袁笃平物理学讲义

基础物理

作者 袁笃平

2013 年春

目录

第一章 物理量……3
第二章 运动学……43
第三章 动力学……116
第四章 基本相互作用……215
第五章 刚体力学……278
第六章 连续介质力学——物体的力学特性……322
第七章 振动与波……387

第 1 章 物理量

如果你能够对某种事物进行测量并能对事物的性质给出量化描述，你对该事物就有一定程度的了解。如果做不到这些，你对该事物的理解程度必然是贫乏的，不能令人满意的。

——威廉·汤姆森[1]

1-1 引言

物理学是研究物质基本结构和物质运动最一般规律的科学。和所有研究物质的其它自然学科一样，物理学对物质的认知和描述并不是物质本身，而是我们观测到的物质特性和行为。物理学从诞生之时就努力将描述和理解物质及其运动建立在定量观测的基础上，这决定了物理学形成为一门数量化的科学。需要注意的是一种测量无论涉及多么深奥的原理、精巧的设计和复杂的技术，从本质上来说都只是一种提供数量化对比的方法或手段。这种比较被认为是科学的基本特征之一：没有测量，我们对事物的认识最多只是一种怀疑和推测。

我们对一个陌生物体的认知过程一般是从感官开始的：通过我们的眼睛构建物体的大小，形状，颜色，透明度等视觉印象；通过触摸了解到它的软硬，弹性，表面的光滑度；通过敲击倾听它发出的声音；如果我们足够勇敢，还会用舌头尝试一下它的滋味。在物理学中，这些直接通过人体感官获得的印象和认知可以通过感知物质特性的方式和过程定义相应的参量—**物理量**，甚至可以利用不同物理量在观测中表现出的关联，设计相应的仪器来探测和量度我们所不能感知的物体特性。例如，我们可以用“尺”量度物体的大小，形状，还能通过一些特别的仪器测量材料的脆性，韧性，弹性、延展性等力学性能，依据电磁现象的感知方式设计仪器来测定物质的电导，磁化等电磁特性...通过比较和分析定量观测的数据获得一些看似完全不同特性之间的内在联系，从而使我们对物质的认识超越人体器官的感知。例如，人眼对物体‘**颜色**’的感知在很大程度上只是半定量的，有时甚至是不准确的：人眼不能分辨单色和复色，而在不同亮度的照明条件下，人眼对同一个物体‘颜色’的判断会有所不同。但当麦克斯韦在计算当时他假想的电磁波传播速度后发现该速度和当时已经测得的光速数据非常接近时，就判断出光波就是电磁波，由此揭示出原来似乎完全没有关联的光和电磁现象具有的相同本原。而将光描述为一种波也就建立了‘颜色’和波的‘频率’之间的联系，由此颜色被定义为人眼对不同频率光波的响应，并设计出可以测量光波频率的仪器来量度颜色，区分人眼所不能分辨的单色和复色。

物理学中，对物质及其运动的认知一般涉及三个过程和层次：感知[2]→测量；

---

[1] 威廉·汤姆森(William Thomson 1824-1907) 出生于北爱尔兰的英国数学物理学家、工程师。他在格拉斯哥大学与 Hugh Blackburn 合作研究电学的数学分析、他是热力学温标的发明人，将第一和第二热力学定律公式化，被称为**热力学之父**。他对电报机的研究给他带来财富和荣誉，由于在横跨大西洋的电报工程中的贡献而被维多利亚女王授予威廉·汤姆森男爵爵位。他是首位进入英国上议院的科学家。政治上他反对爱尔兰自治，因而自称**拉格斯的开尔文男爵**，这个头衔来自于流经他在苏格兰格拉斯哥大学实验室的开尔文河。

[2] 对物体的观察一般只是对物体某种特性的定性感知，而获得定量感知的方法称为测量。测

实验数据→经验公式(两个或多个参量之间的数学关系)；经验公式→物理定律(公式中参量按实际的物理含义分解组合，反映各种因素的内在关联，如因果关系)。而描述这些反映物理特性和行为的物理参量是贯穿这样一个认知过程的最基本的要素。通过物理量可以利用数学运算分析物质的相关特性，综合有关方面性能的物理量拼图成描述该物体相关性能的数学模型，归纳出更为普适的物理定律。我们甚至能依据我们的某些特殊需求，设计并制造出一些具有相应“超性能”的“超物质”。例如，我们看到一个物体的存在是基于这个物体反射光或吸收光(物体吸收可见光使我们在照明光下可以看到物体的黑暗轮廓)这两种特性，用物理量可以表述为该物体对光的反射率或吸收率不为零，如果我们不希望该物体被别人看见，可以设计一种具有特殊结构的超物质，使照射到上面的光波不会被吸收，也不被直接反射，而是绕过物体后再沿原来的路径传播，该物体就会从我们的视觉中消失而具有隐身的特性(见图 1-1-1)。事实上物理学正是通过这种基于物理量所定义的定量观测以及由对观测结果的定量比较分析所延伸的进一步的性能测量，使我们能够不断完善对物质的理解和认识。

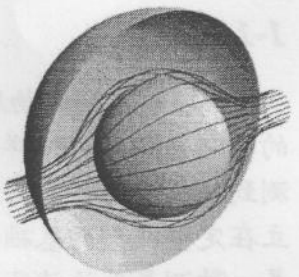


图 1-1-1 照到物体上的光波绕过物体后再沿原来的路径传播，使物体达到隐身的效果。

下面再通过一个简单的临床事例来说明定量观测对描述和理解物质特性的意义。设想你受到某种昆虫的叮咬，你会如何向医生描述你的“痛”感呢？通常我们都是用文字很不准确地表述“痛”的程度，如“很痛”或“有点痛”，有时还会给出一些分类描述，如“钻心般的痛”，“烧灼似的痛”，又是则可能是毫无意义的“说不出来的痛”。医生基于这种缺乏数量对比的表述所做出的分析和诊断往往只是定性的猜测。因判断失误而误诊的情形是经常发生的。医学中有很多症状由于不能找到一个恰当的“医学临床参量”而无法做出任何诊断，事实上，直至今天也还没有一种仪器能数量化地表述或量度“痛”(或“痒”)的程度，以及对不同的“痛”(或“痒”)感之间的差异进行分类。

读者可能会对物质的性质在物理上可以数量化，而在化学，医学，生物学科中做不到这一点感到疑惑，事实上，物理学对物质演化过程的理解和描述通常是基于物质的某个本质不变的层次，讨论不变基元的组合形式或形态变化所呈现的现象[3]，而化学和生物学对物质演化的描述则是大多基于物质本质的改变。例如，在一个碳被氧化为二氧化碳的过程前后，化学家的描述是原来的物质—碳(C)分子和氧分子(O_2)分子消失了，产生了新的物质二氧化碳(CO_2)分子。而在物理学家眼中，反应前后，C 原子和氧原子都没有改变，只是它们彼此之间的相对距离

量首先要确定观察什么，如何观察才能正确反映这种性质。对某种性质的不同测量设计反映的是不同测量属性。没有描述，就无法知道如何进行测量。例如人体器官的各种感知是每个人都能切身体会的。但绝大多数却不能给予恰当的描述。因而也就无法进行测量。一般来说，人们是通过了解物体所展出来的各种性质来认识物体的，而所谓物体的性质是指物体受到外界扰动后所表现出来的响应行为。物体只是其诸多性质集合的一个的载体(物理学家对一个实际物体本身通常并不能说出什么所以然)。

[3] 物理学所基于的哲学信念：世界上所有的一切起源于物质和物质的运动，物质的本原和物质本原之间的基本相互作用只有很少的种类，物质世界所呈现出的多样性以及运动复杂性只是它们相互之间的可能组合具有多样性。

发生了变化。而在生物学中还涉及生长，发育，生殖等更为复杂的变化，但这些在物理学的相关运动方程中，都只是固定数量的原子在电磁作用下的一个解。

作为描述和理解以后各章展开的物理知识及方法的基础，本章并不针对某一个具体物理系统的某一个具体的物理量作深入讨论，而是关注于物理学中所有与物质运动和物质特性相关的**物理量所具有的共同属性：它们的数学形式，量值大小的判断，理解，演绎某种过程**中不同物理量相互之间的可能逻辑关系。这些讨论将是以后各章中讨论具体系统，具体过程时所必须知晓的基本方法。因此这一章是对学习和讨论书中整个内容的一个基本准备。

本章关于物理量的讨论主要涉及以下几个方面：

- ✧ 物理量的测量属性和相关的操作性定义；
- ✧ 物理量的数值表示，其中涉及两个方面：物理量的单位和数值精度；
- ✧ 物理量的数学类型及其运算规则；
- ✧ 实验数据的分类，及得出实验曲线的经验公式；
- ✧ 特定对象的物理量数量级的估计和意义。

1-2 基本物理量

每一个物体都有着许多不同的观测特性，相应地就有着许多的物理量，但这些物理量并不都是独立的。由于每一种特性都基于实际的测量操作，而物体的大多数特性测量都是**间接**的，这意味着这些特性相互关联。这就引出了下列问题：物理学中需要多少独立的基本物理量？什么样的物理量可以成为基本物理量？

自然科学中有一种主流观点：构成物质的最基本组分和它们之间的相互作用类型很少，物质世界及其运动所表现出来的复杂性只是物质基元粒子通过基本相互作用所形成的组合具有丰富而且复杂的表现形式。因此，一个物理系统所处的任何复杂的物理状态或所经历的复杂的物理过程都可以表述为一系列有一定因果关系联系起来的一连串**简单事件**的特定组合。所谓简单事件是指在某一个时刻，物体中的某一物质点处在空间的某一个具体位置。例如，一个自由下落的灯泡在地面破裂成碎片这样一个事故可以理解为由一个特定事件—灯泡上的某一物质点在某一时刻在某一空间位置点与地面物质点发生重合—开始的一系列同类事件的集合(图 1-2-1)。其中的任何一个事件都可以由三个因素：**空间点，时间点和物质点**所完全描述。如果这三个点及其变化都能用数量来表示，任何物理事件就可以进行定量描述，或者说，能够对上述三个因素作定量描述的参量就构成整个物理世界的**基本物理量**。

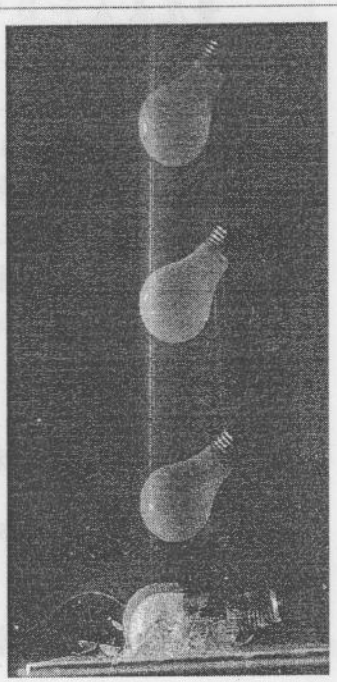


图 1-2-1 灯泡碎裂这一复杂事件可以分解为相继发生的一系列同类简单事件的集合。

上述三个基本物理量各有其自身的特性，如何通过它们有效地描述物质世界的性质和行为依赖于我们从事相关观测的经验积累，并对其主要特征予

以抽象。由此才能赋予这三个物理量以最大可能的普适性，成为真正的基本物理参量。下面首先讨论这三个物理量的基本性质。

1-2-1 基本物理量的观测特性

空间，时间和物质都是我们所能感知的物理量。但感知的直觉程度并不相同，有的非常直接，有的则比较间接，有的在描述时甚至还比较抽象。而且在感知上，三者之间也并不相互独立。

表述空间的特征物理量—长度及其测量特性

我们对空间最直觉的感知是我们周围的物体所处的不同位置及相互之间的方位关系。一个物体所占据的所有位置的集合在我们的意识中形成这一物体的大小概念，而物体的运动范围则汇集成空间的印象：

空间的概念是通过对空间中的点和点之间相互关系的描述建立起来的，这种关系表现为点和点之间的相对距离和相对方位。而这两个特性都可以通过长度这一参量进行实际测量和描述。

首先讨论量度空间任意两个点 A 和 B 间距(长度)L 的实验操作。如果这两个点都在我们所能触及的范围内，长度 L 可通过直接比较的方式进行测量：连接 A 和 B 两点画一条直线，用一根长度 $L_o(<L)$的较短直棒 A_o 和 L 比较长度。比较的结果可表述为“长度 L 是 A_o 长度的若干倍”，表示为

$$L/L_o = l \quad 或 \quad L = lL_o \tag{1-2-1}$$

其中 A_o 就是通常所谓的量尺，L_o 称为该测量的长度单位，数量 l 就是棒 A 的长度测量值。

量度点 A 和 B 之间的相对方位则要复杂一些，空间相对我们自身的取向有上下、前后、左右之分，因此，量度点 A 和 B 的相对方位先要设定方向的标准。选取 A 作为参考(标准)点，任意选取三条通过 A 点的相互正交的有向直线，按右手螺旋法则将这三条有向直线分别标记为 x, y, z 三个轴(图 1-2-2)。再以上面的 L_o 作为长度单位，以 A 为起点沿三条轴分别对其进行标度。然后过 B 点分别对 x, y, z 三个轴作垂线得到三个相交点，这三个相交点与 A 点的距离分别为 L_x, L_y, L_z，则点 A 和 B 之间的相对方位可以定义为

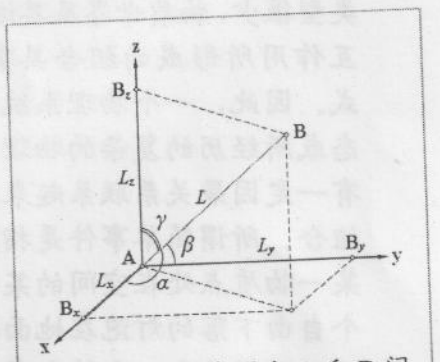


图 1-2-2 空间中两点 A 和 B 间的相互关系：方位和距离的测量和描述定义了空间的特性。

$$\cos\alpha = \frac{L_x}{L},\ \cos\beta = \frac{L_y}{L},\ \cos\gamma = \frac{L_z}{L} \tag{1-2-2}$$

称为 B 点相对于 A 点的方向余弦。以上的长度测量操作以及基于长度测量的方向的定义给出了对空间的一种完整的定量描述(标度)。

时间及其观测特性

我们对时间的直觉认知更多的是对时钟“均匀滴答”的感知。但对时间的测量

操作并不能象长度测量那样可以进行直接的对比。其原因是在我们的直觉中并没有象空间中“尺”那样的实物可用作直接比较时间长短的标准。

事实上，我们对时间的感知和测量从来都是和物体的往复运动联系在一起的。例如，太阳日复一日地经过我们头顶，则相邻两次日照当头的时间间隔就能定义为一个**时间单位**，称为**太阳日**。显然，时间测量中所观测的还是空间位置，即时间测量是以一个作周期运动的物体离开某一空间位置后回归到同一点的时间周期作为“时间量尺”，量度某一个物理过程所经历的时间，实际就是观测相关的物体从一个空间点移动到另一个空间点(或回到原来的空间点)的时间内，“时间量尺”运动所经历的周期数。

时间测量属于间接测量，因而任何时间测量方法中都隐含有物理假设。我们日常测量时间的器具如日晷，沙漏，钟表所内含的假设为“时间量尺”的往复运动具有严格的周期性，而这种周期的稳定性取决于其所受到的外部影响应小到可以忽略。

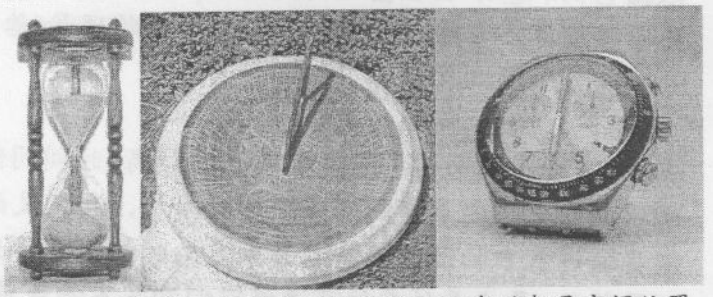


图 1-2-3 用各种时间量具测量时间所观察的都是空间位置。

描述物质的普适特性的参量—质量及其观测特性

基本物理事件的第三个元素是物质，也是事件的主体。由于物质种类的多样性，以及同一种物质可以具有许多不同的性质，对物质的描述相对于时间和空间而言存在更多的复杂性。这也是只关注于特定时空尺度内的各种自然学科对相关领域中的物质对象难以抽象出一个普适的参量进行数学描述的重要原因。

物理学在归纳基本物理事件特性时所关注的是在各种运动中物质所具有的普遍特性。目前我们所能总结出各种物质所具有的共性表现在两个方面，一是所有物质受到外部作用运动时，都会呈现出的不愿意改变其运动状态的特性，即所谓惯性，二是所有物质之间都存在一种相同类型的引力作用，因而这种引力被称为万有引力[4]。历史上将物质的这两种普适特性分别用两个参量—**惯性质量**和**引力质量**来描述。这两个物理量虽然有着不同的测量属性，但在目前测量技术所能达到的测量精度内都表现出严格的正比关系。因而通过对单位的适当标度，可以使这两个“质量”的数量在任何场合都严格相等。这两个量也因此常被不加区分地简称为“质量”[5]，也由此被设定为描述所

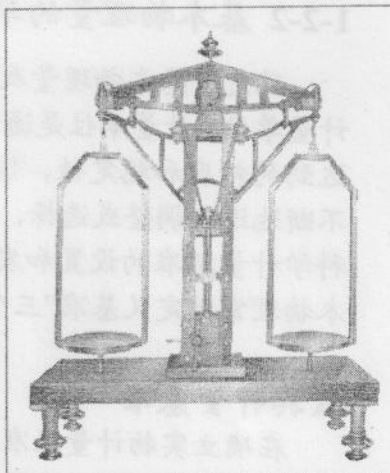


图 1-2-4 比较物体质量的量具—天平。实际观测的参量还是长度。这两个参量之间的关系隐含在天平的设计中。

[4] 关于这两种特性在第三章和第四章有比较详细的讨论。

[5] 质量通常读者在中学的课程中也许还遇到过其他的质量定义，如化学中表示原子量多少的

有物质的一个普适的特征物理量。

我们将在后面看到，这两种质量都有着多种观测属性，但都是通过相应的物理学定律测量可直接观测的长度和时间的读数。如测量一个物体的质量的最常用方法是利用天平和另一个作为质量单位的物体进行重力平衡比较(见图1-2-4)。具体观测的是长度。其中引力质量和长度之间关系隐含在天平的设计之中。

导出物理量

除了长度，时间，质量三个基本物理量，其它物理量大多不能通过这种简单、直接的对比进行测量(有些场合即使长度也无法通过直接比较进行测量，如太阳和地球间的距离)。对这些物理量的测量都是通过对基本物理量实施测量后，再依据它们和基本物理量之间的关系(定义或物理定律)进行计算间接实现的。例如用于量度赛车直线行驶时的快慢程度的物理量(称为速率)的测量属性定义为：

$$\text{速率}=\frac{\text{空间间隔}}{\text{时间间隔}}$$

其中，空间间隔为两个给定空间点的距离，时间间隔是赛车驶过这段距离的时间。通过直接比较测量空间间隔和时间间隔，和定义的关系式得到速率的测量结果。这些通过基本物理量的特定关系定义的物理量称为**导出量**，因此也不具有独立的测量单位，导出量的单位可通过它们和基本物理量的特定数学关系由这些基本物理量的单位(称为**基本单位**)的组合表示，称为**导出单位**。如速率的度量单位可以表示为“长度单位/时间单位”。

希望读者意识到，空间，时间和质量作为物理学(实际上也是所有自然学科)的基本参量意味着我们不能改变借助于时空的思考方式，而我们不能用这种方式来理解的东西，我们也就完全不能理解。

1-2-2 基本物理量的单位及其国际标准定义

确定了基本物理量及其测量属性，紧接着的问题是为这些物理量设立统一的**计量单位**。计量单位是通过实际物体或物理现象体现的，其关键在于测量中所能达到的精度和稳定性，国际计量协会对用作基本物理量单位标准的“**物理原器**”不断地进行调整或选择，基本物理量的测量精度也因此不断“刷新”。至今为止，科学计量标准的设置和发展大致经历了“实物原器基准”，“实验条件基准”和“基本物理常数定义基准”三个阶段。下面对这些历程作一个简短的介绍[6]。

经典计量基准

在确立实物计量基准以前，长度和时间单位都曾选择以地球的特性有关的参量作为参照的实物“原器”，其原因是地球本身及其运动当时被认为是非常稳定的。18世纪法国将子午线的四千万分之一定义为长度计量单位“米”(图 1-2-5a)[7]，

物质质量，这些质量之间都存在着相互关联。我们将在后面的章节中对此作进一步的讨论。

[6] 建议读者不要忽略这一部分。真正理解测量和测量精度的设计有助于物理量直觉的建立和对物理量测量属性的理解，实际感受物理学中的思维方法及物理学中不同的数量级层次。

[7] 从 1792 年起，法国天文学家耗时 7 年测量了从地球的北极经过法国北部的敦克尔克和西

记为 m。人们又将与地球自转相联系的一个昼夜定义为一个太阳日作为时间单位。对于质量单位一开始是将处于 4°C 和一个标准大气压条件下一升(10^{-3} m^3)纯水的质量定义为一千克。

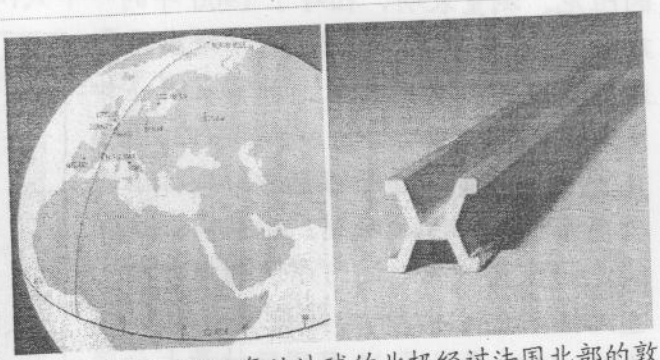


图 **1-2-5** (a) 1791 年从地球的北极经过法国北部的敦克尔克和西班牙的巴塞罗纳到赤道的距离的一千万分之一被定义为一个长度单位"米", (b) 1889 年, 1"米"的长度被固化为用铂铱合金制作的横截面呈X形的器物两端所镌刻的线痕间的距离。

这样定义的物理量单位，准确性不高，量值也不容易复现，而且难以确保世界各地所用的计量单位具有高度的一致性。1875 年，以法国为首成立了国际米制公约组织，米制公约确立了以十进位制，米和千克为基础的国际单位制(简记为 SI)，并决定将计量单位"米"和"千克"用更稳定、更容易复现的实物基准定义。1889 年米制公约的常设机构"国际计量局"(BIPM)用含铂 90%、铱 10%的合金制作了两种"实物基准"。一种是用作长度标准的横截面呈 X 形的"米"原器[8](图 1-2-5b)，1"米"的长度由此固化为在冰的融点温度条件下这根原器的两端所镌刻的线痕间的距离。质量单位的"千克原器"则是一个高度和直径都为 39.17 毫米的铂铱合金圆柱体(见图 1-2-6)。"千克原器"的测量属性是利用天平比较原器和待测物体受到的地球引力(即通常所说的重力)。

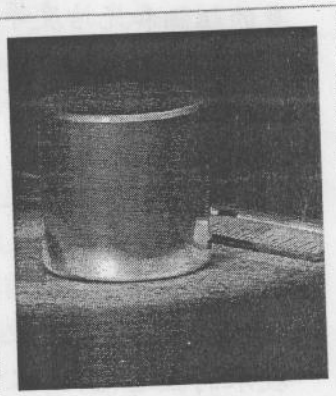


图 1-2-6 千克原器

国际计量局对时间单位从未设定过实物基准，而且直到 20 世纪 50 年代，才正式定义时间单位：

1 秒 ≡ 1 个平均太阳日的 1/86400。

其中太阳日定义为太阳连续两次处于地球某一经线(如格林威治经线)正上方所经历的时间，而所谓平均太阳日是由于一年中每个太阳日的长短存在差异，因而对一年中的太阳日进行了平均。这是由测量方法定义的单位，所谓**测量方法定义**是指在特定的实验环境中，测量对象所表现出来的现象或行为。上述关于时间定义的准确性依赖于地球自转速度的稳定性。自然界中存在着许多会影响地球自转速度的因素，如美国宇航局通过电脑模型评估 2010 年 2 月 27 日智利发生的里氏 8.8 级地震造成的影响结果显示，地震引起地球质量分布像地轴方向集中，地球自转速度加快，导致地球一天的时长缩短了 1.26 微秒[9]。持久，稳定地影响地球自转速度的因素主要是海洋潮汐对地球自转

班牙的巴塞罗纳到赤道的距离，1791 年法国国民议会决定采用这条通过巴黎的地球子午线长度的一千万分之一作为长度单位，并用古希腊文中"metron"一词作为这个单位的名称。

[8] 这种铂铱合金的膨胀系数非常小，而且不容易发生化学反应。而材料做成 X 形状不容易变形而且省料。以后的测量表明，"米"原器较地球子午线的四千万分之一大了 0.2 毫米。这时人们已意识到，随着测量技术的不断进步，新的偏差还会出现，因而不再修改米原器的长度，而直接以这个"米"原器作为统一所有长度计量的基准。

[9] 由于测量地球一天时长至今所能达到的精度约为 20 微秒，智利大地震造成的时长缩短只

的阻尼。测量表明，相对于 1900 年的一个太阳日的 86,400 秒，在 2000 年已经延长为 86,400.002 秒。图 1-2-7 给出的是地球上观测日食的位置。计算表明，如果地球的自转速度不变，公元前 136 年 4 月 15 日发生的日全食在图中红线标示的位置应该都可以观测到。但事实上日食的观测位置从北非的阿尔及利亚移到了巴比伦，说明地球的自转速度正在延缓。为此，1960 年的第 11 届国际计量大会确定将“秒”的定义修改为：

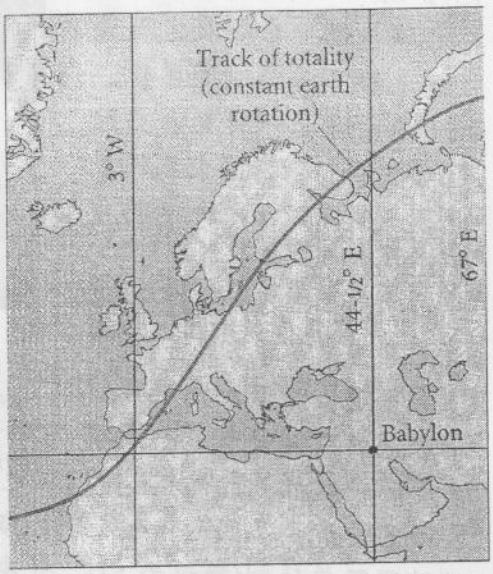


图 1-2-7 图中红线为基于地球维持其自转速度不变计算出的公元前 136 年 4 月 15 日的日全食可观测位置。

1 秒 =历书时 1900 年 1 月 0 日 12 时起算的回归年的 1/31,556,925.9747

这样定义的“秒”称为**历书秒**。其中**回归年**是太阳直射点回归运动周期: 365.2422 日，比**恒星年**(地球绕太阳的公转周期: 365.25636 日)短了 0.01416 日。[10]

上述基于经典物理学原理的实物计量基准设置反映了当时最好的材料及工艺水准。影响米原器精度的因素为刻线本身的粗细和光洁度等原因，其误差值约 0.1 微米，相对误差为 1×10^{-7}。时间单位“平均太阳日”因运动稳定性的限制，精度为 10^{-8} 秒。而当时千克原器的不确定度量级为 10^{-9}，是计量标准中最高的。

早期的物理学主要关注的是宏观物体的机械运动，这也是经典物理的主要研究内容，因而定义基本物理量单位基准的也都是基于宏观物质的机械运动。它所具有的特征是物体质量大，运动速度慢，周期长，短时间内不容易受外部微小扰动的影响，但测量便捷度和测量精度都不高。自 20 世纪 50 年代开始，随着电子技术的快速发展，大部分测量由原来的机械测量方式被电学测量方式替代。电子测量方式的特点是简洁，快速，同时还具有更高的测量精度。例如，石英片(SiO_2)在电场中会发生振荡，其振荡频率与其大小，形状，温度和外加电场有关。而这些外部参量很容易稳定下来，一块石英片产生频率为 10^8 Hz 的振荡，其频率起伏的数量级仅为 10^{-9}。将振荡转换为电脉冲信号，通过对所计量的时间间隔内的电脉冲计数，就能实现相应的时间测量。而对脉冲记数的最大误差只来源于首尾两个脉冲。则由上述晶片振荡测量时间的精度可达到 2×10^{-8}。

是计算结果，无法进行实际测量，而且时长缩短的准确持续时间也无法测量。

[10] 历书时间不同于天文时间。天文中的**恒星年**是以遥远的恒星为参照物，地球绕太阳的公转周期: 365 日 6 小时 9 分 9.5 秒(365.25636 日)。**近点年**是地球两次经过椭圆轨道上近日点的时间间隔: 365.25964 日，其长于恒星年是因为近日点每年东移 11 秒。**回归年**是太阳绕黄道一周经历的时间(也称为**太阳年**)，也是太阳直射点回归运动的周期: 365 日 5 小时 48 分 45.5 秒(365.24220 日)，比恒星年短，这是地球公转和地轴运动结合的结果: 因月球、太阳和行星的引力，椭球形地球的自转轴绕黄轴缓慢移动(相应的春分点沿黄道以每年 50.24 分速度西退，约 71 年西移 1°，约 25800 年移动一周)，即岁差现象。这导致二分二至地球位置在公转轨道上不断西移的动点，导致地球公转一周不等于太阳直射点纬度变化一周。

为了在电磁学，光学，热物理学等其它领域中更便捷地进行相应的计量比较，1960 年第 11 届 CGPM 决定将经典 SI 计量基准扩展到电磁学，光学，热物理学，化学等领域中，为此新定义了的四个物理量的单位基准，由于这些新定义的单位相对于原来的三个单位都是导出单位，因而它们的定义都是基于实验条件所限定的现象，而不是实物原器：

电流：在真空中相距 1 米的两根无限长平行直导线，通以相等的恒定电流，当每米导线上所受作用力为 2×10^{-7}N 时，各导线上的电流定义为**一安培**，记为 A。

温度：温度定义包含两点：一是定义温度零点，SI 单位制中的温度零点称为**绝对零度**，二是某一个具体的以实物的状态定义的温度，SI 单位制中将纯水的**三相点**定义为 273.16 K。其中的 K 为标记温度的计量单位**开尔文**的符号。

物质数量：12 克的碳-12 中所含的原子个数称为**一摩尔**，记为 mol。这个数值称为阿伏加德罗常数 $N_A = 6.0221367(30)\times10^{23}$)。值得注意的是，这样定义的阿伏加德罗常数不是一个被严格定义的数，而是依赖于测量的有误差的数。。

亮度：频率为 540.0154×10^{12} 赫芝(Hz)的单色辐射光源(黄绿色可见光)，在某给定方向的辐射度为(1/683)W/sr 时的发光强度定义为**一坎德拉**，记为 cd。

量子计量基准(原子单位)

实物计量基准存在两个问题：(1)实物原器的宏观尺度决定其难以精确复制，另外其受到外部各种“微小”扰动的影响在短时间可能无足轻重，但长期累积导致的漂移数量难以控制和预测。例如当时制作千克原器时只考虑了铂铱合金具有不易氧化的优点，但其表面仍会吸附一些人眼无法察觉的气体分子和其他杂物，甚至内部也会吸附氢气等气体，通过清洗虽然可以减少吸附的杂物，但反复污染和清洗使千克原器在 100 年内变化了大约 5×10^{-8}。(2)存在因难以预测因素遭受意外破坏等风险。

20 世纪 50 年代，得益于原子物理学的进展，人们发现孤立原子所发射的光谱线波长和频率表现出异乎寻常的精确性和稳定性，并由此提出了更为理想的长度和时间计量标准。这种标准称为**原子计量标准**，或**量子计量标准**。

新的计量标准是一种由实验仪器装置所建立的局部环境和该环境中特有的物质样品所呈现的特定现象。而且与原子运动状态相关的许多特性参量的变化是不连续的。由于不连续变化的物质状态的复现则相对容易，对微弱的外部干扰也不容易受到影响，从而避免了弱干扰导致的缓慢漂移，提高其稳定性和准确度。

1960 年国际计量大会通过了第一个量子计量基准是以氪原子同位素 ^{86}Kr 在两个特定能级($^2P_{10}\rightarrow{}^5d_5$)之间发生量子跃迁时所发射的橙色光谱线的波长 λ_{Kr} 作为长度基准[11]：

1 米 $= 1{,}650{,}763.73\lambda_{Kr}$

该基准的相对误差为 4×10^{-9}，其精度比“米原器”提高了 2 个量级。

1967 年的第 13 届国际计量大会通过了第二个量子计量基准：以铯原子同位

---

[11] 19 世纪末在实验中发现，镉(Cd)有一条红色光谱线波长在 15°C 的干燥空气中有非常好的清晰度和复现性。国际计量大会在 1927 年决定用这条谱线作为光谱学的长度标准，并将米的长度定义为 1 米 $= 1553164.13\lambda_{Cd}$。以后发现 ^{86}Kr 的橙色光谱线比镉红线更优越。因而在 1960 年的第 11 届国际计量大会上，又确定用 ^{86}Kr 橙线代替镉红线作为长度标准。

素 ^{133}Cs 基态的两个超精细能级间跃迁所对应的辐射周期 τ_{Cs} 作为时间基准:

1 秒 $=9{,}192{,}631{,}770\ \tau_{Cs}$

该辐射对应的频率为 9,192,631,770 Hz，每 10 万年误差不超过 1 秒。随着新技术的应用和改进[12]，近年来铯原子钟的精度已达到 10^{-16} 量级(相当于上亿年才相差一秒)，是目前最成功的量子计量基准(比地球运动的稳定性高了 6~7 个量级)。

随着对微观尺度物理现象研究的不断深入，有关电磁学的量子计量标准也出现了突破，特别是 20 世纪下半叶所发现的两种量子电子效应：1962 年发现的**约瑟夫森效应**和 1980 年由德国科学家冯·克利青发现的**量子霍尔效应**。

约瑟夫森效应：如果在由一极薄绝缘层隔开两块超导体构成的特殊器件两端施加一直流电压 V，该器件中会发射频率为 ν 的电磁波，该频率 ν 与器件两端的直流电位差 V 满足关系

$$V=\frac{h}{2e}\nu$$

其中 h 和 e 为两个物理常数，分别称为普朗克常数和电子电量。这意味着可以通过测量辐射频率推算出电压值，则测量精度与频率相当。经典电压标准的精度为 10^{-7}，而频率的精度量级为 10^{-13} 量级，这就将电压标准的准确度提高了 6 个量级。

量子霍尔效应：约束在半导体薄层内流动的电流受到垂直电流方向的强磁场时，薄层内垂直电流的方向会产生电压，低温下该电压与磁场强度之间的关系曲线呈现出台阶式的变化，台阶处的霍尔电阻(即霍尔电压除以电流)的大小为

$$R_H=\frac{V_H}{I}=\frac{h}{e^2 i}\qquad (i=1,2,\ldots)$$

即霍尔电阻只取决于一个基本物理常数 h/e^2 除一个正整数。利用量子霍尔效应建立的量子电阻标准，其精度可达 10^{-10} 量级，比原来传统的标准电池和线绕电阻一类实物基准的精度 10^{-7} 提高了 3 个数量级。

1990 年 1 月 1 日开始实行约瑟夫森量子电压标准和量子化霍尔电阻标准，电学计量由此也进入了量子计量阶段。

在 SI 制规定的七个基本单位中，有两个(安培和摩尔)和质量单位有关，其中对阿伏加德罗常数的测量误差大部份来源于标准“千克”的精度。但质量单位“千克”是目前在量子计量基准中进展最慢的，也是最后一个仍在使用的实物基准的 SI 单位。从计量精度而言，取代千克原器的新基准的不确定度至少需达到 2×10^{-8}。目前正在研制的量子“千克”测量方案大致分成“原子计数”和“电子测量式”两种。

“原子计数”方案是确定一个给定体积的单质物体中原子数后由该类原子的原子量求得物体的质量[13]。这种方法的困难在于至今还没有有效的精确计数原

---

[12] 20 世纪 70 年人们认识到重力时间膨胀会导致不同高度的原子钟有不同的秒，每个原子钟必须修正为平均海平面高度以取得一致的秒(大地水平面的自转约改变$\times10^{-10}$ 的秒长，1977 年开始修正并在 1980 年制度化)。用相对论的术语来说，秒定义为转动的大地水平面上原时。1977 年，BIPM 会议在有关定义中加入新的陈述：铯原子在绝对零度时是静止不动的。修正过的定义似乎暗示理想的原子钟将只有静止的一个铯原子发射出单一的频率。无论如何，该定义意味着在那些原子钟之内的运作和外推的数值，高精度秒应考虑周围温度的补偿

[13] 该方法本质上与确定一摩尔物质中的原子数，即阿伏加德罗常数 N_A 等价。如一摩尔碳同位素 ^{12}C 的质量定义为 12 克给出一个 ^{12}C 原子的质量为 m = 12 克/N_A。清点一摩尔原子中个数的最便利方法为在原子规则排列的晶体中进行。现有的最佳数据给出 $N_A=6.0220\times10^{23}$，

子的方法。现行的方案是测定物质处在晶体状态时的体积和晶格常数，由此推算出物质所含有的原子数目。这种计量方式的主要误差来自实际晶体的不完整性，这种不完整性分为三类：一是晶体内部含有的位错、空穴、杂质原子等缺陷的数量，这导致实际测得的晶体体积并不等于理想硅晶体的体积；二是晶体表面通常吸附有大气中的分子的数量，甚至形成氧化层，这导致质量测量值的误差。三是晶体内共存的同位素原子的含量，而同位素丰度的精密测量也很困难。另外，凝聚成晶体的若干个原子的质量和相同数量的同种自由原子的质量存在区别。

随着以半导体为基础的电子工业的发展，通过硅材料的提纯和拉单晶技术已经能够得到非常理想的硅晶体，人们通过同位素提纯的方法得到了基本上只包含同位素 ^{28}Si(自然界中主要存在三种硅同位素：^{28}Si，^{29}Si 和 ^{30}Si)的硅单晶材料制成一个直径约 10 厘米的非常完美的球体，并准确测出硅球的尺寸(实际上是从多个方向测量球体的直径)，再用 X 射线干涉仪测出硅晶体的晶格常数(硅晶体中位于晶格点阵上的硅原子间的距离)计算出硅球所含有的原子数量。这样定义的质量单位建立在原子质量的基础上，不再是某一具体实物的特性，不会随时间发生量值的漂移，而且可以按同一原理复制而得到相同的质量量值。有着和前面所述的量子计量标准同样的优点。图 1-2-8 所示为德国柏林硅晶体研究所制造出的目前世界上最纯 ^{28}Si 单晶球，其不确定度为 6.8×10^{-8}。

图 1-2-8 最纯硅单晶球

“电子测量式”方式又称为“瓦特天平法”，是用磁场和电功率来定义千克。实际上是将质量和能量联系起来以避开对原子记数中的困难：利用爱因斯坦的质-能关系式 $E = mc^2$，用等效能量来定义质量。有兴趣的读者可以自己查阅有关的测量细节。目前这一方法的测量精度也未达到 10^{-8}。

基本物理常数基准

量子计量精度的主要设计是基于频率测量的高精度，通过某些特殊物理效应将待测量的物理量与频率联系起来，通过频率测量获得待测物理量的高精度。这种方式中所含的误差很大程度上取决于待测物理量与频率的函数关系中的物理常数的精度，如真空中的光速，电子的电荷量，阿伏加德罗常数等。计量科学目前的发展方向是将这些基本单位的定义建立在基本物理常数或者自然量子现象的基础上。将一些原来依赖于实验测量的基本物理常数定义为一个确定的数值。

首先基于基本物理常数定义的计量单位是“米”：得益于激光技术的进步，激光器发射的激光频率的稳定性和精度达到了 10^{-15} 量级。1983 年第十七届国际计量大会将 80 年代用激光测得的光在真空中的传播速度

不确定度约百万分之五。原子质量也可用彼此相对大小描述，其精度比用千克要高得多。因而原子质量常用原子质量单位(u)来表述。例如，一个 ^{12}C 原子质量被赋予标准值 12 u。$1u \approx 1.66057\times10^{-27}$ kg。

$c = 299792458$ 米/秒

定义为一个基准常数，并由此将长度“米”定义为

1 米=光在 1/299,792,458 秒的时间间隔内在真空中传播的距离

即从测量属性而言，c 不再是一个由观测时间和长度得到测量值，而“米”则成了一个导出单位。

这样的定义是基于两个考虑：按照爱因斯坦的相对论，光在真空中的传播速度是不变的基本常数，确定这个常数的数值也就意味着确定了一种度量单位，而且比长度作基本单位具有更好的普适性[14]。二是将长度单位统一到时间上，其测量精度就自动提高到时间的测量精度。例如，按量子基准，米的精度为 10^{-12}，而目前利用光钟技术已经将时间基准的精度提高到 10^{-16} 量级，因此物理常数基准使米的精度也达到了 10^{-16}。

在不久的将来有可能采用的物理常数计量基准及相关的关系为

(1) “摩尔”，“千克”和阿伏伽德罗常数 N_A

将阿伏伽德罗常数定义为 $N_A = 6.022141527\times10^{23}$，则“摩尔”就可明确表述为：$6.022141527\times10^{23}$ 个相同的原子或分子，而“千克”则定义为：

$(6.022141527\times10^{23}/0.012)$个自由的 ^{12}C 原子处于基态时的质量。

(2) “安培”与基本电荷常数 e

将基本电荷的数值定义为 $e = 1.60217653\times10^{-19}$ 库仑，则电流单位“安培”的定义表述为：在流动方向每秒通过 1.60217653×10^{19} 个基本电荷。

(3) 普朗克常数 h 与电磁学测量

将普朗克常量定义为 $h = 6.6260693\times10^{-34}$J·s，结合上面的基本电荷数值 e 的定义，可通过约瑟夫森常量 $K_J = 2e/h$ 及冯·克利青常量 $R_K = h/e^2$ 确定电压标准及电阻标准。另外，通过“瓦特天平”，可将“千克”的定义表述为：使普朗克常数等于 6.6260693×10^{-34}Js 的质量单位。

(4) “开尔文”和玻尔兹曼常量 k_B

玻尔兹曼常量 $k_B=1.3806503(24)\times10^{-23}$J·K^{-1} 和温度的乘积在 SI 制中的单位是能量单位焦耳，而温度单位开尔文一直是按纯水的三相点定义和导出的单位，但纯水的化学纯度几同位素效应都会给该单位的实际测量带来误差。因此用玻尔兹曼常量导出热力学温度单位的可行性也开始得到重视。

通过定义基本物理常数可以提高物理量的单位的表述准确性和基准精度，但因为上述这些物理量造在测量属性上的关联也带来了一些问题。例如，对定义质量单位“千克”的实验装置“硅球”和“瓦特天平”所涉及的基本物理常数分别是阿伏伽德罗常数或普朗克常数。由于普朗克常数与微观世界中几乎所有现象相联系，特别是再同时固定基本电荷 e，将使约瑟夫森常数和冯·克里青常数都得到固定，这对广泛使用的电学计量有利。因此，采用阿伏伽德罗常数还是普朗克常数定义“千克”就需要仔细斟酌。再例如，定义了基本电荷 e 也就确定了电流单位安培。但原来的安培是定义在固定真空中磁导率 μ_o 基础上的，固定了 e，μ_o 只能通过实验确定，从而带来测量误差(相同的情况还有电常数 ε_o 和真空中的特征阻抗 Z_o)。而磁导率 μ_o 是真空的一个性质，应比基本电荷 e 更为普适，因而固定基本电荷似乎并不比固定磁导率更有优势。

---

[14] 按照爱因斯坦的相对论，长度的测量与参考系有关，不是绝对的，而真空中的光速不依赖于参考系，是绝对的。

其它常用的基本物理量单位

除了历史的原因，某些国家和地区仍在使用不同于 SI 制的单位以外，自然科学中的许多领域也仍在使用一些特殊的单位，其主要原因是便利，它包含两个方面：一是大小适配，二是更能反映相关实验操作的测量属性。例如，一般的电磁学实验中 SI 制单位太大，通常使用的是以厘米(cm)、克(g)、秒(s)为基本单位的高斯单位制。而在天体物理领域，SI 制单位对于天体质量及天体间的距离太小，通常以太阳质量为质量单位，对太阳系中的天体间距离以太阳和地球的平均间距作长度单位，称为一个**天文单位**(A.U. = 1.49597870×10^{11} 米)，而对太阳系中的天体间，天文单位还是太小，通常用光在一年中传播的距离作为计量单位，称为光年(1A.U. = 1.58129×10^{-5} 光年)。能直接反映相关实验操作属性的物理量单位，其含义就更为清晰准确。如遥远天体与我们的距离通常是用所谓的**周年视差**方法确定的：当天体与地球的连线垂直于地球的轨道半径时，该天体对日地平均距离 a 的张角 θ 称为该天体的周年视差(图 1-2-9)，因 θ 的角度值通常小于 1°，恒星与太阳的距离可表示为 $r = 206{,}265a/\theta$。天文学中直接用测量值 θ 作长度单位，称为**秒差距**(pc)。如果恒星的周年视差为 1 角秒(1°的 1/3600)，它距离我们 1 秒差距，约为 206,265 天文单位。再例如，在高能物理领域，通常通过加速带电粒子(如电子，质子)使其获得很高的能量后与其它粒子碰撞来观测这些粒子的特性，或产生新的粒子。因此通常用 1 电子伏(eV)表示 1 个单位电荷 e 的粒子在电位差为 1 V 的电场中加速所得到的能量作为能量单位。

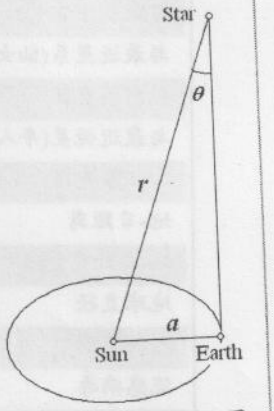


图 1-2-9 遥远恒星的周年视差。

1-2-3 寻求基本物理量测量精度的意义

许多初学者可能不理解为什么物理学家们对这些基本物理量的测量精度有如此吹毛求疵般的苛求，甚至怀疑这是否为从事该领域工作的癖好或态度。事实上，对测量精度的精益求精主要基于以下三个方面的意义和需求。

(1) 通常的直接测量考虑的只是相对误差。许多场合中达到 10^{-3} 就足够令人满意。但有些场合关注的是绝对误差，它会随着观测量值本身的增加而逐渐累加，最终甚至导致实际测量数据变得毫无意义。例如，中国建造的港-珠-澳跨海大桥分为海底隧道和大桥两部分，其中的连接处采用多个直径 22.5 米，高 55 米，重 550 吨的圆钢筒建筑人工岛，每个圆钢筒在大海中的定位要求绝对误差小于 2 厘米。在这种四周没有固定参照物，波涛汹涌的大海中定位只能依赖于全球卫星定位系统 GPS，通过接收由卫星发出的雷达信号，由信号发出和抵达的时间差确定物体的地面位置。读者可以估算一下如此小的地面位置误差所要求的卫星上的所必需时钟精度。

(2) 随着科学的发展进入物质的微观尺度，这些微观粒子不仅空间尺度及质量都非常小，所涉及过程的特征时间也非常短，这就要求在相应的实验观测中需要更小的空间和时间量尺，这意味着只有更高精度的量尺才可能揭示出物质和过程的细节，厘清物理过程中的因果关联。

(3) 探索微小尺度的空间和时间是否与我们所熟悉的宏观尺度的空间和时间有着相同的结构和性质，必须有物质实证才能确认。近代物理对时间和空间的认识的深入影响着我们对测量精度极限的认识，以往我们认为，受制与于技术手段的测量误差会随着技术的进步无限制地趋向于零。但现有的描述微观物质世界的理论表明，对空间和时间的测量都存在一种内在限制，即存在一个最小的可测量长度和可测量时间间隔，分别称为普朗克长度和普朗克时间。虽然它们远远小于目前所能达到的测量精度。但它揭示出我们目前对时-空的理解远不是完善的。

表 1-1 给出了一些典型物理系统的尺度、质量数量级以及一些典型的物理过程的时间数量级。

表 1-1 典型物理系统的尺度、质量数量级以及典型物理过程的时间数量级

物理系统	尺度(米)	物理过程	时间(秒)	物理系统	质量 (公斤)
可观测宇宙	10^{26}	宇宙年龄	4×10^{17}	可观测宇宙	1×10^{53}
与最近星系(仙女座)距离	10^{22}	太阳系年龄	1.4×10^{17}	银河系	4×10^{41}
银河系直径	10^{21}	人类寿命	2.2×10^{9}	太阳	2×10^{30}
与最近恒星(半人马座)距离	10^{17}	太阳日	8.6×10^{4}	地球	6×10^{24}
太阳系直径	10^{13}	自由中子寿命	5×10^{2}	轿车	1.5×10^{3}
地-日距离	10^{11}	人类心跳周期	8×10^{-1}	人类身高	70×10^{0}
地-月距离	10^{9}	可听音最小周期	10^{-4}	雨滴	2×10^{-6}
地球直径	10^{7}	μ粒子半衰期	2.2×10^{-6}	尘粒	$\times10^{-13}$
人的高度	10^{0}	标准铯钟周期	10^{-10}	红血球	9×10^{-14}
流感病毒	10^{-7}	可见光波周期	10^{-15}	牛痘病毒	9×10^{-18}
氢原子	10^{-10}	光穿越原子的时间	10^{-19}	铁原子	9.5×10^{-26}
质子	10^{-15}	最短寿粒子的寿命	10^{-24}	质子	1.7×10^{-27}
普朗克长度	10^{-35}	普朗克时间	4×10^{-43}	电子	9.1×10^{-31}

1-2-4 超越直觉感知的空间、时间和质量的性质和相互关系

基于对空间，时间，物质质量的定量观测所形成的数学模型，我们不仅能够对这三个在逻辑上独立的基本物理事件要素进行更为深入的分析，预言某些超越我们直觉所能感知的性质，并设计相应的实验予以检验，还能回顾和反思这些基于最朴素的直觉感知给出的操作性定义，以及依据这些定义所建立的物理量之间的关联，在理性甚至抽象思维的基础上拓展这些定义及其内涵。

空间

我们对所居住的空间性质的最直接的感知可以综合为以下几个方面:

(1) 具有三个独立的方向;

(2) 具有平直性的几何结构;

(3) 均匀和各向同性的内在;

(4) 连续和无限延伸的。

空间的这些属性可以综合为一个数学模型，称为**坐标系**。从而可以通过数字来表示空间中的任意一点，并通过数学表示空间任意两点之间的相对关系。

从描述物体运动的直观和方便角度出发，可以构建出许多不同形式的满足上述要求的坐标系。其中最常用的是所谓的直角坐标系，又称笛卡儿坐标系[15](见图 1-2-11)：在空间内任意选取一点 O 作为空间基本参考点，任意选取三条通过 O 点的相互正交的有向直线，按这三条右手螺旋法则分别定义为 x, y, z 三个坐标轴，确定一个特定的长度为单位，以 O 为起点对三条坐标轴进行标度。这样空间中任意一点 P 都可以通过比较与原点 O 以及与三个坐标轴的相对方位，用三个相互独立的实数构成的一个唯一数组(x, y, z)所描述，其中的参量 x, y, z 称为位置 P 的**坐标**。而 P 与坐标系原点 O 之间的关系则由勾股定理所定义的 P 与点 O 间的连线的距离

$$r = \sqrt{x^2 + y^2 + z^2} \tag{1-2-3}$$

以及 P 与 O 间的连线与三个坐标轴的夹角(α, β, γ)所确定的方向决定。其中角度与坐标的关系由方向余弦所定义

$$\cos\alpha = \frac{x}{r},\ \cos\beta = \frac{y}{r},\ \cos\gamma = \frac{z}{r} \tag{1-2-4}$$

以上两个关系式给出了符合我们直觉感知的空间的完整定量描述。

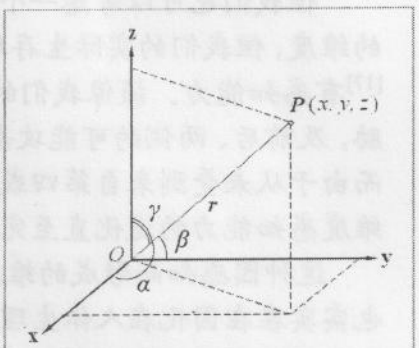


图 1-2-11 直角坐标系

在有些特定场合，为使描述更为简洁、直观，也可采用其它类型的坐标系对空间进行标度。例如，飞机驾驶员更愿意用经度，纬度和高度来描述自己在飞行途中的位置；而对机场指挥塔上的调度人员来说，每一架进出航空港的飞机的位置用方位角，距离地面的高度和相对指挥塔的水平距离来描述更为直观方便。在以上两个例子中，飞行员所用的坐标系类似于**球坐标系**[16]，而机场调度员所用的坐标系称为**柱坐标系**。

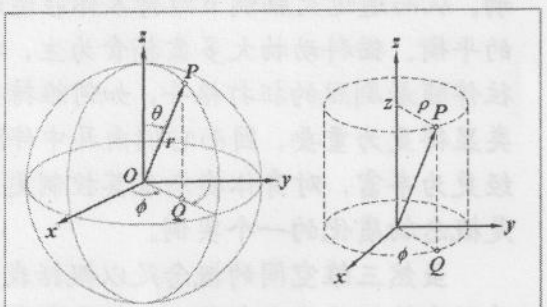


图 1-2-12 (a) 球坐标系，(b) 柱坐标系

在球坐标系中确定一个点的三个独立坐标为(r, θ, φ)，取值范围分别为 $0 \le r < \infty$，$0 \le \varphi < 2\pi$ 以及 $0 \le \theta \le \pi$ (图 1-2-12(a))。空间任意一点 P 的位置在球坐标系中的坐标与笛卡儿坐标系中的坐标之间的相互关系为：

$$x = r\sin\theta\cos\varphi,\quad y = r\sin\theta\sin\varphi,\quad z = r\cos\theta \tag{1-2-5}$$

在柱坐标系中确定一个点的三个独立坐标为(ρ, z, φ)，取值范围分别为 $0 \le r < \infty$，$0 \le \varphi < 2\pi$，$-\infty < z < \infty$ (图 1-2-12b)。空间任意一点 P 的位置在柱坐标系中的坐标表示(r, φ, z)与笛卡儿坐标系中的坐标表示(x, y, z)之间的相互关系为：

$$x = \rho\cos\varphi,\quad y = \rho\sin\varphi,\quad z = z \tag{1-2-6}$$

同一个空间点在不同坐标系中有不同的表述，但不管采用什么坐标系，空间位置的独立坐标数不变，描述这一特性的参量称为空间维度，上述直角坐标系、

---

[15] 笛卡儿(René Descarte 1596-1650) 法国哲学家、数学家、物理学家。因将几何坐标体系公式化而被认为是解析几何之父，对现代数学的发展做出了重要的贡献，他是西方现代哲学思想的奠基人，他的哲学思想影响了几代欧洲人，开拓了所谓"欧陆理性主义"哲学。

[16] 球坐标系的径向从原点 O 开始量度，而飞行坐标系的则以地面为高度的零点。

球坐标系、柱坐标系中的独立坐标数都等于 3，表示相关的空间是 3 维的。需要说明的是：人的视觉、听觉以及触觉可分辨出空间有上下，左右以及前后的区别，但由此推论空间是三维的则是一种信念(而不是基于充分的实验证据)：科学描述应该经济简约(这种信念也称为科学中的美学原理)，三维空间对我们所感知的物质存在以及观察和描述物理过程是必需的，同时也是充分的。或者说，二维空间不足以容纳及解释我们所观测到的世界，同时我们也从未感受到什么物体需要放置在三维以外的空间维度中。

但我们也可以考虑一个有趣也很有意义的想法：也许自然界确实存在着更多的维度，但我们的实际生存环境决定了我们只是对实际空间中的一个三维**子空间**[17]有感知能力。摄像我们的祖先以狩猎为生，同时也受到来自天空，地下的威胁，及前后、两侧的可能攻击，由此我们对这三个维度的感知得到了充分的进化。而由于从未受到来自第四或第五维的猛兽，如剑齿虎的袭击，这导致人体对其它维度感知能力的退化直至完全丧失[18]。

这种因感知而形成的维度的概念不仅仅是我们头脑中的一种抽象意识，而且也实实在在固化在人体生理结构中。图 1-2-13 是人耳的内耳结构图，其中前半部分是三个形状大致相同的弧形软管(称为半规管)分别处于三个两两互相垂直的平面上，构成类似于如图 1-2-11 中的直角坐标系。每个半规管的一端有一个相对隆起的壶腹。壶腹内有壶嵴，其中的毛细胞在人以各种姿态运动时由于滞后作用会向不同方向发生弯曲，并由传入神经向中枢释放神经冲动。从而通过大脑调节维持人体在运动时在空间中的平衡。猫科动物大多靠捕食为生，捕食过程中往往伴随着剧烈的扭打格斗，如何维持身体平衡比人类显得更为重要，因而它们内耳中伴随半规管的神经更为丰富，对身体姿态及其控制更为灵敏。这也是概念物质化的一个实例。

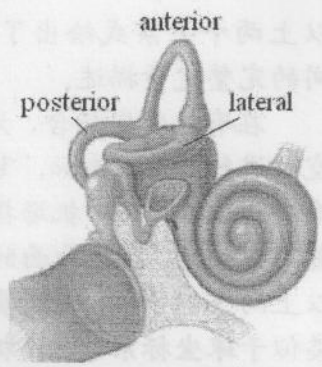


图 1-2-13 人内耳中的三个弧形半规管构成的三维直角坐标系。

虽然三维空间的概念足以概括我们所具有的全部感官的认知，但这并不意味着所有物理系统也都能、或只能感受到空间的三个维度。例如，我们可以制造出一些电子器件，使约束在其中运动的电子感受到的空间是严格二维甚至一维的[19]。这样，对这些系统的性质和行为的完整描述就涉及到具有不同维度的空间。

另外，如何能够更为简洁地描述一个物理系统的性质和行为也是影响我们对空间维度认识的一个重要方面。由于所有描述物理系统的状态和性质的物理量都通过测量和空间，时间以及物质存在着逻辑上的关联，在不同维数的空间中讨论一个实际的物理现象或描述相关的物理过程，相应的数学描述会具有不同的形式。一些物理问题在低维空间中的数学表述非常复杂，甚至很难求解，但在较高

---

[17] 虽然我们经常会讨论一些低维度的问题，例如对于直线运动问题或平面运动问题，我们甚至在二维空间中讨论物体的三维结构。这些低维度空间在数学上称为三维空间的子空间。但那种讨论只是为了说明方便，我们心目中并不会因此认可一维或二维空间的实际存在的。

[18] 我们无法感知许多现象可能是能量强度太低，毕竟生命只能存在于很窄小的物理环境中。

[19] 这种器件称为二维量子器件，第 10 章中有比较详细的讨论。

维度的空间中却有很简洁的表述并且很容易求得解析解。有时甚至可以将物体的性质或行为表述为空间维数的连续函数，从而可通过高维度空间的数学表述来理解物体在三维空间中的行为特征。后面的章节中会涉及一些相关的实际问题。

对空间维度的认识对描述和理解物理世界具有非常重要的影响。在一些低维数的空间中，物体会因运动形式受到限制而丧失某些特性，也可能获得某种功能。例如，任何生物是无法在二维空间生存的，因为生物体最基本的新陈代谢需求，如进食和排泄，呼吸和血液循环都会导致生物体碎裂(见图 1-2-14)。而在更高维数的空间中，很多最基本的微观和宏观的物质世界的运行规律如万有引力和电荷之间的相互作用的表示需要重新认识，否则现有的物质运行秩序就会变得不稳定[20]。

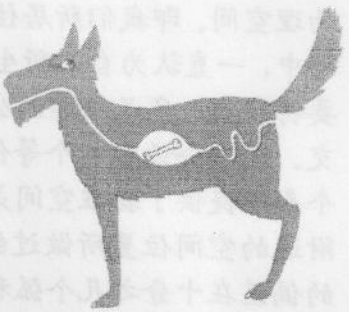


图 1-2-14 生物体新陈代谢功能导致二维生物体碎裂

事实上，在日常生活的有些场合中从不同的角度观察，描述对同一个物体时，我们对它的空间维数也会有不同的感知。例如对一条如图 1-2-15 中的带子所表示的曲线。从其构型上考虑，那显然是一个三维空间中的物体，因为至少在二维空间中不可能形成图中的那样一个结。但是从另一个角度我们也会感觉到这一条带子应该是一维的。它与我们平时所常用的一维坐标轴的差异在于它是弯曲的。考虑这个问题的一个有趣的视角是，如果生活在这条带子所决定的空间内将感受不到带子的弯曲构型，只有在三维空间内才能观测到这根带子全部的几何特征。这意味着空间的维度特征是通过其允许物体运动的自由度的数目表现出来的：限制在一维空间中的物体只有前后运动一种自由度，而不管该空间的具体构型如何；而在二维，三维的世界中，物体的运动自由度数目则分别为 2 和 3。显然对维度的这种理解的视角和讨论方式与我们前面所讨论的情况不同。我们一般说，图 1-2-15 中的这根带子本身是一维的，但构型是三维的，前者一般称为“**拓扑维数**”，而后者称为“**空间维数**”，但两者显然存在一定的关联。

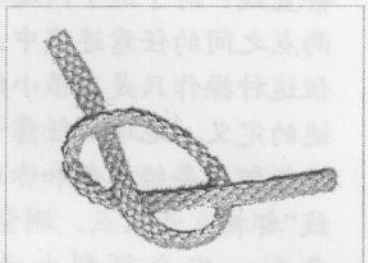


图 1-2-15 一条带子形成的一个结的构型的维度

空间维度还可以有其它的定义，这些定义除了符合我们的通常对空间的感知和认识以外，会有一些新的特性，可用于描述物质的一些特殊的结构和运动形式。例如德国数学家豪斯多夫[21]曾针对一种几何形态定义的维度可以不是整数，甚

---

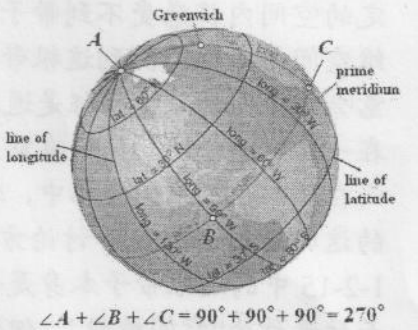
[20] 多于三维的空间中，两个物体之间的引力将随距离衰减得比在三维空间中更快(在三维空间内，如果距离加倍则引力减少到 1/4。在四维空间减少到 1/8，...)，这使绕太阳运行的行星的轨道变得不稳定，地球偏离圆周轨道的最小微扰(例如由于其他行星的引力吸引)都会引起它以螺旋线的轨道向外离开或向内落到太阳上去。而原子里使电子绕着原子核运动的电力行为和引力一样，这样电子或者从原子逃逸出去，或者以螺旋的轨道落到原子核上去。在任一情形下，都不存在我们所知道的原子。

[21] Felix Hausdorff (1868 年 11 月 8 日 – 1942 年 1 月 26 日)，德国数学家。拓扑学的创始人之一，对集合论和泛函分析都有不少贡献。曾以笔名 Paul Mongré 出版哲学和文学作品。纳粹当权后，他的抽象数学研究竟然被批评为属“犹太人”的，没用而且“非德国”。1942 年，为避免被被送往集中营，他与妻子服毒自尽。

至不是有理数，这种被称为豪斯多夫空间和豪斯多夫维度的概念在物理学中被用于描述许多物质结构和物理现象。读者可以阅读本章附录中的相关介绍。

空间的另一个重要性质是它的几何结构。对描述图 1-2-15 中的一条带子的特征所需要的空间维数的讨论会让我们将同样的问题引伸到我们自己所生活的物理空间。即我们所居住的三维空间是平直的还是弯曲的。人们在很长的一段时间中，一直认为自己所生活的空间是平直的。这种被称为**欧几里德空间**的一个重要特征由一条几何学的公设所描述或定义的：该空间中的两条平行线永远不会相交。这一公设有一个等价表述：一个平面三角形的内角之和严格等于 180°。这个表述提供了验证空间是平直的还是弯曲的实验方法。直至目前在地球上或地球附近的空间位置所做过的实际测量表明，一个平面三角形的实测内角和与 180° 的偏差在十分之几个弧秒之内，即在百万分之几的精度内可以认为我们所居住的空间是欧几里德的[22]。

但这里似乎有些问题并不明确，其中的一个问题是，这个三角形是如何在实验中被确认为处于同一个平面上的？一个等价的表述是：我们是如何在实验中实现三个顶点之间的连线为直线的？或者说应该如何从实验操作性的角度定义一条直线？对于这个问题的一个可能的答案是：直线是连接两点间的最短距离，即两点之间的任意连线中，长度最短的即为直线，这也是测量学中**测地线**的概念。但这种操作只是在很小的尺度上才是有意义的，如果测度尺度与地球相当，按上述的定义，地球的任意一条经线(经过两极的最短线段)就是一条标准的直线，由于任何两条经线都和赤道垂直，即所有的经线应该是互相平行的，但这些“平行线”都相交于极点。测量用两条经线和赤道这样的三条“直线”围成的三角形的内角和一定会得到大于 180° 的结果(见图 1-2-16)。就如前面讨论一条曲线的维数时所遇到的问题：被限制在一条曲线上(该世界中所有的物理系统只能呈现出一维空间的行为特征)生活的生物是无从知晓该曲线是否弯曲的。类似地，只有在离开地球一段距离的太空中观察，或借助于离开地面的物体参照，才能感知到地球表面是一个近似的球面。因此，只有在更高维的空间中才能观测到我们所生活的三维空间是否平直或弯曲。



$\angle A + \angle B + \angle C = 90° + 90° + 90° = 270°$

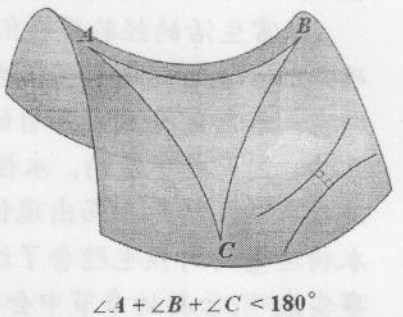
图 1-2-16 地球表面上由赤道和两根相差90°的经线构成的一三角形的内角和等于 180°

我们无法越出三维空间，因而只能考虑在三维空间内通过一些特殊的实验设计来测量我们所处的空间是否平直。但这样的结果会导致空间的几何结构与我们周围的物质分布有关这一结论。例如，我们通常默认一条绷紧的弦线一定是直线，即所谓的“木受绳则直”。但事实上一条绷紧的水平线不可能是直线，因为受到地球的重力作用，一条水平绷紧的弦线在一定程度上是悬挂线；而地面上一条悬挂着重物的铅垂线则是太阳系中的悬挂线。这是因为任何物质在宇宙中都会受到周围物质的引力作用。读者也许会想到用两点间传播的一束光的路径来定义直线，但已有的实验观测结果表明，光的路径是

22 一些基于相对论的见解证据表明欧几里德几何的误差为一百亿之几。

否弯曲以及弯曲的程度仍然与路径附近的物质多寡有关[23]。也许只有在更高维度的空间中，我们才能对自身所处空间的几何结构的观察摆脱对物质的依赖。

上述以三角形内角和大于 180°为标志的几何结构所决定的空间称为**黎曼空间**，相应的几何称为**黎曼几何**。实际上还存在着一种非欧几里德空间：以三角形内角和小于 180°为标志的几何结构所决定的空间，称为**罗巴切夫斯基空间**，相应的几何称为**罗巴切夫斯基几何**。在一种特殊的物质分布中可以实现这种特殊的弯曲空间。这种空间的平面特征类似于图 1-2-17 所示的马鞍面。在该平面上的三角形的内角和小于 180°。



$\angle A+\angle B+\angle C<180^\circ$

图 1-2-17 马鞍面上由三根直线构成的三角形的内角和小于 180°。

时间

由于时间是在不断地流逝，它应类似与空间也具有维度和几何结构。从我们对时间的直觉感受而言，时间只有一维。另外，我们对时间有一个不同于空间的直觉感受：我们可以在空间的任何一个维度中自由地前后运动，但在时间维度中我们只能持续前行。事实上，任何生物都在经历成长，发育，成熟，然后衰老，死亡这样一个演化方向。

我们观察到的时间具有方向性是基于时间的观测方式：我们是通过观测物体运动来量度时间的，而物体的运动受控于物质世界内在的因果关系。从表面上看似乎是时间的流逝方向决定了物质世界的演化方向，但实际上是决定物质演化的因果关系展示出了物体的运动方向，并由此让我们感受到时间的流逝。因此，并不是所有的物质运动都能让我们知道时间的方向。对于可逆运动，由于无法判断因果关系，也就无从知道时间的先后。例如，观看一个关于单摆在无耗散的情况下的摆动过程的记录影片是无法判断该影片的正反放映秩序的。我们习惯于时间具有方向性，是因为我们日常所观察到的宏观运动都是不可逆的。而可逆的微观运动则不为我们的感官所感受到。

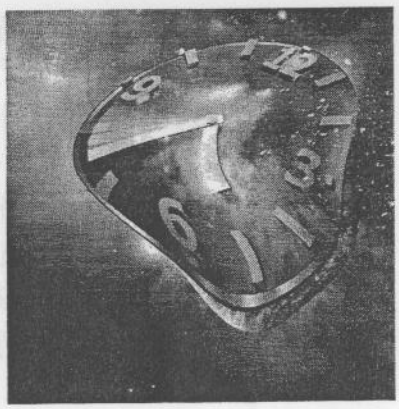


图 1-2-18 艺术家通过对钟的变形来描述时间的弯曲，其中的隐喻是：时间弯曲的概念必须通过空间的弯曲得到描述和理解。

由于时间是一维的，如果时间的概念是绝对的(即独立于空间和物质世界)，我们就无法知道它的几何结构，即时间轴是否是弯曲的。但我们对时间的观察是基于对物质运动的过程观察，这意味着因为时间与物质及物质运动的关联性而有可能发生弯曲。图 1-2-18 是艺术家通过对时钟的空间变形来描述时间的弯曲，该作品形象地表达了这样的概念：时间弯曲的概念必须通过空间的变形予以量度和描述。从目前的理论计算和相关的精确测量结果看来，这种

[23] 参见第四章

弯曲确实存在。

空间、时间和物质之间的关联

日常生活的经验使我们对于时间、空间和物质形成了一种直觉：在任何一个确定的位置和时刻只能发生一个事件，这意味着物质具有相互不可进入的性质，即空间和时间分别是各自的“点”的集合。或者说，空间、时间以及物质是各自独立的，因而是不变的，永恒的，绝对的。空间和时间的某个部分并不会因为其中是否有物质的存在而出现任何变化。我们在设定时间、空间和物质这三个量为基本物理量的时候也隐含了这一假定，并认为这是不言自明的事实。但越来越多的事实表明(后面的章节中会有所介绍)，这三者之间存在着相互关联。例如，时间的弯曲程度不仅与空间的弯曲有关，而且和物质的存在及物质的运动有关。在趋向于一些所谓的极限条件下，这些关联可以呈现出一些有趣的现象。例如，爱因斯坦的相对论理论提出的命题之一则是：当物体在强引力场中以接近光速的速度运动时(比如接近黑洞等等)，观察到的时间也会随之发生变化。

思考题

1-2-1 理解一个物体的特性一般应从哪些方面入手？

1-2-2 一个隐身物体是否能完全隐身？想象一个物体要隐身会涉及那些因素。

1-2-3 探讨一下人眼不能分辨单色和复色这一现象的原因，能否建立一个简单的模型来汇总你的设想和结论。

1-2-4 如何根据月球的运动导出月亮历(阴历)。

1-2-5 设想能否脱离运动来建立空间和时间的概念？

1-2-6 基本测量(直接对比)只依赖于长度，为什么时间测量却能比长度测量有更高的精度。

1-2-7 解决计量基准的稳定性和复现性这两个问题的方向是由离散型量替代模拟量，另外以物理环境基准替代实物基准，说明这两者在物理上是互相关联的。而且变化的离散程度越高，越容易复现，也越稳定。

1-2-8 原子的尺度远小于宏观物体尺度，受外部扰动的响应更为明显，为什么原子计量基准比实物原器计量基准更为稳定。

1-2-9 想象四位空间中的一个结在三维空间中是什么样的？

1-2-10 一个孤立星球上的铅垂线是直线吗？一个人能否用两只手绷直一根线？两个人在外空中能否绷直一根线？

1-2-11 为什么可以从电视荧光屏上看到的图像感受不到频闪，而可以从电视荧光屏上放映其它荧光屏的图像看到明显的频闪？

习题

1-2-1 找出两个空间点距离以及它们与原点的联线间的角度在球坐标系和柱坐标系中表达式。

1-2-2 1900 年，地球旋转一周所需要的时间的 86,400 秒，而在 2000 年，地球旋

转一周所需要的时间为 86,400.002 秒，计算到哪一年，你的生日会整整差一天?

1-2-3 1660 年，伦敦皇家学会曾将秒摆的摆长作为长度单位，约一米摆长的单摆，在地球表面的一次摆动(半周期)的时间约为一秒。讨论该时间单位的精度。

1-2-4 太阳系是银河系中的很小一部分，引入恰当的参量来描述太阳系的特征以及两者之间的关联。(太阳离银河系中心的距离，整个太阳系绕银河系中心转动。可见，太阳系不在宇宙中心，也不在银河系中心。太阳是 50 亿年前由星际云瓦解后的一团小云塌缩而成的，它的寿命约为 100 亿年。)

1-2-5 从中国东部城市上海到美国西部城市洛杉矶的民航飞行路线通常都不是沿图中所示的联接这两个城市的指向近乎正东的直航行程 **R**，的最短距离的路径，而是沿海岸线飞行(以防出现不测?)。粗略地说，首先沿东北飞越东京到达阿拉斯加的城市安克雷奇上空，然后折向东南飞往洛杉矶。如果将这两个位移分别记为 **A** 和 **B** (见本题图)，比较这两种航程的实际飞行距离的差异。

1-2-6 圆柱面上两点的最短距离如何计算。

1-3 物理量的表述

物理量虽然以数学量的形式出现，但其内涵并不等同于相应的数学量。数学量是定义在运算属性上的运算对象，它具有如下两个特性：(1)数学量在运算前后属于同一类数学量，即数学量在运算规则下是封闭的，(2)数学量(至少在定义上)是精确的，例如一个实数对应与数轴上的一个唯一的几何点。而物理量是定义在测量属性上的**测量对象**，测量操作就是与一个作为**标准**的同类物理量进行比较，测得的数值结果具有两个特点：(1)大小只具有相对意义(取决于单位的选择)，(2)具有一定的误差。这两个特点影响着物理量在比较和运算时与数学量的运算出有所不同。本节首先讨论物理量测量值的误差，相应的表示以及运算特点，然后以矢量为主要例子介绍以非单一数值作为物理量描述物理特性的意义和特点。

1-3-1 物理量的误差和表示

除了极个别物理量是定义的准确值(如真空中光速等)，绝大多数物理参量值都是测量值。由于受到测量仪器的精度制约，测量值总是带有一定的**观测误差**。

物理量观测值

考虑用一"量尺"量度(直接对比)某个物体的长度 L，通常 L 并不恰好是"量尺"上的最小分度(单位)a 的整数倍，对介于两个相邻最小分度值之间的剩余长度只能进行估读，因而会不可避免地产生相应的误差，误差的大小和 a 的大小有关。通常将量具上的最小分度值的一半($\frac{1}{2}a$)定义为测量值的**最大绝对误差**，而以 a 的十分之一作为估读单位(即将最小分度分为十份，估算测待测部分占几份)。例如图 1-3-1(a)和(b)中的两条棒的长度测量值分别记为 $15.0a$ 和 $15.2a$。这两个数字的最后一位都是超出量具精度的估计值。这两个数值都称为**有效数字**。这里特别说明有效数值 15.0 和 15 所表示的测量含义不同：15.0 有三位有效数字，小数点前的 5 是精确的，估读部分是小数点后的 0，而 15 只有两位有效数字，小数点前的 5 不是精确的，两个有效数的精度差一个数量级。

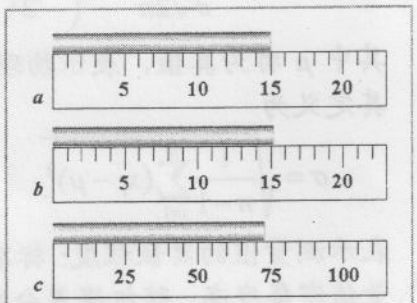


图 1-3-1 测量值估读及误差

测量数值 = 读数值(有效数字) +单位

有效数字 = 可靠数字+估计数字(一位)

上述测量值的估读规则对大多数采用十进制标度的测量仪器比较自然，如一般量尺上的最小分度是毫米，测量值的最后一位对应于 0.1 毫米的估读值。但这种规则对最小分度非十进制标度的量具会产生过多的有效位数，如用最小分度为 5(单位)的量尺测量长度(见图 1-3-1c)，按上述规则，图中的读数就是 $70+0.5\times5=72.5$(单位)，但其中第 2 位的 2 就是估读的，再估读第 3 位的 5 就完全失去了意义。所以测量值应为 72 (单位)。

物理量的科学记数法

一个数值的有效位数取决于具体的测量，而与表示所用的单位无关。例如，将一个用如下三种方法 632.8nm，0.6328μm，0.0000006328m 表示的长度测量值

是完全相同的。但对于给定的单位制，上述物理量的记数方法会涉及两个问题：书写、阅读的便捷和数值的精度。如太阳和地球的平均距离用国际单位制记为150,000,000,000 米，其中大量的"0"不仅书写、阅读不便，也无法辨识出有效位数。这两个问题可采用科学记数法一并予以消除：将数值写成 1 与 10 之间的实数(不包括 10)和 10 的整数幂的乘积。例如，太阳和地球的平均距离可以记为'1.50×10^{11} 米'，不仅形式简洁，也给出了有效位数。用科学记数法还能通过负整数幂给出一个给定单位的很小的数值的简洁表示，如 0.0000006328m 可表示为 6.328×10^{-7}m。

测量误差及其出现的规律

测量误差通常可以分为系统误差和随机误差两类。其中**系统误差**是由测量中的某些固定不变的因素引起的，如仪器的精度，测量的方法、测量者的观测习惯等。例如，一个秒表走得偏快，或者仪器或仪表的零点校准不准确，甚至观测者无意识的某种读数习惯，动态测量中的滞后效应等。各种系统误差有一个共同点，误差的大小及方向(符号)在同一组测量中具有一致性：统一偏大或偏小相同的数量或比例。只要找出导致误差的原因，就能这组对数据进行统一的修正。**随机误差**是由诸多不确定的偶然因素造成的，如果在相同条件下对同一物理量进行多次重复测量，每一次测量给出的误差大小，正负具有不可预知的不确定性，这类误差因此也称为**偶然误差**。随机误差产生的原因在于实验条件和环境因素无规则的起伏变化引起测量值围绕真值出现涨落。例如：电表轴承的摩擦力变动，螺旋测微计测力在一定范围内随机变化，操作读数时的视差影响等。

对各类实验测量数据所作的统计分析表明，随机误差也满足一定的规律：不同大小的随机误差出现的相对频度服从**正态分布**[1](见图 1-3-3)：

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) \tag{1-3-1}$$

其中 μ 称为**真值**，表示物理量在一定实验条件下的客观存在值。σ 称为**标准差**，其定义为

$$\sigma=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\mu)^2} \tag{1-3-2}$$

表示测量值的离散程度：标准差小表示测量值密集度高，随机误差分布范围窄，测量精度高；标准差大表示测量值分散，随机误差分布范围宽，测量精度低。由(1-3-2)式可以计算出任意一次测量值落入区间$[\mu-\sigma,\mu+\sigma]$的概率

$$P=\int_{\mu-\sigma}^{\mu+\sigma}f(x)dx=0.683$$

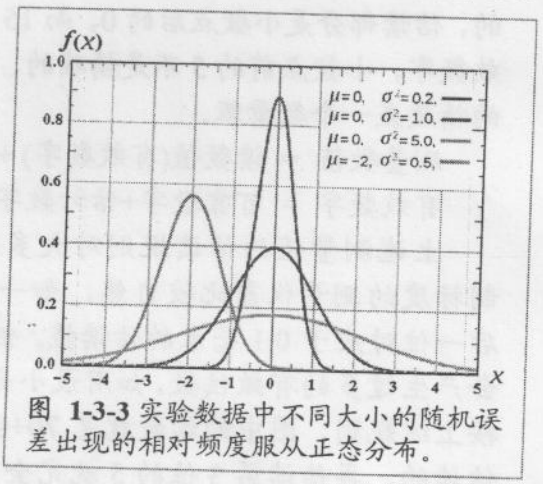


图 1-3-3 实验数据中不同大小的随机误差出现的相对频度服从正态分布。

[1] 由概率论的极限定理可以证明，若影响某一数量指标的随机因素很多，但每个因素所起作用不大的情形，这个指标服从正态分布。无穷多次测量时服从正态分布，n 较小时，偏离正态分布较多，n 较大时，趋于正态分布。

称为置信概率，$[\mu-\sigma, \mu+\sigma]$称为置信区间。扩大置信区间可增加置信概率，如

$$P=\int_{\mu-2\sigma}^{\mu+2\sigma} f(x)dx=0.954 \qquad P=\int_{\mu-3\sigma}^{\mu+3\sigma} f(x)dx=0.997$$

如果对一个物理量进行了 n 次测量，测得的数值为$\{x_1, x_2, \ldots, x_n\}$，可以用算术平均值

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n} x_i \tag{1-3-4}$$

作为真值 μ 的最佳估计值，测量次数 n 趋于无穷时，算术平均值趋于真值。n 有限时，算术平均值与真值的标准偏差为：

$$\sigma_{\bar{x}}=\sqrt{\frac{\sum(\bar{x}-x_i)^2}{n(n-1)}}=\frac{\sigma_x}{\sqrt{n}} \tag{1-3-5}$$

可以理解为待测物理量处于区间$[\mu-\sigma, \mu+\sigma]$内的概率为 0.683。

在重复试验中，可能会出现个别测量值过大或过小的异常情况。这些不正常的**可疑数据**通常是由实验操作不当造成的**过失误差**，如读错，记错数据等。可以利用上述**置信度的概念**予以剔除。例如，当试验次数较多时，可采用 3 倍标准偏差作为可疑数据取舍的标准：当某一测量数据 x_i 符合：$|x_i-\bar{x}|>3\sigma$ 就应予剔除。因为按正态分布，测量值落在$(x-3\sigma, x+3\sigma)$以外的概率仅为 0.27%(近 400 次试验中才有一次)，这种小概率事件在经典物理中通常认为是不可靠的。

1-3-2 物理量的运算

大多数的物理量是通过间接测量获得的，即通过对若干个物理量的测量值经过某种运算间接得到的结果。在不同的数值运算过程中，各数值所带的误差会以不同的方式累积起来，并影响最终结果的准确性，这一现象称为**误差的传递**。在不同的运算中误差具有不同的传递特性，因而在对这些有效数字的运算处理过程中要注意运算规则的合理应用，避免在运算过程中丢失数据的精度。

有效数字运算中误差的传递

我们首先从测量的角度讨论有效数字在初等算术运算中的传递效果。

(1) 加(减)运算对应于实验操作中的两个同类测量的累加效果，运算时首先要统一单位。设两个物理量在同一单位下的真值和测量值分别为 a_o、b_o 和 a、b，即 $a=a_o\pm\Delta a$、$b=b_o\pm\Delta b$，其中用正值 Δa 和 Δb 分别表示 a 和 b 的误差，则 a 和 b 两者的和或差及其误差为

$$a\pm b=(a_o\pm b_o)\pm(\Delta a+\Delta b)$$

a、b 相加(减)后的**误差数量级**取决于 Δa 和 Δb 中数量级较大的一个。

例：有效数字 1234 和 12.34 的误差数量级分别为 10^0 和 10^{-2}，这两个数的和的误差数量级为 1：1234 + 12.34 = 1234 + 12 = 1246。

两个有效数字相加(或减)时，其和(或差)的数值(用科学记数法表示较大的数，较小的数用与较大数的10的整数幂相同的幂表示)在小数点后应保留的位数与被加(或减)数和加(或减)数中小数点后位数最少的一个相同。

例：4.178 与 21.3 的和应表示为$(0.42 + 2.13)\times10 = 25.5$。

求和(或差)结果的数值精度可以理解为由两次用不同量具测量中精度较低

的测量决定。例如，测量两条棒的长度时所用的量具不同，它们的最小分度分别为 1 厘米和 0.1 毫米，测量结果分别为 1234(±1)mm 和 12.34(±0.01)mm。两条棒的总长度为 1234 + 12 = 1246 (mm)，数值精度为 mm。

(2) 乘(除)运算：考虑两个有效数字 $a = a_o \pm \Delta a$ 和 $b = b_o \pm \Delta b$ 的乘积[2]：

$$a \cdot b = (a_o \cdot b_o) \pm (|a_o|\Delta b + |b_o|\Delta a)$$

其中误差"$|a_o|\Delta b+|b_o|\Delta a$"的数量级由两项中数值较大的一个决定。为了估计 $a \cdot b$ 的有效位数，设 a 和 b 以及 Δa 和 Δb 的数量级分别为

$$a \sim 10^m, b \sim 10^n, \Delta a \sim 10^x, \Delta b \sim 10^y$$

则 a 和 b 的有效位数分别为 $m-x+1$ 和 $n-y+1$，$a \cdot b$ 及其误差的数量级表示为

$$a \cdot b \sim 10^{m+n} \pm (10^{m+y} + 10^{n+x}),$$

设 $m+y \geq n+x$，则 $a \cdot b$ 的误差取决于 $a\Delta b$。即 $a \cdot b$ 的有效位数为 $m+n-(m+y)+1 = n-y+1$，和 b 的有效位数相同。因 $m+y > n+x$ 等价于 $m-x+1 > n-y+1$，两个有效数字乘积的有效数字与原来两个数中有效数字较少的一个相同。

例：$a = 4.178$，$b = 10.1$

$$b\Delta a \sim 0.008 \times 10.1 \sim 0.08，\quad a\Delta b \sim 4 \times 0.1 = 0.4$$

$a \cdot b$ 的误差为 $a\Delta b \sim 4 \times 0.1 = 0.4$，因而 $a \cdot b$ 的精度在小数点后一位：

$$4.178 \times 10.1 = 42.2$$

其有效位数为 3，和 10.1 的有效位数字位数相同。

请读者自己证明，两个有效数字 a 和 b 的商

$$a/b = (a_o/b_o) \pm (|a_o|\Delta b + |b_o|\Delta a)/b_o^2$$

所保留的有效数字也与其中有效数字较少的一个相同。例如：$4.178/10.1 = 0.414$。

对更为复杂的有效数字的函数运算：$Q = f(q)$，设 q 所含的误差为 Δq，可以利用函数的微分公式

$$dQ = \left(\frac{df}{dq}\right)_o dq \tag{1-3-6}$$

计算误差的传递特性。读者可以由此验证下面的有效数字的计位规则：

(3) 乘方(或开方)：有效数字与其底的有效数字相同。

(4) 对数函数：运算后的尾数位数与真数位数相同。如：

$\lg 1.938 = 0.2973$，$\lg 1938 = 3 + \lg 1.938 = 3.2973$

(5) 指数函数：运算后的有效数字位数与指数的小数点后的位数相同(包括紧接小数点后的零)。例如：$10^{6.25} = 1.8 \times 10^6$，$10^{0.0035} = 1.008$

(6) 三角函数：运算后的有效数字位数按照角度有效数字而定。如：

$\sin 30°00' = 0.5000$，$\cos 20°16' = 0.9381$

间接测量的不确定度

设一个物理量 Q 是通过直接测量参量 $q_1, q_2, \ldots, q_n$，再由函数关系

$$Q = f(q_1, q_2, \ldots, q_n) \tag{1-3-8}$$

计算获得的间接测量值，其平均值可直接用各量平均值计算，则：

$$\overline{Q} = f(\overline{q}_1, \ldots, \overline{q}_n) \tag{1-3-9}$$

---

[2] 在乘(除)运算并不需要同一个单位表示，

类似于偶然误差的出现及大小都具有随机性，误差的累积也具有随机性：总的误差可能因为单向积累而持续放大，也可能因相互交错而在一定程度上抵消。但对于给定的物理量测量值中的误差分布和一定的运算，运算结果的误差分布也可以计算，这里只讨论误差传递给出的最大误差范围估计。

设(1-3-8)式中各 q_i 所含有的平均误差为 Δq_i，标准误差为 σ_i，利用函数的全微分公式

$$dQ=\left(\frac{\partial f}{\partial q_1}\right)_o dq_1+\left(\frac{\partial f}{\partial q_2}\right)_o dq_2+\cdots+\left(\frac{\partial f}{\partial q_n}\right)_o dq_n$$

其中$(\partial f/\partial q_i)_o$表示偏导数$(\partial f/\partial q_i)$在$(q_1, q_2,\ldots, q_n)$取值$(q_{10}, q_{20},\ldots, q_{n0})$时的数值。用 q_i 的平均误差 Δq_i 代替上式中的 dq_i，可得 Q 的最大累积误差为：

$$|\Delta Q|=\left|\frac{\partial f}{\partial q_1}\right|_o|\Delta q_1|+\left|\frac{\partial f}{\partial q_2}\right|_o|\Delta q_2|+\cdots+\left|\frac{\partial f}{\partial q_n}\right|_o|\Delta q_n| \tag{1-3-10}$$

相对误差为：

$$\frac{|\Delta Q|}{|Q|}=\frac{1}{|f|}\left\{\left|\frac{\partial f}{\partial q_1}\right|_o|\Delta q_1|+\left|\frac{\partial f}{\partial q_2}\right|_o|\Delta q_2|+\cdots+\left|\frac{\partial f}{\partial q_n}\right|_o|\Delta q_n|\right\} \tag{1-3-11}$$

而 Q 的标准误差应为：

$$\sigma_Q=\sqrt{\left(\frac{\partial f}{\partial q_1}\right)^2\sigma_{q_1}^2+\left(\frac{\partial f}{\partial q_2}\right)^2\sigma_{q_2}^2+\cdots} \tag{1-3-12}$$

(1-3-10)式和(1-3-11)式是讨论各种不同的参量关联中误差传递的基本公式。

例 1-3-1 一金属环的相关测量值为：内径 $D_1=2.889\pm0.004$cm，外径 $D_2=3.600\pm0.004$cm，高 $h=2.575\pm0.004$cm。计算金属环的体积。

解：体积和上述个参量之间的关系为

$$\overline{V}=\frac{\pi}{4}(\overline{D}_2^2-\overline{D}_1^2)\overline{h}=9.436(\text{cm}^3)$$

由间接测量的不确定度合成的体积测量误差为

$$u_V^2=\sum_i\left(\frac{\partial f}{\partial x_i}u_{x_i}\right)^2=\left(\frac{\partial V}{\partial D_1}u_{D_1}\right)^2+\left(\frac{\partial V}{\partial D_2}u_{D_2}\right)^2+\left(\frac{\partial V}{\partial h}u_h\right)^2$$

$$u_V=\frac{\pi}{4}\sqrt{(2D_1hu_{D_1})^2+(2D_2hu_{D_2})^2+\left[(D_2^2-D_1^2)u_h\right]^2}=0.08\,(\text{cm}^3)$$

实验结果表示：$V=9.44\pm0.08(\text{cm}^3)$，相对不确定度为：$\Delta V/V=0.8\%$。◀

代数规则和算法规则

以上所讨论的只是依次运算，在实际的测量数据处理过程中往往会涉及大量的数据运算(如多重求和一类)。由于误差在运算中不断累积，运算的最后结果可能会导致有效位数的全部丢失而完全没有意义，而这种情况在实际数值计算过程几乎无所不在，因而在处理数值前应分析可能出现有效位数丢失的可能环节，选择适当的计算方法和流程，以求得可能达到的最佳精度。

例如，1234+0.1234+…+0.1234 (共 1000 个 0.1234)，采用不同的结合可以得

到不同的结果，如果按顺序累加，得到的结果为 1234，而最符合实际的结果是 2468，它是通过对每一对 0.1234 求和得到 50 个 0.2468，再将每一对 0.2468 求和得到 25 个 0.4936...。

上述例子表明有效数字在和、差计算中并不总是满足一般代数运算的交换律和结合律。因此在数值计算过程中要避免相差很大的数直接相加，两个相近的数值的相减。它们都会导致有效位数的丢失。

特别需要注意的是现在许多场合都在通过计算机采用数值方法求解各种方程或模拟实际过程，其中涉及大量的数值运算，由于计算涉及到的数据字具有有限的位数，如果不关心中间数据的误差传递，就完全不能判断最后结果的可靠性。

对于物理常数而言还需要注意另一点：很多基本物理常数彼此并不独立，但它们又能被独立地测定(例如原子量 u 和阿伏加德罗常数 N_A)。为得出内部自洽的常数，通常对实验测量值和有关公式进行最小二乘法平差处理，以得出一组最佳值。由于实验测量不断更新，国际上隔几年就会作一次"常数平差"处理，并依此推荐一组数据。如 1973 年公布的质子质量为 $m_p = 938.296(27)\ \mathrm{MeV/c^2}$，2002 年发布的推荐值则为 $m_p = 938.271\,998(38)\ \mathrm{MeV/c^2}$。选取不同年度的物理常数推荐值时要注意自洽。

1-3-3 物理量的不同数学形式

对于给定的测量单位，物理量的一次测量操作得到的结果是一个实数。例如，我们所熟悉的距离，时间间隔，速率，体积、温度等，这种仅由一个实数确定的参量在数学中称为标量。这类数学量有一个直观的几何表示：任何一个实数值都唯一地对应于数轴上的一个相应的点，其大小由它与原点的距离表示。

但有些物体的特性仅用一个数值及单位并不足以给出其完整的描述，因而要用其它形式的数学量来表示。下面我们主要以矢量为例，介绍如何选择恰当的数学形式来描述我们所期望描述或量度的物质的特性和行为，并根据相关物理现象的测量属性(即最直觉的物理图像)讨论相应数学量的运算规则。

矢量及其运算

设想一位船长要从英国的港口城市朴茨茅斯跨越英吉利海峡前往法国诺曼底海岸的朱诺(见图 1-3-3)。两个城市相隔约 100 英里。船长仅依据这一行驶里程信息显然不足以确定这一航行，他还必须知道航行方向：朱诺位于朴茨茅斯南面略偏东。而向正东行驶相当的里程则是法国的北部城市加莱，这种偏航如果发生在所谓的 D 日(1944 年 6 月 6 日)无疑会灾难性地改变第二次世界大战的进程。船长在航海图上标出的位置的变化在物理学中用一个包含有大小和方向两种属性的物理量表示，称为"位移"。

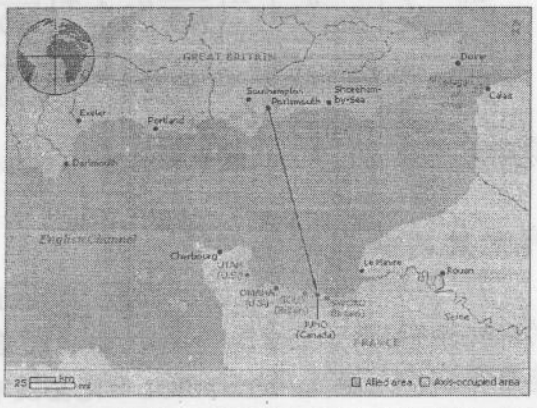


图 1-3-3 从朴茨茅斯到朱诺的跨海峡航行的位移矢量表示。

物理学中有不少类似于位移那样需要大小和方向两个参量才能完整描述的物质运动特性。如速度，加速度，作用力，电场等都属于这类参量，这类参量在数学中称为矢量。

矢量是定义在其运算属性上的，而物理量矢量的运算必须在具体的物理现象中反映出相应的测量属性及其合成效果。我们先从位移和一些比较熟悉的物理矢量的测量操作及其效果的直觉认知出发，引入矢量的运算图像，然后从解析的角度给出矢量代数的定义[3]。

矢量及其运算的几何描述

矢量的两个属性，大小和方向可以用一个带箭头的直线段作为几何表示，其中直线段的长度代表矢量的数值部分，而箭头的指向代表矢量的方向。这种表示没有规定矢量的具体空间，这意味着两个长度和方向相同但处于不同位置的矢量 **A** 和 **B**(见图 1-3-4)是恒等的。这一特点可以对矢量运算的操作性定义带来许多便利。

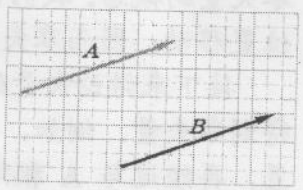


图 1-3-4 处于不同位置的两个相同矢量。

矢量加法

类似于两个标量的求和所表示的是这两个量的累积效果，两个矢量的求和 **R** = **A** + **B** 可通过对一个物体相继施行 **A** 和 **B** 两次位移形成的总位移 **R** 来说明。如图 1-3-5 所示，首先将(*a*)中的矢量 **A** 或 **B** 平移，使 **B** 的尾端与 **A** 的顶端重合。这样连接矢量 **A** 的尾端和矢量 **B** 顶端的矢量 **R** 就构成了物体历经 **A** 和 **B** 两次位移的净位移(见图 1-3-5-*b*)，也就是所谓的两次位移形成的总位移。由于矢量 **A**、**B** 与 **R** 构成三角形，相应的矢量求和法则称为**三角形规则**。

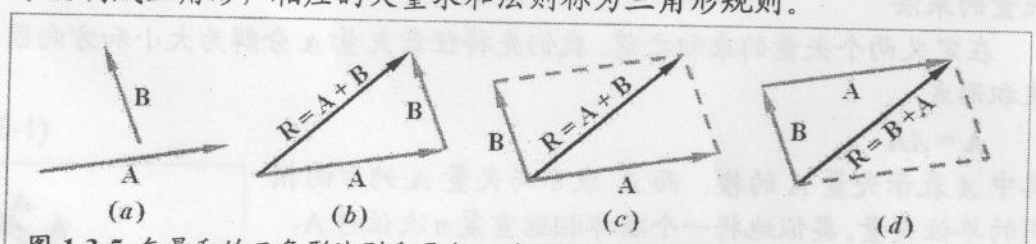


图 1-3-5 矢量和的三角形法则和平行四边形法则的几何表示。

图 1-3-5(*c*)给出了对矢量 **A** 和 **B** 求和的一种等价的操作：平移矢量 **A** 或 **B** 使 **A** 和 **B** 的两个尾端重合，显然，在 **A** 和 **B** 所确定一个平行四边形中分隔 **A** 和 **B** 的对角线矢量和图 1-3-5(*b*)中的和矢量 **R** 恒等。这种基于平行四边形对角线构成的矢量求和法则称为**平行四边形规则**。

类似于标量求和满足交换率和结合率，矢量求和也同样具有这些性质：

加法交换率：如图 1-3-5-*d* 所示，将矢量 **A** 平移，使其尾端与矢量 **B** 的顶端重合，这样由三角形规则构成的和矢量 **B**+**A** 与平行四边形规则构成的矢量和 **A** + **B** 重合，因而有

$$\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A} \tag{1-3-21}$$

[3] 我们希望这样可以让读者理解，数学虽然并不属于自然科学，但其中的各个分支的形成仍然是起源于实际的需求，而不纯粹是抽象的逻辑。

加法结合率：考虑如图(1-3-6)所示的三个任意矢量 **A**, **B** 和 **C** 的累加，总位移效果可以用三角形规则分别分解为 $\mathbf{R}=\mathbf{D}+\mathbf{C}=(\mathbf{A}+\mathbf{B})+\mathbf{C}$ 和 $\mathbf{R}=\mathbf{A}+\mathbf{D}'=\mathbf{A}+(\mathbf{B}+\mathbf{C})$。因而有

$$(\mathbf{A}+\mathbf{B})+\mathbf{C}=\mathbf{A}+(\mathbf{B}+\mathbf{C}) \tag{1-3-22}$$

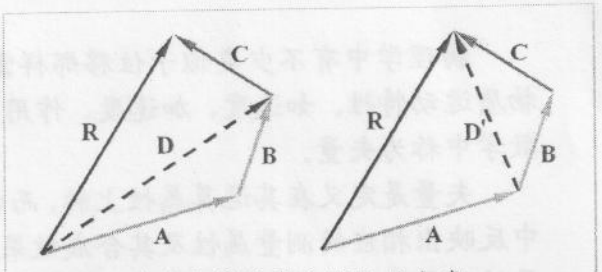


图 **1-3-6** 多个矢量的求和满足结合率。

矢量的减法

对一类数学量定义减法运算的第一步是确定零元素。对于矢量而言需要定义零矢量的概念。我们仍以位移矢量为例进行说明，零位移表示在任何方向都没有位置移动，因而零矢量的几何表示的就是一个几何点(没有尺度，方向不确定)。

接下来讨论矢量的减法运算。矢量减法的最简单例子是一个矢量减去它自身得到零矢量。例如，一个物体经历一个由位移矢量 **A** 描述的平移移动到某一位置，再经过与 **A** 相同距离但相反方向的位移使该物体回到原来的位置，这个使物体返回的位移矢量称为 **A** 的负矢量，记为–**A** (见图 1-3-7-*a*)，则：

$$\mathbf{A}+(-\mathbf{A})=\mathbf{0} \tag{1-3-23}$$

即一个矢量 **A** 减去其自身等效于矢量 **A** 加上其负矢量。这一运算规则可以推广到任意两个矢量 **A** 和 **B** 的减法：

$$\mathbf{A}-\mathbf{B}\equiv\mathbf{A}+(-\mathbf{B}) \tag{1-3-24}$$

其图形规则见图 1-3-7(*b*)和(*c*)。

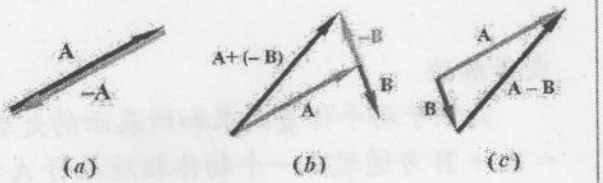


图 **1-3-7** 矢量减法的操作性定义 (a) $\mathbf{A}+(-\mathbf{A})=\mathbf{0}$; $\mathbf{A}-\mathbf{B}\equiv\mathbf{A}+(-\mathbf{B})$; (b) $\mathbf{A}-\mathbf{B}$ 的三角形法则。

矢量的乘法

在定义两个矢量的乘积之前，我们先将任意矢量 **A** 分解为大小和方向因子的乘积形式

$$\mathbf{A}=A\hat{\mathbf{A}} \tag{1-3-25}$$

其中 A 表示矢量 **A** 的模，而 $\hat{\mathbf{A}}$ 表示与矢量 **A** 的方向相同的单位矢量。类似地将一个物体相继重复 n 次位移 **A**，物体的总位移效果(见图 1-3-8)可以表示为

$$\mathbf{A}+\mathbf{A}+\ldots+\mathbf{A}=n\mathbf{A}$$

称为矢量的**数乘**。对于一般的矢量 **A** 而言，一个方向相同，但大小是|**A**|的 a 倍的矢量(a 是一个任意实数)，可以记为 $a\mathbf{A}$。称为标量 a 与矢量 **A** 的数乘。

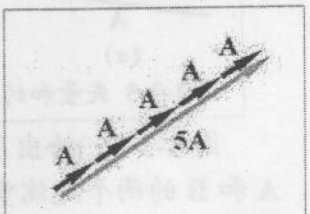


图 **1-3-8** 矢量数乘的操作性定义

对于两个矢量的乘积运算，考虑到两个矢量都有方向属性，因而我们首先面临着如何选择两者乘积的方向属性问题，或者说需要规定这是否是一个封闭性操作，即两个矢量的乘积是否仍然是一个矢量，或者是一个其它类型的数学量。另外我们还需要确定这种“乘法”中“模”的运算。上述这些选择和定义是以描述相应的物理效应为依据的。

标量积

考虑一个物理过程：一个物体在外力 **F** 的作用下经历了一个位移 **r**，如果这两个矢量方向不同，则外力 **F** 对物体做的功是一个标量，而其大小是力 **F** 在位移 **r** 的方向上的分量 F_r 与位移 **r** 的大小 r 的乘积。这样在计算上被分解为两部分：一是计算 **F** 的投影：$F_r \equiv |\mathbf{F}|\cos\theta$，其中 θ 为力矢量 **F** 和位移矢量 **r** 之间的夹角(图 1-3-9)，二是计算力 **F** 在位移 **r** 的方向上的分量 F_r 做的功：$W = F_r \cdot r \equiv |\mathbf{F}||\mathbf{r}|\cos\theta$。在物理上这对应于两个矢量单一的作用，因此可以定义一种乘积运算以统一反映上述物理效果，记为

$$W = \mathbf{F}\cdot\mathbf{r} \equiv |\mathbf{F}||\mathbf{r}|\cos\theta$$

称为力矢量 **F** 和位移矢量 **r** 的标量积。上述标量积运算规则可以推广到对所有的矢量：

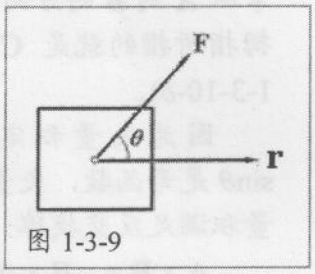


图 1-3-9

定义两个矢量 **A** 和 **B** 之间的一个运算—标量积

$$\mathbf{A}\cdot\mathbf{B} \equiv AB\cos\theta \qquad (1\text{-}3\text{-}26)$$

其中 θ 为矢量 **A** 和 **B** 之间的夹角。为避免歧义，规定其取值范围为 $0° \le \theta \le 180°$。由于标积符号用一个点来标记，矢量的标积通常又称为点积或点乘。矢量标积的一个很重要的应用是可以用来定义或给出自身的模：

$$A = \sqrt{\mathbf{A}\cdot\mathbf{A}}$$

显然，两个矢量 **A**, **B** 的标积反映了它们之间的线性相关程度，或者说这两个矢量是否包含共同的成分。这在几何图像上可以通过将一个矢量向另一个矢量投影，用得到的投影分量描述两者的线性相关程度。

因为标积定义中的因子 $\cos\theta$ 是偶函数，矢量 **A** 和 **B** 的标积满足交换率：

$$\mathbf{A}\cdot\mathbf{B} = \mathbf{B}\cdot\mathbf{A} \qquad (1\text{-}3\text{-}27)$$

矢量标积在物理学中有许多应用的例子。除了上面讨论的计算力 **F** 作用在一个物体上，在物体位移 **r** 的过程中这个力所做的功外，再如，一条磁棒在磁场 **B** 中的势能可以表示为 $W = -\boldsymbol{\mu}\cdot\mathbf{B}$，其中矢量 $\boldsymbol{\mu}$ 表示磁棒的磁矩。所有这些物理量都会在以后的章节中讨论。

矢量积

考虑另一个物理过程：将外力 **F** 作用在一杠杆上使之绕支点转动。日常经验告诉我们，即使作用力大小相同，只要力的方向或作用点不同，产生的效果是不同的：定义由支点到力 **F** 的作用点的连线矢量 **r**，称为力臂。如果 **r** 和 **F** 的夹角为 θ，能够对杠杆转动产生贡献的具有大小和方向两种特性，，效应最大，而相互平行时，效应为零。即只有和 **r** 垂直的 **F** 分量才产生这种效应。其数学表示为 $|\mathbf{F}||\mathbf{r}|\sin\theta$。若转动方向用由 **r** 和 **F** 决定的右手螺旋规定的转动轴来标志，描述外力作用的综合效果可以定义为 **r** 和 **F** 的一种运算：

$$\boldsymbol{\tau} \equiv \mathbf{r}\times\mathbf{F}, \qquad |\boldsymbol{\tau}| = |\mathbf{r}\times\mathbf{F}| \equiv |\mathbf{F}||\mathbf{r}|\sin\theta$$

称为 **r** 和 **F** 的矢量积。由此定义得到的新矢量 $\boldsymbol{\tau}$ 称为力矩。

上述运算规则可以推广到对三维空间中任意两个矢量 **A** 和 **B** 之间的一个运算—矢量积：

$$\mathbf{A}\times\mathbf{B} = \mathbf{C} \qquad (1\text{-}3\text{-}28)$$

由于矢量积的符号用“×”来标记，矢量积因此通常又称为叉乘。这意味着通过两

个不同矢量 **A** 和 **B** 的矢量积定义出一个新的矢量，新矢量的模定义为

$$|\mathbf{A}\times\mathbf{B}|\equiv AB\sin\theta \tag{1-3-29}$$

其中 θ 为矢量 **A** 和 **B** 之间的夹角，为避免歧义，规定其取值范围为 $0°\le\theta\le180°$。矢量积的模有一个明显的几何含义，如果将 A 理解为一个平行四边形的一条底边的长度，则 $B\sin\theta$ 可以理解为该平行四边形的高(见图 1-3-10-*a*)。则式(1-4-9)为矢量 **A** 和 **B** 所决定的平行四边形的面积。而矢量积所定义的新矢量 **C** 的方向由右手法则来确定：将展开的手掌从 **A** 到 **B** 的方向并拢四指，则拇指所指的就是 **C** 的方向(见图 1-3-10-*b*)。

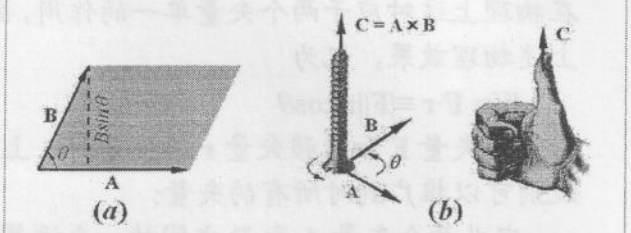


图 1-3-10 (*a*) 矢量积大小的几何意义；(*b*)矢量积的方向由右手螺旋法则定义。

因为矢量积定义中的因子 $\sin\theta$ 是奇函数，矢量 **A** 和 **B** 的矢量积满足反交换律：

$$\mathbf{A}\times\mathbf{B}=-\mathbf{B}\times\mathbf{A} \tag{1-3-30}$$

需要说明的是，矢量积运算仅仅在三维空间才有意义。因为对于二维空间的矢量而言，矢量积所定义的新矢量超出了原来的空间。读者可以自己分析更高维度的空间中定义矢量积所可能出现的问题(见思考课题 1-4-3)。

矢量积在物理学中有许多应用事例。例如，作用在一柄杠杆上的力矩可以表述为，作用在该杠杆上的力矢量和力的作用点相对于杠杆支点的位置矢量(力臂)的矢量积。以后的章节中会讨论到许多物理量矢量积的具体例子。

矢量及其运算的解析描述

矢量的解析(代数)描述是由法国数学家笛卡儿在 1637 年引入的。先选取一个直角坐标系，将空间中我们关注的位移矢量 **A** 平移，使其后端点与坐标系的原点重合。则 **A** 的前端点 P 相对于原点的位置矢量 **r** 就可用以表示位移矢量 **A**。按照矢量的加法运算规则，位置矢量 **r** 分解为坐标系 S 的三个直角坐标分量之和(见图 1-3-11)

$$\mathbf{r}=x\hat{\mathbf{x}}+y\hat{\mathbf{y}}+z\hat{\mathbf{z}} \tag{1-3-31}$$

其中 $\hat{\mathbf{x}}$，$\hat{\mathbf{y}}$ 和 $\hat{\mathbf{z}}$ 分别为对应于三个坐标轴的单位矢量，即位置矢量 **r** 可以唯一地用其顶端的坐标 (x, y, z)表示。需要说明的是，**r** 的三个互相独立的坐标分量并非必须利用(1-3-31)式连接起来，也可以通过定义互相之间的运算规则而视为一种不同类型的数学量，称为行矩阵[4]，其运算规则称为矩阵代数。

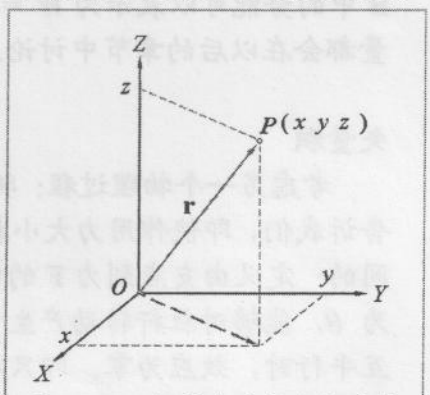


图 1-3-11 矢量在直角坐标系中的解析表示(坐标分量表示)

矢量加法与减法在直角坐标系中的解析表示

利用矢量的加法和减法满足交换律和结合律的性质，重新排列矢量各分量的

[4] 不熟悉线性代数的读者可以从下面的讨论中自己领悟出矩阵的基本运算规则。后面也有一些相应的习题使读者可以通过实际的物理问题来理解矩阵的一些性质。

位置可以将位置矢量 $\mathbf{r}_A$ 和 $\mathbf{r}_B$ 的和或者差分别表示为

$$\begin{aligned}\mathbf{r}_A \pm \mathbf{r}_B &= (x_A\hat{\mathbf{x}} + y_A\hat{\mathbf{y}} + z_A\hat{\mathbf{z}}) \pm (x_B\hat{\mathbf{x}} + y_B\hat{\mathbf{y}} + z_B\hat{\mathbf{z}}) \\ &= (x_A \pm x_B)\hat{\mathbf{x}} + (y_A \pm y_B)\hat{\mathbf{y}} + (z_A \pm z_B)\hat{\mathbf{z}}\end{aligned} \tag{1-3-32}$$

即两个矢量的"和矢量"的每一个直角分量为两个矢量的相应直角分量的和。显然，这里的矢量求和采用的是平行四边形法则。

矢量数乘在直角坐标系中的解析表示

利用标量代数中的乘法分配律，可以将矢量 $\mathbf{A}$ 与标量 k 的数乘表示为

$$k\mathbf{A} = k(A_x\hat{\mathbf{x}} + A_t\hat{\mathbf{y}} + A_z\hat{\mathbf{z}}) = (kA_x)\hat{\mathbf{x}} + (kA_t)\hat{\mathbf{y}} + (kA_z)\hat{\mathbf{z}} \tag{1-3-33}$$

标量积在直角坐标系中的解析表示

利用矢量的不同直角分量互相正交，标量积为零的性质，重新排列矢量各分量的位置得

$$\begin{aligned}\mathbf{A}\cdot\mathbf{B} &= (A_x\hat{\mathbf{x}} + A_y\hat{\mathbf{y}} + A_z\hat{\mathbf{z}})\cdot(B_x\hat{\mathbf{x}} + B_y\hat{\mathbf{y}} + B_z\hat{\mathbf{z}}) \\ &= \sum_{i,j=x,y,z} A_iB_j(\hat{\mathbf{i}}\cdot\hat{\mathbf{j}}) = \sum_{i,j=x,y,z} A_iB_j\delta_{ij}\end{aligned} \tag{1-3-34}$$

其中 δ_{ij} 称为克罗尼克符号，它可以理解为一种特殊函数，其定义为

$$\delta_{ij} = \begin{cases}1 & i = j \\ 0 & i \neq j\end{cases} \tag{1-3-35}$$

由此可得矢量 $\mathbf{A}$ 和 $\mathbf{B}$ 的标量积为

$$\mathbf{A}\cdot\mathbf{B} = A_xB_x + A_yB_y + A_zB_z \tag{1-3-36}$$

矢量积在直角坐标系中的解析表示

利用矢量的不同直角分量因互相正交，矢量积为±1，相同直角分量的矢量积为零的性质，重新排列矢量各分量的位置得

$$\begin{aligned}\mathbf{A}\times\mathbf{B} &= (A_x\hat{\mathbf{x}} + A_y\hat{\mathbf{y}} + A_z\hat{\mathbf{z}})\times(B_x\hat{\mathbf{x}} + B_y\hat{\mathbf{y}} + B_z\hat{\mathbf{z}}) \\ &= \sum_{i,j=x,y,z} A_iB_j(\hat{\mathbf{i}}\times\hat{\mathbf{j}}) = \sum_{i,j,k=x,y,z} \varepsilon_{ijk}A_iB_j\hat{\mathbf{k}}\end{aligned} \tag{1-3-37}$$

其中 ε_{ijk} 称为 Levi-Civita 符号(也称为置换符号)，它也可以理解为一种特殊函数，其定义为

$$\varepsilon_{ijk} = \begin{cases}0 & i,j,k\text{中有两个相同} \\ 1 & i,j,k\text{ 按}i,j,k\text{偶排列} \\ -1 & i,j,k\text{ 按}i,j,k\text{奇排列}\end{cases} \tag{1-3-38}$$

由此可得矢量 $\mathbf{A}$ 和 $\mathbf{B}$ 的矢量积为

$$\mathbf{A}\times\mathbf{B} = (A_yB_z - A_zB_y)\hat{\mathbf{x}} + (A_zB_x - A_xB_z)\hat{\mathbf{y}} + (A_xB_y - A_yB_x)\hat{\mathbf{z}} \tag{1-3-39a}$$

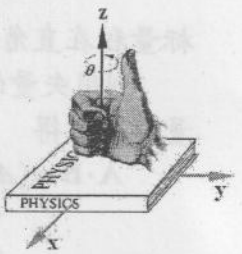
这个表达式可以更简洁地用行列式表示为

$$\mathbf{A}\times\mathbf{B} = \begin{vmatrix}\hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ A_x & A_y & A_z \\ B_x & B_y & B_z\end{vmatrix} \tag{1-3-39b}$$

坐标系的旋转，矢量的一般定义*

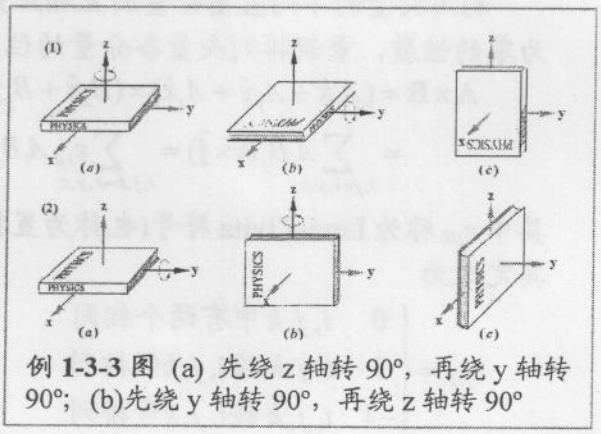
上面关于矢量的性质和运算的讨论都是针对位置矢量进行的，将矢量描述为具有长度和方向几何体以及所得到的结论也符合我们日常经验中对位置所积累的直觉感知和相应的认识。但并不能将位置矢量的特性和运算规则直接移植并应用到任意任何一个具有大小和方向属性的物理量，因为物理量运算所对应的操作特性决定了并不是任何一个具有大小和方向属性的物理量总是能够满足上面所讨论的位置矢量所满足的代数规则，或者说并非任何一个具有大小和方向属性的物理量都能用矢量来表征的。而一些描述完全没有方向属性的物理特性的物理量却可以用矢量代数规则来描述。下面通过一个实际的例子来说明这一点。

例 1-3-3 说明描述一个物体绕任意一个固定轴转动一个角度的物理量是否是一个矢量。



解：首先分析物体(例如你手中的这本书)做这样一次转动所涉及的可观测属性。容易看出这样的转动可由转轴的取向和物体绕该转轴转动的角度两个属性所决定。我们定义一个能概括这两个属性的物理量(称其为角位移)：规定由图中所规定的右手定则所确定的转动轴取向为角位移的方向，角位移的大小定义为物体实际转过的角度。这样定义的角位移直觉上与矢量非常相象，但判断其是否为一个矢量，需要检查它是否满足矢量代数规则。具体考虑物体经两次转动的合成效果。由于两次绕同一个轴转动的效果不涉及方向的合成(只涉及标量运算)，我们讨论物体绕两个不同转轴依次分别转动有限大小角度的总体效果，再交换两个转动的次序。比较两次实验的观察效果的差异，就可得知“角位移”是否满足矢量代数中的加法交换律。

将书本页面设置为直角坐标系的 *xy* 平面，书分别绕 z-轴和 y-轴逆时针旋转 90° 的两个角位移记为 Z 和 Y。将书本从原始取向开始相继转动 Z 和 Y。角位移 Z 和 Y 的各自的操作效果和累加效果见例图 1-3-3(a)。而将书本从原始取向开始按先 Z 后 Y 的秩序相继作两次转动操作的相应操作效果见例图 1-3-3(b)。显然，转动结果因次序不同而完全不一样，即角位移 Z 和 Y 不满足加法交换：



例 **1-3-3** 图 (a) 先绕 z 轴转 90°，再绕 y 轴转 90°；(b)先绕 y 轴转 90°，再绕 z 轴转 90°

$$Z+Y\neq Y+Z$$

这也表明了有限大小的角位移不是矢量。◀

上例中的讨论所涉及的只是加法交换率是否满足的特殊情形，实际上即使符合加法交换率也不足以确定它能满足全部的矢量代数规则。由于对一般的具有方向和长度属性的物理量实际上难以通过穷尽所有可能的实验操作来判断其是否满足必需的矢量运算的代数规则，因此，对判断一个具有大小和方向属性的物理量是否是一个矢量(能否按矢量代数进行运算)需要给出普适、明确的约束条件，

或者说需要给出矢量完整的定义。下面我们通过对位置矢量所满足的矢量运算性质给出普适的数学表述，然后将其推广到一般的物理矢量。

位置矢量的运算特性

我们对位置矢量的认识是和空间的概念联系在一起的，因而空间通常也被称为“位置空间”或“位置矢量空间”。一个物理系统(一个物质粒子或多个物质粒子的集合)与位置相关的特性在位置矢量空间中必须具备两个性质：

(1) 系统中任何一个物质粒子的位置矢量在一个给定的坐标系对应于唯一的一组坐标，而该坐标系中的任何一组坐标也只对应于空间中的唯一的一个位置矢量；

(2) 系统中任何两个物质粒子的相对位置，任何两个位置矢量的夹角与坐标系的选择无关。

其中第一个性质反映空间的单重性，第二个性质反映自由空间的均匀性、各向同性和平直性。这两个性质可以用两个矢量的和与两个矢量的标量积的运算性质加以概括：位置矢量在任何一个给定坐标系中的具有唯一的解析表达式意味着矢量和必须满足交换律和结合律。而系统中任何两个物质粒子的位置是由每个物质粒子的位置矢量的长度以及与任何两个物质粒子的位置矢量的夹角决定的，它们与坐标系选择无关的特性可以通过要求这两个位置矢量的标量积在任何坐标系中不变来实现。为此我们将上述两个条件概括为位置矢量在不同坐标系中的各坐标分量之间的变换方程必须满足的约束条件。位置的坐标系变换分为两类：平移和旋转。由于矢量本身与位置无关(即具有平移不变性)，位移矢量的上述特性只需通过坐标系绕原点转动所引起的各坐标分量的变换方程中得到体现。

为简单起见，我们先考虑处于直角坐标系 $\boldsymbol{S}$ (O-XYZ)中 X-Y 平面内的一个位置矢量 $\mathbf{r}$，相应的坐标为(x, y, $z=0$)，设想由坐标系 $\boldsymbol{S}$ 绕 z-轴逆时针旋转角度 φ 形成的另一个直角坐标系 $\boldsymbol{S}'$(O-$X'Y'Z$)，则位置矢量 $\mathbf{r}$ 在 $\boldsymbol{S}'$中的坐标表示为(x', y', $z=0$) (见图 1-3-12)。矢量 $\mathbf{r}$ 在两个坐标系中的两组坐标(分量)之间满足线性变换关系[5]：

$$x' = x\cos\varphi + y\sin\varphi$$

$$y' = -x\sin\varphi + y\cos\varphi$$

其中的线性变换系数($\cos\varphi$, $\sin\varphi$)和($-\sin\varphi$, $\cos\varphi$)分别是 $\boldsymbol{S}'$中的 $\boldsymbol{X}'$和 $\boldsymbol{Y}'$坐标轴在 $\boldsymbol{S}$ 系中的方向余弦。

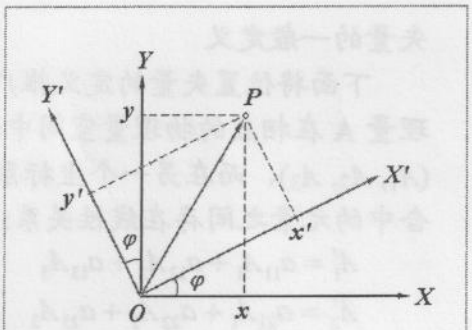


图 1-3-12 位置矢量在坐标系 $\boldsymbol{S}'$和 $\boldsymbol{S}$ 中的表示和相互间的几何关系。

下面通过上述变换关系和 X-Y 平面内的任意两个位置矢量 $\mathbf{r}_A$ 与 $\mathbf{r}_B$ 的标量积考察它们之间的位置相关性是否依赖于坐标系的选择。$\mathbf{r}_A$ 与 $\mathbf{r}_B$ 的标量积在两个坐标系 $\boldsymbol{S}'$和坐标系 $\boldsymbol{S}$ 之间的变换关系为

$$\begin{aligned}\mathbf{r}'_A\cdot\mathbf{r}'_B &= x'_A x'_B + y'_A y'_B \\ &= (x_A\cos\varphi + y_A\sin\varphi)(x_B\cos\varphi + y_B\sin\varphi) + (-x_A\sin\varphi + y_A\cos\varphi)(-x_B\sin\varphi + y_B\cos\varphi) \\ &= x_A x_B(\cos^2\varphi + \sin^2\varphi) + y_A y_B(\cos^2\varphi + \sin^2\varphi) \\ &= x_A x_B + y_A y_B = \mathbf{r}_A\cdot\mathbf{r}_B\end{aligned}$$

[5] 分解的唯一性是由矢量和满足交换律和结合律所保证的。

即两个矢量的相对位置不会因为坐标系的选择而改变。

上述分析可以推广到坐标系 S 相对于 S'有一个更一般的三维转动的情形。为简洁起见，记三个坐标分别为 x_1, x_2 和 x_3。这两个坐标系之间的关系可以用 S' 的各坐标轴在 S 系中的方向余弦$(a_{i1}, a_{i2}, a_{i3}) = (\cos\varphi_{i1}, \cos\varphi_{i2}, \cos\varphi_{i3})$来表示。则位置矢量 $\mathbf{r}$ 在这两个坐标系中的坐标之间的变换关系为

$$\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ x_2' &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ x_3' &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{aligned} \tag{1-3-40}$$

上述线性变换系数变换方程一般可以写为:

$$X' = AX \tag{1-3-41}$$

其中 A 称为**变换算符**。在变换方程的矩阵表示中表示为一个变换矩阵 $A = \{a_{ij}\}$，由于矩阵元 $a_{ij} = a_{ji}$ $(\cos\varphi_{ij} = \cos\varphi_{ji})$，$A$ 称为对称矩阵。而(1-4-21)式可表示为对变换矩阵的约束

$$(\tilde{A}A)_{ij} = \sum_{k=1}^{3} a_{ik}a_{kj} = \sum_{k=1}^{3} \cos\varphi_{ik}\cos\varphi_{jk} = \delta_{ij} \tag{1-3-42}$$

其中利用了关系式: $\cos\varphi_{ij} = \cos\varphi_{ji}$，及方向余弦满足的恒等式: $\cos^2\varphi_{12} + \cos^2\varphi_{23} + \cos^2\varphi_{31} = 1$。这样的变换称为**正交变换**。

而两个位置矢量 $\mathbf{r_A}$ 与 $\mathbf{r_B}$ 的标量积在两个坐标系中的变换关系为

$$\begin{aligned} \mathbf{r}_\mathbf{A}' \cdot \mathbf{r}_\mathbf{B}' &= \sum_{i=1}^{3} x_i'^A x_i'^B = \sum_{i,j,k=1}^{3} x_i^A x_j^B \cos\varphi_{ik}\cos\varphi_{kj} = \sum_{i,j=1}^{3} \delta_{ij} x_i^A x_j^B \sum_{k=1}^{3} \cos^2\varphi_{ik} \\ &= \sum_{i,j=1}^{3} \delta_{ij} x_i^A x_j^B = \sum_{i=1}^{3} x_i^A x_i^B = \mathbf{r}_\mathbf{A} \cdot \mathbf{r}_\mathbf{B} \end{aligned} \tag{1-3-43}$$

即两个空间矢量的相对方位不会因为坐标系的三维转动而改变。

矢量的一般定义

下面将位置矢量的定义推广到一般的物理矢量: 设量度某一种物理特性的物理量 **A** 在相应的物理量空间中的某一坐标系 S_A 内表示为三个独立的元素的集合 (A_1, A_2, A_3)，而在另一个坐标系 S'_A 中则表示为集合(A'_1, A'_2, A'_3)。如果这两个集合中的元素之间存在线性关系:

$$\begin{aligned} A_1' &= a_{11}A_1 + a_{12}A_2 + a_{13}A_3 \\ A_2' &= a_{21}A_1 + a_{22}A_2 + a_{23}A_3 \\ A_3' &= a_{31}A_1 + a_{32}A_2 + a_{33}A_3 \end{aligned} \tag{1-3-44}$$

只要该线性变换满足正交变换的要求:

$$\sum_{k=1}^{3} a_{ik}a_{kj} = \delta_{ij} \tag{1-3-45}$$

则物理量 **A** 为物理量 **A** 空间中的一个矢量。A_1, A_2, A_3 称为 **A** 在坐标系 S_A 中的正交分量。

有些读者可能觉得(1-4-20)和(1-4-22)以及(1-4-20)和(1-4-22)式以及有关的描述完全相同，因而对为什么要强调将矢量分解法则从位移矢量推广到一般的矢量

感到不理解。这里的要点在于我们所强调的物理矢量 **A** 所构建的矢量空间是一种抽象空间，首先它与我们所能感知的位置矢量所描述的并不是同一个空间，该空间中任意两“点”的相互关系完全没有位置空间中相应的方位或间距那样的几何意义。第二，它并不需要我们所能感知的“方向”属性。例如，1-1 节中曾经提到物理量空间的维度并不总是三维的，因此在处理相应的物理特性时就由可能定义超过三维的矢量空间。例如，在第二章中我们将会看到，物质在高速运动时所涉及到的空间和时间的观测效应中，时间和空间分量不是相互独立的，它们之间的相互变换可以用类似上述的正交变换描述。这意味着三个维度的空间分量和一个维度的时间分量形成了一个四维矢量，相应的物理空间称为四维时空。另外需要注意的是，对于矢量的乘积运算而言，我们所熟悉的位置空间中的运算要求的是属于同一空间的矢量，因而乘积具有封闭性，而一般的物理矢量之间的乘积运算反映的是不同的测量属性，并不要求参与测量的两个矢量属于同一空间，因而运算结果一般都不是封闭的。

上述矢量的定义强调的是它的各个分量之间的相互转换的规则，以及可以统一描述、处理的整体特性，而我们原来所强调的矢量的方向属性反而被淡化了，甚至可以说是抛弃了。而且我们将通过下面的事例看到，这样的定义使得矢量可以推广到更高维度的空间以及包含更多分量的不同数学量。

超越三维的矢量和矢量空间

我们首先将矢量的概念推广到超越我们直观感知的 n-维空间(n 为正整数)。尽管从直觉上很难想象一个超越三维空间中的矢量，但上述对三维矢量及其代数规则可以直接推广到 n-维空间中的矢量：

如果一个物理量 **X** 是一个包含了 n 个独立变化的元素的集合，它在直角坐标系 $\boldsymbol{S}_X$ 和 $\boldsymbol{S}'_X$ 中的表示分别为 $\{x_1, x_2, x_3, \ldots, x_n\}$ 和 $\{x'_1, x'_2, \ldots, x'_n\}$，只要物理量 **X** 的这两组分量之间的关系满足线性变换

$$x'_i = \sum_{k=1}^{n} a_{ij} x_j \quad (i = 1, 2, \ldots, n) \tag{1-3-46}$$

其中由系数矩阵 $\{a_{ij}\}$ 所决定的这个变换保证 X 的各分量自身乘积的总和不因坐标系的变换而改变：

$$R^2 = \sum_{k=1}^{n} x_k'^2 = \sum_{k=2}^{n} x_k^2 \tag{1-3-47}$$

则式(1-3-46)所定义的坐标系之间各矢量分量的变换称为正交变换，X 称为定义在 n 维矢量空间的矢量，R^2 称为 X 与自身的标积，R 称为 X 的模或长度。将(1-4-26)式代入(1-4-27)式

$$\sum_{i=1}^{n} x'_i x'_i = \sum_{k=1}^{n} \sum_{j,k=1}^{n} a_{ij} a_{ik} x_j x_k = \sum_{j,k=1}^{n} x_j \left(\sum_{i=1}^{n} \tilde{a}_{ji} a_{ik} \right) x_k$$

其中 $\tilde{a}_{ji} = a_{ij}$ 称为 a_{ij} 的转置。由此可以求出正交变换系数矩阵 $\{a_{ij}\}$ 所满足的条件

$$\sum_{i=1}^{n} \tilde{a}_{ji} a_{ik} = \delta_{ij} \tag{1-3-48}$$

为了方便，矢量 X 可以形式上写成通常的矢量形式

$$\mathbf{X} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n = \sum_{k=1}^{n} x_k\mathbf{e}_k$$

其中$\{\mathbf{e}_k\}$互相正交的基矢量，表示直角坐标系中坐标轴的方向。正交关系表示为

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} \equiv \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

这样，前面所讨论的矢量代数规则可以直接应用到 n-维空间中的矢量。例如，两个矢量 **A** 和 **B** 的和矢量的每一个直角分量$(\mathbf{A}+\mathbf{B})_k$为两个矢量的相应直角分量的和$A_k + B_k$，因此有

$$\mathbf{A} \pm \mathbf{B} = \sum_{k=1}^{n} A_k\mathbf{e}_k \pm \sum_{k=1}^{n} B_k\mathbf{e}_k = \sum_{k=1}^{n} (A_k \pm B_k)\mathbf{e}_k \tag{1-3-49}$$

矢量 **A** 和矢量 **B** 的和与差的行矩阵表示为

$$[a_1, a_2, \ldots, a_n] \pm [b_1, b_2, \ldots, b_n] = [a_1 \pm b_1, a_2 \pm b_2, \ldots, a_n \pm b_n]$$

而利用矢量的不同直角分量互相正交，标量积为零的性质，得矢量 **A** 和矢量 **B** 的标量积为

$$\mathbf{A} \cdot \mathbf{B} = \sum_{i,j=1}^{n} A_i B_j (\mathbf{e}_i \cdot \mathbf{e}_j) = \sum_{i,j=1}^{n} A_i B_j \delta_{ij} = \sum_{i=1}^{n} A_i B_i \tag{1-3-50}$$

我们从微积分课程中知道任何一个在原点无限次可导的函数$f(x)$可以展开为泰勒级数:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$$

上述展开式也可将函数$f(x)$理解为一个(函数)矢量，等式右边的各阶幂函数x^n为一个坐标系的轴向单位矢量，而每一项中的泰勒系数f_n就是矢量$f(x)$在坐标轴x^n上的投影。即矢量$f(x)$在一个特定的无穷维空间(幂函数空间)中展开。读者可以自己分析和探索这样的矢量空间中的代数运算规则。

以上事例虽然看起来都是数学问题，但在以后的章节中将会看到，它们都可用来描述实际的物理现象，即描述那些物理现象的物理量是一种无穷维的矢量。

需要用其它数学量来描述的物理量*

这里我们主要讨论两个例子，希望能让读者体会到物质的一些不能用标量或矢量来描述的属性和行为。并学习自己定义相关的数学量及运算规则。

第一个例子是描述物体的受力形变。我们知道一维的线弹簧在外力 **F** 的拉伸或压缩下发生的形变 **x** 可以用胡克定律描述

$$\mathbf{F} = -k\mathbf{x}$$

其中 k 称为弹性常数。由于作用力和形变共线，无论将作用力和形变表述为矢量还是标量，参量 k 用标量表示就足以表示弹簧的弹性特征。

但如果受力物体是一块弹性固体，形变就会复杂很多。例如沿 z-轴方向

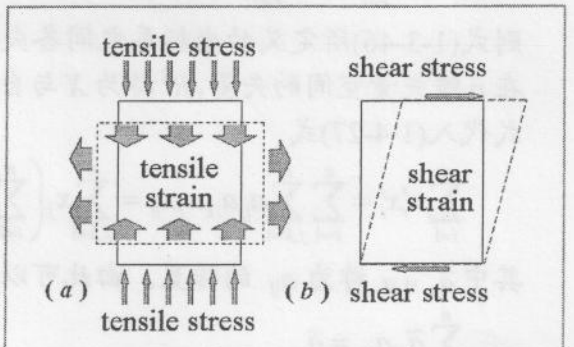


图 **1-3-13** 应变包含两个与方向有关的属性

拉伸或压缩一块橡胶长方体，形变不仅发生在 z 轴方向，也出现在与作用力垂直的 x 轴和 y 轴方向(见图 1-3-13)。这样的应变包含两个与方向有关的属性，这类数学量称为二阶张量[6]。实际上不仅应变是一个二阶张量，作用在物体表面上的作用力(称为应力)可以垂直于作用平面，也可以以两个独立的方向躺在作用平面内，分别称为拉伸应力和剪切应力。因此也存在着两个与方向有关的属性：作用力的方向以及表面的外法线方向，因此需要一个二阶张量来描述。二阶张量之间的运算规则称为**矩阵代数**。

物理量还能具有更多的方向属性，例如弹性固体的表面在受到上述应力而发生应变的情形，如果各外应力分量和全部的应变分量的关联都是线性的[7]，描述外应力 σ 和形变量 ε 之间关联的弹性参量包含有四个方向属性，在数学上称为**四阶张量**。在数学上一般地根据涉及方向的数目 n 的相应物理量称为 n **阶张量**。因而标量和矢量也分别被称为零阶张量和一阶张量。

第二个例子是描述原子中电子的运动行为。我们对电子在原子中的运动行为的描述曾经与行星的绕日轨道运动非常相似：电子绕着原子核旋转，同时也绕着自身的一条轴自转。描述这类运动的物理量是一种称为“角动量”的矢量。但与我们所熟悉的矢量不同，描述电子轨道运动的角动量的虽然有三个分量，但我们却不能将这三个分量合成为一个总的矢量，使定轴转动的角动量可以表现出一个特定的方向，在实验操作上我们从来没有在一个方向上探测到其分量为零。这意味着，这样一个角动量是没有确定方向的。更为奇异的是电子的**自旋角动量**，除了象轨道角动量那样没有确定的自旋轴，实验发现其旋转的角度周期竟然不是 2π 而是 4π，即电子需要自转两圈才能复原回到原来的位置。这种行为特性不能用我们上面介绍过的标量，矢量或各种阶数数的张量所描述，为了理解和描述电子的这种不同寻常的性质，物理学家自创了一个数学量，称其为“**旋量**”。用旋量描述的电子在原子中的行为可以很好地说明元素周期表中各元素的排列特征和化学特性。正是因为原子中亚原子粒子的这些奇异特性和描述这些特性的数学量的特殊运算规则使得牛顿力学在描述原子特性和亚原子粒子行为时不再有效。我们并不可能在这里探讨旋量数学，但会在后面涉及电子行为的章节中简洁地说明旋量的一些性质。

思考题

1-3-1 设想能否脱离运动来建立空间和时间的概念？

1-3-2 求一个半径为 3.520m 的圆的面积。应该认为原始数据为(3520±1)mm，因此圆的面积为 3.89m^2，但有人认为半径是由最小分度值为 cm 的尺得到测量数据，其不确定度为 5mm，即原始数据为 3.520±5mm。因此圆面积为 $3.892\text{m}^2\pm 0.1\text{m}^2$ 只有两位有效数字。应表示为 3.9m^2。你怎样看？

1-3-3 设一个物理量 Q 是通过直接测量参量 $q_1, q_2, \ldots, q_n$，再由函数关系 $Q=f(q_1, q_2, \ldots, q_n)$ 计算获得的间接测量值，是否可以先计算函数值，最后再计算其平均值：$\overline{Q}=\overline{f(q_1,\ldots,q_n)}$

---

6 这里暂时不讨论定义张量所必需的运算规则——张量代数。

7 通常称作用在物体的一个点上的力为作用力，称作用在物体的一个面上的力为应力.

1-3-4 两个单位矢量的和是否可能得到另一个单位矢量?

1-3-6 说明对任意两个矢量 **A** 和 **B** 都有 $\mathbf{A}\cdot(\mathbf{B}\times\mathbf{A})=0$.

1-3-7 说明$(\mathbf{A}\times\mathbf{B})\times\mathbf{C}$ 的几何意义。

1-3-8 能否定义矢量除法?

1-3-9 能否对二维空间中的矢量定义矢量积?能否对高于三维的空间中的矢量定义矢量积?(提示:分析定义矢量积所可能出现的问题。)

1-3-6 如何理解一个三维速度空间中两个点的相互关系。

习题

1-3-1 常见的仪器最小分度有 1(包括 0.1、0.01 等),2(包括 0.2、0.02 等)和 5(包括 0.5、0.05 等),说明应如何估读相应的测量值。例如,实验室常用的 2.5 级安培表的表盘,有两档量程。使用量程为 0.6A 档时,最小分度表示的电流强度为 0.02A,使用量程为 3A 档时,最小分度表示的电流强度为 0.1A。在两种量程情形,观测误差和最大绝对误差各为多少?

1-3-2 证明:a/b 的有效数字与 a, b 中有效数字较少的一个相同。

1-3-3 证明:a^2 及 $a^{1/2}$的有效数字与 a 的有效数字相同。

1-3-4 证明:有效数字 a 的对数函数,$\log a$ 的尾数位数与真数位数相同。

1-3-5 证明,10^a 的有效数字位数与指数 a 的小数点后的位数相同(包括紧接小数点后的零)。

1-3-6 证明,三角函数:取位可以按照角度有效数字而定。

1-3-7 写出位置矢量 **r** 在球坐标系和柱坐标系中的分量形式,说明它们和直角坐标系中的情况有什么不同。

1-3-8 矢量 **A** 的模为 2 个单位,方向与矢量 $2\hat{\mathbf{x}}-3\hat{\mathbf{y}}+4\hat{\mathbf{z}}$ 的方向相同,写出 **A** 的表示式。

1-3-9 利用矢量代数,求解立方体的一对对角线之间的夹角。

1-3-10 一单位矢量的方向指向矢量 $\mathbf{A}=5\hat{\mathbf{x}}-3\hat{\mathbf{y}}$ 和 $\mathbf{B}=4\hat{\mathbf{x}}-7\hat{\mathbf{y}}+5\hat{\mathbf{z}}$ 矢量的方向中间,写出该矢量的表达式。

1-3-11 一单位矢量与矢量 $\mathbf{A}=5\hat{\mathbf{x}}-3\hat{\mathbf{y}}$ 和 $\mathbf{B}=4\hat{\mathbf{x}}-7\hat{\mathbf{y}}+5\hat{\mathbf{z}}$ 垂直,写出该矢量的表达式。

1-3-12 证明矢量 **A, B, C** 满足恒等式:$\mathbf{A}\times(\mathbf{B}\times\mathbf{C})=\mathbf{B}(\mathbf{A}\cdot\mathbf{C})-\mathbf{C}(\mathbf{A}\cdot\mathbf{B})$

1-3-13 证明矢量 **A, B, C** 满足恒等式:$\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C})=\mathbf{B}\cdot(\mathbf{C}\times\mathbf{A})=\mathbf{C}\cdot(\mathbf{A}\times\mathbf{B})$

1-3-14 利用矢量代数,证明三角形的三条高线相交于同一点(称为垂心)。

1-3-15 证明梯度算符:$\nabla\equiv\hat{\mathbf{x}}\dfrac{\partial}{\partial x}+\hat{\mathbf{y}}\dfrac{\partial}{\partial y}+\hat{\mathbf{z}}\dfrac{\partial}{\partial z}$是一个矢量算符。

1-3-16 对于均匀,且各向同性的弹性体,这种特性应如何在应变张量反映出来。

1-3-17 对于均匀,而且各向同性的弹性体,描述外应力 σ 和形变量 ε 之间关联的胡克定律:$\sigma_{ij}=\sum_{k,l=1}^{3}c_{ijkl}\varepsilon_{kl}$,其中弹性参量 c 是一个**四阶张量**。对于均匀,而且各向同性的弹性体,弹性参量 c 如何反映这种特性。

1-3-18 设计一种数学符号形式或规则,将上述胡克定律用矩阵形式表现出来。

综合课题

1-3-1 考虑一个描述随机误差分布的数学模型，该模型应包含大量的小概率不确定的物理过程

求解大量的小颗粒最终在不同位置的分布函数，说明随着 n 趋向于无穷，形成的分布趋向于正态分布。说明什么因素决定了分布的标准差大小。

1-3-2 函数$f(x)$的泰勒展开式

$$f(x)=\sum_{n=0}^{\infty}\frac{f^{(n)}(0)}{n!}x^n$$

可以理解为：函数矢量$f(x)$在一个以各阶幂函数x^n为坐标轴的一个特定函数矢量空间中的分解。其中x^n向单位矢量，展开式每一项中的泰勒系数f_n就是函数矢量$f(x)$在坐标轴x^n上的投影。说明该矢量空间中的代数运算规则，特别是标量积运算。

1-3-3 计算机厂商一直致力于设备组件的微型化，对设计更轻、更薄的计算设备而言，现有键盘的体积已成为一个瓶颈。其中的困难之一是在不影响键盘性能的同时缩减键盘体积。根据你使用键盘的经验，说明应定义怎样的物理参数最能反映你对键盘感受，并说明可以在那些参量上予以改进以缩减键盘体积。(提示：尝试并比较现有的各种普通键盘，电容键盘、机械键盘、薄膜键盘的手感。)

第2章 运动学

运动学研究的是如何恰当描述物体的运动状态及其随时间的变化，但不涉及运动物体自身的结构和内在性质(如物体的质量)以及其运动状态变化可能涉及的原因(如物体与外界可能发生的相互作用)。描述物体的运动过程实际上就是确认由一系列相关事件($\mathbf{r}, t$)构成的集合中各个事件的次序和间隔，它有助于我们理解时间和空间与其中的物体的运动状态的关联形式和程度。例如，我们希望知道物体及其运动状态是否影响其所处的空间和时间的结构和性质，物理学中将这一命题表述为时间和空间是绝对的还是相对的。本章主要讨论与运动相关的四个主要问题：描述物体运动的特征物理量是什么，描述运动的一些不同的视角和方法，运动的最基本特征—相对性原理，以及该原理所决定的对时空的认识和观念。

2-1 参考系—时空的统一描述

在试图对一个物体的运动进行定量描述以前，首先需要找出或定义反映物体静止和运动，以及不同程度运动之间的差异的物理量。我们对机械运动以及运动差异的直觉感知主要来自物体在不同时刻处于不同空间位置的多幅静态的图像的依次替换形成的视觉印象，我们面临的问题是如何从这样的视觉印象中抽象出一个静态的描述。艺术家就这个问题提供了一种方案：通过静态的点和线的特殊组合可以在视觉中戏剧性地重现物体处于剧烈运动的场景。图 2-1-1 是两幅通过不同的手法传达运动视觉印象的作品。上图是一辆处于高速行驶状态的赛车照片，我们从中获得的赛车运动及其剧烈程度的视觉印象是通过画面背景的模糊化形成的。这种描述技法也揭示了作者的观测角度和产生运动效果的原因：相机是以相对于赛车静止的状态拍摄的，模糊的背景是底片经受一段时间持续曝光的结果。艺术家也采用另一种描述运动的视觉印象手法：让相机相对于背景静止，而让底片对运动物体持续曝光一段时间，法国抽象派画家杜尚的作品《下楼梯的裸者》(见下图)所呈现的就是这种观察视角：五个相同的姿态的重复人物构形呈现出正在下楼梯的运动印象，它相当于对下楼梯的模特相隔同样的时间间隔在同一底片上曝光 5 次的结果，这也是物理学中的一种常用的观察记录方法，称为频闪。图 1-2-1 中记录的灯泡从桌面上掉落到地面碎裂这一运动

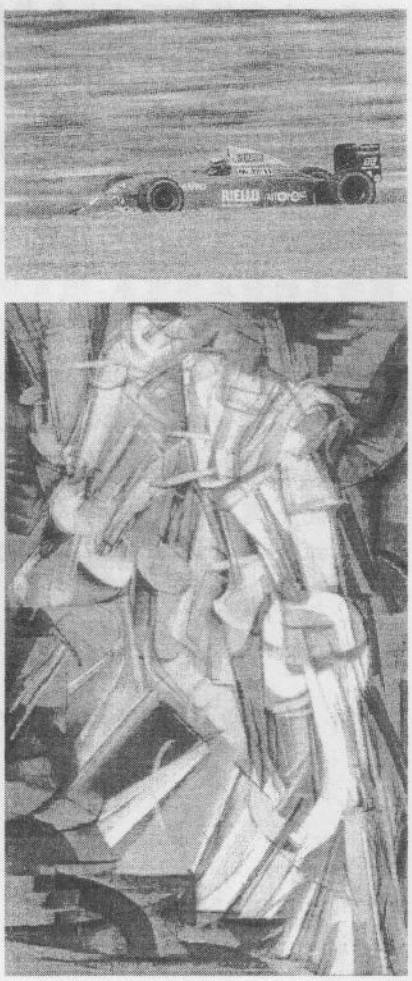


图 **2-1-1** (a) 背景的高度模糊化形成的赛车正在高速疾驶的视觉效果。(b) 马塞尔·杜尚的《下楼梯的裸者》(1912 年)。

过程就是采用这种方法拍摄的。

图 2-1-1 中的两幅作品都是通过一个物体相对于另一个物体的清晰程度来反映物体的运动属性的。这种对运动的表现手法揭示了观察、描述一个物体运动的两个最基本要素：第一，一个物体的运动是相对于另一个物体而言的；第二，观察和描述运动需要有限的时间间隔。我们先讨论第一个要素，第二要素在下一节予以说明。

当我们说物体 **A** 在运动时，意味着即时一定存在着另一个物体 **B** 处于静止状态，或者说"物体 **A** 的运动"的确切含义是"物体 **A** 相对于物体 **B** 的运动"。其中物体 **B** 扮演着观测和描述物体 **A** 运动的参照物角色。

因此对一个物体运动的描述可以通过将该物体在不同时刻 t 相对于另一个物体的位置 **r** 记录下来。这样，一个物理过程可以表述为一系列事件$(\mathbf{r}, t)$的集合，而对物体运动的定量描述实际上就是确认运动物体在任一时刻所处的位置。设想将一个**坐标系**固定在一个物体上，并设想坐标系中各点都设置有一个同步运行的时钟(见图 2-1-2)，经过这样标度的时间和空间称为**参考系**。由此，构成其它物体相对于该参考系运动的每一个事件都可以由参考系的时空坐标$(\mathbf{r}, t)$描述。这种描述显然是相对的，因为时空坐标都是相对于我们所选择的参考系而言的。

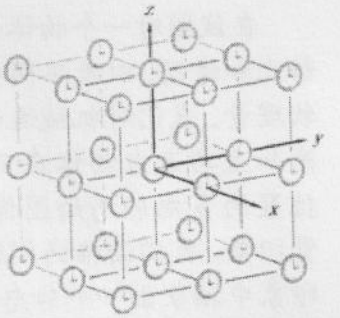


图 2-1-2 各时空点通过坐标系和各处的时钟得到标度，从而可以用作描述物体运动的参照系。

我们也可以将参考系固定在某一个运动物体上观测其它物体相对于该物体的运动。这时物体对于所设置的参考系始终是静止的，这样的参考系称为该物体的**静止参考系**，通常称为相关物体的参考系，如火车参考系，地球参考系等。这里读者应该注意到，与运动物体相关联的参考系中的时钟和静止参考系中的时钟的运行从逻辑上说不存在等同性，但在日常经验中，我们通常都默认这两个时钟"滴答"着同样节奏的时间，物理学中称之为**时间的绝对性**。这样测量时就不需要每个参考系都携带时钟，只要一个参考系配置了时钟，其它参考系中都可以从这个参考系中的时钟读出自身参考系中的时间。

思考题

2-1-1 说明坐标系和参考系之间的差异。

2-2 质点运动的描述

一个宏观物体的运动通常表现为三种基本运动类型的组合：**平动**，**转动**和**形变**。其中平动表现为物体整体的位置变化；转动描述的是物体的各部分围绕着空间一个固定点(或一条固定直线)作圆周运动，相应的固定点或固定直线称为**转动中心**或**转动轴**[1]；而形变反映的是物体中不同部分相对距离的变化。例如，地球的运动是其整体环绕太阳的轨道运动(平动)和绕其自身轴线的自旋运动(转动)；一个雨点在下落(平动)过程中由于各个部分的速度不同而导致其形状发生变化(形变)。在讨论实际物体的某些特性时，通常并不需要计入其全部运动类型。例如，如果我们只关注于地球绕日运行的周期，就没有必要考虑地球的自转行为。而试图理解河流对两岸的不同冲刷程度，就要分析地球的自转行为而不必再考虑描述地球围绕太阳的公转的平动。事实上，**平动是比转动和形变更基本的运动形式**。从物质构成的角度而言，宏观物体总可以细分为大量点状粒子的集合，由于一个点状粒子不具有空间尺度和内部结构，也就不存在空间取向和形变的可能性，从而自动排除了出现转动和形变的可能性。其唯一可能的运动形式只是位置的变化。这意味着宏观物体的转动和形变都可以表示为构成大量点状粒子的不同平动状态的特殊组合。描述物体运动的物理量也相应地可以表述为描述诸多粒子平动的物理量的组合。本章我们只是关注于物体的平动行为，对物体的转动和形变将在以后的章节中讨论。

2-2-1 平动运动模型

对处于平动中的物体而言，其中各点的运动状态及其变化情况完全相同，因此对物体上任何一点的运动状态的描述就足以代表整个物体的运动状态。因而可以建立一个仅仅包含这一特性的物体模型——一个点状粒子。许多教科书上将这种点粒子模型称为**质点**。事实上质点模型不仅要求忽略物体的空间尺度，而且还强调将整个物体的全部质量集聚在一个几何点上。但质量的概念对运动学而言并没有什么意义，它只是在讨论物体的动力学过程中才是必需的。但为了避免以后重复，下面也称这样的点粒子模型为质点。

2-2-2 解析描述

2-1 节中说明了描述物体运动的两个要点：(1) 一个物体的运动是相对于另一个物体而言的，(2) 观察和描述运动必须反映出有限时间的流逝。描述物体(可以观察到的)平动的运动属性是物体的位置变化以及这种变化的快慢程度，我们分别定义“位移”和“速度”来描述这两个运动特性。

按照物理量的测量属性，**位移**定义为一个物体的空间位置的改变。设一个物体(质点)从位置 $\mathbf{r}_1$ 沿某一路径运动到位置 $\mathbf{r}_2$(见图 2-2-1)，则该物体的位移定义为

$$\Delta\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1 \tag{2-2-1}$$

[1] 转轴可以穿过物体内部，也可以在物体外部，相应的情形分别称为自转和轨道运动，另外，转轴本身可以是固定的，也可以发生取向的变化，前者称为定轴转动，后者则称为定点转动。

这里需要注意两点，(1) 我们是在物理参考系而不是在数学坐标系中描述物体运动，对物体从一个位置出发到达另一位置的两个事件包含了全部的时空坐标，(2-2-1)式的实际含义是 $\Delta\mathbf{r}(t_2, t_1) = \mathbf{r}_2(t_2) - \mathbf{r}_1(t_1)$。(2) 运动物体的位移 $\Delta\mathbf{r}$ 与其实际移动过的路程长度(距离)Δs 不同(见图 2-2-1)：位移是一个矢量，反映物体在空间的位置变化，而物体实际移动过的路程长度是一个标量。

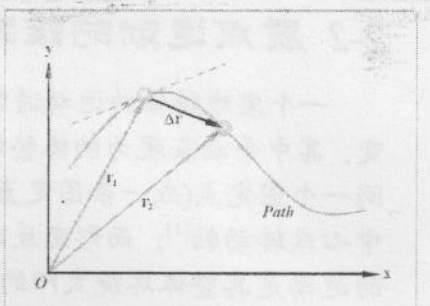


图 2-2-1 物体沿曲线路径从时空点$(\mathbf{r}_1, t_1)$运动到 $(\mathbf{r}_2, t_2)$。

物体在时间间隔 $\Delta t = t_2 - t_1$ 内的位置变化的快慢程度可以由位移矢量相对与时间的变化率来定义

$$\overline{\mathbf{v}}(t) = \frac{\Delta\mathbf{r}}{\Delta t} \tag{2-2-2}$$

称为质点在 Δt 时间内运动的**平均速度**。通常平均速度与时间间隔 Δt 有关，Δt 越小，越能反映物体运动快慢的细节。为了得到对运动的最佳描述，牛顿引入了使(2-2-2)式中的 Δt 趋向于零的“极限”概念：

$$\mathbf{v}(t) = \lim_{\Delta t\to 0}\frac{\mathbf{r}(t+\Delta t)-\mathbf{r}(t)}{\Delta t} = \lim_{\Delta t\to 0}\frac{\Delta\mathbf{r}}{\Delta t} \tag{2-2-3}$$

由此得到的相关速度称为**瞬时速度**。为了进行这种特殊的运算，牛顿甚至开创了一个新的数学领域——微积分。在这个数学框架中，(2-2-3)式中的极限运算称为位移函数 $\mathbf{r}(t)$对时间 t 的**微商**[2],[3]，记为

$$\mathbf{v}(t) = \frac{d\mathbf{r}}{dt} \tag{2-2-4}$$

物理学中还有一个与速度相关的物理量—**速率**，定义为物体在单位时间内运动过度距离

$$\overline{v} = \frac{\Delta s}{\Delta t} \tag{2-2-5}$$

这也是根据操作性要求定义的平均速率。也可以类似地定义**瞬时速率**：

$$v = \lim_{\Delta t\to 0}\frac{\Delta s}{\Delta t} = \frac{ds}{dt} \tag{2-2-6}$$

---

[2] 式(2-2-2)给出的是平均速度是基于测量操作的，但(2-2-3)式给出的瞬时速度的定义却不具有操作属性，因为无法实施对应于$\Delta t\to 0$的相关测量。微积分的运算对象是连续函数，利用微积分描述物体的运动意味着描述运动的位移函数$\mathbf{r}(t)$是时间的连续函数。这虽然与日常观测到的物体行为相符，但这些物体的尺度都远大于目前所能测量的最小长度和最短时间间隔。极限运算的定义是一种逻辑抽象，它在数学中可以通过一些定义使其在一大类数学对象的运算中形成完整，自洽的封闭结构，但不符合物理学中物理量及其运算所必需的操作属性要求。事实上，关于连续位移的可能性在哲学认识上也存在一些不同的观念和争论。这里的特别说明只是希望读者能意识到任何归纳和抽象都可能存在着局限性，它只是就某一类现象所作的。例如，就速度本身而言，它在宏观世界中确实反映了物体的运动属性，但在微观世界中根本就不存在运动速度这样一个物理量

[3] 许多学者受研究运动的影响而否定运动的存在。如公元前五世纪的希腊哲学家巴门尼德论述道，因为不能无中生有，因而不存在变化。他从自然界的永恒和始终如一出发推断所有的变化以至运动都只是假象。公元前五世纪的希腊哲学家赫拉克利特则持有相反的看法，他将变化看成自然的实质。这两个同样著名的观点吸引着许多学者仔细研究自然界中是否存在守恒量。巴门尼德的合作者芝诺提出过一个“飞矢不动”的悖论以反对运动的观点：由于箭在其飞行过程中的任一瞬间都有一个暂时的位置，它在这个位置上和不动没有什么区别。类似的问题在近代量子物理理论中至今仍困惑着许多学者。

瞬时速率和瞬时速率虽然反映的是不同的概念，但两者的数值大小是相等的。

物体的平动属性可以由位移和速度所完全描述，但经典力学中通常还引入一个额外的物理参量—**加速度**来描述物体运动速度相对于时间的变化率。类似于速度的定义，我们按测量操作过程，先定义物体在时间间隔 Δt 内的**平均加速度**

$$\bar{\mathbf{a}} = \frac{\Delta \mathbf{v}}{\Delta t} \tag{2-2-7}$$

然后引入极限运算定义**瞬时加速度**:

$$\mathbf{a} = \lim_{\Delta t \to 0} \frac{\Delta \mathbf{v}}{\Delta t} = \frac{d\mathbf{v}}{dt} = \frac{d^2\mathbf{r}}{dt^2} \tag{2-2-8}$$

其中利用了(2-2-4)式的定义：速度 $\mathbf{v}$ 为位移 $\mathbf{r}$ 是关于时间的一阶微商，因此加速度可以表示为位移 $\mathbf{r}$ 是关于时间的二阶微商[4]。

下面我们讨论质点运动参量的计算。一般而言质点的运动速度可以表示为:

$$\mathbf{v}(t) = \frac{d\mathbf{r}}{dt} = \frac{d(r\hat{\mathbf{r}})}{dt} = \frac{dr}{dt}\hat{\mathbf{r}} + r\frac{d\hat{\mathbf{r}}}{dt} \tag{2-2-9}$$

即位置矢量函数 $\mathbf{r}(t)$对于时间的微商包含两部分：位移大小对时间的变化率以及位移方向对时间的变化率。这两部分所描述的是最典型的两类平动：直线运动和圆周运动。其中，直线运动反映的是运动物体相对于一个固定参考点的距离变化，圆周运动代表的是物体运动中的方向改变。而一般的曲线运动只是上述两种运动的组合。

类似地，可以将加速度表示为

$$\mathbf{a}(t) = \frac{d\mathbf{v}}{dt} = \frac{d^2\mathbf{r}}{dt^2} = \frac{d}{dt}\left(\frac{dr}{dt}\hat{\mathbf{r}} + r\frac{d\hat{\mathbf{r}}}{dt}\right) = \frac{d^2\mathbf{r}}{dt^2} + 2\frac{dr}{dt}\frac{d\hat{\mathbf{r}}}{dt} + r\frac{d^2\hat{\mathbf{r}}}{dt^2} \tag{2-2-10}$$

其中包含了速度大小对时间的变化率以及速度方向对时间的变化率，以及两者的相互耦合。

原则上我们可以在各种坐标系中描述和讨论质点的运动，但通常会根据运动的特点选择最方便的坐标系。例如，描述直线运动的最简便的方法是直角坐标系，而描述圆周运动则利用极坐标系来标度最为便利。而一般的运动可以选用自然坐标系。

质点运动在直角坐标中的描述

位移 $\Delta\mathbf{r}(t)$在直角坐标系中可分解为 $\Delta x(t)\hat{\mathbf{x}} + \Delta y(t)\hat{\mathbf{y}} + \Delta z(t)\hat{\mathbf{z}}$，其中三个直角分量的方向不随时间发生改变，因此速度和加速度计算只需要计算各分量对时间的变

[4] 有些读者可能会认为，类似地定义位移关于时间的三阶导数，四阶导数等(加速度相对于时间的变化率，加速度的加速度等)等物理量可以给出物体运动状态的更多细节描述。但事实上并非如此。物体(质点)位置随时间的变化给出了其运动学的全部描述，只是我们对时间不是直接感知的，因此通过引入速度的概念将时间和空间联系起来，这样质点运动的全部特征就可由位移和速度进行描述，而时间只是一个连接位移和速度数学参量。而引入对应于位移关于时间的高阶导数(包括加速度)就没有必要，因为它们在运动学范畴内不会提供任何新的可观测信息。但动力学研究物体与物体之间的相互作用，由于相互作用无法直接观察，只能借助于观察它所产生的物理效应，即物体受到外界作用时其运动状态会发生变化的程度来描述，而在一定程度上它可以用速度的时间变化率表示。由此可知引入加速度主要是为了量度物体所受力的大小。位移相对于时间的更高阶导数则不存在相对应的可观测物理效应。

化率：

$$\mathbf{v}(t)=\frac{dx}{dt}\hat{\mathbf{x}}+\frac{dy}{dt}\hat{\mathbf{y}}+\frac{dz}{dt}\hat{\mathbf{z}}=v_x(t)\hat{\mathbf{x}}+v_y(t)\hat{\mathbf{y}}+v_z(t)\hat{\mathbf{z}} \tag{2-2-11}$$

$$\mathbf{a}(t)=\frac{d^2x}{dt^2}\hat{\mathbf{x}}+\frac{d^2y}{dt^2}\hat{\mathbf{y}}+\frac{d^2z}{dt^2}\hat{\mathbf{z}}=a_x(t)\hat{\mathbf{x}}+a_y(t)\hat{\mathbf{y}}+a_z(t)\hat{\mathbf{z}} \tag{2-2-12}$$

显然，直角坐标系对描述直线运动而言是最简洁的：只要取质点的运动轨迹为坐标轴，描写运动的参量就只是一维的。

例 2-2-1 一个被弹簧悬挂着的重物作上下振动，它在任一时刻t的位置坐标z可以用数学公式表示为$z = A\sin(\omega t+\varphi)$，其中$A$，$\omega$和$\varphi$都是反映振动特性的常数参量。计算重物在任一时刻$t$的速度和加速度。

解： 描述一维运动只需要在标量的框架中讨论。重物在任一时刻t的运动速度为

$$v=\frac{dz}{dt}=\omega A\cos(\omega t+\varphi)$$

而在相应时刻的加速度为

$$a=\frac{dv}{dt}=-\omega^2 A\sin(\omega t+\varphi)$$

说明： 由于质点的位移描述的是质点的运动情况，而加速度反映的是质点运动状态的变化率。本例中这两个参量随时间的变化具有相同的形式这一事实反映了物理学中的因果关系的内在对称性：物体周期性的演化行为取决于控制因素的周期性特征。◀

如果知道各时间点物体的位移，通过微商可以依次求解出物体的速度和加速度。反过来，如果知道不同时间点物体的加速度，则可以通过积分依次求解出物体的速度和位移：

$$\mathbf{v}(t)=\mathbf{v}_o+\int_{t_o}^{t}\mathbf{a}(t)dt \tag{2-2-13}$$

$$\mathbf{r}(t)=\mathbf{r}_o+\int_{t_o}^{t}\mathbf{v}(t)dt \tag{2-2-14}$$

其中$\mathbf{r}_o$和$\mathbf{v}_o$分别为物体在t_o时刻的位置和速度。上述积分在直角坐标系中简化为

$$\mathbf{v}(t)=[v_{xo}+\int_{t_o}^{t}a_x(t)dt]\hat{\mathbf{x}}+[v_{yo}+\int_{t_o}^{t}a_y(t)dt]\hat{\mathbf{y}}+[v_{zo}+\int_{t_o}^{t}a_z(t)dt]\hat{\mathbf{z}} \tag{2-2-15}$$

和

$$\mathbf{r}(t)=[x_o+\int_{t_o}^{t}v_x(t)dt]\hat{\mathbf{x}}+[y_o+\int_{t_o}^{t}v_y(t)dt]\hat{\mathbf{y}}+[z_o+\int_{t_o}^{t}v_z(t)dt]\hat{\mathbf{z}} \tag{2-2-16}$$

例2-2-2 设想一个物体从高度为z_o的位置以初速度v_o下落。(*a*) 计算物体下落至高度为z的位置处的速度。(*b*) 如果物体在下落过程中受到空气阻力，相应的减速度与物体的下落速度成正比：$\mathbf{a}_r=-\beta\mathbf{v}$，求解物体下落过程中的速度随时间的演化。

解： 这是一个相对简单的一维问题，如图设置向下的坐标轴z。则物体落向地面

所具有的中立加速度表示为：$\mathbf{a}=g\hat{\mathbf{z}}$，其中$\hat{\mathbf{z}}$是表示z-轴的单位矢量。

(*a*) 这里已经知道的是物体自由下落的速度演化方程

$$\frac{dv}{dt}=g$$

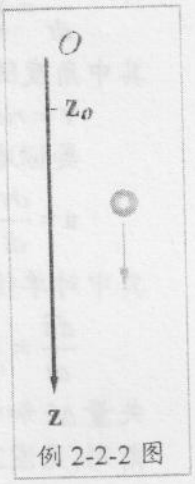
而问题涉及的是速度与位置的关系，可以通过微商的链式法则将速度随时间的演化变换为速度与位置的关联

$$\frac{dv}{dt}=\frac{dv}{dz}\frac{dz}{dt}=\frac{dv}{dz}v=g$$

其中利用了速度的定义：$v=dz/dt$。

$$\int_{v_o}^{v}vdv=g\int_{z_o}^{z}dz$$

由此可得所要求解的：$v^2=v_o^2-2g(z-z_o)$



例 2-2-2 图

(*b*) 若下坠过程中物体空气阻力而具有减速度$\mathbf{a}_r=-\beta\mathbf{v}$，其净加速度为

$$\frac{dv}{dt}=-\beta v+g$$

将微分变量分置于方程两侧，并分别从初始时刻计算定积分

$$\int_0^{v}\frac{dv}{v-g/\beta}=-\beta\int_0^{t}dt \quad \text{即：} \quad \log\left(1-\frac{\beta}{g}v\right)=-\beta t$$

解得：$v=\frac{g}{\beta}(1-e^{-\beta t})$

随着物体速度增加，净加速度持续减小，最终将趋向于一个极限速度$v=g/\beta$。◀

圆运动在极坐标中的描述

我们接着考虑质点的速度量值不变，只是方向发生变化的情形。这种运动中，粒子被限制在一个球面上运动，这时(2-2-9)式中的第一项(粒子的径向速度)为零，速度只是表示为第二项：物体运动方向的变化率，其中单位径向矢量$\hat{\mathbf{r}}$关于时间的变化率可以按定义表示为

$$\frac{d\hat{\mathbf{r}}}{dt}=\lim_{t'\to t}\frac{\hat{\mathbf{r}}'-\hat{\mathbf{r}}}{t'-t}=\lim_{\Delta t\to 0}\frac{\Delta\hat{\mathbf{r}}}{\Delta t}$$

单位矢量$\hat{\mathbf{r}}$经过时间间隔Δt，转过角度$\Delta\theta$成为$\hat{\mathbf{r}}'$，对于很小的时间间隔Δt，相应的位移$\hat{u}_r$和单位矢量$\hat{\mathbf{r}}$及$\hat{\mathbf{r}}'$构成一个平面等腰三角形OPP'(见图2-2-2)。如果不断缩小时间间隔Δt，等腰三角形的顶角$\Delta\theta$也就相应地变得很小，这时单位圆上的弦长$\overline{PP'}$就能用相应的弧长来近似：

$$\overline{PP'}\cong PP'=|\hat{\mathbf{r}}|\,\Delta\theta=\Delta\theta$$

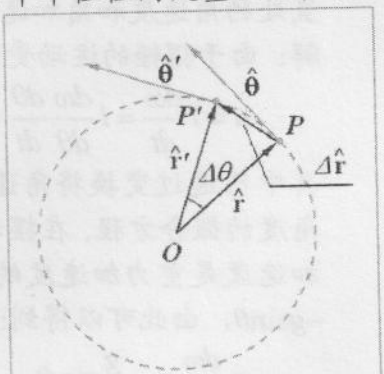


图2-2-2 单位矢量$\hat{\mathbf{r}}$及$\hat{\mathbf{r}}'$和位移$\Delta\hat{\mathbf{r}}$构成等腰三角形。

而矢量$\Delta\hat{\mathbf{r}}$的方向随着$\Delta\theta$趋向于零而趋向于和单位矢量$\hat{\mathbf{r}}$垂直，相应的单位矢量记为$\hat{\boldsymbol{\tau}}$(见图2-2-2)，指向单位圆的切线方向。则单位矢量$\hat{\mathbf{r}}$关于时间的微商可以表示为

$$\frac{d\hat{\mathbf{r}}}{dt}=\lim_{\Delta t\to 0}\frac{\Delta\theta}{\Delta t}\hat{\boldsymbol{\tau}}=\frac{d\theta}{dt}\hat{\boldsymbol{\tau}} \tag{2-2-17}$$

其中角度随时间的变化率称为(瞬时)角速度，记为ω。则速度表达式可以表示为

$$\mathbf{v}=r\omega\hat{\boldsymbol{\tau}} \tag{2-1-18}$$

类似地可以计算加速度—速度随时间的变化率:

$$\mathbf{a}=\frac{d\mathbf{v}}{dt}=r\frac{d\omega}{dt}\hat{\boldsymbol{\tau}}+r\omega\frac{d\hat{\boldsymbol{\tau}}}{dt}$$

其中对单位矢量$\hat{\boldsymbol{\tau}}$的时间微商定义为:

$$\frac{d\hat{\boldsymbol{\tau}}}{dt}=\lim_{t_2\to t_1}\frac{\hat{\boldsymbol{\tau}}'-\hat{\boldsymbol{\tau}}}{t'-t}=\lim_{\Delta t\to 0}\frac{\Delta\hat{\boldsymbol{\tau}}}{\Delta t}$$

矢量$\Delta\hat{\boldsymbol{\tau}}$和两个单位矢量$\hat{\boldsymbol{\tau}}$及$\hat{\boldsymbol{\tau}}'$构成一个等腰三角形(见图2-2-3)。随着时间间隔Δt趋于零，等腰三角形的顶角$\Delta\theta$趋于无穷小量，弦长$|\Delta\hat{\boldsymbol{\tau}}|$就能用单位圆上的相应的弧长表示:

$$|\Delta\hat{\boldsymbol{\tau}}|=|\hat{\boldsymbol{\tau}}|\Delta\theta=\Delta\theta$$

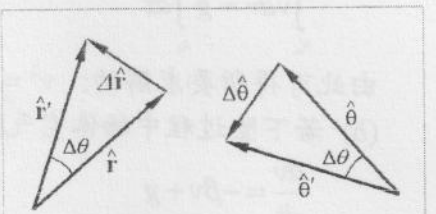


图2-2-3 单位矢量$\hat{\boldsymbol{\tau}}_1$及$\hat{\boldsymbol{\tau}}_2$和位移$\Delta\hat{\boldsymbol{\tau}}$构成等腰三角形。

而矢量$\Delta\hat{\boldsymbol{\tau}}$的方向随着$\Delta\theta$趋向于零而趋向于和单位矢量$\hat{\boldsymbol{\tau}}$垂直，即趋向于$\hat{\mathbf{r}}$的反方向。因此单位矢量$\hat{\boldsymbol{\tau}}$关于时间的微商可以表示为

$$\frac{d\hat{\boldsymbol{\tau}}}{dt}=-\lim_{\Delta t\to 0}\frac{\Delta\theta}{\Delta t}\hat{\mathbf{r}}=-\frac{d\theta}{dt}\hat{\mathbf{r}}=-\omega\hat{\mathbf{r}} \tag{2-2-20}$$

则加速度表示为

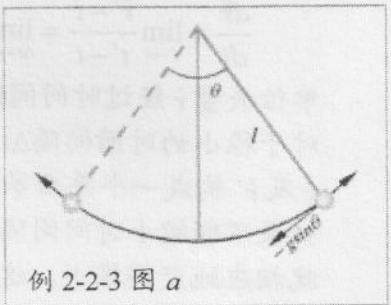
$$\mathbf{a}=-r\omega^2\hat{\mathbf{r}}+r\frac{d\omega}{dt}\hat{\boldsymbol{\tau}} \tag{2-2-21}$$

其径向分量描述的是横向运动，因其方向指向坐标系原点而称为**向心加速度**。加速度的横向分量反映的是角速度的变化率，即角加速度。

例 2-2-3 考虑一个由长度为l的轻杆和摆锤构成的单摆系统。写出摆锤在不同位置处的角速度和角加速度的表达式。

解: 由于摆锤的运动受到轻杆的约束只能是一种圆周运动，其加速度可以表示为

$$a=l\frac{d\omega}{dt}=l\frac{d\omega}{d\theta}\frac{d\theta}{dt}=l\omega\frac{d\omega}{d\theta}$$



例 2-2-3 图 *a*

其中已通过变换将角速度和角加速度表示成关于角度的微分方程。在摆动过程中摆锤所具有的实际加速度是重力加速度的一个分量(见本例图): $a=-g\sin\theta$，由此可以得到:

$$\omega\frac{d\omega}{d\theta}=-\frac{g}{l}\sin\theta$$

对上式积分

$$\int_{\omega_o}^{\omega}\omega\,d\omega=-\frac{g}{l}\int_{\theta_o}^{\theta}\sin\theta\,d\theta$$

其中θ_o和ω_o分别为摆锤的初始时刻的角位置和角速度。可以得到

$$\omega^2-\omega_o^2=2\frac{g}{l}(\cos\theta-\cos\theta_o)$$

上式可以表示为

$$\omega=\frac{d\theta}{dt}=\pm\sqrt{\omega_o^2+2\frac{g}{l}(\cos\theta-\cos\theta_o)}$$

再对上述方程积分可得到摆锤在不同时刻的位置:

$$\int_{\theta_o}^{\theta}\frac{d\theta}{\sqrt{\omega_o^2(l/g)+2(\cos\theta-\cos\theta_o)}}=\pm\sqrt{\frac{g}{l}}t$$

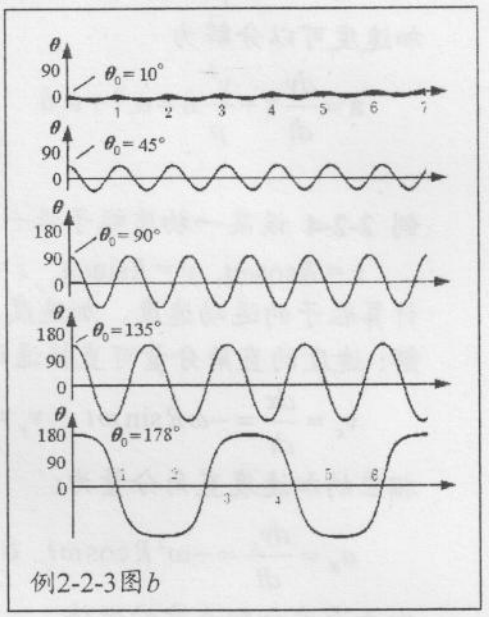
函数$\theta(t)$称为**Sn Jacobi椭圆函数**。初始角速度ω_o为零，摆锤的初始角位置$|\theta_o|\ll 1$所对应的小幅摆动有近似表达式: $2(\cos\theta-\cos\theta_o)\approx\theta_o^2-\theta^2$，上述方程简化为

$$\int_1^{\theta/\theta_o}\frac{dx}{\sqrt{1-x^2}}=\pm\sqrt{\frac{g}{l}}t$$

或

$$\theta=\theta_o\sin\sqrt{\frac{g}{l}}t$$

即退化为简谐振动。例图 b 给出不同摆幅 θ_o 的单摆摆角随时间的演化情况。◀



例2-2-3图*b*

自然坐标系

一般的曲线运动反映的是质点在不同位置处的运动状态变化，从观测属性而言，包含了质点速度的大小和方向的变化率，分别称为**切向加速度**和**法向加速度**。这两个概念可用于描述一般的曲线运动特征，设想将一个直角坐标系固定在运动质点上，其中的一个坐标轴始终指向粒子的运动方向，即指向粒子运动轨迹 ℓ 的切线方向 $\boldsymbol{\tau}$，而另一个轴 $\mathbf{n}$ 的方向维持在运动轨迹的**瞬间平面内**，并与轨迹的方向垂直(见图 2-2-4)。这种分解方法的优点在于所有分量仅与轨道各处的几何特性有关，而与坐标系的选择无关，该坐标系称为**自然坐标系**。粒子的速度

$$\mathbf{v}=\frac{d\mathbf{r}}{dt}=\frac{d\mathbf{r}}{d\ell}\frac{d\ell}{dt}=v\boldsymbol{\tau}\qquad(2\text{-}2\text{-}22)$$

其中$d\ell$是弧元的长度，因而有

$$\boldsymbol{\tau}=\frac{d\mathbf{r}}{d\ell},\quad v=\frac{d\ell}{dt}\qquad(2\text{-}2\text{-}23)$$

将质点的加速度分解为切向和横向两个分量

$$\mathbf{a}=\frac{d\mathbf{v}}{dt}=\frac{dv}{dt}\boldsymbol{\tau}+v\frac{d\boldsymbol{\tau}}{dt}\qquad(2\text{-}2\text{-}24)$$

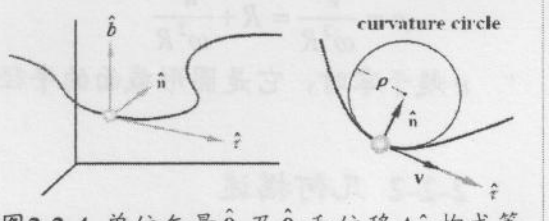


图2-2-4 单位矢量$\hat{\boldsymbol{\theta}}_1$及$\hat{\boldsymbol{\theta}}_2$和位移$\Delta\hat{\boldsymbol{\tau}}$构成等腰三角形。

其中单位矢量 $\boldsymbol{\tau}$ 关于时间的微商可以用类似计算 $d\hat{\mathbf{r}}/dt$ 的方法得到:

$$\frac{d\boldsymbol{\tau}}{dt}=\frac{d\boldsymbol{\tau}}{d\theta}\frac{d\theta}{dt}=\frac{|d\boldsymbol{\tau}|}{d\theta}\omega\mathbf{n}=\rho\omega\mathbf{n}\qquad(2\text{-}2\text{-}25)$$

其中 ρ 称为曲线的内切圆半径，由于曲线的弯曲程度可由在相应位置处的内切圆的弯曲程度表征，通常称该内切圆为**曲率圆**，而内切圆的半径成为曲线在该点的**曲率半径**。而**曲率半径** ρ 的倒数称为曲线在相应位置处的**曲率**，记为 κ。质点的横向加速度分量可以表示为

$$a_n = \rho\omega = \frac{v^2}{\rho} \tag{2-2-26}$$

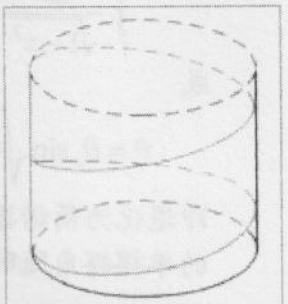
加速度可以分解为

$$\mathbf{a} = \frac{dv}{dt}\hat{\boldsymbol{\tau}} + \frac{v^2}{\rho}\hat{\mathbf{n}} = a_\tau\hat{\boldsymbol{\tau}} + a_n\hat{\mathbf{n}} \tag{2-2-27}$$

例 2-2-4 设某一物质粒子沿一条螺旋轨道(见本例图)运动，运动参数方程为

$$x = R\cos\omega t,\quad y = R\sin\omega t,\quad z = ut$$

计算粒子的运动速度，加速度，以及运动轨道的曲率半径。


例 2-2-4 图

解: 速度的直角分量可直接通过计算关于时间的微商求得:

$$v_x = \frac{dx}{dt} = -\omega R\sin\omega t \qquad v_y = \frac{dy}{dt} = \omega R\cos\omega t \qquad v_z = u$$

相应的加速度直角分量为:

$$a_x = \frac{dv_x}{dt} = -\omega^2 R\cos\omega t \qquad a_y = \frac{dv_y}{dt} = -\omega^2 R\sin\omega t \qquad a_z = 0$$

由于速度和加速度的数值

$$v = \sqrt{v_x^2 + v_y^2 + v_z^2} = \sqrt{\omega^2 R^2 + u^2}$$

$$a = \sqrt{a_x^2 + a_y^2 + a_z^2} = \omega^2 R$$

都是常数，加速度的切向分量为

$$a_\tau = \frac{dv}{dt} = 0$$

则整个加速度都属于向心分量:

$$a_n = \frac{v^2}{\rho} = \omega^2 R$$

由此可得运动轨道的曲率半径

$$\rho = \frac{v^2}{\omega^2 R} = R + \frac{u^2}{\omega^2 R}$$

u 趋于零时，它是圆形截面的半径，螺旋运动退化为平面上的匀速圆周运动。◀

2-2-2 几何描述

物体运动的解析描述可以包含非常完整的信息和特征，但在许多场合它们并不直观。例如，单摆运动中摆锤位置的解析描述(见例 2-2-3)为

$$\int_{\theta_o}^{\theta} \frac{d\theta}{\sqrt{\tilde{\omega}_o^2 + 2(\cos\theta - \cos\theta_o)}} = \pm\omega_s t$$

从这个表达式中，我们很难想象摆锤的位置是怎样随时间变化的，以及对应于不

同的摆幅，其对简谐运动的偏离程度。因而在许多场合，我们更乐意用图像的方法，而不是那些晦涩，难以理解的函数来说明物体的运动情况。

描述一个粒子的运动参量是位置 **r**，速度 **v** 和时间参数 t，粒子的运动的几何描述就是在这三组参量构建的 7 维空间中的曲线。但我们通常只是在两组参数所构筑的子空间中讨论物体的运动，其图像实际上是上述 7 维空间的曲线向相应的特定子空间的投影。例如投影到其中的 **x-y-z** 子空间得到的是粒子运动的轨迹描述，而投影到 **r-***t* 子空间得到的是粒子运动的时空演化图，投影到其中的 **r-v** 子空间得到的是粒子运动的状态描述。

演化图

系统的状态、特性、行为、功能等随着时间的推移而发生的变化，称为系统的演化。只要观察的时间尺度足够大，任何系统都是作为过程而展开的。但在每种具体情况即特定的时间尺度下，有些系统变化显著，必须作为过程来考察，有些系统几乎看不到明显的变化，可以按状态来考察。系统的演化过程是在时间维中展开的，在系统的描述中时间 t 只是一个外在的几何参数。

最基本也是最直观的关于运动的概念之一是运动的演化，我们所讨论的是系统的特性如何随时间变化的。例如，将一颗小石子投入宁静的湖水中所激起的圆形涟漪在湖面上逐次扩散。这一段历史可以在湖面上空俯瞰连续拍摄记录下来。每一张照片上标记拍摄的时间，这样一段历史事件就能通过将这些照片按时间顺序依次叠放的存放方式中保存下来(见图 2-2-5)。这样一份档案记载包含了涟漪在湖面上演化的全部细节。在物理学中称为**时空图**。用所谓时空图描述粒子的运动实际是一种观察它变老的过程，世界线描述了它在逐渐沧桑的过程中的空间轨迹。

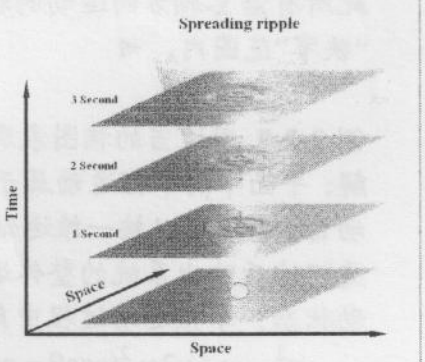


图 2-2-5 湖面上的涟漪演化的时空图

物理学家通常更喜欢将时间看成演化过程中的独立参量，而将其它所讨论的参量看为随时间变化的参量，由此可以得到描绘这些参量的演化过程。如例 2-2-3 图中的单摆摆锤位置和速度随时间演化的情况。读者可以自己绘制摆锤的角速度和角加速度的演化图(习题 2-2-3)。

轨迹图

描述物体运动演化的方程.

$$x = x(t),\quad y = y(t),\quad z = z(t)$$

是一种参数方程，若略去其中的时间参数，方程可以约化为

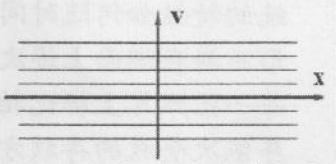
$$f(x, y, z) = 0$$

这实际上就是描述物体运动的空间轨迹方程，如例2-3-4中的螺旋轨迹。

系统的状态及其演化的整体描述—相空间和相图*

2-2-2 节已经说明，描述一个物体运动状态的参量是物体所处的位置(空间位置实际反映的是粒子相对于参考系的数学描述)和粒子的速度(速度反映相对于该参考系的运动)。这两个参数**r**和**v**构成了对该粒子机械运动状态的完全描述[5]。用粒子的位置和速度参量所建构的空间称为系统的**状态空间**，又称为**相空间**，相空间中的每个点对应于粒子的一个状态。粒子的运动状态在相空间中的几何描述称为**相图**。对于一个质点的一般运动而言，该状态空间是一个 6 维的抽象空间，而对于一维的运动，相应的状态空间退化到 2 维，可以通过几何图像在页面上描绘出来。粒子状态的演化是由运动方程 $\mathbf{a} = \mathbf{a}(t)$所决定的，其位置和速度的变化在相空间中勾画出一条**相轨道**或**流线**($\mathbf{a} = d\mathbf{v}/dt = (\mathbf{v}\cdot\nabla)\mathbf{v}$，一维运动情形退化为 $a = v\cdot dv/dx$)。运动方程的每一个解(对应于不同的积分常数)就对应于一条相轨道。

例 2-2-6 画出自由运动粒子的相轨道。

解： 自由粒子的运动演化方程为 $\mathbf{a} = 0$，或 $\mathbf{v}$ = 常矢量。如果我们选取粒子的运动方向为相图中的 **x**-轴，每一条与 **x**-轴平行的直线就对应于一个特定速度的匀速运动。在经典物理学中，所有的可能运动应将填满整个 **x-v** 平面，其中上半平面对应于粒子沿 **x**-轴方向的运动，而下半平面的水平直线描述的是粒子逆 **x**-轴方向的匀速运动。但在近代物理学中。各种物体的运动速度都存在一个共同的上限—光在真空中的传播速度 c (见 2-4 节)，因此所有沿 **x**-轴方向运动的物体只能处于 $-c < v < c$ 的“狭窄”范围内。◀

例 2-2-6 图：自由运动粒子的相轨道

例 2-2-7 用适当的相图表现轻杆单摆运动的各种可能的相轨道。

解： 平面中的单摆运动属于圆周运动。由于物体位置的变化仅限于角度这一个运动自由度，可以按一维运动的方式进行描述，选用角度和角速度作为运动参量能更好地反映出系统的整体运动特征，即通过 θ-ω 平面相图表示出其所有可能的运动状态。单摆运动过程中角速度和角度之间的函数关系(见例 2-2-3)为

$$\omega^2 - \omega_o^2 = 2\omega_s^2(\cos\theta - \cos\theta_o)$$

其中(θ_o, ω_o)分别为初始时刻单摆的摆角和角速度。为简单起见，设单摆的初始摆角为零，则对应于逆时针和顺时针的摆动，单摆的相轨道按其角速度的方向分为正、负两支

$$\omega = \pm\sqrt{\omega_o^2 - \omega_S^2(1-\cos\theta)}$$

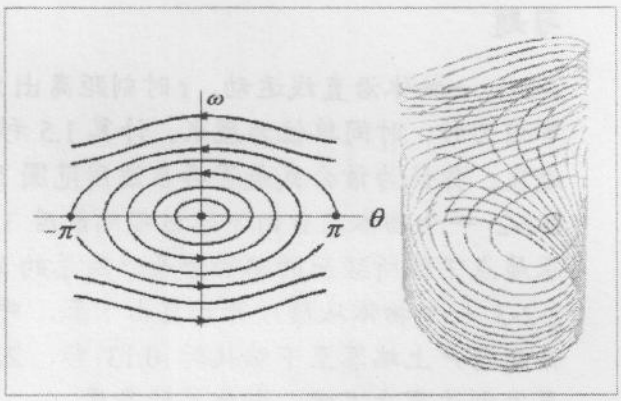
在 ω_o 很小的情况下，单摆的运动范围很小，相轨道近似是一个椭圆。很容易理解具有不同能量的无阻尼单摆运动状态在相图中对应于不同大小的相轨道。所有相轨道的集合对应于各种可能的初始条件，对于摆角小于 180°的运动而言，相

[5] 有些读者可能会认为速度与位移不是互相独立的，因为速度可以通过计算位移相对时间的导数求得。该问题的要点在于，首先时间不是物质系统本身的参量，另外从导数的定义而言，速度至少依赖于两个位置，不管这两个位置如何接近，它们在逻辑上是属于两个不同状态的参量，一个状态的参量与另一个状态的参量有关，这显然与状态参量的定义是相矛盾的。

轨道是一组共心的椭圆轨道。当系统能量达到并超过某一临界值 E_C，对应于摆动达到其角度的极限位置 $\theta_o=\pm180°$，单摆的运动状态将会发生一个突变：从往复摆动突变为单方向(逆时针或顺时针)的圆周运动：即从具有特定能量的单一轨道运动分裂为两个不同的轨道运动。

与粒子自由运动的相轨道不同的是，作圆周运动的粒子相轨道是有界的，它反映出与角度相关的物理观测效应具有周期性的特征。即 $\theta=\pi$ 和 $\theta=-\pi$ 在测量上是等价的。这种相图的描述存在一个于实际观测不同的地方，实际观测到的运动在空间位置上的变化是连续的，但一个作匀速运动的物体在例图 a 中却不是这样，它在从 $\theta=-\pi$ 运动到 $\theta=\pi$ 后必须以不连续的方式跳跃回到 $\theta=-\pi$ 位置，然后再开始新一轮的运动。这显然和坐标系本身不具有周期特性有关。

一个好的相图描述，除了描述物理系统运动状态变量必须满足完备性(状态变量的数量足以描述系统的全部可能状态)和独立性(任一状态变量都不能表示为其他状态变量的函数)以外，还应该尽可能多地反映系统运动的其它状态特征，如运动的延续性，连续性，以及各种对称性等等。因此描述粒子圆周运动的相轨道在直角坐标系构成的相空间内所具有间断性，可以采用柱面坐标系来反映圆周运动运动的连续性和运动的周期性特征。例图-b给出的是刚性轻杆和摆锤构成的单摆系统的各种可能运动状态在柱面坐标相图上的相轨道。◀



下面对状态和描述状态的相图做一些进一步的说明。状态是用以反映系统定性性质的概念。在相图(状态空间)中讨论系统运动特性的目的并不在于求解每一条具体轨道，而在于从系统所有可能的相轨道的集合中了解轨道的类型及其分布，从而在整体上理解系统的运动规律和特性。例如，在相空间中用几何图形直观地表示出系统所有可能的定态，标明定态的类型、个数、分布以及每个定态周围的轨道特性和走向。

状态空间是在控制参量给定的条件下建立的。控制参量的变化有时只是引起系统行为特性的某些量变，而有时则可能导致系统定性性质的改变。以控制参量 $c_1, c_2, \ldots, c_m$ 为坐标轴构造的 m 维空间称为参量空间。参量空间的每个点都对应一个确定的系统，所以，在参量空间研究的是演化方程结构相同的无穷多系统构成的系统族，而不是单个系统，更不是系统的某条轨道。控制参量的变化不改变系统演化方程的数学结构，但可能改变系统的动力学特性，包括定性性质的改变，即系统相图结构的改变。这种变化包括：定态的产生和消失，稳定性的改变，定态的类型、个数及其在相空间分布的改变等。

在状态空间和参量空间中描述系统及其行为，统称为**状态空间方法**。有时需要同时在状态空间和参量空间中研究系统，即以全部状态变量和控制参量为轴支撑起来的高维空间，称为乘积空间。特别地，当只有一个状态变量 x 和一个控制参量 a 时，把系统放在乘积空间即 x–a 平面上研究最为简便有效。

思考题

2-2-1 说明芝诺"飞矢不动"的悖论中国古代的名家惠施也提出过"飞鸟之景，未尝动也"的类似说法。

2-2-2 在例 2-2-2 中下落物体因受到空气阻力而获得的减速度正比于物体的运动速度：$\mathbf{a}_r = -\beta\mathbf{v}$。通过量纲分析探讨比例常数 β 与哪些物理参量有关。实际大气中的 β 数值和什么有关？估计其数量级。

2-2-3 一个一维运动质点的完整轨道是三维空间 x-v-t 中的一条曲线，它在三个坐标平面上的投影分别形成该粒子运动的时空演化图，速度演化图和状态轨道。将例 2-2-2 中的两种落体行为在上述 x-v-t 三维空间中的描绘出来。

2-2-4 讨论一个粒子以不同方向和大小的角速度作圆周运动的相轨道。

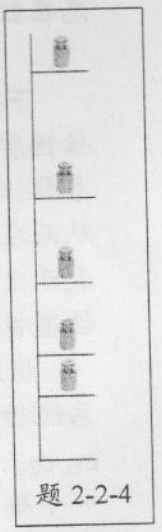
习题

2-2-1 一物体沿直线运动，t 时刻距离出发点 $x = 0.010t^3 - 0.050t^2 + 1.5t$，其中时间单位为秒，时间单位为厘米。计算 1.5 秒及 10 秒时刻，以及开始 10 秒内的平均速度。你认为该公式是否存在适用范围？

2-2-2 一个物体在自由下落过程的最后 1 秒内下落的距离是全部路程的三分之一，求质点下落所经历的路径全长以及总的时间。

2-2-3 一个物体从楼顶开始自由下落，中途经过一高度为 1.5 米的窗户，测得物体自窗户上端落至下端耗时 0.13 秒，2.00 秒后物体由地面反弹(完全弹性碰撞)再次到达窗户下端，求楼房的高度。

2-2-4 讨论一种测量重力加速度的实验设计：将 n 个重物分别放置在不同高度的平台上(如图)，然后将它们同时推出平台边缘使之自由下落，重物落地时可以听到声响。若要听到重物相继落地发出声响的时间间隔相等，应如何设置各个平台的高度 $h_1, h_2, h_3, \ldots$。



题 2-2-4

2-2-5 A sphere of radius r is rotating around its center. Show that the velocity of a particle fixed on the surface of the sphere is $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$ at any moment, where $\mathbf{r}$ is the position vector of the particle, and $\boldsymbol{\omega}$ is the angular velocity at the moment we concern.

2-2-6 如何在球坐标系和柱坐标系中表述质点的运动速度。

Chapter 3 No. 77, 78, 79

2-2-9 简谐振子的运动方程为 $x = A\cos\omega t$，其中参量 A 和 ω 分别为振子的振幅和角频率。绘制振子的相图。

2-2-10 绘制轻杆单摆摆锤的角速度和角加速度的时间演化图，画出单摆的运动周期与摆幅之间的曲线，并做简短评论。

2-2-7 绘制并讨论在一个倒立圆锥内表面上自由运动的质点的各种可能的相轨道(设内表面是光滑的)。

1-4 寻求物理学规律的常用分析方法

科学对自然规律的探索通常起始于观察，但观察和数据一般并不导致科学的结论。物理学对一种物质现象或某些物质特性的认识和理解大都涉及到物质系统对外部扰动的响应或自身的内在演化过程，认识和描述其中所涉及到的各种因素之间的关联通常都会历经下面三个具体过程：

(1) 找出问题所涉及的各种可能的因素，定义或确定相关的物理量，针对这些物理量设计并实施具体测量：通常是改变各种可能影响物质运动行为的外部条件，测量相关物理量的变化情况。

(2) 在实验观测过程中系统对不同的外部扰动所表现出来的不同程度的响应，建立相应的数学关系。这种关系只是对实验现象的唯象描述，因而称为**经验公式**。公式中一般包含有一些具有物理量单位的参数，但并不能说明其在过程中的含义，因而不能明示过程中各种因素间具体的因果关联。建立经验公式的一个很大的优点在于能够将相关参量之间的数学关系外推到实验所不能达到的范围。

(3) 从经验公式中归纳，总结出各种参量之间的因果关系，即物理定律(即从最基本的物理定律得到相关现象的特定物理规律)。或者从更基本的物理规律在特定的约束条件下演绎出实验定律。

上述分析过程在实际研究中大都不是单向的，更多的时候需要经历多次反复[1]。本节中将通过一些具体的例子帮助读者对常用的基本推理步骤获得一些初步的印象和理解。在以后的章节中也会经常予以应用，帮助读者熟悉，并积累经验。

1-4-1 从实验数据归纳经验公式的几何分析方法

分析物理过程中所涉及的物理量之间的关联的最直接方法是从对一组实验测量数据建立经验公式着手，寻找其中的定量关系。

设想物理系统的某种特性 Q 在一个物理过程中可能涉及到的物理量为 q_1, q_2,..., q_n，它们之间的关联在形式上可以表示为：

$$Q=f(q_1, q_2,..., q_n) \tag{1-4-1}$$

为了确定上述函数的具体形式，我们可以在实验中固定其它参量，通过改变物理系统的某一个参量 q_i，测量参量 Q 的相应变化(响应)，从而构成一组反映物理系统某一种性质的实验数据记录列表$\{q_i, Q\}$。列表时通常将自变量(或运算后的自变量)尽可能处理为均匀递增的形式，这样的数据排列可以大致反映出函数的主要的变化特性：如响应是否单调、或是否为周期性变化，最大值、最小值的大概位置及数量大小，不同区域的变化率特征等。也可以通过对原始数据进行特定运算后重新列表，以更清晰地反映两类参量之间的关联。

---

[1] 事实上，对任何物理系统或现象的特性分析和理解都包含定性和定量两个方面，定性描述是定量描述的基础，而借助定量描述能使定性描述更加深刻和精确。单一的定性描述难以对系统的行为和特性有深入准确的理解和把握，而单一的定量描述无论如何精确也无法达到对问题本质的认识。因此，问题的定性特性决定了它的定量特性，但定性特性是通过定量特性表现的。

实验数据的几何表示

由于表格数据的离散特性以及数据本身所含有的误差，一般不容易看出描述响应参量变化规律的准确函数形式。因此，通常先将一组实验数据在特定的直角坐标系 Q-q 中绘制成曲线以直观地反映变量之间的关系 $Q=f(q)$。为了修正实验数据所包含的误差，可以在图像中用长度对应于其误差的线段(称为**误差棒**)来标示实验点，而所绘制的光滑曲线应落在所有数据点的误差棒范围内，并尽可能通过或接近所有的实验点。而从解析的角度上说，应使曲线与实测值之间的误差的平方和达到最小，这一般对应于曲线两边的实验点数接近相等[2]。表 1-4-1 列出了一些固体材料在 20°C 附近测得的热膨胀系数，其中 α 为线膨胀系数，β 为体膨胀系数。图 1-4-1 是这些数据在以 α 为横轴，α 为纵轴的坐标系中的分布，以及根据实验测量值绘制成的曲线。

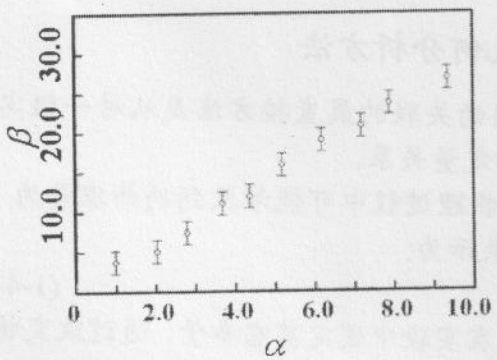


图 1-4-1 固体材料的线膨胀系数 α 与体膨胀系数 β 之间的关系。

表 1-4-1 一些固体材料在 20°C 附近的热膨胀系数，α 为线膨胀系数(10^{-6}/°C)，β 为体膨胀系数(10^{-6}/°C)

Material	α	β
Aluminum	23	69
Benzocyclobutene	42	126
Brass	19	57
Carbon steel	10.8	32.4
Concrete	12	36
Copper	17	51
Diamond	1	3
Ethanol	250	750
Gallium(III) arsenide	5.8	17.4
Gasoline	317	950
Glass	8.5	25.5
Glass, borosilicate	3.3	9.9
Gold	14	42
Indium phosphide	4.6	13.8
Invar	1.2	3.6
Iron	11.8	33.3
Kapton	20	60
Lead	29	87
Magnesium	26	78
Mercury	61	182
Molybdenum	4.8	14.4
Nickel	13	39
Oak	54	162
Pine	34	102
Platinum	9	27
PVC	52	156
Quartz (fused)	0.59	1.77
Rubber	77	231
Silicon	3	9
Silicon Carbide	2.77	8.31
Silver	18	54
Sitall	0.15	0.45
Stainless steel	17.3	51.9
Steel	11.0 ~13.0	33.0~39.0
Tungsten	4.5	13.5
Water	69	207

实验数据的数学表示—经验公式

求解曲线所对应的解析表达式称为**回归分析**。由于得到的公式一般并不完全准确地表达全部相关参量以及实际的适用范围，也不反映这种关系背后的内在物理原因，因而通常被称为**经验公式**。根据一组测量数据建立相应的经验公式通常并不是一件容易的事情，也没有固定的规律可循。在很大程度上取决于个人的经验和判断能力，而且一般要经过多次反复才可能得到与测量数据更接近的公式。

[2] 这种数值拟合方法称为**最小二乘法**，现在由许多数学软件可用于做各种回归分析，但实际上依赖于软件中的函数库和原始实验数据的误差分布，对于误差分布不规则，往往会因为软件包中的函数库缺乏相应的函数类型而得不到正确的结果。

这里仅列出两种对进一步分析有所帮助的常用方法。

直线法则

人眼视觉在辨识一组分布在直角坐标系中的离散点所对应的函数曲线的能力是很弱的，例如，一个单摆只是在摆动辐角很小的情况下，摆锤的角位置和时间的关系才很接近正弦函数曲线 $\theta=\theta_o\sin(\omega t+\varphi)$ (见图 1-4-2a)，但在我们的视觉印象中摆辐达到 135°的摆锤角位置和时间的关系曲线(见图 1-4-2b)与正弦曲线几乎没有什么差异。但我们的视觉对判断一组离散点的分布是否能形成直线却非常敏感，例如，图 1-4-1 中的实验数据点并不严格处在一条直线上，但我们从直觉上倾向于认为属于直线分布。而且很容易绘出这条直线

$$Q=kq+b \tag{1-4-2}$$

使各实验数据点与直线的离散程度完全处在实验的误差范围之内。通过直线量得其斜率和截距，从而确定方程中的参数 k 和 b。

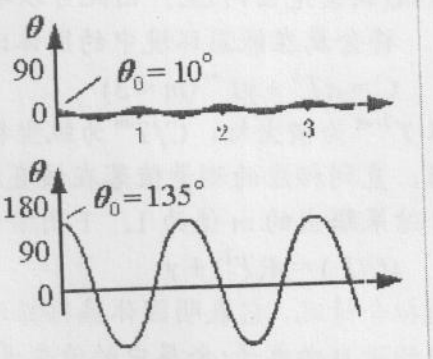
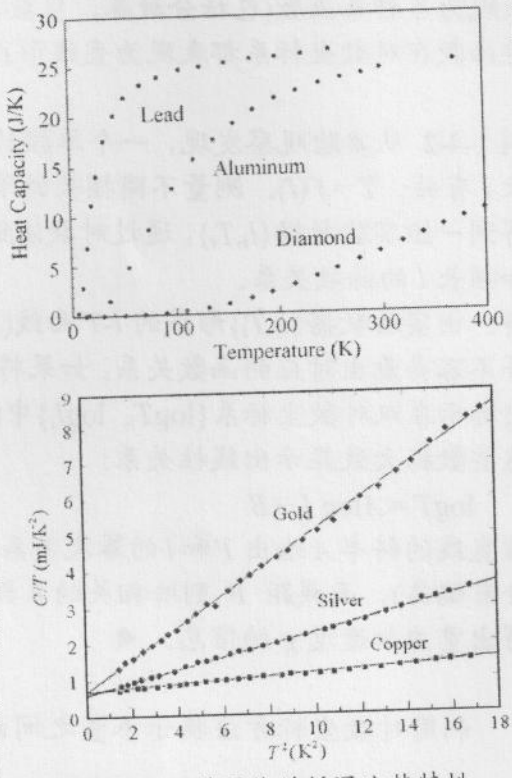


图 1-4-2 单摆摆辐为 3°和 100°时摆锤的角位置和时间的关系曲线。

由于各种物理过程中的两个物理量之间的关联并不总是线性的，而仅仅通过视觉印象很难判断出其属于哪一类函数。这时，如果能通过某些分析大致确定实验数值所勾画出的曲线函数类型：$y=f(x;a,b,\ldots)$，其中 $a,b,\ldots$ 为未知参数，我们就可以通过对原来的变量进行数学变换：$(x,y)\rightarrow(X,Y)$，即重新定义坐标系的坐标，使之成为直线的函数关系。然后再对变换后的实验数组(X,Y)进行直线拟合。这种方法可用于验证原来的分析的正确性，也可以由拟合得到的结果确定其中的参数。下面通过一个具体的例子来说明这种方法的应用。

例 1-4-1 物理量“比热”量度的是单位质量的物质升高温度 1℃ 所需要的热(能)量。实验发现，固态物质在低温区中的比热 C 与温度 T 有关。例图(a)给出了固体铅和铝，以及金刚石在不同温度下的比热测量值。利用直线法则找出图中低温区的比热与温度的关系曲线所对应的函数。

解： 图中的数据在低温区的变化趋势显示出铅和铝的比热随温度的增长比金刚石快很多，以后的变化趋势则趋于相同。由 $C(T)$在低温区的泰勒(幂函数)展开可知，铅和铝的比热关于温度的幂次低于金刚石，考虑到铅和铝与金刚石的



例 1-4-2 图 几种固体的低温比热特性。

差异在于前两者属于金属，其中含有自由价电子，而后者为绝缘体。由此可以推断两者比热对温度响应的差异部分来自价电子，剩余部分是固体晶格的贡献。由于金属中的价电子相当自由，其行为与原子晶格的关联很弱，可以认为两者对比热的贡献彼此独立。因而可以将金刚石在低温下的比热与温度的关系表示为

$$C = AT^n$$

以 T^n 为横坐标绘制 C-T^n 直线，然后试探不同的 n 值，直到相应的测量值落在该直线(误差范围内)上。由此可以确定指数 n。实际拟合结果给出的 n 值为 3。

将金属在低温环境中的固体比热与温度的关系表示为

$$C = AT^3 + \gamma T^m \ (m < 3)$$

以 T^{3-m} 为横坐标，C/T^m 为纵坐标绘制直线 $C/T^m = AT^{3-m} + \gamma$。再试探不同的 n 值，直到相应的测量值落在该直线(误差范围内)上以确定指数 m。实际的直线拟合结果给出的 m 值为 1。下图给出的是几种金属的低温比热实验值与直线

$$(C/T) = A(T^2) + \gamma$$

的拟合情况，它表明固体晶格振动对比热的贡献与温度的三次幂成正比，而固体中的近自由电子(金属中的价电子)的独立运动所贡献的比热与温度成正比[3]。直线法则还给出了斜率 A 和截距 γ 的具体数值，这些参数有助于建立并进一步优化相关的理论模型。◀

对数坐标

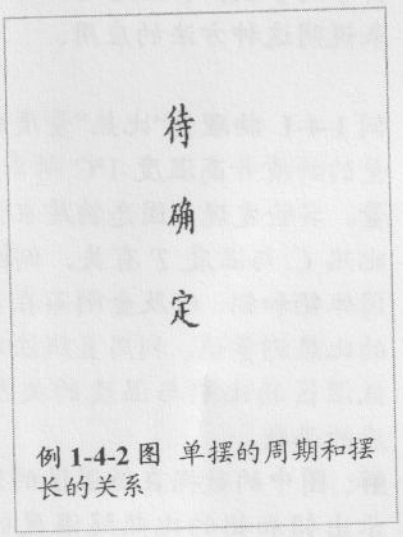
上述利用坐标变换，再应用直线法则确定变量之间函数关系的前提是我们能够对曲线的函数类型做出预判。如果做不到这一点，对实验数据进行对数变换在很多情况下可能会有助于判断函数类型。这是因为物理参量之间的函数关系常常表现为各种幂函数(包括分数幂，负数幂等)，指数函数，对数函数等形式，而这些函数在对数坐标系都表现为直线形式。

例 1-4-2 从实验观察发现，一个单摆的周期 T 和摆长 l 有关：$T=f(l)$。测量不同摆长的单摆周期可以得到一组实验数据 $\{l_i,T_i\}$，通过对数法则确定周期 T 和摆长 l 的函数关系。

解：由实验数据 $\{l_i,T_i\}$ 形成的 l-T 曲线(见本例图 a)并不容易看出对应的函数关系。如果将实验数据 $\{T_i, l_i\}$ 标示在双对数坐标系 $\{\log T_i, \log l_i\}$ 中(见本例图 b)，这些数据大致显示出线性关系：

$$\log T = A\log l + B$$

该直线的斜率 A 给出 T 和 l 的幂次关系½(实际结果会有偏离)，而截距 B 则给相关的系数，从中可以得出重力加速度 g 的信息。◀



例 1-4-2 图 单摆的周期和摆长的关系

利用对数坐标方法揭示参量之间的函数关系的同时也缩小了图上的误差显

[3] 14-4 节中对固体比热有比较详尽的讨论。

示，这导致以此确定的参量可能具有较大的误差，得到的经验公式不是很可靠。实际分析中可以用对数坐标得到的结果作为对可能函数形式的一个初步预判，再用直线方法配合使用来求得经验公式。

对数坐标的另一个用途是能够对物理系统的渐近行为的测量给出比较恰当的自变量取值的设计。对于涉及范围很大的，但变化幅度迅速减小的物体量，均匀采点测量是很低效的，在图形表示中也很难表示非常大的自变量范围。如果采用对数坐标，可以在变化幅度迅速收敛的区域采用逐步增加自变量采点区间的测量方法，而实验曲线的变化也会比较均匀。图 1-4-3 所示为 1897 年，德国物理学家考夫曼(W. Kaufman)在通过放电管中阴极射线时观察到 e/m 值随电子速度的改变而改变，如果横坐标采用 log(v/c)可使测量时采点更有效率，由此确定物体质量随速度变化的函数关系更为准确。

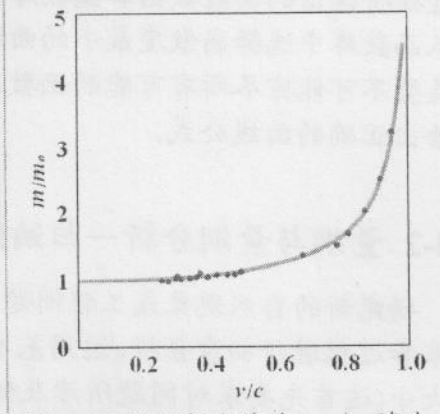


图 1-4-3 阴极射线荷质比 e/m 随速度的变化。

类似曲线的类比

许多不同的自然现象中两种参量之间的变化往往会表现出非常相似的趋势。如果一种曲线的数学表达式是已知的，这一表达式就可以作为拟合出另一条曲线经验公式的出发点(试探公式)。例如：图 1-4-4(*a*)是物体在不同温度 T 下的热辐射在不同波段中强度分布。对图中所示的曲线形状拟合出确切的数学描述并不容易。这项工作以及对热辐射现象内在机制的研究在 19 世纪后期和 20 世纪初曾经吸引了不少当时最杰出的物理学家。1893 年维恩注意到经典热学理论中描述处于热平衡的理想气体中分子速率 v 的分布图形(图 1-4-4*b*)与热辐射谱曲线形态相近，而气体中分子速率 v 的分布函数当时是已知的

$$f(v,T)=Cv^2\exp\left(-\frac{mv^2}{2k_BT}\right)$$

式中 v 为气体分子的速率，T 为气体的温度，k_B 为玻尔兹曼常数。虽然从两种曲线形态上看不出这两个区域的行为在频谱分布和频率的确切函数关系，但形式上可以引入试探函数：$r(\lambda,T)=C_1\lambda^{\alpha}\exp(-C_2/\lambda^{\beta}T^{\gamma})$，通过数值拟合求得相应的参数

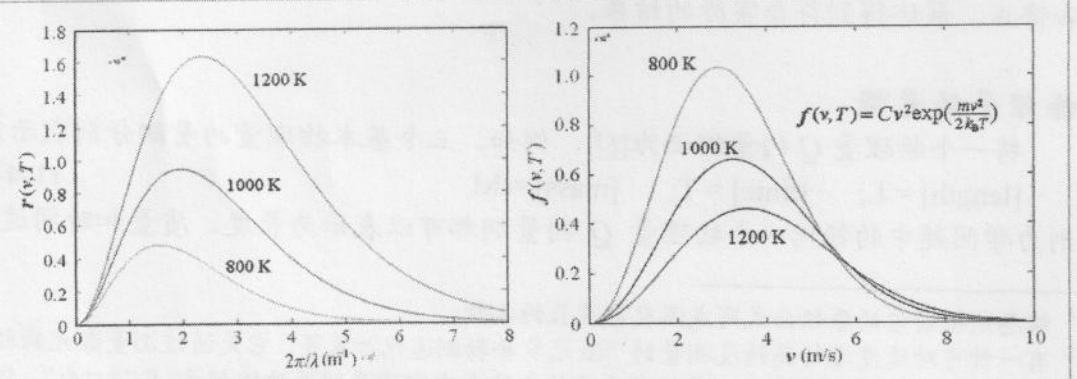


图 1-4-4 (*a*)不同温度的黑体辐射频谱。(*b*)气体分子的速率分布曲线。

值 α, β 和 γ。最后得到的热辐射谱经验公式为[4]

$$r(\lambda, T) = C_1 \lambda^{-5} e^{-C_2/\lambda T}$$

其中 C_1 和 C_2 是通过拟合实验观测值得到的常数。

需要说明的是，现在许多通用数学软件中都备有函数库，很多实验设备也可以直接将读出的实验数据和函数库中的各种函数进行自动对比，再通过最小二乘法从函数库中选择离散度最小的曲线，并给出匹配的参数值。由于函数库中的函数类型不可能穷尽所有可能的函数，甚至可能的初等函数组合，因而并不总能给出给出正确的曲线公式。

1-4-2 量纲与量纲分析—归纳经验公式的代数分析方法

研究新的自然现象或工程问题中的规律，首先要对其中所蕴涵的物理环节、关系和过程进行初步分析，运用基本物理学规律分析相关参量在过程中所起作用的大小(这首先要求对问题所涉及的物理量按属性进行分类，因为只有同类物理量才能比较大小)，确定问题中起控制作用的参量。然后找出不同类别的物理量之间的相互关联。从而得到尽可能明确的函数关系。

在 1-2 节中有关物理量的单位和测量的讨论中曾提到，任何物理量的测量最终都会归结到对长度，时间及质量这三种基本物理量的测量，即任何物理量的测量属性都可以表述为长度，时间及质量三种基本测量的特定组合。例如，物质密度所涉及到的测量是单位体积中的物质质量，其中质量本身是一个基本测量，而体积则分解为三个独立的长度测量。再例如，速度可以分解为长度和时间的两个独立的测量[5]。由此可见，“长度”，“时间”及“质量”在物理量中的意义并不仅仅是为实际测量提供操作的基本量具那样简单，它们实际上揭示、并给出了物理量中隐含的基于其操作性定义所具有的一种能够定量描述的性质，这种属性称为量纲。量纲实际是区别不同物理量属性的指标之一。每一个物理量通过自己特有的量纲表现出在相关物理性质中与其他物理量测量属性之间的内在关联，并在形式上通过将其量度单位以长度，时间及质量单位的特定代数组合表现出来。这意味着对于有些无法借助或采用现有的物理模型和数学方程的复杂现象或问题的研究，量纲分析可能成为分析问题，设计合适的模型实验来揭示问题的物理实质，从而明确因果联系的有效的方法。但运用量纲分析方法必须和对问题的基本物理内涵的分析结合起来。这一方面需要经验以及对现象的洞察力，也需要多次试探和修正，最终得到符合实际的结果。

物理量的量纲

将一个物理量 Q 的量纲记为[Q]，例如，三个基本物理量的量纲分别表示为

$$[\text{length}] = \text{L}; \quad [\text{time}] = \text{T}; \quad [\text{mass}] = \text{M} \tag{1-4-3}$$

则力学问题中的任何一个物理量 Q 的量纲都可以表示为长度、质量和时间这三

---

4 维恩热辐射谱的原始公式写成温度和波长的函数。

5 有一种可对速度进行高精度测量的“激光多普勒测速仪”装置，它是通过测量激光的频率移动或波长变化得到速度的。虽然从速度的表达式中它不再明显地依赖于“尺”和“表”，但频率测量和波长测量涉及的仍然是时间和长度测量(频率为时间周期的倒数)。

个基本量的量纲的幂次表达式[6]:

$$[Q] = \mathrm{L}^{\alpha}\mathrm{T}^{\beta}\mathrm{M}^{\gamma} \tag{1-4-4}$$

即一个物理量的量纲可表示为三个指数的集合(α, β, γ)。例如，面积S的量纲为$[S] = \mathrm{L}^2$，速度的量纲为$[v] = \mathrm{LT}^{-1}$。而对一个与长度、质量和时间均无关的纯数a而言，其量纲$[a]=\mathrm{L}^0\mathrm{T}^0\mathrm{M}^0$。对于任何一个物理问题而言，出现在其中的各个物理量的量纲或者由物理量本身的操作性定义给出，或者由相关的物理定律决定。

读者务必要注意到物理量量纲与物理量单位之间的区别。从概念上说，量纲是物理量的性质，而单位则是量度该物理量的具体量具；而从表述上说，一个物理量的量纲是唯一的，但这个物理量的单位可以有许多。例如速度的量纲为LT^{-1}，但速度的单位可以是$\mathrm{m\cdot s^{-1}}$，也可以是$\mathrm{km\cdot hr^{-1}}$。

量纲分析

通常的自然现象会涉及到多种因素，理解这些现象或探索现象背后的自然规律也就归结为寻找这些物理量之间的相互关联，甚至演化过程中的因果关系。虽然有些现象或问题的研究可以借助已有的数学模型及相应的方程，但通常也会遇到一些没有现成的模型和方程可以利用的新现象、新问题。在这种场合，量纲分析常常是用来分析问题，揭示问题的物理实质，设计合适的实验方案，最终确定因果联系的非常有效的方法。

例 1-4-3 利用量纲分析说明自由落体的速度v和落体的质量无关，并确定落体的速度v和下落距离h之间的关系。

解: 这是一个因物体受到地球作用力的加速而随着下落距离增加，速度越来越快的典型事例，它可以通过直接求解标准的运动微分方程得到结果(见 2-2 节)。但利用量纲分析方法可以通过简单的代数确定自由落体速度v和下落距离h之间确切的函数关系。

用量纲分析求解实际的物理问题，首先需要了解这个问题可能涉及哪些因素，这些因素应该由什么物理量来表征。对于自由落体的最终落体速度v而言，可能想到的相关参量为落体的质量m(我们可能会认为重的物体下落的快一些)，物体所受到的重力(因为落体运动与地球对物体的引力有关)而具有的重力加速度g和下落的距离h。它们的量纲分别为$[m] = \mathrm{M}$，$[g] = \mathrm{LT}^{-2}$和$[h]= \mathrm{L}$，设自由落体的速度v和重力加速度g以及h之间的关系为：$v = f(m, h, g)$。相应的量纲方程为

$$[v]= [m]^{\alpha}[h]^{\beta}[g]^{\gamma}$$

其中各相关基本物理量的指数(α, β, γ)可通过相应的量纲方程指数平衡予以确定:

$$\mathrm{LT}^{-1} = \mathrm{M}^{\alpha}\mathrm{L}^{\beta}\,[\mathrm{LT}^{-2}]^{\gamma} = \mathrm{M}^{\alpha}\mathrm{L}^{\beta+\gamma}\mathrm{T}^{-2\gamma}$$

两边指数平衡给出$\alpha = 0, \beta = ½, \gamma = ½$。由此得到自由落体速度和下落距离的关系

$$v = C\sqrt{gh}$$

其中C是无量纲比例常数。其数值可通过实验或更基本的物理定律确定。◀

例 1-4-4 利用量纲分析确定声波和相关物理量之间可能的函数关系。

---

[6] J. C. Maxwell 在 1871 年就提出了量纲的幂次表示假设。严格的证明见本章的附录。

解：由于声波是介质中传播的机械振动，传播依赖于介质的机械性质，我们从对振动的直觉理解中可以意识到，这些性质依赖于介质的惯性参量(质量)和弹性参量(相互作用)。由于分析波的传播速度时通常考虑无限大的介质，对无限长的弹性弦这样的一维介质而言，相应的参量是线质量密度μ和张力f，其量纲分别为$[\mu]=\mathrm{ML}^{-1}$和$[f]=\mathrm{MLT}^{-2}$，相应的量纲方程为：$[v]=[f]^p[\mu]^q$，则声波表达式为

$$v=\sqrt{T/\mu}$$

无量纲系数C可由实验或理论分析确定。

对于三维介质(如空气、水)而言，相应的参量是质量密度ρ和弹性模量Y，其量纲分别为$[\rho]=\mathrm{ML}^{-3}$和$[Y]=\mathrm{ML}^{-1}\mathrm{T}^{-2}$，相应的量纲方程为：$[v]=[Y]^p[\rho]^q$，可得：

$$v=C\sqrt{Y/\rho}$$

无量纲系数C可由实验或理论分析确定。◀

量纲分析可用于揭示一个物理现象所涉及的物理量之间的关联是基于物理量所具有的两个普适特征：

首先，物理量量纲反映了物理量的基本测量属性，即测量中所包含的基本物理量的测量组合。不同量纲的物理量因所涉及的基本测量不同而无法进行比较(数学上表示为两个具有不同量纲的物理量不能进行加、减运算)，因此一个物理现象中同类型的控制因素是以线性叠加(和差)的形式起作用的，而不同类型的控制因素在物理定律中是以乘积(或商)的组合方式对结果起作用的。这样，它们才能从量纲上表现出与表示最后结果的物理量之间的因果关联。

第二，虽然物理量的定义基于其测量属性，但物理现象背后的物理规律和观测所选用的量度单位(量具)无关。这在数学上表述为任何一个物理过程中所涉及到的物理量之间的函数关系式总能用无量纲方程表示，其中的物理量也能表示为无量纲参数。这些无量纲参数和无量纲函数分别称为**约化参数**和**约化方程**。

以上两点实际上也说明了量纲分析的基本应用：任何一个物理定律所对应的方程(相关物理量之间的表达式)两边的量纲必然是平衡的。比较方程两边的量纲有助于判断该方程的正确与否，也有助于探讨控制相关过程的物理定律[7]，利用现象中所涉及的无量纲参量(即约化参数)确定相关的物理定律。下面对此做进一步的说明。

一个物理问题中所涉及的多个物理量中，总可以将这些物理量分为**基本量**和**导出量**两类。基本量是指相应问题中具有独立量纲的那些物理量，导出量则是指其量纲可以表示为基本量量纲的组合的物理量。问题中所涉及的这些物理量之间的关联一般总是表述为有量纲的物理量方程

$$Q=f(q_1, q_2, \ldots, q_n) \tag{1-4-1}$$

如果选择其中的独立控制参量$q_1, q_2, \ldots, q_k$为该问题中的基本物理量[8]，并以这**些基本量的集合作为问题中各参量的基本单位系统**，则上式可写成无量纲形式

$$P=f(1, 1, \ldots, 1; p_1, p_2, \ldots, p_{n-k}) \tag{1-4-5}$$

---

[7] 否则必须引入带量纲的比例常数，而这通常意味着我们对相关物理过程的理解还不够充分。

[8] $q_1, q_2, \ldots, q_n$中的次序是无关紧要的，我们可以在n个控制参量中先选择k个基本量，并记为$q_1, q_2, \ldots, q_k$

其中 P 是对应于 Q 的无量纲量，而 p_1, p_2, …, p_{n-k} 分别为对应于导出量 q_{k+1}, q_{k+2}, …, q_n 的无量纲约化参量。其一般形式为

$$p_i = \frac{q_{k+i}}{q_1^{\alpha_1} q_2^{\alpha_2} .. q_k^{\alpha_k}} \tag{1-4-6}$$

略去方程(1-4-5)中所有的常数参量 1，而得到对应于方程(1-4-5)的约化方程

$$P = f(p_1, p_2, \ldots, p_{n-k}) \tag{1-4-7}$$

由于方程(1-4-7)中所有的参量大小都与单位无关，相对于(1-4-1)式更能反映问题的本质。它所揭示的真正内涵是，任何物理现象背后的物理定律与观测所选用的度量单位无关。

下面利用矢量相关性的几何图像帮助读者理解上述物理量量纲的特性。

1-4-4 小节中讨论矢量的定义时曾经提到，具有直观的方向属性对一个矢量而言并不是必须的，重要的在于其各分量要满足矢量代数运算规则。由于任何一个物理量的量纲可以表述为三个独立量纲指数的集合(α, β, γ)，因此长度，时间及质量的量纲 L, T, M 可以理解为一个特殊三维空间(称为量纲空间)中的一个直角坐标系的三个坐标轴。所有物理量 Q 的量纲$[Q] = \mathrm{L}^{\alpha}\mathrm{T}^{\beta}\mathrm{M}^{\gamma}$ 都是该空间的矢量，其中的指数 α, β, γ 可以理解为$[Q]$在 L, T, M 轴的分量。通过对数变换可以将物理量 Q 的量纲$[Q]$所具有的这种矢量特性描述的更加直观:

$$\log[Q] = \alpha\log\mathrm{L} + \beta\log\mathrm{T} + \gamma\log\mathrm{M} \tag{1-4-8}$$

其中指数 α, β, γ 就直接表示为三个轴上的投影值(见图 1-4-3)。例如速度的量纲为: $[v] = \mathrm{L}^1\mathrm{T}^{-1}\mathrm{M}^0$，相应的矢量表示为(1,–1,0)。

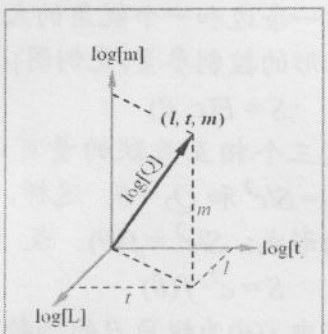


图 1-4-3 量纲矢量在量纲空间中的分解。

由于三维空间中最多只能有三个矢量是相互独立的，如果一个物理问题中所涉及到的物理量有 N 个: $\{Q_i, i = 1, 2, \ldots, N; N > 3\}$，而其中的$[Q_1]$, $[Q_2]$, $[Q_3]$是相互独立的，则其它 $N-3$ 个物理量的量纲可以表示这三个物理量量纲的组合[9]:

$$[Q_k] = [Q_1]^{X_k}[Q_2]^{Y_k}[Q_3]^{Z_k} \qquad (k = 4, 5, \ldots, N) \tag{1-4-9}$$

因此，物理量可以按照其量纲的属性分为两大类。第一类物理量的大小与测量时所选用的单位有关，如速度、加速度、力、能量等，称为**有量纲参量**；另一类物理量的大小与测量时所选用的单位无关，其量纲指数 α, β, γ 都为零(例如两个同类物理量的比值等)，称为**无量纲参量**。对于上面所提到的涉及 N 个物理量$\{Q_i; i = 1, 2, \ldots, N; N > 3\}$的物理问题，这些量至少可以构成 N–3 个无量纲量

$$[P_k] = [Q_k]^{-1}[Q_1]^{X_k}[Q_2]^{Y_k}[Q_3]^{Z_k} \qquad (k = 4, 5, \ldots, N) \tag{1-4-10}$$

上述量纲方程组实际上给出了所讨论问题所涉及的物理定律:

$$Q_k = Q_1^{X_k} Q_2^{Y_k} Q_3^{Z_k} f(p_1, p_2, \ldots, p_{n-k}) \qquad (k = 4, 5, \ldots, N) \tag{1-4-11}$$

在物理学中，有量纲和无量纲物理量所起的作用不同,有量纲参量反映的是特定现象中相关因素的特定性质，而无量纲参量所反映的是相似现象所具有的共同特征和所遵循的普适规律。具体而言，一个有量纲的物理量的大小只具有相对的意义，而无量纲参量的数值却具有某种绝对的含义。例如，一根米尺的长度在数

[9] 若相互独立参量只有两个，甚至只有一个，问题所涉及的量纲空间就是二维或一维的。

量上可以表示为 1,000(毫米), 或者 1(米), 或者 0.001(公里)。不同的数值取决于所选用的测量单位。但对于无量纲参量，其数值大小是确定的，不可能通过改变度量单位进行放大或缩小。从这个意义上说，无量纲参量所揭示的是一个物理系统，或者物理过程中的某种总体或者一般的性质，而有量纲的参量反映的则通常是一个系统或者过程的某种特定的性质。例如，一个三角形的几何结构和性质是由它的“边”和“角”决定的。其中“边”具有长度量纲，而“角”则是无量纲参量(角度可通过两条边的比值来表述)。如果一个三角形的三条边是完全确定的，这个三角形也就被完全确定了。但如果只是确定了三个角(相当于确定了三条边的相对比例)，所确定的是一族形状相同，但大小不确定的相似三角形，即确定的是这一类三角形的总体特性。另外，如果一组无量纲参数满足某种数学关系，这种关系通常反映的是更一般的普适特性。仍以三角形为例，一个三角形的三个内角和(也是一个无量纲参量)的数值小于，还是等于或大于 180°揭示该三角形所处的空间的几何结构，分别对应于罗巴切夫斯基、欧几里德或黎曼空间。

例 1-4-4 利用量纲分析方法证明勾股定理

解: 欧几里德空间中的一个直角三角形 ABC(见图)的面积 S 可以由该三角形的任意一条边和一个锐角的大小所确定。如果选择弦长 c 和锐角 A 的大小 θ 为该三角形的控制参量(见例图)，则三角形 ABC 的面积 S 可以表示为 c 和 θ 的函数

$$S = F(c, \theta)$$

这三个相互关联的量可以构成两个无量纲参量 $Q_1 = S/c^2$ 和 $Q_2 = \theta$。这样，上述函数可写成无量纲的形式: $S/c^2 = f(\theta)$。或

$$S = c^2 \cdot f(\theta)$$

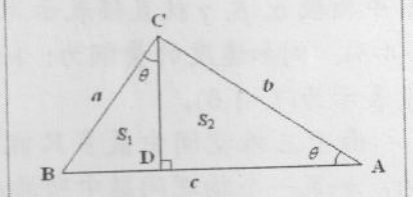
其中 $f(\theta)$为锐角 θ 的函数。由于三角形 ABC 与三角形 CBD 及三角形 ACD 相似，这三个三角形的面积和相应的弦长和角度应满足相同的函数关系:

$$S_1 = a^2 \cdot f(\theta), \quad S_2 = b^2 \cdot f(\theta)$$

因三角形 ABC 的面积是三角形 CBD 及 ACD 的面积之和

$$S = S_1 + S_2 = (a^2 + b^2) \cdot f(\theta) = c^2 \cdot f(\theta)$$

而 $f(\theta)$不等于零，由此可以得到: $a^2 + b^2 = c^2$。◀



例 1-4-4 图 直角三角形的边和角具有不同的量纲属性。

上述辨识几何图形类别的方法可以推广用以辨识物理现象的类别和认识物理问题的规律。虽然描述物理现象或问题的物理量和几何图形不同，除了长度，还有时间、质量等其它属性的物理量，但我们总能在与这类物理现象或问题相关的物理量中，选定一组基本物理量取作单位系统用以量度这类现象中的任何物理量，由此得到的无量纲物理量的大小就反映出这类现象的本质。进一步说，如果在反映问题的物理规律或因果关系中，所有变量都采用上述量度方法得到的无量纲的数值，由此得到的反映无量纲变量之间的函数关系也必然客观地反映了这类现象的本质。下面通过一个简单例子—单摆的周期来说明运用量纲分析的方法分析问题，揭示问题中所涉及的物理量之间的函数关系。

例 1-4-5 单摆周期 单摆是由长度为 l 的摆弦，一端固定，另一端悬挂一个具有一定质量 m 的摆锤构成，如果摆弦的质量及在摆动时摆弦发生的形变都可以忽略，也无须考虑摆锤在运动过程中所受到的空气阻力，利用量纲分析求解单摆在重力作用下的摆动周期与系统各种参量之间的函数关系。

解：由于单摆是悬挂(与摆弦长度 l 有关)的摆锤(与质量 m 有关)在重力场中(与重力加速度 g 有关)摆动(与摆幅 θ 有关)，与摆动周期相关的系统参数有 4 个：摆锤质量 m、摆弦长度 l、重力加速度 g 以及单摆的摆幅 θ，因此形式上有

$$T=f(l,g,\theta,m)$$

这全部的 5 个参量在量纲空间可以构成 5–3=2 个无量纲参量。通过这些参量的量纲对比可以得到这 2 个参量。它们分别为：

$$Q_1=T\sqrt{g/\ell} \quad 和 \quad Q_2=\theta$$

其中 Q_1 和 Q_2 都不包含 m 意味着单摆周期与摆锤的质量无关。由此可以得到单摆周期的约化方程：$T\sqrt{g/\ell}=f(\theta)$。单摆周期为

$$T=\sqrt{\ell/g}\,f(\theta)$$

由于 θ 是无量纲参量，量纲分析无法确定 $f(\theta)$的形式[10]。但从单摆运动图像上可以判断，$f(\theta)$是偶函数。当单摆的摆幅很小($\theta \ll 1$)时，可以将 $f(\theta)$在平衡位置($\theta=0$)附近作泰勒展开：

$$f(\theta)=f(0)+\tfrac{1}{2}f^{(2)}(0)\theta^2+\cdots$$

在一阶近似下，单摆周期可以表示为：

$$T\approx f(0)\sqrt{l/g}$$

无量纲常数 $f(0)$可通过实验或求解运动方程得到，确切的数值为 $f(0)=2\pi$。◀

上述例子体现了量纲分析在建立模型和导出结论过程中的两个要点：一是只能在同类参量之间比较大小，二是物理现象和物理规律与所选用的度量单位无关。虽然量纲分析不能完全确定一个物理定律，但确实给出了确定的相关参量之间的函数关系，从而可以大大提高实验的效率。例如，如果完全通过实验测量来确定单摆周期与四个控制参量的函数关系：$T=f(m,l,g,\theta)$，即使对每个参量都只做大小不同的 10 次测量，一共需要做 10^3 次实验(将 g 视为常量)；而通过量纲分析，只要做 10 次实验以确定 $f(\theta)$就够了。不仅提高了实验效率，也相应提高了测量精度。而对于小摆幅振荡，只要做一次实验就可确定常数 $f(0)$的数值。

1-4-3 物理规律中的相似性—标度律

约化方程(1-4-7)以及其中所涉及的如(1-4-6)形式的约化参量的形式可用于描述物理现象及物理过程，这一事实揭示出物理规律具有的一种普遍性质：一类物理现象或过程中的相关带量纲参量$\{Q_i\}$发生变化时，只要保持约化方程中各无量纲约化参量$\{P_i\}$不变，描述该现象的特征参量 P 就不会发生变化。由于方程(1-4-7)描述的通常是一类特定的物理现象或物理过程，这就意味着只要与该特性相关的

[10] $f(\theta)$只是表示 θ 的某个函数，并不是和前面的 $f(m,l,g,\alpha)$有相同的函数形式。为简洁起见，后面将经常用这种形式表示变量之间所存在的函数关系，而不特指函数 f 的具体形式。

若干参量在发生变化时相互之间满足的某种特定的约束，就可以维持物理系统的相关行为特性不变。物理规律所具有的这种性质称为**相似律**或**标度律**。

最简单、也最直观的相似律是**几何相似律**。上一小节中关于三角形形状的讨论中已经提到，对于不同大小的三角形，只要其三条边 a, b 和 c 所构成的比值对应相等，这些三角形就一定是相似的。这种相似性在实际观察中表现为无量纲量 b/a 和 c/a 的测量值不随观察者与被观察的三角形之间的距离变化而改变。另外，用 a^2 为单位量度该三角形的面积 S 得到的值 S/a^2 也不随观察者距离的远近改变。对几何相似律的认识并付诸实际应用的事例可以追溯到意大利文艺复兴时期。当时的艺术家已经知道如何在画布平面上描绘出三维的视觉景象：根据实际物体与观察距离的远近以一定的比例缩小或放大在画面上安排各个物体的图像可以在视觉上维持物体不出现变形，从而在画面上给出了另人愉悦的空间层次的视觉效果。这种后来被称为“**透视**”技法的要点在于在画面上确定一个唯一的构图视点，将几何学上的三维空间以适当的比例安排到画面上[11]。图 1-4-4 为意大利文艺复兴时期的最具代表性的艺术家达芬奇的最具代表性的作品之一的《最后的晚餐》，以及绘图草稿上将三维空间中不同位置的人和物以中心投射的方式放入平面，从而使每个图形的位置准确，并保持与其他图形的正确比例关系，使整个绘画作品在视觉上构成一个和谐、完美的整体。

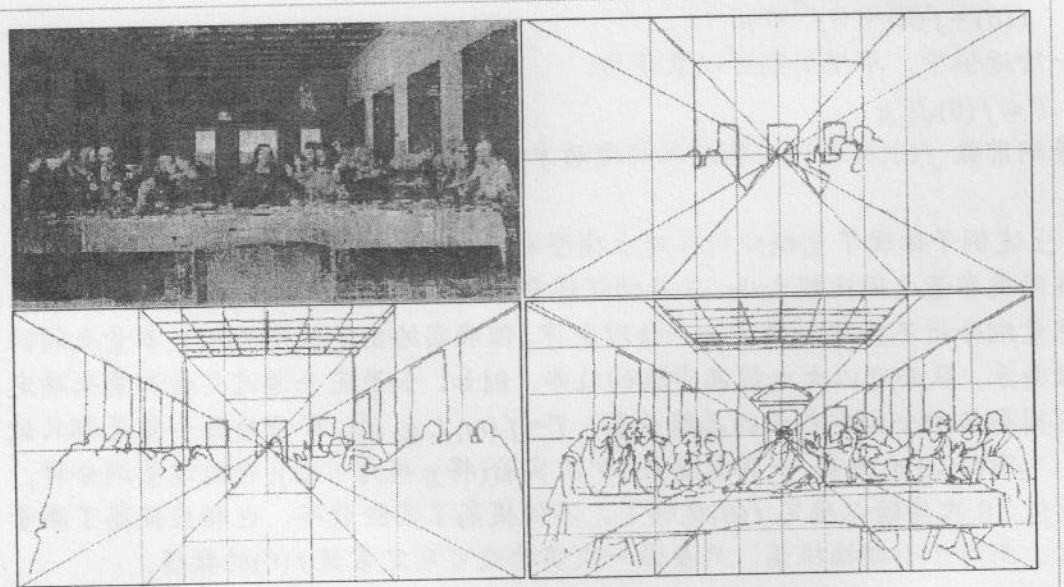


图 1-4-4 为《最后的晚餐》及其构图的透视分解

几何相似性涉及的只是由基本量纲“长度”所描述的单一静态空间结构问题，相关的参量也只有长度，面积和体积。而实际物理问题通常要复杂得多，其内在的物理规律所含的相似性除了几何相似，还涉及**运动学相似**和**动力学相似**。所谓运动相似性所涉及的是由“长度”和“时间”为基本量纲的纯粹运动描述问题，它和几何相似性问题一样，都不涉及物理学中的由相互作用决定的系统演化的因果关联。而动力学相似性则涉及由“长度”、“时间”和“质量”全部三个基本量纲描述的

[11] 历史记载中第一个凭直觉悟出这种绘画技巧的是乔托，他在构图时将视点放在一个静止不动的点上，并由此引出一条水平轴线和一条竖直线，从而在视觉产生了第三维度—深度，在平面画布上开设出欧几里得的三维空间概念。

系统演化问题。事实上，一个物质系统只有在它与外界没有任何相互作用的情况下才能等比例地缩放其空间尺度并维持其物理或“生存”特性不变。即物理相似性表现为几何相似性。由于自然界中并不存在这样的情况，物质能保持其特性的时空尺度的范围就成为能维持一个实际物理系统能维持其几何标度不变性的范围，即系统的几何标度不变性受到物体系统本身相关性质的制约。由此可见，几何相似是运动相似和动力相似的前提，而运动相似是几何相似和动力相似的表现形式。

虽然一般的物理相似性问题要比单纯的几何相似性问题复杂得多(包含了更多的量纲也就牵涉到更多的可能性)，但上述利用几何相似求解问题的方法可以推广用以辨识物理现象的类别和认识物理问题的规律。因为我们总能在控制这类物理现象或问题的物理量中选定一组基本物理量作为单位，以此度量这类现象中的所有物理量，这样得到的无量纲物理量的大小就反映出这类现象的本质，类似于一个物理上“相似形”，如果能找到一个物理问题所涉及的无量纲参量，也就得到了相应的一类物理现象中物理量之间的相互关系。或者说，如果在反映问题的物理规律或因果关系中，所有自变量和因变量都采用上述度量方法得到的无量纲的数值，这些反映无量纲的因变量与自变量之间的因果关系也必然反映了这类现象的本质。这一小节将通过对一些简单的具体事例向读者展示，一个物质系统为了能在实际的物理世界中“生存”而不得不屈服于某些物理约束(运动学或动力学特性)所表现出来的一些有趣的现象。特别地，一些物质系统为了维持物理特性空间画布上的“形状”相似性。不得不自毁其在视觉空间画布上的几何相似性—即自身的形状。

重力效应

首先讨论一下我们直觉中对人体比例的想法。大多数人对自身形体的想法都是出自美学的考虑，希腊雕塑中的人体的比例反映了大多数人对自身体形的一种完美的印象，由于视觉对人在远近在我们的印象中形成的几何相似性，我们在下意识中认为，一个人的比例(高矮胖瘦匀称性)是不应该随着他的高度而有很大变化的。例如乔纳森·斯威夫特 1726 年发表的讽刺小说《格列佛游记》中就有类似的描述。小说中的主人公格列佛在其游历中曾途经一个小人国和一个巨人国。其中的人物与物品与格列佛原来生活的世界的尺度分别以 1/12 或 12 倍的比例分别被缩小或放大。斯威夫特对他心目中的小人国和巨人国的情景描写实际上反映了人们对几何标度不变性的直觉期望，即我们所看到的物体空间尺度比例的不变性。这是由于运动相似性和动力学相似性远不如几何相似性那样直观，而人们的思维惯性又往往倾向于将在一定范围内形成的意识以及适用的知识无限制地外推，物理标度不变性因我们的视觉对周围世界的有限感知而常常被无意识地简单化为单一的几何标度不变性而出现直觉与客观事实的矛盾。

实际上，自然界存在这种巨人的可能性早在斯威夫特发表他的小说的前一个多世纪就被伽利略所否定[12]。与图像中的人物不同，各种物体(包括我们自身)由于受到地球的重力作用而必须具有一定的力学强度才能维持各自特有的几何构型。例如，液体因为不具有这种强度，其形状就只能由承载它的容器所决定，

12 见伽利略所著《关于两门新科学的对话》一书

海蛰等软体动物只能趴在地面上。而一般具有形态的固态物体所承载的负荷如果超过了其本身的强度极限也会瘫痪。从宏观上说，固态物体的强度一般与受力的截面积大小成正比。例如，如果拉断一根绳索所需要的力为 F，两根这样的绳索缠在一起所能承受的最大拉力就会增长一倍，达到 $2F$。物体的强度与截面积大小的正比关系不仅适用于张力，也适用于压力的情况。一根柱子所能承受的负载和冲击极限也与柱子的截面成正比。

斯威夫特小说中巨人的外形比例与格列佛相同，三个维度的尺寸都是格列佛的 12 倍，因此巨人腿骨的截面积及其强度是格列佛的 $12^2 = 144$ 倍。但其腿骨所必须承受的重力(体重)则与其身体的体积成正比，是格列佛的 $12^3 = 1728$ 倍。即巨人们的骨头单位面积所承受的自身负载是格列佛的 12 倍，这相当于格列佛成天背着 11 个和他自身重量相等的人，这种负担显然不是一个正常人所能承受的。

为了说明巨人为了保持其足以支撑自身的体重所必须满足的条件，设想一个动物的形态特征主要由其高度 l 和宽度 d 表征，而其骨头的强度(单位面积所能承受的应力) γ 应正比于其截面积：

$$\gamma = c_1 d^2$$

其中比例常数 c_1 的大小包含了骨架的材料强度的特征，这对不同大小的动物而言是基本相同的。由于骨架的强度必须足以支撑身躯的重量，因此骨架的最小强度 $\gamma_{\min}$ 也应该与动物的质量成正比：

$$\gamma_{\min} = c_2 m$$

由于动物的质量与 l 的立方成正比 $m = c_3 l^3$，这两个特征强度的比值是一个无量纲常数

$$\alpha \equiv \frac{\gamma_{\min}}{\gamma} = \frac{c_2 c_3 l^3}{c_1 d^2} = c\frac{l^3}{d^2} \tag{1-4-12}$$

参量 α 所反映的是一个动物静止站立的**强度稳定系数**。其中的参数 c 与骨头的几何结构特征和物质特性有关。对于大多数哺乳动物而言差不多是一个常数。对于任何物体而言在地面上一相同的强度稳定系数维持一个姿态就意味着比值 l^3/d^2 必须保持为常数。如果长度 l 增加了 k_l 倍，骨头的粗细 d 必须相应地增加 k_d 倍以维持足够的强度以支撑体重。因此(1-4-12)式中的比值是由力学结构而不是几何结构决定的，即

$$\alpha = c\frac{l^3}{d^2} = c\frac{(k_l l)^3}{(k_d d)^2} = c\frac{k_\ell^3 l^3}{k_d^2 d^2} \tag{1-4-13}$$

必须保持不变，由此得 $k_d = k_l^{3/2}$。若身高增长到原来的三倍，其骨头的粗细必须增长到原来的 $3^{3/2} \approx 5$ 倍才足以支撑体重增长(见图 1-4-6)。而对于身高为格列佛的 12 倍的巨人，他的骨头粗细就必须达到格列佛的 $12^{3/2} \approx 42$ 倍。

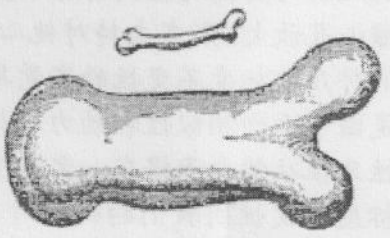


图 1-4-6 为维持骨骼的强度，骨头变长时会变得很粗。

我们大概不会有格列佛那样的运气去见识大人国中巨人们特有的风采，但我们周围却有着足够的事例支持上述分析得到的结论。各种哺乳动物因其体形大小不同而具有迥异的体形比例与上述推断大体一致。例如，羚羊有着纤巧的四肢和脖颈，而河马为了支撑其巨大的头颅和身躯，其颈项和腿骨则显得粗短笨拙(见

图 1-4-7)。

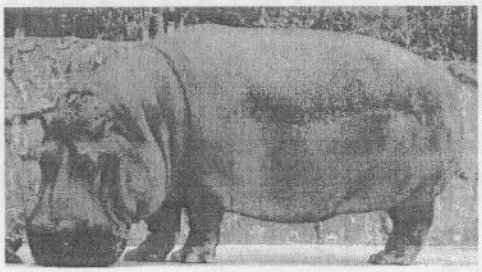


图 **1-4-7** 羚羊与河马的形体比例的巨大差异。

容易理解在重力效应不显著的环境中的动物形态应该不受到这种限制。例如在水中生活的鱼类。由于它们所受到的重力与水的浮力大致抵消了，大多数的鱼类具有非常相近的体形比例。蓝鲸这种地球生态圈中最庞大的动物的长度可以达到 30.5 米，而其长度和粗细的比例与大多数一般尺寸的鱼类基本相同。图 1-4-8 中所示为长度仅为 0.3~0.4 米的鲱鱼和蓝鲸的形体。鲸虽然能在海洋中以其苗条的体形自由，敏捷地做出各种另人惊叹的姿态，但一旦在近海搁浅，其巨大的身躯会立即导致其因肋骨骨折而死亡。

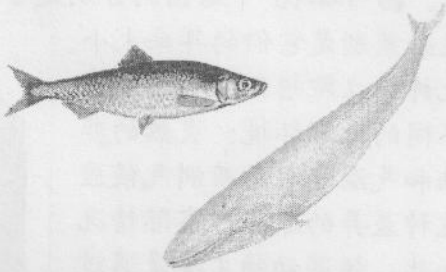


图 **1-4-8** 鲱鱼和蓝鲸的形体比例相近。

陆生哺乳动物的体形比例必须随着其身躯的大小而改变是取决与这样一个事实：式(1-4-12)中描述支撑动物身躯的骨架材料和特征的参数 c 是大致相同的，如果换以其他的参数，就可以突破(1-4-13)式规定的比例，仍能维持 α 不变的标度不变性。例如，我们的骨架如果用高强度的钢材替代，我们就能变得更纤细苗条，行动也能更敏捷，更不容易受伤。一个相应的实际例子是美国纽约的前世贸中心各有 110 层的双塔(见图 1-4-9-a)。每一个楼型比例与人体相当，但高度达到 412 米，是一般人身高的 250 倍左右。楼体截面积不足对几何标度的改变由大楼建筑混凝土墙体中的钢筋骨架得到强度补偿，而这类建材在二、三层楼的一般房屋建筑中是不需要的。这种用高强度材料对建筑的强度补偿作用在 2001 年 911 事件中得到了悲剧性的演示说明。被劫持的大型民航客机自杀式地撞击世贸大楼上部，造成整幢大厦在约一个小时内完全坍塌(见图 1-4-9-b 和-c)。导致大厦坍塌的致命因素并不是飞机对塔楼上层撞击造成的破坏，而是航空燃料的快速燃烧产生的巨大热量所导致的近 1300°C 的高温，楼体内的钢筋骨架在这样的温度下完全失去了强度[13]，不能再对建筑其到所设计的支撑作用。

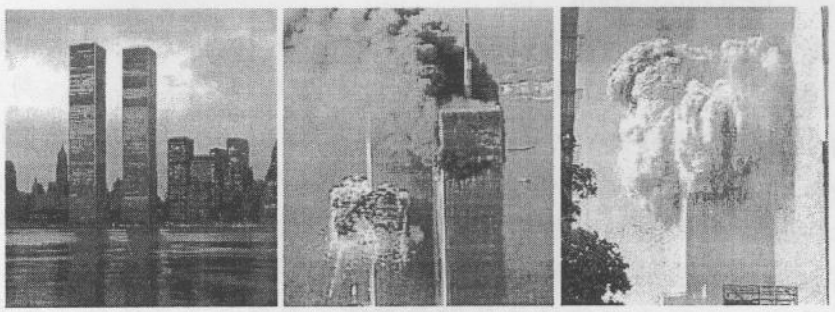


图 **1-4-9** 世贸中心各有 110 层的双塔。

[13] 实际上钢筋在 500°C 就失去了强度，

动物骨架的强度除了支撑自身的体重，还必须承受因运动(加速度)造成的额外的力的冲击。动物的躯体越大，也就越笨重，进行相同类型的运动必然承受更大的作用力。这也说明小型动物往往行动敏捷，而大型动物则往往比较笨拙。从事同一种运动项目的运动员，身躯高大的相对容易受伤，而运动幅度大的部位也更容易受伤。如径赛运动员和球类运动员最容易受伤的位置大都为脚踝部位。

表面效应

生物世界中生物的各种与空间维度相关的功能还有许多。也因此存在着与生物功能所要求的不变性相关的标度律。其中的一个例子是生物体的表面积与体积的关系。图 1-4-10 中给出的分别是非洲象和亚洲象的照片，这两个相同的物种的最显著差别是它们的耳朵大小。读者也许会立即想到两种象所生活的不同的地理环境：炎热的非洲大陆和气温宜人的亚洲气候应该是这种差异的原因。实际情况确实如此。恒温动物主要是通过皮肤散去体内过剩的热量(只有很少一部分是通过呼吸来散热的)，因而散热的效率与动物的表面积成正比。

图 **1-4-10** 非洲象与亚洲象耳朵的差异

但问题是热量的产生却是发生在整个动物体内，即与动物的体积成正比，因此身体的散热效率为：

$$\eta \equiv \frac{\text{heat loss}}{\text{heat production}} = \frac{c_1 \ell^2}{c_2 \ell^3} \Delta T = \frac{c}{\ell} \Delta T \tag{1-4-14}$$

为了维持体内温度的恒定，动物体内产生的热量必须和通过皮肤散发的热量平衡，这就要求无量纲参量 η 不变。由于其中比例常数 $c = c_1/c_2$ 与皮肤的织构和体内的生物学过程有关，对非洲象和亚洲象而言，这个参数差异很小，这意味着只能要求(1-4-14)式中的比值 $\Delta T/\ell$ 保持不变。对非洲象而言，温差的减小可以由减小体型得到体内热量的平衡，或者选择能生长出一些表面积很大而体积却很小的器官以增加总体的 S/V 的比值。在维持体积不变的情况下增加表面积意味着非洲象需要有一些近乎二维的体外附属器官。自然界显然选择了后者(一个明显的理由是个头变小容易成为其他动物的猎物)，它的两个巨大的耳朵充满了丰富的血管，大量的热量可以有血液的循环带到身体表面。同样的理由可以解释在沙漠中生活的狐狸比生活在极地的狐狸有大得多的耳朵(见图 1-4-11)。而实际上常用的散热器也

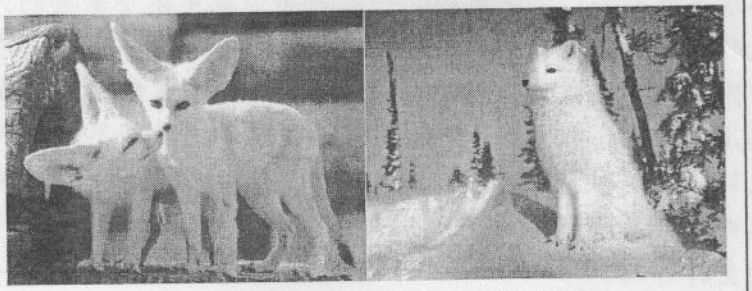


图 1-4-11 沙漠狐狸和极地狐狸的耳朵尺度的差异。

有着同样的几何设计。

大型动物因为其自身的表面积和体积比不够大而为散热效率低痛苦，小型动物却必须为其自身的表面积和体积比过大而为散热过快烦恼。由于不可能通过缩小表面积来减缓散热，一般只能通过改变表面构造(厚实松软的毛皮)以及加速体内产热的效率，因此这些动物往往食欲惊人，每天的食量可占其体重的很大比例。

例 1-4-6 设格列佛的身高为 1.7 米，体重为 70 公斤(与我们的身体相当)，每天的的食物量为 1 公斤。利用标度律估计小人国居民每天所必需的食物数量。

解： 由于食物对于恒温动物而言主要是补充身体散热即运动导致的能量流失。因而他们对食物的最低需求只是因身体和环境的温差导致的散热，热量的流失率正比于身体的表面积。设格列佛和小人国居民身体的质量和所必须的最低食物数量分别为 M 和 m，身体的特征长度为 l，则维持热量平衡有关的标度律可以表示为

$$\frac{m_G}{m_L}=\frac{l_G^2}{l_L^2} \quad 以及 \quad \frac{M_G}{M_L}=\frac{l_G^3}{l_L^3}$$

将这两个方程的两边分别相乘，可以得到

$$\frac{m_L}{m_G}\frac{M_G}{M_L}=\frac{l_L^2}{l_G^2}\cdot\frac{l_G^3}{l_L^3}=\frac{l_G}{l_L}$$

由此得到

$$\frac{m_L}{M_L}=\frac{l_G}{l_L}\cdot\frac{m_G}{M_G}=12\cdot\frac{1}{72}\approx\frac{1}{6}$$

即小矮人每天的最小食量要达到其体重的六分之一。因此他们必然是一帮饕餮之徒，不停地进食就会产生消化问题，也很容易过量，这通常需要依靠运动来消耗掉过剩的能量，因此他们也必然患有类似多动症的行为特征。小矮人的这些行为特征可以通过观察许多大小类似的小动物(例如老鼠)的生活习性得到印证。◀

上面的例子说明，对一个物理系统而言，一个给定的无量纲参数通常决定了该系统所处状态或运动类型的特征。一个物理系统可能会有几种不同的运动形式，每一种运动形式都对应于一个无量纲参数。系统为了维持某一运动状态或特性，就必须保持与其相关的无量纲参量的不变性。这种不变性称为**物理相似性或标度不变性**。物理相似性是指系统所处运动状态的相似性，而不是几何形状的相似性。由于无量纲参量一般可以由若干个带量纲的物理量组合而成，这就意味着如果其中某一个物理量发生了一定程度的变化，为了维持标度不变性，即保持给定的无量纲参数决定的运动类型及其特征不变，其他参量必须发生相应的改变。

物理学中到处可以用到标度律，特别是对于才开始研究的新系统而言是极其有价值的。系统的行为是如何随着系统的大小尺度，运动状态等变化的信息常常是通向细节分析的最佳引导。

1-4-3 相似律的应用—模型实验

理解物理系统的标度行为对于工程学而言非常重要。这里简单介绍一下相似

律的一个非常重要的实际应用领域：**模型实验**。

实验测量在物理学中是揭示自然规律的基本方法，也是验证理论计算结果的终结手段，但许多实验由于各种因素的制约而难以实施。以工程力学中常见的横梁强度问题(常见于各种建筑如桥梁，以及机械构件的设计中)为例，横梁由于受到自身重量和各点的负载承重会发生弯曲变形，甚至断裂。其强度与自身的几何尺度、形状，以及材料的性质、负载的大小和分布都有关系。横梁设计时存在许多事先无法预知的未知参量，理论计算只能通过实验测试提供相关的数据。但又不可能建造一个相同的实体横梁用于测试，因为这就使设计失去了原有的意义。一个可行的办法是制造一个小型模型，这样不仅可以降低建造成本，在实际测试时操控也更为方便。但从前面的标度分析可知，等比例的小尺寸模型和实际尺度的物体在力学性能上可能大相庭径，因而模型实验必须针对相关的性能找到相应的约化参量，在改变尺度的同时，维持约化参量不发生变化，由此从实验得到的约化参量之间的函数关系，即相关性能的约化方程。

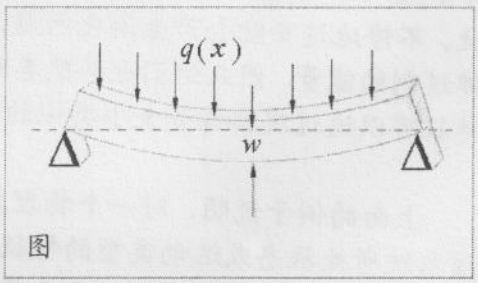
例如，设想建造一个标志型，且具有特定周期的巨型长摆(例如 50 米高)。由于摆的周期 T 与摆长 l 和摆幅 θ_0 有关，我们必须设计确定摆幅的范围，因为它会影响周围的建筑。如果我们不知道 T、l 和 θ_0 之间确切的函数关系，就只能通过实验来确定摆幅的范围。为避免进行完全相同的实体实验，我们可以利用单摆周期的约化方程：$T\sqrt{g/\ell}=f(\theta)$。其中需要通过实验确定的只是 $f(\theta)$的形式。这可以通过 5 米摆，甚至 1 米摆进行实验予以确定。

下面再以简支梁的挠曲(见图 1-4-4)为例说明这种实验设计的要点。

对给定的载荷分布 $q(x)$，梁的挠度 $w(x)$满足控制方程和边界条件：

$$EI\frac{d^4w}{dx^4}=-q(x) \qquad \text{(平衡方程)}$$

$$w=\frac{d^4w}{dx^4}=0 \quad (x=0,l) \qquad \text{(边界条件)}$$



图

其中 l 是梁的长度，E 和 I 分别是梁的杨氏模量和截面矩，它们在方程中以乘积形式 EI 出现，表征梁的抗弯刚度，量纲是$[F][L]^2$($[F]$是力的量纲)，载荷分布 $q(x)$可由特征分布载荷 q_m (量纲是$[F][L]^{-1}$)和表示分布特征的长度 l_q 表征。则挠度应为坐标 x 和上述参数的函数

$$w=f(x,l,EI,q_m,l_q)$$

其中仅含两个独立量纲，取 l 和 EI 为基本量组成单位系统，则

$$w/l=f(x/l,q_ml^3/EI,l_q/l)$$

只要保持 q_ml^3/EI 和 l_q/l 不变，就会有同一个无量纲的挠度分布

$$w/l=f(x/l)$$

可见，在截面矩 q_ml^3/EI 不变的条件下，若模型采用与原型同样的材料，就不再要求模型的截面形状遵守几何相似的条件，方截面梁、工字梁以至空心梁都可互相模拟。

这种基于标度分析的实验设计称为**模型实验**。模型实验可以用来了解和预测真实运动的性质及特征物理量的数值。

采用模型实验来了解和预测真实运动的性质及特征物理量的数值，特别是对

于那些物理模型和数学方程还不清楚的问题是非常有效也是经常采用的手段。通过模型实验或数值模拟得到了关系式

在相关现象中反映物理相似律的约化方程

$P=f(P_1, P_2, \ldots, P_{n-k})$

中，一些在实验中难以实现的参量数值的参量 Q，在相关的物理定律中往往不是单独起作用的，而是通过和其它的物理参量 $Q_1, Q_2, \ldots$ 一起构成的无量纲相似准数 P 来控制或制约相关规律的。因而我们可以设计一个模型，让模型的参量$(Q_i)_m$和原型(p)的自变量 $P_1, P_2, \ldots, P_{n-k}$ 分别对应相等：

$(P_1)_m=(P_1)_p, (P_2)_m=(P_2)_p, \ldots, (P_{n-k})_m=(P_{n-k})_p$

就能保证模型和原型的因变量 P 也相等：

$(P)_m=(P)_p$

上述模型实验或数值模拟所遵循的**相似规律**也称为**模型律**。而其中的无量纲自变量 $P_1, P_2, \ldots, P_{n-k}$ 称为决定问题本质的**相似准数**。

另一个有趣而且著名的例子是分析水在管道中的流动情况。管道中水的流动有时呈现为规则的层流，有时则表现为杂乱的湍流。影响流动形式的因素有管道直径 d，平均流速 v，流体的密度 ρ 和粘性系数 μ。1883 年，雷诺运用量纲分析方法整理模型实验的数据后发现，这些因素并不是单独起作用的，而是通过这些参数构成的无量纲参数 $R_e=\rho vd/\mu$ 作为整体进行制约的，当 R_e 从小增大到某个临界值，水流即从层流转变为湍流。这种将多个影响因素综合为一个无量纲参数进行实验分析，意味着只需采用一种流体，一个管道作为模型，通过调节流速就可模拟多种流体，多种口径的管道中的流动情况。这种将多个影响因素综合为一个无量纲的比值作为判断同一问题中出现不同现象的分析方法，是研究方法论的一个创新。R_e 后来被称为**雷诺数**。

当系统的某些参量的变化程度大到某种异乎寻常的尺度，一些方面对系统某个性能的影响会变得特别重要，而另一些方面则缩减到可以忽略，这为我们提供了一种相对于在常规尺度观察物体某一特性的一种更有针对性，因而也更清楚的视角。从而更容易揭示出某些未曾预料到的物理因果关系。由此，物理学家通过大幅度改变系统的尺度，过程的快慢，系统的温度，以及各种所能实现的不一般的条件或状态下测量系统的某种性能。例如上面提到的横梁的强度随着长、宽尺度的增加是怎样变化的，如何通过变化结构提高的强度，如何通过小型模型在风洞实验测试大型飞机的强度和飞行性能等。

思考题

1-4-1 一个物理量的量纲中的某个指数能否为任意实数?如果让一个物体的位移随时间的 $\sqrt{2}$ 次幂变化，是否可以成为上题的一个例证。

1-4-2 平时我们量度或表示一个角度的大小时都标示有单位的，如弧度，角度。为什么说角度是无量纲的量?

1-4-3 两个不同量纲的物理量是用不同量具测量的，它们的量具标度没有可比性，这两个具有不同有效位数的物理量的乘积的有效位数是否能多一些(如 $s=vt$)?

1-4-4 在 1-3 节介绍矢量的点乘和叉乘时中提到的“功”和“力矩”具有相同的量纲，

它们是否属于相同的物理量，如果不是。应如何区分它们？在中学的物理课程中学到过什么物理量具有相同的量纲，即具有相同的测量属性，却被归类为不同的物理量？

1-4-5 一个数学量是否是矢量的判据在于其各分量要是否满足矢量代数运算规则。一个物理量的量纲可以表述为三个独立量纲指数的集合(α, β, γ)，文中将长度，时间及质量的量纲 L, T, M 可以理解为量纲空间中的一个直角坐标系的三个坐标轴。所有物理量 Q 的量纲$[Q]=\mathrm{L}^{\alpha}\mathrm{T}^{\beta}\mathrm{M}^{\gamma}$都是该空间的矢量，说明这些矢量是如何满足矢量代数运算规则的。

1-4-6 根据你对如下物理词汇的直觉理解，从测量的角度定义相关的物理量，并写出它们的量纲：(1) 气体，固体，液体的压缩特性，(2) 物体的热膨胀特性，(3) 物体的硬度；(4) 光源的发光强度，(5) 光照度，(6) 光亮度，(7) 人眼的视见度，(8) 物体的透明度，(9) 物体的光泽度；(10) 表面的粗糙度。

1-4-7 为什么一般物理定律中包含的变量之间的关联都表现为乘幂(包括负幂)形式，而不是求和形式？

1-4-8 所有可能的物理量的量纲矢量能否填满整个量纲空间中的全部空间点，或填满量纲空间中的某个区域？

1-4-9 在单摆振荡的幅角 θ 不是很小的情形，其振荡周期与幅角有关。从物理上判断，$f(\theta)$ 实际上是 θ 的偶函数，但并不能确定它在 $\theta=0$ 处作泰勒展开的形式为 $f(\theta)=f(0)+A|\theta|+\ldots$ 还是 $f(\theta)=f(0)+B\theta^{2}+\ldots$。通过实验得到若干组实验数值$\{\theta_i, f(\theta_i)\}$，如何选择坐标系，从而在图形中确定上述的函数关系。

(答案：$\{y_i=[f(\alpha)-f(0)]/|\alpha|, \alpha_i^{2}\}$给出。)

1-4-10 一幅肖像画挂在房间的什么位置能达到最佳欣赏效果？

1-4-10 能否用物理标度律说明极地恒温动物通常具有比较大的形体，而许多体形很小的物种如昆虫一般都不是恒温动物的可能原因？

习题

1-4-1 从电荷的测量属性出发写出其量纲。

1-4-2 1950 年麦克斯韦，C. A. Raynold 和 B. Serin 两个研究组各自独立地测量水银同位素的超导临界温度 T_C(实验结果见表)。求这些数据满足的经验公式。

例 1-4-4 表 不同水银同位素的临界温度

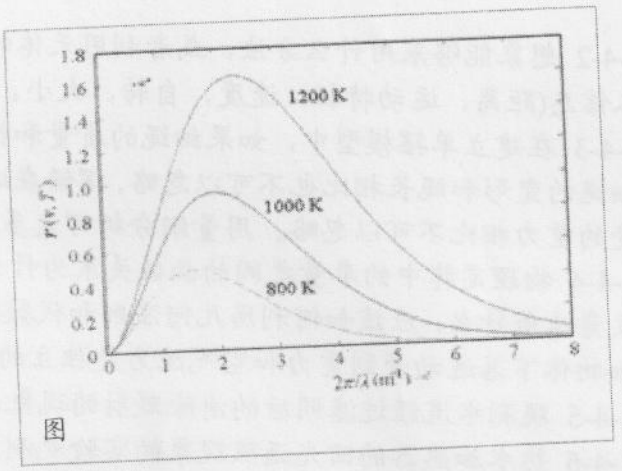
同位素原子质量	200.6	198	203.4	202.1	00.7	997
临界温度(K)	4.156	4.177	4.126	4.143	4.150	4.161

1-4-3 表中所列的数据为开普勒根据第谷的观测计算出的太阳系中各行星的椭圆轨道运动周期 T 和椭圆轨道半长轴 a 之间的关系。利用作图、拟合等方法求这两个轨道运动参量之间的经验公式 $T=f(a)$。

行星	水星	金星	地球	火星	木星	土星	天王星	海王星	冥王星
a (10^8km)	0.5791	1.082	1.50	2.28	7.786	14.34	28.72	44.95	58.70
Period T	88 (d)	225 (d)	365.25d	687d	11.86 y	29.46 y	84.07 y	164.8 y	248.5y

1-4-4 右图为一组物体在不同温度条件下的辐射随不同频率区域内的分布。指出这组曲线的最主要的特征，根据相应参量之间的对应关系，绘制出相应的曲线，

并求解出上述曲线的经验公式。



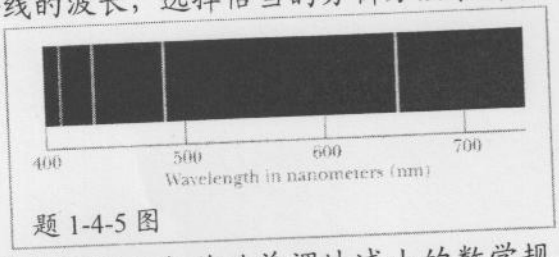
图

1-4-5 一个物体在自由下落过程中受到空气阻力 $\mathbf{F}_D = -A\mathbf{v}$，其中 $\mathbf{v}$ 为物体的速度。物体最后会达到一个恒定速度 $v_f = mg/A$。利用量纲分析估计物体达到终极速度的时间。

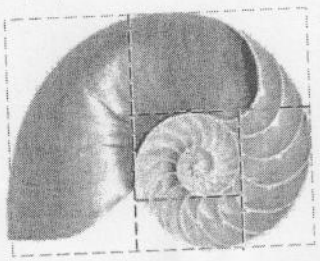
1-4-6 表中数据为声波在不同气体中的传播速度 u(单位 $m \cdot s^{-1}$)，利用你已知的知识对表中的数据进行分类，分析声速和介质的那些性质有关，尽可给出这些关联的定性或半定量的关系。

Gas	T(K)	u	Gas	T(K)	u	Gas	T(K)	u	Gas	T(K)	u
He	273.2	971.9	Cl_2	293.2	219	D_2O	373.2	451	CH_4	273.2	430
Ne	273.2	434	Br_2	331.2	149	CO_2	273.2	259	C_2H_2	273.2	329
Ar	273.2	307.8	I_2	453.2	138	H_2S	273.2	310	C_2H_4	273.2	318
Kr	273.2	213	CO	273.2	337	SO_2	273.2	211	C_2H_6	283.2	308
Xe	273.2	170	NO	283.2	324	N_2O	298.2	268	C_2H_6	304.2	316
H_2	273.2	1286	NO	289.2	334	CS_2	273.2	192	C_3H_8	273.2	238
D_2	273.2	890	HBr	273.2	200	NH_3	273.2	415	C_6H_6	273.2	177
F_2	373.2	332	HI	273.2	157	CCl_4	370.2	145	C_6H_{12}	303.2	181
N_2	273.2	337	H_2O	373.2	473	SF_6	284.2	133	C_2H_3OH	370.2	335
O_2	303.2	332	H_2O	407.2	494				C_2H_5OH	326.2	258

1-4-7 图中为氢原子的光谱线，量出各谱线的波长，选择恰当的分析方法求解表述这些氢原子光谱线波长的经验公式。可见光区域中的四条氢原子的特征光谱线，6562 Å (H_α)，4861 Å (H_β)，4340 Å (H_γ)，和 4101 Å (H_δ)(见图 9-2-2-d)。求出相邻波长的间距随着波长向短波方向移动单调地减小的数学规律性。



题 1-4-5 图



1-4-8 鹦鹉螺壳在平面上的投影给出的轮廓曲线如图，确定该曲线对应的方程，能否说明鹦鹉螺能从这种螺壳的几何外形获得什么益处。

1-4-9 如何简洁地表示一种商品，如笔记本电脑，洗衣机，彩电，一门课程的主要特性？

1-4-10 设计一种实验方法测量人对矩形的长-宽比例满足黄金分割的喜爱程度。

探讨性课题

1-4-1 用颜色作为一个维度描述一个 3-自由度的几何图形，4-自由度的几何图形

1-4-2 想象能够采用什么方法，或者利用天体的那些特性可以测量遥远天体的基本信息(距离，运动特征，速度，自转，大小，质量，温度)。

1-4-3 在建立单摆模型中，如果细绳的质量和摆锤的质量相比不可以忽略，假设细绳的变形和绳长相比也不可以忽略。摆锤在运动过程中所受到的空气阻力和所受的重力相比不可以忽略。用量纲分析可走多远？

1-4-4 物理定律中的参量之间的函数关系为什么表现为积的形式，如果是和的形式意味着什么，应该如何利用几何法则和代数法则分析物理参量之间的关联。例如物体下落运动受到重力和空气阻力是独立的。

1-4-5 观测水流经过滤网后的消除溅射的现象，探讨相关的因素。

1-4-6 热水加热器的回光返照现象的实验观测，应该测量那些相关因素

1-4-7 物理量及其相关实验设计：餐巾纸的特性的物理量定义及其相关实验设计。The most important properties of a tissue are its absorption and strength. How can we quantify these two properties? Introduce physical quantities to describe them.

1-4-8 用量纲分析方法分析风吹绳弦所发出的声调。

1-5 物理量的数量级估计—费米问题

量纲分析所关注的基本上属于或偏向于物理分析定性的方面，它给进一步的定量研究指出了大概的方向，但一个物理现象或过程中通常会包含很多因素，而在不同的空间范围，时间阶段，起控制作用的参量可能不同。如果我们能对各种因素的影响有一个数量级大小的判断，就能在一定的范围中舍弃一些权重较小因素，使相应的物理模型能更加简洁地突出问题的主要方面。例如我们可以将一个物理问题分解为几个过程，每一个过程都能明确地表现出不同的特征。如何决定描述各个阶段的决定参量和分界参量的分析时，有一种所谓的半定量分析非常有用，特别是对一个物理现象或过程所涉及的物理量的大小做出数量级估计往往在分析中起着特殊的作用。这是因为同一个物理量的数值在不同数量级区间内，其控制参量相关过程的内在原因以及所控制的物理规律一般并不相同。(如金属比热)

追求对自然现象和过程的精确描述和透彻理解永远是物理学的主题。但在追求这一目标的具体过程中，我们往往更多地是处于缺乏信息，甚至无法获得相关信息的境地。很多时候我们需要迅速作出估计，例如可行性预判，概述可行的运作路径，或检查一个复杂的数字计算结果的合理性，否决一个侯选的方案等等。虽然这样的结果并不是最终的，只是在数量级上才有意义，但这种定量的判断是非常有价值的。

意大利物理学家费米[1]是特别擅长这一类工作的代表人物，他经常通过做这种事情消遣、取乐，他常常在一张很小的餐巾纸上，甚至一个信封的背面，用几个算式给出精度令人惊异的估计和预判。因而这种计算常被称为费米问题[2]。下面先通过一个实际例子说明问题的求解思路，讨论如何估计一个实际问题中所涉及的一个重要参量的数量级大小。

例 1-4-1 估计一个成人在正常情况下有多少根头发。

解：许多人面对身边的伙伴一头密实的毛发，都会觉得这不是一个可以很容易解决的问题，在一筹莫展之际，甚至会在内心深处排斥这种问题。但对这类问题的畏惧实际上源于被问题的字面意思误导：问题要求回答一个准确的数字，而头皮上头发的分布并不均匀(如头顶的头发密度大于后脑区域)，于是我们通常会立即陷入如何才能尽快准确数出这一数字的纠结之中。但对这类问题答案的真正意义上实际并不是一个准确的数值，因为没有两个人的头发数量是完全一样的，而且也没有有效的办法核对你给出的答案准确与否。其实也没有人在乎这个数字是一

---

[1] Enrico Fermi (1901-1954) an Italian physicist best known for his contributions to nuclear physics and the development of quantum physics in both theory and experiments. Fermi was awarded the Nobel Prize for his physics work on the nuclear process.

[2] A wide spread legendary story in physics circle shows how Fermi estimated the power of the first test atomic bomb at Los Alamos, New Mexico. About nine miles from the blast, Fermi tore a paper into shreds, and at the moment that the shock wave of the blast arrived he let them down freely from his pate. Then he measured the distance of the paper blown off from him by the wave, and within seconds he declared that the energy released by the bomb is equivalent to that of ten thousand tons of TNT dynamite. The prediction was confirmed later in the data analysis recorded by some very sophisticated instruments.

千或一千零五十之间的差异。

这一问题的真正意义在于头发的密度，或者平均密度。对这一问题，你可以很快地目测一下身边伙伴的鬓发区域的毛发间距，一般是稍小于 1 毫米，如果大概估计头发密度约为 $200cm^{-2}$，这个问题就转化为求解头上毛发的覆盖面积，我们可以很粗略地将人头看成是半径约为 10cm 的圆球，而头发的覆盖面积约为该球表面的一半，则头发数量的估计值为

$$\frac{1}{2}\times 4\pi r^2 n \approx \frac{1}{2}\times 4\times 3\times 10(\text{cm})\times 10(\text{cm})\times 200(\text{cm}^{-2}) = 120000$$

中国人的头发密度的统计数值为 175±54 根/平方厘米，按照头皮面积 600 平方厘米估算，头发数约为十万根。◀

上述例子在一定程度上说明了费米问题求解的方法，其要点并不在于知晓有关问题本身的确切信息或知识，而在于如何将一个看似很专门的问题分解成一组相对常识性的问题[3]，这种分解因人而异，原则是分解为我们所熟悉的问题。对问题的分解，最直接的一些属于同类问题的细分(同一类物理量)，例如，估计一幢楼房的高度，可以将楼房分解为楼房有多少层和估算每一层的高度，每一层的高度可用层的门或窗作为单位进行估量，而门或窗的高度可以和我们自己的身高作比较。如果分解中缺乏某一层次的信息而造成信息流断裂，就需要将该问题分解为不同类型但能提供信息的相关问题。

例 1-4-7 右表为美国纽约机场所列由纽约飞往洛杉矶的航班信息。由这些信息估算地球的直径。东部城市纽约飞往西部城市洛杉矶，

纽约—洛杉矶航班表

航班	起飞	到达	机型
WN3578	08:00	13:45	739
NK706	16:45	21:30	319
US1437	14:45	20:23	321
US1437	14:45	20:23	321
UA1870	14:45	20:23	321
NW737	12:35	17:19	320
DL2685	12:35	17:19	320
DL1777	10:00	15:11	EQV

解：各航班提供的飞机飞行时间并不相同，但平均约 5 小时。由于没有直接可以利用的信息，我们试图将地球的周长分解为若干个与飞行时间相关的常识性问题，以求最后归结到问题的解。因飞机在地面上空飞行与地球直径相关的信息是该航班的里程和飞机环地球一周的路程之比。因而我们将问题作如下分解：

(1) 两个城市之间的距离可以从飞机时速和飞行时间(5 小时)的乘积得到，虽然无法直接获知所需要的飞机实际飞行速度，但现在大多数的喷气飞机大多都是以近音速的，我们取估计值为 300m/s。因而飞行里程约为 5 × 3600 × 0.3 = 5400 公里

(2) 由纽约至洛杉矶所跨越 3 个时区(因为旅行信息中飞机起飞和抵达的时间会说明东部时间和西部时间以及两地时差。)

(3) 环球共有 24 个时区，因而环球飞行距离(相当于地球的周长)约为 5400×(24/3) ≈ 43000 公里。

(4) 由此可估计出地球的直径约为 $D = C/\pi \sim 43000/3 \approx 14000$ 千米。

这一结果与地球的实际平均直径 12,742 千米的相对误差约为 10%。◀

[3] 这就是哲学上的还原论

上述估算中引入了很多的近似，如飞行时间，飞行速度，飞机并非沿赤道飞行，甚至不是沿某一确定纬度飞行，飞机在高空飞行，因而环球飞行距离和地球周长并不等同，地球并非圆球，时区的确切位置等等。在估计中无须拘泥于已知相关参量的数值精度，数量级的估计最多只需要一位有效数字，更多的场合甚至只需要知道数值的一定区间范围。因此诸如 $\sqrt{2}$，π 一类的数值只要简单地记为 1 或 3，而对于不了解的数值，可以根据常识进行猜测，如上例中如果不知道两个城市之间的距离，可以从飞机时速的数量级(甚至音速)和飞行时间的乘积得到。另外，简化几何形状，地球可以视为球体，一个人也可以近似为球体，甚至立方体；以熟悉方便的尺度进行度量。

另外，基本的守恒定律，如物质守恒，流量守恒，能量守恒等都是经常可以利用的信息。

例 1-4-8 估计整个地球大气层的质量

解：这个看起来难以下手的问题可以分解为以下几个常识性的问题：

(1) 一个标准大气压约为 1×10^5 帕；

(2) 大气压强 P 源于大气受到地球的引力，重力：$F=Mg$；

(3) 大气所受到的重力于大气压强的关系为 $F=PS$，

(4) 地球的半径约为 6×10^6 米，由此可估计出地球的表面积约为 $S\sim3\times6\times6\times10^{12}\sim1.2\times10^{14}$ 米^2。

则大气层的总质量：$M=PS/g\sim10^5\times1.2\times10^{14}/10\sim10^{18}$ 公斤。◀

费米计算方法的有效性也往往令人疑惑，为什么如此粗略的数据，似乎胡乱拼凑的信息却可以导致如此的精度。另外，读者也许还会问：如果使用更多的信息，更精确的数据是否会得到更多的有效位数，并改善估计的结果。

事实上，上述事例包含了一些运气的成分，但也反映了一个基本事实：形成估算误差的许多因素对估算结果的影响并不是完全单向积累的，忽略一些已知的信息会导致高估或低估，但大多数情况下各种误差因素对最终结果的影响大部分会被相互抵消。这实际上也是构成事物本质难以直接理解的复杂性原因之一。

图 1-4-12 以示意的方式说明测算精度随着计入越来越多的努力的典型的变化情况：在开始阶段，各种假设中的误差会很快相互抵消，这使得计入稍多的因素和努力就可使估算精度得到明显的提高。在该阶段，我们也许会很幸运地得到一个误差很小的估值，但我们仍然只能认为它的精度只具有数量级上的意义，即费米方案中的数值只有一位有效数。但在计算中记入更多的细节并不总是意味着结果的改善，有时也导致精度降低。因为各种原因的误差在计算改善方面并不

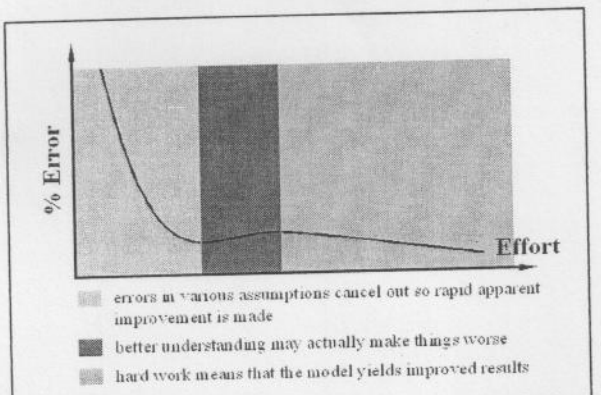


图 1-4-12 A typical curve of a calculation's accuracy versus mental effort

是按相同比例同步或齐头并进的，这就导致各种误差以同等程度减小的可能性减小，其结果就如同阶段 II 中的情况。如果我们计入了所有的误差因素，精度就会随着努力的程度如同 III 阶段那样单调减小。但这种改善的程度往往并不显著，因而在求解费米问题时，没有必要苛刻地精益求精。

费米问题的方法实际上并不只能用于估计一个物理量的数量级，将一个难以直接求解的问题分解成若干个已知的问题的思想和方法在许多方面都是行之有效的，例如在分解一个经验公式，确定其中的参量以便将公式转变为一个真正的物理定律，就需要将这个物理公式分解为若干个因式。每个因式都代表着一个已知的物理量的经验公式，我们会在以后的章节中利用一些实际的例子演示这类问题的求解过程[4]。

[4] 例如，第九章中讨论如何从唯像的普朗克辐射公式得到普朗克辐射定律，如何从原子辐射的谱线公式得到原子的内部运动状态特征。

1A 基本天文测量

1A-1 天体概述

天体是就宇宙间物质的存在形式而言的，是宇宙物质的任何集聚形成的各种天文研究对象的通称，如太阳系中的太阳、行星、小行星、卫星、彗星、流星体、行星际物质，银河系中的恒星、星团、星云、星际物质，以及河外星系、星系团、超星系团、星系际物质等。

通过射电探测手段和空间探测手段所发现的红外源、紫外源、射电源、X射线源和γ射线源也都是天体。人类发射并在太空中运行的人造卫星、空间实验室、月球、行星及行星际探测器等则称为人造天体。

宇宙空间中估计有数以万亿计的恒星，恒星因表面温度不同，发出的光的颜色也不同。依光谱特征可以对恒星进行分类，光谱相同的恒星其表面温度和物质构成均相同。恒星最小的质量大约为太阳的百分之几，最大的约有太阳的几十倍。大质量恒星含氢多，中心温度比小质量恒星高的多，其蕴藏的能量消耗比小的更快，只能存活100万年，而小质量恒星的寿命要长达一万亿年。恒星有半数以上不是单个存在的，它们往往组成大大小小的集团。其中两个在一起的叫双星，三、五成群的叫聚星，几十、几百甚至成千上万个彼此纠集成团的叫做星团，联系比较松散的叫星协。

另外现在科学界普遍认为，宇宙中还存在大量的暗物质(见4A)

1A-2 天体的可观测参量

对于与我们相距遥远的天体，人体感官所能起作用的只是视觉：来自天体(辐射或反射)的光我们可以看到该天体的亮度和颜色(频率)，借助这些光我们还能观测天体的方位，以及天体从一个方位运动到另一个方位所历经的时间。

天体方位与各种球面坐标系

在地球上仰望星空，看到的所有星星似乎都距离我们一样遥远，镶嵌在一个半径无限大的圆球球面天幕上(见图)。这个球面称为天球，地球则处在天球中心。天球本身不会移动，但由于地球的自转，我们看整个天球的视觉效果是天球围绕着地球旋转：我们看到的星星是它们在天球上的投影，它们从东方的地平线升起，到达最高点(中天)后沉向西方。

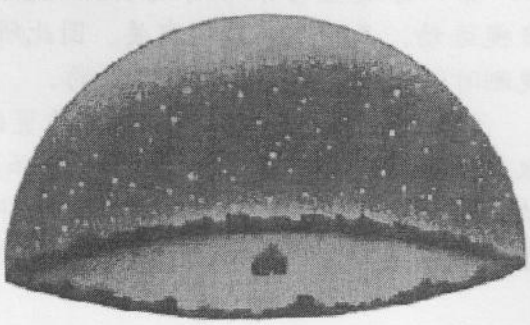


图1A-1 半个天球面

在地球上观测一个天体只能确定其相对方位，因而应该首先选择一个合适的球面坐标系，并对坐标系进行恰当的标度。通常用一个过地球中心的平面与天球面相截，截出

的大圆称为**基圈**。以与基圈平行的天球上的小圆标度第一球面坐标，这样的球面坐标系称为**天球坐标系**[1]。

由于基圈有不同的选择，相应地有着不同的天球坐标系。选择基圈的考虑主要是基于方便实际观测和结果分析。因而最常用的是赤道坐标系，黄道坐标系和地平坐标系三类

赤道坐标系

以无限延伸的地球赤道平面与天球相交截得的基圈称为**天(球)赤道**，垂直于该平面(称为**天赤面**)的地球自转轴无限延伸与天球的两个相交点称为**天极**，对应于地球北极和南极的天极分别称为**天球北极和天球南极**。再用类似于标度地球表面经纬度的方法定义坐标系的两个球面坐标，分别称为**赤纬**及**赤经**。其中，天球上平行于天赤道的圆称为**赤纬圈**，天赤道定为赤纬 0 度，天球北极和南极分别定为+90 度和–90 度。而天球上经过两个天极的大圆称为**赤经圈**。天赤道自西向东分为 24 个赤经小时[2]，一个赤经小时分为 60 分，每分再分为 60 秒。由此得到的坐标系称为**赤道坐标系**。由于坐标分别用小时和角度为单位，赤道坐标系也称为**时角坐标系**。

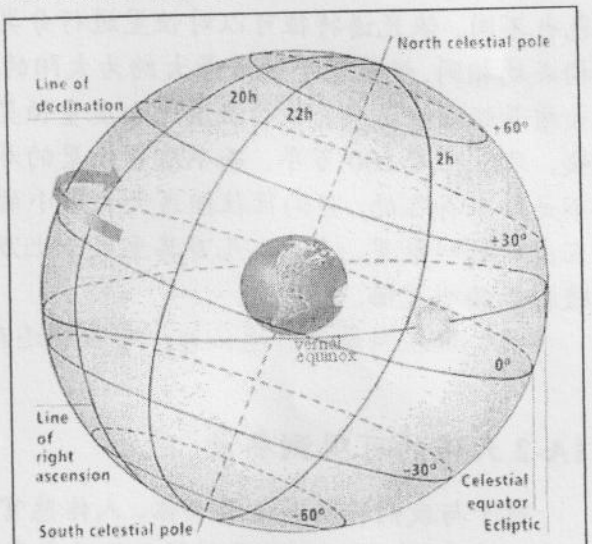


图 1A-2 天球中的赤道与黄道。

确定赤道坐标系还需要在天球(面)上选定一个零点，一般是选择一条赤经(称为**主圈**)与天赤道相交的两个点中的一个，称为**主点**。选择的主点通常也是根据具体的测量需要，不同的主点对应于不同的赤道坐标系。常用的赤道坐标系有下面两种。

第一赤道坐标系是基于观测者所在位置设置的：以地球中心为圆心，观察者正上方处的天球点为**天顶**(即观察者位置对应的天球上的坐标)，正下方处的天球点为**天底**。以过天极和天顶的天球大圆(称为**天子午圈**)为主圈，天赤道与观测者的天顶以南那段子午圈的交点为主点，向西为正，向东为负，从 0°到±180°或从 0h 到±12h 计量。第一赤道坐标系是右旋坐标系。

第一赤道坐标系中的天顶和天子午圈与观测者的地面位置有关，而因周日视运动，天顶也与时间有关。因此所观测天体的时角依赖于观测站的位置和观测时间，但天体的赤纬是不变的。

为避免天体的坐标因观测时空位置的不同而变化，可以选择一个合适的天体取代观测者本身作为参照点。例如选择太阳作为这样的天体，地球绕太阳公转的轨道平面称为**黄道面**，黄道面无限延伸与天球相交的大圆称为**黄道**，黄道和天赤道相交的两点称为**二分点**，其中黄道对于天赤道的升交点称为**春分点**，降

---

[1] 类似于地球表面上用到的地理坐标系中的基础平面是地球的赤道平面。

[2] 和地球经度(由−180 度至+180 度)的标度不同。

交点称为**秋分点**。黄道上与二分点相距 90°的两点称为**二至点**，位于天赤道以北的称为**夏至点**，位于天赤道以南的称为**冬至点**。取过春分点的赤经圈为主圈，从春分点(即主点)起沿天赤道逆时针向量到天球上一点的赤经圈与天赤道交点的弧长为赤经坐标，从 0°到 360°或从 0h 到 24h 计量。以此定义的坐标系称为**第二赤道坐标系**。第二赤道坐标系为左旋坐标系。

在第二赤道坐标系中，天体的周日运动不影响春分点与天体之间的相对位置，因此也不改变天体的赤经和赤纬。各种星表中通常列出的都是天体在第二赤道坐标系中的坐标，第二赤道坐标系也通常被直接称为赤道坐标系。

黄道坐标系

以黄道面为基本平面的天球坐标系称为**黄道坐标系**。黄道是太阳周年视运动轨迹在天球上的投影。经过地心并与黄道面垂直的直线称为**黄轴**，黄轴与天球相交的两点分别称为**北黄极**和**南黄极**，黄极与天极之间的夹角为 23°27，称为**黄赤交角**。经过南北黄极并与黄道垂直的天球大圆称为**黄经圈**，天球上与黄道平行的圆称为**黄纬圈**。坐标系以春分点为主点，过该点的黄经圈为主圈。从春分点起沿黄道逆时针方向(向夏至点)标度黄经：0°~360°。因而黄道坐标系为左旋坐标系。黄道坐标系的基圈和主圈随天球一起作周日运动，但天体的黄道坐标不会因观测时间和观测地点的不同发生变化。

地平坐标系

通常所说的**地平线**指能看到的地面与天空的分际线。天文学中的地平线指观察者所看到的地平面与天球相截形成的大圆，称为地平圈(这里忽略了过地面与过地心的两个平面在整个天球中的观测误差)。以地平圈为基圈、子午圈为主圈、子午圈与地平圈的交点(称为**南点** S)为主点的坐标系称为**地平坐标系**。

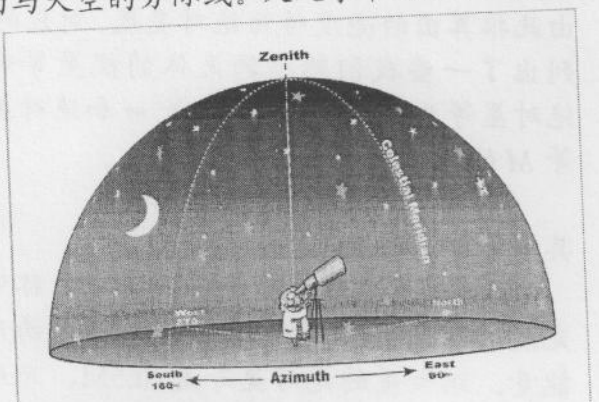


图 1A-3

天体在地平坐标系中的坐标称为**地平纬度**和**地平经度**。过天顶 Z 和天体 X 的大圆与地平圈交于 D，天体 X 的地平纬度定义为天体和观测者所在地的地平线的夹角 DOX，其中以地平线为 0°，天顶为+90°，天底为−90°。该坐标也称为(**地平**)**高度**或**仰角**。天文学更常用夹角 ZOX 表示天体 X 的第一坐标，称为**天顶距**。其中以天顶为 0°，地平线是 90°，天底为 180°。地平经度定义为南点 S 与 D 间的大圆弧所对的角 SOD，也称为或(天文)**方位角**。天文学中习惯从南点起沿顺时针方向量度。对不同的观测者，由于铅垂线方向的不同，就有不同的地平坐标系，同一天体也就有不同的地平坐标。而天体的地平坐标因周日视运动也在不断发生变化。

天体辐射的光谱型，光度及标度

天文学中用光谱型和亮度来说明不同恒星辐射的波段和强度这两种特性。

由于人眼能够观测到的辐射只限于波长为 400nm~700nm 的可见光区域，天文学中的光谱型也就是主要颜色红、黄、蓝的分区。借助于不同的传感器，则可以探测射电、红外、紫外、X 射线及 γ 射线等几乎全部波段的电磁辐射。

天文学上用星等来量度星体的亮度，这一概念最早是由希腊天文学家依巴谷(Hipparchus)于公元前 2 世纪创立的。当时他为辨析在天蝎座中发现的一颗陌生星体而想到除了位置，还可利用天体的亮度来标示天体。为此他绘制一份标有 1000 多颗恒星精确位置和亮度的恒星星图。为描述恒星的亮度，喜帕恰斯将恒星亮暗分成等级：将看起来最亮的 20 颗恒星定为一等星，将眼睛看到最暗弱的恒星定为六等星。其间再分出二等星、三等星、四等星和五等星。这一“星等”概念由此一直沿用到今天。1850 年，英国天文学家普森(M.R.Pogson)利用光度计对上述一等星到六等星的亮度比较中发现，星等相差 5 等的一等星和六等星的亮度之比约为 100 倍。由此他将两个相邻星等的亮度比值规定为 100 的五次根：$100^{1/5} \approx 2.512$ 倍，即一等星的亮度是二等星的 2.51 倍，三等星的 $2.51^2 \approx 6.31$ 倍，…，由此引入了零星等，负星等以衡量更亮的天体，同时，星等也用小数表示。

地球上的直接观测到的天体亮度称为天体的**视亮度**，相应的星等称为**视星等**，记为 m。它们与天体和观测者之间的距离有关。例如一颗亮星可能因为相距遥远而看上去很暗(视星等数值大)；而一颗暗星可能因为距离近而显得很亮(视星等数值小)。为比较不同天体的真实亮度，天文学中设定了一个标准距离：10 秒差距(即 32.62 光年)，即将测得的星体的视星等折算到这个标准距离，由此推算出的亮度称为**绝对亮度**，对应的星等称为**绝对星等**，记为 M。表 1A-1 列出了一些我们熟悉的天体的视星等和绝对星等为 4.75M。目视星等 m 和绝对星等 M 的转换公式

$$M = m+5-5\lg r$$

其中 r 为距离。

表 1A-1 明亮星体光度表

星体	目视星等	绝对星等
太阳	–26.70	4.75
月球(满月)	–13.00	不适用
金星(最亮时)	–4.6	不适用
天狼星	–1.450	1.43
织女星	0.03	0.50
天蝎座	3.8	–9.4

描述恒星实际亮度的标准物理量称为**光度**，描述的是恒星单位时间内辐射的光能量。织女星的绝对星等是 0.5M，光度是太阳的 50 倍。绝对星等为–9.4M 的天蝎座，光度约为是太阳的 50 万倍。最强恒星的光度约为太阳的 100 万倍。二天狼星的伴星光度不到太阳的万分之一。

光度的概念并不涉及辐射的波段，具体测量时对电磁波谱中单一波长或其周围很窄波段测定的光度称为**单色星等**，测量略宽一点的频段范围所得的星等称为**窄频带星等**；而测量频段范围更宽的光度所得的星等称为**宽频带星等**；表征天体在整个电磁波段内辐射总量的星等称为**热星等**。

另外，由于探测器对不同频率的辐射由不同的响应，在一定波段内测定的星等与选用的探测器有关。因而对应不同探测器有各种星等系统。

(1) **目视星等**是人眼测定的星等。因人眼对黄色最敏感，目视星等也称黄星等。哈佛大学天文台以小熊座 λ 星的 6.55m 确定目视星等的零点。由此推得太阳的目视星等为–26.74m。目视星等 1m 的星对地面的照度约为 8.3×10^{-9} 勒克司。

(2) **照相星等**是用蓝敏照相底片测定的星等。

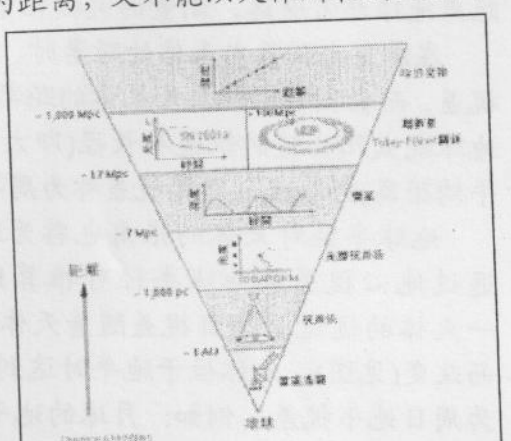
(3) **仿视星等**是用正色底片加黄色滤光片测定的。其分光特性与人眼相近，实际上取代了目视星等。

(4) **光电星等**：用光电倍增管测定的星等。目前最常用的光电星等系统是 *UBV* 测光系统。通过在检测系统前放置标准滤光片实现三个波段的辐射测量，其中 *U*(波长 360 nm 左右)为紫外星等，*B*(波长 440nm 左右)为蓝星等，*V*(波长 550nm 左右)为黄星等(和目视星等相似)。天文学中，不特别说明的星等一般是可见星等。

1A-3 各种天象观测中的信息

天体测量学是天文学中最先发展起来的一个分支，天文学的另一个分支—天体力学也是基于天体测量学发展起来的。早期的观测注意到恒星在天空的方位相对不动，天文学家据此绘制出星图，划分星座和编制星表；进而研究太阳、月亮及行星的运动。今天的观测手段已从可见光观测扩展到射电，红外、紫外、X 射线和 γ 射线等波段；观测到数量更多的、亮度更暗的恒星、星系，测量精度也随着各种精密测量仪器的应用大幅提高，观测精度正在走向千分之一角秒。观测位置已由固定天文台发展为流动站、全球性组网观测以及空间观测。尽管如此，可以直接测量天文参量非常有限，大多数参量的确定都是基于由观测中归纳出来的经验公式的推算。其中有些公式可以通过普适的物理原理演绎出来，有些则不清楚其物理本原：现有的模型推算出的结果都与观测相差很大。

天文中最大的问题是不能直接测量天体的距离，又不能以天体所在的空间对天体进行实验，因而无法通过基本物理学原理确定天体及其运动的位置，质量，(视向)速度这三个最基本的物理量。本节主要介绍与几何学，运动学，动力学，物体辐射相关的一些物理学规律，以及根据观测中归纳出来的经验公式间接地推算出天体距离的方法。例如其中最基础的经验关系和基本物理定律是开普勒根据行星绕日运动的观测数据归纳出的行星运动三大定律，以及牛顿的万有引力定律则奠定了天体力学的基础。事实上天体物理学与天体测量是密不可分，相辅相成的。


银河系内测距

天文学中将天体按距离分为多个等级，并以特定的天体来设定不同距离的指标。用作远距离指标的天体有：新星、发射星云、行星状星云、球状星团、I 型超新星、星系…其中，行星状星云与 I 型超新星的亮度范围明确，亮度高，是较成功的远距离指标。不同类型的天体远近相差悬殊，测量的方法也各不相同。远距离指标经常需要利用近距离指标来校准，不准确也逐渐累积，因此距离愈远的天体，测量的不确定愈高。各种量距离方式的关联性，可以用倒

立的"距离金字塔图"来表示。

雷达测距

20 世纪 40 年代出现的雷达天文学是射电天文学的一个分支，雷达向天体发射电磁波，并接收天体反射的回波。通过提取中回波信号携载的信息研究天体的物理性质和几何结构。例如天体与地球之间的距离可通过测量来自天体表面的反射波时间间隔推算出来。需要说明的是这种看似非常简明的测量也包含着基于假设的物理学原理：光是沿欧几里德直线传播的[3]。

由于雷达发射功率的限制(回波信号强度与距离的四次方成反比)，探测距离局限于太阳系以内。目前主要用于太阳系内天体精确测距和定位等。雷达技术目前已从无线电波段扩展到光波段的激光雷达，是天文观测中精度最高的一种手段。

三角视差

三角视差法是基于欧几里德几何学原理发展出来的测距方法。设想观测者在两个不同位置 B 和 C 观测位于 A 处的物体(如图)，测量出$\angle\beta$、$\angle\gamma$和 B，C 间的距离(称作基线)，就完全确定了三角形 ABC 的全部几何特性。若将待测天体位置视为上述三角形的顶点 A，并在空间另选两个位置(一般是同一时间地球上的两个不同位置或利用地球本身的运动选择不同时刻地球上的同一个位置)作为该三角形的另外两个顶点 B 和 C，通过测量 B 和 C 对 A 的视角和 B，C 间的距离可以推算出天体与地球的间距。在不同位置 B 和 C 观测同一天体的方向差异称为**天体视差**，可用 B，C 间的距离(基线)在天体处的张角表示。因为天体通常距离地球非常遥远，测量的关键在于选择足够长的基线及精确测量角度。

在测定太阳系内天体的距离时，通常以地球半径为基线，所得视差称为**周日视差**。而在测量太阳系外星体的距离时，一般以地球绕太阳公转的轨道半长径(即太阳和地球的平均距离)为基线，所得视差称为**周年视差**。

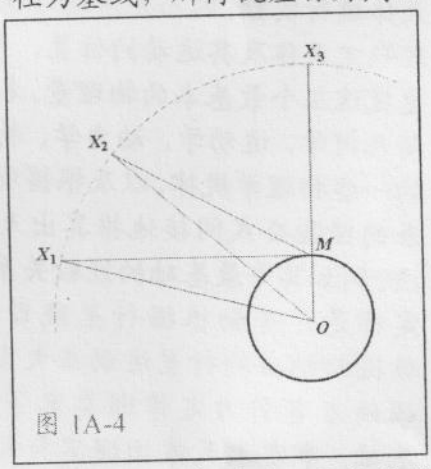


图 1A-4

地球半径对天体的张角也称为**地心视差**，通过地心视差和地球半径可推算地球距离某一天体的远近。周日视差随着天体的高度变化而改变(见图)，天体位于地平时达到极大值，称为**周日地平视差**。例如：月球的地平视差 57′，再由地球半径 6371 km 可推算出月-地距离约为 384 400 km。

因地球是个扁球体，赤道半径大于极半径，当观测者位于赤道时，天体的周日地平视差具有最大值，称为**赤道地平视差**。

地球至太阳的平均距离是天文测量中的最基本单位，称为一个**天文单位**，记为 A.U.，通常就是用太阳的地心视差表示的。但太阳-地球距离远大于地球半径，直接测量太阳的地心视差误差很大，实际测定太阳的周日地平视差往往

3

采用间接的方法，一般是通过测量太阳系内行星的地心视差，再根据天体力学中的开普勒第三定律推算出太阳的地心视差。进行这种观测的最佳时机是利用太阳系的行星(或小行星)最接近地球(如金星凌日(图 1A-3)[4]、火星冲、小行星冲)等天象，测量行星的周日赤道地平视差。1716 年，哈雷建议利用 1761 年和 1769 年金星凌日测量太阳视差。为此天文学家在世界各地组织了不少联合观测。通过对观测资料的综合分析，法国天文学家潘格雷(1711-1796)于 1775 年公布太阳视差为 8"8。这一结果为国际天文界接受、并一直沿用到 1967 年。现代天文学借助空间探测器，雷达和激光对近距离行星进行的直接距离测量推算出金星与地球的最近距离约 0.28A.U.。在金星最接近地球时，通过测量金星表面的雷达回波时间间隔推算出 1A.U.≈149,597,870 公里(误差小于一公里)。1964 年国际天文学联合会天文常数系统中，太阳视差定义为导出常数：arcsin(a e/A) = 879405，其中 A 是光经过一个天文单位距离的时间，在 1976 年的天文常数系统中，该数值取为 8794148。

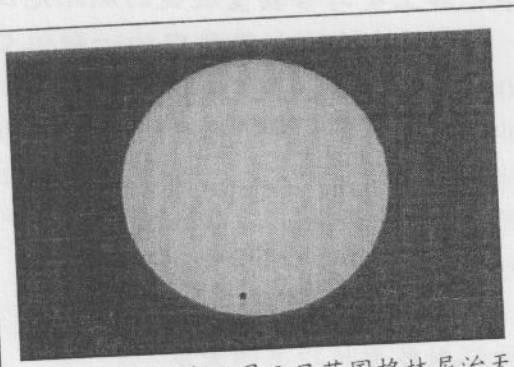


图 1A-3 2004 年 6 月 8 日英国格林尼治天文台于当地时间 5 时 19 分拍摄的"金星凌日"照片。

"天文单位"作为一个导出单位，其测量属性基于的物理定律对某些领域就可能形成困扰。例如按引力定律，"天文单位"被视为一个由视角测量值换算成距离的基本常数：

一质量无限小的颗粒以角频率 0.01720209895/天(公转周期 365.2568983 日，即一高斯年)环绕太阳作牛顿圆轨道转动的半径"。

这一定义方式中最重要的缺陷在于和爱因斯坦的广义相对论相矛盾。广义相对论下的时空是相对的，它取决于观察者的位置。这意味着天文单位也会随观测者的位置而发生变化。德国天文学家赛吉尔·克里诺尔(Sergei Klioner)对比在地球的参考系和木星的参考系中的计算，发现这一数值相差 1000 米甚至更多。这种差异不会对探测器的飞行造成影响，因为它们是直接进行距离测量的，但对构建太阳系模型的行星科学家们造成了困扰。该定义的另外一个重大问题在于太阳系高斯常数与太阳质量有关，因而天文单位也和太阳质量相关联。由于随着时间推移，太阳不断损失质量，这也将造成天文单位数值的缓慢变化。

为解决这些问题，2012 年的国际天文学大会(IAU)决定采用新的定义

1 AU ≡ 1495,9787,0700 米。

其中"米"的单位则是由光在真空中 1/2,9979,2458 分之一秒内通过的距离定义的，这样，上述数值不再受太阳质量变化的影响，另外，由于光速在所有参考系中相同，天文单位的定义值也不再随观察者在太阳系中位置的不同而发生改变。

当远处的恒星与地球的连线垂直地球轨道半径时，恒星对日地平均距离 α 所张的角称为恒星的周年视差。恒星的距离 $r = a/\sin\theta$，由于恒星的周年视差通常小于 1°，$r \approx a/\theta$。由于作为一个天文单位的日-地距离 a 相对于我们与其它 横行的距离很小，天文学中常用 1 角秒(1°的 1/3600)的恒星周年视差作为距离单位，称为秒差距(pc)。1 秒差距约为 206265A.U.。当恒星距离地球约几

[4] 金星轨道在地球轨道内侧，在一些特殊时刻，地球、金星、太阳处于一条直线上，这时从地球上可以看到金星就像一个小黑点一样在太阳表面缓慢移动，天文学称之为"金星凌日"。

百秒差距时，其周年视差的观测误差与视差本身的数值相当，但地面观测受大气视宁度的限制，因而周年视差的适用范围小于 100 秒差距，而地球大气层外的依巴谷卫星与哈勃望远镜的测距范围可延伸到 1000 pc。

贝塞耳选择了天鹅座中的一颗恒星，测出该星半年间在天空的角位移约 1/5000 度。由此推算出这颗恒星与我们相距 94.6 万亿千米。亨德森选择了天空中非常明亮的半人马座 α 星，亨德森判断，如此明亮是因为距离近。他测出其距离地球仅 4.3 光年，是太阳系外离我们最近的恒星。斯特鲁维选择的研究对象是天琴座中的织女星，距离地球约 26 光年，由于视差很小，使当年斯特鲁维的观测结果不太准确。

天体辐射参量与距离的关系

对于太阳系外的天体主要利用其辐射特性测定距离。物理学中物体本身的发光光度 L_o、观测到的亮度 S 以及观测者与物体间的距离 r 满足关系

$$S \propto L_o/r^2$$

在天体测量中分别用视星等(m)和绝对星等(M)来描述 S 与 L_o，它们与距离 r 有关的响应关系即为距离模数关系

$$5\lg r = m - M + 5$$

利用该方法确定距离的关键是确定天体的光度 L_o 或绝对星等。其主要途径大多基于恒星的**赫罗图**。

19 世纪中叶，意大利教授赛奇将恒星按照发射光谱分成白星、黄星、橙红星、深红星 4 类，认识到这样的分类和恒星温度有关；英国的哈斯根通过分析这些恒星的化学组成，指出亮星具有和太阳相同的化学组成。下层炽热物的辐射穿过大气层而向外发射。决定恒星发射的连续光谱特征的物理要素是恒星的温度。实验发现，热辐射体的光度及光谱类型都与其温度有关：其在单位时间内由单位面积所释放的热辐射能量与温度四次方成正比，因而恒星的表面积越大，恒星的光度也越大。如果将恒星看成热辐射体，则只要测得恒星的绝对星等和光谱类型，就能估计其温度和体积。

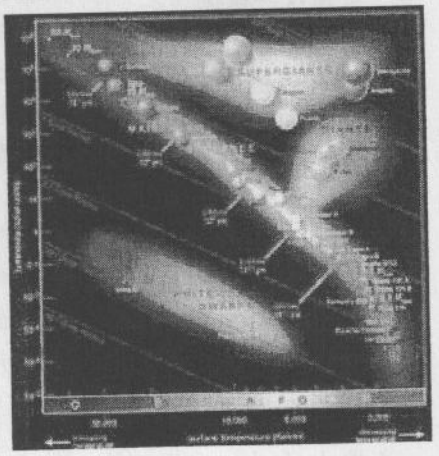


图 1A1-1 赫罗图

1911 年丹麦天文学家赫兹伯仑测定了几个银河星团中恒星的光度和颜色，并将它们标示在以这两个参量为纵横坐标的坐标系中。1913 年美国天文学家罗素通过分析诸多恒星的光度和光谱得到一系列恒星光度和光谱型之间的关系图。由于光谱型与颜色等价，两种分布图所反映的实质都是恒星的发光强度与恒星的表面温度的关系[5]，这一类恒星光谱型-光度关系图被统称

[5] 恒星是一团炽热的，被自身引力束缚的气体，其中心区域密度和温度都特别高，足以产生热核反应。恒星表面的高温使之发射类似黑体辐射一样的光谱。在很宽的频率范围内都有辐射，因此称为连续谱。恒星的光谱型和表面温度之间存在着对应的关系，因而从恒星的颜色可以判断它们的温度。表面温度在绝对温度 30000K 以上的恒星发蓝光，

为赫罗图(见图 1A-1)。

恒星在赫罗图上主要集中在四个区域:

(1) 约 90%的恒星密集分布于图的左上方至右下方大致沿对角线的带中，天文学中称其为**主星序**，带上的恒星称为**主序星**。太阳、牛郎、织女等都是主序星。主星序表明，大多数表面温度高的恒星光度也大，表面温度低的恒星光度也随之减小，这反映出恒星演化的一种规律性。

其余的星形成小群。分别处于

(2) 主星序右上方：该区域的特点是光度大，表面温度却不高(呈红色)，这表明该区域中的恒星体积巨大，因而称为**红巨星**。红巨星上面的恒星点光度更大，因而对应的体积也更大，称为**超巨星**。如北极星(小熊座 α)、大角(牧夫座 α)属于巨星，心宿二(天蝎座 α)为超巨星。巨星的直径是太阳的几十到几百倍，如红超巨星参宿四的直径是太阳的 900 倍。

(3) 主星序左下方：该区域中恒星的特点是光度小，表面温度却很高(呈蓝白色)，这表明该区域中的恒星体积很小，因而称为**白矮星**。如宁静新星和行星状星云的核是已公认的白矮星。天狼 B(天狼星的伴星)是最亮的白矮星。白矮星的直径只有几千千米，中子星的直径则只有 20 千米。需要说明的是，恒星的体积相差非常大，它们的质量却差别不太大。大多数恒星的质量在太阳质量的 0.5~5 倍之间。质量最大的恒星，其质量能比太阳大几十倍。质量最小的恒星，其质量也有太阳质量的几十分之一。

(4) 图中很右边的区域对应于温度非常低的区域，这里通常是低温星际云的位置，由星际云形成的原恒星也在图的右边。

赫罗图是由恒星的光学观测数据构成的，因此中子星和黑洞不能在赫罗图上显现。最初的赫罗图中没有第三和第四个区域，因为那时还没有发现白矮星，也没有讨论恒星的形成。

赫罗图对认识和理解恒星的演化非常重要。特别是随着测量方法的改进及测量精度的提高，赫罗图的光谱型逐渐被色指数替代，色指数再转换为恒星表面温度；而观测得到的恒星视星等则经过距离修正后表示为绝对星等，再转换为光度。由于恒星内部能源的不断消耗，光度和温度都随之变化，由此，恒星的演化就通过恒星在赫罗图上的位置变化反映出来，给出恒星从诞生到主序星、红巨星、变星、新星(超新星)、致密星(白矮星或中子星或黑洞)的演化机制和模型：位于图中极右区域的分子云因引力收缩而温度上升，慢慢移向主序带。恒星临终时会离开主序带往右上方移动进入红巨星及红超巨星区域，经过红巨星但未发生超新星爆炸的恒星会越过主序带移向左下方进入白矮星区域，接着因能量损失渐渐变暗成为黑矮星。

如果认为所有的主序星都具有相同的性质，相同光谱型的恒星就有相同的绝对星等。对照太阳附近恒星的赫罗图可以求出遥远恒星的绝对星等，这种方法称为叠法。

温度在 10000~30000K 的恒星颜色蓝白，温度在 7500~10000K 的恒星颜色纯白，6000~7500K 的恒星呈黄白色，温度在 5000~6000K 时，恒星的颜色发黄，温度在 3500~5000K 时恒星的颜色为红橙，温度在 2000~3500K 的恒星颜色发红。

由恒星的光谱分类与光度分类推算出恒星的绝对星等也是估算恒星距离的有效方法。该方法中视星等很容易测量，而通过对恒星光谱的分析可以得出该恒星的绝对星等。因而该方法称为**光谱视差法**。

因为主序星的分布较集中在带状区域，光谱视差法常以主序星为标的。利用邻近的恒星，校准光谱视差法的测量。另也假设远处的恒星的组成与各项性质，大致与邻近恒星类似。误差常在 25% 以上。(注：银河系直径约 30 Kpc)

例如，测得某恒星的视星等为+15 ，其光谱判定为 G2 V 的恒星‘i，从赫罗图读出该星的绝对星等为+5，由距离模数公式得 $r = 1000\ \text{pc} = 3260$ 光年。

测距适用范围： ~ 7Mpc。

造父变星

位于不稳定带的后主序带恒星，其亮度有周期性的变化，因而称为变星。20 世纪初，美国天文学家李维特(H. Lieavitt)发现其中一类变星的光变周期越长，其光度就越大，具有这种“周光关系”的变星称为**造父变星**(Cepheid Variable)[6]，其周期从几天到几个月不等。

造父变星所具有的周光关系提供了一种简便而相对精确的测量恒星距离方法：通过测量造父变星的光变周期获知其绝对星等，再结合测量到的视星等即可推算出距离。造父变星因而被天文学家们称为“标准烛光”。

造父变星有两种：经典造父变星和室女座 W 型造父变星(I 型与 II 型)，它们有不同的周光关系。天琴座的 RR 型变星也具有特定的周光关系，也可用以测定距离。这种使用变星测距的方法大致可以测量 108 秒差距的恒星。目前发现的最远造父变星在 M100，距离我们约 17 Mpc。

测距适用范围： ~ 17 Mpc。

如果能在一个遥远的星系中找到一颗造父变星，就能大致判断星系的距离。甚至判断宇宙的膨胀速度。

超新星

大部份超新星(I 型与 II 型)的最大亮度很相近，天文学家常假设它们一样，以用作大距离的指标。

以造父变星校准超新星的距离分别找出 I 型与 II 型星的平均最大亮度。由超新星的光度曲线决定其归类。对新发现的超新星，把最大视亮度(mV)与理论最大绝对亮度代入距离模数公式，即可找出超新星的距离。

II 型超新星受外层物质的干扰，平均亮度的不确定性较高，I 型超新星较适合做为距离指标。

测距适用范围： > 1000 Mpc。

---

[6] 造父变星的亮、暗变化周期源于其大气温度产生有规律的冷热变化导致整个星体有规律的膨胀或收缩。

与天体速度由关的经验规律

多普勒展宽

由展宽程度推算出谱线位移量 $\Delta\lambda$，

Tulley-Fisher 关系

漩涡星系因星系自转，其中的的氢 21 公分线形成多普勒展宽 $\Delta\lambda$ 揭示出星系的漩涡臂在视线方向的速度 V_r

$\Delta\lambda/\lambda_o = v_r/c$

Tulley 与 Fisher 发现漩涡星系的光度与自转速率成正比

称为 Tulley-Fisher 关系。

由测量漩涡星系的旋转速率求得漩涡星系的光度，再用距离模数公式就可推知漩涡星系的距离。该方法适用的距离测量范围大致与 I 型超新星同级，可互为对照。现常观测红外线区谱线，以避免吸收。

测距适用范围：> 100 Mpc。

哈勃定律

1929 年，美国天文学家哈勃(Edwin Hubble)对河外星系的视向速度与距离的关系进行了研究。当时只有 46 个河外星系的视向速度可以利用，而其中仅有 24 个有推算出的距离，哈勃得出了视向速度 V_r 与距离 d 之间大致的线性正比关系。现代精确观测已证实这种线性正比关系

$V_r = H \cdot d$

称为哈勃定律。其中 $H = 100h_0\ \mathrm{km \cdot s^{-1}Mpc}(0 < h_0 < 1)$为比例常数，称为哈勃常数。利用该定律可通过测量星系的退行速度推算星系的距离。例如，室女群(Vigro cluster)的径向远离速度 v_r=1180 km/s，推算出与地球的距离 $r = v_r/H$ = 1180/70 = 16.8 Mpc。

哈勃定律方法是目前所知道的最大的量天尺，其尺度可以达到 100 亿光年的数量级。并可以测量到这个可观测宇宙的边缘了。这种距离测量的精度取决于确定哈勃常数 H，事实上，哈勃太空望远镜的一项主要任务就是确定哈勃常数。

恒星的质量一般是依据观测其运动状态，结合牛顿的引力定律推算的。一个实用的方法是找到一个双星系统，通过分析两个星体间的引力所决定的轨道运动决定它们的质量[7]。通过对许多距离我们较近的双星的分析，在赫罗图上标示出这些星体的质量后发现，主序列上的恒星排列也呈现出按质量大小分布的规律性：左上方的高温高亮度恒星质量比较大，而右下方的低温低亮度恒星质量较小。

[7] 双星系统是很常见的。许多天上肉眼可见的星星都是双星，甚至是多聚星系统。

2-3 运动的相对性及伽利略相对性原理

我们测量和描述运动的方式决定了运动的概念是相对的而不是绝对的，即描述物体的运动需要另一个物体作为参照物。这导致表述"物体 A 相对于物体 B 以速度 **u** 运动"与"物体 B 相对于物体 A 以速度–**u** 运动"完全等价，不同的只是两种表述中分别将物体 B 或 A 作为参考系。A 或 B 对同一个现象具有不同，但对称等价的表述形式，这对许多初学者的直觉或日常经验而言似乎平淡无奇，但它所反映的正是运动相对性的本质，对其中内涵的认识程度极为深刻地影响着我们对时间和空间的认识，并由此决定我们将如何理解和描述物质的运动。

2-3-1 惯性参考系

我们首先对用于描述运动的参考系做一个简单的分类和相应的说明。

如果物体 A 在运动中速度不发生变化，物体 A 的运动称为**惯性运动**，以物体 A 为参照物设置的参考系称为**惯性参考系**(简称**惯性系**)。如果物体 A 在运动过程中速度发生变化，物体 A 的运动称为**非惯性运动**，以物体 A 为参照物设置的参考系称为**非惯性参考系**。

显然，判别上述物体 A 的运动速度是否发生变化同样需要先明确参照物体运动速度的变化情况，由此就会陷入一个无限的循环当中。事实上，惯性参考系的定义涉及参照物体所受到的其它物体的作用力，因而在运动学范畴中无法给出惯性系的准确定义。但这里先作一点说明，当一个物体没有受到其周围物质的影响(相互作用)，或者说当这个物体只是一个孤立系统时，其运动称为惯性运动，以该物体为参照物设置的参考系就是惯性系。读者显然会意识到，宇宙中任何物体都不是孤立的，因而并不存在理想的惯性系，惯性系只是描述物体惯性运动以及物体偏离惯性运动程度的理想模型。但就特定的测量而言，自然界中存在着许多很好的近似惯性系。下一章会详细讨论物体间相互作用以及惯性系的概念。

2-3-2 伽利略相对性原理

由于对物质世界中各种物理事件的观测最终总能归结到对运动的描述，上述关于运动相对性的表述和认识可以总结为力学中的一条基本原理—**相对性原理**。相对性原理在不同的场合可以表述为不同的形式，而其中关于运动学的表述是最直观，也是最简单的。在伽利略时代，人们对运动的感知和理解都局限于机械运动的范畴，而且所观测到的运动物体的速度都非常低(远小于 299,792,458 米/秒[1])。实际观测中的局限性深刻地影响到了伽利略对时空的认识，并由此形成了所谓的**伽利略相对性原理**[2]：

所有的惯性参考系都是等价的。处于惯性运动状态的物体 A 或者 B 都可以以另一方作为参照物来说明自己是否处于运动状态，但任何一方都不能推断其

---

[1] 这个速度就是1-2节提到的光在真空中的传播速度。

[2] 相对性原理在物理学中属于不变性原理一类。不变性原理在近代物理学中起着非常重要的作用，我们会在后面的章节中作一些事例的介绍。

中哪一个处于绝对的运动状态。

该原理的一种特殊表述为：**不可能在一个惯性参考系中通过任何一种力学实验测量出参考系自身的运动速度**。例如，对处在一个匀速运动的封闭车厢里的观察者而言，他不可能通过投掷硬币，做自由落体实验，观测单摆运动或其它任何机械运动实验测量出车厢的运动速度。唯一能够让他得知车厢运动状态的方式是观测车厢外的物体(如树木，建筑等)相对于他自身有没有相对运动。

伽利略对相对性原理的表述是建立在定量观测的基础上的，其中的一个很重要的观测结果是速度的叠加：设想有两个惯性参考系 S 和 S'，其中 S' 以速度 $\mathbf{u}$ 相对于 S 运动。如果一个坐在参考系 S' 中的观测者看到一个物体以速度 $\mathbf{v}'$ 运动，而参考系 S 中的观测者看到该物体的运动速度为 $\mathbf{v}$，伽利略将两个惯性系中观察到的物体运动速度之间的关系总结为一个经验公式：

$$\mathbf{v} = \mathbf{u} + \mathbf{v}' \tag{2-3-1}$$

称为**伽利略速度变换方程**。一个常见的例子是一部自动扶梯中的传送带以速度 $\mathbf{u}$ 相对于地面运动，一个乘客在扶梯的传送带上以速度 $\mathbf{v}'$ 运动，则地面上的观测者看到该乘客的运动速度 $\mathbf{v}$ 满足经验公式(2-3-1)。

下面我们试图通过分析揭示出该速度变换方程中蕴涵的时空特征，即通过对经验公式进行变形(数学变换)使其展示出其内在的物理本质。

从观测的角度而言，两个观测者所能测量的只是物体相对于自身所在的参考系中的物理参量。设在某一时刻，两个观测者在各自所处的惯性系中测得的物体位置分别为 $\mathbf{r}$ 和 $\mathbf{r}'$，由图 2-3-1 中的几何关系可以得到位置矢量 $\mathbf{r}$ 和 $\mathbf{r}'$ 之间的变换方程：

$$\mathbf{r} = \mathbf{R} + \mathbf{r}' \tag{2-3-2}$$

其中 $\mathbf{R}$ 为惯性系 S' 的原点相对于惯性系 S 的位置矢量。如果对上述变换方程两边计算相对于时间的微商：

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} = \frac{d\mathbf{R}}{dt} + \frac{d\mathbf{r}'}{dt} \tag{2-3-3}$$

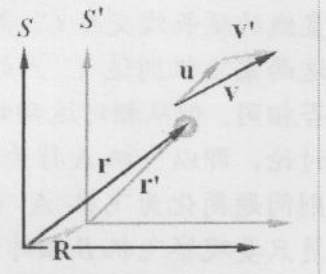


图 2-3-1 惯性系之间的位移和速度变换。

形式上，这似乎就是伽利略的速度变换方程(2-3-1)，但从物理量的测量属性分析，上述推理中含有一些不明确的假设。首先，图 2-3-1 只是一个(瞬间)静态图像，图像本身并没有表现出两个参考系之间的相对运动。而位置矢量 $\mathbf{r}$ 和 $\mathbf{r}'$ 是分别在 S 系和 S' 系中测量的，方程中只有两个参考系原点之间的相对位置矢量 $\mathbf{R}$ 既可以在 S 也可以在 S' 系中测量。或者说，实际测量是 S' 系中的观测者测量出 $\mathbf{r}'$，而 S 系中的观测者测量出 $\mathbf{r}$ 和 $\mathbf{R}$，比较三个测量值发现它们在实验精度范围内满足式(2-3-2)。因而式(2-3-2)隐含有假设：在两个不同的惯性系 S 和 S' 中观测到的两个空间点间的距离相等：

$$\Delta\mathbf{r} = \Delta\mathbf{r}' \tag{2-3-4}$$

这一关系式所反映的就是伽利略关于"**空间是绝对的**"概念的实质所在：

对空间间隔的测量结果与观察者所处的惯性系无关。[3]

另一个类似的假设隐含在方程(2-3-3)中：从速度的操作性定义而言，$\mathbf{v}$ 和 $\mathbf{v}'$ 是分别在 S 和 S' 中各自的时间间隔中 Δt 和 $\Delta t'$ 测量 $\Delta\mathbf{r}$ 和 $\Delta\mathbf{r}'$ 得到的。位移关于

[3] 即与两个惯性参照系之间的相对运动速度无关。

时间的微商应表示为

$$\mathbf{v}=\frac{d\mathbf{r}}{dt}=\frac{d\mathbf{R}}{dt}+\frac{d\mathbf{r}'}{dt}=\frac{d\mathbf{R}}{dt}+\frac{d\mathbf{r}'}{dt'}\frac{dt'}{dt}=\mathbf{u}+\mathbf{v}'\frac{dt'}{dt}$$

伽利略的速度变化关系(2-3-1)成立要求 $dt/dt'=1$，或者

$$\Delta t=\Delta t' \tag{2-3-5}$$

这一关系所反映的是伽利略关于"**时间是绝对的**"观念：

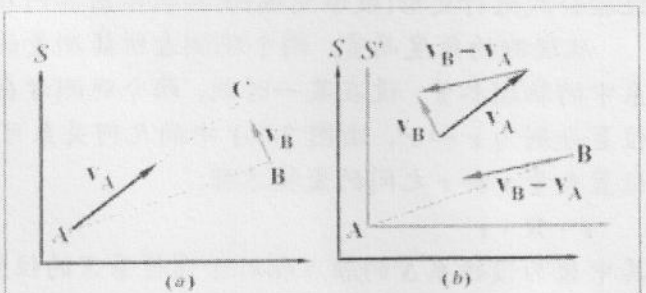
对时间间隔的测量结果与观察者所处的惯性参考系无关。

尽管伽利略关于时间和空间的概念中存在一些不明确的因素，但他的速度方程在力学领域和仅涉及低速运动物体的场合与实验观测符合得相当好。

2-3-3 伽利略相对性原理的应用

利用相对运动的概念可以将一个物体的复杂运动转化或分解为若干个简单运动之和，从而给计算和描述运动带来很多方便。

例 2-3-1 摄像两架飞机在某一时刻分别从位置 A 和 B 以速度 $\mathbf{v}_A$ 和 $\mathbf{v}_B$ 飞行(见图 *a*)，分析它们会不会在某一时刻在某一地点发生碰撞。

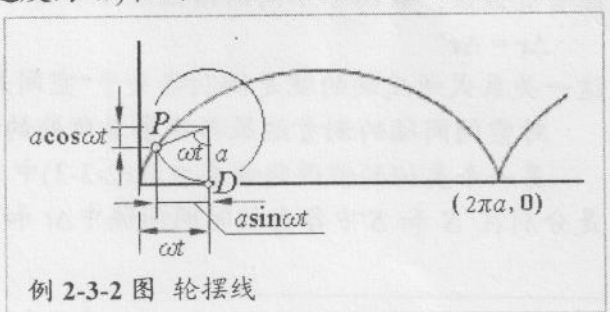


例 2-3-1 图 对运动的不同观察视角。

解：如果以地面为参考系进行计算，首先要确定这两架飞机的可能撞击地点：沿 $\mathbf{v}_A$ 和 $\mathbf{v}_B$ 的两条直线的延长线交点 *C*。然后确定这两架飞机到达 *C* 点的时间是否相同。但从相对运动的视角来讨论，即以飞机 A 作为惯性系，则问题简化为飞机 A 中的飞行员只要观察飞机 *B* 相对于飞机 A 的速度 $\mathbf{v}_B-\mathbf{v}_A$ 的方向是否与飞机 *B* 在飞机 *A* 为原点的系中的位置矢量共线(即飞机 *B* 的飞行方向是否正对着 A，见图 b)。由于相对速度与联线 AB 不在同一方向，A 和 B 不会发生碰撞。

例 2-3-2 一个半径为 *a* 的圆盘沿 **x**-轴在地面上作纯滚动(见图)，滚动角速度为 ω。求圆盘边缘上的一点 *P* 在圆盘滚动过程中绘出的空间轨迹方程，并求 *P* 点相对与地面的速度和加速度。设 $t=0$ 时刻，*P* 点处在坐标系的原点位置。

解：直接求 *P* 点的运动轨迹显然比较复杂。而仔细观察可以发现，*P* 点的运动是 *P* 点绕圆盘中心的匀速圆周运动(角速度为 ω)和圆心沿 **x**-轴方向的匀速直线运动(速度为 $u=a\omega$)的合成。若以地面为惯性参考系 ***S***，圆盘为惯性参考系 ***S'*** (以圆盘中心 *O'*为参考系原点，以圆盘中心的运动所勾勒出的直线为 ***S'***参考系的横轴 *x'*)，两者的相对运动速度为 $u=a\omega$，记 $\theta=\omega t$ 为 *P* 点在 ***S'***参考系中的位置矢量在



例 2-3-2 图 轮摆线

时间 t 中旋转过的角度，则 P 点的坐标可以表示为 P 点相对于 O' 点坐标($x'=-a\sin\theta$, $y'=-a\cos\theta$)和 O' 点相对于 O 点的坐标($x'_o=a\theta$, $y'_o=a$)之和

$$x=x'_o+x'=a\theta-a\sin\theta$$

$$y=y'_o+y'=a-a\cos\theta$$

这是 P 点在 S' 参考系中运动轨迹的参数方程，该轨迹称为**轮摆线**。

P 点相对于 S 参考系的速度为

$$v_x=\frac{dx}{dt}=a\frac{d\theta}{dt}-a\cos\theta\frac{d\theta}{dt}=a\omega(1-\cos\omega t)$$

$$v_y=\frac{dy}{dt}=a\omega\sin\omega t$$

速度大小为

$$v=\sqrt{v_x^2+v_y^2}=a\omega\sqrt{(1-\cos\omega t)^2+\sin^2\omega t}=\left(2a\sin\frac{\omega t}{2}\right)\omega=\overline{PD}\omega$$

上式反映了圆盘作纯滚动时的一个重要特征：圆盘与地面的接触点 D 在所讨论的瞬间的运动速度为零，而圆盘边缘的其它点在该瞬时的运动恰好为相对于点 D 以角速度 ω 的转动。因此，接触点 D 通常被称为圆盘纯滚动时的瞬时转动轴。这也说明圆盘沿 **x**-轴在地面上作纯滚动可以分解为一系列绕瞬时轴的转动。

P 点相对于 S 参考系的加速度为

$$a_x=\frac{dv_x}{dt}=R\omega^2\sin\omega t,\qquad a_y=\frac{dv_y}{dt}=R\omega^2\cos\omega t$$

加速度的大小 $a=\sqrt{a_x^2+a_y^2}=R\omega^2$。这也说明在不同惯性系中观察到的加速度是相同的。◀

例 2-3-3 设一个物体在一个以角速度 ω 顺时针匀速转动的圆台上沿转动圆台的一条直径以速度 v' 匀速运动。求物体运动到距圆台中心 r 的位置时，相对于地面的速度和加速度。

解： 选择极坐标。相对于地面而言，物体运动到距圆台中心 r 处的速度为

$$\mathbf{v}=v'\hat{\mathbf{r}}-r\omega\hat{\boldsymbol{\tau}}$$

在 Δt 时刻物体运动到$(r+\Delta r,\theta+\Delta\theta)$处的速度径向分量增量为

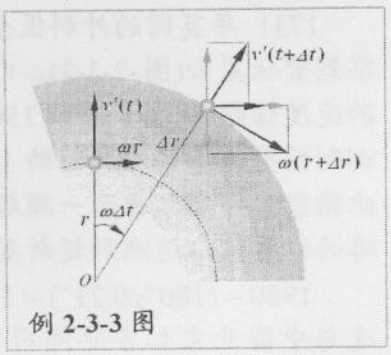
$$\Delta v_r=[v'\cos(\omega\Delta t)-\omega(r+v'\Delta t)\sin(\omega\Delta t)]-v'$$

$$\approx[v'-\omega(r+v'\Delta t)\omega\Delta t]-v'\approx-\omega^2r\Delta t$$

横向速度的变化为

$$\Delta v_\theta=[v'\sin(\omega\Delta t)+\omega(r+v'\Delta t)\cos(\omega\Delta t)]-\omega r'$$

$$\approx[v'\omega\Delta t+\omega(r+v'\Delta t)]-\omega r'\approx 2\omega v'\Delta t$$



例 2-3-3 图

由此得到物体相对于地面的径向和横向加速度分量分别为

$$a_r=\lim_{\Delta t\to0}\frac{\Delta v_r}{\Delta t}=-\omega^2r\qquad\text{和}\qquad a_\theta=\lim_{\Delta t\to0}\frac{\Delta v_\theta}{\Delta t}=2\omega v'$$ ◀

2-3-4 伽利略相对性原理面临的问题

伽利略相对性原理是描述力学领域中物体低速运动的实验定律。由于在关于时间和空间的概念中存在着一些不确定的因素，如果将这个原理应用到伽利略未曾考察过的高速运动的现象，特别是与真空中光的传播速度向关联的电磁现象[4]时，就可能出现理论预言和实验结果的冲突。下面我们通过一个具体的事例来揭示其中的问题。

设想两个球员在进行棒球的投、接球练习(图 2-3-2)。两人相距 l，投手掷出的球的速度为 u。按照伽利略的速度变换方程，接球手看到来自速度为 u 的球的光的速度等于球的速度与光相对于球的速度 c 之和：$v=c+u$。因而接球手应在 $t=l/(c+u)$时刻开始看到来自投掷过来的球的光，比较接球手看到的投手掷球动作的时间 $t=l/c$ 要早一些。

图 2-3-2 接球手先看到投手的掷球动作还是先看到投过来球？

许多读者一定会对这一结果感到诧异，因为如果上述分析成立，基于伽利略速度变换方程的相对性原理可能导致违反因果关系秩序的现象：即结果可以发生在原因之前，从而导致我们认识整个物质世界所基于的内在逻辑关系的崩溃。因此在进一步学习物理学之前，必须先通过实验观测确认这一现象是否真的存在，因果关系是否可能逆转。为此我们先估计上述现象中因果两个事件的时间间隔。设 l 为 15 米，一个专业投手的投出的球速可达每秒 30 米，则相应的时间间隔为

$$\Delta t=\frac{l}{c}-\frac{l}{c+u}=\frac{l}{c}\frac{u}{c+u}\approx l\frac{u}{c^2}=5\times10^{-15}(\mathrm{s})$$

如此小的时间间隔实际上难以观察到，但我们可以通过改变实验中的相应参数 l 和 u 放大该时间间隔。一个很有说服力的著名事例是对距离我们 5000 光年的蟹状星云的观测。

1731 年英国的外科医生同时也是天文学爱好者的贝维斯(John Bevis)首次观察到蟹状星云(图 2-3-3)，以后的持续观察发现这个“大螃蟹”在以每年 0.21 弧秒的速度膨胀(相当于每秒 1500 公里)。1920 年，观测到的星云半径达到了约 180 弧秒。天文学家由此猜想它可能起源于一颗超新星的爆发。通过粗略的计算可以追溯到超新星爆发[5]的大致年份为

$$1920-(180''/0.21'')\approx1920-680=1060$$

这与中国北宋仁宗年间司天监记录的相应天象(强烈闪光)观测时间(1054 年)基本相符。

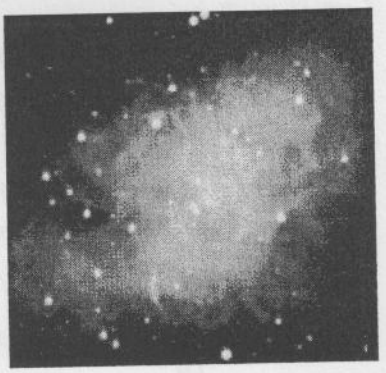


图 2-3-3 蟹状星云

对照按伽利略速度变换公式估计在地球上观测这个爆炸所发出的闪光可以持续的时间，可以检验伽利略相对性原理在涉及光现象时是否正

[4] 由于光被认为是电磁波，与光有关的现象一般被认为属于电磁学领域

[5] 蟹状星云是超新星爆发形成的气态外壳，爆发后的遗迹现在成了星云中心的中子星。

确。对于各向同性的爆炸，碎片的平均速度 u 在各个方向大致相同，这样，在地球上观测飞向地球和背离地球运动的碎片向地球发出的光的速度应分别为 $c+u$ 和 $c-u$。其中爆炸碎片的运动速度约为 1500(km/s)，一个瞬间爆发事件出现的闪光在地球上可以持续观测到的时间为

$$\Delta t = \frac{l}{c-u} - \frac{l}{c+u} = 2\frac{u}{c^2-u^2}l \approx 2\frac{u}{c^2}l \approx 43 \text{ (years)}$$

而根据中国司天监的天象记录，观测到的强烈闪光持续时间还不到两年[6]。远小于伽利略相对性原理给出的预言。如果考虑到超新星爆发本身所持续的可能的时间，实验观测的结果似乎意味着飞向地球和背离地球运动的碎片向地球发出的光的速度在地球上观测几乎是相同的。如果在地面上观察到的光的传播速度与发射光的物体的运动速度无关，则前面投、接球的实验中出现的因果时序反转也就不存在了。

上述事例和分析表明。伽利略变换中所隐含的有关绝对时空的假设实际上并不准确，它在涉及高速运动，特别是速度接近光在真空中的传播速度时会表现出比较显著的偏差。下一节将讨论如何正确地表述相对性原理，及如何更新相关的时间和空间的概念。

思考题

2-3-1 伽利略相对性原理可表述为："不可能在惯性系内部进行任何物理实验来确定该系统作匀速直线运动的速度。"但汽车司机通过车内的车速表就能知道汽车的速度。这是否违背伽利略相对性原理？

2-3-2 从地球上看几乎所有的星体都在远离我们而去，按照现在流行的观点，这是宇宙爆炸引起不断扩张的结果。由于地球并非宇宙中心，如何理解大爆炸的图像？根据大爆炸图像能否说明目前各个天体的总体运动状态？

习题

2-3-1 两辆车沿同一条直线行驶，其中一辆具有恒定的加速度 1 $m\cdot s^{-2}$，另一辆则具有恒定的减速度 2 $m\cdot s^{-2}$。两辆车在 t_o 时刻分别以速率 1 $m\cdot s^{-1}$ 和 16 $m\cdot s^{-1}$ 经过某一位置，分别用解析和几何方法确定它们将在何时何地再次相聚。

2-3-2 设 A 城位于 B 城正西 1200 km，C 城位于 A 城正北 1200 km。一架飞机在由东向西的大风中以速度 600 km/h 飞行，风速为 400 km/h。计算航程 *ABA* 及 *ACA* 所需要的飞行时间。

2-3-3 爆竹在上空炸开后，碎片向四周各方向以相同的速度飞散，求碎片在下落过程中形成的外包络的形状。

2-3-4 设想在一无风的雨天，你要从一个地方走到另一个地方。为尽量减小淋到的雨水数量，应用尽可能快或尽可能慢的速度，还是某一中等的速度行走？设雨点的下落速度为常数。

2-3-5 一质点在惯性系 S 中作匀速圆周运动，轨迹方程为 $x^2+y^2=a^2$，$z=0$，求

6

解该质点在惯性系 S'中观测到的轨迹方程，其中惯性系 S'相对于 S 系以速度 v 沿 x 方向运动。

2-3-6 一个半径为 a 的圆盘沿 x-轴在地面上作纯滚动，滚动角速度为 ω(如例 2-3-1)。求圆盘上距离圆盘中心 O'为 $r\,(\le a)$的一点 P 在圆盘滚动过程中绘出的空间轨迹方程，并求 P 点相对与地面的速度和加速度。设 $t=0$ 时刻，P 点处在 O' 点正下方，且 P 与 O'点的连线过地面坐标系的原点位置。

2-3-7 地球半径为 6.4×10^6m，地球在以轨道半径为 1.5×10^{11}m 绕太阳公转的同时也在绕自身的轴自转，而地球的自转轴与黄轴间有 23°26′的夹角，由于地球自转的惯性作用，外力矩并不能使地轴与黄道轴的夹角发生变化，而是使地轴绕黄道轴由东向西进动(同地球自转和公转方向相反)，进动周期为 25,800 年，相应的进动速度为 50.2564 角秒。(1) What is the speed with respect to the Sun of the point on Earth's equator nearest the Sun? (2) Of the point on Earth's equator farthest from the Sun? (c) Of the two points on Earth's equator, midway between the points in parts (1) and (2)? (3-69)

2-3-8 A softball player hits the ball when it is 0.9 m above home plate. The ball leaves the bat at an elevation angle of 30°. What initial speed must the ball have to clear a fence 1.8 m high located 65 m away in left field? A softball player hits the ball when it is 0.9 m above home plate. The ball leaves the bat at an elevation angle of 30°. What initial speed must the ball have to clear a fence 1.8 m high located 65 m away in left field? (3-72)

2-3-9 一质点从一匀速转动的圆盘边缘上的一点沿圆盘直径以匀速度 v 运动到圆盘的另一边缘点。设圆盘的转动角速度为 ω。求解质点在圆盘上留下的运动轨迹。

2-3-10 设参考系S'相对于惯性参考系S以角速度$\boldsymbol{\omega}$旋转。证明矢量$\mathbf{G}$的时间变化率在两个参考系之间的变换关系满足

$$\frac{d\mathbf{G}}{dt}=\frac{dG}{dt}\hat{\mathbf{G}}+\boldsymbol{\omega}\times\mathbf{G}$$

其中dG/dt为$\mathbf{G}$的大小的时间变化率。

2-4 爱因斯坦相对性原理

对运动物体发射光的观测结果与伽利略变换所给出的结论之间的差异涉及到光的本质及光的传播特性。我们必须说，有关光的这些问题直到目前仍然没有完全解决。但仅仅从光的发射和传播的物理图像而言，光可以理解为是一种交替感应并在空间扩散的电磁场。根据有关电磁感应的实验定律和由此总结出的电磁场理论，空间任何一处的磁场变化会在其周围感应出变化的电场，而在空间任何一处的电场变化则会在其周围感应出变化的磁场，这样的电磁交替感应并不断向四周扩散的交变电磁场称为**电磁波**[1]。而光只是整个电磁波谱中一个特殊频率波段中的电磁波：人的视觉神经细胞只对该频率区域的电磁波产生响应[2]。交替感应的电磁场在向四周扩展时并不要求周围必须存在任何实际物质作为传播介质，因此，电磁波和我们所熟悉的声波、水波、固体弹性波等机械波有一个重要的差异：机械波只能在介质中传播，而电磁波可以在真空中传播[3]，而且在真空中传播的速度高达 2.99792458×10^8 米/秒，远大于空气中声波的传播速度。

2-4-1 爱因斯坦[4]相对性原理

电磁波能够在真空中传播，这一事实引出一个与运动相对性原理有关的问题：机械波(如声波)的传播速度都是相对于介质而言的，或者说以介质作为参考系。那么光在一无所有的真空中传播时，其速度是相对于什么**实物参考系**而言的呢？对这一问题所能想到的可能答案只能是，真空中观测到的光的传播速度是相对于观测者自身的[5]，这类似于物体 A 在空无一物的环境中测量 B 的运动速度时只能以自己作为参考系一样。但这两个情形就伽利略相对性原理而言实际上并不相

---

[1] 我们假定读者在中学中已经大概知道法拉第的电磁感应现象，以及光的本质是电磁波这样的概念。如果读者没有这样的背景知识，可以先接受这些结论，第13章中会对这些相关的问题做比较仔细的讨论。

[2] 不同生物对电磁波的视觉反应并不相同，因此，他们对光的波长范围的定义也和我们不同。

[3] 有一个很简单的实验可以说明光不同于声波，可以在真空中传播：将一个振铃放在一玻璃容器内，我们可以隔着玻璃看见铃的振动，同时听到铃声。如果抽去容器内的空气，我们仍然可以看见铃的振动，但不再能听到铃声。这说明声波的传播依赖于介质(这个例子中是空气)，但光可以在真空中传播

[4] 阿尔伯特·爱因斯坦(Albert Einstein, 1879-1955)美国德裔物理学家。因为对光电效应的解释获得1921年的诺贝尔物理学奖。但他更为世界所了解的工作是狭义相对论和广义相对论。

[5] 相对于观测者本身应该理解为相对于观测者所代表的参考系。对这个问题曾经有过两种解释。其一是认为光在真空中的传播速度是相对于光源而言的，这相当于我们前面对观察投掷棒球，以及超新星爆发的讨论。但这与我们对波的观测结果不一致，因为介质中传播的波(实际是介质的一种运动形式)的速度仅依赖于介质的性质，而与波源的运动状态无关。例如，一艘船在行驶过程中激发的水面波的传播速度并不受到船本身的速度的影响。第二种解释是设想真空中存在一种我们至今还不了解的神秘介质，暂且称之为以太。以太除了承载光的传播，对其他物体及其行为不产生任何影响。根据我们在前一章中对波速度所做的量纲分析，光速应与以太的弹性模量的平方根呈正比，这意味着能使光在其中高速传播的介质必然使非常坚硬的。如果整个宇宙充满着的坚硬的物质却不能被任何方法所测量到，这样的假设显然不是很有说服力的。但人们还是为判断这种物质是否真实存在设计出了一种实验验证的方法：通过假设以太会或不会被运动物体所带动来观察光在运动以太或静止以太中传播所应出现的不同效应，但实验结果发现无论假设以太能被带动或不能被带动，都没有观测到响应的效应。第8章中我们将会从波动的特性讨论这些相关的实验，这里我们仅关注于惯性参考系和时空特性这一主题。

同：对物体 A 或 B 各自的惯性参照系而言，通过互相观察只能说明两者的相对运动，不能指认出其中任何一个处于绝对运动状态，而对光波而言，由于真空中不允许存在局域的电磁振荡，相互交替感应所激发的电磁场在空间的扩展就是绝对的，或者说光的运动是绝对的[6]：自然界中找不到一个能使光静止于其中的惯性系。这样，电磁场理论就与伽利略相对性原理产生了矛盾。两个理论在各自的领域中都与实验观察符合，而它们在解释同一种现象时则出现了冲突[7]。

就电磁场理论和相对性原理所涉及的具体对象而言，前者描述的是电磁场的具体行为，后者则是一个关于惯性参考系等价性的基本原理。矛盾出现在相对性原理对"惯性参考系等价性"判断的表述上。伽利略是用惯性系之间的速度变换方程来表述这种等价性的，而该方程事实上只是从低速机械运动实验中归纳出来的，能否适用于全部可能的速度范围并不具有实验基础，而它所赖以立足的逻辑基础—绝对的时空观念在很大程度上被认为是不言自明的公设。而电磁场理论中基于电磁场交替感应而扩散的电磁波在真空中传播则是一个实验规律。由此可见，解决这个冲突决非在两个理论中选择其中一个那样简单。新的相对性原理如果是一个普适的基本原理，就必须将所有物理规律(包括电磁感应)纳入其中，同时还必须在速度远低于光速的力学运动的范围能退化为伽利略速度变换方程的表述。

为了使新的相对性原理能够包容电磁场理论，爱因斯坦最终放弃用伽利略的速度变换方程作为相对性原理的出发点，或者说放弃了"所有的速度都必须是相对的"这一约束。爱因斯坦的相对性原理表述为：

所有的惯性参照系都是等价的。不可能在某一个惯性参照系中进行任何一种物理实验以测量出该惯性参照系自身的运动速度。

由于新的原理将所有可能的实验现象，包括光传播在内的电磁现象纳入其中，这也就包含了真空中光的运动和运动速度是绝对的这一概念：

光在真空中的运动是绝对的，其传播速度 c 与观测者所处的惯性参照系无关。

真空中的光速与惯性参考系无关通常被列为爱因斯坦相对论的两条假设之一，但实际上应该认为是麦克斯韦的电磁场理论的结果，因而也是爱因斯坦相对性原理所要求的结果。设想一个观察者能够在真空中测得光相对于他所在的惯性系的传播速度 v 与 c 有所不同，该观察者据此就能推知自己是处于某种绝对运动的状态，并可以根据他所测得的光速 v 和 c 的差异推算出他所处的惯性系的绝对速度。这就违反了爱因斯坦相对性原理中关于"不可能在某个惯性参照系中进行任何一种物理实验以测量出该惯性参照系自身的运动速度"的表述。因而在任意惯性系中测得的光在真空中的速度都只能是同样的 c。

---

[6] 爱因斯坦曾回忆少年时代一个假想实验所引起的困惑：如果他能以和光同样的速度与光并行，那么当他张开手掌，伽利略相对性原理就要求光会成为一个站在在他手掌之上的电磁振荡。但这个现象不符合电磁感应定律，因为只要出现电磁振荡，电磁感应就会使该振荡的电磁场向四周传播。即实际上不可能存在一个惯性参考系，在其中能观测到一个静止不动的光。

[7] 事实上，在科学的发展过程中，出现这样的情况还是时有发生的。其原因在于我们对问题的认识和理解一般都会受到我们观察相关现象的视角的影响。开始阶段我们讨论的出发点通常是具体的现象，如果我们试图描述两个并没有呈现出任何关联的自然现象，对这两个现象的描述视角和所建立的相应模型通常是相互独立的。但以后我们往往会将所建立的模型和理论推广，应用到更广的范畴，如果两个独立发展出来的理论或模型被应用发生交汇的领域，就面临着这两个独立的，原来看起来风马牛不相及的理论是否能对同一种现象给出和谐一致的解释。如果这两个理论给出不同的，甚至冲突的结论，这就是通常所谓的佯谬。

爱因斯坦的相对性原理要求真空中的光速是绝对的这一概念和伽利略相对性原理中所有物体的运动速度都是相对的概念完全不同。这种差异必然会影响我们原来绝对的空间和时间观念，以及与空间和时间相关的物理量性质的认识。下面分别讨论与此相关的三个最重要时空概念：“同时性”，“时间间隔”和“空间距离”，以及它们与惯性系运动速度之间的关系。

2-4-2 同时的相对性

“昨天晚上 10 点钟你在什么地方?”

如果你受到警方这样的询问，而你所居住的社区昨晚发生了谋杀事件。对你来说最急迫的是提供一个能够证明你当时不在现场的证人。只要他的证词得到认可，你也就摆脱了相关的嫌疑。你的证人可以是当时与你坐在一起聊天的朋友，或者只是驾车从你家旁边经过看到你坐在自己卧室的窗户旁边的邻居。警方并不会因为你的邻居和你处于不同的惯性系而拒绝他的证词，因为就我们的日常经验而言，同一时刻发生在不同地点的两个事件(有人被谋杀和你坐在家中喝茶)的“同时性”被理所当然地认为是与观察者所处的惯性参照系无关的。这种绝对的“同时性”观念正是伽利略相对性原理的基础。但在爱因斯坦相对性原理的框架中，这种“同时性”却出现了问题。

为了让读者对爱因斯坦相对性原理对绝对“同时性”的影响有一个更为直观的认识，我们考虑另外一个例子。设想在匀速行进中的火车的一节车厢中有一场为了却私人恩怨的决斗。决斗仲裁人让两个决斗者分别站在车厢的两端，而在车厢的正中间设置一盏灯。决斗者看到灯亮就可以拔枪射击。对车厢内的人而言，决斗的公平性(即光信号能分别同时抵达车厢两端两个决斗者的位置)是通过两路光信号所传播的距离相同以及光信号传播的速度相同这两个事实得以保证的。但这种公平性对站在火车外的观察者而言却完全是另外一回事。因为他所看到的是，面向火车头的决斗者由于向着光源运动，光信号从光源到达他的实际位置的距离因而被缩短了，由于光的速度是绝对的，这意味着他能较早地接受到光信号。而背对火车头的决斗者是背离光源运动的，光信号从光源到达他的实际位置的距离拉长了，因而接受到光信号的时间就会晚一些。

由于两个不同参照系中观测者的论证都不存在纰漏，两个观测者对于两个时间是否同时的结论就他们自身所处的惯性系而言都是正确的。这一结果说明，在不同地点发生的两个事件的同时性并不是绝对的，而是和观测者所处的惯性参照系有关。这种与我们的日常经验不相符合的“同时性”概念的相对性显然就是爱因斯坦相对性原理所要求的真空中光速不变的结果(读者可以自己证明，如果采用伽利略的速度变换公式，“同时性”就是绝对的)。

2-4-3 时间间隔的相对性—时间膨胀效应

物理学中，对时间的测量操作是对一个均匀的周期运动过程中的周期进行记数，相应的实验装置(称为时钟)主要由一个振荡器和一个记数器构成。为了理解时钟的运动是否会对时间的流逝快慢产生影响(从观测的角度而言，时钟的运动

是否改变其“滴答”的节律)，我们设想用一个内部构造最简洁的“**光钟**”做一个(理想)实验，这个特殊的时钟的振荡器由上、下两块反射镜和这两个镜面之间来回反射的一个光脉冲(见图 2-4-1)构成。光脉冲每次到达上镜面时触发记数器进行记数。这种装置的简单内在结构可以免于我们考虑各种实际计时装置中的特殊构造所可能带来的复杂性，从而提供了一个可以直接考察运动对时间间隔影响的简单模型。

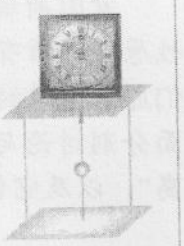


图2-4-1 光钟

设惯性系 S'以匀速度 u 相对于另一个惯性系 S 向右运动，我们考察分别处于这两个惯性系中的观测者测量 S'系中时钟的计数情况。

我们具体讨论光脉冲在下镜面相继发生两次反射的时间间隔(振荡周期)。对于 S'系中的观测者而言，光钟内光脉冲两次在同一镜面反射是发生在同一空间位置的两个事件(见图 2-4-2-*a*)。发生在同一空间位置的两个事件的时间间隔称为原时[8]。但对处于 S 系中的观测者而言，光脉冲的路径是图 2-4-2-*b* 所示的两个全等直角三角形的两条斜边之和，因而光脉冲完成一次来回反射相对于在静止的钟内要走更长的距离。记 S' 系和 S 系中观测到的两个事件的时间间隔分别为 $2t'$和 $2t$，则运动时钟在时间间隔 t 内的运动距离为 ut，而光脉冲从底部到达顶部的距离为 ct'，则由勾股定理可以得到[9]

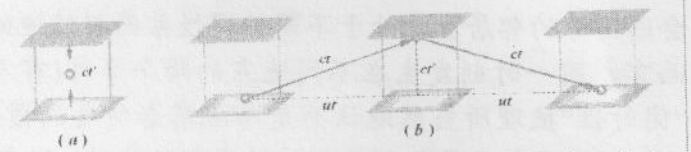


图 2-4-2 在 S'系和 S 系中观测，光钟内光脉冲两次在下镜面反射这两个事件中光脉冲所经历的空间距离和时间间隔不同。

$$(ct)^2 = (ct')^2 + (ut)^2$$

由此可以得到

$$t = \frac{t'}{\sqrt{1-u^2/c^2}} = \frac{t'}{\sqrt{1-\beta_u^2}} = \gamma_u t' \tag{2-4-1}$$

其中 $\beta_u = u/c$ 称为**约化速度**，$\gamma_u = (1-\beta_u^2)^{-1/2} > 1$ 称为**洛仑兹因子**。方程(2-4-1)表明时间原时 t' 短于惯性系 S 中观测到的相应时间间隔 t。即在 S 系观测到 S'系中的时间间隔变长了，或者说运动时钟比静止时钟走得慢。时间膨胀效应的根源显然也源于真空中光速的恒等性，它使得在运动时钟内的光脉冲需要较长的时间才能完成一次振荡，表现为运动时钟的节律变慢。

为了得到时间膨胀效应大小的数量级概念，设想一个钟以接近空气中的声速 $v_s \approx 300$ 米/秒的速度运动，则其“滴答”节律减缓的比例为 5×10^{-13}，相当于 760 多万年走慢 1 秒。这样小的时间膨胀效应在日常生活中是完全微不足道的。也就是说对于远低于光速的运动而言，伽利略的绝对时间概念是和实验结果吻合的。要观察到比较明显的时间膨胀效应，时钟的运动速度就必须达到光速的一定的比例，例如，要观察到 1%的节律减缓的比例，时钟的运动速度就必须达到 $u \approx 0.14c$。

有些读者也许会怀疑这种时间膨胀效应是否和“光钟”的独特设计有关，或者

[8] 例如，你所带的手表给出的读数表示的就是你自身所经历的两个事件的原时间隔。.
[9] 由光速不变原理得到的图2-15b中的几何图形表明，实际物体的运动速度不可能达到真空中的光速的数值，这是因为直角三角形的斜边ct总是大于任意一条直角边ut。

认为自然界也许会存在某种计时装置能够“滴答”出绝对的时间间隔。但很容易证明：爱因斯坦的相对性原理是不允许这种可能性存在的。因为如果存在这样一个能够“滴答”上帝时间的时钟，我们可以将它和“光钟”放在一起。这两个时钟相对于我们静止时，它们的“滴答”速率一致，而它们一旦相对于我们运动时，它们的读数就会出现差异，这样，与这两个时钟处于同一个惯性系的观测者可以利用这种差异来判断其自身的运动状态甚至计算出自身的运动速度。由此可知，我们观察到运动惯性系中的各种自然过程的节奏放慢是一个普遍的现象。这种自然过程可以是摆钟内单摆的振荡周期，也可以是生物体内的某种生命节律。例如地球上的人看来，相对地球做高速运动的宇宙飞船中的宇航员的生命节律会放慢，因而会比我们活得更长。这种效应在 1937 年得到了实验的直接验证，但实验对象不是一个人或某种动物，而是一种称为 μ 子的短寿命微观粒子。

μ 子是寿命仅为 $\tau = 2.197\times10^{-6}$ 秒的放射性介子[10]。μ 子一旦产生，就会很快地衰变为一个电子和两个中微子：

$$\mu \to e^- + \nu_\mu + \tilde{\nu}_e$$

外空中的宇宙射线与地球大气层顶部的分子碰撞会产生大量的高能量 μ 子，其中部分 μ 子能穿越厚度约 25 千米的大气层而被地面的接受装置所探测到。由此产生的问题是，这些 μ 子即使以光速运动，在仅有的 2 微秒左右的寿命内也只能行进 $l = ct \approx 660$ 米的距离，为什么它们能活着穿越整个大气层到达地面？该问题的答案要点在于，μ 子的寿命是由其体内的时钟控制的，即寿命是 μ 子自身所处的惯性系中的时间间隔。对地面上的观测者而言，μ 子体内的时钟会因为它的运动而走慢，因此他所观测到的 μ 子所能行进的距离也相应延长了。μ 子的实际行进距离和它相对与地面的运动速度有关[11]。

例 2-4-1 要在地面上探测到大气层顶部产生的μ子，这些μ子至少需要多大的运动速度？

解： 设μ子相对与地面的运动速度为u时，恰好能在临终前到达地面，即有

$$h = u\Delta t = u\gamma_u\tau$$

其中h为大气层厚度，τ为μ子的寿命，可以解得μ子的速度和真空中光速的比值为

$$\frac{u}{c} = \sqrt{\frac{h^2}{h^2 + c^2\tau^2}} = \sqrt{\frac{(2.5\times10^4)^2}{(2.5\times10^4)^2 + (3.0\times10^8\times2.2\times10^{-6})^2}} = 0.9997$$

μ子的运动速度并没有超越真空中的光速。而如果采用伽利略的速度公式，这些μ子的运动速度至少需要达到$u = (25\text{ km})/(2.2\mu\text{s}) = 38c$才能到达地面。◀

---

[10] 放射性粒子的寿命是一种平均的概念，$N(t) = N_o e^{-t/\tau}$，μ子的寿命$t = 2.197\ \mu$s是通过测量实验室中产生低速μ子的自发衰变半衰期确定的。

[11] 这类实验也可以在实验室中完成，其要点是选择具有与实验条件适配的实验对象。例如在实验室中测得π介子在速率$v = 0.913c$时寿命为6.37×10^{-8}s。由此算得的时间膨胀因子$\gamma = 0.408$，π介子的固有寿命$6.37\times10^{-8}\times0.408 = 2.60\times10^{-8}$s。这个时间尺度就很适合实验测量。首先是它在接近光速运动时能飞行10~20m的路程，可以在实验室中完成测量，第二是它长到足够使$\pi^{\pm}$介子在碰撞产生后能在衰变之前停下来，使我们能测量其固有寿命(测量结果为2.60×10^{-8}s，计算与测量结果一致)

2-4-4 空间间隔的相对性—空间收缩效应

至此，有些读者也许会对μ子在地面参考系中可以通过延年益寿而运动更长的距离的现象感到一些困惑："μ子相对于自身的参考系并不能延长其寿命，那μ子自己如何解释它能够在2个微秒内到达地面的事实呢？"

对μ子而言，这样一次长途跋涉只涉及三个物理量：光速(上限速度)，时间(自身寿命)以及运动距离(大气层厚度)。由于前两个量是确定的，解决该问题的唯一可能性就在于对μ子而言，大气层厚度在μ子运动方向应该有同等比例γ_u的缩小。

但自然界是否真是如此运行的呢？

在不同惯性系中测量沿运动方向的米尺长度

为了理解一个物体的运动状态是否影响它的观测长度，以及是如何影响它的观测长度的，首先需要从实验操作性角度解决应该如何测量一杆匀速运动的米尺的长度这个问题。

设固定在尺上的惯性系S'相对于另一惯性系S以速度u向右匀速运动，为了在两个不同惯性系中测量该米尺的长度，我们在惯性系S中的某一位置P设置一标杆，则S'系中的观测者将看到该标杆以速度u向左匀速运动，相继经过米尺的右和左两端B和A (见图2-4-3a)，这两个事件相隔的时间间隔为

$$\Delta t' = L'/u \tag{2-4-3}$$

其中L'是尺在其自身所处的惯性系S'中被测量到的长度，称为尺的**固有长度**。而对坐在标杆旁的观测者(S系)而言，尺以速度u向右运动(见图2-4-3b)，尺的两端B和A相继经过位置P所用的时间为

$$\Delta t = L/u \tag{2-4-4}$$

其中L是在惯性系S中测量到的尺的长度。由于B和A相继经过位置P这两个事件在惯性系S中是在同一地点发生的，两个事件的时间间隔Δt为时间原时，时间间隔Δt和$\Delta t'$满足关系

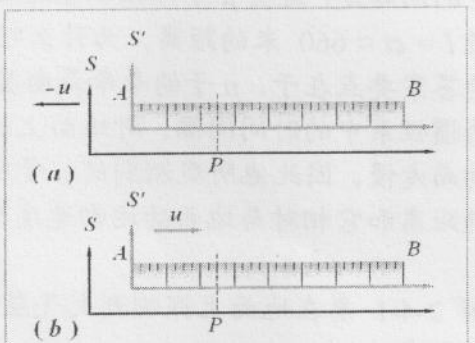


图2-4-3 在不同惯性系中对尺的长度测量

$$\Delta t = \Delta t'/\gamma_u \tag{2-4-5}$$

由此可以得到尺的固有长度L'和观测长度L之间的关系：

$$L = L'/\gamma_u = \sqrt{1-\frac{u^2}{c^2}}L' < L' \tag{2-4-6}$$

即测量到的运动直尺的长度短于该直尺的固有长度。需要说明的是，测量到的运动直尺的缩短与该直尺的性质无关，它反映的是空间本身的性质。空间由于相对运动而导致运动方向上出现比例为γ_u的观测尺度收缩，称为**洛仑兹收缩**。这一效应也就帮助μ子解释了它能够在2个微秒内到达地面这一事实。

日常生活中所能看到的物体的速度都远小于光速，所产生的洛仑兹收缩很小，几乎不能被现有的仪器所探测的。例如，一杆米尺以每秒30米的速度运动，它所产生的洛仑兹收缩仅约5.0×10^{-13}米，还不到一个氢原子大小的百分之一。

运动的相对性要求洛仑兹收缩效应也是相对的。即惯性系S中的观测者看到惯性系S中所有物体在其运动方向上的尺度发生了收缩，而S中的观测者则看到惯性系S中所有物体在其运动方向上的尺度发生了收缩，由于两个不同参照系的物体无法直接比较，因而两个看似不同的结论实际并不矛盾。这两个不同观测者的结论都是正确的。图2-4-4为俄裔美国物理学家伽莫夫写的科幻小说《汤普金斯先生奇遇记》[12]中的一个场景：主人公汤普金斯先生来到一个极限速度很低的城市：真空中的光速仅为30 km/h。汤普金斯先生站在街上看到骑车人因速度由于接近当地的自然极限速度而发生了明显的洛仑兹收缩，但沿街的房屋都没有形变。而当他自己骑车前行时，站在路边的行人及沿街的房屋都发生了收缩。

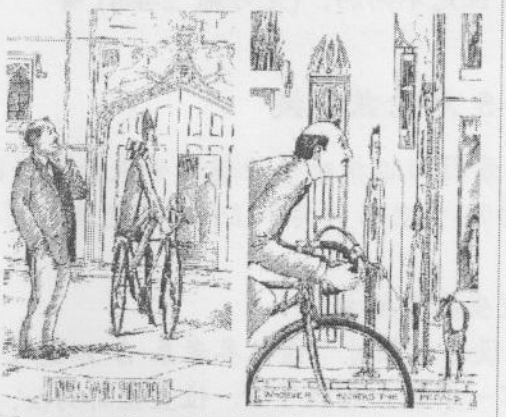


图 2-4-4 洛仑兹收缩效应也是相对的，汤普金斯先生在不同的惯性系中看到周围的物体显示出不同的几何特征。

在不同惯性系中测量沿垂直运动方向的米尺长度

为了说明在垂直于运动方向是否存在空间尺度变化的观测效应，我们可以设计这样一个实验。在两杆相同的米尺两端都安装上刀片，并平行放置(见图2-4-5a)。然后固定其中一杆米尺(设为惯性系S')，将另一杆米尺(设为惯性系S)相对米尺S'以速度u运动。如果能观测到运动米尺的横向收缩效应，则米尺S两端的刀片在经过米尺S'时将在S'上留下刮痕(见图2-4-5b)。但对米尺S而言，运动的是米尺S'，因此，S应观察到米尺S'发生横向收缩效应，则米尺S'两端的刀片在经过米尺S时将在S上留下刮痕(见图2-4-5c)。由于实际在米尺上留下刮痕是一个客观事实，不依赖于处在哪个惯性系观察的问题，上述因横向收缩效应引发的逻辑冲突实际上就排除了因运动产生的横向变形的可能。

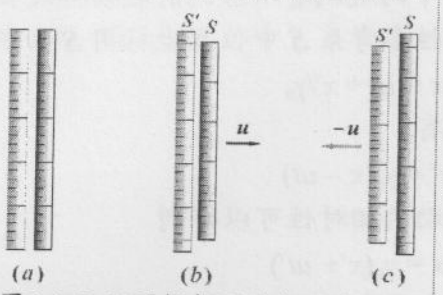


图2-4-5 (a)两根相同的米尺处于同一惯性系中；(b)若运动物体有横向收缩效应，运动米尺S两端的刀片经过静止米尺S'时将在S'上留下刮痕；(c)对米尺S而言，S'运动并有收缩效应，因而S'两端的刀片在经过静止米尺S时将在S上留下刮痕。

例 2-4-2 一根直杆在自身的参考系中与 x-轴成 θ 角，并以此固定位型沿 x-轴方向以恒定速度 v 运动，计算在地面参考系中观察到的直杆取向与水平轴的角度。

解：相对于静止的观测者而言，棒在其运动方向上出现洛仑兹收缩，而在侧向则维持其原来的尺度不变，因此他所观测到的角度应大于棒在其自身惯性系中的相应角度：

$$\tan\theta' = \frac{l'_y}{l'_x} = \gamma_v \frac{l_y}{l_x} = \frac{\tan\theta}{\sqrt{1-v^2/c^2}} > \tan\theta$$

[12] Mr. Tompkins in Wonderland, by George Gamow

爱因斯坦相对性原理要求光在真空中的传播速率对于任何惯性参考系是相等的，但并不意味着光的传播方向不发生变化。光传播方向的相对性导致光行差效应：地球本身的运动和光速的有限性使我们在不同的时候在地球上观测到一个天体的方位是不同的。(见习题 2-4-4) ◀

2-4-5 洛伦兹变换

爱因斯坦相对性原理所带来的关于时空相对性的推论可以用两个惯性系之间的时-空坐标变换方程和速度变换方程予以统一的表述。

时空坐标变换方程

设想在相对于惯性参考系 S 以速度 u 沿 **x**-轴运动的另一个惯性参考系 S' 上有一个静止的粒子 P，位置坐标为(x', y', z')。在伽利略变换中，时间间隔和空间间隔与参考系无关，粒子 P 在惯性参考系 S 中的时空坐标可以表述为：

$$x = ut + x'$$
$$y = y'$$
$$z = z'$$
$$t = t'$$

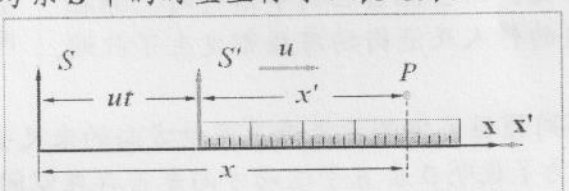


图 **2-4-6** 在不同惯性系中量度 P 的位置

而在爱因斯坦的相对论框架中，参考系 S 中的观测者所看到的运动标尺长度 x'应表现出洛仑兹收缩效应，因此粒子 P 在惯性参考系 S 中位置坐标用 S'系中的时空坐标表示时应改写为

$$x = ut + x'/\gamma_u$$

因而有

$$x' = \gamma_u(x - ut) \tag{2-4-7}$$

从运动的相对性可以得到

$$x = \gamma_u(x' + ut') \tag{2-4-8}$$

从以上两个方程可以解出两个惯性参照系中时间坐标之间的关系

$$t' = \gamma_u(t - ux/c^2) \tag{2-4-9}$$

以及

$$t = \gamma_u(t' + ux'/c^2) \tag{2-4-10}$$

由于垂直于 S'-系运动方向上的空间间隔不变，相应的坐标变换方程保留了伽利略方程形式

$$y = y', \quad z = z' \tag{2-4-11}$$

上述时空坐标的变换方程可以更简洁地表述为

$$\mathbf{r}' = \gamma_u(\mathbf{r}_{\parallel} - \mathbf{u}\,t) + \mathbf{r}_{\perp} \tag{2-4-12}$$

$$t' = \gamma_u(t - \mathbf{u}\cdot\mathbf{r}/c^2) \tag{2-4-13}$$

其中 $\mathbf{r}_{\parallel}$和 $\mathbf{r}_{\perp}$分别为位置矢量 $\mathbf{r}$ 的沿 S'相对于 S运动方向和垂直于该方向的分量。方程(2-4-12) 和(2-4-13)称为**洛伦兹变换方程**[13]。

---

[13] 亨德里克·洛伦兹(Hendrik Antoon Lorentz 1853-1928)，荷兰物理学家。因在磁性对放射的影响方面的研究而获1902年诺贝尔奖。他在爱因斯坦发表关于相对论研究的结果之前就提出了他的变换方程，这些方程后来成为狭义相对论中的基本关系式。

如果两个惯性系之间的相对运动速度远小于真空中的光速: $u/c \ll 1$, 即 $\gamma_u \approx 1$, 则洛伦兹变换方程(2-4-6)和(2-4-7)可近似简化为

$$\mathbf{r}' \approx \mathbf{r}_{\parallel} - \mathbf{u}t + \mathbf{r}_{\perp} = \mathbf{r} - \mathbf{u}t$$

$$t' \approx t - \mathbf{u}\cdot\mathbf{r}/c^2 \approx t$$

这就是伽利略的时空坐标变换方程.

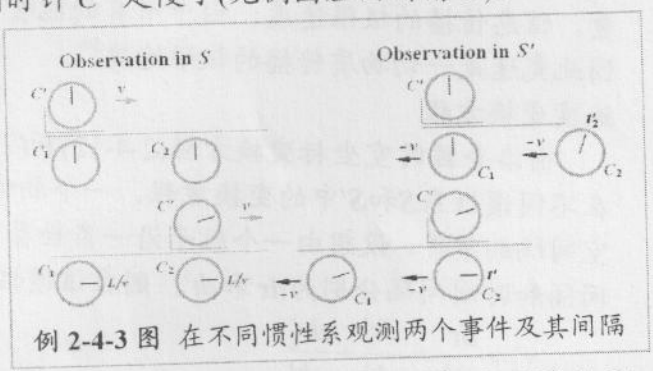
例 2-4-3 时间膨胀对称性的佯谬 设惯性系 S' 相对于另一个惯性系 S 以速度 u 沿其 x 轴方向运动。S 系的 x 轴上固定有两个同样的钟 C_1 和 C_2，两者相距 L，S'系的 x'轴上固定着一个同样的钟 C'。设想在时钟 C' 经过 C_1 的瞬间，将两个钟校准为零，同时，S 系中的观测者将钟 C_2 也校准为零。由此可以提出一个佯谬: 对于 S 系中的观测者而言，因为运动时钟 C'走慢了，当它经过 C_2 的瞬间，直接对比的结果是 C'指示的时间应落后于 C_2。但 S'系中的观测者同样可以认为 C_2 由于运动而走慢了，当它经过 C'的瞬间可以看到 C_2 指示的时间应落后于 C'。应该如何解释这种矛盾？说明 C'和 C_2 相互经过时究竟是哪个钟走慢了？

解: 首先说明，直接对比的结果是观测事实，它是绝对的而不是相对的。另外，直接对比的结果只说明在对比瞬间两个时钟指针的位置，并不说明任何两个时钟"滴答"的相对快慢。而时钟指针在某一时刻的位置与两个因素有关，一是其起始位置，二是其运动速率。在我们的问题中，时钟指针的相对运动速率是明确的，运动时间相比原时要慢一个洛仑兹因子，因此时钟指针在某一时刻的位置就取决于起始位置。

先讨论 S 系中的观测情况。C'由 C_1 的位置运动到 C_2 的位置所需要的时间为 $\Delta t\ (= t = L/u)$。而 S'系中观测到的是，C_1 与 C_2 相继经过 C'所处位置的时间分别为 0 和 t'，因此 $\Delta t' = t'$，由洛伦兹变换可得 S 系中时钟的读数为

$$\Delta t = \gamma_u(t' + 0\times u/c^2) = \gamma_u \Delta t' > \Delta t'$$

因而直接对比 C_2 和 C'，应该看到时钟 C' 走慢了(见例图左面的情形)。



例 2-4-3 图 在不同惯性系观测两个事件及其间隔

下面讨论 S'系中观测的情况。因 C'分别要和 C_1 和 C_2 对比，我们需要确认 C'和 C_1 与 C_2 是否都得到了校准。按照爱因斯坦的相对性原理，在 S 系中 $x = 0$ 和 $x = L$ 两地所发生的触发校准时钟 C_1 和 C_2 这两个事件在惯性系 S'中观察不是同时的，因此 S 系中的观测者认为时钟 C_1 与 C_2 并没有被校准: C_1 在 $t'_1 = 0$ 时被校到零(见例图右面)，而 C_2 被校到零这一事件在 S'系中的时间 t'_2 可以通过洛伦兹变换计算得到

$$t'_2 = \gamma_u(0 + Lu/c^2) = \gamma_u Lu/c^2$$

即时钟 C_2 比 C_1(与 C')抢跑了 t'_2 的时间。因此按照洛伦兹变换，S'系中观察到 C_2 到达 C'的位置的时间为

$$t' = \gamma_u(\Delta t + Lu/c^2) = \gamma_u \Delta t + t'_2$$

扣除 C_2 的抢跑时间 t'_2，C_2 的实际计时为 $\Delta t' = t' - t'_2 = \gamma_u \Delta t$，这正是 S'系中观察到 S 系中发生时间膨胀的结果。说明运动时钟发生的时间膨胀是相对的。◀

例 2-4-4 **不同事件的时间顺序与因果关系** 设在惯性系 S 中相继发生的两个事件 $A(x_A, t_A)$和 $B(x_B, t_B)$，其中 A 事件发生在 B 事件前：$t_A < t_B$。有没有可能存在这样的惯性系 S'，在其中可以观测到这两个事件的时间顺序会发生逆转，即 $t'_B < t'_A$？如果存在，其相应的条件是什么？

解：对于相对于另一个惯性参考系 S 以速度 u 沿其 x 轴方向运动的惯性参考系 S'而言，其中的观测者记录的 A 和 B 两个事件发生的时间间隔可以通过洛伦兹变换方程得到

$$\Delta t' = t'_B - t'_A = \frac{(t_B - t_A) - (x_B - x_A)u/c^2}{\sqrt{1-u^2/c^2}}$$

观测到这两个事件的时间顺序发生逆转($t'_B < t'_A$)意味着

$$\frac{|x_B - x_A|}{|t_B - t_A|} > \frac{c^2}{u} > c \tag{2-4-14}$$

也就是说，只要事件 A 和 B 的空间间隔超越了光信号联系的极限，就能观察到事件 B 发生在事件 A 前面。

讨论：由于两个事件的时间顺序可能和这两个事件的因果关系有关，时间顺序逆转是否意味着两个具有因果关联的事件之间的因果关系也会逆转？事实上，因果关系是我们能够认识自然界的最基本的信念之一，如果我们必须坚守这一信念，这也就意味着真空中的光速是自然界中物质传播的极限速度，这样才能确保在可以用实际物质联系的事件 A 和 B 之间可能的因果关系不发生颠倒。而任意两个时间顺序可以发生逆转的事件之间由于不能用任何实际物质发生联系，它们之间也就不会存在因果关联。能用光信号连接的两个事件的间隔称为**类空间隔**，而超出光信号连接的两个事件的间隔称为**类时间隔**。◀

洛仑兹变换中光速的独特地位实际上已超越了光本身的特殊性质，光速是能量，信息传播的极限速度。由于所有的能量、甚至信息都是依赖于物质传播的，因此光速是一切物质传播的极限速度[14]。

速度变换方程

由洛仑兹时空坐标变换方程(2-4-12)和(2-4-13)可以立即得到物体的运动速度在不同惯性系S和S'中的变换方程。一个物体的瞬时速度涉及两个具有无穷小时空间隔的事件，设想由一个粒子沿一条世界线上的两个事件在惯性系S'中的空间间隔和时间间隔分别为$d\mathbf{r}'$和dt'。则在该惯性系中的速度为

$$\mathbf{v}' = \frac{d\mathbf{r}'}{dt'} = \frac{d\mathbf{r}'_{\parallel}}{dt'} + \frac{d\mathbf{r}'_{\perp}}{dt'}$$

---

[14] 一个有趣的问题是如果在某种介质中对某一类物理过程而言存在着不同于光速的速度上限v，是否存在相应的以v为速度极限的相对论效应呢？目前确实发现了一个这样的实验例证：晶体中的位错的速度上限是该物体中的声速。因此，位错的变换也具有洛仑兹变换的相同形式，只是变换公式中的极限速度被晶体中的声速v_s所取代，相关的介绍可见A.Seeger & P. Schiller, *Kinks in dislocations lines and their effects in internal friction in crystals*, *Physical Acoustics* **3A**, W.P. Mason ed., Academic Press, 1966; Frank R.N.Nabarro, *Theory of Crystal Dislocations*, Oxford University Press, 1967, J.P.Hirth & J.Lothe, *Theory of Dislocations*, McGrawHills Book Company, 1968; L.B.Freund, *Dynamic Fracture Mechanics*, Cambridge U. P., New York (1990). 一个有趣的想法是能否将物质粒子理解为空间，或时空中的位错或位错环

其中$\mathbf{r}'_{\parallel}$和$\mathbf{r}'_{\perp}$分别是粒子在惯性系S'中的平行于和垂直于S'-系相对于S-系的运动方向的位置坐标，通过对(2-4-12)和(2-4-13)式的微分，可以得到用惯性系S中的时空坐标表示的$d\mathbf{r}'$和dt'

$$d\mathbf{r}' = \gamma_u(d\mathbf{r}_{\parallel} - \mathbf{u}\,dt) + d\mathbf{r}_{\perp}$$

$$dt' = \gamma_u(dt - \mathbf{u}\cdot d\mathbf{r}_{\parallel}/c^2)$$

由此可以得到两个惯性系中的速度变换方程:

$$\mathbf{v}' = \frac{d\mathbf{r}'}{dt'} = \frac{\gamma_u(d\mathbf{r}_{\parallel} - \mathbf{u}dt) + d\mathbf{r}_{\perp}}{\gamma_u(dt - \mathbf{u}\cdot d\mathbf{r}_{\parallel}/c^2)} = \frac{(d\mathbf{r}_{\parallel}/dt - \mathbf{u}) + \gamma_u^{-1}\,d\mathbf{r}_{\perp}/dt}{1-(\mathbf{u}/c^2)\cdot d\mathbf{r}_{\parallel}/dt}$$

即

$$\mathbf{v}' = \frac{(\mathbf{v}_{\parallel} - \mathbf{u}) + \mathbf{v}_{\perp}/\gamma_u}{1-\mathbf{u}\cdot\mathbf{v}/c^2} \tag{2-4-15}$$

对于非相对论极限$u \ll c$的情形，变换方程中的分母和由洛仑兹因子都趋于1，洛仑兹速度变换回复到伽利略速度变换(2-3-1)。

$$\mathbf{v}' = \mathbf{v} - \mathbf{u} \tag{2-4-16}$$

若两个惯性系之间的相对速度$\mathbf{u}$的方向沿$\mathbf{x}$-轴的方向，则相应的直角坐标分量为:

$$v'_x = \frac{u_x - v}{1-vu_x/c^2},\quad v'_y = \frac{v_y}{\gamma_u(1-uv_x/c^2)},\quad v'_z = \frac{v_z}{\gamma_u(1-uv_x/c^2)} \tag{2-4-17}$$

上述变换方程表明速度的纵向分量和横向分量是按不同的比例发生变化的。这是因为运动对时间间隔和运动方向上的空间间隔的影响程度都是由同一个因子γ_u描述的，因而在惯性系运动的方向上的速度分量$\mathbf{v}'_{\parallel}$不出现γ_u因子，而在垂直于惯性系运动的方向上没有长度变化的效应，因此相应方向上的速度分量$\mathbf{v}'_{\perp}$中时间间隔中所含的γ_u因子不象$\mathbf{v}'_{\parallel}$中的情形可以消除。

电子在直线加速器受到加速的过程中出现的现象直接验证了相对论的速度叠加法则：斯坦福直线加速器可将电子加速到 20GeV 的能量，电子在一条三公里长的真空管道中飞行的前一半路程中，电子被电场加速到 10GeV 的能量时，其速度达到 $v = (1-0.13\times10^{-8})c$，只比光速小 0.39m/s。而在后一半路程中能量增加到 20GeV，在相对于实验室参考系运动速度为 v 的惯性系中，电子的速率增加也是 v，但在实验室参考系中，其速率只增加了 0.20m/s，电子的速率达到比光速只小 0.10m/s。

速度变换方程还表明，一个以低于光速运动的物体在任何惯性系中观察其速度都不会超越光速。例如，如果一个粒子在某一时刻衰变为两个粒子，它们分别以相对于实验室坐标系的速度$0.98c$互相分离，则按照上述关系式，这两个粒子的相对速度为

$$u_x = \frac{dx}{dt} = \frac{u_x - v}{1-vu_x/c^2}$$

没有超越真空中的光传播速度。上述计算有一个实验观测的证明：π介子的γ衰变：在高能加速器中产生的π介子，其相对于实验室的速率非常接近光速。它在飞行过程中衰变，发出一种波长极短的电磁波—γ射线。根据伽利略相对性原理，γ射线相对于实验室的速度是光速与一非常接近于光速的速度叠加，大约是光速的2倍，而根据相对论，应等于光速。测出这种γ射线的速率可以检验相对论。在一

个实验中，π介子的速率$v=0.00075c$，测得γ射线速率为$(2.9977\pm0.0004)\times10^8$m/s，证实了爱因斯坦的光速不变性原理。

例 2-4-5 设地面上的观测者看到固有长度为L_A的飞船A以速度v_A向东运动，固有长度为L_B的飞船B以速度v_B向西运动。计算飞船A的飞行员观测到两艘飞船互相飞越所经历的时间。以及飞船B的飞行员观测到两艘飞船互相飞越所经历的时间.

解：我们首先在飞船A的惯性系中进行讨论，这个问题涉及到四个方面：

(a) 整个时间段分为两个部分，第一部分是飞船B的头部和尾部相继飞过飞船A的头部。第一部分是飞船B的尾部相继飞过飞船A的头部和尾部。

(b) 由于飞船B相对于地面以速度v_B向西飞行，由速度变换方程可求得它相对于飞船A的速度

$$u_x=\frac{dx}{dt}=\frac{u_x-v}{1-vu_x/c^2}$$

(c) 在飞船A的惯性系中，飞船B发生洛仑兹收缩γ_v，飞船B的头部和尾部相继飞过飞船A的头部所需要的时间为$t_1=L_B/\gamma_v v$

(d) 飞船B的尾部相继飞过飞船A的头部尾部和所需要的时间为$t_2=L_A/v$，则两艘飞船互相飞越所经历的整个时间段为

$$t=t_1+t_2=\frac{1}{v}\left(\frac{L_B}{\gamma_v}+L_A\right)=\frac{1}{v}(\sqrt{1-\beta_v^2}L_B+L_A)$$

$$=\frac{1-\beta_A\beta_B}{c(\beta_A+\beta_B)}\left(\sqrt{1-\left(\frac{\beta_A+\beta_B}{1-\beta_A\beta_B}\right)^2}L_B+L_A\right)$$

在飞船B的惯性系中可以进行类似的讨论，其结果为

$$t=\frac{L_A}{\gamma_v v}+\frac{L_B}{v}=\frac{1}{v}(\gamma_v^{-1}L_A+L_B)=\frac{1-\beta_A\beta_B}{c(\beta_A+\beta_B)}\left(\sqrt{1-\left(\frac{\beta_A+\beta_B}{1-\beta_A\beta_B}\right)^2}L_A+L_B\right)\blacktriangleleft$$

需要说明的是，运动速度不能超过光速是对物质的运动或能量的传播速度而言的，但自然界并不排斥视速度超过光速的可能性。例如快速转动的探照灯投射到遥远云层上产生的亮点移动速度就可能超过光速。又如由一原子核抛射出两块反向运行的碎片，它们相对于该原子核的速度分别为 0.6c 和 0.8c，则对相对于该原子核静止的观测者观察到的两块碎片的相对分离速度为 1.4c。视速度可以超过光速的原因在于两个事件的相关性并不是因果关联的。

时、空间隔的相对性并不是因为光速是最大的速度这一概念导出的，而是两个不同惯性系中光速之间的关系时已知的，而其它物体或信号的运动或传播速度的变换关系则不知道。因而总是用光信号来联系两个事件。

思考题

2-4-1 证明，在2-4-2节中关于车厢内中点发出的光信号能否同时到达车厢两端的例子中，如果采用伽利略的速度变换公式，“同时性”就是绝对的。即车厢内外的观测者都观察到光信号是同时到达车厢两端的。如果用铃声代替光源作决斗信号，同时性的问题应怎样考虑？

2-4-2 在推导时间膨胀效应的“光钟”实验中，如果用空气中的声信号替代光脉冲，分析应该会得到什么结果？

2-4-3 设想一个不是靠发射光脉冲，而是靠下面的一个弹簧装置向上抛射小球的时钟：小球每次上升到顶部速度恰好为零，从不同参考系看它的时间都是一样的，这是否就是所谓上帝的时钟(绝对时钟)。

2-4-4 光在介质中的传播速度为 c/n，如果物质粒子在该介质中的运动速度与光速相等，是否意味着在这个粒子的参考系中可以观察到电磁振荡不会在周围的空间中相互感应、传播的现象？

2-4-5 宇宙中有一些双星体系，其中有一颗是X射线脉冲星。当它被另一颗星遮住时，地球上就接收不到它发射的X射线，称为蚀。能否利用对X射线脉冲双星蚀的测量为对伽利略相对性原理或相对论的正确性提供实验判断。

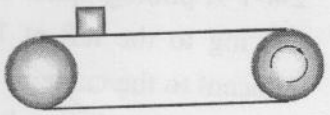
2-4-6 手持一面镜子并向前水平地伸直手臂，可以看到手中镜子中自己的像。设想你以近乎光的速度向前运动。按伽利略或爱因斯坦的相对性原理，你是否还能看到自己在镜中的像。

2-4-7 由于地球按自身的轴转动，赤道上的一个点相对于北极以速率 460m/s 运动。这是否意味着赤道上的钟比放在北极位置的钟走的慢？

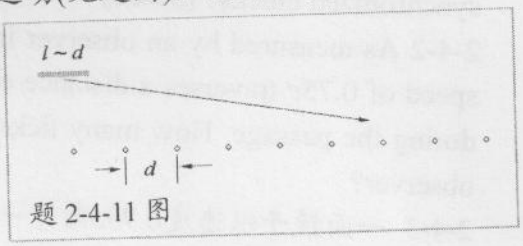
2-4-8 曾经有一种对于运动时钟走慢的解释：设想一人以近乎光的速度离开钟楼，并始终面向钟楼看，将看到指针固定在一个位置。该说法是否正确？

2-4-9 假设一个人在实验中发现某种粒子在真空中的运动速度略大于真空中的光速。我们是否就必须放弃狭义相对论？

2-4-10 一条传送带绷在两个固定间距的转动轮上(见题图)。在地面参考系中，传送带的水平部分因具有速率 v 而可以观测到收缩，这导致绷在转动轮上的传送带会绷紧，但在传送带上的物体看来，两个转动轮具有运动速度，它们之间的距离出现收缩，应导致绷在转动轮上的传送带变松。如何解释其中的矛盾。



2-4-11 设有一水平栅栏，栏杆间距为 d。一根长度与栅栏间距相当($l \approx d$)的棒以水平的姿态和一定的水平但略向下的速度运动(见题图)。就栅栏惯性系而言，运动棒的长度因洛仑兹收缩而能够穿越水平栅栏，但就固定在棒上的惯性系内观察，栏杆间距因洛仑兹收缩而小于棒的长度。因而无法穿越水平栅栏。分析两种分析结果产生冲突的原因，并说明棒究竟能否穿越水平栅栏。



题 2-4-11 图

2-4-12 双胞胎效应：设想有一对双胞胎 A 与 B，A 留在地球上，B 乘宇宙飞船去遥远的行星旅行。根据时间膨胀，A 看到 B 的钟慢了，因而 B 将变得比 A 年

轻。同样，B 看到 A 的钟慢了，因而 A 将变得比 B 年轻。如何解释其中的矛盾。

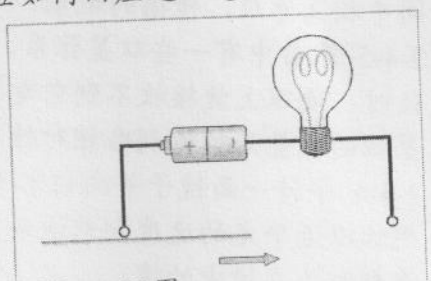
2-4-13 两个分别处于惯性系 S 和 S' 中的观测者试图验证时间膨胀效应，其中 S' 系相对于 S 系的速度为 0.6c，当两个惯性系的原点重合时，他们同时启动时钟计时。两人约定在各自的计时达到 60 分钟时发射一光信号，并在各自收到信号后再发射一确认信号。

(1) 求 S 系中的观测者收到第一个信号的时间和确认信号的时间?

(2) 设置 S 系观测者发射第一个信号，接收信号以及接收到确认信号的时间。比较 S' 中的观测者中的时间表。

2-4-14 针对2-4-2节中匀速行进火车车厢中的决斗案例中关于同时的相对性问题有一种不同的意见: 对车厢内的观测者而言，决斗的公平性表现为两个决斗者死亡的概率各为50%。而对站台上的观测者而言，面向火车头的决斗者能较早地接受到光信号，而背对火车头的决斗者接受到光信号的时间会晚一些，这表现为两个决斗者死亡的概率分别为0和100%。由于死亡的概率是客观事实，不应随着惯性系改变，因此同时的相对性问题是不存在的。应如何回应这一想法。

2-4-15 一段接有灯泡的电路两端间距为 ℓ，一根长度略短于 ℓ 的导体杆以很高的速度向右运动(见题图)。按照爱因斯坦相对性原理，对处于导体杆所在参照系中的观测者而言，导体杆可以同时接触到电路的两端，但对处于灯泡所在的参照系中的观测者而言，导体杆经过电路一侧时不能同时够到电路的两端。说明导体杆经过电路一侧时灯泡是否发光?



题 2-4-15 图

2-4-16 对雷电的观测总是看见闪电在前，听到雷声在后，有没有可能在某一个坐标系中先听到雷声，然后看到闪电?

习题

2-4-1 A photographer takes a flash-illuminated photo of a train car as the car passes moving to the left at 12m·s^{-1}. The flash goes off when the midpoint of the car is adjacent to the camera. Will two observers, one at each end of the 18-m-long train car, receive the flash simultaneously on their synchronized clocks? If not, which observer sees the flash earlier, and what is the difference in arrival time according to their synchronized clocks? (39-68)

2-4-2 As measured by an observer in an inertial frame, a clock moving at a constant speed of 0.75c traverses a distance of 60 km. The moving clock records 10,000 ticks during the passage. How many ticks pass on an identical clock at rest relative to the observer?

2-4-3 一面镜子以速度0.75c离开一光源，计算光源的像相对于光源的退行速度。

2-4-4 斜放的直尺以速度 v 相对于惯性系 S 沿 x 轴方向运动、它的固有长度为 l_0，在与之共动的惯性系 S' 中它与 x′-轴的夹角为 θ'。求 S 系中的观察者看到的直尺长度和直尺与 z 轴的夹角。

2-4-5 惯性系 S'相对于惯性系 S 以速度 v 沿 x 方向运动，在 S'系观测到一个质点的速度矢量 $\mathbf{v}'$在 x′-y′面内与 x′轴成 θ'角。求质点速度相对于 S 系的 x 轴的夹角。

2-4-6 2-4-2 节讨论关于同时的相对性的事例中，站在地面上的观察者看到火车厢内背对火车头的决斗者比面向火车头的决斗者看到光信号晚了多长时间？设火车厢长为 L，火车相对于地面的速度为 u。

2-4-7 运动中的观测者所观测到的天体视方向与静止的观测者所观测到天体的真方向之差称为光行差。光行差本质是由于光速有限以及光源与观察者存在相对运动造成的，类似于运动中的雨滴：下雨的时候，站在原地不动的人感觉到雨滴是从正上方落下的，而向前走的人感觉雨滴是从前方倾斜落下的，因此需要把伞微微向前倾斜。走得越快，需要倾斜得越厉害。光行差的成因与此相似，但需要考虑相对论效应。求解光行差的表达式。

2-4-8 推导两个惯性系中观察到一个物体的加速度的变换公式。

2-4-9 设一光束在一个折射率为 n 的介质中传播(在介质中光相对于介质的传播速度为 c/n)，当介质相对于地面以速度 u ($\ll c$)运动，求在介质中光束相对于地面的传播速度。

2-4-10 利用在两个惯性系 S 和 S'中观察到的光波波面的形状，推导出 S 和 S'之间时空坐标的洛仑兹变换方程，设系 S 和 S'的相对速度为 u。

第3章 动力学

大自然和大自然的戒律
隐匿在冥冥黑暗之中;
上帝说，让牛顿来吧!
一切都由此豁然开朗。

—亚历山大·波普

前一章中对物体运动的讨论在很大程度上仅限于几何视角：对物体运动的演化用位置随时间的变化描述，用速度和加速度来描述物体的运动状态和物体运动状态的变化。但仅仅在运动学的范围内不可能了解物体为什么会能够维持其运动状态，以及物体运动状态发生变化的原因，这些问题所涉及的一个要点是物体之间的**相互作用**，也就是通常所谓的“**作用力**”，这是**动力学**范畴内的最重要的概念和研究课题之一。

对于作用力的认识和描述涉及两个方面：第一是在观测属性上是将与“力”相关的可观测物理量的关系用数学形式表示出来。例如，一个物体在地面附近所受的重力可以表示为 $\mathbf{F}=m\mathbf{g}$，其中 m 为该物体的质量，$\mathbf{g}$ 为当地的重力加速度；再例如，弹簧的弹性力可表示为 $\mathbf{F}=-k\mathbf{x}$，其中 $\mathbf{x}$ 为弹簧的形变(拉伸或压)数量，k 是量度弹簧硬度的参量，通常称为弹簧系数。“力”的**唯像公式**往往是探索和理解相关作用**力的本原和特性**的起点。第二个方面是定量地描述“力”作用在物体上所产生的可观测的效应。这种描述可以让我们通过求解相关的数学问题预言物体的运动演化行为。

本章在讨论受力物体的运动时仍将物体抽象为“物质粒子”，这意味着我们对受力物体的运动仍局限于物体的“平动”。但需要特别提及的是这些“粒子”并非真是体积很小的所谓微观粒子。具体而言，我们不会讨论诸如电子在原子的运动，或者一个质子和一个原子核的碰撞问题。这一类涉及到真正微观物体运动的问题将在第九、和第十章中介绍。另外，本章主要讨论的是低速运动物体的动力学问题，即所谓的“经典动力学”问题，低速运动的物体所表现出来的一些特征相对于高速运动的物体行为在直觉上更容易理解和接受，在数学上的表述也更简单。高速运动物体的动力学，即“相对论动力学”问题将在3-8节中介绍。

3-1 牛顿定律

经典力学理论的建立主要归功于牛顿[1]：这一理论的最基本的要义总结在牛顿的三个著名定律之中。无论从形式的简练还是内涵的广博，牛顿定律都代表着自然科学所取得的伟大成就。但这些对物体运动规律的准确描述却是基于对实验现象的广泛而且仔细的观察，以及对前人工作的富有洞察力的理解。从原始形态回顾牛顿的工作能帮助我们更加清楚地理解如何从对物体运动行为的实际观测

[1] 牛顿 (Isaac Newton 1642-1727) 被誉为人类历史上最有影响力的科学家。其研究领域包括了物理学、数学、天文学、神学、自然哲学和炼金术。牛顿的主要科学工作为：发明微积分，发现万有引力定律，开创经典力学，设计并制造了第一架反射式望远镜等。

中总结、抽象出诸如“惯性”，“惯性系”，“力”等基本概念，而只是被这些定律的华丽数学形式所迷惑，误认为它们只是来自于天才头脑中的意念。

3-1-1 惯性与惯性定律 (牛顿第一定律)

理解一个物体受到环境的影响(作用)后会产生什么可观测的效应大致会涉及到这样一些问题：一个物体在不受环境影响情况下会处于什么样的运动状态？一个物体在什么条件下会从静止状态进入运动状态，能够维持一个物体运动状态不变的条件是什么，什么条件能够改变一个物体的运动状态。这些问题显然不是完全独立的，但应该从哪一个问题，从哪一个角度具体入手却不是容易看清楚的，而且对这些问题的求解也并非基于任何逻辑演绎，而是基于实验观察。

古希腊哲学家亚里斯多德[2]关注于第一个问题，他认为物体在不受外力作用时所处的“**自然状态**”是静止状态。这种见解是基于我们日常生活中非常熟悉的一类现象：持续地拖曳地面上的一个物体才能使其在地面上以恒定的速度运动，帆船在风的持续推动下才能在湖面上平稳地航行。如果撤除这类“**拉**”或“**推**”的作用，物体或帆船就会很快失去原来的速度并最后达到静止状态。这类现象强烈(但错误)地暗示着上述“推”、“拉”的作用(即物理学中的作用力)对维持物体的匀速运动是必须的。亚里斯多德关于“自然状态”的见解很快被他的信徒们所接受，并在亚里斯多德时代，以及以后相当长的一个时期成为大多数人头脑中的一种信念。

历史上第一个对上述概念提出质疑的是伽利略。他利用非常简单的实验设计揭示了导致物体在“推”、“拉”的作用撤除后仍然作用在运动物体上的难以察觉的阻力，如地面作用在物体上的摩擦力，大气作用在船体和船帆上的流体阻力。他**用实验观测的结果作为推理前提，再利用逻辑演绎得到的结论**说服了整个世界(在他的时代，他至少说服了自己)：导致运动物体失去了原有的运动状态是因为存在着阻碍物体运动的作用力，而不是由于缺失外部作用力的持续驱动。

图 3-1-1 是伽利略用以说明其推理过程的斜面实验。让一个小球从左面的斜面滚下去，再从右边的斜面攀爬上来[3]。伽利略从该实验揭示的第一个现象是：无论对于什么样的斜面坡度，球在滚下斜面的过程中总是获得加速，而在向上攀爬时总是被减速。由此可以从逻辑上推论，球在既不向上，也不向下的水平面上运动时就不应该失去速度也不会获得加速，即小球只能维持恒定的速度，也就是说小球会一直滚下去。这种推理不仅避开了无法在实验中验证一个球是否能在水平面上一直滚下去的难题，也说明了一个在水平面上滚动的球会逐渐慢下来并不取决于球的内在性质，而是存在着的眼

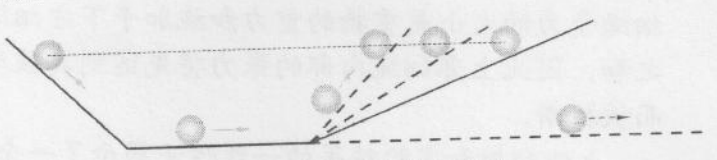


图 3-1-1 伽利略揭示物质惯性的斜面实验。

[2] 亚里斯多德(公元前 384 年－公元前 322 年 3 月 7 日)，他和柏拉图、苏格拉底一起被誉为西方哲学的奠基者。他的著作包括形而上学、逻辑学、伦理学、物理学、生物学、动物学、政治、政府以及诗歌、戏剧等。亚里士多德关于物理学的思想深刻地塑造了中世纪的学术思想，其影响一直延伸到了文艺复兴时期。

[3] 由于滚动摩擦远小于滑动摩擦，实验中选择小球而不是滑块是为了尽可能减小摩擦对运动的影响。但由于滚动增加了运动的复杂性，在实际测量中需要评估小球转动的影响。

睛不能看见的外部因素，如球与地面的摩擦以及空气的阻力。另外伽利略还利用斜面实验演示了另一个我们今天所熟知的现象：不管如何改变右斜面的坡度，球总是在攀爬到它从左斜面开始自由滚落的高度时才会停止下来。如果减小右斜面的坡度，小球要达到这一高度就必须滚动更长的距离。伽利略由此推论：如果右斜面的坡度下降到零，在水平面上滚动的球因为永远到达不了原来的起始高度，它就只能维持其到达地面时的速度无休止地滚下去。这一推理从另一个角度支持了前面得到的分析结果。

注意到处在斜面上的球总是受到沿斜面向下的重力分量，而在水平面上重力分量为零，伽利略由此推论：所有物体都具有一种属性，表现为在不受外力作用的情况下维持原来的运动状态(静止或是匀速运动)不发生改变。伽利略将物体所具的这普适属性称为**惯性**，并将上述现象总结为**惯性原理**[4]：

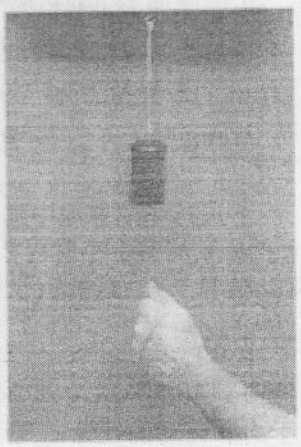
若作用在物体上的合外力为零[5]，**物体将保持其原来的运动状态—静止或匀速运动。**

事实上，物体处于静止或匀速运动这两种状态是等价的。其观测效应之间的差异只是取决于观察者所在的惯性系。处在不同惯性系中的观察者所能达成一致的是一个物体是否处在受力状态，因为这是相对性原理所要求的。否则观察者就能由此测得自身所在惯性系的速度。

伽利略对于惯性的发现和描述后来被牛顿并入他关于力和运动的统一理论框架中，因而惯性原理通常也称为**牛顿第一定律**。

例 3-1-1 一个重物用细绳悬挂在天花板上，一根同样的细绳系在重物的下端，如果突然猛拉下面的细绳，或者往下拉细绳时逐渐增加拉力。说明在这两种情况中，哪根细绳会被拉断。

解：回答这个问题的要点在于注意到重物具有比较大的惯性，即不能很快地改变它原来的运动状态(这里指改变它的静止状态)。这样，作用在下面细绳上的作用力就不能很快地通过重物传递给上面的绳子。如果在很短的时间内施加了一个拉力，由于重物的惯性阻止了其运动状态在短时间内发生显著的改变，重物上端的细绳的受力就没有明显的增加，因而当拉力超过细绳的强度，下面的绳子就被拉断。而如果逐渐增加拉力，重物就有充分长的时间传递下面的拉力，则上面细绳受力的大小是重物的重力和施加于下方细绳拉力之和，因此上部细绳内部的张力将先达到其强度极限而被拉断。



例 3-1-1 图 用不同的方式拉绳子，绳子断裂的位置不同，从而表现出系统的惯性。

上述分析和实验结果的一致性还揭示了一个现象：**力的传递不是瞬时的。** ◀

---

[4] 从这个角度可以看出引进一个概念或建立一个定律的负面因素：科学家们往往在无法进一步解释或说明一个现象及其成因时，就会将这类现象归结为物体的一个性质，并将其中的无知部分粉饰为一个科学定律。

[5] 宇宙中并不存在孤立的物体，因而不存在完全不受到任何外力作用的物体。另外，实验上并不能区分一个物体完全不受外力的作用与受到外力作用但总的外力为零时其行为的差异。

牛顿第一定律适用于各种物体，但并不适用于所有的参考系。例如，一个站在车中的人在汽车刹车时会改变原来的静止状态而突然前倾(而地面上的观测者所看到的是乘客仍然保持着原来的匀速前行的运动状态)。为此，我们将牛顿第一定律适用的参考系称为**惯性参照系**，牛顿第一定律不适用的参考系称为**非惯性参照系**。这样惯性定律就包含有两层内容，它首先揭示了物体的内在属性：物体在不受外力作用的情况下保持其运动状态不发生改变的惯性，第二是定义了能够让物体展现其惯性的一类特殊参考系—惯性参考系。

读者也许会觉得这里存在着循环定义：判断一个物体是否受到外力作用取决于我们观察到它是否处于匀速运动状态，而观察物体是否处于匀速运动状态的先决条件是我们本身必须处在惯性系，但判断我们是否身处惯性系的标准则是观察不受外力作用的物体是否处于匀速运动状态。为了解决这一问题，我们依据其它的观察和经验引进两个假设：一是作用到一个物体上的力一定来自其它物体，二是自其它物体所有类型的作用力都具有随着两个物体之间距离的增加而很快衰减的特征。基于这样的假设，我们可以将惯性系定义为一个远离所有其它物体的一个物体参考系，或者一个周围物体呈现为均匀而且各向同性分布的物体或位置所处的参考系。显然，自然界中不可能真实存在这样的惯性系，因为有限的宇宙中没有一个物体可以无限远离所有其它的物体(这样的物体也无法观测到)，另外，任何物体周围的物质分布也不是各向同性的，因而对它的作用力也无法抵消。由此可见，惯性系只是一个理想模型。但科学界通常基于上述理想惯性系存在的条件，将遥远的恒星，或多个恒星的平均位置作为近似惯性系。目前最好的恒星参考系是一个固定在1 534颗遥远的恒星的平均位置上的参考系，称为**FK4**参考系。

我们平时最常用的近似惯性系是固定在地球表面上的参考系。由于地球围绕着太阳作轨道运动，同时本身又处于自转状态，这样的参考系显然不是一个惯性系。但地球绕日运动的平均向心加速度约为 $a_{orbit} = 0.0059\ \mathrm{m/s^2}$，而地球自转的向心加速度的最大值(赤道位置)为 $a_{spin} = 0.034\mathrm{m/s^2}$，它们都远低于地球表面的重力加速度 $g \approx 9.8\mathrm{m/s^2}$。因而在许多实际场合中，固定在地球上的参照系可以近似为一个很好的惯性参考系。但在一些重力加速度并不起决定作用的场合，地球参考系就必须考虑成是一个非惯性系。例如，在讨论海洋上空贸易风(地球大气环流的主要成分)的行为时就必须计入地球的自转效应。3-1-4 节中将讨论地球作为一个非惯性系所产生的一些常见的观测效应。

3-1-2 力以及力的描述 (牛顿第二定律)

我们日常生活中所用的“拉”或“推”这类动作词汇显然不足以精确、定量地描述作用力的概念，要将“力”这个概念定义为一个物理量，需要赋予“力”一个测量属性。从测量的角度而言，可以从力的起源或力作用在物体上所产生的效果两个不同的角度进行描述。由于力的起因所具有的多样性和复杂性，从更为普适的力的作用效果的角度出发就更具有实际意义，也更为可行。为此我们从物体不受力作用时的惯性运动状态出发展开讨论。

我们首先将“物体不受外力作用时其运动状态不会发生改变”的惯性定律表述从相反的角度等价地表述为：

如果一个物体改变了它原来的惯性运动状态，则它一定受到了外力的作用。

这一新的表述意味着作用力本身虽然并不能被直接观测到，但可以通过物体对受到外力作用所产生的响应(其运动状态所发生的变化情况)反映出来。更具体地说，物体相对于原来运动状态的偏离程度(数量)可以用来量度作用在该物体上的作用力。然而用这种方法定义作用力，首先需要定义一个能够量度一个物体的"惯性运动状态"的物理量。这个物理量在物体不受外力作用时是一个不随时间变化的守恒量，而在物体受到外力作用期间，它的变化数量反映出外力对物体的作用效果。

对许多读者而言，也许很快会想到"速度"大概会是这样的物理量。因为速度确实反映了物体运动状态的重要属性(相图就是以物体的位移和速度来描述物体的运动状态的)，而且速度对时间的变化率，即加速度所具有的大小和方向两个属性也符合"力"应该是一个矢量的要求。但由此产生的一个新的问题是，如果以速度作为对物体的"运动数量"的完整描述，我们就会得到这样一个结论：具有相同速度的不同物体携载的运动数量是相同的，并可以更进一步推论：不同物体受到相同力的作用时所产生的加速度是完全相同的。

惯性质量—惯性的量化

上述推论正确与否并不是能够通过逻辑推理做出判断的，它需要通过专门设计的实验予以验证。为此，我们需要一个能控制施加在物体上作用力大小的装置，让这个装置对不同的物体施加相同的作用力，然后测量物体在受力前后的运动数量的变化。我们可以用一根弹簧作为这种施加作用力的装置，虽然开始时我们还无法对这个力进行标度(这实际上是我们最终需要解决的问题)，但我们对一根被拉伸或压缩的弹簧所施加给物体的力的大小和方向有一个确实的直观感受：一根给定的弹簧被压缩到不同的长度时和对压缩它的物体所施加的作用力大小应该是一一对应的。考虑两个物体 A 和 B 以及一条轻弹簧构成的系统(见图 3-1-2*a*)。当弹簧受到物体 A 和 B 的压缩时，两个物体都受到弹簧的排斥力，而且这两个力大小相等，但方向相反。若在某一时刻突然解除外部对系统的约束，物体 A 和 B 就会在弹簧力的作用下分别向左和向右运动，并在弹簧恢复到原来的松弛状态的瞬间，同时与弹簧脱离接触，而且分别获得各自的最大速度 v_A 和 v_B。然后两个物体分别以该速度作匀速直线运动(见图 3-1-2b)。

图 **3-1-2** 演示不同物体具有不同惯性质量的反冲实验。惯性质量可通过两个反冲物体的末速度的比值进行量度。

若加速度能完整地描述物块受到的弹簧力的特征，受到同样大小压缩的弹簧赋予反冲物体的末速度应该与物体本身的性质无关。例如，受到同样大小压缩的弹簧对木块和铅块施加作用力，它们反冲的末速度就应该相等。但实验结果表明情况并非如此。例如，对相同体积的木块和铅块进行实验，测量发现木块的末速度要比铅块的末速度大十几倍(读者可以自己做一下这样的简单实验)。即不同的物体在同样的力(同样的弹簧压缩量)作用下所获得的速度不同，这说明不同物体具有不同的抵抗自身运动状态被外界改变的能力(牛顿曾称惯性为"抵抗能力")，

而且不同物体的这种差异是可以实际测量的。物体的这种抵抗自身运动状态被外界改变的能力就是伽利略所定义的惯性，上述反冲实验的结果意味着物体的惯性是可以通过测量操作定义的物理(数)量，而不仅仅是一个只能定性说明的概念。

下面通过上述反冲实验的结果定义物体惯性的量度。因物体 A 和 B 的末速度 $\mathbf{v}_A$ 和 $\mathbf{v}_B$ 也分别是这两个物体的速度改变量：$\Delta\mathbf{v}_A=\mathbf{v}_A$，$\Delta\mathbf{v}_B=\mathbf{v}_B$，按惯性的定义，惯性大的物体运动状态不容易改变，因此物体最后获得的反冲速度也比较小。我们用 m 表示描述惯性的物理量，并将物体 A 和 B 的惯性定义为

$$\frac{m_A}{m_B}=\frac{|\Delta\mathbf{v}_B|}{|\Delta\mathbf{v}_A|}=-\frac{v_B}{v_A} \tag{3-1-1}$$

即物体A和B的惯性比与它们的反冲速度比成反比[6]，由于物体A和B的速度反向，为了保证物体的惯性质量的正定特性，(3-1-1)式中定义速度向右为正，向左为负。

定义物理量还需要对该物理量设定一个量度标准。在这里我们需要一个作为惯性单位的实物原器。由于惯性是物质的属性，而我们前面所涉及到的与物质相关的标准原器只有被称为“标准质量”的铂铱合金圆柱体。如果我们以该圆柱体替代上述实验中的物体B，得到

$$m_A=m_S\frac{v_S}{v_A}=\frac{v_S}{v_A}(\mathrm{kg}) \tag{3-1-2}$$

其中 m_S 和 v_S 分别为“标准质量”的惯性和反冲速度。历史上“标准质量”的惯性单位被直接标示为“千克”，由于“千克”是质量的量度，因此描述物体惯性的物理量被称为**惯性质量**，通常简称为**质量**[7]。

需要说明的是，仅仅上面这样一个反冲实验并不能构成惯性的定义，因为这只是在一个特定条件下(对弹簧压缩一个特定的数量)进行的实验，改变上述实验中弹簧的压缩量，物体获得的反冲速度也会随之改变。如果(3-1-2)式是对物体惯性的定义，则对于任何一组物体 A 和 B 在不同弹簧，不同压缩量导致的反冲实验中获得的不同反冲速率 $v_A^{(i)}$ 和 $v_B^{(i)}$ 的比值是一个常数。即一个物体的惯性质量是与物体运动状态无关的常量[8]：$m(v)=m(v=0)$。

另外，惯性质量是物体的一个内在属性，将物体的任何一个特性定义为物理参量的目的之一是可以对同一类的参量进行代数运算。但这种运算的内涵并不是能从逻辑演绎中来推论，而只能从大量的实验观察归纳出来的结果。例如，两个相同的物体A的惯性和物体A的惯性之间关系并不能简单地认为前者是两个后者之和，或这是后者的两倍。我们需要以两个相同的物体A捆绑在一起替代图3-1-2中的物体B以观测物体的物质数量与物体的惯性之间的数量关系。实验测得的反冲速度比为 $v_A/v_{2A}=2$。再按式(3-1-2)的定义得 $m_{2A}=2m_A$，即复合物体的惯性质量

---

[6] 定性地说，惯性只是和反冲速度反相关联的，但通过操作性定义将这种反向的关联定义为数量上的反比关系的原因在于，我们总是希望定义能给出最简单的函数关系，由定义给出的关系式就应该理解为实验结果。这也是物理量因操作性定义所具有的性质之一。

[7] 物体虽然在不受力时也表现出惯性，但只是在受力时才表现出不同物体所具有的惯差异。但在这里，以“千克”作为度量单位很容易引起误解，因为以“标准千克”为单位量度物体的“质量”时采用的是天平平衡，并不涉及物体惯性的大小，因此，一个物体的惯性为 5 千克应理解为该物体的惯性是称为“标准千克”的圆柱体惯性的 5 倍。请读者回顾所学过的有关质量的概念：惯性质量，引力质量，化学中表示物质数量，并思考一下它们之间的可能的差异。

[8] 显然，对于高速运动的物体，例如当物体运动的速度和真空中的光速数量级相当时需要通过实验重新测量这种关系，当然也需要采用不同的实验设计。

满足相加性。这一结果也说明了为什么我们可以将物体的惯性质量和物体所含的物质数量联系在一起，有时甚至不加区分的原因。

运动量—运动的数量化

不同物体具有不同数量的惯性这一事实表明，物体的"运动数量"与物体的惯性和物体的速度都有关系。或者说，能够完整地描述一个物体运动状态的物理量应该是物体的惯性和速度的函数$f(m,\mathbf{v})$。由于物体在不受外力作用时，其运动状态不随时间发生变化，因此，$f(m,\mathbf{v})$是守恒的。显然，对于一个不受外力作用的单质点系统而言，它的质量是常量，实验只能表现出速度守恒。或者说，在一个物体的运动守恒的条件下无法定出$f(m,\mathbf{v})$的函数形式。为此我们试图通过考虑由两个粒子构成的系统在不受外力作用而保持运动守恒的情况来推测这个函数的可能形式。

在前面图 3-1-2 所示的反冲实验中，物体 A 和 B 以及弹簧系统在整个过程中都不受到外力的作用，按牛顿第一定律，物体 A 和 B 在反冲前后，系统的运动总量应该守恒，即:

$$f(m_A, \mathbf{v}_A) + f(m_B, \mathbf{v}_B) = f(m_A, 0) + f(m_B, 0) \tag{3-1-3}$$

为了确定$f(m,\mathbf{v})$的函数形式，我们将定义物体惯性质量的式(3-1-1)改写为:

$$m_A\mathbf{v}_A + m_B\mathbf{v}_B = m_A 0 + m_B 0 \tag{3-1-4}$$

比较式(3-1-3)和式(3-1-4)可以看出，描述物体运动量的函数可以定义为:

$$f(m, \mathbf{v}) = m\mathbf{v} \equiv \mathbf{p} \tag{3-1-5}$$

称为物体的**线动量**，通常简称为**动量**。其国际单位制量度单位为 $kg\cdot m\cdot s^{-2}$。

方程(3-1-4)表示在反冲前后整个系统不受外力影响，它的总体动量是守恒的:

$$\mathbf{p}_i = \mathbf{p}_f \tag{3-1-6}$$

但对单个物体 A 和 B 而言，它们各自在反冲过程中都受到了物体外部的作用力(弹簧的作用)，因而每个物体的动量在反冲前后都发生了变化。仔细分析这一反冲过程可以发现，作用力(这里通过弹簧的形变具体表现出来)在两个物体发生相互作用的短暂过程中扮演着双重角色，其一是在过程中充作经纪人：通过收贷方式改变每一个物体的动量，其二是在两个物体相互动量借贷期间的任一时刻维持系统的总动量不变。

牛顿运动方程

利用动量的概念，我们可以用物体动量变化的数量来量度物体所受外力作用的大小。但实验中表现出来的物体动量的改变不是瞬间完成的，事实上力的瞬间作用并不能产生这种可观测的效应，我们看到的物体动量从$\mathbf{p}_i = 0$到$\mathbf{p}_f = m\mathbf{v}_f$是在一段时间内弹簧力持续作用于物体所产生的累积效果。如果用 $\mathbf{F}$ 表示某一时刻作用在物体上的力，并将 t 时间内力对物体的累积作用(见图 3-1-4)定义为

$$\mathbf{I} = \int_0^{\Delta t} \mathbf{F}dt = \mathbf{p}_f - \mathbf{p}_i \tag{3-1-7a}$$

称为物体在 t 时间内所受到的**冲量**。然后再用这一时间内物体动量的变化数量来量度力对物体的冲量作用

$$\mathbf{I} = \mathbf{p}_f - \mathbf{p}_i \tag{3-1-7b}$$

对于很小的Δt，上式可近似表述为

$$\Delta I \approx \mathbf{F}\Delta t \approx \Delta \mathbf{p}$$

力对物体的冲量导致物体动量的改变所表述的因果关系称为**冲量定理**。

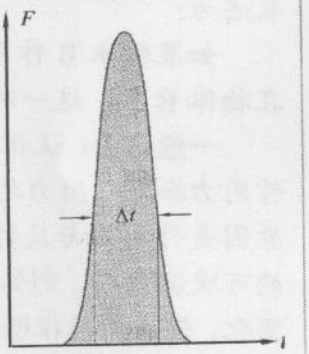


图 **3-1-4** 物体在 Δt 时间内受外力作用的积分效果—冲量的示意图。

这样，一个物体在某一时刻所受到的力就被定义为该物体的动量关于时间的变化率：

$$\mathbf{F} \equiv \lim_{\Delta t \to 0} \frac{\Delta \mathbf{p}}{\Delta t} = \frac{d\mathbf{p}}{dt} \tag{3-1-8}$$

这就是牛顿第二定律的数学表述。力在国际单位制中的量度单位 $kg \cdot m \cdot s^{-2}$ 通常被称为“牛顿”(记为 N)。需要强调的是，(3-1-8)只是量度力作用在一个物体上所产生的效果，因此等式所反映的是一种因果关系，而不是描述力和物质特性之间的恒等关系。

在经典物理范畴，或者说在宏观物体的机械运动过程中，其速度很小，在实际测量中从未观测到物体的惯性质量与物体的运动速度的关联，因此，牛顿第二定律可以简化为

$$\mathbf{F} = m\frac{d\mathbf{v}}{dt} = m\mathbf{a} \tag{3-1-9}$$

这也就是通常所说的物体的**运动方程**，因为只要知道了每一时刻的作用在物体上的力，也就知道了物体的加速度，物体运动状态的演化也就确定了。

3-1-3 作用力与反作用力 (牛顿第三定律)

牛顿第二定律所描述的虽然不是力的本原，但我们在任何动力学实验中总是可以观察到这样的事实：对一个物体施加力的作用一定是通过另一个物体实现的，即在力作用的过程中，除了受力者，还必须有一个施力者。在前面的反冲实验中，只要将 A 和 B 看成是两块不同质量的弹性体(如橡胶)就可以理解，作用在物体 A 上的力是物体 B 施加的，而作用在物体 B 上的力是物体 A 施加的。在反冲实验中，整个系统由于不受外力作用而保持总体动量守恒：

$$\Delta \mathbf{p} = \Delta \mathbf{p}_A + \Delta \mathbf{p}_B = 0 \tag{3-1-10}$$

由于只有两块弹性体(相当于两个质点间的弹簧)处于压缩状态才能对与之接触的物体施加作用，因此 A 和 B 相互作用的时间 Δt 是完全一样的。我们可以将(3-1-10)式改写为

$$\frac{\Delta \mathbf{p}_A}{\Delta t} = -\frac{\Delta \mathbf{p}_B}{\Delta t} \tag{3-1-11}$$

等式两边分别是物体 A 和物体 B 受到的对方的平均作用力，对其中的时间微元 $\Delta t \to 0$，有

$$\mathbf{F}_{B\to A} = \mathbf{F}_{A\to B} \tag{3-1-12}$$

即任一瞬间物体 A 和物体 B 受到的对方的作用力总是大小相等，方向相反的。(3-1-12)式称为牛顿第三定律的数学表述。它意味着力的作用是成对出现的，以大小相等，方向相反的作用力和反作用力的形式表现出来。牛顿第三定律的具体

表述为:

如果物体 B 作用在物体 A 上的力为 F_{AB}，则物体 A 一定同时有力 F_{BA} 作用在物体 B 上，这一对力的大小相等，方向相反：$F_{AB} = -F_{BA}$。

一般认为，认识到物体之间的作用力无以例外都是相互作用，并由此推论出作用力和反作用力之间的定量关系是牛顿完成其力学理论中最为关键的一步。其原因是作用力与反作用力大小相等并不意味着两个受力物体总是会产生的相同的可观测效应，例如，我们在向地面拍球时，球的反弹表现出其运动状态的明显变化，使我们能很明显地意识到球所受到的地面作用力，但我们完全感受不到地面受到球撞击后所产生的运动状态变化，但事实上这种地震是存在的，只是因为相对惯性质量过于悬殊，相应的加速度大小也就过于微弱而已。能够意识到常人所不能，从而也就显示出牛顿的洞察力。而对于学生读者而言，关于作用力和反作用力的概念还需要注意以下一些细节:

(1) 作用力与反作用力起因于同一个物理过程，因而必然属于同一种作用力，例如在地面上运动的物体受到地面的摩擦力作用，同时也以摩擦力的方式反作用于地面，或都属于物体中的张力。
(2) 作用力与反作用力分别作用在不同的物体上，因此两者不能互相抵消。
(3) 作用力与反作用力之间的关系只是在惯性系中成立。

例 3-1-2 一个缓慢靠向码头出小船在离码头仅一个跨步的距离时，站在船头的游客直接跳向码头。但他没能跳上码头，而是落入水中。分析其中的原因。

解: 决定游客能否成功跳上码头的因素是他的速度是否能在确定的时间间隔内能够在水平方向移动他和码头之间的距离。

When rowing boat back to the dock, you tries to jump ashore but find yourself soaked in the water. Explain why?

Solution:

Notice the forces involved when you jump: you exert forces with your muscles on the boat which reacts by exerting the forces you desire.

The boat is not very massive, so when you jump, the forces you exert on the boat cause it to accelerate out from under you.

思考题

3-1-1 为什么说伽利略斜面实验并不是一个精确的实验，它的实验误差来源是什么，估计相应的数量级，速度和高度之间是否存在一一对应的关系。从逻辑上说明伽利略斜面实验设计中所含的演绎前提，条件约束，演绎过程，以及结论的说服力。

3-1-2 伽利略关于惯性推理所基于的实验基础—斜面实验并非无懈可击，例如伽利略并不能真正穷尽所有的斜面角度(特别是极小的坡度)演示物体能够获得速度的增益或损失。为什么我们可以接受伽利略的结论。

3-1-3 为什么将恒星参考系固定在多颗遥远恒星的平均位置上，而不固定在其中最大的恒星上？

3-1-4 有一种表演：将一块大而重的石板压在一个平躺的人身上，然后用大锤猛砸石板，石板砸碎了而人却能安然无恙，解释其中的原因。

3-1-5 重物体重心越高就越不稳定，可杂技表演中演员用手指、前额或肩膀顶一个重物体，物体的重心越高反而越容易平衡。说明其中的原因。

3-1-6 在图 3-1-2 所示的反冲实验中，分析和定义物体的惯性步骤中是否隐含着某些假设？这些假设成立的依据是什么？可能会存在一些什么问题？

3-1-7 为什么力 **F** 在 t 时间内对物体的累积作用定义为 $\mathbf{F}t$，而不是其它的形式。

3-1-8 牛顿定律将作用力通过观测效应定义为

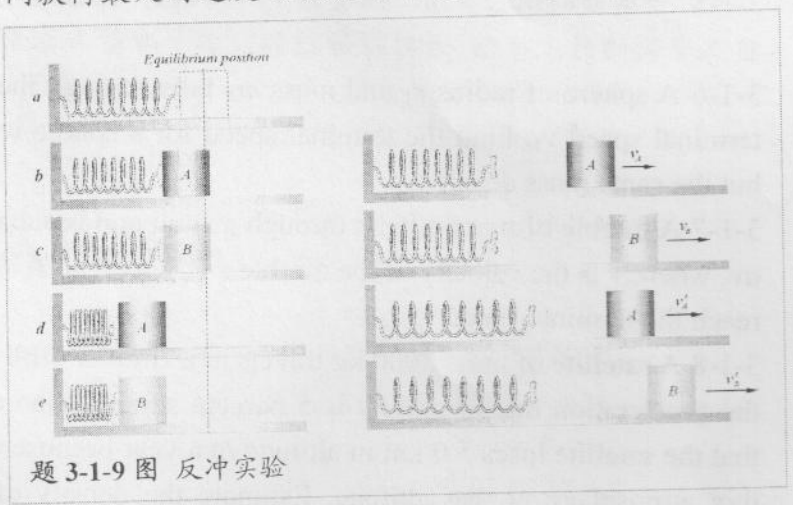
$$\mathbf{F} \equiv \lim_{\Delta t \to 0} \frac{\Delta \mathbf{p}}{\Delta t} = \frac{d\mathbf{p}}{dt}$$

说明这样的定义对力的有效描述可能带来的局限。

3-1-9 如题图所示为光滑水平面上的一个物体与一根一端固定在重墙上的弹簧构成系统，向左推动物块，被压缩的弹簧就会对物块施加一个向右的力。只要一松手，物块就会在弹簧力的作用下向右运动，并在它与弹簧脱离接触(弹簧恢复到原来的松弛状态)的瞬间获得最大的速度 v。然后物体就以该速度作匀速直线运动(见例图 b)。如果分别用两个不同的物体 A 和 B 进行上述实验，两个物体在同样的力(同样的弹簧压缩量)作用下所获得的速度 v_A 和 v_B 不同，是否可以将物体 A 和 B 的惯性质量比定义为

$$\frac{m_A}{m_B} = \frac{\Delta v_B}{\Delta v_A} = \frac{v_B}{v_A}$$

它和式(3-1-1)中定义的物体A和B的惯性质量比是否自洽。



题 3-1-9 图 反冲实验

3-1-10 如果用上面的实验得到的不同物体的末速度大小的比值定义两个物体的惯性之比，可能会出现什么问题？

3-1-11 两个物体的惯性满足相加性在物理上意味着什么？

3-1-12 将两块质量都是 m 的磁铁放在秤上。如果这两块磁铁以不同极相互面对放置或者同极相互面对放置，说明磅秤的读数有何差异。

3-1-13 人在台秤上站起和蹲下的过程中，秤的读数会出现怎样的变化。

3-1-14 为什么坐在车上的人无法推动车，但坐在轮椅上的人却能让车前进？

3-1-15 若一个物体的运动方向平行于它所在坐标系中的位置矢量 $\hat{\mathbf{r}}$。判断以下论述是否正确：(a) 作用在该物体上的力垂直于 $\hat{\mathbf{r}}$。(b) 若物体的速率保持不变，作用在物体上的力为零。(c) 作用在该物体上的力一定沿着的 $\hat{\mathbf{r}}$ 方向。

3-1-16 当小船离码头的距离进入一个人的跳跃程度时，这个人试图从船头跳跃上岸，但结果往往会因为够不到岸而掉进水里，通过过程分析说明导致该结果的

原因。

3-1-17 设计一个在失重的航天飞机中测量体重的实验方法。

3-1-18 考虑一种实验设计用于确定高速运动的物体(例如物体运动的速度和真空中的光速数量级相当)的惯性和速度之间可能存在的联系。

习题

3-1-1 说明 3-1-2 节中的反冲实验中，整个系统在反冲前后，在地面参考系中和随着相对于地面以速度 u 向右匀速运动惯性系中都能观察到系统的总动量守恒。即动量守恒在伽利略变换下具有不变性。

3-1-2 证明冲量定理在伽利略变换下具有不变性。

3-1-3 A 10-g bullet that travels at 300m/s is fired into a 30-kg block. The bullet is found 7 cm into the wood. What was the average force exerted by the wood opposing the bullet's motion? Assume that the deceleration was uniform. (5-1 modified)

3-1-4 A man of 60 kg mass standing on a bathroom scale lifts a 150-kg crate by pulling down on one end of a rope with 400 N of force (as shown in the Figure). What would the scale read in kilogram? (5-2 modified)

3-1-5 一单摆的弦长 3 m，强度为 T_o = 1 kgf，摆锤质量 0.5 kg。开始时单摆处于自然平衡的静止状态。若对摆锤施加一水平冲量 I，求能将摆弦打断的最小 I 值。

3-1-6 A sphere of radius r_1 and mass m_1 falls through the air and is found to have a terminal speed v_1. Find the terminal speed for a sphere with a radius n times greater but the same mass density.

3-1-7 A marble of mass m falls through a fluid and is subject to the drag force $\mathbf{F}_D = -\alpha\mathbf{v}$, where $\mathbf{v}$ is the velocity of the marble。利用量纲分析估计 how long it will take to reach the terminal speed.

3-1-8 A satellite of mass 3000 kg travels in a circular orbit 180 km above Earth, where the acceleration due to gravity is 5 percent smaller than on Earth's surface. Assume that the satellite loses 5.0 km in altitude in a year because of the drag of the extremely thin atmosphere at that altitude. Estimate the density of air to be at that altitude. Assume that the effective area of the satellite is 6.0 m^2 and the drag coefficient C_D is 1.0.

3-1-9 A force of magnitude 8.0 N pushes on a horizontally stacked set of blocks on a frictionless surface (Fig. 4–32) with masses m_1 = 2.0 kg, m_2 = 3.0 kg, and m_3 = 4.0 kg. (a) What is the acceleration of the stack? (b) What are the forces on block 1 as well as the net force on this block? (c) Repeat part (b) for block 2. (d) Repeat part (b) for block 3.

3-1-10 When a small sphere starting from rest falls through a liquid, it is acted by the force of gravity mg and a drag force $\mathbf{F}_D = -b\mathbf{v}$, where b is a constant. (a) find out the velocity of the sphere at moment t. (b) Sketch the speed as a function of time. (c) Sketch the position as a function of time.

3-1-11 The ultimate velocity of an object dropped from an altitude h at the equator is

less than that at the poles. (a) Analyze the origin of the difference and work out the quantity of this difference. The radius of Earth is $R = 6.38\times10^6$ m and that the acceleration at the poles is $g_o = 9.80\text{m/s}^2$. (b) Find out the acceleration of a falling object as a function of the latitude in terms of g_o. If the experiment data deviates from your prediction, what effects could be responsible for that?

3-1-12 A plumb bob suspended from a frame settles to a steady position as the frame slides down a frictionless inclined plane. What is the angle the plumb bob makes with the vertical during the slide?

3-1-13 设单位时间内冲击矿面的水量为 Q，冲击前后水的速率均为 v，求水流在作用在矿面上的冲击力。

课题

The most important properties of tissue are its absorption and strength. How can we quantify these two properties? Introduce physical quantities in operational sense and make a real measurement.

3-2 力的唯像描述

牛顿运动方程可用于求解力学中所涉及的两类问题：力的数学描述以及物体受力后的演化行为。就第一类问题而言，我们可以通过观测受力物体运动状态的变化来量度该物体所受到的作用力，由此得到的描述作用力的经验公式有助于我们理解该作用力的特征，进而探讨作用力的本源。对太阳系中的行星运动规律的认识是求解这类问题的最典型事例之一：开普勒花了十余年的时间分析第谷观察行星运动的数据记录，并最终将其总结为关于太阳系中各大行星绕日运动所遵守的三个经验公式(开普勒三定律)，而牛顿则利用这些经验公式将太阳与行星之间的相互作用总结为适用于所有物质的引力定律[1]。第二类问题是在知道物体的受力和初始运动状态的情况下，可以用牛顿运动方程预测该物体运动状态的演化(即求解描述物体运动状态随时间变化的演化方程)。

从测量属性而言，作用力是通过物体运动状态的变化情况来描述的，而物体的运动状态又可以通过观察物体的运动状态参量—位置和速度进行量度，因此各种作用力通常可以表示为描述施力物体或受力物体的位置和速度的函数[2]。特别地，如果在物体的某个特定的运动状态(x_o, v_o)附近讨论其受力情况，则作用力函数可以表示为这一特定运动状态附近的泰勒级数展开的前几项

$$\begin{aligned} F(x,v) = {} & F_o + A_1(x-x_o) + A_2(x-x_o)^2 + \cdots \\ & + B_1(v-v_o) + B_2(v-v_o)^2 + \cdots \\ & + C_1(x-x_o)(v-v_o) + \cdots \end{aligned} \tag{3-2-1}$$

其中的展开系数一般是系统的物性参数，可以通过实际测量确定。这也是这类力学定律一般被称为经验公式的原因。

经验公式虽然只是对作用力的宏观描述，但有时在一定程度也能揭示出力在作用过程中所包含的某些细节[3]。通过对具体经验公式进行数学变换，可以清楚地反映作用力的内在形式和特性。本节将介绍一些常见的作用力及其经验公式，并对这些作用力可能涉及的原因和物理过程作一些简单的探讨，使读者对基本作用力表现出不同形式的原因及其物理分析方法有一些初步的了解。

对大多数求解物体受力后其运动状态演化的工程和日常的力学问题并没有必要知道作用力的全部微观细节，通常用速度和位置这些状态参量表述的近似公式就足以理解这些特定场合中的宏观物体间的相互作用力并求解相关的演化问题。相关的处理方法将在下一节讨论。

---

[1] 从开普勒定律出发得到万有引力定律的有关过程分析见 4-2 节。

[2] 由于作用力和反作用力的对称性，施力物体和受力物体的运动状态及其变化都可用于对作用力的量度。

[3] 常见的宏观物体之间的相互作用力事实上都可归属于物体间的引力作用和物质内部电荷间的电磁相互作用，而这两种作用力在不同场合中所呈现出不同(有时甚至非常复杂)形式的原因在于物体不同的内在结构和所处的不同运动状态。虽然原则上可以通过计算物质中的所有微观粒子之间的基本相互作用来得到物体的运动状态及其演化过程，但难以想象的超大计算量实际上并不可行，而且在很多场合并不能给出直观的物理图像。

3-2-1 重力

我们在日常生活以及工程问题中最常涉及到的作用力是重力，即地球对地面附近物体的引力。从大量的实验观测结果归纳出重力具有两个明显的特征：首先，一个物体受到的重力在很大范围内和物体相对于地面的高度以及物体的运动速度无关。即重力只对应于(3-2-1)式中的第一个常数项。第二，从重力对物体的作用效果而言，它在同一个空间区域对不同的物体产生的加速度完全相同。这和上一节中所讨论的反冲实验中的弹簧力作用在不同物体上导致物体因自身具有的惯性不同而获得不同的加速度的情形完全不同。

确定重力具有这一特性的实验之一是测量自由落体的位移和时间的关系 $\mathbf{r}(t)$，最早也最具有传奇色彩的相关实验是伽里略的比萨斜塔实验：他让铁球和木球从塔顶同时自由下落，并观测到它们在同一时刻到达地面。伽里略实验中的一些细节虽然受到许多质疑，但它所反映的结果已经被其它精确度更高的实验(如光滑斜面，定滑轮系统，单摆等)所证实[4]：重力对不同大小，形状，材料的物体产生的加速度都是同样大小的常数。1971 年，这个传奇实验情节得以在月球上重演。美国宇航员戴维·斯科特在他月球表面的工作临近结束时，让一个质量为 1.32 千克的铝质地质锤与一羽质量仅为 30 克的猎鹰翎毛在同样的高度(约 1.6 米)分别从他的两只手中同时自由下落(见图 3-2-1)。收看实况转播的电视观众都亲眼目睹了翎毛和铝锤同时落到月球表面这一不同凡响的时刻。

图 **3-2-1** 翎毛和铝锤同时落到月球表面的历史时刻。

根据在地面上空所做的各类自由落体实验获得的加速度数值为[5]

$$a = 9.8\ \mathrm{m/s^2} \tag{3-2-2}$$

通常将这一加速度称为**重力加速度**，记为 **g**。

按牛顿定律，描述地面上物体所受到的重力的经验公式为

$$\mathbf{F}_G = m\mathbf{g} \tag{3-2-3}$$

物体所受到重力作用获得的加速度与物体本身的惯性无关这一结果曾在很长的时间中使物理学家感到困惑。“$a = F_G/m =$ 常数”意味着物体所受到的重力的大小和该物体本身的惯性或与惯性相关的属性有关。由于惯性反映的是物体在受力情况下维持其原有运动状态能力的一种性质，它是只能通过物体运动的变化表现出来物体的内在属性，而作用在物体上的重力的大小则和物体的运动状态无关。即物体的惯性可以在无重力的条件下测量，而重力则可以在不涉及物体运动的情形下测量。例如，在反冲实验中并不出现重力的身影；而标准的测量物体质量的天平秤平衡方法所比较的是两个物体的重力，其中被测量的物体都处于静止状态。这些都意味着没有理由认为重力和物体的惯性有任何的因果关联。唯

---

[4] 近代的这类实验证明惯性质量和引力质量差异小于 10^{-12}.

[5] 地球本身形状不是规则的球形，内部的密度分布也远不是均匀的，g 的数值在不同的地面位置有一定的差异。另外由于地球本身的自转使落体所感受到的惯性力都会使重力加速度出现一定的差异(见例 3-1-5)。平均而言，赤道处的 g 值较两极处的数值小 1%左右。

一的可能是重力与物体中相关的某种特性(通常称为引力质量)和物体的惯性都不是独立的，物体应该具有另一种更基本的属性，它的测量属性现在还不为我们所知道，但它在运动过程或受到重力的情形分别以不同的形式表现出来，并以两者在数值上严格成比例的形式反映出它们之间的密切关系。我们将在下一章的有关章节中对重力的本质做比较深入的讨论。

3-2-2 弹性力

前一节讨论牛顿定律时所涉及的反冲实验中是用弹簧来传递物体之间的相互作用力的。弹簧受到其它物体的压缩或拉伸而发生形变时总会表现出试图恢复到原来长度的行为特征，使得压缩或拉伸弹簧的物体受到一个反方向的作用力。实验表明，弹簧的这种"**恢复力**"与弹簧本身的形变大小有关，两者之间的定量关系，即弹性力的经验公式可以通过牛顿运动方程得到。如果用一个质量为m的物体系在一端固定的弹簧的另一端，移动物体使其离开平衡位置后放开物体，物体将在弹簧的弹性力作用下围绕着弹簧的平衡位置作往复的周期运动，在弹簧拉伸不大的情况下，实际测量给出物体位置随时间的变化近似为简谐振动(见图3-2-2):

$$x = A\cos(\omega t + \varphi) \tag{3-2-4}$$

其中x为物体相对于平衡位置的位移，常数参量A, ω和φ分别称为简谐振动的**振幅**，**圆频率**和**初相位**。由此可得受弹性力作用的物体的运动加速度为

$$a = \frac{d^2x}{dt^2} = -\omega^2 A\cos(\omega t + \varphi) = -\omega^2 x$$

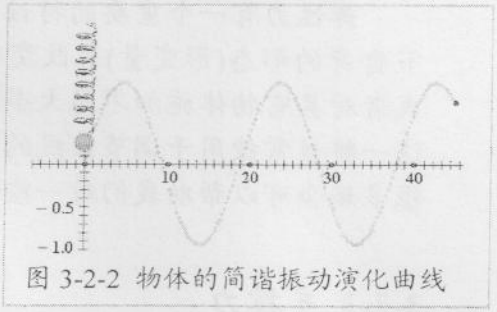


图 3-2-2 物体的简谐振动演化曲线

利用物体的牛顿运动方程可以得到描述弹性力的经验公式: $F = ma = -m\omega^2 x$，或

$$F = -kx \tag{3-2-5}$$

称为**胡克定律**。其中比例系数$k = m\omega^2$是弹簧硬度的量度，称为弹簧的**弹性系数**，在国际单位制中用 $N\cdot m^{-1}$ 作为量度单位。(3-2-5)式中的负号表示弹簧始终与形变的方向相反，说明弹簧力是一种"**恢复力**"。受到外力拉伸或压缩时产生形变，而在外力撤除后又能恢复原状，这种特性称为**弹性**。许多物体都具有弹簧的这种特性，最典型的弹性物体是橡胶，有些通常看起来难以变形的物体，如轴承中的钢球其实也是弹性体，当外力足够大时就会呈现出可观测的形变。物体的弹性的功能在于抵抗外力以避免发生进一步的变形。

例 3-2-2 设计一种方法将一根给定的弹簧在不破坏其本身的条件下让它变得更硬一些，即增加弹簧的有效弹性系数k，如使弹簧的弹性系数k增加一倍。

解: 弹簧的硬度通常与它的几何形状和材料特性有关。调节弹簧的几何形状可以改变弹簧的硬度。为说明这一点，设想一根有 10 圈的均匀弹簧，受外力F作用被拉长了 1 厘米，则其中每一圈的拉伸量就是 1 毫米，由此可知，半根弹簧的弹性常数是一根弹簧弹性系数的两倍。

将弹簧切割成相同的两半是一个解决办法，但由于弹簧系数与弹簧的圈数成反比，我们可以通过将弹簧每两圈固定在一起的方法将弹簧的有效圈数减半来获

得硬度提升一倍的效果。上面的讨论可以总结为：弹簧的硬度和长度 l 成正比，比例系数可表示为 $k=k_0 l$，其 k_0 是弹簧单位长度的弹性系数。◀

在上述分析中通常都默认整根弹簧在任何情形所发生的形变都是均匀的。事实上只有没有质量的弹簧才具有这种理想的特征。但在物体的质量远大于弹簧的情形，弹簧的形变是近似均匀的。第7章中将讨论实际弹簧的非均匀振动情况。

弹簧也是描述物体弹性的常用模型。实际物体的弹性与物体的几何形状和材料本身都有关系。胡克定律仅适用于物体形变很小的情形。图 3-2-3 所示为一根钢制弹簧被拉伸到不同程度的时候所需要的外力。在弹簧拉伸不大时与胡克定律符合得很好。但当形变较大时，外力与形变的关系就会偏离胡克定律中的弹性力和形变之间的线性关系。而当形变超越了一定的限度(称为**弹性极限**)，撤除外力后弹簧不能再回复到原来的形状，即发生了永久形变，这种特性称为**塑性**。

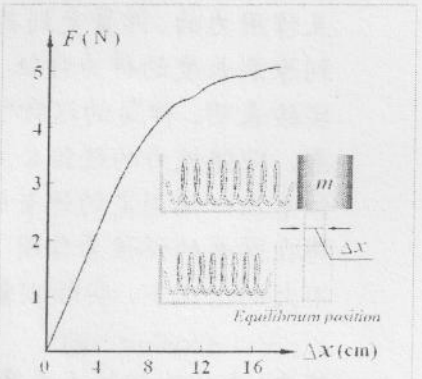


图 3-2-3 弹簧的形变量和弹性力之间的关系。

弹性力有一个重要的特性：物体可以通过调节自身的形态(形变量)来改变对外力的承受程度，或者对其它物体施加不同大小、方向的作用力。这一特性常被用于调节系统的力学平衡，另外在很多场合可以帮助我们对一些作用力建立直观的物理模型。

3-2-3 支撑力

一个放在桌面或地面上的物体能够维持平衡，说明作用在物体上的重力被来自桌面或地面的某种支撑物体的力抵消。实际观测表明该作用力具有两个特点。(1)支撑力不会独立出现，它总是和作用在同一个物体上的其它力的与支撑面垂直的分量平衡(最常见的例子是和重力同时出现)，因此它总是在支撑力方向上处于所受净力为零的状态。(2)支撑力总是与物体所接触的表面垂直。对被抽象成质点的物体，支撑力作用在接触点上，与物体所接触的表面垂直。如图 3-2-4 中斜靠在墙上的长杆受到地面和墙面的支撑力分别垂直于地面和墙面。

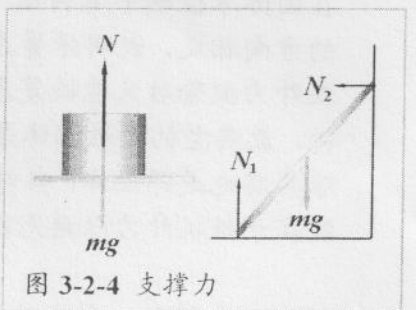


图 3-2-4 支撑力

由于通常观察不到物体在支撑力的作用下其运动状态发生的变化，我们不能简单地将支撑力表示为(3-2-1)式那样的运动参量的函数。但从支撑力总是以平衡力的形式出现的特性可以意识到，它是桌面或地面受到物体的正压力而作用在物体上的反作用力。而且一个物体的表面(例如桌面)在很大的范围内能够根据放在它上面的物体所施加的不同大小的正压力调节支撑力的大小，维持与作用在该物体上的重力达到平衡。从上一小节关于弹性力的讨论中知道，宏观物体受到其它物体的挤压时总会发生形变，因此物体反作用与其它物体的支撑力的本质就是弹性力，只是一般用作支撑物的物体都比较硬，在受到通常大小作用力的挤压下形

变量很小，不容易被眼睛所察觉。而弹性力也能解释同一个支撑面为什么能够根据它上面物体所施加的不同大小的正压力调节支撑力的大小，维持与作用与该物体上的重力达到平衡。由于这一对弹性力都起因于一个物体受到另一个物体的压缩，因此支撑力总是只表现出弹性力的排斥特性。

3-2-4 摩擦力

摩擦力在工程领域和日常生活中所起的积极和消极作用是无论怎样评价都不过分的。一方面摩擦导致材料的磨损，动力的浪费[6]；而另一方面，如果没有摩擦，人类就不会有起源于摩擦取火的文明成就，甚至都不会具有行走的能力。

我们对摩擦的感受和经验来自于这样的事实：摩擦力总是出现在两块具有相对运动的固体接触面之间，并且与相关物体的运动方向相反，因而总是阻碍而不是有助于两个相接触的物体间的相对运动。因而这种摩擦通常称为**滑动摩擦**或**运动摩擦**。但摩擦实际上也会出现在两个固体的接触面之间不发生宏观尺度的相对运动情形。例如，对静止在地面上的重物施加一个水平方向的力 **F**，除非 **F** 足够大，否则重物仍将静止在原来位置。这表明重物和地面间存在另一种总能与 **F** 平衡的阻尼力，称为**静摩擦力**。不同于运动摩擦力的方向总是和物体速度的方向相反，静摩擦力的方向总是与作用在物体上的驱动力在接触面分力的方向相反。

为了得到摩擦力的经验公式，我们需要了解摩擦力可能与受力物体的哪些性质有关。就直觉而言，摩擦力起源于物体表面的相互接触，摩擦力 f 就应该与接触面积 S 和接触面的性质有关，较大接触表面应该产生较大的摩擦力，我们用一个宏观物性参数 a 表征接触面的性质。另外，动力学过程中物体质量 m 也是需要考虑的因素，而对于运动摩擦而言，还应该计入物体间的相对运动速度 v。因此，静摩擦力 f_s 和滑动摩擦力 f_k 可以分别表述为上述参数的函数：

$$f_s = f_s(a, S, m)$$

$$f_k = f_k(a, S, m, v)$$

下面就如何通过实验和分析确定上述的函数形式做一简单的讨论。

观察和测量摩擦力的典型实验设计是斜面实验：一个质量为 m 的物体放置在一个倾角为 θ 的斜面上。若物体能在斜面上保持静止状态，说明静摩擦力和物体所受重力沿斜面的分量 $mg\sin\theta$ 达到了平衡：$f_s = mg\sin\theta$ (见图 3-2-5(a))。缓慢增加桌面倾角，维持物体平衡的静摩擦力随着物体所受重力的沿斜面分量等量增加，直到对应与物体临界平衡状态的最大倾角 θ_c (倾角再有微小的增加将导致物体开始下滑)。这时的静摩擦力 $f_{s\max} = mg\sin\theta_c$ 称为物体与斜面间的**最大静摩擦力**。

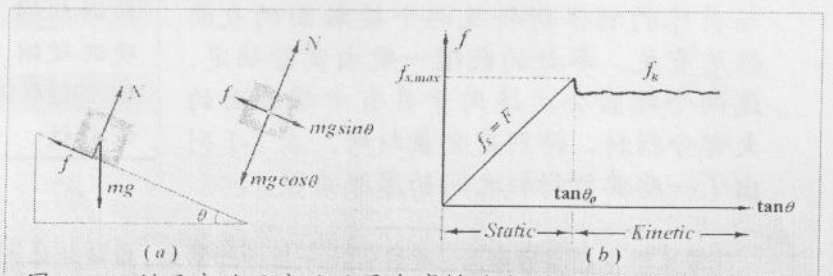


图 3-2-5 斜面实验及实验结果中摩擦力与斜面倾角的关系曲线

由临界角 θ_c 再增添一个无限小量，

[6] 世界上大约 70%的能源被消耗在摩擦中，大量人力和财力被用于研究如何有效地减少摩擦。

物体将相对于斜面下滑。这说明物体获得一个很小但不是无限小的加速度，或者说摩擦力发生了不连续的减小[7]。图 3-2-5 (*b*)所示为摩擦力随驱动力 $mg\sin\theta$ 增加时的变化情况。对摩擦力表现为斜面倾角的不连续函数(严格地说是摩擦力表现为速度的不连续函数)的一种合理解释是物体和斜面处在相对静止和相对运动状态时的摩擦力具有不同的动力学机制[8]。

如果在物体的滑动过程中进一步增加斜面倾角，物体运动速度将会出现持续增加，测量物体的加速度可以获得运动摩擦函数 f_k 和运动速度之间的关系。图 3-2-5(*b*)中滑动摩擦段的实验结果表明滑动摩擦力与物体的运动速度无关[9]。图中滑动摩擦力的微小起伏是不规则的涨落，并不反映确定的阻尼机制。

为确定摩擦力与接触面积 S 之间的关系，达·芬奇[10]曾在斜面实验中用一块长、宽、高三个方向尺度不同的长方体的三个不同面积的侧面与斜面接触做了一组实验，该实验中除了长方体与斜面接触面积不同，另外的两个参数 a 和 m 都没有变化，由此可以得到摩擦力与接触面积 S 之间的关系。但实验结果表明，作用在物体表面上的摩擦力并不因为接触面积不同而发生可观测的改变。

为了获得摩擦力与物块的质量的相关性，达·芬奇曾用两个相同物块叠放在一起以排除实验中因表面差异对摩擦力的影响，实际测量发现两块物体的摩擦力比一块物体的情形增加了一倍，即摩擦力与物体的质量成正比。但接着在不同坡度的斜面重复该实验时又发现，摩擦力不仅与物体的质量相关，而且还和斜面倾角的余弦成正比：$f \propto \cos\theta$。由于在这个实验中，$\cos\theta$ 只可能以某种物理量矢量的分量的形式出现，而唯一的相关的物理量只有重力垂直于斜面的分量。这意味着摩擦力并不是直接和物体的质量发生关联，而是与物体施加在接触面上的正压力 $N = mg\cos\theta$ 成正比。以上这些实验事实可以归纳为：

$$f_{s,max} = \mu_s N \tag{3-2-6}$$

$$f_k = \mu_k N \tag{3-2-7}$$

这就是通常广泛使用的描述摩擦力的经验公式，其中比例常数 μ_s 和 μ_k 分别称为**最大静摩擦系数**和**滑动摩擦系数**，它们与两个接触面的面积以及相对速度无关，但与具体的物体材料及两个接触面的表面性质有关。参数的数值一般由实验确定。这两个经验公式适用于具有干燥表面的大部分材料，特别是金属材料。表 3-1 列出了一些典型材料之间的摩擦系数。

表 3-1 一些材料之间的摩擦系数

材料	μ_s	μ_k
铝-软钢(低碳钢)	0.61	0.47
铸铁-铸铁	1.1	0.15
铜-铸铁	1.05	0.29
铜-软钢	0.53	0.36
玻璃-玻璃	0.9-1.0	0.4
玻璃-镍	0.78	0.56
皮革-橡木(顺纹理)	0.61	0.52
镍-镍	0.7-1.1	0.53
橡木-橡木(顺纹理)	0.62	0.48
橡木-橡木(垂直纹理)	0.54	0.32
软钢-黄铜	0.51	0.44
软钢-铅	0.95	0.95
软钢-软钢	0.74	0.57
硬钢-硬钢	0.78	0.42
钢-钢镀锌膜	0.5	0.45
锌-铸铁	0.85	0.21

[7] 日常生活中的类似经验：所推重物一旦开始滑动，维持物体滑动的推力比原来的略小一些。

[8] 实际上函数的变化趋势突变，即导数不连续也意味着摩擦力的控制机制发生了改变。

[9] 这一结果是库仑首先发现的。

[10] 列奥纳多·达·芬奇 (1452-1519) 意大利文艺复兴时期最负盛名的美术家、雕塑家、建筑家、工程师、机械师、科学巨匠、文艺理论家、哲学家、诗人、音乐家和发明家。

例 **3-2-3** 利用上述斜面实验说明测量物体间最大静摩擦系数和滑动摩擦系数的具体步骤。

解：如图 3-8 所示，作用在斜面上静止物体上的力分别为重力 $m\mathbf{g}$(m 为物体质量)，支撑力 $\mathbf{N}$ 以及静摩擦力 $\mathbf{f}_S$。当斜面倾角增加到临界值 θ_c 时，沿斜面方向的力平衡方程为

$$mg\sin\theta_c = \mu_s N = \mu_{s,max} mg\cos\theta_c$$

临界角 θ_c 的值满足

$$\tan\theta_c = \mu_{s,max} \tag{3-2-8}$$

即临界角的正切函数值就是最大静摩擦系数。

在物体进入滑移状态后，决定物体运动状态的牛顿方程为

$$mg\cos\theta - \mu_k mg\cos\theta = ma$$

其中 $\mu_k mg\cos\theta$ 为滑动摩擦力。由于 $\mu_{s,max} = \mu_k$，可以在物体的滑动过程中通过减小桌面的倾角以减小物体的加速度。当物体进入匀速下滑状态时所对应的桌面倾角使 θ'_c 为另一个临界角。它通过物体匀速下滑的运动方程直接给出了物体的滑动摩擦系数

$$\tan\theta'_c = \mu_k \tag{3-2-9}$$

因实验中难以精确判断物体加速度是否为零，由此得到的 μ_k 值不如 μ_S 精确。◀

摩擦力公式(3-2-6)和(3-2-7)虽然简单，但也包含了许多令人费解的地方。从物理因果关联上分析，摩擦力明显地起因于物体之间的接触，因此，应该表现出与接触面积和外部驱动力相关，而不应该和两个物体之间的弹力(正压力)有什么关联性。另外从数学描述而言，摩擦力矢量和正压力矢量互相垂直，而经验公式中的摩擦系数只是一个无量纲的标量参数，无法在摩擦力公式(3-2-6)和(3-2-7)中标示出各个力的方向关联。所有这些都意味着描述摩擦力的经验公式(3-2-6)和(3-2-7)并没有反映出摩擦过程的实际物理内涵。尽管如此，这些明显的漏洞也变相提供了一些可以继续探讨的方向。我们可以通过对经验公式进行变形，是其呈现出摩擦过程中相关物理量的内在关联。

我们首先按照实际物理图像将摩擦力表述为矢量的相关分量形式，设摩擦力沿 x 轴方向，弹力 N 沿 z 轴方向。为体现这两个力之间的方向关联，摩擦系数应为一个二阶张量，将其表示为 μ_{xz}:

$$f_x = -\mu_{xz} N_z \tag{3-2-10a}$$

其矩阵形式为

$$\begin{pmatrix} f_x \\ f_y \\ f_z \end{pmatrix} = \begin{pmatrix} \mu_{xx} & \mu_{xy} & \mu_{xz} \\ \mu_{yx} & \mu_{yy} & \mu_{yz} \\ \mu_{zx} & \mu_{zy} & \mu_{zz} \end{pmatrix} \begin{pmatrix} N_x \\ N_y \\ N_z \end{pmatrix} \tag{3-2-10b}$$

为了反映摩擦力应与物体接触面积有关的问题，可以设想两个物体的表观接触面积并不是实际的接触面积，真正的有效接触面积 S_z 可能是正压力的函数：$S_z = S_z(N)$，在正压力 N 不是非常大的时候，有效接触面积与正压力成正比[11]

$$S_z = kN \tag{3-2-11}$$

[11] 函数 $S_z = S_z(N)$ 泰勒展开的线性项。

则摩擦力可以改写为。

$$f_x = -\mu'_{xz} S_z(N) \tag{3-2-12}$$

其中 μ'_{xz} 为物体单位面积所受到的侧向力。

科学家们从 15 世纪就开始意识到对摩擦力定律以及对摩擦力成因的解释取决于对物体表面的结构的理解。摩擦的内在机制在不同尺度是不同的。如果物体的接触面在宏观尺度或介观尺度(微米量级)是粗糙不平的[12]，这样的两个表面互相接触时，面上的一些凹凸部分相互啮合(见图 3-2-6-a)，物体沿接触面在相对滑动时的摩擦主要表现为因凸起部分发生碰撞，和断裂而形成物体磨损并伴随有能量损耗。这样一个物体表面粗糙不平的图像同时揭示了一个重要事实：当这样的两个表面叠合在一起时，有效接触面积只占表观接触面积的很小一部分，因此很小部分的接触面实际承受了物体的全部负载—正向压力。正压力较小时仅导致接触部分的顶部发生弹性形变，而压力超过凸起部分的弹性限度(约？)，微凸的顶部因塑性形变而被压平，导致实际接触面积增大(见图 3-2-6(b))。一般的负载情况下，实际接触面积和负载力成正比，这就是(3-2-11)和(3-2-12)给出的结果。

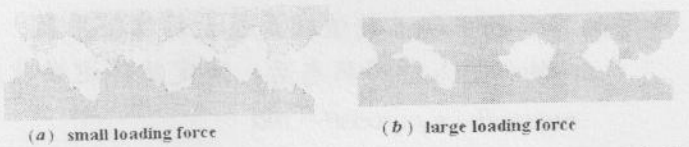


图 3-2-6 表面微小的凹凸不平因物体彼此间的正压力被压平，导致实际接触面积的增加。

微观尺度上的接触意味着两个表面的分子间距在分子引力可以发生有效作用的范围内。这导致两个接触面上的原子或分子发生吸引甚至成键。或者说接触部分实际上已融为一体了。要使两个接触的表面发生相对滑动，就必须对其中的一个表面施加一个切向力以克服分子引力，撕开实际接触区的联结，从而产生对运动的阻力。

对物体表面结构从不同尺度的所作的分析，可以将摩擦力起因归纳为两种层次的效应：宏观和介观尺度上(微米尺度或更大)的表面不平整导致两个物体表面的凹凸部分相互间的锲入，而相对运动将导致锲入部分的断裂，该效应是通常观察到的产生表面“磨损”的直接原因；而两个物体表面在微观尺度上的接触部分由于分子的短程引力作用导致接触点间两个表面的分子发生结合(化学键合)；这些被“点焊”在一起的微小接触面在滑移过程中必须被重新撕开，甚至有些分子被转移到新的位置。

这两种层次的摩擦机制都与表面的有效接触面积有关，有效面积越大，摩擦也越大。有两种办法可以有效地增加表面的微观接触面积，其一是用外力使表面结构产生形变，从而增加两个表面之间的契合以有效扩大微观接触面积。而增加表面的正压力可以有效提高接触百分比，这就是实际测量中摩擦力与正压力成正比的原因。第二种是降低物体表面的粗糙度。从日常经验而言，降低表面的粗糙度似乎总是导致摩擦力减小，但那只是宏观意义上的粗糙度：降低表面的粗糙度会导致互相啮合程度降低，减小磨损。但提高光滑程度在微观上会导致两个表面的大量原子或分子进入有效的相互吸引力程而发生结合，从而使摩擦力急剧上升。

[12] 一些宏观上看非常光滑的机械抛光面，在介观尺度上仍然存在大量的不规则的隆起

一个实际的例子是两块玻璃板之间的摩擦系数要比书本和桌面间的摩擦系数大很多。因为在平板玻璃上推动木块很容易，而在平玻璃板上移动一块相同质量的玻璃就会困难得多。

有效的微观接触面积大小也可以大致解释为什么两个给定的固体表面的静摩擦系数一般要大于它们之间的滑动摩擦系数。这是由于在静止情况下材料之间有足够长的时间得以相互锲入。摩擦力起因两块固体的接触表面间的相对运动。键的弛豫使在运动过程中的没有进入最稳定的键合程度

顺便提一下，尽管摩擦力与接触面面积大小无关对于一大类材料，如金属，是很好的近似，但并非对所有的材料都成立。例如，对塑料和橡胶这类材料而言，摩擦系数依赖于接触面的形状，这显然是和这类材料的长链分子的特性有关。

另外，摩擦力与物体的运动速度无关这一现象只是在一般的速度范围内近似成立。实际上，摩擦力在低速区域是随着速度的增加而下降的(在一定程度上，它和两个物体之间的静摩擦一般要大于它们之间的滑动摩擦的原因类似)，而在高速区域，摩擦力随着速度的增加而升高(见图 3-2-7)。这种变化都涉及到能量耗散机制的变化。如果两个表面的相对速度足够大，在接触区的有些地方可能发生局部融化，尽管整个接触面总体上的温度只是稍微有些发热。但在许多速度不是如此极端的日常工程问题中偏离(3-2-5)式是完全可以忽略的。

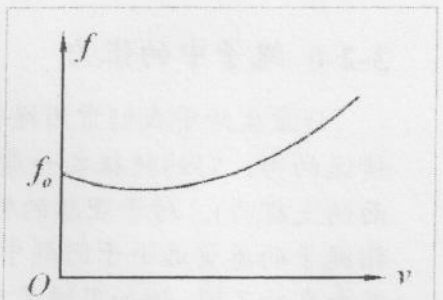


图 3-2-7 摩擦力与速度的关系

从以上分析可知，运动摩擦大多和物体的表面结构破坏以及表面分子的迁移有关。下面再讨论一下静摩擦力的主要来源。从上面讨论的物体的表面结构和表面的接触图像，以及静摩擦力具有可以调节大小以平衡外力的特性可知，静摩擦力应该是一种弹性力，但不同于支撑力的情形，静摩擦力总是处于接触面内，是一种"切向"形变导致的拉伸型弹性力。

在日常生活和工程领域，除了静摩擦和滑动摩擦，有时还会提到"滚动摩擦"，即在滚动问题(如滚动轴承)中的阻力问题。滚动摩擦这个词汇很容易将这种阻力误认为是另一类摩擦力，但其实质与我们所讨论过的静摩擦和滑动摩擦很不相同，可以认为是一种赝摩擦。关于形成滚动摩擦的内在机制将在第 5 章中讨论。

涉及摩擦，磨损以及润滑问题的是一门很大，也是极具实用价值的学科，称为**摩擦学**。该领域中目前已积累了大量的知识和信息，能使我们设计和制造具有特定的摩擦性能的新材料。但对于摩擦的基本理解至今仍然没能得到正确、完整的认识。

3-2-5 流体阻力

固体在固体表面上运动会受到摩擦阻力，如果在两个固体表面用一层很薄的液体或气体分开，就可使运动阻力大大降低，相应的技术称为润滑。但是固体在液体或气体中运动时仍然会受到阻力，称为流体阻力。这一小节我们讨论固体在流体中运动时受到的流体阻力的特征及其唯象描述。

由于均匀流体中的各个位置是等价的，流体阻力只是和运动物体的速度相关。

流体阻力对运动物体的作用表现为它与运动速度的方向相反，但与滑动摩擦不同的是它依赖于物体在流体中的运动速度。在低速运动范围，流体阻力与物体的运动速度成正比，而随着速度增加，流体阻力主要表现为与速度的平方成正比，这意味着流体中开始出现了一些新的效应。而当物体在流体中的运动速度进一步增加，速度的指数高于 v^2 的关联将会变得越来越重要，但这里我们假设它们很小而予以忽略。因而对在给定的流体中运动中的给定物体而言，流体阻力大致可以表示为

$$F_D = -b_1 v - b_2 v^2 \tag{3-2-13}$$

其中系数 b_1 和 b_2 以很复杂的方式依赖于运动物体的大小和形状，以及流体的特性，如流体的密度和黏性。另外，这些系数也包含有流体运动状态的信息，例如当物体处在流体中高速运动状态时，v^2 项变得非常重要(见习题 3-2-6)。需要强调的是方程(3-2-13)只是关于速度的复杂函数的一个近似。导致出现流体阻力的原因很多，分析也非常复杂，第 6 章中会对此做一些简单的介绍。

3-2-6 绳子中的张力

日常生活中我们常用绳子来拖曳物体。张力 **T** 是由柔软的绳子沿绳子的方向传递的力，它的特征之一是总是表现为拉伸力，没有排斥力的成分(对比固体表面的支撑力)。对于理想的轻绳而言，其中的张力分布是均匀的，通常的轻绳是指绳子的质量远小于问题中其它相关物体的质量[13]。张力的变化依据不同的情况会有所不同。例如用绳子在拖动一辆满载的轨道车厢时其中的张力要比拖动空载的车厢时大得多。张力源于使绳子聚在一起又使其柔顺易变形的分子力。值得注意的是绳子中的张力是相同的，即使绳子拐过一个角度，只要转过无摩擦的(理想)轻滑轮。

例 3-2-4 分析绕在绞盘上的一根轻绳中的张力分布。

解： 由于绳子的质量可以忽略，绳子中张力的不同只能来源于绳索和绞盘之间的摩擦力的影响。设绳与绞盘间的摩擦系数为 μ，绳子绕圆柱的张角为 θ。θ 处对圆心张角为 $\Delta\theta$ 的一段线元受到四个力的作用：线元两端张力 $T(\theta)$和 $T(\theta+\Delta\theta)$，法向弹力 ΔN，以及摩擦力 $\mu\Delta N$。在无加速度的情况下合力为 0，切向和法向分量的力平衡方程分别为

切向：$-\mu\Delta N = [T(\theta+\Delta\theta) - T(\theta)]\cos(\Delta\theta/2) \approx \Delta T$

法向：$\Delta N = [T(\theta+\Delta\theta) + T(\theta)]\sin(\Delta\theta/2) \approx T\Delta\theta$

合并得到：$\dfrac{dT}{T} = -\mu d\theta$。方程的解为

$$T_B = T_A e^{-\mu(\theta_B - \theta_A)}$$

可以用小得多的力 T_B 拽住绳子与绳子负荷的巨大拉力 T_A 平衡。若摩擦力可忽略，$\mu \to 0$，$T_B \approx T_A$，两端绳的张力相等。为轻绳跨过无摩擦滑轮的情况。◀

---

[13] When the mass of the rope is not negligible, then the tension will vary along the rope.

张力能在一定的范围内传递大小不同的拉伸力，或平衡外部作用力的行为展示出弹性力的主要特征，因此张力只是弹性力的吸引部分。绳索不能表现出弹性力的抗压缩特性主要在于绳索本身容易发生扭曲，难以使其真正发生有效的压缩(读者可以比较支撑力只是弹性力的排斥部分)，具有类似特征的物体系统还有液体的表面，由此表现出的张力称为**表面张力**[14]。如果能使物体的形态发生有效的压缩，就能使物体表现出弹性力的排斥部分。

3-2-7 非惯性系与惯性力

牛顿的三个定律都只适用于惯性系。按伽利略相对性原理，这些定律必须在两个不同的惯性系 S 和 S' 中以(3-1-14)同样的形式成立，即在惯性系 S 和 S' 中对物体的受力和运动状态演化的观察和描述必须是相同的。设 S' 系相对于 S 系以速度 $\mathbf{u}$ 作匀速运动，这两个惯性系中的观测者通过测量发现某一运动物体相对于自身所处的惯性系 S 和 S' 的速度分别为 $\mathbf{v}(t)$ 和 $\mathbf{v}'(t)$。伽利略的速度变换给出在 S 和 S' 所观测到的运动物体的不同速度间的关系为 $\mathbf{v}=\mathbf{v}'+\mathbf{u}$，则运动物体相对于 S 和 S' 的加速度间的关系可通过对其求关于时间的导数得到：

$$\frac{d\mathbf{v}}{dt}=\frac{d\mathbf{v}'}{dt}+\frac{d\mathbf{u}}{dt} \tag{3-2-14}$$

由于 S 和 S' 之间的相对运动 $\mathbf{u}$ 是常矢量，则 $d\mathbf{u}/dt=0$，因此

$$\frac{d\mathbf{v}}{dt}=\frac{d\mathbf{v}'}{dt}$$

另外，按照牛顿的假设，物体的惯性质量与物体的运动速度无关，在上式两边乘上物体的惯性质量 m，从而得到

$$\frac{d\mathbf{p}}{dt}=\frac{d\mathbf{p}'}{dt} \tag{3-2-15a}$$

即

$$\mathbf{F}=\mathbf{F}' \tag{3-2-15b}$$

这说明牛顿定律在不同的惯性系中是等价的。这个结果实际上也是相对性原理所要求的，否则我们就可以通过实验所显示的力学定律与某一个标准形式的偏差得知自己是否处于绝对运动状态，以及相应的绝对运动速度。

我们在许多实际情形中是在非惯性系中讨论物体运动的，而有时处在特殊的非惯性系中讨论物体运动可能更为方便，如分析加速行驶着的火车中的物体运动，或讨论转动着的地球上的物体运动等。处理这类问题时可先在惯性系中用牛顿定律讨论物体的运动，然后用相对运动的公式将其变换到非惯性系中，以求得物体在非惯性系中的运动。但这样做有时会非常麻烦。另一种处理方法是引进一种被称为惯性力的假想力，然后直接在非惯性系中用牛顿定律求解物体的运动。

设 S 为非惯性系，而 S' 为非惯性系的情形，这时 S' 相对于 S 的运动速度 $\mathbf{u}(t)$ 不是常矢量，观测者 S 和 S' 所观测到的运动物体的加速度之间的关系为

[14]

$$\mathbf{a} = \mathbf{a}' + \frac{d\mathbf{u}}{dt} = \mathbf{a}' + \mathbf{a}_i \tag{3-2-16}$$

在上式两边乘上物体的惯性质量m，注意到惯性质量与参考系武无关，由此得到

$$m\mathbf{a} = m'\mathbf{a}' + m'\mathbf{a}_i \tag{3-2-17}$$

或

$$\mathbf{F}' = \mathbf{F} - m'\mathbf{a}_i \tag{3-2-18}$$

这显然与牛顿定律的形式不同，我们在前一章关于伽利略变换的讨论中知道，相对于牵连速度的变化率是非惯性系中特有的惯性加速度：

$$\frac{d\mathbf{u}}{dt} = \mathbf{a}_i$$

在(3-1-18)式中与之相对应的项为

$$\mathbf{F}_i = -m\mathbf{a}_i \tag{3-2-19}$$

称为**惯性力**。该力的特殊性在于当我们处于非惯性系中时可以观测到相关的力学效应，我们甚至在生理上也能直接感受到这种力的存在，但无法在周围找到施力的物体，当然也就谈不上相应的反作用力。事实上惯性力只是观测者对所处的惯性系中观测到的物体运动演化行为与惯性系中观察到的情形所存在差异的一种抽象描述，并不具有通常的相互作用力所具有的"相互作用"本质。因此惯性力也经常被称为**赝力**。

最典型的非惯性系是相对于惯性系作直线加速运动的参照系和转动参考系。下面分别就通过具体的事例介绍这两类基本非惯性系中相应的惯性力的形式。

例 3-2-5 分析升降电梯在不同的运动过程中，人所感受到的超重和失重效应。

解：为了描述人在升降电梯所感受到的超重和失重效应，设想这个人站在固定在升降电梯地面上的一个秤上，通过从秤上读到的数据来描述人所感受到的超重和失重效应。由于相对于升降电梯，观测者的位置是固定的，决定他的运动状态演化的是受力平衡方程：

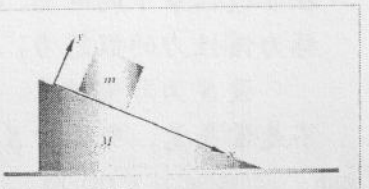
$$\mathbf{N} + m\mathbf{g} + \mathbf{F}_i = 0$$

其中惯性力 $\mathbf{F}_i = -m\mathbf{a}$ 与观测者所处的参考系的加速度方向相反。从秤上读到的数据是作用在观测者双脚上的弹力 $\mathbf{N}$ 的大小：

$$N = m|\mathbf{g} - \mathbf{a}|$$

若物体的加速度与重力方向相同，对应于处在减速向上或加速向下的运动状态，则 N 的读数减小，就会有失重感，如果物体的加速度与重力方向相反，对应于处在加速向上或减速向下的运动状态，则 N 的读数增加，就会有增重感。◀

例 3-2-6 如图所示，一个质量为 M 的三角物块放在水平地面上，在其斜面上放置了一个质量为 m 的立方体。若所有物体的接触面都是光滑的，求这两个物体的加速度。



例 3-2-5 图

解：立方体 m 沿斜面下滑，设斜面的加速度为 a_M，控制立方体 m 相对于斜面运动的演化方程的两个分量分别为

$$mg\sin\theta = m(a_m - a_M\cos\theta)$$

$$N - mg\cos\theta = -ma_M\sin\theta$$

其中$-ma_M$为立方体在斜面参照系所受到的惯性力。由于上述方程组涉及三个未知量：N, a_m, a_M。我们可以利用三角物体相对于地面的运动方程$N\sin\theta = Ma_M$解出三角物体的加速度：

$$a_M = \frac{m\sin\theta\cos\theta}{M + m\sin^2\theta}g.$$

由此可得立方体沿斜面运动的加速度

$$a_m = g\sin\theta + a_M\cos\theta = \left(1 + \frac{m\cos^2\theta}{M + m\sin^2\theta}\right)g\sin\theta = \frac{(M+m)\sin\theta}{M + m\sin^2\theta}g \quad \blacktriangleleft$$

下面讨论匀速转动参照系中的非惯性力。对处于惯性系$\boldsymbol{S}$中的观测者而言，静止在以角速率ω匀速转动的参考系$\boldsymbol{S}'$中的物体具有向心加速度，因而一定受到其它物体的作用力。若该物体位于过原点O且垂直于转动轴的平面上，而且与转轴相距r，该物体受到的力可以表示为

$$\mathbf{F} = -m\omega^2\mathbf{r}$$

为了在形式上用牛顿定律描述物体在该转动参考系内的运动规律，除了计入物体所受的真实力$\mathbf{F}$作用，还必须计入惯性力$\mathbf{F}_i$作用，$\mathbf{F}_i$正好与$\mathbf{F}$相抵消，即

$$\mathbf{F}_{eff} = \mathbf{F}_i + \mathbf{F} = 0$$

故 $$\mathbf{F}_i = \mathbf{F}_{eff} - \mathbf{F} = m\omega^2\mathbf{r} \tag{3-2-20}$$

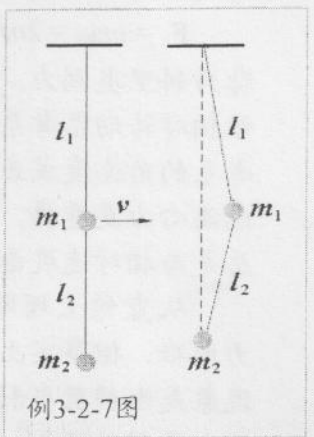
该结果也可直接从(3-2-19)式得到。因为虚拟力$\mathbf{F}_i$的方向沿半径向外，称为**惯性离心力**。若物体并不位于过原点O且垂直于转动轴的平面上，上述惯性离心力应写成

$$\mathbf{F}_i = -m\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}) \tag{3-2-21}$$

式中$\mathbf{r}$是物体的位矢，$\mathbf{F}_i$的方向[$-m\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r})$的方向]仍为离转轴向外。

需要注意的是，与平移惯性力不同，惯性离心力与物体所在位置有关。

例 3-2-7 一个双摆系统由摆长各为l_1与l_2的轻绳和质量各为m_1、m_2的两摆锤组成(见本例图)，当双摆处于平衡位置(竖直下垂)时，对摆锤m_1施加一个向右水平冲量使之具有速度v，求此时摆弦l_1与l_2中的张力T_1与T_2。



例3-2-7图

解：在该时刻摆弦l_1与l_2都处于竖直下垂状态，因而摆锤m_1虽然具有速度v，但尚未发生明显位移，而摆锤m_2则既没有位移也没有速度。因此，摆锤m_1处于绕悬挂点作圆周运动状态，具有向上的向心加速度v^2/l_1，其运动方程为

$$T_1 - T_2 - m_1g = m_1a_1 = m_1v^2/l_1$$

在地面参考系中，摆锤m_2没有位移也没有速度，但有向上的加速度(它在后一时刻将向上运动)，其运动方程为

$$T_2 - m_2g = m_2a_2$$

其中 a_2 可以由约束关系求得：设想经历很短的时间 Δt，m_1 将向右位移 $v\Delta t$，由于摆弦不可伸长，m_2 将向上位移 x_2，由例图不难看出：

$$x_2 = l_1 - \sqrt{l_1^2 - (v\Delta t)^2} + l_2 - \sqrt{l_2^2 - (v\Delta t)^2}$$

因 $v\Delta t \ll l_1, l_2$，将上式按级数展开，仅保留小量 $(v\Delta t/l_i)^2$ 的一次幂：

$$x_2 = \frac{(v\Delta t)^2}{2}\left(\frac{1}{l_1}+\frac{1}{l_2}\right)$$

由上式可得 m_2 的加速度：$a_2 = \dfrac{d^2x_2}{dt^2} = v^2\left(\dfrac{1}{l_1}+\dfrac{1}{l_2}\right)$。由此可解得

$$T_2 = m_2 g + m_2\frac{v^2}{l_1} + m_2\frac{v^2}{l_2}, \qquad T_1 = (m_1+m_2)g + (m_1+m_2)\frac{v^2}{l_1} + m_2\frac{v^2}{l_2}$$

如果在与 m_1 相对静止的平动参考系中考察，m_2 以 v 向左运动，绕 m_1 作圆周运动，其向心加速度为 v^2/l_2，方向向上。但与 m_1 相对静止的平动参考系是非惯性系，相对地面具有向上的加速度 v^2/l_1，故 m_2 除了受到真实力 T_2 和 m_2g 外，还受向下的惯性力 m_2v^2/l_1 的作用，m_2 的运动方程为

$$T_2 - m_2 g - m_2\frac{v^2}{l_1} = m_2\frac{v^2}{l_2}$$

显然，在与 m_1 相对静止的非惯性系中求解比在地球参考系中考察简便得多。◀

例 2-3-2 讨论了一个在匀速转动盘中沿半径方向以速度 v 作匀速运动的粒子所具有的实际加速度为

$$\mathbf{a}_i = \mathbf{a}_r + \mathbf{a}_\theta = -\omega^2\mathbf{r} - 2\mathbf{v}\times\boldsymbol{\omega}$$

在上式两边乘上粒子的惯性质量 m 可以得到匀速转动参照系中的非惯性力为

$$\mathbf{F}_i = -m\mathbf{a}_i = -m\mathbf{a}_r - m\mathbf{a}_\theta = m\omega^2\mathbf{r} + 2m\mathbf{v}\times\boldsymbol{\omega}$$

其中径向分量

$$\mathbf{F}_r = -m\mathbf{a}_r = m\omega^2\mathbf{r} \tag{3-2-22}$$

就是上面讨论过的**惯性离心力**，而横向分量

$$\mathbf{F}_c = m\mathbf{a}_\theta = 2m\mathbf{v}\times\boldsymbol{\omega} \tag{3-2-23}$$

称为**科里奥利力**。$\mathbf{F}_c$ 与相对速度成正比，只有当物体相对转动参考系运动时才会出现。又 $\mathbf{F}_c$ 与转动参考系的角速度成正比，角速度较小时，科里奥利力比离心力更重要；科里奥利力的另一特征是其方向总是与相对速度垂直，故不会改变相对速度的大小。

从直觉上理解科里奥利力通常远比理解离心力困难，但事实上，自然界存在着许多熟悉的自然现象是由科里奥利力所产生的效应。下面我们给出几个典型的事例。

第一个显示地球处于自转状态的在地球上装置是傅科摆。1851 年，法国物理学家傅科(Jean Bernard Foucault 1819-1868)在巴黎万神庙中演示摆

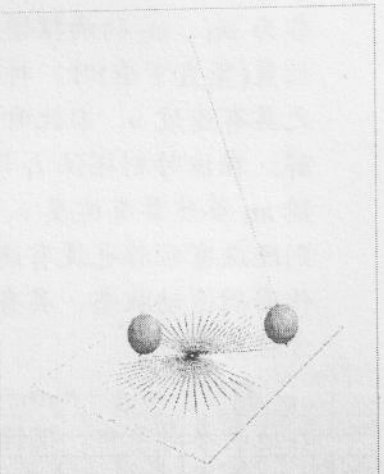


图 3-2-8 揭示地球自转效应的傅科摆。

在摆动时摆平面相对于地面旋转的现象(见图 3-2-8)。他所用的摆长为 67m，测得的摆动周期为 $T = 16.5$s，每一个摆动周期内摆面转动 0.05°，相应的转动周期为 32 小时。由于并不存在影响摆面运动状态的外力，这意味着摆平面的运动原因在于地面本身的自转。傅科摆摆锤在水平面上运动时因受到侧向的科里奥利力作用而使摆动平面旋转。这个力虽然很小，但随着持续的摆动，摆动平面的转动不断积累，从而显示出地球的自转现象。

第二个事例是低纬度区域，尤其是洋面上常年不变由副热带高气压带吹向赤道地区的定向风。因该风每年反复稳定出现而被称为"信风"[15]。信风的形成与大气在地球参考系中流动时受到的科里奥利力的作用有关。由于地球各处受到太阳的辐射并不均匀，其中赤道受热最多，赤道附近空气受热上升并向低纬度地区流动；在北半球高空，受科里奥利力影响，气流逐渐偏东，并在北纬附近形成西风。西风形成一堵墙，阻挡了南来的气流继续向北流动。空气在此堆积、冷却、下沉，形成地面上的副热带高压带。在地面上，空气由此处向南流动，补充赤道带上升的空气，受科里奥利力影响，逐渐右偏，形成稳定的东北信风。而在南半球则成为东南信风(见图 3-2-9)。

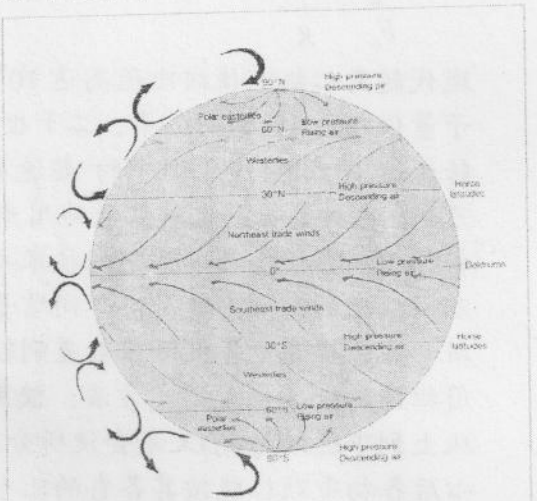


图 3-2-9 信风示意图。

惯性力虽然被称为赝力，但其产生的动力学效应却有许多重要的实际应用。例如，装载在船舰、飞机以及导弹上的惯性导航系统就是利用一种称为加速监测系统来测定自身的加速度的并调节自身的姿态以维持原有的运动状态的。简单的加速监测器主要部件是一个一端固定在参照系上的弹簧振子。当整个系统获得加速时，弹簧被另一端的重物所压缩或拉伸，而弹簧的压缩或拉伸数量和系统本身的加速度成正比。精密的加速监测器可以测得重力加速度大小百万分之一的精度。

惯性离心力的典型应用之一是离心器，由于离心力与物体的质量成正比，利用离心器可用来分离不同密度的各种物质。例如可以把鲜牛奶分离成奶油和较重的脱脂奶，也可以用于使血液中红血球快速沉降。图 3-2-10 为离心机的结构示意图。当相关的装置绕竖直轴高速旋转时，原处于竖直位置的盛有待分离溶液的试管因受离心力而趋近于水平位型。这时处于管内 P 点的溶质颗粒将受向外的离心力和向中心的浮力共同作用，设颗粒密度为 ρ_P，液体密度为 ρ_l，颗粒球的半径为 r_a，则颗粒所受向外合力 F 为

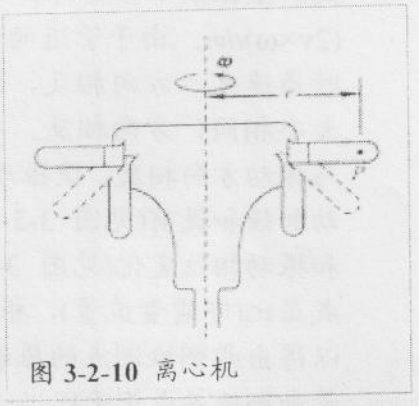


图 3-2-10 离心机

[15] Trade 是来自中古英语，相当于 path、track 的意思，trade wind 原意是"在常轨上的风"，西方古代商人们常借助信风往来于海上进行贸易，导致 Trade wind 有时被译成"贸易风"。

$$F_a = \frac{4}{3}\pi r_a^3(\rho_P - \rho_V)\omega^2 r$$

当 $\rho_P < \rho_V$ 时，F 指向中心，$\rho_P < \rho_V$ 时，F 背离中心。而在重力场中，溶质颗粒在溶液内受到的向下的合力为

$$F_o = \frac{4}{3}\pi r_a^3(\rho_P - \rho_V)g$$

两者之比为

$$\frac{F_a}{F_o} = \frac{\omega^2 r}{g}$$

现代超离心机可使该比值高达 10^4~10^6，在小范围内甚至可达 10^9，可用于分离原子量仅相差 1%的同位素。二十世纪二十年代，瑞典化学家斯维德伯格[16]发明了转速达 10,000 rev/s 以上的"超速离心器"，它产生的离心力比在地面的重力大 90 万倍。这种离心器已被广泛地用于分离并测定许多蛋白质的分子量，成为现代生物化学和分子生物学常用的技术之一。例如，通过交替使用低速和高速转动，用不同强度的离心力使具有不同质量的物质分级分离。这种称为差速离心的方法适用于混合样品中各沉降系数差别较大组分的分离。对于沉降系数接近的物质可采用所谓的密度梯度离心方法：使用一种能使密度形成梯度(在离心管中，其密度从上到下连续增高)又不会使所分离的生物活性物质凝聚或失活的溶剂系统，离心后各物质颗粒能按其各自的比重平衡在相应密度的溶剂中形成区带。常用的密度梯度溶剂是蔗糖或氯化铯溶液。

科里奥利力的一个重要的应用是一种被称为科里奥利质量流量计的计量仪器[17]，可用于测量通过管道的均匀密度的流体流量。科里奥利流量计有多种形式，图 3-2-11 所示为单直管科里奥利流量计的示意图。测量管被激振装置驱动进入振动状态，测量管两端设置有拾振传感器。管道中没有流体流过时，两端的振动状态相同(见图 3-2-11-a)。而在流体流过振动管道时各流体质量元 dm 受到和激振频率相关的科里奥利力：$d\mathbf{F}_c = -dm \cdot \boldsymbol{a}_c = (2\mathbf{v}\times\boldsymbol{\omega})dm$。由于管道的入口段和出口段段的瞬时角速度　方向相反，两端受到的科里奥利力大小相同，方向相反。其中入口段科里奥利力与振动方向相反，使振动减弱和滞后；出口段科里奥利力巧振动方向相同，使振动加强和提前(见图 3-2-11-b)。两种力的作用叠加在测量管上，使管子产生扭曲和振动相位变化(见图 3-2-11-c)。由于科里奥利力和流过管道的流体密度和流速成正比(即质量流量)，根据两端固定支承梁的振动方程和所受的科里奥利力，可以得出两侧检测点的振动信号相位差和质量流量成正比(对于线性系统，质量流量和相位差成正比)：

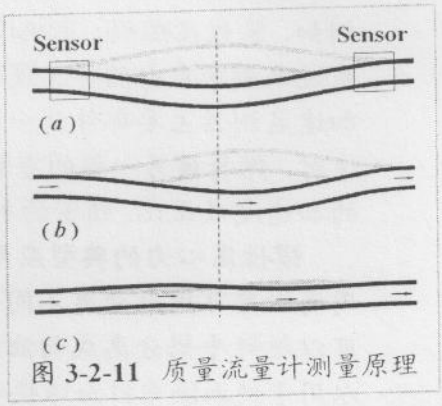


图 3-2-11　质量流量计测量原理

[16] Theodor Svedberg (1884-1971) 瑞典化学家，因其在研究悬浮物体方面，特别是在胶体粒子分散系统方面的成就而获 1926 年诺贝尔化学奖。

[17] The device was introduced in 1977 by Micro Motion Inc. The mass flow meters can be in various designs, but the basic configurations are of straight tube and curve tube.

$$\Delta\theta = (K/E)q_m$$

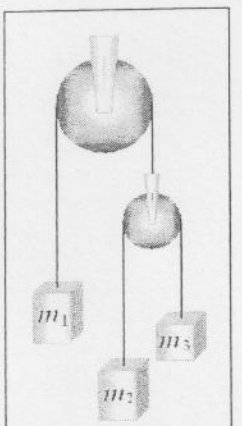
其中 K 为常数；E 为弹性模量；q_m 为质量流量。质量流量计的优点在于它能直接测量管道内流体的质量流量，而不必分别测量流体的体积流量和密度。

思考题

3-2-1 描述作用力的经验公式(3-2-1)可以表示为物理系统在一个特定运动状态附近的泰勒级数展开的前几项，如何理解其中的泰勒系数的物理意义，这种理解对认识作用力有什么帮助？给出具体的例子予以说明。

3-2-2 列举几个测量重力加速度的实验设计，并分析各自的测量误差。

3-2-3 说明利用滑轮组设计(见图)可以用来观测(恒)力和物体加速度之间的关系，并说明这种实验设计对上述关系可能的误差来源。



题 3-2-3 图

3-2-4 摩擦力是否总是和物体运动方向相反？静摩擦力是否总是使物体减速？

3-2-5 摩擦力与接触面积无关，为什么汽车用宽轮胎比用窄轮胎容易刹住车。

3-2-6 人奔跑时加速所需作用力是由鞋底和地面之间的摩擦力提供的，这是否意味着人获得的加速度和他自身的体力无关？

3-2-7 静摩擦力作用期间，物体既没有发生位移，也没有运动速度，静摩擦系数应该用哪些物理参量来描述？

3-2-8 能否用量纲分析、标度律等方法分析摩擦现象的各种可能原因，各种独立导致摩擦力的原因是否都存在相应的无量纲常数。

3-2-9 一个水桶底部装有阀门，阀门下放一只杯子接水。整个装置固定在一个磅秤的托盘上。打开阀门放水和关上阀门断水时，秤的读数可能会出现什么变化？

3-2-10 一艘火箭从北极向赤道飞行，月亮上的观测者看到地球在火箭下方转动，而地球上的观测者看到的是火箭向西偏转。说明地面上的观测者看到的由南极向赤道发射的火箭的情形是怎样的？

3-2-11 中国内陆的河流大多是由西北流向东南的，说明这些河流的东、西两岸受到河流冲刷的程度差异。

3-2-12 按照 3-2-7 节中惯性导航系统原理的介绍，画出简单的加速监测器的构造原理图。

3-2-13 说明为什么发动机安装在汽车前部的情形采用前轮驱动比较好。对于后轮驱动的汽车，为什么冬天在汽车尾部的行李箱中放入沙子会很有用？(5-6)

3-3 物体运动状态的演化

牛顿定律最常用于处理的一类问题是在已知物体的受力环境和初始运动状态的情况下，预报物体以后的运动行为。预测日食，月蚀，哈雷彗星的回归时间等天文现象，分析工程问题中各种机械构件在运行时可能经历的各种状态等都是这一类问题中的典型事例。求解这类问题也是一般力学课程的主题。求解过程一般涉及两个步骤：首先是确定物体在所处环境中实际受到的各种作用力，并列出所有这些力的数学表述(这种表述可以是经验公式，也可以是相应范围内的离散数值列表)，将这些作用在物体上的合力代入牛顿运动方程，使这个系统演化的动力学问题转变为运动学问题。简单而言，就是针对实际问题写出运动方程，然后求解相应的微分方程。本节主要通过一些具体例子讨论这两个问题。

3-3-1 系统的隔离与分割

自然界中的任何物体都不是孤立的，一个物体周围的所有物质(无论是否发生接触)原则上都对这个物体的运动存在影响。用牛顿运动方程求解一个物体运动的演化情况，需要将该物体从其周围的物体环境中分离出来，即将周围物体对它的影响抽象为数量化的具体作用力，用物体的受力环境替代它所处的实际物质环境，从而将物体在物质环境中的运动抽象为在作用力场中的运动。这种方法称为**隔离物体法**。

隔离物体法可直接用于处理简单物质系统的演化问题，这类问题或者仅仅涉及系统的整体运动，或者系统的各个部分在整个运动过程中的运动状态(位移和运动速度)都相同。由于这类系统的运动可以用一个质点的运动情况代表，作用在物体上的外力分布也相应地简化为全部的力作用在物体的同一个点上。

例 3-3-1 一个物体被以初速度 v_0，与地面成 θ 角的方向抛射到空中。物体在空中运动时受到的空气阻力与运动速度的大小成正比。求解物体的运动情况。

解： 由于问题只涉及到物体的整体运动，因而将它看作一个质点。该质点在运动过程中的力学环境只来源于两种物质：重力和空气阻力：$\mathbf{F}_G = -m\mathbf{g}$，$\mathbf{F}_d = -\alpha\mathbf{v}$。设定一个直角坐标系，其 x-轴沿水平方向，y-轴竖直向上，并使 x-y 平面包含初速度矢量，则抛射体的运动将限制在 x-y 平面内，其运动方程也相应地只有 x-分量和 y-分量：

$$F_x = -\alpha v_x = m\frac{dv_x}{dt} \qquad F_y = -g - \alpha v_y = m\frac{dv_y}{dt}$$

以上两个方程的解给出了抛射体的各速度分量随时间的变化关系

$$\int_{v_0\cos\theta}^{v_x}\frac{dv_x}{v_x} = -\frac{\alpha}{m}\int_0^t dt \qquad v_x = \frac{dx}{dt} = v_0\cos\theta\exp\left(-\frac{\alpha}{m}t\right)$$

$$\int_{v_0\sin\theta}^{v_y}\frac{\alpha}{mg+\alpha v_y}dv_y = -\frac{\alpha}{m}\int_0^t dt \qquad v_y = \frac{dy}{dt} = \left(v_0\sin\theta + \frac{m}{\alpha}g\right)\exp\left(-\frac{\alpha}{m}t\right) - \frac{m}{\alpha}g$$

进一步对时间积分可以得到抛射体的轨迹关于时间 t 的参数方程

$$x(t)=\int_0^t v_0\cos\theta\exp\left(-\frac{\alpha}{m}t\right)dt=\frac{mv_0\cos\theta}{\alpha}\left[1-\exp\left(-\frac{\alpha}{m}t\right)\right]$$

$$y(t)=\int_0^t\left\{\left(v_0\sin\theta+\frac{m}{\alpha}g\right)\exp\left(-\frac{\alpha}{m}t\right)-\frac{m}{\alpha}g\right\}dt$$

$$=\frac{m}{\alpha}\left(v_0\sin\theta+\frac{m}{\alpha}g\right)\left\{1-\exp\left(-\frac{\alpha}{m}t\right)\right\}-\frac{m}{\alpha}gt$$

如果不考虑空气阻力，上述方程就退化为

$$x(t)=\lim_{\alpha\to 0}\frac{mv_0\cos\theta}{\alpha}\left[1-\exp\left(-\frac{\alpha}{m}t\right)\right]=v_0t\cos\theta$$

$$y(t)=\lim_{\alpha\to 0}\frac{m}{\alpha}\left\{\left(v_0\sin\theta+\frac{m}{\alpha}g\right)\left[1-\exp\left(-\frac{\alpha}{m}t\right)\right]-gt\right\}=v_0t\sin\theta-\frac{g}{2}t^2$$

即为通常的抛物线方程。从上面结果可以求得射高和射程。◀

如果一个物质系统中的各个部分在运动过程中可能会处于不同的运动状态，隔离物体法就不能直接用于讨论整个系统的运动演化问题。处理这类问题的常用方法是将系统分割为若干个子系统，每个子系统中的各个部分在整个过程(或者某一短暂的时间间隔)中的运动状态大致相同而可以看作质点，然后将该子系统与整个系统的其它部分以及系统外界的物体隔离，即将它们全部处理为外部力学环境后通过求解各子系统所满足的运动方程得到系统各部分的运动演化情况。

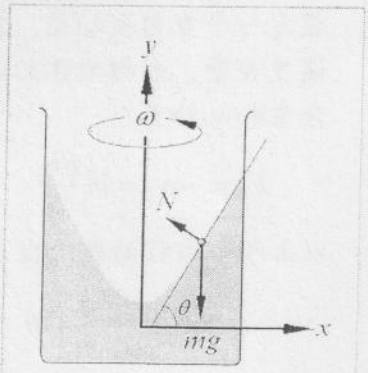
例3-3-2 装有液体的容器绕其对称轴以角速度ω旋转，其中的液体表面会呈现出如图所示的形状。利用牛顿运动方程求解描述液面形状的数学方程。

解：虽然液面形状涉及到全部液体的运动，但液体表面的形状却只需要考虑液面上某一微小液滴的行为。只要确定该液体微粒的位置与运动参量的关系就能确定整个液面的形状。

从系统的旋转对称性可以理解液面应该是一个绕转轴(设为y-轴)的回转面，我们所需要确定的只是包含该转轴的任一平面与该回转面的交线轨迹。从容器外的惯性系中观测液体的转动，液滴所受的力只是重力mg和液面的支撑力N。这两个力的合力维系着液滴在一个固定的高度和离轴间距做圆周运动：

$$N\cos\theta-mg=0$$

$$N\sin\theta=mx\omega^2$$



例 3-3-2 图

由上述方程中的几何参量θ可以得到液滴所在位置的曲线斜率：

$$\frac{dy}{dx}=\tan\theta=\frac{\omega^2}{g}x$$

通过积分可解出液面形状方程

$$y=y_0+\frac{\omega^2}{2g}x^2$$

因此液面形状是一旋转抛物面。该问题也可以在旋转

容器参照系中求解：在旋转参照系中液滴除了重力和液面的支撑力，还受到与非惯性系相关的惯性离心力。相应的运动方程为：

$$N\cos\theta - mg = 0$$

$$N\sin\theta - mx\omega^2 = 0$$

这组方程和在惯性系中得到的方程组在数学形式上完全相同。通过测量液面的高度差可确定系统的旋转角速度。◀

例 3-3-3 一条长为 l，质量为 m 的柔软绳索自由下垂，下端正好与地面接触。如果由静止状态开始释放绳索，当绳索上端降到高度 x 时，求绳索对地面的作用力。

解：这是绳索在重力 mg 和地面弹力 N 两个外力作用下整体动量不断变化的过程。由于体系的动量实际上都由绳索在空中下落的部分所携载，设某一时刻绳索上端下落到高度为 x 的位置时，空中下落部分的速度为 v，整个体系所具有的相应动量为：$p = \rho x v$，其中 $\rho = m/l$ 为绳索的质量密度。Δt 时间后，绳索上端下落到位置 $x - \Delta x$，整个体系动量的变化 Δp 由两部分构成：一是下落部分的绳索受到重力驱动而获得速度增量 $(\rho x)\Delta v$，二是绳索下端长度为 Δx 的微元因与地面接触而丧失其全部的动量 $(\rho\Delta x)v$。则 Δt 时间内的过程系统动量的变化总量为

$$\Delta p = -(\rho\Delta x)v + (\rho x)\Delta v$$

则系统的动量变化率为

$$\frac{dp}{dt} = -\rho v\frac{dx}{dt} + x\rho\frac{dv}{dt}$$

由于下落部分的绳索为自由落体运动，其速度增量仅来源于重力的单一驱动，因而有

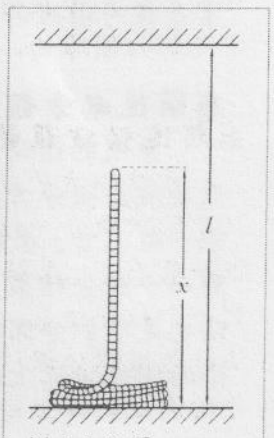
$$v = \sqrt{2g(l-x)}\ ,\quad dv/dt = g$$

代入上述运动方程可得

$$mg - N = -g(l-x)\rho + x\rho g = -2l\rho g + 3x\rho g$$

可得地面对绳索的弹力(即绳索对地面的作用力的大小)

$$N = 3(l-x)\rho g = 3\left(1-\frac{x}{l}\right)mg$$



例 3-3-3 图

说明：本题对系统的空间分割是随时间变化的，分割的要点是对过程的时间进行离散化处理。由于系统各部分的运动状态不同，各部分的受力情况也无法直接给出，对于这种情形应该考虑用系统的动量变化情况来反映系统的受力情况。◀

3-3-2 运动方程的求解程式

一旦所考虑的质点与其周围的物体的相互关联已经被表示为具体力的作用，所需要处理的系统运动问题也就抽象为求解一个具体的二阶微分方程：

$$m\frac{d^2\mathbf{r}}{dt^2} = \mathbf{F}(\dot{\mathbf{r}},\mathbf{r},t) \tag{3-3-1}$$

其中的力函数所描述的是最一般的情况，即物体所受的力可以是物体本身的速度 $\dot{\mathbf{r}}$、位置 $\mathbf{r}$ 和时间的 t 的函数。例如，一个作用于物体的周期性的驱动力则是时间的显函数；弹簧力依赖于弹簧的形变程度，因而是位置的函数；流体阻力起因

于物体和流体的相对运动，因而是速度的函数，而在转动系统中物体所受到的科里奥利力则是速度和位置的函数。

如果(3-3-1)式中的合力的表达式中只是位置，速度的线性函数：

$$F(\dot{r},r,t)=a\dot{r}+br+f(t) \tag{3-3-2}$$

则运动方程就是所谓的线性常微分方程。摩擦力，重力是常量，弹力，低速流体阻力都属于这一类线性函数。在给定的初始条件下，这类系统的演化过程是唯一的，而且由标准的求解模式。这是线性常微分方程理论的所谓"解的存在和唯一性定理"所决定的。

如果力的表达式中含有位置或速度的非线性项，例如高速运动物体所受到的流体阻力 $F=-av^2$，单摆摆锤受到的合力 $F=-l_o\sin\theta$，相应的运动方程为非线性常微分方程。非线性常微分方程不存在所谓"解的存在和唯一性定理"，也没有统一的求解模式。

物体的运动方程通常都是非线性的，但在非线性项很小的情况下，一般可以进行线性化近似，再通过"摄动"法对线性近似得到的结果进行修正。这样做的理由是如果非线性项的数值很小，可以预期，很小的非线性扰动并不会改变系统的总体行为，即系统的实际行为一般不会偏离线性解很远，这有助于我们对物体运动的总体图像获得基本的定性理解。而通过解析方法严格求解一个非线性问题通常是极为困难的。

求解运动方程的解析程式

系统的运动方程原则上确定了系统从任何初始状态出发随时间的演化行为。只要解出微分方程(3-3-1)就可知道系统的所有运动参量 r, v 在任何时间的数值。但一般而言，求解方程并不是一件轻而易举的工作。常规教科书中的做法是寻找到一个积分结果，它通常是一个初等函数或若干个初等函数的组合。这种求解方法也是微积分课程中的主要内容之一，称为微分方程的解析解法。我们上面所讨论的例题实际上都是如此。

求解运动方程的数值方法

物理学教科书中所列举的例题或练习题大多是可以解析求解的，这种安排大多是处于教学考虑，但它可能给学生读者带来一种错误的印象：所有的问题都是存在解析解的。而实际情况却恰恰相反，事实上只有很少的理想化问题才存在解析解。而只要稍稍偏离这些模型化的运动形式，解析求解就会出现很大的困难。例如，在空气中运动的物体只要形状有很小变化就可使其运动的求解变得非常困难。一个上抛物体(质点)的运动在初等微积分课程中和一般的物理课程中都是可以用初等函数表示其解的简单问题，而实际发射一枚火箭则完全是另一回事。空气阻力，重力以及风的影响都是随火箭的高度变化的。实际上必须采用就地调整动力推动的方法以补偿这些变化。因而在许多实际情况中，我们必须借助于数值方法而不是解析计算来预言系统的演化行为。

通用的数值算法可用于求解大多数的运动演化问题，这是因为运动过程被离散化为连贯的足够小的时间片段序列$\{\Delta t_i\}$，控制物体演化的参量—加速度需要历经若干个这样的时间片段才会呈现出显著的变化。这意味着在具体求解过程中可

以将每个时间片段中的加速度近似为常数，只是在进入下一时间片段的瞬间进行微量的调整得到新的加速度，通过对单一时间片段中的简单计算和重复迭代可以由时刻 t_0 时的初始条件给出 $\mathbf{r}(t_0)$ and $\mathbf{v}(t_0)$ 计算出任一时刻的运动状态参量。为了得到上述计算的数学描述，将物体在时间 Δt 前后的位置矢量 $\mathbf{r}_{old}$ 和 $\mathbf{r}_{new}$ 所定义物体速度表达式(图 3-3-1)：

$$\mathbf{v}_i \approx \frac{\Delta \mathbf{r}}{\Delta t} = \frac{\mathbf{r}_{new} - \mathbf{r}_{old}}{\Delta t} \tag{3-3-2}$$

改写为以下形式

$$\mathbf{r}_{new} \approx \mathbf{r}_{old} + \mathbf{v}_{old}\Delta t \tag{3-3-3}$$

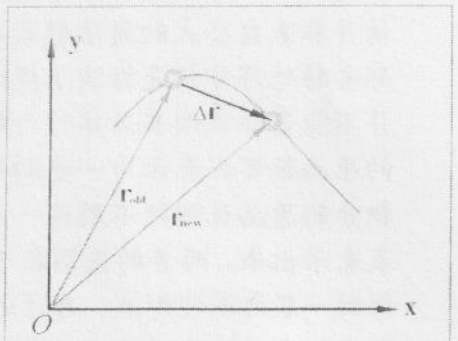


图 3-3-1

这就是利用时刻 t 物体的位置 $\mathbf{r}_i$ 和速度 $\mathbf{v}_i$ 来预言物体在下一时刻 $t+\Delta t$ 的位置 $\mathbf{r}_{i+1}$ 的方程。方程(3-3-1)可反复迭代以预测下一时刻及更下一时刻物体的位置。在这种逐步迭代的计算方法中所必需增添的新内容是不同时间段中的速度 $\mathbf{v}$，这一速度值可通过牛顿运动方程中作用力和物体的加速度关系直接获得：

$$\mathbf{v}_{new} = \mathbf{v}_{old} + \mathbf{a}_{old}\Delta t \tag{3-3-4}$$

$$\mathbf{a}_{new} = \mathbf{F}_{r,old}/m \tag{3-3-5}$$

上述求解过程可以程式化为一个数值迭代循环(见框图)。由大量的数据组(t_{old}, $\mathbf{r}_{old}$, $\mathbf{v}_{old}$)构成了描述所要求解的运动的离散函数。实际运行这一程序时要将运动方程中的所有矢量写出分量形式，而开始运行这一循环还需要输入物体初始时刻的位置和速度(即运动方程的初始条件)：

LET $\Delta t = h$
LET $t_{old} = t_0$
LET $\mathbf{r}_{old} = \mathbf{r}_0$
LET $\mathbf{v}_{old} = \mathbf{v}_0$

```
LOOP
  LET  a = F/m
  LET  r_new = r_old + v_old × Δt
  LET  v_new = v_old + a × Δt
  LET  t_new = t_old + Δt
  LET  t_old = t_new
  LET  r_old = r_new
  LET  v_old = v_new
  SAVE  t_old, r_old, v_old
LOOP until t_old > t_end
END
```

上述算法原则上可以求解出大多数运动的演化过程。其中 $\Delta t = h$ 是计算时实际采用的时间步长，这一般需要综合考虑计算时间和计算精度进行设置。

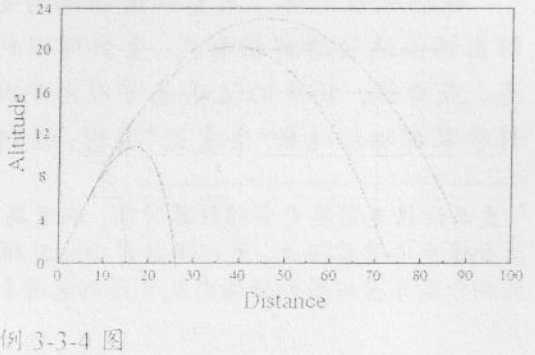
例 3-3-4 利用数值方法重新求解例 3-3-1 的斜抛物体的运动情况。

解： 对数值方法需要先确定控制质点运动演化的参量—加速度，由牛顿运动方程 $\mathbf{F}_G = \mathbf{F}_G + \mathbf{F}_d = -m\mathbf{g} - \alpha\mathbf{v}$。

其中 $\mathbf{v}$ 是质点的瞬时速度，加速度为

$$\mathbf{a} = \mathbf{g} + \mathbf{a}_{air} = \mathbf{g} - k\mathbf{v}$$

其中参量 $k = \alpha/m$ 表示单位质量的物体受到的空气阻尼系数。具体计算时需要对方程中的参量进行无量纲化处理。这里可以用重力加速度作为加速度的量度单位，将上述表达式的 x-分量和 y-分量表示为



例 3-3-4 图

LET $a_x = -(\alpha/mg)^* v_x$

LET $a_y = -1 - (\alpha/mg)^* v_y$

例图给出不同的阻尼参量情形的计算结果，其中选取时间步长为 0.1 秒。◀

习惯于欣赏解析解简洁优美形式的读者可能会不满意数值解那种离散而且庞大的丑陋形式。但这其实是一种偏见。实际上我们追逐解析函数优美形式的热情并非来自公式的简洁形式本身，而是来自它所对应的曲线图像。解析公式中的那些特殊符号只是扮演了提示图像类型的角色。事实上解析函数表示的解的本身并不能带给我们其具体行为的具体印象。所谓的解析解只是运动方程的不定积分的原函数可以表示为一些初等函数的组合。而不存在解析解的运动方程，其不定积分的原函数则做不到这一点。而数值解是将定积分的值在一定的精度内通过列表表示出来。两者的差别在于解析解可以被表达为我们所熟悉的函数形式，即先得到一般的演化形式，再通过初始条件求得各特定时刻的运动参量数值，而数值解只能由初始条件出发，直接得到各特定时刻的运动参量数值。用更简洁的数学语言：解析方法给出了形式上的通解，而数值方法得到的是一个特解。

数值方法求解运动方程中的稳定性问题

运动方程的数值解虽然不再受到力的具体函数形式的局限，但数值方法对运动的预言总是存在着某些不精确性。例如，因为采用有限大小的步长使得原来的微分方程被近似为差分方程，这就可能导致数值解出现精度不够和不稳定的问题[1]。数值解的精度除了与积分步长的大小有关，也与具体的算法有关(一般的微积分课程中所学过的定积分的矩形求和与抛物线求和法则就是不同的算法)。另外，一种数值方法的优劣常常还和具体求解的问题有关。在有些演化方程中，初始值误差或离散化所导致的误差在积分过程中会逐渐增加，由多步计算后得到的结果可能严重偏离方程的真实解，即预报的误差随着积分不断增长，这种解和相应的数值方法被称为是不稳定的。

使用可靠而且有效的算法来求解运动方程是非常重要的。在数值计算领域，目前已有许多的求解微分方程的算法，我们通常可以直接应用这些程序软件(在许多软件库中都作为基本配置提供)来求解运动方程。但需要注意的是，不同的算法对不同方程的适用性并不相同。在本节的附录中有一些在许多场合都可直接应用求解这类问题的简单程序。

3-3-3 确定性系统的内在随机性：演化行为对初始条件的敏感性

根据牛顿定律，只要知道物体的受力情况和它的初始运动状态，即使不能求解出物体运动的解析形式，至少可以利用数值计算法逐点求出物体以后的运动状态。或者说，物体的运动是可以由牛顿运动方程和初始条件所唯一确定的。许多科学家都相信这种"决定论"思想：任何复杂的自然现象的演化也都可以用一组方

[1] 更高的精度需要更多的计算时间，而更高速则得到不太精确的结果。今天计算机的计算能力和速度已非常强大，可以用非常小的时间步长来得到精度很高的结果，经过成千上百万步的积分运算在极短的时间内就可预测遥远未来的行为。

程式和初始条件唯一地确定。法国著名科学家拉普拉斯[2]就曾说过："设想有一位智者，...他能将宇宙间最庞大物体和最轻微原子的运动凝聚到一个公式之中，对他来说没有什么事情是不确定的，未来就像过去一样展现在他的眼前。"

这一概念原则上并没有什么问题，但求解实际问题时却远没有如此简单。因为实际求解时总是要忽略一些次要因素，即使在输入参数时也会引入某些误差[3]。运动方程对物质演化具有全部"决定性"的思想应用到实际的自然界意味着自然界的演化对上述忽略的次要因素和在输入参数时引入某些误差并不敏感。例如，给定运动方程解的具体形式取决于初始条件，但通常我们还会下意识地认为，如果初始值因各种原因带有微小误差，求解出的物体运动与真实运动的偏差也是很小的。或者说，初始值的误差不会随着系统的演化而不断放大。但在许多实际情况中，特别是对非线性的运动方程而言，初始条件的微小变化通常会导致方程所预测的结果出现很大变化。或者说，系统从观测上认为相同的(实际上是观测上无法区分的)初始条件出发，经长期演化后会得到面目全非的结果。

这类问题的一个典型例子是"预报气象"。事实上牛顿定律足以描述预测气象演化这类问题，但至今我们仍无法用解析或数值方法准确地预测明天的天气。1963 年气象学家洛仑兹为了预报天气变化，将大气动力学方程组简化为 12 个方程，并用计算机做数值模拟。图 3-3-2 为计算结果中的一个片段。其中的两条演化曲线对应于相同的演化方程和相同的初始值。洛仑兹只是在输入相同的初始值时分别采用了 5 位有效数或 3 位有效数，其相对误差小于千分之一。而图中的演化曲线表明：只是经过并不很大次数的迭代，两条演化曲线已经变得眉目全非。这种运动演化对初始条件的敏感性通常被生动地称为"**蝴蝶效应**"：一只蝴蝶在日本扇动一下翅膀，可能会在墨西哥湾形成一场风暴。

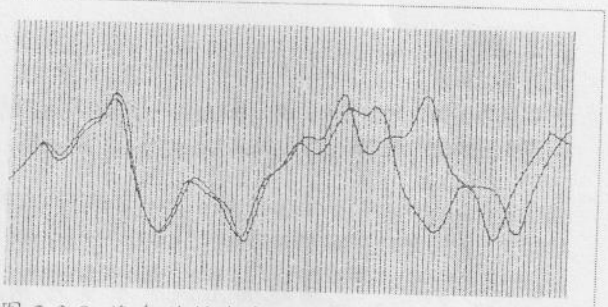


图 3-3-2 洛仑兹的大气动力学方程组数值模拟。非常相近的初始条件给出的模拟结果完全不同。

这类问题的另一个典型事例是一种简化的限制性三体问题：设想三个天体中一个星体的质量比另两个星体质量小很多，因而可以认为两个大质量星体的运动不受小质量星体的影响，其运动可由牛顿定律精确地决定，事实上小星体的运动极具复杂性和不稳定性。19 世纪末，法国数学家庞加莱[4]就认识到，任何微小的扰动都会使小星体的轨道在一段时间后发生显著的偏离。读者可以自己通过计算机模拟来体验小星体的这种复杂且不稳定的运动行为。

原本应该完全确定的运动因为对初始条件的敏感性而变成完全不可预测的

---

[2] 拉普拉斯 (Pierre-Simon Marquis de Laplace 1749-1827)，法国著名数学家和天文学家，天体力学的主要奠基人，天体演化学的创立者之一，分析概率论的创始人，应用数学的先驱。

[3] 例如，对各种实数在具体输入或计算时时只能取有限的位数。

[4] 庞加莱 (Jules Henri Poincaré 1854-1912) 法国数学家，理论物理学家，工程师，科学哲学家，被公认为是继高斯之后对于数学及其应用具有全面知识的最后一个人。他对数学，数学物理，和天体力学做出了很多创造性的基础性的贡献。他提出的庞加莱猜想是数学中最著名的问题之一。他对三体问题的研究中发现混沌现象，并为现代的混沌理论打下了基础。在狭义相对论的研究中，他比爱因斯坦的工作更早一步发表了一个的简略版。

和随机的，这导致不能把一系列短期预言串在一起得到一个长期有效的预报。源于牛顿定律自身的运动对于初始条件的敏感性揭示了我们对运动规律以及由此产生的演化行为之间(即原因与结果之间)关系的一个错误认识。我们曾认为确定性的原因必定产生规则的结果，但实际上可以产生(容易被误解为随机性的)极不规则的结果。我们也曾像当然地认为简单的原因必定产生简单的结果(这意味着复杂的结果必然有复杂的原因)，而实际上简单的原因可以产生复杂的结果。上述讨论说明，知道了系统的演化规律并不意味着一定能够预言未来的行为。

需要说明的是，初始测量的误差会随着系统的演化而放大是与系统行为的高度不规则性密切相关。对初始条件敏感性的确定性系统被称为**混沌**系统。但混沌系统行为也具有其确定性的规律，即混沌所表现出的复杂、无规律的行为背后仍然具有简单的、确定性的秩序。

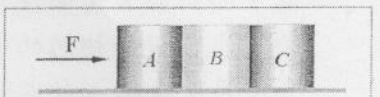
思考题

3-3-1 冰球在冰上的匀速滑动是否具有时间反演不变性？汽车在马路上的匀速行驶以及伞兵在空中匀速下降的过程是否具有时间反演不变性？

3-3-2 流体阻力与物体的运动速度有关，因而不是时间反演不变的。科里奥利力 $\mathbf{f}_C = m\mathbf{v}\times\boldsymbol{\omega}$ 也与速度有关，是否具有时间反演不变性？

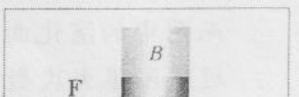
习题

3-3-1 三个质量分别为 m_A、m_B 和 m_C 的物体在光滑平面排成如图所示的一列。若以力 F 推最左面的物体 m_A，计算作用在中间的物体 m_B 所受到的作用力。



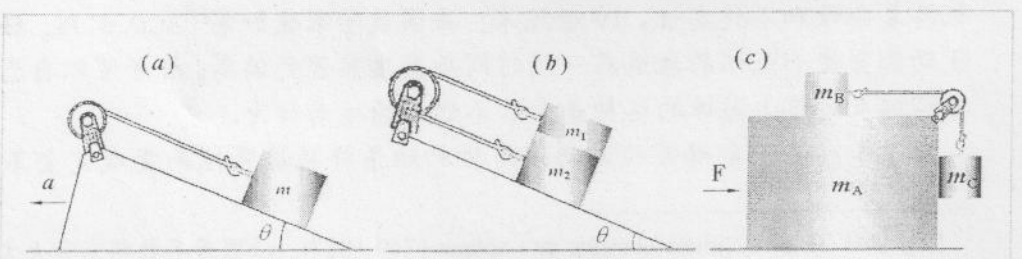
题 3-3-1 图

3-3-2 如图，光滑地面上放有质量为 m_A 的物体 A，A 上物体 B 的质量为 m_B。A 与 B 之间的摩擦系数为 μ_1。用水平力 F 推 A，至少需要多大的 F 才能使 A、B 之间发生相对运动。若 A 与地面间的摩擦系数为 μ_2，F 多大时才可将 A 从物体 B 下抽出？



题 3-3-2 图

3-3-3 质量为 m 的物体用细线拴在倾角为 θ 的光滑斜面上(如图)。若斜面以加速度 a 向左运动，求细线中的张力及物体对斜面的正压力；加速度 a 多大时，物体开始离开斜面？



题图：(a) 题 3-3-3 图，(b) 题 3-3-4 图，(c) 题 3-3-5 图

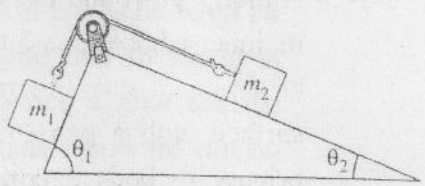
3-3-4 在本题图所示装置中，m_1 与 m_2 及 m_2 与斜面之间的摩擦系数都为 μ，设 $m_1 > m_2$。不考虑滑轮和线的质量及轴承的摩擦以及线的形变。缓慢增大斜面的倾角 θ，θ 达到多大时 m_1、m_2 开始运动。

3-3-5 质量分别为 m_A，m_B 和 m_C 的物块如图放置，设整个系统各处都不存在摩擦现象。水平力 F 多大时能使 m_3 维持其高度不变。

3-3-6 A block of mass 0.70 kg rests on top of another block of mass 0.90 kg, which rests on a frictionless surface. The coefficient of static friction between the blocks is $\mu_s = 0.45$. What is the maximum horizontal force that can be applied to the upper block so that the blocks accelerate together without the sliding between the two blocks? If the horizontal force is applied to the lower block instead, what is the maximum force that will give rise to the same motion?

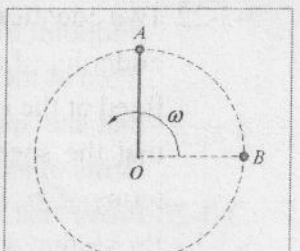
3-3-7 A mass m_1 rests on top of another mass m_2 which in turn rests on a frictionless horizontal surface and is pulled by a force F. Find the acceleration of each block when (a) the surface between them is frictionless. (b) the upper object is sliding on the lower object with kinetic friction coefficient

3-3-8 Two blocks of mass m_1 and m_2 ($m_1 > m_2$) on the two inclines of a triangle block are connected by a light string passing over a light frictionless pulley (see Figure). The coefficient of kinetic friction between the blocks and the planes is μ_k. Find the acceleration, velocity, and displacement of each mass as a function of time. Assume that m_2 initially moves downward at a speed of v_o.



3-3-9 Two blocks A and B of masses 5 and 10 kg are connected by a light spring. A force of 12 N acts on B directly away from A. At a particular instant block A has an acceleration of 3 m/s^2 towards B. What is the acceleration of B at that moment?

3-3-10 如图所示，一长为 l 的细棒绕其一端点 O 在竖直平面内转动，棒的另一端 A 固定有质量为 m 的质点。该质点速率多大才能使棒对它的作用力为 0？设 m = 0.5kg，l = 0.5m，质点速度 v = 0.4m/s，求质点在 B 点时棒对它的切向和法向的作用力。



题 3-3-10 图

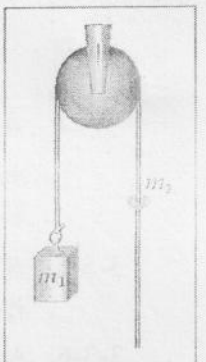
3-3-11 一辆汽车驶入曲率半径为 R = 150 m 的弯道。弯道地面与水平面成倾角 θ = 18°，轮胎与路面间的摩擦系数为 μ = 0.3。求汽车在路面上不作侧向滑动时的最大和最小速率。(5-51)

3-3-12 一劲度系数为 k 的轻弹簧的一端固定在顶角为 2θ 的光滑圆锥面的顶点上，另一端系有一质点 m，绕锥面的轴线旋转。计算质点开始离开锥面时的角速度 ω，以及此时弹簧的伸长量。

3-3-13 一个人质量为 80 kg 顺着一条绳索往下滑，若绳索的最大抗张强度为 600 N，5.0 秒内他最多可以下滑多少距离？

3-3-14 滑轮上跨套有一条不可变形的轻绳，绳索与滑轮间的摩擦可忽略。绳索一

端系有质量为 m_1 的物体，滑轮另一侧的绳索上套有一个质量为 m_2 的小环。小环相对绳索以加速度 a 下落(见图)。求小环与绳索之间的摩擦力。



题 3-3-15 图

3-3-15 一条抛物线形($y = ax^2$)的光滑弯管上套有一小环，弯管绕铅直轴以均匀角速率 ω 转动。ω 多大时小环可在管上的任意位置相对弯管静止。若弯管为圆形，ω 多大时小环可在管上的任意位置相对弯管静止。

3-3-16 一列车在北纬30°处自南向北以90km/h的速率匀速行驶，若车厢重 50 吨，计算铁轨受车轮的旁压力的方向及大小。

3-3-17 在北纬 30°处，计算由 10m 高处落到地面的物体由于科里奥利力的作用而引起的落点水平偏移量。

3-3-18 一个圆盘绕自身的对称轴以角速度 ω 逆时针水平转动，一个质点相对于地面以恒定速度 v 在圆盘面上沿一过圆心的直线运动。求解质点在圆盘上留下的运动轨迹。

3-3-19 讨论在北纬 45°处的一个傅科摆的摆平面的运动情况：运动周期。摆锤投影在地面上的运动轨迹(讨论不同的初始条件)。

3-3-20 A merry-go-round has a circular platform that is 1 m from the central axis at its inner edge and is 5 m from the central axis at its outer edge. The ride turns at a rate of one full rotation every 10 s. A passenger holds himself to the surface with a pair of very sticky shoes and is most comfortable when he orients his body length along the line of the net force on him. Determine the angle his body makes to the vertical (a) 1 m from the axis, (b) 3 m from the axis, and (c) 5 m from the axis.

3-3-21 A small mass slides without friction in a horizontal circular path around the sides of a circular bowl. The bottom of the bowl may be described as a parabola, with the height h above the bottom varying quadratically with the distance r from the axis: $h = br^2$. The mass is observed to move in its circular path with a speed v What is the height of the path?

3-3-22 Two identical objects, each with a mass m, are attached at one end and the midpoint of a string respectively. The string has length of L and its other end is fixed at the centre of the frictionless tabletop. The two are whirled in such way that the speed of the outer mass is v. Calculate the tensions of the string in terms of m, L, and v.

3-3-23 Prove that a horizontal circle on the inside of a slippery conical surface for a small block slides in it is stable.

3-3-24 A block of mass 1.00 kg hangs from a rope placed through a hole in a smooth, frictionless table. At the other end of the rope is attached a puck of mass 400 g. The puck swings in a circular orbit around the hole. With what speed must the puck rotate to keep it 1 m from the hole? Find the tension in the cord. Assume that the length of the cord is 2 m and its mass density is λ.

3-3-25 Two light strings, each of length L, are tethered to a vertical axis with a

separation of L, and a block of mass m at the end of the two strings is whirled about the vertical axis (see Fig.) at a rotational speed , and both strings are taut, so that they and the vertical support form a regular triangle. Calculate the tensions in the two strings.

3-3-26 The road in Example 5–16 is now in better condition, and the coefficient of static friction between the tires and the road is 0.8. With what maximum constant speed can the motorist now negotiate the curve? Friction acts here both to provide a centripetal force and to help oppose the vertical forces. Is there one and only one speed with which the curve can be taken? Why or why not?

3-3-27 A small puck of mass m = 0.1kg moves in a circle of radius 0.3 m on a table top; the puck is tied with a massless string to a tether at the origin. The coefficient of kinetic friction between the puck and the table top is = 0.25. At t = 0 the puck has a velocity in the +y-direction of magnitude 8m/s. What is the tension in the string at t = 0? And what is the tension in the string at the end of one revolution? (5-69)

3-3-28 A pendulum hangs at rest from a hook in a ceiling of a building. The building is located at a latitude such that the radius vector from the centre of Earth to the building makes an angle θ with Earth's axis of rotation. Assume that the force of gravity is the same everywhere and points directly to the center of Earth. What is the angle that the pendulum makes with the vertical because of Earth's rotation? (By definition, Earth rotates about its axis once a day.) (5-70)

3-3-29 两个质量均为 m 的质点，由长度为 $2l$ 的轻绳连接，一恒力 **F** 持续地垂直作用于细绳的中点($x = 0$)处，证明每一个质点在与作用力 **F** 垂直的方向上的加速度的大小为

$$a_x = \frac{F}{2m}\frac{x}{\sqrt{l^2 - x^2}}$$

式中 x 是质点与 **F** 的垂直距离。讨论当 $x = l$ 处的情形。

3-3-30 升降机中一斜面与地板成 θ 角。质量为 m 的立方块从斜面滑下。求以下情形中立方块相对于斜面的加速度：升降机(a)匀速上升、(b)匀速下降、(c)匀加速上升、(d)匀减速下降、(e)缆绳断裂。(f)求在升降机匀加速上升时斜面对立方块的力。

3-3-31 质量为 m 的小立方块置于旋转漏斗内壁(见图)。漏斗的角速度为 ω。设漏斗与水平方向的夹角为 θ，立方块与漏斗表面间的摩擦系数为 μ。求使小立方块不滑动的最大转速 v_{max} 和最小转速 v_{min}。

3-3-32 A hemispherical bowl is placed open end up on a table and rotates around its own vertical axis. A die is allowed to slide down from the edge into the bowl. Describe the motion of the die as seen by a hypothetical observer at the bottom of the bowl, assuming that there is no friction between the bowl and the die.

3-3-33 设一辆四轮汽车的质量为 1000 kg，所受重力均匀分布到四个轮子上，轮子与路面的静摩擦系数为 $\mu_s = 0.8$。如果汽车在水平面上从静止启动，求解

最大的向前加速度 is the greatest forward acceleration that it can have without spinning its wheels?

3-3-34 线密度为 λ 的柔软绳索盘放在水平桌面上。将绳索的一端以恒定速率 v 向上提起，求提到高度为 x 时作用在绳索上端的提拉力。

3-3-35 长为 L、线密度为 λ 的柔软绳索，绳索两端 A, B 合一起，悬挂在支点上，设想 B 端脱离支点自由下落，计算 B 端下落距离为 x 时支点所受的力。

3-3-36 一根质量均匀的绳子一端固定，求绳子在光滑桌面上作圆周运动时的形状。

3-3-37 假设雨滴在下落过程中各点所受到的空气阻力与速度的大小成正比，求雨滴的形状。

3-3-38 估计一个以一定角速度旋转的液滴的形状。

3-3-39 估计一个以一定角速度公转的液滴的形状？与轨道半径的关系。

数值计算

3-3-1 求存在空气阻力时抛射体的射高和射程。当抛体速度较大时，阻力将与速度的平方成正比，求解这类问题。

3-3-2 设一单摆的摆线具有一定的弹性，摆线原长为 L，弹性系数为 k，摆锤质量为 m。将摆锤从一个特定角度of a spring to the ceiling, and a weight on the other end, pull the ball back off to one side, and let go. The behavior of the ball looping around cannot be analyzed with calculus – there is no analytic solution for this motion, but simple to predict by using graphical methods with the aid of computer. The forces simultaneously acting on the weight are the downward gravity $F_G = mg$, and the spring force $F_S = -k(l - l_o)$ pointing the fixed end of the spring, 。利用数值方法求解该系统的演化，并根据摆锤的轨迹讨论系统的演化特征。

3-3-3 设想三个天体的质量分别为 m_1，m_2 和 m，其中 m_1，$m_2 \gg m$，因而可以认为两个大质量星体的运动不受小质量星体的影响。设两个大天体(近似为质点)位置不变，相距 L，天体之间的引力为中心力

$$F = -G\frac{mM}{r^2}$$

读者可以自己设定小天体的初始条件，通过计算机模拟体验小星体的运动演化行为的复杂性和不稳定性。

3-4 功与能

3-1-3 节中探讨物体运动量的函数形式 $f(m,\mathbf{v})$时，通过两个物体彼此相互作用并发生反冲的实验现象，归纳出参量 $\mathbf{p} = m\mathbf{v}$ 可用于描述物体的运动数量。但由此也引出了一系列的问题：$m\mathbf{v}$ 是否是描述和量度物体运动量的唯一函数形式？如果存在其它函数形式，从逻辑上也可用利用它的变化情况定义和量度物体受到的作用力，这些从不同的测量属性定义的作用力是否彼此自洽？

3-4-1 功和动能

上述这些问题显然不可能通过逻辑论证得到解答，只能通过恰当的实验设计，在具体的物理过程中呈现出来。而回顾曾经做过的相关实验的各种细节通常能获得如何开始新的实验探索的一些启发。具体而言，由 3-1-2 节中利用反冲实验过程揭示出动量定义所涉及到运动量与作用力的两个要点：

(1) 反映物体运动量的参量必须在受力过程中表现出可观察的变化，而且这种变化表现为“力”对物体持续作用的累积效应。在观测和量度动量变化累积数量 $\Delta\mathbf{p}$ 的累积参量是时间；相应的作用量参量是冲量 $\mathbf{F}\Delta t$。

(2) 两个物体之间的(内部)相互作用会改变其中每一个物体的运动数量，但不能改变这两个物体的运动数量总和。

上述两点表明，“力”并没有在作用过程中生成或消耗物体的运动数量，它只是在作用过程中将运动数量不断地从一个物体转移到另一个物体。因此，在相互作用过程中的每一时刻，系统整体的运动数量都是守恒的。

如果上述要点对于运动量与作用力的概念是普适的，可以由此设想，如果物体的运动量存在其它函数形式，物体在“力”的作用下应该能呈现出其它形式的可观测效应，而“力”在作用过程中只是作为一个传递运动量的经纪人，物体受到“力”的作用前后，转移的运动量可以表现出不同的形式，但数量必须守恒，因而在恰当的实验设计中可以表现为运动量以不同的形式发生无休止的转换循环。

为了设计出能反映上述特点的实验装置以揭示物体运动量的其它函数形式，我们再仔细回顾反冲实验中的一些细节。反冲实验装置中包括两个部分，一是用于量度物体运动数量的两块物体，二是用来对两块物体实施相互作用的弹簧。而我们在定义“动量”时，更多地关注于物体本身的运动状态变化，而实际过程中弹簧的状态变化也呈现出了和两个物体反冲前后的运动状态变化之间的关联：反冲前弹簧处于压缩状态，物体静止；反冲后弹簧恢复到松弛状态，物体获得了运动量，而且弹簧的压缩程度和物体反冲后获得的运动量之间存在数量上的对应关系：弹簧压缩量越大，物体反冲后的运动数量也越大。这意味着在整个过程中，系统中存在另一种物理量维持不变(守恒)，它可以在系统中物体的运动数量和不依赖于物体运动状态的弹簧压缩参量之间进行等量转换。由于这两个参量具有不同的测量属性，这就为量度物体的运动数量提供了另一个视角：利用弹簧的压缩数量来量度物体的运动数量 $f(m,\mathbf{v})$，再通过两者相同的量纲求解 $f(m,\mathbf{v})$的函数形式。

反冲实验中与“力”相关，而且具有不同测量属性的物理量之间的对应关系也存在于伽利略的斜面实验中：物体从斜面下降失去了其相对于地面的高度，但获

得了速度，物体从斜面上升失去了速度，却由此获得了相应的高度，两者之间也明显存在着数量上的对应关系。这些实验都呈现出了“力”的相互作用特征：“力”在两个物体间作用的过程中扮演的只是一个经纪人的角色，它在整个作用过程中并不改变一个特定运动量的总量，只是通过系统状态的变化改变其表现形式，使我们能够通过测量属性的差异对其进行量度。

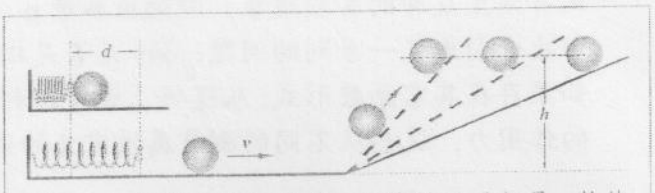


图 3-4-1 弹簧的压缩数量，物体的运动数量，物体相对于地面的数量三者之间的相互转换。

图 3-4-1 将上述两种实验所涉及的物理量联系在一起：一个质量为 m 的物体将一端固定在墙上的弹簧压缩一定的长度 d 后放开，物体在弹簧不断消耗它原有的压缩量的同时持续地获得速度增量。随着压缩量被完全释放，物体被加速到一个特定的速度 v 后进入匀速直线运动状态。而运动物体进入右面的斜面区域后，重力对物体的作用过程则以另一种物理量转移的迹象呈现出来：物体在攀升过程中不断消耗它的运动量，同时获得与高度有关的某种物理量，而在物体到达一定的高度 h 后耗尽其全部运动量。在这两个可量度的运动数量转移过程都是通过力持续作用在物体上累积形成可观测效应的，但在观测和量度其中运动数量 $f(m,\mathbf{v})$ 变化累积效果的参量是空间；相应的作用量参量是 $\mathbf{F}\cdot\Delta\mathbf{r}$，称为作用力对物体做的功。

如果让过程进行下去，则耗尽运动数量的物体将在重力的作用下滚下斜面而重新获得原来的运动量(不是原来的速度和动量，而是和原来反向的速度和动量)，当物体向左压缩弹簧并完全耗尽其运动量所达到的最大压缩量就是 d，这时系统又恢复到初始的状态。然后系统将周而复始地重复以上的运动。在这种循环往复的运动中，物体的运动量在弹簧的压缩或物体的高度的物理量之间不断地转换。而且在物体在“运动量”最大时获得的最大速度和它耗尽其“运动量”时抵达的高度或对弹簧的最大压缩量之间并没有任何变化。

上述实验结果表明，一个具有运动量的物体具有一种能力，它可以将一根弹簧压缩到一定的程度，也可以通过斜面攀升一定的高度。表征运动物体这种能力的运动量可以在力的作用下与弹簧压缩量或斜坡高度之间进行等量的变换。我们用函数 $K(m,\mathbf{v})$ 来量度这个运动量，这种能力转换为对弹簧的压缩，压缩数量显然与弹簧的压缩特性有关，而转换为物体相对于地面的高度时，其数量与地球的性质有关。即对于不处于运动状态的物体而言，这种能力与物体在环境中的位置以及其力学性质有关。我们用一个函数 $U(m,\gamma,\mathbf{r})$ 表示相关参量之间的关系。其中 $\mathbf{r}$ 为物体在系统中的位置，γ 为量度物体所处的环境力学性质的某个参数。显然，理想弹簧的特性参数只是其弹性系数 k，物体在系统中的位置可以表示为相对与弹簧平衡位置的距离，相应的函数为 $U_k(k,d)$。而对于物体攀升到距离地面高度 h 时，相关的特性参数是地球的重力加速度 $\mathbf{g}$ 和物体的质量，相应的函数为 $U_G(g,m,h)$。这些不同形式的物理量在转换过程中的任一时刻维持总量守恒

$$K(m,\mathbf{v})+U_k(k,x)+U_G(g,m,h)=\text{常数} \tag{3-4-1}$$

这三个具有不同函数形式的物理量描述的是同一种物理特性，因而具有相同的量纲，利用量纲分析求解出这些函数的具体形式。设上述三个物理量量纲分别为

$$[K]=[m]^{\alpha}[v]^{\beta},\quad [U_k]=[x]^{\beta'}[k]^{\gamma'},\quad [U_G]=[m]^{\alpha''}[h]^{\beta''}[g]^{\gamma''}$$

由于运动量应该和物质本身的数量成正比，我们取质量的量纲指数 $\alpha = 1$。由量纲平衡可得：

$$\alpha'' = 1,\ \beta = \beta' = 2,\ \beta'' = 1,\ \gamma' = \gamma'' = 1$$

由此得到 $K(m, \mathbf{v}) = amv^2$，$U_k(A, k) = bkx^2$，$U_G(m, h, g) = cmgh$。其中 a, b, c 都是无量纲常数，由式(3-4-1)可得这三个常数之间的关系：

$$amv_{\max}^{\ 2} = bkx_{\max}^{\ 2} = cmgh_{\max} \tag{3-4-2}$$

即这三个常数中只有一个是独立的，可以通过实验确定，也可以按上式得到的作用力的定义与牛顿定义的力在量值上的一致性来确定这个常数。

图 3-4-1 中重力和弹性力对物体作用所产生的可观测的物体运动量的变化 $\Delta K = amv^2$ 都表现为一种力的累积效应。这类似于物体受力作用所观测到的物体的动量变化 $\Delta\mathbf{p} = m\mathbf{v}$ 为时间累积效应 $\Delta\mathbf{I} = \mathbf{F}\Delta t$，有所不同的是这里观察到的是通过空间位置的变化量来呈现出其累积数量的(见思考题 3-4-1)

$$\Delta W = \mathbf{F}\cdot\Delta\mathbf{r} \tag{3-4-3}$$

称为物体在作用力 $\mathbf{F}$ 的方向上发生位移 $\Delta\mathbf{r}$ 的过程中力 $\mathbf{F}$ 对物体做的"**功**"[1]。

下面通过(3-4-3)式中定义的力与牛顿运动方程中对力的定义的一致性来确定 $K(m, \mathbf{v})$ 中的常数 a：

$$W = \int_{\mathbf{r}_1}^{\mathbf{r}_2} \mathbf{F}(\mathbf{r})\cdot d\mathbf{r} = \int_{\mathbf{r}_1}^{\mathbf{r}_2} m\frac{d\mathbf{v}}{dt}\cdot d\mathbf{r} = \int_{\mathbf{v}_1}^{\mathbf{v}_2} m\frac{d\mathbf{r}}{dt}\cdot d\mathbf{v} = \int_{\mathbf{v}_1}^{\mathbf{v}_2} m\mathbf{v}\cdot d\mathbf{v} = \frac{1}{2}mv_2^2 - \frac{1}{2}mv_1^2 \tag{3-4-4}$$

这就确定了物体的运动状态参量 $K(m, \mathbf{v})$ 的函数形式：

$$K = \tfrac{1}{2}m\mathbf{v}\cdot\mathbf{v} = \tfrac{1}{2}mv^2 \tag{3-4-5}$$

称为物体的运动能，通常简称为**动能**。将重力和弹性力的表达式分别代入(3-4-4)式，可以确定常数 $b = \frac{1}{2}$，$c = 1$。由此可得

$$U_k(m, x, k) = \tfrac{1}{2}kx^2 \tag{3-4-6}$$

$$U_G(m, h, g) = mgh \tag{3-4-7}$$

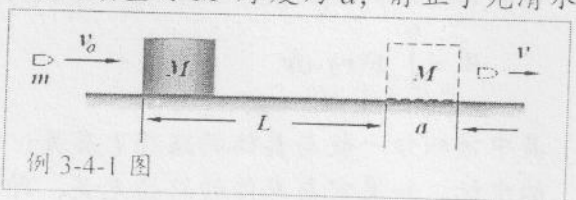
分别称为**弹性能**和**重力能**。显然，"功"所描述的是力将"能量"在系统的不同部分进行转移的过程，而在该过程中"功"充任的是一个簿记，跟踪能量从一个物体迁徙到另一个物体。

作用力对物体做"**功**"导致物体动能的改变所表述的因果关系(3-4-4)通常称为**动能定理**。力对物体在单位时间里所做的功称为**功率**，可以表示为

$$P = \frac{dW}{dt} = \mathbf{F}\cdot\frac{d\mathbf{r}}{dt} = \mathbf{F}\cdot\mathbf{v} = F_x v_x + F_y v_y + F_z v_z \tag{3-4-8}$$

由于功所反映的是力作用在物体上的空间累积效果，因而常用于分析不明显与时间有关的力学分析问题。

例 3-4-1 质量为 m 的子弹以速度 v_0 水平射入质量为 M 厚度为 a，静止于光滑水平面上的木块，然后以速度 v 从木块穿出(见图)。设子弹在木块中受到的阻力为常量。求从子弹进入木块到离开木块的整个过程中木块的位移量 L。



例 3-4-1 图

[1] 功定义为两个矢量的标积是因为相关的实验中没有表现出相关的方向属性。

解：木块是受子弹对木块的摩擦力驱动运动的，求解木块的行进距离需要知道摩擦力以及摩擦力作用时间或木块的末速度 u。由于子弹和木块作为一个系统不受外力作用，其水平方向动量守恒：$mv_o = Mu + mv$。由此可以求得木块的末速度：

$$u = (m/M)(v_o - v)$$

求解摩擦力一般是分析力产生的效果。我们从摩擦力做功的角度出发，设摩擦力 f 对木块作用距离为 L，由动能定理可得

$$fL = \tfrac{1}{2}Mu^2$$

摩擦力同时也作用在子弹上，但对子弹作用的距离为 $L+a$，由动能定理可得

$$-f(L+a) = \frac{m}{2}v^2 - \frac{m}{2}v_o^2$$

将上两式相加可得

$$-fa = \frac{m}{2}v^2 - \frac{m}{2}v_o^2 + \frac{M}{2}u^2$$

由此可求得摩擦力 f

$$f = \frac{1}{a}\left[\frac{m}{2}v_o^2 - \frac{m}{2}v^2 - \frac{m}{2}\left(\frac{m}{M}\right)^2(v_o - v)^2\right] = \frac{m}{2a}\left[v_o^2 - v^2 - \left(\frac{m}{M}\right)^2(v_o - v)^2\right]$$

则相关过程中木块的位移量

$$L = \frac{Mu^2}{2f} = \frac{a(m^2/M)(v_o - v)^2}{m[v_o^2 - v^2 - (m/M)(v_o - v)^2]} = \frac{am(v_o - v)^2}{M(v_o^2 - v^2) - m(v_o - v)^2]}$$

说明：就整个体系(子弹和木块)而言，动摩擦力对体系中的木块做正功，对子弹做负功，但负功大于正功，对整个系统而言，摩擦力作的总功总是负的，使系统的总动能减少。◀

3-4-2 保守力与非保守力

以物体受力后产生的效应描述作用力，其表达式一般是物体的位置和速度以及时间的函数：$\mathbf{F} = \mathbf{F}(t, \mathbf{r}, \mathbf{v})$。例如，汽缸中往复运动的活塞受到的力是时间的周期函数，弹簧振子所受到的弹性力则由其中弹簧的伸长或压缩量来度量，即弹性力是位置的函数，在流体中运动的物体所受到阻力起因于物体相对于流体的运动，因而是速度的函数，而惯性离心力和科里奥利力则是速度和位置的函数。如果物体所受到的作用力仅与它所处的空间位置有关：$\mathbf{F} = \mathbf{F}(\mathbf{r})$，这样的力在空间的分布称为相对于该物体的**力场**。物体在地面上所受到的重力：$\mathbf{F} = m\mathbf{g}$，以及胡克定律所描述的弹性力：$\mathbf{F} = -k\mathbf{x}$ 都是最常见力场的典型例子。

下面通过讨论物体在力场作用下行为说明力场的一些特性。

设想一个物体在力 $\mathbf{F}(\mathbf{r})$的作用下从位置 P_1 沿路径 ℓ 移动到位置 P_2，力 $\mathbf{F}$ 对该物体所做的功在数学上表述为力 $\mathbf{F}(\mathbf{r})$从位置 P_1 沿路径 ℓ 到位置 P_2 的曲线积分：

$$W = \int_{\mathbf{r}_1}^{\mathbf{r}_2}{}_{\ell}\mathbf{F}(\mathbf{r})\cdot d\mathbf{r}$$

其中该积分一般与具体的路径 ℓ 有关：$W = W(\mathbf{r}_1, \mathbf{r}_1, \ell)$，$\mathbf{r}_1$ 和 $\mathbf{r}_1$ 为位置 P_1 和 P_2 的坐标。如果它与具体的路径无关，即仅取决于物体的初始位置和最终位置：

$$\int_{r_1}^{r_2}{}_{\ell}\mathbf{F}(\mathbf{r})\cdot d\mathbf{r}=U_a(\mathbf{r}_2,\mathbf{r}_1) \tag{3-4-9}$$

则力 F 和力场 F(r)分别称为保守力和保守力场。力场 F(r)在驱动物体运动过程中对物体做的功与物体运动路径无关可表示为：

$$\int_{r_1}^{r_2}{}_{\ell_1}\mathbf{F}(\mathbf{r})\cdot d\mathbf{r}=\int_{r_1}^{r_2}{}_{\ell_2}\mathbf{F}(\mathbf{r})\cdot d\mathbf{r}$$

其中 ℓ_1 和 ℓ_2 是连接 P_1 和 P_2 的两条任意任意路径(见图 3-4-2)。交换第二个积分的积分上下限，将方程改写为

$$\int_{r_1}^{r_2}{}_{\ell_1}\mathbf{F}(\mathbf{r})\cdot d\mathbf{r}+\int_{r_2}^{r_1}{}_{\ell_2}\mathbf{F}(\mathbf{r})\cdot d\mathbf{r}=0$$

即物体沿路径 ℓ_1 从位置 P_1 运动到位置 P_2，再沿路径 ℓ_2 的反方向回到位置 P_1，F 对物体所做的总功为零。

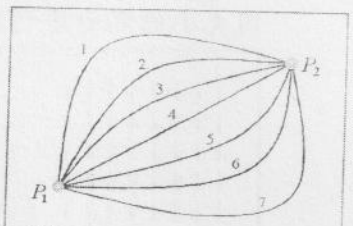


图 3-4-2

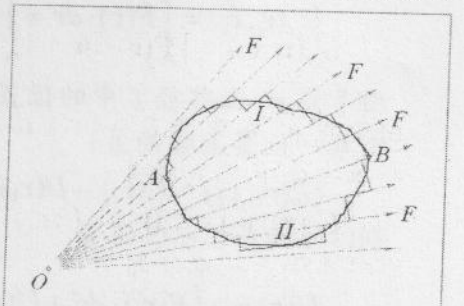
由于 P_1, P_2 以及路径的选取并没有任何限制，上述积分可一般地表示为

$$W=\oint_L\mathbf{F}(\mathbf{r})\cdot d\mathbf{r}=0 \tag{3-4-10}$$

即一个力场是保守的充分必要条件是它的回路积分为零。而对应于使 $W\neq 0$ 的 **F** 和 **F**(r)分别被称为**非保守力**和**非保守力场**。

容易证明重力和弹性力都是保守力，而滑动摩擦力，流体阻力则是典型的非保守力。这是因为重力和弹性力在物体的运动回路中所做的功有正功和负功两部分，因而可以抵消，而作为耗散力的滑动摩擦力和流体阻力总是和物体的运动方向相反，相应的功也总是负值，因而无法在一个封闭路径中发生任何抵消。但需要注意的是，非保守力并非总是做负功，一片在水面上的树叶在非保守的涡旋力的作用下越转越快，说明非保守的涡旋力始终在做正功。

例 3-4-2 如果空间各点的力都是指向或离开同一个空间位置(设为一个坐标系的原点 O)，而且力的大小仅仅取决于力的作用点和 O 点的距离 r (见图 3-4-2)：$\mathbf{F}(\mathbf{r})=F(r)\hat{\mathbf{r}}$，这样的力和力场分别称为**中心力**和**中心力场**。证明中心力场是保守场。



例 3-4-2 图 作用在物体上的力分解为径向和横向分量。

解：考虑中心力场中的一个任意回路积分，将回路中的任一微元分解为该位置沿力场方向(径向)和垂直于力场方向(横向)的两个分量(见图)。因为力场只有径向分量的位移对做功有贡献，因而有

$$\oint\mathbf{F}\cdot d\mathbf{s}=\int_{r_0}^{r_1}F(r)dr+\int_{r_1}^{r_2}F(r)dr+\cdots+\int_{r_n}^{r_0}F(r)dr=\int_{r_0}^{r_0}F(r)dr=0$$

因此中心力场是保守力场。◀

力场的回路积分所描述的力场图像是一个涡旋，由此可以从直观上判断一个力场为保守场的充分必要条件是力场中不存在涡旋。图(3-4-3)给出了几个力场的

几何图像，其中(a)描述的是均匀场，重力场是这类力场的典型例子，图(b)为中心力场，这两类保守力场的图像种都不具有涡旋的迹象。图(c)和图(d)中的弯曲流型给出了明显的涡旋视觉形象，读者可以直接在图上构建合适的回路证明它们都属于非保守力场。但并非只有在力场呈现弯曲流型时才出现涡旋，图(3-4-3-e)给出的力场分布中各点的力都是朝右的，但力的大小往下等比例地减小：$F = kz + f_o$。设想物体沿图中的矩形回路运动，则力场的回路积分对应的功只来自水平分量(力的方向与垂直路径正交，做功为零)，而两条水平路径虽然长度相等，方向相反，但力的大小不等，因此所做的功不能抵消，因而力场是非保守的。(用不同大小的涡旋排列来说明这一问题)

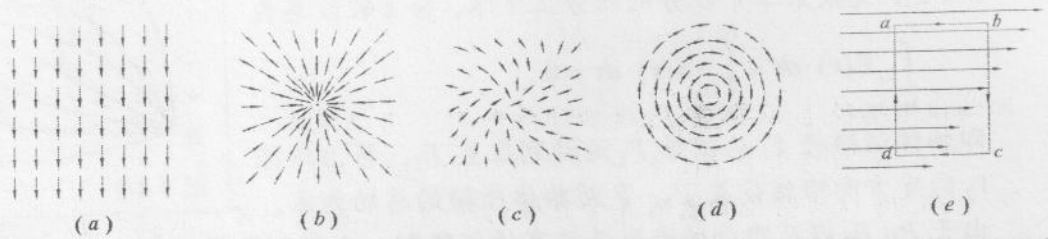


图 3-4-3 几种典型的力场的几何图像，它们可以在一定的程度上给出对力场是否为保守场的直观感受

3-4-3 势能

由于保守力做功和路径无关，力的路径积分就只是初始位置 $\mathbf{r}_1$ 和最终位置 $\mathbf{r}_2$ 的函数：

$$\int_{\mathbf{r}_1}^{\mathbf{r}_2}{}_{\ell}\mathbf{F}(\mathbf{r})\cdot d\mathbf{r} = U_o(\mathbf{r}_2,\mathbf{r}_1)$$

为了确定函数 U 中两个坐标之间的关联，我们将积分路径分为两段：$\mathbf{r}_1 \to \mathbf{r}' \to \mathbf{r}_2$，则 $\mathbf{F}$ 在物体的整个运动过程中所做的功可以表示为

$$U_o(\mathbf{r}_1,\mathbf{r}_2) = \int_{\mathbf{r}_1}^{\mathbf{r}_2}\mathbf{F}(\mathbf{r})\cdot d\mathbf{r} = \int_{\mathbf{r}_1}^{\mathbf{r}'}\mathbf{F}(\mathbf{r})\cdot d\mathbf{r} + \int_{\mathbf{r}'}^{\mathbf{r}_2}\mathbf{F}(\mathbf{r})\cdot d\mathbf{r} = U_o(\mathbf{r}_1,\mathbf{r}') + U_o(\mathbf{r}',\mathbf{r}_2) \qquad (3\text{-}4\text{-}10)$$

由于点 $\mathbf{r}'$ 在路径 ℓ 中的位置是任意的，这就要求 $U_o(\mathbf{r}_1, \mathbf{r}_2)$的函数形式只能是两个单一位置函数的差：

$$U_o(\mathbf{r}_1, \mathbf{r}_2) = U(\mathbf{r}_2) - U(\mathbf{r}_1) \qquad (3\text{-}4\text{-}11)$$

则(3-4-9)式可以改写为

$$U(\mathbf{r}) = -\int_{\mathbf{r}_o}^{\mathbf{r}}\mathbf{F}(\mathbf{r}')\cdot d\mathbf{r}' + U(\mathbf{r}_o) \qquad (3\text{-}4\text{-}12)$$

其中位置函数 $U(\mathbf{r})$具有能量量纲，因而被称为物体在 $\mathbf{r}$ 处具有的**势能**[2]。应该说明的是，(3-4-12)式中所定义的势能只是相对于位置 $\mathbf{r}_o$ 处势能 $U(\mathbf{r}_o)$的差值，其中位置 $\mathbf{r}_o$ 为参考点。对势能参考点所对应的位置 $\mathbf{r}_o$ 的选择主要是考虑计算方便。最常见的选择是以坐标原点或无限远点为势能零点。选定势能零点位置 $\mathbf{r}_o$ 后，表达式(3-4-12)简化为

[2] 势能以前曾被更直接地称为“位能”。

$$U(\mathbf{r}) = -\int_{\mathbf{r}_0}^{\mathbf{r}} \mathbf{F}(\mathbf{r}') \cdot d\mathbf{r}' \tag{3-4-13}$$

例如，一个弹簧振子中的物体质量为 m，弹性系数为 k。通常选择系统平衡位置对应的势能为势能零点，则系统的弹性势能可表示为：

$$U(x) = -\int_0^x (-kx)dx = \frac{1}{2}kx^2$$

如果惯性力满足环路积分为零的条件，形式上也可以理解为保守力场，受惯性力作用的物体也具有相应的势能。例如，在匀速转动参照系中的物体受到的惯性离心力 $\mathbf{F}_I = m\omega^2\mathbf{r}$ 具有中心力形式，满足环路积分为零。该物体相应的势能称为离心势能。若以圆心为离心势能的零点，则在距圆心 r 处的离心势能为 $-\frac{1}{2}mr^2\omega^2$。

3-4-4 机械能守恒

物体的动能和势能都和物体的机械运动有关，因而都称为**机械能**。

如果将物体和它所处的保守力场看成一个系统，由于保守力做功导致物体的动能增加一定是以系统中减少了什么作为补偿的。将方程(3-4-4)表示的功-能原理和(3-4-4)式表示的势能函数结合在一起可知上述的补偿量就是系统的势能：

$$K_2 - K_1 = W = U(\mathbf{r}_1) - U(\mathbf{r}_2)$$

将对应于不同状态的相关参量分别排列在等式的两边：

$$K_2 + U(\mathbf{r}_2) = K_1 + U(\mathbf{r}_1) \tag{3-4-14}$$

这意味着系统的总机械能 $E = K + U$ 在没有受到非保守力场的作用情况下是一个守恒量：

$$E(v_2, \mathbf{r}_2) = E(v_1, \mathbf{r}_1) \tag{3-4-15}$$

例如，理想弹簧振子不受外力作用时，质点的位移：$x = A\cos\omega t$，设其速度为

$$v(t) = \frac{dx}{dt} = -A\omega \sin \omega t$$

系统在 t 时刻的动能和势能分别为

$$U = \frac{1}{2}kx^2 = \frac{1}{2}kA^2\cos^2\omega t = \frac{1}{2}m\omega^2A^2\cos^2\omega t$$

$$K = \frac{1}{2}mv^2 = \frac{1}{2}m\omega^2A^2\sin^2\omega t$$

其中利用了关系式 $k = m\omega^2$。这两种不同形式的能量在振动过程中也在发生相应的交替变化，它们的振幅相等，相位相反。而系统的总机械能不随时间变化的守恒量：$E = K + U = \frac{1}{2}kA^2$。图3-4-4显示了简谐振子在振动过程中其动能、势能及总能量随时间的变化情况。

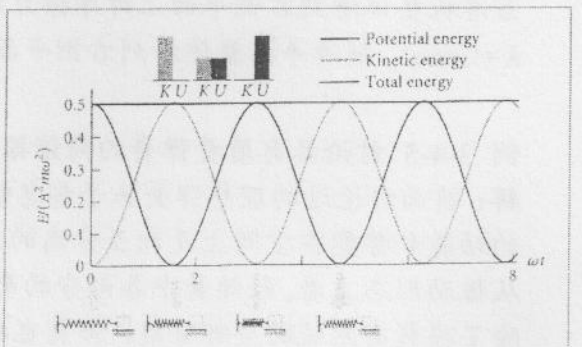


图3-4-4简谐振子的动能、势能及总能量在振动过程中随时间的变化情况。

由于功能原理和保守力场中运动的物理系统的总机械能守恒的表述都将物体的运动速度和物体在力的作用下的位移联系起来

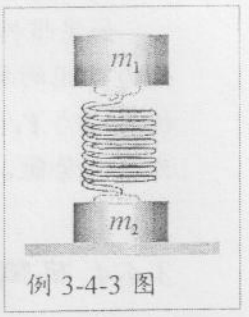
$$\Delta E = \Delta K + \Delta U = 0$$

而由功能原理可知，在系统同时涉及保守力和非保守力的情形，系统的总机械能的变化应等于非保守力对系统做功

$$\Delta E = \Delta K + \Delta U = W$$

这些表达式经常用于求解一些复杂的力学问题而避免细致的受力分析。

例 3-4-3 质量为 m_1 和 m_2 的两个物体用劲度系数为 k 的轻弹簧相连，竖直地放在桌面上。用力 F 压住物体 m_1 后放手，F 多大时反冲的 m_1 能正好拉起 m_2？



例 3-4-3 图

解： 设在力 F 的作用下，弹簧被压缩的总长度为 x，则有

$$F + m_1 g = kx \tag{1}$$

放手后，m_1 仅受弹力和重力作用，其机械能守恒。设 m_1 到达最高点时弹簧的伸长为 x'，此时 m_1 的速度为零，故有

$$\tfrac{1}{2}kx^2 = \tfrac{1}{2}kx'^2 + m_1 g(x + x') \tag{2}$$

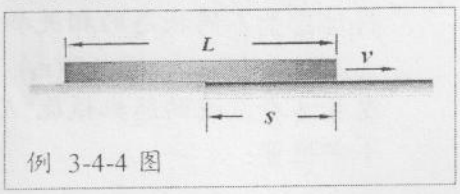
为使 m_2 能被从桌面拉起，弹簧的拉力至少应等于 m_2 的重力，

$$kx' = m_2 g \tag{3}$$

因而由(2)式，$k(x+x')(x-x') = 2m_1 g(x+x')$，即 $k(x-x') = 2m_1 g$。将(1)(3)式代入得：

$$F = (m_1 + m_2)g \quad \blacktriangleleft$$

例 3-4-4 设一块长度为 L，质量为 m 的薄平板以一定的初速度 v_o 从左侧的光滑平面上平移到右侧滑动摩擦系数为 μ 的粗糙平面上。计算薄平板要整体滑到右侧的平面上所需要的最小初速度。



例 3-4-4 图

解： 由于薄平板在运动过程中作用在地面上的压力是均匀的[3]，右侧平面对平板的弹力正比于平板在右侧平面上的比例：

$$N = \begin{cases} mg\dfrac{x}{L} & 0 < x < L \\ mg & x > L \end{cases}$$

设薄平板前缘在右侧平面上平移了距离 $s\ (s \le L)$，则摩擦力对平板做的功为

$$W_f(s) = -\int_0^s f dx = -\int_0^s \mu mg \frac{x}{L} dx = -\frac{\mu mg}{2L} s^2$$

若薄板整体滑到右侧平面上时摩擦力做的功恰好耗尽平板原有的动能：$W_f(L) = K = \tfrac{1}{2}mv_o^2$，则薄平板整体滑到右侧平面上所需要的最小初速度为 $v_o = \sqrt{\mu g L}$。$\blacktriangleleft$

例 3-4-5 讨论具有质量弹簧的简谐振动行为

解： 前面讨论过的理想弹簧振子都忽略了弹簧的质量。该系统的一个特点是系统的动能和势能在空间上是相互分离的。即势能集中于弹簧，而质点携载全部动能。从振动形态上看，轻弹簧中各部分的形变是均匀的。但对于具有质量的实际弹簧，除了携载有系统的全部势能，弹簧也携载有动能，其振动形态上看，弹簧各处的形变可以是均匀的，但也可以是不均匀的(请读者考虑其中的原因)。如果弹簧各

[3] 对有明显高度的物体在运动中对地面的压力可能不均匀，详细的讨论见第 5 章中的例 5-6-3。

处的形变仍然是均匀的，系统就象理想振子一样处于简谐振动状态。设弹簧和振子的质量分别为 m 和 M，弹簧的原长为 ℓ_o，弹性系数为 k。则弹簧长度为 ℓ，振子的运动速度为 v 时的系统能量为

$$E=\frac{1}{2}k(\ell-\ell_o)^2+\frac{M}{2}v^2+\frac{1}{2}\frac{m}{\ell}\int_0^{\ell}dx\frac{x^2}{\ell^2}v^2$$

$$=\frac{1}{2}k(\ell-\ell_o)^2+\frac{M}{2}v^2+\frac{m}{6}v^2=\frac{1}{2}kx^2+\frac{1}{2}\left(M+\frac{m}{3}\right)v^2$$

其中 $x=\ell-\ell_o$ 为弹簧的形变数量，由于整个系统的能量守恒，因而有

$$\frac{dE}{dt}=kx\frac{dx}{dt}+\left(M+\frac{m}{3}\right)\frac{dx}{dt}\frac{d^2x}{dt^2}=0$$

由此可得到

$$\left(M+\frac{m}{3}\right)\frac{d^2x}{dt^2}=-kx$$

它在形式上和一个质量为 $M+\frac{1}{3}m$ 的理想振子运动方程相当，其振动圆频率为 $\omega=k\Big/\sqrt{M+m/3}$ ◀

例 3-4-7 设双原子分子中的相互作用可以用弹性力图像来描述，设弹性系数为k，利用能量分析的方法确定双原子分子，如盐酸分子HCl的振动情况。

解：双原子分子中的两个原子的相对振动与弹簧振子图像的差异是弹簧的两个端点都不是固定的，如果用一个原子(如氢原子)相对于另一个原子(如氯原子)的振动来分析氢原子受力运动情况，还需要考虑惯性力，由于整个分子的总机械能

$$E=\tfrac{1}{2}m_{\mathrm{H}}v_{\mathrm{H}}^2+\tfrac{1}{2}m_{\mathrm{Cl}}v_{\mathrm{Cl}}^2+\tfrac{1}{2}k(x_{\mathrm{Cl}}-x_{\mathrm{H}}-\ell)^2$$

其中ℓ是两个原子的平衡间距，由于分子在振动过程中总动量为零：$m_{\mathrm{H}}v_{\mathrm{H}}+m_{\mathrm{Cl}}v_{\mathrm{Cl}}=0$，两个原子的速度满足关系：$v_{\mathrm{H}}=-(m_{\mathrm{Cl}}/m_{\mathrm{H}})v_{\mathrm{Cl}}$。故上述能量表达式可表示为

$$E=\frac{m_{\mathrm{Cl}}}{2m_{\mathrm{H}}}Mv_{\mathrm{Cl}}^2+\frac{M^2}{2m_{\mathrm{H}}^2}k\left(x_{\mathrm{Cl}}-\frac{m_{\mathrm{H}}}{M}\ell\right)^2$$

其中$M=m_{\mathrm{Cl}}+m_{\mathrm{H}}$为分子质量。由于系统的机械能守恒：

$$\frac{dE}{dt}=0=\frac{m_{\mathrm{Cl}}}{m_{\mathrm{H}}}Mv_{\mathrm{Cl}}\frac{dv_{\mathrm{Cl}}}{dt}+\frac{kM^2}{m_{\mathrm{H}}^2}\left(x_{\mathrm{Cl}}-\frac{m_{\mathrm{H}}}{M}\ell\right)^2\frac{dx_{\mathrm{Cl}}}{dt}$$

或 $\dfrac{d^2x_{\mathrm{Cl}}}{dt^2}+\dfrac{k}{\mu}\left(x_{\mathrm{Cl}}-\dfrac{m_{\mathrm{H}}}{M}\ell\right)=0$，设参量$x=x_{\mathrm{Cl}}-\ell m_{\mathrm{H}}/M$，上式简化为

$$\frac{d^2x}{dt^2}+\frac{k}{\mu}x=0$$

其中 $\mu=m_{\mathrm{Cl}}m_{\mathrm{H}}/M$ 称为分子的**折合质量**。由此可得分子的振动频率为

$$\omega=\sqrt{\frac{k}{\mu}}=\sqrt{k\frac{m_{\mathrm{H}}+m_{\mathrm{Cl}}}{m_{\mathrm{H}}m_{\mathrm{Cl}}}}$$

对分子振动的实际研究过程是相反的：通过分子对特定频率的光的吸收确定分子的特定振动频率，并由此得到分子内部的相互作用特征。◀

3-4-5 作用力的定义

比较保守力的各直角分量对物体做的微功和势能的微小变化之间的关系式

$$dU = -\mathbf{F}\cdot d\mathbf{r} = -F_x dx - F_y dy - F_z dz \tag{3-4-16}$$

和势能的全微分标准数学表达式

$$dU = \frac{\partial U}{\partial x}dx + \frac{\partial U}{\partial y}dy + \frac{\partial U}{\partial z}dz \tag{3-4-17}$$

可以将保守力的各直角分量表示为势能 $U(\mathbf{r})$的函数

$$F_x = -\frac{\partial U}{\partial x} \quad F_y = -\frac{\partial U}{\partial y} \quad F_z = -\frac{\partial U}{\partial z} \tag{3-4-18}$$

上述保守力分量的表达式可统一地用矢量形式表示

$$\mathbf{F} = -\frac{\partial U}{\partial x}\hat{\mathbf{x}} + \frac{\partial U}{\partial y}\hat{\mathbf{y}} + \frac{\partial U}{\partial z}\hat{\mathbf{z}} = -\nabla U \tag{3-4-19}$$

上述力的定义与(3-1-5)和(3-3-4)中表述的力的含义完全不同。牛顿定律反映的是力作用在物体上所产生的某种效果，而势能的负梯度反映的更接近与力的本身。由于物体势能的大小仍然与物体本身的特性有关，受力空间环境应该用一种称为"势场"的物理概念及相关的物理量来描述，它可以定义为作用力除以有关物体的相关特性参量。例如与重力场相关的物体参量是物体的质量，**重力场**可定义为：

$$\mathbf{G} = \frac{\mathbf{F}_G}{m} = \mathbf{g}$$

其操作性定义仍然依赖于实际的可观测效果，即一个实际物体的运动数量的变化来描述。

按照功的定义(3-4-3)，我们也可以将力的效应表述为物体的动能相对于空间位移的变化率。与物理量对时间变化率不同，位移是一个矢量，而我们在第一章的习题中已讨论过定义矢量的商所存在的问题，因此我们只能分别讨论力的每一个分量。(3-4-1)式可以写成

$$\Delta W = F_x\Delta x + F_y\Delta y + F_z\Delta z = \Delta K \tag{3-4-20}$$

则力的各个分量可分别表示为：

$$F_x = \frac{\partial K}{\partial x}, \quad F_y = \frac{\partial K}{\partial y}, \quad F_z = \frac{\partial K}{\partial z} \tag{3-4-21}$$

这三个分量可综合为

$$\mathbf{F} = \nabla K = \frac{\partial K}{\partial x}\hat{\mathbf{x}} + \frac{\partial K}{\partial y}\hat{\mathbf{y}} + \frac{\partial K}{\partial z}\hat{\mathbf{z}} \tag{3-4-22}$$

∇K 称为动能 K 的梯度。力的这种表述在物理学教材中很少提及，因为通常认为其中有一些含糊不清的地方：动能表达式中不显含坐标参量，有时对运动状态的描述中将位置与速度作为两个独立变量来考虑。因此如果用一个物体的动能的空间梯度来描述该物体的受力情况，实际是和物体的受力空间环境联系在一起的。一般认为受力空间环境应该用上述"势场"的物理概念及相关的物理量来描述，但"势场"本身是不可观测的，其操作性定义仍然依赖于实际的可观测效果，即一个实际物体的运动数量的变化来描述。从观察力作用于物体产生的效应而言，"动量"和"动能"的变化都可以作为"力"的操作性定义。这种定义所适用的场合相对

于式(3-4-3)的适用范围更广，如势场梯度只适用于描述保守力，而式(3-4-22)还能描述非保守力。

在 3-8 节相对论动力学框架中，我们将看到物体的动量和动能都是同一个运动量的不同分量。

势能曲线

势能作为位置的函数的几何表示称为**势能曲线**，该曲线的斜率负值就是相应的保守力。势能曲线常用于显示物体在保守力场运动时的一些定性的特征。例如，图 3-4-4 给出的分别是重力势能曲线和弹簧振子系统的势能曲线。图 3-4-5 是典型的双原子分子势能曲线示意图。双原子分子中两个原子的相互作用可以分为排斥和吸引两部分。一种典型的数学表述是

$$U(r)=\frac{\sigma_1}{r^{12}}-\frac{\sigma_2}{r^6}$$

其中 r 为原子间距，σ_1 和 σ_2 分别为排斥和吸引作用的强度参数。图 3-4-5 中的两条虚线分别为排斥和吸引势能。粒子的总能量 E 大小由图中的水平线表示。而粒子处在位置 r 时，E 和 $U(r)$ 之间的差值为粒子的动能。其速度可表示为位置 $\mathbf{r}$ 的函数：

$$v^2=\frac{2}{m}[E-U(\mathbf{r})] \tag{3-4-23}$$

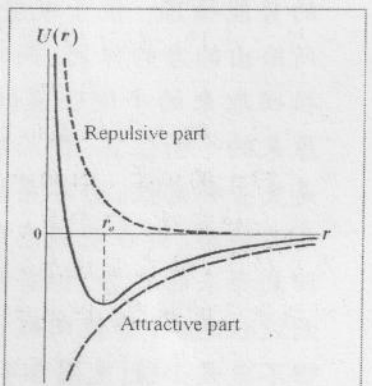


图 3-4-5 一个双原子分子中一个原子的势能曲线

由于速度的平方不能取负值，这意味着系统的势能不会超过总能量：$U(\mathbf{r})\le E$。当系统运动到总能量水平线与势能曲线的相交点 $\mathbf{r}_c$ $(U(\mathbf{r}_c)=E)$，物体因其动能耗尽而只能在力场的作用下返回，因而位置 $\mathbf{r}_c$ 被称为粒子运动的**转折点**。势能曲线上的各转折点 $r=r_{ci}(i=1,2,\ldots)$ 给出了粒子运动区域的边界：质点被限制在两个转折点之间 $U(\mathbf{r})\le E$ 的空间区域，通常称粒子处于一个约束轨道之中。而 $U(\mathbf{r})>E$ 区域称为**禁区**。对于图 3-4-5 中的势能曲线描述的系统，如果粒子的总能量值大于零，粒子向右面的运动就不会受到约束，这时称粒子处于一个非约束轨道。粒子在该势能曲线的底部 $r=r_o$ 位置达到最大速度，由于在该位置曲线的斜率为零：$\partial U/\partial r=-F_r=0$，这意味着 $r=r_o$ 是粒子的一个平衡位置：分子中的每一个原子受到的净力为零。而在该平衡位置的两侧

$$r>r_o \quad F=-\frac{dU}{dr}<0 \quad \text{为吸引力}$$

$$r<r_o \quad F=-\frac{dU}{dr}>0 \quad \text{为排斥力}$$

平衡的稳定性

一个物体在所受外力的总和为零时处于平衡状态，但平衡有几种不同的形态。例如，一个滚珠可以静止(平衡)在一个倒置的碗的顶部，但只要受到一个非常小的扰动，滚珠就会偏离该平衡位置，并在外力(重力和支撑力)作用下，呈现出加速的单向偏离趋势。这种平衡称为**不稳定平衡**。而一个放置在碗内底部的滚珠也可以长时间地平衡在该位置，如果对滚珠作一个小的扰动，滚珠在外力作用下也会偏离该平衡位置，但偏离速度会不断减小，物体偏离的趋势最终回扭转，物体

会回到原来的平衡位置。这种平衡称为**稳定平衡**。

上述例子对平衡稳定性的判断是简单而且直观的，但更多的情况很难看清楚，这时，利用势能曲线可以很方便地判断一个物体所处位置的稳定程度。例如，一个弹簧振子系统因受到压缩或伸长而偏离平衡位置，由于弹力总是将物体拉回平衡位置。系统在自然状态下的平衡是稳定的。这一特性在系统的势能曲线 $U(x)=\frac{1}{2}kx^2$ 中表现得非常清楚。对平衡位置 $x=0$ 的微小偏离 x，势能曲线在该位置变化率所对应的作用力为 $F=-\partial U/\partial x=-kx$ 总是指向 $x=0$ 的位置，即物体虽然并不安居于平衡位置，但也不会加剧偏离该平衡位置，因而称该系统的平衡是**稳定**的。如图 3-4-6 所示，不稳定平衡位置对应于势能曲线中的局部峰值位置(极大值，图中的 B 点)，而稳定平衡位置对应于势能曲线中的局部谷底位置(极小值，图中的 A 和 C 点)，当系统偏离稳定的势能谷底或不稳定的势能峰顶，粒子就会受到由势能曲线的斜率负值所给出的力的作用，而该作用力的方向决定着粒子被拉回原来的平衡位置(恢复力)还是被进一步地推离原来的平衡位置，因而决定着这个里是使系统稳定还是失去稳定性。力学系统的另一种可能的稳定状态称为亚稳态。处在亚稳态的系统受到较小的扰动后仍能回到原来的状态，但若受到较大的扰动后就不再能回到该状态。在势能曲线上亚稳态对应与势能的极小值，但不是最小值(见图 3-4-6 中的 C 点)。

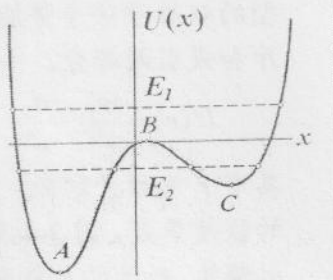


图 3-4-6 势能曲线给出的粒子处在不同位置的各种信息。

思考题

3-4-1 3-1-1 节中的反冲实验中两个物体获得的动量分别为 m_Av_A 和 m_Bv_B，两者相等得到：$m_A/m_B=v_B/v_A$，而两个物体获得的动能分别为 $\frac{1}{2}m_Av_A^2$ 和 $\frac{1}{2}m_Bv_B^2$，两者相等得到：$m_A/m_B=(v_B/v_A)^2$。如何解释其中的矛盾？

3-4-2 如何理解在图 3-4-1 所示的实验中，小球在与弹簧以及斜面发生作用的前后其动能不发生变化，但动量却不守恒？

3-4-3 如果系统中涉及到的作用力可以分解为两个独立的分量，是否对于两个独立的运动方向应该有各自对应的能量守恒定律？为什么？

3-4-4 说明物体的动能相对于不同的惯性系有不同的数值，但动能定理与惯性参照系无关。

3-4-5 人跳高时获得的重力势能来自哪里？获得向上的动量来自哪里？

3-4-6 汽车起动时所获得的动能和动量来自何处，怎样转化过来的？

3-4-7 在地面上拍球时球的动量不守恒，但动能守恒，应如何理解这一现象？

3-4-8 将物体匀速或匀加速地上升到同样的高度，外力对物体做的功是否相同？

3-4-9 两条长度相同的绳子从天花板垂下，其中一条具有弹性，另一条没有弹性。如果两个人各攀一条绳子到顶端，哪个人做的功更多？

3-4-10 一个小球沿着水平面上的一条圆形轨道匀速滚动，作用在小球上的向心力是否是保守力？(Question 7-6)

3-4-11 如果知道一个物体在某一时刻的运动状态(位置和速度)，能否确定物体在

该时刻的动能和势能？如果知道一个物体在某一时刻的动能和势能，能否确定其运动状态？(赵 3-1)

3-4-12 保守力在一维的情形和高于一维的情形是否存在着什么差别？

3-4-13 真实弹簧作用在物体上的力是否是保守力？设计一个可以检验你答案的实验。

3-4-14 任何一个可以表示为仅仅是位置的函数的一维作用力都是保守力，如何解释摩擦力是非保守力这一事实？

3-4-15 从微观尺度上说明流体阻力(如空气阻力)是非保守力。

3-4-16 Is the potential energy zero at any point where the force is zero? Suppose that the potential energy is zero at a given point. Is the force necessarily zero or nonzero at that point?

3-4-17 设两个物体之间存在着某种相互吸引作用，其大小与两者之间距离 r 的平方成反比，能否选择 $r=0$ 或 $r\to\infty$ 作为势能的参考点(零点)？怎样的选择更为合理？

3-4-18 若函数 $U(x)$ 在 $x=x_0$ 处的一阶导数和二阶导数都等于 0，但三阶导数不为 0，则该处称为函数的拐点。讨论势能曲线拐点处平衡的稳定性问题。

3-4-19 列举图 3-4-3 中的各种几何图像所对应的一些实际力场。

3-4-20 用规则排列的相同涡旋表示均匀力场，是否说明均匀力场中存在有涡旋？

3-4-21 由于势能曲线中相对于势能最低点左右并不对称，具有较高动能的基元粒子之间就有比较大的(动态)平衡间距

3-4-22 说明为什么一个橡胶球自由下落，在地面上反弹后通常达不到原来的高度。

习题

3-4-1 弹性系数为 k 的弹性绳上悬挂质点 m，从绳的原长状态对应的位置静止释放。(*a*) 从动力学及能量角度证明：伸长量 Δl 远远小于原长 l 时有 $\Delta l\approx 3mg/k$。(*b*) 证明质点运动到最低点时的速度为 $v=\sqrt{2g(l-3mg/2k)}$，解释该速度小于非弹性绳的速度的原因。

3-4-2 在一弹簧下挂一重物，放开后将迅速下沉，使弹簧拉伸到某一最大长度后回升。如果我们用手托着它缓缓下沉，到达某一高度时它就不动了。比较重物在 A、B、C 三个位置上的总势能(重力势能和弹性势能之和)大小。画出总势能与高度的函数曲线来，并与弹性势能曲线作比较。

3-4-3 一列火车总质量为 M，最后一节车厢质量为 m. 若 m 从匀速前进的列车中脱离然后行进了路程 s 后停下。若机车的牵引力不变，每节车厢所受的摩擦力与速度无关。求脱开的那节车厢停止时与列车后端的距离。(3-1)

3-4-4 一物体由粗糙斜面底部以初速 v_0 冲上去后又沿斜面滑下来，回到底部时的速度减为 v_f。求此物体达到的最大高度。(3-4)

3-4-5 一个质点自半径为 R 的光滑球面的顶点由静止开始下滑，质点下滑到什么位置会脱离球面？求质点的运动轨迹。(3-2)

3-4-6 重物从高度为 h 处沿光滑轨道滑下后，进入半径为 R 的竖立圆环的底部，然后在圆环内作圆周运动。若重物到达圆环顶点不脱离轨道，重物的起始

下滑位置至少要达到多少高度 h_o？如果 $h=0.8h_o$，求质点从最低点出发后的运动轨迹。(3-3)

3-4-7 用细线将一质量为 M 的大光滑圆环悬挂起来。两个质量均为 m 的小圆环套在大圆环。若两小圆环沿相反方向从大圆环顶部自静止下滑，小圆环在下滑到哪个角度 θ 时大圆环开始上升？

3-4-8 弹性系数为 k 的弹簧下挂一质量为 m_1 的物体，另外一个质量为 m_2 的物体用细线连接在 m_1 的下面。系统达到静止状态后，剪断 m_1 和 m_2 之间的连线，求物体 m_1 可以达到的最大速度。

3-4-9 光滑的水平面上，劲度系数为 k 的弹簧一端固定在墙上，另一端系一物体 A，将弹簧压缩 a 后，紧贴 A 旁放一物体 B。撤去外力，求 A、B 分离时 B 的速率以及 A 距起始点的最大距离。设 A 和 B 的质量分别为 m_A 和 m_B。

3-4-10 半径为 R 的大圆环的顶点固定于 A 处，套在大环上的小环 m 通过一劲度系数为 k、自然长度为 $L(<2R)$的弹簧系于 A 点。分析在不同的参数下该装置平衡点的稳定性，画出相应的势能曲线。

3-4-11 质量分别为 m_A 和 m_B 的物体 A 和 B 用一不可伸长的轻绳连接，A 置于摩擦系数为 μ 的水平桌面上，B 通过桌沿的滑轮自然下垂。求物体 B 从静止下降高度 h 后所获得的速度。

3-4-12 *长为 L、线密度为 λ 的柔软绳索放置在光滑的桌面上，设开始时，绳索的一端 A 略伸出桌沿，并在重力作用下带动整条绳索开始运动，说明 A 端下落高度达到 x 时绳索的运动情况。

3-5 角动量

当作用于物体上的合力为零的情形，物体处于速度恒定的惯性运动状态，仔细观测这样一个运动物体的相关参量可以发现，除了动量和动能守恒，物体相对于一个给定的空间位置(如坐标原点 O)的位置矢量 $\mathbf{r}$ 在每一个给定的时间间隔 Δt 内扫过的每个三角形的面积(见图 3-5-1)也是一个不变量：

$$S = \tfrac{1}{2}v\Delta t\cdot a = \tfrac{1}{2}v\Delta t\,(r\sin\theta) = \tfrac{1}{2}\Delta t|\,\mathbf{r}\times\mathbf{v}\,|$$

其中 $\mathbf{v}$ 是物体的运动速度，$a = r\sin\theta$ 是质点运动轨迹与坐标原点 O 的最短距离。这说明自由运动物体还存在一个对应于上述面积变化率的新的守恒"运动量"：物体的位置矢量和动量的矢量积：

$$\mathbf{L} = \mathbf{r}\times\mathbf{p} = m(\mathbf{r}\times\mathbf{v}) \tag{3-5-1}$$

这个与空间位置有关的运动参量称为**角动量**。在国际单位制中其单位为 $\mathrm{kg\cdot m^2/s}$。

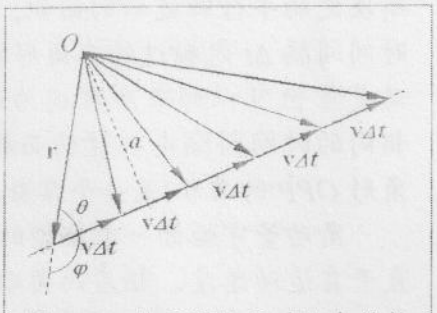


图 3-5-1 处于惯性运动状态的物体相对于一个给定的空间点 O 的位置矢量 $\mathbf{r}$ 在每个给定的时间间隔 Δt 内扫过的每个三角形的面积是一个不变量。

读者可能会认为引进一个自由粒子的角动量以及相关的守恒定律并没有什么实际意义，因为它似乎没有对这个运动状态给出新的信息。但从(3-5-1)的定义中可以看出，角动量守恒相对于动量和动能的守恒成立的条件有了更多的自由度(可以同时改变位置和速度但维持两者的矢量积不变)，这意味着引起角动量发生改变的动力学作用以及角动量守恒的一般动力学条件比动量守恒和动能守恒成立的动力学条件也相应地存在更多的自由度。

通过角动量的时间变化率可以探讨力以什么样的形式作用在物体上能够有效地改变物体的角动量：

$$\frac{d\mathbf{L}}{dt} = \frac{d}{dt}(\mathbf{r}\times\mathbf{p}) = \frac{d\mathbf{r}}{dt}\times\mathbf{p} + \mathbf{r}\times\frac{d\mathbf{p}}{dt} \tag{3-5-2}$$

上式右边的第一项

$$\frac{d\mathbf{r}}{dt}\times\mathbf{p} = \mathbf{v}\times(m\mathbf{v}) = m(\mathbf{v}\times\mathbf{v}) = 0$$

而式(3-5-2)右边的第二项中的动量随时间的变化率就是作用在物体上的力 $\mathbf{F}$。而 $\mathbf{r}\times\mathbf{F} \equiv \boldsymbol{\tau}$ 描述的是外力 $\mathbf{F}$ 作用于物体的一种方式，称为**力矩**，则式(3-5-2)表示力矩作用于物体产生的效应为物体的角动量的变化率：

$$\boldsymbol{\tau} \equiv \mathbf{r}\times\mathbf{F} = \frac{d\mathbf{L}}{dt} \tag{3-5-3}$$

称为**转动定律**，或转动运动方程。在国际单位制中力矩的单位为牛顿·米(N·m)[1]。

物体的角动量守恒的条件是作用在该物体上的合力矩为零。合力矩为零包含有两种可能性：合外力 $\mathbf{F}$ 为零，或者外力与位置矢量 $\mathbf{r}$ 平行或反平行：

$$\mathbf{F} = F\hat{\mathbf{r}} \tag{3-5-4}$$

式(3-5-2)所描述的力称为**中心力**。角动量守恒定律在物理学中是一个重要的守恒

[1] 力矩与能量具有相同的量纲，但注意不要采用相同的单位，如不要用焦耳作为力矩的单位以避免混淆。

律，因为与我们日常生活密切相关的物体之间的两个基本相互作用(引力和库仑力)都是中心力。

处于角动量守恒状态的物体的运动轨道具有两个几何特征。首先，由于角动量的方向不变要求r和p必须始终困守于与角动量垂直的平面内。即对质点而言，角动量守恒的运动状态只能是平面运动[2]。第二，$|\mathbf{r}\times\mathbf{v}|\Delta t = rv\Delta t\sin\theta$ 在几何上对应于由矢量r和$\mathbf{v}\Delta t$所决定的平行四边形的面积，它是位置矢量r在时间间隔Δt内扫过的三角形面积的两倍。因此角动量守恒可以形象地描述为质点的位置矢量在相同的时间间隔内扫过的面积(见图 3-5-2 中三角形 OPP'的面积)是一个常数。

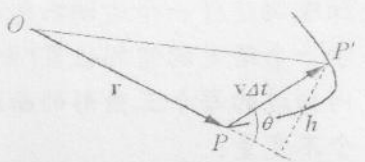
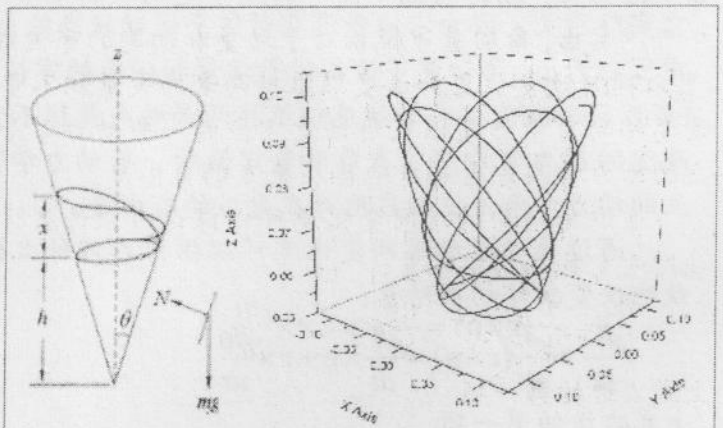


图 3-5-2 角动量守恒意味着为质点的位置矢量在相同的时间间隔内扫过的面积相同。

角动量守恒的一个典型的事例时物体作匀速圆周运动，物体位置矢量总是垂直于其运动速度。相应的角动量方向垂直于轨道平面，其大小 $L = rmv =$ 常数。

不同于功和能量是标量，力矩和角动量都是矢量，相应的守恒律涉及三个分量，即只要力矩的一个分量为零，相应的角动量分量就维持不变。

例 3-5-1 一个质量为 m 的质点在半顶角为 θ 的圆锥的光滑内壁面上运动，设在 $t = 0$ 时刻，质点距离锥顶高度为 h，计算质点能约束在该高度的水平面上运动的临界速度 v_c。若在时刻 $t = 0$，质点速度沿水平方向，大小为上述临界速度的两倍：$v_o = 2v_c$，计算质点在运动过程中受到的约束范围。



例 3-5-1 图

解： (1) 质点在重力 mg 和锥面内壁弹力 N 作用下作圆周运动，因而

$$mg\cot\theta = mv_c^2/R$$

其中圆锥面高度为 h 的圆周半径 $R = h\cdot\tan\theta$，由此可解得物体能稳定在高度 h 处的初速度临界值：$v_o = v_c = \sqrt{gh}$。

(2) 质点初速度大于 v_c 时不再能维持在高度为 h 处的平面上运动，质点往上攀升，但无法停留在某一特定高度的平面上作匀速圆周运动(读者思考其中的原因)，因而它将在一定的高度范围内运动。因质点运动过程中没有耗散效应，机械能守恒。另外，因 N 和 mg 和圆锥对称轴 z 处于同一平面内，质点没有受到 z 方向的外力矩分量：$\tau_z = 0$，因而 $L_z =$ 常量。用 $h+x$ 表示物体所能攀升到的极限高度，由于小球在该高度时的速度只有水平方向，能量守恒方程和角动量守恒方程可分别表示为：

$$\tfrac{1}{2}mv^2 + mg(h+x) = \tfrac{1}{2}mv_o^2 + mgh$$

$$(h+x)\cdot\tan\theta\cdot mv = h\cdot\tan\theta\cdot mv_o$$

[2] 具有尺度的物体

上述联立的方程组给出一个关于 x 的三次方程：

$$2gx^3+(4gh-v_0^2)x^2-2h(v_0^2-gh)x=0$$

其中解 $x=0$ 对应于初速度 v_0 沿水平方向，对于 $v_0=2v_c=2\sqrt{gh}$ 的初始速度，方程的一个正数解为 $x=\sqrt{3}h$。即小球的往返运动高度范围为 h 和 $(1+\sqrt{3})h$ 之间。

读者可以通过数值计算及作图，了解质点在圆锥内表面上的运动轨迹，并探讨周期运动的可能性。◀

需要强调的是，同样作为物体动力学特性的运动参量，动量和动能不依赖于空间位置，而角动量则总是相对于一固定的空间位置而言的。在没有事先指定位置的前提下是无法讨论角动量的概念的。例如，作匀速圆周运动的物体的角动量守恒是相对于圆轨道中心而言的，而相对于其它的位置，则不存在角动量守恒的问题。

对处于圆周运动的质点，外力矩对质点所做的功可以用力矩和质点转动历经的角位移 表示：$\Delta W=\mathbf{F}\cdot\Delta\mathbf{r}=\mathbf{F}\cdot(\Delta\boldsymbol{\theta}\times\mathbf{r})=\Delta\boldsymbol{\theta}\cdot(\mathbf{r}\times\mathbf{F})=\boldsymbol{\tau}\cdot\Delta\boldsymbol{\theta}$。外力矩对质点输入的功率可表示为力矩和质点角速度的点积：

$$P=\frac{dW}{dt}=\boldsymbol{\tau}\cdot\frac{\mathbf{d}\boldsymbol{\theta}}{dt}=\boldsymbol{\tau}\cdot\boldsymbol{\omega} \tag{3-5-5}$$

例 3-5-2 A dragster of mass $m = 10^3$ kg has its speed v = 15 m/s and acceleration a = 27 m/s^2 at an instant. Its engine is turning at 5.0×10^3 rmp. What minimum torque must the engine exert on the dragster's drive train? Why can you only find a minimum necessary torque from the given information rather than the actual torque exerted by the engine?

思考题

3-5-1 Angular momentum is sometimes given in units of joule-seconds. Is this a correct SI unit for angular momentum?

3-5-2 利用角动量以及功和能的概念说明荡秋千的原理。

习题

3-5-1 The earth is traveling in an elliptical orbit around the sun with a perigee distance of 7.02×10^6m and an aphelion distance of 10.3×10^6m (Figure 5-22). At perihelion, the speed of the earth is 8.22×10~ m/s. What is the speed of the earth at apogee?

3-5-2 如本题图，一质量为 m 的质点自由降落，在某时刻具有速度 v，此时它相对于 A、B、C 三参考点的距离分别为 d_1, d_2, d_3。求：(1) 质点对三个点的角动量；(2) 作用在质点上的重力对三个点的力矩。(4-1)

3-5-3 一质量为 m 的粒子位于(弦，锣)处，速度为砂=bε 十句∫，并受到土个沿一 2 方向的力∫. 求它相对于坐标原点的角动量和作用在其上的力矩。

3-5-2 如本题图，在光滑水平面上立一圆柱，在其上缠绕一根细线，线的另一头

系着一个质点。开始时一段线拉直，横向给质点一个冲击力，使它开始绕柱旋转。随着线愈绕愈短，质点的角速度怎样变化？质点的角动量以及动能是否守恒？

3-5-3 圆锥摆的摆弦线穿过一个中空的管子(见图)，可用于调节摆的长度。设摆长为 l_1 时摆锤的线速度为 v_1，将摆长调节到 l_2 时，摆锤的速度 v_2 为多少？圆锥的顶角有什么变化？

3-5-4 如图，在一半径为 R、质量为 m 的水平转台上有一质量½m 的玩具车。起初车在转台边缘，转台以角速度 ω 绕中心轴旋转。汽车相对转台沿径向向里开，求走到½R 处时转台的角速度？动能改变多少？能量从哪里来？

3-5-5 若将地球绕太阳的公转看作是以太阳中心为圆心的圆周运动，求地球相对太阳中心的角动量。已知地球的质量 $M = 6.0\times10^{24}$kg，轨道半径 $R = 1.49\times10^{11}$m。

3-5-6 一颗卫星沿椭圆轨道绕地球运行在近地点，卫星与地球中心的距离为地球半径的 3 倍。而在近地点，卫星的速度为在远地点时的 4 倍 求在远地点时卫星与地球中心之间的距离为地球半径的多少倍。

3-5-7 由火箭将一颗人造卫星送入离地面很近的轨道，进入轨道时，卫星的速度方向平行于地面，其大小为在地面附近作圆运动的速度的 $1.5^{½}$倍。求该卫星在运行中与地球中心的最远距离。

3-5-8 一质子从远处以速度 v_1，瞄准距离为 b 射向一电荷量为 Ze 的重核(如图)。设质子的质量为 m_p，重核质量非常大，可认为静止不动 求质子能接近重核的最近距离 s。

3-5-9 一宇宙飞船绕某一行星作半径为 R_o 的圆轨道运动。

(1) 飞船向其轨道切线方向发射一小型火箭，使飞船在原飞行方向获得一速度增量 v_1，以后便沿半长轴为 $4/3R_o$ 的椭圆轨道运动。求 v_1 与原轨道速度 v_o 的比值。

(2) 若沿轨道半径方向发射小型火箭，使飞船获得一径向的速度 v_2，以后同样沿半长轴为 $4/3R_o$ 的椭圆轨道运动。求 v_2 与 v_o 的比值。

3-5-10 发射一宇宙飞船去考察一质量为 m_o、半径为 R 的行星。当飞船离行星中心 $5R$ 处相对行星静止时，以相对行星的速度 v_o 发射一包仪器(如图)。仪器包的质量 m 远小于飞船的质量，要使此仪器包恰好掠擦行星表面着陆，则发射方向与飞船和行星球心连线的夹角 应是多大？

3-5-11 一质量为 m_a、半径为 r_a 的圆筒 A，与另一质量为 m_b、半径为 r_b 的圆筒 B 同轴，均可各自绕轴自由旋转。在圆筒 A 的内表面上均匀散布了薄薄的一层质量为 m 的砂子，并以 $_o$ 的角速度绕轴匀速旋转，而圆筒 B 则静止。在 $t=0$ 时，A 筒的小孔打开，砂子以恒定的速率 (单位为：kg/s)飞出而贴附在 B 筒的内壁上，若忽略砂子从 A 筒飞到 B 筒的时间，求两筒以后的角速度 $_a$ 和 $_b$ 与时间 t 的关系。

3-5-12 在一长为 $3l$ 的轻杆，距离一端为 l 的位置打一小孔，在杆的两端以及距另一端为 l 处各系一质量为 m_o 的小球。通过此孔将杆悬挂于一光滑的水平细轴 O 上(如图)。开始时轻杆静止，一质量为 m 的小铅粒以 v_o 的水平速

度射入中间的小球，并留在里面。若铅粒相对小球静止时杆的角位移可以忽略，求杆在以后摆动中的最大摆角。

3-5-13 the water flow into the sink usually in a way with combination of rotation are both originated from the Coriolis force. Such effect is even amplified to a very large scale. For example, once air has been set in motion by the pressure gradient force, it undergoes a deflection from its path with respect to the earth. As air moves from high to low pressure in the northern hemisphere, it deflects to the right, while in the southern hemisphere air moving from high to low pressure deflects to the left. The amount of the deflection is dependent on both the speed at which the air is moving and its latitude.

3-6 多粒子系统

本节讨论多粒子系统的运动演化问题。多粒子系统最主要的力学特征是系统内部其它粒子之间的相互作用导致各个粒子的运动之间存在互相关联。这种关联给求解多粒子系统的运动演化问题带来非常大的困难。事实上，当系统中的粒子数目达到三个时，通常就无法严格求解甚至其中任何一个粒子的运动方程，在很多情形中即使近似求解也非常困难。

应该说我们对多粒子系统行为中的规律和秩序的理解至今仍很肤浅，而另一方面，我们所关注的多粒子系统行为大多是一些特定的集合运动形态(模式)，甚至只是系统整体的运动所呈现出的宏观(平均)效应。本节我们将采用最简单和直接的平均方法来讨论系统的整体行为，并利用动量，能量，角动量守恒简化对系统行为的分析。以后的章节中将会介绍一些特定的集合运动形态。

3-6-1 质心和质心运动方程

考虑一个包含有 N 个粒子的系统，其中每一个粒子的行为由相应的运动方程控制

$$\frac{d\mathbf{p}_i}{dt} = \mathbf{F}_i^{\text{ext}} + \sum_{j\neq i}^{N} \mathbf{f}_{i-j} \quad (i = 1,2,\ldots,N) \tag{3-6-1}$$

其中 $\mathbf{p}_i$ 代表粒子 i 的动量，而 $\mathbf{F}_i^{\text{ext}}$ 和 $\mathbf{f}_{i-j}$ 分别代表系统外部和系统内粒子 $j\,(\neq i)$ 作用在粒子 i 上的力。实际求解过程中困难的是对内力 $\mathbf{f}_{i-j}$ 的描述，因为每一个粒子都总是处于运动之中，这导致力 $\mathbf{f}_{i-j}$ 也总在不断变化[1]。

若我们不考虑单个粒子的运动细节，而只关注于系统的整体运动，则问题将得到极大简化。系统整体运动状态的演化可由全部粒子的行为总和描述：

$$\frac{d}{dt}\sum_i \mathbf{p}_i = \sum_i \mathbf{F}_i^{\text{ext}} + \sum_{i\neq j}(\mathbf{f}_{i-j} + \mathbf{f}_{j-i}) \tag{3-6-2}$$

其中我们将系统中每一对质点的相互作用(作用力和反作用力)归并在一个项中。由牛顿第三定律可知，每一对作用力和反作用力的矢量和为零，(3-6-2)式中的作用力因而只剩下外力部分

$$\frac{d\mathbf{p}}{dt} = \sum_i \mathbf{F}_i^{\text{ext}} = \mathbf{F}^{\text{ext}} \tag{3-6-3}$$

其中 $\mathbf{p}$ 是系统中所有粒子的动量总和。(3-6-3)式与单一质点的运动方程形式相同，由于系统内各质点的受力及初始运动状态各不相同，系统各部分的的演化也不相同，因而不能简单地等效于 $\mathbf{F}^{\text{ext}} = m\mathbf{a}$ (m 是体系的总质量)。但实际观测发现，任意一个多质点系统中都存在一个特殊位置C，能使

$$\mathbf{F}^{\text{ext}} = m\mathbf{a}_C \tag{3-6-4}$$

成立，其中 $\mathbf{a}_C$ 是该点的加速度。例如，图 3-6-1 显示了一个跳水运动员在空中的

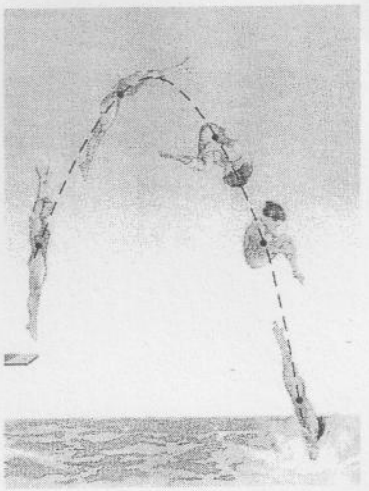


图 3-6-1 跳水运动员的质心运动轨迹与同样速度和角度投射的抛体轨迹完全相同。

[1] 读者可以考虑一个具体的三体系统来体会其中的复杂性。

一系列姿态。其身体每个部分的运动都相当复杂，但其中一点的运动轨迹与同样速度和角度投射的抛体轨迹完全相同。为求解出该特殊点的位置 $\mathbf{r}_C$ 及其含义，将(3-6-3)式写成

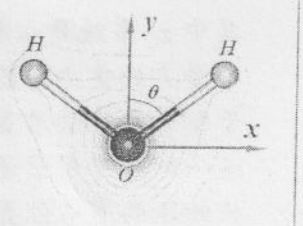
$$\mathbf{F}^{\text{ext}} = \frac{d}{dt}\sum_i \mathbf{p}_i = \frac{d^2}{dt^2}\sum_i m_i\mathbf{r}_i \equiv m\frac{d^2\mathbf{r}_C}{dt^2} \tag{3-6-5}$$

由此得到位置矢量 $\mathbf{r}_C$ 为

$$\mathbf{r}_C = \frac{1}{m}\sum_i m_i\mathbf{r}_i \tag{3-6-6}$$

这个位置矢量表现为系统中所有质点位置矢量的加权平均，其中的权重就是各个位置或区域中的质量占系统的总质量的百分比 m_i/m，因而该位置称为系统的**质量中心**，简称为**质心**。

例3-6-1 水分子中氧原子和两个氢原子构成的V-型(见例图)。独立水分子(处于气态)的两个O-H键的长度为9.57 nm，两个O-H键夹角为104.5°。求水分子质心位置。



例 3-6-1 图

解： 取分子所在平面为坐标系x-y平面，以氧原子的位置为坐标系原点，以两个O-H键夹角的平分线y轴(见图)。由分子的几何对称特征可知，水分子质心一定在y-轴上，具体坐标值满足

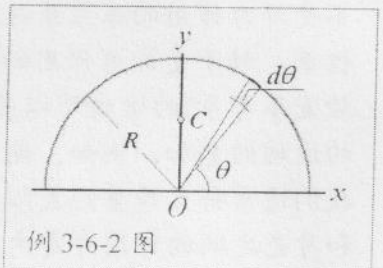
$$y_C = \frac{2m_{\text{H}}\times L_{\text{OH}}\times\cos\theta}{m_{\text{O}}+2m_{\text{H}}} = \frac{2\times 9.57\times\cos 52.2^\circ}{18} = 0.65\text{nm}$$ ◀

虽然实际物体都是由离散的原子构成，但即使对于一枚一分硬币大小的物体所含的原子数也达到 10^{20} 这样的数量级。因此，可以将宏观物体近似看成质量连续分布的系统进行计算。设想将系统可以看成为许多小体积元 ΔV 的集合，其相应的质量为 $\Delta m_i = \rho(\mathbf{r}_i)\Delta V$，其中 $\rho(\mathbf{r}_i)$ 为物体在 $\mathbf{r}_i$ 处的局域密度。则系统质心位置的定义式(3-6-6)可"近似"表示为

$$\mathbf{r}_C = \frac{\int \mathbf{r}dm}{\int dm} = \frac{1}{m}\int_V \mathbf{r}\rho(\mathbf{r})dV \tag{3-6-7}$$

质量分布均匀的物质系统其质心位置就是该物体体积的平均位置。如果系统的形状具有某些对称性，如球形，圆柱形等，该平均位置就是系统的几何中心。

例 3-6-2 计算一半径为 R 的均质半圆弧的质心位置。



例 3-6-2 图

解： 取圆弧所在平面为坐标系 x-y 平面，圆心为坐标系原点。由圆弧的几何对称特征可知其质心一定在 y-轴上，具体坐标值满足

$$y_C = \frac{1}{m}\int_\ell \rho y d\ell = \frac{1}{\pi R}\int_\ell y d\ell$$

对原点所张的角度为 $d\theta$ 的弧元长度可表示为 $d\ell = Rd\theta$，而该弧元在坐标系中的纵坐标为 $y = R\sin\theta$。则上述积分简化为

$$y_C = \frac{1}{\pi R}\int_0^{\pi} R^2 \sin\theta d\theta = \frac{2}{\pi}R$$

质心位置与圆心的距离为 $0.637R$。◀

(3-6-5)式表明，不论体系怎样复杂，体系质心的行为与一个质点相同，从这个意义上说，牛顿定律也适用于多质点系统。但牛顿定律所描绘的不是体系中任意一个质点的运动，而是质心的运动。而质心的存在是任意物体在一定条件下可以看成质点的物理基础。由此可知，质心并不是一个几何学或运动学概念，其性质只有在体系的运动与外力的关系中表现出来，因而是一个动力学概念。外力作用在系统的每一个质点上，单就整体运动的描述而言，相当于作用在质心上，尽管有时在质心位置并没有实际物质存在(如一个圆环和一个球壳的质心)。

当外力为重力时，质心与重心重合。多粒子系统的重力势能可以表示为系统质心位置的函数。多粒子系统的总势能为

$$U = g\sum\nolimits_i m_i z_i = mgz_C \qquad (3\text{-}6\text{-}8)$$

其中 z_C 系统质心相距地面的高度。该表达式与单个质点的重力势能表达式具有相同的形式，相当于整个系统的质量全部集中于整个系统的质心。这一结果可用于通过悬挂方法确定具有不规则形状物体的质心位置(见图 3-6-2)：用弦线的一端固定在扳手上的某一点并将其悬挂起来，当悬挂系统达到平衡时，其重力势能达到最小值，或者说系统质心处于最低位置。因此，质心一定在物体的悬垂点的正下方。在物体上沿悬线(铅垂线)的延长线画一直线，系统的质心一定处在该直线上。然后换一个物体上的悬挂点，并重复上述过程，可以在物体上画出第二条直线，则系统的质心就是这两条直线的相交点。

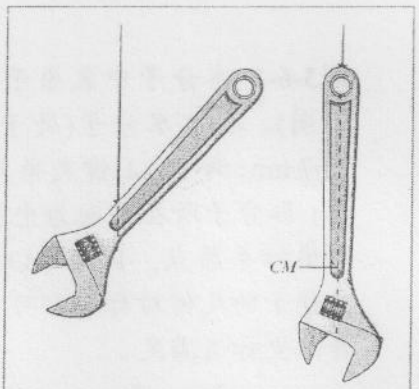


图 3-6-2 悬垂法确定物体质心

3-6-2 多质点系统的动量定理

多质点系统的总动量等于体系总质量与质心速度的乘积，质点系的动量守恒定律表示为

$$\mathbf{p} = m\mathbf{v}_C = \text{常矢量}, \quad \mathbf{v}_C = \text{常矢量} \qquad (3\text{-}6\text{-}9)$$

即对动量守恒的体系，尽管体系内各质点的动量和速度可以有各种变化，但其质心的速度始终不变。当体系动量不守恒时，质心具有加速度。

坐标系原点固定在系统质心上的参考系称为**质心参考系**，简称**质心系**。对于不受外力作用的孤立系统或所受外力的矢量和为零的力学系统，质心参考系是惯性系。对于受外力作用的体系，质心系是非惯性系。物体相对固定参考系(如实验室参考系)的运动可以表示为物体相对质心系的运动与质心系相对固定参考系的运动的叠加。例如，地球和月亮构成的系统围绕着太阳做轨道运动的问题中，我们通常将大质量的太阳看作静止不动的，并将其位置视为坐标系的原点。地球和月亮之间的作用力是内力而相互抵消，地-月系统的运动方程为

$$(m_e + m_m)\mathbf{a} = \mathbf{F}_{s\text{-}e} + \mathbf{F}_{s\text{-}m}$$

即地球和月球因受到太阳的引力，其质心绕太阳作椭圆轨道运动，而地球和月球则围绕着它们质心作椭圆轨道运动。因此，地球和月球的运动轨道都不是真正的椭圆，而是在以椭圆为主体轨道附近晃动，图3-6-3给出的是夸张的月球运动轨道示意图。

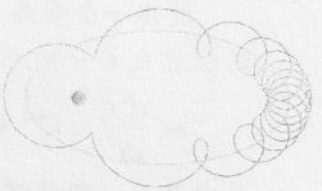
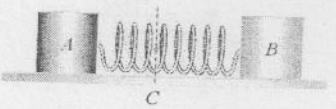


图3-6-3 月球围绕太阳轨道运动的夸张示意图。

例 3-6-3 双原子分子的振动 设双原子分子HCl在作小幅相对振动时，两个原子之间的力可用胡克定律近似描述，弹性系数为k，求该分子振动的特征频率。

解： 在分子HCl的质心坐标系中可以将该分子看成为两个振动频率相同，同步拉伸或压缩的独立弹性振子。每个振子的有效弹性系数因有效长度减小而增加，例如其中的氢原子振子的有效长度ℓ_H与整个分子振子的长度ℓ比满足关系之间满足关系$\ell_H=(m_{Cl}/m)\ell$，则氢原子振子相应的弹性常数k_H与整个分子振子的弹性常数k之间满足关系$k_H=(m/m_{Cl})k$，由此可得氢原子的振动特征频率

$$f_H = \frac{1}{2\pi}\sqrt{\frac{k_H}{m_H}} = \frac{1}{2\pi}\sqrt{\frac{k}{m_H m_{Cl}/m}} = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}} \quad \blacktriangleleft$$



例3-6-3图

3-6-3 多质点系统的角动量定理

3-5 节中我们讨论了单个质点相对于固定点的角动量和角动量定理，对于多质点系统而言，其转动动力学效应由所有质点受到的相对于固定点的总力矩决定：

$$\boldsymbol{\tau} = \sum_i \boldsymbol{\tau}_i \tag{3-6-9}$$

将作用在系统中每一质点上的外力和内力对力矩的贡献分开

$$\boldsymbol{\tau} = \sum_i \mathbf{r}_i \times \mathbf{F}_i = \sum_i \mathbf{r}_i \times \mathbf{F}_i^{ext} + \sum_{i\neq j}(\mathbf{r}_i \times \mathbf{F}_{ij}^{int} + \mathbf{r}_j \times \mathbf{F}_{ji}^{int})$$

粒子之间的相互作用力的方向一般沿两个粒子之间的连线(中心力)，因此一对粒子中每一个粒子受到对方的力矩作用之和为

$$\mathbf{r}_i \times \mathbf{F}_{ij}^{int} + \mathbf{r}_j \times \mathbf{F}_{ji}^{int} = (\mathbf{r}_i - \mathbf{r}_j)\times \mathbf{F}_{ij}^{int} = 0 \tag{3-6-10}$$

即系统中的内力对总力矩没有贡献，系统受到的总力矩全部来自于外力作用：

$$\boldsymbol{\tau} = \sum_i \mathbf{r}_i \times \mathbf{F}_i^{ext} = \sum_i \boldsymbol{\tau}_i^{ext} = \boldsymbol{\tau}^{ext} \tag{3-6-11}$$

由于质点i相对于某一固定点O的角动量变化率可作为作用在该质点上相对于O点的外力矩的度量：

$$\boldsymbol{\tau}_i = \frac{d\mathbf{L}_i}{dt}$$

多质点系统相对于固定点O的总角动量是刚体中所有粒子相对于点O的角动量之和，因此对所有质点的相对于点O的角动量定理求和：

$$\boldsymbol{\tau}^{ext} = \sum_i \frac{d\mathbf{L}_i}{dt} = \frac{d}{dt}\sum_i \mathbf{L}_i = \frac{d\mathbf{L}}{dt} \tag{3-6-12}$$

称为多质点系统相对于固定点O的角动量定理。其中$\mathbf{L}$为整个系统相对于点O的总角动量。

质心是多质点系统的一个特殊点，它具有一些其它空间点位不具有的性质。为此我们讨论一个多质点系统相对于其质心的角动量定理的形式。

设想一个多质点系统中的质点i相对于一个固定坐标系原点O和系统的质心C的位置矢量分别为$\mathbf{r}_i$和$\mathbf{r}'_i$。它在质心参照系中的运动方程为

$$m_i\frac{d^2\mathbf{r}'_i}{dt^2}=\mathbf{F}_i^{\text{ext}}+\mathbf{F}_i^{\text{int}}-m_i\frac{d^2\mathbf{r}_C}{dt^2} \tag{3-6-13}$$

其中最后一项是惯性力。用$\mathbf{r}'_i$从左面叉乘方程式(3-6-13)两边，并对i求和，得

$$\frac{d}{dt}\sum_{i=1}^{N}(\mathbf{r}'_i\times m_i\dot{\mathbf{r}}'_i)=\sum_{i=1}^{N}(\mathbf{r}'_i\times\mathbf{F}_i^{\text{ext}})+\frac{d^2\mathbf{r}_C}{dt^2}\times\sum_{i=1}^{N}m_i\mathbf{r}'_i \tag{3-6-14}$$

其中已用到内力矩互相抵消的结果。等式左边是质点组相对于系统质心的角动量的变化率，等号右边第一项为所有外力对质心的力矩之和，第二项中的求和因子正比于质心系中的质心位置坐标而等于零，表明各质点的惯性力矩互相抵消，因为质心坐标系整体以加速度$\mathbf{a}_C$运动，其中每一质点m_i都受到惯性力$-m_i\mathbf{a}_C$的作用，其合力通过质心。(3-6-14)式由此简化为

$$\boldsymbol{\tau}_C=\frac{d\mathbf{L}_C}{dt} \tag{3-6-15}$$

称为质点组相对于其质心的角动量定理，它与系统相对固定点的角动量定理形式相同。

例 3-6-4 宇宙中的大多数星系处于旋转状态，其形状则呈现为垂直于转轴的扁平型的盘状(见本例图中的银河系)。解释其可能的原因。

解： 设想银河系开始呈球形，因某种原因(请读者想象各种可能的原因)而获得一定的角动量。当银河系中的物质因引力的作用向中心聚集的过程中，因角动量守恒而导致旋转加速，但由此增大的离心力往往比引力增大得更快。因为转速增大导致动能增加，虽然这时引力势能减小，但当半径小到某一临界值后，动能的增加将大于引力势能的减小(见习题 3-6-5)。这使得球状银河系在垂直转轴的方向无法在引力作用下坍缩。但角动量守恒却并不妨碍星系沿着转轴方向的坍缩，因为对这种坍缩，角动量守恒不要求增加转速。这样，星系最终坍缩成垂直于转轴的圆盘状。在沿轴向坍缩过程中减少的引力势能将通过碰撞以辐射等形式释放掉。◀

例3-6-4图 银河系

旋涡与飓风

如果将一个装有一定量的水池(如厨房的水槽，浴缸等)底部的塞子拔去，池子底部的水开始会直接从漏洞口流出去，而上部的水面保持为宁静的平面。但这种情形一般维持不了很久，池子中的水就会进入一种新的流动状态，水面出现旋转，并向中间凹陷，最后出现一个下旋的孔洞，称为**涡旋**(见图 3-6-4)。

涡旋现象起因于水体受到外部扰动而产生一定的角动量，如果不存在人为扰动，水体在向中心流动的过程中也会因地球的自转而受到科里奥利力 $2m\boldsymbol{\omega}\times\mathbf{v}$，

导致水体旋转[2]。由于除了水与水池表面间的微小摩擦效应，并不存在其他作用在水体的外力矩，因而水体旋转的角动量基本守恒。一块小的水体以螺旋形向池底洞的方向移动时所具有相对于该竖直转轴的角动量 $\mathbf{L} = \mathbf{r} \times m\mathbf{v}$ 是守恒的(由于 $\mathbf{r}$ 和 $\mathbf{v}$ 互相垂直，该角动量的数值为 $L = mvr$)。当一小块水体在螺旋行进过程中越来越靠近轴线，即 r 越来越小，它就会越转越快以保证角动量守恒。而提供旋转向心力的是水体元之间的分子力。越靠近转轴的水体，所需要的向心力越大，当所需要的向心力超过了水的抗张强度，水体就会断裂而呈现出涡旋。

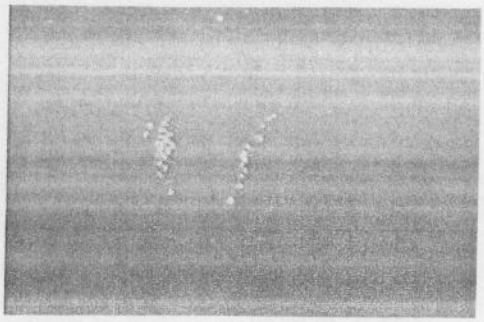


图 3-6-4 水体旋转导致下旋孔洞—涡旋

涡旋也可以发生在大气中，而且可以表现为非常大的尺度。因为大气中经常出现上升气流，如果因为某些原因使上升气流中形成很小的旋转运动，就很容易形成大气中的涡旋，通常称为**气旋**。例如热带海洋上空，由于海水蒸发很快，空气流动剧烈而且复杂，很容易出现热带气旋[3]。飓风登陆后向高纬度地区前进，其能量因地面摩擦而消耗，不像在海面上那样容易得到补充，因此会逐渐减弱并消失。通常一场飓风的寿命为几天到十几天。1981 年的飓风"约翰"在太平洋东北部和西北部共持续了 31 天。目前观测到的维持时间最长的飓风发生在木星上，时间已经延续了 300 多年。图 3-6-7(*b*)为旅行者 1 号飞经木星附近所拍摄的照片，图中的大红斑被认为是木星上的飓风，其大小是地球尺度的三倍。

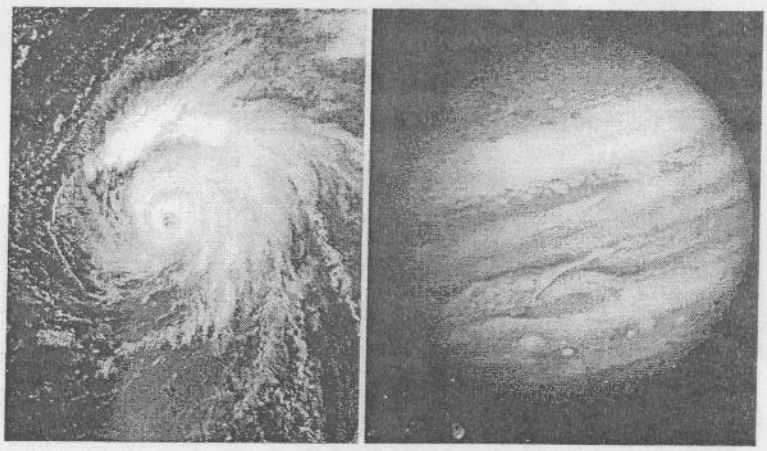


图 3-6-7 (*a*) 比尔飓风是记录到的最大大西洋热带气旋之一，最大直径达 740 公里。左图是 2009 年 8 月 20 日美国宇航局 Terra 卫星拍摄的抵达波多黎各东北部的比尔飓风，持续风速达每小时 193 公里。(*b*) 木星上的飓风所形成的大红斑。

---

[2] 由于地球的自转角速度很小，这种力实际上很微弱，只有在其它扰动都可以忽略的情况下才起作用。由科里奥利力引起的涡旋有一个特征，南、北半球的转动方向相反。

[3] 对气旋通常按其中心附近的风速分级，风速 17 米/秒以下称为**热带低压**，17~33 米/秒的称为**热带风暴**。33 米/秒以上，不同地区有不同的名字：太平洋东部和大西洋称为**飓风**，太平洋西部称为**台风**，印度洋则称为**旋风**。离地球几千公里的卫星上可以清晰地看到这种天气现象，其半径可达几百甚至上千公里。飓风中央是一小块较为平静的区域，称为"眼"。台风眼或飓风眼的直径通常为几十公里，其中风力较弱，一般不下雨。而眼的边缘则是整个飓风中风力最强的部位—"眼壁"(所以描述飓风时称"中心附近最大风力"，而不是"中心最大风力")

3-6-4 多粒子系统的能量

多质点系统的总动能是系统中单的质点的动能的总和：

$$K = \tfrac{1}{2}m_1v_1^2 + \tfrac{1}{2}m_2v_2^2 + \ldots + \tfrac{1}{2}m_nv_n^2 \tag{3-6-16}$$

不同于动量情形：多粒子系统的总动量可表示为一个单一粒子的动量，多粒子系统的总动量通常不能简化为一个具有全部系统质量的单粒子的动能，而是大于这个能量。为说明这一概念，记每一个粒子的速度 $\mathbf{v}_i$ 为该粒子相对于质心的速度 $\mathbf{v}_{iC}$ 与质心相对于某一惯性系的速度 $\mathbf{v}_C$ 的矢量和：

$$\mathbf{v}_i = \mathbf{v}_{iC} + \mathbf{v}_C \quad (i = 1, 2, \ldots, N) \tag{3-6-17}$$

多粒子系统的总动能可以表示为

$$\begin{aligned} K &= \sum_i K_i = \sum_i \frac{1}{2}m_iv_i^2 = \sum_i \frac{1}{2}m_i(\mathbf{v}_{iC} + \mathbf{v}_C)^2 \\ &= \sum_i \frac{1}{2}m_iv_{iC}^2 + \mathbf{v}_C \cdot \sum_i m_i\mathbf{v}_{iC} + \frac{1}{2}v_C^2\sum_i m_i \end{aligned} \tag{3-6-18}$$

其中第一项是系统中所有质点相对于系统质心的动能总和，通常称为系统的内部动能，记为 K_{int}。第二项为零，因为其中的求和因子

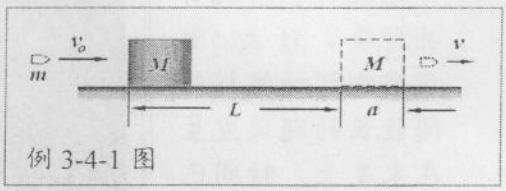
$$\sum_i m_i\mathbf{v}_{iC} = \frac{d}{dt}\sum_i m_i(\mathbf{r}_i - \mathbf{r}_C) = \frac{d}{dt}\left[\sum_i m_i\mathbf{r}_i - \mathbf{r}_C\sum_i m_i\right] = \frac{d}{dt}(m\mathbf{r}_C - m\mathbf{r}_C)$$

系统中全部粒子的总动能为质心的动能 $\tfrac{1}{2}Mv_C^2$ 和质心系中的内部动能 K_{int} 之和

$$K = K_{\text{int}} + K_C \tag{3-6-19}$$

如果作用在多粒子系统中的外力和内力都是保守力，系统具有势能，表示为所有粒子的位置的函数。而系统的总能量守恒。

例 3-6-5 从多质点系统的角度重解例 3-4-1。



例 3-4-1 图

解：驱动木块运动的摩擦力是内力，利用内力做功与参考系无关的性质，在木块参考系上分析摩擦力的功，然后在地面参考系中用质点系动能定理，可立即求得上述结果。对木块和子弹构成的系统而言，求解木块的行进距离需要知道摩擦力以及摩擦力作用时间或木块的末速度 u。由于子弹和木块作为一个系统不受外力作用，其水平方向动量守恒：$mv_o = Mu + mv$。由此可以求得木块的末速度：$u = (m/M)(v_o - v)$。求解摩擦力一般是分析力产生的效果。我们从摩擦力做功的角度出发，设摩擦力 f 对木块作用距离为 L，由动能定理可得

$$fL = \tfrac{1}{2}Mu^2$$

摩擦力同时也作用在子弹上，但对子弹作用的距离为 $L+a$，由动能定理可得

$$-f(L+a) = \frac{m}{2}v^2 - \frac{m}{2}v_o^2$$

将上两式相加可得

$$-fa = \frac{m}{2}v^2 - \frac{m}{2}v_o^2 + \frac{M}{2}u^2$$

由此可求得摩擦力 f

$$f=\frac{1}{a}\left[\frac{m}{2}v_o^2-\frac{m}{2}v^2-\frac{m}{2}\left(\frac{m}{M}\right)^2(v_o-v)^2\right]=\frac{m}{2a}\left[v_o^2-v^2-\left(\frac{m}{M}\right)^2(v_o-v)^2\right]$$

解得相关过程中木块的位移量 L

$$L=\frac{Mu^2}{2f}=\frac{a(m^2/M)(v_o-v)^2}{m[v_o^2-v^2-(m/M)(v_o-v)^2]}=\frac{am(v_o-v)^2}{M(v_o^2-v^2)-m(v_o-v)^2]}$$

说明：就整个体系(子弹和木块)而言，动摩擦力对体系中的木块做正功，对子弹做负功，但负功大于正功，对整个系统而言，摩擦力作的总功总是负的，使系统的总动能减少。◀

3-6-5 两体问题

一般情况下，系统中的相互作用的粒子数目达到三个，相应的动力学问题就难以严格求解，但对于两个相互作用粒子系统可以将问题转化为一个单粒子系统。

一个两体系统的运动方程

$$m_1\frac{d^2\mathbf{r}_1}{dt^2}=\mathbf{F}_{12} \tag{3-6-20a}$$

$$m_2\frac{d^2\mathbf{r}_2}{dt^2}=\mathbf{F}_{21} \tag{3-6-20b}$$

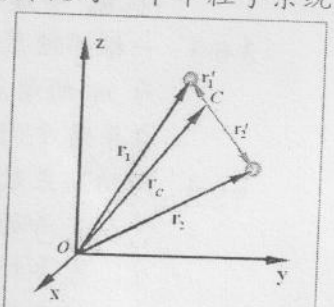


图3-6-3

引入该系统的质心坐标和相对坐标

$$\mathbf{r}_C=\frac{m_1\mathbf{r}_1+m_2\mathbf{r}_2}{m_1+m_2},\ \mathbf{r}=\mathbf{r}_1-\mathbf{r}_2 \tag{3-6-21}$$

将方程(3-6-20a)和(3-6-20b)相加，并利用 $\mathbf{F}_{12}=-\mathbf{F}_{21}$，可得系统的质心运动方程

$$m_1\frac{d^2\mathbf{r}_1}{dt^2}+m_2\frac{d^2\mathbf{r}_2}{dt^2}=\frac{d^2}{dt^2}(m_1\mathbf{r}_1+m_2\mathbf{r}_2)=(m_1+m_2)\frac{d^2}{dt^2}\left\{\frac{m_1\mathbf{r}_1+m_2\mathbf{r}_2}{m_1+m_2}\right\}=m\frac{d^2\mathbf{r}_C}{dt^2}=0$$

而将运动方程(3-6-20a)和(3-6-20b)分别乘上质量因子 m_2 和 m_1 后相减

$$m_1m_2\frac{d^2\mathbf{r}_1}{dt^2}-m_1m_2\frac{d^2\mathbf{r}_2}{dt^2}=m_2\mathbf{F}_{12}-m_1\mathbf{F}_{21}=(m_2+m_1)\mathbf{F}_{12}$$

定义参量

$$\mu=\frac{m_1m_2}{m_1+m_2} \tag{3-6-12}$$

称为两体系统的**折合质量**，可以得到物体 1 相对于物体 2 的运动方程

$$\mu\frac{d^2\mathbf{r}}{dt^2}=\mathbf{F}_{12} \tag{3-6-13}$$

这样就将两体问题简化为一个物体相对于另一个物体的单粒子运动方程。例如，例 3-6-3 中的双原子分子的振动运动方程可以直接写为 $\mu\ddot{x}=-kx$，则相应的振动频率为：$\omega=\sqrt{k/\mu}$。

3-6-6 微扰问题*

思考题

3-6-1 考虑一个三体系统，如太阳、地球、月亮构成的系统的运动方程，体会并说明求解该运动方程所存在的具体困难和其中的复杂性。

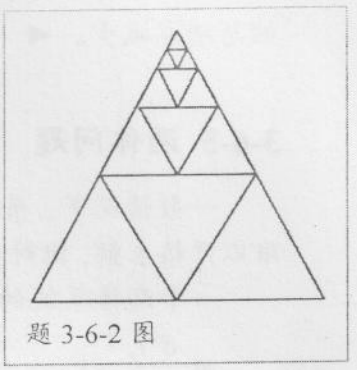
3-6-2 一个国家或地区的人口中心、工业中心和金融中心应该怎样定义。

3-6-3 宇宙中的星系大多处于旋转状态，并呈扁平的圆盘状，盘面垂直于旋转轴。想象星系形成这种状态的各种可能原因。

习题

3-6-1 求均匀半球体的质心。

3-6-2 用粗细均匀的韵致细铁丝制成无限内接等边三角形框架，最外层等边三角形边长为 a，里面的各等边三角形边长分别为 $\frac{1}{2}a$，$\frac{1}{4}a$，…，线密度为 λ。求质心离大三角形定点的距离。(机械 安建平提供)


题 3-6-2 图

3-6-3 一根弹性系数为 k 的弹簧两端各系有一个质量为 m 的质点，并用线绳悬垂于天花板(见图)。在系统平衡的时刻剪断悬挂线，讨论系统以后的运动状态。

3-6-4 用劲度系数为 k 的弹簧将质量为 m_A 和 m_B 的物体 A 和 B 连接，放在光滑的水平面上。A 紧靠墙，在 B 上施力将弹簧压缩 a 长度。撤去外力后，求弹簧和 A 及 B 组成的系统的最大质心加速度，以及质心速度的最大值。

3-6-5 用劲度系数为 k 的弹簧连接的质量为 m_A 和 m_B 的物体 A 和 B 竖直地放置在地面上，与地面接触的是 B。如果用力 F 向下压弹簧上面的物体 A 使弹簧压缩。F 多大可以使物体 B 在撤去 F 后脱离地面，这时 B 的加速度是多大？整个系统的整体加速度什么时候达到最大？什么时候达到 0?

3-6-6 光滑的水平面上有一质量为 M 的三角形木块，木块的斜面与地面之间的夹角为 θ。一质量为 m 的物体从高 h 处自静止沿斜面无摩擦地下滑到地面。分别以 m、M 和地面为参考系，计算下滑过程中 M 对 m 的支撑力 N 及其反作用力 N' 所做的功，并证明二者之和与参考系的选择无关，总是为 0。

3-6-7 求圆心角为 2θ 的一段均匀圆弧的质心。

3-6-8 A box of mass M has an iron block of mass m hanging from its ceiling and a magnetic block of mass m_2 fixed at the bottom (see Figure). Suppose that the magnetic block is set free from the confinement, what is the tension in the rope suspending the box.

3-6-9 用长为 l 的轻绳连结的两个质量分别为 m_1 和 m_2 的小球在光滑的水平面上运动。开始时，线绳正好拉直，m_1 和 m_2 的运动方向相同且垂直于连线，速率分别为 v_1 和 v_2 ($v_1 > v_2$)。求系统相对质心的角动量以及线中的张力。

3-6-10 一根轻绳跨过具有光滑水平轴的定滑轮(质量可忽略)，两个质量分别为

m_1 和 m_2 的人各抓住绳子的一端。开始时，两人与定滑轮水平轴的竖直距离分别为 h_1 和 h_2，他们同时开始向上爬，并同时到达该水平轴。求他们爬绳所经历的时间。

3-6-11 两个滑冰者质量都是 60 kg，在两条相距 10m 的平直跑道上以 6.5m/s 的速率沿相反方向匀速滑行 当他们之间的距离恰好等于 10 m 时，他们分别抓住一根长 10m 的轻绳两端。(1) 求他们抓住绳子前后相对于绳子中点的角动量；(2) 两人都用力往自己这边拉绳子，当他们之间的距离为 5.0m 时，求各自的速率和绳中的张力；(4) 计算每个人在拉绳过程中所作的功。

3-6-12 两质量均为 m 的质点相互吸引，引力的大小与两者之间距离 r 的平方成反比：$F=-k/r^2$，k 为正的常量。开始时，$r_o=a$，一质点静止，另一质点以垂直于两者连线的方向以速率 $v_o=\sqrt{8k/ma}$ 开始运动。求在以后的运动过程中，两者之间的最大距离和最小距离。

3-6-13 银河系中的物质因引力的作用向中心聚集的过程中系统的势能减小，但因角动量守恒也导致旋转加速，动能增加。由此增大的离心力往往比引力增大得更快。证明当半径小到一临界值后，动能增加将大于引力势能的减小。

3-6-14 质量为 m = 0.4 kg 的小球由一根穿过半顶角 $\alpha=37°$ 的光滑圆锥面顶部小孔 O 的细绳连接，并以速率 v_o = 0.5m/s 作匀速圆周运动，小孔到小球的绳长 r_o = 1.5m。现将绳的另一端缓慢向下拉，直至小球与锥面脱离。试求此过程中拉绳力所作的功(g = 10 m/s)。

3-6-15 质量皆为 m 的两珠子可在光滑轻杆上自由滑动，杆可在水平面内绕过 O 点的光滑竖直轴自由旋转。原先两珠子对称地位于 O 点的两边，与 O 相距 a，在 t = 0 时刻，对杆施以冲量矩，使杆在极短时间内即以角速度 ω_0 绕竖直轴旋转，求 t 时刻杆的角速度 ω、角加速度 α 及两珠子与 O 点的距离 r。

3-6-16 质量均为 m 的小球 1、2 用长为 $4a$ 的细线相连，以相同速度 v 沿着与线垂直的方向在光滑水平台面上运动，线处于伸直状态。在运动过程中，线上距离小球 1 为 a 的一点与固定在台面上的一竖直光滑细钉相碰，设在以后的运动过程中两球不相碰。求：(1)小球 1 与钉的最大距离；(2) 线中的最小张力。

3-6-17 水平光滑桌面上有一小孔，穿过小孔的一条绳两端各系一质量为 m 的物体(见图)。开始时，用手握住下面的物体，桌上的物体则以 $v_o=1.5\sqrt{2gr_o}$ 的速率作半径为 r_o(即桌上部分的绳长)的匀速圆周运动，然后放手。求以后的运动过程中桌上部分绳索的最大长度和最小长度。

3-6-18 Two objects of momenta $\mathbf{p}_1 = m_1\mathbf{v}_1$ and $\mathbf{p}_2 = m_2\mathbf{v}_2$, and position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$. Show that in the special case that the center of mass of the two bodies is at rest at the origin (that is, $\mathbf{p} = \mathbf{p}_1 + \mathbf{p}_2 = 0$ and the position of the center of mass is $\mathbf{R} = 0$), the sum of the angular momenta of the two objects about the center of mass equals the angular momentum of a single object of mass $\mu = m_1m_2/(m_1+m_2)$ rotating in circular motion about the origin at a distance $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$.

3-6-19 A man inside a rapidly rotating bottomless cylinder seems to be stuck firmly to

the inner wall. Express the rotational speed ω in terms of the relevant parameters, at which the man begins to slip down.

3-6-20 A conical pendulum is composed of a mass m tied at one end of a light chord whose other end is fixed at a ceiling. Given the proper push, the system can swing in a circle maintaining the same angle θ and height h throughout its motion (see Figure). Express the angular speed in terms of the string angle θ and the string length L.

3-6-21 A white belt of a horizontal conveyor is moving with a constant speed v = 30m/s. A small carbon block placed on the belt with zero initial velocity relative to the ground. The block will slip a bit before moving with the belt, leaving a black mark on the belt. Find the length of the mark if the coefficient of kinetic friction between the carbon block and the belt is 0.20 and the coefficient of static friction is 0.30.

3-7 碰撞

碰撞是一种典型的、也是常见的物体相互作用过程。其典型特征是过程的时间短，相应的作用强度大，系统的运动状态在很短的时间内发生了明显的变化。我们日常所熟悉的各类交通事故，各种体育运动项目无不与碰撞密切相关。通过对碰撞现象的研究可以获得许多关于碰撞物体的内部结构特征及碰撞物体之间相互作用特性的信息，特别是在原子，亚原子为代表的微观领域中，有关微观粒子相互作用(强度，作用范围等)的信息事实上都是通过碰撞的实验数据获得的，碰撞至今仍是探索微观物质的结构和行为的唯一实验手段[1]。例如，原子的核结构模型是卢瑟福分析 α 粒子与金箔中的原子碰撞的实验数据结果发现的，而两个电荷之间的库仑作用在极小尺度内对平方反比定律的偏离则是通过中微子与电子的相互碰撞实验所确定的，而原子核中发生的自发衰变也是一种碰撞现象，对这种现象的深入分析则揭示出了一种全新的相互作用—弱相互作用。

"碰撞"过程中两个物体在极短时间内发生剧烈的相互作用这种特征导致在碰撞过程中很难观察到物体运动状态发生的变化细节，一般能够看到的只是系统碰撞前后运动状态的变化。另外，由于"碰撞"过程极短，能形成可观测效应就意味着相互作用非常强(图3-7-1)，即使系统在过程中可能还受到其它外力(例如重力，各种阻尼力)的作用，这些外力在"碰撞"期间对系统的影响相对于碰撞作用一般都可以忽略。因而可以认为整个系统碰撞前后的总动量守恒:

$$\mathbf{p}_i^{(1)} + \mathbf{p}_i^{(2)} = \mathbf{p}_f^{(1)} + \mathbf{p}_f^{(2)} \tag{3-7-1}$$

而每个物体碰撞前后的动量变化为

$$\Delta\mathbf{p}^{(i)} = \mathbf{I}^{(i)} \equiv \int \mathbf{F}^{(i)} dt = \overline{\mathbf{F}}^{(i)} \Delta t \tag{3-7-2}$$

由于通常并不清楚$\mathbf{F}(t)$的细节，因而动量守恒定律是描述、分析碰撞现象的最基本的方法。

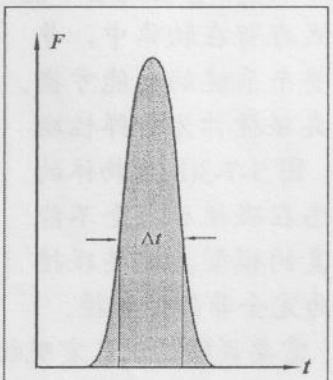


图 3-7-1 物体碰撞过程中相互作用随时间变化示意图。

3-7-1 碰撞的类型及其特征

碰撞过程除了涉及两个物体间的动量转移，还涉及能量的交换和变化。虽然在各种碰撞过程中，系统的总动量和总能量总是守恒的，但从观测角度而言，能量守恒并不总能够对预言碰撞后的物体的运动状态提供有用的信息。因为物体碰撞前所具有的运动能量并不总是能在碰撞后以宏观运动形式完全表现出来。因此，定量区分物体之间不同的碰撞类型还需要利用其它的碰撞特征，即需要有可观察的物理效应来定义相应的物理量。

物体的碰撞可以根据在碰撞前后物体的质量和内部结构及运动状态的变化情况分为两种类型：一是除了运动状态，物体本身在碰撞前后完全没有任何变化；第二是物体本身在碰撞前后有所改变。这种变化可以是物体的质量在过程中发生了转移或亏损，或者物体的内部结构发生了变化。例如，物体的外部形态发生的

[1] 微观粒子的碰撞通常被称为散射，相应的实验称为散射实验。

不同改变，比较极端的则为碰撞导致碰撞物体粘合在一起，或者物体碰撞后碎裂成许多碎片。

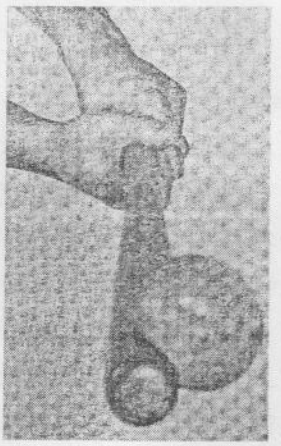


图 3-7-2 球棒和棒球碰撞时球所发生的形变

从物体碰撞时的宏观形态上看，两个物体因碰撞而发生的相互作用总是伴随着物体因互相接触而发生的形变(见图 3-7-2)，因而，碰撞物体之间的相互作用可以利用物体的弹性以及碰撞时的形变程度等特征或参量作唯像描述。如果各种形变的尺度都在碰撞物体的弹性限度以内，物体的形态在碰撞后可以得到完全的恢复。这样的碰撞过程所伴随的是碰撞物体的整体动能和弹性势能之间的转换。物体碰撞后形态的完全复原也就意味着碰撞物体动能的完全恢复。或者说，碰撞前后整个系统的动能守恒。这类碰撞称为**弹性碰撞**。弹性碰撞可以用图 3-7-3(a)中的模型来说明，其中质点物块用以携载物体的动能，无质量轻质弹簧携载碰撞形变时的弹性势能。

如果形变的尺度超出了碰撞物体的弹性限度，物体的形态在碰撞后只能得到部分恢复，甚至完全不能得到恢复，碰撞物体原来的部分动能就会以形变势能的形式存留在物体中，导致整个系统的动能亏损。这类碰撞称为**非弹性碰撞**。图 3-7-3(b)为物体的形态在碰撞后完全不能恢复的模型，这类碰撞称为**完全非弹性碰撞**。

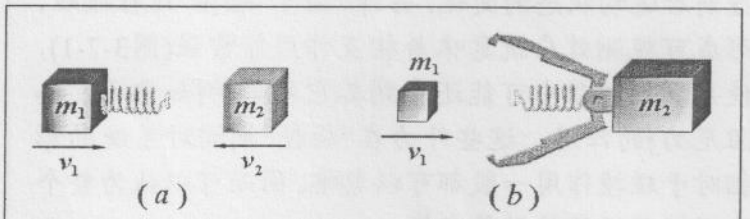


图 3-7-3 完全非弹性碰撞和完全非弹性碰撞模型

需要说明的是，宏观物体的碰撞和微观粒子的碰撞有很大的不同。在速度不是非常大的微观粒子的碰撞过程中，一般不会涉及到碰撞粒子的内部结构和内部运动状态的变化，因而碰撞前后系统的总动能守恒，碰撞是完全弹性的[2]。而宏观物体之间的碰撞通常表现为接触相互作用，碰撞时能量发生转移的具体物理图像可以用物体直接接触时发生的形变来描述。

显然，我们可以根据这种结构的改变在碰撞结束后能有多大的比例得到复元程度来表征各种碰撞之间的差异。下面先对弹性与非弹性这两类碰撞的一些特征做一些分析。为了帮助读者从最简单的例子中获得有关碰撞的一些直觉认识和相关概念，我们先讨论最简单的一维碰撞(对心碰撞)，所得到的结论对下一小节中更一般的碰撞也适用。

弹性碰撞

如果两个物体在碰撞前后没有发生质量转移，系统的运动能量也没有发生变化，这类碰撞称为弹性碰撞。或者说，弹性碰撞是以动量守恒和动能守恒为标志

[2] 微观粒子的结构改变或内部运动状态的变化一般需要很高的能量(见第 9, 10, 11 章)，因此发生动能亏损的非弹性散射情况一般涉及高速粒子的碰撞，而粒子的结构变化也不再能用弹性形变这类连续参数来描述。

的碰撞。

设两个质量分别为 m_1 和 m_2 的质点分别以初速度 u_1 和 u_2 发生弹性碰撞，则两个质点的末速度 v_1 和 v_2 可由碰撞前后所满足的动量守恒和动能守恒所对应的两个方程给出

$$m_1u_1 + m_2u_2 = m_1v_1 + m_2v_2 \tag{3-7-3}$$

$$½m_1u_1^2 + ½m_2u_2^2 = ½m_1v_1^2 + ½m_2v_2^2 \tag{3-7-4}$$

两个物体碰撞后的速度可以用它们的初速度表示出来：

$$v_1 = \frac{m_1 - m_2}{m_1 + m_2}u_1 + \frac{2m_2}{m_1 + m_2}u_2 \tag{3-7-5}$$

$$v_2 = \frac{2m_1}{m_1 + m_2}u_1 + \frac{m_2 - m_1}{m_1 + m_2}u_2 \tag{3-7-6}$$

为了让读者对上述碰撞特征有一个最为直觉的印象，我们考虑一例最简单，也是最典型的碰撞：一个运动物体和静止物体的碰撞：$u_2 = 0$。这时，通常称粒子 m_1 和 m_2 为入射粒子和靶粒子。下面分别就三种质量分布的极端情况进行讨论

(1) $m_1 = m_2$：这时有 $v_1 = u_2 = 0$，和 $v_2 = u_1$。即相同质量物体发生弹性碰撞的结果是两个物体发生速度交换，显然，这也伴随着最大的动能转移过程。桌球中经常可以看见这样的事例。

(2) $m_1 \gg m_2$：这时有 $v_1 \approx u_1$ 和 $v_2 \approx 2u_1$。即入射物体的速度没有什么改变，而散射物体获得了入射物体速度两倍的增量。例如，一辆高速行驶的重型卡车撞上微型小车，或者用高尔夫球杆击球。

(3) $m_1 \ll m_2$：这时有 $v_1 \approx -u_1$ 和 $v_2 \approx u_2 = 0$。即靶粒子几乎维持其原来的静止状态不变，而入射物体则以原来的速率沿反方向弹回。例如壁球撞击墙壁，篮球撞击地球后都表现出这样的特征。

对于碰撞，通常关注的另一个问题是两个物体碰撞前后动能的转移程度，由于动能的大小和观测者所在的参考系有关，我们具体考虑一个粒子撞击另一个静止粒子的射靶参考系，用碰撞后靶粒子 m_2 的动能 ΔK_2 与碰撞前入射粒子 m_1 的动能 K_1 之比来描述这一特性

$$\eta = \frac{\Delta K_2}{K_1} = \frac{m_2v_2^2}{m_1u_1^2} = \frac{4m_1m_2}{(m_1 + m_2)^2} = \frac{4m_1/m_2}{(1 + m_1/m_2)^2}$$

显然，两个碰撞粒子的质量越接近，动能的转移率越高，靶粒子获得的能量也越多。这一结果有许多重要的应用；例如，在核裂变反应中能有效触发核裂变的是速度较低的所谓热中子，而裂变所释放出的中子都是能量很高的快中子，为了在反应堆中让这些快中子触发进一步的核裂变，使核反应能持续进行，需要对这些快中子进行减速，而最有效的减速是选择与中子质量尽可能接近的元素作为靶物质(称为减速剂)，通常使用的靶物质是重水(氘)或石墨(碳)。

例 3-7-1 一个质量为 m 的三角块静止放置在一光滑水平面上，一个相同质量的物体以初速度 v 冲上三角块的斜面。再从斜面滑下。设所有接触面都是无摩擦的光滑面，求物体从斜面滑下后的最终速度。

解：如果从牛顿运动方程的角度分析，会涉及许多力在不同方向的分解问题，甚

至还可能涉及惯性力。但将这一问题理解为完全弹性碰撞，立刻可以知道两者只是交换了速度而已。因此物体从斜面滑下后的最终速度为零。◀

非弹性碰撞

我们先讨论非弹性碰撞中的一个极端的例子：物体因碰撞而造成的形变是完全塑性的，这类碰撞称为**完全非弹性碰撞**。从物体的形态看，两个物体碰撞后以相同的速度运动，即两个物体的接触面不再分开。设两个发生完全非弹性碰撞的质点 m_1 和 m_2 的初速度分别为 u_1 和 u_2，则由动量守恒

$$m_1u_1 + m_2u_2 = (m_1 + m_2)v$$

可以得到两个物体的共同末速度

$$v = \frac{m_1u_1 + m_2u_2}{m_1 + m_2} \tag{3-7-7}$$

另外，系统在这类碰撞中动能损失达到最大，碰撞前后系统的动能损失为

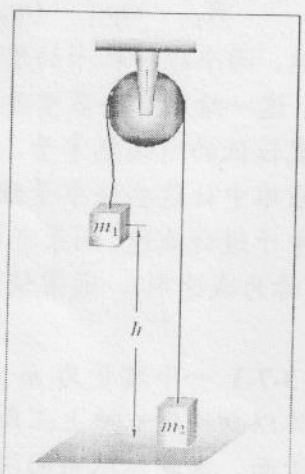
$$\Delta K = \frac{1}{2}m_1u_1^2 + \frac{1}{2}m_2u_2^2 - \frac{1}{2}(m_1 + m_2)v^2 = \frac{m_1m_2}{2(m_1 + m_2)}(u_1 - u_2)^2 \tag{3-7-8}$$

当 $u_2 = 0$ 时，系统的动能损失为

$$\Delta K = \frac{1}{2}m_1u_1^2\frac{m_2}{m_1 + m_2} = K_1\frac{m_2}{m_1 + m_2} = \frac{K_1}{1 + m_1/m_2}$$

非弹性碰撞有两个很重要的应用，一是通过将外部的能量通过碰撞变成另一个物体的内能，例如将一个处于低能量状态的原子通过碰撞激发到高能量状态。二是将系统内部的能量释放出来加以利用。由上式可知，要通过碰撞获得尽可能大的激发能，入射粒子应具有较大的初始能量，而且要求其质量要尽可能小于靶粒子质量。因为在碰撞前后，质心的速度不变，入射粒子的动能不可能完全转化为激发能或电离能(或宏观领域内的热能)。只有相对于质心的动能，才可能转变为其它形式的能量，这部分能量称为**资用能**。显然，使质心动能在系统动能中所占份额尽可能小的方法是让两个粒子实现以相同的动量发生对碰。利用正、负粒子对的对头碰撞($v_2 = -v_1$)获得大质量的新粒子的加速器正是基于这一原理所设计的。

例 3-7-2 两个质量分别为m_1和m_2的重物用一长度为l，绕跨过一轻滑轮的轻绳连接在一起，轻滑轮中心的高度为H。开始时，m_2静止在地面上，m_1从高度$h > l/2$自由下落，求解物体以后的运动。



例3-7-3图

解：求解该问题的要点在于两个方面，一是两个重物是从轻绳绷紧后才发生相互作用，而轻绳绷紧瞬间所发生的作用过程相当于两个重物发生了一次完全非弹性碰撞。二是这两个重物在上述"碰撞"过程中也受到重力的作用，但因"碰撞"过程极短，该过程中重物间的相互作用远大于重力作用。因而可以利用动量守恒求解出系统碰撞后的速度，也就是系统在进入重力作用后，运动状态演化的初速度。

重物m_1从高度h自由下落到轻绳绷紧瞬间的速度为

$$u=\sqrt{2g[(l-H-2\pi R)-(H-h)]}=\sqrt{2g(l-2H-2\pi R+h)}$$

通过碰撞实现重新分配后的速度为

$$v=\frac{m_1u}{m_1+m_2}$$

以后的演化行为是两个物体都以v为初速度，$a=(m_1-m_2)g/(m_1+m_2)$为加速度的匀加速运动。◀

完全非弹性碰撞的一个重要特例是其逆过程：开始紧靠在一起的物体的两个部分突然分别以速度$\mathbf{v}_1$和$\mathbf{v}_2$分开，但保持总的动量为零：$m_1\mathbf{v}_1+m_2\mathbf{v}_2=0$。由于碰撞后体系的动能比碰撞前大，这种碰撞常被称为**超弹性**的。最后出现的动能来自两个物体初始时刻所具有的相互作用势能。该过程相当于炸药爆炸过程，但炸药中的势能是存储在化学物质中的，势能通过爆炸释放出来。

一般非弹性碰撞，碰撞恢复系数

对于宏观物体而言，实际碰撞总是介于弹性和完全非弹性碰撞之间的。最直观的图像是在质心参照系中，碰撞物体的初始动能既不会全部复原，也不会全部丧失。但碰撞中失去的动能并不总是转化为物体的形变势能，它还可能转化为热，声等其它形式的能量，因此，并不能用碰撞物体的形变程度来量度或预言一般的非弹性碰撞中有多少比例的动能会亏损掉。事实上，碰撞中的动能亏损与物体的材料，形状，碰撞的相对取向都有关系。如果能用一个唯像参量来概括碰撞中的形变恢复程度，我们就可以通过观测实际碰撞前后的速度变化情况来获得不同物体的碰撞恢复信息，并在形式上给出各种碰撞后物体的运动参量。

由于形变恢复参量只和碰撞物体的性质以及速度有关，而和观测的惯性参考系无关，该参量就只能是碰撞前、后两个物体的速度差的函数。另外对应于形变能够完全恢复的弹性碰撞和形变完全不能恢复的完全非弹性碰撞，恢复参数应分别取最大值和最小值。这样，最自然、也是最简单的选择就是将这个参量定义为碰撞物体的末速度差值和初速度差值的比值[3]：

$$e=\frac{v_2-v_1}{u_1-u_2} \tag{3-7-9}$$

称为**恢复系数**。容易验证，弹性碰撞和完全非弹性碰撞分别对应于$e=1$和$e=0$。一般的碰撞所对应的恢复系数则介于两者之间：$0<e<1$。这样在形式上可以将一般情况下碰撞的末速度v_1、v_2用初速度u_1、u_2及恢复系数e表示出来：

$$v_1=\frac{m_1-em_2}{m_1+m_2}u_1+\frac{(1+e)m_2}{m_1+m_2}u_2,\qquad v_2=\frac{(1+e)m_1}{m_1+m_2}u_1+\frac{m_2-em_1}{m_1+m_2}u_2 \tag{3-7-10}$$

而各种碰撞过程中所损失的机械能为

$$\Delta E=\frac{1}{2}(1-e^2)\frac{m_1m_2}{m_1+m_2}(u_1-u_2)^2 \tag{3-7-11}$$

现有的实验表明，恢复系数只与碰撞物体的性质有关，而与初速度无关（见思考

[3] 这种按照线性原则来定义参数的方法已经在 3-1-1 节定义物体的惯性时给出了说明。

题 3-7-5)。

3-7-2 二维以及三维空间中的碰撞

上面讨论的一维碰撞通常称为对心碰撞。即碰撞物体质心间的联线和碰撞物体间的相对速度是共线的。但实际上出现真正的对心碰撞是很偶然的，碰撞前两个物体的速度 $\mathbf{u}_1$、$\mathbf{u}_2$ 通常并不在它们的中心连线上。这样的碰撞从形态上说是一种斜碰。两个物体碰撞前的运动轨迹构成的平面称为入射平面，碰撞后两个物体的运动轨迹构成的平面称为散射平面。如果这两个平面重合，相应的碰撞称为二维碰撞，否则就属于三维碰撞。本节主要讨论二维弹性碰撞问题，三维碰撞以及各种非弹性碰撞问题可以用类似的方法处理，读者可以以此作为熟悉这种方法的练习(见习题 3-7-10~3-7-12)。

设两个质点的二维碰撞平面为直角坐标系中的 x-y 坐标面，系统在弹性碰撞前后所受到的受恒定理制约

$$m_1\mathbf{u}_1 + m_2\mathbf{u}_2 = m_1\mathbf{v}_1 + m_2\mathbf{v}_2 \tag{3-7-12}$$

$$\tfrac{1}{2}m_1u_1^2 + \tfrac{1}{2}m_2u_2^2 = \tfrac{1}{2}m_1v_1^2 + \tfrac{1}{2}m_2v_2^2 \tag{3-7-13}$$

具体表示为三个独立方程：x 方向和 y 方向的动量守恒和系统的动能守恒。如果仅仅知道碰撞前两个粒子的初速度，这三个方程显然不足以求解出碰撞后的四个速度分量，它还涉及对碰撞过程中相互作用的细节和更为确定的初始条件。求借碰撞的作用力特征可以仅用若干个参数如相互作用距离和作用强度等进行描述，称为**相互作用势模型**。通常考虑的宏观物体之间的碰撞都是相互作用时间很短的过程，这实际意味着碰撞物体相互作用的距离很小，具体表现为只是在彼此接触时才发生相互作用，如桌球游戏中的硬球之间的完全弹性碰撞。这类相互作用势可以用如图 3-7-4 所示的势能曲线来描写：两个球之间的相互作用距离为两个球的半径之和，而一旦进入相互作用区域，趋于发散的势能代表着近乎于无穷大的排斥作用，相应的作用时间近乎等于零。这种势能曲线描述的相互作用称为**硬球模型**，其中仅含一个参量—硬球半径 a，相应地二维硬球碰撞的初始状态还包含一个参量 b：入射粒子的直线轨道与通过靶心的平行线偏离间距(见图 3-7-5)，这一参量通常被形象地称为瞄准距离。粒子碰撞后的偏转角 θ (称为散射角)显然只是瞄准距离 b 和入射粒子的速率 v 的函数。

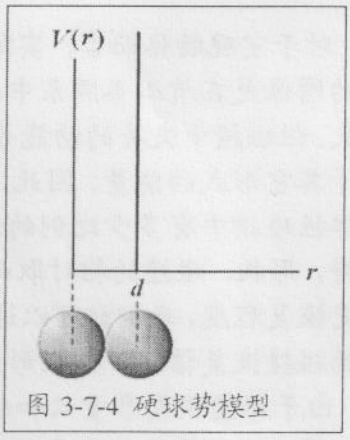


图 3-7-4 硬球势模型

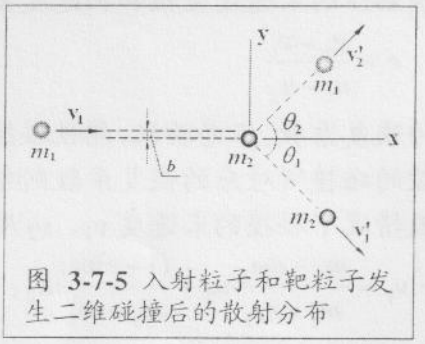


图 **3-7-5** 入射粒子和靶粒子发生二维碰撞后的散射分布

由于两个粒子间的一般(排斥或吸引)相互作用都可以表示为粒子间距离的连续函数，碰撞后粒子的散射方向是逐渐变化的(见图3-7-6(b))。瞄准距离和散射方向都是相对于无穷远处(相互作用衰减为零的区域)的渐近线而言的。散射角θ与瞄准距离b和入射粒子的速率v的具体关联特性取决于相互作用的形式，因而测量$\theta(b,v)$可以获得相互作用的信息。

图3-7-6所示为两种不同相互作用对应的散射结果。其中(a)为硬球相互作用，

可以用两个刚性圆盘接触瞬间的位形求解$\theta(b,v)$。由于碰撞时相对于盘心联线的入射角等于反射角，散射角θ与b满足关系：

$$b = a\sin\varphi = (r_1 + r_2)\cos\frac{\theta}{2} \quad (3\text{-}7\text{-}14)$$

其中r_1和r_2分别为两圆盘的半径。图3-7-7(*b*)为一个以不同瞄准距离b和速度v的小质量荷电粒子射向一个大质量荷电粒子受到静电排斥作用的散射情况，其偏转角θ与瞄准距离b满足：

$$\cot\frac{\theta}{2} = 4\pi\varepsilon_o\frac{mv^2}{2Qq}b \quad (3\text{-}7\text{-}15)$$

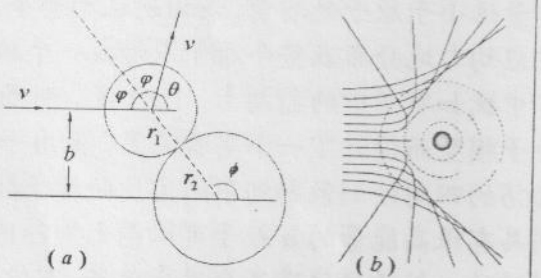
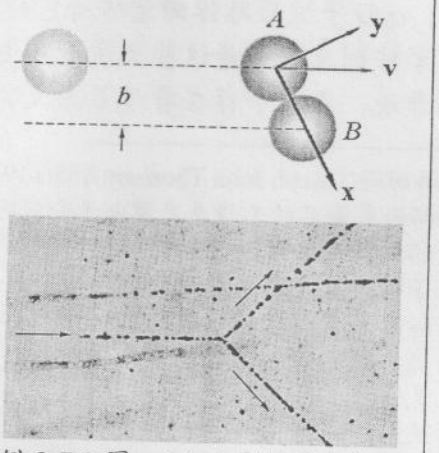


图3-7-6 (*a*)接触相互作用导致的弹性散射，(*b*)静电排斥作用散射。

其中Q是靶粒子电荷，q是入射粒子电荷，m是入射粒子质量。

例 3-7-3 证明质量相同的入射粒子和靶粒子发生弹性散射，两个粒子的出射速度方向一定互相垂直。

解：本例图(*a*)显示的是入射粒子 A 和靶粒子 B 接触瞬间的位型。以两个粒子中心的联线方向作为 x 轴，与之垂直的方向作为 y 轴设立坐标系。则粒子 A 的入射碰撞可以分解为 x 和 y 两个分量。其中两个粒子在 y 方向并没有发生碰撞，因而两个粒子将维持各自原有的 y 方向的速度不变，即粒子 A 保持原有的该方向的速度，粒子 B 则保持该方向的零速度。在x方向两个粒子的碰撞是完全弹性的对心碰撞，由于质量相同粒子的完全弹性碰撞的结果是交换速度，这意味着粒子 A 通过碰撞将原有 x 方向的动量完全转移到原来静止的靶粒子 B 上。因此，碰撞后粒子 A 只具有 y 方向的运动速度，而粒子 B 只具有 x 方向的运动速度，这样，A，B 两个粒子的出射速度方向是互相垂直的。图(*b*)显示的是在照相乳剂中一个质子和另一个静止质子发生碰撞在乳剂中留下的运动轨迹。这两条轨迹线几乎是相互垂直的。◀



例 3-7-3 图 (*a*) 入射粒子与相同质量的靶粒子发生弹性碰撞，两个粒子的出射方向相互垂直。(*b*) 一个质子和一个静止质子发生碰撞在乳剂中留下的运动轨迹。

二维碰撞中一类重要情形是一个质量较小的物体受到一个质量相对大得多的物体的散射。大质量物体的位置大致固定，小质量物体因受到相互作用而相对于原轨道发生偏离。这类碰撞称为小质量物体受到大质量物体的散射。以很高的速度进入太阳系的小星体离开太阳系后发生轨道偏离，以及一个电子从原子旁边经过后运动状态发生了变化都属于这类碰撞。下面通过一个探索原子内部结构的著名的散射实验来说明上述极限散射的重要应用。

最早的原子模型是所谓的葡萄干布丁模型。1895 年，英国物理学家汤姆逊[4]发现了从阴极物质发出的所谓阴极射线是一种质量很小的带电微粒，并将其称为电子。并确定了电子的质量为 9.1×10^{-31}kg。考虑到整个原子的电中性以及电子的质量远小于原子的质量，汤姆逊认为整个原子中携带绝大部分质量和正电荷的物质应均匀地分布在整个原子中形成一个松软的布丁，而有限数量的电子则镶嵌于其中犹如布丁中的葡萄干。1911 年，新西兰物理学家卢瑟福[5]为了检验汤姆逊的原子模型而设计了一个实验装置。他用一束 α 粒子流轰击一张极薄的金箔，并在金箔的四周检测散射到不同角度的粒子数(见图 3-7-6-a)。如果汤姆逊模型正确，则具有很高能量的 α 粒子可以毫无例外地轻易穿过金箔，由于金箔质量分布均匀，穿过的 α 粒子的轨道不会发生偏转。即使考虑到很少一部分 α 粒子有可能受到点状的电子的散射，但由于电子质量只有 α 粒子质量的万分之一数量级，它们对 α 粒子轨道的影响很小(见(3-7-8)式)，因此这部分 α 粒子只可能产生角度很小的散射(见图 3-7-6-b)。但实验结果大出意外：大多数 α 粒子确实以我们期望的方式径直穿过了金箔，但也有一小部分 α 粒子被偏转了很大的角度，有些甚至几乎被直接反弹了回来[6]。这种效应显然不是汤姆逊模型的结果。

为了理解这种不同寻常的大角散射事件，我们特别分析一下粒子被直接反弹回来的极端事例。前面在分析弹性碰撞的事例中已经说明，只有当靶粒子的质量远大于入射粒子质量时的对心碰撞，入射粒子才会被几乎以同样的能量被反射回去。α 粒子被原路弹回意味着它撞到了质量远大于它的物体，而只有极少数的这类事件则又意味着这些大质量正电粒子只能聚集在很小的空间区域(见图 3-7-6c)。或者说，原子中存在有质量很大的硬核，其直径约为原子尺度的 10^{-5}，即 10^{-15}

---

[4] 汤姆逊(Joseph John Thomson 1856-1940)，著名英国物理学家，卡文迪许实险室第三任主任。汤姆逊最著名的工作是发现电子和对同位素的实验研究：他通过使阴极射线在磁场和电场中的偏转测定了电子的荷质比，从而确认阴极射线为带负电荷的物质粒子—电子。基于电子作为原子的组分之一，汤姆逊提出了原子的布丁模型。并用原子内电子的数目和分布解释元素的化学性质。1912 年汤姆逊在对阳极射线的质量分析中发现了氖的同位素。因这些研究的杰出成就而获得 1906 年诺贝尔物理学奖。汤姆逊是一位杰出的物理学家，也是一位杰出的导师。他对学生要求非常严格，强调学生在开始研究前必须掌握所需要的实验技术，而且坚持学生研究使用的仪器必须自己制作。汤姆逊认为大学应培养学生学会思考、有独立工作能力，而不是用"现成机器"投影生产出"死的成品"。因此他要求学生不仅是实验的观察者，更是实验的创造者。在汤姆逊担任卡文迪许实险室主任期间。实验室不仅接二连三地涌现出新的发现。而且培养了众多的杰出人才。他的学生中有九位获得了诺贝尔奖。汤姆逊逝世后与牛顿，达尔文等安葬在一起。

[5] 卢瑟福(Ernest Rutherford 1871-1937)，出生于新西兰的英国物理学家。在放射性和原子结构等方面做出重大贡献，被公认为二十世纪最伟大的实验物理学家，近代原子核物理学之父。卢瑟福 1895 年获得奖学金在英国剑桥大学卡文迪许实验室由汤姆逊指导做博士研究。他原来的研究方向是电磁波，后来接受汤姆逊的建议研究放射性。1898 年经汤姆逊推荐担任加拿大麦吉尔大学物理系教授，在那里发现了放射性的半衰期，证明放射性是原子的自然衰变，并因此获得 1908 年的诺贝尔化学奖。但卢瑟福并不很高兴，他认为"真正的科学只有物理学科，其它只相当于集邮。"而他自己是物理学家。1909 年卢瑟福在英国曼彻斯特大学和学生马斯登通过 α 粒子撞击金箔的实验揭示了原子的有核模型。卢瑟福是一位伟大的导师，他的学生中有十一位先后获得了诺贝尔奖。1925 年卢瑟福获得英国政府颁发功绩勋章，1931 年被封为"纳尔逊的卢瑟福男爵"。卢瑟福在 1937 年因病去世，葬于牛顿和法拉第的墓旁。

[6] 卢瑟福以他特有的风格表达了对散射实验结果的意外："这几乎是我一生中最为意外的事件。就如同你用 15 英寸的火炮轰击一张薄纸，炮弹却被直接弹了回来。"

米的量级，这一区域称为原子核。而为了说明实际观察到的原子尺度，卢瑟福设想原子中的电子围绕着这个荷正电的原子核无休止的旋转(见图 3-7-6d)。这样的原子模型看起来与太阳系非常类似，因而统称被称为行星模型[7]。

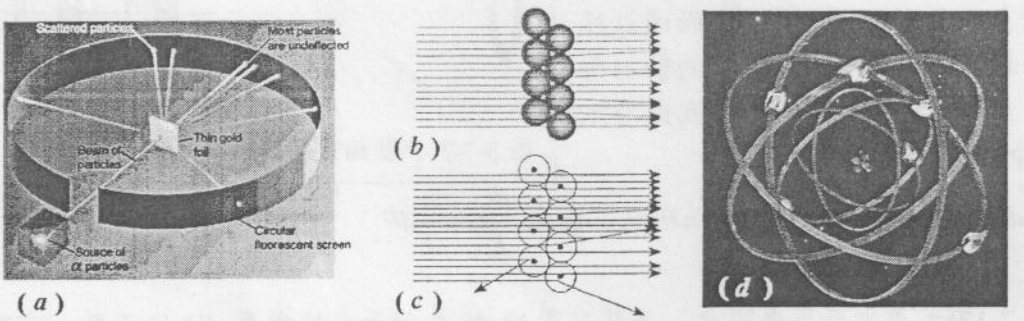


图 3-7-6 (*a*) α-粒子散射实验示意图。(*b*) α-粒子束入射到汤姆逊原子产生的预期散射结果。(c) α-粒子束入射到卢瑟福原子产生的预期散射结果。(*d*)卢瑟福原子行星模型

碰撞截面

对于微观粒子的碰撞而言，粒子的尺度很小，而粒子间的相互作用力程一般都远大于粒子本身的尺度，所以微观粒子之间的碰撞大多也是二维或三维的碰撞。由于微观粒子之间的相互作用随粒子间距快速衰减，一个入射粒子射向一个靶离子而能够发生有效作用的概率也非常小，因此真正的散射实验通常是用一束粒子入射向一个含有大量同类靶粒子的箔材料。入射粒子在接近靶粒子时开始出现相互作用，在较远处掠过靶粒子则因为没有发生相互作用而不会受到散射，则相当于每个靶粒子有一个有效面积，射到该截面上的入射力粒子受到有效散射，而射到该面积以外的入射粒子与靶粒子的相互作用可以忽略，这样的有效面积称为散射截面。记为σ。一个入射粒子与靶粒子发生相互作用而受到散射的概率就等于单粒子的散射截面总和与整个靶面积的比值。

截面是一个用于表达粒子间发生相互作用可能性的术语，显然散射截面与粒子间的相互作用特征，入射粒子的能量等参数有关。

在原子核物理学中，截面这一概念可以很方便地表达特定事件发生的可能性。在统计上，薄膜上原子中心可近似为均匀分布在一个平面上的点集。参与撞击原子的中心与其他原子以 r 为距离通过的概率是确定的。如果平面上 A 区域有 n 个原子中心，这个概率为$(n\pi r^2)/\mathrm{A}$，这仅仅是所有原子中心以 r 为半径的圆的总面积与整个平面的比值，也是入射粒子被被射原子平面阻挡的可能性。如果计算通过的原子，结果则表达为原子的等效截止截面。

例 **3-7-4** 卢瑟福的 α 粒子实验的散射截面。

解：卢瑟福散射的大致特征可用图 3-7-7-b 描述：以瞄准距离 b，和速度 v 的 α

[7] 物理学中一般更愿意称卢瑟福模型为核模型，即模型的意义在于揭示了原子中的大部分质量集中在原子中的一个极小的区域。而电子的绕核作轨道运动的行为几乎在一开始就被认为是不合理的。相关的讨论见第 9-2 节。

粒子受到靶原子的库仑力排斥而改变运动方向[8]，其偏转角 θ 与瞄准距离 b 满足关系(3-7-15)式。则瞄准距离为 b 和 $b+db$ 之间的入射 α 粒子的散射角在 θ 和 $\theta+d\theta$ 之间，而通过环形面积 $d\sigma=2\pi bdb$ 的 α 粒子的散射角在 θ 和 $\theta+d\theta$ 之间的一个立体角元 $d\Omega$ 中(见图 3-7-7)。

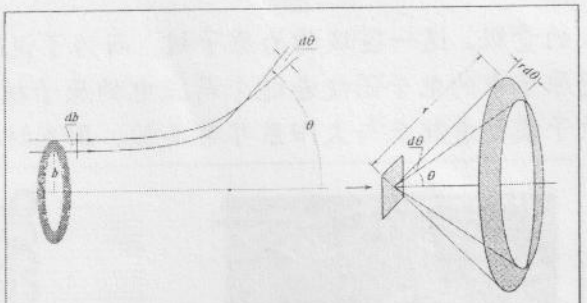


图 3-7-7 卢瑟福散射截面与角度的关系

$$d\Omega=\frac{2\pi r\sin\theta r d\theta}{r^2}=2\pi\sin\theta d\theta=4\pi\sin\frac{\theta}{2}\cos\frac{\theta}{2}d\theta$$

对(3-7-13)式平方后再求微分，并将其中的 $d\theta$ 表示为立体角元 $d\Omega$ 的函数，可得 α 粒子散射到上述立体角 $d\Omega$ 内每个原子的散射截面(又称微分截面)：

$$d\sigma=\left(\frac{1}{4\pi\varepsilon_o}\right)^2\left(\frac{2Ze^2}{mv^{2^2}}\right)^2\frac{d\Omega}{\sin^4\theta/2}$$

为了将实验观察数值与上述卢瑟福散射公式联系起来，设想用一束 α 粒子射向一面积为 A，厚度为 l，原子数密度为 N 的薄膜。薄膜中的原子总数为 $N'=NAl$。对很薄的薄膜，其中的原子核对入射 α 粒子不互相遮蔽，散射到 θ 和 $\theta+d\theta$ 之间的总散射截面是[9]

$$d\Sigma=N'd\sigma=NAld\sigma$$

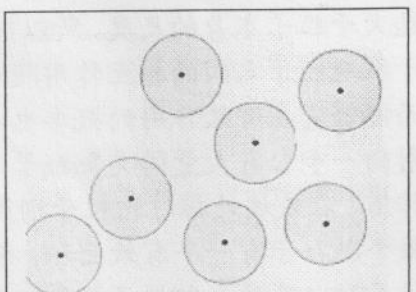


图 3-7-8 截面的示意图

如果射到薄膜的全部面积 A 上的 α 粒子数为 n，其中散射到上述立体角元 $d\Omega$ 中的粒子数为 dn，这些粒子一定落在 $d\Sigma$ 上，因此有

$$\frac{dn}{n}=\frac{d\Sigma}{A}=Ntd\sigma \text{ 则 } d\sigma=\frac{dn}{nNt}$$

即 $d\sigma$ 与 dn/n 成正比，$d\sigma$ 代表 α 粒子被散射到 θ 和 $\theta+d\theta$ 之间的几率。由此得到

$$\frac{d\sigma}{d\Omega}\sin^4\frac{\theta}{2}=\left(\frac{1}{4\pi\varepsilon_o}\right)^2Nnt\left(\frac{2Ze^2}{mv^{2^2}}\right)^2 \qquad (3\text{-}7\text{-}15)$$

对给定的 α 粒子源和散射物而言，等式右边是常数，因此左边的数值应不随 θ 而变。实际测量时，接收 α 粒子的荧光屏在不同方向张了一个小立体角 $d\Omega'$，实际测得的粒子数是在 $d\Omega'$中的 dn'。显然，θ 相同时，$dn'/d\Omega'=dn/d\Omega$。所以(3-7-15)式中克用 $dn'/d\Omega'$代替 $dn/d\Omega$。表 3-7-1 中列出了是金箔对 α 粒子的散射结果。其中在 45°到 150°的散射角范围，$dn'\sin^4(\theta/2)$差异很小，但对小角散射的情况与理论分析不符。

---

[8] 在原子核的质量比 α-粒子的质量大得多的情况下，可以认为前者不会被推动。

[9] 考虑到所用金箔的厚度仅 5×10^{-7} 米，约为一千多个原子的厚度(金原子的直径约 3×10^{-10} 米)，而原子核与原子的半径之比约 $10^{-14}/10^{-10}=10^{-4}$，原子核的几何截面约为原子的 10^{-8}，这样原子核前后遮蔽的机会很小。

上述理论与实验的差异来源于计算中假定通过金属箔的 α 粒子只经过一次散射。这一假设对于不同角度的散射近似程度不同。α 粒子通过金箔实际上将从许多原子核附近经过，但多次非常接近原子核的机会很小。即 α 粒子虽然实际经过了多次散射，但只有瞄准距离 b 很小时才会出现大角散射。更多的散射是小角散射。因小角散射的方向是随机的，多次散射产生的方向改变比一次大角散射小得多。因而观察到的较大的 θ 角是一次大角散射和多次小角散射的累积效应，其中小角散射的影响很小。但对于实际观察到较小的 θ 角，也是多次小角散射累计的。由于对同样的小角散射不能忽略其中任何一个，因而一次散射近似不再适用。这说明表 3-7-1 中的 45°以下的实验结果与理论不符。

表 3-7-1 α 粒子不同角度散射

θ (°)	dn'	$dn' \sin^4(\theta/2)$
150	33.1	28.8
135	43.0	31.2
120	51.9	29.0
105	69.5	27.5
75	211	29.1
60	477	29.8
45	1435	30.8
37.5	3300	35.3
30	7800	35.0
22.5	27300	39.6
15	132000	38.4

读者可能还会想到按原子核外边的电子对 α 粒子的运动可能产生的影响。但因 α 粒子的质量比电子大 7300 倍，α 粒子的运动受电子的影响是微不足道的。

上述关于截面的概念可以延伸到到任何有关入射粒子与靶粒子间的相互作用。例如 α 粒子轰击靶粒子铍会产生中子的可能性，靶区域呈现介于 0 至 100% 反应几率的情况—因为目标中的物质不完全相同，或者因为不均匀的场使之减弱。可以表示为这种原子核反应的反应截面。再例如散射截面用于描述光(或其它形式辐射)被粒子散射的可能性，它不同于几何上的“截面”，与入射粒子的波长和靶粒子的电容率、形状、大小有关。稀疏介质的散射总数决定于散射截面和粒子个数的乘积。考虑到吸收、散射和发光，总散射截面可以用面积表示为

$$\sigma = \sigma_A + \sigma_S + \sigma_L$$

传统的碰撞问题属于两体问题，散射角是两个质点间相对位矢 $\mathbf{r}$ 在散射前后所偏转的角度 θ_r (见图)，可直接观察，虽然两体问题可以简化为单体问题：一个物体以折合质量相对于另一个物体运动，但这样处理得出的散射角是被散射的质点在散射前后所偏转的角度 θ_c(见图)，要由等值单体问题计算。只有当散射粒子在碰撞前后始终静止不动时，θ_r 和 θ_c 才相等。

我们可以在不同的惯性系中分析和求解二维或三维碰撞问题，例如在实验工作中，通常采用射靶坐标系，称为实验室坐标系，而在理论分析时，为简化计算常采用质心坐标系。

微观粒子的完全非弹性碰撞很少出现两个粒子撞击后粘在一起构成复合粒子的情形，但它的逆过程却很常见。如各种自由衰变、放射现象都属于这一类。

3-7-3 持续碰撞—变质量系统的动力学问题

一个在运动过程中质量不断发生变化的物理系统称为变质量系统。这类问题的典型事例是物体持续受到微量物体的完全非弹性碰撞导致的质量连续增长，或

其逆过程—系统连续地抛弃其质量。例如，雨滴在下落过程中因不断积聚水汽而质量不断增大，火箭在飞行过程中因喷射气体而使自身的质量不断减少。如果系统中所有质点的运动情况相同，该体系的运动也可以用一个质点来描写。但这样的质点运动显然不能直接应用牛顿运动方程求解，也不能直接利用多粒子系统的质心动量定理，牛顿运动方程或质心动量定理描述只是给定过程中的一个确定的质点或一个确定组成的体系所遵循的规律。因为只有当体系的组成确定，内力和外力才有确定的意义。

为了使用牛顿定律处理这类问题，我们可以将体系所经历的过程分解成一系列的过程微元，设某个过程微元的起始时刻 t，原来的体系(称为主体)和即将进入主体的物体(称为附体)是分离的，附体在元过程的最后时刻 $t+\Delta t$ 并入主体。在该过程中，主体和附体的组成及质量都是确定的，体系的动量变化服从质点系统的动量定理。由此可以得到主体的运动方程。

设在时刻 t，主体质量为 m，速度为 $\mathbf{v}$，即将进入主体的附体的质量为 Δm，速度为 $\mathbf{u}$，则系统在该时刻的动量为 $m\mathbf{v}+\Delta m\mathbf{u}$，在 $t+\Delta t$ 时刻，Δm 并入 m，并与主体以相同速度$(\mathbf{v}+\Delta\mathbf{v})$一起运动，则体系的动量变为$(m+\Delta m)(\mathbf{v}+\Delta\mathbf{v})$。设体系(主体和附体)受到的外力是 $\mathbf{F}$，则元过程的质点系动量定理给出

$$(m+\Delta m)(\mathbf{v}+\Delta\mathbf{v})-(m\mathbf{v}+\Delta m\mathbf{u})=\mathbf{F}\Delta t$$

用 Δt 除上式两边，并令 $\Delta t\to 0$，可得

$$m\frac{d\mathbf{v}}{dt}=(\mathbf{u}-\mathbf{v})\frac{dm}{dt}+\mathbf{F} \tag{3-7-16}$$

这就是变质量质点(即主体)的运动方程。其中外力 $\mathbf{F}$ 包括主体和附体所受到的外力总和：$\mathbf{F}=\mathbf{F}_m+\mathbf{F}_{\Delta m}$。与质点运动方程比较可以看出，主体运动方程可以写为

$$m\frac{d\mathbf{v}}{dt}=\mathbf{F}_m+\mathbf{F}_{\Delta m\to m} \tag{3-7-17}$$

其中

$$\mathbf{F}_{\Delta m\to m}=\mathbf{F}_{\Delta m}+(\mathbf{u}-\mathbf{v})\frac{dm}{dt} \tag{3-7-18}$$

是附体在并入主体过程中对主体的作用力。

例 3-7-4 以变质量物体运动的观点重新求解例 3-3-5。

解：因为要求的是绳索对地面的作用力，取地面上的绳索作为主体。在绳索上端下落到 x 高度的瞬时，系统主体的速度 v 和加速度都为零，而附体的速度为 $u=\sqrt{2g(l-x)}$，主体质量的增加速率为 $dm/dt=\rho u=\rho\sqrt{2g(l-x)}$，作用在体系上的外力为 $\rho(l-x)g-N$，于是由变质量系统的运动方程(3-7-15)式得到

$$-\rho[2g(l-x)]=\rho(l-x)g-N$$

解得 $N=3\rho(l-x)g$。与例 3-3-5 得到的结果相同。◀

例 3-7-5 质量为 m 的雨点在自由下落过程中不断有水汽凝结到它上面，若单位时间内凝结到它上面的水汽质量为 μ。求不考虑空气阻力的情况下，雨点在 t 秒后所下落的距离。

解：由于水气的运动速度 $\mathbf{u}=0$，故变质量雨点的运动方程为

$$\frac{d}{dt}[(m+\mu t)\mathbf{v}] = (m+\mu t)\mathbf{g}$$

通过对时间积分，并利用雨点的初始质量和速度值可得

$$v = \frac{ds}{dt} = \frac{mt+\mu t^2/2}{m+\mu t}g = \frac{1}{2}gt + \frac{mg}{2\mu} - \frac{m^2 g/2\mu}{m+\mu t}$$

再对时间积分，并利用雨点的初始位置数值可得雨点在 t 秒后所下落的距离为

$$s = \frac{1}{4}gt^2 + \frac{mg}{2\mu}t - \frac{m^2 g}{2\mu^2}\ln\left(1+\frac{\mu}{m}t\right) \quad \blacktriangleleft$$

例 3-7-6 火箭是通过持续向外喷出燃烧过的废气来增加火箭本身运动的速度，设火箭沿着运动物体轨道切线的负方向喷射燃气，且燃气喷射速度的量值不变，求火箭在燃料耗尽时所能获得的最终速率与火箭的运行方式的关系。

解： 火箭主体的运动方程可以写为

$$m\frac{d\mathbf{v}}{dt} = \mathbf{F} - \frac{dm}{dt}v_r\hat{\mathbf{v}}$$

其中 v_r 是燃料相对于火箭的喷射速度。当火箭运动不受外力作用时，其运动方程简化为

$$\frac{dv}{v_r} = -\frac{dm}{m}$$

对上式积分，并利用火箭的起始质量数值 m_o 可得：$v = v_o + v_r\log(m_o/m)$。设 m_G 为初始时刻全部燃料的质量，则在 $v_o = 0$ 的条件下，火箭在燃料耗尽时所能获得的末速度为

$$v_f = v_r\ln(m_o/m_R) \tag{3-7-19}$$

即 v 和 v_r 成正比。当空火箭及燃料的总质量 m_o 与空火箭的质量 m_R 之比按几何级数增加时，燃气耗尽时火箭所能获得的末速度只能按算术级数增加。因此，增加燃气喷射速度比增加燃料的数量更为有效。

如果 $v_r = 2$ km/s，使火箭的 v_f 超过第一宇宙速度所需要的燃料质量要比空火箭的质量大 100 倍左右。如果计及地球引力和空气阻力，这个比值还要大。因此一般采用多级火箭的方案：当某一级火箭里面的燃料烧完后就将其丢弃，以提高火箭的飞行速度。

现在求火箭离开起始点的距离，由式(3-7-19)得 $ds = v_o dt - v_r\ln(m/m_o)dt$，积分得

$$s = v_o t - v_r\int_0^t \ln(m/m_o)dt$$

求火箭离开起始点的距离需要知道 $m(t)$的具体函数形式，两种最常应用的质量变化规律为：

(i) 直线律：$m = m_o(1-\alpha t)$，其中 α 是常数，则每秒钟放出的质量为

$$\frac{dm}{dt} = \frac{d}{dt}[m_o(1-\alpha t)] = -\alpha m_o = \text{const.}$$

等于常数。而反作用

$$F_r = v_r\frac{dm}{dt} = -\alpha m_o v_r$$

因 v_r 等于常数，F_r 也是一个常数。则火箭离开起始点的距离

$$s = v_o t - \frac{v_r}{\alpha}[(1-\alpha t)\ln(1-\alpha t) + \alpha t]$$

(ii) 指数律：$m = m_o e^{-\alpha t}$。由反作用力所引起的附加加速度

$$a_r = \frac{F_r}{m} = \frac{v_r}{m}\frac{dm}{dt} = -\alpha v_r$$

因 v_r 等于常数，指数律用于反作用加速度为常数的情况。火箭离开起始点的距离为

$$s = v_o t - \frac{1}{2}\alpha v_r t^2$$

火箭喷射物质时所经过的路程 s 称为喷射行程，所经历的时间 t_s 为喷射时间。火箭因重力损失的速度正比于 gt_s。因此火箭的加速度不受限制时应使其尽快烧完所携带的燃料以减小 t_s，即减小速度的损失。从以上讨论还可看出：火箭速度和排气性能有关，排气速度 v_r 越大，喷出的质量越多，火箭得到的速度 v 就越大。另外，质量比 m'/m_s 虽不如 v_r 重要，但也影响火箭的速度。因此，最佳的设计方案是除了减小了 t_s 外，还要使 v_r 和 m'/m_s 两者都尽可能的大。这对材料的要求就特别挑剔。除了要使用能产生高温的燃料(如单原子氢或其它固态燃料)外，还要求火箭的外壳用耐高温的难熔轻金属制造。钛是目前一种比较理想的材料。

以上的讨论中火箭被看作是一个质量可变的质点。而实际上火箭具有一定的大小。因此，问题要比所讨论的复杂，例如，将火箭看作刚体，则除了其质心的运动外，还要研究它绕质心的转动问题。

对于持续碰撞，由于碰撞时间长，碰撞期间所受到的累积冲量有限，系统因碰撞受到的作用力并不大。因而系统可能受到的其它外力作用就不能忽略。以地面附近的升空火箭运动为例，除了因受到燃料的推力而获得速度，还因受到地面的引力而减速。因此表达式(3-7-19)需要修改，火箭的速度除了获得因喷射燃料获得的速度 $u_{ex}\ln(m_o/m)$，还受到地面的引力而获得的减速累积 $-gt$，因而得到

$$v = u_{ex}\ln\frac{m_0}{m} - gt \qquad (3\text{-}7\text{-}17)$$

则在速度表达式(3-7-12)中应增加 $-gt$ 项，在距离表达式(3-7-13)中应增加 $-\frac{1}{2}gt^2$ 项。如果考虑空气阻力，问题还要复杂一些。

思考题

3-7-1 两个物体在“碰撞”过程中，系统可能同时受到其它外力(例如重力，各种阻尼力)的作用一般总是远小于“碰撞”的作用强度，这一点应该怎样从实验测量中得到证实或判断。

3-7-2 在非弹性碰撞中损失的机械能，是否与观察的参考系有关？

3-7-3 一个质量为 m 的球以一定的速度与质量为 M 的静止靶球发生非弹性碰撞。以下三种情形中哪一种机械能损失最多？(1) $m \geq M$; (2) $m \leq M$; (3) $m \approx M$。

3-7-4 用一个鸡蛋 A 去撞击另一个相同的鸡蛋 B(两个鸡蛋的撞击部位相同)，为什么结果是鸡蛋 A 碎？能否用一个简单的模型说明你的分析？

3-7-5 说明为什么需要关注碰撞的恢复系数是否与碰撞物体的初速度有关。现有的实验表明，恢复系数只与碰撞物体的性质有关，而与初速度无关，其中是否可能存在什么问题？

3-7-6 定义二维及三维碰撞的恢复系数，并给出碰撞后粒子的末速度的表达式。

习题

3-7-1 质量为 m 的粒子以速度 v_o 与一质量为 M 的静止质点发生完全弹性碰撞。在 $m/M = 0$~10 的范围中画出粒子 m 的末速 v 与比值/M 函数关系图。

3-7-2 质量为 m 的子弹以速率 v_o 水平地射入静止在光滑水平桌面上的质量为 M 的木块内，并与它一起运动。求子弹相对于木块静止后，木块的速率和动量以及在此过程中子弹施于木块的冲量。

3-7-3 在光滑水平面上，两个质量都是 m 的小球由一长为 l 的轻杆相连，另一质量也为 m 的小球 A 以与杆成 $\theta = 45^\circ$ 角的方向以速率 v_o 与杆上某一小球发生弹性碰撞，然后沿垂直于原速度方向运动。求碰撞后小球 A 的运动速率 v_f，以及轻杆系统绕其质心转动的角速度 ω。(郑永令 5-10)

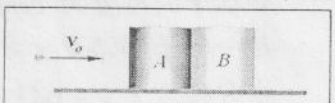
3-7-4 一根不可伸长的绳子跨过一定滑轮，两端各拴质量为 m 和 M 的物体($M > m$)。M 静止在地面上，绳子起初松弛。当 m 自由下落一个距离 h 后绳子开始被拉紧。求绳子刚被拉紧时两物体的速度和此后 M 上升的最大高度 h。

3-7-5 如本题图，质量为 m 的物体放在光滑的水平面上，两边分别与劲度系数为 k_1 和 k_2 的两个弹簧相连，若在右边弹簧末端施以拉力 f，求(1) 拉力非常缓慢地拉了一段距离 L 作的功，(2) 拉到距离 L 后突然不动拉力作的功。

3-7-6 质量为 $M = 980$ g 的木块静止在光滑的水平面上。一质量为 $m = 20$ g 的子弹以速率 $v_o = 800$m/s 水平入射到木块内，并与木块一起运动。求木块对子弹作用力的功；子弹对木块做的功以及耗散掉的机械能。

3-7-7 质量为 m_1 及 m_2 的物体以劲度系数为 k 的弹簧相连静止在光滑的水平面上，质量为 m 的子弹以速率 v_o 水平入射到 m_1 内，求弹簧的最大压缩量。

3-7-8 子弹水平地穿过两个前后并排在光滑水平桌面上的静止木块。木块的质量分别为 m_A 和 m_B，设子弹透过两木块的时间间隔为 t_A 和 t_B。设子弹在木块中所受阻受阻力为恒力 f，求子弹穿过后两木块的运动速度。



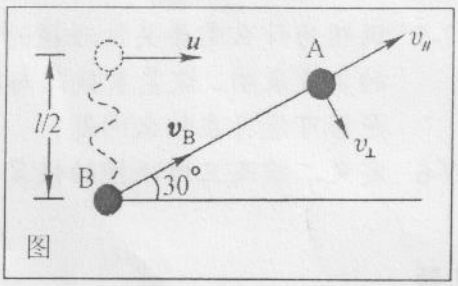
题 3-7-4

3-7-9 相同质量和半径的两个球悬挂于同一高度，静止时两球恰能接触且悬线平行。已知两球碰撞的恢复系数为 e，若球 A 自高度 h_1 释放，求该球碰撞弹回后能达到的高度。

3-7-10 在一铅直面内有一光滑的轨道，轨道左边是光滑弧线，右边是足够长的水平直线。质量为 m_A 的质点 A 自 h 高度自由下滑与静止在水平轨道上的质点 B(质量为 m_B)发生完全弹性碰撞，碰撞后 A 返回到弧线的某一高度再度滑下与质点 B 碰撞。求 A、B 至少发生两次碰撞的条件。

3-7-11 设炮车以仰角 α 发射炮弹，炮身和炮弹的质量分别为 m_o 和 m，炮弹在炮口处相对炮身的速率为 $\mathbf{v}'$，求炮身的反冲速度 $\mathbf{v}_o$。设地面摩擦力可以忽略

3-7-12 质量均为 m 的小球 A、B 用长为 l 的无弹性轻绳相连，放在光滑水平台面上，两球相距½l，绳呈弯曲状。小球 A 因受冲击而获得与两球连线相垂直的速度 **u**，求当绳伸直的瞬时两球的速度。



图

解：当绳伸直的瞬时，小球 B 受到绳的拉力的冲量而沿绳方向运动，小球 A 受到反方向的绳的冲量而使速度发生变化。B 的速度 $\mathbf{v}_B$ 的方向沿绳；将 A 的速度 $\mathbf{v}_A$ 分解为与绳垂直的分量 $v_{\perp}$ 和沿绳的分量 $v_{\parallel}$ 两部分(图 3-2-4)。由于绳不可伸长且无弹性，两球沿绳方向的速度相等：$v_B = v_{\parallel}$。

显然，两球组成的体系动量守恒。将动量分解为与绳垂直的分量和与绳平行的分量两部分，而绳伸直时与 **u** 的方向成 30º 角，注意到 $v_B = v_{\parallel}$，故有 $mu\sin30º = mv_{\perp}$，$mu\cos30º = mv_{\parallel}$，由以上两式解得：$v_{\perp} = ½u$，$v_{\parallel} = u\cos30º$

3-7-13 1932 年詹姆斯·查德威克通过快中子与静止氢核、氮核的弹性碰撞测量中子质量。结果发现，氢核的最大反冲速度为 3.3×10^7m/s，^{14}N 核的最大反冲速度为 4.7×10^6m/s，误差为 10%。求中子质量和所用中子的初速度。氢核，^{14}N 核的质量分别为 1 和 14 个原子质量单位。

3-7-14 一质量为 m_o 速率 v_o 运动的粒子与一质量为 $2m_o$ 静止的粒子碰撞。碰撞后入射粒子偏转了 45º，末速度为½v_o，求质量为 $2m_o$ 的粒子偏转后的速率和方向。动能守恒吗？

3-7-15 两船在静水中都依惯性以速率 6.0m/s 相向匀速而行，它们相遇时，将甲船上的货物搬到乙船上。以后乙船沿原方向继续前进，但速率变为 4.0m/s.，甲船维持原速度不变。设甲船空载时的质量为 500kg，货物的质量为 60kg，求乙船质量。在搬运货物的前后，两船和货物的总动能有没有变化？

3-7-16 三只质量均为 M 的小船鱼贯而行，速率均为 v。若中间船上同时以水平速率 u(相对于船)把两质量均为 m 的物体分别抛到前后两只船上。求此后三只船的速率。

3-7-17 原来静止在光滑水平面上，质量为 m 的物体受到一连串粒子的轰击。每个粒子的质量为 $\Delta m(\ll m)$，速率为 v_o，沿正 x 方向。碰撞是完全弹性的。每一粒子都沿负 x 方向弹回。证明物体经第 n 个粒子碰撞后，得到的速率非常接近于 $v = v_o(1-e^{-an})$，其中 $a = 2\Delta m/m$。讨论该结果对于 $an \le 1$ 和 $an \to \infty$情形的有效性。

3-7-18 站在小船上的人质量 70kg，船的质量为 130kg。若人向船头走 4.0m。人相对于河岸走了多远？

3-7-19 水平地面上停放着一辆小车，车上站着 10 个质量相同的人，每人都以相同的方式、消耗同样的体力从车后沿水平方向跳出。设车的质量远大于 10 个人的质量，而且所有人所消耗的体力全部转化为车与人的动能，在整个过程中可略去一切阻力。为了使小车得到最大的动能，车上的人应一个一个地往后跳，还是 10 个人一起跳？

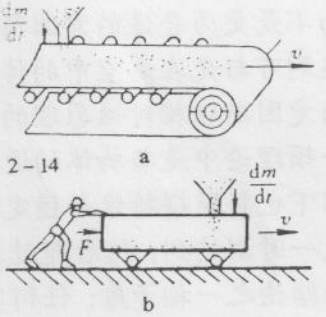
3-7-20 如果 α 粒子撞击的是汤姆逊的布丁原子，计算 α 粒子因受到点状电子的散射而可能产生小角度散射的角度范围。

3-7-21 计算两个粒子发生二维弹性碰撞后散射方向之间的夹角。Figure out the angle between the two trajectories of the scattered particles

3-7-22 一个静止的原子核在放射性衰变时，在两个互相垂直的方向分别放出一个动量为 9.22×10^{-16}g·cm/s 的电子和一个动量为 5.33×10^{-16}g·cm/s 的中微子。求衰变后原子核的动量的大小和方向。

3-7-23 用变质量系统的运动方程重新求解 3-3-36 和 3-3-37。

3-7-24 一水平传送带将沙子从一处运送到另一处，沙子经一垂直的静止漏斗落到传送带上，传送带以恒定速率 v 运动着(见本题图 a)。忽略机件各部位的摩擦。若沙子落到传送带上的速率是 dm'/dt，试问：要保持传送带以恒定速率 v 运动，水平总推力 F 多大？若整个装置是：漏斗中的沙子落进以匀速 v 在平直光滑轨道上运动的货车里(见本题图 b)，以上问题的答案改变吗？



2－14

3-7-25 起始质量为 M_o 的火箭以恒定率 dM/dt = ＿＿＿＿ 排出燃烧过的燃料，排料相对于火箭的速率为 v_o。计算火箭从发射台竖直向上起动时的初始加速度；如果 v_o = 2000m/s，则对于一个质量为 100t 的这种火箭，要给以等于 0.5g 的向上初始加速度，每秒钟必须排出多少公斤燃料？

3-7-26 一个三级火箭的质量如下表所示，不考虑重力，火箭的初速为 0。若燃料相对于火箭喷出速率为 u = 2500 m/s，每级燃料外壳在燃料用完时将脱离火箭主体。设外壳脱离主体时相对于主体的速度为 0，只有当下一级火箭发动后，才将上一级的外壳甩在后边。求第三级火箭的最终速率；若把 48t 燃料放在 12t 的外壳里组成一级火箭，问火箭最终速率是多少。

级别	发射总质量	燃料质量	燃料外壳质量
一级	60t	40t	10t
二级	10t	20/3t	7/3t
三级	1 t	2/3t	

3-7-27 宇宙飞船以恒速 v 在空间飞行，飞行过程中遇到一股微尘粒子流，后者以 dm/dt 的速率沉积在飞船上。尘粒在落到飞船之前的速度为 u，方向与 v 相反，在时刻 t 飞船的总质量为 $M(t)$，试问：要保持飞船匀速飞行，需要多大的力？

3-8 相对论动力学

对物体运动及其演化的定量描述依赖于我们对时空特性的认识，不同的时空观念会影响对物体动力学行为的描述和理解。牛顿的力学理论自建立后获得了空前的成功，以至在17世纪中叶到19世纪末的两百多年中一直被认为是物理学理论中的经典之作和描述自然的真理。但牛顿力学定律及其各种推断和结论都是建立在绝对时空观念基础上的，即牛顿的力学理论所满足的只是伽利略相对性原理，而不是更为普适的爱因斯坦相对性原理。读者因而很容易想到，在描述运动速度达到可与光在真空中的传播速度相比较时的物体行为时，牛顿力学理论会表现出与爱因斯坦相对性原理的明显冲突。一个最明显、直接的例子是，物体的惯性在牛顿理论中是与物体的运动状态无关的，因而该物体即使在极微弱的外力 **F** 作用下也将获得持续而稳定的加速度：$\mathbf{a} = \mathbf{F}/m$，其速度在 **F** 的持续作用下总能在某一时刻达到，然后超过真空中光的传播速度。这显然与爱因斯坦的相对性原理的结论之一相矛盾：任何实验手段都不可能将一个物体通过加速达到光速。另一个直接与牛顿定律冲突的例子则不那样明显：设想两个物体 A 和 B 之间存在与两者距离有关的某种相互作用。若 A 以恒定速度 u 相对于 B 运动，则 A 和 B 之间的相互作用力就是时间的函数。牛顿第三定律对这一问题的表述："在任一时刻物体 A 作用到物体 B 与物体 B 作用到物体 A 上的作用力大小相等，方向相反"是基于对两个粒子的受力"同时"的绝对性。但在相对论框架中，若两个"作用"与"反作用"的相关事件在物体 A 的参照系中是同时的，在物体 B 的参照系中就一定不是同时的，即第三定律在 A，B 两个惯性系中不再是等价的。

描述和理解高速运动物体的动力学行为所面临的无论是修改牛顿的力学理论还是建立全新的力学理论都让我们几乎回到牛顿建立其理论的原点：定义什么样的函数 $f(m,v)$ 使其能够在更大的速度范围中对物体的运动给出普适的描述。本节将在爱因斯坦的相对性原理框架中讨论如何通过修改相关的物体动力学参量定义以得到新的运动方程，以及这些物理量和方程所具有的一些新的特征和属性。

3-8-1 惯性质量和动量

重新定义运动参量需要一个思考问题的起点和一个合适的实验设计。3-1-1节的反冲实验显然不可能使粒子加速到接近真空中的光速，因而也无法通过实际测量两个粒子的反冲速度以检验惯性对运动的无关性和惯性的可加性在高速区域是否仍然成立。但我们可以采用伽利略通过斜面实验揭示惯性定理的论证方法，即基于所能实现条件下的实验结果，利用逻辑推理导出真实的结论。

惯性定律和相对性原理是进行逻辑推理所基于的前提，即仅仅两个物体之间的相互作用并不能改变两个物体的运动总量，而且在任何惯性系中观察到的两个物体运动总量在任何时刻都必须守恒。由于上述基本原理只是一种定性的表述，更接近于一种哲学的表述，为了得到量化分析的实验基础，我们可以利用低速情形的实验结果以及被实验所证实的低速和高速之间的联系桥梁：不同惯性系中的物理参量之间所满足的洛仑兹变换，而且在低速极限下各种物理参量在不同惯性

系之间的洛仑兹变换方程必须还原到伽利略变换方程。

首先我们考虑两个相同质点的完全非弹性碰撞实验，希望系统的动量守恒和相对性原理能够帮助我们求解出动量$\mathbf{p}(m, \mathbf{v})$的相对论形式，或至少能提供一些有用的帮助。

我们先具体考虑系统的质心参考系S_C中，粒子A和B的初速度分别为$\mathbf{u}$和$-\mathbf{u}$(见图3-8-1-a)，碰撞后复合粒子处于静止状态。按牛顿动量的定义，系统在反冲前后的总动量分别为：$\mathbf{P}_i = \mathbf{p}_A + \mathbf{p}_B = m\mathbf{u} - m\mathbf{u} = 0$，$\mathbf{P}_f = 0$，即碰撞前后，系统在$S$系中的牛顿动量是守恒量。

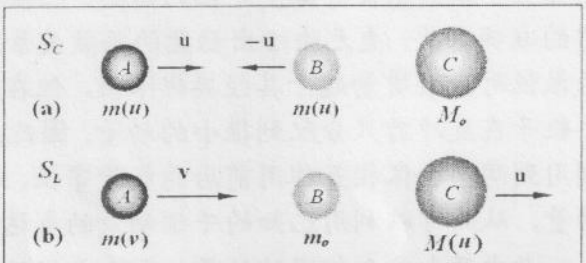


图3-8-1 (*a*) 在质心系S_C中，A和B碰撞前速度分别为$-\mathbf{u}$和$\mathbf{u}$，碰撞后复合粒子速度为零。(*b*) 在实验室系S_L中，A和B碰撞前速度分别为$-\mathbf{u}$和$\mathbf{u}$，碰撞后复合粒子速度的速度为$\mathbf{v}$。

接着考虑质点A和B在碰撞前分别扮演入射粒子和静止靶粒子的所谓实验室参考系S_L，其中粒子A的入射速度为$\mathbf{v}$，碰撞后，复合粒子在S_L中的运动速度为$\mathbf{v}$(图3-8-1-b)。

按伽利略相对性原理，由于系S_L相对于系S_C的运动速度为$\mathbf{u}$，则粒子A的入射运动速度为$\mathbf{v}' = 2\mathbf{u}$。系统在碰撞前后总的牛顿动量分别为

$$\mathbf{P}'_i = \mathbf{p}'_{Ai} + \mathbf{p}'_{Bi} = m\mathbf{v}' = 2m\mathbf{u}, \quad \mathbf{P}'_f = M\mathbf{u} = 2m\mathbf{u}$$

即反冲过程前后所观测到的牛顿动量守恒对于伽利略变换保持不变。

但在爱因斯坦相对性原理的框架中，物体B在惯性系S'中的运动速度为：

$$\mathbf{v}' = \frac{\mathbf{u}+\mathbf{u}}{1+u^2/c^2} = \frac{2\mathbf{u}}{1+\beta_u^2} \tag{3-8-1}$$

则系统在碰撞前后总的牛顿动量分别为

$$\mathbf{P}'_i = \mathbf{p}'_{Ai} + \mathbf{p}'_{Bi} = m\mathbf{u} + m\mathbf{u} = 2m\mathbf{u}, \quad \mathbf{P}'_f = \mathbf{p}_{Bf} = m\mathbf{v}' = \frac{2m\mathbf{u}}{1+\beta_u^2}$$

即反冲过程前后系统的牛顿动量守恒不符合爱因斯坦相对性原理。

考虑到动量守恒必须符合爱因斯坦相对性原理，而在物体运动速度远小于真空中光速的情形，相对论动量的表达式必须回复到牛顿动量的形式，因此两种动量具有相同的量纲$[mv]$。而物体的惯性必须是无量纲参量$\beta_v^2 = (v/c)^2$的函数：$m(v) = m(\beta_v^2)$。因为这才可以使惯性保持原来的量纲，而且与运动方向无关。另外，在低速极限下回到牛顿的惯性定义时，惯性随速度的变化率也非常小，与我们所熟悉的在相当大的速度范围中惯性的变化都小到无法观测的实验结果相符。

关于物体的惯性还需要注意到的一点是，由两个物体构成的复合系统的惯性不能想当然地认为是这两个物体的惯性之和(见3-1-2节中关于物体惯性的讨论)，因此S'系中观测到的复合物体A-B的惯性质量和物体A和B的惯性质量之和不能简单地等同起来。为此我们将反冲前复合物体A-B的系统的惯性质量记为$M(u)$，则系统相应的动量可表示为：

$$\mathbf{P}'_i = M(u)\mathbf{u} \tag{3-8-2}$$

这样S'系中观测到的系统反冲前后的相对论动量守恒$\mathbf{P}'_i = \mathbf{P}'_f$表示为

$$M(u)\mathbf{u} = m(v)\mathbf{v} = -\frac{2m(v)\mathbf{u}}{1+\beta_u^2} \tag{3-8-3}$$

上式中存在两个待求解的变量：$M(u)$和$m(v)$，由于不能将$M(u)$简单地写成$2m(u)$，无法仅仅通过方程(3-8-3)解得物体的惯性质量与运动速度的关系。从实验设计角度而言，反冲实验不能提供足够的方程以提供额外的信息，例如上述过程除了动量守恒，还应包含有能量守恒的信息，但在不清楚动量表达式的情形，无法给出力的准确定义，更无法给出动能的函数关系。另一个可以利用的信息是粒子在速度很低时惯性质量趋于其经典极限值，但在反冲实验中，无法在S'系中使任何一个粒子在反冲后只分配到很小的动量。因此应该设计一个恰当的实验，使我们能利用到两个物体相互作用前后的能量守恒，或只让靶粒子在反冲后分配到很小的动量，从而可以利用已知的牛顿动量的表达式减少未知量的数目。

考虑两个完全相同的粒子 a 和 b 之间的二维完全弹性碰撞，而碰撞前后系统的动量和能量在所有惯性系中观察都是守恒的。设惯性系 S' 是粒子 a 和 b 的质心参考系，碰撞前 a 和 b 相对于 S' 的速度分别为 $\mathbf{u}'_a = \mathbf{u}'$，$\mathbf{u}'_b = -\mathbf{u}'$，$\mathbf{u}'$ 沿 x 轴方向。碰撞后 a 和 b 的散射速度分别为 $\mathbf{v}'_a$ 和 $\mathbf{v}'_b$ (见图 3-8-2)。

相应的动量守恒方程可以表示为

$$m(u')\mathbf{u}' - m(u')\mathbf{u}' = m(v'_a)\mathbf{v}'_a - m(v'_b)\mathbf{v}'_b \tag{3-8-4}$$

对于两个全同粒子的完全弹性碰撞而言，对称性要求两个粒子散射后的速度大小相等，方向相反。由此可以得到：$\mathbf{v}'_a = -\mathbf{v}'_b$，$v'_a = -v'_b = -u'$。即碰撞前后粒子的速度量值都相等，两个粒子的速度方向都相反。S' 系中唯一没有确定的是粒子碰撞后的速度与 x 轴的夹角 θ'，这一角度依赖于散射的碰撞参数。

图 3-8-2 在惯性系 S' 和 S 中两个相同粒子在发生碰撞前后的速度矢量，惯性系 S' 中的表示速度矢量的实线和虚线分别表示通过洛仑兹变换和伽利略变换得到的 S 中两个相同粒子在发生碰撞前后的速度矢量。

我们再考虑粒子 b 静止不动，扮演靶粒子的参考系 S，S 系相对于 S' 系的运动速度为 $-\mathbf{u}'$。由洛仑兹速度变换公式可知，S 系中入射粒子 a 的运动速度为

$$\mathbf{u}_a = \frac{\mathbf{u}' + \mathbf{u}'}{1 + u'^2/c^2} = \frac{2c\boldsymbol{\beta}_{u'}}{1+\beta_{u'}^2} \equiv \mathbf{u} \tag{3-8-5}$$

类似地可以求得两个粒子在碰撞后的散射速度

$$(\mathbf{v}_a)_x = \frac{c\beta_{u'}(1+\cos\theta')}{1+\beta_{u'}^2\cos\theta'} \qquad (\mathbf{v}_a)_y = \frac{c\beta_{u'}\sin\theta'}{\gamma_{u'}(1+\beta_{u'}^2\cos\theta')}$$
$$(\mathbf{v}_b)_x = \frac{c\beta_{u'}(1-\cos\theta')}{1-\beta_{u'}^2\cos\theta'} \qquad (\mathbf{v}_b)_y = \frac{c\beta_{u'}\sin\theta'}{\gamma_{u'}(1-\beta_{u'}^2\cos\theta')} \tag{3-8-6}$$

在 S 系中，散射前后的动量守恒方程 $m(u)\mathbf{u} = m(v_a)\mathbf{v}_a + m(v_b)\mathbf{v}_b$ 的 x, y 分量形式分别为：

$$m(u)\frac{2c\beta_{u'}}{1+\beta_{u'}^2} = m(v_a)\frac{c\beta_{u'}(1+\cos\theta')}{1+\beta_{u'}^2\cos\theta'} - m(v_b)\frac{c\beta_{u'}(1-\cos\theta')}{1-\beta_{u'}^2\cos\theta'}$$
$$0 = m(v_a)\frac{c\beta_{u'}\sin\theta'}{\gamma_{u'}(1+\beta_{u'}^2\cos\theta')} - m(v_b)\frac{c\beta_{u'}\sin\theta'}{\gamma_{u'}(1-\beta_{u'}^2\cos\theta')} \tag{3-8-7}$$

其中动量 y 分量的守恒方程可以表示为

$$m(v_a)=\frac{1+\beta_{u'}^2\cos\theta'}{1-\beta_{u'}^2\cos\theta'}m(v_b) \tag{3-8-8}$$

如果散射过程中这两个粒子只发生很弱的相互作用，S 系中的靶粒子就只能获得很小的运动量，它的惯性质量趋向于它的经典极限值 $m(v_b\simeq 0)\simeq m(0)$，而入射粒子的散射速度相对于其入射速度则只会有很小的变化 $\mathbf{v}_a\simeq\mathbf{u}$。这种情形具体对应于入射粒子掠射过靶粒子的弹性散射，在 S' 系中对应于粒子的散射角 θ' 很小。因而 S 系中上述碰撞实验前后的系统动量 y 分量的守恒方程表示为

$$m(u)=\lim_{\theta'\to 0}\frac{1+\beta_{u'}^2\cos\theta'}{1-\beta_{u'}^2\cos\theta'}m(v_b)=\frac{1+\beta_{u'}^2}{1-\beta_{u'}^2}m(0)$$

利用(3-8-5)式，可以将上述两个惯性质量之间的比例系数表示为

$$\frac{1+\beta_{u'}^2}{1-\beta_{u'}^2}=\frac{1}{\sqrt{1-(u/c)^2}}=\gamma_u$$

记 $m_o=m(0)$，称为物体的**静止质量**。由此得到，在惯性系中观察到一个以速度 $\mathbf{u}$ 运动的物体，其惯性质量表示为

$$m(u)=\gamma_u m_o \tag{3-8-9}$$

为区别于物体的静止质量，称 $m(u)$ 为物体的**运动质量**。

确定了以速度 $\mathbf{u}$ 运动的物体的惯性质量，也就得到了物体的动量表达式

$$\mathbf{p}=m(u)\mathbf{u}=\gamma_u m_o\mathbf{u} \tag{3-8-10}$$

利用这一特殊的实验情形中的能量守恒方程的极限形式

$$E(u_a)+E(u_b=0)=E(v_a\simeq u_a)+E(v_b\simeq 0)$$

我们还可以确定一个物体以速度 $\mathbf{u}$ 运动时的能量函数形式 $E(u)$。由于这会涉及比较多的数学过程，我们将其留给读者作为练习。我们将在 3-8-3 节利用动能定理推导出以速度 $\mathbf{u}$ 运动的粒子的能量函数形式。

读者不应对物体的惯性与速度有关感到惊异，因为惯性是量度改变物体运动状态难易程度的参量，而相对论惯性质量反映的也是这一测量属性，而无论惯性是否与速度有关，都只是描述实验结果的经验参量，不是任何动力学理论的推论。

事实上，早在 1897 年(爱因斯坦的相对论理论发表于 1905 年)，德国物理学家考夫曼(W. Kaufman)在通过放电管中阴极射线偏转测量 e/m 时就已经观察到 e/m 值随电子速度的改变而改变，而且他的测量精度与现代值只差 1%，但他当时没有勇气发表这些结果。1908 年，德国物理学家布赫雷尔(Alfred Heinrich Bucherer)通过测量电子在真空管中以不同速度 v 在均匀磁场 B 中作圆周运动的半径 r，获得的电荷-质量比证实了物体的惯性随着速度增长的预言。电子在与其运动方向垂直的磁场 B 中作圆周运动时的轨道半径与电子的惯性质量 m、电荷 e 及运动速率 v 满足

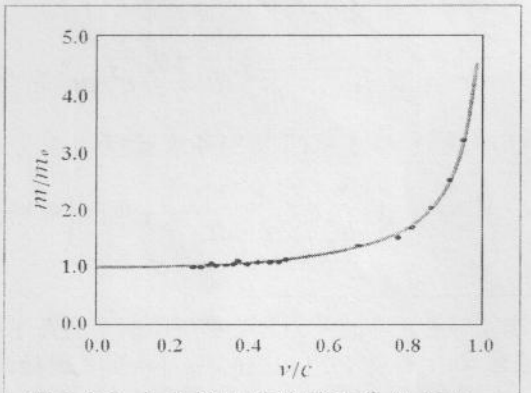


图 3-8-2 电子惯性质量随速度的变化。

$$r = \frac{mv}{eB} \tag{3-8-7}$$

若电子的惯性质量随速度的增加而增长，电子运动的轨道半径会以相同的比例增长。图 3-8-2 所示为电子的运动质量与静止质量在不同速率情形的比值。其中的圆点为实验值，曲线为理论计算结果[10]。由图中的曲线的变化可以看出，即使物体的运动速度达到真空中光速的 10%的时候，惯性质量的相对变化仅为千分之五。

物体惯性质量依赖于运动速度的另一个实验证据是观测高能粒子撞击相同靶粒子后两个粒子的运动轨迹间的夹角。在经典力学中，两个相同质量的粒子发生二维弹性碰撞后其运动轨迹是相互垂直的，而大质量粒子撞击轻质量粒子所形成的散射夹角则小于 90°。由于在相对论中粒子处于运动状态比处于静止状态时质量大，因此高能粒子撞击相同的靶粒子所形成的散射夹角应小于 90°。图 3-8-3 所示为高能电子进入感光乳液中撞击乳液物质原子中的电子所形成的散射径迹照片。其中散射电子和被散射电子的径迹夹角明显地小于 90°。

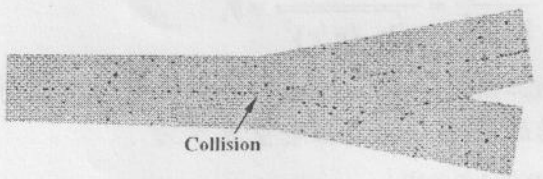


图 3-8-3 高能电子进入感光乳液中撞击乳液原子中的电子所形成的散射径迹。

应该说明上述两个实验都不能从很高的精度上验证物体惯性质量和速度之间的定量关系，但它们很好地显示出如何从利用最简单的力学效应出发，考虑验证一个理论的实验设计思想。关于高精度验证物体惯性质量和速度的实验是测量氢原子谱线的精细结构，有兴趣的读者可以阅读相关的文献。

3-8-2 相对论运动方程

我们仍然用一个物体的动量随时间的变化率来量度其所受到的作用力，则物体的相对论形式的运动方程表示为：

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = \frac{d}{dt}(m_v\mathbf{v}) = m_o\frac{d}{dt}(\gamma_v\mathbf{v}) = m_o\left(\gamma_v\frac{d\mathbf{v}}{dt} + \frac{d\gamma_v}{dt}\mathbf{v}\right) \tag{3-8-8}$$

其中洛仑兹因子γ_v对时间的导数为

$$\frac{d\gamma_v}{dt} = \frac{v}{c^2\sqrt{1-v^2/c^2}^3}\frac{dv}{dt} = \gamma_v^3\frac{v}{c^2}\frac{dv}{dt} \tag{3-8-9}$$

则相对论运动方程(3-8-8)可表示为

$$\mathbf{F} = m_o\gamma_v\left(\frac{d\mathbf{v}}{dt} + \gamma_v^2\frac{v}{c^2}\frac{dv}{dt}\mathbf{v}\right) = m_o\gamma_v\left(\mathbf{a} + \gamma_v^2\beta_v^2\frac{dv}{dt}\hat{\mathbf{v}}\right)$$

---

[10] 读者需要注意的是，实验中所能测量的是半径 r，速度 v 和外磁场 B，实验中所能确定的是在给定 B 的条件下，比值 e/m 和 v/r 之间的函数关系。因此可以就式(3-8-7)中提出一个问题：为什么将测量到的 e/m 随速度的变化归结为电子的惯性而不是电子的电量。事实上有一些实验可以说明物质的电量是与运动速度无关的守恒量。例如，原子中作不同轨道运动的电子具有不同的速度，但物质从未因电子的运动速度不同而失去电中性，下一章将会讨论相关的问题。

将其中是加速度$\mathbf{a} = d\mathbf{v}/dt$分解为平行和垂直于速度$\mathbf{v}$的切向分量$a_\tau$和横向分量$a_n$，并注意到加速度的切向分量只改变物体的速率$a_\tau = dv/dt$，而横向分量$a_n$只改变物体速度的方向，则上式可写为

$$\mathbf{F} = m_o\gamma_v(\mathbf{a}_n + \mathbf{a}_\tau + \gamma_v^2\beta_v^2\,|\,\mathbf{a}_\tau\,|\,\hat{\mathbf{v}}) = m_o\gamma_v[\mathbf{a}_n + (1+\gamma_v^2\beta_v^2)\mathbf{a}_\tau]$$

或

$$\mathbf{F} = m_o\gamma_v^3\mathbf{a}_\tau + m_o\gamma_v\mathbf{a}_n \tag{3-8-10}$$

由于加速度的两个分量中所含的因子不同，作用在物体上的力与物体的加速度的方向通常也不相同[11]。这也是相对论运动方程与牛顿运动方程之间最明显的差别之一。由于$\gamma_v > 1$，改变物体速率比改变物体速度的方向困难得多。随着物体的速度，物体的惯性迅速增长，加速效应也越来越不明显，当物体的速度接近光速时，物体的惯性趋向与无穷大，有限的作用力不再产生加速效应。这表明对于静止质量$m_o > 0$的物体不可能通过力的作用被加速到光速。

当物体的速度远低于光速时，相对论运动方程随着 $\gamma_v \to 1$ 而退化回到牛顿运动方程:

$$\mathbf{F} = m_o\mathbf{a}_\tau + m_o\mathbf{a}_n = m_o\mathbf{a}$$

3-8-3 相对论能量

按照动能定理，外力对物体作用在空间累积(做功)的观察效应表现为物体动能的增加

$$dK = \mathbf{F}\cdot d\mathbf{r} = \frac{d\mathbf{p}}{dt}\cdot d\mathbf{r} = d\mathbf{p}\cdot\frac{d\mathbf{r}}{dt} = d\mathbf{p}\cdot\mathbf{v} = \frac{1}{m}\mathbf{p}\cdot d\mathbf{p} = \frac{1}{2m}dp^2 \tag{3-8-11}$$

由相对论质量的定义可得: $m\sqrt{1-v^2/c^2} = m_o$，两边平方给出

$$p^2 = (mc)^2 - (m_oc)^2 \tag{3-8-12}$$

对上式两边取微分: $dp^2 = 2mc^2dm$，代入(3-8-11)可得: $dK = c^2dm$。通过做功将物体从静止加速到速度v的过程中，物体所增加的能量为

$$K(m,v) = c^2\int_{m_o}^{m_v} dm = m_vc^2 - m_oc^2 \tag{3-8-13}$$

$K(m,v)$称为运动物体的相对论动能。它在形式上与经典力学的动能表达式相差很大，但在低速情况下可以回到原来的经典表示:

$$K = (\gamma_v - 1)m_oc^2 = \tfrac{1}{2}mv^2(1 + \tfrac{3}{4}\beta^2 + \tfrac{3}{4}\tfrac{5}{6}\beta^4 + \ldots) \approx (1 + v^2/2c^2 - 1)m_oc^2 = \tfrac{1}{2}m_ov^2$$

(3-8-13)式中的m_oc^2项与物体的运动无关，称为物体的静止能量，记为

$$E_o = m_oc^2 \tag{3-8-14}$$

相应地将物体的动能与静止能量之和

$$E = K + m_oc^2 = mc^2 = \gamma_vm_oc^2 \tag{3-8-15}$$

称为物体的相对论能量。

从(3-8-12)可以得到物体相对论能量的另一种表述形式

$$E = mc^2 = \sqrt{p^2c^2 + m_o^2c^4} \tag{3-8-16}$$

当物体的速度远小于光速时，该表达式退化为

[11] 如果一定要在形式上将"力"表示为惯性质量与加速度的乘积，则惯性质量应该表示为一个二阶张量。

$$E = m_o c^2 + p^2/2m_o$$

当物体的速度与光速相当时，其静止质量所对应的能量在物体的总能量中只占很小的一部分，物体的能量表达式(3-8-16)简化为

$$E = pc \tag{3-8-17}$$

称为相对论粒子的能量，也是静止质量为零的粒子的能量表达式。

静止的物体具有能量，而相同的物体具有不同运动能量时也就具有不同的惯性质量，这一事实表明物体的能量和惯性质量之间是可以相互转换的。这一概念将一个物体的能量和质量统一起来。因而一个物理系统内部的动能及相互作用能都可以通过物体的惯性形式表现出来，反过来，物体的静止质量也可以通过合适的方式转化为其它形式的能量释放出来。

但对于通常的运动物体而言，这种质能转换的数量和整个物体的静止质量对应的能量相比是非常小的。例如两个质量都为0.1千克的物体以同样的300 m/s的速度发生完全非弹性对撞，所增加的静止质量为$\Delta m_o = \Delta K/c^2 \approx 2(\frac{1}{2}mv^2)/c^2 = m_o\beta_v^2 \approx 1\times10^{-12}$ kg，相对质量变化仅为$\Delta m_o/2m_o \approx 5\times10^{-12}$。再例如将一根弹性系数为$k$ = 10 $N\cdot m^{-1}$的弹簧压缩1厘米，弹簧所具有的弹性势能所对应于弹簧处于松弛状态时有较大的静止质量：$\Delta m_o = (\frac{1}{2}kx^2)/c^2 \approx 5.6\times10^{-19}$ kg。

一般的宏观弹性形变表现为分子间的位型和距离变化。由于分子间的力起源于中性分子极化所产生的静电力，弱于两个电荷之间的库仑作用，因此形变势能所形成的静止质量远小于净电荷间的库仑作用势能导致的静止质量变化。例如一个自由电子($m_e \approx 9.10954\times10^{-31}$ kg)和一个自由质子($m_p = 1.67265\times10^{-27}$ kg)结合形成一个氢原子($m_H = 1.6736\times10^{-27}$ kg)所形成的静止质量亏损为$m = m_p + m_e - m_H \approx 2.4\times10^{-35}$ kg，相对质量变化为$\Delta m/m_H \approx 5\times10^{-8}$比分子作用的强度大3~4个量级。静止质量亏损对应的能量损失以发射光的形式达到平衡，这部分能量称为氢原子的结合能[12]。

比静电作用更强的是核子之间的相互作用[13]，在各种原子核中，具有中等质量数的原子核的核子平均结合能较大，而质量数较大或较小的原子核中核子的平均结合能都较小。所以重原子核发生裂变成为中等质量的原子核时会放出能量，而轻原子核聚合而成较重原子核时也会放出能量。例如一个铀235原子核受到一个中子轻微撞击后所发生裂变反应会以裂变后的各较小原子核所携带的动能的形式释放出约200 MeV的能量(见图3-8-4)

$$^{235}_{92}\mathrm{U} + n \rightarrow {}^{236}_{92}\mathrm{U} \rightarrow \mathrm{X} + \mathrm{Y} + 200\ \mathrm{MeV}$$

单位质量的^{235}U裂变放出的能量是200/235(≈0.85) MeV (计算$\Delta m/m$估计核作用力的强度)。一克^{235}U全部裂变释放出的能量约8×10^{10}J，相当与2.5吨煤的燃烧热，或20吨TNT炸药释放的能量。

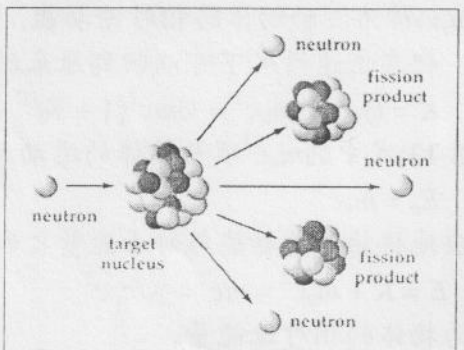


图3-8-4 重原子核的裂变反应。

而典型的聚变反应：

[12] 实验上一般通过测量发射的光的频率确定。

[13] 见4-7节

$^2H + {}^2H \rightarrow {}^3He + n + 3.25\ MeV$

$^2H + {}^2H \rightarrow {}^3H + {}^1H + 4.00\ MeV$

$^3H + {}^2H \rightarrow {}^4He + n + 17.6\ MeV$

$^3He + {}^2H \rightarrow {}^4He + {}^1H + 18.3\ MeV$

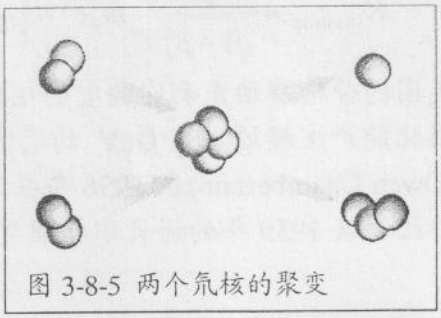


图 3-8-5 两个氘核的聚变

这些反应的原料直接或间接都是氘核^2H。四种反应中共消耗六个^2H，放出43.2MeV的能量，平均每个^2H放出7.2 MeV。则单位质量的^2H可放出3.6 MeV，是单位质量的235U裂变所放出的能量的4倍左右。每克氘经聚变可以放出10^5千瓦小时的能量[14]。

例 3-8-1 高能物理中的一项重要工作是寻找自然界可能存在的一些大质量粒子。这些粒子通常寿命很短，即使它们出现在宇宙射线中，也难于穿越大气层到达地面而被观测到。因而发现这些粒子的途径就是在实验室直接"制造"它们，即通过一些粒子之间的强烈碰撞。这种实验需要满足一定的条件[15]。例如要产生一个反质子[反质子$\bar{p}$和质子具有相同的静止质量(m_pc^2 = 938MeV)，但所携带的电荷与质子的电荷相反]。只能通过质子和质子的碰撞，而守恒定律要求至少要附带产生三个质子才能产生一个反质子，相应的反应过程为

$$p + p \rightarrow p + p + p + \bar{p}$$

为产生这样的反应而建造加速器时的一个基本的设计参数是估算至少需要将入射质子加速到多高的能量来撞击靶质子。入射质子的这个最小动能称为阈值动能，对应于反应后所有的生成粒子就好象是一个复合粒子，都以相同的速度运动。计算该动能阈值。

解: 计算阈值动能的最简单方法是先在质心系计算，再转换到实验室参考系中去。因为在质心参考系中，反应后的总能量只是以产物粒子的静止能量形式出现:

$$E_{Cf} = 4m_pc^2$$

而碰撞前的系统总能量是两个具有同样速度的质子的能量的和:

$$E_{Ci} = E_{C1} + E_{C2} = 2E_{C1}$$

反映前后的能量守恒 $E_{Ci} = E_{Cf}$给出:

$$E_{C1} = E_{C2} = 2m_pc^2$$

设质心参考系中两个反应粒子的速率为 u(这其实也是两个参考系之间的相对运动速率)，则上述能量表达式可写为

$$E_{C1} = \gamma_u m_pc^2 = 2m_pc^2$$

由此可得每一个反应粒子的速率为 $u = (\sqrt{3}/2)c$。通过洛仑兹速度变换公式可以计算在实验室参考系中入射质子的速度:

$$v = \frac{u+u}{1+\beta_u^2} = \frac{2u}{1+\beta_u^2} = \frac{4}{7}\sqrt{3}\,c \approx 0.9897\,c$$

由此可以得到该反应所需要的阈值动能

---

[14] 氢和氘的原子数之比是 1 比 0.00015，按重量计算，氘结成的水约为海水的六千分之一。地球表面海水约 10^{18}吨，所以海水蕴藏的氘所能供给的聚变能约 10^{25} 千瓦小时，按目前世界能量消耗率可以用几百亿年

[15] 见下一章 除了动量，能量守恒，电荷守恒，还有重子数守恒等

$$K_{\text{threshold}} = \frac{m_p c^2}{\sqrt{1-\beta_v^2}} - m_p c^2 = \left(\frac{1}{\sqrt{1-0.9897^2}} - 1\right) m_p c^2 \approx 6 m_p c^2 = 5.6\,(\text{GeV})$$

美国的劳伦斯伯克利实验室的电子感应加速器部分就是为了这一目的所设计的，因此能产生超过 5.6 GeV 的高能质子束。西格雷(Emilio Gino Segrè)和张伯伦(Owen Chamberlain)在 1956 年在该加速器中的质子碰撞实验中发现了反质子，并为此荣获 1959 年的诺贝尔物理奖。◀

3-8-4 动量、能量和力的洛仑兹变换

一个相对于惯性系S以速度$\mathbf{v}$运动的物体，其动量和能量在S系中分别表述为

$$\mathbf{p} = \gamma_v m_o \mathbf{v},\quad E = \gamma_v m_o c^2$$

设惯性系 S'以速度 $\mathbf{u}$ 相对于惯性系 S 运动，则该物体相对于 S'系的速度 $\mathbf{v}'$为

$$v'_{\parallel} = \frac{v_{\parallel} - u}{1 - u v_{\parallel}/c^2},\quad v'_{\perp} = \frac{v_{\perp}}{\gamma_v (1 - u v_{\parallel}/c^2)}$$

其中 $v'_{\parallel}$, $v'_{\perp}$和 $v_{\parallel}$, $v_{\perp}$分别为速度 $\mathbf{v}'$和 $\mathbf{v}$ 平行或垂直于 $\mathbf{u}$ 的分量。而物体在 S'系中的平行或垂直于 $\mathbf{u}$ 的动量分量和能量表达式分别为

$$p'_{\parallel} = \frac{m_o v'_{\parallel}}{\sqrt{1 - v'^2/c^2}} = \frac{m_o (v_{\parallel} - u)/(1 - u v_{\parallel}/c^2)}{\sqrt{1 - [(v_{\parallel} - u)/(1 - u v_{\parallel}/c^2)]^2/c^2}} = \frac{m_o (v_{\parallel} - u)}{\sqrt{(1 - u v_{\parallel}/c^2)^2 - (v_{\parallel} - u)^2/c^2}}$$

$$= \frac{m_o v_{\parallel}}{\sqrt{1 - u^2/c^2}\sqrt{1 - v^2/c^2}} - \frac{m_o}{\sqrt{1 - v^2/c^2}} \frac{u}{\sqrt{1 - u^2/c^2}} = \gamma_v \gamma_u (m_o v_{\parallel} - m_o u)$$

$$p'_{\perp} = \frac{m_o v'_{\perp}}{\sqrt{1 - v'^2/c^2}} = \frac{m_o v_{\perp}/(1 - u v_{\parallel}/c^2)}{\sqrt{1 - [(v_{\parallel} - u)/(1 - u v_{\parallel}/c^2)]^2/c^2}} = \frac{m_o v_{\perp}}{\sqrt{1 - v^2/c^2}}$$

$$E' = m(v')c^2 = m_o c^2 \frac{1 - v_{\parallel} u}{\sqrt{1 - u^2/c^2}}$$

它们可以用 S 系中的动量 p 和能量 E 表示为

$$p'_{\parallel} = \gamma_u p_{\parallel} - \gamma_u \frac{u}{c^2} E,\quad p'_{\perp} = p_{\perp},\quad E' = -\gamma_u u p_{\parallel} + \gamma_u E \tag{3-8-18}$$

其中 $p_{\parallel}$和 $p_{\perp}$分别为该物体相对于惯性系 S 的平行或垂直于 $\mathbf{u}$ 的动量分量。式(3-8-17)表明惯性系 S'中的动量和能量都表示为惯性系 S 中的动量和能量的线性组合，其中组合系数与物体本身的速度无关。因此，当动量和能量在 S 系中守恒，也必然在 S'系中守恒。

需要说明的是，在物体碰撞过程中，S'系中的动量守恒依赖于 S 系中的动量和能量的守恒这一事实并不意味着 S'系中的碰撞必须是完全弹性的。能量守恒对非弹性碰撞仍然成立，但部分动能转化为系统的静止质量所对应的能量。

例 **3-8-2** 设一动能为 K_K = 325 MeV 的 K 介子在其飞行途中衰变为两个 π 介子：$K \rightarrow \pi + \pi$。如果这两个 π 介子的运动轨迹与原来的 K 介子运动轨迹共线，计算各 π 介子的运动动能。其中 K 介子和 π 介子的静止质量能量分别为 $m_K c^2$ = 498 MeV 和 $m_\pi c^2$ = 140 MeV。

解：系统在衰变过程前后的动量和能量分别守恒：

$$\sqrt{(p_1c)^2+(m_\pi c^2)^2}+\sqrt{(p_2c)^2+(m_\pi c^2)^2}=K_K+m_Kc^2=823\,\text{MeV}$$

$$p_1c+p_2c=p_Kc=\sqrt{E_K^2-(m_Kc^2)^2}=655\,\text{MeV}$$

其中 p_1 和 p_2 分别为两个 π 介子的动量，E_K 和 p_K 分别为 K 介子的能量和动量。从上面两个方程可以解得 $p_1c=668$ MeV 和 $p_2c=-13$ MeV。两个 π 介子的动能分别为

$$K_1=\sqrt{(p_1c)^2+(m_\pi c^2)^2}-m_\pi c^2=543\,\text{MeV}$$

$$K_2=\sqrt{(p_2c)^2+(m_\pi c^2)^2}-m_\pi c^2=0.6\,\text{MeV}$$

也可以先在质心坐标系中求解，再将结果用洛仑兹变换变到实验坐标系中(习题3-8-11)。◀

利用动量的变换方程(3-8-17)可推导作用力在上述惯性系 S 和 S' 中的变换关系式：

$$F'_{\parallel}=\frac{dp'_{\parallel}}{dt'}=\gamma_u\left(\frac{dp_{\parallel}}{dt'}-\frac{u}{c^2}\frac{dE}{dt'}\right)=\left(\frac{dp_{\parallel}}{dt}-\frac{u}{c^2}\frac{dE}{dt}\right)\gamma_u\frac{dt}{dt'}\tag{3-8-19a}$$

$$F'_{\perp}=\frac{dp'_{\perp}}{dt'}=\frac{dp_{\perp}}{dt}\frac{dt}{dt'}\tag{3-8-19b}$$

其中：

$$\frac{dt}{dt'}=\frac{1}{\gamma_u}\left(1-\frac{uv_{\parallel}}{c^2}\right)^{-1}\tag{3-8-20}$$

$$\frac{dE}{dt}=\frac{d}{dt}\sqrt{p^2c^2+m_o^2c^4}=\frac{c^2}{E}\mathbf{p}\cdot\frac{d\mathbf{p}}{dt}=\mathbf{v}\cdot\frac{d\mathbf{p}}{dt}=\mathbf{F}\cdot\mathbf{v}\tag{3-8-21}$$

代入(3-8-19)可得到作用力两个分量的变换方程分别为：

$$F'_{\parallel}=\left(1-\frac{uv_{\parallel}}{c^2}\right)^{-1}\left[F_{\parallel}-\frac{u}{c^2}\mathbf{v}\cdot\mathbf{F}\right]\tag{3-8-22a}$$

$$F'_{\perp}=\left(1-\frac{uv_{\parallel}}{c^2}\right)^{-1}\gamma_uF_{\perp}\tag{3-8-22b}$$

例 **3-8-3** 设斜面上的一个物体与斜面间的最大静摩擦系数与斜面的坡度数值大小相同，故在斜面所在的惯性系 S' 中，物体恰好能够静止在斜面上。但在另一相对于斜面系统在 **x** 方向以速度 **v** 运动的坐标系 S 而言，斜面在水平方向发生了收缩，从而导致系统的坡度增加，因此应该观测到物体下滑，如何解释这种分歧。

解：在惯性系 S 中，因所观测到的运动斜面在水平方向发生了收缩，其坡度为

$$\tan\theta=\frac{dy}{dx}=\gamma_u\frac{dy'}{dx'}=\gamma_u\tan\theta'$$

物体所受斜面的摩擦力仅与物体对斜面的正压力有关，弹力的各分量大小分别为 $N_x=N'_x$，$N_y=\gamma_uN'_y$，则总的弹力大小为

$$N=\sqrt{N_x^2+N_y^2}=\sqrt{N_x'^2+\gamma_u^2N_y'^2}=\tan\theta=\frac{dy}{dx}=\gamma_u\frac{dy'}{dx'}=\gamma_u\tan\theta'\ \blacktriangleleft$$

3-8-5 四维空间中的动量-能量矢量和四维空间中的作用力矢量*

动量和能量在不同惯性系之间的洛仑兹变换关系(3-8-17)表明，动量和能量在不同惯性系之间是混合的，这意味着动量-能量在相对论的框架中可以构成一个四维矢量 P_μ，通过协变矢量在惯性系之间变换时的标准的洛仑兹变换关系(2-5-8)得到动量和能量在不同惯性系之间的洛仑兹变换关系。

定义动量-能量四维矢量 P_μ 的要点在于能在物理上找到该矢量与自身标积所对应的不变量。由能量和动量的关系式(3-8-15)可得

$$E^2 - p^2c^2 = m_o^2c^4 \tag{3-8-23}$$

上述关系式右边是一个不依赖于惯性系的常量，这和事件空间中的间隔 $s^2 = c^2t^2 - r^2$ 类似。因而我们可以将式(3-8-23)理解为一个特殊四维矢量 P 与其自身的点积

$$P_\mu P_\mu \equiv p^2 - E^2/c^2 \equiv m_o \tag{3-8-24}$$

容易看出该四维动量矢量可以表述为物体的静止质量和四维速度矢量(见2-5-15)的乘积:

$$P_\mu = (\mathbf{p}, iE/c) = (\gamma_v m_o \mathbf{v}, i\gamma_v m_o c) = \gamma_v m_o(\mathbf{v}, ic) \equiv m_o V_\mu \tag{3-8-25}$$

四维动量矢量在上面讨论过的惯性系 S 和 S' 之间的洛仑兹变换关系为

$$\begin{aligned} P_1' &= \gamma_u(P_1 + i\beta_u P_4) \\ P_2' &= P_2 \\ P_3' &= P_3 \\ P_4' &= \gamma_u(-i\beta_u P_1 + P_4) \end{aligned} \tag{3-8-26}$$

由此可以得到物理动量和能量在这两个惯性系之间的洛仑兹变换关系(3-8-18)。

类似地，我们可以定义四维作用力矢量为物体的四维动量相对于时间原时的变化率:

$$F_\mu = \frac{dP_\mu}{d\tau} = \left(\gamma_u \frac{d\mathbf{p}}{dt}, \gamma_u \frac{i}{c}\frac{dE}{dt}\right) = \left(\gamma_u \mathbf{F}, \gamma_u \frac{i}{c}\mathbf{F}\cdot\mathbf{v}\right)$$

其中的第四分量对应于力作用在物体上的功率的量度:

$$-icF_4 = \frac{dE}{d\tau} = \frac{d}{d\tau}\sqrt{p^2c^2 + m_o^2c^4} = \frac{c^2}{E}\mathbf{p}\cdot\frac{d\mathbf{p}}{d\tau} = \mathbf{v}\cdot\frac{d\mathbf{p}}{d\tau} = \gamma_u \mathbf{F}\cdot\mathbf{v}$$

由于四维作用力矢量满足和时空坐标相同的洛仑兹变换方程:

$$\begin{aligned} F_1' &= \gamma_u(F_1 + i\beta_u F_4) \\ F_2' &= F_2 \\ F_3' &= F_3 \\ F_4' &= \gamma_u(-i\beta_u F_1 + F_4) \end{aligned} \tag{3-8-27}$$

由此可以得到三维力矢量在这两个惯性系之间的洛仑兹变换关系(3-8-22)。

第 4 章 基本相互作用

大自然所呈现出的景象如此缤纷多彩，其根本原因在于构成这个世界的物质之间存在相互作用，即我们通常所说的“力”。一切自然现象，包括自然界本身的结构和演化都取决于这些相互作用。例如，天体之间的相互作用形成了星系的特殊构造以及和星球在其中的运行秩序，细胞内部物质之间的相互作用决定了各种生物的具体形态，并控制着细胞包括生老病死的生理及生命过程。

“力”的概念在牛顿力学中处于中心地位，但它并不是一个可以直接观测的物理量。牛顿定律并没有揭示出力的本质或起源，而只是从因果关系的角度通过“力”作用于物体产生的效应给出了量度“力”的操作性定义和方法。也就是说，牛顿定律并没有给出力学问题的完整描述。事实上，我们至今仍然不能清楚地说明作用力的起源，也不能真正描述两个物体之间力的作用形式，我们所知道的是力作用在物体上可以使物体的运动状态发生可观测的变化，以及物体在“力”的作用下其运动状态发生变化时必须满足的一些**约束规则**，如各种守恒定律。

本章将对前两个问题在经典尺度上作一定程度的探讨性的讨论，通过对一些具体的实例的讨论，为大学生读者提供学习、体会如何通过一般的观察，揣测与论证，模拟与对比，实验设计等探索未知问题的物理学方法。

4-1 相互作用“荷”

我们在处理物体的受力问题时通常采用所谓的隔离物体方法：将物体受到的环境中其它物质的各种影响抽象为作用在该物体上的各种“力”。或者说一个物体如果受到外“力”的作用，其周围就一定存在着其它物体。牛顿第三定律也给出了作用力必须满足的一个规则：作用力和反作用力一定是成对出现的，物体本身既是施力的物体，也是受力的载体。孤立的物体不可能受到任何力的作用。这些概念实际上都揭示出了力的物质起源。另外，发生在两个物体之间的作用力和反作用力总是具有相同的性质，例如摩擦力的反作用力一定是摩擦力，弹性力的反作用力也一定是弹性力。这些概念所反映的是两个具有相互作用的物体中一定存在着某种相同类型的物质属性，我们将这类与某种相互作用对应的物质属性称为“**相互作用荷**”。这样，两个物体之间能够发生摩擦效应的原因就归结为这两个物体内部都具有“摩擦作用荷”，而两个正在发生弹性力作用的物体则可以解释为两个物体中的“弹性力荷”正在彼此作用。这样就将物体之间不同类型的相互作用归类为不同的“相互作用荷”。作用力和反作用力具有相同的本质还说明了不同的“相互作用荷”之间不会发生相互作用。例如一个物体中的“电荷”不可能与另一个物体的“引力荷”发生相互作用。而两个物体之间的相互作用的强弱可以作为它们所携载的“作用荷”大小的量度。例如，太阳和地球之间的相互吸引的原因可以解释为太阳和地球都具有一种称之为“**引力荷**”的物质属性，而太阳和地球之间的引力大小则反映了太阳和地球所具有的引力荷的数量。

如果一个物体可以和其它物体发生多种相互作用，可以理解为该物体具有多种不同的“**相互作用荷**”，但出于人们对于物质构成持有的朴素信念，即世界万物都只是由很少几个基本单元构成(例如门捷列夫的元素周期表揭示出所有宏观物

质都只是由几十种原子调配而成，而二十世纪物理学的重大成就之一是发现所有的物质都可分解为 6 种夸克和 6 种轻子)。看着物质的终极构成越来越简单的同时，人们很自然地意识到使这些物质基元结合在一起，或处于一种运动状态的相互作用种类也一定非常少，甚至可能是单一的。这样，不同的物质可以被理解为相同的基元在相同的相互作用下不同的运动状态的表现形式。事实上，我们在对各种不同表现形式不同的作用力的实际观测和分析中确实发现，具有不同表现形式的作用力可能源于不同的本质，也可能出自相同的起源。例如，下落的物体同时受到的重力和空气阻力属于不同性质的相互作用，而物体之间的摩擦力，弹力，绳索中的张力则起源与同一本质。现代物理学的基本任务之一就是试图找出这些最基本的作用力，以及这些最基本的作用力的特征，其中是否存在共同的属性。物理学在这些领域中已经获得了难以置信的进步：例如，电力和磁力曾在被发现后的很长时间中一直被认为是两种独立的相互作用，而到了十九世纪，人们发现电力和磁力是属于统一本源的作用力。再例如，20 世纪中叶已经发现，自然界的最基本相互作用不会超过四种：万有引力，电磁力，弱作用力和强作用力。相应地也只存在四种相互作用荷。而到了 20 世纪 70 年代，电磁力和弱作用力又被归属于同一种作用荷。物理学们正在做的工作之一就是试图将另外的两个相互作用力也能同样地和谐共处于同一个数学框架，并以此描述整个宇宙的历史进程：它的起源，演化和宿命。统一所有的基本相互作用长期以来一直被许多自然科学家认为是自然科学所追求的终极目标。

对基本相互作用的理解目前仍然是不完善的，而即使对这些尚不完整的认识的描述也要涉及大量的数学工具和高等量子理论，而这些知识都远远超出了这本基础物理教科书的范围，因而我们在本章中无法涉及许多的细节，而只是对这些基本相互作用的主要特征做一些定性、甚至只是唯像的描述。对相关领域的概况有兴趣的读者可以阅读后面所列的一些介绍性的读物。

思考题

4-1-1 物体在非惯性系中受到的惯性力是否作用在物体中的某一类的作用荷上(答：处在非惯性系中的物体意味着

细心的读者也许会意识到这样一个问题：既然我们对力的认识都是来自于力作用在物体上所产生的效应，而所有的这种效应只有两个物体之间的排斥和吸引，我们是如何来判断有四种基本相互作用的呢。我们对基本相互作用的分辨是来自这些作用的大小和力程。如引力和电磁力都是长程力，而强作用和弱作用则是短程力。另外不同的基本相互作用有不同的效应选择性。

由于相互作用强度总是通过其作用在物体上，观测导致物体运动状态的变化的程度来实现度量的。研究作用力的这个实验特点表明，实际上不可能直接测量到相互作用荷的大小，实际测量到的参量是作用荷与惯性质量的比值(**荷质比**)。因此，不同类型相互作用的荷质比是理解相应的相互作用特征的重要参量。

需要说明的是，力并不是在任何场合都是必不可少的角色。在近代物理学中，牛顿运动方程被爱因斯坦方程，薛定谔方程，狄拉克方程取代，在这些方程中抛头露面的物理量是能量，动量，完全没有力的地位。

4-2 引力相互作用

这里所要讨论的“**引力**”是存在于自然界所有物质之间的一种吸引相互作用，因而也常被称为“**万有引力**”。引力是是宇宙中天体运行的秩序维护者，也是地球表面存在生活秩序的最基本条件。

关于引力的发现有许多近乎神话的传奇故事：牛顿因在后院的苹果树下被掉下的苹果砸了脑袋而顿悟出苹果的自由落体运动与月亮的绕地运动源于同一种相互作用。事实上，引力的发现经历了若干世纪持之以恒的艰苦观察和计算，而在这些工作中的每一次突破都与深邃的科学洞察力和坚定的信念联系在一起。

4-2-1 历史

我们现在对引力的理解起源于欧洲人早期对天空的观测。古希腊人在公元前四百年就试图描述这些天体的运动。在我们周围最容易看到的是太阳和行星，一般的观察很容易形成这样的观念，这些星体都是在以不同速度围绕着地球旋转(见图 4-2-1)。这些观念与当时(以人类自我为中心的自然意识)的宗教经典在人们思想和生活中的影响结合，在欧洲古代文化中形成了一种观念：圆是最完美的曲线，因而也应该是圣洁的天体的运动轨道。例如，古希腊最著名的哲学家柏拉图[1]就认为，恒星代表着神圣、永恒的存在，它们自始至终地围绕着地球以均匀的速度做圆周运动[2]。

图 4-2-1 长时间曝光形成的恒星圆轨道运动(“绕”北天极转动)印象。

以后的观察发现许多天体(如我们周围的行星)并不严格沿圆轨道运行。为了将它们的轨道纳入圆轨道模式，希腊天文学家阿波罗尼斯发展出一种用多个

---

[1] 柏拉图(Plato 427-347 B.C.) 古希腊哲学家。苏格拉底的学生，亚里士多德的老师。约公元前 387 年创办雅典学园，建立了以理念论为核心的客观唯心主义体系。认为理念是事物的永恒不变的“范型”，独立于个别事物和人类意识之外的实体。感性的具体事物是不真实的，是理念的不完善的“影子”或“摹本”。认识只是不朽的灵魂对理念世界的回忆，感觉是以个别事物为其对象，因而不可能是真实知识的源泉。柏拉图认为辩证法是最高级的认识，第一次将辩证法提到哲学高度。其伦理学和美学建立在理念论的基础上，认为善的理念是道德的唯一根源，美的事物是相对的、变幻的，美的理念则是绝对的、不变的。前者是后者的摹本，艺术的本质就是模仿，是“摹本的摹本”。其政治思想带有空想性质，他设计了一整套“理想国”方案，主张“哲学王”治理国家，将“哲学王”看作天主的统治者、立法者。

[2] 天文观测对西方早期的宗教和文化的形成和发展都有着深远的影响。恒星的圆轨道被认为是圣洁、庄严、永恒的象征，因而古希腊人认为圣洁的天体的运行轨道应该呈现为所谓的柏拉图圆(实际上宗教经典中的正面人物的画像也都带有这种恒星特有的印记—光圈)，而地上万物受邪恶的沾染，因而它们的运动轨道或多或少是无规则的。但为什么让圣洁的天体环绕着邪恶的地球运动，这一点即使从宗教的角度看也有些难以理解，一种可能的解释是为了让天上的神灵更容易全方位地监视地球上的万物生灵的行为。

圆进行叠加的方法：行星以均匀的速度作圆周运动，这个圆称为**本轮**。本轮再绕着以地球为中心的另一个称为**均轮**的圆匀速运动。如果必要的话，可以继续置入更多的本轮，让小本轮自身又沿着较大的本轮运动。

古希腊的托勒密[3]是第一个将上述天体环绕地球做圆轨道运行的观念和实际的观测统一在一个比较完整的体系中的天文学家。但在他的天体运行体系中，只有太阳和月亮的轨道比较简单，其他行星的运动则显得复杂，甚至没有重复性。尤其引人注目的是这些行星都会出现诡异的**逆行**现象，其中火星由西向东正常运行时所发生的由东向西逆行更为明显(见图 4-2-2-a)。为了将行星的这些行为纳入围绕地球转动的圆轨道，托勒密引入了两个新的方法来描述行星的运动：第一，不再要求地球处于均轮的中心，即行星相对于地球的转动可以是偏心的圆轨道，第二，行星的本轮以均匀的速度相对于空间的一个称为**均衡点**的定点运动，该点既不需要是地球的位置，也不要求是均轮的中心(图 4-2-2-c)。由于托勒密的理论与宗教的论述相符而被教廷认同，在长达 1400 多年的历史中一直是带有官方特许印记的主流理论。

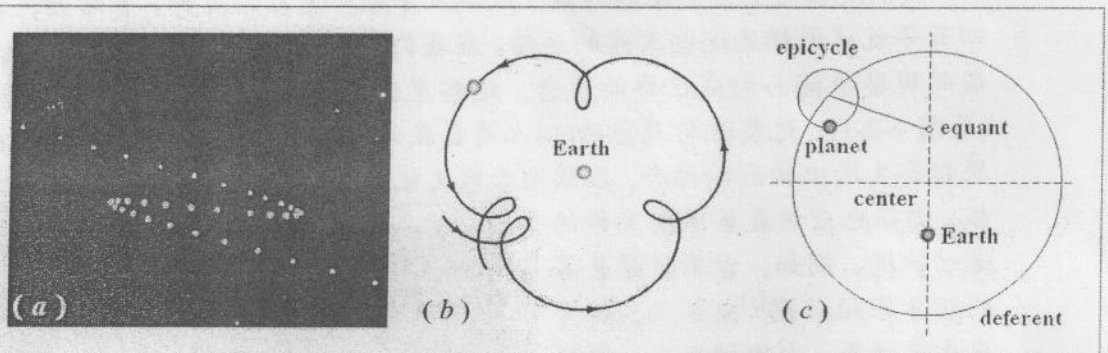


图 4-2-2 (*a*) 在地球上观测到的火星夜空逆行的轨迹，其中相邻两点的拍摄时间间隔为一星期。摄影的期间为 2005 年 7 月底至 2006 年 2 月。北京天文馆网站 (*b*) 在地球上观测到的行星轨迹示意图，(*c*) 托勒密对行星绕地球运动的描述。

虽然托勒密的体系大致给出了这些星体的运动图象描述，但为了获得同当时的观测比较符合的结果需要大约八十个均轮和本轮，而且随着观测精度的提高，对这类小轮的数量需求持续增加。这样烦琐复杂的体系很难让人相信它的物理真实性。事实上，托勒密自己也认为这个体系只是一个计算天体位置的数学方案[4]。波兰天文学家哥白尼[5]受毕达哥拉斯学派的影响，认为宇宙的规律应该以简单的

---

[3] 托勒密(C. Ptolemy Claudius)，科学史上最伟大的人物之一，但对他的生平知道的很少。他活动于公元 2 世纪上半叶，他的著作之一《至大论》(希腊文原名意为"天文学论集") 集当时天文学观测和知识之大成，自其问世直到牛顿时代前的大约 1400 年中一直是天文学领域的权威和标准。所有伟大的西方天文学家，包括对托勒密体系不满意而想有所改进的人如哥白尼、开普勒等，在他们的成长历程都受到过该著作的影响。

[4] 均轮和本轮以及移动本轮中心的方案类似于数学中用不同位置和大小的圆的叠加来逼近一条任意曲线，本质上与泰勒级数中通过将多个不同幂次的曲线叠加来逼近任意一条连续曲线相同。而当所用于叠加的曲线数量很大，即收敛很慢时，一般就会采用其它的函数系列。

[5] 哥白尼(Nicolaus Copernicus 1473-1543)，波兰天文学家，他以著作《天体运行论》中提出的地球及其它行星绕太阳运动的日心说而闻名。由于担心该书会引起各方面的攻击，哥白尼在很长的时间内不愿付诸出版，最终在身染重病的情况下才听从朋友们的劝说中辗转委托教士奥塞安德尔去办理排印工作。他在序言中注明献给教皇保罗三世，期望在这位比较开明的教皇的庇护下该书可以问世。而奥塞安德尔为确保书能安全发行，假造了一篇无署名前言，

几何图形或数学关系来表达。他在分析托勒密体系中的行星运动时注意到每个行星的运动都包含有三种周期运动：一日一周、一年一周和相当于岁差的周期运动[6]。当他试图将这三种运动都归因于地球的运动时，发现托勒密体系中用于描述行星的运行轨道的均轮数目大为减小。因此，他在其 1543 年出版的专著《天体运行论》中提出了太阳静止于宇宙的中心，包括地球在内的行星都绕太阳转动的**日心体系**。但哥白尼的革命性思想在将近一个世纪中并没有得到广泛的注意和认真对待。其原因除了与当时的宗教教义相抵触外，在对行星轨道的描述以及行星绕日运行所应具有的一些观察效应方面都缺乏观察数据的支持或有说服力的解释。例如，哥白尼体系中的行星运行轨道仍然沿用了圆轨道叠加的方案，这使描述行星绕日的真实轨道仍然需要许多大小不等的本轮和均轮，虽然减少了均轮的数目，但随着观测精度的提高均轮的数目仍然会持续增加。另外，日心体系不为当时的所谓科学界接受还有两条重要理由：一是如果地球绕着太阳转动，从其圆轨道直径的两端观测远处恒星的方位应有所改变，但在当时观测不到这种**恒星周年视差**现象(见 A1-2)。二是当时认为，如果地球处于运动状态，垂直于地面上抛物体，其落地点应该偏离原来的位置，这也与实际观测不相符合。历史上首次观测到恒星周年视差是由德国科学家贝塞尔[7]观测恒星天鹅座 61 发现的，但那已经是近三百年后的 1838 年了。而对地面上抛物体的落地点不发生偏离的现象则一直到 17 世纪伽利略阐明了惯性定律以后才得到说明。

丹麦贵族出身的第谷[8]是第一个对太阳系中的行星运动进行长时间系统观测的。基于他自己的观测，第谷一直是托勒密体系的坚定信奉者。他不仅拥有当时最好的皇家天文台和观测设备，而且视力敏锐，善长精确观察，当时虽然还没有望远镜，但他利用自己制造的观测器对行星轨道的观测精度达到了 2 弧分。足以验证哥白尼模型所需要的观测精度，他从 1576 年起 20 年如一日地进行天文观测，对 750 个星体的运行进行了记录。

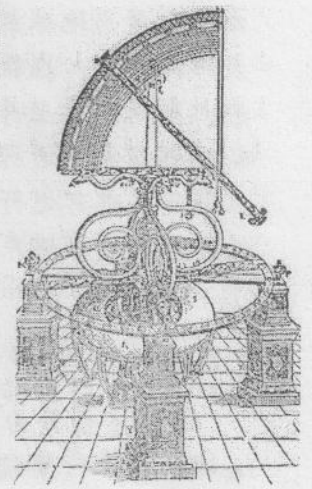


图 4-2-3 Brahe's Azimuth Quadrant 1577

第谷的观测数据表明哥白尼的行星圆轨道模型仍

说明书中的理论不一定代表行星在空间的真正运动，只是为编算星表、预推行星的位置的一种人为设计。这个说明在半个多世纪的时间里蒙蔽了许多人，许多天文工作者真的只把这本书当作编算行星星表的一种方法，因而也未引起罗马教廷的注意。后因布鲁诺和伽利略公开宣传日心说，危及教会的思想基础，《天体运行论》在 1616 年被罗马教廷列为禁书。实际上古希腊时期的阿利斯塔克等也曾提出过的地球绕太阳转动的设想。

[6] 将一个复杂的运动分解为若干个简单运动的组合是一种很有效的分析及理解问题的途径。

[7] 贝塞尔(Friedrich Wilhelm Bessel 1784-1846) 德国数学家、大地测量学家和天文学家。他在 1831-1838 年制定了东普鲁士的弧度测量方案。在地球形状理论的研究中推导出地球椭球参数，后人称 1841 年发表的地球椭球为"贝塞尔椭球"。1837 年提出计算垂线偏差的问题。对绝对重力测量的研究，倒摆的理论和研制都有较大贡献。

[8] 第谷(Tycho Brahe1546~1601) 天文学的奠基者，他在天文观测上取得了辉煌成就，他的工作为开普勒和牛顿的工作奠定了基础。第谷认为从精确测量得到的天文数据出发可以预言以后的天文事件如日食现象等。他曾从丹麦国王那里得到大笔的资助在哥本哈根附近的一个小岛上建造一个大型天文观测建筑。

然只是粗略的近似，因此在 1600 年，第谷请来了数学家开普勒[9]帮助计算行星轨道。与第谷不同的是，开普勒信奉哥白尼学说，他眼睛极为近视，却极富想象力。1601 年第谷逝世，开普勒便开始致力于整理第谷的观测数据。他以行为最离奇的火星作为第一个主要研究目标，希望通过以环绕太阳的正圆编制出的火星运行表，以此求解出描述行星轨道的规律。经过大约 70 多次调整，开普勒的计算与第谷的观察之间仍有 8 弧分的误差，大大超出了第谷的观察误差(2 弧分)。这也最终使开普勒放弃了"天体一定做圆周运动"的传统信念而开始尝试用椭圆轨道描述火星的运动。又经过近 10 年的计算，开普勒最终算出了地球和火星的椭圆轨道。至此，托勒密的本轮和均轮体系寿终正寝，日心体系得到另人信服的证据。在对行星运行轨道的探索过程中，开普勒将第谷 20 余年里观测的几千个数据最终归纳成三条著名的规律(参见图 4-2-4)：

- ✧ 行星的运动轨道是一个椭圆，太阳处在该椭圆的一个焦点上；
- ✧ 行星在运行过程中，相对于太阳在单位时间内扫过的面积相等；
- ✧ 行星公转周期的平方与其椭圆轨道的长轴的立方成正比。

开普勒三定律使哥白尼的日心体系最终获得了实验观测的基础。而火星及其它行星相对于地球发生逆行等不规则景象只是反映地球的绕日运动的正常天文现象：由于火星与地球都绕太阳运行，而地球离太阳比较近，在轨道上运行的速率比火星快，当地球追上并超越火星时，从地球上就观测到火星逆行的景象。太阳系的其它八个行星也有逆行的现象，只是木星以后的行星的逆行现象很不明显。

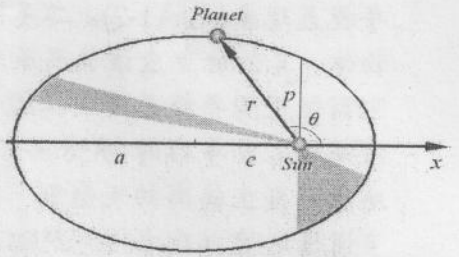


图 4-2-4 行星绕日的椭圆轨道，其中两个阴影是行星位置矢量在相同时间内扫过的面积。

4-2-2 万有引力定律

由于不受外力作用的物体一定处于匀速直线运动状态，行星绕日运行必然是受到外力作用的结果。牛顿从苹果落地联想到同样的作用力能使月球绕地球运行只是一种传奇故事，但牛顿确实由地球对苹果的非接触作用力形式而意识到，重力能从地面延伸到树梢，应该也能一直延伸到更远的地方。如果重力足以延伸到月球，它就能提供月球绕地运行的向心力。这一概念突破的关键不仅在于牛顿跳出了通常思维中作用力必须是接触力的框架束缚，而且意识到月亮的绕地运动也只是自由落体运动的一种形式(图 4-2-5)。更为重

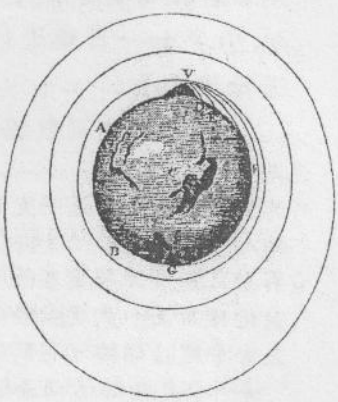


图 4-2-5 足够大的速度能使抛射物体变成一颗地球卫星

[9] 开普勒(Johannes Kepler 1571-1630) 德国天文学家和数学家，是现代天文学和现代实验光学的奠基人。开普勒三定律为哥白尼的日心说提供了最可靠的证据，被誉为"天空立法者"。

要的是，重力既然可以作用在苹果上也能作用在月亮上，这就意味着重力是一种存在于任何两个物体之间的普适作用力。这个概念上的突破使牛顿的注意力转向了描述行星运动的开普勒定律，并试图利用自己描述作用力效应的运动方程得出物体远离地面时的普适的"重力"表达式。下面我们从描述开普勒行星运动定律的三个经验公式出发，结合牛顿力学定律推导出引力定律。

开普勒第一定律的数学表述

开普勒第一定律是对行星运行轨道的描述，为了充分表现出行星以太阳为一个焦点的椭圆轨道的几何特征，我们在极坐标系中描述行星的轨道，并将太阳的位置取为坐标系原点。图 4-2-4 所示的椭圆轨道的极坐标方程为[10]

$$r=\frac{p}{1+e\cos\theta} \tag{4-2-1}$$

其中 p 称为**半焦弦**，它和椭圆的半长轴 a、和半短轴 b 之间的关系为 $pa = b^2$，而式中的 $e\,(=c/a\le 1)$ 称为椭圆的**偏心率**。采用行星和太阳距离的倒数可以将上述椭圆轨道表述为更简洁的形式：

$$u\equiv\frac{1}{r}=\frac{1}{p}+\frac{e}{p}\cos\theta \tag{4-2-2}$$

开普勒第二定律的数学表述

开普勒第二定律"行星在运行过程中，其相对于太阳的位置矢量在单位时间内扫过的面积相等"：

$$\frac{dS}{dt}=\frac{1}{2}|\mathbf{r}\times\mathbf{v}|=\frac{1}{2}r^2\frac{d\theta}{dt}=\frac{D}{2}=\text{const.} \tag{4-2-3}$$

意味着行星在运行过程中相对于太阳的角动量守恒(见 3-5 节)，即

$$L_p=m_p|\mathbf{r}\times\mathbf{v}|=m_p r^2\frac{d\theta}{dt}=m_p D \tag{4-2-4}$$

其中 m_p 是行星的惯性质量，参数 D 是行星在单位时间内扫过面积的两倍。由此可知，行星和太阳之间的引力是中心力：$\mathbf{F}=F(r)\hat{\mathbf{r}}$，因而在以太阳为原点的极坐标系中描述行星运动的牛顿方程只含有径向分量：

$$F(r)=m_p\left[\frac{d^2r}{dt^2}-r\left(\frac{d\theta}{dt}\right)^2\right] \tag{4-2-5}$$

为了从行星的轨道数据得到引力的数学表述，需要消去行星运动方程中的时间参数 t 将其转化为行星的轨道方程 $r=r(\theta)$。为此我们先将 r 关于时间的微商变换为 $u(=r^{-1})$对角度 θ 的导数

$$\frac{dr}{dt}=\frac{dr}{d\theta}\frac{d\theta}{dt}=-\frac{1}{u^2}\frac{du}{d\theta}\frac{d\theta}{dt}=-D\frac{du}{d\theta}$$

其中利用了式(4-2-4)的结果。由于 D 是常数，r 对时间的二阶微商可变换为

[10] 这实际上是圆锥曲线的一般表达式，如果太阳不在焦点，用直角坐标系描写更方便。

$$\frac{d^2r}{dt^2}=\frac{d}{dt}\left(\frac{dr}{dt}\right)=\frac{d}{d\theta}\left(-D\frac{du}{d\theta}\right)\frac{d\theta}{dt}=-D^2u^2\frac{d^2u}{d\theta^2}$$

将上式代入行星的径向运动方程(4-2-5)得到：

$$F=-m_pD^2u^2\left(\frac{d^2u}{d\theta^2}+u\right)$$

再将行星的轨道方程(4-2-2)代入上式可得

$$F=-m_pD^2u^2\left(-\frac{e}{p}\cos\theta+\frac{1}{p}+\frac{e}{p}\cos\theta\right)=-\frac{m_pD^2}{pr^2} \tag{4-2-6}$$

上式揭示出太阳与行星之间的引力与两者间距的平方成反比。但(4-2-6)式依赖于不同行星的具体轨道参量，因而(4-2-6)式只是行星与太阳之间作用力的唯像公式(类似于不同弹簧之间的作用力都可用胡克定律描述，但弹性系数各有不同，事实上，包含有具体的物性参数是唯像公式的主要特征)，没有反映出太阳与所有行星之间作用力实际具有的各种物体之间引力的普适性质。

开普勒第三定律的数学表述

由于太阳和所有行星之间的作用力应具有相同的性质，这意味着各个行星运行特性应该具有某些动力学的共性。这至少可以让我们期待式(4-2-6)能够消除具体的行星轨道参数。开普勒第三定律表明，行星的轨道参数(半长轴)与运动周期的特定函数关系与行星无关，这就给出了描述太阳行星运动共性的一个具体参量。可以通过参数 D(行星在单位时间内扫过面积的两倍)得到椭圆面积和周期的函数关系：$S=\frac{1}{2}DT$，则(4-2-6)式中的轨道参数 D^2/p 可以表示为

$$\frac{D^2}{p}=\frac{4}{p}\frac{S^2}{T^2}=\frac{4}{p}\frac{(\pi ab)^2}{T^2}=\frac{4}{p}\frac{(\pi a\sqrt{ap})^2}{T^2}=4\pi^2\frac{a^3}{T^2} \tag{4-2-7}$$

开普勒第三定律要求这个参数为与行星无关的常数：$D^2/p=k_S$，其中 k_S 仅与太阳有关，称为**太阳常数**。代入式(4-2-6)可以得到太阳和行星之间引力的普遍表达式

$$|\mathbf{F}_{s-p}|=k_s\frac{m_p}{r^2} \tag{4-2-8}$$

由于参量 k_S 仅与太阳有关，(4-2-8)式可以描述各种物体(太阳系中的行星，行星的卫星，流星，甚至太阳系中的各种尘埃)受到太阳的引力，但它仍然只是一个针对特定物体的经验公式。为了得到引力的普遍表达式，牛顿利用他的第三定律求解太阳受到行星引力的表达式。由于太阳作用在行星上的引力和太阳受到行星的引力属于作用力和反作用力，它们必然属于同一种相互作用，也应该具有相同的数学形式，行星作用在太阳上的引力因而可以表示为

$$|\mathbf{F}_{p-s}|=k_p\frac{m_s}{r^2} \tag{4-2-9}$$

其中 m_s 为太阳的惯性质量，k_p 仅与相应的行星有关。作用力和反作用力所具有的对称性要求参数 k_s 和 k_p 分别正比于太阳和行星的质量 m_s 和 m_p：

$$k_s=Gm_s,\quad k_p=Gm_p$$

其中的常数 G 应与太阳和行星无关，称为**引力常数**，由此得到两个物体之间的

引力表达式：

$$\mathbf{F} = -G\frac{m_1 m_2}{r^2}\hat{\mathbf{r}} \tag{4-2-10}$$

这就是牛顿的引力公式，也称为万有引力定律。

引力常数的测量

引力常数 G 的实验测定是由英国物理学家和化学家卡文迪许[11]在 1798 年用他自己设计的扭秤实验(见图 4-2-6)完成的。他用一根 39 英寸(约 1 米)的镀银铜丝悬吊一根 6 英尺(约 1.8 米)长的木杆，杆的两端各固定一个直径 2 英寸(约 5 厘米)的小铅球，另用两颗直径 12 英寸(约 0.3 米)的固定着的大铅球吸引它们。通过测量铅球间的引力引起的摆动周期计算出引力常数。为了防止空气的扰动(当时还没有真空设备)，卡文迪许在悬线上附加一平面镜，并用望远镜在室外远距离操纵和测量。他当时推算出的 G 数值为 $6.754\times10^{-11}\mathrm{N\cdot m^2/kg^2}$，现代测量得到的 G 的数值为

$$G = 6.672\times10^{-11}\mathrm{N\cdot m^2/kg^2} \tag{4-2-11}$$

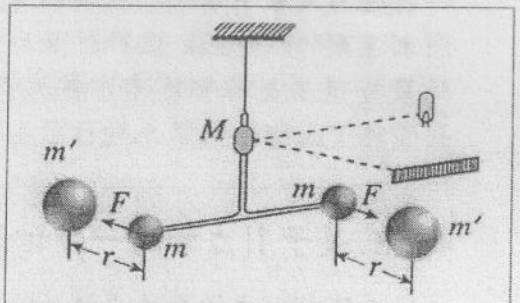


图 4-2-6 卡文迪许实验示意图

引力定律的直接实验验证

牛顿引力理论很快被英国的天文学家哈雷[12]用来研究彗星的运行轨道。不同于太阳系中的行星常年可以看见(其轨道大多接近正圆)，彗星的出现似乎并没有什么规律性，托勒密体系和开普勒系统都没有讨论过彗星的运动轨道。1682 年，哈雷观测到一颗彗星在天空中出现了几个月之久，他将实际测量到的该彗星的小部分轨道与用牛顿理论计算的轨道进行拟合后发现，该彗星的轨道是极扁的椭圆，近日点与太阳的间距为 $8.75\times10^{10}\mathrm{m}$，远日点与太阳的间距达 $5.26\times10^{12}\mathrm{m}$，其半长轴是地球半长轴是的 20 倍。近日点处彗星的速度为 $5.46\times10^{4}\ \mathrm{m/s}$，由此可算出轨道周期约 76 年。哈雷回朔以往的天文观测记录后发现了 1607 年和 1531 年观测到彗星的记录，他由此认为它们所记录的是同一颗彗星。哈雷预言了该彗星下一次的回归时间—1758 年底，1758 年的圣诞节期间确实又一次观测到了这颗彗星。这颗彗星后来被命名为哈雷彗星。现已确认，哈雷彗星的轨道距太阳最近时可进入金星的轨道以内，而最远端则跑出海王星的轨道以外(见图 4-2-7)。哈雷

---

[11] 卡文迪许(Henry Cavendish 1731~1810) 英国物理学家和化学家。在牛顿提出引力定律后的一百多年中，由于不能测定引力常数 G，一直只能由行星的卫星轨道来决定行星质量的相对值。卡文迪许算出的地球密度为水密度的 5.481 倍(地球密度的现代数值为 $5.517\mathrm{g/cm^3}$)，计算出了地球的质量。因而被誉为第一个称量地球的人。英国物理学家 J.H.坡印廷称卡文迪许扭秤的实验构思与设计“开创了弱力测量的新时代”。

[12] 哈雷(Edmund Halley 1656-1742) 英国天文学家，曾任牛津大学几何学教授，第二任格林尼治天文台台长。他正确预言了哈雷彗星的回归运动，发现了天狼星、南河三和大角三颗星的自行，以及月球长期加速现象。哈雷曾涉猎数学、地球物理、考古学等，1693 年哈雷发表的一篇关于死亡年龄分析的文章为英国政府出售寿险时确定合理的价格，提供了坚实基础。据人认为这是有关社会统计学的开创性工作，甚至对后来的人寿保险业影响不小。

彗星最近一次的回归时间是1986年。

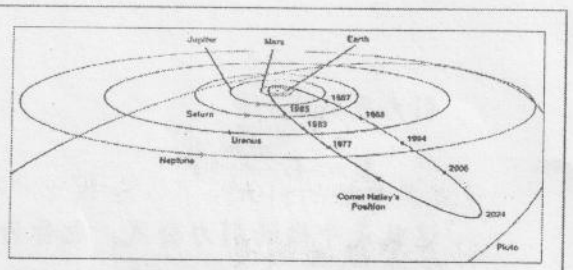


图 4-2-7 哈雷彗星的运行情况和轨道

预见并发现未知的行星一直被认为是牛顿引力理论最成功的传奇例证。自1781年英国天文学家赫歇尔(Frederick William Herschel，1738-1822)发现天王星后，大量的观测数据表明它的运动具有一些小的不规则性，由于这些不规则性不能归因于当时已知的其它星体所产生的摄动效应，人们怀疑在天王星之外还有一颗未知的行星。英国的亚当斯和法国的勒夫里艾在1846年前后各自独立地利用牛顿引力理论计算出了这颗新行星在不同时间的方位。德国天文学家戈勒在与勒夫里艾计算出的位置仅差52′的位置发现了一颗星图上没有的八等星。这颗行星后来被命名为海王星。1930年2月18日，美国的科莱特·多普则根据洛韦尔按海王星自身运动的不规则性计算出的未知行星的位置发现了太阳系的第九颗行星—冥王星[13]。

4-2-3 关于引力的进一步讨论

在万有引力的表达式(4-2-8)及其推导中存在着一些不清楚的概念，甚至混乱的地方。下面对这些地方做一些简单的说明。

(1) **引力质量和惯性质量**：由开普勒第二定律与牛顿运动方程联立得到的表达式(4-2-6)中的 m_p 为行星的惯性质量，而在通过作用力和反作用力的对称关系确定的太阳常数中的 m_s 为太阳的惯性质量。因此在万有引力的表达式(4-2-7)中并不存在4-1节中讨论的引力荷的地位。但在确定引力常数的卡文迪许实验中物体所表现出来的却是引力荷的特性。

(2) **引力常数的量纲和单位**：如果从卡文迪许实验的测量操作而言，引力常数应为无量纲常数，由此，引力荷的量纲应该由力和距离的量纲所决定，即

$$[m] = [\mathrm{F}]^{1/2}[\mathrm{r}] = [\mathrm{M}]^{0.5}[\mathrm{L}]^{1.5}[\mathrm{T}]^{-1}$$

引力方程中因强行赋予引力荷以惯性质量的量纲，导致方程两边的量纲无法平衡，只能给比例常数标记量纲(这种情形以后还会遇到)。

(3) **两体问题中的相对运动方程中的质量修正**：开普勒和牛顿对行星运行轨道的讨论都是建立在太阳的位置固定的前提上的，而作为一个多质点系统，位置恒定的是系统的质心，太阳以及所有的行星都是围绕着这个质心运动。如果仅考虑太阳和一个行星，则两体问题可以简化为行星绕日的单体运动问题。但需要以折合质量替代运动方程中的惯性质量

$$\mu\frac{d^2\mathbf{r}}{dt^2} = -G\frac{m_s m_p}{r^2}\hat{\mathbf{r}} \tag{4-2-12}$$

由于太阳质量远远大于行星质量，折合质量与行星质量非常相近，质心位置也几

[13] 冥王星的发现很不容易，因它质量很小，甚至在2006年8月24日被国际天文学联合会大会投票决定开除出了太阳系的行星行列，而将其列入“矮行星”。“行星”被定义为围绕太阳运转、自身引力足以克服其刚体力而使天体呈圆球状、能够清除其轨道附近其他物体的天体。冥王星由于其轨道与海王星的轨道相交，不符合新的行星定义，因此被降级为“矮行星”

乎与太阳中心的位置重合。因而系统的质心保持静止可以近似理解为太阳的位置是固定的。对于质量相近的双星系统，明显呈现出每一颗星都是围绕着系统的质心以相同的角速度运动的景象(见图 4-2-8)。

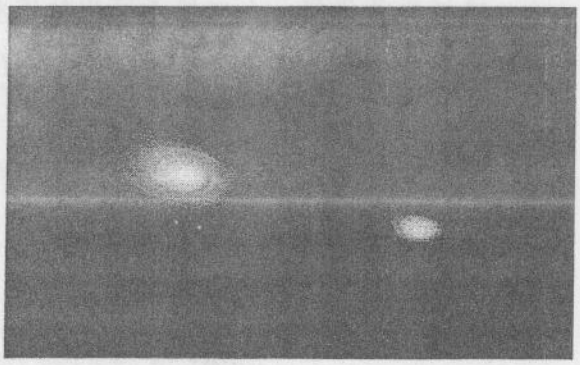


图 4-2-8 两个质量相近的双星系统的轨道运动。

(4) **对平方反比定律的偏离**：以上讨论行星轨道时只考虑了单个行星和太阳之间的相互作用，忽略了该行星和其它行星以及一些比较次要的天体之间的引力作用。可以证明，即使只计入非常微小的其他天体的作用，或者说，如果行星所受到的净作用力稍微偏离严格的平方反比定律，就会导致其运行轨道不再重合(见习题 4-2-5)。偏离最小的情况往往表现为行星的轨道本身也绕着太阳转动，这种运动称为**进动**(见图 4-2-9)。例如水星的绕日运动是水星的椭圆轨道运动和椭圆轨道绕日运动的叠加。

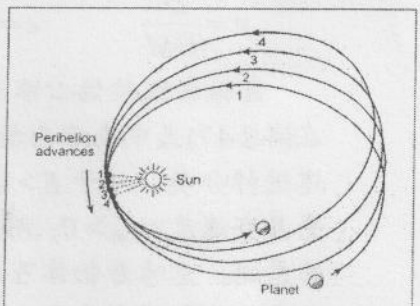


图 4-2-9 行星绕日轨道的进动

(5) **超距作用**：牛顿的引力定律摆脱了作用力必须是接触力的概念束缚，但却因此而陷入了另一个困扰：作用力是如何传递的，以及这种传递的速度问题。另一个相关的问题是万有引力定律的数学表述与相关物体的运动速度无关是严格的，或只是一种近似(例如只近似适用于物体速度远低于光在真空中的传播速度的场合)？万有引力定律的数学表述并没能反映出引力传递的速度信息，因而被称为是一种**超距作用**。引力定律虽然是从开普勒的行星轨道运动导出的，但其真正的适用范围是描述两个相对静止的物体之间的引力。

4-2-4 轨道与能量

一个物体在给定的引力中心作用下，其运动轨道可以有很大的不同。例如，从地面向上斜抛一个物体，物体的运行轨迹更象一条抛物线，而从地面上发射一颗卫星，它不仅能象月亮一样绕着地球作椭圆轨道运行，还可能摆脱地球的引力登陆其它行星，或成为太阳的行星。而太阳能将大行星束缚在以自身为焦点的椭圆轨道上运动，但却抓不住的许多外来的更小的星体，只能看着它们从自己身边掠过。这些现象都说明在相同的引力场中，物体的运行轨道与其自身的能量有关。下面分析轨道和能量之间的函数关系。

万有引力是中心力，因而也是保守力(见 3-4 节)，两个物体之间的引力势能

$$U=-\int_{\infty}^{r} G\frac{mM}{r'^2}dr'=-G\frac{mM}{r} \tag{4-2-13}$$

其中已取两物体相距无穷远时为系统的势能零点。系统总能量在极坐标中表示为：

$$E=K+U=\frac{1}{2}mv^2-G\frac{mM}{r}=\frac{1}{2}m\left[\left(\frac{dr}{dt}\right)^2+r^2\left(\frac{d\theta}{dt}\right)^2\right]-G\frac{mM}{r}$$

通过变量代换，系统能量可用轨道参数表示为(见习题 4-2-1)：

$$E=\frac{1}{2}m\left[\frac{D^2}{r^4}\left(\frac{dr}{d\theta}\right)^2+\frac{D^2}{r^2}-\frac{2GM}{r}\right] \tag{4-2-14}$$

其中参数 D 的定义见方程(4-2-4)。由此得到以系统能量作为参量的轨道微分方程：

$$D\frac{dr}{d\theta}=r\sqrt{\frac{2E}{m}r^2+2GMr-D^2}$$

积分给出：$\sin^{-1}\dfrac{2GMr-2D^2}{r\sqrt{4G^2M^2+8ED^2/m}}=\theta+\dfrac{3\pi}{2}-\theta_o$，整理后得到：

$$r=\frac{D^2/GM}{1+\sqrt{1+2ED^2/G^2M^2m}\cos(\theta-\theta_o)} \tag{4-2-15}$$

这就是极坐标中焦点在原点的标准**圆锥曲线**方程，和方程(4-2-1)比较可以得到该圆锥曲线的半正焦弦和离心率分别为

$$p=\frac{D^2}{GM}\qquad e=\sqrt{1+\frac{2E}{m}\frac{D^2}{G^2M^2}} \tag{4-2-17}$$

圆锥曲线按偏心率大于、等于或小于 1 分为双曲线，抛物线和椭圆，偏心率在(4-2-17)式中表示为物体能量的函数，我们可以用系统能量 E 的符号对运行轨道进行分类。对于 $E>0$ 的系统，物理上意味着物体即使相距引力中心无限远仍然具有速度：$v_\infty>0$，而在式(4-2-17)中对应于离心率大于 1 的双曲线。对于 $E=0$ 的系统，意味着物体可以运动到相距引力中心无限远的位置，但因为 $v_\infty=0$，这也是物体相对引力中心的极限位置，它在式(4-2-17)中对应于离心率等于 1 的抛物线轨道。对于 $E<0$ 的系统，物体不能够无限远离引力中心，对应于离心率小于 1 的椭圆轨道。容易证明，物体绕引力中心作椭圆轨道运动时物体的能量和角动量与轨道参量之间的函数关系为

$$E=-\frac{GMm}{2a} \tag{4-2-18}$$

$$L^2=\frac{GMm^2b^2}{a} \tag{4-2-19}$$

偏心率为零的椭圆轨道是圆轨道，圆轨道的动力学特征是动能和势能分别守恒。而且对于任意半径的圆轨道，物体的动能是势能绝对值的一半(见习题 4-2-5)。

物体受平方反比引力作用时，除了能量和角动量守恒，还存在另一个守恒量

$$\mathbf{A}\equiv\frac{\mathbf{l}\times\mathbf{p}}{GMm^2}+\frac{\mathbf{r}}{r} \tag{4-2-20}$$

称为**隆格-楞茨矢量**。其大小等于偏心率：$|\mathbf{A}|^2=e^2$。

例 4-2-1 宇宙速度 从地面向天空发射航天器，如果能量不够，地球引力最终将使航天器落回地面，而足够的发射能量则可以使航天器环绕地球运动。物体在近地球表面作圆周轨道运动(见图 4-2-10)时必须具备的速度称为**第一宇宙速度**(环

绕速度)。当航天器超过第一宇宙速度达到一定值时可以脱离地球的引力而成为围绕太阳运行的人造行星，该速度称为**第二宇宙速度**，也称为**逃逸速度**。而从地球表面发射航天器飞出太阳系所需要的最小速度，称为**第三宇宙速度**。计算这些特殊速度的数值。

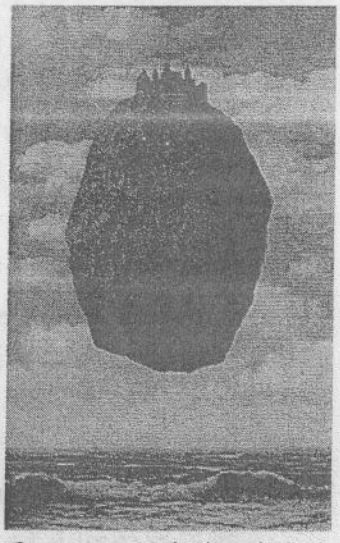


图 4-2-10 具有第一宇宙速度的《比力牛斯城堡》

解：航天器在近地球表面作圆周运动必须具备的速度，即第一宇宙速度 v_1 满足

$$mg = m\frac{v_1^2}{R_e}$$

得到 $v_1 = \sqrt{gR_e} = 7.9\,\mathrm{km\cdot s^{-1}}$。因地球表面存在稠密大气层，航天器不可能近地球表面运动，飞行高度必须达到 150 千米以上。相应的环绕速度为 7.8 千米/秒。

摆脱地球引力场的**第二宇宙速度** v_2 满足：

$$E = \frac{1}{2}mv_2^2 - G\frac{mM}{R_e} = 0$$

则 $v_2 = \sqrt{2GM/R_e} = 11.2\,\mathrm{km\cdot s^{-1}}$。由于地球表面的稠密大气层，航天器难以这样高的初始速度起飞，实际发射时是让航天器先离开大气层，再加速完成脱离，如先抵达近地轨道，在这一高度的脱离速度约为 10.9 千米/秒。航天器的实际飞行速度远低于计算值，只要航天器尾部的喷射器持续地提供向上的推力大于地球对航天器的引力，航天器就能脱离地球的引力场。

第三宇宙速度是地面上发射的航天器摆脱太阳的束缚以飞出太阳系所需的最小初始速度。在地球轨道上脱离太阳引力所需的初始速度为 42.1 千米/秒，但地球的公转使地面上所有物体具有 29.8 千米/秒的速度，若沿地球公转方向发射就只需在脱离地球引力以外再增加 12.3 千米/秒的速度。物体所需的总动能为：

$$\frac{1}{2}mv_3^2 = \frac{1}{2}mv_2^2 + \frac{1}{2}m\Delta v^2$$

由此得知所需速度为 $v_3 = \sqrt{11.2^2 + 12.3^2} = 16.7\,\mathrm{km\cdot s^{-1}}$

它实际上对应于航天器的入轨速度与地球公转速度方向一致的情形：

$$v_3 = \sqrt{2GM_S/R_e} - v_e = 16.7\,\mathrm{km\cdot s^{-1}}$$

航天器按双曲线轨迹飞离地球，而相对太阳来说它将沿抛物线飞离太阳。

类似地可以定义在地球上发射物体以摆脱银河系引力束缚所需的最小初始速度，称为第四宇宙速度。由于至今不清楚银河系的准确大小与质量，只能粗略估算出其数值在 525 千米/秒以上。至今还没有航天器能够达到这个速度。◀

4-2-5 航天器发射技术*

人类的空间技术已能够探索其它大行星：1973 年美国宇航局发射的水手 10 号到达了云山雾罩的金星和水星，1975 年，**海盗号登陆舱**抵达火星，1977 年发射了探索地外大行星—木星，土星，天王星和海王星—的旅行者 2 号。发射这些航天器最直接办法就是利用燃料提供的动力推动火箭，但利用一些特殊技术可大

大减少燃料的消耗。

转移轨道技术

转移轨道技术是先将航天器发射到一个较低的圆轨道(称为**停泊轨道**)作短暂停泊，然后再用火箭将航天器带入一个椭圆轨道，该轨道的远地点与最后要进入的目标轨道相交(见图 4-2-11)。航天器进入转移轨道后由地面控制站对其姿态进行调整。当航天器到达转移轨道的远地点时，启动航天器上的发动机调整其航向和速度，使之达到在目标轨道上运行的要求。这种变换太空船轨道的方法是由德国物理学家瓦尔特·霍曼设计的，因而称为**霍曼转移轨道**。这种技术可以将太空船送往更高的轨道，也可以将太空船送往较低的轨道，另外也用于发射地球同步卫星。

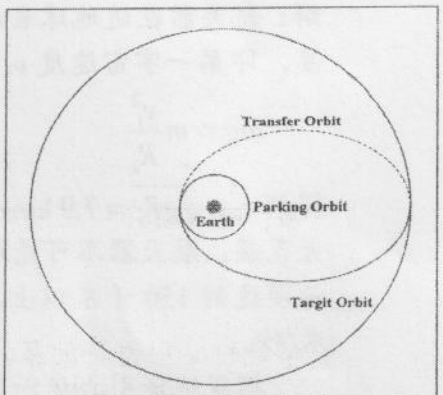


图 4-2-11 临时停泊轨道与目标轨道之间的霍曼转移轨道。

利用转移轨道技术发射行星探测器要求发射时地球和目标行星处于特定的相对位置：地球处在中转轨道的主轴的一端，目标行星处于主轴的另一端，这一时刻称为**发射时机**。由开普勒第三定律"各行星的轨道运动周期 T 与轨道半长轴 a 满足的关系 T^2/a^3= 常数"可知，不同行星的发射时机也有所不同。如金星每隔 19 个月有一次发射时机，火星为每 780 天一次，木星则是每 13 个月一次。

发射探测器除了要考虑发射时机，还必须注意从停泊轨道发射的特定位置，称为**发射窗口**。如果目标是内行星(水星或金星)，发射窗口处在太阳所照到的地球一侧(见图 4-2-12a)；如果目标是外行星，发射窗口则处在相对于太阳的地球背面一侧(见图 4-2-12b)。这是因为飞船的初始速度除了自身绕地球的轨道速度，还包含地球的轨道速度。飞船处于太阳所照到的地球一侧时，这两个速度方向相反，这时在绕地球的轨道运动方向提供额外的火箭推进可使飞船逃离地球轨道，但由于飞船相对于太阳的速度现在小于地球相对于太阳的速度，飞船会落入更靠近太阳的轨道–转移轨道。即尽管飞船轨道相对于地球而言是双曲线的逃逸轨道，但相对于太阳则是椭圆轨道。若目标是外行星，飞船从停泊轨道发射时要处于地球暗面。这时飞船绕地球的轨道方向与地球绕太阳的轨道方向相同，两者的贡献可以累加，再利用火箭的推进可以使飞船逃离地球轨道进入更大的绕日轨道(见图 4-2-12b.)。

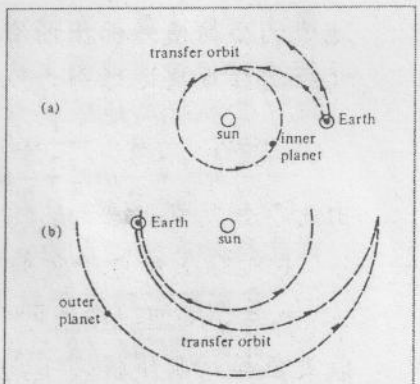


图 4-2-12 (a) 内行星**发射窗口**。(b) 外行星**发射窗口**。

物体的轨道总能量是动能与势能之和，也等于势能(轨道半径为轨道半长轴 a 时的势能)的一半：

$$E=\frac{1}{2}mv^2-G\frac{mM}{r}=-G\frac{mM}{2a}$$

由此可求解得物体的轨道速度 v 为

$$v^2 = GM\left(\frac{2}{r} - \frac{1}{a}\right)$$

假设速度改变是瞬间完成的，霍曼转移所需的两次速度改变分别为：

$$\Delta v = \sqrt{\frac{GM}{r_1}}\left(\sqrt{\frac{2r_2}{r_1+r_2}} - 1\right) \qquad \Delta v' = \sqrt{\frac{GM}{r_2}}\left(1 - \sqrt{\frac{2r_1}{r_1+r_2}}\right)$$

其中 r_1 和 r_2 分别是停泊轨道与目标轨道的半径，其中较大的和较小的值分别对应于转移轨道的远地点和近地点距离。无论前往较高或较低轨道，由开普勒第三定律可解得霍曼转移所需要的时间(即椭圆轨道运动的半周期)为：

$$t_{\mathrm{H}} = \frac{1}{2}\sqrt{\frac{4\pi^2 a_{\mathrm{H}}^3}{GM}} = \pi\sqrt{\frac{(r_1+r_2)^3}{8GM}}$$

其中 a_{H} 是转移轨道的半长轴。

下面说明计算转移轨道中涉及的具体计算过程。为简化计算，突出物理思想，假定飞船绕行星的是圆轨道，并忽略飞船摆脱地球引力的需要及飞船在环绕地球的停泊轨道上的速度。

按上面的讨论，两颗行星的轨道分别在椭圆转移轨道的近日点和远日点与转移轨道相交，确定转移轨道首先要确定椭圆的长轴 $2a$，也就是两颗行星的轨道半径 r_1 和 r_2 之和(参见图 4-2-11)：$2a = r_1 + r_2$。接着需要确定卫星在转移轨道中运动必须具有的速度。因绕日轨道的能量为

$$E = -G\frac{mM}{2a}$$

其中 M 为太阳质量，m 为飞船质量，则飞船的能量

$$E = \frac{1}{2}mv_1^2 - G\frac{mM}{r_1}$$

可解得从转移轨道出发所需要的速度 v_1，从地球的轨道速度可以计算出将飞船送入转移轨道所必需的速度增量或减少。

虽然飞向内行星的探测器在落向太阳的过程中速度不断增加，而飞向外行星的探测器则不断地失去速度，但两者都必需改变飞船的速度以适应行星绕日运动的轨道。因此当飞船抵达目标行星的轨道时，我们需要知道其抵达的速度。这个速度可以利用转移轨道中的运动过程中角动量守恒算出。如果抵达速度为 v_2，由图 4-2-11 中转移轨道中 A, B 两点的角动量守恒：$mv_1r_1 = mv_2r_2$ 可解出 v_2。

利用开普勒第三定律

$$\frac{T^2}{T_E^2} = \frac{a^3}{a_E^3}$$

可解出飞船沿转移轨道飞行所需要的时间，即轨道运动周期的一半。其中 T_E 是地球轨道的周期(1 年)，a_E (= 1 AU = 1.5×10^{11} m)是地球轨道的半长轴，或日-地之间的平均距离。T 是转移轨道的周期。

最后由飞船到达目标行星的轨道时，目标行星也恰好到达这一位置的条件确定发射时目标行星相对于地球的位置。设飞船在转移轨道上的飞行时间为 t，行星的轨道运动周期为 T_p，则参量$(360°)t/T_p$是飞船飞行途中行星的轨道运动所转

果的角度，从这一角度可以计算发射时机。

重力助推技术

飞船在向行星运动的过程中可以利用行星本身的引力场为自己提供额外的推力，称为**重力助推**。1977 年，美国宇航局利用地外的四个气态大行星每隔 175 年会排成一行的时机，向这些大行星发射了旅行者 2 号以完成一次大满贯旅行(见图 4-2-13)。利用从木星获得的重力助推，旅行者 2 号可在 12 年的航行中遍访土星，天王星和海王星。如果直接向土星发射，航行时间就需要再增加 6 年。没有重力助推时，访问天王星需要 12 年，至于海王星则根本不可能到达。

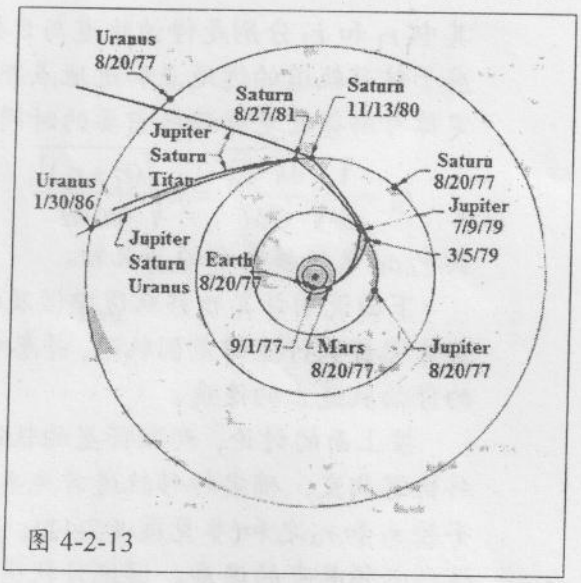


图 4-2-13

飞船飞向木星途中需要考虑太阳和木星的引力，但飞船非常靠近木星时可以忽略太阳的引力作用，将三体问题简化为可解的二体问题。

飞船从极远处以很小的初速度驶向木星时的总能量

$$E=\frac{1}{2}mv^2-G\frac{mM}{r}$$

为零或很小的正数，即环绕木星的轨道是开放的。

其结果就是飞船的速度发生了变化，但尽管在位置 2 处的速度方向与位置 1 处不同，速率却使相同的。这就是从木星上看到的情形。

确定飞船在太阳系中的运行轨道需要从太阳的角度观察这种相互作用。在这种视角中，由于木星使运动的，我们还需要考虑一些其它因素。从木星上看，飞船的轨道为抛物线或双曲线，而从太阳的角度看，飞船还必须被木星拉扯着一起绕太阳作轨道运动。即相对于太阳，飞船得到了一些额外的速度分量(即木星的速度)，使它的动能有所改变。虽然势能未发生变化(飞船与太阳的距离几乎未变)，但动能因重力助推而增加，总能量也因此增加了。由于运动发生在黄道面(太阳系中大部分星星的轨道平面)，这种在黄道平面上增加轨道能量的方式成为泵浦。

重力助推类似于用球棒击打棒球，球棒不仅使球的速度转向，也增加了棒球的速率。击球越有力，就越有机会击出本垒打。类似地，具有大速度的行星可以向飞船提供更多的重力助推。由于整个太阳系的能量守恒，飞船从木星获得的能量来自木星绕日的轨道能量，木星损失了这些助推能量后，其轨道就会稍稍靠近太阳一些。

除了上述获得重力助推的方式，另一种获得重力助推的方法是相对于木星从不同的方向发送航天器，这样，木星就能向飞船作用一个相对于太阳的力矩。该力矩可以改变飞船的角动量方向。因为角动量垂直于飞船的轨道平面，改变角动量方向意味着轨道平面不再重合于黄道平面。以上两种方法都是航天飞行中标准方法。

旅行者 2 号所揭示的激动人心的发现之一是土星的极为复杂的光环(见图 4-2-14)，其中一条从其它环中分离出来的非常有序的外环(称为 F 环)特别窄。加州理工的 Peter Goldreich 在旅行者2号飞经土星前就预言这种特别窄的外环应该是环两边各存在一个卫星造成的结果。而当旅行者 2 号飞经土星环时，确实发现了这两个卫星(见图 4-2-13)。

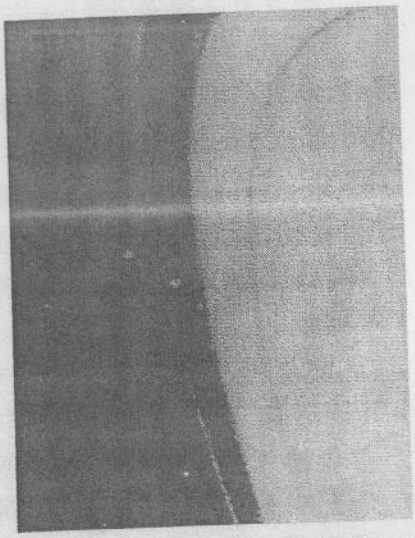


图 4-2-14 旅行者 2 号拍摄得的土星的两个卫星—普罗米修斯和潘多拉放牧 F 环印象。

为了说明为什么 F 环两侧各有一个卫星可以导致 F 环变细，先考虑两个一前一后一起运动发物体受到彼此的引力作用对各自运动的影响：显然其中运动快的物体会受到减速而失去能量，而运动较慢的物体会受到加速而获得动能。F 环内的物质和旁边卫星的相互作用也是如此。因为离行星近的卫星运动得更快，内侧的卫星就比外侧的卫星运动速度快，而环中物质的运动速度则介于两者之间。环中的物质通过与内侧卫星的引力作用而获得能量，而能量的增加使物质进入更高的轨道，即环内侧的物质受到内侧卫星引力作用的效果是向外的推力，类似地，环外侧的物质因与外部的卫星作用失去能量而减速，从而进入较低的轨道，即环外侧的物质受到外侧卫星引力作用的效果是向内的推力，因而就象牧羊人将羊聚集为一个固定的阵形，环 F 环两侧的卫星对环的动力学作用是将环压缩为一个细窄的环形。

4-2-6 潮汐与潮汐力

两个物体之间的引力与距离有关，当物体具有比较大的尺度时，作用在物体各个部分的引力不均匀性可能导致物体发生形变和一些相关的动力学效应。其中一个常见的典型事例就是地球表面上海水每天周而复始的潮汐运动，因而这种因物体各个部分的引力不均匀性导致物体中出现的张力统称为**潮汐力**或**引潮力**。

对潮汐的解释最早是由伽利略提出的。他认为潮汐是月亮对海洋的吸引作用使面向月亮的水面涌起而形成[14]。但按这个解释，潮汐的周期应和地球的自转周期一致，也是 24 小时。而实际上在同一地点一天可以观测到两次海潮，即所谓的“潮”与“汐”。

对潮汐的正确解释是牛顿提出的。如果从外部的惯性系中观察，地球各处的水面距离月亮中心的距离不同，受到的引力大小因此也不相同。设地面上两个位置与月亮的距离分别为 r 和 Δr，则两处相同质量的物体所受到的引力的差异可以用所谓的**潮汐力**描述

[14] 伽利略当时提出对潮汐的解释是希望籍此说服教会接受哥白尼的日心体系，他认为潮汐是地球运动的直接证据：由于地球运动导致海水晃动。他还进一步提出，由于地球围绕着太阳运动，地球各地的海水每天正午会出现一次高潮位。

$$\mathbf{F}_{tide} \equiv \mathbf{F}_G(r+\Delta r)-\mathbf{F}_G(r) \approx \Delta\left(-G\frac{Mm}{r^2}\hat{\mathbf{r}}\right) = G\Delta r\frac{2Mm}{r^3}\hat{\mathbf{r}} = -\frac{2\Delta r}{r}\mathbf{F}_G(r) \quad (4\text{-}2\text{-}18)$$

潮汐力在地表上的分布如图 4-2-14 所示，在 $\theta=0$ 和 π 处(即离月球最近和最远处)潮汐力背离地心而形成海水的高峰；而在 $\theta=\pi/2$ 处指向地心形成海水的低谷。随着地球的自转，一昼夜之间有两个高峰和两个低谷扫过每个地方，形成两次高潮位和两次低潮位。

从地球参照系的观测而言，因地球绕着地-月系统的质心(大约在地-月联线上距离地球中心约三分之二地球半径的位置)旋转，地球上的海水受到惯性离心力 $m\omega^2\mathbf{r}$ 的作用，距离月亮较远一面(背面)的海水受到的惯性力较对着月亮一面的海水受到较大的惯性力。这导致背对着月亮的海面也会出现水面的涌起(见图 4-2-14a)，使得地球完成一次旋转的周期内，一处海面能形成一“潮”一“汐”。另外，地球的自转，以及月亮运动都导致海水的潮位出现相对于地-月联线的东向偏离。月亮绕该点的转动周期为 27.3 天，相应地，两次海水高潮位的时间间隔大于 12 小时。由于月亮导致的海水涨潮早于地球的转动，两次海水高潮位的时间间隔约为 12 小时 25 分钟。

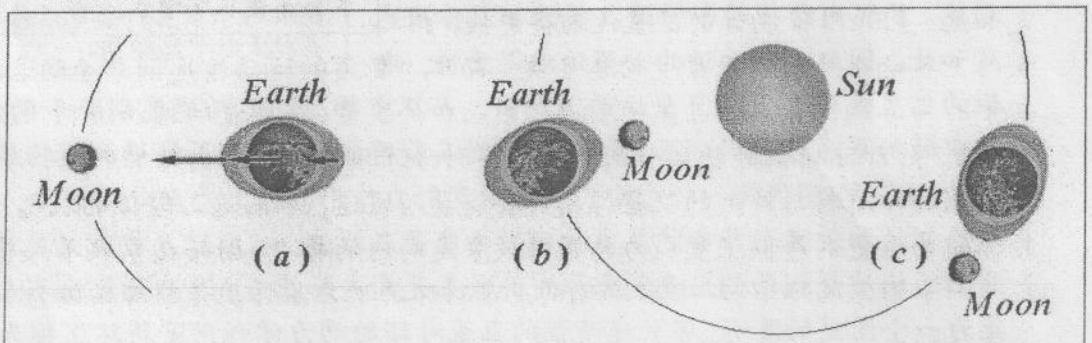


图 4-2-14 (*a*) 地球表面正对和背对着月亮的海水受到月亮引力的差异导致这两个位置海水的“潮”与“汐”。(*b*) 太阳，地球，月亮连成一条直线时，日潮与月潮重合，形成大潮。(*c*) 太阳，地球，月亮构成直角时，日潮和月潮部分抵消，导致小潮。

太阳也能产生海洋潮汐，但强度较月亮所产生的效果小很多。因为潮汐的强度并不取决于作用在海水上的引力(太阳作用在海水上的引力大约是月亮引力的 175 倍)，而是各点相对大小的潮汐力。由于地球的尺度相对于地-日距离只有大约 0.017%，而相对于地-月距离约为 6.7%。按潮汐力的表达式(4-2-18)，太阳与月亮作用在海水上的潮汐力大小比值约为 $175\times0.017/6.7\approx0.44$，不到月亮潮汐力的一半强度。

另外当太阳，地球，月亮连成一条直线(见图 4-2-14-b)时(发生在新月和满月期间)，日潮与月潮重合，形成最大的高潮位，称为大潮。而当三者构成直角(图 4-2-14-c)(发生在上弦月和下弦月)时，日潮和月潮部分抵消导致最小的高潮位，称为小潮。

潮汐力作用在固体上可以产生相应的固体潮，可以表现为固体发生形变或运动状态发生改变。例如，月球的自转和公转周期相等[15]被认为是由固体潮的长期作用(主星作用于伴星上的潮汐摩擦)的结果。下面考虑潮汐力导致固体变形的一种极限情况—破裂。

---

15 即月球总是以它的同一半球对着地球，其他行星的几个卫星也有同样的情况

设想围绕一个大天体(称为**主星**)运动的一个小天体(称为**伴星**)的轨道收缩到某一半径，如果小天体前后两端所受到的引力差(潮汐力)超过其本身的强度，伴星就会被潮汐力撕裂。下面我们对该半径做一简单的估算。

设伴星和所环绕的主星都是刚性圆球形状，伴星表面的最靠近主星的质量为 u 的部分受到伴星的引力和主星作用在 u 上的潮汐力分别为

$$F_G = G\frac{mu}{r^2} \quad 和 \quad F_T = G\frac{2Mu}{d^3}r$$

其中 r 为伴星半径，d 为伴星和主星中心的距离，若伴星恰好处在临界半径，由 $F_G = F_T$ 可得

$$d = r\cdot\sqrt[3]{2M/m} = R\cdot\sqrt[3]{2\rho_M/\rho_m} \approx 1.260R\cdot\sqrt[3]{\rho_M/\rho_m}$$

其中 R 是主星半径，ρ_M 和 ρ_m 分别为主星和伴星的密度。19 世纪法国天文学家洛希[16]研究了伴星的解体过程，解出的临界极限距离为

$$d = 2.44R\cdot\sqrt[3]{\rho_M/\rho_m}$$

若伴星是质量很小的流体团，它在潮汐力作用下很容易形变而被拉成长椭球状，在极限情形，椭球偏心率可达 0.88，洛希解出的临界极限距离

$$r_C \approx 2.423R\cdot\sqrt[3]{\rho_M/\rho_m}$$

称为**洛希极限**。

大部分卫星都不是纯粹的刚体或流体，因而其洛希极限在这两个界限之间。

设主星半径为 R，密度为 ρ_M，自转角速度为 ω，伴星密度为 ρ_m，如果伴星同步自转，该临界半径为

$$r_C = \sqrt[3]{3\rho_M/\rho_m}R = 1.44\sqrt[3]{\rho_M/\rho_m}R$$

如果一个刚体卫星的密度是所环绕的星体的密度两倍以上(例如一个巨大的气体行星和刚体卫星；对于流体卫星来说，则要约 14.2 倍以上)，$d < R$，洛希极限在所环绕的星体之内，这个卫星永远都不会因为所环绕的星体的引力而碎裂。

洛希极限是一个距离。当天体和第二个天体的距离为洛希极限时，天体自身的重力和第二个天体造成的潮汐力相等。如果它们的距离少于洛希极限，天体就会倾向碎裂，继而成为第二个天体的环。一些内部引力较弱的物体，如彗星，就可能在经过洛希极限内时化成碎片。苏梅克-列维 9 号彗星就是一个例子。它在 1992 年经过木星时碎裂，1994 年落到木星上。

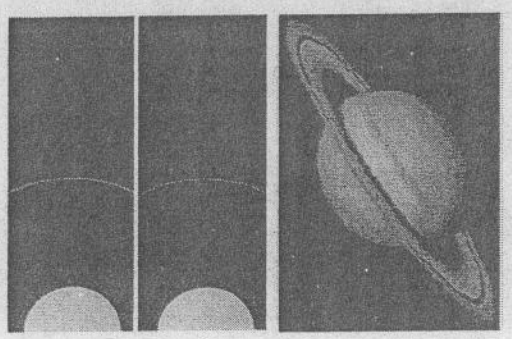


图 4-2-15 (c) 土星环

一个相关的事例是对土星环(图 4-2-10)形成原因的分析，土星环平均半径 r 与土星半径 R 之比为 r/R = 2.31，若土星环中的颗粒物质与土星本身密度相等，

[16] 洛希(Édouard Roche 1820-1883) 他认为一个巨大卫星太接近土星，以致碎散，形成了土星环。他计算了洛希极限。

则该距离已在临界距离之内，因此土星光环被认为可能由于土星的一颗卫星进入极限半径内在土星的潮汐作用下碎裂而形成的。

有些天然和人造卫星，尽管在它们所环绕的星体的洛希极限内，却不至成碎片，因为它们除了引力外，还有其他力的作用。木卫十六和土卫十八是其中的例子，它们和所环绕的星体的距离小于流体洛希极限，它们仍未成为碎片是因为它们并不完全是流体，而且具有弹性。这种情况下，卫星表面的物件可能会被潮汐力扯离卫星，—潮汐力在两个天体中心之间的直线最强。

附录

附录 4-2-1 椭圆的极坐标方程

设椭圆的右焦点处于极坐标系原点，则椭圆的几何参量满足(参见图 A4-2-1)

$$r'^2 = (r\sin\theta)^2 + (r\cos\theta + 2ea)^2$$

或

$$r'^2 = r^2 + 4ear\cos\theta + 4e^2a^2$$

对椭圆的特征关系式：$r' + r = 2a$，两边取平方给出

$$r'^2 = r^2 - 4ar + 4a^2$$

结合上面的两个方程得到

$$r(1 + e\cos\theta) + a(e^2 - 1) = 0$$

由此得到椭圆在极坐标系中的方程

$$r = \frac{a(1-e^2)}{1+e\cos\theta} \quad 或 \quad r = \frac{b^2}{a(1+e\cos\theta)}$$

其中 $b^2 = a^2(1-e^2)$。b 为椭圆的半短轴。

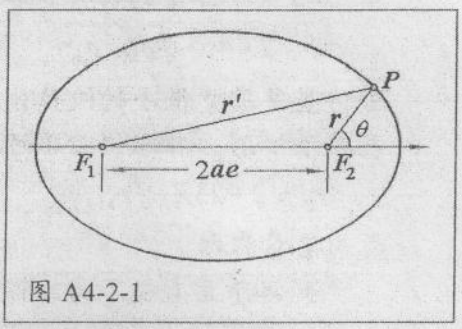


图 A4-2-1

附录 4-2-2 开普勒的火星轨道

开普勒在《新天文学》一书中描述了他确定地球轨道的方法。他的出发点是利用当时已经知道的一个火星年为 687 天。以太阳位置 S、地球位置 E_o 和火星位置 M 三者处于一条直线时火星的位置 M 点为参考点(见图 A4-2-2)，火星在 687 天后回到 M 点时，地球绕太阳转过 687/365 =1.88 圈到达位置 E_1，地球与太阳的连线 S-E_1 与原来地球

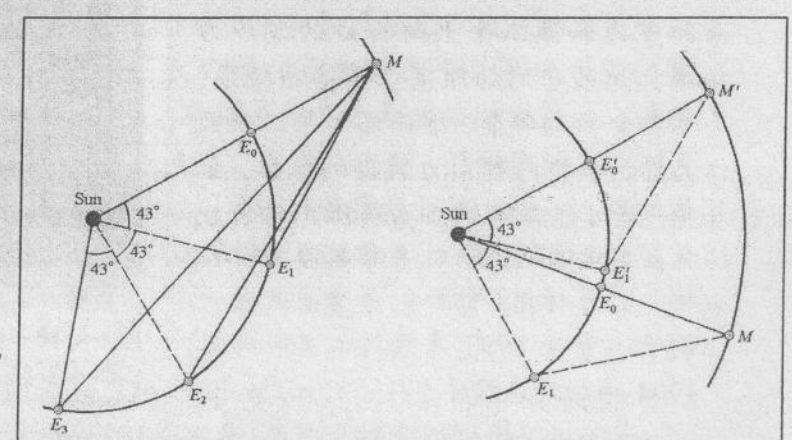


图 A4-2-2 左图：从相继的火星年的火星的固定位置确定地球的轨道，右图：利用地球轨道确定火星的轨道。

与太阳的连线 $S\text{-}E_o$ 的夹角为 43°。当火星再一次到达 M 点时，地球与太阳的连线 $S\text{-}E_2$ 与 $S\text{-}E_1$ 的角度间隔也是 43°。利用火星相继到达位置 M 时地球的位置可求得地球轨道和具体运行参数(到达各点的时间和速度)。开普勒由此发现地球的绕日轨道是一个圆，但太阳的位置有些偏离圆心。他还发现当地球距太阳最近时运动速度最快。通过分析地球在不同轨道中位置的速度，开普勒发现行星在运行过程中，行星与太阳连线在相同时间间隔内扫过的面积相等。即开普勒第二定律。

利用得到的地球轨道及其运行时间表可以反过来分析火星的绕日轨道。开普勒利用每隔一个火星年观测到的地球位置，用三角形法则求得火星轨道的位置(见图 A4-2-2 右图)，其中 M 的确定是通过沿直线 E_oM 和 E_1M sighting，然后选择第二个点，例如，下一次太阳，地球和火星三者处于一条直线时火星的位置 M'，此时地球位置为 E_o'。当火星经过一个火星年回到 M'时，地球到达位置 E_1'，由此可再次用三角形法则确定 M'。开普勒从第谷长达 25 年的观测数据记录中用这种方法确定出火星轨道上的 12 个位置。但他通过计算发现，能和这些位置适配最好的圆轨道也有 8 个弧分的差异[17]。由于第谷的观测误差仅为 2 个弧分，开普勒最终放弃被长期奉为经典的柏拉图圆轨道，并最终用椭圆拟合出了火星的轨道。他由此意识到所有行星的轨道都应该是椭圆，这就是开普勒第一定律。火星轨道是当时所能观测的行星中离心率最大的(参见表 A4-2-1)，选择火星作为第一个研究目标也是开普勒得以成功发现椭圆轨道的重要原因。

表 A4-2-1 行星轨道的离心率

行星	离心率	Comment
水星	0.206	过于靠近太阳，难于观测
金星	0.007	非常接近于圆形
地球	0.017	离心率很小
火星	0.093	当时知道的具有最大离心率的行星
木星	0.048	运动很慢 (周期为 12 年)
土星	0.056	运动很慢 (周期为 30 年)
天王星	0.047	1781 年发现
海王星	0.009	1846 年发现
冥王星	0.249	1930 年发现

基于椭圆的几何特性，开普勒研究了行星的运动周期与椭圆几何参量之间的联系，并最总结出第三定律：

行星绕日周期 T 的平方与其椭圆轨道半长轴 a 的三次幂成正比：

$$T^2 = ka^3$$

其中比例常数 k 对所有的行星都相同。

思考题

4-2-1 说明利用一系列的本轮和均轮系统来拟合行星轨道在本质上相当于一种级数展开。

4-2-2 如果质点所受力的作用线始终通过某个定点，该力称为**有心力**，该定点称为**力心**。而对于行星和太阳的实际情况，行星受到太阳的引力作用线总是通过太阳，但行星和太阳都在太空中运动，为什么仍然称行星所受到的太阳引力是有心力。

[17] 哥白尼时代的观测误差为 10 个弧分，如果哥白尼用这种方法通过火星轨道来适配圆轨道而不是采用本轮系统，他可以获得和当时的观测数据匹配得最好的轨道。

4-2-3 证明质点在有心力作用下，总是在通过力心的平面内运动。

4-2-4 一质点 m 在势能为 $-km/\rho^n$ 的有心力作用下沿半径为 r_0 的圆轨道运动，证明：如果 $n<2$，则圆轨道受到微扰时是稳定的。

4-2-5 在地面上抛射一个物体的轨迹是条抛物线(不计空气阻力)，若考虑到地球是球形的，抛射体的轨迹是什么曲线?

4-2-6 光进入太阳系后的轨道是什么样的?

4-2-7 从地球上观测到太阳系中行星运动轨道的复杂性体会选择惯性系作为观测物质运动出发点的优点。

4-2-8 对于实际的双星系统，开普勒定律需要作什么修正?

4-2-9 说明利用牛顿定律从开普勒公式得出牛顿引力定律时如何混淆了惯性质量和引力荷的概念。如果计入两者测量属性的差异，引力定律应表现为什么形式?

4-2-10 设想惯性质量和引力质量并不相等，两个惯性质量相同而引力质量不同的物体在自由落体运动过程中是否都是匀加速运动？它们的加速度是否相等？它们的重量是否相等？若这两个物体的引力质量相同而惯性质量不同，又会是什么样的情况?

4-2-11 如果单摆的摆锤 $m_I \neq m_G$，求单摆周期的表达式。用铅球和软木球作摆锤的两个摆经历上千次摆动不会出现偏差是否夸张(由于摆的质量不同，摩擦力的作用将是相当显著的)；经历上百次摆动不会出现偏差的说法是否可信。假如经历 100 次摆动后两个摆相差小于圆周的 1/10，则它对于铅球和软木球 m_I/m_G 比值的差别将给出多大限制?

4-2-12 在太空失重的情况下，宇航员怎样测量体重(是靠弹簧测量体重的)。

4-2-13 Is it possible to set the gravitational potential energy between two masses *as* $U(r)=0$ at $r=0$ or $r\to\infty$? Which choice seems more reasonable?

4-2-14 How, if at all, does experiment demonstrate that the gravitational force is conservative'?

4-2-15 为什么环绕地球的停泊轨道总是和地球的转动方向一致?

4-2-16 霍曼转移轨道的两次加速过程不是瞬间完成的，因此需要额外的燃料补偿。使用高推力引擎所需额外燃料较小，低推力引擎则还要以控制推进时间、逐渐提高轨道来逼近霍曼转移轨道。因此要花更多时间，Δv 也更大。

4-2-17 Can you directly feel the gravitational pull of the Earth with your sense organs? (Hint: would you feel anything if you were in free fall?)

4-2-18 Imagine some apple trees growing in the mountains somewhere produce apples of negative mass. In what direction would such an apple fall if it fell off its tree? How would such an apple hang on the tree?

4-2-19 为什么月食总是发生在满月期间，而日食总是发生在新月期间?

4-2-14 Explain why the sidereal day (the time of rotation of the Earth relative to the stars, or 23h 56min 4s) is shorter than the mean solar day (the time between successive passages of the Sun over a given meridian, or 24h). (Hint: The rotation of the Earth around its axis and the revolution of the Earth around the Sun are in the same direction.)

4-3 电相互作用

读者也许非常熟悉图 4-3-1 所示的实验现象：一把塑料梳子在头发上摩擦几次后可以“拾”起地面上的小纸片。就现象而言，梳子和小纸片的作用力和地球对小纸片的引力作用的测量属性相同，但也在作用过程中表现出一些明显的差异：(1) 引力和这种新的作用力的强度存在着巨大的量级差异，这把小巧的塑料梳子实际是将纸片从整个庞大的地球的引力掌握之中轻而易举地掳走的；(2) 物体之间的引力虽然很弱，但总是能够表露无遗，不能被屏蔽，而塑料梳子和小纸片之间一般并不表现出相互作用的迹象。塑料梳子是经过和头发的摩擦后才会吸引小纸片；(3) 纸片之间开始并没有相互作用，小纸片之间的吸引力似乎是由塑料梳子诱发的。

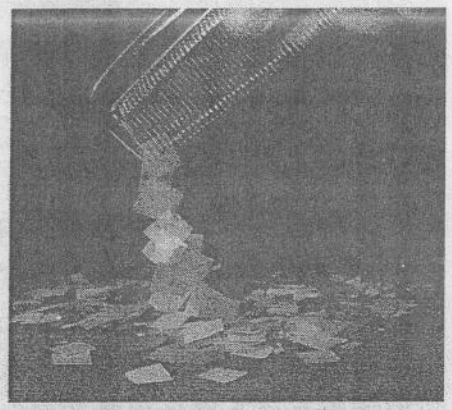


图4-3-1 塑料梳子在头发上蹭上几次就可以“拾起”地上的小纸片。

上述第一个差异让我们意识到梳子和小纸片之间的这种新的作用力与物体之间的万有引力应该属于不同的相互作用类型。或者说，纸片和塑料梳子都带有某种与“引力荷”完全不同的相互作用荷。由于历史的原因，这种“荷”被称为“**电荷**”。而电荷之间的作用力则称为电相互作用。后面两个差异意味着引力荷与电荷之间有着非常不同的特性。本节主要讨论电荷的基本性质和物体之间电相互作用的一些特征。

4-3-1 库仑定律与电荷的定义

为了理解和解释“电荷”所表现出的上述行为，我们利用一些类似的现象，通过类比的方法分析“电荷”所具有的特性。

第一个可以进行类比的是我们所熟悉的摩擦产生“热”的现象，容易想到和头发的摩擦后的塑料梳子表现出能吸引小纸片的现象可能是因为摩擦在物体中产生了电荷。但这一解释显然和电荷只是物质属性的定义相抵触(物质具有守恒性)。

另一种可供类比的现象是3-1-1节中的“反冲”：两个物体A 和B在反冲前静止，总动量为零，反冲后两个物体分别向相反方向运动。我们对此的解释是，弹簧的作用并没有在系统中产生动量，只是在A和B之间进行了动量转移。因此系统在反冲后所具有的净动量仍为零，为了说明A和B在反冲后各自所获得的动量，我们将物体A具有的动量定义为正动量，则物体B具有等量的“负动量”。类似地，设想塑料梳子和头发原来都不具有净电荷，摩擦作用也不能在系统中产生电荷，但将电荷在塑料梳子和头发之间作了转移。假设梳子上的电荷转移到了头发上，为了保证整个系统原来的净电荷量为零，则梳子上就会带有转移到头发上的电荷的等量的“负电荷”。与动量的“正”、“负”表现为物体沿两个相反方向运动不同，“正”、“负”电荷都是体现物体作用力特性的物质参量，其观测属性表现为它们之间的作用力的方向相反。具体而言一个带电荷的物体对另一个物体中的“正电荷”

或“负电荷”的作用力分别表现为吸引或排斥[1]，这样，不同类型电荷所产生的动力学效应的矢量运算就转化为同类或不同类电荷的累积，可以用标量代数的加、减运算进行描述。

另外，电荷是物质的属性这一概念还意味着电荷在物体之间的转移只能是某些携带电荷的物质从一个物体转移到了另一个物体。或者说，物体中携载“正”、“负”电荷的物质组分是可以分离的。而从牛顿定律可知，转移的物体应该是惯性质量较小的物体。按我们现在对物质结构的理解，物体的基本单元是原子，而原子中的带电物质是电子和原子核。因此，摩擦的结果是质量小，带负电的电子发生了转移，从而使失去电子的物体带正电，而获得电子的物体带有等量的负电荷。

电子可以在物体中有一定的移动能力这一事实可以解释为什么经摩擦“带电荷”后的梳子能吸引不带净电荷(电中性)的小纸片，小纸片经过带电梳子的诱惑也可以相互吸引。这是因为摩擦后带负电荷的塑料梳子(关于塑料梳子和头发摩擦后应该带负电结论的分析见思考题4-3-3)会排斥纸屑中的负电荷而导致纸屑中的电荷重新分布，纸屑中的电荷重新分布的现象称为“**极化**”，带电物体导致其周围的物体发生极化的现象称为**静电感应**。由于梳子中的负电荷和纸屑中的正电荷及负电荷之间的距离不同，梳子和纸屑之间的排斥力与吸引力不能完全抵消，从而产生一个残余的吸引力。同样，被极化的小纸片也能使邻近小纸片发生极化，从而表现出相互吸引的效应。

电子在宏观物体中运动的另一个可观察效应是通过电荷在一个宏观金属空腔中进行适当的分布，可以使放置在金属空腔内部的物体不受空腔外部电荷的电相互作用影响，这种效应称为**静电屏蔽**。

由于电子在物体中有一定的移动能力，而物质中的绝大部分质量集中在基本上不运动的原子核部分，因此可以预期，物体之间的电效应和引力效应会表现出很大的不同。这些差异也会极大地影响物质的结构和特性。例如，由于引力荷只有一种，单纯的吸引作用不仅无法屏蔽，而且物体以及引力荷的积聚不存在抑制机制。因此物体因相互吸引积聚越多，对其它物体的吸引越大，而电荷有两种，而且是不同的电荷互相吸引，相同的电荷互相排斥。这就避免了带相同电荷的物质的积聚，或者说，物质的聚集态基本上是电中性的。这两种作用荷的特征实际上决定了整个宇宙的结构和运行状态及秩序：大尺度的物体是电中性的，它们在宇宙中的运行行为是由引力决定的，而物体的结构及内部运动状态则是由电相互作用控制的。从最常见的摩擦后的物体之间相互作用特征出发，利用电荷只是反映物体之间电相互作用的物体属性的概念获得这些认识在历史上历经了整整22个世纪。

---

[1] 由于用相同方法摩擦同一种物体，物体所带的电荷应该是同类的，由此可以通过实验鉴别同类电荷的作用是排斥或吸引。例如实验发现由皮革摩擦过的两根橡胶棒是相互排斥的，两根用丝绸摩擦过的玻璃棒也是相互排斥的，即同类电荷是相互排斥的，这同时也意味着不同类型的电荷之间应该是相互吸引的。例如，用丝绸和玻璃棒摩擦后，两者是相互吸引的，而用皮革和橡胶棒摩擦后，两者也是相互吸引的。实验还发现，由皮革摩擦过的橡胶棒与用丝绸摩擦过的玻璃棒之间也是相互吸引的。这表明皮革摩擦过的橡胶棒与用丝绸摩擦过的玻璃棒所带的也不是同类型的电荷。历史上将丝绸摩擦过的玻璃棒所带的电荷被称为**正电荷**，而皮革摩擦过的橡胶棒所带的电荷被称为**负电荷**。

库仑定律与电荷的定义

按一般的认知未知现象的过程，关于电相互作用，我们接着需要确定两个电荷之间作用力大小和距离之间的关系。由作用力表现为排斥和吸引可以知道，任意两个点电荷[2]之间的作用力都是中心力，按照作用荷的定义，两个电荷之间的相互作用大小取决于两个电荷的强度(分别记为 q_1 和 q_2)以及它们之间的间距，由电荷 q_1 和 q_2 之间作用力和反作用力的对称关系可知，两个电荷的强度在相应的作用力的表达式中一定表现为乘积的关系：$F(q_1, q_2, \mathbf{r}) = q_1 q_2 f(\mathbf{r})$。

知道了引力和作用荷的函数关系以及物体间的引力和电相互作用所呈现出的行为的类似性，就不必重复牛顿从受力物体的运动轨道出发推导作用力表达式的过程，一个更便捷的方法是直接通过类比，定义电相互作用和电荷之间的关系，再通过实验确定作用力的大小与电荷间距离的函数关系。1784 年，法国物理学家库仑[3]猜测电荷间的相互作用与引力荷之间的作用具有同样的函数形式，即作用力与有关的两个静止相互作用荷的乘积成正比，与它们之间的距离成反比：

$$\mathbf{F}_{12} = k_e \frac{q_1 q_2}{r^2} \hat{\mathbf{r}}_{12} \tag{4-3-1}$$

然后直接采用卡文迪许测量引力常数的扭摆设计(图 4-3-2)，在当时的装置精度(与指数 2 的偏离 $\delta < 4\times10^{-2}$)范围内确认了电相互作用与电荷间的距离成平方反比关系。(4-3-1)式后来被称为**库仑定律**。

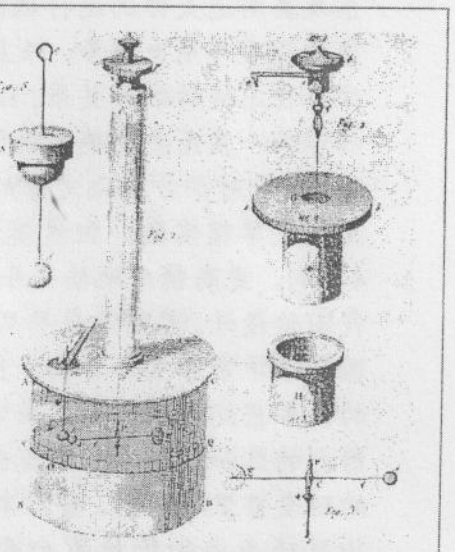


图 4-3-2 库仑验证电荷间作用力与距离呈平方反比定律的扭摆设计。

与引力定律一样，库仑定律给出的不仅是两个带电物体间相互作用的测量，同时也是对“电荷”这个物理量的操作性定义。由于仅仅从对力产生的效应的角度而言，我们并不能分辨作用力是引力还是电力，因此引力荷与电荷具有相同的量纲。为了区分这两种作用荷的差异[4]，只能人为地将它们定义成具有不同量纲的物理量。为此我们在国际单位制中将电荷的单位定义为“库仑”，记为 C。而将引力荷的单位定义为“公斤”(见 4-2 节)。为了保证力学定律的量纲平衡，我们只能象引力常数那样赋予(4-3-1)中的比例系数合适的量纲：

$$[k_e] = [F][Q]^{-2}[r]^2 = [M][L]^2[T]^{-2}[Q]^{-2}$$

在国际单位制中，k_e 通常表述为

$$k_e = 1/4\pi\varepsilon_o = 8.99\times10^9 \mathrm{C^{-2}N\cdot m^2} \tag{4-3-2}$$

其中 $\varepsilon_o = 8.85\times10^{-12}\mathrm{C^2\cdot N^{-1}\cdot m^{-2}}$ 称为**真空介电常数**。

由于对引力荷和电荷定义了不同的量纲，因而无法直接比较两者的强度。一般只能通过自然界中一些稳定的物体中对同时存在着的这两种作用力进行比较，以间接地获得两者强度之间差异的近似数量级。以最简单的原子—氢原子系统为

---

[2] 类比于质点，可以避免电荷分布和电作用的作用点等问题

[3] 库仑 (Charles Augustin Coulomb, 1736–1806) 法国物理学家，研究电学和磁学的先驱。

[4] 我们并没有充分的理由认为两者的本质或起源是完全不同的。

例，原子中电子和质子之间的引力和静电力分别为

$$F_G = G\frac{m_p m_e}{r^2} = 6.67\times10^{-11}\frac{9.11\times10^{-31}\times1.67\times10^{-27}}{(5.29\times10^{-11})^2} \approx 3.6\times10^{-47}N$$

$$F_C = k\frac{e^2}{r^2} = 8.99\times10^{9}\frac{(1.60\times10^{-19})^2}{(5.29\times10^{-11})^2} \approx 8.22\times10^{-8}N$$

静电力的大小比引力大 39 个数量级，这也说明在原子系统中，引力的作用是微不足道的。

平方反比律的特征及其在自然界中的意义

引力荷之间，以及电荷之间的作用力大小都和作用荷的间距的平方成反比。由此引起的一个令人感兴趣的问题是，这是否是现有物质世界秩序必需的要求，或者说如果这两种相互作用对平方反比定律的微小偏离会对现有的物质世界秩序产生多大的影响。

上一节已经提到，中心力偏离平方反比会导致受力物体的约束运动轨道(圆或椭圆)不封闭，例如水星的绕日运动是水星的椭圆轨道运动和该椭圆轨道绕着太阳进动的叠加。由于引力常数很小，行星对椭圆轨道的微小偏离并不会导致太阳系及其附近天体的运行秩序发生剧烈的变化[5]。但对电相互作用而言，由于其强度比引力作用大很多，而且存在有两种电荷，它们之间的屏蔽程度与平方反比定律存在怎样的依赖关系，以及这种偏离会对物质世界的构成和运行秩序产生多大的影响一直是受到关注的问题。

早期对平方反比定律的实验验证大多采用的是卡文迪许扭摆装置，库仑也采用过电单摆实验，但这类涉及直接测量的实验精度都只有 10^{-2} 左右(见思考题 4-3-4)。更高精度地验证平方反比定律需要利用平方反比律的一些特殊的性质进行实验设计。实际上最早从实验中意识到电相互作用可能满足平方反比律的是英国物理学家普里斯特利，1775 年美国科学家富兰克林注意到用一根丝线悬垂着的软木塞处在一个金属带电容器外时会受到容器的强烈吸引，而将软木塞垂入容器内的热和位置都不表现出任何受力的迹象。富兰克林将这一现象告诉了他英国的朋友普里斯特利。普里斯特利重复了富兰克林的实验后，意识到这与放在均匀物质球壳内的物质不会受到来自壳体物质引力作用的情形非常相似(见习题 4-2-5)，他由此猜测电作用力可能与引力一样，作用力大小与两个电荷之间的距离的平方成反比[6]。需要说明的是上述类比并不是完全等同的，因为只有当空腔呈球形，而且质量分布均匀时，空腔内物体受到空腔的引力作用才为零。而富兰克林实验中，任何形状的带电金属容器中的悬空带电体都不会受到容器所带电荷的作用(见思考题 4-3-10)。

1772 年，卡文迪许从普里斯特利的实验和猜测中意识到如果电荷间的作用力

---

[5] 引力偏离平方反比定律所产生的效应一般并不非常重要，因为即使在严格的平方反比定律的引力在很小的扰动下就会导致行星的运动的闭合轨道发生破缺。

[6] 普里斯特利 (Joseph Priestley 1733-1804) 英国牧师和教师。当时以化学研究而知名。他发现了氧气参与了燃烧和呼吸过程，发明了碳酸饮料(汽水)。普里斯特利的上述实验比库仑的实验结果约早 20 年。但这一猜测在当时并未引起足够的重视，普里斯特利自己对能否严格证明这一猜测又缺乏信心，这一发现就此被搁置了。

与距离平方严格成反比，均匀带电的非导体球壳内的点电荷不管处在什么位置受到球壳上电荷的作用力合力都是零；而对其中没有其他带电体的带电导体球壳，球壳的内表面是完全不带电的。若电力与距离平方反比有所偏离，均匀带电的非导体球壳内的点电荷将受到指向或背离球心的作用力，而其中没有其他带电体的带电导体球壳的内表面将带电。他由此提出了能更精确地验证电作用力平方反比律的实验方法：假设电作用力与平方反比律有一很小的指数差异：$F \propto 1/r^{2+\delta}$，计算出带电导体球壳内表面所带电量与 δ 的函数关系，这样就可以通过测量带电导体球壳的内表面所带电量大小确定 δ 的上限。因为预期的实验结果是零，实验误差就直接给出了 δ 理论值的上限。而且随着实验装置和测量技术的进步，验证平方反比律的精度可以不断地提高。这种实验设计称为**"示零"实验**[7]。

若干年后，英国物理学家麦克斯韦[8]了解到了卡文迪许的实验设计[9]，并对其作了改进(见图 4-3-3)，其中 A 和 B 是用绝缘胶木隔开的两个同心导体球壳，外球壳 A 的顶端有一个小孔。金属片 C 的大小刚好等于球壳 A 顶端的小孔，短导线的长度可使合上 C 刚好让内外球壳 A、B 联通，相当于 A 为外表面，B 为内表面。而拉开 C 可使内外球壳 A、B 分离，如果内球壳 B 原来带电，拉开 C 后 B 所带电量应保持不变。可将静电计的电极经小孔与球壳 B 接触检测内球壳 B 的电量。麦克斯韦的具体实验步骤为：

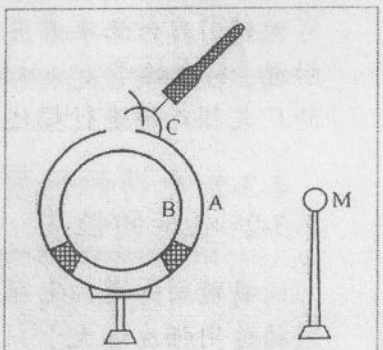


图 4-3-3

(1) 合上 C 联通球壳 A、B，对球壳 A 充电到电势 V；

(2) 充电后打开 C 使 A, B 分离，再将 A 接地放电；

(3) 将静电计电极经外球壳小孔检测 B 的电势或电量。

由麦克斯韦实验得到的平方反比律的精度为 5×10^{-5}。以后随着实验仪器和实验技术的不断改进，目前该实验的精度已经达到 $\delta<10^{-16}$。库仑定律也因此成了物理学中精确度最高的实验定律之一。由于库仑定律是整个电磁学理论的实验基

---

[7] 物理学研究中会涉及到两类典型的实验测量：一是寻找或大致确定两个变量之间的一种函数关联，二是非常精确地确定一种定量关系。前者大多是在一种事先并不清楚或并不完全确定的情况下确定一种趋势，这种实验测量大多是直接的，也并不需要非常高的精度，1%的相对误差一般已经是一个能够接受的结果。而有些定量关系涉及到一些微妙的因素，数值处于某个临界点位置，很小的偏离，就可能导致出现完全不同的效应或结果。这时，通常需要在极高的精度内进行验证，这类实验通常需要极为精致的，富有想象力的设计。而它们的实验结果的呈现形式通常都是以示零实验的形式。这类实验的设计思想大多是利用我们所需要验证的数值所处于特殊位置在物理效应中所能出现的独特的性质。

[8] 麦克斯韦(James Clerk Maxwell 1831-1879)，英国理论物理学家和数学家。电磁学理论的集大成者，也是统计物理学的奠基人之一。他对基础自然科学的主要贡献是建立了麦克斯韦方程组，预言了电磁波的存在，并揭示了光的电磁本质。麦克斯韦被普遍认为是对二十世纪最有影响力的十九世纪物理学家。物理学史上认为牛顿的经典力学打开了机械时代的大门，而麦克斯韦电磁学理论则为电气时代奠定了基石。1931 年，爱因斯坦在麦克斯韦百年诞辰的纪念会上，评价其建树"是牛顿以来，物理学最深刻和最富有成果的工作。"

[9] 卡文迪许是个专心致志于学术研究而不计功名的学者，他的许多研究成果生前都没有发表，1874 年，英国剑桥大学成立了卡文迪许物理实验室，麦克斯韦出任实验室第一任主任并负责整理卡文迪许的遗稿。卡文迪许的许多不为人知重要的研究结果才得以问世。

础[10]，高测量精度使整个电磁学理论及其它的一些重要推论[11]具有了很可靠的实验支持。

另外一个具有实际意义的问题是平方反比律的适用范围，即微观尺度极限和宏观尺度的上限分析是多少。库仑实验、电单摆实验及麦克斯韦实验中有关电荷的间距为几厘米到几十厘米，地球物理的实验表明，在大到 10^7 m 的尺度范围内，库仑力仍然表现为平方反比定律。卢瑟福的 α 粒子散射实验是以 α 粒子受原子核的库仑排斥为依据的，理论和实验的一致性表明在大于原子核的尺度，即 10^{-15}m 的范围内平方反比律仍然适用。更小尺度内库仑力的平方反比率可通过超高能电子与质子碰撞后的散射来进行研究，但得到的结果似乎表明实际电力大于库仑力。该现象现在被解释为**真空极化**。电作用力比预期的要弱(猜测其原因也可能是质子、电子并非点电荷)。

对牛顿引力的平方反比定律的实验验证远没有达到库仑定律的精度。这是因为引力的强度远小于电相互作用，这给提高实验精度带来很大的困难。另外，只有一种引力荷也不具有屏蔽效应。我们现在已经知道，在大引力荷(大质量物体)附近，引力的平方反比定律是失效的，物体在其中运动时行为的必须用爱因斯坦的广义相对论进行描述。

4-3-2 电荷的性质

目前所知道的电荷与引力荷的可观测区别，除了具有正负两种电荷以及电荷间的作用强度远大于引力荷之间的作用强度以外，电荷还具有守恒和量子化两种特性。

电荷守恒定律

前面讨论作用荷作为物质的属性时提到电荷只能在某些物理过程中在物体之间发生转移，但不能凭空产生，也不能被消灭。这就意味着整个自然界中电荷的总量是守恒的[12]。

守恒包含两种含义，一是任何自然过程不可能使自然界中的数量增加或减少，二是其数量应该是相对论不变的，即在不同的惯性参考系中应表现为不变量。就这两点而言，引力荷与电荷有很大的不同，下面分别予以说明。

在一切已知的自然过程中，所涉及的物质系统中各部分的正、负电荷数量可能发生变化，但电荷的总量是不变的。电荷的这种性质与作为引力荷的质量完全不同。例如，在化学反应中，反应物和生成物的质量一般是不相同的，物质守恒的意义只是表现为反应的原子数目在反应前后不变，而在绝对数量上守恒量是爱因斯坦的质-能复合量。这一事实在核反应和高能粒子碰撞中反映得更为明显，核反应中守恒量是更基本的质子数，中子数和电子数。一个典型的核反应的例子是铀核受热中子碰撞后发生的裂变过程：

$$n+{}^{235}_{92}\mathrm{U}\rightarrow{}^{143}_{56}\mathrm{Ba}+{}^{90}_{36}\mathrm{Kr}+3n+\mathrm{energy}$$

---

[10] 在一定的条件下，整个电磁理论可以通过库仑定律和洛仑兹变换导出。

[11] 例如，对平方反比定律的极小偏离会导致光的有效质量不再为零。

[12] 作用荷守恒不同于运动量守恒。运动量是物体的运动属性，而作用荷是物体是内在属性。

该过程中总的92个质子在反应前后没有变化，从而保证了系统总电荷的不变性。但质量的亏损转换为能量释放出来。在高能粒子反应中，通常反应前后的物质已全然不同，(静止)质量可以从有到无，也可以从无到有(见例3-8-3)。物质守恒都是在质能守恒的意义上实现的。而荷电物质在反应中随着新物质产生、旧物质湮灭也随之产生或湮灭，但这些产生或湮灭的荷电物质的电荷都是正负成对的，整个系统的总电荷不会发生任何变化。电子偶素的产生和湮灭则是质量产生和湮灭的一个典型事例。当一个电在和一个正电子(电子的反粒子，它和电子除了电荷相反，其它的各种属性完全相同)发生碰撞，这一对正、反粒子就会消失，粒子原有的静止能量和运动能量转化成为若干个互相飞离的γ光子，例如：

$$e^+ + e^- \rightarrow \gamma + \gamma$$

但在湮灭过程前后整个系统的电荷始终为零，是守恒量。

电荷守恒的另一个含义是指一个物体所携载电荷量大小不会因所观察者所处的参考系不同而发生任何改变。电荷所具有的这种运动不变性也与作为引力荷的质量完全不同：物体的惯性质量只具有相对意义：即惯性质量是物体速度的函数。而引力荷m_G与惯性质量m_I的观测等效性要求对引力的荷质比m_G/m_I守恒。但物体所带的电荷不会随物体的速度而发生变化。这也表明，质量只是伽利略变换的不变量，而电荷则是一个洛仑兹不变量。就现在的实验证据而言，电荷守恒是和动量守恒，质量-能量守恒一样的最基本的物理规律。

有上述讨论可知，物体的荷质比q/m与物体的速度有关。物体所带的电荷与参照系无关的实验验证可以从3-8-1节中关于相对论质量和动量的事例得到说明。另外我们也可以设想由于物质的电中性是依靠物质中含有等量的质子和电子维持的，质子所带的正电荷，电子所带的负电荷非常精确地相等。对于任何一个中性原子，原子核中的质子数与核外的电子数相等，因此未电离的原子和分子内的正电荷与负电荷数量相等，从而中性原子或分子所带电量非常精确地为0。在这一点上，20世纪60年代报导的实验结果的精确度已达10^{-19}至10^{-21}以上。由于原子中的电子和质子处在不同的运动状态，例如氢分子内的电子速率有$(0.01\sim0.02)c$的数量级，铯原子内K壳层电子的速率至少有$0.4c$的数量级，而原子核内的质子和中子的速率具有$(0.2\sim0.3)c$的数量级。这样的运动并未使原子和分子的电中性产生可观测的偏离。这表明电量不受运动影响，电荷是不随参考系的变换而变的。如氢分子内部电子的电量与质子的电量相等的精度至少达到10^{-20}，氦原子内质子和电子数与氢分子相同，也具有严格的电中性，但其中的电子运动状态完全不同。再例如而质子的静止质量远大于电子的静止质量($m_p \approx 1860m_e$)，这意味着，一个物体如果吸收外部能量(如光辐射，加热等)，电子速度的改变会比质子的速度变化大很多，物体会因此失去电中性。现有的生活秩序就不可能存在。再例如任何物体在加热和冷却时，电子的速度比带正电荷的原子核的速度更容易受到影响。虽然每个电子的速度可能变化不大，但物体中电子的数量极大，如果运动确实对电量有影响的话，它可以在物体上获得可观察的电量。但事实上，中性物体在任何温度下总是保持宏观上的电中性，实验中从来没有观察到仅仅通过加热或冷却的方式在物体上获得电量的事实。一个极端的例子是恒星在演化过程中其温度要经历非常显著的变化，因电子热运动速度的变化将远超过质子热运动速度的变化，电子电荷的变化也将远超过质子电荷的变化，这样恒

星整体的电中性便遭到破坏。由于电作用强度远大于引力强度，这将使宇宙中现有的天体运行秩序不复存在。这类事实都证明一个荷电体系的总电量不因自身的运动状态不同而改变。

电荷的量子特性

一个物理量的量子化，是指其数值的变化有一个最小单位，其它的变化只能是该最小单位的整数倍。但历史上似乎并没有明确记载是谁先想到电荷可能是量子化的，以及是什么让他想到了这一点。

汤姆逊在测定 e/m 值后不到二年又分别测定了电子的电荷与质量。他注意到在一定条件下，电荷可以在饱和蒸气中充作凝聚核。测定了雾滴的数目和电荷的总量，可算出电子电荷的平均值，当时他得到的数据为 3×10^{-10} 绝对静电单位(esu)。

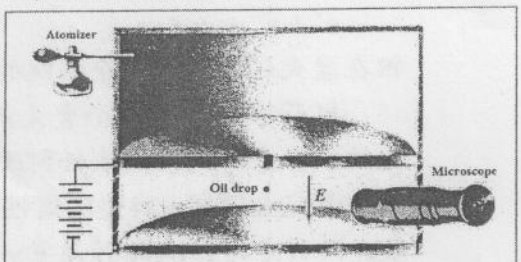


图 4-3-4 密立根实验装置示意图。

电子电荷的精确测定是 1910 年由美国物理学家密立根[13]完成的，他改进并发展了汤姆孙的方法，图 4-3-4 为密立根在 1906 年设计的验证电荷量子化的实验示意图。由油壶喷嘴喷出的油滴在通过狭窄管道时因与管壁摩擦而带有电荷。这些带电油滴因受重力作用掉入两块金属电极之间的区域后会感受到静电力的作用。通过调节金属电极之间的电势差[14]可以改变油滴的运动行为，而油滴的运动情况可以通过侧面的一个显微镜进行测量。只要知道油滴的大小和质量，通过测量油滴的运动速度就能确定油滴所带的电量[15]。这就是著名的**油滴实验**。

经过多年数以百计次的细致的测量和数据分析，，密立根发现确如他所预期的，油滴上的电荷量总可以表示为某一电荷数量的整数倍：

$$Q = ne \quad (n = 1, 2, \ldots) \tag{4-3-6}$$

其中的最大公约数 e 代表物质所能携载的最小电荷，也称为电荷的基本单元。该电荷数值为 4.78×10^{-10} esu (1.59×10^{-19}C)。在很多年中一直被认为是最精确的数值，直到 1929 年才发现来自对空气粘滞性测量的偏离导致约有 1%的误差。电子电荷的现代值为电荷是量子化的，即

$$e = 1.60217733(49)\times10^{-19}\text{C} \tag{4-3-7}$$

精度达千万分之三；括号中的 49 表示最后两位数字的误差(33±49)。现在知道一个电子，一个质子所携载的电荷都是这一基本单元，只是符号相反。每一个已经被粒子探测器所探测到的基本粒子所携带的电荷都是上述电荷量 e 的整数(正或负)倍。绝大多数基本粒子所带的电量为$+e$ 或$-e$，也有很少的粒子带有 $2e$ 的电量。

从实验测到的 e/m_e 及 e 的数值，可以定出电子的质量为：

$$m_e = 9.109\,389\,7(54)\times10^{-31}\text{kg}$$

---

13 密立根(Robert Andrew Millikan 1868-1953) 美国物理学家。他于 1909 年前后对基本电荷量作了精确测定。以后又从实验上肯定了爱因斯坦的光电效应方程，测出当时最好的普朗克常量的数值。由于这两方面的杰出工作荣获 1923 年诺贝尔物理奖。

14

15 关于密立根实验的一些细节，参阅后面的思考题和习题。

至于电荷为什么是量子化的？电子质量为什么是这个数值，物理学至今仍不能回答这些问题。但一个有趣的问题可能与此有关：即是否存在一个物体所能携载的最小电荷量的概念。如果确实存在一个物体所能携载的最小电荷量，则电荷量的变化就不是连续的，而且只能是上述最小电荷量的整数倍。

在20世纪60年代早期，人们曾相信电荷以$+e$和$-e$为单位的**量子化**是电荷的基本特性之一。1961年，美国物理学家盖尔曼[16]预言，基本粒子中的重粒子是由三种更基本的被称为夸克的粒子构成的。夸克的引人注目的特性之一是它们的电荷数量为$\pm\frac{1}{3}e$和$\pm\frac{2}{3}e$。尽管夸克的电荷是基本电荷e的分数，所有由夸克构成的粒子中所含的夸克数目总是使粒子所带的电荷数为基本电荷e的(正或负)整数倍。现在的许多物理学家都相信，夸克只有组合成净电荷为电荷量e的(正或负)整数倍的复合体才可以被分离出来，从而可以观测到。任何能够被孤立出来的电荷称为**自由电荷**，因此电荷量子化是指自由电荷为电荷量e的(正或负)整数倍。

关于电荷量子化的一些讨论

直至今天，还没有什么理论能够说明电荷为什么必须是量子化的。但直观上可以从一切物体具有共同的物质基元来理解电荷的量子化。

保持物质特征的基元是分子，而分子由原子构成，不同元素的原子则是电子和带有等量正电荷的原子核构成(原子必须呈电中性是维持现有宇宙物质秩序的必要条件之一)，原子核中有质子和中子，其中中子是电中性的，质子带有和电子等量但相反的电荷。由于任何两个电子之间，或者两个质子之间并不存在什么差异(基元的定义)，电子和质子的可数性自然导致原子、分子以至于物质的电荷量是电子电荷量的整数倍。

从逻辑上说，电荷是物质的属性，如果物质是由基本单元构成，电荷就一定是量子化的，至于其数值的大小取决于我们能对物质的分割到什么程度。或者说我们在讨论哪一个能量层次的物理过程。

从物质的稳定性而言，电荷具有量子特性目前给物理学家带来的最大好处之一是可以就此回避电子为什么不会因为电荷间的库仑斥力而发生爆炸这样的棘手问题。

在实验观测到的数以百计的微观粒子中，只有质子，电子，光子和中微子是稳定的。电荷守恒也许和物质的稳定性有关(中子被放置在原子核中也是稳定的)

电子是目前所有已知的带电粒子中质量最小的粒子。所有的静止质量来自静电能的贡献。不存在任何衰变途径可以维持电荷守恒和能量守恒。

在20世纪50年代前，质子曾被认为具有类似于电荷的一种守恒特性，而质子是具有这种特性的粒子中质量最小的。但在20世纪60年代，有一些理论(如**大统一理论**)就提出质子是不稳定的，其半衰期约为10^{31}年 (作为对比，宇宙年龄为10^{10}年)。显然直接测量这样长的半衰期是不可行的，但只要10亿摩尔的质子中有若干个质子发生衰变，我们就仍有可能

---

[16] 盖尔曼(Murray Gell-Mann 1929年) 因对基本粒子的分类及其相互作用的发现而获得1969年诺贝尔物理学奖。盖尔曼在研究核子的组分构成时发现了许多粒子质量表现出的某种对称性(规律性)。由该对称性得到的一个结论是具有某种性质的粒子可以被分为两组，分辨具有8或10个粒子。这种分组很有点象门捷列夫周期表中对元素的早期分组。1964年盖尔曼和兹维格从理论上各自独立地发现，如果假设自然界存在三种不同的粒子，称为**夸克**，所有已知的重基本粒子都是这三种夸克构成的。则重粒子的上述对称特性就可以得到简单的解释。

探测到。但至今这些探测器还没有探测到任何质子的衰变迹象。我们现在知道质子的半衰期应大于 10^{32} 年。

4-3-3 典型的电荷系统之间的相互作用力

电荷系统之间的相互作用大小可以表述为不同系统内点电荷之间相互作用的矢量和。设想两个电荷系统内的电荷分布分别为$(q_i(\mathbf{r}_i))$和$(Q_j(\mathbf{r}_j))$

$$\mathbf{F}(\mathbf{r})=\frac{1}{4\pi\varepsilon_o}\sum_{i,j}\frac{Q_j q_i}{|\mathbf{r}_i-\mathbf{r}_j|^2}$$

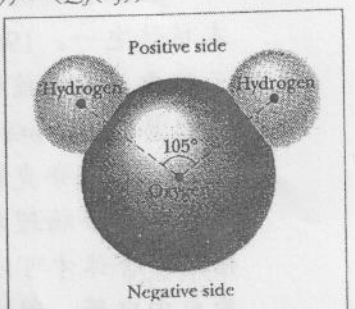


图 4-3-5

最简单的多电荷系统由两个相隔间距为 d 的正负电荷 $+q$ 和 $-q$ 构成，称为**电偶极子**，描述该系统的参量：

$$\mathbf{p}=q\mathbf{d} \tag{4-3-8}$$

称为**电偶极矩**，其方向定义为由负电荷指向正电荷。自然界中称为电偶极子是很常见，也很重要的电荷系统。许多分子由于其中的电荷分布集中于某些区域而形成固有的电偶极子。例如，水分子 H_2O(见图 4-5-3)中的氧原子处于分子 V 型结构的顶点，其总的电偶极矩是两个夹角为 105°的电偶极子的矢量和。实验数据表明其电偶极子的大小$|\mathbf{p}|\approx 6\times10^{-30}$ C·m。

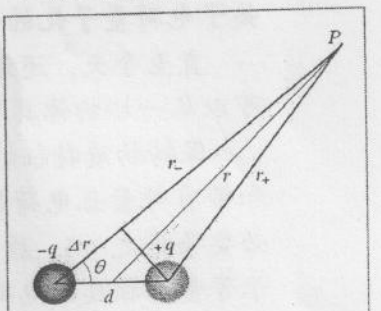


图 4-3-6

下面讨论两个电偶极子 $\mathbf{p}_1$ 和 $\mathbf{p}_2$ 之间的作用力特性。该作用力可以分解为不同电偶极子内电荷之间的作用力。我们首先计算一个点电荷 Q 在这样一个电偶极矩附近受到的作用力。

点电荷 Q 受到的电作用力是电偶极矩中点电荷 $+q$ 和 $-q$ 作用力的矢量和(见图 4-3-6)

$$\mathbf{F}(\mathbf{r})=\frac{Qq}{4\pi\varepsilon_o}\left[\frac{(\mathbf{r}-\mathbf{d}/2)}{|\mathbf{r}-\mathbf{d}/2|^3}-\frac{(\mathbf{r}+\mathbf{d}/2)}{|\mathbf{r}+\mathbf{d}/2|^3}\right]$$

$$=\frac{Qq}{4\pi\varepsilon_o r^3}\left\{\mathbf{r}\left(\frac{r^3}{|\mathbf{r}-\mathbf{d}/2|^3}-\frac{r^3}{|\mathbf{r}+\mathbf{d}/2|^3}\right)-\frac{\mathbf{d}}{2}\left(\frac{r^3}{|\mathbf{r}-\mathbf{d}/2|^3}+\frac{r^3}{|\mathbf{r}+\mathbf{d}/2|^3}\right)\right\}$$

对于 $r\gg d$ 的极限情形，上式括号中的量

$$\frac{r^3}{|\mathbf{r}-\mathbf{d}/2|^3}-\frac{r^3}{|\mathbf{r}+\mathbf{d}/2|^3}=1+3\cdot\frac{|\mathbf{d}/2|}{|\mathbf{r}|}+\cdots-\left(1-3\cdot\frac{|\mathbf{d}/2|}{|\mathbf{r}|}+\cdots\right)\approx 3\cdot\frac{|\mathbf{d}|}{|\mathbf{r}|}$$

$$\frac{r^3}{|\mathbf{r}-\mathbf{d}/2|^3}+\frac{r^3}{|\mathbf{r}+\mathbf{d}/2|^3}=1+3\cdot\frac{|\mathbf{d}/2|}{|\mathbf{r}|}+\cdots+\left(1-3\cdot\frac{|\mathbf{d}/2|}{|\mathbf{r}|}+\cdots\right)\approx 2$$

则点电荷 Q 在离开电偶极子 $\mathbf{p}$ 较远的位置受到 $\mathbf{p}$ 的作用力为

$$\mathbf{F}(\mathbf{r})\approx\frac{Q}{4\pi\varepsilon_o r^3}\left[3(q\mathbf{d}\cdot\hat{\mathbf{r}})\hat{\mathbf{r}}-q\mathbf{d}\right]=\frac{3(\mathbf{p}\cdot\hat{\mathbf{r}})\hat{\mathbf{r}}-\mathbf{p}}{4\pi\varepsilon_o r^3}Q \tag{4-3-9}$$

电偶极子产生的电场与距离的三次方成反比，其衰减比点电荷的电场迅速很多。

读者可以自己证明，两个电偶极子 $\mathbf{p}_1$ 和 $\mathbf{p}_2$ 之间的作用力为

$$\mathbf{F}(\mathbf{r}) \approx \mathbf{F}(\mathbf{r}) - \mathbf{F}(\mathbf{r}) \approx \frac{Q}{4\pi\varepsilon_o r^3}[3(q\mathbf{d}\cdot\hat{\mathbf{r}})\hat{\mathbf{r}} - q\mathbf{d}] = \frac{3(\mathbf{p}\cdot\hat{\mathbf{r}})\hat{\mathbf{r}} - \mathbf{p}}{4\pi\varepsilon_o r^3}Q \tag{4-3-10}$$

虽然，所有的电荷系统原则上都起源于点电荷在空间的分布，但通常对于宏观带电系统的问题时将电荷看成连续分布更容易处理，而且有很高的计算精度。将(4-5-4)中的点电荷替代为体积元 ΔV_i 所带的电荷 $\Delta q_i = \rho(\mathbf{r}_i)\Delta V_i$，其中 $\rho(\mathbf{r}_i)$ 为该体积元所在位置的电荷密度。其在场点 $\mathbf{r}$ 产生的电场为

$$\Delta\mathbf{F}_i(\mathbf{r}) = \frac{(\mathbf{r}-\mathbf{r}_i)}{4\pi\varepsilon_o|\mathbf{r}-\mathbf{r}_i|^3}\Delta Q_i q = \frac{(\mathbf{r}-\mathbf{r}_i)}{4\pi\varepsilon_o|\mathbf{r}-\mathbf{r}_i|^3}\rho(\mathbf{r}_i)\Delta V_i q$$

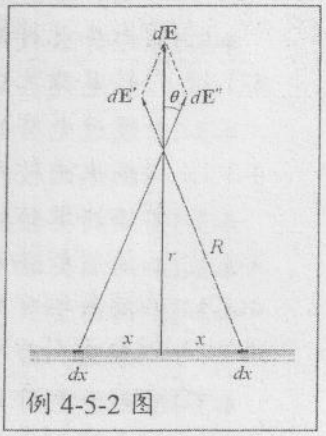
整个电荷系统对场点 $\mathbf{r}$ 处的电荷 q 的作用力为

$$\mathbf{F} = \sum_i \Delta\mathbf{F}_i = \frac{q}{4\pi\varepsilon_o}\int_V dV\rho(\mathbf{r}')\frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|^3} \tag{4-3-11}$$

对于一维或二维的连续电荷分布，电荷微元 $\rho(\mathbf{r})dV$ 分别退化为 $\lambda(\mathbf{r})dl$ 或 $\sigma(\mathbf{r})ds$，其中 $\lambda(\mathbf{r})$ 或 $\sigma(\mathbf{r})$ 分别为线电荷密度或面电荷密度。

例 4-3-1 以长度为 L 的均匀带电细杆所带电荷密度为 λ，计算一个处在垂直与细杆，且过细杆中点的平面上的点电荷受到的作用力。

解： 取细杆为 **x**-轴，细杆中心为坐标系原点(见本例图)。距原点 r 处的电荷 Q 受到原点右侧 x 处的电荷元 λdx 的作用力和其对称位置 $-x$ 处的电荷元 λdx 的作用力的 x 分量恰好抵消。因而由 x 和 $-x$ 处的两个电荷元 λdx 对 Q 的之和只有垂直于细杆的分量



例 4-5-2 图

$$dF_z = \frac{2\lambda dx}{4\pi\varepsilon_o R^2}Q\cos\theta = \frac{2\lambda dx}{4\pi\varepsilon_o r^2}Q\cos^3\theta$$

由几何关系：$x = r\tan\theta$ 将上式中的变量 x 用角度 θ 表示

$$dx = \frac{rd\theta}{\cos^2\theta}$$

则对 dF_z 积分可得

$$F = F_z = \frac{2\lambda Q}{4\pi\varepsilon_o r^2}\int_0^{L/2}\cos^3\theta dx = \frac{\lambda Q}{2\pi\varepsilon_o r}\int_0^{\theta_o}\cos\theta d\theta$$

$$= \frac{\lambda Q}{2\pi\varepsilon_o r}\sin\theta_o = \frac{\lambda LQ}{2\pi\varepsilon_o r\sqrt{4r^2+L^2}}$$

下面讨论两个极限特例：$L \gg r$ 和 $L \ll r$，相应的电场强度分别为

$$\mathbf{F} = \frac{\lambda Q}{2\pi\varepsilon_o r}\hat{\mathbf{z}} \quad 和 \quad \mathbf{F} = \frac{\lambda LQ}{4\pi\varepsilon_o r^2} = \frac{qQ}{4\pi\varepsilon_o r^2}$$

前一个特例中电场强度与距离成反比是由于无限多电荷在一维尺度中作无限均匀分布的结果。而后一个特例中电场强度与平方反比则是有限量的电荷均匀分布在有限区域的结果，其中 q 为细杆上的电荷总和。◀

思考题

4-3-1 两个相同的物体互相摩擦是否会带电？

4-3-2 分析塑料梳子与头发摩擦的过程中电荷是怎样转移的。判断与头发摩擦过

的塑料梳子是带正电还是负电？

4-3-3 两个相互作用荷之间的作用力和反作用力的对称特性为什么要表示为乘积的形式，能否表示为和的形式？

4-3-4 从摩擦并不能产生“电荷”，只是使“电荷”发生了转移的概念出发。你对所熟悉的摩擦产生“热”的现象是否有新的理解？

4-3-5 说明利用扭秤实验来验证电荷之间作用力和距离的平方反比关系时，排斥电荷实验与吸引电荷实验的精度是不同的。估计两者精度差的数量级。

4-3-6 说明利用扭秤实验来验证电荷之间作用力和距离的平方反比关系时，在实验中还存在哪些与引力荷情形不同的误差来源。

4-3-7 电单摆不同距离测量，不同秩序测量得到的比值不同，如何利用

4-3-8 电动摆实验检验静电力的平方反比定律时的主要误差来源(导体球中的电荷移动，绝缘球中的带电均匀性及极化，放电，电动摆实验图)。

4-3-9 麦克斯韦验证电作用力的平方反比定律的示零实验设计中的要点是内层金属中的电荷会跑到外层金属的外表面上，但即使电荷间的作用力不是平方反比定律，外层金属的外表面上

4-3-10 富兰克林-普里斯特利的实验和均匀物质球壳内的物质不会受到来自壳体物质引力作用不是完全等同的，因为只有当空腔呈球形，而且质量分布均匀时，空腔内物体受到空腔的引力作用才为零。而富兰克林实验中，任何形状的带电金属容器中的悬空带电体都不会受到容器所带电荷的作用。说明产生这种差异的原因。

4-3-11 为什么我们很纠结于电荷之间的作用力是否严格平方反比，而不讨论是否和电荷其它幂次的可能性？(答案，电荷是以此定义的)

4-3-12 通过电荷的不同分布说明物体具有不同的硬度和强度，不容易弯折。一个物体比较软。

4-3-13 油滴实验能否用水滴来取代。(答案，水容易蒸发，临界半径很小)

4-3-14 油滴实验中如何处理油滴的质量或体积的测量问题。

4-3-15 但由于引力的强度远小于电力，这在实验的精度上有了很大的难度，另一方面，由于只有一种引力荷，不能产生象电荷一样的屏蔽效应。

4-3-16 由于电荷和质量都是物质的属性，说明电荷的运动不变性可以导致一个重要推论：从逻辑上说，物体所带的电荷与参照系无关也就意味着静止质量为零的微观粒子只能是电中性的，这和目前的实验观测是一致的。

4-3-17 在大引力荷(大质量物体)附近，引力的平方反比定律是失效的，物体在其中运动时行为的必须用爱因斯坦的广义相对论进行描述。想象在巨量电荷之间，公式是否也需要修改？

4-3-18 静电引潮力问题

4-3-19 One of the important facts expressed by Coulomb's law is that the strength of the repulsion between two like charges is exactly same as that of the attraction between two unlike charges. It is not only a meaningless but mathematical definition. Consider an example of

4-3-20 Charge e confined in an electron of size 10^{-15}m doesn't split the electron mean that e is a minimum quantity? If so what does charge of $-e/3$ in quark imply?

4-4 磁相互作用

图 4-4-1 所示是与图 4-3-1 中与头发摩擦过的塑料梳子地面上"拾"起的小纸片非常类似的现象：条形磁铁轻易地吸起地上的钢质别针。而且别针之间也类似纸片的静电感应变得可以相互吸引。但这两种相互作用也表现出明显的差异：磁铁并不需要通过摩擦或其它处理就能吸引钢别针，而一般非磁性物质如纸片、木屑、塑料和大多数金属也不能通过力学处理和磁铁发生这种相互作用。由上述差异可以推断磁铁和钢别针中都具有一种不同于电荷和引力荷的相互作用荷，我们称其为"**磁荷**"，相应的相互作用则称为"**磁力**"。本节主要讨论磁相互作用的一些特征。

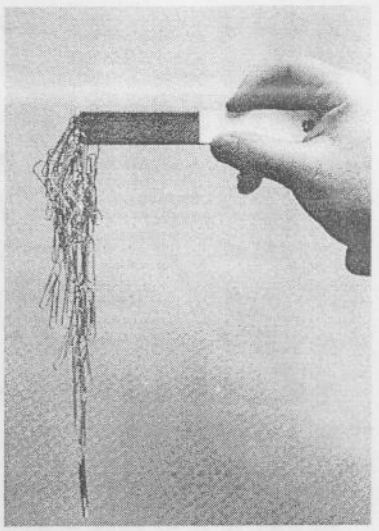


图4-4-1 条形磁铁吸起钢别针。演示了一种新的相互作用

4-4-1 磁荷与磁作用力

我们用类似上节分析电荷和电相互作用特性的方法，根据图4-4-1所示的实验现象分析"磁荷"和磁相互作用的基本特性。

(1) 钢别针之间一开始并不表现出相互吸引作用，因而磁铁吸起地上的钢质别针，以及别针之间变得可以相互吸引的现象显然是经过磁铁的某种作用使钢别针中出现了磁荷转移。为此可以按上节电相互作用导致纸片的静电感应的概念类推：设想磁铁中存在着正、负两种磁荷，磁铁导致了钢质别针中磁荷发生了迁移，这种磁荷在物质中的重新分布的现象称为物质的**磁化**。

(2) 由图4-4-1可以看出，被磁铁吸起的别针主要聚集在条形磁铁的尖端。即对条形磁铁而言，磁铁的尖端就是磁荷的主要集聚区域。另外实验发现，两根条形磁铁的两端靠近时，分别表现为相互排斥和吸引。而如果将条形磁铁水平悬挂起来，通常表现为一端指向地球的南极，另一端则指向地球的南极。这表明地球是一个大磁体，其中一类磁荷集聚于地球北极，另一类磁荷集聚于地球南极(见图4-4-2)，因而通常将指向地球南极的磁铁一端的磁荷称为"**南磁极**"(记为S)，指向地球北极的磁铁一端的磁荷称为"**北磁极**"(记为N)。

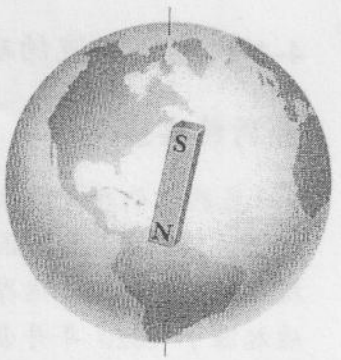


图 4-4-2 地球是一个大磁体，其南磁极在北半球，北磁极在南半球。

(3) 虽然磁极间的磁相互作用与电荷间的电相互作用形式非常相象，但作为物体特性的电荷和磁荷的表现形式却有很大的不同：物体很容易带电而破坏其原有的电中性，例如一般的摩擦可以使大多数的物体带电，有些物质通过加热也可以(因失去电子)带电。但所有的物质都极其精确地呈现为整体的磁中性，至今人们还从未能破坏过一个物体的磁中性而让任何一个物体整体带有净磁荷(称**为磁**

单极)。读者也许会想到两端带有异号磁荷的条形磁铁如果断成两截，每一段应该都应分别带有净的磁荷。但实际结果是两个断面会分别呈现出等量的南磁极和北磁极以保证两段磁棒各自仍然保持磁中性(见图 4-4-3)。如果将这种实验可以不断地延续下去，即使分割到仅剩下了一个电子、一个质子或中子，这些微观粒子仍然表现为一个磁中性的微型磁棒，这种两端分别带有异号磁荷的微小系统通常被称为**磁偶极子**。我们至今并没有观察到过任何孤立磁单极的迹象，但物理学家们却始终兴趣盎然地探索着其存在的可能性。

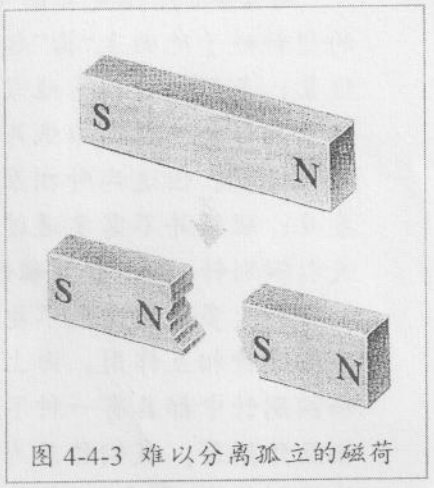


图 4-4-3 难以分离孤立的磁荷

物体中的两种磁荷不能分离也就意味着磁荷不能转移，则物体发生磁化的一种可能的物理图像是可磁化物体中存在有许多小磁针，它们在物体中取向各异，宏观上表现为磁中性。当物体附近存在磁铁时，磁铁中和物体中的磁荷之间的相互作用导致物体中的磁针取向大致趋向一致，从而表现出物质的磁化。

类似于对引力和库仑力的描述，理解磁相互作用首先需要知道它关于磁荷和磁荷间距的依赖关系。由于不能分离出孤立磁荷，我们无法直接通过实验来测量两个磁荷之间的作用力。但通过比较两个小磁针之间的磁作用力和两个电偶极子之间的电作用力的特征(见例 4-3-3)，可以发现两者满足同样的规律(见习题 4-4-2)。由此可推知，两个孤立磁荷 g_1 和 g_2 之间的相互作用力沿两个磁荷的连线方向，力的大小与两个磁荷间距的平方成反比

$$\mathbf{F} = k_m \frac{g_1 g_2}{r^2} \hat{\mathbf{r}} \tag{4-4-1}$$

上式通常称为**磁库仑定律**，同时也是磁荷 g 的定义式。其中的比例系数 k_m 的大小与量纲取决与(4-4-1)式所采用的单位制以及为磁荷强度 g 所定义的量纲。

4-4-2 稳恒电流的磁效应

奥斯特实验

虽然人们对磁现象的认识可以追溯到 13 世纪，但以后的几百年中几乎没有什么进展。其主要原因也和无法分离两类不同磁极有关：两个形影相随的不同磁极相互屏蔽，使自然界展现出的磁现象远不如电现象丰富多彩。对磁学研究的突破起源于 1820 年丹麦物理学家奥斯特[1]发现了电流的磁效应。有一种相关的记载说，当时认为磁棒不会和非磁性的金属导线发生磁相互作用。奥斯特在一次课

[1] 奥斯特(Hans Christian Oersted, 1777-1851)，丹麦物理学家。受康德哲学与谢林的自然哲学的影响，奥斯特相信自然力之间可以相互转化，为此他长期致力于探索电与磁之间的相互关系，他最早(1812 年)提出光与电磁间存在联系。电流磁效应的发现应该说和他对信念的执着有关。奥斯特的研究兴趣广泛，曾研究过声引起的电现象，物质的抗磁性，液体和气体的压缩性，提炼铝的技术等等。奥斯特是一位优秀的重视实验的教师。他曾说："我不喜欢那种没有实验的枯燥的讲课，所有的科学研究都是从实验开始的"。他常常准备一些特别的课程讲述最新的科学发展。电流磁效应的就是他在一次这种课程中无意发现的。

程的实验演示时，向导线中馈入电流，演示电流是否会使下面的磁针偏离原来的指向，开始并没有出现什么迹象，当时有人认为可能是电流太小，奥斯特于是将电流调大，他和学生都目睹了导线下方平行导线放置的磁针发生偏转：磁针的 S 极垂直于由导线和磁针构成的平面向内(即向纸面内)运动，而 N 极则向外运动(见图 4-4-4)。若使导线中的电流反向，则磁针作反向偏转。

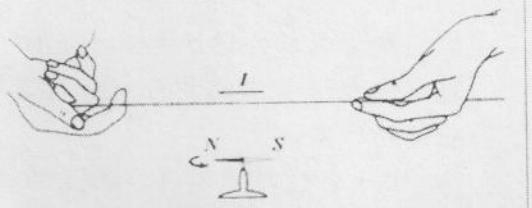


图 4-4-4 奥斯特发现的电流磁效应

这一实验发现改变了人们对磁相互作用的认识，揭示了物质间电相互作用和磁相互作用的内在联系。

毕奥-萨伐尔实验

奥斯特的发现震动了当时整个物理学界，但奥斯特并没有给出任何令人满意的解释，也没有试图用数学描述这一现象。深入理解电流与磁极之间的非接触式相互作用首先需要获得作用力与两者间距的定量关系。奥斯特的发现公布后不久，当时还是学生的两位法国年轻人毕奥和萨伐尔开始就电流对磁极的作用力进行了一系列的测量。由于当时对电流磁效应的认识还局限于电流对磁极的相互作用，磁单极的不可分离性使磁力的测量无法采用卡文迪什的扭秤方法。毕奥与萨伐尔的第一个实验是利用磁针在载流长直导线附近振荡的方法测量载流长直导线对磁极的作用力(见习题 4-4-3)。即测量在载流长直导线垂直平面内的磁针在与载流导线处于不同间距时的振荡周期。实验表明磁针振荡周期的平方正比于导线和磁针间的垂直距离，但与电流大小成反比，由于作用力与振荡周期的平方成反比，由此得到磁极受到的磁力为：

$$F = k\frac{I}{r} \tag{4-4-2}$$

其中 k 为常数。这个结果与均匀带电的长直线对近旁的点电荷的静电作用力非常相似，但由于磁单摆实验的精度并不很高，毕奥和萨伐尔设计了一个示零实验以验证(4-4-2)式中的磁力和距离之间的反比关系：在竖直的长直导线上悬挂一个水平的环状圆盘，沿盘的某一直径对称固定地放置一对条形磁铁(见图 4-4-5)，导线中的电流对每条磁棒的两个磁极的作用形成一对力偶

$$\tau = F_1 r_1 - F_2 r_2$$

如果电流对磁极的作用力与该磁极到长直导线的垂直距离成反比，一条磁棒的两极受到的力矩恰好大小相等、方向相反，圆盘维持力学平衡而不会发生运动，否则每条磁棒所受净力矩都不为零，圆盘就会扭转。实验结果显示导线馈入电流前后，圆盘没有出现扭转的迹象，这就证明了(4-4-2)式中的磁力和距离之间的反比关系。在实验设计中对称地设置两条磁棒，一方面在于维持系统的重力平衡，另一方面是因为两条磁铁所受力矩的方向相同，可以使检测灵敏度比一条磁棒的设计提高一倍。

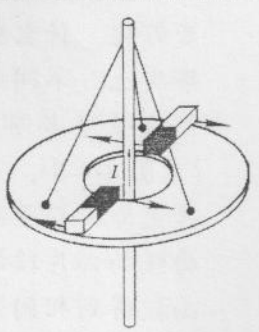


图 4-4-5 毕奥-萨伐特验证载流直导线对磁极的作用力与距离成反比的示零实验。

由于载流长直导线可理解为由许多同方向的电流元构成，毕奥和萨伐尔认为载流长直导线对其附近磁极的作用力应该是构成导线中所有稳恒电流元 $Id\mathbf{l}$ 对该磁荷的作用力的叠加：

$$\mathbf{F} = \int d\mathbf{F} = \int \frac{d\mathbf{F}}{dl} dl \tag{4-4-3}$$

他们希望能求得稳恒电流元 $Id\mathbf{l}$ 对处于位置矢量 $\mathbf{r}$ 的磁极的作用力 $d\mathbf{F}$，然后通过对电流元沿导线的积分就能计算出任意形状的稳恒电流回路对处在任一位置的磁荷的作用。

由于自然界并不存在孤立的稳恒电流元，无法通过实验测量直接获得电流元 $Id\mathbf{l}$ 对磁荷作用力在各个方位的大小。为了得到 $d\mathbf{F}$ 的数学表示，法国数学家拉普拉斯设想将任意空间构型的载流回路 L 中的一个电流元 $Id\mathbf{l}$ 作用于一个给定磁极的作用力表示为电流元 Idl 以及电流元与磁极的距离 r，电流元方向和磁极连线间的夹角 θ 的函数(见图 4-4-6)：

$$d\mathbf{F} = d\mathbf{F}(Idl, r, \theta)$$

则电流元 $Id\mathbf{l}$ 作用于磁荷的作用力在形式上表示为一个任意载流回路对处于 P 点的磁极的作用力函数关于电流元长度参量 dl 的微分：

$$dF = \frac{dF}{dl} dl = \left(\frac{\partial F}{\partial \theta} \frac{\partial \theta}{\partial l} + \frac{\partial F}{\partial r} \frac{\partial r}{\partial l} \right) dl \tag{4-4-4}$$

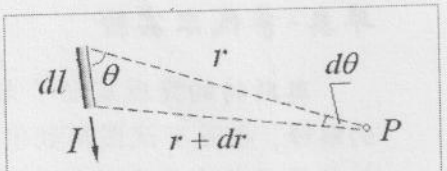


图 4-4-6 拉普拉斯电流元和磁荷之间的几何关系

虽然我们无法通过实际测量获得电流元 $Id\mathbf{l}$ 对 P 点处磁荷的作用力，但如果可以测量整个载流回路 L 作用于 P 点处磁荷的作用力，然后通过改变一个载流回路对给定位置处磁荷的特征距离和特征方位来测量整个回路对磁荷的作用力的变化数量，就能得到 F 关于电流相对于磁荷方位的变化 $\partial F/\partial\theta$ 和 $\partial F/\partial r$ 以及回路相对磁荷的特征距离和特征方位相对与电流元长度参量 dl 的变化 $\partial\theta/\partial l$ 和 $\partial r/\partial l$。其中 $\partial\theta/\partial l$ 和 $\partial r/\partial l$ 可通过图 4-4-6 中的几何关系：$rd\theta = \sin\theta dl$，$dr = -\cos\theta dl$ 得到：

$$\frac{\partial\theta}{\partial l} = \frac{\sin\theta}{r} \qquad \frac{\partial r}{\partial l} = -\cos\theta \tag{4-4-5}$$

通过具体测量确定 $\partial F/\partial\theta$ 和 $\partial F/\partial r$ 的表达式需要设计一个能够计算或测量这两个参量的特殊载流回路，即一个能够同时凸现作用力对距离 r 和方位角 θ 两个因素影响的载流回路。

读者也许会感到困惑，导线的形状可以千变万化，什么样的回路能够满足这样的要求。事实上，不同曲线的总体特性虽然不同，但它们在某些局部的微元可以有完全相同的特性(见图 4-4-8)，这意味着这些局部的微分表达式并完全依赖于函数的整体表达式，更是取决于曲线局部片段的取向。或者说，可以从不同的函数得到相同片段的微分。以长直导线为例，其两端的无限延伸导致其对磁极的作用力表达式(4-4-3)掩盖了其中各电流元方位角 θ 对作用力 F 的关联，因而无法通过改变方位角 θ 来测量 $\partial F/\partial\theta$。为此，毕奥与萨伐尔将长直载流导线折成顶角为 2θ 的“V”型，而将试探磁极放置在“V”

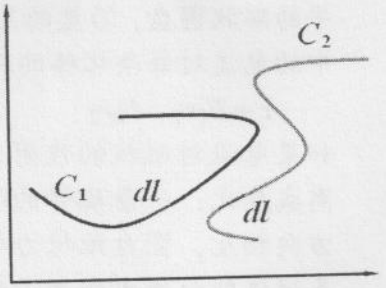


图 4-4-7 不同曲线上的相同微元

型的顶角平分线上距离顶点为 r 的 P 点(见图 4-4-7)。显然，图 4-4-6 中电流元 $Id\mathbf{l}$ 与空间点 P 所构成的位型与图 4-4-7 中顶点一边电流元与位置 P 的数学描述完全相同，可以测量在不同的 r 和不同的 θ 时 P 点的磁荷受到整个载流回路的作用力大小，并由此得到相应的参量。测量结果表明，该作用力随 r 和 θ 的变化可表述为

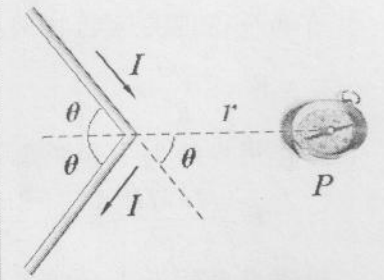


图 4-4-8 毕奥与萨伐尔实验

$$F = k\frac{I}{r}\tan\frac{\theta}{2} \tag{4-4-6}$$

其中比例系数 k 与电流 I 的单位和探测磁荷的强度有关。显然作用力在半顶角 $\theta = 90°$(对应于长直导线)时最大，而在 $\theta = 0°$时(对应于逆向电流互相抵消)消失。分别对(4-4-6)式计算关于 r 和 θ 的偏导数可得

$$\frac{\partial F}{\partial \theta} = k\frac{I}{r}\frac{1}{2\cos^2\theta/2} \qquad \frac{\partial F}{\partial r} = -k\frac{I}{r^2}\tan\frac{\theta}{2} \tag{4-4-7}$$

将(4-4-5)式和(4-4-7)式都代入(4-4-4)式得到

$$dF = k\frac{Idl}{r^2}\sin\theta \tag{4-4-8}$$

上述作用力与距离的平方反比关系可以写成矢量形式[2]

$$d\mathbf{F} = k\frac{Id\mathbf{l}\times\mathbf{r}}{r^3} \tag{4-4-9}$$

其中 $d\mathbf{l}$ 的方向为导线元中的电流方向，而 $\mathbf{r}$ 代表自电流元到位置 P 的位置矢量。力的方向与磁极类型有关。

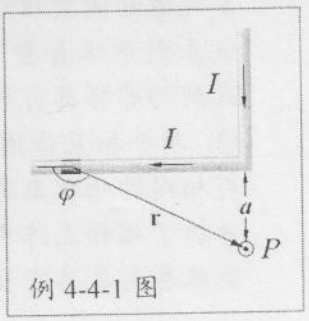
例 4-4-1 将一根载有电流 I 的长直导线折成直角 V-型(见本例图)，计算距离折角正下方 a 处的 P 点处的一个单位磁荷所受到的磁作用力。

解： 由于竖直电流矢量与电流至 P 点的位置矢量共线，其对 P 点处磁荷的作用力为零，因而只要考虑电流水平分量的贡献

$$dF = k\frac{Idx}{r^2}\sin\varphi = -k\frac{I}{a}\sin\varphi\, d\varphi$$

其中 $x = a\cot\varphi$，因所有电流元 Idx 对磁荷的作用力方向都是由纸面向外，总的作用力可以通过对上式积分得到：

$$F = -k\frac{I}{a}\int_{\pi}^{\pi/2} d\varphi\sin\varphi = k\frac{I}{a}$$



例 4-4-1 图

实际上可以不必计算上述积分，直接由式(4-4-2)得到上述结果。◀

例 4-4-2 计算一单位磁荷在半径为 R 的载流线圈轴线上 P 点所受到的磁作用力。

解： 在线圈上选取一段长为 dl 的弧元，它对 P 点处磁荷的作用力方向如图所示。整个线圈对磁单极的作用力的水平分量相互抵消，只有竖直分量，因此

---

[2] 由于载流长直导线对磁极的作用力是横向的(力垂直于由磁极与电流构成的平面)，一个简单而合理的猜测是其中的每个电流元对磁极的作用力也都是横向力。

$$\mathbf{F}=\hat{\mathbf{z}}\oint dF_z=\hat{\mathbf{z}}\oint k\frac{IR^2d\varphi}{\sqrt{R^2+z^2}^3}=\frac{2\pi kIR^2}{\sqrt{R^2+z^2}^3}\hat{\mathbf{z}}$$

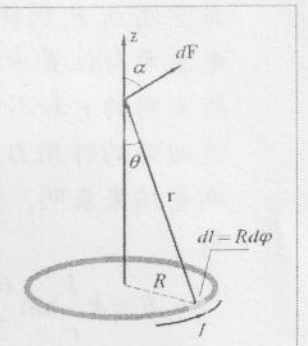
有两个值得关注的特例：$R\gg z$ 和 $R\ll z$，前一个例子中，

$$\mathbf{F}=\frac{2\pi kI}{R}\hat{\mathbf{z}}$$

磁作用力的大小和距离成反比，而后一种情形中

$$\mathbf{F}=\frac{2\pi kIR^2}{z^3}\hat{\mathbf{z}}=k\frac{2\pi}{z^3}\boldsymbol{\mu}$$

磁作用力的大小和距离的三次幂成反比，这和电偶极子与远处点电荷的作用力有相同的距离函数关系。事实上，可以求出单位磁荷在远离载流线圈的任意 P 点($R\ll r$)所受到的磁作用力与处于原点的磁偶极子 $\boldsymbol{\mu}=\mu\hat{\mathbf{z}}$ 对 P 点的磁极的作用力相同



例 4-4-2 图

$$\mathbf{F}=k\frac{3(\boldsymbol{\mu}\cdot\hat{\mathbf{r}})\hat{\mathbf{r}}-\boldsymbol{\mu}}{r^3}\tag{4-4-15}$$

因此电流环通常被称为磁偶极子或磁偶矩。由于不存在磁单极，磁偶矩(电流环)事实上就是磁学中最基本的几何单元，读者后面会发现许多复杂的实际例子中都需要通过电流环的磁学特性来理解。◀

安培定律

仔细考察奥斯特实验可以发现电流磁效应所隐含的两个极不寻常的特点：

(1) 奥斯特实验中导线和磁针开始都处于同一平面内(见图 4-4-8)，即所有的宏观物理内容相对于该平面都是左右对称的，而在导线中馈入电流(处于同一平面内)后，磁针偏离原来磁针与导线所决定的平面使系统失去了实验设置原来所具有的左右对称性，左右对称的实验条件导致左右不对称的实验结果意味着作为自然科学基本规律的因果对称性遭到破坏[3]。而要确保因果对称性成立则意味着整个装置的初始状态并非如我们眼睛所看到的那样具有相对于装置平面的左右对称性。

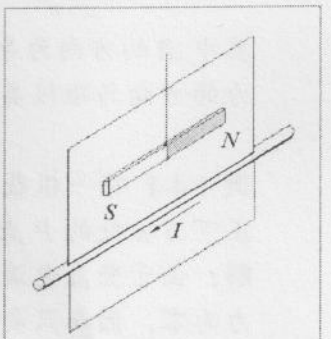


图 4-4-9

(2) 两个相互作用的物体应该具有相同的作用荷，这是作用力和反作用力必须具有相同的物理本原所要求的。法国物理学家安培[4]首先从作用和反作用力的角度分析了磁相互作用。由于电流只是运动的电荷，电荷静止时并不具有磁效应，这意味着与之具有相互作用的磁体中存在着电流。安培由此意识到发生磁相互作用

[3] 奥地利物理学家、哲学家马赫(厄恩斯特·马赫 1838-1916)首先意识到了这一特点。马赫提出的知识是感觉体验的复合这一概念对现代科学与哲学产生过巨大的影响。

[4] 安培 (Andre Marie Ampere, 1775 –1836)，曾从事概率论和积分偏微分方程的研究，1815 年转而研究化学，并多次独立地做出一些重复的发现。如他几乎与 H. 戴维同时认识到氯和碘是元素，比阿伏伽德罗晚 3 年独立地总结出阿伏伽德罗定律。他还曾试图依据元素的性质对元素进行分类。1820 年秋天开始从事电动力学的研究。1827 年，安培将他的研究总结在《电动力学现象的数学理论》一书中，对以后电磁学的发展产生了深远的影响。

的“荷”可能就是电流本身，或者说，电流和电流之间也应该存在相互作用[5]。

1820 年年底，安培首先从实验上证实了两载流直导线之间确实存在这种磁相互作用力，并得到了两条长度都为 L 的平行载流直导线之间的相互作用力的大小与导线间距 r 的关系：

$$F = k\frac{I_1 I_2 L}{r} \tag{4-4-10}$$

其中 I_1 和 I_2 为两条导线中的电流强度，比例常数 k 与单位制的选取有关[6]。在国际单位制中

$$k = \frac{\mu_o}{4\pi} = 10^{-7}(\mathrm{N}\cdot\mathrm{A}^{-2}) \tag{4-4-11}$$

其中 μ_o 称为**真空磁导率**。当电流方向相同时导线之间相互吸引，当电流方向相反时导线之间相互排斥。即该作用力为横向力，它不仅与受力电流的方向垂直，也与施力电流方向垂直。

第一章中曾提到，在国际单位制中，电流强度单位“安培”也被定义为基本物理量单位，其定义就是以(4-4-10)式为依据的：设两条相距为 $r = 1$ 米，长度为 1 米的平行载流导线之间的相互作用力的大小为

$$F = \frac{\mu_o I^2}{2\pi r} = 2\times10^{-7}(\mathrm{N})$$

时，两导线中的电流强度 I 就定义为“1 安培”。

安培对于电流之间存在磁相互作用力的发现使得电流与磁体、磁体与磁体间的相互作用都可以归结为电流与电流之间相互作用。为了说明电流与磁体、磁体与磁体间的相互作用，安培提出了物质磁性的分子电流假说：每个分子都由电流环绕，磁棒的磁性来自其中大量分子电流环(电流环的磁性见例 4-4-2)的方向大致相同而形成了同一方向宏观电流。安培由此认为元电流在物体间的磁相互作用中扮演着磁作用荷的角色。这也就能解释为什么不能将磁性物体的两个磁极分开，以及奥斯特实验中磁针偏转出图 4-4-8 所示的平面导致系统相对于该平面左右对称破坏的原因：磁体中的元电流分布相对于平面并不是左右对称的。

1821~1824 年，安培通过他观察物理现象的出色洞察力以及细致精巧的实验设计，并结合数学分析方法得到了两个电流元之间作用力的数学描述。

两个元电流间的相互作用所涉及的只有三个物理量：电流元 $I_1\mathrm{d}\boldsymbol{\ell}_1$，$I_2\mathrm{d}\boldsymbol{\ell}_2$ 和这两个电流元间的位移矢量 $\mathbf{r}_{12}$：

$$d\mathbf{F}_{12} = d\mathbf{F}_{12}(I_1\mathrm{d}\boldsymbol{\ell}_1, I_2\mathrm{d}\boldsymbol{\ell}_2, \mathbf{r}_{12}) \tag{4-4-12}$$

其中包含着 $\mathrm{d}\boldsymbol{\ell}_1$，$\mathrm{d}\boldsymbol{\ell}_2$ 和 $\mathbf{r}_{12}$ 三个矢量之间的夹角 θ_1 和 θ_2。对比式(4-4-12)和两条载流长直导线单位长度间的相互作用力的公式(4-4-10)关于电流 I_1 与 I_2 的量纲可知，$d\mathbf{F}_{12}$ 只能是电流元 $I_1\mathrm{d}\boldsymbol{\ell}_1$ 与 $I_2\mathrm{d}\boldsymbol{\ell}_2$ 的线性函数，则作用力 $d\mathbf{F}_{12}$ 的大小系可以表示为

$$dF_{12} = I_1 dl_1 I_2 dl_2 f(r_{12}) g(\theta_1, \theta_2)$$

为确定 dF_{12} 对 r_{12} 的函数关系，可以从描述电流元 $I_1\mathrm{d}\boldsymbol{\ell}_1$ 对一个相距 r_{12} 的磁极作

---

[5] 安培在形成这个想法之前曾经做了许多相关的实验，如圆电流对磁针的作用，平行载流直导线间的作用，圆电流之间的作用等实验。这些实验不仅帮助他形成了磁相互作用的本质来源于电流，而且帮助他对这些力和电流之间的距离和方位之间的关系都有了正确的判断。

[6] 读者可以按照引力定律和库仑定律中对比例系数的讨论，对常数 k 进行一些类似的分析。

用力的毕奥-萨伐尔公式得知，电流元 $I_1\mathrm{d}\ell_1$ 对磁极的作用大小应该和 $I_1\mathrm{d}\ell_1$ 与另一个相距 r_{12} 的电流元 $I_2\mathrm{d}\ell_2$ 之间的作用力的大小成正比，也就是说 dF_{12} 应该和这两个电流元之间的距离相距 r_{12} 的平方成反比。事实上，安培从他自己得出的长直载流导线的磁作用力表达式中就已经意识到这个平方反比关系(见思考题 4-4-2)。为了从实验上精确地确认这一关系，安培将三个形状相似，线度比例为 $1:k:k^2$ 的线圈 A、B、C 按 $1:k$ 的间距前后放置(见图 4-4-9)。其中线圈 A 与 C 构成串联回路并固定在一个支架上，而线圈 B 可以活动。当回路 A、C 和 B 中分别通入电流后，检验线圈 A 与 C 对线圈 B 的作用力(线圈 A 与 C 对线圈 B 的作用力的方向相反)。由于 dF_{12} 与电流元 I_1dl_1 和 I_2dl_2 成正比，如果与距离 r_{12} 平方成反比，则线圈 B 所受作用力不变，实验结果证实了这一点，因此有

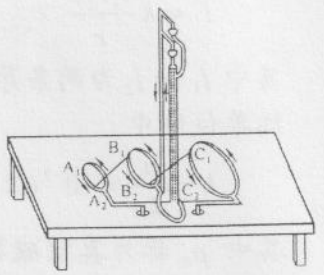


图 4-4-10 安培验证两个电流元之间的磁作用力大小与两者的间距平方成反比的实验设计。

$$dF_{12}=k\frac{I_1d\ell_1I_2d\ell_2}{r_{12}^2}g(\theta_1,\theta_2) \tag{4-4-13}$$

为了确定相互作用对这些矢量方向的依赖关系 $g(\theta_1,\theta_2)$，安培由所有相关实验中电流受到的磁相互作用都与电流的方向垂直这一特征，猜测其中的每个电流元受到的作用力也都是横向力。为了证明这一猜测，安培设计了如图 4-4-10 所示的示零实验：将一段圆弧形导体架在水银槽 A 和 B 上，A 和 B 使圆弧形导体构成一电流元。将与圆弧形导体中点固连的绝缘柄 C 的一端固定在支架上，这样，导体圆弧只能沿自身的切线方向运动，而不能作横向位移(垂直与圆弧形导体的方向运动)。安培用该装置检验各种载流线圈能否使圆弧形导体产生运动。结果发现，任何载流线圈都不能使圆弧形导体发生运动，这说明载流线圈作用在电流元上的力一定与电流元垂直。根据两个电流元的作用力的表达式(4-4-15)，作用在电流元 $I_2\mathrm{d}\ell_2$ 上的力的横向特征可表示为

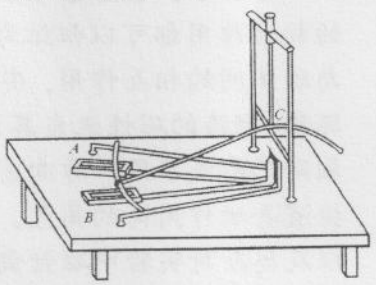


图 4-4-11 安培验证两个电流元之间的磁作用力方向与两个电流元方向都垂直的实验设计。

$$d\mathbf{l}_2\cdot\oint d\mathbf{F}_{12}=0$$

其中的回路积分是产生作用在 $I_2d\ell_2$ 上的力的一个载流回路。而其中的 $d\mathbf{F}_{12}$ 为回路中的电流元 $I_1d\ell_1$ 对 $I_2d\ell_2$ 的作用。这意味着

$$\oint d\mathbf{F}_{12}=I_2d\mathbf{l}_2\times\oint d\mathbf{B}_{12} \tag{4-4-14}$$

其中 $d\mathbf{B}_{12}$ 仅与源电流元 $I_1d\mathbf{l}_1$ 及其延伸到 $I_2d\mathbf{l}_2$ 的位置矢量 $\mathbf{r}_{12}$ 有关，因而可以理解为源电流元 $I_1d\mathbf{l}_1$ 作用在单位电流元上的磁作用力，它和作用在单位磁极上的毕奥-萨伐特磁作用力应该成正比：

$$d\mathbf{B}_{12}=k\frac{I_2d\mathbf{l}\times\hat{\mathbf{r}}_{12}}{r_{12}^2}$$

代入式(4-4-14)可得到：

$$d\mathbf{F}_{12}=k\frac{I_1I_2d\mathbf{l}_2\times(d\mathbf{l}_1\times\hat{\mathbf{r}}_{12})}{r_{12}^2} \tag{4-4-15}$$

若采用国际单位制，常数 $k=\mu_0/4\pi$。上述表达式称为**安培定律**[7]。安培定律应理解为和牛顿万有引力定律以及库仑定律一样的描述基本相互作用荷之间作用力的基本物理定律。

4-4-3 运动电荷之间的磁相互作用

前面曾经提到，牛顿的引力定律及库仑定律都含有超距作用的特性，因此它们都只适用于描述作用物体都处于静止状态时物体之间的作用力。但读者也许会意识到这样一个问题：描述两个稳恒电流元之间的磁作用力的安培定律中的电流元 $Id\mathbf{l}$ 描述的是在 dt 时间间隔中，数量为 dQ 的电荷发生了 $d\mathbf{l}$ 的位移，它可以等价地表示为

$$Id\mathbf{l}=\frac{dQ}{dt}d\mathbf{l}=dQ\frac{d\mathbf{l}}{dt}=q\mathbf{v}$$

这样，安培定律就在形式上表示为两个速度分别为 $\mathbf{v}_1$ 和 $\mathbf{v}_2$ 的匀速运动电荷 q_1 和 q_2 在某一瞬时间距为 r_{12} 的某一瞬时电荷 q_1 和 q_2 的磁相互作用，其中电荷 q_2 作用在电荷 q_1 上的磁力为

$$\mathbf{F}_{21}=\frac{\mu_o}{4\pi}\frac{q_1\mathbf{v}_1\times(q_2\mathbf{v}_2\times\hat{\mathbf{r}}_{21})}{r_{21}^2} \tag{4-4-16}$$

这种将运动物体之间的非接触性相互作用也表示为含有超距作用特征的描述是否成立？更进一步的问题是，处于一般运动形式的电荷之间的相互作用是否也能用上述公式来描述？

显然，如果我们不认同超距作用的概念，安培定律中两个稳恒电流元之间的相互作用就不是该时刻处在相应位置的两个运动电荷之间的安培力，因为在 t 时刻 $\mathbf{r}_1$ 位置上电荷的运动信息传递到 $\mathbf{r}_2$ 位置的时刻为 $t+\Delta t$，这时原来 $\mathbf{r}_1$ 位置上的电荷已经运动到了位置 $\mathbf{r}_1+\Delta\mathbf{r}$，只是对稳恒电流而言，$\mathbf{r}_1$ 位置上的电荷运动状态是不随时间变化的，因而实际测量到的结果相同[8]。因此，安培定律能用于，也只能用于描述稳恒电流之间的磁相互作用。实际的运动点电荷所形成的空间的电流分布只是一个瞬态电流脉冲，因此原则上不能用安培定律来计算运动电荷之间的相互作用。

但如果考虑这样的事实：当我们用牛顿引力定律计算太阳和行星之间的作用力以及它们的运动演化时，行星始终处于运动状态，而我们用描述静态物体之间的作用力公式可以很好地复原开普勒定律，并非常精确地预言行星的运动情况。这表明，在物体运动速度很低的情况，具有超距作用特性的作用力公式仍然是适用的。事实上我们可以设想这样的情形，在物体运动速度非常低的时候，这些超距作用公式是成立的，而对于以任意速度运动的电荷之间的磁相互作用则上可以通过对(4-4-16)式作洛仑兹变换给出

---

[7] 实际上最初在推导电流元之间的相互作用时，安培曾经希望电流元之间的磁相互作用能满足牛顿第三定律，这实际反映了安培对超距作用观点的信念。但安培由此得到的原始公式中存在有明显的反例。参阅陈秉乾等所著的《电磁学专题研究》1-3 节，高等教育出版社 2001。

[8] 实际测量到的是整个回路中电流的贡献。

$$\mathbf{F}_{21}=\frac{\mu_o}{4\pi}\frac{q_1\mathbf{v}_1\times(q_2\mathbf{v}_2\times\hat{\mathbf{r}}_{21})}{\gamma_v^2(1-\beta_v^2\sin^2\theta)^{3/2}r_{21}^3} \tag{4-4-17}$$

其中包含了因电荷运动的信息的有限传播速度所导致的推迟效应。在电荷运动速度很低时，$\beta_v\approx 0$，$\gamma_v\approx 1$，(4-4-17)式就自动回到(4-4-16)式。

事实上，对于典型的行星运动速度 $v \ll c$，以及导线中典型的电子平均运动速度 $v \ll c$ 而言，用引力公式和安培力公式计算所导致的误差都可以忽略。

思考题

4-4-1 说明如何设定(4-4-1)式**磁库仑定律**中的比例系数 k_m 的大小和量纲。比较库仑定律在静电学中的地位与安培定律在静磁学中的地位。

4-4-2 由于物体中的磁极不能分离，物体的磁化和极化可能表现出什么样的可观测效应。

4-4-3 如何从安培的长直载流导线的磁作用力表达式推断出两个电流元之间的作用力满足平方反比关系？

4-4-4 将一根柔软的螺旋形弹簧挂起来，使它的下端和盛在杯里的水银刚好接触形成串联回路，再把它们接到直流电源上，弹簧将发生什么现象？

4-4-5 说明两个电流元之间的磁作用力为横向力在数学上可以直接表示为 $d\mathbf{F}_{21}$ 中三个矢量 $I_1d\boldsymbol{\ell}_1$，$I_2d\boldsymbol{\ell}_2$ 和 $\mathbf{r}_{12}$ 之间的关系为矢量积的关系。

4-4-6 在安培定律的表达式中，若 $r_{12}\to 0$，则 $\mathrm{d}\mathbf{F}_{12}\to\infty$，这一结论是否正确？怎样解释？

4-4-7 设想用一电流元作为检测磁场的工具。若沿某一方向，给定的电流元 I_odl 放在任一空间点都不受力作用，能否由此断定该空间不存在磁场？

4-4-8 可以用一个很小的载流线圈作为定义磁场的检测工具(探测线圈)。怎样用探测线圈来定义 **B** 的大小和方向？

4-4-9 将一电流元依次放置在无限长的载流直导线附近的两点 A 和 B，如果 A 点和 B 点到导线的距离相等，电流元所受到的磁力大小是否一定相等？

习题

4-4-1 求解两个相距为 r 的磁偶矩 $\mathbf{p}_1$ 和 $\mathbf{p}_2$ 之间的相互作用能和相互作用力。

4-4-2 利用磁单极在载流长直导线附近受到的作用力公式(4-4-2)，求在载流长直导线附近的一个小磁针的振荡周期和作用力的关系。

4-4-3 求图(a), (b), (c), (d), (e)和(f)中，圆弧中心 O 处的磁感强度(图中虚线表示恩向无穷远的直导线)。(4-1)

4-4-4 载流 I 的正方形线圈的边长为 a，求线圈轴线上距其中心 O 为 h 处的磁感强度表达式。当 $a=1.0$ cm，$I=5.0$A，$h=0$ 和 20 cm 时的磁感强度。(4-2)

4-4-5 两条无限长平行直导线相距 a，分别载有电流 I_1 和 I_2，当两电流同向及反向时，求距离两条导线分别为 r_1 和 r_2 的空间 P 点的磁感强度？(4-3)

4-4-6 边长为 a 的等边三角形载流回路，电流为 I。求过三角形重心且与三角形平面垂直的轴线上距重心为 r 处的磁感强度。(4-4)

4-5 相互作用力场

本节讨论如何理解和描述两个物体之间力的作用问题。事实上，对这个问题还没有足够另人满意的结论。但我们希望至少可以在概念上克服一些已经意识到的明显的漏洞，并使我们的认识有所深化。

从前几节的讨论知道，引力和电、磁作用力在数学上都被描述为物体之间跨越空间的非接触瞬间相互作用

$$\mathbf{F}_{\mathrm{G}} = -G\frac{m_1 m_2}{r^2}\hat{\mathbf{r}}$$

$$\mathbf{F}_{\mathrm{C}} = \frac{q_1 q_2}{4\pi\varepsilon_o r^2}\hat{\mathbf{r}}$$

$$d\mathbf{F}_{12} = \frac{\mu_o}{4\pi}\frac{I_1 I_2 d\mathbf{l}_2 \times (d\mathbf{l}_1 \times \hat{\mathbf{r}}_{12})}{r_{12}^2} \quad \text{或} \quad \mathbf{F}_{21} = \frac{\mu_o}{4\pi}\frac{(q_1\mathbf{v}_1)\times[(q_2\mathbf{v}_2)\times\hat{\mathbf{r}}_{21}]}{r_{21}^2}$$

其中引力和库伦力都表示为静止作用荷之间的作用力，但我们在实际应用中通常并不顾忌两个载荷的物体是否静止(至少彼此是否存在相对运动)，例如引力定律所基于的开普勒公式描述的就是行星的绕日轨道运动。而在磁作用力公式中虽然计入了电荷的运动参量(电流或速度)，但讨论的也只是稳恒状态，计算作用力时没有必要讨论力的传递及其特征。

但设想两个相互作用之中的物体中的任何一个突然运动状态发生了变化，另一个物体无论相距多远，按照上面的作用力公式必须瞬间感知到，并即时调节自身感受到的作用力以及施加于对方的相互作用。也就是说，物体运动行为发生变化的信息在两个物体之间的传递是瞬间完成的。在 3-8 节中曾提到这种被称为**超距作用**的概念与牛顿第三定律相抵触。这一点在两个电流元之间的磁相互作用中表现的特别明显。例如，就图 4-5-1 所示的两个特定取向的电流元 $I_1 d\mathbf{l}_1$ 和 $I_2 d\mathbf{l}_2$ 而言，$I_1 d\mathbf{l}_1$ 对 $I_2 d\mathbf{l}_2$ 的安培力和 $I_2 d\mathbf{l}_2$ 对 $I_1 d\mathbf{l}_1$ 的安培力分别为

$$d\mathbf{F}_{21} = \frac{\mu_o}{4\pi r^2} I_1 I_2 dl_1 dl_2 \quad \text{和} \quad d\mathbf{F}_{12} = 0$$

因而有：$d\mathbf{F}_{12} \neq d\mathbf{F}_{21}$。事实上牛顿一开始就意识到引入超距作用中所含有的难以逾越的概念困境，但这位数学和物理学的巨人一直无法找到解脱的途径。而对这一问题的突破最终是由没有受过任何数学教育，但充满想象力的英国人法拉第[1]实现的。

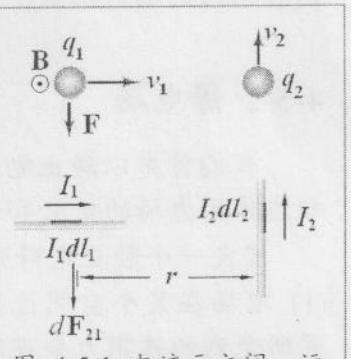


图 4-5-1 电流元之间，运动电荷之间的作用和反作用力不满足牛顿第三定律

[1] 法拉第(Michael Faraday 1791-1867)自学成才的著名英国物理学家和化学家，发电机和电动机的发明者。出生于萨里郡纽因顿一个贫苦铁匠家庭。仅上过几年小学。13 岁时便在一家印书工场当学徒。这份工作使他有机会读到许多科学书籍。他在工作之余，自学化学和电学，并动手做简单的实验，验证书上的内容。1813 年由戴维举荐到皇家研究所任实验室助手。这成了法拉第一生的转折点，他在电磁学方面的研究成就使他成为和麦克斯韦齐名的两位电磁学宗师。1824 年 1 月当选皇家学会会员，1825 年 2 月任皇家研究所实验室主任，1846 年荣获伦福德奖章和皇家勋章。

4-5-1 相互作用力场

为了摆脱用超距作用的概念描述各种非接触相互作用力所引起的困境，法拉第针对电荷之间的相互作用提出了“**电场**”的概念：任何电荷 Q 永远都不会以裸露的形式出现在真实的物理世界中，其周围总是环绕着某种物质—电场(见图4-5-2)。但电场在电荷周围的分布并不是刚性的，它相对于电荷可以发生位移或者形变。例如，当电荷 Q 的状态突然发生了变化(如一个静止点电荷突然得到一个速度)，电荷 Q 的空间电场分布的变化会按相距该电荷的远近有不同的时间滞后。而进入该电场空间的其它带电荷 q 的物质感受到电荷 Q 对它所携载电荷 q 的电相互作用不是由 Q 本身直接作用于 q，而是 Q 通过其在 q 所在位置的电场间接作用的。这样，抽象的作用力场就在概念上摆脱了超距作用带来的问题：如果 Q 发生突然的运动，它在 q 处的电场的相应变动会有一定的延迟。则 q 感受到的作用力变化也就有相应的延迟。

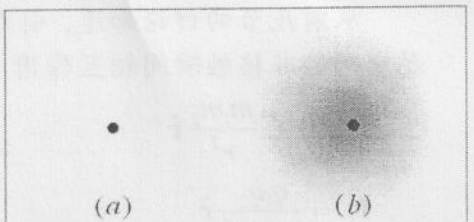


图 4-5-2 (*a*) 与其它电荷直接作用(超距作用)的裸电荷，(*b*)只能与其它电荷间接发生相互作用的电荷与电场的共同体

由于相互作用依赖于物质，法拉第强调，电场并不是一种抽象，而是由物质构成的实际物体。虽然当时并不清楚这是一种什么物质，但“力场”的物质性概念确实可以帮助我们解释前面提到的两个电流元之间相互作用力不同的问题。

显然，这样定义的作用力场的概念对任何相互作用荷都适用。但“力场”的性质(即作用力的方式)与作用荷的特性有关，如描述电荷和引力荷的数学量都是标量，相应的作用力都是中心力，而描述磁作用荷—电流元的数学量是矢量，电流元之间的相互作用是横向的。本节主要以静电场和稳恒磁场为例，讨论这两类不同作用荷的场的形式，以及依赖力场作为中介的“作用荷”之间的相互作用特征。

4-5-2 静电场

我们首先以静止电荷以及静止电荷之间的中心力为例，讨论标量作用荷以及相应作用力场的定义和特性。

定义一个静止电荷分布系统所产生的空间电场分布需要考虑到两点

(1) 电场在某个空间位置是否存在及相关特性只能通过用一个其它电荷在该位置所受到的作用力来观测，即所需要定义的电场的观测属性应该是该试探电荷受到的作用力。

(2) 一个静止电荷分布系统所产生的空间电场分布与其附近是否存在试探电荷应该是无关的，但试探电荷在空间某个位置 $\mathbf{r}$ 感受到的电作用力与试探电荷的大小有关。定义电场就必须扣除试探电荷的特性。为此的电场定义为一个单位试探电荷 q_o 在该位置所受到的库仑力

$$\mathbf{E}(\mathbf{r}) \equiv \frac{\mathbf{F}(\mathbf{r})}{q_o} \tag{4-5-1}$$

但需要注意的是，对于实际用于测量的探测电荷 q_o 本身也所会产生电场，因而会改变产生 q_o 所处场点的电场的电荷系统原来的电荷分布，并反过来影响到原

来场点的电场测量精度。为了减小甚至避免这种效应的影响，应要求探测电荷的电量 q_o 应尽可能小[2]。在国际单位制中，电场的单位为 $\mathrm{N\cdot C^{-1}}$。为了让读者对电场的大小有一个数量级的概念，表 4-5-1 列出了一些典型场合的电场强度。

表 4-5-1　典型的电场 (N/C)

太阳系空间	10^{-3}~10^{-2}
晴空时接近地面大气层	10^{2}
雷暴中	10^{3}
干燥空气的电击穿	3×10^{6}
原子的电子运道轨道以内	10^{9}
强激光发出的电磁辐射	10^{12}
距离铀核中心一个铀核直径	5×10^{20}

原则上通过将试探电荷放置到空间各处并测量出该电荷在各点受到的库仑力，就可以描绘出整个空间中的电场分布。而知道了空间各处的电场分布 $\mathbf{E}$，点电荷 q 在任一位置所受到的静电力可以表示为

$$\mathbf{F}=q\mathbf{E} \tag{4-5-2}$$

上式常用于计算运动电荷在电场受到的作用力，这意味着电场对电荷的作用力与电荷的运动状态无关。由于静电场是由源电荷产生的。源电荷产生的电场对运动的检测电荷的作用力仍然表现为 $\mathbf{F}_{\mathrm{E}}=q_o\mathbf{E}$ 只能看作是实验事实，它并不包括在库仑定律之中。

电场的计算—叠加原理

原则上，所有的静电场都起源于点电荷在空间的分布。设想空间中有 n 个点电荷 $Q_1, Q_2, \ldots, Q_n$ 分别处在位置 $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n$。则探测电荷 q 在空间 $\mathbf{r}$ 处所感受到这些点电荷的作用力是所有这些点电荷对它的作用力 $\mathbf{F}_i(\mathbf{r})$的矢量和

$$\mathbf{F}(\mathbf{r})=\sum_i\mathbf{F}_i(\mathbf{r})=\sum_i\frac{qQ_i(\mathbf{r}-\mathbf{r}_i)}{4\pi\varepsilon_o|\mathbf{r}-\mathbf{r}_i|^3}=q\sum_i\frac{Q_i(\mathbf{r}-\mathbf{r}_i)}{4\pi\varepsilon_o|\mathbf{r}-\mathbf{r}_i|^3}$$

则由电场的定义式可以得到探测电荷 q 所在场点 $\mathbf{r}$ 处的静电场为

$$\mathbf{E}(\mathbf{r})\equiv\frac{\mathbf{F}(\mathbf{r})}{q}=\sum_i\frac{\mathbf{F}_i(\mathbf{r})}{q}=\sum_i\frac{Q_i(\mathbf{r}-\mathbf{r}_i)}{4\pi\varepsilon_o|\mathbf{r}-\mathbf{r}_i|^3} \tag{4-5-3}$$

其中

$$\mathbf{E}_i(\mathbf{r})\equiv\frac{\mathbf{F}_i(\mathbf{r})}{q}=\frac{Q_i(\mathbf{r}-\mathbf{r}_i)}{4\pi\varepsilon_o|\mathbf{r}-\mathbf{r}_i|^3} \tag{4-5-4}$$

为点电荷 Q_i 在位置 $\mathbf{r}$ 处产生的电场。若 Q_i 为正电荷，电场方向为由电荷指向场点位置，若 Q_i 为负电荷，各点电场方向指向电荷所在位置。这些点电荷构成的电荷系统在位置 $\mathbf{r}$ 处产生的电场就是各点电荷在位置 $\mathbf{r}$ 处产生的电场的矢量和

$$\mathbf{E}(\mathbf{r})=\sum_i\mathbf{E}_i(\mathbf{r}) \tag{4-5-5}$$

例 4-5-1 求电偶极子 $\mathbf{p}=q\mathbf{d}$ 在远处产生的电场。

解：由 4-3-3 节中得到的关于电偶极子 $\mathbf{p}$ 与远处点电荷 Q 的作用力表达式(4-3-9)：

$$\mathbf{F}(\mathbf{r})\approx\frac{3(\mathbf{p}\cdot\hat{\mathbf{r}})\hat{\mathbf{r}}-\mathbf{p}}{4\pi\varepsilon_o r^3}Q$$

以及电场的定义可知，电偶极子 $\mathbf{p}$ 在远场点处的电场为

[2] 由于实际上不存在无穷小的电荷，因而探测电荷的标准是对源电荷分布以及它所产生的电场的变化小于对源电荷分布及电场本身的测量误差。

$$\mathbf{E}(\mathbf{r})=\frac{3(\mathbf{p}\cdot\hat{\mathbf{r}})\hat{\mathbf{r}}-\mathbf{p}}{4\pi\varepsilon_o r^3}$$

电偶极子产生的电场与距离的三次方成反比，其衰减比点电荷的电场迅速。◀

所有的静电场原则上都起源于点电荷在空间的分布，但通常对于宏观带电系统的问题时将电荷看成连续分布更容易处理，而且计算结果也能达到很高的精度。将(4-5-3)中的点电荷替代为体积元 ΔV_i 所带的电荷 $Q_i(\mathbf{r}_i)=\rho(\mathbf{r}_i)\Delta V_i$，其中 $\rho(\mathbf{r}_i)$ 为该体积元所在位置的电荷密度。其在场点 $\mathbf{r}$ 产生的电场为

$$\Delta\mathbf{E}_i(\mathbf{r})=\frac{(\mathbf{r}-\mathbf{r}_i)}{4\pi\varepsilon_o|\mathbf{r}-\mathbf{r}_i|^3}Q_i=\frac{(\mathbf{r}-\mathbf{r}_i)}{4\pi\varepsilon_o|\mathbf{r}-\mathbf{r}_i|^3}\rho(\mathbf{r}_i)\Delta V_i$$

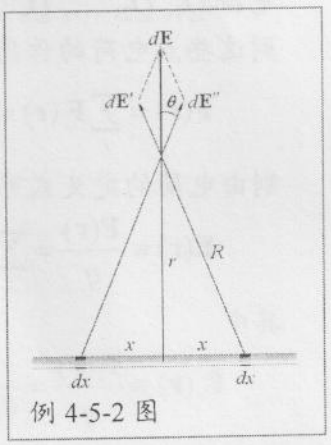
整个电荷系统在场点 $\mathbf{r}$ 产生的电场为上式对电荷系统体积的积分

$$\mathbf{E}=\sum_i\Delta\mathbf{E}_i=\frac{1}{4\pi\varepsilon_o}\int_V dV\rho(\mathbf{r}')\frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|^3} \tag{4-5-6}$$

对于一维或二维的连续电荷分布，电荷微元 $\rho(\mathbf{r})dV$ 分别退化为 $\lambda(\mathbf{r})dl$ 或 $\sigma(\mathbf{r})ds$，其中 $\lambda(\mathbf{r})$ 或 $\sigma(\mathbf{r})$ 分别为线电荷密度或面电荷密度。

例 4-5-2 以长度为 L 的均匀带电细杆所带电荷密度为 λ，计算垂直与细杆，且过细杆中点的平面上的电场分布。

解： 取细杆为 **x**-轴，细杆中心为坐标系原点(见本例图)。因在原点右侧 x 处的电荷元 λdx 在距原点 r 处产生的电场 $d\mathbf{E}'$ 和其对称位置 $-x$ 处的电荷元 λdx 在该处产生的电场 $d\mathbf{E}''$ 的 x 分量恰好抵消。因而由 x 和 $-x$ 处的两个电荷元 λdx 在位置 P 处产生的电场之和只有垂直于细杆的分量



例 4-5-2 图

$$dE_z=\frac{2\lambda}{4\pi\varepsilon_o}\frac{dx}{r^2}\cos\theta$$

由几何关系 $x=r\tan\theta$，将上式中变量 x 用角度 θ 表示

$$\frac{dx}{r^2}=\frac{1}{r^2}\frac{r}{\cos^2\theta}d\theta=\frac{d\theta}{r}$$

则对上式积分可得

$$E_z=\frac{\lambda}{4\pi\varepsilon_o}\int_{-L/2}^{L/2}\frac{dx}{r^2}\cos\theta=\frac{\lambda}{2\pi\varepsilon_o r}\int_0^{\theta_o}\cos\theta d\theta=\frac{\lambda}{2\pi\varepsilon_o r}\sin\theta_o=\frac{\lambda L}{2\pi\varepsilon_o r\sqrt{4r^2+L^2}}$$

下面讨论两个极限特例：$L\gg r$ 和 $L\ll r$，相应的电场强度分别为

$$\mathbf{E}=\frac{\lambda}{2\pi\varepsilon_o r}\hat{\mathbf{z}}\qquad\text{和}\qquad\mathbf{E}=\frac{\lambda l}{4\pi\varepsilon_o r^2}=\frac{Q}{4\pi\varepsilon_o r^2}$$

前一特例中电场与距离成反比是由于无限多电荷在一维尺度中无限均匀分布的结果。而后一特例中电场与距离平方反比则是有限量的电荷均匀分布在有限区域的结果。◀

电场的几何描述—电场线

作用力场是矢量场，但给出直观的图像描述并不容易。虽然表面上似乎可通

过将空间每一点与一个矢量联系起来表示矢量场的概念，其中矢量的方向表示相应场点矢量场的方向，而矢量的长度代表相应场点矢量场的大小。如图 4-5-4 中 a 和 b 分别按这种方法大致画出了两个单位的孤立正电荷和一个单位的孤立负电荷的静电场分布矢量图，而 c 画出的则是上述两个电荷相隔一定的间距构成的系统的静电场的矢量图。容易看出这种矢量场的几何描述既不清晰，也不方便。如此简单的电荷分布就已经使得页面上非常混乱地拥挤了太多的不同长短的矢量。

法拉第通过与流场的类比来描述作用力场的图像。该方法后来经过麦克斯韦改进，最终成为了描述各种矢量场的普适几何方法。这种方法的具体要点是

(1) 在空间中每一个位置都按照矢量场在该点的方向描出一条无限短的线元，能够互相连接的这些线元连成的曲线称为矢量场的**场线**。显然，场线在每一个位置的切线方向代表了该位置的矢量场方向，因此画出一个矢量场所有的场线，该矢量场在所有位置的方向就完全设定了。

(2) 为了描述矢量场在各空间点的大小，在矢量场中任一点 $\mathbf{r}$，设想一个与该点场强的方向相垂直的面积元 ΔS，使得通过 ΔS 的电场线条数 ΔN 满足

$$E(r) \propto \frac{\Delta N}{\Delta S} \tag{4-5-7}$$

即矢量场中任一点处的**场线密度**在数值上与该点处场强的大小成正比：场线密集的区域表示场强较强；场线密度低的区域表示场强较弱。

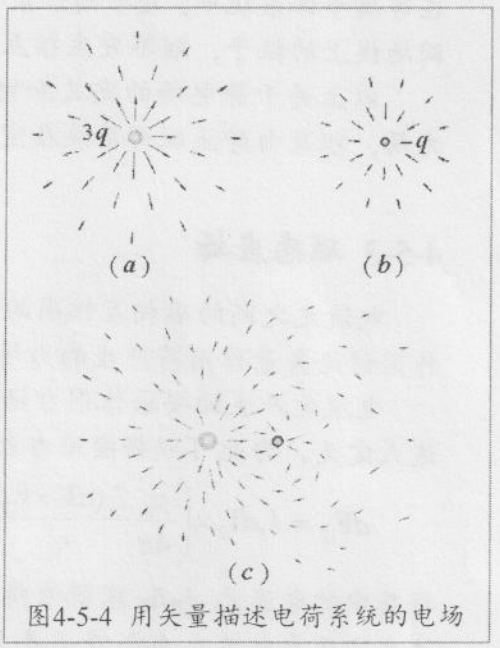


图4-5-4 用矢量描述电荷系统的电场

按照上述规则和电场的定义，静电场的场线具有以下特征：

(1) 由于电场的方向是由正电荷的受力方向定义的，静电场中的任何一条电场线总是起始于正电荷(或无穷远)，终止于负电荷(或无穷远)。静电场线不会在没有电荷的地方中断，也不会回到电场线的起始电荷处而形成闭合线。

(2) 由于除了点电荷所在位置，静电场中的任何一点都只有一个确定的方向，所以任何两条电场线都不可能在空间某一位置相交。

(3) 从正电荷涌出的电场线正比于电荷量，因而如果为某一个电量 Q 设定一个合适的场线的数目 N，则从电荷 q 涌出或没入的电场线数量为 $n=N(q/Q)$。

图 4-5-5 给出了一些简单电荷体系的电场线示意图，其中(*a*)为点电荷，(*b*)为两个相同的点电荷，(*c*)为一个电偶极子，(*d*)为两个点电荷+2Q 和−Q 的复合电荷体。不管系统中的电荷如何分布，紧邻各点电荷 Q_i 的区域 ($|\mathbf{r}-\mathbf{r}_i| \ll \min|\mathbf{r}_j-\mathbf{r}_i|$) 内的电场线分布都是沿径向的均匀分布，即是由区域中心处的点电荷决定的。而远离电荷系统的区域 ($|\mathbf{r}-\mathbf{r}_i| \gg \max|\mathbf{r}_j-\mathbf{r}_i|$) 外的电场线分布也都是沿径向的均

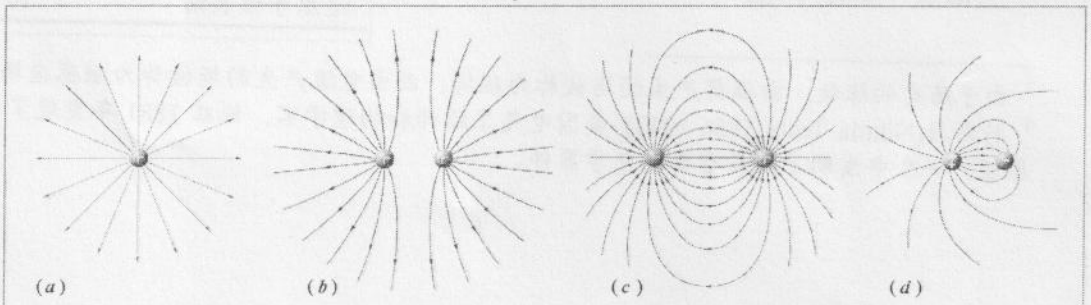


图 4-5-5 一些典型的简单电荷体系的电场线：(*a*) 点电荷，(*b*) 两个相同的点电荷，(*c*)

匀分布，即在远离电荷系统的位置观测，电荷系统相当于一个电量为全部电荷代数和的点电荷产生的电场。如(*d*)中设定出自点电荷+2Q 的电场线数目为 N，则只能有一般数量($\frac{1}{2}N$)的电场线能没入右边的负电荷–Q。剩余的$\frac{1}{2}N$ 的电场线只能走向无穷远，在远离电荷系统的地方电场线的方向趋于径向方向。因而在远离点观测，电荷系统相当于一个电量为 Q 的点电荷。

描述场线的方程是

$$\frac{dx}{E_x}=\frac{dy}{E_y}=\frac{dz}{E_z} \tag{4-5-8}$$

场线的图案能够用来表达某些重要的矢量微积分概念。场线从某一个区域的往外扩散或往内聚敛可以表达散度。场线的涡旋图案可以表达旋度。

虽然场线只是一个数学建构，在某些状况，场线具有实际的物理意义。例如，在等离子体系统中，处于同一条场线的电子或离子会强烈地相互作用；而处于不同场线上的粒子，则不发生相互作用。

以上关于静电场的定义和讨论可以类似地用于定义一个物体系统产生的**引力场**，以及由静止磁荷系统在空间产生的**磁场**(见习题 4-5-5)。

4-5-3 磁感应场

电流元之间的磁相互作用的作用荷是矢量荷，我们以两个电流元之间的相互作用讨论矢量作用荷产生的力场的特性。

电流元产生的安培作用力场可利用两个电流元 $I_1d\boldsymbol{\ell}_1$ 和 $I_2d\boldsymbol{\ell}_2$ 之间的安培力表达式定义。为此可以将安培力表达式(4-4-16)改写为

$$d\mathbf{F}_{12}=I_2d\mathbf{l}_2\times\left(\frac{\mu_o}{4\pi}\frac{I_1(d\mathbf{l}_1\times\hat{\mathbf{r}}_{12})}{r_{12}^2}\right) \tag{4-5-9}$$

将其中的电流元 $I_2d\mathbf{l}_2$ 理解为所在位置的安培力场的探测作用荷，而将括号中表达式定义为电流元 $I_2d\mathbf{l}_2$ 在 $I_1d\mathbf{l}_1$ 所处位置产生的安培作用力场：

$$d\mathbf{B}_{12}=\frac{\mu_o}{4\pi}\frac{I_1d\mathbf{l}_1\times\hat{\mathbf{r}}_{12}}{r_{12}^2} \tag{4-5-10}$$

这个力场由于历史的原因被称为**磁感应强度**[3]。在国际单位制中，其量度单位 $\mathrm{N\cdot A^{-1}\cdot m^{-1}}$ 通常称为**特斯拉**[4]，记为 T。表 4-5-2 列出了一些典型场合的磁感应场强度。但在日常经验中，特斯拉是一个太大的单位，更常用的单位是高斯：1 高斯 $=10^{-4}$ 特斯拉，记为 T。

(4-5-10)式与计算磁荷在电流元附近受到的作用力的毕奥-萨伐尔定律(4-4-9)形式相同，也称为毕奥-萨伐尔定律。(4-5-9)式可以写为

$$d\mathbf{F}_{12}=I_2d\mathbf{l}_2\times d\mathbf{B}_{12} \tag{4-5-11}$$

表 4-5-2 典型的磁感应场 (T)

星际空间	10^{-10}
近地球表面	5×10^{-5}
一般条形磁铁的两极	10^{-2}~10^{-1}
近太阳表面	10^{-2}
最大的稳态磁铁	30
实验室最强电磁脉冲	500~1000
近脉冲星表面	10^{8}
近原子核表面	10^{12}

[3] 由于历史的缘故，由磁荷产生的场被称为磁场，而由电流产生的场被称为磁感应场。

[4] 特斯拉(Nikola Tesla 1856-1943) 美国电气工程师和物理学家，他在 1881 年发现了交流电原理，一生中发明了许多重要的电学器件。

知道了稳恒电流元产生的磁感应场，就可以利用力的叠加原理计算一个稳恒电流(分布)位型在空间各点产生的磁场。

例 4-5-3 计算长度为 L 的直导线段在载流 I 时在其周围产生的磁感应场。

解：按毕奥-萨伐尔定律，载流导线上的电流元 Idl 在导线旁任意一点 P 产生的磁感应场 $d\mathbf{B}$ 的方向同时垂直于导线以及 P 点至导线的垂线 PO，即长直导线周围的磁感应线为在垂直于导线并包含 P 点的平面内构成环绕导线的圆。则直导线段在产生的 P 点磁感应强度为

$$B=\int_a^b dB=\frac{\mu_o I}{4\pi}\int_a^b\frac{\sin\theta}{r^2}dz$$

由例图中的几何关系可知：$z=r\cos(\pi-\theta)=-r\cos\theta$, $r_o=r\sin(\pi-\theta)=r\sin\theta$。消去 r 可得：$z=r_o\cot\theta$，其微分为：

$$dz=\frac{r_o d\theta}{\sin^2\theta}$$

将上面的积分变量 l 换为 θ 后得到

$$B=\frac{\mu_o I}{4\pi}\int_{\theta_1}^{\theta_2}\frac{\sin\theta}{r_o}d\theta=\frac{\mu_o I}{4\pi r_o}(\cos\theta_1-\cos\theta_2)$$

式中 θ_1,θ_2 分别为 θ 角在 a、b 两端的数值。在非常贴近直导线

若导线为无限长，$\theta_1=0$，$\theta_2=\pi$，则

$$B=\frac{\mu_o I}{2\pi r_o}$$

上述磁感应场的表达式可以写成矢量形式

$$\mathbf{B}=\frac{\mu_o}{2\pi}\frac{\mathbf{I}\times\hat{\mathbf{r}}}{r}\qquad(4\text{-}5\text{-}12)$$

自然界中当然不可能遇到真正的无限长直导线。但在非常靠近载流导线($r_o \ll l$)的区域上式成立。这类似于在任何电荷分布的系统中，非常靠近某一点电荷位置的电场几乎就是该点电荷产生的电场。

例 4-5-4 计算半径为 R 的载流线圈在线圈轴线处产生的磁场。

解：在线圈上选取一段长为 dl 的弧元，它在线圈轴线某处产生的磁感应强度为 $d\mathbf{B}$，从对成性的角度考虑，整个载流线圈产生的磁感应场是沿线圈轴线方向 $\mathbf{z}$。

$$dB_z=dB\cos\theta=\frac{\mu_o}{4\pi}\frac{IRd\varphi}{R^2+z^2}\sin\theta=\frac{\mu_o}{4\pi}\frac{IR^2d\varphi}{(R^2+z^2)^{3/2}}$$

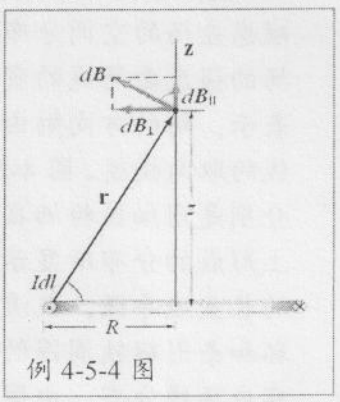
积分得到

$$\mathbf{B}=\oint dB_z\hat{\mathbf{z}}=\frac{\mu_o IR^2}{2(R^2+z^2)^{3/2}}\hat{\mathbf{z}}$$

对于远场问题($R\ll z$)，问题简化为 $\mathbf{B}=\frac{\mu_o}{2\pi}\frac{\boldsymbol{\mu}}{z^3}$，其中

$$\boldsymbol{\mu}=I(\pi R^2)\hat{\mathbf{z}}=I\mathbf{S}\qquad(4\text{-}5\text{-}13)$$

称为**磁偶极子**，其中矢量 $\mathbf{S}$ 的大小定义为平面线圈所



例 4-5-4 图

围的面积，方向则按右手螺旋法则为四指顺着电流的方向时拇指的指向。事实上对于任意位型的稳恒电流环路分布总可以看成大量的微小平面电流环构成的磁矩的矢量和。可以证明磁偶极子在远场一般位置所产生的磁场为(见习题 4-5-11)

$$\mathbf{B}=\frac{\mu_o}{4\pi}\frac{3(\boldsymbol{\mu}\cdot\hat{\mathbf{r}})\hat{\mathbf{r}}-\boldsymbol{\mu}}{r^3} \tag{4-5-14}$$

其中 $\boldsymbol{\mu}=I\mathbf{S}$ 为**磁偶极子**。

运动电荷产生的磁感应场

因自然界不存在孤立的稳恒电流元，无法利用稳恒电流元作为微型探针来测量磁感应强度的空间分布。比较可行的办法是采用测量试探运动电荷在磁场中的动力学行为来了解其受力情况，并由此给出磁场的操作性定义。

设想某一瞬间两个速度分别为 $\mathbf{v}_1$ 和 $\mathbf{v}_2$ 的点电荷 q_1 和 q_2 在其速度都远低于光速的情形，可以将组合参量 $q_1\mathbf{v}_1$ 和 $q_2\mathbf{v}_2$ 分别看成为电流元 $I_1d\mathbf{l}_1$ 和 $I_2d\mathbf{l}_2$，电荷 q_2 作用在电荷 q_1 上的磁力可以表述为

$$\mathbf{F}_{21}=q_1\mathbf{v}_1\times\left[\frac{\mu_o}{4\pi}\frac{q_2\mathbf{v}_2\times\hat{\mathbf{r}}_{21}}{r_{21}^2}\right]=q_1\mathbf{v}_1\times\mathbf{B}_{21} \tag{4-5-15}$$

其中 $\mathbf{r}_{21}$ 是由 q_2 指向 q_1 的位移矢量。则 $q_2\mathbf{v}_2$ 在点电荷 q_1 所处位置产生的磁场为

$$\mathbf{B}_{21}=\frac{\mu_o}{4\pi}\frac{q_2\mathbf{v}_2\times\hat{\mathbf{r}}_{21}}{r_{21}^2} \tag{4-5-16}$$

一般情况下，以速度 $\mathbf{v}$ 运动的点电荷 Q 在其周围产生的磁感应场为

$$\mathbf{B}=\frac{\mu_o}{4\pi}\frac{Q\mathbf{v}\times\mathbf{r}}{\gamma_v^2(1-\beta_v^2\sin^2\theta)^{3/2}r^3} \tag{4-5-17}$$

其中 $\mathbf{B}$ 是与点电荷相距 $\mathbf{r}$ 处的磁场，θ 是 $\mathbf{v}$ 与 $\mathbf{r}$ 的夹角，β_v 和 γ_v 是前面讨论过的相对论因子。磁感应场呈现为以电荷速度为轴线的圆环(见图 4-5-6)。显然，在低速情况($v\ll c$)下，$\beta_v\approx 0$，$\gamma_v\approx 1$，式(4-5-17)退化为式(4-5-16)描述的磁感应场。

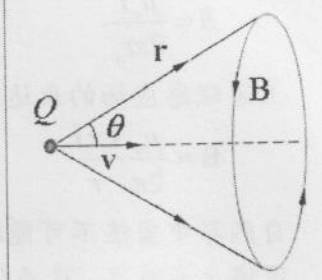


图 4-5-6 运动速度为 v 的电荷 Q 产生的磁感应场。

磁感应场的几何表示—磁感应线

类似于用电场线来描绘电场在空间的分布，我们也可以用磁感应场线来描绘磁感应场的空间分布，场的强度由场线的密度表示，场的方向则由场线的取向描述。图 4-5-7 分别是用细铁粉洒在纸上形成的分布所显示出的长直流导线，载流圆环和条形磁铁周围的磁感应场线分布。由图中

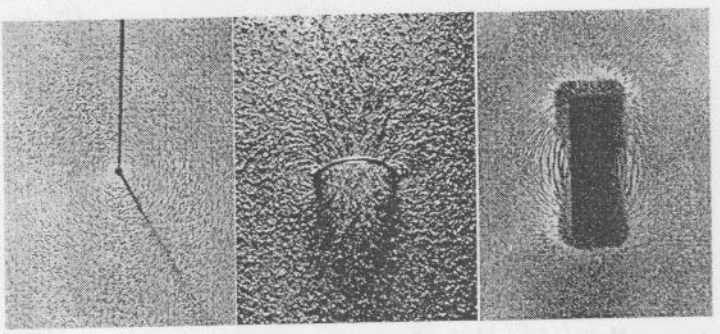


图 4-5-7 将一张白纸放在条状磁铁上面，在白纸上铺洒一堆铁屑，这些铁屑会按磁场的方向排列，形成许多条曲线。在曲线的每一点表现出磁力线的方向。

可以看出，所有的磁感应场线都是封闭的，这对应于自然界中的磁极总是成对出现的。

条形磁铁周围和离载流圆环较远处的磁感应线的分布非常接近，表明一个条形磁铁的磁场可以理解为多个载有稳恒电流的相同线圈沿垂直于线圈平面的对称轴线叠放时形成的磁感应场的矢量和。

类似于电偶极子可以有多种可能的电荷位型，各种位型的载流线圈都可以表述为磁偶极子。对于远场而言，磁偶极子的构型并不重要，磁场只决定于磁偶极矩的强度和相对位置。

4-5-4 洛仑兹力

对磁感应场的定义与静电场很不相同，这是因为不同的场对相应的作用荷的作用力的特征不同。静电场和静磁场作用于相应的作用荷的力都是中心力，而磁感应场作用于运动电荷以及电流元的作用力则为横向力(见式 4-5-12 和 4-5-15)。

可以将一个带电荷 q 的粒子在空间以速度 $\mathbf{v}$ 运动所受到的作用力中可以表述为式(4-5-12)的部分定义为该运动电荷所受到的磁感应场的作用力。运动点电荷在磁感应场中受到的作用力(见式 4-5-17):

$$\mathbf{F} = q\mathbf{v}\times\mathbf{B} \tag{4-5-18}$$

通常称为**洛仑兹力**[5]，上式不仅是运动电荷受到的磁场作用力的表达式，也可以理解为磁感应场的操作性定义。需要说明的是，以往的力场定义在操作上都是通过试探作用荷的静力学平衡实现的，而对磁感应场而言则是通过试探作用荷的动力学行为得到定义的。

洛仑兹力与电荷的运动速度垂直，也垂直于磁感应场的方向，若均匀磁感应场中一个荷电粒子q的运动速度与场垂直，粒子只能在横向的洛仑兹力作用下做圆周运动，其中回旋运动的半径r_c与粒子速度$\mathbf{v}$ 和磁感应场$\mathbf{B}$满足向心力所要求的函数关系：

$$F = |q|vB = \frac{mv^2}{r_c} \tag{4-5-21}$$

图 4-5-8 显示出在气泡室[6]中一个 γ 光子将一个氢原子电离并产生一对正负电子的事件。相关的分析见习题 4-5-5。

下面我们通过讨论一些实际事例使读者进一步理解磁相互作用，这些事例出现在近代物理的各个方面，反映了磁相互作用的各个方面和层次，每一个都曾历经非常激动人心的发现。

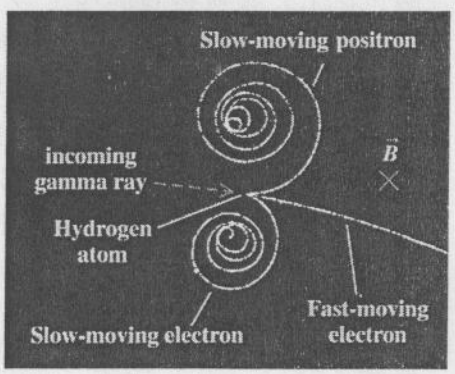


图 4-5-8 在液氢泡室中带电粒子显示出受到均匀磁场的作用而偏转的运动轨迹。

---

[5] 洛仑兹 (Hendrik Lorentz 1853-1928) 荷兰物理学家，因磁性对放射的影响的研究成果荣获 1902 年诺贝尔物理学奖。

[6] 气泡室是一种探测荷电粒子的装置：将盛有接近沸点液氢的容器置入强磁场。当进入泡室的高速电子和氢原子碰撞并使之电离，带电离子就会成为液氢的汽化核，从而在带电粒子的运动径迹和形成许多小气泡，显示出在带电粒子的运动径迹。

回旋加速器与同步加速器

一个在均匀磁感应场中作回转运动的荷电粒子 q 的运动状态由两个特征参量：**回旋半径** r_c 和**回旋频率** ω_c 表示

$$r_c = \frac{mv}{qB}, \qquad \omega_c = |q|\frac{B}{m} \tag{4-5-22}$$

由于回旋频率不依赖于回旋半径，又与粒子的速度无关，这个特征被利用来制造早期的粒子加速器—**回旋加速器**[7]。图 4-5-9 为回旋加速器的结构示意图。它基本构造包含有加速电极、射频系统、离子源、磁极、真空腔、及导引电极等。 回旋加速器的大致工作过程是将经高电压离子化产生的带电粒子引入加速器的真空腔，利用磁场使运动中之带电粒子回转，通过在电极之间交错变换电场的正负，其频率要求与带电粒子于磁场中回转频率相同，则粒子在回转过程中每次经过电极(Dees)之间就会受到电场的加速，带电粒子每回转一圈被加速两次，从而持续获得能量。形成高能粒子束后引出加速器。

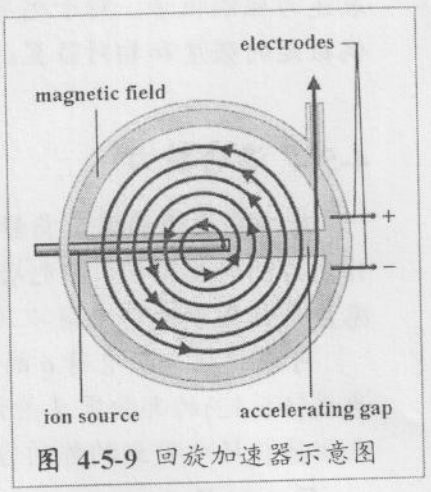


图 4-5-9 回旋加速器示意图

回旋加速器中用于加速的电压并不特别高，粒子是通过在回旋空腔中多次回转，经历多次加速获得高能量的。例如在加速电压为 10^5V 的回旋加速器中，粒子要达到 MeV 的能量需在回旋空腔中回转 100 次。

回旋加速器有两个特征。一是回旋半径随着粒子速度的增加而增大，粒子所能获得的最大动能受限制于加速器的半径 R:

$$K = \frac{1}{2}mv^2 = \frac{q^2B^2R^2}{2m} \tag{4-5-23}$$

例如在 1.5 特斯拉的磁场中将一个质子加速到 30 GeV (0.99998c)，回旋半径达到 65 米，这需要非常大尺度的磁体。另外，粒子加速到一定的能量后其质量随速度很快增加，导致回旋频率随速度很快减小(见式 4-5-22)。这导致粒子逐渐跟不上加速器中的固定的电振荡频率。解决这两个问题的方法是调整振荡频率使 $m\omega_c$ 在粒子加速过程中保持不变以维持共振加速。而同时改变磁场 B 和电振荡频率 ω_o 使得(4-5-13)式得到满足的同时，保证粒子的回旋半径在整个加速过程中为常数，可以用环型磁体来代替巨大的双 D 型磁体。采用这种技术的加速器称为**同步回旋加速器**。显然，回旋加速器运行的关键是保证固定的加速电振动频率 ω_o 与加速粒子的回旋频率 ω_c 相等。位于美国 Illinois Batavia 的质子同步加速器的存储环由 954 块独立的磁体构成,周长达到 6.3 千米。可将质子加速到高达 500GeV 的能量(设计目标为 1000 GeV)。

---

[7] 为探索原子核的基本构成，或者改变原子核的组成而形成新的核，发展出各种加速带电粒子的加速器。Ernest O. Lawrence 于 1929 年提出回旋加速器理论，随后于柏克莱加州大学建造了第一座回旋加速器。

磁学单元"磁偶矩"在磁感应场中的行为

由于自然界不存在孤立的磁单极，也不存在孤立的稳恒电流元，因而静态的磁相互作用的最小磁学单元只能是磁偶矩，而最简单的实际磁偶矩是电流环(表面上看更象磁偶矩的条形磁铁可以分解为多个电流环的叠加)。下面通过分析电流环在磁感应场中受到的作用力和电流环之间的作用力来说明磁偶矩在磁场中的受力特征以及磁偶矩与磁偶矩之间的相互作用。

首先考虑电流环处在均匀场 **B** 中的情形。相对于环平面而言，垂直于该平面的磁感应场分量 $\mathbf{B}_\perp$ 对电流环中每个电流元的作用力在环线中形成了一种张力，但对不会变形的环而言，由此形成的净力和净力矩为零(见图 4-5-10a)，而平行于环平面的磁场分量 $\mathbf{B}_\parallel$ 对环左侧的作用力和对环右侧的作用力都垂直于环平面(图 4-5-10b)，但因方向相反而形成一个力偶。由此可知，载流环受到磁场的净作用力为零，但净作用力矩不为零。

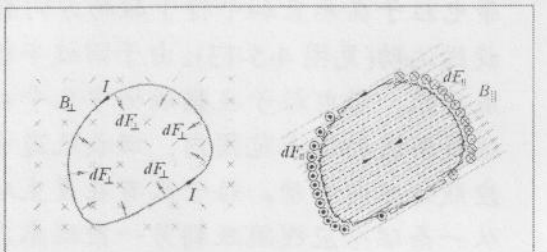


图 4-5-10 不同磁场分量对电流环的作用

为计算该力矩，将载流线圈设想为一系列相互紧靠的细长矩形线圈(见图 4-5-10b)，其中每一矩形线圈只是两端的电流元 $d\mathbf{l}_1$ 和 $d\mathbf{l}_2$ 受到磁场的作用力

$$d\mathbf{F}_1 = Id\mathbf{l}_1\times\mathbf{B}_\parallel = IB_\parallel dl\mathbf{n},\quad d\mathbf{F}_2 = Id\mathbf{l}_2\times\mathbf{B}_\parallel = -IB_\parallel dl\mathbf{n}$$

作用在该细长矩形线圈上的力矩为

$$d\boldsymbol{\tau} = \mathbf{r}_1\times d\mathbf{F}_1 + \mathbf{r}_2\times d\mathbf{F}_2 = I(\mathbf{r}_1-\mathbf{r}_2)\times(d\mathbf{l}_1\times\mathbf{B}_\parallel) = I[\mathbf{L}\times(d\mathbf{l}\times\mathbf{B}_\parallel)] = I(d\mathbf{S}\times\mathbf{B}_\parallel) = d\boldsymbol{\mu}\times\mathbf{B}_\parallel$$

其中 $\mathbf{L}=\mathbf{r}_1-\mathbf{r}_2$ 是矩形线圈的长度，$d\mathbf{S}=\mathbf{L}\times d\mathbf{l}$ 是矩形线圈的面积，而 $d\boldsymbol{\mu}=Id\mathbf{S}$ 称为该载流矩形线圈的磁偶矩。则作用在整个电流环上的净力矩为

$$\boldsymbol{\tau} = \int d\boldsymbol{\tau} = \int d\boldsymbol{\mu}\times\mathbf{B} = \boldsymbol{\mu}\times\mathbf{B} \tag{4-5-25}$$

整个电流环在磁场中受到力矩的作用而发生转动，意味着外力对磁偶矩做功

$$dW = -\boldsymbol{\tau}\cdot d\boldsymbol{\theta} = ISB\sin\theta d\theta$$

做的功以磁势能的形式存储在系统中

$$U = \int dW = -ISB\int d(\cos\varphi)$$

$$= -ISB\cos\varphi = -\boldsymbol{\mu}\cdot\mathbf{B} \tag{4-5-26}$$

其中的负号表示磁偶矩正向偏转 $d\theta>0$ 时，磁力矩做负功。

下面讨论磁偶矩在非均匀磁场中的受力特性。图 4-5-12a 给出一个刚性电流环磁场 $\mathbf{B}=\mathbf{B}(z)$ 中的受力情况(其中将磁感应强度的增长方向定义为 z-轴方向，向上汇聚的磁感应线表示磁感应场沿 z 方向增强)，各个电流元所受到的力在电流环平面方向的分量恰好互相抵消，但垂直于电流环平面方向所共有的分量使得电流环受到磁场汇聚方向的作用力。如果将电流环形式上看为一个

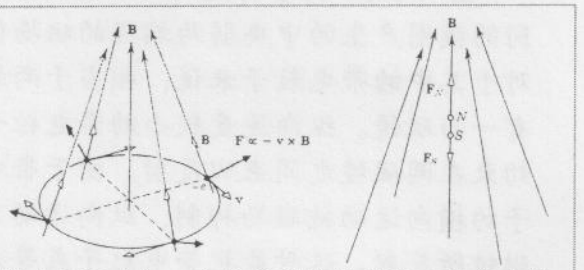


图 4-5-12 (*a*) 一个电流环的各电流元在向上会聚的磁场 $\mathbf{B}=\mathbf{B}(z)$ 中所受的作用力 **F** 分解为沿圆环的径向和磁场 **B** 的增长方向。整个电流元受到的合力沿 **B** 的增长方向。(*b*) 将电流环等效为一个条形磁铁，因 N 极受到的 z 方向的力 $\mathbf{F}_N$ 大于 S 极受到的−z 方向的力 $\mathbf{F}_S$，整个磁偶矩受到的合力沿 **B** 的增长方向。

条形磁铁，由于磁偶矩两端的磁极在非均匀磁场中受到的磁力大小不同(因而受到一个净作用力(见图 4-5-12b，类似于一定尺度的物体在不均匀引力场中受到的引潮力)，其平均值可以表示为

$$\overline{F_z} = -\nabla(\boldsymbol{\mu}\cdot\mathbf{B}) = \mu_z \frac{\partial B_z}{\partial z}$$

处在非均匀磁场中的磁偶矩一方面在沿磁场增加的方向发生位移。另一方面在力矩的作用下绕磁场为轴作进动。

磁约束

前面曾讨论过带电粒子被束缚在垂直于磁场的平面中运动的行为特征。通常带电粒子在垂直和平行于磁场方向都有运动，则其运动形式是沿着磁场方向作螺旋线运动(见图 4-5-13)。由于回旋半径与磁感应强度成反比，带电粒子在强磁场中几乎被约束在一根磁通线附近的很小范围内，回旋轨道中心只能沿磁感应线作纵向移动。粒子只有在发生碰撞时才有可能从一条磁感应线跳跃到另一条磁感应线。这意味着强磁场对带电粒子有着很强的横向约束。

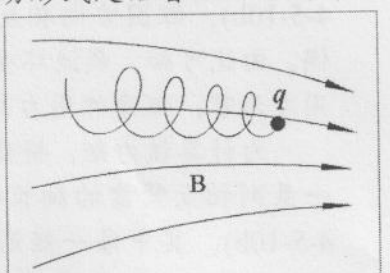


图 4-5-13

在许多场合(如受控热核反应[8])中还希望在纵向也能将粒子约束在一定的空间范围。由于均匀磁场对带电粒子不存在纵向约束作用，只能采用非均匀磁场对磁偶矩的作用力控制带电粒子的运动区域。由于在磁场中作螺旋运动的带电粒子 q 具有磁矩

$$\mu = \frac{q\omega_c}{2}R^2 = \frac{mv_\perp^2}{2B} = \frac{K_\perp}{B} \tag{4-5-28}$$

其中 ω_c 为回旋频率，R 为回旋半径，带电粒子沿磁场方向的运动过程受到减速，从而反映出粒子所受到的纵向约束。

可以证明，在梯度不很大的非均匀磁场中，带电粒子的磁矩是不变的(见习题 4-5-12)。当带电粒子从弱磁场区进入强磁场区时，粒子通过增加横向动能以维持磁矩不变，由于洛伦兹力不做功，粒子为维持总动能不变只能减小纵向速度。这样带电粒子进入强磁场区域就会在纵向减速，如果磁场分布恰当就有可能使粒子在该方向上减速到 0 再沿原路折回。由于这种行为类似于光线的镜面反射，由弱到强的磁场位形通常被称为**磁镜**。

图 4-5-14 所示的是用两个电流方向相同的线圈产生的中央弱两端强的磁场位形，对于其中的带电粒子来说，相当于两端各有一面磁镜。纵向速度较小的带电粒子被约束在两磁镜之间来回反射。由于带电粒子的横向运动被磁场抑制，纵向运动又被磁镜所反射。这种能把带电粒子或等离子

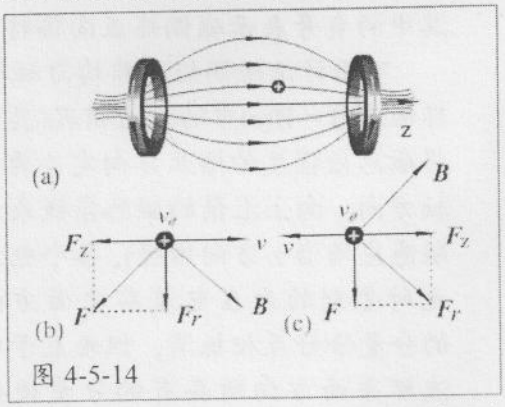


图 4-5-14

[8] 在热核反应的高温下，物质处于等离子态。热核反应的人工控制的困难之一是需要将一定密度的等离子体在高温下约束足够长的时间。等离子体的磁约束和惯性约束是实现受控热核反应的两条主要途径。

体约束在其中的磁场位形称为**磁阱**。

自然界中也存在有天然的磁阱，例如地球磁场两极强、中间弱就形成了一个天然**磁阱**，太阳喷射出的高能荷电粒子(太阳风)，到达地球上空时被地球磁场俘获，然后沿着地磁感应线作螺旋运动沉降进入地球的两极地区。由于两极附近地磁增强，这些螺旋运动的荷电粒子将被减速并折回，这样，带电粒子在两极之间沿磁感应线来回振荡，地磁场的这种特殊分布能捕获来自宇宙空间的大部分各种高能射线或粒子，保护人类和其他生物不受其伤害。从太空中观测，地球磁阱在两磁极间所约束住的带电粒子形成了内、外两个环绕地球的带，就像环绕地球的一对面包圈(见图 4-5-16)，其中内层辐射带距地表约 3000 公里，在赤道上空约 9,400 km (约为地球半径的 1.5 倍)，称为**范·艾仑辐射带**[9]，其所约束住的主要是来自太阳风和地球电离层的能量为几十个兆电子伏的质子和电子，外层辐射带距地表约为 13000 公里内层辐射带高度的三倍，其中所约束的主要是能量为数千电子伏量级的电子。介于两层中的缝隙称为**范·艾伦带缝**。这也是中层轨道卫星可以减少所受的辐射量，安全运行的区域。

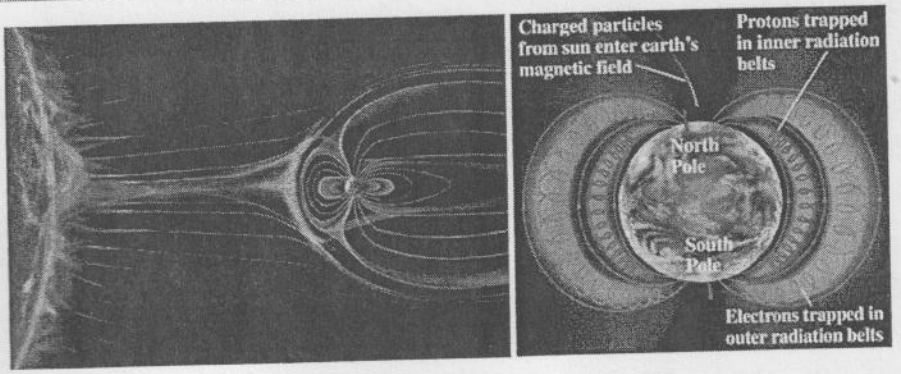


图 4-5-16 范·艾仑辐射带和范·艾伦带缝

4-5-5 电相互作用和磁相互作用的相对性

由于运动电荷周围伴随有磁场，而运动电荷也会受到磁场的作用，由此就产生了一个问题：两个静止电荷之间不存在磁相互作用，但这两个电荷相对于其它惯性参考系都具有运动速度，因而在其它惯性系都可以观察到这两个电荷之间的磁相互作用。为了帮助读者从直观上和理论上理解这一问题中所含的电磁学原理，我们分几个步骤进行讨论。

载流导线对运动电荷的作用力

我们先考虑一条载有稳恒电流 I 的长直导线和一个运动电荷 q 之间相互作用的理想实验。设想点电荷与载流导线相距为 r，并沿着电流方向以速度 u 相对于导线运动。我们可以观察到载流导线产生的磁场对运动电荷 q 的磁作用力吸引电荷向导线方向偏转。但处于点电荷 q 的参照系中的观测者认为电荷 q 与电流之间

[9] 由于带电粒子作加速运动时会辐射电磁波，被磁感应线束缚在一个带中做螺旋运动的大量带电粒子就形成了辐射带。范·艾仑(James Van Allen)辐射带的实际观测是由 1958 年美国航天局发射的探索者 1 号卫星完成。

并不存在磁相互作用，这就需要解释是什么作用力导致点电荷向导线运动的。一个出乎我们意料的结论是，这个力就是载流导线和点电荷之间静电相互作用。事实上，载流导线虽然带有相同数量的正负电荷，却并不是我们通常认为的那样是电中性的，因为导线不携载电流时，其中的正电荷密度和负电荷密度相等保证了整根导线是电中性。但导线载有电流时，导线所处的参照系中构成导线骨架的正离子(原子实)处于静止状态，而由负电荷(价电子)的定向运动构成电流。由于负电荷的运动，相邻负电荷之间会在运动方向出现洛仑兹收缩，这导致导线中负电荷密度增加，导线呈带负电状态，因而吸引附近的正电荷向导线运动。

下面通过具体计算来说明这个例子中电作用和磁作用之间在数量上的关系。按照上面的分析，载流导线一般都不是电中性的。但存在一个特殊的惯性系 S，其中长导线中的正负电荷分别以速率 u 向左和向右运动(见图 4-5-13)。由于正负电荷的运动速率相同，相邻正电荷间距和相邻负电荷间距的洛仑兹收缩率也一样，惯性系 S 中长导线是电中性的。而电荷的运动不变性使得在惯性系 S 中观察到的长导线中的正负电荷密度分别为 $\lambda_+ = \lambda_- = \lambda$。

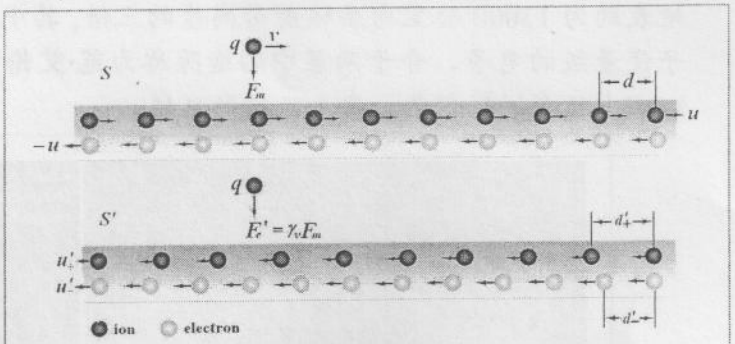


图 4-5-13 (a)S 系中相邻离子间距和相邻电子间距相同，导线呈电中性。(b) S'系中相邻离子间距和相邻电子间距的洛仑兹收缩不同，导线失去电中性。

导线上方的电荷 q 相对于惯性系 S 向右以速度 v 运动(见图 4-5-13-a)。在相对于电荷 q 静止的坐标系 S'中，导线中正负电荷的速度分别为：

$$u'_+ = \frac{u_+ - v}{1 - u_+ v/c^2} = \frac{u - v}{1 - uv/c^2} \qquad u'_- = \frac{u_- - v}{1 - u_- v/c^2} = -\frac{u + v}{1 + uv/c^2} \tag{4-5-29}$$

S'系中导线内部正负电荷的不同速度导致正负电荷间距产生不同程度的洛仑兹收缩。若将惯性系 S 中观察到的导线中的正负电荷间距记为 d，则导线中正电荷间的固有间距 d_{o+}应比 d 长一个洛仑兹因子 γ_u：

$$d_{o+} = \frac{d}{\sqrt{1 - u^2/c^2}} = \gamma_u d \tag{4-5-30}$$

而从惯性系 S'中观察到正电荷的相邻间距相对于 d_{o+}有一个洛仑兹因子 γ_u 的收缩：

$$d'_+ = d_{o+}\sqrt{1 - u'^2_+/c^2} = \frac{d_{o+}}{\gamma_{u'+}} \tag{4-5-31}$$

因此而从惯性系 S'中观察到的正电荷密度为

$$\lambda'_+ = \frac{e}{d'_+} = \frac{e}{d_{o+}}\gamma_{u'+} = \frac{e}{d_+}\frac{\gamma_{u'+}}{\gamma_{u+}} = \lambda\frac{\gamma_{u'+}}{\gamma_u}$$

其中 λ 就是惯性系 S 中观察到的电荷密度。同样可以得到惯性系 S'中观察到的负电荷密度：

$$\lambda'_- = -\frac{e}{d'_-} = -\frac{e}{d_{o-}}\gamma_{u'-} = -\frac{e}{d_-}\frac{\gamma_{u'-}}{\gamma_{u-}} = -\lambda\frac{\gamma_{u'-}}{\gamma_u}$$

这样在惯性系S'中观察到的净电荷密度为(习题4-5-12)：

$$\lambda' = \lambda'_+ + \lambda'_- = \frac{\lambda}{\gamma_u}(\gamma_{u'+} - \gamma_{u'-}) = -\frac{1}{c^2}\gamma_v v(2\lambda u) = -\frac{1}{c^2}\gamma_v vI$$

其中 $I = 2\lambda u$ 是惯性系 S 中观察到的电流。因此在惯性系 S' 中导线不再有电中性。

由于载流导线中的电荷速度仅为 1mm/min 的量级，一些读者可能会认为是否有些小题大做，但对精确的电中性的破坏的结果远比我们的直觉想象要严重得多，静电力的出现总是会伴随一些明显的可观测效应。在上述惯性系 S' 中，试探电荷 q 受到的静电力为：

$$\mathbf{F}' = \frac{q\lambda'}{2\pi\varepsilon_o r'}\hat{\mathbf{r}}' = -qv\gamma_v\frac{\mu_o I}{2\pi r}\hat{\mathbf{r}}' \tag{4-5-32}$$

其中利用了关系式：$\varepsilon_o\mu_o c^2 = 1$[10]。由于静电力 $\mathbf{F}'$ 与速度 $\mathbf{v}$ 以及电流 $\mathbf{I}$ 垂直，上式可表示为：

$$\mathbf{F}' = q\mathbf{v}\times\gamma_v\left(\mu_o\frac{\mathbf{I}\times\hat{\mathbf{r}}}{2\pi r}\right) \tag{4-5-33}$$

惯性系 S' 中水平导线产生的电场对电荷 q 的吸引作用力使它向导线偏转(图 4-5-13b)。电荷 q 向导线偏转这一现象是一个可观测的事实，同样应该在惯性系 S 中观察到。由相对论变换可得到电荷 q 在 S 系中受到导线的作用力

$$\mathbf{F} = \frac{\mathbf{F}'}{\gamma_v} = q\mathbf{v}\times\left(\mu_o\frac{\mathbf{I}\times\hat{\mathbf{r}}'}{2\pi r'}\right) = q\mathbf{v}\times\left(\frac{\mu_o}{2\pi}\frac{\mathbf{I}\times\hat{\mathbf{r}}}{r}\right) \tag{4-5-34}$$

其中考虑到垂直于运动的方向不存在洛仑兹收缩，因而有 $\mathbf{r}' = \mathbf{r}$。由于惯性系 S 中导线是电中性的，运动电荷 q 与导线没有电相互作用，q 受到的这个力只能是来自与导线中电流的相互作用。而上式括号中的正是长直载流导线所产生的磁感应场。因此，(4-5-34)式给出的就是洛仑兹力：$\mathbf{F} = q\mathbf{v}\times\mathbf{B}$。在其它惯性系中，运动点电荷受到的作用力为电场力与磁场力之和：

$$\mathbf{F} = q\mathbf{E} + q\mathbf{v}\times\mathbf{B} \tag{4-5-35}$$

运动电荷之间的作用力

由于稳恒电流只是诸多电荷匀速定向运动形成的结果，稳恒电流与运动电荷之间电磁作用的相互变换也就是运动电荷之间电磁作用的相互变换结果的叠加。电荷之间的作用力所表现出来的观测效果只是电荷之间的排斥或吸引，将不同惯性系中观测到的这种作用力分别称作电作用力或磁作用力只是一种名词术语而已，而力的数学描述形式不同只是在不同惯性系之间互相变换的结果。

现有的相关实验结果表明：电场对检测电荷的作用力与检测电荷的运动状态无关，即库仑定律适用于静止电荷对运动电荷的作用力的描述。这样，分析运动电荷之间的作用力可以先利用库伦定律描述静止电荷对运动电荷的作用力，然后利用惯性系之间的相对论变换得到处于不同运动状态的电荷之间的作用力表示。

我们先分讨论运动电荷对静止电荷的作用力这一比较简单的情况，再讨论两个运动电荷之间的相互作用的一般情况。

(1) 运动电荷对静止电荷的作用力。

设惯性系 S 中有一电荷 q_o 静止于坐标为 $(x, 0, 0)$ 的位置，另一电荷 q 以速度 v

[10] 关系式的有关推导见 12-3 节

沿 x 轴匀速运动，再设电荷 q 在其所处的惯性系 S'中静止于坐标系原点，在 $t'=t=0$ 时刻，两个惯性系的原点重合。

对 S 系中的观察者而言，它在 $t=0$ 时刻看到 q_o 位于 $x=x$ 处以及 q 经过 $x=0$ 是同时发生的两个事件，则 x 即为该时刻 q 与 q_o 的间距。而对 S'系中的观察者而言，这两个事件不是同时发生的：q 位于原点发生在 $t'=0$，q_o 经过 $x'=x$ 的时间为 $t'_1=-\gamma_v\beta_v x/c$。$t'_1$ 时刻 q 对 q_o 的作用力可用库仑定律表示为

$$F'_x=\frac{qq_o}{4\pi\varepsilon_o x'^2}\qquad F'_y=0\qquad F'_z=0$$

其中 x'为 t'_1 时刻 q 与 q_o 在 S'系中的间距，根据力的相对论变换公式(3-8-20)式可得到 S 系内 q 对 q_o 的作用力

$$F_x=F'_x=\frac{qq_o}{4\pi\varepsilon_o x'^2}\qquad F_y=F'_y=0\qquad F_z=F'_z=0$$

上式表明，在 S 系内，$t=0$ 时刻运动电荷 q 对相距 x 的静止电荷 q_o 的作用力与 S'系中 $t'=-\gamma_v\beta_v x/c$ 时刻静止电荷 q 对相距 $x'(>x)$的运动电荷 q_o 的作用力相等。折算到 $x'(>x)$运动电荷对静止电荷的作用力小于静止电荷对运动电荷的作用力：

$$F_x=\frac{qq_o}{4\pi\varepsilon_o x^2}(1-\beta_v^2)$$

运动电荷对静止电荷的作用力不再能用库仑定律描述。

设惯性系 S 中有一电荷 q_o 静止于坐标为$(0,y,0)$的位置，另一电荷 q 以速度 v 沿 x 轴匀速运动，再设电荷 q 所处的惯性系 S'中，q 静止于坐标系原点，在 $t'=t=0$ 时刻，两个惯性系的原点重合。

对 S'系中的观察者，这两个事件不是同时发生的：q 位于原点发生在 $t'=0$，q_o 经过 $x'=x$ 的时间为 $t'_1=-\gamma_v\beta_v x/c$。$t'_1$ 时刻 q 对 q_o 的作用力可用库仑定律表示为

设点电荷 q 以速度 v 相对 S 系沿 **x** 轴运动，在 $t=0$ 时刻经过原点，但检测电荷 q_o 静止在 **y** 轴上。q 与 q_o 的时空坐标分别为$(0,0,0,0)$和$(0,y,0,0)$。而在 q 静止的 S'系中，q 与 q_o 的时空坐标分别为$(0,0,0,0)$和$(0,y'=y,0,0)$，在 S'系中 q 对 q_o 的作用力服从库仑定律

$$F'_x=0\qquad F'_y=\frac{qq_o}{4\pi\varepsilon_o y'^2}\qquad F'_z=0$$

根据力的相对论变换公式(3-8-20)式可得到 S 系内 q 对 q_o 的作用力

$$F_x=0\qquad F_y=\frac{\sqrt{1-\beta_v^2}}{1-\beta_v^2}F'_y=\gamma_v\frac{qq_o}{4\pi\varepsilon_o y'^2}\qquad F_z=0$$

因 $y'=y$，在 S 系内运动的源电荷对静止的检测电荷的作用力为

$$F_y=\gamma_v\frac{qq_o}{4\pi\varepsilon_o y^2}$$

即静止在 **y** 轴上的检测电荷受到电荷 q 以速度 v 沿 **x** 轴运动时的作用大于 q 静止时的作用力。

从上述两种情形可以得到运动电荷作用在任何位置的静止电荷的力。设 S

系中源电荷 q 以速度 v 沿 **x** 轴运动，检测电荷 q_o 静止于坐标为(x, y, z)的位置。在 q 到达 S 系原点的时刻 $t=0$，q 与 q_o 在 q 静止的惯性系 S'中的时空坐标分别为 $q(0, 0, 0, 0)$和 $q_o(\gamma_v x, y'=y, z'=z, -\gamma_v\beta_v x/c)$。$S'$系中的静止电荷在该时刻对运动的检测电荷的作用力可由库仑定律求得

$$F'_x=\frac{qq_o x'}{4\pi\varepsilon_o r'^3}\qquad F'_y=\frac{qq_o y'}{4\pi\varepsilon_o r'^3}\qquad F'_z=\frac{qq_o z'}{4\pi\varepsilon_o r'^3}$$

由力的洛仑兹变换公式，可以求得以速度 v 运动的源电荷对任意位置上的静止的检测电荷的作用力的表达式，

$$\mathbf{F}=\gamma_v\frac{qq_o\mathbf{r}}{4\pi\varepsilon_o}(\gamma_v^2x^2+y^2+z^2)^{-3/2} \tag{4-5-36}$$

作用力方向仍沿该时刻源电荷到检测电荷的连线。若 $v\ll c$，上式回到库仑定律。

当**源电荷与检测电荷都运动时**，两电荷之间同时存在有电作用和磁作用。如果电荷作匀速直线运动，则总能找到源电荷静止的惯性系，不管检测电荷怎样运动，都可以用库仑定律求解源电荷对检测电荷的作用力。再通过力的变换公式就能找到运动的源电荷对运动的检测电荷的作用力。

源电荷 q 以速度 v 相对 S 系沿 **x** 轴作匀速运动，在 $t=0$ 时刻经过坐标原点。源电荷静止的参考系 S'的原点在 $t'=t=0$ 时刻与 S 系的原点重合。设检测电荷 q_o 以速度 u 相对 S 系运动，在 $t=0$ 时刻位于 S 系中的 P 点(x, y, z)，而相对于 S' 系中的时空坐标为$(\gamma_v x, y, z, -\beta_v\gamma_v x/c)$；速度为$(u_x-v, 0, 0, -\beta_v\gamma_v x/c)$。在时刻 $t'=-\beta_v\gamma_v x/c$，静止的源电荷 q 对运动的检测电荷 q_o 的作用力为

$$\mathbf{F}'=\frac{qq_o\mathbf{r}'}{4\pi\varepsilon_o r'^3}$$

式中 r'为此刻由源电荷指向检测电荷的矢径，$\mathbf{F}'$的三个分量为

$$F'_x=\frac{qq_o x'}{4\pi\varepsilon_o r'^3}\qquad F'_y=\frac{qq_o y'}{4\pi\varepsilon_o r'^3}\qquad F'_z=\frac{qq_o z'}{4\pi\varepsilon_o r'^3}$$

变换到 S 系，在 $t=0$ 时刻，q 对 q_o 的作用力为

$$F_x=\frac{F'_x+\beta_v\boldsymbol{\beta}_{u'}\cdot\mathbf{F}'}{1+\beta_v\beta_{u'_x}}=F'_x+\frac{\beta_v(\beta_{u'_y}F'_y+\beta_{u'_z}F'_z)}{1+\beta_v\beta_{u'_x}}=\frac{qq_o\mathbf{r}}{4\pi\varepsilon_o r^3}\left(x'+\beta_v\frac{y'\beta_{u'_y}+z'\beta_{u'_z}}{1+\beta_v\beta_{u'_x}}\right)$$

其中 r 为 S 系中源电荷到检测电荷的距离，记 **r** 与 **v** 之间的夹角为 θ，得

$$F_x=\frac{qq_o\mathbf{r}}{4\pi\varepsilon_o r^3}\frac{1}{\gamma_v^2}(1-\beta_v^2\sin^2\theta)^{-3/2}[x+\beta_v(y\beta_{u_y}+z\beta_{u_z})]$$

$$F_y=\frac{qq_o\mathbf{r}}{4\pi\varepsilon_o r^3}\frac{1}{\gamma_v^2}(1-\beta_v^2\sin^2\theta)^{-3/2}y(1-\beta_v\beta_{u_x})$$

$$F_z=\frac{qq_o\mathbf{r}}{4\pi\varepsilon_o r^3}\frac{1}{\gamma_v^2}(1-\beta_v^2\sin^2\theta)^{-3/2}z(1-\beta_v\beta_{u_x})$$

写成矢量形式为

$$\mathbf{F}=\frac{qq_o}{4\pi\varepsilon_o r^3}\frac{1}{\gamma_v^2}(1-\beta_v^2\sin^2\theta)^{-3/2}[\mathbf{r}+\beta_v(y\beta_{u_y}+z\beta_{u_z})\mathbf{i}-\beta_v\beta_{u_x}(y\mathbf{j}+z\mathbf{k})]$$

F 由两部分组成，一部分与检测电荷的速度无关，方向沿 **r**：

$$\mathbf{F}_{\mathrm{E}}=\frac{qq_o\mathbf{r}}{4\pi\varepsilon_o r^3}\frac{1}{\gamma_v^2}(1-\beta_v^2\sin^2\theta)^{-3/2}$$

对应于电荷间的电场力 $\mathbf{F}_{\mathrm{E}}$。另一部分与检测电荷的速度 $\mathbf{u}$ 有关，它对应于运动电荷作用于检测电荷的磁场力 $\mathbf{F}_{\mathrm{B}}$:

$$\mathbf{F}_{\mathrm{B}}=\frac{qq_o}{4\pi\varepsilon_o r^3}\frac{1}{\gamma_v^2}(1-\beta_v^2\sin^2\theta)^{-3/2}\beta_v[(y\beta_{u_y}+z\beta_{u_z})\mathbf{i}-\beta_{u_x}(y\mathbf{j}+z\mathbf{k})]$$

$$=\frac{qq_o}{4\pi\varepsilon_o r^3}\frac{1}{\gamma_v^2}(1-\beta_v^2\sin^2\theta)^{-3/2}[\boldsymbol{\beta}_u\times(\boldsymbol{\beta}_v\times\mathbf{r})]$$

匀速运动点电荷的电场和磁场

由速度 v 运动的源电荷对任意位置上的静止的检测电荷的作用力的表达式(4-5-36)，可以得到以速度 v 运动的电荷 q 的电场强度

$$\mathbf{E}=\gamma_v\frac{q\mathbf{r}}{4\pi\varepsilon_o}\left(\frac{x^2}{1-\beta_v^2}+y^2+z^2\right)^{-3/2}=\frac{q\mathbf{r}}{4\pi\varepsilon_o r^3}\frac{1}{\gamma_v^2}\left(1-\frac{y^2+z^2}{r^2}\beta_v^2\right)^{-3/2}$$

为了能更清楚地显示出电场线空间分布的特点，可将上式改写成

$$\mathbf{E}=\frac{q\mathbf{r}}{4\pi\varepsilon_o r^3}\frac{1}{\gamma_v^2}(1-\beta_v^2\sin^2\theta)^{-3/2}\tag{4-5-37}$$

其中，θ 为源电荷指向考察点的位置矢量 $\mathbf{r}$ 与源电荷的速度 $\mathbf{v}$ 之间的夹角。运动电荷产生的电场仍在该电荷的径向，但场强的分布不是各向同性的。r 一定时，在电荷运动的方向($\theta=0$)场强最小，电场线分布较疏，而垂直于电荷运动的方向($\theta=\pi/2$)，场强最大，电场线分布较密(见图 4-5-14)。相对这一参考系，匀速运动点电荷存在一个特殊方向，因此电场分布没有球对称性，电荷运动的速度越大，偏离球对称越厉害。磁场对运动电荷的作用力为 $\mathbf{F}_{\mathrm{B}}=q_o\mathbf{u}\times\mathbf{B}$，与(6.3-1)式比较可得以速度 $\mathbf{v}$ 作匀速运动的点电荷产生的磁场的磁感强度为

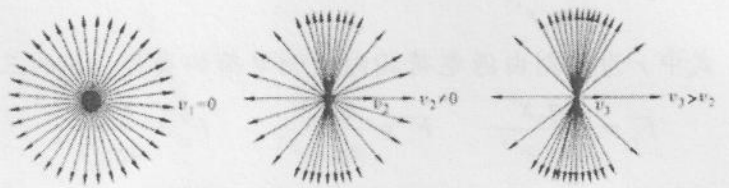


图 4-5-14 几种不同速度的匀速运动点电荷电场的电场线分布

$$\mathbf{F}_{\mathrm{B}}=\frac{q}{4\pi\varepsilon_o r^3c^2}\frac{1}{\gamma_v^2}(1-\beta_v^2\sin^2\theta)^{-3/2}(\mathbf{v}\times\mathbf{r})=\frac{\mu_o}{4\pi}\frac{q\mathbf{v}\times\mathbf{r}}{r^3}\frac{1}{\gamma_v^2}(1-\beta_v^2\sin^2\theta)^{-3/2}$$

当 $v\ll c$ 时，上式与毕奥-沙伐尔定律一致。电荷的电场与磁场满足关系

$$\mathbf{B}=\frac{1}{c^2}\boldsymbol{\beta}_v\times\mathbf{E}$$

即磁感强度 $\mathbf{B}$ 既垂直于电荷运动的速度 $\mathbf{v}$，又垂直于电场强度 $\mathbf{E}$。运动电荷的电场的电场线是从源电荷发出的沿径向的直线，而磁感线是位于垂直于源电荷主动方向的平面内的同心圆，如图 4-5-15 所示。

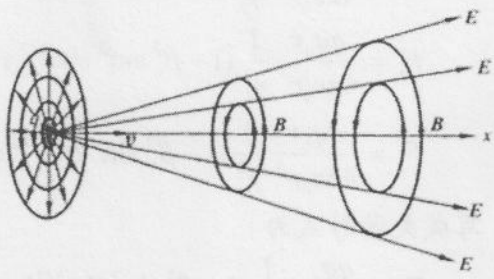


图 4-5-15 沿 $\mathbf{x}$ 方向匀速运动的点电荷的磁场图中给出了垂直于 $\mathbf{x}$ 轴的三个平面内的磁感线的分布

例 4-5-5 计算两匀速运动的点电荷之间的相互作用力。设两电荷的电量都是 q，速度都是 v，运动轨道相距 d。

解：因两电荷之间的距离 $r=d$，运动电荷到检测电荷的矢径与速度方向的夹角 $\theta=\pi/2$，故两电荷间的电力为

$$F_{\mathrm{E}}=\frac{q^2}{4\pi\varepsilon_o d^2}\frac{1}{\gamma_v^2}(1-\beta_v^2)^{-3/2}=\frac{q^2}{4\pi\varepsilon_o d^2}\gamma_v$$

电力是斥力。两电荷间的磁力为

$$F_{\mathrm{B}}=\frac{q^2}{4\pi\varepsilon_o d^3}\frac{1}{\gamma_v^2}(1-\beta_v^2)^{-3/2}\frac{v^2}{c^2}d=\frac{q^2}{4\pi\varepsilon_o d^2}\gamma_v\frac{v^2}{c^2}$$

为引力，但只有电力的 v^2/c^2 倍。两电荷间相互作用力的合力为

$$F=F_{\mathrm{E}}-F_{\mathrm{B}}=\frac{q^2}{4\pi\varepsilon_o d^2}\frac{1}{\gamma_v}$$

两匀速运动的同号电荷间的作用力表现为斥力，因为这作用力主要由电力决定，但是当两平行直导线中通以同向的电流时，两导线间的作用力表现为引力。◀

讨论

读者也许会对于式(4-5-37)表述的含义感到困惑：运动电荷的电场所包含的不只是以有限传播速度形成的推迟效应后的静电场，还包括因电荷的几何尺度导致的推迟的先后，以及空间电场、磁场随时间变化所造成的效应，但式(4-5-37)显示，匀速运动电荷 q 的电场的方向始终沿着相应时刻源电荷指向考察点的矢径 $\mathbf{r}$，似乎 q 所处位置的信息瞬间传到了 q_o 所在的位置。其中的原因在于对于匀速直线运动的点电荷，检测电荷 q_o 在受力时刻可从 q 的运动历史“估计”出 q 在该时刻的位置。其最终效果是 q 产生的电场恰好沿着同一时刻 q 与 q_o 的连线。若源电荷在极短时间内有加速度，场分布就将发生畸变。进一步的讨论见 12-3 节。

从以上的讨论可知，对于两个处于静止状态的物体之间的非超距作用的中心力，只要两个物体出现相对运动，物体之间的作用力就不再是中心力，并和它们的运动速度有关，而且通常不再遵循第三定律。由于两个静止物体之间的万有引力作用也是非超距作用的中心力，因而可以想象，行星和太阳之间的相互作用也应该包含和运动速度相关联的成分，我们通常观测不到与运动相关的“磁引力”作用部分只是因为它所产生的观测效应很小。目前已知的可观测效应是内行星的近日点进动包含有相应的成分。。

第5章 刚体力学

宏观物体是由大量微观粒子构成的多质点系统，因而它的运动一般都会涉及到大量的运动自由度。通常我们将一个宏观物体的运动分解为两部分：物体质心的运动和物体内各质点相对于质心的运动。由于系统中各粒子运动自由度之间存在相互关联，通常我们只是利用质心运动模型计算一个多质点系统的整体平动自由度，而避免涉及系统中各个质点相对于质心的运动问题。

虽然我们甚至无法严格求解一个只含有三个质点的系统的运动方程，但并非对所有的宏观物体运动都束手无策。事实上，通常我们所关心的宏观物体的实际运动问题中，大量的(内部)自由度因各种原因而受到约束，导致物体在运动中所能有效表现出来的自由度并不多，而表现出来的运动自由度往往还呈现出某些规律性，因而可以通过一些特殊的数学方法进行处理。

5-1 刚体模型

物体的内部自由度受到约束的极端事例是物体在运动过程中其大小和形状都不能发生改变，物质的这种特性称为**刚性**，这种约束在数学上表示为：物体中任意两点的间距在运动过程中是一个不变量

$$|\mathbf{r}_j - \mathbf{r}_i| = \text{常数} \tag{5-1-1}$$

其中 $\mathbf{r}_i$ 和 $\mathbf{r}_j$ 表示物体中两个相关质点相对与空间某一固定点 O 的位置矢量(见图 5-1-1)。满足几何约束条件式(5-1-1)的物体称为**刚体**。由于刚体不会发生形变，刚体中各质点相对于质心的运动只能表现为整个物体取向的变化，即系统的整体转动。

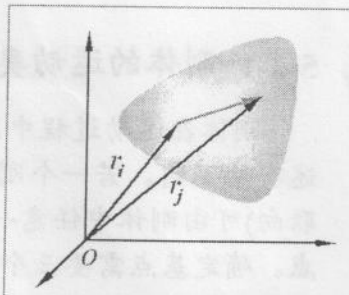


图 5-1-1 刚体模型的特点是物体中任意两点的间距不随其运动而改变。

和质点类似，刚体也只是描述物体某种特定运动形式的模型。自然界中并不存在真正的刚体，即使对最接近刚体模型的固态物体而言，从微观尺度看，构成固态物体的原子或分子始终在各自的平衡位置附近处于无规则的随机运动状态，因而任何一对粒子之间的距离都处于不停的变化之中；而从宏观尺度上说，即使坚硬的金刚石晶粒在足够大的压力负载下也会出现可观察的形变。当一个刚体转动时，其内部不同位置会出现应力作用。设想一飞轮以100 rad/s(约16圈/秒)的角速度旋转，要将一个距离转轴1米的物体保持在原来的位置所需要的力($F=m\omega^2 r$)相当于该物体在地面附近所受重力的1000倍。如果该力超过材料本身的抗张强度，物体就会被撕裂。

另外，绝对刚度的概念与相对论相抵触。设想一根刚性杆绕其一端以一定的角速度转动，只要杆足够长，杆的另一端的运动速度就会超过光速。而即使对一个小尺度的物体，刚度要求任何一点的局部扰动在整个物体中得到瞬时的响应，即要求在刚体中相互作用的传播速度达到无穷大。

物体的刚性很容易和物体的硬度混淆，事实上，两者在概念上反映的是物体对外部扰动的不同响应特性，即分别用于处理不同的力学问题。硬度反映的是物体受到外力(挤压)作用下的形变程度，其本质是物体的内部运动自由度在外力作

用下的响应和协调，即物体的运动或平衡状态涉及到物体的内部自由度，而刚性反映的是物体在外力作用下整体位型(位置和取向)随时间，其实质是物体的外部(即整体)自由度对外力作用的响应和协调。因而在求解运动演化或平衡问题时，如果我们所关注的只是系统的整体行为，刚性假设可使问题得到很大程度的简化，由此可知，只要在运动过程中，物体的形变对所涉及的问题影响很小，刚体模型对于各种尺度物体的运动都很有效，即使对从常规角度观察是比较柔软的物体也是如此。例如，处于自由翻滚下落过程中的海绵块的运动状态就可以作为刚体进行描述。

5-2 刚体运动学

刚体运动学包含两个方面：刚体的整体运动及刚体内各质点的相对运动。由于刚体不能发生形变，刚体内各质点的相对运动就表现为整体的转动，因而刚体运动分为平移和转动，相应的运动参量为位移，速度和加速度以及角位移，角速度和角加速度。

5-2-1 刚体的运动类型及相应的自由度

刚体在运动过程中因受到不同的约束而具有不同的自由度，因而对运动的描述有所不同。若一个刚体在运动过程中不受外部约束，刚体的空间位型(位置和取向)可由刚体中任意一点 A 以及刚体相对于 A 点的取向确定，其中 A 点称为**基点**。确定基点需要三个参量(位置坐标)，而刚体相对于基点的取向可以利用过基点的某一直线的方向及刚体相对于该直线的方位确定，其中确定直线方向需要两个独立变量(方向余弦)，确定刚体相对于该直线的方位需要一个参数(绕该直线转过的角度)。因此不受外部约束的刚体具有三个平移自由度和三个转动自由度。

刚体所受到的外部约束可以分为四类，相应的运动类型和运动自由度分别为

(1) 如果刚体在运动过程中体内任何一条直线保持其取向不变，则刚体中所有质点具有相同的位移、速度和加速度，其整体位置可用其中基点的坐标确定。这种运动就是**平移**。平移刚体方位不变，只有三个平移自由度而没有转动自由度。

(2) 如果运动过程中刚体内只有一点固定不动，称为**定点运动**。设想以该固定点为基点，则处于定点运动的刚体没有平移自由度，而只有三个转动自由度。

(3) 如果刚体在运动过程中体内只有一条直线保持静止，整个刚体就只能围绕该直线转动，称为**定轴转动**，该直线称为**转动轴**。定轴转动中的刚体没有整体平移，转轴方向也不变，因此没有平移自由度及描述转轴方位的两个自由度。刚体位置可用其相对于转轴的取向—角坐标 θ 表示，只有一个转动自由度。

(4) 如果存在一固定平面，刚体在运动过程中体内凡垂直于该平面的任一直线保持取向不变，而且直线上各点运动情况相同。则所有质点的运动都约束在平行于该固定平面的各自平面内，其中任一平面的运动状态可描述整个刚体的运动。这样的运动称为**平面平行运动**：平面中的基点在其所在平面中的平移和整个平面绕过基点且垂直该平面的直线的转动。具有两个平移自由度和一个转动自由度。

刚体的一般运动形式是平移和转动的叠加，图 5-2-1 所显示的是一刚性扳手匀速滑过一光滑的玻璃面的过程中表现出的整体平动和转动的叠加特征。由于平动可以由质点模型描述，刚体模型主要反映的是物体的转动特征[1]。本章主要利用刚体模型来说明实际物体的基本转动特性。

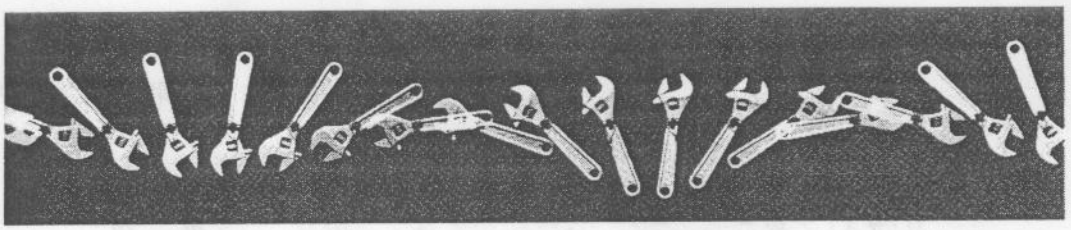


图 5-2-1 刚性扳手的一般运动只是整体平动和转动的叠加。

5-2-2 刚体的转动

刚体的一般转动形式是整个系统围绕着一个固定点 O 的转动，其特征是刚体中的每一个质点在运动过程中各自维持与固定点 O 的距离不变。例如，图 5-2-1 中的立方体的一个顶点被固定于 O 点，则 P 和 P'在转动过程中分别处于以 O 点为中心，立方体的对角线或边长为半径的球面上。如果图 5-2-2 中的立方体上的两点 O 和 O'被固定，则刚体的运动只能是绕 $\overline{OO'}$ 联线的定轴转动，P 和 P'的运动轨迹都是垂直于 $\overline{OO'}$ 轴线的平面上的圆周运动。由于刚体的定轴转动只需要一个自由度参量，下面从这一最简单的转动形式作为讨论的出发点。

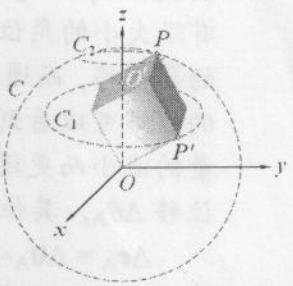


图 5-2-2 刚体的定点转动和定轴转动

定轴转动

刚体绕固定轴转动的运动特征表现为系统中各质点在垂直与转动轴的平面上以不同的半径作圆周运动。虽然刚体在运动过程中各点的位移、速度和加速度不同，但它们与转轴的垂直连线在相同的时间内所转过的角度 θ 都相同，因此转角 θ 就足以描写刚体中所有质点在定轴转动过程中的位置：设刚体上的任何一点 P 在时间Δt 内绕轴(取为 z 轴)转过角度$\Delta\theta$(称为**角位移**)，该点在任意时刻 t 的位置可表示为(见图 5-2-3)

$$x = r\cos\theta(t), \quad y = r\sin\theta(t)$$

其中 r_i 为 P_i 点与 z-轴的垂直距离。因此角位移 θ 实际上是描述刚体定轴转动的唯一参量。

量度角位移变化率的物理量称为“**角速度**”

$$\omega = \lim_{\Delta t \to 0} \frac{\Delta\theta}{\Delta t} = \frac{d\theta}{dt} \tag{5-1-1}$$

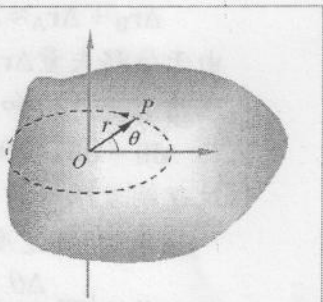


图 5-2-3 刚体的定轴转动

[1] 实际物体的非质心运动可以分为外部运动和内部运动两个部分，(5-1-1)式意味着刚体不具有内部自由度。因而刚体的运动是可以直接观测的外部运动。

其单位是 $\mathrm{rad\cdot s^{-1}}$。类似地还可以定义角速度的变化率：

$$\alpha=\lim_{\Delta t\to 0}\frac{\Delta\omega}{\Delta t}=\frac{d\omega}{dt}=\frac{d^2\theta}{dt^2} \tag{5-1-2}$$

称为**角加速度**，其单位是 $\mathrm{rad\cdot s^{-2}}$。

空间转动参量的矢量性质

定轴转动参量：角位移，角速度和角加速度都是标量，描述定轴转动的运动学参量和描述质点直线运动的运动学参量存在对应：

$$\theta\to x,\quad \omega=\frac{d\theta}{dt}\to v=\frac{dx}{dt},\quad \alpha=\frac{d\omega}{dt}=\frac{d^2\theta}{dt^2}\to a=\frac{dv}{dt}=\frac{d^2x}{dt^2}$$

读者也许会很自然地想到是否可以将这种对应关系推广到三维情形，即建立一个系统的空间转动和三维平动的对应关系。但这种对应要求描述转动的运动学参量也要具有矢量特性。1-5-1 节中曾通过具体事例说明，有限大小的角位移并不满足矢量代数所要求的加法交换律，即物体绕过同一定点的两条不同轴 $\mathbf{n}_1$ 和 $\mathbf{n}_2$ 分别旋转 $\boldsymbol{\theta}_1$ 和 $\boldsymbol{\theta}_2$ 的合成效果与这两个转动的次序有关：$\boldsymbol{\theta}_1+\boldsymbol{\theta}_2\neq\boldsymbol{\theta}_2+\boldsymbol{\theta}_1$。由此可知，有限大小的角位移并不是矢量。但描述刚体转动的(瞬时)角速度只涉及到无穷小的角位移，而通过观察刚体相继绕两个不同转轴的微小转动导致刚体上任一点 P 的位移可以看到，两个角位移之和因次序先后造成的差异随着两个角位移本身数量的减小而更迅速地减小。图 5-2-4 所示为刚体上的 P 点随刚体绕 OA 轴转过角位移 $\Delta\theta_A$，其位置矢量 $\mathbf{r}$ 所经历的位移可表示为

$$\Delta\mathbf{r}_A=\Delta\boldsymbol{\theta}_A\times\mathbf{r} \tag{5-2-3}$$

其中 $\Delta\boldsymbol{\theta}_A$ 的方向沿 OA 轴。再让刚体绕轴 OB 旋转$\Delta\boldsymbol{\theta}_B$，$P$ 点的新位移$\Delta\mathbf{r}_B$ 为

$$\Delta\mathbf{r}_B=\Delta\boldsymbol{\theta}_B\times(\mathbf{r}+\Delta\mathbf{r}_A)\approx\Delta\boldsymbol{\theta}_B\times\mathbf{r}$$

其中略去了二阶无穷小项 $\Delta\boldsymbol{\theta}_A\times\Delta\mathbf{r}_A$。$P$ 点经两次转动后形成的净位移为

$$\begin{aligned}\Delta\mathbf{r}_A+\Delta\mathbf{r}_B&\approx\Delta\boldsymbol{\theta}_B\times\mathbf{r}+\Delta\boldsymbol{\theta}_B\times\mathbf{r}\\&=(\Delta\boldsymbol{\theta}_A+\Delta\boldsymbol{\theta}_B)\times\mathbf{r}\end{aligned}$$

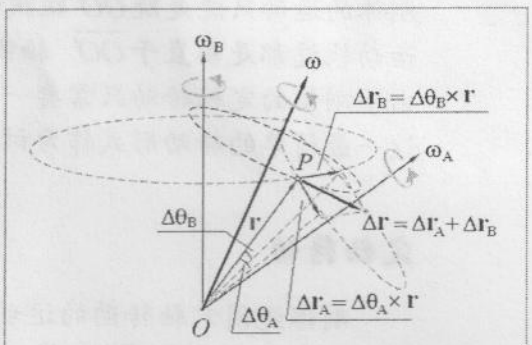


图 5-2-4 刚体无限小定点转动的叠加。

如果变更转动次序，P 点随刚体先绕 OB 轴转 $\Delta\boldsymbol{\theta}_B$，再绕 OA 轴转 $\Delta\boldsymbol{\theta}_A$，$P$ 点相应的净位移为：

$$\Delta\mathbf{r}_B+\Delta\mathbf{r}_A\approx\Delta\boldsymbol{\theta}_B\times\mathbf{r}+\Delta\boldsymbol{\theta}_A\times\mathbf{r}=(\Delta\boldsymbol{\theta}_B+\Delta\boldsymbol{\theta}_A)\times\mathbf{r}$$

由于位移矢量$\Delta\mathbf{r}_A$ 和$\Delta\mathbf{r}_B$ 满足加法交换律：$\Delta\mathbf{r}_A+\Delta\mathbf{r}_B=\Delta\mathbf{r}_B+\Delta\mathbf{r}_A$，在 $\Delta\theta_i\to 0$ 的极限情况下 $\Delta\boldsymbol{\theta}_A$ 和 $\Delta\boldsymbol{\theta}_B$ 也满足加法交换律，记为

$$\mathbf{d\theta}=\mathbf{d\theta}_A+\mathbf{d\theta}_B=\mathbf{d\theta}_B+\mathbf{d\theta}_A \tag{5-2-4}$$

这说明无限小角位移是矢量。

(瞬时)角速度是无限小角位移与相应无限小时间间隔之比。因此也是矢量：

$$\boldsymbol{\omega}=\lim_{\Delta t\to 0}\frac{\Delta\theta}{dt}\mathbf{n}=\frac{\mathbf{d\theta}}{dt} \tag{5-2-5}$$

相应地角加速度也是矢量：

$$\boldsymbol{\alpha}=\lim_{\Delta t\to 0}\frac{\Delta\boldsymbol{\omega}}{dt}=\frac{d\boldsymbol{\omega}}{dt} \tag{5-2-6}$$

定点转动的分解

类似于质点的一般(曲线)运动总可以分解为三个互相独立的直线运动分量的叠加，刚体的定点转动也可以分解为三个独立的定轴转动的叠加[2]。图 5-2-5-b 描述的是飞机在飞行特技表演中复杂的取向变化分解为转圈，翻滚及自转三个相互独立的定轴转动，这三种定轴转动对应的转动轴互相垂直。图 5-2-5-a 所描述的是一个陀螺的定点(与地面的接触点)转动可以分解为绕陀螺本身对称轴(自转轴)的自转，该对称轴绕竖直轴的转动(进动)以及自转轴绕与自转轴和进动轴构成的平面垂直的轴线的转动(章动，表现为自转轴和进动轴夹角的变化)。5-4-3 节将给出刚体空间转动分解为三个独立的定轴转动的数学表示。

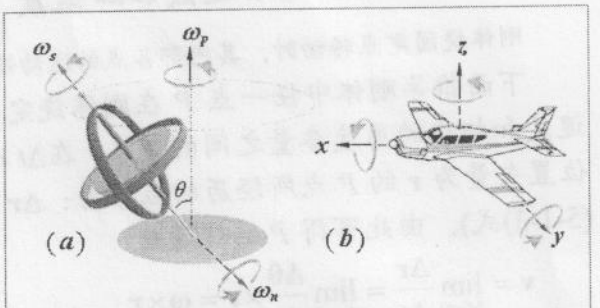


图 5-2-5 刚体定点转动分解：(a) 陀螺定点转动分解；(b) 飞机一般转动的分解。

定点转动中的瞬时转动轴

刚体的转动与质点的平移之间的类似性还表现在质点的一般曲线运动在任一无限小过程中总可以看成一个由瞬时速度描述的直线运动，而处于定点转动的刚体在任一无限小过程中也总能看成绕通过该定点的某一轴线的定轴转动，例如：图 5-2-4 中 P 点的最终位移等于绕 OC 轴转$\Delta\boldsymbol{\theta}$ 角所产生的位移$\Delta\mathbf{r}=\Delta\boldsymbol{\theta}\times\mathbf{r}$。通常，将刚体的定点运动分解为一系列无限小的位置相继变化时，刚体在每一无限小时间间隔内的位置变化都可以描述为刚体绕过该固定点的某一轴线的一个无限小纯转动，该轴线称为**瞬时转轴**。

绕固定点 O 转动的刚体中以 O 点为中心的球面上任选两点 A 和 B(图 5-2-6)，O, A, B 三点确定了整个刚体的位置。设经过Δt 时间 A 和 B 的位置分别运动到 A′和 B′。A, A′, B, B′都处于以 O 为中心，OA 为半径的球面上。在该球面上分别作垂直平分圆弧 AA′和 BB′的两条圆弧，两者相交于 C 点，球面三角形 CAB 与 CA′B′三条对应的边分别相等：AB = A′B′(刚体的任意两点间距不变)，AC = A′C，BC = B′C；因而是全同形。因而有∠ACB = ∠A′CB′，两边都加上∠BCA′，得到

∠ACA′ = ∠BCB′

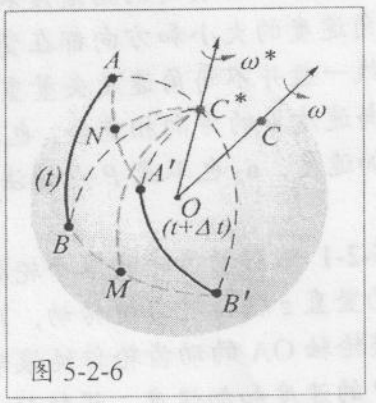


图 5-2-6

即上述 AB 经 Δt 运动到 A′B′的有限位移等价于刚体在时刻绕 OC 转动 $\Delta\theta$ 实现。当 $\Delta t\to 0$，OC^*轴的极限位置 OC 即为该时刻的瞬时转动轴。

---

[2] 这一表述通常称为欧拉(Euler)定理。

由于瞬时轴的方位随时间变化，在相应时刻只要找到刚体上的某一速度为零的点，该点和定点 O 的连线即为该时刻的刚体瞬时转动轴。

由此得到**欧拉定理**：定点运动刚体的任何一个位置移动都可以通过绕着过此定点的某一轴旋转一个角度达到。而定点运动刚体的连续运动则可以描述为它依次围绕一系列瞬时转轴转动无穷小角度的结果。

刚体作定点转动时，每一时刻的运动状态都可由即时的瞬时转动角速度 $\boldsymbol{\omega}$ 描述，其方向即为瞬时轴的方向。刚体运动状态的变化由角加速度 $\boldsymbol{\alpha}$ 描述，其方向和瞬时轴的方向一般并不一致。

刚体内各质点的运动速度和加速度

刚体绕固定点转动时，其内部各点的运动情况一般各不相同。

下面推导刚体中任一点 P 在刚体绕定点 O 转动时(见图 5-2-7)线速度和线加速度和相应的角度参量之间的关系。在Δt 时间内刚体绕某瞬时轴转过角位移$\Delta\boldsymbol{\theta}$，位置矢量为 $\mathbf{r}$ 的 P 点所经历的位移为：$\Delta\mathbf{r} = \Delta\boldsymbol{\theta}\times\mathbf{r}$ (见(5-1-3)式)，由此可得 P 的线速度

$$\mathbf{v} = \lim_{\Delta t\to 0}\frac{\Delta\mathbf{r}}{\Delta t} = \lim_{\Delta t\to 0}\frac{\Delta\boldsymbol{\theta}}{\Delta t}\times\mathbf{r} = \boldsymbol{\omega}\times\mathbf{r} \tag{5-1-7}$$

其大小为 ωh_1，h_1 为质点 P 到瞬时转轴(沿 $\boldsymbol{\omega}$ 的直线)的垂直距离，而线加速度为

$$\mathbf{a} = \frac{d\boldsymbol{\omega}}{dt}\times\mathbf{r} + \boldsymbol{\omega}\times\frac{d\mathbf{r}}{dt} = \boldsymbol{\alpha}\times\mathbf{r} + \boldsymbol{\omega}\times\mathbf{v} \tag{5-1-8}$$

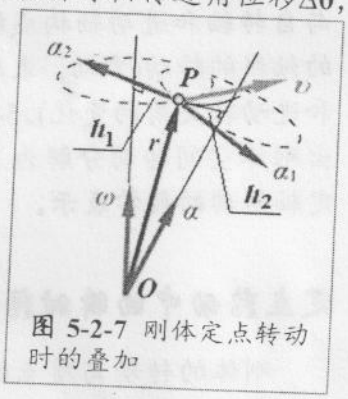


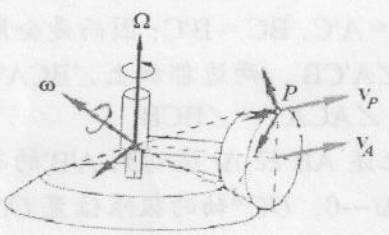
图 5-2-7 刚体定点转动时的叠加

右端第一项 $\mathbf{a}_1 = \boldsymbol{\alpha}\times\mathbf{r}$ 是刚体绕定点转动的角速度 $\boldsymbol{\omega}$ 的变化(包括 $\boldsymbol{\omega}$ 的大小变化和指向变化，$\boldsymbol{\omega}$ 的指向变化就是瞬时轴线的变化)所导致的“切向”加速度，称为**转动加速度**，其方向垂直于 $\boldsymbol{\alpha}$ 和 $\mathbf{r}$ 构成的平面，记质点 P 到沿 $\boldsymbol{\alpha}$ 的直线的垂直距离为 h_2，则 $\mathbf{a}_1$ 大小为 αh_2；第二项 $\mathbf{a}_2 = \boldsymbol{\omega}\times\mathbf{v}$ 是刚体绕瞬时转动轴线转动所导致的“向心”加速度，其方向指向瞬时轴，因而称为**向轴加速度**，其大小为 $\omega^2 h_1$。

瞬时轴上除定点外的各点速度为零。而加速度不为零；刚体上任一点 P 在各特定时刻绕一瞬时轴作圆周运动，而转轴位置的变化使 P 点的轨迹不是圆周。

刚体绕定轴转动时的角速度和角加速度都沿着固定的转轴，而刚体绕定点运动时的角速度的大小和方向都在变化。角加速度沿角速度矢量端点所描绘出的曲线的切线一般并不与角速度矢量重合。在图 5-2-5 中，转动加速度 $\mathbf{a}_1 = \boldsymbol{\alpha}\times\mathbf{r}$ 的方向既不与速度 $\mathbf{v}$ 的方向相重合，也不与向轴加速度 $\mathbf{a}_2 = \boldsymbol{\omega}\times\mathbf{v}$ 垂直，即 $\mathbf{a}_1$ 不是 P 点的切向加速度，$\mathbf{a}_2$ 也不是 P 点的法向加速度。

例 5-2-1 锥形动齿轮的水平轮轴 OA 绕过底齿轮中心的竖直 z 轴转动匀速转动，角速度为Ω (如图)。连接轮轴 OA 的动齿轮作纯滚动。求动齿轮的最高点 P 的速度和加速度。设轮轴长度为 l，动齿轮半径为 r。



解： 动齿轮中心点 A 绕定点 O 在水平面内作圆周运动，其速度为

$$v_A = l\Omega$$

底齿轮和动齿轮母线的啮合意味着动齿轮与地面接触点 C 处的速度为零，因而固定点 O 与 C 点的连线构成动齿轮(复合转动)的瞬时轴，相应的角速度大小为

$$\omega = \frac{\Omega}{\sin\theta}$$

方向沿瞬时轴(见图)。由此可求得齿轮上端 P 点的速度大小

$$|v_P| = |\boldsymbol{\omega} \times \overline{CP}| = \frac{\Omega}{\sin\theta} \cdot 2r\cos\theta = 2l\Omega$$

速度方向垂直于 OPC 平面。因角速度 $\boldsymbol{\omega}$ 只是改变方向，不改变大小，动齿轮相应的角加速度为

$$\boldsymbol{\alpha} = \frac{d\boldsymbol{\omega}}{dt} = \boldsymbol{\Omega} \times \boldsymbol{\omega}$$

其方向沿角速度 $\boldsymbol{\omega}$ 的矢量端点描绘出的水平圆周的切线，而大小为

$$\alpha = \Omega \cdot \omega\cos\theta = \Omega\frac{\Omega}{\sin\theta}\cos\theta = \Omega^2\cot\theta$$

P 点加速度中的转动加速度分量 $\mathbf{a}_1$ 的大小为

$$a_1 = |\boldsymbol{\alpha} \times \overline{OP}| = \Omega^2\cot\theta\frac{l}{\cos\theta} = \frac{l}{\sin\theta}\Omega^2$$

其方向垂直于由 $\boldsymbol{\alpha}$ 和 OP 形成的平面，指向如图所示。向轴加速度 $\mathbf{a}_2$ 的大小为

$$a_2 = \omega^2 \cdot \overline{PE} = \omega^2 2\ell\sin\theta = 2\frac{\ell}{\sin\theta}\Omega^2$$

其方向自 P 指向点 E(在铅直平面 OAC 内)。根据 $\mathbf{a} = \mathbf{a}_1 + \mathbf{a}_2$ 及图中几何关系得

$$a^2 = a_1^2 + a_2^2 - 2a_1a_2\cos 2\theta$$

由几何关系

$$\cot\theta = \frac{\ell}{r} \quad 和 \quad \sin\theta = \frac{r}{\sqrt{r^2+\ell^2}}$$

可得齿轮上 P 点的加速度大小为

$$a = \Omega^2\ell\sqrt{9+\left(\frac{\ell}{r}\right)^2} \blacktriangleleft$$

5-3 刚体动力学

理解一个物体的某种动力学行为首先需要定义描述这一类运动的动力学参量，再根据这些运动参量建立并求解系统运动的演化方程。由于一个多质点系统的动力学行为只是系统中各个质点的平动行为的某种组合，相应的动力学参量也就是系统中各个质点的平动参量的叠加。刚体是一个受到特殊约束的多质点系统，刚体绕固定点或固定轴转动时，其中的每一个质点被约束在相对于该固定点或固定轴的特定球面或圆周上运动，恰当描述这种动力学行为可能需要一些新的概念和相应的物理量，但不应该会出现完全不同于平动的全新物理规律。因此计入描述所有这些质点的圆周运动动力学参量—角动量和转动动能以及转动过程中系统所表现出来的惯性也就足以反映刚体的转动行为及其特性。本节从刚体作为一个受到特殊约束的多质点系统的角度出发，讨论刚体的动力学问题。

5-3-1 描述转动的动力学参量

描述刚体转动的特征运动参量是角位移和角速度，而相关的动力学参量为刚体的角动量和转动动能。下面分别讨论它们的数学表达式。

角动量

首先需要强调的是，无论多质点系统的转动是定点转动，还是定轴转动，计算系统的总角动量时，所有质点都必须相对于一个共同的固定点而言的。对于定点转动，固定点是唯一的，如果只是考虑某一个瞬间的转动，形式上也可以描述为刚体绕过该点的固定轴线(瞬时轴)的转动，但转轴的取向(即角速度 $\boldsymbol{\omega}$ 的方向)是时间的函数。而对于定轴转动，轴线上任何一点都可以作为计算系统角动量的固定参考点，系统转动角速度 $\boldsymbol{\omega}$ 的方向是不变的。

记刚体中质点 m_i 相对于固定点 O 的角动量为 $\mathbf{L}_i = \mathbf{r}_i \times \mathbf{p}_i = \mathbf{r}_i \times (m_i \mathbf{v}_i)$，其中 $\mathbf{r}_i$ 为该质点相对于 O 点的位置矢量，$\mathbf{v}_i$ 为该质点相对于固定坐标系的速度，利用 $\mathbf{v}_i$ 和刚体的转动角速度 $\boldsymbol{\omega}$ 之间的关系：$\mathbf{v}_i = \boldsymbol{\omega} \times \mathbf{r}_i$，则质点 m_i 相对于固定点 O 的角动量可以表示为：$\mathbf{L}_i = m_i[r_i^2 \boldsymbol{\omega} - (\boldsymbol{\omega} \cdot \mathbf{r}_i)\mathbf{r}_i]$，整个刚体相对于 O 点的角动量为：

$$\mathbf{L} = \sum_i \mathbf{L}_i = \sum_i m_i[r_i^2 \boldsymbol{\omega} - (\boldsymbol{\omega} \cdot \mathbf{r}_i)\mathbf{r}_i] \tag{5-3-1}$$

相对于以 O 为原点的固定坐标系 XYZ，角动量的各个分量可以表示为

$$\begin{aligned} L_x &= \sum_i m_i[(x_i^2 + y_i^2 + z_i^2)\omega_x - (\omega_x x_i + \omega_y y_i + \omega_z z_i)x_i] \\ &= \sum_i m_i(y_i^2 + z_i^2)\omega_x - \sum_i (m_i x_i y_i)\omega_y - \sum_i (m_i x_i z_i)\omega_z \equiv I_{xx}\omega_x + I_{xy}\omega_y + I_{xz}\omega_z \\ L_y &= -\sum_i (m_i y_i x_i)\omega_x + \sum_i m_i(z_i^2 + x_i^2)\omega_y - \sum_i (m_i y_i z_i)\omega_z \equiv I_{yx}\omega_x + I_{yy}\omega_y + I_{yz}\omega_z \end{aligned}$$

$$L_z = -\sum_i (m_i z_i x_i)\omega_x - \sum_i (m_i z_i y_i)\omega_y + \sum_i m_i (x_i^2 + y_i^2)\omega_z \equiv I_{zx}\omega_x + I_{zy}\omega_y + I_{zz}\omega_z$$

写成矩阵形式

$$\begin{bmatrix} L_x \\ L_y \\ L_z \end{bmatrix} = \begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{bmatrix} \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix}$$

或写为更加简洁的形式

$$L_i = \sum_j I_{ij}\omega_j \tag{5-3-2}$$

显然，刚体的角动量 $\mathbf{L}$ 与刚体的角速度的方向一般并不相同，式中的参量 I_{ij}

$$\begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{bmatrix} = \begin{bmatrix} \sum_i m_i (y_i^2 + z_i^2) & -\sum_i m_i x_i y_i & -\sum_i m_i x_i z_i \\ -\sum_i m_i y_i x_i & \sum_i m_i (z_i^2 + x_i^2) & -\sum_i m_i y_i z_i \\ -\sum_i m_i z_i x_i & -\sum_i m_i z_i y_i & \sum_i m_i (x_i^2 + y_i^2) \end{bmatrix}$$

称为刚体相对于瞬时转动轴的**转动惯量张量**。张量中的对角分量 I_{ii} 称为刚体的**轴转动惯量**，也常称为刚体相对于转轴的转动惯量，而非对角分量 $I_{ij}(i \neq j)$称为**惯量积**。由以上两式容易看出转动惯量张量是对称张量：$I_{ij} = I_{ji}$，因此总可以通过转动找到一个坐标系 XYZ，在其中所有的非对角分量为零(线性代数中对应于通过线性变换将对称矩阵对角化)，则刚体的角动量约化为：

$$\mathbf{L} = I_x\omega_x\hat{\mathbf{x}} + I_y\omega_y\hat{\mathbf{y}} + I_z\omega_z\hat{\mathbf{z}} \tag{5-3-3}$$

坐标系 XYZ 称为**主轴坐标系**，相应的坐标轴称为刚体的**惯量主轴**，刚体相对于惯量主轴的转动惯量称为**主转动惯量**。

为了帮助读者形成比较直观的转动惯量张量的图像，我们讨论刚体绕固定轴的转动。选择转动轴为固定坐标系 XYZ 的 z 轴，则刚体的转动角速度 $\boldsymbol{\omega} = \omega\mathbf{z}$，由式(5-3-1)将刚体相对于坐标系原点 O (见图 5-3-1a)的角动量表示为：

$$\mathbf{L} = \boldsymbol{\omega}\sum_i m_i\rho_i^2 - \omega\sum_i m_i z_i \boldsymbol{\rho}_i \tag{5-3-4}$$

其中 z_i 和 $\boldsymbol{\rho}_i$ 分别是质点 m_i 的位置矢量 $\mathbf{r}_i$ 沿 z 轴和垂直于 z 轴的平面分量。将上式表示为

$$\mathbf{L} = L_z\hat{\mathbf{z}} + L_\rho\hat{\boldsymbol{\rho}} \tag{5-3-5}$$

角动量 $\mathbf{L}$ 与角速度 $\boldsymbol{\omega}$ 的方向并不一致，其中沿 $\boldsymbol{\omega}$ 方向的分量和 ω 的大小成正比：

$$L_z = I_z\omega \tag{5-3-6}$$

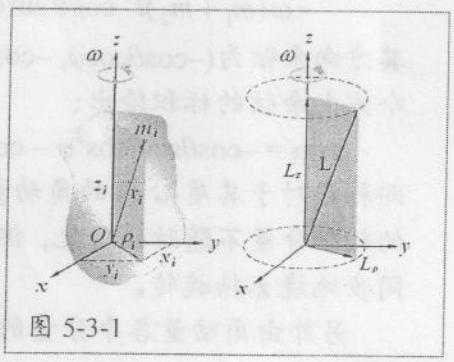


图 5-3-1

比例系数

$$I_z = \sum_i m_i\rho_i^2 \tag{5-3-7}$$

即为刚体关于转轴 z 的**轴转动惯量**，它不随刚体转动发生变化。而垂直于 z 轴平面中的角动量 $\mathbf{L}_\rho$ 含有两个分量

$$\mathbf{L}_\rho = -\omega\sum_i m_i\boldsymbol{\rho}_i z_i = -\omega\hat{\mathbf{x}}\sum_i m_i x_i z_i - \omega\hat{\mathbf{y}}\sum_i m_i y_i z_i = I_{xz}\omega\hat{\mathbf{x}} + I_{yz}\omega\hat{\mathbf{y}} \tag{5-3-8}$$

其中的参量 I_{xz} 和 I_{yz} 即为惯量积。相对于固定坐标系 x-y-z，刚体的质量分布通常随刚体的转动而改变，这意味着 $\mathbf{L}_\rho$ 与刚体的相对方位保持不变，随着刚体绕 z 轴同步旋转。另外从刚体绕定轴转动时的总角动量 $\mathbf{L}$ 的两个分量也可以看出，它们的大小都与时间无关，又都与角速度 $\boldsymbol{\omega}$ 的大小成正比，因而 $\mathbf{L}$ 与 $\boldsymbol{\omega}$ 的夹角 α保持不变，即刚体的角动量 $\mathbf{L}$ 和刚体一起以角速度 $\boldsymbol{\omega}$ 绕z-轴同步旋转(图 5-3-1b)。

如果刚体中的质量分布满足

$$\sum_i m_i z_i \boldsymbol{\rho}_i = 0$$

则 L_ρ 为零，刚体绕定轴转动的角动量只含有轴向分量：

$$\mathbf{L} = L_z \hat{\mathbf{z}} = \boldsymbol{\omega} \sum_i m_i \rho_i^2 = I_z \boldsymbol{\omega}$$

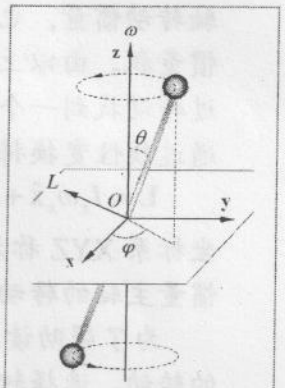
转轴即为刚体的惯量主轴，I_z 即为主转动惯量。

显然，若转轴是刚体的对称轴，刚体中位于$(\boldsymbol{\rho}_i, z_i)$的质点 m_i 在其转动平面中相对于转轴 z 的对称位置$(\boldsymbol{\rho}_j = -\boldsymbol{\rho}_i, z_j = z_i)$一定存在质量相同的质点 $m_j = m_i$，因而

$$m_i z_i \boldsymbol{\rho}_i + m_j z_j \boldsymbol{\rho}_j = 0$$

这就保证了惯量积 $I_{zx} = I_{zy} = 0$。即质量分布具有对称性的刚体的对称轴就是通过其质心的惯量主轴。需要说明的是，对于不存在对称轴的任意形状刚体，也总是存在一组通过其质心互相正交的惯量主轴。

例 5-3-1 一根长度为 $2l$ 的轻杆，两端分别附有质量为 m_1 和 m_2 小球。轻杆与竖直方向成 θ 角的位型绕过其质心 O 的竖直轴以角速度 $\boldsymbol{\omega}$ 转动(见图)，求杆相对于其质心 O 的角动量，以及系统绕 z 轴转动的转动惯量。



例 5-3-1 图

解： 设某一时刻杆两端小球的位置矢量分别为 $\mathbf{r}_1$ 和 $\mathbf{r}_2$，杆相对于直角坐标系的方向余弦为$(\sin\theta\cos\varphi,\ \sin\theta\sin\varphi,\ \cos\theta)$。则杆相对于其质心 O 的角动量为

$$\begin{aligned}
\mathbf{L} &= \mathbf{r}_1 \times m_1(\boldsymbol{\omega}\times\mathbf{r}_1) + \mathbf{r}_2 \times m_2(\boldsymbol{\omega}\times\mathbf{r}_2) \\
&= \boldsymbol{\omega} m_1 r_1^2 - m_1(\boldsymbol{\omega}\cdot\mathbf{r}_1)\mathbf{r}_1 + \boldsymbol{\omega} m_2 r_2^2 - m_2(\boldsymbol{\omega}\cdot\mathbf{r}_2)\mathbf{r}_2 \\
&= [m_1(x_1^2+y_1^2) + m_2(x_2^2+y_2^2)]\omega\hat{\mathbf{z}} \\
&\quad -(m_1 x_1 z_1 + m_2 x_2 z_2)\omega\hat{\mathbf{x}} - (m_1 y_1 z_1 + m_2 y_2 z_2)\omega\hat{\mathbf{y}} \\
&= -\omega(m_1+m_2)l^2\cos\theta\sin\theta\cos\varphi\ \hat{\mathbf{x}} \\
&\quad -\omega(m_1+m_2)l^2\cos\theta\sin\theta\sin\varphi\ \hat{\mathbf{y}} + \omega(m_1+m_2)l^2\sin^2\theta\,\hat{\mathbf{z}}
\end{aligned}$$

其方向余弦为$(-\cos\theta\cos\varphi,\ -\cos\theta\sin\varphi,\ \sin\theta)$。上述两个矢量的夹角 α 的余弦由这两个方向余弦的标积给出：

$$\cos\alpha = -\cos\theta\sin\theta\cos^2\varphi - \cos\theta\sin\theta\sin^2\varphi + \cos\theta\sin\theta = 0$$

即杆相对于其质心 O 的角动量 $\mathbf{L}$ 的方向与杆互相垂直(见图)。由于 θ 是常数，$\mathbf{L}$ 的转轴分量不随时间变化，但由 $\varphi = \omega t$ 可知，$\mathbf{L}$ 的方向随着杆相对于 z-轴的转动同步地绕 z-轴旋转。

另外由角动量各个分量的表达式可以得到系统绕 z 轴转动的转动惯量张量的各个分量：

$$\begin{aligned}
I_z &= m_1(x_1^2+y_1^2) + m_2(x_2^2+y_2^2) = (m_1+m_2)l^2\sin^2\theta \\
I_{z\rho} &= -(m_1 x_1 z_1 + m_2 x_2 z_2)\hat{\mathbf{x}} - (m_1 y_1 z_1 + m_2 y_2 z_2)\hat{\mathbf{y}} \\
&= -(m_1+m_2)l^2\cos\theta\sin\theta\cos\varphi\hat{\mathbf{x}} - (m_1+m_2)l^2\cos\theta\sin\theta\sin\varphi\hat{\mathbf{y}}
\end{aligned}$$

由于系统的质量分布相对于转轴不对称，因而惯量积不为零。而系统的惯量主轴显然是过质量中心，且垂直与轻杆的直线，或轻杆本身。分别对应于 $\theta = \pi/2$ 或 $\theta = 0$。相应的轴转动惯量或惯量积为 $I_z = (m_1+m_2)l^2$ 和 $I_{z\rho} = 0$ 或者 $I_z = I_{z\rho} = 0$。◀

刚体的转动动能

多质点系统的动能是系统中所有单个质点的平动动能之和：

$$K=\sum_i K_i=\frac{1}{2}\sum_i m_i\mathbf{v}_i\cdot\mathbf{v}_i$$

对绕固定点 O 以角速度 $\boldsymbol{\omega}$ 转动的刚体而言，其中质点 m_i 的线速度可表示为 $\mathbf{v}_i=\boldsymbol{\omega}\times\mathbf{r}_i$，则整个刚体的动能可表示为：

$$K=\frac{1}{2}\sum_i m_i\mathbf{v}_i\cdot(\boldsymbol{\omega}\times\mathbf{r}_i)=\frac{1}{2}\boldsymbol{\omega}\cdot\sum_i\mathbf{r}_i\times(m_i\mathbf{v}_i)=\frac{1}{2}\boldsymbol{\omega}\cdot\mathbf{L} \tag{5-3-9}$$

由(5-3-2)式可得

$$K=\frac{1}{2}\sum_{ij}\omega_i I_{ij}\omega_j \tag{5-3-10}$$

在惯量主轴坐标系中，上式简化为

$$K=\frac{1}{2}(I_x\omega_x^2+I_y\omega_y^2+I_z\omega_z^2) \tag{5-3-11}$$

而对于图 5-3-1 所描述的定轴转动，只有和 $\boldsymbol{\omega}$ 同方向的角动量才对绕定轴转动的刚体的动能有贡献，以 $\boldsymbol{\omega}$ 方向为 z 轴，刚体的转动动能为

$$K=\frac{1}{2}I_z\omega^2 \tag{5-3-12}$$

其中 I_z 是刚体相对于 z 轴的轴转动惯量。

5-3-2 转动惯量及其计算

描述刚体转动的两个动力学参量—角动量 $\mathbf{L}$ 和转动动能 K 的表达式中都包含有惯量张量，其中各分量的显式表示分别为

$$I_{xx}=\sum_i m_i(y_i^2+z_i^2)\quad I_{yy}=\sum_i m_i(z_i^2+x_i^2)\quad I_{zz}=\sum_i m_i(x_i^2+y_i^2) \tag{5-3-13a}$$

$$I_{yz}=I_{zy}=-\sum_i m_i y_i z_i\quad I_{zx}=I_{xz}=-\sum_i m_i x_i z_i\quad I_{xy}=I_{yx}=-\sum_i m_i x_i y_i \tag{5-3-13b}$$

刚体绕 z-轴转动时，惯量张量中的分量 $I_{xx}, I_{yy}, I_{zz}, I_{xy}$ 都退化为零。

对于物质连续分布的刚体，可以将质点表示为质量微元 dm，则刚体的惯量张量中的各分量表示为对质量分布的空间积分：

$$I_{xx}=\int dm(y^2+z^2)\quad I_{yy}=\int dm(z^2+x^2)\quad I_{zz}=\int dm(x^2+y^2) \tag{5-3-14a}$$

$$I_{yz}=I_{zy}=-\int yzdm\quad I_{zx}=I_{xz}=-\int xzdm\quad I_{xy}=I_{yx}=-\int xydm \tag{5-3-14b}$$

计算转动惯量

例 5-3-2 计算质量为 m，长度为 l 的细杆相对于垂直于该杆但过质心或者过杆的一个端点的固定轴的轴转动惯量。

解： 将坐标系的 x-轴与细杆重合，以固定转轴作为 z-轴，则细杆相对于垂直于该杆但过杆的一个端点的固定轴的转动惯量为

$$I_{end}=\int_0^l x^2\frac{m}{l}dx=\frac{m}{3}l^2$$

而细杆相对于过杆的质心的固定轴的转动惯量为：

$$I_C = \int_{-l/2}^{l/2} x^2 \frac{m}{l} dx = \frac{m}{12} l^2$$

显然相对于过质心的转动惯量小于过杆的一个端点的固定轴的转动惯量。◀

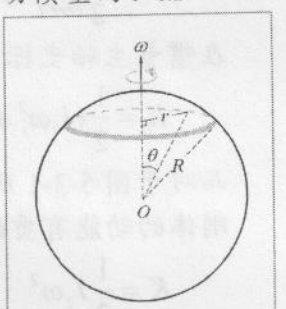
例 5-3-3 计算质量为 m，半径为 R 的圆环相对于垂直圆环平面且过圆心的固定轴的转动惯量，以及质量为 m，半径为 R 的圆球壳和实心圆球相对于其某一直径的轴转动惯量。

解： 相对于垂直圆环平面且过圆心的固定轴的圆环而言，所有质量都与轴的距离相同，因而转动惯量为 $I = mR^2$，对于质量为 m，半径为 R 的球壳绕其直径转动的情形，设 $\sigma = m/4\pi r^2$ 是球壳薄层的质量密度，考虑将球壳分解为一系列围绕直径转轴的圆环(见图)，其中半径为 $r = R\sin\theta$ 的环所贡献的转动惯量为：$dI = r^2 dm = \sigma \cdot 2\pi r R d\theta \cdot r^2 = \frac{1}{2} mR^2 \sin^3\theta d\theta$。对角度积分可得

$$I_{shell} = \frac{m}{2} R^2 \int_0^{\pi} \sin\theta^3 d\theta = \frac{m}{2} R^2 \int_{-1}^{1} (1 - x^2) dx = \frac{2}{3} mR^2$$

对于绕某个直径转动的球体可以分解为诸多的同心球壳，其中半径为 r 的球壳所贡献的转动惯量为：$dI = \frac{2}{3} dm r^2 = \frac{2}{3} \cdot \rho \cdot 4\pi r^2 dr \cdot r^2$，其中 ρ 是球体的质量密度。对半径积分给出

$$I_{sphere} = \frac{2m}{R^3} \int_0^R r^4 dr = \frac{2}{5} mR^2 \quad ◀$$



例 5-3-3 图

实际物体的形状及其质量分布通常比较复杂，转动惯量的计算因而也比较困难。一般情况下，除了可以通过实验直接确定转动惯量，也可以利用转动惯量表达式的函数特性来简化计算。这种特性可以概括为两个很实用的定理：

平行轴定理

设质量为 m 的刚体相对于过刚体质心 C 的固定轴的转动惯量为 I_C，则该刚体相对于相距 C 轴为 d，但与 C-C'轴平行的 O-O' 轴(见图 5-3-2)的轴转动惯量为

$$I = \sum_i m_i r_i^2 = \sum_i m_i (\mathbf{r}_{Ci} + \mathbf{d}) \cdot (\mathbf{r}_{Ci} + \mathbf{d}) = \sum_i m_i r_{Ci}^2 + md^2 + 2\mathbf{d} \cdot \sum_i m_i \mathbf{r}_{Ci}$$

其中第一项正是整个刚体相对于固定轴 C-C'的转动惯量 I_C，第三项中的求和因子则正比于刚体在其质心坐标系中的质心坐标：

$$\mathbf{r}_C = \frac{1}{m} \sum_i m_i \mathbf{r}_{Ci} = 0$$

由此可得所谓的平行轴定理：

$$I = I_C + md^2 \tag{5-3-15}$$

由平行轴定理可知，一个刚体相对与一族相互平行的固定转轴的相应转动惯量中，绕过质心的固定转轴的转动惯量最小($d = 0$)。

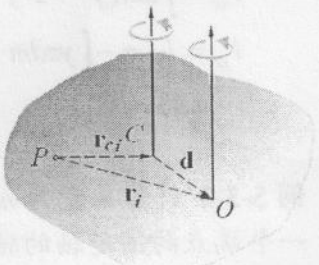


图 5-3-2 分别绕两条平行的轴线转动的轴转动惯量之间的关系。

例 5-3-4 一质量为 m，半径为 R 的圆盘有四个大小相等，相对圆心作对称分布的圆孔(见例图)。圆孔半径为 r，距离圆心的距离为 d。计算该圆盘相对于其对称轴的转动惯量。

解： 首先计算无孔圆盘相对于其对称轴的转动惯量：

$$I_C=\int r'^2dm=\int_0^r \rho(2\pi r'dr')r'^2=\frac{m}{\pi r^2}\cdot 2\pi\cdot\frac{r^4}{4}=\frac{m}{2}r^2$$

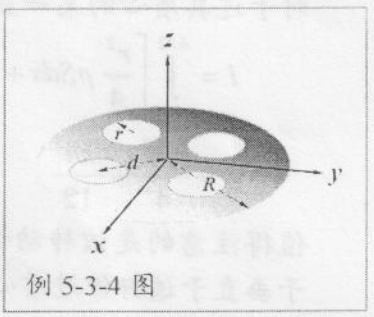
每一个圆盘相对于 z-轴的转动惯量可利用平行轴定理计算：

$$I_{hole}=I_C+md^2=\frac{m}{2}r^2+md^2=\rho\pi r^2\left(\frac{1}{2}r^2+d^2\right)$$

整个系统相对于 z 轴的转动惯量为大圆盘减去四个小圆盘相对于 z 轴的转动惯量：

$$I=\frac{1}{2}(\rho\pi R^2)R^2-4\rho\pi^2\left(\frac{r^2}{2}+d^2\right)=\rho\pi\left(\frac{R^4}{2}-2r^4-4r^2d^2\right)$$

$$=\frac{m}{R^2-4r^2}\left(\frac{R^4}{2}-2r^4-4r^2d^2\right)$$



例 5-3-4 图

垂直轴定理

设一个放置于x-y坐标面中的平面刚体(见图5-3-3)，刚体相对于三个直角坐标轴的转动惯量分别为：

$$I_x=\sum\nolimits_i m_iy_i^2 \qquad I_y=\sum\nolimits_i m_ix_i^2$$

$$I_z=\sum\nolimits_i m_ir_i^2=\sum\nolimits_i m_i(x_i^2+y_i^2)$$

$$=\sum\nolimits_i m_ix_i^2+\sum\nolimits_i m_iy_i^2$$

比较以上三个转动惯量的表达式可以得到三者之间的关系：

$$I_z=I_x+I_y \tag{5-3-16}$$

称为转动惯量的垂直轴定理。

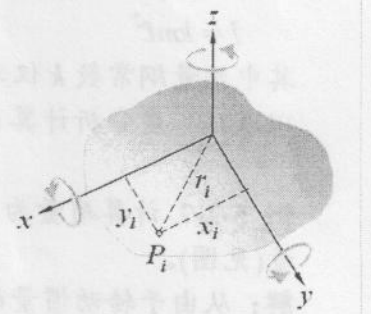


图 5-3-3 平面刚体分别绕三条相互垂直的轴线转动的轴转动惯量之间的关系。

例 5-3-5 计算质量为 m，半径为 R 的圆环相对于其某一直径的转动惯量。

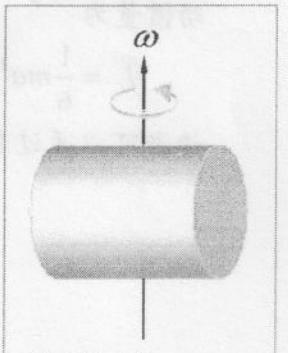
解： 将圆环放置在 x-y 平面上，并使圆心与坐标原点重合。显然圆环相对于 x 轴或 y 轴的转动惯量 I_x、I_y 相等。由于圆环是平面刚体，因此圆环相对于 x 轴、y 轴的转动惯量与相对于 z 轴的转动惯量满足垂直轴定理：

$$I_z=I_x+I_y$$

由此可得圆环相对于某一直径的转动惯量：

$$I_x=I_y=\tfrac{1}{2}I_z=\tfrac{1}{4}mr^2$$

为相对于 z-轴的转动惯量的一半。◀



例 5-3-6 图

例 5-3-6 计算质量为 m，半径为 r，长度为 h 的圆柱体相对于过其质心的横截面某一直径的转动惯量(见图)。

解： 设想将圆柱体分割成一系列的薄圆片，由例 5-3-4 的计算结果可知，每一薄圆片相对于其自身直径的转动惯量为 $I=\frac{1}{4}\rho\ \ Sr^2dr$。利用平行轴定理，圆柱体相对于过其质心的某一直径的转动惯量为

$$I=\int_{-h/2}^{h/2}\left[\frac{r^2}{4}\rho Sdx+x^2\rho Sdx\right]=\frac{r^2}{4}\rho Sh+\rho Sh\frac{h^2}{12}$$

$$=\frac{mr^2}{4}+\frac{mh^2}{12}$$

值得注意的是该转动惯量恰好是圆盘相对于其自身直径的转动惯量与细杆相对于垂直于该杆但过质心的固定轴的转动惯量之和。◀

由刚体转动惯量的量纲[ML^2]可知，刚体对于一个固定转轴的转动惯量形式上可写成

$$I=mR_g^2 \tag{5-3-17}$$

其中 R_g 称为刚体相对于固定转轴的**回转半径**。它相当于整个刚体的质量 m 集中于距离固定转轴 R_g 的位置时的转动惯量。例如将圆柱体相对于其对称轴的转动惯量写成(5-3-17)的形式 $I=\frac{1}{2}mR^2=mR_g^2$，可得圆柱体相对于其对称轴的回转半径为 $R_g=R/\sqrt{2}$。

如果系统具有明显的特征长度 ℓ，则系统的转动惯量形式上可写成

$$I=km\ell^2$$

其中无量纲常数 k 仅与系统的几何位型以及转轴的位置有关。有些情况下我们可以通过标度分析计算系统的转动惯量。

例 5-2-7 计算质量为 m，边长为 a 的立方体相对于过其质心的对称轴的转动惯量(见图)。

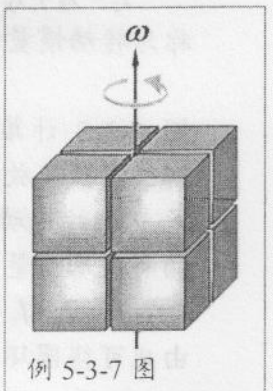
解： 从由于转动惯量的量纲为[m][L]^2，而立方体的特征几何参量为其边长 a，则立方体相对于其对称轴的转动惯量可记为 $I_C=kma^2$。其中 k 为无量纲参数。由于该立方体相对于其对称轴的转动惯量可以表述为 8 个质量为 $\frac{1}{8}m$ 边长为 $\frac{1}{2}a$ 的立方体相对于过自身的一条棱的固定轴的转动惯量之和(见图)。由平行轴定理可得立方体相对于过自身的一条棱的固定轴的转动惯量

$$I_S=kma^2+m(a/\sqrt{2})^2=(k+1/2)ma^2$$

由此可解得 $k=1/6$。故立方体相对于过其质心的对称轴的转动惯量为

$$I_C=\frac{1}{6}ma^2$$

读者可以通过其他方法计算以验证该结果。◀



例 5-3-7 图

5-3-3 刚体转动方程和角动量定理

一个多质点系统相对于空间一固定点的角动量定理为(见式 3-6-15)

$$\boldsymbol{\tau}^{\text{ext}}=\frac{d\mathbf{L}}{dt} \tag{5-3-18}$$

其中 $\boldsymbol{\tau}^{\text{ext}}$ 为系统中所有质点受到的相对于固定点的外力矩总和，$\mathbf{L}$ 为多质点系统相对于固定点的总角动量。对于刚体系统，转动方程(5-3-18)可表示为

$$\boldsymbol{\tau}^{\text{ext}}=\frac{d}{dt}(I\boldsymbol{\omega}) \tag{5-3-19}$$

其中 I 为转动惯量(张量)。由刚体转动方程可以看出，刚体转动惯量在转动时表现出的特性与物体在平动时惯性质量的作用相当：阻碍或抵抗外力矩改变系统转动状态的能力度量。

对于刚体绕固定轴(设为 z 坐标轴)转动的情况，角动量定理中的 $\mathbf{L}$ 和 $\boldsymbol{\tau}^{\text{ext}}$ 都可相应地分解为沿 z 轴和与之垂直的分量：

$$\boldsymbol{\tau}_z=\frac{d\mathbf{L}_z}{dt} \quad \text{和} \quad \boldsymbol{\tau}_\rho=\frac{d\mathbf{L}_\rho}{dt}$$

若转轴恰好是刚体的惯量主轴，即 $\mathbf{L}_\rho=0$，刚体作定轴转动的角动量定理简化为：

$$\tau_z=\frac{dL_z}{dt}=I_z\frac{d\omega}{dt}=I_z\alpha \tag{5-3-20}$$

称为刚体的**定轴转动定律**。其中 α 为刚体相对于固定轴的转动角加速度。

如果刚体转动时所绕的固定转轴不是刚体的惯量主轴，垂直于转轴的角动量分量 $\mathbf{L}_\rho$ 将以相同的角速度随刚体作同步转动。相应的角动量定理的分量为

$$\boldsymbol{\tau}_\rho=\frac{d\mathbf{L}_\rho}{dt}=\left(\sum_i m_i\boldsymbol{\rho}_i z_i\right)\frac{d\boldsymbol{\omega}}{dt}+\left(\frac{d}{dt}\sum_i m_i\boldsymbol{\rho}_i z_i\right)\boldsymbol{\omega}$$

其中第一项与角加速度相联系，而第二项表明只要转轴不是刚体的惯量主轴，即使刚体以恒定角速度 $\boldsymbol{\omega}$ 绕固定轴旋转，也要受力矩作用：

$$\boldsymbol{\tau}'_\rho=\left(\sum_i m_i\frac{d\boldsymbol{\rho}_i}{dt}z_i\right)\boldsymbol{\omega}=\left(\sum_i m_i(\boldsymbol{\omega}\times\boldsymbol{\rho}_i)z_i\right)\boldsymbol{\omega}=\boldsymbol{\omega}\times\left(\sum_i m_i\boldsymbol{\rho}_i z_i\right)\boldsymbol{\omega}=\boldsymbol{\omega}\times\mathbf{L}_\rho$$

方向沿矢量 $\mathbf{L}_\rho$ 的端点所画圆周的切线。

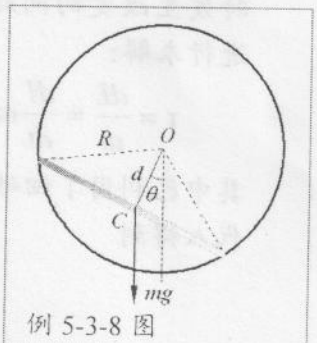
例 5-3-8 描述一根长为 ℓ，质量为 m 的均匀细杆被约束在一半径为 $R\,(2R>\ell)$的竖直光滑圆轨道内的运动。

解： 约束在圆形轨道中滑动的细杆的运动是平动和转动的叠加。但仔细观察可以发现杆上任何一点在运动中相对于圆心 O 的距离保持不变。即杆实际是绕着圆形轨道中心转动，其相应的转动惯量可由平行轴定理得到：

$$I_O=I_C+md^2=\frac{1}{12}m\ell^2+m\left(R^2-\frac{\ell^2}{4}\right)=mR^2\left(1-\frac{\ell^2}{6R^2}\right)$$

而唯一相对于过O点的转轴的力矩有贡献的力是杆的重力矩：

$$\tau_O=-mgd\sin\theta=-mg\sqrt{R^2-(\ell/2)^2}\sin\theta$$



例 5-3-8 图

相对于过O点的转动定律为

$$\frac{d^2\theta}{dt^2}+\frac{3g}{6R^2-\ell^2}\sqrt{4R^2-\ell^2}\sin\theta=0$$

其形式和摆的运动方程完全相同。如果系统的能量很小，以至于细杆只能在其平衡位置附近做往复周期运动，则在$\theta \ll 1$的情况下其角频率ω满足：

$$\omega^2=\frac{3g}{6R^2-\ell^2}\sqrt{4R^2-\ell^2}$$

而当杆的长度缩为一点时，系统的角频率就成为单摆的角频率：$\omega=\sqrt{g/R}$。◀

例 5-3-9 将一张质量为m,半径为R的唱片放到以角速度ω旋转的唱机转台上。设唱片和转台间的滑动摩擦系数为μ。(*a*) 计算唱片受到的总摩擦力矩；(*b*) 计算唱片和转台达到同步转动所需要的时间。

解：唱片和转台之间的摩擦分布在整个接触面上，设想将唱片分解为一系列的同心细圆环，半径为r的圆环受到的相对于圆心的摩擦力矩为：

$$d\tau=r\mu g dm=r\mu g\frac{m}{\pi R^2}(2\pi r)dr=2\mu\frac{m}{R^2}gr^2dr$$

唱片受到的相对于圆心的总摩擦力矩为

$$\tau=\frac{2\mu mg}{R^2}\int_0^R r^2dr=\frac{2}{3}\mu mgR$$

由于转轴恰好是唱片的惯量主轴，唱片的转动方程为

$$\tau=I\frac{d\omega}{dt}=\frac{1}{2}mR^2\frac{d\omega}{dt}$$

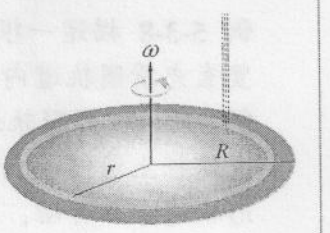
分离变量可得：

$$dt=\frac{2\tau}{mR^2}d\omega$$

代入总摩擦力矩，并对两边积分得：

$$t=\frac{3R}{4\mu g}\int_0^{\omega}d\omega'=\frac{3R}{4\mu g}\omega$$ ◀

例 5-3-10 将一个恒力矩τ作用在一半径为R,质量为M的圆盘上使之绕其自身的对称轴转动。设圆盘上方距转轴为r的一漏斗以均匀的流量k将一缕细砂洒落到圆盘上(见图)。设圆盘的初始角速度为ω_0，计算圆盘的转动角速度随盘上细砂堆积的变化情况。



例 5-3-10 图

解：这是一个涉及物体角动量中转动惯量和角速度同时发生改变的问题，要用原始的动力学方程(5-3-16)进行求解：

$$\tau=\frac{dL}{dt}=\frac{dI}{dt}\omega+I\frac{d\omega}{dt}=\frac{dm}{dt}r^2\omega+I\frac{d\omega}{dm}\frac{dm}{dt}=k\left(r^2\omega+I\frac{d\omega}{dm}\right)$$

其中已利用了细砂均匀流量的关系$dm/dt=k$。将系统的转动惯量$I=\frac{1}{2}MR^2+mr^2$代入得到

$$\frac{\tau}{k}-r^2\omega=\left(\frac{M}{2}R^2+mr^2\right)\frac{d\omega}{dm}$$

分离变量后积分

$$\int_0^{m_o}\frac{dm}{m_oR^2/2r^2+m}=\int_{\omega_o}^{\omega}\frac{r^2d\omega}{\tau/k-r^2\omega}$$

得到

$$\ln\frac{m_oR^2/2r^2+m_o}{m_oR^2/2r^2}=\ln\frac{\tau/k-r^2\omega_o}{\tau/k-r^2\omega}$$

角动量随盘上细砂质量 m 的变化为

$$\omega=\frac{MR^2\omega_o+2m\tau/k}{MR^2+2r^2m}$$

一个特别的例子是外力矩为零的情况：

$$\omega=\frac{MR^2\omega_o}{MR^2+2r^2m}$$

这相当于完全非弹性碰撞的情况。◀

例 5-3-11 求例 5-3-1 中的球-轻杆系统以角速度 ω 匀速转动时转轴所受到的力矩。

解：轻杆系统只有两条对称轴，它们构成一个固定在轻杆上的坐标系 x'-z'，其中 x'轴为轻杆自身，z'轴为过 O 点垂直于轻杆的直线。系统对 O 点的角动量

$$\mathbf{L}=-\omega(m_1+m_2)l^2\cos\theta\sin\theta\cos\varphi\,\hat{\mathbf{x}}$$
$$-\omega(m_1+m_2)l^2\cos\theta\sin\theta\sin\varphi\,\hat{\mathbf{y}}+\omega(m_1+m_2)l^2\sin^2\theta\,\hat{\mathbf{z}}$$

以角速度 $\boldsymbol{\omega}$ 随刚体绕竖直轴同步旋转。$\mathbf{L}$ 的竖直分量 L_z 和水平分量 L_ρ 分别为 $L_{z'}$ 与 $L_{x'}$在转轴方向的投影之和与在水平方向的投影之和：

$$L_z=(m_1+m_2)l^2\omega\sin^2\theta$$
$$L_\rho=\tfrac{1}{2}(m_1+m_2)l^2\omega\sin2\theta$$

由于 $\mathbf{L}$ 的大小与时间无关，只是水平分量随时间在发生变化，相应的力矩为：

$$\tau_\rho=|\boldsymbol{\omega}\times\mathbf{L}_\rho|=L_\rho\omega=\tfrac{1}{2}(m_1+m_2)l^2\omega^2\sin2\theta$$

其方向就是角动量的方向。转轴所受到的力矩是上述力矩的反力矩。因此转轴趋向于向–$\mathbf{F}$ 方向运动，并在旋转时出现抖动，由此形成轴承的额外磨损。只有在转轴为系统的惯量主轴时，转动中不会出现抖动。◀

5-3-5 角动量守恒

当作用在系统上的外力矩为零时，系统的角动量不随时间改变：

$$\boldsymbol{\tau}^{ext}=\frac{d\mathbf{L}}{dt}=0,\quad \mathbf{L}=\text{const.}$$

称为角动量守恒。表面上，系统的角动量守恒非常类似于系统在不受外力作用时的动量守恒，但两者还是存在一些明显的差异。系统的动量是相对于惯性系而言的，而系统的角动量是相对于固定点而言的。在经典力学范畴，系统的质量不会因运动而改变，因此动量守恒通常以系统的速度守恒的形式表现出来。但系统的

质量分布以及系统转动所绕固定点或固定轴的变化都会改变系统的转动惯量，这就意味着角动量守恒会表现出更丰富的效应。

常见的相对于固定轴的物体角动量守恒问题涉及两类系统：一类是孤立刚体，由于孤立刚体的转动惯量(只能绕某一惯量主轴转动)不能改变，角动量守恒的观察效果就是角速度守恒。一个比较典型的例子是我们所观察到的脉冲星(中子星)精准的信号周期。中子星是由恒星因自身引力塌缩形成的密度极高的恒星，其典型的质量相当于2~3个太阳质量，密度与原子核的密度相当，因而其半径约为10公里的数量级。中子星的一个重要特征是存在强度超过10^{12}高斯的极高磁场，当外部电子被磁场俘获，绕磁力线以螺旋轨道运动时将沿着两个磁极向外发射无线电波。因中子星的自转轴和磁轴一般并不重合，沿磁轴向外发射的电磁波束因中子星的自转就像灯塔的灯光一样扫过宇宙(见图5-3-4*a*)。如果地球位于这些宇宙"灯塔"扫过的范围，便会收到脉冲形式的电磁波讯号。孤立中子星的高密度决定了它是一个很好的刚体，在相当长的时间内非常稳定，不会发生显著的形变。因而角动量守恒就表现为其自转速率极为稳定，在地球上观测到极有规律的脉冲信号，称为**灯塔效应**。中子星旋转极快：以太阳现在的速度2 km/s为例，若塌缩到10 km，因角动量守恒，自转周期约为25天~0.001秒[3]。第一个观测到中子星脉冲信号的是1967年由在剑桥大学射电天文台学习的苏珊·乔斯林·贝尔。图5-3-6*b*为当时的观测记录。

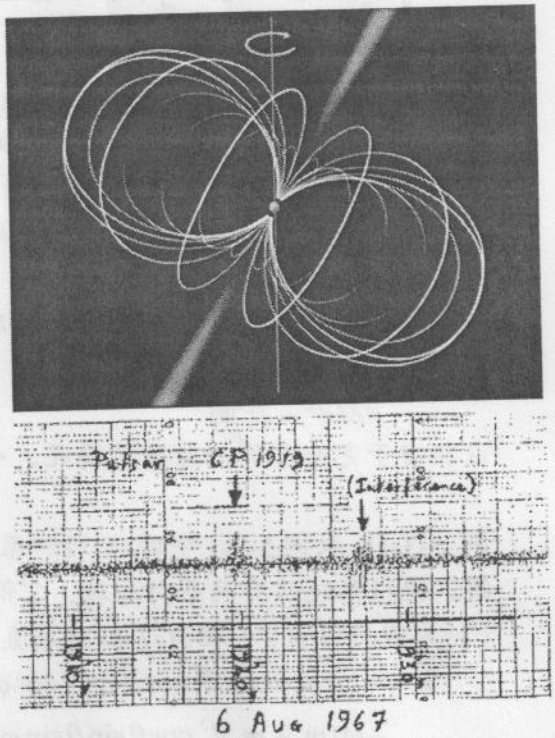


图5-3-4 (*a*) 脉冲射电源运行效果图；(*b*)贝尔观测到的第一个中子星脉冲信号记录

另一类系统是有由相互连接的刚体构成的系统，由于不同部分的相对位型可以变化，从而是整个系统的转动惯量可以改变，这时，角动量的守恒的可观察效应就比较丰富：随着系统构形的变化，系统的角速度必须发生响应的改变以保证整个系统的总角动量不变。例如：花样滑冰表演者往往在开始旋转时张开双臂，进入旋转状态后则并拢双臂以实现更快速的旋转。另一个例子是直升飞机的设计，直升飞机是依靠机顶螺旋桨的快速旋转使飞机顶部的气压减小而使飞机获得升力的。但驱动螺旋桨旋转的力矩属于飞机的内力矩。整个飞机为保持总角

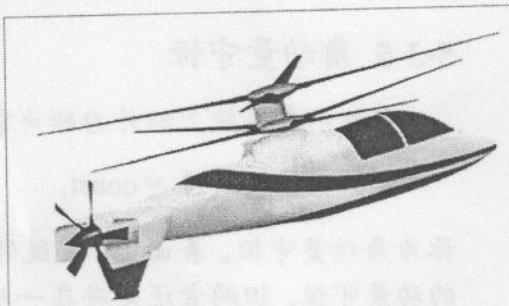


图 5-3-5 直升机的双层顶螺旋桨设计和机尾侧向螺旋桨设计以调节飞机方向。

[3] 为了理解中子星何以具有如此高的自转频率，太阳的5倍。请读者以地球为基准，比较中子星和地球绕各自对称轴自转的转动惯量。

动量守恒会导致与驱动发动机固联的机身发生反向旋转。为避免这种机身旋转一般有两种典型的设计，一是设置两个顶螺旋桨，另一种设计是在机尾加一个侧向旋叶(侧向旋叶的作用类似于机顶螺旋桨，提供侧向的“升力”)。调节两个顶螺旋桨的相对转速或者机尾侧向旋叶的转速就可以维持飞机的取向或改变飞机的航向(见图 5-3-5)。

对于一般的可变形的力学系统，特别是处于绕定点转动的系统，角动量守恒可以表现出更加复杂的形式，我们在 5-3-3 节将讨论过相应的事例。

5-3-5 转动中的功能原理

若刚体受到外力矩的作用而发生了转动状态的变化，可以描述为外力矩对刚体做功改变了刚体的转动动能，按照基本的功能原理，外力 $\mathbf{F}$ 对物体做的功可表示为

$$W=\int \mathbf{F}\cdot d\mathbf{r}=\int \mathbf{F}\cdot(\mathbf{d\theta}\times\mathbf{r})=\int \mathbf{d\theta}\cdot(\mathbf{r}\times\mathbf{F})=\int \boldsymbol{\tau}\cdot\mathbf{d\theta}$$

利用转动方程可将上式改写为

$$\int \boldsymbol{\tau}\cdot\mathbf{d\theta}=\int \frac{d\mathbf{L}}{dt}\cdot\mathbf{d\theta}=\int_{\theta_i}^{\theta_f} I\frac{d\omega}{dt}d\theta=\int_{\omega_i}^{\omega_f} I\omega d\omega=\frac{1}{2}I\omega_f^2-\frac{1}{2}I\omega_i^2 \tag{5-3-21}$$

例 5-2-12 计算例 5-3-8 中唱片从静止到和转台同步旋转一刻，摩擦力矩对唱片作的总功。

解： 利用例 5-3-8 中的计算结果，可得滑动摩擦力矩做的总功为

$$W=\int_0^{\Delta s} Fds=\int_0^{\omega t}\tau d\theta=\frac{2}{3}\mu mgR\omega t=\frac{m}{2}R^2\omega^2=I\omega^2$$

其数值是唱片获得的转动动能$K=\frac{1}{2}I\omega^2=\frac{1}{4}mR^2\omega^2$的两倍，即有一半的功因唱片和转台间在滑动摩擦期间被耗散掉了。◀

思考题

5-2-1 刚体的轴转动惯量是否可以等价地表示为将刚体质量集中在质心上的一个质点的转动惯量？

5-2-2 用手指顶着一根长竹竿或顶一把短尺使之在竖直方位保持平衡，哪一个相对容易一些？为什么？

5-2-3 如何理解一个形状不存在对称轴的刚体也总是存在一组通过其质心互相正交的惯量主轴。

5-2-4 对于定轴转动的刚体，其角动量是否依赖于参考点在轴上的位置？其角动量的轴向分量是否与参考点在轴上的位置有关？

5-2-5 例 5-2-5 中圆柱体相对于过其质心的横截面某一直径的转动惯量可以表示为圆盘相对于其自身直径的转动惯量与细杆相对于垂直于该杆但过质心的固定轴的转动惯量之和，能否将该结果推广到更大的适用范围？

5-2-6 对于一个形状不规则的刚体，可以通过怎样的实验方法测量其绕某一轴线的转动惯量。

5-2-7 有两个形状，大小，质量都相同的圆柱体，其中只有一个中空的，可以通过怎样的简单实验方法确定哪一个是中空的。

5-4 刚体的空间转动

刚体绕固定点的转动(又称为**空间转动**)一般比较复杂，仅在一些很特殊的情形才有解析解。本节将通过对一个典型的定点转动事例—回转仪的转动所表现出的一些与我们通常直觉相悖的行为和动力学效应进行定性或半定量分析，帮助读者建立一些关于定点转动特点的直觉认识和概念。

5-4-1 回转仪的进动

回转仪是一种可以绕自身对称轴自由转动，而且相应的转动惯量非常大的刚体。玩具陀螺是回转仪的一个最简单的例子(图5-4-1)，如果陀螺绕其对称轴快速转动时受外力矩作用，它会表现出不同于我们知觉意识的行为：重力矩 $\boldsymbol{\tau}$ 本应使陀螺向下倾倒，但实际上只是陀螺的对称轴 Oz 绕着竖直轴进动。即绕其对称轴以很大的角速度自转的陀螺似乎并不遵循外力矩的“意志”行事：当外力矩要使它向下(上)运动时，它却表现出向右(左)的运动；而当外力矩要使它向右(左)运动时，它却向下(上)运动。只有在它的自转角速度很小时才会倾倒下来。这种现象叫做**回转效应**。

图 5-4-1 对称陀螺

图5-4-2所示的轮轴系统也是一种很常见的回转仪：一个质量主要集中在轮缘的系统以很高的角速度$\boldsymbol{\omega}$绕轮轴自转。设想开始时轮轴与竖直方向成θ角，而且两端$\boldsymbol{O}$和$\boldsymbol{O'}$都受到支撑而平衡。如果突然撤除$\boldsymbol{O'}$点的支撑，我们从直觉上一般都会觉得系统将发生倾覆。但实际情况却是轮轴系统在保持原有自转的同时开始绕着过支撑点O的竖直轴转动。这种有些出乎我们意料的运动行为称为回转仪的**进动**。

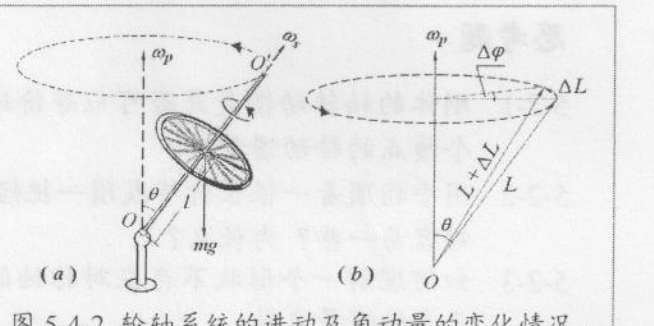


图 5-4-2 轮轴系统的进动及角动量的变化情况。

为了帮助读者理解回转仪发生进动的原因，我们先回顾一下质点在外力作用下所表现出来的类似运动行为。描述质点运动的牛顿方程可以分解为外力与质点的动量共线和垂直的两个分量

$$\mathbf{F}=\frac{dp}{dt}\hat{\mathbf{p}}+p\frac{d\hat{\mathbf{p}}}{dt}=\mathbf{F}_{\parallel}+\mathbf{F}_{\perp} \tag{5-4-1}$$

其中第一项表示与质点动量共线的外力只改变质点动量的大小而不改变动量方

向，第二项表示与质点动量垂直的外力只能改变质点动量方向而不改变质点动量大小，其中 $\hat{\mathbf{p}}$ 随时间的变化率就是 $\mathbf{p}$ 的方向在单位时间内转过的角度，即 $\mathbf{p}$ 的转动角速度 $\boldsymbol{\omega}$(见 2-6 节)

$$\mathbf{F}_{\perp} = p\frac{d\hat{\mathbf{p}}}{dt} = p\boldsymbol{\omega}\times\hat{\mathbf{p}} = \boldsymbol{\omega}\times\mathbf{p} \tag{5-4-2}$$

式(5-4-2)就是描述质点受纯粹向心力作用时的运动方程。

刚体系统作定点转动的一般转动定律可以类似地表述为

$$\boldsymbol{\tau} = \frac{d\mathbf{L}}{dt} = \frac{dL}{dt}\hat{\mathbf{L}} + L\frac{d\hat{\mathbf{L}}}{dt} = \boldsymbol{\tau}_{\parallel} + \boldsymbol{\tau}_{\perp} \tag{5-4-3}$$

其中第一项表示与角动量平行的外力矩只改变系统角动量的大小而不改变其方向，第二项表示与角动量垂直的外力矩只能改变系统角动量方向而不改变其大小，其中 $\hat{\mathbf{L}}$ 随时间的变化率就是刚体角动量方向的变化率，即 $\mathbf{L}$ 的转动角速度 $\boldsymbol{\omega}_p$，因而有

$$\boldsymbol{\tau}_{\perp} = L\frac{d\hat{\mathbf{L}}}{dt} = L(\boldsymbol{\omega}_p\times\hat{\mathbf{L}}) = \boldsymbol{\omega}_p\times\mathbf{L} \tag{5-4-4}$$

对于图5-4-1(*a*)所示的轮轴系统而言，撤去***O'***点的支撑后，作用在系统上的力就只有作用在轮轴中心的重力和枢轴点*O*处的支撑力，而相对于*O*点的力矩就只是重力矩：

$$\boldsymbol{\tau} = \mathbf{l}\times m\mathbf{g} = lmg\sin\theta\,\hat{\mathbf{y}} \tag{5-4-5}$$

其中重力矩方向 $\hat{\mathbf{y}}$ 垂直纸面向里，对应于轮轴与竖直*z*轴构成页面的瞬间。轮轴与*z*轴间的夹角θ也是轮轴的自转角动量$\mathbf{L}_s$与*z*轴间的夹角。为确定$\boldsymbol{\omega}_p$的方向，我们考虑角动量受重力矩作用时的变化情况。系统角动量在Δt时段后变为$\mathbf{L} = \mathbf{L}_s + \boldsymbol{\tau}\Delta t$，因重力矩方向与自转角动量$\mathbf{L}_s$垂直，不改变角动量的大小，角动量$\mathbf{L}$围绕着竖直轴旋转形成一个圆锥面。而整个轮轴系统也就在该圆锥面内以该角速度旋转，这种出现在自转侧向的转动称为**进动**。

由图5-4-1b可以看出，角动量在Δt时间内的变化为$|\Delta\mathbf{L}|=\Delta\varphi(L_s\sin\theta)$，其中$L_s\sin\theta$为$\mathbf{L}$的端点勾勒出的圆轨道的半径，而$\Delta\varphi$为$\Delta t$时间的内$\mathbf{L}$的端点在该圆轨道内转过的角度。由此可得

$$\left|\frac{d\mathbf{L}}{dt}\right| = \frac{d\varphi}{dt}L_s\sin\theta \tag{5-4-6}$$

其中 $d\varphi/dt$ 就是进动角速度 $\boldsymbol{\omega}_p$ 的大小。上述方程可以写成更紧凑的矢量形式：

$$\frac{d\mathbf{L}}{dt} = \boldsymbol{\omega}_p\times\mathbf{L} \tag{5-4-7}$$

与方程(5-4-4)的形式一致。由此可以得到进动角速度的大小为

$$\omega_p = \frac{\tau}{L\sin\theta} = \frac{mgl}{L} = \frac{mgl}{I_s\omega_s} \tag{5-4-8}$$

其中 I_s 为系统绕自转轴的转动惯量。由于进动角速度越小意味着系统运动状态的变化越小，系统也越稳定，而上式中自转角速度和转动惯量的大小对回转仪的进动角速度的影响与我们对转动惯量大，自转速度高的系统不容易改变其运动状态的直觉是一致的。

回转仪可以通过转动惯性来维持其转动状态的特性有许多实际的应用。图

5-4-3 是一个安装在一个支架上以消除所有可能的外力矩的轮轴系统，这类装置称为**陀螺仪**[4]。陀螺仪有两个平衡环，可以环绕三个互相垂直的轴自由转动。当其中的转子高速旋转时，它所具有的很大的转动惯量可抵抗外部对转子轴向的影响，使转子的转轴保持固定的指向。若外界作用或力矩施加在转子的转轴上，转轴会顺着转子旋转向前90°垂直施力方向运动—进动。处于高速旋转状态中的转子能维持自旋轴的方向这一特性可用于导航、定位等系统中。而回转仪在重力矩作用下绕竖直轴进动这一现象可用以分辨竖直程度，如测量一个平面相对于水平面的倾斜度。

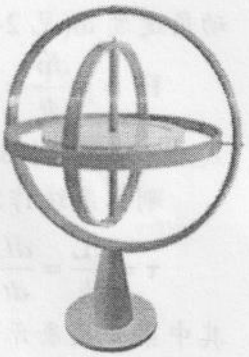


图 5-4-3 陀螺仪

图 5-4-4 给出的是一个利用回转仪原理制作的磁漂浮器的运行情况。将两块永磁体中的一块放在水平面上，另一块则放置在一陀螺仪内部。陀螺竖立时两块磁体的磁极方向相反，相互间的排斥作用使陀螺能悬浮在空中(具体高度由磁力和重力的平衡点确定)。但陀螺的这种姿态是不稳定的，因为两块磁体的磁作用势能处于最大值，陀螺很容易倾覆，并从空中掉落下来，但如果使陀螺高速自转起来，重力矩和磁力矩就无法使陀螺倾覆，即使对它施加小的干扰也只能使陀螺绕竖直轴进动，从而使陀螺能稳定地悬浮在空中。

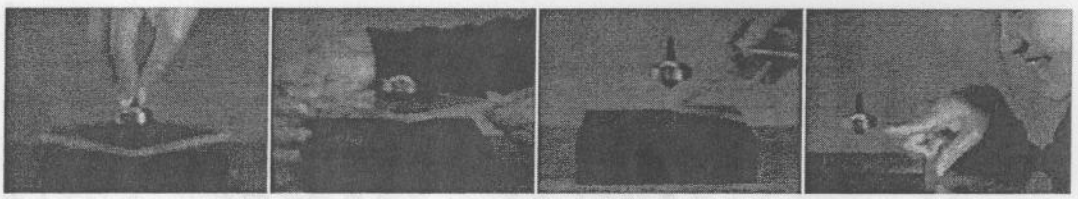


图 5-4-4 磁漂浮器

5-4-2 回转仪的章动-动平衡

上面讨论的轮轴系统在(与自转角动量不平行)重力矩作用下处于平稳、匀速进动的状态事实上只是轮轴高速自转时的一个很特殊的情形。通常自转轮轴在进动过程中，其质心还会围绕着一定的高度上下振荡，从而使轮轴上端点在进动过程中描绘出图 5-4-5 中的摆线。这类特殊的运动形式称为**章动**。实验发现，轮轴自转速度很快时，章动幅度很小，摆动频率也比较快，而且支撑点 O 处存在的摩擦会很快将章动阻尼掉，轮轴系统通常可以很快进入平稳、匀速的进动状态。如果轮轴的自转速度不是很大，轮轴 O' 端的下落的幅度就会比较大，也容易观察到。事实上章动和进动是互相关联的，由于对这种运动过程的精确描

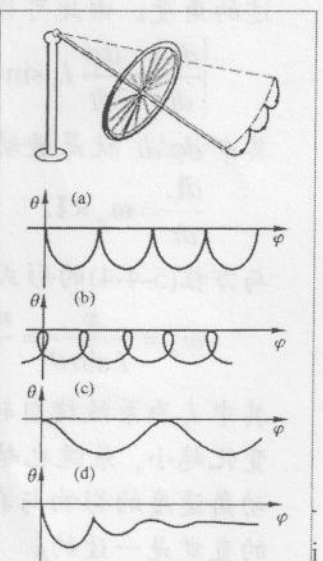


图 5-4-5 轮轴系统在进动时伴随出现的不同形式的章动。

[4] 1850 年法国的物理学家傅科在研究地球自转时首先发现高速转
惯性作用永远指向一固定方向，他用希腊字 gyro (旋转)和 skopein
一字来命名这种仪表。

述比较复杂，这里只能做一些定性的分析。

设想开始时轮轴系统的(自转)轴处于水平状态，轮轴在 O 和 O'端都受到支撑时，系统受到的总外力矩为零，维持着原来的匀速自转状态。突然撤除 O'端支撑的很短时间内，系统质心将如我们的直觉所期待的那样下坠，这是系统在重力和支撑力因暂时大小不相等而失去平衡的结果，也可以理解为出现了相对于支点 O 的重力矩导致系统向重力矩方向转动。由于系统下坠是以绕垂直轮轴且过枢轴支点的水平轴转动的形式出现的，而该下落的角动量增量并不存在该方向的力矩与之对应，要维持垂直方向的角动量不变，系统就必须同时出现一个绕竖直轴的转动以抵消这个角动量，这就是前面所讨论的进动。

由于轮轴系统下倾时伴随着系统发生绕竖直轴的进动，这意味着存在一个竖直向上的回转力矩(称为章动回转力矩)，而进动又产生与重力矩方向相反的回转力矩(称为进动回转力矩)。由于两种回转力矩都与相应的角速度成正比($\boldsymbol{\tau} = \boldsymbol{\omega} \times \mathbf{L}$)，开始时下倾角速度较小，进动角速度也较小，重力矩大于进动回转力矩，系统加速下坠。但随着下坠角速度增大，章动回转力矩也随之增加，这导致进动加速，而随着进动角速度的增大，进动回转力矩也增大。当此回转力矩与重力矩相抵时，回转仪质心下落到平衡高度 h，下倾角速度达到最大，但系统由于惯性继续下坠，同时使进动角速度继续增大，但因进动回转力矩超过重力矩而使下倾运动逐渐减慢而停止，这时进动角速度达到最大。接着因进动回转力矩大于重力矩而使系统上升，同时产生方向竖直向下的章动回转力矩，使进动减慢。这又使进动回转力矩减小，直至与重力矩相等，但系统由于惯性仍将继续上升，同时使进动继续减慢，进动回转力矩也继续减小，最终因重力矩超过进动回转力矩而使上升运动停止，进动角速度也同时减为零，完成一次完整的章动。然后陀螺在重力矩作用下再次下倾并重新发生进动，如此循环往返。

当陀螺自转轴在不同方位以不同的初角速度释放时，形成的章动形式也不同，轮轴端点勾画出的轨迹也不同。如果以进动角 φ 为横轴，章动角 θ 为纵轴，章动所对应的陀螺轴端的一些典型运动轨迹如图 5-4-5 所描绘出的曲线。其中(*d*)是与(*a*)对应但存在有阻尼时曲线。

5-4-3 空间转动的实际事例

地球转动

通常我们想到地球的转动可能只是地球的自转(地球绕日轨道运动属于平动)，但地球实际上就象一个巨大的陀螺，在太阳，月亮和其它行星及其运动的作用下也表现出进动和章动的转动形态。

由于长时间的自转，地球的形状近似于一个以自转轴(**地轴**)为对称轴的扁椭球(其在赤道处的半径比两极处的半径大近 21 公里，相当于地球半径的 1/300)，而地轴与黄道面的垂直轴(**黄道轴**)有 23°26′的夹角[5]，另外，月球绕地球公转的轨道(**白道**)平面，与黄道面和天赤道面都不重合，它与黄道面的夹角为 5°9′。太

[5] 太阳绕地球公转的轨道平面与天球相交的大圆称为**黄道**，垂直于黄道面的轴称为**黄道轴**，这个夹角也是地球赤道面与黄道面的夹角，因而常称为**黄赤交角**

阳，月亮和其它行星对地球赤道带突出部分的不同引力大小形成相对于地球中心的净力矩(见图 5-4-6-b，其中只描绘了太阳潮汐力形成的力矩)，由于地球自转的惯性作用，外力矩并不能使地轴与黄道轴的夹角发生变化，而是使地轴绕黄道轴由东向西进动(同地球自转和公转方向相反)，进动周期为 25,800 年，相应的进动

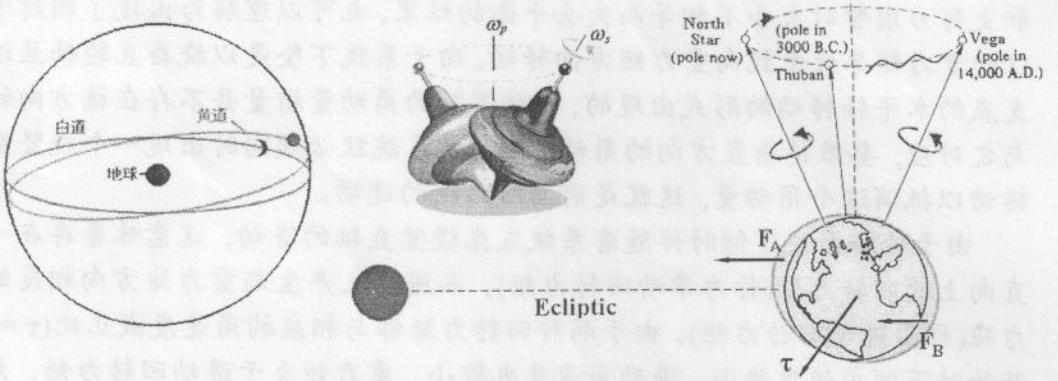


图 5-4-6 太阳对地球赤道带突出部分的不同引力所形成的力矩 τ (其中靠近太阳的部分受到的引力大于地球背面的突出部分受到的太阳引力)

速度为 50.2564 角秒[6]。地轴的进动造成地球的南北两极的空间指向发生相应的改变，例如，公元前 3000 年北极星是天龙座 α 星，目前的北极星在小熊座 α 星附近，到了公元 7000 年，移到仙王座 α 星附近，而到公元 14000 年，织女星将成为北极星(见图 5-4-6)。天极绕黄极的顺时针方向进动导致恒星在天球上的坐标位置(赤经，赤纬)有规律变化，因此天文学家用望远镜搜寻某一颗恒星时必须将其位置从星表历元按照一定的公式改正到观测的历元。

另外，地轴进动使地球的赤道面在空间的倾斜方向发生改变，并引起天赤道相应的变化：天赤道与黄道的交点(春分点和秋分点)在黄道上相应地自东向西移动，与地球公转方向相反，每年移动角度为 50".29[7]。因此年的长度以春分点为参考点周期单位(称为**回归年**)要比以恒定不动的恒星为参考点的周期单位(称为**恒星年**)短 20 分 24 秒。该天文现象称为日月**岁差**[8]。

还需要注意的是，地球在进动过程中，太阳对地球的引力矩作用表现出周期性变化[9]，而月亮的运动也受到太阳和其它行星引力的干扰。因此，月球轨道面在力的作用下发生类似于地球岁差现象的升交点退行。这一现象反过来又影响到地轴的运动，导致天极绕黄极的运行轨迹出现波动或者摆动，即引起地球发生相应周期的章动。具体而言，月球绕地球公转的轨道(平面与天球相交的大圆称为白道)与黄道相交于两点(交角在 4°57′~5°19′ 之间作周期变化，周期约 173 天)。月球沿着白道运行时在黄道两旁穿梭使月球与太阳到地球的距离不断发生变化，

---

[6] 1976 年国际天文学会通过对于标准历元 2000 年的新值为 50.290966

[7] 由于地轴进动而造成的天极、春分点的移动角度相对来讲是很小的，在较长的时间里不会有很大的移动。因而仍然可以说天极和春分点在天球上的位置不变，恒星的赤经、赤纬和黄经也可以近似认为不变，以此建立的星表、星图仍是可以长期使用的。

[8] 牛顿第一个指出产生岁差的原因是太阳和月球对地球赤道隆起部分的吸引。德国天文学家贝塞耳在 1818 年首次得出日月岁差为 5,034".05(历元 1755.0)，今值为 5,025".64。

[9] 因月亮绕地球和地球绕太阳运动的轨道都是椭圆，旋转角速度不是恒定的，它们对地球赤道面的倾角(黄赤交角和白赤交角)也都不是确定不变的

而太阳对月球的引力使两个交点的连线沿黄道与月球运行的相反方向向西移动(称为**交点退行**)，速度为 19°21′/年，相应的周期为 18.6 年。则月球对地球的引力作用也相应地表现出相同周期的变化。即天极(真天极)在绕黄极进动的同时，还围绕其平均位置作周期为 18.6 年的章动[10]。此外，太阳系内其他行星对地球的吸引也会引起黄道面位置的不断变化，这不仅使黄赤交角改变，还使春分点沿赤道产生一个方向与日月岁差相反的微小位移—**行星岁差**，它使春分点沿赤道每年东进约 0".13。转轨道与地球公转轨道相交之处)每年相对于黄道星座西行大约 20 度。月球的轨道面以 18.6 年为周期向西转动(每年 19.4 度)。

事实上，地轴在绕黄道轴进动过程中还伴随有一种周期约为 18.6 年的微小摆动(真天极绕平天极在 18.6 年内描绘出一个小椭圆—章动椭圆，其长半径值为 9".21，称为章动常数[11]。

其它行星的引力也导致黄道面(即地月系重心绕日旋转的轨道面)本身发生错动，这使黄极在天上的位置发生变化，有时接近天极，有时又远离天极，其位移轨迹是一条复杂的螺线。而站在黄极的角度观察，这种改变使得天极绕黄极的轨迹不再是一个圆，而是一个接近圆形的曲线。即天极不仅围绕黄极旋转，而且不断地改变它们的距离，这个距离实际上就是黄赤交角。因此，行星岁差直接导致黄赤交角的改变，这也是回归线漂移的原因之一。

需要说明的是岁差和章动是同源的，只是由于地球自转轴的运动能够分拆为两个周期相差很大的部分，因此看来好像是两个不同的现象。但实际上并不能将岁差纯粹归咎于太阳，而章动则纯粹归咎于月球。

空间转动的实际应用

回转效应在技术上有很多应用。例如由平直枪管射出的子弹在飞行过程中由于受到空气的力矩作用往往会乱翻跟头，这不仅使子弹偏离了原来的方向，大大降低了命中率，而且缩短了飞行距离。以后的枪管内都甚至狙击子弹外表面都刻有螺旋式的来复线，这样射出的枪弹就成为一个绕着自身对称轴快速旋转的回转仪。因旋转的物体不容易改变旋转轴线的方向，枪弹就不会在飞行过程中翻跟头，弹道稳定，从而提高射击精度。子弹在飞行途中的平衡不是静态下的平衡,而是高速飞行时的平衡。

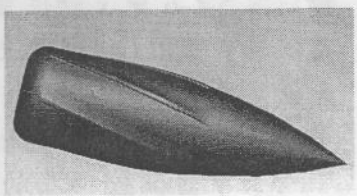


图 5-4-7

这种类似的效应也广泛应用于自动导航。鱼雷里通常装有用电动机带动的回转仪。因风浪或其它原因鱼雷改变方向时，由于回转仪未受到无外力矩作用，其转动轴方向因惯性而保持不变，则回转仪的转动轴和鱼雷的纵轴之间出现偏差，

[10] 地球章动现象首先是由英国天文学家布拉德雷(James Bradley)在 1748 年分析 1727-1747 的 20 年间恒星位置的观测资料后发现的。地球章动周期约为 19a，我国古代历法以 19a 为一章，章动一词或源于此。

[11] 1977 年国际天文学会决定用理论方法求出非刚体地球的章动常数，对于标准历元 2000 年章动常数值为 9".2044，1979 年国际天文学会决定于 1984 年正式采用)。

由此可以触发某些机械装置改变舵的方向，使鱼雷回复到原有航向。火箭也同样通过回转仪来保持航向，并通过改变喷气方向来校正航向。其他如回转罗盘、列车的稳定问题以及轮船上装的回转稳定器等，也都要用到回转效应。需要说明的是，在快速转动的部件上，回转效应所产生的力矩(回转力矩)会给轴承施加很大的附加压力。例如轮船上的巨大涡轮在水平面内绕着对称轴以很大角速度 ω_1 转动，当轮船以角速度 ω_2 转弯时，相当于涡轮机的轴以角速度 ω_2 进动，故轮船一般不应很快地转弯，以免过大的附加压力毁坏涡轮机的轴承。

5-4-4 定点转动及其演化的一般描述—欧勒方程*

刚体作定点转动时，选取固定在空间不动的**固定坐标系** S: O-XYZ 和固定在刚体上随刚体一起转动的**本体坐标系** S': O-xyz(见图 5-4-6a)，S 和 S'的原点都取在定点 O。设初始时刻 S'和 S 重合，S'绕 z 轴逆时针旋转 φ，x 轴转到 Ox'(即 ON)的角位置，y 轴转到 Oy'的角位置(见图 5-4-6b)。再令 S'绕 Ox'转动 θ(见图 5-4-6c)，则 z 轴和 Z 轴分开，S'三个轴现在是 x', y''和 z。z 轴和 Z 轴间的夹角是 θ，xy''平面和 XY 平面间的夹角也是 θ。最后令 S'绕着 z 轴转动 ψ，则 Ox'轴与 ON 分开，

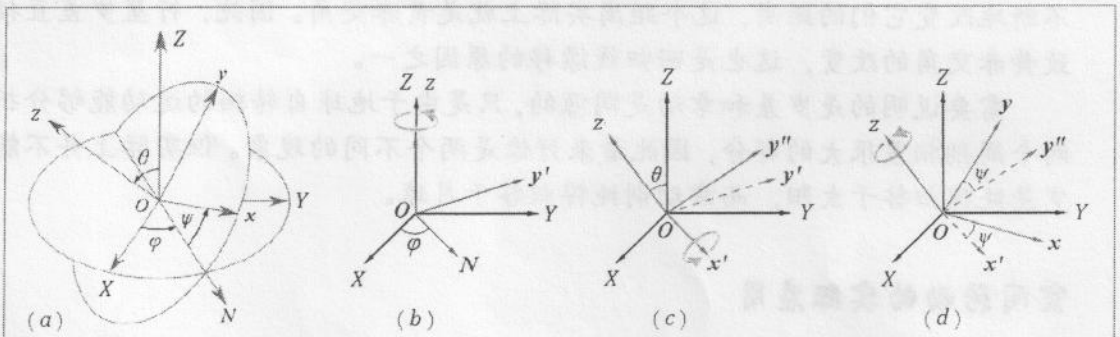


图 5-4-6 欧勒角

转到图 5-4-6 中 x 轴位置，Oy''转到 y 轴的位置(见图 5-4-6d)。x 轴和 Ox'间的夹角以及 y 轴和 Oy''间的夹角都是 ψ。则 S'系转到了所需要的位置。刚体的各种可能位置的**欧勒角**范围为：

$$0 \le \theta \le \pi, \quad 0 \le \varphi \le 2\pi, \quad 0 \le \psi \le 2\pi$$

XY 平面和 xy 平面的交线 ON 称为**节线**，节线和 X 轴之间的夹角 φ 称为**进动角**。节线和 x 轴之间的夹角 ψ 称为**自转角**(ψ 为绕 z 轴转动的角度)。而 z 轴和 Z 轴间的夹角 θ 称为**章动角**[12]。z 轴和 ON 垂直，其方向由 θ 与 φ 确定。这三个独立的角参量(θ, φ, ψ)统称为**欧勒角**。欧勒角(θ, φ, ψ)可确定 S'相对于 S 的位置，因 S'固定在刚体上，由此也就确定了刚体的位置。

欧勒运动学方程

设刚体绕过定点 O 的某一轴线以角速度 $\boldsymbol{\omega}$ 转动，$\boldsymbol{\omega}$ 投影到 S'系各坐标轴上的分量为

$$\dot{\boldsymbol{\omega}} = \omega_x \hat{\mathbf{x}} + \omega_y \hat{\mathbf{y}} + \omega_z \hat{\mathbf{z}} \tag{5-4-9}$$

[12] 这些名称都来自陀螺的转动。

因 $\boldsymbol{\omega}$ 可理解为绕 Z 轴的角速度 $\dot{\varphi}$，绕 ON 轴的角速度 $\dot{\theta}$ 及绕 z 轴的角速度 $\dot{\psi}$ 三者的矢量和，其中 $\dot{\psi}$ 沿着 z 轴，没有 x 轴和 y 轴分量。$\dot{\theta}$ 沿着 ON 轴，而 ON 和 z 轴垂直，$\dot{\theta}$ 没有 z 轴分量，在 x, y 轴上的分量分别为 $\dot{\theta}\cos\psi$ 和 $-\dot{\theta}\sin\psi$。$\dot{\varphi}$ 沿 Z 轴(见图 5-4-4)，其 z 轴分量和 Oy''(因 z 轴, Z 轴及 Oy''在同一平面内，z 轴与 Oy'' 垂直)方向的分量分别为 $\dot{\varphi}\cos\theta$ 和 $\dot{\varphi}\sin\theta$。再将 Oy''上的 $\dot{\varphi}\sin\theta$ 分解到 x 轴和 y 轴上。相应的分量分别是 $\dot{\varphi}\sin\theta\sin\psi$ 和 $\dot{\varphi}\sin\theta\cos\psi$。

比较 $\dot{\theta},\dot{\varphi},\dot{\psi}$ 投影到 x, y, z 轴上的各个分量和(5-4-9)式，可以得出用欧勒角及其对时间的微商来表示角速度沿转动坐标轴 x, y, z 三个分量的表达式：

$$\omega_x = \dot{\varphi}\sin\theta\sin\psi + \dot{\theta}\cos\psi$$

$$\omega_y = \dot{\varphi}\sin\theta\cos\psi - \dot{\theta}\sin\psi$$

$$\omega_z = \dot{\varphi}\cos\theta + \dot{\psi} \tag{5-4-10}$$

称为**欧勒运动学方程**。刚体运动学中主要讨论在已知(θ, φ, ψ)和时间 t 的关系的条件下，利用上式计算角速度 $\boldsymbol{\omega}$ 在 $\boldsymbol{S'}$系上的三个分量；也可通过已知任一时刻 t 的 $\boldsymbol{\omega}$ 的各分量，求解 θ, φ, ψ 和时间的关系，从而确定刚体的运动。

欧勒动力学方程

设 $\mathbf{L}$ 是刚体绕定点 O 转动时的角动量，$\boldsymbol{\tau}$ 为外力对 O 点的**主力矩**。则描述绕定点转动刚体的动力学方程是：

$$\frac{d\mathbf{L}}{dt} = \boldsymbol{\tau}$$

相对于固定在刚体上的坐标系 $\boldsymbol{S'}$系，绕 O 点转动的各转动惯量分量$\{I_{ij}\}$都是常数。若选用刚体的惯量主轴为 $\boldsymbol{S'}$系坐标轴，各惯量积都为零。角动量 $\mathbf{L}$ 进一步简化为：

$L_x = I_1\omega_x$，$L_y = I_2\omega_y$，$L_z = I_3\omega_z$，

其中 I_1, I_2, I_3 是刚体在 O 点上的主转动惯量。因坐标轴的转动角速度就是刚体的角速度 $\boldsymbol{\omega}$，若将 $\mathbf{L}$ 和转动坐标系 $\boldsymbol{S'}$固定在一起，$\mathbf{L}$ 对时间的微商为 $\boldsymbol{\omega}\times\mathbf{L}$。通常 $\mathbf{L}$ 相对于 O-xyz 也有变化，其每一个分量也都是时间的函数，把 $\mathbf{L}$ 写为

$$\mathbf{L} = L_x\hat{\mathbf{x}} + L_y\hat{\mathbf{y}} + L_z\hat{\mathbf{z}} \tag{5-4-11}$$

则 $\mathbf{L}$ 对时间的微商是上述两部分之和：

$$\frac{d\mathbf{L}}{dt} = \dot{L}_x\hat{\mathbf{x}} + \dot{L}_y\hat{\mathbf{y}} + \dot{L}_z\hat{\mathbf{z}} + \boldsymbol{\omega}\times\mathbf{L} \tag{5-4-12}$$

因为我们是从固定坐标系 $\boldsymbol{S}$ 系来观测矢量对时间的微商，式(5-4-12)中的 $\dot{\mathbf{L}}$ 反映了固定坐标系中看到的 $\mathbf{L}$ 的时间变化率。则式(5-4-12)写为

$$I_1\dot{\omega}_x - (I_2 - I_3)\omega_y\omega_z = \tau_x$$

$$I_2\dot{\omega}_y - (I_3 - I_1)\omega_z\omega_x = \tau_y \tag{5-4-13}$$

$$I_3\dot{\omega}_z = \tau_z$$

上面的方程组称为**欧勒动力学方程**[13]。利用欧勒运动学方程消去欧勒动力学方

[13] 欧勒方程的推导过程中用了两次简化：(1)用固定在刚体上的活动坐标系使六个惯量系数 I_{ij} 都成为常数；(2)取 O 点上的惯量主轴为动坐标系的坐标轴以消去惯量积 I_{xy}, I_{yz} 和 I_{zx}，使

程中的 ω_x、ω_y 和 ω_z，可得到欧勒角 θ, φ, ψ 的三个演化方程，求解出 θ, φ, ψ 随时间的变化也就确定了刚体的运动。

力矩 $\boldsymbol{\tau}$ 通常是 $t, \theta, \varphi, \psi, \omega_x, \omega_y$ 和 ω_z 的函数，因而**欧勒动力学方程**一般很复杂，甚至无法求解。下面我们只讨论重力矩作用下陀螺的定点转动。

取固定点为 S 及 S' 系的原点 O，陀螺的对称轴为 S' 系的 z 轴，设陀螺的重心 G 与原点的距离为 l，作用在刚体上的重力 $-mg\hat{\mathbf{Z}}$ 对 O 点的力矩为

$$\boldsymbol{\tau} = l\hat{\mathbf{z}} \times (-mg\hat{\mathbf{Z}}) = -mgl(\hat{\mathbf{z}} \times \hat{\mathbf{Z}})$$

因 $\hat{\mathbf{Z}}$ 与进动角速度方向相同，因而

$$\hat{\mathbf{Z}} = \sin\theta\sin\psi\hat{\mathbf{x}} + \sin\theta\cos\psi\hat{\mathbf{y}} + \cos\theta\hat{\mathbf{z}}$$

而重力矩 $\boldsymbol{\tau}$ 在 S' 系中的表达式为

$$\boldsymbol{\tau} = mgl(\sin\theta\cos\psi\hat{\mathbf{x}} - \sin\theta\sin\psi\hat{\mathbf{y}}) \tag{5-4-14}$$

由式(5-4-13)得对称陀螺的运动方程

$$\begin{aligned} I_1\dot{\omega}_x - (I_1 - I_3)\omega_y\omega_z &= mg\ell\sin\theta\cos\psi \\ I_1\dot{\omega}_y - (I_3 - I_1)\omega_z\omega_x &= -mg\ell\sin\theta\sin\psi \\ I_3\dot{\omega}_z &= 0 \end{aligned} \tag{5-4-15}$$

对第三式积分得到

$$\omega_z = \dot{\varphi}\cos\theta + \dot{\psi} = s(\text{const.}) \tag{5-4-16}$$

即陀螺沿其对称轴的角速度分量不变。用 ω_x 乘方程组(5-4-15)中的第一式，ω_y 乘第二式，ω_z 乘第三式，相加后再将式(5-4-10)中的 ω_x 和 ω_y 的表示式代入所得方程的右端：

$$I_1\dot{\omega}_x\omega_x + I_1\dot{\omega}_y\omega_y + I_3\dot{\omega}_z\omega_z = mgl\dot{\theta}\sin\theta$$

对上式积分，得到系统的能量守恒方程

$$\frac{1}{2}(I_1\omega_x^2 + I_1\omega_y^2 + I_3\omega_z^2) + mgl\cos\theta = E \tag{5-4-17}$$

总能量 E 是常数。将式(5-4-10)中的 ω_x 和 ω_y 的表示式及式(5-4-16)中的 ω_z 的值代入(5-4-17)，得到

$$I_1(\dot{\theta}^2 + \dot{\varphi}^2\sin^2\theta) + I_3s^2 = 2(E - mgl\cos\theta) \tag{5-4-18}$$

因重力对 Z 轴的力矩为零，因而该方向上的角动量分量是常数：

$$\mathbf{L}\cdot\hat{\mathbf{Z}} = \alpha = \text{常数}$$

代入 $\mathbf{L}$ 及 $\hat{\mathbf{Z}}$ 的表达式，得

$$(I_1\omega_x\hat{\mathbf{x}} + I_1\omega_y\hat{\mathbf{y}} + I_3\omega_z\hat{\mathbf{z}})\cdot(\sin\theta\sin\psi\hat{\mathbf{x}} + \sin\theta\cos\psi\hat{\mathbf{y}} + \cos\theta\hat{\mathbf{z}}) = \alpha$$

将(5-4-10)代入上式，整理得

$$I_1\dot{\varphi}\sin^2\theta + I_3s\cos\theta = \alpha \tag{5-4-19}$$

再令 $I_3s = \beta$, $\cos\theta = x$，并用 $\sin^2\theta$ 乘式(5-4-18)，然后由式(5-4-19)解出 ψ，代入式(5-4-18)中，整理后得

$$I_1\left[\dot{x}^2 + \frac{(\alpha - \beta x)^2}{I_1^2}\right] + \frac{\beta^2}{I_3}(1 - x^2) = 2(E - mglx)(1 - x^2) \tag{5-4-20}$$

方程组(3-8-7)多出了 $(I_2 - I_3)\omega_y\omega_z$ 等项。如果取用固定在空间的坐标系，所有转动惯量分量都将是时间的函数，惯量积亦不会永远为零，问题将更复杂，因此采用**欧勒**方法比较方便。但运动最后还是要从固定坐标系来观察。

其中

$$\dot{x}^2=\frac{1}{I_1}\left(2E-\frac{\beta^2}{I_3}-2mglx\right)(1-x^2)-\frac{(\alpha-\beta x)^2}{I_1^2}$$

是 x 的三次多项式，有三个实根：$-1<x_1<x_2<1<x_3$ (特殊情况下，可有一个或多个等号代替不等号)。因此，

$$\dot{x}^2=\frac{2mgl}{I_1}(x-x_1)(x-x_2)(x-x_3) \tag{5-4-21}$$

经变数变换，可把式(5-4-21)化为椭圆积分，$\cos\theta=x$ 可用 t 的椭圆函数表示，而由(5-4-18)及(5-4-16)两式可解出 $\dot{\varphi}$ 及 $\dot{\psi}$ 为 x 的函数：

$$\dot{\varphi}=\frac{\alpha-\beta x}{I_1(1-x^2)},\quad \dot{\psi}=\frac{\beta}{I_3}-x\dot{\varphi} \tag{5-4-22}$$

利用 $x(t)$的关系就可由(5-4-21)两式求出 $\varphi(t)$及 $\psi(t)$，从而完全确定陀螺的运动。$\varphi(t)$的变化称为进动(z 轴绕 Z 轴转)，而 $\theta(t)$的改变(z 轴在 $\theta_1(x=x_1)$与 $\theta_2(x=x_2)$之间来回摆动)称为**章动**。

5-5 刚体的一般运动—平动与转动的合成

刚体的一般运动为转动和整体平动的合成，它与单纯的质点平动和刚体的转动的不同之处涉及两个方面：作用于物体不同位置上的力如何合成，以及物体整体发生平动时，不再有相应的固定点(或固定轴)可用于描写其各种转动参量以及转动方程。本节主要以刚体最简单的一般运动—刚体的平面运动(即一维或二维的平动和一维转动的集合)为例，介绍如何处理上述两个问题。

5-5-1 刚体受力的合成

作用到实际物体上的外力通常是分布在物体的不同位置上的。相同的力作用在物体的不同位置上所产生的效果往往并不相同，这就对问题的求解带来一定的复杂性。但刚体由于不会发生形变，一个力作用在刚体的某些位置上产生的效果可能完全相同，这样就可能将一些力的作用点集中迁移到某些特定位置(例如质心)，从而在一定程度上简化对物体的受力情况分析。

为了说明力的作用点的迁移问题，设想外力 **F** 作用于刚体的某一位置 A，外力 **F**′作用于刚体的另一位置 B，按 **F**′与 **F** 所引起的不同的合成效果对 **F**′的作用形式进行分类，探讨简化受力分析的方法。**F**′与 **F** 的相互关联有以下几种类型：

(1) **F**′ = –**F**，而且 **F**′和 **F** 沿两个作用点 A 和 B 的连线(见图 5-5-1a)，由于刚体不发生变形，无论 **F**′作用在连接 A 和 B 的直线上的任何一点都完全抵消了 **F** 对刚体产生的全部力学效果。这意味着 **F**′对刚体的作用与它作用在作用线上的位置无关[1]。或者说，需要考虑的只是作用在刚体上的力的作用线，而可以忽略作用线上的具体位置。利用这一特性，可以将任意多个**共面**的非平行作用力两两交汇于一点，求出它们的合力(见图 5-5-1b)。再用类似的方法。相继将所有作用在刚体上的非平行力交汇于一点后计算出它们的合力。

(2) **F**′和 **F** 平行或反平行，但不共线。这时可由力的代数和确定合力，而力的作用线可用力矩关系确定：合力对垂直于各个力的轴线的力矩与各分力对同一轴线的力矩的代数和相等。例如，图 5-5-1c 中的两个不沿同一条作用线的平行力 $\mathbf{F}_1$ 和 $\mathbf{F}_2 = -\mathbf{F}_1$ 构成**力偶**。对垂直于在平面的任意轴线的合力矩不为零，P 为 $\mathbf{F}_1$ 和

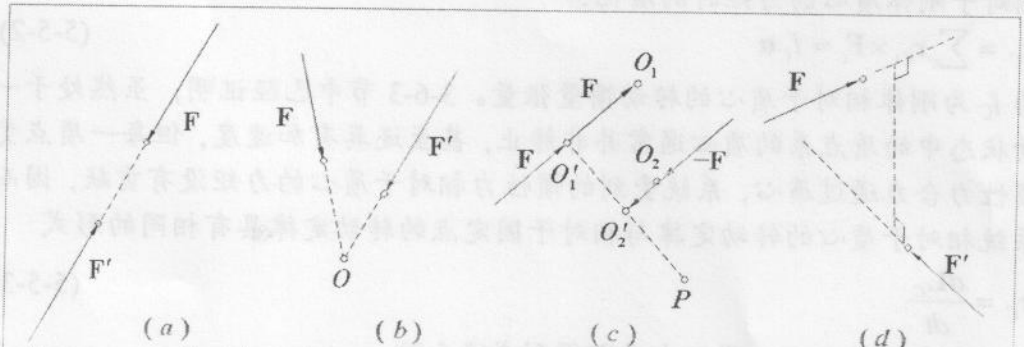


图 5-5-1 (*a*) 作用在刚体上的力所产生的力学效果与力的作用点在作用线上的位置无关，(*b*) 两个**共面**的非平行力交汇于一点后给出合力，(*c*) 不沿同一条作用线的平行力 $\mathbf{F}_1$ 和 $\mathbf{F}_2 = -\mathbf{F}_1$ 构成**力偶**，(*d*) 两个力既不平行又不能交汇于一点。

[1] 刚体力学中的力因此也被称为**滑移矢量**。但对于可变形的物体，力的作用点沿作用线移动所产生的力学效果是不同的(见第 6 章)。

$\mathbf{F}_2$所在平面内的任意一点，$\mathbf{F}_1$和$\mathbf{F}_2$对P点的合力矩为$\tau = F\overline{AP} + F'\overline{BP} = F\overline{O_1'O_2'}$。其中$\mathbf{F}_1$及$\mathbf{F}_2$之间的垂直距离$\overline{O_1'O_2'}$称为**力偶臂**。作为力偶唯一力学效果的**力偶矩**矢量通常用垂直于力偶面的任一直线来表示，其方向用右螺旋法则确定。因此，力偶矩为一**自由矢量**，可作用于力偶面上的任一点。上述力偶矩可写为$\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F}$，其中$\mathbf{r}$是$\mathbf{F}_2$的作用点$O_2$对$\mathbf{F}_1$的作用点$O_1$的位置矢量。

(3) $\mathbf{F}'$和$\mathbf{F}$既不平行也不能交汇于一点(图 5-5-1d)。这两个力无法直接合成。

下面讨论力的作用点的迁移问题。设外力$\mathbf{F}$作用在刚体的上的某一点O，为了将力$\mathbf{F}$的作用点迁移到我们感兴趣的点C上，我们引入一对假想力$\mathbf{F}' = \mathbf{F}$和$-\mathbf{F}'$同时作用于C点(见图5-5-2)。这对假想作用力显然不会对刚体产生任何动力学效应，但它们对刚体的受力情况提供了重新进行组合的可能性：其中的一种是将刚体所受到的外力组合为作用在C点的力$\mathbf{F}'$和作用在相对于C点的力偶矩$\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F}$，其中$\mathbf{r}$为$O$对$C$的位置矢量。

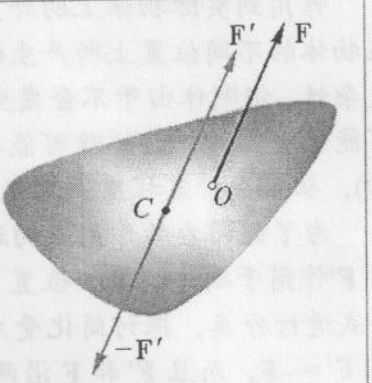


图 5-5-2 作用在刚体任一点上的力可转化为经过质心的一个力和一个力偶。

利用上述方法可以将作用在刚体其它点上的单个力全部转化为经过C点的一个力和一个力偶。由此得到作用在C点上的单个合力和相对于C点的单个合力偶。C点由此称为**简化中心**，合力称为**主矢**，而合力偶矩称为对简化中心的**主矩**。刚体动力学中，通常选择刚体的质心为简化中心，这样，刚体的动力学问题可以按平动和转动分别进行处理。

5-5-2 刚体的一般运动方程

对图5-5-2中刚体的受力分析很容易让读者产生这样的猜想，即可以用多质点系统的质心运动方程(3-6-5)描述整个刚体的平动：

$$\mathbf{F} = \sum_i \mathbf{F}_i = \sum_i m_i \mathbf{a}_i = m\mathbf{a}_C \tag{5-5-1}$$

而用类似于刚体系统的转动方程(5-2-16)描述刚体相对于自身质心的转动状态在受到相对于刚体质心的力矩时的演化：

$$\boldsymbol{\tau}_C = \sum_i \mathbf{r}_{Ci} \times \mathbf{F}_i = I_C \boldsymbol{\alpha} \tag{5-5-2}$$

这里的I_C为刚体相对于质心的转动惯量张量。3-6-3 节中已经证明，虽然处于一般运动状态中的质点系的质心通常并非静止，甚至还具有加速度，但每一质点受到的惯性力合力通过质心，系统受到的惯性力相对于质心的力矩没有贡献，因而整个系统相对于质心的转动定律与相对于固定点的转动定律具有相同的形式

$$\boldsymbol{\tau}_C = \frac{d\mathbf{L}_C}{dt} \tag{5-5-3}$$

对于刚体而言$\mathbf{L}_C = I_C \boldsymbol{\omega}$，代入上式即得到式(5-5-2)。

为避免因刚体定点转动本身的困难给处理刚体一般运动带来的额外复杂性，下面只讨论刚体的平面运动，即刚体的平动限制于一特定的平面内，而刚体的转动受限于绕垂直于该特定平面的轴构成的复合运动。

5-5-3 刚体的平面运动

刚体的平面运动可以分解为刚体绕某一固定方向的轴的转动和刚体整体随着该轴的平动的合成。例如一个轮子在地面上的纯滚动可以描述为轮子绕过其质心的对称轴的转动和质心在轮子平面内的直线运动的叠加。而根据地面与轮子的相互作用情况可以表现出轮子平动和转动的不同程度的关联。

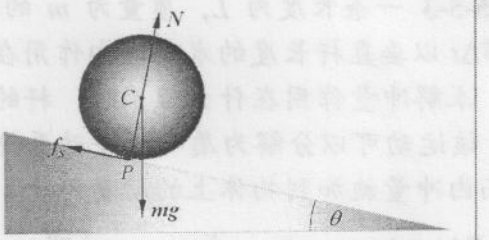
例 5-5-1 一半径为 R，质量为 m 的实心圆柱体无滑动地滚下一个坡度角为 θ 的斜面。计算该圆柱体从静止出发滚动了距离 l 后圆柱体的角速度。

解: 作用在圆柱体上的力和相应的作用点分别为：作用在质心的重力 mg，作用在圆柱体与斜面接触点 P_C 的弹力 N 以及摩擦力 f [2]。其中弹力 N 指向圆柱体质心，而摩擦力 f 的方向与圆柱体的整体运动方向相反(否则刚体就不会发生滚动)。

以圆柱体质心所在的对称轴作为转轴，因重力和弹力的作用线相对于该转轴的距离均为零，决定系统转动的力矩只能来自摩擦力矩。因此决定整个系统演化的质心运动方程和刚体转动方程分别为：

$$mg\sin\theta - f = ma_C$$

$$I_C\alpha = fR$$



例 5-5-1 图 圆柱体在斜面上的纯滚动

以上两个方程不足以求解其中 f, a_C, α 三个未知参数。但刚体纯滚动时平动参量 a_C 和转动参量 α 之间满足几何约束关系：

$$a_C = R\alpha$$

从以上三个方程可以求解出圆柱体中心的加速度：

$$a_C = \frac{mg\sin\theta}{mR^2 + I_C}R^2 = \frac{2}{3}g\sin\theta = \frac{dv_C}{dt} = v_C\frac{dv_C}{ds}$$

则圆柱体质心的末速度为

$$v_C = \sqrt{2a_C s} = \sqrt{(4/3)gl\sin\theta} = \sqrt{(4/3)gh}$$

这一速度小于系统无摩擦地滑下斜面后的速度 $\sqrt{2gh}$，这并不是因为斜面与圆柱体表面之间的摩擦耗散了部分重力势能，而是部分重力势能转化成了刚体的转动能。圆柱体的转动角速度与质心移动过的距离的函数关系为

$$\omega = \frac{v_C}{R} = \sqrt{\frac{4}{3}\frac{\sin\theta}{R^2}gl}$$

上述求解中涉及到摩擦力矩时，我们将摩擦力简化为作用在接触点 P 上。这在概念上会产生一些误解，这涉及到刚体模型本身的局限性。5-7 节会对相关的问题进行讨论。◀

例 5-5-2 将一半径为 R，质量为 m，绕其对称轴以角速度 ω_o 自由转动的均匀圆柱体放到一粗糙的水平面上，圆柱体经短暂的打滑后进入纯滚动状态，设圆柱体和水平面之间的滑动摩擦系数为 μ，求圆柱体进入纯滚动状态所需要的时间以及进入纯滚动状态后的质心速度。

---

[2] 这里的摩擦力与以前所讨论的静摩擦和滑动摩擦力不同，见 5-5 节中的讨论

解：因为滑动摩擦力是一个常数，圆柱体发生打滑的过程终质心的加速度 a 也是不变的，设打滑的时间为 T，则我们有

$$F=\mu mg=ma=mv_c/T$$

另外，绕其对称轴转动的圆柱体的角加速度 α 也时常数，因而转动方程可以表示为

$$\tau=-\mu mgR=I\alpha=-I\frac{(\omega_f-\omega_o)}{T}=\frac{mR^2}{2}\frac{\omega_o-v_C/R}{T}$$

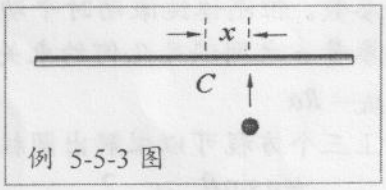
由此解得：$T=\dfrac{\omega_o R}{3\mu_k g}$ 和 $v_c=\dfrac{1}{3}\omega_o R$。

由于 v_C 与参量 m，g，μ 都无关，设想这些参量如果都为零会出现什么情况。值得注意的是，这里机械能，动量还是角动量都不守恒，但它们都取决于动量的变化以及摩擦力。◀

例 5-5-3 一条长度为 L，质量为 m 的刚性杆放置在光滑水平面上。设想有一冲量 $\mathbf{F}\Delta t$ 以垂直杆长度的水平方向作用在与杆中心相距 x 的位置。说明杆以后的运动，求解冲量作用在什么位置时，杆的一端的瞬时加速度为零。

解：该运动可以分解为质心的平动系统围绕质心的转动。质心的初始动量为零，因而由冲量施加到物体上的动量为：

$$\mathbf{F}\Delta t=\Delta\mathbf{p}_C=m\mathbf{v}_C \quad 或 \quad v_C=\frac{F\Delta t}{m}$$



例 5-5-3 图

由于冲量矩应等于冲力作用期间物体角动量的变化量：$F\Delta t=I_C\omega_f$，其中已考虑到杆的初始角动量为零，ω_f 为杆绕质心的最终角速度。记碰撞的作用点与质心的距离为 x，则

$$\omega_f=\frac{F\Delta t}{I_C}x$$

刚性杆一端的速度为 $v=v_c-L\omega/2=F\Delta t/m(1\text{-}6x/L)$。如果冲量作用在位置 $x=L/6$，杆的远端的速度为零。该位置称为**打击中心**。◀

5-5-4 刚体的一般平面运动到纯转动的变换，瞬时轴

考虑在某一瞬间，刚体中任意两点 A、B 的速度分别为 $\mathbf{v}_A$ 和 $\mathbf{v}_B$，而它们的垂线相交于空间 ***O*** 点(见图 5-5-3)，由于系统中各粒子间的几何约束，刚体上其它任一点 P 的速度也一定垂直于该点和 ***O*** 点的连线 PO。而且 P 点的速度和 PO 连线的长度成正比。或者说，在这一瞬间刚体是绕着 ***O*** 点在作纯转动，因而称 ***O*** 点为**瞬时轴**。

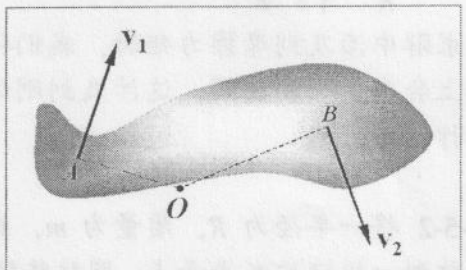


图 5-5-3 刚体中任意两点 A、B 的速度 $\mathbf{v}_A$ 和 $\mathbf{v}_B$ 的垂线相交于 ***O*** 点。

利用瞬时轴概念求解系统的运动方程只需要讨论转动问题而不必再考虑系统的平动。但由于瞬时轴不是固定的。而且一般还具有加速度，因而需要先推导出相对于瞬时轴的转动方程。为此，我们将转动方程式(5-5-3)应用于由于瞬时轴

问题，即选取**瞬心**(瞬时转动中心)作为 S'的原点。由于瞬心的速度 $\mathbf{u}=0$，瞬心参考系中的转动方程为

$$\frac{d\mathbf{L}'}{dt}=\boldsymbol{\tau}'+\mathbf{r}'_C\times\mathbf{F}_I \tag{5-5-4}$$

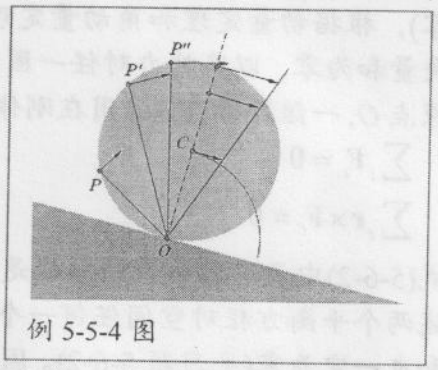
其中 $\boldsymbol{\tau}'$为外力对瞬心的力矩，$\mathbf{r}'_C\times\mathbf{F}_I=-m\mathbf{r}'_C\times\mathbf{a}$ 是作用在质心上的惯性力相对于瞬心的力矩。而瞬心 O 的加速度 $\mathbf{a}$ 可表示为质心加速度 $\mathbf{a}_C$ 及 O 点在质心系中的加速度 $\mathbf{a}'$之和：

$$\mathbf{a}=\mathbf{a}_C+\mathbf{a}' \tag{5-5-5}$$

例 5-5-4 例 5-5-1 在求解时涉及到滚动摩擦的不确定问题，利用瞬心的概念重新求解例 5-5-1 题，说明该求解方案可以避免这一问题。

解： 由例图容易看出，圆柱体作纯滚动运动时其瞬时轴就是圆柱体和支撑面的所有接触点构成的直线。其中圆柱体上任意一点 P 的运动速度的方向与图中的接触点 O 的连线垂直，大小为 $v=\overline{OP}\cdot\omega$($\omega$ 为圆柱体绕其对称轴转动的角速度，见习题 5-5-5)。将 O 点在质心系中的加速度 $\mathbf{a}'$分解为切向和法向两部分：$\mathbf{a}'=\mathbf{a}'_\tau+\mathbf{a}'_n$，其中切向分量与质心相对于斜面的加速度大小相等，方向相反：$\mathbf{a}_C=-\mathbf{a}'_\tau$。由(5-5-5)式可得：

$$\mathbf{a}=\mathbf{a}_C+\mathbf{a}'=\mathbf{a}_C+\mathbf{a}'_\tau+\mathbf{a}'_n=\mathbf{a}'_n$$



例 5-5-4 图

因 $\mathbf{a}'_n$ 沿着 O 点的半径方向，惯性力的合力与 $\mathbf{a}$ 反方向，作用点通过质心，因而 $\mathbf{r}'_C\times\mathbf{F}_I=0$。(5-5-4)式简化为

$$\boldsymbol{\tau}'=I\alpha$$

其中相对于瞬时轴 O 的力矩为重力矩 $\boldsymbol{\tau}'=R\cdot mg\sin\theta$，圆柱体相对于 O 轴的转动惯量为 $I=I_C+mR^2$。相应的转动方程为

$$I\alpha=R\cdot mg\sin\theta$$

由此可解得刚体纯滚动时的转动加速度为

$$\alpha=\frac{mg\sin\theta}{mR^2+I_C}R=\frac{2}{3R}g\sin\theta$$

与例 5-5-1 得到的结果一致。由此可解得刚体纯滚动时的相关运动参量。◀

思考题

5-5-1 在讨论纯滚动的例子中，我们看到纯滚动物体的动能总是大于纯平动物体的动能，是否存在滚动物体(可以存在滑动)的动能不大于纯平动物体的动能的情形?

5-5-2 将两个绕各自对称轴旋转的竖立圆柱缓慢靠近并接触，接触前后的总角动量是否守恒?

(提示：分析总角动量时怎样定义的？总角动量守恒或不守恒的原因，轴上的约束力对角动量是否有影响?)

5-6 刚体静力学

刚体静力学主要研究刚体处在平衡(即相对某个惯性系静止或作匀速直线运动)状态下的受力分布情况，是分析各种工程构件的强度、刚度、稳定性设计的基础，在各类制造工程、建筑等行业有广泛的应用。静力学方程常用于确定各种尺度的建筑所需要各种构件的材料，尺度，以及建筑的安全限度等问题，也可用于分析肌肉，韧带及骨骼在运动过程中的受力情况及可能形成的伤害。

5-6-1 刚体平衡条件

处于平衡(通常指静止)状态的刚体的动量和角动量均不随时间改变(通常等于零)，根据动量定理和角动量定理，刚体平衡的条件有两条：刚体所受的外力的矢量和为零，以及外力对任一固定点的力矩的矢量和为零。取固定点为坐标系的原点 O，一组外力 $\{\mathbf{F}_i\}$ 作用在刚体的不同位置 $\mathbf{r}_i$ 上，则系统的平衡条件表示为：

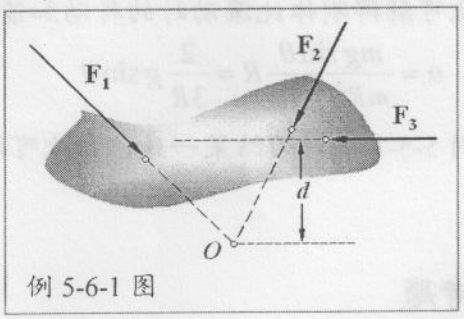
$$\sum_i \mathbf{F}_i = 0 \tag{5-6-1}$$

$$\sum_i \mathbf{r}\times\mathbf{F}_i = 0 \tag{5-6-2}$$

在式(5-6-2)中没有具体说明力矩是相对于哪个固定点或固定轴的。事实上，只要上述两个平衡方程对空间任何一个固定点满足时，外力对任何固定点的力矩的矢量和也一定为零(见习题 5-6-2)。因此，我们实际上可以针对不同的具体问题选择最方便的转动轴来计算力矩。

对于一般的刚体平衡问题，平衡方程(5-6-1)和(5-6-2)涉及 6 个独立的分量方程。为简单起见，我们只讨论平面平衡问题，即作用在刚体上的所有的力都是共面的特殊情形。则(5-6-1)、(5-6-2)式所对应的 6 个方程简化为 3 个。

例 5-6-1 利用刚体的平衡条件方程(5-6-1)和(5-6-2)，说明如果一个刚体受到三个互相不平行的外力作用而平衡，则这三个作用力不仅共面，而且三条力线的延长线必定共点。



例 5-6-1 图

解： 一个受到三个力作用的刚体只有在两种情况中有可能平衡：三个力相互平行，或者是三条互相不平行的力的力线的延长线能相交于空间的某一点。

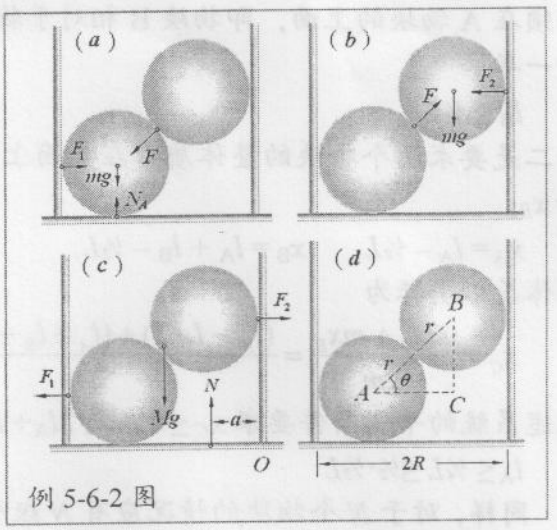
例图中表示的是三个作用在物体上的力互相不平行，而且不能相交于空间的一点的情形：。如果我们选取其中两个共面的力 $\mathbf{F}_1$ 和 $\mathbf{F}_2$ 的作用线交点 O 作为参考点，则相对于 O 点，物体仅受到来自作用力 $\mathbf{F}_3$ 的一个净力矩 $|\boldsymbol{\tau}| = F_3 d$，因而刚体无法达到平衡。◀

5-6-2 刚体系统的平衡问题

由两个或两个以上的刚体通过一定约束方式连接起来的系统称为**刚体系统**。分析刚体系统平衡问题的基本原则与处理单个刚体的平衡问题相同，但关于系统

平衡的概念扩展为：如果系统整体是平衡的，组成系统的每一个局部以及每一个刚体也必然是平衡的。实际计算时有时会遇到这样的情形：作用在系统上的外力似乎满足平衡条件，但由此得到的单个刚体本身的力系却是不平衡的，这意味着遗漏了问题的某些方面。另外在某些刚体系统的平衡问题中，若仅考虑整体平衡，其未知约束力的数目会多于平衡方程的数目。这时需要将刚体系统中的各个部分分开，依次考虑每个刚体的平衡，才能求出全部未知约束力。

例 5-6-2 水平面上一个半径为 R 的无顶无底圆桶，其中放置两个质量为 m、半径为 $r(>\frac{1}{2}R)$的相同光滑圆球(见图)。圆桶质量 M 至少要多大系统才不会倾覆？



例 5-6-2 图

解： 圆桶所受的作用力为 A 球对它的作用力 F_1，B 球对它的作用力 F_2，自身的重力 Mg 以及地面支承力 N。其中 F_1、F_2、Mg 的方向和作用点已知，必须求出 F_1、F_2 和 N 的大小及作用点才能解出 M 的最小值。A 球、B 球以及圆桶的力学平衡条件分别为(见图 b)：

A 球：$F_1 = F\cos\theta$，$N_A = F\sin\theta + mg$

B 球：$F_2 = F\cos\theta$，$F\sin\theta = mg$

圆桶：$F_1 = F_2$，$N = Mg$

圆桶的平衡条件取决于力矩平衡。选择圆桶右侧与地面的接触点 O 为参考点，设 **N** 的作用点距离 ***O*** 点 a 处，相对于 ***O*** 点的力矩平衡条件为

$$F_1 r + Mg(R-a) - F_2(2r\sin\theta + r) = 0$$

由上面四个方程可解得

$$F_1 = F_2 = mg\cot\theta$$

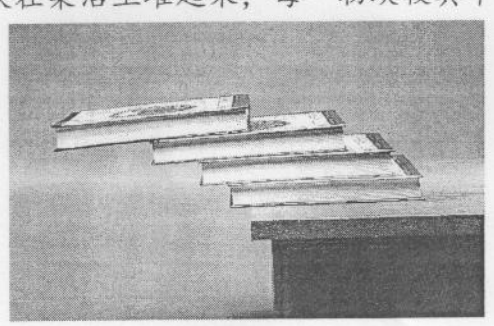
$$N_A = 2mg$$

$$M = \frac{2mr\cos\alpha}{R-a} = \frac{2m(R-r)}{R-a}$$

其中已利用了几何关系(图中 ΔABC)：$\cos\theta = (R-r)/r$。在 $a=0$ 时 M 取最小值：

$$M_{\min} = m\frac{2(R-r)}{R}$$

例 5-6-3 如图将 N 块完全相同的长方体物块在桌沿上堆起来，每一物块较其下面的物块尽可能延伸出一部分，求 N 个物块的最大累计延伸量为多大？



解： 先考虑第一个物块，由于物块下落时是以桌沿为轴转动的，我们以桌沿为参考原点写出平衡方程。在竖直方向，物块受到的重力和桌面的支撑力应达到平衡：

$$\mathbf{N} = -m\mathbf{g}$$

物块受到的重力是作用在质心上的，而桌面的支撑力实际是作用在物块和桌面的

每一个接触点上的。当物块的质心在桌面上方时，而净支撑力也作用在质心上，则系统不发生转动的条件也就得到了满足。

如果物块的质心越出了桌沿，就会破坏不发生转动的条件，因此第一个物块的最大延伸量是其质心正好处在桌沿上。这时各作用力的力臂均为零，而物块在桌面上的部分实际上每一个接触点已没有相互作用，即系统处于临界平衡状态。则物块的最大延伸量 l 为其长度的一半 $l = ½L$.

对两个物块的情况有两组平衡条件。首先是对于 A 上面的 B，要求其质心必须在 A 物块的上面，即物块 B 相对于物块 A 的外沿的延伸量不得超过其长度的一半

$l_B \le ½L$.

第二是要求两个物块的整体质心在桌面上方。设两个物块的质心坐标分别为 x_A 和 x_B，

$x_A = l_A - ½L$，　$x_B = l_A + l_B - ½L$

整体质心坐标为

$$x_C = \frac{mx_A + mx_B}{2m} = \frac{(l_A - L/2) + (l_A + l_B - L/2)}{2} = \frac{2l_A + l_B - L}{2}$$

上述系统的平衡条件要求 $x_C \le 0$，或 $2l_A + l_B \le L$。因为 $l_B \le ½L$，所以

$l_A \le ¼L \le ½ \cdot ½L$

同样，对于 N 个物块的情况应有 N 组平衡条件：对于最上面的物块要求其质心必须在其下面物块的表面上方，而对于从上往下数的第 k 个物块，则要求其上面的 k–1 个物块的整体质心必须在其表面上方。容易推得这对应于各物块的延伸量为

$$l_1 \le \frac{1}{1}\frac{L}{2},\quad l_1 \le \frac{1}{2}\frac{L}{2},\quad \cdots,\quad l_k \le \frac{1}{k}\frac{L}{2}$$

则对于 N 个物块的最大累积延伸量为

$$l_1 + l_2 + \cdots + l_n \le \frac{L}{2}\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}\right)$$

由于括号中的数列和随着 n 的增加不存在上极限，木块的堆积可以延伸到距离桌子边缘任意远的地方。◀

思考题

5-7 刚体模型所引起的运动不确定性问题

刚体模型由于忽略了物体的变形，因而能对物体整体的平动和转动这些外部运动特征给出简洁的描述。但对许多涉及物体内部自由度的真实运动状态和过程以及受力平衡问题，刚体运动方程往往不足以确定和描述物体的状态。从而表现出相应的不确定性[3]。这是因为即使非常坚硬的物体在外力作用下都会发生变形，只是在许多情形中，物体的微小变形难以通过眼睛察觉，因而仅仅刚体的运动方程的数目并不足以确定问题的受力细节。本节我们将通过对几个具体典型事例的讨论，让读者对刚体模型所引起的运动不确定性问题有一个直觉的了解。也作为进入下一章讨论物体内部自由度前的一个衔接和过渡。

需要说明的是，刚体模型对许多不确定问题仍然可以通过引入一些简洁的假设，给出大致合理的结果。例如，在很多情况下通过引入一些近似的约束(通常是一些极限条件)就可以利用刚体模型得到与实际情况相近的结果。

5-7-1 平衡问题

如果知道所有作用在刚体上力的分布，我们就可以确定这个刚体是否处于平衡。但一个处于平衡的物体一般却不足以让我们确定它所受到的约束情况。这在数学上表现为给出的力学平衡方程数目小于需要确定的力的数目。这样的静力学系统也就构成了一个不确定系统。一个系统的**不确定度可以**定义为系统的约束力数目和独立方程的数目的差值。

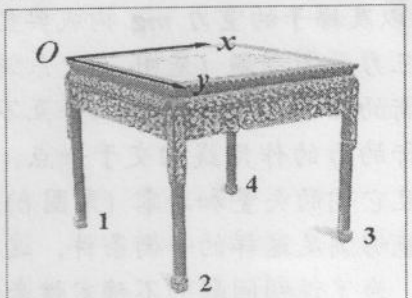
例 5-7-1 一张方桌的桌面边长为 L，质量为 m，四条桌腿的质量相对于桌面而言可以忽略不计，若方桌放在水平地面上处于力学平衡。说明方桌的受力情况。

解：方桌的力学平衡由竖直方向的受力平衡方程和相对于两条相邻桌沿(固定轴)的力矩平衡方程所决定：

$$N_1+N_2+N_3+N_4-mg=0$$

$$LN_3-\frac{L}{\sqrt{2}}mg\sin 45°+\sqrt{2}LN_4\sin 45°=0$$

$$-LN_2+\frac{L}{\sqrt{2}}mg\sin 45°-\sqrt{2}LN_4\sin 45°=0$$



例 5-7-1 图 桌面的平衡

其中 N_1, N_2, N_3, N_4 分别为作用在桌面四个角上的弹力。上述三个方程不足以确定这四个弹力，因而构成不确定问题。求解这类问题需要额外的系统状态信息，一个极端的例子是其中的一条桌腿稍短而导致桌面系统实际只受到三个未知支撑力，若假设桌腿 4 稍短，则 $N_4=0$，由此可以得到

$$N_1=0,\quad N_2=N_3=\tfrac{1}{2}mg$$

显然，这样的平衡对系统是不稳定的。系统质心位置的微小偏移(例如将一个物体放置在桌面的某一个位置上)就会导致桌面受力分布发生明显的改变。

---

[3] 实际上，质点模型也存在同样的问题，只是质点没有空间尺度，没有内部自由度。

通常对处理这类系统引入的额外信息(约束)是假设系统具有正方对称性，在平衡问题中，该几何条件所反映出的力学特性表示为

$$N_1 = N_2 = N_3 = N_4 = N = \tfrac{1}{4}mg$$

从而使问题得到确定。上述对称性假设并不符合一般的真实情况，但得到的结果却与实际情况相差很小。这是因为实际桌面不会是刚性的，即使四条桌腿长短略有差别，并非完全刚性的桌面可以通过发生微小变形使四条桌腿着地，例如，假设桌腿 4 稍短，桌面就会因自身的重力以桌腿 2、3 的对角线为界向两侧弯折，导致桌面的重量能大致均匀分布在四条桌腿上。◀

例 5-7-2 设想一架梯子倚在墙面上，确定梯子平衡时的受力情况。

解：梯子问题是刚体静力学中不确定问题的典型事例。在以梯子与地面的接触点为原点的坐标系中(见例图)，整个系统的平衡方程为

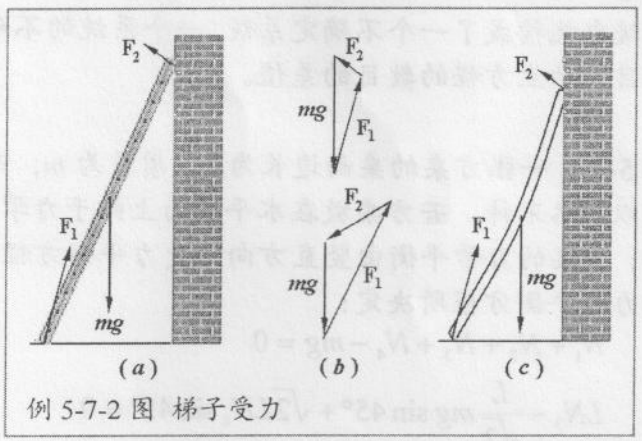
$$\sum F_x = f_G - N_W = 0$$

$$\sum F_y = f_W + N_G - mg = 0$$

$$\sum \tau_O = mgx - f_W l\cos\theta - N_W l\sin\theta = 0$$

其中 mg 为梯子受到的重力，N_G 和 N_W 为地面与墙面对梯子的弹力，f_G 和 f_W 为地面与墙面对梯子的静摩擦力。三个方程中包含了四个未知参数：f_G，f_W, N_G, N_W[4]。

这一问题中的不确定性可通过 5-6-1 节中讨论过的三力平衡需要满足的条件来说明。首先将作用在梯子两端的摩擦力和弹力各合成为一个净接触力 **F**，则梯子与地面和墙面作用的力 $\mathbf{F}_G$ 和 $\mathbf{F}_W$ 以及梯子的重力 $m\mathbf{g}$ 构成典型的三力平衡问题 (见图 *a*)。系统平衡的必要条件就是这三个互不平行的力的作用线相交于一点，而且它们的矢量和为零 (见图 *b*)，由于对于确定的重力，存在着多种接触力的组合能够满足这样的平衡条件，这就使问题变得不确定了。



例 5-7-2 图 梯子受力

为了说明问题的不确定性来源于梯子的刚体模型，我们设想梯子在重力的作用下发生如图 *c* 的形变，对于该问题所有原来的平衡方程都没有发生变化，但梯子的形变程度却提供了求解该问题的额外信息，虽然梯子的形变仍有着许多的可能性，对于一个确定的变形就足以解决方程数量不足的问题[5]。

刚体模型虽然无法确定梯子受力分布问题，但在实际情况中，墙面一般比较光滑，而梯子用于登高时其仰角一般也不大，因而 f_W 很小。这时如果忽略梯子

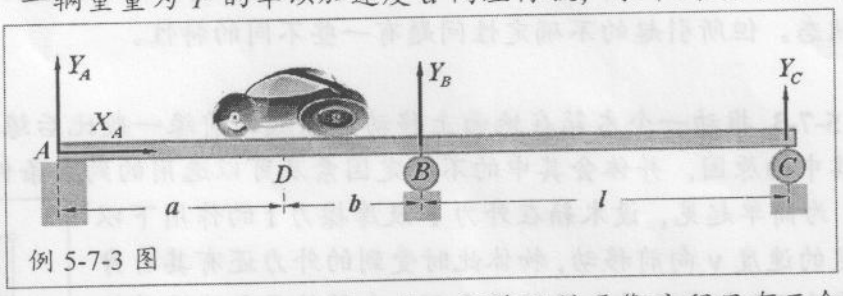
---

[4] 有些读者也许认为摩擦力可以表述为摩擦系数和弹力的乘积，但对于静摩擦，无法确定静摩擦系数

[5] 许多实际工程问题中，为提高结构的强度和刚度或其它工程要求，经常需要在静定结构上再附加一些构件或者约束，从而使作用在刚体上未知约束力的数目多于独立的平衡方程数目。这类问题称为超静定问题，相应的结构称为超静定结构。即对于超静定问题，必须考虑补充某些方程以反映物体因受力而产生的变形，使方程的数目等于未知量的数目。

与墙面的摩擦(可以理解为一种约束)，就可以近似计算梯子临界仰角的数值以及梯子在给定仰角情况下人能够站到的最高位置。◀

例 5-7-3 例图中表示处于力学平衡状态的横梁桥 AC。其中 A 端固定，B 处和 C 端都搁置在辊轴上。一辆重量为 P 的车以加速度 a 向左行驶，求车到达与 A 端相距为 a 的位置 D 时，A, C 两端受到的约束力。



例 5-7-3 图

解： 系统整体所受到的外部约束分别为 A 端的 N_{Ax}, N_{Ay} 和 B 处及 C 端竖直方向的力 N_{By} 和 N_{Cy} 共四个。所涉及的平衡方程只有三个

$$N_{Ax} - f_x = 0$$
$$N_{Ay} - f_y + N_{By} + N_{Cy} = 0$$
$$f_y \cdot a - N_{By} \cdot l - N_{Cy} \cdot 2l = 0$$

求解 A, C 两端未知的约束力需要计入物体材料变形的因素。

最简单的处理方法是将形变集中在梁的中间支点 B，虽然 B 点处的系统形变导致的内力无法出现在结构整体平衡方程中，但可以想象横梁在 B 处分为两半，则对于两个分开的单个刚体而言，B 处的约束表现为外力，由此可以获得额外的约束方程。对于 BC 段刚体梁的平衡，由相对于 B 点的力矩平衡方程

$$N_{By} + N'_{By} + N_{Cy} - W = 0$$
$$N_{Ay} + N_{By} - N'_{By} - W = 0$$
$$N_{Cy} \cdot l - W \cdot \tfrac{1}{2} l = 0$$

求得：$N_{Cy} = \tfrac{1}{2} W$。代入 y 方向的整体平衡方程

$$N_{Ay} - f_y + N_{By} + \tfrac{1}{2} W = 0$$
$$f_y \cdot a - N_{By} \cdot l - Wl = 0$$

可得：$N_{By} = f_y \cdot (a/l) - W$，$N_{Ay} = (1 - a/l) \cdot f_y + \tfrac{1}{2} W$。◀

说明：建筑或桥梁处于静力学平衡时，作用在其每一构件上的合力都必须为零。而计算任一个构件上的合力时必须计入内力，因而在大多数场合中刚体模型并不可行。因为即使如计算一条两端受到支撑的横梁在力学平衡，也会涉及横梁的弯曲。

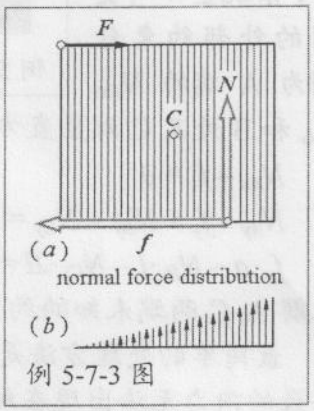
实际处理建筑或桥梁的平衡问题时，可以将刚体系统理解为若干个非刚性连接的刚体构件的集合。例如设想将横梁分解为许多刚体小段，彼此之间用弹簧连接，利用每个刚体构件的平衡条件，可以得到系统的形变后的位型，进而求解出横梁所受到的外部约束。这种方法称为**有限元方法**，其计算精度随着体积元的数目增加可以得到系统性的改善。例如，考虑一根横梁受到纵向的外部作用力，将梁想象为由 N 个弹簧连接起来的重刚体，用弹簧力描述梁抵抗压缩或拉伸的内力。每条弹簧的强度通过匹配测量真实横梁的一些实际特性决定，例如对横梁施加确定的应力后所表现出来的拉伸程度。该模型可用于分析横梁端部受到纵向外力时的响应，然后逐步通过连接弹簧将作用力传递到横梁的各个位置。这些构件可以调整各自的位置而使静力学平衡条件得到满足。因而对每一个构件都有一个方程，而求解 N 个联立方程(simultaneous equations)很适合利用计算机作数值求解。这些构件经调整后的位置反映

了真实的连续体梁的压缩和拉伸的情况。

方程(5-6-1)和(5-6-2)所表述的是系统的加速度和角加速度为零的情况，它们适用于刚体静止的情况，也适用于刚体没有加速度和角加速度为零的所谓动态平衡状态。但所引起的不确定性问题有一些不同的特性。

例 5-7-3 推动一个木箱在地面上移动，箱底的前缘一般比后缘磨损程度高，分析其中的原因，并体会其中的不确定因素及可以选用的约束条件。

解：为简单起见，设木箱在外力 **F** 及摩擦力 **f** 的作用下以恒定的速度 **v** 向前移动，物体此时受到的外力还有其自身的重力 **W** 和地面对木箱的弹力 **N**。木箱处于匀速运动状态意味着所有的四个力达到了平衡。另外，由于物体没有转动，这些力相对于任何一点的合力矩必须平衡。如果以过质心 C 而且垂直页面的水平轴为参考轴线，由于外力和摩擦力相对于质心形成同方向的力矩，则要求作用在接触面上的支撑力的合力就不能过系统的质心，而必须在运动方向的前部，以形成反方向力矩(见图 a)。由于各点的摩擦力与该点的正压力成正比，前半部分受到的摩擦力应大于后半部分，因而前半部分受到的磨损也较后半部分严重。



例 5-7-3 图

但实际的支撑力分布和物体在运动过程中的变形情况有关，以上的分析并不能确定实际的正压力分布。一种方法是将支撑力分布近似为线形的(见图 b)。建议读者自己分析其中的合理性。◀

5-7-2 滚动摩擦

刚体模型引起的不确定性问题的另一个典型事例是滚动摩擦问题。所谓滚动摩擦一般是指一个圆柱或球形物体在另一个物体上滚动时所受到的对滚动的阻碍作用。

滚动摩擦是出现在滚动物体和地面接触点处的阻碍物体相对运动的作用，如果将滚动物体和地面都看成是完全刚性的，则分别从刚体的平动方程和转动方程求解得到的结果表明滚动摩擦在刚体纯滚动中所起的作用就是相悖的。从受力的角度考虑，将滚动摩擦理解为和物体运动方向相反的接触力(见图 5-7-1-a)，因为不存在其他水平方向的力，滚动摩擦是导致物体减速的唯一原因，但从力矩的角度分析，这个向后的摩擦将导致物体围绕质心加速旋转，由于纯滚动的几何约束条件，物体应得到向前的加速度而不是减速度。同样的分析可以发现，如果将滚动摩擦理解为是向前的水平力，则会导致与上述相反的悖论。

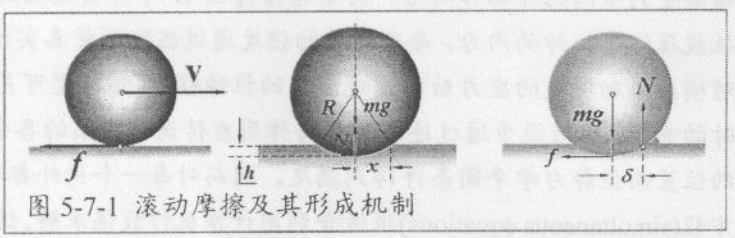


图 5-7-1 滚动摩擦及其形成机制

解决上述悖论的视角是类似于例5-7-3中的情形，

找出系统转动减速的力矩。事实上，作用在系统上的作用力除了摩擦，还有重力和地面对系统的弹力，由于重力作用在质心上，对系统绕质心旋转的力矩没有贡献，而要使系统受到的地面弹力产生于摩擦力矩反向的力矩，系统与地面的接触点必须处在质心的前方。这就意味着正确描述滚动摩擦就必须舍弃系统的刚性特征。设想滚动物体受重力作用而压入支承面(见图 5-7-1-b)，同时本身也受压缩而变形。即两者的接触不仅仅是一个点。当圆柱在水平地面上静止不动时，物体与地面的形变是左右对称的，物体受到的重力 mg 和地面的支持力 N 大小相等，方向相反，且处于同一直线(见图 5-7-1-b)。但物体在向前滚动或具有向前滚动的趋势时，物体与地面的接触部分只是右侧的一段“弧线”，物体所受的支持力也只是分布在右侧的“弧线段”上，这使得支承面对物体的弹力 N 的等效作用点从最低点前移(见图 5-7-1-c)。弹力 N 沿竖直方向和水平方向分解为 N_x、N_y。但 N 一般与竖直方向的夹角非常小，所以 N_x 也很小；而且圆心正下方的 B 点到 N_x 作用线的距离很短，N_x 对 B 点的力矩近似为零。这样 N_y 与物体受到的重力 mg 大小相等，方向相反，因不在同一直线上而形成了一个阻碍滚动的力偶矩。因此，滚动摩擦的效果是用力偶矩来量度的，其大小与正压力成正比：

$$\tau = \delta \cdot N \tag{5-7-1}$$

其中滚动摩擦系数 δ 具有长度量纲，形式上相当于弹力对于滚动物体质心的力臂。物体的形变量越大，δ 也越大，阻力矩也就越大(一个实际的例子是车胎打气不足时形变量很大，滚动阻力也增加许多)，δ 与滚动物体和支承面的材料、硬度及物体的半径等因素有关。以半径为 2 厘米的钢球在木材表面上滚动为例，δ 通常在 0.03-0.04cm 之间。

另外，物体在平面上滚动时除了受到重力和弹力外，在接触处还受到摩擦力 f 作用(纯滚动时是静摩擦力，非纯滚动时是滑动摩擦力)。由于 f 相对于物体的质心产生的力矩 $R \cdot f$ 和转动角速度方向一致，摩擦力是驱动而不是阻碍物体滚动的。

例 5-7-4 力偶矩 $\boldsymbol{\tau}$ 作用在半径为 R 的滑轮上，轮上绕的细绳拉住斜面上半径为 R，质量为 m 的圆柱(见图)。斜面倾角为 θ，圆柱与斜面间的滚动摩阻系数为 δ。求圆柱平衡时力矩 $\boldsymbol{\tau}$ 的最大与最小值，以及能使圆柱匀速纯滚动时的最小静摩擦系数值。

解： 圆柱即将滚动时的临界状态对应于滚动阻力偶矩的最大值：$\tau_{max} = \delta \cdot N$，其方向与滚动趋势相反。而细绳最小或最大拉力对应于圆柱出现向下或向上滚动趋势。

(1) 对应最小的细绳拉力 F_{min}，圆柱受力平衡方程为

$$mg\sin\theta - F_{min} - \tau_{max}/R = 0$$

$$N - mg\cos\theta = 0$$

$$\tau_{max} = \delta \cdot N = \delta \cdot mg\cos\theta$$

求得：$F_{min} = mg(\sin\theta - \delta/R\cos\theta)$

(2) 最大细绳拉力 F_{max}(见图 c)对应于上式中反号的滚阻力矩项

$$F_{max} = mg(\sin\theta + \delta/R\cos\theta)$$

(3) 相对于滑轮而言(见图 d)。由平衡方程

$$FR - \tau_B = 0$$

当细绳拉力分别为 F_{max} 或 F_{min} 时，力偶矩的最大与最小值分别为

$$\tau_{\mathrm{Bmax}} = F_{\max} R = mg(R\sin\theta + \delta\cos\theta)$$

$$\tau_{\mathrm{Bmin}} = F_{\min} R = mg(R\sin\theta - \delta\cos\theta)$$

圆柱平衡时力偶矩 τ_{B} 的取值范围为

$$mg(R\sin\theta - \delta\cos\theta) \le \tau_{\mathrm{B}} \le mg(R\sin\theta + \delta\cos\theta)$$

(4) 求圆柱纯滚动时最小静摩擦系数值 $f_{\min}$，取圆柱为研究对象，当它向上滚动时，受力如图(*c*)。由平衡方程

$$F_S R - \tau_{\max} = 0$$

$$N - mg\cos\theta = 0$$

补充方程 $\tau_{\max} = \delta F_N$，解得：$F_S = (\delta/R)mg\cos\theta$，满足纯滚动的力学条件为 $F_S \le \mu_S N = \mu_S mg\cos\theta$，可得

$$\mu_S \ge \delta/R$$

纯滚动时，最小静摩擦系数的值为 δ/R。由图(*b*)可知，圆柱向下滚动时结果相同。

5-7-4 碰撞问题

3-7 节中讨论各种碰撞事例时都没有涉及过程中的相互作用的细节，只是利用动量守恒和外加的约束条件将碰撞分类成完全弹性碰撞，完全非弹性碰撞和这两种极限之间的一般非弹性碰撞。其中完全弹性碰撞的约束条件是动能守恒，而完全非弹性碰撞的约束条件是末速度相等。如果对于靠在一起排列为一条线的多物体碰撞问题，最终的碰撞结果涉及到一个速度分布问题，显然即使对于完全弹性碰撞，两个守恒定律也不足以确定超过两个以上的粒子系统的速度分布问题。

为让读者对问题先有一个直觉，我们首先讨论一下如图 5-7-2 所示的**牛顿摆**问题。N 个完全相同的钢球摆靠在一起。当最左面的一个钢球摆回到其平衡位置与右面的近邻钢球发生撞击，这种钢球之间的相互作用相继发生，但由于所有的钢球靠在一起，我们无法观测到其中力的传递过程，而最终显示出来的结果是最右边的钢球摆了出去。如果以左面的两个钢球同步摆回到其平衡位置而与右面的钢球发生撞击，则结果是最右边的两个钢球摆了出去。这组实验结果显然是系统动量守恒和动能守恒的一个解，但它反映的实际过程是排列为一条线的多质点碰撞问题，因为排列为一条线的多质点系统，质点之间一定是相互分离的，因而可以简化为一系列的两体碰撞问题：相继的两体碰撞和速度交换[6]。这一过程实际上要求的是两两之间的相互作用是独立的，它可以通过物体之间间隔来实现，也可以通过物体形状导致的物体之间力传递的有效性来达到效果。

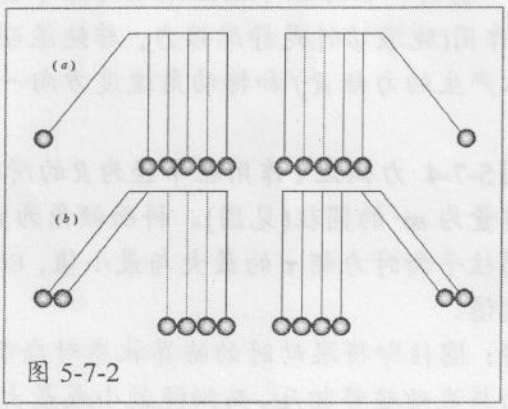


图 5-7-2

对于物体之间的间隔不足以保证物体两两之间的碰撞是独立的情形，如果缺

[6] 而且这个解有一个鲜明特征：运动状态满足时间反演的对称性。即右边的钢球摆回其平衡位置时与左面的钢球发生相继撞击，能使系统恢复到初始状态。

乏在碰撞过程中能量如何在粒子间转移的动力学机制的相关信息，就不能利用能量的转移程度来求解碰撞前后的速度分布，因而这也是一个典型的不确定性问题。

但要从理论上得到上述的解还需要额外的信息，即需要额外的假设或约束。如果我们设想各钢球之间的碰撞是依次进行的，例如一个钢球在完全结束与左边的钢球相互作用之前不会和右边的钢球发生相互作用。由于两个相同质量物体完全弹性碰撞的结果是两者交换动量，因而撞击的钢球将完全失去其原有的速度，而被撞的钢球得到了全部的动量。类似的碰撞过程依次进行，直到最右边的钢球带着全部的动量离开其伙伴。但钢球如果真是理想刚体，上述假设就不会成立。因为刚体中力的传播速度充分大，我们没有理由认为撞击过程的持续时间一定短于作用力在刚体中传播所需要的时间。如果我们将对称性作为对系统的约束，也会存在新的问题。例如为了避免力在钢球中传播的复杂性，我们考虑如图 5-7-3 所示的两个立方钢块和若干个完全相同的紧靠的立方钢块的碰撞问题[7]。按照牛顿摆的结果进行类比，我们期待结果应该是最右面的两个立方体离开。相应的碰撞将依次进行，因而结果是依次发生速度交换。但如果立方钢块之间的接触非常好，有没有其他的可能结果，例如左面的刚体最后仍有一定的前行速度，甚至有后退的可能呢？

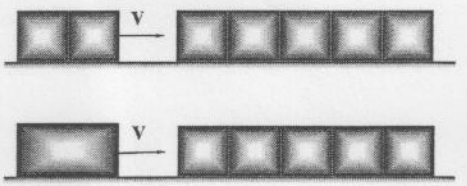


图 5-7-3

显然，这些问题涉及到两个刚体发生撞击时的接触时间，能量交换方式等内部自由度。至少，刚体模型应该不能判断钢块是如何相互连接的。相互紧靠的立方钢块之间的相互作用传递介于两种极限情况之间，一是上述的牛顿摆极限，即钢块之间的作用是相继进行的，而另一个则是撞击钢块组和被撞击钢块组可分别视为一个整体。对于具有一定尺度的系统，后一类碰撞问题并不能用质点碰撞模型来描述。而是用一根短钢棒以一定的速度 v 撞击一根材质和截面积都相同的长钢棒(见图 5-7-4)的模型描述的。通常我们会预计，短棒将以接近原来的速率弹回，而长棒只能获得很小一部分动能。但实际情况却相反，短棒弹回时的速率很小，其大部分动能都成功地转移给了长棒。从而将长棒撞击出去。我们这里不给出问题的答案，而把它留作习题和以后章节中的一个引子。由于物块的线度很大，相互作用从该滑块的一端传递到另一端所经历的时间比大滑块撞击它的时间长，上述结论不再成立。这时，必须考虑相互作用在滑块中以有限速度传播所带来的影响。这属于连续介质的力学问题。

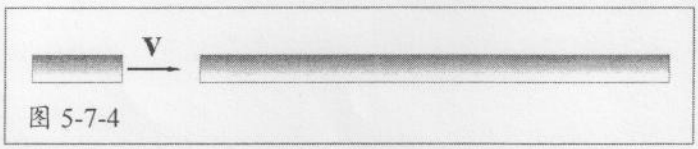


图 5-7-4

另一个解释是不考虑碰撞是否依次进行，而仅仅从两个不同时间的动量和能量守恒的角度考虑问题，如果忽略系统的内部自由度，两个方程可以给出许多种可能的解。其中包括牛顿摆实验所给出的结果。

[7] 物体在碰撞时所表现出的刚性和其形状有很大关系。

第 6 章 连续介质力学—物体的力学特性

这一章讨论涉及物体内部自由度的运动，即物体内部质点之间的相对运动。这一大类运动主要表现为物体受到外部环境(外力)的作用引起的各种内部自由度的响应，而在外部通常表现为物体的各种形变如膨胀，弯折，扭曲甚至断裂等。这类响应通常都被称作物体的特性。固态物体的弹性，韧性，脆性，硬度，液体的流动性，粘性，气体的压缩特性等都是实际物体对外部力学干扰的典型响应事例，统称为物体的力学性质。类似地称物体对外部的电磁扰动，温度变化的响应为物体的电磁性质和热学性质。

物质的上述特性并非取决于系统中某一个质点的运动细节，而是取决于系统内部所有质点的运动及其相互关联，因而相应的运动参量不再是某一个质点的位移和速度，而应该表示为系统内部诸多质点的运动关联和运动参量的某种叠加，即相应的运动参量是系统内部各质点的位置和速度的函数。对于由大量的原子或分子构成的宏观物体而言，其任何一种特性原则上都可以通过求解系统中所有单个粒子的运动方程，再通过叠加给出描述，但如果我们所观察的物体行为和性质所涉及的形变尺度远大于物体中粒子的平均间距，运动演化的时间尺度远大于物体中粒子运动的特征周期或过程时间，该物质系统的内部结构可以近似为大量质量微元的连续分布，这些微元的尺度远小于实际物体的尺度，但远大于物体中原子或分子的尺度，由此可以忽略体积元中的原子或分子的行为和特征细节。这样的物体模型称为**连续介质**。由于体积元可以发生连续的形变，彼此之间也可以有相对的位移[1]，连续介质模型可用于描述宏观的形变甚至非均匀的滑移，研究固体的各种连续形变特性和液体及气体的各种流动性质。

连续介质可以按照能否维持其形状和大小分为两类，如果物体在外力作用下发生形变，外力撤消后则恢复原状，这类物体称为**弹性体**；如果物体在外力作用下发生形变，外力撤除后维持形变后的状态，不再恢复原状，这类物体称为**流体**，流体显然没有确定形状。上述按物体对外力作用的响应进行的分类显然时相对的，有些物体在某些问题中可看成弹性体，而在另一些问题中则应视为流体。例如，一般视固体为弹性体，将液体看作流体。但液体对于外力的压缩通常表现出很完美的弹性，金属在较大的压力和拉伸作用下呈现出很好的延展性，而沥青、松香一类的固态物质则可以在重力的长时间作用下蠕变都是流动特性的表现。

连续介质模型与实际物质的明显区别是物体内部结构中不存在突变，特别对于固态物体而言，里面不存在结构方面的缺陷，而实际物体的一些力学特性，特别大多数的电磁特性和光学特性都是和物体内部的各种缺陷紧密相连的，因此在涉及相关的物体特性时，连续介质模型就不再有效。

从测量属性而言，描述涉及内部自由度的物体特性参量都是与外部扰动相关的响应函数，我们一般是通过实际物体对外部作用和扰动的响应来研究连续介质内部的结构和动力学特性的。外部的干扰通常有两种分类，第一种分类是不破坏或影响结构的微小干扰以及影响甚至破坏结构的强干扰，另一种分类是以物质相关结构的弛豫时为界，外部扰动的变化节奏远小于弛豫时间的称为**低频扰动**或**静**

[1] 体积元指的仅是与物体的体积相比足够小，但仍然大到足以包含为数极多的粒子。

态扰动，而外部扰动的变化节奏大于弛豫时间的称为**高频扰动**。测量和分析这些经验参数可用于辨识响应所涉及的自由度和自由度所受到的约束。

本章首先简单介绍物体在不同聚集形态(物态)中的几何结构，然后讨论物体发生静态或稳态形变时所呈现出的特性。本章只讨论静态扰动及响应，高频扰动及响应，即连续介质对外界扰动所表现出的力学响应在更小时空尺度上的一些典型的表现行为将在下一章中讨论。另外，第 15 章和 18 章中还将将讨论连续介质的电磁特性和热学性质。

6-1 物体的内部结构

我们所处的物质世界的魅力之一在于一个宏观物体在各种外部环境[2]中会以不同的形式出现，物体所表现出来的宏观状态按照构成该物体的原子或分子的聚集形态(通常也称为物态)可以分为**固态**、**液态**、**气态**及**等离子态**。其中固态和液态通常被称为**凝聚态**，液态和气态则常被称为**流体**。处于固、液、气这三种状态的物质的基元粒子都是电中性的，而处于等离子态的物质中的基元粒子则不再是电中性的粒子，而是被离解为电子和不同电离度的原子实离子。从物体的状态而言，等离子态物质也是一种流体。描述其行为特征的学科称为**等离子体物理**。自然界中还存在一种具有极高密度的物质状态，称为**超密态**。其典型的例子是白矮星和中子星中的物质状态，白矮星和中子星因而也被称为致密星。

本节只讨论由中性原子或分子构成的固态、液态和气态物质的内部结构。

6-1-1 作用力，相互作用能与物质结构

物质具体处于固、液和气中哪一种状态主要取决于构成物质的基元粒子(原子、离子或分子)之间的相互作用力(通常称为**结合力**)以及系统所处的外部条件。物质处于凝聚态时，其中的原子或分子处于束缚态(粒子的动能和势能之和小于零)，形成束缚的结合力主要来自原子内部正负电荷分布不完全重合所剩余的库伦作用(见 4-4 节)。不同原子或分子间的结合力存在很大差异，但作用力与粒子间距的关系在定性上非常相似：两个粒子彼此相距很近时会受到很强的排斥，而相距较远时则表现为相互吸引。图 6-1-1 为两个粒子的相互作用势能曲线 $u(r)$示意图，其中对应于曲线最低点的位置 r_o 为粒子间的静态平衡距离。

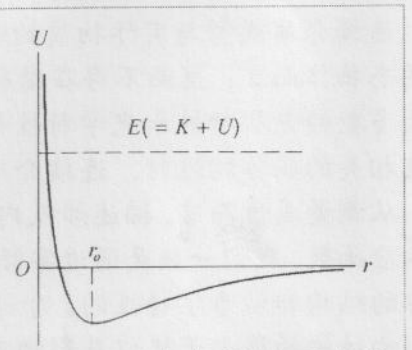


图 6-1-1 两个粒子的相互作用势能曲线示意图。

另外，从分子的空间构型可知，大多数分子都具有电偶极矩或更高的电多极矩，因此，各基元间的作用通常并不是各向同性的[3]。两个分子的相互作用势能曲线还与它们的相对取向有关。因此，两个近邻分子的实际间距在不同方向一般并不相同。但这些不同方向上的差异会随着距离的增加而迅速减小，因而只有在粒子间距较小的

[2] 物质系统与环境的作用主要指外场，外部压强和温度等。具体可参阅第第 15 章和 18 章。
[3] 势能曲线应代之以三维的势能函数

固态物质中才会表现出比较明显的各向异性，而在间距较大的液态或气态物质中一般都是各向同性的。

6-1-2 固体的结构特征及其描述

固态物质的主要特征是其具有保持一定体积(不同于气态)和形状(不同于液态)的能力。这种特性起因于固态物质中的原子或分子的动能相比粒子间的相互作用势能小很多，因而它们基本都被束缚在各自特定的位置(力学平衡位置)附近的很小区域内。固态物质中粒子的典型间距为 10~100 纳米的数量级。在这样小的空间尺度内，原子或分子能感受到相互作用在方向和距离上的细微差异，而这些细节能直接影响原子或分子的空间排列位型。

一个物质系统最稳定的内部结构原则上对应于粒子之间的总相互作用能量最低的位型，而对于大量相同粒子的聚集形态而言，能使相互作用能量达到最低的空间排列一定是某种规则的周期结构(见思考题 6-1-2)，因而大多数由简单基元构成的固态物质都呈现出三维的周期点阵结构。具有这种物态结构的固体称为**晶体**。但实际物质的内部结构还受到外部非力学因素如温度，外场等的影响，这导致同样的粒子系统在不同的外部环境中可以具有不同的结构：如不同的周期点阵，甚至是没有规则的无序结构，这些粒子分布的空间位型不是系统势能最低的稳定状态，而是对应于系统势能出现极小值的亚稳定状态。例如。纯碳原子在常温和常压下可以表现为完全不同的晶型：如石墨、金刚石和碳 60 结构(见图 6-1-2)。这种现象称为**同素异晶体**。不同的周期点阵结构对应于不同的结合能量，但都对应于相互作用能极小的状态。固态系统并不总是处在能量最低的晶态，而很相当稳定地处在能量稍高的状态，是因为这些亚稳态的构型与最低能量的位型之间存在很高的势垒。

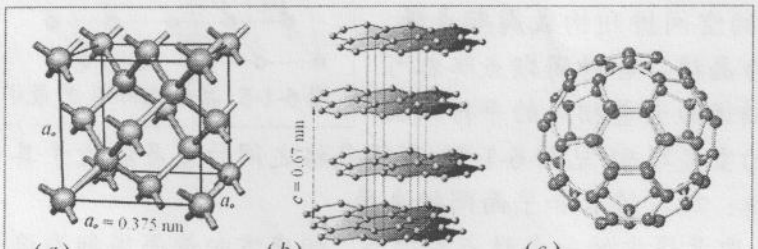


图 6-1-2 目前已知的三种碳的固态有序结构：金刚石，石墨，碳 60。

晶体结构及其对称性的分类和描述

晶体是自然界中固态物体的一个大类，它们通常表现出非常规则的几何外形。

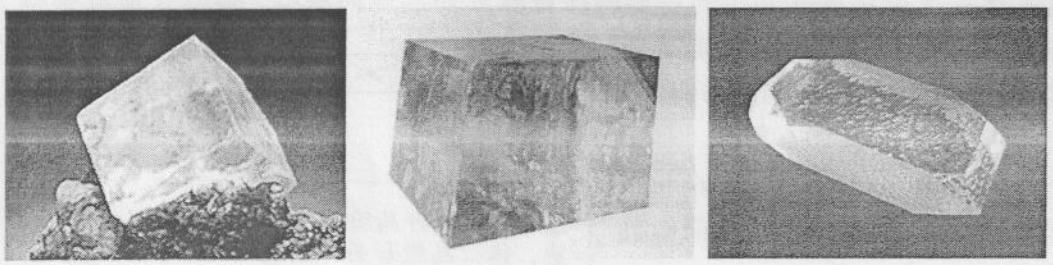


图 6-1-2 从左至右为氯化钠(NaCl)晶体，方解石($CaCO_3$)晶体和石英晶体。

天然晶体大多呈现为由裸露在外的诸多平面(称为**天然解理面**)围成的凸多面体(图 6-1-3)，这些平面也称为**晶面**。晶面间的交线称为**晶棱**，而晶棱的汇集点称为**顶点**。晶体在生长过程中，外界环境可造成其形态呈现出不同程度的畸变，图 6-1-4 为石英晶体的一些常见的外形。但晶体的规则外形具有一种规律：同一种晶体不论其外形如何，裸露的晶面之间或晶棱之间的夹角是确定的，这种现象被称为**晶面角守恒**。氯化钠(NaCl)晶体，方解石($CaCO_3$)晶体都是形状比较简单的晶体，氯化钠晶体外型是立方体，其所有天然晶面之间的角度(晶面角)都是 90°；方解石晶体的外型是斜六面体，晶面间夹角是 78°5′和 101°55′。石英晶体是天然解理面比较复杂的晶体，如图 6-1-3 所示的 a、b 面间的夹角总是 141°47′，b、c 面间的夹角总是 120°00′，a、c 间的夹角总是 113°08′。这说明晶体的大小和形状不反映晶体的特征，而外形晶面之间的夹角才是晶体特征的反映。

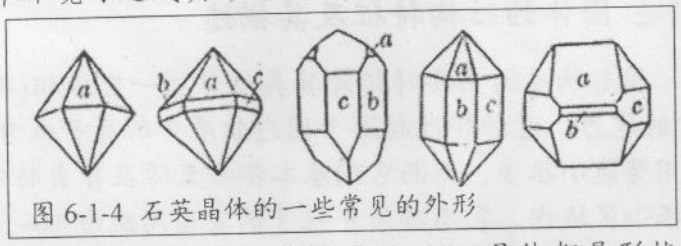


图 6-1-4 石英晶体的一些常见的外形

晶体的宏观规则外型显然来源于同类型的微观单元的堆砌。如果用一个点表示固体中的最小重复单元(原子、离子、分子或原子集团)的某个特征位置(如质心)，这些点(称为**结点**)的空间排列构成**周期点阵**，称为**晶格**。各种周期点阵都可以描述为大量相同的平行六面体的重复堆砌(见图 6-1-5)，不同晶格之间的差异取决于其平行六面体基元的几何特征：相对棱长和平面间的夹角。

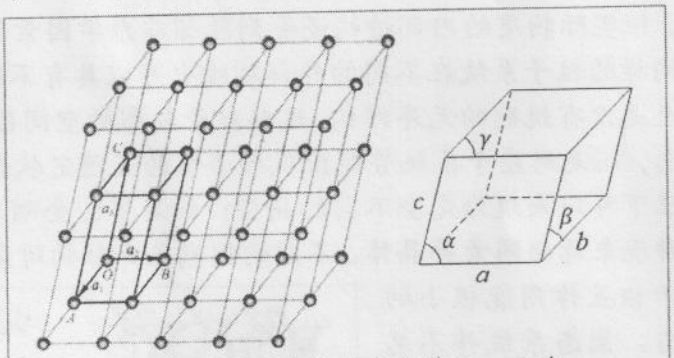


图 6-1-5 晶格点阵及其最小重复单元—元胞。

取晶格中的一个结点为顶点，相应方向最小周期为边长的平行六面体单元称为**原胞**。晶体中体积最小的原胞称为**固体物理学原胞**，对这类原胞而言，结点只处在原胞顶点，原胞体内和面上都没有结点。

晶格的周期性称为平移对称性，即晶格在特定方向上平移一个或多个周期后能与原来的点阵重合。除了具有平移对称性，晶体通常还含有一些如转动，镜面反射，中心反演等几何对称性，一个能反映一个晶体全部这类对称性的最小单元称为**结晶学原胞**(简称**晶胞**)。晶胞是以若干个固体物理学原胞组合成的平行六面体，因而结点可以出现在晶胞的一些特殊对称点如晶胞的**底心**、**体心**或所有**面心**上。结点仅出现在顶角处的晶胞称为**简单晶胞**，记为 P，除了晶胞顶点，结点还分别出现在晶胞的**底心**、**体心**或所有**面心**上的晶胞(见图 6-1-6)相应地称为**底心晶胞**、**体心晶胞**或**面心晶胞**，分别记为 C、I 或 F。

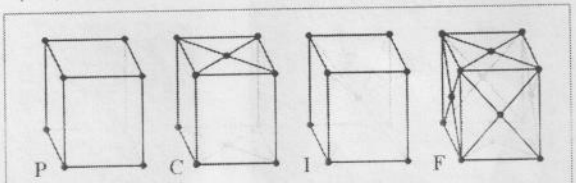


图 6-1-6 四种晶胞类型：简单晶胞、底心晶胞、体心晶胞和面心晶胞

对晶体结构的几何对称性的数学描述是晶胞经过特定的旋转、平移等变换后与原来的结构重合，这种特定的变换称为**对称操作**。常见的几何对称操作分为四

类：**平移**、**旋转**、**镜面反射**和**中心反演**。对称操作前后，晶胞中始终不变的**轴线**、**平面**或**点**称为晶胞的**对称元素**。晶体的周期性所对应的对称元素称为**离散平移对称**操作。而晶胞沿某一选定轴旋转 $2\pi/n$ 后能与原形重合，相应的对称元素称为 ***n* 重旋转对称轴**。由于晶格所满足的旋转对称操作还必须同时满足适配周期性的平移对称性，因而晶体只能具有 1、2、3、4、6 五种旋转对称元素(见习题 6-1-4)。

通常根据晶胞所属的几何对称性对晶体进行分类：取六面体晶胞的一个顶点作为坐标系的原点 O，六面体中过 O 点的三条棱 **a**, **b**, **c** 作为坐标系的坐标轴(见图 6-1-5)。晶胞的几何特性由六个参量(称为晶格结构的**点阵参数**)描述：三条棱的长度 a, b 和 c，以及 **a**, **b**, **c** 之间的夹角 α, β, γ。根据这些参量之间的关系可将晶格分为 7 个晶系，每一种晶系包含有若干种不同的晶胞：除了单胞点阵带心，7 个晶系共有 14 类晶胞(见图 6-1-6)，分别为：

立方晶系：立方晶系是所有晶系中对称性最高的晶系。该晶系的特点是有三个等长且互相垂直的结晶轴，4 个立方体对角线方向的三重轴。点阵参数的限制条件为 $a=b=c$，$\alpha=\beta=\gamma=90°$。立方晶系包含简立方、体心立方和面心立方三种晶胞(见图 6-1-6)：

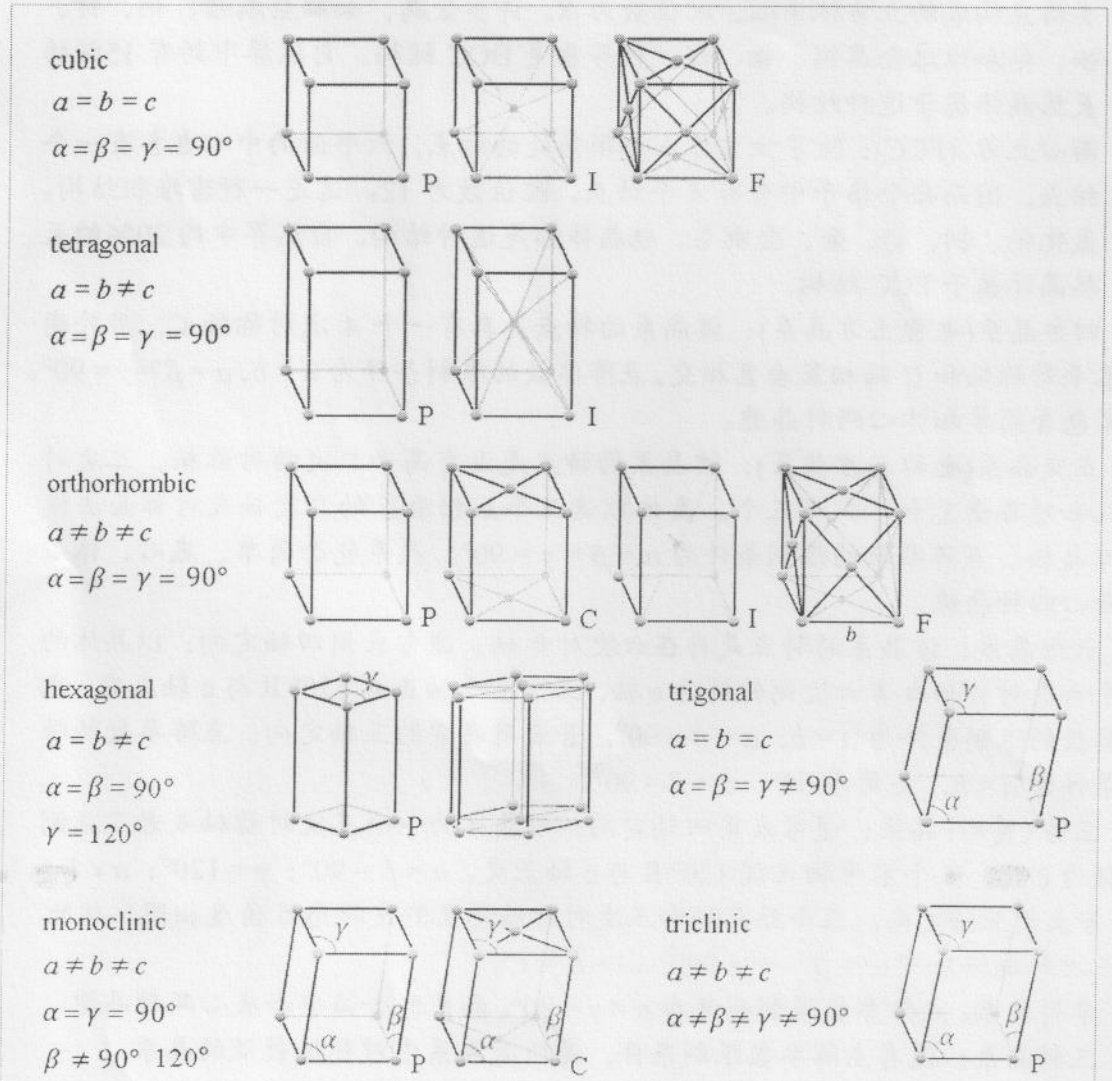


图 6-1-6 晶格分为 7 个晶系，14 类晶胞。

✧ 简单立方 (SC)：结点仅在立方体的八个顶点[4]，每个 SC 格子中包含的结点数为 1。每个结点的最近邻结点数称为配位数，SC 中的配位数为 6。SC 结构的纯元素晶体在自然界中非常罕见，因为该结构通常不对应于能量最低的状态：它对于切变通常是不稳定的。至今在正常条件下发现的唯一例子是钋的 α 相。SC 结构的晶体大多为化合物，如氯化铯，溴化铯，碘化铯，氯化亚钛，溴化亚钛和氯化氨等。图(6-1-7)为氯化铯的点阵结构，其中正负离子 Cs^+ 与 Cl^- 都分别构成 SC 结构。

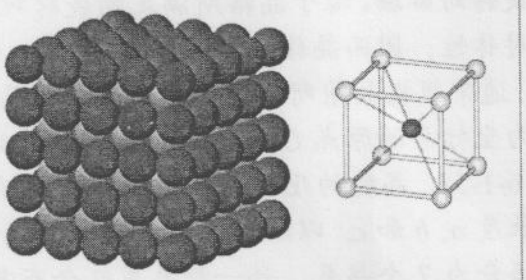


图 6-1-7 氯化铯晶格和晶胞，Cs^+ 与 Cl^- 各自构成 SC 结构。

✧ 体心立方 (BCC)：除八个顶角外，立方体中心也有一个结点，因而每个格子中含有 2 个结点。每个结点在该结构中的位置是完全等价的：都处于 8 个同类格点构成的立方体中心，配位数为 8。许多金属，如碱金属锂、钠、钾、铷、铯和难熔金属钨、钼、铌、钽等都是 BCC 结构。自然界中约有 15%的天然晶体属于这种结构。

✧ 面心立方 (FCC)：除了立方体八个顶角处的结点，六个面的中心也各有一个结点，因而每个格子中含有 4 个结点。配位数为 12。这是一种密堆积结构。氯化钠、铜、铅、金、金刚石、硅晶体都是这种结构。自然界中约 20%的天然晶体属于 FCC 结构。

(2) 四方晶系(也称正方晶系)：该晶系的特点是具有一个 4 次对称轴 C，另外两个水平对称轴和 C 轴相互垂直相交。点阵参数的限制条件为 $a=b, \alpha=\beta=\gamma=90°$。晶系包含简单和体心两种晶胞。

(3) 正交晶系(也称斜方晶系)：该晶系的特点是没有高于二次的对称轴，二次对称轴和对称面总和不少于三个。晶体以这三个互相垂直的二次轴或对称面法线为结晶轴。点阵参数的限制条件为 $\alpha=\beta=\gamma=90°$。晶系包含简单、底心、体心和面心四种晶胞。

(4) 六方晶系：该晶系的特点是存在六次对称轴。通常采用四轴定向：以晶体的一个六次对称轴或者六次倒转轴为 c 轴，三个水平轴正端 120°且与 c 轴正交。点阵参数的限制条件为 $a=b$，$\alpha=\beta=90^0$，若采用通常的三轴定向，点阵参数的限制条件为 $a=b$，夹角为 60°，$\alpha=\beta=90^0$。

(5) 三方(菱形)晶系：通常采用四轴定向：以晶体的一个三次对称轴或者三次倒转轴为 c 轴，三个水平轴正端 120°且与 c 轴正交。$\alpha=\beta=90°$；$\gamma=120°$；$a\neq b\neq c$。若采用三轴定向，三个结晶轴和三次对称轴都呈斜交状态且角度相同。彼此绕三次对称轴分布 $\alpha'=\beta'=\gamma'\neq 90°$，$a'=b'=c'$。

(6) 单斜晶系：点阵参数限制条件为 $\alpha=\gamma=90°$。晶系包含简单和底心两种晶胞。

(7) 三斜晶系：没有点阵参数限制条件，是所有晶系中对称性最低的晶系。

[4] 由于其结构简单，常用于作为估算理想晶体性质的模型。

密堆积

元素金属晶体中的原子电荷分布都具有很好的球对称性，如果将晶体看成大小相同的原子球堆积，可以定义堆积密度f为原子球的体积与其所占有的有效空间体积之比。简单立方结构和体心立方结构的堆积密度分别为$f = \pi/6$，$3^{1/2}\pi/8$。体心立方和简单立方结构都不是最密堆积方式。图 6-1-8(*a*)所示为原子球在一个平面内的最紧密排列方式，称为**密排面**。为了使密排面层层相叠形成堆积最密的晶格，上一层的球心必须对准下一层的球隙。将第一层原子球心位置记为 A，上面各层球体密堆积时可处的所有可能位置只有三类：A 位：第一层球所在位；B 位：三角尖向上的空隙；C 位：三角尖向下的空隙。第二层球只能堆积在第一层球的空隙位置上才最紧密，因而只能全部处在 B 位(见图 *b*)或 C 位，记为 AB 或 AC。这两种结构没有本质区别，因为 AB 旋转 180°与 AC 重合。对 AB 堆积，第二层球形成的三角尖向上的空隙为 C 位，三角尖向下的空隙为 A 位(与第一层球所在位相同)。第三层球可以堆积在 C 位上，形成 ABC 结构，也可堆积在 A 位上，形成 ABA 结构。

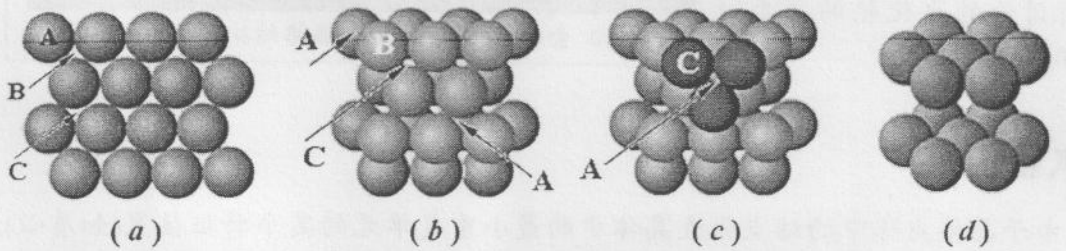


图 6-1-8 (*a*) 密排面及各密排面中原子球可处的三类位置，(*b*) 第二层密排面原子处于 B 类位置，(*c*) 第三层密排面原子处于 C 类位置，(*d*) 第三层密排面原子处于 A 类位置。

以上两种结构可通过多种组合形成不同的三维密堆积结构，如 ABACBABC 的八层堆积结构中，前三层构成 ABA，2-3-4 层构成 ABC，4-5-6 层构成 ABA，第 6-7-8 层构成 ABC 堆积。而对晶体而言，堆积方向的周期性限制了上述组合，只能按...ABCABC...序列或...ABABAB...序列堆积。

按...ABCABC...序列堆积构成面心立方(图 6-1-9*a*)，称为立方密堆积。结点处于立方体顶角和每个面的中心。图中用不同颜色表示密排面，堆积方向沿立方体体对角线方向。每个结点的环境(形式和取向)都相同，每个结点的配位数为 12，$f = 2^{1/2}\pi/8$。很多金属如铜，银，金，铝，镍等晶体具有 FCC 结构。

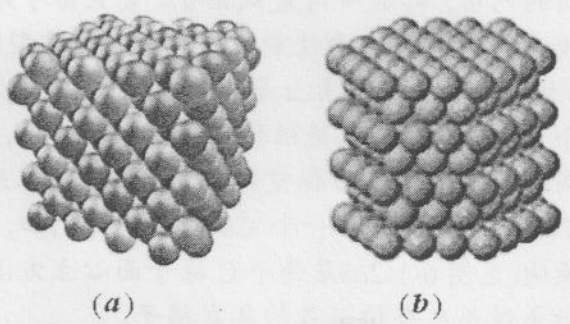


图 6-1-9 (*a*) 立方密堆积，(*b*) 六方密堆积。其中以不同颜色标明不同堆积层。

如果结点按...ABABAB...序列堆积，形成的结构具有六度旋转对称，称为六角密堆(HCP)结构(图 6-1-8*b*)。该结构的配位数为 12，$f = 2^{1/2}\pi/8$。尽管都是同一种原子。但 A 位和 B 位原子周围的原子排布和取向不同，因为从 A 原子看，上、下两层 B 位原子构成的三角形与从一个 B 位原子看，上、下两层 A 位原子构成

的三角形取向相差 180°，因此 A、B 两类原子在该结构中不等价。许多原子，如铍、镁、锌、镉、钛等构成的晶体具有 HCP 结构。

离子晶体中的正负离子一般也有很好的球对称电荷分布，但正负离子大小不同，离子晶体可视为两种不同大小的圆球堆积，大球先按 FCC 或 HCP 方式密堆积，再将小球填充在大球堆积形成的空隙中。图 6-1-10 分别为金和氯化钠的 FCC 晶体，在 NaCl 晶体中，较大的 Cl^-离子先构成立方密积，较小的 Na^+离子充填于所有八面体空隙中。对比金的 FCC 结构图的可看出，Au 原子的分布相当于 NaCl 结构中 Cl^-离子的分布。

对不同尺度的原子球的堆积还需要考虑小球的尺度是否与大球密堆积形成的空隙大小匹配，如果不能匹配就可能破坏负离子形成的最密堆积结构，如前面所讨论的氯化铯的简单立方结构。

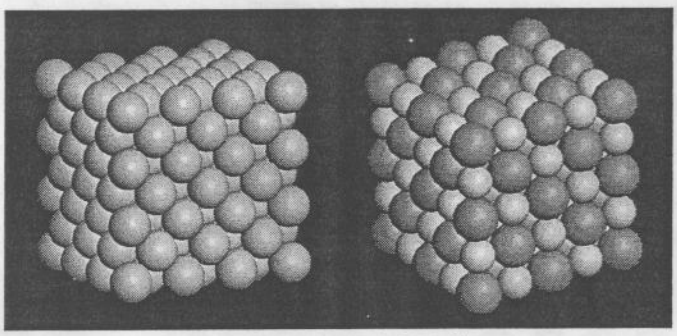


图 6-1-10 金和氯化钠的 FCC 晶格结构。

复式格子

由于晶格点阵中的结点代表晶体中的最小重复单元的某个特征位置(如质心)，它可以包含若干个原子、离子、分子或原子集团，因而每个原子、离子或分子都各自构成完全相同的原胞，称为**复式格子**，即复式格子是由若干相同的原胞互相位移套构而成。

如果晶体由同一种原子组成，而且每个原子周围的情况都相同，该原子构成的网格和结点所组成的点阵相同，称为**布喇菲格子**；如果晶体的基元中包含两个或更多的原子，每个基元中的每个原子在基元堆垛成的晶格中各自构成和晶格点阵相同的网格，形成所谓**复式格子**。复式格子是由若干相同的布喇菲格子相互位移套构而成。例如，氯化铯(CsCl)是一种典型的离子晶体，其结晶学原胞(见图 6-1-7)为立方体，其顶角上是 Cl^-，立方体中心为 Cs^+(也可取顶角为 Cs^+，体心上是一个 Cs^+离子)，氯化铯结构不是“体心立方”，而是 Cl^-和 Cs^+各自组成简立方，是由两个简立方沿立方体空间对角线位移½长度套构而成的复式格子，原胞中包含一个一个 Cs^+离子和一个 Cs^+离子。再例如，金刚石是典型的共价晶体，其结晶学原胞(见图 6-1-2*a*)是两个 C 原子面心立方沿立方晶胞的对角线位移套四分之一的对角线长度套构而成的复式格子。

晶面

由于晶体中的格点分布具有周期性，可以证明，晶体中的任意三个格点构成的平面中一定包含有许多格点，而且该平面上格点的分布也具有周期性，该平面因而称为**晶面**。另外，一个晶体中所有相互平行的晶面的集合一定包含了晶体中的全部格点，而且这些的平行晶面的排列也呈现出周期性。所有这些相互平行晶

面的集合称为**晶面族**。

选取晶体中某一格点为坐标系原点 O，将 O 点为顶点之一的某个晶胞的三个基矢 **a**, **b**, **c** 定义为坐标系的坐标轴方向和相应的长度单位。则一个晶面可用它在三个坐标轴上的截距(图 6-1-5)表示。由于一个晶面族中通常总有两个晶面分别通过基矢的两端，设某一晶面族中的若干晶面将基矢 **a**, **b**, **c** 分别截成 h, k, l 个相等小段。则距 O 点最近的晶面在三个坐标轴上的截距分别为 $a/h, b/k, c/l$，而同族中其它晶面的截距是这组最小截距的整数倍。由此一组互质的整数 h, k, l 不仅可用于表征一个晶面，还可用于表征相应晶面族的(法线)取向，称为该晶面族的**密勒指数**，记为[hkl]。例如，三个截距分别为(4,1,1)(取晶胞基矢为相应的长度单位)的晶面，其截距的倒数为(¼,1,1)，相应晶面族的密勒指数为[1,4,4]。三个截距分别为(2,4,∞)的晶面，截距的倒数为(½, ¼, 0)，晶面族的密勒指数为[2,1,0]，再如三个截距分别为(–3,–1,2)的晶面，截距的倒数为(–⅓, –1, ½)，相应的密勒指数记为$[\bar{2},\bar{6},3]$。

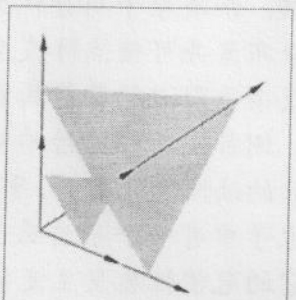


图 6-1-5　晶面和平行晶面族

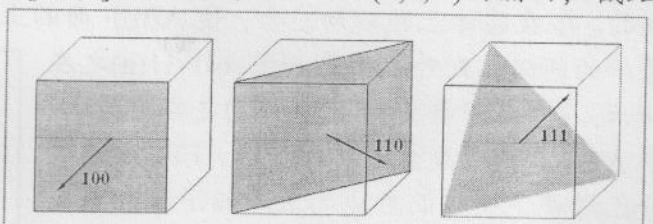


图 8-5-4 立方晶系的 (100), (110)和(111)晶面

简单的密勒指数(如[100], [110], [111]等)所标示的晶面族面间距比较大，因而晶面上的格点密度也大，相应地单位表面能量就比较低。这样的晶面容易解理，晶体生长过程中往往显露为晶体的外表面。但需要注意的是，由于采用晶胞基矢 **a**、**b**、**c** 作为三个方向的最小平移单位，对于简单晶格，以互质整数(hkl)标示的晶面就是相应晶面族中最靠近原点的晶面。但对含有体心、面心这样的复杂晶格而言可能会遗漏晶面族中的部分晶面。或者说，以互质的密勒指数(hkl)所标示的晶面并不一定就是该晶面族中最靠近原点的晶面。例如 FCC 晶格中，(100)晶面族(见图 8-5-5)中距原点最近的晶面并不是 FCC 晶胞的前表面(100)，而是(200)晶面，这正是以 **a**、**b**、**c** 为基矢作平移所遗漏的面心结点所在的晶面。

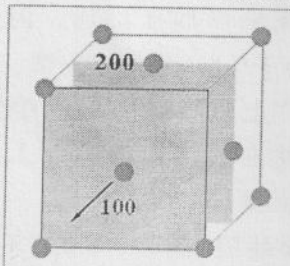


图 8-5-5 FCC 晶格中的(100)晶面族中距原点最近的晶面是(200)而不是(100)。

晶体结构缺陷

理想晶体具有严格、完整的周期性。但实际晶体中的原子的排列或多或少地总会偏离严格的周期性，这种偏离称为**缺陷**。缺陷的存在会影响晶体的力学，电学，光学特性。从能量角度考虑，各种缺陷的状态对应于不同的亚稳态。晶体的缺陷主要可分为三种类型：

点缺陷 若晶体中的原子进入晶格中其它原子的间隙位置形成**填隙原子**，或者跑到表面上而在内部留下**空位**。外来原子也会进入晶体的相应位置，形成为**填隙式杂质**或**替位式杂质**(见图 6-1-10)。这些出现在晶体某些点上的缺陷统称为**点**

缺陷。杂质原子与母体原子的电荷分布差异可使点阵发生畸变，导致杂质附近的势能曲线发生改变。例如半导体的性质对杂质所引起的缺陷十分敏感，平均每 10^9 个原子中有一个杂质原子即足以使它的电学性质发生变化。

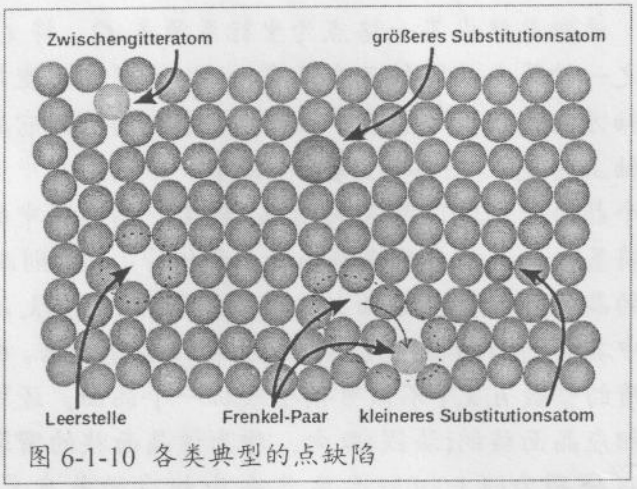
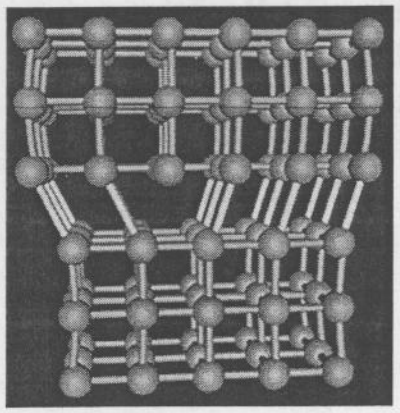


图 6-1-10 各类典型的点缺陷

线缺陷 若晶体中的相邻晶面发生局部相对滑动产生的缺陷将以“线”的形式出现，称为**位错**。常见的位错有刃型位错与螺型位错两种形式，现以刃型位错为例来说明晶体内部如何由相邻晶面相对滑移而产生位错的。若在晶体上施以切应力，使 ABEF 面的上下发生局部滑移，且滑移了一个原子的间距，则外观上看如图 6-1-11(a)之左图所示，内部原子排列情况如图(a)之右图所示。图中 A'B/EF 是已发生滑移的区域，EFGH 左边部分已滑移了一格而右边尚未开始滑移，因而在 EFGH 处多挤进了一层原子，致使在(a)之右图“⊥”处附近，其上部的原子受到压缩，下部的原子被拉伸。但在“⊥”处附近区域以外部分仍能保持周期性排列。因此在“⊥”处发生了位错。由于“⊥”代表一条线，故位错属于线缺陷。在“⊥”上方插进半层原子宛如插进一把刀，而位错就在这把刀的刀口上。因而这种位错称为刃型位错。



图

面缺陷 晶界、相界(不同结构的相边界)、堆垛层错(正常堆垛次序发生错误)等二维缺陷称为晶体中的面缺陷。

固体中具有完整晶格结构的小区域称为晶粒。不同取向的晶粒接触面称为晶界。晶体中的晶粒边界缺陷与晶体的生长过程有密切关系，图 6-1-12 显示的是一个容易生长和观察的简单二维晶体，其中 a 是一完整的晶体，而 b 则含有缺陷。晶体开始生长时，当物体冷却并由液体向固体转化时，晶格在一个区域中沿某一个方向开始生长，而在另一个区域中则沿另一个方向生长。一旦晶格开始生长，它就成为其它原子排列的模版，但如果存在两个不同取向的生长中心，就不会再长成一个单晶体，

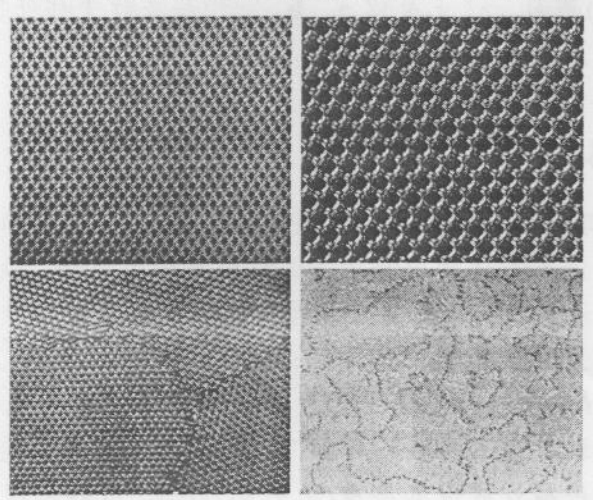
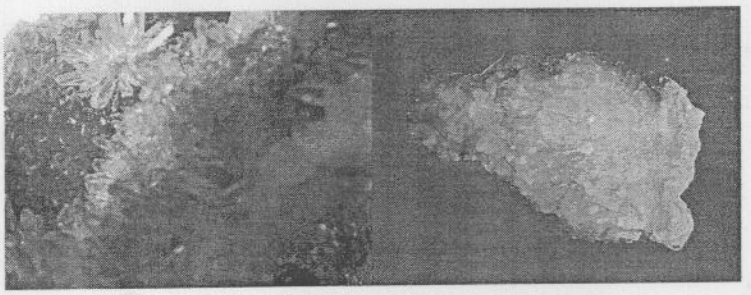


图 6-1-12

而是形成两个晶粒。两个晶粒的边界就形成了一种特殊的缺陷。图 6-1-12c 显示了这类边界。实际的固体是由许多晶粒构成的(见图 6-1-12d)。

多晶体

许多天然晶体矿物(如常见的金属)并不表现出其应有的规则外形。它们是由许多线度约 10^{-4}m~10^{-6}m 的微晶粒组成的。其中每颗晶粒也具有单晶体的上述宏观特征，但微晶粒之间的相互排列方向则毫无规律可言，这样的晶体称**多晶体**。图 6-1-6 是石英晶矿石和方解石晶矿石的照片，相当于多晶体的形貌放大图形。



6-1-6 石英矿及方解石矿

在多晶体中晶粒的分界面附近区域(晶界)，由于相邻晶粒的取向不同。晶界层的原子排列远较晶粒内部不规则和疏松，且易聚集杂质，故晶界可视作晶体中的一种面缺陷。正因为晶界易受侵蚀，多晶体中晶粒的轮廓才得以在显微镜下显示出来。另外，多晶体中大量小晶体彼此间取向的随机分布(称为镶嵌结构)以及晶粒之间的边界(称为晶界)也都是缺陷。

非晶态固体

物质的无规则非晶态结构也称为无定型结构。非晶体中的原子或分子虽然没有周期性那样的长程有序特征，但其分布也表现出一定的规则。由于同一种化学成分构成的晶态材料和非晶态材料(如晶态硅和非晶态硅)中原子的特性是相同的，非晶态中长程序破坏一般体现为结点间距较大，因而

只是因为，取向无序，拓扑无序。

6-1-3 液体的结构及特征

液态与气态不同，它具有一定的体积。与固态不同，液态有流动性，因而没有固定的形状。与非晶态固体一样，液态在物质特性上呈现出各向同性。但初看起来，液体相比于固体似乎更类似于气体，因为它不象固体那样坚硬，容易变形甚至易于流动，但从其微观结构观察，液体更为接近于固体：两者内部的原子都非常靠近，配位结构几乎相等，只是原子之间的距离更大，因而原子之间的聚合力小于固体，分子的运动具有更大的自由度。因而液体容易变形。

6-1-4 气体的结构及特征

气体中分子间距比液体中更大，彼此之间的内聚力可以忽略，因而容易变形，

也容易被压缩，处于容器中时可以充满整个容器。气态物质中的分子动能很大，一般远大于分子间的势能，因此相互作用的过程主要表现为动态的碰撞过程。

由于不同气体的聚集形态结构相同，它们所表现出来的宏观力学特性也非常相近。例如它们都有很大的压缩率，切变模量很小，因而具有很好的流动性。

6-1-5 其它结构形态

以上所讨论的物体都是无机材料，其所涉及单元大都是原子或小分子，它们的三维尺度都很相近。但有机材料通常具有一定的长度，而且两头通常具有极性，这种结构导致一些和通常的无机材料完全不同的结构。这里只讨论两种典型的结构：液晶和高聚物。

液晶

1888 年奥地利植物学家赖尼策尔(F. Reinitzer)将胆甾醇脂类化合物晶体加热到 145.5K 时熔化成一种黏稠的混浊液体，但在光学性能上却呈现出各向异性。继续加热到 1785°C 时，液体变得透明、光学各向异性也随之消失。另外在加热或冷却过程中还呈现出颜色的变化。德国物理学家列曼(O. Lehman)以后陆续发现许多有机化合物也都和粗甾醇脂类似，从固态变为液态的过程中经历一个具有液体流动性，但又呈现光学各向异性的状态[5]，列曼由此将这种特殊的状态称为**液晶态**。

能呈现液晶态的有机化合物都是一些类似于一种不易弯曲的棒状分子(见图 6-1-8)，分子中含有一些呈现极性的基团。由于这些极性分子之间的吸引作用，结晶固体中的这些分子按一定的规律有序地排列。温度升高到熔点 T_1 时。原来的晶体结构受到破坏，但此时处于液体状态的分子在分子力的作用下仍保持一定的有序性：分子按其长轴方向互相平行或接近平行地作三种排列方式：

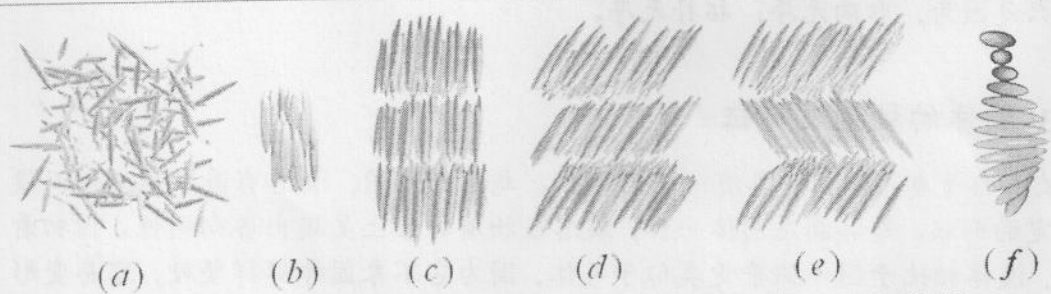


图 6-1-8 (*a*)液晶分子的液相，(*b*)向列形相，(*c*)，(*d*)，(*e*)近晶型相，(*f*)胆甾型相。

(1) 分子按交错平行排列，称为**向列型液晶**。分子除可转动和来回滑动外还可上下滑动，因而黏度比较低，流动性好。

(2) 分子作分层排列，每一层中的分子彼此保持平行，因分子排列较整齐，近似于晶体，称为**近晶型液晶**。分子在层内可自由转动或滑动，但不能在上下层间之移动；因而分子运动受到较大约束，表现为黏度大，流动性差。

[5] 通常称 T_1 为熔点，称 T_2 为清亮点。

(3) 分子也作层状排列，但各层之间的分子取向不同，相邻层之间偏转一固定的角度，使重叠的各层中的排列方向形成螺旋状结构。称为**胆甾型液晶**。

以上三种液晶相中，分子排列都在某一方向上呈现长程有序性。温度升高到清亮点 T_2 时。分子热运动破坏了棒状分子相互平行排列的这种有序性，化合物变为透明的液体。

除了上面介绍的液晶是在一定温度范围内呈现出液晶态(通常称为**热致液晶**)，另有一类液晶是由某些化合物溶解于溶剂中而获得：在某些合适的浓度范围内呈现液晶态，且可在不同的浓度范围内呈现不同的液晶态，这种液晶称为**溶致液晶**。例如肥皂水溶液中的，表面活性分子(见 6-3-2)是由亲水的头部和疏水的尾部组成的条形分子，它们随浓度增加可产生片状或六角柱状液晶态。

高聚物

常见的如塑料、橡胶、纤维等一类的材料是由许多相同的、结构简单的分子单元通过共价键连接而成的**高分子量化合物**，通常简称为**高聚物**。例如聚氯乙烯分子是由许多氯乙烯分子结构单元-CH_2CHCl-相互连接而成。其中的单元的重复数 n 又称为**聚合度**。聚合度很低的(1~100)的聚合物称为**低聚物**，n 达到 10^4-10^6 才称为高聚物。由一种单体聚合而成的聚合物称为**均聚物**，如聚乙烯、聚氯乙烯等。由两种以上单体共聚而成的聚合物称为**共聚物**，如氯乙烯-醋酸乙烯共聚物等。

聚合物的结构可分为**链结构**和**聚集态结构**两大类。

(1) 链结构又分为近程结构和远程结构。近程结构包括构造与构型，构造指链中原子的种类和排列、取代基和端基的种类、单体单元的排列顺序、支链的类型和长度等。构型是指某一原子的取代基在空间的排列。近程结构属于化学结构，又称一级结构。远程结构包括分子的大小与形态、链的柔顺性及分子在各种环境中所采取的构象。远程结构又称二级结构。链结构是指单个分子的形态。

(2) 聚集态结构是指高聚物分子链之间的几何排列和堆砌结构，包括晶态结构、非晶态结构、取向态结构以及织态结构。结构规整或链次价力较强的聚合物容易结晶，例如，高密度聚乙烯、全同聚丙烯和聚酰胺等。结晶聚合物中往往存在一定的无定型区，即使是结晶度很高的聚合物也存在晶体缺陷。结构不规整或链间次价力较弱的聚合物(如聚氯乙烯、聚甲基丙烯酸甲酯等)难以结晶，一般为不定型态。

6-1-5 长程有序与短程有序

下面我们探讨处于非晶态或液态的物体的内部结构中可能存在的规律性及其数学描述。

晶体结构中的规律性主要表现为其在较大的空间尺度内具有平移周期性，即使对于仅 10μm 线度的微晶粒而言，也存在数千个粒子沿直线规则排列，这种规律称为**长程有序**。而在非晶态(固体)及液体中不存在这种特性，但在几个原子的

空间尺度内，其排列也显示出一定的规则性：每一个结点的近邻原子数(配位数)是大致固定的，这种规律称为**短程有序**[6]。

以第 i 个原子为中心，半径在 r 与 $r+\Delta r$ 之间的球壳内 j 原子的数目为 $\Delta N_i(r) = 4\pi r^2 \rho_{ij}(r)\Delta r$，其中 $\rho_{ij}(r)$称为**径向分布密度函数**，相应的物质系统中的微观粒子的排列规则或秩序可以通过 $\rho_{ij}(r)$予以描述。图 6-1-7 是锗在三种不同状态中的径向分布函数。其中离散的短线对应于单晶体锗，振荡曲线对应于非晶态锗，曲线的峰值分别给出最近邻，次近邻与第三近邻的位置。配位数由积分

$$\int_{r_1}^{r_2} 4\pi r^2 \rho(r) dr$$

求得。上述积分实际给出的时半径在 r_1 与 r_2 的球壳内的原子数。图 6-1-6 是用 X 射线衍射方法对液体汞(曲线 a)及晶体汞(曲线 b)测得的径向分布函数。图 6-1-7 的峰值位置为 $r = 2.35$Å。峰的宽度给出原子间距的径向起伏，第一个峰下面的面积为 4，与晶态锗的情形完全相同；第二个峰下的面积为 11.62±0.5，与晶态锗中的次近邻数 12 相近。图中另一曲线代表均匀密度，即 $\rho(r) = \rho_o$ 物质的径向分布函数曲线。当 r 逐渐增大时，曲线振荡的振幅也越来越小，峰值越来越不显著。这些都说明非晶态锗是短程有序，但没有晶态物质的长程序。

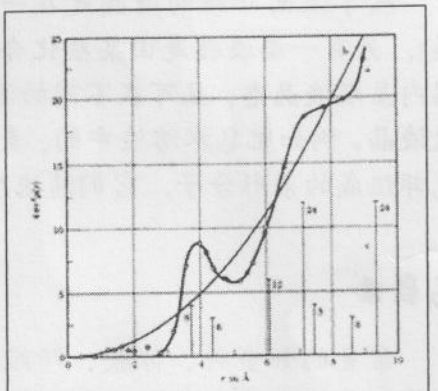


图 6-1-7 锗在三种不同状态中的径向分布函数.

物质的径向分布函数 $\rho(r)$也可以由实验测定，其中一种常用的方法就是利用(X 射线、电子或中子)衍射测量物质的：以任意一个分子的质心为中心画出一系列同心球面，其中相邻两个球面围成的薄层球壳体积相等，每个球壳中的平均分子数与球壳半径 r 的函数关系就是径向分布函数。

晶体熔解时其体积的增加率通常在 10%左右，因此液体分子间平均距离比固体约大 3%。这说明液体分子虽然类似于固体中的情形一个紧挨一个排列而成，但不是严格周期性的堆积，而是一种较为疏松的随机堆积(见图)，与周期点阵比较，这种几何结构最明显的特征是长程无序、短程有序。为简单起见，考虑一个二维系统的例子。对相互作用呈各向同性的原子而言其晶体结构可以描述为由大小相同的刚性球形成的密堆积(图 6-15-a)：每一个粒子周围有六个最近邻粒子。但若先在某个中心粒子周围排列五个粒子，然后由内向外按每一个原子周围均有五个近邻粒子的方式排列(图 b)，得到是比较疏松的结构，而且离开中心粒子越远，粒子的排列也越杂乱，粒子间的空隙也越大。系统仅在中心粒子周围几个粒子直径的线度内反映出排列的有序性(即每个中心原子周围有五个近邻原子)。液体中的情形大致如此。如果将能反映出一定的排列规律性的粒子的群体看成一个单元。液体中同一单元内的粒子排列取向相同，相邻单元中粒子排列取向可以不同，因此液体表现出短程有序、长程无序的特征。

液体的有序结构同样可以由物质的径向分布函数 $\rho(r)$描述：以任意一个分子的质心为中心画出一系列同心球面，其中相邻两个球面围成的薄层球壳体积相等，每个球壳中的平均分子数与球壳半径 r 的函数关系就是径向分布函数。图 6-16

[6] 关于短程有序，在选读材料 6-3 中的非晶态及 6-2-I 的液体微观结构中还要作详细介绍。

是用X射线衍射方法对液体汞(曲线a)及晶体汞(曲线b)测得的径向分布函数。曲线b中出现的峰对应于汞晶体中原子的规则排列：近邻、次近邻及次次近邻…。而液体的曲线a在距离足够大($r>1\times10^{-9}$m)时$\rho(r)$趋向一条水平直线，说明液体确实和气体一样具有长程无序性。但是在$r<0.9\times10^{-9}$m时曲线却有了起伏。第一个峰出现的中心位置与晶体衍射第一个峰出现位置十分相接近。同样第一个谷出现的中心位置与晶体衍射第二个峰出现的位置差不多。这说明在短程(几个分子直径线度)的范围内，液体具有与晶体相类似的有序性。液体在小范围内出现"半晶体状态"的微观结构。每个液体分子周围由最邻近的分子围绕着，形成某种规则的几何构形，但这种规则性只能维持在几个分子直径之内。即使这样的几何构形也是各不相同且变化不定的。不仅不同分子周围的几何构形会有差异，而且任何一种几何构形保持一定时间后均会被破坏。这种具有局部的"结晶结构"的单元有时被称为**类晶胞**。由于液体分子排列得比较松散，液体内部就有许多微小的空隙，这些空隙就为"类晶胞"的拆散与重组创造了条件。

思考题

6-1-1 说明物体的惯性：惯性质量和转动惯量也是物体对外部扰动的响应函数，将其表示为微商或偏微商的数学形式。

6-1-2 为什么一种晶体的解理面总是具有特定的方向，从物理角度解释晶面角守恒定律。

6-1-3 说明能使相同的粒子在空间分布形成相互作用能量达到最低的排列一定是某种规则的周期结构，因而大多数由简单基元构成的固态物质都呈现出三维的周期点阵结构。

6-1-4 怎样从数学上定量说明一种晶胞的对称性比另一种晶胞的对称性高？

6-1-5 如果相同的粒子在空间分布可以形成若干种不同的周期结构，说明高温结构的周期点阵具有较高的对称性。高压结构的周期点阵具有较低的对称性。

6-1-6 为什么单质物体很少采取简单立方结构，却有较多化合物采用这种结构？

6-1-7 说明只有原子或很小的分子才能形成简单结构的单晶体

6-1-8 说明为什么只有高对称的晶系才有体心，面心的晶胞。

6-1-9 你能判断上述两种冰的外部条件差异吗？

6-1-10 辨识金刚石结构的晶胞类型。

6-1-11 说明晶体中不能具有11, 10, 9, 7, 5的配位数，说明配位数是4, 3, 2的晶体结构。

6-1-12 为什么金属的晶体结构大多为面心立方、体心立方的高配位数结构？

6-1-13 为什么离子化合物的晶体结构大多为面心立方、体心立方的高配位数结构？

6-2 固态物质的力学性质

固体的力学性质是指固体材料在外力作用下所呈现出的各种响应行为。这类响应在物体内部直接表现为物质粒子不同程度的相对位置变化，而在宏观上则可呈现为物体不同形式和程度的变形甚至破坏。这说明物体的力学特性原则上都可以表示为固体内部各个基元粒子的位置和运动状态的函数。但这种描述显然很不方便，也无可能。如果所讨论的问题不涉及到物体中基元粒子的差异和排列等细节，可以采用若干宏观参量来量度固体的宏观行为，例如用密度描述物体中的质量分布特性，用声速作为描述高频响应特性的特征参量，而描述静态响应特性的特征参量则是各种不同的弹性模量。而且能通过一些具体条件下的实验结果归纳出描述固体的力学行为的经验规律。实验表明，这种描述对于空间尺度大于10^{-6}m，时间尺度大于10^{-12}s的力学效应都是正确的[7]。

描述系统状态宏观特性的参量可以分为两类：如果将一个处在力学平衡态或稳定状态的系统分为两部分，一个宏观参量对其中一部分的数值与对整个系统的数值相等，该参量称为**强度量**。而一个宏观参量对其中一部分物质的数值等于对整个系统的数值的相应比例，该参量称为**广延量**。显然，广延量与系统的物质数量(或质量)成正比，强度量与物质数量无关。质量m、粒子数N、体积V都是广延量，而压强P、密度ρ、弹性模量Y为强度量。另外，强度量可以用质量、摩尔数或体积等除广延量得到，如质量密度$\rho = m/V$等。下面的讨论可以看到，描述材料特性的参量一般应采用强度量，而描述与边缘效应有关的物体特性通常用广延量。

晶体是各向异性介质，描述各种晶体性质的物理量一般是不同阶数的张量。即使我们不涉及物体内部的原子结构，这种特性一般仍然需要关注。但本节一般只是关注相关力学特性的主体特征及相互之间的关系，我们忽略因各向异性所表现出来的细节差异[8]，将固体看成是各向同性的连续介质。

6-2-1 应力与应变

在连续介质模型中，物体被描述为大量体积元的集合，由于体积元受到外力作用时可以发生变形，而力作用在体积元的不同位置会产生不同的形变，因而在计算形变时不能移动力的作用点位置[9]。作用在体积元上的力有两种不同的类型，一类是作用在体积元内每个质点上的力，如体积元受到的重力，惯性力等，称为**体积力**；另一类是作用在体积元表面质点上力，称为**表面力**。表面力包括作用在整个物体外表面的**外力**和物体内部相邻体积元之间的作用力(内力)。与连续介质形变相关的主要是表面作用力。

---

[7] 该空间尺度和时间尺度由介质中的弹性波相联系，波长大于10^{-6}m的弹性波频率数量级为10^{12}Hz。

[8] 对晶体的各向异性力学特性有兴趣的学生可以阅读相关的书籍。

[9] 研究力对刚体的作用时，只要保持力的大小和方向不变，可以将力的作用点沿力的作用线移动，刚体的运动效应不会发生变化，但这种性质不适用于变形体。例如，在A, B两处施加大小相等，方向相反、沿同一作用线的力F_1和F_2，物体将发生拉伸变形，如果将这两个力沿作用线平移换位，物体将发生压缩变形。因此力的可传性只限于研究力的运动效应

为了量度作用在体积元表面上的力的分布。设想介质中的两个相邻体积元的界面为 ΔS，这两个体积元之间的相互作用力为 ΔF，定义参量：

$$\sigma = \lim_{\Delta S \to 0} \frac{\Delta F}{\Delta S} = \frac{dF}{dS} \tag{6-2-1}$$

描述作用在单位面积上的表面力特征。σ 称为“**应力**”，其单位是“牛顿/平方米”，通常称为“**帕斯卡**”。由于力 ΔF 可以从不同的方向作用在面元 ΔS 上，我们将应力按面元 ΔS 的法线方向和切线方向进行分解。沿 ΔS 法线方向的应力分量称为**正应力**，切向分量称为(**剪**)**切应力**。由此可知，描述整体应力的物理量是一个二阶张量，可以表示为一个三阶方阵$\{\sigma_{ij}\}$，其中的矩阵元(称为张量分量)σ_{ij} 的第一个下标 i 代表力 $\mathbf{F}$ 的直角分量，第二个下标 j 代表面元矢量 $\Delta\mathbf{S}$ 的直角分量。因此，σ_{ii} 代表正应力，$\sigma_{ij}(i \neq j)$代表切应力。

在实际应用中通常还将正应力分为**张力**和**压力**，分别对应于垂直于表面的拉伸或压缩。例如，绷紧的弦中每一个微元受到的是张力，而浸没在水中的固体所受到的应力是与表面垂直的压力(称为**流体静压力**)。另一个常见的正应力例子是沿一根长圆杆的轴拉伸或者压缩(见图 6-2-1a, b)，分别称为**拉伸应力**和**压缩应力**。

常见的切应力例子是作用在物体两个平行单位侧面上的一对大小相等，方向相反的外力(见图 6-2-1c)。

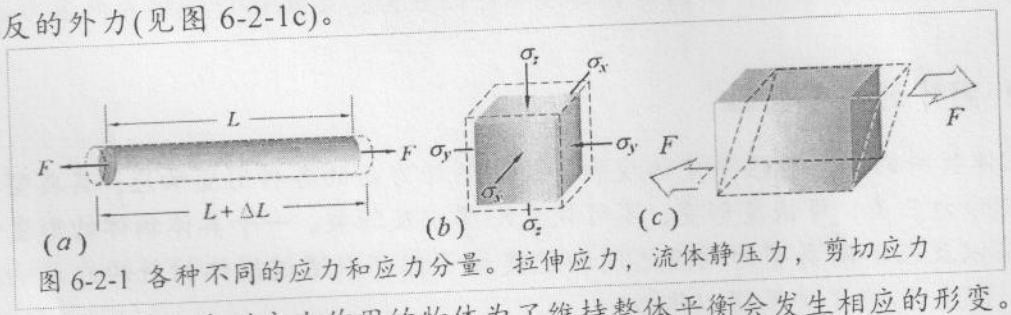


图 6-2-1 各种不同的应力和应力分量。拉伸应力，流体静压力，剪切应力

表面或界面上受到应力作用的物体为了维持整体平衡会发生相应的形变。固体相对原来的大小和形态的变化数量称为**应变**。相应于所受应力的不同类型，固体的应变也分为两种基本类型。第一类为物体在正应力的作用下在应力方向上的尺度相对变化。如一根直杆在轴向正应力的作用下被拉长(见图 6-2-1a)或压缩，直杆的相对伸长或压缩量

$$\varepsilon_{tensile} = \frac{\Delta L}{L} \tag{6-2-2}$$

分别称为**拉伸应变**和**压缩应变**。这两种应变都称为**正应变**。正应变的另一个常见例子为固体在各个方向受到大小相等的正应力[10]时所发生的体积相对变化(见图 6-2-1b)：

$$\varepsilon_{bulk} = \frac{\Delta V}{V} \tag{6-2-3}$$

这种应变称为**体应变**。发生体应变时，各个方向都发生了相应的拉伸或压缩应变，如果各个方向的正应变量都很小：$\Delta L_i \ll L_i$，体应变可表示为：

$$\varepsilon_{bulk} = \frac{1}{L_1 L_2 L_3}[(L_1 + \Delta L_1)(L_2 + \Delta L_2)(L_3 + \Delta L_3) - L_1 L_2 L_3] \approx \frac{\Delta L_1}{L_1} + \frac{\Delta L_2}{L_2} + \frac{\Delta L_3}{L_3}$$

10 实验上容易实现的正应力为各向同性的**流体静压力** P，例如，晶体在自然平衡时所受到大气压强的作用就是各向同性的。

即体应变为三个方向的正应变量之和：

$$\varepsilon_{\text{bulk}} = \varepsilon_x + \varepsilon_y + \varepsilon_z \tag{6-2-4}$$

应变的第二种基本形式是固体的上下底面受到一对大小相等、方向相反的切向应力时发生的形变(见图 6-2-2)，称为**剪切形变**。**剪切应变**定义为平行截面间相对滑动位移 Δw 与截面间垂直距离 L 之比

$$\varepsilon_{shear} = \frac{\Delta w}{L} \tag{6-2-5}$$

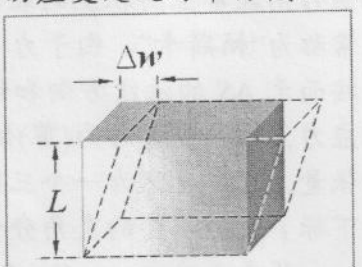


图 6-2-2 剪切形变

由应变的定义可以看出，描述应变的物理量也应该是一个二阶张量，记为$\{\varepsilon_{ij}\}$，其中第一个下标代表位移的直角坐标分量，后一个代表体积元表面的法向分量。因此，ε_{ii}代表正应变，而 $\varepsilon_{ij}\ (i \neq j)$代表剪切应变。应变张量显然也是对称张量，9 个分量中只有 6 个是独立的。

固体材料受到一定的外力作用后会发生相应的形变，而物体也会因各种原因形成的形变导致物体内部出现应力，例如，物体受到正应力被拉伸的同时会出现横向的收缩，这就导致物体内部出现切应力，再例如，物体因内部存在缺陷或温度发生变化发生变形时，其内部会出现相应的应力，这类应力统称为**内应力**。

6-2-2 固体的力学特性

固体材料的力学特性集中体现为物体受到外力后的各种形变响应，其典型特征可以分为三类：可恢复形变，不可恢复形变以及断裂。一个具体物体的形变响应特征以及在不同类型的响应之间的过渡细节都可以通过固体材料的力学响应特性，即应力-应变关系得以体现

The response of a solid to external or internal forces can vary considerably, depending on the magnitude of these forces and the material characteristics. For example, if the forces are great the material may fracture. Lesser values of force may result in material permanent deformation without fracture and, if the forces are low enough, the material may deform only in an elastic way. The treatment of mechanical behavior in this book closely parallels these three possibilities.

固体材料中的应力 σ 和应变 ε 之间的关系一般很复杂，通常和应力的作用方式和过程，持续时间甚至应力加载历史都有关联(见思考题 6-2-5)。在材料科学和工程力学领域，σ-ε 关系通常是对材料在**缓慢加载**的过程中测量在各应力强度环境下的应变数量[11]确定的。

确定材料的各种力学响应试验中最重要的是拉伸试验。图 6-2-4 中的左图给出的是测量材料的拉伸特性时的材料形状和尺度以及测量装置示意图，测量的是材料对静态或缓慢施加的外力的抵抗力。实际测试中通常采用哑铃形试样[12]。在拉伸测试过程中，仪器同时记录施加于样品的作用力(载荷)F和样品的长度 L，再由测得的载荷和位移值及试样的原始横截面面积 A_o 和原始长度 L_o 分别定义工程应力 σ_e 和工程应变 ε_e[13]：

---

11 这里所谓应力的缓慢增加是指应力增加到一个确定的强度后等到形变量稳定

12 陶瓷材料很脆，不易加工成哑铃形的标准样品，一般不作拉伸测试。见 6-2-3 小节。

13 请读者分析工程应力和工程应变与前面所定义的应力和应变之间的差异(见思考题 6-2-3)。

$$\sigma_e = \frac{F}{A_o} \qquad \varepsilon_e = \frac{L - L_o}{L_o} \tag{6-2-6}$$

根据工程应力与工程应变的值后给出应力-应变曲线。图 6-2-4 的中图和右图分别给出的是在对低碳钢和灰铸铁拉伸过程中测得的应力-应变曲线。曲线在不同的应力范围所表现出的迥然不同的形状揭示出材料的不同力学特性。

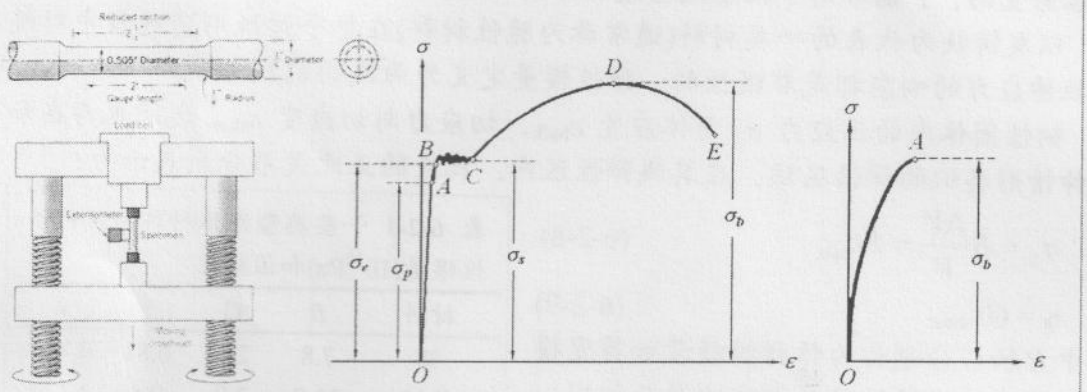


图 6-2-4 低碳钢(*b*)和灰铸铁(*c*)在缓慢加载过程中表现出的拉伸应力-应变曲线。

下面分别就这两类曲线对固体材料主要的力学性能进行简单的介绍并根据响应特性的差异对材料进行分类。

弹性及弹性模量

图 6-2-4 中两条曲线的 *O-B* 段都表现出应力随应变的增加而增加，如果在这一区域中曲线的任一点(状态)逐步减小直至完全撤除外力，物体的形变将沿原路径返回到原来无应变时的形状和尺度 *O*。这种行为呈现出材料具有保持其原有形状和大小，即抵抗形变的能力，这种力学特性称为**弹性**[14]。弹性是固体材料最具有实际用途的力学特性之一。

以低碳钢为代表一类材料(称为韧性材料)的弹性变形分为两个区域，图 6-2-4(*a*)中的 *OA* 直线部分应变和应力成正比，称为**线弹性区**，而 *AB* 区域内应变和应力的关系是非线性的。线弹性区上限 *A* 处的应力值 σ_P 称为**胡克极限**。物体在其**线弹性**区内受到的拉伸应力 σ 和由此表现出的拉伸应变 ε 之间的比例系数 *Y* 称为**弹性模量**。以直杆受到纵向拉伸应力 σ 的情形为例，长度为 *L* 的杆被拉伸 ΔL。其正应变数量与受到的正应力大小成正比：

$$\sigma = Y\varepsilon = Y\frac{\Delta L}{L} \tag{6-2-7}$$

上述现象是由英国物理学家杨[15]在 1800 年首先发现的，因

表 6-2-1 一些固体的杨氏模量

材料	Y(10^{10}Pa)
橡胶	7×10^{-4}
木材	1.4
混凝土	1.7-3.0
骨头	0.9-2.1
玻璃	7.0
铝	7.3
钢	21
钻石	120

[14] 请读者分辨固体的弹性和**刚性**两种概念的差异。

[15] 杨(Thomas Young, 1773-1829) 英国医生、科学家，被誉为“世界上最后一个什么都知道的人”。除了研究物体的弹性，他提出人眼中的晶状体可自动调节以适应所见物体的远近(1793年)，他是第一个研究散光的医生(1801 年)，以后又提出色觉取决于眼中三种不同神经分别感觉红色，绿色，和紫色的理论。杨对血流动力学做出过重要贡献：1808 年提出题为“心脏

此该弹性模量 Y 通常称为**杨氏模量**。表 6-2-1 中列出了一些固态物质的杨氏模量数值。数据表明，钢是弹性极好的材料，因而被广泛用于制造弹簧和建筑中的横梁骨架，甚至直接作为承重的建筑构件，如桥梁，高架中的跨梁。高层建筑中常见的 25 米长直梁(柱)承重达到 10 吨时压缩量仅 1 毫米。由于弹性模量大的材料不容易变形，Y 也常用于描述或量度物体的硬度。

以灰铸铁为代表的一类材料(通常称为脆性材料)在整个弹性形变过程中对缓变拉伸应力的响应都是非线性的，弹性模量定义为曲线的割线斜率。

韧性固体内的正应力 $\sigma_\perp$ 与体应变 $\varepsilon_{\text{bulk}}$，切应力与切应变 $\varepsilon_{\text{shear}}$ 之间也存在和拉伸情形类似的弹性区域，在其线弹性区内，相应的正比关系分别表示为：

$$\sigma_\perp = B\frac{\Delta V}{V} = B\varepsilon_{bulk} \tag{6-2-8}$$

$$\tau_{\parallel} = G\varepsilon_{\text{shear}} \tag{6-2-9}$$

其中 B 和 G 分别称为**体弹性模量**和**剪变模量**。体弹性模量的单位与应力单位相同，剪变模量的单位和杨氏模量相同。

表 6-2-1 一些典型固体材料的三种弹性模量(10^{10}Pa)和泊松比

材料	B	G	Y	ν
铝	7.8	2.5	6.81	0.355
黄铜	13.9	3.8	0.5	0.374
铜	16.1	4.6	12.6	0.37
金	16.9	2.85	8.1	0.42
电解铁	16.7	8.2	21	0.29
铅	3.6	0.54	1.51	0.43
镁	3.6	1.62	4.23	0.306
铂	14.2	6.4	16.8	0.303
银	10.4	2.7	7.5	0.38
不锈钢	16.4	7.57	19.7	0.30
熔凝石英	3.7	3.12	7.3	0.17
聚苯乙烯	0.41	0.133	0.36	0.353

固体在正应力的作用方向出现拉伸或压缩只是形变的一个部分，通常还伴随有横向形变。例如，直杆在轴向应力作用下发生纵向拉伸时伴随有横向收缩，纵向压缩时伴随有横向膨胀(见图 6-2-5)。而且两种形变的数值在固体的弹性限度内有正比关系：

$$\frac{\Delta h}{h_o} = \frac{\Delta w}{w_o} = \nu\frac{\Delta L}{L_o}$$

其中的比例系数 ν 称为**泊松比**。设想一个各向同性的固体受到单轴应力而在受力方向产生应变 ε_1，同时在两个侧向发生相同的应变 ε_2 和 ε_3,满足

$$\varepsilon_2 = \varepsilon_3 = -\nu\varepsilon_1 \tag{6-2-10}$$

由于固体材料受到纵向拉伸(或压缩)时一般伴随有横向形变(式 6-2-6)，固体材料的三种弹性模量 B, G, Y 以及泊松比 ν 互相关联，其中只有两个独立变量。对很多种材料的三种弹性模量以及泊松比 ν 的实验测量结果(见表 6-2-1)表明同一材料在相同温度下都满足关系

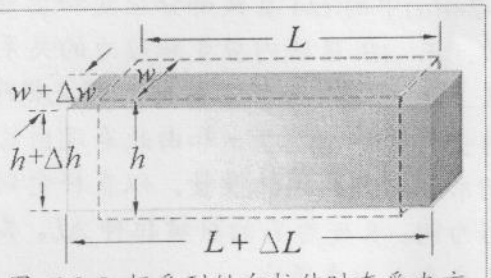


图 6-2-5 杆受到轴向拉伸时在受力方向被拉长，同时侧向发生收缩而变细。

$$G = \frac{Y}{2(1+\nu)},\quad B = \frac{Y}{3(1-2\nu)} \tag{6-2-11}$$

上述关系式也可以从理论上予以证明(见习题 6-2-3)。从这些关系式可以看出，泊松比 ν 必须小于 0.5，否则，弹性模量将出现负值，这将导致一些非物理的效应(见

和血管的功能"的报告；1813 年发表的《医学文学的介绍》和《实践鼻科》；1815 年发表《虚损类病的历史和治疗》等。杨还曾对 400 种语言做了比较，并在 1813 年提出"印欧语系"。

思考题 6-2-2)，典型的弹性介质泊松比数值为½左右。

固体的范性

图 6-2-4 显示，低碳钢在应变超过 B 点后出现了一个较小的近乎平台区 B-C 屈服平台或屈服齿，实验上表现为在维持应力负载几乎不变的情况下材料出现很大的拉伸量，这种现象称为**屈服**，对应这段曲线的最低应力值称为**屈服极限**或**屈服点**。材料的这种特性称为**可延展性**。

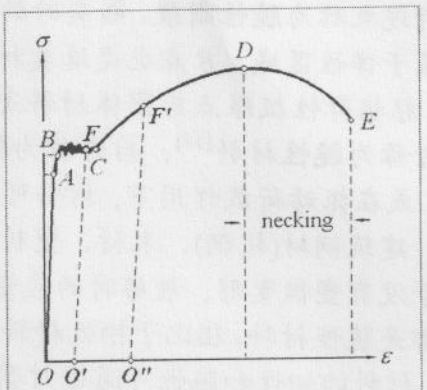


图 6-2-6 低碳钢的范性形变路径及硬化，颈缩时的形变路径。

韧性材料在所受外力超越其屈服点后，对外部应力的响应具有**不可逆性**：如果从某一点 F 开始逐步撤除外力，物体的形变不再能沿原来的拉伸路径 $OABF$ 回到原来的自由状态 O，而是沿平行于 OA 的路径 FO' 减小形变(见图 6-2-6)。外力完全撤去后到达的 O'位置表明材料中残留有剩余形变 OO'。固体的这种力学特性称为**范性**或**塑性**。可逆的弹性形变和不可逆的范性形变的分界点 B 通常称为**弹性极限**。大部分韧性材料的弹性极限与胡克极限非常接近。

如果对已留下范性形变 OO'的材料作再次拉伸，在开始阶段它将沿直线路径 $O'F$ 变化，即服从胡克定律，但外加应力超过 F 点的应力后沿原来的应力-应变曲线的路径 $FF'D$ 变化。若在 F'点撤去外力，它将沿平行于 OA 的直线路径 $F'O''$ 减小形变。使材料的范性形变增加到 OO''。

图 6-2-6 中 BCD 区域都属范性形变范围。$C, F, F'...D$ 各点所对应的持续增长的应力说明材料发生范性形变后会变硬，称为**拉伸硬化**，因而 CD 区域常称为**强化**阶段，其应力的最高限 σ_b 称为**强度极限**或**抗拉强度**。但这一参量在工程设计中用处不大。因为到达极限强度前材料已发生了相当大的塑性形变。塑性形变在多数情况下等同于材料的破坏，但在少数情况下是有利的，如汽车保险扛。

以上讨论的是拉伸引起的范性形变，事实上，其它应力过程，如辗压、锻打都可使材料发生范性形变。

固体的强度

固体的强度是指固体所能承受的最大外部应力，外力达到或超过这个极限，材料就会断裂。不同形式的应力存在相应的强度，如拉伸强度和剪切强度等等。不同的材料在应力达到相应的上限数值时会呈现出不同的断裂行为。

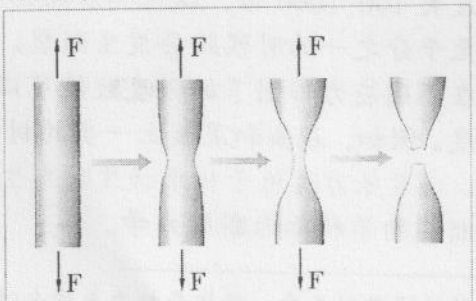


图 6-2-7 颈缩所导致的断裂。

一些如低碳钢、铜等韧性材料在经过应力最大值 D 点后出现局部变形，变形区域内横截面尺寸急剧缩小，这种现象称为**缩颈**。这时物体变形所需拉力相应减小，直至最后

出现塑性断裂(见图 6-2-7)。**断裂状态**对应于图 6-2-6 中的 E 点。

图 6-2-4(b)的灰铸铁应力-应变曲线表现出完全不同于韧性材料的力学响应，灰铸铁在拉伸过程中几乎没有范性形变，并在 B 点毫无先兆地发生突然断裂，这种现象称为**脆性断裂**。断裂时的应力值 σ_b 称为材料的**抗拉强度**。由于 OB 范围都属于弹性区域，B 点也是这类材料的**弹性极限**。

根据弹性极限点后固体材料是发生脆性断裂或范性形变而将固体分为两类，前者称为**脆性**材料[16]，后者称为**韧性**材料。材料的**韧性**通常是指在受到外部冲击、或在振动荷载作用下，材料可吸收较大的能量产生一定的变形而不破坏的性质。建筑钢材(软钢)、木材、塑料等都是比较典型的韧性材料。脆性材料断裂时几乎没有塑性变形，破坏时的应变很小。砖、石材、陶瓷、玻璃、混凝土、铸铁等都是脆性材料。相比于韧性材料，它们对抵抗冲击荷载和承受震动的能力很弱。

材料的韧性和脆性可通过对圆柱体材料受到拉伸时的伸长率 δ 和断面收缩率 ψ 进行定义和量度：

$$\delta=\frac{L-L_o}{L_o}\times 100\% \qquad \psi=\frac{S-S_o}{S_o}\times 100\% \tag{6-2-12}$$

其中 L_o 为圆柱体原长；S_o 为初始横截面面积；L 和 S 分别为材料拉断后长度和断口处最小的横截面面积。伸长率和断面收缩率越大表示材料的韧性越好。工程中一般将 $\delta\geq 5\%$ 的归为韧性材料；$\delta<5\%$ 的归入脆性材料。低碳钢的 δ 为 20~30%，是典型的韧性材料，而铸铁的 $\delta<1\%$，是典型的脆性材料[17]。断裂伸长率 δ 和断面收缩率也常用于量度材料的延展性。

表 6-2-2 一些固体的强度(10^6Pa)

材料	σ_b	σ_{be}	τ
铸铁	170	550	170
钢	500	500	250
黄铜	250	250	200
铝	200	200	200
混凝土	2	20	2
砖块		35	
大理石		80	
花岗岩		170	
松木(平行纹理)	40	35	5
松木(垂直纹理)		10	
尼龙	500		
(肢)骨	130	170	

固体材料所受的拉伸或压缩应力以及剪切应力达到一定大小时都会发生破裂。相应的临界应力称为物体的拉伸或压缩**强度**以及剪切强度。韧性材料在断裂前已发生显著的范性变形，有较好的抵抗拉断能力，其极限应力通常称为**屈服极限**，而脆性材料直至断裂时没有显著的变形，其极限应力通常称为**强度极限**。表 6-2-2 列出了一些典型材料的拉伸强度。值得关注的是，有一些材料对于同样的单轴应力，拉伸应变的弹性模量数值要比拉伸强度大 100~1000 倍，这意味着物体的拉伸量只到达物体本身尺度的百分之一，甚至千分之一的时候就会发生断裂。另外从表 6-2-3 中的列出的一些典型固体材料在不同应力作用下的强度数值可以发现，许多脆性材料的抗压强度远大于抗拉强度。例如，石头和混凝土一类的材料在这两种强度方面的差异达到了一个数量级。以至张力远低于极限强度时发生的所谓断裂称为"低肋强脆性断裂"。研究这种断裂的学科称为**断裂力学**。

---

[16] 脆性通常定义为：物体受拉力或冲击时容易破碎的性质，或材料在断裂前未觉察到的塑性变形的性质。

[17] 材料的塑性或脆性特征并非一成不变，它还与材料所处的温度、应变率和应力状态等条件有关。

工程应力-应变曲线中的应力和应变是以试样的初始尺寸进行计算的，事实上，在拉伸过程中试样的尺寸是在不断变化的，此时的真实应力 S 应该是瞬时载荷(P)除以试样的瞬时截面积(A)，即：S=P/A；同样，真实应变 e 应该是瞬时伸长量除以瞬时长度 de=dL/L。真应力-真应变曲线不像应力-应变曲线那样在载荷达到最大值后转而下降，而是继续上升直至断裂，即金属在塑性变形过程中不断发生加工硬化，从而外加应力必须不断增高，才能使变形继续进行，即使在出现缩颈之后，缩颈处的真实应力仍在升高，这就排除了应力-应变曲线中应力下降的假象。

固体的其它力学特性

以上讨论的固体力学特性都是基于应力-应变曲线的特征，这些曲线特征都隐含有一个基本特征：即物体在受到外部应力后发生应变，并能即刻达到平衡，因此不涉及在外力变化过程中物体尚未达到平衡时的行为，这类与时间有关的响应特性归类为高频响应，将在第 7 章中讨论。

但固体还有一些与时间有关的响应特性并不表现为形变的滞后，而是与应力加载的过程以及材料曾经经受过应力的历史有关。例如应力-应变曲线中会出现图 6-2-7 所示的**回线**结构，这些现象对应的力学过程显然都是不可逆的。

读者可以通过下面一些典型事例"理解"固体材料的这些力学行为及其描述。

疲劳

固态材料或构件在随时间作周期性变化的应力作用下，即使应力峰值远低于极限抗拉应力，经过一段时间后也会出现断裂。这种在重复加载下发生的断裂过程称为**疲劳破坏**[18]。图 6-2-8 分别给出了描述工具钢和合金铝的疲劳特性的应力 S 与破坏周数 N 的关系曲线。其中工具钢曲线在应力降到某一值后成为水平线，即只要应力低于该数值，周期应力持续时间再长，材料也不会受到破坏。该应力值称为**疲劳极限**。通常用材料的拉伸强度作为计量疲劳极限的参考。如铁合金的疲劳极限一般为拉伸强度的一半，铝和其它有色金属的疲劳极限则低得多，有些材料甚至没有疲劳极限，即不存在一个不破坏材料的低应力值。

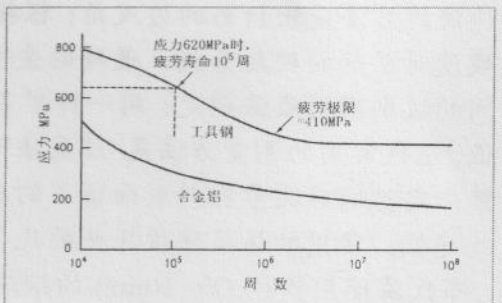


图 6-2-8 工具钢和合金铝的 S-N 曲线

在某一给定的周数中材料不被破坏的应力称为**疲劳强度**。它显然高于或等于疲劳极限。如工具钢的疲劳极限为 410MPa，10^5 周以下的疲劳强度则可达 620MPa。由于实际测量一条 *S-N* 曲线成本昂贵，在材料设计中通常参考描述疲劳性质的另一参数—**耐久比**：疲劳极限与拉伸强度之比。铁合金的耐久比一般为 0.5，其它材料一般在 0.25~0.45 之间。

[18] 如内燃机曲轴、机器传动轴或车轴，在运转过程中都会因疲劳而发生破坏。疲劳与磨损、腐蚀是机械破坏的三种最常见的原因。

蠕变

蠕变是材料在超过弹性限度恒定应力下随时间缓慢塑性形变的过程[19]。**蠕变**进程很慢，但很长的时间也会累积形成明显的形变。蠕变是一种常见现象，如一件厚毛衣经长期悬挂会被拉长，一张揉皱的纸在一本厚书中夹很长时间后可恢复平整；不受支撑的铅会在自身重量的作用下缓慢下垂。蠕变性质也可用于材料成型：将预制件放在模具中，预制件在一定的力学约束下升温至 160°C，通过蠕变充满模具而成形。铝制机翼档板的成型就是利用该方法的典型例子。

有些材料会因蠕变而发生断裂，因此长年负荷运转的部件必须有很高的抗蠕变性。如长年在高温下运转的汽轮发电机桨叶就要考虑因蠕变而发生变形的问题。

蠕变速度可通过蠕变曲线(应变量-时间，图 6-2-9)的斜率描述。通常用“应变百分比/小时”为单位计量材料的蠕变速率，典型值为 10^{-4}%/小时。蠕变曲线可以分为三个区域：减速区，匀速区和加速区。进入加速区后不久材料将发生断裂。研究表明，粗晶粒材料在较高温度下的抗蠕变性优于细晶粒材料；而细晶粒材料在低温下有更好的抗蠕变性。

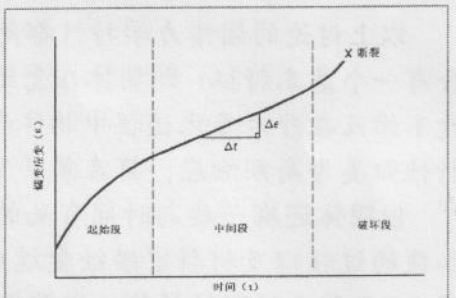


图 6-2-9 蠕变曲线

硬度

固体材料的另一种与时间有关的力学特性是所谓的“抗穿刺能力”，这种性能往往和材料的耐磨性和耐久性有关，通常称为**硬度**[20]。测试硬度有很多方法，最传统的方法是锉材料的边或角：根据锉掉的量来判断硬度。在矿物学中常用**莫式硬度**对矿物的硬度分级：最硬的金刚石定义为十级，最软的滑石定义为一级。中间的级别用划痕法确定：用一种矿石在另一种矿石划痕，被划出痕的材料硬度较低。现代常用的测量方法是“压痕法”：用一个压头以一定的压力压入材料表面，持续一定时间后测量材料表面留下的压痕尺度。压痕面积越大，越深，材料的硬度就越低。常用的压痕法按压头的几何特征分为布氏法、维氏法、洛氏法等。

布氏法使用直径 $D = 10$mm 的钢球或碳化钨球以一定压力 F 压入材料 10~15 秒后，测量压痕的直径 d。布氏硬度(HB)定义为

$$\mathrm{HB} \equiv \frac{F}{S} = \frac{2F}{\pi D(D - \sqrt{D^2 - d^2})} (\mathrm{kg/mm^2}) \tag{6-2-13}$$

式中 S 为压痕的表面积。为得到精确的硬度值，d 必须控制在(0.25~0.5)D。所以对压球直径与压力有一些规定：测定钢材料：$F/D^2 = 30$，测定铜材料，$F/D^2 = 10$。如果材料过软或压力过大可能将球全部压入：$d = D$；而材料硬度过大，会对压球造成损坏。所以布氏硬度的最高测定范围为：钢球：HB 459，碳化钨球：HB600。

维氏法用锥角为 136°的金刚石棱锥作为压头，选择 136°是为了与布氏硬度有对应关系。低于 HB300 的范围内，维氏与布氏硬度基本相等。采用金刚石压头可以不受被测材料的限制。维氏硬度值与硬度有线性关系，即 HV400 的材料硬

[19] 蠕变在任何温度下都会发生，但温度越高过程就越快。

[20] 注意和前面的 stiffness 的测量属性的差异

度恰为 HV200 材料硬度的两倍。维氏硬度的缺点是压痕过小，肉眼难以分辨，表面必须打光，且测定时间较长。

洛氏法采用菱形金刚石压头测定薄层材料的硬度。在同样压力下压痕比维氏的小。洛式法以压痕的深度而不是直径来测定硬度，从而加快了测量时间。洛氏硬度从低到高分成四级，分别记为 HRR，HRM，HRB，HRC，较常用的 HRB 与 HRC 级。其中 HRB 使用直径为 1/16 英寸(约 1.6mm)的钢球，压力为 100 公斤，用于测定有色合金的硬度[图 2-17(d)]；HRC 使用 120°角的金刚石锥压头和 150kg 的压力，用于测定钢的硬度[图 2-17(e)]。

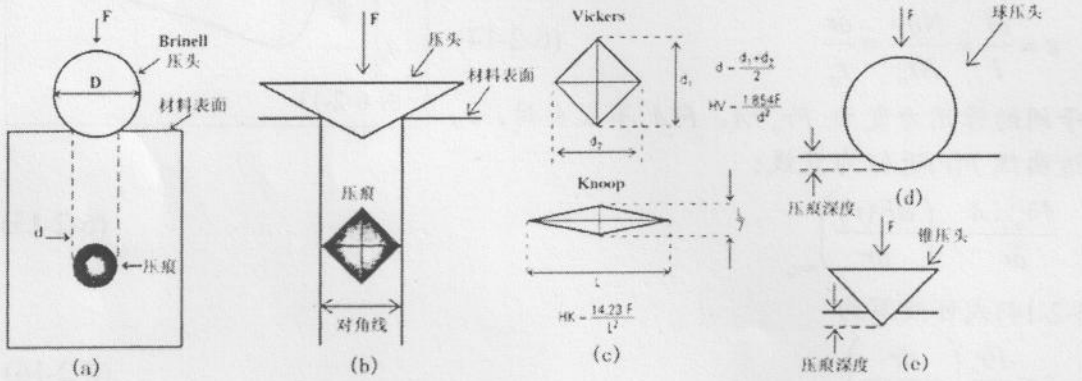


图 6-2-10 (a)布氏硬度测试法 (b)维氏硬度测试法 (c)维氏与洛氏压头比较 (d) HRB (e) HRC

不同的硬度体系方法各异，所得数据也有很大出入。这是因为上述所有硬度都是相对硬度。用微硬度计可以测定**绝对硬度**，其单位不是相对数值，而是 kg/mm^2。其数值是通过压力除以压痕投影面积得到的。

硬度与拉伸强度有一定关联。例如钢的拉伸强度(MPa)大约为布式硬度的 3.4 倍。当然这种对应关系仅是近似的，误差在 10%左右。

6-2-3 固体力学性质的微观解释

固体材料所表现出的各种力学特性都起源于构成固体的原子或分子之间的相互作用特性。这种力主要起源于原子之间的残余静电作用，当物体受到挤压，相邻的分子相互靠近时，分子间的相互作用表现为斥力，而当物体受到拉伸，相邻分子间的相互作用表现为吸引。通过实际测量得到固体的各种力学参量可用于研究固体中的结构特征，甚至结构中的各种可能的缺陷。

晶体弹性

在外力很小的情况下，晶体中的原子只发生了很小的相对位移，晶体中的粒子之间的势能都具有抛物线形的特征。这正是大多数材料具有一定范围的线弹性应变的原因。

当系统中原子间的相对位移比较大的时候，势能曲线就会偏离抛物线形而表现出非线性的弹性形变。下面从微观图像分析晶体的弹性，即用微观参量来描述各种弹性模量。我们具体讨论晶体受纵向外力 F 拉伸时的弹性模量。作为练习，

读者可以自己推导固体的体积模量，剪切模量与弹性势能之间的关系。

设晶体中平行于外力 **F** 的方向共有 n 行原子或离子，每一行有 N 个离子，最近邻粒子间距为 r_o，晶体截面积是 A，则每一行粒子受到平均拉力为 $F/n = Fr_o^2/A$。图 6-2-11 为晶体中最近邻粒子间的作用力曲线。其中 r_o 表示晶体不受外力作用时近邻粒子间的平衡距离(相当于化学键的长度)。晶体两端施加外力 F 后晶体伸长 δl，则每一离子键平均伸长 $\delta l = N\delta r$，相应的线应变为

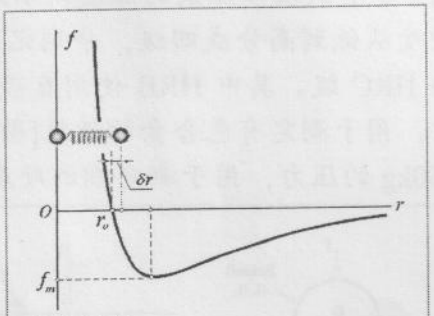


图 6-2-11

$$\varepsilon = \frac{\delta l}{l} = \frac{N\delta r}{Nr_o} = \frac{\delta r}{r_o} \tag{6-2-14}$$

粒子间的作用力变为 Fr_o^2/A。F/A 不太大时，r_o 附近曲线 $F(r)$近似为直线：

$$\frac{Fr_o^2/A}{\delta r} = \left(\frac{dF(r)}{dr}\right)_{r=r_o} \tag{6-2-15}$$

(6-2-14)式可改写为

$$\varepsilon = \frac{Fr_o}{A}\left(\frac{dr}{dF(r)}\right)_{r=r_o} \tag{6-2-16}$$

由线应力 $\sigma = F/A$，可得晶体的弹性模量

$$Y = \frac{1}{r_o}\left(\frac{dF(r)}{dr}\right)_{r=r_o} = -\frac{1}{r_o}\left(\frac{d^2U}{dr^2}\right)_{r=r_o} \tag{6-2-17}$$

其中 U 为近邻粒子之间的相互作用能。

下面以离子晶体 NaCl 为例对上述计算做一数量级估算。离子晶体中最近邻离子间的作用力可表示为

$$F(r) = \frac{mB}{r^m} - \frac{e^2}{4\pi\varepsilon_o r} \tag{6-2-18}$$

对氯化钠晶体而言，其中的常数 $m = 9$，B 可由 $dU(r_o)/dr = 0$ 的条件求得，其中离子间最近邻距离 $r_o = 2.81\times10^{-10}$m，由此可得到

$$U(r) = \frac{e^2}{4\pi\varepsilon_o r_o}\left(-\frac{r_o}{r} + \frac{r_o^9}{9r^9}\right) \quad 及 \quad \frac{d^2U}{dr^2} = \frac{e^2}{4\pi\varepsilon_o r_o}\left(-\frac{2r_o}{r^3} + \frac{10r_o^9}{r^{11}}\right)$$

代入(6-2-16)式，得到弹性模量

$$Y = \frac{8e^2}{4\pi\varepsilon_o r_o^4}$$

计算得出 $Y = 2.95\times10^{11}$Pa。实验推得的相应数值为 5.8×10^{10}Pa[21]，只有计算值的 1/5。如果在上述推导过程中计入次近邻及次次近邻间的相互作用，则实验值可达到上述理论值的 1/2。

如果应力较大，dU/dr 的线性性质不再存在，若外力大到使 $Fr_o^2/A \geq f_m$ (原子间的最大恢复力，见图 6-2-12)，化学键被拉断、材料发生断裂。

---

[21] 在 $T = 0$K 时的值。

屈服和范性形变

实际晶体材料中难免存在一些裂纹和缺陷，有些微小的裂纹和缺陷虽然用高精度的无损探伤仪也难以探测，但这些微量缺陷可极大地改变晶体的力学性质。例如纯铁中掺入微量的碳后(称为钢)要比纯铁坚硬得多。固体中的力学特性如韧性或脆性在很大程度取决于线缺陷—位错。另外，这种微小的裂纹和缺陷会在受力过程中逐渐扩展，

韧性固体发生塑性形变时表明系统已经进入了一系列的亚稳态结构。这时候系统中的缺陷，尤其是线缺陷，如位错和细裂纹起到了举足轻重的作用。例如，通过显微镜可以观察到金属发生范性形变时，其表面上出现一些条纹，称为**滑移带**。表明范性形变是晶体沿某族晶面出现滑移的结果，相应的晶面称为**滑移面**(见图 4-17)。实验表明，对于一定的材料(结构)，容易发生滑移的晶面和晶向往往是一定的。这意味着对一定的晶体材料而言，晶面发生滑移存在一个最小(临界)切应力 τ_c。若假设晶体出现滑移的过程中，晶面发生了整体的相对滑移，τ_c 的理论估算值约 $10^9 N/m^2$ 的数量级，而纯粹金属材料的 τ_c 实验值一般为 $10^5 \sim 10^6 N/m^2$。类似于多节爬虫的爬行模型可解释这一差异：爬虫体上的各节不是同时移动的，而是近尾部部分先缩短，再推动前面部分伸长而向前爬行，这样所需的力会小很多。位错运动模型假设：晶面中的一部分原子先发生滑移，从而推动同一晶面中另外部分原子滑移。循序渐移，最后导致晶面间的滑移。从而使晶面滑移时所需的临界切应力 τ_c 大为减小。

以上讨论的局部滑移部分逐步扩大时，反映在图 6-13(a)中的是位错线“⊥”逐步向右移动，达到晶体表面并消失时，滑移面上下方滑移出一个原子间距。若位错不断地从滑移面左端产生并运动到右端消失，则滑移出很多个原子间距后，就能看到宏观的滑移。范性形变就成形于这种宏观滑移中。以单晶体作拉伸试验为例，当应力超过弹性极限时，范性形变开始发生。通过高倍显微镜的观察发现表面上出现条痕，随着应力的增大，表面上各处都出现条痕，如图 6-13(b)所示。由此可见，范性形变与位错的产生及位错的运动有关。例如，在切应力作用下，位错是从内部逐步移向表面的，在移动时只是位错中心附近的原子重新排列，每作一次移动所需外力很小，滑移所需切应力很小，其弹性极限较低。

拉伸强度

理想固体材料的拉伸强度反映的是材料截面上的诸多连接截面两侧粒子的原子间的相互作用能的总和。但实验发现，由同一种材料制成的不同样品的强度极限值往往相差很大，而且明显地低于完整晶体的理论计算值，有时仅为理论值的数千分之一。前面也曾提到，许多物体的实验强度极限值也只有该物体的杨氏模量数值的百分之一，甚至千分之一，即实际的固体材料的拉伸量只有固体自身尺度的百分之一，甚至千分之一时往往就会断裂，这一数值远小于原子间的相互作用所能承受的伸长量。由于作用在物体上的拉伸应力是分担在材料截面上的诸多连接截面两侧粒子的分子键上的，上述现象的出现意味着材料断裂前，断裂面上有相当比例的原子并没有象在完整晶体中那样构成化学键，这对应于实际固体

材料中存在有微细裂纹。这些裂纹直接导致相应区域中连接截面两侧的分子键数目减小，这导致各分子键实际承受的拉伸作用力大大增加，从而显著降低了物体的拉伸强度。

物体中的裂纹也可用于解释脆性材料的抗拉强度和抗压强度之间的巨大差异，这种差异主要来源于材料中的裂纹对两种应力的不同响应。通常情况下，压缩会导致裂纹数量和尺度的减少，而拉伸则会导致裂纹数量和尺度的增加。(一个极端的例子是，一堆由砖块垒起的墙可以看成内在布满裂纹的一整块大砖。这块大砖的行为所反映正是耐挤压但经不起拉伸的固体的特性.) 铸铁、混凝土一类的材料的行为与砖墙的这种特性非常类似。不同材料的拉伸强度有着很大的差异，这通常也是在不同的用途中选择材料的标准之一， 例如在土木工程中最常用的材料是钢材和混凝土，钢在受到的拉伸力要达到 3000 MN/m^2 时发生断裂，而混凝土构件在拉伸应力仅为 5~10 MN/m^2 时就可能断裂，但混凝土的压缩强度要大很多。如果在混凝土内部植入钢条形成所谓的钢筋混凝土。由于钢条受到拉伸时相对于压缩时不容易变形，而混凝土受到压缩时不容易变形，钢筋混凝土对于拉伸和压缩都不容易变形。

疲劳断裂

目前对疲劳破坏的机理还没有很好的了解。一般认为，在内部缺陷或应力集中的区域，由于局部产生细微的裂纹，裂纹逐步发展，以致在工作应力远小于强度极限下即突然断裂。

如果一个金属曾经历过多个周期的应力变化，这种被称为**疲劳断裂**的现象是由小的裂纹缓慢扩展引起的。

蠕变

是材料在超过弹性限度恒定应力下随时间缓慢塑性形变的过程

6-2-4 高聚物的力学特性

以上讨论的实际上大多是金属材料的性质，某些非金属材料，如橡皮等还有所谓**弹性后效**：加上外力后会立刻出现一部分形变，而在接着的持续受力过程中形变将逐渐增加，并趋于某个极限值；撤除外力后，大部分形变会立即消失，而其余部分则需要很长的时间(若干天)才能逐渐消失。

高分子化合物一般具有优良的弹性、可塑性、抗拉、抗弯、抗冲击等机械性能，不同高聚物之间的性能差别与其分子结构、分子间作用力、分子链的柔顺性、聚合度、分子极性等因素有关。

这类材料对外加应力的响应很不相同：如果增加施加在橡皮带上的应力，并在断裂前撤去外力，其行为不服从胡克定律。它可能达到钢和玻璃大得多的应变量。例如，拉伸量可以达到原长的两倍，即应变量达到 2。

图 6-2-12 中给出的是橡皮筋在橡皮筋受外部拉伸前后的应力应变曲线。受外部拉伸应力的开始阶段 OA，橡胶拉伸的贡献主要来自卷曲着的分子链被展开拉直，因而，在较小的应力下表现出较大的形变响应，当分子链已处于几乎平直的状态，橡胶拉伸的贡献主要来自分子链本身的形变，橡胶显得更难以拉伸，对应于图中的 AB 阶段。撤除外力的过程中，两种形变的比例并非如伸展过程返回，而且最终的前后位型并不相同，这和我们在第三章中讨论摩擦力的情形类似，可以称为在拉伸和收缩的循环中存在有内摩擦过程，外部的机械能量转化为材料的热能(内耗)，这种效应称为**弹性滞后**. The shaded area represents the energy lost per unit volume.

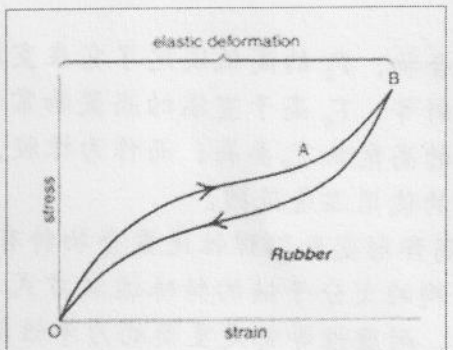


图 6-2-12

高分子链的结构和柔顺性

链型(包括带支链的)高聚物的长链分子通常呈卷曲状，且互相缠绕，分子链间的分子力较大，表现为一定的柔顺性和弹性。它们在加热时会变软，冷却时又变硬，可反复加工成型，称为热塑性高聚物。聚氯乙烯、未硫化的天然橡胶、高压聚乙烯等都是链型高分子化合物。

体型高聚物是链型（含带支链的）高分子化合物分子间以化学键结合而形成的具有空间网状结构的高分子化合物，其力学特性一般表现为弹性和可塑性较小，而硬度和脆性则较大。一次加工成型后不再能熔化，故又称为热固性高聚物。酚醛树脂、硫化橡胶及离子交换树脂等都是体型高聚物。

高分子化合物的物理形态

非晶态链型高分子化合物无确定的熔点，在不同的温度范围内可呈现三种不同的物理形态：**玻璃态**、**高弹态**和**粘流态**。

无定型聚合物在一定负荷和受力速度下可呈现**玻璃态**、**高弹态**和**黏流态**三种力学状态。聚合物处于玻璃态时，整个分子链和链段的运动都被冻结，宏观性质为硬、脆、形变小，只呈现一般硬性固体的普弹形变。聚合物处于高弹态时，链段运动高度活跃，表现出高形变能力的高弹性。当线型聚合物到达黏流温度以上时变为熔融、黏滞的液体，受力可以流动，并兼有弹性和黏流行为，称黏弹性。聚合熔体和浓溶液搅拌时的爬杆现象，挤出物出口模时的膨胀现象以及减阻效应等，都是黏弹行为的具体表现。其他如聚合物的蠕变、应力松弛和交变应力作用下的发热、内耗等均属黏弹行为。

高弹态是高聚物所独有的罕见的一种物理形态。呈高弹态的高聚物当温度降低到一定程度时，可转变成如同玻璃体状的固态，称为玻璃态（如常温下的塑料即处于玻璃态），该转化温度称为玻璃化温度，记为 T_g。天然橡胶的 T_g=−73°C。而温度升高时，高聚物可由高弹态转变成能流动的粘液，称为粘流态。高弹态转变为粘流态的温度叫粘流化温度，记为 T_f。对于非晶态高分

子化合物，T_g 的高低决定了它在室温下所处的状态，以及是否适合作橡胶还是塑料等。T_g 高于室温的高聚物常称为塑料，低于室温的常称为橡胶。用作塑料的高聚物 T_g 要高；而作为橡胶，T_g 与 T_f 之间温度的差值则决定着橡胶类物质的使用温度范围。

高弹形变和黏弹性是聚合物特有的力学性能。这些特性均与大分子的多层次结构的大分子链的特殊运动方式以及聚合物的加工有关。聚合物的强度、硬度、耐磨性等都是重要的力学性能指标。

6-2-5 固体力学特性应用举例

材料设计

固体的各类形变及其分解

上面讨论的都是比较简单的均匀形变(拉伸，压缩，剪切)。而实际问题中出现的形变通常是不均匀的。下面通过两个具体的典型事例：直杆的弯曲和圆杆的扭转来说明非均匀形变问题的求解和计算。

梁的弯曲

例 6-2-1 一高为 h，宽为 w，长为 L 的矩形截面的横梁在梁两端的支撑力 N_1, N_2 与全部载荷(包括自重)W 的作用下平衡。求解横梁的弯曲及其中的应力分布。

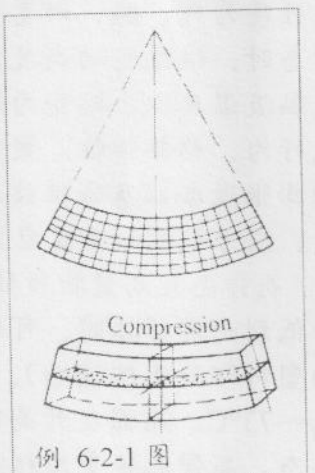
解： 利用系统的对称性和力学平衡条件可设载荷集中在梁的中点，而横梁两端的支撑力满足

$$N_1 = N_2 = \tfrac{1}{2}W$$

设想一个面 $O'O''$从中间将横梁分成两段，每一段分别受到一个力偶矩¼WL，该力偶矩的平衡由横梁的弯曲程度和横截面 $O'O''$上的应力分布所决定。

设想横梁由上至下分为许多薄层，梁发生弯曲时上部各薄层受到不同程度的压缩，下部各薄层则受到不同程度的拉伸。因此中间一定会有一个长度不变的中性层。在 $O'O''$面上的内应力分布如图 b 所示，这种内应力分布所形成的力偶矩 τ_{int} 与外力矩¼WL 平衡。假设横梁弯曲可以近似为一段圆弧，其曲率半径为 R，梁对曲率中心 C 的张角为 $\theta = L/R$。取截面 $O'O''$上中性层处 O 为原点，z 轴沿高度方向，坐标为 z 处一层的长度为 $(R - z)\theta = L(R - z)/R = L - Lz/R$，即 $\Delta L = -Lz/R$，相应的应变 $\varepsilon = \Delta L/L = -z/R$。应力 $\sigma = -Yz/R$(负应力为压力，正应力为张力)。高度为 dz 的一层横截面积为 $dS = bdz$，作用在其上的总力为 $df = \sigma dS = -Ybzdz/R$，对 O 点的力矩为 $d\tau = zdf = -Ybz^2dz/R$，故总力偶矩为

$$\tau_{int} = -\frac{Yb}{R}\int_{-h/2}^{h/2} z^2 dz = -2\frac{Yb}{R}\int_{0}^{h/2} z^2 dz = -\frac{Ybh^3}{12R}$$



例 6-2-1 图

其中负号代表左边半段所受力偶矩沿逆时针方向。平衡

时 $\tau_{int}=\tau_{ext}=¼WL$，梁的曲率为

$$\kappa=\frac{1}{R}=\frac{12\tau_{ext}}{Ybh^3}$$

由于一定的载荷下梁的弯曲程度(曲率)与宽度的一次方和高度的三次方成反比，说明增加梁的高度比增加宽度能更有效地提高其抗弯能力。另外，由于梁的中性层部分对抗弯的总力偶矩 τ_{int} 贡献不大，除去或减少这部分的材料对梁的抗弯能力不会有显著影响。因此工程上广泛采用工字钢、空心钢管等构件，这样可以在保证强度的前提下减轻重量、节约材料。◀

将杆分为若干小段(图 6-2-13 中 A、B, C、...等)，最右端小段 A 在外加力偶的作用下发生转动，改变其相对于 B 的取向。杆的弹性使改变了相对取向的 A、B 之间出现力偶作用，并导致 B 也发生转动，改变了它相对于 C 的取向。而 B 与 C 之间的力偶作用又使 C 发生转动...，杆的各部分发生的大小不等的转动形成杆的弯曲。

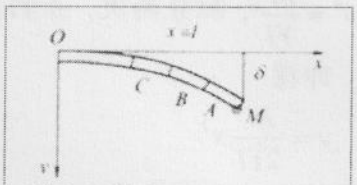


图 6-2-13

弯曲现象持续到杆的每一小段所受力矩达到平衡。即小段 A 受到外加力矩 M 以及 B 所施加的力矩–**M**；B 受到 A 所施加的力矩 M 以及 C 所施加的力矩–**M** ... 都相等。

伯努利假定杆的每一截面在弯曲后仍保持为平面，则上述弹性力矩就是杆的任一截面两侧的相互作用力矩，称为**挠矩**。在单纯弯曲情形，上述各截面上的挠矩相同。

为求解挠矩大小与杆发生弯曲程度之间的关系，具体讨论杆上一段原长为 dx 的微元(见图 6-2-14)，弯曲时杆微元的上半部受到不均匀拉伸。下半部则受到不均匀的压缩，因此两者之间一定存在一层长度不变的面 NN′，称为**中性面**，该面与页面相截形成图中所示曲线称为**中性线**。则杆元的曲率 κ 可用于描述其弯曲程度。设杆元的中性线曲率半径为 R，它对曲率中心 O 所张的角为 $d\theta=dx/R$，而距离中性面 y 的薄层弯曲时的长度由原长 dx 改变为

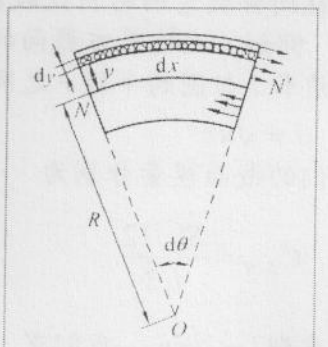


图 6-2-14

$$(R+y)d\theta=(R+y)\frac{dx}{R}=dx+\frac{y}{R}dx \qquad (6\text{-}2\text{-}19)$$

拉伸量为 ydx/R，相应的拉伸应变和拉伸应力分别为：$\varepsilon_{xx}=y/R$，$\sigma_{xx}=Y(y/R)$。

在纯弯曲情形，杆的长度整体上没有变化，因而杆端面上所受力的矢量和为零

$$\int p_{11}dS=\frac{Y}{R}\int ydS=0$$

其中 dS 为前述薄层的截面面积。上式中的积分为零表明质心在上述中性线上，即中性线是各个横截面质心的连线。

矢量和为零表明作用在端面上的许多力归结为力偶，相应的力矩就是所要求解的挠矩。过点 N'作垂直于纸面的垂线，作用在端面上的力相对于该直线的力矩为

$$M=\int yp_{11}dS=\frac{Y}{R}\int y^2dS\ \ M=YI\frac{1}{R} \qquad (11\text{-}4\text{-}4)$$

其中的积分 I 是横截面相对于上述直线的转动惯量。上式称为**伯努利-欧拉定律**。挠矩 τ 正比于中性线的曲率 $1/R$，即正比于杆的弯曲程度，符合我们的直觉。因为杆的自重通常远小于杆中的巨大张力，上述计算没中有计入杆的自重。

在单纯弯曲情形，各截面上的挠矩相同，杆的曲率半径 $1/R = M/YI$ 是常数，即杆的形状是半径 $R = YI/M$ 的圆。由此可计算弛垂度 δ(即右端所下降的距离)。

杆的实际弯曲一般很小，近似有

$$\frac{1}{R}=\left|\frac{y''}{(1+y'^2)^{3/2}}\right|\approx y''$$

则 $y''=\dfrac{M}{YI}$，积分两次，并利用杆左端的端点条件 $y(0)=0$ 和 $y'(0)=0$ 定出积分常数，即得

$$y=\frac{M}{2YI}x^2$$

求得杆右端弛垂度为

$$y|_{x=l}=\frac{M}{2YI}l^2$$

式(11-4-2)表明，距中性线越远处的张应力越大。设 y 的最大值为 h，则最大张应力大小为

$$p_{\max}=Y\frac{h}{R}=\frac{M}{I/h} \tag{11-4-5}$$

通常将参量 I/h 称为截面模量。对于一定的弯曲力矩 τ，杆的截面模量越大，$p_{\max}$ 越小，杆越不容易出现断裂。

但孤立地比较截面模量并没有什么意义。事实上，如保持 h 一定，只要加大截面的面积就能增大截面模量。但增大截面面积意味着需要更多的材料。因此应该就同样数量的材料比较截面模量，即就同样的截面面积来比较截面模量。

例如，比较正方截面的钢杆与圆形截面的钢杆如截面面积 S 相等；则正方形的边长 l 与圆的半径 r 之间的关系为 $l^2=S=\pi r^2$。即

$$l=\sqrt{\pi}r$$

它们的截面模量分别为

$$C_{正方}=\frac{Sl^2/12}{l/2} \qquad C_{圆}=\frac{Sr^2/4}{r}=\frac{Sr}{4}$$

考虑到 $l=\sqrt{\pi}r$，我们有

$$C_{正方}=\frac{\sqrt{\pi}}{6}Sr=0.294Sr>0.25Sr=C_{圆}$$

用同样多材料的正方截面钢杆比圆形截面刚杆更能抵抗弯曲。

正方截面事实上也不是最佳选择。靠近中性面的薄层形变小，张应力也小，远未达到它所能承受的应力。而距中心面较远的薄层形变较大，承受着较大的应力。为充分利用材料的性能，可以减小近中性面的薄层，而节省出来的材料添加到远离中性面的薄层中去，这导致截面呈现出工字形、T 字形、U 字形、圆环形等形状。这就是工程技术上广泛采用工字钢、T 字钢、槽例、空心钢管的原因。生物机体在长期的演化过程中也适应了该力学规律，如鸟的骨头和麦杆都是空心的。

圆柱的扭转变形

各种机械中的传动轴两端经常分别受到主动机构与被动机构所施加的相反的力矩而处于扭转状态。

例 6-2-2 用于机械中的传动轴、旋锲入木材中的螺丝钉以及扭摆装置中的悬丝，在运动过程中都经受到一定的扭转变形。在扭转角度很小的情况下，扭转形变是线弹性的，即系统所受到的**扭矩** $\tau_{\parallel}$ 和**扭转角** θ 成正比：$\tau=-D\theta$。求圆柱体的扭转系数 D 与圆柱体的几何尺度及力学参量之间的关系。

解： 设圆柱的长度为 L，半径为 R，两端分别受到一对大小相等、方向相反的力构成的力偶矩(见例图)，圆柱体的两个端面相对转过一个角度 θ(**扭转角**)。

由于圆柱体扭转时物体内与对称轴不同距离的质量元所发生的剪变角 φ 不同(扭转形变属于剪切形变[22])，圆柱体中半径从 r 到 $r+dr$ 的薄层上剪变角为 $\varphi(r)=r\theta/L$。横截面上的应力是切向的，剪切应力为

$$\tau_{\parallel}(r)=G\varphi(r)=Gr\theta/L$$

作用在该薄层横截面积 $dS=2\pi rdr$ 上的力为

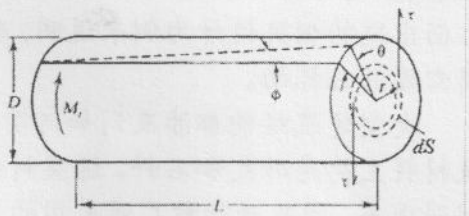
$$dF=\tau_{\parallel}(r)dS=2\pi G\theta r^2dr/L$$

该力对柱轴的力矩为

$$d\tau=rdF=2\pi G\theta r^3dr/L$$

在整个横截面上的总力矩为

$$\tau=\frac{2\pi G\theta}{L}\int_0^R r^3dr=\frac{\pi GR^4}{2L}\theta$$



例 6-2-2 图

总力矩 τ 与扭转角 θ 成正比，由此得到圆柱体的**扭转系数**(相应于扭转弹性模量)

$$D=\frac{\pi GR^4}{2L}$$

正比于半径的四次方，而于长度成反比。显然增加半径比减小长度能更有效地提高圆柱体的抗扭转能力，因此，用于抗扭转的零部件(如传动轴等)都做得比较粗，而用于一些涉及精密测量的实验器具(如卡文迪许实验，灵敏电流计等)中的悬丝等，通常做得很细很长以减小扭转系数，增加仪器的灵敏度。◀

金属材料的强化

大块纯金属晶体的抗变形能力一般不强，强度也不高。体内的位错很容易发生扩展，并导致出现晶面的滑移。要改善金属的这种力学特性，就要阻碍体内的位错的扩展。典型的方法是：

晶粒小型化：晶粒越小，位错就越难发生扩展。熔融的金属在迅速冷却、固化过程中会生成许多小晶粒，晶粒的尺度在材料再次受到加热或机械加工处理也会受到影响。

合金化：合金是不同金属或金属和非金属的混合物。合金中，外来原子往往能阻

[22] 圆柱体表面上的每根母线因扭转都倾斜一个角度 ϕ。如果在柱面上的一个“正方形”面元因扭转而变成了“菱形”，

断位错的扩展，使得滑移更为困难。例如，钢是铁(约占 99%)和碳的合金，有时还会掺入一些其它元素。它比铁的强度更高，不容易变形，也更强韧。

大多数金属在受到拉伸时不会发生脆性断裂，这是因为位错通常位阻止裂纹的成长或扩展。对金属制品进行锻压、辗展，拉伸，弯曲等加工处理都能提高材料的强度。但金属零件在这种机械加工过程中会变硬、变脆。其原因是这些加工过程导致晶体内生成大量的位错，由于位错之间能相互阻碍了位错的扩展，难于产生范性形变，表现为硬度增大，称为**加工硬化**。相反，通过减少晶体中的位错数量可以使金属材料变软，更柔韧，例如可以通过退火[23]处理来实现这一目的。

材料的各种强度在结构设计中的应用

对于实际的结构设计中涉及的构件强度的问题可以分为两类，一类是构件形状(或不同材料组装成的形状)，另一类是选择合适强度的材料。我们以建筑结构方面出现的相关设计为例来说明，当材料的强度不够时如何通过改变形状来达到所需要的目标的。

大型建筑结构都涉及到如何实现两点之间的空间跨接问题。十九世纪前，建筑材料主要是砖瓦和石料。这类材料的特点是抗压强度很高，但抗拉和剪切强度则弱很多。因此开始被广泛采用的是直柱和横梁跨接的希腊式结构，即用两根立柱支撑横梁两端。在古埃及的金字塔以及古希腊神庙建筑中随处可见这类结构的例子(见图 6-2-8)。该结构的最大缺点是需要大尺度的天然石料作横梁，而且由于石料抗拉伸和抗剪切强度不高，因而跨度不能很大，这就需要大量使用支撑横梁的立柱，从而占用大量的室内空间。

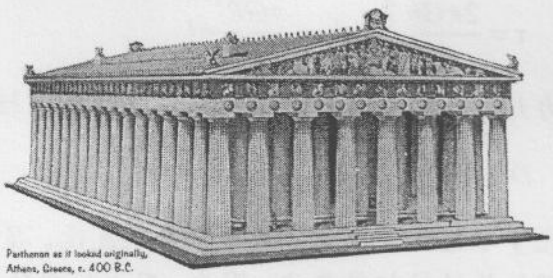


图 6-2-8

对这项结构的一个重大的革新是由许多形状特定的石料跨接形成的弧形拱，该结构在老式的砖石房屋的窗户和门洞设计中很常见，也被广泛应用于桥梁的建设。其特点是施加在拱弧上的所有负载应力被完全转化为压缩应力，没有拉伸和剪切分量，这样圆拱结构至少在理论上可以完全消除空间跨度结构中的拉伸应力，从而可以采用砖石，混凝土等材料实现大尺度的空间跨接。(见图 6-2-9)显示出外部的负载不仅没有削弱，反而加固了这种特殊的结构：砖石之间的互相挤压使得它们甚至不需要任何粘合材料也能也能更好地连

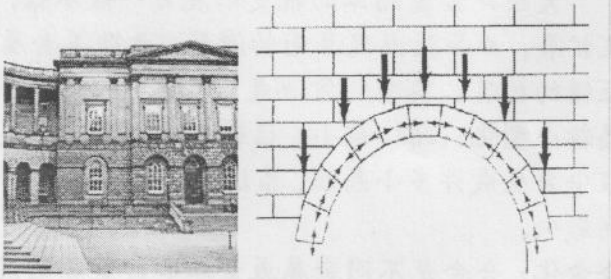
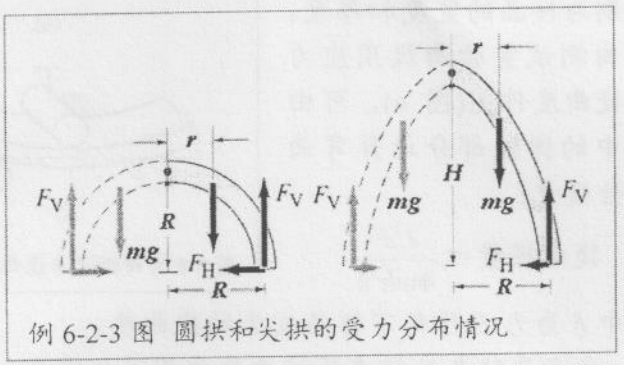


图 6-2-9

[23] 退火是指将金属材料、玻璃及制成品加热保温，然后缓慢冷却的热处理工艺。通过退火可消除材料的内应力，以降低脆性，软化金属，细化金属晶粒，以改善机械性能。

接在一起。而拱部材料中的水平和竖直方向的应力最终都会传递给支撑部分。这要求在两边提供足够的支撑以支撑力的水平分量。即需要采用一些技术方法予以约束，一般采用, either with heavy sides and friction or angled cuts into bedrock or similar.

例 6-2-3 比较具有相同跨径和重量的圆拱和尖拱在基座位置的水平负载分量。



例 6-2-3 图 圆拱和尖拱的受力分布情况

解： 例图所示为具有相同跨径和重量的一个圆拱和一个尖拱。为简单起见，将每一个拱分为两部分，每一部分受到的重力为 mg。维持重力平衡需要向每一部分提供大小为½mg 的竖直作用力，而保持转动平衡，需要在拱的基座位置向每一部分提供水平作用力 F_H。具体考虑右面的拱，要求它所受到的力相对于拱顶的总力矩为零，则相对于圆拱的力矩方程为

$$R\times mg - \tfrac{1}{2}R\times mg - H\times F_H = 0$$

可得

$$F_H = \frac{R}{2H}mg$$

圆拱的高度为 $h=R$，而尖拱的高度为 $h=H>R$ 圆拱，因此尖拱所需的水平支撑力 F_H 因为具有更长的力臂而减小了相同的比例。拱越陡，所需要的水平支撑分力越小，对于几乎竖直的拱，力几乎都作用在基座上。◀

圆拱围绕其对称轴旋转就形成了所谓的圆顶，这种结构用于建筑最早出现于公元1100 年的罗马。图 6-2-10 为罗马的万神庙中的歌特式大教堂是其典型的代表，也是那一时期罗马建筑的标志。它的技术创新在于如何支撑大教堂的塔那样的沉重载荷。圆顶的重量，类似于圆拱的情形产生很大的压缩应力。由于圆顶一般很陡，上面的负载重力几乎是竖直于地面，这就大幅度地减小了为维持力学平衡所需要的水平支撑力。圆顶减小了施加在墙上的负载，因而可以提供更大的空间。

图 6-2-10 罗马万神庙的圆顶

脆性材料挠曲强度

6-2-2 节曾提及，有些脆性材料因表面裂缝太多而无法进行正常的拉伸测试：样品一放上夹具就会开裂。这时可采用图 6-2-11 所示的弯曲测试方法，用挠曲强度与挠曲模量对材料进行描述。

$$\text{挠曲强度} = \frac{3FL}{2wh^2}$$

其中 F 为断裂时的外力，L 为最外两点间的距离，w, h 分别为样品的宽度和厚度。弯曲测试实验曲线用应力对挠曲度作图(图 c)。可由其中的弹性部分计算弯曲弹性模量：

$$\text{挠曲模量} = \frac{L^3F}{4wh^3\delta}$$

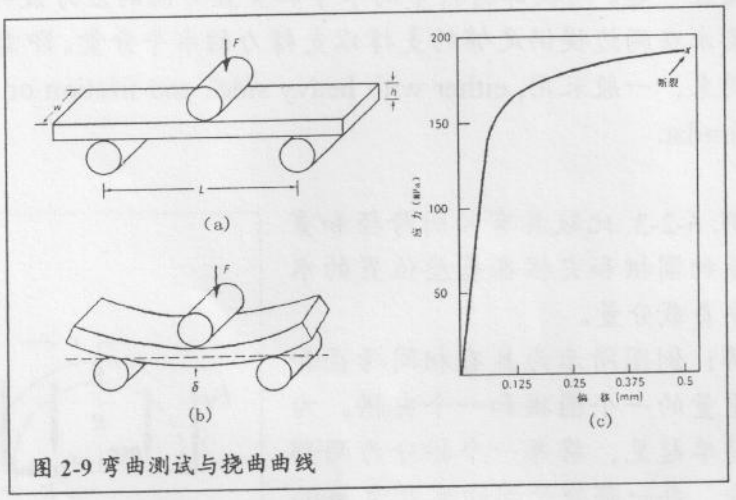


图 2-9 弯曲测试与挠曲曲线

其中 δ 为力 F 作用下样品发生的挠曲度。

由于裂缝与缺陷在压缩力的作用下会闭合，脆性材料也常用压缩方法进行测试。

例 6-2-4 一玻璃纤维增强复合材料的挠曲强度为 315MPa，挠曲模量为 124GPa。一样品宽 12mm，厚 9.5mm ，长 200mm，置于相距 125mm 两圆辊之间。计算使样品断裂所需的力以及样品断裂时的挠曲。假设无塑性形变。

解：将样品尺寸代入挠曲强度公式：

$$315 = \frac{3FL}{2wh^2} = \frac{3F \times 125}{2 \times 12 \times 9.5^2} = 0.173F \qquad F = \frac{315}{0.173} = 1819N$$

又由挠曲模量公式：

$$124 \times 10^3 = \frac{L^2F}{4wh^3\delta} = \frac{125^3 \times 1819}{4 \times 12 \times 9.5^3\delta} \qquad \delta = 0.7\text{mm}$$

6-3 流体的力学特性及其描述

从研究和应用的尺度而言，流体相对于固体更接近于连续介质。由于流体与固体在微观结构上的差异，两者对外力的响应有很大的不同，对流体行为和特性的研究方向和方法也和固体有很大的区别。例如，对固体力学特性的研究大都关注于物体的弹性、断裂等与形变和强度相关的性能和现象，而影响这些性能的一个主要方面是实际结构中的缺陷。而流体由于在剪切应力的作用下可以很容易地发生连续变形，处于静态条件下的流体内部不会出现类似固体中的结构缺陷(不存在结构上的亚稳态)，因而流体对外部的静态扰动的响应比固体简单。但同样因为容易变形，流体对外部动态扰动的响应比固体更为复杂，而且涉及到许多固体问题中所没有的复杂情况。本节首先讨论表征液体性质的特征参量，我们将特别关注两个概念：表面张力和黏滞系数，以及对这两个的不同测量设计所产生的差异。我们还将讨论处于静态和稳恒状态中流体的力学特性及其描述，下一节介绍流体的一些典型的动力学问题。

6-3-1 流体的应力特性

从日常对水的运动行为观测可知，流体非常容易变形，而且这种因外力作用而产生的形变在外力撤除后并不会有任何恢复。如果用固体的力学特性，即固体的应力-应变关系来描述液体的力学行为，可以说液体几乎没有抗切变，抗拉伸的强度可言，或者说流体几乎没有相应形变的弹性区域，在很小的拉伸或剪切应力作用下流体就直接进入断裂状态。

正应力—压强及负压

流体的抗张强度很小，但具有相当的抵抗体压缩能力[24]，因而流体中的正应力通常表现为压力，也称为**压强**，而静止流体中的压强通常称为**流体静压力**。流体的抗压缩能力可以和固体一样用**体积模量**描述：

$$B = -\frac{\Delta P}{\Delta V/V} = -V\frac{dP}{dV} \tag{6-3-1}$$

其中的负号表示增加外部压力总是使流体的体积减小，即 $dP/dV<0$。

液体一般具有比较大的体积模量。例如，在一个大气压下，温度为 15.6°C 的水受到 21.5 MPa 的压强，其体积减小的比例仅为 1%。体积模量取该数量级在液体中是很典型的，这意味着许多液体在一般的工程问题中可以看成是不可压缩的。

流体的抗压缩特性也常用流体的密度 ρ 随压强变化时的相对改变来描述，相应的参量称为**压缩系数**

$$\beta = \frac{1}{\rho}\frac{d\rho}{dP} \tag{6-3-2}$$

显然，压缩系数 β 是体积模量的倒数

[24] 这表明液体内部的分子作用力主要表现为斥力，

$$\beta = \frac{1}{\rho}\frac{d\rho}{dP} = -\frac{1}{V}\frac{dV}{dP} = \frac{1}{B} \tag{6-3-3}$$

流体的体积模量通常随着自身的压缩量增加而增加，说明体积模量时需要注意流体所处的压强区域。不加说明的体积模量一般是指流体处于大气压强的环境。

在某些场合，液体中的正应力也表现为张力。如图 6-3-1 所示，在弹簧秤的下端水平地悬挂一块玻璃，使玻璃片的下表面与水面接触后往上提拉玻璃片，这时弹簧秤的读数会明显大于玻璃片受到的重力。这意味着玻璃片的下表面受到下面紧邻的水分子的吸引作用，相应地附着在玻璃片的下表面的水分子薄层与下面紧邻的水分子之间处于相互拉伸的受力状态。液体中的这种张应力(其微观图像是液体分子之间的吸引力，通常称为**内聚力**)通常称为**负压**。持续增加向上的作用力最终将导致玻璃片与水体脱离。这时可以看到玻璃片的下表面沾有薄层的水，说明两层水分子间的张力达到了最大**抗张强度**(称为"**负压极限**")后发生了断裂，水的负压极限约为 300 atm。读者可以做一个实验，调节水龙头的出水流量，在某一临界状态下可以观察到连续的水柱发生断裂的情况。负压的概念也可用于解释植物从根部吸收的土壤中的养分往上输运到顶部的效应(见习题 6-3-2)。

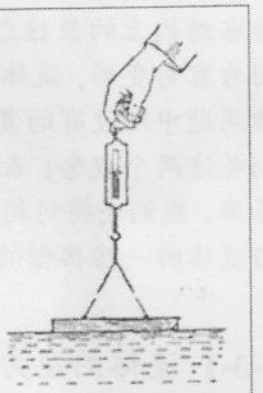


图 6-3-1 演示液体负压的实验。

切应力—黏滞性

从宏观尺度而言，流体与固体在力学特性上的另一个差异在于它们对切应力的响应完全不同：处于静态的固体可以维持在切应变的受力状态，因而称固体具有抗切变的能力(用切变模量描述)，而处于静态的流体质元之间几乎不能维持在切应变的受力状态，即流体几乎不具有抗切变的能力，微小的剪切应力就能使流体质元之间直接发生**断裂**，它表现为液体层之间的相对滑移，这种运动形式称为"**流动**"。但这并不意味着各流层之间不存在切应力，事实上，相邻的两层流体发生相对滑移(即剪切运动)时会出现相反方向的应力以阻碍变形的加剧，只是这种作用力很小，不足以阻止两层流体相对滑移量的增加，因而表现为两层流体发生断裂后出现的相对滑移阻尼[25]。流体的这种特性称为**黏滞性**。相应的应力称为**黏滞力**。由于这种切向应力结果与固体受到切向应力表现出的观测效应不同，描述这种应力和形变之间关系的参量也不同于固体中的情形。

流体的**黏滞张力**所涉及到的是流体的流动状态，而流体的流动通常相当复杂，它不仅与流体的流速，温度、压强等状态参量有关，也与流体通道的形状及表面情况有关。为了能简洁地描述和说明流体的**黏滞**特性，我们考虑一种最为简单，有序的流动：各流体质元的轨迹表现为规则的光滑曲线，相邻流体质元的轨迹线彼此差别很小，彼此不相互混杂。这种流动具有分层平行流动的特点，因而称为**层流**[26]。

---

[25] 用固体之间的摩擦词汇来说，流体之间没有静摩擦，只有滑动摩擦。

[26] 层流发生在流速较小，更确切些说是发生在雷诺数(见 6-4 节)较小时的情形，当雷诺数超过 2300 左右时，流动成为**湍流**，流速随时间和空间坐标呈现出随机的紊乱变化。

具体考虑如图 6-3-2 所示意的实验：两块间隔很小的平行板之间填充某种液体，下平板固定，对上平板施加一水平力 F 使之以较小但恒定的速度 u_o 运动。实验发现，流体受上平板运动的带动而产生流动，达到稳态时形成了许多不同速度的水平薄层流体。其中黏附在上平板上的薄层以速度 u_o 随上平板运动，而次一层流体受顶层的拖曳向前运动过程中出现相对滑移运动。从而使顶层流体受到黏性阻力。这种下一层流体由上一层流体驱动并伴随有相对滑移运动波及到整个流体，附着在静止平板上的最下一层流体也受到一向前的作用力，但该力被固定的下平板对它的作用力平衡而处于静止状态。流体最后达到的**稳定流动**状态是两平行板之间的流体速度随高度的分布 $u(z)$ 呈线性变化关系[27]：

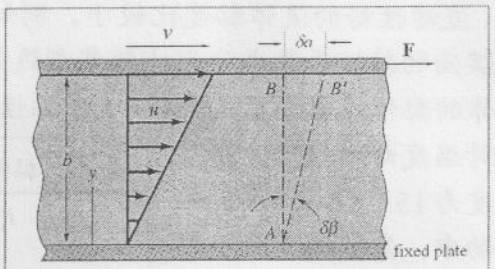


图 6-3-2

$$u(y) = u_o \cdot (y/b) \tag{6-3-4}$$

其中 y 是液体距离下平板的高度，b 是上、下板之间的距离。两板之间每一水平薄层内的流体流速不变，因而上一层流体对它的拖曳力 f 必然等于下一层流体对它作用的阻尼力 f'。而对于两个相邻流层而言，作用在上一流层上的阻力 f' 一定等于它作用在下一流层上的驱动力 f。平行于流速的层面两侧的相邻流层间存在有一对阻止它们相对"滑动"的切向作用力与反作用力，这种**内摩擦力**被称为**黏性力**，其作用效果是使流动较快的流层减速，流动较慢的流层加速。

为了得到这种黏性张力与切向流体形变量之间的关系，考虑在 Δt 时间内，流体中的垂直线 AB 转过角度 $\Delta\beta$(见图 6-3-2)：

$$\tan\delta\beta = \frac{\Delta a}{b} = \frac{u_o \Delta t}{b} \tag{6-3-5}$$

即切应变 $\delta\beta$ 除了与切应力 τ ($\tau = F/A$, F 是控制 u_o 的参量)有关，还是时间的函数。因而无法建立 τ 与 $\delta\beta$ 之间完全确定的联系。为了去除时间的关联性，考虑切应变的时间变化率

$$\dot{\gamma} = \frac{d\beta}{dt} = \frac{u_o}{b} = \frac{du}{dz} \tag{6-3-6}$$

因外力 F 是控制最上层流体速度 u_o 的参量，切应力 $\tau = F/A$ 也与切应变成正比：$\tau \propto \dot{\gamma}$。这表明切应力正比于流体中的速度梯度：

$$\tau = \mu \frac{du}{dz} \tag{6-3-7}$$

上式通常称为**牛顿黏滞定律**，其中 μ 称为**动力学黏性系数**，简称**黏性系数**。该系数与流体密度的比值

$$\nu = \frac{\mu}{\rho} \tag{6-3-8}$$

---

[27] 需要注意的是速度梯度为常数 $du/dy = v/b$ 的情形并非总是成立的。The experimental observation that the fluid "sticks" to the solid boundaries is a very important one 在流体力学中 and is usually referred to as the no-slip condition. 各种流体(液体和气体)都满足这一条件。*Real fluids, even though they may be moving, always "stick" to the sold boundaries that contain them.* 泰勒展开的说明。

称为**运动学黏性系数**，单位是 $N\cdot s\cdot m^{-2}$。

流动性好的流体黏度比较小。例如，水、煤油比糖浆、汽油流动性好，而水与煤油的黏性系数也分别比糖浆与汽油的小。气体相对于液体更容易流动，因而气体的黏性系数小于液体。另外，黏性系数对压强的依赖性很小(一般可以忽略)，而对温度却很敏感。水在温度为 15.6°C ~ 37.8°C 的变动中，其密度下降不到 1%，但黏性系数下降了 40%。因微观机制不同，液体与气体的黏滞性与温度的关系很不相同：气体的黏度随温度升高而增加，液体的黏度随温度升高而减小。表 6-2 列出了一些流体的黏度。

表 6-2 黏滞系数

液体	t(°C)	η (10^{-3}Pa·s)	气体	t(°C)	η (10^{-3}Pa·s)
	0	1.79	空气	20	1.82
水	20	1.01		671	4.2
	50	0.55	水蒸	0	0.9
	100	0.28	汽	100	1.27
水银	0	1.69	CO_2	20	1.47
	20	1.55		302	2.7
酒精	0	1.84	氢	20	0.89
	20	1.20		251	1.3
轻机	15	11.3	氮	20	1.96

下面介绍几种常用的测定气体的黏度的实验设计，

(1) 旋转黏度计

图 6-3-3 所示意为测量流体黏度而设计的一种装置。一外径为 R、长为 L 的圆筒用扭丝悬吊着，筒外同心套有另一个相同长度、内径为 $R+\delta$ 的圆筒($\delta \ll R$)，内、外筒之间充有待测流体。使外筒以恒定角速度 ω 旋转，这时内筒受流体黏性力产生的力矩与扭丝的扭转力矩 τ 所平衡。由装在扭丝上的反光镜 M 的偏转角度可测定 τ。

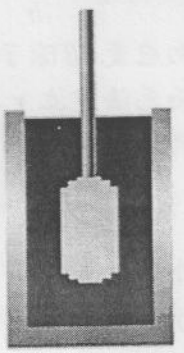


图 6-3-3

因内筒静止，夹层流体有 $v/\delta = \omega R/\delta$ 的速度梯度(因 $\delta \ll R$，可认为夹层内速度梯度为常数)，对内筒表面的黏性力产生对扭丝作用的合力矩为

$$\tau = 2\pi RL\cdot\eta\cdot\frac{\omega R}{\delta}\cdot R = \eta\frac{2\pi R^2 L\omega}{\delta}$$

由此可以得到气体的黏度系数

$$\eta = \frac{\tau\delta}{2\pi R^2 L\omega} \tag{6-3-9}$$

(2) 管流 泊肃叶定律

考虑不可压缩且黏度为 η 的黏性流体在半径为 R、长为 L 的一段水平直圆管中的沿轴匀速流动。要使管内流体作匀速运动，必须有外力来抵消黏性力，这个外力来自圆管两端的压强差 ΔP。对给定的 ΔP，管轴处($r=0$)的流体速度最大，而紧贴管壁的流体速度为 0，即管流中存在速度的径向分布(见图 6-3-3)。为了分析该速度径向分布函数 $v(r)$，考虑流体内的一圆筒状薄

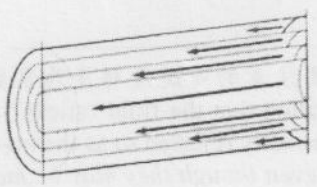


图 6-3-3

流层，其内、外半径分别为 r 和 $r+dr$，因速度梯度 $dv/dr<0$，流层受到的黏滞力的合力为

$$df = f(r+dr) - f(r) = 2\pi\eta L\left[(r+dr)\frac{dv(r+dr)}{dr} - r\frac{dv(r)}{dr}\right]$$

$$= 2\pi\eta L\frac{d}{dr}\left[r\left(\frac{dv}{dr}\right)_r\right]dr$$

对于稳态流动(也称为**定常流动**)，该力应与端面 $2\pi rdr$ 上的压力差平衡

$$(P_b - P_a)2\pi rdr = 2\pi\eta L\frac{d}{dr}\left[r\left(\frac{dv}{dr}\right)\right]dr \quad 或 \quad d\left(r\frac{dv}{dr}\right) = \frac{P_b - P_a}{\eta L}rdr$$

两边从 $r=0$ 到 r 积分，得

$$dv = \frac{P_b - P_a}{2\eta L}rdr \tag{6-3-10}$$

再从 r 积分到 R，并利用管壁上 $v(R)=0$，得

$$v(r) = \frac{P_a - P_b}{4\eta L}(R^2 - r^2) \tag{6-3-11}$$

为旋转抛物面。通过圆环面积 $2\pi rdr$ 的流量为 $dQ = 2\pi vrdr$，故管中的总流量为

$$Q = 2\pi\int_0^R v(r)rdr = \pi\frac{P_a - P_b}{2\eta L}\int_0^R (R^2 - r^2)rdr = \frac{\pi}{8}\frac{P_a - P_b}{\eta L}R^4 \tag{6-3-12}$$

称为**泊肃叶**(J. L. M. Poiseuille)**公式**[28]。泊肃叶公式也可表述为其它形式。管中的平均流速 $\bar{v}$ 可定义为体积流量 Q 除以横截面积 $S=\pi R^2$：$\bar{v} = Q/\pi R^2$，此外，压强差 $P_a - P_b$ 除以管长 L 是沿管的压强梯度，再乘以管的横截面积 S 则为压力梯度：$\Delta F/\Delta L$，由上式可以得到

$$\frac{\Delta F}{\Delta l} = \frac{P_a - P_b}{L}\pi R^2 = \frac{8\eta Q}{R^2} = 8\pi\eta\bar{v} \tag{6-3-13}$$

即克服黏滞力迫使流体沿管流动所需的压力差正比于平均速度和黏滞系数。

利用泊肃叶公式可以制成测量液体黏滞系数的装置。如图 5-53 所示，让液体从接在容器器壁上的水平细管中流出。R 和 L 是已知量，压强差由竖直细管中液面的高度差 h 计算出：$P_a - P_b = \rho\ \ gh$。则可从泊肃叶公式计算出黏滞系数。

(3) 落球测量法 斯托克斯公式

固体在黏滞性流体中作匀速运动时会受到阻尼，英国数学和物理学家斯托克斯[29]于 1851 年导出了球形物体在流体中作匀速运动时所受的阻力

$$F = 6\pi\eta rv \tag{6-3-14}$$

其中 r 为球体半径，v 为球体相对于流体的速度，η 为球体流体的黏滞系数。由此**斯托克斯公式**(6-3-14)提供了一种测量黏滞系数的重要方法。让一个质量为 m、半径为 r 的小球在盛有待测液体的量筒中降落。小球所受的黏滞力因很快与重力达到平衡而匀速降落。测出其降落速度 v 即可得到黏滞系数：

---

[28] 泊肃叶定律仅适用于水平直圆管，且流量不大，呈层流的情况。流量变大而出现湍流时是不适用的。

[29] G. G. Stokes

$$\eta = \frac{mg}{6\pi r v} \tag{6-3-15}$$

满足牛顿黏性定律的流体称为**牛顿流体**，而非**牛顿流体可大致**分为三类：

(1) 其速度梯度和与互相垂直的黏性力间不呈线性函数关系，如泥浆、橡胶、血液等；

(2) 其黏性系数随时间而变或与流体以前的历史过程有关，如油漆等凝胶物质；

(3) 对形变具有部分弹性恢复作用，如沥青等黏弹性物质。

应力张量*

牛顿黏性定律(6-3-7)只适用剪切流动这一简单情形，下面讨论适用于一般情形的广义牛顿定律。

由于流体运动时的应力张量在运动停止后趋于静止流体的应力张量，考虑将应力张量分解为各向同性和各向异性两个部分：

$$\sigma_{ij} = -P\delta_{ij} + \tau_{ij} \tag{6-3-16}$$

其中 P 称为**压力函数**，流体静止时 P 等于流体静压强。τ_{ij} 称为**偏应力张量**，流体静止时 τ_{ij} 等于零。假设偏应力张量的各分量是局部速度梯度张量分量的线性齐次函数[30]：

$$\tau_{ij} = c_{ijkl}\frac{\partial u_k}{\partial x_l} \tag{6-3-17}$$

上式称为**广义牛顿黏滞定律**，其中 c_{ijkl} 为表征流体黏性特征的四阶常数张量，包含 3^4=81 个元素，但其中的独立变数少得多。下面以各向同性流体为例予以说明。

首先将张量$\partial u_k/\partial x_l$分解为对称张量 s_{kl} 和反对称张量 a_{kl} 之和

$$\frac{\partial u_k}{\partial x_l} = s_{kl} + a_{kl} = s_{kl} + \varepsilon_{klm}\omega_m$$

其中 ε_{klm} 定义为

$$\varepsilon_{klm} = \begin{cases} 1 & k,l,m\text{互相不等，而且按右手螺旋排列} \\ -1 & k,l,m\text{互相不等，而且按左手螺旋排列} \\ 0 & k,l,m\text{有相同的情况} \end{cases}$$

为三阶反对称张量，代入(6-3-17)式可得

$$\tau_{ij} = c_{ijkl}s_{kl} + c_{ijkl}\varepsilon_{klm}\omega_m \tag{6-3-18}$$

因 τ_{ij} 是对称张量，c_{ijkl} 中的指标关于 i 和 j 满足交换不变性，则 c_{ijkl} 可表示为

$$c_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \tag{6-3-19}$$

其中只包含两个独立参数 λ 和 μ，代入(6-3-18)式

$$\tau_{ij} = \lambda s_{kk}\delta_{ij} + 2\mu s_{ij} \qquad (i,j,k = x,y,z) \tag{6-3-20}$$

称为**本构方程**，则应力张量可表示为

$$\sigma_{ij} = (-P + \lambda s_{kk})\,\delta_{ij} + 2\mu s_{ij}$$

或

$$\sigma_{ij} = -P\delta_{ij} + 2\mu(s_{ij} - \tfrac{1}{3}s_{kk}\delta_{ij}) + \mu' s_{kk}\delta_{ij} \tag{6-3-21}$$

---

30 应该说，这只是需要实验验证的假设。

其中

$$\mu' = \lambda + \tfrac{2}{3}\mu \tag{6-3-22}$$

为了说明参数 μ 和 μ'的物理含义，回顾并比较牛顿剪切流动实验：$v_x = u(y)$, $v_y = v_z = 0$，则 (6-3-12)式退化为

$$\tau_{xy} = \mu \frac{\partial u}{\partial y} \tag{6-3-23}$$

显然，μ 就是动力学黏性系数。因 表示体积变化，即可压缩流体有两个黏性系数。称为**膨胀黏性系数**，又称为**第二黏性系数**。

由于流体的流动特性，描述流体的宏观力学特性的物理量一般都是时间和空间的函数。

表面应力

在两种不相溶的液体之间或液体与气体之间，液体与固体之间的分界面上所受到应力与液体内部的情况不同，并因此会表现出一些独特的现象，称为**表面张力**效应。图 6-3-4(a)所示为一钢制曲别针在水面上处于静止平衡的情形，这是因为表面张力使液面总是呈现出收缩的趋势，只要曲别针本身的重量不坠破水面，就能维持这种类似于放置在一张张紧的弹性薄膜上的状态。另外，表面张力所导致的液面呈现出收缩的趋势也是漂浮在空中的肥皂泡以及露滴总是呈球状(图 6-3-4b, c)的原因。

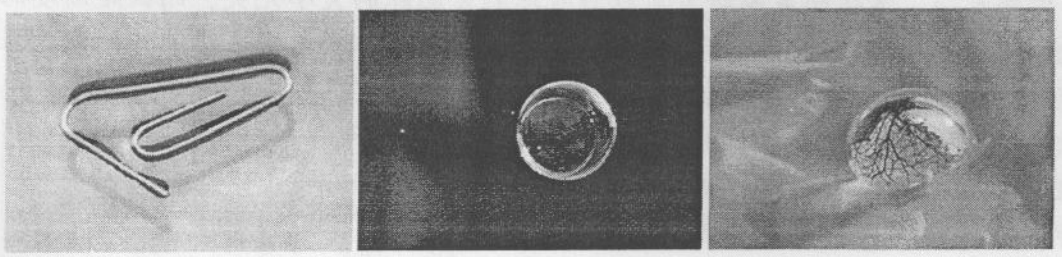


图 6-3-4

下面分别就一种液体与其它流体，以及与固体接触时在界面液体一侧的张力特征，以及由此呈现出的相关效应做一简单的介绍。

气-液界面

表面张力是作用于液体表面上的使液面具有收缩倾向的一种力，能够最简洁地说明这一作用力的事例是气-液界面的情形。

从微观上看，表面张力是由于液体表面的过渡区域(称为**表面层**)内分子力作用的结果。表面层厚度大致等于分子引力的有效作用力程 R_o，其数量级约为 10^{-9}m，即二、三个分子直径的尺度。设分子具有球对称作用势能，以任一分子为中心以 R_o 为半径的分子画一球，则只有在作用球内的其它分子才对该分子有作用，因而称其为**分子作用球**。对处于液体内部的分子而言，其作用球内的其它分子对它的作用力相互抵消，但对液体表面层内的任一分子而言，其作用球中有一部分是密度很低的气体，因此分子所受

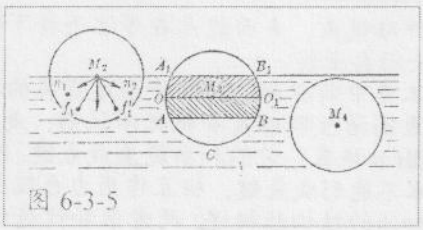


图 6-3-5

到的其它分子的作用力不平衡，其合力垂直于液体表面并指向液体内部(图 6-3-5)。液体在这种合力作用下有尽量缩小其表面积的趋势，使液体表面像拉紧的膜一样。可见表面张力仅是分子间吸引力的一种表现形式。

图 6-3-6 给出一种测量表面张力大小的简单装置。将矩形金属框在肥皂液中浸泡一下使框中形成一肥皂液膜(图 a)。其中金属框的一边可以滑动。当外力 F 在等温条件下拉伸可滑动框边(图 b)，使之移动距离 Δx，则外力 F 因拉伸液膜克服其表面张力做的功为

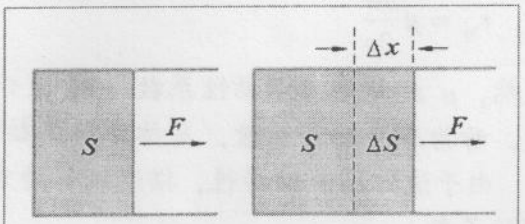


图 6-3-6

$$\Delta W = F\Delta x = 2\sigma L\Delta x = \sigma\Delta S$$

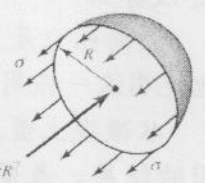
其中 L 为滑动框边的长度，参量 σ 称为**表面张力系数**，表示单位长度的滑动框边两边的液面中其中的一个对该框边的作用力，单位是 $\mathrm{N\cdot m^{-1}}$。我们也可以从另一个角度来理解表面张力系数。由于上式中 $2L\Delta x = \Delta S$，其中 ΔS 为液膜增加的表面积(液膜有两个表面)。液膜的表面积扩大意味着一部分液体内部的分子进入到表面层中，由于进入表面层的每一个分子都需克服分子力的合力(其方向指向液体内部)做功。表面积扩大 ΔS 所需要的功为 $\Delta W = \sigma\Delta S$ 等于表面层中的分子引力势能(称为表面能)的增加：$\Delta U = \sigma\Delta S$。因此表面张力系数也可理解为增加单位面积液体表面时增加的表面能[31]。表面张力系数的单位可写成 $\mathrm{J\cdot m^{-2}}$。

实验发现，液体的表面张力系数与液体的表面积的大小无关，而仅是温度的函数。表面张力系数随温度的升高而降低[32]。表 6-3-4 列出了一些液体的表面张力系数。液体中杂质的存在会影响 σ 的大小。能使 σ 明显变小的物质称为表面活性物质(肥皂就是最常见的能使水的表面张力系数变小的表面活性物质)。若物质 A 能显著降低另一种物质 B 的表面张力，称 A 对 B 具有表面活性。由于水是最重要的溶剂。通常的表面活性都是对水而言[33]。

表 6-3-4 与空气接触液体的表面张力系数

液体			液体		
水	0	75.6	O_2	−193	15.7
	20	72.8	水银	20	465
	60	66.2	肥皂溶液	20	25.0
	100	58.9	苯	20	28.9
CCl_4	20	26.8	乙醇	20	22.3

例 6-3-1 由于表面张力存在，很多液体表面都呈曲面形状，致使液面内外存在压强差(称为曲面附加压强)。计算球形液滴内外的压强差。



解： 设想将液滴对分成相同的两半，它们之间通过表面张力产

[31] 严格说来，表面能是在等温条件下能够转变为机械能的那部分表面内能，在热力学中称为表面自由能。

[32] 本章中所有物质特性于温度有关的物理解释将在 17 章中说明

[33] 表面活性剂都是有机化学试剂，是同时兼有亲水性与疏水性的线型分子：一端带有极性基团(如烃基)，它们和水能生成氢键，因而常称为亲水基团，另一端为非极性基团(如碳氢链)，和水不能形成氢键，相互作用力也较弱，因而常称为疏水基团，疏水基是亲油的。其中疏水基和油的结构越相近，两者亲和性越好。疏水基团的相互作用是由于亲水基团有聚集在一起的倾向，从而反映出疏水基团也会聚合在一起。另外每个烃基应含有 8 个碳原子以上的链才具有较好的表面活性。

生的相互拉力 $2\pi R\sigma$ 应和液滴内、外的压强差达到平衡(见本例图)。内压力作用在半球的大圆面上，$F_i = P_i\pi R^2$，外压力垂直作用在半球面上，其沿轴的分量相当于 P_e 均匀作用在投影面积 πR^2 上。液滴内、外的作用力达到平衡时

$$2\pi R\sigma = \Delta P \cdot \pi R^2$$

则液滴内、外的压强差

$$\Delta P = P_i - P_e = 2\sigma/R$$

即液滴内部的压强大于外部的压强，而该压强差反比于液滴的半径，液滴越小，内外压强差就越大。读者可以自己证明，对于一个球形液膜(如肥皂膜)，只要内、外球面半径相差很小，则膜内压强总比膜外压强高出 $\Delta P = P_i - P_e = 4\sigma/R$。◀

Table 1-6 Physical properties of some common liquids (SI)

液体	T(°C)	ρ(kg·m^{-3})	η(N·s·m^{-2})	σ (N·m^{-1})	B(N·m^{-2})
CCl_4	20	1 590	9.58×10^{-4}	2.69×10^{-2}	1.31×10^{9}
乙醇	20	789	1.19×10^{-3}	2.28×10^{-2}	1.06×10^{9}
汽油	15.6	680	3.1×10^{-4}	2.2×10^{-2}	1.3×10^{9}
甘油	20	1 260	1.5	6.33×10^{-2}	4.52×10^{9}
水银	20	13 600	1.57×10^{-3}	4.66×10^{-1}	2.85×10^{10}
SAE 30 oil	15.6	912	3.8×10^{-1}	3.6×10^{-2}	1.5×10^{9}
海水	15.6	1 030	1.20×10^{-3}	7.34×10^{-2}	2.34×10^{9}
水	15.6	999	1.12×10^{-3}	7.34×10^{-2}	2.15×10^{9}

固-液界面

在液体与固体接触的液体表面上也存在一个表面层，称为**附着层**。附着层的表面能正负决定于液体分子之间及液体分子与邻近的固体分子之间相互作用的相对强弱。若固体分子与液体分子间吸引力的作用半径为 l，而液体分子之间的吸引力作用半径为 R_o，可以取 l 与 R_o 的较大者为附着层厚度。对附着层中某一分子 A 而言，其分子作用球一部分在液体中，另一部分在固体中。因分子作用球内的液体分子的空间分布并非球对称，球内液体分子对 A 分子的吸引力合力(称为**内聚力**)不为零，其方向垂直于液体与固体的接触表面而指向液体内部。而固体分子对 A 分子的吸引力合力(称为**附着力**)则垂直于接触表面指向液体外部。若平均附着力 $f_{附}$<平均内聚力 $f_{内}$，附着层内分子所受到的合力 f 指向液体内部，附着层内分子的引力势能大于液体内部分子的引力势能。引力势能的差值称为附着层分子的**表面能**。这时的表面能显然是正的。若 $f_{附}>f_{内}$，附着层内分子受到的合力指向固体内部，附着层内分子的引力势能小于液体内分子的引力势能，附着层内分子的表面能为负。外界条件一定时，系统总能量最小的状态才是最稳定的，因而在 $f_{附}>f_{内}$ 情形，液体内部分子趋向于进入附着层，扩大附着层的面积，这种现象称为**润湿**。而 $f_{附}<f_{内}$ 情形，附着层内的分子趋向于进入液体内部，减少附着层的面积，这种现象称为**不润湿**。

在固、液、气三者共同接触位置，液面与固体表面切线之间形成的角度 θ 称为**接触角**(图 6-3-5)。显然，θ 为锐角或钝角分别对应于液体润湿或不润湿固体。因而接触角 θ 可用以计量液体对固体的润湿或不润湿程度：习惯上将 $\theta=0$ 或 $\theta=\pi$ 的情形分别称为**完全润湿**或**完全不润湿**。液体能否润湿固体取决于两种物体内部原子或分子的内部作用力以及相互间的作用力的相对大小。例如，水几乎能完全润湿干净的玻璃表面，但不能润湿石腊；水银不能润湿玻璃，但能润湿干净的铜、铁等。

润湿或不润湿有一个常见的观测效应：将内径很细的管子(称为**毛细管**)插入

水中，平衡时管内水面会高于管外的水面。毛细管内径越小，水面升得越高。若将毛细玻璃管插入水银中，毛细管中水银面将低于管外的水银面，内径越小，水银面也降得越低(图 6-3-6)。这种润湿管壁的液体在细管中升高，不润湿管壁的液体在细管中下降的现象称为**毛细现象**。毛细现象由表面张力和接触角所决定。读者可以自己证明(见习题 6-3-？)，若毛细管的半径为 r，水的密度和表面张力系数分别为 ρ 和 σ，接触角为 θ，则毛细管内水柱的高度为

$$h=\frac{2\sigma\cos\theta}{\rho g r} \tag{6-3-24}$$

自然界中有很多现象与毛细现象相联系。植物和动物的大部分组织都是以各种各样管道连通起来的，植物根须吸收的水及无机质物依靠毛细管输送到茎、叶上去。

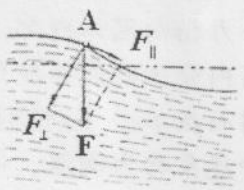
相互不溶的两种液体的接触界面

在两种不同类的液体的接触面上，也都各自有一个表面层。类似于在液体与固体的接触面上可出现正表面能或负表面能的情况，读者可以类似地分析两种不同类的液体的接触面上所可能呈现的表面现象及内在机理。

6-3-2 流体静力学

为简单起见，我们首先讨论处于静态的流体所具有的应力特性。静态流体对应于两种力学状态：整个流体处于静止状态，或者流体处于整体运动状态，但流体中的质量元之间没有相对运动。

由于处于静态的流体不能维持剪切应变，因而只有当流体质元表面和它所受到的作用力垂直时，液体才可能保持静止。设想液面上一流体质元 A 所受的重力 **F** 与液面成一角度(见图 6-3-7)，则质元 A 将在沿液面切线方向上的分力 $F_{\parallel}$ 作用下沿液面移动(改变液面的形状)直到 $F_{\parallel}$ 消失，即作用力与液面垂直时($\mathbf{F}=F_{\perp}$，**正应力**)才停止。因此静止流体体积元上的应力的切向分量恒等于零。正是因为处于静态流体中的各质元所受到的应力没有剪切分量，流体不能维持自身的形状，例如，重力场中的流体形状取决与容器。



流体静力学所涉及的问题主要是处于静止状态的流体内部的压强分布。

弧立流体的内部应力分布

由于流体只有正应力，流体中的压强 P 是各向同性的，即与面元 ΔS 的取向无关。考虑流体内的一直角三角柱体元(图 6-3-8)，其体积为 $\Delta V=\frac{1}{2}\Delta x\Delta y\Delta z$，质量为 $\Delta m=\rho\Delta V$。作用在体积元上的三个压强分量分别为 P_x, P_y, P_z，相应的体积元的三个方向的运动方程为

$$P_x\cdot\Delta y\Delta z-P_n\cos\theta\cdot\Delta l\Delta z=\Delta m a_x$$

$$P_y\cdot\Delta x\Delta z-P_n\sin\theta\cdot\Delta l\Delta z=\Delta m a_y$$

其中 $\Delta m=\rho\cdot\frac{1}{2}\Delta x\Delta y\Delta z$，利用几何关系：$\Delta l\cos\theta=\Delta y$，

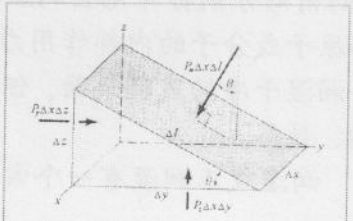


图 6-3-8

$\Delta l\sin\theta=\Delta x$，以及对于流体元有 $\Delta x,\Delta y,\Delta z\to 0$，可得

$$P_x=P_y=P_n$$

同理可证 $P_z=P_x$，$P_z=P_y$。于是得到：

$$P_x=P_y=P_z=P_n \tag{6-3-25}$$

即流体中的压强是各向同性的。这一结果称为**帕斯卡定律**[34]。该定律对于流体静止以及流动的情形都成立。

帕斯卡原理最直接的应用是利用液体传递压强，虽然静止流体内任意两点间可能存在压强差(见下一小节)，若流体中某处压强增大 ΔP，必然导致流体中每点都增大 ΔP 以维持任意两点间的压强差不变，使流体处于宏观静止状态。因此作用在密闭容器中流体上的压强变化能够等值地传递到流体各处和容器壁上。图 6-3-9 所示为油压机或水压机的封闭容器两端的大活塞和小活塞所受到的液体压强均为 P，若小活塞横截面积为 S_1，大活塞横截面积为 S_2，虽然小活塞对流体的作用力仅有 PS_1，而流体对大活塞的作用力却能达到 PS_2。

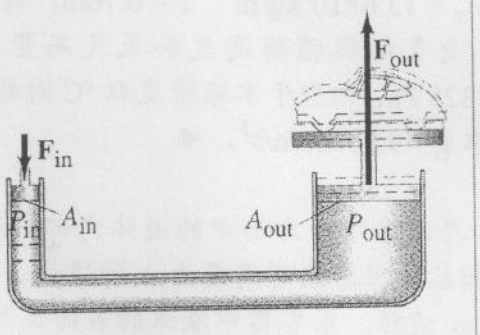


图 6-3-9

外力场中的流体内部应力分布

事实上，孤立流体内部的任一体积元所受到的其它体积元的合力与方向无关，而且所有体积元所受到的合力也完全相同，即孤立流体内部的各点的压强不仅是各向同性的，而且处处相等。但对处在外力场中的流体而言，其内部的各点的压强仍是各向同性的，但各点之间的大小存在差异，下面分别就两种情况予以说明。

(1) 处于重力场的流体内部的压强的分布

考虑图 6-3-10 中流体内部的一柱体形状的流体元处于静力学平衡状态。其中长方体的水平顶面面积为 S，与流体顶表面的距离为 z，高度为 Δz。设流体元顶部的压强为 P，流体元底部的压强为 $P+\Delta P$，则流体元在竖直方向的力学平衡方程为

$$(P+\Delta P)S=PS+\rho gS\Delta z.$$

其中为流体密度。取极限 $\Delta z\to 0$，上述方程变为

$$\frac{dP}{dz}=\rho g \tag{6-3-26}$$

若流体的密度与深度无关，流体内部的压强随深度的分布为

$$P=P_o+\rho gz \tag{6-3-27}$$

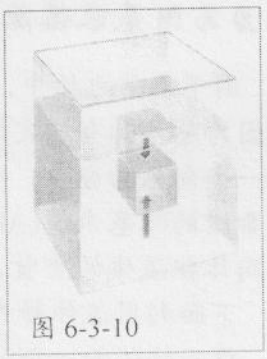
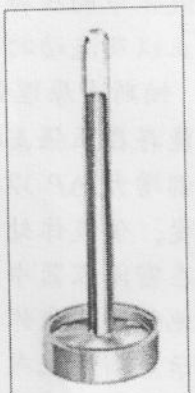


图 6-3-10

[34] **帕斯卡**(Blaise Pascal，1623-1662) 法国数学家、物理学家、宗教哲学家。帕斯卡早期进行自然和应用科学的研究，对机械计算器的制造和流体的研究作出重要贡献，扩展托里切利的工作，澄清了压强和真空的概念。数学上，帕斯卡促成了两个重要的新研究领域。他 16 岁写出一篇题为射影几何的论文，1654 年开始与皮埃尔·费马通信讨论概率论。他在 1654 年末一次神秘主义经历后专注于沉思和哲学与神学写作。其人文思想主要受蒙田的影响。他的宗教论战之作《乡巴佬书信》被奉为法文写作的典范，身后其笔记本被编为《思想录》。

例 6-3-2 1643 年意大利的托里拆利(Torricelli)利用上述原理发明了测量大气压强的水银气压计：将一端封闭的长玻璃管充满水银，然后倒放于水银槽中。管内水银面下降到一定程度即停止，留下的空间除水银蒸汽外没有其它气体。常温下水银蒸汽压可忽略，量得水银柱高 0.76m，求大气压强。


例 6-3-2 图

解： 如例图所示，在管内与槽内水银面等高的点 2 处压强相等：

$$P_{air} = P_m = \rho g z$$

因 $\rho_m = 13.6\times10^3 \text{kg/m}^3$，$z = 0.76\text{m}$，计算得到大气压为 101 Pa。

大气的压强随高度和天气而变，标准大气压(atm)定义为 101325 Pa，相当于水银密度取 ℃ 时的值 13595.1kg/m^3，重力加速度取 g=9.80665m/s^2。◀

(2) 在旋转惯性力场中的液体内部的压强分布。考虑在半径为 R 的圆柱型容器装有高度为 h 的液体。若使容器绕其对称轴以角速度 ω 旋转，求容器中液体的表面在达到稳定后的形状。液面上任一位置的体积元的运动方程为

$$\rho\mathbf{f} - \rho\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}) = \nabla p$$

其中 $\mathbf{f} = -\nabla(gz)$为重力密度，$-\rho\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}) = \nabla(\frac{1}{2}\rho\omega^2 r^2)$为向心力密度，由此得到

$$p + \rho g h - \tfrac{1}{2}\rho\omega^2 r^2 = \text{const.}$$

由于自由表面上 $p = \text{const}$，自由表面的形状的方程为

$$z = z_o + \frac{\omega^2}{2g} r^2$$

由于液面的最高和最低处分别对应于容器的边缘和容器的轴线处

$$h_{\max} - h_{\min} = \frac{R^2}{2g}\omega^2$$

浮力与阿基米德原理

由于在重力场中，流体内部的压强分布是随深度增加的，浸没在流体中的物体因为总是具有一定的体积，上下两个表面受到的压强之间的差异使物体将会受到一个向上的合力，称为**浮力**。物体在流体中受的浮力计算公式是公元前 3 世纪由希腊的阿基米德(Archimedes)提出的：物体在流体中所受的浮力等于该物体排开同体积流体的重量，与物体所在液面下的深度无关。

下面利用流体静力学方程推导出上述**阿基米德原理**。

设一物体的一部分或全部浸没于流体中，其表面因与流体接触而受到流体压力，物体表面各面元所受流体压力的矢量和构成物体所受的浮力。设想物体浮在液体表面上(图 6-3-11)，并忽略液体上大气的影响[35]，物体浸在液体中的部分表面面元 dS 受到的力等于 $dF = \rho g z dS$，其中 ρ 为流体的密度，z 为面元 dS 距液面的深度。计算浮力只需计算液体对物体表面压力的竖直分量，即 $dF\cos\theta$，故浮

[35] 读者自己考虑液面上有大气，或物体全部浸没在液体中的情况

力为

$$F = \rho g z \cos\theta dS / \cos 0 = \rho g z \cos\theta dS = \rho g dV$$

其中 $\cos\theta dS$ 是面元 dS 在水平面上的投影，由于等式右端正好是被排开液体的重量。对浸在液体中的表面 S 上的积分化为对被排开液体体积 V 的积分。

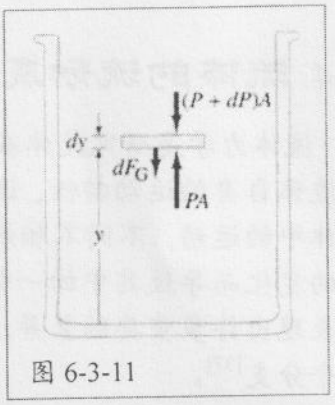


图 6-3-11

图 6-7 表明：在浸没在液体中的物体表面各点受到的液体压力都与该表面元垂直。由于产生作用在相等深度各表面上的压强的力的矢量和为零，而作用在物体底部各表面元上的压强大于作用在物体顶部各表面元上的压强，This net pressure distributed over the bottom of the boulder 形成一个向上的净力—浮力

$$F = \int_{side} P(z)dS + \int_{top} P(z)dS - \int_{bottom} P(z)dS = (P_b - P_t)S$$

$$F_{buoyancy} = \rho_{\text{liquid}} \quad ghS = \rho_{\text{liquid}} Vg = m_{\text{liquid}} g$$

其中 $V = hS$ 是被排开的液体体积[36]。因此等式右端正好是被物体排开的液体重量。即物体在流体中所受到的浮力。

浮力是作用在被物体所排开的同体积的液块的质心(重心)上的，这个点称为浮体的**浮心**。只有浮心 B 高于浮体的质心(重心)G 时(见图 6-3-5)，浮体的姿态才是稳定的。在船舶上，把发动机和货物放在底舱以满足稳定条件。在帆船上则还要设法抵消作用在帆上的力矩，如在船底装上很重的龙骨。在江河湖泊上航行的船只吃水不宜太深，常用宽阔平坦的船底。宽阔的船体也有良好的稳定性，因为当这种船发生倾斜时，浮心就会朝向下倾斜的一侧移动，使浮力与重力组成的力矩能够恢复船体的平衡。

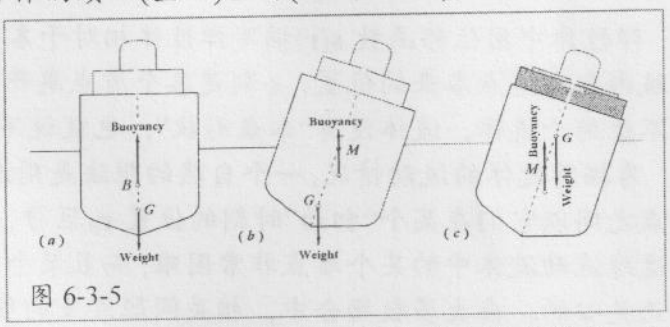


图 6-3-5

36

6-4 流体的流动及其描述

流体力学是研究流体在各种力学环境中演化行为和特性的学科，其中除了讨论流体自身的运动特性，还涉及许多固体情形所没有的复杂情况，如各种固体在流体中的运动，不同不相混合的液体的运动，流体界面的特殊性，甚至因运动状态的变化而导致其中的一部分发生相变等等。另外，液体和气体在流动特性方面也表现出许多重要的差异，流体力学也因此在大类上有水动力学和气体动力学等多个分支[37]。

流体运动形式繁纷多样，如平流，旋涡，波浪和潮汐，动脉中的血液流动，泥石流，烟圈，行星的大气层，星系旋流，甚至核裂变等。令人惊异的是如此多样、复杂的运动类型都能以一般流体运动方程的一个特解的形式呈现出来。只是一般的流体运动方程极难求解，事实上，通常都会根据具体问题所涉及的主要方面以及外部条件引进一些约束性的假设，将变量的数目减少到很少的几个。即使这样，大多也只能通过数值手段求解计算流体的行为和特性。

6-4-1 流动的描述

弹性体中用位移函数 $\mathbf{s}(\mathbf{r})$描写弹性体相对于本来形状的偏离，其中变量 $\mathbf{r}$ 是系统内各个质点本来的位置，$\mathbf{s}$ 则是各个质点离开其本来位置的位移。但这种描写不适用于流体，流体没有"本来形状"，也就说不上"偏离'本来形状'的形变"。

为描写流体的流动情况，一个自然的想法是用流体中各个质点的速度 $\mathbf{v}(\mathbf{r}_o, t)$，质点之间以它们在某个"初始"时刻的位置 $\mathbf{r}_o$ 区分，这种方式称为**拉格朗日描述**。但追踪流动流体中的某个质点非常困难，而且某个质点的命运通常也不是我们所真正关心的。在大多数场合中，相关问题涉及的往往是各个空间点的流速分布，即**速度场** $\mathbf{v}(\mathbf{r}, t)$，这种方式称为**欧拉描述**。如果速度场不随时间而变，相应的流动称为**稳恒流动**(或**定常流动**)。

拉格朗日方法

描述流体运动的最直接的方法是象描述质点运动那样追踪某一个流体质元，求解其运动方程，再由所有流体微元的运动综合出整个流体的相关运动规律。这

---

[37] 流体力学在各个领域都有着广泛的应用。目前已很难找出一个技术部门可以说与流体力学完全没有联系。航空工程和造船工业中的基本问题，如飞机及船的外形设计、操纵性、稳定性等向流体力学提出了广泛的研究课题，在水利工程中，如大型水利枢纽、水库、水力发电站的设计和建造、洪峰的预报工作、河流泥沙等问题都是与流体力学紧密地联系在一起的。在动力机械制造工业中，提高水力及蒸气透平、喷气发动机、压缩机和水泵等动力机械的性能与正确地设计叶片、导流片及其他零件的形状布关。近代热工学在研究如何加速透平机和发动机中的燃烧过程，使灼热的表而冷却等问题中广泛地利用了流体力学的知识。此外，现代气象科学中的基本问题—天气预报，天文学中研究星系螺旋结构，研究组成星云的气状物质的运动，研究气云的膨胀及共相互作用等问题都是宇宙流体力学的内容。近年来，流体力学和其他学科相互渗透，形成了一系列边缘学科，例如电磁流体力学、化学流体力学、高温气体力学、爆炸力学、生物流体力学、地球流体力学、非牛顿流、多相流等。研究非牛顿流体受力和运动规律的学科称为流变学

种描述方法称为**拉格朗日方法**。设想在初始时刻($t=0$)，某一流体质元的位置坐标为 $\mathbf{r}_o$，其运动可表示为

$$\mathbf{r}=\mathbf{r}(\mathbf{r}_o,t) \tag{6-4-8}$$

该流体质元的速度和加速度分别表示为

$$\mathbf{v}=\frac{\partial \mathbf{r}(\mathbf{r}_o,t)}{\partial t} \tag{6-4-9}$$

$$\mathbf{a}=\frac{\partial \mathbf{v}(\mathbf{r}_o,t)}{\partial t}=\frac{\partial^2 \mathbf{r}(\mathbf{r}_o,t)}{\partial t^2} \tag{6-4-10}$$

不同的流体质元可用不同的 $\mathbf{r}_o$ 标示，由此可以得到整个流体的运动图象。在具体的实验观测中可以将不容易扩散的染料滴入流体，观察染色后的流体质元的运动轨迹，流体质元的运动轨迹称为**迹线**。

优点

流体中质元的运动包含平动、转动和形变三种形式，这些运动的特征都可以从微元的运动速度表达式中分解出来：考虑流体中在位置 O 点(坐标为 x, y, z)的某一质元的速度为 $\mathbf{v}_o$，该质元运动到近邻位置 P 点(坐标为 $x+dx, y+dy, z+dz$)的速度可以近似表示为

$$\mathbf{v}=\mathbf{v}_o+\frac{\partial \mathbf{v}}{\partial x}dx+\frac{\partial \mathbf{v}}{\partial y}dy+\frac{\partial \mathbf{v}}{\partial z}dz \tag{6-4-1a}$$

按通常的缩写形式可表示为

$$v_i=v_{oi}+\frac{\partial v_i}{\partial x_j}dx_j \tag{6-4-1b}$$

其中速度矢量关于空间的偏导数$\partial v_i/\partial x_j$共有 9 个分量，构成二阶张量($\nabla\mathbf{v}$)，它可以分解为反对称张量 $a_{ij}=-a_{ji}$ 和对称张量 $s_{ij}=s_{ji}$ 之和：

$$\frac{\partial v_i}{\partial x_j}=\frac{1}{2}\left(\frac{\partial v_i}{\partial x_j}-\frac{\partial v_j}{\partial x_i}\right)+\frac{1}{2}\left(\frac{\partial v_i}{\partial x_j}+\frac{\partial v_j}{\partial x_i}\right)\equiv a_{ij}+s_{ij} \tag{6-4-2}$$

下面分别讨论 a_{ij} 和 s_{ij} 这两个张量的物理含义。反对称张量 a_{ij} 只有三个非零的独立分量(对应于三阶方阵的非对角元)，它们分别对应于速度矢量的旋度的一半：$\boldsymbol{\omega}=\frac{1}{2}\nabla\times\mathbf{v}$。也就是说，这三个分量在形式上构成了一个矢量：

$$\omega_k\equiv\frac{1}{2}\left(\frac{\partial v_i}{\partial x_j}-\frac{\partial v_j}{\partial x_i}\right) \tag{6-4-3}$$

从刚体中任意一点的线速度和刚体整体转动的角速度之间的关系 $\mathbf{v}=\boldsymbol{\omega}\times\mathbf{r}$ 可知，由 a_{ij} 中三个独立分量定义的矢量 $\boldsymbol{\omega}$ 是流体质元中的 P 点绕以 O 点为瞬时轴线转动的角速度。

对称张量 s_{ij} 中的独立元素较多，将其中的对角元和非对角元分别记为

$$\varepsilon_i=\frac{\partial v_i}{\partial x_i},\qquad \theta_k=\frac{\partial v_i}{\partial x_j}+\frac{\partial v_j}{\partial x_i} \tag{6-4-4}$$

下面通过几个简单的特例来说明 ε_i 和 θ_i 这两类分量的物理含义。

(1) 设：$\varepsilon_x\neq 0$，$\varepsilon_y=\varepsilon_z=\theta_x=\theta_y=\theta_z=0$，$P$ 点处流体微元的相应速度分量为：

$v_x=\varepsilon_x\Delta x$, $v_y=v_z=0$

由 $v_x = d(\Delta x)/dt$ 可得 ε_x 的表达式

$$\varepsilon_x = \frac{1}{\Delta x}\frac{d(\Delta x)}{dt} \tag{6-4-5}$$

这说明 ε_x 代表 x 轴线上的线段元 Δx 的相对拉伸速度或相对压缩速度(图 6-4-1)。类似地，$\varepsilon_y, \varepsilon_z$ 分别是 y, z 轴线上线段元 $\Delta y, \Delta z$ 的相对拉伸速度或相对压缩速度。

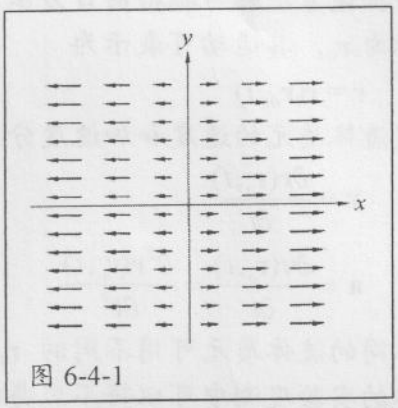


图 6-4-1

(2) 设：$\theta_z \neq 0$，$\varepsilon_x = \varepsilon_y = \varepsilon_z = \theta_x = \theta_y = 0$，则 P 点的相应速度分量为

$$v_x = \tfrac{1}{2}\theta_z \Delta y,\ v_y = \tfrac{1}{2}\theta_z \Delta x,\ v_z = 0$$

记 x 与 y 轴之间的夹角是 γ_{xy}，可得 θ_z 的表达式

$$\theta_z = \frac{v_x}{\delta y} + \frac{v_y}{\delta x} = \frac{d\gamma_{xy}}{dt} \tag{6-4-6}$$

可见 θ_z 代表 x 与 y 轴之间夹角的剪切速度的负值(图 6-4-2)。类似地，θ_x 和 θ_y 分别是 y 与 z 轴，以及 z 与 x 轴之间夹角的剪切速度的负值。

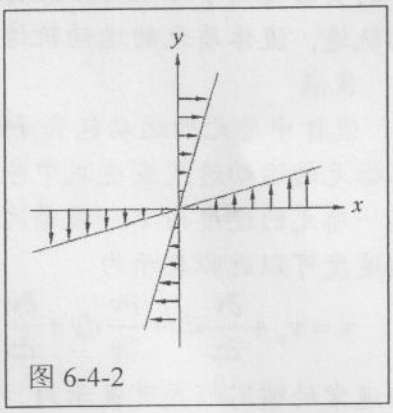


图 6-4-2

由以上事例可知，对称张量 s_{ij} 描述的是物体相对形变的速度。其中对角元描述相对拉伸速度或相对压缩速度，而非对角元描述的是各轴之间夹角的剪切速度。

综合上面的分析，流体微元的运动可通过 P 点的速度(6-4-1)式分解为平动、转动和变形三部分之和：

$$\mathbf{v} = \mathbf{v}_o + \boldsymbol{\omega}\times\Delta\mathbf{r} + \boldsymbol{S}\cdot\Delta\mathbf{r} = \mathbf{v}_t + \mathbf{v}_r + \mathbf{v}_d \tag{6-4-7}$$

上式称为**亥姆霍兹速度分解定理**。

需要说明的是，流体运动中包含的转动成分和刚体运动中的转动成分不同，刚体角速度 $\boldsymbol{\omega}$ 是描述整个刚体转动的一个整体性特征量，而流体的速度旋度只是描述流体微元转动的一个局部性特征量。事实上，刚体速度分解定理对整个刚体成立，是整体性的定理；而流体速度分解定理只是在流体微元内成立，是局部性的定理。

描述流体运动的特征参量是流体微元的速度，其运动特性可以通过流体质元顺应演化过程在某一个时刻到达某一个位置所具有的速度来描述，也可以从进入一个特定时空点的流体质元被赋予一个特定的运动速度这样的视角来考虑。因此流体力学可以从不同流体质元的速度及其演化，或从流体所处空间中不同位置的速度分布这两种视角来描述和讨论流体的运动。

流场方法

另一种描述方式是关注有关流体速度矢量在空间的分布及其随时间的变化而不是特定的流体质元的速度，即用时-空间各点处的速度矢量描述流体的运动状况。这种方法称为**欧勒方法**。

设想在有流体的空间中每一点用相应时刻 t 所在位置的流体质元的速度矢量

$\mathbf{v}$ 标示，这样构成的一个矢量场 $\mathbf{v}(\mathbf{r}, t)$称为**流速场**，或**流场**。我们可以用类似于第四章中用电场线和磁场线描述矢量场的方法，即用场线这种几何方法来描述矢量场的特性。描述流场的场线称为**流线**，流线上每一点的切线方向和流场在该点的速度方向一致(见图 6-4-3)，而流线的密度代表该区域中流体的速度大小。

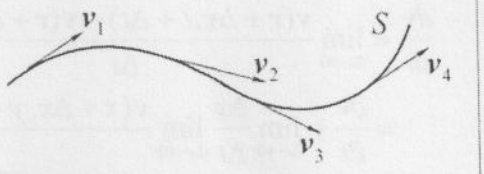


图 6-4-3

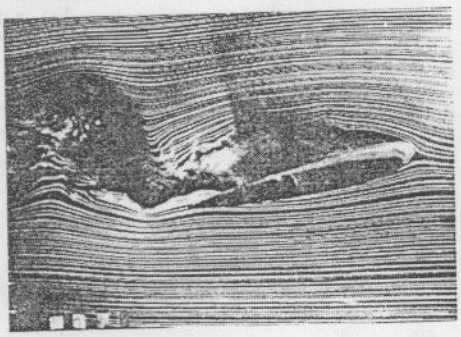


图 6-4-3

不同于描述静态或稳态的作用力场的情形，通过流场的图象要讨论的是流体微元的运动，以及在外力作用下运动状态的改变，这就涉及到如何在流场的描述框架中建立流体的运动方程。

由于流场 $\mathbf{v}(\mathbf{r}, t)$描述的是空间各点的在各个时刻的流体速度分布，其中的参量 $\mathbf{r}$ 所标记的不是特定的流体微元，而是流场中特定的空间位置。用它计算一个流体质元的加速度和用拉格朗日方法的解析方法不同，得到的表达式也会和(6-4-10)式不同。设想在时刻 t，处于位置 $\mathbf{r}$ 的一个流体质元经过 Δt 时间运动到了位置 $\mathbf{r}+\Delta\mathbf{r}$，该流体质元的瞬时加速度应表示为

$$\frac{d\mathbf{v}}{dt}=\lim_{\Delta t\to 0}\frac{\Delta\mathbf{v}}{\Delta t}=\lim_{\Delta t\to 0}\frac{\mathbf{v}(\mathbf{r}+\Delta\mathbf{r},t+\Delta t)-\mathbf{v}(\mathbf{r},t)}{\Delta t} \tag{6-4-11}$$

将其中的分子分解为两个过程

$$\begin{aligned}\Delta\mathbf{v}&=\mathbf{v}(\mathbf{r}+\Delta\mathbf{r},t+\Delta t)-\mathbf{v}(\mathbf{r},t)\\&=\mathbf{v}(\mathbf{r}+\Delta\mathbf{r},t+\Delta t)-\mathbf{v}(\mathbf{r}+\Delta\mathbf{r},t)+[\mathbf{v}(\mathbf{r}+\Delta\mathbf{r},t)-\mathbf{v}(\mathbf{r},t)]\end{aligned}$$

并将[...]中的项在直角坐标系中分解为

$$\begin{aligned}\mathbf{v}(\mathbf{r}+\Delta\mathbf{r},t)-\mathbf{v}(\mathbf{r},t)&=\mathbf{v}(x+\Delta x,y+\Delta y,z+\Delta z,t)-\mathbf{v}(x,y,z,t)\\&=\mathbf{v}(x+\Delta x,y+\Delta y,z+\Delta z,t)-\mathbf{v}(x,y+\Delta y,z+\Delta z,t)\\&\quad+\mathbf{v}(x,y+\Delta y,z+\Delta z,t)-\mathbf{v}(x,y,z+\Delta z,t)\\&\quad+\mathbf{v}(x,y,z+\Delta z,t)-\mathbf{v}(x,y,z,t)\end{aligned}$$

代入(6-4-11)式，可得流体微元的加速度

$$\begin{aligned}\frac{d\mathbf{v}}{dt}&=\lim_{\Delta t\to 0}\frac{\mathbf{v}(\mathbf{r}+\Delta\mathbf{r},t+\Delta t)-\mathbf{v}(\mathbf{r}+\Delta\mathbf{r},t)}{\Delta t}+\lim_{\Delta t\to 0}\frac{\mathbf{v}(\mathbf{r}+\Delta\mathbf{r},t)-\mathbf{v}(\mathbf{r},t)}{\Delta t}\\&=\frac{\partial\mathbf{v}}{\partial t}+\lim_{\Delta t\to 0}\frac{\Delta x}{\Delta t}\lim_{\Delta x\to 0}\frac{\mathbf{v}(x+\Delta x,y+\Delta y,z+\Delta z,t)-\mathbf{v}(x,y+\Delta y,z+\Delta z,t)}{\Delta x}\\&\quad+\lim_{\Delta t\to 0}\frac{\Delta x}{\Delta t}\lim_{\Delta x\to 0}\frac{\mathbf{v}(x+\Delta x,y+\Delta y,z+\Delta z,t)-\mathbf{v}(x,y+\Delta y,z+\Delta z,t)}{\Delta x}\\&\quad+\lim_{\Delta t\to 0}\frac{\Delta y}{\Delta t}\lim_{\Delta y\to 0}\frac{\mathbf{v}(x,y+\Delta y,z+\Delta z,t)-\mathbf{v}(x,y,z+\Delta z,t)}{\Delta y}\\&\quad+\lim_{\Delta t\to 0}\frac{\Delta z}{\Delta t}\lim_{\Delta z\to 0}\frac{\mathbf{v}(x,y,z+\Delta z,t)-\mathbf{v}(x,y,z,t)}{\Delta z}\\&=\frac{\partial\mathbf{v}}{\partial t}+v_x\frac{\partial\mathbf{v}}{\partial x}+v_y\frac{\partial\mathbf{v}}{\partial y}+v_z\frac{\partial\mathbf{v}}{\partial z}\end{aligned}$$

上式可以写成与具体坐标系无关的矢量形式

$$\frac{d\mathbf{v}}{dt}=\frac{\partial\mathbf{v}}{\partial t}+(\mathbf{v}\cdot\nabla)\mathbf{v}\tag{6-4-12}$$

其中第一项表示在一个特定位置(场点)的速度随时间的变化率。利用矢量恒等式：$\mathbf{v}\times(\nabla\times\mathbf{v})=\frac{1}{2}\nabla v^2-(\mathbf{v}\cdot\nabla)\mathbf{v}$，也可将上式改写为

$$\frac{d\mathbf{v}}{dt}=\frac{\partial\mathbf{v}}{\partial t}+\frac{1}{2}\nabla v^2-\mathbf{v}\times(\nabla\times\mathbf{v})\tag{6-4-13}$$

对于任意物理量 P 的空间分布(其中物理量 P 可以是任意矢量，也可以是标量，张量)，其时间变化率都可以通过上述方法表示为

$$\frac{d\mathbf{v}}{dt}=\frac{\partial\mathbf{v}}{\partial t}+(\mathbf{v}\cdot\nabla)\mathbf{v}\tag{6-4-14}$$

称为物理量 P 的随体导数。这在场的分析中是很常用的。

两种方法的说明：

拉格朗日方法所关注的是流体质元，由于所关注的流体质元和流体中其它质元在运动中有相互作用，该方法就面临着和求解多粒子系统一样的运动关联问题。在大多数实际场合中严格求解是不可能的，而且追踪了解每个流体质元的轨迹也没有必要。因而在很多场合都采用欧拉方法来描述流体空间中的速度分布及其演化，以及与之相关的物理现象。例如，在气象观测中所广泛使用的就是欧拉方法：在世界各地星罗棋布地(相当于空间点)设立大量的观测站。根据各气象站在同一时间观测到的气象要素绘制成相应时刻的气象图，并由此做出天气预报。

下面主要采用欧拉方法，即通过流场的时空特性讨论流体运动问题。

6-4-2 流场的空间特性

利用流线分布的几何特征描述流体的流动特性有两个很直接的可观测现象：流量和涡旋。这两种现象在有关矢量场的数学描述中对应于矢量场的通量和环量[38]，它们也都可以通过流场的空间微商来描述各处的局部特征，对应于速度场

[38] 事实上，矢量场的性质都可以由它的通量和旋量来描述，例如，无旋的保守力场。而对

中各个位置的散度和旋度。

流量与连续性原理

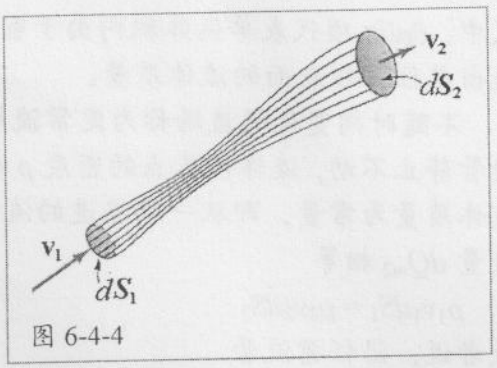


图 6-4-4

在流场的几何描述中，某一局部的流线密度代表该区域中流体的速度大小，它可用于描述在一定时间中流经某一截面的流量。想象在层流流场中的一个面元 dS，通过其周界的全部流线形成了一条管道(见图6-4-4)，称为**流管**，因为流场中每点都有确定的方向，流线不会相交。流体只能通过流管两端的断面 dS_1 和 dS_2 进入或离开这一段流管。因此单位时间内通过某一面元 dS 的流体体积或质量就是以 dS 为断面，长度为 vdt 的一段流管中的流体体积 $d\Phi$ 和质量 dQ

$$d\Phi = dV/dt = v\cos\theta dS = \mathbf{v}\cdot d\mathbf{S} \tag{6-4-15}$$

$$dQ = dm/dt = \rho\mathbf{v}\cdot d\mathbf{S} \tag{6-4-16}$$

这两个参量分别称为体积**流量**或质量**流量**。

设想流场中的一个闭合面 S，其中的流体体积为 V，质量为 m，由于流体质量守恒，封闭面 S 内的质量随时间的变化率等于单位时间内通过该表面流入或流出封闭面 S 的流体总量：

设想在流体中有一封闭面 S，其中的流体体积为 V，质量为 m，流体的质量守恒可以表示为封闭面 S 内的质量随时间的变化率不变：

$$\frac{dm}{dt} = \frac{d}{dt}\int_V \rho\delta\tau = 0 \tag{6-4-17}$$

其中 ρ 为流体密度，根据(6-4-16)式，上式可写为

$$\int_V \frac{\partial\rho}{\partial t}\delta\tau + \int_S \rho v_n \delta S = 0 \tag{6-4-18}$$

称为**连续性方程的积分形式**。利用(6-4-14)式将上式改写为

$$\int_V \left(\frac{d\rho}{dt} + \rho\nabla\cdot\mathbf{v}\right)\delta\tau = \int_V \left(\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v})\right)\delta\tau = 0$$

其中积分区域的体积 V 是任选的，因而被积函数恒等于零，得到

$$\frac{d\rho}{dt} + \rho\nabla\cdot\mathbf{v} = 0 \tag{6-4-19}$$

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v}) = 0 \tag{6-4-20}$$

以上两式都称为连续性方程的微分形式。其中(6-4-19)式可以改写为

$$\frac{1}{\rho}\frac{d\rho}{dt} + \nabla\cdot\mathbf{v} = 0 \tag{6-4-21}$$

于具体的流体而言，相应的物理量是流量和涡旋。

其中第一项表示相对密度变化率，$\nabla\cdot\mathbf{v}$ 表示相对体积变化率。(6-4-21)式说明维持体积元内质量守恒必须要求相对密度变化率等于负的相对体积变化率。而(6-4-20)式中，$\partial\rho/\partial t$ 项代表单位体积内由于密度场不定常性引起的质量变化，$\nabla\cdot(\rho\mathbf{v})$代表流出单位体积表面的流体质量。

不随时间变化的流场称为**定常流场**，相应的流动称为**定常流动**。定常流动中流管静止不动，流体内各点的密度 ρ 也不随时间而改变，因而一段给定流管内的流体质量为常量，即从一端流进的流体质量流量 dQ_{m1} 与从另一端流出流体质量流量 dQ_{m2} 相等

$$\rho_1 v_1 dS_1 = \rho_2 v_2 dS_2 \tag{6-4-22}$$

或者说，沿任意流管

$$\rho v dS = \text{常量}$$

再如果流体具有不可压缩性，则其密度不变：ρ=常量，沿任意流管

$$v dS = \text{常量} \tag{6-4-23}$$

涡旋

描述流场特性的另一个特征是其中是否存在**有旋流**。涡旋提供了最直观的有旋流图像(见图 6-4-5)。数学上通过**环量**来描述有旋流：在流体中作一闭合回路 C，环量 R 定义为流速 $\mathbf{v}$ 沿此回路的线积分：

$$R = \oint_C \mathbf{v}\cdot d\mathbf{l} \tag{6-4-24}$$

环量不恒等于 0 的流动，称为有旋流。

不是所有的有旋流都表现为涡旋，而呈现出涡旋的流场也也并非各处都有旋。从速度分解定理可知，$\nabla\times\mathbf{v}$ 描述流体质元的转动，其方向和大小分别代表质元的瞬时转动轴和两倍的角速度。因此可通过**速度的旋度**是否为零来判断一种流动是否有旋。

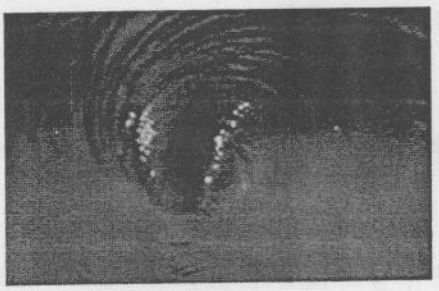


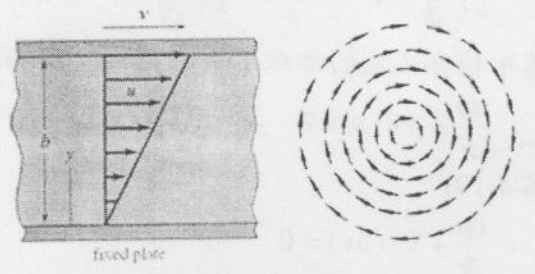
图 6-4-5 涡旋

例 6-4-1 图中显示的流动称为“剪切流”，其速度场为 $v_x = ay$, $v_y = v_z = 0$。其中 a 是常数，图像中没有涡旋的痕迹，判断是否属于“无旋流”。

解： 剪切流动中的质元作直线运动，表面上看似乎是无旋的，但从

$$\nabla\times\mathbf{v} = \left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y}\right)\hat{\mathbf{z}} = -a\hat{\mathbf{z}}$$

可知：剪切流动速度场中各处都是有旋流动。◀



例 6-4-1 图 剪切流动，例 6-4-2 图 点涡运动

例 6-4-2 速度场 $v_\theta = b/r$, $v_r = 0$。其中 b 是常数，流线是以原点为中心的同心圆(见图)。判断该流场中是否各处都属于有旋流动？

解： 表面上看流场中每个质点都绕圆心作圆周运动，似乎应该处处有旋，但实际

上在非原点处，都有。

$$(\nabla\times\mathbf{v})_z=\frac{1}{r}\frac{\partial(rv_\theta)}{\partial r}-\frac{1}{r}\frac{\partial(v_r)}{\partial\theta}=0$$

即除了原点位置，流场中其它位置的流动是处处无旋的，因而这种流动称为**点涡运动**。◀

涡旋环绕的轴线叫做**涡线**。

(1) 涡线、涡面、涡管

对于有旋的流动区域，在任意时刻 t，各场点的 $\Omega=\nabla\times\mathbf{v}$ 也构成矢量场(Ω 称为**涡量**)。因而可以利用矢量场的概念说明涡旋场的一些特性。

涡旋场的场线称为**涡线**，它实际上就是涡旋环绕的轴线，是由同一时刻不同流体质点组成，涡线上各流体微元将绕涡线的切线方向旋转。线上每一点的切线方向和该点的涡量方向重合。涡线的方程为

$$\frac{dx}{\Omega_x(x,y,z,t)}=\frac{dy}{\Omega_y(x,y,z,t)}=\frac{dz}{\Omega_z(x,y,z,t)}$$

其中 $\Omega_x,\Omega_y,\Omega_z$,为 Ω 在直角坐标系的三个分量。

在涡旋场内取一非涡线的曲线，过曲线每一点作涡线，这些涡线组成的曲面称为**涡面**。

如所取的非涡线的曲线 L 封闭，且不自交，则过曲线上每一点作涡线组成一管状曲面，称为**涡管**。若曲线 L 无限小，则称为涡管元。如果在涡管的周围，流体的涡旋皆为零，则称此涡管为孤立涡管(图 2-8-3)。

面积分 $\int_S\mathbf{\Omega}\cdot d\mathbf{S}$ 称为 $\mathbf{\Omega}$ 通过任一截面的**涡通量**。

线积分 $\Gamma=\int_L\mathbf{v}\cdot d\mathbf{r}$ 称为速度向量沿封闭曲线 L 的**环量**。

速度环量和涡通量之间由斯托克斯公式联系起来

$$\int_L\mathbf{v}\cdot d\mathbf{r}=\int_S\mathbf{\Omega}\cdot d\mathbf{S}$$

其中 S 面张于 L 上，法线单位矢量 $\mathbf{n}$ 的正方向与 L 的正方向组成右于螺旋系统(图 2-8-4)。(2-8-2)式写成直角坐标的投影式，有

$$\int_S\mathbf{\Omega}\cdot d\mathbf{S}=\int_L(v_x dx+v_y dy+v_z dz)$$

$$=\int_S\left[\left(\frac{\partial v_z}{\partial y}-\frac{\partial v_y}{\partial z}\right)\cos(n,x)+\left(\frac{\partial v_x}{\partial z}-\frac{\partial v_z}{\partial x}\right)\cos(n,y)+\left(\frac{\partial v_y}{\partial x}-\frac{\partial v_x}{\partial y}\right)\cos(n,z)\right]dS$$

(2-8-2)与(2-8-3)式表明，速度向量 $\mathbf{v}$ 沿封闭曲线 L 的环量等于量 Ω 通过张于 L 上的曲面 s 的通量。

涡通量和速度环量虽然都能表征涡旋强度，但有些情况中，利用速度环量来研究涡旋运动更为方便。因为速度环量是线积分，被积函数是速度本身，而涡通量则是面积分积分，被积函数是速度的偏导数，所以无论是实验和理论利用速度环量常常比利用涡通量简单些。除此之外，有些情况下我们只能利用速度环量的概念来描写涡旋的强度。

(2) 涡旋的物理意义

在 M 点邻域取一与 Ω 垂直的无限小圆，其半径为 α(见图 2-8-5)。写出斯托克斯公式

$$\int_S \boldsymbol{\Omega}\cdot d\mathbf{S}=\int_L \mathbf{v}\cdot d\mathbf{r}$$

其中 L 和 S 分别是小圆的周界及面积。忽略高级小量并引进平均切向速度

$$\bar{v}=\frac{1}{2\pi a}\int_L \mathbf{v}\cdot d\mathbf{r}$$

及平均角速度

$$\overline{\omega}=\frac{\bar{v}}{a}$$

的概念，我们有

$$|\Omega|=\frac{1}{\pi a^2}\int_L \mathbf{v}\cdot d\mathbf{r}=2\frac{\bar{v}}{a}=2\overline{\omega}$$

M 点涡旋矢量大小是流体微元绕该点旋转的平均角速度的两倍，方向与微元的瞬时转动轴线重合。

开尔文涡定理：设想用墨水在理想流体中画上一个闭合回路 C，如果回路不扩散，则无论该回路随流体流到什么地方，其上的环量 C 不变。

开尔文涡定理是角动量守恒的直接结果。设想在流体中有一长 l，半径为 R 的圆柱体，其中的流体绕轴旋转(见图 5-32)，围绕柱体的回路环量为 $\Gamma=2\pi R\mathbf{v}(R)$。若柱体内切向速度的径向分布为 $\mathbf{v}(r)$，则角动量为

$$L=\int \rho v(r) r d^3r=\rho\overline{v(r)rV}$$

其中 $V=\pi R^2 l$ 是柱体的体积，因而

$$\frac{\Gamma}{L}=\frac{2\pi}{\rho V}\frac{Rv(R)}{\overline{rv(r)}}$$

因理想流体内没有粘滞性体的流动可以变形，但因理想流体不可压缩，其体积 V 和密度 ρ 都不变；此外，对于理想流体各层速度的相对分布也是不会变的，即比值 $R\mathbf{v}(R)/r\mathbf{v}(r)$ 不变。故环量 与角动量 L 成正比，两者在流动中皆守恒。

6-4-3 流体运动方程

设想流体中的一个长方体微元，其六个表面的方程分别为

$$x=x_o, y=y_o, z=z_o, x=x_o+\Delta x, y=y_o+\Delta y, z=z_o+\Delta z$$

这六个面所受到的 x 方向的应力分别为：

$$x=x_o:\ -\sigma_{xx}dydz; \qquad x=x_o+dx:\ \left(\sigma_{xx}+\frac{\partial\sigma_{xx}}{\partial x}dx\right)dydz$$

$$y=y_o:\ -\sigma_{xz}dxdz; \qquad y=y_o+dy:\ \left(\sigma_{xz}+\frac{\partial\sigma_{xz}}{\partial y}dy\right)dxdz$$

$$z=z_o:\ -\sigma_{xy}dxdy; \qquad z=z_o+dz:\ \left(\sigma_{xy}+\frac{\partial\sigma_{xy}}{\partial z}dz\right)dxdy$$

其中的负号表示前三个表面的外法线方向和坐标轴的方向相反。六个面的 x 方向

应力总和为

$$F_x=\left(\frac{\partial\sigma_{xx}}{\partial x}+\frac{\partial\sigma_{xy}}{\partial y}+\frac{\partial\sigma_{xz}}{\partial z}\right)dxdydz$$

流体元在 **x** 方向的运动方程为

$$\rho dxdydz\frac{dv_x}{dt}=\rho f_x(x,y,z)dxdydz+\left(\frac{\partial\sigma_{xx}}{\partial x}+\frac{\partial\sigma_{xy}}{\partial y}+\frac{\partial\sigma_{xz}}{\partial z}\right)dxdydz$$

其中 $f(x, y, z)$为流体受到的体积力密度，利用表面应力的表达式(6-3-21)可得 **x** 方向的运动方程为

$$\rho\frac{dv_x}{dt}=\rho f_x(x,y,z)-\frac{\partial P}{\partial x}+2\frac{\partial}{\partial x}\left(\eta\frac{\partial v_x}{\partial x}\right)$$
$$+\frac{\partial}{\partial y}\left[\eta\left(\frac{\partial v_y}{\partial x}+\frac{\partial v_x}{\partial y}\right)\right]+\frac{\partial}{\partial z}\left[\eta\left(\frac{\partial v_z}{\partial x}+\frac{\partial v_x}{\partial z}\right)\right]-\frac{2}{3}\frac{\partial}{\partial x}\left[\eta\left(\frac{\partial v_x}{\partial x}+\frac{\partial v_y}{\partial y}+\frac{\partial v_z}{\partial z}\right)\right]$$

在可压缩流动中，粘滞系数 η 会随温度变化，而温度是时间和空间的函数，因而 η 也是时间和空间的函数。

综合 y 方向和 z 方向类似的受力以及运动方程，可以得到描述流体运动的演化方程的矢量形式：

$$\rho\frac{d\mathbf{v}}{dt}=\rho\mathbf{f}-\nabla P+2\nabla(\eta\nabla\cdot\mathbf{v})$$
$$+\frac{\partial}{\partial y}\left[\eta\left(\frac{\partial v_y}{\partial x}+\frac{\partial v_x}{\partial y}\right)\right]+\frac{\partial}{\partial z}\left[\eta\left(\frac{\partial v_z}{\partial x}+\frac{\partial v_x}{\partial z}\right)\right]-\frac{2}{3}\frac{\partial}{\partial x}\left[\eta\left(\frac{\partial v_x}{\partial x}+\frac{\partial v_y}{\partial y}+\frac{\partial v_z}{\partial z}\right)\right]$$

称为**纳维-斯托克斯方程**[39]。

6-4-4 典型的流体模型

欧勒方法讨论的虽然是速度场，但所涉及的还包括流体的状态参量如密度、压强、温度等在空间的分布及其随时间的变化，即我们关注的不是一个特定的流体微元的演化，由于流体压强 P 和密度 ρ 以及温度 T 等属于物态参量[40]，其中只有两个参量是相互独立的，因此流体的运动可以用流速场 $\mathbf{v}(\mathbf{r}, t)$与流体压强 P 和密度 ρ 的分布用标量场 $P(\mathbf{r}, t)$和 $\rho(\mathbf{r}, t)$得到完整的描述，而流体动力学方程实际上就是这三个场之间的因果及制约关系。流体力学中的主要问题就是求解流体动力学方程。

通常求解一般的流体力学方程相当复杂，但并不是在所有的场合都需要考虑全部的复杂性，许多情况下可以进行不同的简化或近似。

(1) 在压强的变化很小，流动缓慢的情形，流体密度不会有明显的变化：$\rho\approx$ const.。这样的流体称为**不可压缩流体**[41]，纳维-斯托克斯方程可以简化为

---

[39] 纳维(Claude-Louis Navier 1785-1836) 法国工程师与物理学家，主要贡献在力学领域。斯托克斯(George Gabriel Stokes, 1819-1903)，爱尔兰数学家和物理学家，主要贡献在流体动力学，光学和数学物理学。曾任皇家学会秘书和会长。

[40] 有关物态参量的详细讨论见第四部分—热物理学

[41] 需要说明的是，虽然液体比气体难以压缩，却液体并不总可以看作为不可压缩的，而气

$$\frac{d\mathbf{v}}{dt}=\mathbf{f}+\frac{1}{\rho}\nabla P+\frac{\eta}{\rho}\nabla^2\mathbf{v}$$

(2) 完全不可压缩，而且无粘滞性的流体[42]称为**理想流体**。上述方程进一步简化为

$$\frac{\partial\mathbf{v}}{\partial t}+(\mathbf{v}\cdot\nabla)\mathbf{v}=\mathbf{f}+\frac{1}{\rho}\nabla P$$

称为**欧勒方程**。

流体处于静止状态时，$\mathbf{v}=0,\sigma_{ij}=-P\delta_{ij}$，则运动方程退化为平衡方程

$$\rho\mathbf{f}=\nabla P$$

对于均匀流体：$\rho=$ 常数。如果流体质元所受的体积力是保守力，即 $\mathbf{f}=-\nabla V$，V 称为势能密度

$$\nabla(P+\rho U)=0$$

例如，在重力场中：$p-\rho gz=$ 常数

(3) 对于理想流体，而且体积力为保守力：$\mathbf{f}=-\nabla V$（V 称为势能密度）。其运动方程可改写为

$$\frac{\partial\mathbf{v}}{\partial t}+\nabla\left(\frac{1}{2}v^2+V+\Pi\right)+(\nabla\times\mathbf{v})\times\mathbf{v}=0$$

其中 $\Pi=\int\frac{dP}{\rho(P)}$，称为

(i) **定常流动**：$\partial\mathbf{v}/\partial t=0$，运动方程可进一步简化为

$$\nabla\left(\frac{1}{2}v^2+V+\Pi\right)+(\nabla\times\mathbf{v})\times\mathbf{v}=0$$

通过点乘单位速度矢量将上述矢量方程投影到流线方向上

$$\hat{\mathbf{v}}\cdot\nabla\left(\frac{1}{2}v^2+V+\Pi\right)+\hat{\mathbf{v}}\cdot[(\nabla\times\mathbf{v})\times\mathbf{v}]=0$$

得到

$$\frac{\partial}{\partial s}\left(\frac{1}{2}v^2+V+\Pi\right)=0$$

其中 s 为流线，对上式沿流线积分，可以得到

$$\frac{1}{2}v^2+V+\Pi=C(\psi)$$

称为**伯努利积分**。其中积分常数 $C(\psi)$ 通常称为**伯努利常数**，它是流线标识 ψ 的函数，即不同的流线可以有不同的数值。对不可压缩的重流体而言，

$$\Pi=\int\frac{dP}{\rho(P)}=\frac{P}{\rho},\qquad V=gz$$

体也不总被认为是可以压缩的。流体不可压缩近似成立的条件是其流速远小于声音在该介质中的传播速度。空气中的声速约为 300 米/秒，水中的声速约 1500 米/秒。因此在讨论如声波、冲击波类似的情况时，必须计入密度变化的影响。

[42] 在流体非常"稀薄"的情况下通常可以忽略其粘滞性。尽管有的流体粘滞性非常小，但是否完全忽略粘滞性给出的流体行为描述有着重大差别。

伯努利积分的形式为

$$\frac{1}{2}v^2+gz+\frac{P}{\rho}=C(\psi)$$

(ii) 对于**无旋流动**：$\nabla\times\mathbf{v}=0$，速度可以表示为某一个函数的梯度

$$\mathbf{v}=\nabla\varphi$$

运动方程可以简化为

$$\nabla\left(\frac{\partial\varphi}{\partial t}+\frac{1}{2}v^2+V+\Pi\right)=0$$

即括号内的表达式与空间坐标无关，只是时间的函数：

$$\frac{\partial\varphi}{\partial t}+\frac{1}{2}v^2+V+\Pi=f(t)$$

上式称为**拉格朗日积分**。在一个给定时刻，$f(t)$对整个流场取同一数值，这和伯努利积分只是在流线上取同一数值不同。对不可压缩重流体而言

$$\frac{\partial\varphi}{\partial t}+\frac{1}{2}v^2+gz+\frac{P}{\rho}=f(t)$$

流动中不涉及传热以及粘滞性的流体称为**理想流体**

流动中不涉及体积变化的流体称为**不可压缩性流体**。

流动分类

若以运动形式分类，可分为**有旋流动**和**无旋流动**

若以时间分类，可分为定常流动和非定常流动

如果流速场的空间分布随时间变化：$\mathbf{v}=\mathbf{v}(\mathbf{r},t)$，称为**非定常流动**。而空间分布不随时间改变的流速场 $\mathbf{v}=\mathbf{v}(\mathbf{r})$称为流体的**定常流动**。

对于定常流动的伯努利积分，可以理解为是流体处在定常流动中的功能原理的一种表述形式。如图 6-4-4 所示，在作定常流动的理想流体中任取一根流管，用截面 S_1 和 S_2 截出一段流体。在时间间隔 Δt 内，左端的 S_1 从位置 a_1 移到 b_1，右端的 S_2 从位置 a_2 移到 b_2. 令 $a_1b_1=\Delta l_1$, $a_2b_2=\Delta l_2$, 则 $\Delta V_1=S_1\Delta l_1$ 和 $\Delta V_2=S_2\Delta l_2$ 分别是在同一时间间隔内流入和流出的流体体积。对于不可压缩流体的定常流动，$\Delta V_1=\Delta V_2=\Delta V$。流体中没有粘滞，即没有耗散，这段流管内的流体的机械能守恒。在 b_1 到 a_2 一段里虽然流体更换了，但对于定常流动，其运动状态未变，从而动能和势能都没有改变。故考查能量的变化时只需计算两端体元 ΔV_2 与 ΔV_1 之间的能量差。动能的改变为：$\Delta K=\frac{1}{2}\rho\Delta V v_2^2-\frac{1}{2}\rho\Delta V v_1^2$，而重力势能的改变：$\Delta U=\rho g\Delta V(h_2-h_1)$。

现在看外力对这段流管内流体所作的功。设左端的压强为 P_1，作功为 $W_1=F_1\Delta l_1=P_1\Delta V$；右端的压强为 P_2，外力作功为 $W_2=F_2\Delta l_2=P_2\Delta V$。由机械能守恒 $W_{ext}=\Delta K+\Delta U$ 得

$$(P_1-P_2)\Delta V=\frac{1}{2}\rho\Delta V\quad(v_2^2-v_1^2)+\rho g\Delta V(h_2-h_1)$$

或

$$P_1+\frac{1}{2}\rho v_1^2+\rho gh_1=P_2+\frac{1}{2}\rho v_2^2+\rho gh_2$$

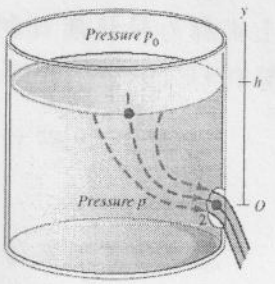
因 1、2 是同一流管内的任意两点，所以上式也可表达为沿同一流线

$P + ½\rho v^2 + \rho gh =$ 常量

上式即为伯努利方程。

再看几个简单的演示实验。向两张平行放置的纸张中间吹气，两张纸就会贴在一起；将一个乒乓球放在倒置的漏斗中间，用口向漏斗嘴里吹气，乒乓球可以贴在漏斗上不坠落(见图 6-4-1)。这都是气流通过狭窄通道时速度加快、压强减少的结果。类似地，两艘同向高速行驶的车、船靠得过近就有相撞的危险[43]。在船长的航海指南里，应当对两条同向并行船舶的速度和容许靠近的距离，加以明确的规定。

例 6-4-1 大桶侧壁有一小孔 A，桶内水面距离小孔的高度差为 h(见例图)，求水从小孔流出的速度。



解： 设想一根从水面 B 到小孔 A 的流管，因圆桶横截面积比小孔面积大很多，水面一端速度几乎为 0，而流管两端的压强都是大气压 P_o，因此伯努利方程的形式为

$$P_o + \frac{1}{2}\rho v^2 + \rho gh = P_o$$

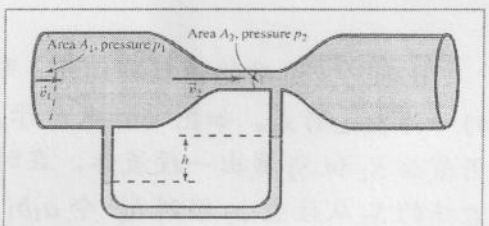
由此得小孔流速为

$$v = \sqrt{2gh}$$

与流体质点从自液面 B 自由下落到小孔的高度时的速度相等，由于实际流体中总是存在一些粘滞性，射流速度略小于理想流体的射流速度。◀

例 6-4-2 一具有缩颈的水平流管(**文托利管**，见图)，，求水从小孔流出的速度。

解： 文丘里(Venhri)流量计通过用 U 形管水银压差计测量出演



图

$$S_1V_1 = S_2V_2$$

对于文丘里管粗细处的，伯努利方程的形式为

$$P_1 + \frac{1}{2}\rho v_1^2 = P_2 + \frac{1}{2}\rho v_2^2$$

文丘里管粗细处的压差 ΔP 来推算流量：

$$\Delta P = P_1 - P_2 = \frac{1}{2}\rho(v_1^2 - v_2^2) = \frac{1}{2}\rho v_2^2[1-(S_2/S_1)^2] = \frac{1}{2}\rho(Q^2/S_2^2)[1-(S_2/S_1)^2]$$

流量

$$Q = S_2\sqrt{\frac{2\Delta P}{\rho[1-(S_2/S_1)^2]}}$$

流管细的地方流速大，压强小。喷雾器、水流抽气机、内燃机中用的汽化器等都

[43] 历史上曾多次发生这样的事故：二十世纪初一支法国舰队在地中海演习，勃林奴斯号装甲旗舰召来一艘驱逐舰接受命令。高速开来的驱逐舰到旗舰附近突然向它的船头方向急转弯，撞在其船头上被劈成两半，1942 年玛丽皇后号运兵船从美国开往英国，与之并行的一艘护航巡洋舰突然向左急转弯，撞在运兵船的船头上而被劈为两半。

利用截面小处流速大、压强小的原理制成。

例 6-4-3 如图所示，水旋转着从圆筒底的漏水口流出。求水面的形状方程。

解：将水近似为理想流体，其环量守恒。当水面的环流向下流动并向中央集中时，积分回路的周长正比于半径 r，因此速度的切向分量按 r^{-1} 的比例增加。若不旋转，沿径向向内并向下的流管的横截面积也正比于 r，速度的径向和向下的分量也按 $1/r$ 的比例增加。在水面上压强处处相等(皆为大气压)，根据伯努利原理，在沿水面的流线上(阿基米德螺线)

$$\rho g h + \tfrac{1}{2}\rho v^2 = \text{const.}$$

式中 $v^2 \propto 1/r^2$，故有

$$h - h_o \propto 1/r^2$$

流体的升力

固体在流体中运动时除了受到平行或反平行于运动方向的阻力或曳引力，还可能受到垂直于相对速度的作用力，例如在空气中向前运动的机翼会受到的向上的作用力。流体力学中，不管这种横向力是否朝上，一般统称为升力。伯努利方程也显示出了这种迹象，图 5-45c 中球上下的流线都来自远方的上游，在那里压强是一样的。球上边流线密，流速大，压强小；球下边流线稀，流速小，压强大。所以球受到向上的升力，使其轨道向上弯曲。1906 年，俄国物理学家茹可夫斯基[44]提出，升力与流速场绕物体的环量成正比。

为简单起见，设想一个刚性长条状物体在 z 方向的线度在 (x, y) 平面内以匀速 $-\mathbf{u}$ 穿过静止流体，这种情形可用二维模型进行分析。相对于物体运动的惯性系，流体总体上以速度 $\mathbf{U} = U\hat{\mathbf{x}}$ 流动。流速场是 $\mathbf{u}$ 和扰动场 $\mathbf{v}(x, y)$ 的叠加：$\mathbf{u} + \mathbf{v}$。

在 (x, y) 平面内取一环绕该物体的闭合曲线 C(不需要紧贴着物体表面，即可圈进部分流体)，取逆时针方向为正，$d\mathbf{l}$ 为 C 上的线元矢量，则 $d\mathbf{l}\times\hat{\mathbf{z}} = d\mathbf{S}$ 为 $\mathbf{z}$ 方向厚度为一个单位的固体侧面 S_c 外法向面元矢量(见图 5-44a)。外部流体作用在面元 dS 上的力由流体静压力 $-PdS$ 和单位时间内流进 dS 的动量两部分构成。设 dt 时间内流进 C 内的流体质量为 $dm = -\rho(\mathbf{u} + \mathbf{v})\cdot d\mathbf{S}dt$，带进的动量为 $(\mathbf{u}+\mathbf{v})dm = -\rho(\mathbf{u} + \mathbf{v})[(\mathbf{u} + \mathbf{v})\cdot d\mathbf{S}dt]$，相应的作用力为 $(\mathbf{u} + \mathbf{v})dm/dt = -\rho(\mathbf{u} + \mathbf{v})[(\mathbf{u} + \mathbf{v})\cdot d\mathbf{S}]$。作用在圈 C 内物质薄层上的力为：

$$\mathbf{F} = -\iint_{S_C} | PdS + \rho(\mathbf{U} + \mathbf{v})[(\mathbf{U} + \mathbf{v})\cdot d\mathbf{S}]| \tag{5-45}$$

取 C_1 紧贴着物体表面，C_2 远离该表面，对定常流场，单位时间里从 C_1 流入的动量与从 C_2 流出的动量相等，C_1, C_2 之间环状体积内的动量不会随时间改变，即对 C_1 或 C_2 计算出来的力是一样的。对回路 C_1 而言，因为没有流体流进物体表面，$(\mathbf{U}+\ \mathbf{v})\cdot d\mathbf{S} = 0$，右端第二项为 0，但压强项中扰动流场 $\mathbf{v}$ 的二次项不好处理。对回路 C_2 而言，无论固体在 (x, y) 平面内的截面是什么形状，从远处看 $\mathbf{v}$ 所引起的扰动流场的分布都渐近地趋于圆柱的情形。计算表明，其径向分量反比于

44

距离的平方[45]，从而 v 的二次项反比于距离的 4 次方。若 C_2 是大致以物体所在位置为中心作一个半径为 R 的大圆，其周长正比于 R。上述积分中含 v 的二次项积分在 $R\to\infty$ 时趋于 0。

先考虑上式右端的压强项。利用伯努利方程，把回路 C 上的变量和上游未受扰动的远处联系起来。在那里压强为 $P_o=$ 常量，流速为 $U=$ 常量。假设高度的影响不大，从而有

$$P+\tfrac{1}{2}\rho(\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v})=P_o+\tfrac{1}{2}\rho U^2$$

即

$$P=P_o-\rho\mathbf{v}\cdot\mathbf{u}-\tfrac{1}{2}\rho v^2$$

其中 P_o 为常数，闭合积分为 0；v^2 项积分时可忽略，故只需保留第二项。

将(5-45)式的被积函数中动量流项展开为：

$$(\mathbf{u}+\mathbf{v})[(\mathbf{u}+\mathbf{v})\cdot\mathrm{d}\mathbf{S}]=\mathbf{u}[(\mathbf{u}+\mathbf{v})\cdot\mathrm{d}\mathbf{S}]+\mathbf{v}(\mathbf{u}\cdot\mathrm{d}\mathbf{S})+\mathbf{v}(\mathbf{v}\cdot\mathrm{d}\mathbf{S})$$

第一项将常量因子 **u** 提到积分号外，$\rho(\mathbf{u}+\mathbf{v})\cdot\mathrm{d}\mathbf{S}$ 的积分是流入 C 的流体质量。对于定常流动，此项为 0。末项为 v 二次项，可忽略。故只需保留第二项。因而有

$$\begin{aligned}\mathbf{F}_{\text{lift}}&=\rho\iint_{S_C}[(\mathbf{U}\cdot\mathbf{v})d\mathbf{S}-\mathbf{v}(\mathbf{U}\cdot d\mathbf{S})]=-\rho\iint_{S_C}\mathbf{U}\times(\mathbf{v}\times d\mathbf{S})\\&=-\rho\oint_C\mathbf{U}\times[\mathbf{v}\times(d\mathbf{l}\times\hat{\mathbf{z}})=\rho\oint_C\mathbf{U}\times[\hat{\mathbf{z}}(\mathbf{v}\cdot d\mathbf{l})-d\mathbf{l}(\mathbf{v}\cdot\hat{\mathbf{z}})]\\&=\rho\oint_C\mathbf{U}(\hat{\mathbf{x}}\times\hat{\mathbf{z}})(\mathbf{v}\cdot d\mathbf{l})=-\rho U\hat{\mathbf{y}}\oint_C\mathbf{v}\cdot d\mathbf{l}\end{aligned}\qquad(5\text{-}46)$$

用一对无限靠近的双线 ab 和 a'b'把两个都围绕着固体的回路 C_1，C_2 连通，形成一个不绕固体的回路(图 5-44b)，因流体内涡度处处为 0，在所有不围绕固体的回路上环量恒为 0：沿双线 ab 和 a'b'的积分抵消，故有 $\Gamma_{c1}=\Gamma_{c2}$，即对围绕固体的回路 C，无论大小，其环量 Γ_c 相等。由于理想流体中环量守恒，若流体中原来没有环量，就产生不出来，则理想流体对在其中作匀速运动的固体不施加任何力(既没有逆向的阻力，也没有横向的升力)。该结论显然与实验事实不符，即使粘滞力趋于 0 的情况也并非如此。这便是著名的**达朗贝尔佯谬**。

运动物体自身的旋转是使运动物体周围形成环流的一个重要原因：只要固体表面和流体之间存在摩擦力，固体的旋转就会形成环流。茹可夫斯基定理虽然是对二维流动而言的，但也定性地适用于三维流动。图 5-45a、b、c 分别是平动而不旋转的球体和只旋转无平动球体周围的流线以及前二者的合成。另外，物体形状的上下不对称也是出现绕体环流和升力的重要机制。例如机翼进入气流伊始[46]，周围流线分布如图 5-46a 所示，其前后各有一个流速为 0 的驻点。因其尾部下面的流速大于上面，这剪切流在尾后具有逆时针方向的涡度(见图 5-46b)。由于机翼和附近的气体在总体上角动量守恒，必有环绕机翼的反向环量出现(见图 5-46b 外部的大环)。机翼尾后的涡度很快就被气流带走，剩下反向环流环绕着机翼。环流与原有的流场叠加，最终形成如图 5-46c 所示的流线分布。其特点是驻点移到机翼后缘，在这附近上下流速趋于一致，不再在翼后形成新涡度，从而使环绕

[45] 如果圆柱能够带动附近的流体转动，则扰动流场的切向分量反比于距离的一次方，v 的二次项反比于距离的 2 次方，$R\to\infty$ 时积分亦趋于 0.

[46] 讨论“伊始”意味着流动尚未达成定常态，这时不能用伯努利原理来分析。

机翼的环量趋于一个恒定值，流场分布也趋于定常。这定常流线分布就像马格纳斯效应的情形那样，上密下疏，流动上快下慢。用伯努利原理就可分析机翼受到的升力了。

很多运动，如掷铁饼、标枪、飞碟和高台滑雪都利用了流体给飞行物体的升力作用。澳大利亚飞镖两臂的横截面如机翼般一边平，一边拱起。投掷者执其一臂，投出时让它在倾斜平面内像车轮那样向前旋转(见图 5-48)。图 5-49 是一臂上装了一个小灯泡的飞镖轨迹照片。飞镖在一面飞行一面旋转的过程中，转到上面的臂端向前运动得快，受到空气的升力大；转到下面的臂端向前运动得慢，从而受到空气的升力小。典型的数据是质心速度为 88 km/h，臂端旋转的线速度为 56km/h，所以上下臂端的合成速度分别为 144km/h 和 32km/h。上下升力差产生一个力矩，要使飞镖的平面侧倾，但由于回转效应，飞镖的平面不侧倾，而是产生进动。于是飞镖就像自行车轮子那样不断朝一边转弯，最后回到原处。

6-4-5 湍流

流速很小时，流体中的流动表现为分层的形式：各层流线互不混合，这就是前面讨论的层流。如果控制流速，使其逐渐增加，流体的流线开始出现波浪状的摆动，摆动的频率及振幅随着流速的增加而增长，这种流动状态称为**过渡流**；流速达到很大的数量时，流线变得很紊乱而难以分辨，而且流场中出现许多小漩涡，这种流动状态称为**湍流**。

图

图 5-58 所示为观测香烟产生的烟的上升轨迹。在最初的几厘米，烟柱是直的，烟的流场是层流，但达到一定高度时随着热空气的向上加速，烟就开始不稳定，变得紊乱起来(见)。这是在热气流加速上升的过程中层流变为湍流。与此类似，空气中污染的扩散就是由湍流主宰的。

为了理解，我们考虑一个气球在空气中的下落速度问题。设气球皮的质量为 $m=10\mathrm{g}$，吹胀后直径达 30 cm，空气的粘滞系数为 $1.8\times10^{-5}\mathrm{Pa\cdot s}$，设空气浮力与气球重量基本平衡，如果采用斯托克斯公式求气球在空气里下降的速度 v，则

$$v=\frac{mg}{6\pi\eta r}=1.9\times10^{3}m/s$$

其数值几乎为大气中声速的 6 倍，显然，斯托克斯公式不再适用这种情形。

到目前为止，对于这些复杂的流动还没有很成功的理论模型可以给出完整的描述。这里我们只是利用量纲分析的方法，寻找并估算液体从层流变化到湍流的过程中的控制参量。并利用一些相应的实验曲线对所观测到的现象做一些定性或半定量的介绍。

第7章 振动与波

实际物体的运动自由度的数量几乎都集中物体内部，因而物质的运动及特性更多涉及的是物质内部运动状态的演化。第6章中所涉及的物体中所有质元对外部扰动的同步响应属于**静态响应**。本章所要讨论的是物体在本身不受破坏的情况下，其内在各部分(质点或质元)最典型的运动形式，在其平衡位置附近的振动及其关联，以及对外部与时间相关的驱动的响应特性。物体中最基本的振动是简谐振动，各种复杂的振动现象都可以分解为不同频率的简谐振动的叠加。

本章首先讨论孤立质点的简谐振动以及一般的孤立线性振动，然后介绍多个振子的耦合振动以及连续介质中的耦合振动—波动现象。

7-1 孤立振动系统

振动所描述的是物体(质点)在某一个位置附近作往复运动，如果这种往复运动表现出重复的循环，这种运动形式称为**周期运动**，循环一周所需要的时间T称为运动的**周期**。单位时间内所完成的运动周期称为周期运动的**频率**，记为ν。显然，频率和周期互为倒数：$\nu = 1/T$。描述周期运动的质点位移函数满足关系

$$\mathbf{r}(t) = \mathbf{r}(t + nT) \quad (n = 0, \pm 1, \pm 2, \ldots) \tag{7-1}$$

称为周期函数。

7-1-1 简谐振动及其特征

最典型、也是最简单的周期运动是所谓的**简谐振动**：运动质点的位移随时间的演化可用**简谐函数**，即正弦或余弦函数表示(见图 7-1-1)：

$$x(t) = A\sin(\omega t + \varphi_o) = A\sin\varphi(t) \tag{7-1-1}$$

其中 $x=0$ 是质点的振动平衡位置；A 称为**振幅**，描述质点在运动中所能偏离平衡位置的最大距离；$\varphi = \omega t + \varphi_o$ 称为**相位**，相位每变化 $\Delta\varphi = \omega\Delta t = 2\pi$ 对应于质点完成一次振动，ω 称为振动的**圆频率**，它和振动周期T之间的关系为 $T = 2\pi/\omega$。式(7-1-1)中的 φ_o 称为**初相位**($t = 0$ 时的相位)。图 7-1-1 为频率、振幅相同，但初相位不同的两个简谐振动的位置演化图。

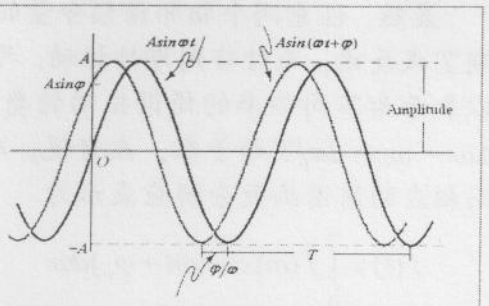


图 7-1-1 频率、振幅相同，但初相位不同的简谐振动。

质点处于简谐振动时的速度和加速度分别可以通过计算质点的位移关于时间的一阶和二阶微商得到：

$$v = \frac{dx}{dt} = \omega A\cos(\omega t + \varphi) \tag{7-1-2}$$

$$a = \frac{d^2x}{dt^2} = -\omega^2 A\sin(\omega t + \varphi) \tag{7-1-3}$$

比较(7-1-1)和(7-1-3)式可知，处于简谐振动中的质点加速度和质点的位移在数值上互相成正比，但方向相反。图7-1-2分别绘出了处于简谐振动中的质点的位移、

速度和加速度随时间变化的情况，其中选择了位移的初相位为零。

虽然正弦型函数具有直观的图像，但有些场合也采用复数形式的指数函数

$$\begin{aligned}\tilde{u}(t) &= Ae^{i\varphi} = Ae^{i(\omega t+\varphi_o)} \\ &= A\cos(\omega t+\varphi_o)+iA\sin(\omega t+\varphi_o)\end{aligned} \tag{7-1-4}$$

其中的虚数部分就是式(7-1-1)所描述的振动。因为这种函数形式描述的振动，其速度以及加速度的函数形式都和描述质点的位移表达式相同(只相差一个常数)：

$$\tilde{v} = \frac{d\tilde{u}}{dt} = i\omega Ae^{i(\omega t+\varphi_o)} = i\omega\tilde{u} \tag{7-1-2'}$$

$$\tilde{a} = \frac{d^2\tilde{u}}{dt^2} = -\omega^2 Ae^{i(\omega t+\varphi_o)} = -\omega^2\tilde{u} \tag{7-1-3'}$$

这对于求解各类振动的运动方程会更为简洁、方便。

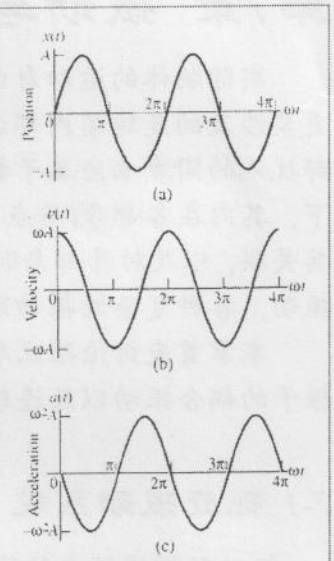


图 7-1-2 质点简谐振时的位移、速度和加速度随时间的变化。

7-1-2 简谐振动的合成与分解

在各种振动的研究中通常都特别强调简谐振动的原因，除了简谐振动能够提供振动的最普遍特征外，还在于各种形态的周期振动 $f(t)$ 都可以分解为多个不同频率的简谐振动的叠加：

$$f(t) = \sum_{n=0}^{\infty}\tilde{f}(\omega_n)\cos[\omega_n t+\varphi_n] \qquad (\omega_o = 2\pi/T,\ \omega_n = n\omega_o) \tag{7-1-5}$$

这种表示称为周期函数 $f(t)$ 的**傅里叶级数展开**，其中 ω_o 称为周期函数 $f(t)$ 的**基频**，而其它简谐频率 ω_n 称为**谐频**，所有简谐分量的集合 $\{\tilde{f}(\omega_n)\}$ 称为 $f(t)$ 的**频谱**。

显然，任意两个相邻谐频分量的频率间隔 $\Delta\omega = \omega_n - \omega_{n-1} = \omega_o$ 与函数 $f(t)$ 的周期 T 成反比，而对非周期性振动，可以理解为周期无限长的周期运动，因而也可以表示为不同频率的简谐振动的叠加，只是其所含的相邻谐频分量的频率间隔 $\Delta\omega = \omega_o = 2\pi/T$ 趋于零，或者说，非周期函数 $f(t)$ 所含的频率分布是连续的，因而相应的简谐函数分解应表示为

$$f(t) = \int_0^{\infty}\tilde{f}(\omega)\cos(\omega t+\varphi_\omega)d\omega \tag{7-1-5'}$$

称为傅里叶积分。因此理解基本的简谐函数及其叠加的特征也是分析一般振动函数的重要方法。

下面先通过一些简单，但典型振动叠加效应的实例，使读者对上面描述的傅立叶展开有一些基本的感性认识。而在以后各小节中做更一般的介绍。

相同方向的简谐振动叠加

图7-1-3中左面三个图形给出的是由两个振动方向相同，频率相同，但振幅和相位不同的两个简谐振动叠加后的结果。其合成后的振动仍然是具有原来频率的简谐振动，只是振幅和初相位出现了变化。图7-1-3右面三个图形给出的是振动方向相同，而频率不同的两个简谐振动叠加后的图像，但两个振动的频率比值为一有理数。图像显示出合成后的振动不再是简谐振动，但仍然是周期振动。其振

动周期就是两个频率所对应的最小公倍数。而对于两个一般的不同频率的简谐振动(两个频率值不存在有限的最小公倍数)的叠加，合成后的振动则不构成周期振动(见习题7-1-1)。

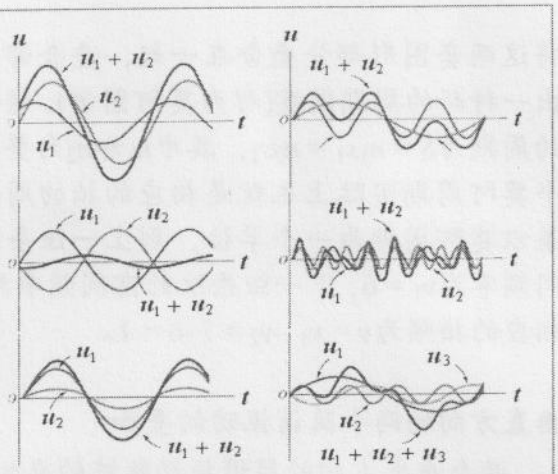


图7-1-3 相同及不同频率的两个简谐振动的叠加。

例7-1-1 讨论两个振幅相同，频率相近的简谐振动叠加后的运动特征，以及可能的应用。

解：设两个相关简谐振动为

$$u_1 = A\sin(\omega_1 t - \varphi_1)$$

$$u_2 = A\sin(\omega_2 t - \varphi_2)$$

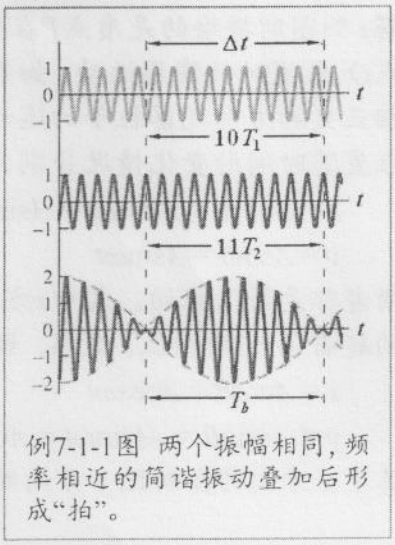
则合成振动可表示为

$$u = u_1 + u_2 = 2A\cos[\tfrac{1}{2}(\Delta\omega t - \Delta\varphi)]\sin(\varpi t - \varphi_a)$$

$$= A(t)\sin(\varpi t - \varphi_a)$$

其中$\varpi = \frac{1}{2}(\omega_1+\omega_2)$为平均频率，$\Delta\omega = \omega_1 - \omega_2$为频率差，$\varphi_a = \frac{1}{2}(\varphi_1+\varphi_2)$，$\Delta\varphi = \varphi_1 - \varphi_2$。

由于$\omega_1 \approx \omega_2$，则$A(t) = 2A\cos[\frac{1}{2}(\Delta\omega t - \Delta\varphi)]$是随时间缓变的函数，因此合成振动是一个振幅随时间缓慢变化的近似简谐振动(见本例图)。这种现象称为**拍**。例图中的u_1在Δt时间内完成了10次振动，而u_2在相同时间内完成了11次振动，它们相应的振动频率分别为$\nu_1 = 1/T_1 = 10/\Delta t$和$\nu_2 = 1/T_2 = 11/\Delta t$。这两个振动的合成构成了一个振幅随时间做周期性的变化的振动，振幅的变化周期为Δt，相应的频率$\nu_b = \nu_2 - \nu_1 = 1/\Delta t$称为**拍频**。需要注意的是以上合成振动的振幅是以$|A(t)|$，而不是$A(t)$来量度的，因此拍频并不是合成振动的函数$A(t)$的变化频率$\frac{1}{2}\Delta\nu$，而是振幅$|A(t)|$的变化频率$\Delta\nu$。实际上$|A(t)|$的变化频率$\Delta\nu$反映的正是拍的测量属性。一般的接收器所能探测的是信号的强弱变化(即信号的能量)，而不能分辨信号的正负值。例如对声音而言，耳朵接收到的是声音的强弱(声能)，正比于振幅的平方：



例7-1-1图 两个振幅相同，频率相近的简谐振动叠加后形成“拍”。

$$I \propto |A(t)|^2 \propto \cos^2[\tfrac{1}{2}(\Delta k x_o - \Delta\omega t - \varphi)] \propto 1 + \cos(\Delta\omega t - \varphi_o)$$

其变化频率为$\omega_b = \Delta\omega$。

“拍”效应有许多实际的应用。例如可用以钢琴的调音：敲击钢琴的一个琴键和一个标准音叉，如果两个声音频率差很小，听到的声音并不是可以清晰分辨的音调，而是两个振动的平均频率。而且响度呈现出有节奏的变化，其频率为钢琴所发出的原来两个声音的频率之差。

上述拍现象是振动叠加形成的**振幅的时间调制**。由于在纸面上无法演示声音及其合成效果，但基于同样的原理，可以通过图形中的周期变化来演示视觉中的“拍”现象。设想两张各印制有等间距平行直线的透明塑料薄膜，两张膜上的相邻条纹间隔分别为s_1和s_2，这实际上就是两套图形各自的周期。图7-1-4所显示的是

将这两套图形部分叠合在一起，叠合部分显示出一种新的周期图案(称为莫阿图案)。该图案中的周期为$S=n_1s_1=n_2s_2$，其中n_1和n_2为整数。一个莫阿周期实际上也就是相应的拍的周期。如果以莫阿周期为一个单位，则上一组条纹的空间频率为$\nu_1=6$，下一组条纹的空间频率为$\nu_2=7$。相应的拍频为$\nu=\nu_1-\nu_2=7-6=1$。

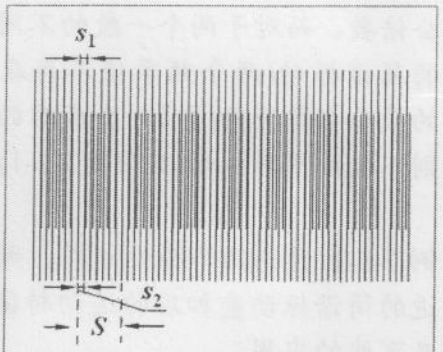
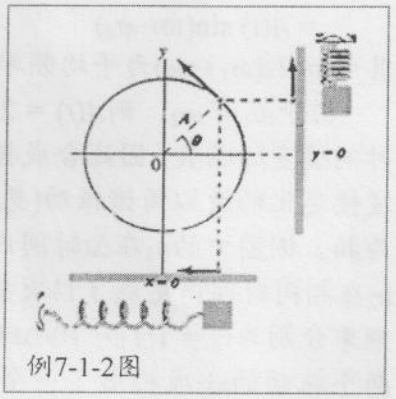


图 7-1-4

垂直方向的两个简谐振动的叠加

两个垂直方向的简谐振动轨迹的叠加可以给出直观的周期图像描述。但在图像中，粒子可以绘出非常丰富多样的运动轨迹，也表现出更多的特征。下面给出几个最简单，但非常典型的振动叠加的例子。

例7-1-2 质点的匀速圆周运动是最简单的二维周期运动，分析运动过程中的简谐振动特性。



例7-1-2图

解：例图所描绘的是质点P在页面内以角速度ω绕原点O做逆时针圆周运动。如果在圆平面内从两个互相正交的方向观察粒子的运动情况，所得到的质点位置随时间的变化情况分别为

$$x=A\cos\theta=A\cos\omega t=A\sin(\omega t+\pi/2)$$

$$y=A\sin\theta=A\sin\omega t$$

两者都是简谐振动。其中x方向的振动比y方向的振动超前了$\pi/2$的相位。显然，绕原点O沿圆周作顺时针运动的质点轨迹参数方程为

$$x=A\cos\theta=A\cos\omega t$$

$$y=-A\sin\theta=-A\sin\omega t=A\cos(\omega t+\pi/2)$$

其中x方向的振动比y方向的振动落后了$\pi/2$的相位。◀

例7-1-3 讨论两个一般垂直简谐振动的合成振动的性质

解：先讨论两个频率相同但初相位不同的简谐振动

$$x=A_x\cos(\omega t-\varphi_x),\quad y=A_y\cos(\omega t-\varphi_y)$$

的合成。上述方程组可以理解为质点运动轨迹的参数方程，消去参数ωt可得到轨迹方程：

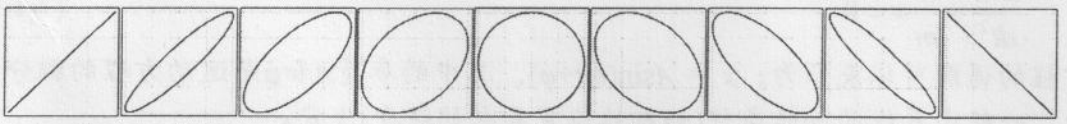
$$\frac{x^2}{A_x^2}+\frac{y^2}{A_y^2}-\frac{2xy}{A_xA_y}\cos\Delta\varphi=\sin^2\Delta\varphi$$

其中$\Delta\varphi=\varphi_x-\varphi_y$为两个振动的相位差。上述方程描述的是一族中心在原点的斜椭圆。不同的相位差对应于椭圆的不同形状，取向和绕行方向，如$\Delta\varphi>0$对应于顺时针绕行的椭圆，而$\Delta\varphi<0$则对应于逆时针绕行的椭圆。其中的几个重要特例为：

1) $\Delta\varphi=0$ 或π，方程退化为直线：$\dfrac{x}{A_x}=\pm\dfrac{y}{A_y}$

2) $\Delta\varphi=\pm\pi/2$，方程退化为正椭圆：$\dfrac{x^2}{A_x^2}+\dfrac{y^2}{A_y^2}=1$

对于$A_x=A_y$(振幅相同)，而且$\Delta\varphi=\pm\pi/2$的特殊情形，运动轨迹退化为圆(见本例图)

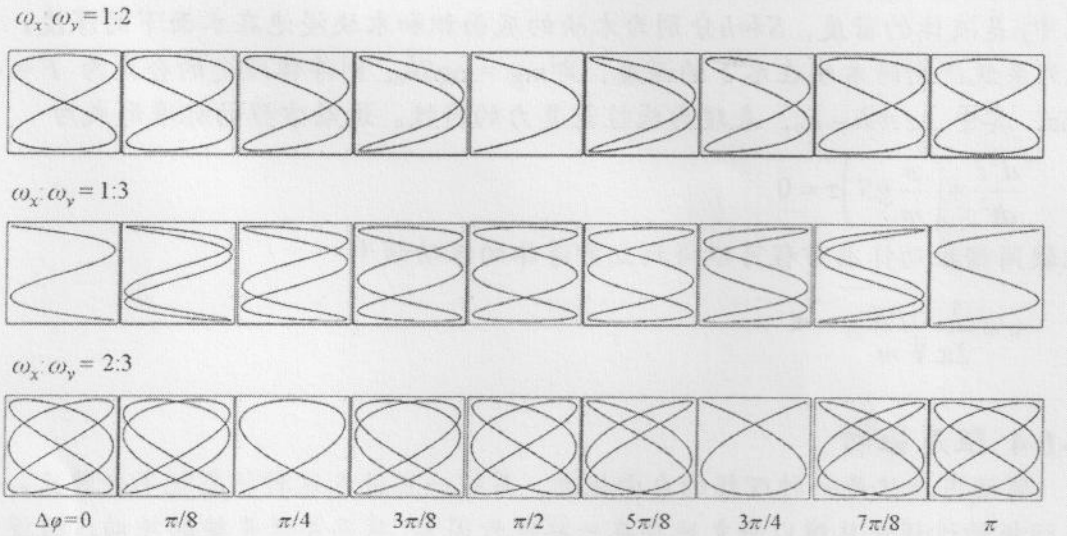


例7-1-2图(*a*)

对于两个振动频率不同，但振动频率比为有理数情形的简谐振动的叠加：

$$\begin{cases} x=A_x\cos(\omega_x t+\varphi_x) \\ y=A_y\cos(\omega_y t+\varphi_y) \end{cases}$$ 其中：$m\omega_x=n\omega_y\ (m,n=1,2,\ldots)$

它们的轨迹是闭合、稳定的。下图是3种最低频率比的质点轨迹(称为李萨如图)。



例7-1-3图

上述分析和图形表明，两个垂直方向的简谐振动的叠加在数学结构上和相同方向的振动合成类似：相同频率的简谐振动的叠加合成为同样频率的简谐振动，而不同频率的简谐振动的叠加只有当两个振动的频率之比是一个有理数，即简单分数的情形，粒子的运动才具有周期性。振动频率比为无理数情形的简谐振动的叠加时，两个独立的垂直振动的合成运动也是没有周期性的，振子的运动轨迹密布在振幅所限定的矩形内。◀

7-1-3 简谐振动的动力学特性

质点处于简谐振动时的位移和加速度的表达式7-1-1和7-1-4表明两者的数量和方向之间具有关系

$$\mathbf{a}=-\omega^2\mathbf{x} \tag{7-1-6}$$

由于质点具有加速度(运动状态变化的一种度量)是物体受到外力的反映：

$$\mathbf{F}=m\mathbf{a}=-m\omega^2\mathbf{x}$$

因而有

$$\mathbf{F}=-k\mathbf{x} \tag{7-1-7}$$

其中比例系数 $k = m\omega^2$ 是正定的。将上述作用力代入牛顿运动方程，可以得到标准的简谐振动方程：

$$\frac{d^2x}{dt^2}+\frac{k}{m}x=0 \tag{7-1-8}$$

方程的通解可以表示为：$x = A\sin(\omega t+\varphi)$，其中的参量$A$和$\varphi$是运动方程的积分常数，一般由系统的初始条件(如初始位置和初始速度)决定。

例 7-1-4 设想一边长为 a 的方木块漂浮在水面上，一个微小的垂直与水面的扰动破坏了木块所受到的重力和浮力的平衡。证明：如果忽略木块在水中运动所受到的耗散性阻力，木块在水中运动是简谐振动，并求出该简谐振动的频率。

解：浮体所受的作用力为浮力和重力：

$$F = mg - \rho gSh$$

其中ρ是液体的密度，S和h分别为木块的底面积和木块浸泡在水面下的深度。设h_o为系统平衡时木块在水下的深度，即$mg = \rho gSh_o$，则浮体所受的合力为 $F = -\rho gSz$，其中，$z = h - h_o$，表现为线性恢复力的特性。运动方程的标准形式为

$$\frac{d^2z}{dt^2}+\left(\frac{\rho}{m}gS\right)z=0$$

比较简谐振动标准方程可以得到立方浮体的振动频率

$$\nu=\frac{1}{2\pi}\sqrt{\frac{\rho g}{m}S}$$ ◀

7-1-4 阻尼振动

简谐振动只是一种理想的自由振动，系统一旦进入这种状态就不会终止。而实际振动过程中总难以避免地存在一些耗散因素，从而导致系统的振动过程逐步衰减。在简谐振动的运动方程中添加阻尼力项就描述这种运动的演化过程，振子受到的最典型的阻尼强度和振子的速度成正比[1]。$F_r = -bv$。其中参量 $b > 0$，以保证 F_r 对任何运动都起着阻尼作用。则振动系统的运动方程为

$$F = -kx - bv = ma \tag{7-1-9}$$

记$\omega_o^2 = k/m$，$\gamma = b/m$，上述运动方程可以写成以下的标准形式

$$\frac{d^2x}{dt^2}+\gamma\frac{dx}{dt}+\omega_o^2x=0 \tag{7-1-10}$$

这类常系数线性齐次常微分方程在数学上有标准解法：令$x = Ae^{at}$，可将关于x的常微分方程(7-2-2)退化为关于参量a的代数方程(见习题7-1-5)。但为了帮助读者形成对振动中各种作用力对振子的影响有更清晰的直觉，我们尝试一种基于物理观察的试探形求解方式，这对以后讨论或者求解更复杂，甚至可能没有解析解的运动方程也有借鉴意义。

首先从经验或直觉上分析阻尼作用将对系统的振动可能产生的影响。从动力学角度而言，阻尼会在两个方面影响系统的振动行为：首先是导致振动能量的耗散，这可以表现为振幅随时间衰减，第二是导致运动减速，这可以表现为系统完

[1] 阻尼作用一般不会是位移的函数(为什么)，而是和振子的速度或位移的高阶导数有关。

成一次振动的时间延长，即系统的振动频率降低。另外我们通常还会有一种直觉：阻尼越大，衰减也越快。

具体求解时，我们考虑两种极限情况。首先考虑阻尼作用远远小于平均弹性力的情况，这时运动方程(7-2-2)退化为：

$$\frac{d^2x}{dt^2}+\omega_o^2x\approx 0 \tag{7-1-11}$$

它描述的就是上一节讨论的简谐振动，得到近似方程的解为：

$$x\approx A\sin(\omega_o t+\varphi) \tag{7-1-12}$$

第二种极限情形是阻尼作用远大于振子弹性力，这时运动方程(7-1-10)退化为：

$$\frac{d^2x}{dt^2}+\gamma\frac{dx}{dt}\approx 0 \tag{7-1-13}$$

两次积分可得振幅随时间衰减的近似解：

$$x\approx Ae^{-\gamma t} \tag{7-1-14}$$

对于介于弱阻尼极限和强阻尼极限之间的一般阻尼振动，其行为应该是振幅随时间衰减的振动，即阻尼振动方程(7-1-10)的解可以表示为特解(7-1-12)式和特解(7-1-14)式的乘积，但考虑到前面所分析的阻尼和振动节奏的相互影响和调制，方程(7-1-10)的解中的动力学参量，即振动频率和指数衰减参数应该不同于极限解(7-1-12)和(7-1-14)中的ω_o和γ，因此设试探解为：

$$x=Ae^{-at}\sin(\omega t+\varphi) \tag{7-1-15}$$

为确定上述参量ω和a，将试探解(7-1-15)代入运动方程(7-1-10)：

$$(\gamma-2a)\omega\cos(\omega t+\varphi)+(a^2-\gamma a-\omega^2+\omega_o)\sin(\omega t+\varphi)=0$$

由于正弦函数和余弦函数两项相互独立，令其分别为零得到

$$a=\tfrac{1}{2}\gamma,\qquad \omega=\sqrt{\omega_o^2-a^2} \tag{7-1-16}$$

阻尼参数a或γ具有时间倒数的量纲，因而可用于表征阻尼振动状态的寿命：

$$\tau=a^{-1}=2\gamma^{-1}=2mb^{-1} \tag{7-1-17}$$

τ 也常被称为振动状态的**弛豫时间**。因而，系统的运动状态可表示为

$$x=Ae^{-t/\tau}\sin(\omega t+\varphi)$$

阻尼振子的振动状态寿命与振子的质量成正比表明惯性大的系统能够有效地降低系统所受到的阻尼。一个常见的例子是一般的摆钟都会采用大质量的摆锤。

下面再讨论阻尼对振动行为的影响以及振动频率对振动寿命的影响。阻尼振动频率表达式(7-2-9)中对阻尼参量而言存在一个临界值：

$$a=\omega_o=a_c \tag{7-1-18}$$

当阻尼参数等于a_c时，振动频率为零，系统振动退化为纯指数衰减：$x=(A\sin\theta)e^{-t/\tau}$，称为**临界阻尼振动**。这时系统能以最短的时间稳定到其平衡位置。

在阻尼参数小于临界值$a<a_c$的范围内，振动频率大于零。系统处于振幅随时间衰减的振动状态，这也是与我们直觉相符合的阻尼振动状态，通常称为**欠阻尼振动**。

对于大阻尼$a>a_c$情形，振动频率退化为虚数：$\omega=i\beta$，系统的运动状态变为

$$x=Ae^{-t/\tau}\sinh(\beta t+i\varphi_o) \tag{7-1-19}$$

上式描述的振动称为**过阻尼振动**，其中A和φ_o是运动方程(7-1-10)解的积分常数，

由系统的初始条件决定。过阻尼振动实际上也已不具有振动的特征，其行为类似于临界阻尼的情形。振动具有过阻尼行为与我们一般的直觉不同，通常我们会认为，增加阻尼能够减小振动幅度，从而使系统更快地回到其平衡位置。但由于$\sinh\beta t$是时间的递增函数，系统回归到平衡位置甚至比临界阻尼状态回到平衡位置所需的时间更长，这意味着过大的阻尼不仅抑止了振荡过程，而且导致系统在向平衡位置回归的过程也受到很大的阻尼。因而在实际问题中，如果希望系统在尽可能长的时间内维持近似简谐振动，应该尽可能减小阻尼，如果希望系统尽可能快地回到平衡状态，需要使阻尼尽可能接近临界阻尼状态，而不是尽可能增加阻尼。

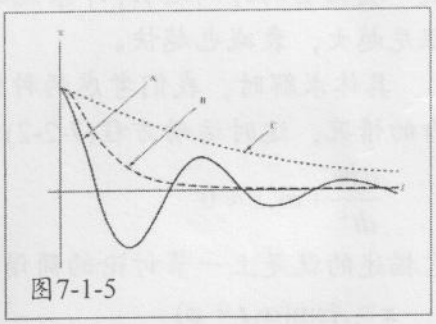


图7-1-5

7-1-5 受迫振动

如果没有外部能量持续、稳定的补充，阻尼振子的振动状态是难以维持的。一种维持系统振动的典型能量补偿方式是对振子施加一简谐力 $F = F_d\cos\omega t$，它可以驱动系统长时间地处于频率为 ω 的简谐振动状态之中，称为**受迫振动**。研究受迫振动的实际意义除了它能揭示系统受到外力作用时振动响应的最基本特征，还由于大部分外力驱动形式可以表示为各种频率的简谐驱动的叠加。通过这种模型可以理解振动系统对各种不同频率驱动的响应特性以及一些相关的物理效应。

受迫振动

受迫振子的运动方程为

$$F = -kx - bv + F_o\cos\omega t = ma \tag{7-1-20}$$

其标准数学形式为

$$\frac{d^2x}{dt^2} + \gamma\frac{dx}{dt} + \omega_o^2 x = \frac{F_o}{m}\cos\omega t \tag{7-1-21}$$

为了得到运动方程的解，我们依然采用前面用过的分步分析方法。

先考虑强驱动的极限情况。这时的运动方程退化为：

$$\frac{d^2x}{dt^2} = \frac{F_o}{m}\cos\omega t$$

直接积分可以解得：$x = C\cos(\omega t + \pi)$，即系统以外部振动频率作受迫简谐振动，但振幅和相位与外部驱动不同(系统的响应落后于外部的扰动)。由此可以认为在一般驱动强度下，系统会作同样的同频率振动，但振幅和相位应有所不同。设运动方程(7-2-14)的试探解为：

$$x = A_o\cos(\omega t - \varphi) \tag{7-1-22}$$

代入方程(7-1-21)可求得试探解中的振幅和相位常数

$$A_o = \frac{F_o}{m\sqrt{(\omega^2 - \omega_o^2)^2 + \gamma^2\omega^2}} \tag{7-1-23a}$$

$$\varphi = \tan^{-1}\frac{\gamma\omega}{\omega^2 - \omega_o^2} \tag{7-1-23b}$$

由于试探解(7-1-22)中所有的参数都由动力学方程确定，这说明它是非奇次方程

(7-1-21)的一个特解。系统的一般解应该是相应奇次方程(7-1-10)的通解(即阻尼振动解)与该特解的和

$$x = Ae^{-t/\tau}\sin(\omega t + \varphi_o) + A_o\cos(\omega t - \varphi) \qquad (7\text{-}1\text{-}24)$$

由于其中的阻尼振动项具有指数衰减的形式，不管初始条件如何，都对应于一个**暂态过程**，而特解则描述了系统的稳定振动过程，称为系统的稳态解。图7-1-6给出了瞬态解、稳态解以及两者的叠加。

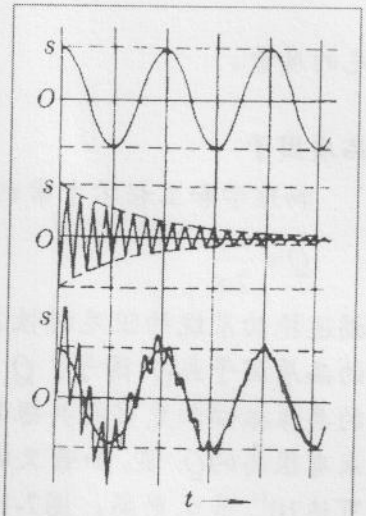


图7-1-6

共振

需要注意的是，描述系统对外界驱动下发生稳定受迫振动的两个参量—振幅和相位都与外界驱动频率有关。式(7-1-23)中的振幅对于不同的外界驱动频率ω存在一个极大值，相应的驱动频率为

$$\omega_r = \sqrt{\omega_o^2 - \gamma^2/2}$$

该现象称为**共振**。相应的频率ω_r称为**共振频率**，该频率就是阻尼振子的固有频率(而不是自由谐振子的固有频率)。

另外，系统**在受破振动过程中**，其振动对外部驱动的响应滞后表现为两者之间存在相位差，式(7-1-23b)表明，相位差随着驱动频率的提高而增长，对于阻尼比较大的振动系统，这种跟随驱动频率的增长主要集中在系统自由振动的固有频率ω_o附近，而对于阻尼很小的振动系统，这种滞后相位随驱动频率的增长比较均匀，但对与阻尼振子的受迫振动而言，无论阻尼大小，都在ω_o处达到$\frac{1}{2}\pi$。其最大的相位滞后数值也都为π(见图7-1-7)。

共振现象也可以通过系统对外界驱动的能量吸收效率来理解。外界的简谐驱动对振子达到稳态受迫振动时的瞬时做功的功率为

$$P(t) = \mathbf{F}(t)\cdot\mathbf{v}(t) = F_o\cos\omega t\cdot\omega A_o\sin(\omega t-\varphi)$$

在一个周期$T=2\pi/\omega$内的平均值(称为**稳态输入功率**)

$$\begin{aligned} P \equiv \overline{P(t)} &= \frac{1}{T}\int_0^T dtF_oA_o\omega\sin\varphi\cos^2\omega t - \frac{1}{T}\int_0^T dtF_oA_o\omega\cos\varphi\sin\omega t\cos\omega t \\ &= \frac{F_oA_o\omega\cos\varphi}{T}\int_0^T dt\cos^2\omega t = \frac{F_oA_o\omega}{2}\cos\varphi = P_o\cos\varphi \end{aligned} \qquad (7\text{-}1\text{-}26)$$

其中P_o是稳态输入功率的极大值。它所对应的是相位部分$\cos\varphi=1$，由

$$\cos\varphi = \frac{\gamma^2\omega^2}{(\omega_o^2-\omega^2)^2+\gamma^2\omega^2} \qquad (7\text{-}1\text{-}27)$$

可知，它所对应的正是发生共振时的频率。由上式也可以解出稳态输入功率取半极大值$P=\frac{1}{2}P_o$，即$\cos\varphi=\frac{1}{2}$(称为**半功率点**)所对应的外部驱动频率值

$$\omega_\pm = \sqrt{\omega_o^2+a^2}\pm a \qquad (7\text{-}1\text{-}28)$$

定义两个半功率点之间的频率间隔为**共振宽度**(也称为共振线宽)

$$\Delta\omega = \omega_+ - \omega_- = \gamma \qquad (7\text{-}1\text{-}29)$$

通常观察振动系统的共振响应比观察其衰减容易，在实验上共振峰宽度可作为阻

尼的度量。

品质因子

物理学和工程学中常利用无量纲参量

$$Q \equiv \frac{\omega_o}{2a} \tag{7-1-30}$$

描述振动系统的阻尼特性及系统相对于共振频率的共振带宽。Q 称为振动系统的**品质因子**或Q **因子**。Q 值高的系统振动持续时间长，受简谐驱动时具有更大的共振振幅和更窄的共振带宽。需要强共振响应或高频率稳定性的共振系统都要求有很高的Q 值。如音叉的Q 值约为 10^3，而原子钟和高品质因子激光器的Q 值可达10^{11} 甚至更高。图7-1-7中分别给出了不同Q 值的振动系统受到外部简谐驱动时其振幅和相位的共振曲线。其中横坐标为驱动频率相对于共振中心频率的偏离，标度单位是$\omega_r/2Q$。

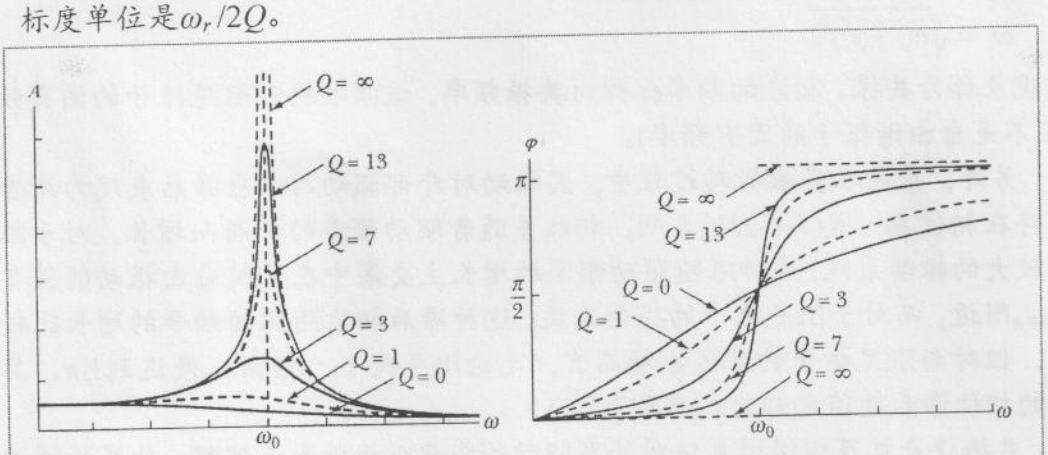


图 7-1-7 振动系统在受到简谐驱动时的振动振幅和相位的共振曲线。

不确定关系

上面的分析表明，含有内在阻尼的振动系统具有两个特性，一是在没有外界能量输入(即无外界驱动)时，其振动状态只能持续有限的时间，描述这一特性的相应参量为平均寿命τ；二是系统只能在有限的频率范围内对外部的简谐驱动(能量输入)产生有效的响应，相应的表征参量是系统的共振峰宽度$\Delta\omega$。因τ和$\Delta\omega$分别正比或反比与振动系统本身的阻尼参量γ：$\tau = \gamma^{-1}$，$(\Delta\omega) = \gamma$，这两个参量互相成反比[2]，因而有：

$$\tau\Delta\omega = 1 \tag{7-1-31}$$

这个关系式揭示了具有内在阻尼的振动系统的一个重要性质：振动衰减越慢的系统对外部周期驱动作用的频率的响应越灵敏，即频率选择性越好。以收音机的稳定性和灵敏度为例，通常为避免所谓的“串台”，收音机需要有较高的频率选择灵敏度，即中心频率以外的载波信号不能得到有效的响应。但由于共振峰宽度$\Delta\omega$窄的系统对信号的衰减很慢，这就使输入信号序列中的脉冲宽度经过系统后被展宽，这导致后续信号与前面的信号发生重叠，使收音机产生啸叫。方程(7-1-31)对只对很窄范围中的驱动频率予以响应滤波器的设计设置了一个极限。滤波器对

[2] 寿命和共振峰宽度和形状有关，也和定义有关，所以这种共性关系只是在数量级上讨论。

频率的选择性得到优化和提高的同时，信号的留存时间越来越长，信号质量也在不断降低，因此实际上只能是在两者之间达到某种妥协。

关系(7-1-31)通常称为不确定关系，它在实验上有很重要的应用。通常观察一个体系的共振响应比观察它的衰减容易，因此可以通过研究体系的共振响应得到响应的频宽来计算衰减的时间。不确定关系在许多领域中都存在，在不同的场合有着不同的物理含义和解释，但原理都是相同的，我们将在后面的相关章节中予以介绍，以帮助读者逐步建立起相关概念的直觉意识。

7-2 耦合振动与驻波

设想一个由若干个振子组成的物理系统，其中每一个振子因偏离平衡位置所受到的恢复力不仅与其本身的位置有关，还与系统中其它振子的位置有关，系统中各个振子的振动由此表现出相互关联，这种系统称为**耦合振子**。从整体振动形态上看，耦合振动有两类特殊的位形，一类表现为局部振子的振动向整个系统中全部的振子扩散，或者说，振动在振子之间传递在系统中形成类似于水面上传播的波，这种振动形态称为**行波**。另一类表现为系统中各个振子独立的振动，但任意两个振子之间的相对位相保持不变，整体上表现为一个不向周边传播的波形，这种振动形态称为**驻波**。这两种振动位型都可以表现出简谐的特性，但即使系统中所有的振子都完全相同，振子之间的关联也完全相同，上述整体的简谐振动却可以有很多个不同的振动频率。一般的耦合振动兼有这两种形态成分，因而系统整体，以及其中每个振子的振动都远不是简谐的。

本节首先通过对若干个具体离散耦合振子的振动分析理解耦合振动的一些基本特征，然后推广到大量的耦合离散振子系统以及连续介质情形，讨论相应的耦合振动行为及特点。

7-2-1 耦合振动的分解—简正模式

图7-2-1所示的是两个由相同质量m 和相同弹性系数 κ 构成的两个简谐振子用一根弹性系数为K的弹簧连接起来的最简单的耦合振动系统。当物体A和B偏离各自平衡位置的位移分别为x_A和x_B时，它们的运动方程分别为

$$m\frac{d^2x_A}{dt^2}=-\kappa x_A-K(x_A-x_B)$$
$$m\frac{d^2x_B}{dt^2}=-K(x_B-x_A)-\kappa x_B \qquad (7\text{-}2\text{-}1)$$

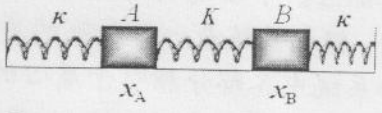


图7-2-1

上述方程表明，由于存在耦合，A和B的振动通常都是非简谐的，即单个物体A和B的振动不能用一个具有确定频率的简谐函数来描述。但这个耦合系统中单个物体是否存在简谐振动的可能性？甚至整个系统处于简谐振动状态的可能性？实际上很容易从图7-2-1中系统的几何对称性看出存在着两个特殊的振动位型：第一个是A和B同步地向左或向右作同方向运动，称为对称振动位型，第二个是A和B分别向左或向右作反方向的同步运动，称为反对称振动位型。在这两种振动位型中，物体A和B总是在同一时刻经过各自的平衡位置。这意味着物体A和B的振动具有相同的频率，因而两个振动具有固定

的相位差。一个耦合振动系统中所有的质点在振动过程中都能在同一时刻经过各自平衡位置，这种振动位型称为耦合系统的**简正振动模式**(简称**模式**)。相应的振动频率称为该模式的**特征频率**或**本征频率**。就图7-2-1所示系统的两种模式而言，在对称模式的振动形态中，因为中间耦合弹簧的长度始终不变，不能表现出耦合的效应，因此，A和B的振动频率就是不存在耦合时各自的自由振动频率

$$\omega_S = \sqrt{\frac{\kappa}{m}} \tag{7-2-2}$$

而在反对称模式中，系统的质心位置不变，每一个振子在偏离平衡位置时所受到的恢复力是边上弹簧和中间半根耦合弹簧的作用力，因此其特征频率为

$$\omega_A = \sqrt{\frac{\kappa + 2K}{m}} \tag{7-2-3}$$

由于不同的简正模式是相互独立的(在数学上表示为两种振动函数是线性无关的)，因此运动方程(7-2-1)的通解就可以表示为这些独立模式的线性组合。上述结果也可以很直接地从系统的运动方程组(7-2-1)所揭示的A和B的位移对称性看出。只要将(7-2-1)中的两个耦合方程相加或相减，并令：$X_1 = x_A + x_B$和$X_2 = x_A - x_B$，可以得到：

$$\begin{aligned} \frac{d^2X_1}{dt^2} &= -\frac{\kappa}{m}X_1 \\ \frac{d^2X_2}{dt^2} &= -\frac{2K+\kappa}{m}X_2 \end{aligned} \tag{7-2-4}$$

即在以X_1和X_2作为运动参量的两种运动模式中，物体A和B的运动是简谐的，而相应的振动频率就是式(7-2-2)和式(7-2-3)所给出的特征频率。X_1和X_2称为**简正坐标**，描述的是集合运动。

由于不同的简正模式是相互独立的(在数学上表示为两种振动函数是线性无关的)，因此运动方程(7-2-1)的通解就可以表示为这些独立模式的线性组合，系统整体以及各个振子的振动形态也因此都失去了简谐特性。

上述耦合系统模式所表现出来的几何对称性对振动系统中是否存在简正模式并不是必须的，耦合系统振动模式所要求的是系统中各部分振动同步的动力学对称性。下面说明这种动力学对称性的特点。

耦合系统是一种特殊的质点组，其运动也可以分为系统的整体运动(质心运动)和系统中各部分相对于质心的内部运动两种。这样各部分振动的适当组合可以形成动力学的同步对称性，如果系统中物体的分布(在这里表现为质量m和弹簧系数κ)具有几何对称性，这种动力学对称性就可以通过振动的几何对称性表现出来。这意味着我们并非必须，或一定能够通过具体的对称性分析得到耦合系统的简正模式，实际上许多具有分布几何对称性的系统所具有的一些简正模式也并不是显而易见的[3]。相反，我们经常通过求解系统简正模式来讨论系统的对称性。仍以上面的系统为例，由于处于简正模式中的所有物体振动频率相同，我们可以将物体A和B在某一个模式中的运动状态表述为

---

[3] 一些特征频率比较低的模式的对称性有时比较明显，而特征频率比较高的模式的对称性通常比较隐蔽。

$$x_j = A_j e^{-i(\omega t-\varphi_j)} = \tilde{A}_j e^{-i\omega t} \quad (j=1,2) \tag{7-2-5}$$

这里采用复数形式的简谐函数是为了简化计算[4]。其中复振幅 $\tilde{A}_j = A_j e^{i\varphi_j}$ 中包含有物体的初位相因子。将(7-2-5)作为试探解代入运动方程组(7-2-1)可将其简化为代数方程组：

$$\begin{aligned}&[\omega^2-(\omega_o^2+\Omega^2)]A_1+\Omega^2 A_2=0\\&\Omega^2 A_1+[\omega^2-(\omega_o^2+\Omega^2)]A_2=0\end{aligned} \tag{7-2-6}$$

其中 $\omega_o^2=\kappa/m$，$\Omega^2=K/m$。上面的奇次代数方程组有非零解的条件是其系数行列式为零：

$$\begin{vmatrix}\omega^2-(\omega_o^2+\Omega^2) & \Omega^2\\ \Omega^2 & \omega^2-(\omega_o^2+\Omega^2)\end{vmatrix}=[\omega^2-(\omega_o^2+\Omega^2)]^2-\Omega^4=0 \tag{7-2-7}$$

这一方程称为确定系统简正模式频率的**特征方程**。由此得到的特征频率和两个物体相应的复数振幅之间的关系为：

$$\begin{aligned}&\omega_1^2=\omega_o^2, && A_1=A_2\\&\omega_2^2=\omega_o^2+2\Omega^2, && A_1=-A_2\end{aligned} \tag{7-2-8}$$

对应于ω_1的第一种模式中有：$A_1=A_2$，表示两个物体的振动有相同的相位和振幅，即上面所说的对称模式；而对应于ω_2的第二种模式中有：$A_1=-A_2=A_2e^{i\pi}$，表示两个物体以相同的振幅作反相的振动(相位差为π)，即反对称模式。这与我们前面讨论的结果一致。

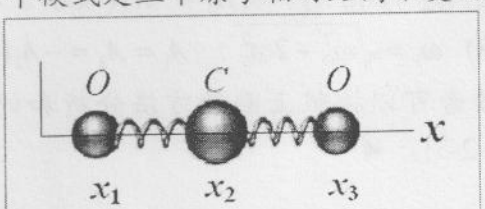
例 7-2-1 二氧化碳 CO_2 是一个线性分子，计算该分子的纵向振动**模式**。

解：就纵向振动而言，该系统的三个原子各有一个自由度，系统相应地有三种振动模式。从系统的运动对称特性考虑，其中一个模式是三个原子相对距离不变的整体平动模式：

$$x_1=x_2=x_3$$

这种运动可以理解为周期无限长的振动，相应的特征频率为零。另外两个模式为系统中所有粒子相对于质心的(内部)运动，因而有

$$m_O x_1+m_C x_2+m_O x_3=0$$



例7-2-1图 CO_2分子的机械结构模型

它包含两根弹簧同相伸缩模式和两根弹簧反相伸缩模式。从对称性考虑，同相伸缩模式中碳原子不动：

$$x_2=0,\quad x_1=-x_3$$

相当于两个氧原子的独立振动，对应的特征频率为 k/m_O。而弹簧反相伸缩的模式中，对称性(质心固定)要求

$$x_1=-x_3,\qquad x_2=-(2m_O/m_C)x_1$$

但这种模式的特征频率不容易从对称性分析中看出。下面我们通过标准的数学程序求解分子的纵向振动模式。各原子的运动方程为

---

[4] 这里我们采用了简谐函数的指数表达形式，因而振幅也是复数(称为复振幅)。这样做是为了数学处理上更加简捷，方便，另外也是为后面量子物理中更抽象波的描述做一些预演。读者可以用更直观的正弦或余弦函数作同样的推导，并先得到的结果是一样的(见习题 7-3-1)。只要在最后的结果中取实部或虚部即可。读者可以利用三角函数作类似的推导

$$m_O\frac{d^2x_1}{dt^2}=\kappa(x_2-x_1-a)$$

$$m_C\frac{d^2x_2}{dt^2}=-\kappa(x_2-x_1-a)+\kappa(x_3-x_2-a)$$

$$m_O\frac{d^2x_3}{dt^2}=-\kappa(x_3-x_2-a)$$

其中a为平衡时碳原子和氧原子之间的距离。记$\omega_O^2=\kappa/m_O$，$\omega_C^2=\kappa/m_C$。设系统的简正振动模式中，原子j的运动为$x_j=A_je^{i\omega t}$ $(j=1,2,3)$，则运动方程组将简化为代数方程组：

$$(-\omega^2+\omega_O^2)A_1-\omega_O^2A_2=0$$

$$-\omega_C^2A_1+(-\omega^2+2\omega_C^2)A_2-\omega_C^2A_3=0$$

$$-\omega_O^2A_2+(-\omega^2+\omega_O^2)A_3=0$$

该奇次方程组的系数行列式为零给出确定简正模式频率的特征方程：

$$\begin{vmatrix}\omega^2-\omega_O^2 & \omega_O^2 & 0\\ \omega_C^2 & \omega^2-2\omega_C^2 & \omega_C^2\\ 0 & \omega_O^2 & \omega^2-\omega_O^2\end{vmatrix}=\omega^2(\omega^2-\omega_O^2)[\omega^2-(\omega_O^2+2\omega_C^2)]=0$$

从而得到三个简正模式的频率：

$$\omega_1=0,\quad \omega_2=\omega_O,\quad \omega_3=\sqrt{\omega_O^2+2\omega_C^2}$$

代入原代数方程组可得到这三个特征频率所对应的模式分别为：

(1) $\omega_1=0$: $A_1=A_2=A_3$ (分子整体的平动)

(2) $\omega_2=\omega_O$: $A_2=0, A_1=-A_3$ (碳原子不动而两个氧原子作同步相对运动)

(3) $\omega_3=\sqrt{\omega_O^2+2\omega_C^2}$： $A_1=A_3=-A_2(\omega_O^2/2\omega_C^2)=-A_2(m_C/2m_O)$

读者可以按照上面的方法分析和计算 CO_2 分子的横向振动简正模式(见习题7-2-3)。◀

7-2-2 多粒子系统的耦合振动

以上的分析及结论可以推广到多个振子的耦合系统：含有N个自由度的系统就存在有N个简正模式，其中包含有反映质心运动的平动(频率为零的振动)模式，而当系统处于某一个反映内部运动的简正模式时，系统中所有的振子总是在相同时间经过各自的平衡位置，因而具有相同的振动频率。如果我们关注于求解相互独立的简正模式，而不是每一个耦合振子的运动状态，这在数学上对应于将涉及N个变量的N个耦合方程变换为N个独立的简谐振动方程。从而使问题的求解得到简化。

物体中振动的传播是由描述局部扰动的物理量的运动方程所决定的。为简单起见，我们考虑由 N 个同一种质点形成的一维(离散)介质系统的纵向振动(见图7-2-2)。设质点质量为 m，系统处

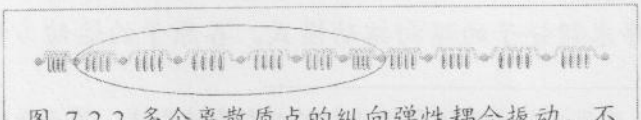


图 7-2-2 多个离散质点的纵向弹性耦合振动。不同简正模式的对称性表现为大小不同的周期性。图中红圈对应的是其中某一个模式振动形态的半个周期。

于力学平衡时相邻质点之间的间距为 a。并设质点之间只存在最近邻的简谐相互作用力，则作用在第 n 个质点上的力取决于它和两边近邻质点之间的间距[5]：

$$F_n = -\kappa(u_n - u_{n-1}) - \kappa(u_n - u_{n+1})$$

其中 u_n 为第 n 个质点偏离其平衡位置的距离，κ 为弹性恢复力的弹性系数，令 $\omega_o^2 = \kappa/m$，则第 n 个质点的运动方程为

$$\frac{d^2u_n}{dt^2} = -\omega_o^2(2u_n - u_{n-1} - u_{n+1}) \tag{7-2-9}$$

下面从系统存在的简正模式角度来分析求解该耦合系统的振动行为。设系统处于特征频率为 ω 的模式时各质点的运动形式为

$$u_n = \widetilde{A}_n e^{-i\omega t} \quad (n = 1, 2, ..., N) \tag{7-2-10}$$

其中 $\widetilde{A}_n = A_n e^{i\varphi_n}$ 称为第 n 个质点的复数振幅。

如果以试探解代入运动方程组，可以将其转化为 N 个齐次线性代数方程构成的方程组，则问题最终归结为求解一个关于特征频率 ω 的 N 阶代数方程的根。对于大数量的 N，这显然不是一个容易解析求解的问题。因此我们试图从振动形态的对称性来简化计算[6]：对处于简正模式振动状态的均匀质点链系统，两个相邻质点的振动应该具有相同大小的振幅，而且各相邻质点之间的振动位相差 $\Delta\varphi = \varphi_n - \varphi_{n-1}$ 也与质点位置(用质点序号 n 标志)无关，则各质点的振动(7-2-13)式可表示为

$$u_n = Ae^{in\Delta\varphi}e^{-i\omega t} \quad (n = 1, 2, ..., N) \tag{7-2-11}$$

将上述试探解代入运动方程(7-2-9)可以得到系统的特征方程

$$-\omega^2 + \omega_o^2(2 - e^{-i\Delta\varphi} - e^{i\Delta\varphi}) = 0$$

特征频率是相邻原子振动相位差的函数：

$$\omega = 2\omega_o \sin\frac{\Delta\varphi}{2} \tag{7-2-12}$$

为了理解相邻质点之间振动位相差 $\Delta\varphi$ 与振动频率 ω 间关系的含义，设想与第 n 个原子与其相距 m 个原子间距的第 $n + m$ 个原子在该模式中的振动行为相同(见图 7-2-2)：

$$u_{n+m} = u_n \tag{7-2-13}$$

则系统振动位型 $\{u_n\}$ 构成了一个离散形式的空间简谐函数，该振动模式的对称性表现为振动的空间周期性，不同的模式表现为不同的空间周期。相对于式(7-2-16)描述的模式，其空间周期为 $\lambda_m = ma$，称为**波长**。则相邻原子之间的振动位相差 $\Delta\varphi_m$ 为：

$$\Delta\varphi_m = \pm 2\pi(a/\lambda_m) = \pm k_m a \tag{7-2-14}$$

其中的参量

$$k_m = 2\pi/\lambda_m \tag{7-2-15}$$

表征空间简谐函数的(角)频率，称为空间(角)**波数**。则模式 u_n 的振动形状表示为：

$$u_n = Ae^{-i\omega_m t}\sin(k_n a + \varphi_o) \qquad (n = 1, 2, ..., N) \tag{7-2-16}$$

---

[5] 读者也许会注意到 $n = 1$ 和 $n = N$ 的情形应该和中间的原子有所不同，这种差异构成表面效应。但这里我们仅关注于体内的振动行为。

[6] 在实际求解同一类的数学问题中，因变量数量的悬殊差异，为了提高求解效率，通常都会通过不同的视角和采用不同的求解策略。

其中通过式(7-2-12)可得到简正模式的振动频率。

$$\omega_m = 2\omega_o \sin(\pi/m) \tag{7-2-17}$$

上述多质点链的耦合振动模型可用于描述由大量原子和分子构成的实际晶体内部的振动行为，因为一个粒子发生较小的相对位移时所受到的作用力主要体现出线性恢复力的特征，这时 N 的有效量度单位是阿伏加德罗常数。这种用大量相互作用的粒子描述介质行为的模型称为**离散介质模型**。

7-2-3 均匀介质中的耦合振动模式—驻波

对于一维均匀质点链中空间周期λ远大于物体中粒子之间的间距a的简正振动模式，由于相邻粒子的振动几乎是同相位的，在可观测到两个振动相位差异的粒子之间相隔有大数量的粒子，这时相邻粒子的空间分布细节不再重要，可以将近似同相振动的相邻的许多质点组合为一个质量元，用连续时空函数 $u(\mathbf{r}, t)$描述介质内部各质量元的运动状态以及由此所决定的介质运动特性。这种描述方法称为**连续介质模型**。

为简单起见，我们讨论将上一节所讨论一维原子链系统在长波模式在长波近似下如何转化为连续介质中的耦合振动问题。设第 n 个原子的振动 $u_n(t)$用相应位置(取其坐标为 x)的质量元的振动 $u(x, t)$描述，用 $u(x+\Delta x, t)$和 $u(x-\Delta x, t)$分别表示第 $n+1$ 和第 $n-1$ 个原子的振动 $u_{n+1}(t)$和 $u_{n-1}(t)$，则由第 n 个原子的运动方程(7-2-12)可得到相应质量元的运动方程

$$\frac{\partial^2 u(x,t)}{\partial t^2} = \frac{\kappa}{m}\{[u(x+\Delta x,t)-u(x,t)]-[u(x,t)-u(x-\Delta x,t)]\}$$

$$\approx \frac{\kappa}{m}\left\{\frac{\partial u(x+\Delta x,t)}{\partial x}-\frac{\partial u(x,t)}{\partial x}\right\}\Delta x \approx \frac{\kappa}{m}\frac{\partial^2 u(x,t)}{\partial x^2}(\Delta x)^2$$

在 $\Delta x \to 0$ 的极限情形下，记 $\mu = m/\Delta x$ 为弹性弦的质量密度，$T = \kappa \Delta x$ 为弹性弦中的张力，则质点链的波动方程组(7-2-12)变形为描述连续弦振动的运动方程：

$$\frac{\partial^2 u(x,t)}{\partial t^2} = v^2 \frac{\partial^2 u(x,t)}{\partial x^2} \tag{7-2-18}$$

方程(7-2-18)称为连续介质的**线性波动方程**，为二阶偏微分方程[7]。其中的参数

$$v = \sqrt{T/\mu} \tag{7-2-19}$$

为振动的**传播速度**，也就是通常所说的波速。

由于波动方程的本质仍然是一个耦合振动系统的运动方程，其通解可以表示为耦合振动系统中各种的简正振动模式的叠加。系统处于简正振动模式时，其中各微元都以相同的频率ω振动，将处于x位置的微元的振动表示为

$$u_\omega(x, t) = \zeta(x)e^{-i\omega t} \tag{7-2-20}$$

其中 $\zeta(x)$为 x 位置的复数振幅[8]。代入波动方程(7-2-19)得到

---

[7] 波动方程(7-2-19)虽然是针对弹性弦推导得出的，但许多线性波动的运动方程都表现为(7-2-19)的形式，所不同的仅仅是表征空间扰动的物理量 $u(x, t)$ (称为**波函数**)不同。空间各点的扰动可以是没有方向属性的标量(如密度、温度等)，由此形成的波场称为**标量波**。如果描述空间各点的扰动是具有方向属性的物理量(如电磁场、速度等)，相应的波场称为**矢量波**。

[8] 这里我们采用了简谐函数的指数表达形式，因而振幅也是复数(称为复振幅)。这样做是为

$$\frac{d^2\zeta}{dx^2}+\frac{\omega^2}{v^2}\zeta(x)=0 \tag{7-2-21}$$

这是一个圆频率为$k=\omega/v$的空间简谐振动方程，其复数型通解为

$$\zeta(x)=\tilde{A}e^{ikx}+\tilde{B}e^{-ikx} \tag{7-2-22}$$

其中的积分常数$\tilde{A}$和$\tilde{B}$代表复数振幅。由此得到波动方程(7-2-19)的简正模式解

$$u_\omega(x,t)=\tilde{A}e^{i(kx-\omega t)}+\tilde{B}e^{-i(kx+\omega t)} \tag{7-2-23}$$

确定简正模式的频率需要知道边界条件，即弦两端的振动情况。我们先以两端固定，长度为L的弦为例讨论系统波动方程的简正模式解。其它形式的边界条件可以做类似的讨论。

为方便起见，将简正模式的空间振动解(7-2-24)式表示为实数型通解形式

$$u_\omega(x,t)=e^{-i\omega t}(A\cos kx+B\sin kx) \tag{7-2-24}$$

系统左端的边界条件$u(x=0)=0$，对波动方程解的限制为

$$A=0$$

而右端边界条件$u(x=L)=0$则意味着：

$$B\sin kL=0 \tag{7-2-25}$$

由于系数B不能等于零(不然整个系统就没有运动)。因此上式实际给出的是对波数k的约束：

$$y_1=A\sin(kx-\omega t)\qquad(n=1,2,\ldots) \tag{7-2-26}$$

利用空间圆频率k(波数)和空间周期λ(波长)的关系$k=2\pi/\lambda$，上式可以表示为

$$L=\tfrac{1}{2}n\lambda_n\qquad(n=1,2,\ldots) \tag{7-2-27}$$

它对应于一个直观的几何图像：弦的长度必须为半波长的整数倍(见图7-2-3)。由此可以解出连续弦两端固定时振动的简正模式特征频率：

$$\omega_n=vk_n=n\frac{\pi}{L}\sqrt{T/\mu}\quad(n=1,2,\ldots) \tag{7-2-28}$$

其中用到了弦中波速v的表达式。对于一个特征频率为ω_n的振动模式，弦的运动位型

$$u_\omega(x,t)=Be^{-i\omega t}\sin nk_0x \tag{7-2-29}$$

介质中的所有质量元以相同的频率ω运动，并在相同的时刻回到平衡位置，各点的差异只是振幅不同，振幅的空间分布为$B\sin nk_0x$，这一特殊的振动位型看上去类似于驻立在原位不向四周传播的简谐波，因而相应的模式被称为**驻波**。简谐波的基本运动特征是具有时空双重周期性运动形式。弦上振幅最大或最小的位置分别称为驻波的**波腹**或**波节**。驻波模式可以用整条弦上的波节数目标记：除了弦的两个端点包含有$n-1$个波节的驻波称为模式n。

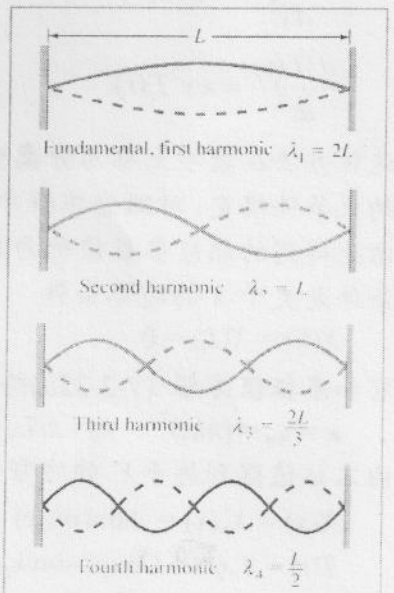


图 7-2-3 几种低频简正模式

了数学处理上更加简捷，方便，另外也是为后面量子物理中更抽象波的描述做一些预演。读者可以用更直观的正弦或余弦函数作同样的推导，并先得到的结果是一样的(见习题 7-80)。

7-2-4 波动方程的本征值问题

以上讨论揭示出了求解线性波动方程的一般数学定式，它可推广用于求解更复杂介质(例如非均匀弹性弦，以及二维或三维弹性介质)中的波动问题。为了使读者能更清楚地理解这一方法，我们用数学语言重新表述弹性弦上的简正振动模式问题的求解过程。

均匀弦振动的本征值问题

前面求解偏微分方程(7-2-19)过程的第一步是设简正模式可表述为 $u_\omega(x, t) = X(x)e^{-i\omega t}$，将试探解取为代入波动方程使之转换为常微分方程。这在数学上意味着将一个多变量函数解分解为两个单变量函数的乘积的特解，我们将其更一般地表示为：

$$u(x, t) = X(x)T(t) \tag{7-2-30}$$

将试探解[9]代入线性齐次偏微分方程(7-2-19)可得到

$$\frac{d^2X(x)}{dx^2}T(t) = \frac{X(x)}{v^2}\frac{d^2T(t)}{dt^2}$$

将与x有关的函数因式和与t有关的函数因式分别改写到等式两边：

$$\frac{1}{X(x)}\frac{d^2X(x)}{dx^2} = \frac{1}{v^2T(t)}\frac{d^2T(t)}{dt^2} \tag{7-2-31}$$

等式两边的这两个不同变量的函数恒等显然意味着它们只能等于同一个常数 κ。这样，偏微分方程(7-2-31)被分解为两个常微分方程

$$\frac{d^2X(x)}{dx^2} = \kappa X(x) \tag{7-2-32a}$$

$$\frac{d^2T(t)}{dt^2} = \kappa v^2T(t) \tag{7-2-32b}$$

这种方法在数学上称为**分离变量法**。其中参数 κ 需要通过实际问题中 X 或 T 的约束条件确定。对微分方程中某些参数数值的各种限制条件通常称为**本征值条件**，由此得到的相应参数值称为**本征值**。例如，对于两端固定的弦振动问题，本征值条件是关于 X 的边界条件

$$X(0) = X(L) = 0 \tag{7-2-33}$$

它和**本征值方程** (7-2-32a)给出本征值的解：

$$\kappa = \kappa_n = (nk_o)^2 \qquad k_o = \pi/L,\ \ n = 1, 2, \ldots$$

由本征值得到关于X 的方程和T 的方程解

$$X(x) = X_n(x) = A\sin(nk_ox)$$

$$T(t) = T_n(t) = C\exp(-inv k_o t) = C\exp(-in\omega_o t)$$

描述本征模式的函数称为本征函数，本征函数系列$\{X_n(x)\}$具有规律：由于 nk_o 描述的是空间周期函数 $X_n(x)$的频率，即除了弦的两个端点，$X_n(x)$含有 $n-1$ 个波节。因而不同本征值的函数可以用函数所含有的波节数目标记。

波动方程的通解可表示为所有可能的简谐驻波解的线性叠加

[9] 能够采用分离变量的方法求解偏微分方程的条件是该偏微分方程是线性和齐次的。

$$u(x,t)=\sum_n A_n e^{-i\omega_n t}X_n(x) \tag{7-2-34}$$

其中的驻波频率为 $\omega_n=nvk_o\ (n=1,2,\ldots)$。

例7-2-2 求解一个半径为R的圆环上的驻波模式。

解： 半径为R的圆环上的波动方程为(习题7-2-10)

$$\frac{\partial^2 u}{\partial t^2}=\frac{v^2}{R^2}\frac{\partial^2 u}{\partial \theta^2}$$

按照上面所用的简正模式求解过程，设方程的试探解(简正模式)为

$$u_\omega(\theta,t)=\Theta(\theta)e^{-i\omega t}$$

代入运动方程得

$$\frac{d^2\Theta}{d\theta^2}+k^2R^2\Theta(\theta)=0$$

其中 $k=\omega/v$。上述方程的通解为 $\Theta(\theta)=e^{\pm ikR\theta}$。这里虽然没有类似于两端固定的弦边界条件来确定本征值 k。但物理上的圆周问题具有特定的周期条件

$$\Theta(\theta+2\pi)=\Theta(\theta)$$

该本征值条件要求本征值 k 满足关系式：$\exp(\pm i2\pi kR)=1$，即

$$k=n\frac{2\pi}{R}\equiv k_n,\quad n=1,2,\ldots$$

由此可求得各驻波模式相应的特征频率：$\omega_n=vk_n=2\pi nv/R\ (n=1,2,\ldots)$。◀

非均匀弦的驻波模式—非简谐驻波

我们将上述方法推广到求解更一般的非均匀介质中的本征值问题。为简单起见，我们具体考虑一根长度为L，两端固定的非均匀弹性弦的沿一个确定方向的横振动问题。设弦的密度分布为$\mu(x)$，由线元Δx 在弹性力T 的作用下的运动方程可得到弦的横向运动方程：

$$\frac{\partial^2 u}{\partial x^2}=v^2\frac{\partial^2 u}{\partial t^2} \tag{7-2-35}$$

其中v是位置的函数$v(x)$，表示弹性波在位置 x 处的转播速度。虽然介质的非均匀性导致偏微分方程变得更复杂，由于仍然存在着弹性弦上各个线元以相同的频率ω作同步振动的本征模式，还是可以将整个介质的振动分布函数$u(x,t)$的时间和空间变量部分分离：

$$u(x,t)=X(x)e^{-i\omega t} \tag{7-2-36}$$

其中弦上的振幅分布函数 $X(x)$就是我们要求解的驻波模式。将试探解代入波动方程(7-2-35)，消去时间因子 $e^{-i\omega t}$ 后得到线性齐次常微分方程

$$\frac{d^2X(x)}{dx^2}+k^2(x)X(x)=0 \tag{7-2-37}$$

其中 $k(x)=\omega/v(x)$，由于 $v(x)$是确定的，弦的边界条件实际上限制了驻波模式频率 ω 的取值，即构成了所谓**本征值条件**。例如对于两端固定的弦，由

$$X(0)=X(L)=0 \tag{7-2-38}$$

可以确定所有允许的本征频率$\{\omega_n\}$，并由此解出与每一个 ω_n 对应的模式函数 X_ω

称为属于本征值 ω_n 的本征函数。

显然，求解非均匀弦的振动问题最终归结为求解方程(7-2-37)和相应边界条件构成的本征值问题。对各种非均匀的质量分布函数 $\mu(x)$ 而言，实际求解上述本征值并不容易，通常都不能得到解析解。如果能从上述方程中得到本征函数族 $\{X_n(x)\}$ 的主要几何特性，不仅有助于理解各种驻波模式的普适特性，也能对求解本征值问题的数值程序流程和方法提供指导。

对有限长度弦上的驻波模式而言，其主要几何特征有三个方面：波节的数目(本征函数的零点)，相邻波节的距离和各个波腹的相对大小。

我们首先就驻波模式的上述三个特征回顾均匀弦的各模式本征函数 $X_n(x)$ 的振动形态。(1) 描述均匀弦上的各驻波模式的本征函数与波节数目都有一一对应的关系：给定的波节数目都对应于一个唯一的本征函数，(2) 由于这些本征函数都是简谐函数，是确定的空间周期函数，这意味着振荡函数的节点分布是均匀的，(3) 各周期中位置的振幅和形状也完全相同，由此可知，其本征值(频率)谱是均匀分布的($k_n = nk_o$，$\omega_n = n\omega_o$)。

对于非均匀弦而言，由于各本征函数不再是空间周期函数，相应的驻波节点的分布是不均匀的，各相邻节点之间的间距，波形，相对波幅也都不尽相同。这也就意味着本征值(频率)谱的分布是不均匀的。为了能大致确定非均匀弦振动模式的上述几何特征，我们可以将任何一条不均匀弦都设想为是由一条均匀弦上的质量分布逐渐连续变形而成的。驻波形状也相应地发生着连续改变[10]。因此弦的质量分布的无限小的变化虽然导致各个驻波模式偏离了简谐形式，但按本征频率谱 $\{\omega_n\}$ 由低到高的分布形成本征函数族 $\{X_n(x)\}$ 对应的各模式振荡函数特征得以保持下来，每一个本征函数与波节数目仍然存在有一一对应的关系，对于给定的模式，本征函数的驻波波节数目不会随着弦的质量分布而改变(在数学中称为拓扑不变性)。例如对于两端固定的各种弹性弦，频率最低的驻波模式 $X_1(x)$ 中都没有波节。

其它驻波模式的频率是随着驻波波节数目的递增单调升高的(均匀弦中两者有简单的正比关系)，这是因为均匀弦上的简谐驻波具有周期性，可以在整根弦上定义空间频率 k，而非均匀弦上的驻波没有周期性，因而只能在各局部区域定义有效空间频率 k_{eff}，而在整根弦上空间频率平均值与均匀弦中的空间频率 k 相等，给出相同的驻波波节数。

而对于驻波模式中的另外两个特征：相邻波节的距离和各个波腹的相对大小可以从本征值方程(7-2-37)中的 $k(x)$ 的意义得到理解。设想将弹性弦所处空间分解为许多很小的区域，每一个区域 $x_i \to x_i + dx$ 中的 $k(x_i)$ 可近似理解为一个常数 k_i，则该区域中方程的解是一个简谐函数，而 k_i 则是该简谐函数的圆频率。显然，k_i 较小的区域中的波形变化比较平缓(波长比较大)，波形得以展开；而 k_i 比较大的区域中的波形相对比较陡峭，波形挤在一起(波长比较小)。因此 k_i 比较小的区域中节点间距比较大，而 k_i 比较大的区域中节点间距比较小。

对于确定的驻波模式，$k_i^2 = \omega^2/v_i^2 = \mu_i \cdot \omega^2/T$ 中的 ω 及 T 都是和位置无关的常数，因此，k_i^2 正比于相应空间区域中的质量密度。k^2 小的区域，质量密度小，因而

[10] 弦的质量分布出现了无限小变化，相应驻波模式的实际变形也只是无限小的。

在物理上对应于较大的振幅；而 k^2 大的区域，质量密度大，对应于较小的振动幅度。

以上的分析表明，对于给定波节数的驻波模式，非均匀弦与均匀弦的振动形态在总体上是一致的(拓扑不变)，但在形态上发生了扭曲：$k(x)$比较小的区域中节点间距比较大，振幅也比较大，而 $k(x)$比较大的区域中节点间距比较小，振幅也比较小。因此也可以将这种非周期的振荡函数理解为是定义在非均匀空间(弦)上的一种谐函数。

驻波本征值问题的数值方法*

对于(7-2-37)形式的方程，求解本征值问题主要还是依赖于数值方法。下面根据各模式对应的本征函数的主要几何特征，通过一种具体算法讨论数值求解中的要点。

二阶线性齐次微分方程可表示为

$$\frac{d^2u}{dx^2} = f(x)u(x) \tag{7-2-39}$$

其中 $f(x)$的函数形式是已知的。将待求解的函数 $u(x)$定义的相关区域均匀分为许多小区间，相应结点的位置记为 $x_1, x_2, \ldots, x_n, \ldots$，将 $u(x)$在 $x = x_n$ 前后两个结点的邻域中作泰勒展开：

$$\begin{aligned} u_{n+1} &= u(x_n + h) \\ &= u(x_n) + hu'(x_n) + \frac{h^2}{2!}u''(x_n) + \frac{h^3}{3!}u^{(3)}(x_n) + \frac{h^4}{4!}u^{(4)}(x_n) + \frac{h^5}{5!}u^{(5)}(x_n) + O(h^6) \\ u_{n-1} &= u(x_n - h) \\ &= u(x_n) - hu'(x_n) + \frac{h^2}{2!}u''(x_n) - \frac{h^3}{3!}u^{(3)}(x_n) + \frac{h^4}{4!}u^{(4)}(x_n) - \frac{h^5}{5!}u^{(5)}(x_n) + O(h^6) \end{aligned}$$

其中 h 为两个相邻结点间的距离。上述两个表达式之和为

$$u_{n+1} + u_{n-1} = 2u_n + h^2u_n'' + \frac{h^4}{12}u_n^{(4)} + O(h^6) \tag{7-2-40}$$

将方程(7-2-39)及其二阶微商

$$u_n'' = f_nu_n \quad \text{和} \quad u_n^{(4)} = f_n''u_n + 2f_n'u_n' + f_nu_n''$$

代入上述方程。忽略 h^6 以上的项后整理得到 u_{n+1} 的表达式

$$u_{n+1} = \frac{(2 - 5/6h^2f_n)u_n - (1 + h^2/12f_{n-1})u_{n-1}}{1 + h^2/12f_{n+1}} \tag{7-2-41}$$

这样，只要知道函数 $u(x)$在两个相邻结点 x_{n-1} 和 x_n 处的函数值 u_n 和 u_{n+1}，就能依次递推出后续所有的 $u(x)$。式(10-4-25)称为 Numerov 递推公式。下面我们以两端固定的非均匀弦为例，利用该递推公式说明驻波模式本征值问题的数值求解过程。

事实上，本征值方程(7-2-37)与微分方程(7-2-39)是不同的，方程(7-2-39)中 $f(x)$的函数是完全确定的，而方程(7-2-37)中的 $k^2(x)$中含有尚未确定的本征值 ω。由于本征函数隶属于本征值，利用(7-2-41)式求解方程(7-2-37)必须先设定本征值。

一般的求解总是从最低频率的模式开始，可以先设定一个试探本征值 $\omega > 0$ 代入方程。利用弦的左端边界条件 $u(x_o = 0) = 0$，再设置一个试探值 $u(x_1 = h) = b >$

0)[11]，就能应用递推公式(7-2-41)进行叠代求解了。由于 $u''(x_1) = -k^2(x_1)u(x_1) < 0$，曲线在 x_1 附近是上凸的，由于最低频率的驻波模式在(0, L)的范围内都没有节点，这种上凸特性就必须一直延续到弦的右端 $x = L$。这就对 ω 的选取产生了限制。如果 ω 取得太小，u 在到达 $x = L$ 处就仍有 $u(x = L) > 0$，而 ω 取得太大，u 在尚未到达 $x = L$ 处就会出现节点 $u(x < L) \le 0$。只有保证整个区域内都有 $u > 0$，并有 $u(x = L) = 0$ 时的 ω 才对应于系统的最低驻波频率。

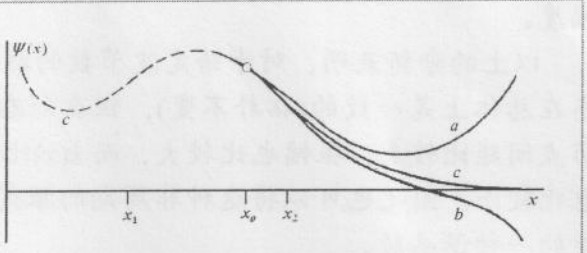


图 7-2-6 由一个试探本征值 ω 的三种初始变化率递推出的本征函数. 其中曲线 a 和 b 都不符合最低频率本征函数的要求。

然后可以同样的步骤求解方程的第二，第三，...，第 n 个模式的本征值问题，对于第 n 个模式的本征值问题，所设置的试探本征值 ω 需要调整到在 $0 < x < L$ 区域的振荡出现 $n-1$ 个波节，并要求 $u(L) = 0$。

其它形式的驻波

上述求解波动的本征值问题的方法还可以推广到二维和三维系统。为简单起见，我们具体讨论一种非常典型的圆鼓面的横向振动驻波模式。描述鼓面的纯横向振动的波动方程为

$$\frac{\partial^2 u}{\partial t^2} = v^2 \nabla^2 u \tag{7-2-42}$$

设其振动特征频率为ω的驻波模式为$u(\mathbf{r}, t) = e^{-i\omega t} f_\omega(\mathbf{r})$，代入上述波动方程可得

$$k^2 f_\omega(\mathbf{r}) = -\nabla^2 f_\omega(\mathbf{r}) \tag{7-2-43}$$

其中的参数$k = \omega/v$。

例7-2-3 求解鼓面振动的模式。

解：鼓面是一个二维圆对称问题，在极坐标系中最为方便，振动方程具体表示为

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 f}{\partial \varphi^2} + k^2 f = 0$$

这是一个线性齐次偏微分方程，我们同样可以采用分离变量法将其转化为本征值问题[12]。设 $f(r,\theta) = R(r)\Phi(\varphi)$，代入上面的方程(7-6-37)可得

$$\frac{r}{R}\frac{d}{dr}\left(r\frac{dR}{dr}\right) + k^2 r^2 = -\frac{1}{\Phi}\frac{d^2\Phi}{d\varphi^2}$$

令方程两边关于r和θ的函数等于同一个常数λ(本征值)，由此将上述方程分解为两个线性齐次常微分方程

---

[11] 由于齐次方程，函数 $u(x)$相差一个任意常数因子，初始值 b 的正负和大小都不会影响解的函数形式。

[12] 虽然这里不再有 r 和 θ 以某种特殊"频率"同步振动的物理图像，但在数学结构上是完全等价的。

$$\frac{d^2\Phi}{d\varphi^2}+\lambda\Phi=0$$

$$r^2\frac{d^2R}{dr^2}+r\frac{dR}{dr}+(k^2r^2-\lambda)R=0$$

确定其中本征值λ条件为周期条件：$\Phi(\varphi+2\pi)=\Phi(\varphi)$，由此得到的本征值问题的解为

本征值：$\lambda=m^2, m=0,\pm1,\pm2,\ldots$

相应的本征函数为：$\Phi_m(\varphi)=e^{im\varphi}$

将上述本征值代入(7-2-42)，并做变量代换：$x=kr$，方程(7-2-42)和边界条件所构成的径向本征值问题可表示为

$$\frac{d^2R}{dx^2}+\frac{1}{x}\frac{dR}{dx}+\left(1-\frac{m^2}{x^2}\right)R=0$$

$$R_m(ka)=0$$

下面我们按上述方法来分析该方程的驻波界的函数形态。首先讨论$m=0$的方程

$$\frac{d^2R_o}{dx^2}+\frac{1}{x}\frac{dR_o}{dx}+R_o=0$$

这在形式上类似于一个阻尼振荡方程，其中固有频率为1，局部阻尼系数为x^{-1}。从半定量的角度来理解，阻尼振荡的特征是振幅逐次衰减，振荡频率随阻尼系数的增加而减小。

对于确定各本征值$\{k_n\}$的式(7-2-46)而言，在于相继减小k值，使R_o的节点数目在$[0, a)$的范围内从0开始持续增长。

对于$m\neq0$的方程可进行类似的分析，其中阻尼系数是x^{-1}，固有频率为$(1-m^2/x^2)$。差异在于对于$m\neq0$的方程，原点处的渐进行为是由

$$\frac{d^2R}{dx^2}-\frac{m^2}{x^2}R=0$$

决定的。这意味着：$R_m(0)=0$

方程(7-2-45)在数学上称为**m阶贝塞尔方程**，其解称为**m阶贝塞尔函数**，记为$J_m(x)$[13]。图7-2-4为$m=1, 2, 3, 4, 5$的贝塞尔函数曲线，它与上面分析所得出的二维各向同性介质中的驻波形式是一致的。

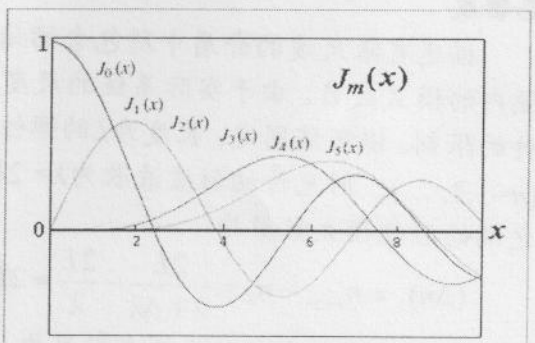


图7-2-4

综合鼓面振动模式的本征值问题，其固定的鼓边缘的边界条件：$R(a)=0$构成确定本征值k的条件：

$$J_m(x_n^{(m)})=J_m(k_na)=0$$

相应的振动模式为

$$f_\omega(r,\varphi)=J_{m,n}(k_nr/x_n^{(m)})e^{im\varphi}$$

---

[13] Bessel function describes harmonic behavior in the case of cylindrical symmetry as the waves produced by throwing a pebble into a pond of water. Bessel functions $J_n(x)$ characteristically begin with a strong peak and then oscillate like a damped harmonic wave as x increases, and therefore are sometimes called cylindrical harmonics, as shown in Figuration 8-29(a).

其中 $x_n^{(m)}$ 为 $J_m(x)$ 的第 n 个零点。由此可确定本征值谱 $\{k_n\}$ 和相应的本征频率 $\{\omega_n\}$。对给定的 m，k_n 按照径向波节数目分布。图7-2-5(*a*), (*c*), (*e*)为三个最低频率的中心对称($m = 0$)驻波模式(0, 0), (1, 0), (2, 0)。(*b*), (*d*), (*f*)则在角度方向有一个节点的驻波模式(0, 1), (1, 1), (2, 1)。

图7-2-5 三个角度方向的驻波模式分别为(0, 0), (0, 1), (1, 0), (1, 1), (2, 0), (2, 1)

需要注意的是虽然驻波模式的频率随波节数的增加而升高，但径向和角度方向是独立的，我们并不能从两种节点数目的总量来判断频率的高低。◀

驻波的形态取决于两类因素，介质的力学特性分布(质量和作用力参数)和边界条件。图 7-2-6 所示为小提琴后板的 6 个低频振动模式的全息记录图。通过各振动图案可以感受到系统振动的各离散频率的间距排列是很复杂的，但频率和波节数目之间单调的正向关联还是很明显的。

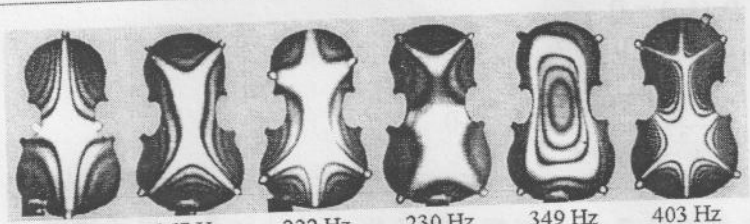


图7-2-6

态密度

描述有限尺度的介质中所包含的稳态驻波模式有一个重要参量：单位频率间隔内的模式数目。由于实际系统的尺度有限，这意味着系统的振动应受到边界条件的限制。以两端固定，长度为 L 的弹性弦为例，驻波的本征值条件为 $L = n(\lambda/2)$，(n=1,2, …)，所允许的驻波波长为 $\lambda = 2L/n$ (n = 1, 2, …). 在波长间隔为 $\lambda \to \lambda+\Delta\lambda$ 之间的驻波模式数目为

$$(\Delta n)_\lambda = n_{\lambda+\Delta\lambda} - n_\lambda = \frac{2L}{\lambda+\Delta\lambda} - \frac{2L}{\lambda} \approx 2L\frac{\Delta\lambda}{\lambda^2}$$

即单位波长间隔内的驻波模式数目为

$$N(\lambda) = \frac{dn}{d\lambda} = \frac{2L}{\lambda^2} \tag{7-2-44}$$

由于模式数目和波长呈负关联，在许多场合中，用频率来讨论状态密度比用波长更为方便。驻波模式的频率可表示为

$$\Delta n = (n+\Delta n) - n = 2L[1/(\lambda+\Delta\lambda) - 1/\lambda] \approx (2L/\lambda^2)\Delta\lambda \tag{7-2-45}$$

其中 c 是介质中波的传播速度。记频率区间 $\nu \to \nu + d\nu$ 内的模式的数目 $N(\nu)d\nu$，其中 $N(\nu)$ 是单位频率间隔内的驻波模式的数目：

$$N(\lambda)d\lambda = (8\pi V/\lambda^4)d\lambda \tag{7-2-32}$$

是一种均匀分布(见图7-2-7)。

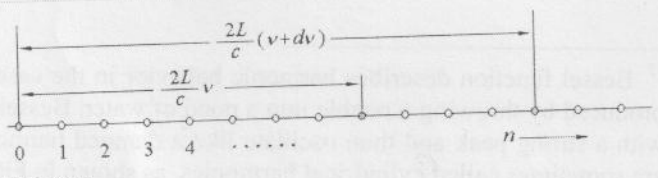


图7-2-7

上述结果的一个重要应用是控制激光器中的模式数目。设激光器谐振腔中的两个平行的平面反射镜间距为L，反射镜横向尺度很小，谐振腔内形成驻波时，相邻的两个模的频率间隔满足

$$\Delta\nu_n=\nu_{n+1}-\nu_n=\frac{c}{2L}$$

其中c为光在谐振腔介质中的传播速度。对L = 20cm的Hi-Ni激光器而言，$\Delta\nu\approx$ 1700MHz，则在上式中的$\Delta\nu$= 750MHz，在线宽内只有一个模。

例7-2-4 计算一尺度为$L_x\times L_y\times L_z$的均质弹性固体在边界不受约束时所具有的驻波模式密度。设弹性固体的力学特性是各向同性的。

解： 由于固体边界上的质点不受约束，而且三个方向的振动互相独立，其驻波模式具有形式：

$$u_\omega(x,y,z;t)=Ae^{-i\omega t}\cos k_x x\cos k_y y\cos k_z z$$

其中$k_x=(\pi/L_x)n_x$，$k_y=(\pi/L_y)n_y$，$k_z=(\pi/L_z)n_z$ (n_x, n_y, n_z为正整数)满足

$$k_x^2+k_y^2+k_z^2=k^2=\left(\frac{2\pi}{\lambda}\right)^2=4\pi\frac{\nu^2}{u^2}$$

其中u为波在介质中的传播速度，将上式改写为

$$\frac{n_x^2}{r_x^2}+\frac{n_y^2}{r_y^2}+\frac{n_z^2}{r_z^2}=1 \qquad \text{其中}\ r_x=\frac{2L_x}{u}\nu,\ r_y=\frac{2L_x}{u}\nu,\ r_z=\frac{2L_x}{u}\nu$$

这是以n_x, n_y, n_z为坐标轴的标准椭球方程，三个半主轴的长度分别为r_x, r_y, r_z。由于描述驻波模式的参量n_x, n_y, n_z都是正定的，模式频率上限为ω的驻波模式数目为椭球中所包含的整数格点总数的1/8，

$$n(\nu)=\frac{1}{8}\times\frac{4}{3}\pi r_x r_y r_z=\frac{4\pi}{3u^3}L_xL_yL_z\nu^3=\frac{4\pi V}{3u^3}\nu^3$$

其中$V=L_xL_yL_z$为固体体积。考虑到沿每个频率有两个横波模式和一个纵波模式，设它们的传播速度分别为u_t和u_h，则

$$n(\nu)=\frac{4\pi V}{3}\left(\frac{2}{u_t^3}+\frac{1}{u_h^3}\right)\nu^3$$

由此可以得到其中频率间隔为$\nu\to\nu+d\nu$内的驻波模式的数目

$$N(\nu)d\nu=\frac{dn}{d\nu}d\nu=4\pi V\left(\frac{2}{u_t^3}+\frac{1}{u_h^3}\right)\nu^2 d\nu$$

由此可得到驻波模式的状态密度$N(\nu)$。◀

7-3 振动形态的简谐分解—傅立叶定理

7-1 节在讨论简谐振动的意义时提到，各种形态的周期振动都可以分解为多个不同频率的简谐振动的叠加，而 7-2 节也说明了一根弦的任意曲线位型也可以表示为多个不同频率的简谐驻波形态的叠加。这一事实说明，可以利用简谐函数及其叠加来分析一般振动的行为特性。这种方法在数学中是建立在傅里叶定理[14]

[14] Baron Jean-Baptiste-Joseph Fourier (1768-1830) was a French mathematician and administrator who also

基础上的。本节我们按周期振动和非周期振动两种类型分别讨论振动形态的简谐分解，并简单说明振动按其它形式振动的分解，即所谓的广义傅立叶分解。最后简单讨论能够在实验测量上实现这种分解的器件及其工作原理。

7-3-1 周期运动的简谐振动分解—傅立叶定理

对于两端固定的弦而言，其一般的振动位型是该系统的所有模式位型的线性叠加：

$$u(x)=\sum_{n=1} a_n \sin\frac{n\pi}{L}x \tag{7-3-1}$$

由于物理上对两端固定的弦的位型除了有界、连续以外并没有其它限制，这就意味着定义在弦所在空间[0, L]上的任意有界、连续的函数都可以表示为空间频率为某一基频 k_o 的整数倍($k_n=nk_o$)的简谐函数的叠加。

需要注意的是，弦的位型函数(7-3-1)只是定义在弦所在的空间[0, L]，而函数(7-3-1)所表示的是一个$(-\infty, +\infty)$的周期函数(见图 7-3-1)，其位型函数为

$$u(x)=u_o(x-mL) \quad (mL<x<(m+1)L) \tag{7-3-2}$$

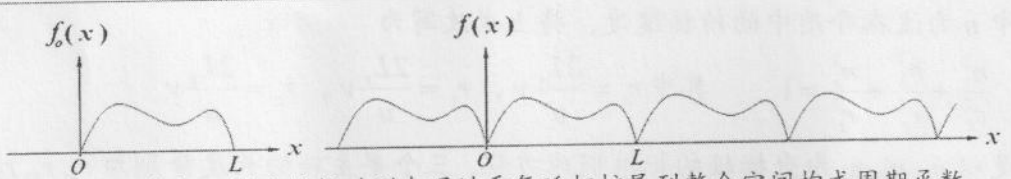


图 7-3-1 长度为 L 的弦的振动形态通过重复延拓扩展到整个空间构成周期函数。

这说明，一个周期函数总是可以表示为一系列频率为$\{k_n=nk_o\}$的简谐函数的叠加。这个表述在数学上称为**傅立叶定理**：

任意一个周期函数 $u(x)$都能够表示为一个频率系列的简谐函数的线性叠加：

$$u(x)=\sum_{n=1} a_n \sin k_n x+\sum_{n=0} b_n \cos k_n x \qquad k_n=nk_o=n\frac{2\pi}{\lambda} \tag{7-3-3}$$

也可以展开为复数形式的级数

$$u(x)=\sum_{n=-\infty}^{+\infty} c_n e^{ik_n x} \quad k_n=nk_o \tag{7-3-4}$$

其中 k_o 称为周期函数的**基频**，而 $k_n=nk_o$ 称为周期函数基频的**谐频**。展开式中的系数 a_n 和 b_n，或 c_n 称为**傅立叶系数**，代表相应傅立叶分量的强度，通常称为函数 $u(x)$的傅立叶频谱。各傅立叶分量之间满足

$$\int_0^{\lambda} \cos mk_o x \cos nk_o x dx=\frac{\lambda_o}{2}\delta_{m,n}$$

$$\int_0^{\lambda} \sin mk_o x \sin nk_o x dx=\frac{\lambda_o}{2}\delta_{m,n} \qquad (m, n=0, 1, 2, \ldots) \tag{7-3-5}$$

$$\int_0^{\lambda} \cos mk_o x \sin nk_o x dx=0$$

---

accompanied Napoleon on his expedition to Egypt. Fourier is best known for his contribution to the analysis of periodic functions and heat conduction in metals. Fourier's problem was, strictly speaking, an engineering problem. Modern science owes a great deal to the nineteenth-century engineers who studied the problems associated with thermal devices.

$$\int_0^{\lambda} e^{imk_o x} e^{-ink_o x} dx = \lambda_o \delta_{m,n} \qquad (m, n = 0, 1, 2, \ldots) \qquad (7\text{-}3\text{-}6)$$

(7-3-5)式和(7-3-6)式在形式上都表现为两个相互正交的两个无穷维数空间的矢量的点乘。由此可以将周期函数 $u(x)$理解为一个抽象矢量，而将简谐函数集$\{\sin nk_o x, \cos nk_o x, (n = 1, 2 \ldots)\}$或$\{e^{inkox}, (n = 1, 2 \ldots)\}$理解为各自构成该矢量所属矢量空间的一组完备的正交基矢量，这样，式(7-3-5)和(7-3-6)所表示的是在给定的一组基矢量$\{\sin nk_o x, \cos nk_o x, (n = 1, 2 \ldots)\}$或$\{e^{inkox}, (n = 1, 2 \ldots)\}$中任意两个基矢量是相互正交的。而 $u(x)$的傅立叶级数表达式(7-3-3)和(7-3-4)可以理解为矢量 $u(x)$在这组正交基上的分解，其中各简谐分量的傅立叶系数 a_n和 b_n，或 c_n分别是 $u(x)$在基$\{\sin nk_o x, \cos nk_o x, (n = 1, 2 \ldots)\}$或$\{e^{inkox}, (n = 1, 2 \ldots)\}$上的投影。由此，我们可直接通过矢量 $u(x)$和基矢量(函数)的点积求得傅立叶系数 a_n和 b_n，或 c_n，例如

$$\int_0^{\lambda} u(x) \sin nk_o x dx = \sum_{m=0}^{\infty} a_m \int_0^{\lambda} \sin mk_o x \sin nk_o x dx + \sum_{m=1}^{\infty} b_k \int_0^{\lambda} \sin mk_o x \cos nk_o x dx$$

其中 $\lambda = 2\pi/k_o$ 为函数 $u(x)$的最小周期。由式(7-3-5)给出的正交关系可知，上述所有积分项中只有被积函数为 $\sin^2 nk_o t$ 的积分不为零，因而有

$$\int_0^{\lambda} u(x) \sin nk_o x dx = a_n \int_0^{\lambda} \sin^2 nk_o x dx = \frac{a_n}{2} \int_0^{\lambda} \left(1 - \cos \frac{4n\pi}{\lambda} x\right) dx = \frac{a_n}{2} \lambda$$

由此可解得

$$a_n = \frac{2}{\lambda} \int_0^{\lambda} u(x) \sin nk_o x dx \qquad (7\text{-}3\text{-}7a)$$

读者可以用同样的方法得到系数 b_o和 b_n 以及 c_n(见习题 7-3-12):

$$b_o = \frac{1}{\lambda} \int_0^{\lambda} u(x) dx$$

$$b_n = \frac{2}{\lambda} \int_0^{\lambda} u(x) \cos nk_o x dx \qquad (7\text{-}3\text{-}7b)$$

另外，同样可以得到

$$c_n = \frac{1}{\lambda} \int_0^{\lambda} u(x) e^{-ink_o x} dx \qquad (7\text{-}3\text{-}8)$$

例7-3-1 求解**周期方波函数的傅立叶级数展开**[15]

解： 方波在电子学中常被用于研究各种电子线路的响应特性。为了简化计算，我们将方波函数的周期标度为2π，则傅立叶展开式中的基函数为$(\sin nx, \cos nx)$，适当选择坐标系原点，使方波函数呈现为奇函数(见例图a):

[15] It is traditional to use the square wave as the first example to show students how a complex wave form can be constructed from harmonic sine waves. This is a bit ironic, because the ideal square wave has discontinuous jumps from one level to another, and therefore does not satisfy Fourier's theorem that any **continuous wave shape** can be made from harmonic sine waves. The result is that if you try to construct an ideal square wave from sine waves, you end up with a small blip at the discontinuity (called the Gibb's effect). Since our focus is experimental data where there is no true discontinuity, we will not encounter this problem.

$$u(x)=\begin{cases}-1 & (2m-1)\pi<x<2m\pi\\ 1 & 2m\pi<x<(2m+1)\pi\end{cases}$$

其中m为整数。由于奇函数的傅立叶展开式只包含正弦分量

$$u(x)=\sum_{n=1}^{\infty}a_n\sin nx$$

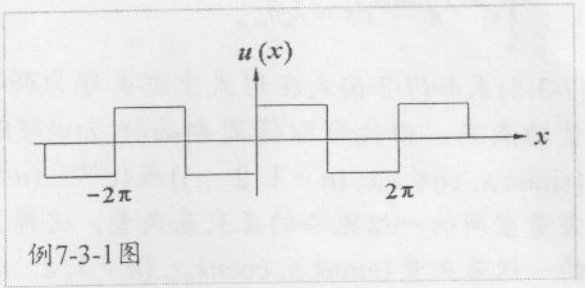
傅立叶系数a_n通过在周期$(-\pi,\ \pi)$中的积分得到:

$$a_n=\frac{1}{\pi}\left\{-\int_{-\pi}^{0}\sin nxdx+\int_{0}^{\pi}\sin nxdx\right\}=\frac{2}{\pi}\int_{0}^{\pi}\sin nxdx=\frac{2}{n\pi}[-(-1)^n+1]$$

其中所有n为偶数的项为零，方波函数的傅立叶分解为:

$$u(x)=\frac{4}{\pi}\sum_{n=0}\frac{1}{2n+1}\sin(2n+1)x$$

例图中给出了前 10 个分量的叠加结果。显然，波形的形貌主要式由低频项给出的，而高频项则给出图形的细节修正，如图形的变化锐度。随着高频分量的增加，得到的图形持续地逼近方波的形状。◀



例7-3-1图

例7-3-2 脉冲型周期方波的傅立叶级数展开，与例7-3-2中方波函数比较展开式的收敛速度。

解: 将函数的周期标度为2π，并选择坐标系原点使脉冲函数呈现为偶函数。为了说明展开式的收敛速度，选脉冲宽度为例7-3-6中的方波脉冲的一半(见例图):

$$u(x)=\begin{cases}-1 & (2m-3/2)\pi<x<2m\pi\\ 1 & 2m\pi<x<(2m+1/2)\pi\end{cases}$$

其中m为任意整数。偶函数的傅立叶展开式只包含余弦分量

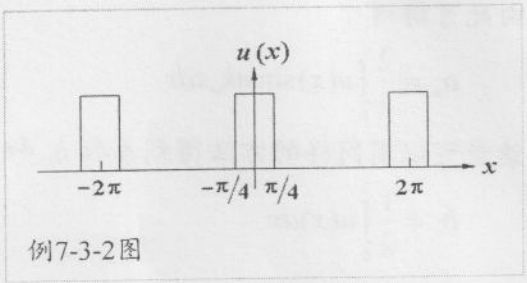
$$u(x)=\sum_{n=1}^{\infty}b_n\cos nx$$

通过下列积分可求得其中的傅立叶系数b_n

$$b_n=\frac{2}{\pi}\int_{0}^{\pi/4}\cos nxdx=\frac{2}{n\pi}\sin n\frac{\pi}{4}$$

例图中给出了前20个分量的叠加结果。比较例7-3-5中的叠加图形可以看到，获得相当的近似程度，对称性低的窄脉冲序列比对称性高的方波序列需要叠加更多的简谐分量。◀

描述一个更窄的波包意味着更多的细节，需要更多的高频分量。



例7-3-2图

7-3-2 波包的简谐分解—傅立叶分析

真正的周期函数在两个方向上都具有无限延展性，既没有出发点，也没有结束点。从这个意义上说，自然界的任何振动和波动在时间上或空间上都是有限的，因此，自然界中并不存在真正的简谐振动或简谐波。我们可以举两个特别的例子:

没有一种乐器可以弹奏出一个绝对的单音，也没有任何一种光源能发射一束绝对单色的光线。

一个有限尺度的波形通常称为**波包**。波包函数由于没有周期性，不符合傅立叶定理的谐函数展开条件。但从物理的角度而言，这些波形应该仍然可以表述为不同频率的简谐驻波模式的叠加，所不同的是，这些定义在无限空间中但尺度有限的波形不再具有有限尺度介质中的边界条件。因此也就不再具有确定驻波模式特征频率的本征值方程。或者说，所有特征频率的驻波模式都是允许的。这样，一个有限尺度的波包的简谐驻波分解可以写为

$$u(t)=\int_0^{\infty} d\omega\,[A(\omega)\cos\omega t+B(\omega)\sin\omega t] \tag{7-3-9}$$

称为函数 $u(t)$的**傅立叶积分**。事实上，一个有限尺度的波包在数学上可以理解为周期无限大的周期函数，这样，函数 $u(t)$仍然可以写成傅立叶级数展开

$$\begin{aligned}u(t)&=\lim_{T\to\infty}\sum_{\omega}^{\infty}\left[\frac{2}{T}\int_0^T u(t)\cos\omega t dt\right]\cos\omega t+\lim_{T\to\infty}\sum_{\omega}^{\infty}\left[\frac{2}{T}\int_0^T u(t)\sin\omega t dt\right]\sin\omega t\\&=\lim_{\Delta\omega\to 0}\sum_{\omega}^{\infty}\left[\frac{\Delta\omega}{\pi}\int_0^T u(t)\cos\omega t dt\right]\cos\omega t+\lim_{\Delta\omega\to 0}\sum_{\omega}^{\infty}\left[\frac{\Delta\omega}{\pi}\int_0^T u(t)\sin\omega t dt\right]\sin\omega t\end{aligned}$$

这实际上就是傅立叶积分(7-3-9)式，其中

$$\begin{aligned}A(\omega)&=\lim_{\Delta\omega\to 0}\sum_{\omega}^{\infty}\left[\frac{\Delta\omega}{\pi}\int_0^T u(t)\cos\omega t dt\right]=\frac{1}{\pi}\int_{-\infty}^{+\infty} dt\,u(t)\sin\omega t\\B(\omega)&=\lim_{\Delta\omega\to 0}\sum_{\omega}^{\infty}\left[\frac{\Delta\omega}{\pi}\int_0^T u(t)\sin\omega t dt\right]=\frac{1}{\pi}\int_{-\infty}^{+\infty} dt\,u(t)\cos\omega t\end{aligned} \tag{7-3-10}$$

称为非周期函数 $u(t)$的**傅立叶变换**，但物理学中常称为波包函数 $u(t)$的**傅立叶频谱**。

类似于实周期函数可以展开为复数形式的傅立叶级数，非周期实函数$u(t)$也可以展开为复数形式的傅立叶积分：

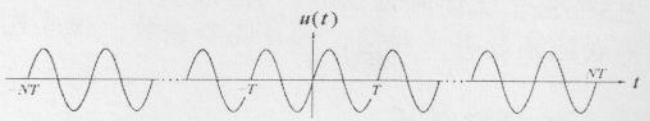
$$u(t)=\int_0^{\infty} d\omega\tilde{u}(\omega)e^{i\omega t} \tag{7-3-11}$$

其中

$$\tilde{u}(\omega)=\frac{1}{\pi}\int_{-\infty}^{+\infty} dt u(t)e^{-i\omega t} \tag{7-3-12}$$

称为函数 $u(t)$的傅立叶变换，或傅立叶频谱。

下面我们讨论两类最典型的波包：有限长度的简谐波列和高斯脉冲的频谱。



例 7-3-3 图有限长的波列

例7-3-3 如图所示的$2N$个简谐振荡构成的时间长度为$2NT$的波列可表示为

$$u(t)=\begin{cases}A\sin\omega_0 t & -NT/2<t<NT/2\\ 0 & \text{otherwise}\end{cases}$$

其中，$\omega_o = 2\pi/T$。求该波列所包含的简谐分量及其大小。

解： 表面上看该波列似乎是一个具有确定频率的简谐波，但实际上并不具有简谐波的周期特征，从整体上看这个波列并没有严格意义上的频率和波长。由于波列为奇函数，其傅立叶分量只包含正弦振荡分量：

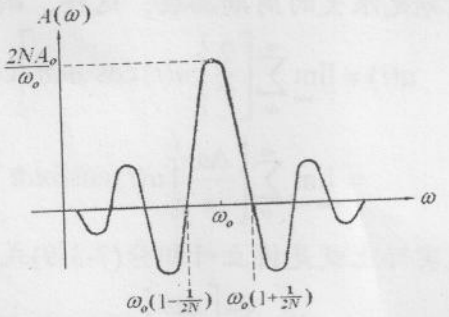
$$u(t) = \int_0^{+\infty} d\omega\, A(\omega)\sin\omega t$$

其频谱由傅立叶变换给出

$$A(\omega) = \frac{2}{\pi}\int_0^{+\infty} dt\, u(t)\sin\omega_o t = \frac{2}{\pi}A_o \int_0^{NT_o/2} dt \sin\omega_o t \sin\omega t$$

$$= -\frac{A_o}{\pi}\int_0^{NT_o/2} dt[\cos(\omega+\omega_o)t - \cos(\omega-\omega_o)t]$$

$$= \frac{2A_o\omega_o}{\pi(\omega^2-\omega_o^2)}\sin\left(2\pi N\frac{\omega}{\omega_o}\right)$$

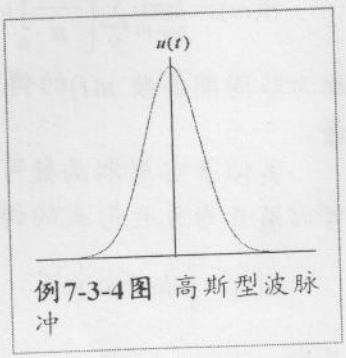


例图给出的上述频谱表明，简谐波列中最主要的频率分量是ω_o，其它比较重要的频率主要集中在$[\omega_o(1-1/2N), \omega_o(1+1/2N)]$区域。当$N$趋于无穷，波列趋于单一频率$\omega_o$的简谐波。◀

例7-3-4 高斯函数(如图)是最常见的波包形状，其数学表达式为

$$u(t) = \frac{}{\sqrt{\sigma\pi}}\exp[-(t-t_o)^2/2\sigma]$$

其中t_o描述波包的中心位置，σ是波包宽度的量度。求其中所包含的简谐频谱函数。



例7-3-4图 高斯型波脉冲

解： 高斯脉冲的频谱由傅立叶变换给出：

$$\tilde{u}(\omega) = \frac{1}{\pi}\int_{-\infty}^{+\infty} dt\, e^{-(t-t_o)^2/2\sigma}e^{-i\omega t} = e^{i\omega t_o}\int_{-\infty}^{+\infty} dt\, e^{-t^2/2\sigma}e^{-i\omega t}$$

$$= e^{i\omega t_o}\int_{-\infty}^{+\infty} dt\, e^{-(t^2+i2\omega\sigma t)/2\sigma} = e^{i\omega t_o}e^{-\omega^2\sigma/2}\int_{-\infty}^{+\infty} dt\, e^{-(t+i\omega\sigma)^2/2\sigma}$$

$$= \sqrt{2\sigma}\, e^{i\omega t_o}e^{-\omega^2\sigma/2}\int_{-\infty}^{+\infty} d\tau e^{-\tau^2} = \sqrt{2\pi\sigma}\, e^{i\omega t_o}e^{-\omega^2\sigma/2}$$

例图b给出了上述频谱图形。高斯函数有一个重要特点：其本身的形状与其频谱的形状的函数形式相同，都是高斯函数，但波包的宽度成互为反比的关系。◀

7-3-3 不确定关系

本节中对各例的振动形态的傅立叶展开都表明，时间宽度较大的波包所包含的主要简谐波的频率分布比较窄，而时间宽度较窄的脉冲所包含的主要简谐波的频率分布比较宽。如在例 7-3-3 中，若以波列长度 $\Delta t = 2NT_o$ 作为脉冲宽度的量度，而以频谱 $A(\omega)$ 中最靠近基频 ω_o 的两个零点之间的区域大小作为主要频率区域：$\Delta\omega = \omega_o/N$。则 Δt 和 $\Delta\omega$ 的乘积为一个常数

$\Delta\omega\Delta t = (\omega_o/N)\cdot 2NT_o = 4\pi$

我们曾在 7-1-5 节中讨论阻尼振子受迫振动时遇到过类似的不确定关系，虽然这两者对应的物理现象并不相同，但确实在数学上存在着共性的成分(见思考题 7-3-3)。有些读者也许会想到，这种关系式是否具有普遍性，或者适用于哪一类现象，其内在的本质又是什么？

不同情形的不确定关系表达式虽然都能以等式的形式表述出来，但其中对 Δt 和 $\Delta\omega$ 的定义在物理测量上并不是一致的，事实上，对不同形式的波包，不一定能从其函数的几何图形对波包宽度和相应的频率宽度给出类似的合理定义。为了探讨不确定关系的普遍性，需要先对任意波形的波包给出普适的 Δt 和 $\Delta\omega$ 的定义。

设波包函数及其频谱函数分别为 $u(t)$和$\tilde{u}(\omega)$，则 Δt 和 $\Delta\omega$ 的定义分别为：

$$\Delta t = \int_{-\infty}^{+\infty}(t-t_o)^2\,|u(t)|^2\,dt \tag{7-3-13a}$$

$$\Delta\omega = \int_{-\infty}^{+\infty}(\omega-\omega_o)^2\,|\tilde{u}(\omega)|^2\,d\omega \tag{7-3-13b}$$

这个定义和确定的误差分布下的测量误差定义相同。

例 7-3-5 分别计算由例 7-3-3 中的 $2N$ 个简谐振荡构成的时间长度为 $2NT$ 的波列和例 7-3-4 中的高斯波包所对应的不确定关系。

解: (1) 对 $2N$ 个简谐振荡构成的波列，有

$$\Delta t = A^2\int_{-NT_o/2}^{NT_o/2} t^2\,|\sin\omega_o t|^2\,dt$$

$$= A^2\int_0^{NT_o/2} t^2(1-\cos 2\omega_o t)dt = A^2\frac{2}{3}\left(\frac{NT_o}{2}\right)^3 - A^2\int_0^{NT_o/2} t^2\cos(2\omega_o t)dt$$

$$= A^2\frac{N^3T_o^3}{12} - \frac{A^2}{2\omega_o{}^2}\frac{NT}{2} = \frac{N\pi A^2}{2\omega_o^3}\left\{\frac{(2N\pi)^2}{3}-1\right\}$$

$$\Delta\omega = \int_{-\infty}^{+\infty}(\omega-\omega_o)^2\left[\frac{2A\omega_o}{\pi(\omega^2-\omega_o^2)}\right]^2\sin^2\left(2\pi N\frac{\omega}{\omega_o}\right)d\omega$$

$$= \frac{4A^2\omega_o^2}{\pi^2}\int_{-\infty}^{+\infty}\frac{\sin^2(NT_o\omega)}{\omega^2-\omega_o^2}d\omega = \frac{4A^2\omega_o^2}{\pi^2}\int_{-\infty}^{+\infty}\frac{\sin^2(NT_o\omega)}{\omega^2-\omega_o^2}d\omega$$

由此给出

$\Delta t\Delta\omega =$

(2) 对高斯波包而言

$$\tilde{u}(\omega) = \int_{-\infty}^{+\infty}dt\,e^{-(t-t_o)^2/2\sigma}e^{-i\omega t} = e^{i\omega t_o}\int_{-\infty}^{+\infty}dt\,e^{-t^2/2\sigma}e^{-i\omega t} = e^{i\omega t_o}\int_{-\infty}^{+\infty}dt\,e^{-(t^2+i2\omega\sigma t)/2\sigma}$$

$$= e^{i\omega t_o}e^{-\omega^2\sigma/2}\int_{-\infty}^{+\infty}dt\,e^{-(t+i\omega\sigma)^2/2\sigma} = \sqrt{2\sigma}\,e^{i\omega t_o}e^{-\omega^2\sigma/2}\int_{-\infty}^{+\infty}d\tau e^{-\tau^2} = \sqrt{2\pi\sigma}\,e^{i\omega t_o}e^{-\omega^2\sigma/2}$$

$$\Delta t = \int_{-\infty}^{+\infty} (t-t_o)^2 e^{-(t-t_o)^2/\sigma} dt = \int_{-\infty}^{+\infty} \tau^2 e^{-\tau^2/\sigma} d\tau = 2(\sigma)^{3/2} \int_{0}^{+\infty} x^2 e^{-x^2} dx$$

$$\Delta\omega = 2\pi\sigma \int_{-\infty}^{+\infty} \omega^2 e^{-\omega^2\sigma} d\omega = \frac{2\pi}{\sqrt{\sigma}} \int_{-\infty}^{+\infty} x^2 e^{-x^2} dx =$$

对应的不确定关系为

$\Delta t\Delta\omega =$

◀

我们分析一下函数和频谱的函数形式相同的高斯函数的 Δt 和 $\Delta\omega$ 的乘积。

从以上的例子可以看出对于脉冲宽度和相应的频率宽度的乘积并不存在普适常数，但从普适的含义上说，存在一个相同数量级的不确定关系式：

$$\Delta\omega\Delta t \approx 2\pi$$

不确定关系式的半定量解释：

设想我们试图用一频率处在范围为 $\Delta\omega = \omega_{max} - \omega_{min}$ 的系列简谐波来合成一个宽度为 Δt 的脉冲，显然对该脉冲有贡献的简谐波应在脉冲的两端相互抵消，即相位相反，而在脉冲中央相互加强，即相位相同。如果对于 Δt，ω_{max} 和 ω_{min} 相对应的周期数目只相差一个：$\Delta\omega = 2\pi/T$，而且这两个谐波在脉冲的两端相位反向，则它们在脉冲的中央同相，而在脉冲尾部反向就能得到满足。于是有

$$\Delta t/T = \Delta t\Delta\omega/2\pi$$

Our condition is then

$$\Delta t(\omega_{max}-\omega_{min})/2\pi \approx 1$$

or

$$\Delta t\Delta\omega \approx 2\pi \qquad (7\text{-}7\text{-}20)$$

Consider an example relevant to the reciprocal relation. 信号和信息传递都必须以脉冲的形式，单纯的简谐波是无限长的，不能用来携载信息。例如，计算机内部或计算机之间的信息传输都是通过二进制的脉冲序列实现的。信息的高速传递要求携载信息的脉冲非常窄， If not, and, in particular, if pulses overlap to make wider pulses, information is lost in the confusion between the pulses. 而按照不确定关系，构建一个很窄的脉冲就必须有一个频谱足够宽的简谐波：脉冲越窄，所需要的频谱就越宽。由于一个简谐波的能量密度和功率正比于简谐波频率 ω 的平方。反比关系意味着妥协：脉冲越窄，信息传输效率越高，所需成本也就越高。

傅立叶定理在物理学和工程学中有许多重要的应用，它与许多不同的边值问题的求解有关。除了上面提到的振动和波的问题，热传播问题，大量的电学和磁学问题的求解都涉及到傅立叶展开。

7-3-4 傅立叶分析的物理实现

傅立叶分析涉及大量的积分，这在实际应用中带来很多的不便。20世纪50年代Cooley和Tukey开发了一种利用计算机进行快速傅立叶分析的算法(称为**快速**

傅立叶变换，简称 **FFT**)可以对各种有实验得到的波形进行快速、精确的积分计算。

在自然界中，存在着许多实际的物质系统对特定频率区间的一些特定的波可以按不同的频率进行分解，从而实现傅立叶变换。例如分光仪器可以将复色光分解为单色光的叠加。这里我们给出对人类的生存最为重要的光波和声波的分解和探测的例子。

光的色散效应

一束复色光分解为单色光的现象称为光的色散，最常见也是最典型的分光器件是棱镜。牛顿在1666年最先利用三棱镜观察到光的色散：一束白光入射到三棱镜一个侧面，经过两次折射，出射光分解为发散的彩色光带(见图7-3-9)。棱镜使光发生色散现象是基于不同频率的光在媒质中的传播速度v (或折射率$n = c/v$)不同这一事实。

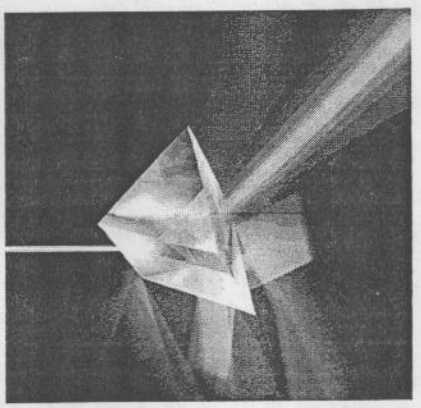


图7-3-2

The dispersion ability originates from a property of the glass, that light of different frequency travel in it at different speed, $v = c/n$, since the refraction index n is a function of the frequency of light. Two refraction processes the beam of sunlight passing through the interface cause the sub-beams of light of different color scatter into various direction.

棱镜对光的色散能力可用于研究光源的光谱组分结构。例如，历史上就是通过太阳的棱镜光谱所获得的太阳光谱中的暗线猜测出相应波长的光被太阳大气中的元素所吸收。利用这种吸收光谱的技术方法可以确定太阳大气中的化学组分，并由此发现了氦元素(氦在希腊语中就称为太阳)。

可以产生光的色散效应的常用器件和装置还有衍射光栅，干涉仪等。

这里的问题是色散不是正比的，需要进行标度。

人耳对声波的分解和辨识

人耳是非常灵敏的声波分解和辨识器官，例如我们一般都能很轻易地辨别演奏同一首乐曲的不同乐器，一些Hi-Fi发烧友甚至能听出同一位大师用不同的小提琴演奏时的细微差别。其原因就在与不同的乐器在演奏同一个单音时所包含的谐频分量的多少和强度是不同的，这说明人耳在音频范围内是非常灵敏精确的傅立叶分析器件。

人耳对声音频率的分解是在内耳的耳蜗(见图 7-3-3)中进行的。耳蜗由一条骨性的蜗管围绕一锥形的蜗轴盘绕 2.5~2.75 周构成。若拉直骨性蜗管可看清骨蜗管内的前庭阶、中阶(膜性蜗管)和鼓阶三个管腔。膜性蜗管是一条充满内淋巴的盲管；而前庭阶和鼓阶内充满外淋巴，它们在蜗顶处通过蜗孔相互交通。声波的感受器官位于基底膜上。感受器官外毛细胞的纤毛顶端嵌入盖膜之中，而内毛细胞的纤毛可能与盖膜没有直接的接触。

基底膜长度约为 31.5mm，但其宽度则自耳蜗底周至耳蜗顶周逐渐增宽。在近镫骨处基底膜的宽度约 0.04mm，至蜗孔处宽度约达 0.5mm。

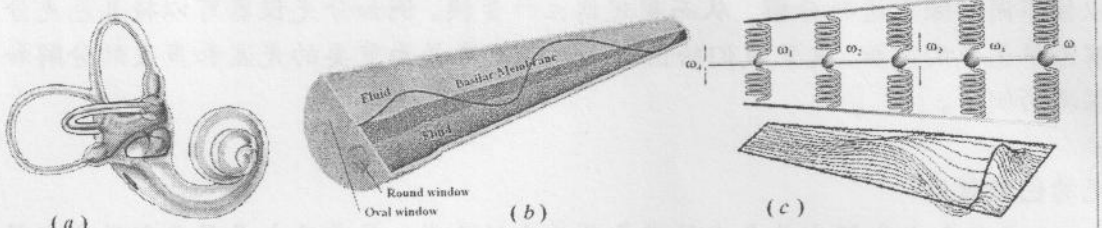


图 7-3-3

毛细胞的长度自耳蜗底周至耳蜗顶周逐渐变长，因此，感受器官的质量可随毛细胞长度的增加而增加。

约 33 毫米长的基膜两端的宽度相差约 5 倍。窄端处的膜绷得比较紧，而另一端的膜则绷得相对比较松。这样的构造及功能可以用一个机械模型进行模拟：一条柔韧的钢带上安装上大小不等的钢球，每一个钢球在侧向用不同劲度系数连接到固定侧壁上(见图 7-3-3)。其中小质量的球和劲度系数大的弹簧配制形成较高频率的振子，大质量的球和劲度系数小的弹簧配制形成较低频率的振子，形成沿着耳蜗向内频率连续减小，共振频率区域范围约 20,000 Hz 到 20 Hz。

设想在钢条的一端施加一个频率为f_o的扰动，该扰动将在钢条中向另一端传播，在另一端反射后在钢条中形成一个频率为f_o的驻波。钢条所连接的刚球，及其两侧的弹簧系统将在钢条的驱动下作受迫振动，其中特征频率接近 f_o的振子会有比其它振子大很多的振幅，由此在该区域形成一个振动的高峰。随着入射声波频率的升高，振动峰的位置向膜的近椭圆窗口的前部移动。因而人耳能够区分声波中很小的频率变化，以及声波中各谐波组分很细微的变化。图7-36给出受到频率为几百赫芝的驱动时的某一时刻振幅分布的示意图。实际的振动振幅小于3微米，但可以被约含有三万左右的毛细胞系统精确探测出来。沿着整个膜，实际测量连续进行，结果被馈入精密的数据分析网络，然后送至大脑。

实验表明，不同频率的声音引起的行波都从基底膜的底部，即靠近卵圆窗膜处开始，但频率不同时，行波传播的远近和最大行波的出现部位各有不同，也就是说，振动频率愈低，行波传播愈远，最大行波振幅出现的部位愈靠近耳蜗顶部的基底膜，而且在行波最大振幅出现后，行波很快消失，不再传播；相反地，高频率声音引起的基底膜振动，只局限于卵圆窗附近耳蜗底部的基底膜。

产生于盖膜和网状层之间的侧向(基底膜横轴方向)的相对位移称**辐射(横向)剪切**(radial shear)。此外，还有一种沿基底膜纵轴方向的位移产生纵向剪切。

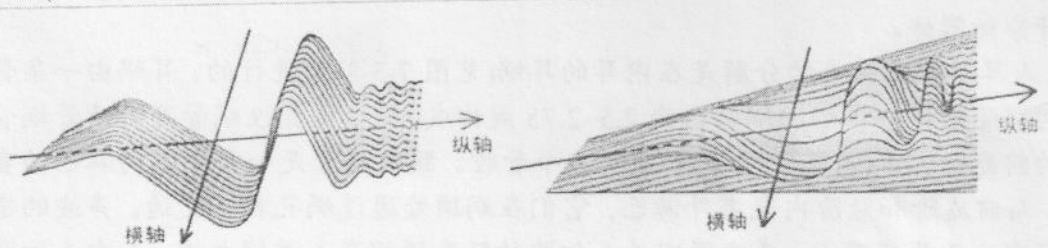


图 行波在丝带上传播表现为沿纵轴的位移。行波在基底膜上传播表现为交替的峰与谷的形式，每一位移处都呈现为基底膜的横轴和纵轴方向的弯曲。

鼓膜与听骨链将空气振动能量传递到外淋巴，完成了第一次阻抗匹配作用。而外淋巴的振动通过在基底膜上的行波方式，将振动转变为盖膜与网状板之间的剪切运动，使毛细胞的纤毛弯曲偏转，对毛细胞产生机械刺激，从而完成了第二次阻抗匹配作用，即外淋巴与毛细胞之间的阻抗匹配。毛细胞顶部纤毛的弯曲或偏转是对声波振动刺激的一种特殊反应形式，是将机械能转变为生物电，引起毛细胞兴奋的开始。纤毛的弯曲使毛细胞顶部机械门控离子通道开放，阳离子内流引起细胞兴奋。

7-4 行波及其传播特性

前面已经提到，耦合振动在整体形态上存在两类基本振动位形，一类表现为系统中各个振子独立的振动，耦合特性表现为任意两个振子之间的相对位相不随时间发生变化，整体上表现为一个不发生传递的波形，因而称为**驻波**。另一类表现为局部振子的振动向整个系统扩散，振动在系统中由近及远地传播，这种振动形态称为**行波**。前两节是从稳态振动，即驻波的角度讨论耦合振动体系的运动特性，本节讨论在多自由度耦合系统，特别是连续介质中振动形态的传播(即行波)的行为特征。

7-4-1 波动方程的行波解

图 7-4-1 显示的是平静的湖面被一颗投入的小石子激发起的一系列同心涟漪，这种常见的水面波非常直观地展现了连续介质波的振动传递行为。波是一种相当普遍而且基本的物质内部运动形式[16]，它以振动的方式在空间传播能量，但携载能量的物质质点或质元并不跟随着振动而迁移。物体内部的这种直接依赖质点或质元之间的耦合来传递振动的运动形态称为**机械波**。

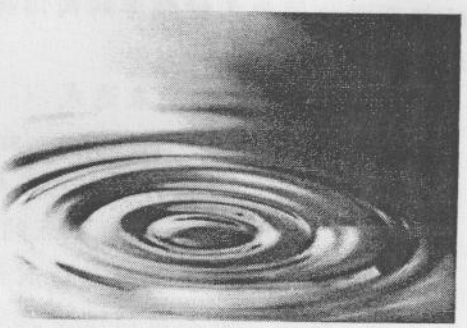


图 7-4-1 向平静的湖面投入一颗小石子，这种局部扰动激发的局部振荡向四周传播形成一系列同心圆形状的涟漪—行波

由于驻波和行波都是耦合振动系统的运动方程的解，因而两者之间不是彼此独立的，其中任何一个都可以表示为另一个的线性叠加。以描述离散质点链的运动方程(7-2-12)

$$\frac{d^2u_n}{dt^2}=-\omega_o^2(2u_n-u_{n-1}-u_{n+1})$$

为例，其简正模式解：$u_n(k)=Ae^{-i\omega t}\sin(kna+\varphi_o)$可以改写为

$$u_n(k)=A_1e^{-i(\omega t-kna)}+A_2e^{-i(\omega t+kna)} \tag{7-4-1}$$

两项描述的分别是随着时间流逝向右或向左迁移的简谐振动位型，分别称为**右行简谐波**和**左行简谐波**。振动位型的迁移速度(这里是指简谐波相位的迁移速度)为

16 这里所说的内部运动是相对于整体运动而言的。

$$v=\frac{\omega}{k}=\frac{\lambda}{T} \tag{7-4-2}$$

称为简谐波的**相速度**。而从一维均匀连续介质的运动方程—波动方程(7-2-18)

$$\frac{\partial^2 u}{\partial t^2}=v^2\frac{\partial^2 u}{\partial x^2}$$

可以直接得到任意物理形态 $f(x)$ 向左或向右迁移的**左行波**和**右行波**解：

$$u_L(x,t)=f(x+vt),\ u_R(x,t)=f(x-vt) \tag{7-4-3}$$

由此从任意时刻的振动位形出发可以得到系统在此后任意时间的振动位形。例如，若弹性弦的向右移动位形函数图像在 $t=0$ 时刻由 $u(x,t=0)=f(x)$ (见图 7-4-2)，则在 T 时间后，该波形以速度 v 向右传播了距离 vT (见图 7-4-2b)，描述其波形的函数为 $u(x,T)=f(x-vT)$。

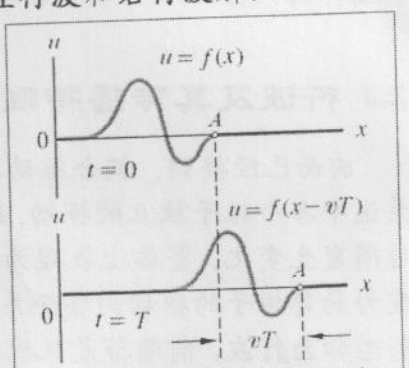


图 7-4-2 一个特定波形的右行波

下面我们主要讨论连续介质中的行波及其主要行为特征[17]。

简谐行波

行波函数 $u(x,t)=f(x\pm vt)$ 是齐次波动方程(7-2-18)的通解，由于任意形态的振动都可以分解为各种频率的简谐振动的叠加，下面我们主要讨论简谐行波的行为特性。

描述一维简谐行波的表达式

$$u(x,t)=u_o\sin[k(x\pm vt)+\varphi_o]=u_o\sin(kx\pm\omega t+\varphi_o) \tag{7-4-4}$$

中的 u_o 称为行波的波幅，而引入参数 k 是为了使简谐函数中的相位 $\varphi=kx-\omega t+\varphi_o$ 成为无量纲参量(φ_o 是处于原点的质元的简谐振动初相位)。类似于 ω 是简谐波关于时间的圆频率[18]，描述单位时间对应的相位延迟，k 描述的是单位传播距离引起的相位延迟，因而是简谐波的空间圆频率，称为简谐波的**角波数**，一般简称为**波数**[19]。

简谐波函数(7-4-5)式表示为空间和时间的周期函数，它可改写为

$$u(x,t)=u_o\sin\left\{2\pi\left(\frac{x}{\lambda}\pm\frac{t}{T}\right)+\varphi_o\right\} \tag{7-4-5}$$

其中时间周期和空间周期分别为

$$T=\frac{2\pi}{\omega}\quad 和\quad \lambda=\frac{2\pi}{k} \tag{7-4-6}$$

空间周期 λ 称为波长，描述同一时刻介质中相隔 λ 的两个空间点具有相同的振动状态。

一维简谐行波也可以用指数型波函数表述为

$$\tilde{u}(x,t)=u_o e^{-i(\omega t\pm kx-\varphi_o)}\equiv\tilde{u}_o e^{-i(\omega t\pm kx)} \tag{7-4-7}$$

---

17 利用连续介质模型讨论波动一是更直观，二是可以在时间和空间都利用微分方程分析。

18 ω 是介质中质元的简谐振动圆频率，由于波是振动的传播，因此也就是简谐波的圆频率。

19 波长的倒数反映了波在单位长度内的空间周期数，称为波的空间频率，有时(光谱学中)也称为波数。

其中 $\tilde{u}_o = u_o e^{i\varphi}$ 为复振幅，其中包含了相位常数。指数型波函数虽然不很直观，但计算波的合成或叠加则会方便很多。

三维空间中最简单的行波是沿一个特定方向传播的简谐波。若将这样一个行波的传播方向取为坐标系 S' 的 z'-轴，则相应的波函数在 S' 坐标系中表示为 $u(z', t) = A\cos(kz' - \omega t)$。而在其它坐标系 $S(x, y, z)$，上述简谐波表示为

$$u(\mathbf{r}, t) = A\cos[\omega t - (kx\cos\alpha + ky\cos\beta + ky\cos\gamma)]$$
$$= A\cos[\omega t - (k_x x + k_y y + k_z z)] \qquad (7\text{-}4\text{-}8)$$

其中 $(\cos\alpha, \cos\beta, \cos\gamma)$ 为 S' 系的 z'-轴在 S 坐标系中的方向余弦。由于波数 $k = 2\pi/\lambda$ 代表沿波传播方向位移一个单位所对应的相位变化(沿 $\mathbf{k}$ 方向移动一个波长 λ，相位增加 2π)，上式中的 k_x, k_y 和 k_z 分别代表波沿着 x、y 和 z 轴移动一个单位长度所对应的相位变化(如沿 x 轴前进 $\lambda/\cos\alpha$，相位增加 2π)。则 $k_x x + k_y y + k_z z$ 可以理解为矢量 $\mathbf{k}$ 和位置矢量的标积：$\mathbf{k}\cdot\mathbf{r}$，称矢量 $\mathbf{k}$ 为波矢量。因而(7-4-8)可表示为

$$u(\mathbf{r}, t) = A\cos(\omega t - \mathbf{k}\cdot\mathbf{r}) \qquad (7\text{-}4\text{-}9)$$

在任一给定的时刻 t，波场中相位相同[20]的点的集合所构成的面称为**等相面**，也称为**波阵面**(或**波面**)，由式(7-4-9)确定的等相面是垂直于 $\mathbf{k}$ 的平面，波面为平面的波称为**平面波**。

波的等相面一般表现为空间三维曲面族。最典型的等相面除了平面、还有球面和圆柱面，相应的波分别称为**球面波**和**柱面波**。球面波通常对应于有限远处理想点源发出的波，柱面波则对应于无限长的线波源发出的波，而实际的平面波对应于无限远处理想点源发出的波(即波面曲率半径趋于无限大时的球面波)。

在各种等相面形状的波中，球面波和平面波问题具有更普遍的意义，因为任何波源都可看成由许多点波源的集合，因而任何一个复杂形状的波面总可以分解为球面波或平面波的叠加。或者说，一个任何复杂的曲面都可以按球面简谐函数或平面简谐函数展开[21]。

均匀介质中的二维和三维波动方程分别为

$$\frac{\partial^2 u}{\partial t^2} = v^2\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) \quad \text{和} \quad \frac{\partial^2 u}{\partial t^2} = v^2\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right) \qquad (7\text{-}4\text{-}10)$$

在具有各向同性特点的传播情形中，它们可在极坐标或球坐标中的表达式分别为

$$\frac{\partial^2 u}{\partial t^2} = v^2 \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) \qquad (7\text{-}4\text{-}11)$$

$$\frac{\partial^2 u}{\partial t^2} = v^2 \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right) \qquad (7\text{-}4\text{-}12)$$

从方程的形式容易看出，即使在各向同性介质中传播的单一频率的柱面波和球面波，其波函数也不再具有空间周期性。空间周期性的丧失包含两个方面，一是波幅的衰减，这一点也可以从在波的传播过程中能量被不断分散的角度来得到理解：

---

[20] 读者应注意分辨相位相同和简谐函数取相同数值时的相位之间的差异。

[21] 这实际上就是上一节所讨论的一般振动形态按特定驻波模式的傅立叶展开，这里的曲面函数按球面简谐函数或平面简谐函数展开所对应的是三维的驻波模式，因而是两种不同的三维傅立叶展开。

相同能量分布在越来越大的波面上，每一点的振幅就会随着与波源的距离增加而减弱。二是两个波峰之间的距离是否相等，对于单频率的柱面波而言，方程解的空间波形实际上就是例 7-2-3 所讨论过的 $m=0$ 阶**贝塞尔函数** $J_o(kr)$，因而在单频柱面波的形态中，相邻波峰之间不是等间距的。而对于单频率的球面行波，相应方程的解有比较简洁的函数形式：

$$u_\omega(r,t)=u_o\frac{e^{-i(\omega t\pm kr)}}{r} \tag{7-4-13}$$

可以看出，相邻波峰之间仍然是等间距的。

横波和纵波

不同于驻波，行波因涉及振动的传播方向，因而讨论行波特征还需要说明振动方向和传播方向之间的关系。这种关系有两种基本类型，振动方向与传播方向相同的波称为**纵波**(见图 7-4-3a)，振动方向与传播方向正交的波动称为**横波**(见图 7-4-3b)。从动力学角度而言，纵波是由于正应力作用在介质内引起的体应变形成的，而横波是因切应力作用在介质内引起的切应变形成的。由于固体材料兼有切变弹性和体变弹性模量，固体中既能传播横波，又能传播纵波。而流体因不具有切变弹性，只有体变弹性，故流体介质内部只能传播纵波，而不能传播横波[22]。纵波的特点是相应空间点的振动状态相对于传播方向具有轴对称性，即在所有通过传播方向所作的截面内具有相同的振动状态。横波的特点是振动状态相对于传播方向不具有轴对称性，即在所有通过传播方向所作的截面内的振动状态可能各不相同，这种振动状态相对于传播方向不对称的现象称为波的**偏振**。对于给定的波传播方向存在有两个独立的偏振方向，两者的合成因相位间的差异使横波往往比纵波呈现出更为丰富的可观测效应。

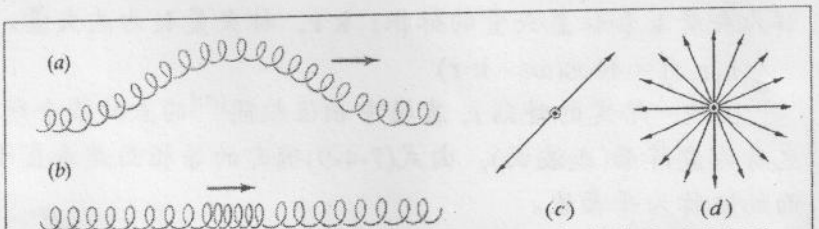


图7-4-3 (a) 纵波，(b) 横波，(c) 一从页面出射的平面偏振波，(d) 从页面出射的非偏振波，其振动方向可理解微是大量线振动方向的随机叠加。

通常的波动是兼有横波和纵波两种成份，例如，水面流体质元的运动轨迹既不是沿水面的直线振动，也不是垂直与水面的直线运动，而是一种椭圆运动(见图 7-4-4)，说明水面波是横波和纵波的叠加。

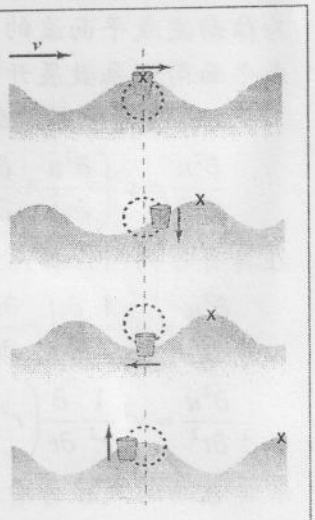


图 7-4-4 水面波中流体质元的运动轨迹。

[22] 水面波与水体中的波不同。水面分子间具有横向的表面张力，故可以产生横波。

波中的能量和能流

弹性波是通过介质的形变来传递能量的，因此描述能量的传递特性是表征波的动力学行为的一个重要方面。我们仍然以一个在弹性均匀弦上的右行横波$u = u(x-vt)$为例展开讨论。设想在特定的时刻t，给定位置x 处的弹性弦微元Δx相对于平衡位置有横向位移u，运动速度为$\partial u/\partial t$，则它所携载的动能为

$$\Delta K \approx \frac{1}{2}\mu\left(\frac{\partial u}{\partial t}\right)^2 \Delta x \tag{7-4-14}$$

其中μ为弦的质量密度。另外，线元的形变量为

$$\Delta l - \Delta x = \sqrt{(\Delta x)^2 + (\Delta u)^2} - \Delta x \approx \Delta x\sqrt{1+(\partial u/\partial x)^2} - \Delta x \approx \frac{1}{2}(\partial u/\partial x)^2 \Delta x$$

其中利用了弦的局部变形很小的假设。利用线弹性近似来计算该弦元因形变而具有的势能

$$\Delta U \approx \frac{1}{2}T\left(\frac{\partial u}{\partial x}\right)^2 \Delta x$$

利用波速度的表达式$v = (T/\mu)^{1/2}$，上述势能可以表示为

$$\Delta U \approx \frac{1}{2}\mu v^2\left(\frac{\partial u}{\partial x}\right)^2 \Delta x \tag{7-4-15}$$

对于一维行波的表示式：$u(x,t) = u(x \pm vt) = u(\varphi)$，可得

$$\frac{\partial u}{\partial t} = \frac{du}{d\varphi}\frac{d\varphi}{dt} = \pm\frac{du}{d\varphi} \qquad \frac{\partial u}{\partial x} = \frac{du}{d\varphi}\frac{d\varphi}{dx} = \frac{du}{d\varphi}$$

分别代入(7-4-14)和(7-4-15)式可得：弦微元的动能和势能在任何时候都相等[23]

$$\Delta K = \Delta U \tag{7-4-16}$$

由此可得到弹性弦的能量密度为

$$\varepsilon = \frac{\Delta E}{\Delta x} = \frac{\Delta K + \Delta U}{\Delta x} = \mu v^2\left(\frac{du}{d\varphi}\right)^2 \tag{7-4-17}$$

对于弹性弦上的简谐波$u(x,t) = u_o\sin(kx \pm \omega t)$而言，其能量密度为

$$\varepsilon = \mu\omega^2 A^2 \sin^2(\omega t - kx) \tag{7-4-18}$$

类似于简谐振子的能量：与振幅的平方和振动的频率的平方都成正比。

通过弦上每一点x传递的功率可以表示为

$$P = \frac{dE}{dt} = \frac{dE}{dx}\frac{dx}{dt} = \varepsilon v = \mu v\omega^2 A^2 \sin^2(\omega t - kx) \tag{7-4-19}$$

功率是时间的函数，其平均值通常称为波的**强度**

$$I = \frac{\omega}{2\pi}\int_0^{\omega/2\pi} P dt = \frac{1}{2}\mu v\omega^2 A^2 \tag{7-4-20}$$

---

[23] 有些读者可能会困惑：简谐振子中的动能和势能始终是反相的，简谐波中局部区域的动能和势能为什么始终是同相的。其中的要点在于单一振子在振动过程中，振子的平衡位置不变，振子的形变势能是以质点的位移数量来量度的，而对弹性弦上的弹性波而言，弹性势能不是质点相对于平衡位置的位移，而是以介质的形变来量度的。质点在动能最大的位置所处的位置虽然是原来的平衡位置，但与相邻质点的距离却相差最大，因而弹性势能也达到最大值。即相对于对空间的与相对于时间的变化率相同。

在三维介质中，波的强度定义为单位时间内穿过单位面积的能量，在各向同性介质中，波面为一组以点波源为中心的同心球面，对两个半径分别为r_1和r_2的同心球面而言，波面的面积分别为$4\pi r_1^2$和$4\pi r_2^2$，如果介质中不存在耗散或吸收，则有$P = 4\pi r_1^2 I_1 + 4\pi r_2^2 I_2$。由此可得球面波的能流与波源距离的平方成反比。

$$I(r) \propto \frac{1}{r^2} \tag{7-4-21}$$

7-4-2 典型波动的例子

为了让读者对上面讨论的波的参量和特性有一个直观的理解，下面针对最为常见的声波、光波和一般的水波做一简单的介绍和讨论。

声波

人耳一般是通过感受空气中的压力起伏听到声音的，频率响应范围在 20~20 000Hz，因而该频率区间内的机械波通常被称为**声波**。频率低于 20Hz 的声波称为**次声波**，高于 20 000Hz 的声波称为**超声波**。不同动物的耳朵对声波的响应频率范围并不一致，如长颈鹿、象、蓝鲸等动物可以感受到次声波，而海豚、狗、蝙蝠等动物则可听到超声波。

由于空气中的弹性只是表现为体积模量，因而气体中传播的声波是纵波[24]，表现为局部体积元出现交替的膨胀或压缩，声音的传播就是介质内密度发生疏-密交替变化的过程。由于气体中的密度与局部压强、温度等参量有关，气体中的声波可以用局部压强、密度、温度等物理量的变化量描述。如果扰动量比较小，声波的传播满足线性波动方程(7-2-18)。其中声音的速度为：

$$v = \sqrt{\frac{B}{\rho}} \tag{7-4-22}$$

其中 B 为空气的体积模量。声扰动在介质内产生的局部压强变化 $\tilde{p} = p - p_0$ 称为**声压**，由于声压比较容易测量，也常用声压描述声波的参量。实际上声场中气体的密度，声压及质点速度等声学量的变化相互关联[25]，通过声压可以求得流体中局部密度变化等其他物理量。通常所说的声压往往是指**有效声压**，即瞬时声压对时间的均方根值

$$p_{\text{eff}} = \sqrt{\frac{1}{T}\int_0^T \tilde{p}^2 dt} \tag{7-4-23}$$

其大小反映声波的强弱。为了使读者对声压大小有一直观的数量级概念，这里给出一些典型的具体例子：人耳对 1 千 Hz 声音的可听阈(即刚能觉察到其存在时的声压)约 2×10^{-5} 帕；微风吹动树叶的声压约 2×10^{-4} 帕；室内的高声谈话声(相距 1 米处)约 0.06~0.1 帕；交响乐演奏声(相距 5~10 米处)约 0.3 帕；飞机的强力发动机发出的轰鸣声(相距 5 米处)约 10^2 帕。

光波

历史上对光本质的波动学描述和微粒学描述几乎是同时开始的，其中对光的

---

24 声波在粘性流体中有一定的横波成分。这里仅讨论理想流体中声波的特性。

25 详细的讨论见第 4 部分，第 14-1 节

波动特性的认识起源于对光在自由空间和各种透明介质中的传播特征的描述。但在很长的一段历史时期内，对光振动所反映的物质特性却并不清楚。即使在发现了光的干涉和衍射以及偏振特性以后，人们也只是从唯像的角度给出更加完整的光的波动图象描述。直到 19 世纪中叶，英国物理学家麦克斯韦统一了电磁场理论，并将光的波动理论纳入了电磁理论体系。该理论揭示了光的电磁本质，说明了光波只是整个电磁波谱中能够引起人眼视觉响应的一个很小的子集，其频率范围约为 8×10^{14}~4×10^{14} Hz (相应的波长是 380~760 nm)。而介质对光波的折射率与其他电磁波同样起因于介质的电磁特性。因为光无需传播介质而可以在真空中传播，是一种不同于机械波的特殊波动[26]。

电磁场理论表明，空间局部区域中的电磁扰动因电场 **E** 和磁场 **B** 的相互感应而在空间中传播。这两种矢量场的每一个直角分量都满足完全相同的运动方程 (7-2-19)，其中 u 可以代表电场 **E** 和磁场 **B** 的任何一个直角分量，因此这些电、磁分量都以相同的速度 $c=(\varepsilon_o\mu_o)^{-1/2}$ 在真空中以波的形式传播。另外光波场中的电振动和磁振动在相位、振幅和偏振方向上有确定的关系，其中电场 **E** 和磁场 **B** 的振动始终保持同相，相互垂直，而且垂直于波的传播方向。因而只需选择其中一个就能代表光场，即可以用标量波理论来描述光的传播。通常选择电场强度矢量 **E**[27] 中一个分量的波动方程描述整个光场的演化特性。

光是纯粹的横波，光振动矢量在垂直于传播方向的平面内可以有许多不同的振动状态(光的偏振态)。通常将光振动方向与光传播方向构成的平面称为偏振面，光的偏振态可由偏振面的不同行为特性分为完全偏振、非偏振和部分偏振三种基本类型。其中完全偏振类型因两个正交偏振分量的相位差可以分为**平面偏振**(又称**线偏振**)和**椭圆偏振**(包括**圆偏振**)。

根据光源的发光机制，来自光源任一原子或分子的任一次辐射波列，都具有确定的振动方向和初相位，由于构成光源的大量原子或分子在空间分布上的各向同性，以及一般探测器的响应时间远大于光振动的周期，由一般光源发出的众多波列构成的光矢量在垂直于传播方向的平面上既具有空间分布的均匀性，又具有时间分布的均匀性。空间分布上的均匀性表明，光矢量的偏振面包含各种方向且在各个方向的平均大小相同；时间分布上的均匀性表明，各个光矢量的初相位取 0 到 2π 之间的任意值。这类偏振面具有各种取向，且相位随机分布的平面偏振光的集合称为**自然光**。

由于沿任何方向振动的光矢量均可分解为沿两个给定的正交方向的分量，而随机性又保证了所有光矢量沿这两个正交方向每一分量的和的时间平均值相等。因此，自然光实际上可分解成两个强度相等、振动方向正交但相位各自随机变化的平面偏振光。设自然光的强度为 I，则所分解的两个正交平面偏振光分量的强度分别为 $I/2$。值得注意的是，这里的两个平面偏振光的相位各自独立地随机变化，因此不能再合成为一个单一矢量。

---

[26] 第九章和第十章将对这类特殊波动作比较仔细的讨论。

[27] 这其中还有一个实际背景：光与物质相互作用过程中扮演主角的是电场，比如光合作用、视觉效应、光电效应和光热效应等，其中发生的物理过程主要是电场与分子、原子或电子的相互作用，因为光频极高，介质的磁化机构几乎冻结。这样，光波传播行为就被简化为以单一矢量波 $\mathbf{E}(P,t)$ 来描述。

若自光源发出的光振动强度沿两个正交方向的时间平均值不相等：其在其中一方向上取极大值 I_{max}，而在其正交方向上取极小值 I_{min}，这类光波称为**部分偏振光**。为描述部分偏振光的这种偏振特性，可以定义参量：

$$P=\frac{I_{max}-I_{min}}{I_{max}+I_{min}} \tag{7-4-24}$$

称为**偏振度**。显然 $0<P<1$，其中 $P=0$ 表示自然光；$P=1$ 则对应于平面偏振光。

水波

水波是最常见到的波动，不同于大气中的声波和真空中的光波，水波一般表现主要在两种介质界面区域传播的水面波。虽然水的粘滞性和可压缩性非常小，通常可以看作没有切变模量的理想流体，但在接近水面的水波中既有横波也有纵波。水体表面在处于平衡时是水平的，图 7-4-5 显示当区域 A 的液面相对于平衡时的自由水平面发生向下位移，液体的不可压缩性使部分液体从区域 A 流向其周边的 B, B′区域，从而使这些区域的液面隆起。液面一边为大气压，另一边因水面隆起有附加的液体压强，两面的压强差作用到质点 P 上的合力 **F** 也就是使液面平直的恢复力。由于 **F** 与液面垂直，也就与质点 P 向低处流动的速度 **v** 垂直，从而构成驱动流体质点作圆周运动的向心力。**F** 的竖直和水平两个分量导致流体质点在竖直和水平方向作简谐振动，这两个振动的合成为圆周运动，两个振动分量的相位差为 90°。

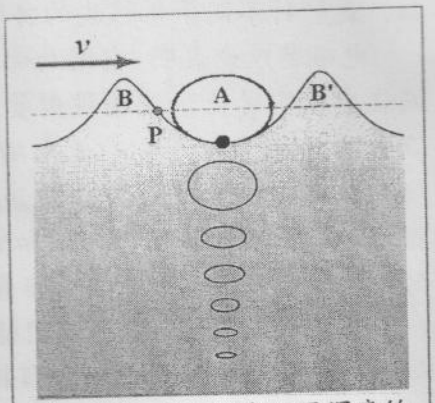


图 7-4-5 水面波中不同深度的流体质元的运动轨迹。

实际上，水面上出现波动时会出现的平复水面平衡的恢复力除了重力，还有表面张力。重力作用导致水面的坡度减小，而表面张力的作用则是通过减小水面的曲率来减小水的表面积。但这两种恢复力导致的两类水面波由具不同的波长尺度：表面张力很小，由此导致的水面扰动的传播表现为波长为几个厘米的涟漪，称为表面张力波，而重力波的波长则可由几米达数百米。下面利用量纲分析估计这两种波的相速度，并由此估算它们各自的波长大小。

因波速与介质中的恢复力和惯性有关，与表面张力波的传播速度 v 相关的参量为表面张力系数 γ 和密度 ρ。考虑到可能的色散效应，还应计入波长 λ。这四个参量可以构成的无量纲参数为 $Q=v^2\rho\lambda/\gamma$，则表面张力波的相速度可以表示为

$$v=C\sqrt{\frac{\gamma}{\rho\lambda}} \tag{7-4-25}$$

其中 C 是无量纲常数。而介质中与重力波相关的恢复力参量和惯性参量分别为重力 ρg 与和惯性密度 ρ，两者之比抵消了 ρ 的效果，因而与波速 v 相关的参量为重力加速度 g，另外考虑到可能的色散效应还需要计入水的深度 h[28]以及波长 λ。这四个相关参量可构成两个无量纲参量：$Q_1=v^2/g\lambda$，$Q_2=\lambda/h$。由此可将重力波的相速度与相关参量的关系表示为

28 水底的存在对水面的运动有阻尼作用，

$$v=\sqrt{g\lambda}\cdot f(\lambda/h) \tag{7-4-26}$$

对于深水波：$\lambda \ll h, f(Q_2)\approx C$。则

$$v=C\sqrt{g\lambda} \tag{7-4-27}$$

对于浅水波：　$\lambda \gg h$, 对应于长波近似，不会出现色散效应，因而有

$$v=C'\sqrt{gh} \tag{7-4-28}$$

通过ω^2(简谐振动系统中，相当于单位质量的物体发生单位长度的位移时受到的恢复力)

$$\omega^2(k)=k^2v^2(k)=k^2v_c^2(k)+k^2v_g^2(k)=gk+\gamma k^3/\rho$$

可以得到一般情形的色散关系

$$\omega(k)=\sqrt{gk+\gamma k^3/\rho}\,f(kh) \tag{7-4-29}$$

为估计两种波的波长，考虑表面张力和重力提供的恢复力相等时$k^2=g\rho/\gamma$所对应的波长，水的相应参量为$\rho=10^3\ \mathrm{kg\cdot m^{-3}}$，$\gamma=0.072\mathrm{N\cdot m^{-1}}$。

$$\lambda_c=\frac{2\pi}{k_c}=2\pi\sqrt{\frac{\gamma}{\rho g}}=0.017m$$

因此表面张力系数只对波长很小的(一般不超过几厘米)水波才有明显的贡献。由此可知湖面上的涟漪主要成分为表面张力波，而海面上波长为几米、甚至几百公里的大波是重力波。

7-4-3 不同惯性系中的简谐波特性的观测—多普勒效应

描述粒子运动的力学参量是粒子的动量和能量，而描述波的物理参量是波矢量 **k** 和时间频率 ω。由于运动是相对的，描述粒子运动的动量和能量与所选择的参照系有关，那么描述波动的空间频率和时间频率是怎样随不同参照系改变的呢？

我们从一个很极端的例子开始讨论。一个在马路上飙车的人在闯红灯后被警察拦下：

“你没有看见红灯吗”

“没有啊，我真的是看见绿灯才开过去的！”司机一脸无辜。

这样的辩解当然不可能逃脱有关的处罚。但就物理学原理而言，两个处于不同惯性系的观察者看同一束光时所看到的颜色确实是不同的，而且向着静止的交通灯运动的观测者所看到的红灯的颜色也确实是向绿色方向偏移的，只是对于一般的车行速度，颜色的变化没有如此显著而已。这一现象是奥地利物理学家多普勒[29]在 1842 年发现的。他当时观察到光源因运动其颜色发生变化的现象，并在光的波动学基础上提出了物理解释。多普勒由此预言这一效应在各种波现象中都存在。1845 年，荷兰的 Christoph Hendrik Diederik Buys Ballot 在实验中证实了声波的多普勒效应：在不同参照系观察同一列简谐波测量到的频率或者波长不同，在同一个惯性系中观测不同运动速度的波源发出的相同的简谐行波，测量到的频率或波长也不相同。本节将分别讨论这两种不同情形的多普勒效应。

---

[29] Christian Johann Doppler (1803-1853)奥地利物理学家，他后来还曾提出恒星的不同颜色可能是由于它们相对于地球的运动速度不同而引起。但被以后的实验观测否定：恒星的不同颜色是因为它们的温度不同。

波源运动而观察者静止

当一个处于均匀介质中的静止的波源向四周发射频率为 f_o 的波时，则波面都是以该波源为中心的同心圆或同心球面。因此在任意方向接收到的波的波长都为 $\lambda_o = v/f_o$，其中 v 为该波在所处介质中的速度。

设想波源相对于介质向右匀速运动，运动速度 v_S 小于波在介质中的传播速度 u。这时每一个球形波面的中心都相继均匀的向右位移(见图7-4-6)。波阵面在波源右侧聚集而在左侧则相对分散。这样处在波源前面的观测者所接收到的信号的波长 λ 要比波源静止时的波长 λ_o 短了一个数量：$\lambda = \lambda_o - v_S/f_o$，相应的频率上升为

$$f = \frac{u}{\lambda} = \frac{u}{u - v_S} f_o = \frac{f_o}{1 - v_S/u} \tag{7-4-30}$$

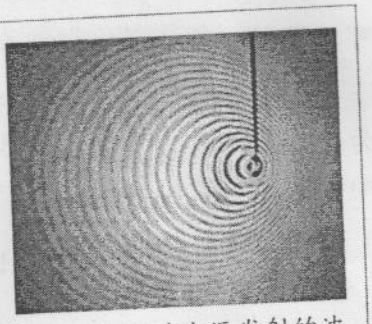


图 7-4-6 运动波源发射的波在介质中形成的波阵面。

类似地，处在波源前的观测者所接收到的波信号的频率要比波源静止时的情形低：

$$f = \frac{f_o}{1 + v_S/u} \tag{7-4-31}$$

而对于其他位置的观测者所接收到的波信号的频率介于以上两个频率之间(见习题 7-4-6)。

波源静止而观察者运动

设想波源相对于介质静止，而观测者相对于介质以速度 v_o 向右运动。每一个球形波面经过观测者的速度为 $v = u + v_R$，而波长没有发生变化。则观测者在单位时间内所接收到的波面的数目要比观测者静止时的情形多了数量 v_R/λ，则观测者所接收到的波的频率为

$$f = \frac{u + v_R}{\lambda} = \frac{u + v_R}{u\lambda/u} = \frac{u + v_R}{u} f_o = (1 + \frac{v_R}{u}) f_o \tag{7-4-32}$$

对于一个背离波源运动的观测者而言，由类似的分析可以知道所接收到的波信号的频率要比静止观测者测量到的频率低：

$$f = \left(1 - \frac{v_R}{u}\right) f_o \tag{7-4-33}$$

而对于其它位置的观测者所接收到的波信号的频率介于以上两个频率之间(见习题 7-4-7)。

波源和观察者均相对于介质运动

如果波源和观测者都相对于介质运动，而且波源和观测者的运动是共线的，运动速度分别为 v_S 和 v_R。则观测者同时可以探测到球形波面经过观测者的速度为 $v = v \pm v_R$，而波长也发生了相应的变化。则观测者所接收到的波的频率为

$$\nu = \frac{u}{\lambda} = \frac{u}{u - v_s} v_o = \left(1 - \frac{v_s}{u}\right)^{-1} v_o \tag{7-4-34}$$

请读者注意分辨其中的正负号所对应的具体物理场景。

例 7-4-1 多普勒流速仪 利用**多普勒效应可以**测量血管中血液的流动速度。将一个超声波源和一个超声波探测器分别放在血管的两旁，波源发射的超声波透入血管后碰到红细胞后被反射，这些反射信号与入射信号的频率差携载着红细胞的运动速度信息，求血液的流速。

解：由于 θ 角很小，超声波的传播方向和血流方向近似一致，红细胞接收到的声频为

$$f=\frac{u-v_O}{u}f_o$$

其中u是声波在血流中的传播速度。接着反射超声波的红细胞又转变角色成为频率为f 的信号发射源，则处于静止状态的探测器接收到的信号频率为

$$f_{\mathrm{R}}=\frac{u}{u+v_O}f=\frac{u-v_O}{u+v_O}f_o$$

由测量到的频率移动可以得到血流的速度v_o。实际上血液中的红细胞的速度并不相同，相应地测量到的反射信号又一个频率分布，由此可以获得速度分布的信息。◀

除了波源和观察者相对于介质运动的情况，另一种特殊的情形是介质相对于波源和观察者所在的惯性系的运动，例如刮风气候中，接受由静止波源发出的信号。读者可以自己进行相应的分析(见习题 7-4-8)。

真空中光的多普勒效应

对于在真空中传播的光而言，由于没有介质作为参照物，相对性原理排除了通过测量多普勒频率移动来区分究竟是光源还是光接收器在运动的可能性，测量到的频率移动只能给出两者的相对运动速度的信息。

为简单起见，这里仅考虑光源$\boldsymbol{S}'$以速度u相对于观测者$\boldsymbol{S}$的运动的情况。读者可以自己分析观测者相对于光源运动时光信号的上述多普勒频移(习题7-4-11)。由于光源运动导致光源前后的观测者所接收到的光波波长收缩或拉长，由此导致的光波频率发生相应的改变为

$$\nu=\frac{u}{\lambda}=\frac{u}{u-v_s}v_o=\left(1-\frac{v_s}{u}\right)^{-1}v_o \tag{7-4-35}$$

需要注意的是，上式中的f'并不是光源所发射的光波的固有频率。因为观测者接收到两个相继到达的波峰的时间间隔T' 相对于光波的固有周期T_o相差一个洛仑兹因子：$T'=\gamma_u T_o$，因而观测者$\boldsymbol{S}$ 接收到的频率f'和固有频率f_o的关系为

$$f'=\frac{1}{T'}=\frac{1}{\gamma_u T_o}=\sqrt{1-u^2/c^2}\,f_o \tag{7-4-36}$$

由此得到观测者测到的光信号频率与固有频率之间的关系为

$$f=\frac{\sqrt{1-u^2/c^2}}{1\pm u/c}f_o=\sqrt{\frac{1\mp u/c}{1\pm u/c}}f_o \tag{7-4-37}$$

其中当光源接近观测者时频率升高，而当光源离开观测者时频率降低。

不同于经典波，光波还具有横向多普勒效应。这一点类似于在不同惯性系中的速度变换方程：在伽利略变换中不存在横向变化，但在洛仑兹变换中就存在横向变化，因为速度描述的是单位时间内物体的位移，而在洛仑兹变换下，运动参照系的时间膨胀是与惯性系的运动方向无关的。因此，即使两个惯性系在横向没有相对位移，但测量到的相应速度分量并不相同。同样，就多普勒效应而言，观测到的运动波源的振动周期发生了洛仑兹膨胀，因此即使波长并没有发生横向压缩，但静止观测者在垂直方向接收到的频率f还是不同于光波的原来频率f_0

$$f=\sqrt{1-u^2/c^2}\,f_0 \tag{7-4-38}$$

这就是所谓的光波的**横向多普勒效应**。

多普勒效应的应用

(1) **穆斯堡尔效应：**一个原子核由激发态跃迁到基态时发射一个γ光子。该光子遇到另一个相同原子核时却不能被共振吸收，这是因为处在自由状态的原子核放射光子时受到反冲，相对于原子核没有反冲的情形光子的能量降低了。而吸收光子的原子核光子同样由于反冲效应，所需要吸收的光子能量相对于理想情况还应有所增加。

1957年底，德国物理学家穆斯堡尔设想将发射和吸收光子的原子核置于固体晶格中，这样原子核发射光子时发生反冲产生的反冲动量就分配到整个晶体。而晶体质量远大于单个原子核质量，则发射光子导致的反冲能量可以忽略，从而实现原子核辐射的无反冲共振吸收。穆斯堡尔再实验中用^{191}Os晶体作γ射线放射源，用^{191}Ir晶体作吸收体。^{191}Os经过β衰变成为^{191}Ir的激发态，^{191}Ir的激发态可以发出能量为129 keV的γ射线，被吸收体吸收。

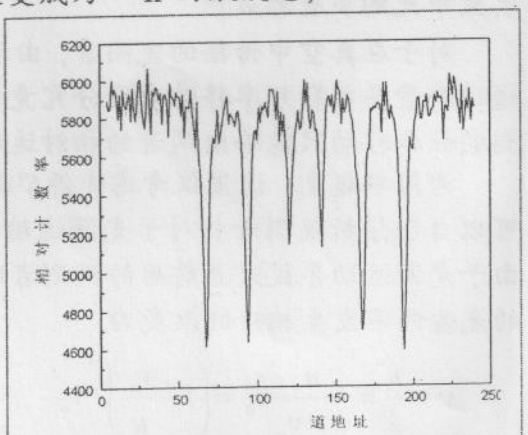


图7-4-7 ^{57}Fe的穆斯堡尔谱呈现6个吸收峰。横轴是道地址，代表相对速度。纵轴是相对计数率。

穆斯堡尔效应对环境的依赖性很高。细微的环境条件差异会对穆斯堡尔效应产生显著的影响。大部分元素只能在低温下才能实现穆斯堡尔效应，为减少热运动对结果的影响，放射源和吸收源都冷却到88K。在室温下只有^{57}Fe、^{119}Sn、^{151}Eu三种同位素能够实现穆斯堡尔效应。其中^{57}Fe的14.4 keV跃迁是最常用、也是研究最多的谱线。为减少环境温度的影响，实验中令γ射线辐射源和吸收体之间具有一定的相对速度v(例如将放射源安装在一个转盘上)，按照多普勒原理，可通过调整v的大小实现对γ射线频率的微调，使其达到共振吸收。透射率与相对速度之间的变化曲线叫做**穆斯堡尔谱**(见图7-4-7)。应用穆斯堡尔谱可以观察到原子核能级的移动和分裂，原子的价态和对称性等方面的信息。

穆斯堡尔谱的宽度很窄，因而分辨本领极高。如^{57}Fe的14.4 keV跃迁，谱宽度与γ射线的能量比为$\Delta E/E\sim10^{-13}$，^{67}Zn的93.3 keV跃迁的$\Delta E/E\sim10^{-15}$，^{107}Ag的93 keV跃迁$\Delta E/E\sim10^{-22}$。因此，穆斯堡尔效应在各种精密频差测量中得到了

广泛的应用。

天文观测中的红移现象，哈勃定律

天文观测中最基本的测量是天体之间的距离和相对运动速度，其中大多数测量都是间接测量。而多普勒效应是其中最重要的测量方法之一。

1848 年，法国物理学家斐索首先发现恒星谱线位置的多普勒红移，1868 年，英国天文学家威廉·哈金斯首次测出了恒星相对于地球的运动速度。Vesto Slipher 在 1912 年及以后的观测中发现，绝大多数的螺旋星云(星系)都有显著的红移。1929 年，美国天文学家哈勃[30]通过研究 24 个星系的谱线发现来自这些遥远星系的光线的红移和它们与地球的距离有关，由于多普勒红移量和光源的退行速率 v 有关，哈勃最终得出，星系的退行速率 v 与的距离 D 成正比：

$$v = H_0 D \tag{7-4-39}$$

这一关系后被称为**哈勃定律**。其中的比例常数 H_0 称为**哈勃常数**，其最佳数值是在 2003 年使用人造卫星威尔金森微波各向异性探测器(WMAP)测得的，数值为 $71\pm4\ \mathrm{km\cdot s^{-1}\cdot Mpc^{-1}}$(百万秒差距)。哈勃定律现在常被用来推算遥远星系的距离。

有关哈勃定律有两点需要说明，一是实验中确定一个光源的光波红移量需要依赖于光源光谱中存在的某些特征，如吸收线、发射线、或光密度分布，并有相似特征的标准光谱作比较(图 7-4-8)。例如，氢原子有确定的特征光谱，比较测量得到的氢光源的发射光谱和氢原子的标准谱线就能得到的红移。二是从红移量可以得知星系远离我们的速度，但它们与我们的距离则是基于这样的假设：星系的总光量大致相同。亮一点的星系离我们较近，暗的则较远，但不同种类和大小的星系有不同的总光量，尤其是数十亿光年外的星系的总光量和近距离星系的总光量有很大分别，所以要观测这类星系的后退速度与距离比例有很大的困难。科学家近来试图通过研究一种称为 Ia 型的超新星来解决这个问题：设想有两颗星互相绕着对方旋转，其中一颗较小的星(伴星)的物质会被慢慢吸引到另一颗较大的星上，这颗星的质量累积到某个临界值时便会发生超新星爆发。所有 Ia 型超新星爆发时的亮度都会达到差不多相同的数值，这样就能通过在地球上观察这类星体爆发的亮度推断它所处的星系与地球的距离。

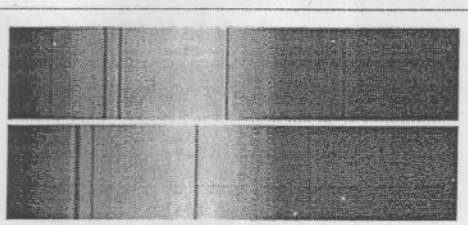


图 7-4-8 上图是遥远星系的可见光谱，与下图的太阳光谱比较可以得到谱线的红移。

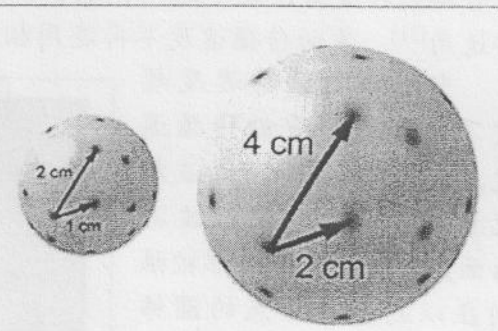


图 7-4-9 星系在膨胀的宇宙中类似于气球表面上的红色标记。当气球膨胀时，两个星系远离对方的速度与它们之间的距离成正比。

哈勃的观测表明所有的星系都在远离我们而去，这意味着宇宙正在膨胀。

[30] Edwin Hubble

在膨胀的宇宙中星系类似于图 7-4-9 中气球表面上的红点。气球膨胀时，两个星系离开对方的速度与它们之间的距离成正比。宇宙膨胀不是等速的，因而不同时期的哈勃常数有所不同。通过观测遥远星系的后退速度与距离的比例可以计算从前宇宙膨胀速度。

利用这种变星的特性进行计算的结果表明，退行速度与距离的比值较小，这意味着以往宇宙膨胀的速度小于现在宇宙膨胀的速度(图 7-4-10)。目前科学界一直在尝试用不同的理论解释这一现象，如果不能提出令人满意的解释，最后不得不承认的结果就是宇宙真的正在加速膨胀。如果真是如此，我们又面临着是什么是导致宇宙加速膨胀的原因这样的问题。这些问题目前都是开放性的，宇宙是否真的加速膨胀有待于更先进的技术，更精确的测量，观察更遥远的星体加以验证。

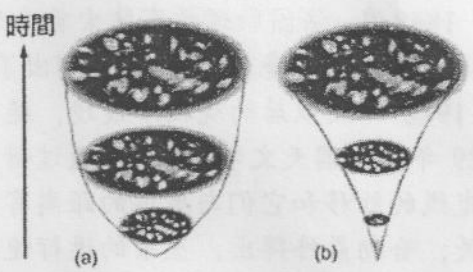


图 7-4-10 宇宙膨胀速度随时间变化的示意图。(a)宇宙膨胀速度随时间减慢,(b)宇宙加速膨胀。

宇宙起源于一次大爆炸这种想法为哈勃定律提供了一个简单明晰的图像，各个星系从大爆炸获得的速度不同，那些在大爆炸后以最快的速度飞离的星系现在离我们最远。一个有趣的问题是，当退行速度接近光速时会发生什么，这似乎也给我们能够观测到的宇宙尺度设置了一个极限。

冲击波

描述波源运动导致的多普勒频率移动的表达式(7-4-31)所适用的条件为波源运动速度小于波在介质中传播的速度，否则由表达式得到的波长或频率都是负值而没有意义。从图像上很容易看到，当波源的运动速度接近于波在介质中的传播速度时，波源前面的向前传播的波的波长趋近于零，因而大量的振动能量都在波源前端的窄小区域堆积起来，导致该区域中波的振幅变得非常大。在这种情况下，多普勒频移公式已经不适用，因为介质中的恢复力和质点位移之间的线性关系不再适用[31]。波的传播速度不再能用相速度来描述，波形也会随时间发生变化。

当波源的运动速度超过了波在介质中的传播速度，波源超越了所有的波前，大量的波面都被抛在波源后面，介质中的振动都被限制在以波源为顶点的圆锥区域以内(见图7-4-8)。该圆锥通常称为**马赫锥**，其半顶角满足关系式

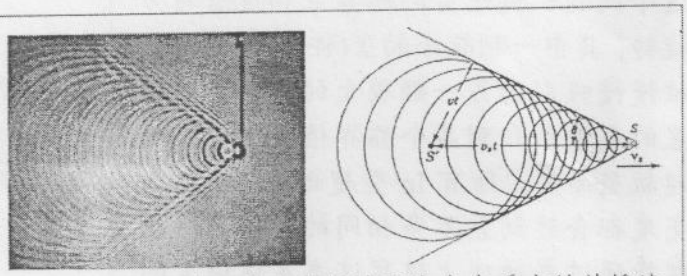


图 7-4-8 波源的运动速度超过了波在介质中的传播速度时介质中的振动都被限制在以波源为顶点的圆锥(马赫锥)区域以内。

[31] The linearity only holds for small perturbation in the media.

$$\sin\theta = \frac{ut}{v_s t} = \frac{u}{v_s} \tag{7-4-39}$$

该数值的倒数 v_S/u 称为**马赫数**。图 7-4-9 中分别为子弹的运动速度达到马赫数为 1.01 和 2.45 时在空气中形成的冲击波波形。在图(a)中波前几乎紧贴到子弹，而在情形(b)中显示出了清晰的马赫锥形状。

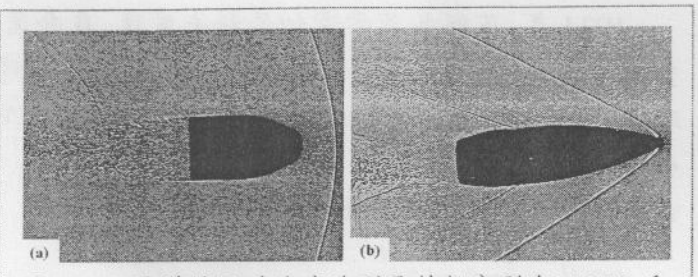


图 7-4-9 子弹的运动速度达到马赫数分别为(a)1.01 和(b)2.45 时在空气中形成的冲击波波形。

当飞机在空中高速飞行时，它对空气产生的扰动以声速向四周传播。当飞机的速度接近声速时，飞机产生的声波波前在飞机前面逐渐密集并堆积起来。这种高强度的压力波前所导致的冲击波被称为**音爆**，也就是当飞机掠过我们头顶时所听到的隆隆声。超音速飞机通常会产生两个音爆，一个在飞机鼻翼处，另一个在飞机尾处，由此导致两次重击。其结果是压力波前具有很高的强度，飞机上很小的振动都将环绕着飞机，从而导致共振的危险。当飞行速度接近音速时，其飞行性能会出现剧烈的变化，喷气机飞行员会感受到存在一种明显的"障碍"，称为"**音障**"[32]。这种障碍是由很强的压强波前形成的，在该压强波前区域飞行会感受到强烈的不规则和紊乱的振动。飞机操纵困难，飞行速度也上不去了。飞机以声速飞行时是很危险的，飞机在飞行接近声速时必须尽快越过这一障碍进入超音速飞行，进入超音速飞行后，声波马赫锥就取代了通常的冲击波。

波源速度超越声速时，在波源前面不能接受到任何声波的信息，只能在波源经过以后可以接收到马赫锥的波前，由于以**马赫角**扩散的线状波前是波源在行进过程中发射出的大量圆形波前的叠加，由此构成的波与简谐波有很大的不同。它呈现出介质中的物质运动突变的传播。这种波因为它的强度大以及突然性而被称为**冲击波**。这也是当喷气机以超音速从我们头顶掠过时能听到犹如爆炸声一般震响的原因。图 7-4-10 为海军少尉 John Gay 在 1999 年 7 月 7 日在航空母舰上所拍摄的海军少尉 Ron Candiloro 驾驶 F/A-18 大黄蜂战机正好穿越音障时所拍摄到的相片，这时水蒸气，密度和温度都恰到好处。图中的那些雾气是飞机突破音障时，空气中的水蒸气受到产生的激波影响，瞬间凝结形成的。

图 7-4-10

类似于物体在介质中的运动速度接近音速时会产生音障，带电粒子在介质中的运动速度接近光在介质中的传播速度 $v = c/n$ 时会产生"**光障**"。而当粒子在介质

[32] Chuck Yeager was the first person to break the sound barrier when he flew faster than the speed of sound in the X-1 rocket-powered aircraft on October 14, 1947.

中运动的速度超过光在介质中的传播速度时，这些高速运动带电粒子发出的辐射，称为切伦科夫辐射[33]。

1934 年，苏联物理学家切伦科夫发现，在介质中作匀速直线运动的带电粒子，当粒子速度大于介质中的光速时会产生一种特殊辐射，这是一种电磁“冲击波”现象，辐射具有明显的方向性和强偏振等特点。它所辐射的电磁波将集中在粒子后方一个圆锥形区域中。粒子位于圆锥的顶点。箭头所指即为辐射传播的方向，它与粒子运动方向之间的夹角 θ_c 满足：

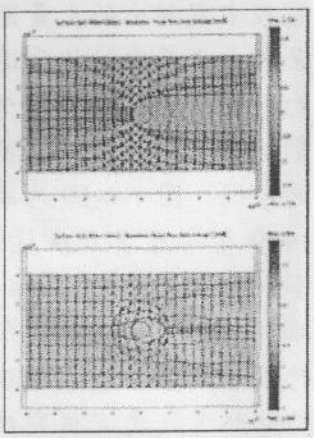
$$\cos\theta_c = \frac{ut}{vt} = \frac{u}{v} = \frac{c}{nv} \tag{7-4-40}$$

其中 u 为辐射在介质中传播的速率，v 为荷电粒子速率，c 为真空中光速，n 为介质折射率。切伦科夫角 θ_c 可用于测量带电粒子的运动速度。

不同于带电粒子加速时的辐射，切伦科夫辐射不是单个粒子的辐射效应，而是运动带电粒子与介质内束缚电荷和诱导电流所产生的集体效应，这种在介质中的电磁冲击波是粒子在其运动轨迹的各点所辐射的波相互干涉的结果。切伦科夫辐射是强偏振辐射，其电矢量在传播方向与粒子运动方向组成的平面内。可用于制成探测高速粒子的切伦科夫计数器。它具有计数率高、分辨时间短、能避免低速粒子干扰、准确测定粒子运动速度等优点。切伦科夫计数器在核物理和粒子物理发展史上起过重要作用，1955 年利用它发现反质子，是高能物理和宇宙线实验中广泛应用的重要计数器之一。

7-4-4 波在不同介质界面处中的传播行为和特性

呈现出不同波面形状的行波实际上是不同初始条件和边界条件下的波动方程特解。波在一个无限大的均匀介质中的传播是一个平凡的问题。人们对波动的兴趣和相关的问题在于介质的各种不均匀性对波的传播的影响，由此可以根据这种影响来理解介质的结构和各种相关特性[34]，另一方面，通过对介质材料，尺度、形状，边界等的设计，可以控制波的传播行为，图 7-4-10 所示为是一例非常有趣的设计：在一种中空圆筒的筒壁上镶嵌着许多特定尺寸的纳米尖刺。当光入射到圆筒上既不会发生反射，也不会向四周散射，而是从其周围绕过去，就像是水流绕过桥墩一样。这样在光路的下游就无法看到圆筒。如果用该圆筒罩住某一物体，圆筒以及所罩住的物体都将变得无法看到。这样的光学器件称为隐形材料[35]。



Optical Cloaking device

波在介质中的传播行为和特性是由波动方程和介质的边界决定的。对于无限大尺度的均匀介质而言，波动方程的解是一个平凡的问题。但实际遇到的情况要

---

[33] Pavel Cerenkov Russian physicist

[34] 对外部扰动的响应是揭示物质性质的最重要，也是唯一的手段，实际上也是由此定义物理量的手段

[35] 可见光可以视为各种颜色的‘混合物’。仪器只能对一种颜色的光线产生作用。我们现在还不清楚如何使物体在任何颜色的光线作用下都达到隐身的效果。

复杂得多，一是因为介质的有限尺度和均匀性会涉及到与各种界面有关的衔接问题，如反射，折射等，另一方面，介质中的物质结构导致波能否有效传播，介质中的阻尼，介质对频率的不同响应特性等。分析介质对波的响应，主要是通过波的变化来研究不同物质的特性，因此，无论在理论或实验中都采用最简单的简谐行波 $u(\mathbf{r},t) = Ae^{-i(\omega t-\mathbf{k}\cdot\mathbf{r})}$ 入射到介质作为入射波以简化分析。这里我们只讨论一维非均匀介质中行波的传播行为。

波的反射与透射

首先考虑介质中的一种最简单的不均匀性：两个无限大均匀介质的界面[36]，而对于垂直与界面入射的平面波可以简化为一维问题，甚至不用特别区分横波和纵波的差异。为了使问题本身及其分析更为直观，我们具体讨论在两根粗细不同的均匀弦A和B拼接而成的组合弦(见图7-4-12))上的简谐行波

$$u(x,t) = Ae^{-i(\omega t-kx)}$$

在衔接点的反射和透射行为。

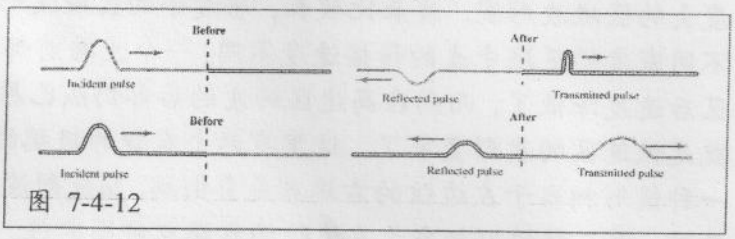


图 7-4-12

首先分析弹性弦上质元受力情况，当波前向右到的达衔接点时，A弦的端点向B弦的端点施加向上的作用力，从而使波前能越过衔接点向右进入B弦的区域。同时A弦的端点受到大小相同，但方向相反的力，由于两边介质存在明显不连续的力学特性(质量密度和弹性张力)，这种力就会在A弦上形成向左传播的行波，称为反射波。设左面弦上的入射行波和反射波以及进入右弦的透射波分别为

$$\begin{aligned} u_i(x,t) &= \widetilde{A}_i e^{-i(\omega t-k_1 x)} \\ u_r(x,t) &= \widetilde{A}_r e^{-i(\omega t+k_1 x)} \\ u_t(x,t) &= \widetilde{A}_t e^{-i(\omega t-k_2 x)} \end{aligned} \tag{7-4-41}$$

其中各个波的振幅都是包含了初相位的复数，即复振幅。

弦的连续性以及不出现突变性的纠结在数学上表示为弦上各质元相对平衡位置的位移的连续性及其空间导数的连续性。将弦的衔接点位置设为坐标系原点，方程的解在衔接点满足以下衔接条件：

$$\begin{aligned} & u_i(0,t)+u_r(0,t) = u_t(0,t) \\ & \frac{\partial u_i(0,t)}{\partial x}+\frac{\partial u_r(0,t)}{\partial x} = \frac{\partial u_t(0,t)}{\partial x} \end{aligned} \tag{7-4-42}$$

将(7-4-30)式中两边弦上的波动代入上述的衔接条件可以得到

$$\widetilde{A}_i + \widetilde{A}_r = \widetilde{A}_t$$

$$k_1\widetilde{A}_i - k_1\widetilde{A}_r = k_2\widetilde{A}_t$$

相应地可以得到波振幅的**反射比**和**透射比**：

[36] Any material in physical world has finite size, and so the behavior of the wave at the boundaries or interfaces of different media is of particular significance.

$$r \equiv \frac{\tilde{A}_r}{\tilde{A}_i} = \frac{k_1 - k_2}{k_1 + k_2} \qquad t \equiv \frac{\tilde{A}_t}{\tilde{A}_i} = \frac{2k_1}{k_1 + k_2} \tag{7-4-43}$$

由于波数与介质惯性的平方根成正比：$k = \omega/v = \omega\sqrt{\mu/T}$，波振幅的反射比和透射比可以表示为弦的密度的函数

$$r = \frac{\sqrt{\mu_1} - \sqrt{\mu_2}}{\sqrt{\mu_1} + \sqrt{\mu_2}} \quad t = \frac{2\sqrt{\mu_1}}{\sqrt{\mu_1} + \sqrt{\mu_2}} \tag{7-4-44}$$

若入射区的弦密度大于透射区的弦密度：$\mu_1 > \mu_2$，r 取正值，说明反射波与入射波是同相的，而当入射区的弦密度小于透射区的弦密度：$\mu_1 < \mu_2$，反射比取负值，反射波与入射波的相位相差π，在波形上相差半个波长，称为**半波损失**。由于透射比是恒正的，表明透射波与入射波总是同相的。

虽然在衔接点入射波与透射波的位相相同，但两个区域的波形并不相同：密度大的区域波形宽，波长比较长，密度小的区域波形窄，波长比较短。其原因是不同密度的区域中波的传播速度不同，一个波的前部经衔接点从高速区进入低速区后速度降低了，而尚在高速区的波的后部仍然已原来的速度运动，其最后结果就是低速区的波形变窄了。这里有两个有趣的极端的情形：$\mu_2 = 0$，或$\mu_2 = \infty$. 第一种情形相当于左边弦的右端点是自由端，而反射波和入射波具有相同的相位和振幅，第二种情形相当于左边弦的右端点被固定住，因而不再有透射波，所有的入射能量被反射回A弦。

波的反射强度与透射强度

If we have an incident pulse that is partly reflected and partly transmitted, then the sum of the energies in these two pulses equals the energy of the incident pulse. If there is no transmitted pulse, as is the case when the new string is either infinitely heavy or infinitely light, the energy in the reflected pulse matches the energy in the incoming pulse. Otherwise there is a sharing, and both the transmitted and reflected waves have an amplitude reduced from that of the incident wave.

To summarize: When a wave arrives at an interface between two media. Some of the wave is reflected and some is transmitted. If the wave speed in the medium containing the incident wave is greater than the wave speed in the medium containing the transmitted wave, the reflected wave is inverted but is otherwise similar in shape to the incident wave. The transmitted wave is always right-side up, but its shape will be stretched if the wave speed is higher in the second medium and compressed if the wave speed is lower in the second medium.

$$\langle P_i \rangle = \frac{1}{2} v_1 \mu_1 \omega^2 A_i^2$$

$$\langle P_r \rangle = \frac{1}{2} v_1 \mu_1 \omega^2 A_r^2 = \frac{1}{2} v_1 \mu_1 \omega^2 \left(\frac{k_1 - k_2}{k_1 + k_2} \right)^2 A_i^2$$

$$\langle P_t \rangle = \frac{1}{2} v_2 \mu_2 \omega^2 A_t^2 = \frac{1}{2} v_2 \mu_2 \omega^2 \left(\frac{2k_1}{k_1 + k_2} \right)^2 A_i^2$$

$$R=\frac{\langle P_r\rangle}{\langle P_i\rangle}=\frac{A_r^2}{A_i^2}=\frac{(k_1-k_2)^2}{(k_1+k_2)^2}$$

$$T=\frac{\langle P_t\rangle}{\langle P_i\rangle}=\frac{v_2\mu_2}{v_1\mu_1}\frac{A_t^2}{A_i^2}=\frac{v_2\mu_2}{v_1\mu_1}\frac{4k_1k_1}{(k_1+k_2)^2}$$

$$\langle P_i\rangle=\langle P_r\rangle+\langle P_t\rangle$$

反射系数 R 的表达式(7-4-30)和透射系数 T 的表达式关于入射波数 k_1 和透射波数 k_2 是对称的，这意味着如果粒子从右边入射，反射系数 R 和透射系数 T 的大小不变。这一结果实际是所有波的共同特性，例如，对于光波而言就表现为光在介质中传播的可逆性。

上面所讨论的一维介质中入射波的反射和透射与平面波垂直入射到二维界面的情形相当。对于一般的平面波非垂直入射到界面的情形，反射和透射除了涉及振幅和相位，还涉及到方向。对于横波而言还需要分别讨论波的不同偏振方向。相关的讨论见下一章的 8-1 节。

7-4-5 波速度和色散

介质中的简谐波是振动的传播，因而波的就是振源的振动频率，与介质的性质无关，而波数($k=\omega/v$，或波长 $\lambda=2\pi v/\omega$)依赖于速度，除了与振源的振动频率有关，也与介质有关。对于振动周期固定的振源，它在任何线性介质中引起波的频率不变，但由于波在不同介质中的传播速度不同，该振源在不同介质中引起的波的波长是不同的。另外，同一种介质对不同频率的波动一般也具有不同的传播速度，因而当含有几种频率成分的混合波进入同一种介质中传播时，不同频率的简谐波因速度不同而相互分离，这种现象称为**色散**。

仅仅从均匀连续介质的波动方程(7-2-19)来看，所有的简谐波都以相同的速度

$$v=\frac{\omega}{k}=\sqrt{\frac{T}{\mu}} \tag{7-4-50a}$$

传播，因而各种形态的波都可以无形变的传播。但事实上，所有的介质都是不均匀的，但不均匀的程度由所不同，对波而言，任何不均匀的尺度只要远小于在其中传播的各简谐波分量的波长，就认为该介质是均匀的，该波在介质中传播时不会发生畸变。实际介质中的不均匀性可以是结构的离散性，物体的尺度，边界等因素造成。例如，在实际晶体中沿一个方向传播的简谐波的速度与频率的关系可以用式(7-4-3)描述

$$v=\frac{\omega}{k}=\frac{2\omega_o}{k}\sin\frac{ka}{2} \tag{7-4-50b}$$

波长大于格点间距时介质不表现出色散行为，而对于波长大于格点间距的弹性波就会表现出色散特性。另一个例子是水面波，其波长远大于水分子尺度的不均匀性，初看起来似乎不具有色散特性，但水面波可以深入到水面下一定的深度，即水面波的振幅沿与表面垂直的方向衰减[37]，其透入表面的深度为一个波长的量级。在深度远小于波长的浅水区域，水面波的振幅在整个深度上几乎是均匀的(满

[37] 重力波的衰减主要原因是流体与水底固体的摩擦(水很深时可忽略)和流体内部的黏滞效应。

足长波近似条件)，因而没有色散。而水面波在深水区就具有明显的不均匀性，从而表现出色散行为。

简谐波在介质中的传播速度与频率或波长的函数关系称为介质的**色散关系**。由于(7-4-50)式所给出的波速和频率及波数的关系，我们通常用函数 $\omega(k)$描述介质的色散关系。例如 $\omega(k)$为线性函数表示介质是无色散的，否则介质在相应的频率范围中就存在色散。

由于各种复杂形状的波都能表示为不同波长或频率的简谐波的叠加，不同波长或频率的简谐波在无色散介质中以相同的速度传播意味着这些波在介质中的传播过程中不会发生形变，事实上，均匀弹性弦的波动方程中由于 v 为常数，因而存在的行波解可以具有各种形态 $u(x,t)=f(x\pm vt)$也说明了这一点。

由于一个波在色散介质中传播时会发生变形，由此也就带来一个很基本的问题：应怎样定义这个波的传播速度。就观测而言，我们可以选择一个时空点(x,t)上的质点(或质量元)相对于其平衡点的位移(形变)，然后不再关注该质点(或质量元)，而是跟踪该位移大小不变的位置在空间连续移动。例如对于简谐波 $u(x,t)=A\sin\varphi(x,t)=A\sin(kx-\omega t+\varphi_o)$，其上的每一位移量都是由相位 $\varphi(x,t)$决定的，则位移大小不变的位置在空间的移动速度由 $d\varphi/dt=0$ 给出

$$v_p\equiv\frac{dx}{dt}=\frac{\omega}{k}$$

称为相速度。对于一般的波函数所描述的波形可以表示为各种频率的简谐波的叠加：

$$u(x,t)=\int_{-\infty}^{+\infty}d\omega g(\omega)e^{i(kx-\omega t)} \tag{7-4-51}$$

其中 $g(\omega)$是波包的频谱。波的传播特性取决于**色散关系** $\omega(k)$或 $k(\omega)$的函数形式，若介质的色散关系在波包中所含简谐分量的频率区间内是线性的

$$k=k_o+\frac{dk}{d\omega}|_{\omega_o}(\omega-\omega_o)=k_o+k'(\omega_o)\cdot(\omega-\omega_o)$$

则介质在该频率范围内不存在色散，

$$\begin{aligned}kx-\omega t&=k_ox+k'(\omega_o)\cdot(\omega-\omega_o)x-\omega_o t-(\omega-\omega_o)t\\&=k_ox-\omega_o t+(\omega-\omega_o)[k'(\omega_o)\cdot x-t]\end{aligned}$$

其中 $k_o=k(\omega_o)$，将上式代入波包的表达式(7-4-51)

$$u(x,t)=e^{i(k_ox-\omega_o t)}\int d\omega g(\omega)\exp\{i[k'(\omega_o)\cdot x-t](\omega-\omega_o)\}$$

它可以理解为一个受调制的载波，积分式前的因子表示频率为 ω_o 的载波，载波速度为其相速度为 $v_p=\omega_o/k_o$。积分描述的是一定频率范围内的调制波，由于波形对时空的关联是以 $k'(\omega_o)\cdot x-t$ 的集约形式，即行波形式呈现的，其波形(即调制波包络线)上某一点(例如取波峰)的传播速度称为调制波的群速度，由固定相位点$[k'(\omega_o)\cdot x-t](\omega-\omega_o)=\text{const.}$确定

$$v_g=\frac{dx}{dt}=[k'(\omega_o)]^{-1}=\frac{d\omega}{dk}|_{\omega_o}$$

非色散频率范围内，$d\omega/dk=\text{const.}$，因此群速度和相速度相等。

介质在色散频率范围内，群速度和相速度与频率的关系通常都比较复杂。事实上对于在传播过程中发生剧烈畸变的波形很难给出其速度的普适定义。一般而