欧拉——拉格朗日方程

小圆滚滚

1 预备定理

1.1 定理1：

$\int_a^b M(x)h(x)dx = 0$

$h(a), h(b) = 0$其中h(x)是任意函数，当其在任一点翘起，为了保证积分为0，那么M在该点必为零，若h在每一点翘起，那么M在每一点都为0.（几何证法）

$\Rightarrow M(x) = 0$

代数证法：

令$h(x) = -M(x)(x-a)(x-b)$

$M \cdot h = M^2(-(x-a)(x-b))$

其中，x-a为正，x-b为负，所以括号表达式恒为正（且不等于0）.M^2必为0，M必为0.

推论：

$\int_a^b [M(x)\eta(x) + N(x)\xi(x)]dx = 0$

令$\eta(x) = -M(x)(x-a)(x-b)$

$\xi(x) = -N(x)(x-a)(x-b)$

$(M^2 + N^2)[-(x-a)(x-b)]$

上式中括号恒为正。左侧可以理解成一个向量恒为0，分量都为0。可以推广到两个变数、三个变数、多个变数的情况。

1.2 定理2：

η/'i:tə/ 伊塔

ξ/ksi/ 柯si

ε/'epsılɔn/ 艾普si隆

穿透公式1

2 泛函、变分

我们先来看一下什么是泛函（functional）。

函数$x_0 \mapsto f(x_0)$

x_0是一个函数的参数（x_0 is an argument of a function）。一个数域到另一个数域的满足某种关系f的映射。实变函数(以实数作为自变量的函数叫做实变函数)f。

泛函一般指从一个**函数空间**到**数域**的映射。

对比$f \mapsto f(x_0)$

同理，对于一个实变函数f，上式左侧由一个数域，变成了一个函数域。这就是一个基本的泛函，它把函数空间映射到了实数域。

一般情况下,泛函式常用积分形式表示：

$J[y(x)] = \int_{x_0}^{x_1} F(x, y, y')\mathrm{d}x$

式中，被积函数$F(x, y, y')$称为核。

拉格朗日量、作用量的对应关系：

拉格朗日函数在欧氏空间配合时间的拉格朗日力学（矢量力学）中，经典物理系统，通常定义为动能减去势能。多个参数在对应时间这个自变量下的变化。而作用量是一个标量值，由拉格朗日函数在某时间段的积分计算得出。

作用量是一个泛函，即函数的函数。使用的函数为时间和空间的函数，进而通过积分变为一个标量。在经典力学中，作用量常表现为研究系统的拉格朗日量关于两个节点之间时间的积分。

$\mathcal{S} \stackrel{\text{def}}{=} \int_{t_1}^{t_2} \mathcal{L}(q, \dot{q}, t)dt = \int_{t_1}^{t_2} (\mathcal{T} - \mathcal{U})dt$

其中L为物体系的拉格朗日量，是广义坐标q，其一阶时间导数$\dot{q}$以及时间t的函数。T为动能，U为势能。

我总结5句话：作为拉格朗日函数的基础知识：

1.不定积分导函数可以得到原函数加常数项。

2.定积分导函数可以得到导函数在区间x轴上方图形面积的绝对值减去x轴下方图形面积的绝对值的差。等于原函数在x_2的值减去原函数在x_1的值（牛顿-莱布尼茨公式、微积分基本公式）.

3.所以原函数在导函数为0（也就是斜率为0）的时候，就是一条平行于x轴的直线。

4.原函数在导数为0的点，就是一个小区域内的极值点（费马引理）。

5.两端点值相等，则必存在罗尔中值，使得斜率为0。(罗尔中值定理)。

上述原理推广到泛函：

现在我们来看一看泛函极值问题：

找到一个满足$y(x_0) = y_0, y(x_1) = y_1$的函数y，使得泛函$J[y] = \int_{x_0}^{x_1} F(x, y, y')\mathrm{d}x$取得极值。

在传统的单变量微积分里的优化问题中，我们一般都会通过寻找导数的零点来寻找函数的极值。在泛函里我们会寻找**变分（variation）**(泛函的导数)的零点来求得泛函的极值。如果我们假设$y(x)$就是最终满足泛函达到极值的函数,那条真正的路径，而函数$\tilde{y}(x)$　$(\tilde{y}(x) = y(x) + \eta(x))$表示的是函数y邻域（罗尔定理中两值相等的中间部分就是一个大邻域，其中的任何一段都是小邻域）内的函数.

其中$\eta(x_0) = \eta(x_1) = 0$以便于所有的$\tilde{y}(x)$,

这样就能保证$\tilde{y}(x)$可以通过两个固定端点。

这时我们可以说，$y(x)$所对应的泛函数J的值是泛函数$\tilde{J} = \int_{x_1}^{x_2} F(x, \tilde{y}, \tilde{y}')dx$的极值。我们可以进一步用$\varepsilon$表示$\tilde{J}$(我们简单一些，这里去掉了$\varepsilon$)：

$\delta J \equiv J[y + \eta] - J[y] = \int_{x_0}^{x_1} [F(x, y + \eta, y' + \eta') - F(x, y, y')]\,\mathrm{d}x$

对于足够“微小”的y，我们可以用泛函核F在$F(x, y, y')$附近的一阶泰勒展开(仅取用前两项)来“估计”$F(x, y + \eta, y' + \eta)$：

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

$$
\begin{aligned}
\delta J &= \int_{x_0}^{x_1} \left[F(x,y,y') + \nabla F \cdot (0,\eta,\eta') - F(x,y,y')\right] \mathrm{d}x \\
&= \int_{x_0}^{x_1} \nabla F \cdot (0,\eta,\eta') \mathrm{d}x \\
&= \int_{x_0}^{x_1} \left[\frac{\partial F}{\partial y}\eta + \frac{\partial F}{\partial y'}\eta'\right] \mathrm{d}x
\end{aligned}
$$

上式中最后一步应该有三项，但是第一项是0.现在我们对被积函数里的第二项进行分部积分，得：

$\delta J = \int_{x_0}^{x_1} \frac{\partial F}{\partial y}\eta \mathrm{d}x + \left[\frac{\partial F}{\partial y'}\eta\right]_{x_0}^{x_1} - \int_{x_0}^{x_1} \eta \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y'}\mathrm{d}x$

根据η满足$\eta(x_0) = \eta(x_1) = 0$的性质，我们可以得到泛函J变分的最简形式：

$\delta J = \int_{x_0}^{x_1} \eta \left[\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y'}\right] \mathrm{d}x$

由于在$J[y]$处取得极值，所以根据“直觉”，我们可以得到$\delta J = 0$ 。我们也不难能看出（变分法基本引理），$\delta J = 0$当且仅当被积函数里η的系数在区间内恒为0。因此，我们不严谨地推导出了：

$\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y'} = 0$

即大名鼎鼎的**欧拉—拉格朗日方程（Euler—Lagrange equation）**，简称E-L方程。知乎有很多其它的推导过程，但是它们大多数都设$u(x) = y(x) + \varepsilon\eta(x), \quad (\varepsilon \ll 1)$然后求泛函对$\varepsilon$的全微分。虽然这样更加的严谨，但是作者个人最容易理解的推导还是以上$u = y + \eta$的方法。

2.1 更严谨的证明

第一节（part1）：

我们称F为拉格朗日函数(Lagrange function)。

F可以是函数$y(x)$和$y(x)$各阶导数的函数(以下$y(x)$均简写成y)。

为了说明方便，我们先姑且设F是y和y'（part1只求到一阶导，高阶请看part2）的函数，所以我们可以进一步将泛函数I写成：

$I(y) = \int_{x1}^{x2} F(y, y'; x)dx$

积分里面我用分号;将x和前面的y隔开代表y和y'是x的函数。一般F和y的函数关系是已知的，所以想要最大或最小化泛函数，实际上是通过选择适当的函数$y(x)$。

为了透彻理解这个概念，我们可以来看一个简单的例子。

Example 1.

一个最简单直观的例子是求两个固定点之间的最短路径。当然大家都知道两点之间直线最短，这里可以用变分法做出解释。

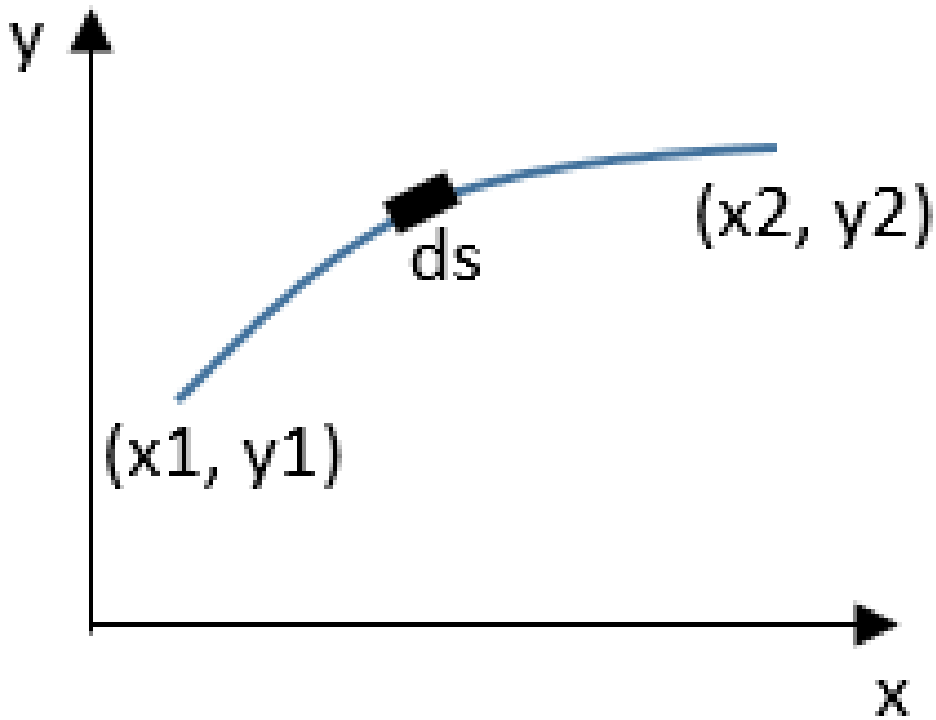

图 1: 两点之间直线最短

如上图所示路径是一任意路径，我们取区中一小段微元ds，可以容易计算微元断的长度为：
$ds^2 = dx^2 + dy^2 = [1 + (y')^2]dx^2$,即：
$ds = \sqrt{1 + (y')^2}dx$
积分得到总的路径长度为：
$L = \int_{x1}^{x2} ds = \int_{x1}^{x2} \sqrt{1 + (y')^2}dx$
这个例子中，L是泛函数，$\sqrt{1 + (y')^2}$是拉格朗日函数F.

这个拉格朗日函数就是核。对积分求导，就是原函数的导函数，也就是现函数。对泛函求ϵ的导数，等式左右却不是同一个自变量，等是右侧自变量是x，所以不会得现函数，也就不是核这个函数，但是因为这两个自变量不相关。微分符号可以穿透积分得到偏分。这就给研究拉格朗日函数，找到它的形状，提供了条件线索，就可以通过这些条件得到一些物理上的真实意义。同时拉格朗日函数又是函数y和y各阶导数的函数。所以就能得到y这个函数的一些特征。帮助我们确定y（或者说是f(x)）。

根据牛顿-莱布尼茨公式，泛函在$[x_1, x_2]$上的值一定的情况下，也就是面积如图2一定的情况下，取值取决于自变量y的形状。泛函的近似值，就是y形状的微变化，并且y在$[x_1, x_2]$上取值相同。泛函的一阶变分，就是泛函在附近点的泰勒展开第二项，（第一项是泛函本身）。一阶变分体现了近似中变化最大的那部分，这个一阶变分也是一个泛函并且是线性的。并且一阶接近说明曲线和原曲线弧平行，不至于抖如筛糠图片3。令一阶变分为0，相当于函数（这里函数就是普通函数，不是泛函）的切线斜率为0.那么泛函将取到极值。在泛函为极值的时候，我们就求得了物理上的意义。比如两点之间弧微分的那个y就是直线，因为结论是y’=0。

从数学的角度有历史沿革：由费马引理到罗尔中值定理，到拉格朗日中值定理，到柯西中值定理，由柯西中值定理到泰勒公式，由柯西中值定理到洛必达法则。推导出柯西。但是倒过来看，泰勒展开

真正解释了对于自变量x因变量的微分是一种一阶导数的近似。所以对于泛函，一阶变分的近似如果为0.也就找到了那条唯一的自变量$f(x)$的那条路。

我们想要找一个函数$y(x)$使得泛函数L的值最小。

泛函数的极值

这里重申下，泛函数I在区间$[x_1, x_2]$上的值取决于积分路径的选择，即取决于函数$y(x)$的选择。我们有理由假设存在一个这样的$y(x)$（这里的y可以指泛函对于自变量ϵ的函数映射，也可以指泛函的自变量$f(x)$本身即$y = f(x)$（后者具备物理意义），虽然$\mathcal{F}(y, y'; x)$还可以包含y的更高阶、更多元，但我们不考虑拉格朗日函数如何消弭这些影响，因为我们不知道系统上拉格朗日函数的具体拼接，只能看到拉格朗日函数产生的值，最终，在自变量$f(x)$选取一个形状时，拉格朗日函数会想办法将泛函的值变成了同一个数值。也就是说泛函的函数值虽然相等，但是自变量$f(x)$却没有确定，扭曲着前进，整个系统肯定是不稳定的。也不是唯一的。现实中封闭系统一定是稳定的，所以过程中的变化路径是一定的。我们要确定这个唯一的路径。那么泛函对于ϵ一定是确定的、$f(x)$一定也是确定的。），可以使得泛函数I取到极值。而在这个$y(x)$附近的任意路径我们记做$\tilde{y}(x)$。$\tilde{y}(x)$是一个研究集合的代表，它的值扫过所有这些集合，一定会开始越来越小，然后又越来越大，真正的那个真相路径就在极值最小的时候出现。雅各布·伯努利的变分法另外，我们假设$y(x)$两阶可微。通过引入一个微小量$\epsilon \ll 1$对于x的函数来讲ϵ就是常数，所以又可以称作放缩系数和一个任意可微函数$\eta(x)$扰动函数（比如sin(x)），我们可以用$y(x)$表示$\tilde{y}(x)$:

$\tilde{y}(x) = y(x) + \epsilon\eta(x)$

如何定义$\eta(x)$？

$\tilde{y}(x) - y(x) = D(x)$

$\epsilon\frac{D(x)}{\epsilon} = \epsilon\eta(x)$ $D(x)$可以吸收ϵ（如果用多项式来代表D(x)，那么多项式的每一项的系数除以这个ϵ常数，得到一个函数，对“$\eta(x)$是一个任意函数”这句话影响不大），对于x，ϵ是一个常数。对于ϵ，x是一个常数。也就是说它们两个没有任何关系。而$\eta(x)$是一个任意函数（比如sin(x)图4）。ϵ是放缩系数。变分思想就是人为的规定一个任意扰动函数，人为的规定一个放缩系数。

这样做的好处是对于一个给定的$\eta(x)$，我们可以通过改变ϵ的值来得到无穷多的路径（扫过），同时对于任何$\eta(x)$，当$\epsilon = 0$的时候，$\tilde{y}(x)$和$y(x)$重合。$\eta(x)$是一个任意函数的意义，就是他随意（但不是0），但是我知道这个扰动很小。任何数M与它相乘$M \cdot \eta$等于0.就可以把它看成$\eta = -M \cdot (x-a)(x-b)$，其中$x \in (a, b)$.可得$M \cdot \eta = M^2(-((x-a)(x-b)))$.因为$-((x-a)(x-b))$恒为正，$M^2$恒为正，所以M必为0.

图像直观表示如下图：

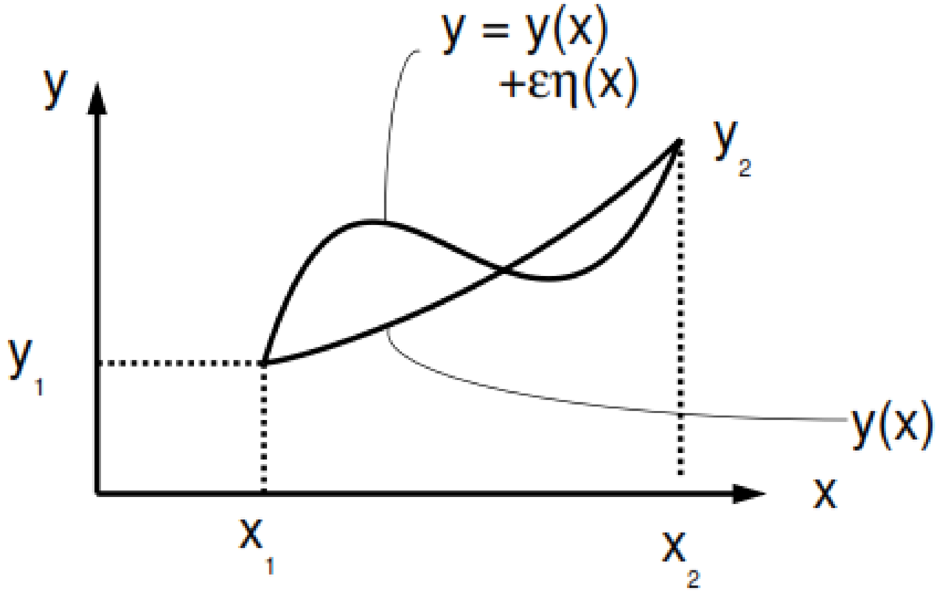

图 2: 选择不同的$y(x)$

数学抽象思维，不管你的数值有多大，只要你相减就不会丢掉公因数。不管你的函数图形有多抖动，只要取足够小的自变量间隔，你要么是凸要么是凹，只有这两种形状。参考拓扑、同胚的一些视频。感受一下这种在自己头脑中任意放大和缩小以及变性的四维力量

两个函数相等，我们判断的准则是x对应取值y，处处相等，两个泛函相等我们如何判断？想象在一条简单平滑的函数$y = x^2$作为自变量$f(x)$，对应的泛函关系$\mathcal{L}$是$y = x+2$，得到一个向上平移的泛函$\mathcal{I}$的图像4。在y轴右侧，在其基础上取两端点，在期间补充$y = sinx$，我们规定叠加的泛函y=(np.sin(x) + (x*0.1)**2) + 2是被干扰之后的泛函$\tilde{\mathcal{I}}$，就会出现图4零阶接近的情况，像藤蔓缠绕一段树枝。

在很多处x值与对应y值都相等。但是中间出现了很多意外的不相等。所以两个泛函不等。如果让它们相等，那么精准判断就要选取一小段，这一小段，两个端点值相等（x,y都相等）。那么这一段，要么是凹、要么是凸。在这个鼓包段研究。两个图形都很平滑。并且总有一条更平滑。当你将其中一条看成是平滑的，那么另一条的变化无非是鼓包偏左还是鼓包偏右。这个问题，在本段不会影响判断的差值，但是我们判断泛函相等是可以任意取段的。当滑动这个取段，还能处处相等时。那么两个泛函一定是同一条。好了，刚才说到取到的一小段，对其积分。可以得到一个图形函数与x轴所夹的面积值，如果这个积分相等，那么由大到小可以确定这一小段泛函图形$\mathcal{I}$就是同一条；拉格朗日函数$\mathcal{L}(y, y'; x)$也是同一条；自变量$y = f(x)_{-}/\text{fi}\Gamma a$。如果处处（任意小段）相等，那么两个泛函相等。这个处处就是极值。如果不是处处，那么在某些区域虽然$\mathcal{I}$的值相等，积分值（面积）相等，但是x对应的f(x)（两端点值）也不相等，中间的图形也不相等。

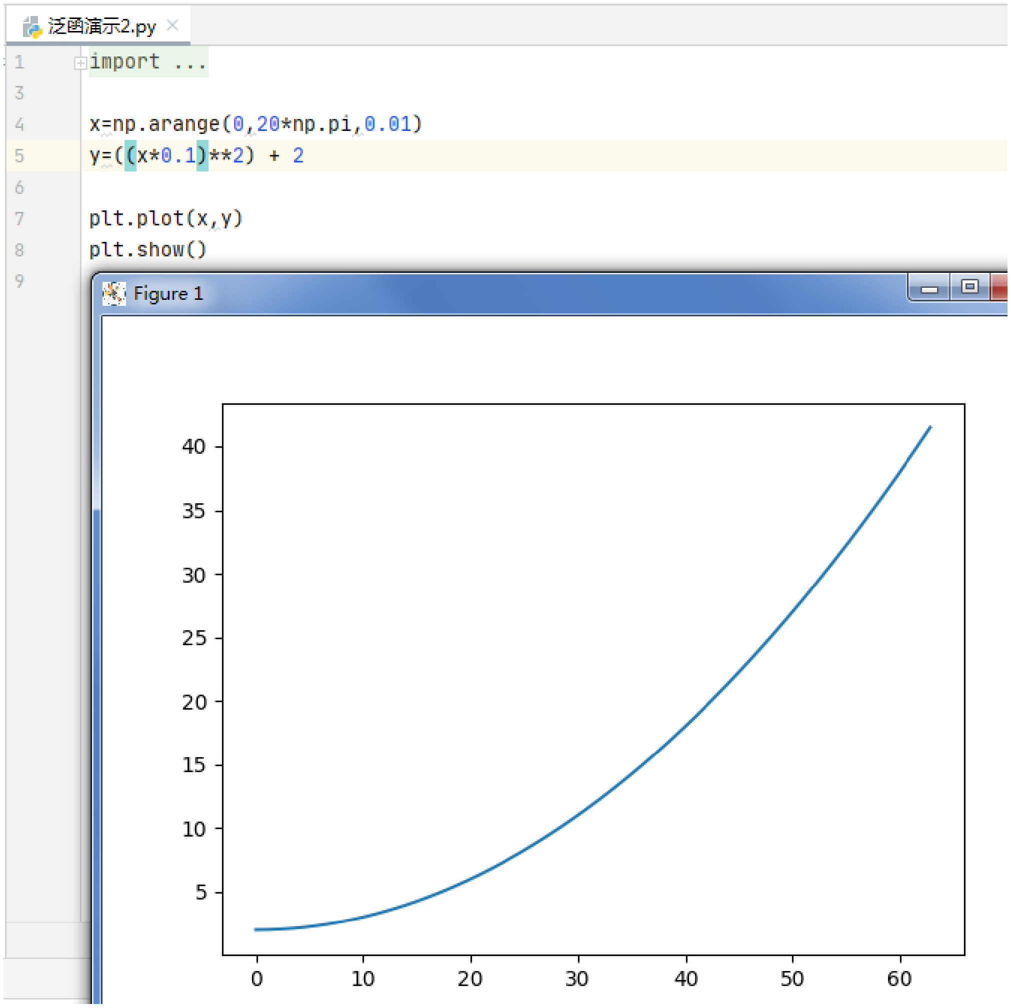
泛函演示2.py
import ...
x=np.arange(0,20*np.pi,0.01)
y=((x*0.1)**2) + 2
plt.plot(x,y)
plt.show()
Figure 1
40
35
30
25
20
15
10
5
0
10
20
30
40
50
60

图 3: 泛函演示1

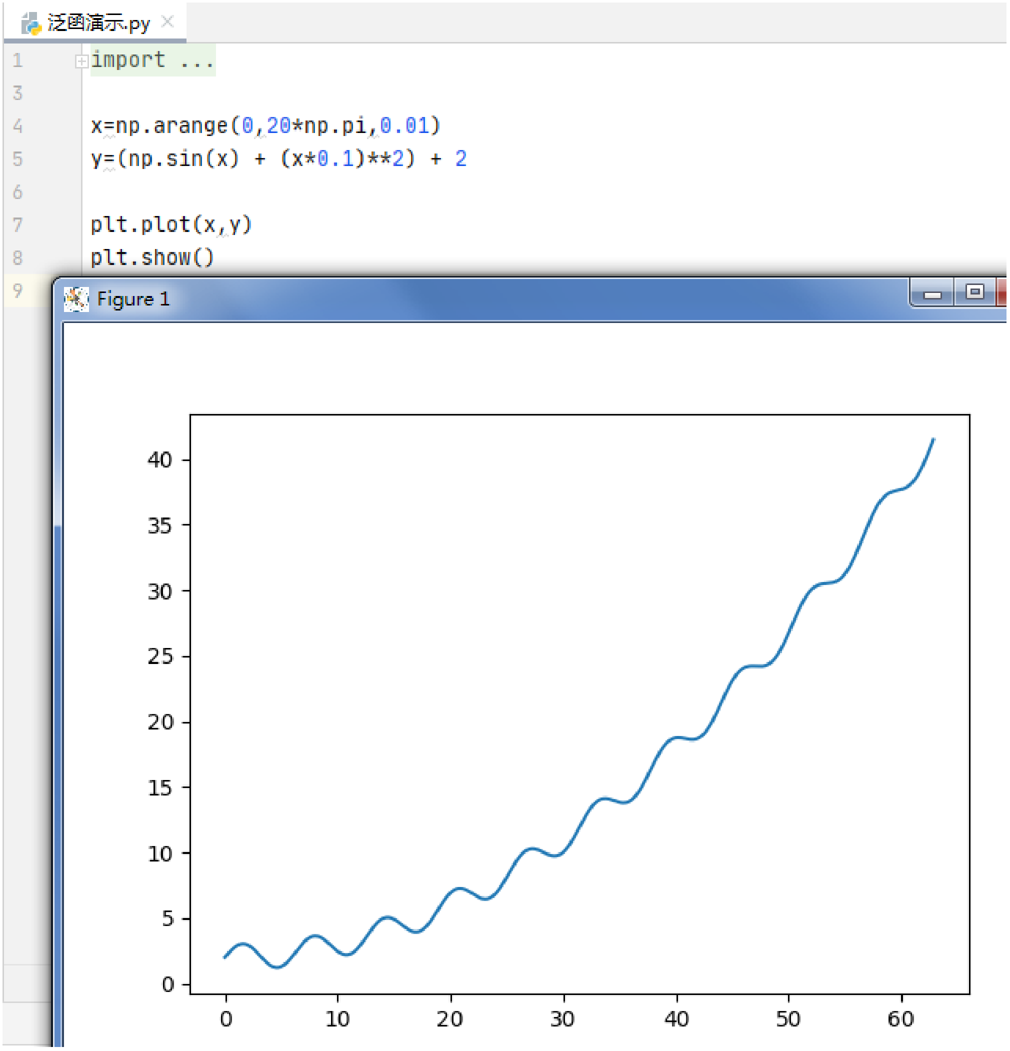

图 4: 泛函演示2

定义3 曲线$x(t)$和$x_1(t)$　k阶接近，若差的绝对值$|x^{(i)}(t) - x_1^{(i)}(t)|, i = \overline{1, k}$极小。这里$x^{(i)}$——为函数$x(t)$的i阶导数。

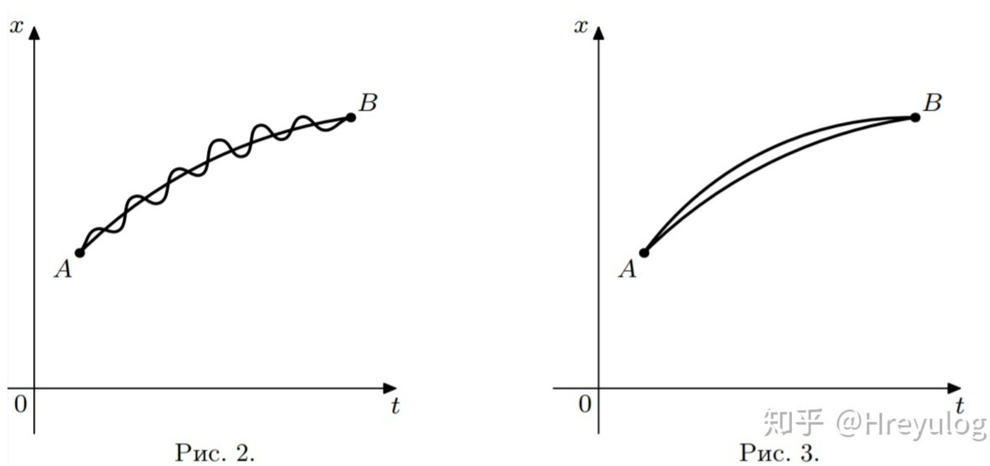

图 5: 零阶接近和一阶接近

图2中的曲线零阶接近，但是一阶不接近。而图3的曲线一阶接近。若曲线k阶接近，则在k以下的任意阶也接近。

由于在边界条件的限制，$\eta(x_0) = \eta(x_1) = 0$。这样就能保证$\tilde{y}(x)$可以通过两个固定端点。

这时我们可以说，$y(x)$所对应的泛函数I的值是泛函数$\tilde{I} = \int_{x_1}^{x_2} F(\tilde{y}, \tilde{y}'; x)dx$的极值。我们可以进一步用$\epsilon$表示$\tilde{I}$：

$\tilde{I} = \int_{x_1}^{x_2} F(\tilde{y}, \tilde{y}'; x)dx = \int_{x1}^{x_2} F(y + \epsilon\eta, y' + \epsilon\eta'; x)dx$

虽然$y(x)$未知，但是根据之前的合理假设，$y(x)$是一个存在的确定函数。所以根据上式，如果给定一个特定的$\eta(x)$，$\tilde{I}$的变化只取决于ϵ的变化。所以我们现在可以把$\tilde{I}$看做是ϵ的函数。用泰勒展开公式将$\tilde{I}$在$\epsilon = 0$处展开得到：

$\tilde{I}(\epsilon) = \tilde{I}|_{\epsilon=0} + (\frac{d\tilde{I}}{d\epsilon})\Big|_{\epsilon=0} \cdot \epsilon + (\frac{d^2\tilde{I}}{d\epsilon^2})\Big|_{\epsilon=0} \cdot \frac{\epsilon^2}{2!} + \cdots = \tilde{I}_0 + \tilde{I}_1\epsilon + \tilde{I}_2\epsilon^2 + \cdots$

很明显，当$\epsilon = 0$时，$\tilde{I}|_{\epsilon=0} = I$,带入上式可得到：

$\tilde{I} - I = \tilde{I}_1\epsilon + \tilde{I}_2\epsilon^2 + \cdots$

这里我们记$\delta I = \tilde{I}_1\epsilon = \frac{d\tilde{I}}{d\epsilon}\Big|_{\epsilon=0} \cdot \epsilon$，并称之为一阶变分。同理二阶变分为$\delta I^2 = \tilde{I}_2\epsilon^2$。

（这里插一句变分和微分的区别。变分在上图的直观解释是$\tilde{y}$和y在竖直方向上的距离，称之为δy，所以这个差是在同一个x上计算的。而微分则是由于x的微小变动引起的y的变动。）变分是在众多函数中挑出那根正确的唯一的线，所以自变量周边的影分身就是干扰项，是虚像。微分是自变量左右邻域运动引起的变化率、变化趋势。

再补充一个本质：从泰勒展开式看导数，在x($x = x_0 + \Delta x$)点上的真实值(x) = 在x_0点上的真实值 + 在x_0点上的导数$\times \Delta x$ + $\frac{\text{在}x_0\text{点上的二阶导数}}{2!} \times \Delta x^2$ + $\frac{\text{在}x_0\text{点上的三阶导数}}{3!} \times \Delta x^3 + \cdots$ ，当函数幂次为2时，等式右侧只有两项。第二项像不像一阶变分？

然后我们可以类比求函数极值时的做法。求函数极值时，我们会令函数的一阶导数为零。这里同样，为了求泛函数$\tilde{I}$的极值，我们令一阶变分$\delta I = 0$。现在我们计算化简δI:

$\delta I = (\int_{x_1}^{x_2} \frac{d\tilde{F}}{d\epsilon}\Big|_{\epsilon=0} dx) \cdot \epsilon$，注意，等式左边是个微分，因为右边$\epsilon$是很小的值，也就是微分的意思

$\frac{d\tilde{F}}{d\epsilon}\Big|_{\epsilon=0} \cdot \epsilon = (\frac{\partial\tilde{F}}{\partial\tilde{y}} \cdot \frac{d\tilde{y}}{d\epsilon} + \frac{\partial\tilde{F}}{\partial\tilde{y'}} \cdot \frac{d\tilde{y'}}{d\epsilon})\Big|_{\epsilon=0} \cdot \epsilon$

因为$\tilde{y}(x) = y(x) + \epsilon\eta(x)$, 不难得到：$\frac{d\tilde{y}}{d\epsilon} = \eta, \frac{d\tilde{y'}}{d\epsilon} = \eta'$,另外我们有$\delta y = \epsilon\eta, \delta y' = \epsilon\eta'$

又因为当$\epsilon \to 0$时，$\tilde{F} \to F, \tilde{y} \to y, \tilde{y}' \to y'$,将这些式子带入原式可以得到：

$\delta I = \int_{x_1}^{x_2} (\frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial y'}\delta y')dx$，这里$\delta y = \partial y - y$，其中蕴含着$\epsilon$，因为$\delta y$的得来正与$\epsilon$对应。所以成功的将一个泛函(函数是I，自变量$\epsilon$)的微分，变成了真实数（在$x$点）与基数（在$x_0$点）的差值。进而变成了另一个自变量函数（拉格朗日函数F，自变量x）的积分的值。

真实值 = 基值 + 导数×增量（一阶变分、微分）

微分 = 导数×增量

真实值 = 基值 + 导数 × 增量 + $\frac{\text{二阶导}}{2!}$增量^2 + $\frac{\text{三阶导}}{3!}$增量$^3 + \cdots$

若$I(\epsilon) = \int_{x_1}^{x_2} F(x, y + \epsilon\eta, y' + \epsilon\eta')dx = \int_{x_1}^{x_2} F(x, u, v)dx$

上式是关于ϵ的微分，用此微分除以它的自变量ϵ，就得到关于ϵ的导数。

根据积分符号下的微分法则(Differentiation of a definite integral. 《Advanced Calculus Frederick S. Woods》)：积分的上下限x和ϵ没有关系，微分符号$d\epsilon$会直接穿透积分符号，变成偏微分的符号$\frac{\partial F}{\partial \epsilon}$

$$\frac{dI}{d\epsilon} = \int_{x_1}^{x_2} \frac{\partial F}{\partial \epsilon} dx = \int_{x_1}^{x_2} (\frac{\partial F}{\partial x}\frac{\partial x}{\partial \epsilon} + \frac{\partial F}{\partial u}\frac{\partial u}{\partial \epsilon} + \frac{\partial F}{\partial v}\frac{\partial v}{\partial \epsilon})dx \tag{1}$$

$\frac{dI}{d\epsilon}\Big|_{\epsilon} = 0$，意思是当$\epsilon = 0$时，等式得0.

因为x不是ϵ的函数，x、ϵ一点关系都没有。所以$\frac{\partial F}{\partial x}\frac{\partial x}{\partial \epsilon}$项等于0

后两项举例：$u = y + \epsilon\eta$，则$\frac{\partial u}{\partial \epsilon} = \eta$，同理有$\eta'$

所以说$\frac{dI}{d\epsilon} = \int_{x_1}^{x_2}(\frac{\partial F}{\partial y}\eta + \frac{\partial F}{\partial y'}\frac{d\eta}{dx})dx$，注意，化成这个式子就是为了提醒等式左边这个导数的自变量是ϵ但是等式右边的值却是对x的积分。最终奥义。对I这个泛函求导，和拉格朗日函数参与的积分只是数值的关系，并不是说对拉格朗日函数求导。

终于到最后一步啦，分部积分一下得到：

$\delta I = \int_{x_1}^{x_2}(\frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial y'}\delta y')dx = \int_{x_1}^{x_2}\frac{\partial F}{\partial y}\delta y dx + \frac{\partial F}{\partial y'}\delta y\Big|_{x_1}^{x_2} - \int_{x_1}^{x_2}\frac{d}{dx}(\frac{\partial F}{\partial y'})dx\delta y = \int_{x_1}^{x_2}(\frac{\partial F}{\partial y} - \frac{d}{dx}(\frac{\partial F}{\partial y'}))\delta y dx + \frac{\partial F}{\partial y'}\delta y\Big|_{x_1}^{x_2}$

令$\delta I = 0$就可以解得最小化泛函数的y啦。我们注意到δI有两个部分。对于第一个积分部分，由于δy是任意的，所以要想使这个部分等于零，需要保证[1]：

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

这就是传说中的欧拉-拉格朗日方程(E-L equation)。

而第二部分等于零则是边界条件。

注[1]:

假设x_1和x_2是给定的常数，$\phi(x)$是一个特定的在$x_1 \leq x \leq x_2$上连续的函数，那么如果对于任意连续可微的函数$\eta(x)$都成立$\int_{x_1}^{x_2}\phi(x)\eta(x)dx = 0$，则$\phi(x) = 0$ $(x_1 \leq x \leq x_2)$。

（任意函数和一个非零的特定函数的乘积仍是任意函数，由于无法保证任意函数的积分是零，所以这个特定函数必须在这个区间上恒等于零使得乘积为零，这样可以保证积分为零。）

第二节（part2）：

上节最后我们得到了泛函数的一阶变分：

$\delta I = \int_{x_1}^{x_2}(\frac{\partial F}{\partial y} - \frac{d}{dx}(\frac{\partial F}{\partial y'}))\delta y dx + \frac{\partial F}{\partial y'}\delta y\Big|_{x_1}^{x_2}$

从上式我们又可以得到欧拉-拉格朗日方程(E-L equation):

$\frac{\partial F}{\partial y} - \frac{d}{dx}(\frac{\partial F}{\partial y'}) = 0, \quad x_1 \leq x \leq x_2$

以及边界条件：

$\frac{\partial F}{\partial y'}\delta y\Big|_{x_1}^{x_2} = 0$

上式左边是由两项相乘得到，一项是$\frac{\partial F}{\partial y'}$，另一项是$\delta y$。很明显如果这两项的乘积为零，那么不是第一项为零就是第二项为零（或者都为零）。如果在边界x_1或x_2处$\frac{\partial F}{\partial y'}$等于零，那么我们称在$x_1$或$x_2$处满足**Natural Boundary Condition**。对应的，如果在边界x_1或x_2处δy等于零，我们就称在x_1或x_2处满足**Essential Boundary condition**。$\delta y = 0$是一个比较模糊的概念，不好判断，所以我们把它等价转化一下。如果y在某一点处的变分为零，说明它在这点的值是确定不变的。换句话说如果y在边界上是确定的值(specified)，那么我们称这个是essential boundary condition。

Solution of Example 1

现在我们终于可以解Part 1.中提出的两点之间直线最短的问题啦。我们回忆下两点之间的路径长度泛函数是：$L = \int_{x1}^{x2}\sqrt{1 + (y')^2}dx$。

$$\frac{dF}{dx}(x, y, y') = \frac{\partial F}{\partial x}\frac{dx}{dx} + \frac{\partial F}{\partial y}\frac{dy}{dx} + \frac{\partial F}{\partial y'}\frac{dy'}{dx}$$

$\frac{dF}{dx}(x, y, y') = \frac{\partial F}{\partial x} \cdot 1 + \frac{\partial F}{\partial y} \cdot y' + \frac{\partial F}{\partial y'} \cdot y'' = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}y' + \frac{\partial F}{\partial y'}y''$

$\because \frac{d}{dx}(y'\frac{\partial F}{\partial y'}) = y''\frac{\partial F}{\partial y'} + y'\frac{d}{dx}(\frac{\partial F}{\partial y'}) = (\frac{dF}{dx} - \frac{\partial F}{\partial x} - \frac{\partial F}{\partial y}y') + y'\frac{d}{dx}\frac{\partial F}{\partial y'}$

上式第二项是分部积分，第三项由公式??得出。

$\therefore \frac{d}{dx}(y'\frac{\partial F}{\partial y'}) = \frac{d}{dx}F - \frac{\partial F}{\partial x} - y'[\frac{\partial F}{\partial y} - \frac{d}{dx}\frac{\partial F}{\partial y'}] = \frac{d}{dx}F - \frac{\partial F}{\partial x} - y' \cdot 0$

上式最后一项是欧拉-拉格朗日第一种形式的结论（为0）

$\Rightarrow \frac{\partial F}{\partial x} - \frac{d}{dx}(F - y'\frac{\partial F}{\partial y'}) = 0$

当$F = \sqrt{1+(y')^2}$时，第二种形式特别有用。即$\frac{\partial F}{\partial x} = 0$(不显含有x，所以相当于对常数求导，等于0).

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

带入欧拉-拉格朗日方程(E-L equation)：$\frac{\partial F}{\partial y} = 0$, $\frac{\partial F}{\partial y'} = \frac{y'}{\sqrt{1+(y')^2}}$（复合函数求导$u = 1 + (y')^2, f = \sqrt{u}$）

E-L: $0 - \frac{d}{dx}(\frac{y'}{\sqrt{1+(y')^2}}) = 0, \quad x \in (x_1, x_2)$

两个边界条件(BC)：$y(x_1) = y_1 \ \ \& \ \ y(x_2) = y_2$,所以在两个边界点上都是Essential Boundary Condition。

然后解一下E-L方程（见青色字部分），可以得出：$\frac{y'}{\sqrt{1+(y')^2}} = constant = C$。进一步化简可以得出：

$(y')^2 = \frac{C^2}{(1-C^2)}$

上式表示y'是常数，导数是常数的曲线是什么？是直线。加上边界条件，两点之间路径最短的曲线就是直线了。这里我们用变分法科学得解释了一个非常直观的问题。

当然上述问题中，F只是y和y'的函数，但是F完全还可能是y"甚至更高阶导数的函数。这时候我们用上节所说的设$\tilde{y}(x) = y(x) + \epsilon\eta(x)$的方法就有点麻烦了（但是也可行）。这里我们提供一个更高效的方法，直接对泛函数求变分。

假如泛函数$I = \int_{x_1}^{x_2} F(y, y', y''; x)dx$,那么它的一阶变分为：

$\delta I = \int_{x_1}^{x_2}(\frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial y'}\delta y' + \frac{\partial F}{\partial y''}\delta y'')dx$

然后再进行分部积分得到：

$\delta I = \int_{x_1}^{x_2}(\frac{\partial F}{\partial y} - \frac{d}{dx}(\frac{\partial F}{\partial y'}) + \frac{d^2}{dx^2}\frac{\partial F}{\partial y''})\delta y dx + \frac{\partial F}{\partial y''}\delta y'\Big|_{x_1}^{x_2} + [\frac{\partial F}{\partial y'} - \frac{d}{dx}(\frac{\partial F}{\partial y''})]\delta y\Big|_{x_1}^{x_2} = 0$

同样我们可以得到欧拉-拉格朗日方程：

E-L: $\frac{\partial F}{\partial y} - \frac{d}{dx}(\frac{\partial F}{\partial y'}) + \frac{d^2}{dx^2}\frac{\partial F}{\partial y''} = 0$

以及边界条件：

@x_1 and x_2: $\frac{\partial F}{\partial y''} = 0$(Natural) or $\delta y' = 0$(Essential) ; $\frac{\partial F}{\partial y'} - \frac{d}{dx}(\frac{\partial F}{\partial y''}) = 0$(Natural) or $\delta y = 0$(Essential)

带约束条件的泛函数极值问题（拉格朗日乘数）

现在我们来讨论下如果我们需要求在满足一定条件下求泛函数的极值，应该如何求解。举个简单的例子，如果给定一个区域的边界长度，我们想最大化该边界所围的区域。这类问题我们需要引进拉格朗日乘数(λ)。

假设约束条件为：$J(y) = \int_{x1}^{x_2} G(y, y'; x) = constant$

那我们引入拉格朗日乘数在原泛函数的基础上构建新的泛函数：

$I^* = I + \lambda J = \int_{x_1}^{x_2}(F + \lambda G)dx = \int_{x_1}^{x_2} F^*(y, y'; x)dx$

然后我们就可以转化成之前的问题来解了。还是用刚才提到的最大化面积问题为例来阐述下这个过程吧：

面积：$A = \int_{x_1}^{x_2} y dx$

边界长度约束：$S = \int_{x_1}^{x_2} \sqrt{1+(y')^2}dx = constant$

则：$I^* = \int_{x_1}^{x_2}[y + \lambda\sqrt{1+(y')^2}]dx$

令一阶变分为零($\delta I^* = 0$):
$\int_{x_1}^{x_2}[\delta y + \lambda \frac{y'}{\sqrt{1+(y')^2}}\delta y']dx = 0$
依旧老朋友分部积分：
$\delta I^* = \int_{x_1}^{x_2}[1 - \lambda \frac{d}{dx}(\frac{y'}{\sqrt{1+(y')^2}})]\delta y dx + B.C.term = 0$
这里图方便我就将边界条件项用B.C.term代替啦。然后欧拉-拉格朗日方程为：
E-L：$1 - \lambda \frac{d}{dx}(\frac{y'}{\sqrt{1+(y')^2}}) = 0 \quad x \in (x_1, x_2)$
现在我们可以解一下E-L方程：
$\frac{d}{dx}(\frac{y'}{\sqrt{1+(y')^2}}) = \frac{1}{\lambda}$
$\Rightarrow \frac{y''}{(1+(y')^2)^{3/2}} = \frac{1}{\lambda}$
也就是说边界是曲率恒为$\frac{1}{\lambda}$的曲线，即半径为λ的圆，这也和我们日常的常识符合一致。即给定一段绳子，能围起来的最大面积将会是圆。

现在引用圆的参数方程：

$$x - x_0 = \lambda \text{cos} t$$

$$y - y_0 = \lambda \text{sin} t \quad t \in [0, 2\pi]$$

带入原约束方程：
$S = \int_{x_1}^{x_2} \sqrt{1+(y')^2}dx = \int_0^{2\pi} \sqrt{(dx/dt)^2 + (dy/dt)^2}dt = 2\pi\lambda$
可以解得半径$\lambda = S/2\pi$

3 欧拉-拉格朗日方程

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：

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

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）**。我们已经将所有需要的工具准备完毕，现在可以开始进入最速降线问题的求解了。

4 用变分法求解最速降线问题

由于我们要找到一个曲线，使得物体沿曲线路径滑落的时间最短，我们需要设计一个描述时间的泛函：

$T = \int_C \frac{\mathrm{d}s}{v}$

利用弧微分公式和前面提到的能量守恒，我们不难得出：

$$\mathrm{d}s = \sqrt{1+y'^2}\mathrm{d}x$$
$$v = \sqrt{2g(h-y)}$$

因此，我们的得到了这样的时间泛函：

$T = \int_{x_0}^{x_1} \underbrace{\frac{\sqrt{1+y'^2}}{\sqrt{2g(h-y)}}}_{F} \mathrm{d}x$

因此，我们的最速降线问题变成变分问题，即找到一个满足边界条件的y，使得泛函T取得极值。

由于泛函T的核F满足$\frac{\partial F}{\partial x} = 0$，

所以我们可以使用更加方便的贝尔特拉米等式来求解。首先，我们求F关于y'的偏导数：

$\frac{\partial F}{\partial y'} = \frac{y'}{\sqrt{2g(h-y)}\sqrt{1+y'^2}}$

代入到贝尔特拉米等式，得到：

$$F - y'\frac{\partial F}{\partial y'} = C$$
$$\frac{\sqrt{1-y'^2}}{\sqrt{2g(h-y)}} - \frac{y'^2}{\sqrt{2g(h-y}\sqrt{1-y'^2}} = C$$
$$\frac{1-y'^2-y'^2}{\sqrt{2g(h-y)}\sqrt{1-y'^2}} = C$$
$$\frac{1}{\sqrt{2g(h-y)}\sqrt{1-y'^2}} = C$$

此时，我们如果令$y' \equiv \frac{\mathrm{d}y}{\mathrm{d}x} = \cot\alpha$，则根据三角恒等式$\csc^2\alpha = 1 + \cot^2\alpha$，原式变成：

$\frac{\sin\alpha}{\sqrt{2g(h-y)}} = C$

我们可以发现这个方程和我们用斯涅尔定律得出的方程完全一样，所以其实伯努利和欧拉的方法都能够说明最速降线问题的答案是摆线。然而，摆线的故事才刚刚开始。摆线不仅是最速降线问题的答案，还是**等时降落问题（Tautochrone problem）**的解。