

理论力学讲义 (2023 秋季)

张芃

中国人民大学物理系

目录

第一章 Lagrange 力学	2
1.1 数学复习：复合函数求偏导的锁链法则	2
1.1.1 复合函数	2
1.1.2 锁链法则	3
1.1.3 偏导数计算应注意的问题	3
1.2 无约束保守系统的 Lagrange 方程	5
1.2.1 无约束保守系统的 Lagrange 方程：直角坐标	5
1.2.2 广义坐标与广义速度	9
1.2.3 无约束保守系统的 Lagrange 方程：广义坐标	11
1.3 完整约束中的 Lagrange 方程	14
1.3.1 完整约束	14
1.3.2 约束力	16
1.3.3 完整约束系统的 Lagrange 方程	23
1.3.4 保守系统的 Lagrange 方程	26
1.4 Lagrange 方程的应用	27
1.4.1 阿特伍德机	27
1.4.2 双摆	28
1.5 广义动量与广义能量	29
1.5.1 广义动量	29
1.5.2 广义能量	30
1.6 与速度有关的广义势	32
1.6.1 广义势	32
1.6.2 Lorentz 力	33
1.6.3 Lorentz 力对应的广义势	34
第二章 微振动	37

2.0.1	二阶线性齐次常微分方程	37
2.0.2	二阶线性非齐次常微分方程	39
2.1	一维简谐振子	40
2.1.1	自由振子	40
2.1.2	受迫振子	41
2.1.3	阻尼振子	44
2.1.4	受迫阻尼振子	45
2.1.5	共振现象总结	47
2.1.6	参数振荡	48
2.2	两个耦合的同频率振子	48
2.3	通过合同变换对角化实对称矩阵	52
2.3.1	单个实对称矩阵的对角化	52
2.3.2	两个正定实对称矩阵的联合对角化	54
2.4	稳定平衡位置附近小振动的一般理论	56
2.4.1	势能的极小点：稳定平衡点	56
2.4.2	稳定平衡位置附近的小振动：直观广义坐标和近似拉氏量。	56
2.4.3	简正坐标	57
2.4.4	简正圆频率	60
2.4.5	简正模式	60
2.4.6	小结	61
2.4.7	求简正坐标与简正频率的其他方法	62
2.5	耦合弹簧振子系统举例	63
2.5.1	两个耦合弹簧振子	63
2.5.2	振子串列	67
第三章	Hamilton 力学	70
3.1	Hamilton 方程	70
3.2	Hamilton 方程的应用	75
3.2.1	相关的问题	75
3.2.2	直角坐标下的无约束保守系统	76
3.3	泊松括号和刘维尔定理	82
3.3.1	泊松括号	82
3.3.2	泊松括号与 Hamilton 方程	84
3.3.3	泊松括号与守恒量	84
3.3.4	相空间体积不变定理和 Liouville 定理	85

第四章 变分法和 Hamilton 原理	88
4.1 集合、映射、函数与泛函	88
4.2 泛函极值与欧拉-拉格朗日方程	90
4.2.1 泛函极值	90
4.2.2 端点固定的极值问题	90
4.2.3 EL 方程的评述	92
4.2.4 EL 方程的应用	93
4.2.5 EL 方程推广到多个函数的情况	94
4.3 Hamilton 原理	95
4.3.1 位形空间的 Hamilton 原理与 Lagrange 方程	95
4.3.2 相空间中的 Hamilton 原理	96
4.4 求泛函极值的其他方法	98
4.4.1 Ritz 变分法—缩小搜索范围	98
4.4.2 Lagrange 不定乘子法	99
第五章 刚体的运动	101
5.1 引言	101
5.2 刚体的定轴转动	102
5.2.1 定轴转动的运动学	102
5.2.2 刚体定轴转动动力学	105
5.3 刚体一般平面运动	108
5.3.1 运动学	109
5.3.2 动力学	110
5.4 刚体的定点运动	113
5.4.1 运动学	113
5.4.2 动力学：角动量和惯量张量	119
5.5 定点运动的刚体动力学	125
5.5.1 Euler 动力学方程	125
5.5.2 Euler-Poinsot 情况	127
5.5.3 高速陀螺的近似理论	128
5.6 刚体动力学小结	129

第一章 Lagrange 力学

本章学习拉格朗日力学. 学习拉格朗日力学的意义有两个.

第一, 从技术上说, 拉格朗日方程对我们求解力学问题有一些好处. 主要包括:

(a) 对于有约束的系统, 拉格朗日方程中不包含约束力. 即, 拉格朗日力学给出了一般性的从“牛顿方程 + 约束方程”构成的方程组中消去未知的约束力的方法. 从这个方程出发, 求解有约束的问题, 要比直接求解“牛顿方程 + 约束方程”构成的方程组容易的多.

(b) 拉格朗日方程的形式与坐标选择无关. 即, 使用极坐标、柱坐标、球坐标等等非直角坐标的拉格朗日方程, 形式上和使用直角坐标的拉格朗日方程相同.

第二, 从理论体系上说, 拉格朗日力学和后来的哈密顿力学的体系, 后来被拓展到了近代物理中. 因而学习拉格朗日力学和哈密顿力学, 为近代物理的学习铺平了道路. 比较有意思的是, 在近代物理中, 人们其实更多的是处理没有约束, 并且使用直角坐标的简单情况.

事实上, 拉格朗日力学的动机和妙处主要在这一点, 而对我们物理系的学生而言, 更重要的其实是第二点. 当然, 我们在学习拉格朗日力学的时候, 仍然要着力理解刚才说的第一点意义中蕴含的逻辑, 巧思和美妙之处.

1.1 数学复习: 复合函数求偏导的锁链法则

1.1.1 复合函数

考虑一个以 x_1, \dots, x_N 为自变量的函数 $f(x_1, \dots, x_N)$. 若 x_1, \dots, x_N 又都是一组新的自变量 y_1, \dots, y_M 的函数, 即:

$$\begin{aligned}x_1 &= x_1(y_1, \dots, y_M); \\x_2 &= x_2(y_1, \dots, y_M); \\&\vdots \\x_N &= x_N(y_1, \dots, y_M),\end{aligned}$$

则以 x_1, \dots, x_N 作为桥梁, f 可视为 y_1, \dots, y_M 的函数.

此时称 f 为复合函数, 可记为 $f[x_1(y_1, \dots, y_M), x_2(y_1, \dots, y_M), \dots, x_N(y_1, \dots, y_M)]$.

1.1.2 锁链法则

复合函数 $f[x_1(y_1, \dots, y_M), x_2(y_1, \dots, y_M), \dots, x_N(y_1, \dots, y_M)]$ 的偏导数满足锁链法则，即：

$$\left. \frac{\partial f}{\partial y_i} \right|_{y_j \neq i} = \sum_{\alpha=1}^N \left. \frac{\partial f}{\partial x_\alpha} \right|_{x_\beta \neq \alpha} \cdot \left. \frac{\partial x_\alpha}{\partial y_i} \right|_{y_j \neq i}. \quad (1.1)$$

在这里，等式左边和等式右侧第二个因式的 “ $\left. \frac{\partial f}{\partial y_i} \right|_{y_j \neq i}$ ” 表示将 $y_1 \cdots y_N$ 视为独立的自变量，并保持除 y_i 以外的自变量不变。类似地，等式右侧第二个因式的 “ $\left. \frac{\partial f}{\partial x_\alpha} \right|_{x_\beta \neq \alpha}$ ” 表示将 $x_1 \cdots x_M$ 视为独立自变量，保持 x_α 以外的自变量均不变。以后这个对不变的自变量的声明将会逐渐略去。

另外，要注意在计算 (1.1) 式中 $\left. \frac{\partial f}{\partial x_\alpha} \right|_{x_\beta \neq \alpha}$ 等项时，应先计算 $\frac{f(x_1 \cdots x_N)}{\partial x_\alpha}$ ，这是 $x_1 \cdots x_N$ 的函数，再将 $x_\alpha = x_\alpha(y_1 \cdots y_N)$, $\alpha = 1, \dots, N$ 代入上述结果。最终的结果是 $y_1 \cdots y_N$ 的函数。

锁链法则的图像理解：

锁链法则的严格证明已经在高等数学课中给出了。此处我们尝试从更直观的数学图像来理解上面得到的数学结果。我们以二元复合函数 $f[x_1(y_1, y_2), x_2(y_1, y_2)]$ 为例，计算其保持 y_2 不变而关于 y_1 的偏导。

我们的思路是将 $F_\delta = f[x_1(y_1 + \delta, y_2), x_2(y_1 + \delta, y_2)]$ 在 $\delta = 0$ 附近做小量展开，则 δ^1 （即 δ 的 1 次方）项前面的系数即为待求偏导 $\left. \frac{\partial f}{\partial y_1} \right|_{y_2}$ 。

$$\begin{aligned} F_\delta &= f \left[x_1(y_1, y_2) + \underbrace{\left. \frac{\partial x_1}{\partial y_1} \right|_{y_2} \cdot \delta + \mathcal{O}(\delta^2)}_{\equiv \Delta x_1}, x_2(y_1, y_2) + \underbrace{\left. \frac{\partial x_2}{\partial y_1} \right|_{y_2} \cdot \delta + \mathcal{O}(\delta^2)}_{\equiv \Delta x_2} \right] \\ &= f[x_1(y_1, y_2), x_2(y_1, y_2)] + \left. \frac{\partial f}{\partial x_1} \right|_{x_2} \Delta x_1 + \left. \frac{\partial f}{\partial x_2} \right|_{x_1} \Delta x_2 + \mathcal{O}(\Delta x_1^2, \Delta x_2^2, \Delta x_1 \Delta x_2) \\ &= f[x_1(y_1, y_2), x_2(y_1, y_2)] + \underbrace{\left. \frac{\partial f}{\partial x_1} \right|_{x_2} \left. \frac{\partial x_1}{\partial y_1} \right|_{y_2} \cdot \delta + \left. \frac{\partial f}{\partial x_2} \right|_{x_1} \left. \frac{\partial x_2}{\partial y_1} \right|_{y_2} \cdot \delta + \mathcal{O}(\delta^2)}_{\delta^1}. \end{aligned} \quad (1.2)$$

此处 $\mathcal{O}(\delta^2)$ 表示 δ 的 2 次方或更高次方项，其他的 $\mathcal{O}(\dots)$ 含义类似。根据上述计算，我们得到

$$\left. \frac{\partial f}{\partial y_1} \right|_{y_2} = \left. \frac{\partial f}{\partial x_1} \right|_{x_2} \left. \frac{\partial x_1}{\partial y_1} \right|_{y_2} + \left. \frac{\partial f}{\partial x_2} \right|_{x_1} \left. \frac{\partial x_2}{\partial y_1} \right|_{y_2}. \quad (1.3)$$

这正是锁链法则的结果。

锁链法则在物理学中应用十分广泛，特别是在理论力学和统计物理部分内容的学习中较重要。此外，锁链法则也在自动微分算法等领域有重要应用。

1.1.3 偏导数计算应注意的问题

在对函数进行偏导数运算时，我们必须搞清楚一个问题，即：“我们是在谁不变（即把哪些自变量视为独立变量）的情况下，对谁求偏导数？” 这里面的两个“谁”，我们都要搞清楚答案。

通常而言，第二个“谁”，也就是对谁求导，是清楚的，它会出现在我们的记号里，比如 $\partial f / \partial x_\alpha$ 就是对 x_α 求偏导数。

但是第一个“谁”，也就是“我们是在谁不变的情况下”求这个偏导数（或者说，哪些自变量视为独立变量），则要格外小心。有时候会把这个问题明确的在符号中说清楚，比如 $\left. \frac{\partial f}{\partial x_\alpha} \right|_{x_{\beta \neq \alpha}}$ ，就表示将 $x_1 \cdots x_M$ 视为独立自变量，求导时保持除了 x_α 以外的自变量均不变。然而，在很多情况下，为避免啰嗦，人们并不专门用符号说这个问题。此时我们一定要搞清楚，否则会导致很多错误和混淆。

我们现在举一个容易造成混淆的例子。考虑函数

$$f(x_1, x_2) = x_1 + x_2. \quad (1.4)$$

再假设 (x_1, x_2) 又是 (y_1, y_2) 的函数，满足

$$x_1 = y_1, \quad (1.5)$$

$$x_2 = y_1 + y_2. \quad (1.6)$$

那么如果我们在 x_2 不变的情况下对 x_1 求偏导数，就有

$$\frac{\partial f}{\partial x_1} = 1 \quad (1.7)$$

同时，如果在 y_2 不变的情况下对 y_1 求偏导数，根据锁链法则有

$$\frac{\partial f}{\partial y_1} = \frac{\partial f}{\partial x_1} \frac{\partial x_1}{\partial y_1} + \frac{\partial f}{\partial x_2} \frac{\partial x_2}{\partial y_1} = 2. \quad (1.8)$$

但是我们有 $x_1 = y_1$ 。也就是说， $\frac{\partial f}{\partial x_1}$ 和 $\frac{\partial f}{\partial y_1}$ 是把 f 对同一个东西求偏导数。之所以结果会不同，就是因为求偏导数的时候，不变的那个东西不同。计算 $\frac{\partial f}{\partial x_1}$ 的时候 x_2 不变，计算 $\frac{\partial f}{\partial y_1}$ 的时候 y_2 不变。这个不同导致了结果不同。这例子告诉我们，计算偏导数的时候，“在谁不变情况下求偏导”，跟“对谁求偏导”同样重要。虽然前者通常不会注明在我们的符号里面，但是心里一定要清楚。

例：

求证：

$$\left. \frac{\partial V(x-y)}{\partial x} \right|_y = - \left. \frac{\partial V(x-y)}{\partial y} \right|_x. \quad (1.9)$$

注意：无论 $V(x-y)$ 是否等于 $V(y-x)$ ，上式都成立。

证明：定义 $z = x - y$ 。则

$$\left. \frac{\partial V(x-y)}{\partial x} \right|_y = \left. \frac{\partial V[z(x,y)]}{\partial x} \right|_y = \frac{dV[z]}{dz} \cdot \left(\left. \frac{\partial z}{\partial x} \right|_y \right) = \frac{dV[z]}{dz} \Big|_{z=x-y}. \quad (1.10)$$

且

$$\left. \frac{\partial V(x-y)}{\partial y} \right|_x = \left. \frac{\partial V[z(x,y)]}{\partial y} \right|_x = \frac{dV[z]}{dz} \cdot \left(\left. \frac{\partial z}{\partial y} \right|_x \right) = - \left. \frac{dV[z]}{dz} \right|_{z=x-y}. \quad (1.11)$$

于是 (1.9) 的左边等于右边，命题得证。

1.2 无约束保守系统的 Lagrange 方程

本节中我们先给出无约束保守系统的拉格朗日方程的推导，从中我们可以体会拉格朗日方程“形式和广义坐标的选择无关”这一特点。在下一节我们给出有约束的一般系统的拉格朗日方程的推导，从而可以体会拉格朗日方程中“不包含约束力”这一特点。

单独讨论无约束保守系统的拉格朗日方程的另一个意义在于：在近代物理中，我们研究的对象大多为微观世界中的系统，而其中以无约束保守系统最为常见。

1.2.1 无约束保守系统的 Lagrange 方程：直角坐标

牛顿力学的回顾。

考虑由 N 个质点（即质点 $1, \dots, N$ ）组成的保守系统。质点 α ($\alpha = 1, \dots, N$) 的质量为 $M^{(\alpha)}$ ，其在惯性系 $O-xyz$ 中的位矢为 $\mathbf{r}_\alpha = x^{(\alpha)}\hat{\mathbf{e}}_x + y^{(\alpha)}\hat{\mathbf{e}}_y + z^{(\alpha)}\hat{\mathbf{e}}_z$ ，所受到的合力为 $\mathbf{F}^{(\alpha)}$ 。

我们的系统为保守系统，这意味着：存在一个势函数 $V(\mathbf{r}_1, \dots, \mathbf{r}_N; t)$ ，每个质点所受合力都满足：

$$\mathbf{F}^{(\alpha)} = -\nabla_\alpha V, \quad (\alpha = 1, \dots, N). \quad (1.12)$$

这里 $\nabla_\alpha = \hat{\mathbf{e}}_x \frac{\partial}{\partial x^{(\alpha)}} + \hat{\mathbf{e}}_y \frac{\partial}{\partial y^{(\alpha)}} + \hat{\mathbf{e}}_z \frac{\partial}{\partial z^{(\alpha)}}$ 。势能函数 V 可能依赖于时间。若 V 依赖于时间，则保守系统的机械能不守恒。

显然，每一个质点的运动都由牛顿方程决定：

$$\begin{cases} M^{(\alpha)} \frac{d^2 x^{(\alpha)}}{dt^2} = \mathbf{F}^{(\alpha)} \cdot \hat{\mathbf{e}}_x = -\frac{\partial V}{\partial x^{(\alpha)}} \\ M^{(\alpha)} \frac{d^2 y^{(\alpha)}}{dt^2} = \mathbf{F}^{(\alpha)} \cdot \hat{\mathbf{e}}_y = -\frac{\partial V}{\partial y^{(\alpha)}} \\ M^{(\alpha)} \frac{d^2 z^{(\alpha)}}{dt^2} = \mathbf{F}^{(\alpha)} \cdot \hat{\mathbf{e}}_z = -\frac{\partial V}{\partial z^{(\alpha)}} \end{cases}, \quad (\alpha = 1 \cdots N). \quad (1.13)$$

整个系统的牛顿方程共计包括 $3N$ 个二阶微分方程。若已知每个质点 0 时刻的位置和速度，则以此作为初始条件，求解上述牛顿方程，可以得到每个质点在任意 t 时刻的位置。

记号的变化。

为了本节后面推导的方便，我们更改一下记号：

(a) $3N$ 个直角坐标 $x^{(1)}, y^{(1)}, z^{(1)}, \dots, x^{(\alpha)}, y^{(\alpha)}, z^{(\alpha)}$ 统一编为 x_1, x_2, \dots, x_n , 这里 $n = 3N$ 。即: $x_1 = x^{(1)}, x_2 = y^{(1)}, x_3 = z^{(1)}, x_4 = x^{(2)}, \dots, x_n = z^{(N)}$ 。不再刻意区分质点的指标和直角坐标轴的指标。这些都统一用 x 的下标表示。

(b) 定义 $m_i (i = 1 \dots n)$ 为“坐标 x_i 的主人”的质量, 即 $m_1 = m_2 = m_3 = M^{(1)}, m_4 = m_5 = m_6 = M^{(2)}, \dots$

在这些约定下, 我们的势函数变为 $V(x_1 \dots x_n; t)$, Newton 方程将改写为:

$$m_i \frac{d^2 x_i}{dt^2} = -\frac{\partial V}{\partial x_i}, \quad (i = 1 \dots n). \quad (1.14)$$

此外, 今后为了书写紧凑起见, 我们有时候也用“ \dot{x} ”来代替“ $\frac{dx}{dt}$ ”, 用“ \ddot{x} ”来代替“ $\frac{d^2 x}{dt^2}$ ”。

上述记号的改变体现了分析力学和牛顿力学的一个差别: 即牛顿力学比较关注每个质点本身, 而分析力学关注每个坐标。在分析力学看来, 同一个质点的两个坐标, 和分属两个不同质点的两个坐标, 地位和处理方法都没有差别。在本课程未来的计算中, 未必全都使用上面的方式来定义坐标符号, 但会始终贯彻“关注坐标而不是关注质点本身”的原则。

力学量的导数运算。

系统的每个力学量 (如机械能、势能、角动量、坐标、动量等) 均可表示为以“所有质点的直角坐标、直角坐标下的速度分量、以及时间”为自变量的函数, 即 $A(x_1, \dots, x_n; \dot{x}_1, \dots, \dot{x}_n, t)$ 。在未来的计算中, 经常会把某个力学量 A 对其某个自变量 (特别是时间) 进行求导。此类运算有两种角度理解:

角度 I: 视 $x_1 \dots x_n, \dot{x}_1 \dots \dot{x}_n, t$ 为独立自变量, 则 A 为 $(2n+1)$ 元函数。保持其他自变量不变, 根据 A 的定义, 对某个特定自变量求偏导数。

角度 II: 关注某一次具体的运动。在此次运动中, 所有的 x_i 和 $\dot{x}_i (i = 1, \dots, n)$ 均是时间的函数, 且 x_i 与 $\dot{x}_i (i = 1, \dots, n)$ 不再独立 (后者由前者对时间的依赖决定), 此时只有一个独立自变量, 即时间 t 。 A 是 t 的复合函数。此角度下可把 A 对时间求导 (此导数也称为全导数)。

上述两个角度的运算在未来计算中都会用到。因此在每一步微分运算中, 我们都要明确该运算是以何种角度进行。一般来说, 在符号中会体现这一点。通常 $\frac{\partial}{\partial x_i}, \frac{\partial}{\partial \dot{x}_i}, \frac{\partial}{\partial t}$ 等符号表示从角度 I 出发的偏导数运算。 $\frac{d}{dt}$ 表示从角度 II 出发的全导数运算。

Lagrange 量和拉格朗日方程。

下面我们定义系统的 Lagrange 量 (Lagrangian, 简称拉氏量)。对我们现在研究的直角坐标系中的无约束保守系统, 拉氏量是“所有直角坐标以及直角坐标系中速度分量, 以及时间”的函数,

其定义为：

$$\begin{aligned} L[x_1, \dots, x_n; \dot{x}_1, \dots, \dot{x}_n; t] &= T - V \\ &= \sum_{i=1}^n \frac{1}{2} m_i \dot{x}_i^2 - V(x_1 \cdots x_n; t). \end{aligned} \quad (1.15)$$

以下我们从牛顿方程推导拉格朗日方程。为此，我们首先计算 $\left. \frac{\partial L}{\partial \dot{x}_j} \right|_{x_i, \dot{x}_{i \neq j}, t}$ ：

$$\begin{aligned} \left. \frac{\partial L}{\partial \dot{x}_j} \right|_{x_i, \dot{x}_{j \neq i}, t} &= \frac{\partial}{\partial \dot{x}_j} \left\{ \sum_{i=1}^n \frac{1}{2} m_i \dot{x}_i^2 - V(x_1 \cdots x_n; t) \right\} \\ &= \frac{1}{2} \sum_{i=1}^n m_i \frac{\partial}{\partial \dot{x}_j} \dot{x}_i^2 \\ &= \frac{1}{2} \sum_{i=1}^n m_i \cdot 2 \dot{x}_i \delta_{ij} \\ &= m_j \dot{x}_j, \end{aligned} \quad (j = 1, \dots, n). \quad (1.16)$$

这里 δ_{ij} 定义为：

$$\delta_{ij} = \begin{cases} 1, & (i = j) \\ 0, & (i \neq j) \end{cases}. \quad (1.17)$$

根据 (1.16) 式，我们有

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}_j} \right) = \frac{d}{dt} (m_j \dot{x}_j) = m_j \ddot{x}_j; \quad (j = 1, \dots, n). \quad (1.18)$$

注意上式中的运算 d/dt 是从角度 II 出发的，而且括号中的偏导数运算 $\frac{\partial L}{\partial \dot{x}_j}$ 则从角度 I 出发。另一方面，拉氏量的定义 (1.15) 还给出：

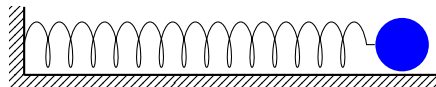
$$\left. \frac{\partial L}{\partial x_j} \right|_{x_{i \neq j}, \dot{x}_k, t} = -\frac{\partial V}{\partial x_j}, \quad (j = 1, \dots, n). \quad (1.19)$$

利用公式 (1.16) 与 (1.19)，我们发现牛顿方程 (1.14) 可以改写成

$$\boxed{\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}_j} \right) - \frac{\partial L}{\partial x_j} = 0, \quad (j = 1 \cdots n).} \quad (1.20)$$

这就是无约束保守系统使用直角坐标的拉格朗日方程。我们再次强调：它和牛顿方程完全等价。该拉格朗日方程是由 n 个耦合的二阶常微分方程构成的常微分方程组。以 0 时刻每个直角坐标以及直角坐标系中的速度分量的取值 $\{x_1(0), \dots, x_n(0), \dot{x}_i(0)\}$ 为初始条件求解此方程组，可得到每个直角坐标在任意时刻的取值 $\{x_1(t), \dots, x_n(t)\}$ 。

例 1：一维简谐振子。



考虑图中的一维弹簧振子，假设振子在 x 轴上运动，选择振子的平衡位置为坐标原点。设振子质量为 m ，弹簧弹性系数为 k 。我们现在写振子的拉氏量和拉式方程。

显然，这是一个 1 维单个质点保守系统，只有一个直角坐标 (即 $n = 1$)，即振子的坐标 x 。振子的动能为 $T = \frac{1}{2}m\dot{x}^2$ ，振子的势能为 $V = \frac{1}{2}kx^2$ ，故振子的 Lagrangian 为

$$L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2. \quad (1.21)$$

振子的 Lagrange 方程为：

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) - \frac{\partial L}{\partial x} = 0. \quad (1.22)$$

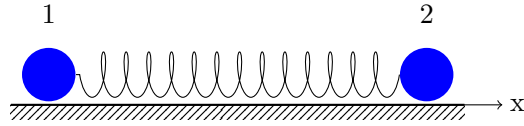
由前面给出的 T 与 V 的表达式，我们显然有

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}} = \frac{d}{dt} m\dot{x} = m\ddot{x}; \quad (1.23)$$

$$\frac{\partial L}{\partial x} = -kx. \quad (1.24)$$

于是振子的 Lagrange 方程 (1.22) 可以改写为 $m\ddot{x} + kx = 0$ ，显然这正是振子的牛顿方程。

例 2: 两个耦合的 1 维振子。



考虑图中的两个振子 1 和 2，假设它们都在 x 轴上运动，坐标分别为 x_1, x_2 ($x_2 > x_1$)，质量分别为 m_1, m_2 。弹簧平衡长度为 l 。我们现在写出该系统的拉格朗日量和拉格朗日方程。

该系统包含两个坐标为 x_1, x_2 。系统的总动能 T 和势能 V 分别为

$$T = \frac{1}{2}m_1\dot{x}_1^2 + \frac{1}{2}m_2\dot{x}_2^2, \quad (1.25)$$

$$V = \frac{1}{2}k(x_2 - x_1 - l)^2. \quad (1.26)$$

(请大家自行验证上述势能满足质点组总势能的要求 (1.12))。于是系统的 Lagrangian 为：

$$L = T - V = \frac{1}{2}m_1\dot{x}_1^2 + \frac{1}{2}m_2\dot{x}_2^2 - \frac{k}{2}(x_2 - x_1 - l)^2. \quad (1.27)$$

系统的 Lagrange 方程为：

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}_1} - \frac{\partial L}{\partial x_1} = 0; \quad (1.28)$$

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}_2} - \frac{\partial L}{\partial x_2} = 0. \quad (1.29)$$

直接的计算给出，上面的两个拉格朗日方程分别可改写为 $m_1\ddot{x}_1 = k(x_2 - x_1 - l)$ 和 $m_2\ddot{x}_2 = -k(x_2 - x_1 - l)$ ，这正是该系统的牛顿方程。

1.2.2 广义坐标与广义速度

广义坐标。

在很多时候，直角坐标并不能让我们最为方便的处理问题。此时我们会选择其他坐标，例如极坐标、柱坐标、球坐标等。一般的，能够描述系统中每个质点位置的一组数字，都称为一组**广义坐标**或正则坐标（Canonical coordinate）。广义坐标的概念是直角坐标概念的推广，它包括但不限于直角坐标。

在无约束的系统中，广义坐标和直角坐标的个数肯定相同。若系统直角坐标为 (x_1, \dots, x_n) ，则广义坐标也有 n 个，可记为 (q_1, \dots, q_n) 。

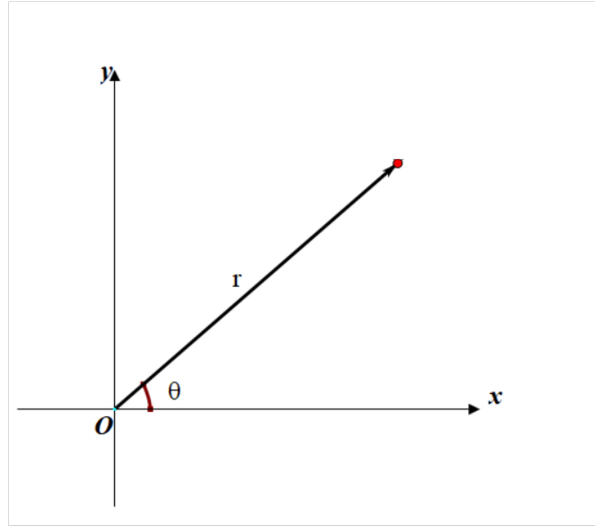


图 1.1: 极坐标

例如，对于一个二维平面上的质点，其位置既可以用直角坐标 (x, y) 描述，也可以用极坐标 (r, θ) 作为广义坐标描述。极坐标就是一组广义坐标。

一般而言，广义坐标是直角坐标和时间的函数。我们可以用坐标变换函数

$$\begin{cases} x_1 = x_1(q_1 \cdots q_n; t) \\ \vdots \\ x_n = x_n(q_1 \cdots q_n; t) \end{cases} \quad (1.30)$$

来描述广义坐标和普通的直角坐标之间的变换关系。当变换关系被明确后，广义坐标也就被定义清楚了。例如，在上述极坐标的例子中，二维直角坐标和极坐标的变换关系即为：

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta. \end{cases} \quad (1.31)$$

广义速度。

系统在运动中，广义坐标会随时间变化。定义 $(\dot{q}_1, \dots, \dot{q}_n)$ 为广义坐标对时间的导数：

$$\dot{q}_j = \frac{dq_j(t)}{dt}, \quad (j = 1, \dots, n). \quad (1.32)$$

$(\dot{q}_1, \dots, \dot{q}_n)$ 被称为广义速度。广义速度刻画了运动中广义坐标随时间变化的快慢。显然，广义速度是直角坐标系下速度 $(\dot{x}_1, \dots, \dot{x}_n)$ 的推广。

根据定义和坐标变换关系 (1.30)，我们可以把直角坐标系下的每个速度分量都写成广义坐标、广义速度以及时间的函数，即：

$$\begin{aligned} \dot{x}_i &= \frac{d}{dt} \left[x_i(q_1, \dots, q_n, t) \right] \\ &= \sum_{j=1}^n \frac{\partial x_i}{\partial q_j} \dot{q}_j + \frac{\partial x_i}{\partial t} \end{aligned} \quad (1.33)$$

$$\equiv \dot{x}_i[q_1, \dots, q_n; \dot{q}_1, \dots, \dot{q}_n, t], \quad (i = 1, \dots, n) \quad (1.34)$$

注意式 (1.33) 中 $\frac{\partial x_i}{\partial q_j}$ 与 $\frac{\partial x_i}{\partial t}$ 均为 (q_1, \dots, q_n) 与 t 的函数。例如，在前面极坐标的例子中，有

$$\begin{cases} \dot{x} = \dot{r} \cos \theta - r \dot{\theta} \sin \theta \\ \dot{y} = \dot{r} \sin \theta + r \dot{\theta} \cos \theta \end{cases}. \quad (1.35)$$

两个有用的公式。

下面我们介绍两个有用的公式，它们将在本课程中反复使用，以后我们将称之为：“有用的公式 1” 和 “有用的公式 2”。

有用的公式 1:

$$\left. \frac{\partial \dot{x}_i}{\partial \dot{q}_j} \right|_{q_1, \dots, q_n; t; \dot{q}_k \neq j} = \frac{\partial x_i}{\partial q_j}, \quad (i, j = 1, \dots, n). \quad (1.36)$$

证明：

$$\text{左边} = \frac{\partial}{\partial \dot{q}_j} \left[\sum_{\alpha=1}^n \frac{\partial x_i}{\partial q_\alpha} \dot{q}_\alpha + \frac{\partial x_i}{\partial t} \right] = \sum_{\alpha=1}^n \frac{\partial x_i}{\partial q_\alpha} \frac{\partial \dot{q}_\alpha}{\partial \dot{q}_j} = \sum_{\alpha=1}^n \frac{\partial x_i}{\partial q_\alpha} \delta_{\alpha j} = \text{右边}. \quad (1.37)$$

上面的证明中用到了如下事实： $\frac{\partial x_i}{\partial q_\alpha}$ 与 $\frac{\partial x_i}{\partial t}$ ($i, \alpha = 1, \dots, n$) 均为 (q_1, \dots, q_n) 与 t 的函数。

有用的公式 2:

$$\left. \frac{\partial \dot{x}_i}{\partial q_j} \right|_{\dot{q}_1, \dots, \dot{q}_n; t; q_k \neq i} = \frac{d}{dt} \left(\frac{\partial x_i}{\partial q_j} \right) = \frac{d}{dt} \left(\frac{\partial \dot{x}_i}{\partial \dot{q}_j} \right), \quad (i, j = 1, \dots, n). \quad (1.38)$$

证明：

首先,

$$\begin{aligned}\left.\frac{\partial \dot{x}_i}{\partial q_j}\right|_{\dot{q}_1, \dots, \dot{q}_n; t; q_k \neq i} &= \frac{\partial}{\partial q_j} \left[\sum_{\alpha=1}^n \frac{\partial x_i}{\partial q_\alpha} \dot{q}_\alpha + \frac{\partial x_i}{\partial t} \right] \\ &= \sum_{\alpha=1}^n \frac{\partial^2 x_i}{\partial q_\alpha \partial q_j} \dot{q}_\alpha + \frac{\partial^2 x_i}{\partial t \partial q_j}.\end{aligned}\quad (1.39)$$

其次,

$$\begin{aligned}\frac{d}{dt} \left(\frac{\partial x_i}{\partial q_j} \right) &= \sum_{\alpha=1}^n \left[\frac{\partial}{\partial q_\alpha} \left(\frac{\partial x_i}{\partial q_j} \right) \right] \dot{q}_\alpha + \frac{\partial}{\partial t} \left(\frac{\partial x_i}{\partial q_j} \right) \\ &= \sum_{\alpha=1}^n \frac{\partial^2 x_i}{\partial q_\alpha \partial q_j} \dot{q}_\alpha + \frac{\partial^2 x_i}{\partial t \partial q_j}.\end{aligned}\quad (1.40)$$

因此, (1.38) 式的第一个等号成立。另一方面, 由 “有用的公式 1”, 可知 (1.38) 式的第二个等号也成立。

在上述极坐标的例子中, 面, 我们可以利用公式 (1.34) 和 (1.35) 直接验证关系 (1.36) 和 ((1.38)).

1.2.3 无约束保守系统的 Lagrange 方程: 广义坐标

我们依旧考虑一个无约束保守系统, 假设该系统直角坐标为 x_1, \dots, x_n , 广义坐标为 q_1, \dots, q_n 。该系统直角坐标满足拉格朗日方程 (1.20), 其中拉氏量 $L = T - V$ 由 (1.15) 式给出。现在推导广义坐标随时间变化满足的方程。

拉氏量可被看成广义坐标和广义速度的函数。

我们首先注意到: 该系统的拉氏量可被看成广义坐标和广义速度的函数, 即:

$$\begin{aligned}L &= T - V \\ &= L[x_1(q_1, \dots, q_n, t), \dots, x_n(q_1, \dots, q_n, t); \dot{x}_1(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t), \dots, \dot{x}_n(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t); t] \\ &\equiv L[q_1, \dots, q_n; \dot{q}_1, \dots, \dot{q}_n; t].\end{aligned}\quad (1.41)$$

例如, 一个在 xy 平面上运动的质点, 若其势能为 $V(x, y, t)$, 则其拉氏量在直角坐标 (x, y) 和极坐标中 (r, θ) 的表达式分别为:

$$\begin{aligned}L &= \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - V(x, y, t) \\ &= \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - V(r \cos \theta, r \sin \theta, t).\end{aligned}\quad (1.42)$$

在推导上式第二个等号时, 我们用到了直角坐标和极坐标之间的坐标变换 (1.31) 式和相应的速度变换 (1.40) 式。

广义坐标下的 Lagrange 方程。

下面我们推导广义坐标下的 Lagrange 方程形式。首先，我们有

$$\begin{aligned}\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) &= \frac{d}{dt} \left[\sum_{\alpha=1}^n \frac{\partial L}{\partial \dot{x}_\alpha} \frac{\partial \dot{x}_\alpha}{\partial \dot{q}_j} \right] \\ &= \sum_{\alpha=1}^n \left[\frac{\partial \dot{x}_\alpha}{\partial \dot{q}_j} \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}_\alpha} \right) + \frac{\partial L}{\partial \dot{x}_\alpha} \frac{d}{dt} \left(\frac{\partial \dot{x}_\alpha}{\partial \dot{q}_j} \right) \right].\end{aligned}\quad (1.43)$$

其次，根据直角坐标中的拉格朗日方程 (1.20)，我们有

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}_\alpha} \right) = \frac{\partial L}{\partial x_\alpha}, \quad (\alpha = 1, \dots, n). \quad (1.44)$$

而根据有用的公式 1 和有用的公式 2，我们又分别有

$$\frac{\partial \dot{x}_\alpha}{\partial \dot{q}_j} = \frac{\partial x_\alpha}{\partial q_j}, \quad (1.45)$$

$$\frac{d}{dt} \left(\frac{\partial \dot{x}_\alpha}{\partial \dot{q}_j} \right) = \frac{\partial \dot{x}_\alpha}{\partial q_j}, \quad (\alpha, j = 1, \dots, n). \quad (1.46)$$

将 (1.44-1.46) 代入 (1.43)，我们最终得到广义坐标下无约束保守系统的 Lagrange 方程：

$$\boxed{\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} = 0, \quad (j = 1 \cdots n).} \quad (1.47)$$

对比公式 (1.20) 和 (1.47)，我们发现：在直角坐标和广义坐标下，无约束保守系统的拉格朗日方程都有相同的形式，而且拉氏量的定义也相同，都是 $L = T - V$ 。这是拉格朗日力学的一个漂亮的地方。

此外，在使用广义坐标的拉格朗日方程处理问题的时候，我们只需要先把直角坐标和直角坐标的速度分量用广义坐标和广义速度表示，即可求出广义坐标的拉氏量。在这个过程中不需要用广义坐标表示加速度，也不需要处理和广义坐标对应的“坐标基矢”。这都是拉格朗日力学比牛顿力学好用的地方。

具体问题的处理程序。

在广义坐标下利用 Lagrange 力学求解无约束保守系统的演化，一般流程如下：

- (1) 明确直角坐标和广义坐标的变换关系，以及相应的速度变换关系。
- (2) 把拉氏量 $L = T - V$ 写成直角坐标 $x_{1,\dots,n}$ 和直角坐标下速度分量 $\dot{x}_{1,\dots,n}$ 的函数。
- (3) 将 (1) 中得出的变换关系代入拉氏量，最终把拉氏量表示成广义坐标 $q_{1,\dots,n}$ 和广义坐标下速度分量 $\dot{q}_{1,\dots,n}$ 的函数。
- (4) 列出广义坐标满足的拉格朗日方程组 (1.47)。
- (5) 利用初始条件 $q_{1,\dots,n}(t=0)$ 和 $\dot{q}_{1,\dots,n}(t=0)$ 求解该方程组，得出每时每刻的广义坐标取值 $q_{1,\dots,n}(t)$ 。

例： xy 平面上单个质点极坐标的 *Lagrange* 方程。

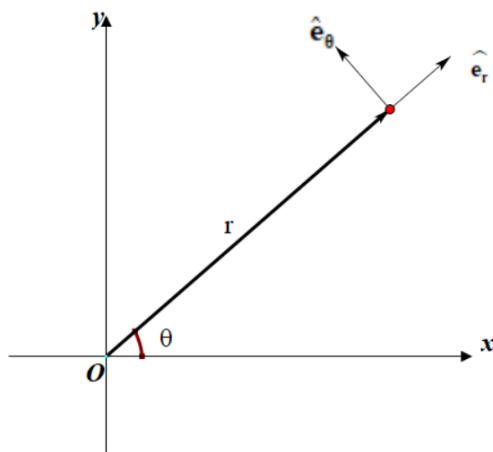


图 1.2: 极坐标和相应坐标基矢

我们仍考虑前面研究过的 xy 平面上单个质点的例子，利用上面的流程推导极坐标满足的拉格朗日方程。

(1) 直角坐标和广义坐标的坐标变换以及速度的变换：

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta \end{cases}, \quad \begin{cases} \dot{x} = \dot{r} \cos \theta - r \dot{\theta} \sin \theta \\ \dot{y} = \dot{r} \sin \theta + r \dot{\theta} \cos \theta \end{cases}. \quad (1.48)$$

(2) 直角坐标下的拉氏量：

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - V(x, y, t). \quad (1.49)$$

(3) 将坐标变换带入 (1.49)，得到极坐标下的拉氏量：

$$L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - V(r \cos \theta, r \sin \theta, t). \quad (1.50)$$

(4) 将 (1.50) 带入 (1.47)，列出拉格朗日方程组：

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{r}} - \frac{\partial L}{\partial r} = 0 \quad \Rightarrow \quad m\ddot{r} = -\frac{\partial V}{\partial x} \cos \theta - \frac{\partial V}{\partial y} \sin \theta + mr\dot{\theta}^2; \quad (1.51)$$

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} = 0 \quad \Rightarrow \quad mr^2\ddot{\theta} + 2mr\dot{\theta}\dot{r} = \frac{\partial V}{\partial x} r \sin \theta - \frac{\partial V}{\partial y} r \cos \theta. \quad (1.52)$$

利用直接的计算可以证明，上述拉格朗日方程 (1.51) 和 (1.52) 等价于极坐标下的牛顿方程 $m\mathbf{a} \cdot \hat{\mathbf{e}}_r = -\nabla V \cdot \hat{\mathbf{e}}_r$ 和 $m\mathbf{a} \cdot \hat{\mathbf{e}}_\theta = -\nabla V \cdot \hat{\mathbf{e}}_\theta$ 。

1.3 完整约束中的 Lagrange 方程

1.3.1 完整约束

在很多情况下,质点系统的运动都处在一定的约束中。比如算盘珠的运动,就被算盘杆约束在一维的轴上,而或者和过山车的运动则被约束在各自的轨道上等等。若系统的直角坐标为 $(x_1 \cdots x_n)$, 而运动中必须满足的约束条件可表示为: 每时每刻的直角坐标都必须满足 m 个 ($m < n$) 方程:

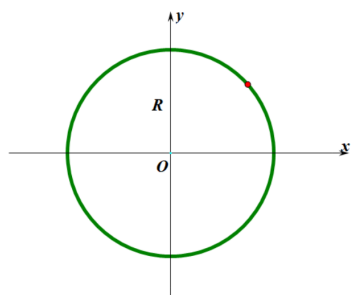
$$f_\alpha(x_1 \cdots x_n; t) = 0 \quad (\alpha = 1 \cdots m) \quad (1.53)$$

则称系统受到**完整约束**。上面的等式称为约束方程, 等号左边的函数 $f_\alpha(x_1 \cdots x_n, t)$ ($\alpha = 1 \cdots m$) 称为约束函数。我们要注意: 完整约束的约束条件都表现为等式 (约束方程), 且约束方程中的约束函数都与速度无关。

进一步地, 若所有约束函数都不显含时间, 即 $\frac{\partial f_\alpha}{\partial t} = 0$ ($\alpha = 1 \cdots m$), 则称系统受**稳定完整约束**。反之, 若存在至少一个约束函数 f_α , 满足 $\frac{\partial f_\alpha}{\partial t} \neq 0$, 则称系统受**不稳定完整约束**。

相应的, 若约束条件包含速度, 或者约束条件表达为某些不等式, 则称系统受到**非完整约束**。例如: 限制粒子在一个有限大小的盒子中运动, 此时的约束条件为 $0 \leq x, y, z \leq L$, 是典型的非完整约束。

完整约束例 1: 固定光滑圆环上的质点



直角坐标: (x, y) , 坐标数 $n = 2$ 。

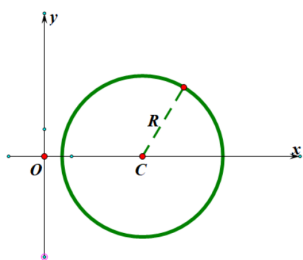
约束条件: $x^2 + y^2 - R^2 = 0$, 约束条件个数 $m = 1$ 。

约束函数: $f(x, y) = x^2 + y^2 - R^2$ 。

这是一个稳定完整约束。

图 1.3: 一个质点被限制在固定光滑半径为 R 的圆环上运动

完整约束例 2: 圆心匀速运动的圆环上的质点



此时圆心坐标为 $x_c = vt, y_c = 0$ ，于是约束条件变为

$$(x - vt)^2 + y^2 - R^2 = 0 \quad (1.54)$$

图 1.4: 一个质点被限制在光滑的半径为 R 的圆环上运动，但这次圆环中心以速度 v 沿 x 轴匀速运动。

这是不稳定完整约束

完整约束系统的广义坐标。

下面我们讨论完整约束系统的广义坐标. 当质点组的运动满足完整约束 (1.53) 的时候，直角坐标 (x_1, \dots, x_n) 彼此之间是相互关联的，他们必须满足方程组 (1.53)，因此只有 $n - m$ 个是独立的。这个时候我们继续使用 (x_1, \dots, x_n) 刻画我们系统的运动，就不方便。此时我们可以选择 $n - m$ 个广义坐标 $(q_1 \cdots q_{n-m})$ 来刻画我们系统的运动， $(q_1 \cdots q_{n-m})$ 可以是直角坐标 (x_1, \dots, x_n) 中的 $n - m$ 个，也可以是这些直角坐标的函数。此时直角坐标和广义坐标之间的变换关系为：

$$\begin{cases} x_1 = x_1(q_1 \cdots q_{n-m}; t) \\ \vdots \\ x_n = x_n(q_1 \cdots q_{n-m}; t). \end{cases} \quad (1.55)$$

例如，在前两个例子中，都只有一个广义坐标，可取为 θ （即质点位置矢量和 x 轴的夹角）。在例 1 中，坐标变换关系为

$$\begin{cases} x = R \cos \theta \\ y = R \sin \theta \end{cases}. \quad (1.56)$$

在例 2 中，坐标变换关系为

$$\begin{cases} x = R \cos \theta + vt \\ y = R \sin \theta \end{cases}. \quad (1.57)$$

这里我们强调：一组在取值范围内的广义坐标 $(q_1 \cdots q_{n-m})$ 代入坐标变换 (1.55) 式后，所得到的直角坐标 $x_1(q_1 \cdots q_{n-m}; t), \dots, x_n(q_1 \cdots q_{n-m}; t)$ 必然能自动满足约束方程。即：若将坐标变换 (1.55) 代入约束函数 $f_\alpha(x_1 \cdots x_n; t)$ ($\alpha = 1, \dots, m$)，则所得的 $f_{1, \dots, m}$ 作为 $q_1 \cdots q_{n-m}$ 的复合函数，不管自变量 $q_{1 \dots n-m}$ 取值如何变化，其函数值总为 0（即： $f_{1, \dots, m}$ 是 $q_{1 \dots n-m}$ 的取值为 0 的常数函数）。写成公式就是：

$$f_\alpha \left[x_1(q_1 \cdots q_{n-m}; t), \dots, x_n(q_1 \cdots q_{n-m}; t); t \right] \equiv 0, \quad (\alpha = 1 \cdots m). \quad (1.58)$$

比如, 在上述例 1 中, 显然有 $f[x(\theta), y(\theta)] = R^2 \cos^2 \theta + R^2 \sin^2 \theta - R^2 \equiv 0$ 。

而式 (1.58) 的一个推论是:

$$\left. \frac{\partial f_\alpha}{\partial q_j} \right|_{q_i \neq j} = \sum_{k=1}^n \frac{\partial f_\alpha}{\partial x_k} \frac{\partial x_k}{\partial q_j} \equiv 0, \quad (\alpha = 1, \dots, m; j = 1, \dots, n - m). \quad (1.59)$$

比如, 例 1 中我们显然有

$$\frac{\partial f}{\partial \theta} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \theta} = 2x \cdot (-R \sin \theta) + 2y \cdot R \cos \theta = -2R^2 \cos \theta \sin \theta + 2R^2 \sin \theta \cos \theta = 0.$$

在未来推导约束系统的拉格朗日方程的时候, 我们会用到 (1.59)。

完整约束例 3: 双摆

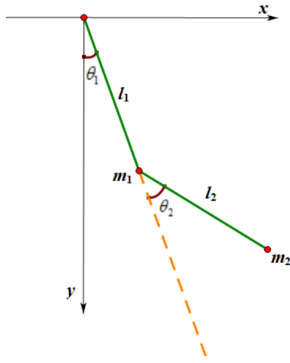


图 1.5: 双摆: 质点 1(质量 m_1) 通过长 l_1 的轻杆和原点相连, 并可以绕着原点在 xy 平面上转动。而质点 2 (质量 m_2) 通过长为 l_2 的轻杆和质点 1 相连, 并可以绕着质点 1 在 xy 平面上转动。

直角坐标: x_1, y_1, x_2, y_2 ($n = 4$)(即两个质点各自的 x, y 坐标),

其约束条件 ($m = 2$):

$$\begin{cases} x_1^2 + y_1^2 - l_1^2 = 0 \\ (x_1 - x_2)^2 + (y_1 - y_2)^2 - l_2^2 = 0 \end{cases} \quad (1.60)$$

广义坐标: θ_1, θ_2 。

坐标变换:

$$\begin{cases} x_1 = l_1 \cos \theta \\ y_1 = l_1 \sin \theta \\ x_2 = l_1 \cos \theta_1 + l_2 \cos \theta_2 \\ y_2 = l_1 \sin \theta_1 + l_2 \sin \theta_2 \end{cases} \quad (1.61)$$

读者可自行验证双摆的广义坐标和直角坐标满足前述结论 (1.58) 和 (1.59)。

1.3.2 约束力

A. 约束力的字面意思

限制质点在满足约束条件的特定轨道上运动, 一定存在一些力的作用。这些起到约束作用, 强迫质点按照约束条件的方式运动的力, 即为本节所要讨论的约束力。

牛顿力学中这些力通常未知, 会增加方程的个数 (此时用牛顿力学, 通常要解 $n+m$ 个相互耦合的方程, 即 n 个牛顿方程 (所有质点的 $\mathbf{F} = m\mathbf{a}$) 和 m 个约束方程 $f_\alpha(x_1, \dots, x_n) = 0, \alpha = 1, \dots, m$)。拉格朗日力学可以消除很多此类型力。

在本节和下一节中, 我们将首先给出 (可被拉格朗日力学消去) 的约束力的数学定义。我们“下定义”的方法如下: 首先, 对稳定完整约束总结这些力的多数情况下的主要共同点, 再将共同点推广到一般的稳定完整约束, 作为约束力的数学定义, 最后将定义拓展到不稳定完整约束 (本小节)。

最后，按照其数学定义，带入拉格朗日方程，进行处理（下一小节）。我们将证明：满足本节给出的数学定义的约束力可以自动从拉格朗日方程中消去。这样就解决了大部分实际问题。即，大部分情况下用拉格朗日力学，只要解 $n - m$ 个方程即可。

B. 从例子中体会约束力的特点

在本节中，我们只考虑稳定完整约束的情况。

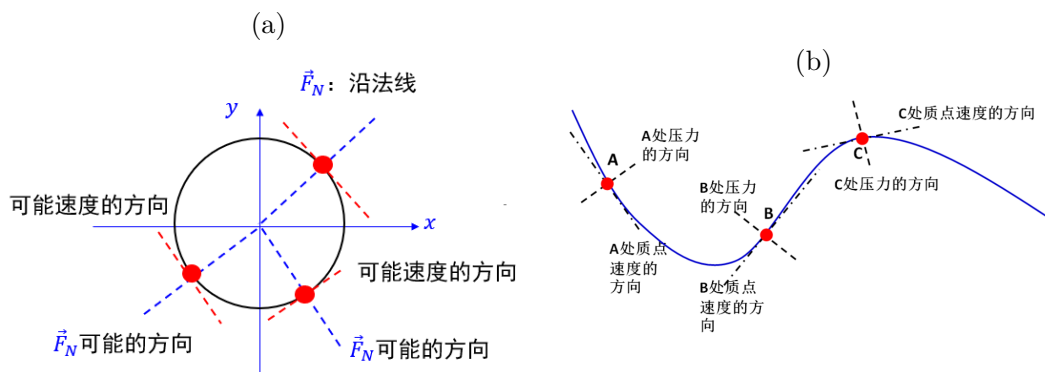


图 1.6: (a) 光滑圆环上的质点及其起约束作用的力 F_N 。(b) 光滑铁丝上的质点及其起约束作用的力 F_N

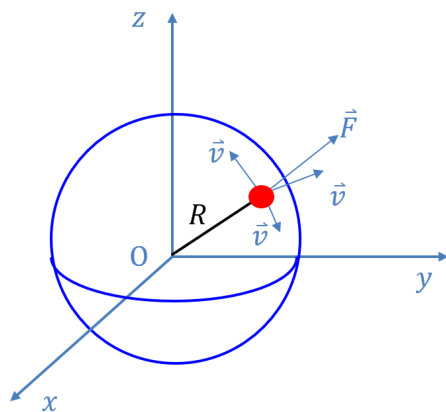


图 1.7: 球面摆上的质点

Example1 光滑圆环上的质点 如图 1.6-(a) 所示的质点不可能凭空便被约束在圆环上，迫使其始终停留在圆环上的原因是圆环向质点施加了支持力 F_N 作为约束力。我们会发现， F_N 总是沿着圆环的法线方向，且质点的所有可能运动方向均沿着圆环的切线方向，这里只要是满足约束条件的运动，在该点的速度就都是该点处可能的速度。由此，只要质点在圆环上运动，那么某点处的 F_N 就一定与质点在该处所有可能的速度方向垂直。

Example2 光滑铁丝上的质点 如图 1.6-(b) 所示, 光滑铁丝上的质点所受到的起约束作用的力是铁丝给质点的压力 \mathbf{F}_N 。质点在铁丝上某处可能的速度沿着该点的切线方向, 而铁丝给质点的压力 \mathbf{F}_N 沿着该位置的法线方向, 和切线方向垂直。因此, 只要质点在铁丝上运动, 某处的 \mathbf{F}_N , 必然与质点在该点处所有可能的速度垂直。

Example3 球面摆上的质点 如图 1.7 所示, 考虑一个长为 R 的质点, 在三维空间中绕原点自由转动, 顶端有一个质点, 构成球面摆。于是, 质点被约束在球面上运动, 其约束作用的力是杆给质点的拉力 (或压力) \mathbf{F} 。而满足约束的运动的所有可能的速度, 都与所在位置的球面相切 (即在切面上)。而起约束作用的力 \mathbf{F} 沿法线方向, 和切面垂直。因此, 只要质点在球面上运动, 某处的 \mathbf{F} 必然与质点在该处所有可能的速度垂直。

这些例子启发我们, 在**稳定完整约束**下, 起到约束作用的力具有一个共同特点: 起约束作用的力的共同特点: 某位置的“起约束作用的力”, 和满足约束条件的运动中在该位置所有可能的速度 (包括但不限于真实速度) 全都垂直。

C. 约束力的数学定义

我们将上面的性质, 总结成可以推广到一般情况下的数学定义。首先我们给出稳定完整约束下约束力的矢量表式。假设系统包含 N 个质点, 其中第 j 个质点 ($j = 1, \dots, N$) 的坐标为 $\mathbf{r}_j = x^{(j)}\mathbf{e}_x + y^{(j)}\mathbf{e}_y + z^{(j)}\mathbf{e}_z$, 约束条件为 $f_\alpha(\mathbf{r}_1 \cdots \mathbf{r}_N) = 0$, 于是在运动中**可能的速度**, 可以通过对约束函数求时间的全导数得到。注意到,

$$\frac{d}{dt}f_\alpha[\mathbf{r}_1(t) \cdots \mathbf{r}_N(t)] = \sum_{j=1}^N \nabla_j f_\alpha[\mathbf{r}_1(t) \cdots \mathbf{r}_N(t)] \cdot \dot{\mathbf{r}}_j \stackrel{!}{=} 0 \quad (1.62)$$

因此, 满足约束条件的运动中, 各个质点的速度 $(\dot{\mathbf{r}}_1, \dots, \dot{\mathbf{r}}_N)$, 一定将要满足

$$\sum_{j=1}^N \nabla_j f_\alpha[\mathbf{r}_1(t) \cdots \mathbf{r}_N(t)] \dot{\mathbf{r}}_j = 0 \quad (\alpha = 1, \dots, N) \quad (1.63)$$

因此, 当各质点处于位置 $(\mathbf{r}_1 \cdots \mathbf{r}_N)$ 时, 一组矢量 $(\mathbf{w}_1, \dots, \mathbf{w}_N)$ 如果满足以下 m 个方程

$$\sum_{j=1}^N \nabla_j f_\alpha[\mathbf{r}_1 \cdots \mathbf{r}_N(t)] \cdot \mathbf{w}_j = 0$$

那么这一组矢量就是满足约束条件的运动在 $(\mathbf{r}_1 \cdots \mathbf{r}_N)$ 处的一组可能的速度。值得注意的是, 在给定的约束和给定的位置下, 可能的速度有不止一种选择。例如对于[Example1](#)的圆环, 我们有约束条件

$$f(\mathbf{r}) = |\mathbf{r}|^2 - R^2 = 0 \quad (1.64)$$

显然在 \mathbf{r} 处 $\nabla f[\mathbf{r}]$ 沿着法线方向, 因此满足 $(\nabla f(\mathbf{r})) \cdot \mathbf{w} = 0$ 的矢量一定沿着切线方向。

下面推广“系统所收到的某一个力 \mathbf{F} ”的概念, 它一般指某一组真实的力, 包括系统中每个质点收到的某一个具体的力, 即

$$\mathbf{F} = (\mathbf{F}^{(1)}, \dots, \mathbf{F}^{(j)}) \quad (1.65)$$

这里 $\mathbf{F}^{(j)}$ 指质点 j 所受到的某一个具体的力，它的物理来源可以是单一的，也可以是几个来源的合力，也可以是零。总之，只要给每一个质点都指定了所受的某一个力，我们就认为已经指定了系统的某一个力。

Example4 固定光滑圆环上的两个质点，我们可以指定一个力 \mathbf{F} ，它包括“质点 1 受到的圆环向外的压力”，以及“质点 2”所受的重力，即

$$\mathbf{F} = (\mathbf{F}^{(1)}, \mathbf{F}^{(2)}) \quad \mathbf{F}^{(1)} = \mathbf{F}_N \quad \mathbf{F}^{(2)} = -m_2 g \mathbf{e}_y \quad (1.66)$$

我们也可以指定另一个力 \mathbf{F} ，它只包括“质点 1 受到的圆环向外的压力”，于是

$$\mathbf{F} = (\mathbf{F}^{(1)}, \mathbf{F}^{(2)}) \quad \mathbf{F}^{(1)} = \mathbf{F}_N \quad \mathbf{F}^{(2)} = 0 \quad (1.67)$$

完成了对于“可能的速度”和“系统所受的一个力”的概念申明后，我们就可以给出矢量表述下约束力的定义

定义 1.3.1 (约束力 [稳定完整约束、矢量表述].) 设有力 $\mathbf{F}_N = (\mathbf{F}_N^{(1)}, \dots, \mathbf{F}_N^{(N)})$ ，若在任意满足约束条件的位置 $(\mathbf{r}_1, \dots, \mathbf{r}_N)$ ，对于满足约束条件的任意可能的速度 $(\mathbf{w}_1, \dots, \mathbf{w}_N)$ ，即任意一组满足

$$\sum_{j=1}^N (\nabla_j f_\alpha) \cdot \mathbf{w}_j = 0, (\alpha = 1, \dots, m) \quad (1.68)$$

的速度 $(\mathbf{w}_1, \dots, \mathbf{w}_N)$ ，该位置的 $\mathbf{F}_N = (\mathbf{F}_N^{(1)}, \dots, \mathbf{F}_N^{(N)})$ 都满足

$$\sum_{j=1}^N \mathbf{F}_N^{(j)} \cdot \mathbf{w}_j = 0 \quad (1.69)$$

则称系统所受的这一个力 \mathbf{F}_N 为约束力。

在很多实际情况下，很容易根据约束条件直接构造所有可能的速度 $(\mathbf{w}_1, \dots, \mathbf{w}_N)$ ，不一定非要带入公式 $\sum_{j=1}^N (\nabla_j f_\alpha) \cdot \mathbf{w}_j = 0$ 来验证。另外，可以看出稳定完整约束中的约束力肯定不做功，但是不做功的力未必是约束力（例如洛伦兹力就不是约束力），这是因为不做功的力只需要和真实速度垂直，而约束力是和所有满足约束条件的可能速度垂直，后者包括但不限于前者。在很多教材中，将可能的速度 \mathbf{w}_j 与“无穷小时间” δt 相乘，定义为“虚位移” $\delta \mathbf{r}_j$ ，之后再虚位移表述上述定义。其内容与我们这里等价。

由于稳定完整约束中大多数起约束作用的力都满足此定义，以后，稳定完整约束的“约束力”专指满足此定义的力。满足我们定义的约束力，可以被拉格朗日力学自动消去。而不满足此定义的力，不管是否未知，是否起约束作用，都不能被拉格朗日方程消去。

例如在前面 Example1 中，固定光滑圆环上的质点的约束条件为 $f(\mathbf{r}) = |\mathbf{r}|^2 - R^2 = 0$ ，而圆环的压力 \mathbf{F}_N 沿着法线方向。如前所述，在 \mathbf{r} 处满足约束条件的任何可能速度 \mathbf{w} 必然沿切线方向，这一点既可以直接判断，也可以用严格判据，即验证得到任意满足 $\nabla f[\mathbf{r}] \cdot \mathbf{w} = 0$ 的矢量 \mathbf{w} 必然沿着切线方向，因此必然有 $\mathbf{F}_N \cdot \mathbf{w} = 0$ ，从而 \mathbf{F}_N 就满足约束力的定义。

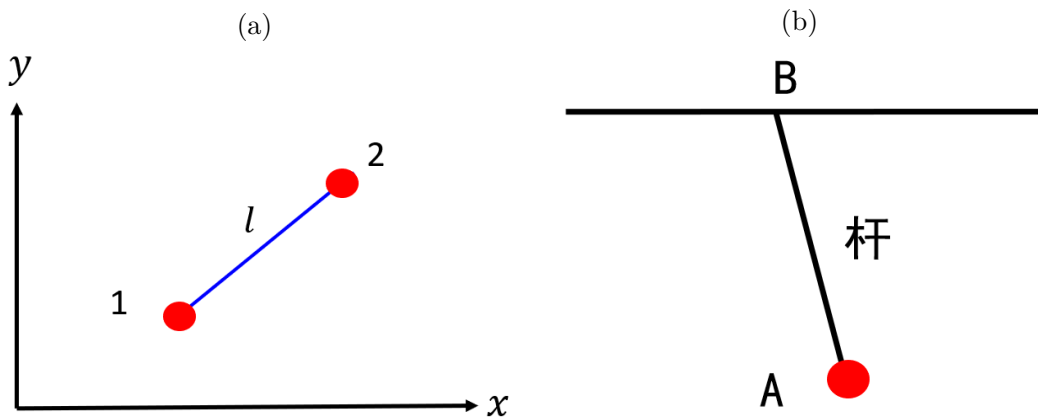


图 1.8: (a) 两个质点构成的刚体。(b) 单摆。

Example5 刚体中的内力 如图 1.8-(a) 所示, 考虑由两个质点 1,2 所构成的刚体, 间距固定为 l , 从而起约束条件为

$$f(\mathbf{r}_1, \mathbf{r}_2) = |\mathbf{r}_2 - \mathbf{r}_1| - l = 0 \quad (1.70)$$

我们取刚体的内力为 $\mathbf{F}_N = (\mathbf{F}_{21}, \mathbf{F}_{12})$, 分别为 2 给 1 的内力和 1 给 2 的内力。在刚体中, 应当有 $\mathbf{F}_{12} = -\mathbf{F}_{21}$, 并且 \mathbf{F}_{12} 和 \mathbf{F}_{21} 与两质点连线方向共线。于是, 在刚体中质点 1,2 可能的速度 $(\mathbf{w}_1, \mathbf{w}_2)$, 就有

$$\mathbf{F}_{12} \cdot \mathbf{w}_2 + \mathbf{F}_{21} \cdot \mathbf{w}_1 = \mathbf{F}_{21} \cdot (\mathbf{w}_1 - \mathbf{w}_2) = 0$$

最后一个等号的成立是因为 \mathbf{F}_{12} 与 $\mathbf{r}_1 - \mathbf{r}_2$ 共线, 而 $\mathbf{w}_1 - \mathbf{w}_2$ 总是与 $\mathbf{r}_1 - \mathbf{r}_2$ 垂直, 从而两质点构成的刚体, 它们的内力成为其约束力。这一结论可以被推广到一般的刚体。

Example5 单摆 如图 1.8-(b) 所示, 当我们选取质点 A 作为研究对象时, 如果杆的质量可以被忽略, 那么杆给 A 的力就成为约束力。但如果杆是有质量的, 杆给 A 的力就不再是约束力。但此时我们可以扩大研究对象, 将杆和 A 都作为研究对象, 则杆和 A 的相互作用力成为了刚体的内力, 从而成为约束力。值得注意的是, 当将杆和 A 都作为研究对象时, 杆在悬挂点 B 处受的力也是约束力。因为 B 点 (作为杆上的点) 位置不变, 这也是约束条件之一。因此满足约束条件的情况下 B 点可能的速度必然为 0。因此 B 点受的任何力与该“可能的速度”点乘后都为 0。所以 B 点受的任何力都是约束力。

前面我们讨论的是稳定完整约束中约束力的定义, 我们可以将其强行推广到不稳定完整约束中

定义 1.3.2 (约束力 [不稳定完整约束、矢量表述].) 设有力 $\mathbf{F}_N = (\mathbf{F}_N^{(1)}, \dots, \mathbf{F}_N^{(N)})$, 若在任意时刻, 任意满足约束条件的位置 $(\mathbf{r}_1, \dots, \mathbf{r}_N)$, 对于满足约束条件的任意可能的速度 $(\mathbf{w}_1, \dots, \mathbf{w}_N)$, 即任意一组满足

$$\sum_{j=1}^N (\nabla_j f_\alpha) \cdot \mathbf{w}_j = 0, (\alpha = 1, \dots, m) \quad (1.71)$$

的速度 $(\mathbf{w}_1, \dots, \mathbf{w}_N)$, 该位置的 $\mathbf{F}_N = (\mathbf{F}_1^{(1)}, \dots, \mathbf{F}_N^{(N)})$ 都满足

$$\sum_{j=1}^N \mathbf{F}_N^{(j)} \mathbf{w}_j = 0 \quad (1.72)$$

则称系统所受的这一个力 \mathbf{F}_N 为约束力。

值得注意的是, 此时满足 $\sum_{j=1}^N (\nabla_i f_\alpha) \cdot \mathbf{w}_j = 0 (\alpha = 1, \dots, N)$ 的速度, 不再是“满足约束条件的可能的速度”了, 因为不稳定完整约束时, 速度满足

$$\frac{d}{dt} f_\alpha[\mathbf{r}_1(t), \dots, \mathbf{r}_N(t), t] = \sum_{j=1}^N \nabla_i f_\alpha[\mathbf{r}_1(t), \dots, \mathbf{r}_N(t), t] \cdot \dot{\mathbf{r}}_j + \frac{\partial f_\alpha}{\partial t} = 0$$

同样是 Example1 中处在光滑圆环上的质点, 但此时如果圆环的圆心正在运动, 此时质点受到不稳定完整约束。满足以上定义的约束力, 必然每时每刻都和该时刻质点所在位置的圆环切线方向垂直, 但这并不和可能的速度垂直, 这是因为由于圆心的速度, 使得可能的速度不再沿着切线方向。

现在, 我们总结一下约束力的判断方法

1. 寻找“候选人”(起约束作用, 稳定完整约束时不做功(不充分))。注意, 每个“候选人”都可以写成 $(\mathbf{F}^{(1)}, \dots, \mathbf{F}^{(N)})$ 的形式
2. 带入严格定义验证
3. 如果有 $\mathbf{F} = (\mathbf{F}^{(1)}, \dots, \mathbf{F}^{(N)})$ 和 $\mathbf{F}' = (\mathbf{F}'^{(1)}, \dots, \mathbf{F}'^{(N)})$ 均为约束力, 那么 $\mathbf{F} + \mathbf{F}' = (\mathbf{F}^{(1)} + \mathbf{F}'^{(1)}, \dots, \mathbf{F}^{(N)} + \mathbf{F}'^{(N)})$ 也是约束力, 从而可以将所有是约束力的“候选人”相加, 得到“总约束力”。

接下来, 我们将约束力的定义写成分量表述的形式。此时, 各个质点的位置 $(\mathbf{r}_1, \dots, \mathbf{r}_N)$ 应当用分量表达为 (x_1, \dots, x_n) , 约束条件 $f_\alpha(\mathbf{r}_1, \dots, \mathbf{r}_N) = 0$ 应当用分量表达为

$$f_\alpha(x_1, \dots, x_n) = 0 \quad \alpha = 1, \dots, m$$

N 个质点可能的速度 $(\mathbf{w}_1, \dots, \mathbf{w}_N)$ 用分量就表达为 $(\mathbf{W}_1, \dots, \mathbf{W}_n)$, 其中 W_j 为 x_j 的“主人”可能的速度在 x_j 方向的分量, 于是可能的速度的条件 $\sum_{j=1}^N \nabla_j f_\alpha \cdot \mathbf{w}_j = 0$ 用分量就表达为

$$\sum_{i=1}^n \frac{\partial f_\alpha}{\partial x_i} W_i = 0 \quad (\alpha = 1, \dots, m) \quad (1.73)$$

而系统的一组受力 $\mathbf{F}_N = (\mathbf{F}_N^{(1)}, \dots, \mathbf{F}_N^{(N)})$ 用分量表述为

$$\mathbf{F} = (F_1, \dots, F_N)$$

其中 F_i 是 x_i 的“主人”所受的 (在 \mathbf{F} 中包含的) 力在 x_i 方向的分量 (若 \mathbf{F} 中不包含任何 x_i 的主人所受的力, 则 $F_i = 0$)。例如对于 [Example 4](#) 中固定在光滑圆环上的质点, 如果 \mathbf{F} 包括“质点 1 受到的圆环向外的压力”和“质点 2 受到的重力”, 在分量表述下

$$x_1 = x^{(1)}, x_2 = y^{(1)}, x_3 = x^{(2)}, x_4 = y^{(2)} \quad (1.74)$$

于是

$$\mathbf{F} = (F_1, F_2, F_3, F_4) \quad (1.75)$$

其中

$$F_1 = F_N \cos \theta_1 \quad F_2 = F_N \sin \theta_1 \quad F_3 = 0 \quad F_4 = -m_2 g \quad (1.76)$$

而如果 \mathbf{F} 只包括“质点 1 受到的圆环向外的压力”, 那么分量表述就有

$$F_1 = F_N \cos \theta_1 \quad F_2 = F_N \sin \theta_1 \quad F_3 = 0 \quad F_4 = 0 \quad (1.77)$$

在如上的分量表述下, 我们就能给出稳定完整约束下约束力定义的分量表述

定义 1.3.3 (约束力 (稳定完整约束, 分量表述)). 设有力 $\mathbf{F}_N = (F_{N1}, \dots, F_{Nn})$, 若在任意满足约束条件的位置 (x_1, \dots, x_n) , 对任意满足

$$\sum_{i=1}^n \frac{\partial f_\alpha}{\partial x_i} W_i = 0 (\alpha = 1, \dots, m) \quad (1.78)$$

的数组 (W_1, \dots, W_n) , 该位置的力 \mathbf{F}_N 都有

$$\sum_{i=1}^n F_i W_i = 0 \quad (1.79)$$

则称 \mathbf{F}_N 是约束力。

这一定义反映了如下物理图像: 所有的直角坐标 (x_1, \dots, x_n) 构成 n 维空间, 系统可视为 n 维空间中运动的一个点。 m 个约束 $f_\alpha(x_1, \dots, x_n) = 0$ 将系统限制在了 n 维空间中的 $n - m$ 维曲面上。此时系统可能的速度 (是一个 n 维空间的矢量) 都与该曲面相切 (都是切面上的矢量)。约束力被定义为与此曲面垂直的力 (沿 $n - m$ 维曲面“法线方向”的力, 和“所有切面上的矢量”都垂直)。

类似地, 我们可以将其定义强行推广至不稳定完整约束

定义 1.3.4 (约束力 (不稳定完整约束, 分量表述)). 设有力 $\mathbf{F}_N = (F_{N1}, \dots, F_{Nn})$, 若在任意时刻, 在任意满足约束条件的位置 (x_1, \dots, x_n) , 对任意满足

$$\sum_{i=1}^n \frac{\partial f_\alpha}{\partial x_i} W_i = 0 (\alpha = 1, \dots, m) \quad (1.80)$$

的数组 (W_1, \dots, W_n) , 该位置的力 \mathbf{F}_N 都有

$$\sum_{i=1}^n F_i W_i = 0 \quad (1.81)$$

则称 \mathbf{F}_N 是约束力。

它的图像和稳定约束时相同：所有直角坐标 (x_1, \dots, x_n) 构成 n 维空间，系统可视为 n 维空间中运动的一个点， m 个约束 $f_\alpha(x_1, \dots, x_n, t)$ 将系统限制在 n 维空间的 $n - m$ 维曲面上，这一曲面的形状和位置将随时间变化。约束力被定义为每时每刻都和瞬时的曲面垂直的力（沿着 $n - m$ 维曲面“瞬时法线方向”的力）

1.3.3 完整约束系统的 Lagrange 方程

系统、约束和受力。

我们现在推导完整约束系统的 Lagrange 方程。设系统的直角坐标为 $(x_1 \cdots x_n)$ ，有 m 个完整约束 $f_\alpha(x_1 \cdots x_n; t) = 0, \alpha = 1 \cdots m$ ，因而有 $n - m$ 个广义坐标分量 $(q_1 \cdots q_{n-m})$ 。并有坐标变换关系：

$$x_i = x_i(q_1 \cdots q_{n-m}; t) \quad (i = 1 \cdots n). \quad (1.82)$$

注意坐标变换关系和约束函数之间存在匹配关系 (1.59)，即：

$$\sum_{i=1}^n \frac{\partial f_\alpha}{\partial x_i} \frac{\partial x_i}{\partial q_j} = 0 \quad (\alpha = 1 \cdots m, j = 1 \cdots n - m). \quad (1.83)$$

我们进一步定义力

$$F = (F_1 \cdots F_n) \quad (1.84)$$

为各个质点所受合力的集合，其中 F_i 是 x_i 所对应的质点所受合力在 x_i 方向的投影。我们将合力 F 写成

$$F = F_Q + F_N, \quad (1.85)$$

这里

$$F_N \equiv (F_{N1}, \dots, F_{Nn}) \quad (1.86)$$

系统受的总约束力，而

$$F_Q = F - F_N \equiv (F_{Q1}, \dots, F_{Qn}) \quad (1.87)$$

被称为主动力。

基本思路。

我们的目标是推导广义坐标 $q_1 \cdots q_{n-m}$ 所满足的 $n - m$ 个方程，且其中不能含有约束力 F_N 。在具体推导前，先介绍一下基本的思路。

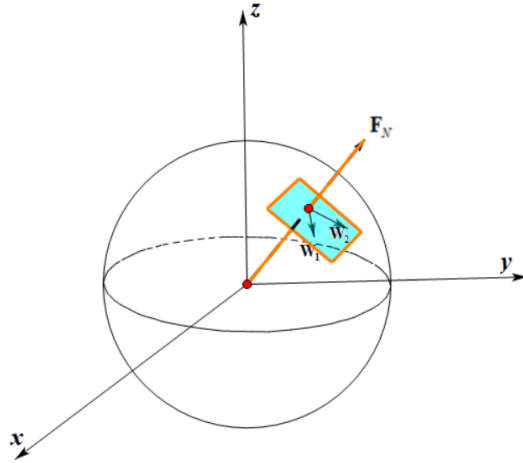


图 1.9: 球面摆

以前面讲过的轻杆上的球面摆（见上图）为例。如前述，该摆受的约束力来自杆，且和杆的方向（即球面的法向）共线，其具体大小和指向未知。为了在运动方程中消去此约束力，我们可以把 Newton 方程 $\mathbf{F} = m\mathbf{a}$ 投影到球面（即满足约束条件的轨道）的切面上。由约束力沿法线，因此其在切面的投影必然为 0。因此投影到切面的牛顿方程必然自动地不包含约束力。且投影后牛顿方程只有两个（因为切面是二维的），正好等于广义坐标的个数。

我们接下来需要做的，就是将上述特例中的投影过程推广到一般情况。

约束力的消去。

我们首先在直角坐标系中写出牛顿方程

$$m_i \ddot{x}_i = F_{Ni} + F_{Qi}, \quad (i = 1 \dots n). \quad (1.88)$$

然后定义数组 $(W_1 \dots W_n)$ 满足

$$W_i = \frac{\partial x_i}{\partial q_1}, \quad (i = 1, \dots, n). \quad (1.89)$$

显然 $(W_1 \dots W_n)$ 的取值依赖于系统的坐标 $(x_1 \dots x_n)$ 。并且根据 (1.59)， W 满足

$$\sum_{i=1}^n \frac{\partial f_\alpha}{\partial x_i} W_i = \sum_{i=1}^n \frac{\partial f_\alpha}{\partial x_i} \frac{\partial x_i}{\partial q_1} = 0. \quad (1.90)$$

因此，根据约束力的定义，我们有

$$\sum_{i=1}^n F_{Ni} W_i = 0. \quad (1.91)$$

进一步，我们把牛顿方程改写成

$$F_{Ni} + F_{Qi} - m_i \ddot{x}_i = 0 \quad (i = 1 \dots n). \quad (1.92)$$

这导致

$$\sum_{i=1}^n (F_{Ni} + F_{Qi} - m_i \ddot{x}_i) W_i = 0. \quad (1.93)$$

再利用 (1.91)，我们得到

$$\sum_{i=1}^n (F_{Qi} - m_i \ddot{x}_i) W_i = 0. \quad (1.94)$$

将 (1.89) 代入上式，即得到

$$\sum_{i=1}^n (F_{Qi} - m_i \ddot{x}_i) \frac{\partial x_i}{\partial q_1} = 0. \quad (1.95)$$

类似的，若数组 $(W_1 \cdots W_n)$ 满足

$$W_i = \frac{\partial x_i}{\partial q_2}, \quad (i = 1, \dots, n), \quad (1.96)$$

则可得到

$$\sum_{i=1}^n (F_{Qi} - m_i \ddot{x}_i) \frac{\partial x_i}{\partial q_2} = 0, \quad (1.97)$$

等等。因此，我们最终有 $n - m$ 个方程：

$$\sum_{i=1}^n (F_{Qi} - m_i \ddot{x}_i) \frac{\partial x_i}{\partial q_\alpha} = 0, \quad (\alpha = 1 \cdots n - m). \quad (1.98)$$

这 $n - m$ 个方程是广义坐标满足的方程，且不含约束力 F_N 。因此到目前位置我们已经完成了任务。

最后给出上述推导背后的图像。我们仍考虑图 1.9 中的球面摆。要将 Newton 方程投影到切平面上，需先找到切平面的两个线性无关基矢 $\mathbf{w}_1, \mathbf{w}_2$ 。则投影后的方程为 $F \cdot \mathbf{w}_1 = m\mathbf{a} \cdot \mathbf{W}_1$ 和 $F \cdot \mathbf{w}_2 = m\mathbf{a} \cdot \mathbf{W}_2$ 。在一般情况下，我们在上面所定义的每个 (W_1, \dots, W_n) (共 $n - m$ 个数组) 都是满足约束的轨道（即 n 维空间中 $n - m$ 维超曲面）的“切平面”的一个基矢。从 Newton 方程 (1.88) 推演得到方程 (1.95) 的过程，即为投影过程。

方程 (1.98) 的化简。

接下来我们化简方程 (1.98)。首先，

$$\begin{aligned} m_i \ddot{x}_i \frac{\partial x_i}{\partial q_\alpha} &= m_i \left(\frac{d\dot{x}}{dt} \right) \left(\frac{\partial x_i}{\partial q_\alpha} \right) \\ &= m_i \frac{d}{dt} \left(\dot{x}_i \frac{\partial x_i}{\partial q_\alpha} \right) - m_i \dot{x}_i \left[\frac{d}{dt} \left(\frac{\partial x_i}{\partial q_\alpha} \right) \right] \\ &= m_i \frac{d}{dt} \left(\dot{x}_i \frac{\partial \dot{x}_i}{\partial \dot{q}_\alpha} \right) - m_i \dot{x}_i \frac{\partial \dot{x}_i}{\partial q_\alpha} \\ &= \frac{d}{dt} \frac{\partial}{\partial \dot{q}_\alpha} \left(\frac{1}{2} m_i \dot{x}_i^2 \right) - \frac{\partial}{\partial q_\alpha} \left(\frac{1}{2} m_i \dot{x}_i^2 \right). \end{aligned}$$

这里我们利用了有用的公式 1 和 2。另一方面，系统的总动能为

$$T = \sum_{i=1}^n \frac{1}{2} m_i \dot{x}_i [q_1, \dots, q_{n-m}, \dot{q}_1, \dots, \dot{q}_{n-m}; t]^2.$$

因此我们有:

$$\sum_{i=1}^n m_i \ddot{x}_i \frac{\partial x_i}{\partial q_\alpha} = \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_\alpha} \right) - \frac{\partial T}{\partial q_\alpha} \quad (\alpha = 1, \dots, n-m). \quad (1.99)$$

我们进一步定义**广义力**($Q_1 \cdots Q_n$), 其中

$$Q_\alpha = \sum_{i=1}^n F_{Q_i} \frac{\partial x_i}{\partial q_\alpha} \quad (\alpha = 1, \dots, n-m).$$

因此, (1.98) 可被改写为

$$\boxed{\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_\alpha} - \frac{\partial T}{\partial q_\alpha} - Q_\alpha = 0} \quad (\alpha = 1 \cdots n-m) \quad (1.100)$$

这就是完整约束系统的 Lagrange 方程。此方程组有 $n-m$ 个方程, 与 Newton 方程等价, 且约束力 F_N 已经自动从公式中被消去。

1.3.4 保守系统的 Lagrange 方程.

一个完整约束系统, 若存在势能函数 $V(x_1 \cdots x_n; t)$, 使系统所有主动力分量都可以被写为

$$F_{Q_i} = - \frac{\partial V}{\partial x_i}. \quad (1.101)$$

则称此系统为保守系统。显然, 一个系统是否保守, 只和系统所受的主动力有关, 而与约束力无关。

若一个完整约束系统为保守系统, 则其广义力分量 Q_α ($\alpha = 1, \dots, n-m$) 可以进一步被改写为:

$$Q_\alpha = \sum_{i=1}^n F_{Q_i} \frac{\partial x_i}{\partial q_\alpha} = - \sum_{i=1}^n \frac{\partial V}{\partial x_i} \frac{\partial x_i}{\partial q_\alpha} = - \frac{\partial V[x_1(q_1 \cdots q_{n-m}; t) \cdots x_n(q_1 \cdots q_{n-m}; t); t]}{\partial q_\alpha} = - \frac{\partial V}{\partial q_\alpha}. \quad (1.102)$$

因此, 完整约束下的 Lagrange 方程 (1.100) 就可以被改写为

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_\alpha} \right) - \frac{\partial T}{\partial q_\alpha} + \frac{\partial V}{\partial q_\alpha} = 0, \quad (\alpha = 1 \cdots n-m). \quad (1.103)$$

我们进一步定义拉格朗日量为 $L = T - V$ 。于是 $\frac{\partial L}{\partial \dot{q}_\alpha} = \frac{\partial T}{\partial \dot{q}_\alpha}$ 。故 (1.100) 可以最终被改写为

$$\boxed{\frac{\partial L}{\partial \dot{q}_\alpha} - \frac{\partial L}{\partial q_\alpha} = 0, \quad (\alpha = 1 \cdots n-m).} \quad (1.104)$$

这就是**完整约束保守系统的 Lagrange 方程**。可以发现, 完整约束下的 Lagrange 方程与无约束形式下完全相同。

具体问题的处理程序。

在广义坐标下利用 Lagrange 力学求解完整约束系统的演化，一般流程如下：

(1) 明确直角坐标和广义坐标的变换关系，以及相应的速度变换关系。

(2) 明确系统受的约束力 F_N 和主动力 F_Q 。

(3) 若系统为保守系统，把拉氏量 $L = T - V$ 写成直角坐标 x_1, \dots, x_n 和直角坐标下速度分量 $\dot{x}_1, \dots, \dot{x}_n$ 的函数。之后将 (1) 中得出的变换关系带入拉氏量，最终把拉氏量表示成广义坐标 q_1, \dots, q_{n-m} 和广义坐标下速度分量 $\dot{q}_1, \dots, \dot{q}_{n-m}$ 的函数。然后列出广义坐标满足的拉格朗日方程组 (1.104)。

(4) 若系统为非保守系统，则列出拉格朗日方程组 (1.100)，并将其写成广义坐标满足的方程。

(5) 利用初始条件 $q_1, \dots, q_n(t=0)$ 和 $\dot{q}_1, \dots, \dot{q}_n(t=0)$ 求解拉格朗日方程组，得出每时每刻的广义坐标取值 $q_1, \dots, q_n(t)$ 。

1.4 Lagrange 方程的应用

1.4.1 阿特伍德机

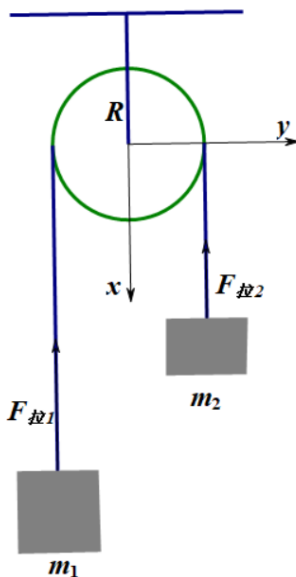


图 1.10:

阿特伍德机由一个光滑的定轻滑轮组成，设滑轮半径为 R 。设轻绳的总绳长为 $l + \pi R$ ，试求重物的加速度。设两个重物的直角坐标分别为 x_1, x_2 ，显然他们受到了一个约束：

$$f(x_1, x_2) = x_1 + x_2 - l = 0. \quad (1.105)$$

我们取广义坐标 $x = x_1$ ，则坐标变换为 $x_2 = l - x, \dot{x}_2 = -\dot{x}_1 = -\dot{x}$ 。

我们首先证明两个质点所受到的拉力为约束力。

证明. 我们记质点 1, 2 所受到的拉力分别为 \mathbf{F}_{pull1} 和 \mathbf{F}_{pull2} 。由于绳子质量忽略不计, 因此有 $\mathbf{F}_{pull1} = \mathbf{F}_{pull2} = -F\mathbf{e}_x$ 。我们进一步定义 \mathbf{F}_N 为该系受到一个力, 具体包括两个质点受的拉力, 于是利用分量表示, 有 $\mathbf{F}_N = (\mathbf{F}_{pull1} \quad \mathbf{F}_{pull2})$ 。设某一向量 (W_1, W_2) 满足

$$\frac{\partial f}{\partial x_1} W_1 + \frac{\partial f}{\partial x_2} W_2 = 0, \quad (1.106)$$

那么将有

$$W_1 + W_2 = 0. \quad (1.107)$$

显然

$$F_{pull,1} W_1 + F_{pull,2} W_2 = -F(W_1 + W_2) = 0 \quad (1.108)$$

因此 F_{pull} 是约束力。 \square

于是系统的主动力只有两个重物所受到的重力, 我们记为 $F_Q = (F_{Q1}, F_{Q2}) = (m_1 g, m_2 g)$ 。显然 F_Q 是保守力, 可以给出对应的势函数 $V = -m_1 g x_1 - m_2 g x_2$, 因而 Lagrangian 为

$$L = T - V = \frac{1}{2} m_x \dot{x}_1^2 + \frac{1}{2} m_2 \dot{x}_2^2 + m_1 g x_1 + m_2 g x_2 = \frac{1}{2} (m_1 + m_2) \dot{x}^2 + m_1 g x - m_2 g (l - x). \quad (1.109)$$

Lagrange 方程 $\frac{d}{dt} \frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0$ 为

$$(m_1 + m_2) \ddot{x} - (m_1 - m_2) g = 0. \quad (1.110)$$

因此, 我们要求的加速度为

$$\ddot{x}_1 = \ddot{x} = \frac{m_1 - m_2}{m_1 + m_2} g. \quad (1.111)$$

1.4.2 双摆

如图1.5所示, 设双摆由两根长为 l_1, l_2 的轻杆连接质量为 m_1, m_2 的两个质点, 设两个质点的直角坐标为 x_1, y_1, x_2, y_2 。

在双摆上, 两个轻杆赋予了这个系统两个约束。即质点 1 到顶点的间距必须保持不变, 数学表达为

$$x_1^2 + y_1^2 - l_1^2 = 0 \quad (1.112)$$

和质点 1, 2 之间的间距必须保持不变, 数学表达为

$$(x_1 - x_2)^2 + (y_1 - y_2)^2 - l_2^2 = 0 \quad (1.113)$$

系统四个直角坐标描述下具有两个约束，因此系统具有两个自由度，需要用两个独立的广义坐标来描述，我们用质点 1 和质点 2 各自与竖直方向的夹角 θ_1, θ_2 ，相应的坐标变换关系为

$$\begin{aligned}x_1 &= l_1 \cos \theta_1 \\y_1 &= l_1 \sin \theta_1 \\x_2 &= l_1 \cos \theta_1 + l_2 \cos \theta_2 \\y_2 &= l_1 \sin \theta_1 + l_2 \cos \theta_2\end{aligned}$$

系统的约束力是杆 1 给质点 1 的力，杆 2 给质点 1 的力和杆 2 给质点 2 的力的总和（显然后两个力可以视为刚体的内力），这一事实我们此处不再严格证明。系统的主动动力是两个质点所受的重力，这是一个保守力，因而可以赋予系统相应的势函数，从而给出保守系统的 Lagrangian 为

$$L = \frac{1}{2}m_1(\dot{x}_1^2 + \dot{y}_1^2) + \frac{1}{2}m_2(\dot{x}_2^2 + \dot{y}_2^2) + m_1gx_1 + m_2gx_2 \quad (1.114)$$

读者可自行请写出双摆系统在广义坐标下的 Lagrange 方程组。

1.5 广义动量与广义能量

1.5.1 广义动量

考虑一个保守系统，有 n 个直角坐标和 m 个约束，广义坐标为 q_1, \dots, q_{n-m} 。系统的拉氏量为 $L[q_1, \dots, q_{n-m}; \dot{q}_1, \dots, \dot{q}_{n-m}, t]$ 。定义与第 i 个广义坐标 q_i 所对应的**广义动量**（后文有时也称其为**正则动量**） p_i 为：

$$p_i \equiv \frac{\partial L}{\partial \dot{q}_i}, \quad (i = 1, \dots, n - m). \quad (1.115)$$

注意：系统的广义动量和广义坐标是一一**对应**的。即，我们可以指着某个广义动量问：“这是与哪个广义坐标对应的广义动量”？而这个问题的答案是唯一的。

根据上述定义，对于使用直角坐标的无约束保守系统，拉氏量为 $L = \sum_{i=1}^n \frac{1}{2}m_i\dot{x}_i^2 - V(x_1 \cdots x_n; t)$ ，与 x_i 对应的正则动量为

$$p_i = \frac{\partial L}{\partial \dot{x}_i} = m_i\dot{x}_i \quad (i = 1, \dots, n). \quad (1.116)$$

而这正是与 x_i 对应的牛顿力学中定义的动量（也称机械动量）分量。因此，一般系统中的正则动量，是机械动量在 Lagrange 力学的推广，在实际具体问题中，正则动量的物理意义未必是机械动量（具体表现在不具有机械动量的量纲）。

广义动量守恒。

此外，显然，保守系统 Lagrange 方程组 (1.104) 可被改写为：

$$\frac{d}{dt}p_i - \frac{\partial L}{\partial q_i} = 0, \quad (i = 1, \dots, n - m). \quad (1.117)$$

因此，如果一个系统的拉氏量不包含某个特定的广义坐标 q_l (即 $\frac{\partial L}{\partial q_l} = 0$)，则 $\frac{d}{dt}p_l = 0$ ，即与该坐标对应的广义动量 q_l 守恒，不随时间变化。

1.5.2 广义能量

仍考虑上一小节中的保守系统。该系统的广义能量 H 的定义为：

$$H \equiv \sum_{j=1}^{n-m} p_j \dot{q}_j - L. \quad (1.118)$$

我们注意到，保守系统的拉氏量为动能和势能的组合，即 $L = T - V$ ，其中势能和广义速度无关。如果系统的动能可以写成广义速度的二次型（大多数系统都是这样），即：

$$T = \sum_{j,k=1}^{n-m} g_{jk}(q_1 \cdots q_s; t) \dot{q}_j \dot{q}_k, \quad (1.119)$$

则广义动量为：

$$p_i = \frac{\partial L}{\partial \dot{q}_i} = \sum_{k=1}^{n-m} (g_{ik} + g_{ki}) \dot{q}_k; \quad (i = 1, \dots, n-m). \quad (1.120)$$

因此

$$\sum_{j=1}^{n-m} p_j \dot{q}_j = \sum_{j,k=1}^{n-m} (g_{jk} + g_{kj}) \dot{q}_k \dot{q}_j = 2T. \quad (1.121)$$

这就意味着广义能量为

$$H = 2T - L = T + V, \quad (1.122)$$

即广义能量等于机械能的表达形式。由此可见，一般情况下的广义能量是机械能概念的推广。

广义能量守恒。

在任何一次具体的运动中，系统的广义坐标、广义速度、广义动量都随时间变化，因此广义能量 H 也随时间变化。我们现在考察在一次具体运动中广义能量 H 对时间的导数：

$$\begin{aligned} \frac{dH}{dt} &= \sum_{i=1}^{n-m} \left(\dot{q}_i \frac{d}{dt} p_i + p_i \ddot{q}_i \right) - \left[\sum_{i=1}^{n-m} \left(\dot{q}_i \frac{\partial L}{\partial q_i} + \ddot{q}_i \frac{\partial L}{\partial \dot{q}_i} \right) + \frac{\partial L}{\partial t} \right] \\ &= \sum_{i=1}^{n-m} \left[\dot{q}_i \left(\frac{d}{dt} p_i - \frac{\partial L}{\partial q_i} \right) + \ddot{q}_i \left(p_i - \frac{\partial L}{\partial \dot{q}_i} \right) \right] - \frac{\partial L}{\partial t} \end{aligned} \quad (1.123)$$

根据 Lagrange 方程 (1.104) 以及广义动量的定义 (1.115)，可以最终得到

$$\frac{dH}{dt} = - \frac{\partial L}{\partial t}. \quad (1.124)$$

这意味着如果我们的 Lagrangian 形式中不显含时间，即 $\frac{\partial L}{\partial t} = 0$ ，则广义能量守恒。

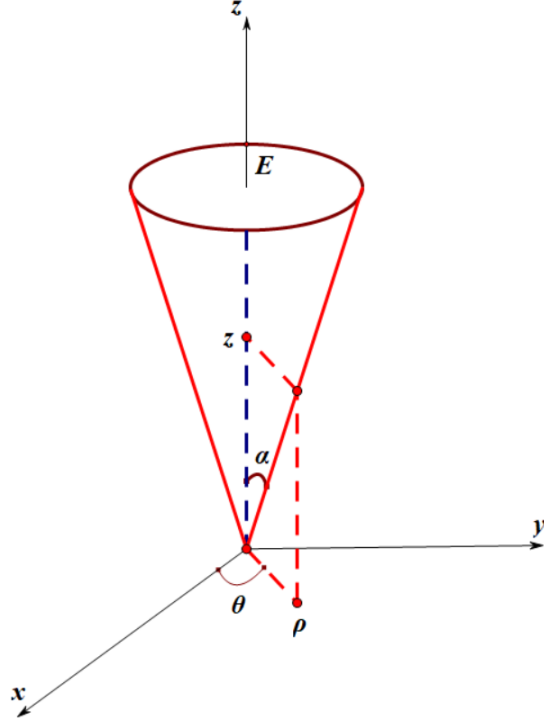


图 1.11: 在光滑圆锥面内部运动的质点

Example1 光滑锥面上的质点 如图 1.11所示, 我们考虑一个在光滑锥面内部运动的质点。设锥面的张角为 α , 质点直角坐标为 x, y, z 。显然, 质点的运动满足约束条件 $z = \sqrt{x^2 + y^2} \cot \alpha$ 。系统有三个直角坐标, 而有一个约束条件, 因而我们需要两个广义坐标。我们取广义坐标为质点在 xy 平面内投影的长度 ρ 与夹角 θ (也就是柱坐标的两个坐标)。其坐标变换关系为

$$\begin{cases} x = \rho \cos \theta \\ y = \rho \sin \theta \\ z = \rho \cot \alpha \end{cases} \quad (1.125)$$

另外, 质点系统中的约束力为锥面施加的支持力, 而主动力是重力。同样, 这是一个保守系统, 因此可以给出其 Lagrangian

$$L = T - V = \frac{m}{2}(\dot{\rho}^2 + \rho^2 \dot{\theta}^2 + \dot{z}^2) - mg\rho \cot \alpha. \quad (1.126)$$

同时我们可以求出两个广义坐标所对应的广义动量为:

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = m\rho^2 \dot{\theta}, \quad (1.127)$$

$$p_\rho = \frac{\partial L}{\partial \dot{\rho}} = m\dot{\rho}(1 + \cot^2 \alpha). \quad (1.128)$$

从 Lagrangian 的形式中, 可以知道:

1. $\frac{\partial L}{\partial \theta} = 0$, 这意味着 p_θ 是守恒量。从它的形式可以看出, 这对应沿 z 轴的角动量 L_z 守恒。

2. $\frac{\partial L}{\partial t} = 0$, 从而 H 是守恒量, 这意味着体系的机械能守恒。

1.6 与速度有关的广义势

到目前为止, 我们对保守系统的定义为: 该系统主动力 F_Q 在直角坐标下的分量 F_{Q_i} ($i = 1, \dots, n$) 都可写成 $F_{Q_i} = -\partial V(x_1, \dots, x_n, t)/\partial x_i$, 其中 $V(x_1, \dots, x_n, t)$ 是一个以直角坐标和时间为自变量的势能函数。如果一个系统是保守系统, 则可以构造该系统的拉氏量。

按照上述定义, 保守系统的主动力只能是系统中各质点位置 and 时间的函数, 而必须和质点的速度 $(\dot{x}_1, \dots, \dot{x}_n)$ 无关。这意味着如果主动力和速度有关 (例如磁场中运动的电荷, 其所受的洛伦兹力是主动力, 但和速度有关), 则该系统就不是保守系统, 无法写出拉氏量。这会给我们处理此类系统带来困难。

在本节, 我们将把保守系统的概念推广到主动力和速度相关的情况。我们将证明, 早主动力 F_Q 与速度有关时, 在某些情况下也构造该系统的拉氏量 L , 使其拉格朗日方程具有 (1.104) 式的形式。而存在洛伦兹力的系统, 恰恰就属于这些情况。因此, 利用本节的理论, 我们将可以为存在洛伦兹力的系统构造拉氏量。

1.6.1 广义势

假设体系中存在一个同时与坐标、速度和时间有关的势函数 $V(x_1, \dots, x_n; \dot{x}_1, \dots, \dot{x}_n; t)$, 使得主动力在直角坐标下的分量满足:

$$F_{Q_i} = -\frac{\partial V}{\partial x_i} + \frac{d}{dt} \left(\frac{\partial V}{\partial \dot{x}_i} \right), \quad (i = 1, \dots, n). \quad (1.129)$$

注意上式中 d/dt 是从前文所提到的角度 Π 进行的运算, 即在具体运动中对时间求全导数。而我们假设在任何一次具体的运动中, 主动力 F_{Q_i} 均满足 (1.129)。

在此情况下, 系统的广义力 Q_j ($j = 1, \dots, n - m$) 可以写成

$$\begin{aligned} Q_j &= \sum_{i=1}^n F_{Q_i} \frac{\partial x_i}{\partial q_j} \\ &= \sum_{i=1}^n \left[-\frac{\partial V}{\partial x_i} \frac{\partial x_i}{\partial q_j} + \frac{\partial x_i}{\partial q_j} \cdot \frac{d}{dt} \left(\frac{\partial V}{\partial \dot{x}_i} \right) \right] \\ &= \sum_{i=1}^n \left[-\frac{\partial V}{\partial x_i} \frac{\partial x_i}{\partial q_j} + \frac{d}{dt} \left(\frac{\partial V}{\partial \dot{x}_i} \frac{\partial x_i}{\partial q_j} \right) - \frac{\partial V}{\partial \dot{x}_i} \cdot \frac{d}{dt} \left(\frac{\partial x_i}{\partial q_j} \right) \right] \end{aligned} \quad (1.130)$$

由于 V 可写成复合函数

$$V(x_1(q_1, \dots, q_{n-m}, t), \dots, x_n(q_1, \dots, q_{n-m}, t), \dot{x}_1(q_1, \dots, q_{n-m}, \dot{q}_1, \dots, \dot{q}_{n-m}, t), \dots, \dot{x}_n(q_1, \dots, q_{n-m}, \dot{q}_1, \dots, \dot{q}_{n-m}, t); t), \quad (1.131)$$

所以我们有

$$\begin{aligned}\frac{\partial V}{\partial q_j} &= \sum_{i=1}^n \frac{\partial V}{\partial x_i} \frac{\partial x_i}{\partial q_j} + \frac{\partial V}{\partial \dot{x}_i} \frac{\partial \dot{x}_i}{\partial q_j}, \\ \frac{\partial V}{\partial \dot{q}_j} &= \sum_{i=1}^n \frac{\partial V}{\partial \dot{x}_i} \frac{\partial \dot{x}_i}{\partial \dot{q}_j}.\end{aligned}$$

因此，广义力可以被写为

$$\begin{aligned}Q_j &= \sum \left[-\frac{\partial V}{\partial x_i} \frac{\partial x_i}{\partial q_j} - \frac{\partial V}{\partial \dot{x}_i} \cdot \left(\frac{d}{dt} \frac{\partial x_i}{\partial q_j} \right) \right] + \sum \frac{d}{dt} \left(\frac{\partial V}{\partial \dot{x}_i} \frac{\partial x_i}{\partial q_j} \right) \\ &= -\sum \left(\frac{\partial V}{\partial x_i} \frac{\partial x_i}{\partial q_j} + \frac{\partial V}{\partial \dot{x}_i} \frac{\partial \dot{x}_i}{\partial q_j} \right) + \sum \frac{d}{dt} \left(\frac{\partial V}{\partial \dot{x}_i} \frac{\partial \dot{x}_i}{\partial \dot{q}_j} \right) \\ &= -\frac{\partial V}{\partial q_j} + \frac{d}{dt} \left(\frac{\partial V}{\partial \dot{q}_j} \right).\end{aligned}$$

这里我们使用了有用的公式 (a) 和 (b)。

根据上述结果，该系统的 Lagrange 方程 (1.100) 可写成

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} + \frac{\partial V}{\partial q_j} - \frac{d}{dt} \frac{\partial V}{\partial \dot{q}_j} = 0; \quad (j = 1, \dots, n - m). \quad (1.132)$$

我们仍定义该系统的拉氏量为

$$L = T - V(x_1, \dots, x_n; \dot{x}_1, \dots, \dot{x}_n; t). \quad (1.133)$$

则 Lagrange 方程 (1.132) 可以被改写为与 (1.104) 相同的形式：

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0; \quad (j = 1, \dots, n - m). \quad (1.134)$$

因此，我们可将保守系统的概念做如下推广：如果一个系统存在一个与位置、速度、时间有关的势函数 $V(x_1, \dots, x_n; \dot{x}_1, \dots, \dot{x}_n; t)$ ，使得系统的主动动力满足 (1.129)，则此系统称为保守系统，相应的势函数被称为**广义势**，系统的 Lagrangian 的形式为 $L = T - V$ 。

下面我们将证明：按照上述定义，电磁场中的带电粒子也是保守系统。

1.6.2 Lorentz 力

考虑一个带电量为 q 的点电荷在真空中的电磁场中运动。电场和磁场的场函数分别为 $\mathbf{E}(\mathbf{r}, t)$, $\mathbf{B}(\mathbf{r}, t)$ 。在国际单位制下，该电荷受到的 Lorentz 力的表达式为：

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (1.135)$$

其中 \mathbf{v} 为该质点的速度。

另一方面，真空中的 Maxwell 方程组给出：

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \quad (1.136)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (1.137)$$

两个方程分别对应 Faraday 电磁感应定律和真空中不存在磁单极子的事实。根据 (1.137)，存在矢量场 $\mathbf{A}(\mathbf{r}, t)$ ，满足

$$\mathbf{B}(\mathbf{r}, t) = \nabla \times \mathbf{A}(\mathbf{r}, t). \quad (1.138)$$

因此，方程 (1.136) 可以被改写为

$$\nabla \times \left[\mathbf{E}(\mathbf{r}, t) + \frac{\partial \mathbf{A}(\mathbf{r}, t)}{\partial t} \right] = 0. \quad (1.139)$$

这意味着存在标量场 $\phi(\mathbf{r}, t)$ ，使得 $\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} = -\nabla \phi$ ，即

$$\mathbf{E}(\mathbf{r}, t) = -\nabla \phi(\mathbf{r}, t) - \frac{\partial \mathbf{A}(\mathbf{r}, t)}{\partial t}. \quad (1.140)$$

因此，Lorentz 力可以写成

$$\mathbf{F} = q \left[-\nabla \phi - \frac{\partial \mathbf{A}}{\partial t} + \mathbf{v} \times (\nabla \times \mathbf{A}) \right]. \quad (1.141)$$

上述矢量场 $\mathbf{A}(\mathbf{r}, t)$ 和标量场 $\phi(\mathbf{r}, t)$ 分别称为电磁场的**矢量势**和**标量势**。一个电磁场，可以用 $\{\mathbf{E}(\mathbf{r}, t), \mathbf{B}(\mathbf{r}, t)\}$ 描述，也可以用 $\{\mathbf{A}(\mathbf{r}, t), \phi(\mathbf{r}, t)\}$ 描述，二者是等价的。要注意，对于给定的电磁场，函数 $\{\mathbf{E}(\mathbf{r}, t), \mathbf{B}(\mathbf{r}, t)\}$ 是唯一的，而 $\{\mathbf{A}(\mathbf{r}, t), \phi(\mathbf{r}, t)\}$ 是不唯一的。两种描述方法之间是“一对多”的关系。不过，在很多情况下，用 $\{\mathbf{A}(\mathbf{r}, t), \phi(\mathbf{r}, t)\}$ 描述更方便。而且，目前大多数物理学家认为 $\{\mathbf{A}(\mathbf{r}, t), \phi(\mathbf{r}, t)\}$ 比 $\{\mathbf{E}(\mathbf{r}, t), \mathbf{B}(\mathbf{r}, t)\}$ 更为本质。

1.6.3 Lorentz 力对应的广义势

如同之前指出的，洛伦兹力不是约束力，是主动力。本节我们构造洛伦兹力对应的，满足 (1.129) 的广义势。

单个带电粒子的情况。

首先考察电磁场 $\{\mathbf{A}(\mathbf{r}, t), \phi(\mathbf{r}, t)\}$ 中的单个带电粒子。假设粒子带电量为 q ，则该粒子受到的 Lorentz 力 (1.135) 的 x 分量为：

$$F_x = -q \frac{\partial \phi}{\partial x} - q \frac{\partial A_x}{\partial t} + q [\mathbf{v} \times (\nabla \times \mathbf{A})] \cdot \mathbf{e}_x. \quad (1.142)$$

由于

$$\begin{aligned} \mathbf{v} \times (\nabla \times \mathbf{A}) \cdot \mathbf{e}_x &= \dot{y} \frac{\partial A_y}{\partial x} + \dot{z} \frac{\partial A_z}{\partial x} - \dot{y} \frac{\partial A_x}{\partial y} - \dot{z} \frac{\partial A_x}{\partial z} \\ &= \frac{\partial}{\partial x} (\dot{y} A_y + \dot{z} A_z + \dot{x} A_x) - \left(\dot{x} \frac{\partial A_x}{\partial x} + \dot{y} \frac{\partial A_x}{\partial y} + \dot{z} \frac{\partial A_x}{\partial z} \right) \\ &= \frac{\partial}{\partial x} (\mathbf{v} \cdot \mathbf{A}) - \left(\frac{dA_x}{dt} - \frac{\partial A_x}{\partial t} \right), \end{aligned} \quad (1.143)$$

Lorentz 力的 x 分量可以被写为

$$\begin{aligned} F_x &= -\frac{\partial}{\partial x}(q\phi - q\mathbf{v} \cdot \mathbf{A}) + \frac{d}{dt}(-qA_x) \\ &= -\frac{\partial}{\partial x}(q\phi - q\mathbf{v} \cdot \mathbf{A}) + \frac{d}{dt} \left[\frac{\partial}{\partial \dot{x}}(q\phi - q\mathbf{v} \cdot \mathbf{A}) \right]. \end{aligned} \quad (1.144)$$

而 Lorentz 力的 y, z 分量也有类似的表达式。因此，可以定义与洛伦兹力对应的广义势为

$$V(\mathbf{r}, t) = q[\phi(\mathbf{r}, t) - \mathbf{v} \cdot \mathbf{A}(\mathbf{r}, t)]. \quad (1.145)$$

显然，洛伦兹力和上述广义势之间满足关系 (1.129)。因此，Lorentz 力是保守力。

N 个带电粒子的情况。

下面我们考虑一般情况。假设系统有 N 个带电质点，无约束地在电磁场中运动。我们取直角坐标为广义坐标，第 α 个质点的质量为 m_α ，直角坐标为 $x_\alpha, y_\alpha, z_\alpha$ ，带电量为 q_α 。电磁场的标势和矢势分别为 $\phi(\mathbf{r}, t), \mathbf{A}(\mathbf{r}, t)$ 。并且所有带电质点除了受到 Lorentz 力以外，还受到某些其他的保守力作用。

此时系统的拉氏量可以写成

$$L = T - V - V', \quad (1.146)$$

这里 V' 是其他保守力的势能，而 T, V 分别代表系统的动能和 Lorentz 力的广义势，表达式为：

$$T = \sum_{\alpha=1}^N \frac{1}{2} m_\alpha (\dot{x}_\alpha^2 + \dot{y}_\alpha^2 + \dot{z}_\alpha^2), \quad (1.147)$$

$$V = \sum_{\alpha=1}^N \left[q_\alpha \phi(\mathbf{r}_\alpha, t) - q_\alpha \mathbf{A}(\mathbf{r}_\alpha, t) \cdot \mathbf{v}_\alpha \right], \quad (\alpha = 1, \dots, N). \quad (1.148)$$

式 (1.148) 中每个带电质点的 Lorentz 力均对应求和中的一项， \mathbf{v}_α ($\alpha = 1, \dots, N$) 是质点 α 的速度。每一项中的 $\phi(\mathbf{r}, t), \mathbf{A}(\mathbf{r}, t)$ 函数形式都相同（都对应同一个电磁场）。该系统满足相应的 Lagrange 方程

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}_\alpha} - \frac{\partial L}{\partial x_\alpha} = 0, \quad (1.149)$$

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{y}_\alpha} - \frac{\partial L}{\partial y_\alpha} = 0, \quad (1.150)$$

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{z}_\alpha} - \frac{\partial L}{\partial z_\alpha} = 0. \quad (\alpha = 1, \dots, N). \quad (1.151)$$

此外，若 V' 与速度无关，则各质点的直角坐标所对应的正则动量将和机械动量有所不同：

$$p_{x_\alpha} = \frac{\partial L}{\partial \dot{x}_\alpha} = m_\alpha \dot{x}_\alpha + q_\alpha A_x(\mathbf{r}_\alpha, t) \neq m_\alpha \dot{x}_\alpha, \quad (1.152)$$

$$p_{y_\alpha} = \frac{\partial L}{\partial \dot{y}_\alpha} = m_\alpha \dot{y}_\alpha + q_\alpha A_y(\mathbf{r}_\alpha, t) \neq m_\alpha \dot{y}_\alpha, \quad (1.153)$$

$$p_{z_\alpha} = \frac{\partial L}{\partial \dot{z}_\alpha} = m_\alpha \dot{z}_\alpha + q_\alpha A_z(\mathbf{r}_\alpha, t) \neq m_\alpha \dot{z}_\alpha. \quad (1.154)$$

例：匀强电磁场中的单个带电粒子。

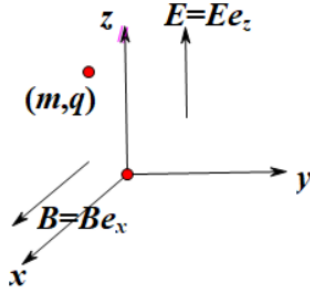


图 1.12:

如图1.12，考虑匀强电磁场中的质量为 m 的单个带电质点，忽略重力。设电场沿 z 方向，磁场沿 x 方向，即 $\mathbf{E} = E\mathbf{e}_z, \mathbf{B} = B\mathbf{e}_x$ 。可以取标势和矢势为 $\mathbf{A} = -Bz\mathbf{e}_y, \phi = -Ez$ 。以直角坐标为广义坐标，则和洛伦兹力对应的广义势为

$$\begin{aligned} V &= q(\phi - \mathbf{A} \cdot \mathbf{v}) \\ &= -qEz + qBz\dot{y}. \end{aligned} \quad (1.155)$$

因此，该系统的拉氏量可以写成：

$$\begin{aligned} L &= T - V \\ &= \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + qEz - qBz\dot{y}. \end{aligned} \quad (1.156)$$

因此，直角坐标满足的 Lagrange 方程为：

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0 \Rightarrow m\ddot{x} = 0, \quad (1.157)$$

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{y}} - \frac{\partial L}{\partial y} = 0 \Rightarrow m\ddot{y} - qB\dot{z} = 0, \quad (1.158)$$

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{z}} - \frac{\partial L}{\partial z} = 0 \Rightarrow m\ddot{z} - qE + qB\dot{y} = 0. \quad (1.159)$$

显然，它们和该质点的牛顿方程等价。

第二章 微振动

二阶线性常微分方程

2.0.1 二阶线性齐次常微分方程

A. 定义

在物理学中，大量的物理规律都会由微分方程来描述。所谓的常微分方程 (ODE) 是关于一个只有一个自变量的未知函数的方程。而所谓的二阶线性齐次常微分方程是指关于单变量未知函数 $x(t)$ 的具有如下形式的方程

$$a_2(t)\ddot{x}(t) + a_1(t)\dot{x}(t) + a_0x(t) = 0 \quad (2.1)$$

这里 $\dot{x}(t) = \frac{dx(t)}{dt}$, $\ddot{x}(t) = \frac{d^2x}{dt^2}$ 。三个系数 $a_0(t), a_1(t), a_2(t)$ 是已知函数，并且允许为常数或者某些为零 (但不能全为零，否则方程是平庸的)。“二阶”是指方程中最多包含对 $x(t)$ 的二阶导数，没有更高阶的导数。“线性”是指在方程中与 $x(t)$ 有关的项，都正比于 $x(t)$ 或其导数。“齐次”是指方程中没有与 $x(t)$ 无关的项。

B. 解的结构

现在我们介绍二阶线性 ODE 的解的结构。假设 $x_\alpha(t)$ 和 $x_\beta(t)$ 是(2.1)的两个线性无关的特解。即 $x_\alpha(t), x_\beta(t)$ 都满足(2.1)，并且对于任意的两个与 t 无关的复数 $b_\alpha, b_\beta \in \mathbb{C}$ ，都有

$$b_\alpha x_\alpha(t) + b_\beta x_\beta(t) \neq 0 \quad (2.2)$$

即上式左端不为在任意 t 取值下都为零的常函数，这被称为两个函数线性无关。可以证明，方程(2.1)的任意解一定可以写成如下形式

$$x(t) = c_\alpha x_\alpha(t) + c_\beta x_\beta(t) \quad (2.3)$$

这里 c_α, c_β 是两个与自变量无关的复数。这意味着方程的任意解都可以写成两个特解的线性组合，亦即方程的所有解构成的集合为

$$S = \{c_\alpha x_\alpha(t) + c_\beta x_\beta(t) | c_\alpha, c_\beta \in \mathbb{C}\} \quad (2.4)$$

这一结论意味着：

- 方程(2.1)的解是不唯一的。
- 对于给定的方程，特解 $x_\alpha(t), x_\beta(t)$ 有无穷多种选择。但是不同的特解所给出的解集合 S 是唯一的。

我们将任意 S 中的函数 $x(t) = c_\alpha x_\alpha(t) + c_\beta x_\beta(t)$ 代入到(2.1)的左侧，就有

$$\text{左侧} = \left\{ a_2(t) + \frac{d^2}{dt^2} a_1(t) + \frac{d}{dt} a_0(t) \right\} [c_\alpha x_\alpha(t) + c_\beta x_\beta(t)] \quad (2.5)$$

$$= c_\alpha [a_2 \ddot{x}_\alpha + a_1 \dot{x}_\alpha + a_0 x_\alpha] + c_\beta [a_2 \ddot{x}_\beta + a_1 \dot{x}_\beta + a_0 x_\beta] = 0 \quad (2.6)$$

这说明 S 中的任意元素都方程(2.1)的解。

Example1 谐振子的运动方程为

$$\ddot{x}(t) + \omega^2 x(t) = 0 \quad \omega > 0 \quad (2.7)$$

这是一个二阶线性 ODE，对照(2.1)，这相当于

$$a_2(t) = 1 \quad a_1(t) = 0 \quad a_0(t) = \omega^2 \quad (2.8)$$

并且这三个系数都与时间 t 无关，从而可以很容易解析求解。我们可以取特解

$$x_\alpha(t) = \cos \omega t \quad x_\beta(t) = \sin \omega t \quad (2.9)$$

当然，也可以取特解

$$x_\alpha(t) = e^{i\omega t} \quad x_\beta(t) = e^{-i\omega t} \quad (2.10)$$

从而解集合为

$$S = \{c_\alpha \cos \omega t + c_\beta \sin \omega t | c_{\alpha,\beta} \in \mathbb{C}\} = \{c_\alpha e^{i\omega t} + c_\beta e^{-i\omega t} | c_{\alpha,\beta} \in \mathbb{C}^G\} \quad (2.11)$$

C. 唯一解的确定

前面已经指出，只有方程不能将方程的解唯一确定下来，那么我们需要补充什么条件才能将解唯一确定下来呢？对于二阶线性齐次 ODE，同时满足原方程和两个额外的初始条件（边界条件），此时解是唯一的。如果我们要寻找同时满足微分方程形式以及两个如下初始条件的 $x(t)$

$$x(t_1) = A \quad \dot{x}(t_2) = B \quad (2.12)$$

完整的步骤是先找到方程的两个特解 $x_\alpha(t), x_\beta(t)$ ，然后得到通解

$$x(t) = c_\alpha x_\alpha(t) + c_\beta x_\beta(t) \quad (2.13)$$

将(2.13)代入到初始条件(2.12)，就有

$$c_\alpha x_\alpha(t_1) + c_\beta x_\beta(t_1) = A \quad (2.14)$$

$$c_\alpha \dot{x}_\alpha(t_2) + c_\beta \dot{x}_\beta(t_2) = B$$

于是这两个方程可以唯一地确定系数 c_α, c_β 。由于有两个系数，所以我们才需要两个初始条件。如果只有一个初始条件，只能缩小系数的取值范围，即缩小 S 集合。注意，尽管特解 $x_\alpha(t), x_\beta(t)$ 的选取是不唯一的，但是最终满足两个初始条件的解 $c_\alpha x_\alpha(t) + c_\beta x_\beta(t)$ 应当是唯一的。

设置初始条件时， t_1, t_2 既可以相等也可以不等。若 $t_1 \neq t_2$ ，则初始条件也可以设置为 $x(t_1) = A, x(t_2) = B$ 或者 $\dot{x}(t_1) = A, \dot{x}(t_2) = B$ 。即，我们可以设置两个不同时刻的 $x(t)$ 或者两个不同时刻的 $\dot{x}(t)$ ，或者同一时刻的 $x(t), \dot{x}(t)$ ，这些都可以构成唯一定解的初始条件。设置 $t_1 = t_2$ 的初始条件，对方程进行数值解会比 $t_1 \neq t_2$ 要容易。

Example2 仍然考察一维谐振子的方程(2.7)，取初始条件为

$$x(0) = A \quad \dot{x}(0) = B \quad (2.15)$$

可以证明，用特解形式(2.9)和(2.10)所给出的运动解是一致的，留作练习。

2.0.2 二阶线性非齐次常微分方程

二阶线性非齐次常微分方程形如

$$a_2(t)\ddot{x}(t) + a_1(t)\dot{x}(t) + a_0(t)x(t) = f(t) \quad (2.16)$$

它总是对应一个令 $f(t) = 0$ 的齐次方程。可以找到对应齐次方程的两个特解 $x_\alpha(t), x_\beta(t)$ 和这一非齐次方程(2.16)的特解 $x_1(t)$ ，于是非齐次方程(2.16)的通解一定可以写成

$$x(t) = C_\alpha x_\alpha(t) + C_\beta x_\beta(t) + x_1(t)$$

从而解的集合为

$$S = \{C_\alpha x_\alpha(t) + C_\beta x_\beta(t) + x_1(t) | C_{\alpha, \beta} \in \mathbb{C}\}$$

同理， $x_\alpha(t), x_\beta(t), x_1(t)$ 是不唯一的，但是 S 是唯一的。可以说明， S 中的任意一个函数都满足(2.16)，于是左侧

$$\begin{aligned} \text{左侧} &= C_\alpha [a_2(t)\ddot{x}_\alpha(t) + a_1(t)\dot{x}_\alpha(t) + a_0(t)x_\alpha(t)] \\ &\quad + C_\beta [a_2(t)\ddot{x}_\beta(t) + a_1(t)\dot{x}_\beta(t) + a_0(t)x_\beta(t)] \\ &\quad + [a_2(t)\ddot{x}_1(t) + a_1(t)\dot{x}_1(t) + a_0(t)x_1(t)] = f(t) \end{aligned}$$

同时满足非齐次方程(2.16)和两个初始条件的解是唯一的，初始条件形式和齐次时一致，于是可以将通解形式代入到初始条件 $x(t_1) = A, \dot{x}(t_2) = B$ 中，就可以将 C_α, C_β 确定下来，从而将解唯一确定下来。

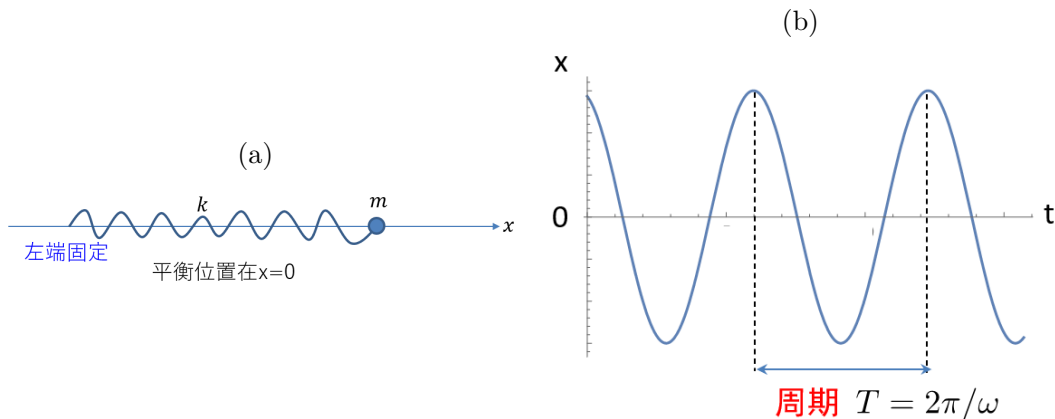


图 2.1: (A) 一个左端固定, 平衡位置在 $x = 0$ 的一自由维振子。(B) 一维自由振子的位置时刻曲线 $x(t)$

2.1 一维简谐振子

2.1.1 自由振子

如图 2.1-(a) 所示, 一个左端固定, 平衡位置在 $x = 0$ 的自由振子, 它的势能为 $V(x) = \frac{1}{2}kx^2$, 从而其拉氏量为

$$L = \frac{1}{2}m\dot{x}^2 - V(x) \quad (2.17)$$

根据前面对拉格朗日方程 (1.22) 的推导, 我们可以给出其拉氏方程为

$$m\ddot{x} = -kx \quad (2.18)$$

定义 $\omega = \sqrt{k/m}$, 于是我们有

$$\ddot{x} + \omega^2 x = 0 \quad (2.19)$$

这里 ω 为振子的固有圆频率, 它与初始条件无关, 可以选取两个特解 $\cos \omega t, \sin \omega t$, 从而振子的运动即为(2.19)的通解形式, 即

$$x(t) = a \cos \omega t + \beta \sin \omega t \quad (2.20)$$

或者写为

$$x(t) = a \cos(\omega t + \varphi) \quad (2.21)$$

这里参数 (a, φ) 由初始条件 $x(0), \dot{x}(0)$, 满足

$$\begin{cases} x(0) = a \cos \varphi \\ \dot{x}(0) = -a\omega \sin \varphi \end{cases} \quad (2.22)$$

这一解形式说明在谐振子势场 $V(x) = \frac{1}{2}kx^2$ 中的振子将做简谐运动, 其运动曲线如图 2.2-(b) 所示, 它的运动周期为 $T = \frac{2\pi}{\omega}$ 。

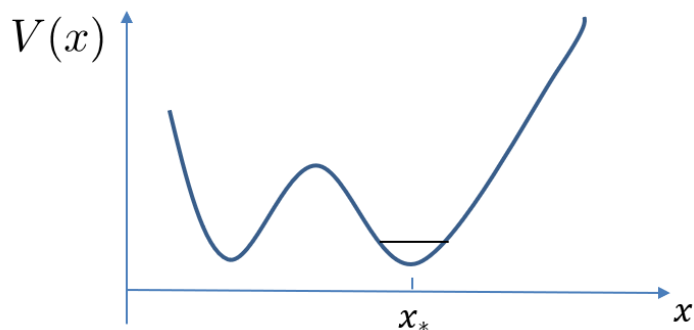


图 2.2: 任意一维质点所处势场形式 $V(x)$, 其中 x^* 是稳定平衡点

事实上谐振子模型的适用范围很广。我们考虑一个一维无约束质点, 并且假定只受一个和速度 \dot{x}, t 都无关的保守力, 从而它的拉格朗日量具有 $L = \frac{1}{2}m\dot{x}^2 - V(x)$ 的形式。如图 2.2 所示, 假定势场 $V(x)$ 在 x^* 最小, 即 x^* 是稳定平衡点。若质点初始位置在 x^* 附近, 而初始机械能仅比 $V(x^*)$ 大一点点, 则质点将在 x^* 附近小幅度振动, 从而可以将 $V(x)$ 在 x^* 附近展开为

$$V(x) = V(x^*) + V'(x^*)(x - x^*) + \frac{1}{2}V''(x^*)(x - x^*)^2 + O((x - x^*)^3) \quad (2.23)$$

x^* 作为平衡点, 将使得 $V'(x^*) = 0$ 。而由于 $V(x^*)$ 是极小点, 因此 $V''(x^*) > 0$ 。而如果假设 $x(0) - x^*$ 以及 $\dot{x}(0)$ 都足够小, 那么在 x^* 附近运动范围很小, 从而可以忽略 $O((x - x^*)^3)$ 高阶项, 以下我们假设这一条件能够满足。于是重新选择势能零点, 并且忽略 $(x - x^*)$ 的高阶项, 我们就有

$$V(x) \approx \frac{1}{2}V''(x^*)(x - x^*)^2 \quad (2.24)$$

从而拉氏量就有

$$L \approx \frac{1}{2}m\dot{x}^2 - \frac{1}{2}V''(x^*)(x - x^*)^2 \quad (2.25)$$

因此任意势场 $V(x)$ 中在平衡位置 x^* 附近运动的质点的运动, 可以认为是平衡位置在 x^* , 劲度系数为 $k = V''(x^*)$ 的振子。而质点的运动性质由 $V''(x^*)$ 决定, 与 $V(x)$ 的其他细节无关, 具有普适性。

2.1.2 受迫振子

现在假定原来的自由振子还同时受到一个时间依赖的驱动力 $F(t)\mathbf{e}_x$, 从而其 Newton 方程 (假定平衡位置在 $x = 0$) 为

$$m\ddot{x} = -kx + F(t) \quad (2.26)$$

运动方程为

$$\ddot{x} + \omega^2 x = \frac{1}{m}F(t) \quad (2.27)$$

这里圆频率仍然为 $\omega = \sqrt{k/m}$, 但是在微分方程中添加了一个非齐次项 $\frac{1}{m}F(t)$ 。

求解二阶线性非齐次常微分方程,一般要先找到对应齐次方程的两个线性无关特解 $x_\alpha(t), x_\beta(t)$, 以及整个非齐次方程的一个特解 $x_1(t)$, 于是实际解就可以被表示为

$$x(t) = \alpha x_\alpha(t) + \beta x_\beta(t) + x_1(t) \quad (2.28)$$

系数 α, β 由初始条件决定。在本例中, 可以将完整解写为 $x(t) = x_0(t) + x_1(t)$, 其中 $x_0(t) = a \cos(\omega t + \varphi)$, $x_1(t)$ 为完整非齐次方程的特解, 系数 (a, φ) 由 $x(0), \dot{x}(0)$ 决定, 满足

$$\begin{cases} a \cos \varphi + x_1(0) = x(0) \\ -a\omega \sin \varphi + \dot{x}_1(0) = \dot{x}(0) \end{cases} \quad (2.29)$$

A. 单一频率振荡的驱动力

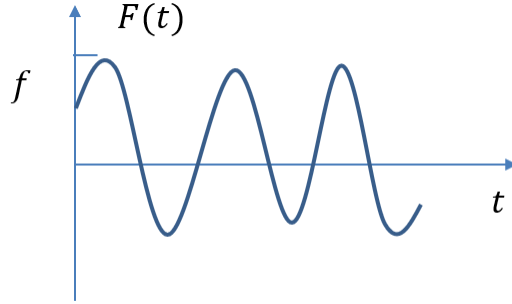


图 2.3: 单一频率振荡的驱动力形式

假设 $F(t) = f \cos(\eta t + \beta)$, 如图 2.3 所示, 并且 $\eta \neq \omega$, 则特解为

$$x_1(t) = \frac{f}{m(\omega^2 - \eta^2)} \cos(\eta t + \beta) \quad (2.30)$$

从而

$$x(t) = a \cos(\omega t + \varphi) + \frac{f}{m(\omega^2 - \eta^2)} \cos(\eta t + \beta) \quad (2.31)$$

此时的运动是圆频率为 ω, η 振动的叠加。假设没有驱动力则振子不动, 于是初始条件为 $x(0) = 0, \dot{x}(0) = 0, \beta = 0$, 从而(2.31)就有

$$x(t) = \frac{f}{m(\omega^2 - \eta^2)} [\cos(\eta t) - \cos(\omega t)] \quad (2.32)$$

可见, 在驱动力作用下, 振子做两个圆频率的振动, 振幅均正比于 $\frac{f}{m(\omega^2 - \eta^2)}$ 。即, 驱动力越强, 振幅越大; 驱动力圆频率和振子固有圆频率越接近, 振幅越大, 驱动力效果越显著, 直到 $\omega = \eta$ 的共振情况。当 $\omega = \eta$, 即当驱动力圆频率和振子固有圆频率精确相等时, Newton 方程的特解为

$$x_1(t) = \frac{f}{2m\omega} t \sin(\omega t + \beta) \quad (2.33)$$

从而振幅将随着时间一直增加。

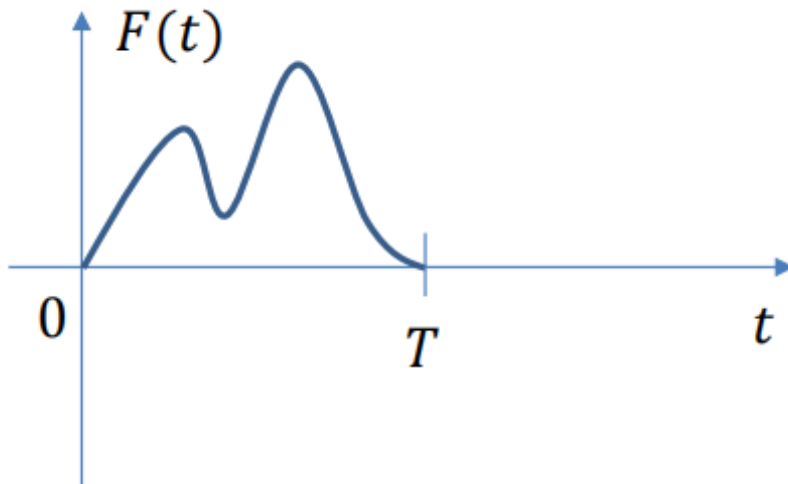


图 2.4: 满足(2.34)的某种驱动力形式

B. 只持续一段时间的驱动力

我们假设驱动力按照时间分段为

$$\begin{cases} F(t < 0) = 0 \\ F(t > T) = 0 \\ \int_{-\infty}^{+\infty} |f(t)| dt < \infty \end{cases} \quad (2.34)$$

例如图 2.4 形式, 则驱动力 $F(t)$ 的行为可以表示为不同圆频率振荡的叠加 (傅里叶变换)

$$F(t) = \frac{1}{\pi} \int_0^{\infty} d\eta |F_{\eta}| \cos(\eta t + \varphi_{\eta}) \quad (2.35)$$

假设振子固有圆频率为 ω , 平衡位置在 $x = 0$, 初始条件为 $x(0) = 0, \dot{x}(0) = 0$, 则在 $t > 0$ 时, 在 $F(t)$ 作用下振子开始运动。而 $t > T$ 时, $F(t) = 0$, 振子在这一时段的运动可以写为

$$x(t > T) = A \cos(\omega t + \Phi) \quad (2.36)$$

而 $F(t)$ 给振子的机械能为

$$\Delta E = \left. \frac{1}{2} m \dot{x}^2(t) + \frac{1}{2} k x^2(t) \right|_{t>T} = \frac{1}{2} k A^2 = \frac{1}{2m} |F_{\eta=\omega}|^2 \quad (2.37)$$

$F_{\eta=\omega}$ 为共振时的驱动力振幅。

我们总结一下这两种情况。在驱动力以单一频率振荡的情形 A 中, 不论 η 是否等于 ω , $F(t)$ 都会有效果, 并且在 η 接近 ω 时效果剧烈。而在驱动力只持续一段时间的情况 B 中, 在 $F(t)$ 施加前阵子不运动, 而在 $F(t)$ 消失以后, 将会由 $F_{\eta=\omega}$ 的分量决定 ΔE 。

2.1.3 阻尼振子

仍然考虑如图 2.1-(a) 所示的一维阵子, 但除了回复力 $-kx\mathbf{e}_x$ 以外, 还受到一个阻力 $-2m\gamma\dot{x}\mathbf{e}_x$, 于是 Newton 方程为

$$m\ddot{x} = -kx - 2m\gamma\dot{x} \quad (2.38)$$

仍然记振子的固有圆频率为 $\omega = \sqrt{k/m}$, 于是

$$\ddot{x} + \omega^2 x + 2\gamma\dot{x} = 0 \quad (2.39)$$

在 $\gamma < \omega$ 的弱阻尼情况, (2.39) 的两个特解为 $e^{-\gamma t} \cos \Omega t, e^{-\gamma t} \sin \Omega t$, 这里

$$\Omega = \sqrt{\omega^2 - \gamma^2} \quad (2.40)$$

从而(2.39)的通解可以写成

$$x(t) = \alpha e^{-\gamma t} \cos \Omega t + \beta e^{-\gamma t} \sin \Omega t \quad (2.41)$$

参数 α, β 由初始条件决定。我们也可以将(2.41)改写为

$$x(t) = a e^{-\gamma t} \cos(\Omega t + \phi) \quad (2.42)$$

并且有

$$a \cos \phi = x(0) \quad -\gamma a \cos \phi - a\Omega \sin \phi = \dot{x}(0) \quad (2.43)$$

它的通解图像如图 2.5 所示。可以看出, 阻尼振子在弱阻尼情形下以圆频率 Ω 振动, 而振幅以正比于 $e^{-\gamma t}$ 的模式衰减。在 $\omega \gg \gamma$ 时, 振动圆频率就有 $\Omega \approx \omega$ 。在长时极限 $t \rightarrow \infty$ 下, 有 $x(t) = 0$, 这与 $x(0), \dot{x}(0)$ 无关, 阻尼将机械能带走了。我们可以认为振动寿命为 $\tau = \frac{1}{\gamma}$, 当 $t \gg \gamma$ 时, $|x(t)| \ll |a|$, 通常就可以认为此时已经满足 $t \rightarrow \infty$ 。

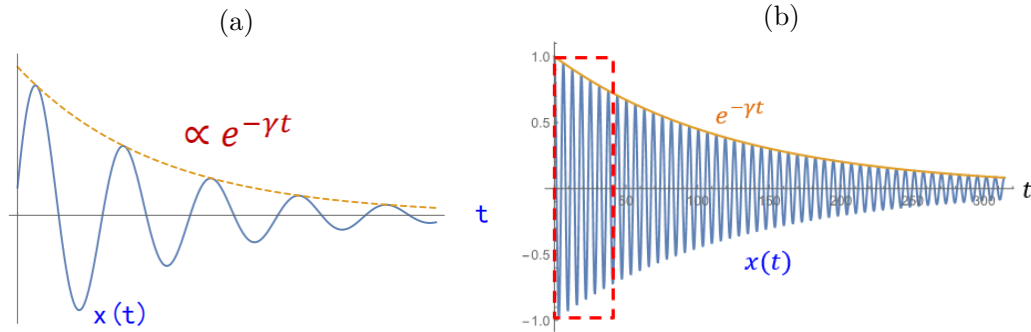


图 2.5: (a) 弱阻尼情形下阻尼振子的运动图像。(b) 当 $\omega \gg \gamma$ 时的弱阻尼振子运动图像

我们可以定义所谓的品质因子 Q 为振子的寿命和固有周期之比再乘以 π , 从而

$$Q = \pi \frac{\tau}{2\pi/\omega} = \frac{\omega}{2\gamma} \quad (2.44)$$

它有如下物理意义：其一，它表示 ω 是 γ 的多少倍；其二，它除以 π 以后，即为在振子寿命以内，如果无阻尼时可以振动的次数；其三，如果 $Q \gg 1$ ，即阻力很弱时，有 $\Omega \approx \omega$ ，从而 Q 与 π 的比值为在振动寿命内，振子的实际振动次数。

当 $t \ll \tau = \frac{1}{\gamma}$ 时，可以直接忽略阻尼，把系统近似为无阻尼自由振子。另一方面，如果 $\omega \gg \gamma$ ，也就是当品质因子 $Q \gg 1$ 时，那么存在一个满足 $\frac{2\pi}{\omega} \ll t \ll \tau$ 的时刻 t ，在 0 到 t 之间，系统可以近似为无阻尼自由振子，并且振荡了多个周期。

另一方面，我们可以这样理解 $x(t \rightarrow \infty) = 0$ ，它反映了阻尼的如下两种作用：其一，阻尼会带走机械能；其二，阻力会抹去对“历史信息”（即初始条件）的记忆，从而才会使得 $x(t \rightarrow \infty)$ 与 $x(0), \dot{x}(0)$ 无关。这两种作用来自于振子和运动表面大量做无规则运动的微观粒子耦合，这种耦合：其一，会产生摩擦力，带走能量；其二，产生大量随机涨落，从而抹去对“历史信息”的记忆，使得 $x(t \rightarrow \infty)$ 和 $x(0), \dot{x}(0)$ 无关。

临界阻尼情形和过阻尼情形 前面我们只讨论了 $\gamma < \omega$ 的情形。当 $\omega = \gamma$ 时，称之为临界阻尼情形，此时的振子运动方程通解为

$$x(t) = (C_1 + C_2 t)e^{-\gamma t} \quad (2.45)$$

当 $\gamma > \omega$ 时，称之为过阻尼情形，此时振子运动方程通解为

$$x(t) = C_1 \exp\left(-\left[\gamma - \sqrt{\gamma^2 - \omega^2}\right]t\right) + C_2 \exp\left(-\left[\gamma + \sqrt{\gamma^2 - \omega^2}\right]t\right) \quad (2.46)$$

在这两种情形下，振子均不再振动，并且与平衡位置的距离随时间衰减至零。临界阻尼情形的衰减要比过阻尼情形更快。

2.1.4 受迫阻尼振子

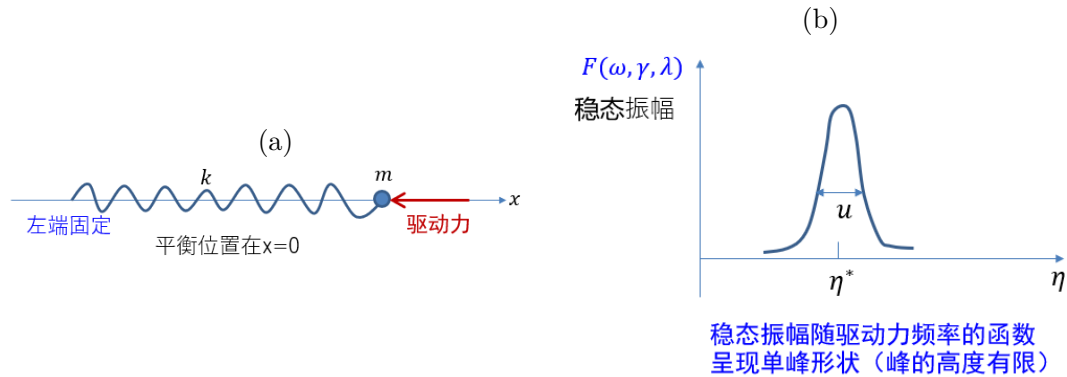


图 2.6: (a) 一个左端固定，受到驱动力的受迫阻尼振子。(b)

如图 2.6-(a) 所示，考虑一个左端固定，平衡位置为 $x = 0$ ，质量为 m 的限制在 x 方向上的一维振子，它同时收到弹簧的回复力 $-kx\mathbf{e}_x$ ，摩擦力 $-m\gamma\dot{x}\mathbf{e}_x$ 以及一个驱动力 $f \cos(\eta t)\mathbf{e}_x$ ，此时 Newton 方程为

$$\ddot{x} + \omega^2 x + 2\gamma\dot{x} = \frac{f}{m} \cos(\eta t) \quad (2.47)$$

$\omega = \sqrt{k/m}$ 仍然是固有圆频率。在本节中，我们只考虑弱阻尼 $\gamma < \omega$ 情形，此时可以给出对应齐次方程的通解形式

$$x_0(t) = ae^{-\gamma t} \cos(\Omega t + \varphi) \quad (2.48)$$

其中 (a, φ) 由初始条件 $x(0), \dot{x}(0)$ 决定，并有

$$\Omega = \sqrt{\omega^2 - \gamma^2}$$

同时可以找到非齐次方程的一个特解

$$x_1(t) = \frac{f}{m\sqrt{(\omega^2 - \gamma^2)^2 + 4\eta^2\gamma^2}} \cos(\eta t + \delta) \quad (2.49)$$

这里 δ 为

$$\delta = \arg(\eta^2 - \omega^2 - 2i\eta\gamma) \quad (2.50)$$

即是复变量 $\eta^2 - \omega^2 - 2i\eta\gamma$ 的幅角，从而最终的通解形式为

$$x(t) = x_0(t) + x_1(t) \quad (2.51)$$

可以看出，受迫阻尼振子的运动是圆频率为 Ω 的振幅衰减震动与频率为 ω 的振动的叠加。我们记

$$F(\omega, \eta, \gamma) = \frac{f}{m\sqrt{(\omega^2 - \eta^2)^2 + 4\eta^2\gamma^2}} \quad (2.52)$$

从而运动方程可以记为

$$x(t) = ae^{-\gamma t} \cos(\Omega t + \varphi) + F(\omega, \eta, \gamma) \cos(\eta t + \varphi) \quad (2.53)$$

当 $t \ll 1/\gamma$ 并且 $|F(\omega, \eta, \gamma)| \ll |a|$ 时，方程退化为

$$x(t) \approx ae^{-\gamma t} \cos(\Omega t + \varphi) \quad (2.54)$$

可以忽略阻尼，从而将系统近似为无阻尼受迫振子。当 $t \gg 1/\gamma$ 时，我们就有

$$x(t) \approx x(t \rightarrow \infty) = F(\omega, \eta, \gamma) \cos(\eta t + \delta) \quad (2.55)$$

于是系统进入稳态，能量的输入与输出平衡，振子将以稳定振幅和驱动力圆频率进行振荡。其中，稳态振幅即为 $F(\omega, \eta, \gamma)$ ，它随着驱动力频率 η 的变化如图 2.6-(b) 所示，呈现有限高度的单峰形状。当 $\gamma \ll \omega$ ，即品质因子 $Q \gg 1$ 时，峰值频率 $\eta^* \approx \omega$ ，达成共振，峰宽 u 和 γ 接近。实验上，通过测量 $F(\omega, \eta, \gamma)$ ，我们就可以从 η 推出 ω ，或者反过来从 ω 推出 η 。而 $\delta = \arg(\eta^2 - \omega^2 - 2i\eta\gamma)$ 反映振子稳态运动和驱动力的相位差，当 $\eta \gg \omega$ 时 $\delta \approx 0$ ，当 $\eta \ll \omega$ 时 $\delta \approx -\pi$ ，当 $\eta \approx \omega$ 时 $\delta \approx -\pi/2$ 。最后，在长时极限 $t \rightarrow \infty$ 下，运动形式 $x(t \rightarrow \infty)$ 与 $x(0), \dot{x}(0)$ 无关，这意味着摩擦力带走了“历史记忆”。这是由于摩擦力来自振子与大量微观粒子耦合。可以看出，无规则混乱运动的大自由度系统对于小系统有两方面影响：其一，会产生摩擦力，带走能量；其二，会产生大量随机涨落，抹去记忆。

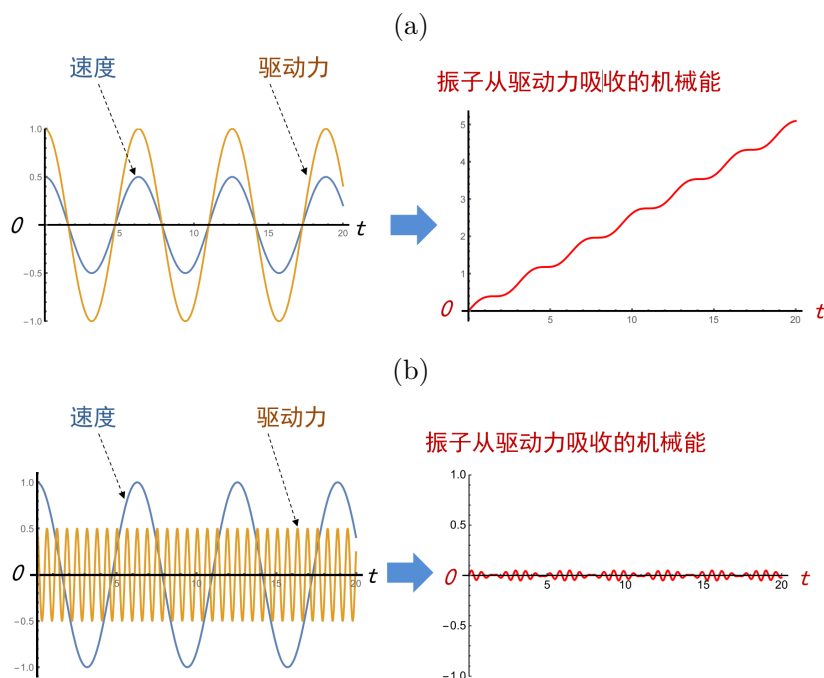


图 2.7: (a) 当驱动力频率和振子的固有圆频率接近时的驱动力、速度变化曲线与机械能随时间变化。(b) 当驱动力频率和振子的固有圆频率差异很大时的驱动力、速度变化曲线与机械能随时间变化

2.1.5 共振现象总结

我们再次审视受迫阻尼振子的“共振”。当受迫振子的驱动力圆频率和振子固有圆频率接近时，驱动力和振子的速度方向将会同步变化，从而外力方向将和振子的速度方向始终相同，从而驱动力总是做正功，振子的机械能将持续增加，如图 2.7-(a) 所示。但当受迫振子驱动力圆频率和固有圆频率差异很大时，驱动力和振子的步调不一致，从而驱动力和振子速度一会儿相同，一会儿相反，驱动力将一会儿做正功，一会儿做负功，相互抵消，从而振子机械能几乎不增加，如图 2.7-(b) 所示。

这意味着，当外力的周期和振子的固有圆频率接近时，外力的效果最明显。此时振子将在外力作用下剧烈振荡。当剧烈程度超过振子的可承受范围时，振子就会被破坏（桥梁垮塌、杯子破碎等等）。而如果外力周期和振子的固有圆频率差异很大（不论是远大于还是远小于），外力效果都不明显，此时振子基本不受外力影响，振荡幅度很小，或者可以忽略。基于这一原理，共振现象可以作为测量周期（频率）的手段。如果振子的周期是已知的，那么我们可以通过振荡幅度来反推出驱动力周期（频率），反之亦然。大量的测量手段，都是基于这样的共振原理。

生物体对于声音的辨别，也是通过共振原理实现的。所谓的“声音”是空气的振动，“声音的大小”即为振动的振幅，“音高”即为振动的频率，我们常说的“音感”即为对振动频率的分辨本领。在听声音时，生物体耳结构中基底膜的肌肉弦作为“振子”，而空气中的声波将充当“驱动力”。基底膜不同位置的组织长短厚度不同，相当于一系列固有频率不同的弦（振子），而特定频率的声音会给这些弦施加该频率的外力，从而让特定位置的弦（即固有频率和声音频率接近的弦）剧烈振

荡，激发特定的神经，从而使得大脑可以分辨声音的频率。

2.1.6 参数振荡

有些时候，振子的频率是可变的，例如对于单摆，如果摆长可以发生变化，那么就相当于单摆的频率在发生振荡。同样的例子还有秋千，如果人在秋千上上下下运动，这也相当于秋千作为振子的频率在发生变化。参数振荡的振子，Newton 方程为

$$\ddot{x} + \omega(t)x^2 = 0 \quad (2.56)$$

注意它和受迫振动不同，此时运动方程仍然是齐次的。系统的拉格朗日量为

$$L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}m\omega^2(t)x^2 \quad (2.57)$$

我们只考虑一个特殊情况，如果频率在 ω_0 附近小幅度地周期振荡，于是

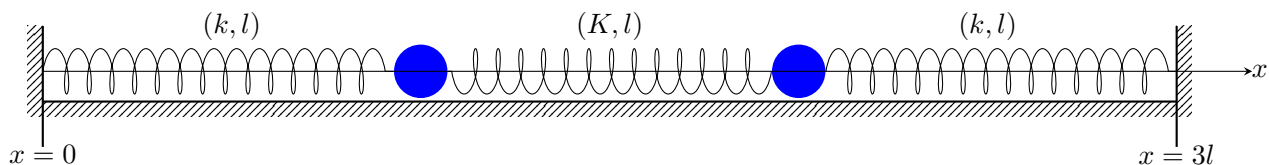
$$\omega^2 \approx \omega_0^2[1 + h \cos \zeta t] \quad (2.58)$$

这里 $h \ll 1$ 是一个小量，当 ζ 接近 $2\omega_0$ 时，频率的振荡效果最为明显，称之为**参数共振现象**，对应的运动方程解为

$$x(t) \approx ae^{hst/4} \cos\left(\frac{1}{2}\zeta t + \varphi\right) \quad (2.59)$$

这里 $s > 0$ 且 s 和 ω_0 接近。可以看出，振动的振幅将被指数放大。¹

2.2 两个耦合的同频率振子



系统和拉格朗日方程。

如上图所示，考虑两个质量为 m 的质点，在 x 轴上运动，按照图中方式用三根弹簧连接。各个弹簧的原长、弹性系数等数据已经标在图中。显然如果没有中间弹簧，那么两个质点将成为独立的两个同频振子。中间弹簧的存在，使得两个振子被耦合起来。

设系统的直角坐标为 x_1, x_2 ，分别为两个质点的坐标。两个质点的稳定平衡位置为 $x_1 = l, x_2 = 2l$ ，我们选取广义坐标为两个质点对平衡位置的偏离：

$$q_1 = x_1 - l, \quad (2.60)$$

$$q_2 = x_2 - 2l. \quad (2.61)$$

¹详细的讨论参加 Landau 《力学》 第五章

该系统拉氏量为：

$$\begin{aligned}
L &= T - V \\
&= \frac{m}{2}(\dot{x}_1^2 + \dot{x}_2^2) - \frac{1}{2}k[(x_1 - l)^2 + (3l - x_2 - l)^2] - \frac{1}{2}K(x_2 - x_1 - l)^2, \\
&= \frac{m}{2}(\dot{q}_1^2 + \dot{q}_2^2) - \frac{1}{2}k(q_1^2 + q_2^2) - \frac{1}{2}K(q_2 - q_1)^2.
\end{aligned} \tag{2.62}$$

因此，Lagrange 方程为：

$$\begin{cases} m\ddot{q}_1 = -kq_1 - K(q_1 - q_2) \\ m\ddot{q}_2 = -kq_2 - K(q_2 - q_1) \end{cases}. \tag{2.63}$$

拉氏量的解耦。

在拉式方程组 (2.63) 中，两个方程相互耦合。为了求解，我们尝试选取一组新的广义坐标 η_1, η_2 ，使得 Lagrangian 具有形式

$$L = L_1(\eta_1, \dot{\eta}_1) + L_2(\eta_2, \dot{\eta}_2). \tag{2.64}$$

若做到这一点，则称 Lagrangian 被解耦。此时 Lagrange 方程为：

$$\frac{d}{dt} \frac{\partial L_1}{\partial \dot{\eta}_1} - \frac{\partial L_1}{\partial \eta_1} = 0, \tag{2.65}$$

$$\frac{d}{dt} \frac{\partial L_2}{\partial \dot{\eta}_2} - \frac{\partial L_2}{\partial \eta_2} = 0. \tag{2.66}$$

即： η_1, η_2 所满足的方程是相互独立的。因此该拉格朗日方程可以被较方便的求解。

我们接下来的目标是寻找使拉氏量 (2.118) 解耦的广义坐标 η_1, η_2 。很容易的发现，拉氏量 (2.118) 只依赖于以下三个量：

$$q_1^2 + q_2^2 = \frac{1}{2}[(q_1 + q_2)^2 + (q_1 - q_2)^2]; \tag{2.67}$$

$$\dot{q}_1^2 + \dot{q}_2^2 = \frac{1}{2}[(\dot{q}_1 + \dot{q}_2)^2 + (\dot{q}_1 - \dot{q}_2)^2]; \tag{2.68}$$

$$(q_1 - q_2), \tag{2.69}$$

即：只依赖于 $(q_1 + q_2)$ 和 $(q_1 - q_2)$ ，以及相应的速度。因此，我们可以选取广义坐标为

$$\eta_1 = \frac{1}{\sqrt{2}}(q_1 + q_2), \tag{2.70}$$

$$\eta_2 = \frac{1}{\sqrt{2}}(q_1 - q_2). \tag{2.71}$$

显然，它们分别正比于质心坐标和相对坐标。我们很容易求出 $q_{1,2}$ 与 $\eta_{1,2}$ 之间的坐标变换和速度变换：

$$q_{1,2} = \frac{1}{\sqrt{2}}(\eta_1 \pm \eta_2), \quad \dot{q}_{1,2} = \frac{1}{\sqrt{2}}(\dot{\eta}_1 \pm \dot{\eta}_2). \tag{2.72}$$

利用上述变换，很容易推导出以 $\eta_{1,2}$ 为广义坐标时的拉氏量：

$$L = \frac{m}{2}(\dot{\eta}_1^2 + \dot{\eta}_2^2) - \frac{1}{2}\lambda_1\eta_1^2 - \frac{1}{2}\lambda_2\eta_2^2, \quad (2.73)$$

这里

$$\lambda_1 = k, \quad (2.74)$$

$$\lambda_2 = k + 2K. \quad (2.75)$$

显然，此拉氏量具有 (2.64) 的形式，因此 (2.70,2.71) 中定义的 $\eta_{1,2}$ ，拉氏量被解耦。

耦合振子的动力学。

利用拉氏量 (2.119)，我们可以得到新的广义坐标 $\eta_{1,2}$ 满足的拉格朗日方程：

$$\begin{aligned} m\ddot{\eta}_1 &= -\lambda_1\eta_1, \\ m\ddot{\eta}_2 &= -\lambda_2\eta_2. \end{aligned} \quad (2.76)$$

这正是两个质量为 m ，弹簧弹性系数分别为 $\lambda_{1,2}$ 的弹簧振子的牛顿方程。因此 η_1 和 η_2 的行为像是两个圆频率不同的独立的振子。方程 (2.76) 的解为：

$$\begin{aligned} \eta_1(t) &= A_1 \cos(\omega_1 t + \phi_1) \\ \eta_2(t) &= A_2 \cos(\omega_2 t + \phi_2). \end{aligned} \quad (2.77)$$

这里圆频率 $\omega_{1,2}$ 为

$$\omega_{1,2} = \sqrt{\lambda_{1,2}/m}, \quad (2.78)$$

显然有 $\omega_1 \neq \omega_2$ 。而系数 $\{A_j, \phi_j\}$ ($j = 1, 2$) 由初始条件 $\{\eta_j(0), \dot{\eta}_j(0)\}$ ($j = 1, 2$) 决定，满足

$$\begin{aligned} \eta_j(0) &= A_j \cos \phi_j, \\ \dot{\eta}_j(0) &= -A_j \sin \phi_j, \end{aligned} \quad (j = 1, 2), \quad (2.79)$$

也即由 $\{q_{1,2}(0), \dot{q}_{1,2}(0)\}$ 决定。

因此，对于实际问题，可通过如下步骤由 $\{q_{1,2}(0), \dot{q}_{1,2}(0)\}$ 求出 $q_{1,2}(t)$ ：

(1) 先利用 (2.70,2.71) 由 $\{q_{1,2}(0), \dot{q}_{1,2}(0)\}$ 求出 $\{\eta_{1,2}(0), \dot{\eta}_{1,2}(0)\}$ ，进而再利用 (2.79) 求出系数 $\{A_j, \phi_j\}$ ($j = 1, 2$)。

(2) 将系数 $\{A_j, \phi_j\}$ ($j = 1, 2$) 代入 (2.77)，得到 $\eta_j(t) = A_j \cos(\omega_j t + \phi_j)$ ($j = 1, 2$)。

(3) 将上述结果代入坐标变换 (2.72)，得到 $q_{1,2}(t)$ ：

$$q_1(t) = \frac{1}{\sqrt{2}}[\eta_1(t) + \eta_2(t)] = \frac{1}{\sqrt{2}}[A_1 \cos(\omega_1 t + \phi_1) + A_2 \cos(\omega_2 t + \phi_2)], \quad (2.80)$$

$$q_2(t) = \frac{1}{\sqrt{2}}[\eta_1(t) - \eta_2(t)] = \frac{1}{\sqrt{2}}[A_1 \cos(\omega_1 t + \phi_1) - A_2 \cos(\omega_2 t + \phi_2)]. \quad (2.81)$$

因此，一般情况下两个直观的坐标 $q_1(t)$ 与 $q_2(t)$ 的运动，各自都包含两个不同圆频率的振动。

简正坐标、简正（圆）频率、简正模式。

能使拉氏量解耦的广义坐标称为**简正坐标**。在这个例子中，直观的坐标是 q_1, q_2 ，而简正坐标 $\eta_{1,2} = \frac{1}{\sqrt{2}}(q_1 \pm q_2)$ 是直观坐标的线性组合，其行为与圆频率为 ω_1, ω_2 的两个独立振子相同。即，虽然两个直观坐标 $q_{1,2}$ 相互耦合，但它们特定的线性组合的行为，就如同一组无耦合的独立振子了。

进一步，简正坐标 $\eta_{1,2}$ 的圆频率 ω_1, ω_2 被称为**简正圆频率**，它完全由系统的性质所决定，而与系统的初始条件无关。

简正模式值得是一类特殊的集体运动，在运动中只有一个简正坐标在随时间演化，其他的简正坐标保持不变。我们的系统有两个简正模式：

简正模式 1: $\eta_1(t) = A_1 \cos(\omega_1 t + \phi_1), \eta_2(t) = 0$ ，即 $q_1(t) = q_2(t) = \frac{1}{\sqrt{2}} A_1 \cos(\omega_1 t + \phi_1)$ 。此时两个振子同频同幅同步调振动。

简正模式 2: $\eta_1(t) = 0, \eta_2(t) = A_2 \cos(\omega_2 t + \phi_2)$ ，即 $q_1(t) = -q_2(t) = \frac{1}{\sqrt{2}} A_2 \cos(\omega_2 t + \phi_2)$ 。此时两个振子同频同幅反步调振动。

显然，只有在特殊的初始条件下，系统才以某个特定的简正模式运动。一般情况下，系统的运动是两种简正模式的叠加（同时包含两种简正模式），而每个简正坐标都随时间变化。原始的广义坐标的运动 $q_1(t), q_2(t)$ 都包含两个频率振荡。

为了更清楚的理解，可以引入运动的矢量描述。定义构型矢量

$$\mathbf{q}(t) = \begin{pmatrix} q_1(t) \\ q_2(t) \end{pmatrix} \quad (2.82)$$

来描述系统 t 时刻的直观广义坐标。引入两个特殊的构型

$$\mathbf{s}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \mathbf{s}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad (2.83)$$

根据 $\eta_{1,2}(t)$ 的定义(2.70)(2.71)，容易证明

$$\mathbf{q}(t) = \eta_1(t) \mathbf{s}_1 + \eta_2(t) \mathbf{s}_2 \quad (2.84)$$

因此 t 时刻的构型可以分解为 \mathbf{s}_1 和 \mathbf{s}_2 的线性组合，组合系数即为简正坐标。从而简正坐标 1 可以写为

$$\mathbf{q}(t) = \eta_1(t) \mathbf{s}_1 = A_1 \cos(\omega t + \phi_1) \mathbf{s}_1 \quad (2.85)$$

由于 $\mathbf{q}(t) = \begin{pmatrix} q_1(t) \\ q_2(t) \end{pmatrix} \propto \mathbf{s}_1 \propto \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ，因此就有 $q_1(t) = q_2(t)$ 。简正模式 2 可以写为

$$\mathbf{q}(t) = \eta_2(t) \mathbf{s}_2 = A_2 \cos(\omega_2 t + \phi_2) \mathbf{s}_2$$

由于 $\mathbf{q}(t) = \begin{pmatrix} q_1(t) \\ q_2(t) \end{pmatrix} \propto \mathbf{s}_2 \propto \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ，所以 $q_1(t) = -q_2(t)$ 。从而在每个简正模式中， q_1, q_2 的比值固定，不随时间改变，总的幅度随时间以单一简正频率简谐振荡。一般情况的运动(2.84)就是两种简正模式的线性叠加（矢量叠加）。

以上的讨论说明，对于双耦合同频振子系统，我们有两种视角来看待。

- 直观的视角：它认为系统的运动即为振子 1 的运动和振子 2 的运动之和，从而关注直观的广义坐标 $q_1(t), q_2(t)$ 。这一视角的优点是直观，但由于 $q_1(t), q_2(t)$ 并不独立，因此难以从 $\{q_{1,2}(0), \dot{q}_{1,2}(t)\}$ “推想”未来的运动 $q_1(t), q_2(t)$ 。
- 基于简正坐标和简正模式的视角，从而 $\mathbf{q}(t) = \eta_1(t)\mathbf{s}_1 + \eta_2(t)\mathbf{s}_2$ 。此时，系统每时每刻将所有直观广义坐标 $q_1(t), q_2(t)$ 视为一个整体 (矢量 $\mathbf{q}(t)$)，进一步将 $\mathbf{q}(t)$ 视为两个特殊构型 $\mathbf{s}_1, \mathbf{s}_2$ 的组合，组合系数为简正坐标 $\eta_{1,2}(t)$ 。在运动中，构型 $\mathbf{s}_{1,2}$ 不随时间演化，而组合系数 $\eta_{1,2}(t)$ 随时间变化。当关注 $\eta_{1,2}(t)$ 时，它们各自相当于一个独立振子，因此系统行为与一组独立振子类似。也可以说，我们把系统整体的运动视为两个简正模式 $\eta_1(t)\mathbf{s}_1, \eta_2(t)\mathbf{s}_2$ 的叠加，关注每个简正模式的幅度和位相。这种视角下，两个简正坐标相互独立，运动规律简单 (相当于独立振子)。容易从 $\{\eta_{1,2}(0), \dot{\eta}_{1,2}(0)\}$ “推想” $\{\eta_{1,2}(t)\}$ 。不过由于简正坐标反映特殊的集体运动振幅，因此如果关心直观坐标 $q_1(t), q_2(t)$ 各自随时间的变化，则需要进行坐标变换。

至此，我们所讨论的仍然是只有两个振子的耦合，并且我们对于简正坐标的选取主要靠猜。我们期望能找到一个能够一般性的方法，来找到任意个振子的耦合下的简正坐标。这是我们后面几节的主要任务。

2.3 通过合同变换对角化实对称矩阵

为了解决任意振子耦合下的简正坐标选取的一般方法，我们将在本节首先介绍矩阵联合对角化的方法，再在下一节介绍其具体的应用实现。

2.3.1 单个实对称矩阵的对角化

一个 $N \times N$ 矩阵

$$W = \begin{pmatrix} W_{11} & \cdots & W_{1N} \\ \vdots & \ddots & \vdots \\ W_{N1} & \cdots & W_{NN} \end{pmatrix} \quad (2.86)$$

如果 $W_{ij} \in \mathbb{R}$ ，并且有 $W_{ij} = W_{ji}$ ，则称矩阵 W 为实对称矩阵。设有数 λ 和一个 N 维矢量 $\mathbf{s} = \begin{pmatrix} s_1 \\ \vdots \\ s_N \end{pmatrix}$ ，若满足

$$W\mathbf{s} = \lambda\mathbf{s} \quad (2.87)$$

则称 λ 是矩阵 W 的本征值， \mathbf{s} 为相应的本征向量。从上式可以看出，本征值 λ 满足

$$\det(W - \lambda I) = 0 \quad \text{也可以写为 } \|W - \lambda I\| = 0 \quad (2.88)$$

这里 $I = \begin{pmatrix} 1 & & \\ & 1 & \\ & & \ddots \\ & & & 1 \end{pmatrix}$ 是 $N \times N$ 单位矩阵。从而 $N \times N$ 矩阵必然有 N 个本征值 $\lambda_1, \dots, \lambda_N$ (其中可能有几个相等), 并且由于 W 是实对称的, 因此所有本征值都是实数。在得到各个本征值 $\lambda_1, \dots, \lambda_N$ 以后, 通过求解方程

$$W \mathbf{s}^{(i)} = \lambda_i \mathbf{s}^{(i)} \quad (2.89)$$

可以得到和本征值 λ_i 对应的本征向量

$$\mathbf{s}^{(i)} = \begin{pmatrix} s_1^{(i)} \\ \vdots \\ s_N^{(i)} \end{pmatrix}$$

注意, $\mathbf{s}^{(i)}$ 不唯一, 至少可以相差一个常数。另外, 如果是一般的 $N \times N$ 矩阵, 未必有 N 个线性无关的本征向量。比如矩阵 $\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$ 的本征值为 $\lambda_1 = \lambda_2 = 1$, 但只有一个线性无关的本征向量。但由于 W 是实对称矩阵, 因此肯定存在一组本征向量 $\mathbf{s}^{(1)}, \dots, \mathbf{s}^{(N)}$ 满足

- $\mathbf{s}^{(i)} (i = 1 \dots N)$ 所有分量都是实数。
- 不同的本征向量彼此正交, 并且每个本征向量都归一, 即

$$(\mathbf{s}^{(i)})^T \mathbf{s}^{(j)} = \sum_{k=1}^N s_k^{(i)} s_k^{(j)} = \delta_{ij} \quad (2.90)$$

我们称满足以上条件的本征向量为 W 的正交归一本征向量。注意, 正交归一本征向量也不唯一 (至少可以相差正负号)

例如对于矩阵 $W = \begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix}$, 于是有

$$\lambda_1 = a \quad \mathbf{s}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (2.91)$$

$$\lambda_2 = -a \quad \mathbf{s}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad (2.92)$$

我们定义 $N \times N$ 矩阵 S 为

$$S = [\mathbf{s}^{(1)} \quad \dots \quad \mathbf{s}^{(N)}] = \begin{pmatrix} s_1^{(1)} & \dots & s_1^{(N)} \\ \vdots & \ddots & \vdots \\ s_N^{(1)} & \dots & s_N^{(N)} \end{pmatrix} \quad (2.93)$$

利用(2.89)(2.90)可以证明

$$S^T S = S S^T = I \quad \Rightarrow \quad S^{-1} = S^T \quad (2.94)$$

$$W S = \begin{bmatrix} \lambda_1 \mathbf{s}^{(1)} & \lambda_2 \mathbf{s}^{(2)} & \dots & \lambda_N \mathbf{s}^{(N)} \end{bmatrix} \quad (2.95)$$

因此,

$$S^T W S = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \end{pmatrix} \quad (2.96)$$

等式右边的矩阵是以 $\lambda_1, \dots, \lambda_N$ 为对角元的对角矩阵, 我们记为 Λ 。

N 维方矩阵 A 的合同变换, 是指用一个 N 维方矩阵 B 对矩阵 A 进行作用 $B^T A B$, 从而得到一个新矩阵。由于 $S^T W S = \Lambda$, 因此一个实对称矩阵可以通过合同变换对角化, 化成对角矩阵 Λ 。对角化 W 的步骤是先找出本征值 $\lambda_1, \dots, \lambda_N$, 再找到每个本征值对应的本征向量 $\mathbf{s}^{(i)} (i = 1, \dots, N)$, 最后再给出 S, Λ , 它们必然满足 $S^T W S = \Lambda$ 。当然我们也有逆变换 $W = S \Lambda S^T$ 。

2.3.2 两个正定实对称矩阵的联合对角化

Thm 1. 对任意两个正定实对称矩阵 M, U 和任意正实数 m , 总存在一个 n 阶方阵 Z , 使得 $Z^T M Z = mI, Z^T U Z = \text{diag}\{\lambda_1 \cdots \lambda_n\}$ 。这里, I 是单位矩阵。需要强调的是, Z 不一定是正交矩阵, $\lambda_1 \cdots \lambda_n$ 也不一定是 U 矩阵的本征值, 它只能保证对角化。

证明. 我们证明的核心在于找到能够同时满足

$$Z^T M Z = mI \quad (2.97)$$

$$Z^T U Z = \text{diag}(\lambda_1, \dots, \lambda_N) \quad (2.98)$$

的矩阵 Z 。证明的第一阶段, 我们做变换使得(2.97)被满足, 即找到一个矩阵 Z' , 使得

$$Z'^T M Z' = mI$$

矩阵 U 也通过 Z' 诱导的合同变换, 变换至 $Z'^T U Z'$, 可以验证它仍然是正定实对称的。因此证明的第二步, 我们可以将 $Z'^T U Z'$ 对角化, 即找到一个矩阵 S , 使得 $S^T = S^{-1}$ 的同时, 有

$$S^T (Z'^T U Z') S = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix} \quad (2.99)$$

而同时, 我们也有

$$S^T (Z'^T M Z') S = S^T mI S = mI \quad (2.100)$$

这时, $Z = Z' S$ 即为我们要寻找的矩阵。

设 M 的本征值为 m_1, \dots, m_n , 对应的正交归一本征向量拼成一个本征向量矩阵

$$K = (\mathbf{k}^{(1)} \cdots \mathbf{k}^{(n)}) \quad (2.101)$$

并且满足 $K^T = K^{-1}$ 。于是

$$K^T M K = \text{diag}\{m_1 \cdots m_n\} \equiv \Gamma \quad (2.102)$$

接下来，定义

$$\gamma = \text{diag} \left\{ \sqrt{\frac{m_1}{m}} \cdots \sqrt{\frac{m_n}{m}} \right\} \quad (2.103)$$

显然满足 $\gamma^T = \gamma$ ，并且

$$\gamma^{-1} = (\gamma^T)^{-1} = \text{diag} \left\{ \sqrt{\frac{m}{m_1}}, \cdots, \sqrt{\frac{m}{m_n}} \right\} \quad (2.104)$$

于是 Γ 矩阵就有

$$\begin{aligned} \Gamma &= \begin{pmatrix} m_1 & & \\ & \ddots & \\ & & m_n \end{pmatrix} = \begin{pmatrix} m & & \\ & \ddots & \\ & & m \end{pmatrix} \begin{pmatrix} \sqrt{\frac{m_1}{m}} & & \\ & \ddots & \\ & & \sqrt{\frac{m_n}{m}} \end{pmatrix} \begin{pmatrix} \sqrt{\frac{m_1}{m}} & & \\ & \ddots & \\ & & \sqrt{\frac{m_n}{m}} \end{pmatrix} \\ &= m\gamma\gamma^T \end{aligned}$$

经过代数变换将有

$$\gamma^{-1}\Gamma\gamma^{-1} = mI = \gamma^{-1}K^T M K \gamma \quad (2.105)$$

令 $Z' = K\gamma^{-1}$ ，则有

$$Z' M Z' = mI \quad (2.106)$$

接下来，利用 Z' 矩阵，我们构造矩阵

$$W = Z'^T U Z' = (\gamma^{-1} K^T) U (K \gamma^{-1}) \quad (2.107)$$

立刻可以验证 $W^T = W$ ，并且 W 矩阵正定，从而可以进行合同对角化。设 W 的本征值为 $\lambda_1, \cdots, \lambda_N$ ，与各个本征值 λ_i 对应的本征向量为 $\mathbf{s}^{(i)}, i = 1, \cdots, N$ ，于是

$$W \mathbf{s}^{(i)} = \lambda_i \mathbf{s}^{(i)} (\lambda_i > 0) \quad (\mathbf{s}^{(i)})^T \mathbf{s}^{(j)} = \delta_{ij} \quad (2.108)$$

于是用本征向量构造矩阵 $S = [\mathbf{s}^{(1)} \cdots \mathbf{s}^{(N)}]$ ，以及本征值构成的对角阵 $\Lambda = \text{diag}\{\lambda_1, \cdots, \lambda_N\}$ ，于是

$$S^T W S = \Lambda \quad S^T = S^{-1} \quad (2.109)$$

从而定义 $Z = Z' S = K \gamma^{-1} S$ ，按照我们前面的思路，这就是我们要找的共同合同对角化矩阵，构造完成。 \square

从证明的步骤中可以看出，要使得两个正定实对称矩阵 M, U 被共同对角化，计算的核心是要找出 K, γ, S 这三个矩阵。

2.4 稳定平衡位置附近小振动的一般理论

2.4.1 势能的极小点：稳定平衡点

考虑一个保守系统。设系统原始广义坐标为 (y_1, y_2, \dots, y_N) ，拉氏量为 $L = T - V$ ，其中 T 为动能， V 为势能。设动能可以写成

$$T = \sum_{i,j=1}^N \frac{1}{2} t_{ij}(y_1, y_2, \dots, y_N) \dot{y}_i \dot{y}_j, \quad (2.110)$$

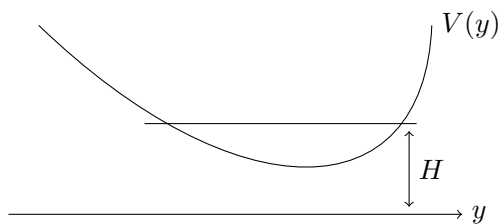
其中 $t_{ij} = t_{ji}$ ($i, j = 1, \dots, N$)。注意：若 (y_1, y_2, \dots, y_N) 与直角坐标关系不含 t ，则 (2.110) 式总成立。同时，设势能可以写成

$$V = V(y_1, y_2, \dots, y_N), \quad (2.111)$$

即 V 不含时间和广义速度。

我们进一步设 V 在 $(y_{10}, y_{20}, \dots, y_{N0})$ 处**极小**，则易证：若系统初始位于此位置，且广义速度为 0，则系统会一直停在此处，不运动。即：若 $y_i(t=0) = y_{i0}; \dot{y}_i(t=0) = 0$, ($i = 1, 2, \dots, N$)，则 $y_i(t) = y_{i0}$, ($i = 1, 2, \dots, N$)。因此，势能的极小点是系统的**平衡点**。

又，由于系统机械能 $H = T + V$ 守恒，且 $T \geq 0$ ，因此系统的只能在势能小于等于总机械能（即 $V \leq H$ ）的范围运动。因此，若系统初始位置不在精确位于上述势能极小点，初始速度也不精确为 0，但初始位置对上述极小点偏离很小（即 $|y_i(t=0) - y_{i0}|$ ($i = 1, \dots, N$) 较小），初始速度也较小 ($|\dot{y}_i(t=0)|$ ($i = 1, \dots, N$) 较小)，则 H 只比 $V(y_{10}, y_{20}, \dots, y_{N0})$ 大一点。因此系统未来会运动，但只会在 $(y_{10}, y_{20}, \dots, y_{N0})$ 附近（即满足 $V \leq H$ 的区域）来回振动。因此，势能的极小点是系统的**稳定平衡点**。下图为上述分析的在系统只有 1 个坐标 ($N = 1$) 的情况下的示意图：



2.4.2 稳定平衡位置附近的小振动：直观广义坐标和近似拉氏量。

我们仍考虑上面的系统。设系统在 $(y_{10}, y_{20}, \dots, y_{N0})$ 附近小振动。定义直观的广义坐标：

$$q_i = y_i - y_{i0}, \quad (i = 1, 2, \dots, N). \quad (2.112)$$

它们具有明确的物理意义：表示系统对平衡位置的偏移。我们假设系统的机械能是如此之低，以至于因此可把 L 在 $\{q_i = 0, \dot{q}_i = 0 | i = 1, 2, \dots, N\}$ 附近展开成 q_i, \dot{q}_i ($i = 1, 2, \dots, N$) 的级数，保留

到 q_i 与 $\dot{q}_i (i = 1, 2, \dots, N)$ 的二次方。因此，动能可近似为：

$$\begin{aligned} T &= \sum_{ij} \frac{1}{2} t_{ij}(q_1 + y_{10}, q_2 + y_{20}, \dots, q_N + y_{N0}) \dot{q}_i \dot{q}_j \\ &= \sum_{ij} \frac{1}{2} [t_{ij}(y_{10}, y_{20}, \dots, y_{N0}) + \mathcal{O}(q_1, \dots, q_N)] \cdot \dot{q}_i \dot{q}_j \\ &\approx \sum_{ij} \frac{1}{2} M_{ij} \dot{q}_i \dot{q}_j, \end{aligned} \quad (2.113)$$

其中

$$M_{ij} = t_{ij}(y_{10}, y_{20}, \dots, y_{N0}), \quad (2.114)$$

且 $M_{ij} = M_{ji}$ 。而势能可以近似为：

$$V \approx V(y_{10}, y_{20}, \dots, y_{N0}) + \sum_{ij} \frac{1}{2} U_{ij} q_i q_j \quad (2.115)$$

其中

$$U_{ij} = \left. \frac{\partial^2 V}{\partial y_i \partial y_j} \right|_{y_1=y_{10}, y_2=y_{20}, \dots, y_N=y_{N0}} \quad (2.116)$$

这里利用了

$$\left. \frac{\partial V}{\partial y_i} \right|_{y_1=y_{10}, y_2=y_{20}, \dots, y_N=y_{N0}} = 0. \quad (2.117)$$

进一步，我们可以通过重新选择能量 0 点去掉势能与 q_1, \dots, q_N 无关的项。于是近似拉氏量的最终形式为：

$$L \approx \frac{1}{2} \sum_{ij} M_{ij} \dot{q}_i \dot{q}_j - \frac{1}{2} \sum_{ij} U_{ij} q_i q_j. \quad (2.118)$$

2.4.3 简正坐标

我们注意到：拉氏量 (2.118) 给出的直观广义坐标 (q_1, q_2, \dots, q_N) 的 Lagrange 方程彼此相互耦合（相当于一组耦合振子）。现在我们寻求让 Lagrange 量解耦成 $L = \sum_{i=1}^N L_i(\eta_i, \dot{\eta}_i)$ 的广义坐标 $(\eta_1, \eta_2, \dots, \eta_N)$ ，也就是简正坐标。

为此，我们定义 N 维列向量 $\mathbf{q} = \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_N \end{pmatrix}$ 。还定义 \mathbf{M} 是以 M_{ij} 为矩阵元的 $N \times N$ 矩阵，

\mathbf{U} 是以 U_{ij} 为矩阵元的 $N \times N$ 矩阵。显然矩阵 \mathbf{M} 和 \mathbf{U} 都是对称矩阵，即 $\mathbf{M} = \mathbf{M}^T, \mathbf{U} = \mathbf{U}^T$ 。进一步，由于动能 T 在 $\dot{q}_1, \dot{q}_2, \dots, \dot{q}_N$ 不都为 0 必然大于零，所以 \mathbf{M} 必然是正定的（为什么？）。由于 V 在 $(y_{10}, y_{20}, \dots, y_{N0})$ 极小，所以 \mathbf{U} 必然是正定的（为什么？）综上所述， \mathbf{M} 和 \mathbf{U} 都是实对称正定矩阵。

利用上述矩阵，我们可以把拉氏量 (2.118) 改写成：

$$L = \frac{1}{2} \dot{\mathbf{q}}^T \mathbf{M} \dot{\mathbf{q}} - \frac{1}{2} \mathbf{q}^T \mathbf{U} \mathbf{q}. \quad (2.119)$$

由此可见， \mathbf{M} 和 \mathbf{U} 事实上的单个弹簧振子的质量和弹性系数在耦合振子系统推广。这有助于理解此二矩阵的正定型。我们现在利用一个合同变换把矩阵 \mathbf{M} 和 \mathbf{U} 同时对角化。**以下使用2.3中的数学定理，所有记号的定义与该定理中相同。**设 m 为任意正数， $m_{1,\dots,N}$ 为 \mathbf{M} 的本征值， $\mathbf{k}^{(1,\dots,N)}$ 为相应的正交归一的本征向量。即：

$$\mathbf{M} \mathbf{k}^{(i)} = m_i \mathbf{k}^{(i)}, \quad \mathbf{k}^{(i)T} \mathbf{k}^{(j)} = \delta_{ij}, \quad (i, j = 1, 2, \dots, N). \quad (2.120)$$

我们进一步定义

$$\mathbf{K} = [\mathbf{k}^{(1)}, \mathbf{k}^{(2)}, \dots, \mathbf{k}^{(N)}], \quad (2.121)$$

$$\gamma = \begin{bmatrix} \sqrt{\frac{m_1}{m}} & 0 & \dots & 0 \\ 0 & \sqrt{\frac{m_2}{m}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sqrt{\frac{m_N}{m}} \end{bmatrix}, \quad (2.122)$$

$$\mathbf{W} = (\gamma^{-1} \mathbf{K}^T) \mathbf{U} (\mathbf{K} \gamma^{-1}). \quad (2.123)$$

进一步，设 $\lambda_{1,\dots,N}$ 是 \mathbf{W} 的本征值， $\mathbf{s}^{(1,\dots,N)}$ 是相应正交归一的本征列向量。即

$$\mathbf{W} \mathbf{s}^{(i)} = \lambda_i \mathbf{s}^{(i)}, \quad \mathbf{s}^{(i)T} \mathbf{s}^{(j)} = \delta_{ij}, \quad (i, j = 1, 2, \dots, N). \quad (2.124)$$

我们再定义

$$\mathbf{S} = [\mathbf{s}^{(1)} \quad \mathbf{s}^{(2)} \quad \dots \quad \mathbf{s}^{(N)}], \quad (2.125)$$

$$\mathbf{Z} = \mathbf{K} \gamma^{-1} \mathbf{S}, \quad (2.126)$$

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_N \end{bmatrix}. \quad (2.127)$$

则有

$$\mathbf{Z}^T \mathbf{M} \mathbf{Z} = \begin{pmatrix} m & 0 & \dots & 0 \\ 0 & m & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & m \end{pmatrix}, \quad (2.128)$$

$$\mathbf{Z}^T \mathbf{U} \mathbf{Z} = \mathbf{\Lambda}, \quad (2.129)$$

以及

$$\mathbf{Z}^{-1} = \mathbf{S}^T \boldsymbol{\gamma} \mathbf{K}^T. \quad (2.130)$$

于是我们有：

$$\mathbf{M} = m(\mathbf{Z}^{-1})^T (\mathbf{Z}^{-1}), \quad \mathbf{U} = (\mathbf{Z}^{-1})^T \boldsymbol{\Lambda} (\mathbf{Z}^{-1}). \quad (2.131)$$

因此，拉氏量 (2.119) 可以被改写成

$$\begin{aligned} L &= \frac{1}{2} \dot{\mathbf{q}}^T \mathbf{M} \dot{\mathbf{q}} - \frac{1}{2} \mathbf{q}^T \mathbf{U} \mathbf{q} \\ &= \frac{m}{2} \dot{\mathbf{q}}^T (\mathbf{Z}^{-1})^T (\mathbf{Z}^{-1}) \dot{\mathbf{q}} - \frac{1}{2} \mathbf{q}^T (\mathbf{Z}^{-1})^T \boldsymbol{\Lambda} (\mathbf{Z}^{-1}) \mathbf{q}. \end{aligned} \quad (2.132)$$

因此我们可以定义新坐标 $(\eta_1, \eta_2, \dots, \eta_N)$ 满足

$$\boldsymbol{\eta} \equiv \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_N \end{pmatrix} = \mathbf{Z}^{-1} \mathbf{q} = (\mathbf{S}^T \boldsymbol{\gamma} \mathbf{K}^T) \mathbf{q}, \quad (2.133)$$

即

$$\eta_i = \sum_j Z_{ij}^{-1} q_j, \quad (i = 1, 2, \dots, N). \quad (2.134)$$

于是 (2.132) 式可进一步被改写为：

$$L = \frac{m}{2} \dot{\boldsymbol{\eta}}^T \dot{\boldsymbol{\eta}} - \frac{1}{2} \boldsymbol{\eta}^T \boldsymbol{\Lambda} \boldsymbol{\eta} \quad (2.135)$$

$$= \sum_{i=1}^N \frac{m}{2} \dot{\eta}_i^2 - \frac{1}{2} \lambda_i \eta_i^2. \quad (2.136)$$

可见，Lagrangian 被解耦。因此 $(\eta_1, \eta_2, \dots, \eta_N)$ 是简正坐标。我们可以同时给出坐标逆变换：

$$\mathbf{q} = \mathbf{Z} \boldsymbol{\eta} = (\mathbf{K} \boldsymbol{\gamma}^{-1} \mathbf{S}) \boldsymbol{\eta}, \quad (2.137)$$

即

$$q_i = \sum_{j=1}^N Z_{ij} \eta_j, \quad (i = 1, 2, \dots, N). \quad (2.138)$$

几种特殊情形

(1) 若 M 是常数，即

$$\mathbf{M} = \begin{pmatrix} m & 0 & \dots & 0 \\ 0 & m & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & m \end{pmatrix} \quad (2.139)$$

则 $\mathbf{K} = \boldsymbol{\gamma} = \mathbf{I}$, 其中 \mathbf{I} 是单位阵, 只须对角化 \mathbf{U} 即可

(2) 若 \mathbf{M} 是对角矩阵

$$\mathbf{M} = \begin{pmatrix} m_1 & 0 & \dots & 0 \\ 0 & m_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & m_N \end{pmatrix} \quad (2.140)$$

则 $\mathbf{K} = \mathbf{I}$, 只须引入 $\boldsymbol{\gamma}$ 、 \mathbf{S} 即可

2.4.4 简正圆频率

简正坐标满足的 Lagrange 方程为:

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\eta}_i} \right) - \frac{\partial L}{\partial \eta_i} = 0, \quad (i = 1, 2, \dots, N), \quad (2.141)$$

即

$$\ddot{\eta}_i + \omega_i^2 \eta_i = 0, \quad (i = 1, 2, \dots, N). \quad (2.142)$$

这里

$$\omega_i = \sqrt{\frac{\lambda_i}{m}} \quad (i = 1, 2, \dots, N). \quad (2.143)$$

与 m 的选择无关, 由系统本身决定, 称为简正圆频率. 系统共 N 个简正圆频率.

方程 (2.142) 的通解形式为:

$$\eta_i(t) = A_i \cos(\omega_i t + \phi_i), \quad (i = 1, 2, \dots, N). \quad (2.144)$$

其中 $\{A_i, \phi_i\}$ 由初始条件决定. 可以看出: $\{\eta_i(t) | i = 1, 2, \dots, N\}$ 的行为如同一组独立的简谐振子。

2.4.5 简正模式

定义 \mathbf{Z} 矩阵的第 i 列为列矢量 $\mathbf{z}^{(i)}, (i = 1, 2, \dots, N)$, 即:

$$\mathbf{z}^{(i)} = \begin{pmatrix} Z_{1i} \\ Z_{2i} \\ \vdots \\ Z_{Ni} \end{pmatrix} \quad (2.145)$$

$$\mathbf{Z} = \begin{bmatrix} \mathbf{z}^{(1)} & \mathbf{z}^{(2)} & \dots & \mathbf{z}^{(N)} \end{bmatrix}. \quad (2.146)$$

则

$$\mathbf{q}(t) = \mathbf{Z}\boldsymbol{\eta}(t) \quad (2.147)$$

$$= \sum_{i=1}^N \eta_i(t) \mathbf{z}^{(i)} \quad (2.148)$$

$$= \sum_{i=1}^N A_i \cos(\omega_i t + \phi_i) \mathbf{z}^{(i)}. \quad (2.149)$$

简正模式是一组 (N 个) 特殊的集体运动。在每个简正模式中, 只有一个简正坐标随时间变化。具体地说, 第 j 个简正模式下, $\eta_{i \neq j}(0) = 0, \dot{\eta}_{i \neq j}(0) = 0$, 因此 $\eta_{i \neq j}(t) = 0, \eta_j(t) = A_j \cos(\omega_j t + \phi_j)$ 。因此, 在第 j 个简正模式下 ($j = 1, \dots, N$) 直观的广义坐标随时间的变化方式为:

$$\mathbf{q}(t) = \eta_j \mathbf{z}^{(j)} = A_j \cos(\omega_j t + \phi_j) \mathbf{z}^{(j)}, \quad (2.150)$$

即

$$q_i(t) = A_j \cos(\omega_j t + \phi_j) Z_{ij}, \quad (i = 1, 2, \dots, N). \quad (2.151)$$

可见, 在具体的某个简正模式中, 不同的直观广义坐标 $q_i (i = 1, \dots, N)$ 之间比例锁定, 不随时间变化。且每个 q_i 都以相同的单一圆频率 (即该模式对应的简正圆频率) 随时间振荡。

一般情况下系统的运动为 N 个简正模式的组合:

$$\mathbf{q}(t) = \eta_1(t) \mathbf{z}^{(1)} + \eta_2(t) \mathbf{z}^{(2)} + \dots + \eta_N(t) \mathbf{z}^{(N)}, \quad (2.152)$$

即

$$q_i(t) = \sum_{j=1}^N A_j \cos(\omega_j t + \phi_j) Z_{ij}, \quad (i = 1, 2, \dots, N). \quad (2.153)$$

可见一般情况下每个直观的广义坐标 $q_i(t) (i = 1, \dots, N)$ 都是以 N 个简谐振荡行为 (对应不同简正圆频率) 的叠加。

2.4.6 小结

一般情况下稳定平衡位置附近小振动问题的处理程序如下:

1. 找到势能极小点的坐标 $(y_{10}, y_{20}, \dots, y_{N0})$ 。定义直观的广义坐标 $q_i = y_i - y_{i0}, (i = 1, 2, \dots, N)$ 。
2. 把动能、势能在附近 $q_i = 0, \dot{q}_i = 0 (i = 1, \dots, N)$ 附近展开, 保留到 2 阶:

$$L = \frac{1}{2} \dot{\mathbf{q}}^T \mathbf{M} \dot{\mathbf{q}} - \frac{1}{2} \mathbf{q}^T \mathbf{U} \mathbf{q}. \quad (2.154)$$

3. 利用2.3中定理, 求 $\mathbf{K}, \boldsymbol{\gamma}, \mathbf{S}, \boldsymbol{\Lambda}, \mathbf{Z}, \mathbf{Z}^{-1}$ 。

4. 简正坐标: $\boldsymbol{\eta} = \mathbf{Z}^{-1} \mathbf{q}, \mathbf{q} = \mathbf{Z} \boldsymbol{\eta}$ 。

5. 简正圆频率: $\omega_i = \sqrt{\lambda_i / m}$ 。

6. 第 j ($j = 1, 2, \dots, N$) 个简正模式: $\mathbf{q}(t) = A_j \cos(\omega_j t + \phi_j) \mathbf{z}^{(j)}$ 。

7. 时间演化:

$$\begin{aligned} \{\mathbf{q}(0), \dot{\mathbf{q}}(0)\} &\Rightarrow \{\boldsymbol{\eta}(0), \dot{\boldsymbol{\eta}}(0)\} \Rightarrow \{A_j, \phi_j | j = 1, 2, \dots, N\} \\ &\Rightarrow \{\eta_j(t) = A_j \cos(\omega_j t + \phi_j), j = 1, 2, \dots, N\} \\ &\Rightarrow \mathbf{q}(t) = \sum_{j=1}^N A_j \cos(\omega_j t + \phi_j) \mathbf{z}^{(j)}. \end{aligned}$$

即

$$q_i(t) = \sum_{j=1}^N A_j \cos(\omega_j + \phi_j) Z_{ij}, \quad (i = 1, 2, \dots, N).$$

2.4.7 求简正坐标与简正频率的其他方法

对于广义坐标 \mathbf{q} 的拉格朗日方程

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \quad i = 1, \dots, N \quad (2.155)$$

可以写成一个矩阵形式

$$M\ddot{\mathbf{q}} + U\mathbf{q} = 0 \quad (2.156)$$

前面已经证明, 系统一定存在 N 个简正模式, 因此, 这一对 \mathbf{q} 的拉格朗日方程将有特解

$$\mathbf{q}(t) = A \cos(\omega t + \phi) \tilde{\mathbf{z}} \quad (2.157)$$

这里 $\tilde{\mathbf{z}}$ 不含时。将这一特解形式代入(2.156), 就有

$$(U - \omega^2 M) \tilde{\mathbf{z}} = 0 \quad (2.158)$$

这一方程有非平庸解, 就意味着

$$\det\{U - \omega^2 M\} = 0 \quad (2.159)$$

解此方程, 就可以求出 N 个简正频率 $\omega_1, \dots, \omega_N$, 代入(2.158), 就可以给出 $\tilde{\mathbf{z}}^{(1)}, \dots, \tilde{\mathbf{z}}^{(N)}$ 。注意应当保证 $\tilde{\mathbf{z}}^{(1)}, \dots, \tilde{\mathbf{z}}^{(N)}$ 之间彼此线性无关 (如果 $\omega_i \neq \omega_j$, 则 $\tilde{\mathbf{z}}^{(i)}, \tilde{\mathbf{z}}^{(j)}$ 线性无关是一定的, 但特别要注意 $\omega_i = \omega_j$ 的简并情况)。于是系统的简正坐标 $(\tilde{\eta}_1, \dots, \tilde{\eta}_N)$ 与广义坐标 (q_1, \dots, q_N) 满足关系

$$\mathbf{q} = \tilde{\eta}_1 \tilde{\mathbf{z}}^{(1)} + \dots + \tilde{\eta}_N \tilde{\mathbf{z}}^{(N)}$$

如果我们记 $\tilde{Z} = \begin{pmatrix} \tilde{\mathbf{z}}^{(1)} & \dots & \tilde{\mathbf{z}}^{(N)} \end{pmatrix}$, 以及 $\tilde{\boldsymbol{\eta}} = \begin{pmatrix} \tilde{\eta}_1 \\ \vdots \\ \tilde{\eta}_N \end{pmatrix}$, 我们就有

$$\mathbf{q} = \tilde{Z} \tilde{\boldsymbol{\eta}} \quad \tilde{\boldsymbol{\eta}} = \tilde{Z}^{-1} \mathbf{q}$$

此方法的计算量和前面的方法是相同的, 因为都要对角化两个矩阵。注意, 此方法要计算到 \tilde{Z}^{-1} 。

2.5 耦合弹簧振子系统举例

2.5.1 两个耦合弹簧振子

A. 简正模式

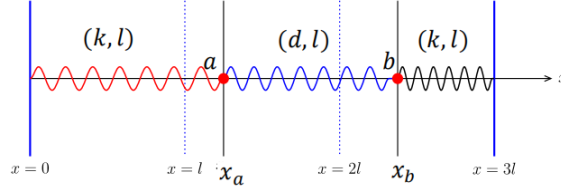


图 2.8: 两个耦合弹簧振子模型

如图 2.8 所示，我们考虑质量分别为 m_a, m_b 并且质量不等 $m_a \neq m_b$ 所组成的耦合弹簧振子。假设两个振子原始坐标为 x_a, x_b ，平衡位置为 $x_a = l, x_b = 2l$ ，于是我们给出直观坐标

$$q_a = x_a - l \quad q_b = x_b - 2l \quad (2.160)$$

在直观坐标下的拉格朗日量为

$$L = T - V = \frac{1}{2}(m_a \dot{q}_a^2 + m_b \dot{q}_b^2) - \frac{k}{2}(q_a^2 + q_b^2) - \frac{d}{2}(q_a - q_b)^2$$

我们定义 $\mathbf{q} = \begin{pmatrix} q_a \\ q_b \end{pmatrix}$ ，于是可以将拉格朗日量写为

$$L = \frac{1}{2} \dot{\mathbf{q}}^T M \dot{\mathbf{q}} - \frac{1}{2} \mathbf{q}^T U \mathbf{q} \quad (2.161)$$

这里矩阵 M, U 分别为

$$M = \begin{pmatrix} m_a & \\ & m_b \end{pmatrix} \quad U = \begin{pmatrix} k+d & -d \\ -d & k+d \end{pmatrix} \quad (2.162)$$

从而我们立刻可以得到 K 矩阵(2.121)在此例中为

$$K = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \quad (2.163)$$

取 $m = 1$ ，于是 γ 矩阵(2.122)在此例中为

$$\gamma = \begin{pmatrix} \sqrt{m_a} & \\ & \sqrt{m_b} \end{pmatrix} \quad (2.164)$$

从而 W 矩阵(2.123)在此例中为

$$W = \gamma^{-1} U \gamma^{-1} = \begin{pmatrix} \frac{k+d}{m_a} & -\frac{d}{\sqrt{m_a m_b}} \\ -\frac{d}{\sqrt{m_a m_b}} & \frac{k+d}{m_b} \end{pmatrix} \quad (2.165)$$

我们将 W 矩阵记为

$$W \equiv \begin{pmatrix} W_{aa} & W_{ab} \\ W_{ba} & W_{bb} \end{pmatrix} \quad (2.166)$$

利用 $\|W - \lambda I\| = 0$ 对角化 W 矩阵, 我们得到两个本征值

$$\lambda_{1,2} = \frac{1}{2} \left[(W_{aa} + W_{bb}) \pm \sqrt{(W_{aa} - W_{bb})^2 + 4W_{ab}^2} \right] \quad (2.167)$$

再利用 $W \mathbf{s}^{(i)} = \lambda_i \mathbf{s}^{(i)} (i = 1, 2)$ 可以得到两个本征矢量

$$\mathbf{s}^{(1)} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \mathbf{s}^{(2)} = \begin{pmatrix} \beta \\ -\alpha \end{pmatrix} \quad (2.168)$$

这里

$$\alpha = \frac{W_{ab}}{\sqrt{W_{ab}^2 + (\lambda_1 - W_{aa})^2}} \quad \beta = \frac{\lambda_1 - W_{aa}}{\sqrt{W_{ab}^2 + (\lambda_1 - W_{aa})^2}} \quad (2.169)$$

由本征矢量组成的矩阵 S (2.125) 在本例中为

$$S = [\mathbf{s}^{(1)} \quad \mathbf{s}^{(2)}] = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix}$$

从而 Z 矩阵 (2.126) 在本例中即为

$$Z = \gamma^{-1} S = \begin{pmatrix} \frac{\alpha}{\sqrt{m_a}} & \frac{\beta}{\sqrt{m_a}} \\ \frac{\beta}{\sqrt{m_b}} & -\frac{\alpha}{\sqrt{m_b}} \end{pmatrix} \equiv [\mathbf{z}^{(1)} \quad \mathbf{z}^{(2)}] \quad (2.170)$$

$$Z^{-1} = S^T \gamma = \begin{pmatrix} \alpha \sqrt{m_a} & \beta \sqrt{m_b} \\ \beta \sqrt{m_a} & -\alpha \sqrt{m_b} \end{pmatrix} \quad (2.171)$$

于是我们得到简正坐标 $\boldsymbol{\eta} = \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = Z^{-1} \mathbf{q}$, 即

$$\eta_1 = \alpha \sqrt{m_a} q_a + \beta \sqrt{m_b} q_b \quad (2.172)$$

$$\eta_2 = \beta \sqrt{m_a} q_a - \alpha \sqrt{m_b} q_b \quad (2.173)$$

而简正频率为

$$\omega_1 = \sqrt{\lambda_1/m} = \sqrt{\lambda_1} \quad (2.174)$$

$$\omega_2 = \sqrt{\lambda_2/m} = \sqrt{\lambda_2} \quad (2.175)$$

时间演化为

$$\eta_i(t) = A_i \cos(\omega_i t + \phi_i) \quad i = 1, 2 \quad (2.176)$$

这里 A_i, ϕ_i 由初始条件决定。对应的两种简正模式为

$$\text{Mode I} : \mathbf{q}(t) = A_1 \cos(\omega_1 t + \phi_1) \mathbf{z}^{(1)} \quad (2.177)$$

$$\text{Mode II} : \mathbf{q}(t) = A_2 \cos(\omega_2 t + \phi_2) \mathbf{z}^{(2)} \quad (2.178)$$

B. 弱耦合情形 I: m_a 接近 m_b

以下考虑振子弱耦合, 即 $d \ll k$ 的情形。首先, 如果两振子质量接近 (此时两个振子的固有频率很接近), 以至于 $|m_a - m_b| \ll d\sqrt{m_a m_b}/k$ (极限情形是 $m_a = m_b$)。此时我们有 $|W_{aa} - W_{bb}| \ll |W_{ab}|$, 于是我们有

$$\lambda_{1,2} = W_{aa} \pm |W_{ab}| \quad (2.179)$$

并且此时 $W_{ab} \ll W_{aa}$, 因此直接导致

$$\omega_1 - \omega_2 \approx 2|W_{ab}| \ll \omega_1 + \omega_2 \approx 2W_{aa} \quad (2.180)$$

对于本征矢量的分量, 就有

$$\frac{\alpha}{\beta} = \frac{W_{ab}}{\lambda_1 - W_{aa}} = -\frac{W_{ab}}{|W_{ab}|} = -1 \quad (2.181)$$

我们取 $\alpha = -\frac{1}{\sqrt{2}}, \beta = \frac{1}{\sqrt{2}}$, 于是简正坐标可简化为

$$\eta_1 \approx \sqrt{\frac{m_a}{2}}(q_b - q_a) \quad \eta_2 \approx \sqrt{\frac{m_a}{2}}(q_b + q_a) \quad (2.182)$$

或者我们有

$$q_a(t) \approx \frac{1}{\sqrt{2m_a}}(\eta_2 - \eta_1) \quad q_b(t) \approx \frac{1}{\sqrt{2m_a}}(\eta_1 + \eta_2) \quad (2.183)$$

如果我们考虑初始条件为

$$q_a(0) = A \neq 0 \quad \dot{q}_a(0) = 0 \quad q_b(0) = 0 \quad \dot{q}_b(0) = 0 \quad (2.184)$$

即认为初始状态下振子 a 从平衡位置移开, 振子 b 仍然在平衡位置, 初始速度都为零, 代入时间演化(2.176), 我们就有

$$\dot{\eta}_1(0) = \dot{\eta}_2(0) = 0 \quad (2.185)$$

$$\eta_1(0) = -A\sqrt{\frac{m_a}{2}} \quad \eta_2(0) = A\sqrt{\frac{m_a}{2}} \quad (2.186)$$

从而有

$$A_1 = -A\sqrt{\frac{m_a}{2}} \quad A_2 = A\sqrt{\frac{m_a}{2}} \quad \phi_1 = \phi_2 = 0 \quad (2.187)$$

从而得到直观坐标的运动模式

$$q_a(t) = A \cos\left(\frac{\omega_1 - \omega_2}{2}t\right) \cos\left(\frac{\omega_1 + \omega_2}{2}t\right) \quad (2.188)$$

$$q_b(t) = A \sin\left(\frac{\omega_1 - \omega_2}{2}t\right) \cos\left(\frac{\omega_1 + \omega_2}{2}t\right) \quad (2.189)$$

这意味着, 两个振子 a, b 都以 $\frac{\omega_1 + \omega_2}{2}$ 为圆频率振动, 而振幅则以 $\frac{\omega_1 - \omega_2}{2}$ 为圆频率交替缓慢变化 (交替振动), 如图 2.9 所示。

注意, 尽管 $d \ll k$, 但只要 m_a 足够接近 m_b , 以至于 $|m_a - m_b| \ll d\sqrt{m_a m_b}/k$ (左右两振子固有频率足够接近), 耦合效果就明显 (即 a 对平衡位置的偏离会引发 b 的振动, 最大振幅和 a 相同)。绝大多数固有频率接近的振子在弱耦合下都有此效应。

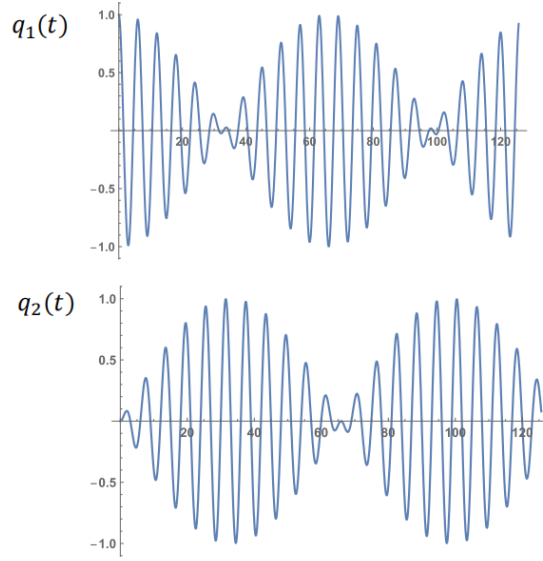


图 2.9: 在 $m_a = m_b = 1, k = 1, d = 0.1$ 参数选取下, 耦合二振子的直观坐标的时间演化图像

C. 弱耦合情形 II: m_a 与 m_b 相差很大

当两个振子质量 m_a, m_b 差别很大 (即所有两个振子固有频率差别很大), 以至于 $|m_a - m_b| \gg d\sqrt{m_a m_b}/k$, 此时就有

$$|W_{aa} - W_{bb}| \gg |W_{ab}| \quad (2.190)$$

我们不妨设 $m_a < m_b$, 则

$$\lambda_1 \approx W_{aa} \quad \lambda_2 \approx W_{bb} \quad (2.191)$$

从而

$$\frac{\alpha}{\beta} = \frac{W_{ab}}{\lambda_1 - W_{aa}} \approx \infty \quad (2.192)$$

因此可以取 $\alpha = 1, \beta = 0$, 于是简正坐标为

$$\eta_1 \approx \sqrt{m_a} q_a \quad \omega_1 \approx \sqrt{W_{aa}} = \sqrt{(k+d)/m_a} \quad (2.193)$$

$$\eta_2 \approx -\sqrt{m_b} q_b \quad \omega_2 \approx \sqrt{W_{bb}} = \sqrt{(k+d)/m_b} \quad (2.194)$$

这说明直观坐标 q_a, q_b 本身也是简正坐标。也就意味着此时两振子独立振动, 几乎没有耦合, 只是振动频率和原固有圆频率稍有不同。

我们考虑初始条件 $q_a(0) = A \neq 0, q_b = 0, \dot{q}_a(0) = \dot{q}_b(0) = 0$, 即仍然是振子 a 从平衡位置移开, b 还在平衡位置, 无初始速度释放, 则

$$q_a(t) \approx A \cos(\omega t) \quad q_b(t) \approx 0 \quad (2.195)$$

这意味着振子 a 对平衡位置的偏离不会引发振子 b 的振动。中间弹簧几乎不关联两个振子的运动, 但是会稍微修正 a, b 各自的频率, 从 $\sqrt{k/m_a}, \sqrt{k/m_b}$ 修正为 $\sqrt{(k+d)/m_a}, \sqrt{(k_d)/m_b}$ 。几乎所有固有频率差别大 (远离共振) 的二振子弱耦合都有这一规律。

2.5.2 振子串列

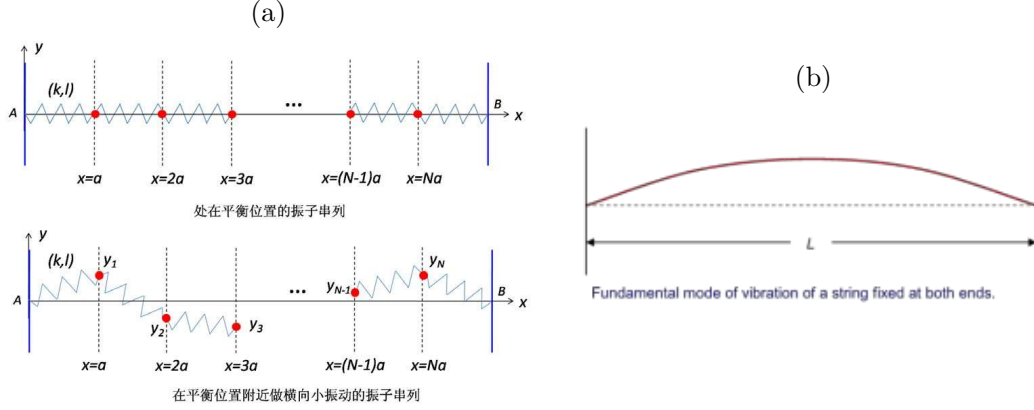


图 2.10: (a) 一个振子串列模型，其中上图表示各个振子处于初始位置情形，下图表示各个振子偏离平衡位置情形。(b) 在 $N \rightarrow \infty, a \rightarrow 0$ 但维持 $L = Na$ 不变的振子串列，此时串列变为一根两端固定，长度为 L 的弦。

考虑由 N 个振子组成的耦合串列，每个振子可以在 y 方向振动，如图 2.10-(a) 所示。这里每个振子之间横向间距 a 要大于弹簧的自然长度 l ，即弹簧始终处于伸长状态。系统的广义坐标为各个振子在 y 方向上的坐标 (y_1, y_2, \dots, y_N) ，则平衡位置就对应于 $y_1 = y_2 = \dots = y_N = 0$ 。于是串列的拉格朗日量为

$$L = T - V = \sum_{n=1}^N \frac{m}{2} \dot{y}_n^2 - \frac{1}{2} \sum_{mn} W_{mn} y_m y_n \quad (2.196)$$

其中

$$W_{mn} = \begin{cases} \frac{2K}{a} & m = n \\ -\frac{K}{a} & m = n \pm 1 \end{cases} \quad (2.197)$$

并且 $K = k(a - l)$ ，将其写为矩阵形式即为

$$W = \frac{K}{a} \begin{pmatrix} 2 & -1 & 0 & 0 & \cdots \\ -1 & 2 & -1 & 0 & \cdots \\ 0 & -1 & 2 & -1 & \cdots \\ 0 & 0 & -1 & 2 & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix} = \frac{K}{a} \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & \ddots & \\ & & \ddots & \ddots & \ddots \end{pmatrix} \quad (2.198)$$

W 矩阵的本征值为

$$\lambda_l = \frac{2K}{a} - \frac{2K}{a} \cos k^{(l)} a \quad k^{(l)} = \frac{l\pi}{(N+1)} a \quad l = 1 \cdots N \quad (2.199)$$

对应的本征向量为

$$\mathbf{s}^{(l)} = \begin{pmatrix} s_1^{(l)} \\ \vdots \\ s_N^{(l)} \end{pmatrix} = \frac{1}{Z} \begin{pmatrix} \sin k^{(l)} a \\ \vdots \\ \sin k^{(l)} Na \end{pmatrix} \quad (2.200)$$

这里 $Z = \sqrt{\sum_{n=1}^N \sin^2 k^{(l)} na}$, 即

$$s_i^{(l)} = \frac{1}{Z} \sin k^{(l)} x_i = \frac{1}{Z} \sin k^{(l)} ai \quad (2.201)$$

因此我们可以给出简正坐标

$$\eta_l = \sum_{n=1}^N s_l^{(n)} y_n \quad l = 1 \cdots N \quad (2.202)$$

以及对应的简正频率

$$\omega_l = \sqrt{\lambda_l/m} \quad l = 1 \cdots N \quad (2.203)$$

于是第 i 个简正模式为

$$\mathbf{y}(t) = A_l \cos(\omega_l t + \phi_l) \mathbf{s}^{(l)} \propto A_l \cos(\omega_l t + \phi_l) \begin{pmatrix} \sin k^{(l)} a \\ \vdots \\ \sin k^{(l)} Na \end{pmatrix} \quad (2.204)$$

它代表了所有振子的集体运动, 这种集体运动下, 不同振子的 y 坐标比值锁定 (为正弦函数在各振子 x 坐标取值的比值), 整体以单一频率振荡。简正坐标相当于将一般构型 \mathbf{y} 分解为各种 “sin 函数形状” 的特殊构型的组合, 每个构型前面的系数即为简正坐标, 如图 2.11-(a) 所示

当我们取极限 $N \rightarrow \infty, a \rightarrow 0$ 但保持 $L = Na$ 不变, 振子串列就变成了一个两端固定的长度为 L 的弦, 如图 2.10-(b) 所示。一般而言, 描述弦这一个 “一维空间场” (等价于无穷多个坐标) 随时间的变化应当使用一个波动方程 (偏微分方程) $\frac{\partial^2 y(x, t)}{\partial t^2} - a^2 \frac{\partial^2 y(x, t)}{\partial x^2} = 0$, 但利用振子串列, 我们相当于用 N 个拉格朗日方程 (常微分方程) 也能够描述这一系统。因此, 利用振子串列的分析, 不同求解弦的波动方程, 也可以得到两端固定弦的简正模式。此时, 振子串列的简正模式就对应于弦的驻波模式, 如图 2.11-(b) 所示。对于简正频率(2.203)(2.199)在 N 很大 l 很小时, 可以做近似 $\cos k^{(l)} a \approx 1 - \frac{1}{2}(k^{(l)} a)^2$, 因此得到

$$\omega_l \propto k_l \propto \frac{l}{(N+1)a} \quad (2.205)$$

因此连续弦最低的几个简正频率, 近似满足

$$\omega_l \propto \frac{l}{L} \quad (2.206)$$

从而可以通过调整弦长 L 来改变最低的几个简正频率。

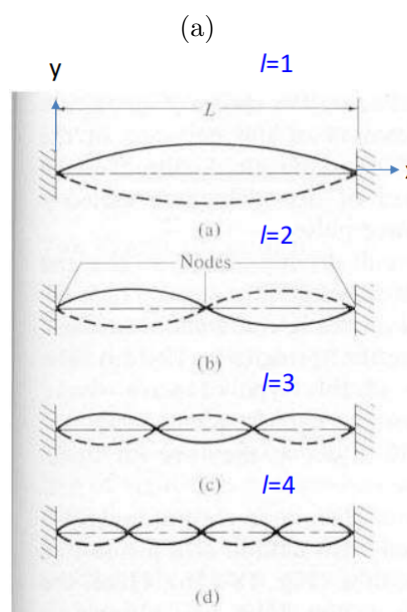
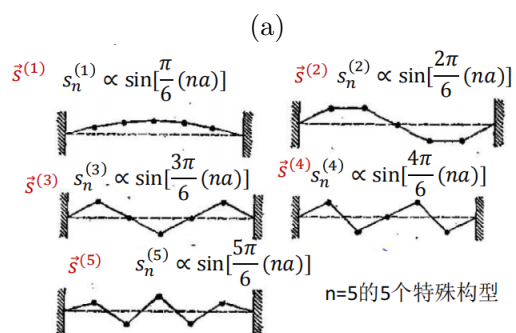


图 2.11: (a) $N = 5$ 时不同简正模式的五个特殊构型。(b) 在 $N \rightarrow \infty, a \rightarrow 0$ 但 $L = Na$ 不变极限下, 各个简正模式的构型

第三章 Hamilton 力学

3.1 Hamilton 方程

Lagrange 力学的回顾。

在 Lagrange 力学中, 我们关心的是系统的广义坐标的时间演化 $q_i(t), i = 1 \dots n$ 。我们使用 Lagrangian 来描述系统, 其形式为 $L = L[q_1, \dots, q_n; \dot{q}_1, \dots, \dot{q}_n; t]$ 。我们用 Lagrange 方程来描述时间演化

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \quad (i = 1, \dots, n) \quad (3.1)$$

定解需要 $2n$ 个初始条件 $q_i(0), \dot{q}_i(0)$

广义动量。

和广义坐标 q_i 共轭的广义动量 (正则动量) 被定义为

$$p_i = \frac{\partial L}{\partial \dot{q}_i} = p_i(q_1 \dots q_n; \dot{q}_1 \dots \dot{q}_n; t), \quad (i = 1, \dots, n). \quad (3.2)$$

因此, 我们可以反过来将广义速度表示成广义坐标的函数, 即

$$\dot{q}_i = \dot{q}_i(q_1 \dots q_n; p_1 \dots p_n; t), \quad (i = 1, \dots, n). \quad (3.3)$$

Hamilton 力学关注的内容。

在 Hamilton 力学中, 我们关心的是广义坐标和广义动量的时间演化 $q_i(t), p_i(t), (i = 1, \dots, n)$ 。要回答此问题, 最直接的办法是利用广义坐标和广义速度的初值 $q_i(0), \dot{q}_i(0), (i = 1, \dots, n)$ 和 Lagrange 方程, 得到广义坐标和广义速度的时间演化 $q_i(t), \dot{q}_i(t), (i = 1, \dots, n)$, 然后带入 (3.2) 直接求得 $p_i(t), (i = 1, \dots, n)$ 。

不过, 我们希望找到 n 个广义坐标 $q_i(t)$ 和 n 个广义动量 $p_i(t)$ 所满足的方程。即, 与 Lagrange 方程等价的, 形如

$$\begin{cases} \frac{dq_i}{dt} = \text{所有 } p, q \text{ 的函数} \\ \frac{dp_i}{dt} = \text{所有 } p, q \text{ 的函数} \end{cases}, \quad (i = 1, \dots, n) \quad (3.4)$$

的方程。

直接的方程。

我们不妨首先考察一个一维系统 (即 $n = 1$ 的情况), 此时 $L = L(q, \dot{q}, t)$, 于是

$$p = \left. \frac{\partial L}{\partial \dot{q}} \right|_{q,t} = p(q, \dot{q}, t) \quad (3.5)$$

于是反解得到

$$\dot{q} = \dot{q}(q, p, t). \quad (3.6)$$

这就意味着

$$\frac{dq}{dt} = \dot{q}(q, p, t) \quad (3.7)$$

此方程已经具有了 (3.4) 中第一行的形式, 因此我们的问题已经解决了一半。注意此方程来自 L 和 p 的定义。

再根据 Lagrange 方程

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = \frac{dp}{dt} - \frac{\partial L}{\partial q} = 0, \quad (3.8)$$

可以得到

$$\frac{dp}{dt} = \left. \frac{\partial L}{\partial q} \right|_{\dot{q},t}. \quad (3.9)$$

注意上式的等号右边 $\left. \frac{\partial L}{\partial q} \right|_{\dot{q},t}$ 是 (q, \dot{q}, t) 的函数。为了强调这一点, 我们形式上定义函数

$$f(q, \dot{q}, t) = \left. \frac{\partial L}{\partial q} \right|_{\dot{q},t}. \quad (3.10)$$

进一步, 显然可利用 (3.6) 将 $f(q, \dot{q}, t)$ 中的 (\dot{q}) 表示为 (p, t) 的函数。因此方程 (3.9) 可以进一步被改写为

$$\frac{dp(t)}{dt} = f[q, \dot{q}(q, p, t), t]. \quad (3.11)$$

此方程已经具有了 (3.4) 中第二行的形式。注意此方程来自 Lagrange 方程。因此, 方程 (3.7) 和 (3.11) 已经满足了我们的要求。

Hamiltonian 和 Hamilton 方程。

我们现在将方程 (3.7) 和 (3.11) 改写为更漂亮的形式。

首先定义 Hamiltonian 为

$$H(q, p, t) = p\dot{q}(q, p, t) - L\left[q, \dot{q}(q, p, t), t\right]. \quad (3.12)$$

我们通常在强调 H 对广义坐标和广义动量的函数关系时, 称其为 Hamiltonian。在强调 H 在某一系统中的具体取值时, 通常称之为广义能量。

我们注意到，Hamiltonian 对广义动量的偏导数为：

$$\left. \frac{\partial H}{\partial p} \right|_{q,t} = \dot{q}(q, p, t) + p \frac{\partial \dot{q}}{\partial t} - \left. \frac{\partial L}{\partial \dot{q}} \right|_{q,t} \frac{\partial \dot{q}}{\partial p} = \dot{q}(q, p, t). \quad (3.13)$$

这里我们用到了 $\frac{\partial L}{\partial \dot{q}} = p$ 。因此式 (3.7) 可以被改写为

$$\frac{dq(t)}{dt} = \frac{\partial H(q, p, t)}{\partial p}. \quad (3.14)$$

另一方面，Hamiltonian 对广义坐标的偏导数为

$$\begin{aligned} \left. \frac{\partial H}{\partial q} \right|_{p,t} &= p \left. \frac{\partial \dot{q}}{\partial q} \right|_{p,t} - \left. \frac{\partial L}{\partial q} \right|_{\dot{q},t} - \left. \frac{\partial L}{\partial \dot{q}} \right|_{q,t} \cdot \left. \frac{\partial \dot{q}}{\partial q} \right|_{p,t} \\ &= - \left. \frac{\partial L}{\partial q} \right|_{\dot{q},t} = -f[q, \dot{q}(q, p, t), t]. \end{aligned} \quad (3.15)$$

因此，方程 (3.11) 可以被改写为

$$\frac{dp}{dt} = - \frac{\partial H}{\partial q}. \quad (3.16)$$

于是，我们最终将广义坐标和广义动量的时间演化方程 (3.7) 和 (3.11) 改写为：

$$\frac{dq(t)}{dt} = \frac{\partial H(q, p, t)}{\partial p}, \quad (\text{来自 Lagrangian 和广义动量的定义}), \quad (3.17)$$

$$\frac{dp(t)}{dt} = - \frac{\partial H(q, p, t)}{\partial q}, \quad (\text{根据 Lagrange 方程}). \quad (3.18)$$

这个方程组被称为 **Hamilton 方程**。它与 Lagrange 方程完全等价，与 Lagrange 方程不同的是，在一维情形下 Lagrange 方程是一个二阶常微分方程，利用广义坐标和广义速度的初值 $\{q(0), \dot{q}(0)\}$ 定解，得到广义坐标的时间演化 $q(t)$ 。而 Hamilton 方程则由两个一阶方程构成，利用初始的广义坐标和广义动量 $\{q(0), p(0)\}$ 来定解，得到广义坐标和广义动量二者的时间演化 $\{q(t), p(t)\}$ 。从数学形式上，Hamilton 方程具有很好的对称性。

例：一维无约束质点

我们以直角坐标作为广义坐标，可以写出其 Lagrangian 和广义动量分别为

$$\begin{aligned} L &= \frac{1}{2} m \dot{x}^2 - V(x) \\ p &= \frac{\partial L}{\partial \dot{x}} = m \dot{x} \end{aligned}$$

我们可以反解其广义速度为 $\dot{x} = \frac{p}{m}$ ，于是

$$H = \dot{x}p - L = \frac{p^2}{m} - \left(\frac{p^2}{2m} - V(x) \right) = \frac{p^2}{2m} + V(x) \quad (3.19)$$

进一步可以得到 Hamilton 方程

$$\begin{cases} \frac{dx}{dt} = \frac{\partial H}{\partial p} = \frac{p}{m} \\ \frac{dp}{dt} = - \frac{\partial H}{\partial x} = - \frac{\partial V}{\partial x} \end{cases} \quad (3.20)$$

上述 Hamilton 方程显然意味着

$$m\ddot{x} = \frac{dp}{dt} = -\frac{\partial V}{\partial x} \quad (3.21)$$

可见, Hamilton 方程与 Newton 方程等价。

n 个自由度的情况。

最后, 我们将上述讨论推广到 n 个自由度的情况, 此时其广义坐标为 $q_1 \cdots q_n$, Lagrangian 为 $L = L[q_1 \cdots q_n; \dot{q}_1 \cdots \dot{q}_n; t]$ 。广义动量为

$$p_i = \frac{\partial L}{\partial \dot{q}_i} = p_i(q_1 \cdots q_n; \dot{q}_1 \cdots \dot{q}_n; t), \quad (i = 1, \dots, n). \quad (3.22)$$

从而

$$\dot{q}_i = \dot{q}_i(q_1 \cdots q_n; p_1 \cdots p_n; t) \quad (i = 1, \dots, n). \quad (3.23)$$

这时的 Hamiltonian 为

$$\begin{aligned} H &= H[q_1 \cdots q_n; p_1 \cdots p_n; t] \\ &= \sum_{i=1}^n p_i \dot{q}_i(q_1 \cdots q_n; p_1 \cdots p_n; t) - L[q_1 \cdots q_n; \dot{q}_1(q_1 \cdots q_n; p_1 \cdots p_n; t) \cdots \dot{q}_n(q_1 \cdots q_n; p_1 \cdots p_n; t); t] \end{aligned}$$

则 n 个广义坐标和 n 个广义动量 $q_i(t), p_i(t), (i = 1, \dots, n)$ 将满足 Hamilton 正则方程

$$\begin{cases} \frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} \Big|_{p_j \neq i, q_1 \cdots q_n, t} \\ \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} \Big|_{q_j \neq i, p_1 \cdots p_n, t} \end{cases} \quad (i = 1, \dots, n) \quad (3.24)$$

与单个自由度的情况相同, 该 Hamilton 方程与 Lagrange 方程等价。 n 个自由度的情况下, Lagrange 方程为 n 个二阶方程, 以 n 个广义坐标和 n 个广义速度的初值作为定解条件, 可以得到 n 个广义坐标的时间演化。而 Hamilton 正则方程恒则是 $2n$ 个以解方程, 由 n 个广义坐标和 n 个广义动量的初值作为定解条件, 可以得到 n 个广义坐标和 n 个广义速度的时间演化。并且 Hamilton 方程形式上更为对称。

这里强调, Hamilton 方程只对保守系统适用。

相空间和世界图景

在哈密顿力学中, 已知外界影响 (例如受力、拉格朗日量、哈密顿量等) 的情况下, 要完全确定系统未来 (以及过去) 的演化, 必须知道一些信息。这些信息被称为系统此时的**状态**。考虑到哈密顿正则方程的形式与变量依赖为

$$\begin{aligned} \frac{dq_i(t)}{dt} &= \frac{\partial H(q_1 \cdots q_n; p_1 \cdots p_n)}{\partial p_i} \\ \frac{dp_i(t)}{dt} &= -\frac{\partial H(q_1 \cdots q_n; p_1 \cdots p_n)}{\partial q_i} \end{aligned} \quad (3.25)$$

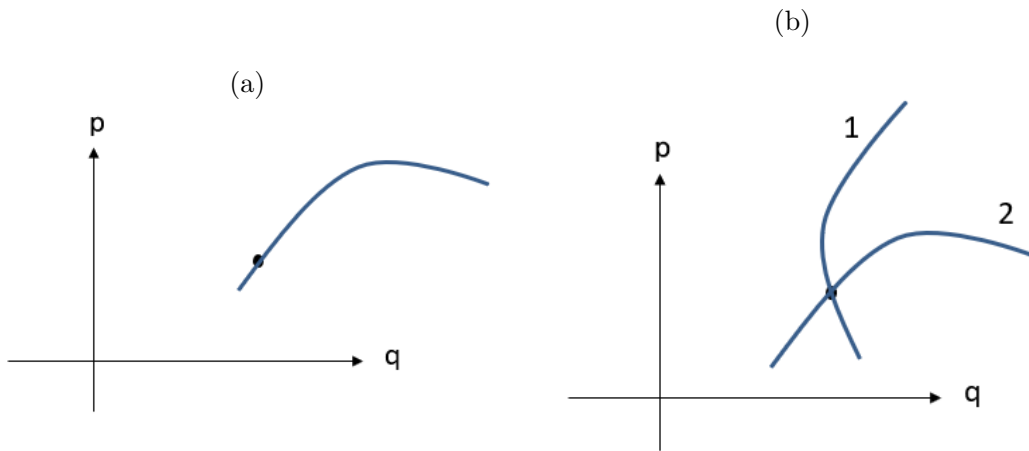


图 3.1: (a) 系统的状态在相空间中，图中的某一点为系统某时刻的状态，给定哈密顿量时，系统状态在相空间的演化给出一条相空间的轨迹。(b) 在相空间的两条轨迹 1,2 出现交叉的可能情形

因此任意 T 时刻下，系统的状态应当为

$$\{q_1(T), \dots, q_n(T); p_1(T), \dots, p_n(T)\}$$

所有数组 $(q_1 \dots q_n \ p_1 \dots p_n)$ 会构成一个 $2n$ 维空间，称之为系统的**相空间**。任何系统无论有多少自由度，每时每刻都占据对应维数的相空间的一个点。于是当给定系统的哈密顿量以后，系统的状态会在哈密顿量的支配下，在相空间中划出一条轨迹，并且系统 $t = 0$ 时刻在相空间的位置将决定系统在 $t > 0$ 和 $t < 0$ 时刻的位置，如图 3.1-(a) 所示。这就意味着相空间是“舞台”，而系统是“舞台”上的“演员”，而哈密顿量就是“导演”，因此在给定“导演”的时候，“演员”某时刻的位置，就会唯一决定了它此前和此后的轨迹（即“整场演出”中的轨迹）。

另外，当 $\frac{\partial H}{\partial t} = 0$ 时，即 $H = H(q_1 \dots q_n; p_1 \dots p_n)$ 而不显含时间。此时，相空间轨迹不会交叉。如果两条轨迹 1,2 在某点出现交叉，如图 3.1-(b) 所示，那么他们一定完全重合，即此时相空间的两条轨迹，要么完全重合，要么完全不交叉。当 $\frac{\partial H}{\partial t} \neq 0$ 时，相空间轨迹不会“同时”交叉，即两条轨迹不会在同一时刻在某一相点上重合。如果轨迹 1,2 在同一时刻的某点交叉，那么这两条轨迹就一定完全重合。

最后，我们将相空间的世界图景和牛顿力学做一个比较。和哈密顿力学不同，牛顿力学是以实空间作为“舞台”，每个质点会占据一个实空间的点，而“导演”是每个质点的受力 $\mathbf{F}(\mathbf{r}_i, \mathbf{v}_i, t)$ 。此时系统的状态为 $(\mathbf{r}_i, \mathbf{v}_i), i = 1 \dots n$ ，运动方程为 $\mathbf{F}(\mathbf{r}_j, \mathbf{v}_j, t) = m_i \frac{d\mathbf{v}_i}{dt}, i = 1 \dots n$ 。而拉格朗日力学是以广义坐标的取值组成的集合作为“舞台”，整个系统占据一个点，此时“导演”是拉格朗日量 $L(q_\alpha, \dot{q}_\alpha, t)$ 。

3.2 Hamilton 方程的应用

3.2.1 相关的问题

Hamiltonian 的构造。

在利用哈密顿力学解决问题时，首要任务是写出系统的正确的哈密顿量。以下介绍两个常用的方法：

(1) “保底”方法：从 Lagrangian 出发。即严格利用上一节的程序，从拉格朗日量出发，写出系统的广义动量，进而构造出哈密顿量。

$$\begin{aligned} L[q, \dot{q}, t] &\Rightarrow p = \frac{\partial L}{\partial \dot{q}} = p(q, \dot{q}, t) \\ &\Rightarrow \dot{q} = \dot{q}(q, p, t) \\ &\Rightarrow H = p\dot{q} - L = H(q, p, t) \end{aligned}$$

对任何系统，此方法都肯定能给出正确的结果，但有时会稍微繁琐一些。在用其他方法拿不准是否正确的时候，可以用此方法。

(2) 前面已经证明，如果拉氏量可以写成 $L = T - V$ ，且动能 T 是广义速度的二次型 $T = \sum_{ij} g_{ij}(q_1, q_2, \dots, q_n, t) \dot{q}_i \dot{q}_j$ ，而势能 V 是广义坐标和时间的函数 $V = V(q_1, q_2, \dots, q_n, t)$ ，与广义速度 $\{\dot{q}_1, \dots, \dot{q}_n\}$ 无关，则 Hamiltonian 必然是动能和势能之和，即必然有

$$H = T + V. \quad (3.26)$$

因此，对于满足上述条件的系统，构造哈密顿量时只需要把动能 T 写成广义坐标和广义动量的函数，之后直接和势能相加即可。

2. 广义动量与广义能量守恒

(1) 由于 $\frac{dp_\alpha}{dt} = -\frac{\partial H}{\partial q_\alpha}$, $(\alpha = 1, 2, \dots, n)$ ，因此若 $\frac{\partial H}{\partial q_\alpha} = 0$ ，则有 p_α 守恒

(2) 注意到：

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

故若 $\frac{\partial H}{\partial t} = 0$ ，则广义能量守恒。

3.2.2 直角坐标下的无约束保守系统

无 Lorentz 力情况下的一般讨论。

在本小节我们对直角坐标下的无约束保守系统在无 Lorentz 力情况下做一般讨论。我们用“从拉氏量出发”的方法构造该系统的 Hamiltonian，并列出哈密顿方程。以后遇到此类系统，将直接使用本小节的结论。

我们考虑由 N 个质点组成的保守系统， N 个质点，第 α 个质点 ($\alpha = 1, \dots, N$) 质量 m_α ，直角坐标 $(x_\alpha, y_\alpha, z_\alpha)$ 。以直角坐标为广义坐标，则本系统广义坐标为 $(x_1, y_1, z_1, \dots, x_N, y_N, z_N)$ 。设该系统势能为

$$V = V(x_1, y_1, z_1, \dots, x_N, y_N, z_N, t) \quad (3.27)$$

这个系统的拉氏量是：

$$L = T - V = \sum_{\alpha=1}^N \frac{m_\alpha}{2} (\dot{x}_\alpha^2 + \dot{y}_\alpha^2 + \dot{z}_\alpha^2) - V.$$

于是直角坐标对应的广义动量为：

$$\begin{cases} p_{x_\alpha} = \frac{\partial L}{\partial \dot{x}_\alpha} = m_\alpha \dot{x}_\alpha \\ p_{y_\alpha} = \frac{\partial L}{\partial \dot{y}_\alpha} = m_\alpha \dot{y}_\alpha, \\ p_{z_\alpha} = \frac{\partial L}{\partial \dot{z}_\alpha} = m_\alpha \dot{z}_\alpha \end{cases} \quad (\alpha = 1, \dots, N). \quad (3.28)$$

显然，此时广义动量即为机械动量。直角坐标下的速度分量可以表示成广义动量的函数：

$$\dot{x}_\alpha = \frac{p_{x_\alpha}}{m_\alpha}, \dot{y}_\alpha = \frac{p_{y_\alpha}}{m_\alpha}, \dot{z}_\alpha = \frac{p_{z_\alpha}}{m_\alpha}, \quad (\alpha = 1, \dots, N). \quad (3.29)$$

因此，利用“从拉氏量出发”可以直接得出本系统的 Hamiltonian：

$$\begin{aligned} H &= \sum_{\alpha=1}^N (\dot{x}_\alpha p_{x_\alpha} + \dot{y}_\alpha p_{y_\alpha} + \dot{z}_\alpha p_{z_\alpha}) - L \\ &= \sum_{\alpha=1}^N \frac{1}{2m_\alpha} (p_{x_\alpha}^2 + p_{y_\alpha}^2 + p_{z_\alpha}^2) + V \\ &= T + V \end{aligned}$$

本系统的 Hamiltonian 为：

$$\begin{cases} \dot{x}_\alpha = \frac{\partial H}{\partial p_{x_\alpha}} = \frac{p_{x_\alpha}}{m_\alpha} \\ \dot{p}_{x_\alpha} = -\frac{\partial H}{\partial x_\alpha} = -\frac{\partial V}{\partial x_\alpha} \end{cases} \quad (3.30)$$

$$\begin{cases} \dot{y}_\alpha = \frac{\partial H}{\partial p_{y_\alpha}} = \frac{p_{y_\alpha}}{m_\alpha} \\ \dot{p}_{y_\alpha} = -\frac{\partial H}{\partial y_\alpha} = -\frac{\partial V}{\partial y_\alpha} \end{cases} \quad (3.31)$$

$$\begin{cases} \dot{z}_\alpha = \frac{\partial H}{\partial p_{z_\alpha}} = \frac{p_{z_\alpha}}{m_\alpha} \\ \dot{p}_{z_\alpha} = -\frac{\partial H}{\partial z_\alpha} = -\frac{\partial V}{\partial z_\alpha} \end{cases} \quad (3.32)$$

显然，上述 Hamiltonian 方程 (3.30, 3.31, 3.32) 分别与 Newton 方程

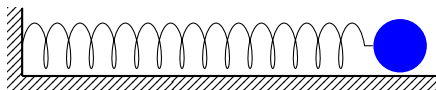
$$m_\alpha \ddot{x}_\alpha = -\frac{\partial V}{\partial x_\alpha}; \quad (3.33)$$

$$m_\alpha \ddot{y}_\alpha = -\frac{\partial V}{\partial y_\alpha}; \quad (3.34)$$

$$m_\alpha \ddot{z}_\alpha = -\frac{\partial V}{\partial z_\alpha}. \quad (3.35)$$

等价。

一维谐振子。



考虑图中的一维振子，假设振子在 x 轴上运动，质量、弹簧弹性系数分别为 m, k ，且弹簧平衡时振子位于 $x = 0$ 。根据前面的讨论，如果以振子的直角坐标 x 为广义坐标，则该振子的 Hamiltonian 为：

$$H = T + V = \frac{p_x^2}{2m} + \frac{1}{2}kx^2. \quad (3.36)$$

相应的，Hamilton 方程为：

$$\frac{dx}{dt} = \frac{p_x}{m}, \quad (3.37)$$

$$\frac{dp_x}{dt} = -kx. \quad (3.38)$$

而方程的解为：

$$\begin{cases} x(t) = A \cos(\omega t + \phi), \text{ 其中 } \omega = \sqrt{k/m} \\ p_x(t) = -m\omega A \sin(\omega t + \phi) \end{cases} \quad (3.39)$$

其中 (A, ϕ) 由 $x(0), p_x(0)$ 决定，满足

$$\begin{cases} x(0) = A \cos \phi \\ p_x(0) = -m\omega A \sin \phi \end{cases} \quad (3.40)$$

下面讨论该振子在相空间中的运动轨迹。显然，该系统的 Hamiltonian 也可以写成：

$$H = \frac{p_x^2}{2m} + \frac{m\omega^2 x^2}{2}. \quad (3.41)$$

而广义能量守恒给出

$$\frac{p_x^2(t)}{2m} + \frac{m\omega^2}{2}x^2(t) = H_0, \quad (3.42)$$

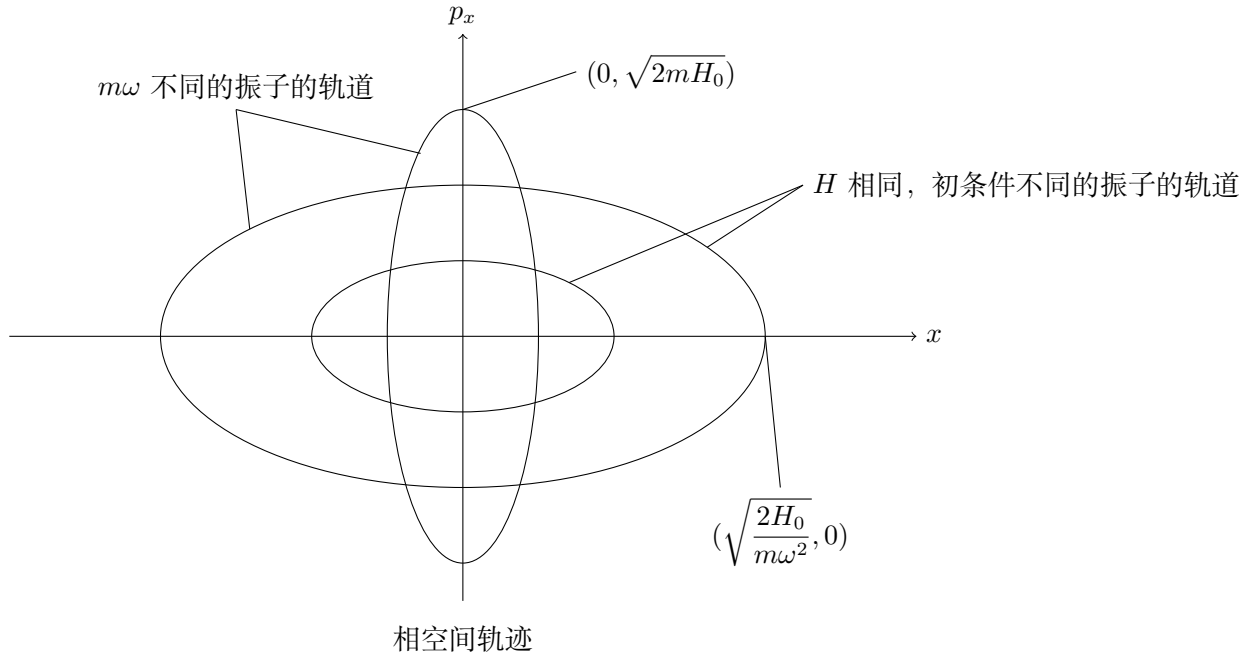
其中

$$H_0 = \frac{p_x^2(0)}{2m} + \frac{m\omega^2}{2}x^2(0) \quad (3.43)$$

是系统 0 时刻的机械能。

方程 (3.42) 可视为振子在相空间的轨道方程。根据该方程，振子在相空间的运动轨道具有以下特征：

- 相空间轨迹：椭圆
- 离心率由 $m\omega$ 决定，与初条件无关
- 大小：与初条件 (H_0) 有关
- 运动方向：顺时针



质心坐标与相对坐标（以 1 维情况为例）。

考虑 x 轴上的质点 A 、 B ，质量为 m_A, m_B 。设该系统为保守系统，且势能和速度无关。以各自直角坐标： x_A, x_B 为广义坐标，则拉氏量、广义动量、哈密顿量分别为：

$$L = \sum_{i=A,B} \frac{m_i}{2} \dot{x}_i^2 - V(x_A, x_B, t) \quad (3.44)$$

$$p_A = m_A \dot{x}_A, p_B = m_B \dot{x}_B \quad (3.45)$$

$$H = \sum_{i=A,B} \frac{p_i^2}{2m_i} + V(x_A, x_B, t). \quad (3.46)$$

定义该系统质心坐标 X 和相对坐标 x 分别为:

$$X = \frac{m_A x_A + m_B x_B}{M}, \quad (3.47)$$

$$x = x_A - x_B, \quad (3.48)$$

其中 $M = m_A + m_B$ 。于是有

$$x_A = X + \frac{m_B}{M}x, \quad (3.49)$$

$$x_B = X - \frac{m_A}{M}x. \quad (3.50)$$

如果以 X 与 x 为广义坐标, 则相应的拉氏量、广义动量分别为:

$$L = \frac{M}{2}\dot{X}^2 + \frac{\mu}{2}\dot{x}^2 - V\left(X + \frac{m_B}{M}x, X - \frac{m_A}{M}x\right) \quad (\text{其中 } \mu = \frac{m_A m_B}{m_A + m_B}), \quad (3.51)$$

$$p_X = \frac{\partial L}{\partial \dot{X}} = M\dot{X} = p_A + p_B \quad (\text{质心动量}), \quad (3.52)$$

$$p_x = \frac{\partial L}{\partial \dot{x}} = \mu\dot{x} = \frac{1}{M}(m_B p_A - m_A p_B) \quad (\text{相对动量}). \quad (3.53)$$

而该系统的哈密顿量为

$$H = \frac{p_X^2}{2M} + \frac{p_x^2}{2\mu} + V\left(X + \frac{m_B}{M}x, X - \frac{m_A}{M}x\right), \quad (3.54)$$

其中 $\frac{p_X^2}{2M}$ 为质心动能, $\frac{p_x^2}{2\mu}$ 为相对运动动能。

此外, 在 $m_A \ll m_B$ 时, 上述个坐标和动量有如下近似结果:

$$\begin{cases} M \approx m_B \\ \mu \approx m_A \\ X \approx x_B \\ x = x_A - x_B \\ p_x \approx m_A(\dot{x}_A - \dot{x}_B) \end{cases}. \quad (3.55)$$

上式中 $X \approx x_B$ 表示 X 位于 AB 位置连线上, 且很靠近 B 的位置。

若干具体情况

氢原子 如图 3.2 所示, 氢原子系统由一个质子和一个电子组成。我们记电子的直角坐标为 x_{e1}, y_{e1}, z_{e1} , 质子的直角坐标为 x_{n1}, x_{n2}, x_{n3} , 于是系统的广义动量为

$$\begin{cases} p_{x_{ei}} = m_e \dot{x}_{ei} \\ p_{x_{ni}} = m_n \dot{x}_{ni} \end{cases} \quad i = 1, 2, 3 \quad (3.56)$$

从而哈密顿量为

$$H = T + V = \sum_{i=1}^3 \left(\frac{p_{ei}^2}{2m_e} + \frac{p_{ni}^2}{2m_n} \right) - \frac{e^2}{4\pi\epsilon_0 \sqrt{\sum_{i=1}^3 (x_{ei} - x_{ni})^2}} \quad (3.57)$$

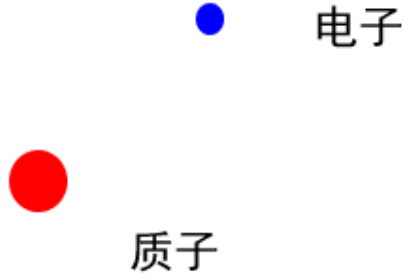


图 3.2: 氢原子系统

这里 ε_0 是真空介电常数。右侧势能项为质子-电子相互作用的库伦势能。

我们定义质心坐标 $\mathbf{X} = (X_1 \ X_2 \ X_3)^T$ 和相对坐标 $\mathbf{x} = (x_1 \ x_2 \ x_3)^T$

$$X_i = \frac{m_e x_{ei} + m_n x_{ni}}{m_e + m_n} \quad (3.58)$$

$$x_i = x_{ei} - x_{ni} \quad (3.59)$$

对应的广义动量分别为质心动量 P_i 和相对动量 p_i

$$P_i = M \dot{X}_i \quad (3.60)$$

$$p_i = \mu \dot{x}_i \quad (3.61)$$

这里 $M = m_e + m_n, \mu = \frac{m_e m_n}{M}$ 。系统的哈密顿量就变为

$$H = T + V = \sum_{i=1}^3 \frac{P_i^2}{2M} + \sum_{i=1}^3 \frac{p_i^2}{2\mu} - \frac{e^2}{4\pi\varepsilon_0 \sqrt{\sum_i x_i^2}} \quad (3.62)$$

这里第一项为质心运动的哈密顿量，而后两项则为相对运动的哈密顿量。

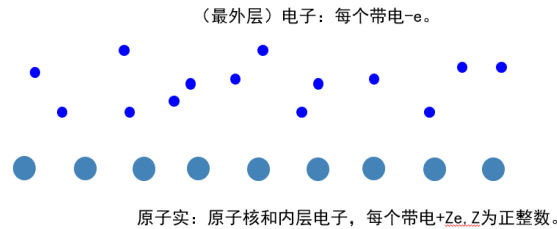


图 3.3: 金属晶格中的电子与原子实模型

金属晶格中的电子和原子实 如图 3.3 所示, 我们考虑金属晶格模型, 每一个晶格中的原子实都带有 Ze 的正点, 而外层电子每一个带电 $-e$ 。我们记第 j 个电子的直角坐标为 $(x_j^{(E)} \ t_j^{(E)} \ z_j^{(E)})$,

第 α 个原子实的直角坐标记为 $(x_\alpha^{(A)} \ y_\alpha^{(A)} \ z_\alpha^{(A)})$ 。系统的哈密顿量中应当包括如下几项：动能 T ，电子电子相互作用 V_{EE} ，电子原子实相互作用 V_{AE} ，原子实-原子实相互作用 V_{AA} ，即

$$H = T + V_{EE} + V_{AE} + V_{AA} \quad (3.63)$$

这几项的具体形式如下所示

$$T = \sum_{\alpha} \frac{1}{2m_A} \left[(p_{x_\alpha}^{(A)})^2 + (p_{y_\alpha}^{(A)})^2 + (p_{z_\alpha}^{(A)})^2 \right] + \sum_j \frac{1}{2m_e} \left[(p_{x_\alpha}^{(E)})^2 + (p_{y_\alpha}^{(E)})^2 + (p_{z_\alpha}^{(E)})^2 \right] \quad (3.64)$$

$$V_{EE} = \frac{1}{2} \sum_{i \neq j} \frac{e^2}{4\pi\epsilon_0 \sqrt{(x_i^{(E)} - x_j^{(E)})^2 + (y_i^{(E)} - y_j^{(E)})^2 + (z_i^{(E)} - z_j^{(E)})^2}} \quad (3.65)$$

$$V_{AE} = \frac{1}{2} \sum_{i, \alpha} \frac{Ze^2}{4\pi\epsilon_0 \sqrt{(x_i^{(E)} - x_\alpha^{(A)})^2 + (y_i^{(E)} - y_\alpha^{(A)})^2 + (z_i^{(E)} - z_\alpha^{(A)})^2}} \quad (3.66)$$

$$V_{AA} = \frac{1}{2} \sum_{\alpha \neq \beta} \frac{Z^2 e^2}{4\pi\epsilon_0 \sqrt{(x_\alpha^{(A)} - x_\beta^{(A)})^2 + (y_\alpha^{(A)} - y_\beta^{(A)})^2 + (z_\alpha^{(A)} - z_\beta^{(A)})^2}} \quad (3.67)$$

电磁场中带电质点 质量为 m ，带电为 q 的质点，在具有矢势 $\mathbf{A}(x, y, z)$ 和标势 $\phi(x, y, z)$ 的电磁场中运动。除了洛伦兹力以外，质点还受到对应于势能 $V'(x, y, z, t)$ 的保守力 \mathbf{F} ，即有

$$F_x = -\frac{\partial V'}{\partial x} \quad F_y = -\frac{\partial V'}{\partial y} \quad F_z = -\frac{\partial V'}{\partial z} \quad (3.68)$$

系统的拉格朗日量为

$$L = T - V - V' \quad (3.69)$$

其中动能项为

$$T = \frac{m}{2} (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) \quad (3.70)$$

势能项为

$$V(x, y, z; \dot{x}, \dot{y}, \dot{z}) = q\phi(x, y, z) - q[\dot{x}A_x(x, y, z) + \dot{y}A_y(x, y, z) + \dot{z}A_z(x, y, z)] \quad (3.71)$$

于是对应的广义动量为

$$\begin{cases} p_x = \frac{\partial L}{\partial \dot{x}} = m\dot{x} + qA_x(x, y, z) \neq m\dot{x} \\ p_y = \frac{\partial L}{\partial \dot{y}} = m\dot{y} + qA_y(x, y, z) \neq m\dot{y} \\ p_z = \frac{\partial L}{\partial \dot{z}} = m\dot{z} + qA_z(x, y, z) \neq m\dot{z} \end{cases} \quad (3.72)$$

注意这里我们给出的广义动量和机械动量并不一致。于是此时哈密顿量为

$$H = p_x \dot{x} + p_y \dot{y} + p_z \dot{z} - L \quad (3.73)$$

$$= \frac{1}{2m} [(p_x - qA_x)^2 + (p_y - qA_y)^2 + (p_z - qA_z)^2] + q\phi(x, y, z) + V' \quad (3.74)$$

$$= T + q\phi(x, y, z) + V' \quad (3.75)$$

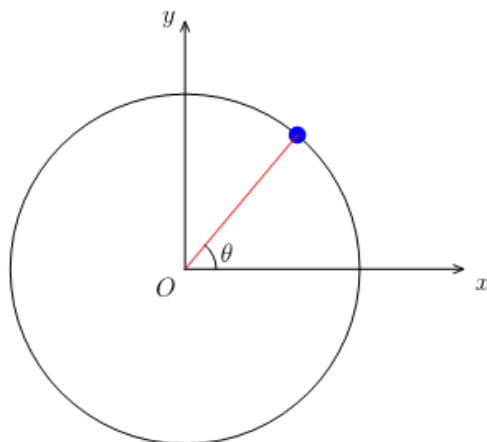


图 3.4: 圆环上的质点模型

圆环上的质点 考虑如图 3.4 所示的质量为 m 的质点，它被约束在 $x-y$ 平面上半径为 R 的光滑铁环上。设 y 方向为重力方向并选择 θ 作为广义坐标。它的拉格朗日量为

$$L = T - V = \frac{m}{2} R^2 \dot{\theta}^2 - mgR \sin \theta \quad (3.76)$$

从而广义动量为

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mR^2 \dot{\theta} \quad (3.77)$$

哈密顿量为

$$H = p_\theta \dot{\theta} - L = \frac{p_\theta^2}{2mR^2} + mgR \sin \theta \quad (3.78)$$

3.3 泊松括号和刘维尔定理

3.3.1 泊松括号

在经典力学中，任何广义坐标、广义动量、时间的（取值为实数的）函数都可称为一个“力学量”。一般的，一个力学量 A 可写成

$$A = A(q_1, \dots, q_n, p_1, \dots, p_n, t). \quad (3.79)$$

特定的广义坐标、特定的广义动量，以及哈密顿量，都是力学量。

若 A, B 都是力学量，则 A, B 的泊松括号也是一个力学量，定义为：

$$\{A, B\} \equiv \sum_{\alpha=1}^n \left(\frac{\partial A}{\partial q_\alpha} \frac{\partial B}{\partial p_\alpha} - \frac{\partial A}{\partial p_\alpha} \frac{\partial B}{\partial q_\alpha} \right). \quad (3.80)$$

泊松括号具有如下的性质（以下大写字母，在不声明的情况下都指力学量）：

性质 1: $\{A, B\} = -\{B, A\}$.

性质 2: 若 C 为常数 (常数也是力学量), 则 $\{A, C\} = 0$.

性质 3: $\{A_1 + A_2, B\} = \{A_1, B\} + \{A_2, B\}$.

上述性质 1-3 很容易通过泊松括号的定义直接证明。

性质 4: $\{A_1 A_2, B\} = \{A_1, B\} A_2 + A_1 \{A_2, B\}$.

证明: 根据泊松括号的定义, 我们有:

$$\{A_1 A_2, B\} = \sum_{\alpha=1}^n \left(\frac{\partial A_1 A_2}{\partial q_{\alpha}} \frac{\partial B}{\partial p_{\alpha}} - \frac{\partial A_1 A_2}{\partial p_{\alpha}} \frac{\partial B}{\partial q_{\alpha}} \right). \quad (3.81)$$

进一步, 我们有

$$\frac{\partial A_1 A_2}{\partial q_{\alpha}} = A_1 \frac{\partial A_2}{\partial q_{\alpha}} + \frac{\partial A_1}{\partial q_{\alpha}} A_2 \quad (3.82)$$

$$\frac{\partial A_1 A_2}{\partial p_{\alpha}} = A_1 \frac{\partial A_2}{\partial p_{\alpha}} + \frac{\partial A_1}{\partial p_{\alpha}} A_2 \quad (3.83)$$

将 (3.82, 3.83) 代入 (3.81), 即可证明性质 4。

性质 5: 雅可比等式 $\{A, \{B, C\}\} + \{B, \{C, A\}\} + \{C, \{A, B\}\} = 0$.

若 f, g 为 $N \times N$ 矩阵, f 与 g 的对易子定义为 $[f, g] \equiv fg - gf$ 。矩阵的对易子运算也有上面的性质 1-5。

性质 6: $\frac{\partial}{\partial t} \{A, B\} = \left\{ \frac{\partial A}{\partial t}, B \right\} + \left\{ A, \frac{\partial B}{\partial t} \right\}$

证明.

$$\begin{aligned} \frac{\partial}{\partial t} \{A, B\} &= \frac{\partial}{\partial t} \sum_{\alpha=1}^n \left[\frac{\partial A}{\partial q_{\alpha}} \frac{\partial B}{\partial p_{\alpha}} - \frac{\partial A}{\partial p_{\alpha}} \frac{\partial B}{\partial q_{\alpha}} \right] \\ &= \sum_{\alpha} \left[\left(\frac{\partial}{\partial q_{\alpha}} \left(\frac{\partial A}{\partial t} \right) \right) \frac{\partial B}{\partial p_{\alpha}} + \frac{\partial A}{\partial q_{\alpha}} \frac{\partial}{\partial p_{\alpha}} \frac{\partial B}{\partial t} \right. \\ &\quad \left. - \left(\frac{\partial}{\partial p_{\alpha}} \left(\frac{\partial A}{\partial t} \right) \right) \frac{\partial B}{\partial q_{\alpha}} - \frac{\partial A}{\partial p_{\alpha}} \frac{\partial}{\partial q_{\alpha}} \frac{\partial B}{\partial t} \right] \\ &= \left\{ \frac{\partial A}{\partial t}, B \right\} + \left\{ A, \frac{\partial B}{\partial t} \right\} \end{aligned}$$

□

性质 7: $\{q_{\alpha}, A\} = \frac{\partial A}{\partial p_{\alpha}}, \quad \{p_{\alpha}, A\} = -\frac{\partial A}{\partial q_{\alpha}}$

证明.

$$\{q_\alpha, A\} = \sum_{\beta=1}^n \left(\frac{\partial q_\alpha}{\partial q_\beta} \frac{\partial A}{\partial p_\beta} - \frac{\partial q_\alpha}{\partial p_\beta} \frac{\partial A}{\partial q_\beta} \right). \quad (3.84)$$

将关系 $\frac{\partial q_\alpha}{\partial q_\beta} = \delta_{\alpha\beta}$ 和 $\frac{\partial q_\alpha}{\partial p_\beta} = 0$ 带入上式, 即可证明性质 7 的前半部分。后半部分证明类似。 \square

性质 8:

$$\{p_\alpha, p_\beta\} = -\frac{\partial p_\beta}{\partial q_\alpha} = 0, \quad (3.85)$$

$$\{q_\alpha, q_\beta\} = \frac{\partial q_\beta}{\partial p_\alpha} = 0, \quad (3.86)$$

$$\{p_\alpha, q_\beta\} = -\frac{\partial q_\beta}{\partial q_\alpha} = -\delta_{\alpha\beta}. \quad (3.87)$$

3.3.2 泊松括号与 Hamilton 方程

借助泊松括号, 我们可以将 Hamilton 正则方程改写为这样的形式:

$$\frac{dq_\alpha}{dt} = \frac{\partial H}{\partial p_\alpha} = \{q_\alpha, H\} \quad (3.88)$$

$$\frac{dp_\alpha}{dt} = -\frac{\partial H}{\partial q_\alpha} = \{p_\alpha, H\} \quad (3.89)$$

其中 $\alpha = 1, \dots, n$.

3.3.3 泊松括号与守恒量

考虑系统的一个力学量 $A = A[q_1, \dots, q_n, p_1, \dots, p_n, t]$ 。前面多次说过, 在一次具体的运动中, 所有广义坐标和广义动量都在哈密顿方程的支配下随时间变化。于是系统在 t 时刻“感受到的 A 的取值”为:

$$A = A[q_1(t), \dots, q_n(t), p_1(t), \dots, p_n(t), t], \quad (3.90)$$

是时间 t 的复合函数。我们可以把该函数对时间 t 求全导数, 其含义为: 运动中系统“感受到的 A 的取值”随时间变化的快慢。之前多次说过, 此全导数记为 dA/dt 。

根据哈密顿方程, 我们有:

$$\frac{dA}{dt} = \sum_{\alpha=1}^n \left(\frac{\partial A}{\partial q_\alpha} \dot{q}_\alpha + \frac{\partial A}{\partial p_\alpha} \dot{p}_\alpha \right) + \frac{\partial A}{\partial t} \quad (3.91)$$

$$= \sum_{\alpha=1}^n \left(\frac{\partial A}{\partial q_\alpha} \frac{\partial H}{\partial p_\alpha} - \frac{\partial A}{\partial p_\alpha} \frac{\partial H}{\partial q_\alpha} \right) + \frac{\partial A}{\partial t}. \quad (3.92)$$

于是我们有

$$\frac{dA}{dt} = \{A, H\} + \frac{\partial A}{\partial t}. \quad (3.93)$$

如前所述, 公式 (3.95) 中等号左边表示

$$\frac{dA[q_1(t), \dots, q_n(t), p_1(t), \dots, p_n(t), t]}{dt}, \quad (3.94)$$

其中 $\{q_1(t), \dots, q_n(t), p_1(t), \dots, p_n(t)\}$ 表示一次具体的运动中 t 时刻的坐标和动量。而 (3.95) 式等号右边表示: 首先忘掉这次具体的运动, 只是根据 A 的定义计算 $\{A, H\}$ 和 $\frac{\partial A}{\partial t}$, 即, 把它们都写成 $\{q_1, \dots, q_n, p_1, \dots, p_n\}$ 的函数。之后把 $\{A, H\}$ 和 $\frac{\partial A}{\partial t}$ (都写成了 $\{q_1, \dots, q_n, p_1, \dots, p_n\}$ 的函数) 的自变量 $\{q_1, \dots, q_n, p_1, \dots, p_n\}$ 替换为此次具体运动中 t 时刻的 $\{q_1(t), \dots, q_n(t), p_1(t), \dots, p_n(t)\}$ 。因此, 在一次具体的运动中, (3.95) 式等号左右两边, 最终都是时间 t 的函数。

根据, (3.95) 式, 若算符 A 满足

$$\{A, H\} + \frac{\partial A}{\partial t} = 0, \quad (3.95)$$

(即根据 A 的定义计算出的 $\{A, H\} + \frac{\partial A}{\partial t}$, 作为 $\{q_1, \dots, q_n, p_1, \dots, p_n\}$ 的函数, 取值与自变量无关, 永远是 0), 则在任何一次运动中, 都有 $\frac{dA}{dt} = 0$, 即 A 是守恒量。

进一步地, 我们给出泊松定理: 若力学量 A, B 都是守恒量 (即它们都满足 (3.95) 式), 则 $\{A, B\}$ 也是守恒量。

证明. 我们记 ① = $\{\{A, B\}, H\}$ ② = $\frac{\partial \{A, B\}}{\partial t}$, 由雅可比等式, 我们有

$$\textcircled{1} = -\{\{B, H\}, A\} - \{\{H, A\}, B\} = \{A, \{B, H\}\} + \{\{A, H\}, B\}. \quad (3.96)$$

利用性质 6, 我们有

$$\textcircled{2} = \left\{A, \frac{\partial B}{\partial t}\right\} + \left\{\frac{\partial A}{\partial t}, B\right\}. \quad (3.97)$$

因此

$$\begin{aligned} \{\{A, B\}, H\} + \frac{\partial \{A, B\}}{\partial t} &= \textcircled{1} + \textcircled{2} \\ &= \left\{A, \{B, H\} + \frac{\partial B}{\partial t}\right\} + \left\{\{A, H\} + \frac{\partial A}{\partial t}, B\right\}. \end{aligned} \quad (3.98)$$

进一步, 由于 A, B 都满足 (3.95) 式, 所以 (3.98) 等号右边为 0, 即 $\{\{A, B\}, H\} + \frac{\partial \{A, B\}}{\partial t} = 0$, 因此 $\{A, B\}$ 是守恒量。□

3.3.4 相空间体积不变定理和 Liouville 定理

现在我们考虑一个力学系统, 它用广义坐标 q_1, q_2, \dots, q_n 来描述。在描述该系统的相空间中, 任意一块儿闭合的区域 Ω , 我们将它的相体积定义为

$$V_\Omega = \int_\Omega dq_1 \cdots dq_n \quad (3.99)$$

现在考虑相空间中的一个闭合区域 Ω , 假设在零时刻 Ω 中每个点上都有一个相同的系统, 随着时间的推移, 这些系统在相空间中的位置将不断变化, 如图 3.5 所示。在 t 时刻, 它们将移动到另外

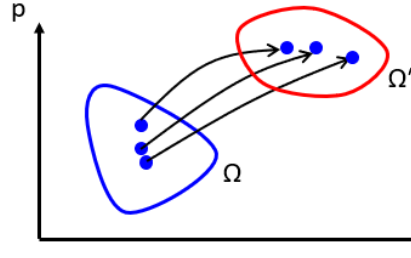


图 3.5: 在相空间的一个闭合区域 Ω , 区域内每一点经过时间演化到达相空间的另一个位置, 所有点的新位置构成一个新的闭合区域 Ω'

一些位置上去, 而这些位置将构成一个新的闭合区域 Ω' 。所谓的**相空间不变原理**是指: 只要这些系统的运动是被相同的哈密顿量所支配的, 那么一定有

$$V_{\Omega} = V_{\Omega'} \quad (3.100)$$

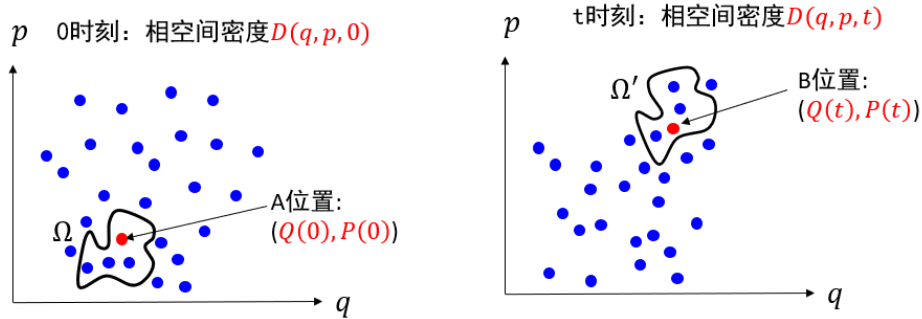


图 3.6: 某一特定系统在哈密顿正则方程的支配下进行时间演化, 初始位于相空间的 A 点, 在 t 时刻演化到相空间的 B 点。

接下来我们介绍 Liouville 定理, 为了方便说明起见, 我们以自由度为 1 的系统 (即 $n = 1$) 举例说明。现在考虑大量具有相同哈密顿量 H 的系统, 它们分布在相空间中, 每个系统都按照这一哈密顿量进行演化。在任意时刻, 这些系统都可以近似认为连续地分布在相空间中, 即可以视为相空间中的“流体”。我们定义相空间密度 $D(q, p, t)$, 从而 t 时刻在相空间某区域 Ω 中的系统数为

$$\int_{\Omega} D(q, p, t) dq dp \quad (3.101)$$

显然, $D(q, p, t)$ 将由初始时刻的 $D(q, p, 0)$ 所决定, 如图 3.6 所示。**Liouville 定理**指出: 如果广义坐标 $q(t)$ 和广义动量 $p(t)$ 满足系统的哈密顿量 H 所支配的哈密顿正则方程, 那么就有

$$D[q(t), p(t), t] = D[q(0), p(0), 0] \quad (3.102)$$

即零时刻图中 A 位置附近的相空间密度, 将等于它 t 时刻遵循哈密顿正则方程演化到达的位置 B 附近的相空间密度, 换言之: 追踪一个特定系统, 会发现该系统在运动中, 其“附近”的相空间密度将不随时间改变。

我们来解释 Liouville 定理证明的物理图像。我们考虑 3.6 中 A 位置附近面积为 S 的小区域 Ω ，假设零时刻其中有 M 个系统，那么在零时刻 A 位置的相空间密度约为

$$D[q(0), p(0), t] \approx \frac{M}{S} \quad (3.103)$$

在哈密顿正则方程所支配的时间演化下， t 时刻这 M 个系统运动到 B 位置附近的区域 Ω' 。根据相空间体积不变定理， Ω' 的面积也为 S ，因此 t 时刻 B 位置附近的相空间密度约为

$$D[q(t), p(t), t] \approx \frac{M}{s} \quad (3.104)$$

因此，我们就有

$$D[q(t), p(t), t] \approx D[q(0), p(0), t] \quad (3.105)$$

值得注意的是，这里的论证和小区域 Ω 的选择是无关的，因此事实上上式的“约等于”可以加强到“等于”。

现在考虑许多具有相同哈密顿量的系统，它们在相空间中的密度为 $D(q, p, t)$ 。由于 $q(t), p(t)$ 遵循哈密顿方程，按照 Liouville 定理就有 $D[q(t), p(t), t] = D[q(0), p(0), t]$ 。由于 t 可以任意小，因此可以有 $\frac{dD[q(t), p(t), t]}{dt} = 0$ ，于是

$$\frac{dD}{dt} = \frac{\partial D}{\partial q} \dot{q} + \frac{\partial D}{\partial p} \dot{p} + \frac{\partial D}{\partial t} = 0 \quad (3.106)$$

于是得到

$$\{D, H\} + \frac{\partial D}{\partial t} = 0 \quad (3.107)$$

这一结论对于任意自由度的情况都是正确的。

假设我们有一个系统，但是该系统在相空间中的位置不确定，而是满足一个概率分布 $\rho[q, p, t]$ ，即 t 时刻该系统处于相空间区域 Ω 中的概率为

$$\int_{\Omega} dq dp \cdot \rho(q, p, t) \quad (3.108)$$

若系统按照哈密顿方程演化，则 t 时刻的概率分布 $\rho(q, p, t)$ 由初始时刻 $\rho(q, p, 0)$ 决定，并且 $\rho(q, p, t)$ 的变化规律与前述相空间密度 $D(q, p, t)$ 相同。从图像上看，这是因为如果有大量 (N 个) 系统，每个都以概率 $\rho(q, p, t)$ 分布在相空间中，那么相空间的系统密度 $D(q, p, t) = N\rho(q, p, t)$ 。因此根据 Liouville 定理，概率分布 $\rho(q, p, t)$ 满足

$$\{\rho, H\} + \frac{\partial \rho}{\partial t} = 0 \quad (3.109)$$

这称之为 Liouville 方程，是哈密顿方程的“概率版”，适用于任意自由度的情况。一个推论是：如果 $\frac{\partial \rho}{\partial t} = 0$ ，即若 $\rho = \rho(q, p)$ ，这意味着系统的概率分布处于稳态，那么一定有 $\{\rho, H\} = 0$ 。例如如果系统处于温度为 T 的平衡态，那么

$$\rho(q, p) = \frac{1}{Z} e^{-H(q, p)/k_B T} \quad (3.110)$$

这里 Z 是一个归一化系数，而 k_B 是 Boltzmann 常数。

第四章 变分法和 Hamilton 原理

4.1 集合、映射、函数与泛函

所谓的集合，指的是“一些事物的总和”，例如我们常见的实数集合 \mathbb{R} 以及复数集合 \mathbb{C} 。它是数学中最基本的概念，无法给出更精确的定义。在现代数学中，总是倾向于将各种研究对象定义为各种“有结构的集合”，例如线性代数的研究对象是“线性空间”（定义了加法和数乘的集合）。

两个集合 A, B 之间的笛卡尔积也是一个集合，它的定义为

$$A \times B = \{(a, b) | a \in A, b \in B\} \quad (4.1)$$

例如 $\mathbb{R} \times \mathbb{R} = \{(x, y) | x \in \mathbb{R}, y \in \mathbb{R}\}$, $\mathbb{R} \times \mathbb{R} \times \mathbb{R} = \{(x, y, z) | x, y, z \in \mathbb{R}\}$, 更一般地，我们有

$$\underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n\text{个}} \equiv \mathbb{R}^n \quad (4.2)$$

这里 \mathbb{R}^n 中的每一个成员都是形如 (x_1, \cdots, x_n) 的 n 维实数数组。

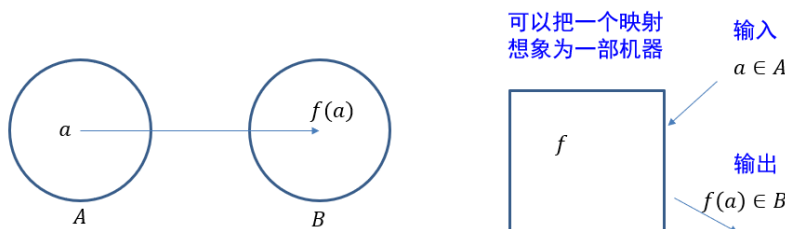


图 4.1: 映射的图示与理解

在集合 A 和集合 B 之间，我们可以定义一个对应关系 f ，使得 $\forall a \in A$ ，都有唯一的一个 $f(a) \in B$ 与之对应，我们称之为定义了一个从 A 到 B 的映射，记为 $f: A \rightarrow B$ 。可以将映射想象成一部机器，如图 4.1 所示，这部机器接受来自集合 A 的输入 a ，而输出一个集合 B 中的成员 $f(a)$ 。定义一个映射需要同时有 A, B 和对应规则这三个要素。

我们常说的函数，是指从集合 \mathbb{R}^n 或者 \mathbb{C}^n ，到集合 \mathbb{R} 或者 \mathbb{C} 的映射。例如 $f(x) = x^2, x \in \mathbb{R}$ 是一个函数 $f: \mathbb{R} \rightarrow \mathbb{R}$ ，而 $f(z) = e^z, z \in \mathbb{C}$ 也是一个函数 $f: \mathbb{C} \rightarrow \mathbb{C}$ 。

最后，定义 Fun 是满足一定条件的函数的集合，于是所谓的泛函就是从集合 Fun 到 \mathbb{R} 或者 \mathbb{C} 的集合。如果我们说函数是一部机器，向它输入一个或者一组数字，将会得到一个数字，那么泛函也是一部机器，向它输入一个函数，将会得到一个数字。

我们举几个泛函的例子。如果定义 $Fun = \{f(x)|_{\substack{x \in \mathbb{R} \\ f(x) \in \mathbb{R}}} f(0) = f(1) = 0, f(x) \text{光滑}\}$ ，则我们可以定义 Fun 上的泛函 \mathcal{F} ，例如

$$\mathcal{F}[f(x)] = \int_0^1 f(x) dx \quad (4.3)$$

注意这里应当注意 Fun 的性质，例如 $f(x) = \frac{1}{x-0.5}$ 就不能作为这一泛函 \mathcal{F} 的输入。

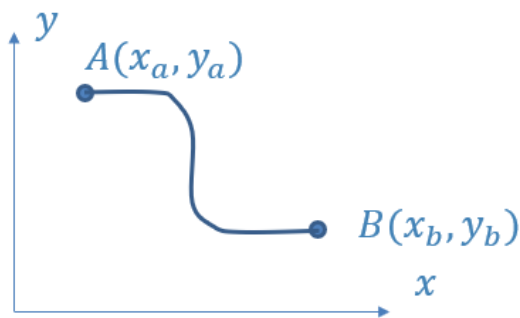


图 4.2: 两点之间的光滑函数

另外一个例子是我们定义 Fun 为点 $A(x_a, y_a), B(x_b, y_b)$ 之间光滑函数的集合

$$Fun = \{f(x)|y(x_a) = y_a, y(x_b) = y_b, y(x) \text{光滑}\} \quad (4.4)$$

如图 4.2 所示。于是长度泛函 \mathcal{L} 为

$$\mathcal{L}[y(x)] = \int_{x_a}^{x_b} \sqrt{1 + (y'(x))^2} dx \quad (4.5)$$

泛函可以视为无穷多自变量的函数，可以这样理解：例如对于 $Fun = \{f(x)|f(0) = f(1) = 0, f(x) \text{光滑}\}$ ，那么 $\forall f(x) \in Fun$ ，有

$$f(x) = \sum_{n=1}^{\infty} C_n \sin(n\pi x) \quad (4.6)$$

上式说明， $f(x)$ 的信息等价于无穷多个数字 (C_1, C_2, \dots) ，于是 Fun 上的泛函等价于以 C_1, C_2, \dots 为自变量的函数。如果是对于 $Fun = \{f(x)|x \in (-\infty, +\infty), f(x) \text{光滑}\}$ ，则我们可以想象， $f(x)$ 的信息等价于一个连续的数组 $\{f(x_1), f(x_2), \dots\}$ ，其中 x_1, x_2, \dots 取遍实轴（即，该数组元素个数“和实数一样多”）， Fun 上的泛函在图像上可以视为“以无穷多个数字 $\{f(x_1), f(x_2), \dots\}$ 为自变量的函数”。注意：将泛函视为无穷多变量的函数，更多的是用于“图像层面”的分析，进行严格的数学推导则需要小心，简单地把普通函数的性质拿给泛函，有时候对，有时候不对。

从上述分析可见，泛函是函数概念的推广，因此可以借鉴函数的分析方法研究泛函问题。

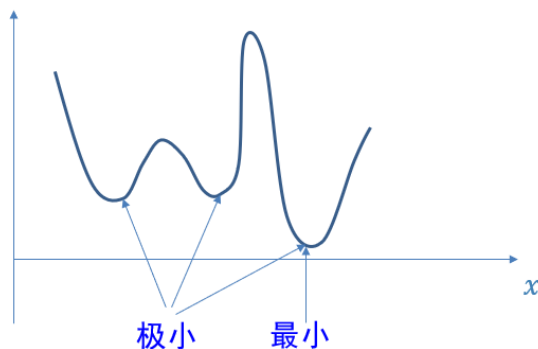


图 4.3: 函数的极小值与最小值的图示

4.2 泛函极值与欧拉-拉格朗日方程

4.2.1 泛函极值

我们知道, 对于函数而言, 函数的最小点是取值比其它点都要小的点, 而极小点是取值比它邻域内其它点都要小的点, 如图 4.3 所示。

类似地可以说明泛函最小值和极小值的含义。设 \mathcal{F} 是 Fun 上的泛函, 则 \mathcal{F} 在 $f_0(x)$ 上取最小值, 就意味着 $\forall f(x) \in Fun$, 都有 $\mathcal{F}[f_0(x)] \leq \mathcal{F}[f(x)]$ 。而 \mathcal{F} 在 $f_0(x)$ 上取极小值则意味着:

对任意 $f(x) \in Fun$ 且 $f(x)$ 在 $f_0(x)$ “附近” (即 $f(x)$ 与 $f_0(x)$ 的形状接近), 都有

$$\mathcal{F}[f_0(x)] \leq \mathcal{F}[f(x)]$$

而变分法则是用来求泛函极值的方法。

4.2.2 端点固定的极值问题

我们已知, 若 $f(x)$ 在 x_0 极小, 那么必要条件是 $\left. \frac{df(x)}{dx} \right|_{x=x_0} = 0$ 。这一必要条件经常用来寻找极值点, 因为极小点的候选人必然是满足这个条件的点, 而这样的点通常不多, 因此利用这一条件可以筛选极小点的范围。现在, 我们的问题是上面的必要条件, 如何推广到泛函极值问题?

我们考虑函数集合 $Fun = \{f(x) | f(x_a) = f_a, f(x_b) = f_b, x \in [x_a, x_b], f(x) \text{ 光滑} \}$, 它是在 x_a, x_b 两点取值被固定的光滑函数 (“端点固定” 的函数), 如图 4.4 所示。假设该集合上的某个泛函被定义为

$$\mathcal{F}[f(x)] = \int_{x_a}^{x_b} F[f(x), f'(x), x] dx \quad (4.7)$$

其中 $f'(x) = \frac{df(x)}{dx}$, $F(\alpha, \beta, \gamma)$ 是三元函数。我们的问题是: 如果 \mathcal{F} 在 $f_0(x)$ 取到极小, 那么 $f_0(x)$ 应当满足什么必要条件? 即如何把 “函数在极小点导数为零” 推广到这一泛函?

我们的方法是将泛函极值问题变为函数极值问题。引入一个新的光滑函数 $\delta f(x)$, 它满足 $\delta f(x_a) = \delta f(x_b) = 0$ (注意 $\delta f(x)$ 不是无穷小)。若 a 是实数, 则显然 $f_0(x) + a\delta f(x)$ 仍然是

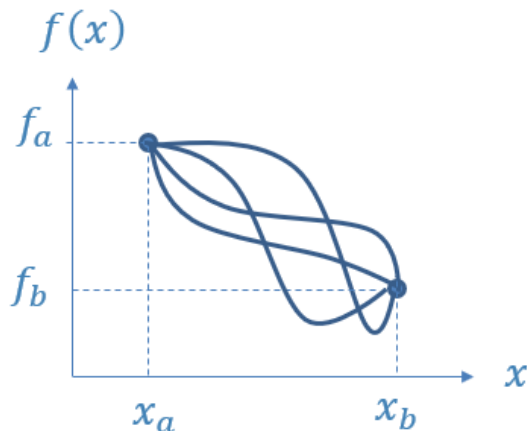


图 4.4: 端点固定的若干光滑函数

Fun 集合中的元素。于是, 对于任意给定的 $f_0(x)$ 和 $\delta f(x)$, 只要 $|a|$ 足够小, 则 $f_0 + a\delta f(x)$ 就“很接近” $f_0(x)$ 。

定义

$$G[a] = F[f_0(x) + a\delta f(x)] \quad (4.8)$$

若 \mathcal{F} 在 $f_0(x)$ 极小, 那么在 $|a|$ 很小时, 有 $\mathcal{F}[f_0(x)] < \mathcal{F}[f_0(x) + a\delta f(x)]$, 也就是 $G[0] < G[a]$ 。所以 \mathcal{F} 在 $f_0(x)$ 取极小, 就意味着 $G[a]$ 在 $a = 0$ 取极小, 它的必要条件就是 $\left. \frac{dG[a]}{da} \right|_{a=0} = 0$ 。即: \mathcal{F} 在 $f_0(x)$ 取极小的必要条件是

$$\left. \frac{dG[a]}{da} \right|_{a=0} = 0 \quad (4.9)$$

根据定义, 我们有

$$G[a] = \int_{x_a}^{x_b} F[f_0(x) + a\delta f(x), f'_0(x) + a\delta f'(x), x] dx \quad (4.10)$$

这里 $\delta f'(x) = \frac{d\delta f(x)}{dx}$ 。于是

$$\left. \frac{dG[a]}{da} \right|_{a=0} = \int_{x_a}^{x_b} \left[\frac{\partial F}{\partial f_0} \delta f(x) + \frac{\partial F}{\partial f'_0} \delta f'(x) \right] dx \quad (4.11)$$

这里 $\frac{\partial F}{\partial f_0}, \frac{\partial F}{\partial f'_0}$ 都是 x 的函数, 它们被定义为

$$\frac{\partial F}{\partial f_0} \equiv \left. \frac{\partial F(\alpha, \beta, \gamma)}{\partial \alpha} \right|_{\substack{\alpha=f_0(x) \\ \beta=f'_0(x) \\ \gamma=x}} \quad (4.12)$$

$$\frac{\partial F}{\partial f'_0} \equiv \left. \frac{\partial F(\alpha, \beta, \gamma)}{\partial \beta} \right|_{\substack{\alpha=f_0(x) \\ \beta=f'_0(x) \\ \gamma=x}} \quad (4.13)$$

又因为

$$\int_{x_a}^{x_b} \frac{\partial F}{\partial f'_0} \delta f'(x) dx = \frac{\partial F}{\partial f'_0} \delta f(x) \Big|_{x=x_a}^{x=x_b} - \int_{x_a}^{x_b} \delta f(x) \frac{d}{dx} \left(\frac{\partial F}{\partial f'_0} \right) dx \quad (4.14)$$

因为

$$\delta f(x_a) = \delta f(x_b) = 0 \quad (4.15)$$

所以

$$\frac{\partial F}{\partial f'_0} \delta f(x) \Big|_{x=x_a}^{x=x_b} = 0$$

因此我们得到

$$\int_{x_b}^{x_a} \frac{\partial F}{\partial f'_0} \delta f'(x) dx = - \int_{x_a}^{x_b} \delta f(x) \frac{d}{dx} \left(\frac{\partial F}{\partial f'_0} \right) dx \quad (4.16)$$

代入(4.11)，我们得到

$$\frac{dG[a]}{da} \Big|_{a=0} = \int_{x_a}^{x_b} \left\{ \left[\frac{\partial F}{\partial f'_0} - \frac{d}{dx} \left(\frac{\partial F}{\partial f'_0} \right) \right] \delta f(x) \right\} dx \quad (4.17)$$

因此对任意满足 $\delta f(x_a) = \delta f(x_b) = 0$ 的光滑 $\delta f(x)$ ，都有

$$\frac{dG[a]}{da} \Big|_{a=0} \quad (4.18)$$

因此我们有

$$\frac{\partial F}{\partial f'_0} - \frac{d}{dx} \left(\frac{\partial F}{\partial f'_0} \right) = 0 \quad (4.19)$$

所以，定义于两端 (x_a 和 x_b) 取值固定的光滑函数集合 Fun 上的泛函

$$\mathcal{F}[f(x)] = \int_{x_a}^{x_b} F[f(x), f'(x), x] dx \quad (4.20)$$

在 $f_0(x)$ 处极小的必要条件是： $f_0(x)$ 必须满足下面的微分方程

$$\boxed{\frac{\partial F}{\partial f'_0} - \frac{d}{dx} \left(\frac{\partial F}{\partial f'_0} \right) = 0} \quad (4.21)$$

其中 $\frac{\partial F}{\partial f'_0} \equiv \frac{\partial F(\alpha, \beta, \gamma)}{\partial \alpha} \Big|_{\substack{\alpha=f_0(x) \\ \beta=f'_0(x) \\ \gamma=x}}, \frac{\partial F}{\partial f'_0} \equiv \frac{\partial F(\alpha, \beta, \gamma)}{\partial \beta} \Big|_{\substack{\alpha=f_0(x) \\ \beta=f'_0(x) \\ \gamma=x}}$ ，这一微分方程被称之为 Euler-Lagrange

方程 (EL 方程)

4.2.3 EL 方程的评述

需要注意几点。首先，需要注意泛函的“定义域” Fun 是两端固定的函数集合

$$Fun = \{f(x) | f(x_a) = f_a, f(x_b) = f_b, x \in [x_a, x_b], f(x) \text{光滑}\} \quad (4.22)$$

其次,满足 EL 方程是泛函取极值的必要但非充分条件。正如对于函数极值问题中,满足 $\left. \frac{df(x)}{dx} \right|_{x=x_0} = 0$ 可以意味着 $f(x)$ 在 x_0 极小,极大或者不是极值点,泛函极值问题中 $f_0(x)$ 满足 EL 方程可以意味着泛函 \mathcal{F} 在 f_0 取极小、极大值或者都不是,只能统称为“ \mathcal{F} 在 $f_0(x)$ 取平稳值”,图像上意味着 \mathcal{F} 在 $f_0(x)$ 处变化缓慢。但我们可以利用 EL 方程筛选泛函 \mathcal{F} 所有可能的“极值点”,然后利用其他信息一一判断。

最后,EL 方程存在其它的推导方法。例如可以将 Fun 杀光泛函视为无穷多个自变量的函数,则极值的必要条件为(不严格的表述)“对每个自变量的导数为零”,由此也可以推出 EL 方程。考虑到 Fun 形式为(4.22),因此所有 Fun 中的元素 $f(x)$ 可以写为

$$f(x) = f_a + f_b - f_a \left(\frac{x - x_a}{x_b - x_a} \right) + \sum_{n=1}^{\infty} C_n \sin \left[n\pi \frac{x - x_a}{x_b - x_a} \right] \quad (4.23)$$

即 $f(x)$ 的信息等价于无穷多个系数 C_1, C_2, \dots 的取值。因此泛函 $\mathcal{F}[f(x)] = \int_{x_a}^{x_b} F[f(x), f'(x), x] dx$ 可以视为无穷多个自变量 C_1, C_2, \dots 的函数。因此,如果泛函 \mathcal{F} 在 $f_0(x)$ (对应参数 C_{1*}, C_{2*}, \dots) 上取极值,则必要条件是 \mathcal{F} 对每个自变量 C_1, C_2, \dots 的导数(也是 C_1, C_2, \dots 的函数)在 $C_1 = C_{1*}, C_2 = C_{2*}, \dots$ 的时候都为零。利用此条件,可以得到 C_{1*}, C_{2*}, \dots 所满足的方程,并由此反推出 $f_0(x)$ 满足的微分方程。这一方法在数学上不是百分百严格,但图像直观。

4.2.4 EL 方程的应用

A. 两点之间直线最短

我们考虑平面上两点 $(x_a, y_a), (x_b, y_b)$, 我们就有长度泛函

$$\mathcal{L}[y(x)] = \int_{x_a}^{x_b} \sqrt{1 + (y'(x))^2} dx \quad (4.24)$$

这意味着

$$F(\alpha, \beta, \gamma) = \sqrt{1 + \beta^2} \quad (4.25)$$

现在我们已知存在两点之间的最短曲线,需要给出曲线方程。我们假设 \mathcal{L} 在 $y_0(x)$ 最小,那么它应当满足 EL 方程,即

$$\frac{d}{dt} \frac{\partial F}{\partial y'_0} - \frac{\partial F}{\partial y_0} = 0 \quad (4.26)$$

这给出

$$y''_0(x) \left[\frac{1}{(1 + (y'_0(x))^2)^{3/2}} \right] = 0 \quad (4.27)$$

从而

$$y''_0(x) = 0 \quad (4.28)$$

因此 $y_0(x)$ 是连接 AB 的直线。

B. 最速降线

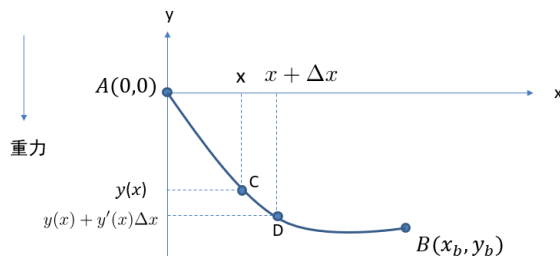


图 4.5: 一种可能的从 A 到 B 的降线。

如图 4.5 所示, 假设质点被约束在形状为 $y(x)$ 的光滑曲线槽上运动, 除约束力外只受重力。初始 $t = 0$ 时刻质点在 A 点, 速度为零。试问曲线 $y(x)$ 是什么形状的时候, 质点能用最短的时间滑到 B 点?

我们考察质点从横坐标 x 到横坐标 $x + \Delta x$ (图中 C, D 两点) 的运动, 此时所走的路程为 $\sqrt{1 + (y'(x))^2} \Delta x + O(\Delta x^2)$, 速度为 $\sqrt{2g|y(x)|} + O(\Delta x)$, 于是运动的时间为

$$\frac{\sqrt{1 + (y'(x))^2}}{\sqrt{2g|y(x)|}} \Delta x + O(\Delta x^2) \quad (4.29)$$

因此, 从 A 到 B 的总时间是关于曲线形状 $y(x)$ 的如下泛函

$$T[y(x)] = \int_0^{x_b} F[y(x), y'(x), x] dx \quad (4.30)$$

其中

$$F[\alpha, \beta, \gamma] = \frac{\sqrt{1 + \beta^2}}{\sqrt{-2g\alpha}} \quad (4.31)$$

因此使得质点在最短时间内滑落的曲线, 应当满足

$$-\frac{d}{dx} \frac{\partial F}{\partial y_0} + \frac{\partial F}{\partial y_0} = 0 \quad (4.32)$$

4.2.5 EL 方程推广到多个函数的情况

考虑“两端固定的函数族”构成的集合

$$Fun = \{y_1(x), y_2(x), \dots, y_N(x) | y_i(x_a) = y_{ia}, y_i(x_b) = y_{ib}, y_i(x) : \mathbb{R} \rightarrow \mathbb{R} \text{ 并且是光滑函数} \} \quad (4.33)$$

假设 Fun 上有泛函

$$\mathcal{F}[y_1(x), \dots, y_N(x)] = \int_{x_a}^{x_b} F[y_1(x), \dots, y_N(x); y'_1(x), \dots, y'_N(x), x] dx$$

这里 $F(\alpha_1, \dots, \alpha_N; \beta_1, \dots, \beta_N; \gamma)$ 是一个 $(2N+1)$ 元函数, 则该泛函在函数组 $[y_{10}(x), \dots, y_{N0}(x)]$ 处取极值的必要条件为: 函数组 $[y_{10}(x), \dots, y_{N0}(x)]$ 满足方程组

$$\frac{\partial F}{\partial y_{i0}} - \frac{d}{dx} \frac{\partial F}{\partial y'_{i0}} = 0 \quad i = 1, \dots, N \quad (4.34)$$

这里

$$\frac{\partial F}{\partial y_{i0}} \equiv \left. \frac{\partial F(\alpha_1, \dots, \alpha_N; \beta_1, \dots, \beta_N; \gamma)}{\partial \alpha_i} \right|_{\substack{\alpha_j = y_{j0}(x), (j=1, \dots, N) \\ \beta_j = y'_{j0}(x), (j=1, \dots, N) \\ \gamma = x}} \quad (4.35)$$

$$\frac{\partial F}{\partial y'_{i0}} \equiv \left. \frac{\partial F(\alpha_1, \dots, \alpha_N; \beta_1, \dots, \beta_N; \gamma)}{\partial \beta_i} \right|_{\substack{\alpha_j = y_{j0}(x), (j=1, \dots, N) \\ \beta_j = y'_{j0}(x), (j=1, \dots, N) \\ \gamma = x}} \quad (4.36)$$

4.3 Hamilton 原理

4.3.1 位形空间的 Hamilton 原理与 Lagrange 方程

以单个自由度力学系统为例, 假设有一个此类系统, 广义坐标为 q , 拉格朗日量为 $L[q, \dot{q}, t]$, 我们定义函数集合 $Fun = \{q(t) | q(t_a) = q_a, q(t_b) = q_b, q(t) : \mathbb{R} \rightarrow \mathbb{R} \text{ 并且光滑}\}$, 它给出 t_a 时刻从 q_a 出发, t_b 时刻到达 q_b 的所有数学上可能的运动轨道 (轨迹), 这些轨迹中只有一条满足 Lagrange 方程, 是系统的真实轨道, 如图 4.6 所示。

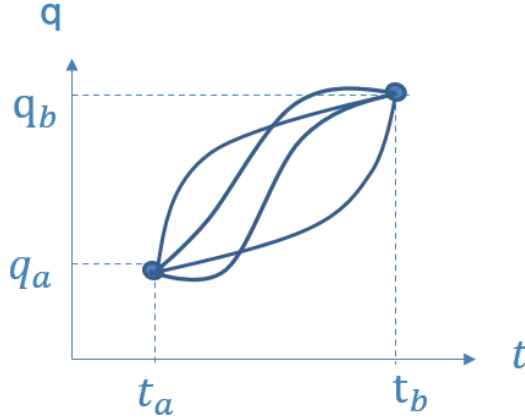


图 4.6: 从 q_a 到 q_b 的一系列可能运动轨道, 它们被包含在函数集合 Fun 中

我们定义 Fun 上的作用量泛函

$$S[q(t)] = \int_{t_a}^{t_b} L[q(t), \dot{q}(t), t] dt \quad (4.37)$$

若 S 在 $q_0(t)$ 取平稳值, 那么 $q_0(t)$ 就满足

$$\frac{d}{dt} \left. \frac{\partial}{\partial \beta} \frac{\partial L(\alpha, \beta, t)}{\partial \beta} \right|_{\substack{\alpha = q_0(t) \\ \beta = \dot{q}_0(t)}} - \left. \frac{\partial L(\alpha, \beta, t)}{\partial \alpha} \right|_{\substack{\alpha = q_0(t) \\ \beta = \dot{q}_0(t)}} = 0 \quad (4.38)$$

显然这一方程即为理论力学给出的该系统的拉格朗日方程, 从而 $q_0(t)$ 即为系统的真实轨迹。由此, 我们可以给出在位形空间的 Hamilton 原理

定理 4.3.1 (位形空间的 Hamilton 原理). 在所有满足 $q(t_a) = q_a, q(t_b) = q_b$ 的轨道中, 系统的真实轨道就是让作用量泛函 $S[q(t)]$ 取平稳值的那一条。

这是一个物理规律，是拉格朗日方程的一个推论，它对任意多自由度系统都成立。对于多自由度系统，可以做如下推广：对于一个具有 N 个广义坐标 $\{q_1, q_2, \dots, q_N\}$ 的系统，我们把由所有 N 维数组 (q_1, \dots, q_N) 构成的集合，称作这个系统的**位形空间**。那么这个系统在位形空间的一条运动轨迹，就可以用 N 个函数 $(q_1(t), \dots, q_N(t))$ 来描述。我们现在考虑所有在 t_0 时刻和 t_1 时刻具有确定位置的轨迹构成的集合，即集合

$$Fun = \{(q_1(t), \dots, q_N(t)) | q_i(t) \text{光滑}, q_i(t_a) = q_{ia}, q_i(t_b) = q_{ib}, i = 1, \dots, N\} \quad (4.39)$$

进而，我们定义这个集合上的作用量泛函

$$S[q_1(t), \dots, q_N(t)] = \int_{t_a}^{t_b} L(q_1(t), \dots, q_N(t); \dot{q}_1(t), \dots, \dot{q}_N(t); t) dt \quad (4.40)$$

这里 $L(q_i; \dot{q}_i; t)$ 是系统的拉格朗日量。我们注意，现在 Fun 是满足条件 $(q_i(t_a) = q_{ia}, q_i(t) = q_{ib}; i = 1, \dots, N)$ 的所有可能的轨迹的集合，这里面当然包含了拉格朗日方程所决定的系统的真实轨迹，也包含了不满足拉格朗日方程的轨迹。试问 Fun 上的作用量泛函 S 什么时候取平稳值？显然，由于集合 Fun 由端点固定的轨道组成，因此，让作用量泛函 S 取平稳值的轨道（我们记为 $(q_{10}(t), \dots, q_{N0}(t))$ ）必然满足 EL 方程

$$\frac{d}{dt} \left(\frac{\partial L[q_1(t), \dots, q_N(t); \dot{q}_1(t), \dots, \dot{q}_N(t); t]}{\partial \dot{q}_{i0}(t)} \right) - \frac{\partial L[q_1(t), \dots, q_N(t); \dot{q}_1(t), \dots, \dot{q}_N(t); t]}{\partial q_{i0}} = 0 \quad i = 1, \dots, N \quad (4.41)$$

而这正是系统真实运动的轨道所满足的拉格朗日方程，**这意味着，在所有满足 $(q_i(t_a) = q_{ia}; q_i(t) = q_{ib}; i = 1, \dots, N)$ 的轨道中，系统真实运动的轨道，恰恰是使作用量泛函 S 取平稳值的轨道，这就是哈密顿原理在位形空间的表述。**

4.3.2 相空间中的 Hamilton 原理

仍然以单个自由度系统为例，假设有一个这样的系统，广义坐标为 q ，广义动量为 p ，哈密顿量为 $H(q, p, t)$ 。定义函数集合

$$Fun = \{(q(t), p(t)) | q(t_a) = q_a, p(t_a) = p_a; q(t_b) = q_b, p(t_b) = p_b; q(t), p(t) \text{是 } \mathbb{R} \rightarrow \mathbb{R} \text{ 上的光滑函数}\} \quad (4.42)$$

定义相空间作用量泛函

$$S_p[q(t); p(t)] = \int_{t_a}^{t_b} \tilde{L}[q(t), p(t), \dot{q}(t), t] dt \quad (4.43)$$

其中

$$\tilde{L}[q(t), p(t), \dot{q}(t), t] = p\dot{q} - H(q, p, t) \quad (4.44)$$

注意在这一定义中 $q(t), p(t)$ 彼此独立，因此一般来说 \tilde{L} 取值不等于系统的拉格朗日量 L 。若 S_p 在 $[q_0(t), p_0(t)]$ 取平稳值，则 $[q_0(t), p_0(t)]$ 必须满足的两个 EL 方程为

$$\frac{d}{dt} \frac{\partial \tilde{L}}{\partial \dot{q}_0} - \frac{\partial \tilde{L}}{\partial q_0} = 0 \quad (4.45)$$

$$\frac{d}{dt} \frac{\partial \tilde{L}}{\partial \dot{p}_0} - \frac{\partial \tilde{L}}{\partial p_0} = 0 \quad (4.46)$$

根据(4.44), 我们有

$$\frac{\partial \tilde{L}}{\partial \dot{q}_0} = p_0(t) \quad \frac{\partial \tilde{L}}{\partial q_0} = - \frac{\partial H(q, p, t)}{\partial q} \Big|_{q=q_0(t), p=p_0(t)} \quad (4.47)$$

于是(4.45)就变为

$$\dot{p}_0(t) = - \frac{\partial H(q, p, t)}{\partial q} \Big|_{q=q_0(t), p=p_0(t)} \quad (4.48)$$

类似地, 我们也有

$$\frac{\partial \tilde{L}}{\partial \dot{p}_0} = 0 \quad \frac{\partial \tilde{L}}{\partial p_0} = \dot{q}_0 - \frac{\partial H(q, p, t)}{\partial p} \Big|_{q=q_0(t), p=p_0(t)} \quad (4.49)$$

于是(4.46)就变成

$$\dot{q}_0(t) = \frac{\partial H(q, p, t)}{\partial p} \Big|_{q=q_0(t), p=p_0(t)} \quad (4.50)$$

于是两个 EL 方程(4.45)(4.46)就可以写为

$$\begin{cases} \dot{q}_0(t) = \frac{\partial H(q, p, t)}{\partial p} \Big|_{q=q_0(t), p=p_0(t)} \\ \dot{p}_0(t) = - \frac{\partial H(q, p, t)}{\partial q} \Big|_{q=q_0(t), p=p_0(t)} \end{cases} \quad (4.51)$$

显然这两个方程就是理论力学给出的该系统的哈密顿方程, 从而 $q_0(t), p_0(t)$ 就是系统在相空间中的真实轨迹。于是我们给出在相空间中的哈密顿原理

定理 4.3.2 (相空间中的哈密顿原理). 在所有 t_a 时刻在 (q_a, p_a) , t_b 时刻在 (q_b, p_b) 的相空间轨道中, 系统的真实轨道就是让相空间作用量泛函 $S_p[q(t), p(t)]$ 取平稳值的那一条。

这是一个物理规律, 是哈密顿方程的一个推论。同样, 这一原理有多自由度版本: 我们考虑由一个具有 N 个自由度系统, 我们知道, 相空间是这个系统所有广义坐标和广义动量可能取值构成的集合, 也就是所有形如 $(q_1, \dots, q_N, p_1, \dots, p_N)$ 的 $2N$ 维数组构成的集合, 跟前面类似, 那么这个系统在相空间的一条运动轨迹, 就可以用 $2N$ 个函数 $(q_1(t), \dots, q_N(t), p_1(t), \dots, p_N(t))$ 来描述, 我们现在考虑所有在 t 时刻和 t 时刻具有确定位置的轨迹构成的集合, 即集合

$$Fun = \left\{ (q_1(t), \dots, q_N(t), p_1(t), \dots, p_N(t)) \mid \begin{array}{l} q_i(t), p_i(t) \text{光滑}; q_i(t_a) = q_{ia}, p_i(t_a) = p_{ia} \\ q_i(t_b) = q_{ib}, p_i(t_b) = p_{ib}; i=1, \dots, N \end{array} \right\} \quad (4.52)$$

我们注意, 这时候集合 Fun 中的每个元素, 也就是相空间中的每条可能的轨迹, 都由 $2N$ 个函数构成, 进而, 我们定义这个集合上的作用量泛函, 也就是相空间中的作用量泛函

$$S_p[q_i(t); p_i(t)] = \int_{x_a}^{x_b} \tilde{L}[q_i(t); p_i(t); \dot{q}_i(t); t] dt \quad (4.53)$$

这里我们有 \tilde{L} 的定义为

$$\tilde{L}[q_i; p_i; \dot{q}_i; t] = \left(\sum_{i=1}^N p_i \dot{q}_i \right) - H(q_i; p_i; t) \quad (4.54)$$

在相空间中所有满足 $(q_i(t_a) = q_{ia}; q_i(t_b) = q_{ib}; p_i(t_a) = p_{ia}; p_i(t_b) = p_{ib}; i = 1, \dots, N)$ 的轨道中，系统真实运动的轨道，恰恰是使相空间中作用量泛函 S_p 取平稳值的轨道，这就是哈密顿原理在相空间的表述。

值得注意的是，存在这样的可能性，即在所有 t_a 时刻在 (q_a, p_a) , t_b 时刻在 (q_b, p_b) 的相空间轨道中，没有一条轨道能让相空间作用量泛函 $S_p[q(t), p(t)]$ 取平稳值 (即，没有一条轨道能满足 Hamiltonian 方程，或者说没有一条轨道是“真实的”)。这对应于如下情况：以 (q_a, p_a) 作为 t_a 时刻的初始条件，求解 Hamiltonian 方程，解出系统在 t_b 时刻的相空间坐标不等于 (q_b, p_b) 。事实上，相空间的哈密顿原理可以修改为：在所有 t_a 时刻坐标为 q_a , t_b 时刻坐标为 q_b 的相空间轨道中，系统的真实轨道就是让相空间作用量泛函 $S_p[q(t), p(t)]$ 取平稳值的那一条。

最后我们对哈密顿原理做一些评述。首先，Hamilton 原理通常又称为“最小作用原理”，但事实上作用量泛函未必在真实轨道取最小值¹。大多数物理学的基本领域 (量子力学、电动力学、广义相对论、几何光学等等) 都有类似 Hamilton 原理的推论，即，系统物理量满足的运动方程，与让某个作用量泛函取平稳值的 EL 方程相同。另一方面，我们既可以把“最小作用原理”作为基本假定，推出运动方程，也可以把运动方程作为基本假定，最小作用原理当做推论。两种观点数学上等价，但很多人接受后一种观点，因为运动方程来源于实验，是更本质的。最后，Hamilton 原理在量子力学中对应 Feynman 的路径积分表述：设 q 为直角坐标，系统无约束，若质点在 t_a 时刻位于 q_a ，则 t_b 时刻质点位于 q_b 的几率为

$$P(t_b, q_b \leftarrow t_a, q_a) = \frac{1}{N} \left| \sum_{q(t) \in Fun} e^{iS[q(t)]} \right|^2 \quad (4.55)$$

这里 N 是一个归一化系数。

4.4 求泛函极值的其他方法

4.4.1 Ritz 变分法—缩小搜索范围

我们的问题是：如果已知函数集合 Fun 上有泛函 \mathcal{F} ，求 Fun 中令 \mathcal{F} 取最小值的函数 $f_0(x)$ 。我们的方法是，先猜测 $f_0(x)$ 近似具有形式 $f(a; x)$ ，其中 a 是一个或一组待定参数，从而函数 $f(a; x)$ 的形式已经确定。接下来，求出让函数 $F(a) = \mathcal{F}[f(a; x)]$ 取最小值的 a ，记为 a_* ，则

$$f_0(x) \approx f(a_*; x) \quad (4.56)$$

并且泛函 \mathcal{F} 的最小值近似为 $\mathcal{F}[f(a_*; x)]$ 。

Ritz 变分法是一种近似方法。其核心在于“搜索范围”从整个函数集合 Fun 缩小到具有特定形式 $f(a; x)$ 的函数的集合，从而简化计算，将泛函极值问题转化为函数极值问题。该近似结果的好坏取决于对 $f_0(x)$ 的形式猜的准不准，即准确的 $f_0(x)$ 是否真的具有 (或者近似具有) 形式 $f(a; x)$ 。

¹ 参见戈德斯坦《经典力学》中关于哈密顿原理的讨论

我们举一个例子。考虑函数集合

$$Fun = \left\{ \psi(x) | x \in \mathbb{R}, \int_{-\infty}^{\infty} |\psi(x)|^2 dx = 1 \text{ 且 } \psi(x) \text{ 及其导数都连续} \right\} \quad (4.57)$$

中的泛函 $\varepsilon[\psi(x)] = \int_{-\infty}^{\infty} \psi^*(x) \left(-\frac{1}{2} \frac{\partial^2}{\partial x^2} + \frac{x^2}{2} \right) \psi(x) dx$, $\psi(x) \in Fun$ 在函数 $\psi_0(x)$ 取极小值, 要求 $\psi_0(x)$ 。我们可以猜测 $\psi_0(x)$ 具有形式

$$\psi(x) = \left(\frac{2c}{\pi} \right)^{1/4} e^{-cx^2} \quad (4.58)$$

这里 c 是参数, 且 $c > 0$ 并且 c 是实数。显然, 泛函 ε 在此类函数上的取值为

$$\varepsilon \left[\left(\frac{2c}{\pi} \right)^{1/4} e^{-cx^2} \right] = \frac{1}{4c} + c \quad (4.59)$$

它可以视为 c 的函数, 并且在 $c = 1/2$ 取最小值。因此我们就有

$$\psi_0(x) = \left(\frac{1}{\pi} \right)^{1/4} e^{-x^2/2} \quad (4.60)$$

4.4.2 Lagrange 不定乘子法

Lagrange 不定乘子法是求条件极值的方法, 它扩大了搜索范围。

首先我们考虑它在函数条件极值问题上的应用。我们的问题是求函数 $f(x, y)$ 在满足 $g(x, y) = 0$ 的曲线上的极小值, 我们不能把 f 直接对 x, y 求偏导数, 因为在曲线上的极小点未必是平面上的极小点。我们的方法是, 定义 $f^*(x, y, \lambda) = f(x, y) - \lambda g(x, y) = 0$, 并求解如下方程组

$$\begin{cases} \frac{\partial f^*(x_0, y_0, \lambda_0)}{\partial x_0} = 0 \\ \frac{\partial f^*(x_0, y_0, \lambda_0)}{\partial y_0} = 0 \\ g(x_0, y_0) = 0 \end{cases} \quad (4.61)$$

若接触的 (x_0, y_0, λ_0) 刚好令 $f^*(x_0, y_0, \lambda_0)$ 在点 (x_0, y_0) 处取极小值, 则 (x_0, y_0) 是函数 $f(x, y)$ 在满足 $g(x, y) = 0$ 的曲线的极小点, 这里 λ_0 就被称为 Lagrange 乘子。

可以理解该结果的正确性。首先, (x_0, y_0) 一定是位于曲线 $g(x, y) = 0$ 上。由于函数 $f^*(x, y, \lambda_0)$ 在点 (x_0, y_0) 处取极小值, 则 $f^*(x, y, \lambda_0)$ 在点 (x_0, y_0) 的取值比平面上在其周围所有点 (如图 4.7 中黄色区域所示, 注意其中包括位于曲线 $g(x, y) = 0$ 上的点) 的取值都小。函数 $f^*(x, y, \lambda_0)$ 在曲线 $g(x, y) = 0$ 上自动变为 $f(x, y)$ 。因此推论是: $f(x, y)$ 在 (x_0, y_0) 的取值, 比 $f(x, y)$ 在曲线 $g(x, y) = 0$ 上 (x_0, y_0) 附近的点 (图 4.7 中红色线段所示) 的取值都小。

这一方法同样可以用于泛函的条件极值。如果已知函数集合 Fun 上有泛函 \mathcal{F}, \mathcal{G} , 要求让 \mathcal{F} 在 Fun 中满足 $\mathcal{G}[f(x)] = 0$ 的子集中取极小值的函数 $f_0(x)$ 。我们的办法是定义

$$\mathcal{F}^*[f(x), \lambda] = \mathcal{F}[f(x)] + \lambda \mathcal{G}[f(x)] \quad (4.62)$$

设函数 $f_0(x)$ 与数字 λ_0 同时满足如下两个条件

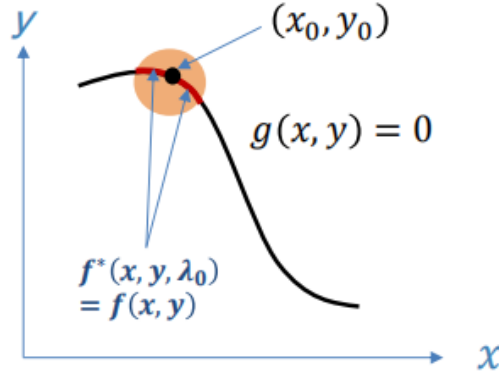


图 4.7: Lagrange 不定乘子法所给出的极值

- 在 Fun 中, 泛函 $F^*[f(x), \lambda_0]$ 在 $f_0(x)$ 处取极小值。
- $\mathcal{G}[f_0(x)] = 0$

Example1 最大熵 考虑一个自由度为 1 的系统, 坐标和动量为 q, p , 哈密顿为 $H(q, p)$ 。系统在相空间的概率密度分布为 $\rho(q, p)$, 则系统的熵为 $\rho(q, p)$ 的泛函

$$S[\rho(q, p)] = - \int \rho(q, p) \ln \rho(q, p) dp dq \quad (4.63)$$

系统的平均能量也为 $\rho(q, p)$ 的泛函

$$E[\rho(q, p)] = \int \rho(q, p) H(q, p) dp dq \quad (4.64)$$

若存在密度分布 $\rho_0(q, p)$ 能在保证平均能量为 E_0 的情况下做到熵最大, 这一分布 $\rho_0(q, p)$ 是什么?

这一问题相当于要在如下两个约束下求泛函的极值

$$E[\rho(q, p)] = \int \rho(q, p) H(q, p) dp dq = E_0 \quad (4.65)$$

$$\int \rho(q, p) dp dq = 1 \quad (4.66)$$

第二条是概率归一化的要求。因此使用拉格朗日不定乘子法, 定义泛函

$$\begin{aligned} S_*[\rho(q, p), \alpha, \beta] &= S - (\alpha + 1) \left(1 - \int \rho(q, p) dp dq \right) - \beta (E[\rho(q, p)] - E_0) \\ &= \int [-\rho(q, p) \ln \rho(q, p) + (\alpha + 1) \rho(q, p) - \beta \rho(q, p) H(q, p)] dp dq - (\alpha + 1) + \beta E_0 \end{aligned} \quad (4.67)$$

使得这一泛函 S_* 在 $\rho_0(q, p)$ 取极值, 这就导致 $\rho(q, p) = e^{-\beta H(q, p)}$ (为什么?), 即统计物理中的正则分布。参数 α, β 由条件(4.65)(4.66)确定。 α 的物理意义为归一化系数 (即配分函数), β 的物理意义为 $\frac{1}{k_B T}$, 这里 T 是温度, k_B 为 Boltzmann 常数。

第五章 刚体的运动

5.1 引言

所谓的刚体，它的本质是一个质点组，但这一质点组满足

- 任意两质点间距不变。记刚体中任意两个质点 i, j 的位置为 $\mathbf{r}_i, \mathbf{r}_j$ ，它们的相对位置记为 $\mathbf{r}_{ij} \equiv \mathbf{r}_i - \mathbf{r}_j$ ，则 $|\mathbf{r}_{ij}|$ 不随时间变化 (约束系统)。
- 满足牛顿第三定律强表述。对于刚体中的任意两质点 i, j ，记 \mathbf{F}_{ij} 为 j 给 i 的内力，而 \mathbf{F}_{ji} 为 i 给 j 的内力，则一定有

$$\mathbf{F}_{ji} = -\mathbf{F}_{ij} \quad (5.1)$$

并且 \mathbf{F}_{ji} 与 \mathbf{r}_{ij} 共线 (平行或反平行)

刚体的所有动力学性质都可以从质点组力学规律和以上两个性质严格给出。从上述两个性质可以得到一个推论：对于任意刚体中的质点 i, j ，我们记 $\mathbf{v}_i = \dot{\mathbf{r}}_i, \dot{\mathbf{v}}_j = \dot{\mathbf{r}}_j, \mathbf{v}_{ij} = \mathbf{v}_i - \mathbf{v}_j = \dot{\mathbf{r}}_{ij}$ ，则一定有 \mathbf{F}_{ij} 与 \mathbf{v}_{ij} 垂直 (如何证明?)，从而刚体的内力一定不做功。因此，刚体 (总的、质心的、质心系中总的) 动量、角动量、动能变化均与内力无关。

刚体的转动 (自转) 设 $O - xyz$ 为惯性系。若刚体运动中，存在一个与刚体固连的刚性轴 l ，其与 $O - xyz$ 系的某一个或某几个坐标轴的夹角随时间变化，则称刚体存在**转动**。若刚体中任意与刚体固连的刚性轴与 $O - xyz$ 的所有坐标轴夹角都不随时间变化，则称刚体无转动 (即，做平动)。刚体是否转动，与惯性系 $O - xyz$ 的选择无关。刚体力学研究的核心就是刚体的转动。

角动量 任何情况下，刚体中任意 A, B 两点，其惯性系中相对速度 $\mathbf{v}_{AB} = \mathbf{v}_A - \mathbf{v}_B$ 和相对位置 $\mathbf{r}_{AB} = \mathbf{r}_A - \mathbf{r}_B$ 均满足

$$\mathbf{v}_{AB} = \boldsymbol{\omega} \times \mathbf{r}_{AB} \quad (5.2)$$

这里 $\boldsymbol{\omega}$ 为角速度矢量，它的取值 (包括方向和大小) 都和 A, B 的选择无关。即对于一个刚体，每时每刻只有一个 $\boldsymbol{\omega}$ 。 $\boldsymbol{\omega}$ 是刻画刚体转动快慢和方向的量，它可以被视为“质点速度”概念的推广。在定轴转动和定点运动情况下，如果左边原点选在转动轴或者固定点 (即坐标原点是刚体上的一个点)，则这一关系(5.2)进一步变为

$$\mathbf{v}_A = \boldsymbol{\omega} \times \mathbf{r}_A \quad (5.3)$$

研究思路 对于刚体的运动，有三种典型情况，分别是定轴转动、一般平面运动和定点运动。其中定轴转动是最简单也最基本的情况，其他情况均可以分解为各种定轴转动。在牛顿力学下，我们的思路是

1. 建立和刚体固连的坐标系 (不是参考系)
2. 确定广义坐标 q_1, \dots, q_n ，最多六个
3. 推导角动量矢量和广义坐标、广义速度的关系 $\omega(q_1 \dots q_n, \dot{q}_1, \dots, \dot{q}_n)$
4. 推导某种角动量 (矢量，或者 z 分量) 和角速度的关系 $L[\omega(q_1 \dots q_n, \dot{q}_1 \dots \dot{q}_n)]$ (其中出现转动惯量或者惯量张量)，从而可以把角动量表示为广义坐标、广义速度的函数
5. 动力学: 利用角动量定理 $\frac{dL}{dt} = M$, 其中 M 为与 L 对应的外力力矩, 建立 $q_1 \dots q_n, \dot{q}_1 \dots \dot{q}_n, \ddot{q}_1 \dots \ddot{q}_n$ 与 M 的联系，即刚体动力学方程

最后，利用拉格朗日力学，正确写出刚体总动能作为 $q_1 \dots q_n, \dot{q}_1 \dots \dot{q}_n$ 的函数，从而写出拉格朗日量，注意刚体内力是约束力。

最后再次指出，要注意刚体的“两面性”。刚体既是由无数质点组成的质点组，也是少数自由度的整体，在处理问题时有时要灵活运动这两个视角。

5.2 刚体的定轴转动

5.2.1 定轴转动的运动学

定轴转动的定义。

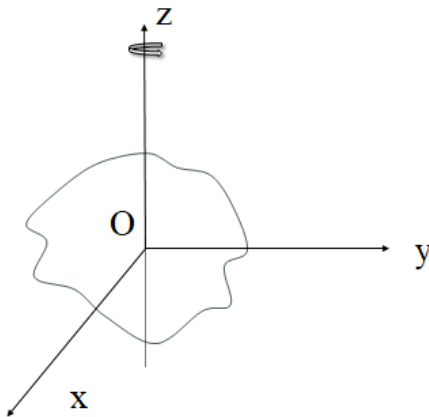


图 5.1: $O - xyz$ 为惯性系。一个刚体绕 z 轴做定轴转动。

如图5.1所示，设 $O - xyz$ 为惯性系。刚体绕 z 轴做定轴转动定义如下：

若在刚体运动中，其上任何一点的：

(1) z 坐标;

(2) 到 z 轴的距离,

都不随时间变化, 则称该刚体绕 z 轴做定轴转动 (或称刚体绕 z 轴转动)。

根据上述定义, 若刚体绕 z 轴转动, 则 z 轴上任意一点到刚体上任意一点的距离都不随时间改变。所以 z 轴 (含原点) 都可视为刚体的一部分。

与刚体固连的坐标系。

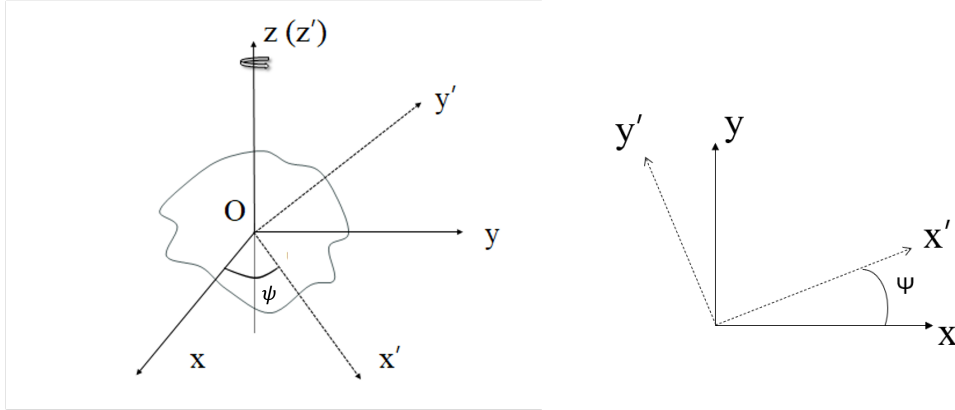


图 5.2: (左): 绕 z 轴转动的刚体。 $O-xyz$ 为惯性系。 $O-x'y'z'$ 为和刚体固连的坐标系。(右): 从 z 轴上方俯视看到的 Oxy 平面的各个坐标轴。

如图5.2左边所示, 我们建立与刚体固连的坐标系 $O-x'y'z'$ 。该坐标系原点和惯性系相同, 也是 O , 且 Oz' 轴与惯性系的 Oz 轴重合。 $O-x'y'z'$ 系的三个坐标轴都焊接在刚体上, 和刚体一起运动。

显然, Ox' 轴与 Oy' 轴也都位于 Oxy 平面。我们进一步定义: ψ 为 Ox' 轴与 Ox 轴的夹角。在刚体运动中 ψ 随时间变化。

记 $O-xyz$ 的坐标基矢为 $e_{x,y,z}$, $O-x'y'z'$ 的坐标基矢为 $e_{x',y',z'}$ 。如图5.2右边所示, 这两组基矢满足如下变换关系:

$$\begin{cases} e_{x'} = \cos \psi e_x + \sin \psi e_y \\ e_{y'} = -\sin \psi e_x + \cos \psi e_y \\ e_{z'} = e_z \end{cases} \quad (5.4)$$

定轴转动刚体的广义坐标。

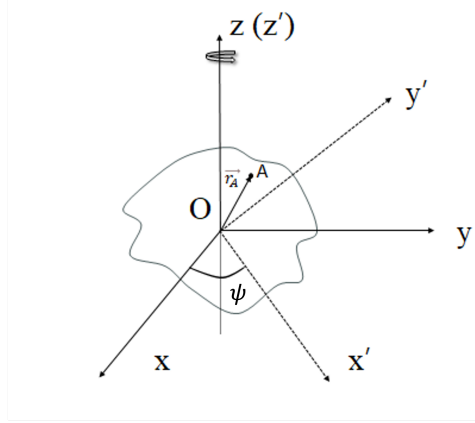


图 5.3: 刚体上某点 A 的位置矢量。

我们现在确定定轴转动刚体的广义坐标。为此，考虑刚体上任意一点 A (见图5.3)，其位置矢量 \mathbf{r}_A 可用 $O - x'y'z'$ 系的坐标基矢分解，写成

$$\mathbf{r}_A = \alpha \mathbf{e}_{x'} + \beta \mathbf{e}_{y'} + \gamma \mathbf{e}_{z'}. \quad (5.5)$$

其中 $\alpha = \mathbf{r}_A \cdot \mathbf{e}_{x'}$, $\beta = \mathbf{r}_A \cdot \mathbf{e}_{y'}$, $\gamma = \mathbf{r}_A \cdot \mathbf{e}_{z'}$ 。由于 \mathbf{r}_A 与 $\mathbf{e}_{x',y',z'}$ 都和刚体固连，因此它们各自的长度以及彼此夹角都不随时间变化，所以它们彼此的点乘，即 $\alpha \beta \gamma$ ，也都不随时间变化。而 $\mathbf{e}_{x'}, \mathbf{e}_{y'}, \mathbf{e}_{z'}$ 都是 ψ 的函数。因此，一旦 ψ 确定了，则刚体上每点的位置就确定了。因此我们有如下结论：

- ψ 是定轴转动的刚体的广义坐标。
- 定轴转动的刚体只有一个自由度。

定轴转动刚体的角速度矢量。

我们定义定轴转动刚体的角速度矢量 $\boldsymbol{\omega}$ 为：

$$\boldsymbol{\omega} = \dot{\psi} \mathbf{e}_z = \dot{\psi} \mathbf{e}_{z'}. \quad (5.6)$$

由基矢变换 (5.4) 可以得到：

$$\begin{cases} \dot{\mathbf{e}}_{x'} = -\dot{\psi} \sin \psi \mathbf{e}_x + \dot{\psi} \cos \psi \mathbf{e}_y = \boldsymbol{\omega} \times \mathbf{e}_{x'} \\ \dot{\mathbf{e}}_{y'} = -\dot{\psi} \cos \psi \mathbf{e}_x - \dot{\psi} \sin \psi \mathbf{e}_y = \boldsymbol{\omega} \times \mathbf{e}_{y'} \\ \dot{\mathbf{e}}_{z'} = 0 = \boldsymbol{\omega} \times \mathbf{e}_{z'} \end{cases} \quad (5.7)$$

考虑刚体上的点 A 。其速度为：

$$\mathbf{v}_A = \frac{d\mathbf{r}_A}{dt} = \alpha \dot{\mathbf{e}}_{x'} + \beta \dot{\mathbf{e}}_{y'} + \gamma \dot{\mathbf{e}}_{z'}. \quad (5.8)$$

将 (5.7) 式代入 (5.8) 式, 我们发现定轴转动中刚体上任意一点 A 的每时每刻的位置和速度都满足一个基于角速度矢量的简单关系:

$$\mathbf{v}_A(t) = \boldsymbol{\omega}(t) \times \mathbf{r}_A(t). \quad (5.9)$$

注意: 在上述关系中, $\boldsymbol{\omega}$ 属于刚体整体 (即每时每刻, 整个刚体只有唯一的一个 $\boldsymbol{\omega}(t)$), 刻画刚体的转动。 $\boldsymbol{\omega}$ 的取值与 A 点的选择无关。即, 对刚体上不同的点, \mathbf{v}_A 和 \mathbf{r}_A 不同, 但, $\boldsymbol{\omega}$ 都相同, 且上述关系总成立。

当然, 关系 (5.9) 有一个直接的推论, 即: 对刚体上任意两点 A, B , 在定轴转动中其相对位置, $\mathbf{r}_{AB} = \mathbf{r}_A - \mathbf{r}_B$ 和相对速度 $\mathbf{v}_{AB} = \mathbf{v}_A - \mathbf{v}_B$ 在每时每刻总满足关系:

$$\mathbf{v}_{AB}(t) = \boldsymbol{\omega}(t) \times \mathbf{r}_{AB}(t). \quad (5.10)$$

我们注意到, 在定轴转动中, 刚体的角速度矢量 $\boldsymbol{\omega}$ 必然与 $z(z')$ 轴平行或者反平行。但其 z 分量 (也是 z' 分量)

$$\omega_z = \omega_{z'} = \dot{\psi} \quad (5.11)$$

的取值则可正可负。 ω_z 的正负号可以通过如下方法判断: 在定轴转动中, 角速度矢量 $\boldsymbol{\omega}$ 的方向与转动方向满足图 (5.4) 中的右手关系。或者, 在右手拇指指向 $+z$ 方向的情况下, 若转动方向沿右手四指方向, 则 $\omega_z = \dot{\psi} > 0$; 反之, 则 $\omega_z = \dot{\psi} < 0$ 。

综上所述, 角速度矢量 $\boldsymbol{\omega}$ 刻画刚体转动的方向和快慢 (有点类似质点的速度), 属于刚体整体, 不依赖于刚体上的具体的点。

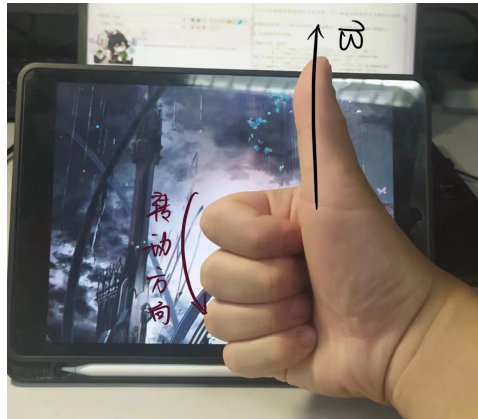


图 5.4: 右手关系

5.2.2 刚体定轴转动动力学

本小节研究定轴转动刚体的动力学。为此, 我们要推导广义坐标 ψ 满足的动力学方程, 而推导的依据是角动量定理。

角动量和转动惯量。

我们首先推导定轴刚体的总角动量的 z 分量 L_z 的表达式。为此，我们记刚体上第 i 个质点的质量为 m_i ，位置矢量为 \mathbf{r}_i 。而 \mathbf{r}_i 可以写成

$$\mathbf{r}_i = x_i \mathbf{e}_x + y_i \mathbf{e}_y + z_i \mathbf{e}_z = x'_i \mathbf{e}_{x'} + y'_i \mathbf{e}_{y'} + z'_i \mathbf{e}_{z'}. \quad (5.12)$$

由于在定轴转动情形下，刚体上质点到 z 轴距离的平方不随时间改变，我们有

$$x_i^2 + y_i^2 = x'^2_i + y'^2_i. \quad (5.13)$$

我们进一步定义刚体对 z' 轴（ z 轴）的转动惯量 I 为：

$$\begin{aligned} I &= \sum_i m_i (x'^2_i + y'^2_i) \\ &= \int dx' dy' dz' \underbrace{\rho(x', y', z')}_{\text{刚体的质量密度}} (x'^2 + y'^2). \end{aligned} \quad (5.14)$$

显然，根据这个定义，刚体绕 z' 轴（ z 轴）做定轴转动时， I 的取值不会随时间变化。

根据 L_z 的定义和前面的结果，直接的计算给出：刚体绕 z' 轴（ z 轴）做定轴转动时

$$\begin{aligned} L_z &= \sum_i m_i (\mathbf{r}_i \times \mathbf{v}_i) \cdot \mathbf{e}_z \\ &= \sum_i m_i [\mathbf{r}_i \times (\boldsymbol{\omega} \times \mathbf{r}_i)] \cdot \mathbf{e}_z \\ &= \sum_i m_i [\mathbf{r}_i \times (\omega_z \mathbf{e}_z \times \mathbf{r}_i)] \cdot \mathbf{e}_z \\ &= \omega_z \sum_i m_i (x_i^2 + y_i^2) \\ &= \omega_z \sum_i m_i (x'^2_i + y'^2_i). \end{aligned} \quad (5.15)$$

这显然意味着：

$$L_z = \omega_z I = \dot{\psi} I. \quad (5.16)$$

即： L_z 是刚体绕 z' 轴转动惯量和定轴转动角速度 z 分量的乘积。

定轴转动刚体的动力学方程。

我们将刚体视为质点组，根据质点组的总角动量定义， L_z 对时间的导数为外力对坐标原点力矩 z 分量之和，即：

$$\frac{dL_z}{dt} = M_z. \quad (5.17)$$

这里 M_z 为外力对坐标原点力矩 z 分量之和，可以表示为：

$$M_z = \left(\sum_i \mathbf{r}_i \times \mathbf{F}_i^{(e)} \right) \cdot \mathbf{e}_z, \quad (5.18)$$

其中 $\mathbf{F}_i^{(e)}$ 为第 i 个质点所受的外力（即来自刚体以外的力）。这里要注意： M_z 不是“合外力的力矩的 z 分量”。

我们将 (5.16) 式代入 (5.17) 式，就得到故可以得到刚体定轴转动动力学方程，即广义坐标 ψ 满足的方程：

$$I \frac{d\omega_z}{dt} = I\ddot{\psi} = M_z. \quad (5.19)$$

由该方程可知，转动惯量 I 刻画了刚体转动的惯性，即刚体保持转动角速度 ω_z (即 $\dot{\psi}$) 不变的能力。在外力力矩相同的情况下， I 越大，则角速度的改变越困难。

定轴转动的动能。

有时候需要用拉格朗日力学处理定轴转动的刚体问题。此时将刚体看成有很强约束的质点组，只有一个广义坐标，即 ψ 。为了写出正确的拉氏量，我们就需要把刚体的总动能 T 写成 ψ 和 $\dot{\psi}$ 的函数。根据总动能的定义，我们有

$$\begin{aligned} T &= \sum_i \frac{1}{2} m_i |v_i|^2 \\ &= \sum_i \frac{1}{2} m |\boldsymbol{\omega} \times \mathbf{r}_i|^2 \\ &= \sum_i \frac{1}{2} m_i |\omega_z \mathbf{e}_z \times (x_i \mathbf{e}_x + y_i \mathbf{e}_y + z_i \mathbf{e}_z)|^2 \\ &= \frac{1}{2} \omega_z^2 \sum_i m_i (x_i^2 + y_i^2). \end{aligned}$$

带入转动惯量的定义，我们就得到 T 的表达式：

$$T = \frac{1}{2} I \omega_z^2 = \frac{1}{2} I \dot{\psi}^2. \quad (5.20)$$

该表达式可用于构造拉氏量。

此外，利用动力学方程 (5.19) 我们还会得到：

$$\frac{dT}{dt} = I_z \omega_z \dot{\omega}_z = \omega_z M_z. \quad (5.21)$$

该方程体现了我们之前说过的刚体的特点：刚体内力不做功，因此动能的改变只与外力有关。

例题：杆摆小振动的圆频率。

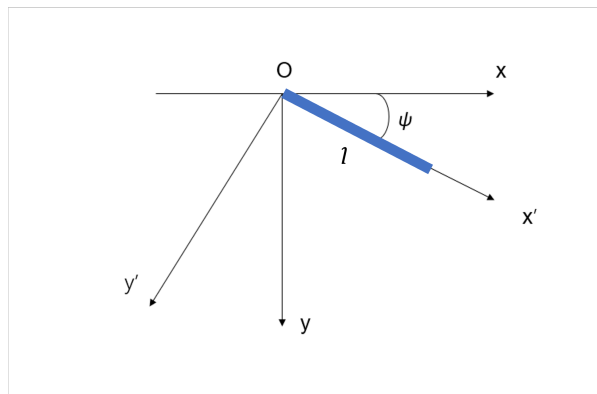


图 5.5: 杆摆。

考虑图5.11所示的杆摆：一段长为 l ，总质量为 m 且质量分布均匀的刚性杆，一端悬挂于 O 点，且能绕该点在 xy 平面内自由转动。除了来自悬挂点的力，杆只受重力，重力沿 $+y$ 方向。若该杆只在平衡位置附近小幅度摆动，求摆动的圆频率。

在本题中，刚性杆绕 z 轴定轴转动。因此可用本节方法处理。如图5.11，建立和刚体固连的坐标系 $O-x'y'z'$ ，其中 Oz' 轴和 Oz 轴重合。广义坐标为 ψ 。我们使用拉格朗日力学，杆受的主动力量是重力，因此可以写出拉氏量

$$\begin{aligned} L &= T - V \\ &= \frac{1}{2}I\dot{\psi}^2 + mg\frac{l}{2}\sin(\psi). \end{aligned} \quad (5.22)$$

其中转动惯量 $I = \frac{1}{3}ml^2$ 。

该系统平衡位置为 $\psi = \frac{\pi}{2}$ （杆沿 y 轴下垂）。由于杆在平衡位置附近小幅度摆动（做小振动），因此定义新广义坐标 q 为广义坐标对平衡位置的偏离：

$$q = \psi - \frac{\pi}{2}, \quad (5.23)$$

并把势能在 $q = 0$ 展开且只保留到 2 阶。则拉氏量近似为

$$L = \frac{1}{2}I\dot{q}^2 - \frac{1}{2}I\left[\sqrt{3g/(2l)}\right]^2 q^2 \quad (5.24)$$

该拉氏量和圆频率为 $\sqrt{\frac{3g}{2l}}$ 的谐振子是一致的（ $L_{\text{谐振子}} = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}m\omega^2 x^2$ ），故平衡位置附近的振动圆频率为 $\sqrt{\frac{3g}{2l}}$ 。当然，我们也可以直接解拉氏方程得到此结论。

5.3 刚体一般平面运动

本节我们介绍所谓的一般平面运动。设 $O-xyz$ 是惯性系，若刚体运动中每一点的 z 坐标都不随时间变化，就称刚体在 $x-y$ 平面内做平面运动。例如滚动的车轮、擦黑板时的板擦，都是刚

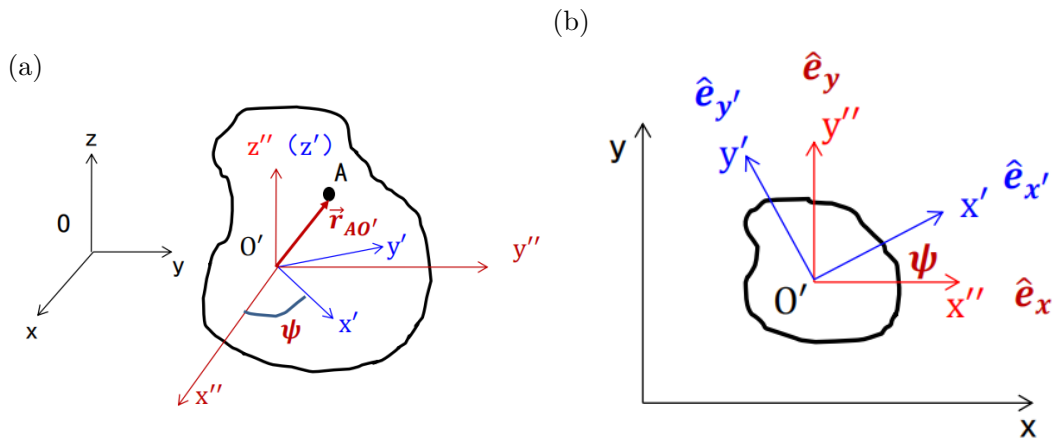


图 5.6: (a) 研究刚体时所用的惯性系 $O-xyz$, 固连系 $O'-x'y'z'$ 以及质心系 $O'-x''y''z''$, ψ 是 $O-x'$ 与 $O-x$, $O'-x''$ 轴的夹角。(b) 为在刚体运动平面 $O-xy$ 上各坐标系基矢的投影。

体 (车轮、板擦) 在做平面运动的实例。

在后面的分析中, 我们要介绍如下三种坐标系, 如图 5.6 所示

- 惯性系 $O-xyz$ 。在这一惯性系中, 我们记刚体的质心为 O' , 不失一般性, 可以选择 $O-xy$ 平面为质心所在平面, 即质心的 z 坐标可以始终选为零。
- 与刚体固连的坐标系 $O'-x'y'z'$ 。其中 $O'z'$ 轴与 Oz 轴同向, 方向不变
- 质心系 $O'-x''y''z''$, 其中坐标基矢都与 $O-xyz$ 惯性系保持平行, 不随刚体转动。注意 $O'-x'y'z'$ 与 $O'-x''y''z''$ 的不同

5.3.1 运动学

记 ψ 是 $O-x'$ 与 $O-x$, $O'-x''$ 轴的夹角, 记 $O-xyz$ 的坐标基矢为 $\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$ (同时也是质心系的坐标基矢), $O'-x'y'z'$ 的坐标基矢为 $\mathbf{e}_{x'}, \mathbf{e}_{y'}, \mathbf{e}_{z'}$, 则

$$\begin{cases} \mathbf{e}_{x'} = \cos \psi \mathbf{e}_x + \sin \psi \mathbf{e}_y \\ \mathbf{e}_{y'} = -\sin \psi \mathbf{e}_x + \cos \psi \mathbf{e}_y \\ \mathbf{e}_{z'} = \mathbf{e}_z \end{cases} \quad (5.25)$$

现在讨论广义坐标。考虑刚体上的任意一质点 A (注意 A 不一定在 $O'-x'y'$ 平面上), 我们有

$$\mathbf{r}_A = \mathbf{r}_{O'} + \mathbf{r}_{AO'} \quad (5.26)$$

其中 $\mathbf{r}_{O'} = x_{O'} \mathbf{e}_x + y_{O'} \mathbf{e}_y$ 。而 $\mathbf{r}_{AO'}$ 可以写为

$$\mathbf{r}_{AO'} \equiv \alpha \mathbf{e}_{x'} + \beta \mathbf{e}_{y'} + \gamma \mathbf{e}_{z'} \quad (5.27)$$

由于 $\mathbf{r}_{AO'}$ 和 $\mathbf{e}_{x'}, \mathbf{e}_{y'}, \mathbf{e}_{z'}$ 都与刚体固连, 因此 α, β, γ 不随时间变化。而 $\mathbf{e}_{x'}, \mathbf{e}_{y'}, \mathbf{e}_{z'}$ 都是 ψ 的函数, 因此一旦确定了 ψ , $\mathbf{r}_{AO'}$ 也就确定了。因此, 一般平面运动的广义坐标为 $\{x_{O'}, y_{O'}, \psi\}$, 这里

有三个自由度，其中两个自由度是质心运动，一个自由度是定轴转动，这也就意味着，一般平面运动相当于质心运动（由 $x_{O'}, y_{O'}$ 描述）和质心系中定轴转动（由 ψ 描述）的合运动。

我们定义角速度矢量为

$$\boldsymbol{\omega} \equiv \dot{\psi} \mathbf{e}_z = \dot{\psi} \mathbf{e}_{z'} \quad (5.28)$$

即 $\omega_z = \omega_{z'} = \dot{\psi}$ 。由于 $\mathbf{r}_{AO'} = \alpha \mathbf{e}_{x'} + \beta \mathbf{e}_{y'} + \gamma \mathbf{e}_{z'}$ ，并且 α, β, γ 不随时间改变，与定轴转动类似，直接计算可以证明

$$\dot{\mathbf{r}}_{AO'} \equiv \frac{d}{dT}(\mathbf{r} - \mathbf{r}_{O'}) = \boldsymbol{\omega} \times \mathbf{r}_{AO'} \quad (5.29)$$

即

$$\mathbf{v}_A = \dot{\mathbf{r}}_A = \dot{\mathbf{r}}_{O'} + \dot{\mathbf{r}}_{AO'} = \mathbf{v}_{O'} + \boldsymbol{\omega} \times \mathbf{r}_{AO'} \quad (5.30)$$

于是对于刚体上任意两个质点 A, B ，可以定义 $\mathbf{r}_{AB} \equiv \mathbf{r}_A - \mathbf{r}_B$ ，以及 $\mathbf{v}_{AB} = \dot{\mathbf{r}}_{AB} = \mathbf{v}_A - \mathbf{v}_B$ ，就有

$$\mathbf{v}_{AB} = \mathbf{v}_{AO'} - \mathbf{v}_{BO'} = \boldsymbol{\omega} \times (\mathbf{r}_{AO'} - \mathbf{r}_{BO'}) = \boldsymbol{\omega} \times \mathbf{r}_{AB} \quad (5.31)$$

这里 $\boldsymbol{\omega}$ 与 AB 的选择无关，它属于刚体整体。

5.3.2 动力学

1. 质心系中总角动量 z 分量

在质心系中，总角动量的 z 分量被定义为

$$L_{O'z} = \sum_i m_i [(\mathbf{r}_i - \mathbf{r}_{O'}) \times (\mathbf{v}_i - \mathbf{v}_{O'})] \cdot \mathbf{e}_z \quad (5.32)$$

与定轴转动类似，它可以被记为

$$L_{O'z} = \omega_z I_{O'} = I_{O'} \dot{\psi} \quad (5.33)$$

这里我们有

$$I_{O'} = \sum_i m_i [(x_i - x_{O'})^2 + (y_i - y_{O'})^2] = \sum_i m_i (x_i'^2 + y_i'^2) \quad (5.34)$$

其中 x'_i, y'_i 是第 i 个质点在 $O' - x'y'z'$ 系坐标，不随时间变化。 $I_{O'}$ 即为刚体对“过质心，与刚体固连，沿 $z(z')$ 方向”的轴的转动惯量，不随时间变化。

2. 质心动力学方程

质心的动力学方程被表达为

$$M \ddot{x}_{O'} = F_x \quad (5.35)$$

$$M \ddot{y}_{O'} = F_y \quad (5.36)$$

这里 F_x, F_y 是刚体所收到合外力的 x, y 分量，即 $F_{x,y} = \sum_i \mathbf{F}_i^{(e)} \cdot \mathbf{e}_{x,y}$ ，其中 $\mathbf{F}_i^{(e)}$ 为第 i 个质点所受合外力。

3. 质心系中角动量定理 (z 方向)

在质心系下, 可以给出 z 方向上的角动量定理为

$$\frac{dL_{O'z}}{dt} = M_{O'z} \quad (5.37)$$

即有

$$I_{O'z}\ddot{\psi} = M_{O'z} \quad (5.38)$$

这里 $M_{O'z}$ 为质心系中外力力矩的 z 分量之和, 即 $M_{O'z} = \left[\sum_i (\mathbf{r}_i - \mathbf{r}_{O'} \times \mathbf{F}_i^{(e)}) \cdot \mathbf{e}_z \right]$ 。于是(5.35)(5.36)(5.38)为一般平面运动刚体的动力学方程, 这三个方程分别给出 x', y', ψ 的描述。立刻可以看出如果 $M_{O'z} = 0$, 那么 $\dot{\psi}$ 将不随时间改变 (例如刚体只受重力的情形)。

4. 动能 (用于 Lagrange 力学)

根据柯尼希定理, 我们有

$$T = T_{O'} + T' \quad (5.39)$$

其中, $T_{O'}$ 为质心动能

$$T_{O'} = \frac{M}{2}(\dot{x}_{O'}^2 + \dot{y}_{O'}^2) \quad (5.40)$$

而 T' 是质心系中的动能, 根据与定轴转动中类似的推导可以给出

$$T' = \frac{1}{2}I_{O'}\omega_z^2 = \frac{1}{2}I_{O'}\dot{\psi}^2 \quad (5.41)$$

5. 无滑滚动

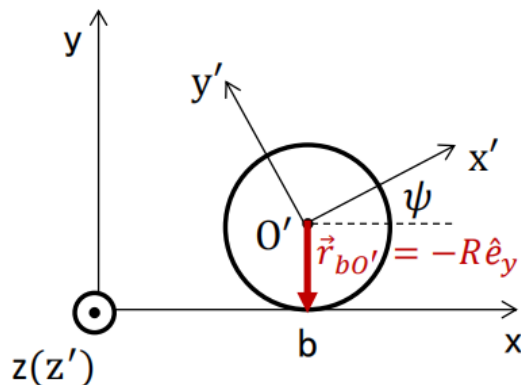


图 5.7: 无滑滚动模型示意图

我们来考察一类特殊的平面运动—无滑滚动。考虑半径为 R 的圆柱体在有摩擦的平面上运动, 如图 5.7 所示, 其中 b 接触点, 而 O' 是圆心。所谓的无滑滚动, 是指在惯性系中 $\mathbf{v}_b = 0$ (由于静

摩擦导致的)。考虑到

$$\mathbf{v}_b = \mathbf{v}_{O'} + \mathbf{v}_{bO'} = \mathbf{v}_{O'} + \boldsymbol{\omega} \times \mathbf{r}_{bO'} = \mathbf{v}_{O'} + \boldsymbol{\omega} \times (-R)\mathbf{unit}_y = \mathbf{v}_{O'} + \dot{\psi}R\mathbf{e}_x \quad (5.42)$$

设 $\mathbf{v}_{O'} = v_{O'}\mathbf{e}_x$ ，无滑滚动就意味着 $v_{O'} + \dot{\psi}R = 0$ ，即

$$v_{O'} = -\dot{\psi}R = -\omega_z R \quad (5.43)$$

其中 $\dot{\psi} > 0$ 对应于向左滚动。因此在无滑滚动时，质心速度分量和转动角速度分量关系被锁定。

6. 一个无滑滚动例子

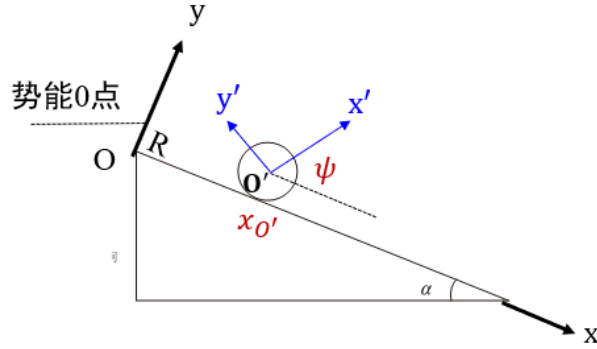


图 5.8: 斜面上进行无滑滚动的圆柱体

如图 5.8 所示，一个质量均匀的圆柱体质量为 M ，半径为 R ，斜面倾角为 α ，在斜面上进行无滑滚动，现试问这一圆柱体的质心加速度。

我们记 O' 为圆心，描述这一圆柱体的坐标可以如图选为 $x_{O'}, \psi$ 。记 $Oxyz$ 为惯性系，而 $O' - x'y'z'$ 为与刚体固连的坐标系。由于 $\dot{x}_{O'} = -\dot{\psi}R$ ，因此有

$$x_{O'} = -\psi R + \psi_0 R \quad (5.44)$$

这里 ψ_0 是在 $x_{O'} = 0$ 时的 ψ 。因此无滑滚动是完整约束 (它让 $x_{O'}$ 与 ψ 相联系，导致只有一个独立的广义坐标)，相应的静摩擦力是约束力。于是选取广义坐标为 O' ，拉格朗日量记为 $L = T - V$ ，其中动能项为

$$L = \frac{1}{2}M\dot{x}_{O'}^2 + \frac{1}{2}I_{O'z}\dot{\psi}^2 = \frac{1}{2}M\dot{x}_{O'}^2 + \frac{I_{O'z}}{2R^2}\dot{x}_{O'}^2 \quad (5.45)$$

这里转动惯量 $I_{O'z} = \frac{M}{2}R^2$ 。另一方面，通过选择合适的势能零点，可以将重力势能表达为 (请读者思考为什么?)

$$V = -mgx_{O'} \sin \alpha \quad (5.46)$$

于是从拉格朗日方程出发，自然可以导出

$$\ddot{x}_{O'} = \frac{2g \sin \alpha}{3} \quad (5.47)$$

5.4 刚体的定点运动

5.4.1 运动学

两条射线夹角的准确定义。

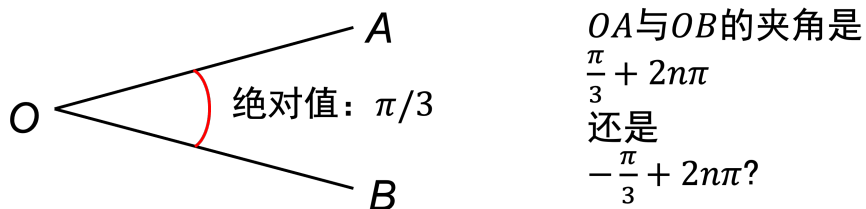


图 5.9: “两条射线的夹角是多少?”的可能答案。

如图 (5.9) 所示, 严格来说, 对两条相交的射线, 问题“它们的夹角是多少?”有至少两个答案, 彼此绝对值相同, 正负号相反。而且还可以在每个答案的基础上再加上 2π 的整数倍。因此, 对给定的两条相交的射线, 我们要约定如下三个要素, 才能够唯一地定义其夹角:

- “右手拇指方向” (与两条射线都垂直的一个方向)。
- 从谁转到谁。
- 角的取值范围 (很多时候可省略)。

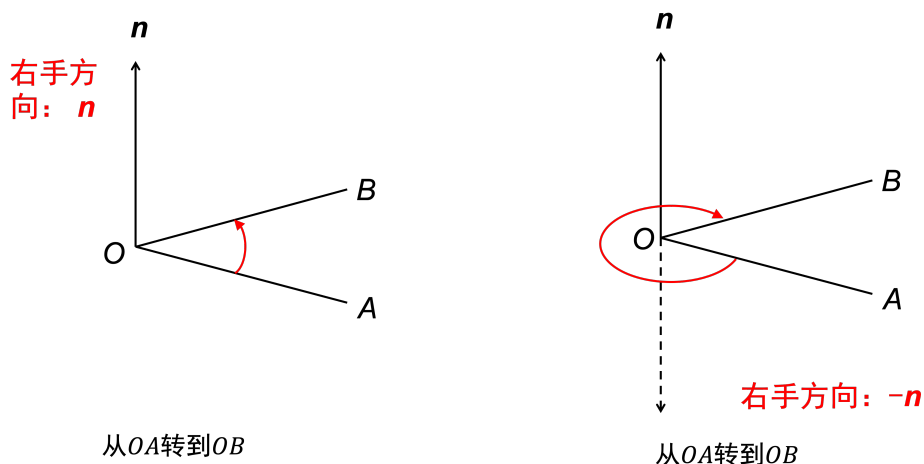


图 5.10: 明确定义夹角的例子

例如, 对图 5.10 中的射线 OA 和 OB , 如果约定 n 为右手拇指方向, 从 OA 转到 OB (见图 5.10 左), 且 $\theta \in [0, 2\pi)$, 则此约定下的 OA 和 OB 的夹角取值为: OA 沿“ n 的右手四指方向”转到 OB , 所转之角度, 再通过加上 2π 的整数倍, 调至 $[0, 2\pi)$ 区间。

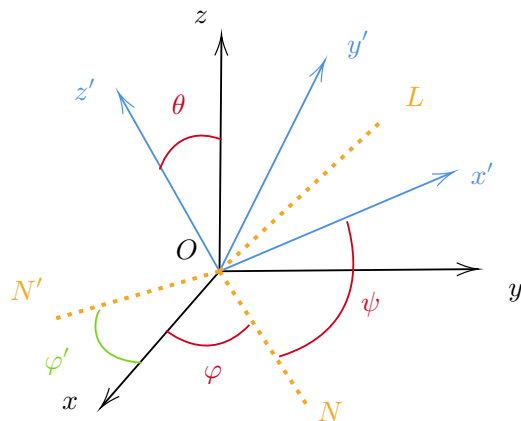


图 5.12: 定点运动刚体的坐标系和欧拉角。

再比如，仍考虑图5.10的情况，若将右手拇指方向的约定修改为 $-\mathbf{n}$ ，仍从 OA 转到 OB ，(见图5.10右) $\theta \in [0, 2\pi)$ ，则 OA 和 OB 的夹角取值会变为： $2\pi -$ “约定修改前的取值”。

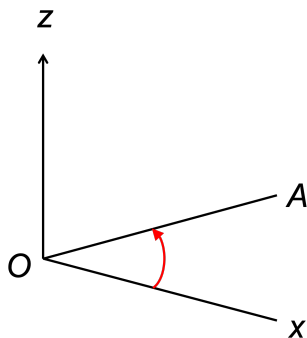


图 5.11: 过去常说的“ OA 与 x 轴的夹角”。

如图5.10所示，我们过去常说的“ OA 与 x 轴的夹角”，其实是默认了“以 z 轴为拇指方向，从 Ox 转到 OA ”的约定。

广义坐标：欧拉角。

我们考虑一种特殊的运动，即**定点运动**。在定点运动中，刚体上的某一点在惯性系中始终不动，我们设其为惯性系坐标原点。

如图5.12，设 O 为定点运动刚体的固定点， $O - xyz$ 为惯性系。与之前的情况类似，我们同时建立与刚体固连，且原点也在 O 的坐标系 $O - x'y'z'$ 。定点运动刚体的广义坐标是三个欧拉角。为了介绍它们的定义，我们首先引入如下的三条射线：

- 射线 ON ：从 O 出发，沿 $\mathbf{e}_z \times \mathbf{e}_{z'}$ 方向 ($\mathbf{e}_{z(z')}$ 为 $z(z')$ 方向的单位矢量)。

由于 $ON \perp Oz$ 且 $ON \perp Oz'$ ，所以 ON 既位于 $O - xy$ 平面，也位于 $O - x'y'$ 平面。

- 射线 ON' : Oz' 轴正半轴在 $O - xy$ 平面的投影。

由于 ON' 位于 Ozz' 面, 所以 $ON \perp ON'$ 。

- 射线 OL : Oz 轴正半轴在 $O - x'y'$ 面的投影。

与 ON' 类似, OL 位于 Ozz' 面, 且 $\therefore OL \perp ON$ 。

下面介绍定点运动刚体的广义坐标, 即欧拉角。显然, 对定点运动的刚体, 只要确定了和刚体固连的三个坐标轴 (Ox' , Oy' , Oz' 轴) 的方向, 就能确定刚体中每个质点的位置。因此, 寻找广义坐标的问题, 归结为如何确定三个固连坐标轴的方向问题。

我们分两步解决此问题, 即: 先确定 Oz' 轴的方向, 再在 Oz' 轴的方向确定的情况下, 确定 Ox' 轴和 Oy' 轴的方向。以下角度定义均参考图5.12。

1、确定 oz' 轴的方向:

oz' 轴的方向可由两个角度确定:

- θ : oz 与 oz' 轴的夹角 (以 ON 为右手拇指方向, 从 oz 转到 oz' , $\theta \in [0, \pi]$)
- φ' : OX 与 ON' 的夹角

出于习惯的原因, 我们将角度 φ' 替换为如下角度 φ :

- φ : OX 与 ON 的夹角 (以 Oz 为右手拇指方向, 从 OX 转到 ON , 通常 $\varphi \in [0, 2\pi]$)

显然, 由于 ON 与 ON' 垂直, 角度 φ 和 φ' 承载的信息是等价的。

2、确定 ox' 轴和 oy' 轴的方向:

在 oz' 轴的方向确定后, ox' 轴和 oy' 轴的方向可由如下角度确定:

- ψ : ON 与 OX 轴的夹角 (以 Oz' 为右手拇指方向, 从 ON 转到 OX' , 通常 $\psi \in [0, 2\pi]$)

(θ, φ, ψ) 三个角度即为**欧拉角**。欧拉角是定点运动刚体的广义坐标。如前所述, 三个欧拉角中 θ, φ 用来确定 Oz' 轴的方向, ψ 决定 Ox' 和 Oy' 轴绕 Oz' 轴的转动。此外, 由于欧拉角是定点运动刚体的广义坐标, $O - x'y'z'$ 系的方向矢量 $e_{x',y',z'}$ 就必然是三个欧拉角的函数。它们可以被写为: $e_{x'}(\theta, \varphi, \psi)$, $e_{y'}(\theta, \varphi, \psi)$ 和 $e_{z'}(\theta, \varphi, \psi)$ 。

角速度矢量。

以下我们将证明: 若刚体做定点运动, 则对 $\forall A \in$ 刚体, 其速度 v_A 和位置 r_A 之间必然满足关系:

$$v_A = \omega \times r_A, \quad (5.48)$$

其中 ω 与 A 的选择无关。

证明. 刚体上点 A 的位置可以写成:

$$r_A = \sum_{j'=x',y',z'} r_{j'} e_{j'}(\theta, \varphi, \psi). \quad (5.49)$$

和定轴转动以及一般平面运动的情况类似, 由于 \mathbf{r}_A 和 $\mathbf{e}_{x',y',z'}$ 均和刚体固连, 故三个分量 $r_{x',y',z'}$ 在运动中都不随时间变化。因此 A 的速度可以表示成:

$$\begin{aligned}\mathbf{v}_A &= \dot{\mathbf{r}}_A \\ &= \sum_{j'=x',y',z'} r_{j'} \dot{\mathbf{e}}_{j'} [\theta(t), \varphi(t), \psi(t)] \\ &= \sum_{j'=x',y',z'} r_{j'} \left[\frac{\partial \mathbf{e}_{j'}}{\partial \theta} \dot{\theta} + \frac{\partial \mathbf{e}_{j'}}{\partial \varphi} \dot{\varphi} + \frac{\partial \mathbf{e}_{j'}}{\partial \psi} \dot{\psi} \right].\end{aligned}\quad (5.50)$$

因此, 对 \mathbf{v}_A 的计算, 可以归结为对如下九个导数的计算:

$$\boxed{\frac{\partial \mathbf{e}_{j'}}{\partial \theta}, \frac{\partial \mathbf{e}_{j'}}{\partial \varphi}, \frac{\partial \mathbf{e}_{j'}}{\partial \psi}, \quad (j' = x', y', z').} \quad (5.51)$$

我们首先计算 $\frac{\partial \mathbf{e}_{x'}}{\partial \theta}$ 。为此, 考虑一个特殊的定点运动, 即绕 ON 的定轴转动。此运动中 φ 和 ψ 都不随时间变化, 只有 θ 随时间变化, 即有 $\dot{\varphi} = \dot{\psi} = 0$ 。于是 $\mathbf{v}_A = \sum_{j'=x',y',z'} r_{j'} \frac{\partial \mathbf{e}_{j'}}{\partial \theta} \dot{\theta}$ 。我们进一步考虑一个特殊的点 A , 它在 $O-X'$ 轴上, 距 O 为 1, 因此满足 $\mathbf{r}_A = \mathbf{e}_{x'}$, 即 $r_{x'} = 1, r_{y'} = r_{z'} = 0$ 。于是在绕 ON 的定轴转动中有

$$\mathbf{v}_A = \frac{\partial \mathbf{e}_{x'}}{\partial \theta} \dot{\theta}. \quad (5.52)$$

另一方面, 由于刚体绕 ON 定轴转动, 故利用定轴转动理论中的结论, 我们还有

$$\mathbf{v}_A = \dot{\theta} \mathbf{e}_N \times \mathbf{r}_A = \dot{\theta} \mathbf{e}_N \times \mathbf{e}_{x'}. \quad (5.53)$$

对比 (5.52) 式和 (5.53) 式, 可以得到:

$$\frac{\partial \mathbf{e}_{x'}}{\partial \theta} = \mathbf{e}_N \times \mathbf{e}_{x'} \quad (5.54)$$

类似, 考虑刚体仍绕 ON 的定轴转动, 但 $\mathbf{r}_A = \mathbf{e}_{y'}$ 和 $\mathbf{r}_A = \mathbf{e}_{z'}$ 的情况, 我们会得到: $\frac{\partial \mathbf{e}_{y'}}{\partial \theta} = \mathbf{e}_N \times \mathbf{e}_{y'}$ 和 $\frac{\partial \mathbf{e}_{z'}}{\partial \theta} = \mathbf{e}_N \times \mathbf{e}_{z'}$, 这就意味着:

$$\frac{\partial \mathbf{e}_{j'}}{\partial \theta} = \mathbf{e}_N \times \mathbf{e}_{j'} \quad (j' = x', y', z'). \quad (5.55)$$

这就计算出了九个导数中的三个。类似的, 如果我们考虑刚体绕 Oz 轴的转动 ($\dot{\theta} = \dot{\psi} = 0$), 和刚体绕 Oz' 轴的转动 ($\dot{\theta} = \dot{\varphi} = 0$), 则会得到另外六个导数:

$$\frac{\partial \mathbf{e}_{j'}}{\partial \varphi} = \mathbf{e}_z \times \mathbf{e}_{j'} \quad (j' = x', y', z'), \quad (5.56)$$

$$\frac{\partial \mathbf{e}_{j'}}{\partial \psi} = \mathbf{e}_{z'} \times \mathbf{e}_{j'} \quad (j' = x', y', z'). \quad (5.57)$$

把 (5.55-5.57) 代入 (5.50), 我们最终得到

$$\mathbf{v}_A = \sum_{j'} r_{j'} \left[\dot{\theta} \mathbf{e}_N \times \mathbf{e}_{j'} + \dot{\varphi} \mathbf{e}_z \times \mathbf{e}_{j'} + \dot{\psi} \mathbf{e}_{z'} \times \mathbf{e}_{j'} \right]. \quad (5.58)$$

定义定点运动刚体的角速度矢量 $\boldsymbol{\omega}$ 为：

$$\boldsymbol{\omega} = \dot{\theta}\mathbf{e}_N + \dot{\phi}\mathbf{e}_z + \dot{\psi}\hat{\mathbf{e}}_{z'}. \quad (5.59)$$

显然， $\boldsymbol{\omega}$ 属于刚体整体，与 A 点选择无关，且有

$$\mathbf{v}_A = \boldsymbol{\omega} \times \mathbf{r}_A. \quad (5.60)$$

故命题得证。 □

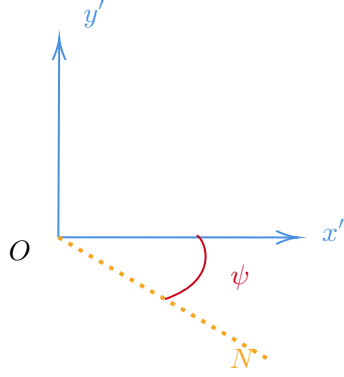
公式 (5.59) 和 (5.60) 还有如下物理意义：定点运动刚体任意“无穷小运动”均可分解成 3 个无穷小的定轴转动，即绕 \mathbf{e}_N 、 \mathbf{e}_z 、 $\mathbf{e}_{z'}$ 轴的定轴转动。每个无穷小的定轴转动分别改变一个欧拉角。

Euler 运动学方程

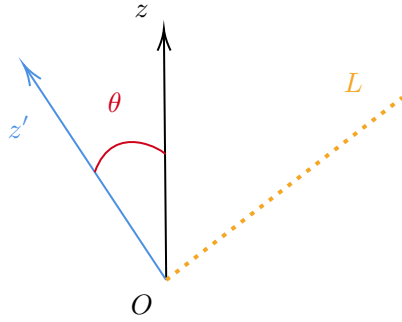
在应用中，常把 $\boldsymbol{\omega}$ 在 $O - x'y'z'$ 系分解，写成

$$\boldsymbol{\omega} = \omega_{x'}\mathbf{e}_{x'} + \omega_{y'}\mathbf{e}_{y'} + \omega_{z'}\mathbf{e}_{z'}. \quad (5.61)$$

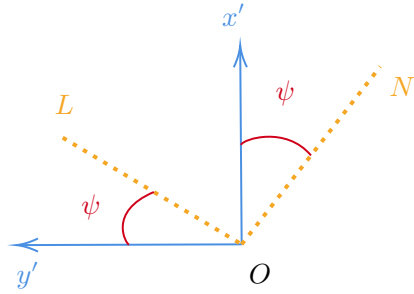
为了利用公式 (5.59) 计算三个分量 $\omega_{x',y',z'}$ ，我们需要将 \mathbf{e}_N 、 \mathbf{e}_z 写成 $\mathbf{e}_{x',y',z'}$ 的线性组合。我们注意到如下关系：



$$\mathbf{e}_N = \cos \psi \mathbf{e}_{x'} - \sin \psi \mathbf{e}_{y'}$$



$$\mathbf{e}_z = \cos \theta \mathbf{e}_{z'} + \sin \theta \mathbf{e}_L$$



$$\mathbf{e}_L = \cos \psi \mathbf{e}_{y'} + \sin \psi \mathbf{e}_{x'}$$

故最终可以得到 $\mathbf{e}_N = \cos \psi \mathbf{e}_{x'} - \sin \psi \mathbf{e}_{y'}$ 以及

$$\mathbf{e}_z = \cos \theta \mathbf{e}_{z'} + \sin \theta \cos \psi \mathbf{e}_{y'} + \sin \theta \sin \psi \mathbf{e}_{x'}. \quad (5.62)$$

将上述结果代入

$$\boldsymbol{\omega} = \dot{\theta} \mathbf{e}_N + \dot{\varphi} \mathbf{e}_z + \dot{\psi} \mathbf{e}_{z'} = \omega_{x'} \mathbf{e}_{x'} + \omega_{y'} \mathbf{e}_{y'} + \omega_{z'} \mathbf{e}_{z'}, \quad (5.63)$$

我们就得到欧拉运动学方程：

$$\begin{cases} \omega_{x'} = \dot{\varphi} \sin \theta \sin \psi + \dot{\theta} \cos \psi \\ \omega_{y'} = \dot{\varphi} \sin \theta \cos \psi - \dot{\theta} \sin \psi \\ \omega_{z'} = \dot{\varphi} \cos \theta + \dot{\psi} \end{cases} . \quad (5.64)$$

5.4.2 动力学：角动量和惯量张量

惯性系中的总角动量。

定点运动的刚体在惯性系中（对坐标原点，即固定点 O ）的总角动量 \mathbf{L} 为：

$$\begin{aligned}\mathbf{L} &= \sum_i m_i \mathbf{r}_i \times \mathbf{v}_i \\ &= \sum_i m_i \mathbf{r}_i \times (\boldsymbol{\omega} \times \mathbf{r}_i) \\ &= \sum_i m_i \left[\boldsymbol{\omega} |\mathbf{r}_i|^2 - \mathbf{r}_i (\boldsymbol{\omega} \cdot \mathbf{r}_i) \right].\end{aligned}\quad (5.65)$$

与定点运动的情况类似，此处 m_i , \mathbf{r}_i 和 \mathbf{v}_i 分别表示刚体上第 i 个质点的质量、位置和速度， \sum_i 表示对刚体上所有质点求和。上述推导过程的第三步可以直接运用叉乘的定义进行证明。

惯量张量。

定义 V 为所有 3 维实矢量的集合，因此对给定的一个定点运动的刚体（固定点为坐标原点），我们通过如下方式定义一个与之对应的 $V \rightarrow V$ 的线性映射 \hat{K} ：对于 V 中的任意矢量 \mathbf{a} ， \hat{K} 作用于 \mathbf{a} 上的结果 $\hat{K}[\mathbf{a}]$ 也是 V 中一个三维矢量，它满足：

$$\hat{K}[\mathbf{a}] \equiv \sum_i m_i [\mathbf{a} |\mathbf{r}_i|^2 - \mathbf{r}_i (\mathbf{a} \cdot \mathbf{r}_i)], \quad \forall \mathbf{a} \in V. \quad (5.66)$$

该线性映射 \hat{K} 称为该定点运动的刚体的**惯量张量**。显然，根据 (5.65)，对一个定点运动的刚体而言，其惯量张量 \hat{K} 作用在其角速度 $\boldsymbol{\omega}$ 上之后所得的矢量，恰恰就是其总角动量 \mathbf{L} 。即：

$$\mathbf{L} = \hat{K}[\boldsymbol{\omega}]. \quad (5.67)$$

因此，一个定点运动刚体的惯量张量 \hat{K} ，是连接其总角动量矢量和角速度矢量的桥梁，像质点的质量。如前所述，定点运动是定轴转动在 3 维情况下的推广。所以，惯量张量其实是定轴转动中转动惯量概念在 3 维情况下的推广。

惯量张量的矩阵表示。

由于 \hat{K} 为 $V \rightarrow V$ 的线性映射，在一组给定的基矢下， \hat{K} 可表示为 3×3 矩阵。例如，对于和刚体固连的 $O - x'y'z'$ 系的坐标基矢 $\mathbf{e}_{x',y',z'}$ ， $\boldsymbol{\omega}$ 和 \mathbf{L} 可分别写成：

$$\boldsymbol{\omega} = \sum_{l'=x',y',z'} \omega_{l'} \mathbf{e}_{l'}; \quad \mathbf{L} = \sum_{l'=x',y',z'} L_{l'} \mathbf{e}_{l'}. \quad (5.68)$$

于是我们有

$$\begin{aligned}L_{l'} &= \mathbf{e}_{l'} \cdot \mathbf{L} \\ &= \mathbf{e}_{l'} \cdot \hat{K} \left[\sum_{m'=x',y',z'} \omega_{m'} \mathbf{e}_{m'} \right] \\ &= \sum_{m'=x',y',z'} \left(\mathbf{e}_{l'} \cdot \hat{K}[\mathbf{e}_{m'}] \right) \omega_{m'}, \quad (l' = x', y', z').\end{aligned}\quad (5.69)$$

我们定义 \hat{K} 的分量

$$K_{l'm'} = \mathbf{e}_{l'} \cdot \hat{K} [\mathbf{e}_{m'}], \quad (l', m' = x', y', z') \quad (5.70)$$

于是 (5.69) 式可以被改写为:

$$L_{l'} = \sum_{m'=x',y',z'} K_{l'm'} \omega_{m'} \quad (5.71)$$

即

$$\begin{pmatrix} L_{x'} \\ L_{y'} \\ L_{z'} \end{pmatrix} = \begin{pmatrix} K_{x'x'} & K_{x'y'} & K_{x'z'} \\ K_{y'x'} & K_{y'y'} & K_{y'z'} \\ K_{z'x'} & K_{z'y'} & K_{z'z'} \end{pmatrix} \begin{pmatrix} \omega_{x'} \\ \omega_{y'} \\ \omega_{z'} \end{pmatrix}. \quad (5.72)$$

我们记 K 为 (5.72) 式中以 $K_{l'm'}(l', m' = x', y', z')$ 为矩阵元的 3×3 矩阵。 K 是惯量张量 \hat{K} 在 $O - x'y'z'$ 系的矩阵表示, 也可简单称之为 “ $O - x'y'z'$ 系转动惯量矩阵”。

惯量张量的矩阵元。

我们来计算惯量张量在和刚体固连的 $O - x'y'z'$ 系的坐标基矢 $\mathbf{e}_{x',y',z'}$ 中的矩阵元 $K_{l'm'}(l', m' = x', y', z')$ 。设刚体上第 i 个质点的位置矢量为

$$\mathbf{r}_i = x_i' \mathbf{e}_{x'} + y_i' \mathbf{e}_{y'} + z_i' \mathbf{e}_{z'} \quad (5.73)$$

如前所述, 由于 $O - x'y'z'$ 系和刚体固连, x_i', y_i', z_i' 不随时间变化。直接利用定义我们可以计算出惯量张量矩阵 K 的一个对角元 $K_{x'x'}$ 和一个非对角元 $K_{x'y'}$ 的表达式:

$$\begin{aligned} K_{x'x'} &= \mathbf{e}_{x'} \cdot \hat{K} [\mathbf{e}_{x'}] \\ &= \mathbf{e}_{x'} \cdot \sum_i \left[\mathbf{e}_{x'} |\mathbf{r}_i|^2 - \mathbf{r}_i (\mathbf{e}_{x'} \cdot \mathbf{r}_i) \right] \\ &= \sum_i \left[|\mathbf{r}'_i|^2 - x_i'^2 \right] \\ &= \sum_i m_i (y_i'^2 + z_i'^2), \end{aligned} \quad (5.74)$$

和

$$\begin{aligned} K_{x'y'} &= \mathbf{e}_{x'} \cdot \hat{K} [\mathbf{e}_{y'}] \\ &= \mathbf{e}_{x'} \cdot \sum_i \left[\mathbf{e}_{y'} |\mathbf{r}_i|^2 - \mathbf{r}_i (\mathbf{e}_{y'} \cdot \mathbf{r}_i) \right] \\ &= - \sum_i m_i x_i' y_i'. \end{aligned} \quad (5.75)$$

其他的矩阵元也可以用类似方法计算。最后我们得到的结果为:

$$K_{x'x'} = \sum_i m_i (y_i'^2 + z_i'^2), \quad (5.76)$$

$$K_{y'y'} = \sum_i m_i (z_i'^2 + x_i'^2), \quad (5.77)$$

$$K_{z'z'} = \sum_i m_i (x_i'^2 + y_i'^2). \quad (5.78)$$

和

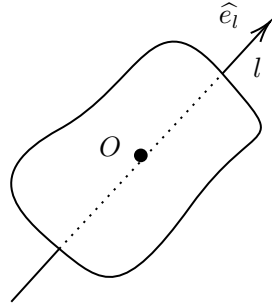
$$K_{y'z'} = K_{z'y'} = - \sum_i m_i z'_i y'_i, \quad (5.79)$$

$$K_{x'z'} = K_{z'x'} = - \sum_i m_i z'_i x'_i, \quad (5.80)$$

$$K_{y'x'} = K_{x'y'} = - \sum_i m_i x'_i y'_i. \quad (5.81)$$

显然，在与刚体固连的坐标系中， \hat{K} 的 9 个分量都不随时间改变。而且， \hat{K} 的三个“对角分量” $K_{x'x'}$ 、 $K_{y'y'}$ 和 $K_{z'z'}$ 就是刚体对三个固连坐标轴 (x' 轴、 y' 轴、 z' 轴) 的转动惯量。

因此，我们还有如下推论 (如下图所示)：定点运动的刚体对任意过固定点且与之固连的轴 l 的转动惯量可表示为 $\mathbf{e}_l \cdot \hat{K} [\mathbf{e}_l]$ ，其中 \mathbf{e}_l 为轴方向的单位向量。



惯量张量矩阵的坐标变换。

设有两个与刚体固连的坐标系，分别为 $O - x'y'z'$ 系 (基矢 $\mathbf{e}_{x'}, \mathbf{e}_{y'}, \mathbf{e}_{z'}$) 和 $O - \alpha\beta\gamma$ 系 (基矢 $\mathbf{e}_\alpha, \mathbf{e}_\beta, \mathbf{e}_\gamma$)。 \hat{K} 在 $O - x'y'z'$ 系分量为

$$K_{l'm'} = \mathbf{e}_{l'} \cdot \hat{K} [\mathbf{e}_{m'}], \quad (l', m' = x', y', z'). \quad (5.82)$$

\hat{K} 在 $O - \alpha\beta\gamma$ 系分量为：

$$K_{\eta\xi} = \mathbf{e}_\eta \cdot \hat{K} [\mathbf{e}_\xi], \quad (\eta, \xi = \alpha, \beta, \gamma). \quad (5.83)$$

我们记 $O - x'y'z'$ 系和 $O - \alpha\beta\gamma$ 系的惯量张量矩阵为 $K^{(x'y'z')}$ 和 $K^{(\alpha\beta\gamma)}$ 。

我们定义：

$$Q_{\eta l'} = \mathbf{e}_\eta \cdot \mathbf{e}_{l'}, \quad (\eta = \alpha, \beta, \gamma; l' = x', y', z'). \quad (5.84)$$

因此， $O - x'y'z'$ 系和 $O - \alpha\beta\gamma$ 系的基矢之间变换关系可以表示为：

$$\mathbf{e}_\eta = \sum_{l'=x',y',z'} Q_{\eta l'} \mathbf{e}_{l'}, \quad (\eta = \alpha, \beta, \gamma). \quad (5.85)$$

于是我们有：

$$\begin{aligned} K_{\eta\xi} &= \mathbf{e}_\eta \cdot \hat{K} [\mathbf{e}_\xi] \\ &= \left(\sum_{l'=x',y',z'} Q_{\eta l'} \mathbf{e}_{l'} \right) \cdot \hat{K} \left[\sum_{m'=x',y',z'} Q_{\xi m'} \mathbf{e}_{m'} \right] \\ &= \sum_{l',m'=x',y',z'} Q_{\eta l'} Q_{\xi m'} \mathbf{e}_{l'} \cdot \hat{K} [\mathbf{e}_{m'}], \quad (\eta, \xi = \alpha, \beta, \gamma). \end{aligned} \quad (5.86)$$

所以，惯量张量 \hat{K} 在 $O - x'y'z'$ 系和 $O - \alpha\beta\gamma$ 系的矩阵元满足变换关系：

$$K_{\eta\xi} = \sum_{l'm'} Q_{\eta l'} K_{l'm'} Q_{\xi m'}, \quad (\eta, \xi = \alpha, \beta, \gamma). \quad (5.87)$$

定义 Q 为以 $Q_{\eta l'} (\eta = \alpha, \beta, \gamma, l' = x', y', z')$ 为矩阵元的矩阵：

$$Q = \begin{pmatrix} \mathbf{e}_\alpha \cdot \mathbf{e}_{x'} & \mathbf{e}_\alpha \cdot \mathbf{e}_{y'} & \mathbf{e}_\alpha \cdot \mathbf{e}_{z'} \\ \mathbf{e}_\beta \cdot \mathbf{e}_{x'} & \mathbf{e}_\beta \cdot \mathbf{e}_{y'} & \mathbf{e}_\beta \cdot \mathbf{e}_{z'} \\ \mathbf{e}_\gamma \cdot \mathbf{e}_{x'} & \mathbf{e}_\gamma \cdot \mathbf{e}_{y'} & \mathbf{e}_\gamma \cdot \mathbf{e}_{z'} \end{pmatrix}. \quad (5.88)$$

于是变换关系 (5.87) 也可以写成：

$$K^{(\alpha\beta\gamma)} = Q K^{(x'y'z')} Q^T. \quad (5.89)$$

惯量主轴

我们注意到：(5.88) 式所定义的 $O - x'y'z'$ 系和 $O - \alpha\beta\gamma$ 系的基矢之间变换矩阵 Q 满足：

$$\begin{aligned} (QQ^T)_{\eta\xi} &= \sum_{l'=x',y',z'} Q_{\eta l'} Q_{\xi l'} \quad (\eta, \xi = \alpha, \beta, \gamma) \\ &= \sum_{l'} (\mathbf{e}_\eta \cdot \mathbf{e}_{l'}) (\mathbf{e}_{l'} \cdot \mathbf{e}_\xi) \\ &= \mathbf{e}_\eta \cdot \mathbf{e}_\xi \\ &= \delta_{\eta\xi} \end{aligned} \quad (5.90)$$

这里第三个等号处用到了：

$$\begin{aligned} \mathbf{A} \cdot \mathbf{B} &= A_{x'} B_{x'} + A_{y'} B_{y'} + A_{z'} B_{z'} \\ &= \sum_{l'=x',y',z'} (\mathbf{A} \cdot \mathbf{e}_{l'}) (\mathbf{e}_{l'} \cdot \mathbf{B}) \end{aligned} \quad (5.91)$$

因此我们有

$$QQ^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (5.92)$$

即

$$Q^T = Q^{-1}. \quad (5.93)$$

也就是说： Q 是实正交矩阵。换句话说：实正交矩阵与和刚体固连的直角坐标变换一一对应。

设 K 为惯量张量在与刚体固连的 $O - x'y'z'$ 系的矩阵形式，由于 K 是实对称矩阵，因此一定存在实正交矩阵 Q 使得

$$QKQ^T = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}. \quad (5.94)$$

其中, $\lambda_{1,2,3}$ 为 K 的本征值。因此, 对于一个给定刚体与一个给定的固定点 O , 存在至少一组特殊的, 与刚体固连的, 以 O 为原点的坐标系 $O - x''y''z''$, 使 \hat{K} 在其中的矩阵 K'' 是对角的矩阵。此时, 称 $O - x''y''z''$ 为**主轴系**, 称 $O - x'', O - y'', O - z''$ 为该刚体对固定点 O 的**惯量主轴**。

主轴系中的惯量张量矩阵形式很简单, 所有非对角元都是 0。进一步, 在主轴系中, 角动量和角速度分量的关系 (5.72) 变成:

$$\begin{pmatrix} L_{x''} \\ L_{y''} \\ L_{z''} \end{pmatrix} = \begin{pmatrix} K_{x''x''} & 0 & 0 \\ 0 & K_{y''y''} & 0 \\ 0 & 0 & K_{z''z''} \end{pmatrix} \begin{pmatrix} \omega_{x''} \\ \omega_{y''} \\ \omega_{z''} \end{pmatrix}, \quad (5.95)$$

即:

$$\mathbf{L} = \omega_{x''} K_{x''x''} \mathbf{e}_{x''} + \omega_{y''} K_{y''y''} \mathbf{e}_{y''} + \omega_{z''} K_{z''z''} \mathbf{e}_{z''}. \quad (5.96)$$

将式 (5.72) 和式 (5.95、5.96) 对比, 我们看到在主轴系中角动量 \mathbf{L} 和角速度 $\boldsymbol{\omega}$ 的关系被大大简化。即: 一般情况下, 如式 (5.72) 所示, 角动量的任何一个分量都和角速度的全部三个分量有关。然而, 在主轴系中, 角动量在某惯量主轴 $l'' (l'' = x'', y'', z'')$ 方向的分量, 只和角速度在同方向的分量有关, 和角速度在另外两个分量无关。

在实际问题中, 通常尽可能把角动量、角速度等矢量在主轴系中分解, 并进行相关计算。所以, 对定点运动的刚体 (记固定点为 O), 我们往往需要找到其惯量主轴。

对均匀对称的刚体, 若 O 点在对称中心, 则惯量主轴记为其对称轴。比如, 对于质量均匀的立方体而言, 三个惯量主轴就是过中心且和三个边平行的坐标轴。对于质量均匀的有厚度的圆盘而言, 三个惯量主轴分别是过圆心且垂直于盘面的轴, 和过圆心, 平行于盘面的两个直角坐标轴。

对一般形状的刚体, 可先写出 \hat{K} 在任意与刚体固连的 $O - x'y'z'$ 系的矩阵 K , 再对角化 K , 即找到满足 $Q = Q^{-1}$ 以及 $QKQ^T = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$ 的 Q , 则 Q 的矩阵元为 $\mathbf{e}_\eta \cdot \mathbf{e}_{l'}$ ($\eta = \alpha, \beta, \gamma; l' = x', y', z'$), 其中, $\mathbf{e}_\alpha, \mathbf{e}_\beta, \mathbf{e}_\gamma$ 为主轴系基矢。

定点运动刚体的动能。

定点运动刚体的总动能可以表示为:

$$\begin{aligned} T &= \frac{1}{2} \sum_i m_i \mathbf{v}_i \cdot \mathbf{v}_i \\ &= \frac{1}{2} \sum_i m_i \mathbf{v}_i \cdot (\boldsymbol{\omega} \times \mathbf{r}_i) \\ &= \frac{1}{2} \boldsymbol{\omega} \cdot [\sum_i m_i (\mathbf{r}_i \times \mathbf{v}_i)] \\ &= \frac{1}{2} \boldsymbol{\omega} \cdot \mathbf{L} \\ &= \frac{1}{2} \boldsymbol{\omega} \cdot \hat{K} [\boldsymbol{\omega}] \end{aligned} \quad (5.97)$$

设 $O - \alpha\beta\gamma$ 系与刚体固连，则总动能可以用角动量、角速度以及惯量张量在该系中的分量表示为：

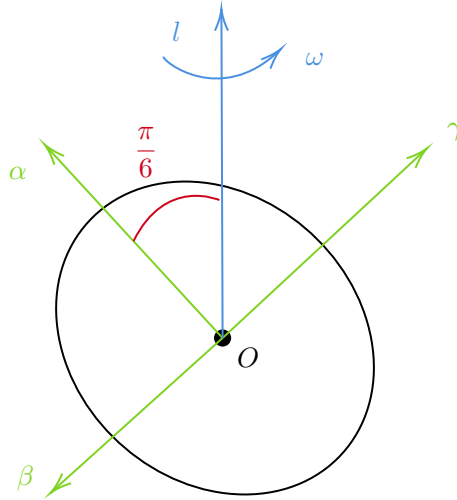
$$\begin{aligned}
 T &= \frac{1}{2} \sum_{\eta=\alpha,\beta,\gamma} \omega_{\eta} L_{\eta} \\
 &= \frac{1}{2} \sum_{\eta,\xi=\alpha,\beta,\gamma} \omega_{\eta} K_{\eta\xi} \omega_{\xi} \\
 &= \frac{1}{2} \begin{pmatrix} \omega_{\alpha} & \omega_{\beta} & \omega_{\gamma} \end{pmatrix} K \begin{pmatrix} \omega_{\alpha} \\ \omega_{\beta} \\ \omega_{\gamma} \end{pmatrix}
 \end{aligned} \tag{5.98}$$

其中， K 为 $O - \alpha\beta\gamma$ 系的惯量张量矩阵。进一步，若 $O - \alpha\beta\gamma$ 为主轴系，则

$$T = \frac{1}{2} \sum_{\eta=\alpha,\beta,\gamma} K_{\eta\eta} \omega_{\eta}^2. \tag{5.99}$$

其中， $K_{\eta\eta}$ ($\eta = \alpha, \beta, \gamma$) 为刚体对 $O - \eta$ 轴的转动惯量。

例题。



如图，考虑一个薄圆盘，质量为 m ，半径为 R 。假设圆盘绕着过圆心且和盘面夹角为 $\frac{\pi}{6}$ 的轴 l 转动，角速度为 ω ，求角动量，动能，以及圆盘绕转动轴的转动惯量。

解：

设 $O - \alpha\beta\gamma$ 为主轴系， $O\gamma$ 与圆盘垂直， $O\alpha, Ol, O\gamma$ 共面。由于刚体绕 l 轴定轴转动，故其角速度矢量为：

$$\begin{aligned}
 \boldsymbol{\omega} &= \omega \mathbf{e}_l \\
 &= \omega \left(\cos \frac{\pi}{6} \mathbf{e}_{\alpha} + \sin \frac{\pi}{6} \mathbf{e}_{\gamma} \right), \\
 &= \omega \left(\frac{\sqrt{3}}{2} \mathbf{e}_{\alpha} + \frac{1}{2} \mathbf{e}_{\gamma} \right)
 \end{aligned} \tag{5.100}$$

即

$$\begin{pmatrix} \omega_\alpha \\ \omega_\beta \\ \omega_\gamma \end{pmatrix} = \omega \begin{pmatrix} \frac{\sqrt{3}}{2} \\ 0 \\ \frac{1}{2} \end{pmatrix}. \quad (5.101)$$

另一方面， $O - \alpha\beta\gamma$ 系中的转动惯量矩阵为：

$$K = \frac{1}{4}mR^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}. \quad (5.102)$$

因此，刚体的角动量矢量可以写成 $\mathbf{L} = L_\alpha \mathbf{e}_\alpha + L_\beta \mathbf{e}_\beta + L_\gamma \mathbf{e}_\gamma$ ，其中

$$\begin{pmatrix} L_\alpha \\ L_\beta \\ L_\gamma \end{pmatrix} = K \begin{pmatrix} \omega_\alpha \\ \omega_\beta \\ \omega_\gamma \end{pmatrix} = \begin{pmatrix} \frac{m}{8}\omega R^2 \sqrt{3} \\ 0 \\ \frac{m}{4}\omega R^2 \end{pmatrix} \quad (5.103)$$

因此可以得到

$$T = \frac{1}{2} \boldsymbol{\omega} \cdot \mathbf{L} = \frac{5}{32} m R^2 \omega^2. \quad (5.104)$$

此外，刚体对 l 轴的转动惯量可以写成：

$$I_l = \mathbf{e}_l \cdot \hat{K} [\mathbf{e}_l]. \quad (5.105)$$

利用

$$\mathbf{e}_l = \frac{\sqrt{3}}{2} \mathbf{e}_\alpha + \frac{1}{2} \mathbf{e}_\gamma, \quad (5.106)$$

我们最终得到：

$$\begin{aligned} I_l &= \left(\frac{\sqrt{3}}{2} \mathbf{e}_\alpha + \frac{1}{2} \mathbf{e}_\gamma \right) \cdot \hat{K} \left[\frac{\sqrt{3}}{2} \mathbf{e}_\alpha + \frac{1}{2} \mathbf{e}_\gamma \right] \\ &= \frac{3}{4} K_{\alpha\alpha} + \frac{1}{4} K_{\gamma\gamma} + \frac{\sqrt{3}}{4} (K_{\alpha\gamma} + K_{\gamma\alpha}) \\ &= \frac{5}{16} m R^2. \end{aligned} \quad (5.107)$$

5.5 定点运动的刚体动力学

仿照定轴转动一样的思路，我们来给出定点转动情形下的刚体动力学规律

5.5.1 Euler 动力学方程

在经典力学中，我们有角动量定理，并且我们已经在 5.2.2 给出了其分量形式。我们再次给出其最一般的形式为

$$\frac{d\mathbf{L}}{dt} = \mathbf{M} \quad (5.108)$$

接下来，我们考虑其在某一个刚体固连坐标系下的具体形式。我们设坐标系 $O - x'y'z'$ 与刚体固连，则刚体的角动量具有形式 $\mathbf{L} = \sum_{l'=x'y'z'} L_{l'} \mathbf{e}_{l'}$ ，则

$$\frac{d\mathbf{L}}{dt} = \sum_{l'=x'y'z'} \dot{L}_{l'} \mathbf{e}_{l'} + \sum_{l'=x'y'z'} L_{l'} \dot{\mathbf{e}}_{l'} \quad (5.109)$$

而另一方面，角动量与角速度之间满足

$$L_{l'} = \sum_{m'=x'y'z'} K_{l'm'} \omega_{m'} \quad (5.110)$$

因此我们就有

$$\dot{L}_{l'} = \sum_{m'=x'y'z'} K_{l'm'} \dot{\omega}_{m'} \quad (5.111)$$

另一方面，我们有

$$\dot{\mathbf{e}}_{l'} = \frac{\partial \mathbf{e}_{l'}}{\partial \theta} \dot{\theta} + \frac{\partial \mathbf{e}_{l'}}{\partial \phi} \dot{\phi} + \frac{\partial \mathbf{e}_{l'}}{\partial \psi} \dot{\psi} \quad (5.112)$$

并且已经证明

$$\frac{\partial \mathbf{e}_{l'}}{\partial \theta} = \mathbf{e}_N \times \mathbf{e}_{l'}, \quad \frac{\partial \mathbf{e}_{l'}}{\partial \phi} = \mathbf{e}_z \times \mathbf{e}_{l'}, \quad \frac{\partial \mathbf{e}_{l'}}{\partial \psi} = \mathbf{e}_{z'} \times \mathbf{e}_{l'}$$

因此有

$$\dot{\mathbf{e}}_{l'} = (\dot{\theta} \mathbf{e}_N + \dot{\phi} \mathbf{e}_z + \dot{\psi} \mathbf{e}_{z'}) \times \mathbf{e}_{l'} = \boldsymbol{\omega} \times \mathbf{e}_{l'} \quad (l' = x', y', z') \quad (5.113)$$

将(5.111)(5.113)代入到(5.109)，我们就得到

$$\frac{d\mathbf{L}}{dt} = \sum_{l'=x'y'z'} \dot{L}_{l'} \mathbf{e}_{l'} + \sum_{l'=x'y'z'} L_{l'} \dot{\mathbf{e}}_{l'} = \sum_{l', m'=x'y'z'} K_{l'm'} \dot{\omega}_{m'} \mathbf{e}_{l'} + \boldsymbol{\omega} \times \mathbf{L} \quad (5.114)$$

这意味着

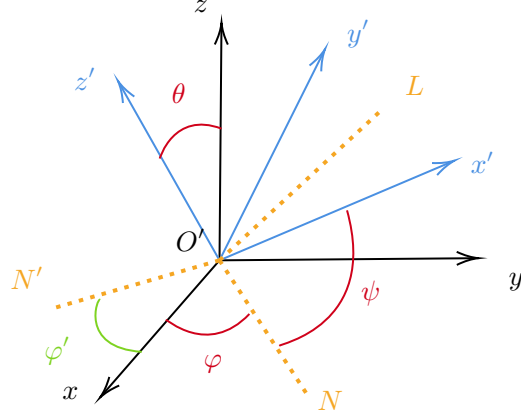
$$\mathbf{M} = \sum_{l', m'=x'y'z'} K_{l'm'} [\dot{\omega}_{m'} \mathbf{e}_{l'} + \omega_{m'} (\boldsymbol{\omega} \times \mathbf{e}_{l'})] \quad (5.115)$$

(5.115)将 $\omega_{l'}$ 和 $\dot{\omega}_{l'}$ 与 \mathbf{M} 联系在一起，在前文中的欧拉运动学方程(5.64)将 $\boldsymbol{\omega}$ 和 θ, ϕ, ψ 以及 $\dot{\theta}, \dot{\phi}, \dot{\psi}$ 联系在一起。于是(5.115)(5.64)就将广义坐标 θ, ϕ, ψ 以及 $\dot{\theta}, \dot{\phi}, \dot{\psi}$ 和 $\ddot{\theta}, \ddot{\phi}, \ddot{\psi}$ 联系在了一起，构成了关于 θ, ϕ, ψ 的动力学方程。如果我们选取 $O - x''y''z''$ 为主轴系，即使得

$$K_{l''m''} = I_{l''} \delta_{l''m''} \quad l'', m'' = x'', y'', z'' \quad (5.116)$$

里 I 承接惯量张量矩阵对角元的含义，是刚体在各个轴上的转动惯量。此时(5.115)就成为

$$\begin{cases} I_{x''} \dot{\omega}_{x''} - (I_{y''} - I_{z''}) \omega_{y''} \omega_{z''} = M_{x''} \\ I_{y''} \dot{\omega}_{y''} - (I_{z''} - I_{x''}) \omega_{z''} \omega_{x''} = M_{y''} \\ I_{z''} \dot{\omega}_{z''} - (I_{x''} - I_{y''}) \omega_{x''} \omega_{y''} = M_{z''} \end{cases} \quad (5.117)$$



现在, 我们考虑最一般运动的刚体, 即同时存在平动和转动的情形。设 O' 为刚体的质心, 我们同时建立两个坐标系, 分别为坐标轴不转动的质心系 $O'-xyz$, 以及与刚体固连的坐标系 $O'-x'y'z'$, 则描述刚体的广义坐标为 $(x_{O'}, y_{O'}, z_{O'}, \theta, \varphi, \psi)$, 则动力学方程包括两部分, 第一部分是描述平动的动力学方程

$$M\ddot{\mathbf{r}}_{O'} = \mathbf{F} \quad (5.118)$$

这里 m 是总质量, \mathbf{F} 是刚体所受的合外力, $\ddot{\mathbf{r}}_{O'}$ 是质心的加速度。第二部分是描述转动的

$$\frac{d\mathbf{L}_{O'}}{dt} = \mathbf{M}_{O'}$$

这里 $\mathbf{M}_{O'}$ 是外力对质心的外力矩。

在下面几节中, 我们讨论几种特殊的刚体运动模式。

5.5.2 Euler-Pansel 情况

如果刚体不受外力矩, 我们称此时刚体处于 Euler-Pansel 情况, 即此时 $\mathbf{M} = \frac{d\mathbf{L}}{dt} = 0$ 。我们称刚体 ψ 变化是自转, ϕ 变化是进动, θ 变化是章动。

设 $O-x'y'z'$ 为主轴系, 且 $I_{x'} = I_{y'} \neq I_{z'}$, 并设 z 方向为刚体的总角动量方向, 即 $\mathbf{L} = L\mathbf{e}_z$, 欧拉运动学方程被表达为

$$\begin{cases} \omega_{x'} = \dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi \\ \omega_{y'} = \dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi \\ \omega_{z'} = \dot{\phi} \cos \theta + \dot{\psi} \end{cases} \quad (5.119)$$

从而角动量

$$\mathbf{L} = I_{x'}\omega_{x'}\mathbf{e}_{x'} + I_{y'}\omega_{y'}\mathbf{e}_{y'} + I_{z'}\omega_{z'}\mathbf{e}_{z'} = L\mathbf{e}_z \quad (5.120)$$

下面我们将第二个等号两边的表达式, 统一分解到 Euler 角辅助线基矢 $\mathbf{e}_{z'}, \mathbf{e}_N, \mathbf{e}_L$ 之下。由

于

$$\begin{cases} \mathbf{e}_z = \cos \theta \mathbf{e}_{z'} + \sin \theta \mathbf{e}_L \\ \mathbf{e}_{x'} = \cos \psi \mathbf{e}_N + \sin \psi \mathbf{e}_L \\ \mathbf{e}_{y'} = -\sin \psi \mathbf{e}_N + \cos \psi \mathbf{e}_L \end{cases} \quad (5.121)$$

因而, (5.120)第二个等号的右边可以被改写为

$$L \mathbf{e}_z = L \cos \theta \cdot \mathbf{e}_{z'} + L \sin \theta \cdot \mathbf{e}_L$$

同时, 联合 Euler 运动学方程 (5.64) 和上面的基矢变换 (5.121), 可以发现 (5.120) 第二个等号左边可以被改写为

$$\mathbf{L} = I_{x'} \omega_{x'} \mathbf{e}_{x'} + I_{y'} \omega_{y'} \mathbf{e}_{y'} + I_{z'} \omega_{z'} \mathbf{e}_{z'} = I_{x'} \dot{\theta} \cdot \mathbf{e}_N + I_{x'} \dot{\phi} \sin \theta \cdot \mathbf{e}_L + I_{z'} (\dot{\phi} \cos \theta + \dot{\psi}) \cdot \mathbf{e}_{z'} \quad (5.122)$$

我们对系数可以得到

$$\begin{cases} \dot{\theta} = 0 & (\theta = \theta_0) \\ \dot{\phi} = \frac{L}{I_{x'}} \\ \dot{\psi} = \left(\frac{1}{I_{z'}} - \frac{1}{I_{x'}} \right) L \cos \theta_0 \end{cases} \quad (5.123)$$

于是我们可以发现在 Euler-Pansel 情况下, 刚体无章动, 绕着对称轴匀速自转, 匀速进动, 进动方向保持为绕 \mathbf{L} 进行。若 $\theta_0 = 0$, 则 z' 轴与 z 轴重合, 此时刚体做匀速的定轴转动, 转动方向在惯性系中保持不变。

5.5.3 高速陀螺的近似理论

最后, 我们来讨论一个高速旋转的陀螺。我们不考虑陀螺所发生的平动, 而只考虑陀螺在地面上原地打转。为了方便起见, 固连坐标系的原点我们不取在质心, 而是取在陀螺与地面的接触点, 并建立主轴系 $O - x'y'z'$, 如图 5.13 所示。我们假设陀螺的质量分布绕着陀螺的旋转轴是对称的, 这意味着其转动惯量满足 $I_{x'} = I_{y'}$ 。我们假设质心到 O 点的距离为 l 。

物理直观上, 高速旋转的陀螺做定点运动, 并且高速自转, 缓慢进动, 无章动, 即 $\dot{\theta} \approx 0, |\dot{\psi}| \gg |\dot{\phi}|$, 以下我们推导进动角速度 $\dot{\phi}$ 和自转角速度 $\dot{\psi}$ 以及 l , 质量和倾角 θ 的关系。

刚体的角速度矢量有

$$\boldsymbol{\omega} = \dot{\phi} \mathbf{e}_z + \dot{\psi} \mathbf{e}_{z'} + \dot{\theta} \mathbf{e}_N \approx \dot{\psi} \mathbf{e}_{z'} \quad (5.124)$$

于是角动量有

$$\mathbf{L} = \hat{K}[\boldsymbol{\omega}] = \hat{K}[\dot{\psi} \mathbf{e}_{z'}] = I_{z'} \dot{\psi} \mathbf{e}_z \quad (5.125)$$

进一步地, 我们对角动量对时间求导

$$\frac{d\mathbf{L}}{dt} = I_{z'} \ddot{\psi} \mathbf{e}_{z'} + I_{z'} \dot{\psi} (\boldsymbol{\omega} \times \mathbf{e}_{z'}) = I_{z'} \ddot{\psi} \mathbf{e}_{z'} + I_{z'} \dot{\psi} [(\dot{\phi} \mathbf{e}_z + \dot{\psi} \mathbf{e}_{z'}) \times \mathbf{e}_{z'}] = I_{z'} \ddot{\psi} \mathbf{e}_{z'} + I_{z'} \dot{\phi} \dot{\psi} \sin \theta \cdot \hat{\mathbf{e}}_N \quad (5.126)$$

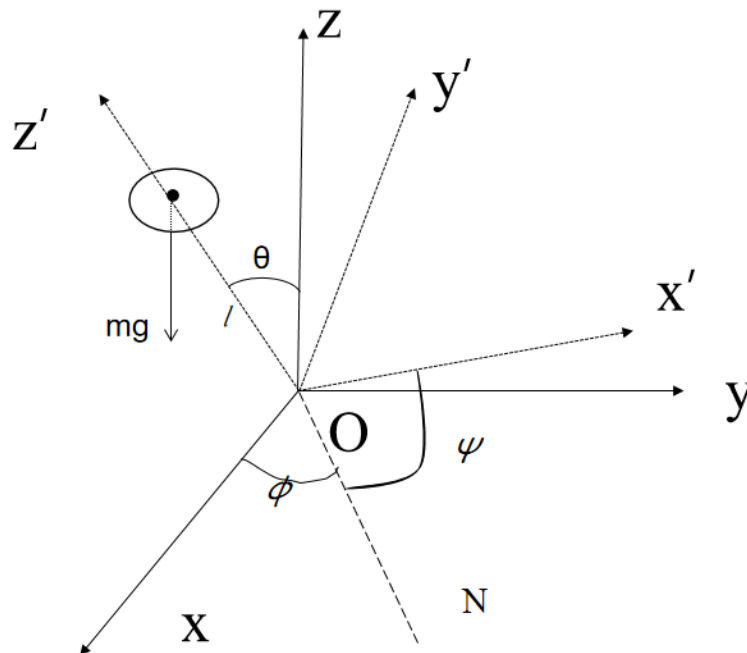


图 5.13: 高速陀螺的坐标系建立及受力分析

另一方面，我们从定义出发写出力矩 \mathbf{M} 的表达式，则

$$\mathbf{M} = \mathbf{r}_{O'} \times (-mg\mathbf{e}_z) = l\mathbf{e}_{z'} \times (-mg\mathbf{e}_z) = mgl \sin \theta \cdot \mathbf{e}_N \quad (5.127)$$

而我们有角动量定理 $\frac{d\mathbf{L}}{dt} = \mathbf{M}$ ，从而

$$I_{z'}\ddot{\psi}\mathbf{e}_{z'} + I_{z'}\dot{\psi}\dot{\phi}\sin\theta\mathbf{e}_N = mgl \sin \theta\mathbf{e}_N \quad (5.128)$$

于是对比各基矢前的分量可以得到

$$\ddot{\psi} = 0 \quad (5.129)$$

$$\dot{\phi} = \frac{mgl}{I_{z'}\dot{\psi}} \quad (5.130)$$

第一个式子意味着高速旋转陀螺的自转是匀速的。第二个式子指出陀螺匀速进动，进动角速度与倾角 θ 无关，而反比于自转角速度。

5.6 刚体动力学小结

最后我们再次回顾整个刚体动力学的研究思路。

在研究刚体时，我们需要建立两套坐标系，分别是保持不变的惯性系 $O - xyz$ 以及随刚体一起转动的与刚体固连的参考系 $O - x'y'z'$

从刚体的广义坐标出发，我们定义了刚体的角速度 $\boldsymbol{\omega}$ 。于是在定轴定点的无平动情形下，刚体上任意一个质点 A 都有 $\mathbf{v}_A = \boldsymbol{\omega} \times \mathbf{r}_A$ 。而在一般的二维或三维运动中，刚体上质点 B 相对于质点 A 则有 $\mathbf{v}_{AB} = \boldsymbol{\omega} \times \mathbf{r}_{AB}$

在完成角速度的定义后，角速度 $\boldsymbol{\omega}$ 可以通过一个线性变换变为角动量 \mathbf{L} 。在二维平面运动夏，刚体做定轴转动，则 $L_z = I_z \omega$ 。这里我们的旋转轴为 z 方向， L_z 是我们对定轴转动时所唯一需要关心的物理量。而在三维转动情形下， $\mathbf{L} = \hat{K}[\boldsymbol{\omega}]$ ，这里的 \hat{K} 是一个惯量张量变换，在确定固连坐标系以后，转动惯量张量可以用一个三阶方阵表示。

刚体的动力学描述则由角动量定理所给出。在刚体的二维运动下，角动量定理给出 $\frac{dL_z}{dt} = M_z$ 。而在三维运动时，则有 $\frac{d\mathbf{L}}{dt} = \mathbf{M}$

在对角动量和动力学进行描述时，若刚体没有平动，我们可以直接选取一个惯性系进行描述。而如果加上了刚体的平动使之进行一般的二维或三维运动，我们一般选取质心系来描述其中的转动。