探星阁理论物理学讲义

理论力学（第二版）

Theoretical Mechanics （2nd Edition）

D.Liang
dl_phy@qq.com

理论力学

（第二版）

参考资料

- *Classical Dynamics* , David Tong

  这是 Cambridge 应用数学与理论物理系 (*Department of Applied Mathmatics and Theoretical Physics*) David Tong 教授的讲义[1]. 这份讲义的质量很高，但是内容不够全面，只涉及了分析力学部分，不符合国内理论力学的教学大纲，最适合拿来入门分析力学；它让你感觉没有一个字是可以删去的，而且字里行间充满了启发性的灵感火花. 果真就像他们说的那样——不读英文资料你会错过很多.

- *理论力学教程（第五版）*, 周衍柏

  这本书是我在学校里用的教材，是不错的一本理论力学入门书籍，内容很全面；不过在内容的讲解方面显得繁琐而不简洁。推荐用它来学习前面的 Newton 力学以及刚体部分。

- *力学（第四版，下册 理论力学）*, 梁昆淼

  这本书就如同梁老师的《数学物理方法》一样高质！对于这本书，我只能说和它相见恨晚。我是直到大二下学期期末复习的时候才开始看这本书，感觉要是从一开始就看这本书就好了。翻阅了这么多的书，这一本是真正把问题掰开揉碎了的，分析十分全面且一针见血；而且内容同样很全面，十分推荐！

- *力学*, Landau, Lifshitz

[1] 他的这些讲义面向的是 Cambridge 的一个数学学位项目——Mathematical Tripos，完全开源，网址是：https://www.damtp.cam.ac.uk/user/tong/teaching.html。我发现这些宝藏讲义是得益于的理论力学老师的推荐，他说每一个学理论物理的人都值得去学一学 David Tong 的讲义. 这本讲义对我的意义不只在于让我彻底明白了什么时候坐标变换中会含有时间 t 等等，而且也是作为我英文化阅读的开端，这一点我很开心. 多希望有一天我也能成为像 David Tong 教授一样，写许多课程的而且很受大家欢迎的讲义，再免费开放给全世界的物理学冒险家们.

Landau 十卷的第一卷，风格精简且严谨，大家行文. 尤其是第一章，利用 Lagrange 方程以及几条合理的假设，就推出了牛顿定律以及其他各结论，令人眼前一亮。适合作为补充读物。

- *大学生理工专题导读——拉格朗日量和哈密顿量*, Patrick Hamill

  这是中文翻译版本[2]，我那本是机械工业出版社 2023 年 1 月第 1 版. 这本书教会了我变分法，其它的内容也是很好的参考. 要相信机械工业出版社的眼光. 如果你也想对变分法及其应用有一个初步了解，去看一看吧！

- 此外，田光善老师写的理论力学讲义第一章、哔哩哔哩上蔡子星老师对于诺特定理的讲解视频、我的理论力学老师讲授连续体系分析力学的 PPT 等，我都有所参考。

[2]译者是井帅教授，他似乎是一位经管学教授. 译者序里说管理科学的学生看一看这本书也会大有脾益. 十分感谢井老师的翻译工作，但毕竟不是物理领域的，因此在某些名词翻译上难免有所疏忽. 比如将“基本粒子物理学”翻译成了“基本质点物理学”. 不过这也算情有可原，因为粒子和质点的英文都是 particle .

第二版前言

大二下学期，为了准备期末考试，笔者写了本讲义的第一版。现如今，第二版正式与大家见面！本次修订，笔者想要努力实现如下两个非常现实的目标：

- **为其他的物理讲义做好铺垫**众所周知，在大学物理的基础专业课中，*力学*是基础，那么我想*理论力学*对于理论物理各个分支也是同样的定位。后续的讲义中，我们将不可避免地要用到（质点系或场论的）Lagrange 力学、Hamilton 力学中的基础知识（包括概念与方法）；而在本次修订中，笔者刻意增添了许多这一类的基础知识在其中。此外，在后续的其他讲义中，为了行文简洁，我们不便展示理论力学基础知识的细节，而是给读者提供索引，引导读者来阅读本讲义中的相关内容；相应地，本讲义中会有大量且精细的概念剖析与数学公式推导，供读者参考。

- **为夏令营笔试做准备**国内大多数一流高校的理论物理保研夏令营都设置有笔试环节，对应试者的基本功有较高的要求。笔者经过这么长时间的摸索，认为把学过的知识以讲义的形式梳理一遍对于理解掌握有极大的帮助。正如上面所提到的，本讲义中会有大量且精细的概念剖析与数学公式推导，这些内容正是精华所在——唯有躬身去讲义式地梳理知识，才会促进对概念及概念之间联系的思考与剖析式理解；唯有自己去动笔推导、计算，才能发现每一个公式背后最关键、最要害的一步在哪。我想，花费这些功夫不论是对于应试，还是对于今后的科研，都是利远大于弊的。

不过，笔者必须承认，本次修订之后，本讲义中对于**“我们为什么需要这些知识”**，或者说**“这些知识有哪些重要科学意义”**的回答依然非常不如人意，而上面两个方面却恰恰应该是一份讲义最有价值的地方。笔者没能做得很好，其中的原因当然是因

为笔者的学识还没有达到那样的高度，见识还不够广、思考也还不够深。仍需努力！希望下一次再版，能够在这一点上有长足的进步。下面具体来谈谈第二版相较于第一版都有哪些更新吧。（1）对于整个讲义的宏观样貌我做了一些修改。我修改了讲义的名称、扔掉了繁琐无聊的章前介绍、将所有的参考资料合为一体并且挪移到了书的最前面、更换了封面等等；（2）修正了绝大多数勘误。第一版讲义的确有太多的勘误之处，原因是我的手写稿经由 AI 转换成 LaTex 代码之后，并没有去做认真的校对，在此应该向广大读者致歉；（3）认真修订了 *Hamilton-Jacobi* 理论的相关内容，力争能够把这一复杂的理论讲得尽量清楚些；（4）将原先的最后一章无限自由度体系分析力学修订为重要的经典场论，为后续的课程奠定必要的基础；（5）在本次修订中，同时在质点动力学与场论的框架下，我详尽地讨论了极其重要的 *Norther* 定理，它揭示了守恒律的本质；（6）在 Hamilton 力学一章中添加了 *Liouville* 定理与从经典到量子的内容，前者不但为统计力学做了铺垫，还能够帮助我们回答为什么说 Hamilton 力学仅适用于描述没有耗散的体系；后者的行文内容虽人十分牵强，却也为我们学习量子力学奠定了不错的理论基础；（7）添加了有趣的附录部分；（8）还有许多不胜枚举的内容补充与修缮。

一路走来，特别感谢爸爸妈妈的认可、支持与鼓励；特别感谢身边的小伙伴们，他们或积极与我进行交流讨论，或认真阅读我写的讲义并提供修改意见，他们是（排名不分先后）：T.R.Cheng、Q.Zhuo、M.Q.Zhang、Z.C.Huang、Z.Y.Yang、Y.R.Wang、Y.J.Xu、Y.T.Wang 、Y.Y.Chen、X.Y.Wang 等；特别感谢我的中学、大学老师们给予我的鼓励与认可；感谢科学先辈们的付出，让我们能够学习到如此优美的理论物理知识！

中国的基础学科正处在快速发展的阶段，我很想为祖国的理论物理事业出一份力，奈何自己还只是一个学生，人微言轻。当下的我，只能尽自己的一点绵薄之力，将自己在学习过程中的所思所感梳理成讲义。如若这份讲义的萤烛之光能够照亮一点点他人的物理求索之路，将会是对我最好的认可！期待你的读后反馈，让我们一起把这份讲义打磨得越来越好！

D.Liang
2026 年隆冬

第一版前言

学习是一个螺旋上升的过程——从我们小学的科学课开始接触重力、弹簧、引力的概念，到初中学习简单的受力分析，再到高中认识到了力的矢量性质以及力和运动之间的密切联系，再到大学普通物理教给你用高等数学求解更加复杂的问题。我们在不断地进步，思想越来越深邃，但其实我们的视角始终没有跳脱出 Newton 力学的框架，我们所用到的基本原理都是 Newton 运动定律。

而这本小册子能带给我们新的力学框架，拥有新的概念、新的原理和新的方程。这新的力学框架包括 **Lagrange 力学**与 **Hamilton 力学**，而我喜欢将它们合称为**分析力学**，因为只要我们掌握了数学分析（或者低配版地讲：微积分），就获得了参观这个新力学框架的入场券；同时我们称 Newton 为**矢量力学**以示比较，因为矢量和受力分析可以说是 Newton 力学的一大核心（虽然 Newton 力学也需要不少微积分的知识）。

Lagrange 力学在解决静力学和动力学问题有自己的一套新的思路，对于处理某一大些具体问题而言 Lagrange 力学明显体现出了相较于 Newton 力学的优越性。而 Hamilton 力学虽然在处理具体问题时不如前两者有实用性，却具有极大的理论价值。Hamilton 力学（以及 Lagrange ）力学是打开后续更高阶理论物理领域（包括量子力学、广义相对论、量子场论等）大门的钥匙。

这本小册子可以算是笔者在大二下学期在学习分析力学时的学习笔记，是笔者在参考了许多教材、讲义之后，经过自己的思考与整合的产物。虽说其中时不时会出现一些笔者自己的新思考，但是大多还是知识的搬运、整合与总结，只能证明笔者学过、了解过这些东西。真正要写出一份更加富有创新精神的小册子，还要笔者继续修炼才行。

分析力学归根结底是力学，而力学在物理学中毕竟是一门工具性的小科目，笔者下一步的计划是在本小册子中添加更多的运用——包括但不限于如何在电动力学、量子力学中体现和运用分析力学。以及添加更多的故事，包括笔者自己的故事、科学先辈们的故事等等。笔者真的是一个很喜欢故事和科学史的人！

我想把这本小册子献给我的父母，他们一直是我坚实的后盾；献给我的老师和同学们，特别是 ZMQ 、WYR 和 HZC 同学，感谢他们在学习过程中同我进行许多的交流讨论，让我受益良多；献给探星阁，梦开始的地方，也希望探星阁越来越好! 献给书院的学辅和学弟学妹们，或许明年我可以拿这一份小册子开讨论班。

由于是第一版且笔者水平有限，本小册子一定会有一大堆的问题。如果读者发现了任何的知识性错误、非知识性错误（比如错别字、符号勘误等），或者你觉得有哪些内容怎样写可以更好，请不要犹豫，欢迎立刻联系我们！[3]

期待你的读后反馈，让我们一起把这份小册子打磨得越来越好！

D.Liang

2025 年盛夏

[3]联系方式：微信公众号搜索 **"探星阁"** 并在聊天框或者发布本小册子推送的评论区留言。

目录

1 Lagrange 力学 1

1.1 基本概念 . . . 2

1.1.1 约束与直角坐标的独立性 . . . 2

1.1.2 位形空间与广义坐标 . . . 3

1.2 静力学问题与虚功原理 . . . 6

1.2.1 两大力学体系处理静力学问题的区别 . . . 6

1.2.2 虚位移与虚功 . . . 6

1.2.3 虚功原理 . . . 7

1.2.4 利用虚功原理处理静力学问题 . . . 8

1.2.5 约束力的求解问题 . . . 9

1.3 动力学问题与 Lagrange 方程 . . . 11

1.3.1 两大力学体系处理动力学问题的区别 . . . 11

1.3.2 Lagrange 方程的一般形式 . . . 11

1.3.3 保守系统的 Lagrange 量与 Lagrange 方程 . . . 13

1.3.4 带电粒子在电磁场中的 Lagrange 量与 Lagrange 方程（非相对论情况） . . . 15

1.4 Lagrange 力学的简单应用 . . . 18

1.4.1 极坐标系下的运动方程 . . . 18

1.4.2 惯性离心力与 Coriolis 力 . . . 19

1.4.3 两体问题 . . . 21

1.4.4 守恒量（运动积分） . . . 22

1.4.5 小振动问题 25

2 变分法与 Hamilton 原理 29

2.1 变分法简介 30

2.1.1 泛函及其变分简介 30

2.1.2 泛函的极值问题与 Euler - Lagrange 方程 31

2.1.3 有约束的泛函极值问题 32

2.2 变分法的简单应用 * 34

2.2.1 折射定律 34

2.2.2 最速降线 35

2.2.3 悬链线 38

2.3 Hamilton 原理 41

2.3.1 Hamilton 原理 41

2.3.2 Norther 定理（质点动力学） 43

3 Hamilton 力学 47

3.1 Hamilton 正则方程 48

3.1.1 Hamilton 量 48

3.1.2 Hamilton 正则方程 49

3.1.3 Hamilton 量版本的 Hamilton 原理 51

3.2 Poisson 括号 53

3.2.1 Poisson 括号及其性质 53

3.2.2 守恒量 54

3.3 正则变换 55

3.3.1 正则变换的母函数 55

3.3.2 Poisson 括号的正则变换不变性 57

3.3.3 正则变换的辛结构 * 58

3.3.4 无穷小正则变换与 Liouville 定理 60

3.4 Hamilton - Jacobi 理论 64

3.4.1 含时的 Hamilton - Jacobi 方程 64

3.4.2 不含时的 Hamilton - Jacobi 方程 65

3.4.3 Hamilton-Jacobi 理论的简单应用 66
3.5 从经典到量子 * 72
3.5.1 正则量子化 72
3.5.2 Schrödinger 方程的“导出” 72

4 经典场论 **75**
4.1 场方程 76
4.1.1 什么是场? 76
4.1.2 场的 Euler-Lagrange 方程 76
4.1.3 举例一：d’Alembert 方程 77
4.1.4 举例二：Maxwell 方程组 83
4.2 Norther 定理（场论） 85
4.2.1 场论中的守恒量 85
4.2.2 场论中 Noether 定理的证明 86
4.2.3 对称性与守恒量 86

A 经典力学简史 **91**

B 我与变分法之间的故事 **93**

Chapter 1

Lagrange 力学

1.1 基本概念

1.1.1 约束与直角坐标的独立性

描述质点（系）的运动需要首先选取**参考系**，参考系有惯性参考系与非惯性参考系之分；然后，要在参考系中建立**坐标系**来确定每一个质点的位置坐标，位置坐标表征了质点在空间中的位置. 对于一个质点系而言，我们说质点系的位置确定，当且仅当所有质点的位置确定. 把各个质点的位置坐标并在一起，够成质点系的位置坐标. 要描述质点系的空间位置，参考系和坐标系缺一不可.

最基本的坐标系是直角坐标系，利用直角坐标可以映射出广义坐标，见（1.1.2）小节. 今后我们用 (x_1, x_2, x_3) 描述一个质点的直角坐标，而用 $(x_1, x_2, \cdots, x_{3N})$ 描述一个质点系的直角坐标. 下面讨论直角坐标的独立性.

首先看单个质点. **自由质点**的三个直角坐标 (x_1, x_2, x_3) 可以独立地随时间改变. 什么叫"独立地随时间改变"? 我的理解是受任意主动力 $\boldsymbol{F}(\boldsymbol{x}, \dot{\boldsymbol{x}}, t)$ 下，$x_1(t)$、$x_2(t)$、$x_3(t)$ 的函数形式可以是任意的.

可是如果质点的运动受到了**约束**，则 x_1、x_2、x_3 三者不再独立. 我们接下来只考虑**几何约束**，即那些约束关系可以表示成坐标之间的代数方程（约束方程）的约束[1]. 如果一个质点受到约束的约束方程为:

$$f(x_1, x_2, x_3) = 0 \tag{1.1}$$

则受任意主动力下，只有其中的某两个坐标，不妨取 $x_1(t)$ 与 $x_2(t)$，可以对应有任意的函数形式，而第三个坐标随时间的变化情况可以通过将 x_1、x_2 代回约束方程反解出来. 尽管 $x_3(t)$ 可能不唯一，但一定是有限个而不是任意个. 从而我们说 x_3 依赖于 x_1、x_2，三者不独立.

在有约束的情况下，质点受到的力包括主动力与约束力. 正是因为有了约束力，才导致了坐标之间不独立.

可以把以上说法推广到质点系. 设质点系有 k 个约束，相应就有 k 个约束方程:

$$f_i(x_1, x_2, \cdots, x_{3N}) = 0 \quad i = 1, 2, \cdots, k \tag{1.2}$$

[1]这包含了两种情况，第一种情况，约束关系就直接是一个坐标之间的代数式，如 $x_2 + y^2 = 1$; 第二种情况是，约束关系可以通过数学计算，转化成坐标之间的方程式，比如纯滚动模型里的 $\dot{x} = \dot{\theta} R$（只不过这里的坐标是我们下面要提到的广义坐标），从而跟第一种情况是等价的.

每一个约束方程都会减去一个独立坐标的个数，因此质点系的独立坐标只剩下 $3N-k$ 个. 我们简记 $3N-k=\alpha$，并不妨设 α 个独立坐标为 $(x_1, x_2, \cdots, x_\alpha)$.

1.1.2 位形空间与广义坐标

接下来我们定义一个空间[2]叫质点系的**位形空间**，它是一个 α 维空间，其中的向量为 $(x_1, x_2, \cdots, x_\alpha)$. 某时刻质点系的位置状态表现为位形空间中的一个点，而质点系在受力下的运动表现为在位形空间中点动成线，这条线是 α 维位形空间的一条超曲线，又叫做该质点系在位形空间的运动轨迹.

广义坐标是通过位形空间坐标的映射得来的. 现在将位形空间的点 $(x_1, x_2, \cdots, x_\alpha)$ 映射到另一个 α 维空间 $(q_1, q_2, \cdots, q_\alpha)$，如果这两个空间坐标之间的变换关系满足:

- 存在连续的映射 $q_i = q_i(x_1, x_2, \cdots, x_\alpha)$ $(i = 1, 2, \cdots, \alpha)$ 且各个坐标 q_i 之间相互独立.
- 存在连续逆映射 $x_i = x_i(q_1, q_2, \cdots, q_\alpha)$ $(i = 1, 2, \cdots, \alpha)$

则称 $(q_1, q_2, \cdots, q_\alpha)$ 构成了体系的另一个位形空间，$(q_1, q_2, \cdots, q_\alpha)$ 称为体系在此新位形空间下的 α 个**广义坐标**. 这样的映射原则上有无数种，所以位形空间原则上有无数个. 而且根据上述定义易知任意两个位形空间都有一个连续的映射与逆映射关系. 所以原则上从任意一个位形空间的广义坐标出发，可以映射出其它任何一个位形空间的广义坐标. 这些映射关系称为位形空间的坐标变换

下面我们证明: **任意一个位形空间中的点，能够表征体系在实际物理空间中的位置，而任一位形空间中的（超）曲线，可以表征体系实际的运动.** 要证明这个，只要证明体系的 $3N$ 个直角坐标都可以用任意位形空间的广义坐标来表示就行了.

利用 $x_i = x_i(q_1, q_2, \cdots, q_\alpha)$ $(i = 1, 2, \cdots, \alpha)$ 我们可以得到那些我们挑选出来的 α 个互相独立的坐标如何用广义坐标表示. 而其余的 $3N-\alpha$ 个坐标，则可利用约束方程（1.2）将它们也用广义坐标表示. 因此，N 个质点的 $3N$ 个直角坐标都可以一一用广义坐标表示出来. 因此，用广义坐标就可以完备地刻画体系的运动.

[2]这个“空间”是数学意义上的空间.

然而，上面的位形空间坐标变换中我们默认了是在同一个参考系下进行的. 质点系在每一个参考系中都有其若干个位形空间，若要进行不同参考系之间的位形空间坐标变换，则会牵涉到时间 t. 设质点系在 S 系下的直角坐标为 $(x_1, x_2, \cdots, x_{3N})$，在 S' 系下变为 $(x'_1, x'_2, \cdots, x'_{3N})$，不同参考系下直角坐标之间的变换关系一般为：

$$
\begin{aligned}
& x'_i = x'_i(x_1, x_2, \cdots, x_{3N}, t) \\
& x_i = x_i(x'_1, x'_2, \cdots, x'_{3N}, t) \\
& (i = 1, 2, \cdots, 3N)
\end{aligned}
\tag{1.3}
$$

代入（1.2），得 S' 系下的约束方程[3]为

$$
f'_i = f'_i(x'_1, x'_2, \cdots, x'_{3N}, t) \quad (i = 1, 2, \cdots, k)
$$

从而 S' 系下看，独立的坐标个数仍为 $\alpha = 3N - k$ 个，不妨设为 $(x'_1, x'_2, \cdots, x'_\alpha)$，且有：

$$
x'_i = x'_i(x_1, x_2, \cdots, x_\alpha, t) \quad (i = 1, 2, \cdots, \alpha)
$$

$(x'_1, x'_2, \cdots, x'_\alpha)$ 也构成一个位形空间.

原先我们位形空间之间的坐标变换局限于同一参考系之下，现在我们考虑参考系之间的变换，得到更普遍的位形空间坐标变换为[4]：

$$
\begin{aligned}
& q_i = q_i(x'_1, x'_2, \cdots, x'_\alpha) = q_i(x_1, x_2, \cdots, x_\alpha, t) \\
& x_i = x_i(q_1, q_2, \cdots, q_\alpha, t) \\
& (i = 1, 2, \cdots, \alpha)
\end{aligned}
\tag{1.4}
$$

显然，从体系在任一参考系下的位形空间的广义坐标出发，可以映射出在任何另一参考系中的任何位形空间的广义坐标. 我们今后说广义坐标不一定局限于同一参考系.

一个体系的广义坐标的个数称为体系**自由度**. 强调一点，当我们说广义坐标时，

[3]和（1.2）相比，这个约束方程含有时间 t. 我们称不含时间 t 的约束为稳定约束，而含有时间 t 的约束为不稳定约束.

[4]解释一下其中的第一行：q_i 与 x'_i 在同一个参考系，第一个等号意味着同一个参考系下的坐标变换；第二个等号进一步叠加上了从这个参考系到另一个参考系之间的直角坐标变换，于是有了 t.

它们之间就一定是互相独立的了，因此自由度也可以定义为: 描述一个体系位置所需要的独立变量的个数. 自由度在任何参考系下都相同，因为约束方程的个数不会因为转换参考系而减少.

1.2 静力学问题与虚功原理

1.2.1 两大力学体系处理静力学问题的区别

Newton 力学的精髓是矢量，故又被称为矢量力学. 分析力学的精髓是对广义坐标的数学分析，故而得名.

矢量力学处理静力学的方法是画受力图（几何法），或者在直角坐标系下进行力的合成与分解. 这是高中生需要掌握的物理知识.

分析力学中处理静力学问题的方法是利用虚功原理.

1.2.2 虚位移与虚功

一般情形下，地面系下质点系中第 i 个质点的位矢用广义坐标表示是：$\boldsymbol{r_i} = \boldsymbol{r_i}(q_1, q_2, \cdots, q_\alpha, t)$，而质点系中第 i 个质点的虚位移定义为 [5]：

$$\delta \boldsymbol{r_i} = \sum_\alpha \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} \delta q_\alpha \tag{1.6}$$

与之相对应的是实际位移：

$$d\boldsymbol{r_i} = \sum_\alpha \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} dq_\alpha + \frac{\partial \boldsymbol{r}_i}{\partial t} dt$$

[5] 其中出现了 $\frac{\partial \boldsymbol{r}_i}{\partial q_\alpha}$，弄清楚其几何意义是很重要的. 我们已知道，在某一个参考系 S' 下，体系在物理空间中的位置可用该参考系下位形空间中的点（$q_1, q_2, \cdots, q_\alpha$）来描述. 设当体系在位形空间的坐标从（$q_1, q_2, \cdots, q_\alpha$）变为（$q_1, q_2, \cdots, q_\alpha + \delta q_\alpha$）时，在 S' 系下看，第 i 个质点在物理空间的位矢从 $\boldsymbol{r}_i$ 变为 $\boldsymbol{r}_i + \delta \boldsymbol{r}_i$，则我们说此时有：

$$\frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} = \frac{\boldsymbol{r}_i + \delta \boldsymbol{r}_i - \boldsymbol{r}_i}{\delta q_\alpha} = \frac{\delta \boldsymbol{r}_i}{\delta q_\alpha} \tag{1.5}$$

这就是 $\frac{\partial \boldsymbol{r}_i}{\partial q_\alpha}$ 的定义，显然它是（$q_1, q_2, \cdots, q_\alpha$）的函数，因为在位形空间中不同的地方，改变相同的 δq_α，物理空间中的位置改变是不同的. 因为经典力学中空间距离是绝对的，所以就算是改变了参考系，上面那个 $\delta \boldsymbol{r}_i$ 依然不变，从而 $\frac{\partial \boldsymbol{r}_i}{\partial q_\alpha}$ 不变，因此它在任何情况下都只是（$q_1, q_2, \cdots, q_\alpha$）的函数.

$\frac{\partial \boldsymbol{r}_i}{\partial q_\alpha}$ 本身作为物理空间中的一个矢量，可以有其分量. 其直角坐标分量为：

$$\frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} = \left(\frac{\partial x_i^1}{\partial q_\alpha}, \frac{\partial x_i^2}{\partial q_\alpha}, \frac{\partial x_i^3}{\partial q_\alpha} \right)$$

还能从下面的分析明白为什么这个矢量仅仅是（$q_1, q_2, \cdots, q_\alpha$）的函数. 由（1.5）易知，由于体系位形空间中的 q_α 坐标改变了 δq_α，而引起的第 i 个质点的位移大小为：

$$|\delta \boldsymbol{r}_i| = \left| \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} \right| \delta q_\alpha = \sqrt{\left(\frac{\partial x_i^1}{\partial q_\alpha} \right)^2 + \left(\frac{\partial x_i^2}{\partial q_\alpha} \right)^2 + \left(\frac{\partial x_i^3}{\partial q_\alpha} \right)^2} \delta q_\alpha = H_{i\alpha} \delta q_\alpha$$

其中 $H_{i\alpha}$ 称为 Lame 系数，它显然是（$q_1, q_2, \cdots, q_\alpha$）的函数，于是我们上面得到的结论成立.

二者的区别是虚位移是假想的位移，是我们脑中的“思想实验”. 此外还可以看到，虚位移和实际位移有一项之差，该怎么理解这一项之差呢？可以这么理解：虚位移是在约束不随时间改变的情况下假想的的实际位移中的一种；或者说虚位移是一种瞬时的位移；或者说，虚位移是在 $(q_1, q_2, \cdots, q_\alpha)$ 所在参考系 S' 系下观察所看到的实际位移的一个（即相对于 S' 系的位移中的一个），而在地面系下看，实际位移除了有相对于 S' 系的位移之外，还应该加上参考系变换的牵连位移（后面对 t 求偏导那一项就是牵连位移）.

对于稳定约束而言，不涉及参考系变换，$\boldsymbol{r_i}$ 与 t 无关，从而实位移最后一项消失，所有的实际位移都可以考虑为虚位移；对于非稳定约束而言，虚位移和实位移没有必然联系.

在虚位移下力所做的功叫做**虚功**，常用符号 δW 表示.

1.2.3 虚功原理

静力学问题处理的对象一般是受到稳定、理想的几何约束的质点系. 所谓**理想约束**是指对于质点系中所有质点的任意虚位移，它们所受到的所有约束力的虚功为 0.

根据 Newton 第三定律，质点系受力平衡时，对于任何一个质点，它受到的主动力与约束力等大反向，即：

$$\boldsymbol{F}_i + \boldsymbol{R}_i = \boldsymbol{0}$$

这里，我们使用了 d'Alambert 的主动力和约束力的记号. 现将上式乘以任意的虚位移 $\delta\boldsymbol{r}_i$ 并对指标 i 求和后，我们得到所有力所做的虚功之和：

$$\sum_i (\boldsymbol{F}_i + \boldsymbol{R}_i) \cdot \delta\boldsymbol{r}_i = 0$$

理想约束的约束力虚功为 0，所以我们得到所有主动力的虚功之和：

$$\sum_i \delta W_i = \sum_i \boldsymbol{F}_i \cdot \delta\boldsymbol{r}_i = 0 \tag{1.7}$$

即是说，**对于一个受到稳定且理想的几何约束的体系而言，其处于受力平衡的充要条件为：对于所有质点的任意虚位移，它们所受的所有主动力所做的虚功之和为 0**. 这

就是**虚功原理**. 它能够帮助我们解决理想、稳定的几何约束体系下的静力学问题.

1.2.4 利用虚功原理处理静力学问题

解决静力学问题就是要找到质点系在平衡时所受到的各个力之间的关系，而且我们往往关注的是主动力而不关注约束力. 我们的思路如下:

- 首先确定体系的广义坐标，原则上一个体系的广义坐标有无数个，因为空间坐标变换的形式不唯一. 但我们一般是依解题方便而选择我们的广义坐标.
- 然后，先在直角坐标系下用体系的所有直角坐标（$x_1, x_2, \cdot, x_{3N}$）写出质点系受到的所有主动力（包括内主动力和外主动力）的虚功 δW 的表达式，并且写为全微分形式:

$$\delta W = \sum_i \boldsymbol{F}_i \cdot \delta \boldsymbol{r}_i = \sum_i \sum_N F_{iN} \delta x_i^N$$

- 然后利用空间坐标变换式以及约束方程（1.2），或者用其他技巧，不论如何，将虚功改写成对广义坐标的全微分，你将得到一个形如下式的式子:

$$\delta W = \sum_\alpha Q_\alpha \delta q_\alpha \tag{1.8}$$

其中 Q_α 定义为第 i 个体系受到的**广义力**的 α 分量. 虚功原理要求对于任意的 $\delta \boldsymbol{r_i}$（从而任意的 δq_α），虚功都为 0，其充要条件是 $Q_\alpha = 0$ 对于任意的 α 都成立，于是你得到了 α 个方程，解这些方程就可以得到平衡条件，不过结果只包含了各个主动力之间的关系，而不包含约束力（因为我们虚功的表达式只包含主动力）.

需要说明的是，上面的方法同时也给出了广义力的求法. 但是广义力的求法不止一种，更直接的求法是对体系虚功中的 $\delta \boldsymbol{r_i}$ 直接进行操作:

$$\delta W = \sum_i \boldsymbol{F}_i \cdot \delta \boldsymbol{r}_i = \sum_i \boldsymbol{F}_i \cdot \sum_\alpha \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} \delta q_\alpha = \sum_\alpha (\sum_i \boldsymbol{F}_i \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha}) \delta q_\alpha \tag{1.9}$$

于是广义力定义为：$Q_\alpha = \sum_i \boldsymbol{F}_i \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha}$. 于是，现在你有两种方法求广义力.

1.2.5 约束力的求解问题

上面我们求解出的结果中只包含了主动力，我们只能得到平衡时各个主动力之间的关系，而无从知道约束力. 这是因为在虚功的表达式中，约束力做的虚功为 0 从而消失了. 本小节介绍分析力学中求解约束力的方法. 需要强调的是，约束力的求解一般放在直角坐标系下来处理，而不利用广义坐标. 这是因为直角坐标更容易描述方向.

我们秉持由易到难的原则，先看单个质点的情形：设一个质点受到约束：

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

且主动力为:$\boldsymbol{F} = (F_x, F_y, F_z)$ 已知. 下面我们求:(1) 它的平衡位置;(2) 约束力. 设该质点的平衡位置为 (x, y, z)，让其在约束允许的范围内有一假想偏移，位置变为 $(x + \delta x, y + \delta y, z + \delta z)$，由于是在约束允许范围内变动，显然有：

$$f(x+\delta x, y+\delta y, z+\delta z) - f(x, y, z) = \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial z}\delta z = 0$$

由虚功原理易知：$F_x\delta x + F_y\delta y + F_z\delta z = 0$ 所以：

$$F_x\delta x + F_y\delta y + F_z\delta z + \lambda(\frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial z}\delta z) = 0$$

即：

$$(F_x + \lambda\frac{\partial f}{\partial x})\delta x + (F_y + \lambda\frac{\partial f}{\partial y})\delta y + (F_z + \lambda\frac{\partial f}{\partial z})\delta z = 0$$

接下来，你可能要先选 x, y 为广义坐标，然后用 $f(x, y, z)$ 消去 z 与 δz，再让 δx 和 δy 系数为零，从而你可以解出 x, y, λ. 不过，你实际上只用到了下面这几个式子：

$$\begin{cases} F_x + \lambda\frac{\partial f}{\partial x} = 0 \\ F_y + \lambda\frac{\partial f}{\partial y} = 0 \\ F_z + \lambda\frac{\partial f}{\partial z} = 0 \\ f(x, y, z) = 0 \end{cases}$$

所以干嘛不直接解这个方程组呢？这个方程组一共有四个方程、x,y,z,λ 四个未知量，解之得平衡位置 (x,y,z). 而知道了 λ，也就知道了约束力. 因为上面前三个方程显然是在说：

$$\boldsymbol{F}+\lambda\nabla f=0$$

而又有

$$\boldsymbol{F}+\boldsymbol{R}=0$$

从而有

$$\boldsymbol{R}=\lambda\nabla f$$

其实上面我们讲的思路换一个理解方式会更简单：$f(x,y,z)=0$ 描述了 $\mathbb{R}^3$ 中的一个曲面，其法线方向就是 ∇f 的方向，约束力就是沿法向的，所以我们设 $\boldsymbol{R}=\lambda\nabla f$. 在平衡时显然有：

$$\begin{cases}\boldsymbol{F}+\boldsymbol{R}=\boldsymbol{F}+\lambda\nabla f=\boldsymbol{0}\\ f(x,y,z)=0\end{cases}\tag{1.10}$$

解之得 x,y,z,λ，从而 ∇f 也知道了，从而 $\boldsymbol{R}$ 也知道了.

下面把上述方法推广至 N 个质点组成的质点系. 设系统有 k 个约束：

$$f_j(x_1,x_2,\cdots,x_{3N})=0\ ,(j=1,2,3\cdots k)$$

设质点系的总约束力为

$$\boldsymbol{R}=\sum_j\lambda_j\nabla f_j$$

平衡时有：

$$\begin{cases}\boldsymbol{F}+\boldsymbol{R}=\sum_i\boldsymbol{F}_i+\sum_j\lambda_j\nabla f_j=\boldsymbol{0}\\ f_j(x_1,x_2,\cdots,x_{3N})=0\,(j=1,2,3\cdots,k)\end{cases}\tag{1.11}$$

这是 $3N+k$ 个方程，$3N+k$ 个未知数（$3N$ 个坐标和 k 个 λ），解之可得 $x_1,x_2,\cdots,x_{3N}$ 与 $\lambda_1,\lambda_2\cdots\lambda_k$，从而得知系统所有信息.

唯一要注意的是，这里的 $\boldsymbol{F}$、$\boldsymbol{R}$、∇f_j 均是 $3N$ 维的向量，是将 N 个质点的力并在一起形成的.

1.3 动力学问题与 Lagrange 方程

1.3.1 两大力学体系处理动力学问题的区别

不论是在矢量力学的体系下，还是分析力学的体系下，求解动力学问题的核心一定是找**运动方程**. 它是一个包含各个（广义）坐标对时间的二阶导数、一阶导数、零阶导数以及常数项的常微分方程. 此外，如果能找到体系的那些不随时间改变的量即**守恒量**，那也将会大大有助于问题的简化与解决. 两大力学体系对于动力学问题的处理的不同之处在于它们导出运动方程和诸守恒量的方式不同.

矢量力学中，导出运动方程以及守恒量的依据是牛顿定律. 而在分析力学的框架下，导出运动方程的途径是求解 Lagrange 方程，而得到守恒量的方法是分析 Lagrange 量的对称性，在（1.4）节我们会有所体会.

1.3.2 Lagrange 方程的一般形式

下面我们从 Newton 第二定律出发来求 Lagrange 方程的一般形式. 瞬时的[6] Newton 第二定律指出:

$$\boldsymbol{F}_i + \boldsymbol{R}_i = m_i \ddot{\boldsymbol{r}}_i$$

两边乘上任意虚位移 $\delta \boldsymbol{r}_i$ 并对 i 指标求和，得:

$$\sum_i (\boldsymbol{F}_i + \boldsymbol{R}_i) \cdot \delta \boldsymbol{r}_i = \sum_i \boldsymbol{F}_i \cdot \delta \boldsymbol{r}_i = m_i \ddot{\boldsymbol{r}}_i \cdot \delta \boldsymbol{r}_i$$

地面系下第 i 个质点的位矢为:

$$\boldsymbol{r}_i = \boldsymbol{r}_i(x_i^1, x_i^2, x_i^3) = \boldsymbol{r}_i(q_1, q_2, \cdots, q_\alpha, t)$$

其中（$q_1, q_2, \cdots, q_\alpha$）不一定与（$x_1, x_2, \cdots, x_\alpha$）在同一参考系，所以带上 t. 下面我们用广义坐标来改写上式:

$$\text{等式左边} = \sum_i \boldsymbol{F}_i \cdot \delta \boldsymbol{r}_i = \sum_i \boldsymbol{F}_i \cdot \sum_\alpha \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} \delta q_\alpha = \sum_\alpha Q_\alpha \delta q_\alpha$$

[6] 一定要强调瞬时，因为在某个参考下看，约束方程可能会随时间改变，从而约束力可能会随时间改变.

$$
\begin{aligned}
\text{等式右边} &= \sum_i m_i \ddot{\boldsymbol{r}}_i \cdot \delta \boldsymbol{r}_i \\
&= \sum_i m_i \ddot{\boldsymbol{r}}_i \cdot \sum_\alpha \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} \delta q_\alpha \\
&= \sum_i \sum_\alpha m_i \ddot{\boldsymbol{r}}_i \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} \delta q_\alpha \\
&= \sum_i \sum_\alpha \left[\frac{d}{dt} (m_i \dot{\boldsymbol{r}}_i \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha}) - m_i \dot{\boldsymbol{r}}_i \cdot \frac{d}{dt} \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} \right] \delta q_\alpha \\
&= \sum_i \sum_\alpha \left[\frac{d}{dt} (m_i \dot{\boldsymbol{r}}_i \cdot \frac{\partial \dot{\boldsymbol{r}}_i}{\partial \dot{q}_\alpha}) - m_i \dot{\boldsymbol{r}}_i \cdot \frac{\partial \dot{\boldsymbol{r}}_i}{\partial q_\alpha} \right] \delta q_\alpha \\
&= \sum_i \sum_\alpha \left[\frac{d}{dt} \frac{\partial (\frac{1}{2} m_i \dot{\boldsymbol{r}}_i^2)}{\partial \dot{q}_\alpha} - \frac{\partial (\frac{1}{2} m_i \dot{\boldsymbol{r}}_i^2)}{\partial q_\alpha} \right] \delta q_\alpha \\
&= \sum_\alpha \left[\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_\alpha} - \frac{\partial T}{\partial q_\alpha} \right] \delta q_\alpha
\end{aligned}
$$

其中 $T = \sum_i \frac{1}{2} m_i \dot{\boldsymbol{r}}_i^2$，为体系总动能. 对比等式两边，我们得到:

$$
\sum_\alpha \left(\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_\alpha} - \frac{\partial T}{\partial q_\alpha} - Q_\alpha \right) \delta q_\alpha = 0
$$

由于 δq_α 为任意的，故只能是:

$$
\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_\alpha} - \frac{\partial T}{\partial q_\alpha} = Q_\alpha \tag{1.12}
$$

这就是**一般形式的 Lagrange 方程**.

下面要对上述推导过程进行一些说明:

- 在直角（地面系）下看，质点 i 的速度为:

$$
\dot{\boldsymbol{r}}_i = \frac{d\boldsymbol{r}_i}{dt} = \sum_\alpha \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} \dot{q}_\alpha + \frac{\partial \boldsymbol{r}_i}{\partial t} dt
$$

  其中 $\dot{q}_\alpha$ 为**广义速度**, 从而有:

$$
\frac{\partial \dot{\boldsymbol{r}}_i}{\partial \dot{q}_\alpha} = \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} \tag{1.13}
$$

  这一关系是在上述推导中关键的一步[7]，值得记住. 此外，值得强调的是，$\dot{\boldsymbol{r}}_i$ 的

[7]（1.13）式也有几何意义，读者不妨思考一下.

第一项是由于广义坐标的改变而引起的速度，或者说是相对于 $(q_1, q_2, \cdots, q_\alpha)$ 所在的参考系的速度. 第二项是从 $(q_1, q_2, \cdots, q_\alpha)$ 所在的参考系变换到地面系时所附带的牵连速度. 不要将对位矢的全倒数与偏导数混淆.

- 我们在（1.2.2）小节的脚注中提到，$\frac{\partial \boldsymbol{r}_i}{\partial q_\alpha}$ 仅是 $(q_1, q_2, \cdots, q_\alpha)$ 的函数，因此：

$$\frac{d}{dt}\frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} = \sum_\beta \frac{\partial}{\partial q_\beta}\left(\frac{\partial \boldsymbol{r}_i}{\partial q_\alpha}\right)\dot{q}_\beta = \frac{\partial}{\partial q_\alpha}\sum_\beta \frac{\partial \boldsymbol{r}_i}{\partial q_\beta}\dot{q}_\beta = \frac{\partial \dot{\boldsymbol{r}}_i}{\partial q_\alpha} \tag{1.14}$$

  这在上述推导中也是关键的一步，也值得记住.

- 我们上述推导的过程并没有明确指出位置矢量 $\boldsymbol{r}_i$ 所处的参考系的类型. 其实不论是惯性系还是非惯性系上述推导都是适用的，只要在考虑主动力 $\boldsymbol{F}_i$ 的时候同时考虑实际力和惯性力即可. 以及，如果我们选择 $\boldsymbol{r}_i$ 是惯性参考系中的位置矢量，也并不影响我们选择广义坐标时取非惯性系中的广义坐标，广义坐标的选择也是任意的.

1.3.3 保守系统的 Lagrange 量与 Lagrange 方程

上面我们导出了 Lagrange 方程的一般形式，下面我们来看一个重要的情形——主动力为保守力.

为什么要强调保守力? 一方面因为我们在宏观、微观领域中的诸多物理模型中都需引入保守力; 另一方面因为保守力系的 Lagrange 方程具有很好的结构, 这一结构是我们找寻其他系统 Lagrange 量的基础. 下面首先推导单个质点受保守力的 Lagrange 方程.

保守力 $\boldsymbol{F}$ 满足 $\nabla\times\boldsymbol{F} = \boldsymbol{0}$, 从而可以引入势函数 $V = V(\boldsymbol{r}, t) = V(q_1, q_2, \cdots, q_\alpha, t)$, 使得:

$$\boldsymbol{F} = -\nabla V(\boldsymbol{r}, t)$$

从而保守力 $\boldsymbol{F}$ 在广义坐标 $(q_1, q_2, \cdots, q_\alpha)$ 下的虚功为:

$$\delta W = \boldsymbol{F}\cdot\delta\boldsymbol{r} = -\sum_\alpha \nabla V\cdot\delta\boldsymbol{r} = -\sum_\alpha \frac{\partial V}{\partial q_\alpha}\delta q_\alpha$$

从而广义力为:

$$Q_\alpha = -\frac{\partial V}{\partial q_\alpha}$$

代入一般形式的 Lagrange 方程中去:

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_\alpha} - \frac{\partial T}{\partial q_\alpha} = -\frac{\partial V}{\partial q_\alpha}$$

即:

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_\alpha} - \frac{\partial (T-V)}{\partial q_\alpha} = 0$$

因为 V 仅与 $\boldsymbol{r}$、t 有关，与 $\dot{q}_\alpha$ 无关，故有 $\frac{\partial V}{\partial \dot{q}_\alpha} = 0$，于是

$$\frac{d}{dt}\frac{\partial (T-V)}{\partial \dot{q}_\alpha} - \frac{\partial (T-V)}{\partial q_\alpha} = 0$$

定义 Lagrange 量[8]:

$$L = T - V$$

所以我们得到了单个质点在保守力系下的 Lagrange 方程:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_\alpha} - \frac{\partial L}{\partial q_\alpha} = 0 \tag{1.15}$$

下面推广至质点系. 质点系中每个质点都有个势函数 $V_i = V_i(q_1, q_2, \cdots, q_\alpha, t)$，从而有:

$$\begin{aligned}
\sum_i \delta W_i &= -\sum_i \nabla V_i \cdot \delta \boldsymbol{r}_i \\
&= -\sum_\alpha \sum_i \frac{\partial V_i}{\partial q_\alpha}\delta q_\alpha \\
&= -\sum_\alpha \frac{\partial \sum_i V_i}{\partial q_\alpha}\delta q_\alpha \\
&= -\sum_\alpha \frac{\partial V}{\partial q_\alpha}\delta q_\alpha
\end{aligned}$$

[8]容易分析出一个保守系的 L 一般为广义坐标、广义速度、时间的函数——因为动能一般为广义坐标、广义速度（如果还涉及参考系变换的话，那么还有时间）的函数，而势能一般为广义坐标和时间的函数. 对于其他的体系，我们也是在默认了 L 是广义坐标、广义速度和时间的函数的前提下，去凑 L 的具体形式.（1.3.4）小节会进一步说明.

从而体系受到广义力为

$$Q_\alpha = -\sum_\alpha \frac{\partial V}{\partial q_\alpha}$$

从而，同样得到：

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_\alpha} - \frac{\partial L}{\partial q_\alpha} = 0$$

其中 Lagrange 量：

$$L = \sum_i (T_i - V_i) = T - V$$

为总动能减总势能.

1.3.4 带电粒子在电磁场中的 Lagrange 量与 Lagrange 方程（非相对论情况）

对于其它系统求 Lagrange 方程的策略是：引入一个广义势 V，它可以是 $\boldsymbol{r}$、$\dot{\boldsymbol{r}}$、t 的函数，使得方程：

$$\frac{d}{dt}\frac{\partial (T-V)}{\partial \dot{q}_\alpha} - \frac{\partial (T-V)}{\partial q_\alpha} = 0$$

在形式上仍旧成立[9]，它就作为这个体系的 Lagrange 方程. 而且当我们找到了这个广义势 V，$T-V$ 就定义为带电粒子在电磁场中运动的 Lagrange 量，这也是我们找一个一般体系 Lagrange 量的步骤.

明确了上面的策略之后，我们接下来的工作其实就是要找满足（1.15）的 L，这只能去凑了，下面一起来凑：

带电量为 q 的单个粒子在电磁场中受到 Lorentz 力作用：

$$\boldsymbol{F} = q(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}) \tag{1.16}$$

这力的虚功为：

$$\begin{aligned}\delta W &= \boldsymbol{F} \cdot \delta \boldsymbol{r} \\ &= q\boldsymbol{E} \cdot \delta \boldsymbol{r} + q(\boldsymbol{v} \times \boldsymbol{B}) \cdot \boldsymbol{v}\delta t\end{aligned}$$

[9]为什么需执着于凑出形如（1.15）这样的形式呢？一方面因为这个形式正是变分法中的 Euler - Lagrange 方程之形式，从而能与变分扯上关系，详见第 2 章；另一方面，这样的方程形式能让我们用统一的手段去分析运动积分，详见（1.4）节.

$$
\begin{aligned}
&= q\boldsymbol{E}\cdot\delta\boldsymbol{r} + q(\boldsymbol{v}\times\boldsymbol{v})\cdot\boldsymbol{B}\delta t \\
&= q\boldsymbol{E}\cdot\delta\boldsymbol{r} = m\ddot{\boldsymbol{r}}
\end{aligned}
$$

电场有静电场与涡旋电场之分，这里的 $\boldsymbol{E}$ 应是总电场，即：

$$
\boldsymbol{E} = -\nabla\varphi - \frac{\partial\boldsymbol{A}}{\partial t}
$$

其中 φ 为电标势，$\boldsymbol{A}$ 为磁矢势，它们都只是广义坐标和时间的函数，与广义坐标对时间的导数即广义速度无关.

“容易”[10]知道：

$$
\begin{aligned}
&\sum_\alpha\left[\frac{d}{dt}\frac{\partial(T-q\varphi)}{\partial\dot{q}_\alpha} - \frac{\partial(T-q\varphi)}{\partial q_\alpha}\right]\delta q_\alpha = -q\frac{\partial\boldsymbol{A}}{\partial t}\cdot\delta\boldsymbol{r} \\
=&-q\frac{d}{dt}(\boldsymbol{A}\cdot\delta\boldsymbol{r}) + q\sum_\beta\frac{\partial\boldsymbol{A}}{\partial q_\beta}\dot{q}_\beta\cdot\delta\boldsymbol{r} + \sum_\alpha q\boldsymbol{A}\cdot\frac{d}{dt}\left(\frac{\partial\boldsymbol{r}}{\partial q_\alpha}\right)\delta q_\alpha \\
=&\sum_\alpha\left[-q\frac{d}{dt}\left(\boldsymbol{A}\cdot\frac{\partial\boldsymbol{r}}{\partial q_\alpha}\right)\delta q_\alpha + q\sum_\beta\frac{\partial\boldsymbol{A}}{\partial q_\beta}\dot{q}_\beta\cdot\frac{\partial\boldsymbol{r}}{\partial q_\alpha}\delta q_\alpha + q\boldsymbol{A}\cdot\frac{\partial\dot{\boldsymbol{r}}}{\partial q_\alpha}\delta q_\alpha\right] \\
=&\sum_\alpha\left[-q\frac{d}{dt}\left(\boldsymbol{A}\cdot\frac{\partial\dot{\boldsymbol{r}}}{\partial\dot{q}_\alpha}\right)\delta q_\alpha + q\sum_\beta\frac{\partial\boldsymbol{A}}{\partial q_\beta}\dot{q}_\beta\cdot\frac{\partial\boldsymbol{r}}{\partial q_\alpha}\delta q_\alpha + q\frac{\partial(\boldsymbol{A}\cdot\dot{\boldsymbol{r}})}{\partial q_\alpha}\delta q_\alpha - q\frac{\partial\boldsymbol{A}}{\partial q_\alpha}\cdot\dot{\boldsymbol{r}}\delta q_\alpha\right] \\
=&\sum_\alpha\left[-q\frac{d}{dt}\frac{\partial(\boldsymbol{A}\cdot\dot{\boldsymbol{r}})}{\partial\dot{q}_\alpha} + \frac{\partial(\boldsymbol{A}\cdot\dot{\boldsymbol{r}})}{\partial q_\alpha}\right]\delta q_\alpha + \sum_\alpha\sum_\beta q\frac{\partial\boldsymbol{A}}{\partial q_\beta}\cdot\frac{\partial\boldsymbol{r}}{\partial q_\alpha}\dot{q}_\beta\delta q_\alpha \\
&-\sum_\alpha\sum_\beta q\frac{\partial\boldsymbol{A}}{\partial q_\alpha}\cdot\frac{\partial\boldsymbol{r}}{\partial q_\beta}\frac{\delta q_\beta}{\delta t}\delta q_\alpha \\
=&\sum_\alpha\left[-q\frac{d}{dt}\frac{\partial(\boldsymbol{A}\cdot\dot{\boldsymbol{r}})}{\partial\dot{q}_\alpha} + \frac{\partial(\boldsymbol{A}\cdot\dot{\boldsymbol{r}})}{\partial q_\alpha}\right]\delta q_\alpha + \sum_\alpha\sum_\beta q\frac{\partial\boldsymbol{A}}{\partial q_\beta}\cdot\frac{\partial\boldsymbol{r}}{\partial q_\alpha}\dot{q}_\beta\delta q_\alpha \\
&-\sum_\alpha\sum_\beta q\frac{\partial\boldsymbol{A}}{\partial q_\alpha}\cdot\frac{\partial\boldsymbol{r}}{\partial q_\beta}\frac{\delta q_\alpha}{\delta t}\delta q_\beta \\
=&\sum_\alpha\left[-q\frac{d}{dt}\frac{\partial(\boldsymbol{A}\cdot\dot{\boldsymbol{r}})}{\partial\dot{q}_\alpha} + \frac{\partial(\boldsymbol{A}\cdot\dot{\boldsymbol{r}})}{\partial q_\alpha}\right]\delta q_\alpha
\end{aligned}
$$

[10]这是我在写某次作业时写下的推导过程，这并不是无厘头的推导，我在凑之前是不知道最后的结果的，但我仍然凑出来了. 凑 Lagrange 量也是有技巧的，比如需要使用到函数乘积的求导公式、比如先确定全导数那一项等等. 其实我在凑这个 Lagrange 的时候受益于对（1.12）推导过程的理解和掌握.

从而，原式变为:

$$
\begin{aligned}
&\sum_{\alpha}\left[\frac{d}{dt}\frac{\partial(T-q\varphi)}{\partial\dot{q}_\alpha}-\frac{\partial(T-q\varphi)}{\partial q_\alpha}\right]\delta q_\alpha\\
=&\sum_{\alpha}\left[-q\frac{d}{dt}\frac{\partial(\boldsymbol{A}\cdot\dot{\boldsymbol{r}})}{\partial\dot{q}_\alpha}+\frac{\partial(\boldsymbol{A}\cdot\dot{\boldsymbol{r}})}{\partial q_\alpha}\right]\delta q_\alpha
\end{aligned}
$$

故，带电粒子在电磁场中运动的 Lagrange 量为:

$$
L=T-q\varphi+q\boldsymbol{A}\cdot\dot{\boldsymbol{r}} \tag{1.17}
$$

使得:

$$
\frac{d}{dt}\frac{\partial L}{\partial\dot{q}_\alpha}-\frac{\partial L}{\partial q_\alpha}=0
$$

仍成立.

1.4 Lagrange 力学的简单应用

这一节我们小试牛刀，介绍一下如何用 Lagrange 力学解决具体的问题. 本节内容展示了 Lagrange 体系的强大威力.

1.4.1 极坐标系下的运动方程

在直角坐标系下，取广义坐标为:

$$\begin{cases} r(x,y) = \sqrt{x_2 + y^2} \\ \theta(x,y) = \arctan \frac{y}{x} \end{cases}$$

你就搭建了极坐标. 逆过来有

$$\begin{cases} x(r,\theta) = r\cos\theta \\ y(r,\theta) = r\sin\theta \end{cases}$$

从而

$$\begin{cases} \dot{x} = \dot{r}\cos\theta - r\sin\theta\dot{\theta} \\ \dot{y} = \dot{r}\sin\theta + r\cos\theta\dot{\theta} \end{cases}$$

故，用广义坐标表示动能为:

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) = \frac{1}{2}m(\dot{r}^2 + \dot{\theta}^2 r^2)$$

由

$$\begin{aligned} \delta W &= F_x\delta x + F_y\delta y \\ &= F_x(\cos\theta\delta r - r\sin\theta\delta\theta) + F_y(\sin\theta\delta r + r\cos\theta\delta\theta) \\ &= (F_x\cos\theta + F_y\sin\theta)\delta r + (F_y\cos\theta - F_x\sin\theta)r\delta\theta \\ &= F_r\delta r + F_\theta r\delta\theta \end{aligned}$$

得广义力为[11]:

$$\begin{cases} Q_r = F_r = ma_r \\ Q_\theta = F_\theta r = ma_\theta r \end{cases}$$

代入一般形式的 Lagrange 方程:

$$\begin{cases} \frac{d}{dt}\frac{\partial T}{\partial \dot{r}} - \frac{\partial T}{\partial r} = ma_r \\ \frac{d}{dt}\frac{\partial T}{\partial \dot{\theta}} - \frac{\partial T}{\partial \theta} = mra_\theta \end{cases}$$

简单运算一下，我们得到:

$$\begin{cases} a_r = \ddot{r} - \dot{\theta}^2 r \\ a_\theta = \ddot{\theta} r + 2\dot{\theta}\dot{r} \end{cases} \tag{1.18}$$

这就是那个不太好记忆的极坐标系下的运动方程，利用 Lagrange 方程，1 分钟以内我们就可以推导出来.

1.4.2 惯性离心力与 Coriolis 力

惯性离心力与 **Coriolis 力**也是让我们在学习力学时感到头疼的两个量，下面展示如何用 Lagrange 方程推导出它们.

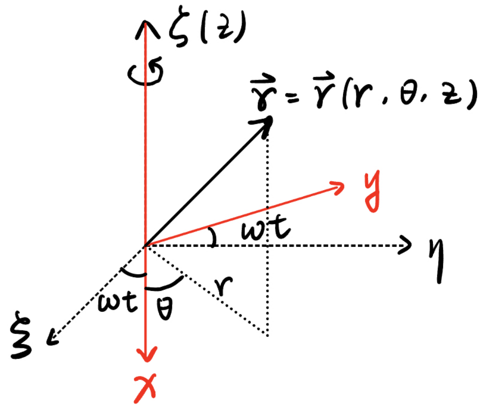

图 1.1: S 系与 S' 系

有一个自由质点在空间运动. 如图 (1.1) 所示，现在有两个参考系 S 与 S'，S 系为惯性系，坐标系为 $\xi-\eta-\zeta$，S' 系为非惯性系，坐标系为 $x-y-z$. 其中 z 轴

[11] 或者你也可以用“瞪眼法”得到广义力是什么.

与 ζ 轴重合，x、y 轴一开始与 ξ、η 轴重合，随后 S' 系的坐标系绕 z 轴转动，角速度 $\boldsymbol{\omega} = \omega\hat{z}$. 在 S 系下看，取 r，θ，z 为广义坐标，则有坐标变换:

$$\begin{cases} \xi(r,\theta,t) = r\cos(\theta+\omega t) \\ \eta(r,\theta,t) = r\sin(\theta+\omega t) \\ \zeta(r,\theta,t) = z \end{cases}$$

故:

$$\begin{cases} \dot{\xi} = \dot{r}\cos(\theta+\omega t) - r(\dot{\theta}+\omega)\sin(\theta+\omega t) \\ \dot{\eta} = \dot{r}\sin(\theta+\omega t) + r(\dot{\theta}+\omega)\cos(\theta+\omega t) \\ \dot{\zeta} = \dot{z} \end{cases}$$

故在 S 系下，用我们选取的广义坐标表示体系的动能为

$$T = \frac{1}{2}m(\dot{\xi}^2+\dot{\eta}^2+\dot{\zeta}^2) = \frac{1}{2}m\left[\dot{r}^2+(\dot{\theta}+\omega)^2r^2+\dot{z}^2\right]$$

代入到自由质点的 Lagrange 方程，有:

$$\begin{cases} \frac{d}{dt}\frac{\partial T}{\partial \dot{r}} - \frac{\partial T}{\partial r} = m\left[\ddot{r}-(\dot{\theta}+\omega)^2r\right] = 0 \\ \frac{d}{dt}\frac{\partial T}{\partial \dot{\theta}} - \frac{\partial T}{\partial \theta} = m\ddot{\theta}r^2 + 2m(\dot{\theta}+\omega)\dot{r}r = 0 \\ \frac{d}{dt}\frac{\partial T}{\partial \dot{z}} - \frac{\partial T}{\partial z} = m\ddot{z} = 0 \end{cases}$$

从而，在 S 系下看质点的加速度为:

$$\begin{cases} a_r = \ddot{r} - \dot{\theta}^2r = \omega^2r + 2\omega\dot{\theta}r \\ a_\theta = \ddot{\theta}r + 2\dot{\theta}\dot{r} = -2\omega\dot{r} \\ a_z = \ddot{z} = 0 \end{cases}$$

从而，在 S' 系下看，质点受力为:

$$\begin{aligned} \boldsymbol{F} &= ma_r\hat{e}_r + ma_\theta\hat{e}_\theta + ma_z\hat{e}_z \\ &= m\omega^2r\hat{e}_r + 2m\omega\dot{\theta}r\hat{e}_r - 2m\omega\dot{r}\hat{e}_\theta \end{aligned}$$

$$
\begin{aligned}
&= m\omega^2 \boldsymbol{r} - 2m \begin{vmatrix} \hat{e}_r & \hat{e}_\theta & \hat{e}_z \\ 0 & 0 & \omega \\ \dot{r} & \dot{\theta} r & \dot{z} \end{vmatrix} \\
&= m\omega^2 \boldsymbol{r} - 2m\boldsymbol{\omega} \times \boldsymbol{v} \qquad (1.19)
\end{aligned}
$$

上式中 $m\omega^2\boldsymbol{r}$ 就是惯性离心力，$-2m\boldsymbol{\omega}\times\boldsymbol{v}$ 一项就是 Coriolis 力[12].

1.4.3 两体问题

在力学课上我们已经学过，质心系下的两体问题可以通过引入**约化质量** μ 的方式转化成单体问题。事实上，Lagrange 力学提供了一个很直白的方式帮助你理解为什么可以这样操作。

假设现在有两个质点，质量分别为 m_1 和 m_2。以质心为原点，二者的位置矢量分别为 $\boldsymbol{r_1}$ 和 $\boldsymbol{r_2}$。紧接着，定义 $\boldsymbol{r} = \boldsymbol{r}_2 - \boldsymbol{r}_1$ 为从质点 1 指向质点 2 的矢量，则利用质心的定义我们可得 $\boldsymbol{r}_1 = -\frac{m_2}{m_1+m_2}\boldsymbol{r}$ 与 $\boldsymbol{r}_2 = \frac{m_1}{m_1+m_2}\boldsymbol{r}$. 然后我们可以写下体系在质心系下的拉格朗日量为

$$
\begin{aligned}
L_C &= \frac{1}{2}m_1\dot{\boldsymbol{r}}_1^2 + \frac{1}{2}m_2\dot{\boldsymbol{r}}_2^2 - V(r) \\
&= \frac{1}{2}\frac{m_1 m_2^2}{(m_1+m_2)^2}\dot{\boldsymbol{r}}^2 + \frac{1}{2}\frac{m_2 m_1^2}{(m_1+m_2)^2}\dot{\boldsymbol{r}}^2 - V(r) \\
&= \frac{1}{2}\mu\dot{\boldsymbol{r}}^2 - V(r) \qquad (1.20)
\end{aligned}
$$

其中 $\mu = \frac{m_1 m_2}{m_1+m_2}$ 为约化质量。这个 Lagrange 量和一个质量为 μ，位置矢量为 $\boldsymbol{r}$ 的在势场 $V(r)$ 中运动的单质点的 Lagrange 量一样，因此所有的动力学分析都是一样的。至此我们成功地将质心系下的两体问题转化为单体问题。

顺便说一句，如果想要进一步在地面系下进行分析、写出在地面系下的 Lagrange 量，我们只需要在上面质心系下的 Lagrange 量中添加上质心相对于地面系的动能项即可。

[12] 相信很多人跟我一样，大一学力学时第一次看到极坐标下的运动方程（1.18）的时候，感觉似乎其中就有惯性离心加速度和 Coriolis 加速度的意思，所以就开始去思考这两者有什么联系. 当时是没想出来，现在有了上述的推导，我看到了它们的联系了. 你看到了吗，我的朋友?

1.4.4 守恒量（运动积分）

这一节我们探讨守恒量。在动力学问题中，我们常常希望寻找体系的守恒量来简化问题的描述与求解，而获得这些守恒量的手段是对一些微分关系式进行积分即**运动积分**. 我们往往也直接称这些守恒量为运动积分. 进行运动积分所用到的这些微分关系可以利用 Lagrange 方程去找，下面来看几个最常用的运动积分——能量积分和广义动量积分.

(1) 能量积分与能量守恒

一个体系的 Lagrange 量一般为 $L=L(q,\dot{q},t)$（这里我们把 $q_1,q_2,\cdots,q_\alpha$ 压缩到一个符号 q 中）易知:

$$
\begin{aligned}
\frac{dL}{dt} &= \sum_\alpha \left(\frac{\partial L}{\partial q_\alpha}\dot{q}_\alpha + \frac{\partial L}{\partial \dot{q}_\alpha}\ddot{q}_\alpha\right) + \frac{\partial L}{\partial t} \\
&= \sum_\alpha \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_\alpha}\dot{q}_\alpha + \frac{\partial L}{\partial \dot{q}_\alpha}\ddot{q}_\alpha\right) + \frac{\partial L}{\partial t} \\
&= \sum_\alpha \frac{d}{dt}\left(\dot{q}_\alpha \frac{\partial L}{\partial \dot{q}_\alpha}\right) + \frac{\partial L}{\partial t}
\end{aligned}
$$

从而

$$
\frac{d}{dt}\left(\sum_\alpha \dot{q}_\alpha \frac{\partial L}{\partial \dot{q}_\alpha} - L\right) + \frac{\partial L}{\partial t} = 0 \tag{1.21}
$$

这就是我们得到的第一个微分关系. 若 L 不显含 t，则 $\frac{\partial L}{\partial t}=0$，从而有

$$
\frac{d}{dt}\left(\sum_\alpha \dot{q}_\alpha \frac{\partial L}{\partial \dot{q}_\alpha} - L\right) = 0
$$

若坐标变换 $\boldsymbol{r}_i=\boldsymbol{r}_i(q_1,q_2,\cdots,q_\alpha)$ 不显含 t [13]，即不涉及参考系变换，而且势能 V 只与广义坐标有关，即势能不是广义势能，则有:

$$
\begin{aligned}
\sum_\alpha \dot{q}_\alpha \frac{\partial L}{\partial \dot{q}_\alpha} - L = \sum_\alpha \dot{q}_\alpha \frac{\partial T}{\partial \dot{q}_\alpha} - T + V &= \sum_i \sum_\alpha \dot{q}_\alpha m \dot{\boldsymbol{r}}_i \cdot \frac{\partial \dot{\boldsymbol{r}}_i}{\partial \dot{q}_\alpha} - T + V \\
&= \sum_i m\dot{\boldsymbol{r}}_i \cdot \sum_\alpha \frac{\partial \dot{\boldsymbol{r}}_i}{\partial \dot{q}_\alpha}\dot{q}_\alpha - T + V = \sum_i m\dot{\boldsymbol{r}}_i \cdot \dot{\boldsymbol{r}}_i - T + V
\end{aligned}
$$

[13]这是必要的，因为若 $\boldsymbol{r}_i=\boldsymbol{r}_i(q_1,q_2,\cdots,q_\alpha,t)$，则 $\dot{\boldsymbol{r}}_i=\sum_\alpha \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha}\dot{q}_\alpha + \frac{\partial \boldsymbol{r}_i}{\partial t}$

$$
\begin{aligned}
&=2T-T+V\\
&=T+V\\
&=E
\end{aligned}
$$

即上面那个被求导的物理量就是总能量 E . 从而 $\frac{dE}{dt}=0$，两边积分，就是**能量积分**，得出总能量 E 是个常数.

(2) 广义动量积分与广义动量守恒

对于一个体系的 Lagrange 量 $L=L(q,\dot{q},t)$，若其中不显含某个坐标 q_α[14]，则有:

$$
\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_\alpha}-\frac{\partial L}{\partial q_\alpha}=\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_\alpha}=0 \tag{1.22}
$$

定义与广义坐标 q_α 共轭的**广义动量**为

$$
p_\alpha=\frac{\partial L}{\partial \dot{q}_\alpha} \tag{1.23}
$$

故

$$
\frac{dp_\alpha}{dt}=0
$$

两边积分就是**广义动量积分**，得出若 L 不显含 q_α，则 p_α 守恒. 这时候，q_α 被成为**循环坐标**或**可遗坐标**.

值得指出，从做题方法的角度来看，用上述两个运动积分来搭建运动方程往往是要比用 Lagrange 方程来转化运动方程要方便的.

这就好比是在中学时期，我们往往利用能量、动量守恒来解题而很少直接用 Newton 第二定律来解题一样. 守恒量不论是在 Newton 体系还是 Lagrange 体系中都是很有力的工具. 在解题的时候要牢记这一点

下面举一些例子:

[14]这样的坐标被称为**循环坐标**或**可遗坐标**.

有心力场

有心力场[15]中质点的角动量守恒. 此时，Lagrange 量为

$$L = \frac{1}{2}m(\dot{r}^2 + \dot{\theta}^2 r^2) - V(r)$$

L 不显含 θ，从而角动量

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta} = \text{const}$$

是守恒量.

电磁场中的带电粒子

下面是带电粒子在电磁场中的运动规律，Lagrange 量为:

$$L = T - \varphi q + q\boldsymbol{A} \cdot \dot{\boldsymbol{r}}$$

首先来看能量守恒

$$\begin{aligned}
E &= \sum_\alpha \dot{q}_\alpha \frac{\partial L}{\partial \dot{q}_\alpha} - L \\
&= \sum_\alpha \dot{q}_\alpha \frac{\partial T}{\partial \dot{q}_\alpha} + q\sum_\alpha \dot{q}_\alpha \frac{\partial (\boldsymbol{A} \cdot \dot{\boldsymbol{r}})}{\partial \dot{q}_\alpha} - T + \varphi q - q\boldsymbol{A} \cdot \dot{\boldsymbol{r}} \\
&= 2T + q\sum_\alpha \boldsymbol{A} \cdot \frac{\partial \boldsymbol{r}}{\partial q_\alpha}\dot{q}_\alpha - T + \varphi q - q\boldsymbol{A} \cdot \dot{\boldsymbol{r}} \\
&= T + \varphi q + q\boldsymbol{A} \cdot \dot{\boldsymbol{r}} - q\boldsymbol{A} \cdot \dot{\boldsymbol{r}} \\
&= T + \varphi q \\
&= \text{const}
\end{aligned}$$

其次是（广义）动量守恒。设 L 中不显含 q_α，则有

$$p_\alpha = \frac{\partial L}{\partial \dot{q}_\alpha} = \frac{\partial T}{\partial \dot{q}_\alpha} + q\boldsymbol{A} \cdot \frac{\partial \dot{\boldsymbol{r}}}{\partial \dot{q}_\alpha}$$

[15]要注意，有心力和保守力是不同的. 有心力特指那些只与坐标 r 有关的保守力，一般的保守力除了与 r 有关，还可能与其它（广义）坐标有关系.

$$
\begin{aligned}
&= m\dot{\boldsymbol{r}} \cdot \frac{\partial \dot{\boldsymbol{r}}}{\partial \dot{q}_\alpha} + q\boldsymbol{A} \cdot \frac{\partial \dot{\boldsymbol{r}}}{\partial \dot{q}_\alpha} \\
&= (m\dot{\boldsymbol{r}} + q\boldsymbol{A}) \cdot \frac{\partial \boldsymbol{r}}{\partial q_\alpha} \\
&= const
\end{aligned}
$$

其中第一项是粒子的 (机械) 动量，第二项为电磁场的电磁动量. 上述结果说明算上了电磁动量的 p_α 才守恒，单单粒子的机械动量是不守恒的.

1.4.5 小振动问题

本节我们看 Lagrange 力学如何帮助我们处理小振动问题。首先看自由度为 1 的小振动问题. 设质点在一维势场 $V(x)$ 中运动. Lagrange 量为:

$$L = \frac{1}{2}m\dot{x}^2 - V(x)$$

由 Lagrange 方程得:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = m\ddot{x} + \frac{dV}{dx} = 0$$

故

$$m\ddot{x} = -\frac{dV}{dx} \tag{1.24}$$

这其实就是 Newton 第二定律，只不过上面我们用 Lagrange 力学将其导出. 利用 $\frac{dV}{dx} = 0$，可以找到质点在这个势场中的平衡位置，并且人为设置它为坐标原点 $x = 0$.

现在想象有一个质量为 m 的质点在平衡位置附近做微小运动，其 $x(t)$ 的值始终为小量. 把 $-\frac{dV}{dx}$ 做 Taylor 展开，得

$$m\ddot{x} = -\left(V'|_{x=0} + V''|_{x=0}\, x + \cdots\right) = -V_0'' x$$

上面我们利用了 $V'|_{x=0} = 0$ 且忽略了 x 的高阶项. 从而得到了小振动方程:

$$\ddot{x} = -\frac{V_0''}{m}x \tag{1.25}$$

其中 $\frac{V_0''}{m}$ 为常量，称之为**回复系数**. 这个方程是很常规的二阶微分方程，并不难解.

下面我们讨论 N 个质点组成质点系的情形，这时体系的自由度不止为 1，每一个质点都在其平衡位置附近做小振动。下面我们首先给出用 Lagrange 力学处理这类问题的步骤

- 选取合适的广义坐标（在小振动问题中，我们选取的广义坐标有自己的名字，叫**简正坐标**）$\boldsymbol{r} = (q_1, q_2, \cdots, q_\alpha)$，注意其中 q 一定为为小量，因为我们在研究小振动。

- 然后利用 Lagrange 方程，外加小量分析舍去高阶项，很快地写出如下形式的线性方程组:
$$\ddot{\boldsymbol{r}} = \boldsymbol{F}\boldsymbol{r}$$
其中 $\boldsymbol{r}$ 为列矩阵，$\boldsymbol{F}$ 为 α 维方块矩阵，其解析形式见后文的证明过程。

- Lagrange 力学保证了 $\boldsymbol{F}$ 的本征值为实数，从而上述方程组的求解很容易.

最后我们来证明为什么 Lagrange 力学保证了 $\boldsymbol{F}$ 的本征值为实数。设质点系有广义坐标 $q_1, q_2, \cdots, q_\alpha$. 对任何一个坐标 q_α，有 Lagrange 方程（注意这里我将 Lagrange 量拆开为动能和势能项）：
$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_\alpha} - \frac{\partial T}{\partial q_\alpha} = -\frac{\partial V}{\partial q_\alpha}$$
将上述 Lagrange 方程用广义力改写:
$$\begin{aligned}
\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_\alpha} - \frac{\partial T}{\partial q_\alpha} = Q_\alpha& \\
&= \sum_i m_i \ddot{\boldsymbol{r}}_i \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} \\
&= \sum_i m_i \sum_\beta \left(\frac{\partial \boldsymbol{r}_i}{\partial q_\beta}\dot{q}_\beta + \frac{\partial \dot{\boldsymbol{r}}_i}{\partial q_\beta}\dot{q}_\beta \right) \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} \\
&= \sum_i m_i \sum_\beta \left(\frac{d}{dt}\frac{\partial \boldsymbol{r}_i}{\partial q_\beta}\dot{q}_\beta + \frac{\partial \dot{\boldsymbol{r}}_i}{\partial q_\beta}\dot{q}_\beta \right) \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} \\
&= \sum_i m_i \left(\sum_\beta \sum_\gamma \frac{\partial^2 \boldsymbol{r}_i}{\partial q_\beta \partial q_\gamma}\dot{q}_\beta \dot{q}_\gamma + \sum_\beta \frac{\partial^2 \boldsymbol{r}_i}{\partial q_\beta \partial t}\dot{q}_\beta \right) \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} \\
&= \sum_\beta \left(\sum_i m_i \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_\beta} \right) \ddot{q}_\beta
\end{aligned}$$

$$= -\frac{\partial V}{\partial q_\alpha}$$

请注意在上述推导过程中，由于 q_β、q_γ 为小量，故 $\dot{q}_\beta$、$\dot{q}_\gamma$ 亦为小量，因此 $\dot{q}_\beta\dot{q}_\gamma$ 项为高阶小量，被我们舍去. 至此我们实际上得到了一个由 α 个二阶微分方程组成的微分方程组，用矩阵形式表达:

$$\begin{pmatrix} \sum_i m_i \frac{\partial \boldsymbol{r}_i}{\partial q_1}\cdot\frac{\partial \boldsymbol{r}_i}{\partial q_1} & \cdots & \sum_i m_i \frac{\partial \boldsymbol{r}_i}{\partial q_1}\cdot\frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} \\ \vdots & \ddots & \vdots \\ \sum_i m_i \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha}\cdot\frac{\partial \boldsymbol{r}_i}{\partial q_1} & \cdots & \sum_i m_i \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha}\cdot\frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} \end{pmatrix} \begin{pmatrix} \ddot{q}_1 \\ \ddot{q}_2 \\ \vdots \\ \ddot{q}_\alpha \end{pmatrix} = -\begin{pmatrix} \frac{\partial V}{\partial q_1} \\ \frac{\partial V}{\partial q_2} \\ \vdots \\ \frac{\partial V}{\partial q_\alpha} \end{pmatrix} \tag{1.26}$$

这个方程和自由度为 1 时的方程（1.24）对应，于是我们相应地可以对等号右边 Taylor 展开:

$$\frac{\partial V}{\partial q_\alpha} = \frac{\partial V}{\partial q_\alpha}(q_1, q_2, \cdots, q_\alpha) = \frac{\partial V}{\partial q_\alpha} + \sum_\beta \frac{\partial^2 V}{\partial q_\alpha \partial q_\beta} q_\beta + \cdots = \sum_\beta \frac{\partial^2 V}{\partial q_\alpha \partial q_\beta} q_\beta$$

这里，我们忽略了高阶小量，并利用了平衡位置处[16] $\frac{\partial V}{\partial q_\alpha} = 0$. 从而:

$$\begin{pmatrix} -\frac{\partial V}{\partial q_1} \\ -\frac{\partial V}{\partial q_2} \\ \vdots \\ -\frac{\partial V}{\partial q_\alpha} \end{pmatrix} = -\begin{pmatrix} \sum_\alpha \frac{\partial^2 V}{\partial q_\alpha \partial q_1} q_\alpha \\ \sum_\alpha \frac{\partial^2 V}{\partial q_\alpha \partial q_2} q_\alpha \\ \vdots \\ \sum_\alpha \frac{\partial^2 V}{\partial q_\alpha \partial q_\alpha} q_\alpha \end{pmatrix} = -\begin{pmatrix} \frac{\partial^2 V}{\partial q_1 \partial q_1} & \cdots & \frac{\partial^2 V}{\partial q_1 \partial q_\alpha} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 V}{\partial q_\alpha \partial q_1} & \cdots & \frac{\partial^2 V}{\partial q_\alpha \partial q_\alpha} \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_\alpha \end{pmatrix}$$

从而，方程组变为:

$$\begin{pmatrix} \sum_i m_i \frac{\partial \boldsymbol{r}_i}{\partial q_1}\cdot\frac{\partial \boldsymbol{r}_i}{\partial q_1} & \cdots & \sum_i m_i \frac{\partial \boldsymbol{r}_i}{\partial q_1}\cdot\frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} \\ \vdots & \ddots & \vdots \\ \sum_i m_i \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha}\cdot\frac{\partial \boldsymbol{r}_i}{\partial q_1} & \cdots & \sum_i m_i \frac{\partial \boldsymbol{r}_i}{\partial q_\alpha}\cdot\frac{\partial \boldsymbol{r}_i}{\partial q_\alpha} \end{pmatrix} \begin{pmatrix} \ddot{q}_1 \\ \ddot{q}_2 \\ \vdots \\ \ddot{q}_\alpha \end{pmatrix} = -\begin{pmatrix} \frac{\partial^2 V}{\partial q_1 \partial q_1} & \cdots & \frac{\partial^2 V}{\partial q_1 \partial q_\alpha} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 V}{\partial q_\alpha \partial q_1} & \cdots & \frac{\partial^2 V}{\partial q_\alpha \partial q_\alpha} \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_\alpha \end{pmatrix}$$

[16] 此外请读者注意，这里为了美观，我们没有用"0"来标注物理量在平衡位置处的取值，读者应该明白这些坐标之前的系数通通都是在平衡位置处的取值，是确定值而不是一个变量.

简记为:

$$\boldsymbol{A}\ddot{\boldsymbol{r}} = \boldsymbol{B}\boldsymbol{r}$$

从而:

$$\ddot{\boldsymbol{r}} = \boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{r} = \boldsymbol{F}\boldsymbol{r} \tag{1.27}$$

其中 $\boldsymbol{F}$ 应该为在平衡位置处的取值，是一个常矩阵，这就是和（1.25）对应的方程，只不过这是一个线性微分方程组. 解这个方程就能够得到各个坐标随时间的小振动关系了，剩下的就是数学功夫了. 解这个方程要用到 $\boldsymbol{F}$ 的本征值，我们最后再证明 $\boldsymbol{F}$ 的本征值总是实数[17]:

设 $\boldsymbol{r}$ 为 $\boldsymbol{F}$ 的一个本征向量，本征值为 f. 即:

$$\boldsymbol{F}\boldsymbol{r} = f\boldsymbol{r}$$

从而:

$$\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{r} = f\boldsymbol{r}$$

从而:

$$\boldsymbol{B}\boldsymbol{r} = f\boldsymbol{A}\boldsymbol{r}$$

从而:

$$\boldsymbol{r}^{\dagger}\boldsymbol{B}\boldsymbol{r} = f\boldsymbol{r}^{\dagger}\boldsymbol{A}\boldsymbol{r}$$

其中 $\boldsymbol{r}^{\dagger}$ 为 $\boldsymbol{r}$ 的 Hermit 共轭，即对 $\boldsymbol{r}$ 转置取共轭. 注意到 $\boldsymbol{A}$、$\boldsymbol{B}$ 均为实对称矩阵，于是 $\boldsymbol{r}^{\dagger}\boldsymbol{A}\boldsymbol{r}$、$\boldsymbol{r}^{\dagger}\boldsymbol{B}\boldsymbol{r}$ 均为实数，从而 f 也为实数，证毕.

[17] $\boldsymbol{F}$ 的本征值为实数让方程组的求解大大简化，因为有现成的通解公式.

Chapter 2

变分法与 Hamilton 原理

2.1 变分法简介

2.1.1 泛函及其变分简介

我们知道，函数是一个数到数的映射. 如果不是把一个数而是一个别的什么东西映射为一个数，那么这个映射就叫**泛函**[1].

我们接下来需用到的泛函是定积分——比如定积分:

$$S(y) = \int_a^b L(y, y', x)dx \tag{2.1}$$

其中 L 是 y，y'，x 的函数，是将一个函数 $y(x)$ 映射为了一个数，因此是一个泛函，记为 $S(y)$. 下面我们提到泛函，总是默认其形式是如（2.1）一样.

什么是泛函的变分呢? 首先回顾一下函数的微分. 对于一个函数 $y = y(x)$，它的微分即:

$$dy = y(x + dx) - y(x)$$

那么，泛函的**变分**就定义为:

$$\delta S = S(y + \delta y) - S(y) \tag{2.2}$$

注意，这里的 $\delta y = \delta y(x)$ 是一个函数. 代入到 (2.1) 式的具体形式中，即:

$$\begin{aligned}
\delta S &= \int_a^b L[y(x) + \delta y(x), y'(x) + \delta y'(x), x]dx - \int_a^b L[y(x), y'(x), x]dx \\
&= \int_a^b \{L[y(x) + \delta y(x), y'(x) + \delta y'(x), x] - L[y(x), y'(x), x]\}\, dx \\
&= \int_a^b \left[\delta\boldsymbol{\xi} \cdot \nabla L + \frac{(\delta\boldsymbol{\xi} \cdot \nabla)^2}{2!} L + \cdots\right] dx
\end{aligned} \tag{2.3}$$

其中 $\delta\boldsymbol{\xi} = (\delta y, \delta y')$ 是为了美观我们引入的一个矢量记号。若忽略高阶小量项，则有:

$$\delta^1 S = \int_a^b (\delta\boldsymbol{\xi} \cdot \nabla L)dx = \int_a^b \left(\frac{\partial L}{\partial y}\delta y + \frac{\partial L}{\partial y'}\delta y'\right) dx \tag{2.4}$$

我们称（2.4）为 S 的**一阶变分**，很多时候直接称其为 S 的变分.

[1] 为什么它叫泛函? 我的理解是因为它是一种泛泛的函数.

2.1.2 泛函的极值问题与 Euler - Lagrange 方程

正如函数有极值问题一样，泛函也有其极值问题：**对于泛函 $S(y)$，试问对于初末值固定的 $y = y(x)$ 是何函数能使泛函 S 取极值?** 实际中的光的折射定律、最速降线问题、悬链线形状问题等均属于此类问题. 下面给出此类问题的基础，然后在（2.2）节中会具体解决这些问题.

要 $S(y)$ 在 y 处取极值，则当 y 有任意小变动 [2] δy 之后，S 的一阶变分等于 0[3]. 即

$$\delta^1 S = 0 \tag{2.5}$$

常记为 $\delta S = 0$，它就是整个变分法的基石，这和函数取极值时的判据是一样的. 于是现在，我们有：

$$\begin{aligned}
\delta S &= \int_a^b \left(\frac{\partial L}{\partial y}\delta y + \frac{\partial L}{\partial y'}\delta y' \right) dx \\
&= \int_a^b \frac{\partial L}{\partial y}\delta y dx + \int_a^b \frac{\partial L}{\partial y'} d(\delta y) \\
&= \int_a^b \frac{\partial L}{\partial y}\delta y dx + \left(\frac{\partial L}{\partial y'}\delta y \right) |_a^b - \int_a^b \frac{d}{dx}\left(\frac{\partial L}{\partial y'} \right) \delta y dx \\
&= -\int_a^b \left(\frac{d}{dx}\frac{\partial L}{\partial y'} - \frac{\partial L}{\partial y} \right) \delta y dx \\
&= 0
\end{aligned}$$

由于 $\delta y(x)$ 是一个任意的函数，所以我们得到了使得 $S(y)$ 取极值的函数应满足如下微分方程：

$$\frac{d}{dx}\frac{\partial L}{\partial y'} - \frac{\partial L}{\partial y} = 0 \tag{2.6}$$

这就是变分法中的 **Euler - Lagrange 方程**[4]. 若欲进一步判断 $S(y)$ 在 y 处是极大

[2] 由于我们一开始要求了 $y(x)$ 的初末值固定，所以这里的 δy 应满足 $\delta y(a) = \delta y(b) = 0$.

[3] 2026.1.29 晚，我和 WYB 同学讨论了这个问题，我们得到了一些心得，在此分享。WYB 同学是工科生，同时辅修数学学位，有数学分析背景和数学专业思维。他对作用量原理一直很兴趣，晚饭的时候我就给他讲什么是作用量原理，经过我们的探讨，得出结论：从数学的严谨的角度分析，真正能够导出 Lagrange 方程的条件应该是

$$\lim_{\delta x \to 0} \frac{S}{\int_a^b \delta y(t) dx} = 0$$

而我们这里给出的判据，是一种不够严谨的、受到物理学家甚至是理论物理学家欢迎的“物理数学”。

[4] 细心的读者应该已经看破天机了.

值还是极小值，则需要判断（2.3）式的高阶项（高阶变分）的正负.

2.1.3 有约束的泛函极值问题

正如有约束的函数极值问题一样，也存在有约束的泛函极值问题. 而它们的处理方式都是一样的，都是用 **Lagrange 乘数法**[5]. 下面我们首先回顾一下函数极值问题中的 Lagrange 乘数法.

设一个多元函数 $f(x_1, x_2, \cdots, x_\alpha)$，在约束条件 $\varphi(x_1, x_2, \cdots, x_\alpha) = 0$ 的限制下求其极值，怎么办? 想象让点 $(x_1, x_2, \cdots, x_\alpha)$ 在约束允许的范围内有一变动，从 $(x_1, x_2, \cdots, x_\alpha)$ 变到了 $(x_1 + dx_1, x_2 + dx_2, \cdots, x_\alpha + dx_\alpha)$，则约束要求由于点变动引起的 φ 的变动为 0，即:

$$
\begin{aligned}
d\varphi &= \varphi(x_1 + dx_1, x_2 + dx_2, \cdots, x_\alpha + dx_\alpha) - \varphi(x_1, x_2, \cdots, x_\alpha) \\
&= \sum_\alpha \frac{\partial \varphi}{\partial x_\alpha} dx_\alpha \\
&= 0
\end{aligned}
$$

上述点位置的变动也会引起 f 有一个变动 df，$(x_1, x_2, \cdots, x_\alpha)$ 是极值点要求:

$$
df = \sum_\alpha \frac{\partial f}{\partial x_\alpha} dx_\alpha = 0
$$

联立二式，可得在平稳点处须满足:

$$
\frac{\partial f / \partial x_1}{\partial \varphi / \partial x_1} = \frac{\partial f / \partial x_2}{\partial \varphi / \partial x_2} = \cdots = \frac{\partial f / \partial x_\alpha}{\partial \varphi / \partial x_\alpha}
$$

设它们都等于 $-\lambda$，于是你现在有 $d+1$ 个方程

$$
\begin{cases}
\frac{\partial f}{\partial x_i} + \lambda \frac{\partial \varphi}{\partial x_i} = 0, (i = 1, 2, \cdots, \alpha) \\
\varphi(x_1, x_2, \cdots, x_\alpha) = 0
\end{cases}
$$

与 $\alpha + 1$ 个未知数: $x_1, x_2, \cdots, x_\alpha, \lambda$. 可解出极值点，从而解出极值.

[5] 其实我们在（1.2.5）小节中已经用到这个方法了.

让我们少一些思维负担：上面说了这么多，其实只是在求

$$L(x_1, x_2, \cdots, x_\alpha, \lambda) = f(x_1, x_2, \cdots, x_\alpha) + \lambda\varphi(x_1, x_2, \cdots, x_\alpha) \tag{2.7}$$

这个 $\alpha + 1$ 维函数的极值点，不是吗？这就是 Lagrange 乘数法，其中 L 是你构造的 Lagrange 函数，λ 是 Lagrange 乘数.

下面将 Lagrange 乘数法推广至求泛函极值的问题中. 问题是：求泛函 $S(y)$ 在约束 $\Phi(y) = 0$ 下的极值（点）. 方法是：构造一个新泛函：

$$\tilde{S}(y) = S(y) + \lambda\Phi(y)$$

其中 λ 为 Lagrange 乘数. 所以问题归结为求 $\tilde{S}(y)$ 的无约束极值（点），这我们已经会求了，就是让 $\tilde{S}$ 的变分为 0. 即：

$$\delta\tilde{S} = 0 \tag{2.8}$$

2.2 变分法的简单应用 *

2.2.1 折射定律

设 、 介质的折射率为 n_1、n_2，A 点在 号介质中，B 点在 号介质中. 现已知在任何一种介质中光均只能走直线（作为一种约束），问光走何种路径从 A 到 B 的所用时间为最短[6]?

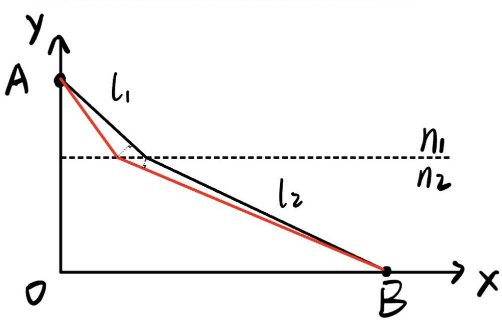

图 2.1: 折射定律问题中的变分

把折射率 n 视为 y 的函数，易知总时间:

$$t=\int_0^{x_B}\frac{ndl}{c}=\int_0^{x_B}\frac{n}{c}\sqrt{1+y'^2}dx$$

这里由于 n 不连续，故不能用 Euler - Lagrange 方程. 但是一阶变分 $\delta S=0$ 仍是成立的.

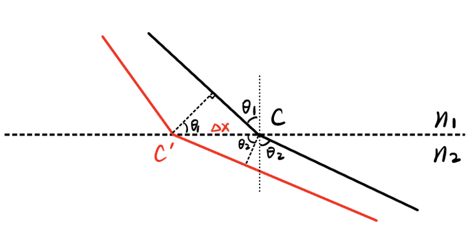

图 2.2: 放大图

设黑线是让时间最短的路径，让它有一个小的变动从而变为红线，且变动后在各

[6]耗时最短符合光学中的 Fermat 原理.

介质中仍是沿直线传播. 过 C 与 C' 点作垂线. 易知这一变动引起的时间 t 的变分为:

$$\delta t = \frac{n_2 \Delta x \sin\theta_2 - n_1 \Delta x \sin\theta_1}{c} = 0$$

于是得到折射定律:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{2.9}$$

2.2.2 最速降线

这是历史上很有名的问题. 现在竖直平面上有两个点 A 和 B，在 AB 之间修一条光滑轨道，求其形状 $y = y(x)$，使得小球无初速度地自 A 滑向 B 用时最短. 这条轨道的曲线就叫做**最速降线**.

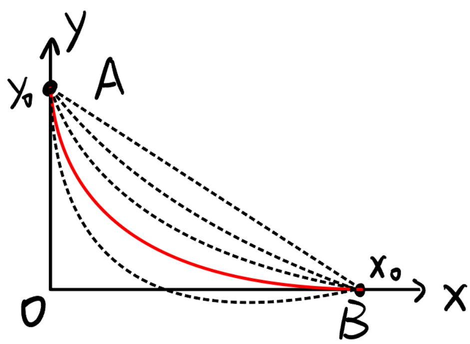

图 2.3: 哪一个轨道最快呢?

容易知道，时间为:

$$t = \int_0^{x_0} \frac{dl}{v} = \int_0^{x_0} \frac{dl}{\sqrt{2g(y_0 - y)}} = \int_0^{x} \frac{\sqrt{1 + y'^2}}{\sqrt{2g(y_0 - y)}} dx$$

于是记

$$L(y, y', x) = L(y, y') = \sqrt{\frac{1 + y'^2}{y_0 - y}}$$

先别着急代入 Euler - Lagrange 方程，我们可以进一步简化这个计算. 注意到 L 不

显含 x，于是有:

$$\begin{aligned}\frac{dL}{dx} &= \frac{\partial L}{\partial y}y' + \frac{\partial L}{\partial y'}y'' \\ &= \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right)y' + \frac{\partial L}{\partial y'}y'' \\ &= \frac{d}{dx}\left(y'\frac{\partial L}{\partial y'}\right)\end{aligned}$$

于是:

$$\frac{d}{dx}\left(y'\frac{\partial L}{\partial y'} - L\right) = 0$$

说明:

$$y'\frac{\partial L}{\partial y'} - L = \text{const} \tag{2.10}$$

这个结论很有用[7]. 于是:

$$\frac{y'}{\sqrt{y_0 - y}} \cdot \frac{y'}{\sqrt{1+y'^2}} - \sqrt{\frac{1+y'^2}{y_0 - y}} = C$$

故:

$$\frac{1}{\sqrt{1+y'^2}} = C\sqrt{y_0 - y}$$

记 $y' = \cot\theta$[8]，故 $y = y_0 - \frac{\sin^2\theta}{C^2} = y_0 + \frac{\cos 2\theta - 1}{2C^2}$，而:

$$y' = \frac{dy}{dx} = -\frac{2\sin\theta\cos\theta}{C^2}\frac{d\theta}{dx} = \cot\theta$$

故:

$$\frac{dx}{d\theta} = -\frac{2\sin^2\theta}{C^2} = \frac{\cos 2\theta - 1}{C^2}$$

故:

$$x = \frac{\sin 2\theta - 2\theta}{2C^2} + C'$$

[7]细心的读者应该已经看破天机了——这不就是第一章讲的能量积分吗?

[8]其实这里一劳永逸的方法是记 $y' = -\cot\theta$ ，这样一来就可以直接解出 $y(x)$ 参数方程的标准形式，而不需要像我下面那样去调参数了. 但是我仍然记 $y' = \cot\theta$ ，是想要表达：我们不管得到什么样的方程，是可以自己去调参数的，这并不会改变方程的物理意义. 参数不同只不过是坐标架的搭建和参数的选择稍微别扭一点而已.

于是我们得到 $y(x)$ 曲线的参数方程：

$$\begin{cases} y = y_0 - C(1-\cos\theta) \\ x = -C(\theta - \sin\theta) + C' \end{cases}$$

其中 θ 为参数，C、C' 为积分常数. 由曲线过 $(0, y_0)$ 得 $C' = 0$. 于是:

$$\begin{cases} y_0 - y = C(1-\cos\theta) \\ -x = C(\theta - \sin\theta) \end{cases}$$

再让 θ 变为 $-\theta$，得:

$$\begin{cases} y_0 - y = C(1-\cos\theta) \\ x = C(\theta - \sin\theta) \end{cases} \tag{2.11}$$

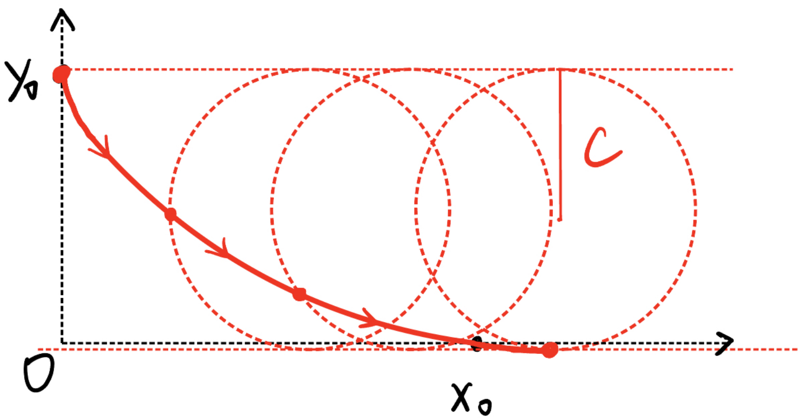

图 2.4: 轨道的形状是滚轮线

这明显是一段半径为 C 的轮子的**滚轮线**，且 $(0, y_0)$ 点为起点[9]. 欲得到 C，只需要把 $(x_0, 0)$ 代入，归结为解如下方程:

$$x_0 = C\left[\arccos\left(1 - \frac{y_0}{C}\right) - \sqrt{1 - \left(1 - \frac{y_0}{C}\right)^2}\right]$$

[9] $(0, y_0)$ 为起点不假，但是 $(x_0, 0)$ 不一定为终点，因为轮子的半径 C 不一定就等于 y_0.

解之即得 C.

最后提一个有趣的事，请读者对比图（2.1）与图（2.3），你有发现相似点吗？再想想 Fermat 原理，那么你能用折射定律找到最速降线的参数方程吗？

2.2.3 悬链线

这也是一个很有趣的问题，一根质量均匀的项链，两端分别悬挂于 A、B 两点，A、B 在同一竖直面但不一定等高，求平衡时项链的形状. 虚功原理告诉我们，平衡

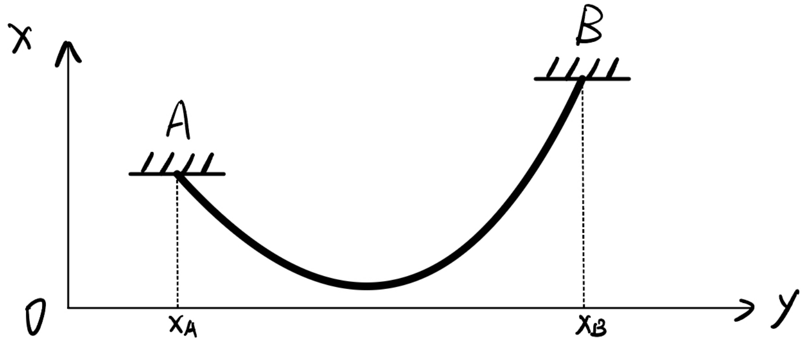

图 2.5: 悬挂着的匀质项链是什么形状呢?

时重力的虚功:

$$\delta W=\sum_i m_i g\cdot\delta\boldsymbol{r}_i=-\sum_i\delta V_i=-\delta V=0$$

因此在平衡时体系势能取极小值. 写出体系的势能:

$$\begin{aligned}V&=\int_{x_1}^{x_2}dmgy=\int_{x_1}^{x_2}\mu gydl\\&=\int_{x_1}^{x_2}\mu gy\sqrt{1+y'^2}dx\end{aligned}$$

其中 μ 是质量线密度. 此题还有一个约束，即链长一定，故有:

$$\int_{x_1}^{x_2}\sqrt{1+y'^2}dx=l$$

即

$$\int_{x_1}^{x_2}\left(\sqrt{1+y'^2}-\frac{l}{x_2-x_1}\right)dx=0$$

在该约束下求泛函 V 的极值点，需要我们构造新泛函：

$$\tilde{V}=\int_{x_1}^{x_2}\left[\mu g y\sqrt{1+y'^2}+\lambda\mu g\left(\sqrt{1+y'^2}-\frac{l}{x_2-x_1}\right)\right]dx\quad=\int_{x_1}^{x_2}L(y,y')dx$$

并令：

$$\delta\tilde{V}=0$$

于是得：

$$\frac{d}{dx}\frac{\partial L}{\partial y'}-\frac{\partial L}{\partial y}=0$$

由于 L 不显含 x，故：

$$y'\frac{\partial L}{\partial y'}-L=C$$

故：

$$(y+\lambda)\left(\frac{y'^2}{\sqrt{1+y'^2}}-\sqrt{1+y'^2}\right)=C$$

记 $\xi=y+\lambda$，于是 ξ 满足的微分方程为：

$$\xi\cdot\left(\frac{\xi'^2}{\sqrt{1+\xi'^2}}-\sqrt{1+\xi'^2}\right)=-\frac{\xi}{\sqrt{1+\xi'^2}}=C$$

从而：

$$\frac{d\xi}{dx}=\sqrt{\frac{\xi^2}{C^2}-1}$$

记 $\frac{\xi}{C}=\cosh\theta$，则：

$$\frac{d\xi}{dx}=\frac{dC\cosh\theta}{dx}=C\sinh\theta\frac{d\theta}{dx}=\sinh\theta$$

于是 $\theta=\frac{x+C'}{C}$，故 $\xi(x)=C\cosh\frac{x+C'}{C}$，从而

$$y(x)=C\cosh\frac{x+C'}{C}-\lambda\tag{2.12}$$

这就是**悬链线**的方程，其中 C、C'、λ 为积分常数. 由绳两端坐标定出 C' 与 λ，由绳长定出 C. 可以看到，悬链线为双曲余弦函数.

2.3 Hamilton 原理

2.3.1 Hamilton 原理

等待很久了.

相信读者早已经看到了数学中的 Euler-Lagrange 方程与分析力学中的 Lagrange 方程之间的联系了[10]. 它们的形式是一样的！那么 Lagrange 量 L 是否能和变分联系起来呢？是可以的，相应的泛函被称为作用量.

一般情况下，体系的 Lagrange 量 $L = L(q, \dot{q}, t)$ 是所有广义坐标、所有广义速度和时间的函数. 现在让体系在位形空间中从 t_1 时刻到 t_2 时刻演化，对应了位形空间中的一条超曲线. 这一运动的**作用量**定义为：

$$S(q_1, q_2, \cdots, q_n) = S(q) = \int_{t_1}^{t_2} L(q, \dot{q}, t)dt \tag{2.13}$$

与之前略有不同的是，泛函 $S(q_1, q_2, \cdots, q_n)$ 的宗量是 n 维而非 1 维的了. 但这并不造成原则性的困难，我们只需将各个 q_α 视为是独立改变的，然后分别研究即可.

Hamilton 原理指出：固定体系在位形空间中的两个时刻 t_1、t_2 时的初末状态（既包括坐标，也包括速度），则体系从 t_1 时状态向 t_2 时状态演化的实际方式是使得作用量取平稳值的方式. 下面让我们看看由 Hamilton 原理可以推得什

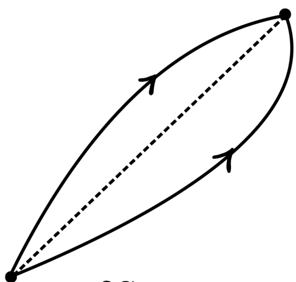

图 2.6: 作用量的一阶变分为 0 说明演化的超曲线有一个任意的变动所造成的作用量改变量的一阶近似为 0 ，这可以作为 Hamilton 原理的一个几何描述.

[10] 我在大二下学期的时候曾经问过田光善老师：Lagrange 当年得出他的方程的时候，难道没有看出和变分法的联系吗？为什么他没有得出作用量原理而是让 Hamilton 得出来了？田老师说 Lagrange 当时并不知道变分法，变分法是 Euler 的贡献，Lagrange 或许知道 Euler 这个人但并不知道变分那一套. 这是田老师的观点. 历史果然如此吗？有待进一步查明.

么.Hamilton 原理显然是在说对于 $\boldsymbol{q}(t) = (q_1(t), q_2(t), \cdots, q_n(t))$ 的一个微小变动 $\delta\boldsymbol{q}(t) = (\delta q_1(t), \delta q_2(t), \cdots, \delta q_n(t))$ 而言，引起 S 的（一阶）变分为 0. 即：

$$\delta S = \int_{t_1}^{t_2} \left(\sum_\alpha \frac{\partial L}{\partial q_\alpha} \delta q_\alpha + \frac{\partial L}{\partial \dot{q}_\alpha} \delta \dot{q}_\alpha \right) dt = 0 \tag{2.14}$$

利用前面几节的结论，易知这说明了[11]：

$$\delta S = \int_{t_1}^{t_2} \sum_\alpha \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_\alpha} - \frac{\partial L}{\partial q_\alpha} \right) \delta q_\alpha dt$$

由于各个 δq_α 是互相独立的，于是得到对任意广义坐标 q_α，有：

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_\alpha} - \frac{\partial L}{\partial q_\alpha} = 0$$

这正是 Lagrange 方程，不过这一次我们是利用新的 Hamilton 原理导出来的.

于是我们从 Hamilton 原理出发推出了 Lagrange 方程，由 Lagrange 方程可以进一步反推出 Newton 第二定律，于是 **Hamilton 原理可以作用第一性原理** [12]. 比如，我们如果把 Hamilton 原理当成第一原理，整个逻辑体系就变成了这样：系统的运动由 Lagrange 量 L 来描述，它是广义坐标、广义速度和时间的函数，它符合最小作用量原理，也就是说符合 Lagrange 方程. 至于 L 具体的形式是什么，要靠实验给出. 一旦你确定了 L 的具体形式，利用 Lagrange 方程就可以推得和 Newton 力学一样的结论了. Landau 的《力学》和 David Tong 的讲义都是以这种逻辑体系铺展的. 我们这一份讲义的逻辑体系实际上沿用了周衍柏老师的《理论力学教程（第五版）》的第五章，即从 Newton 到 Lagrange .

此外，还有很重要的一点；**体系的 L 并不唯一，而是可以相差任何一个关于时间的全导数项.** 因为添加任何一个时间的全导数项不会改变变分的极值点，从而也不

[11]我们还能够将这个变分以更紧凑的形式表出：

$$\delta S = \int_{t_1}^{t_2} \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{\boldsymbol{q}}} - \frac{\partial L}{\partial \boldsymbol{q}} \right) \cdot \delta \boldsymbol{q} dt = 0 \tag{2.15}$$

其中我们引入矢量记号：

$$\begin{cases} \delta \boldsymbol{q} = (\delta q_1, \delta q_2, \cdots, \delta q_n) \\ \frac{\partial L}{\partial \dot{\boldsymbol{q}}} = \left(\frac{\partial L}{\partial \dot{q}_1}, \frac{\partial L}{\partial \dot{q}_2}, \cdots, \frac{\partial L}{\partial \dot{q}_n} \right) \\ \frac{\partial L}{\partial \boldsymbol{q}} = \left(\frac{\partial L}{\partial q_1}, \frac{\partial L}{\partial q_2}, \cdots, \frac{\partial L}{\partial q_n} \right) \end{cases} \tag{2.16}$$

[12]但这有一些前提，比如说不能有耗散

会影响运动方程即 Lagrange 方程.

最后我们说明一下，(2.14) 式的名字有很多，除了我们所采用的"Hamilton 原理"，还有"最小作用量原理"、"作用量原理"、"平稳作用量原理"等等. 其中"最小作用量"原理用得最广泛，但是这一名字很受大家的不满，原因是人们认为作用量的一阶变分为 0 并不能够说明对应的是最小值. 而就我所知道的资料中，只有 Landau 的《力学》明确提出了作用量应该取"最小值"这一结论，而且梁昆淼老师的《力学（下册）》也对作用量原理的称谓做了较详细的区分和说明，感兴趣的读者可以去瞧一瞧. 今后在不至于混淆的情况下，我们还是采用"Hamilton 原理"的称谓.

2.3.2 Norther 定理（质点动力学）

有了作用量的概念，下面简介十分有名的 Noether 定理，它将对称性和守恒量联系了起来，是近代物理学的明星定理。但是需要注意的是，该定理有针对质点动力学的版本，还有关于场的版本我们在后面的章节会介绍，本节内容我们只涉及了质点动力学的版本。

首先，什么是对称性? 数学物理中的对称性有别于我们日常生活中所使用的对称性（不过它们本质上是趋同的）. 数理中体系的对称性是指受到某种操作后仍保持不变的性质[13]。具体到 Lagrange 力学的框架之下，**体系的对某种操作有对称性即是说在该种操作下，运动方程不变**。而要想运动方程不变，我们只要求 L 在该操作下不变或者添加一个关于时间的全导数项就可以了，即是说 $L(q,\dot{q},t) \to L + \delta L = L + \frac{d}{dt}\chi$，其中 χ 是某个广义坐标、广义速度和时间的函数. 其次，守恒量是指体系中不随时间改变的物理量，这不陌生.

Noether 定理：体系每一个连续的对称性都对应一个守恒量，反之亦然. 下面我们来证明它。

- 首先，若我们仅仅对空间广义坐标施加变换，而不涉及时间平移

$$q_\alpha \to q_\alpha + \delta q_\alpha$$

[13] 接下来我们将只考虑连续的操作与连续的对称性

则首先，一定有

$$
\begin{aligned}
\delta L &= \sum_\alpha \left(\frac{\partial L}{\partial q_\alpha} \delta q_\alpha + \frac{\partial L}{\partial \dot{q}_\alpha} \delta \dot{q}_\alpha \right) \\
&= \sum_\alpha \left[\frac{\partial L}{\partial q_\alpha} \delta q_\alpha + \frac{\partial L}{\partial \dot{q}_\alpha} \frac{d}{dt} (\delta q_\alpha) \right] \\
&= \sum_\alpha \left(\frac{\partial L}{\partial q_\alpha} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_\alpha} \right) \delta q_\alpha + \sum_\alpha \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\alpha} \delta q_\alpha \right) \\
&= \sum_\alpha \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\alpha} \delta q_\alpha \right)
\end{aligned}
$$

另一方面，对称变换要求我们能够且必须找到这样的形式 $\delta L = \frac{d\chi}{dt}$，从而我们得到

$$
\frac{d}{dt} \left(\sum_\alpha \frac{\partial L}{\partial \dot{q}_\alpha} \delta q_\alpha - \chi \right) = 0
$$

故我们不但证明了存在一个守恒量，甚至还直接找到了守恒量的具体形式：$\xi = \sum_\alpha \frac{\partial L}{\partial \dot{q}_\alpha} \delta q_\alpha - \chi$.

- 其次，如果我们施加的是时间的平移变换 $t \to t + \delta t$ (δt 为小量) , 首先一定有 $\delta L = \frac{dL}{dt} \delta t$ 另一方面，我们有

$$
\begin{aligned}
\delta L &= \frac{\partial L}{\partial q} \dot{q} \delta t + \frac{\partial L}{\partial \dot{q}} \ddot{q} \delta t + \frac{\partial L}{\partial t} \delta t \\
&= \left[\frac{\partial L}{\partial q} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) \right] \dot{q} \delta t + \frac{d}{dt} \left(\dot{q} \frac{\partial L}{\partial \dot{q}} \right) + \frac{\partial L}{\partial t} \delta t \\
&= \frac{d}{dt} \left(\dot{q} \frac{\partial L}{\partial \dot{q}} \right) \delta t + \frac{\partial L}{\partial t} \delta t
\end{aligned}
$$

时间平移对称要求 Lagrange 量的改变为全导数，那么只能是让 $\frac{\partial L}{\partial t} = 0$，从而 $\delta L = \frac{d}{dt} \left(\dot{q} \frac{\partial L}{\partial \dot{q}} \right) \delta t$，因此我们得到

$$
\frac{d}{dt} \left(\dot{q} \frac{\partial L}{\partial \dot{q}} - L \right) = 0
$$

所以 $\dot{q}\frac{\partial L}{\partial \dot{q}} - L$ 是守恒量，这就是体系的 Hamilton 量。

至此 Norther 定理证毕。我们下面通过举三个最典型的例子来说明。

空间平移对称性

下面我们用直角坐标，且只考虑体系的 x 自由度。假设体系对于空间平移操作 $x \to x + \epsilon \quad$ (ϵ 为小量) 是对称的, 则 L 在该操作下的改变只能是

$$L(x, \dot{x}, t) \to L(x + \epsilon, \dot{x}, t) = L$$

这意味着 $\chi = 0$(同时也意味着 $\frac{\partial L}{\partial x} = 0$, 即 x 是循环坐标). 此时, 根据上面得到的公式，动量 $p_x = \frac{\partial L}{\partial \dot{x}}$ 为守恒量. 因此空间平移对称性对应着动量守恒.

空间旋转对称性

我们只需将广义坐标从 x 换成角度 θ, 重复上述推论即可推知: 若体系具有空间旋转对称性, 则角动量 $p_\theta = \frac{\partial L}{\partial \dot{\theta}}$ 为守恒量. 因此空间旋转对称性对应着角动量守恒.

时间平移对称性

在证明 Norther 定理的时候，我们已经证明了在此种情况下，Hamilton 量（能量）为守恒量。因此时间平移对称性对应着能量守恒。

Chapter 3

Hamilton 力学

3.1 Hamilton 正则方程

3.1.1 Hamilton 量

我们已经定义了一个力学体系的 Lagrange 量，以此出发我们可以进一步定义一个体系的 **Hamilton** 量：

$$H=\sum_{\alpha}\dot{q}_{\alpha}\frac{\partial L}{\partial\dot{q}_{\alpha}}-L \tag{3.1}$$

我们可以将 H 视为 $(q,\dot{q},t)$ 的函数，但我们不这么做，而是视之为 (q,p,t) 的函数，其中：

$$p_{\alpha}=\frac{\partial L}{\partial\dot{q}_{\alpha}}$$

为广义动量，但是在 Hamilton 力学的框架下，我们称之为**正则动量**，同时称广义坐标为 **正则坐标**，二者的地位是等价的，都作为 Hamilton 量的宗量出现[1]。

要想找 H 对 (q,p,t) 的显式的步骤如下：

- 写出体系的 Lagrange 量 $L(q,\dot{q},t)$
- 将 $\dot{q}_{\alpha}$ 用 $q,\frac{\partial L}{\partial\dot{q}},t$ 表示，即找到 $\dot{q}(q,p,t)$
- 代入定义式，即得：

$$H(q,p,t)=\sum_{\alpha}\dot{q}_{\alpha}(q,p,t)p_{\alpha}-L\left(q,\dot{q}(q,p,t),t\right)$$

除此之外，由第一章可知：**若 L 不显含 t，则 $H=T+V$ 为体系总能**.

[1]这里有必要强调一点——在下面的讨论中，我们实际上默认了各个正则坐标和各个正则动量之间都是互相独立的。这一点对于大部分力学体系而言的确是正确的，但是严格来讲并不是普遍成立的。对于 Lagrange 量而言，各个广义坐标（从而广义速度）之间的确是相互独立的，然而我们却无法总是保证利用 Lagrange 量得出的广义动量也是彼此之间相互独立的。这里不妨举一个例子，比如考虑如下形式的 Lagrange 量

$$L=(\dot{x}-\dot{y})^2$$

其中 x 与 y 是相互独立的，但是经过计算不难发现，这个体系的广义动量不是互相独立的，它们互为相反数。

3.1.2 Hamilton 正则方程

下面我们导出 Hamilton 正则方程. 由 Hamilton 量的定义可知:

$$\begin{aligned}
dH &= \sum_\alpha d\dot{q}_\alpha p_\alpha + \sum_\alpha \dot{q}_\alpha dp_\alpha - dL \\
&= \sum_\alpha \frac{\partial L}{\partial \dot{q}_\alpha} d\dot{q}_\alpha + \sum_\alpha \dot{q}_\alpha dp_\alpha - \sum_\alpha \frac{\partial L}{\partial q_\alpha} dq_\alpha - \sum_\alpha \frac{\partial L}{\partial \dot{q}_\alpha} d\dot{q}_\alpha - \frac{\partial L}{\partial t} dt \\
&= \sum_\alpha \dot{q}_\alpha dp_\alpha - \sum_\alpha \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\alpha}\right) dq_\alpha - \frac{\partial L}{\partial t} dt \\
&= \sum_\alpha \dot{q}_\alpha dp_\alpha + \sum_\alpha (-\dot{p}_\alpha) dq_\alpha + \left(-\frac{\partial L}{\partial t}\right) dt
\end{aligned}$$

于是，我们得到:

$$\begin{cases}
\dfrac{\partial H}{\partial q_\alpha} = -\dot{p}_\alpha \\[2ex]
\dfrac{\partial H}{\partial p_\alpha} = \dot{q}_\alpha
\end{cases} \tag{3.2}$$

这就是 **Hamilton 正则方程**. 此外，我们还有:

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \tag{3.3}$$

这也表明若 L 不显含 t，则 H 亦不显含 t.

下面我们举几个简单的实例来进行分析:

一维谐振子

首先我们来构建谐振子的 Hamilton 量。已知一维谐振子的 Lagrange 量为 $L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}\omega^2 x^2$，而广义动量为 $p = \frac{\partial L}{\partial \dot{x}} = m\dot{x}$，从而有 $\dot{x} = \frac{p}{m}$。从而，Hamilton 量为

$$\begin{aligned}
H &= p\dot{x} - L = p\dot{x} - \frac{1}{2}m\dot{x}^2 + \frac{1}{2}\omega^2 x^2 \\
&= \frac{p^2}{m} - \frac{m}{2}\left(\frac{p}{m}\right)^2 + \frac{1}{2}\omega^2 x^2 \\
&= \frac{p^2}{2m} + \frac{1}{2}\omega^2 x^2
\end{aligned} \tag{3.4}$$

我们还可以写出正则方程：

$$\begin{cases}\dot{x}=\frac{\partial H}{\partial p}=\frac{p}{m}\\ \dot{p}=-\frac{\partial H}{\partial x}=-\omega^2 x\end{cases}$$

上面第一式为动量之定义，第二式即 Hooke 定律。

电磁场中的带电粒子

下面我们来看另一个例子，电磁场中的带电粒子的 Hamilton 量。我们已知其 Lagrange 量为（仍假定取直角坐标）

$$L=\frac{1}{2}m\dot{\boldsymbol{r}}^2-q\varphi+q\boldsymbol{A}\cdot\dot{\boldsymbol{r}}$$

广义动量为 $\boldsymbol{p}=\frac{\partial L}{\partial\dot{\boldsymbol{r}}}=m\dot{\boldsymbol{r}}+q\boldsymbol{A}$，从而我们可以用 $\boldsymbol{p}$ 与 $\boldsymbol{r}$ 表示速度为 $\dot{\boldsymbol{r}}=\frac{\boldsymbol{p}-q\boldsymbol{A}}{m}$ 进而得到 Hamilton 量为

$$\begin{aligned}H&=\boldsymbol{p}\cdot\dot{\boldsymbol{r}}-L\\&=\boldsymbol{p}\cdot\frac{\boldsymbol{p}-q\boldsymbol{A}}{m}-\frac{1}{2}m\left(\frac{\boldsymbol{p}-q\boldsymbol{A}}{m}\right)^2+q\varphi-q\boldsymbol{A}\cdot\frac{\boldsymbol{p}-q\boldsymbol{A}}{m}\\&=\frac{(\boldsymbol{p}-q\boldsymbol{A})^2}{2m}+q\varphi\end{aligned}\tag{3.5}$$

进一步，我们可以得到正则方程：

$$\dot{\boldsymbol{r}}=\frac{\partial H}{\partial\boldsymbol{p}}=\frac{\boldsymbol{p}-q\boldsymbol{A}}{m}$$

与[2]

$$\dot{\boldsymbol{p}}=-\nabla H=\frac{1}{m}(\boldsymbol{p}-q\boldsymbol{A})\times(q\nabla\times\boldsymbol{A})+\frac{1}{m}[(\boldsymbol{p}-q\boldsymbol{A})\cdot\nabla](q\boldsymbol{A})-q\nabla\varphi$$

[2]其中要用到矢量分析中的恒等式：

$$\nabla(\boldsymbol{A}\cdot\boldsymbol{B})=\boldsymbol{A}\times(\nabla\times\boldsymbol{B})+\boldsymbol{B}\times(\nabla\times\boldsymbol{A})+(\boldsymbol{B}\cdot\nabla)\boldsymbol{A}+(\boldsymbol{A}\cdot\nabla)\boldsymbol{B}$$

从而

$$\nabla(\boldsymbol{A}\cdot\boldsymbol{A})=2\boldsymbol{A}\times(\nabla\times\boldsymbol{A})+2(\boldsymbol{A}\cdot\nabla)\boldsymbol{A}$$

即

$$m\ddot{\boldsymbol{r}}+q\frac{d\boldsymbol{A}}{dt}=q\dot{\boldsymbol{r}}\times(\nabla\times\boldsymbol{A})+q(\dot{\boldsymbol{r}}\cdot\nabla)\boldsymbol{A}-q\nabla\varphi$$

故我们能得到 Lorentz 力公式

$$\begin{aligned}
m\ddot{\boldsymbol{r}}&=q\dot{\boldsymbol{r}}\times(\nabla\times\boldsymbol{A})-\left[q\frac{\partial\boldsymbol{A}}{\partial t}+q(\dot{\boldsymbol{r}}\cdot\nabla)\boldsymbol{A}\right]+q(\dot{\boldsymbol{r}}\cdot\nabla)\boldsymbol{A}-q\nabla\varphi\\
&=q\dot{\boldsymbol{r}}\times(\nabla\times\boldsymbol{A})-q\nabla\varphi-q\frac{\partial\boldsymbol{A}}{\partial t}\\
&=q\dot{\boldsymbol{r}}\times\boldsymbol{B}-q\boldsymbol{E}\\
&=F_{\text{Lorentz}}
\end{aligned}$$

3.1.3 Hamilton 量版本的 Hamilton 原理

用 Hamilton 量定义一个运动过程的作用量为:

$$S(q,p)=\int_{t_1}^{t_2}\left[\dot{q}p-H(q,p,t)\right]dt \tag{3.6}$$

然后我们定义一个有别于位形空间的新空间——**相空间**，它是由 $(q_1,q_2,\cdots,q_n,p_1,p_2,\cdots,p_n)$ $2n$ 个坐标张成的 $2n$ 维空间，显然相空间里的点表征了体系在某个时刻的运动状态，而相空间中的一条超曲线表征了体系在实际物理空间中的一个运动.

Hamilton 量版本的 Hamilton 原理指出：**体系在相空间中两点之间的真实演化一定使得其作用量取平稳值.** 于是，S 的（一阶）变分为 0，即:

$$\begin{aligned}
\delta S&=\sum_\alpha\int_{t_1}^{t_2}\left(p_\alpha\delta\dot{q}_\alpha+\dot{q}_\alpha\delta p_\alpha-\frac{\partial H}{\partial q_\alpha}\delta q_\alpha-\frac{\partial H}{\partial p_\alpha}\delta p_\alpha\right)dt\\
&=\sum_\alpha\int_{t_1}^{t_2}p_\alpha d(\delta q_\alpha)+\sum_\alpha\int_{t_1}^{t_2}\left(\dot{q}_\alpha\delta p_\alpha-\frac{\partial H}{\partial q_\alpha}\delta q_\alpha-\frac{\partial H}{\partial p_\alpha}\delta p_\alpha\right)dt\\
&=\sum_\alpha p_\alpha\delta q_\alpha\Big|_{t_1}^{t_2}+\sum_\alpha\int_{t_1}^{t_2}\left(-\dot{p}_\alpha\delta q_\alpha+\dot{q}_\alpha\delta p_\alpha-\frac{\partial H}{\partial q_\alpha}\delta q_\alpha-\frac{\partial H}{\partial p_\alpha}\delta p_\alpha\right)dt\\
&=\sum_\alpha\int_{t_1}^{t_2}\left(\dot{q}_\alpha-\sum_\alpha\frac{\partial H}{\partial p_\alpha}\right)\delta p_\alpha dt-\int_{t_1}^{t_2}\left(\dot{p}_\alpha+\frac{\partial H}{\partial q_\alpha}\right)\delta q_\alpha dt\\
&=0
\end{aligned}$$

由于 δp_α 与 δq_α 的任意性，故：

$$\begin{cases} \dfrac{\partial H}{\partial p_\alpha} = \dot{q}_\alpha \\ \\ \dfrac{\partial H}{\partial q_\alpha} = -\dot{p}_\alpha \end{cases}$$

即我们利用 Hamilton 原理推出了正则方程.

我们最后说明一下：为什么这里对 q 与 p 同时进行了变分而在 Lagrange 量的版本只对 q（而未对 $\dot{q}$）进行变分呢？其原因很简单——**因为 Hamilton 原理指定的空间不同！** 当时 Lagrange 版本选用位形空间，坐标只有 q；这里选用相空间，坐标有 q, p. 变分变的是所选空间中的全部坐标.

3.2 Poisson 括号

3.2.1 Poisson 括号及其性质

设一个物理量 $f = f(q, p, t)$，来看其随时间的演化：

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_{\alpha} \frac{\partial f}{\partial q_\alpha}\dot{q}_\alpha + \frac{\partial f}{\partial p_\alpha}\dot{p}_\alpha$$

再代入 Hamilton 正则方程，得：

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_{\alpha} \frac{\partial f}{\partial q_\alpha}\frac{\partial H}{\partial p_\alpha} - \frac{\partial f}{\partial p_\alpha}\frac{\partial H}{\partial q_\alpha} = \frac{\partial f}{\partial t} + \{f, H\} \tag{3.7}$$

这里我们就定义了 **Poisson 括号**，对于两个一般的函数 f, g，其 Poisson 括号为：

$$\{f, g\} = \sum_{\alpha} \frac{\partial f}{\partial q_\alpha}\frac{\partial g}{\partial p_\alpha} - \frac{\partial f}{\partial p_\alpha}\frac{\partial g}{\partial q_\alpha} = \sum_{\alpha} \frac{\partial(f, g)}{\partial(q_\alpha, p_\alpha)} \tag{3.8}$$

这里我们用了 Jacobi 行列式的记号：

$$\frac{\partial(f, g)}{\partial(q_\alpha, p_\alpha)} = \begin{vmatrix} \frac{\partial f}{\partial q_\alpha} & \frac{\partial g}{\partial q_\alpha} \\ \frac{\partial f}{\partial p_\alpha} & \frac{\partial g}{\partial p_\alpha} \end{vmatrix} \tag{3.9}$$

于是，行列式的线性的代数性质同样适用于 Poisson 括号：

- $\{f, f\} = 0$
- $\{f, g\} = -\{g, f\}$
- $\{f, \alpha g + \beta h\} = \alpha\{f, g\} + \beta\{f, h\}$

此外，不难证明，还有一些微分运算性质：

- $\{f, g\}' = \{f', g\} + \{f, g'\}$
- $\{fg, h\} = f\{g, h\} + g\{f, h\}$
- $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$

基本 Poisson 括号

不难证明我们有如下基本 Poisson 括号：

$$\{q_\alpha, q_\beta\} = \{p_\alpha, p_\beta\} = 0 \quad , \quad \{q_\alpha, p_\beta\} = \delta_{\alpha\beta} \tag{3.10}$$

此外，利用上述 Poisson 括号的代数性质不难证明，如果取广义坐标为直角坐标，则角动量 $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$ 及其各分量之间的 Poisson 括号为：

$$\{L_i, L_j\} = \varepsilon_{ijk} L_k \quad , \quad \{L_i, L^2\} = 0 \tag{3.11}$$

这里 $i, j, k = 1, 2, 3$。

3.2.2 守恒量

设两个力学量 $f = f(q, p)$，$g = g(q, p)$ 均不显含 t，且有 $\frac{df}{dt} = \frac{dg}{dt} = 0$，即二者皆为守恒量，于是有：

$$\{f, H\} = 0, \quad \{g, H\} = 0$$

这说明，**守恒量与 Hamilton 量的 Poisson 括号为零**。利用这一性质，结合基本 Poisson 括号，我们可以得出类似于 Lagrange 力学里的结论：**如果 Hamilton 量中不含有某个正则坐标，则与其相共轭的正则动量为守恒量**[3]，**因为该正则动量一定和 Hamilton 量的 Poisson 括号为零**。进一步我们还可以得到

$$\begin{aligned}\{\{f, g\}, H\} &= -\{H, \{f, g\}\} = \{f, \{g, H\}\} + \{g, \{H, f\}\} \\ &= \{f, 0\} + \{g, 0\} \\ &= 0\end{aligned}$$

这说明：

$$\frac{d}{dt}\{f, g\} = 0$$

这就是 **Poisson 定理**，它指出**两个守恒量的 Poisson 括号亦是守恒量.**

[3] 其实，这一点从正则方程本身亦可以直接得到。

3.3 正则变换

3.3.1 正则变换的母函数

正如在 Lagrange 力学中可以进行广义坐标的变换一样，Hamilton 力学中有所谓**正则变换**，其最一般的形式为：

$$\begin{cases} \mathcal{Q} = \mathcal{Q}(q, p, t) \\ \mathcal{P} = \mathcal{P}(q, p, t) \end{cases} \tag{3.12}$$

如果说广义坐标变换是某时刻从一个位形空间到另一个位形空间的变换，那么正则变换则是（某时刻）从一个相空间 (q, p) 到另一个相空间 $(\mathcal{Q}, \mathcal{P})$ 的变换。然而正则变换并不及广义坐标变换那样可以很随意地映射——正则变换冠以"正则"二字是有原因的，因为 **正则变换之后的正则变量 $\mathcal{Q}, \mathcal{P}$ 必须也满足正则方程**：

$$\dot{\mathcal{Q}} = \frac{\partial K}{\partial \mathcal{P}}, \quad \dot{P} = -\frac{\partial K}{\partial \mathcal{Q}} \tag{3.13}$$

其中 $K(\mathcal{Q}, \mathcal{P}, t) = H\left[q(\mathcal{Q}, \mathcal{P}, t), p(\mathcal{Q}, \mathcal{P}, t), t\right]$ 为用新的正则变量 $\mathcal{Q}, \mathcal{P}$ 表示的 Hamilton 量.

既然 $\mathcal{Q}, \mathcal{P}$ 亦满足正则方程，那么其对应的 Hamilton 原理：

$$\delta S'(\mathcal{Q}, \mathcal{P}, t) = \delta \int_{t_1}^{t_2} \left[\sum_\alpha \dot{\mathcal{Q}}_\alpha \mathcal{P}_\alpha - K(\mathcal{Q}, \mathcal{P}, t)\right] dt = 0$$

亦成立. 于是 L' 与 L 之间至多相差一个时间的全导数项，设其为：

$$\frac{dU}{dt} = L - L' = \sum_\alpha \left(\dot{q}_\alpha p_\alpha - H - \dot{\mathcal{Q}}_\alpha \mathcal{P}_\alpha + K\right) \tag{3.14}$$

从而我们能得到关于 U 的全微分

$$dU = \sum_\alpha \left[p_\alpha dq_\alpha - \mathcal{P}_\alpha d\mathcal{Q}_\alpha + (K - H)dt\right]$$

若我们认为 $q_\alpha, \mathcal{Q}_\alpha, t$ 为独立变量，即认为 $U = U_1 = U_1(\mathcal{Q}, q, t)$ 是 $\mathcal{Q}, q, t$ 的函数，则

上面的全微分式指出：

$$\begin{cases} \mathcal{P}_\alpha = -\frac{\partial U_1}{\partial \mathcal{Q}_\alpha} \\ p_\alpha = \frac{\partial U_1}{\partial q_\alpha} \\ K - H = \frac{\partial U_1}{\partial t} \end{cases} \tag{3.15}$$

对于一个给定了具体形式的 $U_1(\mathcal{Q}, q, t)$，利用（3.15）中的前两个式子可求出用 $\mathcal{Q}, q, t$ 表出的 $\mathcal{P}_\alpha, p_\alpha$，再从中反解出用 q, p, t 表出的 $\mathcal{Q}$ 与 $\mathcal{P}$，就能得到正则变换的形式了. **因此，我们说（3.15）与（3.12）是等价的——我们进而可以说 $U_1(\mathcal{Q}, q, t)$ 的一个具体形式决定了一个具体的正则变换，故称 U_1 为正则变换的（第一类）母函数.**

母函数产生正则变换，除了以 $\mathcal{Q}, q, t$ 为宗量的 U_1，母函数还可以以其它正则变量为宗量，分别对应了不同类型的母函数，它们是：$U_2(\mathcal{Q}, p, t)$，$U_2(\mathcal{P}, q, t)$，$U_4(\mathcal{P}, p, t)$. 从数学的角度来看，它们由 U_1 经 Legendre 变换生成。下面我们只看第二类母函数（因为第二类母函数在后文经常出现），其余种类的母函数不难以此类推。记第二类母函数为

$$U_2(q, \mathcal{P}, t) = U_1\left(\mathcal{Q}(q, \mathcal{P}, t), q, t\right) + \sum_\alpha \mathcal{Q}_\alpha(q, \mathcal{P}, t)\mathcal{P}_\alpha \tag{3.16}$$

取全微分，得

$$\begin{aligned} dU_2 &= dU_1 + \sum_\alpha \mathcal{P}_\alpha d\mathcal{Q}_\alpha + \sum_\alpha \mathcal{Q}_\alpha d\mathcal{P}_\alpha \\ &= \sum_\alpha -\mathcal{P}_\alpha d\mathcal{Q}_\alpha + \sum_\alpha p_\alpha dq_\alpha + \sum_\alpha \mathcal{P}_\alpha d\mathcal{Q}_\alpha + \sum_\alpha \mathcal{Q}_\alpha d\mathcal{P}_\alpha + (K-H)dt \\ &= \sum_\alpha (p_\alpha dq_\alpha + \mathcal{Q}_\alpha d\mathcal{P}_\alpha) + (K-H)dt \end{aligned}$$

故我们得到坐标变换关系

$$\begin{cases} p_\alpha = \frac{\partial U_2}{\partial q_\alpha} \\ \mathcal{Q}_\alpha = \frac{\partial U_2}{\partial \mathcal{P}_\alpha} \\ K - H = \frac{\partial U_2}{\partial t} \end{cases} \tag{3.17}$$

3.3.2 Poisson 括号的正则变换不变性

记 $f=f(q,p,t)$，$g=g(q,p,t)$. 正则变换下，二者变为 $f=f(\mathcal{Q},\mathcal{P},t)$，$g=g(\mathcal{Q},\mathcal{P},t)$. 下面我们证明如下关系:

$$\{f,g\}_{qp}=\sum_{\alpha}\frac{\partial(f,g)}{\partial(q_\alpha,p_\alpha)}=\sum_{\alpha}\frac{\partial(f,g)}{\partial(\mathcal{Q}_\alpha,\mathcal{P}_\alpha)}=\{f,g\}_{\mathcal{Q}\mathcal{P}}$$

即 **Poisson 括号具有正则变换下的不变性**。设 $U_2(q,\mathcal{P},t)$ 为母函数，生成正则变换:

$$p_\alpha=\frac{\partial U_2}{\partial q_\alpha},\quad \mathcal{Q}_\alpha=\frac{\partial U_2}{\partial \mathcal{P}_\alpha}$$

接下来通通把 $q_\alpha,\mathcal{P}_\alpha$ 视为 $\mathcal{Q}_\alpha,p_\alpha$ 的函数，这样做是为了在过程中引入 U_2 的二阶偏导数，这是关键一招. 于是:

$$\begin{aligned}
\{f,g\}_{qp}&=\sum_{\alpha}\frac{\partial f}{\partial q_\alpha}\frac{\partial g}{\partial p_\alpha}-\sum_{\alpha}\frac{\partial f}{\partial p_\alpha}\frac{\partial g}{\partial q_\alpha}\\
&=\sum_{\alpha}\sum_{\beta}\left(\frac{\partial f}{\partial p_\beta}\frac{\partial p_\beta}{\partial q_\alpha}+\frac{\partial f}{\partial \mathcal{Q}_\beta}\frac{\partial \mathcal{Q}_\beta}{\partial q_\alpha}\right)\frac{\partial g}{\partial p_\alpha}-\sum_{\alpha}\sum_{\beta}\frac{\partial f}{\partial p_\alpha}\left(\frac{\partial g}{\partial p_\beta}\frac{\partial p_\beta}{\partial q_\alpha}+\frac{\partial g}{\partial \mathcal{Q}_\beta}\frac{\partial \mathcal{Q}_\beta}{\partial q_\alpha}\right)\\
&=\sum_{\alpha}\sum_{\beta}\left(\frac{\partial f}{\partial p_\beta}\frac{\partial^2 U_2}{\partial q_\alpha\partial q_\beta}+\frac{\partial f}{\partial \mathcal{Q}_\beta}\frac{\partial^2 U_2}{\partial q_\alpha\partial \mathcal{P}_\beta}\right)\frac{\partial g}{\partial p_\alpha}\\
&-\sum_{\alpha}\sum_{\beta}\frac{\partial f}{\partial p_\alpha}\left(\frac{\partial g}{\partial p_\beta}\frac{\partial^2 U_2}{\partial q_\alpha\partial q_\beta}+\frac{\partial g}{\partial \mathcal{Q}_\beta}\frac{\partial^2 U_2}{\partial q_\alpha\partial \mathcal{P}_\beta}\right)\\
&=\sum_{\alpha}\sum_{\beta}\left(\frac{\partial f}{\partial p_\beta}\frac{\partial g}{\partial p_\alpha}-\frac{\partial g}{\partial p_\beta}\frac{\partial f}{\partial p_\alpha}\right)\frac{\partial^2 U_2}{\partial q_\alpha\partial q_\beta}\\
&+\sum_{\alpha}\sum_{\beta}\left(\frac{\partial f}{\partial \mathcal{Q}_\beta}\frac{\partial g}{\partial p_\alpha}-\frac{\partial g}{\partial \mathcal{Q}_\beta}\frac{\partial f}{\partial p_\alpha}\right)\frac{\partial^2 U_2}{\partial q_\alpha\partial \mathcal{P}_\beta}\\
&=\sum_{\alpha}\sum_{\beta}\left(\frac{\partial f}{\partial \mathcal{Q}_\beta}\frac{\partial g}{\partial p_\alpha}-\frac{\partial g}{\partial \mathcal{Q}_\beta}\frac{\partial f}{\partial p_\alpha}\right)\frac{\partial^2 U_2}{\partial q_\alpha\partial \mathcal{P}_\beta}
\end{aligned}$$

而:

$$\begin{aligned}
\{f,g\}_{\mathcal{Q}\xi}&=\sum_{\alpha}\frac{\partial f}{\partial \mathcal{Q}_\alpha}\frac{\partial g}{\partial \mathcal{P}_\alpha}-\sum_{\alpha}\frac{\partial f}{\partial \mathcal{P}_\alpha}\frac{\partial g}{\partial \mathcal{Q}_\alpha}\\
&=\sum_{\alpha}\sum_{\beta}\frac{\partial f}{\partial \mathcal{Q}_\alpha}\left(\frac{\partial g}{\partial \mathcal{Q}_\beta}\frac{\partial \mathcal{Q}_\beta}{\partial \mathcal{P}_\alpha}+\frac{\partial g}{\partial p_\beta}\frac{\partial p_\beta}{\partial \mathcal{P}_\alpha}\right)-\sum_{\alpha}\sum_{\beta}\left(\frac{\partial f}{\partial \mathcal{Q}_\beta}\frac{\partial \mathcal{Q}_\beta}{\partial \mathcal{P}_\alpha}+\frac{\partial f}{\partial p_\beta}\frac{\partial p_\beta}{\partial \mathcal{P}_\alpha}\right)\frac{\partial g}{\partial \mathcal{Q}_\alpha}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_\alpha \sum_\beta \frac{\partial f}{\partial \mathcal{Q}_\alpha} \left(\frac{\partial g}{\partial \mathcal{Q}_\beta} \frac{\partial^2 U_2}{\partial \mathcal{P}_\alpha \partial \mathcal{P}_\beta} + \frac{\partial g}{\partial p_\beta} \frac{\partial^2 U_2}{\partial \mathcal{P}_\alpha \partial q^\beta} \right) \\
&- \sum_\alpha \sum_\beta \left(\frac{\partial f}{\partial \mathcal{Q}_\beta} \frac{\partial^2 U_2}{\partial \mathcal{P}_\alpha \partial \mathcal{P}_\beta} + \frac{\partial f}{\partial p_\beta} \frac{\partial^2 U_2}{\partial \mathcal{P}_\alpha \partial q^\beta} \right) \frac{\partial g}{\partial \mathcal{Q}_\alpha} \\
&= \sum_\alpha \sum_\beta \left(\frac{\partial f}{\partial \mathcal{Q}_\alpha} \frac{\partial g}{\partial \mathcal{Q}_\beta} - \frac{\partial f}{\partial \mathcal{Q}_\beta} \frac{\partial g}{\partial \mathcal{Q}_\alpha} \right) \frac{\partial^2 U_2}{\partial \mathcal{P}_\alpha \partial \mathcal{P}_\beta} \\
&+ \sum_\alpha \sum_\beta \left(\frac{\partial f}{\partial \mathcal{Q}_\alpha} \frac{\partial g}{\partial p_\beta} - \frac{\partial f}{\partial p_\beta} \frac{\partial g}{\partial \mathcal{Q}_\alpha} \right) \frac{\partial^2 U_2}{\partial \mathcal{P}_\alpha \partial q^\beta} \\
&= \sum_\alpha \sum_\beta \left(\frac{\partial f}{\partial \mathcal{Q}_\alpha} \frac{\partial g}{\partial p_\beta} - \frac{\partial f}{\partial p_\beta} \frac{\partial g}{\partial \mathcal{Q}_\alpha} \right) \frac{\partial^2 U_2}{\partial \mathcal{P}_\alpha \partial q^\beta}
\end{aligned}
$$

因此

$$\{f, g\}_{qp} = \{f, g\}_{\mathcal{Q}\mathcal{P}}$$

3.3.3 正则变换的辛结构 *

下面来简单看一看正则变换的代数结构——**辛结构**. 首先定义辛矩阵:

$$\boldsymbol{\Omega} = \begin{pmatrix} \mathbf{0} & \boldsymbol{I} \\ -\boldsymbol{I} & \mathbf{0} \end{pmatrix} \tag{3.18}$$

正则变换 $\mathcal{Q} = \mathcal{Q}(q, p, t)$，$\mathcal{P} = \mathcal{P}(q, p, t)$ 的 Jacobi 矩阵为:

$$\boldsymbol{J} = \frac{\partial(\mathcal{Q}, \mathcal{P})}{\partial(q, p)} \tag{3.19}$$

可以证明，正则变换一定使得如下关系成立:

$$\boldsymbol{J}^T \boldsymbol{\Omega} \boldsymbol{J} = \boldsymbol{\Omega} \tag{3.20}$$

Jacobi 矩阵满足 （3.20）式的变换称为**辛变换**. 所以，正则变换是辛变换。这为我们提供了一种思路: **如果我们想要检验一种正则变量的变换是否为正则变换，我们可以检验其 Jacobi 矩阵是否满足辛变换的关系。**

易知，从 $(\mathcal{Q},\mathcal{P})$ 到 (q,p) 的变换，即上述正则变换之逆变换，的 Jacobi 矩阵为[4]：

$$\boldsymbol{J}' = \frac{\partial(q,p)}{\partial(\mathcal{Q},\mathcal{P})} = \left(\frac{\partial(\mathcal{Q},\mathcal{P})}{\partial(q,p)}\right)^{-1} = \boldsymbol{J}^{-1}$$

于是，我们有：

$$(\boldsymbol{J}^{-1})^T\boldsymbol{\Omega}\boldsymbol{J}^{-1} = (\boldsymbol{J}^T)^{-1}\boldsymbol{\Omega}\boldsymbol{J}^{-1} = \boldsymbol{\Omega} \tag{3.21}$$

从而，**正则变换之逆变换也是辛变换，可以证明其也为正则变换.**

此外，设正则变量 (q,p) 经历了两次正则变换，则总变换的 Jacobi 矩阵为：

$$\boldsymbol{J} = \boldsymbol{J}_1\boldsymbol{J}_2$$

从而

$$\boldsymbol{J}^T\boldsymbol{\Omega}\boldsymbol{J} = \boldsymbol{J}_2^T(\boldsymbol{J}_1^T\boldsymbol{\Omega}\boldsymbol{J}_1)\boldsymbol{J}_2 = \boldsymbol{J}_2^T\boldsymbol{\Omega}\boldsymbol{J}_2 = \boldsymbol{\Omega} \tag{3.22}$$

于是**总变换亦为辛变换，可以证明其也为正则变换.**

最后，**正则变换的 Jacobi 行列式的绝对值等于 1**. 因为：

$$|\boldsymbol{J}^T|\,|\boldsymbol{\Omega}|\,|\boldsymbol{J}| = |\boldsymbol{\Omega}|$$

因为 $|\boldsymbol{\Omega}| \neq 0$，故

$$|\boldsymbol{J}^T|\,|\boldsymbol{J}| = |\boldsymbol{J}|^2 = 1 \tag{3.23}$$

这表明，在相空间 (q,p) 中体积元 dV，在变换到相空间 $(\mathcal{Q},\mathcal{P})$ 之后的体积元为：

$$dV' = dV \cdot \left|\frac{\partial(\mathcal{Q},\mathcal{P})}{\partial(q,p)}\right| = dV \tag{3.24}$$

即体积元大小不变，这是一条很好的性质.

[4]欲验证此关系，只需验证：

$$\frac{\partial(q,p)}{\partial(\mathcal{Q},\mathcal{P})}\frac{\partial(\mathcal{Q},\mathcal{P})}{\partial(q,p)} = \boldsymbol{I}$$

这不难验证。

3.3.4 无穷小正则变换与 Liouville 定理

我们首先给定一正则变换的母函数函数 $U_2=\sum_\alpha q_\alpha\mathcal{P}_\alpha+\varepsilon G(q,\mathcal{P},t)$ ，其中 ε 为小量，G 为一个任意形式的连续函数。这个母函数会生成如下正则变换:

$$\begin{cases}\mathcal{Q}_\alpha=\frac{\partial U_2}{\partial\mathcal{P}_\alpha}=q_\alpha+\varepsilon\frac{\partial G(q,\mathcal{P},t)}{\partial\mathcal{P}_\alpha}\\ p_\alpha=\frac{\partial U_2}{\partial q_\alpha}=\mathcal{P}_\alpha+\varepsilon\frac{\partial G(q,\mathcal{P},t)}{\partial q_\alpha}\end{cases}$$

于是我们可知，变换前后正则变量的差值为一阶小量:

$$\begin{cases}dq_\alpha=\mathcal{Q}_\alpha-q_\alpha=\varepsilon\frac{\partial G(q,\mathcal{P},t)}{\partial\mathcal{P}_\alpha}\\ dp_\alpha=\mathcal{P}_\alpha-p_\alpha=-\varepsilon\frac{\partial G(q,\mathcal{P},t)}{\partial q_\alpha}\end{cases}\tag{3.25}$$

于是，如若我们用变换后的正则动量替换变换之前的正则动量，即如若我们将 $G(q,\mathcal{P})$ 换成 $G(q,p)$，产生的误差将为 ε 阶；进而，如若我们将 $\frac{\partial G(q,\mathcal{P},t)}{\partial\mathcal{P}_\alpha}$ 换成 $\frac{\partial G(q,p,t)}{\partial p_\alpha}$，产生的误差将为 ε^2 阶，是高阶小量，可以忽略. 于是，我们可以大胆地将（3.25） 式改写为:

$$\begin{cases}dq_\alpha=\varepsilon\frac{\partial G(q,p,t)}{\partial p_\alpha}\\ dp_\alpha=-\varepsilon\frac{\partial G(q,p,t)}{\partial q_\alpha}\end{cases}\tag{3.26}$$

这同时意味着在小量 ϵ 存在的前提下，母函数 U_2 可以改成:

$$U_2=\sum_\alpha q_\alpha\mathcal{P}_\alpha+\varepsilon G(q,p,t)\tag{3.27}$$

使得（3.26）式仍旧为正则变换，我们称上面这个母函数为 **无穷小正则变换的母函数**. 正常来讲，U_2 的宗量中是不能出现 p 的，是 ε 这一小量赋予了无穷小正则变换这一特权.

如果我们把变换后得到的正则变量仍然放在变换前的相空间去看的话，那么以（3.26）表出的无穷小正则变换有很生动的物理意义：**这一变换可视为将体系在相空间 (q,p) 的点平移一个小距离 $d\boldsymbol{r}=(dq_1,\ldots,dq_n,dp_1,\ldots,dp_n)$，各分量的大小由（3.26）决定.**

特殊地，若令 $G(q,p,t)=H(q,p,t)$（即体系的 Hamilton 量），而 $\varepsilon=dt$，从而

有：

$$dq_\alpha = dt\frac{\partial H}{\partial p_\alpha}, \quad dp_\alpha = -dt\frac{\partial H}{\partial q_\alpha}$$

即得正则方程：

$$\dot{q}_\alpha = \frac{\partial H}{\partial p_\alpha}, \quad \dot{p}_\alpha = -\frac{\partial H}{\partial q_\alpha}$$

说明**体系在瞬间的演化可以视为一无穷小正则变换，而体系随宏观时间尺度的演化可视为一系列无穷小正则变换之累加**。上一小节我们指出，正则变换累加之后仍为正则变换，于是我们有可能用一个含有参数 t 的正则变换表征体系随时间的演化。这个正则变换将初始时刻的正则坐标（常数 q_0）和正则动量（常数 p_0）变换为 t 时刻的正则坐标与正则动量，即：

$$\begin{cases} q_{t\alpha} = q_{t\alpha}(q_0, p_0, t) \\ p_{t\alpha} = p_{t\alpha}(q_0, p_0, t) \end{cases} \tag{3.28}$$

找到这个正则变换，也就解出了体系的运动，而这正是我们下一节要讨论的 Hamilton-Jacobi 理论的出发点；不过在此之前，我们先来看一个有趣的定理。

Liouville 定理

Liouville 定理是无穷小正则变换最生动的直接应用。在前文我们已指出，体系的运动可用一条在相空间的 (超) 曲线来表示。通过设置不同的初始条件（或者通过取拥有不同初始条件的多个体系），可以在相空间中获得多个初始点。如若这些点是连续分布的，则我们能获得相空间中的一个区域。假设我们在 t_1 时刻取相空间的一块区域，经过一段时间的演化之后，在 t_2 时刻，原先区域中的点全部演化到了另一区域。**Liouville 定理表明，这两区域的体积一定是相同的，时间演化不会改变相空间区域的体积**。

对于相空间，我们有一个很生动的物理图像：如若我们将相空间中的各个点联合起来看，就得到了所谓**相流体**，各个点的演化也就成为了相流体的流动。那么，**Liouville 定理显然是在说相流体是不可压缩的**。

下面让我们来证明 Liouville 定理。我们已经证明体系在瞬间的演化可以用一无

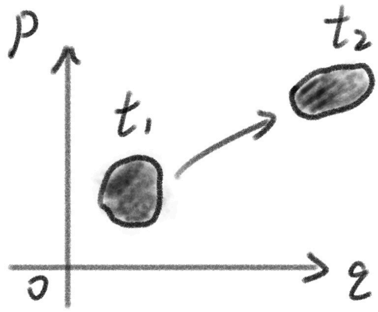

图 3.1: 时间演化不改变相流体的体积

穷小正则变换来表示

$$\begin{cases} q'_\alpha = q_\alpha + \dot{q}_\alpha dt = q_\alpha + \dfrac{\partial H}{\partial p_\alpha} dt \\ p'_\alpha = p_\alpha + \dot{p}_\alpha dt = p_\alpha - \dfrac{\partial H}{\partial q_\alpha} dt \end{cases}$$

既然是正则变换，其 Jacobi 行列式 $\dfrac{\partial(q',p')}{\partial(q,p)}$ 就应该为 1。如此一来，变换之后的区域体积为

$$V' = \int_{V'} dV' = \int_V \frac{\partial(q',p')}{\partial(q,p)} dV = \int_V dV = V$$

这便证明了 Liouville 定理。

为什么 Hamilton 力学不适用于耗散体系?

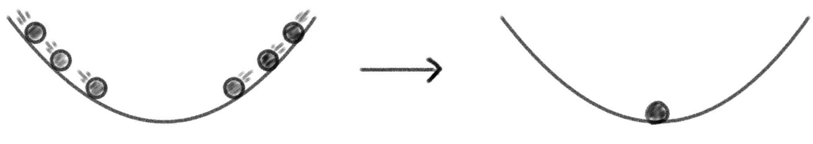

图 3.2: 有摩擦的土坑

利用 Liouville 定理,我们可以回答一个一直困扰着笔者的问题——为什么 Hamilton 力学只能描述没有耗散的体系呢？其原因是因为**耗散体系在相空间中的演化不满足 Liouville 定理，而 Hamilton 力学是一定满足 Liouville 定理的**。

为什么耗散体系不满足 Liouville 定理呢？我们来看一个例子。如图是一个有摩擦的土坑，现在我们在坡上不同位置以不同的初速度释放无穷多个小球，这些小球的初始条件不尽相同，从而初始时刻它们会在相空间中集结成一块区域。然而，由于存在摩擦耗散，这些小球最后都会停止在坡底，而它们在相空间的轨迹最后统统都会交汇于一点。显然它们相应的相流体一定不是不可压缩的，因为相流体的体积逐渐减小，最后变成零，不满足 Liouville 定理。

3.4 Hamilton - Jacobi 理论

Hamilton - Jacobi 理论是 Hamilton 力学的核心延伸理论，用这个理论求解体系运动的核心思想是利用一种特殊的正则变换，它将原先的正则坐标变换为常数，同时满足新的 Hamilton 量为零。

通过人为的令变换后的 Hamilton 量为零，同时用（第二类）母函数 U_2 对旧正则坐标 q_α 的偏导替换掉旧正则动量 p_α，我们可以直接得到 U_2 满足的偏微分方程即 Hamilton - Jacobi 方程。在解方程的过程中我们会遇到若干积分变量，它们都是常数；然后我们再人为地规定这些积分变量常数分别对应了正则变换后新正则动量 $\mathcal{P}_\alpha$；最后，利用第二类母函数对新正则动量求偏导，得到同样为常数的新正则坐标 $\mathcal{Q}_\alpha$，而到了这一步，你其实得到的是一个包含了旧正则坐标 q_α、时间 t 以及一大堆积分常数的式子，即体系运动方程的形式。

下面我们来详细解读上述内容，领略 Hamilton - Jacobi 理论的复杂之妙。

3.4.1 含时的 Hamilton - Jacobi 方程

上节末尾提到的正则变换：

$$\begin{cases} q_t^\alpha = q_t^\alpha(q_0, p_0, t) \\ p_{t\alpha} = p_{t\alpha}(q_0, p_0, t) \end{cases}$$

表征了体系随时间的演化. 而其逆变换：

$$\begin{cases} q_\alpha = q_t^\alpha(q, p, t) \\ p_\alpha = p_{t\alpha}(q, p, t) \end{cases} \tag{3.29}$$

也是正则变换——**这类正则变换很特殊，因为变换后的正则坐标是常数！若能找到这类正则变换，则实际上我们就找到了关于 q, p, t 的方程，从而可以解出体系的演化**. 问题是，如何找？母函数产生正则变换，因此问题自然归结为寻找上述变换的母函数。下面我们利用第二类母函数 $U_2(q, P, t)$ 来进行说明.

由于我们的目标正则变换得到的新正则坐标均为常数，因此它们的时间变化率

自然为 0，从而正则变换后，新的正则方程为：

$$\begin{cases} \dot{\mathcal{Q}}_\alpha = \frac{\partial K}{\partial \mathcal{P}_\alpha} = 0 \\ \dot{P}_\alpha = -\frac{\partial K}{\partial \mathcal{Q}_\alpha} = 0 \end{cases}$$

对任意 α 均成立。考虑最简单的情形，我们令新的 Hamilton 量 $K(P, Q, t) = 0$。于是，由 $K - H = \frac{\partial U_2}{\partial t}$ 可知：

$$H\left(q, p(q, \mathcal{P}, t), t\right) + \frac{\partial U_2(q, \mathcal{P}, t)}{\partial t} = H\left(q, \frac{\partial U_2}{\partial q}, t\right) + \frac{\partial U_2}{\partial t} = 0$$

我们便得到了 U_2 满足的偏微分方程，我们称之为**含时的 Hamilton-Jacobi 方程**，且称其解 U_2 为 **Hamilton 主函数**，换用 S 表示。于是我们重新写一下含时的 Hamilton-Jacobi 方程为

$$H\left(q, \frac{\partial S}{\partial q}, t\right) + \frac{\partial S}{\partial t} = 0 \tag{3.30}$$

下面我们的任务是求解这个微分方程，解出主函数 S[5]。

3.4.2 不含时的 Hamilton - Jacobi 方程

以下我们只考虑 H 不显含 t 的情形，则 Hamilton - Jacobi 方程变为：

$$H\left(q, \frac{\partial S}{\partial q}\right) + \frac{\partial S}{\partial t} = 0$$

[5] 顺便再来看一个神奇的结论. 我们来求 S 对时间的变化率：

$$\begin{aligned} \frac{dS(q, \mathcal{P}, t)}{dt} &= \sum_\alpha \left(\frac{\partial S}{\partial q_\alpha}\dot{q}_\alpha + \frac{\partial S}{\partial \mathcal{P}_\alpha}\dot{P}_\alpha + \frac{\partial S}{\partial t}\right) \\ &= \sum_\alpha \dot{q}_\alpha p_\alpha + 0 - H\left(q, p(q, \mathcal{P}, t), t\right) \\ &= \sum_\alpha \dot{q}_\alpha p_\alpha - H \\ &= L \end{aligned}$$

其中第一行到第二行是因为新的正则动量为常数，且 S 满足含时的 Hamilton-Jacobi 方程。于是：

$$S = \int L\, dt$$

即我们这里的 Hamilton 主函数正是体系的作用量！这就是为什么它要用 S 表示的原因.

然后，分离变量，记 $S(q,\mathcal{P},t)=W(q,\mathcal{P})+f(t)$，于是:

$$H\left(q,\frac{\partial W}{\partial q}\right)=-f'(t)$$

由于等式左右两边宗量不同，因此它们只能等于同一个常数，这个常数常常是能量 E. 于是:

$$\begin{cases}H\left(q,\frac{\partial W}{\partial q}\right)=E\\ f'(t)=-E\end{cases} \tag{3.31}$$

上式第二式给出 $f(t)=-Et$；而第一式系 W 关于广义坐标的一阶非线性偏微分方程，亦称之为**不含时的 Hamilton - Jacobi 方程**，其解 W 称为 **Hamilton 特征函数**。至此我们可以将 Hamilton 主函数表示为 $S=W(q,\mathcal{P})-Et$.

含时的 Hamilton - Jacobi 方程中只出现了对 q 的偏导，而未对 $\mathcal{P}$ 做出限制。我们对 $\mathcal{P}$ 只要求其为常数. 该方程是 $(q_1,q_2,\ldots,q_n,t)$ 共 $n+1$ 个变量的偏微分方程，解 S 应有 $n+1$ 个积分常数 $C_1,C_2,\ldots,C_{n+1}$. 注意到“$S+$ 常数”亦是方程之解，从而 $C_1,\ldots,C_{n+1}$ 中有一个以相加形式出现，不妨设之为 C_{n+1}. 而其余 n 个常数皆以非相加形式出现，**然后，我们人为地要求** $\mathcal{P}_i=C_i(i=1,2,3\cdots n)$，**这一步是神来之笔**。同样的道理，对于不含时的 Hamilton - Jacobi 方程，W 的解的形式中含有 n 个非相加型的积分常数，比如 E 就是其中之一。我们设这些积分常数为 $E,C_2,\ldots,C_n$，**且依次让它们对应于** $\mathcal{P}_1,\ldots,\mathcal{P}_n$.

对于一般的体系，求解 W 是十分困难的，但对于某些特殊的体系，还可以进一步用分离变量法解上述方程. 这样的体系叫可分离体系，进一步，若体系的所有变量均可分离，则称之为完全可分离体系. 下一小节我们举一些简单又典型的例子介绍如何实操.

3.4.3 Hamilton-Jacobi 理论的简单应用

经典谐振子的运动

谐振子的 Hamilton 量为:

$$H(x,p)=\frac{p^2}{2m}+\frac{1}{2}kx^2=E$$

首先写出 Hamilton 主函数：

$$S(x, \mathcal{P}, t) = W(x, \mathcal{P}) - Et$$

其中 W 满足的 Hamilton - Jacobi 方程为：

$$\frac{1}{2m}\left(\frac{\partial W}{\partial x}\right)^2 + \frac{1}{2}kx^2 - E = 0$$

解之得：

$$W(x, \mathcal{P}) = \int \sqrt{2mE - mkx^2}\, dx$$

然后，令 $\mathcal{P} = E$，于是母函数化为：

$$S(x, E, t) = \int \sqrt{2mE - mkx^2}\, dx - Et$$

故

$$X = \frac{\partial S}{\partial E} = \int \frac{m\, dx}{\sqrt{2mE - mkx^2}} - t = \sqrt{\frac{m}{k}} \arcsin\left(\sqrt{\frac{k}{2E}} x\right) - t = C$$

从而

$$x = \sqrt{\frac{2E}{k}} \sin\left(\sqrt{\frac{k}{m}}(t + C)\right)$$

不受力的自由质点

该问题的 Hamilton 量为：

$$H = \frac{1}{2m}\left(p_x^2 + p_y^2 + p_z^2\right)$$

Hamilton 主函数为 $S = W - Et$，其中 W 满足：

$$\frac{1}{2m}\left[\left(\frac{\partial W}{\partial x}\right)^2 + \left(\frac{\partial W}{\partial y}\right)^2 + \left(\frac{\partial W}{\partial z}\right)^2\right] = E$$

进一步分离变量，设 $W = W_1(x, \mathcal{P}_x, \mathcal{P}_y, \mathcal{P}_z) + W_2(y, \mathcal{P}_x, \mathcal{P}_y, \mathcal{P}_z) + W_3(z, \mathcal{P}_x, \mathcal{P}_y, \mathcal{P}_z)$，故:

$$\left(\frac{\partial W_1}{\partial x}\right)^2 + \left(\frac{\partial W_2}{\partial y}\right)^2 + \left(\frac{\partial W_3}{\partial z}\right)^2 = 2mE$$

故

$$2mE - \left(\frac{\partial W_1}{\partial x}\right)^2 = \left(\frac{\partial W_2}{\partial y}\right)^2 + \left(\frac{\partial W_3}{\partial z}\right)^2$$

等式两边宗量不同，要相等必须等于同一常数，设:

$$2mE - \left(\frac{\partial W_1}{\partial x}\right)^2 = \left(\frac{\partial W_2}{\partial y}\right)^2 + \left(\frac{\partial W_3}{\partial z}\right)^2 = C_2$$

这就实现了 x 的变量分离，故:

$$W_1 = \sqrt{2mE - C_2}\, x$$

再来分离 y 与 z，设:

$$C_2 - \left(\frac{\partial W_2}{\partial y}\right)^2 = \left(\frac{\partial W_3}{\partial z}\right)^2 = C_3$$

从而有:

$$W_2 = \sqrt{C_2 - C_3}\, y, \quad W_3 = \sqrt{C_3}\, z$$

于是:

$$S = \sqrt{2mE - C_2}\, x + \sqrt{C_2 - C_3}\, y + \sqrt{C_3}\, z - Et + C$$

共有 E, C_2, C_3 三个不以相加形式出现的积分常数，依次令它们等于 $\mathcal{P}_x, \mathcal{P}_y, \mathcal{P}_z$，利用正则变换得:

$$\begin{cases} X = \frac{\partial S}{\partial E} = \frac{m}{\sqrt{2mE - C_2}}\, x - t = X_0 \\ Y = \frac{\partial S}{\partial C_2} = \frac{1}{2\sqrt{C_2 - C_3}}\, y - \frac{1}{\sqrt{2mE - C_2}}\, x = Y_0 \\ Z = \frac{\partial S}{\partial C_3} = \frac{1}{2\sqrt{C_3}}\, z - \frac{1}{2\sqrt{C_2 - C_3}}\, y = Z_0 \end{cases}$$

于是，我们得到了结果：

$$\begin{cases} x = \sqrt{\frac{2mE-C_2}{m}}\,(t+X_0) \\ y = 2\sqrt{C_2-C_3}\,\left[Y_0+\frac{1}{m}(t+X_0)\right] \\ z = 2\sqrt{C_3}\,\left[Z_0+Y_0+\frac{1}{m}(t+X_0)\right] \end{cases}$$

重力场中质点的运动

Hamilton 量为：

$$H = \frac{p_x^2}{2m}+\frac{p_y^2}{2m}+\frac{p_z^2}{2m}+mgz = E$$

Hamilton 特征函数满足：

$$\frac{1}{2m}\left[\left(\frac{\partial W}{\partial x}\right)^2+\left(\frac{\partial W}{\partial y}\right)^2+\left(\frac{\partial W}{\partial z}\right)^2\right]+mgz = E$$

设 $W = W_x(x,\mathcal{P}_x,\mathcal{P}_y,\mathcal{P}_z)+W_y(y,\mathcal{P}_x,\mathcal{P}_y,\mathcal{P}_z)+W_z(z,\mathcal{P}_x,\mathcal{P}_y,\mathcal{P}_z)$，代入上式得：

$$\left[\left(\frac{\partial W_x}{\partial x}\right)^2+\left(\frac{\partial W_y}{\partial y}\right)^2\right] = 2mE-\left(\frac{\partial W_z}{\partial z}\right)^2-2m^2gz$$

由于等式左右两边宗量不同，必须等于同一常数，设为 C_2^2. 于是：

$$2mE-\left(\frac{\partial W_z}{\partial z}\right)^2-2m^2gz-C_2^2 = 0$$

即：

$$W_z = \int\sqrt{2mE-C_2^2-2m^2gz}\,dz$$

而

$$\left(\frac{\partial W_x}{\partial x}\right)^2+\left(\frac{\partial W_y}{\partial y}\right)^2 = C_2^2$$

进一步分离变量：

$$\left(\frac{\partial W_x}{\partial x}\right)^2-C_2^2 = -\left(\frac{\partial W_y}{\partial y}\right)^2 = -C_3^2$$

于是:

$$\begin{cases}\left(\frac{\partial W_x}{\partial x}\right)^2 = C_2^2 - C_3^2 \\ \left(\frac{\partial W_y}{\partial y}\right)^2 = C_3^2\end{cases}$$

即:

$$\begin{cases}W_x = \sqrt{C_2^2 - C_3^2}\, x \\ W_y = C_3\, y\end{cases}$$

于是 Hamilton 主函数为:

$$\begin{aligned} S &= W - Et \\ &= \sqrt{C_2^2 - C_3^2}\, x + C_3\, y + \int \sqrt{2mE - C_2^2 - 2m^2 gz}\, dz - Et \end{aligned}$$

令 $\mathcal{P}_x, \mathcal{P}_y, \mathcal{P}_z$ 依次为 C_2, C_3, E，故:

$$\begin{aligned} X = \frac{\partial S}{\partial \mathcal{P}_x} = \frac{\partial S}{\partial C_2} &= \frac{C_2}{\sqrt{C_2^2 - C_3^2}} - \int \frac{C_2\, dz}{\sqrt{2mE - C_2^2 - 2m^2 gz}} \\ &= \frac{C_2}{\sqrt{C_2^2 - C_3^2}} x + \frac{C_2}{m^2 g^2}\sqrt{2mE - C_2^2 - 2m^2 g^2 z} \\ &= X_0 \end{aligned}$$

$$Y = \frac{\partial S}{\partial C_3} = -\frac{C_3}{\sqrt{C_2^2 - C_3^2}} x + y = Y_0$$

$$Z = \frac{\partial S}{\partial E} = \frac{1}{mg}\sqrt{2mE - C_2^2 - 2m^2 g^2 z} - t = Z_0$$

于是解得:

$$\begin{cases} x = \sqrt{\frac{C_2^2 - C_3^2}{C_2^2}}\left[X_0 - \frac{C_2}{m}(t + z_0)\right] \\ y = Y_0 + \frac{C_3}{C_2}\left[X_0 - \frac{C_2}{m}(t + z_0)\right] \\ z = \frac{2mE - C_2^2 - m^2 g^2 (t + z_0)^2}{2m^2 g^2} \end{cases}$$

这就是抛体运动的轨迹。

行星轨道

行星绕日运动的 Hamilton 量为:

$$H=\frac{1}{2m}\left(p_r^2+p_\varphi^2\right)-\frac{GMm}{r}=E$$

Hamilton 特征函数 $W(r,\varphi,\mathcal{P}_r,\mathcal{P}_\varphi)$ 满足:

$$\frac{1}{2m}\left[\left(\frac{\partial W}{\partial r}\right)^2+\frac{1}{r^2}\left(\frac{\partial W}{\partial \varphi}\right)^2\right]-\frac{GMm}{r}-E=0$$

分离变量:

$$W=W_r(r,\mathcal{P}_r,\mathcal{P}_\varphi)+W_\varphi(\varphi,\mathcal{P}_r,\mathcal{P}_\varphi)$$

从而:

$$\frac{1}{2m}\left(\frac{\partial W_r}{\partial r}\right)^2+\frac{1}{2mr^2}\left(\frac{\partial W_\varphi}{\partial \varphi}\right)^2-\frac{GMm}{r}-E=0$$

即:

$$r^2\left(\frac{\partial W_r}{\partial r}\right)^2-2GMm^2r-2mEr^2=-\left(\frac{\partial W_\varphi}{\partial \varphi}\right)^2=-C^2$$

于是:

$$\begin{cases}\left(\frac{\partial W_r}{\partial r}\right)^2-\frac{2GMm^2}{r}-2mE+\frac{C^2}{r^2}=0\\ \left(\frac{\partial W_\varphi}{\partial \varphi}\right)^2=C^2\end{cases}$$

$$\begin{cases}W_r=\int\sqrt{2mE+\frac{2GMm^2}{r}-\frac{C^2}{r^2}}\,dr\\ W_\varphi=C\varphi\end{cases}$$

令 $\mathcal{P}_r=E$，$\mathcal{P}_\varphi=C$，Hamilton 主函数为:

$$S=\int\sqrt{2mE+\frac{2GMm^2}{r}-\frac{C^2}{r^2}}\,dr+C\varphi-Et$$

故

$$\begin{cases}R=\frac{\partial S}{\partial \mathcal{P}_r}=\int\frac{m\,dr}{\sqrt{2mE+\frac{2GMm^2}{r}-\frac{C^2}{r^2}}}-t=R_0\\ \Phi=\frac{\partial S}{\partial \mathcal{P}_\varphi}=\int\frac{-C/r^2\,dr}{\sqrt{2mE+\frac{2GMm^2}{r}-\frac{C^2}{r^2}}}+\varphi=\Phi_0\end{cases}$$

3.5 从经典到量子 *

本节我们揭示量子力学是如何脱胎于经典力学的。Hamilton 力学为我们提供了从经典过渡到量子的桥梁——通过将 Poisson 括号量子化，我们得到了量子力学中的算符对易关系与 Heisenberg 方程；通过将 Hamilton-Jacobi 方程量子化，我们得到了量子力学中的 Schrödinger 方程。

3.5.1 正则量子化

正则量子化为我们提供了量子力学中计算力学量算符之间代数的一种方法。它指出**量子力学中力学量之间的对易关系，完全由经典 Poisson 括号经由以下手续进行了“量子化”而得到——将力学量替换为算符、将 Poisson 替换为对易子并添上系数 $\frac{1}{i\hbar}$，即**

$$\{A,B\} \xrightarrow{\text{量子化}} \frac{1}{i\hbar}[\hat{A},\hat{B}] \tag{3.32}$$

有了如上手续，我们可以直接得到算符之间的对易关系，一些典型且重要的关系有

$$[\hat{x}_i,\hat{x}_j]=[\hat{p}_i,\hat{p}_j]=0 \quad , \quad [\hat{x}_i,\hat{p}_j]=i\hbar\delta_{ij}$$

$$[\hat{L}_i,\hat{L}_j]=i\hbar\varepsilon_{ijk}\hat{L}_k \quad , \quad [\hat{L}_i,\hat{L}^2]=0$$

以及，利用经典力学中力学量随时间的演化方程 $\frac{dA}{dt}=\{A,H\}+\frac{\partial A}{\partial t}$ 经过正则量子化，我们得到了描述 Heisenberg 绘景下算符随时间的演化的 **Heisenberg 方程**

$$\frac{d\hat{A}}{dt}=\frac{1}{i\hbar}[\hat{A},\hat{H}]+\frac{\partial\hat{A}}{\partial t} \tag{3.33}$$

3.5.2 Schrödinger 方程的“导出”

下面我们提供一种“导出”量子力学基本方程即 **Schrödinger 方程**的一种理解方式。在前面的章节我们已得到了 Hamilton-Jacobi 方程 (H-J 方程)

$$H(\boldsymbol{x},\nabla S,t)+\frac{\partial S}{\partial t}=0$$

其中我们采用了直角坐标，S 为体系的作用量，满足 $\frac{dS}{dt}=L$。欲令一个经典体系过渡到量子体系，我们首先做一个**变量替换**，用 ψ 去替换 S，二者之间的关系满足

$$S(\boldsymbol{x},t)=-i\hbar\ln\psi(\boldsymbol{x},t) \tag{3.34}$$

将之代入含时的 H-J 方程，我们得到 $H=-\frac{\partial S}{\partial t}=i\hbar\frac{1}{\psi}\frac{\partial\psi}{\partial t}$，即 $i\hbar\frac{\partial\psi}{\partial t}=H\psi$。然后是关键的一步，将其**算符化**（或者说，将其量子力学化）——把力学量替换成相应的量子力学算符，于是我们得到

$$i\hbar\frac{\partial\psi}{\partial t}=\hat{H}\psi \tag{3.35}$$

这即是含时 Schrödinger 方程最一般化的形式，可将其视为将经典 H-J 方程量子化的结果，是经典体系演化方程的量子力学对应。而欲将该方程进一步化为可计算的微分方程，我们还要找出 Hamilton 算符 $\hat{H}$ 的形式。注意到在经典力学中，我们有

$$\boldsymbol{p}=\nabla S$$

用 ψ 代换之后，我们得到 $-i\hbar\nabla\psi=\boldsymbol{p}\psi$，然后仍然是关键的一步，将其**算符化**，我们便找到了动量算符的形式

$$\hat{\boldsymbol{p}}=-i\hbar\nabla \tag{3.36}$$

利用对易关系 $[\hat{x},\hat{p}]=i\hbar$，容易得到位置算符的形式为

$$\hat{\boldsymbol{x}}=\boldsymbol{x} \tag{3.37}$$

由于 Poisson 括号与对易子有相同的代数结构，因此对于力学量 $A(\boldsymbol{x},\boldsymbol{p})$，其算符化之后的结果为 $\hat{A}(\boldsymbol{x},-i\hbar\nabla)$，且该结果与其它算符的对易子一定满足正则量子化手续。特别地，Hamilton 算符为

$$\hat{H}=-\frac{\hbar^2}{2m}\nabla^2+V(\hat{\boldsymbol{x}})$$

于是，我们得到了含时 Schrödinger 方程（在位置表象下）的具体微分方程形式

$$i\hbar\frac{\partial\psi}{\partial t}=-\frac{\hbar^2}{2m}\nabla^2\psi+V\psi \tag{3.38}$$

对应原理

Bohr 的对应原理指出**量子体系在经典极限下（即在 $\hbar$ 可忽略时）会回归经典体系**。下面我们证明，Schrödinger 在 $\hbar \to 0$ 的极限下的确回归经典力学的 H-J 方程。首先，由 Schrödinger 方程可得

$$i\hbar \frac{1}{\psi}\frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{1}{\psi}\nabla^2\psi + V$$

亦即

$$\begin{aligned}
\frac{\partial}{\partial t}\big(i\hbar \ln\psi\big) &= -\frac{\hbar^2}{2m}\frac{\psi\nabla^2\psi}{\psi^2} + V \\
&= -\frac{\hbar^2}{2m}\nabla\cdot\left(\frac{\nabla\psi}{\psi}\right) - \frac{\hbar^2}{2m}\left(\frac{\nabla\psi}{\psi}\right)^2 + V \\
&= \frac{i\hbar}{2m}\nabla^2\big(i\hbar\ln\psi\big) + \frac{1}{2m}\big(\nabla i\hbar\ln\psi\big)^2 + V
\end{aligned}$$

从而

$$H + \frac{\partial S}{\partial t} = \frac{(\nabla S)^2}{2m} + V + \frac{\partial S}{\partial t} = -\frac{i\hbar}{2m}\nabla^2 S$$

于是在 $\hbar \to 0$ 的经典极限下，我们最终回归 H-J 方程

$$H + \frac{\partial S}{\partial t} = 0$$

Chapter 4

经典场论

4.1 场方程

4.1.1 什么是场?

从数学的角度来讲，**场是时间和空间的连续函数**，在这里，坐标和时间都是自变量，而不是像质点动力学中坐标是时间的函数一样。今后我们用 $\phi_a(\boldsymbol{x},t)$ 表示编号为 a 的场量，它可以是标量场或者矢量场的某一分量。而在物理学中，某些场具有生动的物理图像与深刻的物理意义，**可以像实物粒子一样具有能量、动量等属性，而且也可以具有自己的动力学演化方程即场方程**。我们的问题是，该如何寻找这些场方程?原则上讲，我们除了做实验别无他法，不过我们却有办法将各种场方程的导出归纳到一个统一的理论框架下，那就是场论中的作用量原理。下面让我们来看一看。

4.1.2 场的 Euler-Lagrange 方程

我们定义场的 Lagrange 量

$$L = \int_{\mathbb{R}^3} \mathcal{L}\, d^3x \tag{4.1}$$

其中 $\mathcal{L}$ 我们称之为 **Lagrange 密度**。一般而言，$\mathcal{L}$ 是 $\phi_a, \dot{\phi}_a, \partial_i\phi_a, \boldsymbol{x}, t$（其中包含了所有可能的 a）的函数。研究一个场，重要的是找到其场方程，即 ϕ_a 关于 $\boldsymbol{x}$ 与 t 的微分方程。电磁场的 Maxwell 方程就是典型场方程。下面我们介绍用作用量原理导出场方程的方法，其与导出粒子体系运动方程的方法非常类似。

对于弥散在空间中连续分布的场而言，其在 $t_1 \sim t_2$ 时间内随时间演化的作用量为

$$S = \int_{t_1}^{t_2} L\, dt = \int \mathcal{L}\, d^4x \tag{4.2}$$

其中我们用了简化的标记 $\int = \int_{\mathbb{R}^3}\int_{t_1}^{t_2}$，以及 $d^4x = dt\, d^3x$。假设在 t 时刻，$\boldsymbol{x}$ 处的场为 $\phi(\boldsymbol{x},t)$，然后我们让在除 t_1, t_2 时刻以外的任一时刻的场均有一个构型的改变 $\delta\phi_a = \delta\phi_a(\boldsymbol{x},t)$，它满足：1. $|\delta\phi_a|$ 很小；2. $\delta\phi_a(\boldsymbol{x},t_1) = \delta\phi_a(\boldsymbol{x},t_2) = 0$。随后我们来计算作用量在此构型改变之下的变分

$$\delta S = \int \delta\mathcal{L}\, d^4x = \int \sum_a \left[\frac{\partial\mathcal{L}}{\partial\phi_a}\delta\phi_a + \frac{\partial\mathcal{L}}{\partial\dot{\phi}_a}\delta\dot{\phi}_a + \sum_i \frac{\partial\mathcal{L}}{\partial(\partial_i\phi_a)}\partial_i(\delta\phi_a)\right] d^4x$$

其中 $\partial_i\phi_a = \dfrac{\partial\phi_a}{\partial x_i}$ $(i = 1, 2, 3)$。接着我们用分部积分法，并且利用边界条件 $\delta\phi_a(\boldsymbol{x}, t_1) = \delta\phi_a(\boldsymbol{x}, t_2) = 0$ 与局域化条件 $\phi_a(\infty, t) = 0$，得

$$\delta S = \int \sum_a \left[\frac{\partial\mathcal{L}}{\partial\phi_a} - \frac{\partial}{\partial t}\frac{\partial\mathcal{L}}{\partial\dot{\phi}_a} - \sum_i \frac{\partial}{\partial x_i}\frac{\partial\mathcal{L}}{\partial(\partial_i\phi_a)} \right] \delta\phi_a \, d^4x$$

作用量原理指出，对于场的真实演化，对于任意的 $\delta\phi_a$，一定有 $\delta S = 0$ 恒成立。 于是我们得到了**场方程**，亦即 **Euler-Lagrange** 方程。其第 a 号方程为

$$\frac{\partial}{\partial t}\frac{\partial\mathcal{L}}{\partial\dot{\phi}_a} + \sum_i \frac{\partial}{\partial x_i}\frac{\partial\mathcal{L}}{\partial(\partial_i\phi_a)} - \frac{\partial\mathcal{L}}{\partial\phi_a} = 0 \tag{4.3}$$

在场论中，Lagrange 密度的选取也有一个自由度——**如果我们对 Lagrange 密度添加一个标量对时间的偏导数[1]项，或者一个矢量的散度项，利用作用量原理我们仍可以得到同样的场方程。**

上述场方程对于所有场均成立。然而，欲获得更加具体形式的场方程，我们须找到相应的具体形式的 $\mathcal{L}$，代入场方程中计算求解。下我们看几个实例。

4.1.3 举例一：d'Alembert 方程

本节我们研究质量连续分布的物质体系，如一维匀质弦、薄膜、重弹簧，的微小振动场，场量 $\phi(\boldsymbol{x}, t)$ 为 t 时刻位于 $\boldsymbol{x}$ 处质元偏离平衡位置的位移。这类场的场方程就是 d'Alembert 方程。接下来，我们首先用微元法寻找到相应体系的 Lagrange 密度，然后直接代入 Euler- Lagrange 方程求解出场方程的具体形式。

弦的横向振动

将弦分成 N（$N \to \infty$）份，每份长 $dx = \frac{L}{N}$，质量为 dm. 下面我们考虑每一份的动能

$$dT = \frac{1}{2}dm\left(\frac{\partial\phi}{\partial t}\right)^2 = \frac{\rho}{2}\phi_t^2 dx$$

[1] 为什么在场论中，我们可以添加偏导数，而在之前的质点动力学中，只能添加对时间的全导数？这就是因为，在场论中，时间和空间坐标的地位是平等的；而在质点动力学中，我们将（广义）坐标看成是时间的函数，二者地位是不对等的。

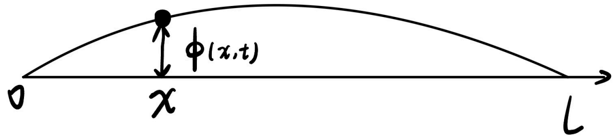

图 4.1: 一维匀质弦的振动

与势能

$$
\begin{aligned}
dV = F_T(dl - dx) &= F_T\left[\sqrt{(dx)^2 + (dy)^2} - dx\right] \\
&= F_T\left(\sqrt{1+\phi_x^2} - 1\right)dx \\
&\approx F_T\left(1 + \frac{1}{2}\phi_x^2 - 1\right)dx \\
&= \frac{F_T\phi_x^2}{2}dx
\end{aligned}
$$

注意，由于在这个体系中不同 x 处的 F_T 都相同，因此绳子的劲度系数不均匀，在计算势能时切不可当作像劲度系数均匀的弹簧一样的体系，然而 $F_T = -\frac{dV}{dl-dx}$ 还是成立的. 从而，我们得到体系的 Lagrange 密度为

$$
\mathcal{L} = \frac{1}{2}\rho\phi_t^2 - \frac{1}{2}F_T\phi_x^2 \tag{4.4}
$$

代入到 Euler- Lagrange 方程中计算，得到

$$
\frac{\partial}{\partial x}\frac{\partial\mathcal{L}}{\partial\phi_x} + \frac{\partial}{\partial t}\frac{\partial\mathcal{L}}{\partial\phi_t} - \frac{\partial\mathcal{L}}{\partial\phi} = \rho\phi_{tt} - F_T\phi_{xx} = 0
$$

即

$$
\phi_{tt} - \frac{F_T}{\rho}\phi_{xx} = 0 \tag{4.5}
$$

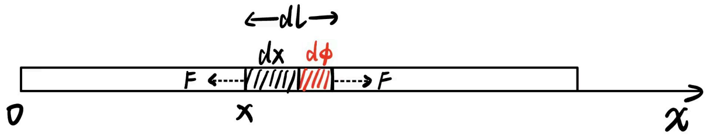

图 4.2: 一维匀质杆的纵振动

杆的纵振动

下面我们考虑一根匀质杆的纵振动。假设在 x 处的微元，原长为 dx，现在被拉伸至 dl 长度，设杆的横截面积为 S。由 Hooke 定律知，应力 $\frac{F}{S}$ 满足关系：

$$\frac{F}{S} = E\frac{dl - dx}{dx}$$

于是内力大小：

$$F = ES\frac{dl - dx}{dx} = ES\frac{dx + \phi(x + dx, t) - \phi(x, t) - dx}{dx}$$

而微元的弹性势能 dV 满足：

$$d\mathcal{V} = \int_0^u ES\frac{d\phi}{dx}d(d\phi) = \frac{1}{2}\frac{ES(d\phi)^2}{dx} = \frac{1}{2}ES\phi_x^2 dx$$

于是我们构造出了 Lagrange 密度：

$$\mathcal{L} = \frac{dT - dV}{dx} = \frac{1}{2}\frac{dm}{dx}\phi_t^2 - \frac{1}{2}ES\phi_x^2 = \left(\frac{1}{2}\rho\phi_t^2 - \frac{1}{2}E\phi_x^2\right)S$$

进而得到场方程：

$$\frac{\partial}{\partial x}\frac{\partial \mathcal{L}}{\partial \phi_x} + \frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \phi_t} - \frac{\partial \mathcal{L}}{\partial \phi} = (\rho\phi_{tt} - E\phi_{xx})\, S = 0$$

故

$$\phi_{tt} - \frac{E}{\rho}\phi_{xx} = 0 \tag{4.11}$$

此即匀质杆的纵振动的 d'Alembert 方程。

薄膜振动

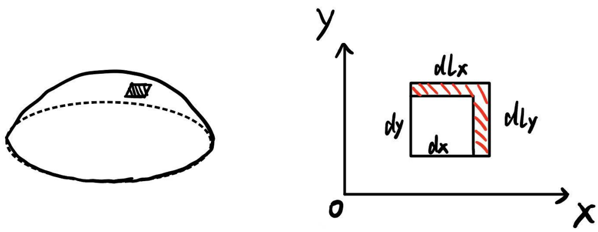

图 4.3: 匀质薄膜的横向振动

然后让我们来看匀质薄膜，设其表面张力系数为 σ，易知在 (x,y) 处的小质元的表面能为:

$$
\begin{aligned}
dV &= F_x(dl_x - dx) + F_y(dl_y - dy) \\
&= \sigma dy\left(\sqrt{dx^2 + (d\phi_x)^2} - dx\right) + \sigma dx\left(\sqrt{dy^2 + (d\phi_y)^2} - dy\right) \\
&= \sigma dxdy\left(\sqrt{1+\phi_x^2} - 1 + \sqrt{1+\phi_y^2} - 1\right) \\
&\approx \frac{\sigma}{2}\left(\phi_x^2 + \phi_y^2\right) dxdy
\end{aligned}
$$

于是，Lagrange 密度为:

$$
\begin{aligned}
\mathcal{L} &= \frac{dT}{dxdy} - \frac{dV}{dxdy} = \frac{1}{2}\frac{\rho dxdy}{dxdy}\phi_t^2 - \frac{\sigma}{2}\left(\phi_x^2 + \phi_y^2\right) \\
&= \frac{1}{2}\rho\phi_t^2 - \frac{1}{2}\sigma\left(\phi_x^2 + \phi_y^2\right)
\end{aligned}
$$

进而得场方程为:

$$
\frac{\partial}{\partial t}\frac{\partial\mathcal{L}}{\partial\phi_t} + \frac{\partial}{\partial x}\frac{\partial\mathcal{L}}{\partial\phi_x} + \frac{\partial}{\partial y}\frac{\partial\mathcal{L}}{\partial\phi_y} - \frac{\partial\mathcal{L}}{\partial\phi} = \rho\phi_{tt} - \sigma(\phi_{xx} + \phi_{yy}) = 0
$$

即:

$$
\phi_{tt} - \frac{\sigma}{\rho}(\phi_{xx} + \phi_{yy}) = 0
$$

这就是平面薄膜横振动的 d'Alembert 方程。

重弹簧

最后我们分析一个稍加复杂的体系——重弹簧的纵振动。首先来看其水平振动如图为一水平振动的重弹簧。设弹簧质量线密度为 ρ，原长为 L，总劲度系数为 k，

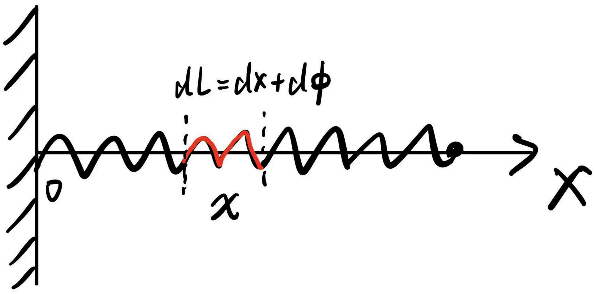

图 4.4: 水平重弹簧的振动

$u=\phi(x,t)$ 为各点相对于原长位置的偏移量。考查 x 处的质元，受力大小为:

$$F=\kappa(dl-dx)=\kappa d\phi$$

其中 κ 为该小段的劲度系数。该小段的势能为:

$$dV=\int_0^{d\phi}Fd(d\phi)=\frac{1}{2}\kappa(d\phi)^2$$

若整条弹簧劲度系数为 k，可视作无穷多质元弹簧串联，故:

$$\frac{1}{k}=\frac{1}{\kappa}+\frac{1}{\kappa}+\cdots+\frac{1}{\kappa}=\frac{L}{dx}\cdot\frac{1}{\kappa}$$

从而我们得到了质元的劲度系数和弹簧劲度系数的关系为[2]:

$$\kappa=\frac{kL}{dx}$$

于是，势能:

$$dV=\frac{1}{2}\cdot\frac{kL}{dx}\cdot(d\phi)^2=\frac{1}{2}kL\phi_x^2dx$$

[2]另一种理解：由应力关系 $\frac{F}{S}=E\frac{\Delta L}{L_0}$，得 $F=\frac{ES}{L_0}\Delta L$，故弹簧劲度系数 $k=\frac{ES}{L_0}$，且 $k\propto\frac{1}{L_0}$。所以对于弹簧微元，有：$\kappa=\frac{L}{dx}\cdot k$

从而我们得到了 Lagrange 量密度：

$$\mathcal{L} = \frac{dT - dV}{dx} = \frac{1}{2}\rho\phi_t^2 - \frac{1}{2}kL\phi_x^2$$

进而得重弹簧水平振动的 d'Alembert 方程：

$$\phi_{tt} - \frac{kL}{\rho}\phi_{xx} = 0$$

这是标准的纵波方程，不要求 ϕ 为小振动量。

现在弹簧吊在天花板上，取天花板为势能零点，ϕ 为相对于原长位置的净位移[3]，则势能包含弹性势能与重力势能：

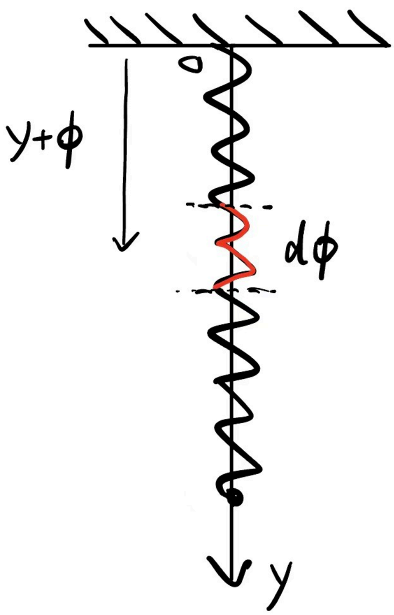

图 4.5: 竖直重弹簧的振动

$$dV = \frac{1}{2}kL\phi_y^2 dy - \rho dy \cdot g(y + \phi)$$

[3]如果读者不想像作者一样在期末考试中扣大分，那么在找竖直重弹簧的 Lagrange 密度时一定要明白这个 ϕ 到底是什么物理量，这样才能写出正确的重力势能项！

从而 Lagrange 量密度为:

$$\mathcal{L}=\frac{1}{2}\rho\phi_t^2+\rho g(y+\phi)-\frac{1}{2}kL\phi_y^2$$

进而得到振动方程

$$\rho\phi_{tt}-kL\phi_{yy}-\rho g=0$$

即:

$$\phi_{tt}-\frac{kL}{\rho}\phi_{yy}=g$$

4.1.4 举例二：Maxwell 方程组

接下来，我们研究另一种经典场——电磁场。下面我们直接给出电磁场的 Lagrange 密度[4]，然后代入到 Euler- Lagrange 方程中，求出场方程，即 Maxwell 方程组。

电磁场的 Lagrange 密度应该是关于电磁势及其时空导数的函数，具体形式为

$$\begin{aligned}\mathcal{L}&=\frac{1}{2}\left(\varepsilon_0\boldsymbol{E}^2-\mu_0\boldsymbol{H}^2\right)-\rho\varphi+\boldsymbol{J}\cdot\boldsymbol{A}\\&=\frac{1}{2}\left[\varepsilon_0\left(\nabla\varphi+\frac{\partial\boldsymbol{A}}{\partial t}\right)^2-\frac{1}{\mu_0}(\nabla\times\boldsymbol{A})^2\right]-\rho\varphi+\boldsymbol{J}\cdot\boldsymbol{A}\end{aligned}\tag{4.6}$$

其中 ρ 与 $\boldsymbol{J}$ 是电荷密度与电流密度。注意，正确的做法是用电磁势 φ 与 $\boldsymbol{A}$ 表示 $\mathcal{L}$，为了实现这一点，在上面的过程中我们实际上利用了如下关系

$$\boldsymbol{E}=-\nabla\varphi-\frac{\partial\boldsymbol{A}}{\partial t}\quad,\quad\boldsymbol{B}=\nabla\times\boldsymbol{A}$$

因此，只要你事先利用了上述关系，通过直接计算，我们便得到了场方程之中的两个

$$\nabla\times\boldsymbol{E}=-\frac{\partial\boldsymbol{B}}{\partial t}\quad,\quad\nabla\cdot\boldsymbol{B}=0$$

下面我们用场论的套路去寻找另两个。将上述 $\mathcal{L}$ 代入 Euler-Lagrange 方程

[4]注意不是电磁场中带电粒子的 Lagrange 密度，而是场本身的 Lagrange 密度。

- 针对 φ 列出方程，我们有

$$\frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot{\varphi}}+\sum_i \partial_i \frac{\partial \mathcal{L}}{\partial(\partial_i\varphi)}-\frac{\partial \mathcal{L}}{\partial \varphi}=\varepsilon_0\sum_i \partial_i\left(\partial_i\varphi+\frac{\partial A_i}{\partial t}\right)+\rho=0$$

这即是

$$\nabla\cdot\boldsymbol{E}=\frac{\rho}{\varepsilon_0}$$

- 然后我们针对 $\boldsymbol{A}$，列出方程 [5]

$$\frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot{A}_j}+\sum_i \partial_i \frac{\partial \mathcal{L}}{\partial(\partial_i A_j)}-\frac{\partial \mathcal{L}}{\partial A_j}=\varepsilon_0\frac{\partial}{\partial t}\left(\partial_j\varphi+\frac{\partial A_j}{\partial t}\right)+\frac{1}{\mu_0}\left(\nabla\times\boldsymbol{B}\right)_j-J_j$$

于是有

$$\nabla\times\boldsymbol{B}=\mu_0\boldsymbol{J}+\varepsilon_0\mu_0\frac{\partial \boldsymbol{E}}{\partial t}$$

[5]在这里的推导中我们跳了很多步，所以或许有些读者看不太懂。如果你已经学习了我们的电动力学讲义，下面的内容会对你有所帮助（请注意，我们在下面启动了 Einstein 求和约定）。首先，我们应该注意到

$$(\nabla\times\boldsymbol{A})^2=\boldsymbol{B}\cdot(\nabla\times\boldsymbol{A})=\varepsilon_{kij}B_k\partial_i A_j$$

因此

$$\partial_i\frac{\partial(\nabla\times\boldsymbol{A})^2}{\partial(\partial_i A_j)}=\partial_i(2\varepsilon_{kij}B_k)=-2\varepsilon_{jik}\partial_i B_k=-2\left(\nabla\times\boldsymbol{B}\right)_j$$

其中第二个式子中，在求第一个偏导数的时候出现了系数 2，这不太好理解。最好的办法是自己去展开指标计算一遍；如果一定要一个概括性的解释，我只能说因为 i 与 j 都是哑指标，因此我们必须要“算两遍”。

4.2 Norther 定理（场论）

在第二章中我们讨论了质点动力学版本的 Noether 定理，下面我们来看一看场论中的 Noether 定理，它指出**场的每一个连续的对称性都对应了一个守恒量**。

4.2.1 场论中的守恒量

在证明该定理之前，不妨先来辨析一下在质点动力学与场论中“守恒量”这一概念有何异同？首先，守恒量就是不随时间改变的量。我们定义**守恒荷** q，**对于一个体系而言，守恒荷的总量不随时间改变，是个守恒量**。总守恒荷我们用 Q 表示。q 在体系中具有一定的分布。对于质点系而言，q 的分布是离散的。假设第 i 个质点带有守恒荷 q_i，则 $Q=\sum_i q_i$ 是守恒量，即 Q 满足

$$\frac{dQ}{dt}=\sum_i \frac{dq_i}{dt}=0 \tag{4.7}$$

换句话说，对于 q_i，我们有

$$\frac{dq_i}{dt}=-\sum_{j\neq i}\frac{dq_j}{dt}=-I_i$$

其中 I_i 为**守恒流强度**，表征了守恒荷从第 i 个质点流出流入的强度，流出为正，流入为负。上面的等式很好地表现出了守恒荷在体系内部的各个质点之间“此消彼长”的过程。将上式稍加变形，我们能得到**守恒方程** [6]

$$\frac{dq_i}{dt}+I_i=0 \tag{4.8}$$

将上述内容推广到场论的情形是很容易的。守恒荷在场中是连续分布的，我们引入**守恒荷密度** ρ 与**守恒荷流密度** $\boldsymbol{j}$，它们与 q,I 的数学关系完全类比于电荷电流体系。如此一来，场论中的守恒方程变成了

$$\frac{\partial\rho}{\partial t}+\nabla\cdot\boldsymbol{j}=0 \tag{4.9}$$

[6] 为什么我们在之前在质点动力学体系下讲 Norther 定理的时候没有强调这个守恒方程呢？因为质点动力学体系下，我们的研究对象很可能只是单个的质点；而即便是研究质点系，我们也会去研究整个体系的守恒量，而不是单个质点带有的守恒荷。在离散体系下我们可以将整体视为研究对象，但是在场论中，自由度是无穷，所以这里的“整体”含有的组分的个数是无穷，原则上不能再将整体视为研究对象（因为这样是没有意义的），而应该研究局域的守恒荷和守恒流。

对于局域场，在无穷远处满足 $\rho=0$, $\boldsymbol{j}=0$。

4.2.2 场论中 Noether 定理的证明

我们的证明策略和在质点系时一样，直接找出那个守恒量。假设体系在操作 $\phi_a\to\phi_a+\delta\phi_a$ 下运动方程保持不变，即上述操作为体系的一个对称操作，则对于 Lagrange 密度，应该有

$$\delta\mathcal{L}=\frac{\partial\chi_0}{\partial t}+\nabla\cdot\boldsymbol{\chi}$$

从而利用作用量原理得到的场方程不会改变。而另一方面，通过直接计算，我们有

$$\begin{aligned}
\delta\mathcal{L}&=\sum_a\frac{\partial\mathcal{L}}{\partial\phi_a}\delta\phi_a+\sum_a\frac{\partial\mathcal{L}}{\partial\dot{\phi}_a}\delta\dot{\phi}_a+\sum_a\sum_i\frac{\partial\mathcal{L}}{\partial(\partial_i\phi_a)}\partial_i(\delta\phi_a)\\
&=\sum_a\left[\frac{\partial\mathcal{L}}{\partial\phi_a}-\frac{\partial}{\partial t}\frac{\partial\mathcal{L}}{\partial\dot{\phi}_a}-\sum_i\frac{\partial}{\partial x_i}\frac{\partial\mathcal{L}}{\partial(\partial_i\phi_a)}\right]\delta\phi_a+\sum_a\frac{\partial}{\partial t}\left(\frac{\partial\mathcal{L}}{\partial\dot{\phi}_a}\delta\phi_a\right)\\
&+\sum_a\sum_i\frac{\partial}{\partial x_i}\left(\frac{\partial\mathcal{L}}{\partial(\partial_i\phi_a)}\delta\phi_a\right)\\
&=\sum_a\frac{\partial}{\partial t}\left(\frac{\partial\mathcal{L}}{\partial\dot{\phi}_a}\delta\phi_a\right)+\sum_a\nabla\cdot\left(\frac{\partial\mathcal{L}}{\partial(\nabla\phi_a)}\delta\phi_a\right)
\end{aligned}$$

从而，我们得到了守恒方程

$$\frac{\partial}{\partial t}\left(\sum_a\frac{\partial\mathcal{L}}{\partial\dot{\phi}_a}\delta\phi_a-\chi_0\right)+\nabla\cdot\left(\sum_a\frac{\partial\mathcal{L}}{\partial(\nabla\phi_a)}\delta\phi_a-\boldsymbol{\chi}\right)=0 \tag{4.10}$$

并找到了守恒荷密度 $\sum_a\frac{\partial\mathcal{L}}{\partial\dot{\phi}_a}\delta\phi_a-\chi_0$ 与守恒流密度矢量 $\sum_a\frac{\partial\mathcal{L}}{\partial(\nabla\phi_a)}\delta\phi_a-\boldsymbol{\chi}$.

4.2.3 对称性与守恒量

本节我们利用 Norther 定理，将最常见的守恒量即能量、动量和角动量与时空的对称性联系起来。通过这一节的讨论我们会明白，经典场中涌现出的上述守恒律正是我们所处的时空是均匀且各向同性的结果。

时空平移对称性与能量-动量张量

假设现在场的构形发生了一个微小的时间平移

$$\phi_a(\boldsymbol{x},t) \to \phi_a(\boldsymbol{x},t+\varepsilon)$$

从而 $\delta\phi_a = \dot{\phi}_a\varepsilon$。在时间平移操作下，从数学角度上来讲 $\mathcal{L}$ **一定会**以如下方式做变化 $\mathcal{L} \to \mathcal{L} + \frac{\partial \mathcal{L}}{\partial t}\varepsilon$。结合先前的讨论，我们能得知一定存在如下守恒荷密度与守恒流密度

$$T_{00} = \sum_a \frac{\partial \mathcal{L}}{\partial \dot{\phi}_a}\dot{\phi}_a - \mathcal{L}, \quad T_{j0} = \sum_a \dot{\phi}_a \frac{\partial \mathcal{L}}{\partial(\partial_j \phi_a)} \tag{4.11}$$

它们满足守恒方程

$$\frac{\partial T_{00}}{\partial t} + \sum_j \partial_j T_{j0} = 0 \tag{4.12}$$

注意，T_{00} 实际上就是 **Hamilton 密度** $\mathcal{H}$，或**能量密度**，它满足

$$\int_{\mathbb{R}^3} T_{00} dV = \int_{\mathbb{R}^3} \mathcal{H} dV = H$$

从而 T_{j0} 应为**能流密度矢量**的第 j 个分量，其中 $j = 1,2,3$。

接下来，我们让场的构形在 x_i 方向上发生一个微小的空间平移 ε_i。从而 $\delta\phi_a = \partial_i\phi_a\varepsilon_i$ 而在此空间平移变换下，在数学上**一定有** $\mathcal{L} \to \mathcal{L} + \frac{\partial \mathcal{L}}{\partial x_i}\varepsilon_i$ ，于是一定存在守恒荷密度与守恒流密度

$$T_{0i} = \sum_a \frac{\partial \mathcal{L}}{\partial \dot{\phi}_a}\partial_i\phi_a, \quad T_{ji} = \sum_a \partial_i\phi_a \frac{\partial \mathcal{L}}{\partial(\partial_j \phi_a)} - \delta_{ji}\mathcal{L} \tag{4.13}$$

它们满足守恒方程

$$\frac{\partial T_{0i}}{\partial t} + \sum_j \partial_j T_{ji} = 0 \quad (i = 1,2,3) \tag{4.14}$$

其中 T_{0i} 的物理意义是 x_i 方向的**动量密度**，T_{ji} 是 x_i 方向的**动量流密度**的第 j 分量。此外要注意，这里我们所说的动量就是线动量，而非广义动量，因为我们的上述平移是在直角坐标系下进行的。

最后，我们可以将上面得到的所有量合并为一个矩阵

$$T_{ij}=\begin{pmatrix}T_{00}&T_{01}&T_{02}&T_{03}\\T_{10}&T_{11}&T_{12}&T_{13}\\T_{20}&T_{21}&T_{22}&T_{23}\\T_{30}&T_{31}&T_{32}&T_{33}\end{pmatrix} \tag{4.15}$$

T_{ij} 就是场的**能量-动量张量**。将其写成写成物理量的形式，即：

$$T_{ij}=\begin{pmatrix}\rho_E&\rho_x&\rho_y&\rho_z\\\boldsymbol{S}&\boldsymbol{\pi}_x&\boldsymbol{\pi}_y&\boldsymbol{\pi}_z\end{pmatrix} \tag{4.16}$$

其中 ρ 为荷密度，$\boldsymbol{S}$ 为能流密度，$\boldsymbol{\pi_i}$ 为动量流密度。

空间旋转对称性与角动量

我们先从一个最简单的情形入手，设场的构形绕 z 轴旋转了 $d\theta$ 的小角度。记则对于空间中的位矢，有变换

$$\boldsymbol{x}\to\boldsymbol{x}+d\theta\hat{e}_z\times\boldsymbol{x}$$

不难知道场构型的变化应该为

$$\delta\phi_a=(x\partial_y\phi_a-y\partial_x\phi_a)d\theta$$

以及 Lagrange 密度的变化为

$$\delta\mathcal{L}=\partial_y(x\mathcal{L})d\theta-\partial_x(y\mathcal{L})d\theta$$

从而我们发现 $\boldsymbol{\chi}=(-y\mathcal{L},x\mathcal{L},0)$。最后我们找到了守恒量

$$\begin{cases}\rho_z=\sum_a\frac{\partial\mathcal{L}}{\partial\dot{\phi}_a}(x\partial_y\phi_a-y\partial_x\phi_a)=xT_{02}-yT_{01}\\\boldsymbol{J}_z=\sum_a\frac{\partial\mathcal{L}}{\partial(\partial_j\phi_a)}(x\partial_y\phi_a-y\partial_x\phi_a)-\boldsymbol{\chi}\end{cases}$$

其中 ρ_z 的物理意义是 z 方向的**角动量密度**[7]，$\boldsymbol{J}_z$ 是相应的流密度。

用完全相同的方法，我们进而可以知道场沿着那个和 i 和 j 不同方向（即 k 方向）的总角动量为

$$\mathcal{Q}_{ij} = \int_{\mathbb{R}^3} (x_i T_{0j} - x_j T_{0i}) d^3x \tag{4.17}$$

一些评述

在上面的推演过程中，我们看似只是进行了一些简单的数学微分操作，但是这些微分操作却直指我们的场具有相应的时空对称性，何解？深究其原因，本质上是因为我们的时空是**平直时空**。

将微分操作与时空对称性联系起来的，其实是因为作用量原理允许 Lagrange 密度有一个对时空全微分形式的选择任意性。那么进一步深究，是什么允许 Lagrange 密度允许有这样形式的任意性？是因为作用量（即对 Lagrange 密度积分）中默认包含的度规是 **Euclid 度规**！

Euclid 度规天然保证了时空是均匀且各向同性的，因此在此度规下的场自然满足时空平移、空间旋转的对称性。

此外，读者应该还能察觉到，在质点动力学框架下的 Norther 定理，和场论框架下的证明与讨论，对于时空是有些不同的。在质点动力学中，时间和空间的讨论被刻意分开来，因为他们的处理方式的确不同；但是在场论中，时间和空间由于地位相同——都是场的宗量——而处理方式显得更加一致。

[7] 可以发现，角动量密度和动量密度的关系，就如同角动量与动量的关系一样，这是很有趣的，也是符合直觉的。

附录 A

经典力学简史

本附录的内容翻译自 *Jagdish Mehra* 与 *Helmut Rechenberg* 所著的 *The historical development of quantum theory, Volume I Part I* 开头中的段落，极其干练地讲述了力学的历史发展，其中所提到的诸多概念与理论都在我们的讲义中出现了。通过阅读这一附录，我们可以对力学知识的科学意义有一个历史性地理解，对我们的学习是大有裨益的。

力学是描述物体机械运动规律最古老、最核心的理论体系，它于 1687 年正式建立——Isaac Newton（1642–1727）在《自然哲学的数学原理》一书中，系统阐述了动力学原理与万有引力定律。

18 世纪，Leonhard Euler（1707–1783）等数学家进一步完善了牛顿力学的细节，并将其应用拓展至刚体与流体领域。Joseph Louis Lagrange（1736–1813）最终以极具适配性且优雅的形式，推导了动力学基本方程，建立起分析力学体系。在这一框架中，他融入了诸如“虚功原理”等内容——Jean le Rond d'Alembert（1717–1783）正是借助这一原理描述了运动方程。

由此，力学作为一套成熟的理论体系延续至 19 世纪；当时人们认为，后续只需借助它来解决物理学与天文学中越来越复杂的系统问题。不过，仍有部分数学家试图对其进行补充。1829 年，Carl Friedrich Gauss（1777–1855）提出了一项适用于所有力学问题（无论静态还是动态）的新原理——“最小应力原理”。该原理指出：任意质点系在约束下的运动，会尽可能接近自由运动，即运动过程中所受应力最小。Gauss

本想借此推广 d'Alembert 的早期原理，但由于"应力"概念较为复杂，最小应力原理并未在实际中取代拉格朗日方程。

相比之下，由 William Rowan Hamilton（1805–1865）提出、Carl Gustav Jacob Jacobi（1804–1851）进一步发展的力学体系则更为成功。Hamilton 以 Pierre Louis Moreau de Maupertuis（1698–1759）在 1747 年引入的"最小作用量原理"为起点，该原理认为：自然界中物体的运动总是以作用量最小的路径发生。Maupertuis 将作用量定义为质量 m 与速度 v 的乘积对路径的积分，即 $\int mv\,ds$，并指出实际运动路径中该积分值最小。

Hamilton 则以另一种更具普适性的方式定义了作用量：它是拉格朗日函数 L（动能与势能之差）对时间的积分。在从时刻 t_1 到 t_2 的积分 $\int_{t_1}^{t_2} L\,dt$ 中，实际运动对应的积分值为最小（Hamilton, 1834, 1835）。他证明这一原理可以推导出一种新的运动方程形式——后来被称为哈密顿方程，它与拉格朗日方程等价。Hamilton 还提出了一种求解这些方程的方法，该方法涉及特征函数或作用函数的偏微分方程。

Jacobi 完善了这一方法，并将其用于求解天文学中尤为重要的动力学多体问题（Jacobi, 1866）。在 Jacobi 及后续对力学系统的研究中，人们认识到了运动常数的作用：这些常数对应着满足守恒定律的物理量，其中最重要的守恒量便是能量。能量守恒的发现，也将力学与物理学的另一分支——热力学联系了起来。

附录 B

我与变分法之间的故事

在本讲义的第二章，我们用变分法解决了三个有趣的应用问题，本节我想来分享一下这三个问题和我之间的一些故事。

- 用变分法去理解折射定律这个问题还有另一个名字：**救生员救落水儿童问题**. 我在 2021 年夏天上 bilibili college ，听"郭爷物理"讲这个救生员救落水儿童问题的时候，第一次了解到了变分这个东西，但当时也没怎么仔细想[1].

- 然后悬链线这个问题，是 2020 年冬天，高一上学期的时候我的数学老师讲双曲三角函数的时候提了一嘴，说这个**项链自然悬挂时候的形状就是双曲余弦函数**. 我很受打动，感觉很是神奇. 于是以后一提到双曲三角函数我就想到了老师给我讲的这个生活中的例子，反之亦然.

  后来，2021 年夏天高一下学期学机械能，有一个题是**用手向下拉一个自然悬挂的绳子，问它的重心是升高了还是降低了**. 这个题很具有迷惑性. 后来我的好朋友来问我这个题，我说这个题你可以这么想：自然悬挂的时候一定是势能最低的形状，所以其它的稳定形状的势能全部都比它高. 我还提到说："上学期在数学课上了解到自然悬挂的绳子的形状是双曲余弦，然后你用手拉绳子得到的形状实际上是一个劈尖型的分段函数，我们可以得到它们各自的曲线方程，然后用重心公式算一算比一比就行了. 不过这个计算肯定很复杂，你先记住结论应付一下期末考试，等考完试了咱俩再商量."

[1] 郭老师还讲了什么狗追兔子问题——狗是一道光.

然后就考完期末了，班里在布置家长会的教室，以及给各科老师做贺卡，我的那个好朋友看我坐在位置上无聊，就来找我讨论这个问题了. 不过当时我压根不懂什么变分，也不知道这个悬链线方程（2.12）长啥样，只是知道和双曲余弦有关系而已. 而且，最主要的是我当时还不会给定曲线求重心呢，只是知道个 Pappus 定理求求简单的重心问题. 所以我抓瞎了，就出了个馊主意说我们去找一个绳子做做实验吧. 我们找了半天也没找到，看到班长脖子上有个红色绳子拴着玉石，就问他借，他不借. 最后没办法，就找了个双面胶搓成一条，这咋做实验啊！

所幸最后那个朋友去忙活做贺卡[2]去了，这个问题也就不了了之了. 不知道那位朋友现在弄明白了悬链重心这个问题没有，希望她能看见我的这份讲义.

- 最后是这个最速降线，我一直都知道这个问题催生了变分法，还有很多传奇的历史故事，但是一直没顾上去深究，就一直没懂.

好在这些问题我们现在已经会了不是吗？没有什么比进步更令人开心的了.

[2]我还记得当时我还在那张贺卡上写给化学老师：我下学期一定好好写化学作业.