广延量的性质，利用热力学第二定律的表达式 $\mathrm{d}S = \mathrm{d}Q/T$(可逆元过程)，可以方便地在不考虑系统微观结构的情况下计算任何系统的熵差. 本节结合几个例子介绍计算熵差的两种热力学方法. 通过例子，我们还将看到，熵增加原理既包含了开氏表述，也包含了克氏表述，的确是热力学第二定律最普遍和最本质的表述.

(一)利用熵函数表达式求熵差

若已知熵作为态参量函数的表达式，将末、始平衡态的态参量代入，立即可得末、始态的熵差. 可见，这种方法的关键是用热力学方法确定系统熵函数的表达式.

作为例子，我们从热力学定律出发推导经典理想气体熵函数的表达式. 显然，态参量选取不同，熵函数有不同表达式. 下面分别以 T,V 和 T,p 为态参量进行推导.

【例 1】 设有 n mol 经典理想气体经历一个元过程，根据热力学第一定律有 $dQ = dU - dW$. 若元过程是可逆过程，则有 $dQ = TdS$ 和 $dW = -pdV$，代入上式，得[亦即热力学基本方程 §7.2.3(3.28)式]

$$dS = (dU + pdV)/T.$$

对经典理想气体有 $dU = C_V dT$ 和状态方程 $pV = nRT$，上式化为

$$dS = C_V dT/T + nRdV/V. \tag{4.1}$$

选取 (T_0,V_0) 为参考态，令其熵为 S_0^*，并设当温区 (T_0,T) 不大时，C_V 是常数，积分上式可得

$$S(T,V) = \int_{T_0}^{T} C_V \frac{dT}{T} + nR\int_{V_0}^{V} \frac{dV}{V} + S_0^*(T_0,V_0) = C_V \ln T + nR\ln V + S_0. \tag{4.2}$$

其中 $S_0 = S_0^*(T_0,V_0) - C_V \ln T_0 - nR\ln V_0$，是由所选参考态决定的一个常数.

(4.2)式是以 T,V 为态参量，(T_0,V_0) 为参考态时，经典理想气体熵函数的表达式. 将末、始态的态参量 (T_f,V_f) 和 (T_i,V_i) 代入(4.2)式，并选取同一个参考态 (T_0,V_0)，然后相减即得末、始态的熵差，而依赖参考态的常数 S_0 被抵消掉.

由经典理想气体的状态方程可得 $dV/V = dT/T - dp/p$，代入(4.1)式有

$$dS = (C_V + nR)\frac{dT}{T} - nR\frac{dp}{p} = C_p\frac{dT}{T} - nR\frac{dp}{p}. \tag{4.3}$$

取 C_p 为常数和 (T_0,p_0) 为参考态，积分上式得

$$S(T,p) = C_p \ln T - nR\ln p + S'_0. \tag{4.4}$$

其中 $S'_0 = -C_p \ln T_0 + nR\ln p_0 + S_0^*(T_0,p_0)$. 类似可得

$$S(p,V) = C_p \ln V + C_V \ln p + S''_0. \tag{4.5}$$

(4.2)，(4.4)和(4.5)式分别是以 (T,V)，(T,p) 和 (p,V) 为态参量时，经典理想气体熵函数的表达式. 要特别注意的是，S_0，S_0' 和 S_0'' 虽可都由同一个参考态决定，但其值不相等.

由以上推导过程看到，只有已知系统的状态方程和热容量，才能求得系统熵函数的表达式. 因此，这种方法的应用范围受到限制.

(二)利用热温比积分求熵差

设系统经某一可逆过程从平衡态 a 到达平衡态 b. 将热力学第二定律的数学表达式 $dS = dQ/T$(可逆过程)用于该过程，并进行积分，则可得 b 态和 a 态的熵差

$$S_b - S_a = \int_a^b dS = \int_a^b \frac{dQ}{T} \quad (\text{可逆过程}). \tag{4.6}$$

设系统经一不可逆过程从平衡态 a 到达平衡态 b，那么如何求熵差 $(S_b - S_a)$ 呢？因为熵是态函数，只要初、末平衡态给定，则熵差有唯一确定值，和系统经由什么过程（可逆或不可逆过程）从初态到达末态无关. 因此，可以设计一个和原不可逆过程具有相同的初、末态，且便于积分的可逆过程，将(4.6)式按所设计的可逆过程计算热温比积分. 所得结果就是原不可逆过程的熵差.

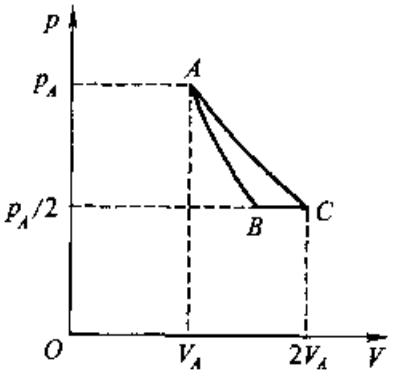


图 1　等温膨胀的熵变

【例 2】　取 n mol 的某种理想气体，初态 A 和末态 C 具有相同的温度，且 C 态的体积是 A 态体积的两倍，如图 1 所示. (1) 试用熵函数表达式(4.5)计算熵差 $(S_C - S_A)$. (2) 按以下两种可逆过程计算熵差：由 A 经等温过程到 C；由 A 经绝热到 B，然后经等压过程从 B 到 C.

解　1) 用方法 I：选取同一参考态，对 A 和 C 用(4.5)式，有

$$S_C = C_p \ln V_C + C_V \ln p_C + S_0'' = C_p \ln 2V_A + C_V \ln (p_A/2) + S_0'',$$

$$S_A = C_p \ln V_A + C_V \ln p_A + S_0'',$$

$$S_C - S_A = (C_p - C_V) \ln 2 = nR \ln 2 > 0.$$

不难证明，用(4.2)和(4.4)式会得到相同的结果.

2) 用方法 II：首先考虑等温可逆过程，这时有

$$S_C - S_A = \int_A^C \frac{\mathrm{d}Q}{T} = \int_A^C \frac{\mathrm{d}U + p\mathrm{d}V}{T} = \int_A^C \frac{p\mathrm{d}V}{T}$$

$$= nR\int_A^C \frac{\mathrm{d}V}{V} = nR \ln \frac{V_C}{V_A} = nR \ln 2.$$

再考虑 $A \to B \to C$ 过程，因为 $A \to B$ 是可逆绝热过程，$S_B - S_A = 0$，所以

$$S_C - S_A = S_C - S_B = \int_B^C \frac{\mathrm{d}Q}{T} = \int_B^C C_p \frac{\mathrm{d}T}{T} = C_p \ln \frac{T_C}{T_B}.$$

设 B 态体积为 V_B，按理想气体绝热方程有 $T_B V_B^{\gamma-1} = T_A V_A^{\gamma-1} = T_C V_A^{\gamma-1}$，或

$$T_C/T_B = (V_B/V_A)^{\gamma-1}.$$

其中 $\gamma = C_p/C_V$. 将等压过程方程 $V_B = (T_B/T_C)V_C$ 代入上式

$$\frac{T_C}{T_B} = \left(\frac{T_B}{T_C}\frac{V_C}{V_A}\right)^{\gamma-1} \quad 或 \quad \frac{T_C}{T_B} = \left(\frac{V_C}{V_A}\right)^{\gamma-1/\gamma}.$$

最后得

$$S_C - S_A = C_p \ln (V_C/V_A)^{\gamma-1/\gamma} = [(\gamma - 1)/\gamma] C_p \ln 2 = nR \ln 2.$$

计算表明，只要末、始态相同，不论沿怎样的可逆路径，熵差都是一样的.

【例 3】　如图 2 所示，容器被分隔为两部分. 开始 A 部充满理想气体，B 部为真空. 整个容器与外界孤立. 当抽去隔板的瞬时，在 A 部的理想气体处于平衡态，但整体（A，B 两部分）处于非平衡态. 以后，由于气体自由膨胀，最终达到平衡态. 试计算理想气体熵的变化.

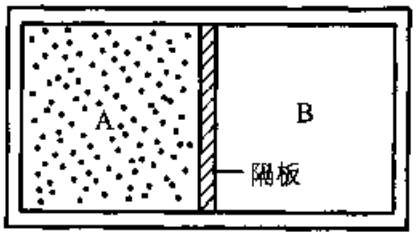


图 2　自由膨胀的熵变

解　因为在自由膨胀中有 $\mathrm{d}Q = 0, \mathrm{d}W = 0$，故 $\mathrm{d}U = 0$. 已知理想气体的内能只是温度的函数，所以在自由膨胀中理想气体的温度保持不变. 过程前后理想气体的态参量分别是状态 $1(T, V_A)$ 和状态 $2(T, V_A + V_B)$. 若以为由 $\mathrm{d}Q = 0$ 可求得 $S_2 - S_1 = \int_1^2 \mathrm{d}Q/T = 0$，那将是错误的. 因为，自由膨胀是孤立系统中的自发不可逆过程，因而 $\mathrm{d}S = \mathrm{d}Q/T$ 不适用. 需要设计一个初、末态仍为 1，2 的可逆过程，才能通过热温比积分求熵差. 显然，选择等温可逆过程最方便. 于是有

$$S_2 - S_1 = \int_1^2 \frac{\mathrm{d}Q}{T} = \int_1^2 \frac{\mathrm{d}U + p\mathrm{d}V}{T} = nR\int_1^2 \frac{\mathrm{d}V}{V} = nR\ln\frac{V_A + V_B}{V_A} > 0.$$

可见，气体自由膨胀后熵增加，的确是一个不可逆过程.

【例 4】　将 1kg 100℃的水和 1kg 0℃的水直接混合，试判断过程的可逆性.

解　直接混合过程可能很复杂，为了计算熵的变化，需要设计两个过程，分别使 100℃的水和 0℃的水可逆地变到共同末温 50℃. 设 373K 和 323K 之间有一系列温差无限小的热源. 首先将 100℃的水依次与 373K − $\mathrm{d}T$，373K − $2\mathrm{d}T$，…，的热源接触，将 $\mathrm{d}Q$ 的热量放给这些热源，如图 3 所示.

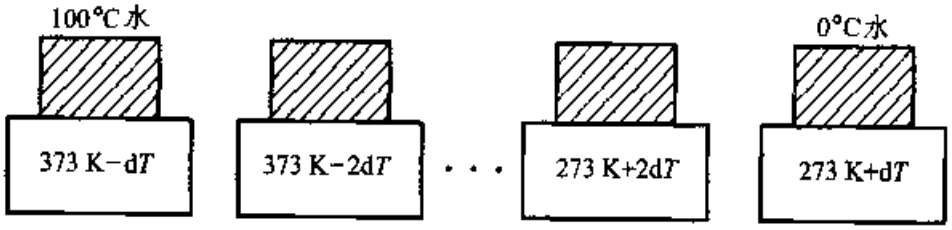


图 3　热传导过程的熵变

由于水和热源的温差无限小，放热过程无限缓慢，水内部各处温度均匀，放热过程是可逆的. 于是 1kg 100℃的水的熵变为

$$\Delta S_1 = \int_{373}^{323} \frac{\mathrm{d}Q}{T} = c_p\int_{373}^{323} \frac{\mathrm{d}T}{T} = c_p\ln\frac{323}{373} = -603\ \mathrm{J\cdot K^{-1}}.$$

类似地，将 0℃的水依次与 273 + $\mathrm{d}T$，273 + $2\mathrm{d}T$，…，的热源接触，最后达到 50℃. 在这一可逆吸热过程中，1kg 0℃的水的熵变为

$$\Delta S_2 = \int_{273}^{323} \frac{\mathrm{d}Q}{T} = c_p \int_{273}^{323} \frac{\mathrm{d}T}{T} = c_p \ln \frac{323}{273} = 704\ \mathrm{J \cdot K^{-1}}.$$

两部分水组成一个大的绝热系，其总熵变为

$$\Delta S = \Delta S_1 + \Delta S_2 = -603 + 704\ \mathrm{J \cdot K^{-1}} > 0.$$

这表明温度不同的水的混合是不可逆过程. 这是热传导过程不可逆的例子.

【例 5】 设有一温度为 T 的热源，一台热机循环工作时，从热源吸收了 ΔQ 的热量全部用来对外作功，试求热源和热机组成的大系统的熵变.

解 热机中的工质经历循环过程，因而熵不变，$\Delta S_1 = 0$. 热源是一个热容为无限大的物体，它有两个特点：首先，不论吸收或放出多少热量，热源的温度保持不变；其次，尽管与热源交换热量的物体可能发生不可逆变化，总可以认为热源经历可逆过程. 因此，热源的熵变是 $\Delta S_2 = \int \mathrm{d}Q/T = -|\Delta Q|/T$. 由热机和热源组成的大绝热系的熵变是

$$\Delta S = \Delta S_1 + \Delta S_2 = -|\Delta Q|/T < 0.$$

结果违反熵增加原理，因而是不可能的. 即从单一热源吸热，全部用来作功而不产生其他影响是不可能的.

以上例子说明，熵增加原理既包括了热力学第二定律的克氏表述（见例 4），又包括了它的开氏表述（见例 5），完整地表达了热力学第二定律的内容，为判断孤立和绝热系中一切不可逆过程进行的方向和限度提供了一个虽然比较抽象，却最普适、最深刻的判据.

§7.2.5 热力学势 麦克斯韦关系及它们的简单应用

至此，我们已从统计物理角度引入研究任何宏观系统热力学性质所需的全部热力学原理. 在应用这些原理解决实际问题的时候，利用数学中变换变量的方法——勒让德变换，引入某些辅助热力学函数，然后利用热力学基本方程导出包括麦克斯韦关系在内的一套对均匀系普遍适用的热力学关系，可使研究方便得多. 由于这些辅助热力学函数在计算作功和判断系统稳定性方面，和保守力系中的势能有相似的作用，所以往往被称为热力学势. 普遍热力学关系的好处是，利用它们可以把均匀系的一切热力学性质和几个可用实验测量的量，主要是热容量和状态方程（或与状态方程有关的系数 α 和 κ）联系起来.

为了简单起见，在引入热力学势和普遍热力学关系时，我们都以均匀封闭可压缩的流体系统为例进行讨论，但所用方法很容易推广到有非压缩功，有多项功的均匀闭系，甚至开放系.

(一)热力学势

在§1.6.1中已经介绍了勒让德变换,见那一节的(1.2)~(1.4)式.在现在的情况下,对应于§1.6.1(1.2)式,我们有均匀封闭流体系统的热力学基本方程

$$dU = TdS - pdV. \tag{5.1}$$

利用勒让德变换可将(5.1)式改写为另外三种等价形式.交换(5.1)式中 p 和 V 的地位,即将 $pdV = d(pV) - Vdp$ 代入,则(5.1)式化为

$$dH = TdS + Vdp. \tag{5.2}$$

其中 H 就是§7.2.1中引入的焓,它的定义是§7.2.1(1.16)式

$$H \equiv U + pV. \tag{5.3}$$

用同样的方法,交换(5.1)式中 T 和 S 的地位,得

$$dF = -SdT - pdV, \tag{5.4}$$

其中 F 叫做亥姆霍兹自由能或简称自由能,它的定义是

$$F \equiv U - TS. \tag{5.5}$$

在(5.1)式中同时交换 T 和 S,p 和 V 的地位,得

$$dG = -SdT + Vdp, \tag{5.6}$$

其中 G 叫做吉布斯自由焓或简称自由焓,它的定义是

$$G \equiv U + pV - TS = H - TS. \tag{5.7}$$

当状态给定时,态参量 p,V,T 及态函数 U,S 都有唯一确定值,因而 H,F 和 G 也有唯一确定值.换句话说,H,F 和 G 都是态函数.因为 U,S 和 V 是广延量,T,p 是强度量,于是,由式(5.3)、(5.5)和(5.7)式定义的 H,F 和 G 都是广延量.显然,式(5.2)、(5.4)和(5.6)都是均匀封闭流体系统的热力学基本方程,它们和(5.1)式等价.

关于 H,F 和 G 的物理意义以及它们为什么叫做热力学势,留到§7.2.7中讨论.

(二)麦克斯韦关系

(5.1)式是以 S 和 V 为独立变量表示的全微分 dU,$U = U(S,V)$,于是可以写出[参看§1.6.1(1.1),(1.2)式]:

$$dU = (\partial U/\partial S)_V dS + (\partial U/\partial V)_S dV.$$

和(5.1)式比较,得到两个偏导数

$$T = (\partial U/\partial S)_V, \qquad p = -(\partial U/\partial V)_S. \tag{5.8}$$

类似地,由 $H = H(S.p)$,$F = F(T,V)$ 和 $G = G(T,p)$,可写出 $dH = (\partial H/\partial S)_p dS + (\partial H/\partial p)_S dp$,$dF = (\partial F/\partial T)_V dT + (\partial F/\partial V)_T dV$ 和 $dG = (\partial G/\partial T)_p dT + (\partial G/\partial p)_T dp$,将它们和(5.2)、(5.4)、(5.6)式比较,得到另外六个热力学偏导数

$$T = (\partial H/\partial S)_p, \qquad V = (\partial H/\partial p)_S, \tag{5.9}$$

$$S = -(\partial F/\partial T)_V, \qquad p = -(\partial F/\partial V)_T, \tag{5.10}$$

$$S = -(\partial G/\partial T)_p, \qquad V = (\partial G/\partial p)_T. \tag{5.11}$$

如果有全微分 $df = adx + bdy$，则必有 $(\partial a/\partial y)_x = (\partial b/\partial x)_y$，称为全微分条件. 将全微分条件用于(5.1)、(5.2)、(5.4)和(5.6)式，得到(5.8)～(5.11)式所给八个热力学偏导数的关系：

$$(\partial T/\partial V)_S = -(\partial p/\partial S)_V, \tag{5.12}$$

$$(\partial T/\partial p)_S = (\partial V/\partial S)_p, \tag{5.13}$$

$$(\partial p/\partial T)_V = (\partial S/\partial V)_T. \tag{5.14}$$

$$(\partial V/\partial T)_p = -(\partial S/\partial p)_T. \tag{5.15}$$

式(5.12)～(5.15)叫做麦克斯韦关系.

以上结果归纳在表Ⅰ中，它们往往被称为均匀封闭系的普遍热力学关系. 掌握和运用这一套普遍热力学关系是热力学方法的核心内容之一. 只要记住热力学基本方程(5.1)式，利用勒让德变换和全微分条件，可以很容易地把整个表的内容推导出来. 为了帮助记忆表Ⅰ的内容，一些教材介绍了各种“记忆图”，读者可查阅有关参考书，也可自己设计方便的“记忆图”.

表Ⅰ　均匀封闭系的普遍热力学关系

热力学势	热力学基本方程	热力学偏导数	麦克斯韦关系
U	$dU = TdS - pdV$	$T = \left(\frac{\partial U}{\partial S}\right)_V \quad p = -\left(\frac{\partial U}{\partial V}\right)_S$	$\left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial p}{\partial S}\right)_V$
$H \equiv U + pV$	$dH = TdS + Vdp$	$T = \left(\frac{\partial H}{\partial S}\right)_p \quad V = \left(\frac{\partial H}{\partial p}\right)_S$	$\left(\frac{\partial T}{\partial p}\right)_S = \left(\frac{\partial V}{\partial S}\right)_p$
$F \equiv U - TS$	$dF = -SdT - pdV$	$S = -\left(\frac{\partial F}{\partial T}\right)_V \quad p = -\left(\frac{\partial F}{\partial V}\right)_T$	$\left(\frac{\partial p}{\partial T}\right)_V = \left(\frac{\partial S}{\partial V}\right)_T$
$G \equiv H - TS$	$dG = -SdT + Vdp$	$S = -\left(\frac{\partial G}{\partial T}\right)_p \quad V = \left(\frac{\partial G}{\partial p}\right)_T$	$\left(\frac{\partial V}{\partial T}\right)_p = -\left(\frac{\partial S}{\partial p}\right)_T$

对于表Ⅰ所列均匀封闭系的普遍热力学关系，我们强调以下几点：

①这套关系是从热力学基本方程推导出来的，因此和基本方程一样，对可逆和不可逆过程皆适用，只要初、末态是平衡态.

②利用这套关系，加上系统的状态方程(或和状态方程等价的系数 β 和 α)以及可由实验测得的热容量 C_p、C_V 之值和几个简单的偏微分公式，就可以很方便地研究均匀闭系的各种宏观性质，而不必考虑系统的微观结构和粒子的运动. 随后的例子将清楚地说明这一点.

③麦克斯韦关系有以下几个特点：首先，它们是“热学量” T,S 对“力学量” p,V 偏导数之间的关系；其次，可以证明，四个麦克斯韦关系并非彼此独立，从其中一个可推出其他三个. 麦克斯韦关系的作用在于，把表面上完全无关的热力学偏导数联系起来，特别是最后两个，它们把不能测量的偏导数 $(\partial S/\partial V)_T$ 和 $(\partial S/\partial p)_T$，和可测量的等容压强系数 $\beta=(1/p)(\partial p/\partial T)_V$ 及等压膨胀系数 $\alpha=(1/V)(\partial V/\partial T)_p$ 联系起来了.

④显然，把 p 和 V 换成其他广义力和广义坐标，表中全部公式依然有效.

⑤在 §7.1.2 中曾强调指出，对于给定系统，描述平衡态的独立变量（态参量）数是确定的，而选择哪几个宏观量作为态参量则是任意的. 改变变量的方法是勒让德变换. 改变的依据则要考虑系统所经过程的条件及处理的方便. 这就是我们把热力学基本方程(5.1)化为其他三个等价形式的指导思想. 在 §7.2.7 中讨论最大功原理及平衡和平衡稳定性判据时，将继续运用这一思想. 实际上，这一思想在整个热力学乃至系综理论中都至关重要.

（三）特性函数

在 §7.1.2 中已经指出，平衡态的特点不仅在于描述系统整体性质的宏观量不再随时间变化，还在于这些宏观量之间存在着函数关系. 只有少数几个是独立态参量，其余宏观量则是态参量的函数. 对于给定的系统，其态参量的数目是确定的，但选哪几个宏观量做为态参量则是任意的. 1869 年马休证明，对于适当选择的独立态参量，只要知道一个态函数，就可以求得全部热力学量，这样的态函数叫做特性函数. 特性函数的概念在热力学和系综理论中都十分有用. 结合表 I 的内容，很容易说明特性函数的概念. 首先以自由能 F 为例来解释.

若已知以 T,V 为独立态参量时自由能的表达式 $F=F(T,V)$，由(5.10)式可求得系统的熵 $S=-(\partial F/\partial T)_V$ 和压强 $p=-(\partial F/\partial V)_T$. 显然，后者是 p,T,V 之间的关系，即系统的状态方程. 将已知的 $F=F(T,V)$ 和求得的 $S(T,V),p(T,V)$ 一道代入 F,G 和 H 的定义式，可得

$$U=F+TS=F-T(\partial F/\partial T)_V, \tag{5.16}$$

$$G=F+pV=F-V(\partial F/\partial V)_T, \tag{5.17}$$

$$H=U+pV=F-T(\partial F/\partial T)_V-V(\partial F/\partial V)_T. \tag{5.18}$$

这说明，只要已知以 T,V 为独立态参量的自由能的表达式 $F=F(T,V)$，就可求得 S、状态方程、p、U、G、H 乃至系统的全部热力学量. 所以，自由能 F 是以 T,V 为独立态参量的特性函数. 在下一章我们将看到这一特性函数在正则系综中的重要作用.

再来证明自由焓 G 是以 T,p 为独立态参量的特性函数. 若已知 $G=G(T,p)$，利用(5.11)式可求得熵和状态方程：$S=-(\partial G/\partial T)_p$，$V=(\partial G/\partial p)_T$. 将已知的 G

$= G(T,p)$ 和求得的 $S(T,p)$，$V(T,p)$ 一道代入 G，H 和 F 的定义式，可得

$$H = G + TS = G - T(\partial G/\partial T)_p, \tag{5.19}$$

$$U = H - pV = G - T(\partial G/\partial T)_p - p(\partial G/\partial p)_T, \tag{5.20}$$

$$F = G - pV = G - p(\partial G/\partial p)_T. \tag{5.21}$$

这说明，已知以 T,p 为独立态参量的自由焓的表达式 $G = G(T,p)$，可求得状态方程、S、H、U、F，乃至系统的全部热力学量.

用类似方法不难证明，内能 U 和焓 H 分别是以 S,V 和 S,p 为独立态参量的特性函数.

〔四〕普遍热力学关系的简单应用

研究系统的宏观性质往往可归结为，求某些有意义的热力学偏导数，如表示绝热膨胀时温度对体积依赖关系的偏导数 $(\partial T/\partial V)_S$；热力学偏导数的关系，如 $C_p - C_V = T[(\partial S/\partial T)_p - (\partial S/\partial T)_V]$；以及热力学函数的全微分，如 $\mathrm{d}S$. 所谓"求"实际上是要把待求的宏观性质用可由实验测定的状态方程及热容量表示出来. 为此目的，除了要用到表Ⅰ中的普遍热力学关系外，常用的还有以下几个偏微分学公式：

循环关系：
$$\left(\frac{\partial x}{\partial y}\right)_z\left(\frac{\partial y}{\partial z}\right)_x\left(\frac{\partial z}{\partial x}\right)_y = -1 \tag{5.22}$$

链式关系：
$$\left(\frac{\partial x}{\partial y}\right)_w = \left(\frac{\partial x}{\partial z}\right)_w\left(\frac{\partial z}{\partial y}\right)_w \tag{5.23}$$

复合函数的偏微分公式：

$$\text{若 } M = M(x,z(x,y)), \text{则}\left(\frac{\partial M}{\partial x}\right)_y = \left(\frac{\partial M}{\partial x}\right)_z + \left(\frac{\partial M}{\partial z}\right)_x\left(\frac{\partial z}{\partial x}\right)_y. \tag{5.24}$$

在学习下面的例子和下节的内容时，希望特别注意分析总结：待求的热力学偏导数有什么特点，根据其特点用什么公式最简便. 凡遇到"热学量" T,S 对"力学量" p,V 的偏导数时，立即利用麦克斯韦关系.

【例 1】 求焦耳系数 $(\partial T/\partial V)_U$.

解 $(\partial T/\partial V)_U$ 的物理意义是，它表示气体自由膨胀时，温度对体积的依赖关系. 虽然自由膨胀是一个不可逆过程，但正如我们已强调的那样，只要初、末态为平衡态，普遍热力学关系仍然适用.

我们首先注意到 $(\partial T/\partial V)_U$ 的特点是，含有热力学势 U，且 U 在偏导数的脚标上. 为了能利用含 U 的热力学基本方程(5.1)式，先利用循环关系将 U 从偏导数的脚标位置转移到偏导数的分子上，即

$$\left(\frac{\partial T}{\partial V}\right)_U = -\left(\frac{\partial U}{\partial V}\right)_T\Big/\left(\frac{\partial U}{\partial T}\right)_V = -\frac{1}{C_V}\left(\frac{\partial U}{\partial V}\right)_T.$$

为求 $(\partial U/\partial V)_T$，考虑 $U = U(T,V)$，写出全微分

$$dU = (\partial U/\partial V)_T dV + (\partial U/\partial T)_V dT.$$

和

$$dU = TdS - pdV = [T(\partial S/\partial V)_T - p]dV + T(\partial S/\partial T)_V dT$$

比较得

$$(\partial U/\partial V)_T = T(\partial S/\partial V)_T - p, \qquad C_V = (\partial U/\partial T)_V = T(\partial S/\partial T)_V.$$

注意 $(\partial S/\partial V)_T$ 是"热学量" S 对"力学量" V 的偏导数，立刻想到利用麦克斯韦关系，将麦克斯韦关系 $(\partial S/\partial V)_T = (\partial p/\partial T)_V$ 代入，得

$$(\partial U/\partial V)_T = T(\partial p/\partial T)_V - p. \tag{5.25}$$

最后得焦耳系数为

$$\left(\frac{\partial T}{\partial V}\right)_U = = -\frac{1}{C_V}\left[T\left(\frac{\partial p}{\partial T}\right)_V - p\right]. \tag{5.26}$$

至此，我们已将要求的热力学偏导数用可由实验测得的 C_V 和可由状态方程求得的量表示出来. 将理想气体的状态方程 $pV = nRT$ 和范氏气体方程 $(p + a/V^2)(V - b) = nRT$ 分别代入，很容易证明有：

理想气体：　　$(\partial T/\partial V)_U = 0$,

范氏气体：　　$(\partial T/\partial V)_U = -a/(V^2 C_V)$.

结果表明，理想气体自由膨胀后温度不变，非理想气体自由膨胀后温度降低. 但这一效应非常微弱，不是致冷的有效方法.

【例 2】　求 $(\partial C_p/\partial p)_T$.

解　所求偏导数的意义是，定压热容量对压强的依赖关系. 根据定义：$C_p = \lim\limits_{\Delta T \to 0}(\Delta Q/\Delta T)_p = (dQ/dT)_p$. 在可逆过程中，有 $dQ = (TdS)_p$，于是 $C_p = T(\partial S/\partial T)_p$. 这类似于 $C_V = T(\partial S/\partial T)_V$.

将 $C_p = T(\partial S/\partial T)_p$ 对 p 求微商

$$\left(\frac{\partial C_p}{\partial p}\right)_T = T\frac{\partial}{\partial p}\left(\frac{\partial S}{\partial T}\right)_p = T\frac{\partial}{\partial T}\left(\frac{\partial S}{\partial p}\right).$$

将麦克斯韦关系 $(\partial S/\partial p)_T = -(\partial V/\partial T)_p$ 代入，得

$$(\partial C_p/\partial p)_T = -T(\partial^2 V/\partial T^2)_p. \tag{5.27}$$

这表明，只要已知系统的温度及状态方程，就可求得该温度下 C_p 随压强的变化. 将上式从初始压强 p_0 积分到任意压强 p，得到

$$C_p = C_{p_0} - T\int_{p_0}^{p}\left(\frac{\partial^2 V}{\partial T^2}\right)_p dp. \tag{5.28}$$

结果表明，只要知道状态方程及任一给定压强下的定压热容量 C_{p_0}，就能计算出给定温度的任意压强下的定压热容量. 也就是说，虽然 C_p 随压强变化，但测量时只要

测得某一压强 p_0 时的值就够了. 作最少的实验得到最多的结果,这正是理论的作用.

【例 3】 求 $C_p - C_V$.

解 在实验中升高温度时,维持体积不变比维持压强不变要困难得多,特别是对固体和液体. 若已知 $C_p - C_V$,则可通过直接测量 C_p 而求得不便测量,但在统计物理中用正则系综便于研究的 C_V. 据 $C_p = T(\partial S/\partial T)_p$, $C_V = T(\partial S/\partial T)_V$,得

$$C_p - C_V = T[(\partial S/\partial T)_p - (\partial S/\partial T)_V].$$

可见,利用复合函数偏导数公式(5.24)会很方便. 将 S 看成 T,V 的函数,而 V 又看成 T,p 的函数,即有 $S = S(T,V(T,p))$,利用(5.24)式得

$$(\partial S/\partial T)_p = (\partial S/\partial T)_V + (\partial S/\partial V)_T(\partial V/\partial T)_p.$$

代入上式得

$$C_p - C_V = T(\partial S/\partial V)_T(\partial V/\partial T)_p.$$

其中 $(\partial S/\partial V)_T$ 仍为不能直接测量的偏导数,但可利用麦克斯韦关系 $(\partial S/\partial V)_T = (\partial p/\partial T)_V$,最后将 $C_p - C_V$ 用可测得的量表示出来:

$$C_p - C_V = T(\partial p/\partial T)_V(\partial V/\partial T)_p = Tp\beta V\alpha = TV\alpha^2/\kappa_T. \tag{5.29}$$

(29)中的 β,α 和 κ_T 分别是等容压强系数,等压膨胀系数和等温压缩系数(见 §7.1.2),它们可由实验测得,或由状态方程计算得到.

将理想气体的状态方程 $pV = nRT$ 代入(5.29)式,可得 $C_p - C_V = nR$,这正是熟知的结果.

因为对已知物质均有 $\alpha > 0$ 和 $\kappa_T > 0$,故总有 $C_p > C_V$ 以及当 $T \to 0$ 时有 $C_p - C_V \to 0$.

(五)巨势

在(一)中所用的方法很容易推广到开放均匀系. 考虑开放均匀流体系的热力学基本方程[见 §7.2.3(3.27)式]

$$\mathrm{d}U = T\mathrm{d}S - p\mathrm{d}V + \sum_k \mu_k \mathrm{d}N_k. \tag{5.30}$$

利用勒让德变换,交换 p 和 V 的地位,利用 H 的定义;交换 T 和 S 的地位,利用 F 的定义;同时交换 T 和 S 及 p 和 V 的地位,利用 G 的定义;立刻得到(5.30)式的三种等价形式:

$$\mathrm{d}H = T\mathrm{d}S + V\mathrm{d}p + \sum_k \mu_k \mathrm{d}N_k, \tag{5.31}$$

$$\mathrm{d}F = -S\mathrm{d}T - p\mathrm{d}V + \sum_k \mu_k \mathrm{d}N_k, \tag{5.32}$$

$$\mathrm{d}G = -S\mathrm{d}T + V\mathrm{d}p + \sum_k \mu_k \mathrm{d}N_k. \tag{5.33}$$

从这些公式可得出

$$\begin{aligned}\mu_k &= (\partial U/\partial N_k)_{S,V,N_{k'}} = (\partial H/\partial N_k)_{S,p,N_{k'}} \\ &= (\partial F/\partial N_k)_{T,V,N_{k'}} = (\partial G/\partial N_k)_{T,p,N_{k'}} \qquad (k' \neq k).\end{aligned} \tag{5.34}$$

也就是说，把 U,H,F,G 这些量中的任何一个对 k 组元粒子数进行微分就可以得到 k 组元的化学势. 但是在每一种情形下，化学势是用不同的变量表示的. 例如，(5.33)式表示 $G = G(T,p,\{N_k\})$，因此，

$$\mu_k = (\partial G/\partial N_k)_{T,p,N_{k'}} = \mu_k(T,p,\{N_{k'}\}) \qquad (k' \neq k). \tag{5.35}$$

显然，还可以用变换参量组合的方法得到更多的热力学势及更多的和(5.30)式等价的方程. 但是，通常感兴趣的要少得多，除 U,H,F 和 G 外，仅还有巨势. 在(5.30)式中利用勒让德变换，同时交换 T 和 S，μ_k 和 N_k 的地位，可得

$$\mathrm{d}\Omega_{巨} = -S\mathrm{d}T - p\mathrm{d}V - \sum_k N_k \mathrm{d}\mu_k. \tag{5.36}$$

其中，巨势 $\Omega_{巨}$ 的定义是

$$\Omega_{巨} \equiv U - TS - \sum_k \mu_k N_k. \tag{5.37}$$

(5.36)式表明，$\Omega_{巨}$ 是以 T,V 和 $\{\mu_k\}$ 为自变量的特性函数. 若已知以 T,V 和 $\{\mu_k\}$ 为自变量的 $\Omega_{巨}$ 的函数形式 $\Omega(T,V,\{\mu_k\})$，可求得

$$\begin{aligned}S &= -(\partial\Omega_{巨}/\partial T)_{V,\{\mu_k\}}, \\ p &= -(\partial\Omega_{巨}/\partial V)_{T,\{\mu_k\}}, \\ N_k &= -(\partial\Omega_{巨}/\partial\mu_k)_{V,T,\mu_{k'}} \qquad (k' \neq k).\end{aligned} \tag{5.38}$$

进而求得全部热力学量. 在下一章中，我们将看到巨势 $\Omega_{巨}$ 在巨正则系综中的重要作用.

已经指出，和内能 U、熵 S 一样，H、F、G 和 $\Omega_{巨}$ 都是广延量或可加量. 可加性意味着粒子数增加多少倍时，这个量也增加多少倍. 换句话说，可加性热力学量是各组元粒子数的一次齐次函数. 考查自由焓 G，则应有

$$G(T,p,\lambda\{N_k\}) = \lambda G(T,p,\{N_k\}).$$

根据齐次函数的欧拉定理[见§1.2.1(1.23)式]有

$$G(T,p,\{N_k\}) = \sum_k (\partial G/\partial N_k)_{T,p,N_{k'\neq k}} N_k = \sum_k \mu_k N_k. \tag{5.39}$$

将(5.39)式代入(5.37)式又可得

$$\Omega_{巨} \equiv U - TS - G = F - G = -pV. \tag{5.40}$$

对单元系，(5.39)式化为

$$\begin{aligned}G(T,p,N) &= \mu N = g(T,p)N, \\ \mu &= (\partial G/\partial N)_{T,p} = g(T,p).\end{aligned} \tag{5.41}$$

即单元系的化学势就是摩尔自由焓. 它仅仅是温度和压强的函数. 而多组元系统

中，某组元的化学势则不仅仅依赖于温度和压强，还依赖其他组元的物质的量，这可从(5.35)式看出. 这些结论在后面的讨论中都要用到.

§7.2.6　气体的节流和绝热膨胀　磁冷却

不论在生产或科研中，低温都愈来愈重要. 气体节流膨胀和绝热膨胀、磁冷却都是获得低温的重要方法. 作为普遍热力学关系的应用，本节对这些过程进行分析.

(一)气体的节流膨胀

1. 实验的描述和结果

图1是节流膨胀过程的示意图. 用绝热材料包扎好的管子，中间有一多孔塞，称为节流阀. 它的作用是阻滞气体通过. 多孔塞的一边维持较高的压强 p_i 和某一温度 T_i，p_i 和 T_i 是可以控制的；另一边维持较低的压强 p_f(通常就是大气压)，相应的温度是 T_f. 由于多孔塞的巨大阻力，气体的流速很小，从而宏观动能可以忽略不计. 气体在绝热条件下，通过这样的多孔塞或节流阀由稳定高压区流到稳定低压区的过程称为节流过程.

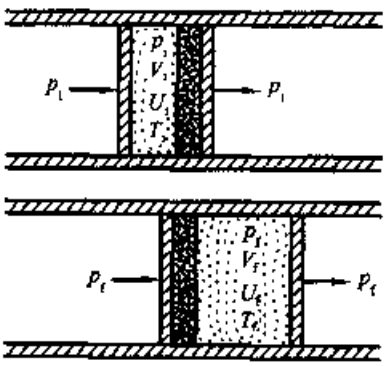


图1　气体节流膨胀示意图

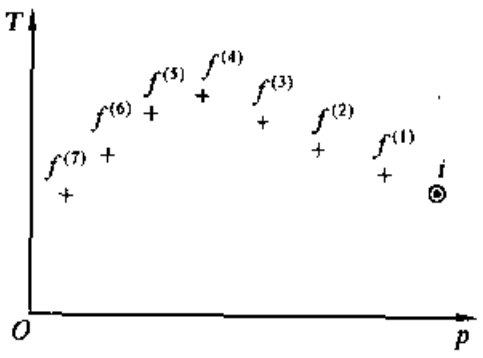


图2　等焓线示意图

由于气体通过多孔塞时要克服阻力做功，此功会变成热，还可能出现涡漩，所以和自由膨胀一样，节流膨胀过程也是不可逆过程. 但因初态和末态是平衡态，故仍可用普遍热力学关系进行研究.

下面证明，节流前后气体的焓相等. 取一定质量的气体，它的压强、体积和内能在节流前后分别为 p_i, V_i, U_i 和 p_f, V_f, U_f. 因为过程是绝热的，根据热力学第一定律有 $\Delta U = \Delta W$，$U_f - U_i = p_iV_i - p_fV_f$，或 $U_f + p_fV_f = U_i + p_iV_i$，所以

$$H_f = H_i, \tag{6.1}$$

即节流前后，气体的焓相等.

实验时，选定一组 (p_i, T_i)，这时气体的焓为 $H_i(p_i, T_i)$，取一个 $p_f^{(1)}$，测得相应的 $T_f^{(1)}$，有 $H_f^{(1)}(p_f^{(1)}, T_f^{(1)})$，取另一 $p_f^{(2)}$，测得相应的 $T_f^{(2)}$，显然有 $H_i(p_i, T_i) =$

$H_f^{(1)}(p_f^{(1)},T_f^{(1)}) = H_f^{(2)}(p_f^{(2)},T_f^{(2)}) = \cdots$，这样可得到一系列等焓的状态代表点 (p_i,T_i)，$(p_f^{(1)},T_f^{(1)})$，$(p_f^{(2)},T_f^{(2)})$，… 在 $T-p$ 图上标出以上各点，并将它们联接起来，就得到一条等焓线，如图 2. 选定另一组 (p_i,T_i)，重复以上步骤，将得到另一条等焓线.

实验表明，节流前后，气体的温度可能升高，可能降低，也可能不变，由气体的性质和节流前的状态决定. 节流过程中气体温度随压强降低而变化的现象称为焦耳-汤姆孙效应. 这一效应可用偏导数 $\mu = (\partial T/\partial p)_H$ 表示，称为焦耳-汤姆孙系数，脚标 H 表示焓不变. 因为节流过程压强总是降低，所以如果 $\mu > 0$，则节流后气体的温度降低，叫致冷效应或正效应. 反之如果 $\mu < 0$，则节流后气体的温度升高，叫致温效应或负效应. 若 $\mu = 0$，气体经节流后温度不变，叫零效应，它对应于等焓线（参看图 3）的极大值. 此时的温度称为相应压强下的转换温度.

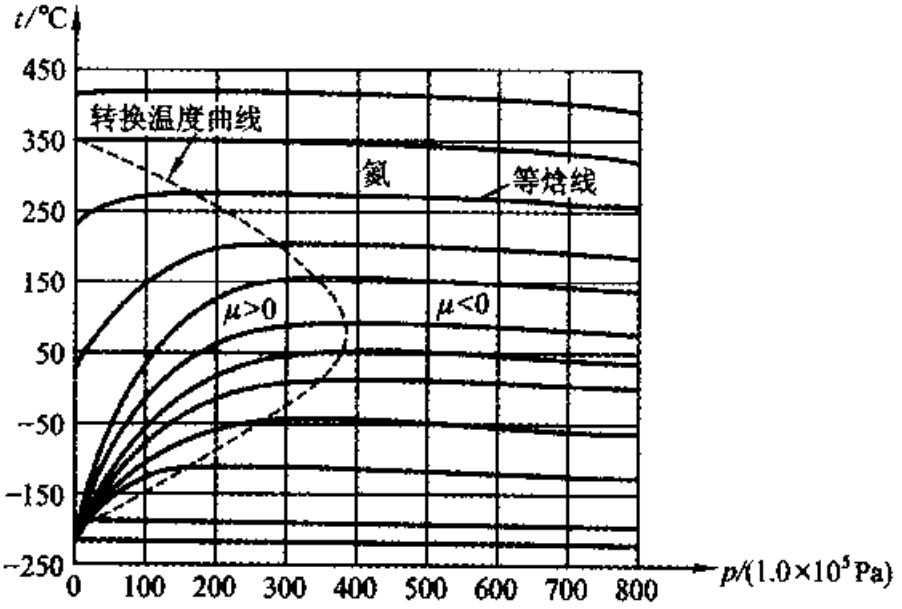


图 3　氮的等焓线和转换温度

图 3 上画出了由实验得到的氮的一组等焓线和转换温度曲线. 由图可见，在虚线以内的区域，各等焓线的斜率 $\mu = (\partial T/\partial p)_H$ 皆大于零. 若气体的初态位于该区，节流膨胀时将得到致冷效应. 反之，在虚线以外的区域，各等焓线的斜率 $\mu < 0$，若气体的初态位于该区，节流膨胀时将得到致温效应. 虚线则对应于 $\mu = 0$ 的状态，叫做转换温度曲线. 转换温度曲线与纵轴的上交点叫做最大转换温度 T_{max}. 在 T_{max} 以上的等焓线，其焦耳-汤姆孙系数 μ 均小于零，属于致温区. 所以，要想通过节流得到致冷效应，节流前的温度必须低于最大转换温度. 如表 I 所示，大多数气体的 T_{max} 都在室温以上，因此，在室温进行节流膨胀就可获得致冷效应. 但氢 H_2 和氦 He 的 T_{max} 低于室温，如果要利用节流膨胀使它们液化，必须先把它们冷却到转换温度以下. 通常的做法是利用其他液化气体（如液态氮，它在一个大气压下的沸点

是77.3K)作预冷却.

表Ⅰ　气体的最大转换温度(K)

气　　体	CO_2	Ar	O_2	空气	N_2	Ne	H_2	He
最大转换温度	≈1500	780	764	659	621	231	202	≈40

2. 焦尔-汤姆孙系数 μ 的计算

焦耳-汤姆孙系数和焦耳系数是同一类型的热力学偏导数,可以用同样的方法计算(见上节例1).先用循环关系将 H 转移到偏导数的分子上,得

$$\mu = (\partial T/\partial p)_H = -(\partial H/\partial p)_T/(\partial H/\partial T)_p = -(\partial H/\partial p)_T/C_p.$$

写出 $dH = (\partial H/\partial T)_p dT + (\partial H/\partial p)_T dp$,和热力学基本方程

$$dH = TdS + Vdp = T(\partial S/\partial T)_p dT + [T(\partial S/\partial p)_T + V]dp$$

比较得 $(\partial H/\partial p)_T = T(\partial S/\partial p)_T + V$.将麦克斯韦关系代入,最后得到

$$\mu = \frac{1}{C_p}\left[T\left(\frac{\partial V}{\partial T}\right)_p - V\right]. \tag{6.2}$$

因为 C_p 总是正的,所以 μ 的符号由上式中方括号的正负决定,即由状态方程和节流前的状态决定.

理论推导与实验结果的一致性还可以用以下方法验证.将气体的范氏方程代入(6.2)式,并令 $\mu = 0$,再和范氏方程联立消去 V,可得转换温度曲线方程:

$$p = (a/b^2)(1 - \sqrt{RTb/(2a)})(3\sqrt{RTb/(2a)} - 1).$$

将上式绘出的图线,及从图上得到的最大转换温度和实验结果比较,定性上是一致的(见习题32).定量上的差别则是由于范氏方程只是实际气体的近似描述.

焦尔-汤姆孙效应非常弱,一般为每个大气压十分之几度的数量级.但由于它是个稳定过程,因而可利用热交换器把冷却效果累积起来,达到致冷的目的.

(二)气体的绝热膨胀

另一种使气体冷却的方法是让气体绝热膨胀.这时的降温效果可用偏导数 $(\partial T/\partial p)_S$ 表示.上节已经看到,利用 C_p, C_V 的定义式和麦克斯韦关系可以把含 S 的一阶偏导数转化成可测的量.因此,为求 $(\partial T/\partial p)_S$,首先要把 S 转移到偏导数的分子上.为此可利用循环关系:

$$(\partial T/\partial p)_S = -(\partial S/\partial p)_T/(\partial S/\partial T)_p = -(T/C_p)(\partial S/\partial p)_T.$$

再利用麦克斯韦关系可得

$$(\partial T/\partial p)_S = (T/C_p)(\partial V/\partial T)_p = \alpha TV/C_p. \tag{6.3}$$

因为 α, T, V 和 C_p 均为正值,故总有 $(\partial T/\partial p)_S > 0$,即在绝热膨胀过程中随着压强的降低,气体温度总是降低,不需要预冷.这是这一方法的优点.但是温度愈低,致

冷效果愈差.通常的做法是将绝热膨胀和节流膨胀结合.这是目前获得4.2K(液氦)低温的有效方法.

(三)磁冷却

如何能获得尽可能低的温度,是科学技术中的一个重要课题.在一个大气压下 ^4He 的沸点为4.2K.用抽气机将液态 ^4He 上的 ^4He 气抽走,使液态 ^4He 在低压下沸腾,可得0.8K的低温.让 ^3He 在低压下沸腾可达0.3K.但是想用这种方法来维持0.1K的低温,所需蒸气压将低到不可想象.这正是卡末林·昂内斯于1908年成功地液化氦气后,企图沿用同样方法获得更低温度时陷入困境的主要原因.1926年德拜和乔克分别在理论上提出了一种完全不同的新方法,并由乔克于1933年成功地进行了第一次实验.该方法的基本原理是,在绝热条件下可逆退磁,使顺磁体的温度降低,所以叫磁冷却.现在利用这种方法已获得0.001K,并能维持数小时之久.因此,磁冷却法是目前获得低温的有效方法之一.

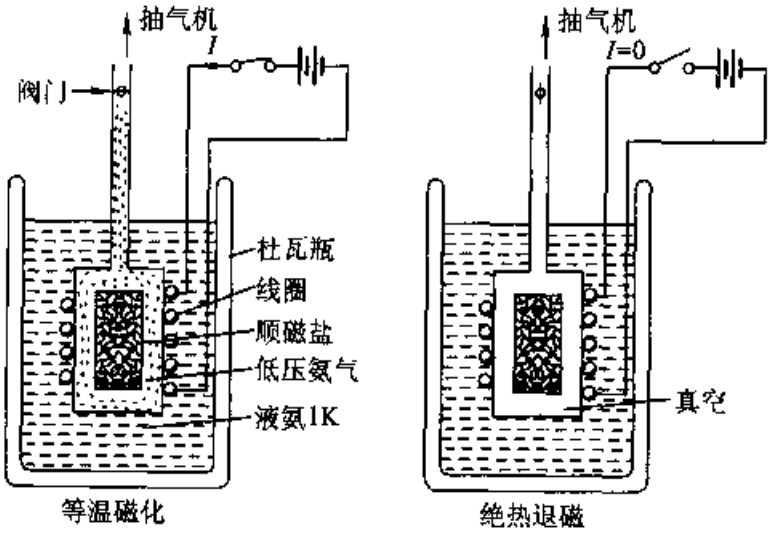


图4　磁冷却法示意图

1.磁冷却实验的描述

磁冷却利用的样品是一些磁性很弱的顺磁盐,如硫酸钆[$Gd_2(SO4)_3 \cdot 8H_2O$],硫酸铁铵[$FeNH_4(SO_4)_2 \cdot 12H_2O$]等,在这类顺磁盐中,由于磁性离子磁矩(如硫酸铁铵中的 Fe^{3+})被许多其他原子隔开,其相互作用非常微弱,以致当温度低到1K左右时,磁性离子磁矩的取向仍然是无规则的,居里定律 $\boldsymbol{M} = c\boldsymbol{H}/T$ 仍然有效.将这种样品进行等温磁化和绝热退磁就可以获得低温.

把顺磁盐样品放在充有低压氦气的小室中如图4,通过氦气与迅速蒸发的液氦热接触,保持1K以下的预冷.加上磁场使顺磁盐磁化,磁化中放出的热量可通过氦气传给液氦,从而保证磁化过程等温.磁化后,抽去氦气使样品绝热,然后使磁场减小到零.在此绝热退磁过程中,顺磁样品温度降低.因为磁化弛豫时间很短,所以虽然退磁在几秒钟内完成,仍可以认为足够缓慢,是可逆过程.

2.热力学分析

现在我们用普遍热力学关系来计算等温磁化过程中样品"吸收"的热量 dQ_T(如果计算结果 $dQ_T < 0$ 则说明实际是放热),和绝热退磁时样品温度的变化

$(\partial T/\partial H)_S$($H$ 是磁场强度).

实验中样品体积的变化可忽略不计.外界对介质所作的唯一功是磁化功 $\boldsymbol{H}\cdot \mathrm{d}\boldsymbol{M}$.只要作以下替换 $p\to -\boldsymbol{H}, V\to \boldsymbol{M}$ 前面所有公式仝部有效❶.如:

自由焓:　　$G = U - TS + pV \to G = U - TS - HM,$

热力学基本方程:　$\mathrm{d}G = -S\mathrm{d}T + V\mathrm{d}p \to \mathrm{d}G = -S\mathrm{d}T - M\mathrm{d}H,$

麦克斯韦关系:　　$-(\partial S/\partial p)_T = (\partial V/\partial T)_p \to (\partial S/\partial H)_T = (\partial M/\partial T)_H,$

状态方程:　　$\mathscr{F}(p,V,T) = 0 \to \mathscr{F}(H,M,T) = 0,$

热容量:　　$C_p = T(\partial S/\partial T)_p \to C_H = T(\partial S/\partial T)_H.$

设选 T 和 H 为态参量,由 $S = S(T,H)$ 可写出

$$\mathrm{d}S = (\partial S/\partial T)_H\mathrm{d}T + (\partial S/\partial H)_T\mathrm{d}H = (C_H/T)\mathrm{d}T + (\partial M/\partial T)_H\mathrm{d}H,$$

或

$$T\mathrm{d}S = C_H\mathrm{d}T + T(\partial M/\partial T)_H\mathrm{d}H, \tag{6.4}$$

于是等温磁化过程中样品吸收的热量为

$$(\mathrm{d}Q)_T = T\mathrm{d}S = T(\partial M/\partial T)_H\mathrm{d}H. \tag{6.5}$$

由居里定律可得 $(\partial M/\partial T)_H = -cH/T^2$,代入(6.5)式得

$$(\mathrm{d}Q)_T = -(c/T)H\mathrm{d}H = -M\mathrm{d}H < 0, \tag{6.6}$$

说明等温磁化($\mathrm{d}H > 0$)时,外场对介质所作的功转变成热被释放出来.

绝热退磁时有 $\mathrm{d}S = 0$,由(6.4)式可得

$$(\partial T/\partial H)_S = -(T/C_H)(\partial M/\partial T)_H = (cH/C_HT) > 0. \tag{6.7}$$

考虑到 $C_H \propto T^{-2}$,则有 $(\partial T/\partial H)_S \propto T$.即绝热退磁时,介质的温度降低.然而温度愈低,降温效果愈差.

§7.2.7　最大功原理　平衡的稳定性条件

在热力学中,功的计算是一个重要问题.根据热力学第一定律,在绝热过程中,功可以用态函数内能的变化来计算,而不必考虑过程的细节.在其他过程,特别是实验中经常遇到的等温、等压过程,功的计算是否也可以通过某个态函数在初、末态的差来求得,从而摆脱过程的细节呢?另一方面,热力学中的一个重要问题是:判断系统中不可逆过程进行的方向,和系统是否已达到平衡.如果使用熵增加原理来判断过程进行的方向和限度,必须把我们感兴趣的系统和不感兴趣的环境构成

❶ 磁场 $\boldsymbol{H}$ 和总磁矩 $\boldsymbol{M}$ 都是矢量.但是,由于顺磁盐各向同性,所以实际上可以只考虑 $\boldsymbol{H}$ 和 $\boldsymbol{M}$ 的大小,而不管它们的方向.在以下公式中可以将 $\boldsymbol{H}$ 和 $\boldsymbol{M}$ 换成它们的大小 H 和 M.例如麦克斯韦关系 $(\partial S/\partial H)_T = (\partial M/\partial T)_H$.

一个大的孤立系，计算孤立系的总熵差. 显然，这并非总是方便的. 对于非孤立系统中发生的不可逆过程，是否可以用别的态函数来判断过程进行的方向和限度呢？下面证明，用 §7.2.5 中引入的热力学势，就可以通过它们的变化来计算特定过程中的功，并能方便地判断特定过程进行的方向和限度. 证明时，所用方法的根据是：任一给定系统的态参量数是确定的，态参量的选取却是任意的，并可通过勒让德变换来完成.

(一)最大功原理

已知对于封闭系统，热力学第一定律为 $\mathrm{d}U = \mathrm{d}Q - \mathrm{d}W_{系}$. 考虑热力学第二定律：$T\mathrm{d}S \geqslant \mathrm{d}Q$，将上式中的 $\mathrm{d}Q$ 用 $T\mathrm{d}S$ 代替，则得

$$\mathrm{d}U \leqslant T\mathrm{d}S - \mathrm{d}W_{系}. \tag{7.1}$$

其中等号对应于可逆过程，不等号对应于不可逆过程.

利用勒让德变换交换(7.1)中 T 和 S 的地位，并利用自由能的定义，得

$$\mathrm{d}F \leqslant -S\mathrm{d}T - \mathrm{d}W_{系}. \tag{7.2}$$

若系统经历一等温可逆过程 $\mathrm{d}T = 0$，由(7.2)式得

$$-\mathrm{d}F = \mathrm{d}W_{系}. \qquad (\text{可逆等温过程}) \tag{7.3}$$

这表明，在等温可逆过程中，系统对外所作的功等于自由能的减少. 正是由于这一原因，人们称 F 是自由能，意思是在可逆等温过程中，只有内能的这一部分 $(U - TS)$ 可以转变为功. 它是自由的，而另一部分 (TS) 似乎被束缚住了，不能变为功.

若系统经历的是不可逆等温过程，则由(7.2)式得

$$-\mathrm{d}F > \mathrm{d}W_{系}. \qquad (\text{不可逆等温过程}) \tag{7.4}$$

这表示，在不可逆的等温过程中，系统对外所作的功小于自由能的减少. 换句话说，在可逆等温过程中，系统所作的功最大. 这叫做最大功原理.

对于 $\mathrm{d}T = 0, \mathrm{d}W_{系} = 0$ 的系统，过程进行的方向和限度由

$$\mathrm{d}F \leqslant 0 \qquad (\text{等温不作功}) \tag{7.5}$$

决定. 当 F 达到极小值时，系统达到平衡态. 即：等温且不作功的系统中进行的过程，系统的自由能绝不会增加，平衡态对应于自由能 F 取最小值的宏观态.

若系统对外的功可分为膨胀功和非膨胀功两部分，即有

$$\mathrm{d}W_{系} = p\mathrm{d}V + \mathrm{d}W'_{系}, \tag{7.6}$$

则(7.1)式化为

$$\mathrm{d}U \leqslant T\mathrm{d}S - p\mathrm{d}V - \mathrm{d}W'_{系}. \tag{7.7}$$

利用勒让德变换同时交换(7.7)式中的 T 和 S，p 和 V 的地位，并考虑自由焓的定义得

$$\mathrm{d}G \leqslant -S\mathrm{d}T + V\mathrm{d}p - \mathrm{d}W'_{系}. \tag{7.8}$$

若系统经历一个可逆的等温、等压过程，即有 $\mathrm{d}T = 0$ 和 $\mathrm{d}p = 0$，由(7.8)式得

$$-\mathrm{d}G = \bar{d}\mathrm{W}'_{系} \quad (可逆等温等压过程). \tag{7.9}$$

这表示，在可逆的等温、等压过程中，系统对外所作的非膨胀功（如电池对外所作电功）等于自由焓的减少.

若系统经历的是不可逆的等温、等压过程，则由(7.8)式得

$$-\mathrm{d}G > \bar{d}\mathrm{W}'_{系} \quad (不可逆等温等压过程). \tag{7.10}$$

这表示，在不可逆的等温等压过程中，系统对外作的非膨胀功小于自由焓的减少. 换句话说，系统在可逆的等温等压过程中所作非膨胀功最大，这也叫做最大功原理.

对于 $\mathrm{d}T = 0, \mathrm{d}p = 0$ 和 $\bar{d}\mathrm{W}'_{系} = 0$ 的系统，过程进行的方向和限度由

$$\mathrm{d}G \leqslant 0 \quad (等温等压无非膨胀功) \tag{7.11}$$

决定. 当 G 达到最小值时，系统达到平衡态. 即：等温等压且无非膨胀功的系统中发生的过程，其自由焓 G 绝不会增加，平衡态对应于自由焓具有最小值的宏观态.

类似的讨论可以证明，对于 $\mathrm{d}S = 0, \mathrm{d}p = 0$ 和 $\bar{d}\mathrm{W}'_{系} = 0$ 的系统，过程进行的方向和限度由

$$\mathrm{d}H \leqslant 0 \quad (等熵等压无非膨胀功) \tag{7.12}$$

决定. 即：这样的系统中发生的过程，焓绝不会增加，平衡态对应于焓具有最小值的宏观态.

从以上的讨论得到，在三种不同条件下，系统分别向 F, G 和 H 减小的方向变化. 平衡态分别对应于 F, G 和 H 的最小值. 这和在重力场中，物体向重力势能减小的方向运动，直到达到势能的最小值相似. 正是由于这种相似性，我们把 F, G 和 H 统称为热力学势.

最后要注意的是，在绝热过程中没有最大功原理，因为绝热过程中，无论过程可逆或不可逆，系统所作的功总是等于内能的减少.

【例 1】 如 §7.2.1 图 3 所示，在金属框上张有一液体膜，长为 l 的金属丝 AA' 可以无摩擦地滑动. 液体膜的表面张力系数为 σ. 当外力 f 等于表面张力 $2\sigma l$ 时，金属丝缓慢地移动，液体膜经历一可逆的等温过程. 根据(7.3)式，外界为克服表面张力所作的功等于系统自由能的增量

$$\mathrm{d}F = 2\sigma l\mathrm{d}x = 2\sigma\mathrm{d}A.$$

A 是表面膜的面积. 考虑到表面张力系数只是温度的函数 $\sigma = \sigma(t)$ 和面积无关，积分上式可得 $F = 2\sigma A + F_0$. 当 $A = 0$ 时 $F = 0$，故

$$\sigma = F/2A.$$

即表面张力系数等于单位面积的自由能. 据此可解释荷叶上水珠呈球形的现象. 若忽略重力的影响，水珠形状的变化是一个等温且和外界无功交换的过程. 按(7.5)式，水珠的自由能，因而表面积最小时达到平衡态. 在体积相同的条件下，球体表面

积最小，因此水珠的稳定状态呈球形.

铁磁体在外磁场中被磁化时，其长度和体积都要发生微小的变化. 这种现象叫做磁致伸缩效应. 可以证明磁致伸缩效应也是为满足总自由能最小而发生的.

【例 2】 简单的生化过程

许多化学反应，包括生物有机体中进行的生化反应，都是在恒温恒压下进行的. 因此，为研究反应进行的方向，用自由焓最方便.

由葡萄糖（六个碳原子组成的环状结构 $C_6H_{12}O_6$）和果糖（六个碳原子组成的另一种环状结构）合成比较复杂的蔗糖是一种最简单的生化过程

$$\text{葡萄糖} + \text{果糖} \rightleftharpoons \text{蔗糖} + H_2O.$$

测量表明，在标准条件下[每种反应物的浓度为 $1\,\text{mol}\cdot\text{l}^{-1}$（每升一摩尔）]沿左向右的反应伴随着自由焓的变化是 $\Delta G = 0.24$ eV. 根据(7.11)式，为了使反应能沿左向右进行，必须与另一个自由焓 $\Delta G'$ 为负且足够大的反应联系起来，使得两个联合反应的总自由焓满足(7.11)式：$\Delta G + \Delta G' \leqslant 0$.

研究表明，生物有机体中广泛存在的能获得负值 $\Delta G'$ 的反应是

$$\text{ATP（腺苷三磷酸）} + H_2O \rightarrow \text{ADP（（腺苷二磷酸）} + \text{磷酸离子}.$$

在标准条件下伴随的 $\Delta G' = -0.30$ eV.

通过中间体（1-磷酸葡萄糖）[这种分子由在葡萄糖环状结构的“1”位上附着一个磷酸基团形成]可以把以上两个反应偶联起来：

$$\text{ATP} + \text{葡萄糖} \rightarrow \text{ADP} + (\text{1-磷酸葡萄糖}),$$

$$(\text{1-磷酸葡萄糖}) + \text{果糖} \rightarrow \text{蔗糖} + \text{磷酸离子},$$

它们给出的净反应是

$$\text{ATP} + \text{葡萄糖} + \text{果糖} \rightarrow \text{蔗糖} + \text{ADP} + \text{磷酸离子}.$$

合成蔗糖时的 ΔG 被 ATP 破裂为较简单的 ADP 时的 $-\Delta G'$ 所补偿.

（二）平衡的稳定性判据和条件

根据热力学第二定律，孤立系统中的真实过程总是向着熵增加的方向进行，即有

$$dS \geqslant 0. \tag{7.13}$$

达到平衡态后，熵有最大值，且不再变化. 设想这时有一个偏离平衡态的虚变动，则必有

$$\Delta S = \delta^{(1)}S + \delta^{(2)}S/2 < 0. \tag{7.14}$$

其中 $\delta^{(1)}S$ 和 $\delta^{(2)}S$ 分别是 S 的一阶和二阶变分. 平衡态的条件是

$$\delta^{(1)}S = 0, \qquad \delta^{(2)}S < 0. \tag{7.15}$$

其中 $\delta^{(1)}S = 0$ 代表极值条件，$\delta^{(2)}S < 0$ 代表极大值条件，在物理上往往称前者为平衡判据，称后者为平衡的稳定性判据.

对于(7.1)、(7.2)、(7.8)式可进行类似讨论. 考虑(7.1)式,对于等熵、无外功的系统,内能取最小值是稳定平衡态,其条件为在偏离平衡态的虚变动中有

$$\Delta U = \delta^{(1)}U + \delta^{(2)}U/2 > 0. \tag{7.16}$$

$$\delta^{(1)}U = 0, \qquad \delta^{(2)}U > 0. \tag{7.17}$$

$\delta^{(1)}U = 0$ 代表极值条件或平衡判据,$\delta^{(2)}U > 0$ 是极小值条件或平衡的稳定性判据.

考虑(7.2)式,对于等温、无外功的系统,自由能取最小值是稳定平衡态,其条件为,在偏离平衡态的虚变动中有

$$\Delta F = \delta^{(1)}F + \delta^{(2)}F/2 > 0. \tag{7.18}$$

$$\delta^{(1)}F = 0, \qquad \delta^{(2)}F > 0. \tag{7.19}$$

$\delta^{(1)}F = 0$ 是平衡判据,$\delta^{(2)}F > 0$ 是平衡的稳定性判据.

考虑(7.8)式,对于等温、等压、无非膨胀功的系统,自由焓取最小值是稳定平衡态,其条件为,在偏离平衡态的虚变动中有

$$\Delta G = \delta^{(1)}G + \delta^{(2)}G/2 > 0. \tag{7.20}$$

$$\delta^{(1)}G = 0, \qquad \delta^{(2)}G > 0. \tag{7.21}$$

$\delta^{(1)}G = 0$ 是平衡判据,$\delta^{(2)}G > 0$ 是平衡的稳定性判据.

显然,(7.15)、(7.17)、(7.19)和(7.21)式彼此等价,各自适用于不同情况,其中熵增加原理(7.15)式是最基本的.

S、U、F 和 G 都不能直接测量,因此有必要把抽象的判据用可直接测量的量表示出来,叫做平衡条件和平衡稳定性条件. 在§7.2.2中,我们运用 $\delta^{(1)}S = 0$,得到了孤立系统各部分达到平衡的具体条件是:具有相同的温度、化学势和压强. 为了研究平衡稳定性的具体条件,我们应该继续运用 $\delta^{(2)}S < 0$. 但是研究发现,用和(7.15)式等价的(7.17)式会更方便.

把 U 看成是 S 和 V 的函数 $U = U(S,V)$,可以把平衡的稳定判据写为

$$\delta^{(2)}U = (\partial^2U/\partial S^2)_V(\delta S)^2 + 2(\partial^2U/\partial S\partial V)\delta S\delta V + (\partial^2U/\partial V^2)_S(\delta V)^2 > 0. \tag{7.22}$$

根据线性代数,这一个正定二次型的系数应该满足下列条件:

$$(\partial^2U/\partial S^2)_V > 0, \tag{7.23}$$

$$\begin{vmatrix} (\partial^2U/\partial S^2)_V & (\partial^2U/\partial V\partial S) \\ (\partial^2U/\partial S\partial V) & (\partial^2U/\partial V^2)_S \end{vmatrix} = \left(\frac{\partial^2U}{\partial S^2}\right)_V\left(\frac{\partial^2U}{\partial V^2}\right)_S - \left(\frac{\partial^2U}{\partial V\partial S}\right)^2 > 0. \tag{7.24}$$

考虑到 $T = (\partial U/\partial S)_V$,(7.23)式化为

$$(\partial^2U/\partial S^2)_V = (\partial T/\partial S)_V = T/C_V > 0.$$

对于 $T > 0$ 的正常系统,这意味着

$$C_V > 0. \tag{7.25}$$

(7.24)式可以写成雅可比的形式[关于雅可比行列式的定义和性质见附录V的第六小节]:

$$D[(\partial U/\partial S)_V,(\partial U/\partial V)_S]/D[S,V] > 0,$$

或

$$D(T,-p)/D(S,V) > 0,$$

变换到以 T,V 为自变量时,有

$$\frac{D(T,-p)}{D(S,V)} = \frac{D(T,-p)}{D(T,V)}\Big/\frac{D(S,V)}{D(T,V)} = -\frac{(\partial p/\partial V)_T}{(\partial S/\partial T)_V} = -\frac{T}{C_V}\left(\frac{\partial p}{\partial V}\right)_T > 0,$$

因为 $T > 0$ 和 $C_V > 0$,上式意味着 $(\partial p/\partial V)_T < 0$,或

$$\kappa_T = -(1/V)(\partial V/\partial p)_T > 0. \tag{7.26}$$

(7.25)和(7.26)式叫做平衡的稳定性条件.(7.25)式叫做热平衡稳定性条件,(7.26)式叫做力学平衡稳定性条件.不满足这两个条件的平衡态是不稳定的,在自然界不可能存在.

稳定性条件的意义很容易理解.设想由于不可避免的涨落,使本来已达到热平衡的系统的温度高于外界的温度,于是将有热量从系统流出,$\mathrm{d}Q_{系} < 0$.若系统的 $C_V = (\mathrm{d}Q_{系}/\mathrm{d}T)_V > 0$,则有 $\mathrm{d}T < 0$.说明系统会由于放热而降温,以致重新回到热平衡.因此 $C_V > 0$ 保证了热平衡的稳定性.反之,若 $C_V < 0$,则将会有 $\mathrm{d}T > 0$,说明系统会由于放热而升温,以致越来越偏离热平衡.换句话说,$C_V < 0$ 会破坏热平衡的稳定性.为理解力学稳定性条件,可作类似的讨论.

我们看到,系统的平衡条件由系统的态参量 (T,p,μ) 决定,而平衡的稳定性条件则由系统的性质决定,这些性质涉及到的是强度量 (T,p) 对其共轭广延量 (S,V) 导数的符号.由此不难预料,扩散平衡的稳定性条件是:$(\partial\mu/\partial N)_{T,p} > 0$.在下一节将给出关于这一条件的证明.

§7.2.8　单元双相系的平衡　相图

宏观系统可以以许多相的形式存在,每种相表现出显著不同的性质.在一定条件下,系统会从一相变到另一相.如气体的液化,晶体的生长,铁磁和顺磁转变,正常流体和超流体的转变,正常导体和超导体的转变,以及在一定温度范围内,一些有机晶体转变为各向异性的液体(液晶)等.相变现象的研究有着愈来愈重要的应用.当然,相变机制的理解必须依赖统计物理,但是用热力学理论也可以方便地得到许多有关相变的普适的重要关系和结论.本节介绍相变热力学的基础知识.首先用已经得到的平衡和平衡稳定性判据研究单元双相系的平衡,并介绍研究相变的

有力工具——相图.

(一)单元双相系的平衡条件和稳定性条件

考虑一单元双相系,用 α 和 β 分别表示这两相,如水和水蒸气.假设系统与一恒温恒压的外源(如大气)接触,并已达到平衡(如图1).也就是说,两相已达热平衡和力学平衡,但粒子仍可从一相转移到另一相,即可发生相变.这正是通常实验的情况.让我们来研究两相平衡共存的平衡条件.

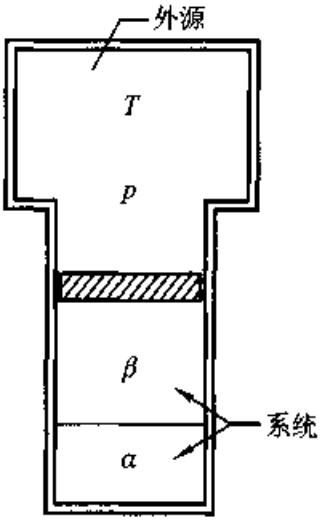


图1 已达到热力学平衡的单元双相系

从上节知道,对于温度和压强保持不变,且无非膨胀功的系统,平衡时自由焓有最小值.因此,对于已达热平衡和力学平衡,但可发生相变的单元双相系统,可利用自由焓判据 $\delta^{(1)}G=0$ 和 $\delta^{(2)}G>0$ 来求两相平衡共存的条件和平衡的稳定性条件.

因为自由焓是广延量,单元双相系的自由焓等于两相自由焓之和

$$G=G^{(\alpha)}(T,p,N^{(\alpha)})+G^{(\beta)}(T,p,N^{(\beta)}),\quad (8.1)$$

$N^{(\alpha)}$ 和 $N^{(\beta)}$ 分别代表 α 相和 β 相中的分子数.在温度和压强不变,而粒子可从一相转移到另一相的情况下,有

$$\begin{aligned}\delta^{(1)}G&=(\partial G^{(\alpha)}/\partial N^{(\alpha)})_{T,p}\delta N^{(\alpha)}+(\partial G^{(\beta)}/\partial N^{(\beta)})_{T,p}\delta N^{(\beta)}\\&=\mu^{(\alpha)}(T,p)\delta N^{(\alpha)}+\mu^{(\beta)}(T,p)\delta N^{(\beta)}.\end{aligned}\quad (8.2)$$

因为是单元系,化学势就是摩尔自由焓,因而 $\mu^{(\alpha)}$ 和 $\mu^{(\beta)}$ 都只是 T,p 的函数[见§7.2.5(5.41)式].根据质量守恒,$\delta N^{(\alpha)}=-\delta N^{(\beta)}$,上式化为

$$\delta^{(1)}G=(\mu^{(\alpha)}-\mu^{(\beta)})\delta N^{(\alpha)}.\quad (8.3)$$

将(8.3)式代入平衡判据 $\delta^{(1)}G=0$,并注意到 $N^{(\alpha)}$ 是独立变量,$\delta N^{(\alpha)}$ 可取任意值,可知 $\delta^{(1)}G=0$ 成立的条件,即相平衡条件是

$$\mu^{(\alpha)}(T,p)=\mu^{(\beta)}(T,p).\quad (8.4)$$

再研究相平衡的稳定性条件.在温度和压强不变的情况下有

$$\begin{aligned}\delta^{(2)}G&=\{(\partial^2G^{(\alpha)}/\partial N^{(\alpha)2})_{T,p}(\delta N^{(\alpha)})^2+(\partial^2G^{(\beta)}/\partial N^{(\beta)2})_{T,p}(\delta N^{(\beta)})^2\}\\&=\{(\partial\mu^{(\alpha)}/\partial N^{(\alpha)})_{T,p}+(\partial\mu^{(\beta)}/\partial N^{(\beta)})_{T,p}\}(\delta N^{(\alpha)})^2.\end{aligned}\quad (8.5)$$

由(8.5)式可见,稳定性判据 $\delta^{(2)}G>0$ 的条件,即相平衡的稳定性条件是

$$(\partial\mu/\partial N)_{T,p}>0.\quad (8.6)$$

(8.6)式的意义不难理解.如果由于某种偶然的原因,使 β 相的化学势高于 α 相,则会有净粒子从 β 相转入 α 相.根据(8.6)式,α 相的化学势将提高,最后消除两相化学势之差,两相重新回到平衡.这样的平衡是稳定的.反之,若 $(\partial\mu/\partial N)_{T,p}$

<0，则 α 相的化学势会因粒子进入而进一步降低，更多的粒子进一步转入 α 相，最终使 α 相成为物质陷阱，以致不可能出现两相稳定平衡共存的情况.

以上的讨论是在认为两相已达热平衡和力学平衡的情况下进行的，因此，全面地说，两相平衡的平衡条件和平衡稳定性条件是

$$T^{(\alpha)}=T^{(\beta)};\quad p^{(\alpha)}=p^{(\beta)};\quad \mu^{(\alpha)}(T,p)=\mu^{(\beta)}(T,p),\tag{8.7}$$

$$C_V>0;\qquad \kappa_T>0;\qquad \left(\partial\mu/\partial N\right)_{T,p}>0.\tag{8.8}$$

(二)相平衡条件的讨论　相图

下面通过对相平衡条件的讨论介绍研究相变的有力工具——相图. 相图在实际中可以根据实验数据绘制而成. 它不仅在物理，而且在化学、冶金中都非常重要.

1. 相平衡曲线

以上讨论表明，不是在任意温度和任意压强下，两相都能平衡共存，只有满足方程

$$\mu^{(\alpha)}(T,p)=\mu^{(\beta)}(T,p).\tag{8.9}$$

的状态 (T,p) 才是两相可能平衡共存的状态. 给定一个压强 p，由此式可解得一个相平衡温度. 例如在 1atm 下，水和汽平衡共存的温度是 100℃. 对于不同的 p，由(8.9)式解出不同的 T，得到以 p,T 为坐标的图上的一条曲线，叫做相平衡曲线，如图 2 所示. 这样的图叫相图. 若两相是气和液，液和固或固和气，相应的相平衡曲线分别叫蒸发曲线、熔解曲线或升华曲线.

2. 单相区

在相平衡曲线以外的点，两相的化学势不相等. 当系统处于图 2 中 A 点(位于 $\mu^{(\alpha)}>\mu^{(\beta)}$ 区域)所代表的压强和温度时，根据

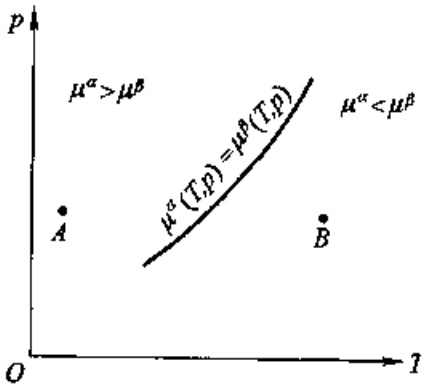


图 2　相平衡曲线

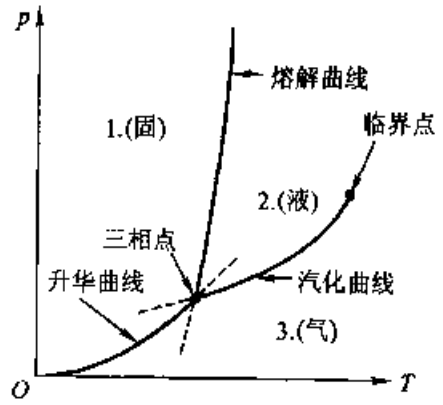


图 3　固液气的相图

$$G=G^{(\alpha)}+G^{(\beta)}=\mu^{(\alpha)}N^{(\alpha)}+\mu^{(\beta)}N^{(\beta)}$$

可知，只有全部粒子都在 β 相时(即 $N^{(\alpha)}=0,N^{(\beta)}=N$)，系统的自由焓才有最小值. 因而在这一区域中，$\beta$ 相是稳定相. 相反，当系统处于 B 点(位于 $\mu^{(\alpha)}<\mu^{(\beta)}$ 区域)所代表的压强和温度时，只有全部粒子都在 α 相时(即 $N^{(\alpha)}=N,N^{(\beta)}=0$)，系统的自由焓才有最小值. 因而在这一区域中，$\alpha$ 相是稳定相. 要注意的是，在稳定的液态区，仍可能存在亚稳的过饱和气；在稳定的气态区，仍可能存在亚稳的过热液

体(见下节).

3. 三相点

如果单元系统有 α、β 和 γ 三个相,在平衡时,每个相的化学势应该相等,即

$$\mu^{(\alpha)}(T,p) = \mu^{(\beta)}(T,p), \quad \mu^{(\beta)}(T,p) = \mu^{(\gamma)}(T,p).$$

这是 (T,p) 的两个方程. 当三个相共同处于平衡态时,这两个方程应该同时成立,要将它们联立求解. 这样就得到 p,T 的一组完全确定的值. 这表明,三相只能在唯一确定的压强和温度下平衡共存. 这一状态 (p_3,T_3) 称为三相点. 从相图上看,这一固定点就是三条平衡曲线的交点,如图 3 所示. 对单元系统,四个相不能同时共存,因为这时将有三个方程,两个未知量,方程无解. 要注意的是,单元系最多只能三相平衡共存,并不意味着单元系只能有三相. 例子见图 4 和图 5. 可以证明,对含 k 个组元的多元系,共存的相数不能超过组元数加 2,即有 $\varphi \leqslant k+2$ (相律).

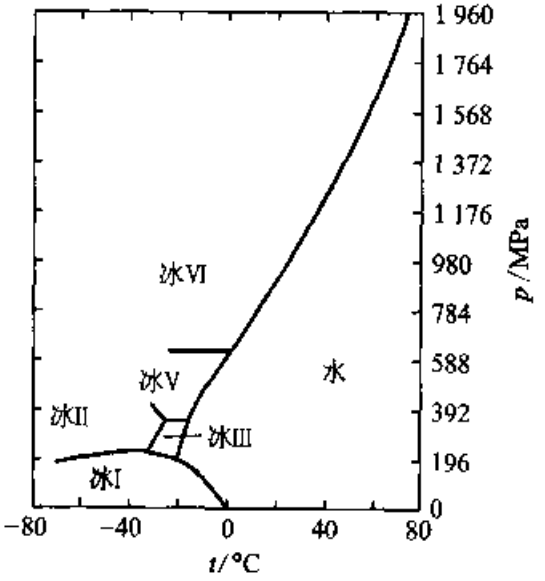


图 4　冰的相图

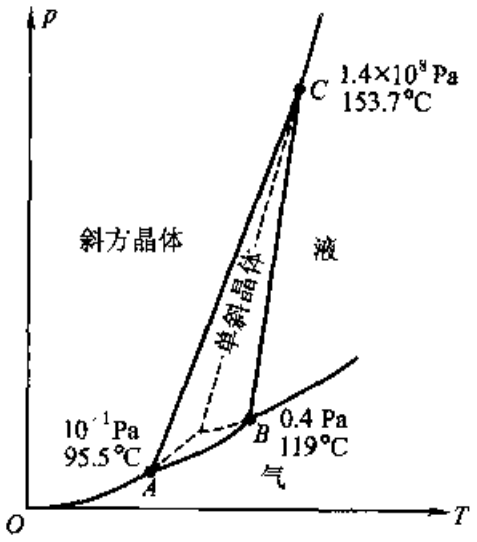


图 5　硫的相图

4. 临界点

1869 年美国物理学家安德鲁斯精确测量了 CO_2 在液态和气态时的密度差,发现在 31℃附近,两者的差别消失,液气转变的潜热也等于零. 这意味着蒸发曲线有一个明确的终点,安德鲁斯称之为临界点. 在图 3 中,用 C 点表示. 对应的温度和压强分别叫临界温度 T_C,和临界压强 p_C. 当温度高于临界温度时,再问物质究竟处于气态或液态没有意义. 而且,由于临界态的存在,从两相中的任意一相的某一状态出发,可以经绕过临界点的任意路径,连续地进行气液过渡,而不必经过两相分离(或两相共存)的状态.

和汽化曲线不一样,熔解曲线没有终点(至今未观察到). 原因是,液相和气相一样,没有晶态那种长程有序结构. 要从具有各向同性的液相连续地过渡到具有一定几何对称性的固相是不可能的.

图 3 是最简单的相图. 由于许多物质的固态可能有多种结构,因而有较复杂的相图. 图 4 是在高压下六种不同的冰,叫冰 Ⅰ,冰 Ⅱ,冰 Ⅲ,冰 Ⅳ(不稳定),冰 Ⅴ 和冰 Ⅵ的相图,它们按结晶而分. 图 5 是硫的相图. 硫的固态有单斜晶体和正交晶体两种结构,因而有三个三相点. 由图可见,每三种相只有一个三相点. 对于液态,一般说来一种物质只有一相,唯一的例外是 ^{4}He 和 ^{3}He. 当 $T = 2.2$ K 时,^{4}He 的液态有两相,分别叫做 He Ⅰ 和 He Ⅱ,前者是正常流体,后者是超流体. 当 $T = 3$ mK 时,^{3}He 也出现超流相.

表 Ⅰ　某些物质的临界点和三相点

物质	临界点		正常汽点		三相点	
CO_2	304K	73atm			216.6K	5.11atm
N_2	126.26K	33.54atm	77.32K	760mmHg	63.14K	94mmHg
H_2	33K	12.3atm	20.37K	760mmHg	13.96K	54.1mmHg
^{4}He	5.20K	1718mmHg	4.216K	760mmHg	2.186K	38.65mmHg
^{3}He	3.35K	890mmHg	3.195K	760mmHg		
H_2O	374.2℃		100℃	760mmHg		
乙醚	193.4℃					

1atm(大气压) = 760mmHg(毫米汞) = 101325Pa,见附录 Ⅶ

前面已经指出,一切物质的相图都需根据实验数据绘制而成. 热力学理论不能给出相图,但利用相平衡条件可以求得相平衡曲线的斜率.

5. 相平衡曲线的斜率　克劳修斯-克拉珀龙方程

考虑相平衡曲线上的两点 $A(T,p)$ 和 $B(T+\mathrm{d}T, p+\mathrm{d}p)$,则根据相平衡条件有

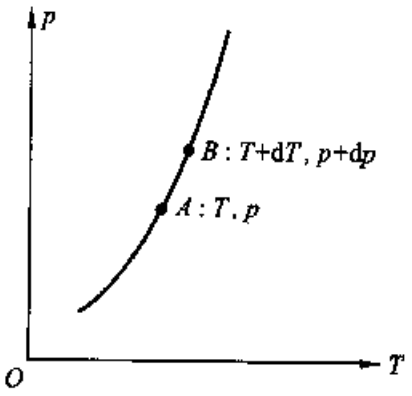


图 6　相平衡曲线的斜率

$$\mu^{(\alpha)}(T,p) = \mu^{(\beta)}(T,p),$$

$$\mu^{(\alpha)}(T+\mathrm{d}T, p+\mathrm{d}p) = \mu^{(\beta)}(T+\mathrm{d}T, p+\mathrm{d}p).$$

两式相减得到

$$(\partial\mu^{(\alpha)}/\partial T)_p\mathrm{d}T + (\partial\mu^{(\alpha)}/\partial p)_T\mathrm{d}p = (\partial\mu^{(\beta)}/\partial T)_p\mathrm{d}T + (\partial\mu^{(\beta)}/\partial p)_T\mathrm{d}p. \tag{8.10}$$

从单元开系的热力学基本方程 $\mathrm{d}G = -S\mathrm{d}T + V\mathrm{d}p + \mu\mathrm{d}N$ 和 $\mathrm{d}G = \mu\mathrm{d}N + N\mathrm{d}\mu$ 可得 $\mathrm{d}\mu = -(S/N)\mathrm{d}T + (V/N)\mathrm{d}p = -s\mathrm{d}T + v\mathrm{d}p$. 于是,单元系的摩尔熵(称为比熵) s

和摩尔体积(称为比容) v 分别为

$$s = -(\partial\mu/\partial T)_p, \qquad v = (\partial\mu/\partial p)_T. \tag{8.11}$$

将(8.11)式代入(8.10)式,最后有

$$\mathrm{d}p/\mathrm{d}T = (s^{(\alpha)} - s^{(\beta)})/(v^{(\alpha)} - v^{(\beta)}). \tag{8.12}$$

$\mathrm{d}p/\mathrm{d}T$ 代表相平衡曲线上一点的斜率,$s^{(\alpha)} - s^{(\beta)}$ 和 $v^{(\alpha)} - v^{(\beta)}$ 则代表在该点 1mol 物质从 β 相变到 α 相时熵和体积的变化. 因为平衡相变时系统的温度不发生变化,若用 $L_{\beta\to\alpha}$ 代表 1mol 物质从 β 相变到 α 相时吸收的热量,即摩尔潜热,则有

$$L_{\beta\to\alpha} = T(s^{(\alpha)} - s^{(\beta)}). \tag{8.13}$$

(8.12)式进一步化为

$$\frac{\mathrm{d}p}{\mathrm{d}T} = \frac{L_{\beta\to\alpha}}{T(v^{(\alpha)} - v^{(\beta)})}. \tag{8.14}$$

这就是一级相变的克劳修斯-克拉珀龙方程.

若将蒸气近似为理想气体,则有

$$v^{(气)} \approx RT/p \gg v^{(液)}. \tag{8.15}$$

因而 $\mathrm{d}p/\mathrm{d}T \approx pL_{液\to气}/T^2R$,或

$$\mathrm{d}p/p \approx L_{液\to气}\mathrm{d}T/RT^2. \tag{8.16}$$

取 $L_{液\to气}$ 为常数并积分上式得

$$p \approx p_0 \mathrm{e}^{-L_{液\to气}/RT}. \tag{8.17}$$

这叫饱和蒸气压方程.

一般说来,熔解总是吸收热量,即 $L_{固\to液} > 0$,或 $s_{液} > s_{固}$(但也有例外,如 $L^{He^3}_{固\to液} < 0$),但 $v_{液} - v_{固}$ 则可能大于零,也可能小于零(后一情况如 H_2O). 前一种情况下,熔解曲线的斜率为正,如果我们在不变的温度下增加压强,将驱使系统更深地进入固相. 在后一种情况下,则熔解曲线的斜率为负,如果在不变的温度下增加压强,将驱使系统从固相变到液相. 当冰刀施压于冰上时,冰化为水,滑冰者可以沿一狭窄的水槽前进,摩擦力很小.

§7.2.9 一级相变和二级相变

相变是自然界普遍存在的一种突变现象,是量变到质变辩证规律的表现. 在长期研究中,人们发现,相变过程大体可分为两类,即一级相变和二级相变. 二级相变又往往叫做连续相变. 二级相变点叫做临界点,因而二级相变时发生的现象叫临界现象. 本节首先以临界温度 T_c 以下的气-液等温转变为例说明一级相变的特点,然后通过无外场时改变温度而导致的正常导体-超导体转变为例引进二级相变的概念. 再讨论相变的分类. 最后简单介绍二级相变的几个例子.

1873 年荷兰人范德瓦耳斯在其博士论文中提出了范德瓦耳斯方程. 从历史看，他是从微观上研究气、液性质和气-液相变的第一人. 从方法看，范氏方程提供了从微观模型研究宏观性质，用热力学判据研究相变的范例. 随着研究的深入，人们发现，范氏方程不仅对气、液性质，气、液转变，而且对整个平衡相变，甚至非平衡相变的研究都有重要的指导意义. 因此，在介绍气液等温转变时，我们还将对范氏方程作较详细的分析. 第八篇将给出范氏方程的微观推导.

关于临界现象及其微观本质，则将在第九篇进一步研究.

(一)气-液等温转变的实验曲线　一级相变的特点

图 1 是气-液等温转变的示意图. 图 a 中的 $KBAN$ 线是 $T < T_c$ 下 1mol 物质的一条等温线，图 b 是这一过程在 $p - T$ 图上的投影.

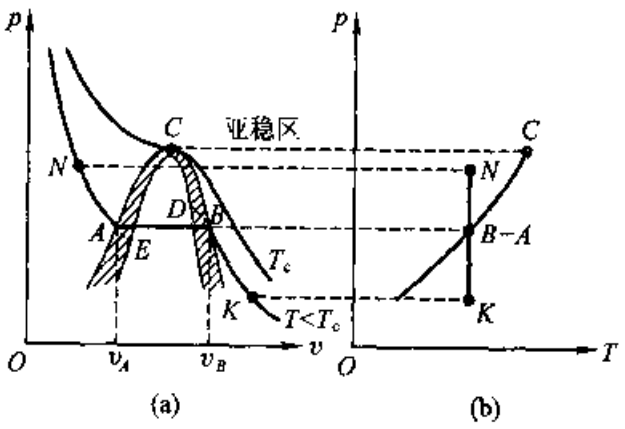


图 1　气-液等温转变的实验曲线

KB 表示等温压缩时，气体压强逐渐增大. 到达 B 点时，气体开始凝聚，同时放出热量. 这时尽管放出热量，温度却不降低，可见放出的热量由气体凝聚而来，是潜热. 继续压缩，凝聚的液体愈来愈多，而压强保持不变，直到 A 点，气体全部液化. 在 p-v 图上，v_B 表示这一温度和压强下气体的比容，v_A 表示液体的比容，水平线 BA 表示相变过程. 在这一过程中，虽然气、液的比例不断变化，但两相一直处于平衡，叫平衡相变，它在 p -T 图上由蒸发曲线上的一点代表. 由于液体的不可压缩性，到达 A 点后继续压缩时，压强迅速增大，在 p -v 图上由 AN 表示，在 p-T 图上则由垂直于 T 轴的 $B(A) - N$ 直线代表.

在临界温度以下 ($T < T_c$)，气 - 液等温转变有以下特点：

1. 在相变点 (T,p)，两相的化学势相等，但比容和比熵不等(因为存在潜热)，故化学势的一阶偏导数不连续. 即

$$\mu^{气}(T,p) = \mu^{液}(T,p), \tag{9.1}$$

$$v^{气}(T,p) \neq v^{液}(T,p), \quad 或 \quad (\partial\mu^{气}/\partial p)_T \neq (\partial\mu^{液}/\partial p)_T, \tag{9.2}$$

$$L = T(s^{气} - s^{液}) \neq 0, \quad 或 \quad (\partial\mu^{气}/\partial T)_p \neq (\partial\mu^{液}/\partial T)_p. \tag{9.3}$$

存在潜热或比熵不等这一事实，和比容不等一样，也可从 p -v 图上看出(见图 2). 因为 $dT = 0$，我们有 $ds = (\partial s/\partial v)_T dv = (\partial p/\partial T)_v dv$，其中，已用了麦克斯韦关系. 从 A 点到 B 点比熵的变化可以用以下积分计算：

$$\Delta s = \int_A^B (\partial p/\partial T)_v \mathrm{d}v. \tag{9.4}$$

考虑温度 T 和 $T+\delta T$ 下的两条曲线，用 δp 表示在一定 v 两条线之间的压强差，则上式可写为

$$\Delta s = \frac{1}{\delta T}\int_A^B \delta p \delta v = \frac{v_A \text{ 和 } v_B \text{ 间两条等温线之间的面积}}{\delta T}. \tag{9.5}$$

此面积（图 2 中阴影面积）为正，因而气相的比熵比液相大，气化时要吸收热量. 由图还可看出，随着温度升高，相应面积减小，说明潜热随温度升高而减小，达到临界点时，潜热为零.

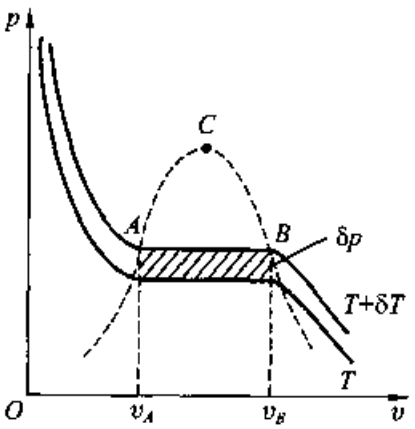


图 2　一级相变存在潜热

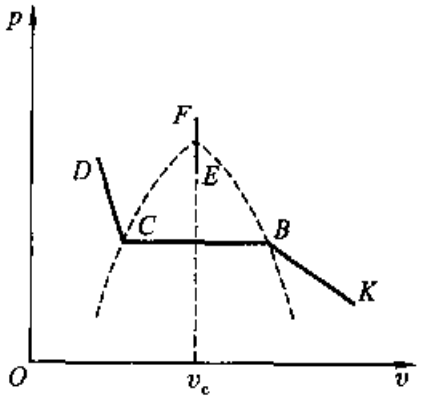


图 3　气液系统的一级和二级相变 KBCD 一级相变　EF 二级相变

2. 在相变点的定压比热为无穷大. 因为在相变过程中，虽然有潜热，但温度不变，根据定义 $c_p = (\mathrm{d}Q/\mathrm{d}T)_p \to \infty$.

3. 在 p-T 图上，相平衡曲线的斜率服从克劳修斯-克拉珀龙方程（见上节）.

4. 存在着亚稳态. 实验发现，对纯净的蒸气进行压缩时，往往会发生这种现象，即相变线已进入液化区（如图 1 中的虚线 BD），可是系统仍保持为气态. 但是，这种状态是不稳定的，稍有扰动立即会凝聚成液体. 这种蒸气由于其压强高于与实际温度对应的饱和压强，而叫过饱和蒸气. 另一方面也会发生这样的情况，相变线已进入气化区（如图 1 中的 AE 线），可是仍然保持为液态. 这种状态也是不稳定的，稍有扰动，马上会变为蒸气. 这种液体由于其温度高于与实际压强对应的平衡温度，而叫过热液体. 过饱和蒸气和过热液体都是亚稳态. 能够存在亚稳态是一级相变的又一特点.

$T < T_c$ 时气-液等温转变的以上性质，实际上为三态转变，同素异晶转变等所共有，其中最基本的性质是第一点. 由于化学势的一级偏导数不连续，故把这类相变叫做一级相变.

需要注意的是，对于气-液系统，除一级相变外，也存在二级相变. 在临界温度以下，保持温度不变增加或减小体积，发生的是一级相变，如图 3 中的 $KBAN$；而在

临界点附近，保持体积为临界体积 v_C 不变而改变温度（因而也改变压强），通过临界点时，气-液系统发生二级相变（图 3 中的 EF），参看 §7.2.9 的表 I 第一列. 低于临界温度，低于临界压强时（图 3 中的 E 点），系统处于气液共存相；而高于临界温度，高于临界压强时（图 3 中的 F 点），系统处于气液不分的均匀状态. 临界点是这个二级相变的相变点. 在 §9.2.1 中将要详细讨论气液系统的二级相变和临界现象. 下面，在第三点里，先以外场为零时的超导相变为例介绍二级相变.

（二）范德瓦耳斯等温线的热力学分析

1 mol 气体的范德瓦耳斯方程 $(p+a/v^2)(v-b)=RT$ 在 p-v 图上的等温线如图 4 所示（见习题 33）. 我们将讨论的重点放在范德瓦耳斯方程的适用性，与实验结果的区别及如何对这种区别用热力学理论进行解释和修正上.

1. 液相区和气相区

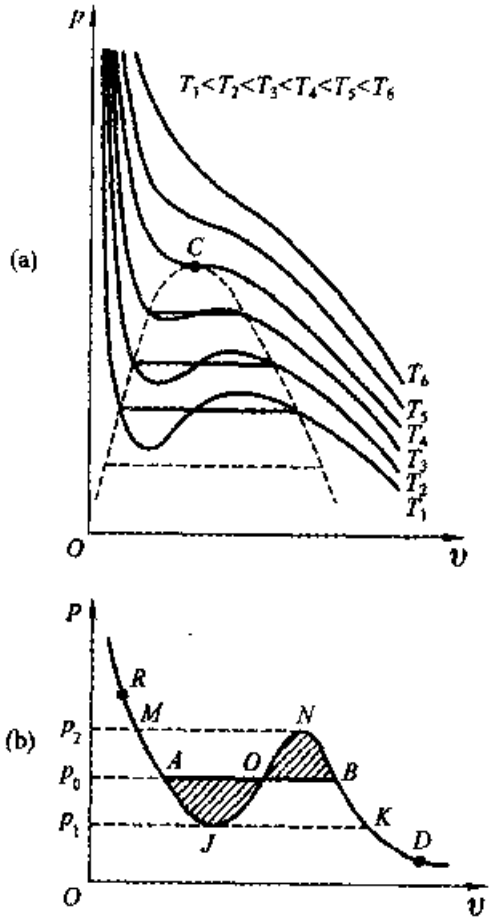


图 4　范德瓦耳斯等温线

由图 4b 可见，当 $p>p_2$ 和 $p<p_1$ 时，和每一个 p 值对应只有一个 v 值，且皆满足力学稳定性条件 $(\partial v/\partial p)_T<0$［§7.2.7(7.26)式］. 因此，这两段等温线代表稳定单相区. 根据气体的 κ_T 大，液体的 κ_T 小的特点，还可以知道 DK 段代表稳定气相，MR 段代表稳定液相，即范氏等温线能较好地描述气相和液相的行为.

2. 不稳定态　等面积法则

再来考虑中等压强，即 $p_1<p<p_2$ 的情况. 这时，与每一 p 值对应有三个 v 值. 在 NOJ 段，由于 $(\partial v/\partial p)_T>0$，不满足力学稳定性条件，所以是实际上不可能实现的状态. 这就可以解释为什么实际的气-液等温转变不能完全沿范氏等温线进行. 不仅如此，用热力学理论还可以进一步确定实际相变中气-液两相共存时的压强. 设 AOB 代表实际的气-液平衡相变过程，则有

$$\mu_A=\mu_B,\qquad 或\qquad \mu_A-\mu_B=0.$$

另一方面，由 1mol 物质封闭系的热力学基本方程 $\mathrm{d}\mu=-s\mathrm{d}T+v\mathrm{d}p$ 知道，化学势沿范氏等温线的变化可由以下积分求得：

$$(\Delta\mu)_T = \int (v\mathrm{d}p)_T. \tag{9.6}$$

当给出积分的上下限时，等式右边代表范氏等温线与 p 轴间的面积，此面积可正可负，由 $\mathrm{d}p$ 的符号决定. 由平衡条件 $\mu_A = \mu_B$ 有

$$\mu_A - \mu_B = \int_{BNOJA} v\mathrm{d}p = 0.$$

这一积分可分为几部分：

$$\begin{aligned}\int_{BNOJA} v\mathrm{d}p &= \int_B^N v\mathrm{d}p + \int_N^O v\mathrm{d}p + \int_O^J v\mathrm{d}p + \int_J^A v\mathrm{d}p \\ &= \left(\int_B^N v\mathrm{d}p - \int_O^N v\mathrm{d}p\right) + \left(-\int_J^O v\mathrm{d}p + \int_J^A v\mathrm{d}p\right) \\ &= \text{面积}(BNOB) - \text{面积}(OJAO) = 0,\end{aligned} \tag{9.7}$$

其中第一项代表直线 BO 和曲线 BNO 所围面积，第二项代表直线 OA 和曲线 OJA 所围面积. 于是，得出结论：决定气-液平衡共存压强 p_0 的水平线 BOA，应画得使面积 $BNOB$ 和面积 $OJAO$ 相等. 这首先由麦克斯韦在推荐范德瓦耳斯的论文时提出，故叫麦克斯韦等面积法则.

3. 亚稳态

在排除等温线上的不稳定态（即 NOJ 段）之后，与同一 p 值对应还有两个可能的 v 值，如图 5 中的 a 点和 b 点. 这两点具有相同的温度和压强. 为确定两者中谁更稳定，可利用自由焓判据，比较它们的化学势. 为此，我们把范德瓦耳斯等温线的 p -v 图转过 90 度并将对应的 μ -p 图画在 v -p 图的下部.

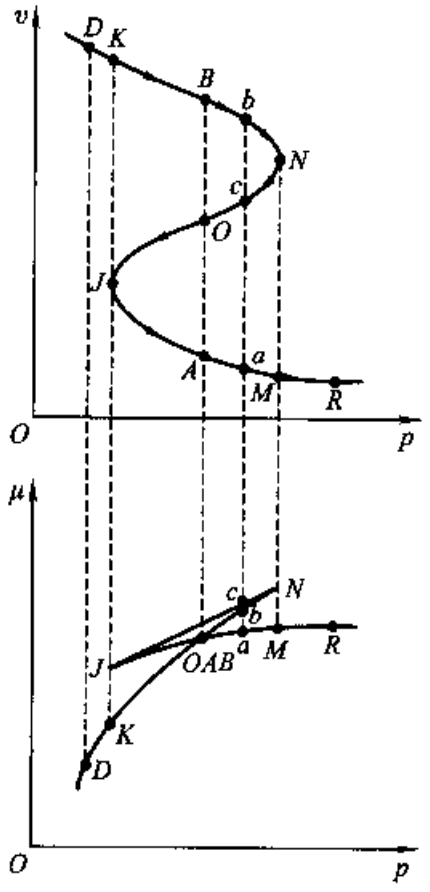


图 5　沿范氏等温线化学势 μ 随 p 的变化

由图看到，当系统的状态沿 $DKBN$ 变化时，曲线与 p 轴所夹面积为正，且逐渐增大. 根据 (9.6) 式知道，这说明系统的化学势增加，直到 N 点. 当沿 NOJ 进行时，面积为负，所以 J 点的化学势较 N 点低，但比 K 点高. 以后沿曲线 $JAMR$ 进行，面积为正，化学势再次增加，趋向 M 点，但 M 点的化学势比 N 点低.

由 μ -p 图可见，对于具有相同温度和压强的 a，b，c 三点，其化学势的关系是

$$\mu_a < \mu_b < \mu_c.$$

所以处于 AM 上的状态（例如 a 点）化学势最低，属于稳定液相；处于 JON 上的状

态(例如 c 点)化学势最高,是不稳定的;而处于 BN 上的状态(如 b 点)化学势较低,是亚稳态,与前面提到的过饱和蒸气对应.类似的比较说明,KB 代表稳定气相,JA 则与亚稳的过热液体对应.而 A 点和 B 点的化学势相等,从热力学上看,它们出现的可能性相同,所以在此将发生相分离.

4.临界点

如图 4 所示,范德瓦耳斯等温线存在着极大值和极小值.随着温度的升高,极大值和极小值逐渐靠拢,当温度达到 T_C 时,两者重合形成拐点:

$$(\partial p/\partial v)_T = 0, \quad \text{和} \quad (\partial^2 p/\partial v^2)_T = 0. \tag{9.8}$$

将范德瓦耳斯方程代入(9.8)式,可得

$$(\partial p/\partial v)_T = -\frac{RT}{(v-b)^2} + \frac{2a}{v^3} = 0,$$

$$(\partial^2 p/\partial v^2)_T = \frac{2RT}{(v-b)^3} - \frac{6a}{v^4} = 0.$$

同范德瓦耳斯方程本身联立求解,可得到临界态的态参量:

$$v_C = 3b, \quad T_C = 8a/(27Rb), \quad p_C = a/(27b^2). \tag{9.9}$$

p_C,v_C 和 T_C 之间有以下关系:

$$RT_C/(p_C v_C) = 8/3 = 2.667. \tag{9.10}$$

这一无量纲的比值叫临界系数.根据范德瓦耳斯方程,此系数对一切物质都等于 2.667.但实际上,对不同物质这一系数有不同值,且与 2.667 相差甚大.这说明,虽然范德瓦耳斯等温线能预言临界态的存在,但定量上却是不精确的.

(三)正常导体和超导体之间的转变

在 §3.2.5 中已介绍过超导体的基本电磁现象,应用二流体模型及 BCS 理论推导了超导体的基本电磁方程,并指出超导态和正常态之间的转变是一种相变.现在我们来研究这种相变的性质.为了方便分析,先复习有关的实验结果,并强调某些细节.

1.实验结果

迄今为止,已发现 28 种金属元素和许多合金甚至天然纯金属和非金属的化合物都具有超导性,即当温度降低到一特定值 T_C 时,样品的电阻会突然下降到零.表 I 给出了一些超导材料的临界温度 T_C.将已处于 $T < T_C$ 的超导体置于外磁场中,当场值升到超过特定值 H_C 时,超导态又会遭破坏而变为正常导体.H_C 称为临界场,和温度有关 $H_C = H_C(T)$,也随材料而异.实验测得临界

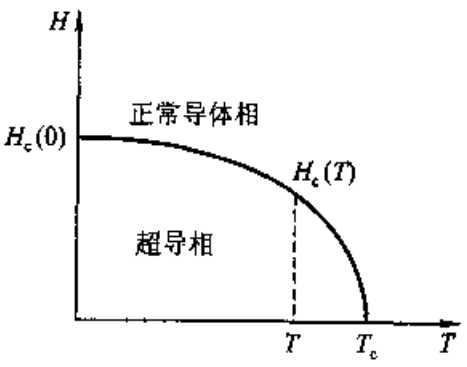


图 6　超导体的临界场曲线

场与温度的关系近似为：

$$H_C(T) \approx H_C(0)[1-(T/T_C)^2], \tag{9.11}$$

如图6所示，$H_C(0)$ 表示绝对零度的临界场. 临界场曲线是超导相和正常导体相的分界线，也是两相平衡共存线，相当于前面的蒸发曲线 $p = P(T)$. 在图7上给出了几种第Ⅰ类超导元素的临界场曲线. 我们看到，对于临界场曲线（共存线）的斜率有：当 $0 < T \leqslant T_C$ 时，$dH_C/dT < 0$，但为一有限值. $T \to 0$ 时，$dH_C/dT \to 0$. 实验表明，在 $T < T_C$ 时，改变磁场而在 $H = H_C(T)$ 下发生正常导体⇌超导转变时有潜热；而在 $H = 0$ 改变温度，在 $T = T_C$ 下发生正常导体⇌超导转变时，没有潜热.

表Ⅰ　超导临界温度

超导材料	T_C (K)	超导材料	T_C (K)
Hg(α)	4.15	Pb	7.20
Nb	9.25	V$_3$Si	17.1
Nb_3Sn	18.1	$Nb_3Al_{0.75}Ge_{0.25}$	20.5
Nb_3Ga	20.3	Nb_3Ge	23.2
$YBaCu_3O_7$	90	$Bi_2Sr_2Ca_2Cu_3O_{10}$	105
$Tl_2Ba_2Ca_2Cu_3O_{10}$	123	$HgBa_2Cu_3O_8$	134

超导体的另一基本性质是理想抗磁性，即超导体内部磁感应强度为零 $B_s = 0$. 因为 $B = H + 4\pi m$，故对超导体有

$$m_s = -H/4\pi. \tag{9.12a}$$

而正常导体一般导磁率 $\mu \approx 1$，故由 $B = \mu H$ 有 $B_n = H$，或

$$m_n = 0. \tag{9.12b}$$

以上两式可视为超导体和正常导体的状态方程. 脚标 s 和 n 分别代表超导相和正常导体相.

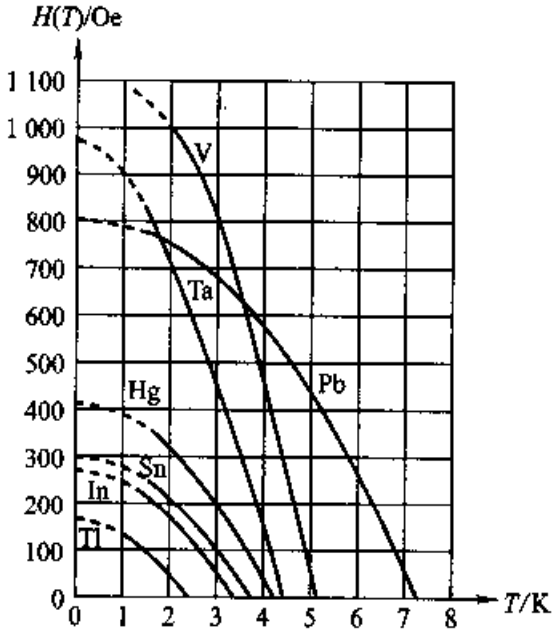


图7　几种金属的临界场曲线

1Oe = 79.5775A · m^{-1}

2. 外场导致的超导相变——一级相变

在忽略体积变化的情况下，1mol 磁介质的热力学基本方程可写为

$$d\mu = -sdT - mdH, \tag{9.13}$$

$$s = -\left(\frac{\partial \mu}{\partial T}\right)_H, \qquad m = -\left(\frac{\partial \mu}{\partial H}\right)_T.$$

在 $T < T_C, H = H_C(T)$ 下，由实验已知，正常导体⇆超导转变中有潜热，这意味着 $s_s \neq s_n$ 或 $(\partial \mu_s/\partial T)_H \neq (\partial \mu_n/\partial T)_H$．又由(9.12a)和(9.12b)可知有 $m_n = 0$ 和 $m_s = -H_C(T)/4\pi$ 或 $(\partial \mu_s/\partial H)_T \neq (\partial \mu_n/\partial H)_T$．

我们再来推导临界场曲线斜率的表达式．根据相平衡条件，在 (T, H_C) 和 $(T + dT, H_C + dH_C)$ 时，两相的化学势相等，因而有

$$d\mu_n = d\mu_s.$$

将(9.13)式代入得 $-s_n dT - m_n dH_C = -s_s dT - m_s dH_C$，因而

$$-dH_C/dT = (s_s - s_n)/(m_s - m_n). \tag{9.14}$$

考虑潜热 $L_{n\to s} = T(s_s - s_n)$，并将(9.12)式代入，最后得

$$\frac{dH_C}{dT} = \frac{4\pi L_{n\to s}}{H_C T}, \qquad 或 \qquad s_s - s_n = \frac{H_C}{4\pi}\frac{dH_C}{dT}. \tag{9.15}$$

(9.15)式是超导体的克劳修斯-克拉珀龙方程．将临界场曲线的斜率值代入，计算出的潜热值 L，和直接由卡路里实验测到的潜热值很好地一致．

在 $T < T_C$ 时，因 $dH_C/dT < 0$，所以 $s_s - s_n < 0$，即和正常态相比超导态是有序相．保持温度不变，改变磁场 H，如图6中的虚线，则当 H 从大于 H_C 降到 H_C 以下时，要放出热量．

以上讨论说明，当 $T < T_C$ 时，外场导致的超导转变是一级相变．

3．无外场时改变温度导致的超导相变——二级相变

在不加外磁场时，将温度降到 T_C 也会发生从正常态向超导态的转变．这一转变和一级相变有完全不同的特点，是二级相变的一个典型例子．

由(9.12)式和(9.15)式可以看到，当 $T = T_C, H_C = 0$，dH_C/dT 有限时，有 $m_s - m_n = 0, s_s - s_n = 0$，即在没有外场作用下让温度降低通过 T_C 使金属从正常态转变为超导态时，没有潜热释放，磁化强度也不发生变化，但比热会发生跃变．为了证明后一点，将(9.15)式对 T 求导并乘以 T：

$$c_s - c_n = T\partial(s_s - s_n)/\partial T = (T/4\pi)[(dH_C/dT)^2 + H_C(d^2H_C/dT^2)],$$

再利用 $H_C = H_0[1 - (T/T_C)^2]$

$$dH_C/dT = -2H_0T/T_C^2, \qquad d^2H_C/dT^2 = -2H_0/T_C^2. \tag{9.16}$$

得到，

$$c_s - c_n = -(H_0^2/2\pi)(1/T_C)[(T/T_C) - 3(T/T_C)^3]. \tag{9.17}$$

当 $T = T_C$ 时，

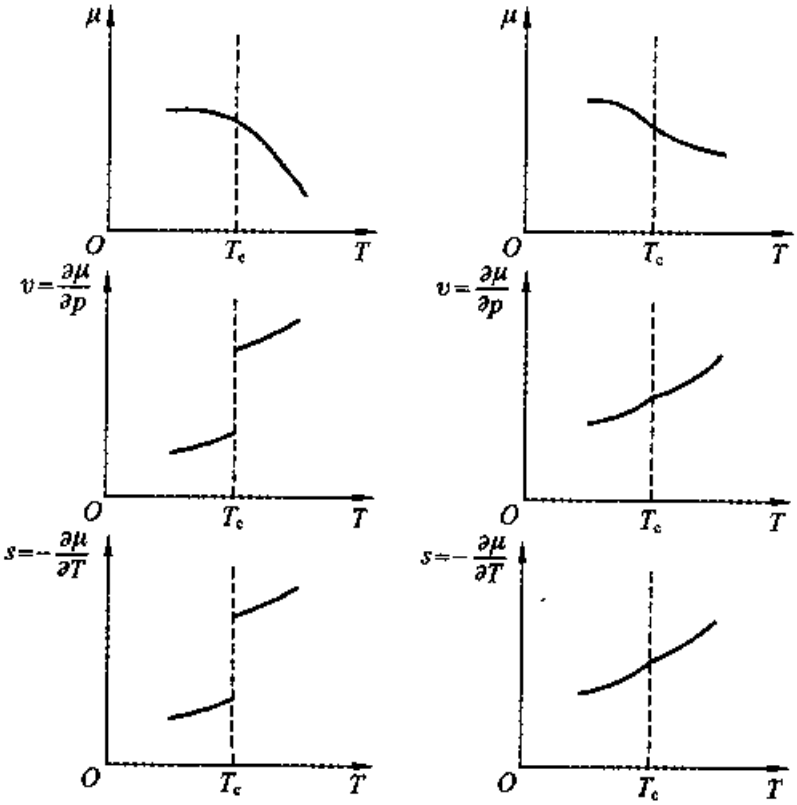
$$(c_s - c_n)_{T=T_C} = H_0^2/(\pi T_C). \tag{9.18}$$

这表明，两相的比热有一个有限的跳跃，且超导相的比热比正常相高．这是二级相

变的特征. 值得注意的是,当 $T\to 0$ 时,因为 $\mathrm{d}H_C/\mathrm{d}T\to 0$,因而 $s_s - s_n \to 0$. 这和热力学第三定律一致(见下节).

(四)相变的分类

前面已经指出,一级相变的特点是化学势连续而化学势的一阶偏导数不连续,即有相变潜热且比容发生改变,见(9.1)~(9.3)式.



(a) 一级相变　　(b) 二级相变

图 8　相变分类

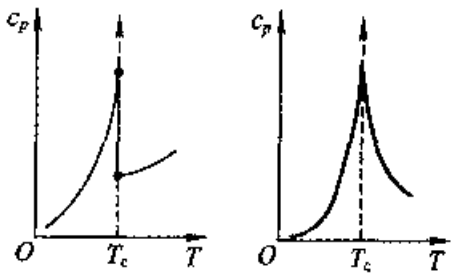
二级相变的特点是,在相变过程中,既没有潜热的吸收或释放,比容也不改变,但物质的比热发生突变,膨胀系数 α 和压缩系数 κ_T 也发生突变,甚至有趋于无穷或尖峰的奇异行为. 如果用热力学关系表示,这类相变的特点是:

$$
\begin{aligned}
&\mu_1 = \mu_2, \\
&v_1 = v_2, \quad \text{即} \quad (\mu_1/\partial p) = (\partial\mu_2/\partial p), \\
&s_1 = s_2, \quad \text{即} \quad -(\partial\mu_1/\partial T) = -(\partial\mu_2/\partial T), \\
&c_{p1} \neq c_{p2}, \quad \text{即} \quad (\partial^2\mu_1/\partial T^2) \neq (\partial^2\mu_2/\partial T^2), \\
&\alpha_1 \neq \alpha_2 \quad \text{即} \quad (\partial^2\mu_1/\partial T\partial p) \neq (\partial^2\mu_2/\partial T\partial p), \\
&\kappa_1 \neq \kappa_2 \quad \text{即} \quad (\partial^2\mu_1/\partial p^2) \neq (\partial^2\mu_2/\partial p^2).
\end{aligned}
\tag{9.19}
$$

就是说,化学势本身和它的一级偏导数连续,但是它的二级偏导数所代表的性质发

生突变，因而这类相变称为二级相变.

由于在相变点，两相的 T,p,μ,s,v 相等，从而态函数 U,F,H 亦相等，即在发生这种相变时，两相的热力学状态是连续变化的，这是把二级相变又叫做连续相变的原因. 一级相变与二级相变的区别可以用图形形象地表示. 为此，我们画出 $\mu,\partial\mu/\partial p,-\partial\mu/\partial T$ 随 T 的变化曲线，如图 8. 因 v 和 s 均随 T 的升高而增大，所以 $\partial\mu/\partial p$ 和 $-(\partial\mu/\partial T)$ 的曲线右正的斜率.



(a)　有限跃变　(b)　趋于无穷

图 9　二级相变的比热

二级相变又分为两种情况，即：在相变点，$c_p=\partial^2\mu/\partial T^2$ 有有限的跃变和趋于无穷，如图 9. 无外场时的超导相变属于前一种情况；铁磁体在温度升高时过渡到顺磁体的相变属于后一种情况. 另外要注意，和一级相变不同，在二级相变点的每一侧，只有一种相能够存在，没有亚稳态.

此外，在二级相变点(临界点)邻域，不同系统的许多热力学特性表现出极大的相似性，某些热力学量的涨落反常地增大，这些将在第九篇介绍.

(五)二级相变的其他例子

(1)He Ⅰ和 He Ⅱ的转变

He 原子有两种稳定的同位素 ^{4}He 和 ^{3}He，它们的原子序数都是 2，但质量数分别为 4 和 3. 两种同位素在自然界中的比例是 $10^5:1$. 它们的物理性质截然不同.

图 10 是由实验得到的 ^{4}He 的相图. 由图可见，若保持 2.25atm 不变而降低温度到 5.25K(He 的临界温度)时，氦气开始液化. 在保持两相共存的条件下继续降低温度，同时改变压强，当达到图 10 上的 λ 点 ($T_\lambda=2.17$K，$p_\lambda=0.058$ atm)时，氦虽保持为液态，性质却发生了惊人的变化. 通常把 λ 点以上的液氦称为 He Ⅰ，λ 点以下的液氦称为 He Ⅱ. He Ⅱ具有超流动性和超导热性. He Ⅱ的粘滞系数大约是 He Ⅰ的 10^{-6} 倍，是水的 10^{-9} 倍，在几 cmHg 的压差下，He Ⅱ通过几 μm 的间隙时，流速之快几乎无法测量. He Ⅱ的导热系数是 He Ⅰ的 10^6 倍，比导热性最好的铜还高上千倍. 图 10 中 BD 线是 He Ⅰ和 He Ⅱ的平衡曲线. 沿 BD 线两相的比熵和比容都相等. 但 c_p 趋于无限，即 He Ⅰ和 He Ⅱ的转变是二级相变. 在不同压强下，He Ⅰ-He Ⅱ相变(二级相变)的临界点发生在不同的温度，形成一根线，称为 λ 线. 利用二级相变在邻近两点 (T,p) 和 $(T+\mathrm{d}T,p+\mathrm{d}p)$ 两相比熵和比容相等的条件(即 $\mathrm{d}s^{(1)}=\mathrm{d}s^{(2)},\mathrm{d}v^{(1)}=\mathrm{d}v^{(2)}$)可导出 λ 线的斜率满足的公式：

$$\frac{\mathrm{d}p}{\mathrm{d}T}=\frac{c_p^{(2)}-c_p^{(1)}}{Tv(\alpha^{(2)}-\alpha^{(1)})},\qquad 和\qquad \frac{\mathrm{d}p}{\mathrm{d}T}=\frac{\alpha^{(2)}-\alpha^{(1)}}{\kappa^{(2)}-\kappa^{(1)}}.\tag{9.20}$$

它们被称为爱仑菲斯特方程.

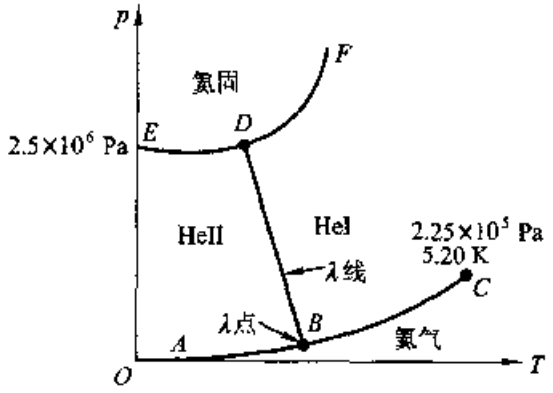


图 10　⁴He 的相图

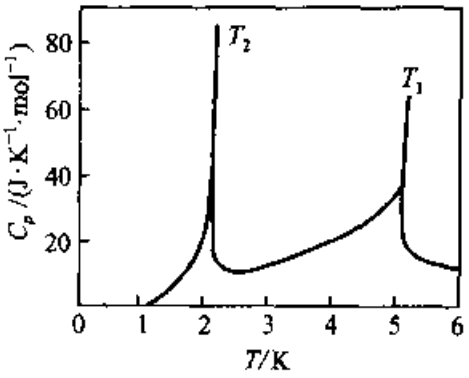


图 11　⁴He 的比热实验曲线

将⁴He 在 λ 点的数据代入上式,计算出 λ 线在 λ 点的斜率和相图上的斜率值很好的符合.

AB 线和 *BC* 线分别代表 HeⅡ,HeⅠ 和氦气的平衡曲线,它们都是一级相变,*AB* 线和 *BC* 线在 λ 点光滑联接, $\mathrm{d}p/\mathrm{d}T\big|_{AB}^{\lambda点} = \mathrm{d}p/\mathrm{d}T\big|_{BC}^{\lambda点}$. 根据克劳修斯-克拉珀龙方程容易看出,这正是 HeⅠ 和 HeⅡ 比熵、比容相等的表现. *EDF* 线是 HeⅡ、HeⅠ 和固态氦的共存曲线. 液、固之间的转变也是一一级相变. 在 25atm 以下,即使温度降低到 0K,氦仍保持液态而不固化. 它是在 $T \to 0$ K 时唯一能保持液态的物质. 由图还可以看到,固-液平衡曲线以零斜率趋向 p 轴,符合热力学第三定律的要求(见下节).

图 11 是⁴He 的比热实验曲线,这条曲线恰巧选在经过气液临界点的等容线上. 在 5K 附近的第一个尖峰对应气液相变的临界点,第二个尖峰对应超流相变的临界点.

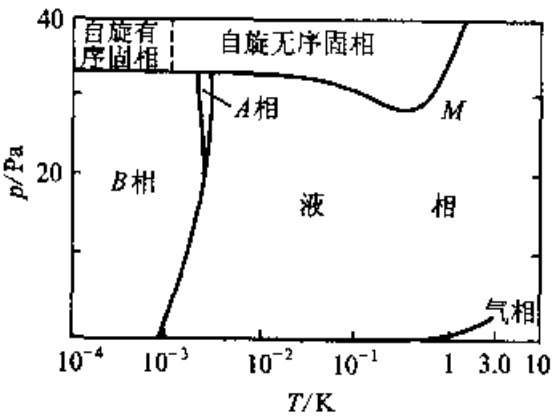


图 12　³He 的相图

图 12 给出了³He 的相图,³He 首次被液化是在 1948 年,比⁴He 晚了 40 年. 在很长时间内,人们以为³He 没有超流相,直到 1971 年才在实验中发现了它的超流相. 由于³He 的原子是费米子,故只有两原子形成束缚对时,才可能出现超流相. 因束缚对自旋取向不同,使得有不止一个超流相. 如图 12 中的 *A*,*B* 为两种不同超流相. 在一定条件下,不同超流相之间,正常相与超流相之间可以互相转变,但这些相变的性质还有待研究.

(2)铁磁体和顺磁体的转变

按照磁化率$\chi = m/H$(m代表磁化强度,即单位体积中的磁矩),物质可分为三类:

若$\chi < 0$,$|\chi|$的值约为10^{-6}左右,称为抗磁质,如惰性气体和非金属.

若$\chi > 0$,且χ和温度的关系可表为$\chi = c/T$,称为顺磁质,如铂、稀土元素等.

若$\chi > 0$,且很高,χ和温度的关系可表为$\chi = c/(T - T_C)$,称为铁磁质,如Fe,Co,Ni,Gd及多种含有这类元素的合金.

铁磁质最显著的特性是自发磁化强度.在特定温度(居里温度T_C)以下,即使没有外场,铁磁体中也存在一些宏观大小的磁畴.磁畴中有自发磁化强度$m_{自}$.纯铁具有体心立方晶格,每个原子有八个最近邻.在铁中,立方轴是最易磁化的轴,即磁畴中的自发磁化强度往往是$\pm x$, $\pm y$, $\pm z$中的一个,由偶然因素决定.单轴铁磁体的自发磁化强度只能平行或反平行于一个特定轴.整个铁磁体的磁化强度是各磁畴中自发磁化强度的矢量和.在没有外场时,这个矢量和为零.

当温度升高到$T > T_C$时,$m_{自}$变为零,磁畴破坏,铁磁性突然消失,铁磁体变为顺磁体.实验表明,这一转变为二级相变.相变温度叫居里温度T_C.Fe和Ni的居里温度分别是1044.0K和631.58K.$H = 0, T = T_C$是相变点或临界点,称为居里点.在居里点附近,热容量呈现出特征性的λ尖,所以往往又称为λ相变.表Ⅱ给出了二级相变的几个例子.

表Ⅱ　二级相变的例子

相变	序参量	例子	T_C(K)
液(气)	$\rho_{液} - \rho_{气}$	H_2O	647.05
铁　磁	磁化强度	Fe	1044.0
反铁磁	子晶格磁化强度	FeF_2	98.26
超　流	He原子的量子概率幅度	^{4}He	1.8~2.1
超　导	电子对的量子概率幅度	Pb	7.19
双　液	二种液体的浓度差	n乙烷-硝基苯	290
二元合金	次晶格中某组元的密度	Cu-Zn	739

§7.2.10 热力学第三定律

在历史上,热力学第三定律是德国的物理化学家能斯特于1906年从许多低温化学实验中总结出来的.由于第三定律的实验基础不如第一、第二定律那样丰富,加之有些实验结果似乎和第三定律不合,以致在一段时间里,许多研究者对第三定律的正确性和普适性曾产生怀疑.而统计物理的发展不仅揭示了第三定律的意义,并能解释某些情况和第三定律不一致的原因,从而结束了关于第三定律的争论,建立了一致的看法.

(一)热力学第三定律的能斯特表述及其统计论证

热力学第三定律有多种表述,通常采用能斯特表述:当绝对温度趋于零时,任何系统的熵趋于不变的极限值,不管其他参量如何(如压强、体积、磁场强度等),也不管处于什么相(如正常导体或超导体,固态氦或液态氦),这个常数可取为零,即

$$\lim_{T\to 0} S = S(0) = 0. \tag{10.1}$$

热容量 $C_V > 0$ 这一事实表明,能量是温度的单调增加函数;反之,当温度降低时,能量单调地减少.在绝对零度时,系统将处于最低能态,即基态.对于很多已知系统(具有不同晶格结构),基态是非简并的,$\Omega_0(E_0, N, V) = 1$,和系统的大小无关.于是,根据熵的玻尔兹曼公式,在绝对零度时,系统的熵为

$$\lim_{T\to 0} S = S(0) = k_B \ln \Omega_0 = 0.$$

如果最低能态不是非简并的,只要 $\Omega_0 \leqslant N$(系统中的总粒子数),在绝对零度仍有

$$\lim_{T\to 0} S = S(0) \leqslant k_B \ln N.$$

由于 k_B 的数值很小,$S(0)$ 实际上也等于零.例如,取 Ω_0 为 $N \approx 10^{23}$ 的数量级时

$$S(0) \sim 23 k_B \ln 10.$$

这与任何宏观量相比实际上也等于零.

(二)热力学第三定律的推论

1.绝对零度下物质的性质

由热力学第三定律可推得绝对零度下物质的一系列性质,其中之一是,当 $T \to 0$ K 时,系统的热容量趋于零.例如定体热容量 $C_V \to 0$.

根据热力学第三定律有

$$S(T,V) = S(0,V) + \int_0^T (C_V/T')\,dT' = \int_0^T (C_V/T')\,dT'.$$

当 $T \to 0$ K 时,等式左边为有限值.为了保证右边有限,C_V 至少应按 T 的一次幂趋于零.这一结论已为迄今所作的一切实验所证实.

另一性质是，当 $T \to 0$ K 时系统的等压膨胀系数 $\alpha \equiv (\partial V/\partial T)_p/V$ 趋于零.

利用麦克斯韦关系，可将 α 化为 $\alpha = (\partial V/\partial T)_p/V = -(\partial S/\partial p)_T/V$. 根据第三定律，当 $T \to 0$ K 时，一切系统的熵趋于零，而和其他参量（如压强）无关，即有 $\lim\limits_{T \to 0}(\partial S/\partial p)_T = 0$，于是得 $\lim\limits_{T \to 0}\alpha = 0$.

显然，以上推导也适用于其他热容量，如 C_p，C_H 和其他偏导数，如 $(\partial p/\partial T)_V$ 和 $(\partial m/\partial T)_H$.

2. 绝对零度不可能达到表述

1912 年能斯特又根据他的定律推出了一个原理：任何过程，经有限步骤达到绝对零度是不可能的. 往往有人把这一推论作为热力学第三定律的又一种表述. 因为这样就把热力学的第一、第二和第三定律都以否定的形式表述出来了.

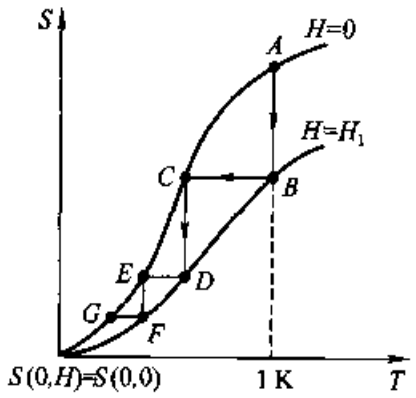


图 1　能斯特表述正确，绝对零度不能通过有限步骤达到

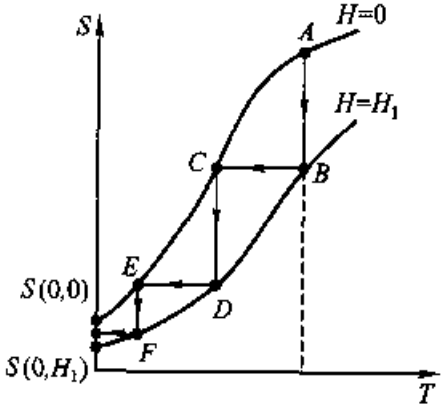


图 2　假设能斯特表述不正确，则绝对零度能通过有限步骤达到

这一表述不难以磁冷却的 T-S 图来证明. 图 1 是实验的典型结果，它可以这样去理解：第一，当外场不变升高温度时，由于会破坏离子磁矩的规则排列，S 会增大，所以各条等场曲线有正斜率，即 $(\partial S/\partial T)_H > 0$. 第二，温度不变，增大外场会促使离子磁矩规则排列，这意味着系统的混乱度减小，因而熵 S 会减小，即有 $(\partial S/\partial H)_T < 0$. 所以 $H = 0$ 的线必在 $H \neq 0$ 的线的左侧. 第三，当 $T \to 0$ K 时，所有等场曲线交于一点，这是热力学第三定律的能斯特表述的要求. 整个磁冷却过程描述如下：从 A 点开始，经等温磁化到达 B 点，然后绝热退磁到达 C 点，$T_C < T_A$. 由图看到，重复以上过程，介质的温度将不断降低，无限趋近于绝对零度，但通过有限步，不可能达到绝对零度. 这正是热力学第三定律“不可能达到表述”的具体体现.

作为对比，在图 2 上画出了假设能斯特表述不正确的情况. 由图可见，在此情况下，绝对零度将能通过有限步骤达到.

从图 1 还可以看出，要取得好的降温效果，要求在 1K 附近，各条等场线下降缓慢，即在低温下离子磁矩仍然无规则排列. 这正是选用离子磁矩相互作用微弱的

顺磁盐为磁冷却介质的原因.

最后要指出的是,能斯特表述不能应用于那些不处于平衡态的物质.例如有些无定形材料或者无序合金,它们在低温下能以很长的弛豫时间作为"冻结"的亚稳态存在.对于这些物质的亚稳态,能斯特表述不成立.

习　题

1.单原子晶体中,每个原子可占据一个格点或一个间隙点.原子占据格点时的能量比占据间隙点时高 ε.设格点数和间隙点数相等,且等于晶体中的原子数 N.

(a)考虑有 n 个原子占据间隙点的宏观态,计算系统处于此宏观态的熵.

(b)设系统达到平衡,问晶体在该态的温度是多少?

(c)若 $\varepsilon = 1\text{eV}$,晶体的温度是 300K,处于间隙点的原子所占比例是多少?

2.考虑橡皮带的简单模型:一个一维链条由 n 个长度为 a 的链环组成,链环沿着 x 轴,但可以重叠(如图 1).链条两个端点之间的距离是 x.系统是孤立的.链环各种方位具有相同能量.证明,当 $x \ll na$ 时,可得到胡克定律.

3.系统由 N 个磁矩为 μ_0、自旋为 1/2 的原子组成.因此,在外场中,每个原子的磁矩只能平行或反平行于外场排列.忽略磁原子间的相互作用.

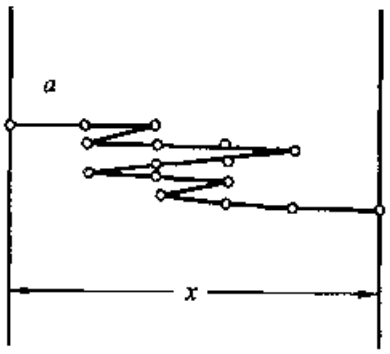


图 1　习题 2 图

(a)求平行于外场的磁原子数为 $(N/2 + l)$ 时,系统的微观态数 $g(N,l)$ 和磁能量.证明当 $1 \ll l \ll N$ 时,$g(N,l) \approx g(N,0)\mathrm{e}^{-2l^2/N}$,$g(N,0) = N!/[(N/2)!(N/2)!]$.

(b)证明系统的化学势为 $\mu(\mathscr{T},H) = \mu(0) - \mu_0^2 H^2/2\mathscr{T}$.其中 $\mathscr{T} = k_B T$,$\mu(0)$ 是磁能为零时的化学势.

(c)关于化学势贡献的来源,由此你可得出什么结论?

4.某种磁性材料,磁化强度 $\boldsymbol{m}$ 与磁场强度 $\boldsymbol{H}$ 的关系为 $\boldsymbol{m} = \chi \boldsymbol{H}$,其中 χ 为磁化率.在弱场及某一温度区域内 $\chi = c/T$,c 为常数.现通过以下两个过程使 $\boldsymbol{m}$ 增加到 $2\boldsymbol{m}$:(a)保持 $\boldsymbol{H}$ 不变,使温度从 T 变为 $T/2$.(b)等温地使 $\boldsymbol{H}$ 增加到 $2\boldsymbol{H}$.在上述两个过程中体积变化均可忽略.试在 $\boldsymbol{H}$-$\boldsymbol{m}$ 图上画出过程曲线,并求外界所作的功.

5.已知水的比热为 $4.18\ \text{J}\cdot\text{g}^{-1}\cdot\text{K}^{-1}$.(a)有 1kg 0℃的水与 100℃的热源接触,当水达到 100℃时,水的熵变了多少?热源的熵变了多少?水与热源的总熵变了多少?(b)若 0℃的水先与 50℃的热源接触达到平衡,然后与 100℃的热源接触达到平衡,则整个系统的熵变了多少?(c)若要使整个系统的熵不改变,水应如何从 0℃变至 100℃?

6.一根均匀杆的温度一端为 T_1,另一端为 T_2,计算在达到均匀温度 $(T_1 + T_2)/2$ 后熵的增加值.

7.(a)两个容器 A 和 B 各储有 N mol 相同的理想气体.开始两容器是绝热的,具有相同的压强,不同的温度,然后让两个容器接触.计算当平衡建立后整个系统的熵变,并证明熵变不可能为负.(b)仍然考虑以上两容器,但开始具有相同的温度和不同的压强.将两容器间的隔板取出.求当平衡建立后,整个系统的熵变,并证明熵变不可能为负.

8. 隔板把 A,B 两种气体分开，它们的物质的量分别为 N_1 mol 和 N_2 mol，压强和温度相同. 今将隔板抽出，两种气体在绝热条件下相互扩散. 求达到平衡后，系统熵的变化.

9. 有两个相同的物体，热容量为常数，初始温度为 T_1，今让一致冷机在此两物体之间工作，使其中一个物体的温度降低到 T_2 为止. 假设物体维持在定压下并且不发生相变. 证明此过程所需的最小功为 $W_{\min}=C_p[(T_1/T_2)^2+T_2-2T_1]$.

10. 有两个相同的物体，初温各为 T_1 和 T_2. 有一热机工作于两物体之间，使两者温度变成相等. 证明热机所能作的最大功为 $W_{\max}=C_p(T_1+T_2-2\sqrt{T_1T_2})$.

11. 证明以下热力学公式：

(a)
$$(\partial H/\partial p)_T=V-T(\partial V/\partial T)_p,$$
$$(\partial U/\partial p)_T=pV\kappa-\alpha VT,$$
$$(\partial U/\partial p)_V=C_V(\partial T/\partial p)_V,$$
$$(\partial U/\partial V)_p=C_p(\partial T/\partial V)_p-p.$$

(b)
$$(\partial S/\partial p)_H<0;\qquad (\partial S/\partial V)_U>0.$$

(c)
$$(\partial T/\partial V)_S=-(T/C_V)(\partial p/\partial T)_V;$$
$$(\partial S/\partial p)_V=(C_V/T)(\partial T/\partial p)_V;$$
$$\kappa_S/\kappa=C_V/C_p\qquad(\kappa_S\text{ 是绝热压缩系数}).$$

(d)
$$(\partial T/\partial p)_S-(\partial T/\partial p)_H=V/C_p,$$
$$\kappa_S-\kappa_T=-(T/V)(\partial V/\partial T)_p^2/C_p.$$

(e)
$$(\partial\alpha/\partial p)_T=-(\partial\kappa_T/\partial T)_p.$$

(f)
$$\alpha/\alpha_S=1-\gamma,\qquad \beta/\beta_S=1-1/\gamma.$$

其中　$\alpha_S=(1/V)(\partial V/\partial T)_S$，　$\beta_S=(1/p)(\partial p/\partial T)_S$，　$\gamma=C_p/C_V$.

(g)
$$C_V=C_V^0+T\int_{V_0}^{V}(\partial^2P/\partial T^2)_V\mathrm{d}V.$$

12. 对一服从范德瓦耳斯方程的气体. (a)求等压膨胀系数. (b)求内能对体积的依赖关系. (c)证明等体热容量仅仅是温度的函数. (d)求 C_p-C_V. (e)求焦尔-汤姆孙系数. (f)求由 $\mathrm{d}V,\mathrm{d}T$ 产生的熵变 $\mathrm{d}S$.

13. 设顺磁质服从居里定律：$m=cH/T$，证明：

(a) U 和 C_m 只是温度的函数；(b) $S=\int C_m\mathrm{d}T/T-cm^2/2+$ 常数；(c) $C_H-C_m=m^2/c$；

(d) $(\partial C_H/\partial H)_T=2cH/T^2$ (H 为磁场强度的大小).

14. 一弹簧在恒温下恢复力 f 与伸长 x 成正比，比例系数 k 是温度的函数. 试证明弹簧的自由能 F、熵 S 和内能 U 的表达式为：

$$F(T,x)=F(T,0)+kx^2/2,$$
$$S(T,x)=S(T,0)-(x^2/2)(\mathrm{d}k/\mathrm{d}T),$$
$$U(T,x)=U(T,0)+(k-T\mathrm{d}k/\mathrm{d}T)x^2/2.$$

15. 求范氏气体的特性函数 $F(T,V)$，并导出内能 $U(T,V)$ 和熵 $S(T,V)$.

16. 在绝对零度附近，一根被伸长的弹性棒的张力和它的长度的关系为 $f\approx aT^2(L-L_0)$，其

中 a 是常数，L_0 是未伸长时的长度. 当 $L = L_0$ 时的热容量 $C_L|_{L=L_0} = bT$，b 是常数. (a)写出此系统的热力学基本方程. (b)求 $(\partial S/\partial L)_T$. (c)已知 $S(T_0, L_0)$，求任意温度和任意长度时的 $S(T, L)$. [最方便的方法是先求 $S(T, L_0)$，因这时的热容量是已知的]. (d)从 $T = T_1, L = L_1$ 出发，将棒准静态绝热拉长到 L_f，最后的温度 T_f 是多少？(e)求长度为 L 时的热容量 C_L.

17. 沿磁场 H 放置长度为 l 的棒，并受到外面的拉力 f. 等温磁化将引起棒长的改变. 由实验可知，当拉力足够大，同时，场强足够弱时，磁矩可表为 m = 常数 · lH/f.

(a)写出热力学基本方程；

(b)定义 $G = U - TS - lf - m \cdot H$，证明 $(\partial l/\partial H)_{f,T} = (\partial m/\partial f)_{H,T}$；

(c)证明当磁场由零增至 H 时棒长的相对变化为 $\Delta l/l$ = 常数 · $H^2(\alpha f - 1)/2f^2$，其中 $\alpha = l^{-1}(\partial l/\partial f)_T$.

18. 对电介质建立热力学基本方程，并证明 $(\partial p_e/\partial p)_{T,E} = -(\partial V/\partial E)_{T,p_e}$. 其中 p_e 为电矩，p 为压强，E 为电场强度，V 为介质体积. 并说明等式的意义.

19. 对电介质证明以下公式：$(\partial T/\partial E)_S = -(T/C_E)(\partial p_e/\partial T)_E$，$C_E$ 为恒电场下的热容量，S 为熵，其他符号同上题.

20. 证明下列平衡判据. (a)在 S, V 不变的情形下平衡态的 U 最小. (b)在 S, p 不变的情形下，平衡态的 H 最小.

21. 试用自由能判据求两相平衡时的力学平衡条件与相平衡条件(两相分界面为平面).

22. 在三相点附近，固、气二相的平衡曲线在 p-T 图上的斜率比液、气两相平衡曲线的斜率陡，试从物理上说明.

23. (a)用自由能判据，而不是自由焓判据证明麦克斯韦等面积法则. (b)用热力学第二定律证明等面积法则.

24. §7.2.9 图 4 范氏等温线 AJ 段上的任一点 Q，和实验等温线 AO 上体积相同的对应点 p 相比，Q 为亚稳态. 试用自由能来说明.

25. 在 p-T 图上范氏气体等温线上的极大点与极小点连成一条曲线 ACB，证明这条曲线的方程为 $pv^3 = a(v - 2b)$，并说明这条曲线分割出的区域 I, II, III 的意义.

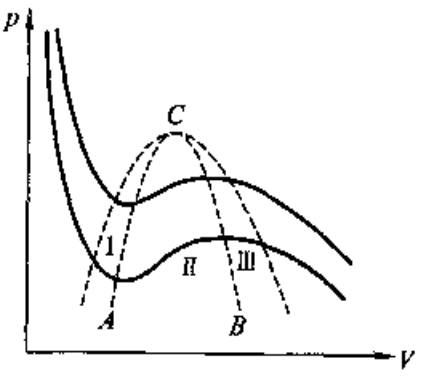


图 2　习题 25 图

26. 设一物质的气相可看作理想气体，液相比容与气相比容相较可以忽略. (a)证明蒸气的"两相平衡膨胀系数"为 $V^{-1}(\mathrm{d}V/\mathrm{d}T) = T^{-1}(1 - L/RT)$，其中 L 为汽化热；(b)当温度较低时，$(\mathrm{d}V/\mathrm{d}T) < 0$，并说明意义.

27. (a)证明绝热曲线的斜率为 $(\partial p/\partial T)_S = C_p/(T\alpha V)$；(b)饱和乙醚蒸气在 35℃ 时绝热膨胀，试判断乙醚是否会凝聚. 已知乙醚蒸气的 $c_p = 1.89\ \mathrm{J \cdot g^{-1} \cdot K^{-1}}$，汽化热 $L = 353\ \mathrm{J \cdot g^{-1}}$. 设饱和乙醚蒸汽近似为理想气体，即 $\alpha = T^{-1}$ 可以应用，及液相体积和气相体积相比可以忽略.

28. 灰锡和白锡在 1atm 下的相变温度是 291K，低于此温度灰锡是稳定的，在此转变中焓的变化是 535cal · $\mathrm{mol^{-1}}$，灰锡和白锡的密度分别是 5.75 和 7.30g · $\mathrm{cm^{-3}}$，相对原子质量是 118.7.

如果系统是在100atm,相变温度改变多少?

29. 一系统的状态方程是 $[p+n^2a/(TV^2)](V-nb)=nRT$. a,b 是实验常数. 用实验常数 a 和 b 表示出系统的临界温度,临界压强和临界体积.

30. 考虑一在零磁场和标准大气压下的金属. 金属在正常态和超导态下的比热分别为 $c_n=\gamma T, c_s=\alpha T^3$, α,γ 是常数. (a)求系数 α,γ 和临界温度 T_c 之间的关系. (b)求在0K时金属在正常态和超导态时内能之差,将结果用 γ 和 T_c 表示.

31. 证明厄仑菲斯特公式

$$\frac{dp}{dT}=\frac{c_p^{(2)}-c_p^{(1)}}{Tv(\alpha^{(2)}-\alpha^{(1)})},\qquad \frac{dp}{dT}=\frac{\alpha^{(2)}-\alpha^{(1)}}{\kappa^{(2)}-\kappa^{(1)}}.$$

32. (a)已知焦尔-汤姆孙系数为

$$\mu=(1/C_p)[T(\partial V/\partial T)_p-V], \tag{1}$$

用 math 解联立方程证明,对于范氏气体,其焦尔-汤姆孙转换温度曲线方程为

$$p=(2/b)\sqrt{2aRT/b}-(3/2)(RT/b)-a/b^2. \tag{2}$$

(b)已知某些范氏气体的常数 a,b 值如下:

气　体	a	b	T_{max}(实验)K
CO_2	3.6	0.043	1500
H_2	0.1975	0.026	202
N_2	1.35	0.039	623

a 的单位是 $atm\cdot l^2\cdot mol^{-2}$, b 的单位是 $l\cdot mol^{-2}$

试用 math 绘出三种气体的转换温度曲线,由所绘曲线确定最大转换温度的近似值,并和实验值进行比较. 绘图时,温度取值能任意低吗? 为什么?

33. 下表给出了范氏气体的 a,b 常数及实验测得的临界温度和临界系数:

气　体	a	b	t℃	RT_c/p_cv_c
CO_2	3.6	0.043	31.10	3.48
H_2	0.1975	0.026	-239.95	3.03
N_2	1.35	0.039	-147.16	3.42

a 的单位 $atm\cdot l^2\cdot mol^{-2}$, b 的单位是 $l\cdot mol^{-2}$

(a)试用 math 就三种气体绘出一系列温度下的 p-v 曲线,并由所绘曲线估算出三种气体的临界温度 T_c 和临界系数 $RT_c/(p_cv_c)$.

(b)绘 p-v 图时, v 的取值可以从零开始吗? 为什么? T 可以任意取值吗? 为什么?

(由 Berchelot 方程 $p=RT/(v-b)-a/(Tv^2)$ 或 Dieterica 方程 $p=[RT/(v-b)]e^{-a/RTv}$ 可给出更好的临界系数,但在描述其他性质方面不如范氏方程.)

第三章　统计系综

通过第一章的学习知道，确定系综的概率是系综理论的关键. 系统所处的宏观条件不同，相应系综的概率就不同. 微正则系综的概率为 $P_\alpha = 1/\Omega(E,V,N)$，求得 $\Omega(E,V,N)$ 以后，利用 $A_{微} = \sum_\alpha P_\alpha A_\alpha$，可以计算出有对应微观量的宏观量；另一方面，像 §7.2.3 例 1 和例 2 处理经典理想气体和理想自旋系统采用的办法那样，利用 $S = k_B \ln \Omega, 1/T = k_B(\partial \ln \Omega/\partial E), p = k_B T(\partial \ln \Omega/\partial x)$ 和 $\mu = -k_B T(\partial \ln \Omega/\partial N)$ 也可以求得理想气体及理想自旋系统的热力学量. 实际上，后一种方法是对任何系统都有效的一般方法，通常称之为微正则系综的统计热力学公式. 但是，由于计算 $\Omega(E,N,V)$ 的实际困难，微正则分布往往不方便直接利用. 而由微正则分布可推得方便运用的其他分布，如正则分布和巨正则分布. 本章将依次介绍如何由微正则分布推导这两种分布，以及由它们求热力学量的方法. 最后讨论三种系综的等价性. 作为正则分布应用的实例，在 §7.3.2 中研究顺磁固体模型和固体比热的爱因斯坦模型. 作为巨正则分布应用的实例，在 §7.3.4 中推导理想气体的三种分布.

§7.3.1　正则系综和正则分布

§7.1.6 中介绍了由孤立系组成的微正则系综和微正则分布. 实际中遇到的大多数系统一般不满足孤立条件，而且能量的测量和严格控制也是不可能的. 相反，温度 T 不仅可以测量也便于控制. 因此，本节研究可以用 (T,N,x) 描述的恒温系统组成的正则系综，x 代表外参量，如 V 或 H.

(一)正则系综和正则分布

设系统 A 在外参量 x 和粒子数 N 不变的情况下和大热源 A' 接触而达到平衡. 由于系统和大热源之间存在热接触，两者可以自由地交换能量，因而系统的能量是可变的. 但是热源很大，交换能量不会改变热源的温度. 两者达到平衡后，系统将和热源具有相同的不变温度，因此把这样的系统叫做恒温系统. 大量相同的恒温系统的集合叫正则系综. 孤立系统的平衡态由约束条件 (E,N,x) 描述，恒温系统的平衡态由 (T,N,x) 描述.

可以用图 1 来表示正则系综. 图中用 ν 标志的微观态是满足定态方程 $\hat{H}|\psi_\nu\rangle = E_\nu|\psi_\nu\rangle$ 的能量本征态.

根据量子刘维尔定理，在能量表象中任何系综的密度矩阵都是对角矩阵，且对角元是能量本征值的函数，见 §7.1.5(5.20)式，即有：

$$\rho_{ml} = \delta_{ml}\rho_l, \qquad \rho_l = P_l(E_l) \qquad \text{(能量表象)}.$$

我们来求正则系综在能量表象中的密度矩阵.具体地说,我们要问,系统处于能量为 E_l 的某一微观态 l(能量本征态)的概率 P_l 是多少,或 P_l 和 E_l 有怎样的函数关系?为了能利用微正则分布,我们考查由系统 A 和大热源 A' 组成的复合系统 $A^{(0)}$,并假定 $A^{(0)}$ 是孤立的.

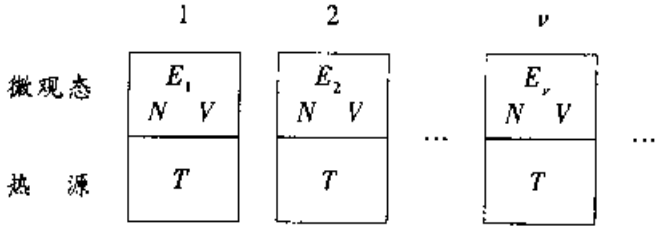


图 1　恒温系统组成的正则系综

假设系统和热源的相互作用微弱,复合系统能量 $E^{(0)}$ 可表为系统能量 E 和热源能量 E' 之和,即

$$E^{(0)} = E + E'. \tag{1.1}$$

既然热源很大,必有 $E \ll E^{(0)}$(严格说来,这个不等式不能对系统可能出现的所有状态都适用,但对于接近平衡态的那些状态是适用的).

当系统处在能量为 E_l 的某一微观态 l 时,热源可以处在能量为 $(E^{(0)} - E_l)$ 的任何一个微观态.以 $\Omega'(E^{(0)} - E_l)$ 表示热源的能量为 $(E^{(0)} - E_l)$ 时的微观态数,则它也就是当系统处在微观态 l 时,复合系统的可能的微观态数.复合系统是孤立系统,根据微正则分布,在平衡态下,它的每一个可能的微观态出现的概率相等,因此,系统处在确定的微观态 l 的概率正比于此时复合系统的微观态数,即有

$$P_l \propto \Omega'(E^{(0)} - E_l). \tag{1.2}$$

考虑熵的玻尔兹曼公式后有 $\Omega'(E^{(0)} - E_l) = \mathrm{e}^{S'(E^{(0)} - E_l)/k_B}$,因而

$$P_l \propto \mathrm{e}^{S'(E^{(0)} - E_l)/k_B}. \tag{1.3}$$

将 $S'(E^{(0)} - E_l)/k_B$ 按 E_l 的幂次展开

$$\frac{1}{k_B}S'(E^{(0)} - E_l) = \frac{1}{k_B}S'(E^{(0)}) - \frac{1}{k_B}\frac{\partial S'(E')}{\partial E'}\bigg|_{E' = E_0} E_l + \frac{1}{2}\frac{1}{k_B}\frac{\partial^2 S'(E')}{\partial E'^2}\bigg|_{E' = E_0} E_l^2 + \cdots. \tag{1.4}$$

根据 §7.2.3(3.2)式 $(\partial S'(E')/\partial E')|_{E' = E_0} = 1/T$,$T$ 是热源的温度(因为系统与热源达到平衡,T 也就是系统的温度),而展开式右方第三项则包含了热源温度的变化.根据热源的定义,这是可以忽略的,于是得到

$$(1/k_B)S'(E^{(0)} - E_l) \approx (1/k_B)S'(E^{(0)}) - \beta E_l, \tag{1.5}$$

这里 $\beta = 1/(k_B T)$,$S'(E^{(0)})$ 是一个与系统无关的常数,(1.3)式化为

$$P_l = Ce^{-\beta E_l}.$$

代入归一化条件 $\sum_l P_l = C\sum_l e^{-\beta E_l} = 1$，可定出常数 $C = 1\Big/\sum_l e^{-\beta E_l}$. 令

$$Z = \sum_l e^{-\beta E_l}, \tag{1.6}$$

最后得到系统处于微观态 l 的概率：

$$P_l = (1/Z)e^{-\beta E_l}. \tag{1.7}$$

(1.6)式右边是对系统的一切微观态求和. 考虑到有些微观态具有相同的能量，用 $\Omega(E_l)$ 表示能量为 E_l 的微观态数(即 E_l 的简并度)，则(1.6)式又可写为

$$Z = \sum_{E_l} e^{-\beta E_l}\Omega(E_l). \tag{1.8}$$

(1.8)式是对系统的一切能级求和. 恒温系统处于能量 E_l 的概率为：

$$P(E_l) = (1/Z)e^{-\beta E_l}\Omega(E_l). \tag{1.9}$$

(1.7)式和(1.9)式都叫做正则分布. (1.7)式表示恒温系统处于能量为 E_l 的某一微观态的概率；而(1.9)式则表示恒温系统处于能量为 E_l 的任一微观态的概率，或具有能量 E_l 的概率. 正则系综在能量表象中的密度矩阵为

$$\rho_{ml} = \rho_l\delta_{ml}, \qquad \rho_l = e^{-\beta E_l}/Z. \tag{1.10}$$

由(1.6)式或(1.8)式定义的 Z 叫做正则系综的配分函数. 由(1.7)式可得系统处于微观态 l_1 和微观态 l_2 的概率之比为

$$P_{l_1}/P_{l_2} = e^{-\beta E_{l_1}}/e^{-\beta E_{l_2}},$$

由此可知，$e^{-\beta E_l}$ 是系统处于微观态 l 的相对概率的量度，文献上往往称为玻尔兹曼因子，而配分函数则是所有微观态的玻尔兹曼因子之和. 因为能量 E_l 不仅与系统中的粒子数，还与外参量 x 有关[例如在体积 V 的容器中运动的自由粒子的能量为 $\varepsilon = h^2(n_x^2 + n_y^2 + n_z^2)/8mV^{2/3}$，磁矩在外场中的能量为 $\varepsilon = -\mu_0 H$]，所以 $Z = Z(T,N,x)$，即配分函数是以 (T,N,x) 为自变量的态函数. 从 §7.2.5 已经知道，自由能 F 是以 (T,N,x) 为自变量的特性函数. 可以预料，配分函数 Z 和自由能 F 有关. 下面将推导它们的关系，并将看到配分函数 Z 在计算系统热力学量时的重要作用.

由于以上的讨论完全未涉及系统的具体性质，所以正则分布适用于满足 (T,N,x) 一定的任何系统. 特别是，不论系统中粒子之间的相互作用是否能够忽略，正则系综都适用. 换句话说，和只适用于近独立子系的最概然统计法比较，系综理论是一个普遍的统计理论.

(二)配分函数和热力学量的关系

对于微正则系综，知道了 $\Omega(E,N,x)$，可以求得以 (E,N,x) 为独立变量的特性函数——熵 $S = k_B\ln\Omega(E,N,x)$，由此可反解出内能 U；根据 $dS = dU/T -$

$Ydx/T - \mu dN/T$，利用 $1/T = (\partial S/\partial U)_{N,x}$；$Y = -T(\partial S/\partial x)_{N,U}$；$\mu = -T(\partial S/\partial N)_{U,x}$ 可计算出系统的温度、状态方程、化学势，进而求得全部热力学量. 与此类似，对于正则系综，知道了 $Z(T,N,x)$，就可以求得以 (T,N,x) 为独立变量的特性函数——自由能 $F(T,N,x)$，利用 $dF = -SdT + Ydx + \mu dN$ 进而求得全部热力学量. 下面进行证明.

前面已经说明，恒温系统是与大热源可自由交换能量而处于热平衡的系统. 因此，恒温系统的内能 U 是能量 E_l 在给定 (T,N,x) 条件下的一切可达微观态上的系综平均，即

$$U = \langle E_l \rangle = \sum_l P_l E_l. \tag{1.11}$$

将正则分布代入得，

$$U = \frac{1}{Z}\sum_l E_l e^{-\beta E_l} = -\frac{1}{Z}\frac{\partial}{\partial\beta}\sum_l e^{-\beta E_l} = -\frac{\partial}{\partial\beta}\ln Z. \tag{1.12}$$

系统哈密顿随外参量的变化定义为外界对系统的广义力 $Y = \partial H(x)/\partial x$. 它在微观态 l 的值是

$$Y_l = \partial E_l(x)/\partial x. \tag{1.13}$$

于是，宏观广义力为

$$Y = \sum_l \frac{\partial E_l(x)}{\partial x}P_l = \frac{1}{Z}\sum_l \frac{\partial E_l}{\partial x}e^{-\beta E_l} = -\frac{1}{\beta Z}\frac{\partial}{\partial x}\sum_l e^{-\beta E_l} = -\frac{1}{\beta}\frac{\partial \ln Z}{\partial x}. \tag{1.14}$$

上式的一个重要情形是压强：

$$p = (1/\beta)(\partial \ln Z/\partial V). \tag{1.15}$$

考虑到 $Z = Z(\beta,N,x)$，对粒子数不变的封闭系统可以写出：

$$\begin{aligned} d\ln Z &= (\partial \ln Z/\partial x)dx + (\partial \ln Z/\partial\beta)d\beta = -\beta Ydx - Ud\beta \\ &= -\beta Ydx - d(\beta U) + \beta dU. \end{aligned}$$

或 $d(\ln Z - \beta\partial \ln Z/\partial\beta) = (dU - Ydx)/(k_B T)$. 与热力学中的 $dS = (dU - Ydx)/T$ 比较，得到熵的表达式

$$S = k_B(\ln Z - \beta\partial \ln Z/\partial\beta). \tag{1.16}$$

(1.16)式可改写为 $S = k_B(\ln Z + \beta U)$，或

$$F = U - TS = -k_B T\ln Z. \tag{1.17}$$

可见 $\ln Z$ 与自由能有简单的关系. 利用自由能的统计表达式(1.17)，则(1.14)、(1.16)和(1.12)式可用更简单的形式表示：

$$Y = \partial F/\partial x, \quad S = -\partial F/\partial T, \quad U = F - T\partial F/\partial T. \tag{1.18}$$

这实际上就是热力学中特性函数自由能的有关公式[见 §7.2.5(三)]. 因此，对于由 (T,N,x) 描述的平衡态恒温系统，只要求得配分函数，就可以得到以 (T,N,x)

为独立变量的特性函数自由能，从而用求偏导数的简单运算求得系统的全部热力学量. 因此(1.6)或(1.8)和(1.17)式是正则系综理论的基本公式.

利用 $U = \sum_l P_l E_l$ 和 $P_l = \mathrm{e}^{-\beta E_l}/Z$ 还可将(1.16)式进行变形：

$$S = k_B\left[\ln Z + \beta\sum_l E_l P_l\right] = k_B\left[\ln Z - \sum_l P_l \ln\left(ZP_l\right)\right]$$

$$= k_B\left[\ln Z - \ln Z\sum_l P_l - \sum_l P_l \ln P_l\right] = -k_B\sum_l P_l \ln P_l. \qquad (1.19)$$

或

$$S = -k_B\langle \ln P_l\rangle. \qquad (1.20)$$

可以证明，(1.19)式或(1.20)式不仅适用于正则系综，而且也适用于微正则和巨正则系综，所以是熵的一般表达式. 它表明系统的熵单值地由概率 P_l 决定. 事实上，式(1.19)或(1.20)是信息熵的表达式，因而具有更广泛的应用.

最后，我们要讨论配分函数的一个重要性质. 在解决许多实际问题时，这一性质会给我们带来很大方便. 考虑由子系统 A 和 B 组成的复合恒温系统 $(A+B)$. 用 E_A 和 E_B 分别表示两个子系统的能量，用 $\Omega(E_A)$ 和 $\Omega(E_B)$ 分别表示相应的微观态数. 若 A 和 B 相互作用十分微弱，以致有

$$E_{A+B} = E_A + E_B, \qquad (1.21)$$

$$\Omega(E_{A+B}) = \Omega(E_A)\cdot\Omega(E_B), \qquad (1.22)$$

则根据(1.9)式，复合恒温系统的配分函数可写为

$$Z = \sum_{E_{A+B}} \mathrm{e}^{-\beta E_{A+B}}\Omega(E_{A+B}) = \sum_{E_A} \mathrm{e}^{-\beta E_A}\Omega(E_A)\cdot\sum_{E_B} \mathrm{e}^{-\beta E_B}\Omega(E_B) = Z_A\cdot Z_B, \qquad (1.23)$$

即：复合恒温系统的配分函数可分解为两个近独立可识别子系统的配分函数之积. 将(1.23)式代入(1.12)，(1.16)和(1.17)式可以得到：

$$U_{A+B} = U_A + U_B, \qquad S_{A+B} = S_A + S_B, \qquad F_{A+B} = F_A + F_B.$$

即：复合系统的内能、熵和自由能等于各个子系统的贡献之和. 显然以上结果可推广到多于两个子系统的情况. 于是，只要分别求得各子系统的配分函数，就很容易计算整个复合系统的热力学性质.

要特别注意的是，系统配分函数可以分解的条件仅仅为：子系统的能量和微观态数满足(1.21)和(1.22)式. 因此，这样的子系统可以是相互作用十分微弱的可识别的不同粒子群或单个粒子；也可以是耦合十分微弱的不同运动形式，而它们属于同一群粒子或同一粒子. 例如，在固体中，有原子在格点附近的振动，原子磁矩在外场中的取向，各种热缺陷的运动，以及自由电子的运动等. 只要它们耦合微弱，就可以分别计算它们的配分函数，研究它们各自对系统热力学性质的贡献.

根据量子力学,(1.7)式可改写为

$$Z = \sum_l e^{-\beta E_l} = \sum_l \langle l | e^{-\beta \hat{H}} | l \rangle = \mathrm{tr}(e^{-\beta \hat{H}}).$$

因为求迹和表象无关,只要知道了系统的哈密顿算符,可以在任何表象中求得配分函数.不过,我们只在能量表象中进行.

(三)恒温系统的能量涨落

由于恒温系统可以与大热源自由交换能量,因此其能量并不总是等于平均值,而是在平均能量上下变化.根据定义,系统能量与平均能量偏差的平方的平均值叫能量的均方涨落.我们用正则分布来进行计算:

$$\begin{aligned}\langle (\Delta E)^2 \rangle &= \langle (E - \langle E \rangle)^2 \rangle = \langle E^2 \rangle - \langle E \rangle^2 \\ &= \frac{1}{Z} \sum_{E_l} E_l^2 e^{-\beta E_l} \Omega(E_l) - \left(\frac{1}{Z} \sum_{E_l} E_l e^{-\beta E_l} \Omega(E_l) \right)^2 \\ &= -\frac{\partial}{\partial \beta} \left(\sum_{E_l} E_l e^{-\beta E_l} \Omega(E_l)/Z \right) = -\frac{\partial U}{\partial \beta} \\ &= k_B T^2 (\partial U/\partial T) = k_B T^2 C_V. \end{aligned} \tag{1.24}$$

能量的相对涨落为

$$\delta E = (\langle (\Delta E)^2 \rangle / \langle E \rangle^2)^{1/2} = (k_B T^2 C_V / \langle E \rangle^2)^{1/2}. \tag{1.25}$$

因为 T 是强度量,与粒子数 N 无关,而 C_V 和 $\langle E \rangle$ 都是广延量,与粒子数 N 成正比,所以有

$$\delta E \propto N^{-1/2}. \tag{1.26}$$

对于宏观系统,$N \sim 10^{23}$,能量的相对涨落完全可以忽略.恒温系统能量的相对涨落可以忽略这一事实告诉我们,恒温系统和孤立系统实际上是等价的.用微正则系综和正则系综研究同一系统,会得到相同的结果.本章第五节还将对此作进一步讨论.

§7.3.2 固体的顺磁性模型和固体比热的爱因斯坦模型

作为正则分布应用的例子,我们研究两个简单但有重要实际意义的模型:固体的顺磁性模型和固体比热的爱因斯坦模型.

(一)固体的顺磁性模型

在§7.2.9中已指出,磁化率 $\chi > 0$ 且 $\chi = m/H = c/T$ 的磁性材料叫顺磁质.居里定律 $m = cH/T$ 就是顺磁质的状态方程.组成顺磁质的原子或分子都有异于零的固有磁矩.为推导顺磁质的状态方程和研究顺磁性的微观本质,在§7.2.3的例2中已经用微正则系综考查过理想顺磁模型,这里再用正则系综研究同一模型.

【例 1】　顺磁固体由 N 个自旋为 1/2，磁矩为 μ_0 的磁偶极子组成，每个磁偶极子位于晶格点阵的格点上. 顺磁固体置于外场中，磁偶极子的相互作用可以忽略. (1)求顺磁固体的磁化强度 $m = \langle M\rangle/V$（$\langle M\rangle$ 为平均总磁矩），和磁化率 χ；(2)求顺磁固体的内能和热容量；(3)求顺磁固体的熵.

解　根据配分函数的可分解性，只要固体内部运动的其他方面（如晶格振动）和磁性质的相互作用十分微弱，就可以分别讨论它们对固体热力学量的贡献. 还是由于配分函数的可分解性，考虑到每个磁偶极子位于格点上，因而可以识别，当忽略磁偶极子间的相互作用时，又可以先计算单个偶极子的配分函数和它对热力学量的贡献，进而研究整个样品的热力学性质.

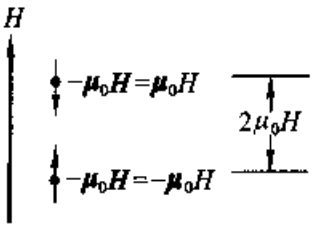


图 1　自旋为 1/2 的磁矩在外场中的取向和能级

(1)磁化强度 m 和磁化率 χ

如图 1 所示，自旋为 1/2 的磁偶极于在外场 H 中有两种可能状态，即平行或反平行于外场，相应的能量为 $\varepsilon = \mp \mu_0 H$. 因而单个偶极于的配分函数为

$$Z_1 = \sum_i \mathrm{e}^{-\beta E_i} = \mathrm{e}^{\beta\mu_0 H} + \mathrm{e}^{-\beta\mu_0 H} = \mathrm{e}^{x} + \mathrm{e}^{-x} = 2\mathrm{ch}\, x, \tag{2.1}$$

其中 $x = \mu_0 H/(k_B T)$. 样品的配分函数为

$$Z = Z_1^N = (2\mathrm{ch}\, x)^N. \tag{2.2}$$

在用配分函数计算介质的热力学性质之前，先作以下说明：

从(2.1)式可知，在计算 Z 或 Z_1 时，我们已将介质在外场中的势能作为系统的内能，而在 §7.2.3 的(3.35)式中我们曾指出，当把磁介质在外场中的势能作为内能的一部分时，介质的热力学基本方程应从 $\mathrm{d}U = T\mathrm{d}S + \boldsymbol{H}\cdot\mathrm{d}\boldsymbol{M}$ 化为

$$\mathrm{d}U^* = \mathrm{d}(U - \boldsymbol{H}\cdot\boldsymbol{M}) = T\mathrm{d}S - \boldsymbol{M}\cdot\mathrm{d}\boldsymbol{H}, \tag{2.3}$$

相应地有

$$\mathrm{d}F^* = \mathrm{d}(U^* - TS) = -S\mathrm{d}T - \boldsymbol{M}\cdot\mathrm{d}\boldsymbol{H}. \tag{2.4}$$

于是和 §7.3.1 中的(1.17)、(1.18)式对应有

$$F^* = U^* - TS = -k_B T\ln Z, \tag{2.5}$$

$$\left.\begin{aligned} \langle M\rangle &= -\partial F^*/\partial H = k_B T(\partial \ln Z/\partial H),\\ S &= -\partial F^*/\partial T = k_B[\ln Z - \beta\partial \ln Z/\partial\beta],\\ U^* &= F^* - T(\partial F^*/\partial T) = -\partial \ln Z/\partial\beta. \end{aligned}\right\} \tag{2.6}$$

于是，样品的平均总磁矩为

$$\langle M\rangle = k_B T(\partial \ln Z_1^N/\partial H) = N\mu_0 \mathrm{th}\, x.$$

磁化强度为

$$m = \langle M\rangle/V = N\mu_0 \text{th}\, x/V. \tag{2.7}$$

高温弱场下($x = \mu_0 H/(k_B T) \ll 1$), $\text{th}\, x \approx x$

$$m \approx N\mu_0 x/V = N\mu_0^2 H/(Vk_B T), \tag{2.8}$$

磁化率为

$$\chi \equiv m/H = N\mu_0^2/(Vk_B T) \propto c/T. \tag{2.9}$$

低温强场情况下($x \gg 1$), $\text{th}\, x \approx 1$

$$m = N\mu_0/V. \tag{2.10}$$

这表明,在低温强场情况下,所有磁偶极子全部沿外场方向排列,顺磁固体磁化达到饱和. $\langle M\rangle/(N\mu_0)$ 对 $\mu_0 H/(k_B T)$ 的依赖关系如图 2 所示.

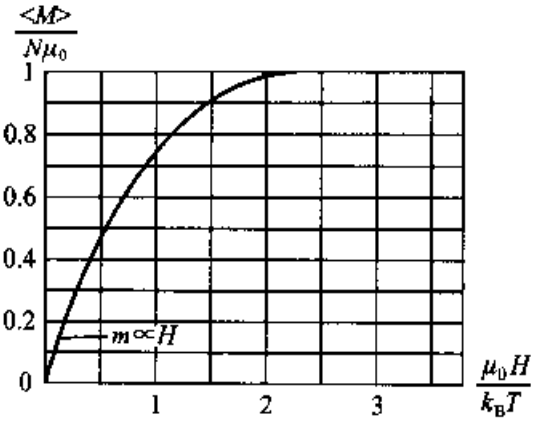


图 2　磁化强度对磁场和温度的依赖关系

理论和实验非常好的一致表明,以上简单模型是合理的. 原因是实际中使用的顺磁盐具有很大的分子,以致携带磁矩的离子被远远隔开,相互作用十分微弱.

(2.7)式也可以直接用求系综平均的方法得到,根据系综理论的第一个基本假设 $A_{测} = \sum_l A_l P_l$ [§7.1.4(4.4)式],磁偶极子平行和反平行于外场的概率分别为

$$P_l = \text{e}^{\pm\mu_0 H/(k_B T)}/Z_1 = \text{e}^{\pm x}/Z_1.$$

磁偶极子沿外场方向的平均磁矩为

$$\langle\mu\rangle = \sum_l P_l\mu_l = \mu_0 P_+ - \mu_0 P_-$$

$$= \mu_0 \frac{\text{e}^x - \text{e}^{-x}}{\text{e}^x + \text{e}^{-x}} = \mu_0 \text{th}\, x.$$

样品的平均总磁矩为

$$\langle M\rangle = N\langle\mu\rangle = N\mu_0 \text{th}\, x.$$

(2)内能 U 和热容量 C_H

将 $Z = Z_1^N = (2\text{ch}\, x)^N$ 代入(2.6)式,则和磁性有关的内能为

$$U^* = -\partial \ln (Z_1^N)/\partial\beta = -N\mu_0 H \text{th}\, x = -\langle M\rangle \cdot H. \tag{2.11}$$

可见 U^* 刚好是顺磁样品在外场中的势能. 外场不变下的热容量为

$$C_H = (\partial U^*/\partial T)_H = Nk_B x^2 \text{sech}\, x. \tag{2.12}$$

根据(2.12)式可绘出 $C_H/(Nk_B)$ 对 $1/x = k_B T/(\mu_0 H)$ 的依赖关系,如图 3. 从图 3

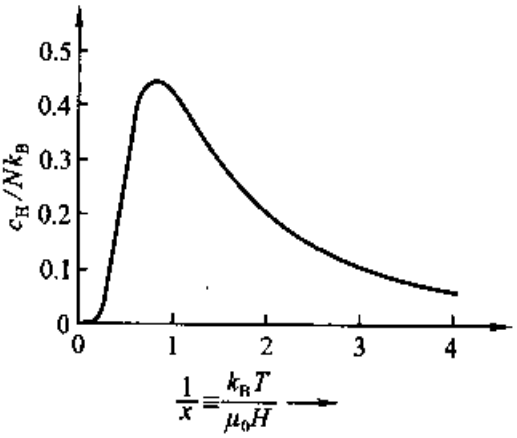


图 3　C_H 对 T 的依赖关系

可以看出，当 $T\to 0$ 时，$C_H\to 0$，和热力学第三定律的推论一致. 其次，结合 $C_H = T(\partial S/\partial T)_H$ 可以看出，系统从完全有序变到无序往往发生在一个很窄的温度范围内，它相应于曲线隆起的部分.

当然，(2.11)式也可由直接求系综平均来得到：

$$\begin{aligned}\langle\varepsilon\rangle &= \sum_i P_i\varepsilon_i = \varepsilon_+ P_+ + \varepsilon_- P_-\\ &= -\mu_0 H\cdot e^{\mu_0 H/(k_B T)}/Z_1 + \mu_0 H\cdot e^{-\mu_0 H/(k_B T)}/Z_1\\ &= -\mu_0 H(e^x - e^{-x})/(e^x + e^{-x}) = -\mu_0 H\,\mathrm{th}\,x.\end{aligned}$$

$$U^* = N\langle\varepsilon\rangle = -N\mu_0 H\,\mathrm{th}\,x.$$

(3) 熵 S

将 $Z = Z_1^N = (2\mathrm{ch}\,x)^N$ 代入(2.6)式，得到

$$S = Nk_B(\ln Z_1 - \beta\partial\ln Z_1/\partial\beta) = Nk_B[\ln 2 + \ln(\mathrm{ch}\,x) - x\,\mathrm{th}\,x].$$

高温弱场 ($x\ll 1$) 情况下，由 $\mathrm{th}x\approx x$，$\ln(\mathrm{ch}\,x)\approx\ln(1+x^2/2)\approx x^2/2$，得到

$$S = Nk_B\ln 2 = k_B\ln(2^N).$$

N 个磁偶极子，每个都有 2 个可能的状态，整个顺磁固体有 2^N 个不同微观态. 与上节的(1.20)式比较，上式意味着，在高温弱场下，所有微观态都是等概率的，即系统处于完全无序的状态.

低温强场 ($x\gg 1$) 下，由 $\mathrm{th}\,x\approx 1$，$\mathrm{ch}\,x\approx e^x/2$，则 $S\approx 0$，说明这时系统只有一个微观态，所有磁偶极子都平行于外场排列.

从这个例子中可以看出，按照系综理论求系统的宏观量，原则上有两种方法. 方法一：直接求相应微观量的系综平均. 这种方法的特点是，对每一个宏观量必须用求和或求积分的方法求其对应微观量的系综平均，然后再利用热力学有关公式求没有对应微观量的宏观量，例如熵. 因而这种方法往往不够方便. 方法二，统计热力学的方法，它可以概括为以下三个步骤：①针对研究的问题中系统可测的宏观性质，如状态方程和热容量，提出一个微观模型，计算该模型的能谱和简并度. 这是一个纯粹的力学问题. 但实际上往往不是解力学问题求得能谱，而是直接提出关于能谱的假设. ②根据模型的能谱和简并度求配分函数 Z 和特性函数 F. 而对由可识别、近独立子系统组成的系统，配分函数分解的原则将带来很大方便. ③利用特性函数的性质通过求配分函数的各种偏微商得到系统的热力学量. 这后一种办法是我们要反复练习的方法.

(二) 固体比热的爱因斯坦模型

关于固体比热，成功的理论必须能够解释以下两个基本实验事实.

1. 接近室温时，大多数固体的 $c_v\approx 3R\approx 6\ \mathrm{cal\cdot mol^{-1}K^{-1}}$，这称为杜隆-珀蒂定

律. 只有少数例外,例如金刚石的 $c_V = 1.46\ \text{cal}\cdot\text{mol}^{-1}\text{K}^{-1}$.

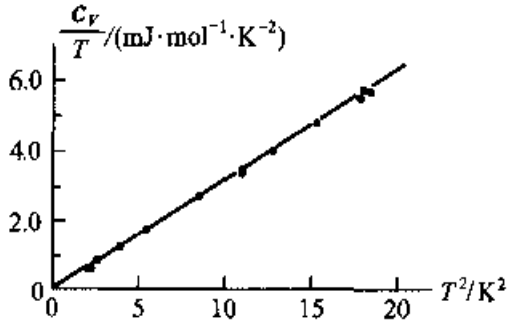


图 4　KCl 的热容量

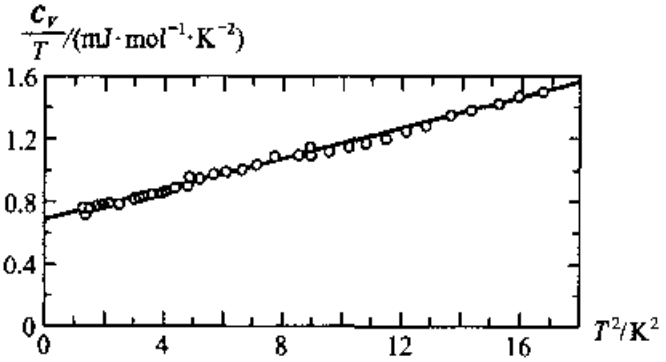


图 5　Cu 的热容量

2. c_V 与 T 有关,在低温下,c_V 随 T 减小,直到变为零,与热力学第三定律符合. 但分两种情况:对于非金属晶体,$c_V = cT^3$ 或 $c_V/T = cT^2$,如图 4 所示;对于处于正常导体状态的金属,$c_V = cT^3 + \gamma T$ 或 $c_V/T = cT^3 + \gamma$,如图 5 所示. 可以证明,其中 T^3 项来源于晶格振动,而 T 项来源于金属中的自由电子. 本节仅研究晶格振动对比热的贡献. 第八篇将研究正常导体中自由电子对比热的贡献.

1907 年爱因斯坦提出了固体比热的爱因斯坦模型,这一模型能得到杜隆-珀蒂定律并能定性说明晶体的 C_V 随着温度降低而减小的性质,但定量结果与实验相差甚远. 1912 年德拜改进爱因斯坦模型,得到了较好的结果. 下面介绍固体比热的爱因斯坦理论. 德拜模型将在下一篇中讲述.

1. 爱因斯坦模型

考虑由 N 个原子组成的晶体. 由于相邻原子间的很强的相互作用,每个原子只能在格点(平衡位置)附近作微振动. 每个原子有 3 个自由度,整个固体有 $3N$ 个自由度. 按照 §1.4.3,在简谐近似下,可以变换到简正坐标,使得这 $3N$ 个耦合的线性谐振子转化为 $3N$ 个独立的线性谐振子. 从线性代数看,这是把一般二次型化为标准二次型或平方和. 原则上,这 $3N$ 个独立振子有不同的频率,但是爱因斯坦模型采用了一种粗略的近似,认为所有这 $3N$ 个振动具有相同的频率 ν,且服从量子力学规律. 这实际上是爱因斯坦(1907 年)对普朗克电磁谐振子能量量子化(1900 年)的大胆推广.

2. 爱因斯坦的比热公式

根据前面所说的统计热力学的三步骤,为求固体比热,首先应按照爱因斯坦模型给出的能谱求出系统的配分函数,然后算出内能. 将内能对温度求微商就得到固体的比热.

在 §5.4.5 中求出了量子谐振子的能量本征值为

$$\varepsilon_n = (n + 1/2)h\nu, \qquad (n = 0,1,2,3,\cdots,) \tag{2.13}$$

每个能级都是非简并的.

根据爱因斯坦模型,由 N 个原子组成的理想晶体的能量为

$$E(n_1,n_2,\cdots,n_{3N}) = \left(n_1+\frac{1}{2}\right)h\nu+\left(n_2+\frac{1}{2}\right)h\nu+\cdots+\left(n_{3N}+\frac{1}{2}\right)h\nu,$$

其中,$n_1,n_2,\cdots,n_{3N}$ 均可独立地取 0,1,2,3,…等值. 系统的微观态由一套量子数 $(n_1,n_2,\cdots,n_{3N})$ 决定,因此系统的配分函数为

$$\begin{aligned} Z &= \sum_{(n_1,n_2,\cdots,n_{3N})} e^{-\beta E(n_1,n_2,\cdots,n_{3N})} \\ &= \sum_{(n_1,n_2,\cdots,n_{3N})} e^{-\beta\left[\left(n_1+\frac{1}{2}\right)h\nu+\left(n_2+\frac{1}{2}\right)h\nu+\cdots+\left(n_{3N}+\frac{1}{2}\right)h\nu\right]}, \end{aligned} \tag{2.14}$$

由于每个 n_i 可以独立取值,故上式可写为

$$Z = \left(\sum_{n_1=0}^{\infty} e^{-\beta\left(n_1+\frac{1}{2}\right)h\nu}\right)\cdot\left(\sum_{n_2=0}^{\infty} e^{-\beta\left(n_2+\frac{1}{2}\right)h\nu}\right)\cdots\left(\sum_{n_{3N}=0}^{\infty} e^{-\beta\left(n_{3N}+\frac{1}{2}\right)h\nu}\right), \tag{2.15}$$

令

$$Z_1 = \sum_{n=0}^{\infty} e^{-\beta\left(n+\frac{1}{2}\right)h\nu}, \tag{2.16}$$

代表单个简正振动的配分函数,则系统的配分函数化为

$$Z = Z_1^{3N}. \tag{2.17}$$

显然,根据配分函数分解的原则,对于由 $3N$ 个独立谐振子组成的爱因斯坦模型,也可直接写出(2.17)式. 利用等式 $1+x+x^2+\cdots=1/(1-x)(|x|<1)$,将(2.16)式中的 $e^{-\beta h\nu}$ 看作 x,可将单个简正振动的配分函数写为

$$Z_1 = e^{-\beta h\nu/2}/(1-e^{-\beta h\nu}), \tag{2.18}$$

于是,

$$Z = [e^{-\beta h\nu/2}/(1-e^{-\beta h\nu})]^{3N}. \tag{2.19}$$

根据 §7.3.1(1.12)式,固体的内能为

$$U = -\partial\ln Z/\partial\beta = 3Nh\nu/2+3Nh\nu/(e^{\beta h\nu}-1). \tag{2.20}$$

其中第一项是与温度无关的零点能. 固体的热容量

$$C_V = \left(\frac{\partial U}{\partial T}\right)_V = 3Nk_B\frac{e^{\beta h\nu}}{(e^{\beta h\nu}-1)^2}(\beta h\nu)^2. \tag{2.21}$$

引入爱因斯坦特征温度 $\Theta_E = h\nu/k_B$(如何确定 Θ_E,如何利用所定 Θ_E 计算 C_V,进而和实验所得 C_V 比较见习题 16)可将热容量表示为:

$$C_V = 3Nk_B\left(\frac{\Theta_E}{T}\right)^2\frac{e^{\Theta_E/T}}{(e^{\Theta_E/T}-1)^2}. \tag{2.22}$$

(2.21)和(2.22)式给出的是定体热容量,其原因是,微商是在频率 ν 不变的条件下进行的,而 ν 与原子间的相互作用力有关,因而和原子间平均距离或体积有关.

下面讨论两种极限情况：

①高温极限，即 $T \gg \Theta_E$，这时有 $e^{\Theta_E/T} - 1 \approx \Theta_E/T$，于是由(2.22)式得

$$C_V \approx 3Nk_B. \tag{2.23}$$

取 $N = N_0$，则有 $C_V = 3R$，和杜隆-珀蒂定律一致.

②低温极限，即 $T \ll \Theta_E$，这时有 $e^{\Theta_E/T} \gg 1$，

$$C_V \approx 3Nk_B(\Theta_E/T)^2 e^{-\Theta_E/T}. \tag{2.24}$$

从定性上看，(2.24)式说明了在低温时，固体比热随温度降低面减小，当 $T \to 0$ 时，$C_V \to 0$ 的事实. 但从定量上看，则与实验结果有较大的差异(见习题 16). (2.24)式预言，当温度降低时，比热会指数减小，这比实验中得到的按 T^3 减小要快得多. 不一致的原因是由于，爱因斯坦模型关于所有简正振动模具有相同频率的假设过于简单. 如果认为每个简正振动模具有不同的频率，则固体内能为

$$U = U_0 + \sum_{i=1}^{3N}\left(\frac{1}{2}h\nu_i + \frac{h\nu_i}{e^{\beta h\nu_i} - 1}\right), \tag{2.25}$$

因而需要计算出所有频率 $\nu_i(i = 1, \cdots, 3N)$. 这过于复杂. 1912 年德拜采用一种近似方法解决了这一问题，这将在下一篇中介绍.

§7.3.3　巨正则系综和巨正则分布

(一)巨正则系综和巨正则分布

考虑一个具有固定体积 V，但与外源既可交换能量又可交换粒子而处于平衡的系统，即开放系统. 因为外源很大，交换能量和粒子不会改变外源的温度和化学势. 达到平衡后，系统和外源具有相同的不变温度和化学势. 所以，处于平衡的开放系统的平衡态由 (T, μ, V) 描述. 大量相同开放系的集合叫做巨正则系综. 三种系综的区别可用示意图 1 表示. 图中单个方框表示系统孤立；实线内框表示可与外源交换能量；虚线内框表示可与外源交换能量与粒子.

因为开放系统的粒子数和能量都是可变的，所以要问：系统的粒子数为 N，处于能量为 $E_{N,l}$ 的某一微观态 (N, l) 的概率 $P(N, l)$ 是多少？我们再一次考虑由系统 A 和外源 A' 组成的复合孤立系统 $A^{(0)}$. 它具有不变的粒子数 $N^{(0)}$ 和能量 $E^{(0)}$. 以 E 和 E' 表示系统 A 和外源 A' 的能量，以 N 和 N' 表示系统 A 和外源 A' 的粒子数. 假设系统和外源相互作用微弱并考虑到孤立系总粒子数不变则有

$$E + E' = E^{(0)}, \qquad N + N' = N^{(0)}. \tag{3.1}$$

因外源很大，所以有 $E \ll E^{(0)}$ 和 $N \ll N^{(0)}$.

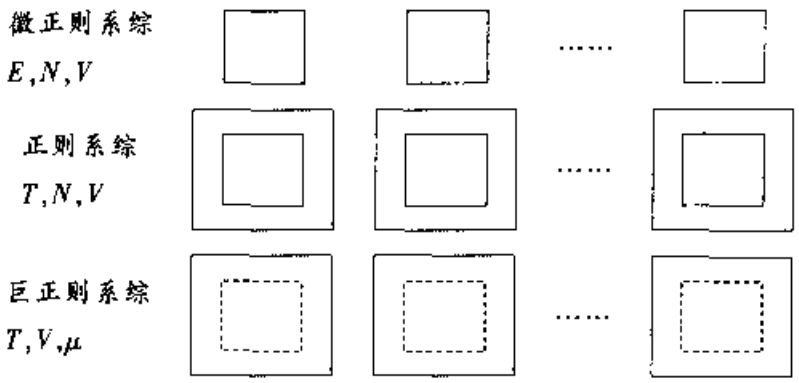


图 1　三种系综的示意图

当系统处于粒子数为 N，能量为 $E_{N,l}$ 的某一微观态 l 时，外源可以处于粒子数为 $(N^{(0)}-N)$，能量为 $(E^{(0)}-E_{N,l})$ 的任何一个微观态，用 $\Omega'(N^{(0)}-N,E^{(0)}-E_{N,l})$ 表示外源粒子数为 $(N^{(0)}-N)$，能量为 $(E^{(0)}-E_{N,l})$ 时的微观态数. 这也就是当系统具有粒子数 N，处于由 (N,l) 标记的微观态时，复合系统的微观态数. 复合系统是一孤立系统，根据微正则分布，在平衡态下，它的每一个可能的微观态出现的概率相等. 所以系统具有粒子数 N，处于能量为 $E_{N,l}$ 的某一微观态 (N,l) 的概率 $P_{N,l}$ 与 $\Omega'(N^{(0)}-N,E^{(0)}-E_{N,l})$ 成正比，即

$$P_{N,l}\propto\Omega'(N^{(0)}-N,E^{(0)}-E_{N,l}). \tag{3.2}$$

或用外源的熵表示

$$P_{N,l}\propto \mathrm{e}^{S'(N^{(0)}-N,E^{(0)}-E_{N,l})/k_{\mathrm{B}}}. \tag{3.3}$$

利用系统远小于外源以及外源的温度和化学势的变化可忽略的事实，将 S' 展成泰勒级数，保留 $E_{N,l}$ 和 N 的线性项

$$\begin{aligned}S'(N^{(0)}-N,E^{(0)}-E_{N,l}) &= S'(N^{(0)},E^{(0)})+\partial S'/\partial E'\big|_{E'=E^{(0)}}(-E_{N,l})\\ &\quad+\partial S'/\partial N'\big|_{N'=N^{(0)}}(-N)+\cdots.\end{aligned}$$

考虑到

$$(\partial S'/\partial E')_{E'=E^{(0)}}=1/T, \tag{3.4}$$

$$(\partial S'/\partial N')_{N'=N^{(0)}}=-\mu/T, \tag{3.5}$$

上式可写为

$$S'(N^{(0)}-N,E^{(0)}-E_{N,l})/k_{\mathrm{B}}\approx S'(N^{(0)},E^{(0)})/k_{\mathrm{B}}-\beta E_{N,l}+\beta\mu N, \tag{3.6}$$

其中 T 和 μ 是外源的温度和化学势. 由于系统与外源达到平衡，T 和 μ 也就是系统的温度和化学势. (3.6)式右方第一项是与系统无关的常数，代入(3.3)式，得

$$P_{N,l}=c\mathrm{e}^{\beta\mu N-\beta E_{N,l}}. \tag{3.7}$$

对上式进行归一化，首先对粒子数取某一特定 N 值时的全部微观态 l 求和，然后对粒子数所有可能值求和，最后可得

$$P_{N,l}=\mathrm{e}^{\beta\mu N-\beta E_{N,l}}/\Xi. \tag{3.8}$$

其中

$$\tilde{Z} \equiv \sum_N \Big(\sum_l e^{\beta\mu N - \beta E_{N,l}} \Big), \tag{3.9}$$

叫做巨配分函数. 定义

$$z \equiv e^{\beta\mu}, \tag{3.10}$$

为逸度,并注意到粒子数为 N 的恒温系统的配分函数为

$$Z_N = \sum_l e^{-\beta E_{N,l}}, \tag{3.11}$$

巨配分函数又可写为

$$\tilde{Z}(T,\mu,V) = \sum_N z^N Z_N(T,V). \tag{3.12}$$

(3.8)式给出了温度为 T,化学势为 μ 的开放系统处于粒子数为 N,能量为 $E_{N,l}$ 的某一微观态 (N,l) 的概率,叫做巨正则分布. (3.12)式表明,巨配分函数是以 (T,μ,V) 为自变量的态函数. 从 §7.2.5 的(3.36)式已经知道巨势 $\Omega_{巨}$ 是以 (T,μ,V) 为自变量的特性函数. 因此可以预料 $\tilde{Z}(T,\mu,V)$ 和 $\Omega_{巨}$ 有关. 下面将推导它们之间的关系,并将看到,和 Z 在恒温系统中的地位相似,$\tilde{Z}$ 对计算开放系统的热力学量有重要作用.

(二)巨配分函数和热力学量的关系

按照上节给出的熵的一般表达式[§7.3.1(1.19)式],在现在情况下有

$$S = -k_B \sum_N \sum_l P(N,l) \ln P(N,l) = -k_B \sum_N \sum_l P(N,l)[\beta\mu N - \beta E_{N,l} - \ln \tilde{Z}].$$

注意到平均粒子数和平均能量为

$$\sum_N \sum_l P(N,l) N = \langle N \rangle, \tag{3.13}$$

$$\sum_N \sum_l P(N,l) E_{N,l} = \langle E \rangle = U. \tag{3.14}$$

上式可写为 $TS = U - \mu\langle N\rangle + k_B T \ln \tilde{Z}$,或

$$-k_B T \ln \tilde{Z} = U - TS - \mu\langle N\rangle = F - G. \tag{3.15}$$

将此式和 §7.2.5(5.37)式比较,可得巨热力学势(或巨势)的统计表达式

$$\Omega_{巨}(T,\mu,V) = -k_B T \ln \tilde{Z}(T,\mu,V). \tag{3.16}$$

由于巨势是以 (T,μ,V) 为独立变量的特性函数,立刻可以写出[见 §7.2.5(5.38)式]

$$\langle N \rangle = -(\partial\Omega_{巨}/\partial\mu)_{T,V} = k_B T (\partial \ln \tilde{Z}/\partial\mu)_{T,V}, \tag{3.17}$$

$$p = -(\partial\Omega_{巨}/\partial V)_{T,\mu} = k_B T (\partial \ln \tilde{Z}/\partial V)_{T,\mu}, \tag{3.18}$$

$$S = -(\partial\Omega_{巨}/\partial T)_{V,\mu}$$

$$= k_B [\ln \tilde{Z} - \beta(\partial \ln \tilde{Z}/\partial\beta)_{\alpha,V} - \mu(\partial \ln \tilde{Z}/\partial\mu)_{\beta,V}]. \tag{3.19}$$

其中 $\alpha \equiv \beta\mu$，将以上有关公式代入 $U \equiv TS + \mu\langle N\rangle + \Omega_{巨}$ 可得

$$U = -(\partial \ln \bar{Z}/\partial\beta)_{V,\alpha}. \tag{3.20}$$

可见，只要求得巨配分函数 $\bar{Z}$，就可以求得以 (T,μ,V) 为独立变量的特性函数——巨势 $\Omega_{巨}$，进而用简单的求偏导数的方法计算出全部热力学量. 巨配分函数 $\bar{Z}$ 及巨势 $\Omega_{巨}$ 在巨正则系综中的作用，和微观态数 Ω 及熵 S 在微正则系综中的作用，以及配分函数 Z 及自由能 F 在正则系综中的作用完全类似，它们都是把微观和宏观、统计物理和热力学联系起来的纽带.

(三)开放系统的粒子数涨落

开放系统粒子数的均方涨落为

$$\begin{aligned}\langle(\Delta N)^2\rangle &\equiv \langle(N-\langle N\rangle)^2\rangle \\ &= \langle N^2\rangle - \langle N\rangle^2.\end{aligned} \tag{3.21}$$

但

$$\begin{aligned}\left(\frac{\partial\langle N\rangle}{\partial\alpha}\right)_{\beta,V} &= \frac{\partial}{\partial\alpha}\left(\sum_N\sum_l N\mathrm{e}^{\alpha N-\beta E_{N,l}}\Big/\sum_N\sum_l \mathrm{e}^{\alpha N-\beta E_{N,l}}\right) \\ &= \sum_N\sum_l N^2\mathrm{e}^{\alpha N-\beta E_{N,l}}\Big/\sum_N\sum_l \mathrm{e}^{\alpha N-\beta E_{N,l}} \\ &\quad -\left(\sum_N\sum_l N\mathrm{e}^{\alpha N-\beta E_{N,l}}\Big/\sum_N\sum_l \mathrm{e}^{\alpha N-\beta E_{N,l}}\right)^2 \\ &= (\langle N^2\rangle - \langle N\rangle^2).\end{aligned}$$

所以

$$\begin{aligned}\langle(\Delta N)^2\rangle &= (\partial\langle N\rangle/\partial\alpha)_{\beta,V} \\ &= k_{\mathrm{B}}T(\partial\langle N\rangle/\partial\mu)_{T,V}.\end{aligned} \tag{3.22}$$

粒子数的相对涨落为

$$\begin{aligned}\delta N &= \left[\frac{\langle(\Delta N)^2\rangle}{\langle N\rangle^2}\right]^{1/2} \\ &= \left[\frac{k_{\mathrm{B}}T}{\langle N\rangle^2}\left(\frac{\partial\langle N\rangle}{\partial\mu}\right)_{T,V}\right]^{1/2}.\end{aligned} \tag{3.23}$$

进一步讨论可以证明

$$\delta N = (k_{\mathrm{B}}T\kappa_T/V)^{1/2}. \tag{3.24}$$

当等温压缩系数 $\kappa_T \equiv -(\partial V/\partial p)_T/V$ 有限时，δN 和 $V^{1/2}(N^{1/2})$ 成反比，因此，对于宏观系统，粒子数的相对涨落完全可以忽略. 巨正则系综和正则系综等价. 但当系统处于临界点时，κ_T 趋于无穷，粒子的相对涨落会很大，这一问题将在第九篇中讨论.

表Ⅰ　三种系综的比较

系统平衡态的描述	分布名称	分布函数	配分函数	特性函数
E,N,V	微正则分布	$\begin{cases}\frac{1}{\Omega} & E\sim E+\delta E\\ 0 & \text{其他}\end{cases}$	$\Omega(E,N,V)$	$S=k_B\ln\Omega(E,V,N)$
T,N,V	正则分布	$\frac{1}{Z}e^{-\beta E_l}$	$Z=\sum_l e^{-\beta E_l}=\sum_{E_l} e^{-\beta E_l}\Omega(E_l)$	$F=-k_BT\ln Z(T,V,N)$
T,μ,V	巨正则分布	$(1/\bar{Z})e^{\beta(\mu N-E_{N,l})}$	$\bar{Z}=\sum_N e^{\beta\mu N}\times\left(\sum_l e^{-\beta E_{N,l}}\right)$	$\Omega_{巨}=-k_BT\ln\bar{Z}(T,\mu,V)$

为了便于比较，将三种系综总结在表Ⅰ上（注意，微正则分布的配分函数 Ω 是微观态数；巨正则分布的特性函数 $\Omega_{巨}$ 是巨热力势，不要混淆）.

（四）巨正则分布的简单应用

吸附有广泛应用，如气体和溶液的净化，石油制品和化学药品生产中的脱色及防毒面具等. 讨论表面吸附往往要提到朗格谬尔（Longmuir）模型. 根据这一模型导出的方程迄今仍有实际意义. 朗格谬尔模型见下例：

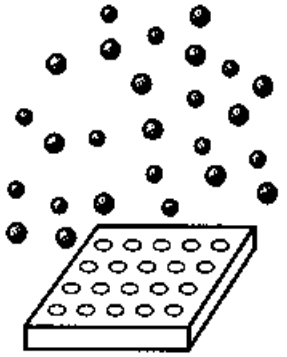


图 2　吸附中心

【例 1】　固体表面可以看成是由规则的分子（或原子、离子）组成，每个分子可以吸引气体分子，叫做吸附中心，见图 2. 与自由态相比，气体分子在吸附中心有负势能 $-\varepsilon_0$. 由于热运动，被吸附的气体分子又能够挣脱吸引回到气体中去，所以在一定的温度下达到平衡时，有一定量的气体分子在吸附点上. 被吸附的平均气体分子数 $\langle N\rangle$ 和吸附中心数 N_0 之比叫吸附率. 假设气体是理想气体，求吸附率 θ 与气体压强的关系.

解　把吸附中心构成的表面看成是一个与外界既可交换粒子，又可交换能量的开放系. 当有 N 个分子被吸附时，系统的能量为 $-N\varepsilon_0$. 因为被吸附的 N 个分子在 N_0 个吸附中心上有 $N_0!/N!(N_0-N)!$ 种不同的分配方法，所以 $Z_N=(N_0!/N!(N_0-N)!)e^{\beta N\varepsilon_0}$. 系统的巨配分函数为

$$\bar{Z}=\sum_N e^{\beta\mu N}Z_N=\sum_N\frac{N_0!}{N!(N_0-N)!}[e^{\beta(\varepsilon_0+\mu)}]^N=[1+e^{\beta(\varepsilon_0+\mu)}]^{N_0}=\bar{Z}_1^{N_0}.$$

这表明，表面的巨配分函数可以分解为吸附中心巨配分函数 $\bar{Z}_1$ 之积（请推导巨配分函数分解的一般条件）. 由(3.17)式，吸附的平均气体分子数为

$$\langle N\rangle = k_B T\left(\frac{\partial \ln \tilde{Z}}{\partial \mu}\right)_{T,V} = N_0 k_B T \frac{\partial}{\partial \mu}\ln\left(1+e^{\beta(\varepsilon_0+\mu)}\right) = \frac{N_0}{\left(1+e^{-\beta(\varepsilon_0+\mu)}\right)}.$$

因为吸附表面与理想气体达到平衡，所以上式中的 T 和 μ 也是理想气体的温度和化学势. 在 §7.2.3(3.18)式的第一项中加上 $k_B T\ln k_B T$ 使它成为 $k_B T\ln(RT/V) = k_B T\ln p$. 第二项减去 $k_B T\ln k_B T$ 成为 T 的另一函数，将它写为 $k_B\psi(T)$. 于是理想气体的化学势可写为：

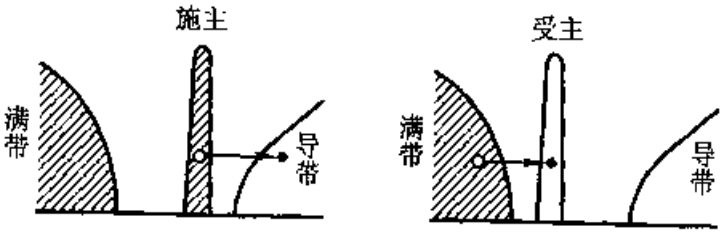
$$\mu = k_B T[\ln p + \psi(T)],$$

或 $e^{-\beta\mu} = e^{-\psi(T)}/p$，代入 $\langle N\rangle$ 得

$$\theta = \langle N\rangle/N_0 = 1/[1+e^{-\beta\varepsilon_0-\psi(T)}/p].$$

令 $e^{-\beta\varepsilon_0-\psi(T)} = p_0$，则吸附率 $\theta = p/(p+p_0)$. 若 $p \ll p_0$，则吸附的气体分子数为 $\langle N\rangle = N_0 p/p_0$. p_0 的意义如下：当吸附率 $\theta = 1/2$ 时，有 $p = p_0$，即 p_0 是平均吸附气体分子数等于吸附中心数的一半时，理想气体的压强值. 它是温度的函数.

【例 2】 杂质半导体有两种类型：N 型半导体和 P 型半导体. N 型半导体(如在硅中掺砷)在 0K 时有被电子填满的杂质能级[图 3(a)]，在较高温度下，电子从这些杂质能级激发到导带，这样的杂质能级称为施主能级. P 型半导体(如硅中掺硼)在 0K 时有未被电子占据的杂质能级[图 3(b)]，在较高温度下，电子从满带激发到这些杂质能级，于是满带中产生了空穴，这样的能级称为受主能级. 半导体中杂质原子数远小于母晶格原子数，平均来说杂质原子相距很大，可视为近独立. 在电场作用下，电子和空穴都可以形成电流，半导体中对电导有贡献的粒子(导带中的电子和满带中的空穴)统称为载流子. 载流子的数目随温度升高而指数增加，从而半导体的电导率也随温度升高而指数增加. 这和金属的电导率随温度升高而减少完全不同，是半导体的主要特性. 可见，研究杂质能级上的平均电子数对温度的依赖有重要意义.



(a)N 型半导体　　(b)P 型半导体

图 3　半导体杂质能级

假设某种 N 型半导体单位体积中有 N 个施主能级，它们的能量是 ε_D. 一个施主能级可以被一个自旋向上的电子占有，也可以被一个自旋向下的电子占有，但不能同时被一个以上电子占有. 求施主能级中的平均电子数.

解　我们以施主能级作为所研究的系统. 施主能级可能被电子占有，也可能不被电子占有，因此是一个粒子数和能量均可变的开放系，而导带则可视为外源. 此

开放系有三个可能的状态：不被电子占有时，$n=0, E_0=n\varepsilon=0$；被自旋向上的电子占有时，$n=1, E_{1,1}=n\varepsilon=\varepsilon_D$；被自旋向下的电子占有时，仍有 $n=1, E_{1,2}=n\varepsilon=\varepsilon_D$. 施主能级的巨配分函数为

$$\tilde{Z}_1=\sum_{n,l}e^{\beta\mu n-\beta E_{n,l}}=e^{\beta\mu\cdot 0-\beta\cdot 0}+e^{\beta\mu\cdot 1}\left[e^{-\beta E_{1,1}}+e^{-\beta E_{1,2}}\right]=1+2e^{-\beta(\varepsilon_D-\mu)}.$$

由(3.17)式，施主能级上的平均电子数为

$$\langle n\rangle=k_BT(\partial\ln\tilde{Z}_1/\partial\mu)_{T,V}=k_BT\frac{\partial}{\partial\mu}\ln\left(1+2e^{-\beta(\varepsilon_D-\mu)}\right)=\left[1+\frac{1}{2}e^{\beta(\varepsilon_D-\mu)}\right]^{-1}.$$

N 个施主能级上的总平均电子数，即此 N 型半导体导带中的电子浓度为

$$N\langle n\rangle=N/(1+e^{\beta(\varepsilon_D-\mu)}/2).$$

当 $\varepsilon_D-\mu\gg k_BT$ 时，近似有

$$\text{电子浓度}\propto e^{-(\varepsilon_D-\mu)/(k_BT)}.$$

§7.3.4　理想气体的三种分布

第一章已经指出，系综理论是一种普遍的统计理论，适用于任何系统. 然而，统计物理最简单和最重要的研究对象是理想气体，即由相互作用可以忽略的非定域粒子组成的系统. 要注意的是，这里所说的相互作用是指动力学相互作用. 当忽略动力学相互作用后，由于粒子的全同性，非定域粒子间仍然存在着量子力学的交换作用，即由于费米粒子的波函数对粒子交换反对称、玻色粒子的波函数对粒子交换对称所导致的相互作用，见 §5.5.1，§6.1.3. 由费米粒子组成的理想气体称为费米气体，由玻色粒子组成的气体称为玻色气体.

作为巨正则分布的应用，本节将首先推导出在平衡状态下，费米粒子和玻色粒子按能量的分布规律，即费米-狄拉克(F-D)分布和玻色-爱因斯坦(B-E)分布. 然后证明，在一定条件下，当交换作用也可忽略时，F-D 分布和 B-E 分布将趋于一共同极限，即麦克斯韦-玻尔兹曼(M-B)分布，或简称为麦－玻分布或玻尔兹曼分布. 关于三种分布的应用将在第八篇中详细讨论.

首先对理想气体的概念及其微观态的描述作进一步说明：

(一)理想气体及其微观态的描述　单粒子态占有数集合

由近独立非定域粒子组成的系统称为理想气体. 理想气体的能量等子各粒子能量之和. 在理解理想气体的定义时，除须注意忽略掉的相互作用只是动力学相互作用外，还要注意的是，近独立粒子间不能完全没有相互作用. 因为，如果近独立粒子间完全没有相互作用，整个气体便不能达到平衡. 所以，组成理想气体的粒子是这样的粒子，它们间的相互作用势能和各粒子本身的能量相比可以略去，但它们之

间又存在着微弱的相互作用，以保证整个气体达到平衡.

对于近独立粒子，单粒子态和单粒子能量均有确切意义. 我们用 $1,2,\cdots,r,\cdots$ 表示各单粒子态；用 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r,\cdots$ 表示各单粒子态的能量；用 $n_1,n_2,\cdots,n_r,\cdots$ 表示各单粒子态上的粒子数. 根据量子力学，属性相同的非定域粒子是不可区分的，所以指出某个粒子处于那个单粒子态毫无意义，有意义的只是单粒子态上的粒子数. 于是可以用一套单粒子态占据数集合 $\{n_1,n_2,\cdots,n_r,\cdots\}$ 代表理想气体的一个微观态，即福克表象，见 §5.5.1. 为了方便，往往缩写为 $\{n_r\}$.

单粒子态：　$1,2,\cdots,r,\cdots$　$\rightarrow \{r\}$

单粒子态的能量：　$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r,\cdots$　$\rightarrow \{\varepsilon_r\}$

单粒子态上的粒子数：　$n_1,n_2,\cdots,n_r,\cdots$　$\rightarrow \{n_r\}$

当然 $\{n_r\}$ 必须满足以下条件：

$$\sum_r n_r = N, \qquad \sum_r n_r\varepsilon_r = E. \tag{4.1}$$

这两个和式是对所有单粒子态求和. 显然，在满足(4.1)式的条件下，N 个粒子在各单粒子态可以有许多不同的分配方法，即可以有许多不同的单粒子态占据数集合 $\{n_r\}$. 也就是说，存在大量可实现的不同微观态，且微观态数和气体的性质(费米、玻色或玻尔兹曼气体)有关. 考查下面的例子.

【例】 设系统由 $N=4$ 个近独立粒子组成，每个粒子可处于七个单粒子态，它们的能量分别为 $0,\varepsilon,2\varepsilon,3\varepsilon,4\varepsilon,5\varepsilon$ 和 6ε，系统的总能量为 E. 试就以下三种情况指出和 $\left(N=4,E=\sum_r n_r\varepsilon_r=6\varepsilon\right)$ 对应的微观态数：(1)粒子可识别；(2)粒子是不可识别的费米子；(3)粒子是不可识别的玻色子.

解 表 I 给出了问题的答案. 表中第一行表示能量分别为 $0,\varepsilon,2\varepsilon$ 和 3ε 的单粒子态上各有一个粒子，总粒子数 $N=4$，总能量 $E=0+1\cdot\varepsilon+1\cdot2\varepsilon+1\cdot3\varepsilon=6\varepsilon$. 若 4 个粒子可以识别，标记为 A,B,C,D，则将它们在四个单粒子态上交换可形成 $4!=24$ 个不同的微观态. 若 4 个粒子不可识别，则将它们在四个单粒子态上交换不形成不同的微观态. 表中第二行表示能量为 0 的单粒子态上有三个粒子，能量为 6ε 的单粒子态上有一个粒子，仍满足 $N=4$ 和 $E=3\cdot0+1\cdot6\varepsilon=6\varepsilon$ 的条件. 若粒子可识别，将它们交换可形成 $4!/3!=4$ 种不同微观态. 若粒子是不可区分的费米子，则根据泡利不相容原理，这是不可能的. 若粒子是不可区分的玻色子，由于不受泡利不相容原理限制，这是可能的，但因粒子不可区分，粒子交换不形成不同微观态. 以下各行类推.

表 I　有固定能量的四个近独立粒子系的微观态

单粒子态							微观态数			单粒子态占有数集合 $\{n_r\}$
0	ε	2ε	3ε	4ε	5ε	6ε	粒子可识别	粒子不可识别		
								费米	玻色	
·	·	·	·				$4!=24$	1	1	$\{1,1,1,1,0,0,0\}$
· · ·						·	$\frac{4!}{3!}=4$		1	$\{3,0,0,0,0,0,1\}$
· ·	·				·		$\frac{4!}{2!}=12$		1	$\{2,1,0,0,0,1,0\}$
· ·		·		·			$\frac{4!}{2!}=12$		1	$\{2,0,1,0,1,0,0\}$
·	· ·			·			$\frac{4!}{2!}=12$		1	$\{1,2,0,0,1,0,0\}$
· ·			· ·				$\frac{4!}{2!\ 2!}=6$		1	$\{2,0,0,2,0,0,0\}$
	· ·	· ·					$\frac{4!}{2!\ 2!}=6$		1	$\{0,2,2,0,0,0,0\}$
·	· · ·						$\frac{4!}{3!}=4$		1	$\{1,0,3,0,0,0,0\}$
	· · ·		·				$\frac{4!}{3!}=4$		1	$\{0,3,0,1,0,0,0\}$
$N=4, E=6\varepsilon$ 的总微观态数							84	1	9	

(二)用巨正则系综推导 F-D 分布和 B-E 分布

我们现在要问的是,当理想气体达到平衡态时,各单粒子态上的平均粒子数和单粒子态的能量有什么关系? 即在平衡态下,理想气体粒子按能量分布的规律是什么? 显然,分布规律和理想气体的性质有关.

根据 §7.3.3(3.17)式,开放系统的平均粒子数为

$$\langle N\rangle = -\partial\Omega_{\text{巨}}/\partial\mu = k_{\text{B}}T\partial\ln\tilde{Z}/\partial\mu = \sum_r \langle n_r\rangle. \tag{4.2}$$

由此可见,计算出系统的巨配分函数 $\tilde{Z}$,就可以求出单粒子态的平均粒子数 $\langle n_r\rangle$.

已知理想气体的微观态由一套单粒子占据数集合 $\{n_r\}$ 表示,该微观态的能量为 $E(\{n_r\}) = \sum_r n_r\varepsilon_r$,所以理想气体的配分函数为

$$Z_N = \sum_{\substack{\{n_r\}\\ \sum_r n_r = N}} \mathrm{e}^{-\beta E_N(\{n_r\})} = \sum_{\{n_r\}} {}' \mathrm{e}^{-\beta(n_1\varepsilon_1+n_2\varepsilon_2+\cdots+n_r\varepsilon_r+\cdots)}. \tag{4.3}$$

右边的求和式表示对各种不同的单粒子占据数集合 $\{n_r\}$,即对不同的微观态求

和.求和上的撇号表示每一种占据数集合必须满足约束条件 $\sum_r n_r = N$.由于这一约束,使得配分函数难以计算.

考虑巨配分函数,据§7.3.3(3.12)式,有

$$\tilde{Z} = \sum_{N=0}^{\infty} e^{\beta\mu N} Z_N = \sum_{N=0}^{\infty} e^{\beta\mu \sum_r n_r} \Big\{ \sum_{\{n_r\}}{}' e^{-\beta \sum_r n_r \varepsilon_r} \Big\} = \sum_{N=0}^{\infty} \sum_{\{n_r\}}{}' e^{-\beta \sum_r n_r (\varepsilon_r - \mu)}.$$

由于第一个求和,即由于总粒子数可取任意值,第二个求和时的约束 $\sum_r n_r = N$ 实际上可以解除,两个求和可以用对每一个 n_r 的独立求和代替,即

$$\tilde{Z} = \Big(\sum_{n_1} e^{-\beta n_1 (\varepsilon_1 - \mu)} \Big) \Big(\sum_{n_2} e^{-\beta n_2 (\varepsilon_2 - \mu)} \Big) \cdots = \prod_r \Big\{ \sum_{n_r} e^{-\beta n_r (\varepsilon_r - \mu)} \Big\}. \tag{4.4}$$

以上的讨论是一般的,下面分别考虑费米气体和玻色气体.对于费米气体,$n_r = 0,1$ 只有两个值,于是

$$\tilde{Z}_{F\text{-}D} = \prod_r (1 + e^{-\beta(\varepsilon_r - \mu)}). \tag{4.5}$$

对于玻色气体,$n_r = 0,1,2,\cdots$,于是

$$\tilde{Z}_{B\text{-}E} = \prod_r \{1 + e^{-\beta(\varepsilon_r - \mu)} + e^{-2\beta(\varepsilon_r - \mu)} + \cdots\} = \prod_r \frac{1}{1 - e^{-\beta(\varepsilon_r - \mu)}}. \tag{4.6}$$

由(4.5)式和(4.6)式得

$$\Omega_{巨} = -k_B T \ln \tilde{Z} = \begin{cases} -k_B T \sum_r \ln [1 + e^{-\beta(\varepsilon_r - \mu)}] & (F\text{-}D), \\ k_B T \sum_r \ln [1 - e^{-\beta(\varepsilon_r - \mu)}] & (B\text{-}E), \end{cases} \tag{4.7}$$

平均粒子数 $\langle N \rangle = \sum_r \langle n_r \rangle = -\partial \Omega_{巨} / \partial \mu$,

$$\langle N \rangle = \begin{cases} \sum_r \dfrac{1}{e^{\beta(\varepsilon_r - \mu)} + 1} & (F\text{-}D), \\ \sum_r \dfrac{1}{e^{\beta(\varepsilon_r - \mu)} - 1} & (B\text{-}E), \end{cases} \tag{4.8}$$

$$\langle n_r \rangle = \begin{cases} \dfrac{1}{e^{\beta(\varepsilon_r - \mu)} + 1} & (F\text{-}D), \quad (4.9) \\ \dfrac{1}{e^{\beta(\varepsilon_r - \mu)} - 1} & (B\text{-}E), \quad (4.10) \end{cases}$$

(4.9)式和(4.10)式是两个重要结果,分别称为费米-狄拉克分布(F-D 分布)和玻色-爱因斯坦分布(B-E 分布).$\langle n_r \rangle$ 代表理想气体达到平衡态时,能量为 ε_r 的单粒子态上的平均粒子数,通常称为分布函数.

由(4.8)式看到,对于粒子数守恒的系统,(4.8)式左边是一个常量,因而当改变系统温度时,(4.8)式右边应保持不变.又已知单粒子态的能量 ε_r 仅仅依赖于系

统的体积 V，而和温度无关，于是化学势必须依赖于温度，即有 $\mu = \mu(T)$.

(4.9)式和(4.10)式对任何能级均适合. 考虑基态 $\varepsilon_r = 0$ ❶，则由(4.10)式得到

$$n_0 = 1/(e^{-\beta\mu} - 1).$$

因为粒子数不可能为负，所以对于玻色气体必须有

$$e^{-\beta\mu} \geqslant 1 \qquad 或 \qquad \mu \leqslant 0, \tag{4.11}$$

即：玻色气体的化学势不可能为正. 以后将看到，$\mu = 0$ 对应于粒子数不守恒的玻色系统，如黑体辐射中的光子气；而 $\mu < 0$ 则对应于粒子数守恒的玻色系统，并在玻色凝聚中起关键作用. 对费米气体则没有这一限制，它的化学势可为正也可为负.[参看 §8.2.5(5.4)式及该章习题 13].

(三)由 F-D 分布和 B-E 分布过渡到 M-B 分布

由(4.9)式和(4.10)式可以看到，当满足条件

$$\langle n_r \rangle \ll 1, \tag{4.12}$$

或

$$e^{\beta(\varepsilon_r - \mu)} \gg 1 \tag{4.13}$$

时，(4.9)、(4.10)式分母中的 1 可以略去，因而 F-D 分布和 B-E 分布趋于一个共同表达式

$$\langle n_r \rangle = e^{\beta(\mu - \varepsilon_r)}. \tag{4.14}$$

利用以下条件

$$\sum_r \langle n_r \rangle = e^{\beta\mu} \sum_r e^{-\beta\varepsilon_r} = N. \tag{4.15}$$

可将(4.14)式改写为

$$\langle n_r \rangle = (N/Z_1) e^{-\beta\varepsilon_r}. \tag{4.16}$$

用 g_r 表示单粒子能级 ε_r 的简并度，则能级 ε_r 中的平均粒子数是

$$\langle N_r \rangle = g_r \langle n_r \rangle = (N/Z_1) e^{-\beta\varepsilon_r} g_r. \tag{4.17}$$

其中 Z_1 叫做单粒子配分函数，定义为

$$Z_1 = \sum_r e^{-\beta\varepsilon_r} = \sum_{\varepsilon_r} e^{-\beta\varepsilon_r} g_r. \tag{4.18}$$

$\sum_r$ 表示对所有单粒子态求和，$\sum_{\varepsilon_r}$ 表示对所有单粒子能级求和. (4.16)和(4.17)式叫麦克斯韦-玻尔兹曼分布，简称为玻尔兹曼分布. 服从麦克斯韦-玻尔兹曼分布(M-B 分布)的气体叫做玻尔兹曼气体或非简并气体. 而费米气体和玻色气体往往又叫做简并气体.

---

❶ 对于宏观系统，单粒子的零点能可以忽略不计.

要强调的是，根据量子力学，自然界只存在两种粒子：费米子和玻色子，分别服从 F-D 分布和 B-E 分布，而 M-B 分布是两者在量子效应可以忽略时的共同极限.

图 1 画出了 B-E 分布，F-D 分布及 M-B 分布. 由图也可以看出，当 $\langle n_r \rangle \ll 1$ 时，B-E 分布和 F-D 分布趋于共同极限——M-B 分布.

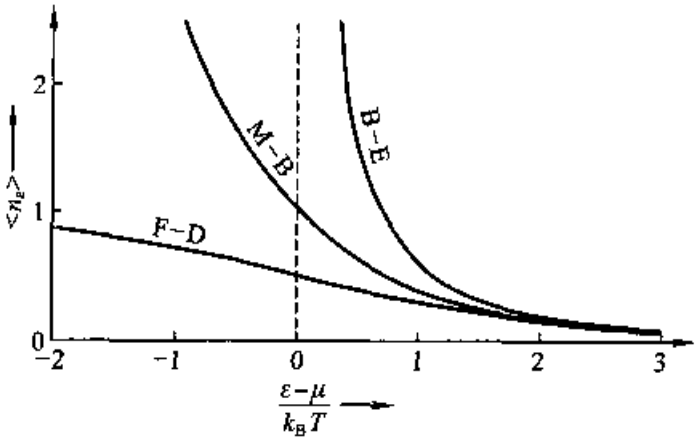


图 1　三种理想气体的分布

要使(4.13)式对任何单粒子态，包括 $\varepsilon_r = 0$ 的基态均适合，必须有

$$e^{-\beta\mu} \gg 1. \qquad (4.19)$$

(4.12)或(4.19)式叫做非简并化条件. 下面对非简并化条件进行讨论，以便明确玻尔兹曼分布的适用范围和非简并条件的物理意义.

(四)对非简并条件的讨论

为了明确 M-B 分布适用于怎样的实际情况，我们假设非简并条件已经满足，看会得到什么结果. 将(4.15)式代入(4.19)式，将非简并条件写为

$$e^{-\beta\mu} = N^{-1} \sum_r e^{-\beta\varepsilon_r} \gg 1. \qquad (4.20)$$

只要算出单粒子配分函数，非简并条件可以具体化. 考虑盛在边长为 L 的容器中的理想气体，根据

$$\varepsilon_r = (n_x^2 + n_y^2 + n_z^2) \cdot h^2/(2mL^2), \qquad (4.21)$$

单粒子配分函数为

$$\begin{aligned} Z_1 &= \sum_{n_x} \sum_{n_y} \sum_{n_z} e^{-\beta h^2 (n_x^2+n_y^2+n_z^2)/(2mL^2)} \\ &= \sum_{n_x} e^{-\beta h^2 n_x^2/(2mL^2)} \sum_{n_y} e^{-\beta h^2 n_y^2/(2mL^2)} \sum_{n_z} e^{-\beta h^2 n_z^2/(2mL^2)}. \end{aligned} \qquad (4.22)$$

设边长为 1cm，则气体分子的能级间隔是 $\Delta\varepsilon \approx h^2/(2m) \approx 10^{-30}$ erg，而在室温(300K)下，能量 ε_r 的典型值是 $k_B T \approx 10^{-14}$ erg. 所以即使在很低的温度下，与 $k_B T$ 相比，总可以认为粒子的平动能量是连续变化的. 与此对应，n_x, n_y, n_z 的变化可以认为是连续的. (4.22)式中的求和可用积分代替，即

$$Z_1 = \left[\int_{-\infty}^{\infty} e^{-\frac{h^2}{2mL^2 k_B T} n_x^2} dn_x\right]\left[\int_{-\infty}^{\infty} e^{-\frac{h^2}{2mL^2 k_B T} n_y^2} dn_y\right]\left[\int_{-\infty}^{\infty} e^{-\frac{h^2}{2mL^2 k_B T} n_z^2} dn_z\right].$$

利用附录(IV.2)式算出积分，得

$$Z_1 = V(2\pi mk_B T/h^2)^{3/2}. \tag{4.23}$$

将(4.23)式代入(4.20)式,则非简并条件可化为

$$(V/N)(\sqrt{2\pi mk_B T}/h)^3 \gg 1. \tag{4.24}$$

(4.24)式表明,如果粒子质量不太小,气体温度不太低,气体的密度不太高,则玻尔兹曼分布适用.考虑粒子间的平均距离为 d,粒子的动量近似为 $p \approx m\bar{v} \approx \sqrt{2\pi mk_B T}$,而相应的热波长近似为 $\lambda_{热} = h/p \approx h/\sqrt{2\pi mk_B T}$,则非简并条件又可写为

$$d \gg \lambda_{热}. \tag{4.25}$$

即当粒子间的平均距离远大子粒子的热波长时,F-D 分布和 B-E 分布过渡到 M-B 分布.

为了获得数量级的概念,考虑两个例子.对于室温和 1atm 下的氮气,有关参量是:平均压强 $p = 10^6 \mathrm{dyn \cdot cm^{-2}}$,$T \approx 300\ \mathrm{K}$,$k_B T \approx 4.1 \times 10^{-14}\ \mathrm{erg}$,分子质量 $m = 6.6 \times 10^{-24}\ \mathrm{g}$,理想气体状态方程给出 $N/V = p/(k_B T) = 2.5 \times 10^{19}$ 分子/$\mathrm{cm^3}$,因此估计值为 $\lambda_{热} \approx 0.14$ Å 及 $d \approx 33$ Å.可见,(4.25)式得到很好的满足.大多数气体的分子质量更大,热波长更短,(4.25)式满足得更好.因此,对于一般的由原子、分子组成的气体,M-B 分布是适用的.

再考虑金属中的自由电子.因为电子的质量比氮原子小 7300 倍,在室温下的热波长为 $\lambda_{热} \approx (0.14) \times \sqrt{7300} \approx 12$ Å.其次,考虑到金属中每一个原子大约有一个自由电子,而原子间距大约为 2Å,故 $d \approx 2$ Å,于是对于金属中的自由电子气直到温度高达 2000~3000K,非简并条件仍不能满足,必须用 F-D 分布处理,而半导体中的电子数较少,往往又可近似用 M-B 分布.

非简并条件 $d \gg \lambda_{热}$ 表明,由于气体极其稀薄,粒子间的平均距离远大于粒子的热波长,波包不会发生重叠.因而在这种情况下,不仅粒子间动力学相互作用可以忽略,量子力学交换作用也可忽略.换句话说,非简并条件的物理意义是,玻尔兹曼气体是一切相互作用均可忽略的气体.

(五)玻尔兹曼气体的配分函数

已经看到,由于存在粒子数守恒的约束,F-D 气体和 B-E 气体的配分函数不能简单地表示出来.但对于 M-B 气体的配分函数却可以做到这一点.为简单具体起见,以由两个不可识别的,近独立粒子组成的系统(假定有 $\langle n_r \rangle \ll 1$)进行证明:

仍然从(4.3)式出发,根据玻尔兹曼气体的特点,可以将 Z 写成以下形式

$$Z = \sum_l \mathrm{e}^{-\beta E_l} = \frac{1}{2!} \sum_r \sum_{s \neq r} \mathrm{e}^{-\beta(\varepsilon_r + \varepsilon_s)}. \tag{4.26}$$

说明如下:① E_l 代表系统处于微观态 l 时的能量,ε_r 和 ε_s 分别代表第一个粒子处

于单粒子态 r 和第二个粒子处于单粒子态 s 时的能量. ②用 $\sum_{r}\sum_{s\atop r\neq s}$ 代替 $\sum_{l}$ 的原因是,由于对 M-B 气体 $\langle n_r\rangle \ll 1$,两个粒子同时处于同一个单粒子态的概率很小. ③当粒子不可识别时,第一个粒子处在单粒子态 r 第二个粒子处于单粒子态 s 和第一个粒子处在单粒子 s 第二个粒子处于单粒子态 r,代表系统的同一个微观态. 为了正确地表示对系统的不同微观态求和,引入了 1/2!.

再来计算单粒于配分函数的乘积

$$Z_1^2 = \Big(\sum_r e^{-\beta\varepsilon_r}\Big)\Big(\sum_s e^{-\beta\varepsilon_s}\Big) = \sum_r e^{-2\beta\varepsilon_r} + \sum_r\sum_{s\atop r\neq s} e^{-\beta(\varepsilon_r+\varepsilon_s)} = \sum_r\sum_{s\atop r\neq s} e^{-\beta(\varepsilon_r+\varepsilon_s)}, \tag{4.27}$$

考虑到 $\langle n_r\rangle \ll 1$, $\sum_r e^{-2\beta\varepsilon_r}$ 已经忽略. 比较(4.26)和(4.27)式,有

$$Z = Z_1^2/2!, \tag{4.28}$$

推广到由 N 个粒于组成的 M－B 气体,于是有

$$Z_{M\text{-}B} = Z_1^N/N!. \tag{4.29}$$

M-B 分布首先是由玻尔兹曼导出的,在他的推导中,近独立粒子被认为是服从经典力学的可识别的粒子,因此往往叫 M-B 分布为经典统计. 玻尔兹曼计算出的配分函数为 $Z = Z_1^N$,由此求得的理想气体的熵不满足广延性的要求,产生著名的吉布斯佯谬. 为了克服这一困难,吉布斯在配分函数的表达式中硬加了一个因子 $(N!)^{-1}$. 从本节的讨论可以看到,事实上,玻尔兹曼气体和费米气体、玻色气体一样,由服从量子力学规律的全同粒子组成. 由于粒子的全同性,吉布斯佯谬自动消失.

作为本节的小结,将三种分布列表如表Ⅱ:

表Ⅱ　理想气体的三种分布

理想气体	分布函数	(巨)配分函数	μ 的取值	适用条件
费米气体	$\langle n_r\rangle = \dfrac{1}{e^{\beta(\varepsilon_r-\mu)}+1}$	$\tilde{Z} = \prod_r (1+e^{-\beta(\varepsilon_r-\mu)})$	$-\infty < \mu < \mu_0$	
玻色气体	$\langle n_r\rangle = \dfrac{1}{e^{\beta(\varepsilon_r-\mu)}-1}$	$\tilde{Z} = \prod_r (1-e^{-\beta(\varepsilon_r-\mu)})^{-1}$	$\mu \leqslant 0$	
玻尔兹曼气体	$\langle n_r\rangle = e^{\beta(\mu-\varepsilon_r)} = N\dfrac{e^{-\beta\varepsilon_r}}{\sum_r e^{-\beta\varepsilon_r}}$	$Z = \dfrac{1}{N!}z_t^N = \dfrac{1}{N!}\Big[\sum_r e^{-\beta\varepsilon_r}\Big]^N$	$\mu < 0$	$\dfrac{V}{N}\Big(\dfrac{\sqrt{2\pi m k_B T}}{h}\Big)^3 \gg 1$

除系综理论外,常用的统计方法还有只适用于近独立子系的最概然统计法. 用

最概然统计法推导三种理想气体的分布见本章习题12.

§7.3.5 三种系综的等价 其他系综

前面我们讨论了三种统计系综,即微正则、正则和巨正则系综.微正则系综描述孤立系统的平衡性质;正则系综描述与大热源平衡的恒温系统的性质;巨正则系综描述与大热源、大粒子源平衡的开放系统的性质.在这节中,我们讨论三种统计系综之间的关系,以及它们和热力学的关系.简单地说,三种统计系综的关系是:它们是等价的,但应用的广泛程度不同,方便应用的条件不同.统计物理中采用不同的系综和热力学中进行勒让德变换,引入不同热力学势或特性函数相对应.

(一)三种系综的等价

三种系综等价的含义是:虽然组成三种系综的系统所处的宏观条件有原则上的区别,但在热力学极限下用三种系综计算同一个宏观系统(临界点除外)的热力学量时,却会得到相同的结果.换句话说,我们可以不管系统所处的实际环境,按照方便,采用任何一种系综进行计算.§7.2.3 例2 和§7.3.2(一)分别将顺磁模型看成孤立系统和恒温系统;本章习题1 和§7.3.2(二)分别将爱因斯坦模型看成孤立系统和恒温系统,用微正则系综和正是系综得到相同结果,就是两个例子.为了加深印象,再看以下简单例子.

【例1】 考虑由 N 个全同定域粒子组成的系统,每个粒子可处于能量为 ε 和0的两个状态之一.(1)用微正则系综求当 n 个粒子处在 ε 状态时系统的熵 $S(n)$,及能量 $E=n\varepsilon$ 作为温度函数的表达式.(2)证明用正则系综和巨正则系综可得到相同的结果.

解 (1)用微正则系综求解:

由简单分析可知

$$\Omega(E,N,V)=\Omega(n,N,V)=N!/n!(N-n)!.$$

$$n=E/\varepsilon.$$

代入熵的玻尔兹曼公式:

$$\frac{S}{k_B}=\ln\Omega(n,N,V)=\ln\frac{N!}{n!(N-n)!}=N\ln\frac{N}{N-n}+n\ln\frac{N-n}{n}.$$

再代入§7.2.3(3.2)式

$$\beta=\left(\frac{\partial\ln\Omega}{\partial E}\right)_{N,V}=\frac{1}{\varepsilon}\left(\frac{\partial\ln\Omega}{\partial n}\right)_{N,V}=\frac{1}{\varepsilon}\ln\left(\frac{N}{n}-1\right),$$

或 $n=N(1+e^{\varepsilon/(k_BT)})^{-1}$.最后得

$$E=n\varepsilon=N\varepsilon/(1+e^{\varepsilon/(k_BT)}).$$

(2)用正则系综和巨正则系综求解　采用正则系综.配分函数是

$$Z=\sum_E \Omega(n,N,V)e^{-\beta E}=\sum_{n=0}^{N}\frac{N!}{n!(N-n)!}e^{-\beta n\varepsilon}=(1+e^{-\beta\varepsilon})^N=Z_1^N.$$

Z_1 是单粒子配分函数.

$$U=-(\partial\ln Z/\partial\beta)_{N,V}=-N(\partial\ln Z_1/\partial\beta)_V=N\varepsilon(1+e^{\beta\varepsilon})^{-1},$$

$$-\beta F=\ln Z=N\ln(1+e^{-\beta\varepsilon}),$$

$$S/k_B=\beta(U-F)=N\beta\varepsilon(1+e^{\beta\varepsilon})^{-1}+N\ln(1+e^{-\beta\varepsilon}).$$

将 $U=\langle n\rangle\varepsilon$ 和 $U=N\varepsilon(1+e^{\beta\varepsilon})^{-1}$ 比较可得 $\langle n\rangle=N(1+e^{\beta\varepsilon})^{-1}$.或

$$1+e^{\beta\varepsilon}=N/\langle n\rangle \quad 或 \quad \beta\varepsilon=\ln[(N-\langle n\rangle)/\langle n\rangle].$$

于是 S/k_B 进一步化为

$$S/k_B=n\ln[(N-\langle n\rangle)/\langle n\rangle]+N\ln[N/(N-\langle n\rangle)].$$

用正则系综和用微正则系综得到相同结果.

再用巨正则系综计算:

$$\tilde{Z}=\sum_N e^{-\beta\mu N}Z_N=\sum_N(zZ_1)^N=(1-zZ_1)^{-1},$$

其中 $z=e^{\beta\mu}=e^{\alpha}$.利用§7.3.3(3.20)和(3.17)式

$$U=-(\partial\ln\tilde{Z}/\partial\beta)_{V,\alpha}=-[zZ_1/(1-zZ_1)](\partial\ln Z_1/\partial\beta)_V$$

$$\langle N\rangle=k_BT\left(\frac{\partial\ln\tilde{Z}}{\partial\mu}\right)_T=zZ_1/(1-zZ_1).$$

再次得到

$$U=-\langle N\rangle(\partial\ln Z_1/\partial\beta)_V=N\varepsilon(1+e^{\beta\varepsilon})^{-1}.$$

(二)三种系综等价的解释　选用不同系综和进行勒让德变换的关系

三种系综等价的原因可以从物理和数学两方面来解释.下面仅以微正则和正则系综的等价来说明,分析的方法不难推广到分析正则和巨正则系综的等价关系.

前面已经计算出正则系综能量的相对涨落为

$$\delta E=\sqrt{\langle(\Delta E)^2\rangle/\langle E\rangle^2}\approx N^{-1/2}.$$

对 N 的数量级为 10^{23} 的宏观系统,能量的相对涨落极小.这说明,虽然由于可以和大热源自由交换能量,和孤立系统不同,恒温系统的能量可以取任意值.但实际上,偏离平均值 $\langle E\rangle$ 的概率微乎其微.换句话说,完全可以把恒温系统当成一个具有 $\langle E\rangle$ 值的孤立系统,而不会带来能观察到的偏差.即正则系综和微正则系综是等价的.

这一结论还可根据正则系综按能量分布的表达式来说明.恒温系统具有能量 E 的概率为

$$P(E)\sim\Omega(E,V,N)e^{-\beta E}. \tag{5.1}$$

这表示恒温系统处于各种能量状态均有一定的概率，而且概率 $P(E)$ 对 E 的依赖关系来源于两个因素：一个因素是 $e^{-\beta E}$，随系统能量 E 的增大而指数下降；另一因素是态密度 $\Omega(E,V,N)$，随 E 的增大而迅速增大，两者的乘积使得 $P(E)$ 在某一能量值 $\langle E\rangle$ 处有尖锐的极大值，如图 1 所示. 和孤立系统概率按能量的分布（图 2）相似.

我们再从数学变换的角度来分析. 对于宏观系统，能谱近似连续，正则系综的配分函数可写为

$$Z(T,V,N) = \int_0^\infty \Omega(E,V,N)\,e^{-\beta E}dE. \tag{5.2}$$

(5.2)式表明，$Z(T,V,N)$ 是 $\Omega(E,V,N)$ 的拉普拉斯变换. 数学中已经证明，拉普拉斯变换是唯一的. 根据这种唯一性，Z 和 Ω 包含着完全相同的信息，将给出相同的结果，即正则系综和微正则系综等价.

已经多次指出，直接应用微正则系综往往比较困难，原因是 $\Omega(E,V,N)$ 的计算必须满足能量守恒和粒子数守恒两个约束条件，以致并非对每一个系统，微正则系综的配分函数 $\Omega(E,V,N)$ 都能精确地求出. 而 $Z(T,V,N)$ 的计算则由于解除了对能量守恒的约束而变得比较容易. 换句话说，虽然理论上微正则和正则系综等价，但实际上，正则系综比微正则系综应用要广泛得多，也方便得多. 巨配分函数 $\tilde{Z}(T,V,\mu)$ 的计算则由于同时解除了能量守恒和粒子数守恒的约束而变得更加容易. 理想气体巨配分函数的计算就是典型例子.

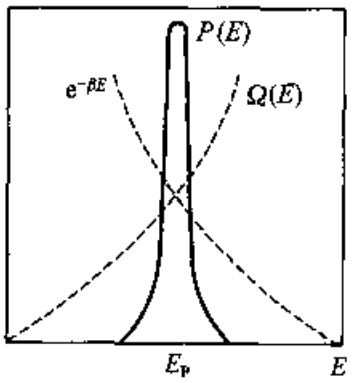


图 1　恒温系统概率按能量的分布

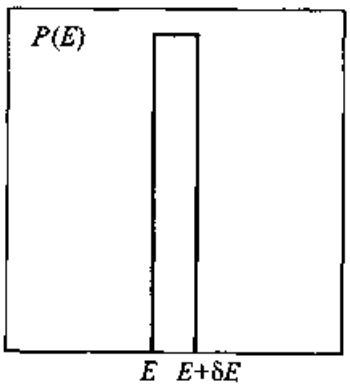


图 2　孤立系统概率按能量的分布

此外，从(5.2)式还可以看出，通过拉普拉斯变换，从微正则系综过渡到正则系综，不过是从以 (E,V,N) 为独立变量的配分函数 $\Omega(E,V,N)$ 变到以 (T,V,N) 为独立变量的配分函数 $Z(T,V,N)$. 即在统计物理中采用不同的系综，和在热力学中进行勒让德变换引入不同热力学势（或特性函数）相对应. 从热力学中已经看到，勒让德变换不会改变系统的性质，却可以简化它们的计算. 这正是既然系综等价却还要引入各种系综的原因.

总之，从理论角度考虑，微正则系综是系综理论的基础，正则分布和巨正则分布由微正则分布导出. 从应用考虑，则三种系综等价. 实际上，巨正则系综应用最

广，而不同系综的方便程度则还要具体问题具体分析.

*（三）T-p 系综

按照前面的分析，对具有三个独立变量的均匀系统，总计可以得到七种不同的系综. 要用统计的方法，推导出全部七种系综的配分函数（及对应的特性函数）及相应的热力学公式是非常麻烦的. 但是，根据从一种系综过渡到另一种系综，是进行拉普拉斯变换（或独立变量变换）的观点，很容易得到从微正则系综的配分函数及热力学公式出发，直接写出其他系综有关公式的简单规则. 下面介绍这种规则，并利用该规则写出一个较为有用的 T-p 系综的基本公式.

孤立系统的宏观条件是 (E,V,N) 一定. 微正则系综的配分函数 Ω 和特性函数 S 的独立变量是 (E,V,N)

$$S(E,V,N) = k_B \ln \Omega(E,V,N). \tag{5.3}$$

以 (E,V,N) 为独立变量的热力学基本方程是

$$\mathrm{d}S = (1/T)\mathrm{d}U + (p/T)\mathrm{d}V - (\mu/T)\mathrm{d}N.$$

或

$$\mathrm{d}U = T\mathrm{d}S - p\mathrm{d}V + \mu\mathrm{d}N. \tag{5.4}$$

恒温系统的宏观条件是 (T,V,N) 一定. 正则系综的配分函数和特性函数的独立变量是 (T,V,N)，并通过对微正则系综的配分函数进行拉普拉斯变换得到：

$$Z(T,V,N) = \int_0^\infty \Omega(E,V,N)\mathrm{e}^{-\beta E}\mathrm{d}E, \tag{5.5}$$

$$F(T,V,N) = -k_B T \ln Z(T,V,N). \tag{5.6}$$

以 (T,V,N) 为独立变量的热力学基本方程，可利用勒让德变换交换(5.4)中的 T 和 S 的地位得到

$$\mathrm{d}F = -S\mathrm{d}T - p\mathrm{d}V + \mu\mathrm{d}N. \tag{5.7}$$

开放系统的宏观条件是 (T,V,μ) 一定. 巨正则系综的配分函数和特性函数的独立变量是 (T,V,μ)，并可通过对正则系综的配分函数进行拉普拉斯变换得到

$$\bar{Z}(T,V,\mu) = \int_0^\infty Z(T,V,N)\mathrm{e}^{\beta\mu N}\mathrm{d}N = \iint \Omega(E,V,N)\mathrm{e}^{-\beta E_N+\beta\mu N}\mathrm{d}E_N\mathrm{d}N, \tag{5.8}$$

$$\Omega_{巨}(T,V,\mu) = -k_B T \ln \bar{Z}(T,V,\mu).$$

相应的热力学基本方程可利用勒让德变换，同时交换(5.4)式中的 T 和 S，μ 和 N 的地位得到：

$$\mathrm{d}\Omega_{巨} = \mathrm{d}(U - TS - \mu N) = -S\mathrm{d}T - p\mathrm{d}V - N\mathrm{d}\mu. \tag{5.9}$$

$T-p$ 系统的宏观条件是 (T,p,N) 一定，E 和 V 都是可变的. 按照以上规则，T-p 系综的配分函数和特性函数的独立变量是 (T,p,N)，而且可以通过以下积分变换得到[对照(5.8)式].

$$Z_{T-p}(T,p,N) = \iint \Omega(E,V,N)\mathrm{e}^{-\beta E_N-\beta pV}\mathrm{d}E_N\mathrm{d}V, \tag{5.10}$$

$$\Omega_{T-p}(T,p,N) = -k_B T \ln Z_{T-p}(T,p,N). \tag{5.11}$$

对应的热力学基本方程可利用勒让德变换，同时交换(5.4)式中的 T 和 S，p 和 V 的地位得到：

$$\mathrm{d}\Omega_{T-p} = \mathrm{d}(U - TS + pV) = -S\mathrm{d}T + V\mathrm{d}p + \mu\mathrm{d}N. \tag{5.12}$$

显然这里写出的 Ω_{T-p} 就是吉布斯自由焓 G. 于是，若求得 $T-p$ 系综的配分函数 Z_{T-p}，很容

易写出各种热力学量的公式：

$$S = -(\partial G/\partial T)_{p,N} = -(\partial k_B T \ln Z_{T-p}/\partial T)_{p,N}, \tag{5.13}$$

$$V = (\partial G/\partial p)_{T,N} = -k_B T(\partial \ln Z_{T-p}/\partial p)_{T,N}, \tag{5.14}$$

$$\mu = (\partial G/\partial N)_{T,p} = -k_B T(\partial \ln Z_{T-p}/\partial N)_{T,p}. \tag{5.15}$$

习　题

1. 有 N 个近独立的频率为 ν 的谐振子组成的系统，每个谐振子的能量为 $\varepsilon = (n + 1/2)h\nu$. (a)求当系统能量为 $E = Nh\nu/2 + Mh\nu$ 时的微观态数和熵. (b)求当系统达到平衡时，此系统能量与温度的关系，并和 §7.3.2 中用正则分布所得结果比较.

2. 设有 N 个独立可识别粒子组成的系统，每个粒子有两个可能状态：一个能量为 0 的状态，一个能量为 ε 的状态. 求系统的配分函数、内能和热容量，并证明当温度很高 $k_B T \gg \varepsilon$ 时，$C_V \propto T^{-2}$；当温度很低 $k_B T \ll \varepsilon$ 时，$C_V \to 0$.

3. 一块晶体包含 N 个原子，原子的自旋磁矩为 μ_0，被置于均匀磁场 H 中. 这些原子可以有三个取向：平行、垂直和反平行于磁场. 求：(a)晶体的配分函数. (b)晶体的磁矩. (c)高温弱场和低温强场下的磁矩. (d)求原子磁矩平行、反平行和垂直于外场时的概率，并由此求晶体的磁矩. (不考虑原子磁矩之间的相互作用.)

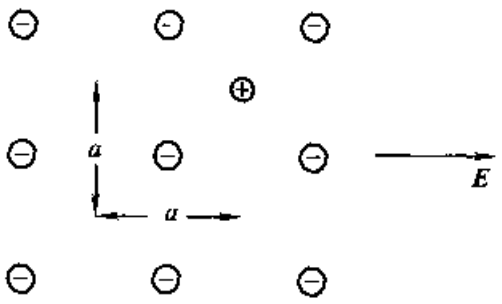


图 1　习题 4 图

4. 绝对温度为 T 的固体，每单位体积内有 N_0 个带负电荷的杂质离子. 这些离子代替了固体中若干正常原子. 整个固体当然是中性的，这是因为每个具有 $-e$ 的负离子，在其邻近都有一个电荷为 e 的正离子. 正离子很小，因而能自由地在晶格间移动. 在没有外场的情况下，它以相等的概率处于静止的负离子周围四个等距离的位置上(见图 1). 如果沿 x 方向加一电场 E，试求极化强度，即单位体积内沿 x 方向的平均偶极矩.

5. 考虑一个由氢原子组成的点阵，假定每个氢原子可处于以下四种状态，求每个原子中的平均电子数为 1 的条件：

状态	电子数	能量
基态	1	$-\Delta/2$
正离子	0	$-\delta/2$
负离子	2	$\delta/2$
激发态	1	$\Delta/2$

6. 假定系统由 N_0 个相同的独立格点对组成(线性聚合体中可能有这种情况)，每一格点对包括 A 格点和 B 格点. A、B 格点分别最多只能被一个分子占有. A 格点上吸附一个分子时，其能量为 $-\varepsilon_A$；B 格点上吸附一个分子时，其能量为 $-\varepsilon_B$. A,B 格点都吸附一个分子时，还有相互作

用势能 u. 不同格点对的分子间没有相互作用,求吸附率 $\theta(T,\mu) = \langle N\rangle/N_0$.

7. 固体表面由 N_0 个相同的,近独立格点组成,每个格点可吸附 s 个分子,$s = 0,1,2,\cdots,m$. 当格点吸附 s 个分子时,格点的能量为 $-s\varepsilon_0$. 用 $a_0,a_1,a_2,\cdots,a_m$ 分别代表吸附了零个分子,一个分子,…,和 m 个分子的格点数. 显然有:

$$\sum_{s=0}^{m} a_s = N_0, \qquad \sum_{s=0}^{m} a_s s = N.$$

N 为吸附的总分子数. 求系统的配分函数,巨配分函数和吸附率. 通过此题,你对正则系综和巨正则系综的应用有何体会?

8. 体积 V 中含有 N 个单原子分子组成的 M－B 理想气体,试由巨正则分布证明,在一小体积 v 中有 n 个分子的概率服从泊松分布,即

$$P_n = \frac{1}{n!}\langle n\rangle^n \mathrm{e}^{-\langle n\rangle}.$$

9. 证明巨正则系综的能量涨落为 $\langle(\Delta E)^2\rangle = k_B T^2\left(\frac{\partial\langle E\rangle}{\partial T}\right)_{\mu,T,V} = \langle(\Delta E)^2\rangle_{\text{正则}} + \langle(\Delta N)^2\rangle(\partial U/\partial N)_{T,V}$.

10. 为了计算石墨的比热采取以下简化模型. 石墨具有高度各向异性的晶格层结构. 每个碳原子作三维简谐振动,在平行于层的方向恢复力很大,因此在 x,y 方向(层的平面内)本征频率均等于 $\omega_{/\!/}$,且很大,有 $\hbar\omega_{/\!/} \gg 300K k_B$. 另一方面,垂直于层的恢复力很小,因此在 z 方向本征频率 $\omega_\perp$ 很小,有 $\hbar\omega_\perp \ll 300K k_B$. 按此模型,石墨在 300K 的摩尔热容量是多少?

11. 由 N 个近独立,定域粒子组成的系统,每个粒子可处于能量为 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_m$ 的任一状态,试计算系统的 Ω 和 Z. 通过此题,你对微正则系综和正则系综的应用有何体会?

12. 由大量全同的近独立粒子组成的理想气体形成一孤立系,具有确定的粒子数 N,能量 E 和体积 V. 用 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_i,\cdots$ 表示粒子能级,用 $G_1,G_2,\cdots,G_i,\cdots$ 表示能级的简并度,用 $N_1,N_2,\cdots,N_i,\cdots$ 表示分布在各能级上的粒子数,即:

能级:	ε_1,	ε_2,	…,	ε_i,	…
简并度:	G_1,	G_2,	…,	G_i,	…
粒子数:	N_1,	N_2,	…,	N_i,	…

数列 $\{N_1,N_2,\cdots,N_i,\cdots\}$ 或 $\{N_i\}$ 称为能级占有数集合或分布. 显然分布 $\{N_i\}$ 必须满足以下两个约束条件:

$$\sum_i N_i = N, \qquad \sum_i N_i\varepsilon_i = E. \tag{1}$$

约束条件只有两个,而 N 很大,故满足约束条件的分布有许多许多.

每个能级有 G_i 个简并态,给定一种分布,只确定了每个能级上的粒子数 N_i,而这 N_i 个粒子在 G_i 个单粒子态上还可以有各种占据方式. N 个粒子在单粒子态上的一种占据方式是系统的一个微观态,所以一种分布包含着许多微观态.

(a)试证明,对于费米气体和玻色气体,与一种分布 $\{N_i\}$ 对应的微观态数分别为

$$\Omega_{F-D}(\{N_i\}) = \prod_i \frac{G_i}{N_i!(G_i - N_i)!}, \tag{2}$$

$$\Omega_{B-E}(\{N_i\}) = \prod_i \frac{(G_i + N_i - 1)!}{N_i!(G_i - 1)!}. \tag{3}$$

(b)证明当对所有能级都有 $\langle n_i \rangle = N_i/G_i \ll 1$ 时，Ω_{F-D} 和 Ω_{B-E} 趋于同一近似表达式(相应于麦-玻气体)：

$$\Omega_{M-B}(\{N_i\}) = \prod_i G_i^{N_i}/N_i!. \tag{4}$$

(c)根据等概率原理，对处于平衡态的孤立系统，每一微观态等概率. 因而所含微观态数最多的分布为最概然分布. 据此，并考虑约束条件(1)，再利用待定乘子法，证明，最概然 M-B 分布为

$$\tilde{N}_i = G_i e^{-\alpha - \beta \varepsilon_i}. \tag{5}$$

(d)利用归一化条件确定 α，进而证明(5)式化为

$$\bar{n}_i = \frac{\tilde{N}_i}{G_i} = \frac{N}{\sum e^{-\beta \varepsilon_i}} e^{-\beta \varepsilon_i} = \frac{N}{z} e^{-\beta \varepsilon_i}. \tag{6}$$

13. 设有 N 个近独立、可识别粒子组成的系统，每个粒子有两个可能状态：一个能量为 ε 的状态和一个能量为 0 的状态.

(a)求粒子的配分函数 Z_1 和处于基态及激发态上的粒子数 N_0 和 N_1，用 math 绘出 N_0/N，N_1/N 对 $k_B T/\varepsilon$ 的曲线，结合所绘曲线讨论高、低温极限下的 N_0 和 N_1.

(b)求系统的 U 和 C_V，用 math 绘出 U 和 C_V 对 $k_B T/\varepsilon$ 的依赖曲线，并利用(a)的结果对此曲线进行解释.

(c)求系统的熵. 用 math 绘出 S/N 对 $k_B T/\varepsilon$ 的依赖曲线，结合曲线讨论高、低温极限下的熵.

14. 有一磁性样品被置于外磁场 H 中，样品由 N 个处于晶格格点，具有固有磁矩 μ_0 的原子组成，每个原子可平行或反平行于外场取向，在两个取向上，原子的能量分别为

$$\varepsilon = \mp [\mu_0 H + k_B \Theta M/(\mu_0 N)]. \tag{7}$$

Θ 是样品的特征温度，M 是样品的总磁矩.

(a)试用系综平均和统计热力学两种方法证明，样品的磁矩可由下式表示：

$$M/(\mu_0 N) = \text{th}\{(k_B T)^{-1}[\mu_0 H + k_B \Theta M/(\mu_0 N)]\}. \tag{8}$$

(b)用 math 绘出当外场为 0 时，$M/(\mu_0 N)$ 对 T/Θ 的依赖曲线(包括 $T > \Theta$ 和 $T < \Theta$).

(c)根据所绘曲线说明方程(7)代表什么磁性材料？你对(7)中 $k_B \Theta M/(\mu_0 N)$ 项的起源怎样理解？

(d)若 $M \ll \mu_0 N$，证明每个原子的磁化率为

$$\chi \equiv M/NH \approx \mu_0^2/[k_B(T - \Theta)], \tag{9}$$

即服从居里-外斯定理.

15. (a)实验测得铜在 60K 和 160K 时的摩尔热容量分别为 2cal · mol^{-1} · K^{-1} 和 5cal · mol^{-1} · K^{-1}. 试根据爱因斯坦模型用计算机进行计算，求与以上数据最佳一致的原子振动频率.

(b)试对银作相同计算，已知银在 37K 和 118K 时的摩尔热容量分别为 2cal · mol^{-1} · K^{-1} 和 5cal · mol^{-1} · K^{-1}.

16. 根据爱因斯坦模型，所有固体晶体振动的摩尔热容量 c_v 是 Θ_E/T 的相同函数

$$c_V = 3R\left(\frac{\Theta_E}{T}\right)^2 \frac{e^{\Theta_E/T}}{(e^{\Theta_E/T}-1)^2} \tag{10}$$

(a)试用计算机按(5.10)式进行数值计算填充下表

Θ_E/T	∞	10.00	5.00	2.00	1.25	0.50	0
T/Θ_E	0	0.10	0.20	0.50	0.80	2.00	∞
c_v	0						

(b)用计算机按(5.10)式绘出 $c_v - T/\Theta_E$ 曲线.

(c)已知银(Ag)在 103.14K 时的 c_v 为 20.09 $J \cdot mol^{-1} \cdot K^{-1}$，试利用(b)中的曲线或(a)中的表确定银的 Θ_E（约为 168K).

(d)将所得 Θ_E 代入(5.10)式，计算下表所列各种温度下的 c_v，并和实验值比较.

T(K)	20	30	40	50	100	200	300
c_V（计算值）							
c_v（实验值）	1.61	4.81	8.49	11.80	19.87	23.39	24.19



第八篇　理想气体

系综理论是适用于任何宏观系统的普遍理论，然而在大多数情况下，推导会遇到严重困难，而不得不去研究实际系统的理想模型. 通常，第一步是建立高度理想化的模型，然后逐步考虑更复杂的因素.

理想气体，即分子之间除碰撞以外完全没有相互作用的系统，就是一种高度理想化的模型. 对理想气体的研究不仅能帮助获得处理问题的数学技巧，而且也为建立实际气体理论打下基础. 特别是，正如前面多次指出的，在一定条件下，相互作用很强的粒子组成的系统可以化为“准粒子”组成的，没有相互作用的理想气体来处理，例如见 §1.4.3、§6.1.5、§ 8.2.2(一)和 § 8.2.3(三).

在第七篇第三章的第四节中，我们已导出了三种理想气体的分布. 本篇讨论这些分布的应用. 第一章是 M－B 分布的应用；第二章和第三章是 B－E 分布和 F－D 分布的应用.

第一章　玻尔兹曼气体

在§7.3.4中已经用数字估算表明，玻尔兹曼分布有广泛应用.应用玻尔兹曼分布大致包括两类问题.第一，通过计算配分函数求玻尔兹曼气体的热力学量；第二，从粒子按能量的分布求粒子按广义坐标和广义动量的分布.在第一、二节中，我们将以单原子理想气体和双原子分子理想气体为例介绍解决第一类问题的基本方法和步骤.在第三节以推导麦克斯韦速度分布等为例说明如何解决第二类问题.在第四节我们导出能量均分定理，该定理在高温极限下，对计算某些系统的热容量极为方便.最后，在第五节里简单地讨论分子之间有相互作用的稀薄实际气体.

§8.1.1　由配分函数求热力学量

系综理论告诉我们，只要求得系统的配分函数，就可以计算全部热力学量.§7.3.4(4.29)式已经给出 $Z_{M-B}=Z_1^N/N!$.可见，为求玻尔兹曼气体的热力学量，首先必须计算分子的配分函数 Z_1.

(一)分子配分函数的分解

按定义，分子的配分函数为

$$Z_1=\sum_{\varepsilon_k} g_k e^{-\beta\varepsilon_k}. \tag{1.1}$$

气体分子的能量往往可分为几部分——质心运动的动能和在外场中的势能，分子转动能及分子中原子间相对振动的能量等.在一定的近似下，可以认为它们是独立的.例如，若忽略由于振动引起的分子转动惯量的改变(因为振动往往比转动快得多)，可以把分子的转动能量和振动能量看成是无关的.我们用 $\varepsilon^{平}$，$\varepsilon^{转}$ 和 $\varepsilon^{振}$ 分别表示分子的平动、转动和振动能量，用 $g^{平}$，$g^{转}$ 和 $g^{振}$ 分别表示这三种能量的简并度，则有

$$\varepsilon_k=\varepsilon^{平}+\varepsilon^{转}+\varepsilon^{振},\qquad g_k=g^{平}\times g^{转}\times g^{振}. \tag{1.2}$$

代入(1.1)式使分子的配分函数 Z_1 分解为三个因子的乘积：

$$Z_1=Z_1^{平}\times Z_1^{转}\times Z_1^{振}. \tag{1.3}$$

$Z_1^{平}$，$Z_1^{转}$ 和 $Z_1^{振}$ 依次为分子平动、转动和振动的配分函数，它们的表达式是：

$$Z_1^{平}=\sum g^{平}e^{-\beta\varepsilon^{平}},\quad Z_1^{转}=\sum g^{转}e^{-\beta\varepsilon^{转}},\quad Z_1^{振}=\sum g^{振}e^{-\beta\varepsilon^{振}}. \tag{1.4}$$

由此不难看出，玻尔兹曼气体的广延热力学量等于各自由度贡献之和.例如，据§7.3.1(1.12)式，玻尔兹曼气体的内能可表示为

$$U=N(\langle\varepsilon^{平}\rangle+\langle\varepsilon^{转}\rangle+\langle\varepsilon^{振}\rangle)$$

$$= Nk_B T^2 \{\partial \ln Z_1^{平}/\partial T + \partial \ln Z_1^{转}/\partial T + \partial \ln Z_1^{振}/\partial T\}. \tag{1.5}$$

其中

$$\langle \varepsilon^i \rangle = k_B T^2 \partial \ln Z_1^i / \partial T. \tag{1.6}$$

这里 i 代表平动、转动和振动自由度，$\langle \varepsilon^i \rangle$ 代表分子平动、转动和振动能量的平均值. 于是，求玻尔兹曼气体热力学量的问题最终归结为求与各种自由度对应的分子配分函数.

(二) μ 空间　分子配分函数在 μ 空间中的表示

为了说明如何计算分子的配分函数，我们不妨将各能级对分子配分函数的贡献作一简略估计. 据(1.4)式，分子任一自由度的配分函数可写为

$$Z_1^i = \sum g^i e^{-\beta \varepsilon^i} = g_0^i e^{-\beta \varepsilon_0^i} + g_1^i e^{-\beta \varepsilon_1^i} + g_2^i e^{-\beta \varepsilon_2^i} + \cdots.$$

取 $\varepsilon_0^i = 0$，则

$$Z_1^i = g_0^i + g_1^i e^{-\beta \Delta \varepsilon_1^i} + g_2^i e^{-\beta \Delta \varepsilon_2^i} + \cdots. \tag{1.7}$$

当能量间隔 $\Delta \varepsilon^i$ 远大于热运动能量典型值 $k_B T$，即 $\Delta \varepsilon^i \gg k_B T$ 时，以上求和中只需保留最前面的几项就足够精确了；而当 $\Delta \varepsilon^i \ll k_B T$，即分子能量可近似为连续变化时，求和可用积分代替，如 §7.3.4 中得到该节(4.23)式所用的方法.

事实上，当分子某一自由度的能量可以看成是连续变化时，普朗克常数的作用已很小，因而分子该自由度的运动可以用经典方法描述. 相应的配分函数的计算可以在 μ 空间中用积分的方法方便地进行.

设一个分子有 f 个自由度，则根据经典力学，确定分子的运动状态要给出它的 f 个广义坐标 $q_1, q_2, \cdots, q_f$ 及相应的 f 个广义动量 $p_1, p_2, \cdots, p_f$. 用 f 个广义坐标 q_i 和广义动量 p_i 为坐标轴所张成的 $2f$ 维空间称为μ 空间. μ 空间中的一点代表分子的一个运动状态. 近独立分子的能量决定于它的状态，即

$$\varepsilon = \varepsilon(q_1, q_2, \cdots, q_f; p_1, p_2, \cdots, p_f). \tag{1.8}$$

按照这样的描述方法，在 μ 空间中任何小区域内包含的分子状态有无穷多个，这是由于在经典过渡时取普朗克常数为零而引起的. 实际上，普朗克常数不论多小，总是个有限量，这对计算分子状态数要起作用. 根据不确定关系[§5.2.4(4.18)式]，分子的广义坐标 q_i 和广义动量 p_i 不能同时精确确定，其不确定度满足：

$$\Delta q_i \Delta p_i \geqslant \hbar/2. \tag{1.9}$$

也就是说，如果我们用广义坐标和广义动量来描述分子的运动，不能把它的状态确定得比 $\hbar/2$ 更精确. 因此，不能用 μ 空间中的一点来描述分子的运动状态. 为了满足不确定关系的要求，我们作如下的修改：将 μ 空间划分为大小相等的相格，每个

相格的相体积为 h^f ❶. 于是每个相格代表分子的一个状态. 在 μ 空间中一个宏观无限小的相体积元 $d\tau = dq_1dq_2\cdots dq_fdp_1dp_2\cdots dp_f$ 中所包含的分子状态数为

$$d\tau/h^f = dq_1dq_2\cdots dq_fdp_1dp_2\cdots dp_f/h^f. \tag{1.10}$$

分子的状态落在某一相格中，分子就有一定的能量. 我们用 $g(\varepsilon)d\varepsilon$ 表示分子的能量落在 $\varepsilon \sim \varepsilon + d\varepsilon$ 之间的状态数，它可以表示为

$$g(\varepsilon)d\varepsilon = \int_{\varepsilon\sim\varepsilon+d\varepsilon}\frac{dq_1dq_2\cdots dq_fdp_1dp_2\cdots dp_f}{h^f}. \tag{1.11}$$

积分号下的 $\varepsilon \sim \varepsilon + d\varepsilon$ 表示积分区间. $g(\varepsilon)$ 代表单位能量间隔中的分子状态数，称为分子态密度. 显然 $g(\varepsilon)d\varepsilon$ 与能级简并度 g 相对应. 于是，对于具有连续变化能量的自由度，分子配分函数可在 μ 空间中用积分计算：

$$Z_1 = \int e^{-\beta\varepsilon}g(\varepsilon)d\varepsilon = \int e^{-\beta\varepsilon}\frac{dq_1dq_2\cdots dq_fdp_1dp_2\cdots dp_f}{h^f}. \tag{1.12}$$

要特别强调的是，对于分子配分函数的计算，比值 $\Delta\varepsilon'/(k_BT)$ 是一个重要的参考量. 只有当此比值远小于 1 时，才可以利用(1.12)式计算分子的配分函数.

一般说来，对于不同的自由度，能量间隔在数量级上是不同的，甚至相差很大. 所以，在同一温度下，有的自由度可以用经典方法处理，有的自由度就不能这样做；同一个自由度，在高温时，能够用经典方法处理，而在低温时又不能用经典方法处理. 下节用双原子分子为例来具体说明.

在 §7.3.4 中已经看到，即使在极低温度下，在宏观尺度容器中，理想气体分子的平动能量总可以看成是连续的，因而平动配分函数的求和总可以用 μ 空间的积分代替.

对于转动自由度，根据量子力学(见第五篇第四章习题 9)，能级和对应的简并度是

$$\varepsilon^{转} = l(l+1)h^2/(8\pi^2I), \quad g^{转} = 2l+1, \quad l = 0,1,2,\cdots. \tag{1.13}$$

$I = m\langle r^2\rangle$ 是分子的转动惯量，m 是约化质量，$\langle r^2\rangle$ 是两原子间距离平方的平均值 $\sqrt{\langle r^2\rangle} \approx 10^{-8}$ cm. 以氢分子为例，$m \approx 10^{-24}$ g，因此能级间隔 $\Delta\varepsilon^{转}$ 约为 $h^2/(8\pi^2I) \approx 10^{-14}$ erg，与室温下 k_BT 有相同数量级，必须考虑量子效应；对于较重的分子，I 变大，量子效应减弱，在室温下往往可用经典方法处理.

对于振动自由度，能级和简并度是

$$\varepsilon^{振} = (n+1/2)h\nu, \quad g^{振} = 1, \quad n = 0,1,2,\cdots. \tag{1.14}$$

能级间隔 $\Delta\varepsilon^{振} = h\nu$. 不同的双原子分子的频率 ν 不同，但数量级相同. 例如，氢分

---

❶ 这相当于每个自由度的相格大小为 h，和玻尔量子化条件 §5.1.2(2.3)式一致，参看 §7.1.3(三)的注.

子 $h\nu \approx 10^{-12}$ erg，比室温下 $k_B T$ 还要高两个数量级，因而量子效应显著，求配分函数不能用 μ 空间中的积分代替求和，只有在非常高的温度下才可用经典方法处理.

最后还要强调指出的是，μ 空间只适用于近独立子系. 因为当粒子间有相互作用时，则粒子的能量不仅依赖于自己的广义坐标和广义动量，还是其他粒子坐标和动量的函数.

【例 1】 在 μ 空间中求分子平动配分函数.

解 描述分子平动的广义坐标和广义动量是 x,y,z 和 p_x,p_y,p_z. 相应的 μ 空间是 6 维的. 在无外场的情况下，平动能量的经典表达式为

$$\varepsilon = p^2/(2m) = (p_x^2 + p_y^2 + p_z^2)/(2m). \tag{1.15}$$

$$g(\varepsilon)\mathrm{d}\varepsilon = \frac{1}{h^3}\int_{\varepsilon-\varepsilon+\mathrm{d}\varepsilon}\cdots\int \mathrm{d}x\mathrm{d}y\mathrm{d}z\mathrm{d}p_x\mathrm{d}p_y\mathrm{d}p_z$$

$$= \frac{1}{h^3}\int_0^L\int_0^L\int_0^L \mathrm{d}x\mathrm{d}y\mathrm{d}z\int_{p-p+\mathrm{d}p}\cdots\int \mathrm{d}p_x\mathrm{d}p_y\mathrm{d}p_z = h^{-3}V4\pi p^2\mathrm{d}p. \tag{1.16}$$

将(1.15)式代入(1.16)式可见，对于非相对论自由运动粒子有

$$g(\varepsilon)\mathrm{d}\varepsilon = 2\pi V h^{-3}(2m)^{3/2}\varepsilon^{1/2}\mathrm{d}\varepsilon. \tag{1.17}$$

由(1.17)式可得

$$Z_1^{平} = \int_0^\infty \mathrm{e}^{-\beta\varepsilon}g(\varepsilon)\mathrm{d}\varepsilon = \frac{1}{h^3}2\pi V(2m)^{3/2}\int_0^\infty \mathrm{e}^{-\beta\varepsilon}\varepsilon^{1/2}\mathrm{d}\varepsilon = V\left(\frac{2\pi mkT}{h^2}\right)^{3/2}. \tag{1.18}$$

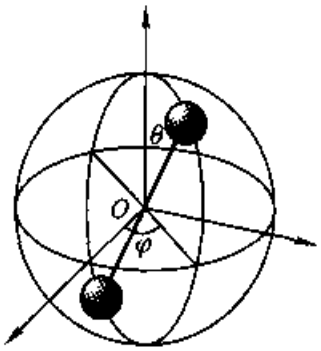


图 1 刚性转子

[积分见附录(Ⅴ.6)式]. 这一结果和 §7.3.4(4.23)式完全一样，说明对能谱连续的自由度，用 μ 空间计算分子配分函数和用量子力学数空间计算结果一样. 因此，今后对在宏观尺度空间中运动的自由粒子，我们总是利用其在 μ 空间中的态密度表达式(1.16)和(1.17)式来求系统的各种宏观量.

当然，将(1.15)式直接代入(1.12)式，会得到相同结果：

$$Z_1^{平} = \frac{1}{h^3}\iiint_V \mathrm{d}x\mathrm{d}y\mathrm{d}z\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{1}{2mkT}p_x^2}\mathrm{d}p_x\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{1}{2mkT}p_y^2}\mathrm{d}p_y \times$$

$$\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{1}{2mkT}p_z^2}\mathrm{d}p_z = V(2\pi mk_BT/h^2)^{3/2}.$$

【例 2】 在 μ 空间中求刚性转子的配分函数.

解 描述刚性转子的广义坐标和广义动量是 θ,φ 和 p_θ,p_φ，其中 θ,φ 是球坐标中相对于转子轴的两个方向角，如图 1 所示. μ 空间为 4 维. 刚性转子能量的经典表达式为

$$\varepsilon^{转} \approx [p_\theta^2 + p_\varphi^2/\sin^2\theta]/(2I). \tag{1.19}$$

代入(1.12)式,得到

$$\begin{aligned} Z_1^{转} &= \frac{1}{h^2}\int_0^{2\pi}\int_0^{\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2IkT}(p_\theta^2+p_\varphi^2/\sin^2\theta)}\mathrm{d}\theta\mathrm{d}\varphi\mathrm{d}p_\theta\mathrm{d}p_\varphi \\ &= \frac{1}{h^2}\int_{-\infty}^{\infty} e^{-p_\theta^2/(2Ik_BT)}\mathrm{d}p_\theta\int_0^{2\pi}\mathrm{d}\varphi\left[\int_0^{\pi}\mathrm{d}\theta\int_{-\infty}^{\infty} e^{-p_\varphi^2/(2Ik_BT\sin^2\theta)}\mathrm{d}p_\varphi\right] \\ &= \frac{2\pi}{h^2}\sqrt{2\pi Ik_BT}\int_0^{\pi}\sqrt{2\pi Ik_BT}\,\sin\theta\cdot\mathrm{d}\theta \\ &= 8\pi^2 Ik_BT/h^2. \end{aligned} \tag{1.20}$$

现在,我们回到本节开始提出的问题,通过计算配分函数求玻尔兹曼气体的热力学量.最简单的例子是单原子分子理想气体.如惰性气体和高温下金属原子组成的气体.

(三)单原子理想气体的热力学量

单原子理想气体的分子只有平动,而平动在任何温度下都可用经典方法处理.所以,根据(1.18)式,单原子理想气体的配分函数为:

$$Z = (1/N!)Z_{平}^N = (1/N!)[V(2\pi mk_BT/h^2)^{3/2}]^N. \tag{1.21}$$

代入§7.3.1(1.15)式,可求得理想气体的压强,

$$p = k_BT(\partial\ln Z/\partial V) = Nk_BT/V. \tag{1.22}$$

玻尔兹曼常数的值就是通过上式与实验测得的理想气体状态方程比较而确定的.

利用§7.3.1(1.12)式和(1.16)式,单原子理想气体的内能和熵为:

$$U = k_BT^2\partial\ln Z/\partial T = 3Nk_BT/2, \tag{1.23}$$

$$\begin{aligned} S &= k_B[\ln Z + \beta U] \\ &= Nk_B\left\{\frac{3}{2}\ln T + \ln\frac{V}{N} + \frac{5}{2} + \frac{3}{2}\ln\frac{2\pi mk_B}{h^2}\right\}. \end{aligned} \tag{1.24}$$

利用(1.22)~(1.24)式中的三个基本热力学量,可以求得全部其他热力学函数.例如定体热容量

$$C_V = (\partial U/\partial T)_V = 3Nk_B/2. \tag{1.25}$$

若气体的物质的量是1mol,则有

$$c_V \approx 3N_Ak_B/2 = 3R/2, \qquad c_p = c_V + R = 5R/2, \tag{1.26}$$

和

$$\gamma = c_p/c_V = 1.667. \tag{1.27}$$

与表Ⅰ所列实验数据比较,可以看出理论与实验很好地符合.这说明把分子平动能量视为连续变化,采用经典描述是恰当的.

表 I 单原子气体的比热比

气 体	温 度(K)	$\gamma = c_p/c_V$	气 体	温 度(K)	$\gamma = c_p/c_V$
氦(He)	291	1.660	氪(Kr)	292	1.689
	93	1.673	氙(Xe)	292	1.686
氖(Re)	292	1.642	钠(Na)	750 ~ 920	1.68
氩(Ar)	288	1.65	钾(K)	660 ~ 1000	1.64
	93	1.69	汞(Hg)	548 ~ 629	1.666

§8.1.2 双原子分子气体的热容量

本节计算双原子分子理想气体的配分函数和热力学量.

上节已经指出,双原子分子理想气体的内能等于平动、转动和振动自由度贡献之和,因而摩尔热容量也等于各自由度贡献之和,即

$$c_V = c_V^{平} + c_V^{转} + c_V^{振}. \tag{2.1}$$

$$c_V^{i} = N_A(\partial\langle\varepsilon^{i}\rangle/\partial T). \quad (i = 平,转,振) \tag{2.2}$$

平动自由度的贡献和单原子理想气体的热容量相同,已经讨论过了.下面来计算转动和振动的配分函数和热容量.

(一)转动和振动的配分函数和热容量

据上节的(1.4)和(1.13)式,双原子分子的转动配分函数为

$$Z_1^{转} = \sum_{l=0}^{\infty} e^{-l(l+1)h^2/(8\pi^2 I k_B T)}(2l+1). \tag{2.3}$$

定义转动特征温度为

$$\Theta^{转} = h^2/(8\pi^2 I k_B). \tag{2.4}$$

它由气体性质决定,可通过远红外区分子光谱实验测得.于是

$$Z_1^{转} = \sum_{l=0}^{\infty}(2l+1)e^{-l(l+1)\Theta^{转}/T}. \tag{2.5}$$

下面分别考虑高温极限和低温极限两种情况.

当 $T \gg \Theta^{转}$ 或 $k_B T \gg h^2/(8\pi^2 I)$ (即 $k_B T \gg \Delta\varepsilon^{转}$)时,能量近似连续变化,(2.5)式中的求和可以用积分代替

$$Z_1^{转} = \int_0^{\infty}(2l+1)e^{-l(l+1)\Theta^{转}/T}dl = T/\Theta^{转} = 8\pi^2 I k_B T/h^2. \tag{2.6}$$

这与上节例 2 中用经典方法处理所得结果完全一样.再将上式代入 §8.1.1(1.6)式,得

$$\langle \varepsilon^{转} \rangle = k_B T^2 (\partial \ln Z_1^{转}/\partial T) = k_B T. \tag{2.7}$$

再代入(2.2)式,则在高温极限下,摩尔转动热容量为

$$c_V^{转} = N_A (\partial \langle \varepsilon^{转} \rangle / \partial T) = N_0 k_B = R. \tag{2.8}$$

当 $T \ll \Theta^{转}$ 时或 $k_B T \ll h^2/(8\pi^2 I)$ (即 $k_B T \ll \Delta\varepsilon^{转}$)时,能量不能看成连续变化,求和不能用积分代替.但作为一级近似,只需保留最前面的两项,故

$$Z_1^{转} \approx 1 + 3e^{-2\Theta^{转}/T}. \tag{2.9}$$

其中第二项是小量,所以

$$\ln Z_1^{转} = \ln (1 + 3e^{-2\Theta^{转}/T}) \approx 3e^{-2\Theta^{转}/T}. \tag{2.10}$$

由此得到低温极限下,分子转动的平均能量和摩尔转动热容量为

$$\langle \varepsilon^{转} \rangle = k_B T^2 (\partial \ln Z_1^{转}/\partial T) = [3h^2/(4\pi^2 I)]\, e^{-h^2/(4\pi^2 I k_B T)}, \tag{2.11}$$

$$c_V^{转} = N_A (\partial \langle \varepsilon^{转} \rangle / \partial T) = 12R(\Theta^{转}/T)^2 e^{-2\Theta^{转}/T}. \tag{2.12}$$

在中间温度区,需要在(2.5)式求和中保留更多的项以保证足够的精度.这时,必须用数值计算.(2.12)式表明,在低温下转动自由度对热容量的贡献随温度降低面减小,最后趋向零,通常叫"转动冻结".

再来考虑双原子分子理想气体的振动热容量.因为按照我们的模型,双原子分子的振动和爱因斯坦模型中的简正振动有相同能谱,所以配分函数必有相同形式,即:

$$Z_1^{振} = \sum_{n=0}^{\infty} e^{-\beta(n+1/2)h\nu} = \frac{e^{-h\nu/2k_B T}}{1 - e^{-h\nu/(k_B T)}}. \tag{2.13}$$

代入(2.2)式,得

$$\begin{aligned} \langle \varepsilon^{振} \rangle &= h\nu/2 + h\nu/(e^{h\nu/(k_B T)} - 1), \\ c_V^{振} &= R[h\nu/(k_B T)^2][\, e^{h\nu/(k_B T)}/(e^{h\nu/(k_B T)} - 1)^2]. \end{aligned} \tag{2.14}$$

引入振动特征温度:

$$\Theta^{振} = h\nu/k_B. \tag{2.15}$$

不同气体有不同的 ν,因面有不同的振动特征温度,其值也可由分子光谱(红外区)实验给出.于是(2.14)式又可写为

$$c_V^{振} = R(\Theta^{振}/T)^2 e^{\Theta^{振}/T} / (e^{\Theta^{振}/T} - 1)^2. \tag{2.16}$$

我们看到,双原子分子理想气体的摩尔振动热容量是 $\Theta^{振}/T$ 的复杂函数.下面分别考虑高温极限和低温极限.

当温度很高,$T \gg \Theta^{振}$ 或 $k_B T \gg h\nu$ (即 $k_B T \gg \Delta\varepsilon^{振}$)时,

$$e^{\Theta^{振}/T} \approx 1 + \Theta^{振}/T.$$

于是

$$\langle \varepsilon^{振} \rangle = h\nu/2 + k_B T \approx k_B T,$$

$$c_V^{振} = R(\Theta^{振}/T)^2(1+\Theta^{振}/T)/(\Theta^{振}/T)^2 \approx R. \tag{2.17}$$

把双原子分子看成是经典谐振子,利用 μ 空间计算 $Z_1^{振}$,会得到与上面相同的结果(见习题4).

在相反的情况下,即 $T \ll \Theta^{转}$ 或 $k_B T \ll h\nu$(即 $k_B T \ll \Delta\varepsilon^{转}$)时,由(2.14)和(2.16)式得到

$$\langle \varepsilon^{振} \rangle = h\nu/2, \qquad c_V^{振} = R(\Theta^{振}/T)^2 e^{-\Theta^{振}/T}. \tag{2.18}$$

即当气体温度远低于振动特征温度时,平均振动能量趋于"零点能",振动对热容量的贡献也会随温度下降而指数减小,最后趋于零,发生"振动冻结".

值得注意的是,由(2.5),(2.13)式和上节的(1.18)式可知,$Z_1^{转}$、$Z_1^{振}$ 和 $Z_1^{平}$ 不同,前两者仅仅是 T 的函数和 V 无关.

(二)对 $c_V^{转}$ 和 $c_V^{振}$ 的讨论及与实验的比较

我们看到,转动自由度和振动自由度对定体热容量的贡献有共同特点,这就是,当气体温度远高于特征温度时,或热运动能量 $k_B T$ 远大于能级间隔,以致能量近似连续变化时,有 $\langle \varepsilon^{转} \rangle = \langle \varepsilon^{振} \rangle = k_B T$,$c_V^{转} = c_V^{振} = R$ 与经典能量均分定理相同(见 §8.1.4).我们说,这时自由度是充分激发的.

与此相反,当气体温度远低于特征温度时,$c_V^{转}$ 和 $c_V^{振}$ 都趋于零,会发生自由度"冻结".为了理解自由度"冻结"的实质,试计算分子振动处于第一激发态和基态的概率之比

$$\frac{P_1}{P_0} = \frac{e^{-\beta\varepsilon_1}/Z_1}{e^{-\beta\varepsilon_0}/Z_1} = e^{-(\varepsilon_1-\varepsilon_0)/(k_B T)} = e^{-h\nu/(k_B T)}.$$

当 $k_B T \ll h\nu$ 时,$P_1/P_0 \to 0$.例如对于 N_2,$h\nu \approx 0.3$ eV,$T = 300$ K 时,$k_B T = 0.025$ eV,因而 $P_1/P_0 \approx e^{-12} \to 0$.即当热运动能量远小于能量间隔 $h\nu$ 时,分子取得 $h\nu$ 的能量跃迁到激发态的概率是极小的.平均说来,几乎全部分子都"冻结"在振动的基态.当气体温度升高时,$\langle \varepsilon^{振} \rangle = h\nu/2$ 不变,或者说振动自由度不吸收热量,因而 $C_V^{振} = 0$.以上讨论说明,振动"冻结"的实质是分子停留在振动的基态.显然以上解释对转动自由度同样适合.

虽然转动自由度和振动自由度对热容量的贡献有以上共同特点,但它们也有明显的区别.我们知道,高温极限和低温极限都是相对于特征温度而言的.表Ⅰ列举了由光谱实验测得的一些双原子分子气体的 $\Theta^{转}$ 和 $\Theta^{振}$.可以看到,同一种气体的 $\Theta^{转}$ 和 $\Theta^{振}$ 相差很大.因此,对于转动自由度和振动自由度来说,高温和低温的概念有很大差别.室温对于振动自由度可看成低温,对于大多数气体的转动自由度应视为高温.在室温下,振动自由度已开始"冻结",而大多数气体的转动自由度却处于充分激发的状态.

表Ⅰ　一些双原子分子气体的特征温度

气体	H_2	D_2	HCl	N_2	O_2	Cl_2	I_2
$\Theta^{振}$(K)	6100		4140	3340	2230		
$\Theta^{转}$(K)	85.4	43	15.2	2.86	2.07	0.347	0.054

根据以上讨论可以预计，在室温下，大多数双原子分子气体应具有刚性转子模型的特点，即有 $c_v = 5R/2$，$c_p = 7R/2$，而 $\gamma = 7/5 = 1.4$．表Ⅱ是一些双原子分子气体 γ 的实验值．我们看到，理论计算和实验结果是符合的．

表Ⅱ　一些双原子分子气体的 γ

气　体	温　度(K)	γ	气　体	温　度(K)	γ
氢 H_2	289	1.407	CO	291	1.396
	197	1.453		93	1.417
	92	1.597	NO	228	1.38
氮 N_2	293	1.398		228	1.39
	92	1.419		193	1.38
氧 O_2	293	1.398	HCl	290～373	1.40
	197	1.411			
	92	1.404			

因为大多数双原子分子气体的转动特征温度低于临界温度，因此早在转动自由度“冻结”之前，气体已经液化，不可能观察到转动“冻结”现象．唯一例外是 H_2，它的转动特征温度是 85.4K，而临界温度是 33K．图 1 是氢(H)和它的同位素氘(D)氚(T)形成的双原子分子气体的 c_p 随 T 的变化曲线．由图看到，正如理论计算所预言的那样，当温度逐渐降低时，随着振动和转动自由度的依次“冻结”，曲线呈明显的阶梯形．

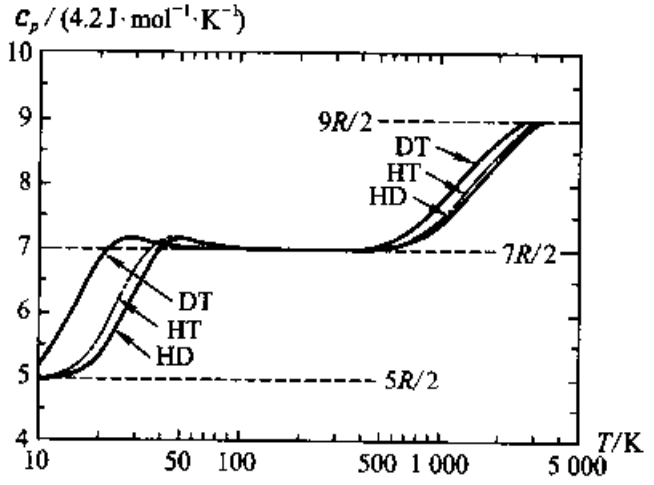


图 1　氢的热容量

最后要指出两点．第一，至今我们没有考虑原子或分子中电子和核对热容量的

贡献.原因如下:要电子从基态过渡到第一激发态,通常要供给大约 15×10^{-19}J 的能量,相应的温度约为 10^5K.核激发态和基态的能级差更大,对应温度约为 10^{10}K.因此,在通常温度下,原子或分子中的电子和核均处于基态,它们对分子配分函数的贡献仅为常数(基态的简并度),从而对热容量的贡献为零.第二,(2.6)式仅适用于由不同原子组成的双原子分子,对子由相同原子组成的双原子分子,当把分子绕垂直于原子核联线的轴转过 180 度时,得到物理上相同的状态.为了得到转动配分函数的正确表达式(每个物理状态只计算一次),应该把对转动角度的一切值积分得到的结果除以对称因子 2.总结两点说明,在温度不特别高的情况下,双原子分子内部配分函数可表为:

$$Z_1^{内}\approx g_e^0(2S_A+1)(2S_B+1))(T/\Theta)Z_1^{振}/\sigma_{AB}.$$

其中 g_e^0 为电子基态的简并度,如对惰性气体 $g_e^0=1$,对卤族金属原子气体 $g_e^0=2$;S_A,S_B 为核 A,B 的自旋量子数;σ_{AB} 为对称因子,当 $A\neq B$ 时,$\sigma_{AB}=1$,当 $A=B$ 时,$\sigma_{AB}=2$.

虽然电子基态简并度和核基态简并度对气体的热容量没有贡献,但在计算气体反应平衡常数(气体质量作用定理)时却必须考虑(见习题 15).

§8.1.3 由玻尔兹曼分布求速度分布和空间分布

从粒子数按能量的玻尔兹曼分布律可以求粒子数按广义坐标或广义动量的分布.

在§8.1.1 中已经指出,当分子某一自由度的能量可视为连续变化时,分子于该自由度的运动可用经典方法描述,并可引入 μ 空间.这时 M－B 分布[§7.3.4 (4.16)和(4.18)式]

$$\langle n_r\rangle=(N/Z_1)e^{-\beta\varepsilon_r},\qquad Z_1=\sum_r e^{-\beta\varepsilon_r}=\sum_{\varepsilon_r}e^{-\beta\varepsilon_r}g(\varepsilon_r)$$

在 μ 空间中应表为以下形式:

$$dN(\boldsymbol{q},\boldsymbol{p})=(N/Z_1)e^{-\beta\varepsilon(\boldsymbol{q},\boldsymbol{p})}\frac{d\boldsymbol{q}d\boldsymbol{p}}{h^f}\tag{3.1}$$

或

$$dP(\boldsymbol{q},\boldsymbol{p})=(1/Z_1)e^{-\beta\varepsilon(\boldsymbol{q},\boldsymbol{p})}\frac{d\boldsymbol{q}d\boldsymbol{p}}{h^f}\tag{3.2}$$

其中

$$Z_1=\int e^{-\beta\varepsilon(\boldsymbol{q},\boldsymbol{p})}\frac{d\boldsymbol{q}d\boldsymbol{p}}{h^f}=\int e^{-\beta\varepsilon(\boldsymbol{q},\boldsymbol{p})}g(\varepsilon)d\varepsilon\tag{3.3}$$

[见§8.1.1(1.12)式]、$d\boldsymbol{q}d\boldsymbol{p}$ 代表 μ 空间中 $(\boldsymbol{q},\boldsymbol{p})$ 附近的小相体积元,$d\boldsymbol{q}d\boldsymbol{p}/h^f$

为该相体积元中的相格数，$\varepsilon = \varepsilon(\boldsymbol{q},\boldsymbol{p})$ 是分子运动状态落在该相体积元中时能量的经典表达式，$\mathrm{d}N(\boldsymbol{q},\boldsymbol{p})$ 代表分子运动状态落在该相体积元中的分子数. $\mathrm{d}P(\boldsymbol{q},\boldsymbol{p})$ 则代表分子的广义坐标落在 $(\boldsymbol{q} \sim \boldsymbol{q}+\mathrm{d}\boldsymbol{q})$，广义动量落在 $(\boldsymbol{p} \sim \boldsymbol{p}+\mathrm{d}\boldsymbol{p})$ 间的联合概率.

在实际问题中，我们往往只对分子按某特定的广义坐标或广义动量分布感兴趣. 按照数学中已知联合概率求边缘概率的方法，这只要将(3.2)式或(3.1)式对我们不感兴趣的广义坐标和广义动量进行积分就可得到.

为具体起见，下面通过三个例子来说明从粒子数按能量的玻尔兹曼分布律求粒子数按广义坐标或广义动量分布的基本方法.

(一)麦克斯韦速度分布

设我们研究的系统是由单原子分子组成的理想气体，则分子的能量只包括质心平动的动能和在外场中的势能. 根据前节的讨论，质心运动的能量可以看成是连续变化的，经典描述有效. 其广义坐标和广义动量为 $\boldsymbol{r}$ 和 $\boldsymbol{p}$，μ 空间的维数是6，分子能量的经典表达式为 $\varepsilon = p^2/(2m) + u(\boldsymbol{r})$. 于是(3.3)式和(3.2)式具体为

$$Z_1 = \int \mathrm{e}^{-\beta\left[\frac{p^2}{2m}+u(\boldsymbol{r})\right]}\frac{\mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{p}}{h^3} = \frac{1}{h^3}\int \mathrm{e}^{-\beta\frac{p^2}{2m}}\mathrm{d}^3\boldsymbol{p}\int \mathrm{e}^{-\beta u(\boldsymbol{r})}\mathrm{d}^3\boldsymbol{r}, \tag{3.4}$$

$$\mathrm{d}P(\boldsymbol{r},\boldsymbol{p}) = \frac{1}{Z_1}\mathrm{e}^{-\beta\left[\frac{p^2}{2m}+u(\boldsymbol{r})\right]}\frac{\mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{p}}{h^3} = \frac{\mathrm{e}^{-\beta\frac{p^2}{2m}}\mathrm{d}^3\boldsymbol{p}\,\mathrm{e}^{-\beta u(\boldsymbol{r})}\mathrm{d}^3\boldsymbol{r}}{\int \mathrm{e}^{-\beta\frac{p^2}{2m}}\mathrm{d}^3\boldsymbol{p}\int \mathrm{e}^{-\beta u(\boldsymbol{r})}\mathrm{d}^3\boldsymbol{r}} \tag{3.5}$$

(3.5)式说明，在平衡态下，分子质心的坐标落在 $(\boldsymbol{r} \sim \boldsymbol{r}+\mathrm{d}\boldsymbol{r})$，同时动量落在 $(\boldsymbol{p} \sim \boldsymbol{p}+\mathrm{d}\boldsymbol{p})$ 的联合概率，等于坐标分布概率和动量分布概率的乘积. 因此，根据独立事件概率相乘的原理知道，质心的动量分布和坐标分布彼此无关，是独立事件.

如果不论分子质心的位置在哪里，而只求分子质心的动量落在 $(\boldsymbol{p} \sim \boldsymbol{p}+\mathrm{d}\boldsymbol{p})$ 的概率，则可将(3.5)式对 $\mathrm{d}^3\boldsymbol{r}$ 积分

$$\begin{aligned}\mathrm{d}P(\boldsymbol{p}) &= \frac{\mathrm{e}^{-\frac{1}{2mk_\mathrm{B}T}(p_x^2+p_y^2+p_z^2)}\mathrm{d}p_x\mathrm{d}p_y\mathrm{d}p_z}{\iiint \mathrm{e}^{-\frac{1}{2mk_\mathrm{B}T}(p_x^2+p_y^2+p_z^2)}\mathrm{d}p_x\mathrm{d}p_y\mathrm{d}p_z} \\ &= [1/(2\pi mk_\mathrm{B}T)^{3/2}]\,\mathrm{e}^{-(p_x^2+p_y^2+p_z^2)/(2mk_\mathrm{B}T)}\mathrm{d}p_x\mathrm{d}p_y\mathrm{d}p_z.\end{aligned} \tag{3.6}$$

将 $p_x = mv_x, p_y = mv_y, p_z = mv_z$ 代入，得到分子速度落在 $(v_x \sim v_x+\mathrm{d}v_x, v_y \sim v_y+\mathrm{d}v_y, v_z \sim v_z+\mathrm{d}v_z)$ 的概率，

$$\mathrm{d}P(\boldsymbol{v}) = [m/(2\pi k_\mathrm{B}T)]^{3/2}\mathrm{e}^{-(m/2k_\mathrm{B}T)(v_x^2+v_y^2+v_z^2)}\mathrm{d}v_x\mathrm{d}v_y\mathrm{d}v_z. \tag{3.7}$$

这就是麦克斯韦速度分布. 如果将速度空间中的体积元 $\mathrm{d}v_x\mathrm{d}v_y\mathrm{d}v_z$ 换为球坐标形式 $v^2\sin\theta\mathrm{d}\theta\mathrm{d}\varphi\mathrm{d}v$，并对 $\mathrm{d}\theta$ 和 $\mathrm{d}\varphi$ 积分，则得到麦克斯韦速率分布：

$$\mathrm{d}P(v) = 4\pi[m/(2\pi k_\mathrm{B}T)]^{3/2}\mathrm{e}^{-mv^2/(2k_\mathrm{B}T)}v^2\mathrm{d}v. \tag{3.8}$$

以上推导表明,外场的存在不影响质心按速度的分布,也不难看出,分子的内部运动也不会影响质心按速度的分布.

(二)重力场中分子随高度的分布

如果不问分子的动量,只求分子质心在外场中的位置分布,则可将(3.5)式对动量积分,得

$$dP(\boldsymbol{r}) = e^{-\beta u(\boldsymbol{r})} d^3\boldsymbol{r} \Big/ \int e^{-\beta u(\boldsymbol{r})} d^3\boldsymbol{r}. \tag{3.9}$$

或

$$dN(x,y,z) = Ne^{-\beta u(x,y,z)} dxdydz \Big/ \int e^{-\beta u(x,y,z)} dxdydz. \tag{3.10}$$

若外场为重力场,且假设重力加速度 g 为常数,则取 z 轴铅直向上时,分子的势能为 $u(x,y,z) = mgz$. 现在问:分子质心落在高度 $z \sim z + dz$ 之间,而不论其 x,y 如何的分子数. 此时可对 $dxdy$ 积分,

$$dN(z) = Ne^{-\beta mgz} dz \Big/ \int e^{-\beta mgz} dz. \tag{3.11}$$

设理想气体装在一截面积为 S 的无限高容器中,则由上式可得

$$n(z) = (1/S)(dN(z)/dz) = (N/S)[mg/(k_B T)]\, e^{-mgz/(k_B T)}.$$

用 n_0 表示 $z = 0$ 处的分子数密度,则上式可写为

$$n(z) = n_0 e^{-mgz/(k_B T)}. \tag{3.12}$$

这就是理想气体分子在重力场中按高度的分布. 考虑理想气体的状态方程 $p = nk_B T$,由(3.12)式可得到大气压强随高度变化的公式

$$p = p_0 e^{-mgz/(k_B T)}. \tag{3.13}$$

(3.13)式说明,可以利用大气压强来估计高度. 但是,由于地面上大气温度不均匀,气体不是处在平衡态,所以这种估计是粗糙的.

由整个推导过程不难看出,(3.12)式不仅适用于重力场中的理想气体,而且适用于在任何均匀外场中的任意近独立非定域粒子系统. 例如,半导体 p - n 结处均匀电场中的电子和悬浮在空气或液体中的布朗微粒. 配兰的布朗运动实验之一就是用显微镜观察悬浮在液体中的胶体微粒,测量不同高度处布朗微粒的相对密度. 他在实验中用了半径相差 50 倍的粒子,温度范围为 10℃ ~60℃,结果很好地验证了(3.12)式,而且由实验数据推算出了玻尔兹曼常量,再由 R/k_B 算出阿伏伽德罗常量,得到了满意的结果.

(三)气体电介质的极化

气体分子的电偶极矩为 $\boldsymbol{p}_e = \sum_i e_i \boldsymbol{r}_i$,求和对分子中所有电荷进行. 显然,在具有对称性的原子和分子中,正、负电荷对称分布,因而 $\boldsymbol{p}_e$ 为零,如 H_2,O_2,CH_4 等.

而非对称性很强的分子，如 HCl, CH_3Cl, H_2O 等则有异于零的电偶极矩 p_e. 为具体起见，设电介质是由具有电偶极矩 p_e 的双原子分子组成的理想气体. 电偶极矩的相互作用可以忽略. 在没有外电场时，分子电偶极矩的取向是混乱的，因而整个气体不显示电性. 在外电场的作用下，电介质发生极化，单位体积中的电偶极矩称为电极化强度. 我们从玻尔兹曼分布来研究在外场中分子电偶极矩沿各种方向的分布和极化强度与外电场的关系.

一个电偶极矩 p_e 在外电场 E 中具有电势能

$$-p_eE\cos\theta. \tag{3.14}$$

θ 为电偶极矩 p_e 与电场 E 的夹角，并取 E 的方向为 z 轴，因而分子的能量为

$$\varepsilon = \frac{p^2}{2m} + \frac{1}{2I}\left(p_\theta^2 + \frac{p_\varphi^2}{\sin^2\theta}\right) - p_eE\cos\theta. \tag{3.15}$$

上式中的第一部分是分子质心平动的动能，第二部分为转动动能. 因为室温下平动和转动动能都可以看成是连续变化的，因而采用了经典表达式. 第三部分是电偶极矩在外场中的电势能. 分子的自由度为 5，μ 空间为 10 维，μ 空间中的相体积元为

$$\mathrm{d}\tau = \mathrm{d}x\mathrm{d}y\mathrm{d}z\mathrm{d}\theta\mathrm{d}\varphi\mathrm{d}p_x\mathrm{d}p_y\mathrm{d}p_z\mathrm{d}p_\theta\mathrm{d}p_\varphi. \tag{3.16}$$

现在求电偶极矩 p_e 的方向落在 $(\theta \sim \theta + \mathrm{d}\theta)$ 和 $(\varphi \sim \varphi + \mathrm{d}\varphi)$ 中的分子数. 为此只要将(3.15)和(3.16)式代入(3.1)式，并对除 θ, φ 以外的其他广义坐标和广义动量积分就行了. 由于对 $\mathrm{d}p_\varphi$ 积分时，被积函数与 θ 有关，因而对于 $\mathrm{d}p_\varphi$ 的积分在分子分母中不能相消，

$$\begin{aligned}\mathrm{d}N(\theta,\varphi) &= N\frac{\mathrm{e}^{\beta p_eE\cos\theta}\left[\int_{-\infty}^{+\infty}\mathrm{e}^{-\beta p_\varphi^2/(2I\sin^2\theta)}\mathrm{d}p_\varphi\right]\mathrm{d}\theta\mathrm{d}\varphi}{\left\{\int_0^\pi \mathrm{e}^{\beta p_eE\cos\theta}\left[\int_{-\infty}^{+\infty}\mathrm{e}^{-\beta p_\varphi^2/(2I\sin^2\theta)}\mathrm{d}p_\varphi\right]\mathrm{d}\theta\right\}\int_0^{2\pi}\mathrm{d}\varphi} \\ &= \frac{N}{2\pi}\frac{\mathrm{e}^{\beta p_eE\cos\theta}\sin\theta\mathrm{d}\theta\mathrm{d}\varphi}{\int_0^\pi \mathrm{e}^{\beta p_eE\cos\theta}\sin\theta\mathrm{d}\theta}.\end{aligned} \tag{3.17}$$

$\sin\theta\mathrm{d}\theta\mathrm{d}\varphi$ 为 (θ,φ) 方向的元立体角，$\mathrm{d}N(\theta,\varphi)$ 就是电偶极矩方向落在该元立体角中的分子数.

令 $x = \beta p_eE$，$y = x\cos\theta$，将(3.17)式对 $\mathrm{d}\varphi$ 积分，得电偶极矩落在 $(\theta \sim \theta + \mathrm{d}\theta)$ 的概率

$$\mathrm{d}P(\theta) = \mathrm{e}^y\mathrm{d}y\Big/\int_x^{-x}\mathrm{e}^y\mathrm{d}y = -(\mathrm{e}^y/2\mathrm{sh}\,x)\mathrm{d}y. \tag{3.18}$$

沿电场方向的平均电偶极矩为

$$\langle p_e\cos\theta\rangle = \int p_e\cos\theta\mathrm{d}p(\theta) = -\frac{p_e}{2x\,\mathrm{sh}\,x}\int_x^{-x}y\mathrm{e}^y\mathrm{d}y = p_e(\mathrm{cth}\,x - 1/x). \tag{3.19}$$

分子电偶极矩的数量级约为 10^{-18}esu·cm. 因此,在气体尚未液化的温度及可达的场强下,总有 $p_e E \ll k_B T$ 或 $x \ll 1$. 因 $\mathrm{cth}\, x \simeq \frac{1}{x} + \frac{x}{3} + \cdots$,故 $\langle p_e \cos\theta \rangle = x p_e/3$. 设单位体积中有 n 个分子,则高温弱场下的电极化强度为

$$\mathscr{P} = \langle n p_e \cos\theta \rangle = n p_e^2 E/(3k_B T). \tag{3.20}$$

电极化率为

$$\chi_E \equiv \mathscr{P}/E = n p_e^2/(3k_B T). \tag{3.21}$$

(3.21)式叫做朗之万公式. 气体电极化率和可直接测量的介电常数 K 有以下关系

$$K = 1 + 4\pi\chi_E,$$

因而利用实验测得的介电常数可以求得气体电介质的电偶极矩 p_e,而 p_e 对判断分子的结构和几何形状有重要作用.

§8.1.4 能量均分定理

从本章的第1、2节中我们看到,双原子分子有三个平动自由度、两个转动自由度和一个振动自由度. 在高温极限下,它们的能量可分别表示为:

$$\varepsilon^{平} = (p_x^2 + p_y^2 + p_z^2)/(2m),$$

$$\varepsilon^{转} = [p_\theta^2 + (p_\varphi^2/\sin^2\theta)]/(2I),$$

$$\varepsilon^{振} = p^2/(2m) + m\omega^2 r^2/(2I).$$

相应的分子平均能量分别为 $3k_B T/2$, $k_B T$ 和 $k_B T$. 也就是说,与分子能量经典表达式中每一平方项对应的平均值是 $k_B T/2$. 在本节里,我们用玻尔兹曼分布律证明,这实际上是普遍的能量均分定理的特例. 在高温极限下,能量均分定理是一个有效的定理,利用它不需计算配分函数,就可以很方便地求得某些系统的内能和热容量.

能量均分定理可表述如下:在给定的温度 T 下,处于平衡态的经典系统中,粒子能量 ε 的表达式中每一个独立平方项的平均值等于 $k_B T/2$.

对"独立平方项"的含义需要作些说明. 一般说来,一个粒子可能有 f 个自由度,其广义坐标和广义动量分别为 $q_1, \cdots, q_f$ 和 $p_1, \cdots, p_f$. 若粒子的能量对 p_i 的依赖关系可以写为

$$\varepsilon = b_i p_i^2/2 + \varepsilon', \tag{4.1}$$

其中 ε' 不包含 p_i,而且 b_i 也与 p_i 无关;或者

$$\varepsilon = c_i q_i^2/2 + \varepsilon'', \tag{4.2}$$

其中 ε'' 不包含 q_i,而 c_i 也与 q_i 无关,则 i 取每一个值的 $b_i p_i^2/2$ 或 $c_i q_i^2/2$ 叫做一个

"独立平方项".要注意的是，b_i（或 c_i）可能是常数，也可能是除 p_i（或 q_i）以外的其他广义坐标和广义动量的函数.能量均分定理可用数学形式写为

$$\langle b_i p_i^2 \rangle /2 = \langle c_i q_i^2 \rangle /2 = k_B T/2, \tag{4.3}$$

证　为了具体起见，我们证明 $\langle b_i p_i^2 \rangle /2 = k_B T/2$.已经知道在高温极限下，粒子能量近似为连续变化时，粒子的配分函数可在 μ 空间中用积分表示

$$Z_1 = h^{-f} \int e^{-\beta \varepsilon} dq_1 \cdots dq_f dp_1 \cdots dp_f. \tag{4.4}$$

又按 §8.1.1(1.6)式

$$\langle \varepsilon \rangle = - \partial \ln Z_1 / \partial \beta. \tag{4.5}$$

将(4.1)式代入(4.4)式

$$Z_1 = h^{-f} \int e^{-\beta (b_i p_i^2/2 + \varepsilon')} dq_1 \cdots dq_f dp_1 \cdots dp_f = h^{-f} \int e^{-\beta \varepsilon'} d\tau' \int_{-\infty}^{\infty} e^{-\beta b_i p_i^2/2} dp_i.$$

其中 $d\tau' = dq_1 \cdots dq_f dp_1 \cdots dp_{i-1} dp_{i+1} \cdots dp_f$.由于

$$\int_{-\infty}^{\infty} e^{-\beta b_i p_i^2/2} dp_i = \sqrt{2\pi/b_i}\, \beta^{-1/2}.$$

所以

$$Z_1 = \beta^{-1/2} h^{-f} \int e^{-\beta \varepsilon'} d\tau' \sqrt{2\pi/b_i}. \tag{4.6}$$

由于 b_i 可能是除 p_i 以外的其他广义坐标和广义动量的函数，所以 $\sqrt{1/b_i}$ 不一定能提到积分号之外.将(4.6)式代入(4.5)式，得到

$$\langle \varepsilon \rangle = \langle b_i p_i^2 \rangle /2 + \langle \varepsilon' \rangle = k_B T/2 + \langle \varepsilon' \rangle, \tag{4.7}$$

其中

$$\langle \varepsilon' \rangle = \int \varepsilon' e^{-\beta \varepsilon'} \sqrt{2\pi/b_i} d\tau' \Big/ \int e^{-\beta \varepsilon'} \sqrt{2\pi/b_i} d\tau'. \tag{4.8}$$

如果独立平方项的变量不是 p_i 而是 q_i，其证明完全相同.　**【证毕】**

对于单原子、刚性双原子分子和准弹性双原子分子，它们能量的经典表达式分别包含 3 个 5 个和 7 个独立平方项.根据能量均分定理，其能量平均值分别为 $3k_B T/2$、$5k_B T/2$ 和 $7k_B T/2$，相应的摩尔定体热容量为 $3R/2$、$5R/2$ 和 $7R/2$.这正是前面通过计算配分函数并考虑高温极限所得结果.

根据能量均分定理还很容易得到高温极限下固体的热容量.设把由 N 个原子组成的理想晶体看成是由 $3N$ 个近独立的经典谐振子组成的系统，每个经典谐振子的能量包含两个独立平方项，平均能量为 $k_B T$.于是，理想固体的摩尔内能为 $3RT$，摩尔定体热容量为 $3R$.根据这一结果，c_V 和原子质量及振动频率无关.在 §7.3.2 中曾经指出这与大多数固体室温下的性质是符合的，对于金刚石，则要在高得多的温度下才有 $c_V = 3R$.

最后,我们对能量均分定理的适用条件作几点说明:

①用能量均分定理计算系统的定体热容量虽然很方便,但是只有在以下两个条件同时满足时才能应用.这两个条件是:i)温度足够高,粒子的能量近似连续以致可以采用经典描述;ii)粒子的能量中可分出独立平方项.

②虽然在这里我们是用玻尔兹曼分布律导出能量均分定理,但是能量均分定理和麦克斯韦速度分布一样可以用正则系综导出,因而不仅对理想气体这样的近独立粒子系统成立,对非理想气体也是适用的.

③能量均分定理不仅对宏观系统中的微观粒子有效,对于作无规则运动的宏观粒子,如布朗微粒、悬线电流计中的线圈等也是适用的.

§8.1.5 稀薄实际气体的状态方程

由第七篇第一章知道,非理想气体最完善的状态方程是卡末林·昂内斯方程,或位力展开式[§7.1.1(1.4)式]:

$$pV = Nk_BT(1 + B_2/V + B_3/V^2 + \cdots). \tag{5.1}$$

其中 B_2,B_3 分别称为第二、第三位力系数.本节将用正则系综来推导这一方程.为了简单,只考虑气体密度相当稀薄的情况.

根据正则系综的公式

$$p = k_BT(\partial \ln Z/\partial V) \tag{5.2}$$

知道,推导卡末林·昂内斯方程的关键是计算非理想气体的配分函数.不妨先回顾一下,在前面我们是怎样计算理想气体的配分函数的.

当分子间相互作用可以忽略不计时,按 $Z_{M\text{-}B} = Z_1^N/N!$,只要知道分子的配分函数,就可以容易地计算出玻尔兹曼理想气体的配分函数.因此那时,我们着重讨论如何计算分子的配分函数.

但是,当分子间的相互作用不能忽略时,系统的能量应表示为

$$E = \sum_{i=1}^{N} \varepsilon_i + u(\boldsymbol{r}_1, \boldsymbol{r}_2, \cdots, \boldsymbol{r}_N), \tag{5.3}$$

u 是分子间相互作用的势能,是 N 个分子坐标的函数.原则上,每个分子的能量都依赖于它和其余 $(N-1)$ 个分子的相对位置,也就是说,单个分子的能量已失去意义,从而,分子的配分函数也随之失去意义.这时,必须把系统作为一个整体来研究,利用最基本的公式

$$Z = \sum_{\{n_i\}} \mathrm{e}^{-\beta E(\{n_i\})} \tag{5.4}$$

来计算系统的配分函数.

对于理想气体中的分子，当经典处理可以适用时，我们曾引入 μ 空间，于是计算分子配分函数的求和可以用 μ 空间的积分代替，计算得以方便地进行. 推广这种方法到整个系统(不论其分子是否有相互作用)，当经典处理可以适用时，需要采用 §7.1.3 中已引入的 Γ 空间. 于是计算系统配分函数的求和式(5.4)也可以用 Γ 空间中的积分代替. 下面先复习什么是 Γ 空间，并说明系统配分函数在 Γ 空间中的表示.

(一) Γ 空间　系统配分函数在 Γ 空间中的表示

考虑由 N 个粒子组成的系统，粒子间可能有相互作用. 每个粒子的自由度为 f，则整个系统的自由度为 $s = Nf$. 系统的广义坐标和广义动量分别为 $q_1, q_2, \cdots, q_s$ 和 $p_1, p_2, \cdots, p_s$. 以 $2s$ 个广义坐标和广义动量为坐标轴所张开的 $2s$ 维空间叫做 Γ 空间. 将 Γ 空间分为大小等于 h^s 的相格，每个相格代表系统的一个微观态. Γ 空间中宏观无限小的相体积元 $\mathrm{d}\Gamma$ 中所含微观态数为

$$\mathrm{d}\Gamma/(N!h^s) = \mathrm{d}q_1\mathrm{d}q_2\cdots\mathrm{d}q_s\mathrm{d}p_1\mathrm{d}p_2\cdots\mathrm{d}p_s/(N!h^s), \tag{5.5}$$

其中 $N!$ 是考虑到粒子全同性而引入的. (5.4)式化为

$$Z = \frac{1}{N!h^s}\int \mathrm{e}^{-\beta E(q_1,\cdots,q_s,p_1,\cdots,p_s)}\mathrm{d}\Gamma. \tag{5.6}$$

(5.6)式是当经典处理适用时，系统配分函数在 Γ 空间中的表示. 相应地，可以写出巨配分函数在 Γ 空间中的表示

$$\Xi = \sum_{N\geqslant 0}\int \mathrm{e}^{-\beta(E_N - N\mu)}\mathrm{d}\Gamma_N/(N!h^{Nf}). \tag{5.7}$$

这里在 $\mathrm{d}\Gamma_N$ 中增加下标 N 以表明它是粒子数为 N 时，$2Nf$ 维 Γ 空间中的相体元.

要注意的是，μ 空间是分子相空间，其中的相格代表分子的运动状态. μ 空间只适用于理想气体. Γ 空间是系统的状态空间，其中的相格代表系统的微观态. Γ 空间适用于任何系统，不论组成系统的粒子间是否有相互作用，只要经典处理是可行的.

为了熟悉 Γ 空间，我们先用(5.6)式计算单原子理想气体的配分函数.

【例】 单原子理想气体的配分函数.

在无外场时，单原子理想气体的总能量是每个分子平动动能之和

$$E = \frac{1}{2m}\sum_{i=1}^{N}(p_{ix}^2 + p_{iy}^2 + p_{iz}^2).$$

Γ 空间中的相体积元中 $\mathrm{d}\Gamma = \mathrm{d}\tau_1\mathrm{d}\tau_2\cdots\mathrm{d}\tau_N$，其中 $\mathrm{d}\tau = \mathrm{d}x\mathrm{d}y\mathrm{d}z\mathrm{d}p_x\mathrm{d}p_y\mathrm{d}p_z$ 是 μ 空间的相体积元. 据(5.6)式，单原子理想气体的配分函数为

$$Z = \frac{1}{N!}\frac{1}{h^{3N}}\int\cdots\int \mathrm{e}^{-\frac{1}{2mk_BT}\sum_{i=1}^{N}(p_{ix}^2+p_{iy}^2+p_{iz}^2)}\mathrm{d}\tau_1\mathrm{d}\tau_2\cdots\mathrm{d}\tau_N$$

$$= (1/N!)\left\{(1/h^3)\int e^{-\frac{1}{2mk_BT}(p_x^2+p_y^2+p_z^2)}\,dxdydzdp_x dp_y dp_z\right\}^N$$

$$= (1/N!)\{V(2m\pi k_B T/h^2)^{3/2}\}^N = Z_1^N/N!.$$

和用 μ 空间所得结果一致.

(二)非理想气体的配分函数 位形积分

非理想气体与理想气体的区别在于分子间有相互作用.若气体密度仍然稀薄,可以认为三个以上分子同时"碰撞"的机会很少.所以整个非理想气体分子间的相互作用势能近似等于所有可能分子对之间的势能 $u_{ij}(r_{ij})$ 之和.于是,稀薄气体的能量为:

$$E = \sum_{i=1}^{N}\frac{p_i^2}{2m} + \sum_{i<j}u_{ij}(r_{ij}), \tag{5.8}$$

r_{ij} 为 i,j 分子间的距离.(5.8)式中除势能外只包含分子平动动能的原因是,只有平动配分函数依赖于气体的体积,因而求状态方程时可以不考虑分子的内部运动,而把分子看成质点.

气体 Γ 空间中的相体积元为(将 $d^3\boldsymbol{r}$ 和 $d^3\boldsymbol{p}$ 简写为 $d\boldsymbol{r}$ 和 $d\boldsymbol{p}$)

$$d\Gamma = d\boldsymbol{r}_1 d\boldsymbol{r}_2\cdots d\boldsymbol{r}_N d\boldsymbol{p}_1 d\boldsymbol{p}_2\cdots d\boldsymbol{p}_N. \tag{5.9}$$

将(5.8)式和(5.9)式代入(5.6)式,得到

$$Z = (1/N!h^{3N})\int\cdots\int e^{-\beta[\sum_i p_i^2/2m+\sum_{i<j}u_{ij}]}d\boldsymbol{r}_1 d\boldsymbol{r}_2\cdots d\boldsymbol{r}_N d\boldsymbol{p}_1 d\boldsymbol{p}_2\cdots d\boldsymbol{p}_N$$

$$= (1/N!h^{3N})\int e^{-\frac{1}{2mk_BT}\sum_{i=1}^{N}p_i^2}d\boldsymbol{p}_1 d\boldsymbol{p}_2\cdots d\boldsymbol{p}_N\int e^{-\sum_{i<j}u_{ij}/(k_BT)}d\boldsymbol{r}_1 d\boldsymbol{r}_2\cdots d\boldsymbol{r}_N$$

$$= (1/N!)[(2m\pi k_B T/h^2)^{3/2}V]^N \times\left[(1/V^N)\int e^{-\sum_{i<j}u_{ij}/(k_BT)}d\boldsymbol{r}_1 d\boldsymbol{r}_2\cdots d\boldsymbol{r}_N\right].$$

上式右方的积分

$$Z_{位形} = \int e^{-\sum_{i<j}u_{ij}/(k_BT)}d\boldsymbol{r}_1 d\boldsymbol{r}_2\cdots d\boldsymbol{r}_N. \tag{5.10}$$

叫做位形积分,由分子相互作用决定.上式可写为

$$Z = Z^{(0)}(Z_{位形}/V^N). \tag{5.11}$$

我们看到,非理想气体的配分函数分解为两个因子相乘.第一个因子就是理想气体的配分函数,第二个因子与分子间的相互作用有关.将(5.11)式代入(5.2)式

$$p = k_B T\partial\ln Z/\partial V = k_B T\partial\ln Z^{(0)}/\partial V + k_B T\partial(\ln Z_{位形}/V^N)/\partial V$$

$$= \frac{Nk_BT}{V} + k_BT\frac{\partial}{\partial V}\left(\ln\frac{Z_{位形}}{V^N}\right). \tag{5.12}$$

第一项正是 M-B 理想气体的压强,第二项则表示由于分子间的相互作用引起的对理想气体压强的修正.

为了求得修正压强，必需计算位形积分. 这里要用到的一个数学技巧是引进迈耶函数 f_{ij}，它定义为

$$f_{ij} \equiv e^{-u_{ij}/(k_B T)} - 1. \qquad (5.13)$$

f_{ij} 的意义如下：当两个分子间的距离 r_{ij} 较大时 u_{ij} 趋于零，分子 i 和分子 j 相互独立，$f_{ij} \to 0$；相反，当两个分子靠近时，r_{ij} 变小，u_{ij} 不等于零，分子 i,j 相互关联，而 f_{ij} 不等于零. 所以 f_{ij} 是反映两个分子是否存在相关性的量，它的大小可作为两个分子相关性的量度. 将(5.13)式代入(5.10)式，得到

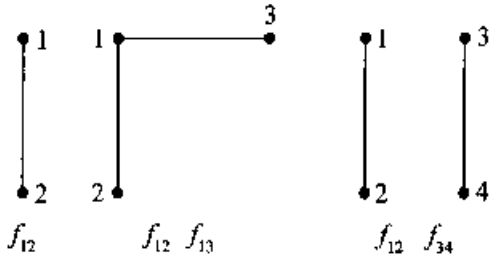


图 1　迈耶函数的意义

$$Z_{位形} = \int \prod_{i<j} (1 + f_{ij}) \, d\boldsymbol{r}_1 d\boldsymbol{r}_2 \cdots d\boldsymbol{r}_N$$

$$= \int \Big\{ 1 + \sum_{i<j} f_{ij} + \sum_{i<j} \sum_{i'<j'} f_{ij} f_{i'j'} + \cdots \Big\} d\boldsymbol{r}_1 d\boldsymbol{r}_2 \cdots d\boldsymbol{r}_N. \qquad (5.14)$$

根据上面所讲的 f_{ij} 的性质，我们来说明展开式中各项的意义. 如图 1 所示，展开式中第一个求和代表一对分子靠近而发生相互作用的项，即代表一对分子发生碰撞的项. 第二个求和中含如 $f_{12}f_{34}$ 的项，它们代表同时有两对分子碰撞的项. 此外，还有另外一些如 $f_{12}f_{13}$，只有在 1,2,3 三个分子同时靠近发生相互作用时，这些项才不为零，它们代表三体碰撞项. 对于稀薄气体，同时有两对碰撞或三体碰撞的概率是很小的，所以除一对碰撞的项外，其余各项都可略去. 于是(5.14)式只剩下前面两项

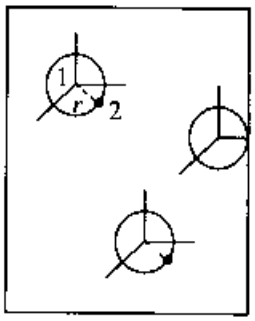


图 2　分子作用球

$$Z^{位形} \approx V^N + \sum_{i<j} \int f_{ij} \, d\boldsymbol{r}_1 d\boldsymbol{r}_2 \cdots d\boldsymbol{r}_N = V^N + \sum_{i<j} V^{N-2} \int f_{ij} \, d\boldsymbol{r}_i d\boldsymbol{r}_j. \qquad (5.15)$$

对于每一对分子，积分 $\int f_{ij} d\boldsymbol{r}_i d\boldsymbol{r}_j$ 相同，而 N 个分子总共可组成 $N(N-1)/2$ 个分子对，于是

$$Z^{位形} = V^N + V^{N-2}[N(N-1)/2] \int f_{12} d\boldsymbol{r}_1 d\boldsymbol{r}_2. \qquad (5.16)$$

因为分子间的相互作用势能 u_{ij} 只是两个分子间相对距离的函数，故可采用相对坐标，将上述积分写为 $\int f_{12} d\boldsymbol{r}_1 d\boldsymbol{r}_2 = \int \Big[\int f(r_{12}) d\boldsymbol{r}_{12} \Big] d\boldsymbol{r}_1$. 注意到，当 r_{12} 大于分子力的力程时，$f(r_{12})$ 等于零. 而分子力又是短程力，所以积分 $\int f(r_{12}) d\boldsymbol{r}_{12}$ 只在分子 1 附近很小的范围内不为零，且与分子 1 的位置无关. 只有当分子 1 靠近器壁时才受到

影响，如图 2 所示. 但边界效应可以忽略，于是积分可以进一步化为 $\int f_{12}\mathrm{d}\boldsymbol{r}_1\mathrm{d}\boldsymbol{r}_2 = V \cdot 4\pi\int_0^\infty f(r_{12})r_{12}^2\mathrm{d}r_{12}$. 代入(5.16)式得

$$Z^{位形}/V^N = 1 + (N^2/2V)\cdot 4\pi\int_0^\infty f(r)r^2\mathrm{d}r. \tag{5.17}$$

如上所述，$f(r)$ 只在 r 小于分子力程时，才不为零，所以 $4\pi\int_0^\infty f(r)r^2\mathrm{d}r$ 的数量级是以分子力程为半径的球，叫做分子作用球 v_c. 对于稀薄气体可以认为

$$\frac{N^2}{V}\cdot 4\pi\int_0^\infty f(r)r^2\mathrm{d}r \approx \frac{Nv_c}{V/N} \ll 1, \tag{5.18}$$

其中，Nv_c 表示全部分子作用球之和，而 V/N 表示每个分子平均分得的空间. 当气体密度很小时，假设 $Nv_c \ll V/N$ 是合理的，故(5.18)式成立. 于是

$$\ln(Z^{位形}/V^N) = \ln\left(1 + \frac{N^2}{2V}\cdot 4\pi\int_0^\infty f(r)r^2\mathrm{d}r\right) \tag{a}$$

$$\approx \frac{N^2}{2V}\cdot 4\pi\int_0^\infty f(r)r^2\mathrm{d}r. \tag{5.19}$$

代入(5.12)式

$$p = \frac{Nk_BT}{V}\left[1 - 2\pi\frac{N}{V}\int_0^\infty f(r)r^2\mathrm{d}r\right]$$

$$= \frac{Nk_BT}{V}\left(1 + \frac{B}{V}\right), \tag{5.20}$$

$$B = -2\pi N\int_0^\infty (\mathrm{e}^{-u(r)/(k_BT)} - 1)r^2\mathrm{d}r \tag{5.21}$$

和(5.1)式比较，可见 B 是第二位力系数，要计算 B 的值，需了解分子作用势 $u(r)$ 的具体形式.

(三)第二位力系数的计算　分子作用的刚球模型

分子相互作用的半经验公式为

$$u(r) = u_0[(r_0/r)^{12} - 2(r_0/r)^6]. \tag{5.22}$$

$u(r)$ 与 r 的关系如图 3 a 所示. r_0 是当势能取最小值 u_0 时分子间的距离，大约是 10^{-8}cm. 当 r 大于 r_c 时，分子间相互作用可忽略不计. r_c 通

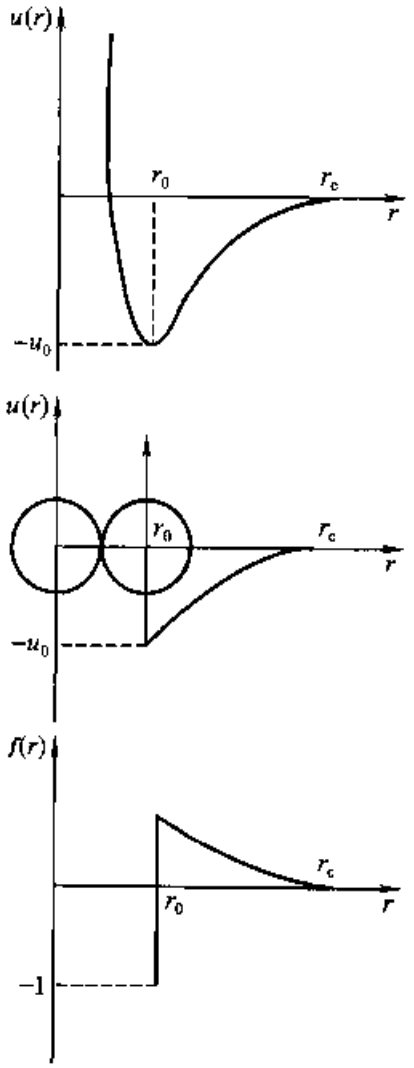


图 3　分子作用势

常叫分子作用半径，其数量级在 $10^{-7}\sim10^{-8}$cm 之间. 可将图 3(a)简化为图 3(b)，这实际上是提出了一个分子作用的刚球模型，即认为分子是有吸引力的刚性小球，其直径为 r_0，当 $r_c > r > r_0$ 时，分子间有吸引力. 但分子只能接近到 r_0 为止. 当 $r \leqslant r_0$ 时，吸引力停止而代之以无限的排斥力. 图 3(c)表示梅逸函数 $f(r)$［见(5.13)式］与 r 的关系.

现在根据这一模型来计算第二位力系数 B. 我们可以把(5.21)式中的积分分为两段，即从 $0\to r_0$ 和 $r_0\to\infty$，

$$B = -2\pi N\int_0^\infty (e^{-u(r)/(k_BT)} - 1)r^2dr$$

$$= -2\pi N\left\{\int_0^{r_0}(e^{-u(r)/(k_BT)} - 1)r^2dr + \int_{r_0}^\infty (e^{-u(r)/(k_BT)} - 1)r^2dr\right\}.$$

因为当 $r \leqslant r_0$ 时，$u(r)\to\infty$，$e^{-u(r)/(k_BT)} - 1 = -1$，所以第一个积分立刻可以求出：

$$B = \frac{2}{3}\pi Nr_0^3 - 2\pi N\int_{r_0}^\infty (e^{-u(r)/(k_BT)} - 1)r^2dr.$$

在一般温度下，$u(r) \ll k_BT$，所以 $e^{-u(r)/(k_BT)} - 1 \approx -u(r)/(k_BT)$，于是，

$$B(T) = \frac{2}{3}\pi Nr_0^3 + \frac{2\pi N}{k_BT}\int_{r_0}^\infty u(r)r^2dr. \tag{5.23}$$

当系统为 1mol 气体时，上式化作

$$B(T) = b - a/RT. \tag{5.24}$$

其中

$$b = 2N_0\pi r_0^3/3, \qquad a = -2\pi N_0^2\int_{r_0}^\infty u(r)r^2dr. \tag{5.25}$$

如将分子看成小球，其直径为 r_0，固有体积为 $\pi r_0^3/6$，则 b 是 1mol 分子固有体积的四倍. 因为在 r_0 到∞范围内，$u(r)$ 是相互吸引能，符号为负，所以 a 为正，是分子吸引力的一种度量.

将(5.25)式代入(5.20)式，并考虑 1mol 气体，

$$pV = RT\left(1 + \frac{B}{V}\right) = RT\left(1 + \frac{b - a/RT}{V}\right) = RT\left(1 + \frac{b}{V}\right) - \frac{a}{V}.$$

由于 $(b/V) \ll 1$，故 $1 + (b/V) \approx [1 - (b/V)]^{-1}$，得到

$$pV = \frac{RT\cdot V}{V - b} - \frac{a}{V}, \qquad 或 \qquad p = \frac{RT}{V - b} - \frac{a}{V^2}$$

$$\left(p + \frac{a}{V^2}\right)(V - b) = RT. \tag{5.26}$$

我们得到了范德瓦耳斯方程，并对方程中的常数 a 和 b 的意义作出了解释.

习 题

1. 表面活性物质的分子在液面上作二维自由运动，可看成一种二维理想气体. (a)试由 $M\text{-}B$ 分布导出二维理想气体的分子速率分布并求平均速率，最概然速率和均方速率. (b)试计算二维理想气体的内能 U、熵 S、压强 p 和热容量 C_s，C_s 的下标 s 为二维理想气体的面积.

2. 在容器中储有 k 种惰性单原子气体组成的混合系统，系统的温度为 T，气体 1 有 N_1 个分子，气体 2 有 N_2 个分子，…，气体 k 有 N_k 个分子. (a)通过计算系统的配分函数求系统的状态方程. (b)系统的总压强与第 i 种气体的分压 p_i (即第 i 种气体在相同温度下占有整个体积时的压强)的关系如何?

3. 有 N 个微弱相互作用的、具有固有磁矩 μ_0 的分子组成的顺磁气体. 每个分子的磁矩的取向是连续变化的，沿 z 轴加一外磁场 H，气体的温度为 T. (a)求每个分子的磁矩落在 $\theta \sim \theta + d\theta$ 之间的概率，并利用这一结果求系统的平均磁矩. (b)求系统的配分函数，并利用配分函数求平均磁矩. (c)证明，当 $\mu_0 H \ll k_B T$ 时，磁化率为 $\chi = N\mu_0^2/(3k_B T)$.

4. 考虑一个质量为 m，圆频率为 ω 的经典谐振子. 试计算它的配分函数，和由 N 个这样的谐振子组成的系统的内能和定体热容量.

5. N 个单原子分子组成的理想气体. 每个分子的质量为 m，气体盛在边长为 L 的立方盒中. 盒子的上下底与地面平行. 重力加速度为 g. (a)每个分子的平均动能是多少? (b)每个分子的平均势能是多少? (c)气体的总能量是多少? (d)求当 $mgL/(k_B T) \ll 1$ 和 $\gg 1$ 时气体的 C_V.

6. 二氧化碳是线性三原子分子，具有四种不同的振动方式(见图 1). 这些振动方式的特征温度 $\Theta^{振} = h\nu/k_B$ 分别是 3360K，1890K，954K 和 954K. 试计算 $T = 312$ K 时的摩尔热容量 c_V，实验值是 $3.53\,R$.

7. 极端相对论下粒子的能量与动量的关系为 $\varepsilon = pc$，求由这种粒子组成的 M-B 理想气体的热容量及状态方程，并与非相对论下的结果比较.

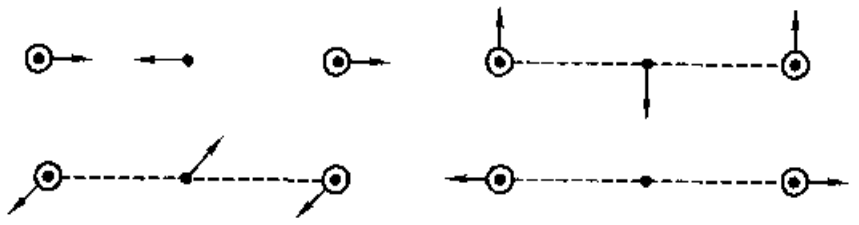


图 1 二氧化碳分子的四种振动方式

8. 质量为 m 的分子组成的理想气体，置于半径为 R 的离心机圆筒中. 圆筒以匀角速度转动，分子的能量为 $\varepsilon = mv^2/2 + m\omega^2 r^2/2$. (a)写出分子配分函数的表达式. (b)写出离轴 r 处，单位体积内分子数的表达式. (c)若圆筒内有质量分别为 m_1 和 m_2 的两种分子组成的混合理想气体，在圆筒壁和靠近离心机轴的地方，圆筒内每单位体积两种气体分子数的比值 $(dn_1/dn_2)_{r=R}$ 和 $(dn_1/dn_2)_{r=0}$ 之比叫分离系数 q：

$$q = (dn_1/dn_2)_{r=R}/(dn_1/dn_2)_{r=0}.$$

求 q，并说明如何利用这种装置测大分子的质量.

9. 半径为 r_0 的一根导线与一半径为 R、长为 L 的金属圆筒同轴，相对于圆筒，导线保持 V 伏的正电势. 系统的温度很高. 从金属发出的电子形成稀薄的气体充满圆筒与之达到平衡. 电子的

数密度很低，以致其静电相互作用可以忽略，从而服从 M-B 统计.（a）利用高斯定理求离开导线距离为 $r(r_0 < r < R)$ 一点的电场强度.（可近似认为圆筒长度 L 很大，以致边界效应可以忽略）.（b）在平衡态中，电子形成不同密度的气体，利用上一问的结果求每单位体积中的电子数对 r 的依赖关系.（c）给出一近似标准，根据它可以确定温度高到什么程度，以致电子数密度充分小，而忽略电子静电相互作用是合理的.

10. 假设爱因斯坦晶体中每个原子只作一维振动，蒸气为理想气体.（a）证明当爱因斯坦晶体与蒸气达到平衡时蒸气的压强为 $p = (k_B T/V)(z_g/z_1)$，其中 z_1 为晶体中一个谐振子的配分函数，z_g 为蒸气分子的配分函数.（b）设谐振子的能量为 $nh\nu(n = 0,1,2,\cdots)$，每个原子结合在晶体格点上时，其能量比在蒸气中低 ε_0，求蒸气压与温度的关系.（c）证明高温极限 $k_B T \gg h\nu$ 和低温极限 $k_B T \ll h\nu$ 下，蒸气压定律分别为

$$\frac{\mathrm{d}p}{\mathrm{d}T} = \frac{\varepsilon_0 + 3k_B T/2}{T}\frac{N_g}{V}, \qquad \frac{\mathrm{d}p}{\mathrm{d}T} = \frac{\varepsilon_0 + 5k_B T/2}{T}\frac{N_g}{V}.$$

11. 试用气体分子在重力场中按高度的分布规律，讨论中学教材中有关大气压强的问题.（a）根据统计理论，压强是由分子无规则运动碰撞产生的，而初中教材中说“大气压强是由大气的重量产生的”.两者是否矛盾？给出你的论证.（b）初中教材中介绍，利用天平称量储有干燥空气的烧瓶，可求得空气的重量.有人指出，烧瓶中的空气分子飞舞在空中，不象黄豆那样沉积在烧瓶底部，为什么能称得空气分子的重量呢？给出你对这一问题的论证.

12. 氮原子能够被金属表面吸附，将氮原子从金属表面移至无穷远处需做一定量的功 Φ.在二维金属表面上氮原子的运动是完全自由的，没有相互作用.如果这样的金属表面与压强为 p 的氮气在温度为 T 的情况下处于平衡，试证明在金属表面单位面积上吸附的平均氮原子数为 $(1/\sqrt{2\pi m k_B T})[hp/(k_B T)]\mathrm{e}^{\beta\Phi}$.

13. 被吸附在液体表面的分子形成一种二维气体.考虑分子间的相互作用，试用正则分布证明二维气体的状态方程为 $pS = Nk_B T(1 + B/S)$，其中 $B = -(N/2)\int(\mathrm{e}^{-u/(k_B T)} - 1)2\pi r\mathrm{d}r$，$S$ 为液体面积.

14. 有某种气体，其两分子间作用可表为方形势阱：

$$\varphi(r) = \begin{cases} \infty & r < a \\ -u_0 & a < r < b \\ 0 & r > b \end{cases}.$$

试计算其第二位力系数.

15. 考虑任一反应过程，可将反应过程写为

$$\nu_1 A_1 + \nu_2 A_2 + \cdots + \nu_s A_s = 0, \tag{1}$$

$A_1, A_2, \cdots$ 代表化学组分，$\nu_1, \nu_2, \cdots$ 代表参加反应的各组分的相对摩尔数.对于反应物有 $\nu_i > 0$；对于生成物有 $\nu_i < 0$（例如 $N_2 + O_2 \rightleftharpoons 2N_2O$，可写为 $N_2 + O_2 - 2N_2O = 0$）.

（a）假定系统由单相组成，且维持温度和压强不变，试用自由焓判据证明，系统化学反应的平衡条件为

$$\sum_i \nu_i \mu_i = 0, \tag{2}$$

(b)假定系统处于气相,且可视为 M-B 气体,试从(2)式出发,证明当达到平衡时,各组元浓度之比为

$$\prod_i c_i^{\nu_i} = K_C(T) = \prod_i \left[\frac{1}{V} z_{i,1}(T,V)\right]^{\nu_i}, \tag{3}$$

其中,$c_i = N_i/V$,是组元 i 的浓度. $z_{i,1}$ 是 i 组元的分子配分函数.(3)式叫质量作用定理,是化学中的重要公式. $K_C(T)$ 叫温度平衡常数.

(c)考虑高温下氢原子的离解过程:$e^- + p \rightleftharpoons H$. 设电离能为 ε_0,并设电子,质子和氢原子均可按 M-B 粒子处理,求温度平衡常数.

(d)在很高温度下,氢分子可离解为氢原子,

$$H + H \rightleftharpoons H_2$$

求此反应的温度平衡常数,要求将结果用分子内部配分函数表示.(已知氢原子核的自旋量子数为 1/2,破坏氢分子链的能量为 ε.)

第二章　玻色气体和费米气体

本章研究 F-D 分布和 B-E 分布的应用. 前四节是 B-E 分布的应用，其中第一节讨论黑体辐射问题，即光子统计；第二节讨论固体比热的德拜理论，即声子统计；第三节讨论液体 ^{4}He 的性质和朗道的超流理论，即声子、旋子统计；第四节讨论玻色-爱因斯坦凝聚. 最后一节讨论费米气体的典型例子，即金属中自由电子气对比热的贡献.

§8.2.1　黑体辐射　光子统计

通过第五篇的学习，我们看到黑体辐射的研究对量子论的建立所起的巨大作用. 正是为了能正确解释从经验公式内插得到的黑体辐射的能谱曲线，普朗克提出了能量量子化的概念，推导出了黑体辐射的普朗克公式. 在这一节里，我们首先复习黑体辐射的实验性质，即能谱曲线和斯特藩-玻尔兹曼公式；然后用 B-E 分布推导能谱曲线的解析表达式，即普朗克公式；进而推导出斯特藩-玻尔兹曼公式. 最后用热力学方法研究黑体辐射.

(一)黑体辐射的能谱曲线　斯特藩-玻尔兹曼公式

在 §5.1.1 中已经指出，一个能将投射在它上面的全部电磁波吸收的物体叫绝对黑体. 实验室中的黑体是一个用不透明材料制成的空腔上开的小孔. 如果使空腔维持在一个不变的温度，则经过一段时间以后，腔体内的电磁波形成平衡的辐射场. 这一辐射场通过所开的小孔射出，叫做黑体辐射. 腔体的温度也就是平衡辐射场的温度.

考虑空腔内一个单位体积元，用 u_ν 表示该体积元中频率在 ν 附近单位频率间隔内的能量，称为单色能密度. 用热力学第二定律很容易证明，达到平衡辐射时，u_ν 仅仅是频率 ν 和温度 T 的函数，与腔壁的材料和形状无关，即

$$u_\nu = u(\nu, T). \tag{1.1}$$

19 世纪末，维恩首先用实验测定了单色能密度 u_ν 与 ν、T 的关系，即黑体辐射的能谱曲线. 实验时，首先使黑体保持在某一温度 T，测得与各种频率 ν 对应的单色能密度 u_ν，得到一条 u_ν-ν 曲线，如图 1[参看 §5.6.1 图 1，注意那个图是双对数坐标]. 然后改变黑体的温度，得到另一条 u_ν-ν 曲线. 由图可见，在给定温度下，在低频(即长波)范围内，u_ν 随 ν 增大而增加；在高频(即短波)范围内，u_ν 随 ν 增大而减小，因而存在一个使 u_ν 取最大值的 ν_m. 当温度升高时，ν_m 向高频(即短波)方向移动，服从维恩位移定则

$$\nu_m/T = 常数. \tag{1.2}$$

黑体辐射的另一个重要实验结果是首先由玻尔兹曼用热力学推导，继而由斯特藩用实验建立的斯特藩-玻尔兹曼公式[见后面的(1.26)式]

$$J = \sigma T^4. \tag{1.3}$$

J 称为辐射通量密度，它代表单位时间，从单位黑体表面向 2π 立体角发射的能量. 斯特藩常数的实验值是 $\sigma^{实} = 5.6686 \times 10^{-8}\ \mathrm{J \cdot m^{-2} \cdot s^{-1} \cdot K^{-4}}$. 由几何光学可以证明：

$$J = cu/4. \tag{1.4}$$

其中

$$u = \int_0^\infty u(\nu, T)\,\mathrm{d}\nu. \tag{1.5}$$

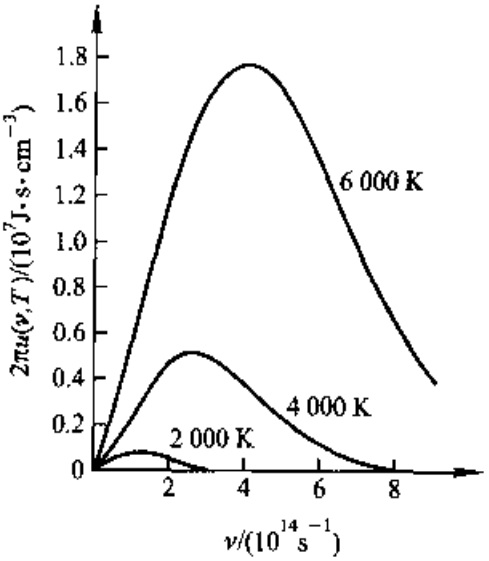


图 1 黑体辐射的单色能量密度

代表黑体辐射的总能密度. 将(1.3)式代入(1.4)式得

$$u = aT^4, \tag{1.6}$$

其中 u 与 σ 的关系是

$$\sigma^{实} = ca^{实}/4. \tag{1.7}$$

显然，由正确的单色能量密度的解析表达式 $u(\nu, T)$，应能得到由实验确定的斯特藩-玻尔兹曼公式和与实验一致的斯特藩常数.

为了导出黑体辐射能谱曲线的解析表达式，可以采取两种不同的观点. 第一种我们不妨称之为波动观点. 在 §3.3.4 中已经证明，由于要满足周期性边界条件，空腔中的电磁场可以分解为具有各种频率 $\nu(0 \sim \infty)$ 和波矢量 $\boldsymbol{k}$ 的单色平面驻波(简正模)的线性叠加. 驻波的能量与具有相同频率的量子谐振子的能量相同，因此整个辐射场的能量可以写为

$$E = \sum_i (n_i + 1/2)h\nu_i, \qquad (n_i = 0, 1, 2, \cdots). \tag{1.8}$$

换句话说，可以把黑体辐射看成是由许多近独立的线性谐振子组成的系统. 这正是普朗克所用的方法. 第二种观点我们称它为光子观点. (1.8)式表明，黑体辐射的能量是不连续的，其变化的最小单位是 $h\nu$，即光子能量. 所以黑体辐射又可以看成是在空腔中的具有各种频率的光子气，其中频率为 ν_i 的光子有 n_i 个. 光子气服从 B-E 分布. 两种物理图像可得到与实验一致的相同结果，这正好说明光的二象性. 作为 B-E 分布的应用，我们按照第二种观点来推导普朗克公式.

(二)光子气的分布函数和普朗克公式的推导

1、光子气的分布函数

我们知道,根据爱因斯坦关系,与光波的频率 ν 、波矢 $\boldsymbol{k}$ 对应,光子的能量为 $\varepsilon = h\nu = pc$,动量为 $\boldsymbol{p} = \hbar\boldsymbol{k}$.由于对应于同一波矢 $\boldsymbol{k}$,电磁振动可以有两个互相垂直的方向,所以光子的运动状态由动量 $\boldsymbol{p}$ 和偏振方向描述.从统计的角度考虑,我们注意光子的以下特点:第一,由于麦克斯韦方程是线性方程,它允许我们把电磁波作线性叠加.这就意味着光子间没有相互作用,光子气是理想气体.第二,由于光子的静止质量为零,所以过渡到 M-B 分布的非简并条件 $(2\pi mk_{\mathrm{B}}T/h^2)^{3/2} \gg N/V$ 总不成立.第三,光子的自旋为 1,是玻色子.第四,由于达到平衡辐射后,仍然存在着腔壁不断吸收和辐射光子的过程,所以与一容器内的气体分子不同,虽然空腔内光子气的能量守恒,但光子数并不守恒.这两者并不矛盾,原因在于发射光子的频率和吸收光子的频率可以不同.例如,当腔壁吸收了一个频率为 $2\nu_0$ 的光子,发射了两个频率为 ν_0 的光子时,空腔内能量不变,而粒子数却发生了改变.空腔中光子数不守恒这一情况使得光子气具有非常重要的特性,即光子气的化学势为零.原因是,由于光子气的光子数 N 是一个变量.因此,根据自由能判据,光子气达到平衡的条件之一是 $(\partial F/\partial N)_{V,T} = 0$,而 $(\partial F/\partial N)_{V,T} = \mu$,所以我们有

$$\mu_{光子} = 0. \tag{1.9}$$

总结以上讨论得到结论,平衡黑体辐射是由具有各种频率的光子组成的,化学势为零的理想玻色气体.因而其分布函数,即每个运动状态上的平均光子数,可以在 §7.3.4(4.10)式中取 $\mu = 0, \varepsilon_\nu = h\nu$ 得到,即

$$\langle n(\nu)\rangle = \frac{1}{\mathrm{e}^{h\nu/(k_{\mathrm{B}}T)} - 1}. \tag{1.10}$$

2.普朗克公式的推导

因为光子的能量和动量满足 $\varepsilon = pc$,同时,给定动量后光子还有两个偏振方向,所以在能量间隔 $\varepsilon \sim \varepsilon + \mathrm{d}\varepsilon$ (或频率间隔 $\nu \sim \nu + \mathrm{d}\nu$)内的运动状态数为

$$g(\varepsilon)\mathrm{d}\varepsilon = 2 \times 4\pi Vp^2\mathrm{d}p/h^3 = [8\pi V/(h^3c^3)]\varepsilon^2\mathrm{d}\varepsilon = (8\pi V/c^3)\nu^2\mathrm{d}\nu. \tag{1.11}$$

辐射场中的单位体积内,频率在 $\nu \sim \nu + \mathrm{d}\nu$ 间的光子数为

$$\frac{1}{V}\langle n(\nu)\rangle g(\nu)\mathrm{d}\nu = \frac{8\pi}{c^3}\frac{\nu^2}{\mathrm{e}^{h\nu/(k_{\mathrm{B}}T)} - 1}\mathrm{d}\nu. \tag{1.12}$$

用 $h\nu$ 乘上式,就得到频率在 $\nu \sim \nu + \mathrm{d}\nu$ 间的能量密度❶

$$u(\nu,T)\mathrm{d}\nu = \frac{8\pi}{c^3}\frac{h\nu}{\mathrm{e}^{h\nu/(k_{\mathrm{B}}T)} - 1}\nu^2\mathrm{d}\nu. \tag{1.13}$$

---

❶ 这里的 $u(\nu,T)$ 即 §5.1.1 中的 ρ_ν ,亦即 §6.2.3 中的 $\rho(\omega)$.

这就是黑体辐射的普朗克公式.

在整个频率范围内,普朗克公式与实验曲线符合得很好.不难看出,在低频区域 $h\nu \ll k_B T$,(1.13)式近似为

$$u(\nu,T)\mathrm{d}\nu = (8\pi/c^3)k_B T\nu^2\mathrm{d}\nu, \tag{1.14}$$

即§5.1.1(1.2)式中提到的瑞利-金斯公式.按照这一公式,u_ν 将随 ν 的增大而趋于发散.这一结论与黑体辐射具有有限能量的实验事实发生尖锐矛盾,也和黑体辐射和腔体达到平衡的事实矛盾(若黑体辐射能量无限,它将不断从腔壁吸收能量,而不能与之达到平衡),历史上称为紫外灾难.正如§5.1.1中所分析的那样,这实际上是经典物理的深刻危机.普朗克正是以此为突破口,创立了量子论.在高频区域 $h\nu \gg k_B T$,(1.13)式过渡到维恩公式[§5.1.1(1.1)式]

$$u(\nu,T)\mathrm{d}\nu = (8\pi/c^3)h\nu^3 \mathrm{e}^{-h\nu/(k_B T)}\mathrm{d}\nu. \tag{1.15}$$

下面,由普朗克公式推导实验给出的维恩位移法则和斯特藩-玻尔兹曼定律.

(三)维恩位移法则和斯特藩-玻尔兹曼定律

将单色能密度 $u(\nu,T)$ 代入极大值条件 $\partial \ln u(\nu)/\partial\nu = 0$ 导致超越方程

$$3\mathrm{e}^x - 3 - x\mathrm{e}^x = 0, \qquad x = h\nu/(k_B T).$$

用计算机很容易求得这个方程的解是(见习题2) $x_m = h\nu_m/(k_B T) = 2.822$,或

$$\nu_m/T = 2.822k_B/h = 常数. \tag{1.16}$$

这正是实验给出的维恩位移法则,这一法则是光测高温计的原理.

将(1.13)式进行积分,可得到黑体辐射中单位体积内的能量密度

$$u(T) = \int_0^\infty u(\nu,T)\mathrm{d}\nu = \frac{8\pi h}{c^3}\int_0^\infty \frac{\nu^3}{\mathrm{e}^{h\nu/(k_B T)} - 1}\mathrm{d}\nu = a^{理} T^4. \tag{1.17}$$

其中的积分可查附录Ⅳ,结果得

$$a^{理} = \frac{8\pi k_B^4}{c^3 h^3}\int_0^\infty \frac{x^3}{\mathrm{e}^x - 1}\mathrm{d}x = \frac{8\pi k_B^4}{c^3 h^3}\frac{\pi^4}{15} = \frac{\pi^2 k_B^4}{15\hbar^3 c^3}. \tag{1.18}$$

将(1.17)式代入(1.4)式,便得到由实验确定的斯特藩-玻尔兹曼定律(1879):

$$J = \sigma^{理} T^4, \tag{1.19}$$

$$\sigma^{理} = \frac{c}{4}a^{理} = \frac{\pi^2 k_B^4}{60\hbar^3 c^2} = 5.6703 \times 10^{-8}\,\mathrm{J\cdot m^{-2}\cdot s^{-1}\cdot K^{-4}}. \tag{1.20}$$

理论计算与实验结果精确吻合.

由(1.17)式立即可得黑体辐射的总能量

$$U = u(T)V = aT^4 V, \tag{1.21}$$

和热容量

$$c_V = 4aVT^3 \propto T^3, \tag{1.22}$$

即在任何温度段,黑体辐射的 c_V 均和 T^3 成正比例.这显然和(1.17)式的积分上限

为无穷大或和黑体辐射具有无穷多个自由度密切相关.

(四)黑体辐射的热力学

在§3.3.2(2.39)式中我们得到了光压与能量密度的关系

$$p = u/3. \tag{1.23}$$

这实际上就是黑体辐射的状态方程.状态方程是用热力学方法研究任何系统时,必须掌握的实验资料.下面把黑体辐射当做热力学系统,选择温度 T 和体积 V 为态参量来计算黑体辐射的内能、熵、自由焓和化学势.

根据实验,黑体辐射的能量密度只是温度的函数,所以黑体辐射的能量可表为

$$U(T,V) = u(T)V. \tag{1.24}$$

利用热力学关系§7.2.5(5.25)式

$$(\partial U/\partial V)_T = T(\partial p/\partial T)_V - p. \tag{1.25}$$

将(1.23)和(1.24)式代入,可得 $u = (T/3)\mathrm{d}u/\mathrm{d}T - u/3$,即 $\mathrm{d}u/u = 4\mathrm{d}T/T$.积分得到

$$u = aT^4, \qquad U = aT^4V. \tag{1.26}$$

a 为积分常数.将(1.26)式代入(1.4)式,可以看到,用热力学方法也能得到斯特藩-玻尔兹曼定律,但是不能确定积分常数.

将(1.23)式和(1.26)式一起代入热力学基本方程 $\mathrm{d}S = (\mathrm{d}U + p\mathrm{d}V)/T$,可得 $\mathrm{d}S = T^{-1}[\mathrm{d}(aT^4V) + (aT^4/3)\mathrm{d}V] = \mathrm{d}[(4aT^3/3)V]$,积分得:

$$S = 4aT^3V/3. \tag{1.27}$$

将(1.23)、(1.26)和(1.27)式代入自由焓的定义式,得

$$G = U - TS + pV = aT^4V - (4/3)aT^4V + (1/3)aT^4V = 0. \tag{1.28}$$

因而 $\mu = 0$,这正是前面用统计物理方法时,考虑到光子数不守恒得到的结果.

从以上计算首先可以看出热力学方法的普适性,即热力学不仅适用于由原子、分子等实物粒子组成的宏观系统,也适用于由场粒子组成的宏观系统.其次,我们还看到了热力学方法的可靠性和简便性.但热力学方法也存在着局限,如不能确定斯特藩常数的值和不能说明 $\mu = 0$ 的物理意义,而统计物理则能提供很好的补充和解释.

§8.2.2　晶格振动比热的德拜理论　声子统计

在§7.3.2中已介绍过固体比热的爱因斯坦理论,并指出该理论在高温下能得出和杜隆-珀蒂定律一致的结论,即 $c_v \approx 6\ \mathrm{cal \cdot mol^{-1} \cdot K^{-1}}$,但在低温下该理论却与实验结果($c_V \propto T^3$)有显著差异.德拜一开始就注意到在足够低的温度下,非金属固体的热容量对温度的依赖和黑体幅射热容量在任何温度段对温度的依赖有

相似的三次幂关系. 于是,在爱因斯坦模型和黑体幅射普朗克理论的启发下,提出了晶格振动比热的德拜模型,爱因斯坦模型认为 $3N$ 个近独立谐振子以相同频率振动,而德拜模型则认为它们具有从 0 到 ν_m 之间的各种频率. 处理黑体辐射的两种观点(波动和光子)自然启发人们想到,晶格振动的比热也可以用两种观点研究,从而引入声子. 本节在声子概念的基础上介绍晶格振动比热的德拜理论. 声子是一种准粒子(又叫元激发). 准粒子概念已成为凝聚态物理中的有力工具.

(一)声子概念的引入

在 §7.3.2 中已经提到,原子在格点附近作微振动时晶格振动能量的表达式可表为

$$E = E_0 + \sum_{i=1}^{3N} (n_i + 1/2)h\nu_i, \qquad (n_i = 0,1,2,\cdots). \tag{2.1}$$

这表明由 N 个有很强相互作用的原子组成的原子晶体,在简谐近似下其晶格振动能量等效于 $3N$ 个近独立简正振动能量之和. 如果把固体看成连续介质(显然,对于那些频率很低,或波长比晶格常数大的简正振动,这种处理是恰当的),这些简正振动就会在固体中传播,与空腔中的电磁波一样. 由于要满足周期性边界条件,结果会形成弹性驻波或声驻波. 于是根据(2.1)式,晶格振动的能量又可看成是 $3N$ 种不同频率驻波的能量之和. 将(2.1)式与上节的(1.8)式比较,我们看到,与电磁场的能量量子化一样,声波场的能量也是量子化的,以 $h\nu$ 为单位来增减能量. 我们把声波场能量变化的最小单位叫做声子. 于是(2.1)式表明,固体中晶格振动的集体表现,等效于一定体积内具有 $3N$ 种不同频率的声子气,其中频率为 ν_i 的声子有 n_i 个. 当某一频率 ν_i 的驻波由 $n_i = 0$ 的能级激发到 n_i 的能级时,我们说产生了 n_i 个能量为 $h\nu_i$ 的声子;反之,我们说消失了 n_i 个能量为 $h\nu_i$ 的声子,参看 §5.4.5 图 1.

与声波的频率 ν 和波矢 $\boldsymbol{k}$ 对应,声子的能量和动量分别为

$$\varepsilon = h\nu, \qquad \boldsymbol{p} = \hbar\boldsymbol{k} = \boldsymbol{e}_n h/\lambda. \tag{2.2}$$

$\boldsymbol{e}_n$ 为动量方向的单位矢量. 根据声波频率和波速的关系 $\nu = v/\lambda$ (v 为声速),可以求得声子的能量和动量的关系为

$$\varepsilon = pv. \tag{2.3}$$

因为固体中有纵波和横波. 用 v_l 和 v_t 分别表示纵波和横波的声速,则对纵声子和横声子分别有:

$$\nu_l = v_l |\boldsymbol{k}| /2\pi, \qquad \nu_t = v_t |\boldsymbol{k}| /2\pi, \tag{2.4}$$

$$\varepsilon_t = |\boldsymbol{p}| v_t, \qquad \varepsilon_l = |\boldsymbol{p}| v_l. \tag{2.5}$$

声子的运动状态由 $\boldsymbol{p}$ 和声振动方向描述. 对于给定的动量 $\boldsymbol{p}$,声子有三个运动状态——一个纵波和两个横波.

从统计的角度考虑，我们注意声子的以下特点：第一，由于在简谐近似下，各种弹性驻波是相互独立的，所以各种能量的声子没有相互作用，即声子气是理想气体. 第二，声子的自旋为零. 第三，由于声子不断产生和消灭，所以声子数不守恒，即声子气的化学势为零. 于是根据 B-E 分布 §7.3.4(4.10)式，在温度 T 时处在能量为 $h\nu$ 的一个运动状态上的平均声子数，即声子的分布函数为

$$\langle n(\nu)\rangle = \frac{1}{e^{h\nu/(k_B T)} - 1}. \tag{2.6}$$

要指出的是，由于声子并不是原子、分子、光子等那样的真实粒子，但又具有真实粒子所具有的基本性质，因此称为准粒子或元激发（参看 §6.1.5 中关于耦合振子系统的讨论）.

根据以上讨论，晶格振动对固体热容量的贡献可以看成是各种不同频率声子贡献的总和. 下面进行具体计算.

（二）德拜频谱和晶格振动的热容量

用 $g(\nu)\mathrm{d}\nu$ 表示频率在 $\nu \sim \nu + \mathrm{d}\nu$ 之间的声子状态总数. 德拜模型的总能量可表为

$$U = U_0 + \int_0^{\nu_{\max}} h\nu\langle n(\nu)\rangle g(\nu)\mathrm{d}\nu. \tag{2.7}$$

因为声子的能量与动量的关系为 $\varepsilon = p\nu$，于是和上节的(1.11)式对照可知，在能量 $\varepsilon \sim \varepsilon + \mathrm{d}\varepsilon$ 之间，纵声子和横声子的状态数分别为

$$\text{纵声子：}\quad 4\pi V p^2 \mathrm{d}p/h^3 = 4\pi V[\varepsilon^2/(h^3 v_l^3)]\mathrm{d}\varepsilon = 4\pi V v_l^{-3}\nu^2\mathrm{d}\nu$$

$$\text{横声子：}\quad 2\times 4\pi V p^2 \mathrm{d}p/h^3 = 8\pi V[\varepsilon^2/(h^3 v_t^3)]\mathrm{d}\varepsilon = 8\pi V v_t^{-3}\nu^2\mathrm{d}\nu.$$

因而在 $\varepsilon \sim \varepsilon + \mathrm{d}\varepsilon$ 或 $\nu \sim \nu + \mathrm{d}\nu$ 之间声子的总状态数为

$$g(\nu)\mathrm{d}\nu = 4\pi V(v_l^{-3} + 2v_t^{-3})\nu^2\mathrm{d}\nu. \tag{2.8}$$

将(2.8)式代入(2.7)式，确定积分限，就可求得德拜模型的总能量.

上节末已指出，由于黑体辐射有无穷多个自由度，所以光子气的频率无限，这导致在任何温度段，黑体辐射热容量均和 T^3 成比例的性质. 但是晶格振动有 $3N$ 个有限自由度，相应地声子气的频率应有上限 $\nu_{\max}$. 但是，在低温时(2.7)式的积分限能视为无穷大，因而仍将得到热容量在低温下与 T^3 成比例的性质. 鉴于以上分析，德拜建议应将声子气的频率在 $\nu_{\max}$ 处截断，并提出在 $0 \sim \nu_{\max}$ 之间的总状态数等于晶格振动的总自由度数，由此来确定 $\nu_{\max}$.

考虑 1mol 固体，用 N_A 表示亚佛加德罗常量有

$$3N_A = \int_0^{\nu_{\max}} g(\nu)\mathrm{d}\nu = \int_0^{\nu_{\max}} 4\pi V(v_l^{-3} + 2v_t^{-3})\nu^2\mathrm{d}\nu = B\int_0^{\nu_{\max}} \nu^2\mathrm{d}\nu. \tag{2.9}$$

其中

$$B = 4\pi V(v_l^{-3} + 2v_t^{-3}). \tag{2.10}$$

由固体的性质决定.将(2.9)式积分得到

$$\nu_{max} = (9N_A/B)^{1/3}. \tag{2.11}$$

(2.10)、(2.11)式给出了最大频率 ν_{max} 与原子密度和声波速度之间的关系.这一结果首先由德拜在 1912 年提出,故称为德拜频谱.声波频谱的截断频率也可以理解为,在晶格点阵中波长小于晶格常数的短波无意义.

1mol 原子晶体中声子气的总能量可表为

$$U = U_0 + \int_0^{\nu_{max}} h\nu\langle n(\nu)\rangle g(\nu)\,d\nu = U_0 + B\int_0^{\nu_{max}} \frac{h\nu^3}{e^{h\nu/(k_BT)} - 1}d\nu. \tag{2.12}$$

其中 U_0 是绝对零度下固体的能量,它包括结合能和零点振动能.引入新的变量 $x = h\nu/(k_BT)$,并定义德拜特征温度

$$\Theta_D = x_{max}T = h\nu_{max}/k_B. \tag{2.13}$$

则(2.12)式化为

$$U = U_0 + 3N_Ak_BT \cdot 3\left(\frac{T}{\Theta_D}\right)^3\int_0^{\Theta_D/T} \frac{x^3}{e^x - 1}dx = U_0 + 3N_Ak_BTD\left(\frac{\Theta_D}{T}\right). \tag{2.14}$$

其中 $D(\Theta_D/T)$ 叫德拜函数,其定义为

$$D(\Theta_D/T) = 3\left(\frac{T}{\Theta_D}\right)^3\int_0^{\Theta_D/T} \frac{x^3}{e^x - 1}dx. \tag{2.15}$$

下面分别考虑两种极限情况.

在高温情况下,即 $x \ll 1$ 或 $T \gg \Theta_D$ 时,有 $e^x \approx 1 + x$,所以有

$$D(\Theta_D/T) \approx 3\left(\frac{T}{\Theta_D}\right)^3\int_0^{\Theta_D/T} x^2dx = 1, \tag{2.16}$$

$$U = U_0 + 3N_Ak_BT, \tag{2.17}$$

$$c_V = 3N_Ak_B = 3R \approx 6\text{cal} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}. \tag{2.18}$$

这正是杜隆-珀蒂定律.

在低温情况下,即当 $x \gg 1$ 或 $T \ll \Theta_D$,积分上限可用 ∞ 代替,我们再次遇到黑体辐射中计算过的积分

$$D(\Theta_D/T) \approx 3\left(\frac{T}{\Theta_D}\right)^3\int_0^{\infty} \frac{x^3}{e^x - 1}dx = \frac{\pi^4}{5}\left(\frac{T}{\Theta_D}\right)^3, \tag{2.19}$$

$$U = U_0 + (3R/5)(\pi^4T^4/\Theta_D^3),$$

$$c_V = (12R/5)\pi^4(T/\Theta_D)^3. \tag{2.20}$$

我们得到了晶格振动低温热容量的 T^3 定律.在一般情况下,热容量的公式可以从(2.14)式得到,即

$$c_V = 36R\left(\frac{T}{\Theta_D}\right)^3\int_0^{\Theta_D/T}\frac{x^3}{e^x-1}dx - 9R\frac{\Theta_D}{T}\frac{1}{e^{\Theta_D/T}-1}. \tag{2.21}$$

可见，如作出 c_V 和 T/Θ_D 的关系图线，则对一切固体应得到同一曲线，即 c_V 满足“对应态定律”.

将实验测得的低温热容量与(2.20)式中的 T^3 的系数比较可以得到德拜温度[当然也可以将任意温度下的实测 c_V 代入(2.21)式来确定 Θ_D]. 另一方面，理论上的德拜温度由(2.11)式和(2.13)给定，可以通过声速(由固体弹性数据得到)计算出来. 表 I 说明，理论计算和实验测定有很好的一致性.

表 I　一些固体的德拜温度

固　体	从热容量得到的 Θ_D (K)	理论计算得到的 Θ_D (K)
NaCl	308	320
KCl	230	246
Ag	225	218
Zn	308	305

虽然固体比热的德拜理论在高温和低温给出了与实验一致的结果，但在中间温度却与实验结果有偏差. 固体比热的精确理论，需要精确算出晶格振动的频谱.

容易证明，1mol 原子晶体的零点振动能为

$$\int_0^{\nu_m}\left(\frac{1}{2}h\nu\right)g(\nu)d\nu = \frac{9}{8}N_A h\nu_{max}.$$

在研究低温下固体对伦琴射线的散射，量度固体的汽化热等实验中，都可以发现这一零点能. 这可以作为量子谐振子零点振动能 $h\nu/2$ 存在的证明. 值得注意的是，在计算黑体辐射时没有计及零点能，这是因为那样将导致 U 发散.

以上关于固体中晶格振动的讨论，对研究相互作用粒子组成的系统是有启发性和代表性的. 下面来作一般性的分析.

根据系综理论，任何系统的平衡性质都可以通过计算巨配分函数来确定，即计算

$$\tilde{Z}(T,V,\mu) = \sum_N\sum_l e^{-\beta(E_{N,l}-\mu N)}. \tag{2.22}$$

其中 $E_{N,l}$ 是 N 个粒子系统的能量. 对于粒子间相互作用很强的系统(如固体)，$E_{N,l}$ 不再是单个粒子能量之和，甚至单个粒子的状态和能量都已失去意义. 因此，严格直接用(2.22)式研究强相互作用粒子系统的热力学量是一条不现实的途径，必须发展近似方法.

从(2.22)式可以看出，由于 $e^{-E_{N,l}/(k_BT)}$ 因子的存在，虽然在任何温度下，高激

发态对配分函数的贡献总是比低激发态小，但是温度越低，高激发态的作用更可以忽略. 换句话说，相互作用系统在低温下的热力学性质，由低激发态能谱决定.

固体中原子处于平衡位置时，系统处于基态. 原子在平衡位置附近作微振动则对应于系统处于低激发态. 固体比热的研究表明，在简谐近似下，利用简正坐标变换(经典力学，见 §1.4.2)和正则变换(量子力学，见 §6.1.5)，可以把相互作用粒子系的低激发态化为一些独立的准粒子或元激发的集合.

$$E = E_0 + \sum_p \varepsilon(p) n(p). \tag{2.23}$$

E_0 表示系统的基态能量，第二项代表低激发能. $\varepsilon(p)$ 表示准粒子或元激发的能量(或称元激发的能谱)，$n(p)$ 表示准粒子数，p 标志准粒子的量子数，如动量 p.

只要知道了元激发的能谱，并能确定准粒子是服从玻色统计还是费米统计，就可利用已知的理想气体分布来处理相互作用粒子系统.

§6.1.5 中已强调指出，准粒子不是组成系统的真实粒子，是系统热运动集体效应的表现. 这里要进一步强调的是，只有在研究相互作用粒子系统的低温性质时，元激发才是一个有效的近似方法，并已得到越来越广泛的应用.

* §8.2.3 液态 ^{4}He 的性质和朗道的超流理论

通过晶格比热的德拜理论的学习，我们看到，理想气体分布不仅可以直接用于相互作用可以忽略不计的真实粒子组成的系统，还可以用于准粒子或元激发系统，而后者是真实粒子间相互作用不能忽略的系统在低温下的等效模型. 元激发的概念在凝聚态物理中已得到越来越多的应用，本节介绍的朗道超流理论是又一个例子.

(一)液体 ^{4}He 的性质

在 §7.2.9 中已经介绍过 ^{4}He 的相图，其突出的特点是：液态 ^{4}He 有两个完全不同的相，即 He Ⅰ 和 He Ⅱ. 其中，He Ⅰ 具有通常液体的性质，而 He Ⅱ 则具有一系列独特的性质：

(1)*超流动性*

实验发现，He Ⅱ 几乎可以毫无阻力地流过直径为 0.1μm 到 1μm 的毛细管. 但当流速超过临界速度 V_c 时(管径愈小 V_c 愈大)，超流性会遭到破坏.

另一方面，如果将细线悬挂的圆盘浸在 He Ⅱ 中，让圆盘作扭转振动，盘的运动将受到阻尼. 用这种方法测得的 He Ⅱ 的粘滞系数和 He Ⅰ 的粘滞系数可以比拟，比用毛细管测得的粘滞系数至少大 10^6 倍. 实验还发现，用扭盘测得的粘滞系数强烈地依赖于温度，且随 $T \to 0$ 而趋于零.

(2)*热力效应和力热效应*

可以用以下示意图来说明热力效应.

两个容器盛有 He Ⅱ，中间用毛细管相连. 如果使两容

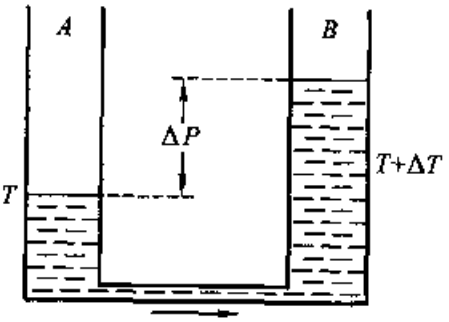


图 1 He Ⅱ 热力效应示意图

器的温度分别保持在 T 和 $T+\Delta T$，He Ⅱ将从温度为 T 的容器流向温度为 $T+\Delta T$ 的容器，直到产生一定的压差来平衡. 这种由温差导致压差的现象称为热力效应.

反之，如果在两容器间维持一定的压差 Δp，He Ⅱ会从压强为 $p+\Delta p$ 的容器经毛细管流向压强为 p 的容器，而使前者升温后者降温. 这种由压差导致温差的现象称为力热效应.

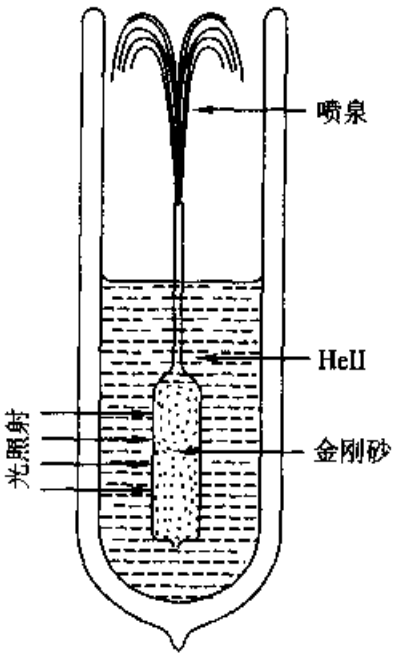


图 2　He Ⅱ的“喷泉”效应

热力效应可通过“喷泉”实验来演示. 如图 2 所示，光照使管中的金刚砂粉末升温，容器中的 He Ⅱ穿过毛细孔，从低温流向高温，产生的“喷泉”可高达 30～40cm.

(3) 良好的导热性

He Ⅱ的热导率非常大，甚至是铜的热导率的近千倍. He Ⅱ的高热导率可以很好地解释一些奇异现象. 如沸腾时的平静和“爬行膜”现象. 1915 年卡末林(Kamerlingh)发现，将一空容器部分浸入 He Ⅱ中时，He Ⅱ会自动地沿器壁爬入容器，直到内、外液面相平为止. 若把容器提起，He Ⅱ又会从容器中沿器壁爬出. “爬行膜”现象可以用 He Ⅱ具有高的导热性和流动性来解释. 通常流体也会沿固体表面形成薄膜，但这层膜在离开液面后会很快消失，原因是膜和液体间存在温差. 但是 He Ⅱ的良好导热性使膜不致消失，兼之 He Ⅱ的良好流动性，会使 He Ⅱ沿膜爬行(见图 3).

(二) He Ⅱ的二流体模型

匈牙利的年轻物理学家迪察(Tisan)首先提出了关于 He Ⅱ的二流体模型. 以后，朗道发展和完善了这一模型. 这是一个唯象的模型，但它能够很好地解释许多实验结果，并预言了 He Ⅱ中的“第二声”.

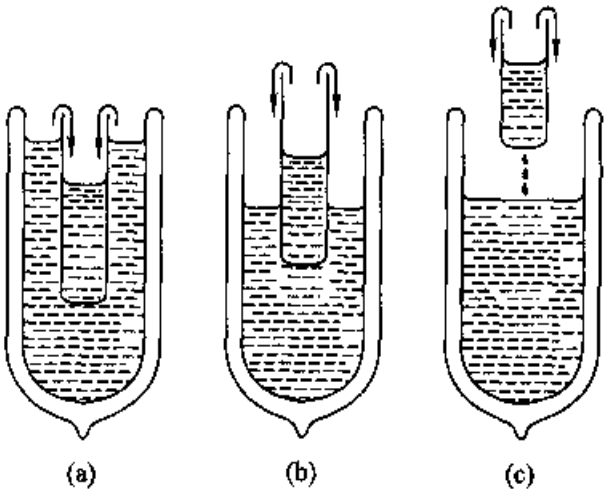


图 3　He Ⅱ的“爬行膜”

二流体模型的基本假设是：

①He Ⅱ是由正常流体和超流体两种成分组成的混合物，即

$$\rho=\rho_s+\rho_n. \tag{3.1}$$

ρ 代表 He Ⅱ的总质量密度，ρ_s 和 ρ_n 分别代表超流成分和正常流体成分的质量密度. 其中超流成分没有粘滞性，熵也为零. 正常流体成分则有粘滞性和熵.

②当 $T=0$K 时，全部 He Ⅱ为超流成分. 当 $T=T_\lambda$ 时，全部 He Ⅱ为正常流体成分. 在 0K 和 T_λ 之间，ρ_s/ρ 和 ρ_n/ρ 是温度的函数，可通过扭盘振动实验来测量.

③用 $v(r)$，$v_s(r)$ 和 $v_n(r)$ 分别代表 He Ⅱ中超流成分和正常流体成分的速度场，则液体的总动量密度为

$$\rho \boldsymbol{v} = \rho_s \boldsymbol{v}_s + \rho_n \boldsymbol{v}_n. \tag{3.2}$$

朗道还假设超流成分的速度场是无旋场，即

$$\nabla \times \boldsymbol{v}_s = 0.$$

两种成分可以相对运动而无内摩擦.

根据二流体模型能解释许多实验结果. 例如，关于 He Ⅱ 通过毛细管时不表现粘滞性可解释为，能穿过毛细管的只是超流成分. 针对这一部分流体来测量，粘滞系数自然为零. 而扭盘实验情况正好相反，只有正常流体成分才对扭盘起阻尼作用，所以这样测得的粘滞系数和 He Ⅰ 的粘滞系数可以比拟.

力热效应是因为只有超流成分才能通过毛细管从高压流向低压. 由于超流成分熵为零，超流部分流过毛细管时，高压部分总质量减少，而总熵不变，故熵密度增大. 根据热力学第二定律，熵密度是无序度的量度. 无序度增加意味着温度升高. 相反，低压部分总质量增加而总熵不变，熵密度减小，从而温度降低. 这样，二流体模型满意地解释了力热效应.

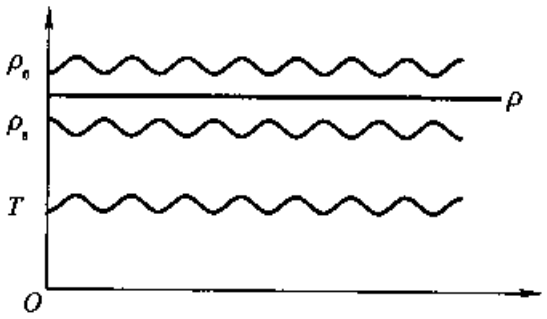


图 4 He Ⅱ 中温度振动的传播

关于 He Ⅱ 的良好导热性，则是由于二流体模型使 He Ⅱ 和普通液体有不同的导热机制. 设想在原来温度均匀的 He Ⅱ 内部，由于不可避免的涨落，某点的温度突然升高，则根据二流体模型，在该点 ρ_n/ρ 会上升，而 ρ_s/ρ 会下降. 为了恢复平衡，热点周围的超流成分会流向该点，该点的正常成分会流离热点，依靠这种进行得很快的“内部运动”来导热，使得 He Ⅱ 比依靠宏观对流来传热的正常液体具有好得多的导热性.

朗道和迪赛还根据二流体模型分别成功地预言了 He Ⅱ 中存在着“第二声”——温度振动（或熵振动）的传播.

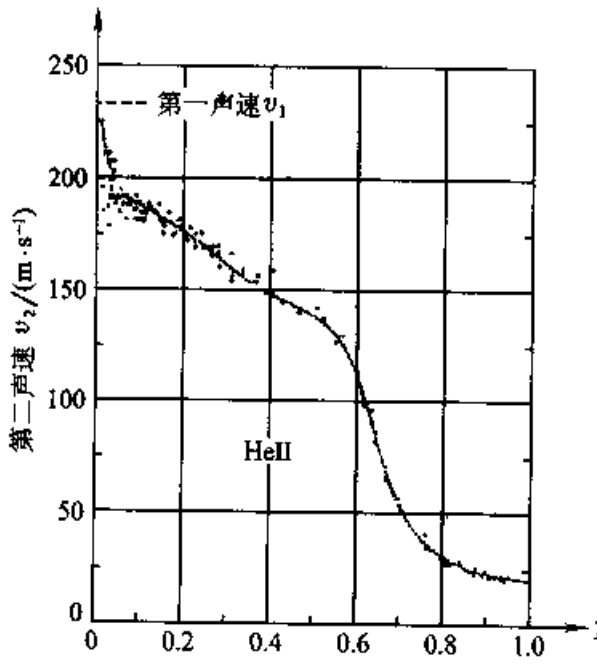


图 5 He Ⅱ 中第二声声速对温度的依赖

因为 He Ⅱ 中存在两种成分 ρ_s 和 ρ_n，朗道认为，假若在各点 ρ_s 和 ρ_n 的振动位相相同，则各点总密度 $\rho = \rho_s + \rho_n$ 会振动，这将产生普通的声波（又叫第一声）. 假若 ρ_s 和 ρ_n 的振动位相相反，则各点总密度基本不变，但是因为超流成分无熵，正常流体成分有熵，故 ρ_n 的振动决定了各处熵密度或温度的振动. 温度振动的传播叫第二声，如图 4 所示.

利用振动的活塞可以在液态 ^{4}He 中产生第一声. 将顺磁盐样品放在稳定的外磁场中，叠加一个交变磁场，顺磁盐的温度振动依靠“内部运动”可以在液 ^{4}He 中产生第二声. 第一个测量第二声的人是

Peshkov. 图 5 上给出的是 De Klerk 等人对第二声速 u_2 的实验结果. 由图可见，在 0.6K 附近，u_2 象其他一些特性一样经历斜率变化. 1K 以上，u_2 基本保持 $18m \cdot s^{-1}$ 不变直到 2K. 这和朗道和迪察的预言（1.5K 左右，u_2 约为 $20m \cdot s^{-1}$）基本相符.

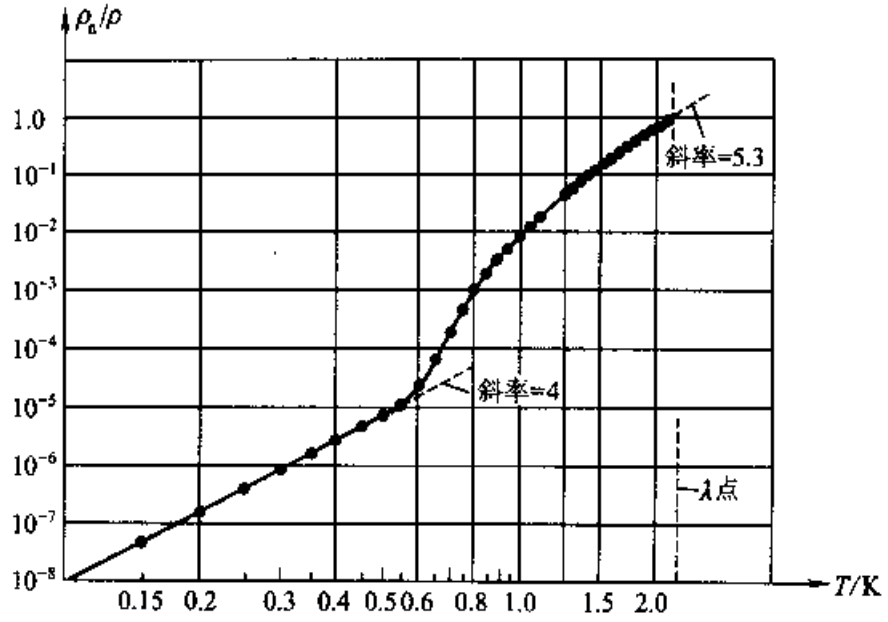


图 6　ρ_n/ρ 随温度的变化

根据二流体模型，仅依靠热力学定律，就可推得第二声速的数学表达式：

$$u_2 = \sqrt{(\rho_s/\rho_n)(TS^2/c_V)}, \qquad S = \int_0^T (c_p dT'/T'). \tag{3.3}$$

将实验测得的 u_2，c_V 和 S 代入上式，可以计算出 ρ_n/ρ 随温度的变化，如图 6 所示. 在 0.6K 附近 ρ_n/ρ 的斜率有不连续变化. 此外，实验还发现在 0.6K 附近，He Ⅱ 的热容量和热传导对温度的依赖都有突然变化.

（三）朗道的超流理论　声子和旋子

二流体模型认为 He Ⅱ 由正常流体成分和超流体成分混合而成，这只是一种形象的描述方法. 不能把 ^{4}He 原子按结构分为两种，但可按所处能态分为两类. 朗道的超流理论对二流体模型作出了微观解释. 朗道超流理论的依据是：首先在 0.6K 附近，$\rho_n/\rho - T$ 曲线的斜率有不连续的变化，c_V 也有突变；其次，在 0K～0.6K 范围中有 $c_V \propto T^3$，这和晶格振动低温比热类似，是声子气的特征，而在 0.6K 附近，c_V 中开始出现和 $e^{-\Delta/T}$ 成比例的附加项，Δ 是一常数.

朗道把温度不十分接近 T_λ 的 He Ⅱ 看成是处于弱激发的玻色系统. 处于基态的原子对应于“超流”成分，而处于低激发态的原子对应于正常流体成分，并可以看成是元激发或准粒子的集合，即玻色理想气体. He Ⅱ 弱激发的能量和动量可表为

$$E = E_0 + \sum_p n(p)\varepsilon(p), \tag{3.4}$$

$$P = \sum_p n(p)p. \tag{3.5}$$

并进一步假设，虽然原则上各种动量 p 的元激发都存在，但在低温热平衡下，元激发主要集中在两个不同的 p 附近，即 $p \to 0$ 和 $p \to p_0$ 附近，如图 7 所示，且在这两个 p 附近，元激发能谱有不同

形式,具体是

$$\varepsilon(p) = \begin{cases} u_1 p, & p \ll p_0, \\ \Delta + (p - p_0)^2/(2m^*), & p \approx p_0. \end{cases} \tag{3.6}$$

u_1 为第一声速,Δ 为常数.朗道称和这两种能谱形式对应的元激发为声子和旋子.朗道的能谱假设已为非弹性中子散射实验证实,见图 7b.根据实验结果确定的参量值为

$$p_0/\hbar = 19.2\text{nm}^{-1}, \qquad \Delta/k_B \approx 8.65\text{K},$$
$$m^* = 0.16 m_{^4\text{He}}. \tag{3.7}$$

其中 $m_{^4\text{He}}$ 为 ^{4}He 原子的质量.由于(3.4)和(3.5)式中的元激发数可以取 0,1,2,…等可能值,因此声子和旋子都服从玻色分布.由于元激发数不确定,它们的化学势为零.在温度为 T 的平衡态,在能量为 ε 的单粒子态上平均元激发数为

$$\langle n(p) \rangle = 1/(e^{\beta\varepsilon(p)} - 1). \tag{3.8}$$

对于声子,上式化为

$$\langle n_{声} \rangle = 1/(e^{\beta u_1 p} - 1). \tag{3.9}$$

对于旋子,在 $T \leqslant 2$ K 的范围内,$\Delta = 8.65 k_B \gg k_B T$,因此(3.8)近似化为

$$\langle n_{旋} \rangle = e^{-\beta\varepsilon(p)}. \tag{3.10}$$

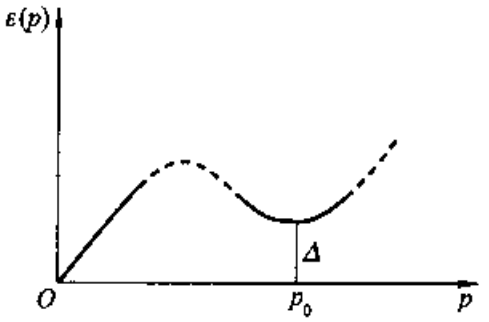
即可近似用玻尔兹曼分布讨论旋子气体.

下面分别计算声子和旋子对热容量 c_V 的贡献.

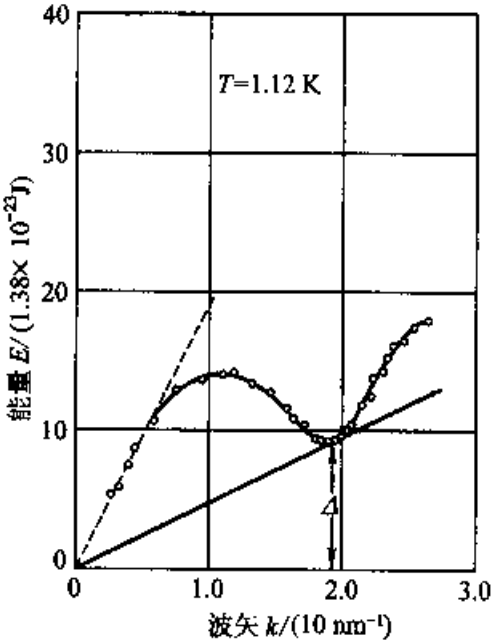
根据系综理论和 B-E 统计知道,对粒子数不守恒的系统($\mu = 0$)有

$$\ln \bar{Z}(T,V,\mu) = -\sum_p \ln(1 - e^{-\beta\varepsilon(p)}),$$
$$\Omega_{\text{巨}} = U - TS - \mu N = F = -k_B T \ln \bar{Z}(T,V,\mu), \tag{3.11}$$
$$U = F - T(\partial F/\partial T)_V, \quad c_V = (\partial U/\partial T)_V. \tag{3.12}$$

只要分别求得声子和旋子的巨配分函数,就可以得到声子和旋子对热容量的贡献.



(a) 朗道假设的 HeII 的元激发能谱



(b) 用中子非弹性散射测得的能谱

图 7 HeⅡ的元激发能谱

考虑热力学极限并注意流体中不存在横声子,因而 $\sum_p$ 可用积分 $(4\pi V/h^3)\int_0^\infty p^2 dp$ 代替,则

$$F_{声} = k_B T(4\pi V/h^3)\int_0^\infty \ln(1 - e^{-\beta u_1 p}) p^2 dp. \tag{3.13}$$

进行分部积分,令

$$X = \ln(1 - e^{-\beta u_1 p}), \qquad p^2 dp = dY,$$
$$dX = \beta u_1/(e^{\beta u_1 p} - 1)\, dp, \qquad Y = p^3/3.$$

并利用 $\int_0^\infty [x^3/(e^x-1)]dx = \pi^4/15$，得到

$$F_{声} = -(4/45)\pi^5 V k_B T[k_B T/(hu_1)]^3.$$

代入(3.12)式最后得

$$U_{声} = (4/15)\pi^5 V k_B T[k_B T/(hu_1)]^3, \tag{3.14}$$

$$c_V^{声} = (\partial U/\partial T)_V = (16/15)\pi^5 V k_B[k_B T/(hu_1)]^3. \tag{3.15}$$

这和德拜模型低温比热对温度有相同的依赖关系. 当 $T \to 0$ 时，$c_V^{声}$ 以 T 的三次幂趋于零.

再计算旋子的热容量，已知旋子服从 M－B 分布，由能量均分定理知，

$$U_{旋} = \langle N_{旋}\rangle(\Delta + k_B T/2), \tag{3.16}$$

$$\langle N_{旋}\rangle = \sum_p \langle n_{旋}(p)\rangle = \sum_p e^{-\beta\varepsilon_{旋}(p)} = \frac{4\pi V}{h^3} e^{-\Delta/(k_B T)} \int_0^\infty e^{-(p-p_0)^2/(2m^* k_B T)} p^2 dp.$$

对积分的贡献主要来自于 $p \approx p_0$ 的部分，故可近似为

$$\langle N_{旋}\rangle \approx \frac{4\pi V}{h^3} e^{-\Delta/(k_B T)} p_0^2 \int_0^\infty e^{-(p-p_0)^2/(2m^* k_B T)} dp.$$

令 $(p-p_0)^2/2m^* k_B T = x^2$ 或 $p = p_0 + \sqrt{2m^* k_B T}\, x$，上式化为

$$\langle N_{旋}\rangle \approx \frac{4\pi V}{h^3} e^{-\Delta/(k_B T)} (2m^* k_B T)^{1/2} p_0^2 \int_{-p_0/\sqrt{2m^* k_B T}}^\infty e^{-x^2} dx.$$

由于在 $T < 2.2\text{K}$ 下有 $p_0^2 \gg 2m^* k_B T$，积分下限可近似取为 $-\infty$，得到

$$\langle N_{旋}\rangle \approx (4\pi V/h^3) e^{-\Delta/(k_B T)} (2\pi m^* k_B T)^{1/2} p_0^2. \tag{3.17}$$

将(3.16)式代入(3.12)式并注意(3.17)式可得

$$c_V^{旋} = \langle N_{旋}\rangle k_B\{3/4 + \Delta/(k_B T) + [\Delta/(k_B T)]^2\}.$$

因为 $\langle N_{旋}\rangle$ 中含有 $e^{-\Delta/(k_B T)}$，所以 $c_V^{旋}$ 随温度的变化主要由此因子决定. 当 $k_B T \ll \Delta$ 时，$c_V^{旋}$ 以指数形式趋于零.

将 $c_V^{声}$ 和 $c_V^{旋}$ 加在一起得 He Ⅱ 的热容量. 具体计算表明在 $T < 0.6\text{K}$ 时，$c_V^{声}$ 起主要作用；在 $T > 1\text{K}$ 时，$c_V^{旋}$ 起主要作用. 在 1.2K 以下，朗道理论的结果和实验符合得很好.

§8.2.4　玻色-爱因斯坦凝聚

现在来研究由质量不为零，粒子数守恒的玻色粒子组成的理想气体. 由于这种粒子不受泡利不相容原理的限制，当 $T \to 0\text{ K}$ 时，几乎所有玻色子会聚集到能量为零，动量为零的基态，这是并不奇怪的. 令人感兴趣的是，研究表明，当温度降低到一个有限的低温 T_C（大约为 3K）时，就会开始有宏观数量的玻色粒子聚集在基态. 这一情况与蒸气凝聚现象有些类似，因而称为玻色-爱因斯坦凝聚（BEC）. 从历史上看，早在 1924—1925 年提出玻色分布时，爱因斯坦就预言了这一奇异现象，但直到 1938 年 F · 伦敦试图把这一现象和 He Ⅰ →He Ⅱ 的相变联系起来，才引起人们的重视. 1995 年科学家利用激光冷却和捕陷技术实现玻色气体中的凝聚后，BEC

再一次引起人们的极大兴趣.

§7.3.4 中已经指出,玻色理想气体化学势的取值范围是 $\mu \leqslant 0$. 对于粒子数不守恒的系统,$\mu = 0$;对于粒子数守恒的系统则有 $\mu < 0$. 再进一步考查粒子数守恒的系统:

$$N = \sum_r \langle n_r \rangle = \sum_{\varepsilon_r} \frac{\Omega(\varepsilon_r)}{e^{(\varepsilon_r-\mu)/(k_BT)} - 1} = \sum_r \frac{1}{e^{(\varepsilon_r+|\mu|)/(k_BT)} - 1}.$$

由于粒子数守恒,上式左边是一常数,子是改变系统温度时,上式右边应保持不变. 已知粒子能级 ε_r 和简并度 $\Omega(\varepsilon_r)$ 只依赖于系统的体积而和温度无关,因而必有 $\mu = \mu(T)$,且当温度降低时,$|\mu|$ 减小或 μ 增大,即有 $\partial\mu/\partial T < 0$. 当温度降低时,化学势将增大,但由于还要满足 $\mu \leqslant 0$,化学势最大只能趋于零. 设当温度降低到某一特殊值 T_C 时,化学势已增大到趋于零,$\mu(T_C) = 0$,继续降低温度,化学势只能保持趋于零而不能再增大(因为 μ 不可能为正),即有 $\mu(T < T_C) \approx 0$. 我们将看到,玻色气体化学势的这一性质是产生 B-E 凝聚的关键.

我们首先计算玻色-爱因斯坦凝聚温度 T_C;然后研究当 $T < T_C$ 时,基态粒子数如何随温度变化和玻色气体的热容量. 最后简单介绍用激光冷却和捕陷技术实现碱金属原子气体 BEC 的进展情况.

(一)凝聚温度 T_C 的计算

为了简单起见,我们考虑自旋为零的玻色子,例如 ^{4}He. 玻色子的能量为 $\varepsilon = p^2/(2m)$,在 $\varepsilon \sim \varepsilon + d\varepsilon$ 之间玻色子的单粒子态数为

$$g(\varepsilon)d\varepsilon = (4\pi V/h^3)p^2dp = 2\pi V(2m/h^2)^{3/2}\varepsilon^{1/2}d\varepsilon. \tag{4.1}$$

考虑到玻色气体的分布函数[§7.3.4(4.10)式],则玻色气体的总粒子数 N 可表为

$$N = \int_0^\infty \langle n(\varepsilon)\rangle g(\varepsilon)d\varepsilon = 2\pi V(2m/h^2)^{3/2}\int_0^\infty \frac{\varepsilon^{1/2}}{e^{(\varepsilon-\mu)/(k_BT)} - 1}d\varepsilon. \tag{4.2}$$

要特别指出的是,因为(4.2)式的被积函数中包含 $\varepsilon^{1/2}$ 的因子,$\varepsilon = 0$ 的能级对积分没有贡献. 即(4.2)式右方代表全部激发态中的态粒子数 $N_{\varepsilon>0}$. 因此,只有当温度比较高,基态粒子数可以忽略不计时,(4.2)式才能成立.

下面我们来分析(4.2)式不能成立的温度条件. 按照前面的讨论,设 $T = T_C$ 时,$\mu(T_C)$ 已增大到最大值,即 $\mu \to 0$. 继续降低温度,μ 将保持趋近于零,(4.2)式不再能成立. 系统的总粒子数大于激发态中的总粒子数,多余的部分必聚集于基态. 所以使化学势 $\mu = 0$ 的温度 T_C 就是(4.2)式成立的最低温度,也就是在总粒子数和体积一定的条件下开始发生 B-E 凝聚的温度. 将 $T = T_C$,$\mu(T_C) = 0$ 代入(4.2)式可得到确定 T_C 的方程,

$$N = 2\pi V\left(\frac{2m}{h^2}\right)^{3/2}\int_0^\infty \frac{\varepsilon^{1/2}}{e^{\varepsilon/(k_B T_C)} - 1}d\varepsilon. \tag{4.3}$$

令 $x = \varepsilon/(k_B T_C)$，上式化为

$$N = 2\pi V\left(\frac{2mk_B T_C}{h^2}\right)^{3/2}\int_0^\infty \frac{x^{1/2}}{e^x - 1}dx. \tag{4.4}$$

其中积分(见附录V)

$$\int_0^\infty \frac{x^{1/2}}{e^x - 1}dx = \frac{\sqrt{\pi}}{2} \times 2.612, \tag{4.5}$$

所以有

$$T_C = [h^2/(2\pi mk_B)][N/(2.612V)]^{2/3}. \tag{4.6}$$

(4.6)式表明，凝聚温度 T_C 由玻色子的质量 m 和数密度 N/V 决定. N/V 愈大，则 T_C 愈高. 以 ^{4}He 为例，其摩尔体积为 $v_{mol} = 27.6\ cm^3 \cdot mol^{-1}$，摩尔质量为 $4g \cdot mol^{-1}$，所以

$$N/V = N_A/v_{mol} = 6.02 \times 10^{23}/27.6 \approx 2.2 \times 10^{22} cm^{-3},$$

$$m = [4/(6.02 \times 10^{23})] \approx 6.6 \times 10^{-24} g,$$

代入(4.6)式得

$$T_C \approx 3.1 K.$$

可将(4.6)式改写为

$$(N/V)_C = 2.612(2\pi mk_B T/h^2)^{3/2}. \tag{4.7}$$

它说明，实现 BEC 有两种途径，除降低给定 N/V 系统的温度外，对温度给定的系统，则可通过提高其密度，使超过临界密度来实现. 在用激光冷却和捕陷技术实现碱金属原子气体的 BEC 实验中，同时考虑了这两种作用. 记住：$\lambda_{热} = (h/\sqrt{2\pi mk_B T})$，$(V/N)^{1/3} = d$，又可把实现 BEC 的条件表为更一般的形式：$\lambda_{热} > d$.

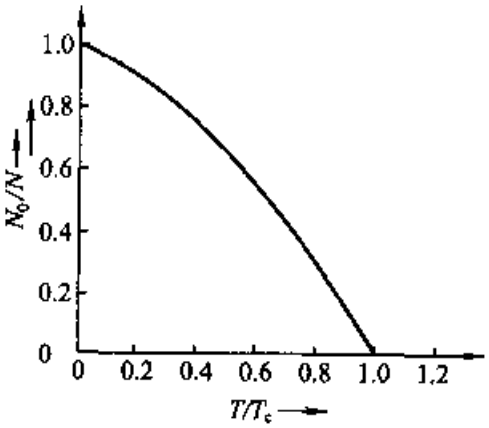


图1　玻色-爱因斯坦凝聚

(二)基态粒子数 N_0 随温度的变化

当 $T < T_C$ 时，基态粒子数已不能忽略，这时代替(4.2)式应有

$$N = N_0 + 2\pi V\left(\frac{2m}{h^2}\right)^{3/2}\int_0^\infty \frac{\varepsilon^{1/2}}{e^{\varepsilon/(k_B T)} - 1}d\varepsilon. \tag{4.8}$$

其中第一项为基态玻色粒子数，第二项为全部激发态中玻色粒子数的总和 $N_{\varepsilon>0}$，其中已取极限 $\mu \to 0$. 首先计算(4.8)式中的第二项. 令 $x = \varepsilon/(k_B T)$，并利用(4.4)式，得

$$N_{\varepsilon>0} = 2\pi V\left(\frac{2mk_BT}{h^2}\right)^{3/2}\int_0^\infty \frac{x^{1/2}}{e^x-1}dx = N\left(\frac{T}{T_C}\right)^{3/2}. \tag{4.9}$$

将(4.9)式代入(4.8)式可求得温度为 T 时处在基态能级上的玻色子数

$$N_0 = N[1-(T/T_C)^{3/2}], \quad (T \leqslant T_C). \tag{4.10}$$

N_0/N 随温度变化的情况如图 1 所示. (4.10)式表明,当 $T < T_C$ 时,玻色子将在基态迅速聚集.凝聚的玻色子数与总玻色子数有相同数量级,T_C 为凝聚温度.要指出的是,B-E 凝聚虽与蒸气凝聚有许多类似之处,但也有根本的区别,这就是 B-E 凝聚是在动量空间而不是在几何空间中的凝聚.聚集在基态的玻色子失去了能量和动量,对压强和粘滞性已无贡献,但仍然分布在全部几何空间中.

(三)玻色理想气体的 c_V

以上说明当 $T < T_C$ 时,虽然已有宏观数量的玻色子聚集在基态,但它们对内能并无贡献.因此,在计算 $T < T_C$ 的内能时,只须考虑激发态中的粒子,即有

$$U = \int_0^\infty \langle n(\varepsilon)\rangle \varepsilon g(\varepsilon)d\varepsilon = 2\pi V(2m/h^2)^{3/2}\int_0^\infty \frac{\varepsilon^{3/2}}{e^{(\varepsilon-\mu)/(k_BT)}-1}d\varepsilon.$$

记住 $T < T_C$ 时有 $\mu \to 0$,上式简化为

$$U = 2\pi V(2m/h^2)^{3/2}\int_0^\infty \frac{\varepsilon^{3/2}}{e^{\varepsilon/(k_BT)}-1}d\varepsilon$$

$$= 2\pi V(2m/h^2)^{3/2}(k_BT)^{5/2}\int_0^\infty \frac{x^{3/2}}{e^x-1}dx. \tag{4.11}$$

将积分求出(查附录Ⅴ),并利用(4.9)式可得

$$U = 0.77Nk_BT(T/T_C)^{3/2}. \tag{4.12}$$

摩尔定体热容量为

$$c_V = 1.925R(T/T_C)^{3/2}. \tag{4.13}$$

当 $T \to 0$ 时,$c_V \to 0$.当 $T < T_C$ 时,$c_V \propto T^{3/2}$.而当 $T = T_C$ 时,有 $c_V = 1.925R > 3R/2$.再考虑到温度很高时,由于 B-E 分布过渡到 M-B 分布,而会有 $c_V = 3R/2$,于是,玻色理想气体的 c_V 随温度变化的情况可用图 2 表示.如果把 $T = T_C$ 时出现的变化看成是发生了某种相变,则在这种相变中 c_V 连续而 c_V 的导数不连续,所以这种相变应属于三级相变.

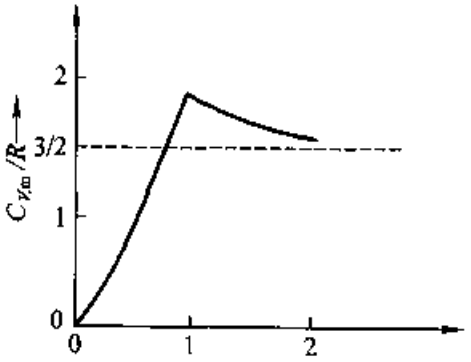


图 2 玻色气体的热容量

(四)BEC 的实验实现

自从 1924 年提出 BEC 以来,科学家一直在寻求 BEC 的实现.首先是 1938 年提出 He Ⅰ↔He Ⅱ可能是 BEC 的一种实现.但由于液氦中粒子间的强相互作用,很难看成是纯的 BEC.1959 年有人提出自旋极化氢原子气体可能是 BEC 的候选者,

但至今未在实验中实现. 近 20 年来,激光冷却和捕陷技术获得了长足发展,人们可以产生较高密度的超冷原子气体,为实现 BEC 提供了新的可能性. 在此基础上,1995 年 JILA,MIT 和 RICE 小组分别用碱金属原子 ^{87}Rb, ^{23}Na 和 ^{7}Li 气体实现了 BEC. 实现 BEC 的实验步骤大体如下:首先利用激光冷却和囚禁技术获得大数目高密度的超冷原子气体,然后将样品装在静磁阱中,再利用射频蒸发冷却技术进一步降低温度,提高密度,最后用光学手段检测是否形成了 BEC.

表 I 列出已报导的三种碱金属原子气体的 BEC 实验参数. 表中下标"C"表示实验中的临界点的值,"0"表示实验中形成的 BEC 的值.

表 I　几种原子气体的 BEC 实验参数.

参数 \ 原子	T_C	N_C	$\left(\frac{N}{V}\right)_C$ (cm^{-3})	$\left(\frac{N}{V}\right)_0$ (cm^{-3})	N_0
^{87}Rb	170nK	2×10^4	2.6×10^{12}	3.2×10^{13}	2×10^3
^{23}Li	300nK	1×10^5	2.0×10^{12}		2×10^4
^{7}Na	2μK	7×10^5	1.5×10^{14}	4.0×10^{14}	5×10^5

§8.2.5　金属中的自由电子气

最简单的金属模型是自由电子气模型. 根据这一模型,金属原子最外层的电子和原子的联系微弱,能够挣脱原子的束缚. 这些脱离了原子束缚的电子在金属中无规则地运动,形成自由电子气体. 我们在 §7.3.4 中已用数字估算表明,由于电子的质量很小,金属中电子数密度很高,在常温下,满足 $\lambda_{热} \gg d$ 的条件,因而金属中的自由电子气是强简并气体. 又因电子的自旋为 1/2,是费米子,所以金属中的自由电子气服从 F-D 分布.

在历史上,金属的自由电子理论是由德鲁德和洛伦兹建立的. 他们把 M-B 分布应用到自由电子气,可以说明金属的导电性和导热性,也可以说明热导率和电导率之比对温度的依赖关系. 但在许多其他问题上,却遇到了严重困难. 例如,根据能量均分定理,每个电子具有 $3k_BT/2$ 的平均能量,1mol 一价金属中有 N_A 个自由电子,总能量为晶格振动能量和自由电子气能量之和,故应为 $3N_Ak_BT + 1.5N_Ak_BT = 4.5RT$. 于是按 M-B 分布的要求,1mol 一价金属的热容量应为 4.5 R. 但实验表明,在室温下,所有金属的热容量均为 3 R(极少例外). 换句话说,在通常温度下几乎测量不到电子气对热容量的贡献. 而且如 §7.3.2 中已指出的那样,金属和非金属比热的比较表明,在低温下,自由电子气对比热的贡献和温度成正比. 1928 年索

末菲提出,代替 M-B 分布,用 F-D 分布来描述自由电子气,成功地解决了上述矛盾,并能很好地解释金属的许多其他特性.

金属中自由电子的能量为 $\varepsilon = p^2/(2m)$,考虑到对应于同一个 p,电子有两个不同的自旋态,所以在能量 $\varepsilon \sim \varepsilon + \mathrm{d}\varepsilon$ 间隔内电子的单粒子态数为

$$g(\varepsilon)\mathrm{d}\varepsilon = 2 \times (4\pi V/h^3)p^2\mathrm{d}p = 4\pi V(2m/h^2)^{3/2}\varepsilon^{1/2}\mathrm{d}\varepsilon. \tag{5.1}$$

根据费米分布,即 §7.3.4(4.9)式,在体积 V 内,自由电子的总数和总能量为

$$N = \int_0^\infty \langle n(\varepsilon)\rangle g(\varepsilon)\mathrm{d}\varepsilon = 4\pi V\left(\frac{2m}{h^2}\right)^{3/2}\int_0^\infty \frac{\varepsilon^{1/2}}{\mathrm{e}^{(\varepsilon-\mu)/(k_BT)}+1}\mathrm{d}\varepsilon, \tag{5.2}$$

$$U = \int_0^\infty \varepsilon\langle n(\varepsilon)\rangle g(\varepsilon)\mathrm{d}\varepsilon = 4\pi V\left(\frac{2m}{h^2}\right)^{3/2}\int_0^\infty \frac{\varepsilon^{3/2}}{\mathrm{e}^{(\varepsilon-\mu)/(k_BT)}+1}\mathrm{d}\varepsilon. \tag{5.3}$$

化学势 μ 由(5.2)式确定为 N,V,T 的隐函数 $\mu = \mu(N,V,T)$. 对于给定的一块金属来说,V 和 N 是确定的,于是化学势 μ 是温度的函数. 定义绝对零度时的化学势为费米能 ε_F:

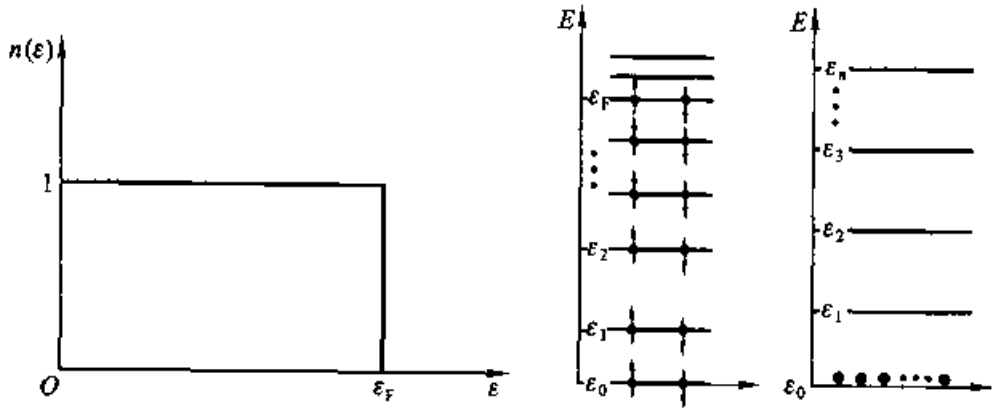
$$\varepsilon_F \equiv \mu(0). \tag{5.4}$$

下面很快会看到 ε_F 的物理意义. 分析(5.2)式可以看出 ε_F 为正 $(\varepsilon_F > 0)$,和 $\mu(T > 0) < \varepsilon_F$,否则将会得到无意义的结果(见习题 13).

(一)0K 下的自由电子气

由 F-D 分布函数 §7.3.4(4.9)式可知,在 0K 时有

$$\langle n(\varepsilon)\rangle = \begin{cases} 1 & \varepsilon < \varepsilon_F, \\ 0 & \varepsilon > \varepsilon_F, \end{cases} \tag{5.5}$$



(a) 电子气的平均占有数　(b) 电子气　(c) 玻色气

图 1　0K 时电子气的平均占有数[图(a)]及电子气和玻色气的能级填充[图(b)、(c)]

如图 1(a)所示. (5.5)式表明在绝对零度下,对于能量小于费米能 ε_F 的那些能级,每个单粒子态的平均电子数为 1,即按照不相容原理,每个单粒子态已被填满. 能量大于 ε_F 的那些能级,单粒子态上的平均电子数为零,即每个单粒子态都是空的,见图 1(b). 这种分布情况与 0K 时玻色子全部聚集在基态大不相同[图 1(c)]. 这

很容易用泡利不相容原理解释. 在 0K 时,电子将尽可能占据能量最低的状态,但由于不相容原理的限制,每个单粒子态最多只能容纳一个电子. 于是,当基态被占据后,电子将被迫填充到第二能级、第三能级…中去,直到全部自由电子填充完毕为止. 所以费米能 ε_F 是 0K 时电子占据的最高能级的能量值. 费米能的大小可由(5.2)式确定. 因为在0K,仅 ε_F 以下的单粒子态被电子占有,且 $\langle n(\varepsilon)\rangle = 1$,故这时(5.2)式化为

$$N = 4\pi V\left(\frac{2m}{h^2}\right)^{3/2}\int_0^{\varepsilon_F}\varepsilon^{1/2}\mathrm{d}\varepsilon = 4\pi V\left(\frac{2m}{h^2}\right)^{3/2}\frac{2}{3}\varepsilon_F^{3/2}, \tag{5.6}$$

或

$$\varepsilon_F = [h^2/(2m)][3N/(8\pi V)]^{2/3}. \tag{5.7}$$

即:费米能 ε_F 与单位体积中自由电子数的 2/3 次方 $(n^{2/3})$ 成正比.0K 时自由电子气的总能量和压强也可用费米能表示出来:

$$U_0 = 4\pi V\left(\frac{2m}{h^2}\right)^{3/2}\int_0^{\varepsilon_F}\varepsilon^{3/2}\mathrm{d}\varepsilon = 4\pi V\left(\frac{2m}{h^2}\right)^{3/2}\frac{2}{5}\varepsilon_F^{5/2} = \frac{3}{5}N\varepsilon_F. \tag{5.8}$$

每个电子的平均能量是

$$\langle\varepsilon_0\rangle = U_0/N = 3\varepsilon_F/5, \tag{5.9}$$

$$C_V = (\partial U_0/\partial T)_V = 0, \tag{5.10}$$

$$p = -(\partial U_0/\partial V)_S = -(\partial U_0/\partial V)_T = (2/5)(N/V)\varepsilon_F. \tag{5.11}$$

(5.11)式中已考虑 $T = 0$ K 时,绝热过程即等温过程.

下面我们以银(Ag)为例对 ε_F 的大小作数字估计. Ag 为一价元素,其相对原子质量为 107.88,密度为 10.49g · cm^{-3},则

$$n = N/V = (10.49/107.88)\times 6.02\times 10^{23}\times 1 = 5.81\times 10^{22}\mathrm{cm}^{-3},$$

代入(5.7)式可得 $\varepsilon_F = 5.4$ eV,与室温(300K)下粒子热运动能量的典型值 $k_BT = 0.025$ eV 相比,约高 200 倍. 若用 v_F 代表与 ε_F 对应的自由电子速度,则可得 $v_F = (2\varepsilon_F/m)^{1/2} \approx 10^8$ cm · s^{-1},比室温下气体分子热运动的平均速度每几百 m · s^{-1}大得多. 若将以上数据代入(5.11)式可以得到 $p_0 \approx 10^4$ atm. 这些数字估算表明,由于泡利不相容原理,即使在绝对零度,自由电子也不是"沉积"在最低能级,而是被迫分布在激发态,以致自由电子气仍然非常活跃,而且自由电子的数密度 n 越大, ε_F 就越高, v_F 和 p_0 也越大.

表 1 一些金属和白矮星的电子密度、费米能、费米温度和相应的自由电子速度

元 素	N/V (cm^{-3})	ε_F (eV)	T_F (K)	v_F ($cm \cdot s^{-1}$)
Li	4.6×10^{22}	4.7	5.5×10^4	1.3×10^8
K	1.34×10^{22}	2.1	2.4×10^4	0.85×10^8
Cu	8.50×10^{22}	7.0	8.2×10^4	1.56×10^8
Au	5.90×10^{22}	5.5	6.4×10^4	1.39×10^8
白矮星	10^{30}	10^7	10^{11}	中心温度 10^7K

通常引入费米温度

$$T_F \equiv \varepsilon_F / k_B, \tag{5.12}$$

表 1 给出了一些金属和白矮星的费米能和费米温度. 金属的费米温度都在 10^4K 以上. 对于室温 T 则有 $T \ll T_F$ 或 $k_B T \ll \varepsilon_F$.

(二) $T < T_F$ 时的自由电子气

当 $T \neq 0$ 时,由 F-D 分布得

$$\langle n(\varepsilon) \rangle = \frac{1}{e^{(\varepsilon-\mu)/(k_B T)} + 1} \begin{cases} > 1/2 & \varepsilon < \mu, \\ = 1/2 & \varepsilon = \mu, \\ < 1/2 & \varepsilon > \mu. \end{cases} \tag{5.13}$$

注意到 $e^{(\varepsilon-\mu)/(k_B T)}$ 按指数规律随 ε 变化,则实际上只在 μ 附近数量级为 $k_B T$ 的范围内自由电子的分布情况才与 0K 时的分布有差异. 对于我们所感兴趣的室温,有 $k_B T \ll \varepsilon_F$ 或 $T \ll T_F$,分布曲线如图 2 所示(见习题 19). 分布曲线的这种变化可以解释如下. 在 ε_F 附近 $\Delta\varepsilon \approx k_B T$ 范围内的电子由于具有 $k_B T$ 的热激发能,可以跃迁到更高的未被占据能级上去,而更低能级上的电子跃迁的概率很小,因而分布曲线只在 ε_F 附近宽 $2k_B T$ 的范围内有变化. 换句话说,只有 ε_F 附近 $k_B T$ 范围内的自由电子才参与热运动. 这是简并费米气最重要的特点.

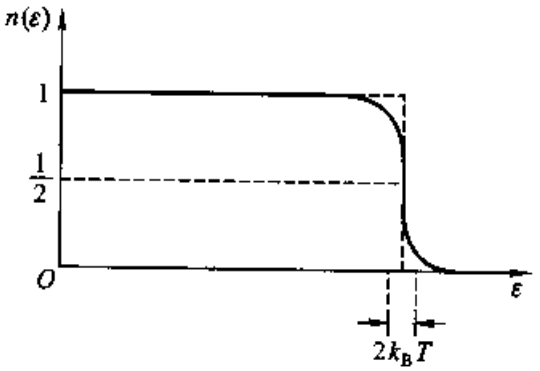


图 2 $k_B T \ll \varepsilon_F$ 时电子气的分布

在定量计算之前,我们先根据这一特点,对自由电子气的比热作定性说明. 从图 2 看到,参与热运动的自由电子数与总自由电子数之比大约为 $k_B T/\varepsilon_F = T/T_F$,因而 1mol 一价金属中对热容量作出贡献的有效自由电子数大致等于 $N_A k_B T/\varepsilon_F =$

$N_A T/T_F$. 再考虑每一个自由电子平均具有 $k_B T$ 的热激发能，故有效自由电子的总激发能为 $U \approx N_A k_B T(k_B T/\varepsilon_F)$，其对金属热容量的贡献大约是

$$c_V^{电} = R(k_B T)/\varepsilon_F = RT/T_F. \tag{5.14}$$

已知室温下有 $T \ll T_F$，故 $c_V^{电} \ll R$. 因而在通常温度下，电子气对热容量的贡献完全可以忽略不计. 但是在低温下，根据德拜理论，晶格振动对热容量的贡献 $c_V^{振}$ 按 T^3 减小，比 $c_V^{电}$ 随温度降低的减小($\sim$T)快得多. 在 3K 以下，$c_V^{电}$ 开始突出，甚至比 $c_V^{振}$ 更大.

下面再来对 $C_V^{电}$ 作定量计算.

考虑(5.7)式后，$\varepsilon \sim \varepsilon + d\varepsilon$ 间的单粒子态数可写为

$$g(\varepsilon)d\varepsilon = (3/2)N\varepsilon_F^{-3/2}\varepsilon^{1/2}d\varepsilon.$$

总电子数和总能量为

$$N = \frac{3}{2}N\varepsilon_F^{-3/2}\int_0^{\infty}\frac{\varepsilon^{1/2}}{e^{(\varepsilon-\mu)/(k_BT)}+1}d\varepsilon, \tag{5.15}$$

$$U = \frac{3}{2}N\varepsilon_F^{-3/2}\int_0^{\infty}\frac{\varepsilon^{3/2}}{e^{(\varepsilon-\mu)/(k_BT)}+1}d\varepsilon. \tag{5.16}$$

从(5.15)式可求得 $\mu(T)$，代入(5.16)式可得 $U(T)$，进而得 C_V.

(5.15)，(5.16)式要求计算同一类型的积分，即

$$I = \int_0^{\infty}\frac{1}{e^{(\varepsilon-\mu)/(k_BT)}+1}f(\varepsilon)d\varepsilon. \tag{5.17}$$

这叫做费米-狄拉克积分. 其中 $f(\varepsilon)$ 是满足 $f(0) = 0$ 的任意函数. 在附录Ⅴ中给出了这一积分的结果：

$$I = \int_0^{\mu}f(\varepsilon)d\varepsilon + \frac{\pi^2}{6}(k_BT)^2f'(\mu) + \frac{7\pi^4}{360}(k_BT)^4f'''(\mu) + \cdots, \tag{5.18}$$

取 $f(\varepsilon) = \varepsilon^{1/2}$，并考虑一级近似，有

$$I = \int_0^{\mu}\varepsilon^{1/2}d\varepsilon + \frac{\pi^2}{6}(k_BT)^2\left(\frac{1}{2}\right)\mu^{-1/2} + \cdots = \frac{2}{3}\mu^{3/2} + \frac{\pi^2}{12}(k_BT)^2\mu^{-1/2} + \cdots.$$

代入(5.15)式得

$$\mu^{3/2} \approx \varepsilon_F^{3/2}[1 + (\pi^2/8)(k_BT/\mu)^2 + \cdots]^{-1},$$

或

$$\mu \approx \varepsilon_F[1 - (\pi^2/12)(k_BT/\mu)^2 + \cdots].$$

采取迭代近似法，将上式括号中的 μ 用 ε_F 替代，最后得低温下自由电子气的化学势为

$$\mu \approx \varepsilon_F[1 - (\pi^2/12)(k_BT/\varepsilon_F)^2 + \cdots]. \tag{5.19}$$

(5.19)式表明，当 T 升高时，μ 减小，但在 $T \ll \varepsilon_F/k_B$ 的条件下，μ 与 ε_F 近似相等，

如 $T = 300\ \mathrm{K}$ 时，$\mu \approx 0.9916\varepsilon_{\mathrm{F}}$. 再取 $f(\varepsilon) = \varepsilon^{3/2}$，则

$$I = \int_0^{\mu} \varepsilon^{3/2} \mathrm{d}\varepsilon + \frac{\pi^2}{6}(k_{\mathrm{B}}T)^2 \frac{3}{2}\mu^{1/2} + \cdots = \frac{2}{5}\mu^{5/2} + \frac{3\pi^2}{12}(k_{\mathrm{B}}T)^2(\mu)^{1/2} + \cdots,$$

代入(5.16)式得

$$\begin{aligned} U &= \frac{3}{2}N\varepsilon_{\mathrm{F}}^{-3/2}\left\{\frac{2}{5}\mu^{5/2} + \frac{3\pi^2}{12}(k_{\mathrm{B}}T)^2(\mu)^{1/2} + \cdots\right\} \\ &= \frac{3}{5}N\left(\frac{\mu}{\varepsilon_{\mathrm{F}}}\right)^{3/2}\mu\left\{1 + \frac{5\pi^2}{8}\left(\frac{k_{\mathrm{B}}T}{\mu}\right)^2 + \cdots\right\} \\ &= \frac{3}{5}N\varepsilon_{\mathrm{F}}\left\{1 + \frac{5\pi^2}{12}\left(\frac{k_{\mathrm{B}}T}{\varepsilon_{\mathrm{F}}}\right)^2 + \cdots\right\}, \end{aligned} \tag{5.20}$$

推导中已利用(5.19)式

$$C_V^{\text{电}} = \frac{\partial U}{\partial T} = \frac{\pi^2}{2}Nk_{\mathrm{B}}\frac{k_{\mathrm{B}}T}{\varepsilon_{\mathrm{F}}}. \tag{5.21}$$

考虑 1mol 一价金属，$N = N_{\mathrm{A}}$，则

$$c_V^{\text{电}} = \frac{\pi^2}{2}R\frac{k_{\mathrm{B}}T}{\varepsilon_{\mathrm{F}}} = \frac{\pi^2}{2}R\frac{T}{T_{\mathrm{F}}}. \tag{5.22}$$

可见，定性估计与理论计算只有系数的差异.

习 题

1. 推导二维空间的普朗克公式，并由此导出二维空间的斯特藩-玻尔兹曼定律及光压公式.

2. 用 $u(\lambda)\mathrm{d}\lambda$ 代表单位体积中，在波长 $\lambda \sim \lambda + \mathrm{d}\lambda$ 范围内，温度 T 时黑体辐射的能量. (a) 问在什么波长时，$u(\lambda, T)$ 具有极大值？[要求用计算机解超越方程.] (b) 太阳辐射的电磁场与黑体辐射具有非常类似的能谱. 如果其 $\lambda_m = 4800\ \text{Å}$，问太阳表面的温度是多少？(c) 从 §5.1.1 图 1 上粗略测出其 ν_m 的值，据此估算该图中的曲线所对应的温度是多少？

3. 利用 $\nu\lambda = c$，可将 $u_\nu(T)\mathrm{d}\nu$ 化为 $u_\lambda(T)\mathrm{d}\lambda$，问能否应用 $\nu_{\max}\lambda = c$ 求得 λ_m（λ_m 由 $\mathrm{d}u_\lambda(T)/\mathrm{d}\lambda = 0$ 得到）？为什么？

4. 试计算光子气的配分函数，并利用配分函数求黑体辐射的内能和压强.

5. 在固态时，碘与碲原子均排成平行的长键. 证明这些物质在低温下的比热与绝对温度 T 成正比. 原子成平面排列的石墨，在低温时的比热又如何呢？

6. 证明二维理想玻色气体不存在玻色-爱因斯坦凝聚现象.

7. 质量为 m 的玻色子组成理想气体，$T < T_{\mathrm{C}}$ 时内能 $U = 0.77Nk_{\mathrm{B}}T(T/T_{\mathrm{C}})^{3/2}$. 求这种气体在 $T < T_{\mathrm{C}}$ 时的熵、自由能和压强，并将压强与 T 的关系与临界温度以下气、液等温转变时的情况进行比较，说明结果的意义.

8. 设在某样品中，当 $\varepsilon > 0$ 时电子的态密度 $g(\varepsilon) = D$（常数），当 $\varepsilon < 0$ 时，电子的态密度 $g(\varepsilon) = 0$，总电子数等于 N. (a) 证明 0K 时的化学势 $\mu(0) = N/D$. (b) 证明当系统强简并（T

$\ll \mu/k_B$）时，化学势为 N/D 与 T 无关，热容量为 $C_V = \pi^2 k_B^2 DT/3$．(c)证明系统非简并（过渡到M-B分布）的条件是 $N \ll Dk_BT$，说明这一条件的意义．

9. 钠的原子量为23，密度为0.95g·cm^{-3}，原子价为1．(a)计算费米温度 $T_F = \varepsilon_F/k_B$．(b)如果用液态 ^{4}He 的蒸发来冷却钠金属，试计算使100cm^3 的钠由1K降低到0.3K需要蒸发多少立方厘米的液体 ^{3}He．已知：在这样低的温度下，晶格振动的比热与电子比热比较可以忽略，蒸发1cm^3 的液体需要0.8J的热量．

10. 在某些星体内部的条件下（如白矮星），电子的平均能量变得可以与 $m_e c^2$ 相比较．（其中 m_e 是电子的质量，c 是光速），以致电子服从相对论力学．电子的能量和动量的关系为 $\varepsilon = pc$，试计算绝对零度时极端相对论理想电子气体的费米能、内能和压强．

11. 由于自旋，电子具有磁矩 μ_B．把金属放在外场 H 中时，对应于自旋平行或反平行于外场，自由电子的能量为 $\varepsilon = p^2/(2m) \mp \mu_B H$．(a)用 N_+ 和 N_- 分别代表自旋平行于和反平行于外场的自由电子数，写出由于电子磁矩而引起的磁化强度 M 的表达式．(b)分别写出 N_+ 和 N_- 的表达式（包含费米分布函数的一个积分）．(c)证明在绝对零度及 $\mu_B H \ll \varepsilon_F$ 的条件下有

$$N_\pm = \frac{4\pi V}{3h^3}(2m\varepsilon_F)^{3/2}\left(1 \pm \frac{3}{2}\frac{\mu_B H}{\varepsilon_F}\right), \qquad M = \frac{3}{2}\frac{N}{V}\frac{\mu_B^2 H}{\varepsilon_F}.$$

12. 为了从金属中逸出，自由电子在金属表面必需克服一高度为 x 的势垒，也就是说，只有当电子垂直于金属表面的速度分量 $v_\perp \geqslant \sqrt{2x/m}$ 时，才能从金属中逸出．证明当温度为 T［$(x-\mu) \gg k_BT$，μ 为电子气的化学势］时，单位时间中从单位金属表面逸出的电子数为 $(4\pi m/h^3)(k_BT)^2 e^{(\mu-x)/(k_BT)}$．

13. 试证明费米能 $\varepsilon_F > 0$ 和 $\mu(T>0) < \varepsilon_F$．

14. 宇宙背景中充满了 $T = 3$ K 的黑体辐射，可把它看成宇宙大爆炸的余迹．试求出光子数密度和 T 的关系，并近似计算 $T = 3$ K 时的光子数密度．

15. 设在一体积 V 的容器中充满温度 $T = 1000$ K 的电磁辐射．将容器的体积 V 可逆绝热膨胀到 $8V$，问电磁辐射的末温是多少？设容器的热容量可忽略不计．

16. 已知 $u_\lambda(T)\mathrm{d}\lambda = 8\pi\dfrac{hc}{\lambda^5}\dfrac{1}{e^{hc/(k_B\lambda T)}-1}\mathrm{d}\lambda$

(a)以 $\ln u_\lambda$ 和 $\ln\lambda$ 为纵、横坐标，就 T =100K，200K 和 400K，用 math 画出普朗克曲线．

提示　能量以J为单位，λ 以μm为单位．

(b)维恩位移法则在此双对数图上如何表现？

(c)从理论上证明 $\ln u(\lambda_m)$ 和 $\ln\lambda_m$ 的关系和(b)问结果一致．

17. 考虑由 $N = 10000$ 个处于三维各向同性谐振子势中的粒子组成的理想气体，粒子的能量可表为 $\varepsilon = (n_x + n_y + n_z)\hbar\omega$，可以证明，和 $\varepsilon = m\hbar\omega$ 对应的简度为 $(m+1)(m+2)/2$．假定粒子服从B-E分布，并取 $\hbar\omega = 1$，试用 math．

(a)确定 $k_BT_C/\hbar\omega$ 的近似值，T_C 代表玻色-爱因斯坦凝聚（BEC）温度．

(b)求当 $T < T_C$ 时，基态（$m = 0$）上的平均粒子数 $\langle N_0\rangle$，并绘出 $\langle N_0\rangle$ 对 T 的依赖曲线．

(c)绘出 $T < T_C$ 时，激发态上的平均粒子数 $\langle N_m\rangle$ 对 m 的依赖曲线．

(d)计算 $T < T_C$ 时，逸度 $z = e^{-\mu/(k_BT)}$ 之值，并绘出 $|\mu|/(k_BT)$ 对 T/T_C 的依赖曲线（包括

$T < T_C$ 和 $T > T_C$).

(e)计算系统的能量 U ,并绘出 U/N 对 T/T_C 的依赖曲线.

(f)根据 (e) 用数据计算并绘出单粒子热容量 C_v/N 对 T/T_C 的依赖曲线.

18. 考查和上题完全相同的系统,但假定粒子服从 M-B 分布.

(a)取 $\hbar\omega = 1, k_B T = 25$ (任意单位),求粒子的配分函数.

(b)求能级 $\varepsilon = m\hbar\omega$ 上的平均粒子数 $\langle N_m \rangle$,并绘出 $\langle N_m \rangle$ 对 m 的依赖曲线.

(c)给出基态($m = 0$)上的平均粒子数 $\langle N_0 \rangle$.

(d)取 $k_B T = 15$,重复(a),(b)及(c)步.

(e)M-B 气体会发生"凝聚"吗? 由 $\langle N_m \rangle \sim m$ 曲线看出,在高温下 M-B 气体和 B-E 气体有怎样的关系?

19. 可将 F-D 分布改写为

$$\langle n(\varepsilon) \rangle = \frac{1}{e^{(\varepsilon-\mu)/(k_B T)} + 1} = \frac{1}{e^{\mu(x-1)/(k_B T)} + 1}, \qquad x = \varepsilon/\mu.$$

(a)先后取 $\mu/(k_B T) = 20, 50$,用 math 绘出 $n(\varepsilon)$ 和 $\mathrm{d}n(\varepsilon)/\mathrm{d}\varepsilon$ 对 ε/μ 的关系曲线.

(b)根据(a)问所给曲线说明 F-D 分布随温度升高而变化的情况.



第九篇　涨落和临界现象的统计理论

涨落是统计物理学的一个重要研究对象. 从理论上看，涨落过程的研究有力地证实了统计物理学的规律性，使人们深刻地认识到热力学量的统计性质；从实验上看，一些自然现象要用涨落理论来解释，涨落在现代科学技术中得到了广泛应用. 研究表明，在平衡态邻域，广延量和强度量的相对涨落都和 $\sqrt{N}$ 成反比（N 是宏观系统中的粒子数）. 因此，对于宏观单相系统，热力学量的相对涨落一般观察不到. 实际上，只有在两种情况下，可以观察到涨落. 第一，系统尺度足够小的情形，如布朗粒子；第二，宏观双相系统处于临界点时，一些物理量的涨落及涨落的关联会反常增大，并被观察到（如气－液系统中的临界乳光就是密度涨落很大的展现）. 这是系统内部进行调整，孕育新相的反映，是关于临界现象的标度理论和重整化群方法的基础.

本篇第一章介绍涨落理论，包括：涨落的准热力学理论和布朗运动理论. 第二章则在 §7.2.9 的基础上学习临界点的特征，进而研究临界点的涨落和涨落关联. 第三章围绕 Ising 模型研究相变的统计理论，并介绍重整化群方法的大意.

第一章　涨落理论

§9.1.1　涨落的准热力学理论

统计物理学的基本观点是，宏观量是相应微观量的统计平均. 但是，当系统处于平衡态时，物理量 A 并非总是等于平均值 $\langle A\rangle$，而是围绕平均值 $\langle A\rangle$ 涨落. 涨落的大小可用均方涨落 $\langle(\Delta A)^2\rangle = \langle(A-\langle A\rangle)^2\rangle = \langle A^2\rangle - \langle A\rangle^2$ 描述，其中 $\langle A^2\rangle = \sum A^2P(A)$，$\langle A\rangle = \sum AP(A)$. 因此，只要知道某个宏观物理量所对应的微观量的概率分布，就可以计算这个物理量的均方涨落. 在 §7.3.1 和 §7.3.3 中从这一基本观点出发，求得了恒温系统的能量涨落，及开放系统的粒子数涨落. 在这一章里，要推导一个对处于任何宏观条件下的系统（如 T,V,N 一定的恒温系统. T,μ,V 一定的开放系统，或 T,p,N 一定的系统）都适合的，基本热力学量涨落的概率分布公式. 利用这个普遍公式，热力学量（包括没有对应微观量的物理量如温度和化学势）的涨落及涨落的关联可以不通过系综平均就方便地计算出来. 这种方法叫涨落的准热力学理论.

（一）非孤立系统基本热力学量涨落的概率分布公式

设我们所研究的系统 1 浸在一个大外源 2 中，并一起构成孤立系. 为简单起见，暂时假设系统 1 和外源 2 只能交换能量和改变体积，而不能交换粒子. 因为复合系统（1 + 2）是孤立系统，所以应有 $\Delta E_1 = -\Delta E_2$，$\Delta V_1 = -\Delta V_2$.

从 §7.2.3 知道，当系统 1 的能量和体积取最概然值（即平均值 $\langle E_1\rangle$ 和 $\langle V_1\rangle$）时，孤立系的微观态数和熵有最大值，且根据熵的玻尔兹曼关系有

$$S_{max} = k_B \ln \Omega_{max}. \tag{1.1}$$

当系统 1 的能量和体积对其最概然值偏离 ΔE_1 和 ΔV_1 时，复合系统的微观态数和熵有较小的值，但仍有

$$S_f = k_B \ln \Omega_f. \tag{1.2}$$

由（1.1）式和（1.2）式得

$$\Omega_f/\Omega_{max} = e^{(S_f - S_{max})/k_B} = e^{\Delta S/k_B}. \tag{1.3}$$

复合系统既然是一孤立系，它的每一个微观态是等概率的. 所以系统 1 的能量和体积有偏差 ΔE_1 和 ΔV_1 的概率 $p \propto \Omega_f/\Omega_{max}$，或

$$P \propto \Omega_f \propto e^{\Delta S/k_B}. \tag{1.4}$$

其中 $\Delta S = S_f - S_{max}$ 是复合系统熵的偏差. 由熵的广延性可知

$$\Delta S = \Delta S_1 + \Delta S_2. \tag{1.5}$$

其中 ΔS_1 和 ΔS_2 分别是系统 1 和外源 2 的熵的偏差. 因设外源很大,即使对系统 1 来说偏差很大,对外源来说偏差仍很小,因而对外源可应用热力学基本方程 $\Delta S_2 = (\Delta E_2 + p\Delta V_2)/T$. 将它用 ΔS_1 和 ΔV_1 表出为

$$\Delta S_2 = -(\Delta E_1 + p\Delta V_1)/T. \tag{1.6}$$

其中 T 和 p 是外源的温度和压强,也是系统的平均温度和平均压强. 将(1.5)式代入(1.4)式,并利用(1.6)式,得 $P \propto e^{-(\Delta E_1 - T\Delta S_1 + p\Delta V_1)/(k_B T)}$. 因为概率已不再依赖于外源的具体性质,所以可以去掉脚标,而写为

$$P \propto e^{-(\Delta E - T\Delta S + p\Delta V)/(k_B T)}. \tag{1.7}$$

(1.7)式适用于任何偏差,无论是比较小还是比较大. 在大多数情况下,偏差是很小的,因此可以将 ΔE 在平衡值 $(\Delta E)_0 = 0$ 附近作泰勒展开. 当保留到二级项时结果为

$$\begin{aligned}\Delta E &= (\partial E/\partial S)_0 \Delta S + (\partial E/\partial V)_0 \Delta V + (1/2)[(\partial^2 E/\partial S^2)_0 (\Delta S)^2 \\ &\quad + 2(\partial^2 E/\partial S\partial V)_0 \Delta S\Delta V + (\partial^2 E/\partial V^2)_0 (\Delta V)^2] + \cdots \\ &= \left(\frac{\partial E}{\partial S}\right)_0 \Delta S + \left(\frac{\partial E}{\partial V}\right)_0 \Delta V + \frac{1}{2}\left[\frac{\partial}{\partial S}\left(\frac{\partial E}{\partial S}\right)_0 \Delta S\Delta S + \frac{\partial}{\partial V}\left(\frac{\partial E}{\partial S}\right)_0 \Delta V\Delta S\right] + \\ &\quad + \frac{1}{2}\left[\frac{\partial}{\partial S}\left(\frac{\partial E}{\partial V}\right)_0 \Delta S\Delta V + \frac{\partial}{\partial V}\left(\frac{\partial E}{\partial V}\right)_0 \Delta V\Delta V\right] + \cdots.\end{aligned} \tag{1.8}$$

利用

$$(\partial E/\partial S)_0 = T, \qquad (\partial E/\partial V)_0 = -p. \tag{1.9}$$

和

$$\begin{aligned}[(\partial T/\partial S)\Delta S + (\partial T/\partial V)\Delta V]\Delta S &= \Delta T\Delta S, \\ [(\partial p/\partial S)\Delta S + (\partial p/\partial V)\Delta V]\Delta V &= \Delta p\Delta V.\end{aligned} \tag{1.10}$$

(1.8)式化为

$$\Delta E = T\Delta S - p\Delta V + [\Delta T\Delta S - \Delta p\Delta V]/2 + \cdots.$$

再代入(1.7)式,最后得

$$P \propto e^{-(\Delta T\Delta S - \Delta p\Delta V)/(2k_B T)}. \tag{1.11}$$

若系统与外源不仅有能量交换和体积交换,还有粒子交换(为简单起见,假设只有一种粒子),(1.11)式可推广为

$$P \propto e^{-(\Delta T\Delta S - \Delta p\Delta V + \Delta\mu\Delta N)/(2k_B T)}. \tag{1.12}$$

(1.12)式就是我们要求的非孤立系基本热力学量涨落的概率分布公式,和(1.7)式不同,(1.11)和(1.12)式只适用于小涨落. 下面将看到,在选定独立变量 x 后,(1.12)式可化为标准高斯分布的形式,这是由中心极限定理决定的:

$$P(\Delta x) \propto e^{-(\Delta x)^2/[2\langle(\Delta x)^2\rangle]}. \tag{1.13}$$

这样,各个热力学量的均方涨落和涨落的统计关联就可以从(1.12)式和(1.13)式的比较中直接写出.

要注意的是,(1.12)式中包括六个基本热力学量的偏差,即三个强度量的偏差 $\Delta T,\Delta p,\Delta\mu$ 和它们的三个共轭广延量的偏差 $\Delta S,\Delta V,\Delta N$,但其中只有三个可以被选作是独立的.当系统受到宏观条件约束时,独立偏差量数还会减少.例如,若系统与外源有能量交换和体积交换,但无粒子交换,则独立偏差量数为2,若系统只与外源有能量交换,而无体积和粒子交换,则独立偏差量数减为1.

应用(1.12)式的基本步骤是:第一,根据宏观约束,确定独立偏差量数.第二,选择独立偏差量,并将其他偏差量用所选独立偏差量表示.第三,将所得(1.12)式和标准高斯分布(1.13)式比较.

为了便于与系综理论已有的结果比较,下面计算恒温系统和开放系统某些热力学量的涨落.

(二)热力学量涨落的计算

【例1】 设系统有固定体积和粒子数,但与热源可交换能量,即恒温系.求 $\langle(\Delta T)^2_{V,N}\rangle,\langle(\Delta p)^2_{V,N}\rangle,\langle(\Delta E)^2_{V,N}\rangle$ 和 $\langle(\Delta T\Delta S)_{V,N}\rangle$ 及 $\langle(\Delta T\Delta p)_{V,N}\rangle$,脚标表示系统的约束条件.

解 由于 $\Delta V=\Delta N=0$,独立偏差量数为1,涨落的普遍公式(1.12)式简化为

$$P\propto e^{-(\Delta T\Delta S)_{V,N}/(2k_BT)}. \tag{1.14}$$

设选择 ΔT 为独立偏差量,则 $\Delta S=(\partial S/\partial T)_{V,N}\Delta T$,

$$(\Delta T\Delta S)_{V,N}=(\partial S/\partial T)_{V,N}(\Delta T)^2_{V,N}. \tag{1.15}$$

将(1.15)式代入(1.14)式,得

$$P\propto e^{-(1/2k_BT)(\partial S/\partial T)_{V,N}(\Delta T)^2_{V,N}}. \tag{1.16}$$

将(1.16)式与高斯分布(1.13)式比较,立刻可以写出

$$\langle(\Delta T)^2_{V,N}\rangle=k_BT(\partial T/\partial S)_{V,N}=k_BT^2/C_{V,N}, \tag{1.17}$$

由(1.15)式和(1.17)式得

$$\langle(\Delta T\Delta S)_{V,N}\rangle=(\partial S/\partial T)_{V,N}\langle(\Delta T)^2\rangle=k_BT. \tag{1.18}$$

(1.18)式表明强度量 ΔT 和它的共轭广延量 ΔS 统计相关.将 Δp 和 ΔE 也用所选独立偏差量 ΔT 表示,

$$\Delta p=(\partial p/\partial T)_{V,N}\Delta T,\qquad \Delta E=(\partial E/\partial T)_{V,N}\Delta T.$$

利用(1.17)式可得:

$$\langle(\Delta p)^2_{V,N}\rangle=\left(\frac{\partial p}{\partial T}\right)^2_{V,N}\langle(\Delta T)^2_{V,N}\rangle=\frac{k_BT^2}{C_{V,N}}\left(\frac{\partial p}{\partial T}\right)^2_{V,N}, \tag{1.19}$$

$$\langle(\Delta E)^2_{V,N}\rangle = (\partial E/\partial T)^2_{V,N}\langle(\Delta T)^2_{V,N}\rangle = k_B T^2 C_{V,N}, \tag{1.20}$$

$$\langle(\Delta T\Delta p)_{V,N}\rangle = (\partial p/\partial T)_{V,N}\langle(\Delta T)^2_{V,N}\rangle = k_B T(\partial p/\partial S)_{V,N}. \tag{1.21}$$

(1.20)式正是利用正则系综所得的§7.3.1(24)式,(1.21)式则表明两个强度量统计相关.

仍考虑以上系统,但选 Δp 为独立变量,将 ΔT 和 ΔS 用 Δp 表示,则有

$$\Delta T = (\partial T/\partial p)_{V,N}(\Delta p)_{V,N},$$

$$\Delta S = (\partial S/\partial p)_{V,N}(\Delta p)_{V,N} = (\partial S/\partial T)_{V,N}(\partial T/\partial p)_{V,N}(\Delta p)_{V,N}.$$

(1.14)式化为

$$P \propto e^{-[1/(2k_B T)][(\partial S/\partial T)_{V,N}/(\partial p/\partial T)^2_{V,N}](\Delta p)^2_{V,N}}.$$

与高斯分布(1.13)式比较,再次得到(1.19)式.将 ΔE 用 Δp 表示,得到

$$\langle(\Delta E)^2_{V,N}\rangle = (\partial E/\partial p)^2_{V,N}\langle(\Delta p)^2_{V,N}\rangle$$

$$= (\partial E/\partial T)^2_{V,N}(\partial T/\partial p)^2_{V,N}\langle(\Delta p)^2_{V,N}\rangle = k_B T^2 C_{V,N}.$$

再次得到(1.20)式.这说明热力学量的涨落由系统的宏观条件决定,与独立变量的选择无关.

【例2】 设系统有固定体积,但与外源可交换能量和粒子,即开放系.求:$\langle(\Delta T)^2_V\rangle$,$\langle(\Delta p)^2_V\rangle$,$\langle(\Delta E)^2_V\rangle$,$\langle(\Delta N)^2_V\rangle$ 及某些涨落的关联函数.脚标 V 表示开放系的约束条件仅仅是 V 不变.

解 由于 $\Delta V = 0$,独立偏差量数为2,(1.12)式化为

$$P \propto e^{-(\Delta T\Delta S + \Delta\mu\Delta N)/(2k_B T)}. \tag{1.22}$$

若选择强度量 ΔT 与它非共轭的广延量 ΔN 作为独立变量来表示 ΔS 和 $\Delta\mu$,则有

$$(\Delta T\Delta S + \Delta\mu\Delta N)_V = (\partial S/\partial T)_{N,V}(\Delta T)^2_V + (\partial S/\partial N)_{T,V}(\Delta N\Delta T)_V$$

$$+(\partial\mu/\partial T)_{N,V}(\Delta T\Delta N)_V + (\partial\mu/\partial N)_{T,V}(\Delta N)^2_V. \tag{1.23}$$

由热力学基本方程 $dF = -SdT - pdV + \mu dN$,有

$$(\partial S/\partial N)_{T,V} = -(\partial\mu/\partial T)_{N,V}. \tag{1.24}$$

用(1.24)式可以消去(1.23)式中的交叉项,使(1.22)式化为

$$P \propto e^{-[1/(2k_B T)][(\partial S/\partial T)_{N,V}(\Delta T)^2_V + (\partial\mu/\partial N)_{T,V}(\Delta N)^2_V]}. \tag{1.25}$$

根据独立事件概率相乘原理,(1.25)式说明,温度的涨落与粒子数的涨落是统计独立的高斯分布.于是可以写出:

$$\langle(\Delta T\Delta N)_V\rangle = \langle(\Delta T)_V\rangle\langle(\Delta N)_V\rangle = 0, \tag{1.26}$$

$$\langle(\Delta T)^2_V\rangle = k_B T(\partial T/\partial S)_{N,V} = k_B T^2/C_{V,N}, \tag{1.27}$$

$$\langle(\Delta N)^2_V\rangle = k_B T(\partial N/\partial\mu)_{T,V}. \tag{1.28}$$

(1.28)式和利用巨正则分布得到的§7.3.3中(3.22)式一致.

利用(1.26)~(1.28)式很容易得出以下关联函数:

$$\langle(\Delta T\Delta S)_V\rangle = (\partial S/\partial T)_{N,V}\langle(\Delta T)_V^2\rangle + (\partial S/\partial N)_{T,V}\langle(\Delta T\Delta N)_V\rangle = k_B T, \tag{1.29}$$

$$\langle(\Delta T\Delta\mu)_V\rangle = (\partial\mu/\partial T)_{N,V}\langle(\Delta T)_V^2\rangle + (\partial\mu/\partial N)_{T,V}\langle(\Delta T\Delta N)_V\rangle = k_B T(\partial\mu/\partial S)_{N,V} = k_B T(\partial T/\partial N)_{S,V}, \tag{1.30}$$

$$\langle(\Delta\mu\Delta S)_V\rangle = 0, \tag{1.31}$$

$$\langle(\Delta\mu\Delta N)_V\rangle = k_B T, \tag{1.32}$$

$$\langle(\Delta S\Delta N)_V\rangle = k_B T(\partial S/\partial\mu)_{T,V} = k_B T(\partial N/\partial T)_{\mu,V}. \tag{1.33}$$

由(1.26)式和(1.29)~(1.33)式可以看到,一个强度量(广延量)和其非共轭广延量(或非共轭强度量)统计独立,而其余的量均统计相关,且不难归纳出其中的规律.

将 Δp 和 ΔE 均用独立变量 ΔT 和 ΔN 表示,利用(1.26)~(1.28)式可求得开放系统的能量涨落和压强涨落:

$$\begin{aligned}\langle(\Delta E)_V^2\rangle &= (\partial E/\partial T)_{N,V}^2\langle(\Delta T)^2\rangle + 2(\partial E/\partial T)_{N,V}(\partial E/\partial N)_{T,V}\langle(\Delta T\Delta N)_V\rangle \\ &\quad + (\partial E/\partial N)_{T,V}^2\langle(\Delta N)_V^2\rangle \\ &= k_B T^2 C_{V,N} + k_B T(\partial N/\partial\mu)_{T,V}(\partial E/\partial N)_{T,V}^2,\end{aligned} \tag{1.34}$$

$$\begin{aligned}\langle(\Delta p)_V^2\rangle &= (\partial p/\partial T)_{N,V}^2\langle(\Delta T)_V^2\rangle + 2(\partial p/\partial T)_{N,V}(\partial p/\partial N)_{T,V}\langle(\Delta T\Delta N)_V\rangle \\ &\quad + (\partial p/\partial N)_{T,V}^2\langle(\Delta N)_V^2\rangle \\ &= k_B T[(\partial p/\partial T)_{N,V}(\partial p/\partial S)_{N,V} + (\partial p/\partial N)_{T,V}(\partial p/\partial\mu)_{T,V}].\end{aligned}$$

利用麦氏关系:$(\partial p/\partial S)_{V,N} = -(\partial T/\partial V)_{S,N}$ 及 $(\partial p/\partial\mu)_{T,V} = (\partial N/\partial V)_{T,\mu}$ 得到:

$$\langle(\Delta p)_V^2\rangle = k_B T[-(\partial p/\partial T)_{N,V}(\partial T/\partial V)_{S,N} + (\partial p/\partial N)_{T,V}(\partial N/\partial V)_{T,\mu}].$$

对于单元系,化学势 $\mu = \mu(T,P)$,因此有 $(\partial N/\partial V)_{T,\mu} = (\partial N/\partial V)_{T,P}$,最后得

$$\begin{aligned}\langle(\Delta p)_V^2\rangle &= k_B T[-(\partial p/\partial T)_{N,V}(\partial T/\partial V)_{N,S} + (\partial p/\partial N)_{T,V}(\partial N/\partial V)_{T,P}] \\ &= k_B T[-(\partial p/\partial T)_{N,V}(\partial T/\partial V)_{N,S} - (\partial p/\partial V)_{T,N}] \\ &= -k_B T(\partial p/\partial V)_{S,N}.\end{aligned} \tag{1.35}$$

(1.34)式正好是巨正则系综的结果,见第七篇第三章的习题9.

通过上面的两个例子可以归纳出以下几点:第一,强度量和其非共轭广延量统计独立,其余均统计相关.第二,将(1.19)和(1.35)式,(1.20)和(1.34)式比较可以看到,同样是压强和能量的均方涨落,但对恒温系统和开放系统有不同结果.即:从原则上说,热力学量的涨落与系统的宏观条件有关.当然也存在着某些热力学量

的涨落与系统宏观条件无关的特殊情况. 例如(1.17)和(1.27)式表明,恒温系统与开放系统有相同的温度涨落. 这是由于 ΔT 和 ΔN 统计独立. 第三,已经指出,涨落由系统的宏观条件决定,与独立变量的选择无关,但推导过程却有简繁之别. 因此,按照正确的原则选取独立变量十分重要. 如例2中选 ΔT 和 ΔN 为独立变量对求 $\langle(\Delta T)^2\rangle$,$\langle(\Delta N)^2\rangle$、$\langle\Delta T\Delta N\rangle$ 很方便,而选 $\Delta\mu$ 和 ΔS 为独立变量对求 $\langle(\Delta\mu)^2\rangle$ 和 $\langle\Delta\mu\Delta S\rangle$ 则更简捷. 第四,因为系统的热容量 $C_{V,N}\propto N$,所以从(1.17)、(1.19)和(1.35)式看到,强度量的均方涨落(如 $\langle(\Delta T)^2\rangle$ 和 $\langle(\Delta p)^2\rangle$,与 N 成反比;从(1.20)、(1.28)和(1.34)式看到,广延量的均方涨落如 $\langle(\Delta E)^2\rangle$,与 N 成正比. 但强度量与广延量的相对涨落均与 $\sqrt{N}$ 成反比. 如:

$$\delta T=\sqrt{\langle(\Delta T)^2\rangle}/T\propto 1/\sqrt{N},\qquad \delta p=\sqrt{\langle(\Delta p)^2\rangle}/p\propto 1/\sqrt{N},$$

$$\delta E=\sqrt{\langle(\Delta E)^2\rangle}/E\propto 1/\sqrt{N},\qquad \delta N=\sqrt{\langle(\Delta N)^2\rangle}/N\propto 1/\sqrt{N}.$$

所以,对于一个宏观单相系统来说,热力学量的相对涨落一般都观察不到,只有临界点附近的密度涨落才大到可以通过光的分子散射而被观察到.

(三)密度涨落和光的分子散射

设系统有确定的温度和粒子数,则由(1.12)和(1.13)式可得

$$P(\Delta V)\propto e^{\Delta p\Delta V/(2k_BT)}=e^{(\partial p/\partial V)_{T,N}[(\Delta V)^2/(2k_BT)]},$$

$$\langle(\Delta V)^2\rangle_{T,N}=-k_BT(\partial V/\partial p)_{T,N},$$

$$(\delta V)_{T,N}=\sqrt{\frac{\langle(\Delta V)^2\rangle_{T,N}}{V^2}}=\left[-\frac{k_BT}{V^2}\left(\frac{\partial V}{\partial p}\right)_T\right]^{1/2}=\left(\frac{k_BT}{V}\kappa_T\right)^{1/2},\qquad(1.36)$$

其中 κ_T 是等温压缩系数. 这说明,体积(或密度)的相对涨落与体积的平方根成反比,与温度和等温压缩系数的平方根成正比. 气体的压缩系数比液体和固体的大,特别是在接近临界点时,压缩系数很大因而密度涨落很大.

一束光射入气体或液体中时,一部分光受到散射而偏离原来方向,使光在原来方向的强度减弱,这就是光的散射现象. 最初发现光的散射现象的是丁多尔,他所观察的现象是光受到悬浮在气体或液体中的杂质的影响而发生的散射. 但是在没有杂质的气体中也有散射现象,这是由于密度涨落而引起的,称为分子散射. 我们知道,物质对光波的折射率与物质的密度有关. 如果由于涨落,气体中某处的密度与平均密度有较大的偏差,则该处的折射率也与平均折射率有较大的偏差. 光通过该处经折射后,其传播方向与平均折射方向不一致,于是发生散射现象.

人们很早就知道,在气液二相的临界点会出现强烈的光的散射——临界乳光. 这是由于在临界点附近,物质的压缩系数很大,很容易造成显著的密度涨落,因而对各种频率的光都产生强烈的散射. 不过在临界点,$(\partial V/\partial p)_T\to\infty$,因而(1.36)式已不适用. 而且在临界点附近那样大的涨落下,物体各部分的涨落将发生相互影

响而产生相关效应，必须加以考虑.

§9.1.2　布朗运动

我们知道，在显微镜下观察悬浮在液体中的微小颗粒（如花粉）时，可以看到这些颗粒不停地进行着无规则的运动. 这一现象是植物学家布朗在 1827 年首先发现的，所以称为布朗运动. 布朗运动实际上是液体分子无规则运动的反映. 我们现在学习布朗运动理论，除因为它在历史上对分子运动论的确立起过重要作用外，还有以下几方面意义. 首先，测量仪器中的活动部分（如分析天平的称盘，悬线电流计的线圈等）在气体分子无规则碰撞下也会产生布朗运动. 随着科学技术的发展，仪器的灵敏度越来越高，布朗运动对灵敏度的影响已成为现代精密测量中一个不可忽视的因素. 其次，扩散现象的本质是布朗运动产生的位移，因此布朗运动理论可用于各类扩散现象. 如半导体中载流子（电子或空穴）的扩散，原子核反应堆中中子的扩散等. 最后，布朗运动是一种典型的随机过程，在随机理论和非平衡统计的发展中占有重要地位. 除物理学、化学、生物学外，布朗运动理论在激光、生命的自组织过程乃至社会经济领域中都有广泛的应用.

布朗粒子经历随机过程. 随机过程理论有两种基本方法：朗之万方程（LE）方法和福克尔-普朗克方程（FPE）方法. 前者从随机过程的动力学方程出发，后者从随机过程概率分布随时间的演化方程出发. 我们将用这两种方法研究布朗运动.

（一）布朗运动的朗之万理论

为简单起见，只考虑布朗粒子的运动在一个水平方向（x 轴）的投影，并且不考虑外力场的作用.

设布朗微粒的质量为 m，在时刻 t，微粒的坐标为 $x(t)$. 周围液体分子施子微粒的净作用力为 $F(t)$. 朗之万把 $F(t)$ 分成两部分. 一部分为液体对布朗微粒的粘滞阻力 $-\alpha v$. 粘滞阻力来自液体分子对运动微粒的碰撞. 当微粒以速度 v 运动时，在其前进方向将与更多的分子碰撞，因此平均而言将受到与其速度方向相反的粘滞阻力. 当 v 不大时，阻力的大小与微粒的速度成正比. 如果应用流体力学近似，将微粒看作半径为 r 的小球，在粘滞系数为 η 的流体中运动，则有

$$\alpha = 6\pi r\eta. \tag{2.1}$$

$F(t)$ 的另一部分是无规作用力 $F'(t)$. 这相当于液体分子对静止的布朗微粒的碰撞作用力. 在长时间间隔中（与分子碰撞的特征时间比较）有

$$\langle F'(t)\rangle = 0, \qquad \langle F'(t_1)F'(t_2)\rangle \propto \delta(t_1 - t_2). \tag{2.2}$$

满足(2.2)式的随机力叫白噪声. 在作出这样的区分后，可以将微粒的动力学方程表示为

$$m\frac{dv}{dt} = -\alpha v + F'(t). \tag{2.3}$$

因为有随机力 $F'(t)$，$v(t)$ 和 $x(t)$ 都是依赖于时间的随机变量.(2.3)式是一个随机微分方程(通常称为朗之万方程 LE).解随机微分方程的常用方法是先对方程求系综平均,得到关于平均值的非随机微分方程,再求解.

为了求布朗运动的位移,我们先将(2.3)式改写为

$$m\ddot{x} = -\alpha\dot{x} + F'(t). \tag{2.4}$$

以 x 乘(2.4)式,并考虑到

$$x\ddot{x} = \frac{d}{dt}(x\dot{x}) - \dot{x}^2 = \frac{1}{2}\frac{d^2}{dt^2}x^2 - \dot{x}^2,$$

可得

$$\frac{1}{2}\frac{d^2}{dt^2}(mx^2) - m\dot{x}^2 = -\frac{\alpha}{2}\frac{d}{dt}x^2 + xF'(t). \tag{2.5}$$

将(2.5)式对大量微粒求平均,即把大群微粒的运动方程相加然后用微粒数去除.注意求平均和对时间求导的次序可以交换,即有

$$\langle d^2(mx^2)/dt^2\rangle = d^2\langle mx^2\rangle/dt^2.$$

随机力 $F'(t)$ 与微粒的位置无关,因此 $xF'(t)$ 的平均值等于 x 的平均值与 $F'(t)$ 的平均值的乘积,即

$$\langle xF'(t)\rangle = \langle x\rangle\langle F'(t)\rangle = 0. \tag{2.6}$$

把布朗微粒看作是服从能量均分定理的大分子,则又有

$$\langle m\dot{x}^2/2\rangle = k_B T/2. \tag{2.7}$$

将(2.6)式和(2.7)式代入(2.5)式,便可得到:

$$d^2\langle x^2\rangle/dt^2 + (\alpha/m)d\langle x^2\rangle/dt - 2k_B T/m = 0. \tag{2.8}$$

(2.8)式是 $\langle x^2\rangle$ 的二阶常系数非齐次线性微分方程,其通解为:

$$\langle x^2\rangle = 2k_B Tt/\alpha + C_1 e^{-\alpha t/m} + C_2. \tag{2.9}$$

我们首先来估计一下 α/m 的数量级:

$$\alpha/m = 6\pi r\eta/[(4/3)\pi r^3\rho] = (9/2)\eta/(r^2\rho).$$

对于水,η 约为 10^{-2}Po(Poise $= g\cdot cm^{-1}\cdot s^{-1}$),微粒(如藤黄粉)的半径和密度约为 $r = 4\times10^{-5}$ cm,$\rho = 1.2\ g\cdot cm^{-3}$,代入可得 $\alpha/m = 10^7\ s^{-1}$,可见(2.9)式中的第二项 $\approx C_1 e^{-10^7 t}$,只要 $t > 10^{-6}$ s 就可忽略不计.其次,若假设当 $t = 0$ 时,所有的微粒均处在 $x = 0$ 处,即 x 描述微粒的位移,便得 $C_2 = 0$,因此(2.9)式简化为

$$\langle x^2\rangle = (2k_B T/\alpha)t. \tag{2.10}$$

(2.10)式指出,在时间间隔 t 内,微粒位移平方的平均值与时间间隔 t 成正比.将(2.1)式代入(2.10)式,得到

$$\langle x^2\rangle = [k_B T/(3\pi r\eta)]t. \tag{2.11}$$

配兰用实验验证了(2.11)式. 因为同时观察大量微粒是不可能的,配兰的方法是在显微镜下跟踪一个布朗微粒. 记下这个微粒在时间间隔 t (例如 30s)内在 x 方向的位移. 例如在时间间隔 0 至 t,t 至 $2t$,$2t$ 至 $3t$,……微粒在 x 方向的位移分别是 $x_1, x_2, x_3, \cdots\cdots$. 由多次观测数据可求得位移平方的平均值 $\langle x^2\rangle$. 结果证实了 $\langle x^2\rangle$ 与时间间隔 t 成正比,与粘滞系数 η 成反比,与温度 T 有关.

利用观测数据还可以计算玻尔兹曼常数. 例如在配兰的一次实验中,用的是 50 粒半径十分相近的藤黄粉粒,其 $r = 2.12 \times 10^{-5}$ cm,悬浮在20℃的纯水中,位移平方是沿一定方向测量的,观察的时间是 30s,测得的结果是 $\langle x^2\rangle = 4.7 \times 10^{-7}$ cm^2. 把这些数据代入(2.11)式得到

$$k_B = \frac{3\pi \times 2.12 \times 10^{-5} \times 0.012 \times 4.7 \times 10^{-7}}{293 \times 30} = 1.26 \times 10^{-16} erg \cdot K^{-1}.$$

这个数值与现今的准确数值 $k_B = 1.38 \times 10^{-23}\ J \cdot K^{-1}$ 相差不到 10%. 鉴于这种实验的困难性,上面所得的准确度是理论与实验的一种很好的符合.

我们再从扩散的观点来研究布朗运动,仍然讨论一维问题. 用 $n(x,t)$ 代表布朗微粒的数密度,以 $J(x,t)$ 代表布朗微粒的通量(即单位时间通过单位面积的微粒数),菲克定律给出:

$$J = -D\nabla n. \tag{2.12}$$

D 是扩散系数. 布朗微粒在液体中的运动满足 $\nabla \cdot J = -\partial n/\partial t$ (连续性方程). 和(2.12)式联立,得到扩散方程:

$$\partial n/\partial t = D\nabla^2 n. \tag{2.13}$$

设 $t = 0$ 时, N 个微粒均在 $x = 0$ 处,即

$$n(x,0) = N\delta(x). \tag{2.14}$$

在此初始条件下,扩散方程(2.13)的解为

$$n(x,t) = (N/\sqrt{4\pi Dt})e^{-x^2/(4Dt)}. \tag{2.15}$$

(2.15)式表明,在任意时间 t ,微粒密度的分布是与 t 有关的高斯分布. 随着 t 的增加,微粒逐渐向四周扩散. 利用(2.15)式可以求得微粒位移平方的平均值[将(2.15)式和归一化高斯分布附录(Ⅲ.32)式比较,立即可得 $\langle x^2\rangle = \sigma^2 = 2Dt$]

$$\langle x^2\rangle = \frac{1}{N}\int_{-\infty}^{+\infty} x^2 n(x,t)\,dx = 2Dt. \tag{2.16}$$

将(2.10)式与(2.16)式比较,得到扩散系数 D 为

$$D = k_B T/\alpha \tag{2.17}$$

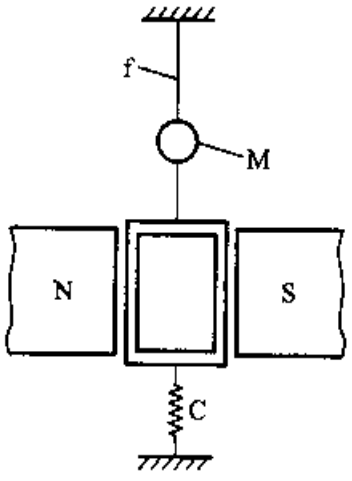


图 1　悬线式电流计示意图

这称为爱因斯坦关系，它说明粘滞及扩散的主要原因在于流体分子不停的碰撞.

下面以悬线式电流计为例，说明布朗运动对仪器灵敏度的影响.

悬线式电流计的主要部分是一个悬挂在很细的石英丝 f 上的可旋转的线圈 C. 石英丝上固定一个反射镜 M，如图 1.

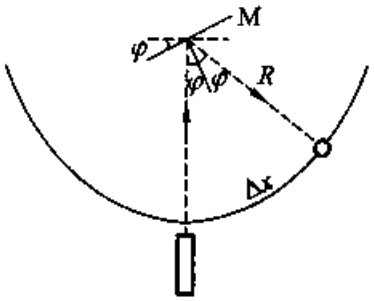


图 2　观测线圈偏转的弧形标尺

当没有电流通过时，线圈应该不偏转而停止在“零点”. 但是，由于空气分子的随机碰撞，在没有电流通过时，线圈也会发生无规则的扭摆式“布朗运动”. 其结果，即使待测电流为零，弧形标尺上的“零点”也会表现微小面突然的跳动.

我们来计算由于布朗运动引起的“零点”的平均偏转大小.

由于石英丝的弹性，线圈有一平衡位置. 当线圈相对于平衡位置偏转一个角度 φ 时，势能增加 $\alpha\varphi^2/2$，其中 $\alpha = \pi^2 r^2 G/(2l)$ 是石英丝的扭转常数，r 是石英丝的半径，l 是石英丝的长度，G 是石英丝的切变模量.

线圈在气体分子无规则碰撞下发生布朗运动. 可以认为，线圈是一个特殊的“大分子”，参予了气体分子的热运动. 其平均能量可由能量均分定理算出为 $\alpha\langle\varphi^2\rangle/2 = k_B T/2$，或 $\sqrt{\langle\varphi^2\rangle} = \sqrt{k_B T/\alpha}$.

下面估计 $\sqrt{\langle\varphi^2\rangle}$ 的大小，石英丝的 $\alpha \approx 10^{-13}\ \mathrm{J\cdot rad^{-1}}$，设 $T = 300\ \mathrm{K}$，代入上式可得 $\sqrt{\langle\varphi^2\rangle} \approx 2\times10^{-4}\ \mathrm{rad}$. 这是一个很小的角度，但在悬线式电流计中却是可以观察到的. 如图 2 所示，设反射镜 M 到弧形标尺 L 的距离 $R = 1\ \mathrm{m}$，当反射镜偏转角度 $\varphi = 2\times10^{-4}\ \mathrm{rad}$ 时，反射光偏转了 2φ 角. 反射光在标尺上的偏转距离是 $\Delta x = 2\varphi R = 0.4\ \mathrm{mm}$. 这是不难观察到的.

显然，电流计能测出的最小电流所引起的偏转必须大子 $\sqrt{\langle\varphi^2\rangle}$，因而 $\sqrt{\langle\varphi^2\rangle}$ 决定着电流计的灵敏度.

*（二）布朗粒子的概率分布　福克尔-普朗克方程

在气体分子的无规则碰撞下，布朗粒子的坐标经历一随机过程，即：它的坐标不断地随机跃变. 这一随机过程的特点是，每一次跃变到不同位置的概率只和当前的位置有关，而与如何跃变到当前位置的历史无关. 这种随机过程称为马尔可夫过程. 用 $P(x,t)$ 表示 t 时刻，布朗粒子位于 x 的概率，则 $P(x,t)$ 随时间的演化方程称为主方程（Master equation）：

$$\begin{aligned}\frac{\partial P(x,t)}{\partial t} &= \text{增益项} - \text{损失项}\\ &= \sum_{x'}\left[P(x',t)w(x'\to x) - P(x,t)w(x\to x')\right]\end{aligned} \tag{2.18}$$

其中 $w(x' \to x)$ 代表单位时间中由 x' 到 x 的概率，叫跃迁概率密度. 考虑到 x 是连续变量，上式写为

$$\frac{\partial P(x,t)}{\partial t} = \int_{-\infty}^{\infty} [P(x',t)w(x' \to x) - P(x,t)w(x \to x')]\mathrm{d}x' \tag{2.19}$$

主方程(2.19)式是一个积分微分方程，求解比较困难. 在只有小跃迁的假设下，可以把它化为一个关于 $P(x,t)$ 的二阶偏微分方程，即福克尔－普朗克方程，其步骤如下.

令 $r = x - x'$，或 $x' = x - r$，可把跃迁概率密度改写为

$$w(x' \to x) = w(x - r, r), \qquad w(x \to x') = w(x, -r)$$

主方程(2.19)式化为

$$\frac{\partial P(x,t)}{\partial t} = \int [w(x - r, r)P(x - r, t) - w(x, -r)P(x,t)]\mathrm{d}r$$

假设：①只有小跃迁发生，即跃迁概率密度是围绕 $r = 0$ 尖锐峰起的函数；② $P(x,t)$ 是随 x 变化而缓慢变化的函数，即有 $P(x + \Delta x, t) \approx P(x,t)$. 于是有

$$w(x - r, r) \approx w(x,r) + \frac{\partial w(x,r)}{\partial x}(-r) + \frac{1}{2}\frac{\partial^2 w(x,r)}{\partial x^2}(-r)^2. \tag{2.20}$$

主方程化为

$$\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x}D_1(x)P(x,t) + \frac{1}{2}\frac{\partial^2}{\partial x^2}D_2(x)P(x,t), \tag{2.21}$$

其中

$$D_1(x) = \int_{-\infty}^{\infty} rw(x,r)\mathrm{d}r = \langle r \rangle \qquad \text{单位时间中跃迁的平均距离，}$$

$$D_2(x) = \int_{-\infty}^{\infty} r^2 w(x,r)\mathrm{d}r = \langle r^2 \rangle \qquad \text{单位时间中跃迁的距离平方的平均值.}$$

(2.21)式就是福克尔－普朗克方程(简写为 FPE). 方程右方第一项叫漂移项，第二项叫扩散项.

现在将福克尔－普朗克方程应用于布朗运动. 当布朗微粒不受外力场时有 $D_1(x) = \langle r \rangle = 0$；又由(2.16)式可知，$\langle r^2 \rangle = \langle x^2 \rangle / t = 2D$，于是布朗粒子的 FPE 具体为

$$\partial P(x,t)/\partial t = D\partial^2 P(x,t)/\partial x^2. \tag{2.22}$$

它和扩散方程(2.13)式具有相同的形式. 其解也有相同形式：

$$P(x,t) = (1/\sqrt{4\pi Dt})\mathrm{e}^{-x^2/(4Dt)}. \tag{2.23}$$

和(2.15)式一致.

习　题

1. 对于与外源有能量和体积交换，但无物质交换的系统. 证明以下各式，并从中寻找规律.

$$\langle (\Delta T)_N^2 \rangle = k_B T(\partial T/\partial S)_{V,N}, \qquad \langle (\Delta p)_N^2 \rangle = -k_B T(\partial p/\partial V)_{S,N},$$

$$\langle (\Delta S)_N^2 \rangle = k_B T(\partial S/\partial T)_{P,N}, \qquad \langle (\Delta V)_N^2 \rangle = -k_B T(\partial V/\partial p)_{T,N},$$

$$\langle (\Delta T \Delta V)_N \rangle = 0, \qquad \langle (\Delta S \Delta p)_N \rangle = 0,$$

$$\langle (\Delta T \Delta p)_N \rangle = -k_B T(\partial T/\partial V)_{S,N} = k_B T(\partial p/\partial S)_{V,N},$$

$$\langle (\Delta S \Delta V)_N \rangle = -k_B T (\partial S/\partial p)_{T,N} = k_B T (\partial V/\partial T)_{p,N}.$$

2. 有半径为 0.4×10^{-4} cm 的粒子，在粘滞系数为 2.18×10^{-2} Po 的 18.8℃液体中作布朗运动. 测得粒子在时间间隔 10 s 中的位移平方平均值为 $\langle x^2 \rangle = 3.3 \times 10^{-8}\ \mathrm{cm}^2$. 根据这些数据求玻尔兹曼常数.

第二章　临界现象

通过§7.2.9的学习知道，相变可以分为一级相变和二级相变（或连续相变）．连续相变的相变点称为临界点．在临界点，一些物理量（如等温压缩系数、等温磁化率、比热）呈现发散尖峰的反常行为，称为临界现象．为了能定量地描述这种奇异性，人们引入临界指数．临界指数的测量表明，在各式各样临界点的邻域，系统的许多热力学性质表现出极大的相似性．在学习临界现象的统计理论之前，在本章第一节里，先以流体系统和铁磁系统为例，介绍临界指数和临界指数之间的关系——标度律．第二节介绍一种普遍适用的二级相变的热力学理论，即朗道－金兹堡理论．第三节讨论临界点邻域序参量的涨落与关联．序参量关联长度趋于无穷是临界点的基本特征，是重刺化群方法的基础．

§9.2.1　临界指数和标度律

（一）流体系统的临界指数

在§7.2.9中已经指出，气液系统除了一级相变外也存在二级相变．在那一节的图1 a 中，以比容和压强为坐标画出了系统的等温线．改以密度和压强为坐标画出的等温线则将如本节图1所示．ρ_C 表示系统在临界点的密度，两侧的虚线分别表示两相平衡下气相和液相的密度 ρ_g 和 ρ_l．以 $t=(T-T_C)/T_C$ 表示离开临界点的相对距离，则实验发现，在临界点邻域有以下几个规律：

（1）指数 β

在 $t\to-0$ 和 $t\to+0$ 时，$\rho_l-\rho_g$ 随 $|t|$ 的变化分别服从：

$$\begin{aligned}&\rho_l-\rho_g\propto|t|^{\beta},\qquad (T<T_C),\\&\rho_l-\rho_g=0,\qquad (T>T_C).\end{aligned}\tag{1.1}$$

在临界温度以上（$T>T_C$），物质处于液气不分的状态．临界指数 β 的实验值约为0.34．

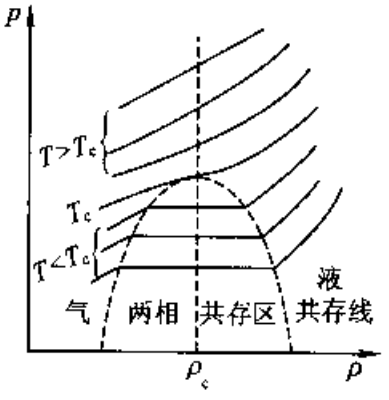


图1　流体系统的等温线

（2）指数 γ,γ'

等温压缩系数 $\kappa_T=-(1/v)(\partial v/\partial p)_T=(1/\rho)(\partial\rho/\partial p)_T$ 随 t 的变化规律为

$$\begin{aligned}&\kappa_T\propto|t|^{-\gamma'},\quad (T<T_c);\\&\kappa_T\propto t^{-\gamma},\qquad (T>T_C).\end{aligned}\tag{1.2}$$

上式中的量在 $T > T_C$ 时，沿临界等体积线 $\rho = \rho_C$ 取值；在 $T < T_C$ 时，沿两相共存线 $\rho = \rho_l(t)$ 和 $\rho = \rho_g(t)$ 取值. γ 和 γ' 的实验值约为 1.2. 两式的比例系数不同. 在临界点 κ_T 发散，这意味着小的压强涨落会导致显著的密度涨落.

(3) 指数 δ

沿着临界等温线 $t = 0$，压强与临界压强之差 $(p - p_C)$ 和密度与临界密度之差 $(\rho - \rho_C)$ 有以下关系

$$p - p_C \propto |\rho - \rho_C|^{\delta}, \qquad (T = T_C). \tag{1.3}$$

δ 值为 $6 \geqslant \delta \geqslant 4$.

(4) 指数 α, α'

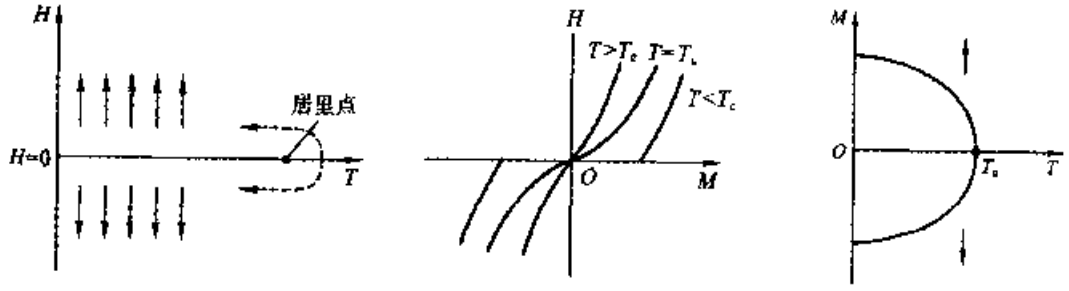
C_V 随 t 的变化规律为

$$\begin{aligned} &C_V \propto |t|^{-\alpha'}, \quad (T < T_C); \\ &C_V \propto t^{-\alpha}, \qquad (T > T_C), \end{aligned} \qquad (\rho = \rho_C). \tag{1.4}$$

两式中的量均在 $\rho = \rho_C$ 下取值. α 和 α' 的实验值约为 0.1. 两式的比例系数不同. 在临界点 C_V 呈对数发散.

(二) 铁磁系统的临界指数

铁磁系统的临界指数 β, γ, δ 和 α 的定义和液体系统类似. 为了便于对照和理解，图 2 画出了单轴铁磁系统相图的示意图. 图 2(a) 画出了单轴铁磁系统的共存曲线和临界点 ($H = 0, T = T_C$). 在临界温度 T_C 以上，外场为零时，磁化强度为零. 在临界点以下，自发磁化强度出现. 直线 $H = 0$ 将两个可能的自发磁化方向分隔开. 图 2(a) 和液体系统的 p-T 图对应. 在图 2b 中我们画出了单轴铁磁系统的一些等温线. 显然图 2(b) 和液体系统的图 1 对应. 图 2(c) 画出作为温度函数的自发磁化强度曲线. 定义各种临界指数时，参考这些曲线将会有帮助.



(a)　　(b)　　(c)

图 2　单轴铁磁系统的　(a)共存曲线　(b)等温线　(c)磁化曲线

(1) 指数 β_m

在 $t \to -0$ 时，自发磁化强度 $m_{磁}$ (简写为 m) 随 $|t|$ 的变化服从

$$m \propto |t|^{\beta_m}, \qquad (T < T_C); \tag{1.1'}$$

在临界温度以上，无自发磁化，

$$m = 0, \qquad (T > T_C).$$

β_m 的实验值约为 1/3.

(2)指数 γ_m, γ'_m

等温磁化率 $\chi_T \equiv (\partial m/\partial H)_T$ 随 t 的变化规律为

$$\begin{aligned} \chi_T &\propto |t|^{-\gamma'_m}, \quad (T < T_C, H \to 0), \\ \chi_T &\propto t^{-\gamma_m}, \qquad (T > T_C, H \to 0). \end{aligned} \tag{1.2'}$$

两式的比例系数不同. γ_m 和 γ'_m 的实验值约为 1.3. 在临界点 χ_T 发散.

(3)指数 δ_m

沿临界等温线($t = 0$),外加磁场和磁化强度的关系为

$$H \propto m^{\delta_m}, \qquad (T = T_C). \tag{1.3'}$$

指数 δ_m 的实验值约为 $6 \geqslant \delta_m \geqslant 4$.

(4)指数 α_m, α'_m

铁磁质的零场比热 $c_{H=0}$ 随 t 的变化规律为:

$$\begin{aligned} c_{H=0} &\propto |t|^{-\alpha'_m}, \quad (T < T_C, H = 0); \\ c_{H=0} &\propto t^{-\alpha_m}, \qquad (T > T_C, H = 0). \end{aligned} \tag{1.4'}$$

指数 α_m 和 α'_m 的实验值约为 0.1. 在临界点 $C_{H=0}$ 呈对数发散.

可以看出:

表 I　一些流体的临界指数

流体	T_C(K)	p_C (MPa)	β	γ	γ'	δ
CO_2	304.20	7.382	0.34 ± 0.15	1.35 ± 0.15	1.1 ± 0.04	5 ± 1
Ar	150.87	4.907	0.35 ± 0.2			
H_2	32.98	1.293	0.375 ± 0.15			
^4He	5.19	0.227	0.325			
^3He	3.324	0.117	0.36 ± 0.02			

表 II　一些铁磁系统的临界指数

铁磁质	T_C(K)	β_m	γ_m　γ'_m	α_m　α'_m	δ_m
Fe	1044.0	0.34 ± 0.02	γ = 1.33 ± 0.15	0.12 ± 0.01	
Ni	631.58	0.33 ± 0.03	γ = 1.32 ± 0.02	0.10 ± 0.03	4.0 ± 0.1
EuO	69.33			0.09 ± 0.01	
$YFeO_3$	643	0.354 ± 0.005	γ = 1.33 ± 0.04 γ' = 0.07 ± 0.01		
Gd	292.5		γ = 1.33		4.0 ± 0.1

①只要将流气系统和铁磁系统作以下对应：

$$\rho_l - \rho_g \leftrightarrow m, \qquad p - p_C \leftrightarrow H,$$

$$\kappa_T \equiv \rho^{-1}(\partial\rho/\partial p)_T \leftrightarrow \chi_T \equiv (\partial m/\partial H)_T, \quad C_{V=V_C} \leftrightarrow C_{H=0}.$$

则上述两个系统在临界点邻域的行为有极大的相似性. 变化规律相同，临界指数值大致相等.

②一般说来，热力学函数 $f(t)$ 的临界指数可定义为

$$\lambda = \lim_{t\to 0}(\ln f(t)/\ln|t|).$$

若 $\lambda > 0$，则对应的 $f(t)$ 在临界点为零，例如 $\rho_l - \rho_g$ 和 m. 这样的 $f(t)$ 常称为序参量.

若 $\lambda < 0$，则对应的 $f(t)$ 在临界点发散，如 κ_T 和 χ_T.

若 $\lambda = 0$，则对应的 $f(t)$ 在临界点有有限的跳跃.

若 $\lambda \neq 0$，但数值很小，根据 $a^x = e^{x\ln a}$，$\lim_{x\to 0} a^x \approx 1 + x\ln a$，则对应的 $f(t)$ 在临界点呈对数发散，例如 c_V.

③在上面介绍的四个临界指数中，除 α 和比热有关外，其他都和序参量有关. 其后，我们还将引入新的临界指数，它们也都和序参量有关.

④各种二级相变在临界点的相似性，不仅表现在临界指数大致相同，更令人吃惊的是指数间存在一些共同关系：

$$\alpha + 2\beta + \gamma \geqslant 2, \qquad \alpha + \beta(\delta + 1) \geqslant 2.$$

最初，这两个关系是作为近似实验规律提出的，以后有人用热力学理论导出了这两个关系. 临界指数之间的关系称为标度律. 确定临界点，推导标度律和计算临界指数是相变理论的主要任务.

§9.2.2　二级相变的朗道-金兹堡理论

临界指数和标度律表明，在临界点不同系统的共性突出，个性退居次要地位，因此发展临界现象的唯象理论是可能的和必要的. 朗道-金兹堡的二级相变理论就是一种普遍适用的唯象理论. 这一理论提出了序参量的概念，总结了前人在解释各种临界现象时的共同精神——平均场思想，得到了在临界点附近，自由焓的非线性表达式. 由此不仅可以证明临界点的存在，证明在临界点化学势和化学势的一阶偏导数连续，二阶偏导数不连续，而且可以计算各种临界指数. 但是，所得到的临界指数和实验值相差较远. 特别是按朗道-金兹堡理论，在临界点比热呈有限跳跃，而不是对数发散. 比热呈有限跳跃，临界指数 $\alpha = 0$ 是一切平均场理论（包括统计理论）的共同特点.

(一)序参量

在讨论铁磁系统和液气系统时,已经看到,自发磁化强度 $m_{自}$ 和液气密度差 $\rho_l - \rho_g$ 处于对等地位,在临界点邻域随温度变化有相同的规律.当系统的温度高于临界温度($T > T_C$)时,自发磁化强度和液气密度差均为零,系统处于对称性较高或有序度较低的相;当系统的温度低于临界温度($T < T_C$)时,自发磁化强度和液气密度差不为零,系统处于对称性较低或有序度较高的相.而当系统的温度从 $T < T_C$ 逐渐升高到 T_C 时,自发磁化强度和液气密度差由不等于零的有限值连续地变到零.自发磁化强度 m 和液气密度差 $\rho_l - \rho_g$ 称为铁磁系统和流体系统的序参量.在已引入的四个临界指数 $\beta,\gamma(\gamma'),\delta$ 和 $\alpha(\alpha')$ 中,除 α 与比热有关外,其余三个均与序参量有关.在下一节中还将引入另外两个临界指数,它们和序参量的涨落及涨落的关联长度有关.

一般说来,对于二级相变,可以引入序参量来定量描写.序参量应该这样选择:在对称性较高的高温相,序参量为零;而在对称性较低的低温相,序参量有不等于零的值.当从有序相趋于临界点时,序参量的值连续地变到零,如图1(b)所示.因此,在临界点附近,序参量是小量.对一级相变,也可以引入描述对称性变化的物理量,但在相变点,该物理量会有不连续的跃变[如图1(a)].二级相变总是伴随着对称性的变化.对称性的变化总是突然的.对称性的突然降低,称为对称性破缺.序参量无穷小变化,引起对称性破缺是二级相变(连续相变)的本质.

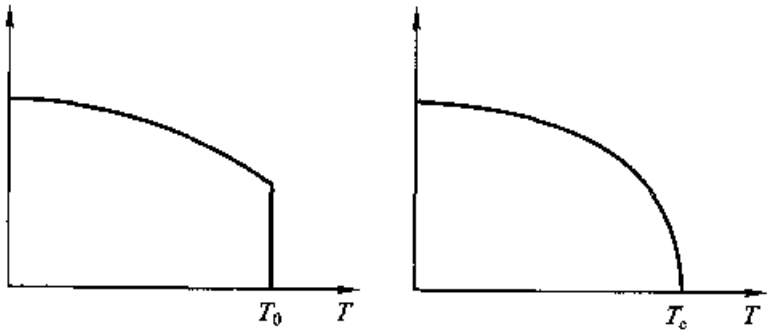


图1　一级相变的参量(图a)和二级相变的序参量(图b)随温度变化的一般特征

§7.2.9 表Ⅱ中列举了几种二级相变的序参量.在不同的相变中,序参量不仅是不同的物理量,而且结构也可以不同.例如在液气、铁磁、双液和二元合金系统中,序参量均是标量,序参量的维数为1.而在超流和超导系统中序参量则是波函数 $\psi = \psi_0 e^{i\varphi}$,$\psi$ 是复数,因此序参量可用复平面的矢量表示,序参量的维数是2.进一步的研究表明,按临界指数的值可以把二级相变分为不同的普适类.同一普适类的临界指数值基本相同,而普适类则由系统的空间维数和序参量的维数决定.

(二)朗道-金兹堡理论

我们以单轴铁磁体为例,介绍朗道-金兹堡的平均场理论.对序参量维数为1的其他系统,以下讨论完全适用.

已知在临界点附近,序参量 m 是小量,因此当外场为零时可以将自由焓 $G(T,m)$ 在 T_C 附近按序参量 m 的幂次展开(因为压强保持不变,故未写出):

$$G(T,m) = G(T,0) + (a(T)/2)m^2 + (b(T)/4)m^4 + \cdots. \tag{2.1}$$

式中 $G(T,0)$ 是 $m = 0$ 时的自由焓.由于自由焓是标量,磁化强度 m 是矢量,所以展开式中不含 m 的奇次幂.系数 $a(T)$ 和 $b(T)$ 和温度 T 有关.

根据在 $T > T_C$ 时顺磁相 ($m = 0$) 是稳定态,而在 $T < T_C$ 时铁磁相 ($m \neq 0$) 是稳定态,可以导出 $a(T)$ 和 $b(T)$ 的函数形式:

$$a(T) = a_0(T - T_C)/T_C = a_0 t \qquad a_0 > 0, \tag{2.2}$$

$$b(T) = b_0(T_C) + b't \approx b_0(T_C) > 0. \tag{2.3}$$

下面对此进行说明.

在稳定平衡态,系统的自由焓有极小值,即应有:

$$\partial G/\partial m = 0, \tag{2.4}$$

$$\partial^2 G/\partial m^2 > 0. \tag{2.5}$$

将(2.1)式代入极值条件(2.4)和极小值条件(2.5),分别得到:

$$\partial G/\partial m = m(a + b_0 m^2) = 0, \tag{2.6}$$

$$\partial^2 G/\partial m^2 = a + 3b_0 m^2 > 0. \tag{2.7}$$

(2.6)式有三个根:$m = 0$ 和 $m = \pm\sqrt{-a/b_0}$,分别代表顺磁相和铁磁相.将 $m = 0$ 代入(2.7)式得到 $a(T) > 0$,可见,为说明顺磁相是高温稳定态应有 $a(T > T_C) > 0$.将 $m^2 = -a/b_0$ 代入(2.7)式得 $-2a(T) > 0$,可见,为说明铁磁相(m 是不为零的实数)是低温稳定态,应有 $a(T < T_C) < 0$ 和 $b_0(T_C) > 0$.当 $T = T_C$ 时,$a(T = T_C) = 0$ 是临界情况.显然,为使系数 $a(T)$ 同时满足 $a(T > T_C) > 0$,$a(T < T_C) < 0$ 和 $a(T = T_C) = 0$,(2.2)式是一种自然而简单的假设.

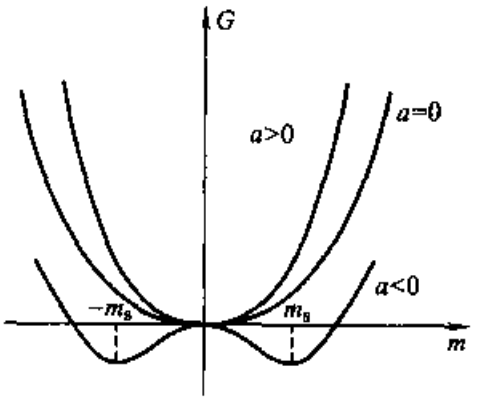


图 2　G-m 的函数曲线

将(2.2),(2.3)式代入(2.1)式得

$$G(T,m) = G(T,0) + a_0 t m^2/2 + b_0 m^4/4 + \cdots, \quad (a_0, b_0 > 0). \tag{2.8}$$

在 $T > T_C(t > 0)$,$T = T_C(t = 0)$ 和 $T < T_C(t < 0)$ 时,由(2.8)式可以绘出 $G - m$ 的函数曲线,如图 2.

由曲线可以更清楚地看出,根据朗道-金兹堡理论在 T_C 邻域关于自由焓 $G(T,m)$ 的表达式可得,当 $T > T_C$ (t>0 或 a>0)时,自由焓的极小值在 $m = 0$ 处;当 T

$< T_C(t < 0$ 或 $a < 0)$ 时，原来单一的极小值分解为两个极小，自由焓的极小值在 $m = m_S = \pm \sqrt{-a/b_0}$ 处，即朗道-金兹堡理论导出了临界点的存在.

再由朗道-金兹堡理论推导临界指数. 将(2.2)式代入 $m = \pm \sqrt{-a/b_0}$，得

$$m = \pm (a_0/b_0)^{1/2}(-t)^{1/2}, \qquad t < 0. \tag{2.9}$$

可见，按朗道-金兹堡理论，序参量 m 对 t 的依赖和上节中的(1.1′)式有相同的形式，但临界指数 $\beta_m = 1/2$.

当存在外场时，磁系统的自由焓为

$$G(T,m,H) = G_0 + am^2/2 + b_0m^4/4 - Hm. \tag{2.10}$$

假设外场很弱，忽略 a, b_0 对 H 的依赖关系，平衡时的 m 值由自由焓的极值条件决定：

$$\partial G/\partial m = am + b_0m^3 - H = 0, \tag{2.11}$$

再将上式对 H 求导，可得磁化率

$$\chi = (\partial m/\partial H)_T = \frac{1}{a + 3b_0m^2} = \begin{cases} t^{-1}/a_0, & t > 0 \\ (-t)^{-1}/(2a_0), & t < 0. \end{cases} \tag{2.12}$$

在最后一步中已将式中的 m^2 近似地用自发磁化值：$m = 0(t > 0)$ 和 $m^2 = -a/b(t < 0)$ 代入，和上节的(1.2′)式比较，有 $\gamma_m(\gamma'_m) = 1$.

在 $T = T_C$ 时，有 $a(T) = 0$，由(2.11)式得

$$H \propto m^3. \tag{2.13}$$

和上节中的(1.3′)比较有 $\delta_m = 3$.

由(2.8)式可得

$$S = -(\partial G/\partial T)_{H=0} = -(m^2/2)(\mathrm{d}a/\mathrm{d}T) - \mathrm{d}G_0/\mathrm{d}T.$$

当 $T > T_C, H = 0$ 时，$m = 0$，因而比热

$$C_{H=0} = T(\partial S/\partial T)_{H=0} = -T(\mathrm{d}^2G_0/\mathrm{d}T^2). \tag{2.14}$$

当 $T < T_C, H = 0$ 时，$m^2 = -a_0t/b_0$，因而

$$C_{H=0} = T(\partial S/\partial T)_{H=0} = -T(\mathrm{d}^2G_0/\mathrm{d}T^2) + Ta_0^2/(2b_0T_C^2). \tag{2.15}$$

(2.15)式减(2.14)式，并于两式中取 $T = T_C$，得到

$$\Delta C_{H=0} = a_0^2/(2b_0T_C).$$

即在临界点，铁磁相的比热大于顺磁相的比热，但 $C_{H=0}$ 有一个有限的跃变，由此可知临界指数 $\alpha = 0$.

朗道-金兹堡理论定性地解释了临界现象，但所得临界指数和实验结果相差较大. 原因是，朗道-金兹堡理论忽略了涨落，采取了平均场近似的思想. 分析如下：我们把铁磁质中的自旋磁矩看成处于外磁场 H 和其他自旋产生的分子磁场 H' 中. H' 应该是一个与铁磁质中的自旋磁矩分布有关，且不断涨落的复杂量. 平均场理

论简单地把分子场 H' 看成是只与磁化强度有关的量，并把 H' 按 m 的幂次展开：

$$H' = a_1 m + a_2 m^3. \tag{2.16}$$

其中 a_1 和 a_2 为常数. 由于 $m = 0$ 时，$H' = 0$，且分子场的方向和磁化强度的方向相同，故展开式只含 m 的奇次幂，于是铁磁质中的总磁场为

$$\mathscr{H} = H + H' = H + a_1 m + a_2 m^3 + \cdots. \tag{2.17}$$

在顺磁态时，磁化强度和 $\mathscr{H}$ 的关系服从居里定理，即 $m = c\mathscr{H}/T$. 将(2.17)式代入得

$$Tm = c[H + a_1 m + a_2 m^3 + \cdots].$$

即

$$H = (T - a_1 c)m/c - a_2 m^3 + \cdots. \tag{2.18}$$

令 $a_1 c = T_C$，$a_2 = -b_0$，则(2.18)式化为

$$H = a_1 tm + b_0 m^3 + \cdots.$$

和(2.11)式完全一致. 这说明朗道-金兹堡理论是一种平均场理论. 平均场思想是处理多体问题的一种常用近似方法.

* §9.2.3 临界点邻域的涨落和关联

已经知道，在临界点，流体系统的等温压缩系数 $\kappa_T = (1/\rho)(\partial\rho/\partial p)_T$ 和比热 $C_V = T(\partial S/\partial T)_V$ 变成无限，意味着密度和熵的涨落很大；磁系统的等温磁化率 $\chi_T \equiv (\partial m/\partial H)_T$ 和比热 $C_H = T(\partial S/\partial T)_H$ 变成无限，意味着磁化强度和熵的涨落很大. 流体系统密度的大涨落可以通过临界乳光直接观察；磁系统磁化强度的大涨落也可以通过中子散射而直接观察. 此外，研究还发现，在临界点，某些量的涨落变成大范围相关. 本节以磁系统为具体例子，研究临界点邻域的涨落和关联，并引进新的临界指数.

如上所述，因临界点邻域 m 的涨落很大，因而 §9.1.1(1.12)式不再适用，而应该用该节(1.7)式进行计算. 此外，在 §9.1.1 中，把系统视为均匀系统，严格说来这也是不恰当的. 本节，我们将采用涨落的准热力学理论中对任意大小的涨落均适用的 §9.1.1(1.7)式，并把系统视为非均匀系，来计算临界点邻域磁化强度 m 的涨落和关联.

首先介绍关联函数的概念，以 $m(\boldsymbol{r})$ 表示 $\boldsymbol{r}$ 点的局域磁化强度，并定义局域磁化强度的空间关联函数为

$$C(\boldsymbol{r}_1, \boldsymbol{r}_2) \equiv \langle \Delta m(\boldsymbol{r}_1) \Delta m(\boldsymbol{r}_2) \rangle = C(\boldsymbol{r}) \qquad \boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2. \tag{3.1}$$

其中 $\Delta m(\boldsymbol{r}_1) \equiv m(\boldsymbol{r}_1) - \langle m(\boldsymbol{r}_1) \rangle$. 当 $\boldsymbol{r}_1 = \boldsymbol{r}_2$ 时，

$$C(\boldsymbol{r}_1, \boldsymbol{r}_1) = C(0) = \langle \Delta m^2(\boldsymbol{r}_1) \rangle \tag{3.2}$$

代表 $\boldsymbol{r}$ 点磁化强度的均方涨落.

如果 $C(\boldsymbol{r}_1, \boldsymbol{r}_2) = \langle \Delta m(\boldsymbol{r}_1) \rangle \langle \Delta m(\boldsymbol{r}_2) \rangle = 0$，则称 $\boldsymbol{r}_1, \boldsymbol{r}_2$ 处磁化强度的涨落不相关；反之，如果 $C(\boldsymbol{r}_1, \boldsymbol{r}_2) \neq 0$，则意味着不同地点的涨落相关.

将 $C(r)$ 展为傅里叶级数

$$C(r) = (1/V)\sum_k C(k)e^{+ik\cdot r}. \tag{3.3}$$

$C(r)$ 的傅里叶分量由逆变换得到

$$C(k) = \int C(r)e^{-ik\cdot r}dr. \tag{3.4}$$

将 $\Delta m(r)$ 展为傅里叶级数

$$\Delta m(r) = (1/V)\sum_k \Delta m_k e^{+ik\cdot r}, \tag{3.5}$$

$$\Delta m_k = \int \Delta m(r)e^{-ik\cdot r}dr. \tag{3.6}$$

对(3.5)式取复共轭,并考虑到 $\Delta m(r)$ 是实量,得

$$\Delta m(r) = (1/V)\sum_k \Delta m_k^* e^{-ik\cdot r} = (1/V)\sum_k \Delta m_{-k}^* e^{+ik\cdot r}.$$

和(3.5)式比较,对 $\Delta m(r)$ 的付氏分量有

$$\Delta m_k = \Delta m_{-k}^*, \qquad 或 \qquad \Delta m_k^* = \Delta m_{-k}. \tag{3.7}$$

因此

$$|\Delta m_k|^2 = \iint dr dr' \Delta m(r)\Delta m(r')e^{-ik\cdot(r-r')}.$$

对上式取统计平均,并注意关联函数的定义(3.1)式,可得

$$\langle|\Delta m_k|^2\rangle = \iint dr' dr C(r-r')e^{-ik\cdot(r-r')} = \int dr' \int dR C(R)e^{-ik\cdot R} = VC(k). \tag{3.8}$$

将上式所得 $C(k)$ 代入(3.3)式,最后得

$$C(r) = (1/V^2)\sum_k \langle|\Delta m_k|^2\rangle e^{ik\cdot r}. \tag{3.9}$$

根据涨落的准热力学理论可以方便地计算 $\langle|\Delta m_k|^2\rangle$,再由(3.9)式便可求得临界点附近磁化强度的关联函数.

对封闭系统,在等温等压情况下,偏离平衡态任意大小的概率 §9.1.1(1.7)式可化为

$$P \propto e^{-\Delta G/(k_B T)}. \tag{3.10}$$

对于非均匀磁系统,可引入自由焓密度 $g(r)$:

$$\Delta G = \int [g(r) - \langle g(r)\rangle]dr = \int \Delta g(r)dr. \tag{3.11}$$

由于非均匀性,自由焓密度不仅和 $\Delta m(r)$ 而且和 $m(r)$ 的梯度 $\partial m(r)/\partial r$ 有关.再考虑磁系统的平移不变性,$g(r)$ 应只和 $\Delta m(r)$ 及 $\partial m(r)/\partial r$ 的偶次幂有关.于是,在无外场情况下,有

$$g(r) = g_0 + a_0 t m^2(r) + b_0 m^4(r) + c_0[\partial m(r)/\partial r]^2 + \cdots. \tag{3.12}$$

由于在临界点平衡态时 $\langle m(r)\rangle = 0$,磁化强度瞬时值对平均值的偏差为

$$\Delta m(r) = m(r) - \langle m(r)\rangle = m(r).$$

忽略(3.12)式中的 $m^4(r)$ 项,则有

$$\Delta g(r) = a_0 t|\Delta m(r)|^2 + c_0|\partial \Delta m(r)/\partial r|^2 + \cdots. \tag{3.13}$$

这里 t 是无量纲的约化温度,a_0 的意义见 §9.2.2.

由(3.5)式得:

$$[\Delta m(r)]^2 = (1/V^2)\sum_{kk'}\Delta m_k \Delta m_{k'}^* e^{i(k-k')\cdot r}$$

$$[\partial \Delta m(r)/\partial r]^2 = (1/V^2)\sum_{kk'}\Delta m_k \Delta m_{k'}^* k\cdot k' e^{i(k-k')\cdot r}.$$

将以上两式代入(3.13)式，

$$\Delta g(r) = \frac{1}{V^2}\sum_{kk'}\Delta m_k \Delta m_{k'}^* (a_0 t + c_0 k\cdot k') e^{-i(k'-k)\cdot r}. \tag{3.14}$$

再将上式代入(3.11)式，并注意到 $(1/V)\int e^{-i(k'-k)\cdot r}dr = \delta_{kk'}$，得到

$$\Delta G = (1/V)\sum_k (a_0 t + c_0 k^2)\mid \Delta m_k \mid^2. \tag{3.15}$$

将(3.15)式代入(3.10)式得

$$P \propto \exp\{[-1/(Vk_B T)]\sum_k (a_0 t + c_0 k^2)\mid \Delta m_k \mid^2\}.$$

我们看到，在波矢空间中，偏差概率的表达式有高斯分布的形式. 这正是我们要对关联函数进行傅氏变换的原因. 将上式和高斯分布比较立刻可写出

$$\langle \mid \Delta m_k \mid^2 \rangle = Vk_B T/[2(a_0 t + c_0 k^2)]. \tag{3.16}$$

将(3.16)式代入(3.9)式，即有

$$C(r) = \frac{k_B T}{2V}\sum_k \frac{1}{a_0 t + c_0 k^2} e^{ik\cdot r}.$$

对于宏观尺度的体积，可将上式的求和化为积分：

$$\sum_k \cdots = \left(\frac{L}{2\pi}\right)^3 \int \cdots dk, \qquad (L^3 = V).$$

于是上式化为

$$C(r) = \frac{k_B T}{2\cdot(2\pi)^3}\int \frac{1}{a_0 t + c_0 k^2} e^{ik\cdot r} dk.$$

利用留数定理计算积分可得

$$C(r) = [k_B T/(8\pi c_0 r)]e^{-r/\xi(t)}. \tag{3.17}$$

其中

$$\xi(t) = [c_0/(a_0 t)]^{1/2}. \tag{3.18}$$

(3.17)式描述相距 r 的两点磁化强度涨落的关联. 当 $r < \xi(t)$ 时，两点涨落的关联显著；当 $r > \xi(t)$ 时，关联函数迅速衰减为零，所以 $\xi(t)$ 是关联的特征长度，叫做关联长度. 而由(3.18)式可知，当 $T \to T_C$ 或 $t \to 0$ 时 $\xi(t) \to \infty$，序参量关联长度趋于无穷是临界点的基本特征，是重整化群方法的基础.

(3.18)式的结果，是从朗道-金兹堡平均场理论得到的. 在普遍情况下，可以把关联长度与 t 的指数关系表为：

$$\xi(t) \propto \begin{cases} t^{-\nu} & T > T_C, \\ \mid t \mid^{-\nu'} & T < T_C. \end{cases} \tag{3.19}$$

对平均场理论来说，$\nu = \nu' = 1/2$. ν 的实验值为 0.6～0.7.

从(3.17)、(3.18)式还可看出，当 $T = T_C$ 或 $t = 0$ 时 $\xi(t) \to \infty$，$e^{-r/\xi(t)} \to 1$，

$$C(r,t=0) \propto 1/r, \qquad (T=T_c).$$

即序参量关联函数 $C(r,t=0)$ 在临界点与 r 成反比，这也是平均场理论的结果．进一步的理论把 $C(r,t=0)$ 与 r 的关系用另一个量 η 表出，即

$$C(r,t=0) \propto 1/r^{d-2+\eta}, \qquad (T=T_C). \tag{3.20}$$

d 是系统的空间维数，对于平均场理论，取 $d=3$，则 $\eta=0$．η 的实验值为 0.1．ν 和 η 是两个新的临界指数，它们分别和序参量的关联长度和序参量的关联函数有关．

习　题

1.(a)证明范氏方程可写成不含参量 a,b 的对应态方程

$$(\bar{p}+3/\bar{V}^2)(3\bar{V}-1)=8\bar{T}, \tag{1}$$

其中

$$\bar{p}=p/p_C, \qquad \bar{V}=V/V_C, \qquad \bar{T}=T/T_C. \tag{2}$$

而临界参量为

$$p_C=a/(27b^2), \qquad V_C=3b, \qquad T_C=8a/(27bR). \tag{3}$$

(b)引入

$$\pi=\bar{p}-1, \qquad t=\bar{T}-1, \qquad \omega=\bar{V}-1, \tag{4}$$

证明：在临界点邻域，对应态方程可近似化为

$$\pi=4t-6t\omega-3\omega^3/2. \tag{5}$$

(c)由(5)式证明，对范氏气体有临界指数

$$\delta=3, \qquad \gamma=1.$$

(d)以 π 为纵坐标轴，ω 为横坐标轴，根据(5)式，用 math 绘出在参数 $t_1=0, t_2=-0.001, t_3=-0.0015$ 下的三条等温线，即在临界点邻域，范氏等温线在 π-ω 平面上的表示．

说明：在 t_2,t_3 两条等温线上，压强为 $\pi=4t$ 的三点的意义及麦克斯韦等面积法则在图上的表示．

(e)由(5)式结合(d)中所绘曲线证明，当 $t<0$ 时，(5)式有实数解

$$\omega_g=(-2t)^{1/2}, \qquad \omega_l=-(-2t)^{1/2}, \tag{6}$$

并利用

$$\Delta\rho \sim 1/V_l-1/V_g \propto (\omega_l-\omega_g)$$

证明临界指数 $\beta=1/2$．

2.已知气体的迭特里奇方程为

$$pe^{a/(RTV)}=RT/(V-b). \tag{7}$$

a,b 为迭特里奇气体常数．

(a)求迭特里奇气体的临界温度 T_C，临界压强 p_C 和临界体积 V_C．

提示　$RT_C/(p_CV_C)=e^2/2$，建议用 math.

(b)由状态方程(7)导出对应态定律．

第三章　临界现象的统计理论

本章介绍临界现象的统计理论.我们将集中讨论一个简单但却有广泛应用的模型——Ising 模型,所用方法有两种.一种是传统的系综理论,即通过直接计算配分函数(近似或严格),进而求临界点及临界指数;另一种是重整化群方法.由于直接计算由有相互作用粒子组成的系统的配分函数的困难,在 Widom 和 Kadanoff 标度假设的启发下,Wilson 开辟了另一途径.即不去直接计算配分函数,而是通过找某种使配分函数保持不变的标度变换,从而确定临界点和临界指数.这种方法叫重整化群(RG)方法.重整化群方法始于六十年代,现已发展成为许多领域工作者的共同语言,如粒子物理、凝聚态物理、宇宙学、生物物理及神经网络等.

§9.3.1 中介绍 Ising 模型并用系综理论处理一维 Ising 模型. §9.3.2 介绍 Widom 和 Kadanoff 的标度假设. §9.3.3 通过一维 Ising 模型介绍重整化群方法的大意.

§9.3.1　Ising 模型

早在 1873 年,范德瓦耳斯用刚刚发展起来的分子运动论推导气体状态方程时,就发现了临界点.此后在相当长的时间中,关于临界现象的研究却没有什么进展.原因是,考虑粒子间存在相互作用时,要精确计算系统的配分函数非常困难.近似计算又往往会舍掉某些重要位形的贡献.可见,关键在于找到一个在数学上便于处理,在物理上又能抓住问题基本特征的模型.1925 年,在他的导师 Lenz 的指导下,Ising 研究了铁磁-顺磁相变的一个简单模型.现在人们认识到,该模型代表了一大类合作现象,其意义已远远超出了铁磁系统.

(一) Ising 模型

Ising 模型由 N 个排列在晶格点阵上的粒子组成,每个粒子只能处于向上或者向下的两个状态.系统的微观态 (ν) 由一套自旋表示,$\nu = \{s_i\}, s_i = \pm 1, i = 1, 2, \cdots, N$.

由于相变只和粒子间的相互作用有关,因此模型不考虑粒子的动能,而且为了简单,只考虑最近邻粒子的相互作用.用 J 表示一对近邻粒子的相互作用能,μ_0 代表粒子的固有磁矩,则模型的能量为

$$E(\{s_i\}) = -\mu_0 H \sum_i s_i - J \sum_{\langle ij \rangle}{}' s_i s_j. \tag{1.1}$$

其中 $J > 0$, $\sum_{\langle ij \rangle}{}'$ 表示对最近邻粒子求和.(1.1)式中的第一项代表所有粒子在外

场中的塞曼能,第二项代表全部最近邻粒子的相互作用能的总和.

根据以上假设,模型中存在两种相反的作用:最近邻粒子间的相互作用要使粒子有序排列,而无规则热运动则要破坏这种有序排列.设当温度达到某一特定值 T_C 时,两种相反作用达到平衡,则当 $T > T_C$ 时,破坏有序排列的作用占有优势,在无外场情况下,粒子将无序排列,以致自发磁化磁矩 $\langle M\rangle_{H=0} = 0$;当 $T < T_C$ 时,促使粒子有序排列的作用占优势,在无外场情况下,粒子仍能有序排列,以致在宏观大小的范围(磁畴)内,自发磁化磁矩 $\langle M\rangle_{H=0} \neq 0$.即在 $T = T_C$ 处,模型经历了一个无序⇌有序转变.确定 T_C 和计算临界指数,是研究 Ising 模型的各种方法的共同目标.由以上分析不难看出,最近邻粒子相互作用愈强,粒子的最近邻粒子数愈多,则 T_C 愈高.

Ising 模型的配分函数为

$$Z_I = \sum_{\{s_i\}} e^{-E(\{s_i\})/(k_B T)} = \sum_{s_1=\pm 1} \cdots \sum_{s_N=\pm 1} e^{-E(\{s_i\})/(k_B T)}. \tag{1.2}$$

如果能计算出配分函数,即可按下列公式求得各热力学量

$$F_I(T,H) = -k_B T \ln Z_I(T,H), \tag{1.3}$$

$$U_I = -T^2(\partial[F_I/T]/\partial T)_H, \tag{1.4}$$

$$C_H = (\partial U_I/\partial T)_H = -T(\partial^2 F_I/\partial T^2)_H, \tag{1.5}$$

$$\langle M\rangle = -(\partial F_I/\partial H)_T. \tag{1.6}$$

若

$$\langle M(T < T_C, H = 0)\rangle \neq 0. \tag{1.7}$$

则表示存在自发磁化.

问题归结为计算 Ising 模型的配分函数.

(二)一维 Ising 模型的矩阵解法

在 1925 年 Ising 用"组合法"给出了一维 Ising 模型的严格解.接着在 1941 年 Kramers 和 Wennier 引进了一维 Ising 模型的矩阵解法,此方法的特点是用矩阵来表示配分函数,不作任何近似求得严格解.虽然一维 Ising 模型不存在铁磁-顺磁相变,但矩阵解法却很有启发性.就在 Kramers 和 Wennier 论文发表后的第三年(1944 年),Onsager 用矩阵方法求得了二维 Ising 模型的严格解.Onsager 的工作显示了精细的数学技巧.本节给出一维 Ising 模型的矩阵解法,并简单介绍二维 Ising 模型矩阵解法的结果.至今尚未得到三维 Ising 模型的严格解.

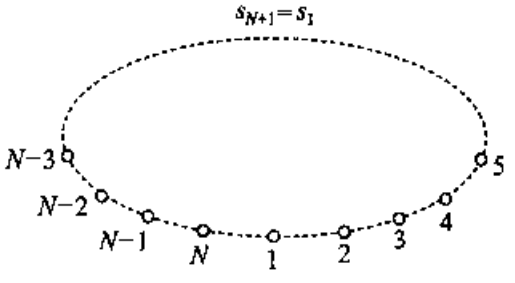


图 1 一维闭合链的 Ising 模型

Kramers 和 Wennier 采用了一维闭合链的模型,如图 1,想象将一维晶格弯成一

个环，从而有

$$s_{N+1} = s_1. \tag{1.8}$$

引入闭合链或周期边界条件(1.8)式的目的是消除边界效应，但又不会影响 $N \to \infty$ 极限时的热力学性质. 我们立刻会看到，正是由于(1.8)式的引入，使得可以把模型的哈密顿写成对称的形式，便于使用矩阵方法.

考虑(1.8)式后，(1.1)式可写为

$$E_I(\{s_i\}) = -J\sum_{\langle ij\rangle}{}' s_i s_j - \mu_0 H\sum_i s_i = -J\sum_{i=1}^{N} s_i s_{i+1} - \frac{1}{2}\mu_0 H\sum_{i=1}^{N}(s_i + s_{i+1}). \tag{1.9}$$

相应的配分函数为

$$\begin{aligned} Z_1(T,H) &= \sum_{s_i} e^{-\beta\sum_{i=1}^{N}\left[-Js_is_{i+1}-\frac{1}{2}\mu_0 H(s_i+s_{i+1})\right]} \\ &= \sum_{s_1=\pm1}\cdots\sum_{s_N=\pm1}\{e^{\beta\left[Js_1s_2+\frac{1}{2}\mu_0 H(s_1+s_2)\right]}e^{\beta\left[Js_2s_3+\frac{1}{2}\mu_0 H(s_2+s_3)\right]}\cdots e^{\beta\left[Js_Ns_1+\frac{1}{2}\mu_0 H(s_N+s_1)\right]}\}. \end{aligned} \tag{1.10}$$

引入一个 2×2 的矩阵 P，其矩阵元为

$$\langle s_i \mid P \mid s_{i+1}\rangle \equiv e^{\beta\left[Js_is_{i+1}+\frac{1}{2}\mu_0 H(s_i+s_{i+1})\right]}. \tag{1.11}$$

由于 s_i, s_{i+1} 可取 ±1 两个值，所以 p 矩阵具体为：

$$P = \begin{bmatrix} e^{\beta(J+\mu_0 H)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-\mu_0 H)} \end{bmatrix}. \tag{1.12}$$

利用 p 矩阵，模型的配分函数可写为

$$\begin{aligned} Z_1(T,H) &= \sum_{s_1=\pm1}\cdots\sum_{s_N=\pm1}\langle s_1 \mid P \mid s_2\rangle\langle s_2 \mid P \mid s_3\rangle\cdots\langle s_{N-1} \mid P \mid s_N\rangle\langle s_N \mid P \mid s_1\rangle \\ &= \sum_{s_1=\pm1}\langle s_1 \mid P^N \mid s_1\rangle = \mathrm{tr}(P^N). \end{aligned} \tag{1.13}$$

注意按(1.12)式定义的 P 是对称矩阵，就知道它可以用正交矩阵 U 对角化，即 UPU^{-1} 是对角矩阵. 将(1.13)式中 N 个连乘的 p 都换成 $PU^{-1}U$，并利用求迹下矩阵乘积可以循环的性质，将最右方的一个 U 移到最左方，上式可改写为

$$\begin{aligned} Z_1(T,H) &= \mathrm{tr}(PU^{-1}UPU^{-1}U\cdots PU^{-1}U) = \mathrm{tr}(UPU^{-1})^N \\ &= \mathrm{tr}\begin{bmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{bmatrix}^N = \lambda_+^N + \lambda_-^N. \end{aligned} \tag{1.14}$$

这里 λ_+ 和 λ_- 是矩阵 P 的本征值，由久期方程

$$\begin{vmatrix} e^{\beta(J+\mu_0 H)} - \lambda & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-\mu_0 H)} - \lambda \end{vmatrix} = 0.$$

决定

$$\lambda_\pm = e^{\beta J}\{\mathrm{ch}(\beta\mu_0 H) \pm \sqrt{\mathrm{ch}^2(\beta\mu_0 H) - 2e^{-2\beta J}\mathrm{sh}(2\beta J)}\}. \tag{1.15}$$

因 $\lambda_+ > \lambda_-$，故有

$$Z_1(T,H) = \lambda_+^N [1 + (\lambda_-/\lambda_+)^N] \underset{N\to\infty}{\to} \lambda_+^N.$$

或

$$\lim_{N\to\infty} N^{-1} \ln Z_1(T,H) \approx \ln \lambda_+. \tag{1.16}$$

即配分函数由 P 矩阵较大的本征值决定. 自由能和磁化强度为

$$F(T,H) = -k_B T \ln Z_1(T,H) = -Nk_B T \ln \lambda_+,$$

$$m = M/N = -N^{-1}\partial F/\partial H = (k_B T/\lambda_+)(\partial \lambda_+/\partial H).$$

将 λ_+ [见(1.15)式]代入，最后得磁化强度为

$$m = \mathrm{sh}(\mu_0 H/(k_B T))/\sqrt{\mathrm{e}^{-4J/(k_B T)} + \mathrm{sh}^2(\mu_0 H/(k_B T))}. \tag{1.17}$$

由(1.17)式不难看出，$m(-H) = -m(H)$；$m(H=0)=0$ 和 $H\to\pm\infty$ 时，$m\to\pm 1$，因而不同温度下，磁化强度对 H 的依赖关系如图 2 所示.

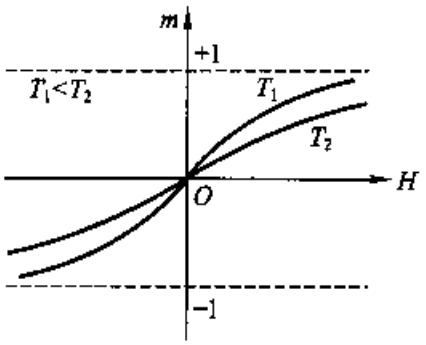


图 2　一维 Ising 模型的磁化强度，无自发磁化

可以看出，对一切 $T>0$，有 $m(T,H=0)=0$. 这表明 Ising 模型在一维情形下没有自发磁化. 换句话说，对于一维 Ising 模型，有 $T_C = 0$.

一维 Ising 模型不存在自发磁化的原因可以通过使模型从基态激发到无序态的能量来理解. 图 3 所示代表一维 Ising 模型的两个基态之一. 这时，能量为 $-NJ$，磁化强度 $m = M/N = \mu_0$. 这是一个有序态.

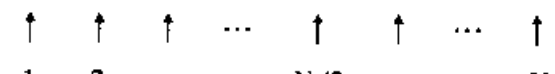


图 3　一维 Ising 模型的基态之一

图 4 所示的无序态的能量为 $(-N+4)J$，磁矩 $M=0$. 从基态激发到此无序态所需能量很小，仅为 J 的数量级. 在温度 T，平均来说，每个粒子具有 $k_B T$ 的热运动动能. 整个模型具有 $Nk_B T$ 的热运动动能，只要满足 $Nk_B T \approx J$，就可以破坏模型的自发磁化状态. 对应的温度为 $T_C \approx J/Nk_B$. 对于宏观的 N，T_C 趋于零. 换句话说，对于一维 Ising 模型，在有限温度下，不存在自发磁化.

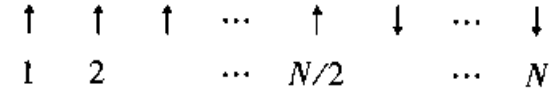


图 4　一维 Ising 模型的一个激发态

然而，对于二维 Ising 模型，从基态激发到无序态所需能量比较大. 考虑如图 5 所示的位形，作上面的类似分析，可

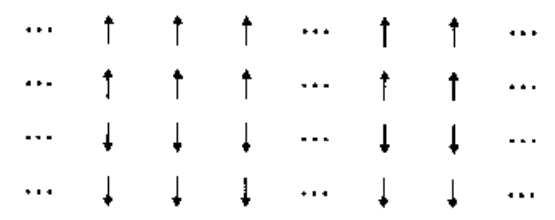


图 5　二维 Ising 模型的一个位形

得 $T_C \approx J/(N^{1/2}k_B)$. 于是,对于二维 Ising 模型,在有限温度下,存在自发磁化,可以发生顺磁-铁磁相变.

求二维 Ising 模型的严格解十分困难. 1944 年 Onsager 得到了二维正方晶格的 Ising 模型在外场为零时的严格解. 主要结果如下:二维正方晶格 Ising 模型的配分函数为

$$Z_1(T,H=0) = [2\mathrm{ch}(\beta J)\mathrm{e}^{I}]^N.$$

其中

$$I = \frac{1}{2\pi}\int_0^{\pi} \mathrm{d}\varphi \ln\left\{\frac{1}{2}[1+(1-\chi^2\sin^2\varphi)^{1/2}]\right\}, \qquad \chi = 2\mathrm{sh}(2\beta J)/\mathrm{ch}^2(2\beta J).$$

进一步可以证明 $T_C = 2.269J/k_B$. 热容量 $C_{H=0}(T)$ 在临界点有对数发散的奇异性(这和实验一致,与平均场理论不同):$\dfrac{C}{N} \approx \dfrac{8k_B}{\pi}(\beta J)^2 \ln\dfrac{1}{T-T_C}$. $T < T_C$ 时,磁化强度的行为是

$$m \approx 常数(T_C - T)^{\beta_m}, \beta_m = 1/8.$$

至今尚未得到三维 Ising 模型的严格解,但数值模拟已经建立,所得 T_C 大约是二维 Ising 模型的二倍,并有

$$C/N \propto |T-T_C|^{-\alpha_m}, \qquad m = M/N \propto |T_C - T|^{\beta_m}, \qquad T < T_C.$$

其中临界指数是 $\alpha_m \approx 0.125$, $\beta_m \approx 0.313$.

§9.3.2　标度理论

从 §9.2.1 中我们看到,物质在临界点附近的性质可以用一些临界指数表示出来. 不同系统其相变性质不同,但几乎有相等的临界指数,而且临界指数满足相同的关系——标度关系. 这说明,标度关系反映了所有物质在临界点的共同特征. 二十世纪六十年代不少学者对标度关系进行了研究. 本节介绍 Widom 和 Kadanoff 关于标度关系的理论推导. Widom 的标度理论是一个纯唯象的理论. Kadanoff 的标度变换理论则为 Widom 的唯象理论提供了一个微观的物理图像,并得到了两个新的标度关系——涉及到关联长度 r_C 的临界指数 ν 和关联函数 $C(r,t=0)$ 的临界指数 η. Kadanoff 的工作的重要性更在于,Wilson 的重整化群理论可以看成是对于 Kadanoff 标度变换思想的推广和执行.

为了阐明标度理论,首先复习广义齐次函数的概念.

(一)广义齐次函数

在 §1.2.1 研究动能的表达式时,曾经用过欧拉的齐次函数定理,见该节的(1.23)式. 现在来对齐次函数进行更仔细的讨论. 若对任意的 λ 值有

$$F(\lambda x) = \lambda^p F(x). \tag{2.1}$$

则称 $F(x)$ 是宗量 x 的 p 次齐次函数. 齐次函数的意义是,当宗量增大 λ 倍时,函数的形式保持不变,而函数值是原值的 λ^p 倍.

用 S 表示正方形的面积, L 表示正方形的边长,则有

$$S(\lambda L) = \lambda^2 S(L).$$

即:正方形的面积是边长的二次齐次函数.

令 $\lambda = x^{-1}$ 或 $\lambda x = 1$,则由(2.1)得

$$F(x) = F(1)x^p,$$

即齐次函数 F 是宗量 x 的幂函数. 齐次函数的这一性质是 Widom 理论的数学依据.

将(2.1)式稍作变形,并推广到有两个宗量的情况,则有

$$F(\lambda^p x, \lambda^q y) = \lambda F(x,y). \tag{2.2}$$

(2.2)式的 $F(x,y)$ 被称为宗量 x,y 的广义齐次函数, p,q 叫广义齐次函数参量. 广义的含义是指, p,q 可以不相等,而且可以是非整数.

考虑由 k 种不同分子组成的系统,取强度量 T,P 及各组元的摩尔数 $N_1,N_2,\cdots,N_k$ 为态参量,显然系统的内能为:

$$U(T,p,\lambda N_1,\lambda N_2,\cdots,\lambda N_k) = \lambda U(T,p,N_1,N_2,\cdots,N_k).$$

即系统的内能是各组元物质的量的广义一次齐次函数.

(二)Widom 的标度理论

Widom 的标度理论是一个普遍理论,并不限于某个特殊的系统或模型. 但为叙述具体,我们仍以铁磁系统为例来介绍.

铁磁系统的临界点是 $T = T_C$, $H = 0$. 用约化温度 $t = (T - T_C)/T_C$ 和磁场强度 H 来表示系统偏离临界点的距离. 以 t,H 为参量,系统在临界点附近时的自由能可表为 $F(t,H)$. Widom 标度理论的核心是:假定在临界点附近,自由能 F 是其宗量的广义齐次函数,即有

$$F(\lambda^p t, \lambda^q H) = \lambda F(t,H), \tag{2.3}$$

其中 λ 为任意参数. 广义齐次函数参量 p,q 是待定参数(标度理论本身不能确定). 从 Widom 的标度假设(2.3)式出发,利用纯热力学公式,就可以将临界指数 β,δ,γ 和 α 都用 p,q 表示出来,进而得到这四个临界指数之间的两个标度关系.

考虑磁介质系统的热力学基本方程

$$\mathrm{d}F = -S\mathrm{d}T - M\mathrm{d}H. \tag{2.4}$$

则有

$$M = -(\partial F/\partial H)_T. \tag{2.5}$$

将(2.3)式两边同时对 H 求偏微商,并注意(2.5)式,得

$$M(t,H) = \lambda^{q-1}M(\lambda^p t,\lambda^q H). \tag{2.6}$$

令 $\lambda = t^{-1/p}$（或 $\lambda^p t = 1$）和 $H = 0$，上式化为 $M(t,0) = t^{(1-q)/p}M(1,0)$．或

$$M(t,0) \propto t^{(1-q)/p}. \tag{2.7}$$

将(2.7)式和§9.2.1(1.1′)式比较，可将临界指数 β 用 p,q 表示出来：

$$\beta = (1-q)/p. \tag{2.8}$$

令 $\lambda = H^{-1/q}$ 和 $t = 0$，(2.6)式化为 $M(0,H) = H^{(1-q)/q}M(0,1)$．或

$$M(0,H) \propto H^{(1-q)/q}. \tag{2.9}$$

将(2.9)式和§9.2.1(1.3′)式比较，可将临界指数 δ 用 p,q 表示出来：

$$\delta = q/(1-q). \tag{2.10}$$

考虑磁化率的定义：$\chi = (\partial M/\partial H)_T$．将(2.6)式两边同时对 H 求偏微商，得

$$\chi(t,H) = \lambda^{2q-1}\chi(\lambda^p t,\lambda^q H). \tag{2.11}$$

令 $\lambda = t^{-1/p}$ 和 $H = 0$，上式化为 $\chi(t,0) = t^{(1-2q)/p}\chi(1,0)$．或

$$\chi(t,0) \propto t^{(1-2q)/p}. \tag{2.12}$$

将(2.12)式和§9.2.1(1.2′)式比较，可将临界指数 γ 用 p,q 表示出来：

$$\gamma = (2q-1)/p. \tag{2.13}$$

由比热的定义 $C_H = T(\partial S/\partial T)_H$ 和热力学基本方程(2.4)式可知有

$$C_H = -T(\partial^2 F/\partial T^2)_H. \tag{2.14}$$

将(2.3)式两边同时对 t 求二阶偏微商，并代入上式得

$$\lambda C_H(t,H) = \lambda^{2p}C_H(\lambda^p t,\lambda^q H).$$

令 $\lambda = t^{-1/p}$ 和 $H = 0$，上式化为 $C_H(t,0) = t^{(1-2p)/p}C_H(1,0)$．或

$$C_H(t,0) \propto t^{(1-2p)/p}. \tag{2.15}$$

将(2.15)式和§9.2.1(1.4′)式比较，可将临界指数 α 用 p,q 表示出来：

$$\alpha = (2p-1)/p. \tag{2.16}$$

由(2.8)，(2.10)，(2.13)和(2.16)式，立即可得

$$\alpha + 2\beta + \gamma = 2, \tag{2.17}$$

$$\gamma = \beta(\delta - 1). \tag{2.18}$$

(2.17)和(2.18)式分别叫做 Rushbroke 标度关系和 Widom 标度关系．

将 p,q 表为临界指数的函数对后面的推导有用：

$$p = 1/[\beta(1+\delta)], \tag{2.19}$$

$$q = \delta/(1+\delta). \tag{2.20}$$

(三) Kadanoff 的标度变换理论

在 Widom 提出他的标度理论的第二年(1966)，Kadanoff 就根据在临界点，序参量涨落的关联长度趋于无穷大的性质，提出了标度变换不变性的概念，试图从微

观上对 Widom 标度假设的物理含义作出解释. Kadanoff 还分析了关联函数和关联长度在标度变换下的变化形式,导出了两个新的标度关系.

考虑一个维数为 d,配位数(即最近邻数)为 r,格点距离为 a,格点自旋总数为 N 的 Ising 模型. 设想把格点晶格化成许多大小相同的块体(或集团),每一块体的边长为 $la(l>1)$,$la \ll \xi$(ξ 为临界点时的关联长度). 每个块体含 l^d 个格点,系统中的块体数为 N/l^d. 图 1 相当于 $d=2, l=3$ 的情况,每个块体(或集团)由 9 个格点自旋组成.

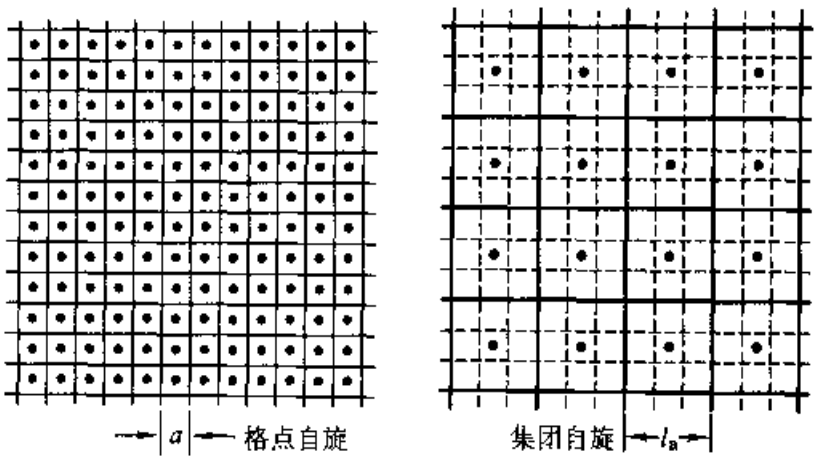


图 1　格点自旋和块体自旋

由于在临界点邻域,关联长度很大,一切有限大小的微观特征长度(如晶格常数 a)的影响都被抹去. 这时系统应具有某种尺度(大于 a,小于关联长度)变换下的不变性. 形象化地说,用放大倍数不同的显微镜来观察处于临界点邻域的系统时,看到的图像是相似的. 这种尺度变换下的不变性具体表现为用格点自旋和用块体自旋表达时,系统的哈密顿具有完全相同的形式:

$$H(\{s_i\}) = -J\sum_{\langle ij\rangle}^{rN/2} s_i s_j - H\sum_{i=1}^{N} s_i. \tag{2.21}$$

$$H(\{S_I\}) = -J_L\sum_{\langle IJ\rangle}^{rNl^{-d}/2} S_I S_J - H_L\sum_{I=1}^{Nl^{-d}} S_I. \tag{2.22}$$

其中,$s_i = \pm 1$ 为格点自旋,而按定义,$S_I = \pm 1$ 称为块体自旋,它正比于块体内自旋之和,见下面的(26)式. 为了简单起见,我们已取格点自旋磁矩 $\mu_0 = 1$. 系统哈密顿的这两种表达式形式上完全相同,只是格点自旋 s_i 由块体自旋 S_I 代替,格点晶格的耦合参数 (J,H) 由块体晶格耦合参数 (J_L,H_L) 代替. 因此可以预料,单个格点的自由能和单个块体的自由能对其宗量的依赖关系也具有相同的形式. 再考虑到单个块体中含有 l^d 个格点,于是有

$$g(t_L,H_L) = l^d g(t,H), \qquad (a \ll la \ll \xi). \tag{2.23}$$

t 是格点晶格的约化温度,H 是作用在格点上的磁场强度;t_L 是块体晶格的约化温度,H_L 是作用在块体上的有效磁场强度.

再来研究在 $a \to la$ 的标度变换下,t 和 t_L,H 和 H_L 的关系. 已知当系统处于临

界点时，序参量涨落的关联长度趋于无穷大. 若系统处于临界点邻域，则关联长度虽然很大，但仍然是一有限值. 当我们用块体晶格常数 la 来代替格点晶格常数 a 进行量度时，有效关联长度减小. 这相当 于把系统从临界点移开，因此有 $t_L > t$. 引入一个正参数 x ，可写出

$$t_L = tl^x, \qquad x \geqslant 0. \tag{2.24}$$

系统在外场中的能量为

$$H\sum_{i=1}^{N} s_i = H\sum_{I=1}^{Nl^{-d}}\sum_{i\in I} s_i = H\sum_{I=1}^{Nl^{-d}} S'_I. \tag{2.25}$$

注意，S'_I 是块体内自旋之和，它和“块体自旋” $S_I = \pm 1$ 不同. 若块体内全部格点自旋都向上，则有 $S'_I = l^d$ ，若块体内全部格点自旋都向下，则有 $S'_I = -l^d$ ，因此一般有 $-l^d \leqslant S'_I \leqslant l^d$. 若仍希望用 $S_I = \pm 1$ 表示块体自旋，则有

$$S'_I = ZS_I, \qquad S_I = \pm 1, \qquad Z \leqslant l^d. \tag{2.26}$$

将(2.26)式代入(2.25)式

$$H\sum_{i=1}^{N} s_i = H\sum_{I=1}^{Nl^{-d}} S'_I = HZ\sum_{I=1}^{Nl^{-d}} S_I = H_L\sum_{I=1}^{Nl^{-d}} S_I. \tag{2.27}$$

于是块体有效磁场强度和格点磁场强度的关系是

$$H_L = HZ.$$

或

$$H_L = l^y H, \qquad (0 \leqslant y \leqslant d). \tag{2.28}$$

将(2.24)和(2.28)式代入(2.23)式有 $g(l^x t, l^y H) = l^d g(t, H)$. 选 $l^d = \lambda$ (或 $l = \lambda^{1/d}$)，上式化为 $g(\lambda^{x/d}t, \lambda^{y/d}H) = \lambda g(t, H)$. 再令

$$x = pd, \qquad y = qd. \tag{2.29}$$

最后得到

$$g(\lambda^p t, \lambda^q H) = \lambda g(t, H). \tag{2.30}$$

即由 Kadanoff 的标度变换不变性推得了 Widom 关于自由能在临界点邻域是广义齐次函数的假设.

(四) Fisher 和 Josephson 标度关系

现在来分析进行标度变换时，关联函数和关联长度变化的形式.

引入块体自旋间的关联函数：

$$C_L(r_L, t_L) = \langle S_I S_J\rangle - \langle S_I\rangle\langle S_J\rangle. \tag{2.31}$$

r_L 表示用块体晶格常数 la 来计量时，块体 I 和块体 J 中心间的距离，t_l 表示块体晶格的约化温度. 利用(2.26)、(2.31)式化为

$$C_L(r_L, t_L) = Z^{-2}[\langle S'_I S'_J\rangle - \langle S'_I\rangle\langle S'_J\rangle] = Z^{-2}\sum_{i\in I}\sum_{j\in J}[\langle s_i s_j\rangle - \langle s_i\rangle\langle s_j\rangle]$$

$$= Z^{-2}(l^d)^2[\langle s_i s_j\rangle - \langle s_i\rangle\langle s_j\rangle] = Z^{-2}l^{2d}C(r,t). \tag{2.32}$$

在推导第三个等号时，已考虑每个块体含 l^d 个格点. 注意，第三个等号后的 s_i、s_j 与第二个等号后的不同，它们分别是块体 I 和块体 J 的中心格点自旋. $C(r,t)$ 代表格点自旋间的关联函数. t 表示格点晶格的约化温度，r 表示用格点晶格常数 a 来计量时，块体 I 和块体 J 的中心格点间的距离. 同一段距离用增大 l 倍的尺子计量，其结果必缩小 l 倍，因此有

$$r_L = r/l. \tag{2.33}$$

由(2.24)和(2.29)式可得

$$t_L = l^{pd}t. \tag{2.34}$$

将(2.33)和(2.34)式代入(2.32)式得 $C_L(l^{-1}r, l^{pd}t) = Z^{-2}l^{2d}C(r,t)$. 注意 $Z = l^y$ 和(29)第二式，上式化为 $C_L(l^{-1}r, l^{pd}t) = l^{2d(1-\eta)}C(r,t)$. 此式对任意 l 均成立，选 $l = r$ 并利用(20)式，得 $C(r,t) = r^{-2d/(1+\delta)}C_L(1, r^{pd}t)$. 令 $t = 0$，$C(r,0) = r^{-2d/(1+\delta)}C(1,0)$. 或

$$C(r,0) \propto r^{-2d/(1+\delta)}. \tag{2.35}$$

实际上这就是 §9.2.3(3.20)式：$C(r, t=0) \propto 1/r^{d-2+\eta}$. 两式比较得

$$d - 2 + \eta = 2d/(1+\delta). \tag{2.36}$$

用增大 l 倍的尺子计量，关联长度也应缩小 l 倍，即有 $\xi_L(t_L) = \xi_L(l^{pd}t) = \xi(t)/l$. 选 $l = t^{-1/(pd)}$，上式化为 $\xi(t) = t^{-1/(pd)}\xi_L(1)$. 或

$$\xi(t) \propto t^{-1/(pd)}. \tag{2.37}$$

将(2.37)式和 §9.2.3(3.19)式($\xi(t) \propto t^{-\nu}$)比较，得 $\nu = 1/(pd)$. 利用本节(2.19)式后得

$$\nu = \beta(1+\delta)/d. \tag{2.38}$$

利用已知的 Rushbrooke 标度关系(2.17)式和(2.18)式，(2.38)式进一步化为

$$\nu d = \beta(1+\delta) = 2 - \alpha. \tag{2.39}$$

(2.39)式被称为 Josephson 标度关系，它涉及到关联长度的临界指数 ν.

从(2.36)和(2.38)式中消除维数 d 可得 $\beta\delta - \beta = \nu(2-\eta)$. 或 $\nu(2-\eta) = \beta(\delta+1) - 2\beta = 2 - \alpha - 2\beta = \gamma$，这里已经用了(2.17)式和(2.18)式. 即

$$\gamma = \nu(2-\eta). \tag{2.40}$$

往往称(2.40)式为 Fisher 标度关系，它涉及到关联函数的临界指数 η.

最后简单小结标度理论的主要结果，有如下几个标度关系：

$$\begin{aligned} &\alpha + 2\beta + \gamma = 2 \quad \text{(Rushbroke)} \\ &\gamma = \beta(\delta - 1) \quad \text{(Widom)} \\ &\gamma = \nu(2-\eta) \quad \text{(Fisher)} \\ &\nu d = 2 - \alpha \quad \text{(Josephson)} \end{aligned} \tag{2.41}$$

由此可见，六个临界指数中只有两个是独立的，六个临界指数都可以用广义齐次函数参量 p,q 表示出来：

$$\alpha = \frac{2p-1}{q}, \qquad \beta = \frac{1-q}{p}, \qquad \gamma = \frac{2q-1}{p},$$
$$\delta = \frac{q}{1-q}, \qquad \eta = d(1-2q)+2, \qquad \nu = \frac{1}{pd}. \tag{2.42}$$

最后要强调的是，标度理论虽然推导了标度关系，但不能确定 p 和 q，因此也不能计算临界指数. 重整化群理论提供了从微观上计算 p,q 的方法，从而可以计算所有的临界指数.

* §9.3.3　重整化群方法简介

在§9.3.1 中我们学习了一维 Ising 模型的精确解. 二维 Ising 模型的精确解由 Onsager 获得，其数学过程十分复杂. 三维 Ising 模型的精确解，虽有很多人进行过尝试，但至今没有还成功. Ising 模型的某些近似解法虽然有数学过程简单的优点，但由于往往舍掉了一些重要位形，而使结果不理想. 为了克服以上困难，在 Kadanoff 标度变换不变性思想启发下，1971 年 Kenneth 和 Wilson 把量子场论中的重整化群理论用到统计物理中来，并因此获得了 1982 年的诺贝尔物理奖. Wilson 的方法是非常一般的，其应用范围远远超出了相变领域. 而在相变领域则可以看成是 Kadanoff 思想的推广和执行. 重整化群方法的基本思想，和 Kadanoff 标度变换理论的基本思想一样，都是：在临界点，关联长度趋于无穷，因此系统应具有尺度变换下的不变性. 但是，重整化群方法不去直接计算系统的配分函数，而是寻找使配分函数不变的尺度变换，从而确定临界点并计算临界指数.

为叙述方便，可以把重整化群方法（RG 方法）大致分为以下四步：第一，划分 Kadanoff 集团或进行尺度变换，即在保证拓扑形式不变的条件下，进行格点归并，把格点自旋晶格化为块体自旋晶格. 格点自旋晶格常数为 a，块体自旋晶格常数为 $l^n a$（n 是归并次数）. 第二步，寻找重整化群方程. 根据系统在临界点具有尺度变换下不变性的假设，找格点自旋耦合参量 k 和块体自旋耦合参量 k' 的关系，即 RG 方程 $k' = U(k)$. 第三，解 RG 方程，求方程的不稳定不动点——临界点和关涉量. 第四，利用关涉量计算临界指数.

虽然一维 Ising 模型不呈现相变，但因为它简单，下面我们仍通过一维 Ising 模型来介绍重整化群方法的大意.

考虑由 N 个 $s_i = \pm 1$ 的自旋粒子组成的一维 Ising 模型.

第一步，划分 Kadanoff 块体，计算块体配分函数 Z_0 和 $\langle s_i' \rangle$.

取三个格点为一个块体，如图 1 所示. 块体中心的距离是 $la = 3a$，系

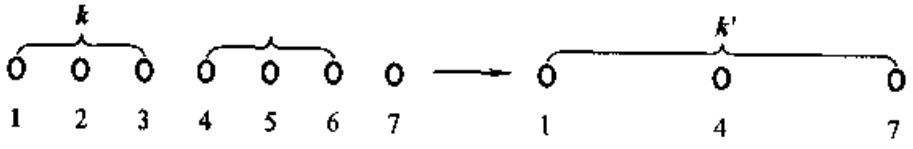


图 1　一维 Ising 模型中的格点归并

统中的块体数为 $N/l^d = N/3$，d 代表系统的维数.

现按少数服从多数的原则定义块体自旋：

$$S_I = \mathrm{sign}(s_1^I + s_2^I + s_3^I). \tag{3.1}$$

这里 s_i^I 表示块体 I 中第 i 个格点的自旋.(3.1)式的意思是，若有两个或三个格点自旋向上，则块体自旋 $S_I = +1$；若有两个或三个格点自旋向下，则块体自旋 $S_I = -1$.

不难看出当块体自旋 $S_I = +1$ 或 $S_I = -1$ 时，块体内具有四个内部自由度，用 $\sigma_\alpha^I(\alpha = 1,2,3,4)$ 表示，以 $S_I = +1$ 为例：

$$S_I = +1:\quad \begin{cases} s_1^I & s_2^I & s_3^I & \sigma_\alpha^I & \mathrm{e}^{-\mathscr{H}_{块}} \\ \uparrow & \uparrow & \downarrow & \sigma_1^I & \mathrm{e}^{K(1\cdot1+1\cdot(-1))} = \mathrm{e}^0 = 1 \\ \uparrow & \downarrow & \uparrow & \sigma_2^I & 1 \\ \downarrow & \uparrow & \uparrow & \sigma_3^I & 1 \\ \uparrow & \uparrow & \uparrow & \sigma_4^I & \mathrm{e}^{2K} \end{cases}$$

于是系统的微观态可用 $|s_i|$ 表示，也可用 $|S_I,\sigma_\alpha^I|$ 表示.

根据系综理论，块体的配分函数可表为

$$Z_0 = \sum_{\sigma_\alpha^I} \mathrm{e}^{-\mathscr{H}_{块}(\sigma_\alpha^I)}, \tag{3.2}$$

$\mathscr{H}_{块}$ 代表和某一内部自由度对应的块体哈密顿(含 β)：

$$\mathscr{H}_{块} = -[J/(k_BT)](s_1^Is_2^I + s_2^Is_3^I) = -K(s_1^Is_2^I + s_2^Is_3^I). \tag{3.3}$$

例如，和内部自由度 σ_1^I 对应，$\mathscr{H}_{块} = -K(1\cdot1+1\cdot(-1)) = 0$，和 σ_4^I 对应，$\mathscr{H}_{块} = -K(1\cdot1+1\cdot1) = -2K$(见上表)，于是有

$$Z_0 = 3 + \mathrm{e}^{2K}. \tag{3.4}$$

将块体当成恒温系统，按平均值公式可计算块体中某一自旋(如 s_1^I)的平均值：

$$\langle s_1^I\rangle = \frac{\sum s_1^I \mathrm{e}^{K(S(S_1^I+S_2^IS_3^I)}}{Z_0}.$$

当 $S_I = +1$ 时，有 $\langle s_1^I\rangle = (1+\mathrm{e}^{2K})/Z_0$；当 $S_I = -1$ 时，有 $\langle s_1^I\rangle = -(1+\mathrm{e}^{2K})/Z_0$，合并可写为

$$\langle s_i^I\rangle = S_I(1+\mathrm{e}^{2K})/Z_0. \tag{3.5}$$

第二步　寻找 RG 方程

以格点晶格为对象时，系统的哈密顿(含 β)和配分函数可表为

$$\mathscr{H}(K,|s_i|,N) = -K\sum_{i\neq j}{}' s_is_j - H\sum_i s_i. \tag{3.6}$$

$$Z(K,N) = \sum_{|S_i|} \mathrm{e}^{-\mathscr{H}(K,|S_i|,N)} = \sum_{S_I}\sum_{|\sigma_\alpha^I|} \mathrm{e}^{-\mathscr{H}(K,\,S_I,\sigma_\alpha^I,\,N)}. \tag{3.7}$$

其中 $K = (K,H) = (J\beta,H'\beta)$ 称为格点耦合参量(含温度，J 是最近邻格点的相互作用，H' 是作用于格点的外场). $\sum'$ 代表只对最近邻求和.

根据 Kadanoff 关于在临界点邻域系统具有标度变换不变性的假设，以块体晶格为对象时，系统的哈密顿和配分函数应和(3.6)、(3.7)式有相同的函数形式，即有

$$\mathscr{H}_L(\boldsymbol{K}_L,\{S_I\},Nl^{-d}) = -K_0 - K_L\sum_{I\neq J}{}'S_IS_J - H_L\sum_I S_I. \tag{3.8}$$

$$Z_L(\boldsymbol{K}_L,Nl^{-d}) = \sum_{\{S_I\}} e^{-\mathscr{H}_L(\boldsymbol{K}_L,\{S_I\},Nl^{-d})}. \tag{3.9}$$

其中 $\boldsymbol{K}_L=(K_L,H_L)=(J_L\beta,H'_L\beta)$ 称为块体耦合参量. J_L 是最近邻块体间的相互作用，H_L 是作用在块体自旋上的有效磁场，Nl^{-d} 是系统中的块体数.

因为不管是以格点晶格还是块体晶格为对象，写出的均是系统的配分函数，因此有 $\sum_{\{S_I\}} e^{-\mathscr{H}_L(\boldsymbol{K}_L,\{S_I\},Nl^{-d})} = \sum_{\{S_I\}}\sum_{\{\sigma_a^I\}} e^{-\mathscr{H}(\boldsymbol{K},\{S_I,\sigma_a^I\},N)}$，或有

$$e^{-\mathscr{H}_L(\boldsymbol{K}_L,\{S_I\},Nl^{-d})} = \sum_{\{\sigma_a^I\}} e^{-\mathscr{H}(\boldsymbol{K},\{S_I,\sigma_a^I\},N)}, \tag{3.10}$$

下面要做的是利用(3.10)式找到块体耦合参量 $\boldsymbol{K}_L$ 和格点耦合参量 $\boldsymbol{K}$ 之间的关系，即重整化群方程. 为此，以格点晶格为对象，将系统的哈密顿分成块体内和块体间两部分.

$$\begin{aligned}\mathscr{H}(\boldsymbol{K},\{s_i\},N) &= -K\sum_{i\neq j}{}'s_is_j - H\sum_i s_i\\ &= -K\sum_I\Big(\sum_{i\in I}\sum_{j\in I}{}'s_is_j\Big) - K\sum_{I\neq J}\Big(\sum_{i\in I}\sum_{j\in J}{}'s_is_j\Big) - H\sum_I\Big(\sum_{i\in I}s_i\Big).\end{aligned} \tag{3.11}$$

令

$$\mathscr{H}_0 = -K\sum_I\Big(\sum_{i\in I}\sum_{j\in I}{}'s_is_j\Big), \tag{3.12}$$

它代表系统中各块体内最近邻相互作用的总和. 令

$$V_{IJ} = -K\sum_{i\in I}\sum_{j\in J}{}'s_is_j, \tag{3.13}$$

它代表块体 I 和块体 J 之间由于最近邻格点自旋而产生的相互作用. 在目前所考虑的例子中，显然有

$$V_{IJ} = -Ks_1^Is_1^J. \tag{3.14}$$

再令

$$S'_I = \sum_{i\in I}s_i = s_1^I + s_2^I + s_3^I. \tag{3.15}$$

它代表块体中格点自旋之和，它和前面按少数服从多数原则定义的块体自旋不同. 引入

$$V = \sum_{I\neq J}V_{IJ} - H\sum_I S'_I, \tag{3.16}$$

利用(3.12)式和(3.16)式后，(3.11)式化为

$$\mathscr{H}(\boldsymbol{K},\{s_i\},N) = \mathscr{H}_0 + \sum_{I\neq J}V_{IJ} - H\sum_I S'_I = \mathscr{H}_0 + V, \tag{3.17}$$

V 代表块体间的相互作用和外场对块体的作用.

将(3.17)式代入(3.10)式，得

$$\begin{aligned}e^{-\mathscr{H}_L(\boldsymbol{K}_L,\{S_I\},Nl^{-d})} &= \sum_{\{\sigma_a^I\}}e^{-\mathscr{H}_0}\cdot e^{-V} = \sum_{\{\sigma_a^I\}}\Big(e^{-V}e^{-\mathscr{H}_0}\Big/\sum_{\{\sigma_a^I\}}e^{-\mathscr{H}_0}\Big)\sum_{\{\sigma_a^I\}}e^{-\mathscr{H}_0}\\ &= \langle e^{-V}\rangle\sum_{\{\sigma_a^I\}}e^{-\mathscr{H}_0} \cong e^{-\langle V\rangle}\sum_{\{\sigma_a^I\}}e^{-\mathscr{H}_0}.\end{aligned} \tag{3.18}$$

因为 $\mathscr{H}_0$ 代表系统中各块体内部哈密顿的总和，即

$$\mathscr{H}_0 = \mathscr{H}_0^I + \mathscr{H}_0^{II} + \cdots + \mathscr{H}_0^{Nl^{-d}}$$

$$\sum_{\{\sigma_0^L\}} e^{-\mathscr{H}_0} = \sum_{\sigma^I} e^{-\mathscr{H}_0^I} \sum_{\sigma^{II}} e^{-\mathscr{H}_0^{II}} \cdots = Z_0^M \tag{3.19}$$

$M = Nl^{-d}$ 代表系统中的块体数.

(3.18)式化为

$$e^{-\mathscr{H}_L(K_L, \{S_I\}, Nl^{-d})} = e^{-\langle V \rangle} Z_0^M,$$

或

$$\mathscr{H}_L(K_L, \{S_I\}, Nl^{-d}) = -M\ln Z_0 + \langle V \rangle. \tag{3.20}$$

注意(3.16)式有

$$\langle V \rangle = \sum_{I \neq J} \langle V_{IJ} \rangle - H \sum_I \langle S'_I \rangle.$$

再注意在本例中对 V_{IJ} 和 S'_I 有(3.14)和(3.15)式,则

$$\langle V \rangle = \sum_{I \neq J} -K\langle s_3^I s_1^J \rangle - H \sum_I [\langle s_1^I \rangle + \langle s_2^I \rangle + \langle s_3^I \rangle] \approx -K \sum_{I \neq J} \langle s_3^I \rangle \langle s_1^J \rangle - 3H\langle s_1^I \rangle. \tag{3.21}$$

将(3.4)式代入(3.5)式,再将(3.5)式代入上式得

$$\langle V \rangle = -K \sum_{I \neq J} [(1 + e^{2K})/(3 + e^{2K})]^2 S_I S_J - 3 \sum_I [(1 + e^{2K})/(3 + e^{2K})] H S_I. \tag{3.22}$$

将(3.22)式代入(3.20)式

$$\mathscr{H}_L(K_L, \{S_I\}, Nl^{-d}) = -M\ln Z_0 - K \sum_{I \neq J} [(1 + e^{2K})/(3 + e^{2K})]^2 S_I S_J$$

$$- 3 \sum_I [(1 + e^{2K})/(3 + e^{2K})] H S_I. \tag{3.23}$$

将(3.23)式和(3.8)式比较得

$$K_L = K[(1 + e^{2K})/(3 + e^{2K})]^2, \quad H_L = 3H[(1 + e^{2K})/(3 + e^{2K})]. \tag{3.24}$$

或

$$K_L = U(K), \tag{3.25}$$

其中 U 表示按(3.24)式由 K 得 K_L 的变换.(3.24)或(3.25)式给出了满足 Kadanoff 标度变换不变性要求的新、旧耦合参量的关系,即重整化群方程(RGE).

第三步,解 RG 方程.求 RG 方程的不稳定不动点——临界态.

(3.24)式是一组复杂的超越方程,利用(3.24)式或(3.25)式,从给定的 K 值可以求得一次归并后的 K';将所得 K' 做为新的 K 值代入(3.25)式右方,可求得二次归并后的 K'';……依次迭代下去.

一般说来,参数空间中一点 K 经变换 U 后将变到另一点 K',如果在某一点 K^* 经 U 变换后不变,即有

$$K^* = U(K^*). \tag{3.26}$$

则称 K^* 为变换 U 的不动点.(3.26)式就是确定不动点的方程.在只有两个参数的情况下,不动点可以简单地分为两类:稳定不动点和不稳定不动点.如果在不动点 K^* 附近重复进行变换时,越来越趋近 K^*,则称 K^* 为稳定不动点;反之,如果从不动点 K^* 附近某点开始重复进行变换后,越来越远离 K^*,则称该 K^* 为不稳定不动点.例如,迭代方程 $x_{n+1} = 2x_n(1 - x_n)(0 \leqslant x < 1)$ 有两个不动点:0 和 1/2.其中"0"是不稳定不动点,"1/2"是稳定不动点.

当系统处于临界状态时，$t \equiv (T - T_C)/T_C = 0$，关联长度 $r_C \to \infty$，进行格点归并(标度变换)，用长短不等的尺度来量度 r_C，结果不变. 因而期望在临界点有 $\boldsymbol{K}' \equiv \boldsymbol{K} \equiv \boldsymbol{K}^*$，换句话说，把临界点和 RG 方程(3.25)的不动点联系起来是合理的. 如果系统不是严格处于临界态，而是处在临界点邻域 $t \neq 0$，关联长度 r_C 为一很大的有限值，进行格点归并或用较大的尺量度时，r_C 的值减小，意味着 t 增大，或远离临界态，因此临界态应该是 RG 方程的不稳定不动点.

再回到我们的例子，不动点方程为

$$K^* = K^* \left(\frac{1 + e^{2K^*}}{3 + e^{2K^*}} \right)^2, \qquad H^* = 3H^* \left(\frac{1 + e^{2K^*}}{3 + e^{2K^*}} \right). \tag{3.27}$$

方程有两组解，分别为

$$K_1^* = 0, \qquad H_1^* = 0;$$

$$K_2^* = \infty, \qquad H_2^* = 0.$$

更详细的讨论和计算机迭代可以证明这两组解都是稳定不动点，不代表临界态. 换句话说，一维 Ising 模型不存在相变. 记住 $K \equiv J/(k_B T)$，则和 $K_1^* = 0$ 和 $K_2^* = \infty$ 对应的 T_C 分别为 $T_C = \infty$ 和 $T_C = 0$. 这也说明一维 Ising 模型不存在相变.

最后第四步应该是利用关涉量计算临界指数. 由于本例不存在相变，故这一步工作不能进行.

习　题

1. 设 N 个 Ising 自旋沿一闭合链排列. 系统的能量为

$$E(\{s_i\}) = -\varepsilon \sum_{i=1}^{N} s_i s_{i+1}, \qquad s_i = \pm 1.$$

利用矩阵法证明，系统的自由能为

$$F = -Nk_B T \ln \{2\mathrm{ch}[\varepsilon/(k_B T)]\},$$

并求比热和温度的关系.



第十篇　非平衡统计基础

到目前为止，我们讨论的是平衡态统计理论及围绕平衡态的微小偏离. 平衡态只是物质运动的一种特殊形态，而非平衡态则更具普遍性. 非平衡统计理论和平衡态统计理论几乎有着同样悠久的历史和相同的开创人. 早在 19 世纪 60－70 年代，玻尔兹曼和麦克斯韦一方面致力于平衡统计理论的研究，导出了以他们的名字命名的麦克斯韦－玻尔兹曼分布，另一方面又致力于非平衡统计理论的研究. 玻尔兹曼从分子运动论的观点出发，通过分析碰撞过程，导出了稀薄气体中单粒子非平衡概率密度随时间演化的玻尔兹曼积分微分方程，以及和熵增加原理对应的 H 定理. 然而，平衡态理论早在 20 世纪中期已发展得比较完备，而非平衡态理论的发展则经历了一个艰难、缓慢的过程，只是在近年来，才有了较大的发展，成为统计物理的研究重心，并与越来越多的领域（如等离子体物理、固体物理、化学反应动力学及生命的基本过程等）发生密切联系，受到广泛的重视.

本篇第一章按照玻尔兹曼原始的方法，推导稀薄气体中单粒子非平衡分布函数（或概率密度）满足的演化方程——玻尔兹曼积分微分方程及 H 定理，并将方程用于输运过程（稳衡过程，即非平衡稳定过程），导出输运系数的微观表达式. 第二章从统计力学的微观基础——刘维尔定理出发，推导出约化概率密度服从的方程序列——BBGKY 序列. 通过引入适当的近似或假设，可以把 BBGKY 序列在某阶 l 处截断，从而得到关于单粒子约化概率密度，双粒子约化概率密度，…，l 粒子约化概率密度的封闭方程组. 解方程组，求系综平均就可以得到系统宏观量的演化方程或宏观量. 从理论上说，玻尔兹曼积分微分方程只适用于稀薄气体，而 BBGKY 序列则是解决稠密气体乃至液体非平衡统计的系统方法. 但后者理论上比较抽象且数学上难度很大. 作为简单应用，我们利用它推导单粒子概率密度的符拉索夫方程，并再一次推导玻尔兹曼积分微分方程.

在本篇的学习中，希望读者特别注意的是，如何通过引入分子混沌性假设，把一个可逆的动力学方程转化为一个不可逆的宏观方程，从而加深对宏观不可逆性起源的认识.

第一章　稀薄气体中的输运过程

§ 10.1.1　玻尔兹曼方程

在平衡态统计理论中，知道了分布函数就可以求得和任何微观量对应的宏观量，而且根据普遍的论据就可以求得分布函数. 建立非平衡态统计理论则要困难得多. 在这一节里讨论比较简单的情况. 第一，只讨论偏离平衡态不远时的非平衡态；第二，只讨论经典稀薄气体；第三，只讨论单组元气体. 已经讲过，当分子热波长远小于分子平均距离时，即

$$h/\sqrt{2m\pi k_B T} \ll (V/N)^{1/3}. \tag{1.1}$$

可将分子看作经典粒子，利用经典力学讨论分子之间的碰撞过程.

我们用

$$f(\boldsymbol{r},\boldsymbol{v},t)\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v} \tag{1.2}$$

表示在 t 时刻，位于坐标 $\boldsymbol{r}$ 和速度 $\boldsymbol{v}$ 附近，$\mathrm{d}\boldsymbol{r} = \mathrm{d}x\mathrm{d}y\mathrm{d}z$ 体积元和 $\mathrm{d}\boldsymbol{v} = \mathrm{d}v_x\mathrm{d}v_y\mathrm{d}v_z$ 速度间隔中的分子数. 所取的 $\mathrm{d}\boldsymbol{r}$ 和 $\mathrm{d}\boldsymbol{v}$，从微观看应足够大，使其中含有大量分子；但从宏观看又足够小，使之可视为宏观的点. 这是可以做到的. 例如取 $\mathrm{d}\boldsymbol{r}$ 为 $10^{-10}\ \mathrm{cm}^3$，无疑这可视为宏观的点，但在标准状态下其中仍含有约 10^9 个分子. $f(\boldsymbol{r},\boldsymbol{v},t)$ 叫做稀薄气体的非平衡态分布函数. 确定了 $f(\boldsymbol{r},\boldsymbol{v},t)$，便可用统计方法求得稀薄气体在非平衡态的性质. 本节用弛豫时间近似方法来建立分布函数 $f(\boldsymbol{r},\boldsymbol{v},t)$ 满足的演化方程，即玻尔兹曼方程.

经过时间 $\mathrm{d}t$ 后，在时刻 $t + \mathrm{d}t$，位于同一体积元 $\mathrm{d}\boldsymbol{r}$ 和速度间隔 $\mathrm{d}\boldsymbol{v}$ 中的分子数将变为 $f(\boldsymbol{r},\boldsymbol{v},t + \mathrm{d}t)\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}$. 当 $\mathrm{d}t$ 很小时，可将上式作泰勒展开，只取头两项，得 $[f(\boldsymbol{r},\boldsymbol{v},t) + (\partial f/\partial t)\mathrm{d}t]\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}$. 减去(1.2)式，得到在 $\mathrm{d}t$ 时间内，$(\boldsymbol{r},\boldsymbol{v})$ 附近 $\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}$ 内增加的分子数为 $(\partial f/\partial t)\mathrm{d}t\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}$. $\partial f/\partial t$ 表示非平衡分布函数随时间的变化率.

$f(\boldsymbol{r},\boldsymbol{v},t)$ 随时间变化有两方面原因. 第一个原因是漂移变化，这又来源于两方面：速度使分子位置发生变化，某些分子由于运动离开了所考虑的 $\boldsymbol{r}$ 附近的体积元 $\mathrm{d}\boldsymbol{r}$，另一些分子由于运动进入了所考虑的 $\boldsymbol{r}$ 附近的体积元 $\mathrm{d}\boldsymbol{r}$；加速度使分子速度发生变化，当存在外场时，分子将被加速，某些分子由于速度变化离开了所考虑的 $\boldsymbol{v}$ 附近的速度间隔 $\mathrm{d}\boldsymbol{v}$，另一些分子由于速度变化进入了所考虑的 $\boldsymbol{v}$ 附近的速度间隔 $\mathrm{d}\boldsymbol{v}$. 第二个原因是碰撞变化. 因为斥力力程很短，碰撞在一个很小的区域中进行，即碰撞前后分子位于同一体积元中. 因此，和漂移变化不同，碰撞并不使 $\boldsymbol{r}$ 附近的 $\mathrm{d}\boldsymbol{r}$ 体积元中的分子数发生变化，而只是使 $\boldsymbol{v}$ 附近的 $\mathrm{d}\boldsymbol{v}$ 速度间隔中的分子数发

生变化.

我们先讨论漂移引起的分布函数的变化.

(1) $(\partial f/\partial t)_{漂}$

如图 1 所示,在 x,y,z 的直角坐标系中,体积元 $d\boldsymbol{r} = dxdydz$ 由三对平面围成,即

$$(x,x+dx),\quad 面积\ dA_x = dydz\quad (垂直于\ x\ 轴),$$
$$(y,y+dy),\quad 面积\ dA_y = dzdx\quad (垂直于\ y\ 轴),$$
$$(z,z+dz),\quad 面积\ dA_z = dxdy\quad (垂直于\ z\ 轴).$$

与此类似,在六维空间中,"体积元" $d\boldsymbol{r}d\boldsymbol{v}$ 则由六对"平面"围成.

$$(x,x+dx)\quad dA_x = dydzdv_xdv_ydv_z,$$
$$\cdots\cdots\qquad\cdots\cdots$$
$$(v_x,v_x+dv_x)\quad dA_{v_x} = dxdydzdv_ydv_z$$
$$\cdots\cdots\qquad\cdots\cdots$$

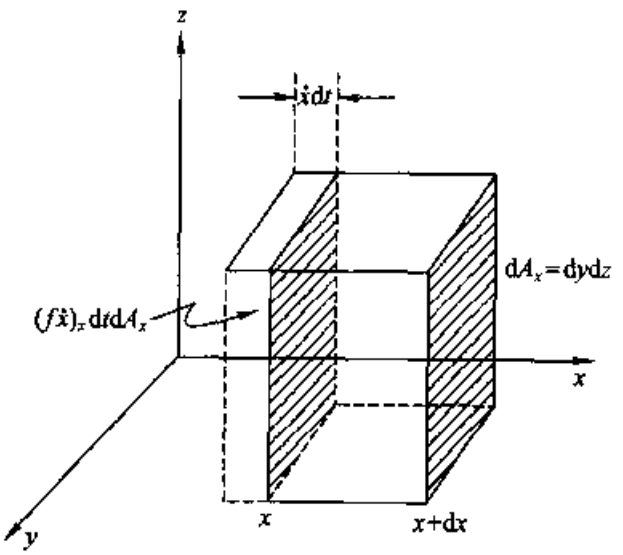


图 1 直角坐标系中的体积元

在 dt 时间内,通过 x "平面"进入 $d\boldsymbol{r}d\boldsymbol{v}$ 的分子数必位于以 dA_x 为底 $\dot{x}dt$ 为高的柱体内.将这一柱体的体积 $\dot{x}dtdA_x$ 乘上分布函数 f 就得到这柱体内的分子数 $(f\dot{x})_x dtdA_x$,其中下标 x 表示 $(f\dot{x})$ 取 x 处的值.在 dt 时间,通过 $x+dx$ "平面"走出 $d\boldsymbol{r}d\boldsymbol{v}$ 的分子数为 $(f\dot{x})_{x+dx}dtdA_x = [(f\dot{x})_x + (\partial(f\dot{x})/\partial x)dx]dtdA_x$.两式相减,得到通过位子 $(x,x+dx)$ 一对"平面"进入 $d\boldsymbol{r}d\boldsymbol{v}$ 的净分子数为

$$-\frac{\partial(f\dot{x})}{\partial x}dxdtdA_x = -\frac{\partial(f\dot{x})}{\partial x}dtd^3\boldsymbol{r}d^3\boldsymbol{v}.$$

经过类似讨论可得,在 dt 时间内,通过 (v_x,v_x+dv_x) 一对"平面"进入 $d\boldsymbol{r}d\boldsymbol{v}$ 的净分子数为

$$-\frac{\partial(f\dot{v}_x)}{\partial v_x}dtd^3\boldsymbol{r}d^3\boldsymbol{v}.$$

在 dt 时间,通过六对"平面"进入 $d\boldsymbol{r}d\boldsymbol{v}$ 的净分子数则为

$$-\left[\frac{\partial}{\partial x}(f\dot{v}_x)+\frac{\partial}{\partial y}(f\dot{v}_y)+\frac{\partial}{\partial z}(f\dot{v}_z)+\frac{\partial}{\partial v_x}(f\dot{v}_x)+\frac{\partial}{\partial v_y}(f\dot{v}_y)+\frac{\partial}{\partial v_z}(f\dot{v}_z)\right]dtd^3\boldsymbol{r}d^3\boldsymbol{v}. \tag{1.3}$$

考虑以下两点后,可将(1.3)式化简.

第一，分子速度与坐标无关，即 $\boldsymbol{r}$ 和 $\boldsymbol{v}$ 是独立变量，因而有 $(\partial v_x/\partial x) = (\partial v_y/\partial y) = (\partial v_z/\partial z) = 0$.

第二，分子所受外力分量与对应的速度分量无关. 用 $\boldsymbol{F}$ 表示单位质量所受的外力，其分量为 X, Y, Z，则根据牛顿第二定律有 $X = \dot{v}_x, Y = \dot{v}_y, Z = \dot{v}_z$. 在一般问题中所遇外力是重力和电磁力，重力和电力皆与分子速度无关，洛伦兹力 $\boldsymbol{v} \times \boldsymbol{B}$ 的分量 $X = (v_y B_z - v_z B_y)$ 也和 v_x 无关. 在这样的条件下，(1.3)式可简化为

$$\left(\frac{\partial f}{\partial t}\right)_{漂} = -\left[v_x \frac{\partial f}{\partial x} + v_y \frac{\partial f}{\partial y} + v_z \frac{\partial f}{\partial z}\right] - \left[\dot{v}_x \frac{\partial f}{\partial v_x} + \dot{v}_y \frac{\partial f}{\partial v_y} + \dot{v}_z \frac{\partial f}{\partial v_z}\right]$$

$$= -\boldsymbol{v} \cdot \nabla_{\boldsymbol{r}} f - \dot{\boldsymbol{v}} \cdot \nabla_{\boldsymbol{v}} f = -\boldsymbol{v} \cdot \nabla_{\boldsymbol{r}} f - (\boldsymbol{F}/m) \cdot \nabla_{\boldsymbol{v}} f. \quad (1.4)$$

(2) $(\partial f/\partial t)_{碰}$

§10.1.3 将对 $(\partial f/\partial t)_{碰}$ 进行详细讨论. 在这里，我们用弛豫时间近似作初步研究.

碰撞是使系统恢复平衡的机制. 设想由于加外场使系统偏离了平衡态，分布函数 f 则偏离了平衡分布 $f^{(0)}$. 若突然撤除外场，系统会由于分子碰撞而趋于平衡. 在外场较弱，偏离 $f - f^{(0)}$ 很小的情况下，可假设由于碰撞恢复平衡的速率 $(\partial f/\partial t)_{碰}$ 与对平衡的偏离 $f - f^{(0)}$ 成比例，即

$$(\partial f/\partial t)_{碰} = -\frac{1}{\tau}(f - f^{(0)}). \quad (1.5)$$

负号表示恢复平衡，$1/\tau$ 是比例系数. 由于平衡分布 $f^{(0)}$ 和时间无关，上式可改写为 $\partial(f - f^{(0)})/\partial t = -(f - f^{(0)})/\tau$. 积分上式，并取撤除外场的瞬间作为起始时刻 $t = 0$，$f(0)$ 表示 $t = 0$ 时的分布函数，则得

$$f(t) - f^{(0)} = [f(0) - f^{(0)}]\mathrm{e}^{-t/\tau}. \quad (1.6)$$

当 $t = \tau$ 时有 $f(\tau) - f^{(0)} = [f(0) - f^{(0)})\mathrm{e}^{-1}$，表明：经 τ 时间后，分布函数对平衡分布的偏离 $(f(\tau) - f^{(0)})$ 减小到初始偏离 $(f(0) - f^{(0)})$ 的 e 分之一. τ 愈大，表示由非平衡趋向平衡愈慢，因而可用 τ 表示趋向平衡的快慢. 趋向平衡的过程叫弛豫过程，τ 称为弛豫时间. 可以证明，弛豫时间依赖于速度，即有 $\tau = \tau(v)$.

将 $(\partial f/\partial t)_{漂}$ 和 $(\partial f/\partial t)_{碰}$ 相加就得到分布函数随时间演化满足的方程：

$$\frac{\partial f}{\partial t} = \left(\frac{\partial f}{\partial t}\right)_{漂} + \left(\frac{\partial f}{\partial t}\right)_{碰} = -\boldsymbol{v} \cdot \nabla_{\boldsymbol{r}} f - \dot{\boldsymbol{v}} \cdot \nabla_{\boldsymbol{v}} f - \frac{f - f^{(0)}}{\tau}$$

$$= -\left[v_x \frac{\partial f}{\partial x} + v_y \frac{\partial f}{\partial y} + v_z \frac{\partial f}{\partial z}\right] - \left[X \frac{\partial f}{\partial v_x} + Y \frac{\partial f}{\partial v_y} + Z \frac{\partial f}{\partial v_z}\right] - \frac{f - f^{(0)}}{\tau}. \quad (1.7)$$

(1.7)式叫做玻尔兹曼方程.

对子输运过程(稳定非平衡过程)，$\partial f/\partial t = 0$，由上式得

$$v_x\frac{\partial f}{\partial x}+v_y\frac{\partial f}{\partial y}+v_z\frac{\partial f}{\partial z}+X\frac{\partial f}{\partial v_x}+Y\frac{\partial f}{\partial v_y}+Z\frac{\partial f}{\partial v_z}=-\frac{f-f^0}{\tau}. \tag{1.8}$$

下面讨论(1.8)式对输运过程的应用.

§10.1.2 气体粘滞系数和金属电导率的计算

输运过程是一种简单的非平衡过程——稳定非平衡过程. 玻尔兹曼方程应用的一个重要方面就是推导各种输运过程系数的微观表达式. 作为例子,在此仅研究气体的粘滞系数和电导率. 我们将看到,近似计算往往仍能给出很有价值的结果.

(一)气体粘滞系数的计算

粘滞现象如图1所示. 气体以宏观速度 v_0 沿 y 方向流动,v_0 是 x 的函数. 考虑 x_0 处的平面,称 $x>x_0$ 的一方为 x_0 的正方,称 $x<x_0$ 的一方为 x_0 的负方. 实验发现,流动速度较快的正方将带动负方使其增速,同时使自己减速. 这种现象称为粘滞现象. 以 p_{xy} 表示正方通过单位面积施于负方的作用力. 脚标 x 代表单位面积的法线方向,脚标 y 代表力的方向. 实验测得 p_{xy} 和宏观流速的梯度成正比,即

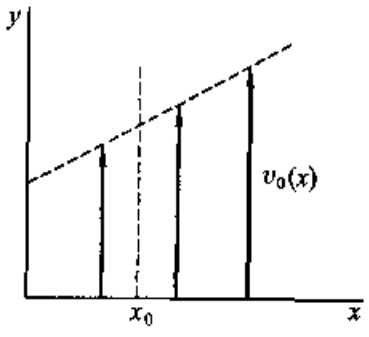


图1 粘滞现象示意图

$$p_{xy}=\eta(\mathrm{d}v_0(x)/\mathrm{d}x). \tag{2.1}$$

(2.1)式称为牛顿粘滞定律,η 称为粘滞系数. 因为流体动量对时间的导数等于力,p_{xy} 也代表单位时间内,穿过单位面积从正方输运到负方的沿 y 方向的净动量.

在 x_0 平面上取单位面积元,考虑速度介于 $v_x\sim v_x+\mathrm{d}v_x$,$v_y\sim v_y+\mathrm{d}v_y$ 和 $v_z\sim v_z+\mathrm{d}v_z$ 中的分子. 显然,要在单位时间中能穿过此单位面积元,由负方进入正方的分子必位于以 v_x 为高,单位面积元为底的柱体内. 该柱体体积为 v_x,其中速度在所述间隔 $\mathrm{d}\boldsymbol{v}$ 中的分子数为

$$f v_x\mathrm{d}v_x\mathrm{d}v_y\mathrm{d}v_z.$$

每一个分子携带沿 y 方向的动量 mv_y. 将各种速度范围的分子所携带动量相加,就得到单位时间中由穿过单位面积元从负方进入正方的分子所输运的沿 y 方向的总动量

$$\int_0^\infty\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}mv_yv_xf\,\mathrm{d}v_x\mathrm{d}v_y\mathrm{d}v_z.$$

同理,在单位时间由穿过单位面积元从正方进入负方的分子所输运到负方的沿 y 方向的总动量

$$-\int_{-\infty}^0\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}mv_yv_xf\,\mathrm{d}v_x\mathrm{d}v_y\mathrm{d}v_z.$$

其中负号是因为只有 v_x 为负的分子才能由正方进入负方. 两式相减就得到,在单位时间中,通过单位面积由正方输运到负方的沿 y 方向的净总动量,即

$$p_{xy} = -\iiint_{-\infty}^{\infty} m v_y v_x f \, \mathrm{d}v_x \mathrm{d}v_y \mathrm{d}v_z. \tag{2.2}$$

其中 f 为非平衡分布函数,由玻尔兹曼方程决定.

在气体作稳定流动的情况下有 $\partial f / \partial t = 0$. 因为没有外力,所以 $X = Y = Z = 0$;只有沿 x 方向存在非均匀性,故又有 $\partial f / \partial y = \partial f / \partial z = 0$,于是上节(1.8)式化为

$$v_x (\partial f / \partial x) = -(f - f^{(0)})/\tau,$$

或

$$f = f^{(0)} - \tau v_x (\partial f / \partial x).$$

采用逐级迭代近似,在第一级近似下,将上式右方修正项中的 f 用 $f^{(0)}$ 代替,得到

$$f = f^{(0)} - \tau v_x (\partial f^{(0)} / \partial x). \tag{2.3}$$

其中 $f^{(0)}$ 是局域平衡分布函数,具有麦克斯韦分布的形式,即 $f^{(0)} = n[m/(2\pi k_B T)]^{3/2} \mathrm{e}^{-[m/(2k_B T)][(v_x - v_{x0})^2 + (v_y - v_{y0})^2 + (v_z - v_{z0})^2]}$,$\boldsymbol{v}_0$ 是气体的宏观流速. 因为是局域平衡,因此一般来说,n, T, v_{x0}, v_{y0} 和 v_{z0} 是 $\boldsymbol{r}$ 的函数. 在目前情况下,局域平衡的分布函数 $f^{(0)}$ 应为

$$f^{(0)} = n[m/(2\pi k_B T)]^{3/2} \mathrm{e}^{-\{m/(2k_B T)\}[v_x^2 + (v_y - v_0)^2 + v_z^2]}. \tag{2.4}$$

这里 n, T, v_x, v_y 和 v_z 都和 $\boldsymbol{r}$ 无关,$f^{(0)}$ 仅通过 v_0 而依赖于 x,因此(2.3)式化为

$$f = f^{(0)} + \tau v_x (\partial f^{(0)} / \partial (v_y - v_0))(\mathrm{d}v_0 / \mathrm{d}x). \tag{2.5}$$

将(2.5)式代入(2.2)式,得

$$p_{xy} = -\iiint_{-\infty}^{\infty} m v_y v_x f^{(0)} \mathrm{d}v_x \mathrm{d}v_y \mathrm{d}v_z - m \frac{\mathrm{d}v_0}{\mathrm{d}x} \iiint_{-\infty}^{\infty} \tau v_y v_x^2 \frac{\partial f^{(0)}}{\partial (v_y - v_0)} \mathrm{d}v_x \mathrm{d}v_y \mathrm{d}v_z.$$

因为 $f^{(0)}$ 是 v_x, v_y 的偶函数,所以第一项积分为零. 一般来说 τ 是 v 的函数,进一步假设 τ 为常数,得

$$p_{xy} = -m\tau \frac{\mathrm{d}v_0}{\mathrm{d}x} \iiint_{-\infty}^{\infty} v_y v_x^2 \frac{\partial f^{(0)}}{\partial (v_y - v_0)} \mathrm{d}v_x \mathrm{d}v_y \mathrm{d}v_z. \tag{2.6}$$

将上式和牛顿粘滞定律(即(2.1)式)比较,得粘滞系数为

$$\eta = -m\tau \iint_{-\infty}^{\infty} v_x^2 \mathrm{d}v_x \mathrm{d}v_z \int_{-\infty}^{\infty} v_y \frac{\partial f^{(0)}}{\partial (v_y - v_0)} \mathrm{d}v_y.$$

利用分部积分

$$\int_{-\infty}^{\infty} v_y \frac{\partial f^{(0)}}{\partial (v_y - v_0)} \mathrm{d}v_y = [f^{(0)} v_y]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f^{(0)} \mathrm{d}v_y = -\int_{-\infty}^{\infty} f^{(0)} \mathrm{d}v_y.$$

$$\eta = m\tau\iiint_{-\infty}^{\infty} v_x^2 f^{(0)}\,\mathrm{d}v_x\mathrm{d}v_y\mathrm{d}v_z = nm\tau\langle v_x^2\rangle.$$

根据能量均分定理 $m\langle v_x^2\rangle/2 = k_B T/2$，因此 η 为

$$\eta = nk_B T\tau. \tag{2.7}$$

取弛豫时间近似为分子飞行平均自由程 $\langle l\rangle$ 所需时间，则可写

$$\langle l\rangle = \tau\langle v\rangle, \tag{2.8}$$

(2.7)式化为

$$\eta = nk_B T\langle l\rangle/\langle v\rangle. \tag{2.9}$$

平均自由程 $\langle l\rangle$ 与单位体积内的分子数 n 成反比，平均 $\langle v\rangle$ 与 $\sqrt{T}$ 成正比，可得

$$\eta \propto \sqrt{T}. \tag{2.10}$$

上式表明，在温度一定时，η 与压强无关，这一结论已得到实验证实.

(二)金属电导率的计算

实验告诉我们，当导体置于一恒定、均匀、沿 x 方向的外电场 $\mathscr{E}$ 中时，将在导体中形成沿 x 方向的稳恒电流，电流密度与场强成正比：

$$j_x = \sigma\mathscr{E}_x. \tag{2.11}$$

σ 叫电导率，和温度成正比. 这就是欧姆定律.

宏观电流形成的微观解释是，在外电场的作用下，金属中的自由电子沿电场的负方向被加速，从而形成电荷的定向运动——电流. 但如果只有这一种作用，电子定向运动的速度，从而电流密度 j_x 会愈来愈大，电阻趋于零. 然而，实际情况是：自由电子将与晶格缺陷不断碰撞发生散射，从而失去由电场加速得到的速度，然后再一次被加速，再一次散射，……. 当外场一定时，最后会达到一个具有稳恒电流的状态. 晶格缺陷包括热振动，晶格破缺及杂质原子等.

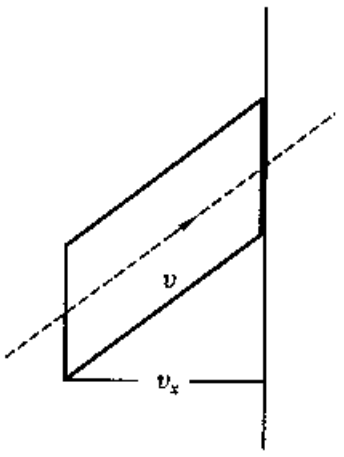


图 2 通过单位面积元的电子

下面我们考虑无破缺、无杂质的理想晶格的热振动，用玻尔兹曼方程导出欧姆定律，从而找到电导率和微观量的关系.

电流密度的定义是

$$\text{电流密度 } j_x = \begin{matrix}\text{单位时间通过垂直于外场方向}\\ \text{单位面积的电子数}\end{matrix} \times(-e). \tag{2.12}$$

在垂直于外场方向(x 方向)取一单位面积元，如图 2. 考虑速度介于 $v_x \sim v_x + \mathrm{d}v_x$，$v_y \sim v_y + \mathrm{d}v_y$，$v_z \sim v_z + \mathrm{d}v_z$ 间隔中的电子. 显然，能在单位时间中通过所取面积元的

电子必位于以所取单位面积元为底，以 $\boldsymbol{v}$ 为轴，以 v_x 为高的斜柱体内，斜柱体的体积为 v_x，其中速度介于 $v_x \sim v_x + \mathrm{d}v_x, v_y \sim v_y + \mathrm{d}v_y, v_z \sim v_z + \mathrm{d}v_z$ 的电子数为 $f(\boldsymbol{r},\boldsymbol{v},t)v_x\mathrm{d}v_x\mathrm{d}v_y\mathrm{d}v_z$. 将上式对 v_x, v_y 和 v_z 积分，得到单位时间通过所取面积元的总电子数. 代入(2.12)式得

$$j_x = -e\iiint_{-\infty}^{\infty} v_x f(\boldsymbol{r},\boldsymbol{v},t)\,\mathrm{d}v_x\mathrm{d}v_y\mathrm{d}v_z. \tag{2.13}$$

当存在均匀不变的外电场时，稳恒状态下的分布函数由 §10.1.1(1.8)式确定，在所讨论的情况下有

$$\partial f/\partial x = \partial f/\partial y = \partial f/\partial z = 0, Y = Z = 0, \quad X = -e\mathscr{E}_x/m, \tag{2.14}$$

(1.8)式化为 $(e\mathscr{E}_x/m)(\partial f/\partial v_x) = (f - f^{(0)})/\tau$，或

$$f = f^{(0)} + (e\mathscr{E}_x/m)\tau(\partial f/\partial v_x). \tag{2.15}$$

设外电场很弱，f 对 $f^{(0)}$ 偏离很小，采用逐级迭代近似，在一级近似下有

$$f = f^{(0)} + (e\mathscr{E}_x/m)\tau(\partial f^{(0)}/\partial v_x). \tag{2.16}$$

$f^{(0)}$ 代表局域平衡态下，$\boldsymbol{r},\boldsymbol{v}$ 附近单位体积元和单位速度间隔中的电子数. $f^{(0)}$ 应等于 $\boldsymbol{r},\boldsymbol{v}$ 附近单位体积元、单位速度间隔中的量子态数乘每个量子态上的平均电子数(即费米分布)，由此可得

$$f^{(0)}(\boldsymbol{r},\boldsymbol{v}) \approx (2m^3/h^3)\langle n(\boldsymbol{p})\rangle. \tag{2.17}$$

上式中 $\langle n(\boldsymbol{p})\rangle = 1/(e^{\beta(p^2/2m-\mu)}+1)$ 是费米分布. 因子2是考虑到电子自旋的两个可能取向而引入的.

将(2.16)式代入(2.13)式得

$$j_x = -e\iiint_{-\infty}^{\infty} v_x\left[f^{(0)} + \frac{e\mathscr{E}_x}{m}\tau\frac{\partial f^{(0)}}{\partial v_x}\right]\mathrm{d}v_x\mathrm{d}v_y\mathrm{d}v_z.$$

因为 $f^{(0)}$ 是 v_x 的偶函数，所以第一项积分为零，再考虑(2.17)式得

$$j_x = -\frac{2m^3}{h^3}\frac{e^2\mathscr{E}_x}{m}\iiint_{-\infty}^{\infty} v_x\tau\frac{\partial\langle n(\varepsilon)\rangle}{\partial\varepsilon}\frac{\partial\varepsilon}{\partial v_x}\mathrm{d}v_x\mathrm{d}v_y\mathrm{d}v_z.$$

对于费米分布 $\partial\langle n(\varepsilon)\rangle/\partial\varepsilon$ 只在 $\varepsilon \approx \varepsilon_F$ 附近不为零，这意味着只有位于费米面附近的电子对电导率有贡献. 因此，可以在上式中近似取 $\tau(v) \approx \tau(v_F) \approx \tau_F$ 并提到积分号外，则

$$j_x \approx -\frac{e^2\mathscr{E}_x}{m}\tau_F\iiint_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} v_x\frac{\partial\langle n(\varepsilon)\rangle}{\partial v_x}\mathrm{d}v_x\right\}\mathrm{d}v_y\mathrm{d}v_z\frac{2m^3}{h^3}.$$

首先利用分部积分法得

$$\int_{-\infty}^{\infty} v_x\frac{\partial\langle n(\varepsilon)\rangle}{\partial v_x}\mathrm{d}v_x = \langle n(\varepsilon)\rangle v_x\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty}\langle n(\varepsilon)\rangle\mathrm{d}v_x = -\int_{-\infty}^{\infty}\langle n(\varepsilon)\rangle\mathrm{d}v_x,$$

$$j_x \approx -\frac{e^2\mathscr{E}_x}{m}\tau_F\iiint_{-\infty}^{\infty}\langle n(\varepsilon)\rangle\mathrm{d}v_x\mathrm{d}v_y\mathrm{d}v_z\frac{2m^3}{h^3}.$$

将(2.17)式代入,得

$$j_x \approx \frac{e^2\mathscr{E}_x}{m}\tau_F\iiint_{-\infty}^{\infty} f^{(0)}(\boldsymbol{r},\boldsymbol{v})\,dv_x dv_y dv_z \approx \frac{e^2\mathscr{E}_x}{m}\tau_F n. \tag{2.18}$$

n 为单位体积中各种速度电子数的总和,即电子数密度.将(2.18)式和(2.11)式比较,可知有

$$\sigma = (e^2/m)\tau_F n. \tag{2.19}$$

要得到进一步结果,需要详细分析自由电子的碰撞机制求出 τ_F.现只作定性讨论.在高温下自由电子主要受离子振动的散射.以 l_F 和 v_F 表示 $\varepsilon \approx \varepsilon_F$ 附近电子的自由程和速率.则有

$$\tau_F = l_F/v_F. \tag{2.20}$$

v_F 对温度只有微弱的依赖关系.电子与离子碰撞的频率与离子振动位移的方均值 $\langle q^2\rangle$ 成正比,因而

$$l_F \propto 1/\langle q^2\rangle. \tag{2.21}$$

又据能量均分定理,离子振动的势能

$$A\langle q^2\rangle/2 = k_B T/2. \tag{2.22}$$

于是由(2.21),(2.20)和(2.19)式得到

$$\tau_F \propto 1/T, \qquad \sigma \propto 1/T. \tag{2.23}$$

与实验结果一致.

对金属电导率的贡献主要来源于费米面附近的电子可作如下说明.

已知金属电子处于平衡态时,分布函数可写为

$$f^{(0)}(\varepsilon) = \frac{1}{e^{\beta[(p_x^2+p_y^2+p_z^2)/2m-\mu]}+1} = f^{(0)}\left(\frac{1}{2m}(p_y^2+p_z^2)+\frac{1}{2m}p_x^2\right). \tag{2.24}$$

即 $f^{(0)}(\varepsilon)$ 对 v_x 具有对称性,且有

当 $\varepsilon < \mu_F, v_x < v_F$ 时 $\quad f^{(0)}(\varepsilon) = 1,$

当 $\varepsilon > \mu_F, v_x > v_F$ 时 $\quad f^{(0)}(\varepsilon) = 0.$

当未加外场时,费米分布的对称中心在 $p_x = 0$,如图3的虚线所示,沿 x 正、负方向运动的电子互相抵消,沿 x 方向的总电流为零.

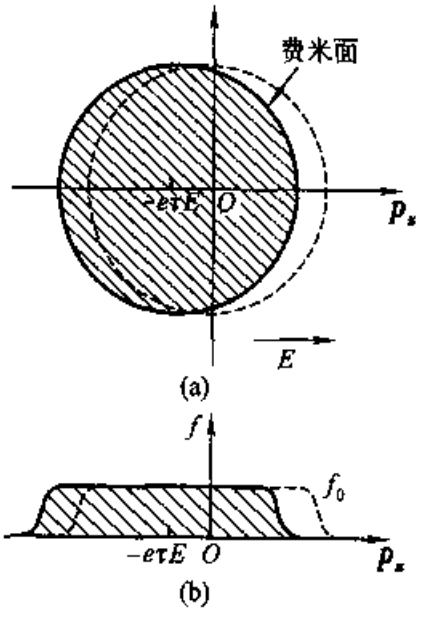


图3 无外场和有外场的费米分布

加外场后的分布函数是(2.16)式.因 τ 是小量,故可按泰勒展开公式将(2.16)式改写为

$$f = f^{(0)}[(p_y^2 + p_z^2)/(2m) + p_x^2/(2m) + \tau e(p_x/m)\mathscr{E}_x]$$
$$\approx f^{(0)}[(p_y^2 + p_z^2)/(2m) + (p_x + \tau e\mathscr{E}_x)^2/(2m)]. \quad (2.25)$$

和(2.24)式比较可知，加外场后的费米分布可看成与原来的费米分布具有相同形式，只是对称中心由原来的 $p_x = 0$ 移至 $p_x = -\tau e\mathscr{E}_x$，如图3的实线所示. 由图不难看出，这时有一部分电子的电流不能抵消，以致形成宏观电流. 换句话说，主要是费米面附近的电子参与导电，从而对电导率作出最大的贡献.

* §10.1.3　玻尔兹曼积分微分方程

在§10.1.1中，我们对分布函数 $f(\boldsymbol{r},\boldsymbol{v},t)$ 的碰撞变化率 $(\partial f/\partial t)_{碰}$ 采用了弛豫时间近似，得到了一个关于分布函数 f 的线性方程，便于求解. 但方程中含有弛豫时间 τ，要从理论上计算 τ，仍然需要分析分子的碰撞过程. 详细分析碰撞过程后，可以得到一个关于 f 的积分微分方程，称为玻尔兹曼积分微分方程.

在以 x,y,z,v_x,v_y,v_z 为坐标的分子相空间中，取某一确定的小相体积元 $\mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v}$（见图1）. t 时刻在该相体积元中的分子数为

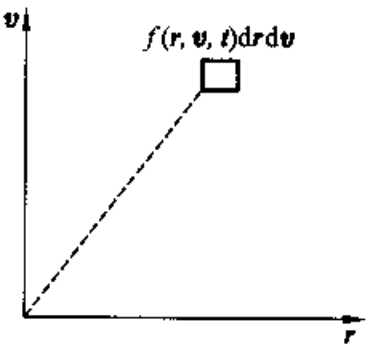


图1　分子相空间

$$f(\boldsymbol{r},\boldsymbol{v},t)\,\mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v}. \quad (3.1)$$

在 δt 时间内，有的分子由于碰撞速度发生变化，从而离开了该相体积元（由于斥力的力程很短，碰撞前后分子坐标不变，所以确切地说是离开了该速度间隔），使该相体积元中的分子数减少. 这部分减少量可表为

$$(\partial f/\partial t)_- \mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v}\delta t. \quad (3.2)$$

当然，也会有些原本不具有速度 $\boldsymbol{v} \sim \boldsymbol{v} + \mathrm{d}\boldsymbol{v}$ 的分子，会由于碰撞改变速度而进入该相体积元，使该相体积元中的分子数增加. 这部分增加量可表为

$$(\partial f/\partial t)_+ \mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v}\delta t. \quad (3.3)$$

将(3.3)式和(3.2)式相减，得到在时间间隔 δt 内，由于碰撞而引起的该相体积元中净增加的分子数

$$(\partial f/\partial t)_{碰}\,\mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v}\delta t = [(\partial f/\partial t)_+ - (\partial f/\partial t)_-]\mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v}\delta t. \quad (3.4)$$

$(\partial f/\partial t)_{碰}$ 就是我们要研究的分布函数的碰撞变化率.

因为碰撞前后分子的性质没有发生变化，玻尔兹曼假设分子碰撞为弹性碰撞. 在具体推求 $(\partial f/\partial t)_{碰}$ 之前，有必要简短复习一下§1.3.4中关于弹性碰撞的一些结论. 为简单起见，考虑同类分子的碰撞.

（一）分子的弹性碰撞

考虑两个弹性碰撞的分子. 在实验室系中，碰撞前后的速度分别为 $\boldsymbol{v}_1,\boldsymbol{v}_2$ 和 $\boldsymbol{v}'_1,\boldsymbol{v}'_2$. 则在质心坐标系中，根据动量守恒和能量守恒可得[见§1.3.4(4.1)、(4.2)和(4.4)式]

$$\boldsymbol{p}_{10}+\boldsymbol{p}_{20}=0,\qquad \boldsymbol{p}_{10}'+\boldsymbol{p}_{20}'=0,$$

$$|\boldsymbol{p}_{10}|=|\boldsymbol{p}_{20}|=|\boldsymbol{p}_{10}'|=|\boldsymbol{p}_{20}'|.$$

若取 $m_1=m_2=m$（同类分子相碰），则以上三式化为

$$\boldsymbol{v}_{10}=-\boldsymbol{v}_{20},\qquad \boldsymbol{v}_{10}'=-\boldsymbol{v}_{20}', \tag{3.5}$$

$$|\boldsymbol{v}_{10}|=|\boldsymbol{v}_{20}|=|\boldsymbol{v}_{10}'|=|\boldsymbol{v}_{20}'|. \tag{3.6}$$

即：在质心系中，碰撞前后两分子的速度总是大小相等方向相反. 碰撞不改变速度的大小，而只是改变了速度的方向.

注意，碰撞前后在实验室系中质心速度分别为 $\boldsymbol{v}_C=(\boldsymbol{v}_1+\boldsymbol{v}_2)/2$，$\boldsymbol{v}_C'=(\boldsymbol{v}_1'+\boldsymbol{v}_2')/2$，两分子的相对速度分别为 $\boldsymbol{u}=(\boldsymbol{v}_1-\boldsymbol{v}_2)$，$\boldsymbol{u}'=(\boldsymbol{v}_1'-\boldsymbol{v}_2')$，于是又可得

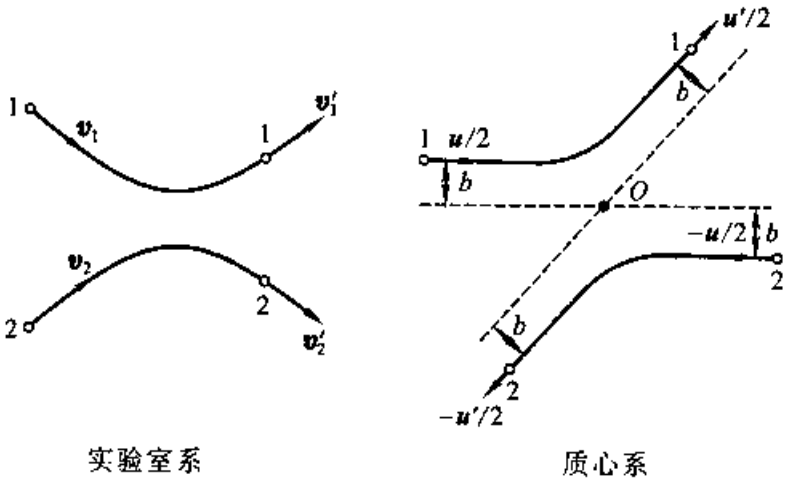
$$\boldsymbol{v}_{10}=\boldsymbol{v}_1-\boldsymbol{v}_C=\boldsymbol{u}/2,\qquad \boldsymbol{v}_{20}=\boldsymbol{v}_2-\boldsymbol{v}_C=-\boldsymbol{u}/2, \tag{3.7}$$

$$\boldsymbol{v}_{10}'=\boldsymbol{v}_1'-\boldsymbol{v}_C'=\boldsymbol{u}'/2,\qquad \boldsymbol{v}_{20}'=\boldsymbol{v}_2-\boldsymbol{v}_C'=-\boldsymbol{u}'/2. \tag{3.8}$$

由(3.5)式和(3.6)式可得

$$|\boldsymbol{u}|=|\boldsymbol{u}'|. \tag{3.9}$$

以上分析如图 2 所示.



实验室系　　质心系

图 2　两个分子的碰撞

由图 2 不难看出，在质心坐标系中我们只要把注意力集中于一个分子就够了，另一个分子的运动总是相反的. 两个分子的弹性碰撞问题可等价于一个分子受到虚拟的中心（由“O”代表）所施的力的散射问题，而该分子以速度 $\boldsymbol{u}$，瞄准距离 b 向“O”运动. 根据角动量守恒有 $b|\boldsymbol{u}|=b'|\boldsymbol{u}'|$，

$$b=b'. \tag{3.10}$$

即：碰撞前后，碰撞参数的大小也不变.

下面来计算 $(\partial f/\partial t)_-\,\mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v}\delta t$ 和 $(\partial f/\partial t)_+\,\mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v}\delta t$

（二）$(\partial f/\partial t)_+\,\mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v}\delta t$

为叙述方便，我们将在 $\boldsymbol{r}$ 处 $\mathrm{d}^3\boldsymbol{r}$ 内，碰撞前后两分子的速度分别为 $\boldsymbol{v},\boldsymbol{v}_1$ 和 $\boldsymbol{v}',\boldsymbol{v}_1'$ 的碰撞称为正碰撞. 而把碰撞前后两分子的速度分别为 $\boldsymbol{v}',\boldsymbol{v}_1'$ 和 $\boldsymbol{v},\boldsymbol{v}_1$ 的碰撞称为逆碰撞，

$$\text{正碰撞：}\quad |\boldsymbol{v},\boldsymbol{v}_1\to\boldsymbol{v}',\boldsymbol{v}_1'|,\qquad \text{逆碰撞：}\quad |\boldsymbol{v}',\boldsymbol{v}_1'\to\boldsymbol{v},\boldsymbol{v}_1|.$$

每发生一次正碰撞，在 $\boldsymbol{r}$ 处 $\mathrm{d}^3\boldsymbol{r}$ 内速度为 $\boldsymbol{v}\sim\boldsymbol{v}+\mathrm{d}\boldsymbol{v}$ 的分子就减少一个；每发生一次逆碰撞，在 $\boldsymbol{r}$ 处 $\mathrm{d}^3\boldsymbol{r}$ 内速度在 $\boldsymbol{v}\sim\boldsymbol{v}+\mathrm{d}\boldsymbol{v}$ 的分子就增加一个.

先把注意力集中于位于 $\boldsymbol{r}$ 处 $\mathrm{d}^3\boldsymbol{r}$ 内速度为 $\boldsymbol{v}$ 的一个分子，把它看成不动，称为靶分子. 以靶分子的中心为球心，画一作用球. 作用球的半径为分子力的有效力程 r_e［见图 3(b)］. 和靶分子

相碰的分子称为入射分子，入射分子以相对速度 $\boldsymbol{u}=\boldsymbol{v}_1-\boldsymbol{v}$ 向靶分子运动. 当入射分子的中心位于作用球面上时，则发生正碰.

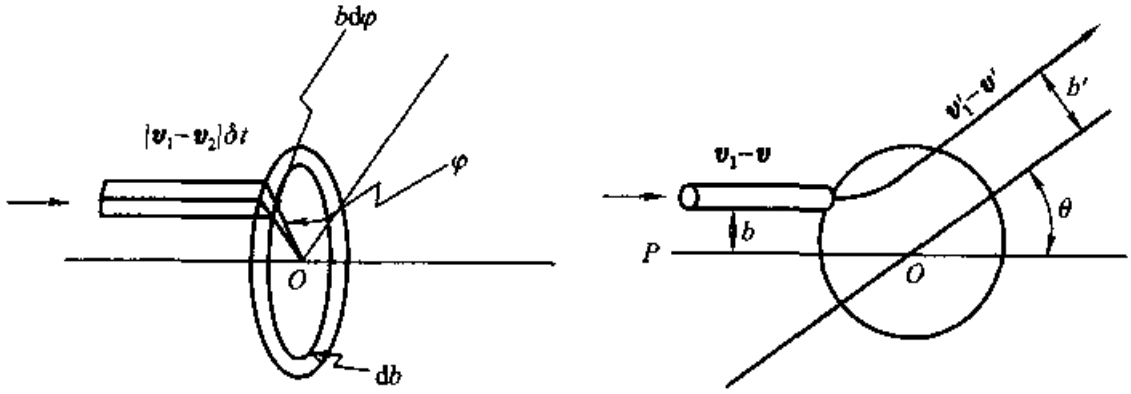


图 3　两分子的正碰撞

设想一个与 $\overline{PO}$ 轴垂直的平面[图 3(a)]. 在此平面上取极坐标 (b,φ)，b 为瞄准距离. 以速度 $\boldsymbol{v}_1$ 运动的，在 δt 时间间隔内能与靶分子发生正碰撞的入射分子，必位于以 $b\mathrm{d}\varphi\mathrm{d}b$ 为底，$|\boldsymbol{u}|\delta t=|\boldsymbol{v}_1-\boldsymbol{v}|\delta t$ 为高的柱体内，该柱体的体积为 $b\mathrm{d}\varphi\mathrm{d}b|\boldsymbol{v}_1-\boldsymbol{v}|\delta t$，该柱体内速度在 $\boldsymbol{v}_1\sim\boldsymbol{v}_1+\mathrm{d}\boldsymbol{v}_1$ 间的入射分子数即为在 δt 时间间隔内和指定靶分子发生的正碰撞数：

$$f(\boldsymbol{r},\boldsymbol{v}_1,t)\mathrm{d}^3\boldsymbol{v}_1 b\mathrm{d}\varphi\mathrm{d}b|\boldsymbol{v}_1-\boldsymbol{v}|\delta t,$$

由于在 $\boldsymbol{r}$ 处 $\mathrm{d}^3\boldsymbol{r}$ 内速度在 $\boldsymbol{v}\sim\boldsymbol{v}+\mathrm{d}\boldsymbol{v}$ 间的靶分子数为

$$f(\boldsymbol{r},\boldsymbol{v},t)\mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v},$$

于是，在时间间隔 δt 内，在 $\boldsymbol{r}$ 处 $\mathrm{d}^3\boldsymbol{r}$ 内，速度为 $\boldsymbol{v}_1\sim\boldsymbol{v}_1+\mathrm{d}\boldsymbol{v}_1$ 间的分子和速度在 $\boldsymbol{v}\sim\boldsymbol{v}+\mathrm{d}\boldsymbol{v}$ 间分子的正碰撞数为上两式的乘积：

$$f(\boldsymbol{r},\boldsymbol{v},t)f(\boldsymbol{r},\boldsymbol{v}_1,t)\mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v}b\mathrm{d}\varphi\mathrm{d}b|\boldsymbol{v}_1-\boldsymbol{v}|\delta t\mathrm{d}^3\boldsymbol{v}_1. \tag{3.11}$$

再考虑入射分子 $\boldsymbol{v}_1$ 的各种情形，对上式进行积分，就得到在 δt 内，由于正碰撞使 $(\boldsymbol{r},\boldsymbol{v})$ 处 $\mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v}$ 内减少的分子数

$$\left(\frac{\partial f}{\partial t}\right)_-\mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v}\delta t=\left[\iiint\mathrm{d}^3\boldsymbol{v}_1\int_0^{r_C}\mathrm{d}b\int_0^{2\pi}\mathrm{d}\varphi b|\boldsymbol{v}_1-\boldsymbol{v}|f(\boldsymbol{r},\boldsymbol{v},t)f(\boldsymbol{r},\boldsymbol{v}_1,t)\right]\mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v}\delta t, \tag{3.12}$$

再来考虑与正碰撞效果相反的逆碰撞：$|\boldsymbol{v}',\boldsymbol{v}_1'\to\boldsymbol{v},\boldsymbol{v}_1|$. 逆碰撞使 $\boldsymbol{r}$ 处 $\mathrm{d}^3\boldsymbol{r}$ 内速度在 $\boldsymbol{v}\sim\boldsymbol{v}+\mathrm{d}\boldsymbol{v}$ 间的分子数增加. 仿照前面的讨论，先把注意力集中于 $\boldsymbol{r}$ 处 $\mathrm{d}^3\boldsymbol{r}$ 内速度为 $\boldsymbol{v}'$ 的一个靶分子，考虑在 δt 内能与之发生碰撞的速度为 $\boldsymbol{v}_1'$ 的入射分子数，再考虑 $\boldsymbol{v}_1'$ 的各种可能情况后得

$$\left(\frac{\partial f}{\partial t}\right)_-\mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v}\delta t=\left[\iiint\mathrm{d}^3\boldsymbol{v}_1'\int_0^{r_C}\mathrm{d}b'\int_0^{2\pi}\mathrm{d}\varphi' b'|\boldsymbol{v}_1'-\boldsymbol{v}'|f(\boldsymbol{r},\boldsymbol{v}',t)f(\boldsymbol{r},\boldsymbol{v}_1',t)\right]\mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v}'\delta t.$$

利用(3.9)式和(3.10)式：$|\boldsymbol{v}_1-\boldsymbol{v}|=|\boldsymbol{u}|=|\boldsymbol{u}'|=|\boldsymbol{v}_1'-\boldsymbol{v}'|$ 和 $b=b'$，及在本节(三)中将要证明的

$$\mathrm{d}^3\boldsymbol{v}\mathrm{d}^3\boldsymbol{v}_1=\mathrm{d}^3\boldsymbol{v}'\mathrm{d}^3\boldsymbol{v}_1', \tag{3.13}$$

上式可化为

$$\left(\frac{\partial f}{\partial t}\right)_+\mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v}\delta t=\left[\iiint\mathrm{d}^3\boldsymbol{v}_1\int_0^{r_C}\mathrm{d}b\int_0^{2\pi}\mathrm{d}\varphi b|\boldsymbol{v}_1-\boldsymbol{v}|f(\boldsymbol{r},\boldsymbol{v}',t)f(\boldsymbol{r},\boldsymbol{v}_1',t)\right]\mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v}\delta t, \tag{3.14}$$

将(3.14)式和(3.12)式代入(3.4)得

$$\left(\frac{\partial f}{\partial t}\right)_{碰} \mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v}\delta t = \mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v}\delta t\iiint \mathrm{d}^3\boldsymbol{v}_1\int_0^{r_C}\mathrm{d}b\int_0^{2\pi}\mathrm{d}\varphi b\,|\,\boldsymbol{v}_1-\boldsymbol{v}\,|\,|f(\boldsymbol{v}')f(\boldsymbol{v}_1')-f(\boldsymbol{v})f(\boldsymbol{v}_1)|. \quad (3.15)$$

将上式代入§10.1.1(1.7)式,最后得

$$\frac{\partial f}{\partial t}+\boldsymbol{v}\cdot\nabla_{\boldsymbol{r}}f+\dot{\boldsymbol{v}}\cdot\nabla_{\boldsymbol{v}}f$$

$$=\iiint \mathrm{d}^3\boldsymbol{v}_1\int_0^{r_C}\mathrm{d}b\int_0^{2\pi}\mathrm{d}\varphi b\,|\,\boldsymbol{v}_1-\boldsymbol{v}\,|\,|f(\boldsymbol{v}')f(\boldsymbol{v}_1')-f(\boldsymbol{v})f(\boldsymbol{v}_1)|, \quad (3.16)$$

其中,$f(\boldsymbol{v})=f(\boldsymbol{r},\boldsymbol{v},t)$,$f(\boldsymbol{v}')=f(\boldsymbol{r},\boldsymbol{v}',t)$,$f(\boldsymbol{v}_1)=f(\boldsymbol{r},\boldsymbol{v}_1,t)$ 和 $f(\boldsymbol{v}_1')=f(\boldsymbol{r},\boldsymbol{v}_1',t)$.(3.16)式就是非平衡分布函数 $f(\boldsymbol{r},\boldsymbol{v},t)$ 所满足的玻尔兹曼积分微分方程.它是一个 f 的非线性积分微分方程,因此要严格求解是困难的,只能针对具体问题作各种近似处理.

早在§7.1.3中就曾经指出,动力学规律(牛顿力学、量子力学)对时间反演是对称的,即将 t 换为 $-t$,方程的形式不变❶.现试考查,当 t 换为 $-t$ 时,玻尔兹曼积分微分方程将发生怎样的变化?$t\to -t$ 时,$\boldsymbol{v}\to-\boldsymbol{v}$,(3.16)式左边改变正负号,而(3.16)式右边的碰撞积分不变.因此玻尔兹曼积分微分方程破坏了时间反演对称.

但是玻尔兹曼积分微分方程的时间发展不可逆性和力学规律的时间发展可逆性并不矛盾.原因是,在推导玻尔兹曼积分微分方程时,玻尔兹曼实际上引入了一个名为分子混沌性的假设,见(3.18)式,因而整个推导不是纯动力学推导.一般说来,在某一时刻 t,两个分子各处于 $\mathrm{d}^3\boldsymbol{r}_1\mathrm{d}^3\boldsymbol{v}_1$ 和 $\mathrm{d}^3\boldsymbol{r}_2\mathrm{d}^3\boldsymbol{v}_2$ 的概率由双粒子联合概率给出

$$\frac{1}{N^2}f(\boldsymbol{r}_1,\boldsymbol{v}_1,\boldsymbol{r}_2,\boldsymbol{v}_2,t)\,\mathrm{d}^3\boldsymbol{r}_1\mathrm{d}^3\boldsymbol{v}_1\mathrm{d}_2^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v}_2. \quad (3.17)$$

如果两个分子的概率统计独立,则上式可分解为单粒子概率的乘积

$$\frac{1}{N}f(\boldsymbol{r}_1,\boldsymbol{v}_1,t)\,\mathrm{d}^3\boldsymbol{r}_1\mathrm{d}^3\boldsymbol{v}_1\,\frac{1}{N}f(\boldsymbol{r}_2,\boldsymbol{v}_2,t)\,\mathrm{d}^3\boldsymbol{r}_2\mathrm{d}^3\boldsymbol{v}_2. \quad (3.18)$$

当两个分子相距足够远时,(3.18)式是可以理解的.但在推导(3.11)式时,两个分子显然在相互作用力程之内,分解就只能看作是一种假设了.

(三)两分子弹性碰撞前后 $\mathrm{d}^3\boldsymbol{v}\mathrm{d}^3\boldsymbol{v}_1=\mathrm{d}^3\boldsymbol{v}'\mathrm{d}^3\boldsymbol{v}_1'$ 的证明

现在补充对(3.13)式的证明.

两个分子组成的系统不受外力时,碰撞前后质心的速度应不变,即有

$$\begin{aligned}\boldsymbol{v}_C &= \boldsymbol{v}_C',\\ d\boldsymbol{v}_C &= \mathrm{d}\boldsymbol{v}_C'.\end{aligned} \quad (3.19)$$

再注意(3.9)式

$$|\,\boldsymbol{u}\,|=|\,\boldsymbol{u}'\,|,$$

则可得

$$\mathrm{d}^3\boldsymbol{v}_C\mathrm{d}^3\boldsymbol{u}=\mathrm{d}^3\boldsymbol{v}_C'\mathrm{d}^3\boldsymbol{u}'. \quad (3.20)$$

若能证得

---

❶ 对于量子力学,时间反演是 $t\to -t,\Psi\to\Psi^*$.

$$c\mathrm{d}^3\boldsymbol{v}_C\mathrm{d}^3\boldsymbol{u} = \mathrm{d}^3\boldsymbol{v}\mathrm{d}^3\boldsymbol{v}_1,$$
$$c\mathrm{d}^3\boldsymbol{v}_C'\mathrm{d}^3\boldsymbol{u}' = \mathrm{d}^3\boldsymbol{v}'\mathrm{d}^3\boldsymbol{v}_1'. \tag{3.21}$$

则最后可证得

$$\mathrm{d}^3\boldsymbol{v}\mathrm{d}^3\boldsymbol{v}_1 = \mathrm{d}^3\boldsymbol{v}'\mathrm{d}^3\boldsymbol{v}_1'.$$

下面来证明(3.21)式.

根据多重积分变量变换的法则附录(Ⅴ.21)式有

$$\mathrm{d}^3\boldsymbol{v}\mathrm{d}^3\boldsymbol{v}_1 = \frac{\mathrm{D}(v_x,v_y,v_z,v_{1x},v_{1y},v_{1z})}{\mathrm{D}(v_{cx},v_{cy},v_{cz},u_x,u_y,u_z)}\mathrm{d}^3\boldsymbol{v}_C\mathrm{d}^3\boldsymbol{u} = c\mathrm{d}^3\boldsymbol{v}_C\mathrm{d}^3\boldsymbol{u},$$

$$\mathrm{d}^3\boldsymbol{v}'\mathrm{d}^3\boldsymbol{v}_1' = \frac{\mathrm{D}(v_x',v_y',v_z',v_{1x}',v_{1y}',v_{1z}')}{\mathrm{D}(v_{cx}',v_{cy}',v_{cz}',u_x',u_y',u_z')}\mathrm{d}^3\boldsymbol{v}_C'\mathrm{d}^3\boldsymbol{u}' = c'\mathrm{d}^3\boldsymbol{v}_C'\mathrm{d}^3\boldsymbol{u}',$$

由(3.7)式和(3.8)式

$$\boldsymbol{v} = \boldsymbol{v}_C + \boldsymbol{u}/2, \qquad \boldsymbol{v}_1 = \boldsymbol{v}_C - \boldsymbol{u}/2,$$
$$\boldsymbol{v}' = \boldsymbol{v}_C' + \boldsymbol{u}'/2, \qquad \boldsymbol{v}_1' = \boldsymbol{v}_C' - \boldsymbol{u}'/2,.$$

不难得 $C = C'$,于是(3.21)式得证.

* §10.1.4　H 定理

克劳修斯和开尔文分别于1850和1851年提出了热力学第二定律的克氏表述和开氏表述. 1868年克劳修斯又提出熵的热力学定义和熵增加原理. 同一时期,气体分子运动论也已由克劳修斯和麦克斯韦发展到相当完美的程度,成功地处理了扩散、热传导和粘滞性等现象. 在这样的背景下,玻尔兹曼向自己提出了以下问题:"第二定律也能在纯粹力学原理的基础上建立起来吗?"这实际上是在对宏观过程的不可逆性和力学原理的可逆性如何协调进行挑战. 如上节所述,通过对气体分子运动的精细研究,玻尔兹曼导出了稀薄气体的非平衡分布函数的积分微分方程,进而于1872年得到了著名的,曾引起很大争议的 H 定理. H 定理和熵增加原理对应,描述了通过分子碰撞使系统从非平衡趋向平衡的过程.

(一) H 定理

定义 $H(t)$ 函数为(以下略去 $\mathrm{d}^3\boldsymbol{r},\mathrm{d}^3\boldsymbol{v}$ 的上标3)

$$H(t) = \iint f(\boldsymbol{r},\boldsymbol{v},t)\ln f(\boldsymbol{r},\boldsymbol{v},t)\,\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}. \tag{4.1}$$

$H(t)$ 随时间的变化为

$$\frac{\mathrm{d}H(t)}{\mathrm{d}t} = \iint(1+\ln f)\frac{\partial f}{\partial t}\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}. \tag{4.2}$$

将玻尔兹曼积分微分方程[上节的(3.16)式]代入,得

$$\frac{\mathrm{d}H(t)}{\mathrm{d}t} = -\iint(1+\ln f)\boldsymbol{v}\cdot\nabla_{\boldsymbol{r}}f\,\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v} - \iint(1+\ln f)\dot{\boldsymbol{v}}\cdot\nabla_{\boldsymbol{v}}f\,\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}$$
$$+\iiiint(1+\ln f)[f'f_1'-ff_1]b\,|\,\boldsymbol{v}_1-\boldsymbol{v}\,|\,\mathrm{d}b\mathrm{d}\varphi\mathrm{d}\boldsymbol{v}_1\mathrm{d}\boldsymbol{v}\mathrm{d}\boldsymbol{r}. \tag{4.3}$$

下面逐项分析(4.3)式右方各项.

第一项：

$$\iint[1+\ln f]\boldsymbol{v}\cdot\nabla_{\boldsymbol{r}}f\,\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}=\int\mathrm{d}\boldsymbol{v}\int\boldsymbol{v}\cdot\nabla_{\boldsymbol{r}}(f\ln f)\,\mathrm{d}\boldsymbol{r}$$

$$=\int\mathrm{d}\boldsymbol{v}\int\nabla_{\boldsymbol{r}}\cdot\boldsymbol{v}(f\ln f)\,\mathrm{d}\boldsymbol{r}=\int\mathrm{d}\boldsymbol{v}\oint\boldsymbol{e}_n\cdot\boldsymbol{v}(f\ln f)\,\mathrm{d}\boldsymbol{\sigma}_r=0.$$

其中已考虑到速度 $\boldsymbol{v}$ 和 $\boldsymbol{r}$ 是独立变量并利用了高斯定理[见附录(Ⅱ.24)]. $\boldsymbol{e}_n$ 为包围系统表面法线方向的单位矢量. 由于系统是封闭的，分子不能通过表面穿出系统，因而有 $\boldsymbol{e}_n\cdot\boldsymbol{v}=0$.

第二项：

$$\iint[1+\ln f]\dot{\boldsymbol{v}}\cdot\nabla_{\boldsymbol{v}}f\,\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}=\int\mathrm{d}\boldsymbol{r}\int\frac{\boldsymbol{F}}{m}\cdot\nabla_{\boldsymbol{v}}(f\ln f)\,\mathrm{d}\boldsymbol{v}$$

$$=\int\mathrm{d}\boldsymbol{r}\int\nabla_{\boldsymbol{v}}\cdot\frac{\boldsymbol{F}}{m}(f\ln f)\,\mathrm{d}\boldsymbol{v}=\int\mathrm{d}\boldsymbol{r}\oint\boldsymbol{e}_{n\boldsymbol{v}}\cdot\frac{\boldsymbol{F}}{m}(f\ln f)\,\mathrm{d}\boldsymbol{\sigma}_{\boldsymbol{v}}=0.$$

其中已考虑到一般情况下，外力与速度无关. 即使是和速度有关的洛伦兹力，仍然满足 $\partial F_x/\partial v_x=\partial F_y/\partial v_y=\partial F_z/\partial v_z=0$. $\boldsymbol{e}_{n\boldsymbol{v}}$ 是速度空间中表面法线方向的单位矢量. 速度空间的表面在无限远处，而当 $v\to\infty$ 时，$f\to0$，故积分值为零.

由于(4.3)式中右方第一项、第二项为零，于是

$$\frac{\mathrm{d}H(t)}{\mathrm{d}t}=\iiiint[1+\ln f][f'f'_1-ff_1]b\,|\,\boldsymbol{v}_1-\boldsymbol{v}\,|\,\mathrm{d}b\mathrm{d}\varphi\mathrm{d}\boldsymbol{v}_1\mathrm{d}\boldsymbol{v}\mathrm{d}\boldsymbol{r}. \tag{4.4}$$

注意其中：

$$f=f(\boldsymbol{r},\boldsymbol{v},t),\qquad f_1=f(\boldsymbol{r},\boldsymbol{v}_1,t),$$
$$f'=f(\boldsymbol{r},\boldsymbol{v}',t),\qquad f'_1=f(\boldsymbol{r},\boldsymbol{v}'_1,t).$$

由于积分的数值和积分变量无关，在(4.4)式中交换 $\boldsymbol{v}\leftrightarrow\boldsymbol{v}_1$，积分值不变，因而有

$$\frac{\mathrm{d}H(t)}{\mathrm{d}t}=\iiiint[1+\ln f_1][f'f'_1-f_1f]b\,|\,\boldsymbol{v}_1-\boldsymbol{v}\,|\,\mathrm{d}b\mathrm{d}\varphi\mathrm{d}\boldsymbol{v}_1\mathrm{d}\boldsymbol{v}\mathrm{d}\boldsymbol{r}. \tag{4.5}$$

将(4.4)式和(4.5)式相加除 2 得

$$\frac{\mathrm{d}H(t)}{\mathrm{d}t}=\frac{1}{2}\iiiint[2+\ln f+\ln f_1][f'f'_1-ff_1]b\,|\,\boldsymbol{v}_1-\boldsymbol{v}\,|\cdot\mathrm{d}b\mathrm{d}\varphi\mathrm{d}\boldsymbol{v}_1\mathrm{d}\boldsymbol{v}\mathrm{d}\boldsymbol{r}. \tag{4.6}$$

再将(4.6)式中的 $\boldsymbol{v}$ 和 $\boldsymbol{v}'$，$\boldsymbol{v}_1$ 和 $\boldsymbol{v}'_1$ 交换，即 $(\boldsymbol{v},\boldsymbol{v}_1)\leftrightarrow(\boldsymbol{v}',\boldsymbol{v}'_1)$ 得

$$\frac{\mathrm{d}H(t)}{\mathrm{d}t}=\frac{1}{2}\iiiint[2+\ln f'+\ln f'_1][ff_1-f'f'_1]b'\,|\,\boldsymbol{v}'_1-\boldsymbol{v}'\,|$$
$$\cdot\mathrm{d}b'\mathrm{d}\varphi'\mathrm{d}\boldsymbol{v}'_1\mathrm{d}\boldsymbol{v}'\mathrm{d}\boldsymbol{r}. \tag{4.7}$$

因为已知 $|\,\boldsymbol{v}_1-\boldsymbol{v}\,|=|\,\boldsymbol{u}\,|=|\,\boldsymbol{v}'_1-\boldsymbol{v}\,|=|\,\boldsymbol{u}'\,|$，$b=b'$，$\mathrm{d}\boldsymbol{v}_1\mathrm{d}\boldsymbol{v}=\mathrm{d}\boldsymbol{v}'_1\mathrm{d}\boldsymbol{v}'$，(4.7)式可改写为

$$\frac{\mathrm{d}H(t)}{\mathrm{d}t}=\frac{1}{2}\iiiint[2+\ln f'+\ln f'_1][ff_1-f'f'_1]b\,|\,\boldsymbol{v}_1-\boldsymbol{v}\,|$$
$$\cdot\mathrm{d}b\mathrm{d}\varphi\mathrm{d}\boldsymbol{v}_1\mathrm{d}\boldsymbol{v}\mathrm{d}\boldsymbol{r}. \tag{4.8}$$

将(4.6)式和(4.8)式相加除以 2 得

$$\frac{\mathrm{d}H(t)}{\mathrm{d}t}=-\frac{1}{4}\iiiint[\ln ff_1-\ln f'f'_1][ff_1-f'f'_1]b\,|\,\boldsymbol{v}_1-\boldsymbol{v}\,|$$
$$\cdot\mathrm{d}b\mathrm{d}\varphi\mathrm{d}\boldsymbol{v}_1\mathrm{d}\boldsymbol{v}\mathrm{d}\boldsymbol{r}. \tag{4.9}$$

其中 $b\,|\,\boldsymbol{v}_1-\boldsymbol{v}\,|$ 为正数. 若 $ff_1-f'f'_1>0$，则 $\ln ff_1-\ln f'f'_1>0$，

$$[\ln ff_1 - \ln f'f'_1][ff_1 - f'f'_1] > 0;$$

若 $ff_1 - f'f'_1 < 0$，则 $\ln ff_1 - \ln f'f'_1 < 0$，也有

$$[\ln ff_1 - \ln f'f'_1][ff_1 - f'f'_1] > 0.$$

因此，除 $ff_1 - f'f'_1 = 0$ 时，被积函数为零外，其他情况下，被积函数都是正数. 由此可知，(4.9)式右边是负数和零，即有：

$$dH(t)/dt \leqslant 0. \tag{4.10}$$

当且仅当积分区内处处

$$ff_1 = f'f'_1 \tag{4.11}$$

满足时，$dH(t)/dt = 0$. (4.11)式称为细致平衡条件.

(4.10)式称为H定理，它表示系统处于非平衡态时，H函数总是随时间单调减小. 当H达到最小值时，便不再随时间变化，系统达到平衡态. 由此可见，H定理和熵增加原理相当，指出了孤立系统中过程进行的方向和限度. 此外，和熵增加原理相比，H定理还给出了趋向平衡的不可逆过程的速率. 不过，在H定理的证明中，分布函数$f(\boldsymbol{r},\boldsymbol{v},t)$随时间的变化率由玻尔兹曼积分微分方程给出. 这就限制了H定理的适用范围. 它只适用于单原子稀薄气体，并且以分子混沌性假设为前提.

下面我们对H定理作进一步讨论.

(二)细致平衡原理

H定理表明，达到平衡态时，分布函数一定满足(4.11)式：$ff_1 = f'f'_1$. 由§10.1.3(3.11)式知道，$ff_1 b\,d\varphi\,db\,|\boldsymbol{v}_1 - \boldsymbol{v}|\,d\boldsymbol{v}d\boldsymbol{v}_1 d\boldsymbol{r}$ 表示在 $d\boldsymbol{r}$ 体积元内，单位时间中，一对速度为$(\boldsymbol{v},\boldsymbol{v}_1)$的分子经碰撞后速度变为$(\boldsymbol{v}',\boldsymbol{v}'_1)$的正碰撞数；$f'f'_1 b'\,d\varphi'\,db'|\boldsymbol{v}'_1 - \boldsymbol{v}'|\,d\boldsymbol{v}'d\boldsymbol{v}_1 d\boldsymbol{r}$ 表示在 $d\boldsymbol{r}$ 体积元内，单位时间中，一对速度为$(\boldsymbol{v}',\boldsymbol{v}'_1)$的分子经碰撞后速度变为$(\boldsymbol{v},\boldsymbol{v}_1)$的逆碰撞数. 因此(4.11)式表明，只有当在系统的任何一个小部分 $d\boldsymbol{r}$ 中，任何一个小速度间隔 $d\boldsymbol{v}$ 内，正碰撞数和逆碰撞数相等，即细致平衡时，系统才能达到总的平衡. 总的平衡只有通过细致平衡才能保证这一命题称为细致平衡原理. 这个原理在分子碰撞的问题上已由H定理证明. 不过这是在特殊的碰撞机制下证明的. 虽然这一原理在许多其他场合也正确，但不是对一切相互作用机制都适用，所以不是自然界的普遍法则.

(三)麦克斯韦速度分布

将(4.11)式取对数，并考虑到给定$\boldsymbol{r}$时，f只是速度的函数后有

$$\ln f(\boldsymbol{v}) + \ln f(\boldsymbol{v}_1) = \ln f(\boldsymbol{v}') + \ln f(\boldsymbol{v}'_1). \tag{4.12}$$

$\boldsymbol{v},\boldsymbol{v}_1$是两个分子碰撞前的速度，$\boldsymbol{v}',\boldsymbol{v}'_1$是两个分子碰撞后的速度. (4.12)式表明$\ln f(\boldsymbol{v})$既是速度的函数，又是碰撞的守恒量.

因为弹性碰撞时有粒子数守恒、动量守恒和能量守恒，因此(4.12)式有五个特解：

$$\ln f = 1,\quad mv_x,\quad mv_y,\quad mv_z,\quad mv^2/2. \tag{4.13}$$

(4.12)式是线性方程，它的通解是特解的线性组合，因此一般可写成如下形式：

$$\ln f(\boldsymbol{v}) = \alpha_0 + \alpha_1 v_x + \alpha_2 v_y + \alpha_3 v_z + \alpha_4 m(v_x^2 + v_y^2 + v_z^2). \tag{4.14}$$

其中$\alpha_0,\alpha_1,\alpha_2,\alpha_3,\alpha_4$为五个和速度无关的常数. 为了使引进的常数有较明确的物理意义，将(4.14)式分别对v_x,v_y,v_z配成平方项，成为

$$\ln f(\boldsymbol{v}) = -\beta[(v_x - v_{x0})^2 + (v_y - v_{y0})^2 + (v_z - v_{z0})^2] + \alpha. \tag{4.15}$$

并引入了五个新的常数 $\beta, v_{x0}, v_{y0}, v_{z0}, \alpha$ 不难看出新常数与原来常数的关系为：

$$-\beta = \alpha_4, \quad 2\beta v_{x0} = \alpha_1, \quad 2\beta v_{y0} = \alpha_2, \quad 2\beta v_{z0} = \alpha_3,$$

$$\alpha - \beta(v_{x0}^2 + v_{y0}^2 + v_{z0}^2) = \alpha_0.$$

(4.15)式可改写为

$$f(\boldsymbol{v}) = A\exp\{-\beta[(v_x - v_{x0})^2 + (v_y - v_{y0})^2 + (v_z - v_{z0})^2]\}. \tag{4.16}$$

其中 $A = e^{\alpha}$.

从 $n = \int f(\boldsymbol{v})\mathrm{d}\boldsymbol{v}$，可得

$$n = A(\pi/\beta)^{3/2}, \quad \text{或} \quad A = n(\beta/\pi)^{3/2}. \tag{4.17}$$

又因 $\langle(v_x - v_{x0})\rangle = \frac{1}{n}\int(v_x - v_{x0})f(\boldsymbol{v})\mathrm{d}\boldsymbol{v} = 0$. 所以，$\langle v_x\rangle = v_{x0}$. 同理有 $\langle v_y\rangle = v_{y0}, \langle v_z\rangle = v_{z0}$，即 v_{x0}, v_{y0}, v_{z0} 为分子速度分量的平均值，也就是气体的宏观流速. 平衡时，v_{x0}, v_{y0}, v_{z0} 和空间坐标无关，因而可以选择坐标使 $v_{x0} = v_{y0} = v_{z0} = 0$，从而(4.16)式化为

$$f(\boldsymbol{v}) = n(\beta/\pi)^{3/2}e^{-\beta v^2}. \tag{4.18}$$

将能量平均值与能量均分定理比较可确定常数 β 值：

$$\frac{1}{2}m\langle v^2\rangle = \frac{1}{n}\int\frac{1}{2}mv^2 f(\boldsymbol{v})\mathrm{d}\boldsymbol{v} = \frac{3}{2}k_B T, \quad \beta = m/(2k_B T).$$

最后得平衡态的分布函数为

$$f(\boldsymbol{v}) = n[m/(2\pi k_B T)]^{3/2}e^{-mv^2/(2k_B T)}. \tag{4.19}$$

这就是麦克斯韦速度分布. 由于玻尔兹曼积分微分方程只适用于单原子分子稀薄气体，因而这样得出的麦克斯韦速度分布不如从正则系综得出的适用广泛. 但从细致平衡，可以更具体地看到，当分子的速度分布函数服从麦克斯韦速度分布时，分子的每一种正碰撞与逆碰撞互相抵消，分子间的碰撞已不会改变分子的速度分布了.

（四）H 函数与熵函数的关系

在 §7.3.1 中已得到熵的普遍公式[见该节(1.19)式，又见 §7.2.2(2.4)式]

$$S = -k_B\sum_r P_r\ln P_r.$$

考虑到对于单原子理想气体，$\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}$ 中的单粒子态数为 $(m^3/h^3)\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}$，则 $\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}$ 中的平均分子数可表为

$$f\,\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v} = NP_r(m^3/h^3)\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}, \quad \text{或} \quad P_r = fh^3/(m^3N),$$

$$s = -k_B\sum\left(\frac{fh^3}{m^3N}\right)\ln\left(\frac{fh^3}{m^3N}\right) = -k_B\iint\left(\frac{fh^3}{m^3N}\right)\ln\left(\frac{fh^3}{m^3N}\right)\frac{m^3}{h^3}\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}$$

$$= -\frac{k_B}{N}\iint f\ln f\,\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v} + k_B\ln\left(\frac{h^3}{m^3N}\right).$$

注意，$H(t) = \int f\ln f\,\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}$ 和系统的 $S = N\cdot s$，得到熵 S 和 H 函数的关系

$$S = N\cdot s = -k_B H(t) + \text{常数}. \tag{4.20}$$

（五）H 定理的统计性质

最后，我们对 H 定理的统计性质进行简单讨论. 首先要说明的是，H 定理是一个统计性的定理. 这是因为：第一，H 函数是一个统计性的函数. 按定义有 $H = \langle \ln f \rangle$，因此 H 是一个统计平均值. 不仅如此，f 本身就是统计性的. 其次，我们是根据分子碰撞的统计性质（混沌性假设）证明 H 定理的，所以 H 随时间单调减小也是统计性的.

其次，我们讨论微观可逆性与宏观不可逆性的问题. 在玻尔兹曼提出 H 定理后，洛施密特在 1876 年提出了质疑. 洛喜密脱指出，根据经典力学定律，在保守力作用下运动是可逆的. 假设一个系统的 H 在减小，如果在某一时刻将所有分子的速度都反过来，由力学规律的可逆性可知，系统将以倒逆的次序经历原先经历过，但速度反向的状态. 由于所有分子的速度反向后系统的 H 不变，这个系统的 H 将增加. 这就跟 H 定理矛盾.

玻尔兹曼对这个问题的解释是，H 定理不是一个力学规律，而是一个统计规律，它指出 H 减少的概率最大. 宏观不可逆性是统计规律的结果，与微观可逆性并不矛盾. H 定理并不排斥 H 的偶然增加，但增加的概率非常小.

习　题

1. 试用弛豫时间近似推导气体的自扩散系数.

2. 试用弛豫时间近似求气体导热系数的微观表达式.

3. 对一半导体中的非简并电子气，出现在玻尔兹曼方程中的 $f^{(0)}$ 可取 M-B 分布，试计算非简并电子气的电导率，假设 $\varepsilon(p) = p^2/(2m^*)$.

4. 镀在绝缘体表面上的银膜可看成二维导体，当外加一稳定弱电场时将产生电流. 试用弛豫时间近似求此薄膜的电导率 σ（可用费米弛豫时间 τ_F 和自由电子面密度 n 表示）.

5. 被吸附的气体分子在表面上作二维运动，试写出二维气体的玻尔兹曼积分微分方程.

6. 若稀薄气体由两种分子组成，可令第一种分子的非平衡分布函数为 $f(\boldsymbol{r},\boldsymbol{v},t)$，第二种分子的非平衡分布函数为 $F(\boldsymbol{r},\boldsymbol{v},t)$，试将 §10.1.3(3.16)式推广，分别写出关于 $f(\boldsymbol{r},\boldsymbol{v},t)$ 和 $F(\boldsymbol{r},\boldsymbol{v},t)$ 的玻尔兹曼积分微分方程.

7. 已知两种分子的混合理想气体的分布函数分别为

$$f(v_1) = n_1[m_1/(2\pi k_B T)]^{3/2} e^{m_1 v_1^2/(2k_B T)}, \qquad f(v_2) = n_2[m_2/(2\pi k_B T)]^{3/2} e^{m_2 v_2^2/(2k_B T)}$$

式中 n_1, n_2, T 都是 $\boldsymbol{r}$ 和 t 的函数. 试证明玻尔兹曼积分微分方程中的碰撞项为零.

8. 试用细致平衡原理导出费米分布.

9. 试用细致平衡原理导出玻色分布.

10. 可以证明，在有外力场情况下，H 定理与无外力场相同，系统达到平衡时，要求细致平衡条件. 而且玻尔兹曼方程的解仍为麦克斯韦分布：$f = Ae^{-m(v_x^2+v_y^2+v_z^2)/(2k_B T)}$ 只不过 A 是分子坐标的函数. 试证明：$A = A_0 e^{-\phi(r)/(k_B T)}$，$\phi(r)$ 是分子的势能，A_0 是常数.

* 第二章　BBGKY 序列

§10.2.1　约化概率密度　BBGKY 序列

在§7.1.5 中已推得了刘维尔定理，并指出刘维尔定理不仅是平衡统计也是非平衡统计的基础. 为方便起见重写如下：

$$\partial\rho_N/\partial t + [H_N,\rho_N]_{PB} = 0. \tag{1.1}$$

在本章里，我们用花括号表示泊松括号[见§1.6.2(2.10)式]：

$$\{A,B\} \equiv [A,B]_{PB} = \sum_{i=1}^{N}\left[(\nabla_{p_i}A)\cdot(\nabla_{q_i}B) - (\nabla_{q_i}A)\cdot(\nabla_{p_i}B)\right]. \tag{1.2}$$

为了更加明确，已将(1.1)式中 N 个粒子的联合概率密度和哈密顿分别表为 $\rho_N = \rho(\boldsymbol{q}_1,\boldsymbol{p}_1,\cdots,\boldsymbol{q}_N,\boldsymbol{p}_N;t)$ 和 H_N. 刘维尔定理包含了统计的全部微观信息，涉及 $(6N+1)$ 个变量. 要从涉及到 $(6N+1)$ 个变量的微分方程中精确得出只涉及几个变量的宏观方程，必须通过引入适当的假设或近似，把包含全部 N 个粒子坐标和动量的概率密度方程化为只含少数粒子坐标和动量的概率密度方程. BBGKY 序列就是这样一种方法，它是分别由 Bogolyubov，Born 和 Green，Kirkwood 和 Yvon 独立引入的.

为了方便，文献中往往采用如下形式的 N 个粒子联合概率密度：

$$V^{-N}\rho_N(\boldsymbol{q}_1,\boldsymbol{p}_1,\cdots,\boldsymbol{q}_N,\boldsymbol{p}_N;t). \tag{1.3}$$

其归一化条件是

$$\int\rho_N(\boldsymbol{q}_1,\boldsymbol{p}_1,\cdots,\boldsymbol{q}_N,\boldsymbol{p}_N;t)\,d\boldsymbol{q}_1 d\boldsymbol{p}_1\cdots d\boldsymbol{q}_N d\boldsymbol{p}_N = V^N. \tag{1.4}$$

V 是系统的体积. 根据概率理论，将(1.3)式对 $6(N-l)$ 个变量 $d\boldsymbol{q}_{l+1}d\boldsymbol{p}_{l+1}\cdots d\boldsymbol{q}_N d\boldsymbol{p}_N$ 积分可得到 l 个粒子的联合概率密度：

$$\rho_l(\boldsymbol{q}_1,\boldsymbol{p}_1,\cdots,\boldsymbol{q}_i,\boldsymbol{p}_i,\cdots,\boldsymbol{q}_j,\boldsymbol{p}_j,\cdots,\boldsymbol{q}_l,\boldsymbol{p}_l;t)$$
$$= V^{(l-N)}\int\rho_N(\boldsymbol{q}_1,\boldsymbol{p}_1,\cdots,\boldsymbol{q}_N,\boldsymbol{p}_N;t)\,d\boldsymbol{q}_{l+1}d\boldsymbol{p}_{l+1}\cdots d\boldsymbol{q}_N d\boldsymbol{p}_N. \tag{1.5}$$

$l = 1,2,\cdots,N$. ρ_l 叫做约化概率密度. 由于粒子是全同的，ρ_l 对 $(\boldsymbol{q}_i,\boldsymbol{p}_i)$ 和 $(\boldsymbol{q}_j,\boldsymbol{p}_j)$ 的交换保持不变.

假设系统中粒子间的相互作用以两体势 $U_{ij} = U(\boldsymbol{q}_i,\boldsymbol{q}_j)$ 的形式出现，且系统还受到外力势 $\phi(\boldsymbol{q}_i)$ 的作用，则系统的哈密顿可表为

$$H_N = \sum_{i=1}^{N}\left(\frac{p_i^2}{2m} + \phi(\boldsymbol{q}_i)\right) + \sum_{1\le i<j\le N}U(\boldsymbol{q}_i,\boldsymbol{q}_j). \tag{1.6}$$

将(1.5)式对 t 求偏微商，并利用刘维尔定理(1.1)可得

$$\frac{\partial\rho_l}{\partial t} = V^{(l-N)}\frac{\partial}{\partial t}\int\rho_N(\boldsymbol{q}_1,\boldsymbol{p}_1,\cdots,\boldsymbol{q}_N,\boldsymbol{p}_N;t)\,d\boldsymbol{q}_{l+1}d\boldsymbol{p}_{l+1}\cdots d\boldsymbol{q}_N d\boldsymbol{p}_N$$

$$= -V^{(l-N)}\int |H_N,\rho_N|\,\mathrm{d}\boldsymbol{q}_{l+1}\mathrm{d}\boldsymbol{p}_{l+1}\cdots\mathrm{d}\boldsymbol{q}_N\mathrm{d}\boldsymbol{p}_N$$

$$= -V^{(l-N)}\int\Big\{\sum_{i=1}^{N}\Big[\frac{p_i^2}{2m}+\phi(\boldsymbol{q}_i)\Big],\rho_N\Big\}\mathrm{d}\boldsymbol{q}_{l+1}\mathrm{d}\boldsymbol{p}_{l+1}\cdots\mathrm{d}\boldsymbol{q}_N\mathrm{d}\boldsymbol{p}_N$$

$$-V^{(l-N)}\int\Big\{\sum_{1\leqslant i<j\leqslant N}U(\boldsymbol{q}_i,\boldsymbol{q}_j),\rho_N\Big\}\mathrm{d}\boldsymbol{q}_{l+1}\mathrm{d}\boldsymbol{p}_{l+1}\cdots\mathrm{d}\boldsymbol{q}_N\mathrm{d}\boldsymbol{p}_N. \tag{1.7}$$

为使上式简化，首先考虑右边第一个积分中 $i>l$ 的某一项

$$\int\Big\{\frac{p_i^2}{2m}+\phi(\boldsymbol{q}_i),\rho_N\Big\}\mathrm{d}\boldsymbol{q}_i\mathrm{d}\boldsymbol{p}_i,$$

利用泊松括号的定义(1.2)式，有

$$-\int\Big\{\frac{p_i^2}{2m}+\phi(\boldsymbol{q}_i),\rho_N\Big\}\mathrm{d}\boldsymbol{q}_i\mathrm{d}\boldsymbol{p}_i$$

$$=\int\Big[(\nabla_{\boldsymbol{q}_i}\phi(\boldsymbol{q}_i))\cdot(\nabla_{\boldsymbol{p}_i}\rho_N)-\frac{\boldsymbol{p}_i}{m}(\nabla_{\boldsymbol{q}_i}\rho_N)\Big]\mathrm{d}\boldsymbol{q}_i\mathrm{d}\boldsymbol{p}_i$$

$$=\int\nabla_{\boldsymbol{q}_i}\phi(\boldsymbol{q}_i)\rho_N\Big|_{p_i=-\infty}^{p_i=\infty}\mathrm{d}\boldsymbol{q}_i-\int\frac{\boldsymbol{p}_i}{m}\rho_N\Big|_{q_i=-\infty}^{q_i=\infty}\mathrm{d}\boldsymbol{p}_i=0. \tag{1.8}$$

最后一个等式是由于，为保证 ρ_N 可归一化，在 Γ 空间的无穷远处 ρ_N 必须为零. 类似地，对于(1.7)式第二个积分中 $i>l$ 的某一项有

$$\int|U(\boldsymbol{q}_i,\boldsymbol{q}_j),\rho_N|\,\mathrm{d}\boldsymbol{q}_i\mathrm{d}\boldsymbol{p}_i\mathrm{d}\boldsymbol{q}_j\mathrm{d}\boldsymbol{p}_j=0. \tag{1.9}$$

利用(1.8)和(1.9)式，可将(1.7)式化为

$$\frac{\partial\rho_l}{\partial t}=-V^{(l-N)}\int\Big\{\sum_{i=1}^{l}\Big[\frac{p_i^2}{2m}+\phi(\boldsymbol{q}_i)\Big],\rho_N\Big\}\mathrm{d}\boldsymbol{q}_{l+1}\mathrm{d}\boldsymbol{p}_{l+1}\cdots\mathrm{d}\boldsymbol{q}_N\mathrm{d}\boldsymbol{p}_N$$

$$-V^{(l-N)}\int\Big\{\sum_{1\leqslant i<j\leqslant l}U(\boldsymbol{q}_i,\boldsymbol{q}_j),\rho_N\Big\}\mathrm{d}\boldsymbol{q}_{l+1}\mathrm{d}\boldsymbol{p}_{l+1}\cdots\mathrm{d}\boldsymbol{q}_N\mathrm{d}\boldsymbol{p}_N$$

$$+V^{(l-N)}\int\sum_{i=1}^{l}\sum_{j=l+1}^{N}(\nabla_{\boldsymbol{q}_i}U_{ij})\cdot(\nabla_{\boldsymbol{p}_i}\rho_N)\mathrm{d}\boldsymbol{q}_{l+1}\mathrm{d}\boldsymbol{p}_{l+1}\cdots\mathrm{d}\boldsymbol{q}_N\mathrm{d}\boldsymbol{p}_N$$

$$=-V^{(l-N)}\int\Big\{\sum_{i=1}^{l}\Big[\frac{p_i^2}{2m}+\phi(\boldsymbol{q}_i)\Big]+\sum_{1\leqslant i<j\leqslant l}U_{ij},\rho_N\Big\}\mathrm{d}\boldsymbol{q}_{l+1}\mathrm{d}\boldsymbol{p}_{l+1}\cdots\mathrm{d}\boldsymbol{q}_N\mathrm{d}\boldsymbol{p}_N$$

$$+V^{(l-N)}(N-l)\int\sum_{i=1}^{l}(\nabla_{\boldsymbol{q}_i}U_{i,l+1})\cdot(\nabla_{\boldsymbol{p}_i}\rho_N)\mathrm{d}\boldsymbol{q}_{l+1}\mathrm{d}\boldsymbol{p}_{l+1}\cdots\mathrm{d}\boldsymbol{q}_N\mathrm{d}\boldsymbol{p}_N.$$

引入 l 粒子的哈密顿

$$H_l=\sum_{i=1}^{l}\Big[\frac{p_i^2}{2m}+\phi(\boldsymbol{q}_i)\Big]+\sum_{1\leqslant i<j\leqslant l}U_{ij}, \tag{1.10}$$

并注意约化概率密度的定义式(1.5)式，最后得

$$\frac{\partial\rho_l}{\partial t}=-\{H_l,\rho_l\}+\frac{(N-l)}{V}\int\sum_{i=1}^{l}(\nabla_{\boldsymbol{q}_i}U_{i,l+1})\cdot(\nabla_{\boldsymbol{p}_i}\rho_{l+1})\mathrm{d}\boldsymbol{q}_{l+1}\mathrm{d}\boldsymbol{p}_{l+1}. \tag{1.11}$$

耦合方程组(1.11)式就是 BBGKY 序列，它给出了约化概率密度 ρ_l 演化方程的一个序列，并精确地等价于刘维尔定理，包含了系统的全部微观信息. 从 BBGKY 序列(1.11)式不难看出，为了

计算单粒子约化概率密度 ρ_1，必须知道双粒子约化概率密度 ρ_2；为计算 ρ_2，又必须知道三粒子约化概率密度 ρ_3 等等，即(1.11)式中的任何一个方程都不封闭. 因此，必须通过引入适当的假设或近似，将 BBGKY 序列截断，得到感兴趣的几个最低阶约化概率密度的封闭耦合方程组.

§10.2.2 符拉索夫方程和玻尔兹曼方程的推导

在实际问题(如输运系数的计算)中，人们最感兴趣的是单粒子约化概率密度 $\rho_1(\boldsymbol{r}_1,\boldsymbol{p}_1,t)$ 和双粒子约化概率密度 $\rho_2(\boldsymbol{r}_1,\boldsymbol{p}_1,\boldsymbol{r}_2,\boldsymbol{p}_2,t)$. 为求得 ρ_1 和 ρ_2，我们讨论 BBGKY 序列中的第一、第二个方程. 由 §10.2.1(1.11)式可得

$$\partial\rho_1/\partial t = -\{H_1,\rho_1\} + (\partial\rho_1/\partial t)_{碰}. \tag{2.1}$$

其中

$$\begin{aligned}\{H_1,\rho_1\} &= \{p_1^2/(2m) + \phi(\boldsymbol{r}_1),\rho_1\}\\ &= -\nabla_{\boldsymbol{r}_1}\phi(\boldsymbol{r}_1)\cdot(\nabla_{\boldsymbol{p}_1}\rho_1) + (\boldsymbol{p}_1/m)\cdot\nabla_{\boldsymbol{r}_1}\rho_1\\ &= [-\nabla_{\boldsymbol{r}_1}\phi(\boldsymbol{r}_1)\cdot\nabla_{\boldsymbol{p}_1} + (\boldsymbol{p}_1/m)\cdot\nabla_{\boldsymbol{r}_1}]\rho_1,\end{aligned} \tag{2.2}$$

$$\begin{aligned}\left(\frac{\partial\rho_1}{\partial t}\right)_{碰} &= \frac{N-1}{V}\int[\nabla_{\boldsymbol{r}_1}U_{1,2}(\boldsymbol{r}_1,\boldsymbol{r}_2)]\cdot(\nabla_{\boldsymbol{p}_1}\rho_2)\mathrm{d}\boldsymbol{r}_2\mathrm{d}\boldsymbol{p}_2\\ &\cong \frac{N}{V}\int[\nabla_{\boldsymbol{r}_1}U_{1,2}(\boldsymbol{r}_1,\boldsymbol{r}_2)]\cdot(\nabla_{\boldsymbol{p}_1}\rho_2)\mathrm{d}\boldsymbol{r}_2\mathrm{d}\boldsymbol{p}_2.\end{aligned} \tag{2.3}$$

在文献中通常称 $(\partial\rho_1/\partial t)_{碰}$ 为碰撞积分.

BBGKY 序列中的第二个方程是

$$\begin{aligned}\frac{\partial\rho_2}{\partial t} = -\{H_2,\rho_2\} + \frac{N-2}{V}\int[&(\nabla_{\boldsymbol{r}_1}U_{1,3})\cdot(\nabla_{\boldsymbol{p}_1}\rho_3)\\ &+(\nabla_{\boldsymbol{r}_2}U_{2,3})\cdot(\nabla_{\boldsymbol{p}_2}\rho_3)]\mathrm{d}\boldsymbol{r}_3\mathrm{d}\boldsymbol{p}_3,\end{aligned} \tag{2.4}$$

其中

$$\begin{aligned}\{H_2,\rho_2\} &= \left\{\frac{p_1^2}{2m} + \frac{p_2^2}{2m} + \phi(\boldsymbol{r}_1) + \phi(\boldsymbol{r}_2) + U_{1,2}(\boldsymbol{r}_1,\boldsymbol{r}_2),\rho_2\right\}\\ &= -\nabla_{\boldsymbol{r}_1}[\phi(\boldsymbol{r}_1) + U_{1,2}(\boldsymbol{r}_1,\boldsymbol{r}_2)]\cdot\nabla_{\boldsymbol{p}_1}\rho_2 + \frac{\boldsymbol{p}_1}{m}\cdot\nabla_{\boldsymbol{r}_1}\rho_2\\ &\quad -\nabla_{\boldsymbol{r}_2}[\phi(\boldsymbol{r}_2) + U_{1,2}(\boldsymbol{r}_1,\boldsymbol{r}_2)]\cdot\nabla_{\boldsymbol{p}_2}\rho_2 + \frac{\boldsymbol{p}_2}{m}\cdot\nabla_{\boldsymbol{r}_2}\rho_2\\ &= -\left[\nabla_{\boldsymbol{r}_1}(\phi(\boldsymbol{r}_1) + U_{1,2})\cdot\nabla_{\boldsymbol{p}_1} - \frac{\boldsymbol{p}_1}{m}\cdot\nabla_{\boldsymbol{r}_1}\right.\\ &\quad \left.+\nabla_{\boldsymbol{r}_2}(\phi(\boldsymbol{r}_2) + U_{1,2})\cdot\nabla_{\boldsymbol{p}_2} - \frac{\boldsymbol{p}_2}{m}\cdot\nabla_{\boldsymbol{r}_2}\right]\rho_2.\end{aligned} \tag{2.5}$$

由于方程(2.1)的右边涉及 ρ_2，方程(2.4)的右边又涉及 ρ_3，因而方程(2.1)和(2.4)都不封闭. 若能在方程(2.1)中把 ρ_2 用 ρ_1 表示出来，则方程(1)得以封闭. 下面我们用推导符拉索夫方程说明这一点. 若能在方程(2.4)中去掉 ρ_3，则方程(2..1)和(2.4)构成一个关于 ρ_1、ρ_2 的封

闭方程组.我们通过推导玻尔兹曼积分微分方程演示这一过程.

(一)符拉索夫方程的推导

我们的目标是找到一个恰当的近似将方程(2.1)中的碰撞积分化为仅含单粒子约化概率密度 ρ_1.换句话说,使之与高阶约化概率密度脱耦.为此,我们假设,在同一时刻,在 $(\boldsymbol{r},\boldsymbol{p})$ 和 $(\boldsymbol{r}_1,\boldsymbol{p}_1)$ 各找到一个粒子的联合概率密度 $\rho_2(\boldsymbol{r},\boldsymbol{p},\boldsymbol{r}_1,\boldsymbol{p}_1)$ 就等于在 $(\boldsymbol{r},\boldsymbol{p})$ 找到一个粒子的概率和在 $(\boldsymbol{r}_1,\boldsymbol{p}_1)$ 找到一个粒子的概率的乘积,即

$$\rho_2(\boldsymbol{r},\boldsymbol{p},\boldsymbol{r}_1,\boldsymbol{p}_1,t) = \rho_1(\boldsymbol{r},\boldsymbol{p},t)\rho_1(\boldsymbol{r}_1,\boldsymbol{p}_1,t). \tag{2.6}$$

于是(2.1)式中的碰撞积分可写为

$$\left(\frac{\partial\rho_1}{\partial t}\right)_{碰} = \frac{N}{V}\nabla_{\boldsymbol{r}}\int U(\boldsymbol{r},\boldsymbol{r}_1)\rho_1(\boldsymbol{r}_1,\boldsymbol{p}_1,t)\,\mathrm{d}\boldsymbol{r}_1\mathrm{d}\boldsymbol{p}_1\cdot\nabla_{\boldsymbol{p}}\rho_1(\boldsymbol{r},\boldsymbol{p},t).$$

引入平均场势

$$\tilde{\Phi}(\boldsymbol{r},t) = \frac{N}{V}\int U(\boldsymbol{r},\boldsymbol{r}_1)\rho_1(\boldsymbol{r}_1,\boldsymbol{p}_1,t)\,\mathrm{d}\boldsymbol{r}_1\mathrm{d}\boldsymbol{p}_1.$$

则碰撞积分化为

$$(\partial\rho_1/\partial t)_{碰} = [\nabla_{\boldsymbol{r}}\tilde{\Phi}(\boldsymbol{r},t)]\cdot[\nabla_{\boldsymbol{p}}\rho_1(\boldsymbol{r},\boldsymbol{p},t)]. \tag{2.7}$$

将(2.7)式代入(2.1)式得到关于 ρ_1 的封闭方程

$$(\partial\rho_1/\partial t) = (-(\boldsymbol{p}/m)\cdot\nabla_{\boldsymbol{r}} + [\nabla_{\boldsymbol{r}}\tilde{\Phi}(\boldsymbol{r}) + \nabla_{\boldsymbol{r}}\phi(\boldsymbol{r})]\cdot\nabla_{\boldsymbol{p}})\rho_1. \tag{2.8}$$

这就是符拉索夫方程.它表示每个粒子都在独立运动,但每个粒子除受外力势 $\phi(\boldsymbol{r})$ 作用外,还受到由其他粒子产生的"平均场势" $\tilde{\Phi}(\boldsymbol{r})$ 的作用.显然符拉索夫方程是关于单粒子约化概率密度 ρ_1 的线性方程.

符拉索夫方程在任意两个粒子间都存在长程库仑力的稀薄等离子体中是很重要的方程.

(二)玻尔兹曼积分微分方程的推导

考虑由通过短程范德瓦耳斯力相互作用的中性分子组成的稀薄气体.

在稀薄中性气体中,当两个粒子处于短程相互作用力程范围 r_0(通常称为相遇)的短暂时间 τ_0 内,第三个粒子也处于这个范围的概率是很小的.换句话说,三分子以上碰撞可以忽略不计.譬如在氮气中,分子相互作用范围大约是 $r_0 = 3$ Å,室温下分子的平均速度大约是 $v = 10^5\ \mathrm{cm\cdot s^{-1}}$.两分子相遇的时间大约是 $\tau_0 = r_0/v \approx 2.2\times10^{-13}$ s.在常压下两次碰撞之间的时间大约为 $\tau_c \approx 2\times10^{-10}$ s(碰撞频率的倒数).于是三分子碰撞在全部碰撞中所占比例大约是 $\tau_0/\tau_c \approx 10^{-3}$.在无化学反应的气体中,忽略三分子碰撞是合理的.

忽略了三分子碰撞意味着在 $|\boldsymbol{r}_i-\boldsymbol{r}_j|\leqslant r_0$ 时有 $\rho_3 = 0$,而如果 $|\boldsymbol{r}_i-\boldsymbol{r}_j| > r_0$ 又有 $U_{ij}\approx 0$(短程力).于是 BBGKY 序列中第二个方程即方程(2.4)中右方的积分可以去掉.方程(2.1)和(2.4)构成仅含约化概率密度 ρ_1 和 ρ_2 的封闭方程组.原则上可以求解.为了书写方便,下面把 $(\boldsymbol{r}_1,\boldsymbol{p}_1)$ 记为 $(\boldsymbol{r},\boldsymbol{p})$,把 $(\boldsymbol{r}_2,\boldsymbol{p}_2)$ 记为 $(\boldsymbol{r}_1,\boldsymbol{p}_1)$,相互作用势仅取决于粒子间距离的绝对值,故记 $U_{12}(\boldsymbol{r},\boldsymbol{r}_1)$ 为 $U(|\boldsymbol{r}_1-\boldsymbol{r}|)$.于是(2.1)式和(2.4)式可写为

$$\frac{\partial\rho_1}{\partial t} = \left[\nabla_{\boldsymbol{r}}\phi(\boldsymbol{r})\cdot\nabla_{\boldsymbol{p}} - \frac{\boldsymbol{p}}{m}\cdot\nabla_{\boldsymbol{r}}\right]\rho_1(\boldsymbol{r},\boldsymbol{p},t) + \frac{N}{V}\int\nabla_{\boldsymbol{r}}U(|\boldsymbol{r}_1-\boldsymbol{r}|)\cdot\nabla_{\boldsymbol{p}}\rho_2(\boldsymbol{r},\boldsymbol{p},\boldsymbol{r}_1,\boldsymbol{p}_1,t)\,\mathrm{d}\boldsymbol{r}_1\mathrm{d}\boldsymbol{p}_1. \tag{2.9}$$

$$\frac{\partial \rho_2}{\partial t} = \Big[\nabla_r [\phi(r) + U(|r_1 - r|)] \cdot \nabla_p - \frac{p}{m} \cdot \nabla_r +$$

$$\nabla_{r_1} [\phi(r_1) + U(|r_1 - r|)] \cdot \nabla_{p_1} - \frac{p_1}{m} \cdot \nabla_{r_1} \Big] \rho_2(r, p, r_1, p_1, t). \tag{2.10}$$

(2.9)式和(2.10)式构成封闭方程组,两式中的 ρ_2 必须是一致的.

(2.10)式还可进一步化简.第一,已假设分子相互作用力是短程力,即仅当 $|r_1 - r| \leqslant r_0$ 时才有 $U(|r_1 - r|) \neq 0$.因此(2.9)式碰撞积分中 ρ_2 的变量必须满足 $|r_1 - r| \leqslant r_0$.而在这样一个距离内,和分子相互作用势比较,外场势可看成常数,因此(2.10)式中 $\nabla_r \phi(r)$ 和 $\nabla_{r_1} \phi(r_1)$ 可以忽略.第二,在碰撞积分中,我们只在两分子相遇的时间 τ_0 内追随 ρ_2,而在 τ_0 内可以令 $\partial \rho_2 / \partial t \approx 0$,于是(2.10)式进一步简化为

$$\{[\nabla_r U(|r_1 - r|) \cdot \nabla_p] + [\nabla_{r_1} U(|r_1 - r|) \cdot \nabla_{p_1}] -$$

$$(p/m) \cdot \nabla_r - (p_1/m) \cdot \nabla_{r_1}\} \cdot \rho_2(r, p, r_1, p_1, t) = 0. \tag{2.11}$$

或

$$\nabla_r U(|r_1 - r|) \cdot \nabla_p \rho_2$$

$$= \Big[\frac{p}{m} \cdot \nabla_r + \frac{p_1}{m} \cdot \nabla_{r_1} - \nabla_{r_1} U(|r_1 - r|) \nabla_{p_1} \Big] \rho_2.$$

将上式代入(2.9)式的碰撞积分,并注意到上式右方第三项对碰撞积分没有贡献,则(2.9)式化为

$$\frac{\partial \rho_1}{\partial t} = \Big[\nabla_r \phi(r) \cdot \nabla_p - \frac{p}{m} \cdot \nabla_r \Big] \rho_1(r, p, t) +$$

$$\frac{N}{V} \int \Big(\frac{p}{m} \cdot \nabla_r + \frac{p_1}{m} \cdot \nabla_{r_1} \Big) \rho_2(r, p, r_1, p_1, t) \, \mathrm{d}r_1 \mathrm{d}p_1.$$

再注意到忽略三分子碰撞后,ρ_2 仅通过 $r_{12} = |r_1 - r|$ 而依赖于 r_1,上式又可改写为

$$\frac{\partial \rho_1}{\partial t} = \Big[\nabla_r \phi(r) \cdot \nabla_p - \frac{p}{m} \cdot \nabla_r \Big] \rho_1(r, p, t) +$$

$$\frac{N}{V} \int \Big(\frac{p_1}{m} - \frac{p}{m} \Big) \cdot \nabla_{r_{12}} \rho_2(r, p, r_1, p_1, t) \, \mathrm{d}r_{12} \mathrm{d}p_1. \tag{2.12}$$

考虑到短程作用,则总可以找到一个距离 $R_0 > r_0$,当 $|r_{12}| \geqslant R_0$ 时,两粒子基本上独立运动,于是有

$$\rho_2(r, p, r_1, p_1, t) = \rho_1(r, p, t) \rho_1(r_1, p_1, t) + A(r, p, r_1, p_1, t). \tag{2.13}$$

A 是一个修正项,它考虑了虽然是短程作用,仍然可以延伸到无限的事实.只有对于真正的有限势或硬球,A 才是零.更详细的分析表明❶$r_0 \leqslant R_0 \leqslant l$.

我们再回到(2.12)式中的碰撞积分,利用高斯定理

$$\int (V \cdot \nabla) a \mathrm{d}\tau = \oint_S (V \cdot n) a \mathrm{d}S.$$

---

❶ 见:H. J. Kreuzer. Nonequilibrium Thermodynamics and its Statistical foundation. Clerandor Press. Oxford. 1981.

将对 $|\boldsymbol{r}_{12}|$ 的体积分化为对半径为 R_0 的球面积分 $\mathrm{d}S(R_0)$，在球面上 ρ_2 按(2.13)式分解，于是(2.12)式中的碰撞积分化为

$$\left(\frac{\partial \rho_1}{\partial t}\right)_{碰} = \frac{N}{V}\int \mathrm{d}\boldsymbol{p}_1 \int_{S(R_0)} \left(\frac{\boldsymbol{p}_1}{m} - \frac{\boldsymbol{p}}{m}\right)\cdot \boldsymbol{n}\rho_1(\boldsymbol{r},\boldsymbol{p}'t)\rho_1(\boldsymbol{r}_1,\boldsymbol{p}'_1,t)\,\mathrm{d}S(R_0)$$

$$+ \frac{N}{V}\int \mathrm{d}\boldsymbol{p}_1 \int \mathrm{d}\boldsymbol{r}_1 \left(\frac{\boldsymbol{p}_1}{m} - \frac{\boldsymbol{p}}{m}\right)\cdot \nabla_{r_{12}} \Lambda(\boldsymbol{r},\boldsymbol{p},\boldsymbol{r}_1,\boldsymbol{p}_1,t). \tag{2.14}$$

把注意力集中于上式右边第一项.考虑积分区的球面 $S(R_0)$.如图1,垂直于 $(\boldsymbol{p}_1 - \boldsymbol{p})$ 方向作一截面 ω,在此平面上引入极坐标 (b,ϕ),$0 \leqslant b \leqslant R_0$,$0 \leqslant \phi \leqslant 2\pi$

$$\mathrm{d}\omega = b\mathrm{d}b\mathrm{d}\phi. \tag{2.15}$$

$S(R_0)$ 在平面 ω 上的投影显然覆盖二次.对于上半球,涉及到离开由 $S(R_0)$ 包围的相互作用体积的粒子,这些粒子具有末动量 $\boldsymbol{p}'$ 和 $\boldsymbol{p}'_1$,我们有

$$(\boldsymbol{p}_1 - \boldsymbol{p})\cdot \mathrm{d}\boldsymbol{\Sigma} = (|\boldsymbol{p}_1 - \boldsymbol{p}|)\mathrm{d}\omega > 0. \tag{2.16}$$

对于下半球,涉及到进入相互作用体积的粒子,这些粒子具有初始动量 $\boldsymbol{p}'$ 和 $\boldsymbol{p}'_1$,并有

$$(\boldsymbol{p}_1 - \boldsymbol{p})\cdot \mathrm{d}\boldsymbol{\Sigma} = -(|\boldsymbol{p}_1 - \boldsymbol{p}|)\mathrm{d}\omega < 0. \tag{2.17}$$

于是(2.14)中的第一项取以下形式:

$$\frac{N}{V}\int \mathrm{d}\boldsymbol{p}_1 \int_0^{R_0} b\mathrm{d}b \int_0^{2\pi} \mathrm{d}\phi\,|\boldsymbol{p}_1 - \boldsymbol{p}|\cdot[\rho_1(\boldsymbol{r},\boldsymbol{p}',t)\rho_1(\boldsymbol{r},\boldsymbol{p}'_1,t) - \rho_1(\boldsymbol{r},\boldsymbol{p}',t)\rho_1(\boldsymbol{r},\boldsymbol{p}'_1,t)]$$

$$= \frac{N}{2V}R_0^2 \int \mathrm{d}\boldsymbol{p}_1 \int_0^{2\pi} \mathrm{d}\phi\,|\boldsymbol{p}_1 - \boldsymbol{p}|\cdot[\rho_1(\boldsymbol{r},\boldsymbol{p}',t)\rho_1(\boldsymbol{r},\boldsymbol{p}'_1,t) - \rho_1(\boldsymbol{r},\boldsymbol{p}',t)\rho_1(\boldsymbol{r},\boldsymbol{p}'_1,t)]. \tag{2.18}$$

将上式代入(2.14)式再代入(2.12)式,最后得

$$\frac{\partial \rho_1}{\partial t} = \left[\nabla_r \phi(r)\cdot \nabla_p - \frac{p}{m}\cdot \nabla_r\right]\rho_1(\boldsymbol{r},\boldsymbol{p},t) + \frac{N}{2V}R_0^2 \int \mathrm{d}\boldsymbol{p}_1 \int_0^{2\pi} \mathrm{d}\phi\,|\boldsymbol{p}_1 - \boldsymbol{p}|$$

$$\cdot[\rho_1(\boldsymbol{r},\boldsymbol{p}',t)\rho_1(\boldsymbol{r},\boldsymbol{p}'_1,t) - \rho_1(\boldsymbol{r},\boldsymbol{p}',t)\rho_1(\boldsymbol{r},\boldsymbol{p}'_1,t)]. \tag{2.19}$$

不难看出,它和 §10.1.3(3.16)式(玻尔兹曼积分微分方程)是一致的.

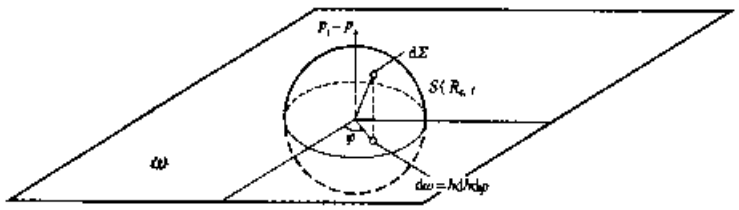
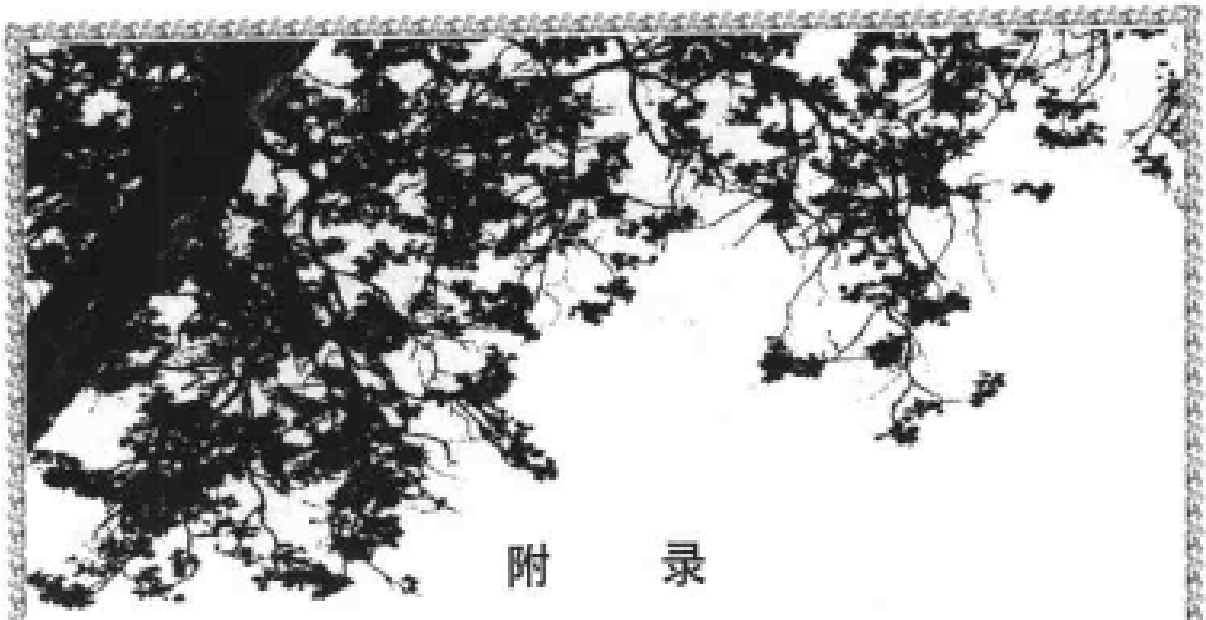


图1 球面 $S(R_0)$ 在垂直于 $(\boldsymbol{p}_1 - \boldsymbol{p})$ 的平面 ω 上的投影



附　录

附录 I　δ 符号和 δ 函数

附录 II　矢量分析与张量计算初步

附录 III　概率的基础知识

附录 IV　拉普拉斯方程的通解

附录 V　一些积分公式

附录 VI　电磁量的国际单位制与高斯单位制

附录 VII　物理常量与单位换算

附录Ⅰ　δ符号与δ函数

(一)克罗内克δ符号

设 $i,j=1,2,\cdots,s$. 定义克罗内克δ符号为

$$\delta_{ij}=\begin{cases}0, & \text{当 } i\neq j\\ 1, & \text{当 } i=j.\end{cases} \tag{Ⅰ.1}$$

当对 j 求和的式子中包含有 δ_{ij} 作为因子时,可以去掉 δ_{ij} 和作和符号,并将其余因子中的脚标 j 改为 i:

$$\sum_{j=1}^{s} b_j\delta_{ij}=b_i. \tag{Ⅰ.2}$$

如果 δ_{ij} 中的 $i,j=1,2,3$,就可以用它写出三维空间中矢量的标积:

$$\boldsymbol{a}\cdot\boldsymbol{b}=\sum_{i=1}^{3}a_ib_i=\sum_{i,j=1}^{3}a_ib_j\delta_{ij}. \tag{Ⅰ.3}$$

利用(Ⅰ.2)式立即可以证明此式.

用 $x_i(i=1,2,3)$ 和 $p_i(i=1,2,3)$ 分别表示坐标和动量的三个笛卡儿分量 (x,y,z) 和 (p_x,p_y,p_z),则由 δ_{ij} 的定义式(Ⅰ.1)显然有:

$$\frac{\partial x_i}{\partial x_j}=\delta_{ij},\qquad \frac{\partial p_i}{\partial p_j}=\delta_{ij}. \tag{Ⅰ.4}$$

(二)三阶全反对称ε符号❶

设 $i,j,k=1,2,3$,定义 ε_{ijk} 对任意两个脚标的交换反对称:

$$\varepsilon_{ijk}=-\varepsilon_{jik}=-\varepsilon_{ikj}=-\varepsilon_{kji}=\varepsilon_{jki}=\varepsilon_{kij}. \tag{Ⅰ.5}$$

因此,ε_{ijk} 的 $3^3=27$ 个分量中,凡是有两个脚标取相同值的分量都是零. 不为零的只是 i,j,k 各不相同,分别取 1,2,3 三个值的六个分量. 这六个分量由(Ⅰ.5)式相联系,独立分量只有一个. 再规定:

$$\varepsilon_{123}=1, \tag{Ⅰ.6}$$

就完全确定了 ε_{ijk}.

利用 ε_{ijk} 可以展开三阶行列式:

$$\begin{vmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{vmatrix}=\sum_{i,j,k}\varepsilon_{ijk}a_ib_jc_k. \tag{Ⅰ.7}$$

证　行列式的一般项是 $\pm a_ib_jc_k(i\neq j\neq k)$. 取正号还是取负号,决定于 ijk 是偶排列还是奇

---

❶ 当脚标取值 1,2,3 时,全反对称符号是三阶的;而当脚标取值 1,2,3,4 时,全反对称符号是四阶的. 令 $\alpha,\beta,\gamma,\delta=1,2,3,4$,则全反对称符号 $\varepsilon_{\alpha\beta\gamma\delta}$ 对任意两个脚标的交换反对称. 它只有一个独立脚标,取为 $\varepsilon_{1234}=1$.

排列. 由 ε_{ijk} 的性质(Ⅰ.5)(Ⅰ.6)式知道,这正好就是(Ⅰ.7)的右边. 【证毕】

ε_{ijk} 和 δ_{ij} 通过以下公式相联系:

$$\sum_k \varepsilon_{ijk}\varepsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}. \tag{Ⅰ.8}$$

证 由等式左边的两个 ε 因子知道,等式左边不为零的条件是

$$i \neq j \neq k, \qquad l \neq m \neq k.$$

此二式同时成立,只有两种可能:

$$i = l, j = m, i \neq j, \qquad \text{或 } i = m, j = l, i \neq j.$$

在前一情况下,排列 ε_{ijk} 和 ε_{klm} 有相同奇偶性,因而 $\varepsilon_{ijk}\varepsilon_{klm} = +1$;在后一情况下,排列 ε_{ijk} 和 ε_{klm} 有相反奇偶性,因而 $\varepsilon_{ijk}\varepsilon_{klm} = -1$:

$$\sum_k \varepsilon_{ijk}\varepsilon_{klm} = \begin{cases} +1, & \text{当 } i = l, j = m, i \neq j, \\ -1, & \text{当 } i = m, j = l, i \neq j. \end{cases}$$

(Ⅰ.8)式右边当 $i = l, j = m, i \neq j$(因而 $i \neq m$)时,只有第一项,因而等于 +1;当 $i = m, j = l$, $i \neq j$(因而 $i \neq l$)时,只有第二项,因而等于 -1. 这正好和上式相符,还剩下一种情况要考虑,即 $i = j = l = m$,此时(Ⅰ.8)式右边两项都等于1,相减为零. 因此(Ⅰ.8)式的右边和左边在所有情况下都相等. 【证毕】

利用 ε_{ijk} 可以将矢量的叉乘缩写为

$$(\boldsymbol{a} \times \boldsymbol{b})_i = \sum_{j,k=1}^{3} \varepsilon_{ijk} a_j b_k. \tag{Ⅰ.9}$$

证 当 $i = 1$ 时,j, k 只能等于2,3,而且 $\varepsilon_{123} = +1, \varepsilon_{132} = -1$. 由此可见,$\varepsilon_{1jk}a_j b_k = a_2 b_3 - a_3 b_2$,而这就是 $(\boldsymbol{a} \times \boldsymbol{b})_1$;对 $i = 2,3$ 也可类似地证明. 【证毕】

将(Ⅰ.9)乘上基矢 $\mathbf{e}_i$ 对 i 求和,得到

$$\boldsymbol{a} \times \boldsymbol{b} = \sum_{i,j,k} \varepsilon_{ijk} \boldsymbol{e}_i a_j b_k. \tag{Ⅰ.10}$$

利用(Ⅰ.3)(Ⅰ.8)(Ⅰ.9)式很容易推导一些矢量公式. 例如:

$$\begin{aligned} \boldsymbol{a} \times [\boldsymbol{b} \times \boldsymbol{c}] &= \sum_{ijk} \boldsymbol{e}_i \varepsilon_{ijk} a_j [\boldsymbol{b} \times \boldsymbol{c}]_k = \sum_{ij} \sum_{lm} \sum_{k} \boldsymbol{e}_i \varepsilon_{ijk}\varepsilon_{klm} a_j b_l c_m \\ &= \sum_{ij} \sum_{lm} \boldsymbol{e}_i (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}) a_j b_l c_m = \sum_i \sum_j [\boldsymbol{e}_i b_i (a_j c_j) - \boldsymbol{e}_i c_i (a_j b_j)] \\ &= \boldsymbol{b}(\boldsymbol{a} \cdot \boldsymbol{c}) - \boldsymbol{c}(\boldsymbol{a} \cdot \boldsymbol{b}). \end{aligned}$$

(三)狄拉克 δ 函数

带电细线上长为 d 的一段,所带电量为1,置于坐标原点,则电荷的线密度为

$$\rho(x) = \begin{cases} 0, & \text{当 } |x| > d/2, \\ 1/d, & \text{当 } |x| \leqslant d/2. \end{cases}$$

它满足

$$\int_{-\infty}^{\infty} \rho(x)\,\mathrm{d}x = 1.$$

当 $d \to 0$ 时,成为点电荷. 此时,电荷密度 $\rho(x)$ 成为 δ 函数,其定义为:

$$\begin{cases}\delta(x) = \begin{cases}0, & \text{当 } x \neq 0,\\ \infty, & \text{当 } x = 0.\end{cases} & (\text{Ⅰ}.11)\\ \int_{-\infty}^{\infty}\delta(x)\mathrm{d}x = 1. & (\text{Ⅰ}.12)\end{cases}$$

$\delta(x)$ 是偶函数，$\delta(-x) = \delta(x)$；它具有以下重要性质：设 $f(x)$ 为任意连续函数，则

$$\int_{-\infty}^{\infty} f(x)\delta(x - x_0)\mathrm{d}x = f(x_0). \qquad (\text{Ⅰ}.13)$$

证　由于 $x \neq x_0$ 时 $\delta(x - x_0) = 0$，故令 ε 为一个小正数，有

$$\int_{-\infty}^{\infty} f(x)\delta(x - x_0)\mathrm{d}x = \int_{-\varepsilon}^{\varepsilon} f(x)\delta(x - x_0)\mathrm{d}x = f(x_0)\int_{-\varepsilon}^{\varepsilon}\delta(x - x_0)\mathrm{d}x$$

$$= f(x_0)\int_{-\infty}^{\infty}\delta(x - x_0)\mathrm{d}x = f(x_0). \qquad \text{【证毕】}$$

比较(Ⅰ.11)、(Ⅰ.12)、(Ⅰ.13)式和(Ⅰ.1)、(Ⅰ.2)式可见，δ函数是δ符号在取分立值的 i,j 过渡为取连续值的 x_0, x 时的极限. 在进行这一过渡时，(Ⅰ.2)式中对 j 的求和过渡为(Ⅰ.13)式中对 x 的积分.

δ函数的一个有用的表达式是：

$$\delta(x - x_0) = \delta(x_0 - x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}k(x - x_0)}\mathrm{d}k. \qquad (\text{Ⅰ}.14)$$

证　第一个等式是由于δ函数是偶函数. 为证第二个等式，取任意连续函数 $f(x)$ 的傅里叶积分变换：

$$f(x_0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} c(k)\mathrm{e}^{-\mathrm{i}kx_0}\mathrm{d}k, \qquad c(k) = \int_{-\infty}^{\infty} f(x)\mathrm{e}^{\mathrm{i}kx}\mathrm{d}x.$$

将后式代入前式，得到

$$f(x_0) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x)\mathrm{e}^{\mathrm{i}k(x - x_0)}\mathrm{d}x\mathrm{d}k = \int_{-\infty}^{\infty}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}k(x - x_0)}\mathrm{d}k\right]f(x)\mathrm{d}x.$$

和(Ⅰ.13)比较可见，方括号中的表达式就是 $\delta(x - x_0)$. 【证毕】

利用(Ⅰ.14)式可以证明以下公式：

$$\delta(ax) = \delta(x)/|a|. \qquad (\text{Ⅰ}.15)$$

证　设 $a > 0$，

$$\delta(ax) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kax}\mathrm{d}x = \frac{1}{2\pi a}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}(ak)x}\mathrm{d}(ax) = \frac{1}{a}\delta(x);$$

设 $a < 0$，则 $a = -|a|$，利用δ函数的偶函数性质，

$$\delta(ax) = \delta(-|a|x) = \delta(|a|x) = \delta(x)/|a|.$$

【证毕】

δ函数的另一个有用的表达式是

$$\delta(x) = \lim_{u\to\infty}\frac{\sin^2 xu}{\pi x^2 u}. \qquad (\text{Ⅰ}.16)$$

证　当 $x \neq 0$ 时，上式中的分子有限而分母随 u 一道趋于无穷，因而极限为零；当 $x = 0$ 时，$\sin^2 xu/(xu)^2 = 1$，故上式正比于 $u \to \infty$. 这样，(Ⅰ.11)式得到满足. 再看积分

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin^2 xu}{x^2u}dx = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin^2 xu}{(xu)^2}d(xu) = 1,$$

因而(Ⅰ.12)式也满足. 【证毕】

三维 δ 函数的定义是:

$$\begin{cases}\delta^3(\boldsymbol{r}) = \begin{cases}0, & \text{当 } r \neq 0,\\ \infty, & \text{当 } r = 0.\end{cases} & (\text{Ⅰ}.17)\\ \int_{-\infty}^{\infty}\delta^3(\boldsymbol{r})\mathrm{d}^3\boldsymbol{r} = 1. & (\text{Ⅰ}.18)\end{cases}$$

它可以看成是三个一维 δ 函数的乘积:

$$\delta^3(\boldsymbol{r}-\boldsymbol{r}_0) = \delta(x-x_0)\delta(y-y_0)\delta(z-z_0). \quad (\text{Ⅰ}.19)$$

利用(Ⅰ.14)得到它的一个表达式:

$$\delta^3(\boldsymbol{r}-\boldsymbol{r}_0) = \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}e^{i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}_0)}\mathrm{d}^3\boldsymbol{k}. \quad (\text{Ⅰ}.20)$$

(Ⅰ.20)式是在无限区间中积分,得到 δ 函数的表达式. 如果是在有限区间中积分就得到 δ 符号的表达式.

$$\delta_{mm'} = \frac{1}{L}\int_0^L e^{i(m-m')2\pi k/L}\mathrm{d}\,k. \quad (\text{Ⅰ}.21)$$

或写 $\varphi = 2\pi k/L$,就有

$$\delta_{mm'} = \frac{1}{2\pi}\int_0^{2\pi}e^{i(m-m')\varphi}\mathrm{d}\,\varphi. \quad (\text{Ⅰ}.22)$$

三维 δ 函数还有一个有用的表达式

$$\delta^3(\boldsymbol{r}) = -\frac{1}{4\pi}\Delta\frac{1}{r}. \quad (\text{Ⅰ}.23)$$

证明见(Ⅱ.66)式. 上式中 Δ 是拉普拉斯算子,其定义见(Ⅱ.18)式.

附录Ⅱ　矢量分析与张量计算初步

(一)梯度、散度、旋度的定义

如果在空间某区域内定义了标量函数 $\varphi(r,t)$ 或矢量函数 $J(r,t)$，则 $\varphi(r,t)$，$J(r,t)$ 分别称为标量场和矢量场，其中，时间 t 是参量，以下不再明显写出.

1. 梯度

设给定标量场 $\varphi(r)$，方程 $\varphi(r)=c$ 决定一个曲面，称为等 φ 值面. 例如，将地图上每一点 r 的高度定义为 $\varphi(r)$，则等 φ 值线就是等高线.

在某一点 P 的 φ 值为 $\varphi(P)$. 离开 P 点，φ 值会变. 沿不同方向离开 P 点，φ 值变化的快慢不同. 令 e_n 为 φ 值增加最快(变化最陡)的方向上的单位矢量，n 为沿此方向量度的离开 P 点的距离，则在 p 点处标量场 $\varphi(r)$ 的梯度定义为：

$$\operatorname{grad}\varphi=\frac{\partial\varphi}{\partial n}e_n. \tag{Ⅱ.1}$$

用 r 表示 P 点的位矢，对于每一个 r 都有一个 $\operatorname{grad}\varphi$，所以 $\operatorname{grad}\varphi$ 是一个矢量场.

用 e_l 表示任一单位矢量，l 为沿 e_l 方向量度的距离. $\operatorname{grad}\varphi$ 在 e_l 方向的投影是

$$(\operatorname{grad}\varphi)_l=\operatorname{grad}\varphi\cdot e_l=\frac{\partial\varphi}{\partial n}e_n\cdot e_l=\frac{\partial\varphi}{\partial n}\cos\theta_{nl}. \tag{Ⅱ.2}$$

式中 θ_{nl} 为 n 与 l 的夹角，但

$$\cos(\theta_{nl})=\lim_{\Delta l\to 0}\frac{\Delta n}{\Delta l}=\frac{\mathrm{d}n}{\mathrm{d}l},\qquad \frac{\partial\varphi}{\partial n}\frac{\mathrm{d}n}{\mathrm{d}l}=\frac{\partial\varphi}{\partial l},$$

故

$$(\operatorname{grad}\varphi)_l=\frac{\partial\varphi}{\partial l}. \tag{Ⅱ.3}$$

用 e_i 分别表示三个笛卡儿坐标方向的单位矢量，$e_i=e_i$($i=1,2,3$ 分别表示 x,y,z)，可以得到 $\operatorname{grad}\varphi$ 在三个坐标轴方向的分量：

$$(\operatorname{grad}\varphi)_i=\partial\varphi/\partial x_i.$$

由此得

$$\operatorname{grad}\varphi=\sum_{i=1}^{3}e_i(\operatorname{grad}\varphi)_i=\sum_{i=1}^{3}e_i\frac{\partial\varphi}{\partial x_i}. \tag{Ⅱ.4}$$

这也可以作为 $\operatorname{grad}\varphi$ 的定义.

定义矢量微分算符：

$$\nabla=\sum_{i=1}^{3}e_i\nabla_i=\sum_{i=1}^{3}e_i\frac{\partial}{\partial x_i}=e_x\frac{\partial}{\partial x}+e_y\frac{\partial}{\partial y}+e_z\frac{\partial}{\partial z}. \tag{Ⅱ.5}$$

则(Ⅱ.4)式可写为

$$\operatorname{grad}\varphi=\nabla\varphi. \tag{Ⅱ.6}$$

利用梯度可写出相距为 $\mathrm{d}r$ 的两点间 φ 的改变量

$$\mathrm{d}\varphi=\sum(\partial\varphi/\partial x_i)\mathrm{d}x_i=\operatorname{grad}\varphi\cdot\mathrm{d}r=\nabla\varphi\cdot\mathrm{d}r. \tag{Ⅱ.7}$$

2. 散度 高斯定理

设给定矢量场 $\boldsymbol{J}(\boldsymbol{r})$. 为了形象化,可以把 $\boldsymbol{J}$ 理解为某种流密度,例如,质量流密度或电流密度. 用 S_0 表示某一曲面, $\mathrm{d}\boldsymbol{\sigma}$ 表示 S_0 上的一个小面积元. 矢量 $\mathrm{d}\boldsymbol{\sigma}$ 垂直于此面积元的法线. 按定义,

$$\int_{S_0} \boldsymbol{J}\cdot\mathrm{d}\boldsymbol{\sigma}$$

是单位时间流过 S_0 面的"荷"(如质量、电荷).

设 ΔV 为空间的某一体积, S 为包围此体积的封闭面,令 $\mathrm{d}\boldsymbol{\sigma}$ 为 S 面上的面积元,指向外法线方向,则

$$I = \oint_S \boldsymbol{J}\cdot\mathrm{d}\boldsymbol{\sigma}$$

是单位时间通过封闭面 S 流出的荷. 如果 ΔV 内有产生这种荷的源,则 $I>0$;如果 ΔV 内有吸收这种荷的区域(这种区域称为尾闾),则 $I<0$;如果 ΔV 内既无产生这种荷的源头,又无吸收这种荷的尾闾,则 $I=0$. 将 I 用 ΔV 除,让 $\Delta V\to 0$,就得到 ΔV 所在位置(设其位矢为 $\boldsymbol{r}$),单位体积的源头(或尾闾)的强度,称为矢量场 $\boldsymbol{J}(\boldsymbol{r})$ 在 $\boldsymbol{r}$ 点的散度:

$$\operatorname{div}\boldsymbol{J} = \lim_{\Delta V\to 0}\frac{1}{\Delta V}\oint_S \boldsymbol{J}\cdot\mathrm{d}\boldsymbol{\sigma}. \tag{Ⅱ.8}$$

由此式可见, $\operatorname{div}\boldsymbol{J}$ 是标量. 在每一点 $\boldsymbol{r}$ 都有这样一个标量,所以 $\operatorname{div}\boldsymbol{J}$ 是一个标量场.

在笛卡儿坐标中, $\operatorname{div}\boldsymbol{J}$ 的表达式是

$$\operatorname{div}\boldsymbol{J} = \frac{\partial J_x}{\partial x}+\frac{\partial J_y}{\partial y}+\frac{\partial J_z}{\partial z} = \sum_{i=1}^{3}\frac{\partial J_i}{\partial x_i} = \nabla\cdot\boldsymbol{J}. \tag{Ⅱ.9}$$

证 考虑平行于坐标面的长宽高分别为 $\mathrm{d}x,\mathrm{d}y,\mathrm{d}z$ 的平行六面体 $\mathrm{d}V=\mathrm{d}x\mathrm{d}y\mathrm{d}z$. 封闭面 S 由六个平面组成. 先来看位于 x 和 $x+\mathrm{d}x$ 处,平行于 yz 面的两个平面,它们的面积是 $\mathrm{d}y\mathrm{d}z$. 外法线方向分别指向 $-x$ 轴和 $+x$ 轴,在这两个面积元上, $\boldsymbol{J}\cdot\mathrm{d}\boldsymbol{\sigma}$ 分别是

$$-J_x(x,y,z)\mathrm{d}y\mathrm{d}z \quad 和 \quad J_x(x+\mathrm{d}x,y,z)\mathrm{d}y\mathrm{d}z. \tag{Ⅱ.10}$$

将这两项相加,得到

$$[J_x(x+\mathrm{d}x,y,z)\mathrm{d}y\mathrm{d}z - J_x(x,y,z)\mathrm{d}y\mathrm{d}z] = \left(\frac{\partial J_x}{\partial x}\mathrm{d}x\right)\mathrm{d}y\mathrm{d}z = \frac{\partial J_x}{\partial x}\mathrm{d}V. \tag{Ⅱ.11}$$

类似地,通过 y 和 $y+\mathrm{d}y$ 处,平行于 zx 面的两个平面流出 $\mathrm{d}V$ 的 $\boldsymbol{J}\cdot\mathrm{d}\boldsymbol{\sigma}$ 是 $(\partial J_y/\partial y)\mathrm{d}V$;通过位于 z 和 $z+\mathrm{d}z$ 处,平行于 xy 面的两个平面流出 $\mathrm{d}V$ 的 $\boldsymbol{J}\cdot\mathrm{d}\boldsymbol{\sigma}$ 是 $(\partial J_z/\partial z)\mathrm{d}V$. 三项相加,得到通过封闭面流出的总量

$$\int_S \boldsymbol{J}\cdot\mathrm{d}\boldsymbol{\sigma} = \left(\frac{\partial J_x}{\partial x}+\frac{\partial J_y}{\partial y}+\frac{\partial J_z}{\partial z}\right)\mathrm{d}V,$$

代入(Ⅱ.8)式即得(Ⅱ.9)式. 【证毕】

现在来证明高斯定理:

$$\int_V \operatorname{div}\boldsymbol{J}\mathrm{d}V = \oint_S \boldsymbol{J}\cdot\mathrm{d}\boldsymbol{\sigma}. \tag{Ⅱ.12}$$

证 将区域 V 分成无限多个无穷小体积元 ΔV_i,用 S_i 表示包围 ΔV_i 的小封闭面,则有

$$\lim_{\Delta V_i \to 0} \operatorname{div} \boldsymbol{J} \Delta V_i = \oint_{S_i} \boldsymbol{J} \cdot \mathrm{d}\boldsymbol{\sigma}.$$

将左边对全部 ΔV_i 求和，得到(Ⅱ.12)的左边. 右边求和的结果，除了得到(Ⅱ.12)的右边外，还有一些内表面项，每个这种内表面同时属于两个相邻体积元，它们的"外"法线方向相反，所以对(Ⅱ.12)的贡献相互抵消. 【证毕】

3. 旋度　斯托克斯定理

由(Ⅱ.9)式可见，任意矢量场 $\boldsymbol{A}(r)$ 的散度是矢量微分算子 ∇ 和 $\boldsymbol{A}$ 点乘得到的一个标量场. 还可以将 ∇ 和 $\boldsymbol{A}$ 叉乘，得到一个矢量场，称为 $\boldsymbol{A}$ 的旋度：

$$\operatorname{rot} \boldsymbol{A} = \nabla \times \boldsymbol{A} = \sum_{ijk} \varepsilon_{ijk} \boldsymbol{e}_i \nabla_j A_k. \tag{Ⅱ.13}$$

明显写出就是

$$(\operatorname{rot} \boldsymbol{A})_x = \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}, \quad (\operatorname{rot} \boldsymbol{A})_y = \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}, \quad (\operatorname{rot} \boldsymbol{A})_z = \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}. \tag{Ⅱ.13′}$$

类似于散度所满足的高斯定理，对于旋度有斯托克斯定理：

$$\int_S \operatorname{rot} \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{\sigma} = \oint_l \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{l}. \tag{Ⅱ.14}$$

其中 S 是一个曲面，l 是包围这一曲面的闭合曲线. 证明留给读者.

(二)高阶导数与乘积的导数

1. 下标运算的一些规则

在采用笛卡儿坐标的分量进行矢量分析的计算时，经常遇到一些带有下标 i,j,k 的量，对于它们有以下规则需要注意：(a)进行了求和的下标称为哑标. 在求和中，哑标取遍 1,2,3 三个值；求和以后，哑标实际上已不存在(类似于定积分的积分变量). 因此，对于哑标可以随意改用任何符号. 但是，在两个求和式子相乘时，哑标不能重复. 同一式子中哑标和不求和的下标(称为自由下标)也不能重复. (b)如果相乘的两个因子中，一个因子对下标 i,j 对称，另一个因子对下标 i,j 反对称，则相乘以后对 ij 求和得零.

证　设 $A_{ij} = A_{ji}, B_{ij} = -B_{ji}$，则

$$\sum_{ij}^{3} A_{ij} B_{ij} = \sum_{ij}^{3} A_{ji} B_{ji} = \sum_{ij}^{3} A_{ij}(-B_{ij}) = 0. \tag{Ⅱ.15}$$

以上第一步利用了规则(a)，将哑标 i,j 互换；第二步利用了 A_{ij} 和 B_{ij} 对 ij 的对称性和反对称性. 最后一步是根据：$a = -a$，则 $a = 0$. 【证毕】

2. 无旋场与无源场

旋度为零的场称为无旋场；散度为零的场称为无源场.

(a)标量场的梯度必为无旋场：

$$\operatorname{rot} \operatorname{grad} \varphi = 0. \tag{Ⅱ.16}$$

证　$\operatorname{rot} \operatorname{grad} \varphi = \nabla \times \nabla \varphi = \sum \boldsymbol{e}_i \varepsilon_{ijk} \nabla_j \nabla_k \varphi$，但是 ε_{ijk} 对 jk 反对称，而 $\nabla_j \nabla_k$ 对 jk 对称，故得零. 【证毕】

(b)无旋场必可表为标量场的梯度，准确到任意相加常数，即：设 $\operatorname{rot} \boldsymbol{f} = 0$，则必有 φ 使 $\boldsymbol{f} =$

$\nabla\varphi$. 对 φ 可以加上任意常数 c，因为 $\nabla c = 0$.

(c)矢量场的旋度必为无源场：

$$\operatorname{div}\operatorname{rot} \boldsymbol{f} = 0. \tag{Ⅱ.17}$$

证 $\operatorname{div}\operatorname{rot}\boldsymbol{f} = \nabla\cdot(\nabla\times\boldsymbol{f}) = \sum\delta_{ij}\nabla_i\varepsilon_{jkl}\nabla_k f_l = \sum\nabla_j\varepsilon_{jkl}\nabla_k f_l = 0$.

【证毕】

(d)无源场必可表为另一矢量场的旋度，准确到任意梯度. 即：设 $\operatorname{div}\boldsymbol{B} = 0$，则必有矢量场 $\boldsymbol{A}$，使 $\boldsymbol{B} = \operatorname{rot}\boldsymbol{A}$. 对 $\boldsymbol{A}$ 可以加上任一标量场的梯度 $\operatorname{grad}\varphi$，因为 $\operatorname{rot}\operatorname{grad}\varphi = 0$. 【证毕】

3. 不为零的二阶导数有：

$$\triangle \equiv \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}, \tag{Ⅱ.18}$$

称为拉普拉斯算子；以及：

$$\begin{aligned}\operatorname{rot}\operatorname{rot}\boldsymbol{a} &= \sum\boldsymbol{e}_i\varepsilon_{ijk}\nabla_j\varepsilon_{klm}\nabla_l a_m = \sum\boldsymbol{e}_i(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\nabla_j\nabla_l a_m \\ &= \operatorname{grad}\operatorname{div}\boldsymbol{a} - \triangle\boldsymbol{a}.\end{aligned} \tag{Ⅱ.19}$$

4. 乘积的导数. 计算方法是：先利用(Ⅱ.4)、(Ⅱ.9)、(Ⅱ.13)写成分量式，再按通常乘积求导计算. 例如：

$$\begin{aligned}\nabla\times(\varphi\boldsymbol{f}) &= \sum\boldsymbol{e}_i\varepsilon_{ijk}\nabla_j(\varphi f_k) = \sum\boldsymbol{e}_i\varepsilon_{ijk}[(\nabla_j\varphi)f_k + \varphi\nabla_j f_k] \\ &= (\nabla\varphi)\times\boldsymbol{f} + \varphi\nabla\times\boldsymbol{f} = (\operatorname{grad}\varphi)\times\boldsymbol{f} + \varphi\operatorname{rot}\boldsymbol{f}.\end{aligned} \tag{Ⅱ.20}$$

$$\begin{aligned}\nabla\times(\boldsymbol{a}\times\boldsymbol{b}) &= \sum\boldsymbol{e}_i\varepsilon_{ijk}\varepsilon_{klm}\nabla_j a_l b_m \\ &= \sum\boldsymbol{e}_i(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})[(\nabla_j a_l)b_m + a_l(\nabla_j b_m)] \\ &= (\boldsymbol{b}\cdot\nabla)\boldsymbol{a} - \boldsymbol{b}(\nabla\cdot\boldsymbol{a}) + \boldsymbol{a}(\nabla\cdot\boldsymbol{b}) - (\boldsymbol{a}\cdot\nabla)\boldsymbol{b}).\end{aligned} \tag{Ⅱ.21}$$

$$\operatorname{grad}(\boldsymbol{a}\cdot\boldsymbol{b}) = \sum\boldsymbol{e}_i\nabla_i a_j b_j = \sum\boldsymbol{e}_i[(\nabla_i a_j)b_j + a_j\nabla_i b_j].$$

为了将等式右边用梯度、散度、旋度表示出来，采用一种技巧，先算第一项，在其中插进两个 δ 符号：

$$\begin{aligned}\sum\boldsymbol{e}_i(\nabla_i a_j)b_j &= \sum\boldsymbol{e}_i\delta_{im}\delta_{jl}(\nabla_m a_l)b_j = \sum\boldsymbol{e}_i(\delta_{il}\delta_{jm} - \varepsilon_{ijk}\varepsilon_{klm})(\nabla_m a_l)b_j \\ &= \sum\boldsymbol{e}_i(\delta_{il}\delta_{jm} + \varepsilon_{ijk}\varepsilon_{kml})(\nabla_m a_l)b_j \\ &= (\boldsymbol{b}\cdot\nabla)\boldsymbol{a} + \boldsymbol{b}\times(\nabla\times\boldsymbol{a}).\end{aligned}$$

第二项也可以类似地计算. 最后得到

$$\operatorname{grad}(\boldsymbol{a}\cdot\boldsymbol{b}) = (\boldsymbol{b}\cdot\nabla)\boldsymbol{a} + \boldsymbol{b}\times(\nabla\times\boldsymbol{a}) + (\boldsymbol{a}\cdot\nabla)\boldsymbol{b} + \boldsymbol{a}\times(\nabla\times\boldsymbol{b}). \tag{Ⅱ.22}$$

(三)积分降维公式与格林定理

1. 将高斯定理和斯托克斯定理推广，得到体积分变面积分和面积分变线积分的公式：

$$\int\mathrm{d}V\,\nabla \leftrightarrow \oint\mathrm{d}\boldsymbol{\sigma}; \qquad \int\mathrm{d}\boldsymbol{\sigma}\times\nabla \leftrightarrow \oint\mathrm{d}\boldsymbol{l}. \tag{Ⅱ.23}$$

这相当于以下六个公式：

$$\int\mathrm{d}V\,\nabla\cdot\boldsymbol{A} = \oint\mathrm{d}\boldsymbol{\sigma}\cdot\boldsymbol{A}; \qquad \int\mathrm{d}\boldsymbol{\sigma}\cdot(\nabla\times\boldsymbol{A}) = \oint\mathrm{d}\boldsymbol{l}\cdot\boldsymbol{A}, \tag{Ⅱ.24}$$

$$\int \mathrm{d}V\,\nabla \times \boldsymbol{f} = \oint \mathrm{d}\boldsymbol{\sigma} \times \boldsymbol{f}; \qquad \int (\mathrm{d}\boldsymbol{\sigma} \times \nabla) \times \boldsymbol{f} = \oint \mathrm{d}\boldsymbol{l} \times \boldsymbol{f}, \tag{Ⅱ.25}$$

$$\int \mathrm{d}V\,\nabla \varphi = \oint \mathrm{d}\boldsymbol{\sigma}\,\varphi; \qquad \int \mathrm{d}\boldsymbol{\sigma} \times \nabla \varphi = \oint \mathrm{d}\boldsymbol{l}\,\varphi. \tag{Ⅱ.26}$$

(Ⅱ.24)式就是高斯定理和斯托克斯定理,令 $\boldsymbol{A} = \boldsymbol{f} \times \boldsymbol{c}$ 或令 $\boldsymbol{A} = c\varphi$,其中 c 为常矢量,代入(Ⅱ.24)可以证明(Ⅱ.25),(Ⅱ.26).

2. 格林定理

$$\oint_S \varphi \nabla \psi \cdot \mathrm{d}\boldsymbol{\sigma} = \int_V [\varphi \triangle \psi + \nabla \varphi \cdot \nabla \psi]\mathrm{d}V, \tag{Ⅱ.27}$$

$$\oint_S (\varphi \nabla \psi - \psi \nabla \varphi) \cdot \mathrm{d}\boldsymbol{\sigma} = \int_V (\varphi \triangle \psi - \psi \triangle \varphi)\mathrm{d}V, \tag{Ⅱ.28}$$

$$\oint_S \varphi \nabla \varphi \cdot \mathrm{d}\boldsymbol{\sigma} = \int_V [\varphi \triangle \varphi + (\nabla \varphi)^2]\mathrm{d}V, \tag{Ⅱ.29}$$

证　由高斯定理

$$\oint_S \varphi \nabla \psi \cdot \mathrm{d}\boldsymbol{\sigma} = \int_V \nabla \cdot (\varphi \nabla \psi)\mathrm{d}V$$

将被积函数展开即得(Ⅱ.27). 在(Ⅱ.27)中令 $\varphi \rightleftharpoons \psi$ 再与(Ⅱ.27)相减,得(Ⅱ.28). 令 $\varphi = \psi$ 得(Ⅱ.29). **【证毕】**

(四)正交曲线坐标中的矢量微分公式

用 $x_i(i = 1,2,3)$ 表示笛卡儿坐标. 设有 x_i 的三个单值连续可微函数

$$q_\alpha = q_\alpha(x_1, x_2, x_3), \qquad \alpha = 1,2,3, \tag{Ⅱ.30}$$

它的反函数

$$x_i = x_i(q_1, q_2, q_3), \qquad i = 1,2,3 \tag{Ⅱ.31}$$

也单值连续可微,则 $q_\alpha(\alpha = 1,2,3)$ 可以代替 $x_i(i = 1,2,3)$ 作为空间点的坐标,称为曲线坐标,例如,球坐标 $(q_1 = r, q_2 = \theta, q_3 = \phi)$. 于是

$$x_1 = q_1 \sin q_2 \cos q_3, \quad x_2 = q_1 \sin q_2 \sin q_3, \quad x_3 = q_1 \cos q_2. \tag{Ⅱ.32}$$

(Ⅱ.31)式可写为

$$\boldsymbol{r} = \boldsymbol{r}(q_1, q_2, q_3), \tag{Ⅱ.33}$$

其中 $\boldsymbol{r} = \sum x_i \boldsymbol{e}_i$ 是空间点的位矢. 让某一个 q_α 变,而保持另外两个 $q_\beta(\beta \neq \alpha)$ 不变,则 $\boldsymbol{r}$ 的终点描出一根曲线,称为坐标线 q_α. 在空间任一点 $\boldsymbol{r}$ 作单位矢量 $\boldsymbol{e}_\alpha$ 和坐标线 q_α 相切,并指向 q_α 增加的方向,如§1.1.1 图2. 这样三个 $\boldsymbol{e}_\alpha(\alpha = 1,2,3)$ 形成一个坐标系的基矢,它们在空间不同点有不同的方向,称为局部标架. §1.1.1 图2 上画出了球坐标系中的局部标架. 如果在空间每一点 P,局部标架的基矢都相互正交

$$\boldsymbol{e}_\alpha(P) \cdot \boldsymbol{e}_\beta(P) = \delta_{\alpha\beta}, \tag{Ⅱ.34}$$

则称这一曲线坐标为正交曲线坐标. 球坐标和柱坐标都是正交曲线坐标.

求(Ⅱ.33)的微分

$$\mathrm{d}\boldsymbol{r} = \sum_{\alpha=1}^{3} \frac{\partial \boldsymbol{r}}{\partial q_\alpha} \mathrm{d}q_\alpha. \tag{Ⅱ.35}$$

如果只有一个 q_α 增加,而另外两个保持不变,则 $\mathrm{d}\boldsymbol{r}$ 指向 $\boldsymbol{e}_\alpha$ 方向,此时(Ⅱ.35)式右边只剩下

$\partial r/\partial q_\alpha$ 一项，故

$$\frac{\partial r}{\partial q_\alpha} = H_\alpha e_\alpha,$$

$$H_\alpha = \left|\frac{\partial r}{\partial q_\alpha}\right| = \sqrt{\sum_{i=1}^{3}\left(\frac{\partial x_i}{\partial q_\alpha}\right)^2}. \tag{Ⅱ.36}$$

H_α 称为拉梅系数. 对于球坐标，由(Ⅱ.32)、(Ⅱ.36)得

$$H_r = 1, \quad H_\theta = r, \quad H_\phi = r\sin\theta. \tag{Ⅱ.37}$$

柱坐标定义为 $q_1 = \rho$, $q_2 = \phi$, $q_3 = z$(参看第一篇第一章习题图1)，它和笛卡儿坐标的关系是：

$$x_1 = q_1\cos q_2, \quad x_2 = q_1\sin q_2, \quad x_3 = q_3. \tag{Ⅱ.38}$$

代入(Ⅱ.36)式得拉梅系数

$$H_\rho = 1, \quad H_\phi = \rho, \quad H_z = 1. \tag{Ⅱ.39}$$

利用(Ⅱ.35)式，得到

$$\mathrm{d}r = \sum_{\alpha=1}^{3}\frac{\partial r}{\partial q_\alpha}\mathrm{d}q_\alpha = \sum_{\alpha=1}^{3} e_\alpha H_\alpha \mathrm{d}q_\alpha, \tag{Ⅱ.40}$$

故 $\mathrm{d}r$ 在局部标架 $e_\alpha(\alpha = 1,2,3)$ 中的分量是

$$(\mathrm{d}r)_\alpha = H_\alpha \mathrm{d}q_\alpha. \tag{Ⅱ.41}$$

例如，在球坐标中，$(\mathrm{d}r)_r = \mathrm{d}r$, $(\mathrm{d}r)_\theta = r\mathrm{d}\theta$, $(\mathrm{d}r)_\phi = r\sin\theta\mathrm{d}\phi$；由它们形成的体积元(见§1.1.1图2)是 $\mathrm{d}\tau = r^2\sin\theta\mathrm{d}r\mathrm{d}\theta\mathrm{d}\phi$. 在柱坐标系中 $(\mathrm{d}r)_\rho = \mathrm{d}\rho$, $(\mathrm{d}r)_\phi = \rho\mathrm{d}\phi$, $(\mathrm{d}r)_z = \mathrm{d}z$；由它们形成的体积元是 $\mathrm{d}\tau = \rho\mathrm{d}\rho\mathrm{d}\phi\mathrm{d}z$.

利用(Ⅱ.41)式得到

$$\mathrm{grad}\,\varphi = \sum_{\alpha=1}^{3}\left(\frac{\partial\varphi}{\partial r}\right)_\alpha e_\alpha = \sum_{\alpha=1}^{3}\frac{1}{H_\alpha}\frac{\partial\varphi}{\partial q_\alpha}e_\alpha. \tag{Ⅱ.42}$$

对于球坐标系和柱坐标系分别有

$$\mathrm{grad}\,\varphi = \frac{\partial\varphi}{\partial r}e_r + \frac{1}{r}\frac{\partial\varphi}{\partial\theta}e_\theta + \frac{1}{r\sin\theta}\frac{\partial\varphi}{\partial\phi}e_\phi; \tag{Ⅱ.43}$$

$$\mathrm{grad}\,\varphi = \frac{\partial\varphi}{\partial\rho}e_\rho + \frac{1}{\rho}\frac{\partial\varphi}{\partial\phi}e_\phi + \frac{\partial\varphi}{\partial z}e_z. \tag{Ⅱ.44}$$

散度按定义(Ⅱ.8)计算. 在局部标架中作平行六面体，类似于高登坐标中的(Ⅱ.10)式，有

$$-J_1(q_1,q_2,q_3)H_2H_3\mathrm{d}q_2\mathrm{d}q_3 \quad 和 \quad J_1(q_1+\mathrm{d}q_1,q_2,q_3)H_2H_3\mathrm{d}q_2\mathrm{d}q_3.$$

注意，H_2H_3 也是 q_α 的函数，在 q_1 和 $q_1+\mathrm{d}q_1$ 取不同值. 将这两项相加，类似于(Ⅱ.11)得到

$$[(J_1H_2H_3)_{q_1+\mathrm{d}q_1} - (J_1H_2H_3)_{q_1}]\mathrm{d}q_2\mathrm{d}q_3 = \frac{\partial(J_1H_2H_3)}{\partial q_1}\mathrm{d}q_1\mathrm{d}q_2\mathrm{d}q_3.$$

用 $\mathrm{d}V = H_1H_2H_3\mathrm{d}q_1\mathrm{d}q_2\mathrm{d}q_3$ 除，得到 $(H_1H_2H_3)^{-1}\partial(J_1H_2H_3)/\partial q_1$. 再和 $\mathrm{d}q_3\mathrm{d}q_1$ 面以及 $\mathrm{d}q_1\mathrm{d}q_2$ 面的类似项相加，得到

$$\mathrm{div}\,\boldsymbol{J} = \frac{1}{H_1H_2H_3}\left\{\frac{\partial(J_1H_2H_3)}{\partial q_1} + \frac{\partial(J_2H_3H_1)}{\partial q_2} + \frac{\partial(J_3H_1H_2)}{\partial q_3}\right\}. \tag{Ⅱ.45}$$

类似地可以得到

$$\text{rot } \boldsymbol{A} = \sum_{\alpha\beta\gamma=1}^{3} \varepsilon_{\alpha\beta\gamma}\left[\boldsymbol{e}_\alpha \frac{1}{H_\beta H_\gamma}\frac{\partial(A_\gamma H_\gamma)}{\partial x_\beta}\right]. \tag{Ⅱ.46}$$

利用(Ⅱ.42),(Ⅱ.45)得

$$\triangle\varphi = \text{div grad } \varphi = \frac{1}{H_1H_2H_3}\left\{\frac{\partial}{\partial q_1}\left(\frac{H_2H_3}{H_1}\frac{\partial\varphi}{\partial q_1}\right)+\frac{\partial}{\partial q_2}\left(\frac{H_3H_1}{H_2}\frac{\partial\varphi}{\partial q_2}\right)+\frac{\partial}{\partial q_3}\left(\frac{H_1H_2}{H_3}\frac{\partial\varphi}{\partial q_3}\right)\right\}. \tag{Ⅱ.47}$$

对于球坐标和柱坐标分别有:

球坐标:

$$\text{div } \boldsymbol{J} = \frac{1}{r^2}\frac{\partial(r^2J_r)}{\partial r}+\frac{1}{r\sin\theta}\left[\frac{\partial(\sin\theta J_\theta)}{\partial\theta}+\frac{\partial(J_\phi)}{\partial\phi}\right]. \tag{Ⅱ.48}$$

$$\begin{cases}(\text{rot } \boldsymbol{A})_r = \dfrac{1}{r\sin\theta}\left[\dfrac{\partial(\sin\theta A_\phi)}{\partial\theta}-\dfrac{\partial A_\theta}{\partial\phi}\right],\\[2ex](\text{rot } \boldsymbol{A})_\theta = \dfrac{1}{r\sin\theta}\dfrac{\partial A_r}{\partial\phi}-\dfrac{1}{r}\dfrac{\partial(rA_\phi)}{\partial r},\\[2ex](\text{rot } \boldsymbol{A})_\phi = \dfrac{1}{r}\left[\dfrac{\partial(rA_\theta)}{\partial r}-\dfrac{\partial A_r}{\partial\theta}\right].\end{cases} \tag{Ⅱ.49}$$

$$\triangle\varphi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\varphi}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\varphi}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2\varphi}{\partial\phi^2}. \tag{Ⅱ.50}$$

柱坐标:

$$\text{div } \boldsymbol{J} = \frac{1}{\rho}\frac{\partial(\rho J_\rho)}{\partial\rho}+\frac{1}{\rho}\frac{\partial J_\phi}{\partial\phi}+\frac{\partial J_z}{\partial z}. \tag{Ⅱ.51}$$

$$\begin{cases}(\text{rot } \boldsymbol{A})_\rho = \dfrac{1}{\rho}\dfrac{\partial A_z}{\partial\phi}-\dfrac{\partial A_\phi}{\partial z},\\[2ex](\text{rot } \boldsymbol{A})_\phi = \dfrac{\partial A_\rho}{\partial z}-\dfrac{\partial A_z}{\partial\rho},\\[2ex](\text{rot } \boldsymbol{A})_z = \dfrac{1}{\rho}\left[\dfrac{\partial(\rho A_\phi)}{\partial\rho}-\dfrac{\partial A_\rho}{\partial\phi}\right].\end{cases} \tag{Ⅱ.52}$$

$$\triangle\varphi = \frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial\varphi}{\partial\rho}\right)+\frac{1}{\rho^2}\frac{\partial^2\varphi}{\partial\phi^2}+\frac{\partial^2\varphi}{\partial z^2}. \tag{Ⅱ.53}$$

(五)张量计算简介

张量是在坐标变换时有一定变换性质的量.最简单的是在坐标变换下保持不变的量,如质量 m,电量 Q.这种量称为零阶张量,又称为标量.矢量也是一种张量,称为一阶张量.矢量(一阶张量)可以用它的分量 $A_i(i=1,2,3)$ 表示,也可以合写为 $\boldsymbol{A}=\sum_{i=1}^{3}A_i\boldsymbol{e}_i$.由于它的分量带有一个下标,所以称为一阶张量.

作为二阶张量的例子,考虑非各向同性导电介质.在这种介质中的某一点加上电场 $\boldsymbol{E}$,产生电流密度 $\boldsymbol{j}$.由于介质各向异性,$\boldsymbol{j}$ 和 $\boldsymbol{E}$ 一般说不平行(即:电流一般不平行于场),所以不能用一个简单的比例系数表明它们的关系,而只能写成一个函数关系:$\boldsymbol{j}=\boldsymbol{f}(\boldsymbol{E})$.当 $\boldsymbol{E}$ 很小时,可以将这一函数对 $\boldsymbol{E}$ 作线性展开,用分量写出就是

$$\left.\begin{aligned} j_1 &= T_{11}E_1 + T_{12}E_2 + T_{13}E_3, \\ j_2 &= T_{21}E_1 + T_{22}E_2 + T_{23}E_3, \\ j_3 &= T_{31}E_1 + T_{32}E_2 + T_{33}E_3. \end{aligned}\right\} \tag{Ⅱ.54}$$

或者合写为

$$j_i = \sum_{j=1}^{3} T_{ij}E_j. \tag{Ⅱ.55}$$

具有两个下标的 T_{ij} 建立了矢量 E_j 和 j_i 之间的线性关系，表征了非各向同性介质在这一点的电导度. 这样的量称为二阶张量，它有 $3^2=9$ 个分量.

正像一阶张量 A_i 可以用整体符号表示为 $\boldsymbol{A} = \sum_{i=1}^{3} A_i \boldsymbol{e}_i$ 一样，二阶张量也可以表示为

$$\vec{\vec{T}} = \sum_{ij=1}^{3} T_{ij}\boldsymbol{e}_i\boldsymbol{e}_j. \tag{Ⅱ.56}$$

这里，两个基矢 $\boldsymbol{e}_i$ 和 $\boldsymbol{e}_j$ 并列，不代表它们的点乘或叉乘，而保持各自独立的矢量意义. 这称为并矢. 利用这种符号，可将(Ⅱ.55)写成

$$\boldsymbol{j} = \vec{\vec{T}} \cdot \boldsymbol{E}. \tag{Ⅱ.57}$$

这里的中圆点表示 $\vec{\vec{T}}$ 中的第二个基矢和 $\boldsymbol{E}$ 中所包含的基矢点乘：

$$\vec{\vec{T}} \cdot \boldsymbol{E} = \sum_{ijk}^{3} T_{ij}E_k\boldsymbol{e}_i(\boldsymbol{e}_j \cdot \boldsymbol{e}_k) = \sum T_{ij}E_k\boldsymbol{e}_i\delta_{jk} = \sum T_{ij}E_j\boldsymbol{e}_i. \tag{Ⅱ.58}$$

代入(Ⅱ.57)式的右边，并注意 $\boldsymbol{j} = \sum j_i\boldsymbol{e}_i$，就得到(Ⅱ.55)式，这就证明了(Ⅱ.55)式和(Ⅱ.57)式一致.

将(Ⅱ.56)式中的哑标改写成 $i'j'$，再用 $\boldsymbol{e}_i$ 和 $\boldsymbol{e}_j$ 分别从左边和右边点乘，得到

$$\boldsymbol{e}_i \cdot \vec{\vec{T}} \cdot \boldsymbol{e}_j = \sum_{i'j'} T_{i'j'}(\boldsymbol{e}_i \cdot \boldsymbol{e}_{i'})(\boldsymbol{e}_{j'} \cdot \boldsymbol{e}_j) = \sum_{i'j'} T_{i'j'}\delta_{ii'}\delta_{jj'} = T_{ij}. \tag{Ⅱ.59}$$

这是由二阶张量的缩写形式(并矢形式) $\vec{\vec{T}}$ 得到它的分量 T_{ij} 的公式. 它类似于一阶张量的公式 $\boldsymbol{e}_i \cdot \boldsymbol{A} = A_i$.

由任意两个矢量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的并矢可以形成一个二阶张量

$$\boldsymbol{AB} = \sum_{ij} A_iB_j\boldsymbol{e}_i\boldsymbol{e}_j. \tag{Ⅱ.60}$$

注意，当 $\boldsymbol{A} \neq \boldsymbol{B}$ 时，并矢 $\boldsymbol{AB} \neq \boldsymbol{BA}$.

单位二阶张量定义为

$$\vec{\vec{e}} = \boldsymbol{e}_x\boldsymbol{e}_x + \boldsymbol{e}_y\boldsymbol{e}_y + \boldsymbol{e}_z\boldsymbol{e}_z = \sum_{ij=1}^{3} \delta_{ij}\boldsymbol{e}_i\boldsymbol{e}_j. \tag{Ⅱ.61}$$

它的分量是 δ_{ij}，它和任意矢量的点积等于这一矢量本身：$\vec{\vec{e}} \cdot \boldsymbol{A} = \boldsymbol{A} \cdot \vec{\vec{e}} = \boldsymbol{A}$.

如果 $T_{ij} = T_{ji}$，则称为对称张量. 对称张量和任意矢量的点积与次序无关：

$$\vec{\vec{T}} \cdot \boldsymbol{f} = \boldsymbol{f} \cdot \vec{\vec{T}} \quad \text{如果 } T_{ij} = T_{ji}. \tag{Ⅱ.62}$$

证 $\vec{\vec{T}} \cdot \boldsymbol{f} = \sum \boldsymbol{e}_i T_{ij} f_j = \sum \boldsymbol{e}_i T_{ji} f_j$，再交换哑标的符号 $i \rightleftharpoons j$，得到 $\vec{\vec{T}} \cdot \boldsymbol{f} = \boldsymbol{f} \cdot \vec{\vec{T}}$.

【证毕】

两个并矢的二次点乘定义为：

$$\vec{\vec{T}} : \vec{\vec{U}} = \sum T_{ij}\boldsymbol{e}_i\boldsymbol{e}_j : \boldsymbol{e}_l\boldsymbol{e}_m U_{lm} = \sum T_{ij}(\boldsymbol{e}_j \cdot \boldsymbol{e}_l)(\boldsymbol{e}_i \cdot \boldsymbol{e}_m)U_{lm} = \sum_m T_{ij}U_{jl}. \tag{Ⅱ.63}$$

二阶张量场 $\vec{\vec{T}}$ 的散度是一阶张量场(矢量场)：

$$\operatorname{div}\vec{\vec{T}} = \nabla \cdot \vec{\vec{T}} = \sum_i \boldsymbol{e}_i \frac{\partial}{\partial x_i} \cdot T_{jk}\boldsymbol{e}_j\boldsymbol{e}_k = \sum \frac{\partial T_{jk}}{\partial x_j}\boldsymbol{e}_k. \tag{Ⅱ.64}$$

高斯定理是：

$$\int_V \nabla \cdot \vec{\vec{T}}\,\mathrm{d}V = \oint_S \mathrm{d}\boldsymbol{\sigma} \cdot \vec{\vec{T}}. \tag{Ⅱ.65}$$

(六)三维 δ 函数的一个表达式

利用高斯定理可以证明三维 δ 函数的一个表达式[即式(Ⅰ.21)]：

$$\delta^3(\boldsymbol{r}) = -\frac{1}{4\pi}\Delta\left(\frac{1}{r}\right). \tag{Ⅱ.66}$$

证　令 V 为以原点为心的一个球，则

$$\int_V \Delta\left(\frac{1}{r}\right)\mathrm{d}^3\boldsymbol{r} = \int_V \operatorname{div}\operatorname{grad}\left(\frac{1}{r}\right)\mathrm{d}^3\boldsymbol{r} = \oint_S \operatorname{grad}\left(\frac{1}{r}\right)\cdot \mathrm{d}\boldsymbol{\sigma},$$

其中 S 为包围 V 的球面，用 $\boldsymbol{e}_r$ 表示 $\boldsymbol{r}$ 方向的单位矢量，则

$$\operatorname{grad}\left(\frac{1}{r}\right) = -\boldsymbol{e}_r/r^2, \qquad \mathrm{d}\boldsymbol{\sigma} = \boldsymbol{e}_r r^2\mathrm{d}\Omega,$$

$$\int_V \Delta\left(\frac{1}{r}\right)\mathrm{d}^3\boldsymbol{r} = \oint_S\left(-\frac{\boldsymbol{e}_r}{r^2}\right)\cdot \boldsymbol{e}_r r^2\mathrm{d}\Omega = \oint_S(-\mathrm{d}\Omega) = -4\pi.$$

利用(Ⅱ.50)得到

$$\text{当 } r \neq 0 \text{ 时} \quad \Delta\left(\frac{1}{r}\right) = \frac{1}{r^2}\frac{\partial}{\partial r}\left[r^2\frac{\partial}{\partial r}\left(\frac{1}{r}\right)\right] = 0. \tag{Ⅱ.67}$$

故上式可写为

$$-\frac{1}{4\pi}\iiint_{-\infty}^{\infty}\Delta\left(\frac{1}{r}\right)\mathrm{d}^3\boldsymbol{r} = 1. \tag{Ⅱ.68}$$

将(Ⅱ.67)、(Ⅱ.68)式和(Ⅰ.17)、(Ⅰ.18)式比较，就证明了(Ⅱ.66)．　**【证毕】**

附录Ⅲ　概率的基础知识

(一)随机事件和概率

1. 随机事件

必然事件:在一组条件实现之下,必然发生的事件叫必然事件.例如,在一大气压下,水温达100℃时,水必然沸腾.

随机事件:在一组条件实现之下,可能发生的事件叫随机事件.如在掷骰子的游戏中,"3"点出现是一个随机事件.

按照几个随机事件之间的关系,往往可分为相斥事件、对立事件和独立事件.

相斥事件:在一组条件实现之下,不能同时发生的事件叫相斥事件或不相容事件,如掷骰子,"3"点和"5"点出现是相斥事件.

对立事件:若两个相斥事件中一定有一个发生,则称此两事件为对立事件.例如掷硬币时,"字"出现和"花"出现是对立事件.对立事件是相斥事件的特殊情况.

独立事件:若两个随机事件发生与否互相独立,则称为独立事件.如同时掷出两颗骰子,则第一颗骰子出现"3"点和第二颗骰子出现"5"点是独立事件.

2. 随机事件的概率和概率的基本公式

为表示随机事件的可能程度引入概率的概念.大量实践表明,在一定的条件下进行一次实验时,实验的某一可能结果(事件A)是否一定发生无法预言.但当实验次数N趋于无穷时,事件A发生的次数$N(A)$与总次数N的比值却趋于稳定的极限值.这一极限值叫做事件A的概率:

$$P(A) = \lim_{N\to\infty} N(A)/N. \tag{Ⅲ.1}$$

由定义可知,$0 \leqslant P(A) \leqslant 1$.

投掷骰子时,有6个可能的结果.设分别用A_1和A_2代表"1"点出现和"2"点出现的事件,它们是相斥事件.把"1"点出现或者"2"点出现算作事件A,称为事件A_1和事件A_2的并,用$A = A_1 \cup A_2$表示.而它们出现的次数分别用$N(A_1)$,$N(A_2)$和$N(A) = N(A_1 \cup A_2)$表示.显然有$N(A_1 \cup A_2) = N(A_1) + N(A_2)$.将事件$A$发生的概率表示为$P(A)$.根据概率的定义,显然有

$$P(A) = P(A_1 \cup A_2) = \lim_{N\to\infty}[N(A_1) + N(A_2)]/N = P(A_1) + P(A_2). \tag{Ⅲ.2}$$

即:两个相斥事件中任出现一个的概率等于两事件概率之和.这称为概率的加法.不难看出,对于任意n个互相排斥的事件$A_1, A_2, \cdots, A_n$的"并",$A = A_1 \cup A_2 \cup \cdots \cup A_n$,有

$$P(A) = P(A_1) + P(A_2) + \cdots + P(A_n). \tag{Ⅲ.3}$$

若$A_1, A_2, \cdots, A_n$表示一次实验中的全部可能结果,如掷骰子时的6种可能情况,则必然有

$$\sum_{i=1}^{n} P(A_i) = 1. \tag{Ⅲ.4}$$

(Ⅲ.4)式表示全部排斥事件中,必有一个出现,因而任出现一个的概率为1,叫做归一化条件.

再考虑同时掷出两颗骰子的游戏,问第一颗骰子出现"1"点(事件A),同时第二颗骰子出现"2"点(事件B)的联合概率是多少?我们把事件A出现并且事件B出现称为事件A和B的

交，用 $C = A \cap B$ 表示，根据定义有：

$$P(C) = P(A \cap B) = \lim_{N\to\infty}[N(A \cap B)]/N.$$

在 N 次实验中，第一颗骰子出现事件 A 的次数为 $N(A) = NP(A)$，在 $N(A)$ 次实验中，第二颗骰子出现 B 的次数为 $N(A \cap B) = N(A)P(B|A)$，其中 $P(B|A)$ 表示在事件 A 出现之下，B 事件发生的概率（称为条件概率）. 若事件 A 和事件 B 毫无联系而彼此独立，因而 B 事件出现与否和附加不附加事件 A 出现为条件无关，则应有 $P(B|A) = P(B)$. 于是有 $N(A \cap B) = N(A)P(B) = NP(A)P(B)$，或

$$P(A \cap B) = P(A)P(B). \tag{Ⅲ.5}$$

(Ⅲ.5)式表示两个独立事件联合发生的概率等于两独立事件出现概率的乘积. 这称为概率的乘法公式. 反之，若一复合事件发生的概率等于两个事件概率相乘，则可以推断，此两事件必然是独立事件. 显然(Ⅲ.5)式可推广到任意多个独立事件.

（二）　随机事件的概率分布　统计平均值　涨落

首先讨论单个随机事件的情况.

1. 单个随机变量的概率分布　平均值和涨落

在一定条件下，能够以确定的概率取一系列不同值的变量叫随机变量. 随机变量可分为离散型和连续型两类. 离散型随机变量的可能取值是数轴上可数的分立值. 例如半导体导带中的电子数，生物群体中的个体数等. 用 X 表示随机变量，$x_1, x_2, \cdots, x_i$ 表示随机变量的可能取值. 用 $P(x_i)$ 表示随机变量 $X = x_i$ 时的概率，设已知随机变量 X 的取值范围及相应的概率为

$$\left.\begin{matrix} x_1 & x_2 & \cdots & x_i \\ P(x_1) & P(x_2) & \cdots & P(x_i) \end{matrix}\right\},$$

则称随机变量 X 的概率分布已知. 随机变量的概率必须满足以下条件：

$$P(x_i) \geqslant 0, \qquad \sum_{i=1}^{i} P(x_i) = 1. \tag{Ⅲ.6}$$

连续型随机变量可以取数轴上某一区间内的一切数值. 例如气体分子的速度分量取值在 $(-\infty, \infty)$. 气体分子的平动动能取值在 $(0, \infty)$. 连续型随机变量的概率分布由概率密度 $\rho(x)$ 描述. 随机变量 X 取值在 $x \sim x + \mathrm{d}x$ 间隔内的概率 $\mathrm{d}P(x)$ 可写为

$$\mathrm{d}P(x) = \rho(x)\mathrm{d}x, \tag{Ⅲ.7}$$

其中 $\rho(x)$ 为变量 X 取值在 x 处附近单位间隔内的概率，叫做概率密度或分布函数. 与(Ⅲ.6)式对应有

$$\rho(x) \geqslant 0, \qquad \int\rho(x)\mathrm{d}x = 1. \tag{Ⅲ.8}$$

积分遍及随机变量 X 的取值范围.

已知随机变量的概率分布，可计算表征该随机变量特征的统计平均值. 对于离散型的随机变量 X，其可能取值为 $x_1, x_2, \cdots, x_i$. 在总实验次数 N 中设测得这些值的次数分别为 $N_1, N_2, \cdots, N_i$，则 X 的算术平均值为

$$(x_1N_1 + x_2N_2 + \cdots + x_iN_i)/N = \left(\sum_{i=1}^{i} x_iN_i\right)/N.$$

当实验总次数无限增大时，X 的平均值的极限叫做 X 的统计平均值，即

$$\bar{X} = \lim_{N\to\infty}\sum_{i=1}^{s}\frac{x_iN_i}{N} = \sum_{i=1}^{s}x_i\lim_{N\to\infty}\frac{N_i}{N} = \sum_{i=1}^{s}x_iP(x_i). \tag{Ⅲ.9}$$

若随机变量是连续的，则可同理推出

$$\bar{X} = \int x\rho(x)\,dx, \tag{Ⅲ.10}$$

积分遍及 x 的取值范围. 显然(Ⅲ.9)和(Ⅲ.10)式对随机变量的任意函数 $f(X)$ 也是适用的.

统计平均值集中地反映了随机变量的统计特征，它又称为期望值，是一次测量中可以期望会出现的结果. 一般说来，每次实验结果并不等于平均值，而是在平均值上下波动. 现在要引入一个量来描述 X 在平均值 $\bar{X}$ 上下波动的平均幅度. 因为 $\overline{\Delta X} = \overline{X_i - \bar{X}} = \sum(x_i - \bar{X})P(x_i) = \sum x_iP(x_i) - \bar{X}\sum P(x_i) = \bar{X} - \bar{X} = 0$. 所以不能用它量度涨落的平均幅度. $\overline{(X_i - \bar{X})^2}$ 恒为正，故可用来量度 X 在平均值上下波动的平均幅度，称为均方涨落，或方差. 常用符号 σ^2 表示：

$$\begin{aligned}\sigma^2 &= (\overline{\Delta X})^2 = \overline{(x_i - \bar{X})^2} = \sum(x_i^2 - 2x_i\bar{X} + \bar{X}^2)P(x_i)\\ &= \sum(x_i^2P(x_i)) - 2\bar{X}\sum x_iP(x_i) + \bar{X}^2\sum P(x_i) = \overline{X^2} - \bar{X}^2.\end{aligned} \tag{Ⅲ.11}$$

σ 称为标准偏差. 对于连续型随机变量，则有

$$\sigma^2 = \overline{X^2} - \bar{X}^2 = \int x^2\rho(x)\,dx - \left(\int x\rho(x)\,dx\right)^2. \tag{Ⅲ.12}$$

随机变量涨落的程度由 σ 和 $\bar{X}$ 之比表示，称为相对涨落：

$$\delta X = \sqrt{\overline{(\Delta X)^2}}/\bar{X} = \sigma/\bar{X}. \tag{Ⅲ.13}$$

2. 随机变量的联合概率　相关矩

以上讨论了单个随机变量的概率分布问题，现在研究若干个随机变量的概率分布规律. 为简单起见，只讨论两个随机变量，但不难推广到多个随机变量的情形.

设有两个随机变量 X 和 Y，用 $dP(x,y)$ 表示随机变量 X 取值在 $x \sim x+dx$，随机变量 Y 取值在 $y \sim y+dy$ 的联合概率，则有

$$dP(x,y) = \rho(x,y)\,dx\,dy, \tag{Ⅲ.14}$$

$\rho(x,y)$ 叫做联合概率密度，它满足以下条件，

$$\rho(x,y) \geqslant 0, \qquad \iint\rho(x,y)\,dx\,dy = 1. \tag{Ⅲ.15}$$

积分遍及 X,Y 的取值范围.

如果已知联合概率密度 $\rho(x,y)$，则可以通过对 y 的积分而得到

$$\rho(x) = \int\rho(x,y)\,dy. \tag{Ⅲ.16}$$

$\rho(x)$ 叫做 X 的边缘概率.

若两个随机变量 X 和 Y 统计独立，则根据概率乘法公式有

$$\rho(x,y) = \rho(x)\rho(y). \tag{Ⅲ.17}$$

利用(Ⅲ.17)可证明两个常用公式. 因 $\overline{XY} = \iint xy\rho(x,y)\,dx\,dy$，若 X,Y 统计独立，则有

$$\overline{XY} = \int x\rho(x)\mathrm{d}x \int y\rho(y)\mathrm{d}y = X \cdot Y, \tag{Ⅲ.18}$$

即：两个独立随机变量乘积的平均值等于每个随机变量平均值的乘积.

若 $Z = X + Y$，且 X, Y 是独立随机变量，则有

$$\Delta Z \equiv Z - \bar{Z} = (X + Y) - (\bar{X} + Y) = \Delta X + \Delta Y,$$

$$\overline{(\Delta Z)^2} = \overline{(Z - \bar{Z})^2} = \overline{(\Delta X)^2} + 2\overline{(\Delta X)(\Delta Y)} + \overline{(\Delta Y)^2}.$$

因 X, Y 统计独立，故 $\overline{(\Delta X)(\Delta Y)} = \overline{(\Delta X)}\,\overline{(\Delta Y)} = 0$，所以最后有

$$\overline{(\Delta Z)^2} = \overline{(\Delta X)^2} + \overline{(\Delta Y)^2}, \quad \text{或} \quad \sigma_Z^2 = \sigma_X^2 + \sigma_Y^2. \tag{Ⅲ.19}$$

即：两个独立随机变量之和的均方涨落等于各变量均方涨落之和.

一般说来，两个随机变量不一定独立而有复杂的关联. 为了描述其关联程度，引入相关矩或协方差，其定义为

$$\mathrm{cov}(X,Y) = \overline{(X - \bar{X})(Y - \bar{Y})} = \overline{XY} - \bar{X} \cdot \bar{Y}. \tag{Ⅲ.20}$$

若 X, Y 统计独立，则据(Ⅲ.18)有 $\mathrm{cov}(X,Y) = 0$. 反之若 X, Y 的相关矩为零，即表示两随机变量没有关联；若相关矩不为零，则表示两随机变量之间存在关联.

〔三〕 二项式分布、泊松分布和高斯分布

有许多特殊的概率分布，但应用最广的是二项式分布及其在两种极限情况下的近似表示——泊松分布和高斯分布.

1. 二项式分布

让我们相继地进行 N 次统计上独立的实验，设每次实验只能出现甲、乙两种结果之一. 令甲结果出现的概率为 p，则乙结果出现的概率为 $q = 1 - p$. 在 N 次实验中，甲结果出现 n 次，乙结果出现 $N - n$ 次的概率为

$$P_N(n) = \frac{N!}{n!(N-n)!} p^n q^{N-n}. \tag{Ⅲ.21}$$

此即二项式分布.

证　以“醉汉荡步”为例来证明.

醉汉在一直线上荡步，步与步之间没有关联. 每一步向前的概率为 p，向后的概率为 q，$p + q = 1$. 求在荡了 N 步后，其中有 n 步向前的概率是多少？

为了计算这一复杂事件的概率. 先考虑某一特定走法的概率，所谓特定走法，是指第一步向前，第二向前，第三步向后…的一种具体走法. 根据独立事件概率相乘公式，这一特定走法的概率为：

$$p \cdot p \cdot q \cdots = p^n q^{N-n}.$$

显然，上式对任何一种特定的向前走 n 步向后走 $N - n$ 步的特定走法皆适合. 每一种特定走法之间是互相排斥的. 但我们要求的不是某一特定走法的概率，换句话说，要求的不是某一确定排列的概率，而是 N 步中有 n 步向前，$N - n$ 步向后的组合的概率. 组合数为 $C_N^n = N!/n!(N-n)!$. 根据排斥事件概率相加的原理，组合的概率为(Ⅲ.21)式.　**【证毕】**

容易证明二项式分布满足归一化条件：

$$\sum_{n=0}^{N} P_N(n) = \sum_{n=0}^{N} \frac{N!}{n!(N-n)!} p^n q^{N-n} = (p+q)^N = 1. \tag{Ⅲ.22}$$

据(Ⅲ.9)式可求得 n 的平均值

$$\begin{aligned} \bar{n} &= \sum_{n=0}^{N} nP_N(n) = \sum_{n=0}^{N} \frac{N!n}{n!(N-n)!} p^n q^{N-n} \\ &= p\frac{\partial}{\partial p}\left[\sum_{n=0}^{N} P_N(n)\right] = p\frac{\partial}{\partial p}(p+q)^N = Np. \end{aligned} \tag{Ⅲ.23}$$

与此类似可求得

$$\overline{n^2} = \sum_{n=0}^{N} n^2 P_N(n) = (Np)^2 + Npq.$$

或

$$\sigma^2 = \overline{(\Delta n)^2} = \overline{n^2} - (\bar{n})^2 = Npq, \tag{Ⅲ.24}$$

$$\delta n = \sqrt{\overline{(\Delta n)^2}/\bar{n}} = \sqrt{Npq}/Np = \sqrt{q/Np}. \tag{Ⅲ.25}$$

2. 泊松分布

泊松分布是 $N \gg 1$ 和 $p \ll q$ 时二项式分布的近似表达式. 因 $p \ll 1$,所以当 n 很大时,p^n 迅速减小以至趋于零. 换句话说,只有在 $n \ll N$ 时,$P_N(n)$ 才有意义. 我们考察这一范围内(Ⅲ.21)式的近似表达式. 首先将(Ⅲ.21)式改写为

$$P_N(n) = \frac{1}{n!} N(N-1)(N-2)\cdots(N-n+1) p^n q^{N-n}. \tag{Ⅲ.26}$$

因 $N \gg 1$ 和 $n \ll N$,故 $(N-1),(N-2),\cdots,(N-n+1)$ 中的每一个因子均可近似用 N 代替,即有

$$N(N-1)(N-2)\cdots(N-n+1) \approx N^n. \tag{Ⅲ.27}$$

又因 $q = 1-p$ 和 $p \ll 1$,利用泰勒展开,忽略高次项可得

$$\ln q^{N-n} = (N-n)\ln(1-p) \approx -Np,$$

或

$$q^{N-n} \approx e^{-Np}. \tag{Ⅲ.28}$$

将(Ⅲ.27)和(Ⅲ.28)代入(Ⅲ.26)式得,$P_N(n) = (Np)^n e^{-Np}/n!$,注意 $Np = \bar{n}$,最后得

$$P_N(n) = (\bar{n})^n e^{-\bar{n}}/n!. \tag{Ⅲ.29}$$

(Ⅲ.29)式就是泊松分布,它适用于实验次数很多,而且结果出现的概率 p 很小的情形. 泊松分布的均方涨落为

$$\sigma^2 = \overline{(\Delta n)^2} = Npq = \bar{n}. \tag{Ⅲ.30}$$

当每次试验独立,试验次数很多,而某一结果出现的平均次数给定时,出现这一结果的次数就服从泊松分布. 这是因为,试验次数多即 $N \to \infty$,而平均次数 $\bar{n} = Np$ 给定,意味着 $p \to 0$.

3. 高斯分布

在实验中,为了提高测量的精度,常常重复进行多次测量,取其平均值. 这样做的根据是:

中心极限定理 设随机变量 X 具有平均值为 $\bar{X}$,方差为 σ_X^2 的概率分布 $P_X(X)$. 对 X 进行 N 次独立测量,取其平均值

$$Y = (X_1 + X_2 + \cdots + X_N)/N. \tag{Ⅲ.31}$$

重复以上步骤,得到一系列 Y.这些 Y 也是随机变量.其概率分布为

$$P_Y(Y) = \frac{1}{\sqrt{2\pi}\sigma_Y} e^{-(Y-\bar{Y})^2/(2\sigma_Y^2)}, \tag{Ⅲ.32}$$

它是归一化的

$$\int_{-\infty}^{\infty} P_Y(Y)\,dY = 1, \tag{Ⅲ.33}$$

Y 的平均值

$$\bar{Y} = \int_{-\infty}^{\infty} Y P_Y(Y)\,dY = \bar{X}. \tag{Ⅲ.34}$$

而方差

$$\sigma_Y^2 = Y^2 - (Y)^2 = \sigma_X^2/N. \tag{Ⅲ.35}$$

以上就是中心极限定理,分布(Ⅲ.32)式称为高斯分布,又叫做正则分布.注意,在上述定理中,并没有规定随机变量 X 的概率分布 $P_X(X)$ 的形式,而只要重复多次测量,平均值 Y 的分布就是高斯分布.因此,高斯分布是具有任意概率分布❶的随机变量 X 多次测量的平均值 Y 的概率分布.特别是,用 Y 来代替 X,平均值不变($Y = \bar{X}$),而方差减小了 N 倍.相应地,标准偏差(σ)和相对涨落(δ)减小了 $\sqrt{N}$ 倍[见(Ⅲ.35)、(Ⅲ.13)式].

证　令 $\varphi(k)$ 和 $\Phi(k)$ 分别为 $P_X(X-\bar{X})$ 和 $P_Y(Y-Y)$ 的付里叶变换:

$$\varphi(k) = \int e^{ik(X-\bar{X})} P_X(X-\bar{X})\,dX. \tag{Ⅲ.36}$$

$$\begin{aligned}\Phi(k) &= \int e^{ik(Y-\bar{Y})} P_Y(Y-\bar{Y})\,dY \\ &= \int e^{i(k/N)[(X_1-\bar{X}_1)+(X_2-\bar{X}_2)+\cdots+(X_N-\bar{X}_N)]} P_X(X_1)P_X(X_2)\cdots P_X(X_N)\,dX_1\,dX_2\cdots dX_N \\ &= [\varphi(k/N)]^N.\end{aligned} \tag{Ⅲ.37}$$

在第二步中用了 Y 的定义(Ⅲ.31)式及独立事件概率相乘的性质

$$P_Y(Y-\bar{Y})\,dY = P_X(X_1)P_X(X_2)\cdots P_X(X_N)\,dX_1\,dX_2\cdots dX_N$$

对 $\varphi(k/N)$ 中的指数函数作泰勒展开,

$$\begin{aligned}\varphi(k/N) &= \int e^{i(k/N)(X-\bar{X})} P_X(X-\bar{X})\,dX \\ &= \int (1 + i(k/N)(X-\bar{X}) - (1/2!)(k/N)^2(X-\bar{X})^2 + \cdots) P_X(X-\bar{X})\,dX \\ &= 1 - (1/2)(k^2/N^2)\sigma^2 + \cdots\end{aligned} \tag{Ⅲ.38}$$

这里利用了分布 $P_X(X-\bar{X})$ 的归一化及对 $X-\bar{X}$ 是偶函数(测量结果大于和小于平均值的几率相等)的性质.

如果

---

❶ 实际上,$P_X(X)$ 需要满足当 X 增加时下降足够快的条件,见(Ⅲ.39)式.但大多数情况下,这一条件能得到满足.

$$P_X(X) \text{ 在 } X \to \infty \text{ 时趋于零足够快,} \tag{Ⅲ.39}$$

则 P_X 的高阶矩是有限的,因而当 $N \to \infty$ 时,可以略去(Ⅲ.38)式中的高阶项.将(Ⅲ.38)式代入(Ⅲ.37)式得到

$$\Phi(k) = \lim_{N\to\infty}\left[1 - \frac{1}{2}\frac{k^2\sigma_X^2}{N^2} + o\left(\frac{k^3}{N^3}\right)\right]^N = \mathrm{e}^{-k^2\sigma_X^2/(2N)}. \tag{Ⅲ.40}$$

进行反付里叶变换,就得到

$$\begin{aligned} P_Y(Y-\bar{Y}) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}k\mathrm{e}^{-\mathrm{i}k(Y-\bar{Y})}\Phi(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}k\mathrm{e}^{-\mathrm{i}k(Y-\bar{Y})}\mathrm{e}^{-k^2\sigma_X^2/(2N)} \\ &= \sqrt{\frac{N}{2\pi}}\frac{1}{\sigma_X}\mathrm{e}^{-(Y-\bar{Y})^2/[2(\sigma_X^2/N)]} \\ &= \sqrt{\frac{1}{2\pi}}\frac{\sqrt{N}}{\sigma_X}\mathrm{e}^{-(Y-\bar{Y})^2/2(\sigma_X/\sqrt{N})^2}. \end{aligned} \tag{Ⅲ.41}$$

这里用了附录(V.2)式.这就证明了中心极限定理(Ⅲ.32)、(Ⅲ.35)式. 【证毕】

附录Ⅳ　拉普拉斯方程的通解

(一)方程的分离变量

将球坐标系中的拉普拉斯算符代入 $\nabla^2\Psi(R,\theta,\varphi)=0$,可得

$$\frac{1}{R}\frac{\partial}{\partial R}\left(R^2\frac{\partial\Psi}{\partial R}\right)+\frac{1}{R^2\sin\theta}\frac{\partial}{\partial\theta}\left((\sin\theta\frac{\partial\Psi}{\partial\theta}\right)+\frac{1}{R^2\sin^2\theta}\frac{\partial^2\Psi}{\partial\varphi^2}=0. \tag{Ⅳ.1}$$

函数 Ψ 应该是处处单值、有限,因而方程(Ⅳ.1)的解应满足

$$\Psi(R,\theta,\varphi)=\Psi(R,\theta,\varphi+2\pi), \tag{Ⅳ.2}$$

$$\Psi\text{有限}\qquad(0\leqslant\theta\leqslant\pi). \tag{Ⅳ.3}$$

对于 R 的条件,由区域的形状决定,根据不同的具体问题写出.例如,如果区域中包含 $R=0$ 的点,则有

$$\Psi\text{有限}\qquad(R=0). \tag{Ⅳ.4}$$

首先在一个特殊的形式下求方程(Ⅳ.1)的解,这种解可以写成表示距离的变量 R 的函数和表示方向的变量 θ,φ 的函数的乘积:

$$\Psi(R,\theta,\varphi)=f(R)Y(\theta,\varphi). \tag{Ⅳ.5}$$

代入(Ⅳ.1)式后,再用 $R^2/f(R)Y(\theta,\varphi)$ 遍乘各项,整理后可得

$$\frac{1}{f(R)}\frac{\mathrm{d}}{\mathrm{d}R}\left[R^2\frac{\mathrm{d}f(R)}{\mathrm{d}R}\right]=-\frac{1}{Y(\theta,\varphi)}\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left[\sin\theta\frac{\partial Y(\theta,\varphi)}{\partial\theta}\right]-\frac{1}{Y(\theta,\varphi)}\frac{1}{\sin^2\theta}\frac{\partial^2Y(\theta,\varphi)}{\partial\varphi^2},$$

方程的左边只是 R 的函数,与 θ,φ 无关;方程的右边只是 θ,φ 的函数,与 R 无关.既然 R,θ,φ 是独立变量,因此只有两边等于同一常数,上式才可能相等.为了今后的方便,将这个常数记为 $l(l+1)$,于是上述方程便分离为两个方程:

$$\frac{\mathrm{d}}{\mathrm{d}R}\left[R^2\frac{\mathrm{d}f(R)}{\mathrm{d}R}\right]-l(l+1)f(R)=0, \tag{Ⅳ.6}$$

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left[\sin\theta\frac{\partial Y(\theta,\varphi)}{\partial\theta}\right]-\frac{1}{\sin^2\theta}\frac{\partial^2Y(\theta,\varphi)}{\partial\varphi^2}+l(l+1)Y(\theta,\varphi)=0. \tag{Ⅳ.7}$$

方程(Ⅳ.6)的解为

$$f_l(R)=CR^l+D/R^{l+1}. \tag{Ⅳ.8}$$

可进一步将方程(Ⅳ.7)中的 θ 与 φ 分开,令

$$Y(\theta,\varphi)=\Theta(\theta)\Phi(\varphi), \tag{Ⅳ.9}$$

代入(Ⅳ.7)式,再用 $\sin^2\theta/\Theta(\theta)\Phi(\varphi)$ 遍乘各项,即有

$$\frac{\sin\theta}{\Theta(\theta)}\frac{\mathrm{d}}{\mathrm{d}\theta}\left[\sin\theta\frac{\mathrm{d}\Theta(\theta)}{\mathrm{d}\theta}\right]+l(l+1)\sin^2\theta=-\frac{1}{\Phi(\varphi)}\frac{\mathrm{d}^2\Phi(\varphi)}{\mathrm{d}\varphi^2}.$$

方程左边仅为 θ 的函数,与 φ 无关;方程的右边仅为 φ 的函数,与 θ 无关.既然 θ 与 φ 为独立变量,只有两边等于同一常数,上式才可能相等.将这常数记为 m^2,即可分解为两个方程:

$$\frac{\mathrm{d}^2\Phi(\varphi)}{\mathrm{d}\varphi^2}+m^2\Phi(\varphi)=0, \tag{Ⅳ.10}$$

$$\sin(\theta)\frac{\mathrm{d}}{\mathrm{d}\theta}\left[\sin\theta\frac{\mathrm{d}\Theta(\theta)}{\mathrm{d}\theta}\right]+[l(l+1)\sin^2\theta-m^2]\Theta(\theta)=0. \tag{Ⅳ.11}$$

由定解条件(Ⅳ.2)得到 Φ 所应满足的条件:

$$\Phi(\varphi+2\pi)=\Phi(\varphi).$$

在此条件下方程(Ⅳ.10)的解可以写成两种等效形式:

$$\Phi(\varphi)=A\cos m\varphi+B\sin m\varphi,\quad \text{其中}\quad m=0,\pm1,\pm2,\cdots, \tag{Ⅳ.12}$$

$$\Phi(\varphi)=Ce^{im\varphi}\quad \text{其中}\quad m=0,\pm1,\pm2,\cdots. \tag{Ⅳ.13}$$

现在求解方程(Ⅳ.11).为了简单起见,用记号 x 代替 $\cos\theta$,y 代替 Θ,将

$$x=\cos\theta,\quad y=\Theta \tag{Ⅳ.14}$$

代入(Ⅳ.11)式,得:

$$(1-x^2)\frac{\mathrm{d}^2y}{\mathrm{d}x^2}-2x\frac{\mathrm{d}y}{\mathrm{d}x}+\left[l(l+1)-\frac{m}{1-x^2}\right]y=0, \tag{Ⅳ.15}$$

称为 l 阶缔合勒让德方程.

如果球坐标系的极轴为对称轴,则 Φ 与 φ 无关,由(Ⅳ.12)式可见,这时 $m=0$,于是方程(Ⅳ.15)变为 l 阶勒让德方程

$$(1-x^2)\mathrm{d}^2y/\mathrm{d}x^2-2x\mathrm{d}y/\mathrm{d}x+l(l+1)y=0. \tag{Ⅳ.16}$$

(二)勒让德多项式

勒让德方程有幂级数解,令

$$y(x)=\sum_{k=0}^{\infty}C_kx^k. \tag{Ⅳ.17}$$

代入(Ⅳ.16)式,将 x 的零次项、一次项与高次项分开,得

$$\left\{2C_2+6C_3x+\sum_{k=0}^{\infty}[(k+2)(k+1)C_{k+2}-k(k-1)C_k]x^k\right\}$$

$$-\left\{2C_1x+\sum_{k=0}^{\infty}2kC_kx^k\right\}+\left\{l(l+1)C_0+l(l+1)C_1x+\sum_{k=0}^{\infty}l(l+1)C_kx^k\right\}=0.$$

令 x 的同次幂系数之和为零,便有

$$2C_2+l(l+1)C_0=0, \tag{Ⅳ.18}$$

$$6C_3+[l(l+1)-2]C_1=0, \tag{Ⅳ.19}$$

$$(k+2)(k+1)C_{k+2}-[k(k+1)-l(l+1)]C_k=0\qquad(k\geqslant2).$$

上式可改写为

$$C_{k+2}=\frac{k(k+1)-l(l+1)}{(k+2)(k+1)}C_k=\frac{(k-l)(k+l+1)}{(k+2)(k+1)}C_k. \tag{Ⅳ.20}$$

这就是系数的递推公式.它说明,幂级数(Ⅳ.17)中 x 的奇次幂和偶次幂的系数之间分别具有递推关系.这样,便可把 l 阶勒让德方程的解(Ⅳ.17)式表示为

$$y(x)=a_0y_0(x)+a_1y_1(x), \tag{Ⅳ.21}$$

其中 $y_0(x)$ 是 x 偶次幂的无穷级数,而 $y_1(x)$ 是 x 的奇次幂的无穷级数.

由(Ⅳ.20)式可见,如果 l 为某个固定的非负整数,$l=0,1,2,\cdots$,则 $C_{l+2}=0$.重复运用(Ⅳ.20)式,则依次有 $C_{l+4}=C_{l+6}=\cdots=0$.当 l 为偶数时,从 C_{l+2} 开始的一切偶次项为零,使得

x 的偶次幂级数 $y_0(x)$ 变为多项式，于是 $y(x)$ 成为多项式 $y_0(x)$ 与无穷级数 $y_1(x)$ 的线性组合. 同理，如果 l 为奇数，则从 C_{l+2} 开始一切奇数项为零，使得 x 的奇次幂级数 $y_1(x)$ 变为多项式，于是 $y(x)$ 是多项式 $y_1(x)$ 与无穷级数 $y_0(x)$ 的线性组合. 在 l 为偶数的情况下，令(Ⅳ.21)式中的 $a_1 = 0$，在 l 为奇数的情况下，令该式中的 $a_0 = 0$，就能得到勒让德方程的多项式解.

可以证明，级数(Ⅳ.17)在 $|x| < 1$ 收敛，而在 $x = 1$ 发散. 这意味着解在 $\theta = 0$ 及 π 方向会发散，违反了定解条件(Ⅳ.3). 发散的根源在于由(Ⅳ.20)式得到的解是无穷级数. 为了能满足定解条件(Ⅳ.3)，必须要无穷级数截断为多项式，而这只有当 l 为正整数或零时才有可能. 因此，方程(Ⅳ.6)、(Ⅳ.7)的系数受以下条件限制❶：

$$l = 0,1,2,\cdots. \tag{Ⅳ.22}$$

对于这样的 l，选取勒让德方程的解为 l 阶多项式：

$$y(x) = \begin{cases} a_0 y_0(x), & \text{当 } l \text{ 为偶数}, \\ a_1 y_1(x), & \text{当 } l \text{ 为奇数}. \end{cases} \tag{Ⅳ.23}$$

在系数应满足的关系式(Ⅳ.18)～(Ⅳ.20)式中有一个系数是可以任取的. 例如，当 l 为偶数时，可以任取 C_0. 而当 l 为奇数时，可以任取 C_1. 通常我们这样选取 C_0(或 C_1)，使得当 $x = 1$ 时，多项式等于 1. 这样选取系数的多项式(Ⅳ.23)称为勒让德多项式，记为 $P_l(x)$.

【例 1】 求 $l = 0,1,2,3$ 时的勒让德多项式.

解 当 $l = 0$ 时，由(Ⅳ.18)式得 $C_2 = 0$，由(Ⅳ.17)式得 $P_0(x) = C_0$，为使 $P_0(1) = 1$，取 $C_0 = 1$，故 $P_0(x) = 1$.

当 $l = 1$ 时，由(Ⅳ.19)式得 $C_3 = 0$，由(Ⅳ.17)式得 $P_1(x) = C_1 x$，为使 $P_1(1) = 1$，取 $C_1 = 1$，故 $P_1(x) = x$.

当 $l = 2$ 时，由(Ⅳ.18)式得 $C_2 = -3C_0$，由(Ⅳ.17)式得 $P_2(x) = C_0 - 3C_0 x^2$，为使 $P_2(1) = 1$，取 $C_0 = -1/2$，故 $P_2(x) = (3x^2 - 1)/2$.

当 $l = 3$ 时，由(Ⅳ.19)式 $C_3 = -5C_1/3$，由(Ⅳ.17)式 $P_3(x) = C_1 x - 5C_1 x^3/3$，为使 $P_3(1) = 1$，应该取 $C_1 = -3/2$，故 $P_3(x) = (5x^3 - 3x)/2$.

利用(Ⅳ.17)～(Ⅳ.19)式以及 $P_l(1) = 1$ 可推出 P_l 的具体表达式为：

$$P_l(x) = \sum_{k=0}^{[l/2]} (-1)^k \frac{(2l-2k)!}{2^l k!(l-k)!(l-2k)!} x^{l-2k}, \tag{Ⅳ.24}$$

其中

$$[l/2] = \begin{cases} l/2, & \text{当 } l \text{ 为偶数}, \\ (l-1)/2, & \text{当 } l \text{ 为奇数}. \end{cases} \tag{Ⅳ.25}$$

利用二项式定理，不难验证勒让德多项式可以用一个微分公式表示为

$$P_l(x) = [1/(2^l l!)]\mathrm{d}^l(x^2 - 1)^l/\mathrm{d}x^l. \tag{Ⅳ.26}$$

称为洛德利格斯公式.

利用(Ⅳ.26)式，可以方便地求出 $P_l(x)$，例如：

---

❶ 附带指出，这就是(Ⅳ.6)、(Ⅳ.7)式中写系数为 $l(l+1)$ 的原因. 如果将系数写成其他形式，则对系数的限制条件会变得很复杂.

$$P_4(x) = \frac{1}{2^4 4!}\frac{d^4}{dx^4}(x^2-1)^4 = \frac{1}{8}(35x^4 - 30x^2 + 3),$$

余类推.

后面还要用勒让德多项式的一些公式,这里不一一推导.其中比较常用的是勒让德多项式的正交性和递推公式:

$$\int_{-1}^{+1} P_l(x)P_k(x)dx = \frac{2}{2l+1}\delta_{lk}, \tag{Ⅳ.27}$$

$$(l+1)P_{l+1}(x) - (2l+1)xP_l(x) + lP_{l-1}(x) = 0. \tag{Ⅳ.28}$$

(三)缔合勒让德函数

l 阶缔合勒让德方程(Ⅳ.15)的解,可以证明为,

$$P_l^{|m|}(x) = (1-x^2)^{|m|/2}\frac{d^{|m|}}{dx^{|m|}}P_l(x), \qquad (m = 0, \pm 1, \cdots, \pm l). \tag{Ⅳ.29}$$

将 $x = \cos\theta$ 代入即为方程(Ⅳ.11)的解. $P_l^{|m|}(x)$ 称为 l 阶缔合勒让德函数.显然,$|m| \leqslant l$,否则(Ⅳ.29)式为零.

缔合勒让德函数的正交性可表示为:

$$\int_{-1}^{+1} P_l^{|m|}(x)P_k^{|m|}(x)dx = \frac{(l+|m|)!}{(l-|m|)!}\frac{2}{2l+1}\delta_{lk}. \tag{Ⅳ.30}$$

(四)球谐函数

现将方程(Ⅳ.10)的解(Ⅳ.13)式与方程(Ⅳ.11)的解(Ⅳ.29)式代入(Ⅳ.9)式,即得:

$$Y_{lm}(\theta,\varphi) = \sqrt{(l-|m|)!(2l+1)/[4\pi(l+|m|)!]}\,P_l^{|m|}(\cos\theta)e^{im\varphi}, \tag{Ⅳ.31}$$

$$l = 0,1,2,\cdots, \qquad m = 0, \pm 1, \pm 2, \cdots, \pm l.$$

它称为球谐函数,式中的系数由归一化条件定出.

球谐函数的正交归一性可表示为:

$$\int_0^{\pi}\int_0^{2\pi} Y_{lm}(\theta,\varphi)Y_{l'm'}^*(\theta,\varphi)\sin\theta d\theta d\varphi = \delta_{ll'}\delta_{mm'}. \tag{Ⅳ.32}$$

头几个球谐函数为:

$$\begin{aligned}
Y_{00} &= \sqrt{1/4\pi},\\
Y_{10} &= \sqrt{3/4\pi}\cos\theta,\\
Y_{1\pm1} &= \sqrt{3/8\pi}\sin\theta e^{\pm i\varphi},\\
Y_{20} &= \sqrt{5/16\pi}(3\cos^2\theta - 1),\\
Y_{2\pm1} &= \sqrt{15/8\pi}\cos\theta\sin\theta e^{\pm i\varphi},\\
Y_{2\pm2} &= \sqrt{15/32\pi}\sin^2\theta e^{\pm i2\varphi}.
\end{aligned} \tag{Ⅳ.33}$$

(五)拉普拉斯方程的通解

将(Ⅳ.8)式和(Ⅳ.31)式代入(Ⅳ.5)式,就得到拉普拉斯方程(Ⅳ.1)式的解:

$$\begin{aligned}
\Psi(R,\theta,\varphi) &= f_l(R)Y_{lm}(\theta,\varphi)\\
&= \sqrt{(l-|m|)!(2l+1)/[4\pi(l+|m|)!]}\,P_l^{|m|}(\cos\theta)e^{im\varphi}.
\end{aligned} \tag{Ⅳ.34}$$

但是，这种解是 R,θ 和 φ 的三个函数的乘积. 用这种特殊形式求解的好处是可以将包含三个变量的偏微分方程(Ⅳ.1)分解成三个单变量的常微分方程(Ⅳ.6)、(Ⅳ.10)、(Ⅳ.11). 然而，拉普拉斯方程的解一般并不具有三个函数乘积的形式. 因而要将(Ⅳ.34)中 l 和 m 取不同值的解叠加起来，才能得到拉普拉斯方程的通解：

$$\Psi(R,\theta,\varphi) = \sum_{lm} \left[a_{lm} R^l + (b_{lm}/R^{l+1}) \right] P_l^{|m|}(\cos\theta)\cos m\varphi + \sum_{lm} \left[c_{lm} R^l + (d_{lm}/R^{l+1}) \right] P_l^{|m|}(\cos\theta)\sin m\varphi. \tag{Ⅳ.35}$$

(六)勒让德多项式的母函数

可以证明

$$\frac{1}{\sqrt{1-2xt+t^2}} = \sum_{l=0}^{\infty} P_l(x) t^l, \qquad |t| < 1, \tag{Ⅳ.36}$$

它称为勒让德多项式的母函数. 将它对 t 求 l 阶导数，除以 $l!$ 再让 $t=0$，就得到 $P_l(x)$.

附录Ⅴ　一些积分公式

(一)高斯型积分

$$\int_{-\infty}^{+\infty} e^{-x^2}dx = \sqrt{\pi}. \tag{Ⅴ.1}$$

证　写上式左方的积分为 I,则

$$I^2 = \int_{-\infty}^{+\infty} e^{-x^2}dx\int_{-\infty}^{+\infty} e^{-y^2}dy = \iint_{-\infty}^{+\infty} e^{-(x^2+y^2)}dxdy,$$

积分遍及整个 xy 平面.引入极坐标 (r,θ),则

$$I^2 = \int_0^{2\pi}\int_0^{\infty} e^{-r^2}rdrd\theta = \pi.$$

注意被积函数是偶函数,故有

$$\int_0^{\infty} e^{-x^2}dx = \frac{1}{2}\sqrt{\pi}. \tag{Ⅴ.2}$$

【证毕】

(二)高斯分布的矩

$$I(n) = \int_0^{\infty} e^{-\alpha x^2}x^n dx \qquad (n = 0,1,2,\cdots).$$

作变量代换 $y = \alpha^{1/2}x$,可得

$$I(0) = \alpha^{-1/2}\int_0^{\infty} e^{-y^2}dy = \frac{1}{2}\left(\frac{\pi}{\alpha}\right)^{1/2}. \tag{Ⅴ.3}$$

$$I(1) = \alpha^{-1}\int_0^{\infty} e^{-y^2}ydy = \frac{1}{2\alpha}. \tag{Ⅴ.4}$$

高阶的 $I(n)$ 可以通过将 $I(0)$ 或 $I(1)$ 对 α 求导得到:

$$I(n) = -\frac{\partial}{\partial\alpha}I(n-2). \tag{Ⅴ.5}$$

例如:

$$\begin{aligned}
I(2) &= \int_0^{\infty} e^{-\alpha x^2}x^2dx = \frac{1}{4}\left(\frac{\pi}{\alpha^3}\right)^{1/2};\\
I(3) &= \int_0^{\infty} e^{-\alpha x^2}x^3dx = \frac{1}{2\alpha^2};\\
I(4) &= \int_0^{\infty} e^{-\alpha x^2}x^4dx = \frac{3}{8}\left(\frac{\pi}{\alpha^5}\right)^{1/2};\\
I(5) &= \int_0^{\infty} e^{-\alpha x^2}x^5dx = \frac{1}{\alpha^3}.
\end{aligned} \tag{Ⅴ.6}$$

(三)玻色分布的积分

$$I(n) = \int_0^{\infty}\frac{x^{n-1}}{e^x-1}dx, \quad \left(n = 2,3,4,\frac{3}{2}\text{ 及 }\frac{5}{2}\right)$$

因为 $\frac{x^{n-1}}{e^x-1}=x^{n-1}\frac{e^{-x}}{1-e^{-x}}=x^{n-1}e^{-x}(1+e^{-x}+e^{-2x}+\cdots)=\sum_{l=1}^{\infty}x^{n-1}e^{-lx}$，故

$$I(n)=\int_0^\infty\frac{x^{n-1}}{e^x-1}dx=\sum_{l=1}^{\infty}\int_0^\infty x^{n-1}e^{-lx}dx=\sum_{l=1}^{\infty}\frac{1}{l^n}\int_0^\infty y^{n-1}e^{-y}dy. \tag{Ⅴ.7}$$

其中 $\sum_{l=1}^{\infty}\frac{1}{l^n}=\zeta(n)$ 叫做黎曼 zeta 函数，且有 $\zeta(2)=\frac{\pi^2}{6}$，$\zeta(3)=1.202$，$\zeta(4)=\frac{\pi^4}{90}$，$\zeta(3/2)=2.612$，$\zeta(5/2)=1.341$，…. 例如：

$$I(2)=\int_0^\infty\frac{x}{e^x-1}dx=\sum_{l=1}^{\infty}\frac{1}{l^2}\int_0^\infty ye^{-y}dy=\sum_{l=1}^{\infty}\frac{1}{l^2}=\frac{\pi^2}{6}\approx 1.645; \tag{Ⅴ.8}$$

$$I(3)=\int_0^\infty\frac{x^2}{e^x-1}dx=2\sum_{l=1}^{\infty}\frac{1}{l^3}=2\times 1.202; \tag{Ⅴ.9}$$

$$I(4)=\int_0^\infty\frac{x^3}{e^x-1}dx=6\sum_{l=1}^{\infty}\frac{1}{l^4}=6\times\frac{\pi^4}{90}=6\times 1.082; \tag{Ⅴ.10}$$

$$I(3/2)=\int_0^\infty\frac{x^{1/2}}{e^x-1}dx=\sum_{l=1}^{\infty}\frac{1}{l^{3/2}}\int_0^\infty y^{1/2}e^{-y}dy=2\sum_{l=1}^{\infty}\frac{1}{l^{3/2}}\int_0^\infty t^2e^{-t^2}dt$$

$$=\frac{\sqrt{\pi}}{2}\sum_{l=1}^{\infty}\frac{1}{l^{3/2}}=\frac{\sqrt{\pi}}{2}\times 2.612; \tag{Ⅴ.11}$$

$$I(5/2)=\int_0^\infty\frac{x^{3/2}}{e^x-1}dx=\frac{3\sqrt{\pi}}{4}\sum_{l=1}^{\infty}\frac{1}{l^{5/2}}=\frac{3\sqrt{\pi}}{4}\times 1.341. \tag{Ⅴ.12}$$

（四）费米分布的积分

$$I=\int_0^\infty(e^{(\varepsilon-\mu)/(k_BT)}+1)^{-1}f(\varepsilon)d\varepsilon,$$

$$=\int_0^\mu f(\varepsilon)d\varepsilon+\frac{1}{6}(\pi k_BT)^2f'(\mu)+\frac{7}{360}(\pi k_BT)^4f'''(\mu)+\cdots. \tag{Ⅴ.13}$$

$$\mu/(k_BT)\gg 1.$$

证　令 $x=(\varepsilon-\mu)/(k_BT)$，则

$$I=k_BT\int_{-\mu/(k_BT)}^{0}\frac{f(\mu+k_BTx)}{e^x+1}dx+k_BT\int_0^\infty\frac{f(\mu+k_BTx)}{e^x+1}dx$$

$$=k_BT\int_0^{\mu/(k_BT)}f(\mu-k_BTx)\left[1-\frac{1}{e^x+1}\right]dx+k_BT\int_0^\infty\frac{f(\mu+k_BTx)}{e^x+1}dx$$

$$\approx\int_0^\mu f(\varepsilon)d\varepsilon+k_BT\int_0^\infty\frac{f(\mu+k_BTx)-f(\mu-k_BTx)}{e^x+1}dx\left(\text{已取上限}\ \mu/(k_BT)\to\infty\right)$$

$$\approx\int_0^\mu f(\varepsilon)d\varepsilon+2k_BT\left[k_BTf'(\mu)\int_0^\infty\frac{x}{e^x+1}dx+\frac{1}{3!}(k_BT)^3f'''(\mu)\int_0^\infty\frac{x^3}{e^x+1}dx+\cdots\right].$$

由于

$$\int_0^\infty\frac{x^{n-1}}{e^x+1}dx=\int_0^\infty x^{n-1}e^{-x}\sum_{l=0}^{\infty}(-1)^le^{-lx}dx=\sum_{l=0}^{\infty}(-1)^l\int_0^\infty x^{n-1}e^{-(l+1)x}dx$$

$$=(n-1)!\sum_{l=1}^{\infty}(-1)^{l-1}\frac{1}{l^n}=\left(1-\frac{1}{2^{n-1}}\right)(n-1)!\sum_{l=1}^{\infty}\frac{1}{l^n}.$$

代入上式,最后得出(Ⅴ.13)式. 【证毕】

(五)Γ 函数

定义

$$\Gamma(n)=\int_0^{\infty}e^{-x}x^{(n-1)}dx. \tag{Ⅴ.14}$$

性质

$$\Gamma(n)=(n-1)\Gamma(n-1)=(n-1)(n-2)\cdots 2\cdot 1\Gamma(1)=(n-1)!. \tag{Ⅴ.15}$$

$$\Gamma\left(n+\frac{1}{2}\right)=\left(n-\frac{1}{2}\right)\left(n-\frac{3}{2}\right)\cdots\frac{3}{2}\cdot\frac{1}{2}\cdot\Gamma\left(\frac{1}{2}\right); \tag{Ⅴ.16}$$

$$\Gamma\left(\frac{1}{2}\right)=\int_0^{\infty}e^{-x}x^{-1/2}dx=2\int_0^{\infty}e^{-y^2}dy=\sqrt{\pi}. \tag{Ⅴ.17}$$

利用Γ函数的性质(Ⅴ.15)式

$$\Gamma(x+1)=x\Gamma(x), \tag{Ⅴ.18}$$

及其渐近展开式

$$\Gamma(x+1)=x^xe^{-x}\left[1+\frac{1}{12x}+\frac{1}{288x^2}+\cdots\right]\qquad(N>10), \tag{Ⅴ.19}$$

得到 N 大时的近似公式

$$\ln N!\approx N\ln N-N\qquad(N>10), \tag{Ⅴ.20}$$

称为斯特林公式.其误差近似为

$$\delta\approx\ln(2\pi N). \tag{Ⅴ.21}$$

在 $N>10$ 以后,相对误差

$$\delta/\ln N!<27\%,\qquad N>10. \tag{Ⅴ.22}$$

(六)雅可比行列式

单变量积分换元时,要用到变元之间的微分导数

$$\int f(y)dy=\int f(y(x))\left|\frac{dy}{dx}\right|dx. \tag{Ⅴ.23}$$

对于多元函数的积分,相应地要用雅可比行列式.例如,二元函数 $f(x,y)$ 的二重积分

$$\int f(u,v)dudv=\int f(u(x,y),v(x,y))\left|\frac{D(u,v)}{D(x,y)}\right|dxdy, \tag{Ⅴ.24}$$

其中 $D(u,v)/D(x,y)$ 是雅可比行列式.

定义 设有 $u=u(x,y),v=v(x,y)$,则由以下偏导组成的行列式叫雅可比行列式,

$$\frac{D(u,v)}{D(x,y)}=\begin{vmatrix}\left(\frac{\partial u}{\partial x}\right)_y & \left(\frac{\partial u}{\partial y}\right)_x\\ \left(\frac{\partial v}{\partial x}\right)_y & \left(\frac{\partial v}{\partial y}\right)_x\end{vmatrix}=\left(\frac{\partial u}{\partial x}\right)_y\left(\frac{\partial v}{\partial y}\right)_x-\left(\frac{\partial u}{\partial y}\right)_x\left(\frac{\partial v}{\partial x}\right)_y. \tag{Ⅴ.25}$$

雅可比行列式有如下性质:

A.
$$\frac{D(u,v)}{D(x,y)} = -\frac{D(v,u)}{D(x,y)} = -\frac{D(u,v)}{D(y,x)}, \tag{Ⅴ.26}$$

B.
$$\frac{D(u,v)}{D(x,v)} = \left(\frac{\partial u}{\partial x}\right)_v, \tag{Ⅴ.27}$$

C.
$$\frac{D(u,v)}{D(x,y)} = \frac{D(u,v)}{D(s,t)}\frac{D(s,t)}{D(x,y)}. \tag{Ⅴ.28}$$

利用定义式(Ⅴ.25),以上性质很容易直接得到.

(七)衍射积分●

$$\int_0^T [t_1(T-t_1)]^{-3/2} e^{\frac{im}{2\hbar}\left[\frac{a^2}{t_1}+\frac{b^2}{T-t_1}\right]} dt_1 = \sqrt{\frac{2\pi^2}{im}} T^{-3/2}\left(\frac{a+b}{ab}\right)^2 e^{\frac{im(a+b)^2}{2\hbar T}} \tag{Ⅴ.29}$$

● 参看 E. G. Harris. Introduction to Modern Theoretical Physics, New York. John Wiley & Sons. 1975: 487.

附录Ⅵ　电磁量的国际单位制与高斯单位制

(一)国际单位制和高斯单位制的基本量和基本单位

1. 国际单位制

国际单位制(MKSA 有理制)的基本量是长度、质量、时间和电流强度,其基本单位分别为 m、kg、s(米、千克、秒)和 A(安培,简称安).其中,安培是根据下述规定来确定的:若真空中相距 1m 的两根通有相等的恒定电流的无限长直导线每米长度的相互作用力为 2×10^{-7}N(牛顿),则其电流为 1A.其他的物理量可按物理关系式导出.如真空磁导率按

$$F=\frac{\mu_0 l I^2}{2\pi a} \tag{Ⅵ.1}$$

求得

$$\mu_0=4\pi\times10^{-7}\text{N}\cdot\text{A}^{-2}(\text{牛顿}\cdot\text{安培}^{-2}). \tag{Ⅵ.2}$$

2. 高斯单位制

高斯单位制是绝对静电单位制(CGSE 单位制)和绝对电磁单位制(CGSM 单位制)的结合,它的所有电学量用 CGSE 单位制量度,所有磁学量用 CGSM 单位制量度.

CGSE 单位制的基本量分别为长度、质量、时间,基本单位分别为 cm、g、s(厘米、克、秒).电量的单位(第一个导出单位)根据真空中的库仑定律

$$F=k\frac{q_1q_2}{r^2}$$

导出,令 $k=1$.若真空中两个电量相等的静止点电荷相距 1cm 时的作用力为 1dyn(达因),则每个点电荷的电量为 1CGSE.

CGSE 单位制中所有电磁学量均用 CGSE 标记.

CGSM 单位制的基本量分别为长度、质量、时间,基本单位也是 cm、g、s.电流强度的单位(第一个导出单位)根据真空中两根无限长平行载流导线的相互作用力公式

$$F=k\frac{2lI_1I_2}{a}$$

导出,令 $k=1$.若真空中两根电流相同,相距 2cm 的平行细导线每厘米长度的相互作用力为 1dyn 时,则每根导线中的电流强度为 1CGSM.

CGSM 单位制中,除磁感应强度用高斯,磁通量用麦克斯韦,磁场强度用奥斯特表示外,其他磁学量均用 CGSM 标记.

这样,高斯单位制中的力学量均由 cm、g、s 及其导出单位标记,电学量由 CGSE 标记,磁学量由 CGSM(除高斯、麦克斯韦、奥斯特以外)标记.当公式中同时包含电学量和磁学量时,其比例系数则由实验测定.

(二)国际单位制与高斯单位制之间的单位换算关系

物理量	国际单位制	高斯单位制
电量 q	1C(库仑)	3×10^{9}CGSE
电流强度 I	1A(安培)	3×10^{9}CGSM
电势 φ	1V(伏特)	(1/300)CGSE
电场强度 E	$1V\cdot m^{-1}$(伏特/米)	$(1/3)\times10^{-4}$CGSE
电阻 R	1Ω(欧姆)	$(1/9)\times10^{-11}$CGSE
电容 C	1F(法拉)	9×10^{11}CGSE
磁感应通量 ϕ	1Wb(韦伯)	10^{8} 麦克斯韦
磁感应强度 B	1T(特斯拉)	10^{4}G(高斯)
磁场强度 H	$1A\cdot m^{-1}$安培/米	$4\pi\times10^{-3}$Oe(奥斯特)

(三)国际单位制与高斯单位制的重要常数及公式对照表

物理量	国际单位制	高斯单位制
真空磁导率 μ_0	$4\pi\times10^{-7}H\cdot m^{-1}$(亨/米)	1
真空介电常数 ε_0	$10^{7}/(4\pi c^{2})F\cdot m^{-1}$(法/米)	1
真空中点电荷的电场 $\boldsymbol{E}$	$(Q/4\pi\varepsilon_0 r^{2})\boldsymbol{e}_r$	$(Q/r^{2})\boldsymbol{e}_r$
真空中线电流的磁场 $\boldsymbol{B}$	$\int\frac{\mu_0 \boldsymbol{J}_f(\boldsymbol{x}')\times\boldsymbol{r}}{4\pi r^{3}}dV'$	$\int\frac{\boldsymbol{J}_f(\boldsymbol{x}')\times\boldsymbol{r}}{c\,r^{3}}dV'$
洛伦兹力 $\boldsymbol{F}$	$Q(\boldsymbol{E}+\boldsymbol{v}\times\boldsymbol{B})$	$Q(\boldsymbol{E}+(\boldsymbol{v}/c)\times\boldsymbol{B})$
电磁能量密度 w	$(\boldsymbol{E}\cdot\boldsymbol{D}+\boldsymbol{H}\cdot\boldsymbol{B})/2$	$(\boldsymbol{E}\cdot\boldsymbol{D}+\boldsymbol{H}\cdot\boldsymbol{B})/(4\pi)$
电磁能流密度 S	$\boldsymbol{E}\times\boldsymbol{H}$	$(c/4\pi)\boldsymbol{E}\times\boldsymbol{H}$
电磁动量密度 $\boldsymbol{g}$	$\varepsilon_0\boldsymbol{E}\times\boldsymbol{B}$	$(1/4\pi c)\boldsymbol{E}\times\boldsymbol{B}$

国际单位制与高斯单位制的重要常量及公式对照表(续)

推迟势 A	$\int \frac{\mu_0 \boldsymbol{J}_f(\boldsymbol{x}', t - r/c)}{4\pi r} \mathrm{d}V'$	$\int \frac{\boldsymbol{J}_f(\boldsymbol{x}', t - r/c)}{cr} \mathrm{d}V'$
推迟势 φ	$\int \frac{\rho_f(\boldsymbol{x}', t - r/c)}{4\pi\varepsilon_0 r} \mathrm{d}V'$	$\int \frac{\rho_f(\boldsymbol{x}', t - r/c)}{r} \mathrm{d}V'$
电磁场张量 $F_{\mu\nu}$	$\begin{bmatrix} 0 & B_3 & -B_2 & -iE_1/c \\ -B_3 & 0 & B_1 & -iE_2/c \\ B_2 & -B_1 & 0 & -iE_3/c \\ iE_1/c & iE_2/c & iE_3/c & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & B_3 & -B_2 & -iE_1 \\ -B_3 & 0 & B_1 & -iE_2 \\ B_2 & -B_1 & 0 & -iE_3 \\ iE_1 & iE_2 & iE_3 & 0 \end{bmatrix}$
四维势矢量 A_μ	$(\boldsymbol{A}, i\phi/c)$	$(\boldsymbol{A}, i\phi)$
四维电流密度 J_μ	$(\boldsymbol{J}_f, ic\rho_f)$	$(\boldsymbol{J}_f, ic\rho_f)$
精细结构常数 α	$\frac{e^2}{4\pi\varepsilon_0\hbar c}$	$\frac{e^2}{\hbar c}$
电子经典半径 r_e	$\frac{e^2}{4\pi\varepsilon_0 mc^2}$	$\frac{e^2}{mc^2}$
辐射阻尼力 F_γ	$\frac{e^2}{4\pi\varepsilon_0 mc^3}\ddot{\boldsymbol{v}}$	$\frac{2e^2}{c^3}\ddot{\boldsymbol{v}}$
麦克斯韦方程组	$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}$ $\nabla \times \boldsymbol{H} = \frac{\partial \boldsymbol{D}}{\partial t} + \boldsymbol{J}_f$ $\nabla \cdot \boldsymbol{D} = \rho_f, \quad \nabla \cdot \boldsymbol{B} = 0$	$\nabla \times \boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t}$ $\nabla \times \boldsymbol{H} = \frac{1}{c}\frac{\partial \boldsymbol{D}}{\partial t} + \frac{4\pi}{c}\boldsymbol{J}_f$ $\nabla \cdot \boldsymbol{D} = 4\pi\rho_f, \quad \nabla \cdot \boldsymbol{B} = 0$
介质电磁性质	$\boldsymbol{D} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P}$ $\boldsymbol{B} = \mu_0(\boldsymbol{H} + \boldsymbol{M})$	$\boldsymbol{D} = \boldsymbol{E} + 4\pi\boldsymbol{P}$ $\boldsymbol{B} = \boldsymbol{H} + 4\pi\boldsymbol{M}$
场强和标势矢势的关系	$\boldsymbol{E} = -\nabla\varphi - \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}$ $\boldsymbol{B} = \nabla \times \boldsymbol{A}$	$\boldsymbol{E} = -\nabla\varphi - \frac{\partial \boldsymbol{A}}{\partial t}$ $\boldsymbol{B} = \nabla \times \boldsymbol{A}$
势方程	$\nabla^2 \boldsymbol{A} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} = -\mu_0 \boldsymbol{J}_f,$ $\nabla^2\varphi - \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} = -\frac{\rho_f}{\varepsilon_0}$	$\nabla^2 \boldsymbol{A} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} = -\frac{4\pi}{c}\mu_0 \boldsymbol{J}_f,$ $\nabla^2\varphi - \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} = -4\pi\rho_f$
洛伦兹条件	$\nabla \cdot \boldsymbol{A} + \frac{1}{c^2}\frac{\partial\varphi}{\partial t} = 0$	$\nabla \cdot \boldsymbol{A} + \frac{1}{c}\frac{\partial\varphi}{\partial t} = 0$
最小相互作用	$\boldsymbol{p} \to \boldsymbol{p} + e\boldsymbol{A}, \quad w \to w + e\phi$	$\boldsymbol{p} \to \boldsymbol{p} + (e/c)\boldsymbol{A}, \quad w \to w + e\phi$

附录Ⅶ　物理常量与单位换算

光速	$c = 2.9979 \times 10^{10}\,\mathrm{cm \cdot s^{-1}}$
万有引力常量	$G = 6.672 \times 10^{-8}\,\mathrm{cm^3 \cdot g^{-1} \cdot s^{-2}}$
阿伏伽德罗常量	$N_A = 6.022 \times 10^{23}\,\mathrm{mol^{-1}}$
玻尔兹曼常量	$k_B = 1.3807 \times 10^{-16}\,\mathrm{erg \cdot K^{-1}}$
普朗克常量	$h = 6.6262 \times 10^{-27}\,\mathrm{erg \cdot s}$
	$\hbar = h/(2\pi) = 1.0546 \times 10^{-27}\,\mathrm{erg \cdot s}$
里德堡常量(氢原子)	$R_H = 109678\ \mathrm{cm^{-1}}$
电子电荷	$e = 4.803 \times 10^{-10}\,\mathrm{esu}$
电子质量	$m_e = 9.1095 \times 10^{-28}\,\mathrm{g} = 0.5110\,\mathrm{MeV \cdot c^{-2}}$
玻尔半径	$a_0 = \hbar^2/(m_e c^2) = 0.5292 \times 10^{-8}\,\mathrm{cm}$
长度单位	1Å(埃) $= 10^{-8}$ cm
	1fm(费米) $= 10^{-13}$ cm
截面单位	1barn(靶恩) $= 10^3$ mb $= 10^6\,\mu$b $= 10^{-24}\,\mathrm{cm^2}$
能量单位	1eV(电子伏特) $= 10^{-6}$ MeV $= 1.6022$ erg(尔格)
	1J(焦耳) $= 10^7$ erg(尔格)
质量单位	1amu(原子质量单位) $= 1.6605 \times 10^{-24}\,\mathrm{g} = 931.5\ \mathrm{MeV \cdot c^{-2}}$
	$1\,\mathrm{eV \cdot c^{-2}} = 10^{-6}\,\mathrm{MeV \cdot c^{-2}} = 1.7827 \times 10^{-33}\,\mathrm{g}$
动量单位	$1\,\mathrm{eV \cdot c^{-1}} = 10^{-6}\,\mathrm{MeV \cdot c^{-1}} = 5.344 \times 10^{-23}\,\mathrm{g \cdot cm \cdot s^{-1}}$
力单位	1N(牛顿) $= 10^5$ dyn(达因)
体积单位	1l(升) $= 10^{-3}\,\mathrm{m^3} = 10^3\,\mathrm{cm^3}$
压强单位	$1\,\mathrm{Pa}$(帕斯卡,简称帕) $= 1\,\mathrm{N \cdot m^{-2}}$
	$1\,\mathrm{Ba}$(巴里,简称巴) $= 1\,\mathrm{dyn \cdot cm^{-2}} = 0.1\,\mathrm{Pa}$
	1atm(大气压) $= 1.013 \times 10^5\,\mathrm{Pa}$
	1Torr(托) $=$ 1mmHg(毫米汞柱) $= 1.333 \times 10^2\,\mathrm{Pa}$
粘滞系数单位	1Po(Poise,简称泊) $= 1\,\mathrm{g \cdot cm^{-1} \cdot s^{-1}} = 1\,\mathrm{Ba \cdot s}$

郑重声明

高等教育出版社依法对本书享有专有出版权。任何未经许可的复制、销售行为均违反《中华人民共和国著作权法》，其行为人将承担相应的民事责任和行政责任，构成犯罪的，将被依法追究刑事责任。为了维护市场秩序，保护读者的合法权益，避免读者误用盗版书造成不良后果，我社将配合行政执法部门和司法机关对违法犯罪的单位和个人给予严厉打击。社会各界人士如发现上述侵权行为，希望及时举报，本社将奖励举报有功人员。

反盗版举报电话：(010) 58581897/58581698/58581879/58581877

传　　真：(010) 82086060

E - mail：dd@hep.com.cn 或 chenrong@hep.com.cn

通信地址：北京市西城区德外大街 4 号
高等教育出版社法律事务部

邮　　编：100011

购书请拨打电话：(010)64014089　64054601　64054588