

力学讲义

陈曦

August 15, 2020

目录

1	物理学构筑了描述世界的模型	1
2	惯性参考系	5
3	力与质量	15
4	不变性	19
5	矢量算符	23
5.1	偏导数与全微分	23
5.2	方向导数、梯度	24
5.3	散度	25
6	场、势和能量	29
6.1	场	29
6.2	保守场	30
6.3	旋度	32
6.4	势	35
6.5	保守系统	38
6.6	一个更一般的框架	39
6.7	非保守场	40
7	谐振子	41
7.1	无阻尼谐振子	41
7.2	阻尼简谐运动	42
7.3	受迫振动	43
7.4	共振	47

8	有心力运动	49
8.1	两体问题约化为单体问题	49
8.2	有心力运动的一般性质	50
8.3	有效势	50
8.4	闭合轨道	52
8.5	行星运动	52
8.6	变轨	57
9	质点系动力学	59
9.1	动量	59
9.2	角动量	60
9.3	多体问题	62
9.4	定轴转动	62
9.5	一般的刚体运动	66
9.6	作为矢量的角速度	68
9.7	回转仪 (Gyroscope)	70
9.8	地球的进动	73
10	非惯性系	75
10.1	匀加速平动非惯性系	75
10.2	定轴匀速转动非惯性系	75
11	线性系统	81
11.1	线性系统	81
11.2	受迫线性系统	82
11.3	Fourier分析	83
12	波动	85
12.1	运动模式	85

12.2 一维环的运动模式	89
12.3 色散关系	92
12.4 长波极限	94
12.5 波动中的能量	95
12.6 横波、纵波	98
12.7 波的一般形式	98
13 狭义相对论的基本假设	101
13.1 假设一	101
13.2 假设二	105
14 相互作用传播速度的上限	109
14.1 光速各向同性实验	109
14.2 光速与光源运动无关实验	110
14.3 光速与频率无关实验	110
15 长度收缩与时间膨胀	113
16 相对论力学	117
17 四矢量和相对论协变性	121

1 物理学构筑了描述世界的模型

我们首先来理解物理学是怎样的一种知识。在探索未知世界的进程中，人类从未止步于利用感觉器官和各种工具来进行观测，而是更进一步依据这些观测结果为世界构筑各种模型。这种异乎寻常的能力似乎仅为具有高级思维能力的人类大脑所特有，我们至今仍然对其机理知之甚少。虽然模型植根于经验，但确是人类大脑的创造。模型是大脑对外部世界的反映但并不等同于外部世界。人们只是希望纷繁复杂的世界能够按照自己构筑的模型所预测的那样运转。在形式上，模型表述为严格的因果关系，这也就意味着我们期望各种可观测的现象都具有可重复性，即相同的条件会得到相同的结果。这种可重复性就是所谓“客观实在”的本质意义和一切科学活动的基础，虽然我们并不知道其根源是什么，甚至不知道是否真的有所谓“可重复性”。物理学就开始于对“客观实在”的经验性认识。虽然原则上这个模型可以是一个记录了万事万物全部因果信息的数据库，但人们更倾向于追求一种简洁的框架。如牛顿体系那样，这个框架只包括为数不多的陈述。这些陈述试图将这个世界上几乎是无穷多的自由度约化为极少的几个约束关系（比如几个公式），并认为这些关系是普适的、基本的和精确的。

必须指出的是，物理学家利用各种实验工具、理论工具和想像力创造出来的模型是用来描述这个世界，而不是解释世界或者发现存在的意义。物理学家所讨论的“客观实在”仅在于“可重复性”，即同样的实验都会得到同样的结果。物理学家以实验观测为基础利用想像力来发明概念和图像去描述“客观实在”。我们没有任何理由认为必定会存在一种包罗万象的简洁理论或这个物质世界的本身一定存在所谓“规律”。即使有所谓“规律”，我们也无法知道我们所写下的模型是否就是那个“规律”。而我们这个世界最为神奇、最为不可思议的地方就是人们迄今为止所作的努力、取得的成功似乎暗示着这个世界存在着一般的秩序和普遍的原则。至于人们是否能够找到一种“终极理论”描述宇宙中的一切现象，则是一个没有答案的问题。

先让我们回到牛顿的原始表述。牛顿的《自然哲学的数学原理》的出版是人类历史上的一个重要时刻。牛顿不仅仅是解决了一系列物理问题，更重要的是开创了一个思想体系、创立了一个优美的框架。他相信可以用普适的、简洁的定律描述自然界。牛顿在他书中给出了著名的三个定律：

定律一 每个物体都保持其静止、或匀速直线运动的状态，除非有外力作用于它迫使它改变那个状态。

定律二 运动的变化正比于外力，变化的方向沿外力作用的直线方向。

定律三 每一种作用都有一个相等的反作用；或者，两个物体间的相互作用总是相等的，而且指向相反。

牛顿创立的体系在今天看来已经变得如此的不言而喻，以至于会使我们低估做出这一贡献的难度。事实上，这就是科学发展过程自身的特点。科学的进步来源于新的工具、新的发现和新的思想的出现。新的工具拓展人的观测能力、发现新的现象，是最常引发突破发生的因素。完成一个科学方向上的突

破之前的尝试往往是复杂、凌乱和胶着的，交织着实验和理论发现、概念和技术创新。其中绝大多数的工作并不会导致重要结果，从而不再被人关注。而突破发生之后出现的新的理解往往是明了而自然的。原本难以理解的事实和概念变得各安其位，似乎一切都如此之自然、如此的显而易见和理所当然，以至于后人很难想象当初为此所作出的智力上超乎寻常的努力。对这一点的认识会使人们对他们正在进行的以发现为目的的智力活动有更深入的理解。

在哥白尼之前，托勒密的本轮—均轮体系取得了很大的成就。但这一体系过于复杂，因此人们期待着一种更简便的模型。哥白尼尝试着将太阳放在中心并作为参考点来描述行星运动并最终取得了成功。哥白尼不是第一个提出日心说的人，古希腊的天文学家也曾做过类似的尝试，不过却因为缺乏证据而未被接受。到了哥白尼的时代，科学技术的进步和观测数据的积累达到了一个新的高度，科学界“已经准备好”接受一次观念的革命。哥白尼的工作将日心说高度定量化，在之后的一百多年的时间里被广泛接受并引发物理学的革命。这恐怕是哥白尼始料不及的。

事实上，在哥白尼的时代，日心说本身并没有明显地优于托勒密体系。尽管哥白尼让地球不再处于模型的中心，但他在处理行星运动的细节上与托勒密是一致的。为了描述很多不规则的运动，哥白尼也需要引入很多本轮。哥白尼的伟大贡献是扩大了人们的视野，使人们发现日心说也是一种可能的选择，为此后伽利略等人的进一步工作打好了基础。

从哥白尼发表《天球运行论》（1543年）到牛顿出版《自然哲学的数学原理》（1687年）的这段时间里，为这场物理学革命做出最重要贡献的有第谷、开普勒和伽利略。第谷做出的一系列精密测量是此后建立严密理论的基础。第谷去世后，接替他的位置的是其助手开普勒。开普勒是哥白尼的坚定支持者，并具有非凡的数学能力。

火星的轨道一直是天文学上空的一朵乌云。无论是托勒密还是哥白尼都没能对其轨道的不规则性做出满意的描述。开普勒在第谷的数据基础上利用了五年的时间尝试了各种圆周的组合后才偶然发现椭圆才是真正的答案。这期间他所构造的圆周组合模型中的许多都要比托勒密和哥白尼的更精确，但是仍然与第谷的精密测量结果有差距。只有椭圆才即简单又精确地给出了行星轨道。这就是开普勒第一定律：行星沿椭圆轨道绕太阳运行，太阳位于椭圆的一个焦点上。只有到这时，哥白尼体系的威力才真正显示出来。由于不再有匀速的圆周运动，开普勒需要为行星寻找新的运动规律。这一次，他发现了开普勒第二定律：对于任何行星，它的径矢在相等的时间内扫过相等的面积。没人知道开普勒是如何发现这个规律的。更为天才的是开普勒的第三定律：行星绕太阳运动轨道半长轴的立方与周期的平方成正比。它建立了不同行星之间的关系，是一个从来没有被觉察到的规律。不知道开普勒是如何想起来去研究不同行星间的关系，更不知道他为什么会关注半长轴与周期的关系，似乎有一种神秘的直觉在引导，完全是神来之笔。有了这些发现以及伽利略的望远镜观测后，哥白尼体系才取得了决定性的胜利。

通常，新的发现或者被遗忘，或者被物理学家小心翼翼地纳入到已有的框架之中，而并不一定会导致突破。只有在形成足够的积累之后，思想的嬗变才能产生普遍的影响。通常发生的是，一个产生重要影响的发现出现后，人们

总是可以在早些时候的研究中找到相同或类似的工作。但这些早期的工作由于没有出现在合适的时间或环境中而未带来革命性的影响。新的发现通常是不完善的、粗糙的，甚至发现者本身可能都没有充分意识到它的重要性。发现的过程往往是曲折迂回地在试探中前进，有时甚至出发点都有可能是错误的。但新的发现会启发和激励一些人继续沿着那个方向探索下去。大部分的情况是，这些后续的工作并没有导致有重要意义的结果，从而不再被人关注。只有一少部分会发展起来。这往往要经历多人的努力。我们今天学习到的都是经过前人长期的完善后留下的精华，可能已经不易看出它的原始的模样，也不会有许多人了解这个过程中曾出现过的曲折。这会给正在学习阶段的学生留下一个不正确的印象，似乎物理学从来在逻辑上和方法上都是一个近乎完美的知识体系。完美的事物使人愉悦，但可能会引导学生只关心那些已经成熟的领域。然而不成熟的领域才会有更多的新发现的机会。

经历了千年的努力，科学在牛顿时代达到了一个顶峰。我们在此不再更多地回顾这段历程，而是基于牛顿体系去仔细审视人与世界的关系并进一步揭示物理学知识体系的结构。这种结构并非是一个物理学知识体系最初的样子，而是之后对已有知识进行系统梳理的结果，其目的在于给出一个清晰的逻辑框架。需要特别注意的是，逻辑固然为科学探索所必须，但远远不是全部。梳理后的框架已经略去了发现的过程，并不能反映发现的全貌。例如，实验会更多地以观测的面目出现逻辑体系中，而事实上实验的作用则更多地体现在发现新的现象、新的效应。新大陆的发现更依赖于实践、眼界、勇气、想法、品味、运气和灵感。但由于力学比较直观化，所以在本课程中我们会更侧重于知识的框架。更为重要的是，这样一个结构总是建立在一系列基本假设基础之上，深刻理解这些基本假设才会知道现有知识的边界在哪里、怎样进一步拓展知识的边界。

模型中要包含一系列可观测量，例如时间、空间、质量、速度、电荷、力、电场、磁场、温度等等。可观测量来源于“客观实在”反映到大脑中的映象又是模型与世界的接口。每一个可观测量都对应具体的测量过程并将测量定量化，否则没有意义。模型的另一部分是一系列定律（也称为假设），例如在惯性参考系中适用于宏观、低速、弱引力场条件下的运动方程 $\vec{F} = m\vec{a}$ 。这些定律是人的创造。如果没有对应具体的测量，公式中的变量，如 \vec{F} 、 m 、 \vec{a} ，也就只有数学意义。如果数学方程解出的结果是可以测量的，那么就可以验证方程的解是否能以足够高的精度符合测量结果，从而判定模型的优劣。这里的一个微妙之处在于既然模型是对世界的描述，那么也应同样适用于测量仪器，亦即模型本身赋予了一个仪器测量结果的意义。例如，牛顿力学断定一个自由粒子（无外力作用）呈现匀速直线运动。那么我们就可以把该粒子当作时钟来计时：时间的数值正比于位移。我们也可以采用其他任何物理学所能描述的运动（如行星、刚体、电磁波）来作为时钟。总而言之，我们能测量什么、怎么测量由模型决定。这里的令人困惑之处是既然模型本身是人发明的，那么诸多可测物理量是否也是人的创造？回答是双重的：可测物理量是人对世界中存在的事物的抽象；而一个量怎么测及其大小依赖于人为的规定。事实上，空间和时间并不是独立于物质运动的背景舞台，只是表明一种次序，没有绝对意义。例如，同样是作为时钟的自由粒子，我们可以规定时间是位移的任意函数 $t = t(l)$ ，只是选择 $t \propto l$ 会让与之匹配的运动方程呈现一个简单的形式 $\vec{F} = m\vec{a}$ 。其他的规定一般地会导致复杂的运动规律。对于一个模型的基本要求是始终满足定律之间的自洽性、定律

与测量之间的一致性以及实验与模型预言高度相符，而追求简单性无疑是一个方便的策略。

2 惯性参考系

我们将对牛顿力学进行进一步的梳理。牛顿力学是宏观、低速、弱引力场条件下的唯像有效理论。尽管物理学家在近代已经发展出比牛顿力学更为成功的框架，但牛顿力学是一个非常典型而又比较直观的范例，可以让我们体验物理学的方方面面。牛顿力学的可观测量至少要包括时间、空间、质量、力。我们需要做的是在宏观意义下定义可观测量的测量方法、给出运动规律。

在当前的国际单位制中，时间的单位是秒，定义为“与铯133原子基态两个超精细能级间跃迁对应的辐射周期的9,192,631,770倍”，其基础是量子力学和电磁相互作用。长度的单位是米，定义为“光在真空中1/299,792,458秒的时间间隔内所经路径的长度”，其基础是相对论和电磁相互作用。然而，如果将牛顿力学作为一个完备的模型看待，我们需要采用不依赖于量子力学和相对论的定义。为此，米仍由国际米原器定义。米原器的材料是90%的铂和10%的铱的铂铱合金，保存在巴黎国际计量局，保存条件是0°C，一个大气压。一个基本要求是米原器必须足够“刚性”。然而，在有长度的定义前，我们并没有先验的办法确定刚性的程度。可以如此规避这个循环论证：原则上任何一根杆都可以作为长度标准，采用不同的杆会导致不同的运动方程，我们挑选的是可以使观测数据用一个简单的运动方程足够好地描述的那一个。时间的定义也是类似的情形。有了长度之后，只要规定了时间与位置的对应关系把时间测量转换为位置测量，原则上我们可以选取任何一个运动来定义时间。也可以采用另一种不依赖于长度定义的方案：把时间测量转换为计数。这时需要的是往复的运动。每当运动回到某一个设定点计数 n 就增加1。可以令时间 t 是 n 的函数。我们发现采用地球自转、公转、单摆等等并规定 $t \propto n$ （在此规定下的往复运动称为周期运动），运动方程可以以一种简单的形式足够好地描述各种观测数据。需要注意的是，在没有定义时间以前所谓周期运动是没有意义的。原则上我们当然可以规定 t 是 n 的任意函数。但发现只有 $t \propto n$ 才给出了简单的运动规律形式。另外，这种简单的正比关系还是存在单位这样一个概念的基础。只有采用了正比关系才可以定义一个量的单位。

即使在牛顿之前，人们就已经在使用各种各样的时钟。但这些时钟只是建立在人的感觉基础之上，而不是以任何物理定律作为依据，可称为经验时钟。牛顿力学开始于对行星运动的研究。这时我们不妨取地球自转作为时钟，每一圈定义为一昼。在此基础上建立的运动方程可以很好地描述行星的运动。反过来，这个运动方程还可以告诉我们地球作为一个不错的刚体，其自转确实很接近周期性运动，因而与将其定义为时钟是自洽的。但我们还需要定义更短的时间单位。为此，我们相信描述行星运动的方程也同样适用于地面上的运动，可以发现单摆等也可以是很好的周期运动并可依此定义更短的时间单位。现代的时间单位定义基于量子力学的原理，可以达到极高的精度和稳定性。一般地，一个新的单位定义方法总是来源于新的原理的运用，否则不会对其精度和稳定性有本质的提高。当基于新原理的测量精度已经优于之前的方法时，单位的定义就需要更新了。

我们作为观测者需要一个参考系来放置实验室，开展实验测量。有了一个参考系作为基准，我们可以通过测量长度和角度来建立一个坐标系，写下空间每一点的坐标，进行空间位置的测量。坐标系可以是任意的，可以是直角坐标也可以是球坐标，甚至是不固定在参考系上，如处理刚体问题时可以选择与刚体一同运动的坐标架。只要能够将空间参量化，如何选择坐标系往往是为了方便。为了测量时间，我们还需要把时钟固定在参考系中。于是一个粒子在某一时刻所处的空间位置就可以确定下来。任何测量都应该是局域的。为了将不同位置的测量协调起来，我们需要一种方便的办法对它们进行同步。显然一个由一群彼此相对静止的粒子的组成的刚体是达到这个目的最好的办法之一。自然界提供了很多非常好的物质体系可以看作刚体。

一个理想的情况是，任何一个刚体都应该可以作为参考系，它们之间不应该有任何区别。没有任何先验的理由告诉我们哪个刚体更适合做参考系，选择哪一个完全是观测者的自由。这种没有任何特殊参考系、任何观测者均应等价的要求只有在广义相对论中才会实现。在牛顿力学中，我们选择了一类称为惯性系的特殊参考系作为测量的基准，并有如下假设。

假设IA

存在这样一个参考系，其中任何自由粒子都做匀速直线运动，称为惯性系。

由于之前已经定义了时间和速度，所以匀速运动有了明确意义。自由粒子是指外力作用为零。迄今为止的发现表明粒子间的相互作用有四种类型：万有引力、电磁作用、弱相互作用和强相互作用。如果只考虑原子核尺度以上的物理，就只需关注引力和电磁作用。电磁作用在很大程度上可以通过屏蔽效应来消减。万有引力随距离呈平方反比关系，因此只要与其他粒子足够远就可以把相互作用降到足够低。另外还可以通过设计合适的电场分布来抵消引力。无论是电磁屏蔽还是用电磁场抵消引力，并不是真正没有力作用在粒子上，只是合外力为零，其实现需要有关于电磁场的详细知识。然而描述电磁现象的麦克斯韦方程组同牛顿定律一样需要在惯性系中成立。于是，一方面需要用惯性系中的知识来实现自由粒子，另一方面又需要用自由粒子来印证惯性系。这种情形只是体现了模型的自洽性：为相互作用写下的形式需要能够实现自由粒子。最终的判据来自实验。设想我们指定了一个参考系为惯性系并已知运动定律，这时如果我们实验发现依此实现的自由粒子并非做匀速直线运动，则有两种可能性：此参考系并非惯性系或者运动定律不准确。

惯性系的存在性是牛顿在建立其理论体系时面临的一个问题。他的解决办法是引入绝对空间的概念。绝对空间是万事万物运动的舞台，是一种可以脱离物质而独立存在的实在。它像一个屏幕，物质的运动就像是上演的电影。绝对空间本身就是一个惯性系。然而，实验并没有发现绝对空间存在的迹象。我们更倾向于认为所谓时间和空间需要基于测量。脱离测量而空泛地谈论时间和空间没有意义。而测量则是基于一个参考系。所以脱离具体的参考系空泛地谈论时间和空间也是没有意义的。

假设IB

任何一个与惯性系相对静止或作匀速直线运动的参考系也是惯性系。

该假设是有关惯性系结构的一个陈述。没有任何先验的理由要求我们必须选择惯性系。最理想的情况下，我们希望所有的参考系都互相等价。但在牛顿力学中我们还无法做到这一点，而是选出了一类特殊的参考系，即惯性系，来作为模型成立的前提。由假设IB可知，这类参考系有无穷多个。这里有一个值得注意的地方：相对静止的参考系仍为惯性系容易理解，但相对于一个惯性系作匀速直线运动的参考系也是惯性系这一点似乎并不显而易见。后面相对论的部分将试图回答这个问题。

假设IC

物理定律在惯性系间满足协变性。

所谓协变性即是有相同的形式。没有理由要求在一切观测者看来物理定律都有相同的形式。一般地，每个参考系中的观测者都可以有自己参考系中的物理定律。不同参考系中的定律在形式上可以不一样。这种情形并不意味着不存在物理学了。我们总可以指定某一个参考系及其观测者作为标准，给出物理定律的一个标准形式。其他观测者给出的定律可以通过变换而得到。我们不能期望所有的观测者都能看到相同的物理定律。但是很幸运，经验表明，我们不是只能选择一个参考系及其观测者作为标准，我们可以有一类（无穷多个）叫惯性系的参考系对于物理定律等价。例如，非惯性系中规律的形式会不同。尽管不是全部的参考系，但有一类无穷多个等价的参考系也已经很好了。这是一个折中的方案，它选取了惯性系这一类参考系，其中的观测者能看到同样的物理定律形式。人们尚不能完全理解为什么恰好存在这样一类参考系。协变性对物理定律的数学形式提出了严格的限制。

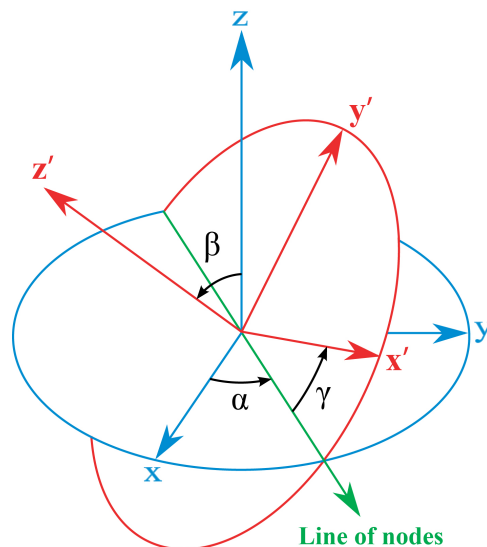


图 2.1: 三维转动

我们先只讨论相对静止的惯性系之间的协变性。简单起见，取直角坐标系对空间参量化，并让两个惯性系S和S'及其坐标架有相对转动并原点重合。一般地，一个规律可能写为如下形式： $\sum_i Q_i = 0$ 。为了保证协变性，每一项 Q_i 在坐标变换下必须保持不变，这样才能使物理规律始终写为 $\sum_i Q_i = 0$ 的形式。这样的量称为标量。

很多时候一个量会随着坐标系的变化而变化。我们有时会发现三维空间中的某三个量之间有一定关联，它们在坐标旋转变换下总是满足一定的变换关系。例如某一个点的三个坐标 (x, y, z) ，对于图2.1那样的变换有

$$\begin{aligned} x' &= (\cos \gamma \cos \alpha - \cos \beta \sin \alpha \sin \gamma)x + (\cos \gamma \sin \alpha + \cos \beta \cos \alpha \sin \gamma)y \\ &\quad + (\sin \gamma \sin \beta)z \\ y' &= -(\sin \gamma \cos \alpha - \cos \beta \sin \alpha \cos \gamma)x - (\sin \gamma \sin \alpha - \cos \beta \cos \alpha \cos \gamma)y \\ &\quad + (\cos \gamma \sin \beta)z \\ z' &= (\sin \beta \sin \alpha)x - (\sin \beta \cos \alpha)y + (\cos \beta)z \end{aligned} \quad (2.1)$$

简化为二维的情形， x - y 平面相对转一个角度 θ ，如图2.2。同一点在两个坐标系的坐标分别为 (x, y) 和 (x', y') ，它们之间的关系是

$$\begin{aligned} x' &= r \cos(\phi - \theta) \\ &= r \cos \phi \cos \theta + r \sin \phi \sin \theta \\ &= x \cos \theta + y \sin \theta \end{aligned} \quad (2.2)$$

$$\begin{aligned} y' &= r \sin(\phi - \theta) \\ &= -r \cos \phi \sin \theta + r \sin \phi \cos \theta \\ &= -x \sin \theta + y \cos \theta \end{aligned} \quad (2.3)$$

其中， $r = \sqrt{x^2 + y^2}$ 。

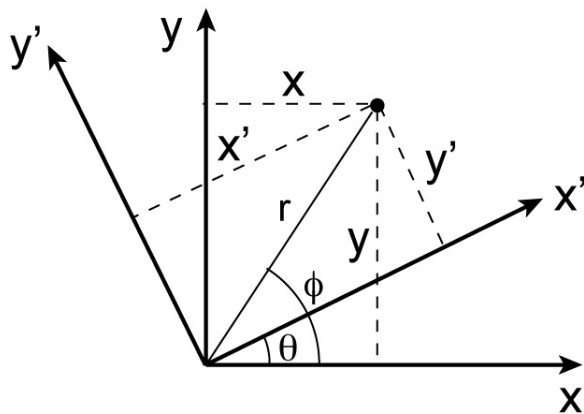


图 2.2: 二维转动

我们把如此关联在一起的三个量称为一个矢量。一般地，一个矢量记为 $\vec{Q} = (Q_x, Q_y, Q_z)$ 。一个矢量的三个分量在坐标变换下与 $\vec{r} = (x, y, z)$ 有相同的变换关系。需要强调的是有大小和方向并不一定是矢量，例如有限转动。矢量必须与 \vec{r} 的变换方式相同。

将变换系数用 λ 代表，三维空间的变换可以记为

$$Q'_x = \lambda_{11}Q_x + \lambda_{12}Q_y + \lambda_{13}Q_z$$

$$\begin{aligned} Q'_y &= \lambda_{21}Q_x + \lambda_{22}Q_y + \lambda_{23}Q_z \\ Q'_z &= \lambda_{31}Q_x + \lambda_{32}Q_y + \lambda_{33}Q_z \end{aligned} \quad (2.4)$$

二维转动的情形

$$\begin{aligned} Q'_x &= Q_x \cos \theta + Q_y \sin \theta \\ Q'_y &= -Q_x \sin \theta + Q_y \cos \theta \\ Q'_z &= Q_z \end{aligned} \quad (2.5)$$

假如一个物理规律在S参考系中可以写为 $\sum_i Q_{ix} = 0$ 、 $\sum_i Q_{iy} = 0$ 和 $\sum_i Q_{iz} = 0$ 。可以证明在S'参考系中

$$\begin{aligned} \sum_i Q'_{ix} &= \sum_i (\lambda_{11}Q_{ix} + \lambda_{12}Q_{iy} + \lambda_{13}Q_{iz}) \\ &= \lambda_{11} \sum_i Q_{ix} + \lambda_{12} \sum_i Q_{iy} + \lambda_{13} \sum_i Q_{iz} \\ &= 0 \end{aligned} \quad (2.6)$$

类似地， $\sum_i Q'_{iy} = 0$ 和 $\sum_i Q'_{iz} = 0$ 。这样，由于每一项 Q 都有相同的变换关系，在S'参考系中物理规律就与S有着相同的形式。此时，我们可以将规律写为更简洁的形式： $\sum_i \vec{Q}_i = 0$ 。这个形式就意味着在任何一个参考系中它的每一个分量都为零。这里的要点是所有的 Q 彼此都按照同样的由 λ 给出的形式进行变换。但这并不一定要求 \vec{Q} 与 \vec{r} 有相同的变换关系。然而通常 \vec{r} 或其导出的量会出现在某一个或一些 \vec{Q} 中，所以将矢量定义为与 \vec{r} 有相同的变换形式就是一个最合理的做法了。协变性表明不同惯性系的观测者是等价的，物理定律在不同的惯性系中有相同的形式。观测者没有办法通过物理定律的形式来区别两个惯性系。协变性对定律的数学结构给出了限制。

举一个物理中的例子。在二维空间，某个粒子的运动方程在S中可以写为

$$\begin{aligned} m \frac{d^2x}{dt^2} + f(r)x &= 0 \\ m \frac{d^2y}{dt^2} + f(r)y &= 0 \end{aligned} \quad (2.7)$$

函数 $f(r)$ 只依赖于 r 。在S'中运动方程变换为

$$\begin{aligned} m \frac{d^2x'}{dt^2} + f(r)x' &= 0 \\ m \frac{d^2y'}{dt^2} + f(r)y' &= 0 \end{aligned} \quad (2.8)$$

其中 m 在参考系转动下保持不变，是标量。该运动方程在不同参考系下有相同的形式，因此具有协变性。具有协变性的方程可以简单地写成一个不依赖于坐标系的矢量形式

$$m \frac{d^2\vec{r}}{dt^2} + f(r)\vec{r} = 0 \quad (2.9)$$

为了方便讨论矢量及参考系间的变换，需要引入一个记号，叫矩阵。三维的变换可以记为

$$\vec{r}' = \Lambda \vec{r} \quad (2.10)$$

其中

$$\vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \vec{r}' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}, \Lambda = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{pmatrix} \quad (2.11)$$

这些符号都是矩阵。

对于一个绕z轴的转动

$$\Lambda = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.12)$$

矩阵的乘法是相应的行和列的乘积。例如

$$\Lambda \vec{r} = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \lambda_{11}x + \lambda_{12}y + \lambda_{13}z \\ \lambda_{21}x + \lambda_{22}y + \lambda_{23}z \\ \lambda_{31}x + \lambda_{32}y + \lambda_{33}z \end{pmatrix} \quad (2.13)$$

又如

$$\begin{pmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} \end{pmatrix} \begin{pmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{pmatrix} = \begin{pmatrix} \omega_{11} & \omega_{12} & \omega_{13} \\ \omega_{21} & \omega_{22} & \omega_{23} \\ \omega_{31} & \omega_{32} & \omega_{33} \end{pmatrix} \quad (2.14)$$

其中

$$\omega_{11} = \gamma_{11}\lambda_{11} + \gamma_{12}\lambda_{21} + \gamma_{13}\lambda_{31} \quad (2.15)$$

...

上式简记为

$$\Gamma \Lambda = \Omega \quad (2.16)$$

矩阵的转置定义为将矩阵的行变为列，列变为行。例如

$$\vec{r}^T = (x \ y \ z) \quad (2.17)$$

$$\Lambda^T = \begin{pmatrix} \lambda_{11} & \lambda_{21} & \lambda_{31} \\ \lambda_{12} & \lambda_{22} & \lambda_{32} \\ \lambda_{13} & \lambda_{23} & \lambda_{33} \end{pmatrix} \quad (2.18)$$

可以证明

$$(\Lambda B)^T = B^T \Lambda^T \quad (2.19)$$

来看一下转置的用途

$$r^2 = \vec{r}^T \vec{r} = (x \ y \ z) \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (2.20)$$

可以得到矢量 \vec{r} 长度的平方。由于矢量的长度在变换下保持不变

$$r'^2 = \vec{r}'^T \vec{r}' = \vec{r}^T \Lambda^T \Lambda \vec{r} = r^2 \quad (2.21)$$

所以任何 Λ 必须满足

$$\Lambda^T \Lambda = I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.22)$$

I 称为单位矩阵。

\vec{r} 本身是矢量。一个有用的方法是使用矢量的长度和沿矢量方向的单位矢量来标记矢量： $\vec{r} = r\hat{r}$ 。

一般地，物理学家从实验中得到一些启发，着手尝试各种可能的物理规律的形式，以便找到一个可以比较好地描述实验观测的方程。协变性为这种寻找提供了一个很强的限制，因而使物理学家更加有目的性。一个常见的构造物理量的方法是从已知的矢量出发，通过运算得到新的矢量或标量。下面就讨论这种可能性。首先让我们找到更多的矢量。速度定义为

$$\vec{v} = \frac{d\vec{r}}{dt} \quad (2.23)$$

可以证明 \vec{v} 是矢量：

$$\vec{v}' = \frac{d\vec{r}'}{dt} = \frac{d(\Lambda\vec{r})}{dt} = \Lambda \frac{d\vec{r}}{dt} = \Lambda\vec{v} \quad (2.24)$$

\vec{v} 和 \vec{r} 满足同样的变换，所以是矢量。同样可以证明加速度 \vec{a} 也是矢量。

矢量的和定义为

$$\vec{a} + \vec{b} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ a_3 + b_3 \end{pmatrix} \quad (2.25)$$

矢量的和仍然是矢量：

$$\vec{a}' + \vec{b}' = \Lambda\vec{a} + \Lambda\vec{b} = \Lambda(\vec{a} + \vec{b}) \quad (2.26)$$

矢量和标量的乘积定义为

$$\vec{b} = \mu\vec{a} \quad (2.27)$$

\vec{b} 的长度是 $b = |\mu|a$ ，方向是 $\hat{b} = \text{sign}(\mu)\hat{a}$ 。容易证明矢量与标量的乘积是矢量：

$$\vec{b}' = \mu'\vec{a}' = \mu\Lambda\vec{a} = \Lambda\mu\vec{a} = \Lambda\vec{b} \quad (2.28)$$

一个例子是动量。

我们可以定义一种运算叫标量积：

$$\vec{a} \cdot \vec{b} = \vec{a}^T \vec{b} = (a_1 \ a_2 \ a_3) \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = a_1 b_1 + a_2 b_2 + a_3 b_3 \quad (2.29)$$

标量积利用两个矢量构造一个标量：

$$\vec{a}' \cdot \vec{b}' = \vec{a}'^T \vec{b}' = \vec{a}^T \Lambda^T \Lambda \vec{b} = \vec{a}^T \vec{b} = \vec{a} \cdot \vec{b} \quad (2.30)$$

$\vec{a} \cdot \vec{b}$ 在坐标转动下不变，是标量。为方便计算，我们需要找出一种计算标量积的更直观的算法。利用 $\vec{a} \cdot \vec{b}$ 是标量的性质，可知对坐标进行变换将得到同样的标量积。变换后，让 \vec{a} 和 \vec{b} 都处在x-y平面上，并使得 \vec{a} 与x轴正方向重合。这样

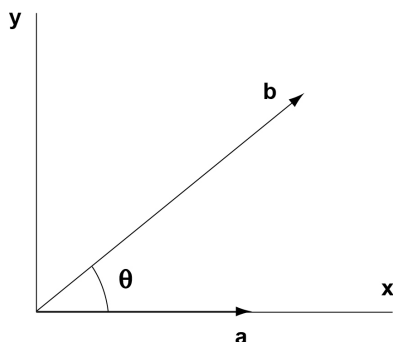


图 2.3: 标量积和矢量积

$$\vec{a} \cdot \vec{b} = (a \ 0 \ 0) \begin{pmatrix} b \cos \theta \\ b \sin \theta \\ 0 \end{pmatrix} = ab \cos \theta \quad (2.31)$$

也就是标量积是两个矢量的长度之积再乘以夹角的余弦。如果 $\vec{a} \cdot \vec{b} = 0$ ，那么 $a = 0$ 或 $b = 0$ 或 $\vec{a} \perp \vec{b}$ 。再有 $\vec{a} \cdot \vec{a} = a^2$ 。从两个矢量出发通过标量积构造一个标量例子是功： $W = \vec{F} \cdot \vec{s}$ 。

从两个矢量出发还可以定义一种运算来得到一个矢量。 $\vec{c} = \vec{a} \times \vec{b}$ 叫矢量积。其中

$$c_1 = a_2 b_3 - a_3 b_2 \quad (2.32)$$

$$c_2 = a_3 b_1 - a_1 b_3 \quad (2.33)$$

$$c_3 = a_1 b_2 - a_2 b_1 \quad (2.34)$$

通过代数的方法可以证明 \vec{c} 是一个矢量。证明较繁，此处略去。我们可以找到计算矢量积更加直观的办法。首先，

$$\vec{a} \cdot \vec{c} = (a_1 \ a_2 \ a_3) \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix} = 0 \quad (2.35)$$

同样 $\vec{b} \cdot \vec{c} = 0$ 。因而 \vec{a} 、 \vec{b} 垂直于 \vec{c} 。下面我们需要找出 \vec{c} 的指向和大小。与处理标量类似，我们可以让坐标系转动（矢量长度不变），使得 $\vec{a} = \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix}$ ， $\vec{b} = \begin{pmatrix} b \cos \theta \\ b \sin \theta \\ 0 \end{pmatrix}$ ，则 $\vec{c} = \begin{pmatrix} 0 \\ 0 \\ ab \sin \theta \end{pmatrix}$ 。于是 \vec{a} 、 \vec{b} 、 \vec{c} 满足右手定则， $c = ab \sin \theta$ 。

事实上，通过矢量积得出的矢量与我们通常的矢量并不一样。我们把通常的矢量叫极矢量 (polar vector)，而矢量积得到的矢量叫轴矢量 (axial vector) 或赝矢量 (pseudovector)。基矢量和轴矢量在坐标转动变化下的行为没有区别，但它们在镜面反射变换下的行为不同。

两个矢量积的例子：

$$\text{角动量: } \vec{J} = m\vec{r} \times \vec{v}. \quad \text{洛仑兹力: } \vec{F} = q\vec{v} \times \vec{B}$$

矢量积的性质：

$$\vec{a} \times \vec{a} = 0 \quad (2.36)$$

$$\vec{a} \times \vec{b} = -\vec{b} \times \vec{a} \quad (\text{无交换律}) \quad (2.37)$$

$$\vec{a} \times (\vec{b} + \vec{c}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{c} \quad (\text{分配律}) \quad (2.38)$$

$$\vec{a} \times (\vec{b} \times \vec{c}) \neq (\vec{a} \times \vec{b}) \times \vec{c} \quad (\text{无结合律}) \quad (2.39)$$

为了方便，我们经常使用基矢量。三维空间的三个基矢量分别沿三个轴，长度为1。

$$\hat{i} \cdot \hat{i} = \hat{j} \cdot \hat{j} = \hat{k} \cdot \hat{k} = 1 \quad (2.40)$$

$$\hat{i} \cdot \hat{j} = \hat{j} \cdot \hat{k} = \hat{k} \cdot \hat{i} = 0 \quad (2.41)$$

$$\hat{i} \times \hat{j} = \hat{k} \quad (2.42)$$

$$\hat{j} \times \hat{k} = \hat{i} \quad (2.43)$$

$$\hat{k} \times \hat{i} = \hat{j} \quad (2.44)$$

一个矢量可以表示为 $\vec{A} = A_x \hat{i} + A_y \hat{j} + A_z \hat{k}$

矢量积可以写为

$$\vec{A} \times \vec{B} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} \quad (2.45)$$

右边的项叫行列式。

以上仅考虑的是两个相对静止的参考系间的变换。对于做匀速直线运动的参考系，有如下假设。

假设ID

惯性系间满足伽利略变换。

为简单起见，使用直角坐标系，让两个坐标系的原点在 $t = t' = 0$ 时重合，两坐标系的坐标轴平行，A' 相对于 A 的速度为 \vec{v} 。这时伽利略变换表示为

$$\vec{r}' = \vec{r} - \vec{v}t \quad (2.46)$$

$$t' = t \quad (2.47)$$

由此还可以得到速度变换关系 $\vec{u}' = \vec{u} - \vec{v}$ 。

之所以伽利略变换不能采用如图2.4所示的方法来得到，是因为 \vec{r}' 和 \vec{r} 不在同一个参考系，它们的坐标由各自参考系中的测量仪器决定，没有理由保证由 (x, y, z) 表示的矢量和由 (x', y', z') 表示的矢量可以简单地按照数学的方式相加。矢量相加的法则只适用于同一个参考系中。

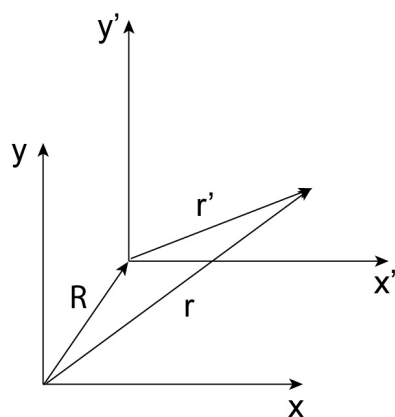


图 2.4: 不可靠的伽利略变换推导

伽利略变换的一个直接推论是两个事件对一个惯性系的同时性也意味着在一切惯性系中的同时性。这个有关同时性的陈述意味着在牛顿力学中时间是绝对的。

3 力与质量

假设I是关于参考系的陈述。我们现在需要一个惯性系中的运动方程。为此，牛顿首先写下 $\vec{F} = m\vec{a}$ ，这里力 \vec{F} 和加速度 \vec{a} 都是有大小有方向的矢量， m 是质量。力代表外界的驱动。加速度是粒子对外界驱动响应。质量刻画了响应的程度，称为惯性。在此，牛顿假设了外界的驱动可以表示为某种形式，使得响应与之成正比。这显然是一种最简单的选择。到此为止我们还需要关于力的详细知识以便描述运动。经过尝试，牛顿发现由以下力的形式可以非常准确地导出太阳系中行星的轨道：

$$\vec{F}_1 = \frac{Gm_1m_2}{r^2}\hat{r}_{12} \quad (3.1)$$

$$\vec{F}_2 = \frac{Gm_1m_2}{r^2}\hat{r}_{21} \quad (3.2)$$

\hat{r}_{12} 和 \hat{r}_{21} 分别是粒子1到2和从2到1方向上的单位位移矢量。两粒子间存在以上形式的相互吸引力，被称为万有引力。力的方向沿两粒子的连线。力的大小正比于粒子的质量，与两粒子间的距离呈平方反比关系。万有引力的两个要点是与距离成平方反比和与质量成正比。此两点就可以完美地导出开普勒的三个定律。

假设万有引力也适用于地球上物体。一个物体在地球表面上受到的地球引力称为重力。重力大小正比于质量。因此利用力的平衡（本质上就是合力为零时加速度为零）就可以很方便地确定两个物体的质量之比。于是人们定义了一个保存于国际计量局由铂铱合金制成的标准质量，称为国际千克原器，大小指定为1 kg。这样就有了质量的操作性定义。在此定义下，

$$G = 6.6742 \times 10^{-11} \text{ N} \cdot \text{m}^2/\text{kg}^2 \quad (3.3)$$

力的单位是牛顿 (N)， $1 \text{ N} = 1\text{kg} \cdot \text{m}/\text{s}^2$ 。

有了重力就可以方便地对各种力进行测量、标定。在牛顿的万有引力定律出现之后的大约一百年的时间里，人们在实验中又发现除质量外一个粒子还需要引入另一个粒子的内禀性质，称为电荷，用来描述实验中发现的另一种相互作用，称为静电作用。在一个惯性系中静止的两个电荷间的静电力大小为（库仑定律）：

$$F = \frac{1}{4\pi\epsilon_0} \frac{|q_1||q_2|}{r^2} \quad (3.4)$$

q 为电荷，单位是库仑 (C)。与引力不同，电荷可正可负，静电力可吸引也可排斥。电荷同性时相斥，异性相吸。1 C的定义是：两个电荷量均为 q 、距离为1 m的静止粒子，如相互作用力为 $8.988 \times 10^9 \text{ N}$ ，则 $q = 1 \text{ C}$ 。 $\epsilon_0 = 8.85 \times 10^{-12} \text{ C}^2 \cdot \text{N}^{-1} \cdot \text{m}^{-2}$ ，称为真空介电常数。

简单的计算就可以表明，对带电粒子（电子、质子等）而言，静电力远大于引力，这时引力可以忽略。两个电子之间的引力：

$$F_G \sim 6 \times 10^{-11} \frac{(10^{-30})^2}{r^2} \quad (3.5)$$

两个电子之间的静电力：

$$F_k \sim 10^{10} \frac{(1.6 \times 10^{-19})^2}{r^2} \quad (3.6)$$

二者的比值

$$\frac{F_G}{F_k} \sim 10^{-17}/10^{-28} = 10^{-43} \quad (3.7)$$

这个比值大致相当于光通过一个质子所需时间与宇宙寿命之比。与静电力相比，引力非常弱。只有在电中性环境下，引力才能显著表现出来。甚至对于中性粒子（原子、分子等），由于极化的存在，相互作用也仍由静电力主导（比如范德华力）。对于中性的天体而言，其巨大的质量使得引力成为起主导的相互作用，静电力可以被忽略。

在迄今为止发现的四种基本相互作用中，引力和电磁作用的有效范围非常之广，从亚原子尺度到实际上无限远。这两种力在我们的日常生活中起决定作用。对于荷电体系，引力通常微不足道，只有电磁力才是唯一重要的力。例如，电磁力是把电子束缚在原子核周围形成原子以及所有原子、分子间相互作用的来源，决定了各种固体、液体、气体的性质，主导着这种化学反应和生物过程。在宏观上，我们日常所见的许多作用力均是静电作用的结果，如弹性力、表面张力、摩擦力等等。这些非基本的作用力在微观上都来源于大量电荷之间的静电力。任何一个描述非基本力看上去简单的公式都是对微观上基本相互作用求和的结果，本质上都是电磁作用。

在讨论万有引力时，有一处很不自然。出现在牛顿运动方程中的质量代表惯性，是一个物体抵抗外力的能力。而万有引力中的质量是决定力大小的内禀性质，与电荷类似。原则上，我们应该赋予每个粒子两个独立的内禀参量：决定惯性的惯性质量和决定引力的引力质量。二者的物理意义完全不同。然而实验发现引力质量在极高的精度内（ 10^{-11} ）与惯性质量成正比。例如，二者的正比关系造成在地球上初始条件（初始位置、初始速度）相同但质量不同的两个粒子，其运动轨迹相同。考虑到惯性质量和引力质量刻画的是完全不同的性质，二者成正比这个实验事实显得极为不平凡。然而虽然很难相信二者之间的这种关系纯属偶然，但直到有了爱因斯坦的广义相对论，其等价性才有了一个自然的解释。既然二者等价，一个方便的做法是选取合适的引力常数 G ，使得引力质量与惯性质量在数值上相等，并有同样的单位。

之前我们曾经用重力来定义惯性质量。然而由于惯性质量与引力质量的物理意义完全不同，我们更希望有一个独立于重力的对惯性质量的定义。于是有了如下假设来代替牛顿运动方程并不依赖于力定义质量。

假设II

在惯性系中观测一个相互作用粒子组成的孤立（无外力）体系，总可以将每一个粒子赋予一个与参考系和相互作用无关的正的常数 $\mu(i)$ ，并有如下性质： $\sum_i \mu(i) \vec{v}(i)$ ， $\vec{v}(i)$ 是速度，对于这个相互作用的粒子系统在任何时刻都有相同的值，而无论粒子间相互作用是什么。

值得注意的是 $\sum_i \mu(i) \vec{v}(i)$ 这个性质满足协变性。该假设隐含着相互作用的传递不需要时间，是瞬时的、超距的。我们可以定义惯性质量，简称

质量。不失一般性，给定两个粒子0和*i*，在相互作用前后， $\mu(i)\Delta\vec{v}(i) = -\mu(0)\Delta\vec{v}(0)$ 。令粒子0的质量为1 kg。粒子*i*的质量大小定义为 $m(i) = \frac{\mu(i)}{\mu(0)}$ 。质量是一个粒子的内禀性质，不依赖于参考系、运动状态和所受到的相互作用。有了质量之后可以定义动量： $\vec{p} = m\vec{v}$ ， \vec{p} 是矢量，并有所谓动量守恒定律：孤立系统动量守恒。

我们的目标是确定体系中每个粒子的运动轨迹。如果粒子间没有相互作用，则每个粒子都会做匀速直线运动。如果存在相互作用，一般地，粒子的速度会随时间改变。为了描述相互作用，我们引入力的概念，定义为 $\vec{F} = d\vec{p}/dt = m\vec{a}$ 。事实上力的定义具有相当的任意性。原则上力可以是 \vec{p} 和*t*的任意函数： $\vec{F} = \vec{F}(\vec{p}, t)$ ，只是这时运动方程和力的表达式通常不再具有简单的形式。有了力的定义后就需要对每一种相互作用通过尝试来写下它的表达形式，以使得运动方程能够足够好地描述粒子的轨迹。例如，牛顿猜测引力是平方反比力，带入牛顿运动方程后给出了开普勒通过实验观测得到的三定律。没有办法能够依靠逻辑先验地给出力的表达式。基于以下的观察，力是个有用的工具。

力可以帮助我们将运动约化成为一个简洁的物理定律。一般地，关于一个体系的完备信息由每个粒子在每个时刻的位置构成。但是我们发现可以有一个等价（即包含了完全相同的信息量）的方式。这种方式需要知道力的表达形式，以及在某个时刻的初始条件即可。牛顿最重要的贡献就在于他发现了这种约化的可能性。之后的科学发展都遵循着这条道路。

实验观测经常会发现如下事实：两个构成相同的体系A和B，如果在某个时刻，A中每一个粒子都与B中相应的粒子有相同的位置和速度，那么A与B在所有时刻的运动状态都会相同。如此的行为暗示着描述相互作用的力应该仅是粒子的位置和速度的函数而不含有更高阶的位移对时间的微分，否则初始条件就不能仅包括位置和速度。

通常力可以叠加。考虑三个粒子组成的体系。作用在第一个粒子上的力 \vec{F}_1 可以分解为 \vec{F}_{12} 和 \vec{F}_{13} 两项之和，并且即使第三个粒子不存在 \vec{F}_{12} 仍然不变，两个粒子之间的力不受其他粒子的影响。一般地， $\vec{F}_i = \sum_{j \neq i} \vec{F}_{ij}$ ，即力可以按照粒子对来分解。 \vec{F}_{ij} 只与*i*和*j*两个粒子有关。更进一步，我们发现对于同一类相互作用， \vec{F}_{ij} 都可以写成同样的形式（通式）而不依赖于具体粒子。这样力所依赖的变量就变得很简单。如果力没有普适性（想象一下每两个粒子之间的力的形式都不同的情形），则我们将可能需要一套完全不同的框架（至少不能用力力的概念）来描述世界。

在宏观上，力是一种方便的唯像工具，接近人的直观感受。这些力的根源往往是电磁力，但用有了力的概念就可以不必考虑相互作用的细节，而只是唯象地给出形式，只要足够准确就可以很方便地描述这些宏观上的非基本相互作用。可以想像如果我们不使用摩擦力，而是对每个原子之间的相互作用求和，那将会是一个灾难。一般地这些力的形式需要通过实验来确定，当然也可以从微观上推导出来（这是统计力学的一个任务）。但由于大量粒子的存在，并不是每种非基本的相互作用都可以方便地从微观上推导出来。

只考虑两个粒子的相互作用,

$$\vec{F}_{12} = \frac{d\vec{p}_1}{dt} \quad (3.8)$$

$$\vec{F}_{21} = \frac{d\vec{p}_2}{dt} \quad (3.9)$$

那么,

$$\vec{F}_{12} + \vec{F}_{21} = \frac{d}{dt}(\vec{p}_1 + \vec{p}_2) = 0 \quad (3.10)$$

于是有,

$$\vec{F}_{12} = -\vec{F}_{21} \quad (3.11)$$

如果两个粒子间有相互作用, 第一个粒子作用到第二个粒子上的力与第二个粒子作用到第一个粒子上的力大小相等, 方向相反。这就是牛顿第三定律。事实上, 第三定律只是对力的性质进行了一些描述, 完全不能与前两条定律放在同等的地位, 况且力并不是牛顿力学中对相互作用的唯一描述方法。更近一步, 第三定律在很多时候与观测不符, 原因是引入了力这种超距、瞬时的方式来描述相互作用不合理。例如考虑两个运动电荷之间的相互作用(如图3.1), Q_1 在 Q_2 处产生的磁场为0, 因而 Q_2 受到的Lorentz力为0。而 Q_2 在 Q_1 处产生的磁场不为0, Q_1 受到的Lorentz力也不为0。在这种情况下作用力不等于反作用力。

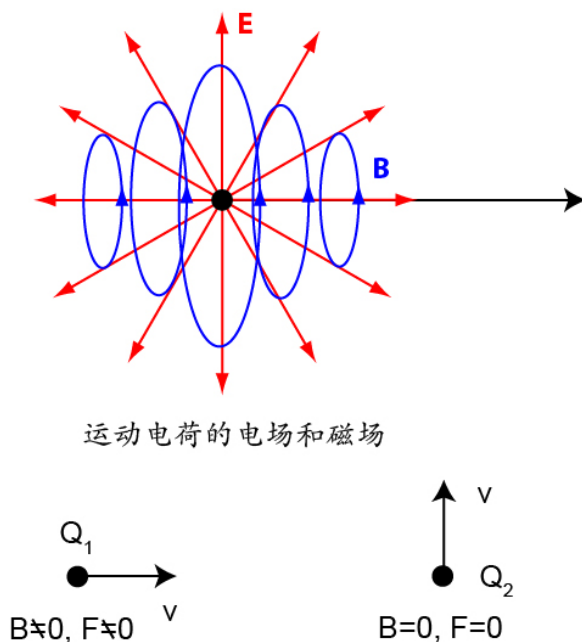


图 3.1: 作用力不等于反作用力

到此我们已经搭建完成了牛顿力学的框架。如果发现牛顿力学与实验不符, 我们能做些什么呢? 可能发生的事情是: 一、参考系可能不是惯性系; 二、孤立体系一直是讨论各种问题的一个重要前提, 也许某个系统并不是那么孤立; 三、某个假设不是很好, 例如用力这种超距、瞬时的方式来描述相互作用。

4 不变性

不变性又称对称性，指的是在一个观测者看来将被研究系统变换后仍然得到一个满足同样物理原理的系统。例如，把一个粒子体系的每一个粒子每一时刻的位置都在空间上平移一个相同的位移后仍然是一个可以实现的体系。不变性并不意味着协变性，协变性也不意味着不变性。我们期待物理世界有很多对称性，如由于空间均匀性所导致的平移不变性和由于空间各向同性导致的转动不变性。这些对称性为物理定律的形式给出了严格的限制。

先考虑空间平移不变性。空间平移不变性说的将被研究的系统在空间上进行平移后仍然满足同样的物理规律。在这里观察者是同一个。具体地讲，考虑两个全同的系统A和A'。A是一个真实的系统，A'暂时是一个假想的系统。A'中每个粒子在任何时刻的坐标都与A中相应粒子的坐标沿某一确定方向平移相同的距离 a 。由于我们坚信空间平移不变性，我们断定A'也是一个可以实现的真实系统。于是观察者为了描述运动而写下的物理规律（方程）应该同样好地描述A和A'的运动。这就为物理规律可能的形式提出了限制，要求物理规律必须有某种特殊的形式。一方面，由于A'是由A衍生出来的，所以与A相比A'的运动除了空间上平移之外没有别的区别，完全是A的复制。因此在数学上将描述A的方程中的坐标都平移 a 就可以描述A'的运动。另一方面，由于A'也同样是一个可以实现的真实物理系统，所以如果将方程中的所有坐标都直接写成A'的坐标也可以得到描述A'运动的方程。但是上述的两个方程必须完全相同。这就要求物理规律必须有某种特殊的形式。否则一般地不能保证两个方程相同。

例如考虑相互作用的两个粒子的一维运动。观察者通过研究决定将描述物理规律的运动方程写为：

$$m \frac{d^2 x_1}{dt^2} + \alpha \frac{dx_1}{dt} + \beta x_1 = F(x_1, x_2) \quad (4.1)$$

$$m \frac{d^2 x_2}{dt^2} + \alpha \frac{dx_2}{dt} + \beta x_2 = -F(x_1, x_2) \quad (4.2)$$

将A系统两粒子的坐标记为 X_1 和 X_2 。将运动方程运用到A上，称为一个实现，这时有

$$m \frac{d^2 X_1}{dt^2} + \alpha \frac{dX_1}{dt} + \beta X_1 = F(X_1, X_2) \quad (4.3)$$

$$m \frac{d^2 X_2}{dt^2} + \alpha \frac{dX_2}{dt} + \beta X_2 = -F(X_1, X_2) \quad (4.4)$$

数学上，在式 4.3 中进行代换 $X'_1 = X_1 + a$ 和 $X'_2 = X_2 + a$ 可以得到A'的方程

$$m \frac{d^2 X'_1}{dt^2} + \alpha \frac{dX'_1}{dt} + \beta X'_1 - \beta a = F(X'_1 - a, X'_2 - a) \quad (4.5)$$

$$m \frac{d^2 X'_2}{dt^2} + \alpha \frac{dX'_2}{dt} + \beta X'_2 - \beta a = -F(X'_1 - a, X'_2 - a) \quad (4.6)$$

另一方面，从式 4.1 应该可以直接写出描述 A' 的运动方程：

$$m \frac{d^2 X'_1}{dt^2} + \alpha \frac{dX'_1}{dt} + \beta X'_1 = F(X'_1, X'_2) \quad (4.7)$$

$$m \frac{d^2 X'_2}{dt^2} + \alpha \frac{dX'_2}{dt} + \beta X'_2 = -F(X'_1, X'_2) \quad (4.8)$$

而式 4.5 和 4.7 必须完全相同。这就要求 $\beta = 0$ 。另外要求对于任何 a 有 $F(X'_1 - a, X'_2 - a) = F(X'_1, X'_2)$ 。一个合理的期望是 $F(x_1, x_2) = F(x_1 - x_2)$ ，而不是 x_1 和 x_2 单独的函数。这样运动方程可以写成

$$m \frac{d^2 x_1}{dt^2} + \alpha \frac{dx_1}{dt} = F(x_1 - x_2) \quad (4.9)$$

$$m \frac{d^2 x_2}{dt^2} + \alpha \frac{dx_2}{dt} = -F(x_1 - x_2) \quad (4.10)$$

下面我们继续说明对称性还要求 $\alpha = 0$ 。

一个重要的对称性是所谓时间反演不变性。时间反演不是时间倒流，而是说把运动反过来。如图 4.1，一个粒子从 0 时刻运动到 T 时刻，这时另一个粒子原路返回，从 T 时刻运动到 $2T$ 时刻，并且 $\vec{r}'(t) = \vec{r}(2T - t)$ 。

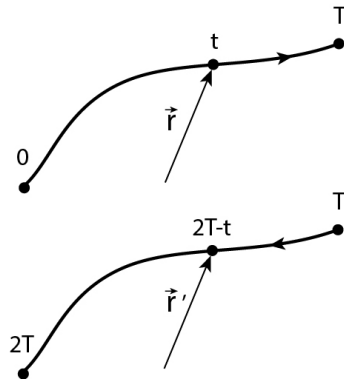


图 4.1: 时间反演

时间反演不变性是说系统 A 的时间反演状态 A' 也是一个真实的物理系统。例如把地球的运动反过来则仍然是可能的运动。于是就要求运动方程必须满足某种形式，否则不能用来描述时间反演不变性的系统。在数学上，将 $X(t)$ 都换为 $X'(2T - t)$ 就得到 A' 系统的运动方程：

$$m \frac{d^2 X'_1(2T - t)}{dt^2} + \alpha \frac{dX'_1(2T - t)}{dt} = F(X_1 - X_2) \quad (4.11)$$

$$m \frac{d^2 X'_2(2T - t)}{dt^2} + \alpha \frac{dX'_2(2T - t)}{dt} = -F(X_1 - X_2) \quad (4.12)$$

即

$$m \frac{d^2 X'_1}{dt^2} - \alpha \frac{dX'_1}{dt} = F(X_1 - X_2) \quad (4.13)$$

$$m \frac{d^2 X'_2}{dt^2} - \alpha \frac{dX'_2}{dt} = -F(X_1 - X_2) \quad (4.14)$$

但另一方面，A'也是一个真实的物理系统，运动方程可以由 4.1 式直接写出：

$$m \frac{d^2 X'_1}{dt^2} + \alpha \frac{dX'_1}{dt} = F(X_1 - X_2) \quad (4.15)$$

$$m \frac{d^2 X'_2}{dt^2} + \alpha \frac{dX'_2}{dt} = -F(X_1 - X_2) \quad (4.16)$$

以上两组方程必须一致，所以要求 $\alpha = 0$ 。于是，对称性要求下运动方程的一种可能形式是

$$m \frac{d^2 x_1}{dt^2} = F(x_1 - x_2) \quad (4.17)$$

$$m \frac{d^2 x_2}{dt^2} = -F(x_1 - x_2) \quad (4.18)$$

一般地，时间反演不变性要求运动方程中不出现奇数阶导数。

一个破坏时间反演对称性的例子是带电粒子在磁场中的运动。但如果把产生磁场的电流也反向，则仍有时间反演对称性。

我们利用对称性还可以说明 $F(x_1 - x_2) = -F(x_2 - x_1)$ ，即 F 是自变量的奇函数。这个对称性是所谓的空间反演不变性。空间反演即空间坐标的反号。这样，运动方程又可以写成

$$m \frac{d^2 x_1}{dt^2} = F(x_1 - x_2) \quad (4.19)$$

$$m \frac{d^2 x_2}{dt^2} = F(x_2 - x_1) \quad (4.20)$$

在某些情形，协变性和不变性可以有一定关系。例如，对应于坐标旋转的协变性与空间各向同性导致转动不变性直接相关。考虑参考系 A 和 A'，A' 由 A 通过转动得到。分别由 A 和 A' 来观测一个粒子系统的运动。现在将作为观测者的 A' 和粒子系统看作一个整体。将 A' 和粒子系统一同作反方向转动，使得 A 和 A' 重合，这是 A 和 A' 不再有区别。由于 A 中的转动不变性，所以转动后仍是可以实现的运动。可以容易地说明协变性所要求的矢量方程会导致转动不变性。但一般地，协变性和对称性是两种不同的性质。例如，相对论中的洛伦兹变换指的是协变性，而量子场论中的规范变换指的是对称性。

5 矢量算符

5.1 偏导数与全微分

考虑一个二元函数, x 和 y 同时变化会发生什么?

$$\begin{aligned}\Delta f &= f(x + \Delta x, y + \Delta y) - f(x, y) \\ &= [f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y)] \\ &\quad + [f(x, y + \Delta y) - f(x, y)] \\ &\simeq \frac{\partial}{\partial x} f(x, y + \Delta y) \Delta x + \frac{\partial}{\partial y} f(x, y) \Delta y \\ &\simeq \frac{\partial}{\partial x} [f(x, y) + \frac{\partial f(x, y)}{\partial y} \Delta y] \Delta x + \frac{\partial}{\partial y} f(x, y) \Delta y \\ &\simeq \frac{\partial}{\partial x} f(x, y) \Delta x + \frac{\partial}{\partial y} f(x, y) \Delta y\end{aligned}\tag{5.1}$$

$f(x, y)$ 的全微分定义为

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy\tag{5.2}$$

例:

$$f(x, y) = xy\tag{5.3}$$

$$\begin{aligned}\Delta f &= f(x + \Delta x, y + \Delta y) - f(x, y) \\ &= (x + \Delta x)(y + \Delta y) - xy \\ &= y\Delta x + x\Delta y + \Delta x\Delta y\end{aligned}\tag{5.4}$$

忽略高阶小量

$$df = ydx + xdy\tag{5.5}$$

还可以定义高阶的偏微分, 如

$$\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) \equiv \frac{\partial^2 f}{\partial x^2}\tag{5.6}$$

$$\frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) \equiv \frac{\partial^2 f}{\partial y \partial x}\tag{5.7}$$

$$\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) \equiv \frac{\partial^2 f}{\partial x \partial y}\tag{5.8}$$

$$\frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right) \equiv \frac{\partial^2 f}{\partial y^2}\tag{5.9}$$

容易证明

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}\tag{5.10}$$

5.2 方向导数、梯度

对于二维空间的函数 $f(x, y)$, 当 x 和 y 分别变化时, 函数的变化分别为 $\frac{\partial f}{\partial x}dx$ 和 $\frac{\partial f}{\partial y}dy$ 。现在想求 x, y 在某单位矢量 \hat{u} 方向上变化时 f 的变化, 并求出变化最快的方向。单位矢量 $\hat{u} = u_x\hat{i} + u_y\hat{j}$ 。其中 $u_x^2 + u_y^2 = 1$, $u_x = \cos\theta$, $u_y = \sin\theta$ 。 Δs 是 \hat{u} 方向上的线段, 则 $\Delta x = u_x\Delta s$, $\Delta y = u_y\Delta s$ 。 f 在 \hat{u} 方向上的变化为

$$\begin{aligned}\Delta f(x, y) &= f(x + u_x\Delta s, y + u_y\Delta s) - f(x, y) \\ &= \frac{\partial f}{\partial x}u_x\Delta s + \frac{\partial f}{\partial y}u_y\Delta s\end{aligned}\quad (5.11)$$

于是

$$\frac{\Delta f}{\Delta s} \simeq \frac{\partial f}{\partial x}u_x + \frac{\partial f}{\partial y}u_y\quad (5.12)$$

取极限 $\Delta s \rightarrow 0$

$$\frac{df}{ds} = \frac{\partial f}{\partial x}u_x + \frac{\partial f}{\partial y}u_y\quad (5.13)$$

称为方向导数。定义一个量, 叫梯度 (后面我们会证明这是一个矢量, 于是我们用标量构造了矢量)

$$\vec{\nabla}f = \frac{\partial f}{\partial x}\hat{i} + \frac{\partial f}{\partial y}\hat{j}\quad (5.14)$$

则:

$$\frac{df}{ds} = \vec{\nabla}f \cdot \hat{u}\quad (5.15)$$

类似, 在三维空间

$$\vec{\nabla}f = \frac{\partial f}{\partial x}\hat{i} + \frac{\partial f}{\partial y}\hat{j} + \frac{\partial f}{\partial z}\hat{k}\quad (5.16)$$

由于

$$\frac{df}{ds} = \vec{\nabla}f \cdot \hat{u} \leq |\vec{\nabla}f|\quad (5.17)$$

我们得到 df/ds 的最大值是 $|\vec{\nabla}f|$, 发生在 \hat{u} 与 $\vec{\nabla}f$ 有相同方向时。在一点附近, 沿 $\vec{\nabla}f$ 方向函数变化最快。这就是梯度的直观几何意义。

把具有相同函数值的区域连接起来构成等高面 (二维时是等高线)。在等高面上, $df = 0$ 。所以如果 \hat{u} 与等高面相切, 则 $\vec{\nabla}f \cdot \hat{u} = 0$ 。因而 $\vec{\nabla}f$ 与等高面垂直, 即在垂直等高面的方向上函数变化最快。

$df = \vec{\nabla}f \cdot \hat{u}ds$, 因而如果 $df > 0$, 那么 \hat{u} 与 $\vec{\nabla}f$ 指向相同, 所以 $\vec{\nabla}f$ 的指向是从低 f 的等高面到高 f 的等高面。也就是说 $\vec{\nabla}f$ 指向 f 增加最快的方向。

在垂直于等高面的方向上

$$|\vec{\nabla}f| \sim \frac{\Delta f}{\Delta s}\quad (5.18)$$

因而等高面越密, 梯度越大。

梯度的一个应用是在高维空间求函数极值。

下面我们说明梯度是矢量

f 是一个标量

$$\Delta f = \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial y} \Delta y + \frac{\partial f}{\partial z} \Delta z \quad (5.19)$$

$(\Delta x, \Delta y, \Delta z)$ 是矢量, 而 Δf 是标量, 因而 ∇f 应该是矢量。

5.3 散度

考虑空间中的一个矢量场 $\vec{\Lambda}(\vec{r})$ 。一个体积 V , 如图5.1, 其表面记为 S 。我们想知道 $\vec{\Lambda}$ 通过 S 的总通量 Φ 。通量定义为 $\vec{\Lambda}$ 对 S 的面积分:

$$\Phi = \int_S \vec{\Lambda} \cdot d\vec{a} \quad (5.20)$$

$d\vec{a}$ 是一个无限小的矢量, 它的大小是 S 上一个面积元的面积, 方向垂直于面积元向外。通量代表该体积内总的源或漏。例如, 已知速度场 \vec{v} 的流体, 单位时间流过面积 $d\vec{a}$ 的流量是 $\vec{v} \cdot d\vec{a}$, 而总流量由上面的积分给出。

下面我们试图将作为通量的面积分与闭曲面包围的体的内部性质相关联。现在设想我们将 V 通过一个隔膜 D 分成两部分, 分别记为 V_1 和 V_2 , 然后分别计算每个体积的面积分。显而易见, 两个面积分之和等于原来整个表面的面积分。

$$\int_{S_1} \vec{\Lambda} \cdot d\vec{a}_1 + \int_{S_2} \vec{\Lambda} \cdot d\vec{a}_2 \quad (5.21)$$

原因是 D 上的一个面积元对第一个和第二个积分的贡献大小相等、符号相反。换句话说, V_1 通过 D 流出的通量又流入了 V_2 , 二者抵消。余下的表面与原来相同。我们还可以对 V 继续划分下去。然而无论如何划分, 我们都会有

$$\Phi = \int_S \vec{\Lambda} \cdot d\vec{a} = \sum_{i=1}^N \int_{S_i} \vec{\Lambda} \cdot d\vec{a}_i = \sum_{i=1}^N V_i \left[\frac{\int_{S_i} \vec{\Lambda} \cdot d\vec{a}_i}{V_i} \right] \quad (5.22)$$

下面我们计算取极限 $N \rightarrow \infty$ 下一个体积元 i 的面积分与体积的比值。该比值代表该体积内源或漏的平均密度。在 V_i 趋于0的极限下表示了矢量场 $\vec{\Lambda}$ 在该点附近的性质, 称为散度, 记为 $\text{div} \vec{\Lambda}$ 。

$$\text{div} \vec{\Lambda} = \lim_{V_i \rightarrow 0} \frac{1}{V_i} \int_{S_i} \vec{\Lambda} \cdot d\vec{a}_i \quad (5.23)$$

$\text{div} \vec{\Lambda}$ 表示的是对于无限小体积元单位体积的通量。如果我们知道了 $\text{div} \vec{\Lambda}$, 我们就可以反过来得到对于任何体积的通量。

$$\int_S \vec{\Lambda} \cdot d\vec{a} = \int_V \text{div} \vec{\Lambda} dV \quad (5.24)$$

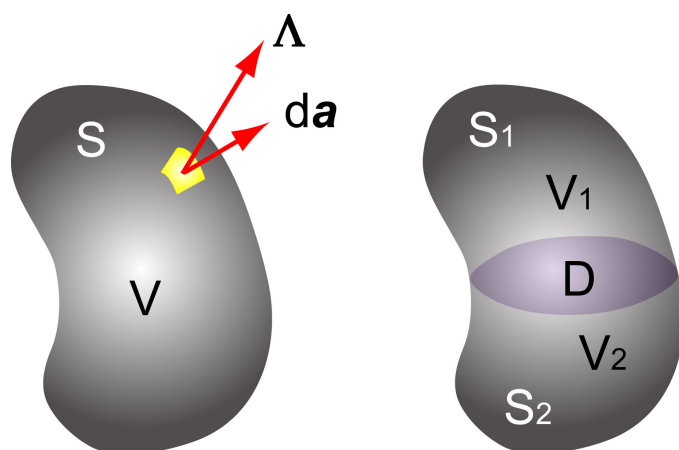


图 5.1: 通量的计算

这叫做Gauss定理，它给出了通量与源的关系。左边是矢量场的通量。如果通量为0就意味者这个体积中没有净的源或漏。 $\text{div}\vec{\Lambda}$ 可以看作源或漏的密度。 $\text{div}\vec{\Lambda}$ 的体积分是该体积中总的源或漏。

下面我们计算散度定义中极限。为此，我们需要在直角坐标系中表示 $\text{div}\vec{\Lambda}$ 。我们计算一个长方体（图5.2）的通量。长方体的中心位于 (x,y,z) ，边长分别为 $\Delta x, \Delta y, \Delta z$ 。首先计算上下两个面。计算这两个面的通量时只需考虑 $\vec{\Lambda}$ 的

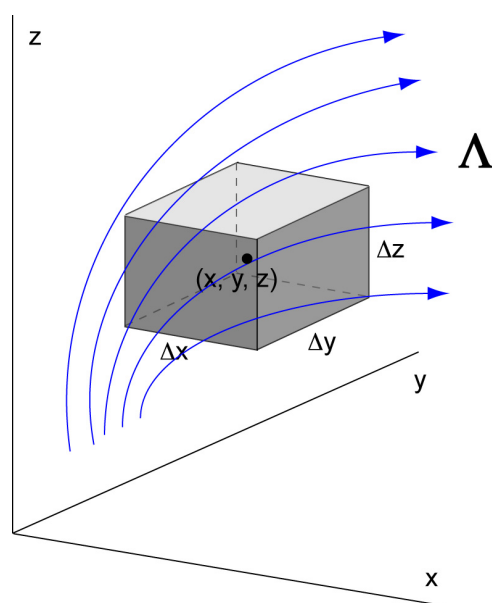


图 5.2: 长方体的通量

z 分量。对通量总的贡献依赖于 Λ_z 在上下表面平均值的差别。对于下表面

$$- \int \Lambda_z \left(x', y', z - \frac{1}{2} \Delta z \right) dx' dy'$$

$$\begin{aligned}
&= - \int \Lambda_z \left(x + (x' - x), y + (y' - y), z - \frac{1}{2}\Delta z \right) dx' dy' \\
&= - \int \left(\Lambda_z(x, y, z) + (x' - x) \frac{\partial \Lambda_z}{\partial x} + (y' - y) \frac{\partial \Lambda_z}{\partial y} - \frac{1}{2}\Delta z \frac{\partial \Lambda_z}{\partial z} \right) dx' dy' \\
&= -\Lambda_z(x, y, z)\Delta x\Delta y + \frac{1}{2} \frac{\partial \Lambda_z}{\partial z} \Delta x\Delta y\Delta z
\end{aligned} \tag{5.25}$$

类似地，上表面的通量是

$$\begin{aligned}
&\int \Lambda_z \left(x', y', z + \frac{1}{2}\Delta z \right) dx' dy' \\
&= \Lambda_z(x, y, z)\Delta x\Delta y + \frac{1}{2} \frac{\partial \Lambda_z}{\partial z} \Delta x\Delta y\Delta z
\end{aligned} \tag{5.26}$$

因此上下表面的通量是和是

$$\frac{\partial \Lambda_z}{\partial z} \Delta x\Delta y\Delta z \tag{5.27}$$

类似的做法可以用到其他面上。长方体总的通量是

$$\Phi = \Delta x\Delta y\Delta z \left(\frac{\partial \Lambda_x}{\partial x} + \frac{\partial \Lambda_y}{\partial y} + \frac{\partial \Lambda_z}{\partial z} \right) \tag{5.28}$$

长方体体积为 $\Delta x\Delta y\Delta z$ 。所以通量对体积的比值是 $\partial \Lambda_x / \partial x + \partial \Lambda_y / \partial y + \partial \Lambda_z / \partial z$ 。于是

$$\operatorname{div} \vec{\Lambda} = \frac{\partial \Lambda_x}{\partial x} + \frac{\partial \Lambda_y}{\partial y} + \frac{\partial \Lambda_z}{\partial z} \tag{5.29}$$

定义梯度算符：

$$\vec{\nabla} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \tag{5.30}$$

这是一个矢量算符，即 $\vec{\nabla}$ 作用在一个标量函数上得到一个矢量。而

$$\vec{\nabla} \cdot \vec{\Lambda}(\vec{r}) = \frac{\partial \Lambda_x}{\partial x} + \frac{\partial \Lambda_y}{\partial y} + \frac{\partial \Lambda_z}{\partial z} \tag{5.31}$$

于是

$$\operatorname{div} \vec{\Lambda} = \nabla \cdot \vec{\Lambda} \tag{5.32}$$

积分形式

$$\int_S \vec{\Lambda} \cdot d\vec{a} = \int_V \nabla \cdot \vec{\Lambda} dV \tag{5.33}$$

6 场、势和能量

我们将讨论如何更好地描述相互作用。

6.1 场

用什么方式来描述相互作用具有很大的任意性。前面我们使用了力，可以方便一些问题的处理。但力往往意味着超距作用。一般地，力都不会有平方反比这样简单的关系。当相互作用的粒子运动时，力的形式会更复杂。例如，一个振荡的电荷产生的效应会传到很远处。过一段时间后，远处的另一个电荷才会感受到它。但这时，产生效应的电荷可能已经停止了运动。当然我们仍然可以用力来描述复杂的情形。但是，为此我们会付出较大的代价，力的形式会极为复杂。所以人们开始寻找更简洁方便的方法。

更成功的描述方式是粒子间通过场作为载体来传递相互作用。一个粒子在其周围建立起场，另外一个粒子感受到场并与之相互作用。核心的想法是将相互作用分为两部分，一部分是场的建立，另一部分是场与粒子的相互作用。我们不再采用瞬时的超距作用，相互作用仅存在于紧邻的点之间，即相互作用都是局域的。场的传播需要时间。一般地，在某一时刻一个粒子感受到的相互作用不由其他粒子的瞬时状态所决定。在这个意义下，场具有和粒子同样的真实性。在使用场描述相互作用时需要两个定律，一个是描述场的建立（场与源的关系），叫场方程，一个是描述粒子对场的响应，叫运动方程。在现代物理中，力的概念不存在，代之以场。

以引力为例，一个质量 M 建立的场是（场方程）

$$\vec{G} = -G \frac{M}{r^2} \hat{r} \quad (6.1)$$

运动方程是

$$m \frac{d^2 \vec{r}}{dt^2} = m \vec{G} \quad (6.2)$$

看似只是将同样的事情用另一种方式写，但当相互作用变得更复杂时，这种分离的方式变得更为现实。例如，描述电磁作用的Maxwell方程组：

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0} \quad (6.3)$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad (6.4)$$

$$\vec{\nabla} \cdot \vec{B} = 0 \quad (6.5)$$

$$\vec{\nabla} \times \vec{B} = \mu_0 \vec{j} + \epsilon_0 \mu_0 \frac{\partial \vec{E}}{\partial t} \quad (6.6)$$

以及运动方程

$$m\vec{a} = q(\vec{E} + \vec{v} \times \vec{B}) \quad (6.7)$$

这里 \vec{E} 和 \vec{B} 是电场和磁场。没有场的概念，超出静电力的电磁相互作用不可想象。

场满足叠加原理，所有的场源产生的总场是单个源的场之和，例如N个粒子总的引力场是

$$\vec{G} = \sum_i -G \frac{m_i}{r_i^2} \hat{r}_i \quad (6.8)$$

考虑最简单的一种一般情形。有一个场，强度为 $\vec{\Lambda}$ 。 $\vec{\Lambda}$ 只与场的源有关。物理原理中的场方程和运动方程分别写为

$$f(\vec{\Lambda}) = 0 \quad (6.9)$$

$$m \frac{d^2 \vec{r}}{dt^2} = \alpha \vec{\Lambda} \quad (6.10)$$

α 叫做耦合常数。在讨论引力问题时，耦合常数是引力质量。电磁作用中耦合常数是电荷。耦合常数描述粒子对场的反应（事实上也描述了建立的场的强度）。这里采用了一种最简单的方式使场与粒子耦合，即由一个参量和场本身的乘积给出。

6.2 保守场

有这样一种场，它对任何闭合回路的积分都为0，称为保守场。例如引力场

$$\oint \vec{G} \cdot d\vec{r} = 0 \quad (6.11)$$

$d\vec{r}$ 是沿一路径的积分元。可以很简单地证明

$$\oint -G \frac{M}{r^2} \hat{r} \cdot d\vec{r} = \oint -G \frac{M}{r^2} dr = 0 \quad (6.12)$$

一般地， $\oint \vec{\Lambda} \cdot d\vec{r} = 0$ ，如果 Λ 是保守场。一个直接推论是： $\vec{\Lambda} \cdot d\vec{r}$ 的积分只依赖于路径的起点和终点，而不依赖于路径。如图6.2

$$\int_{A,1}^B \vec{\Lambda} \cdot d\vec{r} + \int_{B,2}^A \vec{\Lambda} \cdot d\vec{r} = 0 \quad (6.13)$$

$$\Rightarrow \int_{A,1}^B \vec{\Lambda} \cdot d\vec{r} = \int_{A,2}^B \vec{\Lambda} \cdot d\vec{r} \quad (6.14)$$

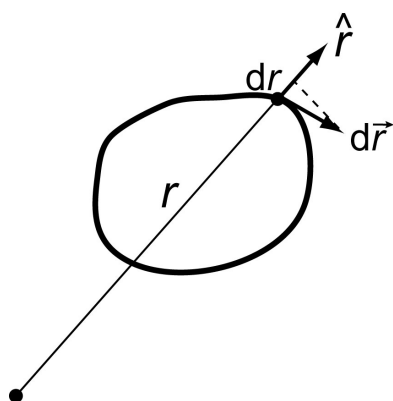


图 6.1: 对引力场的回路积分

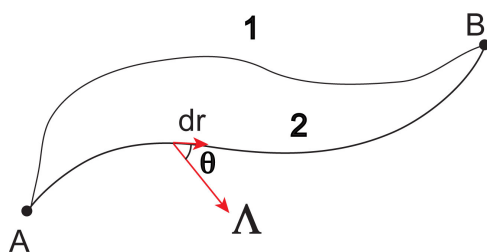


图 6.2: 保守场的积分不依赖于路径

这是保守场的充分必要条件。

一个非保守场的例子

二维空间

$$\vec{\Lambda}(\vec{r}) = \frac{A}{r} \hat{\theta} \quad (6.15)$$

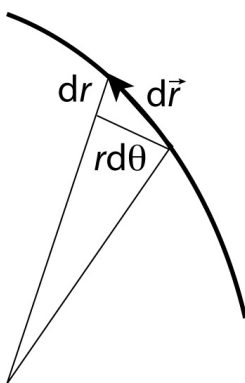
$$\vec{\Lambda}(\vec{r}) \cdot d\vec{r} = \frac{A}{r} \hat{\theta} \cdot d\vec{r} \quad (6.16)$$

$$d\vec{r} = dr \hat{r} + r d\theta \hat{\theta} \quad (6.17)$$

$$\vec{\Lambda}(\vec{r}) \cdot d\vec{r} = \frac{A}{r} r d\theta = A d\theta \quad (6.18)$$

$$\int_1^2 \vec{\Lambda}(\vec{r}) \cdot d\vec{r} = A(\theta_2 - \theta_1) \quad (6.19)$$

如果积分的闭合回路不包含原点，积分为0，如果包含原点，积分为 2π 。



另一个非保守场的例子, $\vec{\Lambda}(\vec{r}) = A(xy\hat{i} + y^2\hat{j})$ 。

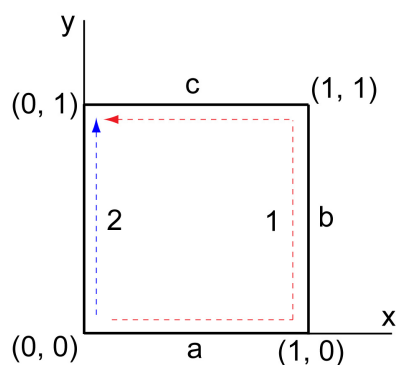


图 6.3: 非保守场的积分依赖于路径

第一条路径 $(0,0) \xrightarrow{a} (1,0) \xrightarrow{b} (1,1) \xrightarrow{c} (0,1)$ 。

积分是 $\int_1 \vec{\Lambda} \cdot d\vec{r} = \int_a \vec{\Lambda} \cdot d\vec{r} + \int_b \vec{\Lambda} \cdot d\vec{r} + \int_c \vec{\Lambda} \cdot d\vec{r} = -\frac{A}{6}$ 。

第二条路径 $(0,0) \rightarrow (0,1)$

积分是 $\int_2 \vec{\Lambda} \cdot d\vec{r} = \frac{A}{3}$ 。

6.3 旋度

我们需要一种更方便的方法判断保守场。积分的方法被微分取代, 将问题转化为只与空间的一点有关。

一个闭合回路包围着一个曲面。可以将面分成多个小片。类似于散度的讨论

$$\oint \vec{\Lambda} \cdot d\vec{r} = \sum_{i=1}^N \oint_{c_i} \vec{\Lambda} \cdot d\vec{r} \quad (6.20)$$

这提示环路积分与面积的比值（见下式）在面积趋于0时的极限应该存在。

$$\lim_{a_i \rightarrow 0} \frac{\oint_{c_i} \vec{\Lambda} \cdot d\vec{r}}{a_i} \quad (6.21)$$

这个极限应该依赖于面积元的指向 \hat{n} 。 \hat{n} 的方向按右手定则确定。

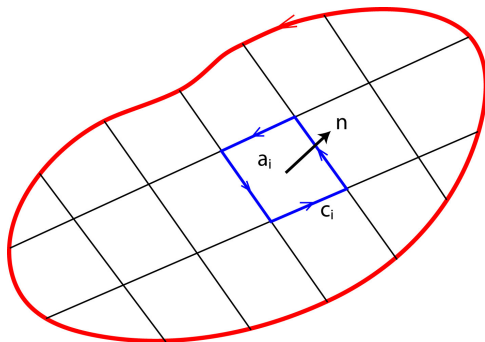


图 6.4: 矢量场对路径的积分

可以定义一个量叫做旋度， $\text{curl} \vec{\Lambda}$ ，它在 \hat{n} 上的投影是

$$\text{curl} \vec{\Lambda} \cdot \hat{n} = \lim_{a_i \rightarrow 0} \frac{\oint_{c_i} \vec{\Lambda} \cdot d\vec{r}}{a_i} \quad (6.22)$$

旋度代表环流密度，表示矢量场在空间中旋转的趋势。我们后面将看到旋度是矢量。利用旋度可以得到Stokes定理

$$\begin{aligned} \oint \vec{\Lambda} \cdot d\vec{r} &= \lim_{a_i \rightarrow 0} \sum_{i=1}^N \oint_{c_i} \vec{\Lambda} \cdot d\vec{r} \\ &= \lim_{a_i \rightarrow 0} \sum_{i=1}^N a_i \frac{\oint_{c_i} \vec{\Lambda} \cdot d\vec{r}}{a_i} \\ &= \lim_{a_i \rightarrow 0} \sum_{i=1}^N a_i \text{curl} \vec{\Lambda} \cdot \hat{n} \\ &= \int_S \text{curl} \vec{\Lambda} \cdot d\vec{a} \end{aligned} \quad (6.23)$$

这样一个矢量对一个回路的积分就表示成为了另一个矢量对该回路张成的任意面的面积分。一个矢量对任意回路的积分为0就意味着该矢量在每一点的旋度都是0，反之亦然。保守场的旋度处处为0。保守则无旋。这就为保守场的判断提供了一个方便的方法，因为旋度只和局域的性质有关。我们还需要给出一个方便的计算旋度的方法。我们将证明

$$\text{curl} \vec{\Lambda} = \nabla \times \vec{\Lambda} \quad (6.24)$$

其中

$$\nabla \times \vec{\Lambda} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \partial_x & \partial_y & \partial_z \\ \Lambda_x & \Lambda_y & \Lambda_z \end{vmatrix} \quad (6.25)$$

为证明上面的关系，先取一个正方形面积元（如图6.5）， $\hat{n} = \hat{k}$ ，包围该面积的回路为 c ，求 $(\text{curl}\vec{\Lambda})_z$ 。

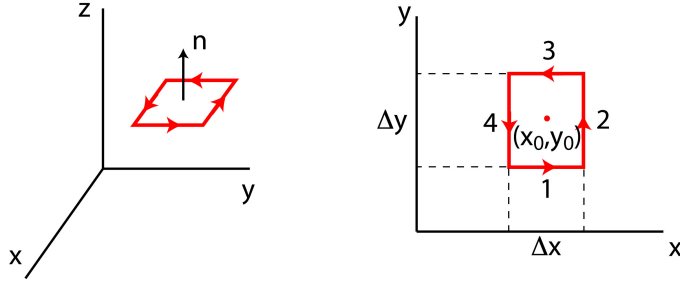


图 6.5: 计算旋度的z分量

$$(\text{curl}\vec{\Lambda})_z = \lim_{a \rightarrow 0} \frac{\oint_c \vec{\Lambda} \cdot d\vec{r}}{a} \quad (6.26)$$

第一段积分

$$\int_1 \Lambda_x(x, y_0 - \frac{1}{2}\Delta y, z_0) dx \quad (6.27)$$

其中

$$\begin{aligned} & \Lambda_x(x, y_0 - \frac{1}{2}\Delta y, z_0) \\ &= \Lambda_x(x_0 + (x - x_0), y_0 - \frac{1}{2}\Delta y, z_0) \\ &= \Lambda_x(x_0, y_0, z_0) + (x - x_0) \frac{\partial}{\partial x} \Lambda_x(x_0, y_0, z_0) - \frac{1}{2}\Delta y \frac{\partial}{\partial y} \Lambda_x(x_0, y_0, z_0) \end{aligned} \quad (6.28)$$

积分

$$\begin{aligned} & \int_1 \Lambda_x(x, y_0 - \frac{1}{2}\Delta y, z_0) dx \\ &= \Lambda_x(x_0, y_0, z_0) \Delta x + 0 - \frac{1}{2}\Delta x \Delta y \frac{\partial}{\partial y} \Lambda_x(x_0, y_0, z_0) \end{aligned} \quad (6.29)$$

第三段

$$\begin{aligned} & \int_3 \Lambda_x(x, y_0 + \frac{1}{2}\Delta y, z_0) dx \\ &= -\Lambda_x(x_0, y_0, z_0) \Delta x - \frac{1}{2}\Delta x \Delta y \frac{\partial}{\partial y} \Lambda_x(x_0, y_0, z_0) \end{aligned} \quad (6.30)$$

1+3得到

$$-\Delta x \Delta y \frac{\partial}{\partial y} \Lambda_x(x_0, y_0, z_0) \quad (6.31)$$

类似地，2+4得到

$$\Delta x \Delta y \frac{\partial}{\partial x} \Lambda_y(x_0, y_0, z_0) \quad (6.32)$$

所以

$$(\text{curl}\vec{\Lambda})_z = \Delta x \Delta y \left(\frac{\partial \Lambda_y}{\partial x} - \frac{\partial \Lambda_x}{\partial y} \right) / \Delta x \Delta y = \frac{\partial \Lambda_y}{\partial x} - \frac{\partial \Lambda_x}{\partial y} \quad (6.33)$$

类似地

$$(\text{curl}\vec{\Lambda})_x = \Delta y \Delta z \left(\frac{\partial \Lambda_z}{\partial y} - \frac{\partial \Lambda_y}{\partial z} \right) / \Delta y \Delta z = \frac{\partial \Lambda_z}{\partial y} - \frac{\partial \Lambda_y}{\partial z} \quad (6.34)$$

$$(\text{curl}\vec{\Lambda})_y = \Delta x \Delta z \left(\frac{\partial \Lambda_x}{\partial z} - \frac{\partial \Lambda_z}{\partial x} \right) / \Delta x \Delta z = \frac{\partial \Lambda_x}{\partial z} - \frac{\partial \Lambda_z}{\partial x} \quad (6.35)$$

所以

$$\text{curl}\vec{\Lambda} = \left(\frac{\partial \Lambda_z}{\partial y} - \frac{\partial \Lambda_y}{\partial z} \right) \hat{i} + \left(\frac{\partial \Lambda_x}{\partial z} - \frac{\partial \Lambda_z}{\partial x} \right) \hat{j} + \left(\frac{\partial \Lambda_y}{\partial x} - \frac{\partial \Lambda_x}{\partial y} \right) \hat{k} \quad (6.36)$$

回忆 $\vec{\nabla}$ 的定义, 我们有

$$\text{curl}\vec{\Lambda} = \nabla \times \vec{\Lambda} \quad (6.37)$$

于是旋度是一个矢量, 它的方向垂直于环流面。

Stokes定理也可写为

$$\oint \vec{\Lambda} \cdot d\vec{r} = \int_S (\nabla \times \vec{\Lambda}) \cdot d\vec{a} \quad (6.38)$$

有了旋度的微分形式就可以方便地计算旋度, 例如平方反比力的场

$$\nabla \times \frac{\vec{r}}{r^3} = \left(\frac{\partial_y z}{r^3} - \frac{\partial_z y}{r^3} \right) \hat{i} + \left(\frac{\partial_z x}{r^3} - \frac{\partial_x z}{r^3} \right) \hat{j} + \left(\frac{\partial_x y}{r^3} - \frac{\partial_y x}{r^3} \right) \hat{k} \quad (6.39)$$

其中

$$\frac{\partial_y z}{r^3} = -\frac{3z}{r^4} \frac{\partial r}{\partial y} = -\frac{3z}{2r^5} \frac{\partial r^2}{\partial y} = -\frac{3yz}{r^5} \quad (6.40)$$

$$\frac{\partial_z y}{r^3} = \frac{3yz}{r^5} \quad (6.41)$$

等等。容易证明旋度为0。

6.4 势

场 Λ 本身没有隐含保守性, 那么如何才能自动保证保守性? 由于场的路径积分不依赖于路径, 所以我们可以定义一个函数 $\phi(\vec{r})$ 叫势, 满足

$$\phi(B) - \phi(A) = - \int_A^B \vec{\Lambda} \cdot d\vec{r} \quad (6.42)$$

积分可以选择从A到B的任何路径。我们可以选任何一点作为 ϕ 的零点。如果把势作为出发点，就可以自动保证保守性。

例：引力场的势

取无限远处为0, $\phi(\infty) = 0$

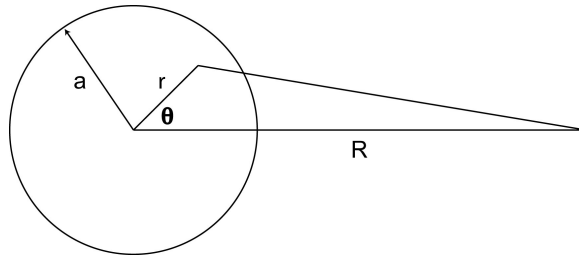
$$\begin{aligned}\phi(r) &= \phi(\infty) - \int_{\infty}^r -\frac{GM}{r^2} \hat{r} \cdot d\vec{r} \\ &= - \int_r^{\infty} \frac{GM}{r^2} dr \\ &= \frac{GM}{r} \Big|_r^{\infty} = -\frac{GM}{r}\end{aligned}\quad (6.43)$$

例：均匀场的势

$$\vec{\Lambda} = -g\hat{k} \quad (6.44)$$

$$\phi(z) - \phi(0) = - \int_0^z (-g) dz = gz \quad (6.45)$$

例：均匀球体的引力势



当 $R > a$ 时

$$\begin{aligned}\phi(R) &= \int -\frac{G}{r'} \rho r^2 \sin \theta dr d\theta d\varphi \\ &= -G\rho \int r^2 dr d\theta d\varphi \frac{\sin \theta}{\sqrt{R^2 + r^2 - 2Rr \cos \theta}} \\ &= -2\pi G\rho \int_0^a dr \int_0^\pi d\theta \frac{r^2 \sin \theta}{\sqrt{R^2 + r^2 - 2Rr \cos \theta}} \\ &= -2\pi G\rho \int_0^a dr \int_0^\pi \frac{r}{2R} \frac{d(-2Rr \cos \theta)}{\sqrt{R^2 + r^2 - 2Rr \cos \theta}} \\ &= -2\pi G\rho \int_0^a dr \frac{r}{2R} \cdot 2 \sqrt{R^2 + r^2 - 2Rr \cos \theta} \Big|_0^\pi\end{aligned}$$

$$\begin{aligned}
&= -2\pi G\rho \int_0^a dr \frac{r}{R} (|R+r| - |R-r|) \\
&= -2\pi G\rho \int_0^a \frac{2r^2}{R} dr \\
&= -\frac{4\pi G\rho}{R} \cdot \frac{1}{3} a^3 \\
&= -\frac{4\pi a^3}{3} \rho \frac{G}{R} = -\frac{GM}{R}
\end{aligned} \tag{6.46}$$

类似，当 $R < a$ 时，也可以求出 $\phi(R)$ 。

利用积分可以从场得到势。那么如何从势得到场？考虑邻近两点间势的差别

$$\begin{aligned}
\Delta\phi &= \phi(\vec{r} + \Delta\vec{r}) - \phi(\vec{r}) \\
&= - \int_{\vec{r}}^{\vec{r} + \Delta\vec{r}} \vec{\Lambda} \cdot d\vec{r} \\
&= -\vec{\Lambda} \cdot \Delta\vec{r} \\
&= -(\Lambda_x \Delta x + \Lambda_y \Delta y + \Lambda_z \Delta z)
\end{aligned} \tag{6.47}$$

同时我们还知道

$$\Delta\phi = \frac{\partial\phi}{\partial x} \Delta x + \frac{\partial\phi}{\partial y} \Delta y + \frac{\partial\phi}{\partial z} \Delta z \tag{6.48}$$

所以 $\vec{\Lambda} = -\vec{\nabla}\phi$ 。我们可以用梯度从势得到场。利用上一章梯度的性质，我们有

- 场垂直于等势面
- 场从势高的地方指向势低的地方
- 场的方向是势变化最快的方向
- 等势面越近场越强

$\vec{\Lambda}$ 和 ϕ 可以相互推出，但 ϕ 更方便，因为 ϕ 的存在已经隐含了 $\vec{\Lambda}$ 是保守场。势自动体现了保守性。容易证明 $\vec{\nabla} \times (\vec{\nabla}\phi) = 0$ 。

$$(\vec{\nabla} \times (\vec{\nabla}\phi))_x = \frac{\partial}{\partial y} \frac{\partial}{\partial z} \phi - \frac{\partial}{\partial z} \frac{\partial}{\partial y} \phi = 0 \tag{6.49}$$

$$(\vec{\nabla} \times (\vec{\nabla}\phi))_y = \frac{\partial}{\partial z} \frac{\partial}{\partial x} \phi - \frac{\partial}{\partial x} \frac{\partial}{\partial z} \phi = 0 \tag{6.50}$$

$$(\vec{\nabla} \times (\vec{\nabla}\phi))_z = \frac{\partial}{\partial x} \frac{\partial}{\partial y} \phi - \frac{\partial}{\partial y} \frac{\partial}{\partial x} \phi = 0 \tag{6.51}$$

6.5 保守系统

我们已经讨论完了场方程，下面讨论运动方程。

运动方程

$$m \frac{d^2 \vec{r}}{dt^2} = \alpha \vec{\Lambda} \quad (6.52)$$

我们希望将 $\vec{\Lambda}$ 换成更基本的 ϕ 。运动方程两边对一路径积分

$$\int_A^B m \frac{d^2 \vec{r}}{dt^2} \cdot d\vec{r} = \alpha \int_A^B \vec{\Lambda} \cdot d\vec{r} \quad (6.53)$$

$$\int_A^B m \frac{d}{dt} \left(\frac{d\vec{r}}{dt} \right) \cdot \frac{d\vec{r}}{dt} dt = -\alpha(\phi(B) - \phi(A)) \quad (6.54)$$

$$\begin{aligned} \text{左边} &= \frac{1}{2} \int_A^B m \frac{d}{dt} \left(\frac{d\vec{r}}{dt} \right)^2 dt \\ &= \frac{m}{2} \int_A^B d \left(\frac{d\vec{r}}{dt} \right)^2 \\ &= \frac{1}{2} m v_B^2 - \frac{1}{2} m v_A^2 \end{aligned} \quad (6.55)$$

$$\Rightarrow \frac{1}{2} m v_B^2 - \frac{1}{2} m v_A^2 = -\alpha(\phi(B) - \phi(A)) \quad (6.56)$$

$$\Rightarrow \frac{1}{2} m v_A^2 + \alpha\phi(A) = \frac{1}{2} m v_B^2 + \alpha\phi(B) \quad (6.57)$$

能量定义为 $E = T + U$ ，其中 $T = \frac{1}{2} m v^2$ 叫动能， $U = \alpha\phi$ 叫势能。在保守场中，如果采用6.1式这用最简单的耦合方式，则可以定义一个函数称为能量，并且能量守恒。这样的系统称为保守系统。对于一些所谓的非保守系统，如有摩擦力的系统，事实上在微观上的仍是保守场在传递相互作用，只是没有将微观上原子的所有运动都考虑进来。牛顿运动方程可以描述所有的事情而不必引入能量。能量的出现本质上是用场和势来描述相互作用的必然结果。

力可以从势能得到

$$\vec{F} = \alpha \vec{\Lambda} = -\vec{\nabla} U \quad (6.58)$$

$$\frac{1}{2} m v_B^2 - \frac{1}{2} m v_A^2 = -(U(B) - U(A)) = \int_A^B \vec{F} \cdot d\vec{r} \quad (6.59)$$

势能也可以从力得到

$$U(B) - U(A) = -\alpha \int_A^B \vec{\Lambda} \cdot d\vec{r} = -\int_A^B \vec{F} \cdot d\vec{r} \quad (6.60)$$

功定义为 $\int_A^B \vec{F} \cdot d\vec{r}$, 因而 $W = T_B - T_A$, 叫功能定理, 场对粒子作的功等于粒子动能的增加。

功率定义为

$$P = \frac{dW}{dt} = \vec{F} \cdot \frac{d\vec{r}}{dt} = \vec{F} \cdot \vec{v} \quad (6.61)$$

势能的例子

引力势能

$$U(r) = -G \frac{mM}{r} \quad (6.62)$$

对于非基本的力, 由于叠加原理, 仍然可以用势和能量来唯象地描写, 如弹性力

$$\vec{F}(\vec{r}) = -k(r - r_0)\hat{r} \quad (6.63)$$

势能

$$\begin{aligned} U(r) - U(0) &= - \int_{r_0}^r -k(r - r_0) dr \\ &= \frac{1}{2}k (r - r_0)^2 \Big|_{r_0}^r \\ &= \frac{1}{2}k(r - r_0)^2 \end{aligned} \quad (6.64)$$

令 $U(r_0) = 0$, 则 $U(r) = \frac{1}{2}k(r - r_0)^2$ 。

6.6 一个更一般的框架

我们需要寻找一个使用势并可以替代牛顿方程的方案。

更一般地, 运动方程可以有如下形式, 称为哈密顿正则方程

$$\begin{cases} \dot{p}_i = -\frac{\partial H}{\partial q_i} \\ \dot{q}_i = \frac{\partial H}{\partial p_i} \end{cases} \quad (6.65)$$

p_i, q_i 称为广义动量、广义坐标。对于守恒系统, 哈密顿量 H 在势能不依赖于速度时就是系统的能量 $H = T + U$ 。可以证明, 哈密顿方程与牛顿方程等价。由所有广义动量和广义坐标张成的空间称为相空间。系统的状态由相空间中的一点代表, 其轨迹刻画系统的运动。

使用场和势的概念有重要优势。在量子力学中, 我们会看到它们更重要的体现。在更深的层次上, 力消失了, 但势存在下来。

6.7 非保守场

保守场导致了保守系统。但保守系统不一定要基于保守场。例如，考虑了全部的电磁效应（非静电）后，电场和磁场不再是保守场。但仍然可以引入势并将Maxwell方程组等效地写为：

$$\nabla^2\phi - \frac{1}{c^2} \frac{\partial^2\phi}{\partial t^2} = -\rho/\epsilon_0 \quad (6.66)$$

$$\nabla^2\vec{A} - \frac{1}{c^2} \frac{\partial^2\vec{A}}{\partial t^2} = -\mu_0\vec{j} \quad (6.67)$$

其中 c 是光速。 \vec{A} 和 ϕ 称为矢量势和标量势，与电场和磁场的关系是 $\vec{E} = -\nabla\phi - \partial\vec{A}/\partial t$ 和 $\vec{B} = \nabla \times \vec{A}$ ，这并不能保证保守场。历史上，势的引入可自动保证某些场的保守性，但势最本质的重要性在于它能彻底地保证相互作用的局域性，而场不能，例如AB效应。

尽管电磁场不再是保守场，但我们可以通过定义场的能量使得场与粒子一起构成的体系的总能量守恒。一般地，基于电磁作用的系统是保守系统。

7 谐振子

很多小振动问题可以归结为谐振子运动。例如，图7.1是双原子分子的势能曲线。 $r \rightarrow \infty$ 时 $U \rightarrow 0$ 。 $E < 0$ ($U + K < 0$) 时是束缚态。 $E > 0$ 时，非束缚态，化学键断裂。

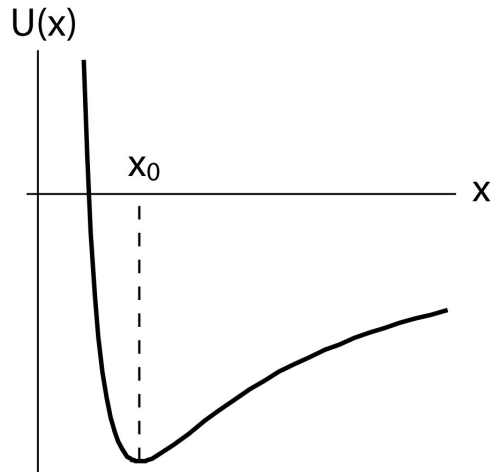


图 7.1: 双原子分子势能图

在平衡点附近分子发生小振动。将势能在平衡点附近进行泰勒展开

$$U(x) = U(x_0) + (x - x_0) \left. \frac{dU}{dx} \right|_{x=x_0} + \frac{1}{2}(x - x_0)^2 \left. \frac{d^2U}{dx^2} \right|_{x=x_0} + \dots \quad (7.1)$$

在平衡点 $x = x_0$ ， $dU/dx = 0$ 。在小振幅时，

$$U(x) = U_0 + \frac{1}{2}(x - x_0)^2 \frac{d^2U}{dx^2} \quad (7.2)$$

这就是一个谐振子的势能曲线。

7.1 无阻尼谐振子

谐振子的运动方程是

$$m\ddot{x} + kx = 0 \quad (7.3)$$

解为

$$x = B \cos \omega_0 t + C \sin \omega_0 t \quad (7.4)$$

$$\omega_0 = \sqrt{\frac{k}{m}} \quad (7.5)$$

B和C由初始位置和速度决定。 ω_0 是角频率，周期 $T = 2\pi/\omega_0$ 。

通常解还可写为：

$$x = A \cos(\omega_0 t + \varphi) \quad (7.6)$$

A叫做振幅， φ 是相角， $A = \sqrt{B^2 + C^2}$ ， $\tan \varphi = -C/B$ 。

解也可以写为：

$$x = C_1 e^{i\omega_0 t} + C_2 e^{-i\omega_0 t} \quad (7.7)$$

由于 x 为实， $C_1^* = C_2$ 。

能量：

$$\text{势能： } U = \frac{1}{2} k x^2 = \frac{1}{2} k A^2 \cos^2(\omega_0 t + \varphi)$$

$$\text{动能： } K = \frac{1}{2} m v^2 = \frac{1}{2} m \omega_0^2 A^2 \sin^2(\omega_0 t + \varphi)$$

$$E = K + U = \frac{1}{2} k A^2 \quad (7.8)$$

7.2 阻尼简谐运动

$$m\ddot{x} = -kx - b\dot{x} \quad (7.9)$$

$$\ddot{x} + \gamma\dot{x} + \omega_0^2 x = 0 \quad (7.10)$$

其中 $\gamma = b/m$ ， $\omega_0^2 = k/m$ 。

弱阻尼的解：

$$x = A e^{-\gamma t/2} \cos(\omega t + \varphi) \quad (7.11)$$

其中

$$\omega = \sqrt{\omega_0^2 - \frac{\gamma^2}{4}} \quad (7.12)$$

弱阻尼的条件是 $\omega_0 \gg \gamma$ 。这时振幅随时间变化很小，在振幅发生显著变化之前已经振荡多次。

动能：

$$K(t) = \frac{1}{2} m \omega^2 A^2 e^{-\gamma t} \sin^2(\omega t + \varphi) + \text{高阶小量} \quad (7.13)$$

势能：

$$U(t) = \frac{1}{2} k A^2 e^{-\gamma t} \cos^2(\omega t + \varphi) \quad (7.14)$$

总能量：

$$E(t) = \frac{1}{2} A^2 e^{\gamma t} (m \omega^2 \sin^2(\omega t + \varphi) + k \cos^2(\omega t + \varphi))$$

$$\begin{aligned}
 &= \frac{1}{2}A^2e^{-\gamma t}[k \cos^2(\omega t + \varphi) + k \sin^2(\omega t + \varphi)] \\
 &= \frac{1}{2}kA^2e^{-\gamma t}
 \end{aligned} \tag{7.15}$$

其中用到 $\omega \simeq \omega_0$ 。于是

$$E(t) = E(0)e^{-\gamma t} = E_0e^{-t/\tau} \tag{7.16}$$

$\tau = 1/\gamma$ 叫阻尼时间。

Q因子

Q因子用来表征阻尼的程度

$$Q = 2\pi \frac{\text{储藏在振子中的能量}}{\text{一个周期内损失的能量}} \tag{7.17}$$

高Q的谐振子意味着低阻尼。在低阻尼条件下

$$\frac{dE}{dt} = -\gamma E_0e^{-\gamma t} = -\gamma E \tag{7.18}$$

一个周期内损失的能量是

$$\Delta E \simeq \left| \frac{dE}{dt} \right| \cdot \frac{2\pi}{\omega} = \gamma E \frac{2\pi}{\omega} \tag{7.19}$$

所以

$$Q = 2\pi E \bigg/ \frac{2\pi\gamma E}{\omega} \simeq \frac{\omega_0}{\gamma} \tag{7.20}$$

7.3 受迫振动

简谐振动是很多线性系统的模型。 ω_0 和 γ 是一个线性系统的两个参量。那么如果给定了一个线性系统，如何测量这两个参量？往往产生周期运动比直接测量周期运动容易（这往往需要时间分辨的测量手段）。一个简单的做法是对系统施加外力，看系统的反应，称为受迫振动。

$$m\ddot{x} + kx = F_0 \cos \omega t \tag{7.21}$$

解为

$$x = A \cos \omega t \tag{7.22}$$

其中

$$A = \frac{F_0}{m} \frac{1}{\omega_0^2 - \omega^2} \quad (7.23)$$

$\omega \rightarrow \infty$ 时, $A \rightarrow 0$ 。 $\omega \rightarrow \omega_0$ 时, $A \rightarrow \infty$ 。 在 $\omega = \omega_0$ 处发生共振。 $\omega = \omega_0$ 时 $A \rightarrow \infty$ 的原因是没有考虑阻尼。 将 x 改写为另一种形式 [见图7.2] ,

$$x = |A| \cos(\omega t + \phi) \quad (7.24)$$

其中

$$\phi = \begin{cases} 0, & \omega < \omega_0 \\ -\pi, & \omega > \omega_0 \end{cases} \quad (7.25)$$

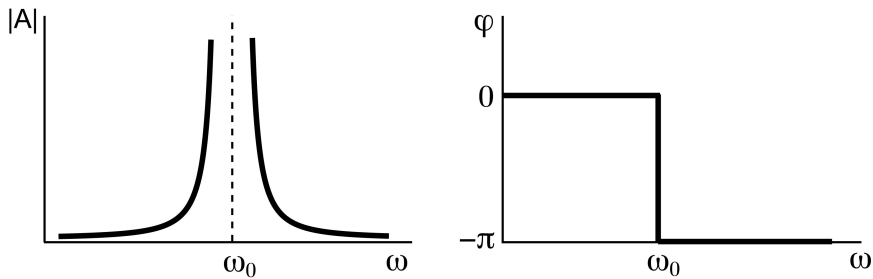


图 7.2: 无阻尼受迫振动

在共振处 $|A|$ 趋于无穷大, 而相角改变 π 。 当考虑了阻尼后, 在共振时, $|A|$ 不再为无穷大, 相角则由0渐变为 $-\pi$ 。

有阻尼时

$$\ddot{x} + \gamma \dot{x} + \omega_0^2 x = \frac{F_0}{m} \cos \omega t \quad (7.26)$$

$$x = A \cos(\omega t + \varphi) \quad (7.27)$$

$$A = \frac{F_0}{m} \frac{1}{[(\omega_0^2 - \omega^2)^2 + (\omega\gamma)^2]^{1/2}} \quad (7.28)$$

$$\varphi = \tan^{-1} \left(\frac{\gamma\omega}{\omega^2 - \omega_0^2} \right) \quad (7.29)$$

如图7.3, 当 $\omega = \omega_0$ 时发生共振, 振幅为 $A = \frac{F_0}{m\omega_0\gamma}$ 。 在 $\gamma \rightarrow 0$ 时 $A \rightarrow \infty$ 。 相角在共振附近由0渐变为 $-\pi$ 。 事实上, A极大值并不是精确出现在 ω_0 , 而是 $\omega^2 = \omega_0^2 - \frac{\gamma^2}{2}$ 。 当 γ/ω 很小时 $\omega \simeq \omega_0$ 。

相角 φ 代表驱动力与位移之间的关系。 一般地, 位移的相位落后于驱动力 $|\varphi|$ 相位 [图7.4]。 当驱动力频率远远低于共振频率时, 位移只是略微

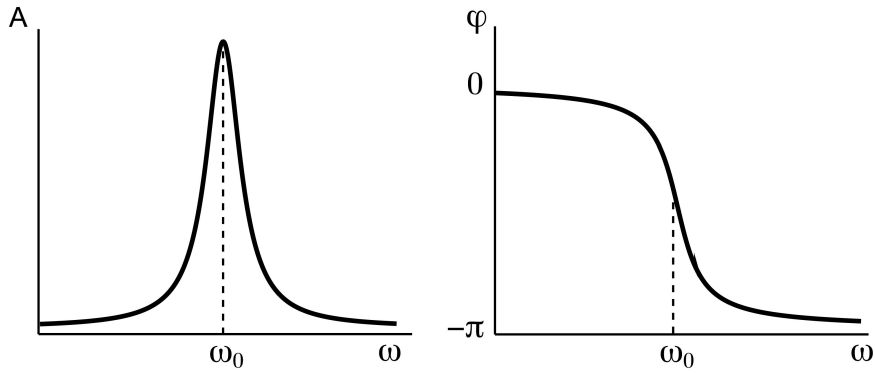


图 7.3: 阻尼受迫振动

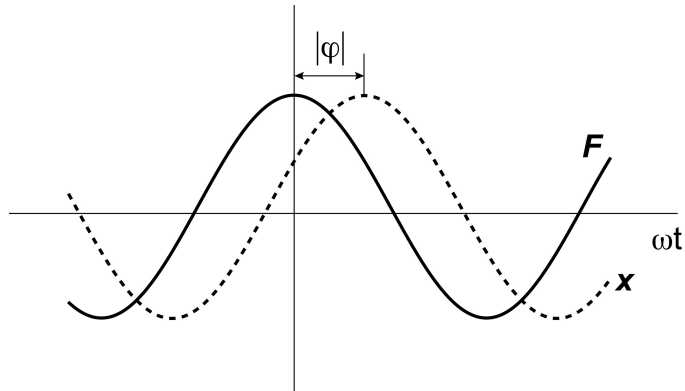


图 7.4: 位移与驱动力间的相位关系

落后于驱动力。在共振时，位移落后相位 $\pi/2$ 。驱动频率远远高于共振频率时，位移与驱动力几乎反相。

我们下面研究谐振子中的能量和能量转化。

$$x = A \cos(\omega t + \varphi) \quad (7.30)$$

$$v = -\omega A \sin(\omega t + \varphi) \quad (7.31)$$

所以

$$K(t) = \frac{1}{2} m \omega^2 A^2 \sin^2(\omega t + \varphi) \quad (7.32)$$

$$U(t) = \frac{1}{2} k A^2 \cos^2(\omega t + \varphi) \quad (7.33)$$

注意这里不再有 $\omega^2 = k/m$ 。有驱动时能量不守恒。我们需要对一个周期求平均，并用到

$$\langle \sin^2(\omega t + \varphi) \rangle = \langle \cos^2(\omega t + \varphi) \rangle = \frac{1}{2} \quad (7.34)$$

于是

$$\langle K \rangle = \frac{1}{4} m \omega^2 A^2 \quad (7.35)$$

$$\langle U \rangle = \frac{1}{4}kA^2 \quad (7.36)$$

$$\langle E \rangle = \frac{A^2}{4}(m\omega^2 + k) = \frac{1}{4}mA^2(\omega^2 + \omega_0^2) \quad (7.37)$$

代入A的表达式

$$\langle E(\omega) \rangle = \frac{1}{4} \frac{F_0^2}{m} \frac{\omega^2 + \omega_0^2}{(\omega_0^2 - \omega^2)^2 + (\omega\gamma)^2} \quad (7.38)$$

还可以计算出一个周期中的能量耗散

$$\begin{aligned} & \frac{1}{T} \int_0^T b\dot{x}^2 dt \\ &= \frac{1}{T} \int_0^T \gamma m (-A\omega \sin(\omega t + \varphi))^2 dt \\ &= \frac{\gamma m \omega^2 A^2}{T} \int_0^T \sin^2(\omega t + \varphi) dt \\ &= \frac{1}{2} \gamma m \omega^2 A^2 \end{aligned} \quad (7.39)$$

能量的耗散由驱动力做功来补偿。可以证明二者相等。驱动力功率的计算如下。

$$\begin{aligned} P &= F \cdot v \\ &= F_0 \cos \omega t \cdot (-\omega A) \sin(\omega t + \varphi) \\ &= -\omega F_0 A \cos \omega t \sin(\omega t + \varphi) \end{aligned} \quad (7.40)$$

对一个周期进行平均，可得到平均功率，

$$\begin{aligned} \langle P \rangle &= -\omega F_0 A \langle \cos \omega t \sin(\omega t + \varphi) \rangle \\ &= -\omega F_0 A \langle \cos \omega t \sin \omega t \cos \varphi + \cos^2 \omega t \sin \varphi \rangle \\ &= -\omega F_0 A (\langle \cos \omega t \sin \omega t \cos \varphi \rangle + \langle \cos^2 \omega t \rangle \sin \varphi) \\ &= -\frac{1}{2} \omega F_0 A \sin \varphi \end{aligned} \quad (7.41)$$

其中用到了 $\langle \cos \omega t \sin \omega t \rangle = 0$ 和 $\langle \cos^2 \omega t \rangle = 1/2$ 。为了与能量耗散比较，需要求 $\sin \varphi$ 。

$$\tan \varphi = \frac{\gamma \omega}{\omega^2 - \omega_0^2}, \quad \varphi \in (-\pi, 0) \quad (7.42)$$

所以

$$\sin \varphi = \frac{-\gamma \omega}{\sqrt{(\omega^2 - \omega_0^2)^2 + (\gamma \omega)^2}} \quad (7.43)$$

于是

$$\langle P \rangle = \frac{1}{2} \omega F_0 A \frac{\gamma \omega}{\sqrt{(\omega^2 - \omega_0^2)^2 + (\gamma \omega)^2}}$$

$$\begin{aligned}
 &= \frac{1}{2} \gamma m \omega^2 \frac{A F_0}{m} \frac{1}{[(\omega_0^2 - \omega^2)^2 + (\omega \gamma)^2]^{\frac{1}{2}}} \\
 &= \frac{1}{2} \gamma m \omega^2 A^2 = \text{能量耗散} \quad (7.44)
 \end{aligned}$$

驱动力对谐振子能量的输入应该等于谐振子本身能量的耗散。测量中往往测功率更方便，而不是直接测A。由功率与频率的关系的测量可以得到 ω_0 和 γ 。

7.4 共振

在 $\omega = \omega_0$ 时，A取最大值， $\sin(-\frac{\pi}{2}) = -1$ ，因而共振时 $\langle P \rangle$ 达到最大。在 $\omega = \omega_0$ 处 φ 有很大变化： $0 \rightarrow -\pi$ 。相位 φ 告诉我们的是位移比力晚多少角度到达极大值。在 ω 小时，位移与驱动力同相。共振时， $\varphi = -\frac{\pi}{2}$ 。共振出现时 $\varphi \neq 0$ ，原因在于振子对驱动力的功率吸收并不直接依赖于与位移的位相差，而是依赖于与速度的位相差 [图7.5]。在共振处， $v = \omega A \sin \omega t$ 。当速度与驱动力同相时，将得到最大的功率吸收。这时，物体刚好在合适时间和地点被推动。位移为0时速度最大，同时这点的力也最大。在转折点，速度改变方向，力同时改变方向。相位 φ 发生 π 的变化是共振的一个重要特征。当 φ 为0和 π 时，没有能量吸收。

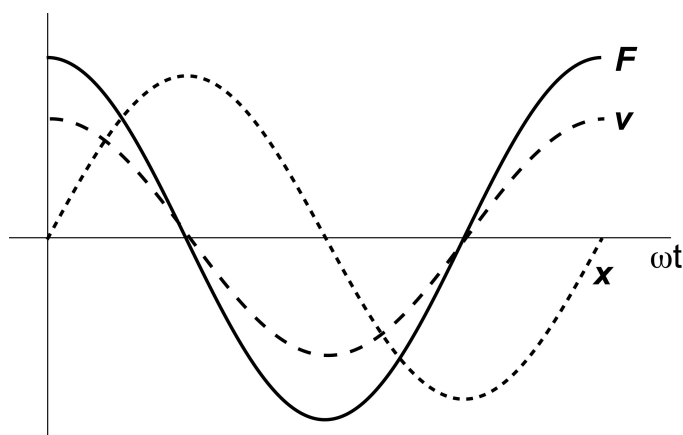


图 7.5: 共振时的驱动力、位移和速度的相位关系

$$\begin{aligned}
 P &= \frac{\gamma F_0^2}{2m} \frac{\omega^2}{(\omega_0^2 - \omega^2)^2 + (\omega \gamma)^2} \\
 &= \frac{\gamma F_0^2}{2m} \frac{1}{\left(\frac{\omega_0^2 - \omega^2}{\omega}\right)^2 + \gamma^2} \quad (7.45)
 \end{aligned}$$

当 $\gamma \ll \omega_0$ 时，吸收功率 P 只在共振处有显著的值。远离共振时 $P \simeq 0$ 。如果我们主要关心共振附近的吸收功率，在做近似后可以得到一个吸收功率的更

简单的结果。令 $\delta\omega = \omega - \omega_0$ 。展开到 $\delta\omega$ 的一阶，有

$$\left(\frac{\omega_0^2 - \omega^2}{\omega}\right)^2 \approx (2\delta\omega)^2 = 4(\omega - \omega_0)^2 \quad (7.46)$$

因而吸收功率可以写为

$$P = \frac{\gamma F_0^2}{8m} \frac{1}{(\omega - \omega_0)^2 + (\gamma/2)^2} \quad (7.47)$$

我们把这个共振曲线称为 Lorentz 线型 (图 7.6)。Lorentz 线型普遍存在于自然界中。其最大高度为 $4/\gamma^2$ 。在半高处, $(\omega - \omega_0) = \pm\gamma/2$, $\Delta\omega = \gamma$ 称为共振宽度或半高宽。这也是位相 ϕ 发生 π 的变化的特征宽度。阻尼的大小决定共振的强度。

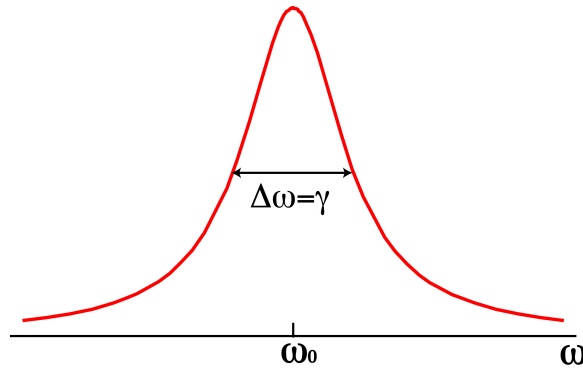


图 7.6: Lorentz 线型

利用共振曲线, Q 可以写为

$$Q = \frac{\text{共振频率}}{\text{共振宽度}} = \frac{\omega_0}{\gamma} \quad (7.48)$$

Q 实际上表征了系统的频率响应, Q 越大共振曲线越窄。

人们利用共振来探测系统的响应, 从而研究系统固有的频率和耗散。耗散可以用阻尼时间 (寿命) 来表达 $\Delta\omega = \gamma$, 而 $\tau = 1/\gamma$, 所以 $\tau\Delta\omega = 1$ 。寿命是系统内在的性质, 它决定了系统对外界的响应的共振宽度。因而共振可以测系统的固有频率和寿命。

共振极为重要, 因为提供了一种研究体系的方法。通常在实验中我们总是用一种外界的驱动力去激发一个系统, 然后看系统的反应, 从而得到线性系统的本征模式。在共振处, 驱动力被强烈地吸收例如用红外吸收谱测定分子振动。振动是分子的指纹, 不同的功能团有不同的振动频率。

8 有心力运动

8.1 两体问题约化为单体问题

考虑在有心力作用下的两个质点 m_1 和 m_2 ，相对位移 $\vec{r} = \vec{r}_1 - \vec{r}_2$ 。两质点的质心是

$$\vec{R} = \frac{m_1\vec{r}_1 + m_2\vec{r}_2}{m_1 + m_2} \quad (8.1)$$

于是

$$\vec{r}_1 = \vec{R} + \frac{m_2}{m_1 + m_2}\vec{r} = \vec{R} + \vec{r}'_1 \quad (8.2)$$

$$\vec{r}_2 = \vec{R} - \frac{m_1}{m_1 + m_2}\vec{r} = \vec{R} + \vec{r}'_2 \quad (8.3)$$

而

$$\vec{r}'_1 = \frac{m_2}{m_1 + m_2}\vec{r} \quad (8.4)$$

$$\vec{r}'_2 = -\frac{m_1}{m_1 + m_2}\vec{r} \quad (8.5)$$

两个质点的运动方程分别为

$$m_1\ddot{\vec{r}}_1 = f(r)\hat{r} \quad (8.6)$$

$$m_2\ddot{\vec{r}}_2 = -f(r)\hat{r} \quad (8.7)$$

用质心和相对位移写出的方程是

$$M\ddot{\vec{R}} = 0 \quad (8.8)$$

$$\mu\ddot{\vec{r}} = f(r)\hat{r} \quad (8.9)$$

$M = m_1 + m_2$ 为总质量， $\mu = m_1m_2/(m_1 + m_2)$ 称为折合质量。这样两体问题就可以映射到一个假想的单体问题。

单体问题的解是一个椭圆。两个质点的轨迹也都是椭圆，如图所示。两个椭圆以质心为共同的焦点，它们的偏心率都与单体的轨迹一样，但大小都按比例缩小，质量大的比例小。如果一个质量很大，另一个质点的轨迹就会几乎与单体问题的轨迹重合，如太阳与行星。在同一时刻，两个质点与质心共线。

假想的单体系统的角动量是两粒子系统在质心系中的角动量：

$$\begin{aligned} \vec{L}_c &= m_1\vec{r}'_1 \times \vec{v}'_1 + m_2\vec{r}'_2 \times \vec{v}'_2 \\ &= \frac{m_1m_2}{m_1 + m_2}\vec{r} \times (\vec{v}'_1 - \vec{v}'_2) \\ &= \mu\vec{r} \times \vec{v} \end{aligned} \quad (8.10)$$

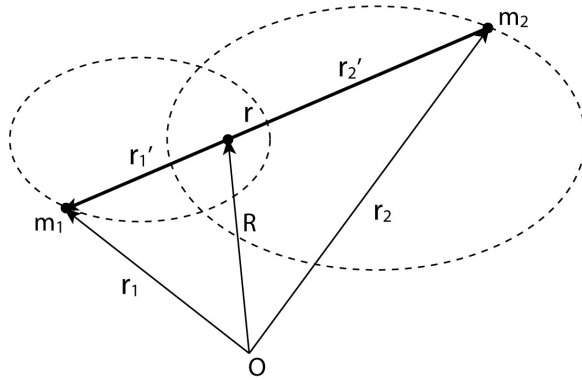


图 8.1: 两体运动约化为单体运动

同样

$$E_c = \frac{1}{2}\mu v^2 + U(r) \quad (8.11)$$

因此在假想的单体系统中求出角动量和能量就可以直接得到真实两体系统在质心系中相应的值。

8.2 有心力运动的一般性质

对于单体问题, 角动量 $\vec{L} = \vec{r} \times \mu \vec{v}$, $\vec{v} \perp \vec{L}$ 。由于 \vec{L} 守恒, 所以运动位于一个平面中。这样就可以使用极坐标来解运动方程

$$\mu(\ddot{r} - r\dot{\theta}^2) = f(r) \quad (8.12)$$

$$\mu(r\ddot{\theta} + 2\dot{r}\dot{\theta}) = 0 \quad (8.13)$$

第二个方程意味着角动量守恒: $dl/dt = 0$, $l = \mu r^2 \dot{\theta}$ 。因此同样时间扫过的面积相同。

8.3 有效势

将系统的能量写为

$$\begin{aligned} E &= \frac{1}{2}\mu v^2 + U(r) \\ &= \frac{1}{2}\mu(\dot{r}^2 + r^2\dot{\theta}^2) + U(r) \\ &= \frac{1}{2}\mu\dot{r}^2 + \frac{1}{2}\frac{l^2}{\mu r^2} + U(r) \end{aligned} \quad (8.14)$$

如果定义有效势

$$U_{\text{eff}} = \frac{1}{2} \frac{l^2}{\mu r^2} + U(r) \quad (8.15)$$

则可以将二维问题约化为一维的问题。这样我们就首先将原来的一个有六个自由度的三维问题化为单体的运动，然后由于角动量守恒化为二维运动，最后利用有效势化为一维问题。由于角动量守恒，角方向的运动可以化为只与 r 有关，成为了有效势的一部分。我们下面判断一下什么样的有心势可以有轨道不过原点的束缚态，即轨道不会到无限远处也不会有在某一时刻穿过原点的状态。这是行星上可以存在生命的必要条件。假设有有心势的一般形式是 $U(r) = \alpha r^n$ 。对应的有效势是

$$U_{\text{eff}} = \alpha r^n + \frac{1}{2} \frac{l^2}{\mu r^2} \quad (8.16)$$

我们可以通过求导数来确定有效势的极值，从而得到有效势的性状。一阶导数

$$U'_{\text{eff}} = n\alpha r^{n-1} - \frac{l^2}{\mu r^3} \quad (8.17)$$

在 r_0 处为0意味着

$$l^2 = n\alpha\mu r_0^{n+2} \quad (8.18)$$

$$U_{\text{eff}}(r_0) = \frac{1}{2}\alpha(n+2)r_0^n \quad (8.19)$$

二阶导数为

$$U''_{\text{eff}} = r^{n-2} \left(n(n-1)\alpha + \frac{3l^2}{\mu r^{n+2}} \right) \quad (8.20)$$

在极值点

$$U''_{\text{eff}} = r_0^{n-2} (n(n-1)\alpha + 3n\alpha) \quad (8.21)$$

如果 r_0 为极小值点，二阶导数需要大于0，即 $n(n-1)\alpha + 3n\alpha > 0$ 进而 $n(n+2)\alpha > 0$ ，否则 r_0 为极大值点。同时还要要求 $r_0 > 0$ ，即 $n\alpha > 0$ ，从而 $n+2 > 0$ 。按照 $\alpha > 0$ 和 $\alpha < 0$ ，得到如图8.2中的前两种情况。

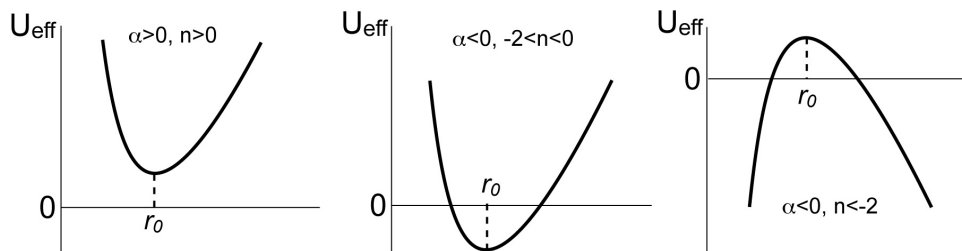


图 8.2: 各种有效势

存在轨道不过原点的束缚态的条件是 $\alpha > 0, n > 0$ 或者 $\alpha < 0, -2 < n < 0$ ，即图中的前两个有效势。对于引力 $\alpha < 0, n = -1$ ，可以有这样的束缚态。当不满足上述有束缚态的条件时也可以存在束缚态。对于图8.2中的第三个有效势，如果 $0 \leq r \leq r_0$ ，则仍然可以有束缚态。只不过这时轨道可能过原点。例如，对于 $U \sim -1/r^2$ 的势能，一种可能的轨道是 $r = ae^{-b\theta}$ ；对于 $U \sim -1/r^3$ 的势能，一种可能的轨道是 $r = a(1 + \cos\theta)$ 。

8.4 闭合轨道

U_{eff} 中的 $\frac{l^2}{2\mu r^2}$ 项实际上是动能的一部分，它代表非径向运动的动能。在当前情况下，我们可以把这部分动能表达为仅是 r 的函数，因而才会将问题转化为一维，定义了 U_{eff} 。 U_{eff} 对于讨论运动的形态有重要意义，下面讨论一般情况下轨道的形态。当 $E < 0$ 时，由有效势曲线可知 r 在 r_1 （最小）和 r_2 （最大）之间振荡，具有周期 T_r 。

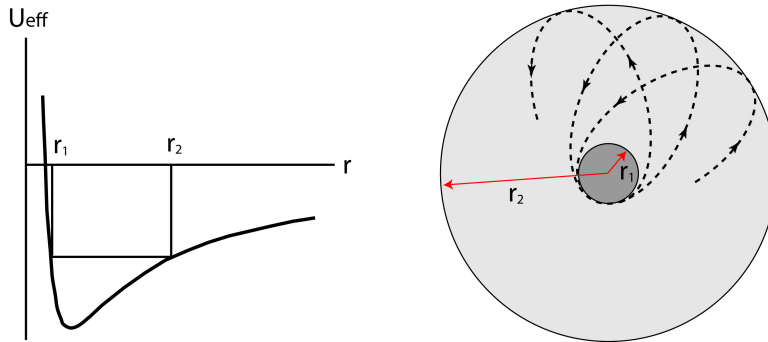


图 8.3: 有心力作用下轨道的一般形式

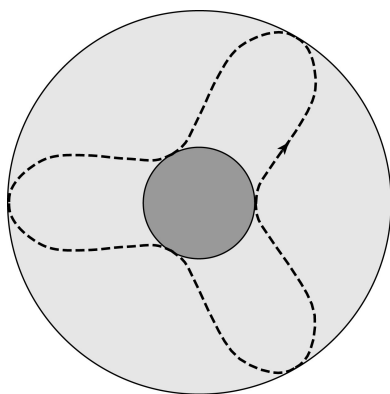
质点总是在如图8.3的两个圆的中间进行运动。 r_1 , r_2 代表径向运动的转折点，在这两点上径向速度为0，但切向速度不为0，因为角动量不为0。径向速度为0的点，其轨道的切线垂直于 \vec{r} ，因为 $\vec{v} \cdot \vec{r} = (r\dot{\theta}\hat{\theta}) \cdot r\hat{r} = 0$ 。而 r_1 和 r_2 两个圆的切线也垂直于 \vec{r} ，所以轨道在 r_1 , r_2 处与两个圆相切。在一个一般的轨道上，质点从与小圆相切到与大圆相切，再与小圆相切， r 完成一个周期运动，周期为 T_r 。同时 θ 也在 $(0, 2\pi)$ 间变化，周期为 T_θ 。如果 $T_r = T_\theta$ ，则每次 r 完成一个周期， θ 也变化了 2π ，轨道闭合。一般地如果 T_r/T_θ 能够表示为两个整数的比值，那么轨道闭合，否则不闭合。图8.4是一个 $T_\theta = 3T_r$ 的例子。可以证明平方反比力导致闭合轨道。事实上有一个定理表明对于 r^n 的力的形式只有 $n = 1$ 和 -2 才能对任意的 E 和 l 都有闭合轨道。在天文观测中发现天体的运动在第一级近似下都有闭合轨道。通过这种观测就可以得知天体间的作用是平方反比力（ $n = 1$ 的力不现实）。轨道的形式确定了力的形式。广义相对论效应可导致非闭合轨道，如水星近日点进动，

8.5 行星运动

由能量守恒和角动量守恒

$$E = \frac{1}{2}\mu\dot{r}^2 + U_{\text{eff}}(r) \quad (8.22)$$

$$l = \mu r^2 \dot{\theta} \quad (8.23)$$

图 8.4: $T_\theta = 3T_r$

可得到

$$\frac{dr}{dt} = \sqrt{\frac{2}{\mu}(E - U_{eff})} \quad (8.24)$$

和

$$\frac{d\theta}{dt} = \frac{l}{\mu r^2} \quad (8.25)$$

两式相除有：

$$\frac{d\theta}{dr} = \frac{l}{\mu r^2} \frac{1}{\sqrt{\frac{2}{\mu}(E - U_{eff})}} \quad (8.26)$$

上式可以给出轨道的方程： $r(\theta)$ 。对于引力， $U(r) = -\frac{k}{r}$ ，所以

$$\begin{aligned} \frac{d\theta}{dr} &= \frac{l}{\mu r^2} \frac{1}{\sqrt{\frac{2}{\mu}(E + \frac{k}{r} - \frac{l^2}{2\mu r^2})}} \\ &= \frac{l}{r\sqrt{2\mu r^2 E + 2\mu k r - l^2}} \end{aligned} \quad (8.27)$$

积分得到

$$\theta = l \int \frac{dr}{r\sqrt{2\mu r^2 E + 2\mu k r - l^2}} + \theta_0 \quad (8.28)$$

对 r 积分后得到

$$\theta - \theta_0 = \arcsin \left(\frac{\mu k r - l^2}{r \sqrt{\mu^2 k^2 + 2\mu E l^2}} \right) \quad (8.29)$$

将 r 表达成 θ 的函数有（按习惯 $\theta_0 = -\frac{\pi}{2}$ ）：

$$r = \frac{r_0}{1 - \epsilon \cos \theta} \quad (8.30)$$

其中 $r_0 = \frac{l^2}{\mu k}$, $\epsilon = \sqrt{1 + \frac{2El^2}{\mu k^2}}$ 。

从 r 和 θ 的关系得出行星的轨道为椭圆形。绕太阳运动的行星的能量位于有效能最低点和0之间： $-\frac{\mu k^2}{2l^2} \leq E < 0$, 因而 $0 \leq \epsilon < 1$ 。当 E 正好等于有效势的最低点时, $\epsilon = 0$, 轨道方程为 $r = r_0$ 为圆轨道, r_0 就是角动量为 l 时通过改变 E 得到的圆轨道的半径。一般地 $0 \leq \epsilon < 1$ 时, 轨道为椭圆, 太阳位于椭圆的一个焦点上, ϵ 为椭圆的偏心率。 l 一定时, 最低的 E 对应的轨道是圆轨道。

远日点在 $\theta = 0$ 处: $r_{\max} = \frac{r_0}{1 - \epsilon}$

近日点在 $\theta = \pi$ 处: $r_{\min} = \frac{r_0}{1 + \epsilon}$

$$\epsilon = \frac{r_{\max} - r_{\min}}{r_{\max} + r_{\min}} \quad (8.31)$$

长轴长度 $A = 2a = r_{\min} + r_{\max} = \frac{2r_0}{1 - \epsilon^2} = \frac{k}{-E}$

短轴长度 $= 2b = \frac{2r_0}{\sqrt{1 - \epsilon^2}} = \frac{2l}{\sqrt{-2\mu E}}$

$b = \sqrt{r_{\max} \cdot r_{\min}}$

长轴长度只能与能量有关, 短轴长度与能量和角动量都有关。

轨道的性质可以通过有效势曲线来理解。

$$U_{eff} = -\frac{k}{r} + \frac{l^2}{2\mu r^2} \quad (8.32)$$

近日点和远日点分别是能量与势能曲线的两个交点, 而 $A = r_{\min} + r_{\max}$ 。 r_{\min} 和 r_{\max} 可以由下式确定

$$E = -\frac{k}{r} + \frac{l^2}{2\mu r^2} \quad (8.33)$$

即

$$2\mu E r^2 + 2\mu k r - l^2 = 0 \quad (8.34)$$

由二次方程两根之和满足的关系知 而 $A = r_{\min} + r_{\max} = -\frac{2\mu k}{2\mu E} = -\frac{k}{E}$ 和 $r_{\min} \cdot r_{\max} = -\frac{l^2}{2\mu E}$ 。

考虑这样一种情况：切向动能一定，改变径向动能，例如沿径向推动卫星。对于一定的 l ，不同的 E 对应于不同的轨道形状，如图 8.5。

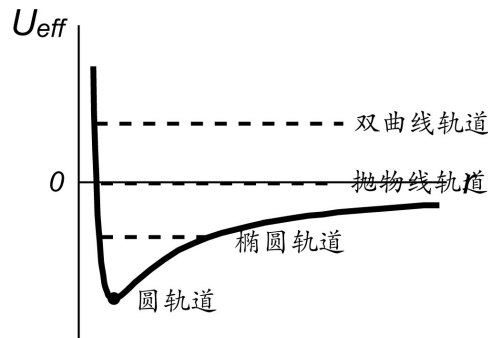


图 8.5: 一定的角动量、不同的能量

对于一定的 E ，可以有不同的 l 。 l 越大，有效势越浅。如图 8.6 中 $l_1 > l_2$ ， l 的最大的可能只是 l_1 ，这时势能曲线与 E 只有一个交点，对应于圆轨道。任何小于 l_1 但大于 0 的 l 与 E 都有两个交点，从曲线上可以得知 l 越小， r_{\min} 和 r_{\max} 的差别越大。 l 越大，分配给径向的动能越少，轨道越圆。

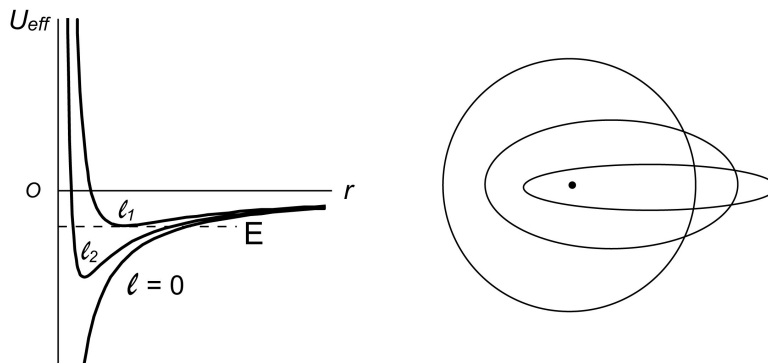


图 8.6: 一定的能量、不同的角动量

如图，能量一定时，角动量越小，椭圆越扁。角动量最大时，对应于圆。但对于所有这些轨道，它们的长轴都相同。

类似从有效势还可以定出 $r_{\max} - r_{\min}$ ，进而定出 $\epsilon = \frac{r_{\max} - r_{\min}}{r_{\max} + r_{\min}}$ 。

$$\begin{aligned} \epsilon &= \frac{r_{\max} - r_{\min}}{r_{\max} + r_{\min}} \\ &= \left(1 - \frac{4r_{\max}r_{\min}}{r_{\max} + r_{\min}} \right)^{\frac{1}{2}} \end{aligned}$$

$$\begin{aligned}
 &= \left(1 + \frac{4l^2 \cdot 2\mu E}{4\mu^2 k^2} \right)^{\frac{1}{2}} \\
 &= \sqrt{1 + \frac{2El^2}{\mu k^2}} \quad (8.35)
 \end{aligned}$$

而 r_0 是 U_{eff} 极小值的位置 $l^2/\mu k$ 。所以如果是椭圆轨道， r_0 和 ϵ 都可由势能曲线定出，进而定出轨道方程。

行星的开普勒定律

一、行星在椭圆轨道上运动，太阳在椭圆焦点上，此定律上面已知证明。

二、太阳到行星的径向矢量在相等时间内扫过相同面积。

面积 $\Delta A = \frac{1}{2}r \cdot r\Delta\theta = \frac{1}{2}r^2\Delta\theta$ 。面积被扫过的速率： $\frac{dA}{dt} = \frac{1}{2}r^2\frac{d\theta}{dt} = \frac{1}{2}\frac{l}{\mu}$ 。所以第二定律是角动量守恒的直接结果

三、行星公转周期 T 与椭圆长轴 A 有以下关系： $T^2 = kA^3$ ， k 对所有行星相同。

$l = \mu r^2 \frac{d\theta}{dt}$ 给出 $\frac{l}{2\mu} dt = \frac{1}{2} r^2 d\theta$ ，右边是 dt 时间内扫过的面积。因而对于一个周期 $\frac{l}{2\mu} T = \pi ab$ ， πab 是椭圆的面积，进而 $T^2 = \frac{4\mu^2}{l^2} \pi^2 a^2 b^2 = \frac{\pi^2 \mu k^2}{-2E^3} = \frac{\pi^2 \mu}{2k} A^3$ 。而 $\mu = \frac{Mm}{M+m}$ ， $k = GMm$ ，所以 $T^2 = \frac{\pi^2}{2(M+m)G} A^3$ 当 $M \gg m$ 时，第三定律准确。第三定律可用于确定行星或太阳的质量。例如对于地球可以用月球的轨道参数来算出质量，而月球的轨道参数可以用三角形法或雷达确定。

$\epsilon \geq 1$ 时，不能形成束缚态，轨道为抛物线（ $\epsilon = 1$ ）或双曲线（ $\epsilon > 1$ ）。例如彗星。考虑 $E > 0$ （ $\epsilon > 1$ ）时的情形，称为散射问题。

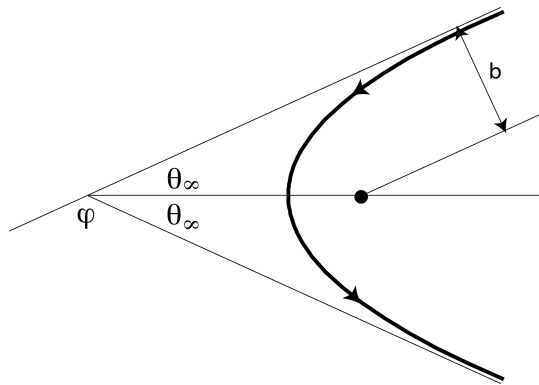


图 8.7: 散射

质量 m （ $\ll M$ ）在 M 作用下轨道偏转 φ 角， b 称为碰撞参数。

$$r = \frac{r_0}{1 - \epsilon \cos \theta} \quad (8.36)$$

我们需要求出 $r \rightarrow \infty$ 时 θ 的大小, 这时

$1 - \epsilon \cos \theta = 0 \Rightarrow \cos \theta_\infty = \frac{1}{\epsilon}$, 对应于入射 ($t = -\infty$) 和出射 ($t = \infty$) 的两个角度, 而 $\varphi = \pi - 2\theta_\infty$ 。

$$\epsilon = \sqrt{1 + \frac{2El^2}{mk^2}} \quad (8.37)$$

其中 $E = \frac{1}{2}mv_\infty^2$, $l = mbv_\infty$, $k = GmM$

则

$$\epsilon = \sqrt{1 + \left(\frac{bv_\infty^2}{GM}\right)^2} \quad (8.38)$$

所以

$$\cos \theta_\infty = \frac{1}{\sqrt{1 + \left(\frac{bv_\infty^2}{GM}\right)^2}} \quad (8.39)$$

可以将 θ_∞ 写为更为简洁的形式:

$$\begin{aligned} \cos^2 \theta_\infty &= \frac{1}{1 + \left(\frac{bv_\infty^2}{GM}\right)^2} \\ \Rightarrow \frac{1}{\cos^2 \theta_\infty} - 1 &= \left(\frac{bv_\infty^2}{GM}\right)^2 \\ \Rightarrow \tan \theta_\infty &= \frac{bv_\infty^2}{GM} \end{aligned} \quad (8.40)$$

$$\varphi = \pi - 2 \arctan \frac{bv_\infty^2}{GM} \quad (8.41)$$

$\arctan x \sim \frac{\pi}{2} - \frac{1}{x}$, 当 $x \rightarrow \infty$ 时。因而当 $v_\infty \rightarrow \infty$ 时, $\varphi \rightarrow 0$, 高速意味着小散射。

8.6 变轨

在地球轨道上的卫星通过自身的动力可以实现变轨。

根据火箭的公式:

$$\Delta \vec{v} = -\vec{u} \ln \left(\frac{m}{m - \Delta m} \right) \quad (8.42)$$

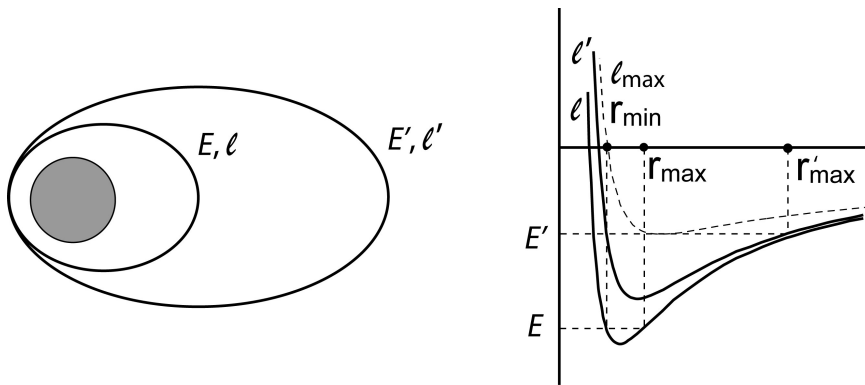


图 8.8: 变轨

Δv 越小, 需要的燃料也越少。一般地

$$\begin{aligned}
 E' - E &= \frac{1}{2}mv'^2 - \frac{1}{2}mv^2 \\
 &= \frac{1}{2}m(\vec{v} + \Delta\vec{v})^2 - \frac{1}{2}mv^2 \\
 &= m\vec{v} \cdot \Delta\vec{v}
 \end{aligned} \tag{8.43}$$

为了用最少的燃料, 得到最大的能量变化, 需要 v 最大, 并且 \vec{v} 和 $\Delta\vec{v}$ 平行。这意味着变轨应发生在近地点 (v 最大), 并且冲量与轨道相切。

利用有效势曲线可以证明, 可以将卫星轨道通过上述办法变到具有任意的远地点的轨道。

初始的轨道有 E 和 l , 近地点 r_{\min} 。利用上述变轨方案, r_{\min} 不变, 新的远地点为 r'_{\max} 。需要证明的是存在一条曲线, 角动量为 l' , 这条曲线与 r_{\min} 和 r'_{\max} 两条垂直线相交在相同的高度上。这一高度即为新的能量。注意这条曲线的 r_{\max} 为 ∞ , 而 l 对应的曲线的 r_{\max} 小于 r'_{\max} , 所以总会有一条曲线 l' 满足要求。 E 和 l 可通过下式确定:

$$r_{\max} + r_{\min} = -\frac{k}{E} \tag{8.44}$$

$$r_{\max} \cdot r_{\min} = -\frac{l^2}{2\mu E} \tag{8.45}$$

9 质点系动力学

9.1 动量

在多质点体系中，如果我们知道所有的外力和质点间的相互作用，就可以详细描述每个质点的运动。但是详细描述往往过于复杂，人们一般采用的办法是分层次解决问题。

首先是最基本的层次。作用在第 i 个质点上的力和该质点的动量满足

$$\vec{f}_i = \frac{d\vec{p}_i}{dt} \quad (9.1)$$

力分为由质点系外的源产生的力和质点系内质点间的力：

$$\vec{f}_i = \vec{f}_i^{\text{ext}} + \vec{f}_i^{\text{int}} \quad (9.2)$$

所以

$$\vec{f}_i^{\text{ext}} + \vec{f}_i^{\text{int}} = \frac{d\vec{p}_i}{dt} \quad (9.3)$$

上式对系统中的所有质点求和：

$$\sum \vec{f}_i^{\text{ext}} + \sum \vec{f}_i^{\text{int}} = \frac{d}{dt} \sum \vec{p}_i \quad (9.4)$$

而 $\sum \vec{f}_i^{\text{int}} = 0$ 。令 $\vec{F}^{\text{ext}} = \sum \vec{f}_i^{\text{ext}}$ ， $\vec{p} = \sum \vec{p}_i$ 为总外力和总动量，则，

$$\frac{d\vec{p}}{dt} = \vec{F}^{\text{ext}} \quad (9.5)$$

以下外力简记为 \vec{F} 。方程只描述了质点系的整体平动。体系总的动量变化率只与外力有关，而与质点间相互作用细节无关。外力的合力为0则总动量守恒。

一般地，总动量是各质点速度的函数，即 $\vec{p} = \vec{p}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_N)$ ，共 $3N$ 个自变量。但是我们发现自变量数目实际上可以减少到3个。定义

$$\vec{R} = \frac{1}{M} \sum m_i \vec{r}_i \quad (9.6)$$

\vec{R} 对应的点叫质心。其中 M 是总质量。容易看到 $\vec{p} = \vec{p}(\vec{V})$ ，是3个自变量的函数，其中 $\vec{V} = d\vec{R}/dt$ 。运动方程可写为

$$\vec{F} = M\ddot{\vec{R}} \quad (9.7)$$

9.2 角动量

用质心描述运动往往很粗糙，如同用平均值来描述一组数，而更多一些的信息可由方差给出。我们需要引入比动量复杂的新的量，叫角动量。角动量很重要，是矢量、守恒量，方便于理解转动。对于一个粒子，角动量的定义是

$$\vec{L} = \vec{r} \times \vec{p} \quad (9.8)$$

角动量依赖参考点的选取，如同方差依赖于参考点的选择一样。需要注意角动量是相对于一个固定在惯性系中的参考点而言，而不是一个轴。

对于一个粒子，

$$\begin{aligned} \frac{d\vec{L}}{dt} &= \frac{d}{dt}(\vec{r} \times \vec{p}) \\ &= \frac{d\vec{r}}{dt} \times \vec{p} + \vec{r} \times \frac{d\vec{p}}{dt} \\ &= \vec{v} \times \vec{p} + \vec{r} \times \vec{F} \\ &= \vec{r} \times \vec{F} \end{aligned} \quad (9.9)$$

$\vec{r} \times \vec{F} = \vec{\tau}$ 叫做力矩。

$$\frac{d\vec{L}}{dt} = \vec{\tau} \quad (9.10)$$

对于一个粒子，动量的方程已经给出了完整的动力学， $d\vec{L}/dt = \vec{\tau}$ 没有给出比 $d\vec{p}/dt = \vec{F}$ 更多的信息。

例

$$\vec{L} = m\vec{r} \times \vec{v} = mlv\hat{k} \quad (9.11)$$

$$\vec{\tau} = \vec{r} \times \vec{f} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ x & -l & 0 \\ -f & 0 & 0 \end{vmatrix} = -lf\hat{k} \quad (9.12)$$

$$\frac{d\vec{L}}{dt} = \vec{\tau} \Rightarrow -lf\hat{k} = ml \frac{dv}{dt} \hat{k} \Rightarrow \frac{dv}{dt} = -f \quad (9.13)$$

与 $d\vec{p}/dt = \vec{F}$ 给出相同的方程。

不同于单个质点，对于质点系，角动量给出了更多的信息。

$$\vec{L} = \sum_i \vec{r}_i \times \vec{p}_i \quad (9.14)$$

$$\frac{d\vec{L}}{dt} = \sum \frac{d\vec{r}_i}{dt} \times \vec{p}_i + \sum \vec{r}_i \times \frac{d\vec{p}_i}{dt} = \sum \vec{r}_i \times \vec{F}_i \quad (9.15)$$

$$\vec{F}_i = \vec{F}_i^{\text{ext}} + \vec{F}_i^{\text{int}} \quad (9.16)$$

所以

$$\begin{aligned} \frac{d\vec{L}}{dt} &= \sum \vec{r}_i \times (\vec{F}_i^{\text{ext}} + \vec{F}_i^{\text{int}}) \\ &= \sum \vec{r}_i \times \vec{F}_i^{\text{ext}} + \sum \vec{r}_i \times \vec{F}_i^{\text{int}} \end{aligned} \quad (9.17)$$

内力总是成对出现。假如*i*和*j*两质点间的力平行于 $\vec{r}_i - \vec{r}_j$ ，则有 $\vec{r}_i \times \vec{F}_{ij} + \vec{r}_j \times \vec{F}_{ji} = (\vec{r}_i - \vec{r}_j) \times \vec{F}_{ij} = 0$ ，因此

$$\frac{d\vec{L}}{dt} = \sum \vec{r}_i \times \vec{F}_i^{\text{ext}} = \vec{\tau} \quad (9.18)$$

如外力矩为0，则角动量守恒。

角动量的变化率只与外力的力矩有关。 \vec{L} 满足一个简单的方程，比 \vec{p} 更进一步给出质点系更多的信息。一般地还可以定义比 p 和 L 更高阶的物理量，但它们一般不再只与外力有关。

有关角动量的讨论可以进一步简化。可以证明如果以质心作为参考基点，角动量和角动量满足的运动的方程可以有简单的形式。

$$\begin{aligned} \vec{L} &= \sum m_i \vec{r}_i \times \dot{\vec{r}}_i \\ &= \sum m_i (\vec{r}'_i + \vec{R}) \times (\dot{\vec{r}}'_i + \dot{\vec{R}}) \quad \vec{R} \text{是质心的位置矢量} \\ &= \sum m_i \vec{r}'_i \times \dot{\vec{r}}'_i + (\sum m_i \vec{r}'_i) \times \dot{\vec{R}} + \vec{R} \times (\sum m_i \dot{\vec{r}}'_i) + \sum m_i \vec{R} \times \dot{\vec{R}} \end{aligned} \quad (9.19)$$

根据质心的定义有 $\sum m_i \vec{r}'_i = 0$ 和 $\sum m_i \dot{\vec{r}}'_i = 0$ ，所以

$$\vec{L} = m \vec{R} \times \dot{\vec{R}} + \sum m_i \vec{r}'_i \times \dot{\vec{r}}'_i = m \vec{R} \times \dot{\vec{R}} + \vec{L}_0 \quad (9.20)$$

\vec{L}_0 是以质心为参考点的角动量，第一项是质心的角动量。再来看力矩

$$\begin{aligned} \vec{\tau} &= \sum \vec{r}_i \times \vec{f}_i \\ &= \sum (\vec{r}'_i + \vec{R}) \times \vec{f}_i \\ &= \sum \vec{r}'_i \times \vec{f}_i + \vec{R} \times \vec{F} \end{aligned} \quad (9.21)$$

一般地

$$\begin{aligned} \frac{d\vec{L}}{dt} &= \frac{d\vec{L}_0}{dt} + m \frac{d}{dt} (\vec{R} \times \dot{\vec{R}}) \\ &= \frac{d\vec{L}_0}{dt} + m (\dot{\vec{R}} \times \dot{\vec{R}}) + m \vec{R} \times \ddot{\vec{R}} \end{aligned}$$

$$\begin{aligned}
 &= \frac{d\vec{L}_0}{dt} + m\vec{R} \times \ddot{\vec{R}} \\
 &= \frac{d\vec{L}_0}{dt} + \vec{R} \times \vec{F}
 \end{aligned} \tag{9.22}$$

而 $d\vec{L}/dt = \vec{\tau}$, 所以

$$\frac{d\vec{L}}{dt} = \frac{d\vec{L}_0}{dt} + \vec{R} \times \vec{F} = \sum \vec{r}'_i \times \vec{f} + \vec{R} \times \vec{F} \tag{9.23}$$

因而

$$\frac{d\vec{L}_0}{dt} = \sum \vec{r}'_i \times \vec{f} = \vec{\tau}_0 \tag{9.24}$$

9.3 多体问题

我们来开始考虑多体问题, 来看一看角动量如何为我们带来多体动力学的更多的信息。我们讨论的多体问题仅限于刚体, 一个物体可能有各种各样的运动, 如弯曲、抖动等等, 但我们要考虑的是一个理想化的物体, 叫刚体。刚体中原子间的相对作用力如此之强, 以致外力不会使物体变形。刚体中任何两点的距离始终保持不变。从经验可以得知, 刚体除了平动, 还可以有一种运动叫转动。平动有三个自由度, 转动也有三个自由度。因为转动总是绕一个轴, 轴的指向用两个量描述, 沿该轴转过的角度用一个量描述, 所以转动一共三个自由度。由 N 个质点组成的刚体, 其质点之间有刚性连接, 这样名义上就会产生 $C_N^2 = N(N-1)/2$ 个约束。这个约束的数目一般地会远远大于系统最多可能的自由度 $3N$ 。问题来源于这些约束并不独立。事实上, 只需要三个不共线的点就可以确定一个刚体。三个点最多可能的自由度有九个, 再减去它们之间的三个约束, 得到六个自由度。因此刚体的自由度由 N 个质点的 $3N$ 个减为 6 个。而动量和角动量的动力学方程一共 6 个, 所以应该正好够描述刚体的一般运动。当然, 一般的多体体系仅使用动量和角动量远远不够。但是刚体, 作为很多体系一个很好的近似, 恰好可以这样做。讨论刚体运动时, 角动量已经够用, 我们无须深入到比角动量更深的层次, 到角动量即可截断。正如如果一群数据间存在某种关联, 则平均值和方差就可以给出充分的描述, 而无须使用高阶距。许多物体可以看作刚体。刚体假设成立的一个基本的条件是相互作用传播的速度可以看作无穷大。下面我们来关注角动量是否是一个描述刚体转动方便的物理量。

9.4 定轴转动

首先考虑二维转动。二维转动意味着刚体中每个点的运动都位于某一平面中, 而所有这些平面互相平行。二维转动问题只有三个自由度: 两个平动自由度, 一个转动自由度。我们将会看到垂直于这些相互平行平面的角动量分量在二维转动中扮演重要角色。

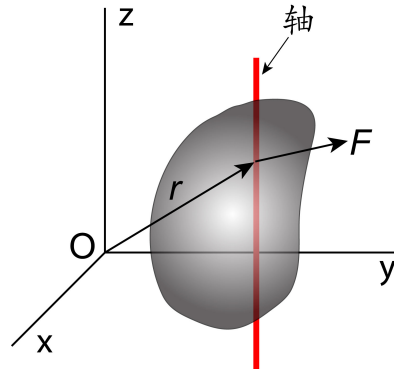


图 9.1: 固定轴转动原点的选取

我们还可以对二维转动继续加约束，即加一根固定的轴。刚体绕这根固定的轴转动。理想化的轴的半径趋于0。将轴固定后，物体运动的自由度降为1。这时原则上即可以用质心的圆周运动也可以用角动量的方程来描述刚体的行为，二者应该给出同样的描述。但是轴给予刚体的力比较复杂，因而质心的运动方程不方便。那么来看角动量的方程。计算角动量和力矩时，需要选择一个点。首先我们任意选择一个点来进行计算，并把坐标原点放在该点上，且让刚体的转动轴平行于 z 轴（图 9.1）。在这个坐标系中，轴上任一点的坐标记为 (x_0, y_0, z_0) 。我们的目的是选择一个合适的坐标原点，使得转轴的力矩不出现。转轴上某一点对刚体的力是外力，为 F ，力矩是：

$$\begin{aligned} & \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ x_0 & y_0 & z_0 \\ F_x & F_y & F_z \end{vmatrix} \\ & = (y_0 F_z - z_0 F_y) \hat{i} + (z_0 F_x - x_0 F_z) \hat{j} + (x_0 F_y - y_0 F_x) \hat{k} \\ & = \dot{L}_x \hat{i} + \dot{L}_y \hat{j} + \dot{L}_z \hat{k} \end{aligned} \quad (9.25)$$

即

$$\dot{L}_x = y_0 F_z - z_0 F_y \quad (9.26)$$

$$\dot{L}_y = z_0 F_x - x_0 F_z \quad (9.27)$$

$$\dot{L}_z = x_0 F_y - y_0 F_x \quad (9.28)$$

但是我们发现只要将原点取在轴上，即 x_0 和 y_0 为0，轴的力矩在 z 方向恒为0。所以对于固定轴的问题，一个方便的做法是将参考点放在轴上，这时轴对于物体的力的力矩沿轴方向为0， L_z 的方程中不出现轴的力。一个 L_z 的方程已足够给出问题的解。在方程 $dL_z/dt = \tau_z$ 中， τ_z 只包含其他力的力矩。由于问题只有一个自由度，因而上式正好可以完全地描述运动。但这并不意味着 $L_x, L_y = 0$ ，只是我们不需要 L_x, L_y 的方程。实际上由于固定轴的存在， L_x, L_y 的方程会较复杂。对于一个自由度的问题，除 $dL_z/dt = \tau_z$ 外的方程不会给出新的物理。还需要强调，角动量是绕一个点而不是绕一个轴定义的。之所以看上去像是对一个轴定义的角动量，是因为 L_z 和 τ_z 与原点的位置无关。

考虑下图中的问题。由图可知 \vec{L} 的大小为 $2ml \cdot l\omega \sin \theta = 2m\omega l^2 \sin \theta$ ，方向如图。写成分量形式：

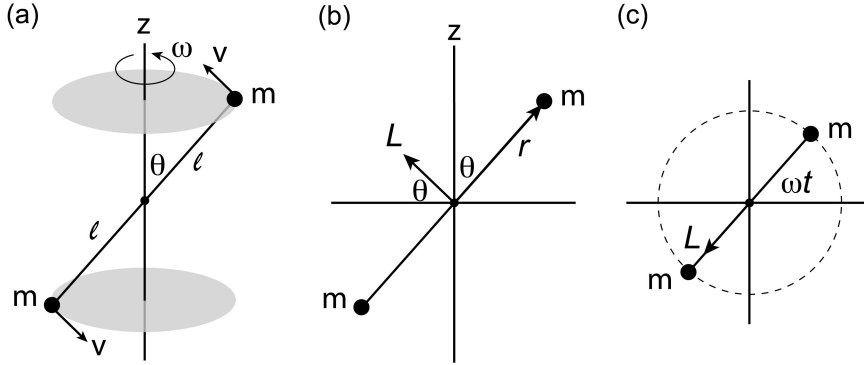


图 9.2: (a) 两质点绕固定轴的转动 (b) 侧视图 (c) 俯视图

$$\begin{aligned}\vec{L} &= 2m\omega l^2 \sin \theta \cos \theta (\cos(\omega t + \pi) \hat{i} \\ &\quad + \sin(\omega t + \pi) \hat{j}) + 2m\omega l^2 \sin^2 \theta \hat{k} \\ &= -2m\omega l^2 \sin \theta \cos \theta (\cos \omega t \hat{i} + \sin \omega t \hat{j}) + 2m\omega l^2 \sin^2 \theta \hat{k}\end{aligned}\quad (9.29)$$

\vec{L} 的 z 分量是常值，不随时间变化，因为力矩为0。但 L_x , L_y 对时间的导数并不为0，轴会施加力矩在系统上，力矩是：

$$\vec{\tau} = 2m\omega^2 l^2 \sin \theta \cos \theta (\sin \omega t \hat{i} - \cos \omega t \hat{j})\quad (9.30)$$

总之，对于固定转轴的问题，只须考虑 $dL_z/dt = \tau_z$ 即可。

下面我们要对一般情况具体写出 L_z 的形式。转动自由度可以用转角 ϕ 来表示，以体现一个自由度。 $d\phi/dt = \omega$ 是角速度。为了得到 ϕ 满足的微分方程，我们希望用 ω 来表示 L_z 。考虑刚体中的第 i 个质点。

$$L_z(i) = (\vec{r}_i \times m_i \vec{v}_i)_z\quad (9.31)$$

$$\vec{r}_i = \vec{\rho}_i + \vec{z}_i\quad (9.32)$$

$$\begin{aligned}L_z(i) &= m_i (\vec{\rho}_i \times \vec{v}_i)_z + m_i (\vec{z}_i \times \vec{v}_i)_z \\ &= m_i v_i \rho_i \\ &= m_i \rho_i^2 \omega\end{aligned}\quad (9.33)$$

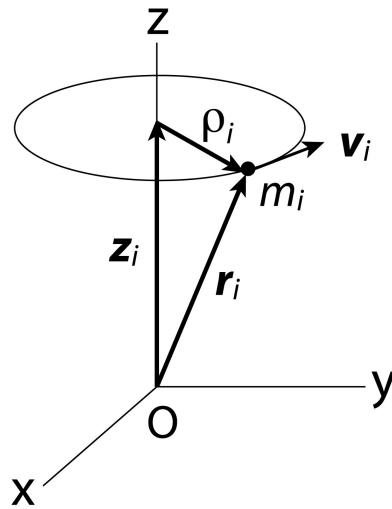


图 9.3: 转动惯量

这里用到了

$$v_i = \omega \rho_i \quad (9.34)$$

因而

$$L_z = \sum_i m_i \rho_i^2 \omega = \left(\sum_i m_i \rho_i^2 \right) \omega = I \omega \quad (9.35)$$

I 称为该刚体绕 z 轴的转动惯量。这样角动量就与描述刚体转动的运动学量联系在一起了。 L_z 与 O 点在轴上的位置无关，因为 $\hat{z} \times v$ 在 z 上无分量。因此看上去 L_z 像是对一个轴定义的。 $dL_z/dt = \tau_z$ 可以写为

$$I \alpha = \tau_z \quad (9.36)$$

α 是角加速度。力矩也可写成关于转轴的形式

$$\tau_z = (\vec{r} \times \vec{F})_z = ((\vec{\rho} + \vec{z}) \times \vec{F})_z = (\vec{\rho} \times \vec{F})_z \quad (9.37)$$

转动行为不仅仅依赖于质量，还依赖质量的分布。质量不变，但 I 依赖于转轴位置，因而可变。下面我们来看怎样计算转动惯量。一般地，

$$\begin{aligned} I &= \int \rho^2 dm \\ &= \int (x^2 + y^2) dm \\ &= \int (x^2 + y^2) w dV \end{aligned} \quad (9.38)$$

w 是密度。

刚体的动能也可以用 I 和 ω 表示：

$$K = \sum \frac{1}{2} m_i v_i^2 = \sum \frac{1}{2} m_i \rho_i^2 \omega^2 = \frac{1}{2} I \omega^2 \quad (9.39)$$

上面考虑的是二维转动中的定轴转动。下面我们放松约束，来研究轴不固定的情形。这时有三个自由度，需要质心平动的方程和一个有关角动量的方程。由于没有了固定的轴，所以原则上取任一点作为参考点都是方便的。

9.5 一般的刚体运动

以前面的几节为基础，我们可以讨论对一般的刚体运动的描述。讨论一个粒子系统的运动时，普遍地讲我们需要解决两个问题。第一、用什么几何参量来方便地描述运动，即运动学。比如，对于单粒子的运动需要用位移随时间的变化来描述。第二、运动学量满足怎样的动力学方程。如单粒子的牛顿方程。这里的运动学量不一定是前面提到的描述运动的参量本身，但可以从这些参量导出。描述刚体运动的两个普遍规律是：系统总动量的变化率等于外力的合力；系统总角动量的变化率等于外力的合力矩。但是具体用哪些参量来写下动量、角动量、力矩等力学量则有很大的任意性。 $d\vec{p}/dt = \vec{F}$ 和 $d\vec{L}/dt = \vec{\tau}$ 都来源于牛顿方程，是对牛顿方程的约化，但包含不同层次的信息。前者类似于平均值，而后这类似于方差。对于刚体这两个方程恰好给出足够的信息。一般的三维刚体运动有六个自由度，通过引入欧拉角可以证明以下定理，

Chasles定理

刚体最一般的位移可以分解为随刚体上任选基点的平动位移和绕过基点的某个轴的转动。选择刚体上不同的点为基点时，平动位移的大小和方向将改变，而转动轴的方向和转角不依赖于基点的选择。

这个定理只是一个几何学的结果，不涉及到任何动力学。至此我们有三点：惯性系中坐标架的原点，是计算角动量等时的参考点；质心；刚体中任意一个代表点，称为基点。这里的基点并不是计算角动量时用的参考点。由该定理可知，六个自由度可用基点的三个坐标、轴指向的两个角度和转角共六个参量描述。这时相应的六个动力学方程（三个动量、三个角动量的方程）中的每一个一般地都与全部六个参量有关。找到一个合适的基点可使得六个参量在动力学中解除耦合，即三个动量的方程只与基点的三个坐标有关，而三个角动量的方程只与三个角度有关。这样的基点是存在的，就是质心。

描述刚体需要6个运动学量 $(X, Y, Z, \alpha, \beta, \phi)$ 。 (X, Y, Z) 是任意的一点。 α, β 给出通过该点的轴的指向，轴随时间变化，因此这两个角度一般是时间的函数。 ϕ 是绕轴的转角。一般地， $\vec{P} = \vec{P}(X, Y, Z, \alpha, \beta, \phi)$ ， $\vec{L} = \vec{L}(X, Y, Z, \alpha, \beta, \phi)$ 。所以关于动量和角动量的两个方程一般都与六个变量有关。原则上，选择任一点作为基点都可以。但如果将运动学量的基点取为质心，则动量的方程只与该基点的坐标有关，而与角度无关。同样可以在角动量的方程中去掉与基点坐标有关的项，从而角动量方程只与绕基点的转动，即三个角度有关。这

样就大大地简化了问题的分析。即力学量简化为

$$\vec{P} = \vec{P}(X, Y, Z) \quad (9.40)$$

$$\vec{L} = \vec{L}_0(\alpha, \beta, \phi) + \vec{R} \times \vec{P} \quad (9.41)$$

我们还曾证明

$$\frac{d\vec{L}_0}{dt} = \vec{\tau}_0 \quad (9.42)$$

一个重要的例外是如果刚体至少有一个点固定，则选择一个固定的点作为基点是最方便的做法。这时虽然动量方程一般会包含角度项，但角动量方程已经足以描述刚体，因而不再需要动量的方程。同时轴或定点对刚体的力所产生的力矩为零，简化了角动量的方程。

这就是我们需要的结果。这是个非常普遍的结果，不仅限于刚体，更不仅限于二维转动。由于该定理的存在，在讨论转动问题时可以将运动分为质心的平动加上绕质心的转动。 $d\vec{L}/dt = \vec{\tau}$ 对于惯性系中的一个固定点成立，同时也对质心成立。至于选择哪个点作为参考点视问题方便而定。如果刚体有一个固定的转轴或有一个固定的点，则通常将参考点选在该轴或点上比较方便，因为轴或点产生的有关的力矩为零。

类似地，对于动能：

$$\begin{aligned} T &= \frac{1}{2} \sum m_i \dot{r}_i^2 \\ &= \frac{1}{2} \sum m_i (\dot{r}'_i + \dot{R})^2 \\ &= \frac{1}{2} \sum m_i \dot{r}'_i{}^2 + \frac{1}{2} \sum m_i \dot{R}^2 + (\sum m_i \dot{r}'_i) \cdot \dot{R} \\ &= T_0 + \frac{1}{2} m \dot{R}^2 \end{aligned} \quad (9.43)$$

T_0 是相对于质心的动能。

再回到前面的例子。如果以质心为基点，则 $bf = I_0\alpha$ 。平动的方程是 $mg \sin \theta - f = ma$ ，同时 $a = \alpha b$ 。从而得到 $mg \sin \theta - \frac{1}{2}ma = ma$ ，即 $a = \frac{2}{3}g \sin \theta$ 。计算得到简化。

Kleppner的6.18是一个很好的利用能量来分析问题的例子，希望大家都能看一下。

例

圆柱与板无滑动地相对运动。圆柱质心的运动方程是

$$m \frac{dv_c}{dt} = f \quad (9.44)$$

圆柱绕质心的转动方程是

$$fR = I \frac{d\omega}{dt} = \frac{1}{2} m R^2 \frac{d\omega}{dt} \quad (9.45)$$

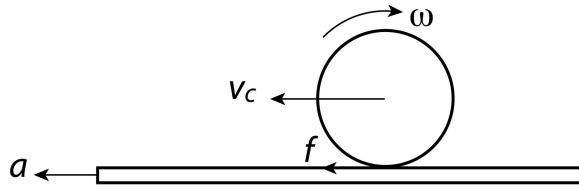


图 9.4: 参考点的选择

几何约束

$$\frac{dv_c}{dt} + R \frac{d\omega}{dt} = a \quad (9.46)$$

于是

$$m \frac{dv_c}{dt} = \frac{1}{2} m R \frac{d\omega}{dt} = \frac{1}{2} m \left(a - \frac{dv_c}{dt} \right) \quad (9.47)$$

从而

$$\frac{dv_c}{dt} = \frac{a}{3} \quad (9.48)$$

但如果取接触点为轴就会得到 L 不变的错误结果。

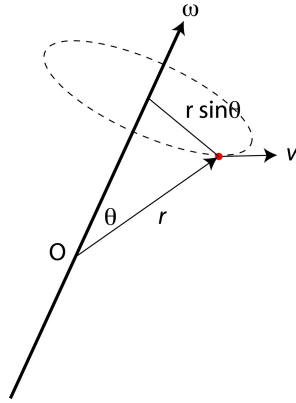
9.6 作为矢量的角速度

怎样用几何参数写出角动量？通过上一节可知刚体的运动总是可以理解为其中一点的平动加上绕这一点的转动。转动必然有一个转轴的存在，尽管转轴一般会随时间而改变方向。角动量可以写为与转动有关的三个角度的函数。在转角无限小时，这三个角度事实上形成了一个矢量，叫做角速度，记为 $\vec{\omega}$ 。其方向是转轴的方向，指向通过转动方向由右手定则确定。大小是转角对时间的导数 $d\phi/dt$ 。前面讨论的转动的转轴总是沿一个方向，因此不需要讨论角速度的方向。然而一般地，角速度是一个矢量。刚体中每点的速度可以方便地用角速度来表达，因而也可以将角动量用角速度这样一个描述转动的量来表达。角速度是描述转动的一个方便的物理量。

相对于基点，刚体中任一点的速度为 $\vec{v} = \vec{\omega} \times \vec{r}$ ，这个结果可以图9.5得到验证，注意 \vec{v} 的大小是 $\omega r \sin \theta$ ，而方向总是垂直于 $\vec{\omega}$ 和 \vec{r} 张成的平面。

前面的一个例子9.2曾经计算过两个粒子体系的角动量。现在我们再利用 $\vec{\omega}$ 的矢量性重新计算 \vec{L} ，需要注意 \vec{L} 和 $\vec{\omega}$ 一般不平行。将 $\vec{\omega}$ 分解为 $\vec{\omega} = \vec{\omega}_{\parallel} + \vec{\omega}_{\perp}$ ，两项分别平行于和垂直于杆，由于质点没有大小，因此 $\vec{\omega}_{\parallel}$ 不会产生角动量，而 $\vec{\omega}_{\perp}$ 导致角动量为 $2ml^2\omega_{\perp} = 2ml^2\omega \sin \theta$ ，与前面的结果一致。

上节的定理中的基点实际上时可以任意选取，对于任何基点都能够把运动分解为该点的平动加上绕该点的转动。一个有趣的事实是不管选取那个点 $\vec{\omega}$ 都不变。图9.6中： $\vec{R} = \overline{OP}$ ， $\vec{R}_1 = \overline{OO_1}$ ， $\vec{R}_2 = \overline{OO_2}$ ， $\vec{a} = \overline{O_1O_2}$ ，

图 9.5: $\vec{v} = \vec{\omega} \times \vec{r}$

$\vec{r}_1 = \overline{O_1P}$, $\vec{r}_2 = \overline{O_2P}$ 。取 O_1 作为基点时, P 的速度是

$$\vec{v} = \frac{d\vec{R}_1}{dt} + \frac{d\vec{r}_1}{dt} = \vec{v}_1 + \vec{\omega}_1 \times \vec{r}_1 \quad (9.49)$$

取 O_2 作为基点时, P 的速度又可以写为

$$\vec{v} = \frac{d\vec{R}_2}{dt} + \frac{d\vec{r}_2}{dt} = \vec{v}_2 + \vec{\omega}_2 \times \vec{r}_2 \quad (9.50)$$

把 \vec{v}_2 用 \vec{v}_1 表示出来, 有:

$$\vec{v}_2 = \frac{d\vec{R}_1}{dt} + \frac{d\vec{a}}{dt} = \vec{v}_1 + \vec{\omega}_1 \times \vec{a} \quad (9.51)$$

从以上两式可得

$$\vec{v} = \vec{v}_1 + \vec{\omega}_1 \times \vec{a} + \vec{\omega}_2 \times \vec{r}_2 \quad (9.52)$$

同时又有 $\vec{r}_1 = \vec{r}_2 + \vec{a}$, 于是从 9.49 式可得

$$\vec{v} = \vec{v}_1 + \vec{\omega}_1 \times \vec{a} + \vec{\omega}_1 \times \vec{r}_2 \quad (9.53)$$

即

$$(\vec{\omega}_1 - \vec{\omega}_2) \times \vec{r}_2 = 0 \quad (9.54)$$

由于 P 点的任意性, 故有

$$\vec{\omega}_1 = \vec{\omega}_2 \quad (9.55)$$

我们可以用角速度来表示角动量。

$$\begin{aligned} \vec{L} &= \sum m_i \vec{r}_i \times \vec{v}_i \\ &= \sum m_i \vec{r}_i \times (\vec{\omega} \times \vec{r}_i) \end{aligned}$$

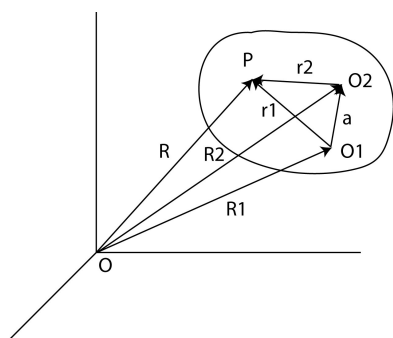


图 9.6: 角速度的唯一性

$$\begin{aligned}
 &= \sum m_i ((\vec{r}_i \cdot \vec{r}_i) \vec{\omega} - (\vec{r}_i \cdot \vec{\omega}) \vec{r}_i) \\
 &= \sum m_i (r_i^2 \vec{\omega} - (\vec{r}_i \cdot \vec{\omega}) \vec{r}_i) \\
 &= (\sum m_i r_i^2) \vec{\omega} - \sum m_i \vec{r}_i (\vec{r}_i \cdot \vec{\omega})
 \end{aligned} \tag{9.56}$$

这里用到了 $\vec{A} \times (\vec{B} \times \vec{C}) = (\vec{A} \cdot \vec{C})\vec{B} - (\vec{A} \cdot \vec{B})\vec{C}$ 。写成分量形式

$$L_x = \sum m_i (r_i^2 - x_i^2) \omega_x - (\sum m_i x_i y_i) \omega_y - (\sum m_i x_i z_i) \omega_z \tag{9.57}$$

L_y 和 L_z 类似。一般地

$$L_x = I_{xx} \omega_x + I_{xy} \omega_y + I_{xz} \omega_z \tag{9.58}$$

$$L_y = I_{yx} \omega_x + I_{yy} \omega_y + I_{yz} \omega_z \tag{9.59}$$

$$L_z = I_{zx} \omega_x + I_{zy} \omega_y + I_{zz} \omega_z \tag{9.60}$$

角速度和角动量是线性关系，但角动量在一个方向上的分量不仅只与该方向上的角速度有关。九个 I 组成了转动惯量张量。

当外力矩为零时角动量守恒。

9.7 回转仪 (Gyroscope)

从演示中可以看到一个转动的回转仪能够绕竖直轴转动，这种运动叫进动 (precession)。

首先一个问题是，放开回转仪后，为什么它不会在重力作用下倒下，而是进动。一个向下的重力为什么可以使回转仪向侧向运动。事实上，在手松开的时候，回转仪确实向下落。但是当回转仪下落时， \vec{L}_0 就会发生变化， $\Delta \vec{L}$ 向下。但是重力的力矩是侧向的，因而不能产生一个向下的角动量变化，为了使角动量守恒，需要回转仪有一个向上的角动量来抵消 ΔL ，这就造成了回转仪必需产生角速度为 Ω 的侧向运动。回转仪在重力作用下的下落不会持续下去，因为 θ 越大， Ω 就需要越大，于是终于会有一个角度使得下落获得的

重力势能完全转化成为侧向转动的动能，下落的速度降为0。然后， θ 会减小， Ω 也会降低，动能再转化为势能，如此周期进行下去，如图9.7所示。绕竖直轴的转动称为进动，在垂直面内的摆动称为章动。让我们来分析一下这种运动。

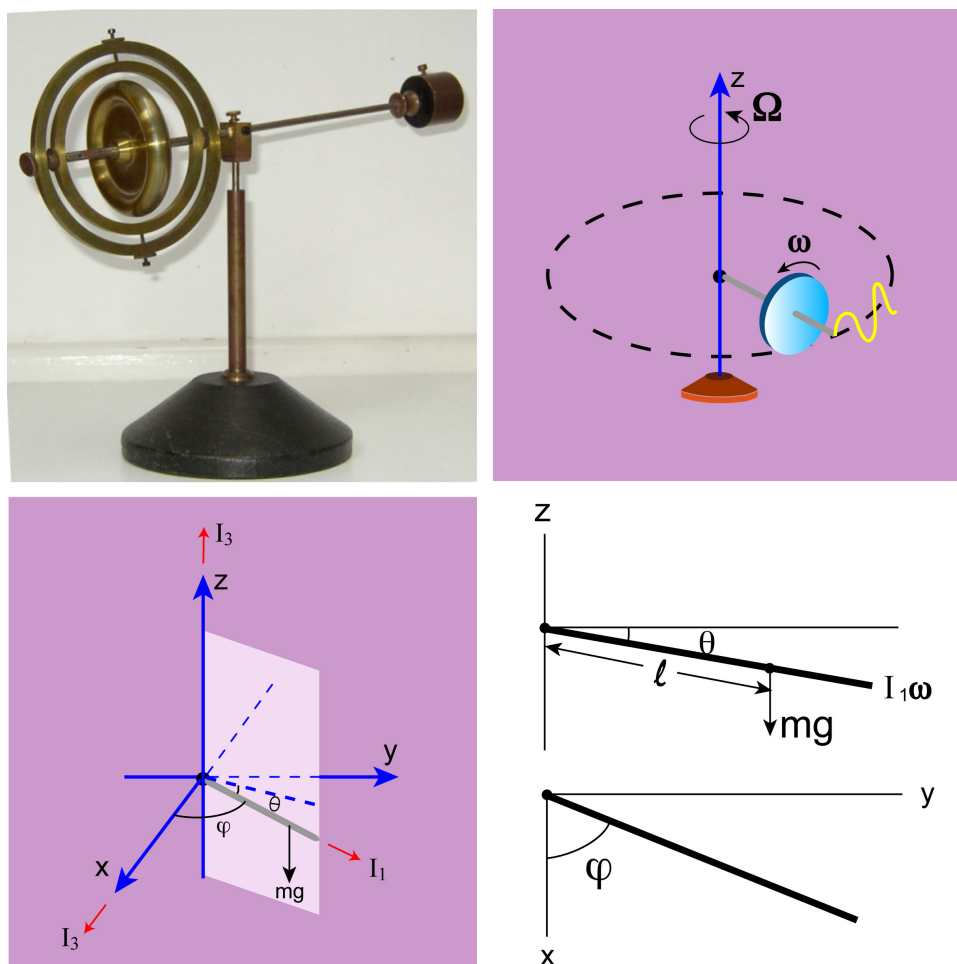


图 9.7: 回转仪

为方便起见，假设手松开时，回转仪水平。在重力作用下，回转仪下落角度为 θ ，以此造成的 ΔL 需要 Ω 来补偿。假设 θ 比较小。回转仪沿自转轴的转动惯量为 I_1 ，沿竖直轴和水平轴的转动惯量均为 I_3 （即与进动和章动有关的转动惯量都为 I_3 ）。于是， $L_0 = I_1\omega$ 。 \vec{L}_0 的变化为 $I_1\omega\theta$ ，而侧向运动的角动量为 $I_3\Omega$ 。由竖直方向角动量守恒得知， $I_3\Omega = I_1\omega\theta$ 。所以

$$\Omega = \frac{d\varphi}{dt} = \frac{I_1\omega}{I_3}\theta = \alpha\theta \quad (9.61)$$

其中 $\alpha = I_1\omega/I_3$ 。下面还需要考虑能量守恒。动能由两部分组成：进动动能（ $I_3\Omega^2/2$ ）和章动动能（ $I_3\dot{\theta}^2/2$ ）。转换为动能的重力势能为 $mgl\theta$ ，与 θ 成正比。进动的动能与 θ^2 成正比。如图 9.8所示，势能与进动动能之差为章动动能。两条线必然会在 $\theta \neq 0$ 的位置有一个交点。在交点处，势能完全转化为进动动能，章动动能为0。这时，章动开始折返。

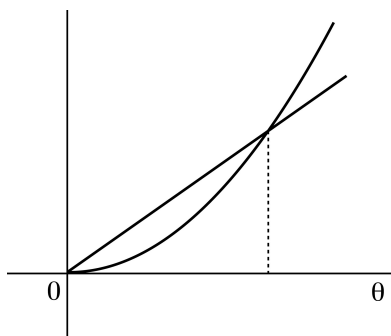


图 9.8: 进动中的能量转换

下面我们仔细研究势能与动能之间的转换。由于能量守恒，

$$\frac{1}{2}I_3\dot{\theta}^2 + \frac{1}{2}I_3\Omega^2 = mgl\theta \quad (9.62)$$

将式9.61代入得到

$$\frac{1}{2}I_3\dot{\theta}^2 + \frac{1}{2}\frac{I_1^2\omega^2}{I_3}\theta^2 - mgl\theta = 0 \quad (9.63)$$

上式左边后两项构成了章动的有效势，而总能量为0（上式右边）。式9.63可写为

$$\frac{1}{2}I_3\dot{\theta}^2 + \frac{1}{2}\frac{I_1^2\omega^2}{I_3}(\theta - \theta_0)^2 = \frac{1}{2}\frac{I_1^2\omega^2}{I_3}\theta_0 \quad (9.64)$$

其中 $\theta_0 = mglI_3/I_1^2\omega^2$ 。有效势为抛物线，见图9.9。于是 θ 以 θ_0 为平衡位置作

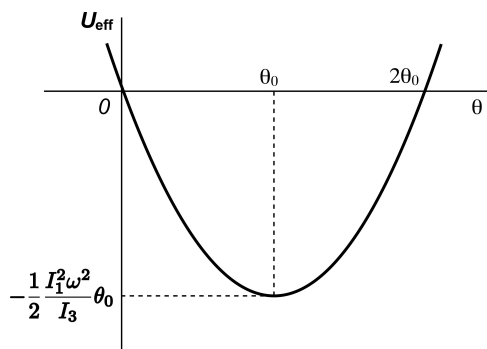


图 9.9: 章动的有效势

简谐振动，振幅为 θ_0 。考虑到初始条件 $t = 0$ 时， $\theta = 0$ ，得到

$$\theta = \theta_0(1 - \cos \alpha t) \quad (9.65)$$

θ 在0和 $2\theta_0$ 之间振荡，频率为 α 。这样还可以得到 Ω 和 φ ，

$$\Omega = \alpha\theta_0(1 - \cos \alpha t) \quad (9.66)$$

$$\varphi = \theta_0(\alpha t - \sin \alpha t) \quad (9.67)$$

φ 有一个线性项, 说明回转仪绕竖直轴持续运动。叠加在这个匀速转动之上的是一个以 θ_0 为振幅的震荡, 几个典型的运动轨迹如图9.10。该轨迹与一个滚动的轮子上的一点相同:

$$y = R - R \cos \omega t \rightarrow \theta \tag{9.68}$$

$$x = \omega R t - R \sin \omega t \rightarrow \varphi \tag{9.69}$$

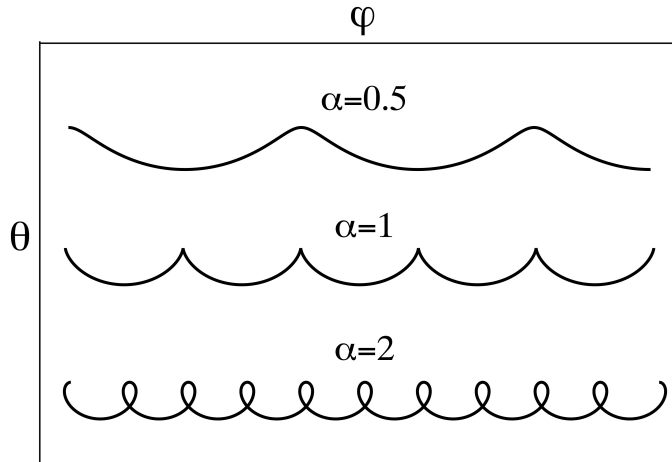


图 9.10: 进动和章动

一般来说 ω 很大, 而章动的频率 $\alpha \sim \omega$, 所以章动很快。由于轴上摩擦力的缘故, 章动很快就会停下来, 而只剩下侧向运动的进动, 其角速度是 $\alpha\theta_0 = \frac{mgl}{I_1\omega} = \Omega$ 。

当然如果我们知道最后达到的稳定态是进动, 那么 Ω 还可以更简单的求出, 具体见Kleppner 7.3节。回转仪的应用见Kleppner 7.4节。

9.8 地球的进动

何时引力的力矩可以把力看作作用在质心上? 如果任何情况下都可以, 则相对于质心引力的力矩总是零。实际上一般只有两种情况可以这样做, 一种是均匀的引力场, 一种是点源中的球对称刚体。

只考虑太阳造成的进动。将太阳的看作点源, 由于地球不是严格的球对称, 因此相对地球质心的力矩不为0。在此力矩作用下, 地球会进动。将问题简化为如图的模型。两个质量为 μ 的质点组成一个刚体。 $\mu = \alpha m$, m 为地球质量, α 为地球偏离球形的那部分质量的一半与地球质量的比值。地球半径为 a , 地球与太阳的距离为 R , $a \ll R$ 。 $\theta = 23.5^\circ$ 。在质量为 M 的太阳的引力场中受到的力矩是

$$\tau = \frac{G\mu Ma}{R^2 + a^2 - 2Ra \cos \theta} \sin(\theta + a \sin \theta / R)$$

$$\begin{aligned}
 & -\frac{G\mu Ma}{R^2 + a^2 + 2Ra \cos \theta} \sin(\theta - a \sin \theta/R) \\
 = & \frac{3G\mu Ma^2}{R^3} \sin \theta \cos \theta
 \end{aligned} \tag{9.70}$$

地球的公转周期为 T ，则 $GM/R^3 = 4\pi^2/T^2$ 。于是

$$\tau = \frac{12\pi^2 \alpha m a^2}{T^2} \sin \theta \cos \theta \tag{9.71}$$

$\Omega = 2\pi/\Gamma$ 为地球进动角速度，则

$$L\Omega \sin \theta = \frac{12\pi^2 \alpha m a^2}{T^2} \sin \theta \cos \theta \tag{9.72}$$

$L = \frac{2}{5} m a^2 \frac{2\pi}{\lambda}$ ， λ 为地球自转周期。于是

$$\Gamma = \frac{T}{\lambda} \frac{2}{15\alpha \cos \theta} T = \frac{53}{\alpha} T \tag{9.73}$$

地球两轴半径相差 Δa ， $\Delta a/a \sim 300$ 。因此 $\alpha = \frac{1}{2} \frac{\Delta V}{V} = \frac{1}{2} \frac{\Delta a}{a} = \frac{1}{600}$ 。于是 $T \sim 30000$ 年。

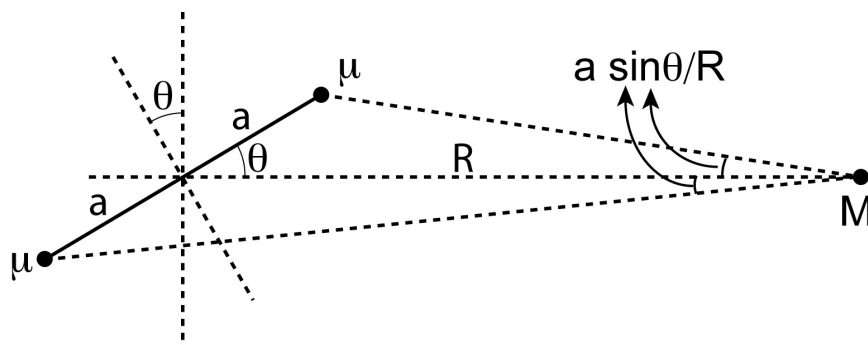


图 9.11: 地球的进动

10 非惯性系

为了对惯性系有更进一步的理解，我们来看一下非惯性系。参考系的基本结构是一个刚体。参考系中有观察者和测量工具，测量工具用来观测物理过程。本质上讲测量工具也是作为参考系基本结构的刚体的一部分。只有刚性连接的一系列质点才能作为观测的参考基准，并记录测量结果。而坐标系只是为了描述方便而为时空点引入的一套参数。坐标系的选择具有随意性，完全是为了讨论问题的方便。比如可以选择与参考系固定在一起的直角坐标系，也可以选择与质点一起运动的极坐标系。物理规律都是相对于参考系而言，用不同的坐标系，可以写出不同的具体形式。牛顿力学的原理只在惯性参考系中成立。但是对于一些问题引入非惯性系会简化计算。运用非惯性系时需要引入虚拟力。

一个刚体的一般运动是平动加转动。在一个惯性参考系中看另一个参考系，作为一个刚体其可能运动也是平动加转动。一般地，非惯性参考系是两种运动的组合。下面我们只考虑非惯性参考系的两种最简单的情形，匀加速平动和定轴匀速转动。

10.1 匀加速平动非惯性系

α 为一惯性参考系， β 为一相对 α 作匀加速平动的非惯性参考系，加速度为 \vec{A} 。对于空间某一点，在 α 和 β 中分别用 \vec{r}_α 和 \vec{r}_β 描写， $\vec{r}_\beta = \vec{r}_\alpha - \vec{R}$ ，进而 $\vec{a}_\beta = \vec{a}_\alpha - \vec{A}$ 。 $\vec{F} = m\vec{a}_\alpha \Rightarrow \vec{F} = m\vec{a}_\beta + m\vec{A}$ 或 $\vec{F} - m\vec{A} = m\vec{a}_\beta$ ，将 $-m\vec{A}$ 定义为虚拟力 $\vec{F}_{fict} = -m\vec{A}$ ，则有 $\vec{F} + \vec{F}_{fict} = m\vec{a}_\beta$ 。匀加速参考系中的虚拟力与重力类似，作用在多粒子系统的虚拟力可认为是作用在质心上。

重新考虑前面的一个例题（图9.4）。

在非惯性系中 $f - F_{fict} = Ma'$ 。其中 $F_{fict} = MA$ 。而 $Rf = -I_0\alpha'$ ，其中 $I_0 = \frac{1}{2}MR^2$ ，以及 $\alpha'R = a'$ 。于是 $a' = -\frac{2}{3}A$ ，即 $a = A + a' = \frac{1}{3}A$ 。

10.2 定轴匀速转动非惯性系

考虑矢量 \vec{B} 在惯性和非惯性参考系中的变化。为此需要仔细地考虑测量的过程。对一个矢量的测量事实上就是用作为刚体的参考系记录下每一时刻的矢量。任何测量只能参照自己的参考系。在非惯性系中的该记录随时间而变化（在惯性系中观察）： $\vec{B} \rightarrow \vec{B}'$ 。因而在非惯性系中测量时， $\Delta\vec{B}' = \vec{B}'(t + \Delta t) - \vec{B}'(t)$ 。而

$$\vec{B}'(t) = \vec{B}(t) + \vec{\Omega} \times \vec{B}\Delta t \quad (10.1)$$

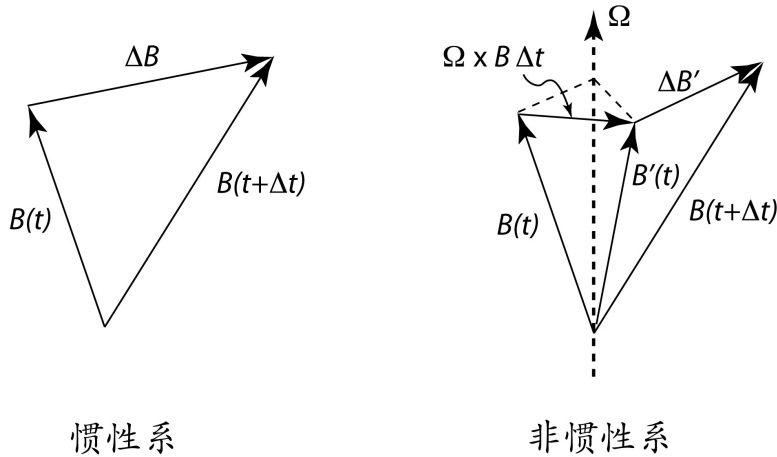


图 10.1: 定轴匀速转动非惯性系中的矢量

所以

$$\frac{\Delta \vec{B}'}{\Delta t} = \frac{\Delta \vec{B}}{\Delta t} - \vec{\Omega} \times \vec{B} \quad (10.2)$$

或者

$$\left(\frac{d\vec{B}}{dt} \right)_{\text{in}} = \left(\frac{d\vec{B}}{dt} \right)_{\text{rot}} + \vec{\Omega} \times \vec{B} \quad (10.3)$$

矢量在任一瞬时，在两个参考系中都一样，但时间变化率不同。这样我们可以求出非惯性系中的速度和加速度：

$$\vec{v}_{\text{in}} = \vec{v}_{\text{rot}} + \vec{\Omega} \times \vec{r} \quad (10.4)$$

以及

$$\begin{aligned} \vec{a}_{\text{in}} &= \left(\frac{d\vec{v}_{\text{in}}}{dt} \right)_{\text{rot}} + \vec{\Omega} \times \vec{v}_{\text{in}} \\ &= \left[\frac{d}{dt} (\vec{v}_{\text{rot}} + \vec{\Omega} \times \vec{r}) \right]_{\text{rot}} + \vec{\Omega} \times (\vec{v}_{\text{rot}} + \vec{\Omega} \times \vec{r}) \\ &= \left(\frac{d}{dt} \vec{v}_{\text{rot}} \right)_{\text{rot}} + \vec{\Omega} \times \left(\frac{d\vec{r}}{dt} \right)_{\text{rot}} + \vec{\Omega} \times \vec{v}_{\text{rot}} + \vec{\Omega} \times (\vec{\Omega} \times \vec{r}) \\ &= \vec{a}_{\text{rot}} + 2\vec{\Omega} \times \vec{v}_{\text{rot}} + \vec{\Omega} \times (\vec{\Omega} \times \vec{r}) \end{aligned} \quad (10.5)$$

由于

$$\vec{F} = m\vec{a}_{\text{in}} = m(\vec{a}_{\text{rot}} + 2\vec{\Omega} \times \vec{v}_{\text{rot}} + \vec{\Omega} \times (\vec{\Omega} \times \vec{r})) \quad (10.6)$$

于是

$$m\vec{a}_{\text{rot}} = \vec{F} - 2m\vec{\Omega} \times \vec{v}_{\text{rot}} - m\vec{\Omega} \times (\vec{\Omega} \times \vec{r}) \quad (10.7)$$

虚拟力

$$\vec{F}_{\text{fict}} = -2m\vec{\Omega} \times \vec{v}_{\text{rot}} - m\vec{\Omega} \times (\vec{\Omega} \times \vec{r}) \quad (10.8)$$

第一项叫Coriolis力，垂直于 $\vec{\Omega}$ 和 \vec{v} ，运动时才存在。

第二项叫离心力，方向为径向向外，大小为 $m\Omega^2\rho$ 。

利用转动的非惯性系会为分析问题带来方便。下面讨论几个由于地球转动带来的效应。

重力加速度

离心力

$$F_{\text{cent}} = m\omega^2 R \sin \theta = m\omega^2 R \cos \lambda \quad (10.9)$$

λ 为纬度。于是朝向地心的重力加速度的改变量为

$$\begin{aligned} g(\lambda) &= g - \omega^2 R \cos^2 \lambda \\ &= (g - \omega^2 R) + (\omega^2 R - \omega^2 R \cos^2 \lambda) \\ &= g_0 + \omega^2 R \sin^2 \lambda \end{aligned} \quad (10.10)$$

g_0 是赤道上的加速度。

$$g(\lambda) = 9.8(1 + 0.0035 \sin^2 \lambda) \quad (10.11)$$

自由落体的偏离

自由落体下落过程中由于速度的存在，会受到Coriolis力。科氏力为 $-2m\vec{\omega} \times \vec{v}$ ，方向指向东。与一个重锤所指的点相比，落体偏东。落体速度 $v = gt$ ，因而科氏力的大小为 $2m\omega gt \cos \lambda$ 。所以在向东的方向上的运动方程是

$$m \frac{d^2 x}{dt^2} = 2m\omega gt \cos \lambda \quad (10.12)$$

积分得到

$$x = \frac{1}{3} \omega g t^3 \cos \lambda \quad (10.13)$$

而总的下落时间 $t = \sqrt{\frac{2h}{g}}$ ，因此

$$\begin{aligned} x &= \frac{2\sqrt{2}}{3} \frac{\omega \cos \lambda}{g^{1/2}} h^{3/2} \\ &= 2 \times 10^{-5} h^{3/2} \cos \lambda (m) \end{aligned} \quad (10.14)$$

$h = 50 \text{ m}$ 时，在北（或南）纬 45° ， $x \sim 5 \text{ mm}$ 。

在非惯性系中得到落体偏东，事实上在惯性系中可以得到同样的结论。为方便，仅考虑赤道上的情形，在地面上 h 处，物体具有比地面更大的向东的速度： $v_{0x} = \omega(R+h)$ ， R 是地球半径。在 x 方向上引力的分量是

$$F_x = -\frac{x}{R} mg \quad (10.15)$$

这里用到 $x \sim \omega R t$ 当 $h \ll R$ 。于是

$$\frac{d^2x}{dt^2} = -g\omega t \quad (10.16)$$

积分得到

$$\frac{dx}{dt} = v_{0x} - \frac{1}{2}g\omega t^2 = \omega(R+h) - \frac{1}{2}g\omega t^2 \quad (10.17)$$

再积分

$$x = \omega(R+h)t - \frac{1}{6}g\omega t^3 \quad (10.18)$$

但是 O 点也在运动, 所以偏离的距离是 $\omega h t - \frac{1}{6}g\omega t^3$ 。而 $h = \frac{1}{2}gt^2$, 因而得到与用科氏力同样的结果。

傅科摆

傅科摆的摆动平面缓慢转动, 能用来作为地球本身在转动的证据。

参考 Kleppner 例 8.11 的图。 \vec{v} 在 $(\hat{r}, \hat{\theta})$ 构成的平面中。 ω 的 \hat{z} 分量是 $\omega \sin \lambda \hat{z}$ 。 \vec{v} 的 $\hat{\theta}$ 分量很小, 可忽略。

$$\begin{aligned} \vec{F}_{CH} &= -2m\omega \sin \lambda \hat{z} \times \vec{v} \\ &= -2m\omega \sin \lambda \begin{vmatrix} \hat{r} & \hat{\theta} & \hat{z} \\ 0 & 0 & 1 \\ \dot{r} & 0 & 0 \end{vmatrix} \\ &= -2m\omega \sin \lambda \dot{r} \hat{\theta} \end{aligned} \quad (10.19)$$

\vec{F}_{CH} 垂直于摆动平面, 引起平面发生转动。 $\hat{\theta}$ 方向的运动方程:

$$m(r\ddot{\theta} + 2\dot{r}\dot{\theta}) = -2m\omega \sin \lambda \dot{r} \quad (10.20)$$

即

$$r\ddot{\theta} + 2\dot{r}\dot{\theta} = -2\omega \sin \lambda \dot{r} \quad (10.21)$$

解为 $\dot{\theta} = \text{常数}$ 。这时 $\dot{\theta} = -\Omega \sin \lambda$, 沿顺时针方向进动, 周期为 $T = 2\pi/\dot{\theta} = 2\pi/\Omega \sin \lambda = 24\text{hours}/\sin \lambda$ 。

天气系统

参考 Kleppner 例 8.10 的图。

当大气中有低气压或高压中心时, 气体向内或向外流动。由于科氏力的作用, 会形成大气环流。在北半球低气压中心形成逆时针环流, 高压中心形成顺时针环流。而在赤道附近, 科氏力为零, 没有大气环流形成。

由于气压差, 大气会向中心运动。但由于科氏力, 大气开始旋转。稳定后, 压强差和科氏力提供旋转的向心力。

$$m \frac{v^2}{r} = S\Delta P - 2m\omega v \sin \lambda \quad (10.22)$$

$$m = \rho S \Delta r \quad (10.23)$$

ρ 是密度, 则

$$\frac{v^2}{r} = \frac{1}{\rho} \frac{dP}{dr} - 2v\omega \sin \lambda \quad (10.24)$$

风速

$$v = \sqrt{(r\omega \sin \lambda)^2 + \frac{r}{\omega} \frac{dP}{dr}} - r\Omega \sin \lambda \quad (10.25)$$

可达到 $> 100 \text{ km/h}$ 。

11 线性系统

我们开始探讨自然界普遍存在的波动行为。为此先补充一些数学知识。

11.1 线性系统

谐振子是线性系统的特例。一个自由谐振子的微分方程是

$$m \frac{d^2x}{dt^2} + \gamma m \frac{dx}{dt} + m\omega_0^2 x = 0 \quad (11.1)$$

一般地，一个线性系统的微分方程是

$$a_n \frac{d^n x}{dt^n} + a_{n-1} \frac{d^{n-1} x}{dt^{n-1}} + \dots + a_1 \frac{dx}{dt} + a_0 x = 0 \quad (11.2)$$

我们可以使用符号 $L(x)$ ，叫做算子，来代表线性系统。上面的方程简记为

$$L(x) = 0 \quad (11.3)$$

一个线性系统满足

$$L(x + y) = L(x) + L(y) \quad (11.4)$$

$$L(ax) = aL(x) \quad (11.5)$$

线性系统很重要，因为我们处理的大部分问题都可以近似地看作线性系统，而且我们也只能对线性系统进行有效的分析。我们经常看到和使用三角函数和指数函数，其原因就是我们在处理线性系统。三角函数和指数函数是线性系统的解。我们将使用指数函数。指数函数优于三角函数之处是能够用一个函数形式统一表示解，而不需要 \sin 、 \cos 、 \sinh 和 \cosh 四个函数。让我们来解方程 11.2。一个解是

$$x = Ae^{i\omega t} \quad (11.6)$$

将解代入方程 11.2 得到

$$a_n(i\omega)^n + a_{n-1}(i\omega)^{n-1} + \dots + a_1(i\omega) + a_0 = 0 \quad (11.7)$$

这个方程有 n 个根： $\omega_1, \omega_2, \dots, \omega_n$ 。方程 11.2 的普遍解是

$$x(t) = \sum_{j=1}^n c_j e^{i\omega_j t} \quad (11.8)$$

一般地， ω 是复数。其实部代表周期振荡，虚部代表指数衰减或增加。

如果我们有一个 n 阶的微分方程，就会有 n 个独立的解。独立的解的数目叫做系统的自由度。独立的解的任意线性组合仍然是一个解。

我们要求系数 a_i 为实数，则一个解的复共轭也是解，而所有的解与实轴成对称分布。

11.2 受迫线性系统

当一个线性系统受到外力时，方程变为

$$L(x) = F(t) \quad (11.9)$$

如果我们有了解了一个受迫系统的解，叫做 x_F （受迫解），我们可以再加上一个自由解（方程 11.8）。通常一个线性系统都有阻尼项，自由解最终会被衰减。因此自由解的部分叫做瞬态解。受迫解不会衰减，因为有外力驱动。起始时，依赖于初始条件，系统的运动会因情况不同而不同。但最终，经过一段长时间后，解变得唯一。

假设我们有一个力 F_a 作用在系统上。这时的受迫解是 x_a 。假如我们还有另外一个力 F_b ，其对应的解为 x_b 。如果两个力同时作用，则解是 x_a 和 x_b 之和，因为

$$L(x_a + x_b) = L(x_a) + L(x_b) = F_a(t) + F_b(t) \quad (11.10)$$

只是线性系统叠加原理的一个例子。它意味着如果我们有一个复杂的力，我们可以先试图把这个力分解为几个简单的力之和。所谓简单是指我们可以解相应的方程。于是系统的解就是这些解的叠加。例如，受迫谐振子的驱动力是 $F \cos \omega t$ ，阻尼项 $\gamma \dot{x}$ 会引入 $\sin \omega t$ ，导致解方程的过程变得复杂。但是我们可以先把力分解为

$$F \cos \omega t = \frac{F}{2} e^{i\omega t} + \frac{F}{2} e^{-i\omega t} = F_+(t) + F_-(t) \quad (11.11)$$

其中 $F_+(t) = (F/2)e^{i\omega t}$ ， $F_-(t) = (F/2)e^{-i\omega t}$ 。

我们可以很容易地找到 F_{\pm} 对应的解。对于 F_+ ，

$$\frac{d^2x}{dt^2} + \gamma \frac{dx}{dt} + \omega_0^2 x = \frac{F}{2m} e^{i\omega t} \quad (11.12)$$

解的形式是

$$x_+ = A_+ e^{i\omega t} \quad (11.13)$$

A_+ 满足

$$-\omega^2 A_+ + i\omega\gamma A_+ + \omega_0^2 A_+ = \frac{F}{2m} \quad (11.14)$$

解出 A_+ 得到

$$A_+ = \frac{F}{2m} \frac{1}{\omega_0^2 - \omega^2 + i\omega\gamma} \quad (11.15)$$

类似地，

$$A_- = \frac{F}{2m} \frac{1}{\omega_0^2 - \omega^2 - i\omega\gamma} \quad (11.16)$$

于是 x 的解是

$$x = x_+ + x_-$$

$$\begin{aligned}
&= \frac{F}{2m} \left(\frac{e^{i\omega t}}{\omega_0^2 - \omega^2 + i\omega\gamma} + \frac{e^{-i\omega t}}{\omega_0^2 - \omega^2 - i\omega\gamma} \right) \\
&= \frac{F}{m} \Re \left(\frac{e^{i\omega t}}{\omega_0^2 - \omega^2 + i\omega\gamma} \right) \tag{11.17}
\end{aligned}$$

这与前一章的结果相同。

我们观察到无论怎样对 $e^{i\omega t}$ 取导数，都保持同样的函数形式。如果 ω 是系统的共振频率，则 $L(e^{i\omega t}) = 0$ 。如果不是，则有 $L(e^{i\omega t}) \sim e^{i\omega t}$ 。因而，我们可以试图先将周期驱动力分解为指数函数形式，然后再将各自的响应合成。

使用代数的语言， $e^{i\omega t}$ 是算符 L 的本征函数。可以将任意函数对本征函数展开。

11.3 Fourier分析

最方便的分解驱动力的方法是用Fourier序列。我们假设驱动力 $F(t)$ 是时间的周期函数，周期为 T 。

$$F(t) = \sum_{-\infty}^{\infty} f_n e^{i\omega_n t} \tag{11.18}$$

这里 $\omega_n = 2n\pi/T$ 。序列中的每一项都是以 T 为周期的周期函数。我们使用指数函数展开的原因是，正如上一节看到的，容易找到线性系统的解。一般地只要 $F(t)$ 不是非常奇异，我们都可以有上面的展开。另外一个使用指数函数而不是 \sin 和 \cos 的好处是，我们只需要一类函数作为展开的基。当然，指数函数的自变量需要从 $-\infty$ 到 ∞ 。如果 F 为实， $f_n = f_{-n}^*$ 。

如何才能找到 f_n ？我们先看下面的事实：

$$\frac{1}{T} \int_0^T e^{i\omega_m t} e^{-i\omega_n t} dt = \delta_{mn} \tag{11.19}$$

其中

$$\delta_{mn} = \begin{cases} 1, & n = m \\ 0, & n \neq m \end{cases} \tag{11.20}$$

我们称这些指数函数是正交的。如果我们把式 11.18 两边乘以 $e^{-i\omega_m t}$ ，然后从0到 T 积分，可以得到

$$f_m = \frac{1}{T} \int_0^T F(t) e^{-i\omega_m t} dt \tag{11.21}$$

如同我们熟悉的线性空间一样，指数函数作为基矢量，也张成了一个线性空间。内积由积分给出。这些函数的正交性由式11.19体现。

由于我们的世界是线性的，所以Fourier变换及其有用。它将外部的激发分解成不同的频率，而线性系统对每个频率的响应都是简单的。

一个Fourier变换的例子：一个周期函数为

$$F(t) = \begin{cases} 1, & 0 < x < T/2 \\ -1, & -T/2 < x < 0 \end{cases} \quad (11.22)$$

容易证明

$$f_n = \frac{i}{m\pi}((-1)^n - 1) \quad (11.23)$$

这样

$$\begin{aligned} F(t) &= \frac{i}{\pi} \sum_{n=-\infty}^{\infty} \frac{1}{n}((-1)^n - 1)e^{i\omega_n t} \\ &= \frac{4}{\pi} \left(\sin \omega t + \frac{1}{3} \sin 3\omega t + \frac{1}{5} \sin 5\omega t + \dots \right) \end{aligned} \quad (11.24)$$

其中 $\omega = 2\pi/T$ 。

如果令 $t = T/4$ ，则有

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots = \frac{\pi}{4} \quad (11.25)$$

按照(11.18)式对力进行分解后，容易得到受迫谐振子的解为

$$x = \sum_{n=-\infty}^{\infty} \frac{f_n}{m} \frac{1}{\omega_0^2 - \omega_n^2 + i\gamma\omega_n} e^{i\omega_n t} \quad (11.26)$$

当 ω_n 接近 ω_0 时发生共振。

12 波动

这里先讨论机械波。

12.1 运动模式

我们要讨论多个粒子通过简谐力联系在一起时的运动行为。由于很多系统在每个粒子的平衡位置附近，粒子间的作用力可用简谐力近似，因而以下讨论具有普遍意义。

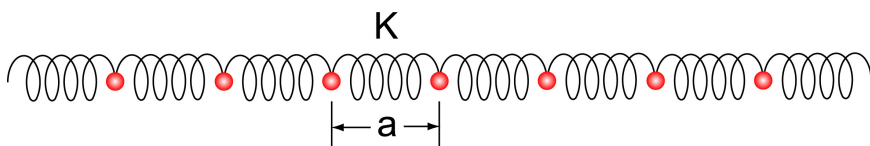


图 12.1: 一维链的运动模式

一个有 n 个粒子的一维链共有 n 个自由度。因此只要我们能够找到一组包括 n 个独立解的集合，就可以得到问题的一般解。一般解为这 n 个独立解的线性组合。我们发现一组方便的独立解是这样构成的：相应于每一个解，链上的每个粒子都以相同的频率振动，每个粒子的振幅和相位可以不同，但有关联而不能随意。这种解称为集体运动。这相当于 n 个独立的解张成了一个完备的线性空间。这些解是空间的基矢量。任何矢量可以用基矢量展开。

先考虑两个粒子的情况

$$m \frac{d^2 u_1}{dt^2} = -K(u_1 - u_2) \quad (12.1)$$

$$m \frac{d^2 u_2}{dt^2} = -K(u_2 - u_1) \quad (12.2)$$

将两运动方程相加：

$$m \frac{d^2 (u_1 + u_2)}{dt^2} = 0 \quad (12.3)$$

相应于两个质点一同平动，可以看作 $\omega_1 = 0$ 。两式相减：

$$m \frac{d^2 (u_2 - u_1)}{dt^2} = -2K(u_2 - u_1) \quad (12.4)$$

是两质点的相对运动，角频率 $\omega_2 = \sqrt{\frac{2K}{m}} = \sqrt{2}\omega_0$ 。 ω_1 和 ω_2 都是两质点运动的模式。可以看到集体运动即是粒子的坐标进行线性组合得到新的坐标。

零频率代表平动。考虑一个简谐振动, $u = A \sin \omega t$ 。 $t = 0$ 时, $\dot{u} = v$, 则 $A = v/\omega$ 。 于是 $u = \frac{v}{\omega} \sin \omega t = vt \cdot \frac{\sin \omega t}{\omega t}$ 。 当 $\omega \rightarrow 0$, $u \rightarrow vt$ 。

利用矩阵我们可以将这种线性组合的做法变得更加形式化。运动方程可以写成矩阵形式

$$\frac{d^2}{dt^2} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} -\omega_0^2 & \omega_0^2 \\ \omega_0^2 & -\omega_0^2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \quad (12.5)$$

令

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \quad \Omega = \begin{pmatrix} -\omega_0^2 & \omega_0^2 \\ \omega_0^2 & -\omega_0^2 \end{pmatrix} \quad (12.6)$$

Ω 是对称矩阵。运动方程可以记为

$$\frac{d^2 u}{dt^2} = \Omega u \quad (12.7)$$

利用矩阵形式, 粒子坐标的线性组合可表示为 $u' = Su$ 。 u' 满足方程

$$\frac{d^2 u'}{dt^2} = S \Omega S^{-1} u' \quad (12.8)$$

如果 u' 的每个分量都代表一个集体模式, 则 $S \Omega S^{-1}$ 必须是对角矩阵。因此寻找集体模式的过程即是将对角化的过程。线性代数可以告诉我们对于 Ω 这样一个对称矩阵, 总可以找到一个矩阵 S 使其对角化, 因此总可以找到两个模式。例如对于上述两粒子问题,

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad (12.9)$$

利用一种集体模式中每个粒子的频率相同这样一个特点可以假设解为

$$u_1 = A_1 e^{i\omega t} \quad (12.10)$$

$$u_2 = A_2 e^{i\omega t} \quad (12.11)$$

于是,

$$(m\omega^2 - K)A_1 + KA_2 = 0 \quad (12.12)$$

$$KA_1 + (m\omega^2 - K)A_2 = 0 \quad (12.13)$$

两个解分别为 $\omega_1 = 0$ 和 $\omega_2 = \sqrt{\frac{2K}{m}} = \sqrt{2}\omega_0$ 。 对应于 ω_1 有 $A_1 = A_2$, 代表两个粒子一起平动。 对应于 ω_2 有 $A_1 = -A_2$, 代表两个粒子以同样频率相向运动。 ω_1 和 ω_2 是系统的两个独立的运动模式, 一般的解是两个模式的线性组合。

一个一般地得到 A_1 和 A_2 的做法是使用矩阵。这需要用到一个重要的结论: 如果 $Ax = 0$ 有 x 的非零解, 则矩阵 A 的行列式 $|A|$ 必为零。有关 A_1 和 A_2 的方程可以写成矩阵的形式

$$\begin{pmatrix} \omega_0^2 & -\omega_0^2 \\ -\omega_0^2 & \omega_0^2 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = \omega^2 \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \quad (12.14)$$

或者

$$\begin{pmatrix} \omega_0^2 - \omega^2 & -\omega_0^2 \\ -\omega_0^2 & \omega_0^2 - \omega^2 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = 0 \quad (12.15)$$

行列式

$$\begin{vmatrix} \omega_0^2 - \omega^2 & -\omega_0^2 \\ -\omega_0^2 & \omega_0^2 - \omega^2 \end{vmatrix} = 0 \quad (12.16)$$

即

$$(\omega_0^2 - \omega^2)^2 = \omega_0^4 \quad (12.17)$$

解为

$$\omega_1 = 0, \quad A_1 = A_2 \quad (12.18)$$

$$\omega_2 = \sqrt{2}\omega_0, \quad A_1 = -A_2 \quad (12.19)$$

以上的过程不仅仅适用于两粒子的情形。考虑三个粒子的运动

$$m \frac{d^2 u_1}{dt^2} = -K(u_1 - u_2) \quad (12.20)$$

$$m \frac{d^2 u_2}{dt^2} = -K(u_2 - u_1) - K(u_2 - u_3) \quad (12.21)$$

$$m \frac{d^2 u_3}{dt^2} = -K(u_3 - u_2) \quad (12.22)$$

试探解: $u_i = A_i e^{i\omega t}$ (为了方便取复数解) 代入运动方程

$$(m\omega^2 - K)A_1 + KA_2 = 0 \quad (12.23)$$

$$KA_1 + (m\omega^2 - 2K)A_2 + KA_3 = 0 \quad (12.24)$$

$$KA_2 + (m\omega^2 - K)A_3 = 0 \quad (12.25)$$

如果 $K \neq m\omega^2$, 则

$$A_1 = -\frac{K}{m\omega^2 - K} A_2 \quad (12.26)$$

$$A_3 = -\frac{K}{m\omega^2 - K} A_2 \quad (12.27)$$

带入第二个方程得到

$$\left(\frac{2K^2}{m\omega^2 - K} + (2K - m\omega^2) \right) A_2 = 0 \quad (12.28)$$

于是

$$(m\omega^2)^2 = 3K(m\omega^2) \quad (12.29)$$

解为

$$\omega_1 = 0, \quad \omega_3 = \sqrt{\frac{3K}{m}} = \sqrt{3}\omega_0 \quad (12.30)$$

当 $\omega_1 = 0$ 时, $A_1 = A_2 = A_3$, 三个粒子同时平动。

当 $\omega_3 = \sqrt{3}\omega_0$ 时, $A_1 = A_3 = -\frac{1}{2}A_2$ 。

如果 $K = m\omega^2$, 代入方程, 发现也是解, 这时 $\omega_2 = \omega_0$, $A_2 = 0$, $A_1 = -A_3$ 。

我们得到三个集体运动的解, 频率分别为 $\omega_1 = 0$, $\omega_2 = \omega_0$ 和 $\omega_3 = \sqrt{3}\omega_0$ 。每个模式中各个粒子的振幅如图12.2所示。

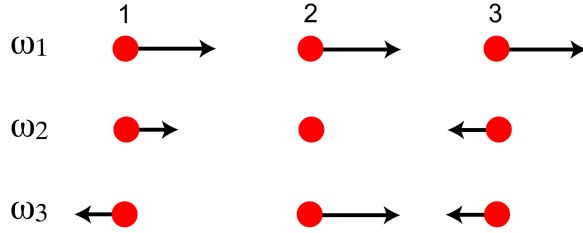


图 12.2: 三粒子的运动模式

也可以用矩阵得到解

$$\begin{pmatrix} \omega_0^2 & -\omega_0^2 & 0 \\ -\omega_0^2 & 2\omega_0^2 & -\omega^2 \\ 0 & -\omega_0^2 & \omega_0^2 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} = \omega^2 \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} \quad (12.31)$$

由

$$\begin{vmatrix} \omega_0^2 - \omega^2 & -\omega_0^2 & 0 \\ -\omega_0^2 & 2\omega_0^2 - \omega^2 & -\omega^2 \\ 0 & -\omega_0^2 & \omega_0^2 - \omega^2 \end{vmatrix} = 0 \quad (12.32)$$

得到

$$\omega^2(\omega_0^2 - \omega^2)(\omega^2 - 3\omega_0^2) = 0 \quad (12.33)$$

也可得到同样的三个解: 0 , ω_0 和 $\sqrt{3}\omega_0$ 。

也可以利用矩阵对角化的方式得到集体模式。可以证明对于三个粒子

$$S = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \\ 2 & -1 & 2 \end{pmatrix} \quad (12.34)$$

这时三个集体模式的坐标分别为

$$\begin{aligned} u'_1 &= u_1 + u_2 + u_3 \\ u'_2 &= u_1 - u_3 \\ u'_3 &= 2u_1 - u_2 + 2u_3 \end{aligned} \quad (12.35)$$

而

$$S\Omega S^{-1} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \omega_0^2 & 0 \\ 0 & 0 & 3\omega_0^2 \end{pmatrix} \quad (12.36)$$

于是 u'_1 、 u'_2 和 u'_3 三个集体坐标的频率分别为 0 、 ω_0 和 $\sqrt{3}\omega_0$ 。

集体运动一个模式意味着每个粒子都以同样的频率运动。但由于粒子之间有关联，所以每个粒子的运动的振幅和相位有关联，不能随意。

集体运动模式是特解，每个模式有单一的频率。线性系统一般的运动是集体模式的线性组合。一般的，运动是各个集体运动模式的线性叠加，每个模式都有单一频率。叠加后，每个模式仍独立存在，叠加能产生无穷多种运动行为。

叠加的一个例子

例如在三个粒子时，不考虑 ω_1 ，另两个模式的运动分别为

$$\omega_2 = \omega_0: \quad u_1 = Ae^{i\omega_2 t}, \quad u_2 = 0, \quad u_3 = -Ae^{i\omega_2 t} \quad (12.37)$$

$$\omega_3 = \sqrt{3}\omega_0: \quad u_1 = -\frac{1}{2}Be^{i\omega_3 t}, \quad u_2 = Be^{i\omega_3 t}, \quad u_3 = -\frac{1}{2}Be^{i\omega_3 t} \quad (12.38)$$

一般地，将两种模式叠加起来，取实部（假设A和B为实）

$$u_1 = A \cos \omega_2 t - \frac{1}{2}B \cos \omega_3 t \quad (12.39)$$

$$u_2 = B \cos \omega_3 t \quad (12.40)$$

$$u_3 = -A \cos \omega_2 t - \frac{1}{2}B \cos \omega_3 t \quad (12.41)$$

第一和第三两个粒子的运动都包含 ω_2 ， ω_3 的成份，第二个粒子只有 ω_3 。如果 $B = 0$ ，只有第一种模式，如果 $A = 0$ ，只有第二种模式。假如 $A = 1$ ， $B = 2$

$$u_1 = \cos \omega_2 t - \cos \omega_3 t = -2 \sin \frac{\omega_2 + \omega_3}{2} t \sin \frac{\omega_2 - \omega_3}{2} t \quad (12.42)$$

$$u_3 = -\cos \omega_2 t - \cos \omega_3 t = -2 \cos \frac{\omega_2 + \omega_3}{2} t \cos \frac{\omega_2 - \omega_3}{2} t \quad (12.43)$$

在 u_1 为0时， u_3 达到最大值，反之亦然，能量在1和3之间传递，这是一种波动行为。

我们发现两个粒子有两个模式，三个粒子有三个模式。利用前面关于矩阵对角化的讨论可以知道N个粒子的系统会有N个模式。这样的系统推广到无穷多粒子应该有无穷多模式（图12.3）。

12.2 一维环的运动模式

一维链模型的一个不方便的地方是端点的两个粒子与其他粒子所处环境不同。为了讨论问题方便我们将两个端点用同样的弹簧连接起来形成环。这样

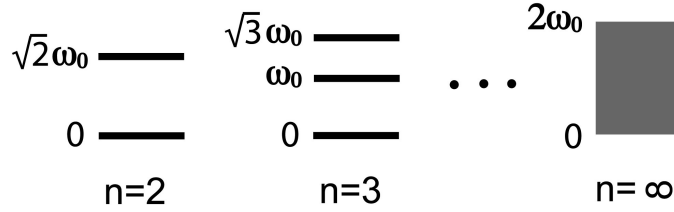


图 12.3: 不同粒子数的模式数

的系统不再有端点。如果N很大，则开放和闭合环在运动模式上的差别可以忽略。

先看三个粒子的情形。这时需要解的方程是

$$\begin{pmatrix} 2\omega_0^2 & -\omega_0^2 & -\omega_0^2 \\ -\omega_0^2 & 2\omega_0^2 & -\omega^2 \\ -\omega_0^2 & -\omega_0^2 & 2\omega_0^2 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} = \omega^2 \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} \quad (12.44)$$

该方程为线性齐次方程。只有当 ω 为某些特殊值才有非零解。对于这样的一个 ω ，方程的任一个解可由几个线性无关的解表示出来。这几个线性无关的解称为基础解系。线性无关解的个数称为基础解系的维度。

可以证明三个模式分别为

$$\omega_1 = 0 : A_1 = 1, A_2 = 1, A_3 = 1 \quad (12.45)$$

$$\omega_2 = \sqrt{3}\omega_0 : A_1 = 1, A_2 = e^{-i\frac{2}{3}\pi}, A_3 = e^{-i\frac{4}{3}\pi} \quad (12.46)$$

$$\omega_3 = \sqrt{3}\omega_0 : A_1 = 1, A_2 = e^{i\frac{2}{3}\pi}, A_3 = e^{i\frac{4}{3}\pi} \quad (12.47)$$

对于 ω_1 ，基础解系的维度为1。后两种模式频率相同，称为简并。对应的情况是当 $\omega = \sqrt{3}\omega_0$ 时，基础解系的维度为2。

将第j个模式中第n个粒子的位移表示为统一的形式： $u_n^j = A \exp(i\omega_j t - i\phi_n^j)$ 。 $n = 0, 1, \dots, N-1$ 。 ϕ_n^j 表示第n个粒子在以第j个模式振动时的相位。可以发现在同一个模式中，相邻粒子的相位差固定。例如三个粒子的系统中，令 $\Delta\phi^j = \phi_1^j - \phi_2^j = \phi_2^j - \phi_3^j = \phi_3^j - \phi_1^j$ 。第一个模式中 $\Delta\phi^1 = 0 \times 2\pi/3$ ，第二个模式中 $\Delta\phi^2 = 1 \times 2\pi/3$ ，第三个模式中 $\Delta\phi^3 = (-1) \times 2\pi/3$ 。都是把 2π 分成三等分后再乘以一个整数。

对于四个粒子，运动方程对应于

$$\begin{pmatrix} 2\omega_0^2 & -\omega_0^2 & 0 & -\omega_0^2 \\ -\omega_0^2 & 2\omega_0^2 & -\omega^2 & 0 \\ 0 & -\omega_0^2 & 2\omega_0^2 & -\omega_0^2 \\ -\omega_0^2 & 0 & -\omega_0^2 & 2\omega_0^2 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{pmatrix} = \omega^2 \begin{pmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{pmatrix} \quad (12.48)$$

$$\omega_1 = 0 : A_1 = 1, A_2 = 1, A_3 = 1; A_4 = 1 \quad (12.49)$$

$$\omega_2 = \sqrt{2}\omega_0 : A_1 = 1, A_2 = e^{-i\frac{1}{2}\pi}, A_3 = e^{-i\pi}; A_4 = e^{-i\frac{3}{2}\pi} \quad (12.50)$$

$$\omega_3 = \sqrt{2}\omega_0 : A_1 = 1, A_2 = e^{i\frac{1}{2}\pi}, A_3 = e^{i\pi}; A_4 = e^{i\frac{3}{2}\pi} \quad (12.51)$$

$$\omega_4 = 2\omega_0 : A_1 = 1, A_2 = e^{-i\pi}, A_3 = e^{-i2\pi}; A_4 = e^{-i3\pi} \quad (12.52)$$

令 $\Delta\phi^j = \phi_1^j - \phi_2^j = \phi_2^j - \phi_3^j = \phi_3^j - \phi_4^j = \phi_4^j - \phi_1^j$ 。第一个模式中 $\Delta\phi^1 = 0 \times 2\pi/4$ ，第二个模式中 $\Delta\phi^2 = 1 \times 2\pi/4$ ，第三个模式中 $\Delta\phi^3 = (-1) \times 2\pi/4$ ，第四个模式中 $\Delta\phi^4 = 2 \times 2\pi/4$ 。都是把 2π 分成四等分后再乘以一个整数。

以上的行为具有普遍性。一个 N 个粒子的环中，相邻粒子的相位差固定，第 n 个粒子的相位 $\phi_n^j = n\Delta\phi^j$ ， $\Delta\phi^j$ 是 $2\pi/N$ 的整数倍。定义 $k^j = \Delta\phi^j/a$ ，是单位长度上的相位差，称为波数 (wave number)。波数表示一个模式在空间上的分布。例如，对于三个粒子的三个模式，波数分别为 $k_1 = 0$ ， $k_2 = 2\pi/3a$ 和 $k_3 = -2\pi/3a$ ；对于四个粒子的四个模式，波数分别为 $k_1 = 0$ 、 $k_2 = \pi/2a$ 、 $k_3 = -\pi/2a$ 和 $k_4 = \pi/a$ 。定义了波数，第 n 个粒子在第 j 个模式中的位移可以统一写为

$$u_n^j = Ae^{i(\omega_j t - k_j n a)} \quad (12.53)$$

k 通过 j 与 ω 有函数关系，称为色散关系。容易看出对于三个粒子， (ω_j, k_j) 三个点都落在 $\omega = 2\omega_0 \left| \sin \frac{ka}{2} \right|$ 这样一条曲线上。对于四个粒子， (ω_j, k_j) 四个点也落在同一条曲线上。

还可以尝试比如六个粒子， $(\omega_1, k_1) = (0, 0)$ 、 $(\omega_2, k_2) = (\omega_0, \pi/3a)$ 、 $(\omega_3, k_3) = (\omega_0, -\pi/3a)$ 、 $(\omega_4, k_4) = (\sqrt{3}\omega_0, 2\pi/3a)$ 、 $(\omega_5, k_5) = (\sqrt{3}\omega_0, -2\pi/3a)$ 和 $(\omega_6, k_6) = (2\omega_0, \pi/a)$ 。这六个点也落在 $\omega = 2\omega_0 \left| \sin \frac{ka}{2} \right|$ 这样一条曲线上。

现在考虑无穷多个粒子。粒子的编号为 $x_n = na$ ， $n = -\infty, \dots, \infty$ 。粒子只沿链方向运动， u_n 是粒子 n 的位移。前面提到无穷长链中有无穷多种模式。第 n 个粒子的运动方程是：

$$m \frac{d^2 u_n}{dt^2} = K(u_{n+1} - u_n) - K(u_n - u_{n-1}) \quad (12.54)$$

或

$$\frac{d^2 u_n}{dt^2} = \omega_0^2 (u_{n+1} - 2u_n + u_{n-1}) \quad (12.55)$$

我们仍然试图求出每个粒子的运动都有相同频率的模式。方程的解为：

$$u_n(t) = A_n e^{i\omega t} \quad (12.56)$$

于是

$$-\omega^2 A_n = \omega_0^2 (A_{n+1} - 2A_n + A_{n-1}) \quad (12.57)$$

可以验证解为

$$A_n = Ae^{-ikna} = Ae^{-ikx_n} \quad (12.58)$$

其中波数 k 仍然是单位长度上相位差，与 ω 满足色散关系

$$\omega = 2\omega_0 \left| \sin \frac{ka}{2} \right|, \quad -\frac{\pi}{a} < k < \frac{\pi}{a} \quad (12.59)$$

于是有

$$u_n(t) = Ae^{i(\omega t - kna)} \quad (12.60)$$

模式用波数 k 来标记。一对 ω 和 k 对应一个解、一个模式。这个模式中每个粒子都在以相同的频率振动，但不同位置的粒子，振动的相位 $(\omega t - kna)$ 不同。对于同一时刻相邻的两个粒子相位差是 ka (图12.4)。因而 k 为单位长度上相位的变化。如果 $k > 0$ ， $n+1$ 粒子比 n 粒子相位落后 ka ，需要 ka/ω 的时间 $n+1$ 的粒子的相位才能赶上。 ω 越大， k 就越大，相位差也就越大。这是因为频率越大，下一个粒子就越跟不上前一个粒子的运动。如果 $k < 0$ ，那么左边粒子的相位落后于右边粒子的相位。所以在 $k > 0$ 时相位向右传播， $k < 0$ 时相位向左传播。 k 给出相位传播方向和速度。相位沿 k 方向传播。相位的传播速度为：

$$v_p = a \left/ \frac{ka}{\omega} \right. = \frac{\omega}{k} \quad (12.61)$$

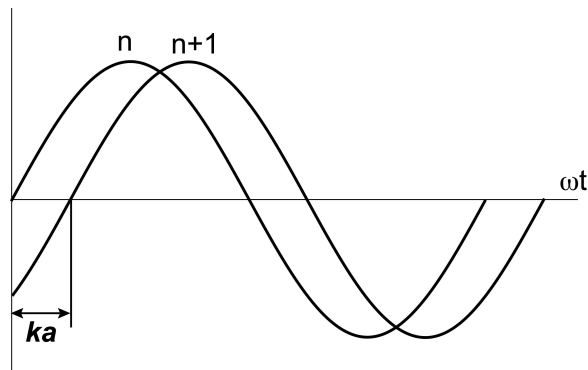


图 12.4: k 与相位的关系

再考虑某一确定时刻，运动在空间中的分布：

$$u_n(t) = A \cos(\omega t - nka) = A \cos(\omega t - kx_n) \quad (12.62)$$

对于一定频率，相邻粒子的相位差一定，隔一段距离后相位差会改变 2π 。每经过空间上 $2\pi/k$ 的距离， $u_n(k)$ 重复一次。因而空间上具有 $\lambda = 2\pi/k$ 的周期， λ 叫波长。 λ 越小 ω 越大，因为小 λ 意味着粒子间的强耦合。

我们把上面解出的运动模式叫行波解。在一个模式中，每个粒子有相同的振动频率，空间上波形是一个正弦波。该正弦波沿某方向移动，时间上的周期是 $2\pi/\omega$ ，空间上的周期为 $\lambda = 2\pi/k$ 。在一个模式中，每个粒子在平衡位置附近做简谐振动，其振幅相同，相位不同。位移成正弦分布。前后两个粒子的相位差为一确定值。

12.3 色散关系

$$\omega = 2 \left| \omega_0 \sin \frac{ka}{2} \right| \quad (12.63)$$

以 $2\pi/a$ 为周期。

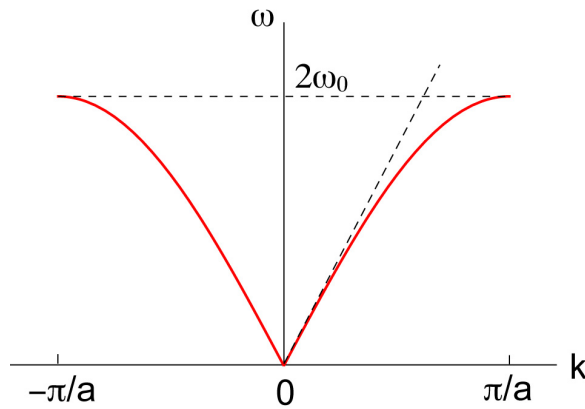


图 12.5: 一维链的色散关系

原则上 k 的取值范围是 $-\infty$ 到 ∞ 。但按照上式， k 和 $k+2\pi/a$ 给出同样的 ω ，实际上只有 $-\pi/a$ 到 π/a 的 k 才是独立的。对于 $k+2\pi/a$ ， $u_n(t) = Ae^{i(\omega t - n(k + \frac{2\pi}{a})a)} = Ae^{i(\omega t - nka)}$ ，给出与 k 相同的位移。因而 k 与 $k + 2\pi/a$ 描述的是同一个波。例如， $k = \frac{\pi}{a}$ ($\lambda = 2a$) 和 $k = \frac{\pi}{a} + \frac{2\pi}{a} = \frac{3\pi}{a}$ ($\lambda = \frac{2a}{3}$) 两个波所对应的粒子的位移完全相同 (图 12.6)。在链上波长最短为 $2a$ 。比 $2a$ 小的波长总可以找到一个比 $2a$ 大的波长与之对应，二者有相同的振动行为。相邻粒子的相位可差 2π 的整数倍，描述相同的运动。当 $k = \pi/a$ 时，相邻粒子的运动反向，能量不再传递。

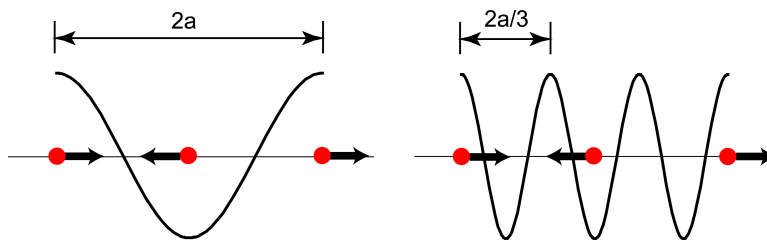


图 12.6: $k = \pi/a$ 和 $k = 3\pi/a$ 是相同的运动

由色散关系还可以得知， ω 的最大值是 $2\omega_0$ 。当 $\lambda = 2a$ 时， ω 达到最大。对于无限多的粒子，集体运动模式的频率展宽为 $2\omega_0$ 。

$2\omega_0$ 称为截止频率 (cut-off frequency)。用小于 $2\omega_0$ 的频率去激发，运动会传播。如果 $\omega > 2\omega_0$ 会发生什么？假如在链的一端用大于 $2\omega_0$ 的频率去激发，方程 $d^2u_n/dt^2 = \omega_0^2(u_{n+1} - 2u_n + u_{n-1})$ 不会有 $u_n(t) = Ae^{i(\omega t - nka)}$ 形式的解。但如果让 k 为虚数 $k = -i\kappa$ ，则 $u_n(t) = Ae^{i\omega t - n\kappa a}$ 仍为上式的解。因此，为使 $\omega > 2\omega_0$ ， k 必须为虚数。但是，

$$\begin{aligned} \omega^2 &= \omega_0^2(e^{\kappa a} - 2 + e^{-\kappa a}) \\ &= \omega_0^2(e^{\kappa a/2} - e^{-\kappa a/2})^2 \end{aligned}$$

$$= \omega_0^2 \cdot 4 \sinh^2 \frac{\kappa a}{2} \quad (12.64)$$

即：

$$\omega = 2\omega_0 \sinh \frac{\kappa a}{2} \quad (12.65)$$

这时 ω 大于 $2\omega_0$, $u_n(t)$ 在空间上呈指数衰减, 大于截止频率的波不能传播, 质点形成的链相当于一个低通滤波器。截止频率来源于系统的分立结构。当波长与 a 可比拟时, 色散关系的曲线开始下弯曲, 在 $\lambda = 2a$ 时达到 ω 的最大值。再短的波长不再给出更高的频率。能量传递需要时间。下一个粒子需要一定的反应时间。在下一个粒子反应过来前, 如前一个粒子的运动已发生很大的变化, 会发生反向推动, 总的能量传递为0。如前所述, 在 $\lambda \gg a$ 时, 空间中粒子的分立结构不再重要。这时, 可以将波传播的媒质看成是连续的。当 $\lambda \gg a$ (或 $ka \ll 1$) 时, $\omega \simeq \omega_0 ka$, 相速度 $v_p \simeq \omega_0 a = c$ 为常数。这时色散关系近似为一条直线, 斜率为 v_p , 这类似于空气中的声波。空气中的声波对于不同的频率有相同的相速度, 即没有色散, 否则利用声音进行通讯是不可想象的。

12.4 长波极限

在长波极限下 ($\lambda \gg a$), 分立的运动方程变为连续。重新考虑运动方程

$$\frac{d^2 u_n}{dt^2} = \omega_0^2 (u_{n+1} - 2u_n + u_{n-1}) \quad (12.66)$$

$\lambda \gg a$ 时, 相邻粒子的 u 差别不大, 利用Taylor展开,

$$\begin{aligned} & u_{n+1} - 2u_n + u_{n-1} \\ &= (u(x_n + a, t) - u(x_n, t)) - (u(x_n, t) - u(x_n - a, t)) \\ &= \frac{\partial u}{\partial x} a + \frac{1}{2} \frac{\partial^2 u}{\partial x^2} a^2 - \frac{\partial u}{\partial x} a + \frac{1}{2} \frac{\partial^2 u}{\partial x^2} a^2 \\ &= \frac{\partial^2 u}{\partial x^2} a^2 \end{aligned} \quad (12.67)$$

因而运动方程成为：

$$\frac{\partial^2 u}{\partial t^2} - a^2 \omega_0^2 \frac{\partial^2 u}{\partial x^2} = 0 \quad (12.68)$$

即

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0 \quad (12.69)$$

$c = \omega_0 a$ 就是前面讨论相速度在 $\lambda \gg a$ 极限下的值。上面方程的单频解为 $u(x, t) = Ae^{i(\omega t - kx)}$ 。代入得到 $\omega = kc$, 与前面的分析一致。 $\lambda \gg a$ 时, v_p 为常数, $c = \sqrt{\frac{K}{m}} a$ 。利用这个 c 的公式可以写出杆中的声速。杆的密度为

ρ , $\rho = m/Sa$, S 是截面积。而弹簧系数 K 正比于 S/a , 这可以用串联弹簧和并联弹簧得出, 因而 $K = Y(S/a)$, Y 叫杨氏模量。所以 $c = \sqrt{\frac{Y}{\rho}}$ 。一般地, $Y \sim 10^{10} N/m^2$, $\rho \sim 10^3 - 10^4 kg/m^3$, 则 $c \sim 10^3 - 10^4 m/s$ 。

12.5 波动中的能量

在波动行为中, 波的载体并不传播, 但波这种运动形式在传播。在前面的例子中, 每个粒子都在平衡位置周围振动, 但是相位和能量可以沿 k 的方向传播。相邻的两个粒子间有相位差。这意味着一个粒子可以对它周围的粒子做功, 从而产生能量的传播。如两个粒子完全同相, 彼此不做功。

粒子 n 作用在 $(n+1)$ 上的力为

$$\begin{aligned} & -K(\Re(u_{n+1}) - \Re(u_n)) \\ & = -KA(\cos(\omega t - (n+1)ka) - \cos(\omega t - nka)) \end{aligned} \quad (12.70)$$

第 $(n+1)$ 个粒子的速度为

$$\frac{d(\Re(u_{n+1}))}{dt} = -\omega A \sin(\omega t - (n+1)ka) \quad (12.71)$$

粒子 n 作用在 $(n+1)$ 上的平均功率为($\omega = 2\pi/T$)

$$\begin{aligned} \langle P \rangle & = \frac{1}{T} \int_0^T KA(\cos(\omega t - (n+1)ka) \\ & \quad - \cos(\omega t - nka))\omega A \sin(\omega t - (n+1)ka) dt \\ & = \frac{\omega KA^2}{T} \left(\int_0^T \cos(\omega t - (n+1)ka) \sin(\omega t - (n+1)ka) dt \right. \\ & \quad \left. - \int_0^T \cos(\omega t - nka) \sin(\omega t - (n+1)ka) dt \right) \end{aligned} \quad (12.72)$$

其中第一个积分为0, 第二个积分是

$$\begin{aligned} & \frac{1}{T} \int_0^T \cos(\omega t - nka) \sin(\omega t - (n+1)ka) dt \\ & = \frac{1}{T} \int_0^T \cos(\omega t - nka) (\sin(\omega t - nka) \cos ka \\ & \quad - \cos(\omega t - nka) \sin ka) dt \\ & = -\frac{1}{T} \int_0^T \cos^2(\omega t - nka) \sin ka dt \\ & = -\frac{1}{2} \sin ka \end{aligned} \quad (12.73)$$

所以，将色散关系代入得到：

$$\begin{aligned}
 \langle P \rangle &= \frac{\omega K A^2}{2} \sin ka \\
 &= \omega_0 K A^2 \sin \frac{ka}{2} \sin ka \\
 &= 2\omega_0 K A^2 \sin^2 \frac{ka}{2} \cos \frac{ka}{2}
 \end{aligned} \tag{12.74}$$

这是粒子 n 作用在 $n+1$ 上的平均功率。这里 $k > 0$ 。当 $k = \pi/a$ 时， $\langle P \rangle = 0$ 。原因是 $k = \pi/a$ 时，相邻粒子的相位差 π ，两粒子完全相向或相反运动，总的功率为0，尽管这时相速度为 $v_p = 2\omega_0/k = 2\omega_0 a/\pi$ ，但无能量传播。得到功率后，下面我们计算能量的传播速度。从前面的讨论得知能量不是以相速度 v_p 传播。当 $k = \pi/a$ 时，甚至没有能量传播。原因是一个周期内作的功（ n 对 $n+1$ ）有一部分是正的，一部分是负的，而没有利用全部时间让 n 对 $n+1$ 作正功。所以在这种情况下，能量传播速度小于相速度。能量传播速度的计算可参照流体情形。在单位时间内通过一截面的质量叫做通量。而单位时间内通过一截面的质量等于速度 v 乘以单位长度内流体的质量 ρ ，所以速度 v 等于通量除以流体线密度 ρ 。完全类似，能量传播的速度等于功率（相当于能量通量）除以能量密度。前面已经计算了功率，下面计算能量密度。第 n 到 $n+1$ 这一段的平均动能：

$$\begin{aligned}
 \langle E_k \rangle &= \frac{1}{2} m (\dot{u}_n^2 + \dot{u}_{n+1}^2) / 2 \\
 &= \frac{1}{4} m \omega^2 A^2 (\langle \sin^2(\omega t - nka) \rangle \\
 &\quad + \langle \sin^2(\omega t - (n+1)ka) \rangle) \\
 &= \frac{1}{4} m \omega^2 A^2 \\
 &= K A^2 \sin^2 \frac{ka}{2}
 \end{aligned} \tag{12.75}$$

第一行的 $1/2$ 是由于每个粒子都属于左右两段。

平均势能：

$$\begin{aligned}
 \langle E_p \rangle &= \frac{1}{2} K \langle (u_{n+1} - u_n)^2 \rangle \\
 &= \frac{1}{2} K A^2 \langle (\cos(\omega t - (n+1)ka) - \cos(\omega t - nka))^2 \rangle \\
 &= \frac{1}{2} K A^2 \langle \cos^2(\omega t - (n+1)ka) + \cos^2(\omega t - nka) \\
 &\quad - 2 \cos(\omega t - (n+1)ka) \cos(\omega t - nka) \rangle \\
 &= \frac{1}{2} K A^2 (1 - 2 \langle \cos(\omega t - (n+1)ka) \cos(\omega t - nka) \rangle) \\
 &= \frac{1}{2} K A^2 (1 - 2 \langle \cos^2(\omega t - nka) \rangle \cos ka \\
 &\quad + 2 \langle \sin(\omega t - nka) \cos(\omega t - nka) \rangle \sin ka)
 \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} K A^2 (1 - \cos ka) \\
&= K A^2 \sin^2 \frac{ka}{2}
\end{aligned} \tag{12.76}$$

与平均动能相同

单位长度内的平均能量

$$w = \frac{\langle E_k \rangle + \langle E_p \rangle}{a} = 2 \frac{K A^2}{a} \sin^2 \frac{ka}{2} \tag{12.77}$$

于是能量传播速度

$$\begin{aligned}
v_g &= \frac{\langle P \rangle}{w} \\
&= \frac{2\omega_0 K A^2 \sin^2 \frac{ka}{2} \cos \frac{ka}{2}}{2 \frac{K A^2}{a} \sin^2 \frac{ka}{2}} \\
&= \omega_0 a \cos \frac{ka}{2}
\end{aligned} \tag{12.78}$$

与相速度 $v_p = \frac{2\omega_0}{k} \sin \frac{ka}{2}$ 不同。当 $k = \frac{\pi}{a}$ 时, $v_g = 0$ 与前面的分析相同。

更普遍的求 v_g 的方法是跟踪波包。平面波传播能量, 但我们没有好的办法跟踪能量的运动。而如果我们制造非均匀的波, 如波包, 我们则可以通过跟踪波包来定出能量的速度。波包的传播就是一包能量的传播, 制造波包的简单办法是用两个 ω 相近的波的叠加 (图 12.7)

$$\begin{aligned}
&\cos(\omega_1 t - k_1 x) + \cos(\omega_2 t - k_2 x) \\
&= 2 \cos \left(\frac{\omega_1 + \omega_2}{2} t - \frac{k_1 + k_2}{2} x \right) \cos \left(\frac{\omega_1 - \omega_2}{2} t - \frac{k_1 - k_2}{2} x \right)
\end{aligned} \tag{12.79}$$

当 ω_1 与 ω_2 接近时, 第一个 \cos 相当于以平均 ω 和 k 传播的波, 第二个 \cos 相当于加在振幅上的调制, 在空间上和时间上都有很长的周期。

考虑振幅的调制。调制项按照 $\cos(\frac{\omega_1 - \omega_2}{2} t - \frac{k_1 - k_2}{2} x)$ 运动。这表明上面的波包以 $\frac{\omega_1 - \omega_2}{k_1 - k_2}$ 的速度运动 (参照前面相速度的讨论)。由于 ω_1 与 ω_2 很接近, 所以 $v_g = d\omega/dk$ 。

对于一维链, $\omega = 2\omega_0 \sin \frac{ka}{2}$, 所以

$$v_g = \frac{d\omega}{dk} = 2\omega_0 \frac{a}{2} \cos \frac{ka}{2} = \omega_0 a \cos \frac{ka}{2} \tag{12.80}$$

与前面一致。

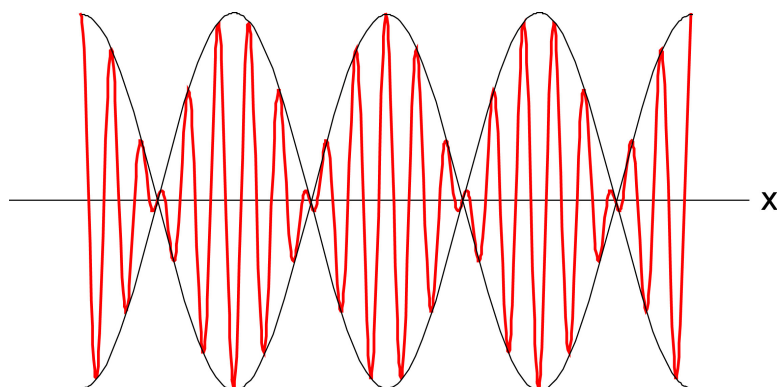


图 12.7: 拍

12.6 横波、纵波

以上讨论的都是纵波，即粒子运动方向与波传播方向平行。在流体中通常只有纵波，因为不能有切向形变。例如空气、水中的声波。但固体中可以有横波，可以有切变。横波中粒子的运动方向与波传播方向垂直。

12.7 波的一般形式

$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0$ 的解 $u(x, t) = Ae^{i(\omega t - kx)}$ 或 $A \cos(\omega t - kx)$ 描述的是无限媒质中单一频率，单一 k 的波，叫行波。行波给出了介质中波的全部模式，但与前面三个粒子的分析相同，波的一般形式是各种模式的一个叠加，如同 Fourier 变换一样。行波解给出了一组展开的基矢量。行波有很多应用，易于产生，更重要的是有单一频率和动量，可用于探测物质中的激发。但一般的解，如脉冲也有重要的用途，如用短脉冲可研究动力学、测寿命等。上面的波动方程描述了无色散的波动，一般解为 $u = f(x - ct) + g(x + ct)$ ， f 和 g 为任意函数。证明如下： $\frac{\partial f}{\partial t} = -cf'$ ， $\frac{\partial^2 f}{\partial t^2} = c^2 f''$ ， $\frac{\partial^2 f}{\partial x^2} = f''$ ，所以 u 是解。在运动中，通解始终可用 f 和 g 表示，波形不变，原因是没有色散。通解都可以表示一系列行波的叠加。

驻波

考虑行波模式的一种特殊的叠加

$$\begin{aligned} u(x, t) &= \frac{A}{2} \cos(\omega t - kx) - \frac{A}{2} \cos(\omega t + kx) \\ &= A \sin \omega t \sin kx \end{aligned} \quad (12.81)$$

这是两个相反运动的行波的叠加。从图像上看每一点都以 ω 频率振动，但波形不会移动。 $\sin kx$ 代表了一个固定的波形，叫驻波。驻波每一点的振幅只与 x 有关，而不随 t 变化，驻波不传递能量。

在一个两端固定的弦中, $u(0) = u(l) = 0$, 所以 k 必须满足: $k = \frac{n\pi}{l}$, $\lambda = \frac{2l}{n}$ 。

$$u(x, t) = A \sin \omega t \sin \frac{n\pi x}{l} \quad (12.82)$$

$n = 1, 2, 3, \dots$ 是弦中的模式。在驻波的节点处位移永远是零。

13 狭义相对论的基本假设

13.1 假设一

牛顿力学中一个非常不自然的特征是相互作用是超距的、瞬时传播的。这时每个粒子所受到的其他粒子的力仅由那些粒子瞬时的运动状态有关。我们更希望相互作用的传播需要时间。当一个体系中的一个粒子的状态发生变化时，经过一段时间后另一个粒子才会受到影响。两粒子的距离除以这个时间就是相互作用的传播速度。这个传播速度需要是有限的，存在上限。狭义相对论最核心的目的就是解决超距作用的问题。而电磁现象是一个特例。在假设I中需要改变的是ID。

假设ID

相互作用的传播速度存在上限。

相互作用不是一个抽象的概念，而是需要一个载体。这个载体需要有一个物理定律来描述。由于惯性系的协变性，这个定律的方程在各惯性系中有相同的形式，因此传播速度上限在各惯性系中相同，是普适的。至于这个普适的速度是什么则是一个实验问题。之后我们会讨论实验发现这个上限就是真空中的光速 c 。这个速度的一个重要用处是使一个参考系中的时钟同步。只有时钟同步之后时间测量才有意义。相互作用的传播速度存在上限也意味着任何物体的运动速度也必须以此为上限，否则利用该物体作为载体就可以实现一个传播速度超过上限的相互作用。

由假设ID我们可以推导出两个参考系间的坐标变换关系。取直角坐标系，A与A'的坐标轴平行。在 $t = t' = 0$ 时，两坐标系原点重合。A'相对A的速度沿x轴方向，大小为 v 。这时我们只需考虑 x 和 t 两个坐标的变换关系，而 $y = y'$ ， $z = z'$ 。在此我们不准备根据空间的均匀性、各相同性等给出更严格的推导，而是不加证明地引入一些结论。我们首先认为有如下的线性关系

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \begin{pmatrix} \lambda_{11}(v) & \lambda_{14}(v) \\ \lambda_{41}(v) & \lambda_{44}(v) \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} \quad (13.1)$$

由此可以解出A'到A的变换

$$\begin{pmatrix} x \\ t \end{pmatrix} = \begin{pmatrix} \frac{\lambda_{44}(v)}{\lambda_{11}(v)\lambda_{44}(v) - \lambda_{14}(v)\lambda_{41}(v)} & \frac{-\lambda_{14}(v)}{\lambda_{11}(v)\lambda_{44}(v) - \lambda_{14}(v)\lambda_{41}(v)} \\ \frac{-\lambda_{41}(v)}{\lambda_{11}(v)\lambda_{44}(v) - \lambda_{14}(v)\lambda_{41}(v)} & \frac{\lambda_{11}(v)}{\lambda_{11}(v)\lambda_{44}(v) - \lambda_{14}(v)\lambda_{41}(v)} \end{pmatrix} \begin{pmatrix} x' \\ t' \end{pmatrix} \quad (13.2)$$

还有另一种方法得到A'到A的变换。在A'看来，A以 $-v$ 的速度运动。因此从A'到A的变换可以写为

$$\begin{pmatrix} x \\ t \end{pmatrix} = \begin{pmatrix} \lambda_{11}(-v) & \lambda_{14}(-v) \\ \lambda_{41}(-v) & \lambda_{44}(-v) \end{pmatrix} \begin{pmatrix} x' \\ t' \end{pmatrix} \quad (13.3)$$

为了求出 $\lambda(-v)$ ，我们将 v 反向。这时变换的形式仍不变，但 v 、 x 和 x' 都反

号, 即

$$\begin{pmatrix} -x' \\ t' \end{pmatrix} = \begin{pmatrix} \lambda_{11}(-v) & \lambda_{14}(-v) \\ \lambda_{41}(-v) & \lambda_{44}(-v) \end{pmatrix} \begin{pmatrix} -x \\ t \end{pmatrix} \quad (13.4)$$

与13.1式比较, 得到

$$\lambda_{11}(-v) = \lambda_{11}(v), \lambda_{14}(-v) = -\lambda_{14}(v), \lambda_{41}(-v) = -\lambda_{41}(v), \lambda_{44}(-v) = \lambda_{44}(v) \quad (13.5)$$

带入到13.3就有了

$$\begin{pmatrix} x \\ t \end{pmatrix} = \begin{pmatrix} \lambda_{11}(v) & -\lambda_{14}(v) \\ -\lambda_{41}(v) & \lambda_{44}(v) \end{pmatrix} \begin{pmatrix} x' \\ t' \end{pmatrix} \quad (13.6)$$

将之与13.2比较, 得到

$$\begin{aligned} \lambda_{11} &= \frac{\lambda_{44}}{\lambda_{11}\lambda_{44} - \lambda_{14}\lambda_{41}} \\ \lambda_{14} &= \frac{\lambda_{14}}{\lambda_{11}\lambda_{44} - \lambda_{14}\lambda_{41}} \\ \lambda_{41} &= \frac{\lambda_{41}}{\lambda_{11}\lambda_{44} - \lambda_{14}\lambda_{41}} \\ \lambda_{44} &= \frac{\lambda_{11}}{\lambda_{11}\lambda_{44} - \lambda_{14}\lambda_{41}} \end{aligned} \quad (13.7)$$

化简后有

$$\lambda_{11} = \lambda_{44} \quad (13.8)$$

$$\lambda_{11}\lambda_{44} - \lambda_{14}\lambda_{41} = 1 \quad (13.9)$$

这时四个待定系数减少为两个。进一步, A'的原点相对A沿x方向以速度v运动, 即当 $x = vt$ 时 $x' = 0$ 。所以有

$$\lambda_{11}v + \lambda_{14} = 0 \quad (13.10)$$

以上三式将待定系数减少为了一个, 这使得13.1式成为了

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \begin{pmatrix} \lambda_{11} & -v\lambda_{11} \\ \frac{1-\lambda_{11}^2}{v\lambda_{11}} & \lambda_{11} \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} \quad (13.11)$$

下面定出最后一个待定系数 λ_{11} 。首先, 当 $v = 0$ 时, $(x, t) = (x', t')$, 于是 $\lambda_{11}(0) = 1$ 。之后求速度的变换。

$$u' = \frac{dx'}{dt'} = \frac{u - v}{\frac{1-\lambda_{11}^2}{v\lambda_{11}}u + 1} \quad (13.12)$$

可以证明 u' 是 u 的单调增函数:

$$\frac{du'}{du} = \frac{(v\lambda_{11})^2}{(1 - u\lambda_{11}^2 + v\lambda_{11}^2)^2} > 0 \quad (13.13)$$

将 u 取为普适的相互作用传播速度的上限 c 。如果 u' 不等于 c ，则必须小于 c 。于是 A' 中的最大速度 c 在变换到 A 后一定会大于 c 。这与 c 是普适的速度上限矛盾。因此如果 $u = c$ ，则 $u' = c$ 。这样如果实验中能够发现某个载体的速度在不同惯性系中保持不变，这个载体就是一个可以实现最快传播相互作用的候选者。令

$$c = \frac{c - v}{\frac{1 - \lambda_{11}^2}{v \lambda_{11}^2} c + 1} \quad (13.14)$$

这样

$$\lambda_{11}^2 = \frac{c^2}{c^2 - v^2} \quad (13.15)$$

由于 $\lambda_{11}(0) = 1$ ，我们最后得到

$$\lambda_{11} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (13.16)$$

由此可见，任何速度 v 不可以大于 c ，否则会导致 λ_{11} 为虚数这样一个非物理的结果。这样就自动保证了 c 是速度上限。

为了方便，定义 $\beta = v/c$ ， $\gamma = 1/\sqrt{1 - \frac{v^2}{c^2}} = 1/\sqrt{1 - \beta^2}$ 。洛伦兹变换表示为

$$x' = \gamma(x - vt) \quad (13.17)$$

$$y' = y \quad (13.18)$$

$$z' = z \quad (13.19)$$

$$t' = \gamma(t - \beta x/c) \quad (13.20)$$

实验中为了发现相互作用传播速度的上限可以去寻找某个可以承载传递相互作用功能的载体，其速度不依赖于参考系，是一个普适的常数。这样一个载体就是一个可能的选择。如果将这样的速度指定为传播速度的上限，则惯性系间需满足洛伦兹变换。实验发现电磁场具有如此的特征。真空中电磁场以光速 c 传播，不依赖于参考系。描述电磁场的方程需要具有洛伦兹协变性。这就是麦克斯韦方程组所满足的。麦克斯韦方程组导出的电磁波在不同惯性系中都有相同的速度。

洛伦兹变换清楚地表明同时性不再具有绝对意义。一个参考系中同时发生的两个事件，在另一个参考系看来可能并不同时。与之相反，绝对时空意味着判断事件的关系有一个统一的、普适的标准。绝对时间的意义是指用唯一的一套时钟就可以测量任意参考系的时间，每一个时刻都可以有一个固定的值而不论该参考系相对时钟有怎样的运动。因此，在任何参考系中测量两个事件的间隔都是相同的，在一个参考系中同时的两个事件在任意参考系中也都是同时的。绝对空间意味着对空间每一点都可以给一个固定的坐标，这个固定的坐标不随参考系而改变。当然一个参考系可以有自己的坐标架，但这种局域的坐标与绝对坐标的差别仅仅是原点。如果坐标轴彼此都平行，则两者的差别为原点的绝对坐标。在相对论中，每个参考系中时间和空间的度量必须由固定在该参考系中的时钟和尺来完成。其中时钟通过光来同步。世界上没有一个统一的、普适的绝对时间和空间。存在绝对时空意味着测量可以

独立于参考系进行，而相对意味着测量依赖于参考系。时间的测量需要有与参考系固定在一起的一套同步的时钟来完成，空间的测量需要有与参考系固定在一起的尺子，如图13.1。观测者的作用是：判断、读数和记录。对于一个观测者，他有与其固化在一起的这套时钟，他观察到的一个事件发生的时间就定义为该事件发生处他的时钟的读数。我们发现的规律都是以这样的时间测量为依据的。如果时间和空间只是用来表明事件之间关系，那么这种关系就不是抽象的，而是应该依赖于观测者（参考系）。每个观测者应该有他自己的工具。没有先验的理由要求不同的观测者（参考系）能看到相同的关系，不同参考系可以有不同关系。每个人的立场不同，结论当然会不一样。

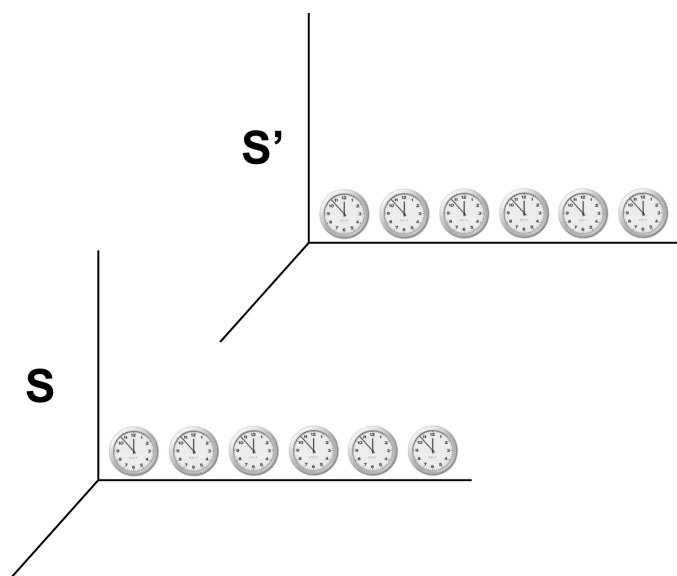


图 13.1: 固定于参考系的时钟和坐标架用于测量时间和空间

对于一个观测者（参考系），他有与其固化在一起的一套时钟，他观察到的一个事件发生的时间就定义为位于该事件发生处的这个观测者的时钟的瞬时读数。换句话说，一个事件发生时，可以有无穷多个时钟处于该事件的位置，每个时钟有不同的速度和读数，而某一个时钟的读数就正好是在它所静止的那个参考系中这个事件发生的时间。

在同一个参考系内关于时间的测量没有任何疑问。一个参考系内有一系列时钟，它们已经互相同步。每个事件发生的时间必须用当地的时钟测量。再考虑两个参考系的问题，如图13.1。S'以速度 v 沿 x 方向运动。当两参考系的原点在空间上重合时，位于两参考系原点的时钟的读数均为0。假设有两个事件1和2，在S系中它们发生在同一时刻 $t_1 = t_2 = t$ ， $\Delta t = t_2 - t_1 = 0$ 。位置分别为 x_1 和 x_2 ， $\Delta x = x_2 - x_1$ 。按照洛伦兹变换，它们发生在S'系中的位置和时间分别为

$$x'_1 = \gamma(x_1 - vt) \quad (13.21)$$

$$x'_2 = \gamma(x_2 - vt) \quad (13.22)$$

$$t'_1 = \gamma(t_1 - \beta x_1/c) \quad (13.23)$$

$$t'_2 = \gamma(t_2 - \beta x_2/c) \quad (13.24)$$

于是有 $\Delta x' = \gamma \Delta x$ 和 $\Delta t' = \gamma(\Delta t - \beta \Delta x/c) = -\gamma \beta \Delta x/c$ 。这样在 S' 系两个事件不再同时，同时不再具有绝对的意义。一个特殊情况是，如果 S 中两个事件的空间位置也相同，即 $\Delta x = 0$ ， $\Delta t = 0$ ，则它们在 S' 中也会有 $\Delta x' = 0$ ， $\Delta t' = 0$ 。

由于同时的相对性，在一个参考系中同步的时钟，在另一个参考系看来，这些时钟一般不再同步。考虑这样一个问题，在参考系 S 中的 O 时刻（每个时钟的读数都是 0 ）观察 S' 参考系的时钟。两参考系的原点在空间上重合时，位于两参考系原点的时钟的读数均为 0 。在绝对时空的意义下，在 S 系看 S' 的每个时钟都会有相同的值，且与 S 相同。但相对论会告诉我们，实际上在 S 中看 S' 时， S' 每个时钟的读数都不同。考虑瞬时位于 S 系中 x 点的一个属于 S' 的时钟。这里要研究的事件是该时钟与坐标 x 重合。这个事件发生在 S 系中的时间为 0 。而该时钟的读数就是这个事件发生在 S' 中的时间，即 $-\gamma \beta x/c$ 。所以在 S 中看 S' 的时钟时，这些时钟在正 x 方向都依次落后。

再看另一个问题，一个速度为 v 的运动的时钟在 $t = 0$ 时位于 S 系的原点，即 $x = 0$ ，求在 t 时刻时钟的读数。与上面的讨论类似，我们可以得到读数应为 $\gamma(t - \gamma \beta x/c)$ 。

洛伦兹变换会导致与伽利略变换不同的速度变换

$$dx' = \gamma(dx - vdt) \quad (13.25)$$

$$dy' = dy \quad (13.26)$$

$$dz' = dz \quad (13.27)$$

$$dt' = \gamma(dt - \beta dx/c) \quad (13.28)$$

所以

$$u'_x = \frac{u_x - v}{1 - vu_x/c^2} \quad (13.29)$$

$$u'_y = \frac{u_y}{\gamma(1 - vu_x/c^2)} \quad (13.30)$$

$$u'_z = \frac{u_z}{\gamma(1 - vu_x/c^2)} \quad (13.31)$$

光速在变换下不变。

13.2 假设二

牛顿力学框架下的假设二不具有洛伦兹协变性。因而在相对论中我们需要一个新的假设 II。实验表明：

假设 II

在惯性系中，总有可能将每一个粒子赋予一个与参考系无关的常数 $\mu(i)$ ，并有如下性质：如果两个或多个粒子相互作用，则 $\sum_i \gamma(u(i)) \mu(i) \vec{u}(i)$ ， $\vec{u}(i)$ 是速度，在相互作用前后都有相同的值，而无论粒子间相互作用是什么。

类似于牛顿力学，我们可以由此定义惯性质量 $m(i)$ 。一个值得注意的细节是这里我们不再提“在任何时刻”而仅是“在相互作用前后“都有相同的值。其原因在于相互作用不再是瞬时的。

可以定义一个粒子的动量

$$\vec{p} = \gamma m \vec{u} \quad (13.32)$$

假设II表明动量守恒。

假设II（或动量守恒）要与洛伦兹变换相容。在S参考系中的初末动量为

$$\vec{P}^i = \sum_j \gamma(u^i(j)) m(j) \vec{u}^i(j), \quad \vec{P}^f = \sum_j \gamma(u^f(j)) m(j) \vec{u}^f(j) \quad (13.33)$$

S'参考系中的初末动量为

$$\vec{P}'^i = \sum_j \gamma(u'^i(j)) m(j) \vec{u}'^i(j), \quad \vec{P}'^f = \sum_j \gamma(u'^f(j)) m(j) \vec{u}'^f(j) \quad (13.34)$$

为简单起见，只考虑二维运动，两参考系间的相对速度 v 沿x方向。为了求动量在不同参考系间的变换形式，需要得到 $\gamma(u)$ 、 $\gamma(u')$ 和 $\gamma(v)$ 三个量之间的关系。容易证明

$$\gamma(u') = \gamma(v) \gamma(u) (1 - vu_x/c^2) \quad (13.35)$$

于是

$$\frac{1}{1 - \frac{vu_x}{c^2}} = \frac{\gamma(u) \gamma(v)}{\gamma(u')} \quad (13.36)$$

从而

$$\begin{aligned} u'_x &= \frac{\gamma(u) \gamma(v)}{\gamma(u')} (u_x - v) \\ u'_y &= \frac{\gamma(u)}{\gamma(u')} u_y \\ u'_z &= \frac{\gamma(u)}{\gamma(u')} u_z \end{aligned} \quad (13.37)$$

对于动量有

$$\begin{aligned} p'_x &= \gamma(u) \gamma(v) m (u_x - v) \\ &= \gamma(v) (p_x - \gamma(u) m v) \\ p'_y &= p_y \\ p'_z &= p_z \end{aligned} \quad (13.38)$$

S'系中动量守恒，即

$$\vec{P}'^i = \vec{P}'^f \Rightarrow \sum_j \vec{p}'^i(j) = \sum_j \vec{p}'^f(j) \quad (13.39)$$

代入动量的变换关系，在x方向有

$$\sum_j \gamma(v)(p_x^i(j) - \gamma(u^i(j))m(j)v) = \sum_j \gamma(v)(p_x^f(j) - \gamma(u^f(j))m(j)v) \quad (13.40)$$

而S系中动量守恒，即

$$\vec{P}^i = \vec{P}^f \Rightarrow \sum_j \vec{p}^i(j) = \sum_j \vec{p}^f(j) \quad (13.41)$$

从而 $\sum_j \gamma(u(j))m(j)$ 必须是守恒量。定义能量为

$$E = \gamma mc^2 \quad (13.42)$$

假设II的洛伦兹协变性要求如此定义的能量守恒。

在相对论中能量守恒，但质量不守恒。即使粒子不运动，也有能量 mc^2 ，称为静止能量。动能定义为 $K = \gamma mc^2 - mc^2$ 。在 $v \ll c$ 时， $K \sim mv^2/2$ 。当物体的静止能量发生变化时，质量也会变化。例如，将一物体内能增加 ΔE ，则质量也会增加 $\Delta E/c^2$ 。在相对论中不再区分能量的各种形式。例如，非弹性碰撞，动能损失了，并转化为内能，但总能量守恒。两粒子对撞并发生完全非弹性碰撞，动能完全转化为静止能量，碰撞后粒子的质量要大于碰撞前两粒子质量之和。能量总是在静止能量和动能间转化，但总能量守恒。只有静止能量与质量成正比。实际上质量不是量度物质的量的物理量，而是量度惯性的量，因此没有理由守恒。惯性，即质量，是粒子对外界的响应，这种响应很自然地应该依赖于系统内部的状态（由静止能量描述）。

下面我们将通过实验找到一个普适的速度作为传递相互作用的速度上限。我们还将讨论相对论的一系列预言及其实验验证。这些预言表明相对论是一个很好的理论框架，但这并不意味着没有其他的可能。

14 相互作用传播速度的上限

现代物理学认为基本的相互作用都是通过交换相应的粒子实现的。例如，光子传递电磁作用、中间玻色子传递弱相互作用、胶子传递强相互作用。存在粒子运动速度的上限也就意味着存在作用传播速度的上限。实验中没有发现运动速度超过光速的粒子。例如，实验中可以将电子加速到很高的能量，例如20 GeV。这个能量在牛顿力学中意味着大约 10^{11} m/s的速度。但实验中测到电子的速度仍然比光速低大约0.07 m/s。又例如，宇宙射线中的质子其能量可达5 J，对应牛顿力学中的速度大约是 2.4×10^{14} m/s。但是实验测到的质子速度要低于光速。1987年天文学家观测到大麦哲伦星云中的超新星爆发，距离地球 1.6×10^{21} m。爆发产生的光传播到地球的时间大约是 5.4×10^{12} s。在这个时间之前的大约 10^4 s，地球上还观测到了中微子的束流。这个时间与理论预言的超新星爆发过程中中微子先于光子产生的时间相符。这个观察结果表明，中微子在 10^{-8} 不确定度下与光子有相同速度。

所有的实验都表明相互作用传播的速度存在上限。而实验又发现这个速度的上限就是真空中的光速。这意味着光在真空中以恒定的速度传播。其实验验证大致可分为三类：光速各向同性实验、光速与光源运动无关实验、光速与频率无关实验。

14.1 光速各向同性实验

这类实验基本上都是基于干涉仪。Timmer等人用类似图14.1所示的原理测量光速的各向异性 (W. S. N. Trimmer, R. F. Baierlein, J. E. Faller, and H. A. Hill, Phys. Rev. D 8, 3321 (1973))。 v 是实验室参考系的速度。

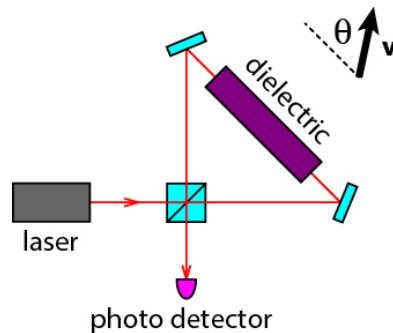


图 14.1: 三角形干涉仪

光速的各向异性表示为

$$\frac{c_0}{c(\theta)} = 1 + b_1 P_1(\cos \theta) + b_3 P_3(\cos \theta) \quad (14.1)$$

其中 P_l 是 Legendre 多项式: $P_1(x) = x$, $P_3(x) = \frac{1}{2}(5x^3 - 3x)$ 。Timmer 等人的实验表明 $b_1 = (0.1 \pm 8.4) \times 10^{-11}$ 、 $b_3 = (2.3 \pm 1.5) \times 10^{-11}$ 。实验仪器中插

入一块电介质的原因是使得顺时针和逆时针运动的光有所不同。目前的测量已经将光速各向同性的上限降低到 10^{-17} 的水平。

14.2 光速与光源运动无关实验

典型的实验包括天体光源实验、运动的反射镜实验、运动介质实验和 γ 射线源实验等。Alvager等人 (T. Alvager et al., Phys. Lett. 12 260 (1964)) 在CERN测量了高速运动的 π^0 介子衰变产生的光子的速度，如图14.2所示。

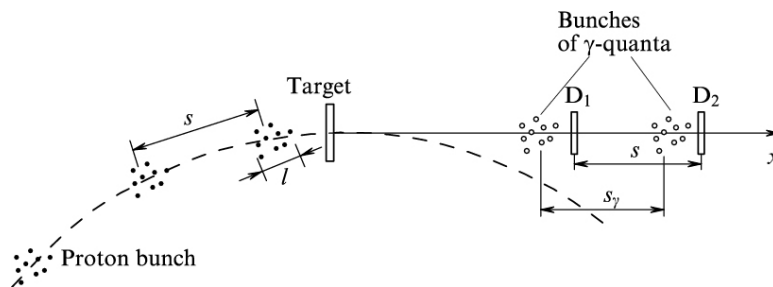


图 14.2: π^0 介子的衰变

被加速的质子束团打到Be靶上会产生各种二次粒子，其中包括中性的 π^0 介子。其他的带电粒子在磁场作用下会偏转，而 π^0 介子则会沿切向以接近光速运动。 π^0 介子的寿命极短，只有 0.83×10^{-16} s。一个 π^0 介子在产生后的很短距离内就衰变为一对 γ 光子。 γ 光子的脉冲由相距31.45米远的两个探测器 D_1 和 D_2 进行测量，由此可以计算光子的速度。实验表明以近乎光速运动的 π^0 介子其衰变产生的光子的速度为 $(2.9977 \pm 0.0004) \times 10^8$ m/s，在实验精度范围内没有发现与源的速度有关的证据。

14.3 光速与频率无关实验

光速与频率无关意味着光在真空中无色散。光在真空中相对一切惯性系的速度恒定，因此不存在光子静止的参考系，等价于光子的质量为零。实验通过多种手段测量光子的质量。相关结果的综述见A. S. Goldhaber and M. M. Nieto, Rev. Mod. Phys. 82, 939 (2010)。

Proca将Maxwell方程组推广到光子有质量的情形。Proca方程表明电磁波的群速度为 $c\sqrt{1 - \mu^2 c^2 / \omega^2}$ 。 μ 是光子的质量。 ω 是频率。不同频率的电磁波传播速度不同。在 $\omega \rightarrow \infty$ 的极限情况下，群速度趋于常量 c ，可以当作相互作用传播速度的上限。 μc 是电磁波的下截止频率，只有 $\omega > \mu c$ 时电磁波才能无衰减地传播出去。实验测到的电磁波的最低频率可达8 Hz左右 (Kroll, N. M., Phys. Rev. Lett. 27, 340 (1971))。这给出了光子质量的上限为 4×10^{-49} kg。

Proca方程给出的静电场的形式为

$$V(r) = \frac{Q}{4\pi\epsilon_0 r} \exp(-\mu r) \quad (14.2)$$

光子质量为零时得到库伦定律。迄今对库伦定律最精确的检验由Williams等人完成 (E. R. Williams, J. E. Faller, and H. A. Hill, Phys. Rev. Lett. 26, 721 (1971).)。基本的实验原理是使用相互连接的内外两个金属球壳, 在给外球壳充电后断开连接, 测量内球壳上的带电。严格的库伦定律要求内球壳上的电荷为零。实验给出的光子质量的上限为 2×10^{-50} kg。

15 长度收缩与时间膨胀

一静止时长度为 l_0 的物体，在S系中以速度 v 运动，我们需要测量该物体的长度 l 。测量长度意味着需要同时记录下物体两端点的坐标 x_1 和 x_2 ， $l = x_2 - x_1$ 。由洛伦兹变换，

$$x'_1 = \gamma(x_1 - vt) \quad (15.1)$$

$$x'_2 = \gamma(x_2 - vt) \quad (15.2)$$

而物体静止在S'中，所以 $l_0 = x'_2 - x'_1$ 。于是， $l = l_0/\gamma$ ，长度收缩。

两个静止时一样长的物体A和B，以相对速度 v 运动。在A的静止系中看，B的长度小于A。但这并不意味着在B的静止系中会认为A长于B，事实上在B的静止系中看，A会短于B。B<A并不意味着A>B，原因是B<A是A的观点，只对A成立，对B不成立，即在B看来，不能从A中的B<A得到B中的A>B。

在S'中同一点 x' 先后发生的两个事件，时间分别为 t'_1 和 t'_2 ， $\tau = t'_2 - t'_1$ ，称为原时。在S系，两个事件发生的时间分别为

$$t_1 = \gamma(t'_1 + \beta x'/c) \quad (15.3)$$

$$t_2 = \gamma(t'_2 + \beta x'/c) \quad (15.4)$$

于是 $\Delta t = t_2 - t_1 = \gamma\tau$ ，时间膨胀。在这个效应中我们比较的是单一时钟和两个不同时钟的读数。

对于任何两个事件，时间和空间间隔分别为 Δt 和 Δx ，可以定义

$$(c\tau)^2 = (c\Delta t)^2 - (\Delta x)^2 \quad (15.5)$$

$c\tau$ 称为时空间隔。有意思的是 $c\tau$ 在洛伦兹变换下不变。

如果 $(c\tau)^2 > 0$ ，则称两事件的间隔为类时的，这时总可以找到一个参考系使得这两个事件发生在同一地点，而先后时间不同，这时 $\Delta x = 0$ ， $\Delta t = \tau$ 。如果两个事件发生在同一个物体上，则这两个事件一定时类时的，因为物体的速度不能超过 c 。

如果 $(c\tau)^2 < 0$ ，则称两事件的间隔为类空的，这时总可以找到一个参考系使得这两个事件发生在同一时间，而地点不同。

事件间隔的类时和类空的划分是绝对的，因为时空间隔 $c\tau$ 不依赖于参考系。考虑一个发生在一个参考系的时间和空间原点的事件O，如图15.1。图中的虚线为另一个参考系的坐标轴。这个时空图上的每一点称为那个事件的世界点。一个事件的演化在时空图上表现为世界点的一个轨迹，称为世界线。图15.1中两条斜的世界线表示在 $t = 0$ 时刻位于 $x = 0$ 的匀速运动的粒子，速度分别为 $+v$ 和 $-v$ 。这两条世界线的斜率的绝对值一定不小于1，因为 $v \leq c$ 。现取 $v = c$ ，时空图被两条世界线划分为四个区域。在I区，相对于事件O，每个点代表的事件都是类时的，并且无论在哪个参考系每个点的事件发生的

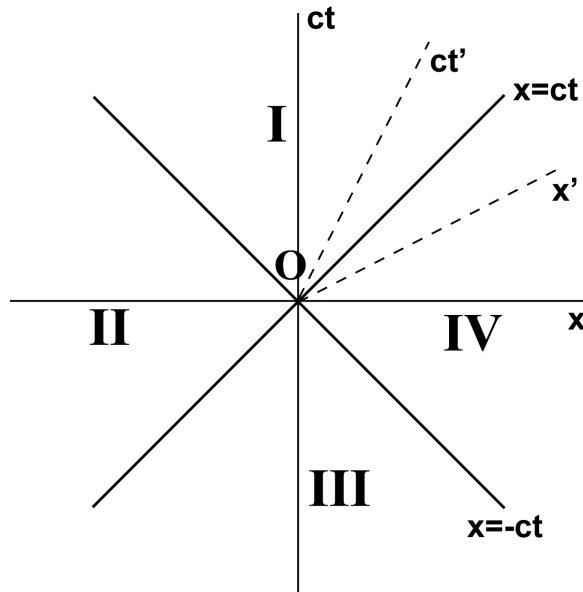


图 15.1: 类时与类空事件

时间都晚于 O 。因此称I区为“未来”。类似地，称III区为“过去”。在II和IV两个区，每个事件都是类空的。对每个事件都存在参考系使其早于、晚于 O ，或与 O 同时。这两个区称为“别处”。这两个区中的事件与 O 都不能有因果关系，否则因果关系就失去了意义。

有关时间膨胀和长度收缩的预言都可以在加速器实验上得到验证。如图15.2是CERN的实验装置的示意图。

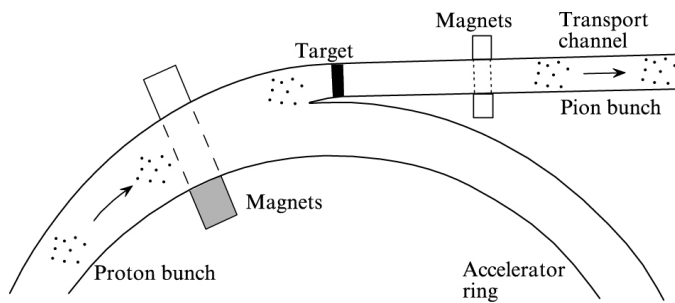


图 15.2: 加速器实验示意图

每个质子的束团包含 10^8 个质子。速度为 $v = 0.999999991c$ 。一个束团的长度为3到5厘米，横向尺寸为16微米。相邻两个束团的平均距离是90米。质子打到靶上后会产生多种二次粒子的束团。不同的粒子可以用磁场分开。考虑带正电的 π 介子。在静止参考系中， π 介子的寿命 $\tau = 2.56 \times 10^{-8}$ s。在 t' 时刻，粒子的数目为

$$N(t') = N(0) \exp(-t'/\tau) \tag{15.6}$$

于是可以求出在静止参考系中，当粒子数目由 $N(0)$ 衰减为 $N(t')$ 是，经历的

时间是

$$t' = \tau \frac{N(0)}{N(t')} \quad (15.7)$$

在实验中， π 介子的速度非常接近光速， $\gamma = 857$ 。在实验室参考系中，飞行长度为 $L = 100 \text{ m}$ ，飞行时间为 $t = L/c = 3.33 \times 10^{-7} \text{ s}$ 。在transport channel的起点实验测得 $N(0) = 1.5 \times 10^7$ 。 π 介子衰变为 μ 子和中微子。通过测量 μ 子的数量 N_μ 就可以得到在终点位置的 π 介子数。实验测得 $N_\mu = 2.27 \times 10^5$ ，于是 $N(t') = N(0) - N_\mu = 1.477 \times 10^7$ 。这样可以得到 $t' = 3.90 \times 10^{-10} \text{ s}$ 。这样 $t/t' = 854$ ，与理论预言的 $\gamma = 857$ 的误差是0.35%。

多普勒效应也可以给出时间膨胀效应的实验验证。光源相对观测者以速度 v 运动，如图15.3所示。

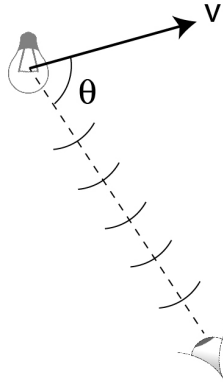


图 15.3: 多普勒效应

相对论给出

$$\nu = \nu_0 \frac{\sqrt{1 - v^2/c^2}}{1 - \frac{v}{c} \cos \theta} \quad (15.8)$$

其中 ν_0 和 ν 分别是光源静止和运动的参考系中测到的电磁波的频率。相对论效应仅存在于分子上，来源于时间膨胀，是 v/c 的二阶效应。

在验证相对论多普勒效应的一个实验中，电离的氢分子离子被加速到大约 10^6 m/s 的速度，然后与静止的氢分子碰撞产生处于激发态的快速运动氢原子。运动氢原子的光谱线由于多普勒效应而发生蓝移或红移。实验采用图15.4的构型。直接进入光谱仪的光线发生蓝移，波长为

$$\lambda_B \approx \lambda_0(1 - \beta \cos \theta + \beta^2/2) \quad (15.9)$$

经过反射镜进入光谱仪的光线发生红移，波长为

$$\lambda_B \approx \lambda_0(1 + \beta \cos \theta + \beta^2/2) \quad (15.10)$$

实验中使用的是氢原子的 H_α 线，波长 $\lambda_0 = 6562.793 \text{ \AA}$ 。二阶效应出现在 $\lambda_0\beta^2/2$ 这样一项中。实验观测结果是 $(0.498 \pm 0.025)\lambda_0\beta^2$ ，见H. I. Mandelberg and L. Witten, J. Opt. Soc. Am. 52, 529 (1962)。

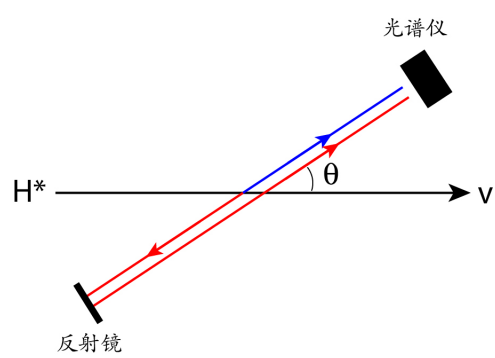


图 15.4: 多普勒效应实验示意图

16 相对论力学

为简单起见, 只考虑二维运动。在S系中,

$$E = \gamma(u)mc^2 \quad (16.1)$$

$$p_x = \gamma(u)mu_x \quad (16.2)$$

$$p_y = \gamma(u)mu_y \quad (16.3)$$

S'系中,

$$E' = \gamma(u')mc^2 \quad (16.4)$$

$$p'_x = \gamma(u')mu'_x \quad (16.5)$$

$$p'_y = \gamma(u')mu'_y \quad (16.6)$$

其中

$$u'_x = \frac{u_x - v}{1 - vu_x/c^2} \quad (16.7)$$

$$u'_y = \frac{u_y}{\gamma(v)(1 - vu_x/c^2)} \quad (16.8)$$

由 $\gamma(u)$ 、 $\gamma(u')$ 和 $\gamma(v)$ 三个量之间的关系

$$\gamma(u') = \gamma(v)\gamma(u)(1 - vu_x/c^2) \quad (16.9)$$

有

$$E' = \gamma(v)(\gamma(u)mc^2 - v\gamma(u)mu_x) \quad (16.10)$$

$$= \gamma(v)(E - vp_x) \quad (16.11)$$

类似地,

$$p'_x = \gamma(v)(p_x - vE/c^2) \quad (16.12)$$

$$p'_y = p_y \quad (16.13)$$

有趣的是 (E, p) 与 (t, x) 有完全相同的变换关系。另外, 能量和动量间满足

$$E^2 - c^2p^2 = (mc^2)^2 \quad (16.14)$$

因此 $E^2 - c^2p^2$ 是相对论不变量。这类似于 $(c\Delta t)^2 - (\Delta x)^2$ 。

光子的质量为零, 因而 $E = pc$ 。能量与频率的关系是 $E = h\nu = \hbar\omega$, 其中 h 是普朗克常数, $\hbar = h/2\pi$ 。而动量 $\vec{p} = \hbar\vec{k}$, 其中 $k = 2\pi\nu/c = 2\pi/\lambda$, λ 是波长。

在牛顿力学中, 为了描述相互作用引入了力。相对论中不再有超距作用。这时需要用场来描述相互作用。我们也可以引入力的概念, 但这时力仅仅是将运动方程的右边耦合常数与场的乘积换一种写法。

四种相互作用的都需要用场来描述。描述引力需要广义相对论，强和弱相互作用也超出本课程范围。所以我们只研究一下电磁相互作用。实验发现，电磁现象由电场 \vec{E} 和磁场 \vec{B} 描述，场满足协变的麦克斯韦方程组，运动方程由下式给出

$$\frac{d\vec{p}}{dt} = q(\vec{E} + \vec{v} \times \vec{B}) \quad (16.15)$$

在不同参考系， \vec{p} 不同， \vec{E} 和 \vec{B} 也不同，但会保证运动方程的协变性。

可以用以下方式定义力

$$\vec{F} = \frac{d\vec{p}}{dt} \quad (16.16)$$

“作用力等于反作用力”在相对论中没有地位。作用力和反作用力需要同时作用在两个位于空间不同位置的粒子，但同时性在相对论中不再有绝对意义。两个粒子之间有相互作用，但我们只能一个一个地描述作用在某个粒子上的力。一般地，作用力和反作用力不再大小相等方向相反。粒子的动量必须加上场的动量才守恒。

相对论的质能关系是普适的。让我们考虑光子与原子的相互作用。一个处于基态的原子，其质量为 m 。当原子处于激发态时，它的质量会有所增加。记 $\hbar\omega_0$ 为原子中电子处于激发态与基态之间的能量差。激发态原子的质量为 $m + \hbar\omega_0/c^2$ 。初态时原子的速度为 \vec{v}_i ，动量 $\vec{p}_i = m\vec{v}_i$ 。光子的初始动量和能量是 $\hbar\vec{k}$ 和 $\hbar\omega$ 。对于光的吸收过程，动量守恒给出

$$\vec{p}_i + \hbar\vec{k} = \vec{p}_f \quad (16.17)$$

能量守恒要求

$$\sqrt{p_i^2 c^2 + m^2 c^4} + \hbar\omega = \sqrt{p_f^2 c^2 + (mc^2 + \hbar\omega_0)^2} \quad (16.18)$$

将上式两边平方得到

$$\hbar\omega \sqrt{p_i^2 c^2 + m^2 c^4} = mc^2 \hbar\omega_0 + \hbar c^2 \vec{p}_i \cdot \vec{k} + \frac{1}{2} \hbar^2 \omega_0^2 \quad (16.19)$$

于是

$$\omega = \left(mc^2 \omega_0 + c^2 \vec{p}_i \cdot \vec{k} + \frac{1}{2} \hbar \omega_0^2 \right) \frac{1}{\sqrt{p_i^2 c^2 + m^2 c^4}} \quad (16.20)$$

对于 $v_i \ll c$ 的情形

$$\frac{1}{\sqrt{p_i^2 c^2 + m^2 c^4}} \approx \frac{1}{mc^2} \left(1 - \frac{v_i^2}{2c^2} \right) \quad (16.21)$$

得到

$$\begin{aligned} \omega &= \left(\omega_0 + \vec{v}_i \cdot \vec{k} + \frac{\hbar \omega_0^2}{2mc^2} \right) \left(1 - \frac{v_i^2}{2c^2} \right) \\ &= \omega_0 + \vec{v}_i \cdot \vec{k} - \frac{\omega_0 v_i^2}{2c^2} + \frac{\hbar \omega_0^2}{2mc^2} + \dots \end{aligned} \quad (16.22)$$

最后一步略去了更高阶的项，其中用到了 $k = \omega/c$ 。

上式的第二项和第三项都来源于原子运动所引起的多普勒效应。最后一项表示动量守恒导致的原子反冲，是相对论效应。利用反冲可以实现原子冷却。

反冲效应可以利用精密光谱手段进行测量。实验中测量的是反冲引起的光子吸收与发射之间的频率差，大小为

$$\frac{\Delta\nu}{\nu} = \frac{\hbar\omega_0}{mc^2} \quad (16.23)$$

例如，对于 ^{40}Ca 原子， $m = 39.96u$ ， $1u = 1.66 \times 10^{-27} \text{ kg}$ 。 $^1S_0 \rightarrow ^3P_1$ 跃迁谱线的波长是657 nm。计算得到 $\Delta\nu = 23.1 \text{ kHz}$ 。实验结果见图16.1。

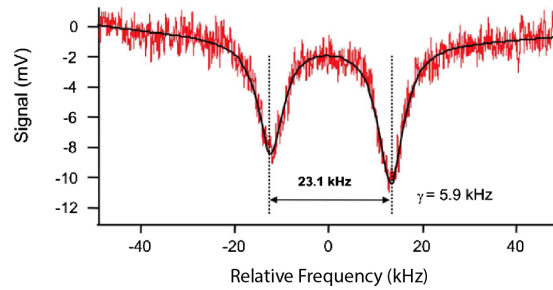


图 16.1: ^{40}Ca 原子高分辨光谱

17 四矢量和相对论协变性

牛顿力学中的时间和空间相互独立，动量与能量代表的是不同的属性。但相对论有着不同的观点。为此我们需要一个新的数学结构。在此我们将用更一般的代数语言来叙述狭义相对论。数学给出普遍的结构，物理是一种实现。我们关心的第一件事是时空结构。时空构成了一个线性空间。线性空间是一个集合（与坐标架无关）。线性空间中需要定义加法和与纯数的乘法。线性空间这个集合中的每一个元素称为矢量。一个方便的研究线性空间的办法是找出一组元素（矢量）组成基。基也是该空间中的矢量。这样任一矢量可对这组基展开，系数称为坐标。

引入一组基 (η_1, \dots, η_m) 和另一组基 (ξ_1, \dots, ξ_n) 。两组基间的关系是 $(\xi_1, \dots, \xi_n)L = (\eta_1, \dots, \eta_m)$ 。L称为过渡矩阵。在两组基下一矢量的坐标分别为 $(x_1, \dots, x_n)^T$ 和 $(x'_1, \dots, x'_n)^T$ ，二者间的关系是

$$\begin{pmatrix} x'_1 \\ x'_2 \\ \dots \\ x'_n \end{pmatrix} = L \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} \quad (17.1)$$

简记为 $x' = Lx$ 。于是引入基就可以用矩阵来表示线性空间。

内积为线性空间引入更多的结构，用集合中两元素的双线性函数 $\langle x, y \rangle$ 给出，具有以下性质：

1. $\langle x, y \rangle = \langle y, x \rangle$
2. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
3. $\langle cx, y \rangle = c\langle x, y \rangle$, $c \in \mathbf{R}$
4. $\langle x, x \rangle \geq 0$, 当且仅当 $x = 0$ 取等号

定义了内积的有限维实线性空间称为欧几里德空间。

引入基 (η_1, \dots, η_m) 后，内积可表示为

$$\langle x, y \rangle = (x_1 \ x_2 \ \dots \ x_n) \begin{pmatrix} g_{11} & g_{12} & \dots & g_{1n} \\ g_{21} & g_{22} & \dots & g_{2n} \\ \dots & \dots & \dots & \dots \\ g_{n1} & g_{n2} & \dots & g_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} \quad (17.2)$$

简记为 $\langle x, y \rangle = x^T G y$ 。G为实对称正定矩阵，称为度规。

$$G = \begin{pmatrix} \langle \eta_1, \eta_1 \rangle & \langle \eta_1, \eta_2 \rangle & \dots & \langle \eta_1, \eta_n \rangle \\ \langle \eta_2, \eta_1 \rangle & \langle \eta_2, \eta_2 \rangle & \dots & \langle \eta_2, \eta_n \rangle \\ \dots & \dots & \dots & \dots \\ \langle \eta_n, \eta_1 \rangle & \langle \eta_n, \eta_2 \rangle & \dots & \langle \eta_n, \eta_n \rangle \end{pmatrix} \quad (17.3)$$

度规矩阵在两组基下的关系是 $G' = L^T G L$ 。内积不依赖于基，但度规矩阵依赖于基。

一个参考系相当于取了一组基。参考系间的变换不是线性空间中的线性变换，而是一个过渡矩阵。不同参考系在物理上是等价的。等价性即规律形式上的不变。因此不同参考系下度规矩阵应相同。

我们前面曾经发现在洛伦兹变换下 $(c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2$ 不变。这预示着我们可以对时空 (ct, x, y, z) 引入以下度规

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (17.4)$$

但是这种度规虽然满足度规的前三个要求，但不是正定的。这样定义的内积称为赝内积，定义了这种赝内积的空间称为闵可夫斯基空间。

在S系中取坐标为 $\eta_1 = (1, 0, 0, 0)^T$, $\eta_2 = (0, 1, 0, 0)^T$, $\eta_3 = (0, 0, 1, 0)^T$, $\eta_4 = (0, 0, 0, 1)^T$ 的四个事件为基。四维时空是事件作为元素组成的线性空间。取S'系的基为该参考系中坐标为 $\xi_1 = (1, 0, 0, 0)^T$, $\xi_2 = (0, 1, 0, 0)^T$, $\xi_3 = (0, 0, 1, 0)^T$, $\xi_4 = (0, 0, 0, 1)^T$ 的四个事件。根据洛伦兹变换， η_1 事件在S'中的坐标为 $(\gamma, -\gamma\beta, 0, 0)$, η_2 的坐标为 $(-\gamma\beta, \gamma, 0, 0)$, η_3 的坐标为 $(0, 0, 1, 0)$, η_4 的坐标为 $(0, 0, 0, 1)$ 。于是

$$(\xi_1 \ \xi_2 \ \xi_3 \ \xi_4) = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} (\eta_1 \ \eta_2 \ \eta_3 \ \eta_4) \quad (17.5)$$

令

$$L = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (17.6)$$

两参考系中同一事件的坐标间满足 $x' = Lx$ 。

过渡矩阵 L 满足 $G = L^T G L$ ，物理规律在洛伦兹变换下形式不变，称为相对论协变性。

我们还曾发现，在洛伦兹变换下 $E^2 - c^2 p^2$ 不变， (E, cp) 有相同的变换关系，也构成了一个四矢量。我们还可以定义标量，即洛伦兹变换下不变的量，如质量。