Chapter 1

Variational Method

变分法是数学物理中的一种重要方法，它研究某类特殊的变量—泛函的极大值和极小值问题.这里所讨论的只是变分法的一些最基本的内容以及一些应用.

§1.1 泛函和泛函的极值问题

1.1.1 泛函的概念

先从几个最简单的例子引进泛函的概念.

例 1 设已给 x 轴上两点 $x=x_0$ 和 $x=x_1$，$y=y(x)$ 是定义在区间 $[x_0,x_1]$ 上的有连续一阶导数的函数，则曲线 $y=y(x)$ 的长为

$$l[y(x)]=\int_{x_0}^{x_1}\sqrt{1+y'^2}\,\mathrm{d}x\,, \tag{1.1.1}$$

当函数 $y(x)$ 改变成另一函数 $y_1(x)$ 时，$l[y(x)]$ 也随之改变成代表函数 $y_1(x)$ 的曲线

在$[x_0,x_1]$上的弧长$l[y_1(x)]$.这就是说，由(1.1.1)式定义的变量l依赖于“整个函数”$y(x)$.

定义 1 具有某种共同性质的函数集，称为**函数类**.

例如，把在区间$[x_0,x_1]$上连续的函数的集合记为$C[x_0,x_1]$.把在区间$[x_0,x_1]$上有连续一阶导数的函数称为在$[x_0,x_1]$上的C_1类函数，这类函数的集合记为$C_1[x_0,x_1]$.以此类推，把在区间$[x_0,x_1]$上有连续n阶导数的函数称为此区间上的C_n类函数，这类函数的集合记为$C_n[x_0,x_1]$.如果一个函数$y(x)$在区间$[x_0,x_1]$上n次连续可导，它就属于$C_n[x_0,x_1]$，简记为$y(x)\in C_n$.这一记号也用于多元函数，例如$z(x,y)\in C_2(D)$，就意味着函数$z(x,y)$在域D上有连续的二阶偏导数.

例 2 设D是平面上的已给区域，函数$z(x,y)\in C_1(D)$，那么与之相应的曲面面积为

$$S[z(x,y)]=\iint_D\sqrt{1+z_x'^2+z_y'^2}\,\mathrm{d}x\mathrm{d}y \tag{1.1.2}$$

显然，变量S也是依赖于“整个函数”$z(x,y)$的.

现在引进**泛函**的概念.

定义 2 设R是一数域，设Y是已给定的某函数集，这一集合记为$\{y(x)\}$，如果对于Y中的，每一个函数$y(x)$，有变量$J\in R$的值与之对应，那么我们就说变量J是函数$y(x)$的**泛函**，记之为$J=J[y(x)]$.而此函数集称为泛函$J[y(x)]$]的**定义域**，有时也称为泛函的**容许函数**.简言之，泛函是函数集Y到数域R上的一个映射，映射的自变元是一个函数，而属于Y的每一个函数$y(x)$称为容许函数.读者不难自己类似的给出依赖于多个函数的泛函的定义.

按此定义，积分(1.1.1)和(1.1.2)分别是在$C_1[x_0,x_1]$和$C_1(D)$上的泛函.

例 3 傅里叶变换

$$F[f(x)]=\int_{-\infty}^{\infty}f(x)\mathrm{e}^{-\mathrm{i}kx}\mathrm{d}x$$

就是一个泛函.这个泛函以k为参量.一旦确定了k之后，对于每一个函数$f(x)$都确定了泛函的一个数值.函数$f(x)$就是泛函的容许函数.可以把这个泛函看成是自变量k的函数，记作$F(k)$，但是$F(k)$的数值并不是由函数的

定义来给出，而是按照上式根据函数 $f(x)$ 的具体形式来给出.

1.1.2 泛函的极值问题

1. 有关泛函极值的概念

变分法的基本问题是关于泛函的极值问题.例如 1696 年由约翰 伯努利提出的，并且对变分法的发展有过重大影响的**最速降线问题**(又称**捷线问题**)：在铅直平面内，所有连结两定点 A, B 的曲线中，求出一条曲线来，使初速为零的质点，在重力作用下，自 A 点沿着这条曲线下落到 B 点所需时间为最短(介质的摩擦和阻力不计).

首先若单纯从路程的角度，A 到 B 点的直线路程是最短的，但是因为下滑过程并不能较快获得很大的速度，所以所需要到时间不是最短的.

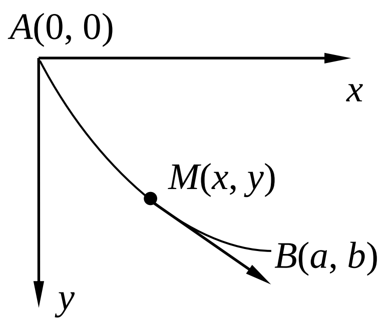

图 1.1

如图 1.1，以 A 点为坐标原点，Ox 轴取在水平方向，Oy 轴铅直向下.设 $y = y(x)$ 是连接点 $A(0,0)$ 和 $B(a,b)$ 的一条光滑曲线，质点沿这条曲线下滑.由于曲线是光滑的，在质点运动的切线方向上不受力.而法向作用力只改变运动速度的方向，不改变速度的大小.因初速度为零，故质点下滑到任意点 $M(x, y)$ 的速率为

$$v = \sqrt{2gy} \qquad (1.1.3)$$

其中 g 是重力加速度.

若以 S 表示曲线的弧长，$\mathrm{d}t$ 表示时间的微分，则

$$v = \frac{\mathrm{d}S}{\mathrm{d}t} = \frac{\sqrt{1+y'^2}\,\mathrm{d}x}{\mathrm{d}t} \qquad (1.1.4)$$

所以

$$\mathrm{d}t=\frac{\sqrt{1+y'^2}\mathrm{d}x}{v}=\frac{\sqrt{1+y'^2}}{\sqrt{2gy}}\mathrm{d}x \tag{1.1.5}$$

由此可得质点沿着曲线 $y=y(x)$ 自 A 点下滑到 B 点所需的时间 T 为：

$$T[y(x)]=\int_0^a\frac{\sqrt{1+y'^2}}{\sqrt{2gy}}\mathrm{d}x \tag{1.1.6}$$

这样，最速降线问题的数学提法是：求一条满足边界条件

$$y(0)=0,y(a)=b \tag{1.1.7}$$

的曲线 $y=y(x)$，使得泛函 $T[y(x)]$ 取最小值.

下面给泛函极值以类似于函数的极大极小值的定义.先给出曲线 $y=y(x)$ 的 ε-邻域的概念.

定义 3 定义在 $[x_0,x_1]$ 上的曲线 $y=y(x)$ 的 **ε-邻域**是指适合下列条件的一切可能的曲线 $y=y_1(x)$ (图 1.2)：它在整个区间 $[x_0,x_1]$ 上，满足不等式

$$|y_1(x)-y(x)|\le\varepsilon\,, \tag{1.1.8}$$

此时也称 $y_1(x)$ 与 $y(x)$ 有**零级 ε-接近度**.对于 C_1 类函数，曲线 $y=y(x)$ 的 ε-邻域的概念，有时除要求不等式(1.1.8)外，还要求满足下列不等式

$$|y_1'(x)-y'(x)|\le\varepsilon\,, \tag{1.1.9}$$

此时称 $y_1(x)$ 与 $y(x)$ 有**一级 ε-接近度**.完全类似的可以给出 **n 级 ε-接近度**的定义.固定一条曲线 $y=y_0(x)$ 之后，与 $y_0(x)$ 具有 n 级 ε-接近度的所有曲线的的集合称为 $y_0(x)$ 的 **n 级 ε-邻域**.

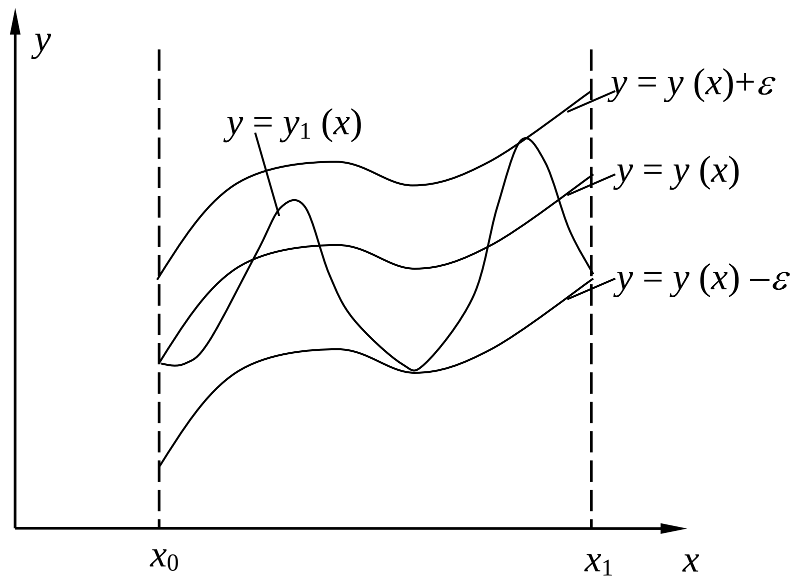

图 1.2

定义 4 设 $J[y(x)]$ 是定义在某个函数类 $\{y(x)\}$ 上的泛函，如果存在 ε >0，使得它在函数 $y_0(x)$ 上的值不小于它在函数类 $\{y(x)\}$ 中且与 $y_0(x)$ 有某确定级数的 ε-接近度的任何函数 $y_1(x)$ 上的值，即

$$J[y_0(x)] \geq J_1[y_1(x)]， \tag{1.1.10}$$

则称泛函 $J[y(x)]$ 在函数 $y_0(x)$ 上达到**相对极大值**.如果泛函 $J[y(x)]$ 在函数 $y_0(x)$ 的零级 ε-邻域，(1.1.10)式总是成立，那么称 $J[y(x)]$ 在函数 $y_0(x)$ 上达到**相对强极大值**，简称**强极大值**.如果泛函 $J[y(x)]$ 在函数 $y_0(x)$ 的一级 ε-邻域，(1.1.10)式总是成立，那么称 $J[y(x)]$ 在函数 $y_0(x)$ 上达到**相对弱极大值**，简称**弱极大值**.类似地，将(1.1.10)式的≥号改为≤号，就可以给出泛函**相对极小值**、**强极小值**、**弱极小值**等的定义.下面，我们把这些名词一律简称为**极值**.使泛函 J 取极值的曲线 $y_0(x)$ 称为**极值曲线**.

2. 变分学基本引理

为了以后的需要，现在我们给出**变分学的基本引理**.

引理 1 设 $f(x)$ 是区间 $[x_0, x_1]$ 内的连续函数，若对于任何在 $x = x_0$ 和 $x = x_1$ 两点为零的 C_2 类函数 $\eta(x)$，积分

$$\int_{x_0}^{x_1} f(x)\eta(x)\mathrm{d}x = 0\,, \tag{1.1.11}$$

则 $f(x)\equiv 0$.

证明 用反证法.设在某点 $\xi(x_0<\xi<x_1)$ 处 $f(x)$ 不等于零，例如 $f(\xi)>0$，因 $f(x)$ 是连续函数，故必存在 ξ 的邻域 $\xi_1<\xi<\xi_2$，在此邻域内 $f(x)>0$，命

$$\eta(x)=\begin{cases} 0 & x_0\le x\le \xi_1 \\ (x-\xi_1)^4(x-\xi_2)^4 & \xi_1\le x\le \xi_2 \\ 0 & \xi_2\le x\le x_1 \end{cases}$$

它满足引理中的一切条件，事实上，由定义有 $\eta(x_0)=\eta(x_1)=0$，不难验证 $(x-\xi_1)^4(x-\xi_2)^4$ 和它对 x 的一阶和二阶导数在 $x=\xi_1$ 及 $x=\xi_2$ 都为 0，而在区间 $[\xi_1,\xi_2]$ 之外 $\eta(x)$ 恒等于零，由此可见这函数和它的一阶及二阶导数在整个区间 $[x_0,x_1]$ 的连续性.但对此函数，积分

$$\int_{x_0}^{x_1} f(x)\eta(x)\mathrm{d}x = \int_{\xi_1}^{\xi_2} f(x)\eta(x)\mathrm{d}x > 0 \tag{1.1.12}$$

而与所设矛盾.因此 $f(x)=0$.**证明完毕**.

注意此引理的条件是，对于任何在 $x=x_0$ 和 $x=x_1$ 两点为零的 C_2 类函数 $\eta(x)$，(1.1.11)式都成立，才有引理的结论.若只对其中特定几个满足条件的 $\eta(x)$，从(1.1.11)式不能得出 $f(x)=0$ 的结论的.

引理 2 设 $f(x,y)$ 在域 D 内连续，若对任何在边界 C 上为零的域 D 内的 C_2 类函数 $\eta(x,y)$，有

$$\iint\limits_D f(x,y)\eta(x,y)\mathrm{d}x\mathrm{d}y = 0$$

那末 $f(x,y)=0$.

证明 若在 D 内某点 (a,b) 有 $f(a,b)>0$，则在以 (a,b) 为圆心，充分小 ρ 为半径的圆 K 内有 $f(x,y)>0$，令

$$\eta(x,y)=\begin{cases} 0 & (x-a)^2+(y-b)^2\ge \rho^2 \\ [(x-a)^2+(y-b)^2-\rho^2]^4 & (x-a)^2+(y-b)^2<\rho^2 \end{cases}$$

显然$\eta(x,y)$满足引理中的要求，但对这个η，有

$$\iint_D f(x,y)\eta(x,y)\mathrm{d}x\mathrm{d}y=\iint_K f(x,y)\eta(x,y)\mathrm{d}x\mathrm{d}y>0$$

这与所设矛盾，故 $f(x, y)$在 D 中处处为零.**证明完毕**.

显然，引理 1 和引理 2 分别针对一维和二维情况.对于三重积分或多重积分的情况，也有类似的引理.

§1.2 泛函的变分和最简单情形的欧拉方程

1.2.1 泛函的变分

设已给泛函

$$J[y(x)]=\int_{x_0}^{x_1}F(x,y,y')\mathrm{d}x \tag{1.2.1}$$

其中 F 是三个变元 x,y,y' 的连续函数，且有连续的二阶偏导数，又 $y(x)\in C_2$.

由于泛函的变分在研究泛函极值曲线的必要条件时的作用，相当于导数(或微分)在研究函数极值时的作用，下面先给出泛函变分的概念.

设想函数 $y(x)$ 稍有变动，变为 $y(x)+\delta y(x)$，这里 $\delta y(x)$ 表示一个函数，而不是 δ 乘 $y(x)$， $\delta y(x)$ 称为函数 $y(x)$的**变分**.如果 y 和 $y+\delta y$ 都是泛函的容许函数，我们来研究泛函(1.2.1)的值的增量(为方便起见，记 $\delta y'=(\delta y(x)')$.

$$\begin{aligned}\Delta J&=J[y+\delta y]-J[y]=\int_{x_0}^{x_1}[F(x,y+\delta y,y'+\delta y')-F(x,y,y')]\mathrm{d}x\\&=\int_{x_0}^{x_1}[\frac{\partial F}{\partial y}\delta y+\frac{\partial F}{\partial y'}\delta y']\mathrm{d}x+\int_{x_0}^{x_1}(\varepsilon_1\delta y+\varepsilon_2\delta y')\mathrm{d}x\end{aligned} \tag{1.2.2}$$

其中

$$\lim_{\delta y\to0,\delta y'\to0}\varepsilon_1=0,\ \lim_{\delta y\to0,\delta y'\to0}\varepsilon_2=0. \tag{1.2.3}$$

我们把(1.2.2)最后一个等式中的第一个积分称为泛函 $J[y]$ 在“点” $y(x)$ 的变分，记为δJ，即

$$\delta J=\int_{x_0}^{x_1}(\frac{\partial F}{\partial y}\delta y+\frac{\partial F}{\partial y'}\delta y')\mathrm{d}x \tag{1.2.4}$$

显然，变分δJ 关于δy 是线性的.再考虑泛函值的增量和变分之差 $\Delta J-\delta J$ ，如果记

$$\|\delta y\|=\max_{x_0\le x\le x_1}\left\{|\delta y(x)|,|\delta y'(x)|\right\}, \tag{1.2.5}$$

则

$$\left|\Delta J-\delta J\right|=\left|\int_{x_0}^{x_1}(\varepsilon_1\delta y+\varepsilon_2\varepsilon y')\mathrm{d}x\right|\le\max(\left|\varepsilon_1\right|+\left|\varepsilon_2\right|)(x_1-x_0)\left\|\delta y\right\|. \qquad (1.2.6)$$

则由(1.2.3)式可知，$\Delta J-\delta J$ 是$\left\|\delta y\right\|$的高阶无穷小量.这样，我们就可以说变分δJ是泛函增量ΔJ的线性主部.这与函数微分学中的情况是一致的，函数的微分是函数无穷小增量的线性主部.

例如，设 $J[y]=\int_{x_0}^{x_1}(y^2+y'^2)\mathrm{d}x$，则

$$\begin{aligned}\Delta J&=J[y+\delta y]-J[y]=\int_{x_0}^{x_1}[(y+\delta y)^2+(y'+\delta y')^2]-\int_{x_0}^{x_1}(y^2+y'^2)\mathrm{d}x\\&=\int_{x_0}^{x_1}(2y\delta y+2y'\delta y')\mathrm{d}x+\int_{x_0}^{x_1}[(\delta y)^2+(\delta y')^2]\mathrm{d}x\end{aligned} \qquad (1.2.7)$$

上式中第一个积分是与δy成线性的，第二个积分是$\left\|\delta y\right\|$的高阶无穷小量，所以

$$\delta J=\int_{x_0}^{x_1}(2y\delta y+2y'\delta y')\mathrm{d}x \qquad (1.2.8)$$

当然，这个结果用(1.2.4)式可以立即得到.

正如函数 $f(x)$ 的微分可以表示为对参数的导数，即

$$\mathrm{d}f=\frac{\mathrm{d}}{\mathrm{d}\alpha}f(x+\alpha\Delta x)\big|_{\alpha=0} \qquad (1.2.9)$$

证明：

$$\mathrm{d}f=\frac{\mathrm{d}f(x)}{\mathrm{d}x}\mathrm{d}x,$$

$$\mathrm{d}f=\frac{\mathrm{d}}{\mathrm{d}\alpha}f(x+\alpha\Delta x)\big|_{\alpha=0}=\frac{\mathrm{d}f(x+\alpha\Delta x)}{\mathrm{d}(x+\alpha\Delta x)}\frac{\mathrm{d}(x+\alpha\Delta x)}{\mathrm{d}\alpha}\big|_{\alpha=0}$$

$$=\frac{\mathrm{d}f(x+\alpha\Delta x)}{\mathrm{d}(x+\alpha\Delta x)}\Delta x\big|_{\alpha=0}=\frac{\mathrm{d}f(x)}{\mathrm{d}x}\Delta x=\frac{\mathrm{d}f(x)}{\mathrm{d}x}\mathrm{d}x$$

泛函 $J[y]$的变分δJ也可以表示为对参数α的导数，即

$$\delta J=\frac{\mathrm{d}}{\mathrm{d}\alpha}J[y+\alpha\delta y]\big|_{\alpha=0} \qquad (1.2.10)$$

由于现在 y 是个函数，证明步骤有所不同。事实上

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\alpha}J[y+\alpha\delta y]\big|_{\alpha=0}&=\int_{x_0}^{x_1}\frac{\mathrm{d}}{\mathrm{d}\alpha}F(x,y+\alpha\delta y,y'+\alpha\delta y')\mathrm{d}x\big|_{\alpha=0}\\&=\int_{x_0}^{x_1}(\frac{\partial F}{\partial y}\delta y+\frac{\partial F}{\partial y'}\delta y')\mathrm{d}x=\delta J\end{aligned} \qquad (1.2.11)$$

下面我们还会用(1.2.10)式进行推导.

泛函的变分(1.2.4)式也可写成如下形式.

$$\delta J=\int_{x_0}^{x_1}\delta F\mathrm{d}x \qquad (1.2.12)$$

这样，泛函的变分写成对于被积函数的变分.写成这个形式的时候，我们把函数 F 视为当 x 取固定值时依赖于函数 y 和 y' 的泛函.比较(1.2.12)与(1.2.4)式可知

$$\delta F=\frac{\partial F}{\partial y}\delta y+\frac{\partial F}{\partial y'}\delta y' \tag{1.2.13}$$

回顾一个二元函数

$$f=f(x,y)$$

它的微分是

$$\mathrm{d}f=\frac{\partial f}{\partial x}\mathrm{d}x+\frac{\partial f}{\partial y}\mathrm{d}y$$

可见，函数的一阶变分的形式与函数的一阶微分形式是相同的.只不过前者是函数的函数，而后者是自变量的函数.我们用字母 d 表示微分，而用字母δ表示变分.把泛函 F 对它所依赖的某个函数 y 求偏导数时的 $\partial F/\partial y$，与一个函数 f 对其一个自变量 x 的求偏导的操作是一样的.

再看二阶变分.把 δJ 看做是一个泛函，δF 是一个与 F 同样意义上的泛函，那么有

$$\delta^2 J=\int_{x_0}^{x_1}\delta^2 F\mathrm{d}x \tag{1.2.14}$$

把(1.2.13)再应用一遍，有

$$\delta^2 F=\delta(\delta F)=\frac{\partial\delta F}{\partial y}\delta y+\frac{\partial\delta F}{\partial y'}\delta y'=(\delta y\frac{\partial}{\partial y}+\delta y'\frac{\partial}{\partial y'})\delta F$$

$$=(\delta y\frac{\partial}{\partial y}+\delta y'\frac{\partial}{\partial y'})^2 F$$

$$=\frac{\partial}{\partial y}(\frac{\partial F}{\partial y}\delta y+\frac{\partial F}{\partial y'}\delta y')\delta y+\frac{\partial}{\partial y'}(\frac{\partial F}{\partial y}\delta y+\frac{\partial F}{\partial y'}\delta y')\delta y'$$

$$=(\frac{\partial^2 F}{\partial y^2}\delta y+\frac{\partial^2 F}{\partial y\partial y'}\delta y')\delta y+(\frac{\partial F}{\partial y}\delta y+\frac{\partial F}{\partial y'}\delta y')\delta\frac{\partial}{\partial y}y+(\frac{\partial F}{\partial y}\delta\frac{\partial}{\partial y}y+\frac{\partial F}{\partial y'}\delta\frac{\partial}{\partial y}y')\delta y$$

$$+(\frac{\partial^2 F}{\partial y'\partial y}\delta y+\frac{\partial^2 F}{\partial y'^2}\delta y')\delta y'=\frac{\partial^2 F}{\partial y^2}(\delta y)^2+2\frac{\partial^2 F}{\partial y\partial y'}\delta y'\delta y+\frac{\partial^2 F}{\partial y'^2}(\delta y')^2$$

再回顾二元函数的二阶微分.

$$\mathrm{d}^2 f=\frac{\partial^2 f}{\partial x^2}(\mathrm{d}x)^2+2\frac{\partial^2 f}{\partial x\partial y}\mathrm{d}x\mathrm{d}y+\frac{\partial^2 f}{\partial y^2}(\mathrm{d}y)^2=(\mathrm{d}x\frac{\partial}{\partial x}+\mathrm{d}y\frac{\partial}{\partial y})^2 f$$

因此，对函数的变分完全可以像对函数的微分那样来进行.

同理，容易证明，两个函数乘积的变分和函数倒数的变分分别为

$$\delta(F_1F_2)=F_2\delta F_1+F_1\delta F_2 \tag{1.2.15}$$

和

$$\delta(\frac{1}{F})=-\frac{1}{F^2}\delta F \tag{1.2.16}$$

现在来证明对于泛函变分的相应的公式,先看两个泛函 J_1 和 J_2 的乘积的变分，其中

$$J_1=\int_{a_1}^{b_1}F_1(x_1,y_1,y_1')\mathrm{d}x_1$$

和

$$J_2=\int_{a_2}^{b_2}F_2(x_2,y_2,y_2')\mathrm{d}x_2$$

那么，

$$\begin{aligned}\delta(J_1J_2)&=\int_{a_2}^{b_2}\mathrm{d}x_2\int_{a_1}^{b_1}\mathrm{d}x_1\delta(F_1F_2)=\int_{a_2}^{b_2}\mathrm{d}x_2\int_{a_1}^{b_1}\mathrm{d}x_1(F_2\delta F_1+F_1\delta F_2)\\&=\int_{a_2}^{b_2}F_2\mathrm{d}x_2\int_{a_1}^{b_1}\mathrm{d}x_1\delta F_1+\int_{a_2}^{b_2}\delta F_2\mathrm{d}x_2\int_{a_1}^{b_1}\mathrm{d}x_1F_1=J_2\delta J_1+J_1\delta J_2\end{aligned} \tag{1.2.17}$$

对于泛函(1.2.1)式的倒数$1/J$ 的变分可以用如下方式来证明.

$$\begin{aligned}\delta\left(\frac{1}{J}\right)&=\frac{1}{\int_{x_0}^{x_1}F(x,y+\delta y,y'+\delta y')\mathrm{d}x}-\frac{1}{\int_{x_0}^{x_1}F(x,y,y')\mathrm{d}x}\\&=\frac{1}{\int_{x_0}^{x_1}F(x,y,y')\mathrm{d}x+\int_{x_0}^{x_1}(\frac{\partial F}{\partial y}\delta y+\frac{\partial F}{\partial y'}\delta y')\mathrm{d}x}-\frac{1}{J}\\&=\frac{1}{J+\delta J}-\frac{1}{J}=\frac{1}{J}(\frac{1}{1+\delta J/J}-1)\end{aligned}$$

如果记$\|\delta y\|\equiv\max\limits_{x_0\le x\le x_1}\left\{|\delta y|,|\delta y'|\right\},\varepsilon_1=\max\limits_{x_0\le x\le x_1}\left\{\left|\frac{\partial F}{\partial y}\right|\right\},\varepsilon_2=\max\limits_{x_0\le x\le x_1}\left\{\left|\frac{\partial F}{\partial y'}\right|\right\}$，则

$\delta J\le|x_1-x_0|(\varepsilon_1+\varepsilon_2)\|\delta y\|$，即$\delta J$ 是$\|\delta y\|$的高一阶小量.

这样将$\left(1+\frac{\delta J}{J}\right)^{-1}$做泰勒展开至一阶小量，得到

$$\delta\left(\frac{1}{J}\right)=\frac{1}{J}\left[\left(1+\frac{\delta J}{J}\right)^{-1}-1\right]=\frac{1}{J}\ \ (1-\frac{\delta J}{J}-1)=-\frac{1}{J^2}\delta J$$

可见

$$\delta(\frac{1}{J})=-\frac{1}{J^2}\delta J \tag{1.2.18}$$

1.2.2 最简单情形的欧拉方程

1. 欧拉方程

定理 1 设 $y(x)$ 是泛函(1.2.1)的极值曲线，则在 $y=y(x)$ 这个泛函的变分

$$\delta J=0. \tag{1.2.19}$$

证明 由于这里要讨论的是极值曲线 $y(x)$ 所应满足的必要条件，所以我们可以选取特殊形式的函数 y^* 来比较 $J[y]$ 和 $J[y^*]$，例如，取

$$y^*(x)=y(x)+\alpha\delta y(x), \tag{1.2.20}$$

这里 α 是小参数(从而 $y^*(x)$ 与 $y(x)$ 就有指定的 ε-接近度)，$\delta y(x)$ 是任意固定的 C_2 类函数.

由所设条件，α 的函数 $J[\alpha]=J[y^*]=J[y+\alpha\delta y]$ 在 $\alpha=0$ 取极值，故根据(1.2.10)式，必有

$$\delta J=\frac{\mathrm{d}}{\mathrm{d}\alpha}J[y+\alpha\delta y]\Big|_{\alpha=0}=0 \tag{1.2.21}$$

即泛函 $J[y]$ 在 $y=y(x)$ 的变分 $\delta J=0$.**证明完毕**.

定理 2 设 $y(x)$ 是泛函(1.2.1)的极值曲线，则函数 $y=y(x)$ 必满足微分方程

$$F_y-\frac{\mathrm{d}}{\mathrm{d}x}F_{y'}=0 \tag{1.2.22}$$

或

$$F_{y'y'}y''+F_{yy'}y'+F_{xy'}-F_y=0 \tag{1.2.23}$$

其中，例如 $F_{yy'}$ 是 F 对 y 及 y' 的二阶偏导数.

证明 现在取更特殊形式的函数

$$y^*=y+\alpha\delta y$$

来比较 $J[y]$ 和 $J[y^*]$，这里要求 y^* 与 y 在点 $x=x_0$ 和 $x=x_1$ 两点处取相同的值，即

$$y^*(x_0)=y(x_0), y^*(x_1)=y(x_1)$$

亦即

$$\delta y(x_0)=0, \delta y(x_1)=0 \tag{1.2.24}$$

利用前面所得到的变分 δJ 的表达式(1.2.4)，并对积分的后一项进行分部积分，得

$$\delta J=\int_{x_0}^{x_1}(F_y\delta y+F_{y'}\delta y')\mathrm{d}x=F_{y'}\delta y(x)\big|_{x_0}^{x_1}+\int_{x_0}^{x_1}(F_y-\frac{\mathrm{d}}{\mathrm{d}x}F_{y'})\delta y\mathrm{d}x \tag{1.2.25}$$

由式(1.2.24)式知，第一项为零，因此

$$\delta J=\int_{x_0}^{x_1}(F_y-\frac{\mathrm{d}}{\mathrm{d}x}F_{y'})\delta y\mathrm{d}x \tag{1.2.26}$$

再由变分学基本引理知，极值曲线 $y(x)$ 必满足微分方程

$$F_y - \frac{\mathrm{d}}{\mathrm{d}x} F_{y'} = 0$$

证明完毕.

二阶常微分方程(1.2.22)或(1.2.23)称为泛函(1.2.1)的极值问题的**欧拉方程**，它的通解中含有两个任意常数.通常讨论泛函的极值时，对容许函数在边界 x_0 和 x_1 处的值总要加上一些附加的要求.例如，要求两端点固定，即

$$y(x_0) = y_0, y(x_1) = y_1 \tag{1.2.27}$$

其中 x_0, y_0, x_1, y_1 都是常数，从边界条件(1.2.27)确定出欧拉方程的解中的常数后，就可求得泛函(1.2.1)的可能的极值曲线.

2. 两种特殊情况

下面就两种特殊的情形讨论欧拉方程.

第一种情况.设 F 不依赖于 y，即 $F = F(x, y')$，因 $F_y = 0$，故此时欧拉方程为

$$\frac{\mathrm{d}}{\mathrm{d}x} F_{y'} = 0 \tag{1.2.28}$$

从而得**首次积分**

$$F_{y'}(x, y') = C_1 \tag{1.2.29}$$

这是一个不显含 y 的一阶微分方程，它可用解出 y' 再积分的方法或适当选择参数的方法求解.

例 1 求泛函

$$J = \int_{x_0}^{x_1} \frac{\sqrt{1+y'^2}}{x} \mathrm{d}x$$

且满足边界条件

$$y(x_0) = y_0, y(x_1) = y_1$$

的极值曲线.

解答 因 F 不依赖于 y，故其欧拉方程的首次积分为

$$F_{y'} = \frac{y'}{x\sqrt{1+y'^2}} = C_1$$

两边平方

$$y'^2 = C_1^2 x^2 (1 + y'^2)$$

解出 y' 得

$$y' = \frac{C_1 x}{\sqrt{1 - C_1^2 x^2}}$$

$$\mathrm{d}y=\frac{C_1x\mathrm{d}x}{\sqrt{1-C_1^2x^2}}=-\frac{1}{2C_1}\frac{\mathrm{d}(1-C_1^2x^2)}{\sqrt{1-C_1^2x^2}}=-\frac{1}{C_1}\mathrm{d}(1-C_1^2x^2)^{1/2}$$

再积分得

$$x^2+(y-C_2)^2=1/C_1^2,(C_1\neq 0)$$

这是中心在纵轴上的圆族，再由所给边界条件确定 C_1 和 C_2 即可得极值曲线.

第二种情况.设 F 不依赖于 x，即 $F=F(y,y')$，因 $F_{xy'}=0$，故由(1.2.23)式，此时欧拉方程成为

$$y''F_{y'y'}+y'F_{y'y}-F_y=0 \tag{1.2.30}$$

而由

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}(F-y'F_{y'})&=F_yy'+F_{y'}y''-y''F_{y'}-y'^2F_{yy'}-y'y''F_{y'y'}\\&=-y'(y''F_{y'y'}+y'F_{yy'}-F_y)=0\end{aligned} \tag{1.2.31}$$

可知方程(1.2.30)有首次积分

$$F-y'F_{y'}=C_1 \tag{1.2.32}$$

再积分此一阶微分方程，并由边界条件定出常数，就能求出可能的极值曲线.

例 2 在前节中，我们已把最速降线问题归结为：求泛函

$$T[y(x)]=\int_0^a\frac{\sqrt{1+y'^2}}{\sqrt{2gy}}\mathrm{d}x$$

的满足边界条件

$$y(0)=0,y(a)=b$$

的极值曲线.求出这一极值曲线.

解答 因为 $F=\sqrt{\dfrac{1+y'^2}{2gy}}$ 不依赖于 x，故此时欧拉方程有首次积分

$$F-y'F_{y'}=C$$

即

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

化简上式，并记 $C_1=1/2gC^2$，则得

$$\frac{1}{\sqrt{2gy}}\frac{1}{\sqrt{1+y'^2}}=C,\sqrt{y}\sqrt{1+y'^2}=1/C\sqrt{2g}$$

$$y(1+y'^2)=C_1.$$

我们用参数法解此方程：令

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

于是方程化为

$$y=\frac{C_1}{1+y'^2}=C_1\sin^2\frac{\theta}{2}=\frac{C_1}{2}(1-\cos\theta)$$

$$\mathrm{d}x=\frac{\mathrm{d}y}{y'}=\frac{C_1\sin(\theta/2)\cos(\theta/2)\mathrm{d}\theta}{\cot(\theta/2)}=\frac{C_1}{2}(1-\cos\theta)\mathrm{d}\theta$$

将后一等式积分得 $x=\frac{C_1}{2}(\theta-\sin\theta)$，故所求积分曲线的参数方程为

$$x=\frac{C_1}{2}(\theta-\sin\theta)+C_2,y=\frac{C_1}{2}(1-\cos\theta)$$

由边界条件 $y(0)=0$，可得 $C_2=0$，如此得到

$$x=\frac{C_1}{2}(\theta-\sin\theta),y=\frac{C_1}{2}(1-\cos\theta)$$

此为一族圆滚线(或称旋轮线)，滚圆的半径为 $C_1/2$，常数 C_1 可由圆滚线通过点 B 这个条件来确定.因此，所求的最速降线为通过 A、B 两点的圆滚线.

圆滚线的一个性质：从最高起点到第一个最低点之间，任取一点作为起始位置，质点的起始速率为零，则质点由此点沿圆滚线下降到最低点所用的时间是固定的。这个数值与从最高起点下降到最低点的时间一样。

这里我们再次指出，正如方程 $f'(x)=0$ 的根，只是函数 $y=f(x)$ 取极值的必要条件，根据已给条件定出的欧拉方程的特解，也只是泛函(1.2.1)取极值的必要条件，而不是充分条件，但是对于许多实际问题，特别是工程技术，力学，物理学中引出的变分问题，这样求得的函数 $y=y(x)$，常常就是使泛函 $J[y(x)]$ 取极值的函数.以下几节通过变分得到函数所应满足的方程，也都是泛函取极值的必要条件.

§1.3 多个函数和多个自变量的情形

1.3.1 多个函数

当泛函依赖于多个函数时，也不难给出欧拉方程.下面就含两个未知函数的泛函

$$J[y(x),z(x)]=\int_{x_0}^{x}F(x,y,y',z,z')\mathrm{d}x \tag{1.3.1}$$

进行讨论.设 F 对所有变元有连续二阶偏导数，又设 $y(x),z(x)\in C_2$.现假定泛函(1.3.1)在 $y(x)$， $z(x)$ 处取极值，讨论函数 $y(x)$， $z(x)$ 应满足的必要条件.

下面的讨论完全与只依赖于一个函数的泛函相同.取比较曲线 $y^*=y+\alpha\delta y, z^*=z+\alpha\delta z$，这里$\alpha$是小参数，且

$$\delta y(x_0)=\delta y(x_1)=\delta z(x_0)=\delta z(x_1)=0 \tag{1.3.2}$$

把 y^*，z^*代入(1.3.1)式，得到含参数α的函数

$$J(\alpha)=\int_{x_0}^{x_1}F(x,y+\alpha\delta y,y'+\alpha\delta y',z+\alpha\delta z,z'+\alpha\delta z')\mathrm{d}x$$

由于当$\alpha=0$时， $J(\alpha=0)$ 取极值，故有

$$J'(\alpha)\big|_{\alpha=0}=\int_{x_0}^{x_1}(F_y\delta y+F_{y'}\delta y'+F_z\delta z+F_{z'}\delta z')\mathrm{d}x=0$$

我们把

$$\delta J=\int_{x_0}^{x_1}(F_y\delta y+F_{y'}\delta y'+F_z\delta z+F_{z'}\delta z')\mathrm{d}x \tag{1.3.3}$$

称为泛函(1.3.1)的**变分**.由上可知，泛函(1.3.1)在 $y(x)$，$z(x)$ 取极值的必要条件是：在 $y(x)$， $z(x)$ 处泛函的变分为零.

如果把(1.3.3)式的积分中的第二项和第四项进行分部积分，并利用(1.3.2)式，即得

$$\begin{aligned}\delta J&=[F_{y'}\delta y+F_z\delta z]\big|_{x_0}^{x_1}+\int_{x_0}^{x_1}\left[\left(F_y-\frac{\mathrm{d}}{\mathrm{d}x}F_{y'}\right)\delta y+\left(F_z-\frac{\mathrm{d}}{\mathrm{d}x}F_{z'}\right)\delta z\right]\mathrm{d}x\\&=\int_{x_0}^{x_1}\left(F_y-\frac{\mathrm{d}}{\mathrm{d}x}F_{y'}\right)\delta y\mathrm{d}x+\int_{x_0}^{x_1}\left(F_z-\frac{\mathrm{d}}{\mathrm{d}x}F_{z'}\right)\delta z\mathrm{d}x=0\end{aligned} \tag{1.3.4}$$

特别取 $\delta z=0$，则

$$\int_{x_0}^{x_1}(F_y-\frac{\mathrm{d}}{\mathrm{d}x}F_{y'})\delta y\mathrm{d}x=0$$

再由基本引理，得

$$F_y-\frac{\mathrm{d}}{\mathrm{d}x}F_{y'}=0$$

同理

$$F_z-\frac{\mathrm{d}}{\mathrm{d}x}F_{z'}=0$$

综合上述可知，如果 $y(x)$， $z(x)$ 给泛函(1.3.1)以极值时，它们必须满足二阶微分方程组

$$\begin{cases} F_y - \dfrac{\mathrm{d}}{\mathrm{d}x} F_{y'} = 0 \\ F_z - \dfrac{\mathrm{d}}{\mathrm{d}x} F_{z'} = 0 \end{cases} \tag{1.3.5}$$

解此微分方程组，并利用已给边界条件，例如，在两端点固定的极值问题中，有

$$\begin{aligned} & y(x_0) = y_0, y(x_1) = y_1 \\ & z(x_0) = z_0, z(x_1) = z_1 \end{aligned} \tag{1.3.6}$$

这样确定出常数后，就可以求得可能的极值曲线.

对于两个函数 $y(x)$ 和 $z(x)$ 的情况，可以考虑这样的几何图像，建立坐标系，设想 x 是 y 和 z 的函数，空间有一条曲线，这条曲线实际上包含了 $y(x)$ 和 $z(x)$ 两个函数.

例 1 求泛函

$$J = \int_0^{\pi/2} (y'^2 + z'^2 + 2yz)\mathrm{d}x$$

满足边界条件

$$y(0) = 0, y\left(\frac{\pi}{2}\right) = 1; z(0) = 0, z\left(\frac{\pi}{2}\right) = -1$$

的极值曲线.

解答

$$F = y'^2 + z'^2 + 2yz$$

由(1.3.5)式，它的欧拉方程组为

$$\begin{cases} y'' - z = 0 \\ z'' - y = 0 \end{cases}$$

这是一个二阶线性联立方程组，消去 z 得

$$y^{(4)} - y = 0$$

容易求得其通解为

$$y = C_1\mathrm{e}^x + C_2\mathrm{e}^{-x} + C_3\cos x + C_4\sin x$$

再由 $z = y''$，得

$$z = C_1\mathrm{e}^x + C_2\mathrm{e}^{-x} - C_3\cos x - C_4\sin x$$

利用边界条件，求得

$$C_1 = 0, C_2 = 0, C_3 = 0, C_4 = 1$$

故所求极值曲线为

$$y = \sin x, z = -\sin x$$

如果有 M 个函数，一个自变量，$y_\alpha(x), (\alpha = 1, 2, \cdots, M)$，我们用$\{y_\alpha(x)\}$来表示这一组 M 个函数.泛函是

$$J[\{y_\alpha(x)\}] = \int_{x_0}^{x} F(x, \{y_\alpha(x)\}, \{y'_\alpha(x)\}) \mathrm{d}x \tag{1.3.7}$$

那么，容易将(1.3.5)式扩展为以下包含 M 个方程的方程组.

$$\frac{\partial F}{\partial y_\alpha} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y'_\alpha} = 0, (\alpha = 1, 2, \cdots, M) \tag{1.3.8}$$

1.3.2 多个自变量

现在来考虑多个自变量的情形.为简单起见，我们以二元函数$u(x, y)$的泛函

$$J[u] = \iint_D F(x, y, u, u_x, u_y) \mathrm{d}x\mathrm{d}y \tag{1.3.9}$$

为例来讨论重积分所表示的泛函的极值问题.这里 D 为 x-y 平面上的某给定区域，函数 F 对所有的变元有连续二阶偏导数，且$u(x, y) \in C_2(D)$.

设$u = u(x, y)$是泛函(1.3.9)取极值的函数(极值曲面)，为求出它应满足的必要条件，如前，取比较函数为

$$u^* = u(x, y) + \alpha\eta(x, y)$$

其中$\eta(x, y)$是任意的C_2类函数，但满足边界条件$\eta(x, y)|_C = 0$，α是小参数，显然，u^*和 u 在边界上取相同的值，将 u^*代入泛函(1.3.9)，得

$$J[u^*] = \iint_D F(x, y, u + \alpha\eta, p + \alpha\eta_x, q + \alpha\eta_y) \mathrm{d}x\mathrm{d}y = J(\alpha)$$

为书写简便，这里令 $p = u_x$， $q = u_y$.于是，与前相同，极值曲面$u(x, y)$应满足

$J'(0) = 0$，由此得

$$J'(0) = \iint_D (F_u\eta + F_p\eta_x + F_q\eta_y) \mathrm{d}x\mathrm{d}y$$

$$= \iint_D (F_u - \frac{\partial}{\partial x}F_p - \frac{\partial}{\partial y}F_q)\eta(x, y) \mathrm{d}x\mathrm{d}y$$

$$+ \iint_D [\frac{\partial}{\partial x}(F_p\eta) + \frac{\partial}{\partial y}(F_q\eta)] \mathrm{d}x\mathrm{d}y = 0 \tag{1.3.10}$$

利用格林公式，可将(1.3.10)式的后一个二重积分化为线积分，即

$$\iint_D [\frac{\partial}{\partial x}(F_p\eta)+\frac{\partial}{\partial y}(F_q\eta)]\mathrm{d}x\mathrm{d}y=\int_C (F_p\eta\mathrm{d}y-F_q\eta\mathrm{d}x) \tag{1.3.11}$$

由边界条件$\eta(x,y)|_C=0$，故这一线积分为 0，所以

$$J'(0)=\iint_D (F_u-\frac{\partial}{\partial x}F_p-\frac{\partial}{\partial y}F_q)\eta(x,y)\mathrm{d}x\mathrm{d}y=0$$

再利用基本引理，就得到

$$F_u-\frac{\partial}{\partial x}F_p-\frac{\partial}{\partial y}F_q=0\,. \tag{1.3.12}$$

这就是函数$u(x,y)$必须满足的方程，它称为**奥斯特洛格拉得斯基方程**，亦称欧拉方程.这个方程中含有未知函数的偏导数，它是一个偏微分方程.

例 2 泛函

$$J[u]=\iint_D (u_x^2+u_y^2)\mathrm{d}x\mathrm{d}y$$

的奥氏方程为

$$\Delta u=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0$$

这是有名的二维拉普拉斯方程. $J[u]$的满足已知边界条件$u(x,y)|_C=u_0(x,y)$的极值曲面，就是区域D内的狄里克雷问题或者第一边值问题

$$\begin{cases}\Delta u=\dfrac{\partial^2 u}{\partial x^2}+\dfrac{\partial^2 u}{\partial y^2}=0,(x,y)\in D\\ \quad u(x,y)|_C=u_0(x,y)\end{cases}$$

的解.

回顾以上几个定理，证明的思路都是：将泛函写成为某一个参量α的函数，然后对于参量求导取极值.如果需要，还可以将泛函设成是α和β甚至更多参量的函数，然后对各参量求导取极值.

如果有 N 个自变量$x_i,(i=1,2,\cdots,N)$，我们用$\{x_i\}$来表示这一组 N 个自变量，一个函数$y=y(x_1,x_2,\cdots,x_N)=y(\{x_i\})$，泛函就是

$$J[y]=\int_R F(x_1,x_2,\cdots,x_N,y,\frac{\partial y}{\partial x_1},\frac{\partial y}{\partial x_2},\cdots,\frac{\partial y}{\partial x_N})\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_N \tag{1.3.13a}$$

此式也简记为

$$J[y]=\int_R F(\{x_i\},y,\{\frac{\partial y}{\partial x_i}\})\,\mathrm{d}^N x \tag{1.3.13b}$$

容易将(1.3.12)式扩展为以下方程.

$$\frac{\partial F}{\partial y}-\sum_{i=1}^{N}\frac{\partial}{\partial x_i}\frac{\partial F}{\partial(\partial y/\partial x_i)}=0 \tag{1.3.14}$$

最后，若是 N 个变量 $x_i,(i=1,2,\cdots,N)$，M 个函数 $y_\alpha(\{x_i\}),(\alpha=1,2,\cdots,M)$，泛函就是

$$J[\{y_\alpha(\{x_i\})\}]=\int_R F(\{x_i\},\{y_\alpha(\{x_i\})\},\{\frac{\partial y_\alpha}{\partial x_i}\})\mathrm{d}^N x \tag{1.3.15}$$

结合(1.3.8)和(1.3.14)，得到多变量多函数时的欧拉方程组如下.

$$\frac{\partial F}{\partial y_\alpha}-\sum_{i=1}^{N}\frac{\partial}{\partial x_i}\frac{\partial F}{\partial(\partial y_\alpha/\partial x_i)}=0,(\alpha=1,2,\cdots,M) \tag{1.3.16}$$

这是包含有 M 个方程的方程组.

§1.4 泛函的条件极值问题

类似于求函数的条件极值问题，在许多泛函的极值问题中，容许函数也受到一些附加条件的限制.下面我们将讨论两种类型的泛函的条件极值问题.

1.4.1 等周问题

这种问题起源于下面的几何问题：如图 1.3，求一条通过点 A、B 且有定长 l 的曲线，使曲边梯形 $ABCD$ 的面积取极大值.设 $y=y(x)$ 为所求曲线，不难看出，问题归结为求泛函

$$S[y]=\int_{x_0}^{x_1}y(x)\mathrm{d}x \tag{1.4.1}$$

满足边界条件 $y(x_0)=y_0,y(x_1)=y_1$ 和条件 $\int_{x_0}^{x}\sqrt{1+y'^2}\mathrm{d}x=l$ 的极大值.

由于题中要求曲边梯形 $ABCD$ 的周长为定值，因之这个问题叫做**等周问题**.

一般地，提出下列问题：在使泛函

$$J_1=\int_{x_0}^{x}G(x,y,y')\mathrm{d}x=l\ (常数) \tag{1.4.2}$$

的一切曲线中，确定出这样一条曲线，它使泛函

$$J=\int_{x_0}^{x}F(x,y,y')\mathrm{d}x \tag{1.4.3}$$

取极值.这类问题统称为**等周问题**，条件(1.4.2)称**等周条件**.

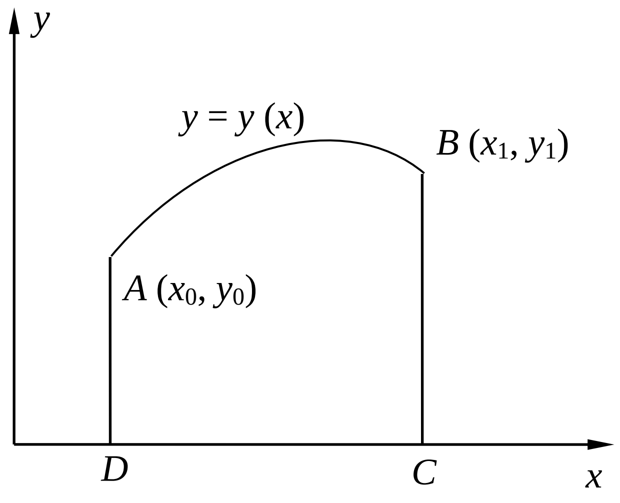

图 1.3

类似于求多元函数条件极值，拉格朗日乘子法，对等周问题也有下面的定理，它把条件极值的变分问题归结为无条件极值的变分问题.

定理 1 (拉格朗日) 若曲线 $y=y(x)$ 是所提问题的极值曲线，但 $y=y(x)$ 不是泛函 J_1 的极值曲线，则必存在一常数λ，使得 $y=y(x)$ 是泛函

$$\int_{x_0}^{x_1} H(x,y,y')\mathrm{d}x \qquad (1.4.4)$$

的极值曲线，其中

$$H=F+\lambda G. \qquad (1.4.5)$$

故 $y=y(x)$ 必满足欧拉方程

$$H_y-\frac{\mathrm{d}}{\mathrm{d}x}H_{y'}=0. \qquad (1.4.6)$$

证明 取比较函数

$$y^*(x)=y(x)+\alpha\eta_1(x)+\beta\eta_2(x) \qquad (1.4.7)$$

其中 α 和 β 是小参数， $\eta_1(x)$ 和 $\eta_2(x)$ 是任意固定的 C_2 类函数，且 $\eta_1(x_0)=\eta_1(x_1)=\eta_2(x_0)=\eta_2(x_1)$=0.将 $y^*(x)$分别代入(1.4.2)和(1.4.3)式，得

$$J(\alpha,\beta)=\int_{x_0}^{x} F(x,y+\alpha\eta_1+\beta\eta_2,y'+\alpha\eta_1'+\beta\eta_2')\mathrm{d}x \qquad (1.4.8)$$

和

$$J_1(\alpha,\beta)=\int_{x_0}^{x} G(x,y+\alpha\eta_1+\beta\eta_2,y'+\alpha\eta_1'+\beta\eta_2')\mathrm{d}x=l \qquad (1.4.9)$$

因 $y(x)$ 是在条件(1.4.2)之下泛函 $J[y]$ 的极值曲线，而 $\alpha=\beta=0$ 时 $y^*=y$，所以当 $\alpha=\beta=0$ 时二元函数 $J(\alpha,\beta)$ 在条件 $J_1(\alpha,\beta)=l$ 之下取极值.于是由普通函数的条

件极值问题的拉格朗日乘子法，有

$$\left.\begin{aligned}\frac{\partial J}{\partial\alpha}+\lambda\frac{\partial J_1}{\partial\alpha}=0\\ \frac{\partial J}{\partial\beta}+\lambda\frac{\partial J_1}{\partial\beta}=0\end{aligned}\right\},(\alpha=\beta=0) \tag{1.4.10}$$

其中λ是待定常数.把 J 和 J_1 代入这两个方程，得

$$\left.\begin{aligned}\int_{x_0}^{x_1}[(F_y+\lambda G_y)\eta_1+(F_{y'}+\lambda G_{y'})\eta_1']\mathrm{d}x=0\\ \int_{x_0}^{x_1}[(F_y+\lambda G_y)\eta_2+(F_{y'}+\lambda G_{y'})\eta_2']\mathrm{d}x=0\end{aligned}\right\} \tag{1.4.11}$$

再分别将这两个积分的第二项进行分部积分，并注意到 $\eta_1(x_0)=\eta_1(x_1)=\eta_2(x_0)=\eta_2(x_1)=0$，得

$$\int_{x_0}^{x_1}[(F_y+\lambda G_y)-\frac{\mathrm{d}}{\mathrm{d}x}(F_{y'}+\lambda G_{y'})]\eta_1\mathrm{d}x=0 \tag{1.4.12}$$

$$\int_{x_0}^{x_1}[(F_y+\lambda G_y)-\frac{\mathrm{d}}{\mathrm{d}x}(F_{y'}+\lambda G_{y'})]\eta_2\mathrm{d}x=0 \tag{1.4.13}$$

由所设条件，$y(x)$ 不满足方程 $G_y-\frac{\mathrm{d}}{\mathrm{d}x}G_{y'}=0$，故可选取 $\eta_2(x)$ 使

$$\int_{x_0}^{x_1}(G_y-\frac{\mathrm{d}}{\mathrm{d}x}G_{y'})\eta_2\mathrm{d}x\neq 0 \tag{1.4.14}$$

从而由(1.4.13)式，得

$$\lambda=\frac{\int_{x_0}^{x_1}(F_y-\frac{\mathrm{d}}{\mathrm{d}x}F_{y'})\eta_2\mathrm{d}x}{\int_{x_0}^{x_1}(G_y-\frac{\mathrm{d}}{\mathrm{d}x}G_{y'})\eta_2\mathrm{d}x} \tag{1.4.15}$$

因η_2 与η_1 是相互独立的，故λ与η_1 无关.同理，λ也与η_2 无关.所以,λ是与η_1 和η_2 都无关的常数.最后再对(1.4.12)式应用基本引理，得

$$F_y+\lambda G_y-\frac{\mathrm{d}}{\mathrm{d}x}(F_{y'}+\lambda G_{y'})=0 \tag{1.4.16}$$

证明完毕.

式(1.4.16)是一个含参数λ的二阶常微分方程，它的通解中含三个待定常数 C_1、C_2、λ，它们可由等周条件(1.4.2)和已给边界条件，例如，$y(x_0)=y_0$，$y(x_1)=y_1$ 来确定.

例 1 求泛函

$$S=\int_{x_0}^{x_1}y\mathrm{d}x, y(x_0)=y_0, y(x_1)=y_1$$

满足等周条件

$$\int_{x_0}^{x_1}\sqrt{1+y'^2}\mathrm{d}x=l$$

的极值曲线.

解答 作辅助泛函

$$J=\int_{x_0}^{x_1}(y+\lambda\sqrt{1+y'^2})\mathrm{d}x$$

令 $H=y+\lambda\sqrt{1+y'^2}$ ，其欧拉方程为

$$H_y-\frac{\mathrm{d}}{\mathrm{d}x}H_{y'}=0$$

因 H 不含 x，故其首次积分为

$$H-y'H_{y'}=C_1$$

即

$$y+\lambda\sqrt{1+y'^2}-\frac{\lambda y'^2}{\sqrt{1+y'^2}}=C_1$$

或

$$y-C_1=-\frac{\lambda}{\sqrt{1+y'^2}}$$

用参数法解此微分方程，令 $y'=\tan t$ ，得

$$y-C_1=-\lambda\cos t,\mathrm{d}x=\frac{\mathrm{d}y}{y'}=\frac{\lambda\sin t\mathrm{d}t}{\tan t}=\lambda\cos t\mathrm{d}t,x=\lambda\sin t+C_2$$

于是所求极值曲线的参数方程是

$$x-C_2=\lambda\sin t,y-C_1=-\lambda\cos t$$

消去参数 t，得 $(x-C_2)^2+(y-C_1)^2=\lambda^2$，这是一族圆，$C_1,C_2$ 及 λ 可由条件

$$y(x_0)=y_0,y(x_1)=y_1,\int_{x_0}^{x_1}\sqrt{1+y'^2}\mathrm{d}x=l$$

确定.

1.4.2 测地线问题

求两个函数 $y(x)$ 及 $z(x)$，它们给泛函

$$J=\int_{x_0}^{x_1}F(x,y,y',z,z')\mathrm{d}x \tag{1.4.17}$$

以极值，且满足附加条件

$$G(x,y,z)=0 \tag{1.4.18}$$

这个问题，在力学中是所谓约束问题，从几何上看就是求一条曲线，它位于曲面(1.4.18)内且给泛函(1.4.17)以极值.注意现在是一个自变量 x 和两个函数 $y(x)$

和 $z(x)$.考虑约束曲面时，可建立坐标系，设想 x 是 y 和 z 的函数，随着 y 和 z 的变化，x 的数值构成一个曲面，极值曲线就在这个曲面上，它包含了 $y(x)$ 和 $z(x)$ 两个函数.

解这个问题的一个很自然的想法就是：从方程(1.4.18)确定 z 为 x, y 的函数，且将这个函数代入(1.4.17)式，这样我们就归结为只有一个待求函数 $y(x)$ 的没有附加条件的寻常的变分问题.下面我们就沿着这个想法来求出极值函数 $y(x)$ 及 $z(x)$ 应当满足的方程.

设 $G_z \neq 0$，则依隐函数存在定理，可以由(1.4.18)式解出 $z=\varphi(x,y)$，代入(1.4.17)式得

$$J=\int_{x_0}^{x_1} F(x,y,y',\varphi,\varphi_x+\varphi_y y')\mathrm{d}x \tag{1.4.19}$$

为方便起见，令 $F^*(x,y,y')=F(x,y,y',\varphi,\varphi_x+\varphi_y y')$，于是 $y(x)$ 应满足欧拉方程

$$F_y^* - \frac{\mathrm{d}}{\mathrm{d}x}F_{y'}^* = 0 \tag{1.4.20}$$

因

$$\begin{aligned} &F_y^* = F_y + F_z\varphi_y + F_{z'}(\varphi_{xy}+\varphi_{yy}y'),\\ &F_{y'}^* = F_{y'} + F_{z'}\varphi_y,\\ &\frac{\mathrm{d}}{\mathrm{d}x}F_{y'}^* = \frac{\mathrm{d}}{\mathrm{d}x}F_{y'} + \varphi_y\frac{\mathrm{d}}{\mathrm{d}x}F_{z'} + F_{z'}(\varphi_{xy}+\varphi_{yy}y') \end{aligned} \tag{1.4.21}$$

这样(1.4.20)式就成为

$$F_y + \varphi_y\left(F_z - \frac{\mathrm{d}}{\mathrm{d}x}F_{z'}\right) - \frac{\mathrm{d}}{\mathrm{d}x}F_{y'} = 0 \tag{1.4.22}$$

又 $\varphi_y = \dfrac{\partial z}{\partial y} = -\dfrac{G_y}{G_z}$，代入(1.4.22)式，得

$$\frac{1}{G_y}\left(F_y - \frac{\mathrm{d}}{\mathrm{d}x}F_{y'}\right) = \frac{1}{G_z}\left(F_z - \frac{\mathrm{d}}{\mathrm{d}x}F_{z'}\right) \tag{1.4.23}$$

这个等式的两端沿着极值曲线是同一个函数，记之为 $-\lambda(x)$，于是得

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}x}F_{y'} - [F_y + \lambda(x)G_y] = 0 \\ \dfrac{\mathrm{d}}{\mathrm{d}x}F_{z'} - [F_z + \lambda(x)G_z] = 0 \end{cases} \tag{1.4.24}$$

这就是极值函数 $y(x)$，$z(x)$ 所应满足的二阶微分方程组.

综合上述，我们得到了下面的定理.

定理 2 设函数 $y(x)$， $z(x)$ 使泛函(1.4.17)取极值，并满足方程(1.4.18)，则必有适当的因子 $\lambda(x)$ 存在，使得 $y(x)$，$z(x)$满足泛函 $J^*=\int_{x_0}^{x_1} H(x,y,y',z,z')\mathrm{d}x$ 的欧拉方程

$$\begin{cases}\dfrac{\mathrm{d}}{\mathrm{d}x}H_{y'}-H_y=0\\ \dfrac{\mathrm{d}}{\mathrm{d}x}H_{z'}-H_z=0\end{cases} \tag{1.4.25}$$

其中 $H=F+\lambda(x)G$.

注意(1.4.24)与(1.4.25)的差别，后者的第一项中写的是 $F_{y'}$ 而不是 $H_{y'}$，这是因为现在的约束条件(1.4.18)中，G 不含 y' 和 z'.

例 2(短程线问题) 求在曲面 $\varphi(x,y,z)=0$ 上两定点 $A(x_0,y_0,z_0)$ 和 $B(x_1,y_1,z_1)$ 间的最短距离.

解答 我们知道，两点间的距离是由公式

$$l=\int_{x_0}^{x_1}\sqrt{1+y'^2+z'^2}\,\mathrm{d}x$$

表示的.因此，这个问题就是要在条件 $\varphi(x,y,z)=0$ 下求泛函 l 的极小值.作辅助泛函

$$l^*=\int_{x_0}^{x_1}[\sqrt{1+y'^2+z'^2}+\lambda(x)\varphi(x,y,z)]\mathrm{d}x$$

写出它的欧拉方程

$$\lambda(x)\varphi_y-\frac{\mathrm{d}}{\mathrm{d}x}\frac{y'}{\sqrt{1+y'^2+z'^2}}=0$$

$$\lambda(x)\varphi_z-\frac{\mathrm{d}}{\mathrm{d}x}\frac{z'}{\sqrt{1+y'^2+z'^2}}=0$$

由这两个方程及约束条件 $\varphi(x,y,z)=0$ 就可以确定出因子 $\lambda(x)$ 及函数

$$y=y(x) \text{ 和 } z=z(x)$$

而已给边界条件 $y(x_0)=y_0, y(x_1)=y_1$ 和 $z(x_0)=z_0, z(x_1)=z_1$，可用来确定通解中的常数.

在力学中，常要用到更一般形式的约束问题，现将结果叙述如下：求含有 n 个变量函数的泛函

$$J[y_1,y_2,\cdots,y_n]=\int_{x_0}^{x_1}F(x;y_1,y_2,\cdots,y_n;y_1',y_2',\cdots,y_n')\mathrm{d}x \tag{1.4.26}$$

在 m 个约束条件

$$\varphi_k(x,y_1,y_2,\cdots,y_n)=0,k=1,2,\cdots m,\ (m<n) \tag{1.4.27}$$

下的极值问题，归结为泛函

$$J^*=\int_{x_0}^{x_1}[F+\sum_{k=1}^{m}\lambda_k(x)\varphi_k]\mathrm{d}x \tag{1.4.28}$$

的欧拉方程组

$$\frac{\mathrm{d}}{\mathrm{d}x}F_{y_i'}-\left[F_{y_i}+\sum_{k=1}^{m}\lambda_k(x)\frac{\partial\varphi_k}{\partial y_i}\right]=0,i=1,2,\cdots,n \tag{1.4.29}$$

§1.5 自然边界条件

前面我们讲了泛函的极值曲线或极值曲面必满足欧拉方程，它是一个常微分方程或偏微分方程.为了把极值曲线或极值曲面完全确定出来，对它们就要附加一定的边界条件，我们在前面所举的例子中，所加的边界条件，都是预先给定极值函数在边界上所取的值.例如条件(1.2.27)式.相应地，对两端点变分的结果是零，见(1.2.24)式.

但是在实际问题中所遇到的边界条件却要广泛得多，甚至连边界本身都可以变动，仍以最简单的典型泛函

$$J[y]=\int_{x_0}^{x_1}F(x,y,y')\mathrm{d}x \tag{1.5.1}$$

为例讨论.一般说来，它的积分限 x_0,x_1 和容许函数在边界上的值 $y(x_0)$ 和 $y(x_1)$ 都是可变的，即曲线的端点可变动.为了简单起见，下面只讨论，当(1.5.1)式中的积分限 x_0,x_1 不变，而 $y(x_0)$ 和 $y(x_1)$ 的值可变的情形.从几何上看，就是容许曲线的两端点可以在两条平行于 y 轴的直线 $x=x_0$ 和 $x=x_1$ 上变动，见图 1.4.

泛函极值的基本条件是，其变分为零.

$$\delta J=0 \tag{1.5.2}$$

我们只能从此式出发来进行讨论.

由(1.2.25)式我们知，

$$\delta J=F_{y'}\delta y(x)|_{x_0}^{x_1}+\int_{x_0}^{x_1}(F_y-\frac{\mathrm{d}}{\mathrm{d}x}F_{y'})\delta y\mathrm{d}x=0 \tag{1.5.3a}$$

注意，前面的欧拉方程是根据固定端点的(1.2.24)式的条件而得到.此处不能利用这一条件了，因为端点是变动的.不过，由于变动端点泛函的容许曲线包括了固定端点泛函的容许曲线，因而只要 $y(x)$ 是泛函(1.5.1)的极值曲线，那么它一定是

满足欧拉方程的.因为假定我们找到了这一泛函问题的解 $y=y^*(x)$，则端点值 $y(x_0)$ 和 $y(x_1)$ 为固定的常数，于是可将这一边界条件增加到原来的变分问题中去，构成固定边界变分问题，显然 $y^*(x)$ 也就必然是这一固定变分问题的解.因而 $y^*(x)$ 必满足欧拉方程

$$F_y-\frac{\mathrm{d}}{\mathrm{d}x}F_{y'}=0$$

也就是说，无论端点是否固定，极值曲线总是满足欧拉方程的.因此式(1.5.3a)就只剩下第一项.我们有

$$\delta J=(F_{y'}\delta y)_{x_1}-(F_{y'}\delta y)_{x_0}=0 \tag{1.5.3b}$$

因为函数在端点的改变量(图 1.4) $(\delta y)_{x_0}$ 和 $(\delta y)_{x_1}$ 是独立的，故必有

$$F_{y'}\Big|_{x=x_0}=0, F_{y'}\Big|_{x=x_1}=0 \tag{1.5.4}$$

这就是 $y(x)$ 在端点 $x=x_0$ 和 $x=x_1$ 所应满足的条件，称为**自然边界的条件**.

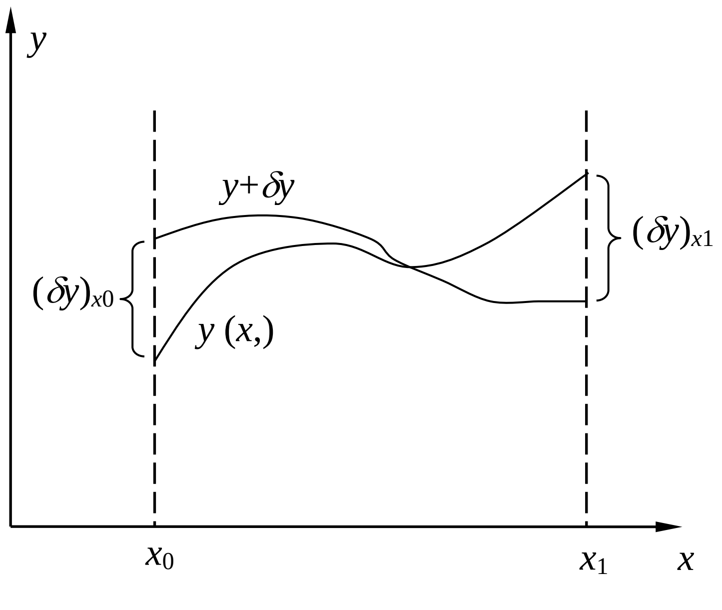

图 1.4 积分限 x_0, x_1 不变.但是端点的位置，也就是在 x_0, x_1 处的值 $y(x_0)$ 和 $y(x_1)$ 可以变化.

特别的，如果容许曲线的一端固定，例如左端点固定，即

$$y(x_0)=y_0 \tag{1.5.5}$$

右端点在定值线 $x=x_1$ 上变动，则此时 $\delta y(x)=0$，而 $\delta y(x_1)$ 是任意的，于是由(1.5.3)式可得自然边界条件

$$F_{y'}\Big|_{x=x_1}=0 \tag{1.5.6}$$

例 1 在§1.2 节中，我们已求得最速降线问题的欧拉方程的通解是圆滚线族

$$x=C_1(\theta-\sin\theta)+C_2, y=C_1(1-\cos\theta)$$

如果设最速线的左端固定，即 $y(0)=0$，而右端点可在直线 $x=x_1$ 上变动.此时，由 $y(0)=0$ 仍得 $C_2=0$，而由(1.5.6)来确定 C_1，因 $F=\sqrt{\dfrac{1+y'^2}{2gy}}$，

$$F_{y'}\big|_{x=x_1}=\frac{y'}{\sqrt{2gy(1+y'^2)}}\Big|_{x=x_1}=0$$

由此有 $y'(x_1)=0$，因此，所求的圆滚线应当和直线 $x=x_1$ 相交成直角(图 1.5)，由 $y'(x_1)=0$，不难算得顶点 B 对应于 $\theta=\pi$，所以

$$x_1=C_1\pi, C_1=\frac{x_1}{\pi}$$

于是，可能的极值曲线只能是

$$x=\frac{x_1}{\pi}(\theta-\sin\theta), y=\frac{x_1}{\pi}(1-\cos\theta)$$

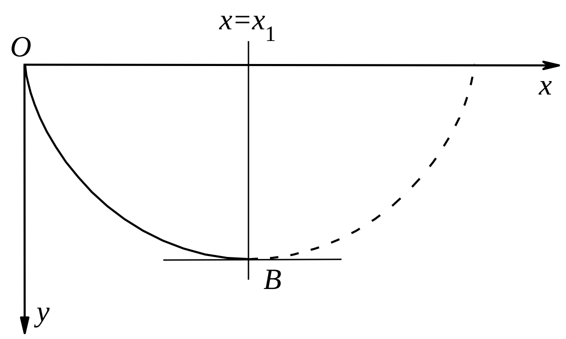

图 1.5 圆滚线起点固定在原点，终点在直线 $x=x_1$ 上变动.则圆滚线和直线 $x=x_1$ 一定直角相交.

§1.6 变分原理

数学中的变分学理论在物理上的应用，一般称为变分原理.变分原理是从自然现象中归纳得到，然后以原理的形式表现出来，作为推导物质的运动或者存在的状态所应该满足的方程的出发点.变分原理是一种属于科学假设范畴的极值原则，它在各种学科中表现为不同的形式.

例如，在光学中就表现为**光行最速原则—费马原则**：光从 A 点传到 B 点，永远沿时间为最少的路线传播.设光在某平面介质内沿曲线 $y=y(x)$ 进行传播，用 $v(x,y)$ 表示在介质内点(x,y)的光速，则类似于最速降线问题的推导，可知光线沿曲线 $y=y(x)$ 由 $A(x_0,y_0)$ 点传播到 $B(x_1,y_1)$ 点所需的时间 T 为

$$T[y]=\int_{x_0}^{x_1}\frac{\sqrt{1+y'^2}}{v(x,y)}\mathrm{d}x \tag{1.6.1a}$$

根据费马原则，确定光的路线问题，就变成确定一条曲线，使泛函 $T[y]$ 取最小值.

如果用时间 t 作为自变量，那么光前进的路径上每一点坐标 x,y 都是时间的函数，$x=x(t),y=y(t)$.则(1.6.1a)是可写成，

$$T[x(t),y(t)]=\int_{t_0}^{t_1}\frac{\sqrt{\dot{x}^2+\dot{y}^2}}{v(x,y)}\mathrm{d}t \tag{1.6.1b}$$

在经典力学中变分原理主要是指**哈密顿最小作用原理**，简称**哈密顿原理**.

哈密顿原理是指：一个力学系统，如果将系统的动能 T 和势能 U 相减构成系统的**拉格朗日量**，也简称**拉氏量**，

$$L=T-U \tag{1.6.2}$$

其中动能和势能可以是广义坐标和广义动量的函数，而广义坐标和广义动量是随时间而变化的，那么构造如下的泛函：

$$J=\int_{t_0}^{t_1}L\ \mathrm{d}t \tag{1.6.3}$$

系统在一段时间内的一切可能的(即与约束相容的)运动中，实现的是这样一种运动，它使泛函 J 的变分为零

$$\delta J=0 \tag{1.6.4}$$

决定.也就是泛函到达极值.在一般常遇到的力学问题中，这极值总是最小值，所以也称之为最小作用原理.式(1.6.3)的 J 也称为**作用量**.

下面以保守力场为例，说明哈密顿原理.

设已给质点系的第 i 个质点的质量为 m_i，坐标为 $(x_i,y_i,z_i),i=1,2,3\cdots,n.$ 又设在质点 m_i 上作用着以 $U(t,x_1,y_1,z_1,\cdots\cdots,x_n,y_n,z_n)$ 为势函数的力 $\boldsymbol{F}_i(i=1,2,\cdots n)$：

$$F_{i,x}=-\frac{\partial U}{\partial x_i},F_{i,y}=-\frac{\partial U}{\partial y_i},F_{i,z}=-\frac{\partial U}{\partial z_i}. \tag{1.6.5}$$

我们知道，质点系的运动是用牛顿方程组描写的，即各质点的运动曲线方程 $x=x_i(t),y=y_i(t),z=z_i(t)$ 满足

$$\begin{cases} m_i\ddot{x}_i = F_{i,x} = -\dfrac{\partial U}{\partial x_i}, \\ m_i\ddot{y}_i = F_{i,y} = -\dfrac{\partial U}{\partial y_i}, \\ m_i\ddot{z}_i = F_{i,z} = -\dfrac{\partial U}{\partial z_i}. \end{cases} \tag{1.6.6}$$

质点系的位能为 U，动能为

$$T = \frac{1}{2}\sum_{i=1}^{n} m_i(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2) = T(\dot{x}_1, \dot{y}_1, \dot{z}_1, \cdots, \dot{x}_n, \dot{y}_n, \dot{z}_n)$$

假定质点系的起始位置 $x_1(t_0)$，$y_1(t_0)$，$z_1(t_0)$，$\cdots$，$z_n(t_0)$ 和终了位置 $x_1(t_1),\cdots,z_n(t_1)$ 固定，考虑泛函

$$J_1 = \int_{t_0}^{t_1} T(\dot{x}_1,\cdots\dot{z}_n)\mathrm{d}t$$

和

$$J_2 = \int_{t_0}^{t_1} U(t_1, x_1,\cdots,z_n)\mathrm{d}t$$

计算这两个泛函的变分：

$$\delta J_1 = \int_{t_0}^{t_1}\left(\frac{\partial T}{\partial \dot{x}_1}\delta\dot{x}_1 + \cdots + \frac{\partial T}{\partial \dot{z}_n}\delta\dot{z}_n\right)\mathrm{d}t$$

把每项分部积分，并注意

$$\delta x_1(t_0) = \delta x_1(t_1) = \cdots = \delta z_n(t_0) = \delta z_n(t_1) = 0$$

可得

$$\delta J_1 = -\int_{t_0}^{t}\left[\left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial \dot{x}_1}\right)\delta x_1 + \cdots + \left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial \dot{z}_n}\right)\delta z_n\right]\mathrm{d}t$$

$$= -\int_{t_0}^{t}(m_1\ddot{x}_1\delta x_1 + \cdots + m_n\ddot{z}_n\delta z_n)\mathrm{d}t$$

其次

$$\delta J_2 = \int_{t_0}^{t}(\frac{\partial U}{\partial x_1}\delta x_1 + \cdots + \frac{\partial U}{\partial z_n}\delta z_n)\mathrm{d}t$$

由牛顿方程组(1.6.6)，可得

$$\delta J_1 = \delta J_2$$

即

$$\delta(J_1 - J_2) = 0$$

由上讨论可见：如果函数 $x_1(t), y_1(t), z_1(t),\cdots, x_n(t), y_n(t), z_n(t)$，在时间 $t_0 \le t \le t_1$ 内

描述质点系的运动，那末在这些函数处，泛函

$$J_1 - J_2 = \int_{t_0}^{t_1} (T-U)\mathrm{d}t = \int_{t_0}^{t_1} L\mathrm{d}t \qquad (1.6.7)$$

取极值.这就是说，上述质点系是遵循哈密顿原理的.

反之，一旦把系统的泛函(1.6.7)写出来之后，使这一泛函取极值时的相应的欧拉-拉格朗日方程组就是牛顿第二定律的运动方程(1.6.6)式.在前几节中，我们一直称之为欧拉方程的方程组，在物理学中应用时，也常称为**欧拉-拉格朗日方程组**，甚至更有简称为**拉格朗日方程组**的.

动能和势能一般是广义坐标和广义动量的函数.如果有 n 对广义坐标(用 $\{q_i(t)\}$ 表示)和广义动量(用 $\{\dot{q}_i(t)\}$ 表示)，那么根据 1.3.1 小节对于多个函数的情况的讨论知，此时的拉格朗日方程组可按(1.3.8)式写成

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, (i=1,2,\cdots,n) \qquad (1.6.8)$$

动能是不显含时间的，而且一般说来，它是广义动量的二次函数.假如势能 U 是不依赖于时间 t 的，那么 L 就不显含 t，根据 1.2.2 小节中对于不含自变量的特例的讨论知，有一个首次积分.这个首次积分的常数 E 就是体系的总能量，它等于动能加势能

$$E = T + U$$

因此，势能不显含时间意味着系统的总能量守恒.如果势能不显含某一广义坐标 q_i，那么由(1.6.8)式知，拉格朗日函数对于相应的广义动量的求导是一个常数，

$$\frac{\partial L}{\partial \dot{q}_i} = P_i \qquad (1.6.9)$$

这个常数 P_i 是与广义动量成正比的.因此势能不显含坐标 q_i 意味着系统相对于这个坐标的动量守恒.

现在假定上述质点系的坐标还受到 m 个约束条件

$$\varphi_j(t, x_1, y_1, z_1, \cdots, x_n, y_n, z_n) = 0, \ j = 1,2,\cdots,m \qquad (1.6.10)$$

利用哈密顿原理，很容易导出这个质点系的运动方程.

事实上，由哈密顿原理，各质点的运动曲线

$$x_i = x_i(t), y_i = y_i(t), z_i = z_i(t)$$

应使泛函

$$J = \int_{t_0}^{t_1} (T-U)\mathrm{d}t = \int_{t_0}^{t_1} \left[\frac{1}{2}\sum_{i=1}^{n} m_i(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2) - U \right]\mathrm{d}t$$

取极值，这是一个条件极值的变分问题.作辅助泛函

$$J* = \int_{t_0}^{t_1} \left[\frac{1}{2}\sum_{i=1}^{n} m_i(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2) - U + \sum_{j=1}^{m} \lambda_j(t)\varphi_j \right]\mathrm{d}x$$

这个泛函的欧拉方程

$$\left.\begin{aligned}
m_i\ddot{x}_i &= -\frac{\partial U}{\partial x_i}+\sum_{j=1}^{m}\lambda_j(t)\frac{\partial \varphi_j}{\partial x_i}\\
m_i\ddot{y}_i &= -\frac{\partial U}{\partial y_i}+\sum_{j=1}^{m}\lambda_j(t)\frac{\partial \varphi_j}{\partial y_i}\\
m_i\ddot{z}_i &= -\frac{\partial U}{\partial z_i}+\sum_{j=1}^{m}\lambda_j(t)\frac{\partial \varphi_j}{\partial z_i}
\end{aligned}\right\}(i=1,2\cdots n)$$

就是质点系的运动微分方程.

现在将哈密顿原理推广到一般的系统.如果能够写出一个系统的拉格朗日量，那么作用量就由(1.6.3)构造.一般将作用量写成 S：

$$S=\int_{t_0}^{t_1}L\mathrm{d}t \tag{1.6.11}$$

系统的运动方程就由作用量的变分为零 $\delta S=0$ 决定.由于通常情况下， $\delta S=0$ 使得作用量 S 达到极小值，因此也常称为最小作用量原理.(1.6.11)式中的 L 也常被称为拉格朗日密度.

在一般的系统中，拉格朗日量有可能不是动能减势能.后面我们会给出相对论力学系统的一个例子.

例 1 一个粒子在一个半径为 R 的球面，受到重力的作用，求其运动方程.

解答 由于是在球面上运动的，因此采用球坐标 (r,θ,φ) 是比较方便的.以球心为原点，可写出粒子的运动动能为

$$T=\frac{1}{2}m(\dot{r}^2+r^2\dot{\theta}^2+r^2\sin^2\theta\dot{\varphi}^2)$$

它的势能为

$$U=mgz=mgr\cos\theta$$

这样，就写出拉格朗日量

$$L=\frac{1}{2}m(\dot{r}^2+r^2\dot{\theta}^2+r^2\sin^2\theta\dot{\varphi}^2)-mgr\cos\theta \tag{1.6.12}$$

由于要求粒子只能在球面上运动，即 $r=R$，这就给出了一个约束条件

$$G(t,r,\theta,\varphi)=r-R=0$$

此例是三个函数(三个坐标)，一个自变量(时间).

根据§1.4节的介绍，现在实际上是要求泛函

$$\int_{t_0}^{t_1}F(t,r,\theta,\varphi)\mathrm{d}t$$

的极值问题.其中

$$F(t,r,\theta,\varphi)=L+\lambda(t)G$$

$$=\frac{1}{2}m\ (\dot{r}^2+r^2\dot{\theta}^2+r^2\sin^2\theta\dot{\varphi}^2)-mgr\cos\theta+\lambda(t)(r-R)$$

注意，类似于测地线问题，此处的λ可能是时间 t 的函数.

运用欧拉-拉格朗日方程，即仿照(1.4.29)式，写出各微分方程.对于r，有，

$$m\ddot{r}-(mr\dot{\theta}^2+mr\sin^2\theta\dot{\varphi}^2)+mg\cos\theta-\lambda(t)=0 \qquad (1.6.13a)$$

对于θ，有

$$mr^2\ddot{\theta}-mr^2\sin\theta\cos\theta\dot{\varphi}^2-mgr\sin\theta=0 \qquad (1.6.13b)$$

函数F中不含角度φ，因此有一个首次积分.仿照(1.2.28)式写出首次积分为

$$mr^2\sin^2\theta\dot{\varphi}=p_\varphi \qquad (1.6.13c)$$

我们用p_φ来表示这个首次积分的常数，它具有角动量的量纲，它正是与广义坐标φ对应的广义动量.约束条件给出了第四个方程.

$$r=R \qquad (1.6.13d)$$

现在(1.6.13)的四个等式可以用来求解$r(t),\theta(t),\varphi(t),\lambda(t)$四个函数.

实际上，(1.6.13d)已经明确给出了$r(t)$的运动方程.由(1.6.13c)，

$$\dot{\varphi}=\frac{p_\varphi}{mR^2\sin^2\theta}$$

函数$\varphi(t)$的表达式有赖于$\theta(t)$的表达式.把(1.6.13c)代入(1.6.13b)式，得到

$$mR^2\ddot{\theta}=\frac{{p_\varphi}^2}{mR^2\sin^3\theta}\cos\theta+mgR\sin\theta$$

此式两边乘以$\dot{\theta}\mathrm{d}t$，即可进行积分，得到

$$\frac{1}{2}mR^2\dot{\theta}^2=-\frac{{p_\varphi}^2}{2mR^2\sin^2\theta}-mgR\cos\theta+C_2 \qquad (1.6.14)$$

如果写成以下形式，

$$\frac{1}{2}mR^2\dot{\theta}^2+\frac{{p_\varphi}^2}{2mR^2\sin^2\theta}+mgR\cos\theta=C_2$$

则容易看到，等式左边分别是两个角向的动能和势能，因此积分常数C_2实际上就是粒子的总能量.事实上，拉格朗日函数(1.6.12)是不显含时间的，因此这个系统的总能量是守恒的.

函数$\theta(t)$的形式不容易求得解析解.由(1.6.13a)式，

$$\lambda(t)=-mR\dot{\theta}^2-mR\sin^2\theta\dot{\varphi}^2+mg\cos\theta \qquad (1.6.15)$$

此式的前两项分别是两个角向运动的离心力的负值，第三项是重力的负号.把(1.6.14)与(1.6.13c)式代入(1.6.15)，得到

$$\lambda(t)=\frac{{p_\varphi}^2}{mR^3\sin^2\theta}+2mg\cos\theta-2\frac{C_2}{R}-\frac{{p_\varphi}^2}{mR^3\sin^2\theta}+mg\cos\theta$$
$$=\frac{1}{R}(3mgR\cos\theta-2C_2)$$

本例的两个积分常数确实表现出系统总能量守恒和一个广义动量守恒.一般的物理系统中的积分常数都是有着明确的物理意义的.

§1.7 变分法在物理学中的应用

如上一节所述，变分原理是推导系统的状态或者运动应该满足的方程的一个重要方法.原则上，任何物理学中的运动方程都可以在合理的物理前提下，利用变分法推得，尽管实际上有许多方程是按照变分原理以外的思路得到的.

物理学中，一个系统的物理量随时间和空间的分布一般由一组函数$\phi_\alpha(r,t)$来描写.其中 r 表示空间位置的变量，t 表示时间.下标α的取值范围是由系统的内部性质决定.例如，对于电磁场，α是指电磁场的三个空间坐标的分量.对于量子力学的单粒子系统，α是本征状态的标志，一般称为量子数.例如，固体中电子的系统，标志状态的量子数常采用波矢 $\boldsymbol{k}$.函数$\phi_\alpha(r,t)$满足一组偏微分方程组.可以把这样的函数统称为场，它们满足的微分方程称为场方程.

要得到场方程，首先要选取合适的场的拉格朗日函数密度 L.在经典力学的范围内，根据分析力学，它应该是动能减去势能：$L=T-U$.不过，只有用空间坐标和动量作为自变量时，写出来的才是真正的动能和势能.采用广义坐标和广义动量时，写出来的是广义的动能和势能.例如电磁场的情况.在非经典力学的系统中，例如相对论和量子力学中，拉格朗日量不是简单的动能减势能.

一般说来，拉格朗日函数密度具有

$$L=L(\{x_i\},\{\phi_\alpha(\{x_i\})\},\{\partial_i\phi_\alpha(\{x_i\})\}) \tag{1.7.1}$$

的形式.此处用英文字母 i 来标记自变量，α标记函数.可以有 N 个变量和 M 个函数.式(1.7.1)中简写$\partial_i=\dfrac{\partial}{\partial x_i}$.

构造作用量

$$S=\int_R L\mathrm{d}^N x \tag{1.7.2}$$

之后，最小作用原理要求其变分为零.这就是 1.3 节介绍的多变量多函数的变分问题.可按照(1.3.16)写出欧拉-拉格朗日方程：

$$\frac{\partial L}{\partial\phi_\alpha}-\sum_{i=1}^{N}\frac{\partial}{\partial x_i}\frac{\partial L}{\partial_i\phi_\alpha}=0,(\alpha=1,2,\cdots,M) \tag{1.7.3}$$

这也就是物理量所满足的场方程.

写出一个拉氏量的时候，要明确其中的函数$\{\phi_\alpha\}$是什么，自变量$\{x_i\}$是什么.

前几节由变分法得到的都是微分方程.一般情况下，由变分法得到的方程也可能是积分方程，或者积分微分方程.这完全取决于拉氏量的构造.

需要说明的是，满足变分原理的拉格朗日函数密度并不唯一.

例如，设(1.7.2)式的积分区域固定，在区域边界上函数$\{\phi_\alpha\}$的数值是确定的，如果 L 再附加一项$\sum_{j=1}^{N}\partial_j\Gamma_j(\{x_i\},\{\phi_\alpha(\{x_i\})\})$，其中$\Gamma_j$可以是任意的可导函数，则由最小作用原理，得到

$$\delta\int_R[L+\sum_{j=1}^{N}\partial_j\Gamma_j(\{x_i\},\{\phi_\alpha(\{x_i\})\})]\mathrm{d}^N x=\delta\int_R L\mathrm{d}^N x+\delta\oint\sum_{j=1}^{N}\Gamma_j\mathrm{d}l_j$$

其中后一项是在区域边界上的线积分，积分结果取决于函数$\{\phi_\alpha\}$在边界上的数值.而这数值已知是确定的，因此第二项的变分总是为零.即

$$\delta\int_R[L+\sum_{j=1}^{N}\partial_j\Gamma_j)]\mathrm{d}^N x=\delta\int_R L\mathrm{d}^N x \tag{1.7.4}$$

所以 L 和 $L+\sum_{j=1}^{N}\partial_j\Gamma_j$ 是等效的，描述同一物理体系.

变分原理没有告诉我们如何选取拉氏密度量，拉氏密度量的选取是依据现有的物理方程为基础，带有很大的“猜”的成分，检验正确与否的标准就是其导出的方程与实际的物理方程是否一致，或者，由导出的方程计算的理论结果是否与实验一致.量子力学的薛定谔方程的导出就是属于后一种情况.

1.7.1 在经典物理中的应用

1. 麦克斯韦电磁场方程组

由电磁学已有的方程，电磁场的拉氏量可以写为

$$L=-\frac{1}{4\mu}\sum_{ij}F_{ij}F_{ij}+\sum_i J_i A_i \tag{1.7.5}$$

其中

$$F_{ij}=\frac{\partial A_j}{\partial x_i}-\frac{\partial A_i}{\partial x_j} \tag{1.7.6}$$

为四维电磁场反对称张量的分量.四维向量$\{A_i\}$和$\{J_i\}$由下式定义：

$$\{A_i\}\equiv\{A_1,A_2,A_3,i\varphi/c\}\equiv\{\boldsymbol{A},\mathrm{i}\varphi/c\} \tag{1.7.7a}$$

$$\{J_i\}\equiv\{j_1,j_2,j_3,ic\rho\}\equiv\{\boldsymbol{j},\mathrm{i}c\rho\} \tag{1.7.7b}$$

其中 $\boldsymbol{A}$ 和φ分别是电磁场的矢势和标势，$\boldsymbol{j}$ 和ρ分别是电流密度和电荷密度.四维时空坐标则是

$$\{x_i\}\equiv\{x_1,x_2,x_3,ict\}\equiv\{\boldsymbol{r},\mathrm{i}ct\} \tag{1.7.7c}$$

式(1.7.5)符合洛仑兹变换的不变性.在相对论性系统中，作用量的洛仑兹不变性是寻找系统拉格朗日量的重要依据.

写出了拉氏量，代入拉格朗日方程就可以得到场方程.(1.7.5)式中，代入 $\boldsymbol{E}$ 和 $\boldsymbol{B}$，可以改写为

$$L=\frac{1}{2}\varepsilon\boldsymbol{E}^2-\frac{1}{2\mu}\boldsymbol{B}^2+\boldsymbol{j}\cdot\boldsymbol{A}-\rho\varphi \tag{1.7.8}$$

与经典力学中的 $L=T-U$ 对比，对此形式的拉氏量可以理解为：前两为广义的动能(注意不是电磁场能量，电磁场能量对应于由拉氏量导出的哈密顿量)，后两项为与速度有关的广义势能.现在的拉氏量是四维电磁势$\{A_i\}$的函数.而自变量则是空间坐标$\{x_i\}$.因此这是一个多函数多自变量的变分的问题.要求的是电磁场所满足的方程.

将(1.7.6)代入(1.7.5)，得到

$$\begin{aligned}
L&=-\frac{1}{4\mu}\left(\frac{\partial A_j}{\partial x_i}-\frac{\partial A_i}{\partial x_j}\right)\left(\frac{\partial A_j}{\partial x_i}-\frac{\partial A_i}{\partial x_j}\right)+J_iA_i\\
&=-\frac{1}{4\mu}\frac{\partial A_j}{\partial x_i}\left(\frac{\partial A_j}{\partial x_i}-\frac{\partial A_i}{\partial x_j}\right)+\frac{1}{4\mu}\frac{\partial A_i}{\partial x_j}\left(\frac{\partial A_j}{\partial x_i}-\frac{\partial A_i}{\partial x_j}\right)+J_iA_i\\
&=-\frac{1}{4\mu}\frac{\partial A_j}{\partial x_i}\left(\frac{\partial A_j}{\partial x_i}-\frac{\partial A_i}{\partial x_j}\right)-\frac{1}{4\mu}\frac{\partial A_i}{\partial x_j}\left(\frac{\partial A_i}{\partial x_j}-\frac{\partial A_j}{\partial x_i}\right)+J_iA_i\\
&=-\frac{1}{2\mu}\frac{\partial A_j}{\partial x_i}\left(\frac{\partial A_j}{\partial x_i}-\frac{\partial A_i}{\partial x_j}\right)+J_iA_i
\end{aligned}$$

对于每一项中相同的下标要进行求和。其中，第一个圆括号因子中的第二项把指标 i 和 j 交换之后，结果与第一项相同。然后计算以下求导。

$$\begin{aligned}
\frac{\partial L}{\partial(\partial A_i/\partial x_j)}&=-\frac{1}{2\mu}\frac{\partial}{\partial(\partial A_i/\partial x_j)}\left[\frac{\partial A_l}{\partial x_k}\left(\frac{\partial A_l}{\partial x_k}-\frac{\partial A_k}{\partial x_l}\right)\right]\\
&=-\frac{1}{2\mu}\left[\left(\frac{\partial A_i}{\partial x_j}-\frac{\partial A_j}{\partial x_i}\right)+\left(\frac{\partial A_i}{\partial x_j}-\frac{\partial A_j}{\partial x_i}\right)\right]=-\frac{1}{\mu}\left(\frac{\partial A_i}{\partial x_j}-\frac{\partial A_j}{\partial x_i}\right)
\end{aligned}$$

有了这样的结果，欧拉-拉格朗日方程(1.7.3)就成为

$$\frac{\partial L}{\partial A_i}-\sum_{j=1}^{4}\frac{\partial}{\partial x_j}\frac{\partial L}{\partial(\partial A_i/\partial x_j)}=\frac{\partial L}{\partial A_i}+\frac{1}{\mu}\sum_{j=1}^{4}\frac{\partial}{\partial x_j}\left(\frac{\partial A_i}{\partial x_j}-\frac{\partial A_j}{\partial x_i}\right)=0$$

现在可以来求各广义坐标分量的方程了。对于前三个广义坐标，得

$$j_1+\frac{1}{\mu}[\frac{\partial}{\partial x_2}(\frac{\partial A_1}{\partial x_2}-\frac{\partial A_2}{\partial x_1})+\frac{\partial}{\partial x_3}(\frac{\partial A_1}{\partial x_3}-\frac{\partial A_3}{\partial x_1})+\frac{\partial}{\partial(\mathrm{i}ct)}(\frac{\partial A_1}{\partial(\mathrm{i}ct)}-\frac{\partial(\mathrm{i}\varphi/c)}{\partial x_1})]$$
$$=j_1-\frac{1}{\mu}[\frac{\partial}{\partial x_2}B_z-\frac{\partial}{\partial x_3}B_y-\frac{1}{c^2}\frac{\partial}{\partial t}E_x]$$
$$=j_1-\frac{1}{\mu}[(\nabla\times\boldsymbol{B})_1+\frac{1}{c^2}\frac{\partial E_x}{\partial t}]=0.$$

因此

$$\frac{1}{\mu}(\nabla\times\boldsymbol{B})_1=j_1+\varepsilon\frac{\partial E_x}{\partial t} \tag{1.7.9a}$$

同理，得到

$$\frac{1}{\mu}(\nabla\times\boldsymbol{B})_2=j_2+\varepsilon\frac{\partial E_y}{\partial t} \tag{1.7.9b}$$

$$\frac{1}{\mu}(\nabla\times\boldsymbol{B})_3=j_3+\varepsilon\frac{\partial E_z}{\partial t} \tag{1.7.9c}$$

其中后两式可以从前一式取指标轮换得到，即取指标 $1\to2\to3\to1$ 和 $x\to y\to z\to x$. 以上三式合起来就是

$$\frac{1}{\mu}\nabla\times\boldsymbol{B}=\boldsymbol{J}+\varepsilon\frac{\partial\boldsymbol{E}}{\partial t} \tag{1.7.9}$$

对于第四个广义坐标，得

$$\frac{\partial L}{\partial A_4}-\sum_{j=1}^{4}\frac{\partial}{\partial x_j}\frac{\partial L}{\partial(\partial A_4/\partial x_j)}=-\frac{\partial(\rho\varphi)}{\partial(\mathrm{i}\varphi/c)}$$
$$+\frac{1}{\mu}\left[\frac{\partial}{\partial x_1}(\frac{\partial\mathrm{i}\varphi}{c\partial x_1}-\frac{\partial A_x}{\partial(\mathrm{i}ct)})+\frac{\partial}{\partial x_2}(\frac{\partial\mathrm{i}\varphi}{c\partial x_2}-\frac{\partial A_y}{\partial(\mathrm{i}ct)})+\frac{\partial}{\partial x_3}(\frac{\partial\mathrm{i}\varphi}{c\partial x_3}-\frac{\partial A_z}{\partial(\mathrm{i}ct)})\right]$$
$$=\mathrm{i}c\rho+\frac{1}{\mathrm{i}c\mu}(\frac{\partial E_x}{\partial x_1}+\frac{\partial E_y}{\partial x_2}+\frac{\partial E_z}{\partial x_3})=\mathrm{i}c\rho+\frac{1}{\mathrm{i}c\mu}\nabla\cdot\boldsymbol{E}=0$$

由此得到

$$\varepsilon\nabla\cdot\boldsymbol{E}=\rho \tag{1.7.10}$$

式(1.7.6)本身表示以下两个方程.

$$\boldsymbol{B}=\nabla\times\boldsymbol{A},\quad \boldsymbol{E}=-\nabla\varphi-\frac{\partial\boldsymbol{A}}{\partial t}$$

由此得到

$$\nabla\cdot\boldsymbol{B}=\nabla\cdot(\nabla\times\boldsymbol{A})=0 \tag{1.7.11}$$

$$\nabla\times\boldsymbol{E}=\nabla\times(-\nabla\varphi-\frac{\partial\boldsymbol{A}}{\partial t})=-\nabla\times\nabla\varphi-\frac{\partial}{\partial t}\nabla\times\boldsymbol{A}=-\frac{\partial\boldsymbol{B}}{\partial t} \tag{1.7.12}$$

(1.7.9-12)四式就是我们熟知的麦克斯韦方程组：

$$\begin{cases}\nabla\times\boldsymbol{H}=\boldsymbol{J}+\dfrac{\partial\boldsymbol{D}}{\partial t}\\ \nabla\times\boldsymbol{E}=-\dfrac{\partial\boldsymbol{B}}{\partial t}\\ \nabla\cdot\boldsymbol{B}=0\\ \nabla\cdot\boldsymbol{D}=\rho\end{cases}$$

可见，麦克斯韦电磁场方程组可以从拉氏量(1.7.5)是出发，利用变分法的欧拉-拉格朗日方程组推导出来.

3. 电磁场中的带电粒子

拉格朗日方程

$$\frac{\partial L}{\partial q_\alpha}-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial\dot{q}_\alpha}=0 \tag{1.7.29}$$

现在我们要求的是一个带电粒子的运动，拉氏量中的函数$\{q_\alpha\}$就是粒子的空间坐标$\boldsymbol{r}(t)$，自变量是时间 t.矢势 $\boldsymbol{A}$ 和标势φ也都是$\boldsymbol{r}(t)$的函数.因此属于三个函数和一个自变量的情况.要注意与式(1.7.5)拉氏量的区别.

1) 非相对论情形

拉格朗日量为

$$L=\frac{1}{2}mv^2-e(\varphi-\boldsymbol{v}\cdot\boldsymbol{A}) \tag{1.7.30}$$

运用拉格朗日方程.把空间坐标的三个方程都写出来。

$$\frac{\partial}{\partial x}L-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial v_x}L=0,\frac{\partial}{\partial y}L-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial v_y}L=0,\frac{\partial}{\partial z}L-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial v_z}L=0$$

配上单位矢量之后，写成统一的矢量式。

$$\boldsymbol{i}(\frac{\partial}{\partial x}L-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial v_x}L)=0,\boldsymbol{j}(\frac{\partial}{\partial y}L-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial v_y}L)=0,\boldsymbol{k}(\frac{\partial}{\partial z}L-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial v_z}L)=0$$

$$\nabla L-\frac{\mathrm{d}}{\mathrm{d}t}\nabla_v L=0$$

把拉格朗日量代入

$$\begin{aligned}&(\nabla-\frac{\mathrm{d}}{\mathrm{d}t}\nabla_v)[\frac{1}{2}m\boldsymbol{v}\cdot\boldsymbol{v}-e(\varphi-\boldsymbol{v}\cdot\boldsymbol{A})]\\&=-e\nabla\varphi+e\nabla(\boldsymbol{v}\cdot\boldsymbol{A})-\frac{\mathrm{d}}{\mathrm{d}t}(m\boldsymbol{v}+e\boldsymbol{A})\\&=-\frac{\mathrm{d}}{\mathrm{d}t}m\boldsymbol{v}-e\nabla\varphi+e\nabla(\boldsymbol{v}\cdot\boldsymbol{A})-e\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}\end{aligned}$$

用公式

$$\nabla(\boldsymbol{v}_1\cdot\boldsymbol{v}_2)=\boldsymbol{v}_2\times(\nabla\times\boldsymbol{v}_1)+\boldsymbol{v}_1\times(\nabla\times\boldsymbol{v}_2)+(\boldsymbol{v}_2\cdot\nabla)\boldsymbol{v}_1+(\boldsymbol{v}_1\cdot\nabla)\boldsymbol{v}_2$$
$$\nabla(\boldsymbol{v}\cdot\boldsymbol{A})=\boldsymbol{v}\times(\nabla\times\boldsymbol{A})+(\boldsymbol{v}\cdot\nabla)\boldsymbol{A}$$
$$\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}=\frac{\partial\boldsymbol{A}}{\partial t}+(\boldsymbol{v}\cdot\nabla)\boldsymbol{A}=\frac{\partial\boldsymbol{A}}{\partial t}+\nabla(\boldsymbol{v}\cdot\boldsymbol{A})-\boldsymbol{v}\times(\nabla\times\boldsymbol{A})$$
因此得到
$$(\nabla-\frac{\mathrm{d}}{\mathrm{d}t}\nabla_v)[\frac{1}{2}m\boldsymbol{v}\cdot\boldsymbol{v}-e(\varphi-\boldsymbol{v}\cdot\boldsymbol{A})]$$
$$=-\frac{\mathrm{d}}{\mathrm{d}t}m\boldsymbol{v}-e\nabla\varphi-e\frac{\partial\boldsymbol{A}}{\partial t}+e\boldsymbol{v}\times(\nabla\times\boldsymbol{A})=0$$
再利用其中， $\boldsymbol{E}=-\nabla\varphi-\frac{\partial}{\partial t}\boldsymbol{A}$， $\nabla\times\boldsymbol{A}=\boldsymbol{B}$.最后得到
$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{p}=e(\boldsymbol{E}+\boldsymbol{v}\times\boldsymbol{B})\tag{1.7.31}$$
这就是电磁场中的带电粒子满足的运动方程.

书上是对于每一个分量式分别计算。
$$\frac{\partial}{\partial x}L-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial v_x}L=-e(\frac{\partial}{\partial x}\varphi-\frac{\partial}{\partial x}\boldsymbol{v}\cdot\boldsymbol{A})-\frac{\mathrm{d}}{\mathrm{d}t}mv_x-e\frac{\mathrm{d}}{\mathrm{d}t}A_x=0$$
其中
$$\boldsymbol{v}\cdot\frac{\partial}{\partial x}\boldsymbol{A}-\frac{\mathrm{d}}{\mathrm{d}t}A_x=v_x\frac{\partial}{\partial x}A_x+v_y\frac{\partial}{\partial x}A_y+v_z\frac{\partial}{\partial x}A_z-\frac{\mathrm{d}}{\mathrm{d}t}A_x$$
$$=(\boldsymbol{v}\cdot\nabla)A_x-\frac{\mathrm{d}}{\mathrm{d}t}A_x-v_y\frac{\partial}{\partial y}A_x-v_z\frac{\partial}{\partial z}A_x+v_y\frac{\partial}{\partial x}A_y+v_z\frac{\partial}{\partial x}A_z$$
$$=-\frac{\partial}{\partial t}A_x+v_y(\frac{\partial}{\partial x}A_y-\frac{\partial}{\partial y}A_x)-v_z(\frac{\partial}{\partial z}A_x-\frac{\partial}{\partial x}A_z)$$
$$=-\frac{\partial}{\partial t}A_x+v_yB_z-v_zB_y=-\frac{\partial}{\partial t}A_x+(\boldsymbol{v}\times\boldsymbol{B})_x$$
2) 相对论情形

拉格朗日量为
$$L=-mc^2\sqrt{1-v^2/c^2}-e(\varphi-\boldsymbol{v}\cdot\boldsymbol{A})\tag{1.7.32}$$
运用拉格朗日方程(1.7.29)，得到的运动方程也是(1.7.31)式的形式，只是现在的动量是相对论动量的表达式.注意(1.7.32)的第一项不是粒子的动能，也不是粒子总能量，还是一个负值.此时的拉格朗日量不是动能减势能.

1.7.2 在量子力学中的应用

1. 薛定谔方程

薛定谔的最初的论文确实是利用变分法得到了量子力学的基本方程.该方程就以他的名字命名.

分析力学中有哈密顿-雅可比微分方程

$$H(q,\frac{\partial S}{\partial q})=E \tag{1.7.33}$$

哈密顿量就是能量的数值.S 是作用量.设一个未知量ψ，作用量可以用ψ通过以下方式来表示：

$$S=K\ln\psi\ , \tag{1.7.34}$$

即$\psi=\exp(S/K)$. 显然，ψ应该是一个无量纲的量. K 是个常数，它应具有作用量的量纲.现在，要让ψ函数负担的物理上的功能是：它描述一个微观粒子的运动状态或者存在的方式.相应地，对这个函数数学上的要求是，它在整个空间内是单值、有限、连续、有一阶和二阶导数.那么(1.7.34)写成

$$H(q,\frac{K}{\psi}\frac{\partial\psi}{\partial q})=E \tag{1.7.35}$$

薛定谔是这样考虑的：哈密顿量是动能加上势能，$H=T+U$.势能是坐标的函数.动能部分是ψ及其导数的函数.并设动能是ψ及其导数的二次函数.此处有一个观念上的改变.**在经典力学中，动能是与广义坐标对时间的导数有关.而现在动能则是和一个函数对于空间坐标的导数有关！**这样，在直角坐标系中，(1.7.35)式写成如下形式.

$$\frac{K^2}{2m\psi^2}[(\frac{\partial\psi}{\partial x})^2+(\frac{\partial\psi}{\partial y})^2+(\frac{\partial\psi}{\partial z})^2]+V(\boldsymbol{r})=E$$

或者

$$\frac{K^2}{2m}[(\frac{\partial\psi}{\partial x})^2+(\frac{\partial\psi}{\partial y})^2+(\frac{\partial\psi}{\partial z})^2]-(E-V(\boldsymbol{r}))\psi^2=0 \tag{1.7.36}$$

写出一个$1/2m$ 的因子是为了接近经典力学的动能的形式.式(1.7.36)看上去可以直接求解函数ψ，但由于这是关于ψ的二次的方程，解出来的函数ψ可能不是唯一的.为了得到可以求出唯一解的方程，借助于变分法.

把式(1.7.36)左边的量不看成是零，而是一个函数，它就是一个微观粒子的拉格朗日量.构造泛函

$$J=\iiint \mathrm{d}x\mathrm{d}y\mathrm{d}z[(\frac{\partial\psi}{\partial x})^2+(\frac{\partial\psi}{\partial y})^2+(\frac{\partial\psi}{\partial z})^2-\frac{2m}{K^2}(E-V(\boldsymbol{r}))\psi^2] \tag{1.7.37}$$

令ψ变化，求此泛函的极值.拉氏量中，有一个函数ψ和三个自变量(x,y,z) .现在的变分问题是：

$$\delta J=\delta\iiint \mathrm{d}x\mathrm{d}y\mathrm{d}z[(\frac{\partial\psi}{\partial x})^2+(\frac{\partial\psi}{\partial y})^2+(\frac{\partial\psi}{\partial z})^2-\frac{2m}{K^2}(E-V(\boldsymbol{r}))\psi^2]=0$$

注意，其中势能$V(\boldsymbol{r})$ 虽然也是空间坐标的函数，但是它的数值在各点是固定的，

而不是可变的.因此对于$V(\boldsymbol{r})$不施行变分.根据变分的步骤运作，

$$\delta J=2\iiint \mathrm{d}x\mathrm{d}y\mathrm{d}z[\psi_x\delta\psi_x+\psi_y\delta\psi_y+\psi_z\delta\psi_z-\frac{2m}{K^2}(E-V(\boldsymbol{r}))\psi\delta\psi]=0$$

对动能项做分部积分.

$$\delta J=\int\delta\psi\frac{\partial\psi}{\partial n}\mathrm{d}s+\iiint \mathrm{d}x\mathrm{d}y\mathrm{d}z[-\psi_{xx}-\psi_{yy}-\psi_{zz}-\frac{2m}{K^2}(E-V(\boldsymbol{r}))\psi]\delta\psi=0$$

变分的结果是

$$-\nabla^2\psi-\frac{2m}{K^2}(E-V(\boldsymbol{r}))\psi=0 \tag{1.7.38}$$

和

$$\int\delta\psi\frac{\partial\psi}{\partial n}\mathrm{d}s=0 \tag{1.7.39}$$

其中$\mathrm{d}s$是无限接近被积区域表面的面积元.

式(1.7.38)就是函数ψ应该满足的方程.它与式(1.7.36)比较，ψ的平方项不出现了.然后，薛定谔把氢原子的势能代入(1.7.38)式，采用球坐标系求解微分方程得到能量的本征值，与实验结果的巴耳末公式对照，得到常数 K 恰好就是普朗克常数.这说明方程(1.7.38)是正确的.该方程也就被命名为定态薛定谔方程.这个方程的出现，标志着量子力学的诞生！

函数ψ被称为一个微观粒子的**波函数**.薛定谔方程描述的是单个微观粒子的情况.

最后我们简单提一下另外几个利用变分方法来求出系统应满足的微分方程的例子.

计算固体中电子结构的密度泛函方法的孔恩-沈吕九方程也是用变分法得到的.其中的能量泛函都只写成电子密度的函数，而与电子密度的梯度无关.这个假设对于电子密度不太高的情况是合适的.当电子密度比较高时，能量泛函就应该与电子密度的梯度有关的.这时把对于密度函数的一阶导数也要考虑进去.

在超导体中，当温度接近超导转变温度的时候，可以定义一个序参量ψ，它具有超导电子波函数的物理意义.金兹堡和朗道写出了此时在电磁场中的超导体的自由能.然后根据变分原理推导出了序参量所应该满足的金兹堡-朗道方程组.

希尔伯特用变分法推出了爱因斯坦的宇宙学方程.当然这是在爱因斯坦得到了该方程以后的事情.

本章重点

最简单情形的欧拉方程(1.2.22).

多个函数的欧拉方程组(1.3.8)。两个函数就是(1.3.5)。

多个自变量的欧拉方程(1.3.14)。两个自变量就是(1.3.12)。

等周条件的欧拉方程(1.4.16)。

测地线问题的欧拉方程(1.4.25)。

轻点：

变分学的两个基本引理。

自然边界条件：在单边情况下是欧拉方程加上条件(1.5.6)。

变分原理的基本思想。

小贴士

变分法是至关重要的数学方法。没有变分法，现代理论物理无法进行。

在基本粒子物理和广义相对论中，现在的研究方法都是首先写出一个系统的拉格朗日量，然后由变分法求出系统的运动方程和其它的一些量。也就是说，拉格朗日量可以看做是系统的最基本的量。其它的内容都可以有拉格朗日量导出。例如，可以由拉氏量导出哈密顿量。写出这个基本的拉氏量有猜测的成分，主要是根据系统的物理表现写出有关的项。当然，一定要跟已有的正确的理论融洽，并得到实验的验证。

本章的内容中，泛函的概念将在第五章中再次提到。其它的内容后面不再涉及。所以是比较独立的一章内容。

习题

2. 求以下泛函的全增量和变分.

(1) $J[y]=\int_a^b \sqrt{1+y'^2}\,\mathrm{d}x$.

(2) $J[y]=\int_{x_0}^{x_1} (x^2 y+y^2+yy')\,\mathrm{d}x$.

(3) $J[y]=\int_0^1 y^3 y'^2\,\mathrm{d}x$,条件是量端点固定.

3. 设两端点固定，求下列泛函的极值曲线：

(1) $J[y]=\int_{x_0}^{x_1} \sqrt{y(1+y'^2)}\,\mathrm{d}x$.

(2) $J[y]=\int_{x_0}^{x_1} \dfrac{1+y^2}{y'^2}\,\mathrm{d}x$.

4. 求以下泛函的带端点条件的极值曲线，并通过考察全增量判断所得曲线使泛函 J 取极小值.

(1) $J[y]=\int_0^2 (xy'+y'^2)\,\mathrm{d}x$， $y(0)=1, y(2)=0$.

(2) $J[y]=\int_0^1 (12xy+yy'+y'^2)\,\mathrm{d}x, y(0)=1, y(1)=4.$

6. 求下列泛函的极值曲线所应满足的微分方程.并求得解.

(1) $J[y]=\int_{x_0}^{x_1} (16y^2-y''^2+x^2)\,\mathrm{d}x$.

(2) $J[y]=\int_{x_0}^{x_1} (y'''^2+2xy)\,\mathrm{d}x$.

7. 求下列泛函的极值曲线.

(1) $J[y,z]=\int_{x_0}^{x_1}(2yz-2y^2+y'^2-z'^2)\mathrm{d}x$.

(2) $J[y,z]=\int_{x_0}^{x_1}(y'^2+z'^2+y'z')\mathrm{d}x$.

8. 写出下列泛函的奥氏方程：

(1) $J[z(x,y)]=\iint\limits_D\left[\left(\frac{\partial z}{\partial x}\right)^2-\left(\frac{\partial z}{\partial y}\right)^2\right]\mathrm{d}x\mathrm{d}y$.

(2) $J[u(x,y,z)]=\iiint\limits_V\left[\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2+\left(\frac{\partial u}{\partial z}\right)^2+2uf(x,y,z)\right]\mathrm{d}x\mathrm{d}y\mathrm{d}z$.

10. 求下列等周问题的极值曲线.

(1) 泛函 $J[y]=\int_0^1(y'^2+x^2)\mathrm{d}x, y(0)=0, y(1)=0$ 等周条件：$\int_0^1 y^2\mathrm{d}x=2$.

(3) $J[y,z]=\int_0^1(y'^2+z'^2-4xz'-4z)\mathrm{d}x, y(0)=z(0)=0; y(1)=z(1)=1;$
$\int_0^1(y'^2-xy'-z'^2)\mathrm{d}x=2.$

(提示：含多个函数的泛函的等周问题的解法，与一个未知函数的情形相同.)

11. 求圆柱面 $r=R$ 上的短程线.

(提示：用柱面坐标 r,φ,z 求解较便利.)

12. 写出在条件 $\int_0^{x_1}r(x)y^2\mathrm{d}x=1; y(0)=0, y(x_1)=0$ 下，泛函 $J[y]=\int_0^{x_1}[p(x)y'^2+q(x)y^2]\mathrm{d}x$ 的极值曲线所应满足的微分方程. 其中 $r(x), p(x), q(x)$ 是三个固定的已知函数.

13. 设函数 $y(x)$ 的左端固定在原点，即 $y(0)=0$，另一边界点在直线 $x=\pi/4$ 上滑动.试求泛函 $J[y]=\int_0^{\pi/4}(y^2-y'^2)\mathrm{d}x$ 的极值曲线.

15. 给定空间某点坐标 (x_0,y_0)，例如 $(0,1)$，画出通过该点的各中参量数值的悬链线，这些线都相切于一根包络线.

16. 求泛函 $J[y]=\int_0^1(y'^2-yy'+y^2)\mathrm{d}x$ 满足下列边界条件的极值曲线.

(1) $y(x)$ 过点 $P_1(0,1)$，在 $x=1$ 处满足自然边界条件.

(2) $y(x)$ 过点 $P_2(1,2)$，在 $x=0$ 处满足自然边界条件.

(3) $y(x)$ 在 $x=0$ 和 $x=1$ 都满足自然边界条件.

17. 证明：(1.7.5)式的拉格朗日量可用电磁场来表达，即写成(1.7.8)式.
18. 从(1.7.32)式的拉格朗日量，得到粒子的运动方程(1.7.31)式.
19. 证明：如果取ψ和ψ^*作为独立的场变量，且场的实拉格朗日密度为

$$L=\frac{\hbar^2}{2m}\nabla\psi\cdot\nabla\psi^*+U(r,t)\psi\psi^*-\frac{i\hbar}{2}(\dot{\psi}\psi^*-\psi\dot{\psi}^*)$$

其中∇为梯度算符，$\dot{\psi}$上的一点表示对时间求偏导，$\dot{\psi}=\frac{\partial\psi}{\partial t}$.将这个拉格朗日密度代入欧拉-拉格朗日方程之后，导致一个什么方程?

20. 设F不显含y，即$F=F(x,y',z,z')$是否有首次积分？设F不显含y和z，即$F=F(x,y',z')$是否有首次积分？设F不显含x，即$F=F(y,y',z,z')$是否有首次积分？每一种情况，若有首次积分，首次积分的表达式是什么？

(扩展题：设F不显含x，即$F=F(y,u,u_x,u_y)$是否有首次积分？设F不显含x和y，即$F=F(u,u_x,u_y)$是否有首次积分？设F不显含u，即$F=F(x,y,u_x,u_y)$是否有首次积分？每一种情况，若有首次积分，首次积分的表达式是什么？可查书籍文献。)

附录 1A 函数的极值问题

泛函变分的公式和以下函数求导的公式有着对应关系.

一元函数$f(x)$在$x=x_0$处取极值的必要条件是

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x}\Big|_{x=x_0}=0 \tag{1A.1}$$

先根据此式计算得到x_0.再由此点的二阶导数是大于零、小于零或者等于零来判断此点处函数是取极小值、极大值还是拐点.

对于多元函数，就要考虑**无约束的极值问题**和**有约束的极值问题**.

(1) 无约束的极值问题

二元函数$f(x,y)$在$x=x_0,y=y_0$处取极值的必要条件是

$$\frac{\partial f(x,y)}{\partial x}\Big|_{x=x_0,y=y_0}=0,\frac{\partial f(x,y)}{\partial y}\Big|_{x=x_0,y=y_0}=0 \tag{1A.2}$$

先根据此式计算得到(x_0,y_0).再由此点的二阶导数来判断此点处函数是否取到极值.

对于n元函数$f(x_1,x_2,\cdots,x_n)$在$x_i=x_{i,0},(i=1,2,\cdots,n)$点处取极值的必要条件

是

$$\frac{\partial f(x_1, x_2, \cdots, x_n)}{\partial x_i}\bigg|_{x_i = x_{i,0}, (i=1,2,\cdots,n)} = 0, (i = 1, 2, \cdots, n) \qquad (1A.3)$$

(2) 有约束的极值问题

有约束的极值问题也称为**条件极值问题.**

设有一个二元函数

$$u = u(x, y) \qquad (1A.4)$$

限制要求其在给定条件

$$g(x, y) = 0 \qquad (1A.5)$$

时的极值问题.此时，构造一个新的函数

$$L = u + \lambda g \qquad (1A.6)$$

其中λ是一个待定的参量.如果取极值时 $\lambda = \lambda_0$，那么，在给定条件下函数 u 在 (x_0, y_0, λ_0) 处取得极值的必要条件是函数 L 对 x 和 y 的一阶导数

$$\frac{\partial L}{\partial x}\bigg|_{x=x_0, y=y_0, \lambda=\lambda_0} = 0, \frac{\partial L}{\partial y}\bigg|_{x=x_0, y=y_0, \lambda=\lambda_0} = 0 \qquad (1A.7)$$

由此式和

$$g(x_0, y_0) = 0 \qquad (1A.8)$$

约束条件一起，有三个方程，可以解出极值点 (x_0, y_0) 和参数λ的数值 $\lambda = \lambda_0$.

这个求解过程也可以从另一个角度来看.把原来的约束条件下求极值的问题看作为就是要对三个自变量 (x, y, λ) 的函数 L 求极值.于是，根据前面多元函数求极值的必要条件，应该有：数 L 对 (x, y, λ) 三个自变量的一阶偏导数在 (x_0, y_0, λ_0) 处为零.前两个条件就是(A.7)，后一个条件就是(A.8)式.

函数 L 常称为**拉格朗日函数**，参量λ称为**拉格朗日乘数**或**拉格朗日乘子**.这一方法称为**拉格朗日乘数法**.

一般地，如果要求一个 n 元函数 $f(x_1, x_2, \cdots, x_n)$ 在 m 个约束条件

$$g_j(x_1, x_2, \cdots, x_n), (j = 1, 2, \cdots, m) \qquad (1A.9)$$

下在 $x_i = x_{i,0}, (i = 1, 2, \cdots, n)$ 点处取极值，那么构造一个带有 m 个拉格朗日乘数 $\{\lambda_j\}$ 的函数 L 如下

$$L = f(x_1, x_2, \cdots, x_n) + \sum_{j=1}^{m} \lambda_j g_j(x_1, x_2, \cdots, x_n) \qquad (1A.10)$$

取极值的必要条件是：函数 L 对 n 个自变量的一阶偏导数在自变量为 $x_{i,0},(i=1,2,\cdots,n)$ 处和参量为 $\lambda_j=\lambda_{j,0},(j=1,2,\cdots,m)$ 处的值应为零.

$$\frac{\partial L}{\partial x_i}\Big|_{x_i=x_{i,0},(i=1,2,\cdots,n);\lambda_j=\lambda_{j,0},(j=1,2,\cdots,m)}=0,(i=1,2,\cdots,n) \tag{1A.11a}$$

或者，更确切地，写成

$$\left[\frac{\partial f(x_1,x_2,\cdots,x_n)}{\partial x_i}+\sum_{j=1}^{m}\lambda_j\frac{\partial g_j(x_1,x_2,\cdots,x_n)}{\partial x_i}\right]_{x_i=x_{i,0},(i=1,2,\cdots,n);\lambda_j=\lambda_{j,0},(j=1,2,\cdots,m)}=0,\quad (i=1,2,\cdots,n) \tag{1A.11b}$$

式(1A.11)有 n 个方程，加上(1A.9)的 m 个方程，决定了 n 个数 $x_{i,0},(i=1,2,\cdots,n)$ 和 m 个参数 $\lambda_{j,0},(j=1,2,\cdots,m)$ 的数值.