

# 分析力学

巴鯤兹

2025 年 1 月 20 日





# 前言

这是我关于分析力学的一份讲义。写这份讲义的目的是对费曼学习法的实践。当然,在巩固和深化理解的同时,这份讲义也有一些新的结构和内容。

这份讲义与通常的分析力学讲义不同,不从牛顿力学过渡到 Lagrange 力学再过渡到 Hamilton 力学,而是从态矢量和相空间讲起,给出分析力学的基本方法和思路。随后直接讲解相空间中 Hamilton 力学的理论结构,最后再过渡到位形空间的 Lagrange 力学。在讲完 Hamilton 力学和 Lagrange 力学的基本框架之后,再分别用这两套方法讨论对称性和守恒量的关系,有助于对这两套语言进行对比,并深入理解对称性与守恒量的关系。最后,利用分析力学的语言简单讲解中心力场,微振动和刚体以及连续体系的分析力学。通过这套讲法,有助于建立起不局限于力学的现代的分析力学的理论框架,并且采用假设-理论-实验验证的思路讲解,可以很好地过渡到电动力学,统计力学和量子力学。

这份讲义其实没有什么特定的受众,因为主要动机是写给自己的。但是,这份讲义或许对想要提纲挈领地了解分析力学的读者有帮助:理论内容完整且讲解清晰,但对于一些技术没有进行介绍(例如微扰法等内容),换来了较短的篇幅。内容更丰富的讲义推荐高显教老师的《经典力学》或者李岩松老师的《分析力学讲义》<sup>1</sup>,如果读者对分析力学中更深入的数学结构有兴趣,推荐陈童老师的《经典力学新讲》。陈童老师的讲义可以在他的网站上免费获取。

在完成这份讲义时,我相对注重物理图像的构建。经过一段数学推导后,要充

---

<sup>1</sup>李岩松老师的讲义目前还没有开源,对能够获取的读者推荐。这本书是我学习分析力学的教材,内容非常丰富。

分理解每一步在干什么, 不能对物理图像失去跟踪, 否则数学推导很难进行下去.

目前这份讲义是第一版, 可能有许多错误, 欢迎读者来进行交流. 这份讲义的后续更新会在我的个人网站上发布, 可能会补充更多的例子或者内容 (但尽量保持简洁) 并修改一些错误.

我的个人网页: <https://kunziba.github.io/>.



# 目录

前言	i
<b>1 分析力学的基本概念</b>	<b>1</b>
1.1 广义坐标与系统的态	1
1.1.1 广义坐标	1
1.1.2 系统的态	2
1.2 相空间	3
1.3 约束	4
1.3.1 约束的分类和系统的自由度	5
1.3.2 完整约束的判断	7
1.3.3 关于约束的一些讨论	9
<b>2 最小作用量原理和 Hamilton 正则方程</b>	<b>11</b>
2.1 相空间的最小作用量原理	11
2.1.1 作用量和 Hamiltonian	11
2.1.2 最小作用量原理	14
2.2 Hamilton 正则方程	14
<b>3 Hamilton 力学及其正则形式</b>	<b>17</b>
3.1 Poisson 括号	18
3.1.1 Poisson 括号的几何解释	18

3.1.2	力学量的导数 . . . . .	19
3.1.3	Poisson 括号的性质 . . . . .	21
3.1.4	基本力学量的 Poisson 括号 . . . . .	22
3.2	Liouville 定理 . . . . .	24
3.3	正则变换 . . . . .	26
3.3.1	正则变换的几何含义 . . . . .	27
3.3.2	生成函数 . . . . .	29
3.4	Hamilton-Jacobi 方程 . . . . .	32
3.4.1	Hamilton-Jacobi 方程 . . . . .	32
3.4.2	Hamilton 主函数的确定 . . . . .	34
3.4.3	Hamilton-Jacobi 方程的分离变量 . . . . .	35
<b>4</b>	<b>Lagrange 力学</b>	<b>37</b>
4.1	位形空间的最小作用量原理 . . . . .	37
4.1.1	位形空间 . . . . .	37
4.1.2	位形空间的最小作用量 . . . . .	38
4.2	Lagrange 方程 . . . . .	39
4.3	Lagrange 力学与 Hamilton 力学 . . . . .	40
4.3.1	Legendre 变换 . . . . .	40
4.3.2	Routh 函数 . . . . .	42
4.4	Lagrange 力学与 Newton 力学 . . . . .	43
4.4.1	d'Alembert 原理 . . . . .	43
4.4.2	第一类 Lagrange 方程 . . . . .	45
4.5	Lagrange 力学的广义协变性 . . . . .	47
<b>5</b>	<b>对称性与守恒量</b>	<b>51</b>
5.1	循环坐标 . . . . .	51
5.2	无穷小正则变换 . . . . .	52

5.2.1	对称性 . . . . .	52
5.2.2	无穷小正则变换 . . . . .	54
5.3	诺特定理和规范对称 . . . . .	56
5.3.1	Lagrange 力学的规范对称性 . . . . .	56
5.3.2	诺特定理 . . . . .	56
<b>6</b>	<b>分析力学的应用</b>	<b>61</b>
6.1	有心力场 . . . . .	61
6.1.1	二体约化和守恒量 . . . . .	61
6.1.2	开普勒问题 . . . . .	62
6.1.3	LRL 矢量 . . . . .	62
6.2	微振动 . . . . .	62
6.2.1	一维谐振子 . . . . .	62
6.2.2	多自由度体系在平衡位置附近的运动 . . . . .	62
6.2.3	参数共振 . . . . .	62
6.3	刚体的运动 . . . . .	62
6.3.1	转动 . . . . .	62
6.3.2	惯量张量 . . . . .	62
6.3.3	刚体运动的两种可解情况 . . . . .	62
<b>7</b>	<b>连续体系的分析力学</b>	<b>63</b>
<b>A</b>	<b>变分法</b>	<b>65</b>
A.1	E-L 方程 . . . . .	65
A.2	多自由度变分 . . . . .	67
A.3	约束变分 . . . . .	67
A.4	几何变分 . . . . .	68

<b>B 微分形式</b>	<b>71</b>
B.1 微分形式的基本概念 . . . . .	71
B.2 微分形式的计算与 Stokes 定理 . . . . .	71





# Chapter 1

## 分析力学的基本概念

### 1.1 广义坐标与系统的态

#### 1.1.1 广义坐标

分析力学的创立的初衷是想构造一套与牛顿力学等价的由标量描述方程的力学体系. 矢量方程可以转化为分量的标量方程, 所以如果仅仅是建立某种标量方程, 那么相较牛顿力学就体现不出任何优势. 所以, 用更准确的话说, 分析力学是要建立一套有着相对更简单的标量方程的理论. 牛顿矢量力学处理一些机械的动力学问题带来的麻烦是, 约束会带来更多的未知数和方程. 但从物理直观上讲, 约束越多, 系统可能的行为就越少, 进而可能更好求解.

所以为此, 分析力学引入了**广义坐标**, 是任意可以描述系统的参数, 进而量纲也是不固定的. 常见的广义坐标有角度  $\theta$  (用于处理球对称体系) 和弧长  $s$  (用于处理单自由度运动). 后来发现, 广义坐标的引入是伟大的, 因为广义坐标不仅可以描述质点位置以求解各种机械的动力学, 更是可以推广到其他物理现象. 比如说, 电路体系中的电荷  $q$  就可以作为一种广义坐标. 更进一步地, 电磁场都可以视作一种广义坐标. 这使得分析力学的威力不仅仅局限于机械力学, 更是在电磁场理论, 热学乃至更深入的其他物理领域发挥着作用.

回到质点的力学. 对一组多质点的体系 (也可以包含刚体), 我们引入广义坐标  $q^1, \dots, q^s$ . 使得任意一个质点的位置矢量可以被写成  $q_1, \dots, q_s$  的函数:

$$\mathbf{r}_i = \mathbf{r}_i(q^1, \dots, q^s, t) \quad (1.1)$$

对于广义坐标有如下几个要求: 首先, 不同的广义坐标必须是完全独立的, 即

$$\frac{\partial q^i}{\partial q^j} = 0 \quad i \neq j \quad (1.2)$$

第二个要求是一组广义坐标必须能唯一确定系统的位形.

对于满足上述要求的广义坐标的个数我们目前还无法讨论. 但是通过单摆作为例子可以发现, 我们只用两个角度 (纬度角和经度角) 作为满足要求的广义坐标就可以完全确定摆球的位置, 比直角坐标的  $(x, y, z)$  要少一个参量. 这说明对于存在一定约束的体系, 适当的广义坐标的选取可以简化为系统位形的描述.

### 1.1.2 系统的态

如果想要描述系统在某一个瞬间的状态, 我们需要同时确定系统的位形和运动状态 (这是经典力学的基本假设, 即确定一个系统的态, 我们就知道态的瞬时变化率, 进而知道下一个瞬间的态, 从而完成对系统之后运动的预测). 对于系统的运动状态, 一个比较自然的想法是用系统中每个质点的速度来表示, 如果用广义坐标写的话:

$$\frac{d\mathbf{r}_i}{dt} = \frac{\partial \mathbf{r}_i}{\partial t} + \sum_{k=1}^s \frac{\partial \mathbf{r}_i}{\partial q^k} \dot{q}^k = \mathbf{v}_i(q^1, \dots, q^s, \dot{q}^1, \dots, \dot{q}^s) \quad (1.3)$$

这样, 对于一个系统, 我们通过  $2s$  个变量就可以确定系统的运动状态. 或者我们写成一组数  $(q^1, \dots, q^s, \dot{q}^1, \dots, \dot{q}^s)$ , 可以确定某个  $2s$  维空间的一个点. 换句话说, 这个空间中的某个点对应系统的一个运动状态. 这个空间就是我们所说的**速度相空间**, 它是描述 Lagrange 力学所需要的空间. 但注意, 这组数可能并不构成一个矢量. 因为在直角坐标下广义坐标的表达式可能并不是线性的, 这会导致对两个

“矢量”的加法不能定义, 进而, 我们所说的这个  $2s$  维空间也不一定是一个线性空间, 而是一个流形.

但是, 在引入广义坐标的时候我们就已经提到, 分析力学可以用于描述更广泛的物理系统. 对于非机械的系统, 速度可能是没有什么物理含义的. 进而, 如果我们想描述系统的“运动状态”, 能够描述不同系统及其相互作用的动量是一个更好的选择. 因此, 现代物理中, 我们广泛使用位置和动量来描述系统的状态. 由广义坐标和广义动量 (我们现在还没有定义, 但可以描述其性质, 即描述不同系统的相互作用, 并且目前与广义坐标独立) 所组合起来的一组数

$$(q^1, \dots, q^s, p_1, \dots, p_s) \quad (1.4)$$

被称为系统的态 (state). 在量子力学中, 这对应系统的态矢量  $|\psi\rangle$  (这时候我们可以说这是一个 Hilbert 空间中的矢量, 因为这是量子力学的基本假设之一).

## 1.2 相空间

刚刚提到, 系统的状态可以表示为一个空间中的代表点. 这个空间就是系统的相空间. 系统的状态随时间演化, 在相空间中这些连续的代表点构成的集合就是相轨迹. 我们知道, 按照经典力学“预测”的观点, 如果给定了系统运动所遵守的规律, 那么给定一个初始状态, 其之后的演化一定是唯一的. 也就是说, 不同的相轨迹是不相交的. 否则取这个交点, 那么系统随后的演化状态将会由两种可能性, 这不符合我们的生活常识.

**Example 1.1.** 谐振子的相轨迹 由于我们现在还没有在分析力学的框架下导出任何的规律, 所以假设我们知道牛顿运动方程, 并且知道动量和速度的关系:

$$p = m\dot{x} \quad (1.5)$$

容易知道谐振子的解:

$$\begin{cases} x = A \sin \omega t, \\ p = m\dot{x} = \omega m A \cos \omega t \end{cases} \quad (1.6)$$

消去三角函数我们就得到了相轨迹的方程:

$$\frac{p^2}{2m} + \frac{m\omega^2 x^2}{2} = \frac{m\omega^2 A^2}{2} \quad (1.7)$$

容易发现, 随着  $A$  的变化, 这是一族离心率确定的椭圆. 并且容易验证, 这一族椭圆中不同的椭圆是没有交点的.

在之后的分析中, 相空间都将是主要的“阵地”, 并且随着更多系统的运动规律的导出, 我们会发现相空间还有具有其他的结构的性质, 这部分内容将在之后讨论.

### 1.3 约束

一个很自然的问题是, 目前我们讨论的相空间的任意一个点是否都能代表系统的一个状态, 或者说, 我们关注的系统遍历所有可能的位形, 能否跑遍相空间中的所有点? 并且, 在之前, 我们在描述广义坐标的个数的时候一直用  $s$  表示. 对于一个陌生的体系, 我们甚至无法给出相空间的维数. 所以本节我们讨论分析力学中极为重要的概念——约束.

通过单摆的例子我们知道, 对于存在一定约束的体系, 独立的广义坐标的个数要比三维直角坐标系下描述粒子的坐标的  $(x, y, z)$  要少. 通过数学化约束的表达, 我们可以探讨  $s$  的值.

### 1.3.1 约束的分类和系统的自由度

从物理上讲, 约束一般是对系统的位形进行的约束. 也就是一些广义坐标和广义速度需要满足一定的条件. 我们可以同意用数学上的函数来表达:

$$\Phi(q, \dot{q}, t) = 0 \quad (1.8)$$

写下这个式子后, 有人会 argue 说, 有些物理上的约束无法写成这个形式. 比如说连接在一端固定的软绳下的小球, 绳子未绷紧的情况是完全有可能发生的, 这时候约束就应该写成:

$$\sqrt{x^2 + y^2 + z^2} \leq l \quad (1.9)$$

这个约束就不能写成等式的形式. 类似的例子还可以举出很多. 一般来说, 我们可以将这种约束写成:

$$\Phi(q, \dot{q}, t) \geq 0 \quad (1.10)$$

从这个角度我们就可以发现, 不同情况下的约束, 其性质可能是截然不同的. 这个情况中的取等与否, 就可以引出约束的一个分类: 对形如式 (1.8) 的约束, 我们称之为**双侧约束**, 而对于形如式 (1.10) 的约束, 我们称之为**单侧约束**. 从形式上就可以看出, 双侧约束比单侧约束要好处理很多, 因此我们之后主要讨论双侧约束.

接下来我们来看约束中是否含  $t$ . 从物理直观上讲, 若一个系统的约束含  $t$ , 那么这个系统就需要受到一个运动“机器”来构造这种约束. 例如说一个质点的约束有如下形式:

$$y = vt \quad (1.11)$$

我们就发现, 质点似乎处在一个正在以匀速  $v$  向  $y$  轴正方向的管子里. 如果说我们真的要做实验来构造这样一个系统, 就需要建造一个能够匀速运动的管子. 但是, 对于一个单摆的约束:

$$x^2 + y^2 + z^2 = l^2 \quad (1.12)$$

这时候我们就完全不需要再去人为地干预这个系统. 用更物理的一点语言来说, 是否“人为干预”造成的区别就是, 人为干预会与系统构成能量交换, 而非人为

干预的系统的能量是守恒的. 也就是从能量的表现形式上, 这两种约束下的系统的行为有着区别. 这样, 对于不显含时间的约束:

$$\frac{\partial \Phi}{\partial t} = 0 \quad (1.13)$$

我们称之为**定常约束**, 反之, 若约束满足

$$\frac{\partial \Phi}{\partial t} \neq 0 \quad (1.14)$$

我们称之为**非定常约束**.

从物理直观上看, 约束可以减少系统自由度 (我们还没有严格定义这个概念, 粗略地说这是一个衡量系统可能行为多少的指标) 的个数, 也就是会减少独立的广义坐标的个数. 但是, 从数学上说, 只有形如

$$\Phi(q^1, \dots, q^s, t) = 0 \quad (1.15)$$

这样的关于广义坐标的方程才能切实地减少系统广义坐标的个数. 因此对于这样的约束, 我们称之为**完整约束**. 反过来, 如果约束形如

$$\Phi(q, \dot{q}, t) = 0 \quad (1.16)$$

那么, 从数学上说我们不保证能够减少广义坐标的个数. 这种约束我们称之为**非完整约束**. 这种约束会带来这样一种很微妙的情况, 我们需要一些广义坐标, 不妨先令他们的个数为  $m$ , 但这些广义坐标却不是完全独立的. 然而, 由于约束并非完整约束, 我们又不能继续减少广义坐标的个数的同时满足对广义坐标的要求 (唯一确定系统的位形). 其原因在于, 从数学上考虑  $\dot{q}^s, q^s$  是没有代数上的关系的 (粗略地说, 加减乘除或者其他函数能够表示这二者的关系), 但是  $\dot{q}^s$  又确实是  $q^s$  的导数.

这样, 我们会更喜欢完整约束, 因为其能直接地减少系统广义坐标的个数. 这样, 通过讨论完整约束的个数, 我们就能严格地给出系统自由度的概念. 系统的**自由度**是指, 能够唯一确定系统位形的最少的广义坐标的个数. 从数学上具体计算,

若完全不考虑约束, 系统的自由度的个数为  $F$  (例如一个  $N$  例子系统, 这个值为  $3N$ ), 所有**双侧完整约束**的个数为  $r$ , 则考虑约束的系统的自由度为:

$$f = F - r \quad (1.17)$$

根据定义, 我们就知道了之前的  $s = f$ . 以后我们就把广义坐标下角标的最大值改写为  $f$ .

### 1.3.2 完整约束的判断

一个例子告诉我们, 并不是说只要约束中存在  $\dot{q}^s$ , 那么这种约束就一定是非完整约束. 考虑一个二维的纯滚的轮子, 纯滚的约束为

$$\dot{x} = R\dot{\theta} \quad (1.18)$$

那这样一个约束是否是完整约束呢? 直观上讲不是, 但是如果我们对这个式子两边积分, 就得到

$$x = R\theta + C \quad (1.19)$$

是一个完整约束! 也就是说, 如果一个非完整约束可以写成一个只与  $(q, t)$  有关的函数的全导数:

$$\Phi(q, \dot{q}, t) = \frac{d\Pi(q, t)}{dt} = 0 \quad (1.20)$$

那我们就可以通过积分将其化成一个完整约束.

对于一个一般的约束, 判断其是否是完全可积是较为困难的. 好在我们最常遇到的一类带有广义速度的约束关于广义速度是线性的. 进而约束就可以写成:

$$\omega = \Phi dt = \sum_i \Phi_i(q, t) dq_i + \Phi_0(q, t) dt \quad (1.21)$$

这是一个 1-Form. 根据之前介绍的微分形式的知识, 如果  $\omega = \Phi dt = dF$ , 那么必然有

$$d(\Phi dt) = 0 \quad (1.22)$$



但是在这个问题中, 我们并不一定要求  $\Phi$  是某个函数的全微分. 这是因为  $\Phi dt = 0$ , 我们完全可以再乘上任意一个函数 (或者说是积分因子)  $Q$  使得

$$Q\Phi dt = dF(q, t) \quad (1.23)$$

这样也可以将这个约束化为完整约束. 因此, 根据外微分的知识可以证明, 这样的积分因子的存在的充要条件是

$$\omega \wedge d\omega = 0 \quad (1.24)$$

我们下面通过具体的计算给出一个相对普遍的公式. 设约束对应的 1-Form 为

$$\omega = \sum_{i=0}^s a_i dq^i \quad (1.25)$$

其中我们规定  $t = q^0$  (这个规定在讨论之后的相空间中也会用到, 从推导中我们发现很多时候时间似乎广义坐标有着相同的地位).

$$\begin{aligned} d\omega &= d \left( \sum_i a_i dq^i \right) \\ &= \sum_{i \neq j} \frac{\partial a_i}{\partial q_j} dq_j \wedge dq_i \end{aligned} \quad (1.26)$$

继续计算其与  $\omega$  的外积,

$$\omega \wedge d\omega = \left( \sum_i a_i dq_i \right) \wedge \left( \sum_{j \neq k} \frac{\partial a_k}{\partial q_j} dq_j \wedge dq_k \right) \quad (1.27)$$

计算过程中, 我们对三个指标的不同排列分别计算 6 次, 而其通过排列可以合并, 只不过换序一次指标要加一个负号, 上述求和可以写成:

$$\sum_{0 \leq i < j < k \leq s} \left( a_i \left( \frac{\partial a_j}{\partial q_k} - \frac{\partial a_k}{\partial q_j} \right) + a_j \left( \frac{\partial a_i}{\partial q_k} - \frac{\partial a_k}{\partial q_i} \right) + a_k \left( \frac{\partial a_j}{\partial q_i} - \frac{\partial a_i}{\partial q_j} \right) \right) dq_i \wedge dq_j \wedge dq_k \quad (1.28)$$

因而可积的条件为:

$$a_i \left( \frac{\partial a_j}{\partial q_k} - \frac{\partial a_k}{\partial q_j} \right) + a_j \left( \frac{\partial a_i}{\partial q_k} - \frac{\partial a_k}{\partial q_i} \right) + a_k \left( \frac{\partial a_j}{\partial q_i} - \frac{\partial a_i}{\partial q_j} \right) = 0 \quad (1.29)$$

其中  $i, j, k$  为任意三个按大小排列的指标. 一个约束是完全可积的要求这满足所有的方程.



### 1.3.3 关于约束的一些讨论

我们对约束的分类暂且到这. 事实上, 约束还有更多的分类, 就是理想约束与非理想约束. 但这里就需要额外引入力学结构来对其进行讨论. 我们目前对约束的讨论只局限于系统的位形上, 不涉及到更深入的力学结构. 通过对约束的讨论, 我们可以确定系统的自由度, 进而确定研究一个系统所需要的广义坐标的个数.

在讨论约束中的分类的过程中, 我们见到了很多性质比较差的约束, 对于这些约束我们缺乏讨论. 客观上的一个原因是, 这些性质比较差的约束很难讨论. 但这样会引发人的不安: 那我们该如何用分析力学的范式来求解这些体系呢? 这里就涉及到对分析力学一个观念上的看法. 分析力学不是以求解问题为导向的 (尽管 Lagrange 力学在处理一些性质良好的力学体系相对牛顿力学有较大的优势), 而是以建立一种能够广泛描述物理体系所建立的, 这一点在开篇我们就提到过. 这时, 我们所要讨论的是物理现象和规律. 我们发现, 目前一些性质较差的约束的本质不过是基本作用的宏观体现或者是人为创造的一些系统有着不那么“自然”的约束. 但是, 这些性质较差的约束在我们研究最基本的物理规律时并不存在. 我们研究的基本规律与体系都是性质良好的.

这段讨论也可以解答后续的一些读者可能的疑惑, 我们会对系统的性质逐渐地进行限制. 消除对系统性质作出要求这种不安的有力工具, 就是逐步地建立起上述对分析力学的看法. 至此, 我们已经搭建好了分析力学中“观念”性的基本看法, 之后我们就会建立在这套语言下的规律, 并探讨这些规律带来的结构.



## Chapter 2

# 最小作用量原理和 Hamilton 正则方程

## 2.1 相空间的最小作用量原理

### 2.1.1 作用量和 Hamiltonian

我们说过, 系统随时间的演化可以表征为相空间中代表点的运动, 这会在相空间中形成一条相轨迹. 这条相轨迹从数学上说是一条一维曲线, 因此一个比较便捷的方式是引入一个自由参数, 将  $p, q$  都表示成这个自由参数的函数. 因此, 从数学上说, 相轨迹就是:

$$\begin{cases} p_i = p_i(t) & i = 1, \dots, f \\ q^i = q^i(t) & i = 1, \dots, f \end{cases} \quad (2.1)$$

这个参数理论上是可以任意选取的, 但从物理上我们通常选择时间, 这样相轨迹的参数方程会有比较明确的物理含义——就是系统的动量和坐标随着时间变化的函数.

真实世界中, 根据经典力学的假设, 我们给定一个系统的初始态, 那么我们就可以通过一套规律相应地确定系统后续的演化状态. 牛顿力学告诉我们, 系统的

初始态就是系统的二阶微分方程组的初始条件. 在分析力学中, 我们选取的确定微分方程组解的条件是边界条件, 即一段运动的初始和结束时系统的位形. 这样, 我们可以断言, 对一段给定初始和结尾时位形的系统在这个过程中的运动是唯一确定的.

下面我们的任务就是找出这个唯一确定的运动所满足的方程. 在之前我们引入了泛函的数学工具, 这个工具是说, 我们输入一个函数, 让某个这个函数对应的泛函取得极值的条件就能给出这个这个输入函数需要满足的条件. 我们注意到这个问题和我们需要解决的问题很像: 我们输入一个轨迹函数, 接下来我们要找其满足条件 (这个条件就是要求我们解出来的相轨迹要与真实世界的运动符合) 的微分方程. 这样就引发一个构想, 我们是否可以找到一个相轨迹的泛函  $S$ , 使得相轨迹为真实运动的情况下这个泛函取极值.

当然, 这只是一种构想, 满足这样的泛函完全有不存在的理由. 但好在自然足够美, 其运动规律能让我们找出这样一个泛函来 (但又有一些系统的性质足够差, 例如这个系统的约束是未完整的或者单侧的, 我们不能用 Hamilton 力学描述这样的系统). 这个泛函就是我们所说的**作用量**. 至于这个泛函是怎样找到的, 必须要诚实的说, 历史上我们是根据牛顿力学总结出的规律经过数学上的推导找到的, 接下来推广到麦克斯韦的电磁学, 通过麦克斯韦的方程组我们又可以通过数学上的推导找到电磁场的作用量. 对于量子力学就不再细说了. 但这三个领域的成功经验告诉我们, 一个陌生的物理体系应当 (这只是一种归纳) 存在一个作用量. 这也是分析力学的普遍之处, 现代物理研究问题都是考虑如何写出系统的作用量. 话说回来, 我们用数学的语言来描述作用量, 这个泛函可以写成

$$S = \int_{t_1}^{t_2} L dt \quad (2.2)$$

其中  $L$  是一个我们目前未知的函数 (这里要提醒熟知 Lagrange 力学的读者, 这个量和 Lagrangian 的关系有些微妙, 并不完全等同, 再之后介绍 Lagrange 力学的时候我们会进行相对详细的讨论). 一旦确定  $L$ , 我们就可以找到一个与之对应

的函数  $H$ , 使得

$$H = \sum_i \dot{q}^i p_i - L \quad (2.3)$$

进而将作用量写成

$$S = \int_{t_1}^{t_2} \left( \sum_i \dot{q}^i p_i - H \right) dt \quad (2.4)$$

而不改变任何规律的表达. 至于我们为什么不直接讨论  $L$  而要通过  $H$ , 是因为我们发现  $H$  具有更为人熟知的物理含义——能量. 在分析力学中我们称这个函数为 **Hamiltonian**(哈密顿量).

目前我们所作的一切都是数学上的一些等价变换. 接下来我们给出一些系统的 Hamiltonian 的例子, 试图构造一些符合牛顿力学的物理图像. 首先, 自由粒子的 Hamiltonian 可以写成

$$H = \frac{p^2}{2m} \quad (2.5)$$

其中  $m$  是一个正的数, 牛顿力学中的能动量关系告诉我们这代表了粒子的质量. 对于存在相互作用 (这种相互作用能用势能函数表达) 的多粒子系统的 Hamiltonian 是:

$$H = \sum_i \frac{p_i^2}{2m_i} + V(q^1, \dots, q^f) \quad (2.6)$$

以及相对论下自由粒子的 Hamiltonian:

$$H = \sqrt{p^2 c^2 + m^2 c^4} \quad (2.7)$$

可以看出, 我们举出的例子中  $H$  就是牛顿力学中的能量函数. 后续, 随着见到的“奇怪”的体系越来越多, 我们会发现  $H$  并不一定等于牛顿力学中系统的能量函数 (但量纲上一定是能量的量纲). 但是, 对于  $H$  的物理图像上, 将其理解为系统的能量是完全没有问题的.

### 2.1.2 最小作用量原理

这样, 根据之前的讨论, 当我们找到一个系统的 Hamiltonian 后, 物理规律就可以表示为

$$\delta S = \delta \int_{t_1}^{t_2} \left( \sum_i \dot{q}^i p_i - H(p, q, t) \right) dt = 0 \quad (2.8)$$

这就是物理中的**最小作用量原理**. 这个名字有一个奇怪的地方, 就是从数学上  $\delta S = 0$  不能推出  $S$  取最小值, 甚至我们都无法说它取极值, 只能说它取驻值. 但是, 这种叫法比较符合我们的“物理直觉” (当然这种直觉的根据并不充分), 或者说对自然规律的美的追求, 就是系统自然的运动会选取一条“代价”最小的路, 这很符合我们的行为逻辑<sup>1</sup>.

接下来有必要强调的一点是, 我们是如何取变分的. 我们取的变分是, 系统在  $t_1$  和  $t_2$  处**位形确定**下的变分, 这时候  $p$  在初始和末尾的值都是不固定的. 并且, 取变分时  $t_1, t_2$  是固定的.

## 2.2 Hamilton 正则方程

可以说, 我们对 Hamilton 力学的基本原理的介绍已经结束了. 但是, 这套语言很显然不是很方便, 毕竟我们要做的是泛函的变分. 相比之下我们还是更喜欢解微分方程. 接下来我们就通过取变分来导出 Hamilton 力学体系下最重要的一套方程. 我们对  $S$  取变分:

$$\delta S = \sum_i \int_{t_1}^{t_2} \left( p_i \delta \dot{q}^i + \dot{q}^i \delta p_i - \frac{\partial H}{\partial q^i} \delta q^i - \frac{\partial H}{\partial p_i} \delta p_i \right) dt \quad (2.9)$$

利用变分和导数的可交换性我们有

$$\delta \dot{q}^i = \frac{d}{dt} \delta q^i \quad (2.10)$$

---

<sup>1</sup>最小作用量原理的正确性还有一个原因, 就是最小作用量原理满足物理规律的协变性. 这一点需要在 Lagrange 力学中讨论.

这样根据我们导出 Lagrange 方程时的一个 trick:

$$p_i \delta \dot{q}^i = p_i \frac{d}{dt} \delta q^i = \frac{d}{dt} p_i \delta q_i - \dot{p}_i \delta q_i \quad (2.11)$$

带回就有:

$$\begin{aligned} \delta S &= \sum_i \int_{t_1}^{t_2} \left( \frac{d}{dt} p_i \delta q_i - \left( \frac{\partial H}{\partial q^i} + \dot{p}_i \right) \delta q^i - \left( \frac{\partial H}{\partial p_i} - \dot{q}^i \right) \delta p_i \right) dt \\ &= \sum_i p_i \delta q_i \Big|_{t_1}^{t_2} - \sum_i \int_{t_1}^{t_2} \left( \left( \frac{\partial H}{\partial q^i} + \dot{p}_i \right) \delta q^i - \left( \frac{\partial H}{\partial p_i} - \dot{q}^i \right) \delta p_i \right) dt \\ &= - \sum_i \int_{t_1}^{t_2} \left( \left( \frac{\partial H}{\partial q^i} + \dot{p}_i \right) \delta q^i - \left( \frac{\partial H}{\partial p_i} - \dot{q}^i \right) \delta p_i \right) dt \end{aligned} \quad (2.12)$$

这其中我们已经利用了  $\delta q^i(t_1) = \delta q^i(t_2) = 0$  的要求. 这样, 由于所有的不同指标的  $\delta q^i$  和  $\delta p_i$  都是无关且任意的, 这要求

$$\begin{cases} \dot{p}_i = -\frac{\partial H}{\partial q^i} \\ \dot{q}^i = \frac{\partial H}{\partial p_i} \end{cases} \quad \forall i = 1, \dots, f \quad (2.13)$$

这就是著名的 **Hamilton 正则方程**, 他给出关于系统的  $2f$  个微分方程.

接下来对 Hamilton 方程进行一些直观的理解. 首先看第一个方程. 我们取自由粒子的 Hamiltonian, 代入后就可以得到

$$\dot{x} = \frac{\partial H}{\partial p_x} = \frac{p_x}{m} \Rightarrow p_x = m\dot{x} \quad (2.14)$$

这给出经典自由粒子的动量关于速度的关系. 之前在引入动量后我们只给出了直观的理解, 但并没有给出其具体形式. 这是因为在知道 Hamilton 正则方程一些, 动量和坐标是完全独立的, 只有通过 Hamilton 方程我们才可以给出关于动量和坐标导数之间的关系.

仍然考虑自由粒子, 第二个方程给出

$$\dot{p} = -\frac{\partial H}{\partial q} = 0 \quad (2.15)$$

结合第一个式子我们就知道, 自由粒子静止或做匀速直线运动. 这就是牛顿第一定律. 或者, 我们认为粒子在外场  $V(x, y, z)$  中, 那么这个粒子的 Hamiltonian 改写为

$$H = \frac{p_x^2 + p_y^2 + p_z^2}{2m} + V(x, y, z) \quad (2.16)$$

$$\dot{p}_x = -\frac{\partial H}{\partial x} = -\frac{\partial V}{\partial x} \quad (2.17)$$

其他两个分量也有一样的方程. 如果我们定理势能的负梯度为外场施加给粒子的力, 那么三个 Hamilton 方程写成矢量形式就是

$$\dot{\mathbf{p}} = -\nabla V = \mathbf{F} \quad (2.18)$$

这就是牛顿第二定律. 这样, 我们发现由 Hamilton 方程 (假设知道 Hamiltonian) 就可以导出牛顿力学 (读者可以自行思考 Hamilton 方程是否可以“导出”牛顿第三定律?). 之后再介绍过 Lagrange 力学之后, 我们会发现也可以从牛顿力学出发推导出整个分析力学体系. 因此我们说分析力学的经典部分和牛顿力学是等价的.



## Chapter 3

# Hamilton 力学及其正则形式

在本章中我们主要讨论由 Hamilton 力学衍生出的相空间的一些结构.

在物理学的问题研究中, 守恒量是极为重要的概念. 对于问题求解来说, 守恒量能直接带给我们代数方程而非微分方程; 从物理概念上来说, 守恒量意味着系统中不变的量, 这种不变量其实充分标记了这个系统重要的特征. 从数学上来说, 一个矢量函数  $\mathbf{f}$  或者标量函数  $f$  是守恒的, 意味着其关于时间的导数为零, 即

$$\frac{dF}{dt} = 0 \quad (3.1)$$

更进一步地, 根据经典力学的假设, 我们关注守恒的**动力学量**, 即

$$F = F(p, q, t) \quad (3.2)$$

其中不出现广义坐标和动量的导数. 也就是说, 我们给定系统的一个态和对应时间, 动力学量是完全确定的. 这在热力学统计物理中就是所谓的态函数. 我们之所以更关心动力学量, 是因为动力学量只与系统的态有关, 进而其是系统固有性质 (无关运动规律) 的一个体现. 通过研究动力学量的性质, 就可以在某一个方面研究系统的性质. 因此我们接下来对相空间结构探索的动机就是为了找出守恒量.

第一节的 Poisson 括号源自于对守恒量所要求的几何结构; 第二节的 Liouville 定理中我们找出了相空间中相流的重要守恒量, 这在处理多体系统中非常重

要; 第三节我们试图通过对广义坐标和动量进行某种不破坏系统动力学性质的变换得到守恒量, 这就是所谓的“正则变换”. 我们发现得到守恒量后能够极大减少我们求解方程的数量, 最终我们只需要求解 Hamilton-Jacobi 方程.

## 3.1 Poisson 括号

### 3.1.1 Poisson 括号的几何解释

从几何上说, 我们该如何判断一个力学量是否在真实轨迹上守恒呢. 我们注意到, 力学量是定义在相空间上的多元函数, 守恒代表着真实的相轨迹是该多元函数的等值曲线. 进一步转化为数学一些的语言是说, 力学量在相空间上的梯度应该垂直于相轨迹的演化方向, 也就是相流.

关注到, Hamilton 正则方程天然就有 Hamiltonian 在相空间中梯度的结构. 所以我们定义

$$\mathbf{j} = (\dot{q}^1, \dots, \dot{q}^f, \dot{p}_1, \dots, \dot{p}_f)^T \quad (3.3)$$

为相流矢量 (仍然提醒, 态不是矢量, 但由于相空间是一个  $2f$  维的流形, 因此其切空间是一个线性空间, 这样态的导数就是一个矢量), 以及 Hamiltonian 的梯度

$$\nabla H = \left( \frac{\partial H}{\partial q^1}, \dots, \frac{\partial H}{\partial q^f}, \frac{\partial H}{\partial p_1}, \dots, \frac{\partial H}{\partial p_f} \right)^T \quad (3.4)$$

那么 Hamilton 正则方程就是

$$\mathbf{j} = \omega \nabla H \quad (3.5)$$

其中  $\omega$  是一个矩阵, 其具有如下的形式

$$\omega = \begin{bmatrix} 0 & \mathbf{I}_f \\ -\mathbf{I}_f & 0 \end{bmatrix} \quad (3.6)$$

如果我们计算关于经过该矩阵变换前后的标准内积, 即

$$\sum_{\alpha, \beta} v_\beta \omega^{\alpha\beta} v_\alpha = 0 \quad (3.7)$$

如果我们利用这个矩阵定义一个内积 (广义的) 的话, 要求相同矢量和自己的内积为零, 我们发现这似乎更像是通过内积得到一个面积 (目前先不深入讨论内积出面积的深刻内涵, 这其实表明相空间具有“辛结构”, 也就是带有面积的结构). 在理解这个矩阵的几何意义之后, 我们发现 Hamilton 正则方程是说, Hamiltonian 的梯度垂直于相流的方向, 也就是说 Hamiltonian 守恒! 但注意, 我们进行上述推论的基础是 Hamiltonian 只是  $p, q$  的函数, 这样 Hamiltonian 不显含  $t$ , 但普遍定义的 Hamiltonian 是可以含时的. 这样我们就得出了一个结论, 不含时的 Hamiltonian 守恒. 我们之后会用一个完全不同的视角得到这个结果.

这样, Hamiltonian 利用其正则方程给我们竖了一个“标杆”, 只要  $F$  的梯度和 Hamiltonian 梯度张成的面积为零, 那么  $F$  的梯度就自然垂直于相流的轨迹, 进而得到  $F$  守恒. 具体进行计算, 这个面积为

$$[F, H] = \sum_{\alpha, \beta} (\nabla F)_\beta \omega^{\alpha\beta} (\nabla H)_\alpha \quad (3.8)$$

我们将其定义为 Poisson 括号. 对于任意两个力学量的 Poisson 括号, Poisson 括号的几何意义就是两个力学量梯度张成面积的大小.

### 3.1.2 力学量的导数

我们在这节进行对 Poisson 括号数学上更严格一些的证明. 根据定义直接计算动力学量的导数:

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + \sum_i \frac{\partial F}{\partial q^i} \dot{q}^i + \sum_i \frac{\partial F}{\partial p_i} \dot{p}_i \quad (3.9)$$

动力学量的导数应该在系统的真实演化中计算, 所以我们代入 Hamilton 正则方程:

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + \sum_i \left( \frac{\partial F}{\partial q^i} \frac{\partial H}{\partial p_i} - \frac{\partial F}{\partial p_i} \frac{\partial H}{\partial q^i} \right) \quad (3.10)$$

可以发现, 右边式子的求和项具有优美的反对称性, 我们定义其为力学量  $F$  和 Hamiltonian  $H$  的 **Poisson 括号**. 即

$$[F, H] = \sum_i \left( \frac{\partial F}{\partial q^i} \frac{\partial H}{\partial p^i} - \frac{\partial F}{\partial p_i} \frac{\partial H}{\partial q^i} \right) \quad (3.11)$$

这个定义和第一节中的定义完全一致:

$$\sum_{\alpha, \beta} \frac{\partial F}{\partial x_\alpha} \omega^{\alpha\beta} \frac{\partial F}{\partial x_\beta} \quad (3.12)$$

根据  $\omega$  的形式我们发现只有

$$\frac{\partial F}{\partial q^i} \frac{\partial H}{\partial p^i}, \quad \frac{\partial F}{\partial p_i} \frac{\partial H}{\partial q^i} \quad (3.13)$$

这样的项保留了下来, 并且前后差系数  $-1$ .

这样力学量的导数就可以写成:

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + [F, H] \quad (3.14)$$

如果说  $F$  不显含时间且  $[F, H] = 0$ , 那么该力学量守恒. 这样, 对不显含时间的力学量我们就可以直接计算其与 Hamiltonian

更广义地说, 任意两个力学量的 Poisson 括号定义为

$$[F, G] = \sum_{i=1}^f \left( \frac{\partial F}{\partial q^i} \frac{\partial G}{\partial p_i} - \frac{\partial G}{\partial q^i} \frac{\partial F}{\partial p_i} \right) \quad (3.15)$$

在量子力学中这直接对应**对易子**, 并且有着更简洁的形式

$$[\hat{F}, \hat{G}] = \hat{F}\hat{G} - \hat{G}\hat{F} \quad (3.16)$$

并且对算符对时间的导数也有完全相同的式子

$$\frac{d\hat{F}}{dt} = \frac{\partial \hat{F}}{\partial t} + [\hat{F}, \hat{H}] \quad (3.17)$$

### 3.1.3 Poisson 括号的性质

本小节介绍 Poisson 括号的一些性质, 这些性质可以极大简化我们的运算. 在定义 Poisson 括号时, 中间的矩阵  $\omega$  是一个反对称矩阵, 因此当我们交换  $F, G$  两个力学量的顺序的时候, 其 Poisson 括号应该相差符号. 也就是说, Poisson 括号具有反对称性:

$$[F, G] = -[G, F] \quad (3.18)$$

因为偏导和求和都是线性映射, 所以 Poisson 括号满足线性性的条件:

$$[c_1 F_1 + c_2 F_2, G] = c_1 [F_1, G] + c_2 [F_2, G] \quad (3.19)$$

这对第二个分量也成立.

随后, 我们注意到 Poisson 括号的偏导运算使其也有导数的性质. 因此 Poisson 括号也满足对应的莱布尼茨法则, 即

$$[F_1 F_2, G] = F_1 [F_2, G] + F_2 [F_1, G] \quad (3.20)$$

若  $\lambda$  是  $F, G$  共有的参数, 那么

$$\frac{\partial}{\partial \lambda} [F, G] = \left[ \frac{\partial F}{\partial \lambda}, G \right] + \left[ F, \frac{\partial G}{\partial \lambda} \right] \quad (3.21)$$

以及链式法则:

$$[F(f), G] = \frac{dF}{df} [f, G] \quad (3.22)$$

最后是 Poisson 括号一个重要的恒等式: **Jacobi 恒等式**. 描述了 Poisson 括号的一种循环对称性.

$$[F_1, [F_2, F_3]] + [F_2, [F_3, F_1]] + [F_3, [F_1, F_2]] = 0 \quad (3.23)$$

就是按照 1, 2, 3 的顺序的三种轮换, 求和起来是 0.

利用 Jacobi 恒等式以及前面的链式法则, 我们就有 **Poisson 定理**. 对于两个守恒的力学量, 他们的对时间的偏导数和他們与 Hamiltonian 的 Poisson 括号的关系为:

$$\frac{\partial F}{\partial t} = -[F, H] = [H, F] \quad (3.24)$$

这样, 我们就可以用  $F_1, F_2, H$  三个力学量的 Poisson 括号写出  $[F_1, F_2]$  对时间的导数:

$$\frac{d}{dt}[F_1, F_2] = \frac{\partial}{\partial t}[F_1, F_2] + [[F_1, F_2], H] \quad (3.25)$$

将式 (3.21) 中的参数  $\lambda$  换成时间  $t$ , 就有

$$\frac{d}{dt}[F_1, F_2] = [F_1, [H, F_2]] + [F_2, [F_1, H]] + [H, [F_1, F_2]] \quad (3.26)$$

我们发现等号最右边就是关于  $F_1, F_2, H$  三个量的 Jacobi 恒等式, 进而  $[F_1, F_2]$  对时间的全导数为零. 也就是说, 两个守恒的力学量的 Poisson 括号也是守恒的. 但是, 这个定理并不能帮助我们无限地生成新的守恒量. 一个系统守恒量的个数不应该超过其自由度的个数.

### 3.1.4 基本力学量的 Poisson 括号

一些复杂的力学量可以表为基本力学量的函数. 这样, 计算一些基本力学量的 Poisson 括号有助于我们后续的计算.

最基础的两个力学量就是广义坐标和广义动量. 对于不同的广义坐标:

$$[q^i, q^j] = \sum_m \frac{\partial q^i}{\partial q^m} \frac{\partial q^j}{\partial p_m} - \frac{\partial q^j}{\partial q^m} \frac{\partial q^i}{\partial p_m} = 0 \quad (3.27)$$

也就是说任意两个广义坐标的 Poisson 括号为零. 完全类似的计算也可以得到:

$$[p_i, p_j] = 0 \quad (3.28)$$

接下来就是广义坐标和广义动量之间的 Poisson 括号.

$$[q^i, p_j] = \sum_m \frac{\partial q^i}{\partial q^m} \frac{\partial p_j}{\partial p_m} - \frac{\partial p_j}{\partial q^m} \frac{\partial q^i}{\partial p_m} = \sum_m \frac{\partial q^i}{\partial q^m} \frac{\partial p_j}{\partial p_m} \quad (3.29)$$

根据不同广义坐标和广义动量的独立性, 我们有

$$\frac{\partial q^i}{\partial q^m} = \frac{\partial p_i}{\partial p_m} = \delta_{im}, \quad (3.30)$$

带回得到

$$[q^i, p_j] = \sum_m \delta_{im} \delta_{mj} = \delta_{ij} \quad (3.31)$$

这个式子说明, 对于一组共轭的广义坐标和广义动量, 其 Poisson 括号不为零.

接下来我们计算任意力学量与广义坐标和广义动量的 Poisson 括号.

$$[q^i, F] = \sum_m \frac{\partial q^i}{\partial q^m} \frac{\partial F}{\partial p_m} - \frac{\partial F}{\partial q^m} \frac{\partial q^i}{\partial p_m} = \sum_m \delta_{im} \frac{\partial F}{\partial p_m} = \frac{\partial F}{\partial p_i} \quad (3.32)$$

以及

$$[p_i, F] = \sum_m \frac{\partial p_i}{\partial q^m} \frac{\partial F}{\partial p_m} - \frac{\partial F}{\partial q^m} \frac{\partial p_i}{\partial p_m} = - \sum_m \delta_{im} \frac{\partial F}{\partial q_m} = - \frac{\partial F}{\partial q^i} \quad (3.33)$$

以上两个式子在量子力学中的一定条件下也成立. 若观测量  $\hat{F}$  可以展开成  $\hat{p}, \hat{q}$  的幂级数, 那么:

$$\frac{1}{i\hbar} [\hat{q}, \hat{F}] = \frac{\partial \hat{F}}{\partial \hat{p}}, \quad \frac{1}{i\hbar} [\hat{p}, \hat{F}] = - \frac{\partial \hat{F}}{\partial \hat{q}} \quad (3.34)$$

可以看出, 量子力学中的对易子  $[\cdot, \cdot]$  和经典力学种的 Poisson 括号有着对应关系:

$$\frac{1}{i\hbar} [\cdot, \cdot] \leftrightarrow [\cdot, \cdot]_{\text{classical}} \quad (3.35)$$

这样, 如果我们将  $H$  取作 Hamiltonian, 就有

$$[q^i, H] = \frac{\partial H}{\partial p_i}, \quad [p_i, H] = - \frac{\partial H}{\partial q_i} \quad (3.36)$$

利用式 (3.14) 就发现这两个式子其实就是 Hamilton 正则方程.

最后我们来看角动量的 Poisson 括号. 计算角动量的 Poisson 括号主要是为量子力学中的角动量理论作一些准备. 我们取广义坐标为直角坐标, 这样角动量可以写成:

$$l_k = \sum_{i,j} \varepsilon_{ijk} x_i p_j \quad (3.37)$$

我们根据之前得到的力学量和坐标, 动量间 Poisson 括号的关系可以得到角动量和坐标, 动量的 Poisson 括号:

$$[x^m, l_k] = \frac{\partial l_k}{\partial p_m} = \sum_i \varepsilon_{imk} x^i, \quad [p_m, l_k] = - \frac{\partial l_k}{\partial x_m} = \sum_i \varepsilon_{imk} p_i \quad (3.38)$$

这样, 我们可以计算两个角动量间的 Poisson 括号:

$$[l_i, l_j] = \sum_{m,n} \varepsilon_{mni} [x^m p_n, l_j] = \sum_{m,n} \varepsilon_{mni} (x^m [p_n, l_j] + p_n [x^m, l_j]) \quad (3.39)$$

这里利用  $\varepsilon$  符号的缩并公式:

$$\sum_i \varepsilon_{ijk} \varepsilon_{imn} = \delta_{jm} \delta_{kn} - \delta_{jn} \delta_{km} \quad (3.40)$$

可以得到:

$$[l_i, l_j] = \mathbf{x} \cdot \mathbf{p} \delta_{ij} - x^j p_i + x^i p_j - \mathbf{x} \cdot \mathbf{p} \delta_{ij} = x^i p_j - x^j p_i \quad (3.41)$$

对于  $i = j$ , 右边显然为零. 对于  $i \neq j$ , 右边则是第三个方向的角动量. 因此, 用  $\varepsilon$  符号可以将这个结果写成更紧凑的形式:

$$[l_i, l_j] = \sum_k \varepsilon_{ijk} l_k \quad (3.42)$$

接下来计算  $l^2$  与  $l_i$  的 Poisson 括号. 计算这个 Poisson 括号的动机纯粹与量子力学有关, 在经典力学中这个结论不太常用.

$$[l^2, l_i] = \sum_k 2[l_k, l_i] = 2 \sum_{m,k} \varepsilon_{kim} l_m = 0 \quad (3.43)$$

最后一个等号是因为, 在求和过程中, 只有  $k \neq m \neq i$  的项才会被保留. 而  $k, m$  求和会得到  $\varepsilon_{kim} + \varepsilon_{mik}$  的情况, 利用  $\varepsilon$  符号的反对称性就发现求和得零. 在量子力学中, 这个结果说明  $\hat{L}^2$  与  $\hat{L}_z$  两个力学量是对易的, 因而有共同的本征态.

## 3.2 Liouville 定理

在统计力学中, 我们研究大量粒子组成的体系. 每个粒子的运动状态在相空间中可以用一个相点表示. 对于粒子数在  $N_A$  量级的系统, 所有粒子的相点也构成一种流体. 我们考虑流体的密度, 密度和每个粒子的运动状态有关, 所以相空间中相流体的密度也是体系的一种力学量. 利用连续性方程, 我们有:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{j}) = 0 \quad (3.44)$$



接下来我们计算  $\nabla \cdot (\rho \mathbf{j})$ . 根据矢量分析中的公式:

$$\nabla \cdot (\rho \mathbf{j}) = (\nabla \rho) \cdot \mathbf{j} + \rho \nabla \cdot \mathbf{j} \quad (3.45)$$

代入式 (3.3) 并与式 (3.8) 对比发现, 第一项就是  $\rho$  和  $H$  的 Poisson 括号, 而第二项中出现的算子  $\nabla \cdot \boldsymbol{\omega} \nabla = 0$ . 也就是说,

$$\frac{\partial \rho}{\partial t} + [\rho, H] = 0 \quad (3.46)$$

而这个式子的含义正是  $\rho$  作为力学量对时间的全导数. 也就是说, 相流的密度不随时间改变, 是一个守恒量. 这就是 **Liouville 定理**, 是统计力学中非常重要的结论.

接下来我们对这个定理进行一些辨析和讨论. 首先要理解什么是相空间的相流体密度. 在研究流体力学中有两种语言, 一种是欧拉方法, 另一种是 Lagrange 方法. 欧拉法关注空间中的固定点, 也就是说, 我们在空间中框出一块确定的范围, 研究这里面流体的运动. 对于密度来说, 任意一个时刻  $t$  下流体有空间分布, 不同的  $t$  下有不同的空间分布, 这样密度可以写成:

$$\rho = \rho(t, \mathbf{x}) \quad (3.47)$$

如果我们框出一块范围研究这个区域内流体的密度, 那么这个区域内的流体密度关于时间的变化应当为  $\frac{\partial \rho}{\partial t}$ . 但 Liouville 定理是说  $\rho$  关于时间的全导数为零.  $\rho$  作为一种随时间变化的空间分布, 全导数的物理含义由 Lagrange 方法给出. 即我们研究一块确定的流体元的运动. 这块流体元的各种力学量随时间的变化就是这些力学量关于时间的全导数. 因此 Liouville 定理是说, 当我们考虑一些确定的相点, 这些相点的数密度不随时间改变. 而因为我们研究的是确定的一组相点, 数密度不变也就是说这组相点在相空间中所占据的体积不变. 利用更数学一些的语言来说, 就是体系在真实运动轨迹下, 轨迹上每一点的切空间是保体积的. 数学中我们一般用面积来描述, 也就是保辛:

$$\iint_{\Sigma} d\mathbf{q} \wedge d\mathbf{p} \quad (3.48)$$

而 Liouville 定理给出的相流的保辛性给出: 对于一个区域  $\Sigma$ , 其边界上的所有点开始随时间演化到区域  $\Sigma'$ , 在这前后有:

$$\iint_{\Sigma} d\mathbf{q} \wedge d\mathbf{p} = \iint_{\Sigma'} d\mathbf{q} \wedge d\mathbf{p} \quad (3.49)$$

利用 Stokes 定理就得到

$$J = \oint_{\Sigma} \mathbf{q} \cdot d\mathbf{p} = \oint_{\Sigma'} \mathbf{q} \cdot d\mathbf{p} \quad (3.50)$$

这就是 **Poincare-Cartan 积分不变量**. 其几何含义是相流管的截面积不变 (其实也就是 Liouville 定理).

通过 Liouville 定理, 我们发现相空间中面积的结构是极为重要的. 在下一小节中我们讨论体系的保面积的变换来寻找力学体系更多的守恒量.

### 3.3 正则变换

在系统真实的运动中, 我们希望系统的广义动量守恒:

$$p_k = \beta_k \quad (3.51)$$

这样, 由于广义动量守恒, 利用 Hamilton 正则方程就有:

$$\frac{\partial H}{\partial q^k} = -\dot{p}_k = 0 \quad (3.52)$$

也就是说, Hamiltonian 中不显含  $q_k$ . 进而 Hamiltonian 是所有广义动量的函数, Hamiltonian 守恒. 我们再利用另一个 Hamilton 正则方程:

$$\alpha_k = \frac{\partial H}{\partial p_k} = \dot{q}^k \quad (3.53)$$

式子的左边是 Hamiltonian 关于  $p_k$  的偏导数. 因为 Hamiltonian 是所有广义动量的函数, 所以求偏导后仍然是广义动量的函数, 是一个常数. 所以  $q^k$  满足的微分方程及其简单, 我们可以直接给出:

$$q^k = \alpha_k t + q_0^k \quad (3.54)$$

但是真实的运动中, 系统的性质不一定有我们期待的这么理想. 因此我们考虑, 如果对广义坐标和广义动量进行一种变换:

$$\begin{cases} Q_i = Q_i(q, p) \\ P_i = P_i(q, p) \end{cases} \quad (3.55)$$

使得变换得到的  $Q_i, P_i$  仍满足 Hamilton 正则方程且  $P_i$  是守恒的, 这样我们就可以简单地找出系统的守恒量并求解系统的运动. 很显然, 并不是所有的变换都能使得变换后的广义坐标和广义动量满足 Hamilton 正则方程. 我们需要对变换作出一些要求. 我们要求变换是一种**正则变换**, 即不破坏正则方程的变换.

### 3.3.1 正则变换的几何含义

这里我们需要再深入理解所谓的满足正则方程. 所谓的满足正则方程是指. 在  $\{q, p\}$  这套正则坐标下, 系统的有一个真实运动的轨道:

$$q = q(t), \quad p = p(t) \quad (3.56)$$

如果我们的变换是连续的, 那么这条轨道就可以映射到新的相空间  $\{Q, P\}$  下的一条轨道. 这样, 我们在新的相空间中如果可以定义一个与原来的 Hamiltonian 一一对应的新的 Hamiltonian  $H'$ ,

$$H'(Q(q, p), P(q, p)) = H(q, p) \quad (3.57)$$

使得  $Q, P$  这条由原轨道映射过来的新轨道满足 Hamilton 正则方程, 那么根据开头的讨论我们就可以容易地解出  $Q(t), P(t)$  并反解出  $q, p$ , 达到我们的目的.

我们就需要确认, 究竟是相空间的什么性质使之可以用 Hamilton 正则方程来描述, 即:

$$(\dot{Q}, \dot{P}) \equiv \mathbf{V} = \omega \nabla' H' \quad (3.58)$$

为了形式上的简洁, 我们统一将广义坐标和广义动量记作  $x^i, i = 1, \dots, 2f$ . 变换后的广义坐标和广义动量用  $X$  表示. 新老坐标下 Hamiltonian 梯度的

$$\nabla H = \frac{\partial H}{\partial x^i} = \sum_j \frac{\partial H}{\partial X^j} \frac{\partial X^j}{\partial x^i} = \mathbf{J}^T \nabla' H' \quad (3.59)$$

其中  $\mathbf{J}$  是坐标变换的 Jacobi 矩阵. 而相速度的变换公式为:

$$\dot{\mathbf{X}} = \sum_j \frac{\partial X^i}{\partial x^j} \dot{x}^j = \mathbf{J} \dot{\mathbf{x}} \quad (3.60)$$

进而就有:

$$\dot{\mathbf{X}} = \mathbf{J} \boldsymbol{\omega} \mathbf{J}^T \nabla' H' \quad (3.61)$$

满足 Hamilton 正则方程的条件要求:

$$\mathbf{J} \boldsymbol{\omega} \mathbf{J}^T = \boldsymbol{\omega} \quad (3.62)$$

这就是正则变换的**辛条件**. 我们将其写成分量形式:

$$\sum_{m,n} \frac{\partial X^i}{\partial x^m} \omega_{mn} \frac{\partial X^j}{\partial x^n} = \omega_{ij} \quad (3.63)$$

代入  $X, x$  和  $\omega$  矩阵的具体形式, 我们发现这个式子的意义就是  $Q, P$  的 Poisson 括号 (也称为正则对易关系) 不变:

$$[Q_i, Q_j] = [P_i, P_j] = 0, \quad [Q_i, P_j] = \delta_{ij} \quad (3.64)$$

由于基本力学量的 Poisson 括号不变, 进而所有力学量的 Poisson 括号都不改变. 我们之前说过, Poisson 括号的几何意义就是面积, 进而正则变换是一种保辛变换: 在相空间中, 任意一个区域的面积在变换后不发生变化, 这就是正则变换的几何含义.

根据这样的几何含义, 再根据我们之前证明的 Liouville 定理, 我们发现系统的真实演化:

$$q(t) \rightarrow q(t + dt), \quad p(t) \rightarrow p(t + dt) \quad (3.65)$$

就是一种正则变换. 只不过这种正则变换是无穷小的. 针对无穷小的正则变换我们留到之后讨论.

### 3.3.2 生成函数

虽然正则变换的辛条件有着更清晰的几何含义, 但是其形式并不能帮我们生成正则变换. 我们还是回到满足正则方程的条件上来. 利用这个条件, 我们允许一种更广泛的正则变换. 即

$$\begin{cases} Q_i = Q_i(q, p, t) \\ P_i = P_i(q, p, t) \end{cases} \quad (3.66)$$

在这种变换下, 我们就不要要求变换前后的相空间对应点上有着相同的 Hamiltonian, 因为这种变换还有时间作为参数. 这时候的几何意义相对不含时的正则变换就相对没那么明确<sup>1</sup>.

利用变换前后真实运动都满足 Hamilton 正则方程的条件, 我们可以给出二者的作用量的关系:

$$\delta S = \delta S' = 0 \quad (3.67)$$

仔细回忆变分的过程, 我们发现如果要求运动方程不变还允许一个额外的自由度. 比如说,

$$S = \int_{t_1}^{t_2} \frac{dF(q, t)}{dt} dt = F(q, t) \Big|_{t_1}^{t_2} \quad (3.68)$$

那么无论  $F$  的具体形式, 取变分都会得零. 如果将泛函导数视作一种广义的导数, 那么此时这个  $\dot{F}(q, t)$  的角色就像是一个任意常数. 这个自由度的存在就是所谓的**规范对称性**. 规范对称性在场论中是极为重要的, 我们在第十章就会看到. 对于正则变换, 变分的时候还要求  $\delta Q = 0$ , 因此  $F$  可以有更多的自由度:

$$F = F(q, Q, t) \quad (3.69)$$

这样, 我们就得到:

$$\sum_k P_k \dot{Q}^k - H' + \frac{dF(q, Q, t)}{dt} = \sum_k p_k \dot{q}^k - H \quad (3.70)$$

---

<sup>1</sup>数学上可以证明, 如果一种变换满足我们接下来给出的条件, 那么这种变换也是保辛的. 不过具体的证明需要更系统地引入辛几何, 但对物理图像的构建帮助不大. 因此这里我们就不再具体说明.

这就是正则变换应当满足的条件. 我们发现, 如果给出  $F$  的具体形式, 那么写出  $F$  的全微分:

$$dF = \sum_k \frac{\partial F}{\partial q^k} dq^k - \sum_k \frac{\partial F}{\partial Q^k} dQ^k + \frac{\partial F}{\partial t} dt \quad (3.71)$$

并于式 (3.70) 比较可以发现:

$$dF = \sum_k p_k dq^k - \sum_k P_k dQ^k + (H' - H)dt \quad (3.72)$$

$$\frac{\partial F}{\partial q^k} = p_k, \quad \frac{\partial F}{\partial Q^k} = -P_k \quad (3.73)$$

进而从中反解出  $Q^k = Q^k(q, p, t)$ ,  $P_k = P_k(q, p, t)$  就确定了一个正则变换. 同时我们也确定了新的 Hamiltonian 的形式:

$$H' = H + \frac{\partial F}{\partial t} \quad (3.74)$$

也就是说, 我们可以通过给出  $F$ , 就生成了一种正则变换并给出了这个正则变换的所有信息. 因此我们称函数  $F$  为正则变换的**生成函数**.

#### 四种生成函数

我们知道, 对多元函数进行 Legendre 变换可以改变其变元. 我们对生成函数进行 Legendre 变换就可以改变生成函数的变元. 对于  $q, Q, p, P$ , 两两组合一共有 4 种组合, 因此我们可以给出四组不同的生成函数. 我们以最开始得到的  $F(q, Q, t) = F_1(q, Q, t)$  作为出发点.

取

$$F_2 = \sum_k Q^k P_k + F_1 \quad (3.75)$$

这样

$$dF_2 = \sum_k P_k dQ^k + \sum_k Q^k dP_k + dF_1 = \sum_k p_k dq^k + \sum_k Q^k dP_k + (H' - H)dt \quad (3.76)$$

可以看到, 这时我们的变元成为了  $q^k, P_k$ . 因此  $F_2 = F_2(q, P, t)$ , 这就是第二类生成函数. 对比全微分关系我们可以得到第二类生成函数带来的正则变换:

$$p_k = \frac{\partial F_2}{\partial q^k}, \quad Q^k = \frac{\partial F_2}{\partial P_k} \quad (3.77)$$

进行类似的操作可以给出第三类和第四类生成函数以及其对应的正则变换关系:

$$\begin{cases} F_3(Q, p, t) = F_1 - \sum_k p_k q^k \\ P_k = -\frac{\partial F_3}{\partial Q^k}, \quad q^k = -\frac{\partial F_3}{\partial p_k} \end{cases} \quad (3.78)$$

$$\begin{cases} F_4(p, P, t) = F_2 - \sum_k p_k q^k \\ Q_k = -\frac{\partial F_4}{\partial P_k}, \quad q^k = -\frac{\partial F_4}{\partial p_k} \end{cases} \quad (3.79)$$

但是无论生成函数如何选取, 其关于时间的偏导数项都没有变化:

$$H' - H = \frac{\partial F_1}{\partial t} = \dots = \frac{\partial F_4}{\partial t} \quad (3.80)$$

接下来我们通过正则变换求解谐振子.

### Example 3.1. 谐振子

谐振子的 Hamiltonian 为

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2$$

作正则变换:

$$p = m\omega f(P) \cos Q, \quad q = f(P) \sin Q$$

这时  $p = m\omega q \cot Q$ , 这提醒我们要用第一类生成函数.

$$\frac{\partial F}{\partial Q} = -P, \quad \frac{\partial F}{\partial q} = m\omega q \cot Q$$

求解得到

$$F = \frac{1}{2}m\omega q^2 \cot Q \Rightarrow P = \frac{m\omega q^2}{2 \sin^2 Q}$$

由此得到:

$$f(P) = \sqrt{\frac{2P}{m\omega}}$$

因为我们的生成函数不含时, 所以新的 Hamiltonian 就是

$$H' = H = \omega P$$

符合我们期待的不含广义坐标. 因此根据之前的讨论就有:

$$Q = \omega t + \varphi, \quad P = \frac{E}{\omega}$$

反解得到

$$q = \sqrt{\frac{2E}{m\omega}} \sin(\omega t + \varphi), \quad p = \sqrt{2m\omega E} \cos(\omega t + \varphi)$$

## 3.4 Hamilton-Jacobi 方程

### 3.4.1 Hamilton-Jacobi 方程

相空间的正则变换给我们一种更大胆的猜想: 通过正则变换我们可以将所有的广义动量变为积分常数, 广义坐标变为时间的线性函数. 那么我们是否可以通过一种正则变换, 使得变换后的所有广义动量和广义坐标均为一个常数? 在这种情况下, 新的相空间中的相点应当是一个静止不动的点, 而只需要要求 Hamiltonian 则只与时间有关. 那么变换后的系统就满足 Hamilton 正则方程, 也就是说这种变换是一个正则变换. 因此我们只需要说明, 存在一种变换使我们找到  $2f$  个常数.

这种变换是存在的. 现设想我们求解系统的 Hamilton 正则方程. 为了完全确定唯一的确定轨道, 我们需要  $2f$  个初始条件. 因此, 系统的真实轨道可以写成



关于初始条件和时间的函数:

$$q = q(t, q(0), p(0)), \quad p = p(t, q(0), p(0)) \quad (3.81)$$

这样, 我们总可以通过具体的表达式把  $2f$  个初始条件反解出来, 而这个反解的过程形成一种变换. 变换后得到的初始条件当然是  $2f$  个常数. 因此我们就找到了满足要求的变换, 根据之前的讨论, 这种变换也是一种正则变换. 因此我们最初的设想可以达成.

接下来, 我们设这个变换的生成函数为  $F$ . 根据上节的讨论, 我们知道这个生成函数应当满足如下的条件:

$$\begin{aligned} H' = H'(t) &= H + \frac{\partial F}{\partial t} \\ \frac{\partial F}{\partial q^k} &= p_k, \quad \frac{\partial F}{\partial Q^k} = -P_k \end{aligned} \quad (3.82)$$

第二个式子中的  $P_k, Q^k$  均为常数, 我们用加粗的  $\mathbf{Q}, \mathbf{P}$  来表示. 这样,  $F$  实际上就可以看作是  $q, \mathbf{Q}, t$  的函数, 但其中只有  $q, t$  是变元.

接下来是  $H'(t)$  的选取. 在讨论正则变换时, 我们对  $H'$  的唯一要求是  $H'$  能够使新的相空间中的真实运动满足 Hamilton 正则方程, 只要能够满足 Hamilton 正则方程,  $H'$  的选取实际上是任意的. 因此为了简便, 我们直接取  $H'(t) = 0$ . 这样就得到了:

$$H + \frac{\partial F}{\partial t} = 0 \quad (3.83)$$

这时候, 由于我们已经给定了  $H', F$  就相对应的被这个方程决定 (允许有任意的常数存在). 我们称满足要求的生成函数为 **Hamilton 主函数**, 记作  $S^2$ . 此时, 因为  $S$  是  $q, t, \mathbf{Q}$  的函数, 所以由生成函数条件,

$$p_k = \frac{\partial S}{\partial q^k} = p_k(q, t) \quad (3.84)$$

即广义动量也是  $q, t$  的函数. 因此我们应将 Hamiltonian 改写为

$$H(q, p, t) = H\left(q, \frac{\partial S}{\partial q}, t\right) \quad (3.85)$$

---

<sup>2</sup>之后会看到这个记号的意义

带回到式 (3.83) 就可以得到:

$$H\left(q, \frac{\partial S}{\partial q}, t\right) + \frac{\partial S}{\partial t} = 0 \quad (3.86)$$

这就是 **Hamilton-Jacobi 方程**. 对于一个物理系统, 我们可以确定其 Hamiltonian 的形式, 进而从这个式子中可以解出  $S(q, t)$ , 其中包含  $f + 1$  个积分常数. 而其中一个常数我们已经可以知道其来源: 方程中没有出现  $S$  本身, 因此  $S' = S + S_0$  也是方程的解. 剩下  $f$  个包含物理信息的常数  $\alpha$ . 这个实际上就是我们正则变换后得到的广义坐标常数. 根据正则变换条件, 我们可以推出广义动量对应的常数:

$$P_k = \beta_k = -\frac{\partial S(q, \alpha, t)}{\partial \alpha_k} \quad (3.87)$$

根据这个方程, 我们可以直接反解出

$$q = q(\alpha, \beta, t) \quad (3.88)$$

包含  $2f$  个常数. 最后利用式 (3.84) 我们就可以解出所有的广义动量的真实轨道.

### 3.4.2 Hamilton 主函数的确定

我们写出 Hamilton 主函数的全微分:

$$dS = \sum_k p_k dq^k - \sum_k P_k dQ^k + (H' - H)dt \quad (3.89)$$

因为正则变换后的  $P, Q$  均为常数, 在真实运动中不发生变化, 因此右边的第二项为零. 我们再把  $H' = 0$  带回得到:

$$dS = \sum_k p_k dq^k - H dt \quad (3.90)$$

因此:

$$S = \int \sum_k p_k dq^k - H dt = \int \left( \sum_k p_k \dot{q}^k - H \right) dt \quad (3.91)$$

就是系统运动的作用量. 也就是说, Hamilton 主函数其实就是作用量. 但这个作用量与我们最开始讨论的作用量有细微的差别, 这个差别在于最开始的作用量是一个泛函, 而这里的作用量是给定初始条件的情况下系统随着真实轨道演化的作用量. 因此这个作用量只与末尾的时间  $t$  有关. 在初始条件确定的情况下, 这个作用量只是时间  $t$  的一元函数. 为了与泛函的作用量进行区分, 我们称这个作用量  $S(t)$  为经典作用量.

### 3.4.3 Hamilton-Jacobi 方程的分离变量

Hamilton-Jacobi 方程将  $2f$  个正则常微分方程改为了一个含  $f+1$  的独立变量的偏微分方程. 求解偏微分方程的一个常用的技术是分离变量的方法.

对于分离变量, 我们观察到, 如果一个系统的 Hamiltonian 不显含时间, 那么 Hamilton-Jacobi 方程中与时间有关的项就只有  $\partial S/\partial t$  的项. 进而如果这一项与时间无关, 那么整个偏微分方程的求解就完全与时间无关, 我们就完成了对时间的分离变量. 具体写出来就是:

$$S = W(q) - Et \quad (3.92)$$

这样

$$\frac{\partial S}{\partial t} = -E, \quad \frac{\partial S}{\partial q^k} = \frac{\partial W}{\partial q^k} \quad (3.93)$$

Hamilton-Jacobi 方程就变成:

$$H\left(q, \frac{\partial W}{\partial q}\right) = E \quad (3.94)$$

其中,  $W(q)$  被称为系统的 **Hamilton 特征函数**.

至此我们就基本完成了对哈密顿力学及其正则形式的主要内容的叙述. 以上的所有内容主要是一个理论框架, 我们基本上没有讨论如何具体地写出系统的 Hamiltonian, 但是相应地, 我们目前得到的所有结论有着较强的普遍性. 接下来我们从相空间移步到更为具体的位形空间来介绍 Lagrange 力学. Lagrange 力学在解决具体问题上比 Hamilton 力学更为有力, 在某种程度上也能够帮助我们写

出系统的 Hamiltonian. 具体的比较我们在结束对 Lagrange 力学的介绍之后进行. 以上的内容我们没有对周期运动的作用-角变量进行介绍, 这部分内容留到分析力学的应用这一章中.



# Chapter 4

## Lagrange 力学

### 4.1 位形空间的最小作用量原理

#### 4.1.1 位形空间

在描述相空间的最小作用量原理时, 我们说作用量为

$$S = \int_{t_1}^{t_2} \left( \sum_k p_k \dot{q}^k - H \right) dt \quad (4.1)$$

在当时, 这个形式的引入较为不自然. 我们解释说这是因为这样会使得  $H$  有更为明确的物理意义. 如果我们更愿意从直接从被积分的函数:

$$L(q, \dot{q}, p, t) = \sum_k p_k \dot{q}^k - H \quad (4.2)$$

进行研究, 就会得到另外一套截然不同的力学体系: Lagrange 力学.

Lagrange 力学是建立在位形空间之上的. 所谓**位形空间**就是由  $\{q^i\}$  为变量的空间, 是一个  $f$  维的流形. 在位形空间中, 用时间作为参数所描述的一条一维的轨迹:

$$q^i = q^i(t) \quad (4.3)$$

就是系统运动的一条轨道.

位形空间和相空间的重大区别在于, 位形空间中的一个点不能代表系统的一个态, 因为我们缺失了系统的动量信息. 因此, 即使是位形空间中的真实轨迹, 也是可以相交的, 因为即使两个系统处于位形空间的同一点, 也有可能有不同的态, 进而产生不同的演化方向.

#### 4.1.2 位形空间的最小作用量

一个自然的想法是, 位形空间是否有最小作用量原理? 也就是说, 是否存在一个定义在位形空间上的作用量:

$$S = \int_{t_1}^{t_2} L(q, \dot{q}, t) dt \quad (4.4)$$

使得取真实轨道的形况下作用量取得驻值? 这个问题涉及到多元泛函的极值问题. 在相空间中, 作用量是一个多元泛函

$$S = S[q, p] \quad (4.5)$$

泛函取极值时要求

$$\frac{\delta S}{\delta q} = 0, \quad \frac{\delta S}{\delta p} = 0 \quad (4.6)$$

这时候, 我们如果先求解第二个方程, 或者说利用 Hamilton 正则方程中的一组

$$\frac{\partial H}{\partial p} = \dot{q} \quad (4.7)$$

反解出  $p = p(q, \dot{q}, t)$ . 此时, 泛函就从  $2f$  元泛函变为了  $f$  元泛函. 我们只需要使得

$$\frac{\delta S}{\delta q} = 0 \quad (4.8)$$

即可. 因此, 如果我们可以不通过相空间的最小作用量原理直接确定  $p = p(q, \dot{q}, t)$ , 这样就自动满足第二组 Hamilton 正则方程, 进而我们发现泛函取驻值的两个条件就完全相同. 因此, 我们称带入了  $p = p(q, \dot{q}, t)$  后的  $f$  泛函的作用量为**位形空间的作用量**, 坐标空间中的最小作用量原理成立, 且在确定  $p = p(q, \dot{q}, t)$  的条件下和相空间中的最小作用量原理是等价的.

在利用坐标空间的最小作用量原理之前,我们先讨论一下 Lagrange 力学和 Hamilton 力学的区别. 在量子力学中我们更多的使用 Hamilton 力学的方法来讨论 (即正则量子化方法), 因为在量子力学中  $\dot{q}$  很难有明确的定义, 而且我们对  $\hat{p}, \hat{q}$  是未知的, 只知道  $[\hat{q}, \hat{p}] = i\hbar$ . 但是, 这并不说明 Lagrange 力学一定就不如 Hamilton 力学. 对于经典体系, 我们知道  $p = mv$ , 利用 Lagrange 力学就能很快地求解问题. 从对称性角度来说, Lagrange 力学有着规范对称性, 并且有着协变性, 这两点都是 Hamilton 力学不具备的, 而恰恰是这两点在场论中是极为重要的. 因此在相应的问题中, 我们会更倾向于采用 Lagrange 力学来解决问题. Hamilton 力学和 Lagrange 力学有着各自的优势所在.

## 4.2 Lagrange 方程

将  $p = p(q, \dot{q}, t)$  带回到作用量的表达式中就得到了位形空间的作用量:

$$S = \int_{t_1}^{t_2} L(q, \dot{q}, t) dt \quad (4.9)$$

其中,  $L$  被称为系统的 **Lagrangian**(拉格朗日量). 在相空间中的变分我们要求初末的广义坐标固定, 在位形空间中也是如此. 我们对  $S$  取变分:

$$\delta S = \int_{t_1}^{t_2} dt \left( \sum_i \frac{\partial L}{\partial q^i} \delta q^i + \frac{\partial L}{\partial \dot{q}^i} \delta \dot{q}^i \right) \quad (4.10)$$

利用变分与导数的可交换性并进行分部积分:

$$\frac{\partial L}{\partial \dot{q}^i} \delta \dot{q}^i = \frac{\partial L}{\partial \dot{q}^i} \frac{d}{dt} \delta q^i = \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}^i} \delta q^i \right) - \delta q^i \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} \quad (4.11)$$

带回得到

$$\delta S = \sum_i \left. \frac{\partial L}{\partial \dot{q}^i} \delta q^i \right|_{t_1}^{t_2} - \int_{t_1}^{t_2} dt \sum_i \delta q^i \left( \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} \right) \quad (4.12)$$

第一项根据端点固定的条件得零, 第二项要求在任意变分下积分恒为零, 只能有

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = 0 \quad (4.13)$$

这就是 **Lagrange 方程**.

### 4.3 Lagrange 力学与 Hamilton 力学

#### 4.3.1 Legendre 变换

Lagrange 力学是位形空间的力学, 而 Hamilton 力学则是相空间中的力学. 前者以  $q, \dot{q}, t$  为变量而后者以  $q, p, t$  为变量. 从数值上, Lagrangian 和 Hamiltonian 的关系为

$$L = \sum_k p_k \dot{q}^k - H \quad (4.14)$$

如果我们目前只有 Hamiltonian, 那么必须要通过 Hamilton 正则方程解出  $p = p(q, \dot{q}, t)$  代入, 才可以得到位形空间上的 Lagrangian. 那如果我们知道 Lagrangian 的形式, 如何能够得到 Hamiltonian? 这里我们就需要用到微分关系. 考虑系统在  $dt$  时间内, 广义坐标变动  $dq$ , 广义速度变动  $d\dot{q}$ , 广义动量由于  $q, \dot{q}$  的变化而变动  $dp$ , 则 Lagrangian 的变化为

$$L = \sum_k p_k d\dot{q}^k + \sum_k \dot{q}^k dp_k - dH \quad (4.15)$$

系统的变动是真实变动, 因此仍然满足 Hamilton 正则方程, 因此  $H$  的变化可以写为

$$dH(q, p(q, \dot{q}, t), t) = \sum_k \frac{\partial H}{\partial q^k} dq^k + \sum_k \frac{\partial H}{\partial p_k} dp_k + \frac{\partial H}{\partial t} dt \quad (4.16)$$

代入 Hamilton 正则方程:

$$dH(q, p(q, \dot{q}, t), t) = - \sum_k \dot{p}_k dq^k + \sum_k \dot{q}^k dp_k + \frac{\partial H}{\partial t} dt \quad (4.17)$$

带回得到

$$dL = \sum_k p_k d\dot{q}^k + \sum_k \dot{p}_k dq^k - \frac{\partial H}{\partial t} dt \quad (4.18)$$

对比  $L$  对应的全微分关系就可以得到:

$$p_k = \frac{\partial L}{\partial \dot{q}^k} \quad (4.19)$$



这就是广义动量  $p$  关于  $q, \dot{q}, t$  的函数关系. 也就是说, 当知道 Lagrangian 以后, 我们就可以通过这个式子求出广义动量, 再反解出  $\dot{q} = \dot{q}(q, p, t)$ . 利用

$$H = \sum p_k \dot{q}^k - L \quad (4.20)$$

就可以得到 Hamiltonian. 通过对比全微分关系, 我们还可以得到两个副产品:

$$\dot{p}_k = \frac{\partial L}{\partial q^k} \quad (4.21)$$

代入  $p = p(q, \dot{q}, t)$  我们发现, 这其实就是 Lagrange 方程; 以及

$$\frac{\partial L}{\partial t} = -\frac{\partial H}{\partial t} \quad (4.22)$$

也就是说, Lagrangian 与 Hamiltonian 是否含时是统一的.

上述过程从数学上来说是一个 **Legendre 变换** 的过程. Legendre 变换作为一种技术在物理中有广泛的应用, 因此我们也将作为一种技术进行介绍, 不过多地探讨其几何含义. 对于一个多元函数  $f$ , 其全微分为

$$df = \sum_i \frac{\partial f}{\partial x_i} dx_i \quad (4.23)$$

如果我们希望改变函数的变元, 使得  $\partial f / \partial x_k$  代替  $x_k$  与其他变量共同作为变元. 这时候, 为了凑出全微分的形式, 进行一个分部积分的操作:

$$\frac{\partial f}{\partial x_k} dx_k = d\left(\frac{\partial f}{\partial x_k} x_k\right) - x_k d\frac{\partial f}{\partial x_k} \quad (4.24)$$

移项得到:

$$d\left(\frac{\partial f}{\partial x_k} x_k - f\right) = -\sum_{i \neq k} \frac{\partial f}{\partial x_i} dx_i + x_k d\frac{\partial f}{\partial x_k} \quad (4.25)$$

这样我们得到了一个新的函数, 使得从全微分关系上看, 其变元符合我们的要求. 这个新的函数

$$\tilde{f} = \frac{\partial f}{\partial x_k} x_k - f \quad (4.26)$$

就是  $f$  的 Legendre 变换. 以上的讨论是纯数学的, 不过不难看出, Lagrangian 和 Hamiltonian 互为 Legendre 变换, 我们所替换的变元就是广义速度和广义动量.

### 4.3.2 Routh 函数

从式 (4.21) 中可以看出, 当 Lagrangian 中不显含某个广义坐标时, 其对应的广义动量守恒. 因此, 当写下一个 Lagrangian 的时候, 我们就已经清楚系统中守恒的广义动量了. 这时候的一个想法就是, 我们能否先求出这些广义动量:

$$\beta_k = p_k = \frac{\partial L}{\partial \dot{q}^k} \quad (4.27)$$

不妨设  $1 \sim s$  的广义动量是守恒的. 这样我们就可以反解出

$$\dot{q}^k = f(q^i, \dot{q}^i, \beta_{1 \sim s}), \quad k = 1, \dots, s; i = s+1, \dots, f \quad (4.28)$$

这样, 重新得到的 Lagrangian 中就只剩下  $f-s$  个广义坐标及其导数, 我们就可以求解更少的方程. 但这种做法是错误的. 我们观察  $L$  的全微分:

$$\sum_i \frac{\partial L}{\partial \dot{q}^i} d\dot{q}^i + \sum_i \frac{\partial L}{\partial q^i} dq^i + \frac{\partial L}{\partial t} dt = dL = \sum_i p_i dq^i - \sum_i \dot{p}_i dq_i - \frac{\partial H}{\partial t} dt \quad (4.29)$$

左边的等号来源于  $L$  全微分的定义, 而右侧的等号来源于真实轨道的 Hamilton 正则方程以及 Legendre 变换. Lagrange 方程来自将这两个等号连起来. 也就是说, 要给出正确的物理方程, 需要  $L$  有正确的函数形式而非数值. 当我们将  $\dot{q}$  反解再带回, 相当于保持 Lagrangian 的数值不变但改变了其函数形式, 因此我们不一定能得到正确的物理方程. 从最小作用量原理来说, 我们作的数值上不变的反解操作, 只在真实轨道上正确, 但我们却改变了泛函的结构, 因此前后的泛函极值条件不一定等价.

但是, 毕竟系统确实存在守恒的广义坐标, 因此我们所希望的简化方程求解的操作还是存在的. 我们发现, 通过 Legendre 变换, 我们变元从所有的广义动量改为了广义速度. 我们能否再通过 Legendre 变换, 将守恒的那些广义动量的共轭的广义速度变回到广义动量. 根据之前提到过的 Legendre 变换的方法, 这个新的函数就是

$$R = \sum_{i=1}^s p_s \dot{q}^s - L(q, \dot{q}, t) \quad (4.30)$$

被称为 **Routhian**(罗斯函数). 我们取这个函数的全微分发现:

$$dR = \sum_{i=1}^s \dot{q}^i dp_i - \sum_{k=s+1}^f \left( \frac{\partial L}{\partial \dot{q}^k} d\dot{q}^k + \frac{\partial L}{\partial q^k} dq^k \right) - \frac{\partial L}{\partial t} dt \quad (4.31)$$

而在真实演化的过程中,  $dp_s = 0$ , 这样, 对比其和  $R$  全微分的定义就可以得到 **Routh 方程**:

$$\frac{d}{dt} \frac{\partial R}{\partial \dot{q}^k} - \frac{\partial R}{\partial q^k} = 0, \quad k = s+1, \dots, f \quad (4.32)$$

其中, 我们已经省略了下角标:

$$\frac{\partial R}{\partial Q} = \left( \frac{\partial R}{\partial Q} \right)_{p_1, \dots, p_s} \quad (4.33)$$

因此, 从实际上的操作来说, 我们就是先观察出守恒的广义动量, 将对应的这部分广义速度反解出来, 再代入 Routhian. 这样的 Routhian 给出的 Routh 方程就是正确的.

## 4.4 Lagrange 力学与 Newton 力学

这一节我们主要探讨 Lagrange 力学与 Newton 力学的关系. 通过这一节的讨论, 我们就基本上可以用 Lagrange 力学以及 Hamilton 力学来处理大部分的经典力学体系. 并且我们能够看到分析力学可以描述更广义的体系.

### 4.4.1 d'Alembert 原理

对于经典力学系统, 系统的直角坐标下的 Hamiltonian 可以写成:

$$H = \sum_i \frac{\mathbf{p}_i^2}{2m_i} + V(\mathbf{r}_i) \quad (4.34)$$

根据 Hamilton 正则方程我们知道  $\mathbf{p}_i = m_i \mathbf{v}_i$ , 这样带回就可以得到经典力学系统的 Lagrangian:

$$L = \sum_i \frac{1}{2} m_i \mathbf{v}_i^2 - V = T - V \quad (4.35)$$

即系统的动能减去势能. 这样, 最小作用量原理就可以写成:

$$\delta S = - \int_{t_1}^{t_2} \sum_i \left( m\ddot{\mathbf{r}}_i + \frac{\partial V}{\partial \mathbf{r}_i} \right) \cdot \delta \mathbf{r}_i dt = 0 \quad (4.36)$$

Newton 力学中力的定义为:

$$\mathbf{F}_i = - \frac{\partial V}{\partial \mathbf{r}_i} \quad (4.37)$$

被积函数在任意  $\delta \mathbf{r}_i$  下均成立, 只能有

$$\sum_i (m\ddot{\mathbf{r}}_i - \mathbf{F}_i) \cdot \delta \mathbf{r}_i \quad (4.38)$$

其中根据最小作用量中我们约定的变分,  $\delta \mathbf{r}_i$  为约束允许情况下的任意的等时的  $\mathbf{r}_i$  的无穷小变动. 式 (4.38) 就是著名的 **d'Alembert 原理**, 是连接 Newton 力学与最小作用量原理的桥梁. 从 d'Alembert 原理中我们可以看出, 分析力学描述的体系都是**理想约束**的体系. 因为在给出 d'Alembert 原理时, 我们将力定义为势能的负梯度, 这只能描述质点与其他质点以及外场的相互作用, 但不能描述约束对系统造成的力. 事实上, 我们在构建分析力学体系时就默认了这一点. 理想约束的定义为

$$\sum_i \mathbf{N}_i \cdot \delta \mathbf{r}_i = 0 \quad (4.39)$$

即约束带来的相互作用不对作用量的变动贡献. 或者从量纲上看, 这个式子似乎是约束力在位置改变下做的功. 因为这样的功并不会在真实演化下有任何的物理意义, 因此我们将其称为**虚功**. 所以换句话说, 理想约束就是约束力的虚功之和为零. 更进一步地, 我们可以用虚功的概念重新看 d'Alembert 原理. 在体系平衡时, d'Alembert 原理就是说, 所有除约束外的相互作用带来所作的虚功为零:

$$\sum_i \mathbf{F}_i \cdot \delta \mathbf{r}_i = 0 \quad (4.40)$$

这就是所谓的**虚功原理**.

### 4.4.2 第一类 Lagrange 方程

如果将 d'Alembert 原理作为分析力学中最基本的假设 (而非最小作用量原理), 我们就可以从 Newton 力学开始构建整个分析力学体系. 我们将 d'Alembert 原理广义坐标化, 则根据式 (1.3) 我们可以给出等时变分的虚位移的表达式, 只需要去掉  $\frac{\partial \mathbf{r}}{\partial t} \delta t$  项:

$$\delta \mathbf{r}_i = \sum_k \frac{\partial \mathbf{r}_k}{\partial q^k} \delta q^k \quad (4.41)$$

这样 d'Alembert 原理就变成了

$$\sum_i m_i \ddot{\mathbf{r}}_i \cdot \left( \sum_k \frac{\partial \mathbf{r}_k}{\partial q^k} \delta q^k \right) = \sum_i \mathbf{F}_i \cdot \left( \sum_k \frac{\partial \mathbf{r}_k}{\partial q^k} \delta q^k \right) \quad (4.42)$$

观察式 (1.3) 可以发现,

$$\frac{\partial \mathbf{r}_k}{\partial q^k} = \frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}^k} \quad (4.43)$$

取该式的全微分可以看到

$$\frac{d}{dt} \frac{\partial \mathbf{r}_k}{\partial q^k} = \frac{\partial \dot{\mathbf{r}}_i}{\partial q^k} \quad (4.44)$$

这样将这两个式子带回到左边:

$$\begin{aligned} \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_k}{\partial q^k} &= \frac{d}{dt} \dot{\mathbf{r}}_i \cdot \frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}^k} - \dot{\mathbf{r}}_i \cdot \frac{\partial \dot{\mathbf{r}}_i}{\partial q^k} \\ &= \frac{1}{m_i} \left( \frac{d}{dt} \frac{\partial T}{\partial \dot{q}^k} - \frac{\partial T}{\partial q^k} \right) \end{aligned} \quad (4.45)$$

其中  $T$  为体系的总动能. 对式 (4.42) 的右边, 我们定义广义力:

$$\sum_i \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q^k} = Q_k \quad (4.46)$$

就可以将 d'Alembert 原理改写为

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}^k} - \frac{\partial T}{\partial q^k} = Q_k \quad (4.47)$$

这就是**第一类 Lagrange 方程**. 进而, 如果我们可以将广义力  $Q_k$  写成

$$Q_k = \frac{d}{dt} \frac{\partial U}{\partial \dot{q}^k} - \frac{\partial U}{\partial q^k} \quad (4.48)$$

这样定义系统的 Lagrangian 为  $L = T - U$ , 就得到了我们所熟知的 Lagrange 方程. 第一类 Lagrange 方程的优势在于, 我们不用要求势能必须只能是位置的函数:  $V = V(q)$ , 而是可以定义广义势:

$$U = U(q, \dot{q}, t) \quad (4.49)$$

取描述带有速度的力.

#### Example 4.1. Lorentz 力

取广义势为

$$U = -q \sum_i A_i(\mathbf{x}) \dot{x}_i + q\phi \quad (4.50)$$

则

$$\begin{aligned} \frac{d}{dt} \frac{\partial U}{\partial \dot{x}_k} &= -q \sum_i \frac{\partial A_k}{\partial x_i} \dot{x}_i \\ \frac{\partial U}{\partial x_k} &= q \frac{\partial \phi}{\partial x_k} - q \sum_i \frac{\partial A_i}{\partial x_k} \dot{x}_i \end{aligned}$$

这样

$$Q_k = q \sum_i \left( \frac{\partial A_i}{\partial x_k} - \frac{\partial A_k}{\partial x_i} \right) \dot{x}_i - q \frac{\partial \phi}{\partial x_k}$$

定义矢量场  $\mathbf{B}, \mathbf{E}$  为

$$\varepsilon_{ijk} B_k = \frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j}, E_i = -\frac{\partial \phi}{\partial x_i} \quad (4.51)$$

这样用矢量写出广义力就是

$$\mathbf{F} = q(\mathbf{v} \times \mathbf{B} + \mathbf{E}) \quad (4.52)$$

这就是带电粒子受到的 Lorentz 力. 熟悉电动力学的读者容易看出,  $\mathbf{B} = \nabla \times \mathbf{A}$  为磁感应强度,  $\mathbf{E} = -\nabla \phi$  为电场.  $\mathbf{A}$  为磁矢势,  $\phi$  为电势. 也就是说, 带电粒子在电磁场中的 Lagrangian 为

$$L = \frac{1}{2} m \mathbf{v}^2 + q \mathbf{A} \cdot \mathbf{v} - q\phi \quad (4.53)$$

可以得到对应的正则动量

$$\mathbf{p} = m\mathbf{v} + q\mathbf{A} \quad (4.54)$$

由此得到对应的 Hamiltonian 为

$$H = \frac{(\mathbf{p} - q\mathbf{A})^2}{2m} + q\phi \quad (4.55)$$

## 4.5 Lagrange 力学的广义协变性

我们前面说过, Lagrange 力学相较 Hamilton 力学的一个优势就是具有协变性. 或者更准确的说, 是具有**广义协变性**. 在 Hamilton 力学中, 我们对系统的广义坐标和广义动量进行某种变换, 只有当这种变换是一个正则变换的时候, 变换后的系统才能仍然满足 Hamilton 正则方程. 但是, Lagrange 力学中, 我们对坐标进行任意的点变换, 也就是说采用不同的广义坐标对系统进行描述, 系统都自动满足 Lagrange 方程. 这就是 Lagrange 力学的广义协变性.

在探讨拉格朗日力学的广义协变性之前, 我们先给几个著名的结论. 对于某一个只与坐标有关的量 (可以是矢量或者标量)  $Q(q, t)$ . 我们求其对时间的导数:

$$\dot{Q} = \frac{\partial Q}{\partial t} + \sum_k \frac{\partial Q}{\partial q^k} \dot{q}^k \quad (4.56)$$

由于  $Q$  只和  $q, t$  有关, 因此我们有

$$\frac{\partial \dot{Q}}{\partial \dot{q}^k} = \frac{\partial Q}{\partial q^k} \quad (4.57)$$

随后我们求

$$\frac{\partial \dot{Q}}{\partial q^k} = \frac{\partial^2 Q}{\partial q^k \partial t} + \sum_i \frac{\partial^2 Q}{\partial q^k \partial q^i} \dot{q}^i \quad (4.58)$$

根据混合导数的可交换性, 我们发现这其实就是  $\partial Q / \partial q^k$  关于时间的全导数. 也就是说

$$\frac{\partial \dot{Q}}{\partial q^k} = \frac{d}{dt} \frac{\partial Q}{\partial q^k} \quad (4.59)$$

式 (4.57) 和 (4.59) 就是著名的 **Lagrange 关系**. 这一点在我们以 Newton 力学为出发点来推导 Lagrange 方程时也有用处, 可以看到, 式 (4.43) 和式 (4.44) 就是 Lagrange 关系的一个特殊情况.

接下来我们就来说明 Lagrange 力学的广义协变性. 我们设  $Q, q$  为两组不同的广义坐标, 这两组广义坐标都可以唯一确定系统的位形. 这样两组广义坐标之间就存在非奇异的变换关系:

$$Q^j = Q^j(q, t), \quad \dot{Q}^j = \dot{Q}^j(q, \dot{q}, t) \quad (4.60)$$

我们假设在  $Q$  作为广义坐标时, Lagrange 方程成立. 如果我们用  $q$  作为广义坐标来描写 Lagrangian, 就有

$$L(q, \dot{q}, t) = L(Q(q, t), \dot{Q}(q, \dot{q}, t), t) \quad (4.61)$$

那么根据偏导数的链式法则, 我们有

$$\frac{\partial L}{\partial q^i} = \sum_j \frac{\partial Q^j}{\partial q^i} \frac{\partial L}{\partial Q^j} + \sum_j \frac{\partial \dot{Q}^j}{\partial q^i} \frac{\partial L}{\partial \dot{Q}^j} \quad (4.62)$$

$$\frac{\partial L}{\partial \dot{q}^i} = \sum_j \frac{\partial \dot{Q}^j}{\partial \dot{q}^i} \frac{\partial L}{\partial \dot{Q}^j} \quad (4.63)$$

我们在写关于  $q$  的 Lagrange 方程时, 还要求这项关于时间的全导数:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} = \sum_j \frac{d}{dt} \left( \frac{\partial \dot{Q}^j}{\partial \dot{q}^i} \right) \frac{\partial L}{\partial \dot{Q}^j} + \sum_j \frac{\partial \dot{Q}^j}{\partial \dot{q}^i} \frac{d}{dt} \frac{\partial L}{\partial \dot{Q}^j} \quad (4.64)$$

于是利用 Lagrange 关系,  $q$  的 Lagrange 方程为

$$\begin{aligned} & \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} \\ &= \sum_j \frac{d}{dt} \left( \frac{\partial \dot{Q}^j}{\partial \dot{q}^i} \right) \frac{\partial L}{\partial \dot{Q}^j} + \sum_j \frac{\partial \dot{Q}^j}{\partial \dot{q}^i} \frac{d}{dt} \frac{\partial L}{\partial \dot{Q}^j} - \sum_j \frac{\partial Q^j}{\partial q^i} \frac{\partial L}{\partial Q^j} - \sum_j \frac{\partial \dot{Q}^j}{\partial q^i} \frac{\partial L}{\partial \dot{Q}^j} \\ &= \sum_j \frac{\partial \dot{Q}^j}{\partial q^i} \frac{\partial L}{\partial \dot{Q}^j} + \sum_j \frac{\partial Q^j}{\partial q^i} \frac{d}{dt} \frac{\partial L}{\partial \dot{Q}^j} - \sum_j \frac{\partial Q^j}{\partial q^i} \frac{\partial L}{\partial Q^j} - \sum_j \frac{\partial \dot{Q}^j}{\partial q^i} \frac{\partial L}{\partial \dot{Q}^j} \\ &= \sum_j \frac{\partial Q^j}{\partial q^i} \left[ \frac{d}{dt} \frac{\partial L}{\partial \dot{Q}^j} - \frac{\partial L}{\partial Q^j} \right] \\ &= 0 \end{aligned} \quad (4.65)$$



自动成立. 从这个角度看我们也可以看到 Lagrange 方程和 Newton 力学的等价性. 我们只需要取

$$L = T - V, \quad Q = \boldsymbol{r} \quad (4.66)$$

这样  $Q$  和 Lagrange 方程就是 Newton 运动定律, 而左边是我们选取的广义坐标的 Lagrange 方程.





# Chapter 5

## 对称性与守恒量

在第四章中我们以寻找守恒量为动机构建了 Hamilton 力学的主要结构, 在经过对 Lagrange 力学的讨论后, 我们在这 Hamilton 力学和 Lagrange 力学这两种视角下重新对系统的守恒量进行一个相对系统的讨论. 我们发现, 这两种力学都给出一个结论: 系统的变换对称性与守恒量之间有着深刻的联系.

### 5.1 循环坐标

首先考虑两个最简单的动力学量:  $p_i, q^i$ . 他们的导数由 Hamilton 正则方程直接给出. 根据 Hamilton 正则方程, 若

$$\frac{\partial H}{\partial q^i} = 0 \quad (5.1)$$

即写 Hamilton 函数时我们发现这里头没有  $q^i$ . 这会导致

$$\dot{p}_i = 0 \quad (5.2)$$

说明  $p_i$  是守恒的. 这时候, 我们称  $q^i$  是一个循环坐标.

我们发现, 位于我们研究中心地位的 Hamiltonian  $H$  本身就是一个态函数. 计算其导数:

$$\frac{dH}{dt} = \sum_i \frac{\partial H}{\partial q^i} \dot{q}^i + \sum_i \frac{\partial H}{\partial p_i} \dot{p}_i + \frac{\partial H}{\partial t} \quad (5.3)$$

代入我们已经导出的 Hamilton 正则方程, 就发现

$$\dot{H} = \sum_i \frac{\partial H}{\partial q^i} \frac{\partial H}{\partial p_i} + \sum_i \frac{\partial H}{\partial p_i} \left( -\frac{\partial H}{\partial q^i} \right) + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t} \quad (5.4)$$

这个结果是非平凡的, 它告诉我们说, 一旦 Hamiltonian 不显含时间, 那么 Hamiltonian 就是守恒的! 从物理上这个结果可以进行直观理解. 在讨论非定常约束时我们说, 若一个约束显含时间, 那么这个系统就需要受到一个运动的“机器”来构造这种约束. 对于 Hamiltonian 来说也同理, 如果我们要构造一个能量显含时间的系统, 那么就需要类似于构造非定常约束的那样一个“机器”, 而这个机器会与系统本身存在能量的交换, 这样 Hamiltonian 就不守恒. 如果一个系统的 Hamiltonian 不显含时间, 那么这个系统就是“自然”的, 会有能量守恒.

对于 Lagrange 力学也可以有相同的结论, 写下 Lagrange 方程

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}^i} \right) = \frac{\partial L}{\partial q^i} \quad (5.5)$$

如果 Lagrangian 不显含某个广义坐标, 就有与其对应的广义动量

$$\frac{\partial L}{\partial \dot{q}^i} = p_i \quad (5.6)$$

守恒. 如果 Lagrangian 不显含时间, 根据之前导出的关系式

$$\frac{\partial L}{\partial t} = -\frac{\partial H}{\partial t} \quad (5.7)$$

就知道 Hamiltonian 守恒.

## 5.2 无穷小正则变换

### 5.2.1 对称性

上述的讨论给我们一种感觉, 就是当 Hamiltonian 不显含某个量时, 就能找到一个量守恒. 我们仔细分析 Hamiltonian 不显含某个量的定义, 就是当这个量进行一个无穷小的变化:

$$Q \rightarrow Q + \varepsilon, \quad \varepsilon \ll 1 \quad (5.8)$$

则变换后的 Hamiltonian

$$\bar{H} = H(Q + \varepsilon) = H(Q) + \frac{\partial H}{\partial Q} \varepsilon = H(Q) \quad (5.9)$$

与变换前的 Hamiltonian 相同.

从物理的观点上看, 如果我们让

$$q \rightarrow q + \varepsilon \quad (5.10)$$

我们实际上进行了在这个广义坐标上的一个无穷小的平移. 关于平移, 这里需要进行一个额外的解释. 我们之前说过, 相空间不一定是一个线性空间, 这导致对于广义坐标一个有限的变化其实并没有任何意义. 但是我们还提到过, 相空间是一个  $2f$  维的流形, 而流形在一个很小的邻域内是微分同胚与一个  $2f$  维的欧式空间的, 不严格地说就是只要取的邻域足够小, 我们就感受不到空间的弯曲. 这样, 这个无穷小量就可以认为是在该线性空间中的一个平移. 一个很好的例子是以旋转角为广义坐标. 在角度足够小的时候, 我们就可以用一个矢量来表达这个无穷小的旋转, 进而得到角速度矢量.

回到对物理的讨论. Hamiltonian 在这个广义坐标上的一个无穷小的平移下不变, 实际上是一种对称性. 这里的对称性是广义的对称性, 指的是系统在某种变换下不变的特性. 这是对狭义对称的一个延拓, 在广义的对称性下, 狭义的对称性就是说在反射变换下不变的性质. 如果将时间视作一个  $q^0$  的广义坐标, 那么我们就将之前的规律总结为, 若 Hamiltonian 关于某个广义坐标存在无穷小平移对称性, 那么存在一个与之对应的守恒量. 我们发现守恒量似乎能够刻画系统的对称性. 从解决问题的角度看, 如果我们注意到了系统的某种对称性, 就可以找到一个守恒量来方便我们对系统性质的研究.

在 Hamilton 力学中, 广义坐标和广义动量的地位是一致的, 所以广义的无穷小对称变换应当是把  $2f$  个广义坐标和广义动量变换到  $2f$  个新的广义坐标和动量. 这让我们回想起了正则变换, 而对称性指的是在连续变换下的不变性, 因此我们在这里讨论的就是无穷小正则变换.

### 5.2.2 无穷小正则变换

接下来我们认为  $t$  是序号为零的广义坐标, 将广义坐标和广义动量统一写成  $x_i$ . 最普遍的正则变换是

$$x_i = f_i(x_{0 \sim 2f}) \quad (5.11)$$

但这样一个变换是点变换, 变换没有任何的自由度. 这样, 这个变换要么是恒同映射, 要么不是无穷小的. 所以考虑一个二维的旋转变换, 我们发现在描述这个变换时需要引入一个连续变化的参量  $\theta$ , 并且  $\theta = 0$  时为恒同变换. 再考虑三维的变换, 我们需要不止一个连续变化的参量. 所以总结下来, 对无穷小变换的描述我们需要引入多个连续变化的参量, 且在参量全部为零的时候变换为恒同变换. 这样我们就可以利用生成函数写出带参数的正则变换. 这里我们取第二类生成函数:

$$G = G(q, P, \alpha) \quad (5.12)$$

我们要求在  $\alpha_1 = \dots = \alpha_r = 0$  时  $G$  是恒同正则变换的生成函数, 因此对其进行一阶展开就可以得到

$$G = \sum_{k=1}^f q^k P_k + \sum_{l=1}^r \left. \frac{\partial G}{\partial \alpha_l} \right|_{\alpha=0} \Delta \alpha_l \quad (5.13)$$

为了简洁, 我们记

$$\left. \frac{\partial G}{\partial \alpha_l} \right|_{\alpha=0} = F_l(q, P) \quad (5.14)$$

这样我们就构造出了无穷小正则变换. 由此:

$$p_k = P_k + \sum_l \frac{\partial F_l}{\partial q^k} \Delta \alpha_l, \quad Q^k = q^k + \sum_l \frac{\partial F_l}{\partial P^k} \Delta \alpha_l \quad (5.15)$$

因为正则变换前后的广义坐标和广义动量差无穷小, 因此在忽略高阶无穷小的情况下, 我们可以直接将生成函数中的  $P$  代换称  $p$ :

$$\begin{cases} Q^k = q^k + \sum_l \frac{\partial F_l}{\partial p_k} \Delta \alpha_l \\ P^k = p^k - \sum_l \frac{\partial F_l}{\partial q^k} \Delta \alpha_l \end{cases} \quad (5.16)$$

**Example 5.1.** 真实演化作为无穷小正则变换

我们之前提到过, 系统的真实演化可以看成是一种无穷小的正则变换. 现在经过更详细的讨论我们知道这种无穷小的正则变换的参数就是  $t$ . 进而, 根据 Hamilton 正则方程我们就知道

$$F = H \quad (5.17)$$

即 Hamiltonian.

现在我们考察系统的对称性. 经过无穷小正则变换后, 系统的 Hamiltonian 的改变为零:

$$0 = \delta H = \sum_{k=1}^f \frac{\partial H}{\partial q^k} \Delta q^k + \frac{\partial H}{\partial p_k} \Delta p_k \quad (5.18)$$

代入具体的正则变换的形式:

$$0 = \delta H = \sum_{k,l} \left( \frac{\partial H}{\partial q^k} \frac{\partial F_l}{\partial p_k} - \frac{\partial H}{\partial p_k} \frac{\partial F_l}{\partial q^k} \right) \Delta \alpha_l = \sum_l [F_l, H] \Delta \alpha_l \quad (5.19)$$

而  $\alpha$  是相互独立的任意的无穷小的参数, 因此一定有  $[F_l, H] = 0$ . 而  $F_l$  不含时间, 所以根据式 (3.14) 我们就发现了  $r$  个守恒量  $F_l$ .

**Example 5.2.** 量子力学中的对称与守恒

量子力学中, 一种无穷小变换可以写成么正算符:

$$\hat{U} = \mathbb{1} + i\varepsilon \hat{F}$$

要求变换前后仍满足薛定谔方程 (其实这就是对正则变换满足 Hamilton 正则方程的要求):

$$i\hbar \frac{d}{dt} \hat{U} |\psi\rangle = \hat{U} \hat{H} \hat{U}^\dagger \hat{U} |\psi\rangle$$

且 Hamiltonian 不变:

$$\hat{U} \hat{H} \hat{U}^\dagger = \hat{H}$$

代入计算就可以得到

$$[\hat{F}, \hat{H}] = 0$$

即力学量  $\hat{F}$  守恒.

因此, 通过无穷小正则变换, 我们发现, 如果系统有  $r$  个相互独立的连续正则变换的对称性, 那么就一定有与之对应的  $r$  个守恒量. 这就是广义的**诺特定理**. 一般的诺特定理是说系统在时空上有某种连续对称性对应的守恒量. 这一点我们在下节用 Lagrange 力学讨论.

## 5.3 诺特定理和规范对称

### 5.3.1 Lagrange 力学的规范对称性

考虑从一种变换后的作用量

$$\bar{S} = \int_{\bar{t}_1}^{\bar{t}_2} \bar{L} d\bar{t} \quad (5.20)$$

如果  $S \equiv \bar{S}$ , 这里的恒等符号是指, 变换后的任意一个相同的轨迹  $\bar{q}(\bar{t})$  下, 变换前后的作用量相同. 这样对两边同时取变分得零就可以得到相同得运动方程. 再具体一些就是, 我们将原来导出的微分方程中的  $q, \dot{q}, t$  全部换成  $\bar{q}, \dot{\bar{q}}, \bar{t}$ , 这个微分方程仍然能描述系统的真实运动. 根据式 (3.68) 的讨论, 如果存在某种无穷小的变换, 使得

$$\bar{S} = \int_{\bar{t}_1}^{\bar{t}_2} \bar{L} d\bar{t} = S + \int_{t_1}^{t_2} \frac{dF(q, t)}{dt} dt \quad (5.21)$$

那么系统就会满足相同的运动方程. 进而具有某种不变性.

### 5.3.2 诺特定理

我们之前对对称性与守恒量的规律只总结到了对系统进行平移变换的情况. 一个很自然的想法是, 一种更普遍的无穷小的变换是否能够对应一个守恒量. 通



过下面的推导, 我们发现对这个问题的答案是肯定的. 对应的结论就是著名的**诺特定理**.

根据之前的讨论, 我们知道在位形空间一种无穷小的连续变换可以写成可以写成

$$\bar{q}^k = T^k(q^{0\sim f}, \alpha_{1\sim r}) \quad (5.22)$$

那么对所有参量在 0 处做泰勒展开就得到无穷小变换的表达式:

$$\bar{q}^k = q^k + \sum_{l=1}^r \frac{\partial T^k}{\partial \alpha_l} \Big|_{\alpha_1=\dots=\alpha_r=0} \Delta \alpha_l \quad (5.23)$$

为了少写些重复的东西, 我们记

$$\frac{\partial T^k}{\partial \alpha_l} \Big|_{\alpha_1=\dots=\alpha_r=0} = T_l^k(q^{0\sim f}) \quad (5.24)$$

这样就可以写成

$$\bar{q}^k = q^k + \sum_{l=1}^r T_l^k \Delta \alpha_l \quad (5.25)$$

注意这里的  $T$  系数并不是任取的. 随后我们要确定  $F$  全导数这项的无穷小. 我们考虑任意一个函数  $F(q^{1\sim f}, t, \alpha_{1\sim r})$ , 对  $\alpha_i$  的函数的要求与变换一样. 这样在  $\alpha_{1\sim r} \rightarrow 0$  时, 就有

$$\Delta F = \sum_{l=1}^r \frac{\partial F}{\partial \alpha_l} \Delta \alpha_l = \sum_l F_l \Delta \alpha_l \quad (5.26)$$

之后我们就将变换关系带回到运动方程不变的条件中. 虽然在变换中  $t$  可以视为一个广义坐标, 但是在计算作用量时其作为积分变元, 地位还是相对特殊. 我们首先考虑积分变元和上下限的改变. 对时间的变换取微分得到:

$$d\bar{t} = \left(1 + \frac{d\Delta t}{dt}\right) dt \quad (5.27)$$

接下来, 因为积分函数的变化为一阶小, 积分上下限的变化也为一阶小, 因此积分上下限带来的积分值的变化为可忽略的二阶小量, 这样我们就可以把  $S$  的变化化成相对紧凑的形式:

$$\bar{S} - S = \int_{t_1}^{t_2} dt \left[ L(\bar{q}, \dot{\bar{q}}, t) \left(1 + \frac{d\Delta t}{dt}\right) - L \right] \quad (5.28)$$

为了形式的简洁, 我们先不引入变换的具体形式, 只记

$$\sum_l T_l^k \Delta \alpha_l = \Delta q^k \quad (5.29)$$

这样, 我们分别写出要求的  $d\bar{q}^i$  和  $\bar{H}$ :

$$d\bar{q}^i = dq^i + d\Delta q^i \quad (5.30)$$

$$\bar{L} = L + \sum_i \frac{\partial L}{\partial q^i} \Delta q^i + \frac{\partial L}{\partial \dot{q}^i} \Delta \dot{q}^i \quad (5.31)$$

带回原来的式 (5.28) 就有

$$= \int_{t_1}^{t_2} dt \left[ \sum_i \left( \frac{\partial L}{\partial q^i} \Delta q^i + \frac{\partial L}{\partial \dot{q}^i} \Delta \dot{q}^i \right) + \frac{\partial L}{\partial t} + L \frac{d\Delta t}{dt} \right] \quad (5.32)$$

这里进行的两个凑微分:

$$\frac{\partial L}{\partial \dot{q}^i} \Delta \dot{q}^i = \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}^i} \Delta q^i \right) - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}^i} \right) \Delta q^i \quad (5.33)$$

$$L \frac{d\Delta t}{dt} = \frac{d}{dt} (L \Delta t) - \left( \sum_i \left( \frac{\partial L}{\partial \dot{q}^i} \dot{q}^i + \frac{\partial L}{\partial \ddot{q}^i} \ddot{q}^i + \frac{\partial L}{\partial t} \right) \right) \Delta t \quad (5.34)$$

需要再进行一步凑微分

$$\frac{\partial L}{\partial \ddot{q}^i} \ddot{q}^i = \frac{d}{dt} \frac{\partial L}{\partial \ddot{q}^i} \dot{q}^i - \dot{q}^i \frac{d}{dt} \left( \frac{\partial L}{\partial \ddot{q}^i} \right) \quad (5.35)$$

分别带回得到

$$\begin{aligned} \Delta S = & \int_{t_1}^{t_2} dt \left[ \sum_i \left( \frac{\partial L}{\partial q^i} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}^i} \right) \right) (\Delta \dot{q}^i + \dot{q}^i \Delta t) + \dots \right. \\ & \left. \dots + \frac{d}{dt} \left[ \frac{\partial L}{\partial \dot{q}^i} \Delta q^i + L \Delta t - \frac{\partial L}{\partial \ddot{q}^i} \dot{q}^i \Delta t \right] \right] \end{aligned} \quad (5.36)$$

要求相差  $F(q, t)$  关于时间的全导数且积分值在任何轨道下相等. 这样就有

$$\frac{d}{dt} \left[ \frac{\partial L}{\partial \dot{q}^i} \Delta q^i + L \Delta t - \frac{\partial L}{\partial \ddot{q}^i} \dot{q}^i \Delta t - \Delta F \right] = \sum_i \left( \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}^i} \right) - \frac{\partial L}{\partial q^i} \right) (\Delta \dot{q}^i + \dot{q}^i \Delta t) \quad (5.37)$$

我们试图找出的守恒量应当实在真实运动下守恒的, 所以代入 Lagrange 方程得到右边的项都得零, 这样我们就导出了一个量  $J$

$$J = \sum_i p_i \Delta q^i - H \Delta t - \Delta F \quad (5.38)$$

满足条件

$$\frac{dJ}{dt} = 0 \quad (5.39)$$

因此  $J$  在真实运动下守恒. 现在我们可以带回对称变换的具体形式:

$$J = \sum_l \left( \sum_i p_i T_l^i - H T_l^0 - F_l \right) \Delta \alpha_l \quad (5.40)$$

由于  $\Delta \alpha_l$  均是任意选取的常数, 所以我们只能要求

$$\sum_i p_i T_l^i - H T_l^0 - F_l = \text{Const.} \quad \forall l = 1, \dots, r \quad (5.41)$$

这样, 我们就通过一种有  $r$  个独立参量的对称变换导出了  $r$  个守恒量.



## 参考文献

- [1] 李岩松 《分析力学讲义》
- [2] 高显 《经典力学》
- [3] 陈童 《经典力学新讲》

