Lagrange 乘子法

前面我们讨论了完整约束体系，对于这样的体系不独立坐标可以通过约束方程（或者引入广义坐标）加以消除。如果约束是非完整的，我们就无法通过约束方程达到消除不独立坐标的目的。没有一般性的方法处理非完整约束问题，但是，对于一类特殊的非完整约束，也就是线性微分约束，尽管无法通过约束方程消去不独立坐标，但是我们却可以通过引入 Lagrange 乘子消去不独立的方程。

现在我们就考虑受到约束方程有如下形式的一个力学体系

$$A_{\nu k}\dot{q}_k + A_{\nu t} = 0, \quad \text{or} \quad A_{\nu k}dq_k + A_{\nu t}dt = 0 \quad \left(\nu = 1, 2, \cdots, m\right) \tag{1}$$

这里，n 是坐标的数目，m 是约束的数目，并且 $n > m$；$A_{\nu k}$ 和 $A_{\nu t}$ 都只是坐标和时间的函数。

无论上面的方程是否可积都不会影响我们后面的讨论：这些讨论对于完整约束和非完整约束同样有效。因此，如果不方便把所有的坐标都化为独立变量，或者如果你想要考虑约束效应，那么，下面推导的 Lagrange 乘子方法也适用于完整约束体系。例如，对于约束 $f_\nu\left(q,t\right) = 0$，两边对时间求微商就变为（可积）微分约束的形式 $\sum_k \left(\partial f_\nu / \partial q_k\right)\dot{q}_k + \partial f_\nu / \partial t = 0$（这也正是我们把完整约束有时也称为可积约束的原因）。这时 $A_{\nu k} = \partial f_\nu / \partial q_k$，$A_{\nu t} = \partial f_\nu / \partial t$。

象上一节推导完整约束体系的 Lagrange 方程一样，我们仍然从最小作用原理出发

$$\delta S = \delta \int_{t_1}^{t_2} L\left(q, \dot{q}, t\right) dt = 0 \tag{2}$$

我们假定它对于非完整约束也是正确的。因此，完全类似于以前的推导，我们有

$$\int_{t_1}^{t_2} \left[\left(\frac{\partial L}{\partial q_k} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} \right) \delta q_k \right] dt = 0 \tag{3}$$

这里 q_i 不是独立的，从而变分（或者虚位移）δq_k 也不能任意取值，所以我

们从上面的积分等于零也不再能得到通常形式的 Lagrange 方程。跟上一节不同的是，你无法通过约束方程把不独立的变量消去。为了把变分约化为独立的形式，我将在下面引入可任意选取的m个 Lagrange 乘子λ_v，$\left(v=1,2,\cdots,m\right)$，一般情况下，$\lambda_v$是坐标$q_i$、速度$\dot{q}_i$和时间$t$的函数。由于变分运算是在固定的时间操作的（等时变分），也就是说$\delta t=0$，因此，变分满足下面的条件（正如完全约束中我们得到虚位移所满足的方程一样的道理）

$$A_{vk}\delta q_k=0 \tag{4}$$

实际上，这个条件仅仅对于完整约束（当然也就包含可积分的微分约束）才是成立的，而对于一般的不可积线性微分约束，你可以认为我们仅仅考虑限于使得上式成立的那些想象路线，而所谓最小也是在满足个关系的所有路径当中而言的。这个方程左右两边都乘上任意函数λ_v也是成立的，把所有这样得到的方程相加（对v求和）

$$\lambda_v A_{vk}\delta q_k=0 \tag{5}$$

或者把求和重新安排并从t_1到t_2对时间积分

$$\int_{t_1}^{t_2}\left[\left(\lambda_v A_{vk}\right)\delta q_k\right]dt=0 \tag{6}$$

现在我手上由两个等于零的积分，(3)和(6)，把这两个积分相加显然也是等于零的

$$\int_{t_1}^{t_2}\left[\left(\frac{\partial L}{\partial q_k}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k}+\lambda_v A_{vk}\right)\delta q_k\right]dt=0 \tag{7}$$

这个积分涉及到n个变量q_k：其中m个q_k是不独立的，它们通过约束方程与其它的q_k相联系。我们将假设这m个不独立变量为q_k（指标从1到m取值），而$n-m$个独立变量为q_k（指标从$m+1$到n取值，这样做并不影响下面分析得到的结论）。由于m个 Lagrange 乘子λ_v可以取任何函数，因此，我总是可以找到适当的λ_v使得下面的方程成立

$$\frac{\partial L}{\partial q_k}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k}+\sum_{v=1}^{m}\lambda_v A_{vk}=0 \qquad (v=1,2,\cdots,m) \tag{8}$$

即方程(7)积分中对应于前面m个不独立坐标变分的系数为零。

这样选取λ_v之后，方程(7)剩下的积分就是

$$\int_{t_1}^{t_2} dt\sum_{k=m+1}^{n}\left(\frac{\partial L}{\partial q_k}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k}+\sum_{v=1}^{m}\lambda_v A_{vk}\right)\delta q_k=0 \tag{9}$$

由于这些q_k（对于$k=m+1,\cdots,n$）不再受到任何限制（根据我们的假定它们是独立的），因此，正如以前所熟悉的，上面积分中每个q_k前面的系数都等于零。这些方程连同前面得到的不独立变量的m个方程就可以统一表示为

$$\frac{\partial L}{\partial q_k}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k}+\lambda_v A_{vk}=0 \qquad (v=1,2,\cdots,n) \tag{10}$$

这n个方程涉及$n+m$个未知量：n个坐标以及m个 Lagrange 乘子。m个约束方程正好为我们提供了所需要的其余m个方程。

Lagrange 乘子λ_v有什么样的物理含义呢？我们知道每一个质点的 Newton 方程可以写为

$$\vec{F}_a+\vec{N}_a-\frac{d\vec{p}_a}{dt}=0 \tag{11}$$

其中$\vec{F}_a=\partial L/\partial \vec{r}_a=-\nabla_a U$，而$\vec{p}_a=\partial L/\partial \dot{\vec{r}}_a=m_a\dot{\vec{r}}_a$，因此上式也可以写为

$$\frac{\partial L}{\partial \vec{r}_a}-\frac{d}{dt}\frac{\partial L}{\partial \dot{\vec{r}}_a}+\vec{N}_a=0 \tag{12}$$

如果把这个等式两边与$\partial \vec{r}_a/\partial q_k$作标量积并对所有的粒子（即对$a$）求和，那么

$$\left(\frac{\partial L}{\partial \vec{r}_a}-\frac{d}{dt}\frac{\partial L}{\partial \dot{\vec{r}}_a}\right)\cdot\frac{\partial \vec{r}_a}{\partial q_k}+\vec{N}_a\cdot\frac{\partial \vec{r}_a}{\partial q_k}=0 \tag{13}$$

上一节我们已经知道第一项等于

$$\frac{\partial L}{\partial q_k}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} \tag{14}$$

而第二项则是广义约束力

$$Q_k' = \vec{N}_a \cdot \frac{\partial \vec{r}_a}{\partial q_k} \tag{15}$$

而方程(13)必然与我们前面得到的方程(10)是一样的，所以

$$Q_k' = \lambda_\nu A_{\nu k} \tag{16}$$

也就是说，Lagrange 乘子λ_ν确定了广义约束力，它本身也是问题解的一部分。

最后请大家注意这样一点：关系式(5)意味着

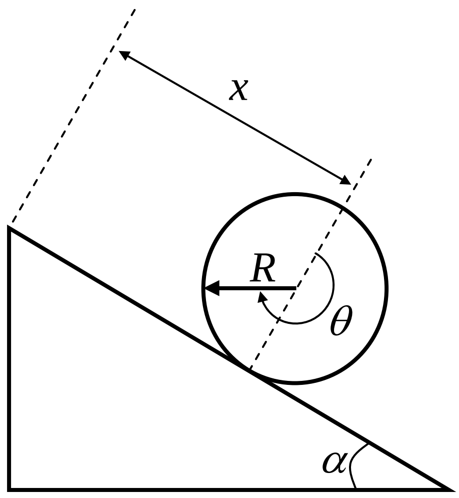
$$Q_k' \delta q_k = 0 \tag{17}$$

也就是说所有约束力的虚功等于零。这可以看作是理想约束假设的一般性的证明。

举一个例子。考虑一个圆盘在倾斜平面上的纯滚动。设圆盘的质量为m，半径为R。圆盘的动能分为两部分：平移动能和转动动能（实际上就是质心动能和相对于质心的动能，这一点我们在质点组部分已经证明）：

$$T=\frac{1}{2}m\dot{x}^2+\frac{1}{2}I\dot{\theta}^2=\frac{1}{2}m\dot{x}^2+\frac{1}{4}mR^2\dot{\theta}^2 \tag{18}$$

第二部分动能也可以直接计算得到，它等于



$$\begin{aligned}\int\frac{1}{2}v^2dm &= \int\frac{1}{2}\rho\left(r\dot{\theta}\right)^2 rdrd\theta \\ &= \frac{1}{2}\rho\dot{\theta}^2\int r^3drd\theta \\ &= \frac{1}{2}\rho\dot{\theta}^2\cdot\frac{1}{4}R^4\cdot 2\pi \\ &= \frac{1}{4}\cdot\left(\rho\pi R^2\right)R^2\cdot\dot{\theta}^2=\frac{1}{4}mR^2\dot{\theta}^2\end{aligned}$$

势能为

$$U=-mgx\sin\alpha \tag{19}$$

这里我们假设在斜面顶端圆盘的势能等于零。因此 Lagrange 函数为

$$L=T-U=\frac{1}{2}m\dot{x}^2+\frac{1}{4}mR^2\dot{\theta}^2+mgx\sin\alpha \tag{20}$$

约束方程为

$$f\left(x,\theta\right)=x-R\theta=0 \tag{21}$$

由于圆盘作纯滚动，这个体系只有一个自由度，因此我们可以选择 x 或者 θ 作为广义坐标，并利用约束方程将另一个变量消去。但是，我们也可以把 x 和 θ 都当作广义坐标，而利用 Lagrange 乘子法求解这个问题。此时 Lagrange 方程为

$$\begin{aligned}0&=\frac{\partial L}{\partial x}-\frac{d}{dt}\frac{\partial L}{\partial \dot{x}}+\lambda\frac{\partial f}{\partial x}=mg\sin\alpha-m\ddot{x}+\lambda\\0&=\frac{\partial L}{\partial \theta}-\frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}}+\lambda\frac{\partial f}{\partial \theta}=0-\frac{1}{2}mR^2\ddot{\theta}-\lambda R\end{aligned} \tag{22}$$

这两个方程连同约束方程 $x=R\theta$ 就可以完全确定未知量 x 、θ 和 λ 。把约束方程对时间求导得到

$$\ddot{\theta}=\ddot{x}/R \tag{23}$$

代入前面两个方程并作简单的组合就得到

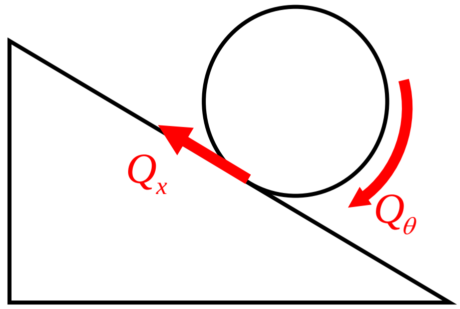
$$\ddot{x}=\frac{2}{3}g\sin\alpha,\quad \ddot{\theta}=\frac{2g}{3R}\sin\alpha,\quad \lambda=-\frac{1}{3}mg\sin\alpha \tag{24}$$

我们注意到如果圆盘沿着斜面无摩擦地向下滑动，那么 $\ddot{x}=g\sin\alpha$ 。因此，滚动约束使得加速度的值减小一个因子 $2/3$ 。

根据前面的定义，广义约束力等于

$$\begin{aligned}Q_x&=\lambda\frac{\partial f}{\partial x}=\lambda=-\frac{1}{3}mg\sin\alpha\\Q_\theta&=\lambda\frac{\partial f}{\partial \theta}=-\lambda R=\frac{1}{3}mgR\sin\alpha\end{aligned} \tag{25}$$

注意这里Q_x正是滚动约束力，它的方向沿着斜面向上（与x增加的方向相反），而Q_θ则是这个力相对于质心的力矩（其效应使得θ增加）。它们是使得圆盘保持沿着平面纯滚动所需要的广义约束力。



如果我们通过方程$\dot{\theta}=\dot{x}/R$把$\dot{\theta}$从 Lagrange 函数中消去

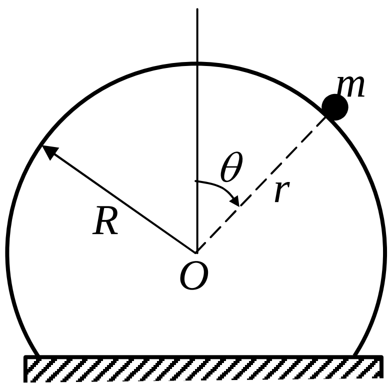
$$L=\frac{3}{4}m\dot{x}^2+mgx\sin\alpha \tag{26}$$

这时 Lagrange 函数就用一个独立的广义坐标表示了出来，这个坐标的运动方程马上可以得到

$$\frac{\partial L}{\partial x}-\frac{d}{dt}\frac{\partial L}{\partial \dot{x}}=mg\sin\alpha-\frac{3}{2}m\ddot{x}=0 \tag{27}$$

与我们前面得到的方程是一样的。尽管这个方法更加简单，但是它却不会告诉你有关约束力的任何信息。

另举一例。考虑固定圆球顶端的一个静止的小珠，小珠的质量为m，圆球的半径为R，并假设圆球表面是光滑的。现在如果给小珠一个很小的扰动，它就会运动，由于球面的存在，它不能自由下落，而是会沿着球面滑下。但是，在重力场中，小珠也不可能永远在球面上运动，在某个时刻，即当约束力等于零的时候，它就会开始脱离球面，而我们的问题就是找出开始脱离球面时小珠的位置。



这是一个单自由度的问题，因此用一个变量（如角度θ）就可以完全确定小珠的位置，但是，为了得到约束力的信息，我们就必须把不独立的那个变量（如径向距离r）也保留在 Lagrange 函数中，即

$$L=T-U=\frac{1}{2}m\left(\dot{r}^2+r^2\dot{\theta}^2\right)-mgr\cos\theta \tag{28}$$

而这两个变量所满足的约束方程简单地就是

$$f(r,\theta) = f(r) = r - R = 0 \tag{29}$$

写出带乘子的 Lagrange 方程

$$\begin{aligned}
0 &= \frac{\partial L}{\partial r} - \frac{d}{dt}\frac{\partial L}{\partial \dot{r}} + \lambda \frac{\partial f}{\partial r} = mr\dot{\theta}^2 - mg\cos\theta - m\ddot{r} + \lambda = 0 \\
0 &= \frac{\partial L}{\partial \theta} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}} + \lambda \frac{\partial f}{\partial \theta} = mgr\sin\theta - \frac{d}{dt}\left(mr^2\dot{\theta}\right) + 0 = 0
\end{aligned} \tag{30}$$

将约束方程$r = R$代入上式就得到了

$$\lambda = mg\cos\theta - mR\dot{\theta}^2, \qquad R\ddot{\theta} = g\sin\theta \tag{31}$$

而第二个方程又可以写为

$$R\dot{\theta}\frac{d\dot{\theta}}{d\theta} = g\sin\theta \tag{32}$$

将此式积分并利用初始条件$\theta(t=0) = \dot{\theta}(t=0) = 0$，就得到了

$$R\dot{\theta}^2 = 2g(1-\cos\theta) \tag{33}$$

将此关系式再代回到(31)中的第一个方程，我们得到

$$\lambda = mg(3\cos\theta - 2) \tag{34}$$

在我们的问题中，广义约束力Q_r就是球面对小珠所施加的约束力N_r，即

$$Q_r = \vec{N}\cdot\frac{\partial \vec{r}}{\partial r} = \vec{N}\cdot\hat{r} = N_r = \lambda\frac{\partial f}{\partial r} = \lambda = mg(3\cos\theta - 2) \tag{35}$$

因此，当$N_r = 0$或者说

$$\cos\theta = \frac{2}{3} \quad \text{or} \quad \theta = \arccos\frac{2}{3} \tag{36}$$

时小珠开始脱离球面。至于另一个广义约束力则是这个约束力相对于球心的力矩，当然由于约束力完全是径向的，因此，这个力矩也就是等于零的：

$$Q_\theta = \vec{N}\cdot\frac{\partial \vec{r}}{\partial \theta} = r\vec{N}\cdot\hat{\theta} = \lambda\frac{\partial f}{\partial \theta} = 0 \tag{37}$$

附录：关于虚位移所满足的方程(4)的一点说明

为简单起见，考虑单个约束的情况：

$$g(q,\dot{q},t)=\sum_{k=1}^{n}A_k(q,t)\dot{q}_k+B(q,t)=0 \tag{A1}$$

想象路线满足约束意味着

$$\begin{aligned}\delta g&=g(q+\delta q,\dot{q}+\delta\dot{q},t)-g(q,\dot{q},t)\\&=(\delta A_k)\dot{q}_k+A_k\delta\dot{q}_k+\delta B\\&=\frac{\partial A_k}{\partial q_i}\dot{q}_k\delta q_i+A_k\delta\dot{q}_k+\frac{\partial B}{\partial q_k}\delta q_k=0\end{aligned} \tag{A2}$$

注意到第二项又可以写成

$$\begin{aligned}A_k\delta\dot{q}_k&=A_k\frac{d\delta q_k}{dt}=\frac{d}{dt}(A_k\delta q_k)-\frac{dA_k}{dt}\delta q_k\\&=\frac{d}{dt}(A_k\delta q_k)-\left(\frac{\partial A_k}{\partial q_i}\dot{q}_i+\frac{\partial A_k}{\partial t}\right)\delta q_k\end{aligned} \tag{A3}$$

将它代入(A1)时得到

$$\begin{aligned}\delta g&=\frac{\partial A_k}{\partial q_i}\dot{q}_k\delta q_i-\frac{\partial A_k}{\partial q_i}\dot{q}_i\delta q_k-\frac{\partial A_k}{\partial t}\delta q_k+\frac{\partial B}{\partial q_k}\delta q_k+\frac{d}{dt}(A_k\delta q_k)\\&=\left(\frac{\partial A_i}{\partial q_k}-\frac{\partial A_k}{\partial q_i}\right)\dot{q}_i\delta q_k-\left(\frac{\partial A_k}{\partial t}-\frac{\partial B}{\partial q_k}\right)\delta q_k+\frac{d}{dt}(A_k\delta q_k)=0\end{aligned} \tag{A4}$$

对于可积线性微分约束，由于有（可积条件，这里t也应看作一个变量）

$$\frac{\partial A_i}{\partial q_k}=\frac{\partial A_k}{\partial q_i}\quad\text{and}\quad\frac{\partial A_k}{\partial t}=\frac{\partial B}{\partial q_k} \tag{A5}$$

因此(A4)就告诉我们

$$\frac{d}{dt}(A_k\delta q_k)=0 \tag{A6}$$

或者说

$$A_k\delta q_k=\text{constant} \tag{A7}$$

特别是，由于端点处δq_k是等于零的，因此这个常数实际上也是等于零的，即

$$A_k \delta q_k = 0 \tag{A8}$$

这个结论没什么稀奇的，它就是我们在上一节对于完整约束所得到的关系

$$\delta f = f\left(q+\delta q, t\right) - f\left(q, t\right) = \frac{\partial f}{\partial q_k}\delta q_k = 0 \tag{A9}$$

这里

$$A_k = \frac{\partial f}{\partial q_k}, \quad B = \frac{\partial f}{\partial t} \tag{A10}$$

重要的是，对于不可积线性微分约束，上面的过程实际上表明：想象路线满足约束并不意味着虚位移满足关系(A8)或者(4)！这并不就说明我们前面的讨论出错了，而是说，对于这样的体系，其最小作用原理应该重新表述为：真实运动是沿着满足(4)的想象路线当中作用量取最小值的那条路线行进的。用数学表示就是

$$\begin{cases} \delta S = \delta\int_{t_1}^{t_2} L\left(q, \dot{q}, t\right) dt = 0 \\ \delta q_k\left(t_1\right) = \delta q_k\left(t_2\right) = 0 \quad \text{and} \quad A_{vk}\delta q_k = 0 \end{cases} \tag{A11}$$