变分学总结

tron

2023 年 1 月 20 日

目录

1 经典理论

简单地说, 变分学是研究泛函极值 (以及更一般的临界值的一个数学分支). 设 $\Omega \subseteq \mathbb{R}^n$ 是一个有界开集. 给定函数 $L \in C^1(\overline{\Omega} \times \mathbb{R}^N \times \mathbb{R}^{nN})$ (通常称 L 为 Lagrange 函数, 即 Lagrangian), 变分学主要研究如下形式的泛函:

$$I(u) = \int_\Omega L(x, u(x), \dot{u}(x))\, \mathrm{d}x, \tag{1}$$

其中 $u \in M$, M 是一个函数集合, 例如 $C^1(\overline{\Omega})$ 等, 同时 M 由一些边值条件所规定. 有时 I 的积分中还可以含有高阶导数项, 此时 M 也应做相应的改变.

变分问题的重要组成部分: **约束集合 + 目标泛函**.

例 1.1 (最速降线)**.** *在平面上给定两点 $A = (x_1, y_1)$ 和 $B = (x_2, y_2)$, 其中 $x_1 < y_1, x_2 < y_2$. 一个质点沿着一条连接这两点的“光滑”曲线仅受重力下滑. 设初速度为零, 问沿怎样的一条曲线滑行时间最短?*

- 约束集合. 为了问题的简单, 我们假设这条曲线有着显式表达式 $y=u(x)$, 其中 $u\in C^1[x_1,x_2]$. 注意到这条曲线的起点和终点都是固定的, 由此我们设定约束函数集合为
$$M=\{u\in C^1[x_1,x_2]\colon u(x_i)=y_i, i=1,2\}.$$
- 目标泛函. 我们要在 M 中选取合适的函数 u^*, 使得总时间 $T=T(u)$ 达到极小. 直接列出动力学方程:
$$\begin{cases}\dfrac{1}{2}mv^2=mg(y_1-u(x)),\\ v=\dfrac{\mathrm{d}s}{\mathrm{d}t}.\end{cases}$$
联立得
$$T=T(u)=\int \mathrm{d}t=\frac{1}{\sqrt{2g}}\int_{x_0}^{x_1}\sqrt{\frac{1+|\dot{u}(x)|^2}{2g(y_1-u(x))}}\,\mathrm{d}x.$$

至此, 我们便将最速降线问题转化成了一个变分学问题:
$$\begin{aligned}&\text{minimize } T(u)\\ &\text{s.t. } u\in M.\end{aligned}$$

在以下几节中, 若无特别说明, 均假设 $n=1$.

1.1 必要条件: Euler-Lagrange 方程

函数极值的必要条件. 设 Ω 为 $\mathbb{R}^N$ 中的开集, 函数 $f\in C^1(\Omega)$ 在 $x^*\in\Omega$ 处达到极小. 于是对任意的 $h\in\mathbb{R}^n\setminus\{0\}$, 存在充分小的 $\varepsilon(h)>0$, 使得当 $0<|\varepsilon|<\varepsilon(h)$ 时, $x^*+\varepsilon h\in\Omega$, 且 $f(x^*+\varepsilon h)\geq f(x^*)$. 令 $g_h(\varepsilon)=f(x^*+\varepsilon h)$, 由上述分析可知, $\varepsilon=0$ 是单变量函数 g_h 的极小点, 故
$$0=\dot{g}_h(0)=\nabla f(x^*)\cdot h.$$
再由 h 的任意性, 即得 $\nabla f(x^*)=0$.

1.1.1 Euler-Lagrange 方程

对于泛函的情形我们也可以进行类似的处理. 首先做一些设定:

给定区间 $J=[t_0,t_1]\subseteq\mathbb{R}$ 和开区域 $\Omega\subseteq\mathbb{R}^N$, 取 $L=L(x,u,p)\in C^1(J\times\Omega\times\mathbb{R}^N)$. 考虑约束集合
$$M=\{u\in C^1(J)\colon u(t_i)=P_i, i=1,2\}$$
以及目标泛函
$$I(u)=\int_J L(t,u(t),\dot{u}(t))\,\mathrm{d}t.$$
称 u^* 是 I 的**极小点**, 若存在 u^* 的邻域 U, 使得对任意的 $u\in U\cap M$, 有 $I(u)\geq I(u^*)$.

- 函数的情形引入非零向量 $h \rightsquigarrow$ 泛函的情形引入测试函数 φ. 注意到对于函数自身正则性以及边值的约束, 一般取 $\varphi \in C_0^1$.
- 邻域的刻画: 函数的情形引入充分小的 $\varepsilon \rightsquigarrow$ 泛函的情形是否能引入充分小的 $\varepsilon = \varepsilon(\varphi)$ 来刻画? 事实上, 注意到 $u(J)$ 是紧的, 故 $\varepsilon(\varphi)$ 的存在性可由 $\mathbb{R}^n$ 的 T4 分离公理保证.

引理 1.2 (du Bois-Reymond, 变分法基本引理)**.** 若 $\psi \in C[t_0,t_1]$, 且

$$\int_{t_0}^{t_1} \psi(t)\dot{\lambda}(t)\,\mathrm{d}t = 0, \quad \forall \lambda \in C_0^1[t_0,t_1],$$

则 $\psi = \text{const}$.

证明. 令

$$c = \frac{1}{t_1 - t_0}\int_{t_0}^{t_1} \psi(t)\,\mathrm{d}t, \lambda(t) = \int_{t_0}^{t} (\psi(s) - c)\,\mathrm{d}s.$$

注意到此时 $\lambda \in C_0^1[t_0,t_1]$, 故利用题设条件可得

$$\int_{t_0}^{t_1} (\psi(t) - c)^2\,\mathrm{d}t = \int_{t_0}^{t_1} \dot{\lambda}(t)(\psi(t) - c)\,\mathrm{d}t = -c\int_{t_0}^{t_1} (\psi(t) - c)\,\mathrm{d}t = 0.$$

从而有 $\psi = c$. □

注 1.3. *上述 du Bois-Reymond 引理还有一些简单的推广形式. 如将 $\psi \in C(J)$ 换成 $\psi \in L^\infty(J)$ 或 $\psi \in L^1(J)$, 相应地把 $\lambda \in C_0^1(J)$ 换成 $\lambda \in \mathrm{AC}_0(J)$ 或 $\lambda \in C_c^\infty(J)$.*

命题 1.4. *设 $u^* \in M$ 是泛函 I 在 M 上的一个极小点, 则它满足下列* ***积分形式的 Euler-Lagrange 方程*** *(简称* ***E-L 方程****):*

$$\boxed{-\int_{t_0}^{t} L_{u_i}(s,u^*(s),u^*(s))\,\mathrm{d}s + L_{p_i}(t,u^*(t),u^*(t)) = \text{const}, \qquad \forall t \in J, 1 \le i \le n.}$$

证明. 由前述分析可知, 对任意的 $\varphi \in C_0^1[t_0,t_1]$, 存在充分小的 $\varepsilon(\varphi) > 0$, 使得当 $0 < |\varepsilon| < \varepsilon(\varphi)$ 时, 有 $I(u^* + \varepsilon\varphi) \ge I(u^*)$. 令 $g_\varphi(\varepsilon) = I(u^* + \varepsilon\varphi)$. 题设条件表明 $\dot{g}_\varphi(0) = 0$ (称 $\dot{g}_\varphi(0)$ 为 I 对 φ 的**一阶变分**, 记作 $\delta I(u^*,\varphi)$). 利用分部积分公式, 进一步计算得

$$\begin{aligned} 0 = g_\varphi'(0) &= \sum_{i=1}^{N} \int_J (L_{u_i}(t,u^*(t),\dot{u}^*(t))\varphi_i(t) + L_{p_i}(t,u^*(t),\dot{u}^*(t))\dot{\varphi}_i(t))\,\mathrm{d}t \\ &= \sum_{i=1}^{N} \int_J \left(-\int_{t_0}^{t} L_{u_i}(s,u^*(s),\dot{u}^*(s))\,\mathrm{d}s + L_{p_i}(t,u^*(t),\dot{u}^*(t))\right)\dot{\varphi}_i(t)\,\mathrm{d}t. \end{aligned}$$

由于上述结果对任意的 $\varphi \in C_0^1[t_0,t_1]$ 均成立, 故利用 du Bois-Reymond 引理, 我们便得到了所证结论. □

特别地, 若假设 L 和 u^* 具有更高的正则性, 如 $L \in C^2$ 以及 $u^* \in C^2$, 则 u^* 满足如下**微分形式的 E-L 方程**:

$$\boxed{L_{u_i}(t,u^*(t),\dot{u}^*(t)) - \frac{\mathrm{d}}{\mathrm{d}t}L_{p_i}(t,u^*(t),\dot{u}^*(t)) = 0.}$$

若 $L \in C^1$ 且 $u^* \in C^1$, 上述等式在广义导数意义下成立.

注 1.5. *事实上, E-L 方程可适用于更广的函数类. 例如, 考虑 Lipschitz 函数类* $\mathrm{Lip}(J)$. *注意到包含关系* $\mathrm{Lip}(J) \subseteq \mathrm{AC(J)}$, *故其导数几乎处处存在, 此时目标泛函中的积分按照 Lebesgue 积分意义理解. 利用控制收敛定理和 du Bois-Reymond 引理, 我们仍可以导出积分形式的 E-L 方程.*

特别地, 逐段 C^1 的函数是 Lipschitz 连续的, 因此积分形式的 E-L 方程对于逐段 C^1 的连续函数也成立.

例 1.6 (质点运动方程)**.** *设 $\mathbb{R}^3$ 中某质量为 m 的质点受外力 F 的作用, 其位置坐标为 $x=(x_1,x_2,x_3)$, 则其动能 $T=m|\dot{x}|^2/2$. 若更设 F 有位势, 即存在函数 V 满足 $-\nabla V = F$, 我们称*

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

为 Lagrange 函数[1]*. 适当确定定义域 M, 考虑泛函*

$$I(x) = \int_{t_1}^{t_2} L(x(t),\dot{x}(t))\,\mathrm{d}t.$$

通过直接计算可知, I 对应的 E-L 方程为 $F = m\ddot{x}$, 即 Newton 第二定律所确定的运动轨道.

1.1.2 变分导数

从函数极值的角度来看, 若 x^* 是 f 的极小值点, 则我们只需要考虑 x^* 的一个小邻域内 f 的行为. 类似的, 尽管 E-L 方程的推导是在整个区间上进行的, 但对于任意的 $\tau \in \mathrm{Int}\,J$, 在这一点的 E-L 方程只依赖于这点附近 L 的行为. 粗糙地说, 如果 u^* 在某点的某一邻域内达到 "最优", 那么 u^* 便满足整个区间上在这一点的 E-L 方程.

从以下极限过程看这种**局部性**: 引入 **Euler-Lagrange 算子** E_L:

$$\boxed{E_L(u)(t) := L_{u_i}(t,u^*(t),\dot{u}^*(t)) - \frac{\mathrm{d}}{\mathrm{d}t}L_{p_i}(t,u^*(t),\dot{u}^*(t)).}$$

以 $N=1$ 为例. 设 $L \in C^2$, $u \in C^2$. 取 $c \in (t_0,t_1)$, $\varphi \in C_0^1[t_0,t_1]$, 其中 $\mathrm{supp}\,\varphi \subseteq B_h(c)$. 直接计算得

$$\begin{aligned}
\frac{I(u+\varphi)-I(u)}{\Delta\sigma} &= -\frac{\displaystyle\int_{c-h}^{c+h}\left(\int_{t_0}^{t} E_L(u+\theta\varphi)(s)\,\mathrm{d}s\right)\dot{\varphi}(t)\,\mathrm{d}t}{\Delta\sigma} \\
&= \frac{\displaystyle\int_{c-h}^{c+h} E_L(u+\theta\varphi)(t)\varphi(t)\,\mathrm{d}t}{\Delta\sigma},
\end{aligned}$$

[1]这是分析力学中的专有名词.

其中 $\theta\in(0,1)$, $\Delta\sigma=\int_{B_h(c)}\varphi\,\mathrm{d}t$. 由简单的估计可知, 当 $h\to 0$ 且 $\sup_{\overline{B_h(c)}}|\dot\varphi|\to 0$ 时, 上式趋近于 $E_L(u)$. 因此, 我们称 E-L 算子对函数 u 作用后在 t 点的值为 I 在 t 的**变分导数**.

1.1.3 例

如下考虑 E-L 方程的具体求解 ($N=1$):

- L 不含 u, 即 $L=L(t,p)$. 此时 E-L 方程简化为

$$\frac{\mathrm{d}}{\mathrm{d}t}L_p(t,\dot u(t))=0.$$

  若能从上述方程中解出 $\dot u$, 那么也就可以通过积分得到 u.

- L 不含 p, 即 $L=L(t,u)$. 此时 E-L 方程化为函数方程 $L_u(t,u)=0$, 其解是一条或多条曲线.

- (自守系统) L 不含 t, 即 $L=L(u,p)$. 在这种情况下, 引入 **Hamilton** 量 $H(u,p)=pL_p(u,p)-L(u,p)$. 通过直接计算, 我们有如下结果:

命题 1.7. *设 $L\in C^2$ 且与 t 无关. 又设 $u^*\in C^2$ 是对应 E-L 方程的解, 则*

$$H(u^*(t),\dot u^*(t))\equiv\text{const},\qquad\forall\, t.$$

利用上述命题, 我们可以在一些特殊情况下求解出 u^*.

例 1.8 (最速降线-续)**.** *此时*

$$L(u,p)=\frac{1}{2g}\frac{\sqrt{1+p^2}}{\sqrt{y_1-u}}$$

与 t 无关. 利用上述命题, 我们有 $(pL_p-L)|_{(u,\dot u)}=\text{const}$, 即存在常数 c 使得

$$-\frac{\sqrt{1+\dot u^2}}{\sqrt{y_1-u}}+\frac{\dot u^2}{\sqrt{(1+\dot u^2)(y_1-u)}}=c.$$

化简得

$$c^2(1+\dot u^2)(y_1-u)=1.$$

令 k 为一待定常数, 引入参变量 θ, 令

$$\begin{cases}x=x(\theta),\\ u=u(\theta)=y_1-k(1-\cos\theta),\end{cases}$$

则有

$$c^2\left(1+k^2\frac{\sin^2\theta}{\dot x(\theta)}\right)k(1-\cos\theta)=1.$$

联想到半角公式, 故取 $k=1/2c^2$, 则可以解出

$$\dot{x}(\theta)=k(1-\cos\theta).$$

从而有

$$\begin{cases} x(\theta)=x_1+k(\theta-\sin\theta), \\ u(\theta)=y_1-k(1-\cos\theta), \end{cases} \qquad \theta\in[0,\Theta],$$

其中 k 与 Θ 通过

$$\begin{cases} x(\Theta)=x_2, \\ y(\Theta)=y_2 \end{cases}$$

确定.

1.2 充分与必要条件: Legendre-Hadamard 条件, Jacobi 场

一阶导/梯度 $\rightsquigarrow$ 一阶变分 $\Rightarrow$ E-L 方程; 二阶导/Hessian 阵 $\rightsquigarrow\cdots$

设 Ω 为 $\mathbb{R}^N$ 中的开集, $f\in C^2(\Omega)$ 在 $x_0\in\Omega$ 处取得极小值. 由多元微积分的知识可知, 此时 Hessian 阵 $(\partial_{ij}f)|_{x=x_0}$ 是半正定的; 反之, 若 $(\partial_{ij}f)|_{x=x_0}$ 是正定的, 则 f 在 x_0 处取到极小值.

1.2.1 二阶变分, Legendre-Hadamard 条件

设 $L\in C^2(J\times\mathbb{R}^N\times\mathbb{R}^N)$,

$$I(u)=\int_J L(t,u(t),\dot{u}(t))\,\mathrm{d}t.$$

又设 $u^*\in M$ 是对应 E-L 方程的解. 同 1.1 节的处理方法, 对任意的 $\varphi\in C_0^1(J)$, 考虑单变量函数 $g_\varphi(\varepsilon)=I(u^*+\varepsilon\varphi), 0<|\varepsilon|<\varepsilon(\varphi)$. 则 g_φ 以 $x=0$ 为极小值点, 从而有 $\ddot{g}_\varphi(0)\geq 0$. 直接计算得

$$\ddot{g}(0)=\sum_{i,j=1}^N\int_J(L_{u_iu_j}(t,u^*(t),\dot{u}^*(t))\varphi_i(t)\varphi_j(t)+ \\ L_{u_ip_j}(t,u^*(t),\dot{u}^*(t))\varphi_i(t)\dot{\varphi}_j(t)+L_{p_ip_j}(t,u^*(t),\dot{u}^*(t))\dot{\varphi}_i(t)\dot{\varphi}_j(t))\,\mathrm{d}t.$$

将上述等号右侧的表达式记为 $\delta^2I(u^*,\varphi)$, 称为 I 在 u^* 处的**二阶变分**. 若引入函数矩阵

$$\begin{cases} A_{u^*}=(L_{p_ip_j}(t,u^*(t),\dot{u}^*(t))), \\ B_{u^*}=(L_{p_iu_j}(t,u^*(t),\dot{u}^*(t))), \\ C_{u^*}=(L_{u_iu_j}(t,u^*(t),\dot{u}^*(t))). \end{cases}$$

则二阶变分的表达式可以简化为

$$\int_J(\dot{\varphi}^tA_{u^*}\dot{\varphi}+2\dot{\varphi}^tB_{u^*}\varphi+\varphi^tC_{u^*}\varphi)\,\mathrm{d}t.$$

由上述分析, 我们有必要条件

$$u^*\text{极小} \Rightarrow \delta^2 I(u^*,\varphi) \geq 0, \forall \varphi \in C_0^1(J).$$

上述必要条件可以进一步简化. 事实上, 三个矩阵 $A_{u^*}, B_{u^*}, C_{u^*}$ 在判断 u^* 成为极小点中的地位是不平等的, 具体体现在当 $\|\varphi\|_C$ 变化不大时, $\|\varphi\|_{C^1}$ 可以有很大的变化; 或者说, $\dot{\varphi}$ 比 φ 的影响更大.

命题 1.9 (必要条件 1)**.** *设 $L \in C^2(J \times \mathbb{R}^N \times \mathbb{R}^N)$. 若 $u^* \in M$ 是 I 的极小点, 则 A_{u^*} 是半正定的, 即*

$$\boxed{\xi^t A_{u^*} \xi = \sum_{i,j=1}^N L_{p_i,p_j}(\tau, u^*(\tau), \dot{u}^*(\tau))\xi_i \xi_j \geq 0, \qquad \forall \tau \in J, \xi \in \mathbb{R}^N.} \tag{2}$$

称(2)为 ***Legendre-Hadamard* 条件**.

证明. 取 $v \in C^1(\mathbb{R})$: $v|_{|s|\geq 1} = 0, \|\dot{v}\|^2_{L^2(\mathbb{R})} = 1$. 对任意的 $\tau \in \mathrm{Int}(J), \xi \in \mathbb{R}^N$ 以及充分小的 $\mu > 0$, 令

$$\varphi(t) = \xi \mu v\left(\frac{t-\tau}{\mu}\right).$$

将上述构造的 φ 代入至二阶变分的表达式中, 即得

$$0 \leq \delta^2 I(u^*,\varphi) = \mu(\xi^t A_{u^*}\xi) + o(\mu) \qquad (\mu \to 0^+).$$

上式即表明 A_{u^*} 是半正定的. □

1.2.2 Poincaré 不等式

命题 1.10 (充分条件 1)**.** *若 $u^* \in C_0^1(J)$ 满足 E-L 方程, 且存在 $\lambda > 0$ 使得*

$$\delta^2 I(u^*,\varphi) \geq \lambda \int_J (|\varphi|^2 + |\dot{\varphi}|^2)\,\mathrm{d}t, \forall \varphi \in C_0^1(J),$$

则 u^* 是 I 的一个严格极小点.

证明. 对任意的 $\varphi \in C_0^1(J)$, 令 $g_\varphi(\varepsilon) = I(u^* + \varepsilon\varphi), 0 < |\varepsilon| < \varepsilon(\varphi)$. 注意到 u^* 满足 E-L 方程, 故有

$$g(\varepsilon) - g(0) = g(\varepsilon) - g(0) - s\dot{g}(0) = \frac{s^2}{2}\ddot{g}(\theta\varepsilon) = \frac{s^2}{2}(\ddot{g}(\theta\varepsilon) - \ddot{g}(0)) + \frac{s^2}{2}\ddot{g}(0),$$

其中 $\theta = \theta(\varphi) \in (0,1)$. 再注意到对于 $\|\varphi\|_{C^1} \leq 1$, 当 $s \to 0$ 时有

$$|A_{u^*+\varepsilon\varphi} - A_{u^*}| + |B_{u^*+\varepsilon\varphi} - B_{u^*}| + |C_{u^*+\varepsilon\varphi} - C_{u^*}| = o(1).$$

因此当 $|\varepsilon| > 0$ 充分小时, 存在 $\eta < \lambda$ 使得

$$\ddot{g}(\theta\varepsilon) - \ddot{g}(0) \geq -\eta \int_J (|\varphi|^2 + |\dot{\varphi}|^2)\,\mathrm{d}t.$$

结合题设条件, 我们便有

$$I(u^*+\varepsilon\varphi)-I(u^*)=g(\varepsilon)-g(0)\geq\frac{\varepsilon^2}{2}(\lambda-\eta)\int_J(|\varphi|^2+|\dot\varphi|^2)\,\mathrm{d}t>0,$$

从而 u^* 是 I 的严格极小点. □

注 1.11. 由上述命题, 如果矩阵

$$\begin{pmatrix} A_{u^*} & B_{u^*} \\ B_{u^*} & C_{u^*} \end{pmatrix}$$

是正定的, 那么 *E-L* 方程的解 u^* 必是极小点.

命题1.10给出的充分条件还可以被进一步地简化.

引理 1.12 (Poincaré 不等式)**.** 设 $\varphi\in C_0^1[t_0,t_1]$, 则

$$\int_J|\varphi|^2\,\mathrm{d}t\leq\frac{(t_1-t_0)^2}{2}\int_J|\dot\varphi|^2\,\mathrm{d}t. \tag{3}$$

证明. 注意到

$$\varphi(t)=\int_{t_0}^t\dot\varphi(s)\,\mathrm{d}s,$$

故由 Cauchy-Schwarz 不等式可知,

$$|\varphi(t)|^2\leq\left(\int_{t_0}^t|\dot\varphi(s)|^2\,\mathrm{d}s\right)^2\leq(t-t_0)\int_J|\dot\varphi(s)|^2\,\mathrm{d}s.$$

积分得

$$\int_J|\varphi(t)|^2\,\mathrm{d}t\leq\frac{(t_1-t_0)^2}{2}\int_J|\dot\varphi|^2\,\mathrm{d}t.$$

□

注 1.13. 若 $\varphi\in\mathrm{AC(J)}$, $\dot\varphi\in L^2(J)$ 且 $\varphi(a)=0$, 则 *Poincaré* 不等式(3)仍成立.

由 Poincaré 不等式可知, 若存在 $\lambda>0$ 使得

$$\delta^2I(u^*,\varphi)\geq\lambda\int_J|\dot\varphi|^2\,\mathrm{d}t,\qquad\forall\varphi\in C_0^1(J),$$

那么命题1.10成立.

命题 1.14 (充分条件 2)**.** 给定一个足够光滑的 *L*, 设沿其 *E-L* 方程的解 u^* 满足**严格的 *Legendre-Hadamard* 条件**, 即 A_{u^*} 是正定的. 若存在 $\mu>0$ 使得

$$\delta^2I(u^*,\varphi)\geq\mu\int_J|\varphi|^2\,\mathrm{d}t,\qquad\forall\varphi\in C_0^1(J),$$

则存在 $\lambda>0$ 使得

$$\delta^2I(u^*,\varphi)\geq\lambda\int_J(|\varphi|^2+|\dot\varphi|^2)\,\mathrm{d}t,\qquad\forall\varphi\in C_0^1(J).$$

从而 u^* 是 I 的一个严格极小值点.

证明. 由于 A_{u^*} 是正定的, 故存在 $\alpha>0$ 使得

$$\dot{\varphi}^t A_{u^*}\dot{\varphi} \geq \alpha|\dot{\varphi}|^2, \qquad \forall \varphi \in C_0^1(J).$$

从而存在正常数 C_1, C_2 使得

$$\begin{aligned}
\alpha \int_J |\dot{\varphi}|^2 \,\mathrm{d}t &\leq \delta^2 I(u^*,\varphi) + \int_J (2|\dot{\varphi}^t B_{u^*}\varphi| + |\varphi^t C_{u^*}\varphi|)\,\mathrm{d}t \\
&\leq \delta^2 I(u^*,\varphi) + C_1\left(\left(\int_J |\dot{\varphi}|^2\,\mathrm{d}t\right)^{1/2}\left(\int_J |\varphi|^2\,\mathrm{d}t\right)^{1/2} + \int_J |\varphi|^2\,\mathrm{d}t\right) \\
&\leq \delta^2 I(u^*,\varphi) + \frac{\alpha}{2}\int_J |\dot{\varphi}|^2\,\mathrm{d}t + C_2\int_J |\varphi|^2\,\mathrm{d}t,
\end{aligned}$$

这里我们用到了加权的初等不等式 $\sqrt{ab} \leq (ka+b/k)/2, a,b>0, k>0$. 由此即得

$$\int_J |\dot{\varphi}|^2\,\mathrm{d}t \leq \frac{2}{\alpha}(1+C_2\mu^{-1})\delta^2 I(u^*,\varphi), \qquad \forall \varphi \in C_0^1(J).$$

这表明 u^* 是一个极小点. □

1.2.3 Jacobi 场, 共轭点

1.2.2 节中列出的充分条件中仍含有任意函数 φ, 还需要寻找更为精确的充分条件. 以下我们建立充分条件与严格 Legendre-Hadamard 条件之间的联系.

注意到二阶变分的具体表达式. 设 $L \in C^3, u^*$ 是 E-L 方程的解. 令

$$\boxed{\Phi_{u^*}(t,\xi,\eta) := \eta^t A_{u^*}\eta + 2\xi^t B_{u^*}\eta + \xi^t C_{u^*}\xi, \qquad \forall (\xi,\eta) \in \mathbb{R}^N \times \mathbb{R}^N.}$$

我们称其为**附属的 (accessory)Lagrange 函数**. 此时 $\delta^2 I(u^*,\varphi)$ 可以看作是附属 Lagrange 函数相关的变分积分:

$$\delta^2 I(u^*,\varphi) = Q_{u^*}(\varphi) = \int_J \Phi_{u^*}(t,\varphi(t),\dot{\varphi}(t))\,\mathrm{d}t.$$

若设 u^* 是一个极小点, 则有 $Q_{u^*}(\varphi) = \delta^2 I(u^*,\varphi) \geq 0, \forall \varphi \in C_0^1(J)$. 注意到 $Q_{u^*}(0)=0$, 故 0 是 Q_{u^*} 的极小点. 更一般地, 我们将 Q_{u^*} 的定义域扩充到 $\mathrm{Lip}_0(J)$ 上, 导出它的积分形式的 E-L 方程:

$$A_{u^*}\dot{\varphi}(t) + B_{u^*}\varphi(t) - \int_{t_0}^t (B_{u^*}\dot{\varphi}(t) + C_{u^*}\varphi(t))\,\mathrm{d}t = \mathrm{const}.$$

事实上, 若 L 沿 u^* 满足满足严格的 Legendre-Hadamard 条件, 那么上述方程的解 $\varphi \in C^2(J)$. 因此 φ 满足微分形式的 E-L 方程:

$$\boxed{J_{u^*}(\varphi) = \frac{\mathrm{d}}{\mathrm{d}t}(A_{u^*}\dot{\varphi}(t) + B_{u^*}\varphi(t)) - (B_{u^*}\dot{\varphi}(t) + C_{u^*}\varphi(t)) = 0, \quad t \in J.}$$

称此方程为 **Jacobi 方程**, 算子 J_{u^*} 为沿 E-L 方程的解 u^* 的 **Jacobi 算子**, Jacobi 算子的任意一个 C^2 解为沿轨道 u^* 的一个 **Jacobi 场**.

命题 1.15 (Jacobi 场的刻画)**.** *设 φ^* 是沿 u^* 的一个 Jacobi 场, 则 $Q_{u^*}(\varphi^*)=0$; 反之, 若 $\varphi^*\in \mathrm{Lip}_0(J)$ 满足 $Q_{u^*}(\varphi^*)=0$, 而且 $Q_{u^*}(\varphi^*)\geq 0, \forall \varphi\in C_0^1(J)$, 则 φ^* 是沿 u^* 的一个 Jacobi 场.*

证明. 若 φ^* 是沿 u^* 的一个 Jacobi 场, 通过直接计算可得

$$
\begin{aligned}
Q_{u^*}(\varphi^*) &= \int_J ((\dot{\varphi}^*)^t(A_{u^*}\dot{\varphi}^*+B_{u^*}\varphi^*)+(\varphi^*)^t(B_{u^*}\dot{\varphi}^*+C_{u^*}\varphi^*))\,\mathrm{d}t \qquad (4)\\
&= \int_J (\varphi^*)^t\left(-\frac{\mathrm{d}}{\mathrm{d}t}(A_{u^*}\dot{\varphi}^*+B_{u^*}\varphi^*)+(B_{u^*}\dot{\varphi}^*+C_{u^*}\varphi^*)\right)\mathrm{d}t \qquad (5)\\
&= -\int_J (\varphi^*)^t J_{u^*}(\varphi^*)\,\mathrm{d}t = 0. \qquad (6)
\end{aligned}
$$

反之, 利用光滑函数逼近，我们有

$$
Q_{u^*}(\varphi)=0, \qquad \forall \varphi\in \mathrm{Lip}_0(J).
$$

于是 φ^* 是 u^* 的一个极小点. 由前述分析可知, φ^* 也满足微分形式的 E-L 方程, 即 $J_{u^*}(\varphi^*)=0$. □

定义 1.16. *设 u^* 是 I 的 E-L 方程的一个解. 称 $(a,u^*(a))$ 与 $(b,u^*(b))$ 是轨道 $(t,u^*(t))$ 上的一对**共轭点**, 如果存在一个沿 $u^*(t)$ 的非零 Jacobi 场 $\varphi\in C_0^1[a,b]$. 若在轨道 $\{(t,u^*(t)) : t\in (t_0,t_1]\}$ 上 $(t_0,u^*(t_0))$ 没有共轭点，则称 u^* 没有共轭点.*

命题 1.17 (必要条件 2)**.** *设 u^* 是 I 的 E-L 方程的一个解, 且 L 沿 u^* 满足严格的 Legendre-Hadamard 条件. 若 $Q_{u^*}(\varphi)\geq 0, \forall\varphi\in C_0^1(J)$, 则不存在 $a\in \mathrm{Int}(J)$ 使得 $(a,u^*(a))$ 共轭于 $(t_0,u^*(t_0))$.*

证明. 若不然, 设存在 $a\in\mathrm{Int}(J)$ 使得 $(a,u^*(a))$ 是 $(t_0,u^*(t_0))$ 的共轭点, 即存在 u^* 的非零 Jacobi 场 $\xi\in C^2[t_0,a]$, 并满足 $\xi(a)=\xi(t_0)=0$. 现令

$$
\widetilde{\xi}(t)=\begin{cases}\xi(t) & t\in[t_0,a],\\ 0 & t\in(a,t_1],\end{cases}
$$

则 $\widetilde{\xi}\in\mathrm{Lip}(J)$, 且满足 $\widetilde{\xi}(t_0)=\widetilde{\xi}(t_1)=0$, 而且

$$
Q_{u^*}(\widetilde{\xi})=\int_{t_0}^{a}\Phi_{u^*}(t,\xi(t),\dot{\xi}(t))\,\mathrm{d}t=0.
$$

由命题1.15可知, $\widetilde{\xi}\in C^2(J)$ 且满足 Jacobi 方程, 即 $J_{u^*}(\widetilde{\xi})=0$. 由常微分方程初值问题的唯一性可知, $\widetilde{\xi}=0$, 矛盾. □

当 $N=1$ 时, 上述命题的逆也成立.

命题 1.18 (充分条件 3, $N=1$)**.** *设 u^* 是 E-L 方程的一个解, L 沿 u^* 满足严格的 Legendre-Hadamard 条件. 若在 J 上存在一个正的 Jacobi 场 ψ, 则 u^* 是一个严格极小点。特别地, 若 u^* 没有共轭点, 则在 J 上存在一个正的 Jacobi 场.*

证明. 证明分三步进行.

Step 1. 若 u^* 没有共轭点, 那么便存在一个正的 Jacobi 场. 事实上, 设 λ 是一个 Jacobi 场, 其中 $\lambda(t_0)=0, \dot{\lambda}(t_0)=1$. 由假设条件可知, $\lambda(t)>0, t\in(t_0,t_1]$. 再根据微分方程对初值的连续依赖性, 故存在一个 Jacobi 场 ψ 使得 $\psi(t)>0, t\in J$.

Step 2. 对任意的 $\varphi\in C_0^1(J)$, 有

$$Q_{u^*}(\varphi)=\int_J A_{u^*}\psi^2\left(\frac{\dot{\varphi}}{\psi}\right)^2\mathrm{d}t.$$

记 $\alpha=\varphi/\psi$, 则 $\varphi=\alpha\psi, \dot{\varphi}=\dot{\alpha}\psi+\alpha\dot{\psi}$. 从而有

$$A_{u^*}\dot{\varphi}^2+2B_{u^*}\dot{\varphi}\varphi+C_{u^*}\varphi^2=\alpha^2(A_{u^*}\dot{\psi}^2+2B_{u^*}\dot{\psi}\psi+C_{u^*}\psi^2)+2\dot{\alpha}\alpha(A_{u^*}\dot{\psi}+B_{u^*}\psi)+\dot{\alpha}^2A_{u^*}\psi^2.$$

注意到 ψ 满足 Jacobi 方程, 故

$$\begin{aligned}Q_{u^*}(\varphi)&=\int_J\left(\frac{\mathrm{d}(\psi\lambda^2)}{\mathrm{d}t}(A_{u^*}\dot{\psi}+B_{u^*}\psi)+\psi\lambda^2\frac{\mathrm{d}(A_{u^*}\dot{\psi}+B_{u^*}\psi)}{\mathrm{d}t}+A_{u^*}\dot{\lambda}^2\psi^2\right)\mathrm{d}t\\&=(\psi\lambda^2(A_{u^*}\dot{\psi}+B_{u^*}\psi))\Big|_{t_0}^{t_1}+\int_J A_{u^*}\dot{\lambda}^2\psi^2\,\mathrm{d}t\\&=\int_J A_{u^*}\dot{\lambda}^2\psi^2\,\mathrm{d}t.\end{aligned}$$

Step 3. 记 $\beta=\inf_J(A_{u^*}\psi^2)>0$. 对任意的 $\varphi\in C_0^1(J)$, 由 Poincaré 不等式可得

$$\begin{aligned}Q_{u^*}(\varphi)=\int_J A_{u^*}\psi^2\left(\frac{\dot{\varphi}}{\psi}\right)^2\mathrm{d}t\geq\beta\int_J\left(\frac{\dot{\varphi}}{\psi}\right)^2\mathrm{d}t&\geq\frac{2\beta}{|J|^2}\int_J\left(\frac{\varphi}{\psi}\right)^2\mathrm{d}t\\&\geq\frac{2\beta}{|J|^2}\inf_J\left(\frac{1}{\psi^2}\right)\int_J\varphi^2\,\mathrm{d}t.\end{aligned}$$

结合命题1.14的结论, 我们便证得了所需结论. □

例 1.19. 设 $M=\{u\in C^1[0,1]\colon u(0)=1, u(1)=b\}$. 考虑以下泛函

$$I(u)=\int_0^1(t\dot{u}+\dot{u}^2)\,\mathrm{d}t.$$

解. 直接计算得 $L_u=0, L_p=2p+t$, 故其 E-L 方程 $2\ddot{u}+1=0$ 有满足初值条件的解

$$u^*(t)=-\frac{t^2}{4}+\left(b-a+\frac{1}{4}\right)t+a.$$

由于 $L_{pp}=2>0$, 并且对应的 Jacobi 方程 $\ddot{\varphi}=0$ 在 $[0,1]$ 上有一个正解, 所以 u^* 是一个严格极小点.

1.3 强极小与极值场

C^1 拓扑 $\rightsquigarrow$ 考虑的因素更多, "弱极小点". $C=C^0$ 拓扑 $\rightsquigarrow$ 考虑的因素更少, "强极小点".

定义 1.20. 设 $J=[t_0,t_1], L\in C^1(J\times\mathbb{R}^N\times\mathbb{R}^N)$. 称 $u\in C^1(J)$ 为

$$I=\int_J L(t,u(t),\dot{u}(t))\,\mathrm{d}t$$

的**强 *(*弱*)* 极小点**, 如果存在 $\varepsilon>0$ 使得对一切满足

$$\|\varphi\|_C<\varepsilon \qquad (\|\varphi\|_{C^1}<\varepsilon)$$

的 $\varphi\in C_0^1(J)$ 都有 $I(u+\varphi)\geq I(u)$. 函数类 C^1 可以换成 Lip, 并且用 Lip 模代替 C^1 模.

前两节所讨论的极小点是弱极小点. Lip 意义下的弱极小点也是 C^1 意义下的弱极小点; 强极小点是弱极小点, 但反之不然.

例 1.21. 设 $M=\mathrm{Lip}_0[0,1]$,

$$I(u)=\int_0^1(\dot{u}^2+\dot{u}^3)\,\mathrm{d}t.$$

注意到 $I(0)=0$, 且当 $\|u\|_{\mathrm{Lip}}<1/2$ 时, 有

$$I(u)-I(0)=\int_0^1\dot{u}^2(1+\dot{u})\,\mathrm{d}t\geq\frac{1}{2}\int_0^1\dot{u}^2\,\mathrm{d}t\geq 0.$$

故 $u=0$ 是弱极小点, 但不是强极小点. 事实上, 对充分小的 $h>0$, 令

$$u_h(x)=\begin{cases}-\dfrac{x}{h} & x\in[0,h^2],\\ \dfrac{h(x-1)}{1-h^2} & x\in[h^2,1].\end{cases}$$

一方面, $\|u_h\|_C\leq h$; 另一方面, 直接计算得

$$(\dot{u}_h^2+\dot{u}_h^3)(x)=\begin{cases}\dfrac{1}{h^2}-\dfrac{1}{h^3}\leq-\dfrac{1}{2h^3} & x\in[0,h^2),\\ \left(\dfrac{h}{1-h^2}\right)^2+\left(\dfrac{h}{1-h^2}\right)^3\leq 2 & x\in(h^2,1].\end{cases}\tag{7}$$

因此

$$I(u_h)-I(0)\leq 2-\frac{1}{2h}\to-\infty \quad (h\to 0^+).$$

从而 $u=0$ 不是强极小点.

1.3.1 必要条件: Weierstrass 过度函数

设 $u^* \in C^1(J)$ 是 E-L 方程的解. 以下探究 u^* 成为强极小点的必要条件. 为此我们把 $I(u^*)$ 与 I 在 u^* 的 C 拓扑临近的函数上的值作比较, 即构造适当的 $\varphi \in C_0^1(J)$.

命题 1.22 (必要条件)**.** 若 $u^* \in C^1(J)$ 是 I 的一个强极小点, 则

$$\mathfrak{E}_L(t,u(t),\dot u(t),\dot u(t)+\xi) \geq 0, \qquad \forall \xi \in \mathbb{R}^N, \tau \in J.$$

这里

$$\boxed{\mathfrak{E}_L = \mathfrak{E}_L(t,u,p,q) := L(t,u,p) - L(t,u,q) - (q-p)\cdot L_p(t,u,p),}$$

并称之为 ***Weierstrass* 过度函数**.

证明. 对任意的 $\xi \in \mathbb{R}^N, \tau \in \mathrm{Int}(J)$, 当 $\lambda > 0$ 充分小时, 可使得 $[\tau-\lambda^2, \tau+\lambda] \subseteq (t_0,t_1)$. 令

$$\psi_\lambda(x) = \begin{cases} 0 & s \in (\infty, -\lambda^2] \cup [\lambda, \infty), \\ s+\lambda^2 & s \in [-\lambda^2, 0], \\ -\lambda s + \lambda^2 & s \in [0,\lambda] \end{cases} \tag{8}$$

和 $\varphi_\lambda(t) = \xi \psi_\lambda(t-\tau)$. 注意到 $\|\varphi_\lambda\| = O(\lambda^2)$. 若 u^* 是强极小点, 当 $\lambda > 0$ 充分小时有 $I(u+\varphi_\lambda) - I(u) \geq 0$.

$$\begin{aligned} 0 &\leq \int_J (L(t,u(t)+\varphi_\lambda(t),\dot u(t)+\dot\varphi_\lambda(t)) - L(t,u(t),\dot u(t)))\,\mathrm{d}t \\ &= \int_J (L(t,u(t)+\varphi_\lambda(t),\dot u(t)+\dot\varphi_\lambda(t)) \\ &-L(t,u(t),\dot u(t)) - \varphi_\lambda(t)\cdot L_u(t,u(t),\dot u(t)) - \dot\varphi_\lambda(t)\cdot L_p(t,u(t),\dot u(t)))\,\mathrm{d}t \\ &= \int_J F(t)\,\mathrm{d}t. \end{aligned}$$

注意到 F 的具体表达式, 我们将上述积分拆成两个部分:

$$\int_J F(t)\,\mathrm{d}t = \left(\int_{\tau-\lambda^2}^{\tau} + \int_\tau^{\tau+\lambda}\right) F(t)\,\mathrm{d}t.$$

一方面, 当 $t \in [\tau, \tau+\lambda]$ 时, 注意到 $\|\varphi_\lambda|_{[\tau,\tau+\lambda]}\|_C = O(\lambda)$ 以及 $\|\dot\varphi_\lambda|_{[\tau,\tau+\lambda]}\|_C = O(\lambda), \lambda \to 0$. 故由 Taylor 展开即得 $F|_{[\tau,\tau+\lambda]} = O(\lambda^2) = o(1)$, 从而有

$$\lim_{\lambda\to 0} \frac{1}{\lambda^2} \int_\tau^{\tau+\lambda} F(t)\,\mathrm{d}t = 0.$$

另一方面, 我们有

$$\lim_{\lambda\to 0} \frac{1}{\lambda^2} \int_{\tau-\lambda^2}^{\tau} F(t)\,\mathrm{d}t = L(\tau,u(\tau),\dot u(\tau)+\xi) - L(\tau,u(\tau),\dot u(\tau)) - \xi\cdot L_p(\tau,u(\tau),\dot u(\tau)),$$

因此

$$L(\tau,u(\tau),\dot u(\tau)+\xi) - L(\tau,u(\tau),\dot u(\tau)) - \xi\cdot L_p(\tau,u(\tau),\dot u(\tau)) \geq 0.$$

上述不等式左侧的表达式等于 $\mathfrak{E}_L(t,u(t),\dot u(t),\dot u(t)+\xi)$. $\square$

1.3.2 充分条件: 极值场

基本思想: 对于一个给定的 C^1 函数 u, 我们将其 "嵌入" 至一组 "极值曲线" 中, 这组 "极值曲线" 具有 "统一的方向". 注意到对所有的 $\varphi \in C_0^1(J)$, u 与 $u+\varphi$ 具有相同的起点和终点, 如果能证明积分与道路的无关性, 我们可以将差值 $I(u+\varphi)-I(u)$ 的表达式进一步地简化, 从而得到较为简洁的充分条件.

设 u^* 是 E-L 方程的解. 称 u^* 对应的图像 $\{(t,u^*(t))\colon t\in J\}$ 为一条**极值曲线**. 现设 u^* 可以延拓到更大的区间 $J_1=(a,b)\supseteq J$ 上, 又设 $\{(t,u(t,\alpha))\colon t\in J_1, \alpha\in B_{\varepsilon_1}(0), \varepsilon_1>0\}$ 是 I 的一族足够光滑的极值曲线.

定义 1.23. *设 Ω 是 $\{(t,u(t,\alpha))\colon t\in J_1,\alpha\in B_\varepsilon(0)\}$ $(0<\varepsilon<\varepsilon_1)$ 的一个单连通的开邻域, $\psi=\psi(t,u)\in C^1(\Omega)$ 是向量场. 如果*

- *对任意的 $u=u(t,\alpha)$, u 满足方程 $\partial_t u=\psi(t,u)$;*
- *$\det(\partial_{\alpha_i}u_j(t,\alpha))\neq 0$;*
- *对任意的 $(t_1,u_1)\in\Omega$, 存在唯一的 $\alpha_1\in B_{\varepsilon_1}(0)$ 使得 $u(t_1,\alpha_1)=u_1$;*
- *$u(t,0)=u^*(t)$,*

*那么称 Ω 为一个**极值场 (或临界场)**, 并称 ψ 为其**方向场 (或流)**.*

例 1.24. *设 $L_1=p^2/2$, 则*

$$\Omega_1=\{(t,mt+\lambda)\colon (t,\lambda)\in\mathbb{R}\times\mathbb{R}\},\psi_1(t,u)=m$$

分别是 L_1 的一个极值场和方向场. 再考虑 $L_2=(p^2-u^2)/2$. 对任意的开区间 O, 令

$$\Omega_2=\{(t,\sin(t+\lambda))\colon (t,\lambda)\in O\times(-1,1)\}.$$

虽然极值曲线充满了整个 Ω_2, 但 Ω_2 中每一点都有两个极值曲线通过, 所以 Ω_2 不是极值场.

Analysis: 设极值曲线 $\gamma^*=\{(t,u^*(t))\colon t\in J\}\subseteq\Omega$ 满足方程 $\dot u=\psi(t,u)$, 其中 ψ 是极值场 Ω 上的一个方向场. 我们选取与 γ^* 邻近的, 端点相同的 C^1 曲线 $\gamma=\{(t,u(t))\colon t\in J\}$ 作比较. 将在区间上的积分转化为转化为路径积分, 如果积分与路径无关,那么就有

$$\begin{aligned}
I(u^*) &= \int_J L(t,u^*,\dot u^*)\,\mathrm{d}t \\
&= \int_{\gamma^*}(L(t,u^*,\psi(t,u^*))-\psi(t,u^*)\cdot L_p(t,u^*,\psi(t,u^*))\,\mathrm{d}t+L_p(t,u^*,\psi(t,u^*)))\,\mathrm{d}u \\
&= \int_{\gamma}(L(t,u,\psi(t,u))-\psi(t,u)\cdot L_p(t,u,\psi(t,u))\,\mathrm{d}t+L_p(t,u,\psi(t,u)))\,\mathrm{d}u \\
&= \int_J (L(t,u,\psi(t,u))+(\dot u-\psi(t,u))\cdot L_p(t,u,\psi(t,u)))\,\mathrm{d}t,
\end{aligned}$$

于是

$$\begin{aligned}I(u)-I(u^*)&=\int_J\big(L(t,u,\dot u)-L(t,u,\psi(t,u))-(\dot u-\psi(t,u))\cdot L_p(t,u,\psi(t,u))\big)\,\mathrm{d}t\\&=\int_J\mathfrak{E}_L(t,u,\psi(t,u),\dot u)\,\mathrm{d}t.\end{aligned}$$

因此, 若对任意的 $(t,u,p)\in\Omega\times\mathbb{R}^N$ 有 $\mathfrak{E}_L(t,u,\psi(t,u),p)\geq 0$, 那么 u^* 是一个强极小点.

在上述分析中用到了两个事实:

1. **u^* 对应的极值曲线 γ^* 可以嵌入到一个极值场中**. 具体地, 所谓 "嵌入" 是指, 存在开区间 $J_1\supseteq J$ 以及 $u=u(t,\alpha)\in C^1(J_1\times B_\varepsilon)$, 使得对任意的 $\alpha\in B_\varepsilon, u(t,\alpha)$ 是一条极值曲线, 其中 $u^*(t)=u(t,0)|_J$, 而且 $\{(t,u(t,\alpha))\colon t\in J_1,\alpha\in B_\varepsilon(0)\}$ 是一个极值场.

2. **积分与路径无关**.

以下我们验证这两个事实, 首先是第一个.

命题 1.25. 设 $L\in C^3(J\times\mathbb{R}^N\times\mathbb{R}^N), u^*\in C^2(J)$ 是其 *E-L* 方程的一个解. 又设 L 沿 u^* 满足严格 *Legendre-Hadamard* 条件. 如果沿对应于 u^* 的极值曲线 γ^* 没有共轭点, 那么 γ^* 可以嵌入到一族极值曲线中, 并且由这族曲线确定的单连通区域 Ω 是一个极值场.

证明. 先验证 $N=1$ 的情形. 首先, 由题设条件可知, u^* 可以延拓到更大的区间 $J_1=(a,b)\supseteq J$ 上. 相对于 $\alpha\in\mathbb{R}$, 当 $|\alpha|<\varepsilon_0$ 充分小时, 考虑初值问题

$$\begin{cases}E_L(\varphi(\cdot,\alpha))=0,\\\varphi(a,\alpha)=u^*(a),\\\varphi_t(a,\alpha)=\dot u^*(a)+\alpha.\end{cases}$$

由此我们得到一族解 $\{\varphi(t,\alpha)\}$, 其中 $t\in J_1,|\alpha|<\varepsilon_0$. 再根据初值问题的唯一性, $\varphi(t,0)=u^*(t)$.

现定义

$$\Omega_\varepsilon=\{(t,\varphi(t,\alpha))\colon t\in J_1,|\alpha|<\varepsilon\},$$

其中 $\varepsilon<\varepsilon_1$. 通过直接计算可知,

$$\xi(t)=\partial_\alpha\varphi(t,\alpha)|_{\alpha=0}$$

是一个沿 u^* 的 Jacobi 场, 同时 $\xi(a)=0,\dot\xi(a)=1$. 注意到 u^* 没有共轭点, 故我们可以选取合适的 a 和 b, 使得 $\xi(t)>0,t\in(a,b)\supseteq J_1$. 根据微分方程对于初值的连续依赖性, 故存在 $0<\varepsilon_1<\varepsilon_0$, 使得

$$0<\partial_\alpha\varphi(t,\alpha)\neq 0,\qquad\forall|\alpha|<\varepsilon,t\in J_1.$$

利用上述关系,我们还可以引用隐函数定理,故存在 $0<\varepsilon_2<\varepsilon_1$,使得对任意的 $(t,u)\in\Omega_{\varepsilon_2}$,方程 $u=\varphi(t,\alpha)$ 存在唯一的 C^1 解 $\alpha=w(t,u)\in B_{\varepsilon_2}(0)$. 今令 $\Omega=\Omega_{\varepsilon_2}$, 显然 $\gamma^*\in\Omega$ 且 Ω 是单连通的. 最后我们只需寻找 Ω 对应的方向场 ψ. 事实上, 令

$$\psi(t,u)=\partial_t\varphi(t,w(t,u)),$$

则 ψ 在 Ω 内处处有定义, 并且当 $u=\varphi(t,\alpha)$ 时, 有

$$\dot{u}=\partial_t\varphi(t,\alpha)=\partial_t\varphi(t,w(t,u))=\psi(t,u).$$

由此表明 ψ 是 Ω 的一个方向场. 这便完成了 $N=1$ 的情形的证明.

$N>1$ 的情形是类似的. 对模长充分小的 $\alpha\in\mathbb{R}^N$, 考虑满足初值条件

$$\partial_{\alpha_i}\varphi_j(a,\alpha)=0,\quad \partial_{\alpha_i}\partial_t\varphi_j(a,\alpha)=\delta_{ij},\qquad 1\le i,j\le N$$

E-L 方程的解. 再令 $w_i(t)=\partial_{\alpha_i}\varphi(t,\alpha)|_{\alpha=0}, i=1,\cdots,N$. 可以验证 w_i 是一个 Jacobi 场. 注意到 $w_i(a)=0,\partial_t w_i(a)=e_i, i=1,\cdots,N$, 且 u^* 没有共轭点, 故总可以找到一个充分小的 $\varepsilon^*>0$, 使得 $\det(\partial_{\alpha_i}\varphi_j(t,\alpha))\neq 0,\forall(t,\alpha)\in J_1\times B_{\varepsilon^*}(0)$. 其余部分的证明是相同的. □

现在考虑第二个事实. 定义

$$\begin{cases} R_i(t,u)=L_{p_i}(t,u,\psi(t,u)),\\ H(t,u)=\psi(t,u)\cdot L_p(t,u,\psi(t,u))-L(t,u,\psi(t,u)) \end{cases}$$

以及 1-形式

$$\boxed{\omega=\sum_{i=1}^N R_i\,\mathrm{d}u_i-H\,\mathrm{d}t}.$$

称 ω 为 **Hilbert 积分不变量**. 显然, ω 是闭形式 $\Rightarrow$ 积分与路径无关. 以下引入更多概念来刻画这一条件.

定义 1.26. 称极值场 Ω 是 ***Mayer 场***, 如果它满足如下相容性条件:

$$\boxed{\partial_{u_i}L_{p_j}(t,u(t),\psi(t,u(t)))=\partial_{u_j}L_{p_i}(t,u(t),\psi(t,u(t))),\qquad \forall 1\le i,j\le N.}$$

命题 1.27 (Mayer 场的等价刻画)**.** *若 $L\in C^2(J\times\mathbb{R}^N\times\mathbb{R}^N)$, 那么 (Ω,ψ) 是一个 Mayer 场, 当且仅当 $\mathrm{d}\omega=0$, 即 $\partial_t R_i=-\partial_{u_i}H, 1\le i\le N$.*

证明. 记 $\widetilde{L}=\widetilde{L}(t)=L(t,u(t),\psi(t,u(t)))$. 类似地记 $\widetilde{L_{u_i}},\widetilde{L_{p_i}}$. 利用条件 $\dot{u}(t)=\psi(t,u(t))$ 和 E-L 方程 $\widetilde{L_u}=\partial_t\widetilde{L_p}$, 可以得到 $D_\psi\widetilde{L_p}=\widetilde{L_u}$, 其中

$$D_\psi=\partial_t+\sum_{i=1}^N\psi_i\partial_{u_i}+\sum_{i=1}^N\left(\partial_t\psi_i+\sum_{k=1}^N\psi_k\partial_{u_k}\psi_i\right)\partial_{p_i}.$$

再通过直接计算可得

$$\left.\begin{aligned}\partial_t R_i &= \left(\partial_t + \sum_{j=1}^N \partial_t \psi_j \partial_{p_j}\right) \widetilde{L_{p_i}},\\ \partial_{u_i} H &= \sum_{j=1}^N \psi_j \partial_{u_i} \widetilde{L_{p_j}} - \widetilde{L_{u_i}},\end{aligned}\right. \qquad \forall 1 \le i \le N.$$

可以验证, 相容性条件成立 $\Leftrightarrow \partial_t R_i + \partial_{u_i} H = 0 = D_\psi \widetilde{L_{p_i}} - \widetilde{L_{u_i}}$. □

综上所述, **若 u^* 对应的极值曲线 γ^* 能够嵌入到一个 Mayer 场中**, 则上述两个事实均成立.

注 1.28. 由上述分析可知, 对于给定的 *Mayer* 场 (Ω,ψ), 若设 γ 是连接 $(t_0,u^*(t_0))$ 与 $(t,u^*(t))$ 的任意一条曲线 $(t_0 \le t \le t_1)$, 则线积分

$$\boxed{S(t,u) = \int_\gamma L_p\,\mathrm{d}u + (L - \psi\cdot L_p)\,\mathrm{d}t}$$

与 γ 无关. 称此积分为 ***Hilbert* 不变积分**.

命题 1.29. 设 I 的 *E-L* 方程的解 u^* 对应的极值曲线 γ^* 能嵌入到一族曲线中去, 且这族曲线可以定义一个 *Mayer* 场 (Ω,ψ). 若

$$\mathfrak{E}(t,u,\psi(t,u),p) \ge 0, \qquad \forall (t,u,p) \in \Omega\times\mathbb{R}^N,$$

则 u^* 是 I 的一个强极小点.

注意到**当 $N=1$ 时, 任何极值场都是 Mayer 场**, 故我们对上述充分条件有着更精准的刻画:

命题 1.30 (充分条件, $N=1$)**.** 设 $L\in C^3$, 并设其 *E-L* 方程的解 u^* 没有共轭点. 设 (Ω,ψ) 为关于 u^* 的极值场. 若 L 沿 u^* 满足严格 *Legendre-Hadamard* 条件, 则 u^* 是 I 的强极小点.

证明. 注意到对任意的 $(t,u,p)\in\Omega\times\mathbb{R}$, 我们有

$$\begin{aligned}\mathfrak{E}(t,u,\psi(t,u),p) &= L(t,u,p) - L(t,u,\psi(t,u)) - (p-\psi(t,u))L_p(t,u,\psi(t,u))\\ &= L_{pp}(t,u,v) \ge 0,\end{aligned}$$

其中 v 介于 p 和 $\psi(t,u)$ 之间. 由此足以说明 u^* 是 I 的一个强极小点. □

事实上, 对于高维的情形, 类似于命题1.30的结论也是成立的.

Analysis: 给定 Lagrange 函数 L 和一族极值曲线 $\{(t,\varphi(t,\alpha))\}\subseteq\Omega$, 其中 Ω 是对应的极值场. 记 $\overline{L}(t,\alpha) = L(t,\varphi(t,\alpha),\dot\varphi(t,\alpha))$, 其中 $\dot\varphi(t,\alpha) = \partial_t\varphi(t,\alpha)$. 类似地记 $\overline{L_{u_i}},\overline{L_{p_i}}$. 直接计算得

$$\mathrm{d}\omega = \sum_{i,\ell=1}^N (\overline{L_{u_i}} - \partial_t \overline{L_{p_i}})\partial_{\alpha_\ell}\mathrm{d}\alpha_\ell\wedge\mathrm{d}t + \sum_{m,i,\ell=1}^N \partial_{\alpha_m}\overline{L_{p_i}}\partial_{\alpha_\ell}\varphi_i\mathrm{d}\alpha_m\wedge\mathrm{d}\alpha_\ell.$$

现引入 **Lagrange 括号**

$$\boxed{[\alpha_\ell,\alpha_m]:=\sum_{i=1}^{N}(\partial_{\alpha_\ell}\overline{L_{p_i}}\partial_{\alpha_m}\varphi_i-\partial_{\alpha_m}\overline{L_{p_i}}\partial_{\alpha_\ell}\varphi_i).}$$

从而有

$$\mathrm{d}\omega=\sum_{i,\ell=1}^{N}E_L(\varphi)_i\partial_{\alpha_\ell}\varphi_i\mathrm{d}\alpha_\ell\wedge\mathrm{d}t+\sum_{1\le\ell<m\le N}[\alpha_\ell,\alpha_m]\mathrm{d}\alpha_\ell\wedge\mathrm{d}\alpha_m.$$

结合对 Mayer 场的等价刻画, 我们有

引理 1.31. *设 $L\in C^3(\Omega\times\mathbb{R}^N\times\mathbb{R}^N)$. 又设 (Ω,ψ) 是由一族极值曲线 $\{\varphi(t,\alpha)\}$ 决定的极值场. 则 (Ω,ψ) 是一个 Mayer 场, 必须且只需*

$$E_L(\varphi(\cdot,\alpha))=0,\quad\forall\alpha\in\mathbb{R}^N\quad\text{和}\quad[\alpha_\ell,\alpha_m]=0,\quad\forall 1\le\ell,m\le N.$$

此外, 对于 Lagrange 括号, 有如下结果:

引理 1.32. *设 $L\in C^3(J\times\mathbb{R}^N\times\mathbb{R}^N)$, 且 (Ω,ψ) 是一个极值场, 则*

$$\partial_t[\alpha_\ell,\alpha_m]=0,\qquad\forall 1\le\ell,m\le N.$$

证明. 利用 E-L 方程, 我们有

$$\partial_t[\alpha_\ell,\alpha_m]=\sum_{i=1}^{N}\left(\partial_{\alpha_\ell}\overline{L_{u_i}}\partial_{\alpha_m}\varphi_i+\partial_{\alpha_\ell}\overline{L_{p_i}}\partial_{\alpha_m}\dot{\varphi}_i-\partial_{\alpha_m}\overline{L_{u_i}}\partial_{\alpha_\ell}\varphi_i-\partial_{\alpha_m}\overline{L_{p_i}}\partial_{\alpha_\ell}\dot{\varphi}_i\right).$$

将上式中的 $\partial_{\alpha_\ell}\overline{L_{u_i}},\partial_{\alpha_m}\overline{L_{u_i}},\partial_{\alpha_\ell}\overline{L_{p_i}},\partial_{\alpha_m}\overline{L_{p_i}}$ 展开, 即证得所需结论. □

命题 1.33 (充分条件, $N>1$)**.** *设 $L\in C^3(J\times\mathbb{R}^N\times\mathbb{R}^N)$. 如果其对应的 E-L 方程的解 u^* 没有共轭点, L 沿 u^* 满足严格 Legendre-Hadamard 条件, 且*

$$\mathfrak{E}(t,u,\psi(t,u),p)\ge0,\quad\forall(t,u,p)\in\Omega\times\mathbb{R}^N.$$

那么 u^* 是 I 的强极小点.

证明. 在命题1.25的证明过程中我们构造了一个极值场 (Ω,ψ), 若能证明此极值场是 Mayer 场, 那么命题的结论成立. 事实上, 注意到初值条件 $\partial_{\alpha_i}\varphi_j(a,\alpha)=0,\forall i,j$, 则有 $[\alpha_\ell,\alpha_m](a,\alpha)=0,\forall\ell,m$. 再根据引理1.32, 则有 $[\alpha_\ell,\alpha_m]=0$. 这表明 (Ω,ψ) 是一个 Mayer 场. □

例 1.34. 设

$$I(u)=\int_1^2(\dot{u}+t^2\dot{u}^2)\,\mathrm{d}t,\quad u\in M,$$

其中 $M=\{u\in C^1[1,2]: u(1)=1,u(2)=2\}$. 验证

$$u^*(t)=-\frac{2}{t}+3$$

是 I 的强极小点.

解. 直接计算得 $L_u = L_{uu} = L_{pu} = 0, L_p = 1+2t^2 p, L_{pp} = 2t^2$. 故其对应的 E-L 方程为

$$\frac{\mathrm{d}}{\mathrm{d}t}(1+2t^2 p) = 0.$$

显然 u^* 是满足 E-L 方程和边值条件的解, 且对应的 Jacobi 方程 $t^2\dot{\varphi}(t) = \text{const}$ 存在正解. 现令

$$\Omega = \left\{ \left(t, -\frac{2}{t} + \alpha \right) : (t,\alpha) \in (0,+\infty) \times \mathbb{R} \right\}, \quad \psi(t) = \frac{2}{t^2}.$$

可以验证, (Ω, ψ) 是一个极值场, 且包含了 u^* 所对应的极值曲线. 注意到 $L_{pp}(t,u,p) > 0, \forall (t,u,p) \in \Omega \times \mathbb{R}$, 因此由命题1.30可知, u^* 是 I 的一个强极小点. □

1.4 Hamilton-Jacobi 理论

1.4.1 Hamilton 方程组

Hamilton 方程组:

$$\begin{cases} \dot{\xi} = -\partial_u H \\ \dot{u} = \partial_\xi H. \end{cases}$$

其中 $H = H(t, u(t), \xi(t))$. 利用 Legendre 变换, 上述方程组可以从变分的角度导出, 且其与 E-L 方程法有着深刻的联系.

定义 1.35. *设 X 是赋范线性空间, 函数 $\varphi\colon X \to (-\infty, +\infty]$ 满足*

$$\{x \in X\colon \varphi(x) < +\infty\} \neq \varnothing.$$

称函数

$$\boxed{\varphi^*\colon X^* \to (-\infty,+\infty], f \mapsto \sup_{x\in X}[\langle f, x\rangle - f(x)]}$$

为 φ 的 ***Legendre 变换*** *(或称为 φ 的***共轭函数***), 其中 X^* 代表 X 的对偶空间.*

以上是 Legendre 变换的一般定义, 在这里我们只考虑标准欧式空间的情形: 若函数 $f \in C^2(\mathbb{R}^N)$ 且其梯度 $\xi = \nabla f(x)$ 有逆映射 ψ, 则 f 的 Legendre 变换有表达式

$$\boxed{f^*(\xi) = \xi\cdot x - f(x) = \xi \cdot \psi(\xi) - (f\circ\psi)(\xi).}$$

注 1.36. *Legendre 变换有着如下的几何意义: 记*

$$G(f) = \{(x,y) \in \mathbb{R}^N \times \mathbb{R}\colon y = f(x)\}$$

为 f 的图像, 它在点 $P = (x,y)$ 处的切平面

$$S = \{(\alpha,\beta)\colon \beta - f(x) = \nabla f(x)\cdot(\alpha - x)\}.$$

因此, 任取 S 上的一点 $Q = (\alpha,\beta)$, 我们有

$$\beta - \nabla f(x)\cdot\alpha = f(x) - \nabla f(x)\cdot x,$$

从而有 $\beta - \xi\cdot\alpha = -f^*(\xi)$. 这表明 $-f^*(\xi)$ 是超平面 S 在 β 轴上的截距.

命题 1.37. *Legendre 变换有着如下简单性质:*

1. *若 $f \in C^k$, 则 $f^* \in C^k$;*
2. *$f^{**} = f$, 即 Legendre 变换是***自反的***;*
3. *$\left.(\partial_{\xi_i\xi_j} f^*(\xi))\right|_{\xi=\nabla f(x)} = (\partial_{x_i x_j} f(x))^{-1}$.*

证明. 1. 注意到 $x = \psi(\xi) \in C^{k-1}$, 因此 $f^* \in C^{k-1}$. 另一方面, 注意到等式

$$\nabla f^*(\xi) = \psi(\xi) + \xi \cdot \nabla\psi(\xi) - \nabla f(\psi(\xi)) \cdot \psi(\xi) = \psi(\xi),$$

由此表明 $f^* \in C^k$.

2. 由 1 的证明过程可知 $x = \psi(\xi) = \nabla f^*(\xi)$, 故有

$$f^{**}(x) = \xi \cdot x - f^*(\xi) = f(x).$$

3. 注意到 $x = (\nabla f^*)(\nabla f(x))$, 等式两边取梯度, 即得

$$I_N = \left.(\partial_{\xi_i\xi_j} f^*(\xi))\right|_{\xi=\nabla f(x)} \cdot (\partial_{x_i x_j} f(x)).$$

□

上述命题中展现了一些对称性的结果:

$$\boxed{f(x) + f^*(\xi) = \xi \cdot x, \quad \xi = \nabla f(x), \quad x = f^*(\xi).}$$

例 1.38. 设 $f(x) = x^p/p, x \geq 0, p > 1$. 直接计算得 $f^*(\xi) = \xi^{p'}/p', \xi \geq 0$, 其中 $1/p + 1/p' = 1$. 注意到 *Legendre* 变换的定义, 我们便有

$$x\xi \leq \frac{1}{p}x^p + \frac{1}{p'}\xi^{p'}, \qquad x, \xi \geq 0.$$

这便是经典的 *Young* 不等式.

以下我们利用 Legendre 变换来推导 Hamilton 方程组. 给定 $L \in C^2(\mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^N)$, 并假设 $\det(L_{p_i p_j}(t,u,p)) \neq 0$. 令 $\xi = \xi = L_p(t,u,p)$, 根据隐函数定理, 我们可以局部地解出 $p = \varphi(t,u,\xi)$, 即

$$p_i = \varphi_i(t,u,\xi), \quad 1 \leq i \leq N.$$

现固定 (t,u), 将 L 看作是 p 的函数, 并对其 (关于 p) 作 Legendre 变换:

$$\boxed{H(t,u,\xi) := L^*(t,u,\xi) = \left.(\xi \cdot p - L(t,u,p))\right|_{p=\varphi(t,u,\xi)}.}$$

称 H 是 **Hamilton 函数** (Hamiltonian). 从上述分析中可以看出, H 无非是 L 的 Legendre 变换. 注意到 Legendre 变换是自反的, 故 L 也可以看作是 H 的 Legendre 变换.

接下来考虑 L 对应的 E-L 方程, 并将其改写成方程组的形式:

$$\begin{cases}\dot{u}(t)=p(t),\\ \dfrac{\mathrm{d}}{\mathrm{d}t}L_p(t,u(t),p(t))-L_u(t,u,p(t))=0,\end{cases}$$

设其解为 $(u(t),p(t))$. 在等式 $H(t,u,\xi)=(\xi\cdot p-L(t,u,p))|_{p=\varphi(t,u,\xi)}$ 两边作微分, 即得

$$H_t\,\mathrm{d}t+H_u\cdot\mathrm{d}u+H_\xi\cdot\mathrm{d}\xi=-L_t\mathrm{d}t-L_u\cdot\mathrm{d}u+p\cdot\mathrm{d}\xi,$$

从而有

$$H_t=-L_t,\quad H_u=-L_u,\quad H_\xi=p.$$

若令 $\xi(t)=L_p(t,u(t),p(t))$, 结合上述等式, 便有

$$\dot{\xi}(t)=L_u(t,u(t),\dot{u}(t))=L_u(t,u(t),p(t))=-H_u(t,u(t),\xi(t))$$

与 $\dot{u}(t)=p(t)=H_\xi(t,u(t),\xi(t))$. 综上所述, 我们有

$$\boxed{\begin{cases}\dot{\xi}(t)=-H_u(t,u(t),\xi(t)),\\ \dot{u}(t)=H_\xi(t,u(t),\xi(t)).\end{cases}}$$

这便是经典的 **Hamilton 方程组**, 简称 H-S.

注 1.39. *在上述分析中, 我们先假设 $(u(t),p(t))$ 是 E-L 方程的解, 从而推导出 $(u(t),\xi(t))$ 是 H-S 的解. 反之, 对于给定的 H-S 的解 $(u(t),\xi(t))$, 注意到关系 $p(t)=\dot{u}(t)$ 和 $\xi(t)=L_p(t,u(t),p(t))$, 我们有*

$$\frac{\mathrm{d}}{\mathrm{d}t}L_p(t,u(t),\dot{u}(t))=\dot{\xi}(t)=-H_u(t,u(t),\xi(t))=L_u(t,u(t),\dot{u}(t)).$$

这表明 $(u(t),p(t))$ 是 E-L 方程的解. 综上, 我们得到了 E-L 方程和 H-S 二者之间的一个一一对应关系.

以下考察 H-S 对应的变分积分. 可以验证, H-S 是泛函

$$\boxed{F(u,\xi)=\int_J(\dot{u}\cdot\xi-H)\,\mathrm{d}t}$$

所对应的 E-L 方程, 相应的 1-形式是

$$\boxed{\alpha=\xi\cdot\mathrm{d}u-H\,\mathrm{d}t,}$$

我们称其为 **Poincaré-Cartan 积分不变量**. 注意到 H 是 L 的 Legendre 变换, 因此泛函 F 与 I 表达式中的被积函数实际上是**同一个函数在不同变量下的表示**, 而 Poincaré-Cartan 积分不变量则就是上一节中提到的 Hilbert 积分不变量.

注 1.40. *Hamilton* 方程组对应的泛函 F 不是下方有界的, 从而没有最小值, 因此 *H-S* 的解是相应泛函的 "临界点". 在实际问题中, 我们视实际情况从而选用 *E-L* 方程或 *H-S* 进行求解.

例 1.41. 对于有 n 个自由度的质点组, 记位置坐标 $q=(q_1,\cdots,q_N)$, 则其动能 $T=T(q)=\sum_{i,j=1}^N a_{ij}\dot q_i\dot q_j/2$, 其中 (a_{ij}) 是正定阵. 设位能 $V=V(q)$, 则 *Lagrange* 函数 $L=T-V$. 通过直接计算可知, L 对应的 *E-L* 方程为

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{j=1}^n a_{ij}\dot q_j=-\partial_{q_i}V(q)\quad (i=1,\cdots,N).$$

此时 *Hamilton* 函数

$$H(q,\xi)=\frac12\sum_{i,j=1}^N a^{ij}\xi_i\xi_j+V(q)$$

是这个质点组的能量, 其中 $(a^{ij})=(a_{ij})^{-1}$. 对应的 *Hamilton* 方程组为

$$\begin{cases}\dot\xi_i=-\partial_{q_i}V(q),\\ \dot q_i=\sum_{j=1}^N a^{ij}\xi_j,\end{cases}\qquad i=1,\cdots,N.$$

一般地, 若设 $(u(t),p(t))$ 是 *E-L* 方程的解, 直接计算得

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}H(u(t),\xi(t))=0.}$$

即 ***Hamilton* 方程组的解曲线都保持在同一个等值面上**. 将此结果运用到上述分析中, 即得: 在运动过程中质点组的能量守恒.

1.4.2 Hamilton-Jacobi 方程

对于给定的 Hamilton 函数 $H=H(t,u,\xi)$, 称一阶偏微分方程

$$\boxed{\partial_t S(t,u)+H(t,u,\nabla_u S(t,u))=0}$$

为 **Hamilton-Jacobi 方程** (简称 **H-J 方程**), 其中 $S=S(t,u)$ 是定义在 $\mathbb{R}^{1+N}$ 上的函数. 以下我们结合 Mayer 场和 Legendre 变换等概念, 导出此方程, 并探究其与 H-S 之间的联系.

先引入一些概念. 给定一个 Lagrange 函数 L. 对于一个极值场 (Ω,ψ), 其上有一个对应的 1-形式, 即 Hilbert 不变积分因子:

$$\omega=L_p(t,u,\psi(t,u))\cdot\mathrm{d}u-(\psi(t,u)\cdot L_p(t,u,\psi(t,u))-L(t,u,\psi(t,u)))\mathrm{d}t.$$

我们已经知道, (Ω,ψ) 是一个 Mayer 场, 当且仅当 ω 是闭的. 因此, 在一个 Mayer 场上我们可以由 ω 定义出 Ω 上的一个单值函数 g:

$$\boxed{g(t,u):=g(t_0,u_0)+\int_\gamma\omega,}$$

其中 γ 是连接 (t_0,u_0) 与 (t,u) 的任意一条曲线. 我们称此单值函数 g 为**程函**. 此外, 由程函定义可知, g 满足如下方程组:

$$\boxed{\begin{cases}\nabla_u g(t,u)=L_p(t,u,\psi(t,u)),\\ \partial_t g(t,u)=L(t,u,\psi(t,u))-\psi(t,u)\cdot L_p(t,u,\psi(t,u)).\end{cases}}$$

称此方程组为 **Carathéodory 方程组**.

例 1.42. 设 $\gamma=(t,u(t))\subseteq(\Omega,\psi)$ 是一条极值曲线. 直接计算得

$$g(t_2,u(t_2))-g(t_1,u(t_1))=\int_\gamma\omega=\int_J L(t,u(t),\dot u(t))\,\mathrm{d}t.$$

由此表明, **程函在同一极值曲线上两点的差等于 *Lagrange* 函数沿这条曲线的积分**. 在光学中, *Lagrange* 函数表示光在传播中瞬时走过的路程除以速度, 沿这条曲线的积分就等于光线从 $(t_1,u(t_1))$ 传播到 $(t_2,u(t_2))$ 所经历的时间. 由上述分析可知, 程函的等值面 $\{g(t,u)=\mathrm{const}\}$ 可以用来表示从一点发出的一束光线的等时面, 即波阵面.

以下我们导出 H-J 方程. 给定 Mayer 场 (Ω,ψ), 其对应的程函 g 满足 Carathéodory 方程组. 将 $\xi=L_p(t,u,\psi(t,u))$ 代入至 Carathéodory 方程组中, 即得

$$\partial_t g(t,u)+H(t,u,\nabla_u g(t,u))=0,$$

其中 H 是 L 的 Legendre 变换, 即 Hamilton 函数. 这表明 g 满足 H-J 方程.

注 1.43. 注意到 H 是 L 的 *Legendre* 变换, 对于给定的 *Hamilton* 函数 H 和对应的 *H-S* 的一组解 $(u(t),\xi(t))$, 令 $p(t)=H_\xi(t,u(t),\xi(t))$, 则 $(u(t),p(t))$ 便是 *E-L* 方程的解. 再利用等式

$$L(t,u(t),p(t))=\dot u(t)\cdot\xi(t)-H(t,u(t),\xi(t)),$$

我们便可以写出 *Lagrange* 函数 L. 因此

$$g(t,u)=g(t_0,u(t_0))+\int_{t_0}^t L(t,u(t),p(t))\,\mathrm{d}t$$

便是 *H-J* 方程的一个解. 上述分析表明, 我们可以从任取初值得到的 *H-S* 的所有解导出 *H-J* 的解.

事实上, 我们也可以从 H-J 方程的解导出 H-S 的解.

定义 1.44. 设 $g=g(t,u;\lambda_1,\cdots,\lambda_N)$ 是 *H-J* 方程一族依赖于 N 个参数 $(\lambda_1,\cdots,\lambda_N)\in\Lambda$ 的解, 其中 $\Lambda\in\mathbb{R}^N$ 是一个区域. 如果 $\det(g_{u_i\lambda_j})\neq0$, 那么称 g 为一个**完全积分**.

定理 1.45 (Jacobi). 设 C^2 函数 $g=g(t,u;\lambda_1,\cdots,\lambda_N)$ 是 *H-J* 方程的一个完全积分. 若依赖于 $2N$ 个参数 $(\alpha,\beta)=(\alpha_1,\cdots,\alpha_N,\beta_1,\cdots,\beta_N)$ 的函数

$$\begin{cases} u=U(t,\alpha,\beta), \\ p=P(t,\alpha,\beta) \end{cases}$$

满足方程

$$\begin{cases} g_{\alpha_i}(t,U(t,\alpha,\beta),\alpha)=-\beta_i, \\ P_i(t,\alpha,\beta)=g_{u_i}(t,U(t,\alpha,\beta),\alpha), \end{cases} \qquad i=1,\cdots,N, \tag{9}$$

那么 (U,P) 便是 *H-S* 的一族解.

证明. 先在 H-J 方程的两端对 α_i 求偏导, 即得

$$g_{t,\alpha_i}+\sum_{k=1}^{N}H_{\xi_k}(t,u,\nabla_u g)g_{u_k,\alpha_i}=0, \quad i=1,\cdots,N.$$

将等式 $u=U(t,\alpha,\beta)$ 代入至上式, 并利用(9)的第二个等式, 便有

$$g_{t,\alpha_i}(t,U,\alpha)+\sum_{k=1}^{N}H_{\xi_k}(t,U,P)g_{u_k\alpha_i}(t,U,\alpha)=0. \tag{10}$$

再对(9)的第一个方程等式两边对 t 求偏导, 我们有

$$g_{t,\alpha_i}(t,U,\alpha)+\sum_{k=1}^{N}g_{\alpha_i u_k}(t,U,\alpha)\dot{U}_k(t,\alpha,\beta)=0. \tag{11}$$

联立(10)和(11), 并注意到 g 是完全积分, 从而有

$$\dot{U}_k(t,\alpha,\beta)=H_{\xi_k}(t,U(t,\alpha,\beta),P(t,\alpha,\beta)), \quad k=1,\cdots,N.$$

这是 H-S 的一组方程. 另一方面, 对 H-J 方程等式两边对 u_i 求偏导, 并将等式 $u=U(t,\alpha,\beta),P(t,\alpha,\beta)=\nabla_u g(t,U(t,\alpha,\beta),\alpha)$ 代入, 得

$$-H_{u_i}(t,U,P)=g_{t,u_i}(t,U,\alpha)+\sum_{k=1}^{N}g_{u_iu_k}(t,U,\alpha)\dot{U}_k(t,\alpha,\beta).$$

对(9)的第二个方程等式两边对 t 求偏导, 我们有

$$\dot{P}_i(t,\alpha,\beta)=g_{u_i,t}(t,U,\alpha)+\sum_{k=1}^{N}g_{u_iu_k}(t,U,\alpha)\dot{U}_k(t,\alpha,\beta).$$

从而我们得到

$$\dot{P}_k(t,\alpha,\beta)=-H_{u_k}(t,U,P), \quad k=1,\cdots,N.$$

这便是 H-S 的另一组方程. □

由 Jacobi 定理可知, 我们可以用 H-J 方程的解写出 H-S 的解. 具体方法如下: 设 g 是一个完全积分, 先解 N 个函数方程

$$g_{\alpha_i}(t,u,\alpha) = -\beta_i, \quad i=1,\cdots,N,$$

由此得到

$$u = U(t,\alpha,\beta). \tag{12}$$

因此

$$p = P(t,\alpha,\beta) = \nabla_u g(t,U(t,\alpha,\beta),\alpha). \tag{13}$$

从而 (u,p) 就是 H-S 的解.

注 1.46. 注意到完全积分与通解的意义是不同的. 若考虑 *H-J* 方程的 *Cauchy* 问题, 由唯一性可知, 它的通解里应该含有一个任意函数 $\varphi = \varphi(u)$, 而不仅仅是 $2N$ 个独立参数. 但是, *H-S* 初值问题的解可以由 *H-J* 方程的一个完全积分 g 所确定. 具体地, 我们考虑 *H-S* 方程的初值问题

$$\begin{cases} \dot{u} = H_\xi(t,u,\xi), \\ \dot{\xi} = -H_u(t,u,\xi), \\ u(0) = u_0, \xi(0) = \xi_0, \end{cases} \tag{14}$$

其中 u_0, ξ_0 是任意常数. 如果 g 是一个完全积分, 那么 $\det(g_{u_i\alpha_j}) \neq 0$, 我们因此可以使用隐函数定理对方程

$$\xi_0 = \nabla_u g(0,u_0,\alpha),$$

解出 $\alpha_0 = \alpha(u_0,\xi_0)$. 再令

$$\beta_0 = -\nabla_\alpha g(0,u_0,\alpha_0),$$

并将 (α_0,β_0) 作为初值代入至(12)和(13)中, 我们便得到了初值问题(14)的解.

1.4.3 例

例 1.47 (光在介质中的传播)**.** 设在介质中一点 $(t,u) \in \mathbb{R}\times\mathbb{R}^N$ 的介质密度是 $\rho = \rho(t,u)$. 以真空光速为单位, 若在此点的光速为 $1/\rho(t,u)$, 则对应的 *Lagrange* 函数为

$$L(t,u,p) = \rho(t,u)\sqrt{1+p^2}.$$

从而有

$$H(t,u,\xi) = -\sqrt{\rho(t,u)^2 - \xi^2}.$$

此时程函 g 满足 *H-J* 方程

$$\partial_t g = \sqrt{\rho^2 - |\nabla u|^2},$$

对应的方向场

$$\psi(t,u) = H_\xi(t,u,\nabla_u g) = \frac{\nabla_u g}{\sqrt{\rho^2 - |\nabla_u g|^2}} = \frac{\nabla_u g}{\partial_t g}.$$

即有

$$(\dot{t},\dot{u}) = (1,\dot{u}) = (\partial_t g)^{-1}(\partial_t g, \nabla_u g).$$

由上述结果可知, 积分曲线 (t,u) 空间中沿波阵面 $\{g(t,u) = \mathrm{const}\}$ 的法方向. 上述分析表明: 光线垂直于波阵面.

例 1.48 (简谐振动)**.** 给定 *Lagrange* 函数

$$L = \frac{1}{2}(mp^2 - ku^2),$$

其中 m 与 k 都是正常数. 通过直接计算可知, 其对应的 *Hamilton* 函数

$$H(t,u,p) = \frac{1}{2}\left(\frac{p^2}{m} + ku^2\right),$$

且对应的 *H-S*

$$\begin{cases} \dot{u} = \dfrac{p}{m}, \\ \dot{p} = -ku \end{cases} \tag{15}$$

有解

$$\begin{cases} u = C\sin\left(\sqrt{\dfrac{k}{m}}(t+t_0)\right), \\ p = C\sqrt{mk}\cos\left(\sqrt{\dfrac{k}{m}}(t+t_0)\right), \end{cases}$$

其中 t_0, C 是任意常数. 我们现在利用 *Jacobi* 定理, 通过 *H-J* 方程把(15)的解写出来. 考虑一个特殊的 $g(t,u,\alpha) = \varphi(u,\alpha) - \alpha t$, 其中 α 是一个参数, φ 是一个待定的函数. 将此代入至 *H-J* 方程中, 得

$$\frac{1}{2}\left(\frac{\varphi_u^2}{m} + ku^2\right) = \alpha,$$

即 $\varphi_u = \sqrt{m(2\alpha - ku^2)}$. 解出来有

$$g(t,u,\alpha) = \int_0^u \sqrt{m(2\alpha - kv^2)}\,\mathrm{d}v - \alpha t.$$

此时 $g_{\alpha u} = 2m \neq 0$. 现考虑方程

$$-\beta = g_\alpha(t,u,\alpha) = \sqrt{\frac{m}{k}}\arcsin\left(\sqrt{\frac{k}{2\alpha}}u\right) - t,$$

解得

$$u = \sqrt{\frac{2\alpha}{k}}\sin\left(\frac{k}{m}(t-\beta)\right).$$

将上式代入至 g 中, 即得

$$p = g_u = \sqrt{2\alpha m}\cos\left(\frac{k}{m}(t-\beta)\right).$$

这便是(15)带有两个参数 α, β 的解.

1.5 含多重积分的变分问题

在这一节中, 我们将前几节中所讨论的结果推广到高维的情形. 先引入如下记号:

$$
\begin{aligned}
x &= (x_i)_{1\le i\le n} = (x_1,\cdots,x_n) \in \mathbb{R}^n,\\
u &= (u^m)_{1\le m\le N} = (u_1,\cdots,u_N) \in \mathbb{R}^N,\\
p &= (p_i^m)_{1\le i\le n, 1\le m\le N} \in \mathbb{R}^{nN},\\
\nabla u &= (\partial_i u^m)_{1\le i\le n, 1\le m\le N} \in \mathbb{R}^{nN}.
\end{aligned}
$$

给定 $\mathbb{R}^n$ 中带有 C^1 边界的有界区域 Ω, Lagrange 函数 $L = L(x,u,p) \in C^2(\overline{\Omega}\times\mathbb{R}^N\times\mathbb{R}^{nN})$ 以及边界上的函数 $\Phi \in C^1(\partial\Omega)$. 考虑泛函

$$
I(u) = \int_\Omega L(x,u(x),\nabla u(x))\,\mathrm{d}x
$$

在边值条件 $u \in M = \{v \in C^1(\overline{\Omega}) : v|_{\partial\Omega} = \Phi\}$ 下的极小值.

1.5.1 Euler-Lagrange 方程

类似于单重积分的情形 $(n=1)$, 称 $u^* \in M$ 是 I 在 M 上的**极小点**, 如果

$$
I(u) \ge I(u^*), \qquad \forall u \in U \cap M,
$$

其中 U 是 u^* 在 M 中的一个邻域. 取 C^1 拓扑时, 称 u^* 为**弱极小点**; 取 C 拓扑时, 称 u^* 为**强极小点**.

引理 1.49 (du Bois-Reymond, 变分学基本引理)**.** *设 $u \in L^1_{\mathrm{loc}}(\Omega)$, 且对任意的 $\varphi \in C_c^\infty(\Omega)$ 有*

$$
\int_\Omega u(x)\varphi(x)\,\mathrm{d}x = 0,
$$

则 $u(x) = 0$, a.e. $x \in \Omega$.

证明. 设 $\{\eta_n\}$ 是一族光滑化子. 令 $g_n = g * \eta_n$, 其中 $g \in L^\infty(\mathbb{R}^n)$ 且 $\operatorname{supp} g \subseteq \Omega$. 显然 $g \in L^1(\mathbb{R}^n)$, 且当 n 充分大时, 有 $g_n \in C_c^\infty(\Omega)$, 从而

$$
\int_\Omega u(x)g_n(x)\,\mathrm{d}x = 0.
$$

此外, 由恒等逼近的理论可知, $\|g_n - g\|_{L^1} \to 0, n \to \infty$. 因此存在子列, 不妨记为 $\{g_n\}$, 使得 $g_n \to g$, a.e. 注意到

$$
\|g_n\|_{L^\infty} = \|g * \eta_n\|_{L^\infty} \le \|\eta_n\|_{L^1}\|g\|_{L^\infty} = \|g\|_{L^\infty},
$$

故由控制收敛定理可得

$$
\int_\Omega u(x)g(x)\,\mathrm{d}x = 0. \tag{16}
$$

今在(16)中取

$$g(x)=\begin{cases}\operatorname{sgn} u(x) & x\in K,\\ 0 & x\in\mathbb{R}^n\setminus K,\end{cases}$$

其中 K 是 Ω 内的紧集. 由此我们可以得到 $u(x)=0$, a.e. $x\in K$. 注意到 K 是任意的, 故 $u=0$, a.e. □

以下利用高维形式的变分法基本引理来推导出 E-L 方程. 设 $L\in C^2, u^*\in C^2$. 对任意的 $\varphi\in C_0^1(\Omega)$, 考虑一元函数 $g(\varepsilon)=I(u^*+\varepsilon\varphi)$. 由于 u^* 是极小点, 则一阶变分 $\delta I(u^*,\varphi)=\dot g(0)=0$, 即

$$\begin{aligned}0&=\sum_{m=1}^{N}\int_\Omega\left(L_{u^m}(\tau)\varphi^m(x)+\sum_{i=1}^{n}L_{p_i^m}(\tau)\partial_i\varphi^m(x)\right)\mathrm{d}x\\&=\sum_{m=1}^{N}\int_\Omega\left(L_{u^m}(\tau)-\sum_{i=1}^{n}\partial_iL_{p_i^m}(\tau)\right)\varphi^m(x)\,\mathrm{d}x,\end{aligned}$$

其中 $\tau=(x,u^*(x),\nabla u^*(x))$. 再利用变分法基本引理, 我们便得到

$$\boxed{L_{u^m}(\tau)-\sum_{i=1}^{n}\partial_iL_{p_i^m}(\tau)=0,\qquad 1\le m\le N,}$$

即 **Euler-Lagrange 方程**. 上式还可以等价地写为

$$\boxed{L_{u^m}(\tau)-\operatorname{div} L_{p^m}(\tau)=0,\qquad 1\le m\le N.}$$

或直接简记为 $L_u(\tau)-\operatorname{div} L_p(\tau)=0$. 由此可以看出，当 $n=1$ 时, 上述 E-L 方程与我们第一节所导出的 E-L 方程是相同的. 若 u^* 的光滑性较差, 上述 E-L 方程应在广义导数的意义下理解.

类似地, 称算子 $E_L\colon u\mapsto v=(v_1,\cdots,v_n)$, 其中

$$\boxed{v_m=L_{u^m}(\tau)-\sum_{i=1}^{n}\partial_iL_{p_i^m}(\tau),\qquad m=1,\cdots,N}$$

为关于 L 的 **Euler-Lagrange 算子**.

例 1.50. *设 $N=1, L(p)=|p|^2/2=(p_1^2+\cdots+p_n^2)/2$. 对于泛函*

$$I(u)=\int_\Omega L(p)\,\mathrm{d}x=\frac{1}{2}\int_\Omega|\nabla u|^2\,\mathrm{d}x,$$

其对应的 E-L 方程为 $\nabla\cdot\nabla u=0$, 即

$$\Delta u(x)=0,\qquad \forall x\in\Omega.$$

这便是 Laplace 方程.

例 1.51. 用 $\mathbb{R}\times\mathbb{R}^3$ 表示时空连续统, 时空中任意一点的坐标是 $(t,x)=(t,x_1,x_2,x_3)$, 其中 t 表示时间, $x=(x_1,x_2,x_3)$ 表示空间位置. 如果我们用 $u=u(t,x)$ 表示在时空区域 $\Omega\subseteq\mathbb{R}\times\mathbb{R}^3$ 内弹性波的位移, 那么弹性波的动能是

$$T(u)=\frac{1}{2}\int_\Omega|\partial_t u(t,x)|^2\,\mathrm{d}t\,\mathrm{d}x,$$

势能

$$U(u)=\frac{1}{2}\int_\Omega|\nabla_x u(t,x)|^2\,\mathrm{d}t\,\mathrm{d}x,$$

由此对应的 *Lagrange* 函数

$$I(u)=T(u)-U(u)=\frac{1}{2}\int_\Omega(|\partial_t u(t,x)|^2-|\nabla_x u(t,x)|^2)\,\mathrm{d}t\,\mathrm{d}x.$$

通过直接计算可知, I 对应的 *E-L* 方程为

$$\Box u:=\partial_t^2 u-\Delta u=0.$$

这便是经典的波动方程. 类似地, 如果还有内力或外力存在, 那么在势能中可以再添加一些其他项. 例如:

$$U(u)=\frac{1}{2}\int_\Omega(|\nabla_x u(t,x)|^2+m^2|u(t,x)|^2)\,\mathrm{d}t\,\mathrm{d}x,$$

其中 $m>0$ 是一个常数. 此时对应的 *E-L* 方程为

$$\Box u-m^2u=0.$$

这是 *Klein-Gordon* 方程. 又如

$$U(u)=\int_\Omega\left(\frac{1}{2}|\nabla_x u(t,x)|^2+\frac{1}{4}|u(t,x)|^2\right)\mathrm{d}t\,\mathrm{d}x.$$

此时对应的 *E-L* 方程为

$$\Box u+u^3=0.$$

这是一个非线性波动方程.

例 1.52 (极小曲面)**.** 设 $\Omega\subseteq\mathbb{R}^n$. 给定函数 $u\in C^1(\overline{\Omega})$, 其对应的超曲面 $\{(x,u(x))\colon x\in\overline{\Omega}\}$ 的面积是

$$A(u)=\int_\Omega\sqrt{1+|\nabla u(x)|^2}\,\mathrm{d}x.$$

在给定边值 $u|_{\partial\Omega}=\Phi$ 的条件下, 我们要寻求 u 使得面积 $A=A(u)$ 达到极小. 将 A 看作是关于 U 的泛函, A 对应的 *E-L* 方程为

$$\operatorname{div}\frac{\nabla u(x)}{\sqrt{1+|\nabla u(x)|^2}}=0,\qquad\forall x\in\Omega.$$

这便是极小曲面所满足的方程. 注意到平均曲率有表达式

$$H=\frac{1}{n}\operatorname{div}\frac{\nabla u(x)}{\sqrt{1+|\nabla u(x)|^2}},$$

由此可知, 平均曲率等于零的方程就是极小曲面的方程. 我们有时候也将此作为极小曲面的定义, 即平均曲率为零的曲面.

1.5.2 必要条件: Legendre-Hadamard 条件

二阶变分 $\delta^2 I(u^*,\varphi)=\ddot{g}(0)$. 直接计算得

$$\delta^2 I(u^*,\varphi)=\sum_{m,n=1}^{N}\int_\Omega F(x)\,\mathrm{d}x,$$

其中

$$F(x)=L_{u^m u^\ell}(\tau)\varphi^m(x)\varphi^\ell(x)+2\sum_{i=1}^{n}L_{u^m p_i^\ell}(\tau)\varphi^m(x)\partial_i\varphi^\ell(x)+\sum_{i,j=1}^{n}L_{p_i^m p_j^\ell}(\tau)\partial_i\varphi^m(x)\partial_j\varphi^\ell(x),$$

而 $\tau=(x,u^*(x),\nabla u^*(x))$. 为了书写的简便, 引入记号

$$\begin{aligned}A_{u^*}&=(a_{ij}^{m\ell})=(L_{p_i^m p_j^\ell}(\tau)),\\ B_{u^*}&=(b_j^{m\ell})=(L_{u^m p_j^\ell}(\tau)),\\ C_{u^*}&=(c^{m\ell})=(L_{u^m u^\ell}(\tau)),\end{aligned}$$

并记

$$Q_{u^*}(\varphi)=\delta^2 I(u^*,\varphi)=\int_\Omega(A_{u^*}(\nabla\varphi,\nabla\varphi)+2B_{u^*}(\nabla\varphi,\varphi)+C_{u^*}(\varphi,\varphi))\,\mathrm{d}x.$$

显然, 若 u^* 是一个 (弱) 极小点, 则对任意的 $\varphi\in C_0^1(\Omega)$ 有 $Q_{u^*}(\varphi)\geq 0$.

命题 1.53. *设 $\Omega\subseteq\mathbb{R}^n$ 是一个有界区域, Lagrange 函数 $L\in C^2(\overline{\Omega}\times\mathbb{R}^N\times\mathbb{R}^{nN})$. 若 u^* 是对应的 E-L 方程的解, 则* ***Legendre-Hadamard* 条件**

$$\boxed{\sum_{m,\ell=1}^{N}\sum_{i,j=1}^{n}L_{p_i^m p_j^\ell}(\tau)\xi^m\xi^\ell\eta_i\eta_j\geq 0,\quad\forall(x,\xi,\eta)\in\Omega\times\mathbb{R}^N\times\mathbb{R}^n.}\tag{17}$$

成立.

证明. 对任意的 $x_0\in\Omega$, 取向量值函数 $v\in C_c^\infty(B_1(0))$. 当 $\mu>0$ 充分小时, 令

$$\varphi(x)=\mu v\left(\frac{x-x_0}{\mu}\right).$$

将其代入至二阶变分的具体表达式中, 即得

$$\begin{aligned}0\leq\mu^n\int_{B_1(0)}&(A_{u^*}(x_0+\mu y)(\nabla v(y),\nabla v(y))\\&+2\mu B_{u^*}(x_0+\mu y)(\nabla v(y),v(y))+\mu^2 C_{u^*}(x_0+\mu y)(v(y),v(y)))\,\mathrm{d}y.\end{aligned}$$

令 $\mu\to 0$, 即得

$$\sum_{m,\ell=1}^{N}\sum_{i,j=1}^{n}A_{u^*}(x_0)\int_{B_1(0)}\partial_i v^m(y)\partial_j v^n(y)\,\mathrm{d}y\geq 0.\tag{18}$$

现取函数 $\rho \in C_c^\infty(B_1(0))$ 满足 $\|\rho\|_{L^2(B_1(0))} = 1$. 对任意的 $t > 0, \xi \in \mathbb{R}^N, \eta \in \mathbb{R}^n$, 分别将

$$v_1(y) = \xi \cos(t\eta \cdot y)\rho(y),$$

和

$$v_2(y) = \xi \sin(t\eta \cdot y)\rho(y)$$

代入至(18)中再相加, 即得

$$\sum_{m,\ell=1}^{N} \sum_{i,j=1}^{n} A_{u^*}(x_0)\xi^m \xi^\ell \eta_i \eta_j + O(t^{-1}) \geq 0 \qquad (t \to +\infty),$$

从而有

$$\sum_{m,\ell=1}^{N} \sum_{i,j=1}^{n} A_{u^*}(x_0)\xi^m \xi^\ell \eta_i \eta_j \geq 0.$$

这便是 Legendre-Hadamard 条件(17). □

注 1.54. *若使用秩 1 矩阵的符号*

$$\pi = (\pi_i^m) = (\xi^m \eta_i),$$

那么(17)可以等价写为

$$\sum_{m,\ell=1}^{N} \sum_{i,j=1}^{n} L_{p_i^m p_j^\ell}(\tau)\pi_i^m \pi_j^\ell \geq 0, \qquad \forall \pi, \mathrm{rank}(\pi) = 1.$$

类似地, 若存在 $\lambda > 0$ 使得

$$\boxed{\sum_{m,\ell=1}^{N} \sum_{i,j=1}^{n} L_{p_i^m p_j^\ell}(\tau)\xi^m \xi^\ell \eta_i \eta_j \geq \lambda |\xi|^2 |\eta|^2, \quad \forall (x,\xi,\eta) \in \Omega \times \mathbb{R}^N \times \mathbb{R}^n.}$$

那么我们称其为**严格 Legendre-Hadamard 条件**. 利用秩 1 矩阵的符号, 上式可以等价写为

$$\sum_{m,\ell=1}^{N} \sum_{i,j=1}^{n} L_{p_i^m p_j^\ell}(\tau)\pi_i^m \pi_j^\ell \geq \lambda \|\pi\|^2, \qquad \forall \pi, \mathrm{rank}(\pi) = 1.$$

这里 $\|\pi\|$ 代表 π 的 Frobenius 范数:

$$\|\pi\| = \left(\sum_{m,\ell=1}^{N} \sum_{i,j=1}^{n} (\pi_i^m)^2 \right)^{1/2}.$$

1.5.3 充分条件

先将 Jacobi 场的概念推广到高维情形: 设 $L \in C^3$, u^* 是一个极小点. 将 $Q_{u^*}(\varphi)$ 看作是关于 φ 的变分积分, 写出其对应的 E-L 方程:

$$\boxed{J_{u^*}(\varphi) = \sum_{\ell=1}^{N}\left(\sum_{i=1}^{n}\partial_i\left(\sum_{j=1}^{n}a_{ij}^{m\ell}\partial_j\varphi^\ell + b_i^{m\ell}\varphi^\ell\right) - \left(\sum_{j=1}^{n}b_j^{m\ell}\partial_j\varphi^m + c^{m\ell}\varphi^\ell\right)\right) = 0,}$$

其中 $j = 1, \cdots, N$. 这是一个齐次二阶偏微分方程组. 称此方程为 **Jacobi 方程**, 并称 J_{u^*} 为沿 u^* 的 **Jacobi 算子**. Jacobi 方程的任一 C^2 解为沿 u^* 的 **Jacobi 场**.

以下探究 u^* 成为强极小点的充分条件:

命题 1.55 (充分条件 1)**.** *设 $u^* \in M$ 满足 E-L 方程, 并且存在 $\lambda > 0$ 使得*

$$Q_{u^*}(\varphi) \geq \lambda \int_\Omega (|\varphi|^2 + |\nabla\varphi|^2)\,\mathrm{d}x, \quad \forall \varphi \in C_0^1(\Omega), \tag{19}$$

则 u^* 是 I 的一个严格极小点.

证明. 与 $n = 1$ 的情形一样. □

对于 1 维的情形, 引入了共轭点的概念, 利用严格 Legendre-Hadamard 条件和 Poincaré 不等式来对条件(19)进行简化. 对于高维的情形, 由于共轭点的概念无法推广到高维, 故我们需要采取其它的手段. 以下我们旨在给出命题1.14在高维情形的推广.

引理 1.56 (Gårding 不等式)**.** *设 $(a_{ij}^{m\ell}(x))$ 是 $\overline{\Omega} \subseteq \mathbb{R}^n$ 上的一致连续函数, 且存在 $\sigma > 0$ 使得*

$$\sum_{m,\ell=1}^{N}\sum_{i,j=1}^{n} a_{ij}^{m\ell}(x)\xi^m\xi^\ell\eta_i\eta_j \geq \sigma|\xi|^2|\eta|^2, \quad \forall (x,\xi,\eta) \in \Omega \times \mathbb{R}^N \times \mathbb{R}^n.$$

则存在 $\beta, C_0 > 0$ 使得

$$\int_\Omega \sum_{m,\ell=1}^{N}\sum_{i,j=1}^{n} a_{ij}^{m\ell}(x)\partial_i\varphi^m\partial_j\varphi^\ell\,\mathrm{d}x \geq \beta\int_\Omega |\nabla\varphi(x)|^2 - M\int_\Omega |\varphi(x)|^2\,\mathrm{d}x, \quad \forall \varphi \in C_0^1(\Omega).$$

证明. 我们分为三种情况进行证明.

当 $N = 1$ 时, 结论是显然的.

当 $N > 1$ 且 $(a_{ij}(x))$ 恒为常数使, 我们让 φ 在 Ω 外定义为零, 使其在全空间上由定义, 从而我们可以考虑 φ 的 Fourier 变换:

$$\widehat{\varphi}(\xi) = \int_{\mathbb{R}^n} \varphi(x)\mathrm{e}^{-2\pi\mathrm{i}x\cdot\xi}\,\mathrm{d}x.$$

注意到等式 $(\partial_i\varphi)^\wedge(\xi)=(2\pi\mathrm{i}\xi_i)\widehat{\varphi}(\xi),\forall i$, 从而由题设条件和 Plancherel 定理可得,

$$\begin{aligned}\sum_{m,\ell=1}^{N}\sum_{i,j=1}^{n}\int_\Omega a_{ij}^{m\ell}\partial_i\varphi^m\partial_j\varphi^\ell\,\mathrm{d}x &\geq \sum_{m,\ell=1}^{N}\sum_{i,j=1}^{n}\int_\Omega a_{ij}^{m\ell}\widehat{\partial_i\varphi^m}\overline{\widehat{\partial_j\varphi^\ell}}\,\mathrm{d}\xi\\ &=4\pi^2\sum_{m,\ell=1}^{N}\sum_{i,j=1}^{n}\int_\Omega a_{ij}^{m\ell}\xi_i\xi_j\partial_i\widehat{\varphi}^m\partial_j\widehat{\varphi}^\ell\,\mathrm{d}\xi\\ &\geq 4\pi^2\sigma\int_{\mathbb{R}^n}|\xi\widehat{\varphi}(\xi)|^2\,\mathrm{d}\xi\\ &=\sigma\int_{\mathbb{R}^n}|\widehat{\nabla\varphi}(\xi)|^2\,\mathrm{d}\xi=\sigma\int_{\mathbb{R}^n}|\nabla\varphi(x)|^2\,\mathrm{d}x.\end{aligned}$$

最后我们考虑变系数的情形. 由 $(a_{ij}(x))$ 的一致连续性可知, 对任意的 $\varepsilon>0$, 存在 $\delta>0$, 使得当开集 $U\subseteq\overline{\Omega}$ 的直径 $\operatorname{diam}U<\delta$ 时, 有

$$\sup_{x,y\in U}|a_{ij}^{m\ell}(x)-a_{ij}^{m\ell}(y)|<\varepsilon.$$

对任意的 $x\in\overline{\Omega}$ 取球 $B_{\delta/2}(x)$. 显然有 $\bigcup_{x\in\overline{\Omega}}B_{\delta/2}(x)\supseteq\overline{\Omega}$. 再由 $\overline{\Omega}$ 的紧性可知, 存在 $x_1,\cdots x_k$ 使得 $\bigcup_{\alpha=1}^{k}B_{\delta/2}(x_\alpha)\supseteq\overline{\Omega}$. 令 $B_\alpha=B_{\delta/2}(x_\alpha)\cap\Omega$, 从而有 $\Omega=\bigcup_{\alpha=1}^{k}B_\alpha$, 其中 $\operatorname{diam}B_\alpha<\delta,\forall\alpha$. 我们现在取 $\{B_\alpha\}_{\alpha=1}^{k}$ 对应的一个单位分解, 即函数族 $\{w_\alpha\}_{\alpha=1}^{k}$ 满足如下条件:

- $w_\alpha\in C_c^\infty(B_\alpha),0\leq w_\alpha\leq 1,\forall\alpha$;
- $\sum_{\alpha=1}^{k}w_\alpha^2=1$.

从而有

$$\begin{aligned}\sum_{m,\ell=1}^{N}\sum_{i,j=1}^{n}\int_\Omega a_{ij}^{m\ell}(x)\partial_i\varphi^m\partial_j\varphi^\ell\,\mathrm{d}x &=\sum_{m,\ell=1}^{N}\sum_{i,j=1}^{n}\sum_{\alpha=1}^{k}\int_\Omega a_{ij}^{m\ell}(x)w_\alpha^2(x)\partial_i\varphi^m\partial_j\varphi^\ell\,\mathrm{d}x\\ &=S_1+S_2,\end{aligned}$$

其中

$$\begin{aligned}S_1&=\sum_{m,\ell=1}^{N}\sum_{i,j=1}^{n}\sum_{\alpha=1}^{k}\int_\Omega a_{ij}^{m\ell}(x_\alpha)w_\alpha^2(x)\partial_i\varphi^m\partial_j\varphi^\ell\,\mathrm{d}x,\\ S_2&=\sum_{m,\ell=1}^{N}\sum_{i,j=1}^{n}\sum_{\alpha=1}^{k}\int_\Omega (a_{ij}^{m\ell}(x)-a_{ij}^{m\ell}(x_\alpha))w_\alpha^2(x)\partial_i\varphi^m\partial_j\varphi^\ell\,\mathrm{d}x.\end{aligned}$$

一方面,

$$S_2=o(1)\int_\Omega(|\varphi|^2+|\nabla\varphi|^2)\,\mathrm{d}x\quad(\varepsilon\to 0).$$

另一方面, 由前述常系数的情形的分析, 有

$$\begin{aligned}S_1 &= \sum_{m,\ell=1}^{N}\sum_{i,j=1}^{n}\sum_{\alpha=1}^{k}\int_\Omega a_{ij}^{m\ell}(x_\alpha)\partial_i(\varphi^m w_\alpha)\partial_j(\varphi^\ell w_\alpha)\,\mathrm{d}x - S_3\\ &\geq \sigma\int_\Omega\sum_{\alpha=1}^{k}|\nabla(w_\alpha\varphi)|^2\,\mathrm{d}x + O(\|\nabla\varphi\|_{L^2}\|\varphi\|_{L^2}+\|\varphi\|_{L^2}^2)\\ &\geq \sigma'\int_\Omega|\nabla\varphi|^2\,\mathrm{d}x - M\int_\Omega|\varphi|^2\,\mathrm{d}x.\end{aligned}$$

综上所述, 我们有

$$\sum_{m,\ell=1}^{N}\sum_{i,j=1}^{n}\int_\Omega a_{ij}^{m\ell}(x)\partial_i\varphi^m\partial_j\varphi^\ell\,\mathrm{d}x \geq (\sigma'+o(1))\int_\Omega|\nabla\varphi(x)|^2 - M\int_\Omega|\varphi(x)|^2 \qquad (\varepsilon\to 0).$$

这便完成了引理的证明. □

命题 1.57 (充分条件 2). *设 $L\in C^2$ 满足严格 Legendre-Hadamard 条件. 若 u^* 是 E-L 方程的一个解, 且存在 $\mu>0$ 使得*

$$Q_{u^*}(\varphi)\geq\mu\int_\Omega|\varphi|^2\,\mathrm{d}x, \quad \forall\varphi\in C_0^1(\Omega),$$

则存在 $\lambda>0$ 使得

$$Q_{u^*}(\varphi)\geq\lambda\int_\Omega(|\varphi|^2+|\nabla\varphi|^2)\,\mathrm{d}x, \quad \forall\varphi\in C_0^1(\Omega).$$

从而 u^* 是 I 的一个极小点.

证明. 利用 Gårding 不等式, 参照命题1.14的证明过程即可. □

最后我们再从特征值的角度给出极小值点的刻画. 设 $u^*\in C^1(\overline{\Omega})$ 是 E-L 方程的解. 称

$$\boxed{\lambda_1 = \inf\left\{Q_{u^*}(\varphi)\colon\ \varphi\in C_0^1(\Omega), \int_\Omega|\varphi|^2\,\mathrm{d}x = 1\right\}}$$

为 Jacobi 算子的**第一特征值**.

命题 1.58. *设 $L\in C^2$ 满足严格 Legendre-Hadamard 条件, 又设 $u^*\in M$ 是 I 的一个弱极小点, 则 $\lambda_1\geq 0$; 反之, 若 $\lambda_1>0$, 则 u^* 是 I 的一个严格弱极小点.*

1.6 约束极值问题

1.6.1 等式约束

例 1.59 (等周问题). *在平面上给定封闭曲线的弧长, 问什么样的曲线围成的面积最大? 为了问题的简单, 我们假设这条封闭曲线有参数表示*

$$r = r(t) = (x(t), y(t)),$$

其中 $t\in[0,2\pi]$. 此曲线围成的面积

$$S=S(r)=\frac{1}{2}\int x\,\mathrm{d}y-y\,\mathrm{d}x=\frac{1}{2}\int_0^{2\pi}(x\dot{y}-y\dot{x})\,\mathrm{d}t,$$

其长度为

$$L=\int_0^{2\pi}\sqrt{\dot{x}^2+\dot{y}^2}\,\mathrm{d}t.$$

如果给定长度为 l, 那么我们的问题就是在约束 $L=l$ 之下, 求曲线 $(x(t),y(t))$ 使得面积 S 达到极大值.

函数的条件极值问题: **Lagrange 乘子法**. 设 $\Omega\subseteq\mathbb{R}^n$ 是一个开集, 函数 $f,g\in C^1(\overline{\Omega})$. 又设 $g^{-1}(0)\neq\varnothing$ 且 $\dot{g}(x)\neq 0$. 如果存在 $x_0\in\Omega$ 使得 f 在约束条件 $g(x)=0$ 下达到极小值, 那么便存在一个 Lagrange 乘子 $\lambda\in\mathbb{R}$, 使得

$$\nabla f(x_0)+\lambda\nabla g(x_0)=0.$$

对于泛函的情形, 我们也有类似的 Lagrange 乘子法可用.

命题 1.60. *给定 $L,G\in C^2(\overline{\Omega}\times\mathbb{R}^N\times\mathbb{R}^{nN})$ 和 $\Phi\in C^1(\partial\Omega)$, 定义 M 上的泛函*

$$I(u)=\int_\Omega L(x,u(x),\nabla u(x))\,\mathrm{d}x$$

和

$$N(u)=\int_\Omega G(x,u(x),\nabla u(x))\,\mathrm{d}x.$$

*设 $N^{-1}(0)\cap M\neq\varnothing$. 若 $u^*\in M$ 是 I 在约束 $N(u)=0$ 下的极小点, 且存在 $\varphi^*\in C_0^1(\Omega)$ 使得 $\delta N(u^*,\varphi^*)\neq 0$, 则存在 $\lambda\in\mathbb{R}$(此时 λ 也被称作 **Lagrange 乘子**) 使得*

$$\delta I(u^*,\varphi)+\lambda\delta(u^*,\varphi)=0,\quad\forall\varphi\in C_0^1(\Omega).$$

等价地, 若记 $Q=L+\lambda G$ 为调整后的 Lagrange 函数, 则 u^* 满足 Q 对应的 E-L 方程:

$$\boxed{\operatorname{div}Q_p(x,u^*(x),\nabla u^*(x))=Q_u(x,u^*(x),\nabla u^*(x)).}$$

证明. 对任意的 $\varphi\in C_0^1(\Omega)$, 考虑通过 u^*, 并由 φ 和 φ^* 张成的平面:

$$\pi=\{u^*+\varepsilon\varphi+\tau\varphi^*:\,(\varepsilon,\tau)\in\mathbb{R}^2\}$$

以及函数

$$\Phi(\varepsilon,\tau)=I(u^*+\varepsilon\varphi+\tau\varphi^*),$$
$$\Psi(\varepsilon,\tau)=N(u^*+\varepsilon\varphi+\tau\varphi^*).$$

注意到 $\Psi(0,0)=N(u^*)=0, \partial_\tau\Psi(0,0)=\delta N(u^*,\varphi^*)\neq 0$, 故由隐函数定理可知, 当 $r>0$ 充分小时, 方程 $\Psi(\varepsilon,\tau)=0$ 在 $B_r(0,0)$ 上有唯一 C^1 解: $\tau=\tau(\varepsilon)$. 由此表明 $N^{-1}(0)\cap\pi\neq\varnothing$. 现令

$$g(\varepsilon)=\Phi(\varepsilon,\tau(\varepsilon))=I(u^*+\varepsilon\varphi+\tau(\varepsilon)\varphi^*).$$

由题设条件可知, u^* 是 I 在 $N^{-1}(0)\cap M$ 上的极小点. 对应地, 0 是 g 的极小点, 从而有

$$\begin{aligned}
0=\dot{g}(0)&=\partial_\varepsilon\Phi(0,0)+\partial_\tau\Phi(0,0)\dot{\tau}(0)\\
&=\delta I(u^*,\varphi)+\delta I(u^*,\varphi^*)\left(-\frac{\delta N(u^*,\varphi)}{\delta N(u^*,\varphi^*)}\right)\\
&=\delta I(u^*,\varphi)+\lambda\delta N(u^*,\varphi),
\end{aligned}$$

其中

$$\lambda=-\frac{\delta I(u^*,\varphi_0)}{\delta N(u^*,\varphi_0)}$$

是一个常数. 这便得到了所需结论. □

同理, 我们也可以考虑多个约束的泛函极值问题.

例 1.61 (等周问题-续)**.** *此时调整后的 Lagrange 函数为*

$$Q=\frac{1}{2}(x\dot{y}-y\dot{x})+\lambda\sqrt{\dot{x}^2+\dot{y}^2},$$

其中 $\lambda\in\mathbb{R}$. 对应的 E-L 方程为

$$\begin{cases}\dot{x}=\lambda\dfrac{\mathrm{d}}{\mathrm{d}t}\dfrac{\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}},\\ \dot{y}=\lambda\dfrac{\mathrm{d}}{\mathrm{d}t}\dfrac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}}.\end{cases}$$

由此解出

$$\begin{cases}x-c_1=-\lambda\dfrac{\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}},\\ y-c_2=\lambda\dfrac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}},\end{cases}$$

其中 c_1,c_2 是常数. 显然, 这是圆的方程:

$$(x-c_1)^2+(y-c_2)^2=\lambda^2,$$

其半径 $r=\lambda=l/2\pi$, 圆心为 (c_1,c_2).

上述我们考虑的是**积分形式的约束**, 以下我们考虑**等式约束**. 具体地, 给定函数 $F\in C^1(\overline{\Omega}\times\mathbb{R}^N\times\mathbb{R}^{nN})$, 我们要在约束

$$F(x,u(x),\nabla u(x))=0,\quad \forall x\in\Omega$$

下, 求泛函

$$I(u)=\int_\Omega L(x,u(x),\nabla u(x))\,\mathrm{d}x,\quad x\in M$$

的极值. 事实上, 对于**完整 (holonomic) 约束**, 即 F 只依赖于 u 的情形, 我们也有类似的 Lagrange 乘子法可用.

命题 1.62. *设 Ω 是 $\mathbb{R}^n$ 中的有界区域. 设 $L\in C^2(\overline{\Omega}\times\mathbb{R}^N\times\mathbb{R}^{nN}),F\in C^2(\mathbb{R}^N)$. 又设 $u^*\in M$ 是在约束 $F(u(x))=0$ 下的极小点, 并且 u^* 在有限个逐片 C^1 的 $(n-1)$ 维超曲面之外是 C^2 的. 若对于任意的 $x\in\overline{\Omega},\nabla F(u^*(x))\neq 0$, 那么存在 $\lambda\in C(\overline{\Omega})$, 使得 u^* 满足对应于调整后的 Lagrange 函数 $Q=L+\lambda F$ 的 E-L 方程:*

$$\boxed{L_{u_m}+\lambda M_{u_m}=\sum_{i=1}^n\partial_{x_i}L_{p_i^m},\quad 1\le m\le N.}\tag{20}$$

证明. 同命题1.60的证明思路类似, 我们先利用隐函数定理, 将有约束问题转化为无约束问题, 从而构造出局部定义的连续函数 λ, 最后再将它们粘连起来, 成为一个整体定义的连续函数.

任意固定 $x_0\in\Omega$, 我们选取充分小的 $r>0$ 使得 $\nabla_xF(u^*(x))\neq 0,\forall x\in B_r(x_0)\subseteq\Omega$. 注意到

$$\nabla_xF(u^*(x))=\nabla_uF(u^*(x))\cdot\nabla u^*(x),$$

因此 $\nabla F(u^*)\neq 0,\forall u\in u^*(B_r(x_0))$. 不失一般性, 设 $F_{u^N}(u^*)\neq 0,\forall u\in u^*(B_r(x_0))$. 利用隐函数定理, 我们可以局部地解出 u^N: $u^N=U(\widetilde{u})$, 其中 $\widetilde{u}=(u^1,\cdots,u^{N-1}),U\in C^2$. 现令

$$\Lambda(x,\widetilde{u},\widetilde{p})=L(x,\widetilde{u},U(\widetilde{u}),\widetilde{p},p^N),$$

其中

$$\widetilde{p}=(p_i^m)_{1\le i\le n,1\le m\le N-1},$$
$$p^N=(p_i^N)_{1\le i\le n}=\left(\sum_{m=1}^{N-1}U_{u^m}p_i^m\right)_{1\le i\le n}.$$

由题设条件可知, 若 u^* 是在约束 $F(u^*)=0$ 下 I 的极小点, 则 $\widetilde{u^*}$ 一定是

$$J(\widetilde{u})=\int_{B_r(x_0)}\Lambda(x,\widetilde{u}(x),\nabla\widetilde{u}(x))\,\mathrm{d}x$$

的极小点. 后者对应的 E-L 方程为

$$L_{u^m}+L_{u^N}U_{u^m}+\sum_{i=1}^nL_{p_i^N}\partial_{u^m}p_i^N=\sum_{i=1}^n\partial_{x_i}(L_{p_i^m}+L_{p_i^N}U_{u^m}),\quad\forall m=1,\cdots,N-1.\tag{21}$$

直接计算可得

$$\partial_{u^m}p_i^N=\sum_{\ell=1}^{N-1}\partial_{u^m}U_{u^\ell}p_i^\ell=\partial_{x_i}U_{u^m},$$

从而有
$$\sum_{i=1}^{n}\partial_{x_i}(L_{p_i^m}+L_{p_i^N}U_{u^m})=\sum_{i=1}^{n}(\partial_{x_i}L_{p_i^m}+U_{u^m}\partial_{x_i}L_{p_i^N}+L_{p_i^N}\partial_{u^m}p_i^N).$$
因此(21)式化为
$$L_{u^m}+U_{u^m}\left(L_{u^N}-\sum_{i=1}^{n}\partial_{x_i}L_{p_i^N}\right)=\sum_{i=1}^{n}\partial_{x_i}L_{p_i^m}. \tag{22}$$
注意到
$$U_{u^m}=-\frac{F_{u^m}}{F_{u^N}},$$
因此, 若在 $B_r(x_0)$ 上定义
$$\lambda_{B_r(x_0)}=\frac{1}{F_{u^N}}(\mathrm{div}\,L_{p^N}-L_{u^N}),$$
则(22)可写为
$$L_{u^m}+\lambda_{B_r(x_0)}F_{u^m}=\mathrm{div}\,L_{p^m},\quad m=1,\cdots,N-1.$$
这便是 (在 $B_r(x_0)$ 上) 局部的 E-L 方程.

最后, 由上述的构造可知, $\Omega=\bigcup_{x\in\Omega}B_r(x)$, 且每个球 $B_r(x)$ 对应的函数 $\lambda_{B_r(x)}$ 都是连续的. 进一步地, 当两个这样的小球, 设为 $B_{r_1}(x_1)$ 和 $B_{r_2}(x_2)$, 相交非空时, 显然有
$$\lambda_{B_{r_1}(x_1)}=\lambda_{B_{r_2}(x_2)}.$$
因此, 由粘接引理可知, 存在 $\lambda\in C(\Omega)$, 使得 $\lambda|_{B_r(x)}=\lambda_{B_r(x)},\forall x\in\Omega$, 且满足(20)式. 注意到 u^* 在边界处的值是已知的, 故由(20)的具体表达式可知, 我们可以将 λ 延拓到 $\overline{\Omega}$ 上, 使之成为 $\overline{\Omega}$ 上的连续函数. 显然延拓后的 λ 即为所求。 □

与积分形式的约束类似, 我们也可以考虑不止一个约束函数的情形. 与积分约束的情形相比, 这里的 Lagrange 乘子 λ 不是常数, 而是定义在 $\overline{\Omega}$ 上的连续函数.

例 1.63 (球面上的测地线)**.** *我们以条件约束的观点来讨论球面上的测地线问题. 具体地, 设曲线有参数表示 $u=u(t)=(x(t),y(t),z(t)),a\le t\le b$, 约束为*
$$f(x,y,z)=x^2+y^2+z^2=1,$$
我们要在此约束下求解泛函
$$I(u)=\int_a^b\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}\,\mathrm{d}t$$
的极值. 令 $u=(x,y,z),p=(\xi,\eta,\zeta)$. 引入 Lagrange 乘子 $\lambda=\lambda(t)$, 从而得到调整后的 Lagrange 函数
$$Q=\sqrt{\xi^2+\eta^2+\zeta^2}+\lambda(x^2+y^2+z^2-1)=|p|+\lambda(|u|^2-1).$$

注意到约束 f 的表达式只依赖于 u, 故由命题(1.62)可知, λ 满足 Q 所对应的 *E-L* 方程:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\dot{u}}{|\dot{u}|}=2\lambda u,$$

其中 $|u|=1$. 进一步地, 我们有

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\dot{u}}{|\dot{u}|}\times u\right)=\left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\dot{u}}{|\dot{u}|}\right)\times u+\frac{\dot{u}}{|\dot{u}|}\times\dot{u}=0,$$

即 $v=\frac{\dot{u}}{|\dot{u}|}\times u$ 是常向量. 由此可知, u 必须位于垂直于常向量 v 且过原点的平面上. 因此 u 必是大圆的一部分.

例 1.64 (到球面的调和映射)**.** 设 $\mathbb{B}$ 是 $\mathbb{R}^3$ 中的单位球, $\mathbb{S}=\partial\mathbb{B}$ 是单位球面, $u=(u_1,u_2,u_3)\colon\mathbb{B}\to\mathbb{S}$. 若 u^* 是下列约束问题

$$\min\{I(u)\colon M(u)=0\},$$

的解, 其中

$$I(u)=\int_\Omega|\nabla u|^2\,\mathrm{d}x=\sum_{m=1}^{3}\sum_{i=1}^{3}\int_\Omega|\partial_{x_i}u_m|^2\,\mathrm{d}x,$$
$$M(u)=|u|^2-1=u_1^2+u_2^2+u_3^2-1,$$

那么我们称 u^* 为从 $\mathbb{B}$ 到 $\mathbb{S}$ 的**调和映射**. 以下我们导出 u^* 满足的方程. 首先写出调整后的 *Lagrange* 函数对应的 *E-L* 方程:

$$-\Delta u=\lambda u.\tag{23}$$

其中 $\lambda\in C(\mathbb{B}),\Delta u=(\Delta u_1,\Delta u_2,\Delta u_3)$. 在等式 $u\cdot u=1$ 两边求导, 我们有

$$u\cdot\partial_m u=0,$$

其中 $\partial_m u=(\partial_{x_m}u_1,\partial_{x_m}u_2,\partial_{x_m}u_3)$. 再次求导, 并将得到的等式相加, 即得

$$u\cdot\Delta u+|\nabla u|^2=0.\tag{24}$$

联立(23)和(24), 得

$$\lambda=-u\cdot\Delta u=|\nabla u|^2.$$

由此我们导出了调和映射所满足的方程:

$$-\Delta u=u|\nabla u|^2.$$

1.6.2 不等式约束

设 Ω 是 $\mathbb{R}^n$ 上的有界区域, M 为定义在 $\overline{\Omega}$ 上的 C^1 函数的集合. 给定 M 的一个凸子集 C 和 Lagrange 函数 L, 我们考虑如下约束问题:

$$\min\left\{I(u)=\int_\Omega L(x,u(x),\nabla(u(x)))\,\mathrm{d}x\colon u\in C\right\}.$$

若 $u^* \in M$ 是一个极小点, 那么对于任意的 $v \in C$, 由于 C 是凸集, 则 $tv+(1-t)u \in C, \forall t \in [0,1]$, 从而有

$$I(tv+(1-t)u) \geq I(u), \quad \forall t \in [0,1].$$

由此可以推出

$$\delta I(u, v-u) = \lim_{t \to 0^+} \frac{I(u+t(v-u)) - I(u)}{t} \geq 0,$$

即对任意的 $v \in C$, 我们有

$$\boxed{\int_\Omega (L_u(x,u(x),\nabla u(x))(v(x)-u(x)) + L_p(x,u(x),\nabla u(x)) \cdot (\nabla v(x) - \nabla u(x)))\,\mathrm{d}x \geq 0.}$$

称上述不等式为**变分不等式**.

例 1.65 (障碍问题)**.** 设 $\Omega \subseteq \mathbb{R}^2$ 是一个有界区域. 给定函数 $\varphi \in C^1(\partial\Omega), \psi \in C^1(\overline{\Omega}), f \in C(\overline{\Omega})$. 在 Ω 上我们考虑一张薄膜 u, 它的边界固定: $u|_{\partial\Omega} = \varphi$, 并受外力 f 的作用, 但不能越过 "障碍", 即 $u(x) \leq \psi(x), \forall x \in \overline{\Omega}$. 具体地, 我们要寻求薄膜的平衡位置

$$u \in M = \{u \in PWC^1(\overline{\Omega})\colon u|_{\partial\Omega} = \varphi\}$$

在不等式约束

$$u(x) \leq \psi(x), \quad \forall x \in \overline{\Omega}, u \in M$$

下, 使薄膜的能量

$$I(u) = \int_\Omega \left(\frac{1}{2}|\nabla u(x)|^2 - f(x)u(x)\right)\mathrm{d}x$$

达到极小值. 注意到 $C = \{u \in PWC^1(\overline{\Omega})\colon u|_{\partial\Omega} = \varphi, u(x) \leq \psi(x), \forall x \in \overline{\Omega}\}$ 是一个凸集, 故该变分问题对应的变分不等式为

$$\int_\Omega (\nabla u \nabla(v-u) - f(v-u))\,\mathrm{d}x \geq 0, \quad \forall v \in C.$$

1.7 应用: Noether 定理

为了表达式的简洁, 本节我们在公式推导有时会使用 Einstein 求和约定.

1.7.1 一阶变分的推广

我们先对经典的一阶变分 $\delta I(u^*, \varphi)$ 进行推广. 在前几节我们讨论的变分问题中, M 中的函数均具有相同的定义域. 事实上, 这种限制不是必需的.

设 Ω 为 $\mathbb{R}^n$ 中的有界区域, 并给定 $u \in C^1(\Omega)$ 和 Lagrange 函数 $L \in C^2(\overline{\Omega} \times \mathbb{R}^N \times \mathbb{R}^{nN})$. 现引入一族单参数微分同胚 $\eta_\varepsilon = \eta(\cdot, \varepsilon)\colon \Omega \to \Omega_\varepsilon$, 其中 $\Omega_\varepsilon = \eta_\varepsilon(\Omega) \subseteq \mathbb{R}^n, \eta_0 = \eta(\cdot, 0) = \mathrm{id}$. 若设 $\partial_\varepsilon \eta_\varepsilon(x)|_{\varepsilon=0} = \bar{X}(x)$, 那么

$$\eta_\varepsilon(x) = x + \varepsilon \bar{X}(x) + o(\varepsilon) \qquad (\varepsilon \to 0). \tag{25}$$

我们再考虑一族从 Ω_ε 映到 $\mathbb{R}^N$ 的函数 $v_\varepsilon = v(\cdot,\varepsilon), |\varepsilon| < \varepsilon_0$, 其中 $v_0 = u$. 设 $\partial_\varepsilon \eta_\varepsilon(x)|_{\varepsilon=0} = \bar{X}(x)$, 则有

$$v_\varepsilon(x) = u(x) + \varepsilon\varphi(x) + o(\varepsilon) \qquad (\varepsilon \to 0).$$

现记

$$I = I(u,\Omega) = \int_\Omega L(x,u(x),\nabla u(x))\,\mathrm{d}x,$$

我们考虑 $(u_\varepsilon,\Omega_\varepsilon)$ 在 I 上的取值. 即

$$\begin{aligned}\Phi(\varepsilon) = I(v_\varepsilon,\Omega_\varepsilon) &= \int_{\Omega_\varepsilon} L(y,v_\varepsilon(y),\nabla v_\varepsilon(y))\,\mathrm{d}y \\ &= \int_\Omega L(\eta_\varepsilon(x),v_\varepsilon(\eta_\varepsilon(x)),\nabla_y v_\varepsilon(\eta_\varepsilon(x)))\det(\partial_{x_i}\eta_\varepsilon^j)\,\mathrm{d}x.\end{aligned}$$

进一步地, 由表达式(25)和行列式求导法则可知,

$$\det(\partial_i\eta_\varepsilon^j)|_{\varepsilon=0} = 1, \qquad \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\det(\partial_{x_i}\eta_\varepsilon^j)\bigg|_{\varepsilon=0} = \partial_i\bar{X}^i = \operatorname{div}\bar{X},$$

从而当 $u \in C^2$ 时, 我们有

$$\begin{aligned}\dot{\Phi}(0) &= \int_\Omega (\partial_i L(\tau)\bar{X}^i(x) + L_{u^m}(\tau)\varphi^m(x) + L_{p_i^m}(\tau)\partial_i\varphi^m(x) + L(\tau)\operatorname{div}\bar{X}(x))\,\mathrm{d}x \\ &= \int_\Omega (\operatorname{div}(L\bar{X})(\tau) + L_{u^m}(\tau)\varphi^m(x) - \partial_i L_{p_i^m}(\tau)\varphi^m(x) + \operatorname{div}(L_{p^m}\varphi^m)(\tau))\,\mathrm{d}x,\end{aligned}$$

其中 $\tau = (x,u(x),\nabla u(x))$. 化简得

$$\boxed{\dot{\Phi}(0) = \int_\Omega (E_L(u)^m\varphi^m + \operatorname{div}(L\bar{X} + L_{p^m}\varphi^m))\,\mathrm{d}x.} \tag{26}$$

(26)式也可以看作是一种推广形式的一阶变分. 我们将其记为 $\delta^* I(u;\varphi,\bar{X})$.

例 1.66 (Pohozaev 恒等式)**.** *给定有光滑边界的有界区域 $\Omega \subseteq \mathbb{R}^n$ 和函数 $g \in C^1(\mathbb{R})$, 考虑下列非线性椭圆方程:*

$$\begin{cases} -\Delta u = g(u) & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{27}$$

则当 $n \geq 3$ 时, 它的解满足下列 ***Pohozaev 恒等式****:*

$$\frac{n-2}{2}\int_\Omega |\nabla u|^2\,\mathrm{d}x - n\int_\Omega G(u)\,\mathrm{d}x + \frac{1}{2}\int_{\partial\Omega}\left|\frac{\partial u}{\partial\nu}\right|^2 (x\cdot\nu)\,\mathrm{d}S = 0,$$

其中 G 是 g 的原函数, $G(0) = 0$; ν 是 $\partial\Omega$ 的单位外法向量. 事实上, 取 Lagrange 函数

$$L = \frac{1}{2}p^2 - G(u)$$

以及 $M = C_0^1(\Omega)$. 容易验证, L 对应的 E-L 方程正是(27). 不妨假设 $0 \in \Omega$. 考虑一族单参数微分同胚 $\eta_\varepsilon\colon \Omega \to \Omega_\varepsilon, x \mapsto (1+\varepsilon)x$, 其中 $\Omega_\varepsilon = \eta_\varepsilon(\Omega)$. 显然有 $\eta_0 = \mathrm{id}$. 现设 $u \in M$ 是(27)的

解. 令 $v_\varepsilon = u\circ\eta_\varepsilon$. 显然有 $v_0 = u$. 通过直接计算可知, $\partial_\varepsilon\eta_\varepsilon|_{\varepsilon=0} = 1, \partial_\varepsilon v_\varepsilon|_{\varepsilon=0} = -x\cdot\nabla u$. 利用 $\delta^* I(u;\varphi,\bar{X})$ 的具体表达式, 我们有

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}I(v_\varepsilon,\Omega_\varepsilon)\bigg|_{\varepsilon=0} = \int_\Omega \operatorname{div}\left(\left(\frac{1}{2}|\nabla u|^2 - G(u)\right)x - \nabla u(x\cdot\nabla u)\right)\mathrm{d}x.$$

一方面, 注意到 u 满足方程(27), 故等式右端可以化为

$$\begin{aligned}
&\int_\Omega \left(\frac{n}{2}|\nabla u|^2 - nG(u) + x\cdot\nabla\left(\frac{1}{2}|\nabla u|^2 - G(u)\right) - \Delta u(x\cdot\nabla u) - \nabla u\cdot\nabla(x\cdot\nabla u)\right)\mathrm{d}x\\
&\quad= \int_\Omega \left(\frac{n}{2}|\nabla u|^2 - nG(u) + \frac{x}{2}\cdot\nabla(|\nabla u|^2) - \frac{x}{2}\cdot\nabla(|\nabla u|^2) - |\nabla u|^2\right)\mathrm{d}x\\
&\quad= \int_\Omega\left(\frac{n-2}{2}|\nabla u|^2 - nG(u)\right)\mathrm{d}x.
\end{aligned}\tag{28}$$

这里我们用到了公式

$$\nabla(f\cdot g) = f\times(\nabla\times g) + g\times(\nabla\times f) + (f\cdot\nabla)g + (g\cdot\nabla)f,$$

其中 f,g 均为向量场. 另一方面, 注意到条件 $G(0)=0$ 且 $u|_{\partial\Omega}=0$, 故由散度定理可得

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\varepsilon}I(v_\varepsilon,\Omega_\varepsilon)\bigg|_{\varepsilon=0} &= \int_{\partial\Omega}\left(\frac{1}{2}|\nabla u|^2(x\cdot\nu) - (\nabla u\cdot\nu)(x\cdot\nabla u)\right)\mathrm{d}S\\
&= \int_{\partial\Omega}\left(\frac{1}{2}\left|\frac{\partial u}{\partial\nu}\right|^2(x\cdot\nu) - (\nabla u\cdot\nabla u)(x\cdot\nu)\right)\mathrm{d}S\\
&= -\frac{1}{2}\int_{\partial\Omega}\left|\frac{\partial u}{\partial\nu}\right|^2(x\cdot\nu)\,\mathrm{d}x,
\end{aligned}\tag{29}$$

其中 ν 是 $\partial\Omega$ 的单位外法向量. 联立(28)和(29), 即得欲证恒等式.

1.7.2 Noether 定理

粗略地说, Noether 定理表明: 若变分积分 I 在某单参数变换群下保持不变, 则对于 I 的极值点 u^* 有某种守恒律成立.

在前一节的讨论中, 我们固定了 u, 让 x 在一定范围内作形变. 现在我们考虑更一般的情况, 即在相空间 (x,u) 中作形变. 取 $\mathbb{R}^n$ 中的有界区域 Ω. 给定向量场[2]

$$X = X^i(x,u)\frac{\partial}{\partial x_i} + U^m(x,u)\frac{\partial}{\partial u^m},\tag{30}$$

我们可以找到一族定义在 $\Omega\times\mathbb{R}^N$ 上的单参数变换群 $\{\phi_\varepsilon\}_{|\varepsilon|<\varepsilon_0}$, 其中 $\phi_0 = \mathrm{id}$. 进一步地, 若设 $\phi_\varepsilon(x,u) = (Y(x,u,\varepsilon), W(x,u,\varepsilon))$, 则有

$$\begin{cases}X(x,u) = \partial_\varepsilon Y(x,u,\varepsilon)|_{\varepsilon=0},\\ U(x,u) = \partial_\varepsilon W(x,u,\varepsilon)|_{\varepsilon=0}.\end{cases}$$

[2]这里特指微分流形中的概念.

现对任意的 $u\in C^1(\overline{\Omega})$, 我们令

$$\begin{cases}\eta(x,\varepsilon)=Y(x,u(x),\varepsilon),\\ \omega(x,\varepsilon)=W(x,u(x),\varepsilon).\end{cases}$$

则有 $\eta(x,0)=x,\omega(x,0)=u(x)$. 再令

$$\begin{cases}\bar{X}(x)=\partial_\varepsilon\eta(x,\varepsilon)|_{\varepsilon=0}=X(x,u(x)),\\ \bar{U}(x)=\partial_\varepsilon\omega(x,\varepsilon)|_{\varepsilon=0}=U(x,u(x)),\end{cases}\tag{31}$$

从而有

$$\eta_\varepsilon(x)=\eta(x,\varepsilon)=x+\varepsilon\bar{X}(x)+o(\varepsilon)\qquad(\varepsilon\to 0).$$

若记 $\Omega_\varepsilon=\eta_\varepsilon(\Omega)$, 则由上述表达式可知, $\eta_\varepsilon:\Omega\to\Omega_\varepsilon$ 是一族微分同胚, 其中 $|\varepsilon|<\varepsilon_0,\varepsilon_0>0$ 充分小. 此外, 注意到 η_ε 有逆映射 $\xi_\varepsilon=\eta_{-\varepsilon}$, 从而有

$$\xi_\varepsilon(x)=x-\varepsilon\bar{X}(x)+o(\varepsilon),\quad \varepsilon\to 0.$$

由此我们可以导出一族从 Ω_ε 映到 $\mathbb{R}^N$ 的映射:

$$v_\varepsilon(x)=\omega(\xi_\varepsilon(x),\varepsilon).$$

且有 $v_0(y)=\omega(\xi_0(y),0)=\omega(x,0)=u(x)$. 若设 $\partial_\varepsilon v(x,\varepsilon)|_{\varepsilon=0}=\varphi(x)$, 则有

$$\bar{U}(x)=\partial_\varepsilon v(\eta_\varepsilon(x),\varepsilon)|_{\varepsilon=0}=\varphi(x)+\sum_{i=1}^n\partial_i u(x)\bar{X}^i(x).$$

即

$$\varphi=\bar{U}-\sum_{i=1}^n\partial_i u\bar{X}^i.\tag{32}$$

这样一来, 结合(31)和(32), 并将式子中出现的 $\bar{X}$ 和 φ 代入至 $\delta^*I(u;\varphi,\bar{X})$ 的表达式中, 我们便得到了在**一般的局部单参数变换群下所满足的恒等式**. 进一步地, 若变分积分 $I(u_\varepsilon,\Omega_\varepsilon)$ 与 ε 无关 (此时我们也称 I 关于 $\{\phi_\varepsilon\}$ 是不变的), 那么 $\delta^*I(u;\varphi,\bar{X})$. 特别地, 对于任意的 $x\in\Omega$, 我们取 $\Omega=B_r(x)$, 其中 $r>0$ 充分小, 从而有

$$\frac{1}{|B_r(x)|}\int_{B_r(x)}(E_L(u)^m\varphi^m+\operatorname{div}(L\bar{X}+L_{p^m}\varphi^m))\,\mathrm{d}x=0.$$

在上述等式两边令 $r\to 0$, 并注意到 $x\in\Omega$ 的任意性, 我们便有

$$\boxed{E_L(u)^m\varphi^m+\operatorname{div}(L\bar{X}+L_{p^m}\varphi^m)=0.}$$

称上述恒等式为 **Noether 恒等式**, 它便是守恒律的体现.

定理 1.67 (Noether 定理)**.** 设局部单参数变换群 $\{\phi_\varepsilon\}$ 是由向量场(30)生成的, 又设泛函

$$I(u)=\int_\Omega L(x,u(x),\nabla u(x))\,\mathrm{d}x.$$

则对任意的 $u\in C^2(\Omega)$, 有恒等式(26)成立, 其中

$$\begin{cases}\bar{X}(x)=X(x,u(x)),\\ \bar{U}(x)=U(x,u(x)),\\ \varphi(x)=\bar{U}(x)-\displaystyle\sum_{i=1}^n \partial_i\bar{X}^i(x).\end{cases}$$

进一步地, 若 $\{\phi_\varepsilon\}$ 关于 I 是不变的, 那么有如下的 *Noether* 恒等式成立 L

$$E_L(u)^m\varphi^m+\operatorname{div}(L\bar{X}+L_{p^m}\varphi^m)=0.$$

注 1.68. 当 $u\in C^2$ 是 I 的弱极小点时, 由 *Noether* 恒等式可知, 微分形式

$$\omega=\sum_{i=1}^n\left(L\bar{X}^i+\sum_{m=1}^N L_{p_i^m}\left(\bar{U}^m-\sum_{j=1}^n\partial_j u^m\bar{X}^j\right)\right)\mathrm{d}x_1\wedge\cdots\mathrm{d}x_{i-1}\wedge\mathrm{d}x_{i+1}\wedge\cdots\mathrm{d}x_n$$

是闭的, 即 $\mathrm{d}\omega=0$. 这也是守恒律的一种体现.

例 1.69. 考虑 k 个质点所构成的系统, 其质量分别为 $m_1,\cdots,m_k$. 位置坐标 $X=(X_1,\cdots,X_k)$, 其中 $X_i=(x_i,y_i,z_i)$ 是第 i 个质点的空间坐标. 则系统的总动能为

$$T=\frac{1}{2}\sum_{i=1}^k m_i|\dot{X}_i(t)|^2=\frac{1}{2}\sum_{i=1}^k m_i(\dot{x}_i^2+\dot{y}_i^2+\dot{z}_i^2),$$

势能为

$$V=-k\sum_{i<j}\frac{m_im_j}{|X_i-X_j|}=-k\sum_{i<j}\frac{m_im_j}{((x_i-x_j)^2+(y_i-y_j)^2+(z_i-z_j)^2)^{1/2}}.$$

Lagrange 函数 $L=T-V$, 对应的变分积分为

$$I(X)=\int_{t_0}^{t_1}L(X(t),\dot{X}(t)).$$

- **空间平移群**. 设 $\{S_\varepsilon\}$ 是空间坐标依赖于参数 ε 的一族变换:

  $$\widetilde{t}=t,\quad \widetilde{x_i}=x_i+\varepsilon,\quad \widetilde{y_i}=y_i,\quad \widetilde{z_i}=z_i\qquad(1\le i\le k).$$

  注意到 L 在这组变换下有相同的表达式, 故 I 关于 $\{S_\varepsilon\}$ 是不变的. 此时对应的向量场 $X=0$, 而 $U=(e_1,\cdots,e_1)$, 其中 $e_1=(1,0,0)$. 此时由 *Noether* 恒等式可得

  $$\sum_{i=1}^k m_i\dot{x}_i=\text{const}.$$

同理, 对 y,z 方向作平移, 我们也可以得到类似的等式. 由此可得

$$\sum_{i=1}^{k} m_i \dot{X}_i(t) = \sum_{i=1}^{k} m_i(\dot{x}_i(t), \dot{y}_i(t), \dot{z}_i(t)) = \text{const.}$$

即**动量守恒**.

- **时间平移群**. 设 $\{T_\varepsilon\}$ 是空间坐标依赖于参数 ε 的一族变换:

$$\widetilde{t} = t + \varepsilon, \widetilde{X}_i = X_i \qquad (1 \le i \le k).$$

由于 I 与 t 无关, 故 I 关于 $\{T_\varepsilon\}$ 是不变的. 此时 $X = 1, U = 0$, 故由 *Noether* 恒等式得到

$$H = pL_p - L = \text{const.}$$

这是**能量守恒**.

- **单参数转动群**. 设 $\{R_\varepsilon\}$ 是时空坐标依赖于 ε 的一族变换:

$$\widetilde{t} = t, \quad \widetilde{x_i} = x_i \cos\varepsilon + y_i \sin\varepsilon, \quad \widetilde{y_i} = -x_i \sin\varepsilon + y_i \cos\varepsilon, \quad \widetilde{z_i} = z_i \ (1 \le i \le k).$$

可以验证, I 关于 $\{R_\varepsilon\}$ 是不变的. 直接计算得

$$X = 0, U = (Z_1, \cdots, Z_k),$$

其中 $Z_i = (y_i, -x_i, 0), 1 \le i \le k$. 因此, 由 *Noether* 恒等式可知

$$\sum_{i=1}^{k} m_i(y_i \dot{x}_i - x_i \dot{y}_i) = \text{const.}$$

类似地, 对于平面 yOz 和 zOx 上的转动, 也有类似的等式. 从而有

$$\sum_{i=1}^{k} m_i X_i \times \dot{X}_i = \text{const.}$$

上述等式即代表**角动量守恒**.

1.7.3 内极小

本节我们从另一个角度来探究泛函极值的必要条件.

Motivation: Noether 定理 $\rightsquigarrow$ 自变量 x 可以作 "形变". $u + \varepsilon\varphi \rightsquigarrow$ E-L 方程, etc; $u \circ \eta_\varepsilon \rightsquigarrow$?

具体地, 设 $\eta_\varepsilon : \overline{\Omega} \to \overline{\Omega}$ 是一个微分自同胚:

$$y = \eta_\varepsilon(x) = x + \varepsilon \bar{X}(x) + o(\varepsilon),$$

其中 $\bar{X}|_{\partial\Omega}=0$. 对于给定的函数 $u\in C^1(\overline{\Omega})$, 令 $v_\varepsilon = u\circ\xi_\varepsilon$, 从而有

$$\begin{aligned} I(v_\varepsilon,\Omega) &= \int_\Omega L(y, v_\varepsilon(y), \nabla u(y))\,\mathrm{d}x, \\ &= \int_\Omega (\eta_\varepsilon(x), u(x), \partial_y\xi_\varepsilon(y)\nabla u(x))\det(\partial_i\eta_\varepsilon^j)\,\mathrm{d}x. \end{aligned}$$

因此, 对于任意的 $u\in C^2$, 我们有

$$\begin{aligned} \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}I(v_\varepsilon,\Omega)\right|_{\varepsilon=0} &= \int_\Omega (\partial_i L\bar{X}^i - L_{p_j^m}\partial_j\bar{X}^i\partial_i u^m + L\partial_i\bar{X}^i)\,\mathrm{d}x \\ &= \int_\Omega (\partial_i L - \partial_i(L) + \partial_j(L_{p_j^m}\partial_i u^m))\bar{X}^i\,\mathrm{d}x \\ &= -\int_\Omega E_L(u)\cdot\left(\partial_i u\bar{X}^i\right)\,\mathrm{d}x. \end{aligned}$$

上述推导过程中无非用到了分部积分公式和复合求导法则. 注意到在第三个表达式中, $\partial_i L$ 代表取值, 而 $\partial_i(L)$ 代表求导.

定义 1.70. *称 $u\in C^1(\overline{\Omega})$ 为 I 的一个***内极小点***, 如果对于任意的 $\bar{X}\in C_0^1(\Omega)$, I 在 v_ε 的变换下满足*

$$\boxed{\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}I(v_\varepsilon,\Omega)\right|_{\varepsilon=0} = 0.}$$

从上述推导过程中我们可以看出, 若 $u\in C^2$ 是一个内极小点, 则有

$$\boxed{E_L(u)\cdot\partial_i u = 0, \qquad \forall 1\le i\le n.}$$

这便是**内极小点的必要条件**. 此外, 我们还可以得出, C^2 的内极小点一定是弱极小点.

1.7.4 例

例 1.71. *设 $L(t,u,p)=t^2(p^2-u^6/3)$. 又设 $\phi_\varepsilon\colon \mathbb{R}\times\mathbb{R}^N\to\mathbb{R}\times\mathbb{R}^N, (t,u)\mapsto(Y,W)$, 其中*

$$Y(t,u,\varepsilon)=(1+\varepsilon)t, \quad W(t,u,\varepsilon)=\frac{u}{\sqrt{1+\varepsilon}}.$$

求证:

1. *L 对应的变分积分*

$$I(u)=\int_0^1 L(t,u,p)\,\mathrm{d}t$$

   是 $\{\phi_\varepsilon\}$ 不变的.

2. *若设 u 为 I 的 E-L 方程的解, 则有*

$$\frac{t^3}{3}u^6 + t^3\dot{u}^2 + t^2 u\dot{u} = \text{const.}$$

证明. 1. 由题设条件, 直接计算可得, 形变 $\eta_\varepsilon(t)=Y(t,u(t),\varepsilon)=(1+t)\varepsilon$, 其诱导的映射

$$v_\varepsilon(y)=W(\eta_\varepsilon^{-1}(y),u(\eta_\varepsilon^{-1}(y)),\varepsilon)=\frac{u\left(\frac{y}{1+\varepsilon}\right)}{\sqrt{1+\varepsilon}}.$$

从而有

$$\begin{aligned}I(v_\varepsilon,\Omega_\varepsilon)&=\int_0^{1+\varepsilon}y^2\left(\frac{\dot u\left(\frac{y}{1+\varepsilon}\right)}{(1+\varepsilon)^3}-\frac{1}{3}\frac{u\left(\frac{y}{1+\varepsilon}\right)^6}{(1+\varepsilon)^3}\right)\mathrm{d}y\\&=\int_0^1 t^2\left(\dot u(t)-\frac{1}{3}u(t)^6\right)\mathrm{d}t=I(u,\Omega).\end{aligned}$$

由此表明 I 关于 $\{\phi_\varepsilon\}$ 是不变的.

2. 直接计算得 $\bar X(t)=\partial_\varepsilon\eta_\varepsilon(t)|_{\varepsilon=0}=t$,

$$\varphi(t)=\partial_\varepsilon W(x,u(x),\varepsilon)|_{\varepsilon=0}-\dot u\bar X=-\frac{1}{2}u(t)-t\dot u(t).$$

将上述计算结果代入至 Noether 恒等式中, 并注意到 u 是 E-L 方程的解, 我们有

$$t^3\left(\dot u^2-\frac{1}{3}u^6\right)-2t^2\dot u\left(\frac{1}{2}u+t\dot u\right)=\text{const}.$$

化简即得欲证等式. □

2 直接方法

2.1 引言

传统方法 $\rightsquigarrow$ E-L 方程, etc. 但仍有许多缺陷:

- E-L 方程为 ODE/PDE, 大多数情况下无法写出其解析解;
- 前述理论大多都是基于 E-L 方程的解存在, 且满足一定正则性的前提下的; 如 C^1,C^2,PWC^1.

直接方法: 直接从定义出发, 找极小化序列, 并验证极小点的存在性.

2.1.1 Dirichlet 原理

例 2.1. *考虑 Laplace 方程*

$$\begin{cases}\Delta u=0 & \text{in }\Omega,\\ u|_{\partial\Omega}=\varphi & \text{on }\partial\Omega.\end{cases}\tag{33}$$

其可以看作 Dirichlet 积分

$$D(u)=\int_\Omega|\nabla u|^2\,\mathrm{d}x\tag{34}$$

在 $M=\{u\in C^2(\overline{\Omega})\colon u|_{\partial\Omega}=\varphi\}$ 上的 E-L 方程.

证明 Laplace 方程边值问题(33)有解 $\rightsquigarrow$ Dirichlet 积分(34)在 M 上存在极小值. 为说明极小点的存在性, Riemann 使用了如下的 **Dirichlet 原理**:

- 因为 D 有下界, 所以必有下确界. 现取一列函数 $\{u_n\} \subseteq M$ 使得 $D(u_n) \to \inf_{u\in M} D(u)$. 因为 $\{u_n\}$ 是有界的, 所以有收敛子列 $u_{n_k} \to u^*$. 这个 u^* 就是要找的极小点: $D(u^*) = \inf_{u\in M} D(u)$.

1870 年, Weierstrass 举出了以下反例:

例 2.2. *考虑下列极值问题:*

$$I(u) = \int_{-1}^{1} x^2 \dot{u}^2 \,\mathrm{d}x, \qquad M = \{u \in C^1[-1,1]\colon u(-1) = -1, u(1) = 1\}.$$

显然有 $I \geq 0$. 另一方面, 对 $\varepsilon > 0$ 取

$$u_\varepsilon(x) = \frac{\arctan \dfrac{x}{\varepsilon}}{\arctan \dfrac{1}{\varepsilon}}.$$

我们有

$$I(u_\varepsilon) = \frac{\varepsilon^2}{\arctan^2 \dfrac{1}{\varepsilon}} \int_{-1}^{1} \frac{x^2}{(x^2+\varepsilon^2)^2} \,\mathrm{d}x \to 0 \qquad (\varepsilon \to 0).$$

由此表明 $\inf_M I = 0$. 但若存在 $u^* \in M$ 使得 $I(u^*) = 0$, 则由 I 的具体表达式可知, $\dot{u} = 0$, 从而有 $u^* = \mathrm{const}$, 这与边值条件矛盾!

事实上, 有界集不一定是列紧集, 且若极小化序列存在收敛子列, 其极限点也未必是极小点.

- 问题的提出: 给定一个拓扑空间 X, 设函数 $f\colon X \to \mathbb{R}$ 下方有界, 从而有下确界 $m = \inf_{x\in X} f(x)$. 取极小化序列 $\{x_j\} \subseteq X$ 使得 $f(x_j) \to m$. 在什么条件下 $\{x_j\}$ 存在收敛子列收敛到极小值点?

Analysis: 考虑下方水平集

$$f_t = \{x \in X\colon f(x) \leq t\},$$

其中 $t \in \mathbb{R}$. 若存在 $t > m$ 使得 f_t 是列紧的, 那么由于当 N 充分大时有 $\{x_j\colon j \geq N\} \subseteq f_t$, 故有子列 $x_{k_j} \to x^* \in f_t \subseteq X$. 进一步地, 若 f 在 $x = x^*$ 处还满足条件

$$f(x^*) \leq \varliminf_{k\to\infty} f(x_{j_k}), \tag{35}$$

则有 $f(x^*) = m = \inf_M f$. 特别地, 若 f 是**序列下半连续的**, 即条件 $x_n \to x$ 蕴含 $f(x) \leq \varliminf f(x_n)$, 则不等式(35)成立.

- 下半连续性. 一般定义: 称 f 是下半连续的, 若对任意的 $t\in\mathbb{R}$, 下方水平集 f_t 是闭的. 特别地, 对于度量空间, 下半连续 $\Leftrightarrow$ 序列下半连续.

- 列紧性. 对于有限维赋范线性空间的情形, 闭 + 有界 $\Rightarrow$ 列紧. 因此, 对于下半连续函数 f, 只要添加**强制性条件**:

$$f(x)\to+\infty\qquad(\|x\|\to+\infty),$$

  则可保证 f_t 是有界的. 对于无限维的情形, 有界闭集不一定是列紧的, 如单位球面 $\mathbb{S}=\{\|x\|=1\}$.

- 选取合适的空间. C^1 不是合适的空间. 一方面, 从 $D(u_n)$ 有界不一定能推出 $\{u_n\}$ 有界; 另一方面, 即使点列是 C^1 有界的, 也未必存在 C^1 收敛的点列.

2.1.2 泛函分析初步: 弱拓扑

定义 2.3. *设 X 是赋范线性空间. 称序列 $\{x_n\}$ 在 X 中***弱收敛***到 x(记作 $x_n\rightharpoonup x$), 如果对任意的 $f\in X^*$, 有 $\langle f,x_n-x\rangle\to 0$, 其中 X^* 为 X 的对偶空间. 称序列 $\{f_n\}$ 在 X^* 中***弱 *收敛** *到 f(记作 $f_n\rightharpoonup^* f$), 如果对任意的 $x\in X$, 有 $\langle f_n-f,x\rangle\to 0$.*

我们可以通过如下方式来定义弱拓扑和弱 * 拓扑:

- 对任意的 $f\in X^*$, 定义映射 $\varphi_f\colon X\to\mathbb{R}\colon x\mapsto\langle f,x\rangle$. 那么 X 上的弱拓扑即为使得映射族 $\{\varphi_f\}_{f\in X^*}$ 均连续的最粗糙的拓扑; 类似地, 对任意的 $x\in X$, 考虑映射 $\phi_x\colon X^*\to\mathbb{R}\colon f\mapsto\langle f,x\rangle$. 则 X^* 上的弱 * 拓扑即为使得映射族 $\{\phi_x\}_{x\in X}$ 均连续的最粗糙的拓扑.

- 容易验证, 在上述定义下, X(或 X^*) 上的弱收敛 (或弱 * 收敛) 的具体表现形式即为定义2.3中所定义的的那样.

- 取 $x_0\in X$, 若设 U 为 x_0 的在弱拓扑意义下的邻域, 则存在 $f_1,\cdots,f_k\in X^*$ 以及 $\varepsilon>0$, 使得

$$U=\{x\in X\colon|\langle f_j,x-x_0\rangle|<\varepsilon,\forall j=1,\cdots,k\}.$$

  类似地, 若设 V 为 $f_0\in X^*$ 在弱 * 拓扑意义下的邻域, 则存在 $x_1,\cdots,x_l\in X$ 以及 $\delta>0$, 使得

$$V=\{f\in X^*\colon|\langle f-f_0,x_m\rangle|<\delta,\forall m=1,\cdots,l\}.$$

注 2.4. *在 X^* 中既有弱收敛, 又有弱 *收敛的概念. 注意到连续嵌入 $X\hookrightarrow X^{**}$, 所以弱收敛蕴含着弱 *收敛. 特别地, 若 X 是自反的, 则 X^* 中的弱收敛和弱 *收敛没有区别.*

例 2.5. 设 $D=[0,1]^N$ 是 $\mathbb{R}^N$ 中的立方体. 设 $\varphi\in L^p(D), 1\le p\le\infty$. 将 φ 作周期延拓, 并令 $\varphi_n(x)=\varphi(nx), \forall n\in\mathbb{Z}_{\ge 0}$, 以及

$$\bar{\varphi}=\int_D \varphi(x)\,\mathrm{d}x,$$

则

$$\varphi \rightharpoonup \bar{\varphi} \text{ in } L^p(D) \qquad (1\le p<\infty),$$

以及

$$\varphi \rightharpoonup^* \bar{\varphi} \text{ in } L^\infty(D).$$

有界不一定蕴含列紧性, 但若设原空间是自反的, 则有界性蕴含弱列紧性.

定理 2.6 (Banach-Alaoglu)**.** 设 X 是 *Banach* 空间, 则 X^* 中的闭单位球

$$\mathbb{B}_{X^*}=\{f\in X^*\colon \|f\|\le 1\}$$

是弱 * 紧的.

命题 2.7. 设 X 是 *Banach* 空间. 则 X^* 是可分的, 当且仅当 $\mathbb{B}_X$ 在弱拓扑下是可度量化的. 对偶地, X 是可分的, 当且仅当 $\mathbb{B}_{X^*}$ 在弱 * 拓扑下是可度量化的.

由上述结果可知, 若设 X 是可分的 Banach 空间, 此时由于 $\mathbb{B}_{X^*}\subseteq X^*$ 是可度量化的, 故 $\mathbb{B}_{X^*}$ 是列紧的, 即 $\mathbb{B}_{X^*}$ 中的任意点列均有弱 * 收敛子列. 特别地, 若 X 还是自反的, 注意到 X 可以看作是 X^* 的对偶空间, 故 X 中的闭单位球 $\mathbb{B}_X$ 也是列紧的. 因此, **若 X 是自反且可分的 Banach 空间, 则它的任意有界序列必有一弱收敛子序列**.

事实上, 上述论断中对 X 的可分性假设可以去掉. 我们先利用 Banach-Alaoglu 证明以下结论:

定理 2.8. 设 X 是自反的 *Banach* 空间, 那么 $\mathbb{B}_X$ 是弱紧的.

证明. 设 J 为 X 到 X^{**} 的典范映射. 由于 X 是自反的, 故 $\mathbb{B}_{X^{**}}=J(\mathbb{B}_X)$. 再根据定理2.6可知, $\mathbb{B}_{X^{**}}$ 是弱 * 紧的, 因此我们只需证明 J^{-1} 是从 X^{**}(赋予弱 * 拓扑) 到 X(赋予弱拓扑) 的连续映射. 事实上, 注意到对任意的 $f\in X^*$ 和 $\xi\in X^{**}$, 有

$$\langle f, J^{-1}\xi\rangle=\langle \xi, f\rangle,$$

由此即可说明 J^{-1} 是连续的. 这便证得了所需结论. □

综上所述, 我们便有如下结果:

推论 2.9. 自反 *Banach* 空间 X 中的任意有界序列必有一弱收敛的子序列.

证明. 设 $\{x_n\}$ 为 X 中任意有界序列. 令

$$M_0=\operatorname{span}\{x_n\}=\left\{y\in X: y=\sum_J a_ix_i, a_i\in\mathbb{R}, \#J<\infty.\right\}$$

由题设条件可知, $M=\overline{M_0}$ 是自反的. 进一步地, 若记 M' 为 $\{x_n\}$ 在 $\mathbb{Q}$ 上张成的线性子空间, 则显然 M' 在 M_0 稠密, 由此表明 M 还是可分的. 故 M^* 也是可分, 且自反的. 若记 $\mathbb{B}_M$ 为 M 中的闭单位球, 由上述分析可知, $\mathbb{B}_M$ 是可度量化的. 此外, 由定理2.8可知, $\mathbb{B}_M$ 还是弱紧的, 故 $\mathbb{B}_M$ 是弱列紧的. 由此足以说明 $\{x_n\}$ 存在弱收敛子序列. □

特别地, 任一 Hilbert 空间和 L^p $(1<p<\infty)$ 空间均是自反的; 在这些空间上, 由推论2.9可知, 有界性蕴含着弱列紧性.

2.1.3 Dirichlet 原理-续

以下是直接方法的基本定理:

定理 2.10. *设 X 是自反的 Banach 空间, $E\subseteq X$ 是弱序列闭的非空子集. 若 $f: E\to\mathbb{R}$ 是弱序列下半连续, 并且是强制的, 则 f 在 E 上达到极小值.*

证明. 取极小化序列 $\{x_n\}$: $f(x_n)\to\inf_E f$. 由于 f 是强制的, 则 $\{x_n\}$ 是有界的. 根据推论2.9, $\{x_n\}$ 有弱 * 收敛子列: $x_{k_n}\rightharpoonup^* x^*$. 再由题设条件可知, $x^*\in E$. 注意到 f 是弱序列下半连续的, 从而有

$$f(x^*)\le\varliminf_{n\to\infty} f(x_{k_n})\le\inf_E f.$$

由此表明 $f(x^*)=\inf_E f$. □

若要运用上述定理解决变分问题, 我们需要注意如下几点:

- 选取合适的函数空间. 我们通常选取自反的 Banach 空间, 或更一般地, 可分 Banach 空间的对偶空间.
- 对应的泛函对于此函数空间上的拓扑是弱序列下半连续, 且是强制的.

以下我们使用定理2.10来验证 Dirichlet 原理.

首先是选取合适的空间. 注意到 $C^1(\overline{\Omega})$ 不是自反的, 且 Dirichlet 积分 $D(u)$ 按 C^1 拓扑也不是强制的, 故我们需要选取其它的空间. 注意到 $D(u)$ 的具体表达式, 我们在 $C^1(\overline{\Omega})$ 上引入如下范数

$$\|u\|_{H^1}:=\left(\int_\Omega(|\nabla u|^2+|u|^2)\,\mathrm{d}x\right)^{1/2}.$$

$C^1(\overline{\Omega})$ 在此范数下不是完备的. 我们将其完备化空间记为 $H^1(\Omega)$. 可以验证, 这是一个 Hilbert 空间, 其上的内积定义为

$$(u,v)_{H^1}:=\int_\Omega(\nabla u\cdot\nabla v+uv)\,\mathrm{d}x.$$

对应地, 我们把 $C_0^1(\Omega)$ 在 $H^1(\Omega)$ 中的闭包记为 $H_0^1(\Omega)$. 注意到 $D(u)$ 不是 $H^1(\Omega)$ 上的范数, 然而, 下述引理表明, 当 Ω 是一个有界区域时, $D(u)$ 可以看作是 $H_0^1(\Omega)$ 上的范数:

引理 2.11 (Poincaré 不等式)**.** *设 $\Omega \subseteq \mathbb{R}^n$ 是一个有界区域, 则对任意的 $u \in C_0^1(\Omega)$, 存在常数 $C = C(\Omega) > 0$, 使得*

$$\|u\|_{L^2} \le C\|\nabla u\|_{L^2}.$$

证明. 取 $\mathbb{R}^n$ 中的立方体 $D = \times_{i=1}^n [a_i, b_i]\,(a_i, b_i \in \mathbb{R})$, 使得 $D \supseteq \Omega$. 对任意的 $\varphi \in C_0^1(\Omega)$, 令

$$\widetilde{\varphi}(x) = \varphi(x)\chi_\Omega(x), \qquad x \in D.$$

记 $x = (x_1, \widetilde{x})$. 由引理1.12可知, 存在只依赖于 Ω 的常数 C_1, 使得

$$\int_J |\widetilde{\varphi}(x_1, \widetilde{x})|\,\mathrm{d}x_1 \le C_1 \int_J |\partial_{x_1}\widetilde{\varphi}(x_1, \widetilde{x})|^2\,\mathrm{d}x_1,$$

其中 J 是 D 沿 x_1 方向的投影. 在上述不等式两侧依次对 $x_2, x_3, \cdots, x_n$ 积分, 并重复使用引理1.12, 注意到 $\widetilde{\varphi}$ 的具体表达式, 即证得所需结论. □

以下我们在 $H^1(\Omega)$ 上验证 Dirichlet 原理的正确性. 对给定的 $\varphi = C^1(\overline{\Omega})$, 在 $M = \varphi + H_0^1(\Omega)$ 上, 其 Dirichlet 积分有表达式

$$D(u) = D(\varphi) + D(v) + 2\int_\Omega \nabla u_0 \cdot \nabla\varphi\,\mathrm{d}x.$$

其中 $u = \varphi + v \in M$.

- $D(u)$ 是强制的. 由上述表达式可以看出, 只需说明当 $D(v) \to +\infty$ 时, 有 $D(u) \to +\infty$. 事实上, 由 Schwarz 不等式和加权的初等不等式可得

  $$\left|\int_\Omega \nabla v \cdot \nabla\varphi\,\mathrm{d}x\right| \le \sqrt{D(v)D(\varphi)} \le D(v) + \frac{1}{4}D(\varphi),$$

  从而有

  $$D(u) \ge \frac{1}{2}D(v) - D(\varphi).$$

  由此即表明 D 是强制的.

- $D(u)$ 是弱序列下半连续的. 在 $H^1(\Omega)$ 中取序列 $\{u_n\}$, 其中 $u_n = \varphi + v_n, v_n \in H_0^1(\Omega)$. 显然,

  $$u_n \rightharpoonup u \text{ in } H^1 \Leftrightarrow v_n \rightharpoonup v \text{ in } H_0^1,$$

  这里 $u = \varphi + v$. 注意到

  $$\psi \mapsto \int_\Omega \nabla\psi \cdot \nabla\varphi\,\mathrm{d}x$$

  可以看作是 $H_0^1(\Omega)$ 的连续线性泛函, 由弱收敛的定义以及 Riesz 表示定理可知,

  $$\int_\Omega \nabla v_n \cdot \nabla\varphi\,\mathrm{d}x \to \int_\Omega \nabla v \cdot \nabla\varphi\,\mathrm{d}x.$$

同理有

$$\int_\Omega \nabla v_n \cdot \nabla v \,\mathrm{d}x \to \int_\Omega \nabla v \cdot \nabla v \,\mathrm{d}x = D(v).$$

此时 Schwarz 不等式给出

$$D(v) = \lim_{n\to\infty} \int_\Omega \nabla v_n \cdot \nabla v \,\mathrm{d}x \le \varliminf_{n\to\infty} D(v)^{1/2} D(v_n)^{1/2},$$

即 $D(v) \le \varliminf D(v_n)$. 由此足以说明 $D(u)$ 的弱序列下半连续性.

综上所述, 我们利用定理2.10证明了 Dirichlet 积分可以在 $H^1(\Omega)$ 上达到极小值, 即验证了 Dirichlet 原理.

最后指出, 一个微分方程的解并不能总是用直接方法求得. 以下反例属于 Hadamard: 令

$$u(r,\theta) = \sum_{m=1}^{\infty} \frac{r^{m!}\sin m!\theta}{m^2}, \qquad (r,\theta) \in [0,1] \times [0,2\pi].$$

可以验证, u 是单位圆内的调和函数, 但由于 $\|\nabla u\|_{L^2} = \infty$, 其对应的 Dirichlet 积分 $D(u) = \infty$.

2.2 Sobolev 空间初步

从上一节的分析中可以看出, 若要使用定理2.10来验证极小点的存在性, 选取的函数空间应该是某个 Banach 空间的对偶空间 (特别地, 自反空间), 这要求该空间至少是完备的; 此外, 由于泛函是含导数的变分积分, 则此空间应至少包含函数的一阶导数 (或更弱意义下的导数) 信息. 以下介绍的 Sobolev 空间便满足这些要求.

2.2.1 基本定义和性质

Motivation: 分部积分公式. 若 $u \in C^k(\Omega)$, 则对任意的 $\alpha \in \mathbb{Z}^n_{\ge 0}$: $|\alpha| \le k$ 和 $\varphi \in C_0^\infty(\Omega)$, 我们有

$$\int_\Omega \partial^\alpha u \varphi \,\mathrm{d}x = (-1)^{|\alpha|} \int_\Omega u \partial^\alpha \varphi \,\mathrm{d}x.$$

定义 2.12. 设 $u, v \in L^1_{\mathrm{loc}}(\Omega)$, $\alpha \in \mathbb{Z}^n_{\ge 0}$. 称 v 为 u 的 α 阶**广义导数**, 如果对于任意的 $\varphi \in C_0^\infty(\Omega)$, 有

$$\boxed{\int_\Omega v\varphi \,\mathrm{d}x = (-1)^{|\alpha|} \int_\Omega u \partial^\alpha \varphi \,\mathrm{d}x.}$$

记作 $v = D^\alpha u$.

定义 2.13. 设 $p \in [1,\infty]$, $k \in \mathbb{Z}_{\ge 0}$. 定义

$$\boxed{W^{k,p}(\Omega) := \{u \in L^p(\Omega) \colon D^\alpha u \in L^p(\Omega), \forall \alpha \in \mathbb{Z}_{\ge 0} \colon |\alpha| \le k\}.}$$

并规定其上的范数为

$$\|u\|_{k,p} := \begin{cases} \left(\sum_{|\alpha|\le k} \int_\Omega |D^\alpha u|^p \, \mathrm{d}x\right)^{1/p} & 1 \le p < \infty, \\ \operatorname*{esssup}_\Omega \sum_{|\alpha|\le k} |D^\alpha u| & p = \infty. \end{cases}$$

容易验证, $(W^{k,p}(\Omega), \|\cdot\|_{k,p})$ 是赋范线性空间, 且是完备的. 称此空间为 ***Sobolev* 空间**.

注 2.14. 我们有时也用下式定义 $W^{k,p}(\Omega)$ 上的范数:

$$\|u\|'_{k,p} := \begin{cases} \sum_{|\alpha|\le k} \left(\int_\Omega |D^\alpha u|^p \, \mathrm{d}x\right)^{1/p} = \sum_{|\alpha|\le m} \|D^\alpha f\|_{L^p} & 1 \le p < \infty, \\ \sum_{|\alpha|\le k} \operatorname*{esssup}_\Omega |D^\alpha u| = \sum_{|\alpha|\le m} \|D^\alpha f\|_{L^\infty} & p = \infty. \end{cases}$$

可以验证, $\|\cdot\|_{k,p}$ 和 $\|\cdot\|'_{k,p}$ 是等价的.

命题 2.15. *Sobolev* 空间 $W^{k,p}(\Omega)$ 具有如下简单性质:

1. 对于有界区域, 有

$$W^{k,\infty}(\Omega) \subseteq W^{k,q}(\Omega) \subseteq W^{k,p}(\Omega) \subseteq W^{k,1}(\Omega), \qquad 1 \le p \le q \le \infty,$$
$$W^{m,p}(\Omega) \subseteq W^{l,p}(\Omega), \qquad 0 \le l \le m.$$

2. 若 $\Omega_1 \subseteq \Omega_2$, 则对任意的 $u \in W^{k,p}(\Omega_2)$, 有 $u|_{\Omega_1} \in W^{k,p}(\Omega_1)$.

3. 设 $u \in W^{k,p}(\Omega), \psi \in C_0^\infty(\Omega)$, 则有 $\psi u \in W_0^{k,p}(\Omega)$, 并且对任意的 $\alpha \in \mathbb{Z}_{\ge 0}$: $|\alpha| \le k$, 有 $\operatorname{supp}(D^\alpha(\psi u)) \subseteq \operatorname{supp} \psi$,

$$D^\alpha(\psi u) = \sum_{\beta\le\alpha} \binom{\alpha}{\beta} D^\beta \psi D^{\alpha-\beta} u,$$

   其中 $\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!(\alpha-\beta)!}$.

4. $W_0^{k,p}(\mathbb{R}^n) = W^{k,p}(\mathbb{R}^n)$.

证明. 1 和 2 是显然的. 对于 3, 先利用数学归纳法归结于 $k=1$ 的情形, 再使用广义导数的定义和分部积分公式即可.

4. 显然有 $W_0^{k,p}(\mathbb{R}^n) \subseteq W^{k,p}(\mathbb{R}^n)$. 另一方面, 注意到 $C_0^\infty(\mathbb{R}^n)$ 在 $L^p(\mathbb{R}^n)$ 中稠密, 且 $W_0^{k,p}(\mathbb{R}^n)$ 是闭的, 故反包含关系也成立. □

例 2.16. 设 $J=[a,b]\subseteq\mathbb{R}$, 则 $W^{1,1}(J)=\mathrm{AC(J)}$, 并且

$$Du(x)=\dot{u}(x),\qquad a.e.\ x\in J.$$

事实上, 对于任意的 $u\in\mathrm{AC(J)}$, 则其导数 $\dot{u}$ 几乎处处存在, 并且属于 $L^1(J)$, 因此 $u\in W^{1,1}(J)$. 此外, 注意到对任意的 $x,y\in J$, 我们有

$$u(x)=\int_y^x\dot{u}(t)\,\mathrm{d}t+u(y),$$

从而对任意的 $\varphi\in C_0^\infty(J)$, 有

$$\int_J u(x)\dot{\varphi}(x)\,\mathrm{d}x=-\int_J\dot{u}(x)\varphi(x)\,\mathrm{d}x,$$

此即 $Du=\dot{u}$, *a.e.*

另一方面, 对任意的 $u\in W^{1,1}(J)$, 令

$$\varphi_n(t)=\begin{cases}n(t-a) & t\in\left[a,a+\frac{1}{n}\right],\\ 1 & t\in\left[a+\frac{1}{n},x-\frac{1}{n}\right],\\ -n(x-t) & t\in\left[x-\frac{1}{n},x\right],\\ 0 & t\in[x,b]\end{cases}\qquad(n=1,2,\cdots),$$

再将 φ_n 磨光, 即对任意 φ_n, 取 $\xi_{n,k}\in C_0^\infty(J),\|\xi_{n,k}\|_{C^1}\leq 2n$, 使得 $\xi_{n,k}$ 在 J 上一致收敛到 φ_n, 且 $\dot{\xi}_{n,k}\to\dot{\varphi}_n$, *a.e.* $t\in J$. 在等式

$$\int_J u(t)\dot{\xi}_{n,k}(t)\,\mathrm{d}t=-\int_J Du(t)\varphi_n(t)\,\mathrm{d}t$$

两端先令 $k\to\infty$, 即得

$$\int_J u(t)\dot{\varphi}_n(t)\,\mathrm{d}t=-\int_J Du(t)\varphi_n(t).$$

再令 $n\rightharpoonup\infty$, 有

$$u(x)-u(a)=\int_a^x Du(t)\,\mathrm{d}t,\qquad\forall x\in J.$$

这表明 $u\in\mathrm{AC(J)}$, 并且 $\dot{u}=Du$, *a.e.* $x\in J$.

此外, 利用绝对连续函数的 *Newton-Leibniz* 公式, 我们还可以证明 $W^{1,\infty}(J)=\mathrm{Lip}(J)$.

最后我们来探究 $W^{k,p}(\Omega)$ 的对偶空间具有何种形式. 这对于研究其上的弱收敛是有帮助的. 考虑如下映射:

$$i\colon W^{k,p}(\Omega)\to\underset{|\alpha|\leq m}{\times}L^p(\Omega),u\mapsto(D^\alpha u)_{|\alpha|\leq m},$$

其中乘积空间 $\times_{|\alpha|\leq k}L^p(\Omega)$ 上的范数规定为

$$\|u\|=\sum_{|\alpha|\leq k}\|D^\alpha u\|_{L^p}.$$

容易验证,在此范数下, $\times_{|\alpha|\le k}L^p(\Omega)$ 是 Banach 空间,其对偶空间 $(\times_{|\alpha|\le k}L^p(\Omega))^*\cong\times_{|\alpha|\le k}L^{p'}(\Omega)$. 这里的同构具体表现为: 规定 $(\times_{|\alpha|\le k}L^p(\Omega))^*$ 中的元素均具有形式 $f=(f_\alpha)_{|\alpha|\le k}$, 且

$$\langle f,u\rangle=\sum_{|\alpha|\le k}\langle f_\alpha,D^\alpha u\rangle.$$

那么有 $f_\alpha\in L^{p'}(\Omega)$, 且 $\|f\|=\max_{|\alpha|\le k}\|f_\alpha\|$.

由上述分析可知, i 是等距嵌入映射. 再根据 Hahn-Banach 定理和 Riesz 表示定理, 则有: $f\in(W^{k,p}(\Omega))^*$, 当且仅当存在 $(\psi_\alpha)_{|\alpha|\le k}\in\times_{|\alpha|\le k}L^{p'}(\Omega)$, 使得

$$\langle f,u\rangle=\sum_{|\alpha|\le k}\int_\Omega D^\alpha u\psi_\alpha\,\mathrm{d}x.$$

从而 $W^{k,p}(\Omega)$ 中的弱收敛可表现为:

$$\boxed{u_j\rightharpoonup u\text{ in }W^{k,p}\iff\sum_{|\alpha|\le k}D^\alpha(u_j-u)\psi_\alpha\,\mathrm{d}x\to 0,\forall(\psi_\alpha)\in\times_{|\alpha|\le k}L^{p'}(\Omega).}$$

推论 2.17. *$W^{k,p}(\Omega)$ 是自反且可分的 Banach 空间.*

证明. 注意到 i 是等距嵌入, 从而 $W^{k,p}(\Omega)$ 可以看作是 $\times_{|\alpha|\le m}L^p(\Omega)$ 的闭子空间. 由此足以证得所需结论. □

2.2.2 延拓, 逼近与嵌入

- **延拓**. 基本问题是: 给定 $u\in W^{k,p}(\Omega)$, 是否总能存在 $\widetilde{u}\in W^{k,p}(\mathbb{R}^n)$, 使得 $\widetilde{u}|_\Omega=u$? 事实上, 对于 $W_0^{k,p}(\Omega)$ 型空间, 不论区域 Ω 如何选取, 延拓总是可能的, 且此延拓是有界的. 我们只需将 u 在 Ω 外定义为零即可. 但对于 $W^{k,p}$ 型空间, 由于函数在边界处可能趋于无穷, 故我们不能采取零延拓的方式. 然而, 若 Ω 具有充分光滑的边界, 那么延拓是可能的:

定理 2.18. *设 Ω 是 $\mathbb{R}^n$ 中的有界区域, 其中 $\partial\Omega$ 是一致 C^k 的. 那么对任意的 $l\in[0,k], p\in[1,\infty)$, 存在有界线性算子 $T: W^{l,p}(\Omega)\to W^{l,p}(\mathbb{R}^n)$, 使得 $Tu(x)=u(x)$, a.e. $x\in\Omega$.*

- **逼近**. 我们已经知道 $C_c^\infty(\Omega)$ 在 $L^p(\Omega)$ 中稠密. 特别地, $C^\infty(\Omega)\cap L^p(\Omega)$ 在 $L^p(\Omega)$ 中稠密. 以下逼近定理将此结果推广到了 Sobolev 空间上:

定理 2.19 (Serrin-Meyers)**.** *设 Ω 是 $\mathbb{R}^n$ 中的有界区域. 若 $p\in[1,\infty)$, 则 $C^\infty(\Omega)\cap W^{k,p}(\Omega)$ 在 $W^{k,p}(\Omega)$ 中稠密.*

证明. 我们先证明局部的逼近, 再利用单位分解得到整体的逼近.

取一族光滑化子 $\{\eta_\varepsilon\}_{\varepsilon>0}$. 令

$$u_\varepsilon(x)=(\eta_\varepsilon*u)(x),\qquad x\in\Omega_\varepsilon,$$

其中 $\Omega_\varepsilon=\{x\in\Omega\colon \operatorname{dist}(x,\partial\Omega)>\varepsilon\}$. 显然对任意的 $\varepsilon>0$, $u_\varepsilon\in C^\infty(\Omega_\varepsilon)$. 再根据恒等逼近的理论可知, 当 $\varepsilon\to 0$ 时, u_ε 在 $W^{k,p}_{\mathrm{loc}}(\Omega)$ 中收敛到 u.

现取 Ω 的一族开覆盖 $\{\Omega_i\}$, 使得 $\Omega=\bigcup_{i=1}^\infty\Omega_i$, 且对任意的 i, 有 $\overline{\Omega_i}\subseteq\Omega_{i+1}$. 再令 $V_i=\Omega_{i+1}\setminus\overline{\Omega_{i-1}}, i=1,2,\cdots$, 其中 $U_0=\varnothing$. 显然有 $\Omega=\bigcup_{i=1}^\infty V_i$. 且对任意的 $x\in\Omega$, 只有有限个 V_i 包含 x. 今取 $\{V_i\}$ 对应的一族单位分解 $\{\zeta_i\}$. 对任意的 $u\in W^{k,p}(\Omega)$, 显然有 $\zeta_i u\in W^{k,p}(\Omega)$, 且 $\operatorname{supp}\zeta_i\subseteq V_i$, 根据前述的局部逼近, 对任意的 $\eta>0$, 我们总可以选取足够小的 $\delta_i>0$, 使得

$$\|\eta_{\delta_i}*(\zeta_i u)-\zeta_i u\|_{k,p}\leq\frac{\eta}{2^i}\qquad(i=1,2,\cdots).$$

记 $u^i=\eta_{\delta_i}*(\zeta_i u)$. 令 $v=\sum_{i=1}^\infty u^i$. 一方面, 注意到对任意的 $x\in\Omega$, $\sum_{i=1}^\infty u^i$ 是有限和, 故 $v\in C^\infty(\Omega)$. 另一方面, 取 $V\subseteq\Omega$ 使得 $\overline{V}\subseteq\Omega$, 我们有

$$\|v-u\|_{W^{k,p}(V)}\leq\sum_{i=1}^\infty\|u^i-\zeta u\|_{W^{k,p}(V)}\leq\eta.$$

注意到 V 是任意的, 故有 $\|v-u\|_{W^{k,p}(\Omega)}\leq\eta$. 由此同时表明 $v\in W^{k,p}(\Omega)$. □

利用光滑函数在 Sobolev 空间中的稠密性, 我们可以轻松地将 Poincaré 不等式推广到 $W_0^{1,p}(\Omega)(1\leq p<\infty)$ 上:

推论 2.20 (Poincaré 不等式)**.** *设 $\Omega\subseteq\mathbb{R}^n$ 是有界区域. 若 $u\in W_0^{1,p}(\Omega), 1\leq p<\infty$, 那么存在常数 $C=C(p,\Omega)>0$, 使得*

$$\|u\|_{L^p}\leq C\|\nabla u\|_{L^p}.$$

回顾前几节中对于空间 H^1 的定义: $C^1(\overline{\Omega})$ 在范数 $\|\cdot\|_{H^1}$ 下的完备化. 注意到 $W^{1,2}(\Omega)$ 是包含 $C^1(\overline{\Omega})$ 的完备空间, 因此 $H^1(\Omega)\subseteq W^{1,2}(\Omega)$. 另一方面, 根据上述定理逼近可知, $W^{1,2}(\Omega)\subseteq H^1(\Omega)$. 从而

$$\boxed{H^1(\Omega)=W^{1,2}(\Omega).}$$

这也是 $H^1(\Omega)$ 的一种等价定义.

此外, 由 Poincaré 不等式可知,

$$\boxed{u\mapsto\left(\int_\Omega|\nabla u|^p\,\mathrm{d}x\right)^{1/p}}$$

是 $W_0^{1,p}(\Omega)$ 上的一个等价范数, 其中 Ω 是 $\mathbb{R}^n$ 中的有界区域, $1\leq p<\infty$.

- **嵌入.** 利用嵌入定理, 我们可以将某些特别的 Sobolev 空间看作是某些常见函数空间的闭子空间. 特别地, 若该嵌入还是紧的 (有界集映到列紧集), 则我们可以使用一些特殊的定理, 如 Arzelà-Ascoli 定理, 来判断有界集合的列紧性.

定理 2.21 (Sobolev 嵌入定理)**.** 设 Ω 是 $\mathbb{R}^n$ 中具有一致 C^m 边界的有界区域, $1 \le q < \infty$, $k \in \mathbb{Z}_{\ge 0}$. 则对任意的 $j \in \mathbb{Z}_{\ge 0}$, 有嵌入关系:

$$\begin{aligned} W^{k,p}(\Omega) &\hookrightarrow L^r(\Omega), \qquad \frac{1}{r} \ge \frac{1}{p} - \frac{k}{n} \ \left(k < \frac{n}{p}\right), \\ W^{k+j,p}(\Omega) &\hookrightarrow C^{j,\lambda}(\overline{\Omega}), \qquad 0 < \lambda \le k - \frac{n}{p} \ \left(k > \frac{n}{p}\right), \end{aligned}$$

其中 $C^{j,\lambda}(\overline{\Omega})$ 是 *Hölder* 型空间.

最常用的是 $k = 1$ 的情形: 记 $p^* = \frac{np}{n-p}$ 为 q 的 **Sobolev 型共轭指标**, 则

$$\begin{aligned} W^{1,p}(\Omega) &\hookrightarrow L^r(\Omega), \qquad 1 \le r \le p^* \ (n > p), \\ W^{1,p}(\Omega) &\hookrightarrow C(\overline{\Omega}) \qquad (p > n). \end{aligned}$$

注 2.22. 当 $k = n = 1, \Omega = (a,b) \subseteq \mathbb{R}$ 时, 嵌入定理的结论很容易从 *Hölder* 不等式推出. 注意到此时 $p^* = p'$, 且广义导数和几乎处处导数是一致的, 故对任意的 $x, y \in (a,b)$, 有

$$|u(x) - u(y)| = \left| \int_a^b \dot{u}(t)\,\mathrm{d}t \right| \le |x-y|^{1/p'} \|\dot{u}\|_{L^p},$$

定理 2.23 (Rellich-Kondrachov)**.** 设 Ω 是 $\mathbb{R}^n$ 中具有一致 C^m 边界的有界区域, $1 \le p \le \infty$, $m \in \mathbb{Z}_{\ge 0}$, 则如下嵌入

$$\begin{aligned} W^{k,p}(\Omega) &\hookrightarrow L^r(\Omega), \qquad 1 \le r < \frac{np}{n-kp} \ \left(k < \frac{n}{p}\right), \\ W^{k,p}(\Omega) &\hookrightarrow C(\overline{\Omega}), \qquad \left(k > \frac{n}{p}\right) \end{aligned}$$

均是紧的.

以下是一种常见的特殊情形:

命题 2.24 (Rellich)**.** 设 Ω 为 $\mathbb{R}^n$ 中的有界区域, $1 \le p < \infty$, 则 $W_0^{1,p}(\Omega)$ 中的单位闭球是 $L^p(\Omega)$ 中的列紧集.

证明. 记 $\mathbb{B}$ 为 $W_0^{1,p}(\Omega)$ 中的单位闭球. 以下证明对任意的 $\varepsilon > 0$, 在 L^p 模下, $\mathbb{B}$ 是完全有界的, 即存在有限的 ε-网.

令 $S = C_0^\infty(\Omega) \cap \mathbb{B}$. 取一族光滑化子 $\{\eta_\delta\}_{\delta > 0}$. 对任意的 $\delta > 0$, 记 $S_\delta = \{v_\delta \colon v \in S\}$, 其中 $v_\delta = v * \eta_\delta$. 由恒等逼近的理论可知, 对任意的 $v \in S$ 和 $\varepsilon > 0$, 存在充分小的 $\delta_0 = \delta(\varepsilon) > 0$, 使得

$$\|v - v_\delta\|_{L^p} < \varepsilon, \qquad \forall \delta \in (0, \delta_0].$$

现固定 $\delta = \delta_0$. 一方面, 由卷积不等式可知, S_{δ_0} 是完全有界的; 另一方面, 注意到

$$|\nabla v_\delta(x)| = |(v * \nabla \eta_\delta)(x)| \le \|v\|_{L^p},$$

其中 C 是一个不依赖于 δ 的常数, 由此表明 S_{δ_0} 还是等度连续的. 根据 Arzelà-Ascoli 定理, 我们可以找到 $\{w_i\}_{i=1}^{l} \subseteq S_{\delta_0}$, 使得对任意的 $v_{\delta_0} \in S_{\delta_0}$, 总存在 w_i, 使得

$$\|v_{\delta_0} - w_i\|_C < \frac{\varepsilon}{5|\Omega|}. \tag{36}$$

注意到此时 w_i 不一定在 $\mathbb{B}$ 内. 然而, 每个 w_i 都对应着一个 $v^i_{\delta_0}$ 使得不等式(36)成立, 其中 $v^i \in S$. 若 $v^i_{\delta_0} \notin \mathbb{B}$, 那么我们选取充分小的 $\delta_i \in (0, \delta_0]$, 使得 $w_i' = v^i_{\delta_i}$ 在支集落在 Ω 内. 此时显然有 $w_i' \in \mathbb{B}$, 且

$$\|w_i - w_i'\|_{L^p} \leq \|w_i - v^i_{\delta_0}\|_{L^p} + \|v^i_{\delta_0} - v_i\|_{L^p} + \|w_i' - v^i\|_{L^p} < \frac{3\varepsilon}{5}.$$

今对任意的 $u \in B$, 我们选取 $v \in S$ 使得 $\|u - v\|_{1,p} \leq \varepsilon/5$, 从而有

$$\|u - w_i'\|_{L^p} \leq \|u - v\|_{L^p} + \|v - v_{\delta_0}\|_{L^p} + \|w_i - w_i'\|_{L^p} < \varepsilon.$$

由此表明 $\{w_i'\} \subseteq \mathbb{B}$ 是 $\mathbb{B}$ 的 L^p 意义下的有限 ε-网. □

2.2.3 Euler-Lagrange 方程

在这一节中, 我们旨在将 E-L 方程推广到 $W^{1,p}(\Omega)$ 上. 给定 Lagrange 函数 $L \in C(\Omega \times \mathbb{R}^N \times \mathbb{R}^{nN})$, 其中 $\Omega \subseteq \mathbb{R}^n$ 为一有界区域, L 是可微的. 考虑泛函

$$I(u) = \int_\Omega L(x, u(x), \nabla u(x))\,\mathrm{d}x.$$

为了使泛函 I 是良定义的, 且 E-L 方程式有意义的, 我们还需要对 L 添加如下假设:

1. L, L_u, L_p 是连续的.
2. $|L(x,u,p)| \leq C(1 + |u|^q + |p|^q)$.
3. $|L_u(x,u,p)| + |L_p(x,u,p)| \leq C(1 + |u|^q + |p|^q)$.

命题 2.25. *在上述假设下, 对任意的 $u \in W^{1,q}(\Omega)$ 和 $\varphi \in C_0^1(\Omega)$, 一阶变分有表达式:*

$$\delta I(u^*, \varphi) = \int_\Omega (L_u(x, u(x), \nabla u(x))\varphi(x) + L_p(x, u(x), \nabla u(x))\nabla\varphi(x))\,\mathrm{d}x. \tag{37}$$

证明. 直接计算即可. 其中极限号和积分号交换顺序的合理性由控制收敛定理保证. □

我们还可以将 L 的假设条件再放松一些:

- 连续性假设. 事实上, 若 L, L_u, L_p 满足如下 **Carathéodory 条件**:

  $$\forall (u,p) \in \mathbb{R}^N \times \mathbb{R}^{nN}, x \mapsto L(x,u,p) \text{ 是可测的},$$
  $$\text{对 a.e. } x \in \Omega, (u,p) \mapsto L(x,u,p) \text{ 是连续的},$$

  则命题2.25的结论仍成立.

• 增长性假设. 利用嵌入定理, $|L_u|+|L_p|$ 的增长幂次可以放松为:

$$|L_u(x,u,p)|+|L_p(x,u,p)| \leq \begin{cases} C(1+|u|^r+|p|^q) & (r \leq q^*) \quad q<n, \\ C(1+|u|^r+|p|^q) & (r \geq 1) \quad q=n, \\ C(1+|p|^q) & \quad q>n. \end{cases} \tag{38}$$

因此, 若 L, L_u, L_p 满足 Carathéodory 条件, 且增长幂次满足条件2和(37), 则一阶变分仍具有如(38)所示的表达式. 特别地, 当 $n=1, \Omega=(a,b)$ 时, 若设 $u^* \in M = \varphi + W_0^{1,q}(\mathbb{R}^n)$ 是 I 在 M 中的极小点, 其中 $\varphi \in W^{1,q}(\Omega)$, 根据变分学基本引理, u^* 还满足如下积分形式的 E-L 方程:

$$\boxed{\int_a^t L_u(t,u^*(t),\nabla u^*(t))\,\mathrm{d}t - L_p(t,u^*(t),\nabla u^*(t)) = \text{const.} \qquad a.e.}$$

注 2.26. 相较于前述转化成一元函数的情形, 我们还可以直接考虑 *Banach* 空间上的微分.

定义 2.27. 设 X 是一个 *Banach* 空间, $U \subseteq X$ 是一个开集. 给定 U 上的函数 $f \in C(U)$. 称 f 在 $x_0 \in U$ 处是 ***Gâteaux* 可微**的, 如果对任意的 $h \in X$, 存在 $c \in \mathbb{R}$, 使得

$$|f(x_0+th)-f(x_0)-tc| = o(t) \qquad (t \to 0).$$

若上式成立, 则称实数 c 为 f 在 x_0 处的 ***Gâteaux* 导数**, 记为 $\mathrm{d}f(x_0,h)$.

由上述定义可知, 若考虑 $W^{1,q}(\mathbb{R}^N)$ 上的泛函 I, 则 *Gâteaux* 导数 $\mathrm{d}I(u,\varphi)$ 中的 φ 属于 $W_0^{1,q}(\Omega)$, 而不是变分 $\delta I(u,\varphi)$ 中的 $C_0^1(\Omega)$. 为使 $\mathrm{d}I(u,\varphi)$ 中的积分有定义, 根据嵌入定理, 我们需要规定 $|L_u|+|L_p|$ 的增长条件:

$$|L_u(x,u,p)|+|L_p(x,u,p)| \leq \begin{cases} C(1+|u|^{r-1}+|p|^{q-1}) & (r \leq q^*) \quad q<n, \\ C(1+|u|^r+|p|^{q-1}) & (r \geq 1) \quad q=n, \\ C(1+|p|^{q-1}) & \quad q>n. \end{cases} \tag{39}$$

在此约束下, 若更设 L 满足条件*1*和*2*, 则 $\mathrm{d}I(u,\varphi)$ 的表达式与 $\delta I(u,\varphi)$ 的表达式相同; 如(37)所示. 注意到增长条件(39)比(38)更强, 故在变分学中我们一般不使用 *Gâteaux* 导数, 而直接使用变分导数.

2.3 弱下半连续性

回顾下半连续性的定义: 称函数 $f\colon X \to \mathbb{R}$ 是**弱下半连续**的, 如果对任意的 $t \in \mathbb{R}$, 下方水平集

$$f_t = \{x \in X : f(x) \leq t\}$$

是闭的. 以下我们主要探究如何判断弱序列下半连续性.

2.3.1 凸性与弱下半连续性

注意到对于 Banach 空间中的凸子集而言, (强) 闭等价于弱闭. 以下命题是对前述结论的推广:

命题 2.28. *设 C 是 Banach 空间 X 中的凸子集, 则 C 是闭的, 当且仅当 C 是弱序列闭的.*

证明. 只需证明弱序列闭 $\Rightarrow$ 闭. 现设 C 是弱序列闭的. 取 $x \in X \setminus C$. 若 C 不是闭集, 则对任意的 $r > 0$, $B_r(x) \cap C \neq \varnothing$. 现对任意的 $n \in \mathbb{Z}_{>0}$, 取 $x_n \in B_{1/n}(x) \cap C$, 则显然有 $x_n \to x$, 从而有 $x_n \rightharpoonup x$. 再根据 C 的弱序列闭性, 故有 $x \in C$, 矛盾. □

我们再来回顾凸函数的概念. 称函数 $f\colon X \to \mathbb{R}$ 是凸的, 如果对于任意的 $x, y \in X$ 和 $t \in [0,1]$, 有

$$f(tx + (1-t)y) \le t f(x) + (1-t) f(y).$$

可以证明, f 是凸函数 $\Rightarrow$ f_t 是凸集. 结合命题2.28的结论和凸函数的定义, 我们便有如下结果:

命题 2.29. *设 X 是 Banach 空间. 若 $f\colon X \to \mathbb{R}$ 是凸函数, 则*

$$f \text{ 序列下半连续} \iff f \text{ 弱序列下半连续}.$$

特别地, 凸 + 强连续 $\Rightarrow$ 弱序列下半连续.

例 2.30. *设 Ω 是 $\mathbb{R}^n$ 中的有界区域. 考虑定义在 $W_0^{1,p}(\Omega)$ 上的泛函*

$$I(u) = \int_\Omega (|\nabla u(x)|^p + c(x)|u(x)|^r)\,\mathrm{d}x,$$

其中 $1 \le p < \infty, 1 \le r \le p^*, c \in L^\infty(\Omega), c \ge 0$. 容易验证, I 是连续凸的, 所以也是弱下半连续的. 若去掉 c 的非负性限制, 此时 I 不再是凸的, 因此 I 未必是弱序列下半连续的. 然而, 若进一步假设 $r < p^*$, 那么 I 仍是弱序列下半连续的. 事实上, 若 u_j 在 $W_0^{1,p}(\Omega)$ 中收敛到 u, 则

$$\int_\Omega |\nabla u(x)|^p\,\mathrm{d}x \le \varliminf_{j\to\infty} \int_\Omega |\nabla u_j(x)|^p\,\mathrm{d}x.$$

我们还有

$$\int_\Omega c(x)|\nabla u(x)|^r\,\mathrm{d}x \le \varliminf_{j\to\infty} \int_\Omega c(x)|\nabla u_j(x)|^r\,\mathrm{d}x,$$

由此足以说明 I 的弱序列下半连续性. 因若不然, 则存在 $\varepsilon_0 > 0$ 以及子列 $\{u_{k_j}\}$, 使得

$$\int_\Omega c(x)|\nabla u(x)|^r\,\mathrm{d}x + \varepsilon_0 > \int_\Omega c(x)|\nabla u_{k_n}(x)|^r\,\mathrm{d}x.$$

根据紧嵌入定理, 存在子列 $\{u_{j'}\} \subseteq \{u_{k_j}\}$, 使得 $u_{j'}$ 在 L^r 中收敛到 u, 从而有

$$\int_\Omega c(x)|u_{j'}(x)|^r\,\mathrm{d}x \to \int_\Omega c(x)|u(x)|^r\,\mathrm{d}x,$$

矛盾!

我们现在回到对变分问题的讨论上去. 要使 $I = I(u)$ 对于 u 是凸的, 则 Lagrange 函数 $L = L(x,u,p)$ 对于 (u,p) 是凸的. 下述命题表明, 对 u 的凸性要求可以由一定的增长性条件替代:

定理 2.31 (Tonelli-Morrey)**.** 设 $L\colon \overline{\Omega}\times\mathbb{R}^N\times\mathbb{R}^{nN}\to\mathbb{R}$, 且满足

- $L\in C^1(\overline{\Omega}\times\mathbb{R}^N\times\mathbb{R}^{nN})$,
- $L\geq 0$,
- 对任意的 $(x,u)\in\Omega\times\mathbb{R}^N$, $p\mapsto L(x,u,p)$ 是凸的,

则 $I(u)=\int_\Omega L(x,u(x),\nabla u(x))\,\mathrm{d}x$ 在 $W^{1,q}(\Omega)$ $(1\leq q<\infty)$ 上是弱下半连续的.

证明. 设 u_j 在 $W^{1,q}$ 中弱收敛到 u. 由紧嵌入定理可知, 存在子列, 不妨记为 $\{u_j\}$, 使得 u_j 在 L^q 中收敛到 u. 再由 Riesz 定理, 存在子列的子列, 仍记为 $\{u_j\}$, 使得 $u_j(x)\to u(x)$ a.e. $x\in\Omega$. 现对任意的 $\varepsilon>0$, 选取紧集 $K\subseteq\Omega$ 使得 $|\Omega\setminus K|<\varepsilon$, 且

- u_j 在 K 上一致收敛到 u (Egorov 定理),
- u 和 ∇u 在 K 上是连续的 (Luzin 定理),
- 若 $I(u)<+\infty$, 则
$$\int_K L(x,u(x),\nabla u(x))\,\mathrm{d}x\geq\int_\Omega L(x,u(x),\nabla u(x))-\varepsilon$$
(Lebesgue 积分的绝对连续性). 若 $I(u)=+\infty$, 则取
$$\int_K L(x,u(x),\nabla u(x))\,\mathrm{d}x>\frac{1}{\varepsilon}.$$

现利用 L 的凸性, 我们有

$$\begin{aligned}
I(u_j)&\geq\int_K L(x,u_j(x),\nabla u_j(x))\,\mathrm{d}x\\
&\geq\int_K L_p(x,u_j(x),\nabla u(x))(\nabla u_j(x)-\nabla u(x))\,\mathrm{d}x+\int_K L(x,u_j(x),\nabla u(x))\,\mathrm{d}x\\
&=\int_K L(x,u_j(x),\nabla u(x))\,\mathrm{d}x+\int_K L_p(x,u(x),\nabla u(x))(\nabla u_j(x)-\nabla u(x))\,\mathrm{d}x\\
+&\int_K (L_p(x,u_j(x),\nabla u(x))-L_p(x,u(x),\nabla u(x)))(\nabla u_j(x)-\nabla u(x))\,\mathrm{d}x\\
&=I_1+I_2+I_3.
\end{aligned}$$

对于 I_1, 注意到 u_j 在 K 上一致收敛到 u, 故由 L 的连续性可知,

$$I_1\to\int_K L(x,u(x),\nabla u(x))\,\mathrm{d}x.$$

对于 I_2, 由于 $L_p \in L^\infty(\Omega)$, 从而有 $\chi_K L_p \in L^\infty(\Omega) \subseteq L^{q'}(\Omega)$. 利用条件 $u_j \rightharpoonup u$ in $W^{1,q}(\Omega)$, 我们有 $\nabla u_j \rightharpoonup \nabla u$ in $L^q(\Omega)$. 由此表明

$$\lim_{j\to\infty} I_2 = \lim_{j\to\infty} \int_K L_p(x,u(x),\nabla u(x))(\nabla u_j(x) - \nabla u(x))\,\mathrm{d}x = 0.$$

最后, 注意到弱收敛列是有界列, 故存在 $C_1, C_2 > 0$ 使得

$$\|\nabla u_j - \nabla u\|_{L^1} \le C_1(\|\nabla u_j - \nabla u\|_{L^q}) \le C_1(\|\nabla u_j\|_{L^q} + \|\nabla u\|_{L^q}) \le C_2.$$

再注意到 $L_p(x,u_j(x),\nabla u(x))$ 在 K 上一致收敛到 $L_p(x,u(x),\nabla u(x))$, 从而有

$$\lim_{j\to\infty} I_3 = 0.$$

综上所述, 当 $I(u) < +\infty$ 时, 有

$$\varliminf_{j\to\infty} I(u_j) \ge \int_K L(x,u(x),\nabla u(x)) \ge I(u) - \varepsilon.$$

由 $\varepsilon > 0$ 的任意性, 即可表明 I 的弱序列下半连续性. $I(u) = +\infty$ 的情形的证明是类似的. □

注 2.32. *事实上, 上述定理中对 L 可微性的假设可以减弱为:*

- *对 a.e. $(x,u) \in \Omega \times \mathbb{R}^N$, $p \mapsto L(x,u,p)$ 是可微的.*
- *L 和 L_p 满足 Carathéodory 条件.*

推论 2.33 (存在性)**.** *设 Ω 是 $\mathbb{R}^n$ 中的有界区域. $\varphi \in W^{1,q}(\Omega)$ $(1 < q < \infty)$. 又设*

- *L 和 L_p 满足 Carathéodory 条件,*
- *存在 $a \in L^1(\Omega)$ 和 $b > 0$, 使得对任意的 $(x,u,p) \in \Omega \times \mathbb{R}^N \times \mathbb{R}^{nN}$, 有 $L(x,u,p) \ge -a(x) + b|p|^q$,*
- *对任意的 $(x,u) \in \Omega \times \mathbb{R}^N$, $p \mapsto L(x,u,p)$ 是凸的,*

那么泛函

$$I(u) = \int_\Omega L(x,u(x),\nabla u(x))\,\mathrm{d}x$$

在 $\varphi + W_0^{1,p}(\Omega)$ 上达到极小值.

证明. 在自反的 Banach 空间 $W_0^{1,q}(\Omega)$ 上考虑泛函 $v \mapsto I(\varphi + v)$. 将 Tonelli-Morrey 定理运用到泛函 $I' = I + \int_\Omega a(x)\,\mathrm{d}x$ 上, 即可说明 I' 是弱序列下半连续的, 从而 I 也是弱序列下半连续的. 为证明极小值点的存在性, 现只需证明 I 是强制的. 事实上, 对任意的 $v \in W_0^{1,q}(\Omega)$, 由 Poincaré 不等式可知, 存在 $\alpha, \beta > 0$ 使得

$$I(\varphi + v) \ge -\int_\Omega a(x)\,\mathrm{d}x + b\int_\Omega |\nabla(u_0 + v)|^q\,\mathrm{d}x \ge \alpha\|v\|_{1,q}^q - \beta.$$

由此表明 I 是强制的. □

例 2.34. 给定 $\mathbb{R}^n$ 上的有界区域 Ω. 那么对任意的 $f\in L^2(\Omega)$, 存在 $u\in H_0^1(\Omega)$, 满足等式 $-\Delta u=f$, 即 *Poisson* 方程

$$\begin{cases}-\Delta u=f & \text{in } \Omega,\\ u=0 & \text{on } \partial\Omega\end{cases}$$

存在广义解. 事实上, 先将 *Poisson* 方程转化成等价的泛函

$$I(u)=\int_\Omega\left(\frac{1}{2}|\nabla u|^2-fu\right)\mathrm{d}x.$$

容易验证, I 满足推论2.33的假设条件, 故 I 存在极小值点.

2.3.2 拟凸性

Motivation: 降低对 Lagrange 函数凸性的要求. 例如, 设 Ω 是 $\mathbb{R}^n$ 中的有界区域. $u=(u^1,\cdots,u^n)\in C^1(\Omega)$, $f\in C^1(\mathbb{R})$ 是凸函数. 考虑 u 的 Jacobi 行列式 $A=\det(\nabla u)$, 定义 Lagrange 函数

$$L:\mathbb{R}^{n\times n}\to\mathbb{R}, A\mapsto f(\det A).$$

可以验证, L 对 p 不是凸的.

考虑特殊情形: L 只依赖于 p.

命题 2.35. 设 $\Omega\subseteq\mathbb{R}^n$ 是一个区域, $L\in C(\mathbb{R}^{n\times N})$. 如果

$$I(u)=\int_\Omega L(\nabla u(x))\,\mathrm{d}x$$

在 $W^{1,\infty}(\Omega)$ 上是弱 * 下半连续的, 那么对于任意的立方体 $D\subset\subset\Omega$ 和 $A\in\mathbb{R}^{n\times N}$, 有

$$\int_D L(A+\nabla\varphi(x))\,\mathrm{d}x\geq L(A)|D|,\quad \forall\varphi\in W_0^{1,\infty}(\mathbb{R}^N).$$

证明. 不妨设 $D=[0,1]^n$. 对任意的 $k\in\mathbb{Z}_{>0}$, 将 D 作 2^k 等分: $D=\bigcup_{l=1}^{2^{kn}}D_l^k$, 其中 D_l^k 是边长为 2^{-k}, 中心在 $c_l^k=2^{-k}(y_1^l+1/2+\cdots y_n^l)$ 的立方体, 其中 $(y_1^l,\cdots,y_n^l)$ 遍历 $(0,1,\cdots,2^k-1)^n$. 对任意的 $v\in W_0^{1,\infty}(D)$, 将其周期扩展到 $\mathbb{R}^n$ 上. 令

$$w_k(x)=\frac{1}{2^k}v(2^k(x-c_l^k)),\quad x\in D_l^k, \forall l=1,2,\cdots,2^{kn}.$$

容易验证,

$$\begin{cases}w_k\rightharpoonup 0,\\ \nabla w_k\rightharpoonup^* 0\end{cases}\quad \text{in } L^\infty(D).$$

现对任意的 $A\in\mathbb{R}^{n\times N}$, 定义 $u_k(x)=Ax+w_k(x), k=1,2,\cdots$, 并且令其在 D 外为 0. 由前述分析可知, $u_k \rightharpoonup^* u=Ax$. 一方面, 我们有 $I(u)=|\Omega|L(A)$; 另一方面,

$$\begin{aligned}
I(u_k) &= \int_D L(A+\nabla w_k(x))\,\mathrm{d}x+\int_{\Omega\setminus D} L(A)\,\mathrm{d}x \\
&= \sum_{l=1}^{2^{kn}}\int_{D_l^k} L(A+\nabla v(2^k(x-c_l^k)))\,\mathrm{d}x+L(A)|\Omega\setminus D| \\
&= \int_D L(A+\nabla v(x))\,\mathrm{d}x+L(A)|\Omega\setminus D|.
\end{aligned}$$

再利用 I 的弱 * 序列下半连续性, 我们便证得了所需结果. □

定义 2.36. 函数 f 称为**拟凸的**, 如果对任意的 $A\in\mathbb{R}^{n\times N}$ 和立方体 $D\subseteq\mathbb{R}^n$, 有

$$\boxed{|D|f(A)\le\int_D f(A+\nabla u(x))\,\mathrm{d}x,\qquad \forall v\in W_0^{1,\infty}(D).}$$

- 设 $p\mapsto L(p)$ 是凸的, 则对任意的 $\varphi\in W_0^{1,\infty}(\Omega)$, 由 Jensen 不等式可得:

$$L(p)=L\left(|D|^{-1}\int_D(p+\nabla\varphi(x))\,\mathrm{d}x\right)\le|D|^{-1}\int_D L(p+\nabla\varphi(x))\,\mathrm{d}x.$$

  这表明 L 也是拟凸的. 即拟凸的确是凸的推广.

- 事实上, 当 $n=1$ 或 $N=1$ 时, 拟凸和凸是等价的.

  - $n=1$. 取 $\xi,\eta\in\mathbb{R}^N$. 对任意的 $\lambda\in[0,1]$, 令

    $$\xi_1=\xi+(1-\lambda)\eta, \xi_2=\xi-\lambda\eta,$$

    即有

    $$\xi=\lambda\xi_1+(1-\lambda)\xi_2, \eta=\xi_1-\xi_2.$$

    定义函数

    $$\varphi(t)=\begin{cases} t(1-\lambda)\eta & t\in[0,\lambda),\\ (1-t)\lambda\eta & t\in[\lambda,1],\end{cases}$$

    再利用 L 的拟凸性, 有

    $$\begin{aligned}
    L(\xi)\le\int_0^1 L(\xi+D\varphi(t))\,\mathrm{d}t &= \int_0^\lambda L(\xi_1)\,\mathrm{d}t+\int_\lambda^1 L(\xi_2) \\
    &= \lambda L(\xi_1)+(1-\lambda)L(\xi_2).
    \end{aligned}$$

    这表明 $p\mapsto L(p)$ 是凸的.

例 2.37. 事实上, 存在不是凸的拟凸函数. 例如, 设 $n=N=2$, 考虑函数 $f\colon \varphi \mapsto \det(\nabla\varphi)$, 则 f 不是凸的. 注意到

$$\det(\nabla\varphi)=\partial_1\varphi^1\partial_2\varphi^2-\partial_1\varphi^2\partial_2\varphi^1=\partial_1(\varphi^1\partial_2\varphi^2)-\partial_2(\varphi^1\partial_1\varphi^2),$$

所以

$$\int_D \det(\nabla\varphi)\,\mathrm{d}x=0,\qquad \forall\varphi\in W_0^{1,\infty}(\Omega).$$

从而对任意的 $A=(a_{ij})$, 有

$$\begin{aligned}|D|^{-1}\int_D f(A+\nabla\varphi)\,\mathrm{d}x=&|D|^{-1}\int_D(\det A+\det(\nabla\varphi)+a_{11}\partial_2\varphi^2+a_{22}\partial_1\varphi^1\\&-a_{12}\partial_1\varphi^2-a_{21}\partial_2\varphi^1)\,\mathrm{d}x\\&=f(A).\end{aligned}$$

这表明 f 是拟凸的.

下述定理表明, 拟凸性同样可以保证泛函的弱序列下半连续性.

定理 2.38 (Morrey-Acerbi-Fusco)**.** 当 $1\le p<\infty$ 时, 若

$$I(u)=\int_\Omega L(\nabla u(x))\,\mathrm{d}x$$

在 $W^{1,p}(\Omega)$ 上是弱序列下半连续的 (或当 $p=\infty$ 时, I 是弱 * 序列下半连续的), 则 L 是拟凸的. 反之, 若 L 是拟凸的, 且满足如下增长性条件:

$$\begin{cases}|L(A)|\le\alpha(1+|A|) & p=1,\\ -\alpha(1+|A|^q)\le L(A)\le\alpha(1+|A|^p) & 1\le q<p<\infty,\\ |L(A)|\le\eta(|A|) & p=\infty,\end{cases}$$

其中 $\alpha>0$ 是常数, η 是一个递增的连续函数. 则当 $1\le p<\infty$ 时, I 在 $W^{1,p}(\Omega)$ 是弱序列下半连续的 (或当 $p=\infty$ 时, I 是弱 * 序列下半连续的).

推论 2.39 (存在性)**.** 设 $L\in C(\mathbb{R}^{n\times N})$ 是拟凸的, 并存在常数 $C_2>C_1>0$ 使得

$$C_1|A|^p\le L(A)\le C_2(1+|A|^p),\qquad 1<p<\infty.$$

那么泛函

$$I(u)=\int_\Omega L(\nabla u(x))\,\mathrm{d}x$$

在 $M=\varphi+W_0^{1,p}(\Omega)$ 上达到极小值, 其中 $\varphi\in W^{1,p}(\Omega)$.

2.4 正则性 $(n=1)$

在得到极小点的存在性后, 还需要验证极小点是 E-L 方程的经典解. 这便是**正则性**问题.

以下恒假设 Lagrange 函数 $L\in C^2$.

命题 2.40. *设 $u^*\in C^1(J)$ 是泛函*

$$I(u)=\int_J L(t,u(t),\dot u(t))\,\mathrm{d}t$$

的极小点. 若对任意的 $t\in J$, $\det(L_{p_ip_j}(t,u^*(t),\dot u^*(t)))\neq 0$, 那么 $u^*\in C^2(J)$.

证明. 由积分形式的 E-L 方程可知, 存在常数 C, 使得

$$L_p(t,u^*(t),\dot u^*(t))=\int_{t_0}^t L_u(\tau,u^*(\tau),\dot u^*(\tau))\,\mathrm{d}\tau-C.$$

将上式等号右侧的函数记为 $q=q(t)$. 令

$$\varphi(t,p)=L_p(t,u^*(t),p)-q(t)\qquad ((t,p)\in J\times\mathbb{R}^N).$$

显然, $\varphi\in C^1$, 且有 $\varphi(t,u^*(t))=0$. 此外, 由于

$$\det(\partial_p\varphi(t,p))=\det(L_{p_ip_j}(t,u^*(t),\dot u^*(t)))\neq 0,$$

故由隐函数定理可知, 对任意的 $t\in J$, 存在 t 的邻域 U 以及唯一的 $\lambda\in C^1(U)$, 等式

$$\varphi(t,\lambda(t))=0$$

对任意的 $t\in U$ 成立. 注意到整体上有等式 $\varphi(t,u^*(t))=0$, 故 $\dot u\in C^1$, 从而 $u\in C^2$. □

例 2.41. *若条件 $\det(L_{p_ip_j}(t,u^*(t),\dot u^*(t)))\neq 0$ 不被满足, 则存在这样的泛函 I, 其极值函数是 C^1 的, 而不是 C^2 的. 事实上, 设 $M=\{u\in C^1[-1,1]\colon u(-1)=0,u(1)=1\}$, Lagrange 函数 $L=L(t,u,p)=u^2(p-2t)^2$. 容易验证, L 对应的变分积分有极小点*

$$u^*(t)=\begin{cases}0 & t<0,\\ t^2 & t\geq 0.\end{cases}$$

此时 $u^*\in C^1\setminus C^2$, 而当 $t<0$ 时, $L_{pp}(t,u^*(t),\dot u^*(t))=2(u^*(t))^2=0$.

定理 2.42. *设 L 满足如下增长条件:*

$$\begin{cases}|L|+|L_u|+|L_p|\leq C(1+|p|^r) & 1<r<\infty,\\ \text{None} & r=\infty.\end{cases}$$

又设对任意的 $(t,u,p)\in J\times\mathbb{R}^N\times\mathbb{R}^N$, $(L_{p_ip_j}(t,u,p))$ 是正定的. 若 $u^*\in W^{1,r}(J)$ 是泛函

$$I(u)=\int_J L(t,u(t),\dot u(t))\,\mathrm{d}t$$

的极小点, 则可以改变 u^* 在一个零测集上的值, 使得 $u^*\in C^2$.

证明. 只需证明, 改变 u^* 在一个零测集上的值以后, 有 $u^* \in C^1$.

由题设条件可知, 存在常数 C, 使得

$$L_p(t, u^*(t), \dot{u}^*(t)) = q(t), \qquad \text{a.e. } t \in J,$$

其中

$$q(t) = \int_{t_0}^{t} L_u(\tau, u^*(\tau), \dot{u}^*(\tau))\,\mathrm{d}\tau - C.$$

注意到对 L 的增长性假设, 我们还有 $q \in \mathrm{AC(J)}$. 定义函数

$$\varphi\colon J \times \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}, (t, u, p, q) \mapsto L_p(t, u, p) - q.$$

根据隐函数定理, 方程 $\varphi(t, u, p, q) = 0$ 存在唯一的局部 C^1 解 $p = \lambda(t, u, q)$. 尽管 $\dot{u}$ 可能不是连续的, 但若能说明上述隐函数定理得到的局部解还是整体唯一的, 则有

$$\dot{u}(t) = \lambda(t, u^*(t), q(t)), \qquad \text{a.e. } t \in J,$$

再注意到嵌入关系 $W^{1,r}(J) \hookrightarrow C(\overline{J})$, 故 $\dot{u}^*$ 是几乎处处连续的. 事实上, 设 p_1 和 p_2 是方程 $\varphi = 0$ 的整体解, 那么

$$q = L_p(t, u, p_1) = L_p(t, u, p_2),$$

从而有

$$0 = (L_p(t, u, p_1) - L_p(t, u, p_2)) \cdot (p_1 - p_2) = (B(p_1 - p_2)) \cdot (p_1 - p_2),$$

其中

$$B = \int_0^1 L_{pp}(t, u, p_1 + \tau(p_2 - p_1))\,\mathrm{d}\tau.$$

由题设条件可知, B 是正定的, 从而 $p_1 = p_2$. □

例 2.43. *若去掉定理2.42中对于矩阵 $(L_{p_i p_j}(t, u^*(t), \dot{u}^*(t)))$ 整体的正定性假设, 则极小点未必是 C^1 的. 考虑 Lagrange 函数 $L = L(p) = (p^2 - 1)^2$, 以及定义域*

$$M = \{u \in \mathrm{Lip}[0,1]\colon u(0) = u(1) = 0\}.$$

容易验证, L 对应的变分积分有极小值 0, 此时 $L_{pp} = 4(3p^2 - 1)$ 不是正定的. 若 $u \in C^1$ 是极小点, 则有 $\dot{u}(t) = \pm 1$, 但无论处于何种情况, 都不可能有满足边值条件 $u(0) = u(1) = 0$ 的解.

总结:

$$\text{直接方法} \rightsquigarrow \text{找极小化序列} \begin{cases} \text{存在性} \begin{cases} \text{合适的空间: Sobolev 空间} \\ \text{弱下半连续性} \end{cases} \\ \text{正则性} \end{cases}$$

最后, 我们以几个具体的变分问题为例, 探究其极小点的存在性与正则性.

例 2.44 (两点边值问题)**.** 记区间 $J=(t_0,t_1)$, 环面 $\mathbb{T}^2=\mathbb{R}^2/\mathbb{Z}^2$. 给定 $a_0,a_1\in\mathbb{T}^2, F\in C^2(J\times\mathbb{T}^2)$. 考虑如下方程

$$\ddot{u}(t)=\nabla_u F(t,u(t)) \tag{40}$$

的 C^2 解, 且满足边值条件 $u(t_i)=a_i, i=0,1$.

先将 F 周期延拓到全空间上去, 即对任意的 $t\in J$, 定义

$$F(t,u_1+1,u_2)=F(t,u_1+1,u_2)=F(t,u_1,u_2).$$

将扩张后的函数仍记为 F, 定义泛函

$$I(u)=\int_J\left(\frac{1}{2}|\dot{u}(t)|^2+F(t,u(t))\right)\mathrm{d}t$$

以及定义域 $M=u^*+H_0^1(J)$, 这里

$$u^*(t)=\frac{a_0(t_1-t)+a_1(t-t_0)}{t_1-t_0}.$$

容易验证, I 对应的 E-L 方程正是(40). 在 $H_0^1(J)$ 上赋予等价范数 $v\mapsto\left(\int|\dot{v}|^2\,\mathrm{d}x\right)^{1/2}$.

- 显然 M 是弱序列闭的. 此外, I 在 M 上还是强制的.
- 设 u_n 在 $H^1(\Omega)$ 上弱收敛到 u, 则 $\{u_n\}$ 是有界集. 注意到嵌入 $H^1(J)\hookrightarrow C(\bar{J})$ 是紧的, 且 F 是连续的, 故 I 还是弱序列下半连续的 (参考例2.30中使用的方法).

综上所述, I 在 M 上有极小点 u^*. 此时 L 满足如定理2.42中所列出的增长性限制, 且 $(L_{p_ip_j})$ 为单位阵, 当然是正定的. 因此 $u^*\in C^2$, 从而 u^* 是(40)的经典解.

例 2.45 (强迫振动的周期解)**.** 设 e 为周期为 T, 且在 $[0,T]$ 上的平均值为零的连续函数, 即

$$\int_0^T e(t)\,\mathrm{d}t=0. \tag{41}$$

求下列方程周期为 T 的解:

$$\ddot{u}(t)+a\sin u(t)=e(t).$$

记 $H^1_{\mathrm{per}}(0,T)$ 为周期为 T 的 C^∞ 函数在 $H^1(0,T)$ 中的闭包. 在其上考虑泛函

$$I(u)=\int_0^T\left(\frac{1}{2}|\dot{u}(t)|^2+a\cos u(t)+e(t)u(t)\right)\mathrm{d}t.$$

容易验证, I 对应的 E-L 方程即为(42)式. 注意到 I 按照 H^1 范数并不是强制的, 故我们需要在 $H^1_{\mathrm{per}}(0,T)$ 上选取其它的等价范数. 事实上, 若将 I 改写为

$$I(u)=\int_0^T\left(\frac{1}{2}|\dot{u}(t)|^2+a\cos u(t)-E(t)u(t)\right)\mathrm{d}t,$$

其中
$$E(t)=\int_0^t e(\tau)\,\mathrm{d}\tau,$$
则 I 有估计
$$I(u)\geq\frac{1}{2}\|\dot{u}\|_{L^2}^2-|a|-C\|E\|_{L^\infty}\|\dot{u}\|_{L^2},$$
其中 C 是一个正常数. 由上式可以看出, 若能说明 $u\mapsto\|\dot{u}\|_{L^2}$ 是 $H^1_{\mathrm{per}}(0,T)$ 上的一个等价范数, 则 I 便是强制的.

引理 2.46 (Wirtinger 不等式)**.** 设 $u\in H^1_{\mathrm{per}}(0,T)$. 若
$$\overline{u}=\int_0^T u(t)\,\mathrm{d}t=0,$$
那么
$$\int_0^T|\dot{u}(t)|^2\,\mathrm{d}t\geq\frac{4\pi^2}{T^2}\int_0^T|u(t)|^2\,\mathrm{d}t. \tag{42}$$
证明. 先对光滑函数证明(42)式. 由于 $\overline{u}=0$, 故 u 有 Fourier 展开
$$u(t)=\sum_{k=1}^{\infty}\left(a_k\cos\frac{2\pi k}{T}t+b_k\sin\frac{2\pi k}{T}t\right).$$
从而
$$\dot{u}(t)=\frac{2\pi}{T}\sum_{k=1}^{\infty}\left(-ka_k\sin\frac{2\pi k}{T}t+kb_k\cos\frac{2\pi k}{T}t\right).$$
再利用 Parseval 等式, 便有
$$\int_0^T|\dot{u}(t)|^2\,\mathrm{d}t=\frac{4\pi^2}{T^2}\sum_{k=1}^{\infty}k^2(|a_k|^2+|b_k|^2)\geq\frac{4\pi^2}{T^2}\sum_{k=1}^{\infty}(|a_k|^2+|b_k|^2)=\int_0^T|u(t)|^2\,\mathrm{d}t.$$

最后, 对任意的 $u\in H^1_{\mathrm{per}}(0,T)$ 且满足 $\overline{u}=0$, 取一族周期为 T 的光滑函数 $\{u_n\}$, 使得 u_n 在 $H^1(0,T)$ 中收敛到 u. 注意到此时 $\overline{u_n}$ 不一定等于零. 令 $v_n=u_n-\overline{u_n}$, 则 $\overline{v_n}=0$. 此外, 由 Hölder 不等式容易推出, $\overline{u_n}\to\overline{u}=0$, 从而 $v_n\to u$. 利用第一段中证明的结果, 我们便在一般情况下证明了(42)式. □

为使用 Wirtinger 不等式, 对任意的 $u\in H^1_{\mathrm{per}}(0,T)$, 作分解 $u=\widetilde{u}+\overline{u}$. 显然, $\overline{\widetilde{u}}=0$. 此外, 注意到 $I(u)=I(u+2\pi)$, 这表明我们不需在整个 $H^1_{\mathrm{per}}(0,T)$ 上考虑 I, 而是取集合
$$M=\{u=\xi+\eta:\xi\in H^1_{\mathrm{per}}(0,T),\overline{\xi}=0,\eta\in[0,2\pi]\},$$
并将 I 限制在 M 上. 这样做的好处是 $\overline{u}$ 只在 $[0,2\pi]$ 上变化.

最后, 我们在 M 上考虑 I. 显然 M 是弱序列闭的. 由上述分析可知, I 在 M 上还是强制的. 此外, 容易验证 I 是弱序列下半连续的, 故由存在性定理可知, I 在 $M\subseteq H^1_{\mathrm{per}}$ 上存在极小点. 再利用正则性定理, 此极小点还是 C^2 的.

2.5 专题一: 正交投影

对于某类对应于线性微分方程的特殊变分问题,其极小点的存在性可以由 Hilbert 空间的某些性质所保证,而不需要使用弱拓扑的工具.

例 2.47. *考虑 Poisson 方程*

$$\begin{cases} -\Delta u = f, & in\ \Omega, \\ u = 0, & on\ \partial\Omega, \end{cases}$$

其中 $\Omega \subseteq \mathbb{R}^n$ 为有界区域, $f \in L^2(\Omega)$. 在等式 $-\Delta u = f$ 两边与 $\varphi \in H_0^1(\Omega)$ 相乘后在积分, 并利用 Green 公式和边值条件, 即得

$$(\varphi, u)_{H_0^1} = \int_\Omega \nabla u \cdot \nabla \varphi \, \mathrm{d}x = \int_\Omega f \varphi \, \mathrm{d}x.$$

由 Cauchy-Schwarz 不等式和 Poincaré 不等式, 有

$$\left| \int_\Omega f \varphi \, \mathrm{d}x \right| \le \|f\|_{L^2} \|\varphi\|_{L^2} \le C \|f\|_{L^2} \nabla \varphi\|_{L^2},$$

其中 C 是一个常数. 这表明 $\varphi \mapsto \int_\Omega f \varphi \, \mathrm{d}x$ 可以看作是 $H_0^1(\Omega)$ 上的有界线性泛函. 按 Riesz 表示定理, 存在 $u^* \in H_0^1(\Omega)$, 使得

$$\int_\Omega f \varphi \, \mathrm{d}x = (\varphi, u^*)_{H_0^1}.$$

从而对任意的 $\varphi \in H_0^1(\Omega)$, 有 $(\varphi, u)_{H_0^1} = (\varphi, u^*)_{H_0^1}$. 故 $u = u^* \in H_0^1(\Omega)$ 便是 Poisson 方程的解.

在上述的推导过程中, Riesz 表示定理起到了决定性的作用, 而 Riesz 定理的证明依赖于下述正交投影定理:

定理 2.48 (正交投影)**.** *设 M 为 Hilbert 空间 H 的闭子空间, 则对任意的 $x \in H$, 存在 $y \in M$, 使得 $(x-y) \perp M$.*

值得注意的是, Hilbert 空间 H 中正交投影的存在性可以化归为一个变分问题:

$$\begin{aligned} &\min \|x - z\|, \\ &\text{s.t. } z \in M. \end{aligned}$$

这里 $x \in H$. 事实上, 若存在 $z^* \in M$ 使得 $\|x - z^*\| = \min_{z \in M} \|x - z\|$, 对任意的 $y \in H$, 考虑定义在 $B_r(0)$ 上的函数

$$g(t) = \|x - (ty + (1-t)z^*)\|^2 = \|y - z^*\|^2 t^2 - 2(x - z^*, y - z^*)t + \|x - z^*\|^2.$$

这是一个关于 t 的二次函数, 它在 $t = 0$ 处达到极小值. 从而有

$$\dot{g}(0) = -2(x - z^*, y - z^*) = 0,$$

即 $(x-z^*)\perp M$.

极小点 $z^*\in H$ 的存在性也可以用初等方法证明. 具体地, 令 $m=\inf_{z\in M}\|x-z\|$, 取极小化序列 $\{z_j\}\subseteq M$, 使得

$$\|z_j-x\|<m+\frac{1}{j},\qquad \forall j.$$

对任意的 $\varepsilon>0$, 当 j,k 充分大时, 平行四边形法则给出

$$\|z_j-z_k\|^2=2(\|z_j-x\|^2+\|z_j-x\|^2)-4\left\|\frac{z_j+z_k}{2}-x\right\|^2\le 4(m+\varepsilon)^2-4m^2,$$

由此表明 $\{z_j\}$ 是 Cauchy 列, 从而存在 $z^*\in H$, 使得 $z_j\to z^*$. 又因为 M 是闭的, 故 $z^*\in M$, 且有 $\|z^*-x\|=m$.

综上所述, 利用 Hilbert 空间的完备性和几何性质, 我们验证了 Dirichlet 原理, 而不引入弱拓扑的概念.

2.6 专题二: 特征值问题

设 $\Omega\subseteq\mathbb{R}^n$ 是一个有界区域. 考虑边值问题 (算子 $-\Delta$ 的**特征值问题**)

$$\begin{cases}-\Delta u=\lambda u & \text{in } \Omega,\\ u=0 & \text{on } \partial\Omega.\end{cases}\tag{43}$$

问对于哪些 $\lambda\in\mathbb{R}$, 上述特征值问题存在非零解? 称对应着非零解的数 λ 为**谱**. 若这个非零解 $u\in L^2(\Omega)$, 则称其为**特征函数**, 并称相应的 λ 为**特征值**.

利用约束变分方法, 我们有如下结果:

定理 2.49. 设 Ω 为 $\mathbb{R}^n$ 中的有界区域, 则方程(43)有一列特征值 $\{\lambda_k\}$, 且满足

- $0<\lambda_1\le\lambda_2\le\cdots\le\lambda_i\le\cdots;$
- $\lambda_i\to+\infty,$

以及对应的特征函数 $\{\varphi_i\}\subseteq H_0^1(\Omega)$, 且满足

- $-\Delta\varphi_j=\lambda_j\varphi_j, \|\varphi_j\|_{L^2}=1, \forall i=1,2,\cdots;$
- $(\varphi_j,\varphi_k)_{H_0^1}=0=(\varphi_j,\varphi_k)_{L^2}, \forall i\neq j.$

进一步地, $\{\varphi_i/\sqrt{\lambda_i}\}$ 是 $H_0^1(\Omega)$ 中的一组正规正交基, 从而 $\{\varphi_i\}$ 是 $L^2(\Omega)$ 中的一组正规正交基.

证明. 考虑泛函

$$I(u)=\int_\Omega|\nabla u|^2\,\mathrm{d}x$$

以及

$$N(u)=\int_\Omega |u|^2\,\mathrm{d}x.$$

容易验证, 若 $\varphi_1\in C^2(\overline{\Omega})$ 是变分问题

$$\min_{u\in H_0^1(\Omega)\cap N^{-1}(1)} I(u) \tag{44}$$

的极小点, 则由 Lagrange 乘子法可知, 存在 $\lambda_1\in\mathbb{R}$ 满足 $-\Delta\varphi_1=\lambda_1\varphi_1$. 以下验证(44)解的存在性. 显然 I 是强制的, 并且是弱序列下半连续的. 结合存在性和正则性定理, 以下只需验证集合

$$M=H_0^1(\Omega)\cap N^{-1}(1)$$

是弱序列闭的. 具体地, 设 $\{u_j\}$ 在 H_0^1 中收敛到 u, 注意到紧嵌入 $H_0^1(\Omega)\hookrightarrow L^2(\Omega)$, 故存在子列 $\{u_{j'}\}\subseteq\{u_j\}$, 使得 $u_{j'}$ 在 L^2 中收敛到 u. 由于 $\|u_{j'}\|_{L^2}=1$, 从而有 $\|u\|_{L^2}=1$. 这表明 $u\in M_1$, 故 M_1 是弱序列闭的.

再取集合

$$M_2=\left\{u\in M_1:\ \int_\Omega u\varphi_1\,\mathrm{d}x=0\right\},$$

并考虑变分问题 $\min_{u\in M_2}I(u)$. 若 $\varphi_2\in C^2$ 是极小点, 则存在 $\lambda_2,\mu_2\in\mathbb{R}$ 使得

$$-\Delta\varphi_2=\lambda_2\varphi_2+\mu_2\varphi_1. \tag{45}$$

一方面, 在等式 $-\Delta\varphi_1=\lambda_1\varphi_1$ 两边乘以 φ_2 再积分, 并利用 Green 公式, 即得

$$\int_\Omega \nabla\varphi_1\cdot\nabla\varphi_2\,\mathrm{d}x=-\int_\Omega\Delta\varphi_1\varphi_2\,\mathrm{d}x=\lambda_1\int_\Omega\varphi_1\varphi_2\,\mathrm{d}x=0;$$

另一方面, 在(45)等式两侧乘以 φ_1 后积分, 得

$$\int_\Omega\nabla\varphi_1\cdot\varphi_2\,\mathrm{d}x=\mu_2\int_\Omega|\varphi_1|^2\,\mathrm{d}x=\mu_2.$$

故 $\mu_2=0$. 这表明 λ_2 是一个特征值, φ_2 是相应的特征函数, $\varphi_1\neq\varphi_2$, 且

$$\lambda_2=I(\varphi_2)=\min\{I(u):\ u\in M_2\}\geq\lambda_1.$$

以下只需要说明 $\varphi\in C^2$ 的极小点. 同理, 我们只需说明 M_2 是弱序列闭的, 而这一点可以由紧嵌入定理保证.

如此进行下去, 在第 i 步时, 令

$$M_i=\left\{u\in M_{i-1}:\ \int_\Omega u\varphi_{i-1}\,\mathrm{d}x=0\right\}.$$

可以证明它是弱序列闭的, 从而变分问题 $\min_{u\in M_i}I(u)$ 有解 $\varphi_i\neq 0$, 满足

$$-\Delta\varphi_i=\lambda_i\varphi_i+\sum_{j=1}^{i-1}\mu_j\varphi_j.$$

同理, 我们有 $\mu_1 = \cdots = \mu_{i-1} = 0$, 由此即表明 λ_i 是一个特征值, φ_i 是对应的特征函数, 且

$$\lambda_i = I(\varphi_i) = \min\{I(u) \colon u \in M_i\} \geq \lambda_{i-1}.$$

以下说明 $\lambda_i \to +\infty$. 若存在 $C > 0$ 使得 $\lambda_i \leq C, \forall i$, 那么

$$\int_\Omega |\nabla \varphi_i|^2 \,\mathrm{d}x = \lambda_i \int_\Omega |\varphi_i|^2 \,\mathrm{d}x = \lambda_i \leq C, \qquad \forall i.$$

由此表明 $\{\varphi_i\}$ 是 $H_0^1(\Omega)$ 中的有界列, 从而有弱收敛子列 $\{\varphi_{j'}\}$. 一方面, 根据紧嵌入定理, 有子列 $\{\varphi_{j''}\}$ 在 L^2 中收敛到 φ; 另一方面, 注意到 $(\varphi_j, \varphi_k)_{L^2} = 0, j \neq k$, 故有

$$\int_\Omega |\varphi_j - \varphi_k|^2 \,\mathrm{d}x = \int_\Omega (\varphi_j^2 + \varphi_k^2) \,\mathrm{d}x = 2, \qquad i \neq j.$$

从而有

$$\int_\Omega |\varphi^* - \varphi_{j'}|^2 \,\mathrm{d}x = 2,$$

矛盾.

最后证明 $\{\varphi_i\}$ 是完备的, 即对任意的 $u \in H_0^1(\Omega)$, 部分和

$$s_m = \sum_{i=1}^m c_i \varphi_i$$

在 H_0^1 中收敛到 u, 这里 $c_i = (u, \varphi_i)_{H_0^1}$. 一方面, 由正交性直接计算得

$$\int_\Omega |\nabla (u - s_m)|^2 \,\mathrm{d}x = \int_\Omega |\nabla u|^2 \,\mathrm{d}x - \sum_{n=1}^m |c_n|^2.$$

另一方面, 注意到 $u - s_m \in M_{m+1}$, 故有

$$\int_\Omega |\nabla (u - s_m)|^2 \,\mathrm{d}x \geq \lambda_{m+1} \int_\Omega |u - s_m|^2 \,\mathrm{d}x.$$

因此

$$\int_\Omega |u - s_m|^2 \,\mathrm{d}x \leq \frac{1}{\lambda_{m+1}} \int_\Omega |\nabla u|^2 \,\mathrm{d}x \to 0 \qquad (m \to \infty),$$

即 s_m 在 L^2 意义下收敛到 u. 进一步地, 由 Bessel 不等式可知, 当 $m > n$ 时, 有

$$\int_\Omega |\nabla (s_m - s_n)|^2 \,\mathrm{d}x = \sum_{i=n+1}^m |c_i|^2 \to 0 \qquad (n \to +\infty).$$

结合上述分析可知, s_m 也在 H_0^1 中收敛到 u. □

例 2.50. *在区间 $J = [a,b] \subseteq \mathbb{R}$ 上给定 $p \in C^1(J), q \in C(J)$, 并假设存在 $\alpha \in \mathbb{R}$ 使得 $p(x) \geq \alpha$, 且 $q(x) \geq 0$. 考虑具有 Dirichlet 边界条件的 Sturm-Liouville 问题*

$$\begin{cases} \mathcal{L}u = -(pu')' + qu = \lambda u & \text{in } J, \\ u(a) = u(b) = 0. \end{cases}$$

和定理2.49中的讨论过程类似, 在空间 $H_0^1(J)$ 上考察泛函

$$I(u)=\frac{1}{2}\int_J(p|u'|^2+q|u|^2)\,\mathrm{d}x,$$

此时 $H_0^1(\Omega)$ 上的内积规定为

$$(u,v)=\int_J(pu'v'+quv)\,\mathrm{d}x.$$

通过逐次引入约束 $M_1=\{u\in H_0^1(J)\colon\|u\|^2=1\}$, $M_2=\{u\in M_1\colon(u,\varphi_1)=0\}\cdots$, 同样可以得到一列恒正递增区域无穷的特征值 $\{\lambda_i\}$, 以及对应的特征函数 $\{\varphi_i\}\subseteq C^2(J)$. 此外, $\{\varphi_i\}$ 也是 $H_0^1(\Omega)$ 中的一组正交基.

注 2.51. 上述讨论均涉及的是 *Dirichlet* 边界的情形. 对于 *Neumann* 边界, 由于 *Poincaré* 不等式在 $H^1(\Omega)$ 上不再成立, 对应的泛函未必是强制的. 然而, 注意到如下结果:

命题 2.52 (Poincaré 不等式)**.** 设 Ω 为 $\mathbb{R}^n$ 中的有界区域, 其边界是 C^1 的. 那么对任意的 $u\in W^{1,p}(\Omega)$ $(1\le p\le\infty)$, 存在常数 $C=C(n,p,\Omega)>0$, 使得

$$\left\|u-\fint_\Omega u\,\mathrm{d}x\right\|_{L^p}\le C\|\nabla u\|_{L^p}.$$

由此可知, 我们可以将对应的泛函限制到子空间

$$X=\left\{u\in H^1(\Omega)\colon\int_\Omega u\,\mathrm{d}x=0\right\}$$

上, 此时 *Poincaré* 不等式在 X 上仍成立, 我们可以在 X 上 (相当于添加一个约束条件, 其对应的 *Lagrange* 乘子事实上为零) 求对应泛函的极小值.

在前述的讨论中, 方程(43)的特征值 $\{\lambda_i\}$ 是通过递归给出的. 下述定理直接给出了 λ_i 的刻画:

定理 2.53 (Courant 极小极大定理)**.** 记 $\{\lambda_n\}$ 为方程(43)的特征值, 则有

$$\lambda_n=\max_{E_{n-1}}\min_{u\in E_{n-1}^\perp\setminus\{0\}}\frac{\displaystyle\int_\Omega|\nabla u|^2\,\mathrm{d}x}{\displaystyle\int_\Omega|u|^2\,\mathrm{d}x},$$

其中 E_{n-1} 是 $H_0^1(\Omega)$ 中任意 $(n-1)$ 维线性子空间.

证明. 记 $E_{n-1}=\operatorname{span}\{v_1,\cdots,v_n\}$ 是 $H_0^1(\Omega)$ 中任一 $(n-1)$ 维线性子空间, 其中 $\{v_n\}\subseteq H_0^1(\Omega)$ 线性无关. 再记

$$\mu(E_{n-1})=\min_{u\in E_{n-1}^\perp\setminus\{0\}}\frac{\displaystyle\int_\Omega|\nabla u|^2\,\mathrm{d}x}{\displaystyle\int_\Omega|u|^2\,\mathrm{d}x}.$$

一方面, 由于 $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$,

$$\mu(E_{n-1}) \le \frac{\displaystyle\int_\Omega |\nabla u|^2\,\mathrm{d}x}{\displaystyle\int_\Omega |u|^2\,\mathrm{d}x} = \frac{\displaystyle\sum_{i=1}^n \lambda_i |c_i|^2}{\displaystyle\sum_{i=1}^n |c_i|^2} \le \lambda_n;$$

另一方面, 令 $\{\varphi_1,\cdots,\varphi_n\}$ 为前 n 个特征函数, 注意到

$$(E_{n-1}^{\perp} \setminus \{0\}) \cap \mathrm{span}\{\varphi_1,\cdots,\varphi_n\} \ne \varnothing,$$

取 $\widetilde{E_{n-1}} = \mathrm{span}\{\varphi_1,\cdots,\varphi_{n-1}\}$, 便有

$$\max_{E_{n-1}} \mu(E_{n-1}) \ge \mu(\widetilde{E_{n-1}}) = \lambda_n.$$

□

2.7 专题三: Ekeland 变分原理

下述的 Ekeland 变分原理给出了选取一串近似极小点的方法. 此外, 若将其与近代变分学常用的紧性条件 Palais-Smale 条件结合起来, 则提供了一个求解更广意义下的极小点 (即临界点) 的方法.

定理 2.54 (Ekeland)**.** *设 (X,d) 是一个完备的度量空间. 又设 $f\colon X \to (-\infty,+\infty]$, 且 $f \not\equiv +\infty$. 若 f 是下方有界且下半连续的, 则对任意的 $\varepsilon > 0$ 和 $x_\varepsilon \in X$ 使得 $f(x_\varepsilon) < \inf_X f + \varepsilon$, 存在 $y_\varepsilon \in X$, 使得*

- $f(y_\varepsilon) \le f(x_\varepsilon)$;
- $d(y_\varepsilon, x_\varepsilon) \le 1$;
- $f(x) > f(y_\varepsilon) - \varepsilon d(y_\varepsilon, x), \forall x \in X \setminus \{y_\varepsilon\}$. *即对于给定的 y_ε, 函数 $x \mapsto f(x) + \varepsilon d(y_\varepsilon, x)$ 以 y_ε 为严格极小点.*

证明. 先选择 $u_0 = x_\varepsilon$. 假设 u_i 已经选定, 令

$$S_n = \{x \in X\colon f(x) \le f(u_n) - \varepsilon d(x,u_n)\}.$$

显然 S_n 是非空的. 今选取 $u_{i+1} \in S_n$ 满足

$$f(u_{i+1}) - \inf_{S_n} f \le \frac{1}{2}\left(f(u_i) - \inf_{S_n} f\right). \tag{46}$$

由此我们便得到了 X 中的一组序列 $\{u_i\}$. 此外, 由 $\{u_i\}$ 的构造方式可知,

$$\varepsilon d(x,u_i) \le f(u_i) - f(u_{i+1}), \qquad \forall i = 0,1,\cdots.$$

将上述不等式累加, 并利用三角不等式, 即得

$$\varepsilon d(u_i,u_j) \le f(u_i) - f(u_j), \qquad \forall j \ge i. \tag{47}$$

注意到 $\{f(u_i)\}$ 是递减的, 故上式表明 $\{u_i\}$ 是 Cauchy 列. 记其极限为 u^*, 以下验证 u^* 满足定理中列出的三条性质.

首先, 由于 $\{f(u_i)\}$ 是递减的, 从而有

$$f(u^*) \le f(u_i) \le f(u_0) = f(x_\varepsilon).$$

其次, 由(47)可得,

$$\varepsilon d(x_\varepsilon,u^*) \le f(x_\varepsilon) - f(u^*) \le f(x_\varepsilon) - \inf_X f < \varepsilon,$$

即 $d(x_\varepsilon,u^*) \le 1$. 最后利用反证法证明 u^* 满足第三条性质. 设存在 $x \ne u^*$ 使得

$$f(x) \le f(u^*) - \varepsilon d(u^*,x). \tag{48}$$

首先注意到(47)给出不等式

$$f(u^*) \le f(u_i) - \varepsilon d(u_i,u^*). \tag{49}$$

联立(48)和(49), 即有

$$f(x) \le f(u_i) - \varepsilon d(u^*,x), \qquad \forall i.$$

这表明 $x \in S_i, \forall i$. 在(46)不等式两侧取下极限, 并利用 f 的下半连续性, 我们有

$$f(u^*) \le_{i\to\infty} \inf_{S_n} f \le f(x)$$

这与(48)矛盾. □

推论 2.55. 设 (X,d) 是一个完备的度量空间. 又设 $f\colon X \to (-\infty,+\infty]$ 是下方有界和下半连续的, 且 $f \not\equiv +\infty$. 那么对任意的 $\varepsilon > 0$, 存在 $y_\varepsilon \in S$, 使得对任意的 $x \ne y$, 有 $f(x) > f(y_\varepsilon) - \varepsilon d(x,y_\varepsilon)$.

以下引入 Palais-Smale 条件. 首先回顾一些概念:

定义 2.56. 设 f 是 *Banach* 空间 X 上的实值函数, $x_0 \in U \subseteq X$, 其中 U 是一个开集. 称 f 在 x_0 处是 ***Fréchet* 可微**的, 如果存在 $\xi \in X^*$, 使得

$$|f(x) - f(x_0) - \langle \xi, x - x_0\rangle| = o(\|x - x_0\|) \qquad (x \to x_0).$$

称 ξ 为 f 在 x_0 处的 ***Fréchet* 导数**, 记为 $f'(x_0)$.

若 *Fréchet* 导数 $f'(x)$ 处处存在, 并且 $x \mapsto f'(x)$ 是连续的, 那么称 f 是连续可微的, 记作 $f \in C^1$.

容易看出, Gâteaux 导数和 Fréchet 导数分别是欧氏空间中方向导数和全微分在 Banach 空间中的推广. 此外, 若 f 有 Fréchet 导数 $f'(x_0)$, 那么它必有 Gâteaux 导数 $\mathrm{d}f(x_0,h)$, 且

$$\mathrm{d}f(x_0,h) = \langle f'(x_0),h\rangle, \qquad \forall h\in X.$$

同时有

$$\|f'(x_0)\| = \sup_{0\neq h\in X} \frac{|\mathrm{d}f(x_0,h)|}{\|h\|}$$

反之, 设 f 在 x_0 的一个邻域 U 内处处有 Gâteaux 导数 $\mathrm{d}f(x,h)$ $(x\in U)$, 并且有 $\xi=\xi(x)\in X^*$ 满足

$$\langle \xi(x),h\rangle = \mathrm{d}f(x,h), \qquad \forall x\in U, h\in X.$$

如果 $x\mapsto \xi(x)$ 是连续的, 那么 f 在 x_0 处有 Fréchet 导数 $f'(x_0)$.

以下定义是对经典极小点概念的推广:

定义 2.57. 称满足条件 $f'(x_0)=0$ 的点 x_0 为**临界点**, 并称相应的函数值 $f(x_0)$ 为**临界值**.

由此可以看出, 在变分问题中, 极小点便是临界点, 而一切临界点都是 E-L 方程的解.

定义 2.58. 设 X 是一个 *Banach* 空间, $f\in C^1(X)$. 若对任一满足条件

$$f(x_i)\to c, \qquad \|f'(x_i)\|\to 0 \tag{50}$$

的序列 $\{x_i\}\subseteq X$ 都有收敛的子列, 那么称 f 在 c 处满足 ***Palais-Smale* 条件**, 记作 $\mathrm{PS_c}$. 此外, 称满足(50)的序列为一个 ***Palais-Smale* 序列** *(*简称 ***PS* 序列***)*.

推论 2.59. 设 X 是一个 *Banach* 空间. 设 $f\in C^1(X)$, 并且是下方有界的. 记 $c=\inf_X f$. 若 f 满足 $\mathrm{PS_c}$, 那么 f 能达到极小值.

证明. 根据 Ekeland 变分原理, 我们可以选取 X 中的一组序列 $\{x_i\}$, 满足条件

$$\begin{cases} f(x) > f(x_i) - \dfrac{1}{i}\|x-x_i\| & (\forall x\neq x_i), \\ c\le f(x_i) < c+\dfrac{1}{i}. \end{cases} \qquad \forall i.$$

第一个不等式表明

$$\|f'(x_i)\|\le \frac{1}{i}\to 0 \qquad (i\to+\infty).$$

第二个不等式表明 $f(x_i)\to c$. 按 $\mathrm{PS_c}$, $\{x_i\}$ 有收敛子列 $\{x_{j_i}\}$. 记其极限为 x^*. 由 f 的连续性, 便有 $f(x^*)=c=\inf_X f$. □

最后我们用 Ekeland 变分原理来推导一个最基本的临界点定理, 即山路定理.

问题的提出: 在一个四面环山的盆地, 从山外地面上一点 p_1 出发想要进入盆地中的一点 p_0. 人们希望走的山路是这样一条连接 p_0 和 p_1 的道路, 其最高点不高于任何临近

道路的最高点. 这条山路上的最高点未必是极值点, 一个自然的问题是: 该最高点是否是临界点?

具体地, 设 Ω 是 $\mathbb{R}^n$ 中的有界区域. 给定 $p_0 \in \Omega, p_1 \in X \setminus \overline{\Omega}$. 设函数 $f \in C^1(X)$ 满足

$$\alpha = \inf_{\partial\Omega} f(x) > \max\{f(p_0), f(p_1)\}. \tag{51}$$

令

$$\Gamma = \{\gamma \in C[0,1]\colon \gamma(i) = p_i, i = 0,1\} \tag{52}$$

以及

$$c = \inf_{\gamma\in\Gamma} \sup_{t\in[0,1]} f\circ\gamma(t). \tag{53}$$

问 c 是否是 f 的临界值? 即是否存在 $x \in X$, 使得 $f'(x_0) = 0$ 以及 $f(x_0) = c$? 下述定理回答了这个问题:

定理 2.60 (山路定理)**.** *设 X 是一个 Banach 空间, $f \in C^1(X)$. 又设 $\Omega \subseteq X$ 是一个有界区域. 给定 $p_0 \in \Omega, p_1 \in X \setminus \overline{\Omega}$ 满足(51), 又按(52)和(53)定义 c. 若 f 满足 $\mathrm{PS_c}$, 那么 $c \geq \alpha$ 是 f 的一个临界值.*

证明. 在 Γ 上引入度量

$$d(\gamma_1,\gamma_2) = \max_{t\in[0,1]} \|\gamma_1(t) - \gamma_2(t)\|.$$

容易验证 (Γ, d) 是一个完备的度量空间. 令

$$I(\gamma) = \max_{t\in[0,1]} f\circ\gamma(t).$$

由假设条件可知, $I \geq \alpha$, 即 I 是下方有界的. 此外, 注意到

$$\begin{aligned}|I(\gamma_1) - I(\gamma_2)| &\leq \max_{t\in[0,1]} |f\circ\gamma_1(t) - f\circ\gamma_2(t)|\\ &\leq \max_{t\in[0,1]} \|\dot{f}(\theta\gamma_1(t) + (1-\theta)\gamma_2(t))\|\|\gamma_1(t) - \gamma_2(t)\|\\ &\leq Cd(\gamma_1,\gamma_2),\end{aligned}$$

故 I 是 Lipschitz 连续的, 自然也是下半连续的. 应用 Ekeland 变分原理于 I, 我们可以选取 Γ 中的一组序列 $\{\gamma_i\}$, 使得对任意的 i, 有

$$c \leq I(\gamma_i) < c + \frac{1}{i}, \tag{54}$$

$$I(\gamma) > I(\gamma_i) - \frac{1}{i} d(\gamma,\gamma_i) \qquad (\gamma \neq \gamma_i), \tag{55}$$

现令

$$M(\gamma) = \{t \in [0,1]\colon f\circ\gamma(t) = I(\gamma)\}.$$

容易验证, $M \subseteq (0,1)$ 是非空紧集. 记

$$\Gamma_0 = \{\gamma \in C[0,1]\colon \gamma(i) = 0, i = 0,1\}.$$

对任意的 $h \in C[0,1]$, 任取一列递减趋于零的序列 $\{\lambda_j\}$ 以及序列 $\{\xi_j\} \subseteq M(\gamma_i + \lambda_j h)$, 由(54)可得

$$\lambda_j^{-1}(f \circ (\gamma_i + \lambda_j h)(\xi_j) - f \circ \gamma_i(\xi_j)) \geq -\frac{1}{i}.$$

由于 $\{\xi_j\}$, 故存在收敛子列. 设收敛子列的极限为 η_i, 从而有

$$\mathrm{d}f(\gamma_i(\eta_i), h(\eta_i)) \geq -\frac{1}{i}. \tag{56}$$

若能证明存在 $\eta_i^* \in M(\gamma_i)$, 使得

$$\mathrm{d}f(\gamma(\eta_i^*), x) \geq -\frac{1}{i}, \qquad \forall x \in X, \|x\| = 1, \tag{57}$$

令 $x_i = \gamma_i(\eta_i^*)$, 从而有

$$c \leq f(x_i) = f \circ \gamma_i(\eta_i^*) < c + \frac{1}{i},$$
$$\|f'(x_i)\| = \sup_{\|x\|=1} |\mathrm{d}f(x_i, x)| \leq \frac{1}{i}.$$

注意到 f 满足 PS_c, 故 $\{x_i\}$ 有收敛子列. 若记子列的极限为 x^*, 显然有 $f'(x^*) = 0$, 且 $f(x^*) = c$. 即 c 是 f 的一个临界值.

以下证明(57). 若不存在 η_i^* 使得(57)成立, 则对任意的 $\eta \in M(\gamma_i)$, 存在 $y_\eta \in X, \|y_\eta\| = 1$, 使得

$$\mathrm{d}f(\gamma_i(\eta), y_\eta) < -\frac{1}{i}.$$

根据 f 的连续性, 存在 η 的一个邻域 $O_\eta \subseteq (0,1)$, 使得

$$\mathrm{d}f(\gamma_i(\xi), y_\eta) < -\frac{1}{i}, \qquad \forall \xi \in O_\eta.$$

注意到 $M(\gamma_i)$ 是紧的, 故开覆盖 $\{O_\eta\} \supseteq M(\gamma_i)$ 存在有限子覆盖 $\{O_{\eta_j}\}_{j=1}^m$, 其对应着 $v_{\eta_j} \in X$, 满足 $\|y_{\eta_i}\| = 1$, 且

$$\mathrm{d}f(\gamma_i(\xi), y_{\eta_i}) < -\frac{1}{i}, \qquad \forall \xi \in O_{\eta_j}, j = 1, \cdots, m.$$

现取 $\{O_{\eta_j}\}$ 对应的一组单位分解 $\{\rho_j\}$, 并令

$$y = y(\xi) = \sum_{j=1}^m \rho_j(\xi) y_{\eta_j},$$

从而有

$$\mathrm{d}f(\gamma_i(\xi), y(\xi)) < -\frac{1}{i}, \qquad \forall \xi \in M(\gamma_i),$$

这与(56)矛盾. □

例 2.61. 在山路定理中, 若 f 不满足 *Palais-Smale* 条件, 则定理结论可能不成立. 以下反例来自 *Brezis* 和 *Nirenberg*.

在 $\mathbb{R}^2$ 上考虑函数

$$f(x,y)=x^2+(1-x)^3y^2$$

令 $\Omega=B_{1/2}(0)$, $c=\inf_{\partial\Omega}f>0$. 显然 $f(0,0)=0, f(4,1)=-11$. 但直接验证可知, f 只有一个临界点 $(0,0)$.

最后, 我们利用山路定理来求解变分问题.

例 2.62. 设 a 是 $\mathbb{R}$ 上周期为 T 的连续函数. 定义函数

$$V(t,x)=-\frac{1}{2}|x|^2+\frac{a(t)}{p+1}|x|^{p+1}$$

假设 $p>1, a(t)\geq\alpha>0$. 求满足方程

$$\ddot{x}+\partial_x V(t,x)=0 \tag{58}$$

的非平凡 C^2 周期解.

在 $H^1_{\text{per}}(0,T)$ 上定义泛函

$$I(x)=\int_0^T\left(\frac{1}{2}(|\dot{x}|^2+|x|^2)-\frac{a(t)}{p+1}|x|^{p+1}\right)\mathrm{d}t.$$

而(58)是 I 对应的 *E-L* 方程. 为使用山路定理, 取 $p_0=0$, $p_1=\lambda\xi\sin\frac{2\pi}{T}t$, 其中 $\xi=(1,1,\cdots,1)$, $\lambda>0$. 事实上, 直接计算得 $I(p_0)=0$, 且

$$I(p_1)=O(\lambda^2)-O(\lambda^{p+1}),$$

从而我们可以选取充分大的 $\lambda>0$, 使得 $I(p_1)<0$. 定义

$$\Gamma=\{\gamma\in C[0,1]\colon \gamma(i)=p_i, i=0,1\}$$

以及

$$c=\inf_{\gamma\in\Gamma}\sup_{t\in[0,1]}I\circ\gamma(t).$$

以下验证 f 满足 $\mathrm{PS_c}$. 具体地, 设 $\{x_i\}$ 是一个 PS 序列, 即满足条件

$$\begin{cases}I(x_i)\to c,\\ \|I'(x_i)\|=\sup_{\|x\|=1}|\mathrm{d}I(x_i,x)|\to 0.\end{cases}$$

直接计算得

$$\begin{cases}I(x_i)=\int_0^T\left(\frac{1}{2}(|\dot{x}_i|^2+|x_i|^2)-a(t)\frac{|x_i|^{p+1}}{p+1}\right)\mathrm{d}t\to c,\\ \mathrm{d}I(x_i,x_i)=\int_0^T(|\dot{x}_i|^2+|x_i|^2-a(t)|x_i|^{p+1})\,\mathrm{d}t=o(\|x_i\|).\end{cases}$$

从而有

$$\left(\frac{1}{2}-\frac{1}{p+1}\right)\int_0^T (|\dot{x}_i|^2+|x_i|^2)\,\mathrm{d}t = C+o(\|x_i\|),$$

其中 C 是一个常数. 上式表明 $\{x_i\}$ 是有界序列, 从而存在子列, 仍记为 $\{x_i\}$, 使得 $x_i \rightharpoonup x^*$ in $H^1_{\mathrm{per}}(0,T)$. 特别地, 结合 PS 序列的假设, 我们有

$$\int_0^T a(t)|x_i|^{p-1}x_i\varphi\,\mathrm{d}t \to \int_0^T (\dot{x}^*\dot{\varphi}+x^*\varphi)\,\mathrm{d}t = \int_0^T \left(\frac{\mathrm{d}^2}{\mathrm{d}t^2}+1\right)x^*\varphi\,\mathrm{d}t.$$

此外, 直接利用分部积分公式可得

$$\int_0^T \left(\left(\frac{\mathrm{d}^2}{\mathrm{d}t^2}+1\right)x_i - a(t)|x_i|^{p-1}x_i\right)\varphi\,\mathrm{d}t \to 0,$$

综上分析, 我们有 $x_i \to x$ in $H^1_{\mathrm{per}}(0,T)$, 故 f 满足 $\mathrm{PS_c}$. 由山路定理, 方程(58)有非平凡周期解 u^*. 再由正则性定理, 此解还是 C^2 的.

2.8 专题四: 对偶最小作用量原理

Hamilton 方程组对应的泛函不是下方有界的, 故不能采用直接求极值的方法. 然而, 若 H 是凸的, 我们可以通过 Legendre 变换将问题转化. 这便是**对偶最小作用量原理**.

2.8.1 凸分析初步: Legendre 变换

首先回顾一些关于 Legendre 变换的基本概念. 在第五节中, 我们主要考虑的是 C^1 函数的情形. 在本节中, 我们集中于 $\mathbb{R}^n$ 上的 Legendre 变换, 而对函数的可微性不做进一步的要求. 具体地, 设 $f\colon \mathbb{R}^n \to (-\infty,+\infty]$, 且 $f \not\equiv +\infty$ (此时称 f 为**真**函数). 定义

$$\boxed{f^*(\xi) = \sup_{x\in\mathbb{R}^n}\left((\xi,x)_{\mathbb{R}^n} - f(x)\right)}.$$

称 f^* 为 f 的 **Legendre 变换.**

命题 2.63. *设 $f\colon \mathbb{R}^n \to (-\infty,+\infty]$ 是真函数. 记 f^* 为 f 的 Legendre 变换. 那么*

1. *f^* 是下半连续的凸函数;*

2. *若 f 是真的, 下半连续的凸函数, 则 f^* 是真函数;*

3. *若 $f \le g$, 则 $g^* \le f^*$;*

4. *(Young 不等式) $(x,\xi) \le f(x)+f^*(\xi)$, 其中等号成立当且仅当 $\xi \in \partial f(x)$. 这里 $\partial f(x)$ 为 f 的在 x 处的次微分:*

$$\partial f(x) = \{\xi \in \mathbb{R}^n\colon f(y) \ge f(x)+(\xi,y-x), \forall y \in D(f)\}.$$

证明. 1. 注意到 f^* 可以看作是一族凸的, 连续的 (从而也是下半连续的) 函数之上确界, 而凸且下半连续的函数族的上确界还是凸且下半连续的.

2. 由于 f 是真的, 故可以选取 $x_0 \in \mathbb{R}^n$ 使得 $f(x_0) < +\infty$. 由假设可知, 上图

$$\text{epi } f = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} \colon f(x) \le t\}$$

是非空凸闭集. 取 $t_0 < f(x_0)$, 则 $(x_0,t_0) \notin \text{epi } f$. 再利用凸集分离定理, 存在 $\Phi \in (\mathbb{R}^n \times \mathbb{R})^*$ 和 $\alpha \in \mathbb{R}$, 使得超平面 $\{\Phi = \alpha\}$ 严格分离 epi f 和 (x_0,t_0). 利用 $\mathbb{R}^n$ 的自共轭性, 我们可以找到 $(\xi,\lambda) \in \mathbb{R}^n \times \mathbb{R}$, 使得对任意的 (x,t), 有

$$\Phi((x,t)) = (\xi,x)_{\mathbb{R}^n} + \lambda t.$$

综上所述, 我们有

$$(\xi,x)_{\mathbb{R}^n} + \lambda t > \alpha > (\xi,x_0)_{\mathbb{R}^n} + \lambda t_0, \qquad \forall (x,t) \in \text{epi } f.$$

特别地,

$$(\xi,x_0)_{\mathbb{R}^n} + \lambda f(x_0) > \alpha > (\xi,x_0)_{\mathbb{R}^n} + \lambda t_0,$$

从而 $\lambda > 0$, 并且

$$\left(-\frac{1}{\lambda}\xi, x\right)_{\mathbb{R}^n} - f(x) < -\frac{\alpha}{\lambda},$$

故有

$$f^*\left(-\frac{\xi}{\lambda}\right) < -\frac{\alpha}{\lambda} < +\infty.$$

由此即表明 $f \not\equiv +\infty$.

3. 显然.

4. Young 不等式可以由 Legendre 变换的定义直接得出. 此外, 我们有

$$\begin{aligned}
\xi \in \partial f(x) &\iff (\xi, y-x)_{\mathbb{R}^n} \le f(y) - f(x), \quad \forall y \in \mathbb{R}^n, \\
&\iff (\xi,y)_{\mathbb{R}^n} - f(y) \le (\xi,x)_{\mathbb{R}^n} - f(x), \quad \forall y \in \mathbb{R}^n, \\
&\iff f^*(\xi) \le (\xi,x)_{\mathbb{R}^n} - f(x).
\end{aligned}$$

□

定理 2.64 (Fenchel-Moreau). [3] 若 f 是一个真的, 下半连续的凸函数, 则 $f^{**} = f$.

推论 2.65. 对于一个真的, 下半连续的凸函数 f, 我们有

$$\xi \in \partial f(x) \iff x \in \partial f^*(\xi).$$

证明. 利用 Fenchel-Moreau 定理和命题 2.63(4)即可. □

[3]该定理结论对一般的 Legendre 变换也成立 (利用凸集分离定理证明, 即 Hahn-Banach 定理的几何形式).

推论 2.66. 设 $f\colon \mathbb{R}^n \to (-\infty,+\infty]$ 是真函数, 则

$$f^{**} = \operatorname{conv} f = \sup\{\varphi\colon \varphi \textit{ proper and convex}, \varphi(x) \le f(x), \forall x \in \mathbb{R}^n\}.$$

证明. 一方面, 设 $g \le f$ 是凸的, 则它还是真的. 由命题 2.63(3)和 Fenchel-Moreau 定理, 即得

$$g = g^{**} \le f^{**}.$$

另一方面, 注意到 f^{**} 本身便是凸的, 由此我们便证得了所需结论. □

2.8.2 对偶最小作用量原理

Question: 给定一个凸的 Hamilton 函数 $H \in C^1(\mathbb{R}^N \times \mathbb{R}^N)$, 并满足如下的增长条件:

$$0 \le H(u,\xi) \le C(|u|^2 + |\xi|^2).$$

求满足下列 Hamilton 方程组的周期解 $(u(t),\xi(t))$:

$$\begin{cases} \dot{\xi}(t) = H_u(u(t),\xi(t)), \\ \dot{u}(t) = -H_\xi(u(t),\xi(t)). \end{cases} \tag{59}$$

Solution. 先求解下列方程的以 2π 为周期的解 $(v(t),\eta(t))$:

$$\begin{cases} \dot{\eta}(t) = \lambda H_v(v(t),\eta(t)), \\ \dot{v}(t) = -\lambda H_\eta(v(t),\eta(t)). \end{cases} \tag{60}$$

事实上, 若已求得满足方程(60)的解 $(v(t),\eta(t))$, 令

$$u(t) = v(\lambda^{-1}t), \quad \xi(t) = \eta(\lambda^{-1}t),$$

则 $(u(t),\xi(t))$ 满足(59), 它以 $2\lambda\pi$ 为周期 $(\lambda > 0)$.

现将(60)转化成一个约束极小值问题: 求泛函

$$I(w,\rho) = \int_0^{2\pi} H^*(\dot{\rho}(t), -\dot{w}(t))\,\mathrm{d}t$$

在约束

$$G(w,\rho) = \int_0^{2\pi} (-\dot{\rho}(t)\cdot w(t) + \dot{w}(t)\cdot \rho(t))\,\mathrm{d}t = -\pi$$

下的极小值. 这里 H^* 代表 H 的 Legendre 变换. 事实上, 若 (w_0,ρ_0) 是该约束极值问题的极小点, 则存在 $\lambda \in \mathbb{R}$, 使得

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}(H^*_\rho(\dot{\rho}_0(t), -\dot{w}_0(t)) - \frac{\lambda}{2}w_0(t)) = \frac{\lambda}{2}\dot{w}_0(t), \\ \frac{\mathrm{d}}{\mathrm{d}t}(-H^*_w(\dot{\rho}_0(t), -\dot{w}_0(t)) + \frac{\lambda}{2}\rho_0(t)) = -\frac{\lambda}{2}\dot{\rho}_0(t). \end{cases}$$

再利用2.65, 上式等价于

$$\begin{cases}\dot{\rho}_0(t)=H_w(\lambda w_0(t),\lambda\rho_0(t)),\\ \dot{w}_0(t)=-H_\rho(\lambda w_0(t),\lambda\rho_0(t)).\end{cases}$$

令 $\eta_0=\lambda\rho_0, v_0=\lambda w_0$, 容易验证 (η_0,v_0) 满足(60).

以下利用直接方法来验证极小点 (w_0,ρ_0) 的存在性. 显然, I 是连续的, 从而也是弱序列下半连续的. 其次, 由 H 的增长性条件可知

$$H^*(w,\rho)\geq\frac{1}{C}(|w|^2+|\rho|^2),$$

从而有

$$I(w,\rho)\geq\frac{1}{C}\int_0^{2\pi}(|\dot{w}|^2+|\dot{\rho}|^2)\,\mathrm{d}t.$$

这表明 I 是下方有界, 且是强制的. 此时定义域 $M=G^{-1}(-\pi)$ 上的范数规定为

$$\|(w,\rho)\|=\left(\int_0^{2\pi}(|\dot{w}|^2+|\dot{\rho}|^2)\,\mathrm{d}t\right)^{1/2}.$$

最后, 设 $\{(w_i,\rho_i)\}\subseteq M$ 在 $H^1_{\text{per}}(0,2\pi)$ 上弱收敛到 (w^*,ρ^*). 由紧嵌入 $H^1\hookrightarrow L^2$ 可知, w_j 和 ρ_j 在 $L^2(0,2\pi)$ 中分别收敛到 w^* 和 ρ^*. 从而有

$$\begin{aligned}-\pi=\lim_{j\to\infty}G(w_j,\rho_j)&=\lim_{j\to\infty}\int_0^{2\pi}(-\dot{\rho}_j\cdot w_j+\dot{w}_j\cdot\rho_j)\,\mathrm{d}t\\&=\int_0^{2\pi}(-\dot{\rho}^*\cdot w^*+\dot{w}^*\cdot\rho^*)\,\mathrm{d}t=G(w^*,\rho^*).\end{aligned}$$

这表明 $(w^*,\rho^*)\in M$, 故 M 是弱序列闭的. 综上所述, 我们利用直接方法验证了极小点 (w_0,ρ_0) 的存在性.

最后, 为使得到的极小点 (w_0,ρ_0) 是非平凡的, 我们还需说明 $\lambda>0$. 使用记号

$$\langle(w,\rho),(u,\xi)\rangle=\int_0^{2\pi}(w\cdot u+\rho\cdot\xi)\,\mathrm{d}t.$$

在前述推导中我们得到了等式

$$\nabla H^*(\dot{\rho}_0,-\dot{w}_0)=\lambda(w_0,\rho_0).$$

从而

$$\langle\nabla H^*(\dot{\rho}_0,-\dot{w}_0),(\dot{\rho}_0,-\dot{w}_0)\rangle=-\lambda G(w_0,\rho_0)=\lambda\pi.$$

利用 H 的增长条件, 易得 $\nabla H(0,0)=(0,0)$, 故

$$H^*(0,0)=-H(0,0)=0.$$

由于 H 是凸的, 从而有

$$H^*(0,0)-H^*(\dot{\rho}_0,-\dot{w}_0)\geq-\langle\nabla H^*(\dot{\rho}_0,-\dot{w}_0),(\dot{\rho}_0,-\dot{w}_0)\rangle,$$

即

$$\lambda\pi = \langle \nabla H^*(\dot{\rho}_0, -\dot{w}_0), (\dot{\rho}_0, -\dot{w}_0)\rangle \geq H^*(\dot{\rho}_0, -\dot{w}_0) \geq 0.$$

这表明 $\lambda \geq 0$. 以下验证 $\lambda \neq 0$. 若 $\lambda = 0$, 那么必有

$$\nabla H^*(\dot{\rho}_0, -\dot{w}_0) = (0,0).$$

从而

$$\begin{cases} \dot{\rho}_0 = H_w(0,0), \\ \dot{w}_0 = -H_\rho(0,0). \end{cases}$$

再利用 H 的增长条件, 即得 $(\dot{\rho}_0, \dot{w}_0) = (0,0)$, 而这与约束条件矛盾!