Classical Mechanics Notes

经典力学笔记

惠中懿

2021.1

经典力学笔记

惠中懿

2021 年 3 月 27 日
公开第一版

前面的话

0.1 声明

由于我的学识浅薄、目光短浅，再加上时间紧张，本笔记中存在着大量的错误、偏见，请谨慎阅读．本笔记中大量摘抄引用其他书籍、个人的观点，原创性很低．请不要期望从本笔记中学习到知识，有无数更加优秀的书籍材料都是更好的选择．本笔记还没有写完，在我以后的学习生涯中会不断地将我的收获总结下来．本笔记最主要的价值就是我自己的思考与总结，对于读者的帮助可能非常有限．另外，有些内容是我在学习的早期写下的，可能目前我自己都不会再认同了．但是由于时间比较紧张，我也没有再重新阅读一遍将其删去．

如果发现错误或想与我交流，请联系我，我的邮箱为huizhy@mail2.sysu.edu.cn

0.2 参考教材

1. 高显《经典力学讲义》，这是我最主要的教材．我对于经典力学的知识大多是从这里得到的．

   我很喜欢经典力学，也很喜欢高显老师．高显老师来给我上经典力学是我的幸运．

2. 朗道《力学》，除了老师的讲义以外，我看的最多的是朗道《力学》．

3. 戈德斯坦《经典力学》，内容很全，例子很多，可以作为参考书．

4. 阿诺德《经典力学的数学方法》，我从中学到了时空观的数学表述．

0.3 有感而发

学习应当学到什么程度？thinkthink 这些问题都需要想．用你的大脑去分析．

慢慢想，想到哪里就写到哪里．首先学习应当学到什么程度这个问题是由学习的目的所决定的．比如学理论力学的目的是，

目录

第一部分

时空观

第一章　基本公设：时空观

上下四方曰宇，往古來今曰宙.

《尸子》

这一章会介绍经典力学的基本公设，所谓基本公设就是指时空观. 首先介绍经典力学的基本公设，然后用数学语言严谨地表述基本公设. 最后我们还会学习一点相对论时空观.

1.1　经典力学的基本公设

公设 1.1.1: 时空的数学结构

我们所在的空间是三维欧氏空间，时间则是一维的.

公设 1.1.2: 伽利略相对性原理

存在一些参考系（坐标系）称为惯性系，它们有以下两个性质：

1. 一切自然规律在任何时刻，在所有惯性系中都相同.
2. 相对于一个惯性系作匀速直线运动的一切参考系也都是惯性系.

1.2　基本公设的数学表述

中国诗歌可以翻译成英文，但是译作总是不如原本的诗句. 如果我们将自然规律类比成诗歌，那么章节1.1中的文字表述就相当于被翻译后的诗歌. 而我相信原汁原味的诗歌一定是用最简洁的数学语言创造的. 为此，我们来学习基本公设的数学表述.

1.2.1　位形空间

定义 1.2.1: 位形空间

位形空间是一个物理系统可能处于的所有状态的空间.

这里需要解释一下"状态"是什么. 比如对于经典力学来说，粒子的位置肯定是状态，但是粒子的速度是状态吗？我觉得不是的，因为我们研究的只是粒子位置的状态. 如果抽象一点说，我们想研究的只是 $\boldsymbol{r} \in \mathbb{R}^3$ 的空间中，$\boldsymbol{r}$ 如何变化.

位形是位置这一概念的推广．但是我们为什么要放弃直观形象的位置，而去研究位形呢？因为我们不想只研究位置的变化，我们还想研究电场磁场、弦、膜等等物理系统的变化！而这些物理系统是没有简单的“位置”的概念的．[1]

1.2.2 仿射空间

空间中的点经常会用坐标来描述，而坐标想要有意义，就一定需要原点．可是，既然所有惯性系都是平权的，那么以每一个点为宇宙的原点都是等价的．这就说明了，我们的宇宙是没有原点的．因此我们希望找到一种描述点的方式，能够不用坐标，也就不需要原点了．而这种思想恰恰和数学中的仿射空间一致．

我们简单描述一下仿射空间，这并不是严格的定义．

定义 1.2.2: 仿射空间

若点集 A 关联了一个线性空间 $\vec{A}$, 它们之间定义二元运算加法 $+$ 如下

$$A \times \vec{A} \to A$$

$$(p, v) \mapsto p + v$$

即线性空间中的元素 $T \in \vec{A}$ 构成对点 $p \in A$ 的变换, 变换到 A 中的另外一点 $p + v$. 如果这里的加法 $+$ 对应着加群 $\vec{A}(0, +)$ 使得加群的各种性质在以上变换中也能保持, 即 $\forall v_1, v_2 \in \vec{A}$ 以及取加群的单位元 $0 \in \vec{A}(0, +)$ 有:

1. 单位元: $p + 0 = p$
2. 结合律: $(p + v_1) + v_2 = p + (v_1 + v_2)$
3. 交换律: $(p + v_1) + v_2 = (p + v_2) + v_1, \forall p \in A$

则 $v \mapsto p + v$ 是 $\vec{A} \to A$ 的单满映射.
这样的点集 A 称为仿射空间 (affine space), 而线性空间 $\vec{A}$ 在以上变换上构成仿射空间的平移 (translation) 变换群, 群运算即 $\vec{A}$ 中的加法.

简单来说，仿射空间就是不需要原点的线性空间．比如说势能，我们可以任意选择势能零点，但是势能同样满足实数的加法、数乘．

定义 1.2.3: 距离

$$\rho(x, y) = ||x - y|| = \sqrt{(x - y) \cdot (x - y)}$$

仿射空间带上这个距离函数就叫做欧氏空间，记作 E^n

定义 1.2.4: 标架

$$\text{标架} = \text{一个点} + \text{基向量组}$$

假设有两个标架 $(P_0, u_1, u_2, u_3), (Q_0, v_1, v_2, v_3)$，由于 Q_0, v_1, v_2, v_3 都可以由 P_0, u_1, u_2, u_3 表示，反

[1] 普适的代价是抽象

之也是如此，那么两个标架之间就可以相互表示. 设

$$\begin{aligned} u_1 &= \gamma_{11}v_1 + \gamma_{12}v_2 + \gamma_{13}v_3 + 0P_0 \\ u_2 &= \gamma_{21}v_1 + \gamma_{22}v_2 + \gamma_{23}v_3 + 0P_0 \\ u_3 &= \gamma_{31}v_1 + \gamma_{32}v_2 + \gamma_{33}v_3 + 0P_0 \\ Q_0 &= \gamma_{41}v_1 + \gamma_{42}v_2 + \gamma_{43}v_3 + \mathrm{P}_0 \end{aligned} \tag{1.1}$$

用矩阵可以表示成

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ P_0 \end{bmatrix} = \begin{bmatrix} \gamma_{11} & \gamma_{21} & \gamma_{31} & \gamma_{41} \\ \gamma_{12} & \gamma_{22} & \gamma_{32} & \gamma_{42} \\ \gamma_{13} & \gamma_{23} & \gamma_{33} & \gamma_{43} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ Q_0 \end{bmatrix} \tag{1.2}$$

这样的标架变换就是一个仿射变换，上式的矩阵定义了仿射变换.

1.2.3 伽利略时空构造

定义 1.2.5: 伽利略时空构造

伽利略时空构造包含以下三要素：

1. 宇宙——这是一个四维仿射空间 A^4. A^4 中的点称为世界点或事件. 宇宙 A^4 中的平移构成一个矢量空间 $\mathbb{R}^4$.
2. 时间——这是一个线性映射 $t:\mathbb{R}^4 \to \mathbb{R}$ ，由宇宙的平移矢量空间映到实的“时间轴”，由事件 $a \in A^4$ 到事件 $b \in A^4$ 的时间间隔即数 $t(b-a)$. 若 $t(b-a)=0$ ，就说事件 a 和 b 是同时的.
3. 同时事件的距离 $\rho(a,b) = ||a-b||$

定义 1.2.6: 伽利略空间

具有伽利略时空构造的空间 A^4 就是伽利略空间

定义 1.2.7: 伽利略变换

伽利略群就是伽利略空间中保持伽利略时空构造的一切变换所成的群. 这个群的元素称为伽利略变换

一个伽利略时空构造下有很多个参考系，而这些参考系的伽利略时空构造都相同. 因此这些参考系之间的变换是保持伽利略时空构造的变换，即伽利略变换.

定义 1.2.8: 惯性系

根据伽利略相对性原理，在物理时空中有一个特定的伽利略构造，力学规律对于这个伽利略构造的伽利略变换协变，我们称这个伽利略构造下的参考系为惯性参考系.

时间平移是伽利略变换，因此惯性系中的力学规律具有时间平移不变性，意味着在惯性系中自然法则是恒定的.

时间反演是伽利略变换，因此惯性系中的力学规律具有时间反演不变性，意味着在惯性系中自然法则是可逆的.

空间平移是伽利略变换，因此惯性系中的力学规律具有空间平移不变性，意味着在惯性系中空间是均匀的.

空间旋转是伽利略变换，因此惯性系中的力学规律具有空间旋转不变性，意味着在惯性系中空间是各向同性的.

由上述四点，可以得到惯性系的重要性质.

定理 1.2.1: 惯性系的性质

1. 时间均匀性
2. 时间可逆性
3. 空间均匀性
4. 空间各向同性

1.3 相对论时空观

1.3.1 相对论的基本公设

相对论的基本公设是光速不变和狭义相对性原理. 光速不变描述了时空的数学结构，而狭义相对性原理则和伽利略相对性原理完全一致. [2]我们从这里可以看出相对性原理的重要性，以及经典力学和相对论的相似性——二者仅仅是时空的数学结构不同，除此之外完全相同.

1.3.2 相对论基本公设的数学表述

根据基本公设，我们可以构造狭义相对论的时空. 在狭义相对论的公设下，我们的宇宙是一个闵可夫斯基空间.

1.3.3 度规

未完待续 $\cdots$

[2]除了惯性系变换的区别，这个区别也是由于是空的数学结构不同导致的.

第二章　数学预备知识

最小作用量原理需要一些数学的预备知识，比如多元函数微分学、泛函、变分法，我们先来了解一下.

2.1　多元微分学

多元微分学对于理解拉格朗日量有很大帮助，此外理论力学中的函数多是多元函数，电动力学也会用到多元微分学，因此完全有必要好好学一下多元微分学.

多元微分学的内容就是：处理从一个欧氏空间到另一个欧氏空间的映射 $f:\mathbb{R}^n\to\mathbb{R}^m$，并试图理解：这种映射的导数是什么？既然欧氏空间对于多元微分学如此重要，那么我们一定会需要线性代数的知识.

2.1.1　映射

定义 2.1.1: 线性变换

从欧几里得空间 $\mathbb{R}^n$ 到欧几里得空间 $\mathbb{R}^m$ 的函数 T, 如果满足下述两公理, 就叫作线性变换.

1. (加性) 对于每两个 $\boldsymbol{x},\boldsymbol{x}'\in\mathbb{R}^n$, 有 $T\left(\boldsymbol{x}+\boldsymbol{x}'\right)=T\boldsymbol{x}+T\boldsymbol{x}'$
2. (齐性) 对于每个 $x\in\mathbb{R}^n$ 和每个 $c\in\mathbb{R}$, 有 $T(cx)=cTx$

定理 2.1.1: 矩阵与线性变换

$\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性变换与 $m\times n$ 之间具有一一对应关系．即每个矩阵都是一个线性变换，且每一个线性变换都由一个矩阵给出．且两个线性变换的复合等价于对应的两个矩阵的乘积.

2.1.2　映射的导数

首先回顾映射 $:R\to R$ 的导数，即一元函数的导数．不正式地说，我们有这个式子：

$$f(x)-f(x_0)\approx f'(x_0)(x-x_0)$$

对于多元函数的导数，我们当然希望上述式子的成立，即

$$f(\boldsymbol{x})-f(\boldsymbol{x_0})\approx f'(\boldsymbol{x_0})(\boldsymbol{x}-\boldsymbol{x_0})$$

其中 $f(\boldsymbol{x})-f(\boldsymbol{x_0})$ 是 $\mathbb{R}^m$ 中的矢量，而 $\boldsymbol{x}-\boldsymbol{x_0}$ 是 $\mathbb{R}^n$ 中的矢量，那么 $f'(\boldsymbol{x_0})$ 就必须是一个线性变换，或者说矩阵．可以证明导数的唯一性，于是 $f'\left(x_0\right)$ 是从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的满足下式的唯一的线性变换

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0;\boldsymbol{x}\neq\boldsymbol{x}_0}\frac{\left\|f(\boldsymbol{x})-\left(f\left(\boldsymbol{x}_0\right)+f'\left(\boldsymbol{x}_0\right)\left(\boldsymbol{x}-\boldsymbol{x}_0\right)\right)\right\|}{\left\|\boldsymbol{x}-\boldsymbol{x}_0\right\|}=0$$

有时我们也把 f' 称为 f 的全导数，区别于下面介绍的偏导数和方向导数．全导数也和下面介绍的导数矩阵有着紧密的联系．

2.1.3 偏导数和方向导数

定义 2.1.2: 方向导数

设 E 是 $\mathbb{R}^n$ 的子集合, $f: E \to \mathbb{R}^m$ 是函数, 设 x_0 是 E 的内点．并设 v 是 $\mathbb{R}^n$ 中的非零向量．如果极限

$$\lim_{t\to 0; t>0, \boldsymbol{x}_0+t\boldsymbol{v}\in E} \frac{f\left(\boldsymbol{x}_0+t\boldsymbol{v}\right)-f\left(\boldsymbol{x}_0\right)}{t}$$

存在, 则说 f 在 $\boldsymbol{x}_0$ 沿方向 v 可微, 并把上面的极限记作 $\boldsymbol{D}_{\boldsymbol{v}} f\left(\boldsymbol{x}_0\right)$

$$\boldsymbol{D}_{\boldsymbol{v}} f\left(\boldsymbol{x}_0\right) := \lim_{t\to 0; t>0} \frac{f\left(\boldsymbol{x}_0+t\boldsymbol{v}\right)-f\left(\boldsymbol{x}_0\right)}{t}$$

2.2 变分法

最小作用量原理完全依赖于变分法，此外，变分法在物理学中还有很多应用．变分法是求泛函极值的操作，其中会涉及到变分以及泛函导数的计算等等．

2.2.1 泛函

物理学中经常需要函数到数的映射，这就是泛函．

定义 2.2.1: 泛函

将具备某种性质的函数的集合记作 D．对于 $\forall f\in D$，存在唯一的 $S\in R$ 与之对应．我们把 S 叫做 $f(x)$ 的泛函，记作 $S=S[f]$．

根据定义，$\int_{x_1}^{x_2} f(x)\mathrm{d}x$ 是 $S[f]$．设 $g(x)=f'(x)$，$\int_{x_1}^{x_2} g(x)$ 也肯定是 $S[g]$，问题是 $\int_{x_1}^{x_2} g(x)$ 是 $S[f]$ 吗？答案是肯定的，根据定义可知，对于 $\forall f\in D$，存在唯一的 $\int_{x_1}^{x_2} f'(x)$ 与之对应，满足定义．

绝大多数情况下，物理中所遇到的泛函都可以写成积分形式

$$S[f]=\int_{x_1}^{x_2} L(x, f(x), f'(x), \cdots, f^{(n)}, \cdots)$$

这里被积函数 $L=L\left(x, f(x), f'(x), f''(x), \cdots\right)$ 是 $x, f(x)$ 及其导数的某个组合．在经典力学中，我们要研究的拉格朗日量（Lagrangian，简称拉氏量）就是这种形式，此时 S 被称为作用量（action）．

物理中还有一些泛函，比如卡诺循环中的做功．做功是泛函意味着做功的大小和路径的函数有关，也就是和过程有关，这也说明了做功不是态函数．

2.2.2 函数的变分

变分的概念

泛函的自变量相当于函数，因变量相当于实数．我们总是要研究自变量随着因变量的变化，因此我们就必须研究自变量的变化．泛函的自变量的变化就是函数的变化，我们称之为变分．

定义 2.2.2: 变分

对于泛函 $S[f]$，定义变分为 $\delta f = f_2 - f_1$.

变分是两个函数之差，因此变分也是函数.

变分的运算规则

2.2.3 连续泛函

回顾函数的连续性定义，对于 $f(x)$，若当 Δx 充分小时，Δf 可以任意小，就称 $f(x)$ 是连续的. 同样的，我们也有泛函连续性的定义.

定义 2.2.3: 连续泛函

对于泛函 $S[f]$，若当 $\max\{\delta f\}$ 充分小时，ΔS 可以任意小，就称 $S[f]$ 是连续的

定义中的 $\max\{\delta f\}$ 用来要求 δf 的大小. 直观上看，这种方法也很合理，毕竟如果 $\max\{\delta f\}$ 充分小 δf 肯定也是更小的.

2.2.4 线性泛函

出于许多原因，物理学家关心函数或者泛函的线性部分或是一阶项. 比如函数中的导数，一阶近似等等. 这就需要我们有线性的概念. 在函数中，x 这样的一阶项就是天然的线性函数，不需要额外的定义了. 可是在泛函中，我们没有这种自然的概念，这就需要我们来自己定义线性泛函. [1]

定义 2.2.4: 线性泛函

若泛函 $S[f]$ 与 $f(x)$ 的关系是线性的，即 S 满足以下条件

1. $S[cf] = cS[f], c \in R$
2. $S[f_1 + f_2] = S[f_1] + S[f_2]$

则称 $S[f]$ 为线性泛函

例如 $S[f] = \int_a^b f(x)\mathrm{d}x$ 是线性泛函，而 $S[f] = \int_a^b f^2(x)\mathrm{d}x$ 就不是线性泛函.

2.2.5 泛函的变分

“泛函的变分”是“函数的微分”的推广.

我们正在研究泛函的变化，我们总是希望研究泛函变化的主要部分. 在函数中，$f(x)$ 的微分 $\mathrm{d}f$ 是函数增量 $\Delta f(x)$ 的主要部分，这个主要部分是满足线性性的. 泛函的变分也是泛函增量的主要部分，这个主要部分也是线性性的.

我们来举一个例子，直观地理解一下泛函的变分. 对于泛函

$$S[f] = \int_a^b f^2(x)\mathrm{d}x \tag{2.1}$$

[1] 这也许说明了线性是比一次更本质的说法.

当函数有一个变分是，泛函的变化量为

$$\begin{aligned}\Delta Q &= Q[f+\delta f]-Q[f]\\ &= \int_a^b (f(x)+\delta f(x))^2\mathrm{d}x - \int_a^b f(x)\mathrm{d}x\\ &= \int_a^b 2f(x)\delta f(x)\mathrm{d}x + \int_a^b (\delta f(x))^2\mathrm{d}x\end{aligned} \tag{2.2}$$

考察第一项 $\int_a^b 2f(x)\delta f(x)\mathrm{d}x$，根据定义，第一项是线性泛函.

再考察第二项 $\int_a^b (\delta f(x))^2$，经过一定的计算，可以看出

$$\lim_{\max\{\delta f\}\to 0}\frac{\int_a^b(\delta f(x))^2}{\max\{\delta f\}}=0 \tag{2.3}$$

我们可以说第二项是 δf 的高阶无穷小量.

我们能从中看出，泛函的增量可以被分解为两部分，第一部分是 δf 的线性泛函，第二部分是比 δf 更高阶的无穷小量. 这样我们就能够对泛函的变化的主要部分，也即线性部分——泛函的变分来进行定义了

定义 2.2.5: 泛函的变分

对于泛函 $S[f]$，给 $f(x)$ 以增量 δf，则泛函 $S[f]$ 有增量 $\Delta S=S[f+\delta f]-S[f]$. 如果 ΔS 可以表示成

$$\Delta S = T[f,\delta f]+\beta[f,\delta f]$$

其中 $T[f,\delta f]$ 是 δf 的线性泛函，且

$$\lim_{\max\{\delta f\}\to 0}\frac{\beta[f,\delta f]}{\max\{\delta f\}}\to 0$$

则，$T[f,\delta f]$ 称为泛函的变分，记作 δS

可见泛函 $S[f]$ 的变分 δS 是增量 ΔS 的线性部分.

2.2.6 泛函导数

我们先从比较不严谨的方式来引入泛函导数. 首先我们回顾函数的导数

$$\mathrm{d}f = f'(x)\mathrm{d}x \tag{2.4}$$

因此我们可以将导数 $f'(x)$ 理解成一个从 $\mathrm{d}x$ 到 $\mathrm{d}f$ 的映射.

同样地，我们也可以将泛函导数理解成一个从 δf 到 δS 的映射. 可是 δS 是实数，而 δf 是函数，因此泛函导数可以理解成是一个泛函. 下面我们就来具体地求一下这个泛函是什么.

未完待续 ···

2.2.7 泛函极值

泛函取极值的条件是泛函导数为零. 未完待续 ···

第二部分

拉格朗日力学

第三章　最小作用量原理

我发现它是绝对会令人神往的，并且自那时以来，我发现它总是那么引人入胜.

Feynman

要描述我们宇宙的物理规律，首先要给出宇宙的数学构造，这相当于背景．我在第1.1章中已经给出了时空观．其次就要给出物体在时空中的运动规律，也就是运动方程，这便是我们这一章要做的事情．在经典力学力学中，运动方程由最小作用量原理给出．我们会先学习最小作用量原理的表述，然后找出作用量的具体形式，最后根据最小作用量原理得到运动方程.

3.1　最小作用量原理

3.1.1　最小作用量原理

力学系统运动规律的最一般表述由最小作用量原理给出．根据这个原理, 每一个力学系统都可以用一个确定的函数

$$L\left(q_1, q_2, \cdots, q_s, \dot{q}_1, \dot{q}_2, \cdots, \dot{q}_s, t\right)$$

或者简记为 $L(q, \dot{q}, t)$ 所表征, 并且系统的运动还要满足下面的条件.

公设 3.1.1: 最小作用量原理

假设在时刻 $t=t_1$ 和 $t=t_2$ 系统的位置由两组坐标 $q^{(1)}$ 和 $q^{(2)}$ 确定. 那么, 系统在这两个位置之间的运动（世界线）使得积分

$$S=\int_{t_1}^{t_2} L(q, \dot{q}, t) \mathrm{d} t$$

取最小值．函数 L 称为给定系统的拉格朗日函数，简称拉氏量，积分 S 称为作用量.

根据最小作用量原理，作用量 S 应取极值，因此

$$\delta S=\delta \int_{t_1}^{t_2} L(q, \dot{q}, t) \mathrm{d} t=0 \tag{3.1}$$

为了求泛函导数，我们对积分号内求变分，

$$\begin{aligned}&\delta\int_{t_1}^{t_2}L(q,\dot{q},t)\mathrm{d}t\\=&\int_{t_1}^{t_2}\delta L(q,\dot{q},t)\mathrm{d}t\\=&\int_{t_1}^{t_2}\left(\frac{\partial L}{\partial q}\delta q+\frac{\partial L}{\partial \dot{q}}\delta\dot{q}\right)\mathrm{d}t\end{aligned}\tag{3.2}$$

其中

$$\frac{\partial L}{\partial \dot{q}}\delta\dot{q}=\frac{\partial L}{\partial \dot{q}}\frac{\mathrm{d}}{\mathrm{d}t}(\delta q)=\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\delta q\right)-\delta q\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right)=-\delta q\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right)\tag{3.3}$$

因此

$$\delta S=\int_{t_1}^{t_2}\delta q\left(\frac{\partial L}{\partial q}-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right)\right)\mathrm{d}t\tag{3.4}$$

泛函导数为

$$\frac{\partial L}{\partial q}-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right)\tag{3.5}$$

泛函取极值的条件是泛函导数取零，得到拉格朗日方程.

定理 3.1.1: 欧拉-拉格朗日方程

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}}-\frac{\partial L}{\partial q}=0$$

乍一看，最小作用量原理十分神奇，仿佛它告诉了我们很多事情，可是实际上最小作用量原理只是想告诉我们一个很简单的事实. 如果在数学上，每一个微分方程的解（运动方程）都能表达成某个函数取极值的条件（最小作用量），那么上式其实仅仅说明了运动方程是一个二阶微分方程！求解二阶微分方程需要坐标和坐标一阶导数（速度）的初值. 因此最小作用量原理等价于牛顿决定性原理.

公设 3.1.2: 牛顿决定性原理

一个力学系的初始状态（即其中各点在某一时刻的位置与速度）唯一地决定它的运动.

最小作用量原理的确仅仅说明了运动方程是一个二阶微分方程，看起来最小作用量原理好像什么都没说. 真的什么都没说吗？首先最一般地想，描述物理量演化的为何就不能是积分-微分方程呢，为何最遥远地过去不能对下一刻的物理量期望值产生最直接的影响呢. [1]其次，我们的确通过最小作用量原理得到了很多结论. 比如根据最小作用量原理，我们从时空对称性可以推出守恒量. [2]

3.1.2 题外话：如何理解最小作用量原理

前几天韦乐瑶同学问我：“你是如何理解最小作用量原理的，你是如何理解作用量的？”这个问题虽然不太具体，但是我觉得是值得思考的，我认为学物理的人应当多想想这些问题. 我也思考了一番：

我们学物理的人总是会认为：宇宙规律是很美的. 哪怕你不能欣赏这种美，但是起码也会说：宇宙规律是独一无二的，宇宙规律是唯一的——宇宙规律总是在某一方面是最特别的.

[1] 整理自https://www.zhihu.com/question/407148717/answer/1343973715回答和评论区

[2] https://www.zhihu.com/question/28316575/answer/40328382

如果我们把上帝看成一个艺术家，现在上帝开始设计这个宇宙的规律．上帝说：我要设计一个最美妙的宇宙．那么上帝总是要有一个标准的：如何衡量美妙的程度？先不要试图理解上帝的审美，假设上帝已经有了一个衡量美妙程度的物理量："美"．

那么我们的宇宙规律其实很简单

公设 3.1.3: 最美宇宙原理

宇宙规律是："美"总取最大值．

好奇的人类，想要探寻这个世界的真相，想要知道上帝是如何设计这个宇宙的．费马、欧拉、拉格朗日，这些伟大的人，他们经过无数的努力，终于窥见了一点：原来上帝对于美的标准是"负的作用量"．于是他们以自己的语言写下了宇宙的规律：作用量总是取最小值，"美" $=-$ 作用量，即"美"总是取最大值．

我就是这样理解最小作用量原理的，我就是这样理解这个世界的．

3.2 拉氏量的性质

我们要分析拉氏量的一些性质，为了更好地对操作拉氏量．

3.2.1 拉氏量的可加性

设力学系统由 A, B 两部分组成，如果每个部分都是封闭的，拉氏量分别为 L_A, L_B. 当两个部分的相互作用可以忽略的极限情况下，系统的拉氏量趋于二者之和．

定理 3.2.1: 拉格朗日函数的可加性

当系统 A,B 的距离足够远以至于相互作用可以忽略时，$\lim L_{A+B} = L_A + L_B$.

由于已经假设了相互作用可以忽略，这条性质是不证自明的．拉氏量的这种可加性反映了：每一个独立部分的运动方程不可能包含与另一部分相关的物理量．

3.2.2 拉氏量可乘任意常数

定理 3.2.2: 拉格朗日函数可以乘任意常数

将所有力学系统的拉格朗日函数乘任意常数，不会改变运动的微分方程．

孤立系统的拉氏量乘以任意常数都不会改变欧拉-拉格朗日方程．然而由于拉氏量的可加性，所有系统的拉氏量必须乘以相同的常数，这归结于物理度量单位的自然任意性．

3.2.3 拉氏量可以加上时间的全导数项

定理 3.2.3: 拉格朗日函数可以加上 $\frac{\mathrm{d}}{\mathrm{d}t}f(q,t)$

将拉格朗日函数加上一个对时间和坐标的函数的时间全导数，不会改变运动的微分方程．

证明. 我们要证明两个拉格朗日函数得到的拉格朗日方程的解相同. 思路是证明两个作用量相差一个常数，因此变分之后拉格朗日方程相同. 设

$$L'(q,\dot{q},t) = L(q,\dot{q},t) + \frac{\mathrm{d}}{\mathrm{d}t}f(q,t) \tag{3.6}$$

则有

$$\begin{aligned} S' &= \int_{t_1}^{t_2} L'(q,\dot{q},t)\mathrm{d}t = \int_{t_1}^{t_2} L(q,\dot{q},t)\mathrm{d}t + \int_{t_1}^{t_2} \frac{\mathrm{d}}{\mathrm{d}t}f(q,t)\mathrm{d}t \\ &= S + f\left(q^{(2)},t_2\right) - f\left(q^{(1)},t_1\right) \end{aligned} \tag{3.7}$$

其中 $f\left(q^{(2)},t_2\right) - f\left(q^{(1)},t_1\right)$ 为常数，因此 $\delta S' = \delta S$，即拉格朗日方程相同.

本质上是因为作用量是拉氏量对时间的积分，所以将拉氏量加上时间的全导数项相当于作用量加上常数. □

3.3 经典力学下作用量的具体形式

上一节我们学习了最小作用量原理，但是最小作用量原理其实并没有告诉我们作用量，拉氏量到底是什么. 神奇的是，我们仅仅需要知道最小作用量原理，就能推导出作用量和拉氏量的具体形式了！

我们分别从经典力学和相对论两种不同的时空观出发，推导出经典力学下的拉氏量和相对论下的拉氏量.

3.3.1 自由质点的拉格朗日函数

首先我们来推导惯性系中的自由质点的拉格朗日函数. 本小节大量摘抄朗道《力学》，我认为这是经典力学中非常精彩的地方.

我们现在只知道最小作用量原理和伽利略相对性原理，我们将从这两点出发.

根据最小作用量原理，最一般的拉氏量的形式为

$$L = L(t, \boldsymbol{x}, \boldsymbol{v}) \tag{3.8}$$

根据伽利略相对性原理，时间是均匀的，因此拉氏量不应包含 t. 根据伽利略相对性原理，空间是均匀的，因此拉氏量不应包含 $\boldsymbol{x}$. 根据伽利略相对性原理，空间是各向同性的，因此拉氏量仅与 $\boldsymbol{v}$ 的大小有关，与 $\boldsymbol{v}$ 的方向无关. 综上所述，拉氏量仅是 v^2 的函数. 即

$$L = L(v^2) \tag{3.9}$$

对拉氏量变分得到欧拉拉格朗日方程

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \boldsymbol{v}} = 0, \tag{3.10}$$

因此 $\frac{\partial L}{\partial \boldsymbol{v}}$ 为常数. 而 $\frac{\partial L}{\partial \boldsymbol{v}}$ 仅为速度的函数，因此

$$\boldsymbol{v} = 常数 \tag{3.11}$$

这样我们得到了惯性定律

定理 3.3.1: 惯性定律

在惯性参考系中，自由运动质点速度的大小和方向都不改变，这就是惯性定律.

我们继续推导拉氏量的形式．我们选取两个参考系应用伽利略相对性原理．如果惯性参考系 K 以无穷小速度 $\boldsymbol{\epsilon}$ 相对于另一惯性参考系 K' 运动，则有 $\boldsymbol{v}' = \boldsymbol{v} + \boldsymbol{\epsilon}$．根据伽利略相对性原理，在参考系 K, K' 的物理规律完全相同，即运动方程完全相同．再根据最小作用量原理，运动方程相同等价于拉氏量仅可以相差一个关于时间和坐标的函数的全导数．因此我们就来求一下二者的拉氏量．

$$L' = L(v'^2) = L(v^2 + 2\boldsymbol{v}\cdot\boldsymbol{\epsilon} + \epsilon^2). \tag{3.12}$$

将这个表达式展开成 ϵ 的幂级数并忽略一阶以上的无穷小量得到：

$$L(v'^2) = L(v^2) + 2\frac{\partial L}{\partial v^2}\boldsymbol{v}\cdot\boldsymbol{\epsilon}. \tag{3.13}$$

因此 $2\frac{\partial L}{\partial v^2}\boldsymbol{v}\cdot\boldsymbol{\epsilon}$ 必须是时间的全导数．由于 $\boldsymbol{v}$ 就是一个时间的全导数，因此 $2\frac{\partial L}{\partial v^2}\boldsymbol{\epsilon}$ 必须不依赖于速度，即 $\frac{\partial L}{\partial v^2}$ 不依赖于速度．而 $L = L(v^2)$，因此拉氏量必须与速度平方成正比！即

$$L = \frac{1}{2}mv^2 \tag{3.14}$$

其中 m 是为了满足成正比所乘的常数，根据作用量要取最小值而不是最大值，这个常数是非负的．我们把这个常数称为质量．并将 $\frac{1}{2}mv^2$ 称为动能 T，注意到这里的动能和牛顿力学中的动能相同．

这样我们就得到了自由质点的拉氏量．

3.3.2 自由质点系的拉格朗日函数

根据最小作用量原理，在不考虑相互作用的时候，拉格朗日函数具有可加性．因此自由质点系的拉氏量为

$$L = \sum_a \frac{1}{2}m_a v_a^2 \tag{3.15}$$

3.3.3 质点系的拉格朗日函数

现在我们来求考虑相互作用的质点系的拉氏量．我们要描述质点间的相互作用，也就是与速度无关的项．自然的想法是：我们在质点系的拉氏量中增加坐标的某一函数（根据相互作用的性质确定）．将这个函数记为 $-U$，并将 U 称为势能，则有

$$L = \sum_a \frac{m_a v_a^2}{2} - U(\boldsymbol{r}_1, \boldsymbol{r}_2, \cdots). \tag{3.16}$$

其中 $\boldsymbol{r}_a$ 是第 a 个质点的径矢．这是封闭质点系拉格朗日函数的一般形式．从这个形式上看，U 仅依赖于所有质点在同一时刻的位置．这件事情并不显然，它等价于“相互作用是瞬时传递的”，下面我们来具体说明一下．

我们用反证法．如果相互作用不是瞬时传递的，而是以一个有限速度传递的，那么根据速度的叠加原理[3]，不同参考系中的相互作用传播速度就会不同，这就违背了伽利略相对性原理．因此相互作用是瞬时传递的，反映相互作用的函数 U 也只和 $\boldsymbol{r}_1, \boldsymbol{r}_2, \cdots$ 有关．

那么 U 的具体形式是什么呢？

根据最小作用量原理的欧拉拉格朗日方程，将(3.16)带入欧拉-拉格朗日方程中，得到

$$m_a\frac{\mathrm{d}\boldsymbol{v}_a}{\mathrm{d}t} = -\frac{\partial U}{\partial \boldsymbol{r}_a}. \tag{3.17}$$

从这个式子可以根据相互作用得到势能的具体形式．从这个式子也可以看出，我们讨论的势能和牛顿力学中的势能相同．

[3]可以由时空的数学结构推出

3.3.4 经典力学下作用量的一般形式

综上所述，我们得到了拉氏量的一般形式

$$L = T - U, \tag{3.18}$$

T, U 的含义都在前几小节中讨论过了，它们与牛顿力学中的动能和势能相同，因此也可以通过牛顿力学来计算 T, U，从而计算 L.

因此作用量的一般形式为

$$S = \int_{t_1}^{t_2} (T - U)\mathrm{d}t. \tag{3.19}$$

其中 T, V 分别指代牛顿力学下的动能和势能.

3.4 相对论下的作用量的具体形式

同样地，我们只知道相对论下时空的数学结构、狭义相对性原理和最小作用量原理，我将只通过它们推导出自由粒子的拉氏量．这里可以再一次看出相对论和经典力学的相似．因此以下的推导对于不同的时空都是普适的，只要换一下度规即可.

根据相对性原理，物理定律在一切参考系中都具有相同的形式，因此“最小作用量原理”在一切参考系中都应具有相同的形式，也即作用量在广义坐标变换下不变，且作用量在广义坐标变换下始终最小.[4]即作用量必须是洛伦兹标量．而对于一个自由粒子，我们所能造出的唯一的洛伦兹标量[5]，只能与世界线的长度成正比．这样一来，对于一个自由粒子，作用量积分必须取下面的形式：

$$S = -\alpha \int_a^b \mathrm{d}s \tag{3.20}$$

其中 $\mathrm{d}s$ 是世界线的微元，α 是一个常数．可以证明，世界线的长度有最大值，无最小值．根据作用量要取最小值而非最大值，α 应取非负数.

我们下面将世界线的长度写成拉氏量的形式．即将 $\int_a^b \mathrm{d}s$ 写成 $\int_a^b L\mathrm{d}t$ 根据世界线长度与速度的定义，我们有

$$S = -\int_{t_1}^{t_2} \alpha c\sqrt{1 - \frac{v^2}{c^2}}\mathrm{d}t \tag{3.21}$$

为了与经典力学中的物理量 m 相对应，我们把质量定义为 $\frac{\alpha}{c}$.

这样相对论时空观下的拉氏量的形式为

$$L = -mc^2\sqrt{1 - \frac{v^2}{c^2}} \tag{3.22}$$

在非相对论极限中，拉氏量变成了

$$L = -mc^2 + \frac{1}{2}mv^2 + \frac{3}{8}m\frac{v^4}{c^2}\cdots \tag{3.23}$$

舍去常数以及高阶无穷小项，我们得到了自由质点在非相对论极限下的拉氏量为

$$L = \frac{1}{2}mv^2. \tag{3.24}$$

这与经典力学下的作用量形式完全相同.

[4]的确应该不变，如果只能保证作用量最小，但是可变的话，也是不满足相对性原理的！

[5]这是为什么？希望我以后能够学会如何证明．姑且完全接受吧！

第四章　守恒定律

连续对称性和守恒定律是一一对应的.

Noether

在这一章节中，我们将根据时空的连续对称性来推导出几个重要的守恒量. 所谓连续对称性是与镜像对称等相对的对称性.

4.1　能量

4.1.1　时间平移对称性

根据时间平移对称性，物理规律在不同的时间下具有相同的形式. 即运动方程在时间平移下相同. 我们考虑一个无穷小平移 ϵ_t，这个无穷小平移的改变使拉氏量产生的变化为

$$\delta L = \frac{\partial L}{\partial t}\epsilon_t \tag{4.1}$$

不考虑时间的全导数项[1]，对于任意的 ϵ_t 要求 $\delta L = 0$ 等价于

$$\frac{\partial L}{\partial t} = 0 \tag{4.2}$$

4.1.2　时间平移对称性下的守恒量

所谓守恒量是不随时间变化的量，即守恒量 A 一定满足

$$\frac{\mathrm{d}}{\mathrm{d}t}A = 0 \tag{4.3}$$

根据时间平移对称性得到的结论

$$\frac{\partial L}{\partial t} = 0, \tag{4.4}$$

以及欧拉-拉格朗日方程，我们就能够找到时间平移对称性下的守恒量.

我们要找出 $\frac{\mathrm{d}}{\mathrm{d}t}$，因此我们首先求一下拉氏量对于时间的全导数

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \sum_i \frac{\partial L}{\partial t} + \sum_i \frac{\partial L}{\partial q_i}\dot{q}_i + \sum_i \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i \tag{4.5}$$

将(4.4)带入到(4.5)，得到

$$\sum_i \frac{\partial L}{\partial q_i}\dot{q}_i + \sum_i \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i = 0 \tag{4.6}$$

[1] 我没有认真地思考，以后再想想.

根据欧拉-拉格朗日方程，我们还有

$$\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} = 0 \tag{4.7}$$

带入得

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \sum_i (\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{p}_i} + \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i) \tag{4.8}$$

因此我们得到守恒量

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_i \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L\right) = 0 \tag{4.9}$$

定义 4.1.1: 能量

定义能量 E 为：通过时间均匀性推出的守恒量

$$E = \sum_i \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L$$

带入拉格朗日量的具体形式，得到

$$E = T + U \tag{4.10}$$

4.2 动量

4.2.1 空间平移对称性

由空间平移对称性可得，封闭力学体系在空间中整体平移时，其力学规律保持不变，即运动方程不变．与时间平移对称性思路一致，我们考虑无穷小位移 $\boldsymbol{\epsilon}_r$

$$\delta L = \sum_a \frac{\partial L}{\partial \boldsymbol{r}_a} \cdot \delta \boldsymbol{r}_a = \boldsymbol{\epsilon}_r \cdot \sum_a \frac{\partial L}{\partial \boldsymbol{r}_a} \tag{4.11}$$

对于任意 $\boldsymbol{\epsilon}_r$ 要求 $\delta L = 0$ 等价于

$$\sum_a \frac{\partial L}{\partial \boldsymbol{r}_a} = 0 \tag{4.12}$$

4.2.2 空间平移对称性下的守恒量

根据拉格朗日方程我们有

$$\frac{\partial L}{\partial \boldsymbol{r}_a} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \boldsymbol{v}_a} \tag{4.13}$$

根据空间平移对称性，上式为零

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \boldsymbol{v}_a} = 0 \tag{4.14}$$

即

$$\frac{\partial L}{\partial \boldsymbol{v}_a} \tag{4.15}$$

是不随时间变化的守恒量．

我们定义空间平移下的守恒量为动量．

定义 4.2.1: 动量

力学体系中某个广义坐标 q^a 对应的广义动量为

$$\boldsymbol{p}_a \equiv \frac{\partial L}{\partial \dot{q}^a} \equiv \frac{\partial L}{\partial v^a} \tag{4.16}$$

4.3 角动量

4.3.1 空间转动对称性

与能量、动量的思路一致，我们先找到空间转动对称性能给出什么样的条件.

对于一个无穷小转动 $\delta\boldsymbol{\phi}$，它的大小等于转角，方向沿着转动轴，因此有

$$\delta\boldsymbol{r} = \delta\boldsymbol{\phi} \times \boldsymbol{r}. \tag{4.17}$$

两边同时求导，得到

$$\delta\boldsymbol{v} = \delta\boldsymbol{\phi} \times \boldsymbol{v}. \tag{4.18}$$

带入拉氏量的变化（变分）中，得到

$$\begin{aligned}\delta L &= \frac{\partial L}{\partial \boldsymbol{r}} \cdot \delta\boldsymbol{r} + \frac{\partial L}{\partial \boldsymbol{v}} \cdot \delta\boldsymbol{v} \\ &= \dot{\boldsymbol{p}} \cdot \delta\boldsymbol{r} + \boldsymbol{p} \cdot \delta\boldsymbol{v} \\ &= \dot{\boldsymbol{p}} \cdot (\delta\boldsymbol{\phi} \times \boldsymbol{r}) + \boldsymbol{p} \cdot (\delta\boldsymbol{\phi} \times \boldsymbol{v}) \\ &= \delta\boldsymbol{\phi} \cdot (\boldsymbol{r} \times \dot{\boldsymbol{p}} + \dot{\boldsymbol{r}} \times \boldsymbol{p}) \\ &= \delta\boldsymbol{\phi} \cdot \frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{r} \times \boldsymbol{p})\end{aligned} \tag{4.19}$$

其中

$$\delta\boldsymbol{\phi} = (\delta\phi)\boldsymbol{e}_n \tag{4.20}$$

$\boldsymbol{e}_n$ 表示转轴方向.

体系在沿着 $\boldsymbol{e}_n$ 方向转动对称性，即有 $\delta L = 0$，因此

$$\boldsymbol{e}_n \cdot (\boldsymbol{r} \times \boldsymbol{p}) = 0 \tag{4.21}$$

4.3.2 空间转动对称性下的守恒量

这样我们就找到了空间转动对称性下的守恒量.

定义 4.3.1: 角动量

定义角动量

$$\boldsymbol{J} = \boldsymbol{r} \times \boldsymbol{p},$$

如果体系在沿着 $\boldsymbol{e}_n$ 方向具有转动不变性，则有体系在 $\boldsymbol{n}$ 方向上的总角动量守恒.

$$\boldsymbol{e}_n \times \boldsymbol{J} = 常数$$

4.4 诺特定理

总结一下，在寻找能量、动量、角动量的时候，我们都是先根据连续对称性得到一些条件．连续对称性允许我们取无穷小变化，这会给出有用的条件．然后根据这些条件我们找到了守恒量．

已经可以看出，时空对称性与守恒量之间有着重要的联系．然而事实却更出乎我们的意料：每一个连续对称性都对应了一个守恒量，这便是诺特定理．

4.4.1 一般的连续对称性

假设力学系统在某种一般的连续变化下是对称的．这个一般的连续变化可以表示为广义坐标的变换

$$q^a(t) \to q^a(t) + \delta q^a(t) \tag{4.22}$$

在这个变化下，通过计算变分，拉氏量的变化为

$$\delta L = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^a}\delta q^a\right) \tag{4.23}$$

力学系统是对称的意味着拉氏量可以相差一个时间的全导数项，即

$$\delta L = \frac{\mathrm{d}J}{\mathrm{d}t} \tag{4.24}$$

联立(4.23)和(4.24)得到

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^a}\delta q^a - J\right) = 0 \tag{4.25}$$

4.4.2 一般变换对称性下的守恒量

这样我们就得到了一般变换对称性下的守恒量

$$\frac{\partial L}{\partial \dot{q}^a}\delta q^a - J \tag{4.26}$$

这就是诺特定理．

4.5 运动积分

未完待续 $\cdots$

第五章　两体问题

我曾测量天空，现在测量幽冥．灵魂飞向天国，肉体安息土中．

Kepler

高显老师说过：理论力学中能有精确解的问题不多，一个是两体问题，另一个是简谐振动．其他没有解析解的运动只能由两体问题和简谐振动合成近似．这足以说明两体问题和简谐振动的重要性，我们现在就来学习一下两体问题．

5.1　两体系统

5.1.1　两体系统的拉氏量

所谓两体系统指的是：两个相互作用的粒子组成的封闭体系．我们这一章研究研究两体系统的运动问题，即所谓的两体问题．

我们将用拉格朗日力学研究两体系统．为此我们将求出两体系统的拉氏量．动能很好求，我们来求势能．势能分为相互作用的势能和外场的势能．由于外场的势能不包括相互作用，因此外场的总势能等于两个粒子各自在外场的势能之和．

另外，根据空间平移对称性，两粒子之间的相互作用只与 $\boldsymbol{r}_1-\boldsymbol{r}_2$ 有关．

因此，两体系统的拉氏量的一般形式为

$$L=\frac{1}{2}m_1\dot{\boldsymbol{r}}_1^2+\frac{1}{2}m_2\dot{\boldsymbol{r}}_2^2-V^{in}(\boldsymbol{r}_1-\boldsymbol{r}_2)-V^{ex}(\boldsymbol{r}_1)-V^{ex}(\boldsymbol{r}_2) \tag{5.1}$$

5.1.2　两体系统的解耦

一般来说，我们只能求解分离变量的微分方程，而两体系统中相互作用势能将 $\boldsymbol{r}_1,\boldsymbol{r}_2$ 耦合在了一起，我们需要将其解耦．由于外场的存在，即便我们将内场解耦了，外场又会重新耦合，因此下面的问题我们将忽略外场效应，这样两体系统总是可以解耦的．这样总势能就只有相互作用势能 V^{in} 了，我们将其记为 V．

$$L=\frac{1}{2}m_1\dot{\boldsymbol{r}}_1^2+\frac{1}{2}m_2\dot{\boldsymbol{r}}_2^2-V(\boldsymbol{r}_1-\boldsymbol{r}_2) \tag{5.2}$$

由于我们不知道 V 的具体形式，要想解耦就必须把 $\boldsymbol{r}_1-\boldsymbol{r}_2$ 当作一个整体．因此我们令 $\boldsymbol{r}=\boldsymbol{r}_1-\boldsymbol{r}_2$．系统有两个自由度，因此我们还要求出另一个广义坐标．注意到此时动能又偶合在了一起，因此我们对动能配方．用 $\boldsymbol{r}$ 替换 $\boldsymbol{r}_1$，得到

$$L=\frac{1}{2}m_1\dot{\boldsymbol{r}}^2+\frac{1}{2}(m_1+m_2)\dot{\boldsymbol{r}}_2^2+m_1\dot{\boldsymbol{r}}\cdot\dot{\boldsymbol{r}}_2-V(\boldsymbol{r}) \tag{5.3}$$

配方得到另一个广义坐标为

$$\boldsymbol{r}_c = (\boldsymbol{r}_2 + \frac{m_1}{m_1+m_2}\boldsymbol{r}) = \frac{m_1\boldsymbol{r}_1 + m_2\boldsymbol{r}_2}{m_1+m_2} \tag{5.4}$$

这样，拉氏量变为

$$L = \frac{1}{2}\frac{m_1 m_2}{m_1+m_2}\dot{\boldsymbol{r}}^2 + \frac{1}{2}(m_1+m_2)\dot{\boldsymbol{r}}_c^2 - V(\boldsymbol{r}) \tag{5.5}$$

这样我们就将两体系统的拉氏量解耦了．

回顾一下，我们的方法是：坐标变换．坐标变换在理论力学中是一个重要的方法，或者说是最普遍的方法，大多东西都已固定，我们还能做出什么变化呢？可能只有坐标变换了．坐标变换让我们将两体系统解耦，还可以让我们得到哈密顿方程，哈密顿-雅可比方程．正则变换也是相空间的坐标变换．

5.1.3 两体运动约化成单体运动

对于一个忽略外场的孤立两体系统，它的质心应该是匀速或静止的．因此我们可以取对于质心相对静止的参考系．此时拉氏量为

$$L = \frac{1}{2}\left(\frac{m_1 m_2}{m_1+m_2}\right)\dot{\boldsymbol{r}}^2 - V(r) \tag{5.6}$$

那么现在拉氏量中就只有广义坐标 $\boldsymbol{r}$ 了，也就可以等效地看成是单体运动．

5.2 中心势场

既然两体运动可以约化成单体运动，我们下面就直接来研究单体运动，其中，中心势场是一种重要的体系．中心势场是指两体系统的相互作用势能只与两体之间距离的大小有关，与距离的方向无关，即

$$V(\boldsymbol{r}) = V(r) \tag{5.7}$$

5.2.1 中心势场的运动

由于中心势场的对称性，我们选取球坐标 r, θ, ϕ 来表示 $\boldsymbol{r}$．故拉氏量变为

$$L = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\phi}^2\right) - V(r) \tag{5.8}$$

5.2.2 角动量守恒

这一段高显老师推导的很漂亮，我直接照搬过来了．

拉氏量不显含 $\dot{\phi}$，相应的角动量 $\boldsymbol{J}_z$ 守恒．由于空间相对于中心势场的中心有各向同性，当我们把原点选在中心时，我们就可以任意地选择坐标轴的方向了．而不论如何选择，拉氏量都不变（因为各向同性），因此 $\boldsymbol{J}_z$ 都守恒．角动量在 z 轴方向的投影都是常数．这种情况只有一种可能，即中心势场中粒子角动量矢量的所有 3 个分量都是常数，即

$$\boldsymbol{J} = \boldsymbol{r}\times\boldsymbol{p} = \text{常矢量} \tag{5.9}$$

角动量守恒可以推导出中心势场的粒子必做平面运动．因此

$$L = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) - V(r) \tag{5.10}$$

从三维两体问题的非解耦变成解耦，再变成三维单体问题，再变成二维单体问题，这简直太强大了．

未完待续 $\cdots$

5.3 开普勒问题

与距离成平方反比的中心势场是最常见也是最重要的一种势场，例如牛顿引力势，静电势能. 这时的力学问题称为开普勒问题.

未完待续 ···

第六章　微扰论与小振动

> 如果世界上的 PDE 都有性质良好的解，那么我们就不会有湍流，不会有混沌，很可能也不会有生命．正因为我们的世界的物理规律本身就如此复杂，而 PDE 很大一部分就是在刻画物理规律、建立相应数学模型，所以我们就不能指望 PDE 都会是我们想象的那么美好啊．PDE 的复杂性就是我们这个世界本身的复杂性的一种反映．
>
> Yuhang Liu

6.1　微扰展开

6.1.1　从泰勒展开到微扰展开

我们花了很多篇幅来学习如何得到运动方程，自然我们也希望得到运动方程的解．运动方程一般都是偏微分方程．可是可悲[1]的是，人类目前只能解出少数几个线性偏微分方程，非线性偏微分方程更是解不出，然而现实世界是高度非线性的．

因此，想要探究非线性的运动系统，就只能通过将非线性的系统线性化．什么是线性化呢？我们知道级数是线性的．回顾 $f(x)$ 的泰勒级数

$$f(x)=\sum_{n=0}^{\infty}\frac{f^{(n)}(a)}{n!}(x-a)^n, \tag{6.1}$$

$f(x)$ 本身可能并不是线性函数，但是通过泰勒展开，却可以将 $f(x)$ 在 $x=a$ 附近展开成线性函数！这就将 $f(x)$ 线性化了！

这就是微扰论的思想基于这个思想，我们要对一个非线性的物理系统做线性化展开，才能解出运动方程．因此，很自然地，我们对作用量进行展开，然后得到线性的作用量，然后根据最小作用量原理得到线性的运动方程，最后得到运动方程的解

6.1.2　微扰展开

假设我们已知运动方程的某个特解 Q_0，称为背景．我们希望得到 Q_0 附近的现在考虑系统对这个特定解的偏离，记作

$$Q=Q_{(0)}+\epsilon q \tag{6.2}$$

[1] 可悲么？

这里 q 即表征系统对背景的偏离，被称作扰动．ϵ 是用来表征扰动大小的小量

根据泰勒展开的思想，我们想求出 $S[Q_0+\epsilon q]$，它总是可以表示成

$$\begin{aligned}
&S\left[Q_{(0)}+\epsilon q\right] \\
=&\int \mathrm{d}tL\left(t,Q_{(0)}+\epsilon q,\dot{Q}_{(0)}+\epsilon\dot{q}\right) \\
\equiv&\int \mathrm{d}t\left[L\left(t,Q_{(0)},\dot{Q}_{(0)}\right)\right. \\
&+\epsilon\left(\left.\frac{\partial L}{\partial Q}\right|_{(0)}q+\left.\frac{\partial L}{\partial\dot{Q}}\right|_{(0)}\dot{q}\right) \\
&+\epsilon^2\left(\frac{1}{2}\left.\frac{\partial^2 L}{\partial Q^2}\right|_{(0)}q^2+\left.\frac{\partial^2 L}{\partial Q\partial\dot{Q}}\right|_{(0)}q\dot{q}+\frac{1}{2}\left.\frac{\partial^2 L}{\partial\dot{Q}^2}\right|_{(0)}\dot{q}^2\right) \\
&\left.+\cdots\right] \\
\equiv&S_0+\epsilon S_1+\epsilon^2S_2+\epsilon^3S_3+\cdots
\end{aligned} \tag{6.3}$$

下面来分析一下展开的每一阶的意义：

- 第一项 S_0 表示背景作用量．
- 第二项 ϵS_1 表示作用量的一阶项．回顾泰勒展开，我们知道函数在极值处的泰勒展开的一阶项系数为零．由于最小作用量原理要求背景作用量取极值，因此 $S_1=0$，经计算也是如此．
- 第三项 ϵ^2S_2 是展开的第一个非零项，因此也是最重要的一项．微扰后的运动方程是由这一项决定的．

我们下面研究展开的第三项在单自由度和多自由度的形式．

6.1.3 单自由度

可以计算出

$$S_2[q]=\int\mathrm{d}t\left(\frac{1}{2}K\dot{q}^2-\frac{1}{2}Mq^2\right) \tag{6.4}$$

其中

$$K:=\left.\frac{\partial^2L}{\partial\dot{Q}^2}\right|_{(0)} \tag{6.5}$$

$$M:=\left.\left[-\frac{\partial^2L}{\partial Q^2}+\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial^2L}{\partial Q\partial\dot{Q}}\right)\right]\right|_{(0)} \tag{6.6}$$

6.1.4 多自由度

经计算

$$S_2\left[q^a\right]=\int\mathrm{d}t\left(\frac{1}{2}K_{ab}\dot{q}^a\dot{q}^b+\frac{1}{2}F_{ab}q^a\dot{q}^b-\frac{1}{2}M_{ab}q^aq^b\right) \tag{6.7}$$

其中

$$K_{ab}:=\left.\frac{\partial^2L}{\partial\dot{Q}^a\partial\dot{Q}^b}\right|_{(0)} \tag{6.8}$$

$$F_{ab}:=\left.\left(\frac{\partial^2L}{\partial Q^a\partial\dot{Q}^b}-\frac{\partial^2L}{\partial Q^b\partial\dot{Q}^a}\right)\right|_{(0)} \tag{6.9}$$

$$M_{ab} := \left[-\frac{\partial^2 L}{\partial Q^a \partial Q^b} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial^2 L}{\partial Q^a \partial \dot{Q}^b} \right) \right] \bigg|_{(0)} \tag{6.10}$$

注意 K_{ab} 和 M_{ab} 是对称矩阵，而 F_{ab} 是反对称矩阵.

计算的时候不一定要套这个公式，也许直接计算会更容易一些.

6.2 稳定平衡位形附近的微扰展开

我们在上一节中研究了一般情况下的二阶微扰展开，但是并非所有位形附近的展开都是值得研究的，在这里我们研究最简单也是最普遍的情况：在稳定平衡位形附近的微扰展开.

6.2.1 平衡位形

背景 Q_0 是平衡位形，意思是 Q_0 是系统运动方程的一组静态解，即 Q_0 不随时间变化，即

$$\dot{Q}_0 = \frac{\mathrm{d}Q_0}{\mathrm{d}t} = 0 \tag{6.11}$$

我们将其应用到运动方程中，求一下平衡位形的等价条件.

根据运动方程，我们有

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{Q}}\bigg|_{(0)} - \frac{\partial L}{\partial Q}\bigg|_{(0)} = 0 \tag{6.12}$$

其中第一项可以写成

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{Q}}\bigg|_{(0)} = \frac{\partial}{\partial Q}\left(\frac{\partial L}{\partial Q}\right)\bigg|_{(0)} \dot{Q}_{(0)} + \frac{\partial}{\partial \dot{Q}}\left(\frac{\partial L}{\partial Q}\right)\bigg|_{(0)} \ddot{Q}_{(0)} \equiv 0 \tag{6.13}$$

因此平衡条件为

$$\frac{\partial L}{\partial Q}\bigg|_{(0)} = 0 \tag{6.14}$$

带入 $L = T - V$，平衡条件为

$$\frac{\partial V}{\partial Q}\bigg|_{(0)} = 0 \tag{6.15}$$

6.2.2 稳定

背景 Q_0 是稳定的，意思是：围绕背景的小扰动 q 不会随着时间无限增大．这就要求扰动 $q(t)$ 的运动方程是振动方程.

稳定条件是指势能的二阶导数大于零．即

$$\frac{\partial^2 V}{\partial Q^2}\bigg|_{(0)} > 0 \tag{6.16}$$

因此稳定平衡的条件是：势能一阶导为零，二阶导大于零.

6.3 自由振动：微扰展开后运动方程的解

系统在稳定、平衡位形附近的展开的作用量，做变分的解就是自由振动的运动方程.

6.3.1 一维情况

经过数学推导，可以得到，自由振动的通解为

$$q(t) = A\sin(\omega t) + B\cos(\omega t) \tag{6.17}$$

其中 ω 称为**本征频率**.

将(6.17)带入到运动方程中，就会得到关于本征频率的方程，我们称之为**本征方程**.

6.3.2 二位情况

$$\begin{aligned} q(t) &= A_1\sin(\omega_1 t) + A_2\cos(\omega_1 t) \\ q(t) &= B_1\sin(\omega_2 t) + B_2\cos(\omega_2 t) \end{aligned} \tag{6.18}$$

一般来说 $\omega_1 \neq \omega_2$.

非简并

当 $\omega_1 \neq \omega_2$ 的时候，

简并

第三部分

哈密顿力学

第七章　哈密顿方程

诗与数学是近亲.

Hamilton

理论力学有两大理论体系：拉格朗日力学和哈密顿力学. 拉格朗日力学是用广义坐标和广义速度来描述系统的状态. 然而，我们同样可以通过广义坐标和广义动量来描述系统的状态，而且这种描述方式具有一系列优点，比如得到的方程形式非常简洁. 我们这章来学习如何用广义坐标和广义动量来描述力学规律.

7.1　勒让德变换

描述力学规律主要是要建立运动方程. 因此我们希望把欧拉-拉格朗日方程的变量变换成广义坐标和广义动量. 数学中的勒让德变换就可以从一组独立变量变换到另一组独立变量.

我们首先不考虑物理，来学习一下什么是勒让德变换.

7.1.1　勒让德变换的目的

我们希望换一组新的变量来描述原来的函数. 勒让德变换就可以帮助我们达到这个目的.

考虑 s 个变量 $\{v^a\}$ 的函数

$$L = L(v^a) = L(v^1, v^2, \cdots, v^s), \tag{7.1}$$

我们希望将 L 中的 $\{v^a\}$ 换成一组新的变量 $\{p_a\}$，这里 p_a 定义为

$$p_a = \frac{\partial L}{\partial v^a} = p_a(v^a), \quad i = 1, \cdots, s \tag{7.2}$$

可是，如果仅仅是将(7.2)直接带入拉氏量中，并没有换成一组新的变量. 虽然它的形式上变成了

$$L = L(p^a) \tag{7.3}$$

但是，我们对它求全导数，会发现还是有 v^a 项的. 这说明仅仅将 p^a 带入是不行的，我们还需要将拉氏量进行变换，得到一个只以 p^a 为变量的函数. 这个变换就是勒让德变换.

因此，勒让德变换的目的是要找到一个函数 H，使得

$$\frac{\mathrm{d}H}{\mathrm{d}v^a} = 0 \tag{7.4}$$

7.1.2 勒让德变换

我们直接给出勒让德变换的定义.

定义 7.1.1: 勒让德变换

函数 $L = L(v^a)$ 的勒让德变换 H 定义为

$$H = p_a v^a - L(v^a)$$

下面我们来证明，这样定义勒让德变换的确满足我们对于勒让德变换的期望

$$\begin{aligned} \mathrm{d}H &= \mathrm{d}\left(p_a v^a - L(v^a)\right) \\ &= v^a \mathrm{d}p_a + p_a \mathrm{d}v^a - \frac{\partial L}{\partial v^a}\mathrm{d}v^a \\ &= v^a \mathrm{d}p_a \end{aligned} \tag{7.5}$$

这个方程告诉了我们：哈密顿量只是 p_a 的函数．即 $H = H(p_a)$.

7.1.3 勒让德变换的可逆性

可逆性是一个重要的概念，大体上来说：一个变换是否可逆可以理解为这个变换是否包含了变换前的全部信息．根据隐函数定理，勒让德变换可逆的充要条件为黑塞矩阵

$$\frac{\partial p_a}{\partial v^b} = \frac{\partial^2 L}{\partial v^a \partial v^b} \tag{7.6}$$

非退化.

7.2 哈密顿正则方程

7.2.1 卡农

正则的英文是 Canon，即卡农，意思是规律的、规则的．有一首旋律非常规则的曲子就叫卡农《卡农》. 我个人很喜欢《卡农》，我可能会拉《卡农》已经有十年了，在我的记忆里，这是我自学的第一首非考级曲目．物理和音乐是这个世界上两大美妙的事物，我始终认为，人活在世界上要去追求美好．恰巧，正则方程和《卡农》都很符合我的审美．愿我在写下正则方程时，耳边会响起卡农，愿我在演奏卡农的时候，眼前会浮现正则方程.

7.2.2 正则方程

我们希望将物理规律的描述从坐标空间转移到相空间上，而勒让德变换恰可以让我们完成这一目的，因此我们做勒让德变换．根据定义，有

$$H = p\dot{q} - L \tag{7.7}$$

正如我们所料，对方程两边求全微分

$$\begin{aligned} \mathrm{d}H &= \dot{q}\mathrm{d}p + p\mathrm{d}\dot{q} - \frac{\partial L}{\partial t}\mathrm{d}t - \frac{\partial L}{\partial q}\mathrm{d}q - \frac{\partial L}{\partial \dot{q}}\mathrm{d}\dot{q} \\ &= -\frac{\partial L}{\partial t}\mathrm{d}t - \frac{\partial L}{\partial q}\mathrm{d}q + \left(p - \frac{\partial L}{\partial \dot{q}}\right)\mathrm{d}\dot{q} + \dot{q}\mathrm{d}p \end{aligned} \tag{7.8}$$

根据动量的定义，我们有

$$p-\frac{\partial L}{\partial \dot{q}}=0 \tag{7.9}$$

根据欧拉-拉格朗日方程，我们有

$$\dot{p}=\frac{\partial L}{\partial q} \tag{7.10}$$

因此哈氏量的全微分变成了

$$\mathrm{d}H=-\frac{\partial L}{\partial t}-\dot{p}\mathrm{d}q+\dot{q}\mathrm{d}p \tag{7.11}$$

根据全微分的定义

$$\mathrm{d}H=\frac{\partial H}{\partial t}\mathrm{d}t+\frac{\partial H}{\partial q}\mathrm{d}q+\frac{\partial H}{\partial p}\mathrm{d}p \tag{7.12}$$

我们就得到了哈密顿正则方程

定理 7.2.1: 哈密顿正则方程

$$\dot{q}_i=\frac{\partial H}{\partial p_i},\quad \dot{p}_i=-\frac{\partial H}{\partial q_i}$$

可以证明，正则方程等价于拉格朗日方程．由于方程的形式简单且对称，也称为正则方程．

注意到我们还有

$$\frac{\partial H}{\partial t}=-\frac{\partial L}{\partial t}=\frac{\mathrm{d}H}{\mathrm{d}t} \tag{7.13}$$

证明. 对(7.11)两边同时除以 $\mathrm{d}t$，再加上正则方程易证． □

7.3 Routhian 函数

7.3.1 Routhian 函数

上一节我们将全部的广义速度用广义动量代替．但是在某些情况下，将仅为部分的而不是全部的广义速度用广义动量来代替会更加方便．这就是这一节我们要做的．

为了简化公式，首先假设仅有两个广义坐标，用 q 和 ξ 表示．我们将进行从 $q,\xi,\dot{q},\dot{\xi}$ 到 $q,\xi,p,\dot{\xi}$ 的 Routhian 变换，其中 p 为相应于广义坐标 q 的广义动量．

拉格朗日函数 $L(q,\xi,\dot{q},\dot{\xi})$ 的微分为

$$\begin{aligned}\mathrm{d}L&=\frac{\partial L}{\partial q}\mathrm{d}q+\frac{\partial L}{\partial \dot{q}}\mathrm{d}\dot{q}+\frac{\partial L}{\partial \xi}\mathrm{d}\xi+\frac{\partial L}{\partial \dot{\xi}}\mathrm{d}\dot{\xi}\\&=\dot{p}\mathrm{d}q+p\mathrm{d}\dot{q}\end{aligned} \tag{7.14}$$

可得

$$\mathrm{d}(L-p\dot{q})=\dot{p}\mathrm{d}q-\dot{q}\mathrm{d}p+\frac{\partial L}{\partial \xi}\xi+\frac{\partial L}{\partial \dot{\xi}}\mathrm{d}\dot{\xi} \tag{7.15}$$

因此我们定义 Routhian 函数为

$$R(q,p,\xi,\dot{\xi})=p\dot{q}-L \tag{7.16}$$

则

$$-\mathrm{d}R=\dot{p}\mathrm{d}q-\dot{q}\mathrm{d}p+\frac{\partial L}{\partial \xi}\xi+\frac{\partial L}{\partial \dot{\xi}}\mathrm{d}\dot{\xi} \tag{7.17}$$

由此可得

$$\dot{q}=\frac{\partial R}{\partial p},\quad \dot{p}=-\frac{\partial R}{\partial q} \tag{7.18}$$

$$\frac{\partial L}{\partial \xi}=-\frac{\partial R}{\partial \xi},\quad \frac{\partial L}{\partial \dot{\xi}}=-\frac{\partial R}{\partial \dot{\xi}} \tag{7.19}$$

这样我们就仅仅将部分坐标变换到了相空间中.

7.3.2 Routhian 函数处理循环坐标

循环坐标是指不出现在拉氏量中的广义坐标，用 Routhian 函数处理循环坐标要比拉格朗日力学更加方便.

Lagrange 函数和 Routhian 函数都没有循环坐标，但是在 Lagrange 函数中会出现循环坐标的速度，循环坐标对应的广义速度在 Rouhian 函数中是常数. 因此 Routhian 函数要比 Lagrange 函数更加方便.

用 Routhian 函数表达的运动方程为：

非循环坐标部分:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial R}{\partial \dot{q}^a}-\frac{\partial R}{\partial q^a}=0,\quad a=1,2,\cdots,m \tag{7.20}$$

循环坐标部分:

$$\dot{q}^a=\frac{\partial R}{\partial \alpha_{a-m}},\quad \dot{p}_a=-\frac{\partial R}{\partial q^a}\equiv 0,\quad a=m+1,\cdots,s \tag{7.21}$$

其中非循环坐标部分满足标准的拉格朗日方程. 而循环坐标部分可以立即积出

$$q^a=\int \mathrm{d}t\frac{\partial R}{\partial \alpha_{a-m}},\quad p_a=\text{ 常数},\quad a=m+1,\cdots,s. \tag{7.22}$$

所以问题归结为求解这 m 个非循环坐标的拉格朗日方程.

第八章　正则变换

哈密顿力学就是相空间的几何学.

《经典力学的数学方法》

物理学中有许多空间，比如坐标空间、相空间. 我们很熟悉坐标空间，在坐标空间我们会附加一些数学结构，比如度规、内积、长度等等，它们可以方便我们在坐标空间中更好地表达我们的理论. 而哈密顿力学是在像空间上的，因此我们在相空间上也要附加一些数学结构.

我们还会经常对于空间进行变换. 对于一般的变换，是不保上述的数学结构的；不过既然那些结构很有用，我们就希望那些结构在变换下是不变的. 这种保结构的变换就是这章要研究的重点.

为此，我将从最熟悉的坐标空间出发，介绍上面的数学结构、坐标变换. 再从坐标空间推广至相空间，介绍上面的辛形式、泊松括号以及正则变换. 最后介绍一下如何用生成函数得到正则变换.

8.1　坐标空间的结构

8.1.1　度规

目前我无法给出度规的严格定义. 可以暂时将度规理解为一个对称矩阵 g_{ab}.

带有度规的空间被称为度量空间.

8.1.2　内积

定义了度规我们就能够定义内积.

定义 8.1.1: 内积

$$\boldsymbol{A}\cdot\boldsymbol{B} := g_{ab}A^aB^b$$

举例

1. 欧氏空间的内积为 $\boldsymbol{A}\cdot\boldsymbol{B} = \delta_{ab}A^aB^b = A^1B^1 + A^2B^2$.
2. 闵氏空间的内积为 $\boldsymbol{V}\cdot\boldsymbol{W} = \eta_{\mu\nu}V^\mu W^\nu = -V^0W^0 + V^1W^1 + V^2W^2 + V^3W^3$

带有内积的空间被称为内积空间.

此外，定义了内积我们就能定义长度.

8.1.3 内积括号

定义度量空间中两个标量函数

$$f = f(t, q^a), \tag{8.1}$$

$$g = g(t, q^a), \tag{8.2}$$

的内积括号为，f, g 的梯度

$$A_a = \frac{\partial f}{\partial q^a}, B_b = \frac{\partial g}{\partial q^b} \tag{8.3}$$

的内积

$$\langle f, g \rangle := g^{ab} \frac{\partial f}{\partial q^a} \frac{\partial g}{\partial q^b} \tag{8.4}$$

其中 q^a 表示一组基向量. [1]

由于内积括号使用的是标量函数、而标量是不随坐标系而变换的.. 使用内积括号可以让我们更方便地看出内积是在什么坐标下的，这让我们更方便地进行坐标变换. 另外，内积括号和泊松括号也有很强的类似性. 这些都是我们引入内积括号的原因.

8.1.4 坐标空间的变换

所谓坐标空间的变换即坐标空间的映射 $(M_1, d_1) \to (M_2, d_2)$，$M$ 表示向量空间，d 表示度规. 对于最一般的变换，我们没有任何的要求，向量空间和度规都是可以改变的.

一般来说，我们关注两种变换.

1. 度规不变，但是内积会变.
2. 内积不变，但是度规会变.

一般说的线性变换，只变了空间（虽然还是同构的），度规没变. 对于没有改变度规的变换，我们可以讨论其是否是保内积，保长度的.（不然都不知道度规是啥，咋讨论内积.）对于没有改变内积（长度）的变换，我们可以讨论其是否是保度规的.

8.1.5 正交变换

定义 8.1.2: 正交变换

对于度规不变、改变内积的变换，正交变换的定义是：保度规的变换；
对于内积不变、改变度规的变换，正交变换的定义是：保内积的变换.

第一种正交变换的几何意义是旋转.

第二种正交变换就是我们说的不改变空间结构（度规）的变换.

8.2 相空间的几何

这一节很多东西都是我杜撰的，请谨慎阅读.

[1] 内积括号是高显老师杜撰的.

8.2.1 一般辛形式

一般辛形式可以理解成是一个反对称的矩阵，与度规相对应．[2]

辛形式是一种特殊的一般辛形式，可以与任何一种特定的度规（比如欧氏度规）相对应．辛形式 ω 的矩阵表示为

$$\omega^{\alpha\beta} := \begin{pmatrix} \mathbf{0}_{s\times s} & \mathbf{1}_{s\times s} \\ -\mathbf{1}_{s\times s} & \mathbf{0}_{s\times s} \end{pmatrix} \tag{8.5}$$

8.2.2 一般泊松括号

一般泊松括号是将向量括号中的度规换成一般辛形式．

8.2.3 泊松括号

与内积括号完全类似，只是将度规换成了辛形式，我们就得到了泊松括号．

定义 8.2.1: 泊松括号

定义辛向量空间中两个标量函数

$$f = f(t, \xi^\alpha) = f(t, q^a, p_b), \tag{8.6}$$

$$g = g(t, \xi^\alpha) = g(t, q^a, p_b), \tag{8.7}$$

的梯度的泊松括号为，

$$[f, g] := \omega^{ab} \frac{\partial f}{\partial q^a} \frac{\partial g}{\partial q^b} \tag{8.8}$$

8.2.4 相空间中的变换

一般来说，我们关注两种变换．

1. 一般辛形式不变，但是一般泊松括号的数值会变．
2. 一般泊松括号的数值不变，但是一般辛形式会变．

对于辛向量空间，上述的两种变换变成

1. 辛形式不变，但是泊松括号的数值会变．
2. 泊松括号的数值不变，但是辛形式会变．

在相空间中，很重要的一点是：相空间的变换满足 $\frac{\mathrm{d}f}{\mathrm{d}t}$ 是不变的，而它又是一个泊松括号，因此泊松括号的数值不变．所以我们讨论的变换都是泊松括号的数值不变，辛形式会改变．而在这中间有一种特殊的变换，不会改变辛形式，也就是正则变换．

[2]一般辛形式这个词是我杜撰的，目的只是为了和度规相对应．

8.2.5 正则变换

定义 8.2.2: 正则变换

相空间中的正则变换的定义是，保辛形式的变换.

正则变换的意义就是保辛向量空间的结构.

坐标空间与相空间的对比与总结见表8.1

表 8.1: 坐标空间与相空间的对比与总结

坐标空间	相空间
度规	一般辛形式
欧氏度规	辛形式
内积	一般泊松括号
欧氏空间中的内积	泊松括号
正交变换	正则变换

现在我们明确了正则变换是保辛形式的变换. 而所谓保泊松括号也指的是保泊松括号的辛形式保辛形式. 因此正则变换等价于保泊松括号.

8.3 泊松括号

8.3.1 泊松括号的性质

根据定义可以得到泊松括号的一些的性质

1. $[f,g]=-[g,f]$
2. $[f,c]=0, c\in\mathbb{R}$
3. $[f_1+f_2,g]=[f_1,g]+[f_2,g]$
4. $[f_1f_2,g]=f_1[f_2,g]+f_2[f_1,g]$
5. $\frac{\partial}{\partial t}[f,g]=\left[\frac{\partial f}{\partial t},g\right]+\left[f,\frac{\partial g}{\partial t}\right]$
6. $[f,q_k]=\frac{\partial f}{\partial p_k},\quad [f,p_k]=-\frac{\partial f}{\partial q_k}$

8.3.2 基本泊松括号

根据泊松括号的定义，相空间中的 ξ^α,ξ^β 有

$$[\xi^\alpha,\xi^\beta]=\omega^{\alpha\beta} \tag{8.9}$$

即我们计算一个相空间的基本一般泊松括号，就得到了这个相空间的一般形式. 如果得到了辛形式，就说明我们在辛向量空间中，之前的一般泊松括号也就是泊松括号. 而我们知道，泊松括号的数值是不

变的，这意味着，我们可以在任何一个空间当中计算这个基本泊松括号，得到的都是我们想求的辛形式！即对于一个变换 $\xi \to \Xi$，我们想计算 Ξ 空间中的辛形式是什么，只需要计算 Ξ 的基本泊松括号，即计算

$$[\Xi^\alpha, \Xi^\beta]_\Xi \tag{8.10}$$

而我们又知道

$$[\Xi^\alpha, \Xi^\beta]_\xi = [\Xi^\alpha, \Xi^\beta]_\Xi \tag{8.11}$$

因此

$$[\Xi^\alpha, \Xi^\beta]_\xi = \omega_\Xi^{\alpha\beta} \tag{8.12}$$

8.3.3 用泊松括号表示正则方程

$$\dot{q} = [q, H], \dot{p} = [p, H] \tag{8.13}$$

8.3.4 泊松括号与运动积分

用泊松括号计算运动积分

设 $f(p, q, t)$ 是坐标、动量和时间的某个函数．它对时间的全导数为

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \sum_k \left(\frac{\partial f}{\partial q_k}\dot{q}_k + \frac{\partial f}{\partial p_k}\dot{p}_k \right) \tag{8.14}$$

带入正则方程(7.2.1)可得

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \sum_k \left(\frac{\partial H}{\partial p_k}\frac{\partial f}{\partial q_k} - \frac{\partial H}{\partial q_k}\frac{\partial f}{\partial p_k} \right) \tag{8.15}$$

真是太巧了，我们有

$$[H, f] = \sum_k \left(\frac{\partial H}{\partial p_k}\frac{\partial f}{\partial q_k} - \frac{\partial H}{\partial q_k}\frac{\partial f}{\partial p_k} \right) \tag{8.16}$$

那么 f 的时间全导数可写成

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + [f, H] \tag{8.17}$$

我们知道，如果动力学变量的某个函数当系统运动时保持不变，则称之为运动积分．因此，f 是运动积分的条件可写成

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + [f, H] = 0 \tag{8.18}$$

若运动积分不显含时间，则

$$[f, H] = 0 \tag{8.19}$$

因此，我们要判断 f 是否为运动积分，只需要计算一个泊松括号就好了，这非常方便！

泊松定理

泊松括号的一个重要性质是泊松定理

定理 8.3.1: 泊松定理

如果 f 和 g 是系统的两个运动积分，则两者的泊松括号 $[f,g]$ 也是系统的运动积分，即

$$\frac{\mathrm{d}f}{\mathrm{d}t}=0,\quad \frac{\mathrm{d}g}{\mathrm{d}t}=0,\quad \Rightarrow\quad \frac{\mathrm{d}[f,g]}{\mathrm{d}t}=0$$

证明. 证明见高显老师的讲义. □

8.3.5 泊松括号与正则方程

在辛相空间中，我们有正则方程

$$\dot{q}=[q,H] \tag{8.20}$$

$$\dot{p}=[p,H] \tag{8.21}$$

对于一个坐标变换 $q,p\to Q,P$，上述方程变成

$$\dot{Q}=[Q,K] \tag{8.22}$$

$$\dot{P}=[P,K] \tag{8.23}$$

当且仅当方程中的一般泊松括号为泊松括号时，这个方程才具有辛形式，才是正则方程.
因此我们证明了：保泊松括号等价于保正则方程.
终于，我们说明了

定理 8.3.2: 正则变换的等价表述

正则变换 = 保辛形式 = 保泊松括号 = 保基本泊松括号 = 保正则方程

这里的保泊松括号指的是保形式，保基本泊松括号指的是保数值.

8.3.6 点变换是正则变换

点变换就是坐标空间中的位形变换. 显然，点变换也可以作为相空间中的变换. 而这个变换就是正则变换.

为了验证这个事实，我们要首先找出这个变换的形式：只知道坐标的变换还不够，我们还要知道动量的变换. 然后再验证点变换后的基本泊松括号可以得到这个结论.

8.4 生成函数

现在我们来介绍得到正则变换的方法：生成函数.
考虑正则变换 $\{q,p\}\to\{Q,P\}$，根据拉氏量可以差一个时间的全导数项，我们有

$$p\dot{q}-H=P\dot{Q}-K+\frac{\mathrm{d}F}{\mathrm{d}t} \tag{8.24}$$

其中 F 的形式没有要求．改写成微分形式，有

$$\mathrm{d}F = p\mathrm{d}q - P\mathrm{d}Q + (K - H)\mathrm{d}t \tag{8.25}$$

方程(8.25)是生成函数方法的基本出发点．它说明了 F 的自变量和偏导数关系

$$F = F(t, q, Q) \tag{8.26}$$

$$\frac{\partial F}{\partial t} = K(t, Q, P) - H(t, q, p) \tag{8.27}$$

$$\frac{\partial F}{\partial Q} = -P \tag{8.28}$$

$$\frac{\partial F}{\partial q} = p \tag{8.29}$$

这几个方程完全描述了一个正则变换．**也就是说一个生成函数完全描述了一个正则变换．**

8.5 演化是单参数正则变化

演化当然是正则变换，系统随时间的演化又不会改变空间的辛形式！

第九章　哈密顿-雅可比理论

9.1　完美的正则变换

我们做正则变换的目的是为了简化问题，很自然地，我们会想，什么样的正则变换是能使问题最简化的——最完美的正则变换．显然，最完美的正则变换是：将坐标和动量映射成常数！

$$\{q(t), p(t)\} \Rightarrow \{Q, P\} \tag{9.1}$$

其中 $\{Q, P\}$ 为常数．其中一种很有趣且特别的正则变换就是演化的逆映射，演化是从起始点经过正则变换到了 t 时刻的坐标，它的逆映射就是将坐标和动量全都映射成常数．

而这样的正则变换是要满足一些条件的．根据新的正则方程

$$\dot{Q} = \frac{\partial K}{\partial P} = 0, \quad \dot{P} = -\frac{\partial K}{\partial Q} = 0, \tag{9.2}$$

我们有

$$\frac{\partial K}{\partial P} = \frac{\partial K}{\partial Q} = 0 \Rightarrow K = K(t) \tag{9.3}$$

即新的哈氏量是与新的相空间坐标无关的量．显然，最最简单（且完美）的选择是

$$K = 0. \tag{9.4}$$

因此这个最完美的正则变换 $\{q, p\} \to \{Q, P\}$ 要满足的条件就是

$$\begin{cases} \dot{Q} = 0 \\ \dot{P} = 0 \\ K = 0 \end{cases} \tag{9.5}$$

而 $K = 0$ 又可以推导出前两个条件，因此最完美的正则变换所满足的条件只是

$$K = 0. \tag{9.6}$$

9.2　哈密顿雅可比方程

下面我们来求这个正则变换的具体形式，我们用生成函数来求．

我们选用第二类生成函数 $F_2(t, q^a; \alpha_b)$，并将其记为 S，即

$$F_2(t, q^a; \alpha_b) := S(t, q^a; \alpha_b) \tag{9.7}$$

S 也称为哈密顿主函数．理论上来说，求出来了哈密顿主函数我们就能求出这个最完美的正则变换了．

根据第二类生成函数的正则变换关系，我们有

$$p_a = \frac{\partial S\left(t, q^b; \alpha_c\right)}{\partial q^a} \tag{9.8}$$

$$Q^a = \frac{\partial S\left(t, q^b; \alpha_c\right)}{\partial P_a} \tag{9.9}$$

$$K\left(t, Q^a, P_b\right) = H\left(t, q^a, p_b\right) + \frac{\partial S\left(t, q^a; \alpha_b\right)}{\partial t} \equiv 0 \tag{9.10}$$

根据哈密顿主函数应满足的条件(9.6)，我们有

$$H\left(t, q^a, p_b\right) + \frac{\partial S\left(t, q^a; \alpha_b\right)}{\partial t} \equiv 0 \tag{9.11}$$

我们就要通过这个条件来求解哈密顿主函数，可是哈密顿主函数是 q^a 的函数而非 p_b 的函数，因此我们必须将上式中的 p_b 替换掉，利用(9.8)，我们得到

定理 9.2.1: 哈密顿雅可比方程

$$H\left(t, q^a, \frac{\partial S\left(t, q^c; \alpha_d\right)}{\partial q^b}\right) + \frac{\partial S\left(t, q^a; \alpha_b\right)}{\partial t} = 0$$

这就是哈密顿-雅可比方程，也就是最完美的正则变换的条件.

更神奇的是：哈密顿-雅可比方程就是薛定谔方程的经典极限！这太神奇了.

9.3 求解哈密顿雅可比方程——分离变量

我们来通过变量分离的方式求解哈密顿雅可比方程，只要老的哈密顿函数不显含时间，那么我们总能哈密顿主函数变量分离，下面我们来说明这件事.

作试探解

$$S(t, q^a) = W(q^a) + V(t), \tag{9.12}$$

带入(9.2.1)，得到

$$H\left(q^a, \frac{\partial W(q^c)}{\partial q^b}\right) = -\frac{\partial V(t)}{\partial t} \tag{9.13}$$

方程左边仅为 q^a 的函数，右边仅为 t 的函数，**这就说明我们变量分离成功了.** 要等式成立，唯一地可能是两边取共同的常数，根据哈氏量的意义，这个常数自然为 E. 即

$$H\left(q^a, \frac{\partial W(q^c)}{\partial q^b}\right) = -\frac{\partial V(t)}{\partial t} = E \tag{9.14}$$

于是我们求出

$$V(t) = -Et \tag{9.15}$$

因此哈密顿主函数可以写成

$$S(t, q^a) = W(q^a) - Et \tag{9.16}$$

这里的 $W = W(q^a)$ 也被称为哈密顿特征函数，这样我们把哈密顿主函数的方程：哈密顿雅可比方程，化简成了关于哈密顿特征函数的方程

$$H\left(q^a, \frac{\partial W(q^c)}{\partial q^b}\right) = E \tag{9.17}$$

求解出哈密顿特征函数也就求解出了哈密顿主函数.

9.4 经典作用量

作用量是路径的泛函，每一条路径都有一个作用量．但是根据最小作用量原理，实际的运动只是沿着作用量取最小值的路径．因此确定了两个端点，路径也就确定了．这个确定的路径的作用量称为经典作用量，它是端点的函数．

神奇的是，哈密顿主函数就是经典作用量．我无法理解，这太神奇了．

第十章　转动理论

现在来更加系统地探究一下关于转动的理论.

我将转动定义为保空间数学结构的坐标变换. 所谓数学结构指的就是度规、辛形式，这一点在第八章中有比较详细的讨论.

10.1　转动矩阵

首先我们要找出一般转动的数学形式.

10.1.1　转动是线性变换

考虑一个度量空间，有度规 g_{ab}. 现在做转动变换 $\{q^a\} \to \{Q^a\}$. 根据转动的定义，转动是保度规的，我们有

$$g_{cd}\frac{\partial Q^c}{\partial q^a}\frac{\partial Q^b}{\partial q^d} = g_{ab} \tag{10.1}$$

将(10.1)两边同时对 q^c 求偏导，有

$$g_{cd}\frac{\partial^2 Q^c}{\partial q^e \partial q^a}\frac{\partial Q^d}{\partial q^b} + g_{cd}\frac{\partial Q^c}{\partial q^a}\frac{\partial^2 Q^d}{\partial q^e \partial q^b} \equiv 0 \tag{10.2}$$

将(10.2)中 e, a 指标对换，得到

$$g_{cd}\frac{\partial^2 Q^c}{\partial q^a \partial q^e}\frac{\partial Q^d}{\partial q^b} + g_{cd}\frac{\partial Q^c}{\partial q^e}\frac{\partial^2 Q^d}{\partial q^a \partial q^b} \equiv 0 \tag{10.3}$$

将(10.2)中 e, b 指标对换，得到

$$g_{cd}\frac{\partial^2 Q^c}{\partial q^b \partial q^a}\frac{\partial Q^d}{\partial q^e} + g_{cd}\frac{\partial Q^c}{\partial q^a}\frac{\partial^2 Q^d}{\partial q^b \partial q^e} \equiv 0 \tag{10.4}$$

将前两式相加并减去第三式，注意到二阶偏导可交换次序，我们得到

$$g_{cd}\frac{\partial^2 Q^c}{\partial q^a \partial q^e}\frac{\partial Q^d}{\partial q^b} = 0 \tag{10.5}$$

又因为坐标变换是可逆的，$\frac{\partial Q^b}{\partial q^d} \neq 0$，因此

$$\frac{\partial^2 Q^c}{\partial q^a \partial q^e} = 0 \tag{10.6}$$

即 $Q^a(q^b)$ 是 q^b 的线性函数！即转动是线性变换！欧氏空间中的转动和闵可夫斯基空间中的洛伦兹变换都是线性变换.

线性变换意味着我们可以将转动写成矩阵的形式了！

10.1.2 正交矩阵

我们下面接着来讨论一般转动的数学形式，我们已经知道转动可以表示成矩阵了．根据转动的定义，转动矩阵还需要满足

$$\boldsymbol{R}^T\boldsymbol{R}=\boldsymbol{R}\boldsymbol{R}^T=\boldsymbol{1} \tag{10.7}$$

满足这个关系的矩阵被称为正交矩阵．容易得到，正交矩阵有

$$\det\boldsymbol{R}=\pm 1 \tag{10.8}$$

10.1.3 转动群

注意到，任意两个正交矩阵的乘积仍然是正交矩阵．从代数的角度来说，这意味着 D 阶正交矩阵的集合在矩阵乘法的意义下构成群，这个群被称为正交群或转动群，也被称为李群．其中 $\det\boldsymbol{R}=+1$ 的转动被称为正常转动，其对应的群被称为特殊正交群，记作 $\mathrm{SO}(D)$．

可以证明，一个 D 阶正交矩阵有 $\frac{1}{2}D(D-1)$ 个独立分量．这意味着 D 维欧氏空间中的转动有 $\frac{1}{2}D(D-1)$ 个自由度．等价地，

$$\dim\mathrm{SO}(D)=\frac{1}{2}D(D-1) \tag{10.9}$$

10.2 无穷小转动

“线性”在物理中始终是一个重要且特别的存在．线性化、线性空间，这些都与线性密不可分，我目前无法完全参透它的意义，但是我们至少是希望进行线性化的．转动矩阵在转动群中而非线性空间中，转动矩阵并不是线性的．因此我们将转动矩阵通过指数对应关系线性化，得到在线性空间中的无穷小转动矩阵．

10.2.1 指数映射

设某个无穷小转动矩阵的生成元为 $\boldsymbol{J}$，设参数 ϕ 是有限的，则一个一般的无穷小转动矩阵为

$$\lim_{N\to\infty}\frac{\phi}{N}\boldsymbol{J} \tag{10.10}$$

将无穷小转动做 N 次也就成了有限转动，因此有限转动 $\boldsymbol{R}$ 可表示为

$$\boldsymbol{R}=\lim_{N\to\infty}\left(1+\frac{\phi}{N}\boldsymbol{J}\right)^N\equiv e^{\phi J} \tag{10.11}$$

而根据指数函数的定义，我们有

$$\boldsymbol{R}=\sum_{n=0}^{\infty}\frac{1}{n!}\phi^n\boldsymbol{J}^n \tag{10.12}$$

由于 $\boldsymbol{J}$ 是无穷小转动的生成元，因此 $\boldsymbol{J}^n$ 一般都有规律，这样就能求出 $\boldsymbol{R}$ 的一般形式了．

10.2.2 无穷小转动矩阵的性质

无穷小转动矩阵具有很多好的性质，让我们有充分的理由去使用它

1. 无穷小转动矩阵是反对称的．

2. 无穷小转动矩阵在矩阵加法的意义下构成线性空间.

无穷小转动矩阵都是反对称矩阵.

无穷小转动矩阵在矩阵加法的意义下构成线性空间. 而有限转动矩阵在矩阵乘法的意义下构成群. 矩阵加法要比矩阵乘法更有优越性. 线性空间要比群更有优越性，因为在线性空间中我们可以选取基了. 可以选取基是非常有用的，这意味着我们只需要研究有限个基就可以了. 另外这些基之间还有一种代数结构可以互相转换，这种结构就称为李括号. 由于基都是反对称的，因此李括号（对易子、泊松括号）就是反对称的，这都在量子力学中有很重要的应用. 其实李代数和李括号是等价的，因为所有的反对称矩阵都有李括号，李括号也可以推出反对称矩阵，因此只要有李括号就一定是李代数.

10.2.3 李群与李代数

总结：

1. 正交矩阵在矩阵乘法的意义下构成了一个群，被称为李群.
2. 无穷小转动矩阵被称为生成元. 每个无穷小转动矩阵都可以生成有限转动矩阵.
3. 无穷小转动矩阵在矩阵加法的意义下构成了一个线性空间，这个空间被称为李代数.
4. 这个线性空间有一组基，基之间可以通过李括号互相变换，李代数要有李括号.
5. 生成元在指数映射下是有限转动矩阵. 这意味着李代数在指数映射下是李群.

10.3 角速度

10.3.1 角速度矩阵

对于一个正在做旋转运动的点 $\boldsymbol{A}(t)$，它在 $\boldsymbol{A}(t+\mathrm{d}t)$ 时刻一定可以表示成

$$\boldsymbol{A}(t+\mathrm{d}t)=\boldsymbol{A}(t)+\boldsymbol{\Omega}(t)\boldsymbol{A}(t)\mathrm{d}t \tag{10.13}$$

其中 $\boldsymbol{\Omega}(t)$ 是一个反对称的转动矩阵，$\boldsymbol{\Omega}(t)\mathrm{d}t$ 自然是一个无穷小的转动矩阵. 这样就有

$$\frac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t}=\boldsymbol{\Omega}(t)\boldsymbol{A}(t) \tag{10.14}$$

我们关系其中的转动部分，因此我们定义角速度矩阵：

定义 10.3.1: 角速度矩阵

对于正在旋转的点 $\boldsymbol{A}(t)$，一定有

$$\frac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t}=\boldsymbol{\Omega}(t)\boldsymbol{A}(t)$$

其中 $\boldsymbol{\Omega}(t)$ 是一个反对称矩阵. 我们定义 $\boldsymbol{\Omega}(t)$ 为角速度矩阵.

10.3.2 角速度矩阵与转动矩阵的关系

由于 $\boldsymbol{A}(t)$ 一直在转动，因此 $\boldsymbol{A}(t)$ 一定可以用一个转动矩阵 $\boldsymbol{R}(t)$ 表示

$$\boldsymbol{A}(t)=\boldsymbol{R}(t)\boldsymbol{A}_0 \tag{10.15}$$

对两边同时求导得到

$$\begin{aligned}\frac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t}&=\frac{\mathrm{d}\boldsymbol{R}(t)}{\mathrm{d}t}\boldsymbol{A}_0\\&=\frac{\mathrm{d}\boldsymbol{R}(t)}{\mathrm{d}t}\boldsymbol{R}^{-1}(t)\boldsymbol{R}(t)\boldsymbol{A}_0\\&=\frac{\mathrm{d}\boldsymbol{R}(t)}{\mathrm{d}t}\boldsymbol{R}^T(t)\boldsymbol{A}(t)\end{aligned} \tag{10.16}$$

根据角速度矩阵的定义，我们有

$$\boldsymbol{\Omega}(t)=\frac{\mathrm{d}\boldsymbol{R}(t)}{\mathrm{d}t}\boldsymbol{R}^T(t) \tag{10.17}$$

第十一章　刚体

11.1　刚体

刚体可以看成是有无穷多个粒子所构成的一个粒子系. 其特点是任意两个粒子之间的距离都不变. 刚体这一概念从本质上来说就是非相对论性的.

我们下面思考一下如何像描述平动一样描述转动. 回顾对于平动的描述，我们大致可以这样来总结：

1. 首先我们定义速度.
2. 然后我们定义质量是动能中速度二次项的系数.
3. 下面我们定义动量是速度与质量的乘积.
4. 接着我们定义力是动量的导数.

同样地，对于转动，我们也可以这样描述：

1. 首先我们定义角速度.
2. 然后我们定义转动惯量是动能中角速度二次项的系数.
3. 下面我们定义角动量是角速度与转动惯量的乘积.
4. 接着我们定义力矩是角动量的导数.

这大概就是整个刚体动力学的框架.

11.2　转动惯量

11.2.1　惯量张量

受到动能与速度的关系

$$T = \frac{1}{2}mv^2 \tag{11.1}$$

的启发，我们希望将刚体转动的动能写成角速度二次型的形式. 经过计算发现角速度前的系数无法写成像质量一样的常数，只能写成张量矩阵的形式，即

$$T = \frac{1}{2}I_{ij}\omega^i\omega^j \tag{11.2}$$

其中的 I_{ij} 称为转动惯量．下面我们来求转动惯量具体的形式

$$\begin{aligned}
T &= \sum \frac{1}{2} m v^2 \\
&= \frac{1}{2} \sum m (\boldsymbol{\omega} \times \boldsymbol{r})^2 \\
&= \frac{1}{2} \sum m (\omega^2 r^2 - (\boldsymbol{\omega} \cdot \boldsymbol{r})^2) \\
&= \frac{1}{2} \sum m (\omega_i^2 r_l^2 - \omega_i r_i \omega_j r_j) \\
&= \frac{1}{2} \sum m (\omega_i \omega_j \delta_{ij} r_l^2 - \omega_i \omega_j r_i r_j) \\
&= \frac{1}{2} \omega_i \omega_j \sum m (r^2 \delta_{ij} - r_i r_j)
\end{aligned} \tag{11.3}$$

对比(11.2)，可得

$$I_{ij} = \sum_{\alpha=1}^{N} m_\alpha ((\boldsymbol{r}_\alpha)^2 \delta_{ij} - (\boldsymbol{r}_\alpha)_i (\boldsymbol{r}_\alpha)_j) \tag{11.4}$$

11.2.2 投影张量

我们来分析一下这个惯量张量的意义．将(11.4)中的 $\boldsymbol{r}_\alpha^2$ 提出来，得到

$$\Pi_{ij} = \delta_{ij} - \widehat{r}_i \widehat{r}_j \tag{11.5}$$

第一项 δ_{ij} 是单位张量，它不改变乘上的向量．

第二项是将将向量 $\boldsymbol{A}$ 投影到 $\boldsymbol{r}$ 的方向（平行方向）上，

$$\widehat{r}_i \widehat{r}_j A^j = \widehat{r}_i (\widehat{\boldsymbol{r}} \cdot \boldsymbol{A}) = \boldsymbol{A}^{\parallel} \tag{11.6}$$

而原向量减去平行方向的向量就只剩下了垂直方向的向量．因此张量 Π_{ij} 的意义是将一个向量投影到 $\boldsymbol{r}$ 的垂直分量上．

11.2.3 平行轴定理

转动惯量中的 $\boldsymbol{r}$ 显然和原点有关，一般的情况下，我们选择质心作为原点，此时的转动惯量为 I_C．当我们选取 P 作为原点，P 相对于质心的位矢为 $\boldsymbol{P}$，则 I_P, I_C 之间的关系为

$$(I_P)_{ij} = (I_C)_{ij} + M(P^2 \delta_{ij} - P_i P_j) \tag{11.7}$$

即

$$I_P = I_C + \text{总质量 M 在 P 处的转动惯量} \tag{11.8}$$

这个关系被称为平行轴定理．

证明.

$$\begin{aligned}
(I_P)_{ab} &= \sum_\alpha m_\alpha \left[(\boldsymbol{r}_\alpha - \boldsymbol{P}) - (\boldsymbol{r}_\alpha - \boldsymbol{P})_i (\boldsymbol{r}_\alpha - \boldsymbol{P})_j \right] \\
&= \sum_\alpha m_\alpha \left\{ r_\alpha^2 \delta_{ij} - (r_\alpha)_i (r_\alpha)_j + \left[-2\boldsymbol{r}_\alpha \cdot \boldsymbol{P} \delta_{ij} + (r_\alpha)_i P_j + (r_\alpha)_j P_i \right] + \left(P^2 \delta_{ij} - P_i P_j \right) \right\}
\end{aligned} \tag{11.9}$$

上式中括号正比于 $\boldsymbol{r}_\alpha$，根据质心的定义，其求和为零．因此

$$(I_P)_{ij} = (I_C)_{ij} + M(P^2\delta_{ij} - P_iP_j) \tag{11.10}$$

□

11.2.4 转动惯量的计算

将(11.4)写成矩阵形式，即

$$I_{ij} = \sum_{\alpha=1}^{N} m_\alpha \begin{pmatrix} y_\alpha^2 + z_\alpha^2 & -x_\alpha y_\alpha & -x_\alpha z_\alpha \\ -y_\alpha x_\alpha & x_\alpha^2 + z_\alpha^2 & -y_\alpha z_\alpha \\ -z_\alpha x_\alpha & -z_\alpha y_\alpha & x_\alpha^2 + y_\alpha^2 \end{pmatrix} \tag{11.11}$$

显然，转动惯量是对称的．实对称矩阵都可以对角化．即存在正交矩阵 $\boldsymbol{R}$，是的

$$\boldsymbol{R}\boldsymbol{I}\boldsymbol{R}^T = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \tag{11.12}$$

换句话说，我们选择合适的基矢量

$$\boldsymbol{e}_a \to \boldsymbol{R}_a{}^b \boldsymbol{e}_b \tag{11.13}$$

使得转动惯量是对角化的．这样的基矢量被称为惯量主轴，其中本征值 I_1, I_2, I_3 被称为主轴转动惯量．我们不妨将惯量主轴称为 x, y, z 轴，这样转动惯量为

$$\begin{pmatrix} I_x & 0 & 0 \\ 0 & I_y & 0 \\ 0 & 0 & I_z \end{pmatrix} \tag{11.14}$$

根据转动惯量的定义，即其与动能的关系，我们可以得到，对 x, y, z 轴都有

$$T_x = I_x \omega_x^2, \tag{11.15}$$

注意，对于每一个粒子，它们的 ω_x 都是相同的．这样我们有

$$\begin{aligned} I_x\omega_x^2 =& \frac{1}{2} m_\alpha (v_\alpha)_x^2 \\ I_x\omega_x^2 =& \frac{1}{2} m_\alpha (\omega_x x)^2 \\ I_x =& \frac{1}{2} m_\alpha x^2 \end{aligned} \tag{11.16}$$

上式用积分形式写成

$$I_x = \int_a^b x^2 \rho \mathrm{d}\tau \tag{11.17}$$

因此，在我们找到了合适的坐标系后，可以用这个公式方便地计算转动惯量分量．

11.3 角动量

定义 11.3.1: 角动量

角动量是转动惯量与角速度的乘积，即

$$J_i = I_{ij}\omega^j$$

通过定义可以看出，角动量的方向与角速度的方向一般不同．经过计算可以得到

$$\boldsymbol{J} = \sum \boldsymbol{r} \times \boldsymbol{p} \tag{11.18}$$