理论力学



北京大学物理学院

VERSION 1 ∘ 4

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

开篇说明

理论力学，特别是分析力学一向是我认为“最为优美”的理论物理课程。因此，我一直希望能够有机会欣赏一下。正是基于这种诱惑，我才愿意在 2005 年春尝试讲授《理论力学》课程。

但准备这个课程的过程中发现，诱惑有时是很害人的。具体的备课过程告诉我，这个课程实际上相当具有挑战性。挑战性之一是分析力学的许多原理、方法和术语实际上是为整个理论物理起到了奠基的作用（例如：最小作用量原理、哈密顿量等等）。分析力学中很多地方都蕴含着与理论物理其他分支的紧密联系。例如，最小作用量原理一直到量子场论中都是重要的原理；泊松括号对于经典力学的描述与对易括号对于量子力学的描述是十分类似的；哈密顿－雅克比方程与光学、波动力学（ Schrödinger 方程）的内在联系；还有经典的连续场论（流体力学、弦与膜的振动）等等。挑战之二是经典动力学本身也有十分迅猛的发展，特别是在非线性、混沌等现象的研究中。这些内容本身就是相关领域的科研的前沿。挑战之三在于经典力学中的许多理论方法还与现代数学物理有着十分紧密的联系。例如，在分析力学中可以很直观地引入流形、tangent bundle、differential forms、symplectic geometry 等等。这些数学概念在现代的数学物理中起着重要的作用。以这样的一个观念来看待理论力学，我就发现它变为一个很庞大的体系。

在我看来理论力学的课程应当起到两方面的作用：第一，使同学们掌握理论力学－特别是分析力学－的基本原理和方法，并能够运用它们解决经典力学的具体问题；第二，将分析力学的原理和方法作为一个窗口和桥梁，使之能够联结和透视到以后其他课程的学习和研究中。这第一方面的作用仅仅局限在本课程，固然是重要的；但我觉得第二方面的作用似乎更加重要。为此，我的课程讲义采取了以下两个步骤。第一，我努力强调了从超出经典力学本身的最基本的原理出发来阐述分析力学的内容。为此，我们讨论分析力学的起点是狭义相对论的时空观和最小作用量原理。常见的一种讲述方法是从牛顿力学出发来讨论拉格朗日形式的分析力学。这样的方法固然符合人类的认识规律，但缺点是没有突出基本原理的崇高地位。我采用的讲述方法基本上与 Landau 书 [1] 中的逻辑类似，只不过我们是从相对论性的时空观出发的，而朗道是从非相对论时空观出发的。第二，在这个讲义中我尽量在一些可能的地方，将所学的分析力学的内容进行引申，点出它们与其他课程的联系。希望能够给大家一片更加广阔的天地。我希望多数同学在学过这个课程一段时间后，比如说一年以后，能够留下一些重要的印象：你可以忘了怎么解一个具体的力学问题（就像我现在一样），但希望你能够记住分析力学的基本思想。

¶如上所述，由于与分析力学有关联的物理分支实在是太过广泛，显然要将所有的内容都一一尽述是不可能的，同时也无法与一个 3 学分的课程相容。因此，我必须对内

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

容有所取舍。在制备这个讲义的过程中，我主要参考了三本书：Landau 的 Mechanics [1]、Goldstein 的 Classical Mechanics [3] 和 José 的 Classical Dynamics [4]。本讲义的基本框架与 Landau 的书 [1] 类似。这样的选择有两个原因，一是 Landau 的书是经典著作。它从基本原理出发的讲述方式与我们课程的初衷比较吻合；二是 Landau 的书比较薄（167 页），这样我们可以基本上在一个学期讲完。当然，我也舍去了其中的一些内容，同时又补上了一些其他的内容。这些内容包括：从狭义相对论出发的讨论；非线性和混沌的讨论；连续介质力学的讨论等等。这些内容我们都只是给了简要的介绍。同学们有兴趣可以从 Goldstein 的书 [3] 和 José & Saletan 的书 [4] 中找到更为详尽和丰富的内容，尽管它们出现的顺序与我的讲义中不尽相同。Goldstein 的书也是经典力学的一部经典著作。但是所包含的内容太多（600 多页），很难在一个学期消化。José 的书是一部用现代观点描述经典力学的好书。它介绍了大量的非线性力学中的现象和理论（例如：混沌、KAM 定理等等）。同时，书中还讨论了大量与经典力学密切相关的数学概念的描述。有兴趣的同学可以读读。它的问题也是内容太多（也是 600 多页），同时恐怕并不是每一个同学都对相关的数学如此热衷。总之，上述两部书很难用来做真正的教材，只适合作为参考书。

我并没有忽视中文的教材。但是，在翻阅了所有我能够找到的中文理论力学教材之后发现，它们绝大多数的内容并不与本课程的初衷相吻合。我们这个课程的主旨是讲述分析力学的主要思想方法。绝大多数国内的理论力学教材都不是以分析力学为主体的。其原因在于理论力学课程的主要受众除了我们物理学专业的学生外，还有众多工程专业的学生。后者可能在人数上是绝大多数。因此，绝大多数国内的理论力学教材讲述的主要内容是矢量力学（牛顿力学）。

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

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

¶在这个十分粗略的讲义出炉之际，我要感谢北大理论物理所的陈晓林、张大新、李克平等几位教授无私地与我分享他们多年来教授理论力学的经验。这些经验对于我这个理论力学课程的“新手”来说是十分宝贵的。特别要感谢的是陈晓林教授给了我他们授课的全部 Powerpoint 文件。这些课件的准备无疑需要大量的时间投入和精心的准备。陈晓林还给了我理论力学课程的习题。尽管这些题目我很多都已经不会做了，但仍然使我受益颇多。我也感谢郑汉青、朱世琳、马中水等其他教授对我的一贯支持。最后，感谢我的家人（特别是我的妻子韦丹和儿子咪咪小同学）对我的鼓励和支持。

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

刘川，二零零四年十二月

¶第一次的版本是 Postscript 版本的。考虑到这类文件已经逐步淡出江湖，我对部分格式做了调整（包括部分的字体以及讲义的封面）并产生 PDF 版本讲义。目前版本号定为：1.0ϵ。

刘川，二零零九年九月

¶2013 年，我被调整为理论力学课程主持人。因此，讲述理论力学课程是迟早的事情了。这距离上一次讲述理论力学已经有近 10 年了，时间过得真是太快了。本次的改动还是蛮大的，主体结构没动的情况下几乎各章都有增删。例如，第三章中关于水外行星影响水星进动的讨论；第六章中关于 Liouville 定理的讨论等等。另外，原先讲义的最后一章是关于连续介质和场的经典力学。我发现无论如何都没有时间去涉及，因此干脆删掉了。同时，由于旧版 LaTeX 版本也逐渐不再使用，本讲义因此改用 XeLaTeX 排版。目前暂定版本号为：1.1。

刘川，二零一四年一月

¶2014 年春，我讲授了理论力学课程。根据具体讲述的过程中增加了一些内容：主要包括关于刚体的运动学的部分。这部分的增加是原先没有预期的。例如我原以为关于三维转动矩阵的知识多数同学在线性代数课程中已经涉及（事实上这个假设不正确）；我原以为同样是线性代数课程中多数同学已经有一些群论的知识，实际上并非如此。因此，我增加了关于群论基本知识的一个附录，供有兴趣的同学参考。另外在哈密顿-雅各比方程的部分增加了相当一部分例子。这些内容的增加使得讲义的页数陡增，对此我只能深表遗憾了。目前的版本号为 1.2。

刘川，二零一四年八月

¶2018 年秋，我再次讲授理论力学 B 课程。主要增加的内容有受限三体问题和拉格朗日点的讨论，见第 17节。这是一个相当古老，但近年来又变得很热门的课题。此外，公式的格式上也进行了更新。目前的版本号为 1.4。

刘川，二零一八年八月

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

目 录

第一章 楔子 **1**

1 力学的简要历史 1

2 关于分析力学与矢量力学 5

3 关于约束 8

4 虚功原理和达朗伯原理 11

5 关于有耗散的系统 13

第二章 力学体系的作用量与运动方程 **16**

6 狭义相对论的时空观 16

7 最小作用量原理 20

8 相对论性自由粒子的作用量 25

9 粒子与外场的相互作用 27

10 非相对论极限 30

11 对称性 32

11.1 时间平移 33

11.2 空间平移 33

11.3 空间转动 34

11.4 尺度变换 35

11.5 分立对称性 36

12 非惯性系的力学 38

13 非完整约束系统的最小作用量原理 39

第三章 有心力场中粒子的运动 **42**

14 中心力场 42

15 开普勒问题 44

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

16 潮汐现象 . . . 49
17 受限三体问题与拉格朗日点 . . . 51
17.1 拉格朗日点 . . . 52
17.2 拉格朗日点附近有效势能 . . . 57
17.3 拉格朗日点附近的运动模式 . . . 60
18 相对论效应造成的近日点进动 . . . 64

第四章 小振动 68
19 一维振子 . . . 69
20 多自由度系统的小振动与简正模式 . . . 70
21 非谐效应 . . . 78
22 参数共振 . . . 80

第五章 刚体的运动 85
23 刚体的运动学 . . . 85
23.1 三维转动矩阵的一些基本性质 . . . 86
23.2 无穷小转动与角速度 . . . 89
23.3 欧拉角的描述 . . . 91
23.4 Cayley-Klein 参数 . . . 93
24 刚体的动能、角动量与惯量张量 . . . 97
25 刚体的动力学 . . . 98
26 刚体的动力学的矩阵描述 . . . 100
27 自由不对称陀螺 . . . 101
28 对称陀螺的定点运动 . . . 104

第六章 哈密顿力学 109
29 哈密顿正则方程 . . . 110
30 ξ 符号与 Liouville 定理 . . . 114
31 泊松括号 . . . 115
32 作用量作为端点的函数 . . . 117
33 正则变换 . . . 121
34 哈密顿 -雅克比方程 . . . 123
34.1 一般的讨论 . . . 123

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

34.2 分离变量法 124
34.3 哈密顿 -雅克比方程与波动力学 130
35 绝热不变量与正则变量 132
35.1 绝热不变量 132
35.2 正则变量 134
35.3 完全可分离系统的正则变量 135
36 决定性中的混沌 140

附 录 A 对称性与群 143

1 对称性与群的定义 143
2 两个群之间的那些事儿 145
3 群的例子 148
4 群的表示 149

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

第一章　　楔子

本章提要

- 力学的简要历史 (1)
- 分析力学与矢量力学 (2)
- 约束 (3)
- 虚功原理和达朗伯原理 (4)
- 有耗散的力学系统 (5)

这一章是最无聊同时也是最有趣的一章。它包含了关于力学的鳞鳞总总各种无法合适地放在以后各章中的一些内容。有的是一些历史的回顾；有的是一些细节的讨论；有的是一些个人的体会；有的是一些不知道什么的东西。总之它们都与这门课或多或少有些联系，但是这些联系又不足以使它们成为这个课程的正式的组成部分。于是，它们被汇总到了这里。不知道叫它们什么合适，姑且称之为楔子或者八卦吧。总之，这一章是所有非正式内容的集中营。既然是八卦，就可以有原则地胡说。所谓有原则，是指它们决不是虚构的、招摇撞骗的东西。但是，由于加上了比较多的个人的好恶或评论，因此很难称之为客观。另外，它们也不象标准的教科书的内容之顺乎逻辑。逻辑的跳跃是经常出现的。因此，如果第一次阅读时有些摸不着头脑，这是完全可以理解的。也许在学完了后面的章节以后，最后再来阅读这一章体会会更多一些。无论如何，我希望大家在第一次阅读本章时能够有一个大概的印象就足够了。

1　力学的简要历史

¶每一门课之前，我总是要稍微说一些相关的历史。这不仅仅是因为个人对于历史的偏好，而是有它的理由的。当我们仔细考察一个学科的发展历史的时候，我们往往能够体

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

会到一些超出该学科的具体内容、甚至是超出历史的认识。我个人觉得这种认识对于充分理解一个学科是十分有好处的。正如人们常说的，以史为鉴总是可以使我们进步，忘记历史的人往往不得不去重复历史。

¶经典力学是一门十分古老的学说。如果从古希腊的老祖宗亚里士多德 (Aristotle) 算起来，已经有了两千多年的历史。我们并不想将历史上溯得这么远，因为太过早期的那些历史与我们的这个课程并不是直接有关。我们对于力学历史的回顾将按照时间的顺序，选取一些我认为重要的代表人物来加以说明。这种讲述从一个名叫伽利略 (Galileo Galilei, 1564-1642) 的物理学家开始是合适的。

伽利略生于意大利的 Pisa，就是造塔总是造不正的那个意大利城市。他在 Pisa 大学受到教育并任教到 1592 年。后来他转到 Padua 大学 (当时属于 Republic of Venice) 任教直到 1610 年。这一段时间是他人生中相当平静的年代，也是他很有创造性的年代。他发明望远镜，制造指南针，研究金星和土星，发现了木星的四颗卫星，进行各种力学实验。值得一提的是，为了说明日心说比地心说靠谱，1595 年他提出了解释潮汐的理论。尽管这个理论实际上是不对的 (我们在第 16 节中会给出一个正确的解释)。1611 年他去了罗马。他的动荡生活开始于 1612 年。教廷开始明确地反对哥白尼的日心说，至少在“道义上”是如此。伽利略经常被作为科学与宗教做坚决斗争的典范来加以歌颂。实际上他并不是一个实足的斗士。真正的斗士是诸如布鲁诺等人，早就被教廷消灭了 (比如，可以用火烧)。[1] 伽利略是一个信仰十分坚定，但还是相当善于妥协的人。他一直与教廷在斗争中共存直到 1632 年。在这个期间，教廷允许他宣讲他的日心说，但是必须说明这只是一个假说。1632 年，伽利略的著名著作《对话》，Dialogo sopra i due massimi sistemi del mondo (Dialogue Concerning the Two Chief World Systems) 在佛罗伦萨出版。这本书真正激怒了教廷。教廷成立了专门的委员会调查这本书。同一年他被囚禁在 Siena。不过，没有多久在伽利略服罪后 (1633 年) 他就被释放回家 (实际上是软禁在家)。随后他的生活相当贫困，晚年还几乎失明。伽利略的最大贡献是奠定了现代力学，甚至是现代物理学乃至现代科学的基础。称他为现代科学之父应当不算是过分的。我们课程中会经常提到的所谓伽利略相对性原理也是力学中十分重要的原理。[2] 这个原理到了牛顿的时期就提得更加明确了。

¶第二个要提到的重量级人物就是牛顿 (Sir Isaac Newton, 1643-1727)。这位“大牛”的一生也是十分传奇的。1669 年以前，他基本上是个未成年人。他的父亲在他出生之前就过世了 (可怜的孩子)，因此他的童年基本上是相当凄惨的。1661 年他以“勤工俭学生”(sizar) 的身份进入剑桥大学“三一学院”(Trinity College)。1665 年，年轻的牛顿刚刚显示出一些天赋，一场瘟疫使得剑桥大学关闭，[3] 而牛顿不得不回到老家林肯郡

[1] 有趣的是，直到 400 年后，教皇约翰保罗二世等才正式宣布，当年教廷对待伽利略、布鲁诺和日心说的做法是令人遗憾的。这也算是教廷正式为当年的“冤案”平反吧。参见：http://www.sciencemag.org/news/2000/03/vatican-regrets-burning-cosmologist。

[2] 该原理也出现在伽利略著名的著作《对话》(Dialogo) 中。

[3] 不象我们，在 SARS 来的时候还坚持战斗。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

(Lincolnshire)。随后的两年多的时间里，牛顿创造了人类科学史上的奇迹。他在数学、物理、天文等诸多领域进行了革命性的创造。一般认为，微积分和牛顿三定律的雏形在这时已经基本形成。1667 年，他重新回到剑桥大学。他的辉煌从 1669 年延续到大约 1687 年，也就是他在剑桥大学任教的时间。1687 年以后，牛顿基本上是一个实足的皇家官员，在科学上完全再没有什么创造。相反地，他在沽名钓誉方面的行为倒是愈演愈烈。[4] 当然，我们也不能苛责牛顿，他已经为人类贡献得足够多了。我想称他为有史以来最伟大的科学家都不无道理。他的贡献被集成在它的著名著作《原理》，Philosophiae naturalis principia mathematica 中。这部书被许多科学史专家称为人类有史以来所写成的最伟大的科学著作。这部书奠定了经典力学的基础。牛顿一生的三大贡献可以总结为：微积分的发明、经典力学牛顿定律的确立和万有引力定律的确立。[5] 上述三大发现，只要有一项就足以使其成为伟大的科学家。能集三项发现于一身，的确可以说是前无古人、后无来着。当然，他在光学、热学等其他领域也有相当重要的贡献。他的后半生被各种宗教、政治以及学术纠纷困扰，再也没有精神来研究科学。需要指出的是，尽管经典力学的基础是由牛顿所奠定的。但是目前我们大学中所教授的经典力学 (包括普物力学和理论力学) 的系统实际上与牛顿的处理方法 (特别是符号上) 不是很类似。牛顿的方法和符号实际上很少在现代的教科书中见到，因为它们实在是太难用啦。我们目前大学中经典力学的处理方式以及各种符号的运用实际上在很大程度上应归功于一批欧洲大陆的数学家，例如：John Bernoulli, Euler, d'Alembert 等人的工作。

¶ 一个比较奇怪的现象是，牛顿力学的发扬光大并不是在牛顿的手中实现的。它的伟大成就更多的是在一群欧洲大陆的数学家、物理学家的努力下实现的。这里我们就要提到我们介绍的第三位人物：瑞士数学家、物理学家欧拉 (Leonhard Euler, 1707-1783)。即使是在数学领域，欧拉也堪称是有史以来最伟大的几个数学家之一。他是一个数学天才，一个高产出的伟大数学家。可以毫不夸张地说，欧拉几乎主宰了十八世纪的欧洲的数学。在力学方面，欧拉的贡献也是惊人的。他是系统地将微积分运用到力学的第一人。他与拉格朗日一道也是首先将变分法运用到力学上的人。不幸的是，欧拉的晚年是在失明的痛苦中度过的。[6] 但是他几乎一半的重要工作都是在他失明以后得到的，这简直就是一个神人。在我们这个课程中，大家会接触到欧拉 - 拉格朗日方程。另外，欧拉对于刚体、流体力学等方面的研究也是开创性的。无怪乎拉普拉斯后来说："Read Euler, read Euler, he is the master of us all." 总之，欧拉可以说是划时代的人物，一个当时所有数学、物理特别是力学知识的创造者和集大成者。

¶ 第四个要提到的人物实际上与我们的这个课程关系更为密切，那就是

[4] 例如，胡克同志声称牛顿曾经剽窃了他的光学成果。又例如，他与莱布尼兹关于微积分的发明权的白热化争论。

[5] 虽然据称万有引力定律是牛顿受到苹果落地的启发而突发奇想，但实际上牛顿更主要地是受到了开普勒、惠更斯的结果的影响。牛顿注意到了，如果行星的轨道是圆的，那么结合开普勒的第三定律以及一个匀速圆周运动的物体的加速度的表达式：$a = v^2/r$ （这来自惠更斯），那么吸引这个行星的力一定是与距离平方成反比。

[6] 造成欧拉失明的主要原因之一是 St. Petersbourg 恶劣的气候。要知道在当年那里毕竟仍是苦寒之地。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

拉格朗日 (Joseph-Louis Lagrange, 1736-1813)。拉格朗日可以说是法国物理学家，也可以说是意大利物理学家。他生于意大利城市都灵，但是他的父亲的家族实际上是法国人。1754 年，年轻的拉格朗日提出了变分法的思想。他把他的成果写信告诉 Euler，Euler 十分欣赏他的才能。由于 Euler 的推荐，年仅 19 岁的拉格朗日担任了都灵皇家炮兵学校 (奇怪的学校吧) 的数学教授。随后，拉格朗日与欧拉又有许多通信来往，其中拉格朗日运用了最小作用量原理和变分法讨论了力学的规律 (这正是我们第一章要讨论的主要内容，这些方程后来被称为欧拉－拉格朗日方程)。欧拉为此说服莫佩督 (Maupertuis)，他当时是普鲁士柏林科学院的主席，要把拉格朗日挖到普鲁士来。拉格朗日是牛顿的崇拜者。但是与牛顿不同的是，他是一个极为谦逊的人。而且对于名利看得比较淡。他的主要心思都在学问上。所以尽管欧拉给他的 Offer 十分诱人，他还是没有接受。直到后来 (1766 年)，欧拉去了圣彼德堡，在达朗伯的大力劝说下，拉格朗日才接受柏林科学院的邀请，担任柏林科学院数学分部的领导。1788 年，拉格朗日在力学方面的著名著作 Mécanique analytique 在柏林出版。这部著作总结了牛顿以来的所有力学成果，特别是提出了有别于牛顿的力学处理方法 –即采用以 能量 为根本，而不是以力为根本的分析力学的研究方法。当时的数学家、物理学家都认为这是一部十分重要的著作。拉普拉斯、勒让德等大家都亲自为书做编辑工作。在书的绪言中拉格朗日自豪地写到：“这部书中你不会看到任何的图。我所提出的处理力学的方法不需要任何几何或力学的帮助，只需要代数的操作就足够了。”[7] 因此，经过拉格朗日的改造，力学已经 几乎 变成了分析的一个分支，所以它被后人称为 分析力学。当然，物理学并不是数学。所谓的分析力学仍然不是纯数学。它依赖于一些基本的物理学原理，这些原理仍然是需要实验验证的。因此，尽管 形式上 讲分析力学很像数学，但是从骨子里讲它仍然是物理学。

¶最后要提一下的物理学家是 哈密顿 (William Rowan Hamilton, 1805-1865)。哈密顿是一位爱尔兰数学家、物理学家。他在数学上的最主要的贡献是提出了所谓的 四元数 (quaternions)。这是哈密顿本人最为得意的贡献，他认为这是对复数的一个推广，也是人类首次处理不对易的数。从数学上讲，四元数等价于四个矩阵 (例如可以取为泡利矩阵和单位矩阵)。在物理学方面，哈密顿的主要兴趣在光学方面。他进行了大量的光学的理论研究和实际观测而力学的研究只是他的副业。但实际上他在物理学方面的贡献主要是在力学方面。哈密顿基本上建立了一套与拉格朗日力学平行的、等价的分析力学体系。这种力学体系很自然地又被称为 哈密顿力学，它可以很容易地推广到非力学的领域 (例如统计物理，还有哈密顿所钟爱的光学领域)。不仅如此，哈密顿的力学表述实际上可以很直接地推广到量子力学。因此，哈密顿力学可以说是经典力学与量子力学之间的桥梁。从研究的方法来看，哈密顿力学也是属于分析力学的范畴。这一点它与拉格朗日力学类似而不同于传统的牛顿力学。

[7] 很不幸的是，我的这个讲义中还是有许多的图。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

2 关于分析力学与矢量力学

¶传统的牛顿力学的理论体系中，力是处于中心位置的物理量。力是一个矢量，因此我们可以称这种以“力”为核心概念的力学体系为 矢量力学 。[8] 在拉格朗日和哈密顿建立的力学体系中，“力”并不处于核心的地位，体系的核心物理量是拉格朗日量和哈密顿量。这两个物理量实际上都是与 能量 密切相关的。拉格朗日力学和哈密顿力学被称为 分析力学 。在分析力学的理论体系中，处于核心地位的是能量，“力”反而处于从属的地位。下面我们简要说明一下分析力学与矢量力学之间的关系。同时，我们也讨论一下为什么需要利用分析力学的方法来研究力学问题。

¶我们先来简要说明一下矢量力学与分析力学之间的关系。最为重要的一点是：在经典力学的范畴内，牛顿的矢量力学与拉格朗日或哈密顿的分析力学是 完全等价 的。也就是说，对于一个经典力学问题而言，用任何一种方式来处理，得到的物理结果都是相同的。其次我们必须说明，牛顿的矢量力学表述方式一般只能够用于纯经典力学的范畴，很难将其推广到物理学的其他领域；相反，分析力学的表述方式可以轻易地推广到经典的场系统、经典的电磁学、经典的光学。甚至在考虑了量子力学的基本原理后，可以方便地推广到量子力学。也就是说，分析力学的语言是一种在整个物理学中更为通用的语言。随着理论物理课程的深入，同学们会越来越少地遇到“力”的概念。相应的，会越来越多地遇到拉格朗日量、哈密顿量、欧拉－拉格朗日方程、哈密顿方程等等这些分析力学的概念。因此，建立起分析力学的基本概念，掌握分析力学的基本方法是本课程的 中心任务 。我觉得这个任务远比具体解一两道力学题目要重要得多。这些基本概念和方法将会在随后的其他课程中多次被利用到。事实上，分析力学中的许多概念已经成为现代理论物理学的基础概念。

作为理论力学这门课程，我们当然仍侧重于经典力学问题的分析。但即使是在纯经典力学的范畴之内，我们会发现分析力学也具相当的优势，特别是在处理比较复杂的力学问题的时候。

作为一个例子，我们考虑一个均匀重力场中的平面双摆 (如图 1.1所示) 问题。这是一个只有两个自由度 (例如可以取为图中所示的两个角度 ϕ_1，ϕ_2) 的力学系统。这个问题并不是不能够用普通的矢量力学来解决，只是有点复杂罢了。因为为了得到两个质点的运动方程，我们必须分别分析两个质点的受力情况。显然，在第一个质点与长度为 l_2 的绳子的结合的地方，有着一个随时间变化的力。这就给问题的解决带来了一定的麻烦。我们会看到，这样一个问题利用拉格朗日分析力学可以很容易得到解决。因为在拉格朗日力学中，我们不必再去考虑复杂的力的分析。我们只需要直接写出系统的 拉格朗日量 。这是

[8] 历史上，是马赫首先指出，牛顿力学的力学体系实际上是一个循环逻辑的体系。具体来说， 力 和 质量 这两个概念是相互定义的。关于牛顿力学的框架可以参考 [4] 的 §1.2 节。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

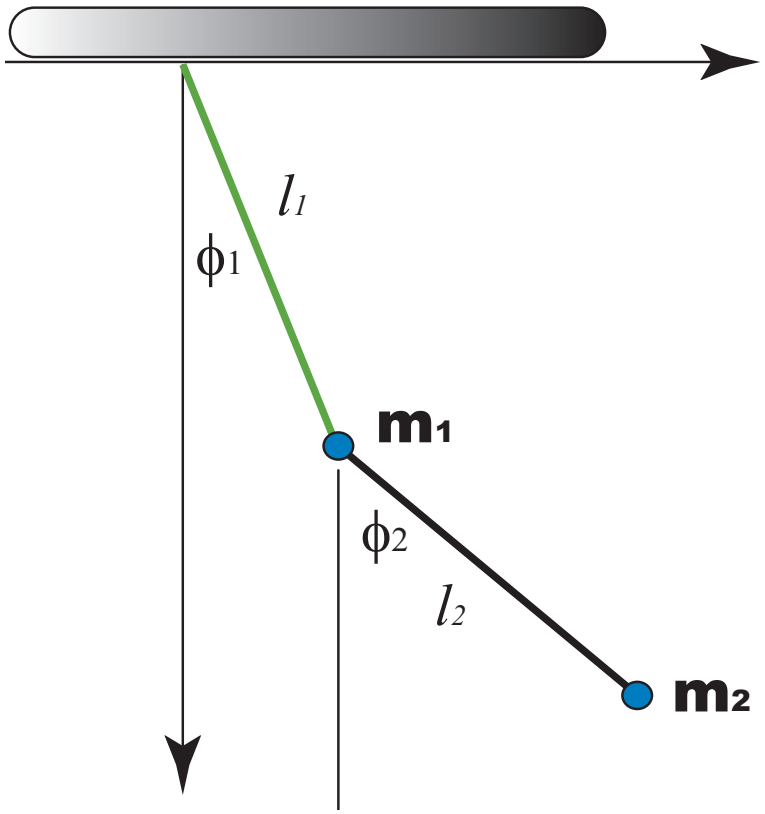

图 1.1: 均匀重力场中的双摆。一个质量为 m_1 的质点，用长度为 l_1 的绳子 (不可拉伸) 悬挂于天花板上。它的下面用长度为 l_2 的绳子再悬挂一个质量为 m_2 的质点。整个系统处于重力加速度为 g 的均匀重力场中。

每一个力学系统的特征物理量。一旦有了一个力学系统的拉格朗日量，我们直接就可以写出这个系统的运动方程。[9]

如果我们考虑的力学问题中牵涉到所谓的约束，那么这些约束往往会使得通常的牛顿矢量力学的方法变得十分复杂，特别是在复杂的约束情况下。这种复杂性体现在约束往往是减少了系统的自由度数目，但同时约束也增加了由于约束带来的约束力。这些约束力的大小和方向往往不是事先知道的，而是必须在解完这个力学问题之后才能知道。因此，约束在减少自由度数目的同时，实际上反而增加了牛顿方程中未知数的个数。当然，在正确的力的分析之后，我们总是可以列出正确的牛顿方程。这些方程中包含约束力作为未知数。在消去了这些未知数之后，我们就可以得到体系的运动方程。求解了体系的运动方程之后，我们也可以得到相应的约束力。但是，如果我们并不那么关心约束力的情况，而是仅仅关心系统的真实自由度的运动情况。这个时候上面描述的矢量力学的求解方法就显得很不方便了。

[9] 这里我们并没有求解这个问题。关于平面双摆的小振动问题的解，可以参考第 四章第 20节中的例题 5.1。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

与矢量力学的方法不同的是，分析力学的处理方法是直接从体系的真实自由度出发，写出系统的拉格朗日量，然后就可以直接得到系统真实自由度的运动方程。直接求解这些方程，就可以得到系统的运动状况。这从根本上避免了复杂的力的分析这个中间环节。以我们前面讨论的双摆为例，如果我们根本不关心 m_1 所受到的力的情况，而仅仅关心整个系统的运动状况，我们就可以直接写出系统的动能 T 和势能 V：

$$\begin{aligned} T &= \frac{1}{2}(m_1+m_2)l_1^2\dot{\phi}_1^2 + \frac{1}{2}m_2 l_2^2\dot{\phi}_2^2 + m_2 l_1 l_2 \dot{\phi}_1\dot{\phi}_2\cos(\phi_1-\phi_2)\ , \\ V &= -(m_1+m_2)gl_1\cos\phi_1 - m_2 g l_2\cos\phi_2\ . \end{aligned} \tag{1.1}$$

写出了系统的动能和势能，我们可以定义一个新的物理量，它称为系统的拉格朗日量。目前这个情况下，系统的拉格朗日量可以取为其动能减去势能：[10]

$$L(\phi_1,\phi_2;\dot{\phi}_1,\dot{\phi}_2) = T - V\ . \tag{1.2}$$

我们随后的讨论会说明，只要写出了系统的拉格朗日量作为其广义坐标 ϕ_1，ϕ_2 以及相应的广义速度 $\dot{\phi}_1$，$\dot{\phi}_2$ 的函数，我们就可以直接 (利用最小作用量原理和变分法) 得到系统的两个自由度 ϕ_1，ϕ_2 的运动方程。由于系统的能量 (动能、势能以及拉格朗日量) 是标量，写出它的形式远比分析各个物体之间的受力要容易。因此在分析力学中，我们可以完全不必考虑系统中复杂的约束相互作用力，可以直接写出体系自由度的运动方程。

¶ 分析力学的另一个优势是它与对称性的紧密联系。分析力学将一个广义的力学系统的几乎所有力学性质都集成在系统的作用量 (拉格朗日量对时间的积分) 之中。从系统的拉格朗日量 (作用量) 出发，就可以直接得到其运动方程。因此，如果系统的拉格朗日量具有某种对称性，这种对称性必然地反映在其运动方程中。特别是对于那些反映时空基本性质对称性，例如时间平移、空间平移、空间转动、空间反射 (宇称)、时间反演等等 (或者它们的相对论性推广)，分析力学都可以直接得到重要的守恒定律。这种与对称性的紧密联系在量子力学中将体现得更为突出。当然，对于纯力学体系，如果利用传统的矢量力学也可以得到这些守恒定律，只不过并不是非常直接，同时也无法推广到其他物理领域。

¶ 综上所述，我们课程中要主要讲述的分析力学具有如下的优点：第一，它在纯经典力学的范畴内与牛顿矢量力学完全等价；第二，它对于经典力学问题的处理更加简洁优美；第三，它能够更深刻而直接地体现系统的各种对称性及其守恒定律；第四，它的基本原理具有更广泛的适用范围，能够推广到经典力学之外的多个物理领域。这么好的东东，希望同学们能够好好消化哦！

[10] 这一点目前大家只好先接受。关于这点的初步解释可以参考第 4节的公式 (1.24)。关于如何写出一个力学系统的拉格朗日量，以及如何由其拉格朗日量直接写出其运动方程恰恰是我们下一章要讨论的主要内容。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

3 关于约束

¶这一节中我们讨论一些更为具体的问题，那就是关于有约束的经典力学系统。约束是很“恶心”的东西。[11] 所以，我实在不想把它放到理论力学的正式叙述中。但遗憾的是，如果理论力学课程中一点儿也不提及约束似乎也实在是说不过去。所以，我们在这里简要地介绍一下。这一节中，我们将从传统的牛顿力学出发，来考察有约束的力学体系。我们会发现，很多情形下 (我下面会解释何谓很多情形) 我们自然地需要引入系统的广义坐标；同时系统的运动方程也自然地具有拉格朗日方程的形式。事实上，许多的理论力学书籍就是从这一线索出发来建立拉格朗日分析力学体系的。在这个课程的主体部分中，我选择了一种不从牛顿力学出发的讲述方法，主要是因为这样可以突显最小作用量原理 (又称哈密顿原理) 的崇高地位，从而说明它是高于牛顿力学的物理学普遍原理。

¶现在让我们暂时从牛顿力学 (矢量力学) 出发，考察一系列粒子 (它们的质量我们记为 m_i) 的运动方程。我们用直角坐标 $\mathbf{x}_i$ 表示第 i 个粒子的位置矢量，那么它的牛顿力学运动方程可以形式地写为：

$$m_i\ddot{\mathbf{x}}_i = \mathbf{F}_i^{(e)} + \sum_j \mathbf{F}_{j\to i}\ , \tag{1.3}$$

其中 $\mathbf{F}_i^{(e)}$ 是第 i 个粒子所感受到的外力；$\mathbf{F}_{j\to i}$ 是第 j 个粒子作用于第 i 个粒子的内力。初看起来，方程 (1.3) 已经完全确定了这个力学体系的运动。只要我们已知所有的力（方程的右边），并且知道了各个粒子的初条件，剩下的问题只是求解微分方程的数学问题了。实际上并不是所有力学问题都如此简单。一类常见的复杂情况出现在力学系统的坐标或速度必须满足一些约束的时候。有约束的力学问题大量出现在结构力学、材料力学研究中。这些约束条件有时可以是十分复杂的。我们将主要讨论一类最为简单的约束。如果系统的约束可以表达为联系系统各个坐标和时间的 (一组) 方程：

$$f_m(\mathbf{x}_1, \mathbf{x}_2, \cdots, t) = 0\ ,\quad m = 1, 2, \cdots, k\ , \tag{1.4}$$

那么我们称这样的约束为完整约束 (holonomic constraint)。当然，并不是所有的约束条件都可以写成完整约束的形式。例如，有些约束包含粒子的速度 (或更坐标的更高阶时间导数)；有些约束甚至不是以等式的形式出现的 (例如以不等式的形式出现)。所有的不属于完整约束的约束都被统称为非完整约束。本课程中，我们将主要处理完整约束的情形。不过，为了给大家一个大致的概念，我们下面会举一个典型的例子来说明什么是非完整约束，这也可以让我们认识一下约束的复杂性。如果约束可以写为一系列方程并且这

[11] 其“恶心”程度在你打算将它量子化时会更加膨胀！

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

些方程中不显含时间，这样的约束被称为稳定约束 (scleronomous constraint)；反之则称为不稳定约束 (rheonomous constraint)。

即使是最为简单的完整约束也会给求解力学问题 (1.3) 带来麻烦。第一，所有质点的直角坐标 $\mathbf{x}_i$ 将不再是独立的。它们除了满足各自的运动方程 (1.3) 之外，还必须满足约束条件所要求的约束方程 (1.4)。第二，一旦有了约束，一般就会有相应的约束力的出现。例如，如果我们限制粒子在一个给定的曲面上运动，那么粒子在这个曲面上运动时会感受到约束力，这些约束力的大小在问题的求解之前一般是未知的。我们仅仅知道它们一定沿着曲面的法线方向 (假定没有切向的摩擦力)。只有在整个力学问题完全求解以后，我们才能知道这些约束力的大小。也就是说，每个粒子的运动方程 (1.3) 右边的力中有一部分是未知的。

为了解决完整约束带来的第一个问题，我们可以引入广义坐标。这是一个完整约束力学系统所有独立的自由度的一个最小集合。如果我们考察的力学体系的完整约束条件一共有 k 个 (如方程 (1.4) 所示)；那么整个力学体系 (由 N 个质点构成的体系) 的自由度数目就是 $3N-k$。完整约束的条件意味着：

$$\begin{aligned} \mathbf{x}_1 &= \mathbf{x}_1(q_1, q_2, \cdots, q_{3N-k}, t) \,, \\ &\vdots \qquad\qquad \vdots \\ \mathbf{x}_N &= \mathbf{x}_N(q_1, q_2, \cdots, q_{3N-k}, t) \,, \end{aligned} \tag{1.5}$$

这样的一组参数方程等价于 k 个完整约束的方程 (1.4)。这样一组完整约束同时隐含着对于体系速度的约束，因为上述各个约束对于时间的导数仍然成立：

$$\dot{\mathbf{x}}_i = \frac{\partial \mathbf{x}_i}{\partial q_l}\dot{q}_l + \frac{\partial \mathbf{x}_i}{\partial t} \,, \tag{1.6}$$

这里 $i = 1, 2, \cdots, N$，$l = 1, 2, \cdots, 3N-k$，同时我们启用了爱因斯坦 求和规则：即重复的指标 (例如上式右边的指标 l) 意味着对其可能的取值进行求和。

这里需要特别强调指出的是，只有对于完整约束的力学体系，我们才能够仅仅取其真实的自由度为广义坐标，这时系统的广义坐标的数目等于 其自由度数目。对于非完整约束，这一点一般是做不到的。原因是非完整约束条件中包含有速度，而且这些速度是无法消去的。因此，我们无法利用含有速度的约束条件消去多余的自由度。所以，这个时候我们选取的广义坐标的数目一般大于 系统的自由度的数目（见下例）。为了说明各种约束 (特别是非完整约束) 对于力学问题的影响，我们考虑一个典型例子就是一个在二维平面上做纯滚动 (无滑动的滚动) 的圆盘。

例题 1.1　二维平面上垂直纯滚的均匀圆盘 考虑一个半径为 a 的均匀圆盘，它在二维平面 (取为 $x-y$ 平面) 上无滑动地纯滚。为了简单起见，假定圆盘中心 (质心)C 点与圆盘和

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

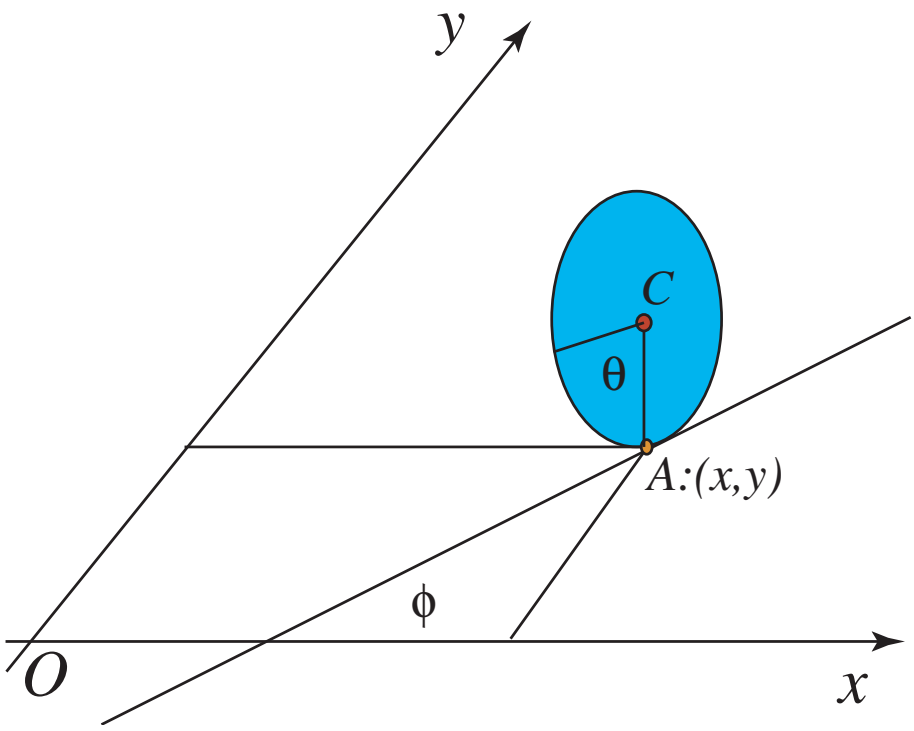

图 1.2: 一个在二维平面上纯滚动 (无滑动的滚动) 的圆盘。这是一个非完整约束的典型例子。

平面的接触点 A 之间的连线永远垂至于 $x-y$ 平面，参见图 1.2。下面来分析它的力学自由度和约束的情况。

解： 我们可以取 A 点的坐标 (x,y) 为两个广义坐标来描写圆盘质心的位置。但是这还不足以完全确定圆盘的位置。另外一个需要知道的物理量是圆盘平面在 $x-y$ 平面的投影与 x 轴之间的夹角 ϕ。除此以外，我们如果选定圆盘上面一个固定的点 B，圆盘中心与 B 的连线 (CB) 与圆盘中心到 $x-y$ 平面的垂线 (也就是 CA) 之间的夹角 θ。如果圆盘永远与 $x-y$ 平面垂直运动，那么显然给定了 x，y，ϕ，θ 之后，圆盘的位置就唯一地确定了。我们可以选取这 4 个参数为圆盘的 4 个广义坐标。

但是，上面的讨论还没有考虑到纯滚的条件。如果圆盘不能够在平面上滑动，那么显然上述几个广义坐标所对应的广义速度之间有联系。这个条件可以表述为:

$$\dot{x}=a\dot{\theta}\cos\phi\,,\quad \dot{y}=a\dot{\theta}\sin\phi\,,$$

或者等价地写为:

$$dx=ad\theta\cos\phi\,,\quad dy=ad\theta\sin\phi\,, \tag{1.7}$$

也就是说，广义坐标 x 和 y 的微分完全由广义坐标 θ 的微分以及另外一个广义坐标 ϕ 确定。它们不是独立的。于是，这个力学系统中，尽管我们可以选取 4 个广义坐标，它的独立的自由度数目只是 $4-2=2$ 个。最为糟糕的是，公式 (1.7) 所体现的约束无法积分

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

出来。也就是说，它无法转换成仅仅包含坐标的一个约束。这个约束中势必包含广义速度。这也就是我们前面提到的 非完整约束。因此，对于这个系统，尽管系统的自由度数目是 2，我们没有办法仅仅取两个广义坐标，而必须选取 4 个广义坐标，因为两个约束条件 (1.7) 中含有速度 (非完整约束)。应当说，这个例子中讨论的仍然是非完整约束中比较好处理的一类。事实上，这个问题是可以利用拉格朗日方程和推广的最小作用量原理来处理的 (参见第 13节的讨论，那里我们会给出这个力学问题的完全解)。实际的应用中可能还会出现更加恶心的约束，例如由不等式描写的约束。

如果我们令 ϕ 永远只能取常数，这时约束条件 (1.7) 可以积分出来，使之变为完整约束。这时，圆盘实际上只是在一条一维的直线上纯滚，它的自由度数目是 1。这时，实际上我们也可以只取一个广义坐标。在另一个极端下，如果圆盘可以发生滑动，那么上面讨论的约束条件根本不存在，系统变成完全没有约束的力学体系，其自由度数目是 4。

暂时搞定

我们这里仅仅以这个例子说明广义坐标、自由度、完整约束、非完整约束的概念。这里并没有完全解这个力学问题。这个例子的完全解我们会在后面给出，参见第 13节中的例题 2.1。

4 虚功原理和达朗伯原理

¶ 前一节对于约束的各种情况作了简要的介绍。通过这些简单的例子我们发现：约束除了造成各个坐标不独立之外，一般还会引入未知的约束力。为了要简化约束力造成的不便，我们可以应用所谓的 虚功原理。

为此，我们考虑一个处于 力学平衡 的、有约束的力学体系。考虑某个时刻 t 系统坐标的一个微小的、与运动方程和约束条件都兼容的虚拟位移：$\delta\mathbf{x}_i$。这称为该力学系统的一个 虚位移。如果作用在质点 i 上的力为 $\mathbf{F}_i$，由于每一个质点都处于力学平衡，显然 $\mathbf{F}_i = 0$。因此我们有：

$$\sum_i \mathbf{F}_i \cdot \delta\mathbf{x}_i = 0\,. \tag{1.8}$$

我们现在将作用于粒子 i 上的力 $\mathbf{F}_i$ 分为两个部分：$\mathbf{F}_i = \mathbf{F}_i^{(a)} + \mathbf{F}_i^{(c)}$，其中 $\mathbf{F}_i^{(a)}$ 是所谓的 主动力，也就是除去约束引起的力之外的所有的力；$\mathbf{F}_i^{(c)}$ 称为 约束力，它完全是由于约束条件引起的力。于是我们有：

$$\sum_i \mathbf{F}_i^{(a)} \cdot \delta\mathbf{x}_i + \sum_i \mathbf{F}_i^{(c)} \cdot \delta\mathbf{x}_i = 0\,,$$

现在我们假设我们的约束永远满足：

$$\sum_i \mathbf{F}_i^{(c)} \cdot \delta\mathbf{x}_i = 0\,. \tag{1.9}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

也就是说约束力的虚功之和为零。这包含了相当多的一类完整约束。例如，如果粒子只能在一个曲面上运动，那么约束力一定沿曲面法线方向 (假定没有滑动摩擦力)，而虚位移一定沿切向，因而其虚功为零。在此条件下我们得到：

$$\sum_i \mathbf{F}_i^{(a)} \cdot \delta \mathbf{x}_i = 0 \ . \tag{1.10}$$

即所有主动力的虚功之和也为零。这个结论被称为静力学中的虚功原理。注意，由于所有的虚位移 $\delta \mathbf{x}_i$ 不再是独立的 (有约束)，因此我们并不能由此得出所有的主动力都为零。虚功原理的用处在于，它仅仅涉及到主动力 (已知的) 的虚功，而不涉及未知的约束力。因此，这个原理可以方便地运用到有约束的静力学问题中。

¶ 如果我们考察的力学系统并不处于力学平衡，那么我们可以将力学的系统的运动方程 $\mathbf{F}_i = 0$ 替换为牛顿方程 $\mathbf{F}_i - \dot{\mathbf{p}}_i = 0$。于是我们得到：

$$\sum_i (\mathbf{F}_i^{(a)} - \dot{\mathbf{p}}_i) \cdot \delta \mathbf{x}_i = 0 \ . \tag{1.11}$$

这个公式通常被称为达朗伯原理 (d'Alembert's principle)。这个原理可以看成是分析力学的一个基本原理。

¶ 现在假定我们仅仅考虑完整约束的力学体系。正如我们前面提到的，对这类力学体系我们可以选取独立的广义坐标 q_i，其数目恰好等于体系的自由度数目，即公式 (1.5) 成立。从上述达朗伯原理出发，利用公式 (1.5) 进行换元，我们可以将达朗伯原理中的虚位移用广义坐标 q 的变分写出：

$$\delta \mathbf{x}_i = \frac{\partial \mathbf{x}_i}{\partial q_j} \delta q_j \ . \tag{1.12}$$

其中我们引用了爱因斯坦求和规则，即重复的指标意味着求和。于是主动力的虚功为：

$$\mathbf{F}_i^{(a)} \cdot \delta \mathbf{x}_i = \mathbf{F}_i^{(a)} \cdot \frac{\partial \mathbf{x}_i}{\partial q_j} \delta q_j = Q_j \delta q_j \ , \tag{1.13}$$

其中我们定义了广义力：

$$Q_j = \mathbf{F}_i^{(a)} \cdot \frac{\partial \mathbf{x}_i}{\partial q_j} \ . \tag{1.14}$$

达朗伯原理中的另外一项涉及动量的时间导数。我们也将它换为广义坐标及其时间导数表达：

$$\dot{\mathbf{p}}_i \cdot \delta \mathbf{x}_i = m_i \ddot{\mathbf{x}}_i \cdot \frac{\partial \mathbf{x}_i}{\partial q_j} \delta q_j \tag{1.15}$$

现在注意到：

$$m_i \ddot{\mathbf{x}}_i \cdot \frac{\partial \mathbf{x}_i}{\partial q_j} = \frac{d}{dt}\left(m_i \dot{\mathbf{x}}_i \cdot \frac{\partial \mathbf{x}_i}{\partial q_j} \right) - m_i \dot{\mathbf{x}}_i \cdot \frac{d}{dt}\left(\frac{\partial \mathbf{x}_i}{\partial q_j} \right) \ . \tag{1.16}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

另一方面：

$$\mathbf{v}_i = \frac{d\mathbf{x}_i}{dt} = \frac{\partial \mathbf{x}_i}{\partial q_j}\dot{q}_j + \frac{\partial \mathbf{x}_i}{\partial t} , \tag{1.17}$$

从而我们有：

$$\frac{\partial \mathbf{v}_i}{\partial \dot{q}_j} = \frac{\partial \mathbf{x}_i}{\partial q_j} . \tag{1.18}$$

这些关系带入到公式 (1.16) 中，我们得到发现达朗伯原理可以写为：

$$\left[\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} - Q_j\right]\delta q_j = 0 . \tag{1.19}$$

对于完整约束，由于我们取的各个 q_j 是独立的变量，因此达朗伯原理要求：

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} = Q_j . \tag{1.20}$$

这就是著名的 欧拉 - 拉格朗日方程 (的一种形式)。如果主动力是由一个不依赖于速度的势能 V 给出的：

$$\mathbf{F}_i = -\frac{\partial V}{\partial \mathbf{x}_i} , \tag{1.21}$$

其中 V 只是各个坐标的函数，那么上面定义的广义力可以写为：

$$Q_j = -\frac{\partial V}{\partial q_j} , \tag{1.22}$$

于是前面给出的欧拉 -拉格朗日方程可以写为：

$$\boxed{\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0 .} \tag{1.23}$$

其中我们定义的系统的 拉格朗日量 L 为系统的动能与势能的差：

$$L = T - V , \tag{1.24}$$

这部分回答了前面提及的事情，对一个保守体系，它的拉格朗日量可以取为动能减去势能 (参见第 2节)。于是我们看到，只要得到了系统的拉格朗日量，我们就可以直接写出系统的含有真实自由度的运动方程 (1.23)。

5 关于有耗散的系统

¶我们这个课程中绝大多数情况下讨论的都是没有耗散的系统，这类系统大体可以分为两类：一类是所谓的 保守系统 (封闭系统)；这类系统的能量 (在纯力学范畴中就是机械

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

能) 是守恒的，时间反演也是对称的。另一类无耗散的系统是在 完全给定 外力作用下的力学系统 (典型的例子是受迫振动的振子)。这类系统的能量虽然不守恒，但是它随时间的变化完全是给定的时间的函数，并不依赖于系统本身的性质。当然，我们的课程中 偶尔 还会遇到所谓 耗散系统 。这类体系会通过耗散 (例如摩擦力) 损失能量。一般来说，一个有耗散的系统的问题已经不是一个纯粹的力学问题，因为耗散往往使这些机械能转化为体系的内能 (热力学问题)。如果这种耗散并不是非常剧烈 (何谓剧烈往往很难普遍地定义，而必须依赖于具体系统的力学和热力学性质)，那么我们可以在力学的范畴之内 唯象地 来讨论它，其代价是引入另一个特征函数。

考虑最为简单的一个单粒子的耗散情况。如果粒子所受到的耗散不是很剧烈，那么它可以用一个与粒子速度成正比的摩擦力 (或者说粘滞力)$\mathbf{F}^{(D)} = (F_1^{(D)}, F_2^{(D)}, F_3^{(D)})$ 来描写：

$$F_i^{(D)} = -k_i v_i \ , \quad i = 1, 2, 3 \ , \tag{1.25}$$

或者我们可以将它写为：

$$\mathbf{F}^{(D)} = -\frac{\partial \mathcal{F}}{\partial \mathbf{v}} \ , \tag{1.26}$$

其中函数 $\mathcal{F} = (1/2)(k_1 v_1^2 + k_2 v_2^2 + k_3 v_3^2)$ 被称为该粒子的 瑞利耗散函数 。如果我们有 N 个粒子，那么整个体系的瑞利耗散函数为：

$$\mathcal{F} = \sum_{i=1}^{N} \frac{1}{2} \left(k_1 v_{i1}^2 + k_2 v_{i2}^2 + k_3 v_{i3}^2 \right) \ . \tag{1.27}$$

瑞利耗散函数的物理意义十分明显：在 dt 时间内力学系统抵抗耗散力所做的功为:

$$dW = -\mathbf{F}^{(D)} \cdot d\mathbf{x} = 2\mathcal{F} dt \ . \tag{1.28}$$

当耗散存在的时候，我们在前一节推导的公式 (1.22) 将不再成立。广义力 Q_j 这时应当包含两个部分：一个是来自于纯力学势的力，这一部分可以吸收到拉格朗日量的定义中；另一部分则来自于耗散力，它可以用瑞利耗散函数表达为：

$$Q_j = \sum_i \mathbf{F}_i^{(D)} \cdot \frac{\partial \mathbf{x}_i}{\partial q_j} = -\frac{\partial \mathcal{F}}{\partial \dot{q}_j} \ . \tag{1.29}$$

从而有耗散的系统的拉格朗日方程可以写为：

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} + \frac{\partial \mathcal{F}}{\partial \dot{q}_j} = 0 \ . \tag{1.30}$$

因此，对于一个有耗散的系统我们必须知道系统的拉格朗日量和瑞利耗散函数这两个特性函数之后，才能够完全确定它的运动方程。这也是为什么我们说，有耗散的系统原则上讲已经不是一个纯粹力学的系统，仅仅知道拉格朗日量已经不足以确定其运动行为了。最后我们还应指出，能够利用瑞利耗散函数来描写的系统仅仅是有耗散系统中的一小部分。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

还有许多系统的耗散是不能用一个简单的耗散函数来确定的，只不过我们这个课程绝对不会去涉及罢了。



相关的阅读

历史上最早的力学书大概是亚里士多德的《物理学》，尽管其中有不少错误的物理，但是毕竟是第一部物理书。关于力学的历史，有兴趣的同学可以参考 Truesdell 的书籍 [5]。另外，许多有关物理学史的信息可以在网络上找到。当然，请看那些比较正式的网站，不要轻信一些非正式网站或者 BBS 上面的八卦。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

第二章　力学体系的作用量与运动方程

本章提要

- 狭义相对论的时空观 (6)
- 最小作用量原理 (7)
- 相对论性粒子的作用量 (8)
- 粒子与外场的相互作用 (9)
- 非相对论近似 (10)
- 对称性 (11)
- 非惯性系中的力学 (12)
- 非完整约束系统的最小作用量原理 (13)

6　狭义相对论的时空观

力学与时空观是密切联系在一起的。如果我们不考虑爱因斯坦的广义相对论，那么一个力学体系的运动规律是在一个时空背景上加以描述的。这个时空背景就被称为参照系。力学规律虽然是通过一个具体的参照系来加以描述的，但是物理的规律本身实际上是不依赖于参照系的选取的。以牛顿力学为例，它的力学规律是建立在所谓的伽利略时空观基础上的；而所谓的相对论力学则是以狭义相对论的时空观为基础的。在不同的参照系之间，各种物理量的数值并不一定相同，但是它们所体现出来的基本物理规律应当是共同的。这一点就称为相对性原理。这一原理在伽利略时空观和狭义相对论时空

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

观中同样成立。这一节中，我们简要复习一下狭义相对论的基本要点。这将是我们后面引入相对论性分析力学的基础。

爱因斯坦的狭义相对论是建立在相对性原理和光速不变原理基础上的时空观。[1] 在这种时空观中，时间和空间不可分割地统一在一起，构成了所谓的闵氏时空(Minkowski spacetime)。不同参照系之间的变换—也就是著名的洛伦兹变换—可以看成是闵氏时空中坐标系之间的变换(或者说“转动”)。

¶ 在洛伦兹变换下具有“确定”变换规则的物理量统称为张量。张量可以按照其独立指标个数的多少(这又被称为张量的阶)分为不同阶数的张量。最简单的张量–也就是零阶张量–在洛伦兹变换下不变，它们又被称为洛伦兹标量。例如，两点(或者说两个事件)之间的不变间隔的平方：Δs^2 就是一个洛伦兹标量。[2]

在狭义相对论时空观中，空间和时间坐标一起构成了洛伦兹四矢量，也就是一阶张量。与通常的三维欧几里德空间不同的是，在四维闵氏空间中我们将区别两种不同变换规则的洛伦兹四矢量。[3] 我们将闵氏空间中的时空坐标用一个四维空间的矢量来标记。具体地说，我们令：

$$x^0 = ct\ ,\ x^1 = x\ ,\ x^2 = y\ ,\ x^3 = z\ . \tag{2.1}$$

我们将使用 x^μ 来统一标记 (x^0, x^1, x^2, x^3)，称它为一个逆变 四矢量(或逆变矢量)。用它的时间分量与空间分量表达，一个逆变四矢量可以写成：

$$x^\mu = (x^0, \mathbf{x})\ . \tag{2.2}$$

我们约定：一个逆变四矢量的指标出现在代表它的符号的右上角。这些在物理量右上角的张量指标因此也被称为逆变指标或者上标。与逆变四矢量相对应，我们称

$$x_\mu = (x^0, -\mathbf{x})\ . \tag{2.3}$$

为一个协变 四矢量。一个协变四矢量的指标是在其符号的右下角。它和相应的逆变四

[1] 相应的，伽利略时空观是建立在相对性原理和信号传递最大速度是无穷大的基础之上的。

[2] 这实际上是光速不变原理的直接要求。

[3] 从原则上讲，在狭义相对论中也可以在所有四矢量的零(时间)分量中引入纯虚数单位 i。这样一来可以不必区分协变四矢量和逆变四矢量，也可以不必引入度规张量 $\eta_{\mu\nu}$。不过，虽然对于狭义相对论来说这也许比较方便，但对于广义相对论来说，引入度规是不可避免的。所以我们采用了引入度规和两种四矢量的讲述方法。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

矢量的时间分量相同，空间分量相差一个负号。这些处于物理量符号的右下角的张量指标因此被称为协变指标 或者下标。

从形式上讲，一个协变四矢量 x^μ 和与其相应的逆变四矢量 x_μ 之间可以通过升高或降低指标 的操作来相互转换：[4]

$$x_\mu = \eta_{\mu\nu} x^\nu \ , x^\mu = \eta^{\mu\nu} x_\nu \ . \tag{2.4}$$

其中 $\eta_{\mu\nu}$ 称为闵氏空间的度规张量，而 $\eta^{\mu\nu}$ 为度规张量的逆，也就是说它们满足：

$$\eta_{\mu\beta}\eta^{\beta\nu} = \delta^\mu_\nu \ , \tag{2.5}$$

其中 δ^μ_ν 为 Kronecher 符号，它在两个指标相同时为 1；不同时为零。在狭义相对论的闵氏时空中，度规张量 $\eta_{\mu\nu}$ 和它的逆 $\eta^{\mu\nu}$ 的每一个分量其实都相等，也就是说这时 $\eta_{\mu\nu}$ 自己就是自己的逆。它们的表达式为：[5]

$$\eta_{00} = \eta^{00} = 1 \ , \eta_{ii} = \eta^{ii} = -1 \ , \ \ i = 1, 2, 3. \tag{2.6}$$

它们的其余分量皆为零。大家不难验证，公式 (2.4) 其实与我们前面给出的逆变、协变四矢量的定义是完全一致的，只不过换了一种更为“文明”的写法而已。

¶ 在公式 (2.4) 中，我们将度规张量 $\eta_{\mu\nu}$ 中的一个协变指标与逆变四矢量 x^ν 的逆变指标取为相同并且 (按照爱因斯坦求和规则) 求和，这样的操作被称为指标的缩并 (contraction)。需要指出的是，在引入了协变矢量和逆变矢量后，所有的缩并一定是 在一个协变指标和一个逆变指标之间进行。也就是说，总是一个上标和一个下标缩并，而绝不会有两个上标或两个下标之间的缩并。被缩并的一对指标由于已经被求和掉了，因此它们实际上已经不再具有矢量指标的含义了。正因为如此，它们也被称为傀标。一对傀标原则上可以随意换成任意的字母，只要仍然是重复的指标 (即保持被求和的状态) 就可以了。

指标的缩并一大优势是可以用来构造洛伦兹标量。例如，利用度规张量、逆变四矢量以及它相应的协变四矢量，闵氏空间中的两个无限接近的点 (事件) 之间的不变间隔平方 ds^2 可以写成下列等价形式中的任何一种：

$$ds^2 = dx^\mu dx_\mu = \eta_{\mu\nu} dx^\mu dx^\nu = \eta^{\mu\nu} dx_\mu dx_\nu \ . \tag{2.7}$$

[4] 这里我们再次运用爱因斯坦求和规则，即对于重复的指标求和。
[5] 闵氏时空中的度规的定义并不统一，有的书中采用的 $\eta_{\mu\nu}$ 的定义与我们这里的定义正好相差一个负号。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

它在任意的洛伦兹变换下是不变的标量。它可以看成是闵氏空间中无穷接近的两点 (两个事件) 之间的“距离”的平方，只不过这种距离是以 $\eta_{\mu\nu}$ 为度规的，因此距离的平方并不一定总是正的。

将任何一个协变四矢量 A_μ 与任意一个逆变四矢量 B^μ 相乘并且缩并它们的指标，我们也得到一个洛伦兹标量，它被称为这两个四矢量的内积：

$$A \cdot B = A^\mu B_\mu = A_\mu B^\mu \ . \tag{2.8}$$

我们前面提到的不变间隔的平方 (2.7) 就是坐标间隔四矢量 dx 与它自己的内积。因此从数学上讲，逆变四矢量的矢量空间与协变四矢量的矢量空间实际上构成了相互对偶的矢量空间。用逆变指标和协变指标的语言来说，只要我们将一个逆变指标和一个协变指标缩并，假定我们的物理量中再没有其他指标，我们就一定得到一个洛伦兹标量。一般来说，将一个任意多个指标的张量的一个上标与一个下标缩并，我们就得到了一个阶数减少 2 的张量。

¶ 前面已经提到，狭义相对论中的四矢量具有明确的变换规则。以 x^μ 为例，在不同参照系之间的洛伦兹变换下我们有：

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \ , \tag{2.9}$$

其中 $\Lambda^\mu{}_\nu$ 是洛伦兹变换的矩阵。[6] 一个任意的洛伦兹变换十分类似于一个“四维空间”中的广义转动。它可以分解为六种基本“转动”的合成。这六种“转动”分别对应于在 $0-1$，$0-2$，$0-3$ 平面内的转动 (它们又被称为在相应方向的 boost) 和在 $1-2$，$1-3$，$2-3$ 平面内的转动。这后三种转动其实就构成了普通三维空间内的转动。[7]

事实上，我们可以将变换规则 (2.9) 视为 (逆变) 四矢量的定义。换句话说，凡是在洛伦兹变换下按照这个形式变换的量就可以称为 (逆变) 四矢量。利用度规张量降低指标，我们可以得到其相应的协变四矢量 (x_μ) 的变换规则。类似地，一个具有任意个上标或下标的张量在洛伦兹变换下的变换规则就是它的每一个上标 (下标) 都按照相应的逆变 (协

[6] 请注意洛伦兹变换矩阵 $\Lambda^\mu{}_\nu$ 的两个指标的安排。我特别注意将上下两个指标错开，不要排在一列上 (如 Λ^μ_ν)。当然，我们可以利用 $\eta_{\mu\nu}$ 或 $\eta^{\mu\nu}$ 将第一或第二个指标进行升降 (例如：$\Lambda_{\mu\nu}$，或 $\Lambda^{\mu\nu}$)，但是由于一般 $\Lambda^{\mu\nu} \neq \Lambda^{\nu\mu}$，因此两个指标的前后是有关系的。

[7] 对于同学们熟悉的一个参照系相对于另一个参照系沿 x 轴方向匀速运动的情况 (我们称之为沿着 x 轴的一个 boost)，这里的矩阵 $\Lambda^\mu{}_\nu$ 具有十分简单的形式：

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} , \tag{2.10}$$

其中 $\beta = v/c$，$\gamma = 1/\sqrt{1-\beta^2}$。这也就是大家在普通物理力学中经常见到的形式。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

变) 四矢量的变换规则来变。例如，一个张量 $A^{\mu\nu}{}_{\rho\sigma\kappa}$ 具有两个上标和三个下标，那么它在洛伦兹变换下的规则应当是：

$$A'^{\mu'\nu'}{}_{\rho'\sigma'\kappa'} = \Lambda^{\mu'}{}_{\mu}\Lambda^{\nu'}{}_{\nu}\Lambda_{\rho'}{}^{\rho}\Lambda_{\sigma'}{}^{\sigma}\Lambda_{\kappa'}{}^{\kappa}A^{\mu\nu}{}_{\rho\sigma\kappa} \ . \tag{2.11}$$

请大家特别注意这个式子中的上下标安排。它符合我们前面所说的，一定是一个上标和一个下标缩并。同样的，这个式子实际上可以看成是型如 $A^{\mu\nu}{}_{\rho\sigma\kappa}$ 的张量的定义。

前面提到的任意两个四矢量的内积都是在洛伦兹变换下不变的事实对于洛伦兹变换矩阵 Λ 有所限制。这个限制的数学描述就是：

$$\Lambda_{\mu\alpha}\Lambda^{\mu\beta} = \delta^{\beta}_{\alpha} \ . \tag{2.12}$$

事实上，所有满足这个条件的变换矩阵构成一个群，它被称为**洛伦兹群**。

¶ 如果一个物理量是时空点的函数，我们称之为**场**。场按照其在洛伦兹变换下的变换性质又可以分为标量场、矢量场、张量场等等。一个非常重要的性质就是对于时空的偏微商算符可以作用在场上，构成多一个指标 (协变或逆变) 的场。例如，利用偏微商的锁链法则和洛伦兹变换所满足的性质 (2.12) 很容易证明：

$$\partial_{\mu} \equiv \frac{\partial}{\partial x^{\mu}} \ , \tag{2.13}$$

实际上是一个具有协变指标 (下标) 的微分算符。因此，如果我们有一个标量场 $\Phi(x)$，那么 $\partial_{\mu}\Phi(x)$ 就是一个协变四矢量场。当然，我们可以利用度规张量升高指标得到具有逆变指标的微分算符：$\partial^{\mu} = \eta^{\mu\nu}\partial_{\nu} = \partial/\partial x_{\mu}$。关于张量场的微积分，由于我们这个课程并不需要很多，我们就不再深入了。

7 最小作用量原理

¶ 前一节实际上只是涉及了狭义相对论时空的几何学，现在我们来开始讨论狭义相对论时空下的动力学。在经典力学的范畴中，如果一个力学体系的位型可以完全由一系列变量：$q_1, q_2, \cdots, q_f$ 来唯一地描述，我们称这一组变量为该力学体系的**广义坐标**；称 f 为该力学体系的**自由度数目**。[8] 一般来说，如果给出了一个力学体系在某个时刻的广义坐标，同时，我们又知道这些广义坐标对于“时间”的微商，那么这个力学体系在任何时刻的位型就可以完全确定了。为了简化记号，我们将用：$q = (q_1, q_2, \cdots, q_f)$ 来统一标记一个力学体系的一系列广义坐标。各个广义坐标对于时间的微商：$\dot{q}_1, \dot{q}_2, \cdots, \dot{q}_f$ 也被称为**广义速度**。[9]

[8] 我们这里假定系统是完整约束系统。非完整约束的系统将在随后讨论。

[9] 这里需要特别强调指出，本节的广义坐标以及广义速度的下标并不代表洛伦兹四矢量的下标。事实上，系统的广义坐标完全可以不是直角坐标，因此它们原则上并不一定构成洛伦兹矢量 (或张量)。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

一个力学体系的所有广义坐标构成一个“空间”，我们称之为该力学体系的位型空间。这个空间的维数就等于其独立广义坐标的个数。如果仅仅考虑完整约束的情形，也就是系统的自由度数目 f。需要特别强调的一点是，力学体系的位型空间一般都不是一个平直的线性矢量空间，而且它的拓扑结构也往往不是平庸的。一个简单的例子就是在球面上面运动的质点，它的位型空间是一个二维空间。但是这个二维空间不是一个平直的二维欧几里德空间，而是一个二维球面。在数学上说，这种“空间”往往被称为微分流形(differentiable manifold)。[10] 所以一个在二维球面上运动的质点的位型空间是一个具有球面拓扑结构的、二维微分流形。因此，在比较新的教科书中都开始改称“位型流形”(configuration manifold)。这实际上是更精确的称呼，不过我们仍然将沿用了比较老的称呼。

一个力学体系的广义速度 $\dot{q}$ 实际上也构成一个“空间”。最为重要的一点是应当注意到，这个空间原则上是与该力学体系的位型空间不同的空间。例如，在上面提到的在一个二维球面上运动的粒子，它的位型空间是一个二维球面，它的广义速度对应的空间实际上是在球面上每一点的切平面，而不是原先广义坐标的流形 (二维球面)。这两种东西结合起来构成的一个数学结构被称为力学体系位型空间 (流形) 的切丛(tangent bundle)。当然，如果我们仅仅考虑一个在三维平直空间的粒子，它的广义坐标 (如果取为其直角坐标的话) 所在位型空间和其广义速度的空间完全一样，都是三维的欧几里德空间。

¶现在我们将叙述经典力学中最为重要的力学原理。经典力学体系的运动规律可以完全概括在下面这个无限重要的原理中，这就是所谓的最小作用量原理，又称为哈密顿原理。它在经典力学乃至整个物理学中的重要地位怎么强调都不为过。

定理 2.1 *力学体系具有一个与其运动相关的物理量称为作用量 S，它是一个洛伦兹标量。如果一个力学体系在给定的时刻 t_1 和 t_2 分别由给定的广义坐标 $q^{(1)}$ 和 $q^{(2)}$ 描写；[11] 那么该力学体系的作用量 S 可以表达为联结这两个位型之间的各种可能轨道的泛函：*

$$S = \int_{t_1}^{t_2} L(q, \dot{q}, t) dt \ . \tag{2.14}$$

这里的函数 $L(q, \dot{q}, t)$ 称为体系的拉格朗日量，该力学体系在时刻 t_1 和 t_2 之间联结广义坐标 $q^{(1)}$ 和 $q^{(2)}$ 的真实运动轨道是使得体系的作用量 S 取极小值的轨道。[12]

[10] 我们这里不想严格地去定义流形，只是给一个大概的概念。如果要求数学上严格的定义，一个微分流形必须满足一系列的基本公理。

[11] 这里我们用 q 和 $\dot{q}$ 来代表力学体系所有的广义坐标和广义速度。如果需要写出每一个具体的自由度，我们将用 q_i 和 $\dot{q}_i$ 来表示，其中 $i = 1, 2, \cdots, f$，f 是力学体系的自由度 (我们这里首先讨论的是完整约束的力

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

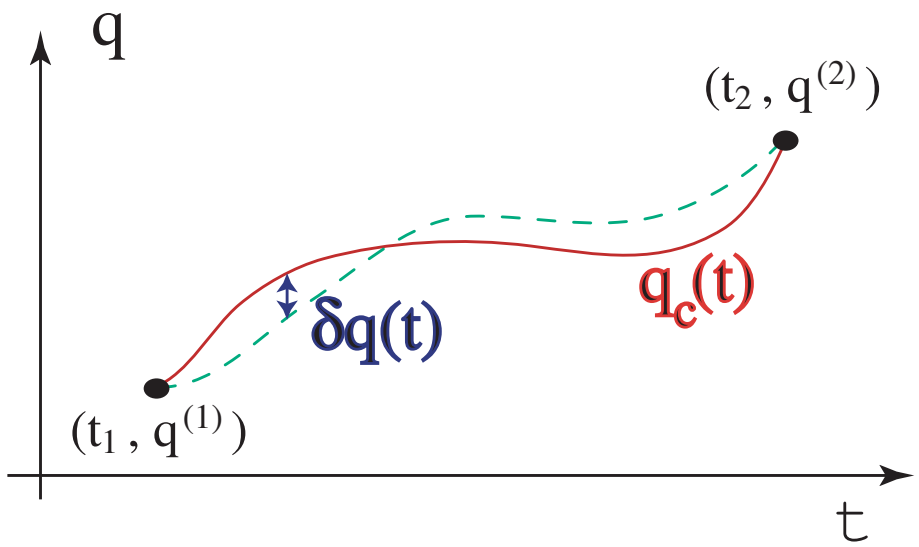

图 2.1: 最小作用量原理示意图。对于给定系统在位型空间的起始点和终止点，系统真实的 (也就是运动方程所描述的) 轨道 $q_c(t)$(图中红色轨道) 一定使得作用量取极小值。也就是说，如果我们考虑真实轨道 $q_c(t)$ 附近的一个无穷小变分 $\delta q(t)$，作用量对于轨道的一级变分为零。

最小作用量原理是分析力学 (乃至整个理论物理) 最重要的原理之一。因此，我们有必要将它的含义更为详细地阐述一下，参考图 2.1。这个原理实际包含了以下三重含义：

☞ 第一，对于任何一个力学体系都可以写出它的一个作用量 S，它是一个洛伦兹标量。这个条件实际上是与后面提到的力学体系的真实运动一定使得作用量取极小相一致的。经典力学中，力学体系的真实运动是唯一的 (决定论)。因此，在一个参照系中得到的力学体系的运动轨道，变换到另一个参照系中应当也是真实的运动轨道。要保证这一点的最简单的选择就是作用量在洛伦兹变换下是一个不变的标量。这个假设使得由最小作用量原理推出的经典力学方程自动地与狭义相对论兼容。

☞ 第二，在初始和终止时刻 t_1，t_2 和位型空间中的位置 $q^{(1)}$ 和 $q^{(2)}$ 给定的情形下，体系的作用量是系统各种可能轨道 $q(t)$ 的泛函，它是系统的拉格朗日量对时间的积分。由于时间在狭义相对论中并不是洛伦兹标量，因此一个力学体系的拉格朗日量本身也不是洛伦兹标量。但是，在非相对论极限下，时间与空间分离，这时系统的作用量和拉格朗日量都是 (三维意义下的) 标量。值得指出的是，系统的拉格朗日量只依赖于广义坐标和广义速度，不依赖于广义坐标的更高阶的时间微商，这一点实

学体系)。

[12] 事实上，仅仅就导出运动方程而言，我们只需要它是一个极值点。但是为了保证至少在自由粒子情形下的质量为正值，这才要求该极值是极小值而不是极大值。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

际上是沿袭了牛顿力学的思想：即给定力学体系的初始位置、初始速度就足以确定一个力学体系以后的运动。

☞ 第三，在起始点和终止点都固定的情况下，系统真实的运动轨道 $q_c(t)$，如果我们将其与其他所有可能的轨道相比，一定是使得系统的作用量 S 取极小值的轨道 (参考图 2.1)。下面变分法的计算和讨论告诉我们，这个轨道一般是唯一的，这正是经典力学决定论的体现。

¶ 从上述最小作用量原理出发，利用变分法就可以求出力学体系的运动方程。为此我们将力学体系的真实运动轨道记为 $q_c(t)$。显然，$q_c(t_1)=q^{(1)}$，$q_c(t_2)=q^{(2)}$。我们考虑力学体系的一个对于真实轨道的、假想的微小偏离：$q(t)=q_c(t)+\delta q(t)$。由于该力学体系的初始和终止位型已经确定，我们考虑的微小偏离满足：$\delta q(t_1)=\delta q(t_2)=0$。力学体系的轨道对于给定的真实轨道的微小偏离 $\delta q(t)$ 在数学上称为对真实轨道的变分。它不同于微分之处在于它本身是时间的任意函数。

如果系统对于其真实轨道的变分 $\delta q(t)$ 是一个 (一阶) 无穷小；那么最小作用量原理要求，这种变分所带来的系统的作用量的 (一阶) 变化必定为零，也就是说，真实轨道对于系统的轨道的一阶变分来说是极值点 (extremum)：

$$\delta S \equiv \int_{t_1}^{t_2} L(q_c+\delta q, \dot{q}_c+\delta\dot{q}, t)dt - \int_{t_1}^{t_2} L(q_c,\dot{q}_c,t)dt = 0\,. \tag{2.15}$$

当 δq 为无穷小量，我们有：[13]

$$\delta S = \int_{t_1}^{t_2} dt \left(\frac{\partial L}{\partial \dot{q}_i}\delta\dot{q}_i + \frac{\partial L}{\partial q_i}\delta q_i\right)\,,$$

这里拉格朗日量被看成是 q 和 $\dot{q}$ 的函数，因此上式中的偏微分也应当在此意义下来理解。同时，由于我们考虑的是系统在其真实轨道附近的变分，因此上面公式中的偏微商应当取其真实轨道 $q_c(t)$ 处的值。现在注意到：

$$\frac{\partial L}{\partial \dot{q}_i}\delta\dot{q}_i = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\delta q_i\right) - \left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\right]\delta q_i\,.$$

带入后第一项的完全微分变为在边界点的函数值之差：

$$\delta S = \frac{\partial L}{\partial \dot{q}_i}\delta q_i\Big|_{t_1}^{t_2} + \int_{t_1}^{t_2} dt \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right)\delta q_i = 0\,. \tag{2.16}$$

在端点处由于我们要求：$\delta q(t_1)=\delta q(t_2)=0$，因此上式中第一项的贡献为零。由于第二项必须对于任意的 $\delta q_i(t)$ 都等于零，又因为各个 $\delta q_i(t)$ 是完全独立的变分，唯一的可能是

[13] 这里我们运用了爱因斯坦求和规则：即所有重复的指标隐含着求和。所以，下面这个公式中的指标 i 要对系统所有自由度数求和。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

上式括号中的量 (对每一个 i) 都恒等于零，[14] 于是我们就得到了力学体系的真实运动的轨道所满足的方程：

欧拉 -拉格朗日方程 (Euler-Lagrange Equation)

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = 0 \,, \qquad i = 1, 2, \cdots, f \,. \tag{2.17}$$

这就是著名的欧拉 - 拉格朗日方程 (Euler-Lagrange equation)。它是一个经典力学体系的广义坐标所满足的二阶常微分方程。因此，当初始的广义坐标和广义速度给定后，经典力学体系的运动就完全由其运动方程 (欧拉 -拉格朗日方程) 确定了。

最小作用量原理和变分法告诉我们，一旦给定了一个力学体系的拉格朗日量 (作用量)，它的经典运动方程就可以很方便地得到。所以，我们可以说经典力学体系的性质完全由其拉格朗日量 (作用量) 所确定。正是在此意义下我们说一个经典力学体系的拉格朗日量 (作用量) 集成了 该力学体系的所有重要力学信息。

现在我们定义该力学体系中与某个广义坐标 q_i 共轭的广义动量，它又称为正则动量：

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \,. \tag{2.18}$$

注意，与一个广义坐标 q_i 共轭的广义动量 p_i 恰好就是在变分计算时边界项中 δq_i 前面的系数，参见公式 (2.16)。利用定义的广义动量，力学体系的运动方程也可以写成：

$$\frac{dp_i}{dt} = \frac{\partial L}{\partial q_i} \,. \tag{2.19}$$

由于这个式子与通常的牛顿力学方程的类似性，方程的右边，即 $\partial L/\partial q_i$，又被称为相应的广义力。

¶ 在经典力学中，所有的物理都体现在体系的经典运动方程中。也就是说，即使系统的拉格朗日量或作用量改变了，只要体系的经典运动方程仍然保持不变，经典的物理就没有改变。[15] 这个事实意味着一个经典力学体系的拉格朗日量本身的数值并没有什么绝对的意义，只是由它导出的经典运动方程 –即欧拉 -拉格朗日方程 (2.17)–才有意义。例如，我们完全可以在拉格朗日量上加上一个给定 函数对于时间的全微分：

$$L'(q, \dot{q}, t) = L(q, \dot{q}, t) + \frac{df(q,t)}{dt} \,, \tag{2.20}$$

显然，对于 L' 来说，其作用量比原先的作用量多出一些仅仅依赖于初始和终止时刻和坐标的项。但是由于我们考虑的变分 $\delta q(t)$ 在端点处为零，因此并不会改变最小作用量原理

[14] 这个事实在数学上被称为变分学基本引理。
[15] 我们之所以强调这点是因为它在量子体系中并不一定正确。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

所确定的力学体系的运动方程。也就是说，L' 与 L 所确定的运动方程是完全一样的，两者在经典力学范畴内是完全等价的。这个事实我们在后面 (见第 9节) 还会多次用到。

8 相对论性自由粒子的作用量

¶ 前两节我们讨论了狭义相对论的时空观，同时又讨论了经典力学中具有普遍性的最小作用量原理。现在我们将这两者结合起来，简单讨论一下相对论性的力学问题。从最小作用量原理出发来讨论经典力学的最大优势是，将最小作用量原理与洛伦兹不变性相结合，再加上其他的一些对称性考虑，我们几乎可以唯一地确定出一个自由粒子的作用量 (拉格朗日量)。

在相对论时空观中，一个粒子运动的轨道被称为世界线。一个世界线可以用 (四个) 参数方程：$x^\mu = x^\mu(\tau)$ 来给出。[16] 这里 x^μ 是粒子坐标的四矢量；τ 是描写世界线上的一个参数 (一个洛伦兹标量)。[17] 对称性的考虑告诉我们，它只可能具有下列形式：

$$S = -mc\int ds = -mc\int d\tau \left(\frac{dx^\mu}{d\tau}\frac{dx_\mu}{d\tau}\right)^{1/2}. \tag{2.21}$$

其中 $ds = (dx_\mu dx^\mu)^{1/2}$ 为该粒子的不变间隔。它也可以写成对世界线参数 τ 的微分形式。另外一种说法是，自由粒子的作用量正比于它的世界线的“长度”。[18] 一个自由粒子的作用量必定具有这种形式的原因是，时空平移不变性（自由粒子）要求拉格朗日量不能依赖于时空坐标；洛伦兹标量表明几乎我们必须选择 ds，因为这是一个自由粒子唯一具有的标量。前面的因子 mc 完全是为了使得作用量具有大家约定俗成的量纲。[19] 这里 c 代表真空中的光速 (洛伦兹标量)，它是一个与狭义相对论时空观相关的常数；参数 m 必须也是一个洛伦兹标量，它是一个与所考虑的粒子的性质有关的物理量。就其力学性质而言，一个自由的相对论性的点粒子所具有的标量物理量一定与它的静止质量有关。我们马上会说明它实际上就是该粒子的静止质量，它的大小体现了该粒子的惯性的大小。我们随后还会说明它一定是大于零的。

¶ 下面我们来考察作用量 (2.21) 所给出的相对论性自由粒子的运动方程。这个运动方程可以由两种方式来获得。一种是采用三维的形式，首先写出拉格朗日量，然后利用普遍的欧拉 -拉格朗日方程 (2.17) 直接写出；另一种方法是采用四维协变的形式，直接对作用量 (2.21) 取变分。为了让大家能够熟悉这两种方法，下面我们将分别利用这两种方法来推导自由粒子的运动方程。

[16] 如果你愿意，你可以从 $x^0 = ct = x^0(\tau)$ 中将 τ 反解出来代入到 $x^i(\tau)$ 中，就给出粒子的三维坐标作为时间的轨道方程。但是在相对论中，实际上运用四维参数形式的世界线方程更为方便，因为在相对论的情形下时间并不是洛伦兹标量。

[17] 这里 τ 可以取为粒子的“原时”：$d\tau = ds/c$，但也可以是任何其他的世界线参量。

[18] 这里的长度是在闵氏空间中以 $\eta_{\mu\nu}$ 为度规的长度。

[19] 具体来说，作用量具有能量乘以时间的量纲；因此拉格朗日量具有能量的量纲。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

首先利用三维形式的拉格朗日量来讨论。为此我们注意到：$ds = \sqrt{c^2dt^2 - d\mathbf{x}^2} = cdt\sqrt{1-\mathbf{v}^2/c^2}$，因此我们得到一个相对论性自由粒子的拉格朗日量为：

$$L(\mathbf{v}) = -mc^2\sqrt{1-\frac{\mathbf{v}^2}{c^2}}\,. \tag{2.22}$$

我们立刻得到粒子的广义动量为：

$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}} = \frac{m\mathbf{v}}{\sqrt{1-\frac{\mathbf{v}^2}{c^2}}}\,. \tag{2.23}$$

这正是大家所期待的一个相对论性粒子的动量。于是，粒子的运动方程为：

$$\frac{d\mathbf{p}}{dt} = 0\,. \tag{2.24}$$

它的的解实际上就是匀速直线运动。这里我们可以立刻证明，代表匀速直线运动的轨道的确使得系统的作用量取极小值，只要参数 $m > 0$。[20] 也就是说，粒子的静止质量一定是非负的。[21]

下面我们利用四维协变的形式来讨论运动方程。利用变分法，我们有：

$$\begin{aligned}\delta S &= -mc\int \delta\sqrt{dx_\mu dx^\mu} \\ &= -mc\int \frac{(d\delta x_\mu dx^\mu)}{\sqrt{dx_\mu dx^\mu}}\,, \\ &= mc\int ds\delta x_\mu \frac{d}{ds}\frac{dx^\mu}{ds} = 0\,,\end{aligned}$$

其中我们在分部积分时已经扔掉了边界项。由于 δx_μ 是任意的变分，因此我们得到粒子的运动方程的四维协变形式为：

$$\frac{d^2x^\mu}{ds^2} = 0\,. \tag{2.25}$$

读者不难验证，这个形式实际上与 (2.24) 是完全一致的。利用四维协变形式的变分法，我们还得到一个副产品，那就是相对论性粒子的四动量。它是与四维广义坐标 x_μ 所对应的广义动量。正如第 24页提到的，这可以从分部积分的边界项中 δx^μ 前面的系数得到：

$$p^\mu = mc\frac{dx^\mu}{ds} = \left(\frac{E}{c}, \mathbf{p}\right)\,, \tag{2.26}$$

其中的空间分量就是前面得到的粒子的相对论性动量 (2.23)；它的零分量是该粒子的相对论性能量 E 除以光速：

$$E = \frac{mc^2}{\sqrt{1-\frac{\mathbf{v}^2}{c^2}}}\,. \tag{2.27}$$

[20] 如果 $m < 0$，那么运动方程给出的轨道则变成极大值了。

[21] 零质量的粒子原则是是允许的。它们永远以光速运动。当然，要利用最小作用量原理讨论零质量粒子，我们不能直接利用公式 (2.21)。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

正因为如此，粒子的四动量又被称为 能量-动量四矢量。显然，它们构成了具有确定变换规则的四矢量。[22] 一个静止质量是 m 的粒子的四动量 p^μ 与自身的内积是一个洛伦兹标量。简单的计算告诉我们，这个标量就是：$p^\mu p_\mu = (E/c)^2 - \mathbf{p}^2 = m^2c^2$。

¶ 以上的讨论限于一个自由粒子的情形。下面我们考虑多个自由粒子组成的系统的作用量。所谓多个自由粒子，乃是指这些粒子之间的运动完全是独立的。也就是说，每一个粒子的运动方程中应当完全不包含其他粒子的广义坐标或广义速度。要实现这一点最简单的选择就是假定无相互作用的粒子体系的作用量是各个粒子作用量的简单相加：

$$S = -\sum_i m_i c \int \sqrt{\eta_{\mu\nu} dx_i^\mu dx_i^\nu} \,. \tag{2.28}$$

这样一来，每个粒子的运动方程都是我们前面得到的自由粒子运动方程的形式。

9 粒子与外场的相互作用

¶ 现在我们来考虑一个相对论性的粒子与外场的相互作用。这时，我们需要区分外场在洛伦兹变换下的性质。如果外场本身是一个洛伦兹标量 (这样的外场被称为标量场)，那么一个相对论性的粒子与它的相互作用可以写成：

$$S = -mc \int ds e^{\Phi(x)} \,, \tag{2.29}$$

这里积分是沿着粒子的世界线 $x^\mu(s)$，$\Phi(x)$ 是一个外加的、无量纲的标量场。我们现在运用变分法，就可以得到粒子的运动方程：

$$\frac{d^2x^\mu}{ds^2} + \frac{\partial\Phi}{\partial x^\nu}\frac{dx^\nu}{ds}\frac{dx^\mu}{ds} = \frac{\partial\Phi(x)}{\partial x_\mu} \,. \tag{2.30}$$

显然，如果外场 $\Phi(x)$ 恒等于零我们回到自由粒子的情况。在一般的标量场情形下，我们得到的系统的拉格朗日量为：

$$\boxed{L = -mc^2\sqrt{1 - \frac{\mathbf{v}^2}{c^2}} e^{\Phi(\mathbf{x},t)} \,.} \tag{2.31}$$

从这个拉格朗日量出发，我们也可以得到三维形式的粒子运动方程。

[22] 由于也是四矢量，因此能量动量四矢量的变换规则与时间坐标四矢量的变换规则完全相同，也是按照我们所熟悉的洛伦兹变换 (2.9) 来变。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

¶如果我们加上的外场不是标量场，而是一个四矢量场，那么为了得到洛伦兹标量的作用量，我们可以写下如下的形式：

$$S = -mc\int ds - \frac{e}{c}\int A_\mu(x)dx^\mu \ , \tag{2.32}$$

这里 e 代表了粒子与矢量场的耦合强度，$A_\mu(x)$ 是外加的一个矢量场。我们知道的电磁场与带电粒子的作用量就具有这种形式。[23] 因此，我们姑且把常数 e 称为该粒子所带的电量。我们要求它是一个洛伦兹标量，但是对它的符号则没有限制。[24] 代表矢量场 (电磁场) 的四矢量 $A_\mu(x)$ 被称为电磁场的四矢量势。由于四矢量势是一个四矢量，因此它可以用时空分量的形式写成：

$$A^\mu(x) = (\Phi(x), \mathbf{A}(x)) \ , \tag{2.33}$$

其中的时间分量 $\Phi(x)$ 被称为电磁场的标量势，而空间分量 $\mathbf{A}(x)$ 被称为电磁场的矢量势。如果把作用量表达式 (2.32) 中的积分用对时间的积分表达，我们就可以写出电磁场中一个相对论性粒子的拉格朗日量：

$$L = -mc^2\sqrt{1-\frac{\mathbf{v}^2}{c^2}} + \frac{e}{c}\mathbf{v}\cdot\mathbf{A} - e\Phi \ . \tag{2.34}$$

它与自由粒子拉格朗日量的区别就在于加上了与外电磁场的相互作用项。得到了带电粒子的拉格朗日量，我们立刻可以写出粒子的广义动量 (又被称为正则动量)$\mathbf{P}$：

$$\mathbf{P} = \frac{m\mathbf{v}}{\sqrt{1-\mathbf{v}^2/c^2}} + \frac{e}{c}\mathbf{A} = \mathbf{p} + \frac{e}{c}\mathbf{A} \ . \tag{2.35}$$

需要注意的是，系统的正则动量 $\mathbf{P}$ 并不是粒子的相对论性动量 $\mathbf{p}$，又称为机械动量。两者之间差一个与矢势成正比的项。

在外电磁场中，一个相对论性粒子的运动方程可以写成三维形式或者四维协变形式。它的推导与推导自由粒子的方法没有本质区别，只不过更为复杂一些。我们这里直接给出结果。四维协变形式的运动方程具有如下的形式：

$$mc\frac{d^2x_\mu}{ds^2} = \frac{e}{c}F_{\mu\nu}\frac{dx^\nu}{ds} \ . \tag{2.36}$$

[23] 很不幸的是，电磁相互作用的作用量的具体形式依赖于单位制的选取。这里给出的单位制是所谓的高斯单位制。关于不同电磁单位制之间的转换规则，可以参考 [7] 的附录。

[24] 相应于参数 e 的符号，我们称该粒子带正电或带负电。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

其中我们引入了电磁场场强的二阶（反对称）张量 $F_{\mu\nu}$，它的定义为：

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \ . \tag{2.37}$$

这个四维协变形式的运动方程很容易表达成三维的形式。另外一种得到三维形式的运动方程的方法是直接从粒子的拉格朗日量 (2.34) 出发写出其欧拉 -拉格朗日方程。经过一些矢量微分的演算，我们得到：

$$\frac{d\mathbf{p}}{dt} = e\mathbf{E} + \frac{e}{c}\mathbf{v}\times\mathbf{B} \ , \tag{2.38}$$

大家应当可以认出，公式 (2.38) 就是一个 (相对论性的) 带电粒子在电磁场中所受到的洛伦兹力的公式。这里 **p** 是粒子的相对论性动量，而 **E** 和 **B** 分别称为电磁场的电场强度和磁感应强度，它们与电磁势的具体关系是：

$$\mathbf{E} = -\nabla\Phi - \frac{1}{c}\frac{\partial\mathbf{A}}{\partial t} \ , \ \mathbf{B} = \nabla\times\mathbf{A} \ . \tag{2.39}$$

这个电磁场与电磁势之间的关系实际上就是四维形式 (2.37) 的三维对应物。[25]

¶ 方程 (2.38) 告诉我们，影响粒子运动方程的并不是电磁势 A_μ 本身，而是由此派生出来的电磁场场强张量 $F_{\mu\nu}$，或者等价地用三维形式表达的电场强度 **E** 和磁感应强度 **B**。因此，在经典力学的水平上，$A_\mu(x)$ 并不具有直接的可测量的物理效果，因为带电粒子直接感受到 (测量到) 的是场强张量 (电场强度和磁感应强度)。这一点的更为明确的表述就是电磁相互作用具有所谓的规范不变性。这种对称性在分析力学的体系中体现得特别直接而简洁。我们记得前面强调过 (见第 7节末尾)，系统的拉格朗日量中可以加上一个任意函数对时间的全微分而不改变运动方程。因此我们可以将粒子与电磁场相互作用的作用量变为：

$$S_{int} = -\frac{e}{c}\int\left(A_\mu(x) + \frac{\partial\Lambda}{\partial x^\mu}\right)dx^\mu \ , \tag{2.40}$$

其中第二项是一个全微分，因此积分后变为边界项从而对粒子的运动方程没有任何影响。因此，从粒子的运动方程来看，这个作用量与我们前面给出的作用量完全等价。也就是说，在外场 $A_\mu(x)$ 在如下的变换下：

$$A_\mu(x) \to A_\mu(x) + \partial_\mu\Lambda(x) \ , \tag{2.41}$$

粒子的运动方程并不会改变，其中 $\Lambda(x)$ 是任意一个给定函数。这种对称性被称为电磁场的规范对称性。上面的电磁势的变换被称为规范变换。 将变换 (2.41) 代入到场强张量的定义 (2.37) 之中很容易验证场强张量 $F_{\mu\nu}$ 在规范变换下是不变的。

[25] 具体来说我们有：$F_{0i} = \mathbf{E}_i$，$F_{ij} = \epsilon_{ijk}\mathbf{B}_k$。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

¶通过以上简单的讨论我们看到，相对论性的粒子在标量场和电磁场 (矢量场) 中的运动方程可以十分简洁地从基本原理 (最小作用量原理、作用量的洛伦兹不变性等) 导出。事实上，不光电磁场中的带电粒子的运动可以这样得到，就连电磁场本身所满足的运动方程 (就是著名的麦克斯韦方程组) 也可以从这些基本原理导出。不过，这些内容应当不是我们理论力学课程的主题，它们会在电动力学课程中详细地加以讨论。我们这里顺便提一下，如果我们讨论一个相对论性粒子与一个二阶张量场耦合，同时加上张量场本身的运动，我们实际上就会得到广义相对论 (这也是其他课程的内容)。我们这里只是希望指出：最小作用量原理是一个十分基本的原理。将它与其他重要的物理原理结合，我们就可以得到各种经典力学体系的运动规律。

10 非相对论极限

¶虽然我们的课程是从相对论性的时空观出发的，但实际上我们的主体内容涉及的是粒子在低速运动下的动力学问题，这被称为非相对论极限。[26] 粗略地来说，在上一节得到的相对论性的公式中，如果我们令 $\mathbf{v}^2/c^2 \to 0$，我们就可以得到非相对论极限下的相应公式。在我们这样做之前，先从时空观的角度来分析一下非相对论极限的特点是有益的。

在非相对论极限下，闵氏时空变为传统的伽利略时空。在伽利略时空中，时间具有绝对的意义 (同时的绝对性)，换句话说，时间与空间发生了分离。相应的，所有在洛伦兹变换下具有确定变换规则的张量的时间分量与空间分量也都会发生分离。以四矢量 V_μ 为例，它的时间分量 V_0 在三维空间下变为三维的标量，而空间分量 $V_i, i = 1, 2, 3$ 则变为三维空间的三矢量。由于四矢量的三个空间分量完全是相同的并且与时间分量分离，因此矢量的上标和下标也就没有区分的必要了。为了方便起见，我们将时间指标写为下标。对空间部分的指标，我们也将统一把它们写为下标。类似的讨论可以扩展到更高阶的张量。例如一个二级张量 $D_{\mu\nu}$ 在非相对论极限下可以分解为：D_{00}，这是一个三维的标量；D_{0i}(或者 D_{i0})，这是一个三维空间的矢量；D_{ij}，这是一个三维空间的二阶张量。

洛伦兹变换在非相对论极限下发生退化。也就是说，时间根本不变，而一个三维矢量的三个分量按照三维空间旋转的矩阵来变：

$$x_i' = A_{ij} x_j \ . \tag{2.42}$$

这里 x_i 和 x_i' 分别表示变换前后的一个三维矢量；A_{ij} 是一个三维转动所对应的正交矩阵 (用张量的语言来说就是二级张量)。它满足：$A_{ij}A_{kj} = \delta_{ik}$。由此变换规则不难得到更高阶三维张量的变换规则。对于一般的三维转动矩阵，我们在后面讨论刚体的运动 (见第 五章) 时会更多更详细地涉及。

¶现在我们来考虑力学体系的拉氏量。在非相对论极限下，它本身是伽利略意义下的标量，即三维空间的标量。例如，在非相对论极限下，我们得到一个自由粒子的拉格朗日

[26] 相对论性的粒子的动力学问题将主要在电动力学课程中进行处理。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

量可以取为：

$$L = \frac{1}{2}m\mathbf{v}^2 \ . \tag{2.43}$$

这里我们略去了常数 $-mc^2$。这个拉格朗日量也可以直接从空间平移不变 (要求拉格朗日量不依赖于空间坐标) 和各向同性 (要求拉格朗日量不依赖速度的方向) 直接得到。它说明一个非相对论性的自由粒子的拉格朗日量可以取为其**动能**。容易验证，要使得相应的作用量取极小值而不是极大值，我们要求粒子的惯性质量参数 $m > 0$。

如果一个粒子与标量外场相互作用，其相互作用能量我们记为 $V(\mathbf{x}, t)$，那么我们可以在公式 (2.29) 中取：$\Phi(\mathbf{x}, t) = V(\mathbf{x}, t)/(mc^2)$。由于在非相对论极限下，一切能量都远远小于粒子的静止能量 mc^2，因此：

$$e^{\Phi(\mathbf{x},t)} = e^{V(\mathbf{x},t)/mc^2} \sim 1 + \frac{V(\mathbf{x}, t)}{mc^2} \ ,$$

我们因此得到：

$$L = \frac{1}{2}m\mathbf{v}^2 - V(\mathbf{x}, t) \ . \tag{2.44}$$

这里 $V(\mathbf{x}, t)$ 被称为**势能** (具有能量量纲)。也就是说，我们可以取一个外场中的粒子的拉格朗日量为它的动能减去势能。[27] 它所对应的运动方程为：

$$\frac{d}{dt}(m\mathbf{v}) = -\frac{\partial V(\mathbf{x}, t)}{\partial \mathbf{x}} \ . \tag{2.45}$$

这就是一个质点在外势场 $V(\mathbf{x}, t)$ 中的**牛顿方程**。

对于一个与矢量场 (电磁场) 耦合的非相对论性粒子，它的拉格郎日量可以取为：

$$L = \frac{1}{2}m\mathbf{v}^2 + \frac{e}{c}\mathbf{v} \cdot \mathbf{A}(\mathbf{x}, t) - e\Phi(\mathbf{x}, t) \ . \tag{2.46}$$

这个拉格郎日量对应的运动方程就是一个非相对论性的粒子在洛伦兹力作用下的牛顿方程。特别需要指出的是，如果外磁场为零，那么一个粒子与矢量场的标量势相互作用的拉格郎日量与一个粒子与标量场相互作用的拉格郎日量在形式上没有区别。这一点并不奇怪，因为正如本节开始所说的，在非相对论极限下一个四矢量的时间分量 (标量势) 的确变成一个三维空间的标量。

¶ 在非相对论极限下，N 个具有相互作用的粒子组的拉格郎日量可以取为：

$$L = \sum_i \frac{1}{2}m_i\mathbf{v}_i^2 - V(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N, t) \ . \tag{2.47}$$

[27] 从这个讨论我们看到，力学中所谓的势能的概念仅仅在非相对论极限下才有意义。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

其中的第一项就是各个粒子的动能之和；第二项的函数 $V(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N, t)$ 被称为这一组粒子的势能。由于取了非相对论极限，因此我们假定势能中的所有与粒子速度相关的项都可以略去。如果系统具有时间平移不变性，那么势能不能显含时间。这样的系统被称为保守系统。保守系统的一个重要性质就是其能量守恒，这一点我们下一节会更为细致地讨论。一个保守体系的运动方程可以写成：

$$m_i \frac{d\mathbf{v}_i}{dt} = -\frac{\partial V(\mathbf{x}_1, \cdots, \mathbf{x}_N)}{\partial \mathbf{x}_i} , \tag{2.48}$$

¶ 拉格郎日量 (2.47) 是利用直角坐标表达的。有的时候我们需要更为普遍的广义坐标。由于广义坐标与直角坐标之间的函数关系仅仅依赖于各个坐标而不依赖于速度，而且我们假定势能中也不包含速度，因此经过这个变换以后拉格郎日量一定可以写为：

$$L = \sum_{ij} \frac{1}{2} a_{ij}(q) \dot{q}_i \dot{q}_j - V(q_1, q_2, \cdots, q_f) . \tag{2.49}$$

这里的 $a_{ij}(q)$ 可以依赖于所有的广义坐标 q_i，但是它一定是一个对称的、正定的矩阵。势能 $V(q_1, \cdots, q_f)$ 则仅仅依赖于所有的广义坐标。

¶ 现在我们总结一下我们目前已经得到的结论。从最小作用量原理和基本的洛伦兹不变性出发，我们得到了相对论性粒子的拉格朗日量，这包括自由粒子的情形和与外场相互作用的情形。我们简单讨论了相对论性粒子的运动方程。在非相对论极限下，我们可以得到非相对论性系统的拉格朗日量。我们看到，这时力学系统的运动方程与我们通常熟悉的牛顿方程是完全一致的。这就是我们在前一章所提到的：在纯力学的、非相对论体系中，拉格朗日力学提供了一种与牛顿力学完全等价的描述。但是我们同时也看到，以最小作用量原理为出发点的拉格朗日力学具有更加广泛的适用性。我们这里简单地讨论了相对论性的力学。这些结果完全是与牛顿力学不同的。我们还说明了，拉格朗日力学可以直接处理（甚至是相对论性的）电磁问题。这也是传统的牛顿力学所无法胜任的。因此，最小作用量原理应当被看成是超出牛顿力学之外的一条更为普遍的原理，而非相对论性的牛顿力学仅仅是它的一部分推论。当然，作为我们这个课程本身，非相对论性的纯力学体系是我们讨论的主要对象。因此，除了在个别的章节我们会再次考虑相对论性的问题之外，我们这个课程以后涉及的绝大多数都将是非相对论性的纯力学问题。[28] 除非特殊声明，我们都假设无须再考虑相对论效应。

11 对称性

¶ 从微分方程的角度来看，一个具有 f 个自由度的力学系统的拉格朗日方程一般具有 $2f$ 个独立的初积分 (例如，这 $2f$ 个常数可以取为系统的初始位置和速度)。也就是说，这

[28] 一个例外是第 18节中讨论的行星近日点的进动问题。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

些初积分都是力学系统的守恒量。但是并不是每一个初积分都具有重要的物理意义。力学系统的初积分中只有少数几个是与时空的对称性密切联系着的，这些守恒量也是物理上最为重要的。这一节中，我们将主要讨论时间平移、空间平移、空间转动和尺度变换的对称性。另一方面，空间反射和时间反演这两种分立对称性也在力学中也扮演重要的角色，我们也一并加以讨论。

11.1 时间平移

¶ 一个力学系统，如果它的拉格郎日量不显含时间，这时系统的运动方程在任何时刻都是相同的。我们称该力学系统具有时间平移不变性。这时，系统拉格郎日量对于时间的变化率为：

$$\frac{dL}{dt} = \frac{\partial L}{\partial q_i}\dot{q}_i + \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i \,. \tag{2.50}$$

同时，让我们考虑 $p_i\dot{q}_i$ 的时间变化率，其中 p_i 是与 q_i 共轭的广义动量：

$$\frac{d(p_i\dot{q}_i)}{dt} = \frac{dp_i}{dt}\dot{q}_i + p_i\frac{d\dot{q}_i}{dt} \,. \tag{2.51}$$

按照拉格朗日方程，这两个表达式的右边实际上是完全相同的。因此，如果我们定义：

$$E = p_i\dot{q}_i - L \,, \tag{2.52}$$

那么我们发现：E 是力学系统的守恒量。这个守恒量是直接与系统时间平移不变联系起来的。它就是这个系统的能量。例如，对于前面提到的保守系统：

$$E = \sum_i \frac{1}{2}m_i\mathbf{v}_i^2 + V(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N) \,. \tag{2.53}$$

第一项是各个粒子的动能和，第二项则是粒子之间相互作用的势能。也就是说，系统的总能量可以写成动能与势能之和。这个结论直接来自于动能是速度的二次齐次函数的事实。

11.2 空间平移

¶ 下面考察直角坐标中的空间平移不变性。这个对称性要求在变换：

$$\mathbf{x}_i \to \mathbf{x}_i + \mathbf{x}_0 \,, \tag{2.54}$$

下系统的拉格郎日量不变。因此我们得到：

$$\sum_i \frac{\partial L}{\partial \mathbf{x}_i} = 0 \,.$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

利用拉格朗日方程，我们得到：

$$\frac{d\mathbf{P}}{dt}=0\ ,\quad \mathbf{P}\equiv\sum_i\frac{\partial L}{\partial\mathbf{v}_i}=\sum_i\mathbf{p}_i\ . \tag{2.55}$$

也就是说，空间平移不变性要求系统的**总动量守恒**。

对于一个总动量守恒的体系，我们总是可以利用伽利略不变性来进行一个参照系变换，使得在新的参照系中系统的总动量恒等于零。假定在某个参照系中系统的总动量为 $\mathbf{P}$，它是一个守恒的矢量。我们选取一个相当于这个参照系以速度 $\mathbf{V}=\mathbf{P}/\sum_i m_i$ 运动的新的参照系。读者很容易验证，在新的参照系中，原来力学体系的总动量等于零。这样的一个参照系被称为该力学体系的**质心系**。在原参照系中：

$$\mathbf{R}=\frac{\sum_i m_i\mathbf{x}_i}{\sum_i m_i}\ , \tag{2.56}$$

被称为体系的质心坐标。显然利用质心坐标可以将体系的总动量写为：$\mathbf{P}=\sum_i m_i\dot{\mathbf{R}}$。换句话说，系统的总动量好像是所有质量都集中在其质心，速度是其质心坐标变化率的一个质点的动量。

¶上面关于空间平移不变性的讨论是在直角坐标中进行的。我们看到，如果系统的拉格朗日量在直角坐标中具有空间平移不变性，那么该力学系统的平动动量是守恒的。类似的讨论实际上可以推广到任意的广义坐标。如果一个力学系统的拉格朗日量 $L(q,\dot{q},t)$ 不显含某一个广义坐标 q_1，那么与 q_1 所对应的广义动量 $p_1=\partial L/\partial\dot{q}_1$ 一定守恒。力学系统的拉格朗日量中不出现的广义坐标被称为**循环坐标**。因此我们可以说，力学系统的循环坐标所对应的广义动量都是守恒的。

11.3 空间转动

¶最后，我们来考察直角坐标中系统的**转动不变性**。对于一个固定的原点，我们绕通过原点的一个轴进行一个无穷小的转动：$\delta\boldsymbol{\phi}$。这里的轴矢量 $\delta\boldsymbol{\phi}$ 指向右手法则所确定的转轴方向；它的大小等于转动的角度 $\delta\phi$。在这样一个无穷小的转动下，原来每个粒子的位置、速度等都相应地发生一个无穷小的变化。简单的几何考虑告诉我们：

$$\delta\mathbf{x}_i=\delta\boldsymbol{\phi}\times\mathbf{x}_i\ ,\quad \delta\mathbf{v}_i=\delta\boldsymbol{\phi}\times\mathbf{v}_i\ . \tag{2.57}$$

于是体系的拉格郎日量不变导致：

$$\frac{d\mathbf{L}}{dt}=0\ ,\quad \mathbf{L}=\sum_i\mathbf{x}_i\times\mathbf{p}_i\ . \tag{2.58}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

即体系的总角动量守恒。

¶ 一个体系的角动量一般来说依赖于坐标原点的选取。如果我们将坐标原点平移一个常矢量 $\mathbf{x}_0$，新的坐标系中所有位置矢量都相应平移：$\mathbf{x}' = \mathbf{x} + \mathbf{x}_0$，那么一个体系的总角动量变化为：

$$\mathbf{L}' = \mathbf{L} + \mathbf{x}_0 \times \mathbf{P} \,, \tag{2.59}$$

其中 $\mathbf{P}$ 是体系的总动量。一个特例是在体系的质心系中，这时系统的总动量 $\mathbf{P} = 0$，因而体系的总角动量不依赖于坐标原点的选取。

¶ 同样的，我们还可以讨论一个体系的总角动量在伽利略参照系变换下的性质。如果考虑一个相对于参照系 K 以速度 $\mathbf{v}_0$ 平动的另一个参照系 K'，这时构成系统的每个质点的平动速度有一个平移：$\mathbf{v} = \mathbf{v}' + \mathbf{v}_0$，于是系统的总角动量变为：

$$\mathbf{L} = \mathbf{L}' + \left(\sum_i m_i\right) \mathbf{R} \times \mathbf{v}_0 \,, \tag{2.60}$$

其中 $\mathbf{R}$ 是系统质心的位置。进一步如果 K' 正好是系统的质心系，那么 $(\sum_i m_i)\mathbf{v}_0$ 就是系统在 K 系中的总动量。

11.4 尺度变换

¶ 这一小节我们讨论一下在物理学中具有重要意义的尺度变换。在这种变换中，空间的坐标被乘以一个常数因子 λ_1：$\mathbf{x} \to \mathbf{x}' = \lambda_1 \mathbf{x}$。这相当于用不同的尺子来度量长度。因子 λ_1 称为该尺度变换的变换因子。类似地，我们可以将时间变换一个因子：$t \to \lambda_2 t$。

在时间和空间的尺度变换下，有一类函数具有特别重要的意义。这就是所谓的齐次函数。这类函数在尺度变换下只是乘以变换因子的某个幂次。例如，对于一个多个粒子组成的闭合系统，在上述尺度变换下，每一个粒子的速度一定乘以一个因子：λ_1/λ_2，因而系统的动能乘以一个因子：$(\lambda_1/\lambda_2)^2$。如果该系统的势能正好是各个粒子坐标的 k 次齐次函数，也就是说它满足：

$$U(\lambda_1 \mathbf{x}_1, \lambda_1 \mathbf{x}_2, \cdots, \lambda_1 \mathbf{x}_N) = \lambda_1^k U(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N) \,, \tag{2.61}$$

那么，我们可以选取适当的 $\lambda_2 = \lambda_1^{1-k/2}$，从而系统的拉格朗日量在尺度变换下正好乘以一个因子。由于拉格朗日量乘以一个常数因子并不改变系统的运动方程，因此变换以后的系统的运动方程仍然形式上与原先的系统相同。这意味着这个系统具有几何上类似的轨道，在这些轨道上运行的特征时间与轨道尺度之间的比例也是固定的。具体来说，我们有：

$$\frac{t'}{t} = \left(\frac{l'}{l}\right)^{1-k/2} \,, \tag{2.62}$$

其中 t'，l' 和 t，l 分别是两个相似轨道上的特征时间和特征尺度。这个公式的一个重要特例就是开普勒的第三定律。在太阳系的行星的周期运动中，势能恰好是坐标的齐次函数

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

($k=-1$)。因此，上面的结论就是行星周期的平方之比等于其轨道尺寸的立方之比，这就是著名的开普勒第三定律。

¶尺度变换的另外一个重要例子是所谓的维里定理。这个定理在统计力学中也会用到。我们考虑一个多粒子的力学体系并且假定体系局限在有限的空间范围内运动 (例如：太阳系中行星的运动，容器中空气分子的运动等等)。由于体系的动能是速度的二次齐次函数，我们有：

$$2T=\sum_i \mathbf{v}_i\cdot\mathbf{p}_i=\frac{d}{dt}\left(\sum_i \mathbf{p}_i\cdot\mathbf{x}_i\right)-\sum_i \mathbf{x}_i\cdot\dot{\mathbf{p}}_i\ . \tag{2.63}$$

现在我们将上式中的两边在很长的时间间隔中平均。右边的第一项是时间的全微商，因此时间平均化为上式圆括号中的量在两个时间之间的差在除以时间间隔。由于系统局限在有限的区域运动，因此这个平均在长时间极限下为零：

$$\lim_{\tau\to\infty}\frac{1}{\tau}\int_o^\tau \frac{d}{dt}\left(\sum_i \mathbf{p}_i\cdot\mathbf{x}_i\right)=0\ .$$

于是，我们得到了重要的维里定理：

$$\boxed{2\langle T\rangle=\left\langle\sum_i \mathbf{x}_i\cdot\frac{\partial U}{\partial\mathbf{x}_i}\right\rangle\ ,} \tag{2.64}$$

其中我们用 $\langle\cdots\rangle$ 来表示对力学量的长时间平均值。如果系统的势能是其坐标的 k 次齐次函数，按照著名的欧拉定理，我们有：

$$2\langle T\rangle=k\langle U\rangle\ , \tag{2.65}$$

也就是说，系统的动能的平均与势能的平均有着简单的比例关系。这个关系的两个重要特例是有心力场中的开普勒问题 ($k=-1$) 和谐振子 ($k=2$)。对于开普勒问题，动能平均是势能平均 (负的) 的一半，从而总能量也是势能平均的一半。对于谐振子，动能平均与势能平均相等。

11.5 分立对称性

¶分立对称性在物理学中也起着十分重要的作用。我们这里主要讨论空间反射 (又称为宇称变换) 和时间反演变换。

空间反射变换代表了如下的操作：

$$\mathbf{x}\to-\mathbf{x}\ , \tag{2.66}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

显然，两次空间反射变换等效于没有变换，因此物理量在空间反射反射变换下往往只能不变或者变一个符号。我们称前一类物理量为具有正的**宇称**的物理量，后一类物理量则称为具有负宇称的物理量。如果结合（三维）空间的转动下物理量的变换规则，我们可以将各种物理量按照其变换的行为分为几类：标量是在空间转动和空间反射下都不变的物理量；**赝标量**是在 空间转动下不变但在空间反射下变号的物理量。矢量是在空间转动下按照直角坐标一样变换在空间反射下变号的物理量；**轴矢量** (又称**赝矢量**) 是在空间转动下与矢量一样变换但是在空间反射下不变的物理量。

一个力学的体系的拉格朗日量 (作用量) 一般来说一定是一个标量。对于一个多粒子组成的闭合系统，它的动能和势能都是标量。其运动方程显然在空间反射变换下保持同样的形式 (两边都变号)。矢量的例子如坐标、速度、动量等等；赝矢量的例子如角速度、角动量、磁场强度等等。例如，从角动量的定义可以看出，在空间反射下位置矢量和动量矢量都变号，因此两者的叉乘不变，也就是说角动量矢量是一个赝矢量。如果我们取一个赝矢量与一个矢量的点乘，我们就可以得到一个赝标量。

¶ 下面我们讨论时间反演变换。时间反演变换是指将时间 t 变为 $-t$ 的变换。也就是说，在时间反演变换下，时间的走向发生反转。一个力学系统，在时间反演变换下，它的所有广义速度都变号。因此它的拉格朗日量的变化行为为：

$$L(q,\dot{q},t) \to L(q,-\dot{q},-t)\ , \tag{2.67}$$

如果系统的拉氏量满足：

$$L(q,\dot{q},t) = L(q,-\dot{q},-t)\ , \tag{2.68}$$

那么，考察拉格朗日方程我们发现，系统的运动方程在时间反演变换下形式不变。最为常见的形式是拉格朗日量本身并不显含时间，同时拉格朗日量中只包含广义速度的偶数次幂 (二次幂)。时间反演对称性意味着力学系统的运动完全是可逆的。如果，某个系统的拉格朗日量中含有广义速度的一次幂，那么这个系统的运动方程在时间反演下不是不变的。这意味着时间有着一个确定的走向。一个典型的例子就是阻尼振子。

¶ 最后，在结束这一节时我们简要总结一下我们这节中讨论的对称性。本节所讨论的时空对称性可以分为连续对称性和分立对称性。连续对称性的对称变换是由若干个连续变量描写的 (例如平移的矢量、尺度变换的因子等等)；而分立对称性由分立的对称操作构成。事实上，本节中讨论的这些对称性 (连续的和分立的)，再加上其他一些特殊的对称性操作可以构成时空变换的一个群，它被称为**共形群** (conformal group)。我们这里只是列举了共形群的对称变换中比较常见的几个。共形群的对称性在讨论低维量子场论问题中会起十分重要的作用。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

12 非惯性系的力学

¶我们首先讨论一个质点在非惯性系中的运动方程。在一个惯性系 K_0 中，一个质点的拉格朗日量可以写为：

$$L_0 = \frac{1}{2}m\mathbf{v}_0^2 - V \ . \tag{2.69}$$

现在我们假定有另外一个参照系 K'(不一定是惯性系)，它相对于惯性参照系 K_0 以速度 $\mathbf{V}(t)$ 平动。于是，我们有：

$$\mathbf{v}_0 = \mathbf{v}' + \mathbf{V}(t) \ , \tag{2.70}$$

其中 $\mathbf{v}'$ 是粒子在参照系 K' 中的速度，而 $\mathbf{V}(t)$ 是一个给定的时间的函数。于是，质点的拉氏量可以写为：

$$L' = \frac{1}{2}m\mathbf{v}'^2 + m\mathbf{v}'\cdot\mathbf{V}(t) + \frac{1}{2}m\mathbf{V}^2 - V \ .$$

这个式子中的 $\frac{1}{2}m\mathbf{V}^2$ 是个给定的时间的函数的时间微商，可以从拉格朗日量中去掉；类似地，我们可以将：$m\mathbf{V}\cdot\dot{\mathbf{v}}'$ 换成 $-m\dot{\mathbf{V}}\cdot\mathbf{x}'$。于是，在参照系 K' 中一个质点的拉格朗日量可以等价地写为：

$$\boxed{L' = \frac{1}{2}m\mathbf{v}'^2 - m\frac{d\mathbf{V}(t)}{dt}\cdot\mathbf{x}' - V \ .} \tag{2.71}$$

这个拉格朗日量给出的运动方程为：

$$\boxed{m\frac{d\mathbf{v}'}{dt} = -\frac{\partial V}{\partial \mathbf{x}'} - m\frac{d\mathbf{V}(t)}{dt} \ .} \tag{2.72}$$

这个方程的右边的第二项 $-m\dot{\mathbf{V}}$ 被称为惯性力。

• • •

现在我们考察另一个参照系 K，它的原点与参照系 K' 重合，但是相对于参照系 K' 以给定的角速度 $\mathbf{\Omega}(t)$ 转动。因此，参照系 K' 中质点的速度可以写为：$\mathbf{v}' = \mathbf{v} + \mathbf{\Omega}\times\mathbf{x}$，其中 $\mathbf{x}$ 和 $\mathbf{v}$ 分别是该质点在参照系 K 中的位置和速度矢量 (注意，参照系 K 和 K' 的位置矢量是重合的，因为我们假定两个参照系原点重合)。于是，质点的拉格朗日量变为：

$$\boxed{L = \frac{1}{2}m\mathbf{v}^2 + m\mathbf{v}\cdot\mathbf{\Omega}\times\mathbf{x} + \frac{1}{2}m(\mathbf{\Omega}\times\mathbf{x})^2 - m\dot{\mathbf{V}}\cdot\mathbf{x} - V \ .} \tag{2.73}$$

这就是一个任意参照系中一个质点的拉格朗日量。分别计算：$(\partial L/\partial\mathbf{v})$ 和 $(\partial L/\partial\mathbf{x})$ 并带入拉格朗日方程，我们就得到了任意参照系中的一个质点的运动方程：

$$\boxed{m\frac{d\mathbf{v}}{dt} = -\frac{\partial V}{\partial\mathbf{x}} - m\dot{\mathbf{V}} + m\mathbf{x}\times\dot{\mathbf{\Omega}} + 2m\mathbf{v}\times\mathbf{\Omega} + m\mathbf{\Omega}\times(\mathbf{x}\times\mathbf{\Omega}) \ .} \tag{2.74}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

与惯性系中的运动方程比较我们发现，这个运动方程中多出了四项惯性力。第一项 $-m\dot{\mathbf{V}}$ 是由于非匀速的平动造成的；第二项 $m\mathbf{x}\times\dot{\boldsymbol{\Omega}}$ 是由于非匀速的转动引起的惯性力；第三项 $2m\mathbf{v}\times\boldsymbol{\Omega}$ 是著名的科里奥利力；[29] 第四项 $m\boldsymbol{\Omega}\times(\mathbf{x}\times\boldsymbol{\Omega})$ 是 惯性离心力。

现在让我们考察一个相对于惯性系没有平动，只有均匀转动的参照系。这时的拉格朗日量简化为：

$$L=\frac{1}{2}m\mathbf{v}^2+m\mathbf{v}\cdot\Omega\times\mathbf{x}+\frac{1}{2}m(\Omega\times\mathbf{x})^2-V\ .$$

粒子的广义动量为：

$$\mathbf{p}=\frac{\partial L}{\partial\mathbf{v}}=m\mathbf{v}+m\boldsymbol{\Omega}\times\mathbf{x}\ , \tag{2.75}$$

于是粒子的能量为：

$$E=\mathbf{p}\cdot\mathbf{v}-L=\frac{1}{2}m\mathbf{v}^2-\frac{1}{2}m(\boldsymbol{\Omega}\times\mathbf{x})^2+V\ . \tag{2.76}$$

这里的第二项能量：$-\frac{1}{2}m(\boldsymbol{\Omega}\times\mathbf{x})^2$ 被称为离心势能。这个能量表达式还可以表达为：

$$E=E_0-\mathbf{L}\cdot\boldsymbol{\Omega}\ , \tag{2.77}$$

其中 E_0 是粒子在惯性系 K_0 中的能量，而 $\mathbf{L}$ 则是粒子的角动量 (两个系中相同)。这个表达式也可以推广到多个粒子组成的系统。这个能量的表达式将在讨论转动物体的统计力学中会用到。

13 非完整约束系统的最小作用量原理

¶ 到目前为止，本章的所有讨论实际上都仅仅涉及完整约束的系统。因为我们假定描写一个力学系统的所有广义坐标都是独立的 自由度。这一点只有对于完整约束系统才成立。这一节中，我们试图将前面讨论的最小作用量原理推广到具有非完整约束 的力学系统。我发誓，这基本上是我们课程中最后一次讨论非完整约束。

首先，并不是所有的非完整约束系统都可以用最小作用量原理来加以描述。我们将只讨论一类特殊的非完整约束系统，即系统的广义坐标 (假定有 s 个) 和时间的微分存在着一系列线性关系：

$$a_{lj}dq_j+a_{lt}dt=0\ ,\quad l=1,2,\cdots,m\ ,j=1,2,\cdots,s\ . \tag{2.78}$$

其中 a_{lj} 和 a_{lt} 都可以是坐标 q_j 和时间的给定的函数，m 代表了这类约束的个数，s 是系统的广义坐标的个数。需要注意的是，由于是非完整约束，因此广义坐标的个数 s 一般大

[29] 地球上许多大尺度现象 (例如台风、洋流、季风等) 的主要成因都受到了科里奥利力的影响。但对小尺度现象，特别值得一提的是家中浴盆、马桶等下水的旋转方向，其实并不是科里奥利力主导的结果。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

于系统真正的自由度数目 f。这时，系统的自由度数目是 $s-m$。显然，所有的完整约束都可以表达成这种形式 (只不过 $m=0$)。另外，一些非完整约束也可以写成这种形式。大家也许还记得，我们在第 3节介绍约束时讨论的圆盘问题 (例题 1.1) 的约束 (一个非完整约束) 就是这种形式，参见式 (1.7)。

现在我们给每个广义坐标一个变分 δq_j 并且要求它们满足：

$$a_{lj}\delta q_j = 0\ ,\quad l=1,2,\cdots,m\ . \tag{2.79}$$

这些约束可以利用拉格朗日乘子的方法加以考虑。因而，我们利用带有拉格朗日乘子的变分法得到：

$$\int dt\left(\frac{\partial L}{\partial q_j}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j}+\lambda_l a_{lj}\right)\delta q_j = 0\ , \tag{2.80}$$

其中 λ_l，$l=1,2,\cdots,m$ 是 m 个拉格朗日乘子。在引入了拉格朗日乘子之后各个 δq_j 可以看成是独立变动的。于是我们得到体系的运动方程：

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j}-\frac{\partial L}{\partial q_j}=\lambda_l a_{lj}\ . \tag{2.81}$$

这些方程，与约束方程 (2.78) 一道 (一共 $s+m$ 个方程) 可以解出所有的广义坐标 (s 个) 和拉格朗日乘子 (m 个)。为了说明这种方法的运用，我们来继续求解第 3节中讨论了一半的圆盘问题 (例题 1.1)。

例题 2.1　二维平面上垂直纯滚的均匀圆盘（续） 利用本节的方法讨论具有非完整约束的圆盘问题，相应的图参见图 1.2。

解： 广义坐标 (x,y,θ,ϕ) 的选取与例题 1.1中的相同，它们满足的约束条件为：

$$\dot{x}=a\cos\phi\dot{\theta}\ ,\quad \dot{y}=a\sin\phi\dot{\theta}\ . \tag{2.82}$$

显然，圆盘的势能永远保持是常数。因此，我们可以将圆盘的拉格朗日量就取为它的动能。为此，我们可以得到：

$$L=\frac{1}{2}M(\dot{x}^2+\dot{y}^2)+\frac{1}{4}Ma^2\dot{\theta}^2+\frac{1}{8}Ma^2\dot{\phi}^2\ . \tag{2.83}$$

现在我们可以按照普遍的方程 (2.81) 写出系统的四个广义坐标的运动方程：

$$\begin{cases}\ddot{x}=\lambda_1\ ,\\ \ddot{y}=\lambda_2\ ,\\ a\ddot{\theta}=-2\lambda_1\cos\phi-2\lambda_2\sin\phi\ ,\\ \ddot{\phi}=0\ .\end{cases} \tag{2.84}$$

我们如果将前面约束的式子 (2.82) 对时间再取一次导数得到：

$$\ddot{x} = a\ddot{\theta}\cos\phi-\sin\phi(a\dot{\theta})\dot{\phi}\ , \tag{2.85}$$

$$\ddot{y} = a\ddot{\theta}\sin\phi+\cos\phi(a\dot{\theta})\dot{\phi}\ , \tag{2.86}$$

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

将此两式带入运动方程的前两式并与第三式比较得到:

$$\lambda_1 \cos\phi + \lambda_2 \sin\phi = 0 \;\Rightarrow \ddot{\theta} = 0\,. \tag{2.87}$$

另一方面，(2.84) 指出：$\ddot{\phi} = 0$。因此，我们得到了一个相当平庸的解 (然而却是正确的解): 所有角度广义坐标的加速度为零，或者说 θ，ϕ 都随时间线性变化。将这个结果带入约束方程 (2.82) 并对时间积分，我们就得到了广义坐标 x 和 y 对时间的依赖关系。显然，这个关系预示着圆盘的质心作匀速圆周运动，因为 $\dot{x}^2 + \dot{y}^2 = a^2\dot{\theta}^2$ 为常数。 搞定



相关的阅读

这一章是我们这个课程中最最重要的一章。本章中我们讨论了描写一个经典力学系统的最基本的原理 -最小作用量原理，又称哈密顿原理。与通常理论力学的讨论不同的是，我们的出发点是狭义相对论的时空观而不是伽利略时空观。当然，后者可以作为前者的近似 (非相对论极限)。这样讨论的最大好处是突出了最小作用量原理是一个与狭义相对论完全兼容的基本原理。它不仅仅适用于纯力学体系，也适用于电磁系统；它不仅适用于粒子，也适用于场系统。虽然我们的讨论起源于狭义相对论，但是我们并不要求同学们熟练的掌握狭义相对论的动力学。这应当是电动力学课程的主要任务。但是对于它的基本描述至少是有益的。对于狭义相对论十分不熟悉的同学，可以参考相关的教科书。

如果你实在对于这种从相对论出发的理论框架不习惯，这里我简单说明一下一般的 (非相对论性) 理论框架。在这个纯粹非相对论性的框架中，伽利略时空观取代了狭义相对论时空观。于是，分析力学的起始点可以从本讲义的第 7节开始。然后可以直接跳到本讲义的第 10节。利用空间平移不变性、空间各向同性可以直接建立一个非相对论性自由粒子的拉格朗日量 (2.43)。后面的所有讨论，都可以按照本讲义的顺序进行了。事实上，朗道的书 [1] 基本上就是按照这个顺序来讲述的。至于说本讲义的第 6 节、第 8 节、第 9 节的内容往往是放在电动力学课程中讲述。

从具体内容来说，本章的第 6节的内容可以参考 [7] 的第七章。本章的第 7节的内容可以在任何分析力学的书籍中找到，例如朗道书 [1] 的第 §2 节。这一节中有关力学系统的位型空间 (流形) 和切空间、切丛的讨论可以参考 [4] 的 2.4 节，或者参考更加数学化的经典著作 [6]。本章的第 8节和第 9节的内容可以在 [2] 中的 §8，§9，§15，§16，§17，§18 中找到，或者参考 Goldstein 书 [3] 的第七章的第 8，第 9 两节。本章的第 11节的主要内容可以在朗道书 [1] 的 §6 到 §10 几节中找到，尽管关于宇称和时间反演的讨论是我加上的。本章的第 12节的内容可以在朗道书 [1] 的第 §39 节找到。本章的第 13节关于非完整约束的讨论内容可以参考 Goldstein 书 [3] 的第二章的第 4 节。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

第三章　有心力场中粒子的运动

本章提要

- 中心力场 (14)
- 开普勒问题 (15)
- 潮汐 (16)
- 开普勒问题的相对论修正 (18)

上一章中我们从狭义相对论的时空观和最小作用量原理出发讨论了力学体系的经典运动方程以及对称性的影响。这一章中我们将着重讨论一个有心力场中的经典动力学问题。特别重要的是所谓开普勒问题，也就是两个质点之间由万有引力相互作用的力学问题。这个问题一直是伴随着力学的成长不断发展的问题。我们将首先讨论最为一般的两体问题，说明它如何分解为质心的运动和相对运动。随后我们会讨论开普勒问题。这之后我们将讨论对经典开普勒问题的拓展讨论。这包括几个不同的因素的影响，例如潮汐问题和相对论效应对于行星近日点进动的修正等。

14　中心力场

¶考虑两个质量分别为 m_1 和 m_2 的粒子，通过一个有心势相互作用的力学体系。所谓有心势是指两者相互作用的势能仅仅是两个粒子之间距离的函数。因此，体系的拉格朗日量可以写为：

$$L = \frac{1}{2}m_1\mathbf{v}_1^2 + \frac{1}{2}m_2\mathbf{v}_2^2 - V(|\mathbf{x}_1 - \mathbf{x}_2|)\,, \tag{3.1}$$

其中 $\mathbf{v}_1$，$\mathbf{v}_2$ 分别是两个粒子的速度；$\mathbf{x}_1$，$\mathbf{x}_2$ 则是两个粒子的位置；$V(r)$ 是两个粒子之间的势能。地球人都知道，这样的一个两体问题可以完全分解为两个一体问题。由于这个力

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

学体系的总的动量是守恒的，因此两个粒子组成的体系的质心的速度是常矢量。为了方便我们取一个特定的惯性系使得两个粒子组成体系的质心速度为零。这样的参照系被称为质心系，并且质心就位于坐标架的原点。同时，我们引入两者的相对坐标 $\mathbf{x}=\mathbf{x}_1-\mathbf{x}_2$，于是：

$$\mathbf{x}_1=m_2\mathbf{x}/(m_1+m_2)\ ,\quad \mathbf{x}_2=-m_1\mathbf{x}/(m_1+m_2)\ . \tag{3.2}$$

因此，只要求出了 $\mathbf{x}(t)$ 我们就可以分别求出两个粒子的运动轨道。利用相对坐标 $\mathbf{x}$，拉格朗日量变为：

$$\boxed{L=\frac{1}{2}m\dot{\mathbf{x}}^2-V(r)\ ,} \tag{3.3}$$

其中 $m=m_1m_2/(m_1+m_2)$ 为两个粒子的约化质量 (也有称为折合质量 的)，$r=|\mathbf{x}|$ 为两个粒子的相对距离。我们看到，这个两体问题在质心的运动与相对运动完全分离之后变为一个单粒子在有心力场中的运动问题。

¶ 转动不变性告诉我们，这样的一个有心力场问题中粒子的角动量一定是守恒的。因此，我们可以将粒子的角动量取为沿 z 轴的方向。这样一来，粒子的运动轨道将完全处在 xy 平面内。采用更为方便的极坐标 (r,ϕ)，粒子的拉格朗日量可以写为：

$$L=\frac{1}{2}m(\dot{r}^2+r^2\dot{\phi}^2)-V(r)\ . \tag{3.4}$$

这个拉格朗日量不显含坐标 ϕ，因此与它相应的广义动量守恒，它实际上就是粒子沿 z 方向的角动量：

$$J=mr^2\dot{\phi}=\text{constant}\ . \tag{3.5}$$

这实际上就是著名的开普勒第二定律：行星单位时间扫过的面积是常数。我们看到，这个定律实际上不依赖于相互作用势 $V(r)$ 的具体形式，只要是有心势 (角动量守恒) 就可以了。

另外一个守恒的物理量是粒子的能量：

$$E=\frac{1}{2}m(\dot{r}^2+r^2\dot{\phi}^2)+V(r)=\frac{1}{2}m\dot{r}^2+\frac{J^2}{2mr^2}+V(r)\ . \tag{3.6}$$

于是我们得到：

$$\frac{dr}{dt}=\sqrt{\frac{2}{m}[E-V(r)]-\frac{J^2}{m^2r^2}}\ . \tag{3.7}$$

另一方面，角动量的方程可以写成：

$$\frac{d\phi}{dt}=\frac{J}{mr^2}\ , \tag{3.8}$$

这两个方程中如果消去时间 t 并积分，原则上就可以得到粒子的轨道。

$$\phi=\int^r dr'\frac{(J/r'^2)}{\sqrt{2m\left[E-V(r')\right]-\frac{J^2}{r'^2}}}+\phi_0\ . \tag{3.9}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

其中 ϕ_0 为一个常数。关于径向运动的微分方程 (3.7) 看上去很像一个粒子在一维空间的运动，只不过粒子感受到的有效势能为：

$$V_{\text{eff}}(r) = V(r) + \frac{J^2}{2mr^2} \,. \tag{3.10}$$

这里的第二项正是所谓的离心势能。

15 开普勒问题

¶ 如果有心势中势的形式是反比于距离，这时的力学问题称为开普勒问题。这类力学体系出现在以万有引力相互作用的行星与太阳之间，也可以存在于有库仑相互作用的电荷之间。我们首先来讨论相互吸引的情形，这时系统的有效势的形式为：

$$V_{eff}(r) = -\frac{\alpha}{r} + \frac{J^2}{2mr^2} \,. \tag{3.11}$$

其中 $\alpha > 0$ 为一个正的常数。这个有效势的特点是在 $r \to 0$ 时，它趋于正的无穷；在 $r \to \infty$ 时，它从负的值趋于零。在区间 $(0, \infty)$ 中间，有效势有一个极小值点。因此，粒子的径向运动可以分为三种情况：如果粒子的能量 $E > 0$，那么粒子可以跑向无穷远，也就是说粒子的运动并不是束缚在力心周围有限的空间内。对于严格的 $1/r$ 势，这时粒子的轨道是双曲线的一支。如果粒子的能量 $E < 0$，那么粒子的径向距离一定局限在一个有限的区间：$[r_{min}, r_{max}]$ 之内，这时粒子的运动一定被束缚在力心周围的一个有限区域内，不可能逃逸到无穷远。对于形如 $1/r$ 型的势，我们下面可以证明粒子的轨道是一个闭合的椭圆。如果粒子能量 $E = 0$，这时粒子的运动仍然可以逃逸到无穷，但运动轨道是一条抛物线。

¶ 要得到粒子运动的轨道方程，可以从能量和角动量守恒的方程中消去时间并积分。得到的结果是：

$$\boxed{\frac{p}{r} = 1 + e\cos\phi \,,} \tag{3.12}$$

其中参数 p 和 e 由下式给出：

$$\boxed{p = \frac{J^2}{m\alpha} \,, \quad e = \sqrt{1 + \frac{2EJ^2}{m\alpha^2}} \,.} \tag{3.13}$$

方程 (3.12) 是极坐标中标准的圆锥曲线的方程，极坐标的原点位于圆锥曲线的焦点 (focus)。这里我们选取了积分常数使得距离原点最近的点对应于 $\phi = 0$。这个点在天文学中称为近日点 (perihelion)。参数 $2p$ 被称为正焦弦 (latus rectum)，e 被称

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

为偏心率。平面解析几何的知识告诉我们，当 $e>1$ 时，轨道为双曲线；当 $e=1$ 时为抛物线；当 $0<e<1$ 时，轨道为椭圆；当 $e=0$ 时，轨道是正圆。[1]

对于椭圆形的轨道，这里的结果正是开普勒第一定律。椭圆的半长轴 a 和半短轴 b 的值分别为：

$$a=\frac{p}{(1-e^2)}=\frac{\alpha}{2|E|}\,,\quad b=\frac{p}{\sqrt{1-e^2}}=\frac{L_z}{\sqrt{2m|E|}}\,. \tag{3.14}$$

椭圆轨道的近日点距离 r_{min} 和远日点距离 r_{max} 分别为：

$$r_{min}=\frac{p}{1+e}=a(1-e)\,,\quad r_{max}=\frac{p}{1-e}=a(1+e)\,. \tag{3.15}$$

椭圆形轨道的周期可以用椭圆的面积除以面积速度 (一个正比于粒子角动量的常数) 来加以确定：

$$T=2\pi a^{3/2}\sqrt{m/\alpha}\,. \tag{3.16}$$

正如第 11.4小节中指出的一样，周期的平方之比等于半长轴立方之比。这就是著名的开普勒第三定律。

对于 $E\geq 0$，粒子的运动是无穷的。对于 $E>0$ 的情形，轨道为双曲线的一支，力心位于该支双曲线的内焦点。其近日点的距离 r_{min} 和双曲线的半长轴 a 分别为：

$$r_{min}=\frac{p}{e+1}=a(e-1)\,,\quad a=\frac{p}{e^2-1}=\frac{\alpha}{2E}\,. \tag{3.17}$$

对于 $E=0$ 的情形，偏心率等于 1，近日点距离 $r_{min}=p/2$。

对于排斥的势：$V(r)=\alpha/r$，其中 $\alpha>0$，我们得到的轨道永远是双曲线。与吸引势不同的是，现在力心位于双曲线一支的外焦点上。类似于前面的公式，我们有：

$$\boxed{\frac{p}{r}=-1+e\cos\phi\,,} \tag{3.18}$$

其中参数 p 和 e 仍然由公式 (3.13) 给出。近日点距离的表达式为：

$$r_{min}=\frac{p}{e-1}=a(e+1)\,. \tag{3.19}$$

¶最后，我们简单讨论一下开普勒问题的特殊性。对于形如 $1/r$ 的吸引势，如果粒子的运动是局限在有限区域的 (也就是说粒子的总能量小于零)，那么它的轨道一定是闭合的椭圆。这种轨道的闭合性并不是对于所有的有心势都成立的。它实际上是开普勒问题所特有的。另外一个已知的、能够构成闭合轨道的有心势是三维谐振子势。[2] 对于绝大多数

[1] 如果我们在轨道方程 (3.12) 中令 $\phi\to\pi+\phi$(这等价于重新选择极轴) 就可以改变轨道方程中 e 前面的符号。因此，不失一般性，我们总可以假定 $e\geq 0$。

[2] 这个结论又被称为 Bertrand 定理。详细的讨论和证明可以参考 [3] 中的附录。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

的有心势，即使粒子的能量小于零，粒子的运动局现在有限的区域，它的轨道也不一定是闭合的。这种不闭合性的标志就是其近日点的进动。如果有心势只是稍微的偏离 $1/r$ 的形式，那么能量小于零的粒子的轨道几乎是闭合的，只是其近日点有微小的进动。

¶开普勒问题中轨道的闭合性实际上意味着这个问题具有更多的对称性。这种对称性并不是由于时空基本对称性而是由于势能的特殊形式造成的，因此被称为动力学对称。动力学对称性的存在直接的后果就是有多余的守恒量。如果我们定义，

$$\mathbf{M} = \mathbf{p} \times \mathbf{L} + m\alpha\mathbf{x}/r \,, \tag{3.20}$$

其中 $\mathbf{p}$ 是粒子的动量，它感受到的中心势为 $V(r) = \alpha/r$，α 可以大于零或小于零。我们下面会直接验证，上式定义的矢量实际上是一个守恒量，它被称为拉普拉斯-龙格-楞次矢量 (Laplace-Runge-Lenz vector)。它的方向是沿着力心到近日点的方向。这个矢量是常矢量意味着力心到近日点的矢量的方向是不变的，也就是说，没有近日点的进动。这直接联系着轨道的闭合性。

要验证这一点，我们直接对它取时间的微商，记住角动量是守恒的，因此，

$$\frac{d}{dt}(\mathbf{p} \times \mathbf{L}) = \dot{\mathbf{p}} \times \mathbf{L} = \frac{m\alpha}{r^3}[\mathbf{x} \times (\mathbf{x} \times \dot{\mathbf{x}})] = \frac{m\alpha}{r^3}[\mathbf{x}(\mathbf{x} \cdot \dot{\mathbf{x}}) - r^2\dot{\mathbf{x}}] = -m\alpha\frac{d}{dt}\left(\frac{\mathbf{x}}{r}\right) \,,$$

其中第二步我们运用了粒子的运动方程 $\dot{\mathbf{p}} = \mathbf{F} = (\alpha/r^3)\mathbf{x}$；第三步我们运用了 $\mathbf{x} \cdot \dot{\mathbf{x}} = r\dot{r}$。所以我们一定有 $d\mathbf{M}/dt = 0$，即 $\mathbf{M}$ 是一个守恒的矢量。

例题 3.1　利用 Laplace-Runge-Lenz 矢量求轨道 利用上面给出的 Laplace-Runge-Lenz 矢量求出 Kepler 问题中轨道的形状。

解： 上面已经证明，$\mathbf{M}$ 是一个守恒的矢量。我们将其取为极坐标中极轴的方向。由于，

$$\mathbf{x} \cdot \mathbf{M} = rM\cos\phi \,, \tag{3.21}$$

其中 $M = |\mathbf{M}|$ 是该矢量的大小，ϕ 是 $\mathbf{x}$ 与 $\mathbf{M}$ 之间的夹角。另一方面，将 $\mathbf{x}$ 直接点乘 $\mathbf{M}$ 的定义式 (3.20) 我们得到，

$$\mathbf{x} \cdot \mathbf{M} = \mathbf{x} \cdot (\mathbf{p} \times \mathbf{L}) + m\alpha r = J^2 + m\alpha r \,. \tag{3.22}$$

原则上将上述两式联立就可以获得轨道。例如，对于吸引的问题，$\alpha = -|\alpha|$，我们有，

$$r\left(1 + \frac{M}{m|\alpha|}\cos\phi\right) = \frac{J^2}{m|\alpha|} \tag{3.23}$$

换句话说，我们发现 $p = J^2/(m|\alpha|)$ 以及偏心率满足 $e = M/(m|\alpha|)$，这与公式 (3.12) 以及 (3.13) 一致。同样，对于排斥势，$\alpha > 0$，我们得到的轨道方程为

$$r\left(-1 + \frac{M}{m\alpha}\cos\phi\right) = \frac{J^2}{m\alpha} \tag{3.24}$$

这与公式 (3.18) 一致。 **搞定**

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

例题 3.2　近日点的进动 如果中心势偏离开普勒形式很小，找出近日点进动的表达式。

解： 这个问题中我们假定:

$$V(r) = -\frac{\alpha}{r} + \delta V(r)\,, \tag{3.25}$$

其中 $\delta V(r)$ 是一个很小的势能修正。正如我们前面提到的，只要势能偏离了 $1/r$ 的形式，行星的轨道一般就不是闭合的。我们希望求出轨道近日点的进动，用 $\delta V(r)$ 来表达。

出发点是上一节的公式 (3.9)。对于一个一般的势能，行星距离从近日点运行一周又回到近日点时的进动角度 $\Delta\phi$ 为:

$$\Delta\phi = 2\int_{r_{min}}^{r_{max}} \frac{dr(J/r^2)}{\sqrt{2m\left[E - V(r)\right] - \frac{J^2}{r^2}}}\,. \tag{3.26}$$

如果 $\delta V(r) = 0$，那么可以证明上面的这个公式正好给出 2π。这意味着轨道是闭合的，而且系统的两个自由度 ϕ 和 r 的周期完全一致。现在，假定 $\delta V(r)$ 是小量，我们可以对它进行泰勒展开，仅仅保留到一阶得到：$\Delta\phi = 2\pi + \delta\phi$，其中的 2π 就是 $1/r$ 势的贡献。我们关心的是 $\delta\phi$，它是由于 $\delta V(r)$ 引起的近日点进动:

$$\delta\phi = 2m\frac{\partial}{\partial J}\int_{r_{min}}^{r_{max}} \frac{dr\delta V(r)}{\sqrt{2m\left[E + \frac{\alpha}{r}\right] - \frac{J^2}{r^2}}}\,. \tag{3.27}$$

只要给出了具体的 $\delta V(r)$ 的形式，这个表达式就可以给出行星一个周期中近日点的进动数值。需要注意的是，由于我们仅仅关心准确到 δV 的一阶的修正，因此上面这个公式中的积分上下限都可以用 $\delta V = 0$ 时候的值代入。如果愿意，上式也可以利用对角度 ϕ 的积分表达出来:

$$\delta\phi = \frac{\partial}{\partial J}\left(\frac{2m}{J}\int_0^{\pi} r^2\delta V(r)d\phi\right)\,, \tag{3.28}$$

其中被积函数中的 r 必须用星体的开普勒轨道 (3.12) 来带入。这些就是考虑近开普勒问题中近日点进动的公式。在下面的例子中，我们首先考虑其他行星对水星近日点进动的影响；在第 18节中，我们会考虑狭义相对论对于近日点进动的修正。 **搞定**

例题 3.3　水外行星势造成的水星近日点进动 简单估计水外行星的运动对水星近日点进动的影响。

解： 我们知道水星近日点的进动号称是对广义相对论的经典检验之一。其实，相对论的修正 (包括狭义相对论和广义相对论) 只是水星近日点进动的一小部分而已。要正确地理解水星近日点进动的问题，首先必须计算出水外行星 (其实就是所有行星，因为水星是距离太阳最近的行星) 对水星近日点进动的贡献。本题中我们就来近似地估计一下这个贡献。

初看起来这是个不可能完成的任务，因为它实际上涉及多体之间的引力相互作用。但是作为一个简单的估计，我们假设太阳的引力作用仍然是主导的 (这至少在太阳系这个地方总是大体正确的吧)。因此，所有的行星在零级近似下仍然按照开普勒规律运动。其

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

次，由于水星的进动十分微小，而它的周期最短，因此我们将假定它感受到的其他行星对它的作用是一个平均的效果。具体来说，我们将假设另一个质量是 M，距离太阳为 a 的水外行星的所有质量都均匀地分布在其轨道之上。同时，为了简化计算，我们假定其他所有行星 (除了水星以外) 的轨道都是圆轨道并且都在同一个平面之内。[3] 换句话说，我们将假定水星感受到的 –除了太阳引力之外 –实际上是一个质量均匀分布的圆环对它的引力作用。

一个半径是 a，总质量 M 均匀分布的圆环在其内部距离圆心为 r 的点所产生的引力势可以直接计算出来:

$$\Phi(r) = -\frac{GM}{2\pi a}\int_{-\pi}^{\pi}\frac{ad\theta}{\sqrt{a^2+r^2-2ar\cos\theta}}\,. \tag{3.29}$$

如果我们假定 $a \gg r$，那么这个积分可以近似为:

$$\Phi(r) = -\frac{GM}{a}\left(1+\frac{1}{4}\frac{r^2}{a^2}+\cdots\right)\,. \tag{3.30}$$

因此除了一个常数之外我们得到的对开普勒问题中势能的偏离为:

$$\delta V(r) = -\frac{GMm}{4a^3}r^2\,. \tag{3.31}$$

将这个表达式带入我们前面例题中的一般表达式 (3.27)，我们得到:

$$\delta\phi = -\frac{GMm}{2a^3}\sqrt{\frac{m}{2|E|}}\frac{\partial}{\partial J}\int_{r_{min}}^{r_{max}}\frac{r^2dr}{\sqrt{(r-r_{min})(r_{max}-r)}}\,. \tag{3.32}$$

类似地，再令 $r = 1/u$ 之后这个积分可以化为复的 u 平面上的一个围道积分:

$$I \equiv \int_{r_{min}}^{r_{max}}\frac{r^3dr}{\sqrt{(r-r_{min})(r_{max}-r)}} = \frac{1}{2}\sqrt{u_1u_2}\oint\frac{du}{u^4\sqrt{(u-u_2)(u_1-u)}}\,,$$

其中 $u_1 = 1/r_{min} \equiv 1/r_1$，$u_2 = 1/r_{max} \equiv 1/r_2$。现在我们可以将积分化为绕 u 复平面原点的积分 (或者等价地说，计算其在原点的留数)，结果得到:

$$I = \frac{5\pi}{16}(r_1+r_2)^3 - \frac{3\pi}{4}(r_1+r_2)r_1r_2\,. \tag{3.33}$$

在带入前面公式 (3.15) 中给出的 r_1，r_2 与能量、角动量等的关系后，经过整理我们可以计算出这个贡献:

$$\delta\phi = \frac{3\pi}{2}\sqrt{1-e_{\mathrm{☿}}^2}\left(\frac{M}{M_\odot}\right)\left(\frac{a_{\mathrm{☿}}}{a}\right)^3\,. \tag{3.34}$$

[3] 对于圆轨道近似，只有水星和冥王星的轨道比较明显地偏离圆轨道。这两个行星的偏心率分别为：0.21 和 0.25。其余行星的偏心率都小于 0.1，因此基本上是圆轨道。同样，也只有水星和冥王星这两个行星运行的轨道平面与地球轨道平面的夹角最大，分别为 7° 和 17°。考虑到冥王星已经被“开除”出行星行列，它对水星近日点的进动的影响应当很小。因此我们下面的近似应当说是不错的。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

其中 a 和 M 表示我们所考虑的行星的轨道半径和质量，$M_\odot$ 则表示太阳的质量，$a_☿$ 和 $e_☿$ 则代表水星轨道的半长轴和偏心率。这个结果其实并不奇怪。我们前面看到，水外行星对水星近日点进动的影响是通过比值 M/a^3 进入本问题的。将这个比值无量纲化 (因为 $\delta\phi$ 是个无量纲的数)，我们知道结果一定是正比于 $(M/M_\odot)(a_☿/a)^3$。当然，要计算出前面的系数仅仅靠量纲分析是不够滴。

现在我们可以来分析各个行星的影响了。在表 3.1中我们列出了有重要贡献的行星的相关数据以及按照按照公式 (3.34) 计算的进动值。其余未在列的行星的贡献基本都可以忽略。特别需要说明的是，我们这里的计算基于一个模型假设，即行星的所有质量均匀分布在其轨道之上。这个假定对于距离水星比较近的行星 –特别是金星 –并不太适合。这导致按照这个计算出的金星的数据比实际的 (278"/ 世纪) 要小。将所有行星的贡献相加，我们发现水外行星对水星近日点进动的贡献大约在 530"/ 世纪。相对论的修正一般比这个还要小大约一个量级。例如，著名的广义相对论给出的修正大约为 43"/ 世纪。我们在后面会利用狭义相对论来计算水星近日点的进动 (参见第 18节)，结果只是广义相对论的 1/6。

表 3.1: 各行星对于水星近日点进动的影响。表中仅仅列出了贡献最大的几个行星。未列的行星贡献可以忽略。第三行列出了按照公式 (3.34) 计算的进动值。为了得到这个数值，我们还利用了地球与太阳质量之比：$M_♁/M_\odot = 3\times10^{-6}$ 以及水星的周期：$T_☿ = 0.24$ 年。

	水星 (☿)	金星 (♀)	地球 (♁)	火星 (♂)	木星 (♃)	土星 (♄)
质量 M [地球质量]		0.815	1	0.107	317.8	95.152
半长轴 a[天文单位]	0.387	0.723	1	1.52	5.20	9.58
$\delta\phi$[秒/世纪]		151.2	70.2	2.1	158.3	7.6

搞定

16 潮汐现象

¶这一节中我们讨论一下日常生活中遇到的潮汐现象。我们考虑两个球形的、靠万有引力束缚在一起的星体，例如地球与太阳，或者地球与月亮。这个两体问题可以将其质心的坐标运动分出，并且取质心系 (一个惯性系)，在这个参照系中两个星体系统的质心是静止的。我们取坐标的原点位于质心。为了进一步简化讨论，我们假定两个星体的运动轨道是一个正圆 (这一点对于地日系统、地月系统差不多都是成立的)。如果两个星体公转的角速度为 Ω。现在我们取另外一个相对于质心系旋转的参照系，其旋转的角速度正好等于 Ω。在这个参照系中 (非惯性系)，两个星体是相对静止的。简单的计算表明：

$$\Omega^2 = \frac{G(m_1+m_2)}{R^3}\,, \tag{3.35}$$

其中 R 代表两个星体中心的距离，m_1，m_2 为两个星体的质量。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

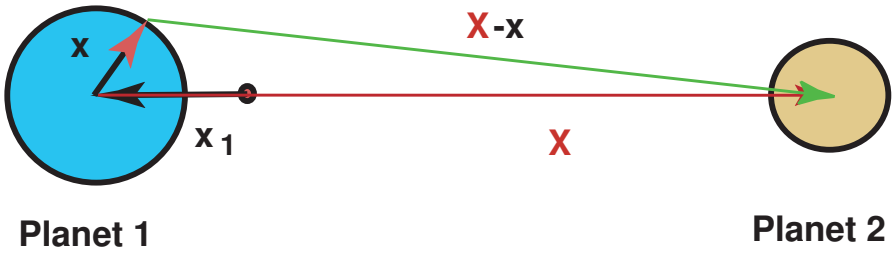

图 3.1: 两个星体潮汐问题的示意图。

处理这个问题的最简单的方法是利用第 12节中给出的旋转参照系中的能量的表达式 (2.76)。我们考察第一个星体表面任意一点处单位质量的能量。图 3.1中显示了两个星体的相对位置关系。为此，我们设第一个星体的表面任意一点相对于第一个星体的中心的位置矢量为 $\mathbf{x}$，那么在这个非惯性系中，位于点 $\mathbf{x}$ 处的单位质量的势能可以写为：

$$\Phi(\mathbf{x}) = -\frac{Gm_1}{r} - \frac{Gm_2}{|\mathbf{X}-\mathbf{x}|} - \frac{1}{2}\Omega^2(\mathbf{x}+\mathbf{x}_1)^2 \,, \tag{3.36}$$

其中前两项分别是两个星体的万有引力造成的势能；第三项是非惯性系中的离心势能。需要特别注意的是，离心势能中的距离是到质心的距离。这个公式中的矢量：$\mathbf{X} \equiv \mathbf{x}_2 - \mathbf{x}_1$ 就是从第一个星体的中心指向第二个星体中心的位置矢量。它的大小我们记为：$R = |\mathbf{X}|$。利用这个矢量，两个星体的位置矢量分别为：

$$\mathbf{x}_1 = -\frac{\mathbf{X}m_2}{m_1+m_2} \,, \quad \mathbf{x}_2 = \frac{\mathbf{X}m_1}{m_1+m_2} \,. \tag{3.37}$$

现在我们假定 $R = |\mathbf{X}| \gg |\mathbf{x}|$，于是我们可以将 (3.36) 中的第二项展开：

$$\frac{1}{|\mathbf{X}-\mathbf{x}|} \simeq \frac{1}{R}\left[1 + \frac{\hat{\mathbf{X}}\cdot\mathbf{x}}{R} - \frac{\mathbf{x}^2}{2R^2} + \frac{3}{2}\left(\frac{\hat{\mathbf{X}}\cdot\mathbf{x}}{R}\right)^2 + \cdots\right] \,, \tag{3.38}$$

其中 $\hat{\mathbf{X}} \equiv \mathbf{X}/R$ 代表沿着 $\mathbf{X}$ 方向的单位矢量。将这个展开式带入 (3.36) 中，并且将 (3.37) 带入，经过一些计算我们得到：

$$\boxed{\Phi(\mathbf{x}) = -\frac{Gm_1}{r} - \frac{Gm_2}{R} - \frac{1}{2}\Omega^2(\mathbf{x}^2+\mathbf{x}_1^2) - \frac{3}{2}\frac{Gm_2}{R}\left(\frac{\hat{\mathbf{X}}\cdot\mathbf{x}}{R}\right)^2} \,. \tag{3.39}$$

得到这个公式的过程中，我们还利用了关系 (3.35)。特别要提醒大家注意的是，势能 (3.36) 的第二项和第三项进行展开时都会出现线性依赖于 $\hat{\mathbf{X}}\cdot\mathbf{x}/R$ 的项，但是两者正好相消 (在利用了关系 (3.35) 后)。因此，最终的势能表达式中只剩下了平方项。

表达式 (3.39) 中的前三项都不依赖于点 $\mathbf{x}$ 在第一个星体表面的位置。如果第一个星体表面是球体，这些项都是恒定的。但是，最后一项却依赖于点 $\mathbf{x}$ 在球面的位置。具体来

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

说，当 **x** 位于两个星体的连线与第一个星体表面的交点时 (也就是说矢量 **x** 与单位矢量 $\hat{\mathbf{X}}$ 平行或反平行时)，这一项的贡献会取最小值 (势能最低)；相反，当 **x** 与连线方向 $\hat{\mathbf{X}}$ 垂直时，这一项取最大值 (为零)。由于最后一项与前几项相比是小的，因此，如果不考虑最后一项的贡献，势能在第一个星体表面的分布是均匀和各向同性的。如果考虑到最后一项的修正，那么势能在沿着两个星体中心连线的方向会比其他方向更低一些。因此，如果第一个星体表面具有流动性的液体物质，比如地球表面的海洋，那么这些流动性的液体就会比较倾向于集中到沿着两个星体中心连线的方向上去，这就造成了 潮汐 。与这一项势能相对应的力被称为 引潮力 。公式 (3.39) 告诉我们，引潮力取决于第二个星体的质量 (与之成正比) 以及它与第一个星体的距离 (与其三次方成反比)。具体的分析表明，太阳和月亮都会对地球造成潮汐，但是由于月亮比较近，因此月亮的影响反而要更重要一些，尽管它的质量远小于太阳的质量。[4]

¶ 虽然液体的潮汐现象最为明显，但实际上固体也受潮汐的影响。这种影响在没有液体的星体表面就突显出来了。例如，构成月球表面的固体由于受到地球的引潮力的影响，会使得月球正对着和背对着地球的一面的固体发生微小的突起，这被称为 固体潮 。这个效应的长期的积累会使得月球的自转消失。也就是说，月球总是以同一面对着地球。对地球来说，由于我们表面有海洋，因此这种效应仅仅造成潮汐现象，对地球的自转的影响远没有地球对月球影响那么强烈。[5]

17 受限三体问题与拉格朗日点

¶ 前面的讨论告诉我们，经典引力影响下的两体问题已经完全获得解决。体系的运动可以分解为质心的运动 (匀速直线运动) 和相对运动，后者就是典型的 Kepler 问题。其轨道可以由圆锥曲线描写。在成功解决了两体问题后，人们尝试求解类似的三体问题。在经典牛顿力学的框架中，所谓 三体问题 (three-body problem) 是指任意质量的三个质点，假定相互之间只有牛顿万有引力的相互作用，该三体系统在任意初始条件下运动轨迹的求解问题。这个问题实际上是远比两体问题复杂得多。事实上，对于一般的三体问题至今也没有完整的解析解。[6]

一般的经典三体问题无法解析求解不意味着在一些特殊的构型下可以求解。下面我们就来讨论一大类非常有实际应用背景的情形：即三个物体的质量，分别记为 m_1，m_2 和 m，有非常大的差别的情形。其中两个物体的质量远大于第三个物体的质量:

$$m_1 \gg m\,, m_2 \gg m\,. \tag{3.40}$$

[4] 公式 (3.39) 还告诉我们，对地球上一个固定的地点来说，引潮力一天之内会有两次 (而不是一次！)、相隔大约半天时间的极大值。这正是当年困扰伽利略的最主要的事情。

[5] 其实也是有影响的，只不过还不足以让地球自转消失罢了。事实上，地球也是越转越慢滴。据估算，现在的一天比起 100 年前多了大约 1.7ms。

[6] 关于这一点的证明首先是数学家 Heinrich Bruns 和 Henri Poincaé 在 1887 年给出的。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

在这种情形下，最轻的质量所产生的引力场对于另外两个大质量物体的运动之影响完全可以忽略。也就是说，这个问题其实可以化简为两个大质量物体 m_1 和 m_2 的两体问题，它们的轨道由我们前面讨论的 Kepler 问题的解给出；而最轻的质量 m 只是在 m_1 和 m_2 所形成的固定的引力场之中运动而已。这类问题的最典型的应用就是考虑地球、月球和人造卫星之间的三体问题，或者太阳、地球、卫星之间的三体问题。无论是哪一类，人造卫星的质量都比太阳、地球或月球要小很多，因此它的存在并不会对日地或者月地的正常轨道有任何可观测的影响。我们称这类问题为**受限三体问题** (restricted three-body problem)。

尽管对于 m_1 和 m_2 构成的两体系统来说，其任意的圆锥曲线的轨道都是可能的，为了与实际更接近些，我们还是加上两者实际上按照圆轨道运行。[7] 这个假设不仅仅对于月地系统是正确的，而且对于太阳系的各个行星与太阳构成的系统基本上也都是对的。[8] 这样我们基本上回到了上一节中讨论的情形，参考图 3.1，即两个质量分别为 m_1 和 m_2 的星体相互之间靠万有引力吸引在圆轨道上运行。从上节的讨论我们知道，选取一个与两个星体一同转动的非惯性系是方便的。这个转动系的原点可以选择在两个星体的质心位置而转动的角速度由公式 (3.35) 给出。

除了忽略掉质量 m 对 m_1 和 m_2 的运动的影响之外，结合具体的情况我们还可以假设两个星体的质量也有比较大的差距。我们约定，

$$m_1 \gg m_2 \gg m \,. \tag{3.41}$$

注意，这个假设虽然原则上并不是必须的。但是由于实际的应用中往往都可以得到满足，因此它可以大大简化我们的计算过程。例如，对于月地系统，较小质量的月球和较大质量的地球的质量比为 $m_2/m_1 = M_{☽}/M_{♁} \sim 0.0123$。对于日地系统来说，$m_2/m_1 = M_{♁}/M_{☼} \sim 3.0 \times 10^{-6}$。即便是对于质量最重的木星而言，它与太阳的质量比也是非常小的一个数值：$m_2/m_1 = M_{♃}/M_{☼} \sim 10^{-3}$。因此，对于任何有实际应用的受限三体系统而言，我们都可以令

$$\varepsilon = \frac{m_2}{m_1} \ll 1 \,, \tag{3.42}$$

是一个小量。这将极大地简化我们的计算。本节中我们将对这类特殊的三体问题做一个简要的介绍。希望了解进一步细节的读者可以参考专门的书籍。[9]

17.1 拉格朗日点

选取两个星体的质心为坐标原点并假定两个星体的运动在 xy 平面内。现在我们转换到与两个星体一起均匀旋转的非惯性系之中。为了简化讨论，我们选择两个星体的连线为

[7] 无论是地球绕太阳的轨道，还是月球绕地球的轨道都非常接近于圆轨道。

[8] 除了水星之外，其他行星的轨道都非常接近于圆轨道。

[9] 可以参考 Marsden 等人的专著：W.S. Koon, M.W. Lo, J.E. Marsden, S.D. Ross, "Dynamical systems, the three-body problem and space mission design", 2006.`http://www.cds.caltech.edu/~marsden/books/Mission_Design.html`.

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

x 轴，第一、第二个星体的坐标分别为 $\mathbf{x}_1 = (r_1, 0, 0)$，$\mathbf{x}_2 = (-r_2, 0, 0)$，显然，

$$r_1 = \frac{m_2}{m_1 + m_2} R \,, \quad r_2 = \frac{m_1}{m_1 + m_2} R \,.$$

对于这个特殊的问题选取一个合适的单位制是方便的。我们首先选取长度单位，使得 $R = 1$。换句话说，所有的长度量都按照 R 来度量。[10] 我们同时还可以选择时间的度量单位，使得 $\Omega = 2\pi$。[11] 这样一来，上面的 r_1 和 r_2 都变成无量纲的量，

$$r_1 = \frac{m_2}{m_1 + m_2} \,, \quad r_2 = \frac{m_1}{m_1 + m_2} \,, \tag{3.43}$$

而角速度的关系 (3.35) 则变为，

$$\Omega^2 = \frac{G(m_1 + m_2)}{R^3} = G(m_1 + m_2) = 4\pi^2 \,. \tag{3.44}$$

由于我们考虑受两个星体引力影响的另一个质量为 m 的质点的运动，我们可以方便地将该质点的质量取为单位，即 $m = 1$。这样我们就有了所有的单位设定，下面出现的所有力学量都变为无量纲的了。在需要恢复量纲的时候我们只需要将任何的无量纲物理量乘以相应的量纲即可。

我们将进一步假定可以仅仅考虑该质点在两个星体的轨道平面内运动。[12] 记质点的坐标为 $\mathbf{x} = (x, y)$。按照我们前面所讲的公式，具体来说就是第 12节的公式 (2.73)，在旋转参照系中质点的拉格朗日量应当写为 (其中我们已经设 $m = 1$)，

$$\begin{aligned} L &= \frac{1}{2}\dot{\mathbf{x}}^2 + \dot{\mathbf{x}} \cdot \boldsymbol{\Omega} \times \mathbf{x} + \frac{1}{2}(\boldsymbol{\Omega} \times \mathbf{x})^2 - V(\mathbf{x}) \,, \\ &= \frac{1}{2}\dot{\mathbf{x}}^2 + \dot{\mathbf{x}} \cdot \boldsymbol{\Omega} \times \mathbf{x} - V_{\text{eff}}(\mathbf{x}) \end{aligned} \tag{3.45}$$

其中 $V(\mathbf{x})$ 是质点感受到的来自两个星体的引力势能，$\boldsymbol{\Omega} = \Omega \hat{z}$ 是随两个星体一同旋转的参照系的角速度。同时我们还定义了有效势能，

$$\boxed{V_{\text{eff}}(\mathbf{x}) = -\frac{1}{2}(\boldsymbol{\Omega} \times \mathbf{x})^2 + V(\mathbf{x}) \,,} \tag{3.46}$$

其中第一项是离心势能，第二项是引力势能。粒子运动的拉格朗日运动方程可以从其拉格朗日函数 (3.45) 获得，

$$\ddot{\mathbf{x}} + 2\boldsymbol{\Omega} \times \dot{\mathbf{x}} = -\frac{\partial V_{\text{eff}}(\mathbf{x})}{\partial \mathbf{x}} \,, \tag{3.47}$$

其中 $2\boldsymbol{\Omega} \times \dot{\mathbf{x}}$ 的项是科里奥利力的贡献，$\partial V_{\text{eff}}(\mathbf{x})/\partial \mathbf{x}$ 的项则体现了有效势的贡献，它又包含惯性离心力和两个星体对质点的引力之和。在有效势能的极值点处，$\partial V_{\text{eff}}(\mathbf{x})/\partial \mathbf{x} = 0$。

[10] 具体来说，对于日地系统而言，这就是所谓的天文单位 (astronomical unit)。

[11] 对于日地系统而言，这就是以年为时间单位。

[12] 这可以大大简化下面的讨论。关于更加复杂的三维的运动的讨论，有兴趣的同学可以参考前面引过的 Marsden 等人的专著。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

因此，如果一个质点在某个时刻恰好静止于这些极值点处，那么它将永远停留在那里。因为在这些点处惯性离心力和引力刚好平衡相消。如果粒子初始没有速度，也不存在科里奥利力。所以该质点将永远停留在这些极值点处。这些使得有效势能取极值的点被统称为拉格朗日点。我们下面会更仔细地考察这些点。

• • • • •

将有效势能 (3.46) 具体写出来为，

$$V_{\text{eff}}(x,y) = -\frac{1}{2}\Omega^2(x^2+y^2) - \frac{Gm_1}{[(x-r_1)^2+y^2]^{1/2}} - \frac{Gm_2}{[(x+r_2)^2+y^2]^{1/2}} \,. \tag{3.48}$$

下面我们寻找有效势能 $V_{\text{eff}}(\mathbf{x})$ 的极值点 $\mathbf{x}_L = (x_L, y_L)$ 所满足的方程。对上面的有效势能求导数并令其为零得到，

$$\begin{cases} \Omega^2 x_L = \dfrac{Gm_1(x_L-r_1)}{[(x_L-r_1)^2+y_L^2]^{3/2}} + \dfrac{Gm_2(x_L+r_2)}{[(x_L+r_2)^2+y_L^2]^{3/2}} \\ \Omega^2 y_L = \dfrac{Gm_1 y_L}{[(x_L-r_1)^2+y_L^2]^{3/2}} + \dfrac{Gm_2 y_L}{[(x_L+r_2)^2+y_L^2]^{3/2}} \end{cases} \tag{3.49}$$

我们看到拉格朗日点大致可以分为两类：一类是位于两个星体连线 (即 x 轴) 上的，满足 $y_L = 0$；另一类则是不在两星体连线上的。

首先看 $y_L = 0$ 的拉格朗日点。这时候极值方程 (3.49) 的第二个方程自动得到满足。第一个方程在令 $y_L = 0$ 后化简为，

$$x_L = \left(\frac{Gm_1}{\Omega^2}\right)\frac{(x_L-r_1)}{|x_L-r_1|^3} + \left(\frac{Gm_2}{\Omega^2}\right)\frac{(x_L+r_2)}{|x_L+r_2|^3} \,. \tag{3.50}$$

现在注意到，

$$\frac{Gm_1}{\Omega^2} = \frac{Gm_1}{G(m_1+m_2)} = \frac{m_1}{m_1+m_2} = r_2 \,, \quad \frac{Gm_2}{\Omega^2} = r_1 \,.$$

其中我们利用了公式 (3.43) 和公式 (3.44)。如果我们定义一个函数 $f(x)$

$$f(x) = r_2\frac{(x-r_1)}{|x-r_1|^3} + r_1\frac{(x+r_2)}{|x+r_2|^3} \,. \tag{3.51}$$

那么满足 $y_L = 0$ 的拉格朗日点满足的方程可以由下列非线性方程给出：[13]

$$x_L = f(x_L) \,. \tag{3.52}$$

很容易验明函数 $f(x)$ 满足如下的性质：

$$\begin{aligned} &f(-r_2+0^+) = +\infty \,, \quad f(+r_1+0^+) = +\infty \,, \\ &f(-r_2-0^+) = -\infty \,, \quad f(+r_1-0^+) = -\infty \,, \\ &f(\pm\infty) = 0 \,. \end{aligned} \tag{3.53}$$

[13] 实际上是一个关于 x_L 的 5 次方程。它一定有三个实数根和一对互为复共轭的复根。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

而且它是关于 x 的一个减函数。因此当求解 (3.52) 时，一定存在三个实数解：一个位于 $(-\infty, -r_2)$，一个位于 $(-r_2, +r_1)$，另一个位于 $(+r_1, +\infty)$。也就是说位于两个星体连线上的拉格朗日点共有三个，一个位于左边星体的左边，一个位于两个星体之间，一个位于右边星体的右边。按照通常的约定，L_1 是位于两个星体之间的那个拉格朗日点；L_2 是在较小质量星体外侧的拉格朗日点；L_3 则是较大质量天体外侧的拉格朗日点。

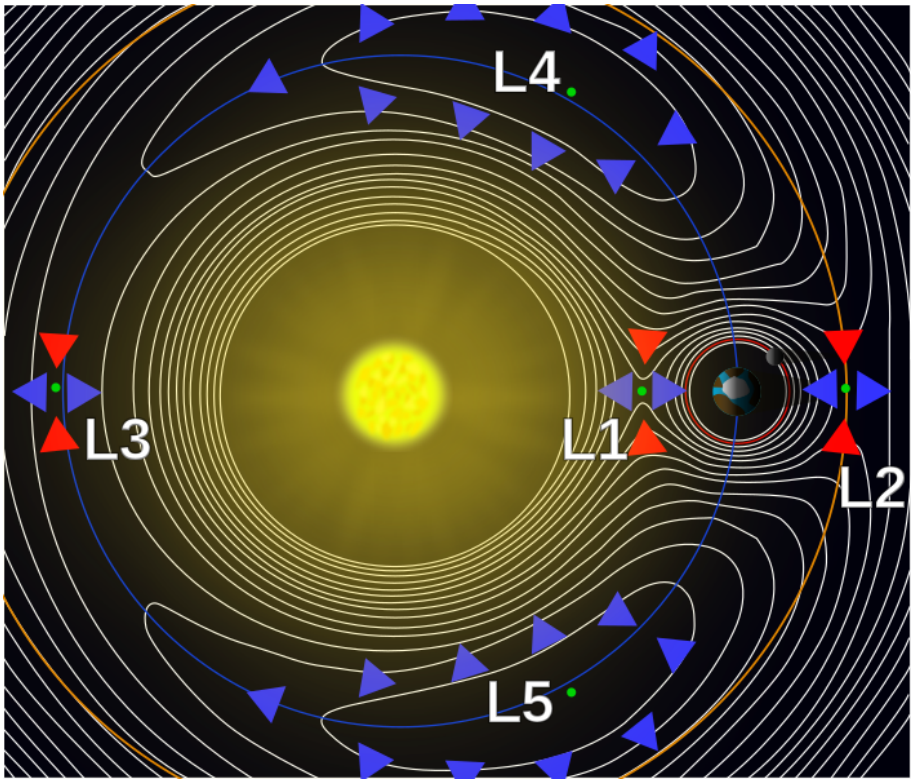

图 3.2: 日地系统中的五个拉格朗日点的示意图 (比例并不正确)。图中还显示了有效势能 $V_{\text{eff}}(\mathbf{x})$ 的等高线。需要注意的是，这个图中 x 轴的设定与我们文中是相反的，即如果要与我们文中讨论一致，x 轴的正方向应当向左。图片取自 https://en.wikipedia.org/wiki/Lagrangian_point。

在 $\varepsilon \to 0$ 的极限下，我们可以近似地来求解前面给出的五次方程 (3.52)。这个时候我们对 r_1 和 r_2 有如下的近似，

$$\begin{aligned} r_1 &= \frac{\varepsilon}{1+\varepsilon} \sim \varepsilon - \varepsilon^2 + \cdots , \\ r_2 &= \frac{1}{1+\varepsilon} \sim 1 - \varepsilon + \varepsilon^2 - \cdots . \end{aligned} \tag{3.54}$$

我们可以将相应的解表达为 ε 的级数。事实上，对于 L_1 和 L_2 来说，它们满足的方程可以写为，

$$x_L = -\frac{r_2}{(x_L - r_1)^2} \pm \frac{r_1}{(x_L + r_2)^2} , \tag{3.55}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

其中 $+/-$ 分别对应于 $L_{1,2}$。在零级近似下，$r_1=0$，$r_2=1$，这时候显然解为

$$x_L^{(0)}=-1\,. \tag{3.56}$$

这是完全符合预期的。因为如果忽略小星体的质量，则质心位置基本与大星体位置重合，这点我们取为原点。小星体位于绕大星体距离为 1(在我们的特定的单位制中) 的圆轨道上。现在我们将 ε 很小但不等于零时的完整解表达为：

$$x_L=x_L^{(0)}+\delta x_L\,, \tag{3.57}$$

其中 δx_L 在 $\varepsilon\to 0$ 时也趋于零，我们将假设对于小的 ε 来说，

$$\delta x_L\sim\varepsilon^{\alpha}\,, \tag{3.58}$$

其中 $\alpha>0$ 是一个待定参数。将这些假设带入方程 (3.55) 并且只取领头阶的贡献我们发现 $\alpha=1/3$ 而相应的 δx_L 为，

$$\delta x_L=\pm\left(\frac{\varepsilon}{3}\right)^{1/3}\equiv\pm r_H\,, \tag{3.59}$$

其中 $\pm$ 分别对应于 L_1 和 L_2。也就是说 L_1 和 L_2 分别位于小质量星体的两侧大约相等的距离，该距离正比于质量比的 $1/3$ 次幂。

上面公式中出现的物理量，

$$\boxed{r_H=R\left(\frac{\varepsilon}{3}\right)^{1/3}=R\left(\frac{m_2}{3m_1}\right)^{1/3}\,,} \tag{3.60}$$

被称为两个星体系统的希尔半径 (Hill radius)。由于我们都生活在太阳系之中，因此太阳的存在被认为是缺省设置，常常被默认为是大质量物体。在这种情形下，我们又直接称 r_H 为小质量星体的希尔半径。以小质量星体的中心为原点，其希尔半径为半径的球体称为该小质量星体的希尔球 (Hill sphere)。粗略来说，希尔球刻画了太阳系中一个行星周围单独受其影响的引力作用范围。也就是说在距离一个行星的希尔半径范围内，我们可以假设一个质点仅仅受这个行星的引力影响而不会受到太阳或者其他行星的引力影响。以地球为例，它 (与太阳一起构成的双星系统) 的希尔半径大约是 0.01AU。这意味着在地球的希尔球体内的物体都将主要受到地球引力的影响。月球距离地球的距离大约为 38 万公里，或者说 $0.00257AU\simeq 0.257r_H$ 远小于地球的希尔半径。正因为如此我们可以将月球视为地球的一个卫星，而不是另一颗绕太阳运行的小一些的行星。

由于所有行星的质量都远小于太阳，因此所有行星的希尔半径都远小于它到太阳的距离，以我们上面取的单位制来说，这意味着 $r_H\ll 1$。到非平庸的第一阶，我们可以明确地写出 L_1，L_2 和 L_3 的位置如下，

$$x_{L_{1,2}}=-1\pm r_H\,,\quad x_{L_3}\simeq 1+\frac{5}{12}\varepsilon\,. \tag{3.61}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

对于日地系统而言，$\varepsilon \sim 3.0 \times 10^{-6}$，因此 L_1 和 L_2 位于距离地球大约 0.01AU 的地方。另外一个共线的拉格朗日点 L_3 基本上位于与地球隔着太阳相对的地方 (也就是地球半年之后或者半年之前所在的地方)。

¶ 下面再考虑 $y_L \neq 0$ 的拉格朗日点。方程 (3.49) 的第二个方程中可以两边消去一个 y_L 的因子。由于剩下的方程的对称性，对任何的解 y_L 来说，$\pm y_L$ 都是解。然后将该方程乘以 x_L 并与第一个方程相减就得到，

$$\begin{cases} 0 = \dfrac{-r_1 r_2}{[(x_L - r_1)^2 + y_L^2]^{3/2}} + \dfrac{r_1 r_2}{[(x_L + r_2)^2 + y_L^2]^{3/2}} \\ 1 = \dfrac{r_2}{[(x_L - r_1)^2 + y_L^2]^{3/2}} + \dfrac{r_1}{[(x_L + r_2)^2 + y_L^2]^{3/2}} \end{cases} \tag{3.62}$$

这是关于分母上两个根号的线性方程组。对其求解我们发现 (x_L, y_L) 必定满足

$$\begin{cases} (x_L - r_1)^2 + y_L^2 = 1 \,, \\ (x_L + r_2)^2 + y_L^2 = 1 \,. \end{cases} \tag{3.63}$$

也就是说，这两个解的点与 m_1 和 m_2 所在的位置恰好构成等边三角形。因此这是两个关于 x 轴对称的拉格朗日点，一般记为 L_4 和 L_5。具体来说，如果考虑质量较小的星体绕质量较大的星体旋转，位相上大约领先较小星体 60° 的拉格朗日点称为 L_4，另一个位相落后于较小星体大约 60° 的则称为 L_5。

¶ 总结来看，对于两个相互做匀速圆周运动的两个星体而言，它们的轨道平面内一共存在 5 个拉格朗日点：三个位于两个星体的连线上 (说的更具体些，一个位于两者连线的内侧，另外两个则在两侧)；另外两个位于轨道平面内与两星体呈等边三角形的对称位置上。历史上拉格朗日点 L_1，L_2 和 L_3 首先是由欧拉发现的 (1765 年)。随后拉格朗日发现了 L_4 和 L_5 的存在 (1772 年)。不过现在它们都被通称为拉格朗日点。L_1，L_2 和 L_3 由于位于两个星体的连线之上因此又被称为 共线的 (colinear) 拉格朗日点；L_4 和 L_5 则被称为 不共线 的拉格朗日点，又称为 等边的 (equilateral) 拉格朗日点，因为它们与两个星体恰好构成等边三角形。对于日地系统的这五个拉格朗日点我们也显示在图 3.2之中。

17.2 拉格朗日点附近有效势能

¶ 上面讨论的五个拉格朗日点处都是体系有效势能 $V_{\text{eff}}(\mathbf{x})$ 的极值点。在这些点处一个质量为 m 的质点感受到的有效力为零。下面我们讨论这些点附近的运动的稳定性问题。也就是说，如果所考虑的质点 m 偏离这些拉格朗日点一点，该质点是否还会在其附近振荡还是说会漂移走。这个问题实际上是一个多自由度系统的小振动问题。虽然通常的多自由度系统的小振动问题我们会在下一章才涉及 (具体说是第 20节)，不过我们这里还是可以简单介绍一下多自由度系统的小振动问题的梗概，以及一个质点在拉格朗日点附近运动与我们在第 20节中将要系统讨论的、一个通常的多自由度系统在其平衡位置附近小振动的异同。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

对于一般的一个惯性系中的多自由度系统来说，如果体系的拉格朗日量就是其动能减去势能，那么在其势能的极值点处体系所受的广义力为零。一般来说，如果该极值点是一个极小值点，也就是说其广义坐标偏离该点的任意无穷小位移都会使得势能上升，那么一般来说系统会在该点附近按照一定的频率做小振动。但是如果该极值点是一个鞍点，即沿某些方向是极小，沿另外一些方向是极大，那么体系的运动一般会越来越偏离该平衡位置。换句话说，只有对应于势能极小值点的平衡点才是稳定的。其他的情况一般都是不稳定的。拉格朗日点附近的运动最大的不同在于，除了有效的势能之外，非惯性系的科里奥利力也会起作用。我们将会看到，恰恰由于有依赖于速度的科里奥利力的存在，使得原先惯性系中的不稳定的点，甚至是势能的极大值点也可以是稳定的。

我们需要的是将有效势能 $V_{\text{eff}}(\mathbf{x})$ 在其极值点 $\mathbf{x}_L = (x_L, y_L)$ 附近展开到二阶。首先我们需要其在任意一点的一阶导数，

$$\begin{cases} \dfrac{1}{4\pi^2}\partial_x V_{\text{eff}} = -x + \dfrac{r_2(x-r_1)}{[(x-r_1)^2+y^2]^{3/2}} + \dfrac{r_1(x+r_2)}{[(x+r_2)^2+y^2]^{3/2}} \\ \dfrac{1}{4\pi^2}\partial_y V_{\text{eff}} = -y + \dfrac{r_2 y}{[(x-r_1)^2+y^2]^{3/2}} + \dfrac{r_1 y}{[(x+r_2)^2+y^2]^{3/2}} \end{cases} \tag{3.64}$$

如果令 $\partial_x V_{\text{eff}} = \partial_y V_{\text{eff}} = 0$，我们就回到了拉格朗日点所满足的方程 (3.49)。我们现在需要的是对上述公式再次求导以获得其二阶导数 (矩阵)。对于混合偏微商结果是，

$$\frac{1}{4\pi^2}\partial_x\partial_y V_{\text{eff}} = -\frac{3r_2(x-r_1)y}{[(x-r_1)^2+y^2]^{5/2}} - \frac{3r_1(x+r_2)y}{[(x+r_2)^2+y^2]^{5/2}} \tag{3.65}$$

类似地，相应的二阶偏微商为，

$$\begin{cases} \dfrac{1}{4\pi^2}\partial_x^2 V_{\text{eff}} = -1 + \dfrac{r_2[y^2-2(x-r_1)^2]}{[(x-r_1)^2+y^2]^{5/2}} + \dfrac{r_1[y^2-2(x+r_2)^2]}{[(x+r_2)^2+y^2]^{5/2}} \\ \dfrac{1}{4\pi^2}\partial_y^2 V_{\text{eff}} = -1 + \dfrac{r_2[(x-r_1)^2-2y^2]}{[(x-r_1)^2+y^2]^{5/2}} + \dfrac{r_1[(x+r_2)^2-2y^2]}{[(x+r_2)^2+y^2]^{5/2}} \end{cases} \tag{3.66}$$

现在我们需要各个拉格朗日点所满足的几何关系。我们下面会发现，三个共线的与两个不共线的拉格朗日点是有差别的，需要分开讨论。

对于共线的三个拉格朗日点来说，我们可以将 $y_L = 0$ 带入，而且 x_L 则需要满足方程 (3.52)。首先我们看到，对三个共线的拉格朗日点来说，其有效势能的二阶导数矩阵必定为对角的。其中 x 方向的的对角元为，

$$\frac{1}{4\pi^2}\partial_x^2 V_{\text{eff}} = -1 - \frac{2r_2}{|x_L-r_1|^3} - \frac{2r_1}{|x_L+r_2|^3} \equiv -\frac{\Omega_x^2}{4\pi^2} < 0\,, \tag{3.67}$$

这必定是负的。这意味着对于 x_L 的微小改变而言，三个共线的拉格朗日点都是有效势能的极大值。考察对于 y 方向的二阶偏微商我们发现，

$$\frac{1}{4\pi^2}\partial_y^2 V_{\text{eff}} = -1 + \frac{r_2}{|x_L-r_1|^3} + \frac{r_1}{|x_L+r_2|^3} \equiv +\frac{\Omega_y^2}{4\pi^2}\,, \tag{3.68}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

这个的正负需要利用 x_L 所满足的方程 (3.52) 来判断。这个方程告诉我们，

$$\left[-1+\frac{r_2}{|x_L-r_1|^3}+\frac{r_1}{|x_L+r_2|^3}\right]=\left(\frac{r_1r_2}{x_L}\right)\left[\frac{1}{|x_L-r_1|^3}-\frac{1}{|x_L-r_1|^3}\right]\,. \tag{3.69}$$

可以证明，上面这个表达式的右边对于 L_1，L_2 或者 L_3 来说都是正的。因此对于 y 方向的二阶导数而言，共线的三个拉格朗日点对应于有效势能的极小值。因此，它们实际上都对应于所谓鞍点的情形。这一点也显示在图 3.2之中。

如果我们运用小 ε 近似，我们可以对上述二阶导数的具体数值给出一个估计。仍然是对于三个共线的拉格朗日点来说，利用前面得到的估计，

$$x_{L_{1,2}}=-1\pm r_H\,,\ x_{L_3}=1+\frac{5}{12}\varepsilon\,. \tag{3.70}$$

我们可以得到如下的参数估计，

$$\begin{aligned}&\left(\Omega_x^2\right)_{L_{1,2}}=36\pi^2\left(1\pm\frac{2}{3})r_H\right)\,,\left(\Omega_x^2\right)_{L_3}=12\pi^2\left(1+\frac{7}{12}\varepsilon\right)\,.\\&\left(\Omega_y^2\right)_{L_{1,2}}=12\pi^2(1\pm r_H)\,,\left(\Omega_y^2\right)_{L_3}=\frac{7\pi^2}{32}\varepsilon\,.\end{aligned} \tag{3.71}$$

对于两个不共线的拉格朗日点 L_4 和 L_5 来说，其二阶导数矩阵并不是对角的。我们需要它们附近的下列几何关系：

$$x_L-r_1=-\frac{1}{2},\ x_L+r_2=\frac{1}{2},\ y_L=\frac{\sqrt{3}}{2}\,. \tag{3.72}$$

于是位于拉格朗日点 L_4 和 L_5 处的二阶导数矩阵可以写为，

$$\left[\begin{matrix}\partial_x^2V_{eff}&\partial_x\partial_yV_{eff}\\\partial_x\partial_yV_{eff}&\partial_y^2V_{eff}\end{matrix}\right]\Bigg|_{\mathbf{x}_L}=-3\pi^2\left[\begin{matrix}1&\sqrt{3}\lambda\\\sqrt{3}\lambda&3\end{matrix}\right]\,, \tag{3.73}$$

其中

$$\lambda=\frac{m_1-m_2}{m_1+m_2}=\frac{1-\varepsilon}{1+\varepsilon}\,, \tag{3.74}$$

是一个绝对值小于 1 的参数。因此上述矩阵实际上是负定的，也就是说不共线的两个拉格朗日点对于任何的偏离来说都对应于有效势能的极大值而不是极小值。

因此，如果我们令

$$\mathbf{x}=\mathbf{x}_L+\mathbf{X}\,, \tag{3.75}$$

其中 $\mathbf{x}_L$ 是某个拉格朗日点的位移矢量，$\mathbf{X}$ 则是偏离拉格朗日点的位移。如果我们要求 $|\mathbf{X}|\ll|\mathbf{x}_L|$。在这些拉格朗日点附近我们有，

$$V_{\text{eff}}(\mathbf{x})=V_{\text{eff}}(\mathbf{x}_L)+\delta V_{\text{eff}} \tag{3.76}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

其中偏离拉格朗日点的有效势能的修正可以写为：

$$\begin{aligned}\delta V_{\text{eff}} &= +\frac{1}{2}\mathbf{X}^T\begin{pmatrix} -\Omega_x^2 & 0 \\ 0 & +\Omega_y^2 \end{pmatrix}\cdot\mathbf{X}\,,\\ \delta V_{\text{eff}} &= -\frac{3\pi^2}{2}\mathbf{X}^T\begin{pmatrix} 1 & \sqrt{3}\lambda \\ \sqrt{3}\lambda & 3 \end{pmatrix}\cdot\mathbf{X}\,,\end{aligned} \tag{3.77}$$

其中前一个式子适用于三个共线的拉格朗日点 (L_1，L_2 和 L_3) 而后一个式子则适用于 L_4 和 L_5。这里的 Ω_x^2 和 Ω_y^2 都是正实数并且由公式 (3.71) 给出。

17.3 拉格朗日点附近的运动模式

¶本小节中我们将讨论拉格朗日点附近质点的运动模式，这又将分为三个共线的拉格朗日点和两个不共线的拉格朗日点来分别进行处理。

在三个共线的拉格朗日点附近的运动方程可以写为，

$$\ddot{\mathbf{X}} + 4\pi\hat{\mathbf{z}}\times\dot{\mathbf{X}} + \begin{pmatrix} -\Omega_x^2 & 0 \\ 0 & +\Omega_y^2 \end{pmatrix}\cdot\mathbf{X} = 0\,, \tag{3.78}$$

其中我们已经带入了 $\boldsymbol{\Omega} = 2\pi\hat{\mathbf{z}}$。现在我们可以假定如下的尝试解，

$$\mathbf{X}(t) = \mathbf{V}e^{\nu t}\,, \tag{3.79}$$

其中 $\mathbf{V}$ 是一个常矢量，ν 是待定的本征值，我们有，

$$\begin{pmatrix} \nu^2 - \Omega_x^2 & 4\pi\nu \\ -4\pi\nu & \nu^2 + \Omega_y^2 \end{pmatrix}\cdot\mathbf{V} = 0\,,$$

要求 $\mathbf{V}$ 有非零解我们必须要求上面矩阵的行列式为零。这给出，

$$(\nu^2 - \Omega_x^2)(\nu^2 + \Omega_y^2) + 16\pi^2\nu^2 = 0\,. \tag{3.80}$$

或者等价地写为，

$$(\nu^2)^2 + (\Omega_y^2 - \Omega_x^2 + 16\pi^2)\nu^2 - \Omega_x^2\Omega_y^2 = 0\,. \tag{3.81}$$

这意味着我们可以解出 ν^2，它一定是实的。并且一定是一个正实数，一个负实数，因为两者的乘积是小于零的。如果取小质量星体的质量趋于零的极限，即 $\varepsilon\to 0$，那么这两个频率的平方的解为，

$$\nu^2 \simeq 4\pi^2(1\pm\sqrt{28})\,. \tag{3.82}$$

小于零的根会给出纯虚的指数因子因此是一个震荡解；但是大于零的一个解会给出随时间指数增加的 (以及随时间指数衰减的) 解。由于一般来说很难控制质点的初始条件使得它仅仅包含震荡的解，因此质点的运动一般来说是不稳定的。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

上面的讨论仅仅局限于质点在轨道平面内的运动。事实上，要更为全面的理解在三个共线的拉格朗日点附近的质点的运动，需要将其垂直于轨道平面的运动也考虑进来。在 1968 年的博士论文中，美国人 Robert W. Farquhar 首先考虑了地月系统 L_2 点附近的所谓晕轨道 (halo orbit)。[14] 这是一个在 L_2 点附近的三维周期轨道。当时美国正在进行 Appolo 登月计划。Farquhar 当时的研究是希望利用地月的 L_2 点来进行远程的通信，因为这个点可以同时连接地球和月球的背面。这可以作为后续在月球背面着陆的先期准备。不过这个计划后来并没有真正实施，美国也没有再推进诸如在月球背面登月的计划。最近，中国的嫦娥登月计划也准备利用类似的设计在月球背面登月。

与晕轨道类似的另一类三维的轨道称为李萨如轨道 (Lissajous orbit)。与晕轨道不同的是，这类轨道并不是严格周期的，而只是准周期的。该轨道在两个星体轨道平面内的投影按照所谓的李萨如图形进行，因此得名。总体来说，在共线的拉格朗日点 L_2 附近，一般需要飞行器进行少许轨道位置调整 (orbital station-keeping)。这类轨道已经多次运用在日地的拉格朗日点 L_2 附近。最为著名的是欧洲航天局发射的 Planck 卫星。其主要任务是探测宇宙微波背景辐射 (cosmic microwave background, CMB) 的各向异性。CMB 可以看成是宇宙早期的一张照片，为我们了解宇宙的起源提供了重要的信息。

总体来说，三个共线的拉格朗日点并不是稳定的，但是在它们的附近人造的星体只需要较小的能量消耗就可以回复到平衡点的附近。同时，由于这些拉格朗日点距离地球比较近 (相较于不共线的两个拉格朗日点 L_4 和 L_5 而言)，因此也被广泛地运用在空间探索之上。例如我国的探月工程就运用了地月系统的拉格朗日点 L_2。有关世界上各个国家正在开展或者拟开展的计划的信息读者可以请参考`https://en.wikipedia.org/wiki/Lagrangian_point`。

¶ 在两个不共线的拉格朗日点 L_4 和 L_5 附近，质点的运动方程可以表达为，

$$\ddot{\mathbf{X}} + 2\boldsymbol{\Omega} \times \dot{\mathbf{X}} - 3\pi^2 \begin{bmatrix} 1 & \sqrt{3}\lambda \\ \sqrt{3}\lambda & 3 \end{bmatrix} \cdot \mathbf{X} = 0 \,. \tag{3.83}$$

我们假定指数形式的解：$\mathbf{X}(t) = \mathbf{V}e^{\kappa(2\pi t)}$，我们发现 κ 满足，

$$\kappa^4 + \kappa^2 + \frac{27}{16}(1 - \lambda^2) = 0 \,. \tag{3.84}$$

对于多数有实际应用的系统来说，λ 的数值非常接近于 1。例如对于地月系统我们有，

$$1 - \lambda^2 = \frac{4m_1 m_2}{(m_1 + m_2)^2} \simeq 4\varepsilon \simeq 0.04908$$

设其为小量于是我们获得 κ^2 的两个解为，

$$\kappa_+^2 = -\frac{27}{4}\varepsilon \,, \quad \kappa_-^2 = -1 + \frac{27}{4}\varepsilon \,. \tag{3.85}$$

[14] Farquhar 后来成为 NASA 著名的任务设计师，参与了多项空间探索计划。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

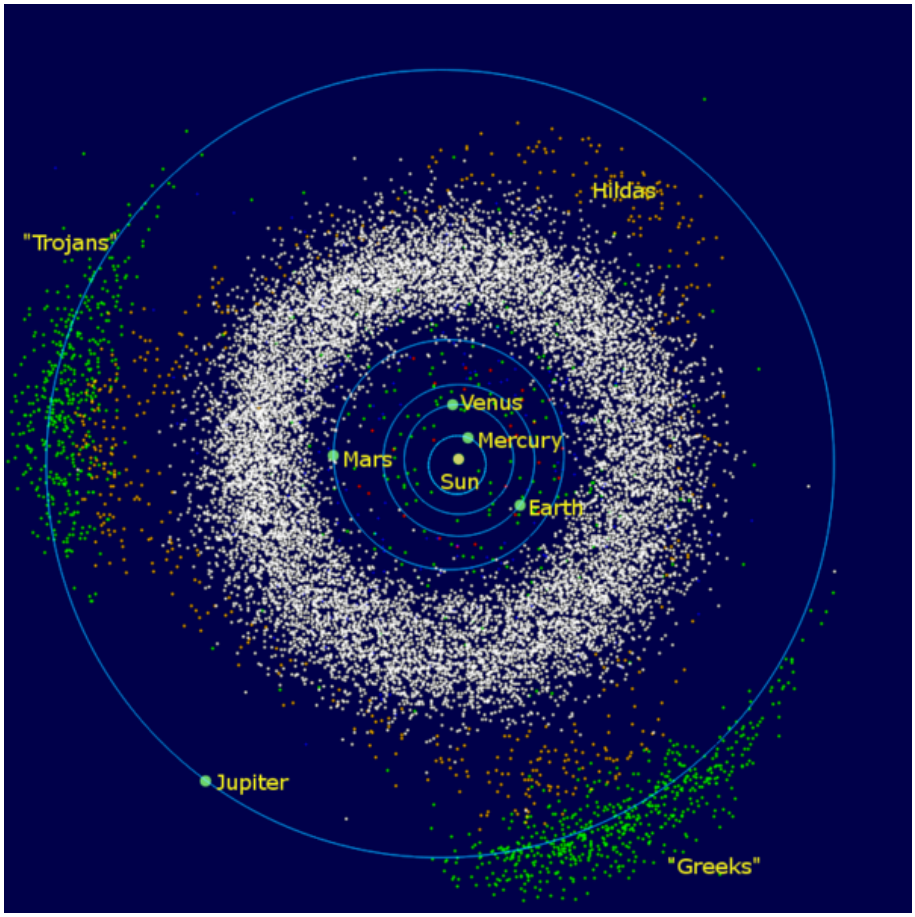

图 3.3: 太阳系中太阳 -木星系统中位于拉格朗日点 L_4 和 L_5 附近的特洛伊星体 (包括特洛伊阵营和希腊阵营) 示意图。图片取自https://en.wikipedia.org/wiki/Trojan_(astronomy).

这两个解都是小于零的实数。因此，相应的 κ 都是纯虚的，这说明 $\mathbf{X}(t)$ 存在振荡的解，其振荡的频率分别为 Ω 和 0.288Ω，并没有趋于无穷的解。换句话说，这时的拉格朗日点 L_4 和 L_5 竟然对小的扰动是稳定的。这是一个相当奇葩的结果。我们看到，尽管拉格朗日点 L_4 和 L_5 是有效势能的极大值点，但是由于科里奥利力的作用，它们确实对于小的扰动是稳定的。这一点也可以从运动方程 (3.83) 中看出。如果仅仅有第三项 (来源于有效势能对坐标的梯度)，方程的解显然是不稳定的。但是第二项科里奥利力的存在使得最后的解是稳定的。

¶ 拉格朗日点 L_4 和 L_5 对于小的扰动稳定的特性使得它们的附近有可能出现大量的小行星。太阳系之中由于木星是最大质量的行星，因此在太阳 -木星的轨道平面的拉格朗日点 L_4 和 L_5 处就聚集了数目众多的小行星。这些小行星在日木旋转参照系中的运动基本上是稳定的，其振荡的圆频率可以由上面的估计给出。两类圆频率分别为 Ω 和

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

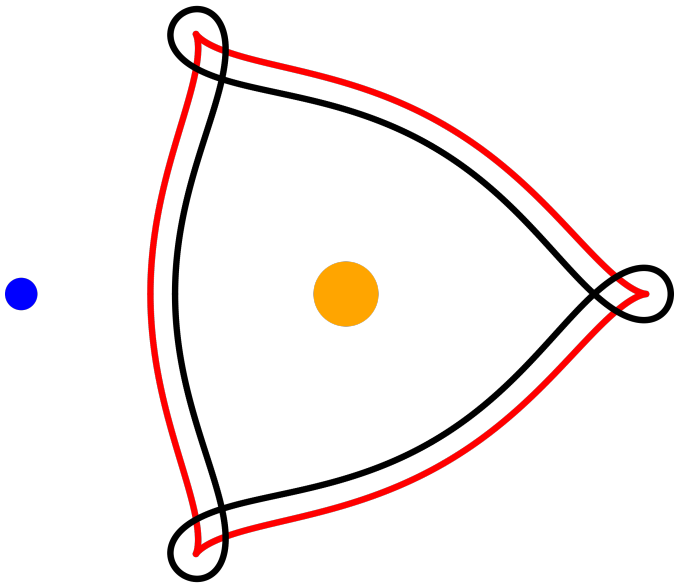

图 3.4: 太阳 -木星旋转系统中所看到的希尔妲星体的运动轨迹示意图。其中金黄色的大圆点代表太阳，蓝色的稍小的圆点代表木星。两者是固定不动的。黑色和红色的曲线代表在木星看来希尔妲星体的运动轨迹，其中黑色和红色的线所对应的偏心率分别为 0.31 和 0.21。图片取自https://en.wikipedia.org/wiki/Hilda_asteroid.

0.08Ω，这里 Ω 是木星绕太阳的圆频率。[15] 这些小行星一般通称为特洛伊星体 (Trojan)。具体到太阳和木星的系统来说，位相领先于木星位于拉格朗日点 L_4 附近的小行星被称为 (特洛伊星体的) 希腊阵营 (Greek camp)；位相落后于木星位于拉格朗日点 L_5 附近的特洛伊星体则被称为 特洛伊阵营 (Trojan camp)。相应的，这两个阵营中的小行星也分别按照各自的阵营中参与了当年特洛伊战争的英雄的名字来命名：例如希腊营中的阿克琉斯 (588 Achilles) 以及特洛伊营中的帕里斯 (3317 Paris) 等等。具体的构型参见图 3.3。目前为止，两个阵营中都发现了数千个小行星，其总数与太阳系中位于火星和木星之间的所谓 小行星带 (asteroid belt) 中的小行星数目相当。一般认为，在这两个拉格朗日点附近直径超过 1 公里的小行星会超过 1 百万颗。不仅仅木星拥有特洛伊星体，其他的行星 (包括地球) 也拥有自己的特洛伊星体，只不过数目远没有木星那么夸张罢了。

为什么木星的拉格朗日点附近有这么多小行星而其他没有呢？这个也可以从前面的估计中看出来。其中一个圆频率约等于相应系统的圆频率 Ω，另一个的圆频率为 $\sqrt{27\varepsilon/4}\Omega$，其中 ε 是行星与太阳的质量比。我们看到，如果 $\varepsilon \to 0$ 那么其中一个圆频率趋于零。这意

[15] 其中我们运用了木星与太阳的质量比：$M_{♃}/M_{☼} = 1/1047$。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

味着如果拉格朗日点附近有小星体，它很可能会有零频的分量从而使得该星体漂移走。木星由于其质量比较大，因此在太阳系中木星具有最大的 ε 数值。这就是木星的拉格朗日点附近聚集了最多的小行星的一个重要原因。

¶围绕着木星的三个拉格朗日点 L_3，L_5 和 L_4 还存在着另外一类小型星体，它们的周期是木星的 2/3，它们的偏心率也较大，可以达到 0.3 左右。由于其半长轴一般在 3.7-4.2AU 左右，因此在近日点可以达到 2.7AU 左右。这时候也是地球上观察它们的好时机。它们的远日点会依次经过上述三个拉格朗日点。这类小行星一般统称为希尔妲家族 (Hilda family)，或者希尔妲小行星。这个家族也包含了几千颗小性星体。[16] 图 3.3中也显示了希尔妲家族的一系列小行星。不过更为形象的显示这些小行星以及木星关系的是下面脚注中给出的链接中的动图。[17] 需要注意的是，在太阳的参考系来看，所有的星体都在围绕太阳做 Kepler 运动，包括我们前面提到的特洛伊小行星以及希尔妲小行星。只不过在木星保持不动的参照系看来，特洛伊小行星只是在两个拉格朗日点附近做往复的小振动，而希尔妲星体则是沿着一个相当规则但有些复杂的曲线运动，如图 3.4所示。由于周期是木星周期的 2/3，因此在木星的两年 (即两个周期) 之中，一个希尔妲星体会沿着它的轨迹走三个周期。

18 相对论效应造成的近日点进动

¶这一节中我们讨论相对论性的修正对于开普勒问题的影响。我们这里将仅仅讨论狭义相对论的修正，没有考虑广义相对论的修正。在第 15节结束时我们指出，只要一个粒子感受到的势能只要稍微偏离 $1/r$ 的形式，它的轨道就不会是闭合的，其近日点就会进动。事实上，如果我们考虑狭义相对论的修正，当我们将粒子的运动方程与开普勒问题的运动方程比较时，我们就会发现它有微小的改变，这种改变就会造成星体近日点的进动。我们这一节就来简要讨论一下这个问题的一个近似解。在狭义相对论框架下讨论行星的运动往往被称为相对论性开普勒问题。这个问题在经典力学中是可以严格求解的。首先获得这个解的是索末菲 (Sommerfeld)。我们将在后面利用哈密顿力学讨论这个解，参见第 29节的例题 6.2。这里我们将仅仅满足于这个问题的一个近似解。[18] 为了简化中间的计算，我们将采用光速为速度的单位，这等效于在以往的公式中令 $c=1$。

在狭义相对论中，粒子的拉格朗日量的非相对论展开 (在扣除了常数 $-mc^2=-m$ 之后) 为：

$$L=\frac{m}{2}\mathbf{v}^2\left(1+\frac{1}{4}\mathbf{v}^2+\cdots\right)+\frac{\alpha}{r}\,, \tag{3.86}$$

[16] 命名自其中的一颗小行星 153 Hilda。这个名字是奥地利的著名天文学家 Theodor von Oppolzer 用自己女儿的名字来命名的。我因此冒昧地将其翻译为希尔妲。

[17] 参见：`https://www.exploremars.org/trojan-asteroids-around-jupiter-explained`。

[18] 就实际的数值上而言，其实近似解已经足够好了，因为这些相对论修正实在是小得令人发指啊。因此更高阶的修正完全没有实际的必要。不过如果能够获得完全解，为什么不呢？

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

其中的 $\cdots$ 包含了 $\mathbf{v}^2$ 的更高次幂。它们都是更高阶的相对论修正。由于我们仅仅感兴趣最低阶的修正，我们可以将它们忽略掉。需要注意的是，这个问题仍然是一个平面的二维问题。因此，在利用极坐标 (r,ϕ) 表达后 $\mathbf{v}^2$ 的形式为，

$$\mathbf{v}^2 = \dot{r}^2 + r^2\dot{\phi}^2 \,. \tag{3.87}$$

我们定义扣除了粒子静止能量 $mc^2 = m$ 之后的能量为 E。那么我们的能量 E 角动量 $p_\phi \equiv J$ 仍然是守恒量。对于角动量而言我们有：

$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} \equiv J \simeq mr^2\dot{\phi}\left(1+\frac{\mathbf{v}^2}{2}\right) \,, \tag{3.88}$$

类似地，我们得到与 r 共轭的动量为，

$$p_r = \frac{\partial L}{\partial \dot{r}} \simeq m\dot{r}\left(1+\frac{\mathbf{v}^2}{2}\right) \,, \tag{3.89}$$

而粒子的能量可以写为：

$$\begin{aligned} E &= \dot{r}p_r + \dot{\phi}p_\phi - L = m\mathbf{v}^2(1+\mathbf{v}^2/2) - L \\ &= \frac{1}{2}m\mathbf{v}^2\left(1+\frac{3}{4}\mathbf{v}^2\right) - \frac{\alpha}{r} \,. \end{aligned} \tag{3.90}$$

由于我们仅仅感兴趣轨道的形状而不是时间依赖关系，因此我们将上式换成 $dr/d\phi$ 来表达，其中的 $\dot{\phi}$ 则必须用守恒量 J 来表达。按照上面的公式 (3.88)，我们有，

$$\dot{\phi} = \frac{J}{mr^2}\left(1-\frac{\mathbf{v}^2}{2}\right) \,. \tag{3.91}$$

注意，这个公式原则上并没有将 $\dot{\phi}$ 解出来，因为 $\mathbf{v}^2$ 里面仍然有 $\dot{\phi}$。不过将它看成是一个近似的解是没有问题的，其中相对论修正的部分则应当用问题的零级近似 –也就是没有任何相对论修正时开普勒问题的解 –带入。下面各个式子中的相对论修正原则上也应当这样处理。另一方面由于，

$$\frac{dr}{dt} = \frac{dr}{d\phi}\frac{d\phi}{dt} \,,$$

将其中的 $d\phi/dt = \dot{\phi}$ 用前面的表达式 (3.91) 表达并换成变量 $u = 1/r$，我们就得到：

$$\dot{r} = -\frac{J}{m}\frac{du}{d\phi}\left(1-\frac{\mathbf{v}^2}{2}\right) \,. \tag{3.92}$$

因此，我们可以得到粒子动能的近似表达式

$$\frac{1}{2}m\mathbf{v}^2 = \frac{J^2}{2m}\left[\left(\frac{du}{d\phi}\right)^2 + u^2\right]\left(1-\mathbf{v}^2\right) \,, \tag{3.93}$$

再利用前面的能量表达式 (3.90)，我们就得粒子总能量的形式：

$$E = \frac{J^2}{2m}\left[\left(\frac{du}{d\phi}\right)^2 + u^2\right] - \frac{1}{8}m(\mathbf{v}^2)^2 - \alpha u \,. \tag{3.94}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

现在注意到这个能量表达式中只有第二项是相对论修正，它是非常小的量。按照我们前面的理解，我们应当利用零级近似的 $\mathbf{v}^2$ 的表达式来替代它。我们利用：

$$\frac{1}{2}m\mathbf{v}^2 \simeq E + \frac{\alpha}{r} = E + \alpha u \,, \tag{3.95}$$

来代换相对论修正项中的 $\mathbf{v}^2$，我们就得到相对论修正为，

$$-\frac{m}{8}(\mathbf{v}^2)^2 = -\frac{1}{2m}(E+\alpha u)^2$$

将这个关系代入到包含相对论修正的能量表达式 (3.94) 之中，相比较零级近似，我们多了三项：一个常数项，一个线性依赖于 u 的项，一个正比于 u^2 的项。常数项仅仅改变能量的定义，线性依赖于 u 的项仅仅改变轨道的尺度，只有正比于 u^2 的项会改变轨道的闭合性。零级近似下轨道之所以闭合，主要的原因就是上面这个微分方程中 $(du/d\phi)^2$ 的系数与 u^2 的系数完全相同。只要这点成立，轨道必定是闭合的椭圆。[19] 但是相对论修正中正比于 u^2 的项将改变 u^2 的系数从而使轨道不再闭合。具体来说这时得到的表达式为：[20]

$$E\left(1+\frac{E}{2mc^2}\right) + \alpha u\left(1+\frac{E}{mc^2}\right) = \frac{J^2}{2m}\left[\left(\frac{du}{d\phi}\right)^2 + \left(1-\frac{\alpha^2}{J^2c^2}\right)u^2\right] \,. \tag{3.96}$$

如果我们将这个结果与完全非相对论的结果比较，我们立刻就可以发现行星运行一周的进动角为

$$\boxed{\delta\phi = \frac{\pi\alpha^2}{J^2c^2} = \frac{\pi GM}{a(1-e^2)c^2} \,.} \tag{3.97}$$

这个狭义相对论的修正比起广义相对论的结果要小。事实上只有广义相对论结果的 1/6。需要特别指出的是，无论是广义相对论还是狭义相对论，它们预言的水星近日点的进动都是非常小的。例如，广义相对论的数值大约是每个世纪进动 43”。这个数值远比实际上观测到的水星近日点的总进动值（大约每世纪 570”）要小。正如我们在第 15节的例题 3.3中看到的，对于水星的近日点的进动最主要作用的并不是相对论的修正，而是其他行星对水星的影响，特别是金星、木星和地球，它们三个加起来占总观测量的 91%。人们说水星近日点的进动是广义相对论的一个经典检验是因为在扣除了其他行星的影响以后，水星近日点每个世纪仍然有大约 43” 的进动无法用其他已知的理论解释，恰恰是这一部分修正与广义相对论的预言十分接近。



相关的阅读

[19] 最简单的做法是将该方程再对 ϕ 求一次导数，我们就得到一个频率为 1 的谐振子方程：$u'' + u = \mathrm{C}$，其中 C 为一常数。

[20] 为了便于比较，我们在这个表达式中恢复了光速 c 的恰当幂次。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

本章中我们讨论了有心力场中的运动。这是一个十分古老的问题，它几乎伴随着牛顿力学的诞生。由于最为基本的开普勒问题大家都已经熟知（如果你还不熟悉，可以参考任意一本普通物理力学的教科书），我们仅仅是作了简单的总结性介绍（利用拉格朗日分析力学的语言）。我们着重讨论了一些稍微复杂的应用，例如开普勒问题的相对论修正；潮汐问题等等。事实上，在天体力学中，讨论各种因素对于行星轨道的摄动（也就是微扰）是一个很实际的应用。由于篇幅的限制，我们这里没有很详细地讨论。但在第 15节的例题 3.3中，我们简要地讨论它对水星近日点进动的影响。

从本章各节的具体内容来说，第 14 节和第 15 节所讨论的内容可以在朗道书 [1] 的第 §13，§14，§15 中找到，或者参考 Goldstein 书 [3] 的第三章。本章中关于潮汐的讨论（第 16 节）取自某个力学的习题（具体出处已不可考）。第 17节关于受限三体问题的讨论是 2018 年新加入的一节，内容相当丰富。加入这一节的一个主要的原因是近年来关于中国航天计划的报道之中曾多次提及拉格朗日点。我想这里对相关的现象做一个介绍是恰当的，而且其处理方法与前面潮汐的讨论非常接近，都需要转换到两个相对做匀速圆周运动的星体的转动参照系 (一个非惯性系) 之中。虽然关于小振动的讨论下一章才真正涉及，不过这里简要讨论一下各个拉格朗日点附近的运动对于多数同学还是比较好接受的。在狭义相对论框架下关于水星近日点进动的讨论（第 18 节）在朗道经典场论的书 [2] 中有更严格的处理，只不过那里他运用了哈密顿 -雅克比理论的方法（见本讲义的第 34 节）。这里给出的是经过我自己“改造”的、更加简单、初等的处理方法。



由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

第四章　　小振动

本章提要

- 一维谐振子 (19)
- 多自由度系统的简正模 (20)
- 非线性振动 (21)
- 参数共振 (22)

小振动是力学中十分常见的运动形式。一个力学体系，只要它偏离其稳定平衡位置很小，那么它的典型运动就是小振动。在位移不太大的情况下，这种振动基本上可以看成是简谐振动。对于一个多自由度的体系，我们可以选取系统自由度的适当线性组合构成的简正坐标，这些简正坐标的简谐振动是相互独立的，并且它们的振动频率也是由系统的基本力学性质所决定的，它们因此被称为系统的本征频率。力学体系也可以在周期外力的作用下参与受迫振动，这时如果外力的频率与系统的某个本征频率几乎相同，系统的相应简正模就会发生共振。共振也可以发生在一个力学系统的参数周期变化的情形下。这些都是我们这一章中希望介绍的内容。经典力学中讨论的小振动可以十分直接地推广到量子的情形，当然本章中我们仅仅介绍小振动的经典理论。

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

19 一维振子

¶ 我们首先来复习一下一个一维的振子运动规律。[1] 对于一个没有阻尼的一维振子，它的拉格朗日量为：

$$L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2 \,, \tag{4.1}$$

这里我们假定粒子的平衡位置是 $x = 0$。它的运动方程可以表达为：

$$\ddot{x} + \omega^2 x = 0 \,. \tag{4.2}$$

其中 $\omega = \sqrt{k/m}$ 为振子的频率。这个微分方程的解可以普遍写为：

$$x(t) = x(0)\cos\omega t + \frac{v(0)}{\omega}\sin\omega t \,, \tag{4.3}$$

其中 $x(0)$ 和 $v(0)$ 代表粒子的初始位置和初始速度。这个解也可以等价地写成：

$$x(t) = A\cos\left(\omega t + \phi\right) \,, \tag{4.4}$$

其中 A 称为小振动的振幅。一个谐振子的总能量可以利用振幅表达为：

$$E = \frac{1}{2}m\omega^2 A^2 \,. \tag{4.5}$$

它是谐振子的一个运动积分 (守恒量)。

¶ 如果一个一维运动的振子除了感受到其平衡位置附近的谐振子势能外，还受到已知外力 $F(t)$ 的作用，这样的振子被称为受迫振子，相应的振动行为被称为受迫振动。这时粒子的拉格朗日量可以表达为：

$$L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2 + xF(t) \,. \tag{4.6}$$

而相应的运动方程为：

$$\ddot{x} + \omega^2 x = F(t) \,. \tag{4.7}$$

这个非齐次方程的通解可以表达为一个特解再加上原来齐次方程 (4.2) 的通解。如果外加的力 $F(t)$ 也是周期的：

$$F(t) = f\cos(\Omega t + \beta) \,, \tag{4.8}$$

我们可以寻找形如：$\cos(\Omega t + \beta)$ 的特解。我们得到的结果是：

$$x(t) = A\cos\left(\omega t + \phi\right) + \frac{f}{m(\omega^2 - \Omega^2)}\cos(\Omega t + \beta) \,. \tag{4.9}$$

[1] 这些内容大家应当在普通物理力学课程中接触过，所以我们只是利用分析力学的语言简单总结一下。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

当外力的频率与振子的固有频率十分接近时，这个系统呈现出共振的行为。这时：

$$x(t) = A\cos(\omega t + \phi) + \frac{ft}{2m\omega}\sin(\omega t + \beta)\ , \tag{4.10}$$

也就是说，粒子的振幅会随时间线性增加，直到小振动的假定不再成立 (从而上面的方程都需要修改)。

¶前面讨论的简谐振子拉格朗日量具有时间反演不变性。也就是说，粒子的拉格朗日量只包含其速度的二次幂。如果系统的时间反演对称性被破坏，例如存在阻尼的情况下，这时严格来说已经不是一个纯力学问题了。因为阻尼一定伴随着机械能向其他形式的能量 (这时是热能) 的转化。因此，一个有阻尼的运动问题一般来说不可能仅仅从粒子的拉格朗日量得到。在耗散 (阻尼) 比较轻微的时候，这类系统的运动可以用系统的拉格朗日量 (纯力学的) 和一个瑞利耗散函数 $\mathcal{F}$(热力学的) 来描写 (参见第 一章第 5节中的讨论)。瑞利耗散函数的两倍代表了力学系统的机械能转变为其他形式能量的功率。对于一个一维的振子，它的瑞利函数可以取为：

$$\mathcal{F} = \gamma m\dot{x}^2\ . \tag{4.11}$$

于是，一维阻尼振子运动方程 (参考第 5节的公式 (1.30)) 变为：

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = 0\ . \tag{4.12}$$

这个方程的解可以分为三大类情况进行讨论。如果 $\gamma < \omega_0$，它的解为：

$$x(t) = Ae^{-\gamma t}\cos(\omega t + \phi)\ , \tag{4.13}$$

这是一个典型的阻尼振动的解。它的振动频率为：

$$\omega = \sqrt{\omega_0^2 - \gamma^2}\ . \tag{4.14}$$

如果 $\gamma > \omega_0$，方程 (4.12) 的解将不再是周期函数，而是随时间指数衰减的函数。这种情况称为过阻尼。如果阻尼参数 $\gamma = \omega_0$，那么方程 (4.12) 的解是一个指数衰减的函数再乘以一个线性增长的函数。这种情况被称为临界阻尼。

20 多自由度系统的小振动与简正模式

¶现在我们讨论多个自由度力学系统的小振动情况。考虑一个一般的保守系统：[2]

$$L = \frac{1}{2}a_{ij}(q)\dot{q}_i\dot{q}_j - U(q)\ ,$$

[2]我们这里假定重复的指标隐含着求和，即启用所谓的爱因斯坦求和约定。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

我们假定系统在其稳定平衡位置 q_{i0} 附近运动。我们可以令：$x_i = q_i - q_{i0}$ 为小量并将势能对其展开。由于 q_{i0} 是系统的稳定平衡位置，我们一定有：

$$U(q) \simeq U(q_0) + \frac{1}{2}k_{ij}x_i x_j \ .$$

另一方面，动能的系数 $a_{ij}(q)$ 中，在第一级近似我们可以取 $q_i = q_{i0}$。因此，一个多自由度的保守系统在其稳定平衡位置附近进行小振动的拉格朗日量可以写成：

$$\boxed{L = \frac{1}{2}m_{ij}\dot{x}_i\dot{x}_j - \frac{1}{2}k_{ij}x_i x_j \ .} \tag{4.15}$$

为了写得更为紧致一些，我们可以引入矩阵和矢量的符号。记 $m, k \in \mathbb{R}^{n\times n}$ 为 $n\times n$ 的实矩阵，它们的矩阵元分别为 m_{ij} 和 k_{ij}。相应的矢量 $x = (x_1, \cdots, x_n)^T \in \mathbb{R}^n$。那么体系的拉格朗日量可以写为

$$L = \frac{1}{2}\dot{x}^T m\dot{x} - \frac{1}{2}x^T kx \ . \tag{4.16}$$

我们下面会分别按照矩阵形式或者分量形式来写这些方程。

这个系统的运动方程为：

$$\boxed{m_{ij}\ddot{x}_j + k_{ij}x_j = 0 \ , \quad m\cdot\ddot{x} + k\cdot x = 0 \ .} \tag{4.17}$$

如果我们假定这些微分方程组的解的形式为：$x(t) = \eta e^{-i\omega t}$ 带入到方程 (4.17) 中，要求其存在非零解就得到：

$$\boxed{\det[k - \omega^2 m] = 0 \ .} \tag{4.18}$$

也就是说，系统振动的 本征频率 ω^2 是作为上述方程的 本征值 求出的。这个方程有时又被称为 久期方程 (secular equation)。对于每一个本征值 ω^2，都有一个相应的 本征矢量，这些本征矢量实际上就按照相应的本征频率进行简谐振动。线性代数的知识告诉我们，不同的本征频率对应的本征矢量一定是线性无关的。当存在某两个本征频率相等时，需要更为细致的考虑。可以证明这时仍然可以找到线性无关的一组完备的本征矢量。这些本征矢量实际上与系统的 简正坐标 是联系在一起的。

¶ 另外一种观点来考察简正坐标是从将体系的拉格朗日量 (4.15) 这个二次型对角化的角度来分析。这里我们要利用线性代数中一个熟知的定理：

定理 4.1 **任何一个厄米的矩阵总可以通过一个幺正变换将其对角化。利用到实对称矩阵的情形就是，任何一个实对称矩阵，总可以找到一个正交变换将其对角化。**

因此，我们总可以通过一个正交变换 $x = P_1\cdot y$ 使得动能的二次型变为对角的：

$$T = \frac{1}{2}\sum_{i=1}^{n} m_i\dot{y}_i^2 \ . \tag{4.19}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

由于动能总是正定的，因此 $m_i > 0$ 是正的实数。注意，在这个变换下势能：$V = (1/2)y^T K' y$，其中 $K' = (P_1)^T k P_1$ 仍然是实对称的正定的矩阵。我们现在可以令 $z_i = \sqrt{m_i} y_i$，或者等价地写为 $z = (M_D)^{1/2} y$，其中 M_D 是对角的矩阵，其矩阵元就是上述各个 m_i。这样一来动能部分变为 $T = (1/2)z^T z$，而势能部分则由一个实对称正定矩阵 K'' 描写：

$$T = \frac{1}{2} z^T z \ , \quad V = \frac{1}{2} z^T K'' z \ , \tag{4.20}$$

其中 $K'' = M_D^{-1/2} K' M_D^{-1/2}$。由于 K'' 仍然是实对称矩阵，我们可以进一步寻找将 K'' 对角化的正交矩阵 P_2。也就是说，我们令，

$$z = P_2 \cdot Q \ . \tag{4.21}$$

并要求它将势能对角化。由于势能部分也是正定的，不失一般性我们令其对角元为 $\omega_i^2 > 0$。由于动能部分已经化为单位矩阵的形式，因此它在任何的正交矩阵变换下是不变的，仍然保持原形式。在上述一系列变换下–准确地说，$Q = (P_2)^T M_D^{1/2} (P_1)^T \cdot x = A^{-1} \cdot x$–我们将动能和势能同时对角化了：

$$\boxed{L = \sum_{i=1}^{n} \frac{1}{2} \left(\dot{Q}_i^2 - \omega_i^2 Q_i^2 \right) \ ,} \tag{4.22}$$

其中的矩阵 $A = P_1 M_D^{-1/2} P_2$ 通常被称为**模态矩阵** (modal matrix)。简正坐标与原先坐标之间的关系为，

$$x = (P_1 M_D^{-1/2} P_2) \cdot Q = A \cdot Q \ , \quad A \equiv P_1 M_D^{-1/2} P_2 \ . \tag{4.23}$$

因此，利用简正坐标 Q 来表达，多自由度的小振动问题完全简化为独立的单自由度系统的小振动问题。因此，求解多自由度系统的小振动问题实际上最主要的步骤就是寻找其简正模和相应的特征频率。在系统的拉格朗日量给定后，这实际上是一个单纯的线性代数问题。这样的一组 Q_i 就被称为小振动系统的**简正坐标**，又被称为**简正模**。[3]

¶ 下面我们简要说明一下如何求解一个多自由度体系的小振动的初值问题。上面的线性代数证明仅仅是从存在性上。具体操作层面更为简单的方法是直接求解本征方程的非零本征矢。

从体系的运动方程 (4.17) 出发令 $x = \eta e^{-i\omega t}$ 带入我们得到，

$$(\omega^2 m - k)_{ij} \cdot \eta_j = 0 \ . \tag{4.24}$$

其中 $i, j = 1, \cdots, n$ 等标志体系的不同自由度。正如前面已经提到的，要获得非零的解 η，我们需要矩阵 $(\omega^2 m - k$ 奇异，即，

$$\det(\omega^2 m - k) = 0 \ . \tag{4.25}$$

[3] 一般来说，我们称坐标 Q_i 为系统的简正坐标。但是有的书籍中也将坐标 $\tilde{Q}_i$ 称为简正坐标。需要注意的是，简正坐标 Q_i 一般并不具有长度的量纲。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

这是关于 ω^2 的一个 n 次方程。按照前面的分析，一定具有 n 个非负的实数解 ω_i^2，$i=1,2,\cdots,n$。

对应于每一个本征方程的解 ω_i^2 而言，我们都可以找到一个非零的本征矢 $\eta^{(i)}\neq 0$，它满足，

$$(\omega_i^2 m-k)\cdot\eta^{(i)}=0\,. \tag{4.26}$$

对于另外一个本征值 ω_j^2 而言，我们有类似的方程

$$(\omega_j^2 m-k)\cdot\eta^{(j)}=0\,.$$

我们可以将前一个方程左乘以 $(\eta^{(j)})^T$，后一个方程左乘以 $(\eta^{(i)})^T$ 再相减就得到，

$$(\omega_i^2-\omega_j^2)\left(\eta^{(j)}\right)^T\cdot m\cdot\eta^{(i)}=0\,, \tag{4.27}$$

其中我们运用了矩阵 m 和 k 为对称矩阵的事实。因此，对于不同的本征频率 $\omega_i^2\neq\omega_j^2$，它们相对应的本征矢一定满足如下的正交关系：

$$\left(\eta^{(j)}\right)^T m\eta^{(i)}=0\,. \tag{4.28}$$

对于一个所有本征频率没有简并的系统而言，我们可以将相应的本征矢进行归一化，即进行如下的代换，

$$\eta_j^{(i)}\to\tilde{\eta}_j^{(i)}=\frac{\eta_j^{(i)}}{\sqrt{(\eta^{(i)})^T m\eta^{(i)}}}\,. \tag{4.29}$$

这样一来归一化之后的本征矢满足，

$$\tilde{\eta}^{(i)T}m\tilde{\eta}^{(j)}=\delta_{ij}\,. \tag{4.30}$$

因此 $\{\tilde{\eta}^{(i)}:i=1,\cdots,n\}$ 构成了 n 维空间一组正交归一完备的基。我们待求的模态矩阵 A 则由下式给出：

$$A_{ji}=\tilde{\eta}_j^{(i)}\,. \tag{4.31}$$

这一点很容易看出，我们注意到对于任意的 j 而言：

$$\begin{aligned}0&=(\tilde{\eta}^{(i)})^T(\omega_j^2 m-k)\tilde{\eta}^{(j)}\\&=\omega_j^2\delta_{ij}-(\tilde{\eta}^{(i)})^T k\tilde{\eta}^{(j)}\end{aligned} \tag{4.32}$$

其中我们利用了正交归一关系 (4.30)。再利用上面的矩阵 A 与本征矢 $\tilde{\eta}^{(i)}$ 之间的关系 (4.31)，我们发现上式意味着矩阵 A 恰好可将矩阵 k 对角化并且其对角元为各个 ω_j^2：

$$A^T kA=\mathrm{Diag}(\omega_1^2,\cdots,\omega_n^2)\,. \tag{4.33}$$

这就完成了对模态矩阵求解的全过程。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

那么求得了模态矩阵 A 之后如何求解一个一般的多自由振动体系的初值问题呢。我们从如下的事实出发：

$$x = AQ \ , \tag{4.34}$$

并且我们有如下的矩阵关系

$$A^T m A = I \ , \quad A^T k A = \text{Diag}(\omega_1^2, \cdots, \omega_n^2) \ . \tag{4.35}$$

每一个简正模 Q_i 都是以确定的频率 ω_i 进行简谐振动：

$$Q_i(t) = C_i \cos(\omega_i t) + D_i \sin(\omega_i t) \ . \tag{4.36}$$

因此最终 $x(t)$ 的解为，

$$\boxed{x_j(t) = \sum_{i=1}^{n} A_{ji} \left[C_i \cos(\omega_i t) + D_i \sin(\omega_i t) \right] \ .} \tag{4.37}$$

其中的 C_i 和 D_i 是由初条件确定的 $2n$ 个参数，它们由初始条件来确定。

要求出这些常数我们首先假设初始的广义坐标 $x_j(0)$ 以及初始的广义速度 $\dot{x}_j(0)$ 是已知的常数。在公式 (4.37) 中令 $t = 0$ 以及对时间求导后再令 $t = 0$ 我们就得到，

$$x_j(0) = A_{ji} C_i \ , \quad \dot{x}_j(0) = A_{ji} f_i \ . \tag{4.38}$$

这两个公式都可以写成矩阵形式。如果我们令 $C = (C_1, \cdots, C_n)^T$，$f = (\omega_1 D_1, \cdots, \omega_n D_n)^T$，那么我们有，

$$x(0) = A \cdot C \ , \quad \dot{x}(0) = A \cdot f \ . \tag{4.39}$$

从而它们的解为

$$C = A^{-1} \cdot x(0) = (A^T m) \cdot x(0) \ , \quad f = A^{-1} \cdot \dot{x}(0) = (A^T m) \cdot \dot{x}(0) \ , \tag{4.40}$$

更明确地写出来就是，

$$\boxed{C_i = A^T_{ik} m_{kk'} x_{k'}(0) \ , \quad \omega_i D_i = A^T_{ik} m_{kk'} \dot{x}_{k'}(0)} \tag{4.41}$$

这就给出了多自由度体系小振动的一般通解。

例题 4.1　平面双摆 考虑第一章第 2节中的重力场中的平面双摆的小振动 (参见图 1.1)。为了简化起见，假定 $l_1 = l_2 = l$，$m_1 = m_2 = m$。给出体系小振动的本征频率以及模态矩阵 A。同时对于特定的初始条件，$\phi_1(0) = -\phi_2(0) = \phi_0 \ll 1$，$\dot{\phi}_1(0) = \dot{\phi}_2(0) = 0$，给出任意时刻体系的解 $\phi_1(t)$ 和 $\phi_2(t)$ 的明显表达式。

解： 以 ϕ_1，ϕ_2 为广义坐标，我们首先写下这个系统的拉格朗日量:

$$\begin{aligned} L' &= \frac{1}{2}(m_1 + m_2) l^2 \dot{\phi}_1^2 + \frac{1}{2} m_2 l^2 \dot{\phi}_2^2 + m_2 l^2 \dot{\phi}_1 \dot{\phi}_2 \cos(\phi_1 - \phi_2) \\ &+ (m_1 + m_2) g l \cos\phi_1 + m_2 g l \cos\phi_2 \ . \end{aligned} \tag{4.42}$$

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

我们现在假定 $\phi_1 \ll 1$，$\phi_2 \ll 1$，于是我们得到小振动近似下的拉格朗日量:

$$L = ml^2\dot{\phi}_1^2 + \frac{1}{2}ml^2\dot{\phi}_2^2 + ml^2\dot{\phi}_1\dot{\phi}_2 - mgl\phi_1^2 - \frac{1}{2}mgl\phi_2^2 \,. \tag{4.43}$$

这意味着矩阵 m 和 k 由下式给出:

$$m = \begin{pmatrix} 2ml^2 & ml^2 \\ ml^2 & ml^2 \end{pmatrix} \,, \quad k = \begin{pmatrix} 2mgl & 0 \\ 0 & mgl \end{pmatrix} \,, \tag{4.44}$$

于是我们得到矩阵 $(\omega^2 m - k)$ 的表达式如下，

$$\omega^2 m - k = ml^2 \begin{pmatrix} 2(\omega^2 - \omega_0^2) & \omega^2 \\ \omega^2 & \omega^2 - \omega_0^2 \end{pmatrix} \,, \tag{4.45}$$

其中 $\omega_0 \equiv \sqrt{g/l}$。令上面矩阵的行列式为零我们得到本征频率满足的方程及其解为，

$$2(\omega^2 - \omega_0^2)^2 - \omega^4 = 0 \quad \Rightarrow \quad \omega_{1,2}^2 = (2 \pm \sqrt{2})\omega_0^2 \,. \tag{4.46}$$

这给出了系统的两个本征频率 $\omega_{1,2}$，其中 1 对应于 “+” 号而 2 对应于 “-” 号。

下面我们求相应的本征矢并进而求出模态矩阵 A。通过求解本征方程 $(\omega_i^2 m - k)\eta^{(i)} = 0$ 我们发现，

$$\eta^{(1)} = C_1' \begin{pmatrix} 1 \\ -\sqrt{2} \end{pmatrix} \,, \quad \eta^{(2)} = C_2' \begin{pmatrix} 1 \\ +\sqrt{2} \end{pmatrix} \,, \tag{4.47}$$

其中 C_1' 和 C_2' 是两个归一化系数。很容易验证这两个本征矢满足正交关系: $(\eta^{(2)})^T m \eta^{(1)} = 0$。归一化条件可以确定出 C_1' 和 C_2' 为，

$$C_{1,2}' = \tfrac{1}{2}\sqrt{(2 \pm \sqrt{2})/(ml^2)} \,,$$

因此最后获得的模态矩阵为，

$$A = \begin{pmatrix} \eta_1^{(1)} & \eta_1^{(2)} \\ \eta_2^{(1)} & \eta_2^{(2)} \end{pmatrix} = \frac{1}{2\sqrt{ml^2}} \begin{pmatrix} \sqrt{2+\sqrt{2}} & \sqrt{2-\sqrt{2}} \\ -\sqrt{4+2\sqrt{2}} & +\sqrt{4-2\sqrt{2}} \end{pmatrix} \,. \tag{4.48}$$

因此我们得到体系的两个简正模式 $Q_{1,2}$ 分别按照 $\omega_{1,2}$ 进行简谐振动，而体系的广义坐标 $x = (\phi_1, \phi_2)^T$ 与简正坐标 $Q = (Q_1, Q_2)^T$ 之间的关系为，

$$Q = A^{-1} \cdot x \,, \tag{4.49}$$

其中 $x = (\phi_1, \phi_2)^T$ 为两个振动角构成的矢量。至此，我们已经找到了系统的所有简正模和本征频率。

如果我们要求解体系的某个特定的初值问题，比如说对于题目中所要求的初条件:

$$x(0) = \phi_0(1, -1)^T \,, \quad \dot{x}(0) = (0, 0)^T \,. \tag{4.50}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

我们只需要确定系数 C_i 和 D_i 即可。由于初始的速度为零，因此按照公式 (4.41) 我们得知 $D_{1,2}=0$，而系数 C_i 也可以求出，

$$C_i = A^T_{ik} m_{kk'} x_{k'}(0)\ ,\quad i=1,2\ . \tag{4.51}$$

将前面求得的矩阵 A 带入我们得到，

$$\begin{aligned}\begin{pmatrix} C_1 \\ C_2 \end{pmatrix} &= \frac{(ml^2)\phi_0}{2\sqrt{ml^2}} \begin{pmatrix} \sqrt{2+\sqrt{2}} & -\sqrt{4+2\sqrt{2}} \\ \sqrt{2-\sqrt{2}} & +\sqrt{4-2\sqrt{2}} \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ &= \frac{\phi_0\sqrt{ml^2}}{2} \begin{pmatrix} \sqrt{2+\sqrt{2}} & -\sqrt{4+2\sqrt{2}} \\ \sqrt{2-\sqrt{2}} & +\sqrt{4-2\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \phi_0 \frac{\sqrt{ml^2}}{2} \begin{pmatrix} \sqrt{2+\sqrt{2}} \\ \sqrt{2-\sqrt{2}} \end{pmatrix}\end{aligned} \tag{4.52}$$

按照一般通解的公式 (4.37)，我们可以写出这个初值问题的通解，

$$\begin{aligned} x(t) &= A\cdot \begin{pmatrix} C_1\cos(\omega_1 t) \\ C_2\cos(\omega_2 t) \end{pmatrix} \\ &= \frac{\phi_0}{4} \begin{pmatrix} \sqrt{2+\sqrt{2}} & \sqrt{2-\sqrt{2}} \\ -\sqrt{4+2\sqrt{2}} & +\sqrt{4-2\sqrt{2}} \end{pmatrix} \begin{pmatrix} \sqrt{2+\sqrt{2}}\cos(\omega_1 t) \\ \sqrt{2-\sqrt{2}}\cos(\omega_2 t) \end{pmatrix} \\ &= \frac{\phi_0}{4} \begin{pmatrix} (2+\sqrt{2})\cos(\omega_1 t) + (2-\sqrt{2})\cos(\omega_2 t) \\ -(2+2\sqrt{2})\cos(\omega_1 t) + (2\sqrt{2}-2)\cos(\omega_2 t) \end{pmatrix}, \end{aligned} \tag{4.53}$$

其中的 $\omega_{1,2} = \sqrt{2\pm\sqrt{2}}\omega_0$。这就完全求解了这个小振动问题。 搞定

¶ 前面我们提到，如果久期方程 (4.18) 定出来的本征频率正好出现相同的情形，这时我们称系统的这两个本征频率出现**简并** (degenerate)。当系统出现简并时需要更多的细致的考虑。一般来说，我们可以将出现简并的情况分为两类：一类是由于所研究的力学系统具有某种对称性而造成的简并；另一类则不是由于对称性造成的。第二类简并因而又被称为**偶然简并**。对于第一类由于对称性所造成的简并往往是我们所特别关心的。为了说明简并与对称性的关系，我们以下面一个例子来说明。[4]

例题 4.2　苯环的振动 考虑由六个等质量的原子 (C 原子，质量设为 m) 构成一个环。相邻的原子之间等效地用一个经典的、倔强系数为 k 的弹簧彼此联结 (参见图 4.1)。每个原子只允许沿着圆环的切向无摩擦地运动。讨论这个经典系统的小振动问题。

解： 原则上讲，我们只需要写出系统的拉格朗日量并且求解其运动方程就可以了。但这将是六个联立的微分方程组。要求出相应的简正模虽然是直接了当的，但还是有些烦琐的 (需要进行 6×6 矩阵的计算)。一个比较快捷的方法是利用对称性直接来**构造**系统的简正模。这一点我们显示在图 4.1中。

[4] 简并与对称性的这种深刻联系我们这里只能意会，不可言传。要明确揭示这种关联需要用到群表示的理论。事实上，对应于一个简并本征频率的所有线性无关的简正坐标构成了系统对称群的一个表示。

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

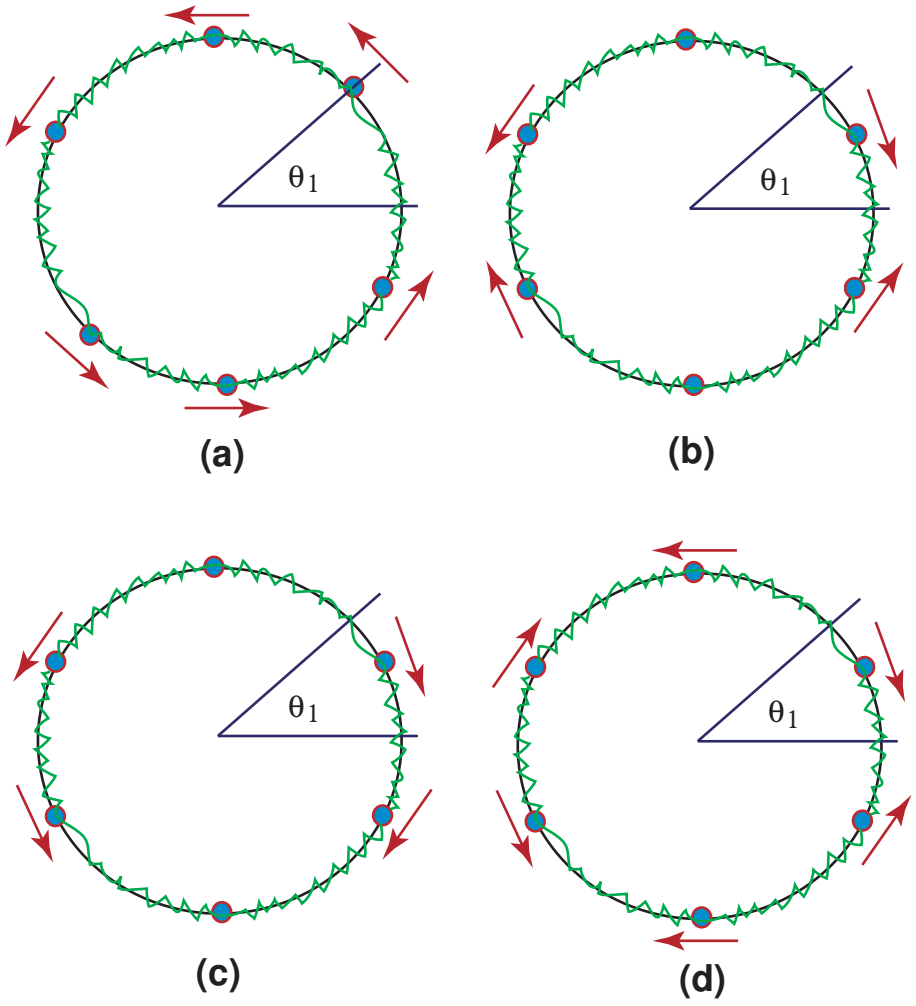

图 4.1: 苯环小振动的经典力学处理。

显然这个系统具有一种模式，那就是所有的原子都向同一方向位移相同的数值，这种模式是体系的一个简正模，它的本征频率显然为：$\omega_1 = 0$。这种模式显示在图 4.1的 (a) 中。另外一种本征模式是将圆环上某两个相对的原子固定不动，另外两对原子相对等位移地振动，如图 4.1的 (b) 所示。这个模式的本征频率很容易确定出来为：$\omega_2 = \sqrt{3k/m}$。第三种简正模式和第二种十分类似，只不过两对原子不是相对振动，而是向一个方向振动，如图 4.1的 (c) 所示。这个模式的本征频率也可以确定出来为：$\omega_3 = \sqrt{k/m}$。第四种模式可以选为相邻的两个原子相对振动，如图 4.1的 (d) 所示。这个模式的本征频率也可以确定出来为：$\omega_4 = \sqrt{4k/m}$。

除了上述四个简正模式以外，另外还存在两个线性独立的振动模式。它们实际上可以选为上面提到的第二种和第三种模式旋转 $\pi/3$ 以后的模式。由于在经典力学中粒子是可以分辨的，因此它们给出不同的简正模 (图中没有显示出)。但是由于对称性，显然它们具

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

有与第二种和第三种模式相同的本征频率。所以我们得到：$\omega_5 = \omega_2$，$\omega_6 = \omega_3$。至此，我们已经找到了这个系统的所有本征频率和振动简正模。 搞定

例题 4.3 **一维固体的振动** 考虑由 N 个等质量的原子（质量设为 m）构成一个环。相邻的原子之间等效地用一个经典的、倔强系数为 k 的弹簧彼此联结。每个原子只允许沿着圆环的切向无摩擦地运动。讨论这个经典系统的小振动问题。

解： 这个问题是上个问题的推广版本。我们用 x_n 来标记第 n 个粒子偏离其平衡位置的坐标，其中 $n = 1, 2 \cdots, N$。那么体系的拉格朗日量可以写成：

$$L = \sum_{n=1}^{N} \frac{m}{2} \dot{x}_n^2 - \frac{k}{2} \sum_{n=1}^{N} \left[(x_n - x_{n-1})^2 \right] . \tag{4.54}$$

注意，周期边条件意味着我们必须保持：$x_{N+n} \equiv x_n$。体系的运动方程为：

$$\ddot{x}_n = \omega_0^2 (x_{n-1} + x_{n+1} - 2x_n) . \tag{4.55}$$

其中 $\omega_0^2 = k/m$。为了求解这个方程，我们寻求如下形式的解：

$$x_n(t) = A e^{-i\omega t + ipn} , \tag{4.56}$$

其中 p 为某个待定参数。[5] 注意，周期边条件意味着参数 p 必须满足：

$$e^{ipN} \equiv 1 \mapsto p = \frac{2\pi}{N} l , \tag{4.57}$$

其中 $l = 1, 2, \cdots, N$。事实上，l 可以取任意连续的 N 个整数值，并不一定要从 1 开始。也就是说真正有意义的是 $\mod(l, N)$，因为具有相同 $\mod(l, N)$ 值的不同 l 实际上给出完全等价的 x_n。因此，一个对称的选择是：$l \in (-N/2, N/2]$。将这个形式代入我们发现：

$$\omega^2(p) = 2\omega_0^2 (1 - \cos(p)) . \tag{4.58}$$

其中 $p = (2\pi/N)l$，$l \in (-N/2, N/2]$。

以 $N = 6$ 为例，我们发现对应于 $l = \pm 2, \pm 1, 0, 3$ 的本征频率的平方分别为：$3\omega_0^2, \omega_0^2, 0, 4\omega_0^2$，其中前两个为二重简并的解，这与上例得到的结果完全一致。

这个例子可以看成是一个一维固体振动的经典模型。在固体物理中，同学们会处理更为复杂的振动模式 (更为复杂的晶格结构，更高的维数等等)。 搞定

21 非谐效应

¶非线性振子或者称为非谐振子是一个十分有趣的研究对象。从小振动的角度来看，如果我们将系统的拉格郎日量在系统的平衡位置附近展开，只要系统偏离平衡位置足够

[5] 我们这里为了方便利用了复数表示。其含义是真正物理的解是相应复数表示的实部，即：$x_n = Re(A e^{-i\omega t + ipn})$。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

小，系统的运动就可以运用简谐振动来加以描述。由于这时的运动方程是线性的微分方程，所以简谐振动有被称为线性振动。如果系统偏离平衡位置的位移 x 不是很小，我们一般需要在拉格郎日量中考虑 x 的高阶 (高于二阶) 的修正。这时相应系统的运动方程就变成非线性微分方程。因此这种振动又被称为非线性振动。当非线性效应出现时，系统会出现一些新的、线性振动中所没有的特点。这一节中我们简单介绍一下这些效应。

考虑一个非谐振子，它的拉格朗日量可以写为：

$$\begin{aligned} L &= \frac{1}{2}\left(m_{ij}\dot{x}_i\dot{x}_j - k_{ij}x_ix_j\right) \\ &+ \frac{1}{2}n_{ijk}\dot{x}_i\dot{x}_jx_k - \frac{1}{3}l_{ijk}x_ix_jx_k \ , \end{aligned} \tag{4.59}$$

其中我们引入了两个非谐项：第一项的形式是 $\dot{x}_i\dot{x}_jx_k$，它是从动能项中的系数对 x 展开所产生的；第二项是势能项展开到位移 x_i 的三次幂所得到的。当然，展开式中应当还存在 x_i 的更高幂次，不过我们假定它们足够小可以忽略。同时，我们也将假定非谐项与谐振项比较是小量，因此我们可以将它们看成是微扰来处理。

现在我们利用谐振项所确定的系统的简正坐标来表达拉格朗日量，它的形式为：

$$L = \frac{1}{2}\left(\dot{Q}_\alpha\dot{Q}_\alpha - \omega_\alpha^2Q_\alpha Q_\alpha\right) + \frac{1}{2}\lambda_{\alpha\beta\gamma}\dot{Q}_\alpha\dot{Q}_\beta Q_\gamma - \frac{1}{3}\mu_{\alpha\beta\gamma}Q_\alpha Q_\beta Q_\gamma \ , \tag{4.60}$$

其中系数 $\lambda_{\alpha\beta\gamma}$ 关于前两个指标是对称的；系数 $\mu_{\alpha\beta\gamma}$ 关于所有指标都是对称的。

非线性振子系统的运动方程可以写成：

$$\ddot{Q}_\alpha + \omega_\alpha^2Q_\alpha = f_\alpha(Q,\dot{Q},\ddot{Q}) \ . \tag{4.61}$$

其中函数 f_α 是一个二次齐次函数。它包含各种可能的 Q，$\dot{Q}$，$\ddot{Q}$ 的二次项。我们将假定所有的这些非线性项是小的。因而我们可以利用逐阶展开的方法来求解这个非线性方程：

$$Q_\alpha(t) = Q_\alpha^{(0)}(t) + Q_\alpha^{(1)}(t) + Q_\alpha^{(2)}(t) + \cdots \ , \tag{4.62}$$

其中 $Q_\alpha^{(0)}(t)$ 就是线性振动的解。因此到第一级我们有：

$$\ddot{Q}_\alpha^{(1)} + \omega_\alpha^2Q_\alpha^{(1)} = f_\alpha(Q^{(0)},\dot{Q}^{(0)},\ddot{Q}^{(0)}) \ . \tag{4.63}$$

注意到这个等式的右边的函数 f 是其宗量的二次函数。因此简单的“三角关系”告诉我们，一个形如：$Q_\alpha^{(0)}Q_\beta^{(0)}$ 的非线性项一定会产生具有频率：$\omega_\alpha \pm \omega_\beta$ 的周期运动。因此，到第一阶 $Q_\alpha^{(1)}(t)$ 中一定会包含具有频率 $\omega_\alpha \pm \omega_\beta$ 的项，这些被称为组合频率。

如果我们考虑到第二阶，那么更多的组合频率会进入。另外，还会出现一个与原先频率相同的项出现在方程 (4.63) 的右边，因为 $\omega_\alpha \equiv \omega_\alpha + \omega_\beta - \omega_\beta$。这样一来，这一项会对方程 (4.63) 的左边形成共振 (表面上会造成无穷大的解)，这需要特别的处理。真实的情况是，到第二阶系统的原先的频率也会有所改变：

$$\omega_\alpha = \omega_\alpha^{(0)} + \Delta\omega_\alpha \ . \tag{4.64}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

只要频率的移动不为零，就不会出现无穷大的解。因此，要求到第二阶能够解出自恰的解就能够定出频率的移动 $\Delta\omega_\alpha$。事实上，频率的移动也可以逐渐展开。将频率的展开式和简正坐标的展开式同时代入系统的运动方程，比较同阶的项就可以逐阶确定非线性振动的微扰解。

¶最后我们指出，上面讨论的情形是假定非线性的影响足够小的情况。这时我们可以将非线性项看成是微扰，逐渐求解运动方程。如果非线性项的贡献本身就不是微扰，那么一般来说求解非线性振动的方程是十分复杂的。很多时候只能够依赖于数值的方法。另外需要注意的一点是，非线性方程的解不具有线性的叠加性。也就是说，非线性微分方程的两个解的线性组合一般不再是方程的解。这也正是非线性振动比较复杂的一个重要原因。有关这方面的详细讨论和更具体的例子可以参考 [4] 的第七章。

22 参数共振

¶考虑一个固有频率依赖于时间的一维振子，它的运动方程可以写为：

$$\ddot{x} + \omega^2(t)x = 0 \,. \tag{4.65}$$

如果函数 $\omega^2(t)$ 满足周期性条件：

$$\omega^2(t+T) = \omega^2(t) \,, \tag{4.66}$$

那么这样的振子被称为参数振子(parametric oscillator)。这个方程又被称为希尔方程(Hill equation)。因为首先是美国天文学家兼数学家 George William Hill 在 1886 年首先研究了它的应用 (主要用于月球精细运动的分析)。其实这个方程的数学理论早年就已经存在了，这就是所谓的 Floquet 理论，由法国数学家 Gaston Floquet 在 1883 年建立。不同的是 Floquet 理论研究的是更为一般的问题，即具有周期性特点的矩阵微分方程：

$$\dot{x} = A(t) \cdot x \,, \tag{4.67}$$

其中 $x \in \mathbb{R}^n$ 是具有 n 个分量的矢量，而 $A(t) \in \mathbb{R}^{n\times n}$ 则是具有周期性的一个 $n \times n$ 实矩阵：$A(t+T) = A(t)$。很显然，希尔方程属于 Floquet 理论的一种特殊情况。尽管一般的 Floquet 理论的讨论也是可以理解的，不过为了明确起见，我们还是将仅仅满足于适用于希尔方程的特殊情况。

方程 (4.65) 是一个二级线性微分方程。因此它有两个线性独立的解：$x_1(t)$ 和 $x_2(t)$。注意，尽管频率函数 $\omega^2(t)$ 是时间的周期函数，但是这个条件一般并不保证方程的解 $x_1(t)$，$x_2(t)$ 也是时间的周期函数。但是可以保证时间平移的解 $x_1(t+T)$ 和 $x_2(t+T)$ 一定也是方程的解。因此一般来说，它们可以写为原先的两个解的某个线性组合：

$$x_i(t+T) = R_{ij}x_j(t) \,, \tag{4.68}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

这里我们引用了爱因斯坦求和规则。上面出现的矩阵 R 可以视为某个抽象的算符 –一般称为 Floquet 算符 –在解空间的矩阵表示。它的具体形式依赖于这两个解的初条件的选择。通常的选择是所谓的 标准基 (standard basis)，它们满足：

$$\begin{cases} x_1(0) = 1\,, \quad \dot{x}_1(0) = 0\,, \\ x_2(0) = 0\,, \quad \dot{x}_2(0) = 1\,, \end{cases} \tag{4.69}$$

这样一来它们分别类似于 $\cos(t)$ 和 $\sin(t)$ 的表现。于是我们很容易验证，对于标准解而言，Floquet 算符 R 的矩阵表达式为，

$$R = \begin{pmatrix} x_1(T) & \dot{x}_1(T) \\ x_2(T) & \dot{x}_2(T) \end{pmatrix}\,. \tag{4.70}$$

我们还可以将 $x_1(t)$ 和 $x_2(t)$ 联合写成两分量矢量的形式。也就是令 $\mathbf{x}(t) = (x_1(t), x_2(t))^T$。这样一来 $\mathbf{x}(x+T)$ 与 $\mathbf{x}(t)$ 之间的关系 (4.68) 就可以简洁地写为：

$$\mathbf{x}(x+T) = R \cdot \mathbf{x}\,, \tag{4.71}$$

其中 R 就是公式 (4.70) 中给出的矩阵。因此就这两个解随时间的变化规律而言，我们仅仅需要知道它们在一个周期之内的行为，随后的依赖都可以通过不停地将矩阵 R 作用于 $x(t)$ 上而得到：$\mathbf{x}(t+nT) = (R)^n \cdot \mathbf{x}(t)$，其中 n 是任意的正整数。

显然，$\mathbf{x}(t)$ 的长时间行为依赖于矩阵 R 的本征值 μ。这个本征值满足的本征方程为，

$$\mu^2 - (x_1(T) + \dot{x}_2(T))\mu + \det(R) = 0\,, \tag{4.72}$$

其中 $\det(R)$ 是矩阵 (4.70) 的行列式。注意到这个行列式实际上是相应二级常微分方程的 Wrongski 行列式 –一个不随时间变化的量，因此 $\det(R(T)) = \det(R(0)) = 1$。因此上述本征方程的常数项实际上等于 1。于是我们可以解出两个本征值：

$$\mu_{1,2} = \frac{x_1(T) + \dot{x}_2(T) \pm \sqrt{[x_1(T) + \dot{x}_2(T)]^2 - 4}}{2}\,. \tag{4.73}$$

由于 $x_{1,2}(T)$ 以及 $\dot{x}_{1,2}(T)$ 都是实数并且两个本征值的乘积永远等于 1，因此基本上它会分化为两类解：第一类是互为复共轭且模为 1 的复数解：

$$\mu_1 = \mu_2^*\,, \quad |\mu_1| = |\mu_2| = 1\,, \tag{4.74}$$

而第二类则是实数解解满足：

$$\mu_1 = 1/\mu_2\,, \quad \mu_1 = \mu_1^*\,, \quad \mu_2 = \mu_2^*\,. \tag{4.75}$$

显然，第一类解出现在 $|x_1(T) + \dot{x}_2(T)| < 2$ 的情况下而第二类解则出现在 $|x_1(T) + \dot{x}_2(T)| > 2$ 的情况下；还有一种边缘情况就是当 R 的两个本征值简并的情况，这出现在 $|x_1(T) + \dot{x}_2(T)| = 2$ 的时候。下面我们分别就这三类情况做一个说明。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

- **第一类情况：** $|x_1(T)+\dot{x}_2(T)|<2$，两个互为复共轭的根
  这时两个本征值是互为复共轭的相因子。我们可以令

$$\mu_{1,2}\equiv\mu_{\pm}=\exp(\pm i\lambda T)\ , \tag{4.76}$$

  其中 $\lambda\in\mathbb{R}$ 可以取为一正实数，称为 Floquet 特征频率。我们将其相对应的本征矢记为 $e_{\pm}(t)$。于是我们有

$$e_{\pm}(t+T)=e^{\pm i\lambda T}e_{\pm}(t)\ . \tag{4.77}$$

  现在我们将上面的解尝试写为如下的形式：

$$e_{\pm}(t)=e^{\pm i\lambda t}u_{\pm}(t)\ . \tag{4.78}$$

  我们发现，

$$e_{\pm}(t+T)=e^{\pm i\lambda t}e^{\pm i\lambda T}u_{\pm}(t+T)\ , \tag{4.79}$$

  与上面的两个式子进行比较我们发现：$u_{\pm}(t+T)=u_{\pm}(t)$。这意味着，在目前这个情况下，相应的本征解可以表达为 (4.78) 的形式 –即一个周期仍然为 T 的周期函数 $u_{\pm}(t)$ 再乘以一个随时间周期振荡的相因子。注意，即使在这种情况下这两个本征解一般也不一定是周期函数。具体来说，$u_{\pm}(t)$ 一定是具有与原先的 $\omega^2(t)$ 相同周期的周期函数；但是它前面随时间振荡的相因子的周期为 $2\pi/\lambda$，这个周期并不一定与 T 的比例是有理数，因此原则上整个函数 $e_{\pm}(t)$ 是非周期函数，尽管整个函数一定是有界的。如果两个周期之比 (或者等价地说两个频率之比) 是有理数，我们称它们是相互公度 的 (commensurate)，只有在这个时候，两个本征解的确是周期函数 (相应的周期是两个公度周期的最小公倍数，而不一定是原先的周期)；反之如果两个周期的比值为无理数，我们则称它们相互是非公度 的 (incommensurate)。[6]

- **第二类情况：** $|x_1(T)+\dot{x}_2(T)|>2$，两个互为倒数的实根
  这时两个 (不相等的) 本征值都是实数且互为倒数。由于两者的乘积为 1，因此它们必定是同号的。无论如何，总是有一个的模会比 1 大。这时微分方程 (4.65) 的两个本征解可以写成：

$$e_1(t+T)=\mu^{t/T}u_1(t)\ ,\quad e_2(t+T)=\mu^{-t/T}u_2(t)\ . \tag{4.80}$$

  其中 $|\mu|>1$，$u_{1,2}(t)$ 仍然是时间的周期函数 (周期仍为 T)。这时第二个解则是指数衰减的，第一个解则是随着时间指数增加的。这意味着平衡位置是不稳定的。这种现象被称为参数共振 。它的典型特点是，如果一个振子的参数 (例如其特征频率 ω) 是时间的周期函数，那么振子的振幅很可能会指数地增加。

[6] 这里讨论的方程都对时间的二阶常微分方程。如果我们将时间换为空间，希尔方程实际上与一维周期势中的 Schrödinger 方程一致。这个问题在固体物理中也需要研究。那里公式 (4.78) 所体现的结论一般被称为布洛赫定理 (Bloch's Theorem)。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

- **第三类情况：** $|x_1(T)+\dot{x}_2(T)|=2$，两个简并的实根 (必定是 ±1)
  第三类是本征值方程出现简并的情形并且我们必定有

$$\mu_1=\mu_2=\mu=\pm1\,. \tag{4.81}$$

  也就是说，矩阵 R 的本征值要么是 $+1$，要么是 -1。这又可以分为两个子类：第一个子类中 $R=(\pm1)\mathbb{1}_{2\times2}$，第二个子类则是 $R=\left[\begin{smallmatrix}\mu&\rho\\0&\mu\end{smallmatrix}\right]$，其中 $\mu=\pm1$ 而 $\rho\neq0$。这时矩阵 R 是不可对角化的，这个形式其实就已经是它的所谓 Jordan 标准型。第一个子类实际上可以视为前面第一类情况的一个极限，即体系的解基本上是有界的。第二个子类恰恰需要额外的考虑。

  显然对于这种情况下的 R，简单的矩阵代数告诉我们，

$$R=\begin{bmatrix}\mu&\rho\\0&\mu\end{bmatrix}\quad\Rightarrow\quad R^n=\begin{bmatrix}\mu^n&n\mu^{n-1}\rho\\0&\mu^n\end{bmatrix}(n\geq2)\,. \tag{4.82}$$

  这个矩阵的对角元基本上是正负震荡的，但是其非对角元的模是按照 $O(n)$ 代数增长的。这虽然不是指数地增加，但是也仍然会导致相应的解趋于发散。

上面的讨论基本上从定性的角度穷尽了希尔方程的解在长时间下的各种可能行为。下面我们来看一个参数共振的定量的例子。参数共振现象在日常生活中还是经常能够遇到的。一个大家一定都十分熟悉的例子就是我们儿时玩儿过的秋千。[7] 在荡秋千时，你需要在秋千上面不断周期性地蹲下和站起。其结果是振子的频率随之发生周期性的改变，因而荡秋千的力学系统可以近似地用参数振子方程 (4.65) 来描述。由于荡秋千时我们对振子频率的改变与振子原来的频率比较是小量，因此我们可以近似地令：

$$\omega^2(t)=\omega_0^2(1+h\cos\gamma t)\,, \tag{4.83}$$

其中 $0<h\ll1$ 是一个小量。我们荡秋千的时候，应当尽量使参数改变的频率 γ 与原来振子的固有频率的两倍 $2\omega_0$ 相近，这时参数共振的现象才是最为有效的。[8] 因此我们假定：$\gamma=2\omega_0+2\epsilon$，其中 $\epsilon\ll\omega_0$。于是，参数振子方程 (4.65) 的解可以写为：

$$x(t)=a(t)\cos(\omega_0+\epsilon)t+b(t)\sin(\omega_0+\epsilon)t\,, \tag{4.84}$$

其中 $a(t)$，$b(t)$ 与上式中的三角函数比较是时间的缓变函数。将这个尝试解带入方程 (4.65)，同时假定 $\dot{a}\sim\epsilon a$，$\dot{b}\sim\epsilon b$，我们寻求指数形式的解：$a(t)\propto\exp(st)$。经过一些简单的计算我们发现：

$$s^2=\frac{1}{4}\left[\frac{1}{4}h^2\omega_0^2-4\epsilon^2\right]\,. \tag{4.85}$$

因此要求存在指数增长的解的条件是：

$$|\epsilon|<\frac{1}{4}h\omega_0\,. \tag{4.86}$$

[7] 话说你们的儿时还玩儿秋千吗？？

[8] 如果你的确荡过秋千，应当还记得这一诀窍吧。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

所以够存在参数共振的频率范围正比于 h；而且当 $\epsilon = 0$ 时，共振增长效应最为明显，这时随着时间指数增长的指数 $s = h\omega_0/4$。



相关的阅读

本章中我们讨论了力学系统的小振动的问题。这类问题所涉及的范围很广。我们的重点是在多自由度系统的线性振动。对于非线性振动 (或者说非谐振动) 只是作了简单的介绍。对于这方面有额外兴趣的同学，可以参考 [4] 的第七章。

就本章的具体内容来说，第 19节取自朗道书 [1] 的第 §21，§22，§25 节。第 20节的讨论取自朗道书 [1] 的第 §23 节。第 21节的讨论取自朗道书 [1] 的第 §28 节。第 22节的讨论则取自朗道书 [1] 的第 §27 节。类似的关于小振动的讨论也可以参考 Goldstein 书中的第六章或者 José 的书 [4] 的第四章。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

第五章　　刚体的运动

本章提要

- 刚体的运动学 (23)
- 刚体的惯量张量 (24)
- 刚体的动力学 (25)
- 自由不对称陀螺 (27)
- 重力场中的对称陀螺 (28)

这一章中我们讨论理论力学中经典的理论 –刚体的力学运动规律。我们的讨论将是比较简略的。这一部分的有些内容实际上在普通物理力学之中都已经涉及了。我们这里将从拉格朗日力学的体系出发，重点讨论一些以前没有涉及的问题。

刚体相关的内容大体可以分为刚体的运动学 (kinematics) 和刚体的动力学 (dynamics) 两个部分。我们也将分为两个部分来讨论。运动学主要涉及刚体运动的描写 –其中的重点是三维转动的描写，动力学则是侧重于刚体的动力学方程的建立和求解。

23　刚体的运动学

¶ 刚体 可以看成是由无穷多质点构成的 一个经典力学体系。它具有固定的尺寸和总质量，而且刚体内 任意 两个质点之间的距离都不随时间改变。也就是说这个物体可以看成是纯粹刚性的。[1] 真实的物体当然都不是真正的刚体，但是如果其形变对运动的影响不在我们的考虑之内，那么可以近似地将其视为刚体从而大大简化计算。

[1] 我们知道，三维的距离并不是一个洛伦兹不变量，只有四维间隔才是。因此，并不存在与狭义相对论完全兼容的“刚体”的概念。在非相对论牛顿力学中，时空对称性由洛伦兹对称性退化为伽利略对称性。这样

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

一个一般的刚体具有六个力学自由度。例如，我们可以选择刚体的质心坐标 (3 个)、从质心 C 到刚体上任意一个固定点 A 的位移矢量 $\mathbf{R}_0$ 的方向 (2 个) 再加上刚体绕着沿 $\mathbf{R}_0$ 方向的一个轴转动的角度 (1 个)。当然某些刚体的自由度数目低于 6。对于经典的刚体，这仅仅出现在刚体本身实际上是低维的情况。例如，一个一维的刚体杆的自由度数目只有 5 个，因为它不存在绕自身轴的转动自由度。

在刚体的运动学描述中，我们常常会选择一个随着刚体一起运动的坐标架来描写刚体的运动。[2] 我们通常称这样固着在刚体上的坐标架为刚体的体坐标架 (body axis)。这里特别需要强调的一点是，不要把这个想象成是取随着刚体一起运动的参照系。这样一来，由于刚体的速度、角速度随时都在变化，这个参照系实际上是一个复杂的非惯性系。比较简单的想法是仅仅把体坐标架想象成随着刚体运动的一些坐标轴。我们仍然在空间固定的惯性系中来考察刚体的运动方程。这个运动方程当然会涉及到表征刚体的一些矢量，例如角速度、角动量、力矩等等。而采用体坐标架仅仅是在每一个时刻都将这些矢量向体坐标架的坐标轴投影而已。一般来说，这个体坐标架的原点总是选择在刚体的质心，这样可以带来诸多的便利。

本节中，我将首先复习一下三维转动矩阵的一些基本性质 (第 23.1小节)；随后我们讨论无穷小转动和刚体的角速度 (第 23.2小节)；然后我们介绍刚体转动的欧拉角的描述 (第 23.3小节)。最后，对转动矩阵的另外一种参数化描述，即所谓的 Cayley-Klein 参数化 (第 23.4小节)，我们也会做个简单的介绍。

23.1 三维转动矩阵的一些基本性质

¶ 这一小节中我们将着重讨论 (复习) 一下与刚体转动有关的三维转动矩阵的一些基本性质。

线性代数的知识告诉我们，对三维空间中的任意一个矢量 $\mathbf{x} \in \mathbb{R}^3$，我们可以将其表达为：

$$\mathbf{x} = x_i \mathbf{e}_i \ , \tag{5.1}$$

其中 $\{\mathbf{e}_i, i = 1, 2, 3\}$ 为三维矢量空间中的一组正交归一的基矢，x_i 则称为矢量 $\mathbf{x}$ 在这组基矢下的坐标，又称为该矢量的分量。在不至于引起误会的前提下，我们常常又直接用它来代替矢量 $\mathbf{x}$ 本身。需要注意的是，一个给定的矢量 $\mathbf{x}$ 并不依赖于基矢的选取，但是它的坐标 x_i 却依赖于基矢的选取。基矢的选取具有相当的任意性，从这个意义上讲，矢量的坐标 (分量) 并不代表真实的物理，仅仅是一种描述方式而已。如果一个矢量 $\mathbf{x}$ 在某组基矢 $\{\mathbf{e}_i, i = 1, 2, 3\}$ 下的坐标 (分量) 为 x_i，在另外一组基矢 $\{\mathbf{e}'_i, i = 1, 2, 3\}$ 下的坐标

一来，三维的转动 (它保持空间中任意两点间三维距离不变) 与时间完全剥离，这时可以自洽地定义刚体的概念。因此，本章的所有讨论仅在非相对论力学中才有意义。

[2] 这样做的优点在下面几节过后就会变得十分显然了。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

(分量) 记为 x_i'，那么我们有：

$$x_i' = A_{ij}x_j \ , \tag{5.2}$$

其中矩阵 A_{ij} 是一个 3×3 的实正交矩阵。它标志了不同基矢的选取之间的转换关系。事实上我们有：

$$\mathbf{e}_i = \mathbf{e}_j' A_{ji} \ . \tag{5.3}$$

读者不难验证，对于任意一个矢量 $\mathbf{x} = x_i\mathbf{e}_i = x_iA_{ji}\mathbf{e}_j' = x_j'\mathbf{e}_j'$，因此这与坐标的变换规则 (5.2) 一致。正是在这个意义下，我们常常说变换规则 (5.2) 可以视为三矢量的定义。换句话说，凡是各个分量 x_i 在坐标变换下 (即不同基矢的选取下) 按照规则 (5.2) 变换的量就称为 (三维) 矢量。

¶ 在讨论转动的过程中，可以采用两种等价的描述方法。一种观点是认为矢量保持不变，转动坐标架 (或者说选择不同的基矢)，这又称为被动观点；另一种等价的描述方法是转动矢量 (或者更复杂的张量) 本身而坐标架保持不动，这又称为主动观点。我们前面说的三维转动是采用前一种方法，即被动观点。有的时候，特别是在讨论刚体问题的时候，我们会相互转换这两种观点。很显然，一种观点中的转动角与另一种观点的转动角刚好相反。

¶ 在物理上，每个实正交矩阵 (在被动观点中) 都等价于一个坐标架的三维转动，因此又称为三维转动矩阵。如果 A 的转置矩阵记为 A^T，那么它们满足：

$$A^T\cdot A = A\cdot A^T = \mathbb{1} \ , \tag{5.4}$$

其中 $\mathbb{1}$ 为 3×3 单位矩阵。这个条件实际上来源于三维转动不改变任何两个三矢量内积的事实。容易证明，两个正交矩阵的乘积仍然是一个正交矩阵。因此，所有的三维转动矩阵 (或者称为正交矩阵) 在矩阵乘法下构成一个群的结构，称为三维转动群或三维正交群，记为 $O(3)$。将这个式子两边取行列式并利用 $\det(A) = \det(A^T)$，我们就得到：$\det(A) = \pm 1$。因此，三维转动矩阵可以按照其行列式的两个不同取值分为互不联通的两支。满足 $\det(A) = +1$ 的转动我们称为正常转动；满足 $\det(A) = -1$ 的转动我们称为非正常转动，它实际上可以看成是一个正常转动再乘以宇称变换矩阵：

$$P = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \ . \tag{5.5}$$

显然，宇称变换矩阵 P 满足 $\det(P) = -1$。所有的正常转动也构成一个群，称为特殊三维正交群，或特殊三维转动群，记为 $SO(3)$。显然，$O(3)$ 群可以分为互不连通的两支，其中一支由所有的正常转动构成，即 $SO(3)$ 群；另一支则由所有的非正常转

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

动构成，它们本身不能构成群的结构，因为两个非正常转动的乘积是一个正常转动。关于对称性与群的概念在经典力学里面的应用，有兴趣的同学可以参考本讲义后面的附录 A以及那里给出的进一步的参考文献。

¶关于三维正常转动矩阵我们有下面的重要结论，它又经常被称为欧拉定理：

定理 5.1 **一个正常转动所对应的三维转动矩阵一定存在一个本征值为 $+1$，与其对应的本征矢量实际上对应于该转动的转动轴。**

这个结论有多种方法可以证明。一个最简单而直接的方法是利用恒等式：

$$(A-\mathbb{1})A^T = \mathbb{1} - A^T \,,$$

并在两边取行列式，利用 $\det A^T = \det A = +1$，因此我们有：

$$\det(A-\mathbb{1}) = (-)^3 \det(\mathbb{1}-A) = \det(\mathbb{1}-A) \,.$$

这意味着 $\det(A-\mathbb{1}) = 0$，即 A 一定有一个本征值为 $+1$。

另外一种证明方法是直接列出矩阵 A 的特征方程：

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix} = 0 \,. \tag{5.6}$$

明确地写出来就是

$$\begin{aligned} &(a_{11}-\lambda)(a_{22}-\lambda)(a_{33}-\lambda) + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} \\ &-a_{13}a_{31}(a_{22}-\lambda) - a_{23}a_{32}(a_{11}-\lambda) - a_{12}a_{21}(a_{33}-\lambda) = 0 \end{aligned}$$

这个等式左边是关于 λ 的多项式，其结构是这样的：最高幂次为 $-\lambda^3$，平方项为：$(a_{11}+a_{22}+a_{33})\lambda^2$；常数项显然就是矩阵 A 的行列式 $\det(A) = +1$；唯一需要仔细一点考察的是关于 λ 的线性项。具体写出来是：

$$\lambda\left(a_{13}a_{31} - a_{11}a_{33} + a_{12}a_{21} - a_{11}a_{22} + a_{23}a_{32} - a_{22}a_{33}\right)$$

这个表达式括号里面的项恰好是原先矩阵行列式的相应代数余子式。按照线性代数的知识，这正比于矩阵的逆矩阵的元素。又由于正交矩阵的转置正好是它的逆，因此实际上就是 $\lambda(a_{22}+a_{33}+a_{11}) = \lambda Tr(A)$。因此矩阵 A 的特征方程可以化为

$$(\lambda-1)\left[\lambda^2 - Tr(A)\lambda + 1\right] = 0 \,. \tag{5.7}$$

这个表达式说明，A 一定存在一个 $+1$ 的本征值。另外两个本征值满足方程 $\lambda^2 - Tr(A)\lambda + 1 = 0$。虽然貌似另外两个本征值一般也可以是实数，但是由于它们的乘积被固定为 $+1$，同时注意到转动矩阵不能改变任意一个矢量的长度，所以最为一般的情形实际上只能是：

$$\lambda_0 = +1 \,, \quad \lambda_{1,2} = e^{\pm i\Theta} \,. \tag{5.8}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

即另外两个本征值为互为复共轭的相因子，这时 $Tr(A) = 1 + 2\cos\Theta$。我们看到，对于都是实根的情形一般只能是：$\lambda_{1,2} = +1$(对应于 $\Theta = 0$) 或 $\lambda_{1,2} = -1$(对应于 $\Theta = \pi$)，其中前者对应于完全不转的情形而后者对应于绕某个轴旋转 π。

基于上面的分析，我们可以写出绕 z 轴旋转任意一个角度 Θ 的正常转动所对应的转动矩阵。它具有如下的形式，

$$A = \begin{pmatrix} \cos\Theta & \sin\Theta & 0 \\ -\sin\Theta & \cos\Theta & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{5.9}$$

读者不难验证，它就具有我们前面分析的特征：具有一个本征值为 $+1$，其余两个本征值为 $e^{\pm i\Theta}$。

¶ 按照欧拉定理，所有三维正常转动都可以看成是绕着某个轴的转动。因此，刻画一个正常转动我们需要 $(\Theta, \mathbf{n})$ 这一对参数，其中 $\mathbf{n}$ 是转轴方向的单位矢量，Θ 则是绕该轴的转动角。我们将用 $R(\Theta, \mathbf{n})$ 来标记这样的转动。有了这个记号，我们可以来讨论所有正常转动构成的群 $SO(3)$ 的拓扑结构，它也是一个一般的经典刚体的转动部分的 相流形 (phase manifold) 的拓扑结构。为此，我们考虑三维空间中一个半径为 π 的球体。我们约定球体的中心 O 对应于不转动 (单位矩阵)；球体内任意一点 A 对应于以原点与该点的连线 OA 为转动轴，以该点到原点的距离为转动角 Θ 的一个转动。对于转动角在 $(-\pi, 0)$ 之内的转动 $R(\Theta, \mathbf{n})$，我们利用 $R(\Theta, \mathbf{n}) = R(-\Theta, -\mathbf{n})$ 将它化为一个正的角度的转动，这使得我们可以控制转动角度 Θ 的取值在 $(0, \pi)$ 以内。同时，对于位于球面上面的点，它们都对应于绕某个轴的转角为 $\pm\pi$ 的转动。对应球面上相对的两个点，由于 $R(\pi, \mathbf{n}) = R(-\pi, -\mathbf{n})$，因此这两个点必须认为是完全相同的。因此我们必须将球面上相对的两个点等同起来 (或者说粘结起来)。这个颇为复杂的构造就是 $SO(3)$ 群流形的拓扑结构，在数学上它被称为 三维影射空间 (three-dimensional projective space)，记为 RP^3。

23.2 无穷小转动与角速度

¶ 转动矩阵之间一般是不可对易（或者说不可交换）的。因此虽然任何两个转动所对应的转动矩阵 (比如说 A 和 B) 之乘积仍然对应于一个转动矩阵 AB，它代表首先进行转动 B 然后进行转动 A，但是一般来说，$AB \neq BA$。这在物理上对应于一般的转动也是不可交换的。但是如果两个转动属于无穷小的转动，那么它们实际上是可对易的。正是这个事实表明了角速度矢量的定义是有意义的。正如前面提到的，所谓无穷小转动是无穷接近不转动的一种转动，它是一种 正常转动，即与其对应的转动矩阵的行列式为 $+1$。这类无穷小转动又可以分为三种独立的转动的叠加：在 xy 平面、yz 平面、xz 平面。对一个无穷小的转动，我们可以将其转动矩阵表达为：

$$A \simeq \mathbb{1} + d\boldsymbol{\theta} \cdot (i\mathbf{S}) + \cdots , \tag{5.10}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

其中 $d\boldsymbol{\theta} = (d\theta_1, d\theta_2, d\theta_3)$ 为一实的无穷小转动角矢量，而 $i\mathbf{S} = (i\mathbf{S}_1, i\mathbf{S}_2, i\mathbf{S}_3)$ 为下列三个反对称 3×3 矩阵：[3]

$$i\mathbf{S}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & +1 & 0 \end{pmatrix}, \; i\mathbf{S}_2 = \begin{pmatrix} 0 & 0 & +1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \; i\mathbf{S}_3 = \begin{pmatrix} 0 & -1 & 0 \\ +1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{5.11}$$

这三个矩阵一般被称为三维转动的 生成元 (generator)，它们分别“生成了”2-3，1-3，1-2 平面中的无穷小转动。一个任意的转动总可以看成是上述三种独立的无穷小转动“积分”后的结果。前面提到过三维的正常转动构成一个群 $SO(3)$，这个群实际上包含无穷多的元素，是一个 李群 。相应的生成元就是上面给出的这三个矩阵。这些矩阵本身构成了相应李群的 李代数 。容易发现三个生成元之间并不可对易，这体现了一般的转动是不可交换的。上述三个生成元之间的基本对易关系为：

$$[\mathbf{S}_i, \mathbf{S}_j] = i\epsilon_{ijk}\mathbf{S}_k \, . \tag{5.12}$$

这恰好是 (量子力学中) 角动量之间的对易关系，说明角动量的确是与空间转动有着密不可分的关系。[4]

在三维空间中一个任意的无穷小 (正常) 转动下，一个矢量的改变可以表达为：

$$d\mathbf{x} \equiv \mathbf{x}' - \mathbf{x} = [d\boldsymbol{\theta} \cdot (i\mathbf{S})] \cdot \mathbf{x} = \mathbf{x} \times d\boldsymbol{\theta} \, . \tag{5.13}$$

其中最后一步我们将它写为两个矢量的叉乘。需要注意的是，上面这个关系式是利用所谓的 被动观点 得到的。因此，如果我们希望用所谓 主动观点 来重新写这个关系式，那么我们需要将其中的 $\boldsymbol{\theta}$ 替换为 $-d\boldsymbol{\theta}$。现在我们考虑刚体具有的某个矢量 $\mathbf{G}$ 在某个无穷小时间间隔 dt 内相对于空间固定坐标系的变化 $(d\mathbf{G})_{\text{space}}$，那么我们有：

$$(d\mathbf{G})_{\text{space}} = (d\mathbf{G})_{\text{body}} + (d\mathbf{G})_{\text{rot}} \, , \tag{5.14}$$

其中 $(d\mathbf{G})_{\text{body}}$ 是矢量 $\mathbf{G}$ 相对于刚体体坐标架的变化而 $(d\mathbf{G})_{\text{rot}}$ 则是由于刚体本身 (以及固着在其上的体坐标架) 的转动所带来的变化。如前所说，我们这时需要用 $-d\boldsymbol{\theta}$ 来替代 $d\boldsymbol{\theta}$，因此我们有：

$$(d\mathbf{G})_{\text{rot}} = \mathbf{G} \times (-d\boldsymbol{\theta}) = d\boldsymbol{\theta} \times \mathbf{G} \, . \tag{5.15}$$

也就是说，

$$(d\mathbf{G})_{\text{space}} = (d\mathbf{G})_{\text{body}} + d\boldsymbol{\theta} \times \mathbf{G} \, . \tag{5.16}$$

由此我们可以定义刚体的 角速度矢量 $\boldsymbol{\Omega}$(这是一个轴矢量)，

$$\boldsymbol{\Omega} = \frac{d\boldsymbol{\theta}}{dt} \, , \tag{5.17}$$

[3] 注意矩阵 $i\mathbf{S}$ 是实的、反对称矩阵，但是它们并不是厄米的，而是反厄米的。$\mathbf{S}$ 则是厄米的，符合量子力学中关于物理量的要求。

[4] 如果大家还没有学过量子力学，就当我什么也没说。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

这样一来上式可以写为对时间的变化率的形式：

$$\left(\frac{d\mathbf{G}}{dt}\right)_{\text{space}} = \left(\frac{d\mathbf{G}}{dt}\right)_{\text{body}} + \boldsymbol{\Omega} \times \mathbf{G} \,. \tag{5.18}$$

这个式子的一个具体的应用就是大家所熟知的 (通过简单的几何分析也可以给出的)，在刚体上任意一点相对于一个空间固定坐标架 (一个惯性系) 的速度 $\mathbf{v}$ 可以表达为：

$$\mathbf{v} = \mathbf{v}_C + \boldsymbol{\Omega} \times \mathbf{r} \,. \tag{5.19}$$

这里 $\mathbf{v}_C$ 表示刚体质心的平动速度；$\boldsymbol{\Omega}$ 是刚体转动的角速度矢量；$\mathbf{r}$ 是从质心 C 到所考虑点的位置矢量。

23.3 欧拉角的描述

¶ 描述一个刚体的三个转动自由度有许多方法。一个经常用到的描述方式就是利用著名的欧拉角。为此我们考虑在惯性系空间固定的一个坐标架 XYZ。我们再考虑一个原

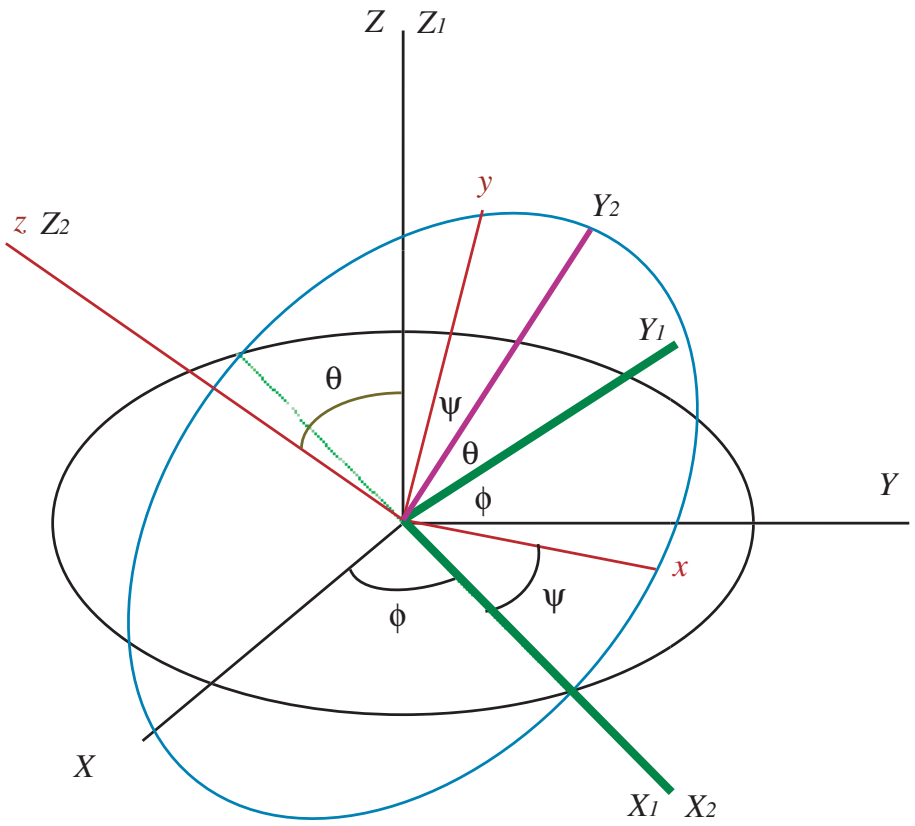

图 5.1: 刚体的欧拉角。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

点位于刚体质心、同时随着刚体运动的坐标架 xyz，参见图 5.1。[5] 那么，刚体的位置实际上可以完全由其质心坐标 (在固定坐标架 XYZ 中的坐标) 和体坐标架的三个坐标轴相对于固定坐标架的三个角度所唯一确定。这三个欧拉角，ϕ，θ，ψ，可以认为是下列三个顺序转动的叠加：[6]

1. 绕原先的 Z 轴，逆时针旋转角度 ϕ 到新的坐标系 $X_1Y_1Z_1$；注意，由于是绕 Z 轴旋转，因此新的 Z_1 轴与老的 Z 轴是完全重合的。这一步可以简记为：$XYZ \xrightarrow{R(\phi,\hat{Z})} X_1Y_1Z_1$。
2. 绕新的 X_1 轴，逆时针旋转角度 θ 到新的坐标系 $X_2Y_2Z_2$; 同样，由于是绕 X_1 轴旋转，因此新的 X_2 轴与老的 X_1 轴也是重合的。这一步可以简记为：$X_1Y_1Z_1 \xrightarrow{R(\theta,\hat{X})} X_2Y_2Z_2$。
3. 绕新的 Z_2 轴，逆时针旋转角度 ψ 到新的坐标系 $X_3Y_3Z_3$，这个坐标系我们也称为 $xyz \equiv X_3Y_3Z_3$。这一步可以简记为：$X_2Y_2Z_2 \xrightarrow{R(\psi,\hat{Z})} xyz$。

这三个角度中，θ 的取值范围是从 0 到 π；ϕ，ψ 的取值范围都是从 0 到 2π。这三个角度一起被称为刚体运动的欧拉角。

为了明确写出与欧拉角对应的转动矩阵的显式，我们将最初的 XYZ 坐标系中某个矢量的坐标记为 $\mathbf{x}^{(0)}$。相应的坐标系 $X_1Y_1Z_1$ 中的坐标记为 $\mathbf{x}^{(1)}$；坐标系 $X_2Y_2Z_2$ 中的坐标记为 $\mathbf{x}^{(2)}$。最后，坐标系 $X_3Y_3Z_3 \equiv xyz$ 中的坐标记为 $\mathbf{x}^{(3)}$。对于第一个旋转我们有：

$$\mathbf{x}^{(1)} = D \cdot \mathbf{x}^{(0)}\ ,\quad D = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}\ . \tag{5.20}$$

对第二个旋转我们有

$$\mathbf{x}^{(2)} = C \cdot \mathbf{x}^{(1)}\ ,\quad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}\ . \tag{5.21}$$

注意第二个转动是绕 x 轴的，因此矩阵的形式也需要做适当调整。第三个转动又是绕 z 轴的，因此我们有

$$\mathbf{x}^{(3)} = B \cdot \mathbf{x}^{(2)}\ ,\quad B = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}\ . \tag{5.22}$$

[5] 由于我们仅仅关心动坐标架 xyz 相对于静止坐标架 XYZ 的角度，因此我们完全可以将它们的原点画在一起。

[6] 我们这里选择的是比较常用的 ZXZ 约定，即首先绕 Z 轴旋转一个角度，然后绕 (新的)X 轴转第二个角度，最后再绕 (新的)Z 轴转第三个角度。其他使用的约定还有 ZXY 约定等。关于更多的约定以及各个约定之间的关系，有兴趣的读者可以参考 Wikipedia 上面的网页：`https://en.wikipedia.org/wiki/Euler_angles`。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

将上述三个转动结合，我们就得到了与欧拉角对应的转动矩阵的表达式：

$$\mathbf{x}^{(3)} = (B \cdot C \cdot D) \cdot \mathbf{x}^{(0)} = A \cdot \mathbf{x}^{(0)} \ , \quad A \equiv BCD \ . \tag{5.23}$$

将上述三个矩阵明确地乘出来我们就得到矩阵 A 的明显表达式，

$$A = \begin{pmatrix} \cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi & \cos\psi\sin\phi + \cos\theta\cos\phi\sin\psi & \sin\psi\sin\theta \\ -\sin\psi\cos\phi - \cos\theta\sin\phi\cos\psi & -\sin\psi\sin\phi + \cos\theta\cos\phi\cos\psi & \cos\psi\sin\theta \\ \sin\theta\sin\phi & -\sin\theta\cos\phi & \cos\theta \end{pmatrix} . \tag{5.24}$$

大家可以验证这个矩阵满足 $A^T = A^{-1}$。

¶ 一个重要的结果就是利用三个欧拉角及其时间微商来表达刚体的角速度在动坐标架 xyz 上面的投影。参考图 5.1我们看到 $\dot{\phi}$，$\dot{\theta}$，$\dot{\psi}$ 分别沿着 $Z(Z_1)$，$X_1(X_2)$，$z(Z_2)$ 轴，于是经过简单的几何考虑我们可以得到的结果是：

$$\begin{cases} \Omega_x = \dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi \ , \\ \Omega_y = \dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi \ , \\ \Omega_z = \dot{\phi}\cos\theta + \dot{\psi} \ . \end{cases} \tag{5.25}$$

这个结果我们后面给出刚体的动能时会用到。

23.4 Cayley-Klein 参数

¶ 另外一种十分巧妙的描写三维转动的数学方法是使用 2×2 的复矩阵。前面的讨论指出，一个任意的三维正常转动矩阵 A 由三个实参数描写 (例如三个欧拉角)。本小节中我们试图建立一种对应关系，这种对应关系将实的三维正常转动矩阵与 2×2 的复矩阵建立某种对应关系，这就是著名的 Cayley-Klein 参数。

首先对应三维空间中任意一个矢量 $\mathbf{x} = (x_1, x_2, x_3)^T$，我们可以引入一个无迹的 2×2 厄米矩阵 P：

$$P = x_i\sigma_i = \mathbf{x}\cdot\boldsymbol{\sigma} = \begin{pmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 \end{pmatrix} , \tag{5.26}$$

其中 $\boldsymbol{\sigma}$ 为下列三个泡利矩阵：

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{5.27}$$

公式 (5.26) 实际上建立了一个任意的三维实矢量 $\mathbf{x}$ 与一个无迹的 2×2 厄米矩阵 P 之间的一一对应关系，或者说等价关系：$P \sim \mathbf{x}$。两者中只要知道了其中一个就可以完全确定另一个。虽然定义式 (5.26) 给出的是给定 $\mathbf{x}$ 后的矩阵 P，但是反过来的关系也很容易得到：

$$x_i = \frac{1}{2}Tr(P\sigma_i) \ . \tag{5.28}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

因此无论我们知道了哪一个，另外一个也可以完全确定。

现在我们注意到 2×2 复矩阵 P 的行列式恰好与矢量 $\mathbf{x}$ 的模方相联系：

$$\det(P) = -\mathbf{x}\cdot\mathbf{x} = -|\mathbf{x}|^2 \,. \tag{5.29}$$

我们知道任意一个矢量的模方 $|\mathbf{x}|^2$ 恰恰是在三维转动下保持不变的量。因此，如果我们希望一个三维转动矩阵 A 能够对应于一个 2×2 的复矩阵 Q，那么相应的变换矩阵 Q 作用于 $\mathbf{x}$ 的对应物 P 的时候，应当保持 $\det(P)$ 不变。要实现这点，我们可以取 2×2 的特殊幺正矩阵 $Q\in SU(2)$。所谓二维特殊幺正矩阵 Q 是指行列式为 $+1$ 的 2×2 幺正矩阵，它们满足：

$$Q^\dagger Q = QQ^\dagger = \mathbb{1}_{2\times 2} \,,\quad \det(Q) = +1 \,. \tag{5.30}$$

如果我们定义变换后的矩阵 P' 为

$$P' = Q\cdot P\cdot Q^\dagger \,. \tag{5.31}$$

按照对应关系 (5.26) 这等价地定义了一个“转动了的”矢量 $\mathbf{x}'$。由于 $\det(P') = \det(Q)\det(P)\det(Q)^* = \det(P)$，因此变换 (5.31) 并没有改变相应矢量 $\mathbf{x}$ 的模方，即 $|\mathbf{x}'|^2 = |\mathbf{x}|^2$。我们可以认为变换 (5.31) 诱导了一个三维的正常转动 A，即：$Q\sim A$。我们这小节的任务就是确定这种对应关系。

$SU(2)$ 矩阵的表达方法有很多，最为简洁的一个是：

$$Q = q_0\mathbb{1} + i\mathbf{q}\cdot\boldsymbol{\sigma} = \begin{pmatrix} q_0+iq_3 & iq_1+q_2 \\ iq_1-q_2 & q_0-iq_3 \end{pmatrix} \,, \tag{5.32}$$

其中实的参数 q_0 以及实矢量 $\mathbf{q}$ 满足约束：$q_0^2+\mathbf{q}^2=1$，这等价于 $\det(Q)=+1$。这个表达式实际上就是哈密顿当年津津乐道的四元数 (quaternion)。满足约束条件 $\det Q = 1$ 的四元数又称为单位四元数 (unit quaternion)。显然，所有的 $SU(2)$ 矩阵在矩阵乘法下构成一个群，这个群也称为 $SU(2)$ 群。

前面的讨论相当于建立了三维正常转动矩阵 $A\in SO(3)$ 与 $SU(2)$ 矩阵之间 (或者说 $SO(3)$ 与 $SU(2)$ 这两个群之间) 的一个对应关系。我们知道三维空间的正常转动矩阵 A 会将空间中的任一矢量 $\mathbf{x}$ 变换为 $\mathbf{x}' = A\mathbf{x}$。与此对应，在 2×2 无迹的厄米矩阵空间中，对于三维空间中的每一个矢量 $\mathbf{x}$ 都可以定义其对应 $P = \mathbf{x}\cdot\boldsymbol{\sigma}$。我们再定义一个 $SU(2)$ 的变换矩阵 $Q\sim A$，使得变换后的矢量所对应的矩阵为 $P' = QPQ^\dagger \sim \mathbf{x}'$。用公式表达就是：

$$\begin{aligned} \mathbf{x}\cdot\boldsymbol{\sigma} = P \quad &\sim \quad \mathbf{x} \,, \\ \mathbf{x}'\cdot\boldsymbol{\sigma} = QPQ^\dagger = P' \quad &\sim \quad \mathbf{x}' = A\cdot\mathbf{x} \\ SU(2)\ni Q \quad &\sim \quad A\in SO(3) \end{aligned} \tag{5.33}$$

那么很容易证明这些对应实际上建立起了两个群 ($SO(3)$ 和 $SU(2)$) 之间的一个同态 (homomorphism)。同态是两个群之间的一个对应关系，它必须保持群的乘法。也就

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

是说，如果给定两个三维转动矩阵 $A_1, A_2 \in SO(3)$，我们知道它们的复合变换 –即两个相应变换矩阵的乘积所对应的变换 –仍然是一个三维转动矩阵 $A = A_2A_1 \in SO(3)$；同态的关系要求两个群元之间的对应关系必须保持这种群的乘法，即与 A_1，A_2 对应的如果分别记为 Q_1，Q_2，我们必须有 $Q = Q_2Q_1$ 一定与 $A = A_2A_1$ 对应：$Q \sim A$。这一点实际上很容易验证。为此考虑任意的一个 $\mathbf{x}$ 经过一个三维转动矩阵 A_1 后变为 $\mathbf{x}' = A_1\mathbf{x}$；另一个转动矩阵 A_2 将 $\mathbf{x}'$ 变为 $\mathbf{x}'' = A_2\mathbf{x}'$。那么这个变换的净效果是转动矩阵 $A = A_2 \cdot A_1$。在二维复矩阵方面，令 $A_1 \sim Q_1$，$A_2 \sim Q_2$；$\mathbf{x} \sim P$，$\mathbf{x}' \sim P'$，$\mathbf{x}'' \sim P''$，那么我们有：$P'' = Q_2P'Q_2^\dagger = (Q_2Q_1)P(Q_2Q_1)^\dagger$。因此 $A = A_2A_1 \sim Q_2Q_1$。所以这种对应关系的确是保持群的乘法的。

需要注意的是，同态并不要求对应关系是一一的。[7] 实际上两个 $SU(2)$ 中的元素对应于同一个三维转动矩阵。这可以从变换关系 $P' = QPQ^\dagger$ 中看出，事实上 Q 和 $(-Q)$ 对应于同一个 P'，因此按照定义它们必定对应于同一个转动矩阵 A。换句话说，更为准确的描述是三维正常转动矩阵 A 与一对 $SU(2)$ 矩阵 $(Q, -Q)$ 之间建立了一一的对应关系。因此 $SU(2)$ 又称为 $SO(3)$ 的 双重覆盖群。

建立三维正常转动群 $SO(3)$ 与二维特殊幺正群 $SU(2)$ 之间的关系有助于我们更好地理解和描述三维正常转动群。在经典力学的层面上，这有助于我们了解 $SO(3)$ 群相流形的很多性质。前面曾经提到过，$SO(3)$ 群的结构之所以重要是因为对于一个经典的三维刚体而言，它的 相流形 的转动部分 (即扣除其质心平动部分之后) 就是 $SO(3)$。这个流形是联通的，但不是单联通的。建立起它与 $SU(2)$ 的对应关系以后有一个方便的地方在于 $SU(2)$ 是一个单联通的流形。$SU(2)$ 群的拓扑结构也是十分简单的。考虑到表示 (5.32) 以及约束条件 $q_0^2 + \mathbf{q}^2 = 1$，$SU(2)$ 群流形等价于一个四维欧氏空间中的三维球面，因此它具有三维球面 S^3 的拓扑。当然，在经典力学层面 $SU(2)$ 群所作用的 矢量空间 貌似仅仅是抽象的一个数学构造，也许与真实的三维转动并没有直接的对应。但在量子力学的层面上，这种对应关系会更加深入一些，因为 $SU(2)$ 实际上还包含了半奇数的角动量 (即所谓的费米子) 的情形。在量子力学中，$SU(2)$ 所作用的空间 –由二维的复矢量 (又称为 旋量) 构成 –是有物理含义的，它们对应于费米子自旋部分的波函数。

¶ 类似于前面关于 3×3 实转动矩阵的表示 (5.24)，我们也可以将欧拉角对应的转动矩阵用 2×2 的复 $SU(2)$ 矩阵表示出来。我们需要的只是对应于三个欧拉角转动的 Q 矩阵：

$$Q_\phi = \begin{pmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{pmatrix}, \quad Q_\theta = \begin{pmatrix} \cos\frac{\theta}{2} & i\sin\frac{\theta}{2} \\ i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}, \quad Q_\psi = \begin{pmatrix} e^{i\psi} & 0 \\ 0 & e^{-i\psi} \end{pmatrix}. \tag{5.34}$$

这三个矩阵分别对应于第 23.3小节中的矩阵 D，C 和 B，参见公式 (5.20)，(5.21) 和

[7] 如果同态是一一对应，这样的同态又称为 同构 (isomorphism)。从代数上讲，同构的两个群被认为是完全一样的。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

(5.22)。与转动矩阵对应的是上面三个 Q 矩阵的乘积，

$$Q = Q_\psi Q_\theta Q_\phi = \begin{pmatrix} e^{i(\psi+\phi)/2}\cos\frac{\theta}{2} & ie^{i(\psi-\phi)/2}\sin\frac{\theta}{2} \\ ie^{-i(\psi-\phi)/2}\sin\frac{\theta}{2} & e^{-i(\psi+\phi)/2}\cos\frac{\theta}{2} \end{pmatrix} . \tag{5.35}$$

它与公式 (5.24) 中的 3×3 实转动矩阵 A 对应。这个表示就被称为转动矩阵的 Cayley-Klein 表示，或者 Cayley-Klein 参数。它的另外一个名称是单位四元数表示。

¶读者也许要问，我们已经有了关于三维转动欧拉角的描述了，为什么我们还需要一个四元数描述呢？这个问题的答案分为两个层面。首先，当然是因为三维的转动实在是太太太重要啦！它在各个方面的应用广泛。对于一个应用如此广泛的东东值得拥有多个备份。从另一个角度来说，虽然欧拉角可以给出三维转动的一个准确无误的描述，但是按照第 89页的描述，所有三维转动的相流形实际上是 RP^3。于是，我们可以将欧拉角表示视为从 T^3 到 RP^3 的一个映射。这个映射构成了两个拓扑空间之间的一个局域的地图。但是由于两者拓扑结构上的区别，我们不可能找到一个照顾到全局的所谓 covering map。这个映射的导数矩阵在某些点必定不是满秩的。这个数学上的结论的一个直接的后果是，这个映射中存在所谓的 万向节锁死 (gimbal lock) 问题。它的具体体现是，对于某些特定的角度 (例如对于 $\theta = 0$)，欧拉角表示仅仅依赖于两个角度的某个组合而不是两个独立的角度。换句话说，维数从三维退化为了二维。这一点我们可以从公式 (5.24) 中直接发现，当 $\theta = 0$ 时我们有，

$$A = \begin{pmatrix} \cos(\psi+\phi) & \sin(\psi+\phi) & 0 \\ -\sin(\psi+\phi) & \cos(\phi+\psi) & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{5.36}$$

这表示沿着 z 轴的一个转动，但是转动角仅仅依赖于 $\psi+\phi$。这就是万向节锁死问题。这个问题是一个拓扑问题，对欧拉角表示来说是无法克服的，无论我们用什么约定。相比较来说，单位四元数表示要好很多。首先，它的相流形 S^3 非常接近于 RP^3。只要我们将球面上相对的两点等同起来，就变成了 RP^3。当然，它毕竟不是 RP^3。这个体现在它的映射是二对一的，而不是一一的。但是除此之外，它是所能找到的最理想的表示。

万向节锁死问题的关键是，它不一定出现在什么时候。设想我们在飞行模拟过程中，任何情况都可能出现。这时利用一个没有万向节锁死问题的表述是十分重要的。同样的问题也出现在 3D 游戏之中。游戏者经常要通过游戏操纵杆或者鼠标变换视角，如果遇到万向节锁死问题，将是非常恼人的。另外，游戏中还经常用到从一个画面过渡到另一个画面的连续变换视角操作 (相信玩儿过一些 RPG 游戏的同学应当都有印象吧)。这时利用单位四元数表示是非常合适的。在计算机图像中应用很广的所谓 slerp(spherical linear interpolation) 算法就是这样一个函数。有兴趣的同学可以参考相关的文献。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

24 刚体的动能、角动量与惯量张量

¶ 前面漫长的一节涉及的是刚体的运动学的描述。这一节中我们讨论一个刚体的动能、角动量和惯量张量。这是计算刚体的拉格朗日量的重要一步，也属于介于运动学和动力学之间的衔接部分。我们会看到，刚体的动能以及角动量都与它的惯量张量密切联系。

一个刚体的动能可以通过计算组成它的各个质点动能之和得到。为此，我们利用刚体中任意一点的速度公式 (5.19)：

$$T = \frac{1}{2}\sum \delta m \left(\mathbf{v}_c + \mathbf{\Omega}\times\mathbf{r}\right)^2 \ . \tag{5.37}$$

这个式子中的求和可以认为是对刚体上所有的质点 -其质量为 δm-的求和。$\mathbf{r}$ 是该质心到被求和的质点的位置矢量。因此，求和过程中 $\mathbf{r}$ 会随着不同的质点而变化。事实上，$\mathbf{r}$ 会遍及刚体中所有的质点。但是对一个刚体而言，它的质心速度 $\mathbf{v}_c$ 和角速度 $\mathbf{\Omega}$ 是唯一的，不会随求和时质点的变化而改变。这点需要大家特别注意。将这个式子展开后得到：

$$T = \frac{1}{2}m\mathbf{v}_c^2 + \frac{1}{2}\sum \delta m \left[\Omega^2 r^2 - (\mathbf{\Omega}\cdot\mathbf{r})^2\right] \ . \tag{5.38}$$

其中 $m = \sum \delta m$ 表示刚体的总质量。这个式子中没有交叉项是因为：$\sum \delta m \mathbf{v}_c \cdot (\mathbf{\Omega}\times\mathbf{r}) = \mathbf{v}_c \cdot (\mathbf{\Omega}\times(\sum \delta m \mathbf{r})) = 0$。这是由于我们取了刚体的质心为坐标原点。这个表达式的物理意义十分明显。第一项代表刚体整体的平动动能；第二项则是刚体绕质心的转动动能。刚体的转动能可以更加明确地写成：

$$\boxed{T_{rot} = \frac{1}{2}\Omega_i \Omega_j I_{ij} \ ,} \tag{5.39}$$

其中我们引入了刚体的转动惯量张量，也简称为惯量张量：

$$I_{ij} = \sum \delta m (\mathbf{r}^2 \delta_{ij} - \mathbf{r}_i \mathbf{r}_j) \ . \tag{5.40}$$

显然 I_{ij} 是一个对称的二阶张量。因此，我们可以将刚体的总动能表达为：

$$\boxed{T = \frac{1}{2}m\mathbf{v}_c^2 + \frac{1}{2}I_{ij}\Omega_i\Omega_j \ .} \tag{5.41}$$

这就是一个刚体的总动能的表达式。它二次地依赖于刚体的质心速度和刚体的角速度。这些二次型前面的系数由刚体的惯性特性所决定。对于一个刚体来说，它的惯性特性由其总质量 m 和绕质心的惯量张量 I_{ij} 所完全描述。[8]

[8] 具体的刚体的惯量张量的计算大家应当在普通物理力学中已经涉及，我们这里不再赘述。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

对于对称的惯量张量，我们总可以适当地选取坐标架使得它被对角化。这样的方向被称为刚体惯量张量的主轴方向。[9] 相应的对角化后的数值：I_1，I_2，I_3 则被称为主轴转动惯量。这时刚体的转动动能部分可以简洁地写为：

$$T_{rot} = \frac{1}{2}(I_1\Omega_1^2 + I_1\Omega_2^2 + I_1\Omega_3^2) \ . \tag{5.42}$$

特别需要注意的是，主轴坐标是相对于刚体本身固定的坐标架，它一般不是在空间固定的坐标架 (因为刚体在运动)。现在我们明白为什么上一节中我们关心一个固定在刚体上面的动坐标架与固定坐标架之间的关系了。刚体的转动能在相对于刚体固定的主轴坐标架中是最为简洁的。例如，如果我们选取上一节中讨论的附着在刚体上的坐标架 xyz 为其主轴坐标架，那么刚体的转动能就可以利用其欧拉角表达出来：

$$T_{rot} = \frac{I_1}{2}(\dot\phi\sin\theta\sin\psi + \dot\theta\cos\psi)^2 + \frac{I_2}{2}(\dot\phi\sin\theta\cos\psi - \dot\theta\sin\psi)^2 + \frac{I_3}{2}(\dot\phi\cos\theta + \dot\psi)^2 \ . \tag{5.43}$$

这个公式对于一个对称陀螺 ($I_1 = I_2 \neq I_3$) 可以进一步简化为：

$$\boxed{T_{rot} = \frac{I_1}{2}(\dot\phi^2\sin^2\theta + \dot\theta^2) + \frac{I_3}{2}(\dot\phi\cos\theta + \dot\psi)^2 \ .} \tag{5.44}$$

我们在后面讨论对称陀螺的动力学时候会用到这个公式。

¶ 一个刚体绕其质心的角动量也可以利用上面引入的惯量张量简单地表达出来。为此我们注意到角动量的表达式：

$$\begin{aligned} \mathbf{L} &= \sum \delta m \mathbf{r}\times[\mathbf{\Omega}\times\mathbf{r}] \\ &= \sum \delta m[\mathbf{\Omega}r^2 - \mathbf{r}(\mathbf{r}\cdot\mathbf{\Omega})] = \overleftrightarrow{I}\cdot\mathbf{\Omega} \ . \end{aligned}$$

其中我们用 $\overleftrightarrow{I}$ 来表示惯量张量，即 $(\overleftrightarrow{I})_{ij} \equiv I_{ij}$。这样我们得到：

$$\boxed{\mathbf{L}_i = I_{ij}\mathbf{\Omega}_j \ .} \tag{5.45}$$

25 刚体的动力学

¶ 我们首先在空间固定坐标系 (假定这是一个惯性系) 中来表述刚体的动力学。刚体具有六个力学自由度，它们可以选为质心的坐标 (三个) 和固着在刚体上的动坐标架相对于空间固定坐标架的三个角度 (例如三个欧拉角) 来表征。

[9] 显然，如果刚体本身具有某种几何的对称性，那么它的主轴方向总是沿着它的对称轴的。但是即使是完全没有任何对称性的刚体也是存在主轴的。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

一般来说，一个刚体的拉格朗日量可以表达为它的动能 T 与势能 $V(\mathbf{r})$ 之差：$L = T - V$，其中的动能 T 又可以分为质心的平动动能和绕质心的转动动能。刚体的势能的具体表达式则取决于具体的问题。

描写刚体质心运动的方程就是：

$$\frac{d\mathbf{P}}{dt} = \mathbf{F}\,, \tag{5.46}$$

其中 $\mathbf{P} = m\mathbf{v}_c$ 就是刚体的总动量，$\mathbf{F}$ 是刚体所受到的外力总和。刚体各个质点之间的内力可以完全不必考虑因为它们对于质心运动没有影响。这个方程完全描述了刚体的质心运动规律。

刚体的另外三个自由度（转动自由度）的运动规律可以通过几种方式来描述。如果利用拉格朗日分析力学的方式，我们可以写下：

$$\frac{d}{dt}\frac{\partial L}{\partial \boldsymbol{\Omega}} = \frac{\partial L}{\partial \boldsymbol{\phi}}\,,$$

由于：$\partial L/\partial \Omega_i = I_{ij}\Omega_j = \mathbf{L}_i$，同时势能 $V(\mathbf{r})$ 对于角度的变化可以表达为：

$$\begin{aligned} \delta V &= -\sum \mathbf{f}\cdot\delta\mathbf{r} = -\sum \mathbf{f}\cdot(\delta\boldsymbol{\phi}\times\mathbf{r}) \\ &= -\delta\boldsymbol{\phi}\cdot\sum \mathbf{r}\times\mathbf{f} = -\mathbf{N}\cdot\delta\boldsymbol{\phi}\,. \end{aligned} \tag{5.47}$$

因此：$-\partial V/\partial\boldsymbol{\phi} = \mathbf{N}$，这里 $\mathbf{N}$ 代表作用在刚体上的总力矩。同样的，这里仅仅需要考虑外力，因为刚体内部的内力的力矩为零。于是我们可以将刚体转动自由度的运动方程写为：

$$\boxed{\frac{d\mathbf{L}}{dt} = \mathbf{N}\,.} \tag{5.48}$$

方程 (5.46) 和方程 (5.48) 就完全描述了一个刚体的动力学规律。这两个方程也是大家在普通物理力学中已经得到过的的结果。这里需要注意的是，这两个方程中各个矢量的分量都是相对于空间固定坐标架 (假定是一个惯性系) 来定义的。

¶现在我们试图用相对于刚体静止的坐标架来讨论刚体的动力学问题。我们前面已经看到了，这种坐标架有它的方便之处，因为我们可以选取刚体的惯量主轴使得刚体的惯量张量具有对角的形式。刚体所具有的的任意一个矢量 $\mathbf{G}$ 相对于一个固着在刚体上的运动坐标架的时间微商与空间固定坐标架的时间微商之间的关系我们在前面已经得出了，即公式 (5.18)。我们现在将这个关系运用到运动方程 (5.48)，我们就得到：

$$\frac{d'\mathbf{L}}{dt} + \boldsymbol{\Omega}\times\mathbf{L} = \mathbf{N}\,. \tag{5.49}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

现在假定我们取固着在刚体上面的坐标架为刚体的惯性主轴，那么写成分量我们有：

$$\begin{cases} I_1\dfrac{d\Omega_1}{dt}+(I_3-I_2)\Omega_2\Omega_3=\mathbf{N}_1\ , \\ I_2\dfrac{d\Omega_2}{dt}+(I_1-I_3)\Omega_3\Omega_1=\mathbf{N}_2\ , \\ I_3\dfrac{d\Omega_3}{dt}+(I_2-I_1)\Omega_1\Omega_2=\mathbf{N}_3\ . \end{cases} \tag{5.50}$$

这一组方程就是关于刚体转动的著名的欧拉方程。需要特别注意的是，欧拉方程中各个矢量的分量都是按照刚体惯量主轴投影的分量。在以下几节中，我们将利用刚体动力学的基本方程来讨论几个刚体运动的典型例子。

例题 5.1　平面上纯滚的球 考虑一个水平平面上做纯滚的刚体球体，给出它最一般的运动。假定我们已知球体的质量为 M，半径为 a，绕过球心的轴的转动惯量为 I。

解：

我们在空间固定坐标架 (假定为一惯性系) 中求解这个问题。取水平平面为 xy 平面，z 轴垂直向上。

对质心运动而言，显然球体如果有加速度，一定沿着 xy 平面；换句话说球体在 z 轴的力一定是平衡的。因此我们只需要列出 xy 方向的牛顿方程：

$$F_1=M\ddot{x}_1\ ,\quad F_2=M\ddot{x}_2\ . \tag{5.51}$$

其中的 F_1，F_2 是平面作用于球体上的力的 x 和 y 分量。它们的作用点一定位于球体与平面的接触点。它们也提供了所有可能的绕球心的力矩。因此力矩的方程给出：

$$N_1=F_2a=I\dot{\omega}_1\ ,\ N_2=-F_1a=I\dot{\omega}_2\ ,\ N_3=0=I\dot{\omega}_3\ , \tag{5.52}$$

现在我们可以利用纯滚条件：$\dot{x}_1=\omega_2 a$，$\dot{x}_2=-\omega_1 a$，于是综合上述方程我们发现：

$$F_1=M\ddot{x}_1=Ma\dot{\omega}_2=-(I/a)\dot{\omega}_2\ , \tag{5.53}$$

这显然只能够在 $\dot{\omega}_2=0$ 时成立。类似地，我们可以证明 $\dot{\omega}_1=0$。因此，我们一定有 $F_1=F_2=0$，所有的 ω_1，ω_2，ω_3 均为常数。需要注意的是，ω_3 并不一定为零。其实大家如果仔细观察过滚动着的足球的同学应当都注意到 (通过足球上面的花纹的变化)，一般来说足球很少是 $\omega_3=0$ 的滚动，而是以 $\omega_3\neq 0$ 进行滚动。 搞定

26　刚体的动力学的矩阵描述

¶前面曾经提到，刚体的六个自由度中质心部分的自由度对应的相流形一般是三维欧氏空间 $\mathbb{R}^3$ 而其转动部分的相流形则对应于 $SO(3)$ 群的流形。上一节中我们利用欧拉角的描述来刻画 $SO(3)$ 的动力学。事实上，也可以直接用 $SO(3)$ 矩阵随时间的变化来描写刚体转动部分的动力学。这一节中我们就简要地介绍这种方法，更多详细的介绍可以参考 [4] 的 §8.2。

由 CPHOWIKI 小组整理自互联网　仅供学习交流使用　更多资料请访问 https://pan.cpho.wiki

27 自由不对称陀螺

¶这一节中，我们首先来讨论最为简单的情形，那就是一个自由的(不受任何力或力矩的) 刚体的运动规律。[10] 一个刚体的惯量张量可以在选择其主轴后变为对角的形式。这时，我们可以按照刚体惯量张量的对角元的情形分为：不对称陀螺，即 $I_1 \neq I_2 \neq I_3$；对称陀螺，即 $I_1 = I_2 \neq I_3$ 和 球形陀螺，即 $I_1 = I_2 = I_3$。[11] 显然，不对称陀螺是最为一般的情形。我们下面来讨论它的自由运动，也就是没有任何外力情形下的运动。为了明确起见，我们首先假定：$I_1 < I_2 < I_3$。

如果一个刚体 (陀螺) 不受任何外力和外力矩，那么它的质心显然保持匀速运动。因此，不失一般性，我们可以假定它的质心是静止的。由于陀螺不受外力矩，描写它运动的欧拉方程 (5.50) 为：

$$\begin{cases} I_1 \dfrac{d\Omega_1}{dt} = (I_2 - I_3)\Omega_2\Omega_3 \,, \\ I_2 \dfrac{d\Omega_2}{dt} = (I_3 - I_1)\Omega_3\Omega_1 \,, \\ I_3 \dfrac{d\Omega_3}{dt} = (I_1 - I_2)\Omega_1\Omega_2 \,. \end{cases} \tag{5.54}$$

这一组联立的常微分方程至少已经有了两个初积分：刚体的能量和角动量。利用这两个初积分我们可以得到：

$$\frac{L_1^2}{2EI_1} + \frac{L_2^2}{2EI_2} + \frac{L_3^2}{2EI_3} = 1 \,, \tag{5.55}$$

$$L_1^2 + L_2^2 + L_3^3 = L^2 \,. \tag{5.56}$$

这个公式中 (5.55) 来自于能量守恒；而 (5.56) 则来自于角动量守恒。需要注意的是，角动量守恒保证角动量矢量的每个分量在空间固定 坐标架上的每个分量都是守恒量，但这并不意味着角动量矢量在随刚体一起运动的坐标架 (例如惯性主轴坐标架) 的分量是守恒量。事实上，它们都是随着时间变化的。当然，三个分量的平方和永远等于角动量大小的平方，这当然是常数。

从几何上讲，公式 (5.55) 在主轴坐标架中确定了一个中心在原点，三个半长轴分别是：$\sqrt{2EI_1}$，$\sqrt{2EI_2}$ 和 $\sqrt{2EI_3}$ 的椭球 (分别对应于图 5.2中的 Π_1, Π_2 和 Π_3 轴)。公式 (5.56) 则确定了一个中心在原点，半径是 L 的球。一个不对称陀螺运动时由于要同时满足这两个方程，因此它的角动量矢量 $\mathbf{L}$ 的端点一定在上述椭球和球的相交线上运动。由于我们一定有：

$$2EI_1 < L^2 < 2EI_3 \,, \tag{5.57}$$

[10] 下面大家会看到，它实际上并不是那么简单。

[11] 不对称陀螺的典型例子如网球拍；对称陀螺的一个例子是地球（两极稍扁、赤道稍凸出）。注意，球形陀螺不一定非得是球形的，只要它的三个主轴惯量相等即可。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

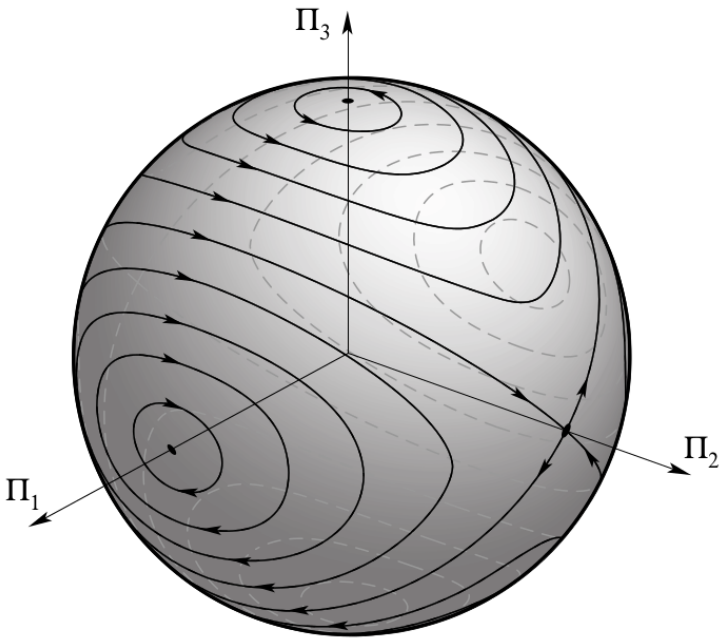

图 5.2: 自由不对称陀螺的运动。公式 (5.55) 和 (5.56) 的交线会在球和椭球的表面描绘出像篮球一样的花纹。

因此，球体的半径一定介于椭球的最短半轴长和最长半轴长之间。也就是说，球体与椭球一定会相交。具体来说，如果 L^2 仅仅比 $2EI_1$ 大一点儿，显然球体 (5.56) 会与椭球的最短轴附近 (也就是沿着 x_1 轴) 相交出两个闭合的曲线，如果 $L^2 \to 2EI_1$，那么这两条封闭曲线会缩成 $x_1 = \pm\sqrt{2EI_1}$ 的两极。如果 $L^2 > 2EI_1$ 并且我们不断增加 L^2，那么上面在 x_1 两极附件的闭合曲线会不断长大。当 $L^2 = 2EI_2$ 时，这两条闭合曲线变成相交于 x_2 两极的两个椭圆。如果 L^2 进一步增大以至于 $L^2 > 2EI_2$，那么球体与椭球的相交线会变成环绕着 x_3 两极的两个闭合曲线。当 $L^2 \to 2EI_3$ 时，环绕着 x_3 两极附近的两条闭合曲线会缩成 x_3 轴的两个极点。由于这些相交线都是闭合的，因此矢量 $\mathbf{L}$ 相对于刚体主轴的投影都是时间的周期函数。另外，对于一个不对称陀螺，它的椭球上的三对极点的性质是不同的。在 x_1 和 x_3 轴的两极 (即转动惯量最大或最小的两个主轴)，陀螺的运动是稳定的，而围绕 x_2 轴 (即转动惯量居中的主轴) 两极点的转动是不稳定的。这个结论有着不同的名字，例如网球拍定理 (tennis racket theorem)，或者 Dzhanibekov 效应，据说是这位前苏联宇航员首先观察到的。[12]

¶以上的讨论主要是定性的。为了能够得到角速度矢量 (或者等价地说，角动量矢量) 分量随时间变化的具体规律，我们必须利用能量守恒、角动量守恒的方程中消去 Ω_1^2 和

[12] 有兴趣的同学可以谷歌一下 “Dzhanibekov effect”。然后可以跟随结果中非 Youtube 的链接 (我假定大家都很乖，没有翻墙)，去观看到这个效应的视频。例如参见http://www.ebaumsworld.com/video/watch/82442953/。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

Ω_3^2：

$$\Omega_1^2 = \frac{[(2EI_3 - L^2) - I_2(I_3 - I_2)\Omega_2^2]}{I_1(I_3 - I_1)}, \tag{5.58}$$

$$\Omega_3^2 = \frac{[(L^2 - 2EI_1) - I_2(I_2 - I_1)\Omega_2^2]}{I_3(I_3 - I_1)}, \tag{5.59}$$

将此式带入到 Ω_2 的欧拉方程中得到：

$$\sqrt{I_1 I_3} I_2 \frac{d\Omega_2}{dt} = \sqrt{[(2EI_3 - L^2) - I_2(I_3 - I_2)\Omega_2^2][(L^2 - 2EI_1) - I_2(I_2 - I_1)\Omega_2^2]}. \tag{5.60}$$

将这个式子两面积分后原则上就可以得到 $\Omega_2(t)$。这个积分是一个椭圆积分，并可以化为其标准的形式。为此，我们先假定 $L^2 > 2EI_2$，否则的话只需要将下列各个公式中的指标 1 和 3 对调即可。我们定义：

$$\tau = t\sqrt{\frac{(I_3 - I_2)(L^2 - 2EI_1)}{I_1 I_2 I_3}},$$

$$s = \Omega_2\sqrt{\frac{I_2(I_3 - I_2)}{2EI_3 - L^2}}. \tag{5.61}$$

$$k^2 = \frac{(I_2 - I_1)(2EI_3 - L^2)}{(I_3 - I_2)(L^2 - 2EI_1)}. \tag{5.62}$$

显然可以证明：$0 < k^2 < 1$。我们于是得到：

$$\tau = \int_0^s \frac{ds}{\sqrt{(1 - s^2)(1 - k^2 s^2)}}. \tag{5.63}$$

这就是椭圆积分的标准形式。它实际上定义了雅克比椭圆函数：[13]

$$s \equiv \mathrm{sn}\tau. \tag{5.64}$$

类似地，我们可以定义另外两个椭圆函数：

$$\mathrm{cn}\tau = \sqrt{1 - \mathrm{sn}^2\tau}, \quad \mathrm{dn}\tau = \sqrt{1 - k^2\mathrm{sn}^2\tau}. \tag{5.65}$$

利用这些定义，我们可以将不对称陀螺的角速度表达成：

$$\Omega_1(\tau) = \sqrt{\frac{(2EI_3 - L^2)}{I_1(I_3 - I_1)}}\mathrm{cn}\tau,$$

$$\Omega_2(\tau) = \sqrt{\frac{(2EI_3 - L^2)}{I_2(I_3 - I_2)}}\mathrm{sn}\tau,$$

$$\Omega_3(\tau) = \sqrt{\frac{(L^2 - 2EI_1)}{I_3(I_3 - I_1)}}\mathrm{dn}\tau. \tag{5.66}$$

[13] 椭圆函数的发现被认为是雅克比（C. G. J. Jacobi, 1804-1851）一生最伟大的数学发现。雅克比是一位犹太血统的德意志数学家、物理学家。他对于数论 (特别是利用雅克比椭圆函数讨论数论问题)、行列式、微分方程等等都有重要贡献。这里讨论的自由陀螺的解和下一节中讨论的重力场中的陀螺运动问题都是雅克比首先处理的（1829 年）。他在分析力学方面的贡献最重要的要数我们将在第 六章的第 34节中讨论的哈密顿 -雅克比理论。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

按照雅克比椭圆函数的定义，这些函数都是 τ 的周期函数，其周期为 $4K$，这里 K 是第一类完全椭圆积分：

$$K = \int_0^1 \frac{ds}{\sqrt{(1-s^2)(1-k^2s^2)}} , \tag{5.67}$$

只要给定数值 k^2，它的数值很容易通过数值计算 (或者查表) 得到。因此，如果用时间 t 来表示，一个自由陀螺的周期为：

$$T = 4K\sqrt{\frac{I_1I_2I_3}{(I_3-I_2)(L^2-2EI_1)}} . \tag{5.68}$$

¶ 需要注意的是，前面的讨论仅仅涉及到角速度矢量在主轴上面的投影随时间的变化规律。现在我们来讨论一下在空间固定坐标架中一个自由不对称陀螺的运动。在空间固定坐标架中，我们可以将 Z 轴取在 (守恒的) 角动量的方向。这样刚体的运动可以完全由欧拉角来描述。我们有：

$$L\sin\theta\sin\psi = L_1 = I_1\Omega_1 , \ L\sin\theta\cos\psi = L_2 = I_2\Omega_2 , \ L\cos\theta = L_3 = I_3\Omega_3 . \tag{5.69}$$

由此我们得到：

$$\cos\theta = \frac{I_3\Omega_3}{L} , \ \tan\psi = \frac{I_1\Omega_1}{I_2\Omega_2} . \tag{5.70}$$

只要将主轴坐标架中的解 (5.66) 代入到这个式子中，我们就可以得到欧拉角 θ 和 ψ 对于时间的依赖关系。同样，它们也都是时间的周期函数 (周期也是公式 (5.68) 给出的 T)。另外一个欧拉角 ϕ 对于时间的依赖关系比较复杂，我们这里就不再讨论了。详细的计算表明 [1]，它是两个时间周期函数的和，其中一个的周期仍然是公式 (5.68) 给出的 T，但是另外一个的周期等于 T'，并且 T'/T 不是有理数。[14] 因此，这使得一个自由的不对称陀螺在固定坐标架中的运动不是严格周期的。这一点似乎有一点与我们最初的预感不一致。至此我们看到，即使是对于不受任何外力和外力矩的自由陀螺，它的运动方程的解都足够的复杂。这一点应当说是相当变态的，也是我们起初所没有想到的吧。

28 对称陀螺的定点运动

¶ 这一节中我们讨论一个重力场中对称陀螺的定点运动。我们假定一个对称陀螺的对称轴的下端固定在 O 点。我们自然取 O 点为空间固定坐标架 XYZ 的原点，并且我们取 Z 轴垂直向上。注意，点 O 同时也是主轴坐标的原点。假定陀螺的质心到点 O 的距离为 l，那么一个重力场中的对称陀螺的拉格朗日量为：

$$L = \frac{1}{2}(I_1+ml^2)(\dot\theta^2+\dot\phi^2\sin^2\theta) + \frac{I_3}{2}(\dot\psi+\dot\phi\cos\theta)^2 - mgl\cos\theta , \tag{5.71}$$

其中各个物理量的定义可以参考图 5.3。对于一个陀螺的运动，我们称角度 ψ 的运动

[14] 我们称这样的两个周期 (频率) 是非公度 的。两个周期非公度的周期函数之和本身不是周期函数，但是它是一个准周期函数。参考第 35.3小节的讨论。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

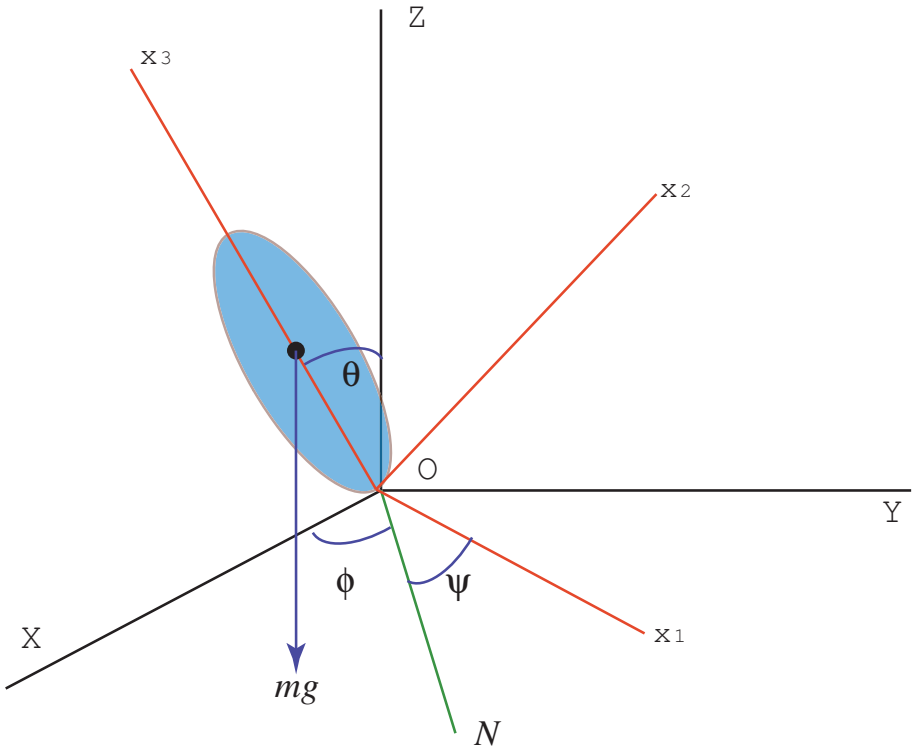

图 5.3: 重力场中的对称陀螺的定点运动。

为自转，因为它代表了一个对称陀螺绕其对称轴 (图中的 x_3 轴) 的转动；我们称角度 ϕ 的运动为进动 (precession)；称角度 θ 的运动为章动 (nutation)。[15]

由于拉格朗日量 (5.71) 中并不显含角度 ϕ 和 ψ，我们得知与它们共轭的广义动量都是守恒量：

$$p_\psi = I_3(\dot{\psi} + \dot{\phi}\cos\theta) = L_3\ ,\quad p_\phi = (I_1'\sin^2\theta + I_3\cos^2\theta)\dot{\phi} + I_3\dot{\psi}\cos\theta = L_Z\ , \tag{5.72}$$

其中 $I_1' = I_1 + ml^2$。L_3 和 L_Z 则代表了刚体绕原点 O 的角动量沿着 x_3 和 Z 方向的分量，它们都是守恒量。另外一个守恒的物理量就是陀螺的能量：

$$E = \frac{I_1'}{2}(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta) + \frac{I_3}{2}(\dot{\psi} + \dot{\phi}\cos\theta)^2 + mgl\cos\theta\ . \tag{5.73}$$

角动量的守恒方程可以给出：

$$\dot{\phi} = \frac{L_Z - L_3\cos\theta}{I_1'\sin^2\theta}\ ,\quad \dot{\psi} = \frac{L_3}{I_3} - \frac{L_Z - L_3\cos\theta}{I_1'\sin^2\theta}\cos\theta\ . \tag{5.74}$$

[15]章原指一段音乐的终结，拆字为从音从十，十是数字的最后一个引申为终结。说文曰：乐竟为一章。古人观天象，以十九年为一个天象周期，谓之一章。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

将此式带入到能量守恒的方程中，我们就可以得到只含章动角 θ 的等效能量守恒方程：

$$\begin{aligned} E' &= \frac{I_1'}{2}\dot{\theta}^2 + V_{eff}(\theta)\ . \\ E' &= E - \frac{L_3^2}{2I_3} - mgl\ , \\ V_{eff}(\theta) &= \frac{(L_Z - L_3\cos\theta)^2}{2I_1'\sin^2\theta} - mgl(1-\cos\theta)\ . \end{aligned} \tag{5.75}$$

这个方程中我们可以积分给出章动角 θ 对时间的依赖关系。同样的，这又会给出一个椭圆积分。将章动角 θ 的解代入到公式 (5.74) 中并积分，就可以得到另外两个欧拉角对时间的依赖关系。由于这些牵涉到复杂的椭圆函数，我们就不明确给出了。

定性讨论一下章动角 θ 和进动角 ϕ 的变化是有意义的。另外一个欧拉角只是反映了陀螺的自转，往往并不是十分重要。因此下面我们只关心陀螺的对称轴随时间的变化规律。这时比较直观的一种几何描述就是看这个对称轴与一个中心位于原点 O 的一个球面的交点在球面上所划出的轨迹。这个轨迹在球面上的“经度”和“纬度”正好对应于欧拉角 ϕ 和 θ。

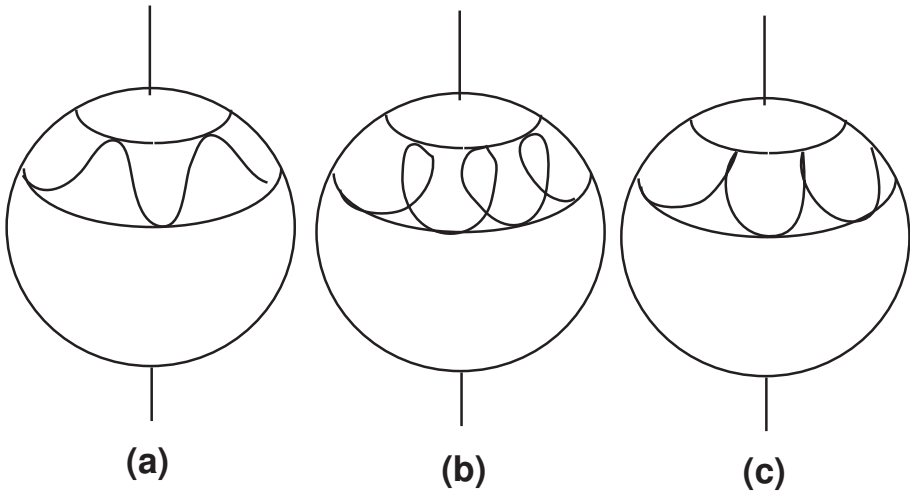

图 5.4: 重力场中的对称陀螺的章动与进动运动。

首先要注意到的一个重要事实是，角度 θ 只能够在一个范围：$[\theta_1, \theta_2]$ 之内变动，这里 θ_1，θ_2 由方程 $E' = V_{eff}(\theta)$ 给出。在章动角变化的过程中，进动角 ϕ 也按照方程 (5.74) 中的第一式变化。因此陀螺对称轴的进动存在两个极限圆，分别对应于 θ_1 和 θ_2。如果在章动角 θ 变化的范围内，$\dot{\phi}$ 并不改变符号 (也就是说 $(L_Z - L_3\cos\theta)$ 保持同一符号)，那么进动角 ϕ 的变化是单调的。这种情况被称为 正常章动 (见图 5.4中的情形 (a))。这时陀螺的对称轴在两个极限圆处的进动是同方向的。反之，如果 $\dot{\phi}$ 在章动角变化的范围内正好改变一次符号，这时的进动角的变化在两个极限圆处将是反方向的 (见图 5.4中的情形 (b))。如果恰好 $\dot{\phi}$ 在一个极限圆的地方为零，那么陀螺的对称轴在该极限圆处正好没有进动，

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

这使得对称轴在该极限圆处形成尖点 (见图 5.4 中的情形 (c))。上述三种情形都定性地显示在图 5.4中。

¶ 作为这一节的最后注解我们顺便指出，地球是一个相当不错的对称陀螺 (事实上，在第一级近似下它甚至可以看成是一个球形陀螺)。它沿着两极方向稍微扁一些。由于这种对于球形的微小偏离，太阳或月亮对于地球的引力会产生一个很微小的力矩。于是，地球的运动可以利用类似于本节处理的方法来研究。这种微小的力矩实际上会造成地球对称轴的方向绕着它的轨道平面的垂直方向缓慢的进动。之所以它是缓慢的，主要是由于地球十分接近于一个球体。换句话说，如果地球是一个严格的球体，那么太阳或月亮对地球的力矩等于零，这时这种缓慢的进动也不复存在。这正是地球总角动量守恒的要求，因为地球的自转轴正是它角动量的方向。如果没有外力矩的话，它的方向是不会改变的。在图 5.5中我们显示了地球 -太阳系统的情况。我们选取地球绕太阳公转的轨道平面 –通常称为 黄道面 (ecliptic)–的法向为空间固定坐标架的 Z 轴。地球的主轴中的对称轴是沿着它的自转方向的 (x_3 轴)，它与 Z 轴并不重合而是有一个大约 23.5° 的夹角 (图中的 θ 角)。[16] 正是这个角度的存在造就了我们地球上的春夏秋冬四季。由于地球偏离理想球体，因此太阳

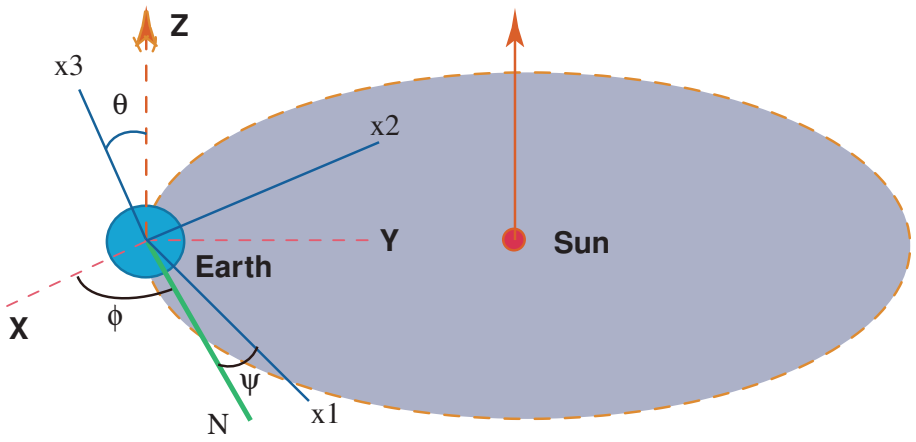

图 5.5: 地球作为一个对称陀螺会受到太阳 (和月亮) 的微小力矩的影响。这里显示了地球和太阳系统的情形。地球绕太阳公转的轨道平面 (黄道面) 的法向可以取为空间固定坐标架的 Z 方向。图中显示了地球的欧拉角，其中它的自转方向沿 x_3 轴，它与 Z 轴之间的夹角大于为 23.5 度。

对于它有微小的力矩，因此它的自转轴会缓慢进动。等效地来说，地球主轴运动的拉格朗日量可以写为：

$$L = \frac{1}{2}(I_1 + ml^2)(\dot{\theta}^2 + \dot{\phi}^2 \sin^2\theta) + \frac{I_3}{2}(\dot{\psi} + \dot{\phi}\cos\theta)^2 - V(\cos\theta)\,, \tag{5.76}$$

这里的 $V(\cos\theta)$ 当然与公式 (5.71) 中的不同，它是由于我们提到的微小的力矩所引起的势能。[17] 具体的计算表明 [3]，大约要过几万年它的方向会完成一个周期。这种变化实际上

[16] 这也就是南北回归线对应的纬线角度。

[17] 事实上，这个势能可以表达为：$V(\cos\theta) = -GM(I_3 - I_1)P_2(\cos\theta)/r^3$，这里 M 是太阳质量，G 是牛

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

体现了 春分秋分点 (equinox) 的 移动 。

粗看起来似乎太阳的影响是最为主要的。但是细心的读者应当注意到了，与第 16节中讨论潮汐一样，这个有效的势能 $V(\cos\theta)$ 是正比于 M/r^3 的。因此，实际上月亮对于地球的影响反而要更加重要一些。真正的地球进动的效果是由月亮和太阳的共同影响造成。太阳系中其他的星体的贡献几乎可以忽略。另外一点复杂的情况是月亮绕地球的轨道平面并不与日地轨道平面 (黄道面) 完全重合 (大约有 5° 的偏离)，因此这种不一致性会造成地球的进动的不规则性，这在天文学中称为地球的 天文章动 (astronomical nutation)。



相关的阅读

本章中我们讨论了刚体的运动学与动力学。这是一个相当技术性的章节，也是理论力学中独具特色的一部分内容。毫无疑问，刚体的动力学是理论力学不可缺少的内容，但是它与其他课程的联系并不是十分密切。当然，角动量、转动惯量、惯量主轴、欧拉角等等概念是十分重要的。在以后的量子力学课程中，同学们还会遇到它们。

从本章的具体内容来说，第 23 节和第 24 节的内容可以参考朗道书 [1] 的第 §31，§32，§35 节。第 25 节的讨论取自朗道书 [1] 的第 §34，§36 节。第 27 节的讨论取自朗道书 [1] 的第 §37 节。第 28 节的讨论则取自朗道书 [1] 的第 §35 节的例题。类似的讨论也可以参考 Goldstein 书中的第五章的 5-7 节。特别是关于地球自转轴的进动与章动的讨论主要取自 Goldstein 的第五章的 5-8 节。

顿引力常数，I_3 和 I_1 分别是地球的主轴惯性矩，r 是日地距离，$P_2(\cos\theta) = (3\cos^2\theta - 1)/2$ 是二阶勒让德多项式。关心其推导细节的同学可以参考 [3] 的 5-8 节。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

第六章　　哈密顿力学

本章提要

- 哈密顿正则方程（ 29）
- ξ 符号与 Liouville 定理（ 30）
- 泊松括号（ 31）
- 作用量作为端点的函数（ 32）
- 正则变换（ 33）
- 哈密顿－雅克比理论（ 34）
- 作用量－角度变量（ 35）

前面几章中我们主要建立了所谓的拉格朗日形式的分析力学体系并且讨论了它的一些重要的典型应用。这一章中，我们将重新回到分析力学的理论体系的讨论中。我们要介绍另一种与拉格朗日力学等价的分析力学描述方式－哈密顿力学。

大家也许会感到奇怪，对于经典力学我们已经有了牛顿的矢量力学体系和拉格朗日的分析力学体系。为什么还需要一个哈密顿力学体系呢？对于这个问题的答案实际上需要超出经典力学之外才能够看得更加清楚。正像我们在本课程一开始所讨论的那样，牛顿的矢量力学与拉格朗日的分析力学在纯经典力学的范畴内是完全等价的。如果只是在纯力学的范畴内考察，我们顶多能够说拉格朗日分析力学有时候更加方便而已，特别是对于有约束的力学系统。事实上，拉格朗日分析力学的优势只有在超出纯力学问题时才显得更加明显。同样，我们需要哈密顿力学的原因实际上也是超出纯经典力学的。在纯经典力学范畴之内，它只是拉格朗日分析力学的另外一种等价的描述。但是，在超出经典力学之外，特别是进入量子力学时，哈密顿力学就显示出它的优点。事实上，在历史上量子力学最初

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

的建立很大程度上是受到哈密顿力学的启发。[1] 同时，哈密顿力学还显示出与波动光学、统计力学等学科的密切联系。因此，对于哈密顿力学的讨论实际上将使我们对于经典力学的理论体系的认识更加深入，同时也为其他后续课程奠定基础。

29 哈密顿正则方程

¶考虑一个由拉格朗日量 $L(q,\dot{q},t)$ 描述的经典力学系统。为了简化记号，我们用 q 和 $\dot{q}$ 来代表该系统所有的广义坐标 q_i 和广义速度 $\dot{q}_i$，其中 $i=1,2,\cdots,f$，f 是系统的自由度数目。我们可以写出拉格朗日量的全微分：

$$dL=\frac{\partial L}{\partial t}dt+\frac{\partial L}{\partial q_i}dq_i+\frac{\partial L}{\partial \dot{q}_i}d\dot{q}_i\ ,$$

其中我们运用了爱因斯坦求和规则，即对于重复的指标意味着求和 (从 1 到 f)。对于每一个广义坐标 q_i，我们曾定义了与之共轭的广义动量 p_i(参见第 7节中的公式 (2.18))：

$$\boxed{p_i=\frac{\partial L}{\partial \dot{q}_i}\ .} \tag{6.1}$$

利用广义动量表达，并结合系统的拉格朗日运动方程，拉格朗日量的全微分变为：

$$dL=\frac{\partial L}{\partial t}dt+p_i d\dot{q}_i+\dot{p}_i dq_i\ . \tag{6.2}$$

现在我们定义新的的一个物理量，它是函数 $L(q,\dot{q},t)$ 的勒让德变换：

$$\boxed{H(p,q,t)=p_i\dot{q}_i-L\ .} \tag{6.3}$$

我们称以广义坐标和共轭的广义动量为变量的函数 $H(p,q,t)$ 为系统的哈密顿量。哈密顿量的物理意义实际上我们在第 二章的第 11节中讨论过。对于一个具有时间平移对称性的系统，哈密顿量的数值就是系统守恒的能量。虽然哈密顿量在数值上就等于系统的能量，但是它与能量是有区别的。特别需要强调的一点是，哈密顿量是作为力学体系的广义坐标以及广义动量的函数。它与拉格朗日量不同，拉格朗日量是作为广义坐标和广义速度的函数。也就是说，按照定义 (6.3)，当我们已知系统的拉格朗日量后，需要按照定义 (6.3) 构造系统的哈密顿量，同时，我们必须利用广义动量的定义 (6.1) 将公式 (6.3) 中的所有广义速度利用广义坐标和广义动量来表达。这样最终得到的以广义坐标和广义动量为变量的函数才是系统的哈密顿量。也就是说，当我们说给定了力学系统的

[1] 如果我们考虑量子力学的描述问题，基于哈密顿力学的正则量子化和基于拉格朗日力学的路径积分量子化是量子力学两种等价的描述。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

哈密顿量的时候，强调的绝不仅仅是它的数值，更重要的是它对于广义坐标和广义动量的函数依赖关系。容易验证，哈密顿量的微分为：

$$dH = \dot{q}_i dp_i - \dot{p}_i dq_i - \frac{\partial L}{\partial t} dt \,. \tag{6.4}$$

这个关系可以从拉格朗日函数的微分 (6.2) 以及哈密顿量的定义 (6.3) 直接取微分得到。从拉格朗日函数 $L(q, \dot{q}.t)$ 变换到哈密顿函数 $H(p, q, t)$ 的变换被称为勒让德变换。这种变换在热力学中会经常遇到。

将哈密顿量的微分 (6.4) 与它的完全微分进行比较，我们立刻得到：

哈密顿正则方程 (Hamilton's Canonical Equations)

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \,, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \,. \tag{6.5}$$

这一组方程就是著名的哈密顿方程。它们又被称为正则方程或者哈密顿正则方程。从数学上讲，这一组微分方程与系统的欧拉 -拉格朗日方程完全等价，只不过拉格朗日方程是二级常微分方程，而哈密顿方程是一阶常微分方程。因此，只要给定初始的广义坐标和广义动量，力学系统的运动就被方程 (6.5) 所完全确定了。

特别需要注意的是，如果我们考察哈密顿量随时间的变化率，我们有：

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} + \frac{\partial H}{\partial q_i}\dot{q}_i + \frac{\partial H}{\partial p_i}\dot{p}_i \,.$$

利用哈密顿正则方程 (6.5) 代入，我们发现：

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \,. \tag{6.6}$$

所以，如果系统的哈密顿量不显含时间 (例如，如果系统的拉格朗日量不显含时间，那么它的哈密顿量也不显含时间)，这个力学系统的哈密顿量是一个守恒量，而其守恒的数值正是系统的能量，见第 二章公式 (2.52)。

¶ 哈密顿方程是关于广义坐标和广义动量对 (p, q) 的一阶常微分方程。我们随后会看到，在哈密顿分析力学的理论框架中，所谓“坐标”和“动量”的概念其实不是一成不变的。事实上，两者完全可以“互换”。因此，更准确的称呼是称它们互为共轭变量。对于一个自由度数目是 f 的力学系统，我们有 f 对共轭变量，而系统的运动完全由这 $2f$ 个变量所描述。这 $2f$ 个独立的变量 (f 对变量) 为坐标轴可以构成一个 $2f$ 维的“空间”，它通常称为这个力学系统的相空间。[2] 一个力学系统在任意时刻的力学状态可以用其相空间中的一个点来表示，它称为这个力学系统在其相空间中的代表点。随着时间的推

[2] 更现代一些的称呼是相流形 (phase manifold)。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

移，力学系统的代表点也会在其相空间中移动从而划出一条轨迹，它称为这个力学系统的相轨道。从相空间中一个给定的点出发，力学系统的相轨道完全由其哈密顿正则方程所唯一确定。

利用相空间、相轨道这些“几何术语”来描述一个力学系统的运动有时是方便的。特别在讨论力学系统定性行为的时候。另外，在统计物理中，我们也会频繁地运用这些概念。特别需要强调指出的是，一个力学系统的相空间的几何结构 (拓扑性质) 可以是相当复杂的。它通常不是一个简单的 $2f$ 维的平直欧几里德空间。一个典型的例子就是当系统的某个广义坐标是角度的时候。例如，一个单摆的相空间是一个圆柱面。一些多自由度力学体系的相空间可以是具有十分丰富拓扑结构的弯曲流形。这时要利用几何的描述来讨论力学问题就必须运用微分几何的术语，[3] 事实上，相空间是力学系统位型流形上的余切丛 (cotangent bundle)。有关这方面更加细致的讨论可以参考 [4] 的第 5.2 节，或者参考 Arnold 的书 [6]。

¶只要给定了系统的拉格朗日量，我们就可以按照标准的程式得到它的哈密顿量。对于一个保守系统我们有 (参见第 10节中的公式 (2.49))：

$$L = \frac{1}{2}a_{ij}(q)\dot{q}_i\dot{q}_j - V(q_1, q_2, \cdots, q_f) \ . \tag{6.7}$$

按照定义，与 q_i 共轭的广义动量 p_i 为：

$$p_i = a_{ij}(q)\dot{q}_j \ . \tag{6.8}$$

这个公式中可以反解出 $\dot{q}_i$ 作为广义动量的函数：[4]

$$\dot{q}_i = a^{-1}_{ij}(q)p_j \ , \tag{6.9}$$

其中 $a^{-1}(q)$ 表示对称矩阵 $a(q)$ 的逆矩阵。简单的计算表明系统的哈密顿量为：

$$H = \frac{1}{2}a^{-1}_{ij}(q)p_ip_j + V(q_1, q_2, \cdots, q_f) \ . \tag{6.10}$$

也就是说，保守系统的哈密顿量可以写成动能项[5] 加上势能项。

例题 6.1 **电磁场中粒子的哈密顿量** 考虑一个相对论性的粒子与外加电磁场的相互作用。它的拉格朗日量由第 9节中的公式 (2.34) 给出。求它的哈密顿量。

解： 按照定义，粒子的正则动量 $\mathbf{P}$ 为：

$$\mathbf{P} = \frac{m\mathbf{v}}{\sqrt{1-\mathbf{v}^2/c^2}} + \frac{e}{c}\mathbf{A} = \mathbf{p} + \frac{e}{c}\mathbf{A} \ . \tag{6.11}$$

[3] 与哈密顿正则方程相应的几何往往被称为辛几何 (symplectic geometry)。

[4] 我们假定对称矩阵 $a(q)$ 不奇异。

[5] 不难验证，这个哈密顿量中的动能项 (第一项) 在数值上与拉格朗日量中的动能项 (第一项) 是相等的。只不过哈密顿量的动能项必须用广义坐标和广义动量表达，而拉格朗日量中的动能项是用广义坐标和广义速度表达的。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

于是，哈密顿量可以直接写出 (注意将原来拉格朗日量中所有的速度都换成正则动量 **P**):

$$H = \sqrt{m^2c^4 + c^2\left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right)^2} + e\Phi \ . \tag{6.12}$$

这就是一个带电粒子在外加电磁场中的哈密顿量 (相对论性的)。如果我们取非相对论近似，那么上式的根号中我们可以假定：$m^2c^4 \gg c^2[\mathbf{P} - (e/c)\mathbf{A}]^2$，因而我们展开根号并且略去静止能量的常数项得到:

$$H = \frac{1}{2m}\left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right)^2 + e\Phi \ . \tag{6.13}$$

这就是非相对论近似下一个带电粒子与外电磁场相互作用的哈密顿量。 搞定

例题 6.2　相对论性 Keppler 问题 利用哈密顿力学的方法求解一个相对论性的粒子在 $V(r) = -\alpha/r$ 的吸引势中的轨道问题。

解： 为了方便我们取 $c = 1$ 的单位，这时粒子的哈密顿量可以表达为:

$$H = \sqrt{\mathbf{p}^2 + m^2} - \frac{\alpha}{r} \ . \tag{6.14}$$

与非相对论的情形一样，这实际上是个二维问题。在选择了极坐标之后，我们有 $\mathbf{p}^2 = p_r^2 + p_\theta^2/r^2$，因此:

$$H = \sqrt{p_r^2 + \frac{p_\theta^2}{r^2} + m^2} - \frac{\alpha}{r} \ . \tag{6.15}$$

哈密顿量不显含时间也不显含 θ，因此机械能 E(哈密顿量的数值) 以及角动量 $p_\theta \equiv J$ 都是守恒量。因此我们有:

$$\left(E + \frac{\alpha}{r}\right)^2 = p_r^2 + \frac{J^2}{r^2} + m^2 \ .$$

另一方面我们可以验证：$p_r = \gamma m\dot{r}$ 以及 $p_\theta = J = \gamma m r^2\dot{\theta}$。因此我们有:

$$p_r = \frac{p_\theta}{r^2}\frac{dr}{d\theta} = \frac{J}{r^2}r' \ . \tag{6.16}$$

因此在进行了通常的替换 $u \equiv 1/r$ 之后，粒子径向的运动方程可以写为:

$$(E + \alpha u)^2 = J^2\left[(u')^2 + u^2\right] + m^2 \ ,$$

其中 $u' = du/d\theta$。如果你对这个方程还看不出如何处理的话，最简单的方法是将其对 θ 再微商一次后得到

$$u'' + \left(1 - \frac{\alpha^2}{J^2}\right)u = \frac{\alpha E}{J^2} \ . \tag{6.17}$$

这看上去是一个典型的“谐振子”型的微分方程，因此它的解为:

$$u \equiv \frac{1}{r} = \frac{1}{p}\left(1 + e\cos\left[\Omega(\theta - \theta_0)\right]\right) \ , \tag{6.18}$$

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

其中 θ_0 是一个无关紧要的常数 (标志极轴的选取) 而其他的各参数由下式给出:

$$\begin{cases} \Omega^2 = 1 - \dfrac{\alpha^2}{J^2}\ , \quad p = \dfrac{J^2}{\alpha E}\left(1 - \dfrac{\alpha^2}{J^2}\right)\ , \\ e^2 = 1 + \dfrac{\left(1 - \frac{\alpha^2}{J^2}\right)\left(1 - \frac{m^2}{E^2}\right)}{(\alpha^2/J^2)}\ . \end{cases} \tag{6.19}$$

需要注意的是，这里的能量 E 是包含了粒子静止能量的总能量。如果我们要将上述各式与第 三章第 15节中非相对论情形下的相应公式进行比较的话，我们需要令: $E \simeq m + \hat{E}$, 其中 $|\hat{E}| \ll m$(注意，对应束缚的情形，$\hat{E} < 0$)。同时，非相对论极限下 $\alpha^2 \ll J^2 \equiv J^2c^2$。于是大家可以比较容易地验证，在非相对论极限下这些公式的确都回到第 15节中相应的公式。有兴趣的同学还可以将这里的公式与第 18节中获得的同样问题的近似解进行比较。

这个问题在历史上是 Sommerfeld 在 1916 年首先研究的。利用所谓旧量子论的量子化条件，他成功地获得了可以描述氢原子精细结构的能级。这个结果实际上还是相当令人吃惊的，因为标准的基于量子力学的推导需要用到关于电子的狄拉克方程，而这是到 1928 年后才出现的。无怪乎当时爱因斯坦就认为量子化条件必定包含了一些重要的信息。[6] 由于具体的计算涉及到所谓的绝热不变量和角度 -作用量变量，我们将在后面 (见第 35.3小节的、例题 6.12) 讨论这个问题。 搞定

30 ξ 符号与 Liouville 定理

¶在哈密顿正则体系里面广义坐标和广义动量变得更加平等了。为此，我们可以将两者集合在一个变量里面。我们定义一个变量 ξ:

$$\begin{cases} \xi^j = q^j\ , \quad j = 1, \cdots, n, \\ \xi^j = p_{j-n}\ , \quad j = n+1, \cdots, 2n. \end{cases} \tag{6.20}$$

容易验明哈密顿正则方程可以改写为:

$$\dot{\xi}^j = \omega^{jk}\partial_k H\ , \tag{6.21}$$

其中 $\partial_j = \partial/\partial\xi^j$ 而 ω^{jk} 为下列 $2n \times 2n$ 反对称矩阵:

$$\omega = \begin{pmatrix} \mathbb{0}_{n\times n} & \mathbb{1}_{n\times n} \\ -\mathbb{1}_{n\times n} & \mathbb{0}_{n\times n} \end{pmatrix}\ . \tag{6.22}$$

而它的逆矩阵为 $\omega^{-1} = -\omega = \omega^T$，其矩阵元我们将用下标表示：$\omega_{jk}$，因此哈密顿正则方程亦可写为：$\omega_{jk}\dot{\xi}^k = \partial_j H$。

[6] 更多细节参见 Peter Vickers, "Historical Magic in Old Quantum Theory", European Journal for Philosophy of Science, 2011

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

考虑在 $t=0$ 时刻相空间的一个小体积元：$d^{2n}\xi(0)$，现在让该体积元中的点按照哈密顿正则方程演化。经过一段时间 t 之后，它变为体积元 $d^{2n}\xi(t)$。我们感兴趣的是这个体积元是如何随时间演化的。按照微积分它们之间相差一个 Jacobian 行列式：

$$J(\xi(t);\xi(0))=\det\left(\frac{\partial\xi(t)}{\partial\xi(0)}\right)=\det\left(\frac{\partial(\xi^1(t),\cdots,\xi^{2n}(t))}{\partial(\xi^1(0),\cdots,\xi^{2n}(0))}\right)\ . \tag{6.23}$$

容易证明它的时间演化规律为：

$$\frac{dJ}{dt}=J(\partial_j\dot{\xi}^j)\ , \tag{6.24}$$

由于 ω^{jk} 的反对称性我们显然有：$\partial_j\dot{\xi}^j=0$，这意味着相空间的体积元随时间是不变的。这个结论被称为 Liouville 定理。这个定理对于统计力学来说非常有启发性。

注意，对于任何的哈密顿系统，其运动方程由哈密顿正则方程给出。用等价的 ξ 符号描述，即由公式 (6.21) 给出。因此我们看到，Liouville 定理对所有的哈密顿系统都是成立的。特别是对非保守的哈密顿系统 (其哈密顿量显含时间，从而其能量并不守恒) 也是成立的。对于非哈密顿系统，我们可以有 $\partial_j\dot{\xi}^j\neq0$。这时我们可以将上述方程积分给出：

$$J(\xi(t);\xi(0))=\exp\left(\int_0^t ds\kappa(\xi(s))\right)\ . \tag{6.25}$$

这意味着我们可以引入一个权函数 $W(\xi(t))$，同时，$J(\xi(t);\xi(0))=\exp(W(\xi(t))-W(\xi(0))$。于是我们可以重新定义：$\sqrt{g(t)}\equiv e^{-W(\xi(t))}$，这样一来我们有：

$$\sqrt{g(t)}d^{2n}\xi(t)=\sqrt{g(0)}d^{2n}\xi(0)\ . \tag{6.26}$$

也就是说，即使在非哈密顿系统中，只要我们适当引入权函数，我们仍然可以得到一个守恒的相空间体积元。[7] 在这样的一个系统中，它的相空间看起来更像是具有弯曲度规的。

31 泊松括号

¶ 我们现在来考察一个力学量 $f(p,q,t)$，它可以是广义动量、广义坐标和时间的任意函数。它随时间的变化率可以写成：

$$\frac{df}{dt}=\frac{\partial f}{\partial t}+\frac{\partial f}{\partial q_i}\dot{q}_i+\frac{\partial f}{\partial p_i}\dot{p}_i\ .$$

将正则方程 (6.5) 代入得到：

$$\boxed{\frac{df}{dt}=\frac{\partial f}{\partial t}+[f,H]\ ,} \tag{6.27}$$

[7] 更多细节可以参考：“On the classical statistical mechanics of non-Hamiltonian systems”, M.E. Tuckerman, C.J. Mundy and G.J. Martyna, Europhys. Lett., 45 (2), pp. 149-155 (1999).

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

其中我们定义了两个力学量的泊松括号：[8]

$$[g,f]\equiv\frac{\partial g}{\partial q_i}\frac{\partial f}{\partial p_i}-\frac{\partial g}{\partial p_i}\frac{\partial f}{\partial q_i}\,. \tag{6.28}$$

注意，这个定义也可以利用前面 (6.22) 中定义的 ω 写为：

$$[g,f]\equiv(\partial_j g)\omega^{jk}(\partial_k f)\,. \tag{6.29}$$

如果一个力学量 $f(p,q,t)$ 随时间的变化率为零，它被称为力学体系的一个运动积分或守恒量。按照公式 (6.27)，一个力学量 $f(p,q,t)$ 是运动积分的条件可以表达为：

$$\frac{\partial f}{\partial t}+[f,H]=0\,. \tag{6.30}$$

特别对于不显含时间的力学变量 $f(p,q)$，它是一个运动积分 (守恒量) 的条件是：

$$[f,H]=0\,. \tag{6.31}$$

也就是说，它与哈密顿量的泊松括号必须等于零。

¶上面定义的两个力学量的泊松括号 (6.28) 满足一些基本的性质，它对于两个量是反对称的同时是双线性的：

$$[c_1f_1+c_2f_2,g]=-[g,c_1f_1+c_2f_2]=c_1[f_1,g]+c_2[f_2,g]\,, \tag{6.32}$$

其中 c_1，c_2 是任意常数。另外，两个函数乘积的泊松括号可以表达为：

$$[f_1f_2,g]=f_1[f_2,g]+f_2[f_1,g]\,. \tag{6.33}$$

另外一个重要的关系是所谓的雅克比等式：

$$[f,[g,h]]+[g,[h,f]]+[h,[f,g]]=0\,. \tag{6.34}$$

利用上述这些关系，任意两个已知函数的泊松括号都可以化简为基本的泊松括号：

$$[q_i,q_j]=[p_i,p_j]=0\,,\quad[q_i,p_j]=\delta_{ij}\,. \tag{6.35}$$

例题 6.3　角动量的泊松括号 考虑一个质点的角动量 $\mathbf{L}=\mathbf{r}\times\mathbf{p}$。计算角动量各个分量之间的泊松括号。

解： 直接利用基本的泊松括号的性质，我们可以得到：

$$[L_1,L_2]=L_3\,,\ [L_2,L_3]=L_1\,,\ [L_3,L_1]=L_2\,. \tag{6.36}$$

[8] 注意，有的教科书中泊松括号的定义与我们这里的相差一个负号，例如 [1]。这里的定义与 [3, 4] 一致。

由 CPHOWIKI 小组整理自互联网　仅供学习交流使用　更多资料请访问 https://pan.cpho.wiki

当然，任何一个分量与自身的泊松括号恒等于零。这三个等式可以统一地写成：

$$[L_i, L_j] = \epsilon_{ijk} L_k \ . \tag{6.37}$$

搞定

泊松括号的一个重要的性质是：如果 f，g 是力学系统的两个运动积分，那么它们的泊松括号 $[f,g]$ 一定也是该力学系统的运动积分。这个结论被称为 泊松定理。如果函数 f，g 都不显含时间，这个结论的证明是十分简单的。只要在雅克比等式中令：$h = H$，同时利用 $[f,H] = [g,H] = 0$，我们立刻得到：$[H,[f,g]] = 0$，即 $[f,g]$ 也是力学体系的运动积分。[9] 泊松定理的一个应用就是可能产生新的运动积分。如果我们已知系统的两个运动积分 f 和 g，如果 $[f,g]$ 并不是常数，或者简单的 f 和 g 的函数，那么 $[f,g]$ 就给出了系统的一个新的运动积分。

¶特别值得一提的是，这一节对于运动积分 (守恒量) 的讨论完全可以移植到量子力学中。我们所需要做的只是将经典的泊松括号换成量子力学中的 对易括号 就搞定了。

32 作用量作为端点的函数

¶在第 二章的第 7节讨论最小作用量原理的时候，我们考虑了作用量：

$$S = \int_{t_1}^{t_2} L(q, \dot{q}, t) dt \ . \tag{6.38}$$

为了得到力学体系的运动方程，我们考虑了 端点固定 的所有可能的轨道。最小作用量原理告诉我们，力学体系真实的运动轨道是所有这些端点固定的轨道中使得作用量取极小值的那个轨道。它的方程可以利用变分法给出，也就是欧拉－拉格朗日方程。现在，我们来考虑另外一种情形。我们假设体系处在它的真实轨道上，也就是说其轨迹满足运动方程，我们要考察作用量 S 对端点的依赖关系。具体来说，我们希望了解，在下端点仍然保持固定的情况下，作用量对于上端点的依赖关系。

¶为此我们假设体系起始点 t_0 仍然固定，但是分别沿着两个无穷接近的轨道 $\eta_1(t)$ 和 $\eta_2(t)$ 运动，即：

$$\eta_1(t_0) = \eta_2(t_0) = q^{(0)} \ .$$

我们下面将分别考虑这两条无穷接近的轨道所对应的作用量，只不过一条是到 $(t, q(t) = \eta_1(t))$，另一条则是到时间 $(t + dt, \eta_2(t + dt) = q(t) + dq(t))$，两者的出发点是共同的，都是 $(t_0, q^{(0)})$，参见图 6.1。因此我们实际上是在研究作用量作为后端点 (t, q) 的函

[9] 对于 f，g 显含时间的情况，也可以证明泊松定理。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

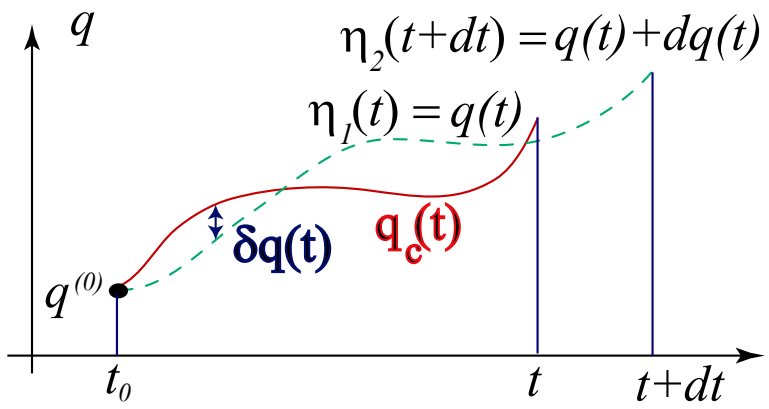

图 6.1: 作用量作为后端点的函数。

数依赖关系。上述两条无限接近的路径所对应的作用量之差为：

$$\begin{aligned} dS &= \int_{t_0}^{t+dt} dtL[\eta_2(t'),\dot{\eta}_2(t'),t'] - \int_{t_0}^{t} dtL[\eta_1(t'),\dot{\eta}_1(t'),t'] \,, \\ &= L[\eta_2(t),\dot{\eta}_2(t),t]dt + \int_{t_0}^{t} dt\delta L[\eta_1(t'),\dot{\eta}_1(t'),t'] \,, \end{aligned}$$

上式的第二项可以利用通常的变分法的方法进行计算，结果为

$$\left(\frac{\partial L}{\partial \dot{q}_i}\delta q_i\right)\Big|_{t_0}^{t} + \int_{t_0}^{t} dt\left(\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right)\delta q_i \,,$$

其中的 $q(t)=\eta_1(t)$ 而 $\delta q(t)=\eta_2(t)-\eta_1(t)$。我们现在要求 $\eta_1(t)=q_c(t)$ 满足欧拉 -拉格朗日方程从而上式的第二项为零。现在注意到：

$$\begin{aligned} dq(t) &= \eta_2(t+dt)-\eta_1(t) = \eta_2(t+dt)-\eta_2(t)+\eta_2(t)-\eta_1(t) \,, \\ &\simeq \dot{\eta}_1(t)dt + \delta q(t) \,. \end{aligned} \tag{6.39}$$

或者说 $\delta q(t)=dq(t)-\dot{q}dt$。于是我们得到，

$$dS = p_i dq_i + (L - p_i\dot{q})dt \,,$$

因此我们可以将作用量的微分写为：

$$\boxed{dS = p_i dq_i - Hdt \,.} \tag{6.40}$$

这个微分式表明了作用量作为上端点的时间 t 和广义坐标 q_i 的函数依赖关系。它的等价描述是

$$\frac{\partial S}{\partial q_i} = p_i \,, \quad \frac{\partial S}{\partial t} = -H \,, \tag{6.41}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

其中第二个式子可以表达为，

$$\frac{\partial S}{\partial t} + H = 0 \,. \tag{6.42}$$

这个方程就是著名的哈密顿-雅克比方程。关于它的进一步的讨论与具体运用我们在后面 (第 34节) 会更细致地涉及。

事实上，如果我们将上面的作用量的微分式子积分，我们就得到系统的作用量：

$$S = \int \left(p_i dq_i - H dt\right) \,. \tag{6.43}$$

对于这个作用量运用变分法，假定广义坐标 q_i 和广义动量 p_i 为独立变量，同时哈密顿量 $H(p,q,t)$ 是广义坐标和广义动量的函数，我们就得到哈密顿正则方程 (6.5)。为此我们注意到，

$$\begin{aligned}
\delta S &= \int_{t_1}^{t_2} dt \delta \left[p_i \dot{q}_i - H\right] \\
&= \int_{t_1}^{t_2} dt \left[\dot{q}_i \delta p_i + p_i \delta \dot{q}_i - \frac{\partial H}{\partial p_i}\delta p_i - \frac{\partial H}{\partial q_i}\delta q_i\right] \\
&= \int_{t_1}^{t_2} dt \left[\left(\dot{q}_i - \frac{\partial H}{\partial p_i}\right)\delta p_i - \left(\dot{p}_i + \frac{\partial H}{\partial q_i}\right)\delta q_i\right] = 0 \,.
\end{aligned} \tag{6.44}$$

令被积函数中 δq_i 和 δp_i 的系数为零，我们就获得了正则方程。

¶ 本节中讨论的作用量作为终点的函数可以改换原始的最小作用量原理的表述形式。在这种形式中，我们更关心粒子所经历的路径的形状，而不关心粒子在什么时间走过该路径。考虑初始点为 $(t_0, q^{(0)}(t_0))$，终止点为 $q(t)$，但允许终止时间有个虚变动 δt 的那些能量守恒的轨道，那么：$\delta S = -E\delta t$，其中 E 为体系守恒的能量值。由于一般来说 (即公式 (6.40) 只不过将微分换为变分)：$\delta S = p_i \delta q_i - E\delta t = 0$，于是我们发现这种情况下最小作用量原理可以等价地表述为：

$$\delta S_0 = 0 \,, \tag{6.45}$$

其中 S_0 为系统的简约作用量 (abbreviated action)：

$$S_0 = \int p_i dq_i \,. \tag{6.46}$$

因此，这个利用简约作用量表述的最小作用量原理为：在所有满足能量守恒、固定初始坐标和时间并且 (在某个时刻) 通过终点的运动路径中，真实的运动使得简约作用量取极小值。这种形式的最小作用量原理通常被称为莫佩督原理 (Maupertuis principle)。[10] 当然，要利用这个原理我们必须将简约作用量中的动量利用坐标和坐标的微分表达。这一点可以通过广义动量的定义：

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \,, \tag{6.47}$$

[10] 虽然通常以莫佩督原理来命名它，但是它实际上首先出现在欧拉和拉格朗日的工作中。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

以及能量守恒条件：

$$E(q,\dot{q}) = E\ . \tag{6.48}$$

得到。具体来说，只要从能量守恒条件中将 dt 表达为坐标及其微分，带入到广义动量的定义之中我们就得到了动量作为坐标及其微分的表达式。

以一个通常保守系统为例，如果它的拉格朗日量为：

$$L = \frac{1}{2}a_{ij}(q)\dot{q}_i\dot{q}_j - V(q)\ . \tag{6.49}$$

它的广义动量为：$p_i = a_{ij}(q)\dot{q}_j$。能量为：

$$E = \frac{1}{2}a_{ij}(q)\dot{q}_i\dot{q}_j + V(q)\ . \tag{6.50}$$

因此我们得到：

$$dt = \sqrt{\left[\frac{a_{ij}(q)dq_idq_j}{2(E-V(q))}\right]}\ . \tag{6.51}$$

于是我们可以构造出系统的简约作用量：

$$S_0 = \int\sqrt{[2(E-V(q))a_{ij}(q)dq_idq_j]}\ . \tag{6.52}$$

一个特别有意思的情况是一个简单的粒子的运动。这时的莫佩督原理变为：

$$\delta\int ds\sqrt{2m(E-V)} = 0\ , \tag{6.53}$$

其中 ds 是该粒子在空间轨迹上的一个微分线元。这种形式的最小作用量原理 (莫佩督原理) 又被称为雅克比原理。[11]

• • • • •

¶需要指出的是，在莫佩督原理或者雅克比原理中，时间变量被消去了。因此，它更适合直接得到运动轨道的形状（方程）。这在我们不关心时间依赖，只关心轨道形状的时候是方便的。如果要将时间的依赖也求出来，我们可以运用莫佩督原理或者雅克比原理所确定的轨道方程，再结合公式 (6.51)，这样就可以完全确定力学系统的所有运动信息。

例题 6.4　利用雅克比原理求一个粒子在势中的运动轨道 考虑在势 $V(\mathbf{r})$ 中的一个质点的运动，我们将利用雅克比原理 (6.53) 来求解它的轨道。

解： 利用雅克比原理 (6.53) 直接取变分：

$$\delta\int\sqrt{E-V(\mathbf{r})}ds = \int\sqrt{E-V(\mathbf{r})}\delta ds - \int\frac{\partial V}{\partial\mathbf{r}}\cdot\frac{\delta\mathbf{r}}{2\sqrt{E-V(\mathbf{r})}}ds\ . \tag{6.54}$$

[11] 细心的读者可能已经发现这个形式与几何光学中的费马原理的相似性。这一点我们会在第 34.3节中讨论。

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

由于 $ds^2 = d\mathbf{r}\cdot d\mathbf{r}$，因此我们有：$\delta ds = d\mathbf{r}/ds\cdot\delta d\mathbf{r}$。带入上式并将第二项分部积分，我们就可以令正比于 $\delta\mathbf{r}$ 的系数为零：

$$\frac{d^2\mathbf{r}}{ds^2} = \frac{\mathbf{F} - (\mathbf{F}\cdot\mathbf{t})\mathbf{t}}{2[E - V(\mathbf{r})]} . \tag{6.55}$$

这里 $\mathbf{t} = d\mathbf{r}/ds$ 是质点运动轨道上切向的单位矢量；$\mathbf{F} = -\partial V(\mathbf{r})/\partial\mathbf{r}$ 是粒子所受到的力。我们发现等式右边的分子上的矢量正好是粒子所受到力的法向分量。另外，二级导数 $d^2\mathbf{r}/ds^2 = d\mathbf{t}/ds = \mathbf{n}/R$，其中 $\mathbf{n}$ 是粒子轨道法向的单位矢量，R 是轨道的曲率半径。因此，这个方程实际上就是：$\mathbf{F}\cdot\mathbf{n} = mv^2/R$，这正如我们所预料的一样。只不过它的形式仅仅涉及粒子运动的轨道 (纯几何) 性质，不直接涉及时间。 搞定

例题 6.5　利用莫培督 -雅克比原理求 Kepler 问题中的轨道形状

我们可以利用莫培督原理来求解 Kepler 问题中的轨道的形状问题。

解： 请验证，Kepler 问题中的解恰好满足上题中的方程 (6.55)。为此，我们首先将二维极坐标中一条曲线 $r = r(\theta)$ 的切向量、法向量等表达出来。

搞定

33　正则变换

¶ 我们在本章一开始提到过，哈密顿形式的分析力学与拉格朗日形式的分析力学实际上是等价的。但是哈密顿形式的分析力学体现出一种优势，那就是它的广义坐标和广义动量是被看作独立变量处理的，因此两者完全可以独立选择。事实上，如果一个力学体系的广义坐标和广义动量分别为 q 和 p，我们可以将它们变换为一组新的的广义坐标和广义动量 Q 和 P：

$$Q_i = Q_i(q,p,t) , \quad P_i = P_i(q,p,t) . \tag{6.56}$$

我们试图寻找一类变换，它使得变换后的正则方程仍然不变：

$$\dot{Q}_i = \frac{\partial H'}{\partial P_i} , \quad \dot{P}_i = -\frac{\partial H'}{\partial Q_i} , \tag{6.57}$$

其中 $H'(P,Q,t)$ 是变化后的哈密顿量。要寻找这类变换最为简单的做法是从最小作用量原理出发。我们上一节末尾提到了，利用广义坐标和广义动量来表述，最小作用量原理可以表述为：

$$\delta\int (p_i dq_i - H dt) = 0 . \tag{6.58}$$

我们在上一节看到了，对这个作用量的变分正好给出哈密顿正则方程。因此，如果要求新的广义坐标、广义动量和哈密顿量也满足正则方程，我们一定有：

$$\delta\int \left(P_i dQ_i - H' dt\right) = 0 . \tag{6.59}$$

由 CPHOWIKI 小组整理自互联网　仅供学习交流使用　更多资料请访问 https://pan.cpho.wiki

这两个表述完全等价的条件是被积函数仅仅相差一个任意函数的完全微分：

$$p_i dq_i - H dt = P_i dQ_i - H' dt + dF \ .$$

稍微整理一下我们得到：

$$\boxed{dF = p_i dq_i - P_i dQ_i + (H' - H) dt \ .} \tag{6.60}$$

这种形式的变换被称为 正则变换。相应的函数 F 被称为该正则变换的生成函数，或者母函数。正则变换 (6.60) 可以视为母函数 $F(q, Q, t)$ 的微分展开式。这种类型的正则变换的母函数是看成为旧的坐标 q，新的坐标 Q 和时间 t 的函数，有的书上又称之为第一型 (Type 1) 正则变换。正则变换的母函数总是可以选择表达为新/旧的坐标或者新/旧的动量以及时间的函数，因此按照这种分类我们会有四类正则变换：

- 第一型（Type 1）：(q, Q, t) 为独立变量；
- 第二型（Type 2）：(q, P, t) 为独立变量；
- 第三型（Type 3）：(p, Q, t) 为独立变量；
- 第四型（Type 4）：(p, P, t) 为独立变量；

注意，母函数一定是包含一个旧的变量和一个新的变量。第一型正则变换 (6.60) 可以等价地写为：

$$p_i = \frac{\partial F}{\partial q_i} \ , \quad P_i = -\frac{\partial F}{\partial Q_i} \ , \quad H' = H + \frac{\partial F}{\partial t} \ . \tag{6.61}$$

也就是说，如果生成函数 $F(q, Q, t)$ 的函数形式已知，上式中的第三式就给出了新的的哈密顿量；第二式给出了新的广义动量 P_i；第一式经过反解，可以给出新的广义坐标 Q_i。

有的时候生成函数并不是表达成广义坐标和时间的函数，而是 (比如说) 依赖于旧的广义坐标和新的广义动量。这时的正则变换可以通过 勒让德变换 得到：

$$d(F + P_i Q_i) = p_i dq_i + Q_i dP_i + (H' - H) dt \ . \tag{6.62}$$

这样我们就得到第二型的正则变换：

$$p_i = \frac{\partial \Phi}{\partial q_i} \ , \quad Q_i = \frac{\partial \Phi}{\partial P_i} \ , \quad H' = H + \frac{\partial \Phi}{\partial t} \ , \tag{6.63}$$

其中 $\Phi(q, P, t) = F + Q_i P_i$ 是这种情形下的生成函数。显然，通过勒让德变换我们还可以构造其他形式 (例如用新的广义坐标和旧的广义动量、新旧广义动量等等) 的生成函数。

经过正则变换，力学系统的哈密顿正则方程的形式不变。但是由于在正则变换中原来的广义坐标和广义动量可能相互混合，因此对于变换以后的广义坐标和广义动量而言，它

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

们可能已经完全失去了“坐标”或者“动量”的原始意义 (甚至连量纲都不一定保持)。换句话说，在哈密顿力学体系中，广义坐标和广义动量是完全等价的、相互共轭的一对独立变量。更为科学的称呼应当将 (q_i, p_i) 这一对变量统一称为正则共轭变量。事实上，如果取生成函数 $F = q_i Q_i$，简单的计算表明：$P_i = -q_i$，$Q_i = p_i$。也就是说，这个正则变换刚好将原先的坐标与动量互换。

正则变换的另外一个重要性质是不改变力学量的泊松括号：

$$[f, g]_{p,q} = [f, g]_{P,Q} \,. \tag{6.64}$$

这个结论可以通过直接计算加以验证。

34 哈密顿 -雅克比方程

34.1 一般的讨论

¶ 在第 32节中，我们曾经考虑了一个力学系统的作用量对于终点坐标的依赖。我们得到了一个重要的方程 (参见方程 (6.42)):

$$\boxed{\frac{\partial S}{\partial t} + H\left(q, \frac{\partial S}{\partial q}, t\right) = 0 \,.} \tag{6.65}$$

这个关于函数 $S(q,t)$ 的一阶偏微分方程就是著名的 哈密顿 - 雅克比方程 。

哈密顿 -雅克比方程是一个 (一般是非线性的) 一阶偏微分方程。它的解一般来说可以依赖于任意的函数。在力学上讲，我们感兴趣的是方程的所谓 完全积分 。这类解中包含与独立自由度数目相同的任意常数。对于一个自由度数目为 f 的力学系统，这类解包含 $(f+1)$ 个独立常数。[12] 由于哈密顿－雅克比方程仅仅依赖于 $S(q,t)$ 的偏微商，所以其中一个独立常数显然是一个相加常数。也就是说，我们可以将哈密顿－雅克比方程的完全积分写为：

$$\boxed{S = S(q_1, \cdots, q_f; Q_1, \cdots, Q_f; t) + S_0 \,,} \tag{6.66}$$

其中 $Q_1, \cdots, Q_f$ 是 f 个任意常数，S_0 是一个任意相加常数。

要给另外的常数 $Q_1, \cdots, Q_f$ 一个物理的解释我们注意到哈密顿 -雅克比方程与我们在第 33节中讨论的正则变换的类似之处。将哈密顿－雅克比方程与公式 (6.61) 比较我们发现，如果我们取 Q 为正则变换后新的广义坐标，取作用量 $S(q,Q,t)$ 为正则变换的母函数，那么哈密顿 -雅克比方程告诉我们，这个正则变换以后的哈密顿量正好恒等于零：$H' = 0$。对于变换以后的系统，由于其哈密顿量恒等于零，它的运动方程的解是平庸的。

[12] 其中的 $+1$ 来自于时间零点的选取。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

也就是说，新的广义坐标 Q 和新的广义动量 P 都是常数。由此我们看到，我们完全可以将函数 $S(q_1,\cdots,q_f;Q_1,\cdots,Q_f;t)$ 中的任意常数 $Q_1,\cdots,Q_f$ 解释为以 S 为生成函数的第一型正则变换后的新的广义坐标。[13] 相应的新的广义动量可以由作用量对于新的广义坐标的偏微商得到：

$$P_i=-\frac{\partial S}{\partial Q_i}\,, \tag{6.67}$$

这正是公式 (6.61) 所指出的。[14] 因此，利用哈密顿 -雅克比方程求解力学问题的过程可以分为以下几个步骤：

1. 根据系统的哈密顿量写下哈密顿－雅克比方程 (6.65)。
2. 求出方程形如公式 (6.66) 的所有完全积分，其中包含一系列任意常数 $Q_1,\cdots,Q_f$。
3. 以 $Q_1,\cdots,Q_f$ 为新的广义坐标，由公式 (6.67) 定义新的广义动量。新的广义坐标 $Q_1,\cdots,Q_f$ 和新的广义动量 $P_1,\cdots,P_f$ 都是不随时间变化的常数。
4. 从公式 (6.67) 可以反解出系统变换前的广义坐标 q_i 作为 $2f$ 个任意常数 (即新的广义坐标 $Q_1,\cdots,Q_f$ 和新的广义动量 $P_1,\cdots,P_f$) 和时间 t 的函数。
5. 系统原先的广义动量则可以由 $p_i=\partial S/\partial q_i$ 得到。至此我们就得到了一个力学系统运动的完全解。

由此我们看到：利用哈密顿 -雅克比方程求解力学问题的关键一步 (往往也是最困难的一步) 是求出形如 (6.66) 的完全积分，即上述步骤中的第 2 步。

¶如果系统的哈密顿量不显含时间那么哈密顿 -雅克比方程的形式可以稍微简化。这时，系统的哈密顿量 $H(q,p)$ 是一个守恒量 $H(q,p)=E$，其中 E 是系统的能量。这时，作用量对于时间的依赖可以完全分离出来：$S(q,t)=S'(q)-Et$。这样一来，我们就得到 (下面我们略去 $S'(q)$ 中的一撇)：

$$\boxed{H\left(q_1,\cdots,q_f;\frac{\partial S}{\partial q_1},\cdots,\frac{\partial S}{\partial q_f}\right)=E\,.} \tag{6.68}$$

这就是一个保守系统的哈密顿 -雅克比方程。

34.2 分离变量法

¶哈密顿 -雅克比方程是一个一阶非线性偏微分方程。它的一般解是十分复杂的。但是，对于一些重要的物理情况，哈密顿 -雅克比方程的解可以由 分离变量法 给出。本节中我们将结合这些实例，简单介绍一下这种方法。

[13] 当然，如果愿意，也可以将常数 Q_i 视为正则变换后的广义动量。这时我们也许应当用 P_i 来标记它们，那么哈密顿 -雅克比方程也可以视为第二型的正则变换的结果。

[14] 事实上，由于 Q_i 和 P_i 都是常数，因此，很多时候我们也将公式 (6.67) 中等式右边的负号略去。这时 P_i 的物理含义是变换后的正则动量的负值。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

一类最为常见的情形是某一个广义坐标和相应的广义动量只出现在一个与其他广义坐标和广义动量分离的组合之中。也就是说，哈密顿 -雅克比方程可以写为：

$$\Phi\left[\phi\left(q_1,\frac{\partial S}{\partial q_1}\right);q_i,\frac{\partial S}{\partial q_i};t,\frac{\partial S}{\partial t}\right]=0\,. \tag{6.69}$$

这里 $\phi(q_1,\partial S/\partial q_1)$ 是任意的已知函数；q_i 和 $\partial S/\partial q_i$ 则代表其他的 (也就是 $i\neq 1$ 的) 广义坐标和相应的偏微商 (广义动量)。这个时候我们可以寻求如下形式的解：

$$S=S'(q_i,t)+S_1(q_1)\,. \tag{6.70}$$

这时哈密顿 -雅克比方程可以分离成下列两个方程：

$$\phi\left(q_1,\frac{dS_1}{dq_1}\right) = \alpha_1\,. \tag{6.71}$$

$$\Phi\left[\alpha_1;q_i,\frac{\partial S'}{\partial q_i};t,\frac{\partial S'}{\partial t}\right] = 0\,. \tag{6.72}$$

其中的 α_1 为一个常数。这个方程组的第一个方程是一个一阶常微分方程。我们可以求解它得到 $S_1(q_1)$。剩下的关于 S' 的哈密顿 -雅克比方程是减少了一个自由度的方程。一个最为常见的例子是如果某个广义坐标根本不出现在哈密顿量之中。这样的坐标变量被称为循环坐标。这时，$\phi(q_1,\frac{dS_1}{dq_1})=dS_1/dq_1$，于是我们很容易积分得到：$S_1(q_1)=\alpha_1 q_1$，从而：

$$S=S'(q_i,t)+\alpha_1 q_1\,. \tag{6.73}$$

注意，这个论证其实同样适用于时间变量。如果哈密顿量不显含时间，也就是说时间变量是一个循环变量，这时哈密顿 -雅克比方程的解的形式就是：$S=S'-Et$，这正是公式 (6.68)。因此，哈密顿量如果具有对称性，同时这种对称性如果意味着循环坐标变量的存在，那么我们就可以将哈密顿－雅克比方程分离变量，从而将循环坐标的对称性考虑在内。显然，广义坐标的选取有时候是十分重要的。正确的选取有可能直接简化循环坐标的辨认，从而大大简化一个力学问题的求解。作为一个例子，我们这类将利用哈密顿－雅克比方程求解一个 $1/r$ 中心势中的粒子的运动 (开普勒问题)。

例题 6.6　哈密顿 -雅克比方法求解开普勒问题 考虑在 $1/r$ 中心势中一个非相对论性的质点的运动，我们将利用哈密顿 -雅克比方法来求解这个问题。

解： 显然，这是一个二维问题。我们选取极坐标来讨论这个问题。这时两个独立的广义坐标为 (r,θ)，其中 r 是质点到力心的距离，θ 是相对于某个固定方向的极角。系统的哈密顿量可以写为：

$$H=\frac{1}{2m}\left(p_r^2+\frac{p_\theta^2}{r^2}\right)+V(r)\,, \tag{6.74}$$

其中对于开普勒问题我们有：

$$V(r)=\frac{\alpha}{r}\,. \tag{6.75}$$

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

由于哈密顿量不显含时间，因此我们可以直接利用不含时的哈密顿 - 雅克比方程 (6.68):

$$\frac{1}{2m}\left[\left(\frac{\partial W}{\partial r}\right)^2+\frac{1}{r^2}\left(\frac{\partial W}{\partial \theta}\right)^2\right]+V(r)=E\,. \tag{6.76}$$

由于势能 $V(r)$ 中不含有坐标 θ(循环坐标)，因此，函数 W 可以进一步分解为:

$$W(r,\theta)=W_1(r)+W_2(\theta)\,. \tag{6.77}$$

将此式带入哈密顿 -雅克比方程得到:

$$r^2\left(\frac{\partial W_1}{\partial r}\right)^2+2mr^2[V(r)-E]=-\left(\frac{\partial W_2}{\partial \theta}\right)^2\,.$$

这个等式的左边只是 r 的函数而右边只是 θ 的函数。要使它能够成立必须两边都等于一个常数。因此:

$$\begin{aligned}\frac{dW_2}{d\theta} &= J\,,\\ \frac{dW_1}{dr} &= \sqrt{2m\left[E-V(r)\right]-\frac{J^2}{r^2}}\,.\end{aligned} \tag{6.78}$$

积分后给出:

$$\begin{aligned}W_2(\theta) &= J\theta\,,\\ W_1(r) &= \int^r dr'\sqrt{2m\left[E-V(r')\right]-\frac{J^2}{r'^2}}\,.\end{aligned} \tag{6.79}$$

这个式子的第一行告诉我们: $p_\theta=J$ 是一个常数 (角动量)。对于 $V(r)=\alpha/r$，第二个式子原则上可以积分给出 $W_1(r)$。因此，我们得到哈密顿 -雅克比方程的完全解:

$$S=J\theta+\int^r dr'\sqrt{2m\left[E-V(r')\right]-\frac{J^2}{r'^2}}-Et\,. \tag{6.80}$$

这里面两个常数 (因为自由度数目是 2)$Q_1=E$，$Q_2=J$。于是，我们可以利用公式 (6.67) 来定义新的“动量”，它们也都是常数。[15] 例如，对于 $Q_2=J$ 的偏微商我们得到:

$$\theta_0=\theta-\int^r dr'\frac{(J/r'^2)}{\sqrt{2m\left[E-V(r')\right]-\frac{J^2}{r'^2}}}\,. \tag{6.81}$$

大家可以将这里利用哈密顿 -雅克比方程得到的结果与我们在第 14节中的结果 (3.9) 比较一下，就可以发现它们是完全一致的。对于开普勒问题，带入势能 $V(r)=\alpha/r$，这个积分表达式可以直接积出得到圆锥曲线的轨道方程。 搞定

[15] 事实上，这些新的所谓“动量”并不具有动量的量纲。因此，在正则变换的体系下，所谓坐标和动量实际上已经很难区分了。它们构成一对正则变量。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

例题 6.7 用哈密顿 -雅克比方法求解均匀磁场中一个非相对论性的带电粒子的运动 考虑在沿 z 轴有一个均匀磁场 $\mathbf{B}=B\hat{z}$。考虑一个非相对论性的带电粒子 (电荷为 e，质量为 m) 在该均匀磁场中的运动。我们将利用哈密顿 -雅克比方法来求解这个问题。

解： 首先我们将选取 $c=1$ 的单位制。这样一来，该粒子的哈密顿量为:

$$H=\frac{1}{2m}(\mathbf{p}-e\mathbf{A})^2\,, \tag{6.82}$$

其中 $\mathbf{p}$ 是粒子的正则动量，[16] $\mathbf{A}$ 是与均匀磁场对应的磁矢势。它的选择依赖于规范。一个常用的规范选取是:

$$A_x=0\,,\ A_y=Bx\,,\ A_z=0\,.$$

大家很容易验证: $\mathbf{B}=\nabla\times\mathbf{A}=B\hat{z}$。于是上面的哈密顿量可以明显地写为

$$H=\frac{p_x^2}{2m}+\frac{1}{2m}(p_y+bx)^2+\frac{p_z^2}{2m}\,.$$

其中参数 $b\equiv -eB$。显然，哈密顿量不显含 t，因此粒子的总能量 $E\equiv E_\parallel+E_\perp$ 是守恒的。同时，粒子沿 z 轴方向的动量 $p_z=Q_3$(因此相应的动能 $E_\parallel=Q_3^2/(2m)$) 也是守恒的。我们可以直接求解不含时的哈密顿 -雅各比方程:

$$\begin{cases}\dfrac{1}{2m}\left(\dfrac{\partial S}{\partial x}\right)^2+\dfrac{1}{2m}\left(\dfrac{\partial S}{\partial x}+bx\right)^2+E_\parallel=E\,,\\[2ex] \qquad\qquad\qquad\dfrac{1}{2m}\left(\dfrac{\partial S}{\partial z}\right)^2=E_\parallel\end{cases} \tag{6.83}$$

由于哈密顿量中不出现 y，因此与其共轭的动量 p_y 也守恒。因此我们总体上分离变量为: $S=X(x)+Y(y)+Z(z)-Et$，我们有

$$X(x)=\int\sqrt{2mQ_1-(Q_2+bx)^2}dx\,,\ Y(y)=Q_2y,\ Z(z)=Q_3z\,.$$

而总的能量 $E=E_\parallel+E_\perp$，$E_\perp=Q_1$，$E_\parallel=Q_3^2/(2m)$。这些积分可以暂时不急于积出，因为下面很快就要计算其导数。

按照哈密顿 -雅克比方程的解法，我们令 $P_i=-\partial S/\partial Q_i$ 并且可以积出相应的积分后得到:

$$\begin{cases}P_1=-\dfrac{\partial S}{\partial Q_1}=t-\dfrac{m}{b}\sin^{-1}\left(\dfrac{Q_2+bx}{\sqrt{2mQ_1}}\right)\\[2ex] P_2=-\dfrac{\partial S}{\partial Q_2}=-y-\dfrac{1}{b}\sqrt{2mQ_1-(Q_2+bx)^2}\\[2ex] P_3=-\dfrac{\partial S}{\partial Q_3}=-z+\dfrac{Q_3}{m}t\,.\end{cases} \tag{6.84}$$

[16] 读者请注意这里与我们前面讨论中一些记号上的变化。在例题 6.1中我们用 **P** 来代表粒子的正则动量而用 **p** 表示它的机械动量。这里由于我们一般用 **P** 来标记正则变换后的新的动量，因此我们将用 **p** 来标记粒子的正则动量。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

这些式子可以等价地化为下列式子:

$$\begin{cases} x(t) = x_0 + R_0 \sin\left[\omega_0(t-t_0)\right] \,, \\ (y-y_0)^2 + (x-x_0)^2 = R_0^2 \,, \\ z = z_0 + (Q_3/m)(t-t_0) \,, \end{cases} \tag{6.85}$$

其中的常数为: $x_0 = -Q_2/b$, $R_0 = \sqrt{2mQ_1}/b, t_0 = P_1$, $\omega_0 = b/m$, $z_0 + P_3 = Q_3 t_0/m$。上式中的 z 分量的运动 (又称为纵向运动) 就是匀速直线运动; 垂直于磁场的横向运动是匀速圆周运动: (x_0, y_0) 为匀速圆周运动的圆心, R_0 为半径, 运动的圆频率称为回旋频率, 它直接与带电粒子的荷质比以及外磁场的磁感应强度大小成正比: $\omega_B = \omega_0 = b/m = -eB/m$。注意其正负号其实只影响带电粒子转向 (顺时针、逆时针), 真正的圆频率是其绝对值。

带电粒子在均匀磁场中的运动对于经典物理和量子物理来说都是十分重要的案例。经典物理中它是粒子加速器上探测以及调制带电粒子束流的最主要手段; 量子物理中它是研究磁场中固体磁性质 (特别是所谓的量子霍尔效应) 的重要手段。 搞定

例题 6.8 同样的问题, 柱坐标系中求解 我们将在柱坐标系中利用哈密顿 -雅克比方法来求解同样的问题。

解: 在柱坐标中哈密顿量可以表达为:

$$H = \frac{1}{2m}\left[(p_r - eA_r)^2 + \left(\frac{p_\theta}{r} - eA_\theta\right)^2 + (p_z - eA_z)^2\right] \,. \tag{6.86}$$

注意其中的第二项的形式。$\mathbf{A}$ 的形式仍然依赖于规范的选取。这时一个比较方便的选择是:

$$A_r = 0 \,, \quad A_\theta = \frac{1}{2}Br \,, \quad A_z = 0 \,. \tag{6.87}$$

大家可以验证, 这个选取给出沿 z 方向的均匀磁场 (check this please!): [17]

$$\mathbf{B} = \nabla \times \mathbf{A} = \frac{1}{r}\begin{vmatrix} \hat{e}_r & r\hat{e}_\theta & \hat{e}_z \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial z} \\ A_r & rA_\theta & A_z \end{vmatrix} = B\hat{e}_z \,. \tag{6.88}$$

因此我们可以直接写出哈密顿 -雅克比方程:

$$\frac{\partial S}{\partial t} + \frac{1}{2m}\left[\left(\frac{\partial S}{\partial r}\right)^2 + \left(\frac{1}{r}\frac{\partial S}{\partial \theta} - \frac{e}{2}Br\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2\right] = 0 \,.$$

[17] 一般来说, 一个三维正交曲线坐标系中如果矢量的微分 $d\mathbf{r} = h_i(q)dq_i\mathbf{e}_i$, 其中 $\mathbf{e}_i$ 是三个正交的单位矢量, $h_i(q)$ 一般可以依赖于各个 q_i。这时的矢量场 $\mathbf{A}(\mathbf{r})$ 的旋度可以写为: $\nabla\times\mathbf{A} = (1/h_1h_2h_3)\epsilon_{ijk}(h_i\mathbf{e}_i)\partial_j(h_kA_k)$。具体到柱坐标, 如果选 r, θ, z 分别对应于 1,2,3 的方向, 我们发现 $h_1 = h_3 = 1$, $h_2 = r$, 于是矢量 $\mathbf{A}$ 的旋度可以写为公式 (6.88) 的形式。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

进一步我们可以写出其完全解: $S = W(r) + Q_2\theta + Q_3 z - Et$。如果我们令 $Q_1 = E - Q_3^2/(2m)$ 为横向的能量，那么函数 $W(r)$ 的解为

$$W(r) = \int dr \sqrt{2mQ_1 - \left(\frac{Q_2}{r} + \frac{1}{2}br\right)^2} .$$

这个积分实际上是初等积分，不过我们这里就不去进一步完成它了。建议大家自行完成并与上一个例题的结果进行一下比较。一般来说，一个问题可以在直角坐标又可以在柱坐标中分离变量求解往往意味着它具有多余的对称性。这个问题也不例外。 搞定

例题 6.9 共焦椭圆坐标中的双力心问题 前面我们看到同样的问题在不同坐标系中求解。下面我们来看一个为求解二维中两个力心的引力问题而特制的坐标系 -共焦椭圆坐标系。

解： 所谓共焦椭圆坐标系 (confocal elliptic coordinates) 是指一对坐标 (ξ, η)，它们与通常的笛卡尔坐标 (x, y) 的关系为:

$$\begin{cases} x = c\cosh\xi\cos\eta \\ y = c\sinh\xi\sin\eta \end{cases} \tag{6.89}$$

其中 $\xi \in (-\infty, +\infty)$，$\eta \in [0, 2\pi)$ 称为 共焦椭圆坐标。这个名称的由来是因为如果我们考察 ξ 为常数的 (x, y) 平面的曲线，我们发现它们是:

$$\frac{x^2}{A^2} + \frac{y^2}{B^2} = c^2 ,$$

其中 $A = \cosh\xi$，$B = \sinh\xi$。这是 (x, y) 平面的一系列椭圆；而且由于 $A^2 - B^2 \equiv 1$，因此它代表了一系列拥有共同焦点的椭圆，它们的共同焦点都位于 $(\pm c, 0)$。所以对于固定的 ξ，不同的 η 描写了一系列具有共同焦点的椭圆。类似地，如果我们固定 η，我们发现它代表了 (x, y) 平面的具有共同焦点 (焦点仍然是位于 $(\pm c, 0)$!) 的一系列双曲线。因此这样的坐标系被称为共焦椭圆坐标。

这个坐标系可以用于处理具有两个力心的二维引力问题。显然，我们将两个无穷重的力心选择在两个焦点上。这时在两个力心的联合引力场中运动的一个质量为 m 的粒子的哈密顿量为:

$$H = \frac{\mathbf{p}^2}{2m} - \frac{\alpha_1}{r_1} - \frac{\alpha_2}{r_2} .$$

其中 r_1，r_2 是粒子到两个力心 (即两个焦点) 的距离。之所以运用这个坐标系是因为上式中的势能部分可以比较简单地用 (ξ, η) 来表达:

$$V(r_1, r_2) = -\frac{\alpha_1}{r_1} - \frac{\alpha_2}{r_2} = -\frac{1}{c}\frac{\alpha\cosh\xi - \alpha'\cos\eta}{\cosh^2\xi - \cos^2\eta} , \tag{6.90}$$

其中 $\alpha = \alpha_1 + \alpha_2$，$\alpha' = \alpha_1 - \alpha_2$。同样的，粒子的动能也可以用 (ξ, η) 表达出来:

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) = \frac{m}{2}c^2(\cosh^2\xi - \cos^2\eta)(\dot{\xi}^2 + \dot{\eta}^2) . \tag{6.91}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

我们可以进一步写出粒子的哈密顿量:

$$H = \left(\frac{1}{2c^2 m}\right) \frac{p_\xi^2 + p_\eta^2 - \alpha \cosh\xi + \alpha' \cos\eta}{\cosh^2\xi - \cos^2\eta} \,. \tag{6.92}$$

剩下的就是写出哈密顿 -雅克比方程并且利用 $S = W_\xi(\xi) + W_\eta(\eta) - Et$ 的形式来求出完全解。我们得到:

$$\begin{cases} (W_\xi')^2 = Q + 2Emc^2 \cosh^2\xi + \alpha \cosh\xi \,, \\ (W_\eta')^2 = -Q + 2Emc^2 \cos^2\eta - \alpha' \cos\eta \,, \end{cases} \tag{6.93}$$

其中 Q 是分离变量时出现的常数。这两个方程可以分别积分 (虽然会涉及到椭圆积分) 给出最后的解。 搞定

34.3 哈密顿 -雅克比方程与波动力学

¶ 这一小节中，[18] 我们来讨论哈密顿 -雅克比方程与量子力学的波动力学描述以及与光学的联系。这种联系初想起来是十分优美而出人意料的，但是实际上恰恰是这种形式上的联系，在量子论的初期催生了量子力学。至于说与光学的联系，早在哈密顿原始的工作中就已经有相当的讨论了。[19] 事实上，从惠更斯的波动光学到几何光学的过渡与从量子力学到经典力学的过渡是完全类似的。

我们考虑一个粒子在三维空间运动，它感受到的势能是 $V(\mathbf{r},t)$。系统的作用量作为坐标和时间的函数：$S(\mathbf{r},t)$。我们构造一个 波函数 :

$$\psi(\mathbf{r},t) = \sqrt{\rho(\mathbf{r})} e^{iS(\mathbf{r},t)/\hbar} \,. \tag{6.94}$$

我们将假定：波函数的振幅 $\rho(\mathbf{r})$ 几乎是一个常数；在波函数的相因子中，$\hbar$ 是一个具有作用量量纲的常数。我们将假定对于经典的力学系统，$S(\mathbf{r},t) \gg \hbar$。这样一来，波函数的相因子是一个随位置和时间快速振荡的函数而它的模则是随位置缓变的函数。现在我们 要求 波函数 $\psi(\mathbf{r},t)$ 满足著名的 薛定谔方程 :

$$i\hbar \frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi + V(\mathbf{r},t)\psi \,. \tag{6.95}$$

将 (6.94) 代入上式，简单的计算表明函数 S 所满足的方程为:

$$\left[\frac{(\nabla S)^2}{2m} + V\right] + \frac{\partial S}{\partial t} = \frac{i\hbar}{2m}\nabla^2 S \,. \tag{6.96}$$

如果我们假定:

$$\hbar\nabla^2 S \ll (\nabla S)^2 \,, \tag{6.97}$$

[18] 可以参考 [3] 的第 10 – 8 节的讨论。

[19] 大家也许还记得，哈密顿本人最为热衷的不是力学，而是光学。它是在对于光学的观测和研究中发现哈密顿力学的。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

那么我们可以略去公式 (6.96) 的右边从而得到 S 所满足的方程，它正好具有哈密顿 -雅克比方程的形式。粗略地说，如果我们令 $\hbar \to 0$，那么薛定谔方程就变成经典的哈密顿 -雅克比方程。事实上，哈密顿 -雅克比方程可以看成是 $\hbar \to 0$ 极限下薛定谔方程的零阶近似。如果我们将作用量 S 按照 $\hbar$ 的幂次展开并且比较同阶项，我们就得到了量子力学中所谓的 准经典近似，又被称为 WKB 近似。这个近似足够好的条件 (6.97) 在物理上实际上是说粒子的动量在其德布洛意波长的范围内的改变相对于其动量本身是一个无穷小量。由于粒子的动量在一个范围内的改变依赖于势能函数，因此这个条件很多时候可以等价地描述为势能函数在粒子的德布洛意波长内变化不大。

上面所说的从量子力学到经典力学的过渡体现了一种从波动的观点 (量子力学) 向粒子的观点 (经典力学) 的过渡。这种过渡实际上我们在纯粹经典物理的范畴之内也遇到过，那就是从 波动光学 到 几何光学 的过渡。

在惠更斯的波动光学的描述中，波函数 $\phi(\mathbf{r}, t)$ 满足波动方程：[20]

$$\nabla^2\phi - \frac{n^2}{c^2}\frac{\partial^2\phi}{\partial t^2} = 0\,, \tag{6.98}$$

其中 $n(\mathbf{r})$ 是光学中的 折射率，它可能是空间位置的函数。如果折射率是不依赖于位置的常数，这个方程具有简单的平面波解：

$$\phi = \phi_0 e^{i\mathbf{k}\cdot\mathbf{r}-\omega t}\,, \tag{6.99}$$

其中波数 $k = |\mathbf{k}|$ 由折射率和频率给出：[21]

$$k = \frac{n\omega}{c}\,. \tag{6.100}$$

这称为相应介质的色散关系。如果折射率存在空间依赖，那么平面波并不是波动方程的解。现在我们着重考虑折射率随空间是缓慢变换的情形。所谓缓慢是指折射率在一个光波波长的范围内几乎没有显著的变化。这时，我们可以寻求波动方程下列形式的解：

$$\phi = e^{A(\mathbf{r})}e^{ik_0(L(\mathbf{r})-ct)}\,, \tag{6.101}$$

其中 $k_0 = \omega/c$ 是真空中的波数，$A(\mathbf{r})$ 和 $L(\mathbf{r})$ 都是位置的实函数。$L(\mathbf{r})$ 在光学中被称为 光程，有时又称为 光程函 或 程函。将公式 (6.101) 代入波动方程分别令实部虚部等于零，我们得到：

$$\nabla^2 A + (\nabla A)^2 + k_0^2[n^2 - (\nabla L)^2] = 0\,, \tag{6.102}$$

$$\nabla^2 L + 2\nabla A\cdot\nabla L = 0\,. \tag{6.103}$$

[20] 事实上，要完全描述一个光波 (电磁波)，我们需要一个矢量波函数。惠更斯的波动理论只是该理论 (实际上就是麦克斯韦电磁理论) 一个近似的、标量描述而已。

[21] 原则上讲，折射率 n 会依赖于波的频率 (色散)，不过这种复杂性并不影响我们这里的讨论。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

现在我们假定介质的变化在一个光波波长范围内是可以忽略的。也就是说我们取短波近似，这个近似在光学上称为 几何光学近似。在这个近似下，上述两个公式的第一个式子中含有 k_0^2 的项是主要的，因此我们得到光程函数 $L(\mathbf{r})$ 所满足的方程：

$$(\nabla L(\mathbf{r}))^2 = n^2(\mathbf{r}) \,. \tag{6.104}$$

另一方面，我们知道一个势场 $V(\mathbf{r})$ 中单粒子的哈密顿－雅克比方程是

$$(\nabla W)^2 = 2m(E - V(\mathbf{r})) \,. \tag{6.105}$$

显然，方程 (6.104) 与方程 (6.105) 的形式是高度相似的。光学中的程函 $L(\mathbf{r})$ 对应于单粒子哈密顿 -雅克比方程的作用量函数 $W(\mathbf{r})$；光学中的折射率 $n(\mathbf{r})$ 则对应于哈密顿 -雅克比方程中的速度函数：$\sqrt{2m(E - V(\mathbf{r}))}$。特别应当指出的是，几何光学中的 费马原理：

$$\delta \int n ds = 0 \,, \tag{6.106}$$

恰恰对应于雅克比形式的最小作用量原理 (参见第 32节中的公式 (6.53))：

$$\delta \int \sqrt{2m(E - V(\mathbf{r}))} ds = 0 \,. \tag{6.107}$$

由此我们看到：波动光学在几何光学极限下可以给出费马原理，而这样一个事实实际上与哈密顿 -雅克比形式的最小作用量原理是等价的。我们现在也许可以理解为什么当年著名的哈密顿同志发表的论文有一个奇怪的名称：”Theory of Systems of Rays”，原来在哈密顿看来，这一切都是统一的。[22]

35 绝热不变量与正则变量

35.1 绝热不变量

¶ 首先考虑一个处在外场中的一维周期运动的粒子。我们假定粒子只在有限的范围之内运动 (周期的)，同时粒子的运动受外界的影响。这种影响体现在某个参数 λ 之中。我们假定外参数 $\lambda(t)$ 仅仅随时间缓慢地变化：

$$T\frac{d\lambda(t)}{dt} \ll \lambda(t) \,, \tag{6.108}$$

其中 T 是该粒子运动的周期。满足这种条件的缓变参数常常被称为 绝热变化参数。

当参数 λ 随时间不改变的时候，力学系统的哈密顿量不显含时间从而它的能量 E 守恒。如果现在参数 λ 随时间缓慢地变化，那么系统的能量一般也会随着时间缓慢变化。我

[22] 该文成文于 1827 年，发表于 1828 年的 Transactions of the Royal Irish Academy。该文不仅仅统一了经典的力学和几何光学，而且在当时客观上帮助了波动光学的建立。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

们可以将其能量的变化率 $\dot{E}(t)$ 在系统的一个周期 T 中平均，这样得到的结果：$\langle \dot{E} \rangle$ 将与 $\dot{\lambda}$ 成正比。事实上，我们有：

$$\frac{dE}{dt} = \frac{\partial H}{\partial t} = \frac{\partial H}{\partial \lambda}\frac{d\lambda}{dt} \,. \tag{6.109}$$

这里系统的哈密顿量 $H(q,p;\lambda)$ 依赖于绝热变化参数 λ。在一个周期之中平均我们得到：

$$\left\langle \frac{dE}{dt} \right\rangle = \left\langle \frac{\partial H}{\partial \lambda} \right\rangle \frac{d\lambda}{dt} \,. \tag{6.110}$$

这里我们假定 $\dot{\lambda}$ 也是缓变的函数，也就是说在一个周期之中它几乎不变。同样，这个等式右边的平均过程中，λ 也可以看成是常数，因为它在一个周期之中几乎不变。现在注意到：

$$\left\langle \frac{\partial H}{\partial \lambda} \right\rangle = \frac{1}{T}\int_0^T \frac{\partial H}{\partial \lambda} dt \,,$$

同时，我们利用哈密顿正则方程可以将上式中对时间的积分换成对于坐标的积分：

$$\left\langle \frac{\partial H}{\partial \lambda} \right\rangle = \frac{\oint \frac{\partial H}{\partial \lambda}\frac{dq}{\partial H/\partial p}}{\oint \frac{dq}{\partial H/\partial p}} \,. \tag{6.111}$$

这个公式中对于坐标的积分遍及它的一个闭合周期。由于在一个周期中我们可以将能量 E 看成是常数，因此我们有：

$$\frac{\partial H/\partial \lambda}{\partial H/\partial p} = -\frac{\partial p}{\partial \lambda} \,.$$

这里动量 p 可以在一个周期中可以看成是坐标 q、能量 E 和参数 λ 的函数，即：$p = p(q;E,\lambda)$。由此我们得到：

$$\left\langle \frac{dE}{dt} \right\rangle = -\frac{d\lambda}{dt}\frac{\oint(\partial p/\partial \lambda)dq}{\oint(\partial p/\partial E)dq} \,. \tag{6.112}$$

或者等价地写为：

$$\boxed{\left\langle \frac{dI}{dt} \right\rangle = 0 \,, \quad I \equiv \frac{1}{2\pi}\oint p dq \,.} \tag{6.113}$$

也就是说，如果参数 λ 仅仅发生绝热的缓慢变化，那么物理量 I 在一个周期内的平均值不随时间变化。具有这样性质的物理量被称为绝热不变量。这个结果告诉我们，公式 (6.113) 中定义的量 I 是体系的一个绝热不变量。它有一个十分简单的几何解释。对于一个一维系统，如果系统的运动是周期的，那么它的轨道在其相空间中是一条闭合的曲线。公式 (6.113) 所定义的作用量变量 I 恰好正比于其该闭合曲线所包围的面积。因此，我们的结果也可以表达为：一个一维周期运动的力学体系相空间中闭合相轨道所包围的面积在系统的参数随时间发生缓慢变化时是不变的。

例题 6.10　一维谐振子的作用量变量 考虑一个一维简谐振子的作用量变量 I。

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

解：直接利用作用量变量 I 正比于其闭合相轨道所包围的面积的事实，同时记住谐振子的相轨道是一个椭圆，我们可以得到：

$$I = E/\omega \,, \tag{6.114}$$

其中 ω 是谐振子的圆频率。 搞定

上面讨论的绝热不变量 I 与系统的简约作用量相当，因此它被称为 作用量变量 。它实际上是系统能量的函数。作用量变量 I 对于能量的偏微商正好正比于系统的周期：

$$2\pi\frac{\partial I}{\partial E} = \oint \frac{\partial p}{\partial E} dq = T \,. \tag{6.115}$$

或者用圆频率写成：

$$\boxed{\frac{\partial E}{\partial I} = \omega \,.} \tag{6.116}$$

显然，这个等式对于一维谐振子 (参见公式 (6.114)) 是平庸地成立的。

35.2 正则变量

¶现在我们假定 λ 不随时间改变。这样一来，系统的能量也是常数。这时，系统的简约作用量：

$$S_0(q, E; \lambda) = \int p(q, E; \lambda) dq \,, \tag{6.117}$$

实际上也可以看成是 q、I、λ 的函数，因为这时能量 E 只是作用量变量 I 的函数。现在我们以 $S_0(q, I; \lambda)$ 为生成函数，以作用量变量 I 为新的动量进行一个正则变换。为此，我们可以参考第 33节中的公式 (6.63)。相应的变换后的广义坐标记为 w，我们有：

$$p = \frac{\partial S_0}{\partial q} \,, \quad w = \frac{\partial S_0}{\partial I} \,. \tag{6.118}$$

这个公式的第一个式子等价于公式 (6.117)，第二个式子则确定了变化后的新的坐标 w。坐标 w 被称为 角度变量 。它与作用量变量 I 一起构成一对正则共轭变量，它们被称为 正则变量 。

由于原来的哈密顿量以及正则变换的生成函数都不显含时间，因此正则变换后的哈密顿量与变换前的哈密顿量在数值上是相等的，只不过它应当表达为新的正则变量 (w, I) 的函数。我们前面刚刚论证了，对于一个不显含时间的系统，它的能量仅仅依赖于作用量变量，因此利用新的正则变量表达，哈密顿运动方程为：

$$\dot{I} = 0 \,, \quad \dot{w} = \frac{dE(I)}{dI} \equiv \omega(I) \,. \tag{6.119}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

这组方程中的第一个就是绝热不变性的特例；第二个方程说明角度变量随时间的变化率也是常数。所以我们有：

$$w(t) = \omega(I)t + \text{const.} . \tag{6.120}$$

也就是说，角度变量随着时间线性增加。另外一个结论是 $S_0(q, I)$ 实际上是坐标 q 的多值函数，当坐标走遍一个周期时，函数 $S_0(q, I)$ 会改变：

$$\Delta S_0(q, I) = 2\pi I , \tag{6.121}$$

或者等价地说，在粒子走遍一个周期的时候，它的角度变量改变为：

$$\Delta w = 2\pi . \tag{6.122}$$

这就是为什么我们称 w 为角度变量。这个结论可以推广到任意坐标和动量的单值函数 $F(q, p)$。如果我们用作用量 -角度变量来表达，这个函数实际上是角度变量 w 的周期函数，其周期为 2π。

35.3 完全可分离系统的正则变量

¶ 前面关于绝热不变量以及作用量 -角度正则变量的讨论都限于一维系统。现在我们试图将这种讨论推广到具有多个自由度的系统。对于具有多自由度的力学系统，我们仅仅讨论所谓完全可分离的系统，有时又被称为完全可积的力学系统。对一般力学系统的讨论我们就不再涉及了。

考虑一个 f 个自由度的保守力学系统。我们同时假定系统的每一个广义坐标 q_i 都限制在有限的区域之内。我们进一步假定这个系统是完全可分离的。我们这里所说的完全可分离是在哈密顿 -雅克比意义下的。也就是说，如果一个多自由度系统的哈密顿 -雅克比方程可以完全分离变量，我们有：

$$S_0 = \sum_i S_i(q_i) , \tag{6.123}$$

而对于每一个 i 我们有：[23]

$$S_i = \int p_i dq_i = 2\pi I_i , \tag{6.124}$$

所有的作用量变量 I_i 都是绝热不变量。当哈密顿量不随时间改变的时候，它们都是常数。因此，我们可以以 I_i 为新的动量，与它们共轭的角度变量为：

$$w_i = \frac{\partial S_0}{\partial I_i} . \tag{6.125}$$

用新的作用量－角度变量来描写哈密顿方程的解为：

$$I_i = \text{constant}, \quad w_i = \frac{\partial E(I)}{\partial I_i} t + \text{constant} . \tag{6.126}$$

[23] 注意，这里重复的指标 i 不求和。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

当力学系统的某个坐标 q_i 走遍一个周期后，相应的角度变量 w_i 改变 2π，而系统的作用量改变 $2\pi I_i$。因此，任何 q、p 的单值函数 $F(q,p)$ 一定是各个角度变量的周期函数。因此我们可以做傅里叶展开：

$$F=\sum_{n_i}A_{n_1,\cdots,n_f}\exp\left[it\left(n_1\frac{\partial E}{\partial I_1}+\cdots+n_f\frac{\partial E}{\partial I_f}\right)\right]\ . \tag{6.127}$$

由此我们看到，任意坐标与动量的单值函数可以表达成一系列周期函数的级数。这个级数的每一项的频率都是一系列基本频率的整数倍之和。这些基本频率的个数等于系统自由度数目，并且可以由下式给出：

$$\omega_i(I)=\frac{\partial E(I)}{\partial I_i}\ ,\quad i=1,2,\cdots,f\ . \tag{6.128}$$

特别需要指出的是，公式 (6.127) 一般并不一定意味着对时间的周期性。也就是说，对于完全可分离的、在有限区域运动的多自由度系统，一般来说它的运动不一定随时间是周期的。这是一个多自由度系统与前面讨论的单自由度系统的重要区别。对于一个单自由度保守力学系统，只要它的运动是局限在有限区域的，它一定是周期的。但是对于多自由度系统，这一点并不成立。一个多自由度的系统，对时间的周期性只是出现在上式中各个频率之比都是有理数的情形下。这时，我们称系统的各个频率 ω_i 是 公度的 (commensurate)；反之，我们则称它们是 非公度的 (incommensurate)。

因此，一般来说一个完全可分离的、多自由度力学系统的运动不是周期的，尽管它的每一个坐标都局域于一个有限的区域内。如果系统的任意两个基本频率都是非公度的，那么系统的运动在任何坐标上的投影也不是周期性的。这时，虽然系统从一个坐标出发不会完全周期地回到初始点，但是它可以任意接近地回到初始点附近。因此，我们称系统的运动是 准周期的 。如果系统的两个或更多的基本频率是公度的，这时我们称这个力学系统是 简并的 。如果系统的所有基本频率都是公度的，我们称该力学系统是 完全简并 的。如果系统是完全简并的，那么显然它的运动是周期的，或者说它的轨道是闭合的。如果系统只是部分简并的，那么系统的运动向那些公度的频率所对应的运动模式的投影将是周期的，尽管系统整体的运动并不是。

一般来说，系统的能量 $E(I)$ 会分别依赖于所有的作用量变量 I_i。但是当系统出现简并的情形下，系统的能量实际上只依赖于更少的独立变量。例如，如果频率 ω_1 与 ω_2 出现简并，即：

$$n_1\frac{\partial E}{\partial I_1}=n_2\frac{\partial E}{\partial I_2}\ , \tag{6.129}$$

其中 n_1，n_2 是两个整数。注意这个式子必须对于任意的 I 都成立，因此，这就要求能量实际上只能依赖于：$n_2I_1+n_1I_2$ 这个线性组合而不能分别独立地依赖于 I_1 和 I_2。这个结论的逆命题也是对的。也就是说，如果系统的能量仅仅依赖于某两个作用量变量的特殊线性组合，那么系统中一定出现了简并。简并的出现可以分为偶然的和必然的两类。所谓必

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

然的简并，主要是指由于系统的对称性所造成的简并。另一类我们则称为偶然简并。其中特别有意思的例子是所谓的动力学简并。这是由于势能的某种特殊形式所造成的偶然简并。典型的例子就是有心力场中的开普勒问题和高维各向同性谐振子。

例题 6.11 **利用作用量 -角度变量讨论开普勒问题** 考虑在 $1/r$ 中心势中的质点的运动，我们将利用作用量 - 角度变量这个问题。

解： 在第 34.2小节中的例题 6.6中，我们已经论证了开普勒问题在哈密顿 -雅克比理论框架中是完全可分离的。由此我们得到：$S_0 = W_1(r) + W_2(\theta)$，而 W_1 和 W_2 由公式 (6.79) 给出。于是:

$$I_\theta = \frac{1}{2\pi}\int_0^{2\pi} \frac{\partial W_2}{\partial \theta} d\theta = J \ . \tag{6.130}$$

也就是说 I_θ 就等于粒子的角动量 (守恒量)。类似地，我们可以构造出 I_r

$$I_r = \frac{1}{2\pi}\oint dr \sqrt{2m\left[E - V(r)\right] - \frac{J^2}{r^2}} \ . \tag{6.131}$$

这个积分可以利用复变函数中残数定理的方法求出 (可能是最为方便、快捷的方法)，结果是:

$$I_r = -I_\theta + \alpha\sqrt{\frac{m}{2|E|}} \ . \tag{6.132}$$

由于我们已经知道角动量 I_θ 是守恒量，因此根据这个式子 I_r 是守恒等价于能量是守恒量。我们可以得到粒子的能量作为作用量变量的函数:

$$E = -\frac{m\alpha^2}{2(I_r + I_\theta)^2} \ . \tag{6.133}$$

这个能量表达式仅仅依赖于组合：$I_r + I_\theta$ 而不分别依赖于 I_r，I_θ。这意味着系统存在简并。对应于坐标 r 和坐标 θ 的基本频率正好相等，因此粒子的轨道一定是周期的 (闭合的)。

另外值得注意的一点是，粒子的轨道参数 (正焦弦和偏心率)p，e 只与 I_r 和 I_θ 有关:

$$p = \frac{I_\theta^2}{m\alpha} \ , \quad e^2 = 1 - \left(\frac{I_\theta}{I_r + I_\theta}\right)^2 \ . \tag{6.134}$$

由于我们知道 I_r 和 I_θ 都是绝热不变量，因此，开普勒问题中粒子的轨道的偏心率是一个绝热不变量。即使在某些参数 (例如粒子的能量) 缓慢变化时，轨道的偏心率也是不变的。

搞定

例题 6.12 **相对论性 Keppler 问题** 同上题，只不过讨论的是相对论性的 Keppler 问题。

解： 相对论性的 Keppler 问题我们在前面讨论过 (见例题 6.2)。那里我们提到过这个问题的历史背景，它最早是 Sommerfeld 在 1916 年研究原子能级时得到的。这里我们来回顾一下当年 Sommerfeld 的一些推导，同时说明它对于旧量子论的作用。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

我们仍然取 $c=1$ 的单位制。体系在是极坐标中的哈密顿量为:

$$H=\sqrt{p_r^2+\frac{p_\theta^2}{r^2}+m^2}-\frac{\alpha}{r}\,. \tag{6.135}$$

显然，p_θ 是守恒量，与之对应的绝热不变量为

$$I_\theta=\frac{1}{2\pi}\oint p_\theta d\theta=p_\theta=J\,, \tag{6.136}$$

另外一个绝热不变量 I_r 的计算要稍微复杂一些。一个比较方便的方法是利用 p_θ 与 p_r 之间的关系:

$$p_r=\frac{p_\theta}{r^2}\frac{dr}{d\theta}=\frac{J}{r^2}r'\,. \tag{6.137}$$

于是我们可以将 I_r 表达为:

$$I_r=\frac{1}{2\pi}\oint p_r dr=\frac{1}{2\pi J}\oint u^{-2}[J^2(u')^2]d\theta\,. \tag{6.138}$$

这里的 $u\equiv(1/r)$ 必须用轨道方程来替代，对 θ 的积分从 0 到 2π。注意到我们有

$$J^2(u')^2=(E+\alpha u)^2-J^2-m^2$$

因此上面的积分可以化为:

$$I_r=\frac{1}{2\pi J}\oint\left[\frac{E^2-m^2}{u^2}+\frac{2\alpha E}{u}+(\alpha^2-J^2)\right]d\theta$$

其中的轨道方程可以取为:

$$\frac{1}{u}=r=\frac{p}{1+e\cos\Omega\theta}\,,$$

其中的 p 和 e 由例题 6.2处的公式给出。于是，上面的积分化为:

$$\begin{aligned} I_r &= \frac{\alpha^2-J^2}{J}+\frac{2(J^2-\alpha^2)}{J}\frac{1}{2\pi}\int_0^{2\pi}\frac{d\theta}{1+e\cos\Omega\theta} \\ &+ \frac{(E^2-m^2)p^2}{J}\frac{1}{2\pi}\int_0^{2\pi}\frac{d\theta}{(1+e\cos\Omega\theta)^2} \end{aligned} \tag{6.139}$$

通常的做法是令 $z=e^{i\theta}$ 并试图在 z 的复平面完成这个积分。这个积分对于任意的 Ω 还是有一定难度滴：由于 $\Omega\neq1$，这使得 $z=0$ 实际上是被积函数的一个枝点。当然，在非相对论极限下，$\Omega\simeq1$ 因此在这个极限下这些积分还是可以比较容易地算出来的。下面我们直接令 $\Omega=1$，分别计算上式中的两个积分。

第一个积分为:

$$I_1 \equiv \frac{1}{2\pi}\int_0^{2\pi}\frac{d\theta}{1+e\cos\theta} \tag{6.140}$$

$$= \left(\frac{2}{e}\right)\frac{1}{2\pi i}\oint\frac{dz}{z}\frac{1}{(2/e)+(z+z^{-1})} \tag{6.141}$$

$$= \left(\frac{2}{e}\right)\frac{1}{2\pi i}\oint\frac{dz}{(z-z_1)(z-z_2)} \tag{6.142}$$

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

其中 z_1 和 z_2 是 $z^2+(2/e)z+1=0$ 的两个根，即

$$z_{1,2}=-(1/e)\pm\sqrt{(1/e)^2-1}\,. \tag{6.143}$$

其中 z_1 是根号前取加号；z_2 则对应于取减号。注意这两个根的乘积恒等于 1，因此一个 (z_1) 位于单位圆内，另一个 (z_2) 则位于单位圆外面。所以积分只会挑出那个在单位圆内的奇点。计算残数后给出：

$$I_1=\left(\frac{2}{e}\right)\cdot\frac{1}{z_1-z_2}=\frac{1}{\sqrt{1-e^2}} \tag{6.144}$$

第二个积分为：

$$I_2\equiv\frac{1}{2\pi}\int_0^{2\pi}\frac{d\theta}{(1+e\cos\theta)^2}=\frac{1}{(1-e^2)^{3/2}}$$

因此，在将例题 6.2处的公式代入后得到：

$$I_r=\frac{J^2-\alpha^2}{J}\left(-1+\frac{1}{\sqrt{1-e^2}}\right)\,. \tag{6.145}$$

注意，上面我们是直接令 $\Omega=1$ 得到的结果，但是实际上 $\Omega\simeq 1$。如果考虑到这点，上面式子中只需要将分母上的 J 替换为 $\sqrt{J^2-\alpha^2}$ 即可。因此我们得到：

$$I_r=\sqrt{J^2-\alpha^2}\left(-1+\frac{1}{\sqrt{1-e^2}}\right)\,. \tag{6.146}$$

这就是当年 Sommerfeld 得到的神奇的公式。

在 Sommerfeld 的量子论假设中，两个绝热不变量都满足一定的量子化条件：

$$I_r=n_r\hbar\,,\quad I_\theta\equiv J=n_\theta\hbar\,. \tag{6.147}$$

其中 n_r，n_θ 都是非负整数。利用它们我们可以反解出能量的表达式：[24]

$$\frac{E}{m}=\left[1+\frac{\alpha^2}{\left(n_r+\sqrt{n_\theta^2-\alpha^2}\right)^2}\right]^{-1/2} \tag{6.148}$$

需要注意的是，这个能量中包含了粒子的静止能量的部分。真正粒子的束缚能量是上式再减去 1 的结果 (一个负值)。这是一个超越时代的结果！如果我们将其对 α 进行展开，我们得到：

$$E=m-\frac{m\alpha^2}{2n^2}+\frac{m\alpha^4}{2n^4}\left(\frac{3}{4}-\frac{n}{n_\theta}\right)+\cdots\,, \tag{6.149}$$

其中所谓的主量子数 $n\equiv n_r+n_\theta$。这个式子的第一项仅仅是粒子的静止能量；第二项就是氢原子的主能级，它只与主量子数 n 有关，并不分别依赖于 n_r 和 n_θ。由于有根号的存在，Sommerfeld 自然地要求 $n_\theta\geq 1$。对于另一个量子数，Sommerfeld 要求它从 0 开始：

[24] 我们这里取了自然单位制：$\hbar=c=1$，因此对于氢原子而言，下式中的 α 可以理解为精细结构常数。对于带有电荷 Z 的类氢原子，我们只需要将 α^2 替换为 α^2Z^2 即可。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

$n_r = 0, 1, 2\cdots$，因此主量子数 $n = 1, 2, 3, \cdots$。当我们取 $n = 1$ 的时候，第二项正好给出著名的氢原子的基态能量 -13.6eV。但是，当我们考察第三项的时候我们发现，它不仅仅依赖于主量子数 n，还直接依赖于 n_θ。换句话说，对于 $n = 2, n_\theta = 1$ 和 $n = 2, n_\theta = 2$ 这两种情形，第三项给出的数值是不同的。由于它仅仅是一个非常小的修正，因此它称为氢原子的精细结构。

这个结果实际上与狄拉克方程的结果十分类似: 搞定

36 决定性中的混沌

¶在这一节中，我们稍微介绍一下动力学系统中的非线性问题。这是一个十分广泛的问题。它的内容足够单独写成一部几百页的著作。我们在这个课程中不可能完全地介绍。有兴趣的同学可以参考 [4] 的第七章作为对该问题的一个导引式的介绍。我们这里只能够更为简略地介绍一些比较重要的概念。

在哈密顿分析力学框架中，特别是在讨论了正则变换以后，我们知道所谓坐标和动量实际上很难严格区分，它们实际上是共生的整体。因此，我们可以引入变量：

$$\xi = (q_1, \cdots, q_f; p_1, \cdots, p_f) \,. \tag{6.150}$$

来统一地表达这些对共轭变量。利用变量 ξ，[25] 一个力学系统的哈密顿正则方程可以统一写成：

$$\dot{\xi} = f(\xi, t) \,. \tag{6.151}$$

也就是说，力学系统的运动方程实际上是一组一阶常微分方程。微分方程的理论告诉我们，只要给定初始的相空间位置：$\xi(t = t_0) = \xi_0$，系统以后的运动就被方程 (6.151) 唯一地确定了。这个性质常常被称为经典力学的 决定性。但是，一般来说微分方程 (6.151) 并不是线性微分方程。因此，一般来说它的求解并不是简单的。但是，只要你愿意，利用各种数值的方法总是可以做到的，尽管其难度取决于函数 $f(\xi, t)$ 的具体行为。

上面讨论的唯一性定理的一个推论就是，微分方程 (6.151) 所确立的相轨道在相空间中是不可能相交的。否则在轨道的交点就会存在两条不同的轨道，这直接与唯一性定理矛盾。既然同样的初始点一定产生同样的相轨道，我们可以考察两个无线接近的初始点的轨道随时间如何演化。假定在某个时刻 (取为 $t = 0$ 时刻) 系统的初始位置为 ξ_0，它确立了一个轨道：$\xi(t; \xi_0)$，我们希望了解一个初始位置在 $\xi_0 + \delta\xi_0$ 的轨道如何随时间演化，其中 $\delta\xi_0$ 是一个无穷小量。这可以通过将两者的微分方程相减得到：

$$\dot{\delta\xi} = \left(\frac{\partial f}{\partial \xi}\right)_{\xi_0} \delta\xi \,, \tag{6.152}$$

[25] 为了简化记号，我们一般用 ξ 来统一标记系统的所有广义坐标和广义动量。如果需要表明其分量，我们会加一个下标：ξ_i 来表示，这里 $i = 1, 2, \cdots, 2f$。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

其中 $\delta\xi(t)=\xi(t;\xi_0+\delta\xi_0)-\xi(t;\xi_0)$，注意，这个微分方程现在对于 $\delta\xi$ 来说是线性的。它告诉了我们两个在 $t=0$ 时刻无线接近的轨道，在以后的时间内 (其实是不太长的时间内) 其相对位置将如何变化。相对于 $\delta\xi$ 而言，矩阵：

$$D_{ij}(\xi_0)\equiv\left(\frac{\partial f_i}{\partial\xi_j}\right)_{\xi=\xi_0}\ , \tag{6.153}$$

实际上是一个常矩阵。如果我们能够选取适当的 $\delta\xi_i$ 的线性组合 η_i 将矩阵 (6.153) 对角化，那么它的本征值中一般会存在实部为正的本征值。于是我们发现，对于这种模式，微分方程 (6.152) 告诉我们 $\eta_i(t)$ 会随着时间指数增加。事实上，在这些本征值中存在一个实部最大的本征值 λ_1，与其对应的本征矢量为 $\eta^{(1)}$，我们假定：$Re(\lambda_1)>0$，于是我们得到：

$$\eta^{(1)}(t)\propto e^{Re(\lambda_1)t}\ . \tag{6.154}$$

由于 λ_1 是实部最大的本征值，因此它所对应的本征模式随时间增长最快。因此，经过一段时间后，它几乎完全左右了 $\delta\xi(t)$ 的长时间行为。也就是说，如果有这样的 λ_1 存在，那么初始位置无限接近的两个轨道会随着时间的增加而指数地分离，分离的速率在长时间后完全由动力系统的矩阵 $D_{ij}(\xi_0)$ 的具有最大实部的本征值 λ_1 所控制。本征值 λ_1 的实部被称为这个动力学系统的**第一李雅普诺夫指数**。对于一般的非线性动力学系统，它的李雅普诺夫指数经常都是大于零的。这个效应有时又被夸大地称为**蝴蝶效应**。一般来说，系统的李雅普诺夫指数依赖于所考虑的位置 ξ_0。当 ξ_0 发生变化后，一般它的李雅普诺夫指数也会变化。在相空间中第一李雅普诺夫指数小于零的区域，我们称系统的轨道是稳定的。这时对任何一个轨道的微小偏移会随时间指数衰减。反之，在相空间中第一李雅普诺夫指数大于零的区域，我们称系统的轨道是不稳定的。

动力学系统轨道的不稳定性往往是与所谓的**混沌**联系在一起的。混沌实际上并没有很一致的定义。例如，有人称我们上面提到的轨道的不稳定性为混沌；也有人称我们前一节提到的系统准周期运动为混沌。无论如何，对于轨道的指数偏离意味着任何初始小的偏差都会随时间而放大。这一点对于利用数值方法求解微分方程 (6.151) 的时候特别应当引起注意。我们知道几乎任何数值的计算都必定存在舍入误差 (roundoff errors)。我们必须保证这些误差在我们能够控制的范围之内。由于这些误差很可能随着时间指数增加，因此这一点必须考虑在相应的算法中。

总之，尽管动力学系统是完全决定性的。但是它的相轨道并不一定是十分“规则的”。由于相空间中不稳定区域的存在，其相轨道在相空间中看上去可以是相当混乱的。事实上，如果我们让两个初始位置无限接近的轨道演化足够长的时间，我们会发现它们很快就“忘记”了对方，分别在相空间中四处徜徉。这就体现了所谓的决定性中的混沌。当然，混沌还会出现在所谓的分立映射 (discrete maps) 过程中。不过这些内容与我们理论力学的内容相距比较远，我们就不讨论了。



相关的阅读

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

本章中我们讨论了哈密顿形式的分析力学。这种讨论很大程度上是纯形式的。但是这种形式上的发展，由于褪去了实际的、具体的内容，因而具有更大的适用性，这使它成为很多其他物理学分支可以借鉴的源泉。例如，只要将泊松括号换成量子力学的对易括号，其形式几乎可以完全照搬到海森堡的量子力学描述中；哈密顿－雅克比方程与薛定谔的波动方程有着天然的联系等等。因此，哈密顿分析力学的重要性恰恰不在于它能够解决多少具体的纯力学问题，而在于它提供的这种形式可以运用到多个物理学的领域。这是我们这一章所希望传递的最重要的信息。

从本章的具体内容来说，第 29 节和第 31 节的内容可以参考朗道书 [1] 的第 §40，§42 节。第 30 节是新加入的一节，主要介绍 ξ 符号和 Liouville 定理，这对于统计力学非常重要。这部分内容主要参考了 [1] 和 [4] 的相关章节。第 32 节的讨论取自朗道书 [1] 的第 §43 节。第 33 节的讨论取自朗道书 [1] 的第 §45 节。第 34 节的讨论则取自朗道书 [1] 的第 §47，§48 节。这一节中的关于哈密顿－雅克比方程与波动力学的关系，以及波动光学与几何光学的关系的讨论主要取自 Goldstein 书 [3] 的第十章的 10-8 节。最后，第 35 节的内容取自朗道书 [1] 的第 §49，§50，§52 节。有关哈密顿力学的讨论也可以参考 Goldstein 书 [3] 中的第八、九、十章相关的讨论。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

附 录 A　　对称性与群

正像物理学（特别是理论物理）的许多课程一样，我们这个课程中经常地遇到的一个概念就是对称性。从数学上讲，描述对称性的数学理论是群论。也正是因为如此，相比于其他的数学分支，群论在物理学中有着更为广泛的应用。在这个附录中，我们将结合经典力学中的例子，简要地介绍对称性与群的概念。这对于进一步理解正文中的一些概念（特别是刚体运动学部分）是有帮助的。[1] 我们的介绍将是十分初步的。希望进一步了解这部分内容的同学，可以参考相关的书籍 [9]。我们的介绍也是比较“物理的”，远没有达到数学上严格的程度。对此有偏好的同学应当去参考相关的数学书籍。

1　对称性与群的定义

¶ 通俗地说，对称性实际上是指当我们对所研究的物理对象进行某种操作之后，对象的某些性质并不发生改变。这时我们就说它具有某种对称性。说简单一点，就是所谓的“做了等于没做”。那些我们做了的事情就被称为对称操作，对某个物理对象的所有对称操作集合起来就构成了一个群。

举例来说，当我们考察一个二维的正三角形的时候，直觉告诉我们它具有某种对称。如果要准确描述这种对称性，实际上需要这样来描述：考虑垂直于该正三角形的平面通过其中心的一个轴，那么如果我们将它绕该轴旋转 $2\pi/3$，三角形正好回到原初的位置。也就是说，“转了 ($2\pi/3$ 角度但效果) 等于没转”。我们就称上述的旋转 $2\pi/3$ 的操作是正三角形的一个对称操作。这时我们称正三角形具有一个三重对称轴。类似的，可以定义所谓的 n-重对称轴。对一个特定的客体，它的对称操作的“数目”是给定的。当然这个数目可以是有限多，也可以是无穷多。将所有的对称操作集合起来，我们就得到这个客体的对称群。在所有对称操作中有一个特别特殊，那就是“不操作”。按照对称操作的定义，这个“不操作”当然也属于对称操作。我们称之为对称群的单位元，记为 $\mathbb{1}$。

群作为一个代数结构必须满足一定的数学公理。下面我们就结合我们对于对称操作和对称群的定义来说明这些“公理”的合理性。

[1] 当然并不是不可或缺的。所以如果实在没有兴趣，也完全可以忽略。

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

1. 首先群 G 是一个非空集合，群上面定义了二元的运算称为群的“乘法”，记为“$\cdot$”。群在乘法下满足闭合性，即 $\forall A_1, A_2 \in G$，$A \equiv A_2 \cdot A_1 = A_2 A_1 \in G$。

   【说明】：这其实就是两个连续的对称操作的“联合操作”。假定有两个对称操作 A_1 和 A_2，那么我们先进行对称操作 A_1，再进行对称操作 A_2，这个联合的操作显然也是对称操作，因此也是对称群的元素。我们将其记为两个对称操作的乘积：$A = A_2 \cdot A_1 = A_2 A_1 \in G$。

2. 群的乘法满足结合律：$A_3(A_2 A_1) = (A_3 A_2) A_1$。

   【说明】：按照对称操作来理解，它们都对应于依次进行 A_1，A_2，A_3 的对称操作。

3. 群必须包含一个特殊的元素，称为单位元，记为 $\mathbb{1}$。它满足 $\forall A \in G$，$A\mathbb{1} = \mathbb{1}A = A$。

   【说明】：由于 $\mathbb{1}$ 代表什么也不做，因此无论 $A\mathbb{1}$ 还是 $\mathbb{1}A$ 都代表只进行一个对称操作 A。需要注意的是，无论任何物体，即使是看起来毫无对称性的物体，也一定具有一个对称操作，那就是“不操作”。因此最为简单的群就是只包含一个元素 -单位元 -的群。这称为平庸群。所以，所有物体都具有的对称性就是平庸群的对称性。这当然是最低的对称性了。

4. 对每一个群元 $A \in G$，都存在一个它的逆，记为 $A^{-1} \in G$ 满足 $AA^{-1} = A^{-1}A = \mathbb{1}$。

   【说明】：这实际上是说每一个对称操作 A 都有一个“逆操作”，它对应于将操作 A 的效果撤销。例如前面提到的旋转 $2\pi/3$ 的逆操作就是旋转 $(-2\pi/3)$。如果前者是对称操作，后者一定也是。

我们看到群的数学定义就像是给对称操作“量身定制”的一样，具有且仅有所有必需的性质，既不多也不少。顺便说一下，第 23.4小节中我们曾经介绍了 Cayley-Klein 参数。这里给出的群的现代定义（或者说抽象的定义）恰恰是英国数学家的 A. Cayley 首先给出的。在他之前，数学家们认为群就是特指置换群。

特别需要注意的是，群的乘法一般来说并不能交换次序。也就是说，它一般不满足交换律。如果某个群的乘法满足交换律，我们就称这样的群为交换群，或者阿贝尔群。反之则称为非阿贝尔群。如果群 G 的元素个数为有限多，我们称为有限群，其元素的个数称为群的阶，记为 $|G|$。如果群的元素个数为无穷，则称为无限群。当然这里面原则上还可以按照其无穷大的“多少”分为可数多的无限群和不可数多的无限群。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

2 两个群之间的那些事儿

¶ 同态与同构

物理学家一般想象群都是给出比较具体的例子。但是数学家更喜欢讨论抽象的群。讨论抽象的群的好处是，如果我们可以建立起不同群之间的对应关系，那么我们会发现许多群实际上是十分类似的，甚至于完全相同的。在讨论两个群之间关系时有两个概念十分重要，一个是群的同态，另一个是群的同构。如果有两个群（及其相应的乘法）$(G,\circ)$ 和 $(H,\cdot)$，以及一个由 G 到 H 的映射 $h: G\mapsto H$ 满足 $\forall u,v\in G$ 有：

$$h(u\circ v)=h(u)\cdot h(v)\,, \tag{A.1}$$

那么 h 就称为从群 G 到群 H 的一个同态（homomorphism）。换句话说，同态是两个群之间保持乘法的映射。容易证明，同态一定将 G 中的单位元映射为 H 中的单位元，并且将任意一个群元的逆映射为群元的逆（即保逆性）：$h(u^{-1})=[h(u)]^{-1}$。注意同态不一定是一一的。如果一个同态是一一的，我们就称之为同构，记为 $G\cong H$。由于同构的两个群之间可以建立起一一的对应关系且这种关系保持群中的所有运算，因此从代数上讲同构的两个群可以认为是完全相同的。

¶ 子群与陪集

另外一个涉及两个群之间关系的概念是子群的概念。如果群 $(G,*)$ 的一个子集 H 在同样的群乘法“$*$”下也构成群，我们就称 H 是 G 的子群，记为 $H\subseteq G$。任何群都包含平庸群以及自身为其子群。除去这两个极端的例子以外，即 $H\subseteq G$ 但 $H\neq G$，我们一般称由较大的一个群 G 所描写的系统具有较高的对称性；反之由其真子集 H 描写的系统则具有较低的对称性。

当一个群 G 存在一个子群 H 时，我们可以对每个 $g\in G$ 定义一个左陪集：

$$gH=\{gh: g\in G, h\in H\}\,. \tag{A.2}$$

由于 g 存在逆 $g^{-1}\in G$，因此映射 $\phi: H\to gH$ 是一个一一的映射。不仅如此，我们可以证明，群 G 中的每个元素，出现且仅出现在一个左陪集之中。[2] 我们可以利用其子群 H 将群 G 进行所谓的陪集分解：

$$G=H+g_1H+g_2H+\cdots\,, \tag{A.3}$$

上式中每个左陪集（其中 H 可以认为是与单位元对应的左陪集）都包含相同多且不重复的元素。因此，如果 G 和 H 都是有限群，我们发现子群 H 的阶 $|H|$ 必定是群 G 的阶 $|G|$

[2] 这可以简单证明如下。如果某个元素 $g_1h_1=g_2h_2$ 同时出现在两个不同的左陪集 g_1H 和 g_2H 之中，我们有 $g_1=g_2h_2(h_1)^{-1}$，但 $h_2(h_1)^{-1}\in H$，这意味着存在 g_1,g_2 之间有关系 $g_1=g_2h'$，$h'\in H$。因此它们一定属于同一左陪集。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

的一个因子。这又被称为拉格朗日定理，尽管在拉格朗日那个年代群的现代定义还没有出现。因此，对于元素数目不是太大的有限群，判别它的子群其实并不太困难。我们只需要将其阶数 G 进行质因子分解就可以猜测出它可能的子群的阶数。完全类似的，对于每个 $g \in G$，我们也可以定义所谓的右陪集：$Hg = \{hg : g \in G, h \in H\}$。需要指出的是，一个群的左陪集和右陪集一般是不同的。但是如果两者完全相同那么这个子群就被称为正规子群(normal subgroup)，记为 $H \lhd G$。换句话说，一个子群 H 是 G 的正规子群当且仅当：$\forall h \in H, g \in G, ghg^{-1} \in H$。正规子群的重要性最早是天才数学家 Galois 意识到的。

¶ 共轭关系与共轭类

群上面还可以定义一种等价关系，称为共轭关系。群 G 中的两个元素 a 和 b 如果对某个 $g \in G$ 满足 $gag^{-1} = b$，我们就称 a 与 b 共轭，记为 $a \sim b$。可以证明这是一种数学上所谓的等价关系。因此它将群的所有元素配分为互不相交的等价类，每个类中包含同等数目的元素。这种等价类因此也称为共轭类。注意，如果群元素是矩阵（或者说是群的表示），那么共轭关系实际上就是线性代数中说的矩阵的相似性。

¶ 直积与半直积

最后，当我们有两个群 $(G, \circ)$ 和 $(H, \cdot)$，我们可以构建两个群的直积群，记为 $G \times H$：

$$G \times H = \{(g, h) : g \in G, h \in H\}\,, \tag{A.4}$$

而其乘法“$*$”的定义为：$(g_1, h_1) * (g_2, h_2) = (g_1 \circ g_2, h_1 \cdot h_2)$。换句话说，新的群元素是一对群元，它们分别来自原先的两个群就像两个不同的坐标一样。两个群的直积群 $G \times H$ 自动包含两个子群，记为 $\tilde{G}$ 和 $\tilde{H}$，它们分别与原先两个群 G 和 H 同构，前者由形如 $\tilde{G} = \{(g \in G, \mathbb{1}_H)\}$ 而后者由形如 $\tilde{H} = \{(\mathbb{1}_G, h \in H)\}$ 的元素构成。显然，$\tilde{G} \cong G$，$\tilde{H} \cong H$，并且可以证明它们都是 $G \times H$ 的正规子群。

除了这里谈到的直积之外，还可以构建两个群的半直积。在半直积中原先的两个子群只有一个是正规子群，另一个不是。典型的例子是刚体的所有运动（包括质心平动）对应的相流形 $\mathbb{Q}$。它实际上是平动的相流形 $\mathbb{R}^3$ 与转动部分的相流形 $SO(3)$（这两个流形其实也是群，只不过平动部分的所谓“群乘法”实际上是矢量的加法）的半直积，记为：$\mathbb{Q} = \mathbb{R}^3 \rtimes SO(3)$，其中 $\mathbb{R}^3$ 是 $\mathbb{Q}$ 的正规子群而 $SO(3)$ 不是。

要说明这点最为快捷的方式是利用下面的给出的表示。我们将三维空间中的任意一个矢量 $\mathbf{x} \in \mathbb{R}^3$ 提升为下面的具有四个分量的矢量：

$$\mathbf{x} \Leftrightarrow x \equiv \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix}\,. \tag{A.5}$$

显然这样的一个“四矢量”其实与原先的三矢量 $\mathbf{x}$ 是一一对应的。于是三维空间的一个任

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

意的转动加平动可以用下面的矩阵表达：

$$g = \begin{pmatrix} R & \mathbf{t} \\ 0 & 1 \end{pmatrix}, \tag{A.6}$$

其中 $R \in SO(3)$ 是一个正常转动而 $\mathbf{t} \in \mathbb{R}^3$ 是任意一个三矢量。大家可以验证，这个矩阵 g 作用于我们前面的 x 的结果为：

$$g \circ x = \begin{pmatrix} R & \mathbf{t} \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} = \begin{pmatrix} R \cdot \mathbf{x} + \mathbf{t} \\ 1 \end{pmatrix} \tag{A.7}$$

换句话说，当 g 作用于 x 后将其相应的三矢量 $\mathbf{x}$ 变为了 $R \cdot \mathbf{x} + \mathbf{t}$，即一个任意的平动加上一个转动。所以公式 (A.6) 的确给出了三维刚体最一般运动的一个表示。容易证明所有公式 (A.6) 给出的矩阵构成一个群，我们称之为 G，这个群其实就是我们前面提到的平移群与正常转动群 $SO(3)$ 的半直积。在这个表示下，群的单位元由

$$\mathbb{1} = \begin{pmatrix} \mathbb{1}_{3\times 3} & \mathbf{0} \\ 0 & 1 \end{pmatrix} \tag{A.8}$$

给出而 g 的逆元素则由

$$g^{-1} = \begin{pmatrix} R^{-1} & -R^{-1}\mathbf{t} \\ 0 & 1 \end{pmatrix} \tag{A.9}$$

给出。通过对上面各个式子的直接计算，大家可以很容易地证明这些结论。群 G 显然包含两个子群，一个对应于正常转动（即 $SO(3)$），另一个对应于纯粹的平动，它们分别由下列矩阵给出：

$$g_R = \begin{pmatrix} R \in SO(3) & \mathbf{0} \\ 0 & 1 \end{pmatrix}, \quad g_T = \begin{pmatrix} \mathbb{1}_{3\times 3} & \mathbf{t} \in \mathbb{R}^3 \\ 0 & 1 \end{pmatrix}, \tag{A.10}$$

大家很容易验证，所有的 g_R 构成一个群我们暂时称之为 H，它其实就是三维正常转动群；所有的 g_T 也构成一个群，我们称之为 N，它其实就是三维的平移群（同构于 $\mathbb{R}^3$）。显然，H 和 N 都是 G 的子群。下面我们来论证：$G = N \rtimes H$。

需要论证的实际上就是 N 是 G 的正规子群而 H 则不是。我们利用前面正规子群的定义进行计算。对于子群 N，我们有：

$$g g_T g^{-1} = \begin{pmatrix} R & \mathbf{t} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbb{1}_{3\times 3} & \mathbf{t} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} R^{-1} & -R^{-1}\mathbf{t} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \mathbb{1}_{3\times 3} & R \cdot \mathbf{t} \\ 0 & 1 \end{pmatrix} \in N$$

同样对于子群 H 我们有：

$$g g_R g^{-1} = \begin{pmatrix} R & \mathbf{t} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A & \mathbf{0} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} R^{-1} & -R^{-1}\mathbf{t} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} RAR^{-1} & \mathbf{t} - RAR^{-1} \cdot \mathbf{t} \\ 0 & 1 \end{pmatrix}$$

我们发现对于子群 H 来说，这样计算的结果一般来说并不属于 H。造成这个的原因是三维转动一般是不可交换的。因此 H 不是 G 的正规子群。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

3 群的例子

¶本节中我们举一些物理学中常见的群的例子，供大家参考。

- 最为平庸的例子就是前面提到的平庸群，它只包含单位元：$G=\{\mathbb{1}\}$。
- 一个稍微不那么平庸的例子是所谓的 Z_2 群，它包含两个元素：$Z_2=\{1,-1\}$，其中 $+1$ 就是单位元，-1 是另外的一个元素，群的乘法就是通常意义的乘法，显然它们构成群。这是一个具有两个元素的阿贝尔群。Z_2 群在物理学中十分重要是因为多数的分立对称性（例如宇称变换）都可以利用这个群描写。
- 类似的包含 n 个元素的推广是所谓的 Z_n 群：$Z_n=\{1,\omega,\omega^2,\cdots,\omega^{n-1}\}$，其中 $\omega=e^{2\pi i/n}$ 为 1 在复数域的 n 次方根。
- 回到我们本附录开始提到的例子，一个正三角形的对称性。显然，它具有绕中心旋转 $2\pi/3$ 的对称性。我们将这个对称操作记为 C_3。那么显然 $C_3^2=(C_3)^2$ 也是一个对称操作，对应于旋转 $4\pi/3$（等价于旋转 $-2\pi/3$）。事实上将不转，转 $2\pi/3$，转 $4\pi/3$ 这三个操作集合起来就构成了群 $C_3=\{\mathbb{1},C_3,C_3^2\}$。[3] 它实际上同构于我们前面提及的群 Z_3。但是这其实还不是一个正三角形全部的对称性。正三角形关于通过其一个顶点的中垂线的反射也是不变的。这个对称操作我们记为 σ_v。那么我们还可以将原先的群扩大一倍，构成群 $C_{3v}=\{\mathbb{1},C_3,C_3^2,\sigma_v,\sigma_vC_3,\sigma_vC_3^2\}$。这才是一个正三角形全部的对称性。这个群实际上同构于两个群的直积：$C_{3v}\cong C_3\times Z_2$。
- 一个稍微复杂一点的例子是一个三维立方体的对称性所对应的群。如果仅仅考虑旋转对称性的部分，它由一个阶数为 24 的群 O 描写。如果加上反射对称性的话，它扩展为阶数为 48 的群 O_h。这两个群都称为立方体群，都是非阿贝尔的有限群。
- 一个极端一些的非阿贝尔有限群是阶数为 $2^{27}3^{14}5^37^211$ 的魔方群（Rubik's Cube group），它包含了大家儿时玩儿的魔方的所有操作。

¶具有无穷多元素的群的例子也是很多的。

- 一个简单的例子是二维平面上的所有转动构成的一个群，称为二维转动群 $SO(2)$。它可以看成是上面提到的 C_n 群（或者 Z_n 群）在 $n\to\infty$ 时的极限。二维的转动是可交换的，因此这是一个阿贝尔群。事实上由于二维的坐标可以利用复数表示，这个群同构于 $U(1)$ 群 -由所有的相因子构成的群。
- 另一个例子就是我们在第 五章讨论刚体转动是提到的三维（正常）转动群或者说三维空间的特殊正交群 $SO(3)$。我们也论证了 -藉由 Cayley-Klein 表示 -它实际上与二维特殊幺正群 $SU(2)$ 之间有个同态关系。这两个群是非阿贝尔群的典型代表。事实

[3] 不幸的是，这个群的符号与生成它的元素 -旋转 $2\pi/3$ 的对称操作 -的符号正好相同了，不过大家应当可以从上下文判断。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

上，上面这些群的特点是其群元素解析地依赖于一系列参数（例如欧拉角），因此它们实际上属于 李群 。

- 上面提到的这些群实际上都是所谓的 一般线性变换群 $GL(V)$ 的特例。所谓一般线性变换群 $GL(V)$ 由矢量空间 V 上的所有非奇异的矩阵构成，相应的群乘法就是矩阵的乘法。

4 群的表示

¶ 如果群的同态是从群 G 到某个矢量空间 V 上的一般线性变换群 $GL(V)$ 的映射，我们称之为群 G 的 表示 （representation）。换句话说，群表示实际上是将抽象的群元素与矢量空间 V 上的 矩阵 对应起来，并且这种对应是保持乘法的。相应的矢量空间的维数也被称为相应的群表示的维数。

说实话，多数物理学家脑子里所想象的群实际上都是数学家脑子里所设想的“抽象的群”的某个表示。例如，我们前面讨论的 $SO(3)$ 群和 $SU(2)$ 群的定义，实际上可以认为是相应的“抽象的群”在相应的矢量空间上的矩阵表示：$SO(3)$ 群是 3×3 的正交矩阵而 $SU(2)$ 群为 2×2 复幺正矩阵。物理学家更热衷于群的表示不是没有原因的。对他们来说，承载这些矩阵的矢量空间本身往往具有实在的物理意义。因此，矢量空间的作用也绝不仅仅是实现某个群的表示而已。比如本课程中讨论的刚体的三维转动，我们讨论的矢量空间就是刚体实实在在“生活的空间”。我们前面也专门提到，虽然对于复 2×2 的 $SU(2)$ 群的 Cayley-Klein 表示所对应的矢量空间在经典力学中貌似缺乏明显的物理意义，但是它实际上在量子力学中对应于自旋 $1/2$ 粒子的角动量的希尔伯特空间。因此，从这个意义上讲，如果群是所有对称操作的集合，那么群表示所对应的矢量空间实际上就是被操作的物理对象所存在的空间。这个矢量空间中的基矢也被称为群表示的 基矢 ，或者简称 基 。

群的表示又可以分为可约的和不可约的两类。粗略地说，如果群的所有元素对应的矩阵都不具有统一的分块结构，这种表示就被称为 不可约 的（irreducible representation, irrep）；反之，如果所有群元的矩阵都具有统一的分块结构，我们就称这种表示是 可约 的（reducible representation）。分块结构意味着我们能够从目前完备的基矢中选出一部分来，在它们所张成的子空间中，群元的矩阵表示是不可约的。将不可约表示进行 直和 ，我们就得到一个可约的表示。反之，我们可以将一个可约的表示分解为一系列不可约表示的直和。 两个群表示 ρ_1 和 ρ_2 除了可以进行直和外还可以进行直积。这个直积表示的定义是：$(\rho_1\otimes\rho_2)(v\otimes w)=\rho_1(v)\otimes\rho_2(w)$，它又称为 ρ_1 和 ρ_2 构成的张量表示。容易证明两个表示的直和或者直积表示仍然是一个表示。如果原先两个表示是不可约的，它们的直和当然是可约的，并且可以分解为原先的那两个表示的直和；但两个不可约的表示的直积表示一般是可约的。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

群表示论中一个十分重要的指标是每个群元对应的矩阵的对角元之和。这被称为该表示的特征标（character）。即如果表示 $\rho : G \mapsto GL(n, C)$，那么一个群元 $g \in G$ 的特征标定义为：

$$\chi(g) = \sum_{i=1}^{n} [\rho(g)]_{ii} = Tr\,[\rho(g)] \ . \tag{A.11}$$

由于表示一定将群的单位元映射到单位矩阵，因此我们有 $\chi(\mathbb{1}) = n$ 就是该表示的维数。由于求迹运算的循环性质，特征标是群的共轭类上的函数。也就是说，如果 $a \sim b$ 属于同一共轭类，它们一定具有相同的特征标：$\chi(a) = \chi(b)$。

特征标满足一系列重要的性质。这些性质可以帮助我们将一个任意的群表示分解为一系列不可约表示的直和。这一点对于讨论量子力学中分子能级的对称性具有十分重要的意义。关于这点我们这里就不深入了。感兴趣的同学可以参考相关的参考书 [8]。

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

参考书

[1] L. D. Landau and E. M. Lifschitz, *Mechanics*, 3rd ed., Pergamon Press, 1994.

[2] L. D. Landau and E. M. Lifschitz, *The Classical Theory of Fields*, 4th ed., Pergamon Press, 1994.

[3] H. Goldstein, *Classical Mechanics*, 2nd ed., Addison-Wesley, 1980.

[4] Jorge V. José and Eugene J. Saletan, *Classical Dynamics: a contemporary approach*, Cambridge University Press, 1998.

[5] C. Truesdell, *Essays in the history of mechanics*, Springer-Verlag, New York, 1968.

[6] Arnold V. I., *Mathematical methods of classical mechanics*, Springer-Verlag, New York, 1978.

[7] J.D. Jackson, *Classical Electrodynamics*, John Wiley & Sons, Inc., 1975.

[8] L. D. Landau and E. M. Lifschitz, *Quantum Mechanics*, Pergamon Press, 1994.

[9] S. Sternberg, *Group theory and physics*, Cambridge University Press, 1994.

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

索引

$O(3)$ 群, 75
$SO(3)$ 群, 75
　　的拓扑结构, 77
$SU(2)$ 群, 82

Bertrand 定理, 47

Cayley-Klein 参数, 81

Dzhanibekov 效应, 90

Floquet 算符, 69

Liouville 定理, 103

slerp 算法, 85

WKB 近似, 119

阿贝尔群, 132

爱因斯坦求和规则, 9, 12, 98

半直积, 134

保守系统, 32

被动观点, 75

苯环的小振动, 66
本征频率, 56, 59
本征矢量, 59

变分, 23
变分法, 23
变分学基本引理, 24

标量场, 27
标量势, 28
标准基, 69

波动光学, 119
泊松定理, 105
泊松括号, 104

不对称陀螺, 89
不稳定约束, 9
布洛赫定理, 70

参数共振, 71
参数振子, 68
参照系, 17

场, 20
场强张量, 29

潮汐, 2, 52

程函, 119

尺度变换, 35

春分秋分点的移动, 96

磁感应强度, 29

达朗伯原理, 12

代表点, 99

单位四元数, 82

电场强度, 29
电量, 28

动力学对称, 47

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

动量守恒, 34
动能, 31

度规张量, 18

对称群, 131
对称陀螺, 89
对易括号, 105

多个自由粒子的作用量, 27

泛函, 21

非阿贝尔群, 132
非公度, 70, 92
非公度的, 124
非惯性系的力学, 38
非完整约束, 8, 11, 39
非完整约束的力学系统, 39
非线性振动, 67
非相对极限, 30
费马原理, 120

分离变量法, 113
分立对称性, 37
分析力学, 4, 5

刚体, 73
刚体的自由度, 74

公度, 70
公度的, 124
共轭, 24
共轭类, 134
共焦椭圆坐标, 117
共形群, 38
共振, 57

固体潮, 52

惯量张量, 85
惯性, 26
惯性离心力, 39
惯性力, 38

光程, 119
光程函, 119
光速不变原理, 17
广义动量, 24
广义力, 12, 24
广义速度, 7, 21
广义坐标, 7–9, 20

规范变换, 30
规范不变性, 30
规范对称性, 30

过阻尼, 58

哈密顿, 4
哈密顿 -雅克比方程, 107, 111
哈密顿－雅克比方程
　　保守系统的, 112
哈密顿方程, 99
哈密顿力学, 4, 97
哈密顿量, 98
哈密顿原理, 8
哈密顿正则方程, 99

耗散系统, 14

黄道面, 95

回旋频率, 116

混沌, 129

基矢, 74
机械动量, 29
几何光学, 119
几何光学近似, 120

伽利略, 2
伽利略时空, 30

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

伽利略相对性原理, 2

简并, 64, 124
简约作用量, 107
简正模, 60
简正坐标, 56, 60

焦点, 45
交换群, 132
角动量守恒, 35, 44
角度变量, 122
角速度矢量, 79

解的叠加性, 68

进动, 93
近日点, 46
近日点的进动, 47

精细结构, 102
经典力学的决定性, 128
静止质量, 26

久期方程, 59
旧量子论, 102

绝热变化参数, 120
绝热不变量, 121

开普勒第二定律, 44
开普勒第三定律, 36, 46
开普勒第一定律, 46
开普勒问题, 45, 113, 125
    相对论性的, 52, 101

科里奥利力, 39

空间反射, 37
空间平移不变性, 34
空间转动不变性, 34

傀标, 18

拉格朗日, 4
拉格朗日乘子, 40
拉格朗日方程, 8
拉格朗日量, 6, 13, 22
    电磁场中带电粒子的, 28
拉普拉斯 -龙格 -楞次矢量, 47

勒让德变换, 98, 99, 110

离心势能, 39, 45
李代数, 78
李群, 137

临界阻尼, 58

洛伦兹变换, 17
洛伦兹标量, 17
洛伦兹群, 20

闵氏空间, 17
闵氏时空, 17

模态矩阵, 60
魔方群, 136
莫佩督原理, 107

母函数, 110

内积, 19

能量 -动量四矢量, 27

逆变矢量, 17
逆变四矢量, 17
逆变指标, 17

牛顿, 2
牛顿方程, 31

欧拉, 3
欧拉 -拉格朗日方程, 3, 13, 24
欧拉定理, 36, 76
欧拉方程, 88

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

欧拉角, 79
偶然简并, 125

陪集分解, 133

偏心率, 46

平动动能, 85
平庸群, 132

切丛, 21

球形陀螺, 89

群, 75, 131
    闭合性, 132
    表示, 137
        基矢, 137
    不可约表示, 137
    乘法, 132
    单位元, 132
    阶, 132
    结合律, 132
    可约表示, 137
    拉格朗日定理, 134
    逆元素, 132
群表示的直和, 137

瑞利耗散函数, 14, 58

三维影射空间, 77
三维正交群, 75
三维转动群, 75

上标, 17

折合质量, 44

生成函数, 110
生成元, 78

时间反演变换, 37
时间平移不变性, 33
时空观, 17
矢量场, 28
矢量力学, 5
矢量势, 28
世界线, 25
势能, 31, 32

守恒量, 104
受迫振动, 57
受迫振子, 57

双重覆盖群, 83

四动量, 27
四矢量, 17
四矢量势, 28
四元数, 4, 82

缩并, 18

特殊三维正交群, 75
特殊三维转动群, 75
特征标, 138

体坐标架, 74

天文章动, 96

同构, 133
同态, 83, 133

完全积分
    哈密顿 -雅克比方程, 111
完全简并, 124
完全可积系统, 123
完整约束, 8, 21

网球拍定理, 90

微分流形, 21
维里定理, 36
位型空间, 21
位型流形, 21

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki

稳定约束, 9

无限群, 132

希尔方程, 68

狭义相对论, 17
下标, 18

线性振动, 67

相对性原理, 17
相轨道, 100
相空间, 99
相流形, 99
　　刚体的, 77, 83

协变矢量, 18
协变四矢量, 18
协变指标, 18

辛几何, 100

虚功原理, 11, 12
虚位移, 11

薛定谔方程, 118

循环坐标, 34, 113

雅克比等式, 104
雅克比原理, 108

赝标量, 37
赝矢量, 37

一般线性变换群, 137
一维固体的振动, 66

引潮力, 52

有限群, 132
有效势能, 45
有心势, 43

右陪集, 134

余切丛, 100
宇称, 37
宇称变换, 37

约化质量, 44
约束, 8
约束力, 9, 11

运动积分, 104

章动, 93
张量, 17
张量的阶, 17

振幅, 57

正常章动, 94
正规子群, 134
正焦弦, 46
正则变换, 110
正则动量, 24
　　电磁场中的, 28
正则方程, 99

直积群, 134
置换群, 132
质心系, 34, 44

轴矢量, 37

主动观点, 75
主动力, 11
主轴方向, 86
主轴转动惯量, 86

转动动能, 85

准经典近似, 119
准周期运动, 124

自由度数目, 20

由 CPHOWIKI 小组整理自互联网　　仅供学习交流使用　　更多资料请访问 https://pan.cpho.wiki

自转, 93

阻尼, 58
阻尼振动, 58
阻尼振子, 58
组合频率, 67

最小作用量原理, 8, 21

左陪集, 133
作用量, 7, 21
作用量变量, 122

由 CPHOWIKI 小组整理自互联网 仅供学习交流使用 更多资料请访问 https://pan.cpho.wiki