

# 最速降线、摆线推导

小圆滚滚

## 1 泛函

简单的说，泛函的定义域是函数集，值域是数集，也就是说，泛函是从函数空间到数域的一个映射。

实际上，推广开来，函数实际上是一种特殊的二元关系，泛函的二元关系是二阶笛卡尔积，所以函数集实际上是一个向量空间，所以泛函也可以说是从向量空间到标量的一个映射。

简而言之，泛函就是函数的函数。

泛函和函数的区别是：函数是变量和变量的关系，而泛函是函数和变量的关系。

## 2 引例

### 2.1 最短路径

众所周知，两点之间，线段最短。这是欧式几何的公理之一。从几何的角度很容易证明（反证）。但是，有没有解析的办法来证明呢？

答案是有的。

设平面上存在两点 $X_1, X_2$ ， $y = f(x)$ 是平面上经过这两点的任意曲线。我们的目的，是求一个距离最短的 $f(x)$ ，这就是函数和数之间的二元关系。显然，这里的定义域是 $f(x)$ ，也就是函数，而值域是距离，是一个数。所以这种关系是一个泛函，记为 $A[f]$ 。

弧微分的几何意义是用一条线段的长度来近似代表一段弧的长度。MT的长度即为弧MM'的微分，由此联系勾股定理可得弧微分公式：

$$\begin{aligned}(ds)^2 &= (dx)^2 + (dy)^2 \\ &= (dx)^2 + (y'dx)^2\end{aligned}$$

$$\text{故}(ds) = \sqrt{1 + y'^2} dx$$

由弧微分公式易知：

对微分方程定积分，可得原函数在某一阶段的应用函数，记为： $A[f] = \int_{x_1}^{x_2} \sqrt{1 + y'^2} dx$ 。

### 2.2 最速降线

最速降线问题是伽利略提出的著名问题：一个质点在重力作用下，从一个给定点到不在它垂直下方的另一点，如果不计摩擦力，问沿着什么曲线滑下所需时间最短。换句话说，就是一个质点不是垂直的下落，只有重力做功，沿什么曲线下落所需时间最短。

这里的问题关键，是找到一条最优的曲线。下落时间是一个数值，所以这这也是一个泛函问题。

根据牛顿第二定律： $F = ma$

外力做功： $W = Fl = Pt$

动量定理： $F\Delta t = m\Delta v = mv_{\text{末}} - mv_{\text{初}}$

动能定理： $E_k = \frac{1}{2}mv^2$

重力势能： $E_p = mgh$

二维运动的总功： $W = Fs \cdot \cos\alpha$ (初中阶段，力方向与位移方向的夹角为 $0$ ，即 $\alpha = 0^\circ$ ， $\cos 0^\circ = 1$ ，所以 $W = Fs$ )

路程等于速度乘以时间：

$$\begin{aligned} S &= \bar{v}t, v_0 = 0, v_t = at \\ &= \frac{0 + v_t}{2} \cdot t \\ &= \frac{0 + at}{2} \cdot t \\ &= \frac{at^2}{2} \end{aligned}$$

由于只有重力做功 $\frac{1}{2}mv^2 = mgh$ ，所以下落 $y$ 之后的下落速度为 $\sqrt{2gy}$ 。

选取这条曲线上很小的一段，这一段中，速度可以看成不变。

易知通过这一段的时间(距离除以速度)是 $\sqrt{\frac{1+y'^2}{2gy}}$ ，分子是这一小段的弧长(利用弧微分公式)，分母是速度，所以这个值是通过这一段的时间。做一次积分，可以得到：

$$A[f] = \int_{x_1}^{x_2} \sqrt{\frac{1+y'^2}{2gy}} dx$$

这就是总的时间，要求一个最优的 $f$ ，使得该时间(该泛函)的值最小。

### 3 欧拉-拉格朗日定理(方程)

因为涉及到光子的运动和能量级传输。所以在波粒二象性、动能势能转换、自旋方面。有着重要理论基础。参数方程代替可分离变量的微分方程(或者说将 $x$ 替换成一个仅与 $x$ 相关的函数)，以拉格朗日中值定理为基础可以得到柯西中值定理。

多元函数代替一元函数(在分析研究时可先讨论经典力学(可分离任意维度来研究，比如 $x$ 轴、1维)，量子场论需要增加所有空间维度、当然必有默认的时间维度才能使方程成立)，泛函代替标量对标量的映射，用偏导数、投影来解释的拉格朗日中值定理，就是欧拉-拉格朗日方程。

#### 3.1 拉格朗日函数

首先介绍拉格朗日函数，在经典牛顿物理学中，一个系统的拉格朗日函数是系统的总动能减去总势能(做功方程(微分方程))。代表着一个匀速直线运动或静止的物体，自身稳定状态的能量。假设静止(方便画图分析)，是一个基于两个自变量：一个事 $x$ 轴的位置，一个是相对于 $x$ 轴的速度，来刻画系统状态、预测路径分析的函数。当一个物体由 $A$ 点移动到 $B$ 点，整个过程由静到动，再由动到静。是有各种外力各个方向(函数)作用的结果。虽然发生了位移。但是总功为 $0$ 。简单举例，正向推了多少米、加速到了多少速度，你就需要反向推维持相同的时间(自然米数也就一样了)。以既成事实来反观物体运动的过程，拉格朗日函数正是对应于时间的功率函数(包含两个自变量，这两个自变量又同时和时间相关)。按时间积分拉格朗日函数可以得到对整个系统的作用量值(做功的面积值(积分方程))。

## 3.2 作用量函数

将上述的作用量值作为因变量，将路径作为自变量，可以得到作用量函数。该函数描述了一个系统，总功只取决于系统的初始状态和最后状态，和它们之间的路径无关。也就是拉格朗日函数值发生了改变，但是可以造成很多种表现形式。最后物体的位置和速度有可能大不相同。若想猜测唯一的路径，需要一个函数来描述不同理想（路径）和作用量值的关系。如果路径改变了，那么作用量值也将对应发生变化。

$F = ma$ 微分运动方程数学等价于其对应的积分运动方程，这具有很重要的哲学意义。微分方程和积分方程当然是微积分的一体两面了。

有了拉格朗日函数，你可以知道在你研究的分量上，比如x轴，此刻物体的位置，以及下一刻预测的物体的位置。如果该函数在某一段时间内受到轻微扰动，那么物体运动的路径一定会发生变化。那么在什么时刻扰动造成的影响最小呢（事半功半和事半功倍的切入点）？

### 3.2.1 最小作用量原理

当作用量函数 $f$ ，作用量值（是一个面积值以下简称作用量）是因变量 $y$ ，路径变化是自变量 $x$ ，斜率为0时，说明物体能量没有变化，不受外力，没有扰动。当然，作用量函数在最低点（最小值）时候肯定满足这个条件。用这个方法求路径就是最小作用量原理。但是更精确的说法不光包含最低点，还有所有斜率为0的点（平稳作用量原理）。

意义：路径变化越大，说明作用量变化越大。比如，你用1牛的力，正推一分钟，再反向推一分钟，物体位移的距离，和你用2牛的力，正推一分钟，再反向推一分钟，物体位移的距离这个变化就是路径的变化。单位时间内作用力变化就是作用量变化。越大的作用量值，会导致越大的路径变化。

举例：匀速直线运动中，拉格朗日函数中，自变量1（物体在x轴的位置）随时间变化，自变量1的值会匀速增加。自变量2（x轴方向的速度）将会为一个常数不变。假设移动自变量1的一个小球，那么对应自变量2将有两个小球相应作出改变。

1.此时拉格朗日函数在该点（x轴方向）发生的变化量与拉格朗日函数关于x的偏导数相等。作用量的增量也是此变化的量值（上升值）。

2.此时拉格朗日函数在该点（x速率轴方向）发生的变化量却不与拉格朗日函数关于x速率的偏导数相等。因为两个小球一个使拉格朗日值上升，另一个使拉格朗日值等值下降，净变化为0。那么谁来抵消1的增长值呢？就是这两个小球引发的下降值。这个拉格朗日函数下降值正是这个方向偏导数的斜率。作用量的减量也是此变化的量值（下降值）。

3.作用量的增加就是1和2两值绝对值相减.那么作用量的增量趋近于0，则说明路径变化趋近于不变.

当作用量函数在一个很小的稳定段（斜率为0）的时候被扰动，路径改变最小的时候，说明该物体路径被改变到被恢复的波动也最小。这个很小的稳定段，小到某一刻，也就是欧拉-拉格朗日方程成立的时候。自变量1（x轴的位置）的偏导数的增加，将会引起自变量2（x轴方向的速度）的偏导数对于时间的导数（偏斜率的斜率，不是偏导数的原因是因为时间对于三维空间上的三个变量是共同作用的）的减少。当作用量函数在这个很小的稳定段的时候，两者一增一减保持平衡。这个平衡并不意味着作用量函数在此刻的函数值是0.只是斜率为0.

当扰动为0的时候，关于速度的偏导数的导数也为0.可知2式，表明两点间最短曲线为一直线。

以下讨论基础均对于连续函数。可导的条件是函数切线不曾垂直（代表了除数不能为0），但可水平（代表了被除数可以为0）。

欧拉-拉格朗日方程是泛函中非常重要的方程，也是非常经典的能量泛函极小化的方法，不论在物理还是计算机领域，应用非常广泛。所谓能量泛函，是指微分的范数平方再积分。

它的最初的思想来源于微积分中“可导的极值点一定是稳定点（临界点）”。它的**精髓思想**在于：假定当前泛函的解已知，那么这个解必然使得泛函取得最小值（假定是最小值）。换言之，只要在泛函中加入任何扰动，都会使泛函的值变大，所以扰动为0的时候，就是泛函关于扰动的一个极小值。所以当扰动的能量趋近于0，泛函关于扰动的导数也是0。关键是扰动如何表示。答案是扰动用一个很小的数 $\varepsilon$ 乘上一个连续函数。当 $\varepsilon$ 趋近于0,意味着扰动也趋近于0。所以当 $\varepsilon$ 为0的时候，泛函对 $\varepsilon$ 的导数也为0。这就非常巧妙的把对函数求导的问题转化成了一个单因子变量求导的问题。这就是这个思想的伟大之处。

先用引例对方程做一个简单的推导（不是证明）。

函数 $f$ 至少需为一阶可微的函数。若 $f_0$ 是一个局部最小值，而 $f_1$ 是一个在两头端点 $x_1$ 、 $x_2$ 取值为零（积分值为0）并且至少有一阶导数的函数，则可得到以下的式子：

$$A[f_0] \leq A[f_0 + \varepsilon f_1]$$

其中 $\varepsilon$ 为任意接近0的数字。

因此 $A[f_0 + \varepsilon f_1]$ 对 $\varepsilon$ 的导数（ $A$ 的一阶导数）在 $\varepsilon = 0$ 时必为0，将 $A[f_0 + \varepsilon f_1]$ 对 $\varepsilon$ 求导，得到下式：

$$\int_{x_1}^{x_2} \frac{f_0(x)f_1(x)}{\sqrt{1 + [f_0'(x)]^2}} dx = 0 \tag{1}$$

### 3.3 图片展示

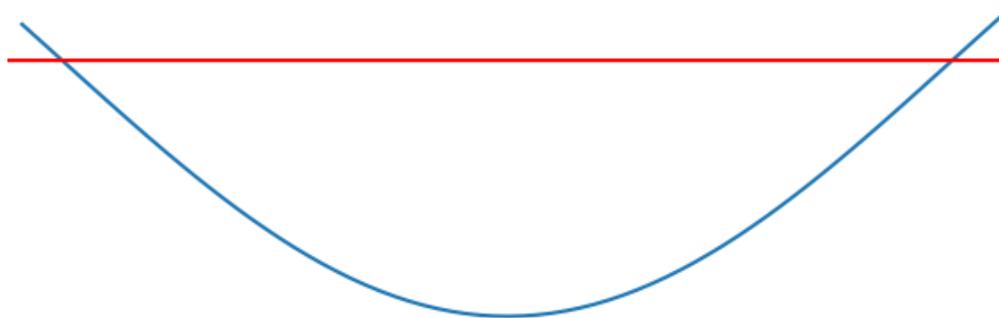


图 1: 欧拉-拉格朗日方程

根据图片1可知aaa**欧拉-拉格朗日方程**aaa

意义在于，以最小值为基础，加上一个放缩系数乘以一个比最小值大的函数值，与这一段弧（最小值到该值的弧微分）做除法（得到切线、导数）。那么这个求导根据罗尔定理应该得0。

首先定义  $A$  集合中的元素（也就是函数了）。 $A$ 集合元素的自变量我们记作  $x$ ，元素记作  $f(x)$ 。我们定义泛函  $\mathcal{F}(f) \in (\mathbb{R} \mapsto \mathbb{R}) \mapsto \mathbb{R}$  .

我们考虑从一种最常见的函数到一个数的运算：定积分运算。因此， $\mathcal{F}$ 的表达式是某种定积分运算，我们表示成对一个与 $f(x)$ 有关的表达式进行定积分运算。这个和 $f(x)$ 有关的表达式，被称作 Lagrangian. 我们记作

$$\mathcal{F}(f) = \int_{x_1}^{x_2} \mathcal{L}(x, f(x), f'(x)) dt$$

我们现在考虑如何计算该定积分的极值。我们假设该定积分取得极值时，自变量的取值为 $\hat{f}$ 。对于  $A$  集合中的其他函数，我们表示成  $f = \hat{f} + \varepsilon \eta$ ，也就是说，我们可以将其他自变量看作是在极值点取得的 $\hat{f}$ 的基础上，加一个扰动函数 $\eta \in \mathbb{R} \mapsto \mathbb{R}$ 。其中 $\varepsilon$ 是个放缩系数。这样，我们的泛函可以写成：

$$\mathcal{F}(f) \int_{x_1}^{x_2} \mathcal{L}(x, f(x), f'(x)) dx = \int_{x_1}^{x_2} \mathcal{L}(x, \hat{f}(x) + \varepsilon\eta(x), \hat{f}'(x) + \varepsilon\eta'(x)) dx$$

这个表达式非常巧妙地将原本不能使用费马引理的情况，变成了可以使用费马引理的情况。当 $\mathcal{F}$ 取得了极值，对扰动的导函数应该等于0。而这里的扰动被刻画成了 $\varepsilon\eta(x)$ ，这里的 $\varepsilon$ 可以看作是扰动的幅度。这样，我们可以将 $\mathcal{F}(f)$ 对 $\varepsilon$ 求导，然后令得到的导函数等于0，就得到了这个泛函的极值条件了。

### 3.4 最短路径问题代入

该式(1)对于任意的满足条件的 $f_1$ 都成立。此条件可视为在可微分函数的空间中， $A[f_0]$ 在各方向的导数均为0。若假设 $f_0$ 二阶可微，则利用分部积分法（关于 $x$ 的两个函数相乘求导）可得：

$$\begin{aligned} & \int_{x_1}^{x_2} \frac{f_0(x)f_1(x)}{\sqrt{1+[f_0'(x)]^2}} dx = 0 \\ \Rightarrow & \int_{x_1}^{x_2} \frac{f_0(x)}{\sqrt{1+[f_0'(x)]^2}} d(f_1(x)) = 0 \\ \text{又} \because & \int u dv = uv - \int v du \\ \Rightarrow & \left[ f_1(x) \frac{f_0(x)}{\sqrt{1+[f_0'(x)]^2}} \right]_{x_1}^{x_2} - \int_{x_1}^{x_2} f_1(x) \frac{d}{dx} \left( \frac{f_0(x)}{\sqrt{1+[f_0'(x)]^2}} \right) dx = 0 \\ \Rightarrow & \int_{x_1}^{x_2} f_1(x) \frac{d}{dx} \left( \frac{f_0(x)}{\sqrt{1+[f_0'(x)]^2}} \right) dx = 0 \end{aligned}$$

扰动需要保证起始点和结束点为0，即 $\eta(x_1) = \eta(x_2) = 0$ ，所以上式左边被减数等于0。

其中 $f_1$ 为在两端点皆为0的任意二阶可微函数。记

$$I = \int_{x_1}^{x_2} f_1(x) H(x) dx = 0$$

$$H(x) = \frac{d}{dx} \left( \frac{f_0(x)}{\sqrt{1+[f_0'(x)]^2}} \right)$$

若存在：使 $H(x) > 0$ ，即在周围有一区间内， $H$ 为正值。

可以选择 $f_1$ 在此区间外为0，在此区间内为非负值，因此 $I > 0$ ，和前提不合。

若存在使 $H(x) < 0$ ，也可证得类似的结果。

因此可得到以下的结论：

$$\frac{d}{dx} \left( \frac{f_0(x)}{\sqrt{1+[f_0'(x)]^2}} \right) = 0$$

由结论可推得下式：

$$\frac{d^2 f_0(x)}{dx^2} = 0 \tag{2}$$

二阶导数为零说明函数的斜率没有变化。如果整个函数的二阶导数都是0，说明整个函数斜率没有变化，则成为直线，如 $y = kx + b$ 。

这表明：两点间最短曲线为一直线。

### 3.5 导数、微分的基本原理

#### 3.5.1 导数的定义

考虑定义在开集 $U \subset \mathbb{R}$ 上的函数 $f : U \rightarrow \mathbb{R}$ 和 $U$ 上的一点 $x_0 \in U$ 。如果极限 $\lim_{h \rightarrow 0} \frac{f(x_0+h) - f(x_0)}{h}$ 存在，我们把这个极限称为 $f$ 在点 $x_0$ 处的导数，拉格朗日记号记为 $f'(x_0)$ 。

斜率：表示一条直线相对于横坐标轴的倾斜程度。

《高数课本》非匀速直线运动的速度和切线的斜率都归结为如下的极限：

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0},$$

这里  $x \rightarrow x_0$  和  $f(x) - f(x_0)$  分别是函数  $y = f(x)$  的自变量的增量  $\Delta x$  和函数（值或者因变量）的增量  $\Delta y$ ：

$$\Delta x = x - x_0$$

$$\Delta y = f(x) - f(x_0) = f(x_0 + \Delta x) - f(x_0)$$

导数：
$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}$$

如果  $f$  在  $U$  上处处可导，定义在  $U$  上的函数： $x \mapsto \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$  称为  $f$  的导数。

导数的根本意义就是切线的斜率，并且不能垂直（天然选择二维空间中表达就要分上下）。也就是说不能两头无穷但可以等于0。微分则是由导数引出，用切线段代替曲线段的一种近似，所以要分清自变量和因变量。

函数值对  $x$  的微分是导数，对  $y$  的微分是1。

自变量发生变化、函数值也发生变化。而微分是，自变量在这一段变化中取很小的一截（比如从起始值向右侧（ $x$ 轴正向）的一段增量），那么函数的导数（也就是斜率）乘以自变量的这一小截，就可以近似的描绘这一小截自变量变化过程导致的函数值的改变量。微分刻画的就是函数值的改变量。而对于因变量来说。微分一定是和因变量同步的。因变量变化多少，函数值铁定变化多少，所以  $\frac{d}{dy}(y) = 1$

### 3.5.2 微分的定义

在各种微积分教材里，「函数的微分」紧接着在「导数」之后出现。这个概念经常使初学者混乱。函数在某点处「可导」与函数在某点处「可微」是一回事，英文都是differentiable。

「函数在某点处的导数(derivative)」与「函数在某点处的微分(differential)」是两个不同的概念，但两者的关系非常简单：那就是某个实数  $a$  与这个实数所定义的  $\mathbb{R}$  上的线性变换  $h \mapsto ah$  的关系。

用花哨点话说就是一维向量空间  $\mathbb{R}^1$  上的元素与其对偶空间  $(\mathbb{R}^1)^*$  上的元素的关系。

如果考虑可导 (differentiable) 函数  $f : U \rightarrow \mathbb{R}$  ( $U$  为实数域里的开集，开集就是只有填色不包含描边) 以及  $\mathbb{R}$  上的一点  $p$ ，那么这个关系就是实数  $f'(p)$  与线性变换  $h \mapsto f'(p)h$  的关系。

如果定义  $dx := h$ ，和  $(df)_p := f'(p)h$ ，那么有  $(df)_p = f'(p)dx$ 。将  $p$  换成  $x$ ，并且省略  $df$  的下标，写成  $df = f'(x)dx$  或者  $dy = f'(x)dx$ ，就是很多微积分或数分教材上写的函数微分。

$y'$  对  $y$  求导不同于  $y'$  对  $x$  求导。 $dx$  这个符号的含义是：在约定  $y = f(x)$  的函数中  $x$  是自变量， $y$  是因变量，那么

$$y \text{ 的微分表示为 } dy = f'(x) \cdot \Delta x, \quad \frac{dy}{dx} = f'(x).$$

$$x \text{ 的微分表示为 } dx = \Delta x, \quad \frac{dx}{dx} = 1$$

$$\text{同理 } dy = \Delta y, \quad \frac{dy}{dy} = 1$$

这是为了保证最内层的变量同样满足一阶微分形式不变性而做出的定义。所以  $dy/dy=1$  就是在考虑  $y = f(y)$  对  $y$  的导数。

但是

$$\frac{d}{dy}(y') = \frac{d}{dy} \left( \frac{d}{dx}(y) \right) = \frac{d}{dx} \left( \frac{d}{dy}(y) \right) = \frac{d}{dx}(1) = \infty$$

上式的意义： $\lim_{h \rightarrow 0} \frac{f(x_0+h) - f(x_0)}{\Delta y} = \lim_{h \rightarrow 0} \frac{\Delta y}{\Delta y} = \lim_{h \rightarrow 0} \frac{1}{h} = \infty$  函数（以数列为理解）发散，极限不存在。

这又和  $y = C$  ( $C$  标识常数)， $\frac{d}{dx}(y) = \frac{d}{dx}(C) = \lim_{h \rightarrow 0} (C) = 0$  完全不同。

### 3.6 最速降线问题代入

一般地，考虑这样的泛函

$$A[f] = \int_{x_1}^{x_2} L(x, f, f') dx$$

形如这种形式的泛函，称为**简单泛函**。其中 $f$ 二阶导连续。在这种情形下，满足**欧拉-拉格朗日方程**（简称E-L方程）：

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

上式就是 $\mathcal{F}(f)$ 取得极值的条件，也即 Euler-Lagrange Equation

这里用E-L方程推导一下最速降线。

质点下落，其速度 $v$ 与它的纵坐标的关系

$$v = \sqrt{2gy}$$

曲线上质点的运动速度

$$v = \frac{ds}{dt} = \sqrt{1 + y'^2} \frac{dx}{dt}$$

$$F(x, y, y') = \sqrt{\frac{1+y'^2}{2gy}} = F(y, y')$$

假定 $y$ 二阶导连续，则它满足E-L方程：

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

虽然E-L方程可以解决大多数泛函极值问题，但是对于满足 $\frac{\partial \mathcal{F}}{\partial x} = 0$ 的泛函，直接代入E-L方程会容易产生比较复杂的微分方程。因此，我们决定再简化一下难度，对E-L方程进行移项

$$\frac{\partial \mathcal{F}}{\partial y} = \frac{d}{dx} \left( \frac{\partial \mathcal{F}}{\partial y'} \right)$$

对两侧乘上 $y'$ 并积分，得到：

$$\begin{aligned} \int \frac{\partial \mathcal{F}}{\partial y} y' dx &= \int y' \frac{d}{dx} \frac{\partial \mathcal{F}}{\partial y'} dx \\ \int \frac{\partial \mathcal{F}}{\partial y} y' dx &= y' \frac{\partial \mathcal{F}}{\partial y'} - \int \frac{\partial \mathcal{F}}{\partial y'} y'' dx \\ \int \left[ \frac{\partial \mathcal{F}}{\partial y} y' dx + \frac{\partial \mathcal{F}}{\partial y'} y'' dx \right] dx &= y' \frac{\partial \mathcal{F}}{\partial y'} + C \\ \int \frac{\partial \mathcal{F}}{\partial x} dx &= y' \frac{\partial \mathcal{F}}{\partial y'} + C \\ \mathcal{F} - y' \frac{\partial \mathcal{F}}{\partial y'} &= C \end{aligned}$$

因此，我们可以在优化满足 $\mathcal{F} = \mathcal{F}(y, y')$ 的泛函使用我们得到的结论，

即贝尔特拉米等式（Beltrami's identity）。

根据一元函数与多元函数复合的情况（拉格朗日函数是一个复合函数 $z = f[\varphi(t), \psi(t)]$ ，对应两个一元函数：位置函数 $u = \varphi(t)$ 、速度函数 $v = \psi(t)$ ）

$$\frac{dz}{dt} = \frac{\partial z}{\partial u} \frac{du}{dt} + \frac{\partial z}{\partial v} \frac{dv}{dt}$$

∴使函数 $F(y, y')$ 对 $x$ ，利用一元函数与多元函数复合公式求导：

$$\begin{aligned} \frac{d}{dx} F(y, y') &= \frac{\partial F}{\partial y} y' + \frac{\partial F}{\partial (y')} (y')' \\ &= \frac{\partial F}{\partial y} y' + \frac{\partial F}{\partial y'} y'' \end{aligned}$$

使函数 $\frac{\partial F}{\partial y'}$ 也对 $x$ ，利用函数积的求导法则

$(uv)' = u'v + uv'$ , 可得

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

假设我们先不知贝尔特拉米等式 (Beltrami's identity), 仅仅有这么个式子突然出现, 则

$$\begin{aligned} & \because \frac{d}{dx} \left[ F(y, y') - y' \frac{\partial F}{\partial y'} \right] \\ &= \frac{\partial F}{\partial y} y' + \frac{\partial F}{\partial y'} y'' - y'' \frac{\partial F}{\partial y'} - y' \frac{d}{dx} \left( \frac{\partial F}{\partial y'} \right) \\ &= \frac{\partial F}{\partial y} y' - y' \frac{d}{dx} \left( \frac{\partial F}{\partial y'} \right) \end{aligned}$$

上式去掉 $y'$ 不就是欧拉-拉格朗日方程么。

$$\text{所以 } \frac{d}{dx} \left[ F(y, y') - y' \frac{\partial F}{\partial y'} \right] = 0$$

$\therefore F(y, y') - y' \frac{\partial F}{\partial y'} = C$ . 欧拉-拉格朗日方程的第二种形式。

代入 $F$ 的表达式, 得

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

化简得

$$y(1+y'^2) = \frac{1}{2gC^2} = 2r$$

不妨设

$$y' = \cot \frac{\theta}{2}$$

代入上式, 得

$$y = 2r \sin^2 \frac{\theta}{2} = r(1 - \cos \theta)$$

再对 $\theta$ 求导

$$y' \frac{dx}{d\theta} = r \sin \theta$$

即

$$\frac{dx}{d\theta} = \frac{r \sin \theta}{y'} = \frac{2r \sin \frac{\theta}{2} \cos \frac{\theta}{2}}{\cot \frac{\theta}{2}} = 2r \sin^2 \frac{\theta}{2} = r(1 - \cos \theta)$$

两边对 $\theta$ 求积分

$$x = r(\theta - \sin \theta)$$

所以, 得到的曲线是

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

这就是摆线方程。所以最速降线就是摆线。(p.s. 最速降线也可以用微分方程的办法解)