

# Methods of Mathematical Physics

## §5 Calculus of Variations

Lecturer: 黄志琦

<http://zhiqihuang.top/mmp>

# Outline

Fermat's Principle

Calculus of Variations

Homework

Appendix: Hamiltonian

Appendix: Extensions

# 费马原理

光的路径使运动时间取极值。

# 几何光学的费马原理

在光的路径上任取两点A, B。在所有连接A, B的假想路径中, 光的真实路径一定是使运动时间

$$\int_A^B dt$$

取极值的路径。

# 思考题



用费马原理证明光在真空（或同一介质）中沿直线传播

# 思考题



用费马原理证明光的反射定律

# 思考题



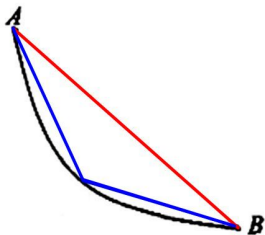
用费马原理证明光的折射定律

# 变分法

$$\frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial y'} = 0$$



# 思考题



如图，从A滑落到B，有直线（红）、折线(蓝)、凹线（黑）三种不同的滑梯设计方案。从哪种滑梯上滑下来的时间更短？

# 最速降线问题(Brachistochrone Problem)

在两个固定点之间造什么形状的滑梯，能使质点滑下来所需时间最短？

## 问题的数学描述

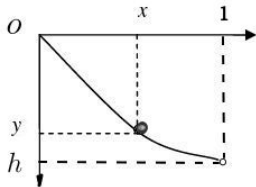
不妨设AB间水平距离为1，垂直距离为 $h$ ，重力加速度为 $g$ 。  
以A点为原点，水平向右为 $x$ 轴建立坐标系，**竖直向下**方向为 $y$ 轴建立坐标系。设滑梯对应函数为 $y(x)$ ，则有**限制条件**

$$y(0) = 0, \quad y(1) = h$$

要求使下滑时间

$$T = \int_0^1 \frac{\sqrt{1 + y'^2}}{\sqrt{2gy}} dx.$$

最小化。



注：这里的速率 $v = \sqrt{2gy}$ 是用能量守恒算出来的； $\sqrt{1 + y'^2} dx = \sqrt{dx^2 + dy^2}$ 是长度元（其中 $y' = dy/dx$ ）。

# 泛函

这里的下滑时间依赖于函数 $y(x)$  (在形式上, 它还依赖于 $y'(x)$ , 虽然 $y'(x)$ 是由 $y(x)$ 决定的)。这样的“依赖于函数的函数”称作**泛函(functional)**。

例如,

$$T = \int_0^1 \frac{\sqrt{1 + y'^2}}{\sqrt{2gy}} dx.$$

既可以看作 $y(x)$ 的泛函 (写成 $T(y(x))$ ), 也可以看作 $y(x)$ 和 $y'(x)$ 的泛函 (写成 $T(y(x), y'(x))$ )。

## 变分问题

在给定的限制条件下，求泛函的极值的问题称为**变分问题**。

例如最速降线问题中，我们要在 $y(0) = 0, y(1) = h$ 的限制下，计算泛函

$$T(y, y') = \int_0^1 \frac{\sqrt{1 + y'^2}}{\sqrt{2gy}} dx$$

的极小值（以及使其取到极小值的函数 $y(x)$ ）。

## 变分法的本质

先撇开最速降线问题不谈，我们考虑一个更一般的问题，设 $y(t)$ 是未知的函数， $L(\cdot, \cdot)$ 是给定的二元函数。我们要求

$$S = \int_a^b L(y, y') dt$$

在某些给定的边界条件（即函数 $y(t)$ 在 $a, b$ 处的属性）下的极值问题。

这里把自变量 $x$ 换成了 $t$ ，并不是为了故意混淆。在理论力学中，这个问题有着清晰的物理意义：泛函 $S$ 是作用量，函数 $L$ 是拉氏量， $y$ 是广义坐标， $t$ 是时间。

# 思路

我们在学习傅立叶变换时了解到了物理学中的函数总是能看成高维空间矢量。那么泛函其实也就和多元函数是一样的。求解变分问题的方法——变分法，其本质和多元函数求极值一样。

# 离散化

把区间 $[a, b]$ 均匀划分成 $N$ 份 ( $N$ 巨大)。记 $dt = \frac{b-a}{N}$ ;

$t_0 = a, t_1 = a + dt, t_2 = a + 2dt, \dots, t_N = b$ ;

记 $y_0 = y(t_0), y_1 = y(t_1), \dots, y_N = y(t_N)$ 。

求解函数 $y(t)$ 的问题被转化为求解 $N + 1$ 元变量 $(y_0, y_1, \dots, y_N)$ 。  
而泛函成了求和

$$S = dt \sum_i L \left( y_i, \frac{y_{i+1} - y_{i-1}}{2dt} \right).$$

当 $dt$ 很小时，我们用离散导数 $\frac{y_{i+1} - y_{i-1}}{2dt}$ 替代了 $y'(t_i)$ 。



## 求极值

取极值的条件是对所有  $y_k$ ,  $\frac{\partial S}{\partial y_k} = 0$ 。注意到对固定  $k$ , 求和中涉及  $y_k$  的就只有  $i = k - 1, k, k + 1$  三项。于是用链式法则求导后得到:

$$\begin{aligned} & \left. \frac{\partial L}{\partial y} \right|_{y=y_k, y'=\frac{y_{k+1}-y_{k-1}}{2dt}} \\ & - \frac{1}{2dt} \left[ \left. \frac{\partial L}{\partial y'} \right|_{y=y_{k+1}, y'=\frac{y_{k+2}-y_k}{2dt}} - \left. \frac{\partial L}{\partial y'} \right|_{y=y_{k-1}, y'=\frac{y_k-y_{k-2}}{2dt}} \right] \\ & = 0. \end{aligned}$$

第二行蓝色部分可以看成是从  $t_{k-1}$  到  $t_{k+1}$  时刻  $\frac{\partial L}{\partial y'}$  的变化量和时间  
的变化量 ( $2dt$ ) 之比, 也就是  $\frac{\partial L}{\partial y'}$  在  $t_k$  时刻的时间导数, 由此得出  
大名鼎鼎的欧拉-拉格朗日 (E-L) 方程——

## 欧拉-拉格朗日方程(E-L equation)

$$\frac{\partial L}{\partial y} - \frac{d}{dt} \frac{\partial L}{\partial y'} = 0.$$

注意和大多数教材或文献上的处理不同，我们的推导并不依赖于边界条件。另外，通过这种离散化的处理方式，我们揭示了变分问题的本质：变分法其实就是多元函数极值问题。

## 最速降线的求解(I)

在最速降线问题中,

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

(我们省略了对极值问题无影响的常数因子。)

$$\frac{\partial L}{\partial y} = -\frac{\sqrt{1 + y'^2}}{2y^{3/2}}$$

$$\frac{\partial L}{\partial y'} = \frac{y'}{\sqrt{y(1 + y'^2)}}$$

代入E-L方程

## 最速降线的求解(II)

$$\frac{\sqrt{1+y'^2}}{2y^{3/2}} + \frac{d}{dt} \left[ \frac{y'}{\sqrt{y(1+y'^2)}} \right] = 0$$

我们要寻找满足边界条件

$$y(0) = 0, y(1) = h$$

的解 $y(x)$ 。但是，这个微分方程看起来并不好解。

## 思考题



证明前述微分方程

$$\frac{\sqrt{1+y'^2}}{2y^{3/2}} + \frac{d}{dt} \left[ \frac{y'}{\sqrt{y(1+y'^2)}} \right] = 0$$

两边同乘以 $-y'$ 后, 可以转化为

$$\frac{d}{dt} \left[ \frac{1}{\sqrt{y(1+y'^2)}} \right] = 0.$$

它的解是

$$y(1+y'^2) = C$$

其中 $C$ 为待定常数。

## 思考题



证明由参数方程

$$t = a(\theta - \sin \theta), \quad y = a(1 - \cos \theta)$$

给出的滚轮线满足微分方程

$$y(1 + y'^2) = 2a.$$

由此得出最速降线为滚轮线（调节参数 $a$ 可以使滚轮线经过 $(1, h)$ ）。

## 一些留待思考的问题

- ▶ 最速降线问题中把E-L方程转化为一个“守恒方程”  
 $\frac{1}{\sqrt{y(1+y'^2)}} = C$ ，好像需要非常高超的数学技巧。有没有一般的方法完成这一步？
- ▶  $L$ 依赖于函数的二阶以上导数怎么办
- ▶  $L$ 依赖于多个函数怎么办
- ▶  $L$ 依赖于多元函数怎么办
- ▶ 限制条件是个积分方程怎么办

这些问题的部分解答请参考附录。关于变分的更广泛的知识可以参考你手中的课外读物。

# Homework

- ▶ 试用费马原理解决最速降线问题。
- ▶ 一根长为 $\ell$ 的均匀软绳两端悬挂在两个在同一水平线上，相距为 $h < \ell$ 的钉子上。求绳子保持静止时的形状。（参考附录中的拉格朗日因子法）



# 附录A: 哈密顿量

$$H = y' \frac{\partial L}{\partial y'} - L$$

# 哈密顿量

理论力学中，拉氏量( $L$ )形式上依赖于广义坐标( $y$ )和广义坐标的时间导数( $y'$ )。由此可以定义和 $y$ 对偶的广义动量

$$p_y = \frac{\partial L}{\partial y'}$$

以及哈密顿量

$$H = y' p_y - L$$

哈密顿量虽然数值上由上式定义，但是注意最后一定要把哈密顿量形式上写成依赖于广义坐标( $y$ )和广义动量( $p_y$ )。

## 拉氏量和哈密顿量的区别

拉氏量和哈密顿量虽然数值上显然不同，但更重要的是——在做任何形式推导之前，拉氏量要写成依赖于广义坐标和广义坐标的时间导数的形式；哈密顿量要写成依赖于广义坐标和广义动量的形式。

这绝对是学习理论力学时值得牢记的一件事情。

## 例子

经典力学系统中，拉氏量一般是动能和势能之差，对固定势场中的一维单粒子而言：

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

其中 $x$ 是粒子的坐标， $\dot{x} = dx/dt$ ， $V(x)$ 是势能函数。注意我们十分遵守规范：把 $L$ 写成了 $x$ 和 $\dot{x}$ 的函数。由此得出粒子的动量：

$$p_x = \frac{\partial L}{\partial \dot{x}} = m\dot{x}$$

哈密顿量

$$H = p_x \dot{x} - L = m\dot{x}^2 - L = \frac{1}{2}m\dot{x}^2 + V(x) = \frac{p_x^2}{2m} + V(x)$$

注意，在最后我们把 $H$ 写成了 $x$ 和 $p_x$ 的函数，这是行车规范。  
(温馨提示你：行车不规范，考试泪两行)

## 最速降线中的哈密顿量

哈密顿量的物理意义是总能量。在不显含时间的保守系统中，哈密顿量就是我们要求的“神奇的守恒量”。在最速降线问题中

$$L = \frac{\sqrt{1 + y'^2}}{\sqrt{y}}$$

容易算出广义动量

$$p_y = \frac{\partial L}{\partial y'} = \frac{y'}{\sqrt{y(1 + y'^2)}}$$

哈密顿量

$$H = y' p_y - L = -\frac{1}{\sqrt{y(1 + y'^2)}}$$

等等，行车规范呢？！

# 为什么没翻车

由于我们之后只用到哈密顿量守恒，再也没开车（做形式推导），所以不规范也没翻车。

# 附录B：变分法的其他情形

凭着感觉模仿就行

## 高阶时间导数

拉氏量里显含高阶时间导数，在理论力学里没有什么市场，但是作为纯数学问题有可能出现。

仅仅举个例子，设拉氏量 $L$ 形式上依赖于 $y(t), y'(t), y''(t)$ 。则E-L方程变为：

$$\frac{\partial L}{\partial y} - \frac{d}{dt} \frac{\partial L}{\partial y'} + \frac{d^2}{dt^2} \frac{\partial L}{\partial y''} = 0.$$

它同样不依赖于边界条件，你可以用离散化的办法给出证明。

对于更高阶的导数的情况想必你也会举一反三了。



## 多个函数

设 $L$ 形式上依赖于多个广义坐标 $y_1(t), y_2(t), \dots$ 和它们的时间导数, 则E-L方程对每个广义坐标分别成立:

$$\frac{\partial L}{\partial y_i} - \frac{d}{dt} \frac{\partial L}{\partial y_i'} = 0, \quad (i = 1, 2, \dots)$$

广义动量也会有多个

$$p_i = \frac{\partial L}{\partial y_i'}$$

哈密顿量则并不会会有多个, 而是求和:

$$H = \sum_i p_i y_i' - L.$$

也就是说, 多个广义坐标的情况, 单单考虑哈密顿量守恒也许并不足以解决问题了。

## 多元函数

迄今为止我们仅仅考虑了“时间”的函数 $y(t)$ 。在最速降线问题中， $t$ 其实是空间坐标，但这并不妨碍我们把 $t$ 当作时间坐标并用标准化的方法来解决数学问题。

在更一般的量子场论问题中，则需要考虑同时依赖于多维空间或时间坐标的广义坐标 $y(x_1, x_2, \dots)$ （更精确地讲，这时候我们称之为场，它弥漫于整个空间，而不再是一个单一的随时间变化的数）。这时候由于场弥漫于整个空间，讨论的对象也变成了“拉氏量密度”（单位时空体积内的拉氏量） $\mathcal{L}$ ：它依赖于 $y$ 和 $y$ 的各个偏导数： $\frac{\partial y}{\partial x_i}$

这时E-L方程成为：

$$\frac{\partial \mathcal{L}}{\partial y} - \sum_i \frac{\partial}{\partial x_i} \left[ \frac{\partial \mathcal{L}}{\partial \left( \frac{\partial y}{\partial x_i} \right)} \right] = 0.$$

## 积分限制条件

这次，我们仍然回到

$$S = \int_a^b L(y, y') dt$$

的极值问题。但是限制条件换成了

$$\int_a^b f(y, y') dt = c.$$

解决的办法就是引入拉格朗日乘子 $\lambda$ ，考虑

$$\tilde{S} = S + \lambda \left( \int_a^b f(y, y') dt - c \right)$$

的极值问题。把 $\lambda$ 也当成一个广义坐标。要求 $\tilde{S}$ 取极值则自然给出需要的限制条件 $\int_a^b f(y, y') dt - c = 0$ 。

再结合 $y$ 的E-L方程，原则上可以求解问题。