arXiv:2305.03897v2 [math.OC] 5 Sep 2024

希尔伯特空间中约束优化问题的精确增广拉格朗日函数 II：应用曾道人

M.V. Dolgopolik*

2024年9月9日

摘要

这项分为两部分的研究旨在分析所谓的精确增广拉格朗日函数。Di Pillo 和 Grippo 最初针对有限维优化问题引入了这一概念；本研究则将其用于希尔伯特空间中的优化问题。在本研究的第二部分，我们将一般的精确增广拉格朗日函数理论应用于若干约束变分问题和最优控制问题，包括带有额外边界约束的变分问题、等周问题、含有非完整等式约束（偏微分方程约束）的问题，以及线性演化方程的最优控制问题。我们为这些问题的增广拉格朗日函数给出全局精确／完全精确的充分条件；也就是说，给出相应条件，使得约束变分问题／最优控制问题等价于同时对原变量和对偶变量中的相应精确增广拉格朗日函数进行无约束极小化。

1 引言

2 预备知识

我们引入一些记号，并回顾我们研究第一部分 [14] 中的若干辅助定义，以及本文通篇将要用到的 Sobolev空间理论。

2.1 精确增广拉格朗日函数

设 X 和 H 为实希尔伯特空间，$\langle\cdot,\cdot\rangle$ 表示 X、H 或 $\mathbb{R}^n$中的内积，具体取决于上下文，$|\cdot|$ 表示欧几里得范数。另设函数f、$g_i\colon X\to\mathbb{R}$，其中 $i\in M:=\{1,\dots,m\}$，且 $F\colon X\to H$为给定函数。记 $g(\cdot)=(g_1(\cdot)\dots,g_m(\cdot))$。

*俄罗斯科学院机械工程问题研究所，俄罗斯圣彼得堡

约束优化问题的经典拉格朗日函数

$$\min \ \ f(x) \quad \text{subject to} \quad F(x)=0, \quad g_i(x)\le 0, \quad i\in M \tag{$\mathcal{P}$}$$

定义为

$$L(x,\lambda,\mu)=f(x)+\langle\lambda,F(x)\rangle+\langle\mu,g(x)\rangle, \quad \lambda\in H, \quad \mu\in\mathbb{R}^m.$$

为问题（$\mathcal{P}$）定义精确增广拉格朗日函数，选择一个凸的、非递减的下半连续（l.s.c.）函数 $\phi\colon [0,+\infty)\to[0,+\infty]$，使得 $\phi(t)=0$ 当且仅当 $t=0$ 且 $\operatorname{dom}\phi\neq\{0\}$。还需选择一个连续可微的凹函数 $\psi\colon [0,+\infty)^m\to\mathbb{R}$，使得零是 ψ 的全局最大值点，且 $\psi(0)>0$（关于这些函数的一些附加假设以及若干具体示例，见 [14]）。

对于任意向量 $y,z\in\mathbb{R}^m$，用 $\max\{y,z\}\in\mathbb{R}^m$ 表示向量 y 和 z 的逐坐标最大值。引入函数

$$b(x)=\psi\big(\max\{g(x),0\}\big), \quad p(x,\mu)=\frac{b(x)}{1+|\mu|^2} \quad \forall x\in X,\ \mu\in\mathbb{R}^m$$

并记 $\Omega=\{x\in X\mid b(x)>0,\ \phi(\|F(x)\|^2)<+\infty\}$。

设函数 f, F 和 $g_i, i\in M$ 连续 Fr´echet 可微，$DF(x)[\cdot]\colon X\to H$ 为 F 在 x 处的 Fr´echet 导数，$\nabla_x L(x,\lambda,\mu)$ 为函数 $x\mapsto L(x,\lambda,\mu)$ 的梯度。问题（$\mathcal{P}$）的精确增广拉格朗日函数定义如下：

$$\begin{aligned}\mathscr{L}(x,\lambda,\mu,c)&=f(x)+\langle\lambda,F(x)\rangle+\frac{c}{2}\big(1+\|\lambda\|^2\big)\phi(\|F(x)\|^2)\\&\quad+\Big\langle\mu,\max\Big\{g(x),-\frac{1}{c}p(x,\mu)\mu\Big\}\Big\rangle+\frac{c}{2p(x,\mu)}\Big|\max\Big\{g(x),-\frac{1}{c}p(x,\mu)\mu\Big\}\Big|^2\\&\qquad+\eta(x,\lambda,\mu),\end{aligned} \tag{1}$$

如果 $x\in\Omega$，且 $\mathscr{L}(x,\lambda,\mu,c)=+\infty$，否则。这里 $\lambda\in H$ 和 $\mu\in\mathbb{R}^m$ 是拉格朗日乘子，$c>0$ 是罚参数，并且

$$\begin{aligned}\eta(x,\lambda,\mu)=&\frac{1}{2}\Big\|DF(x)\big[\nabla_x L(x,\lambda,\mu)\big]\Big\|^2\\&+\frac{1}{2}\sum_{i=1}^m\Big(\langle\nabla g_i(x),\nabla_x L(x,\lambda,\mu)\rangle+g_i(x)^2\mu_i\Big)^2.\end{aligned} \tag{2}$$

为简便起见，下文我们仅考虑 $\phi(t)\equiv t$ 和 $\psi(y)\equiv 1$ 的情形，尽管对于本文所讨论的问题，基于精确增广拉格朗日函数的优化方法在选择其他函数时可能表现得更好。

对于任意 $\gamma\in\mathbb{R}$ 和 $c>0$，引入次水平集。

$$\begin{aligned}S_c(\gamma)&=\Big\{(x,\lambda,\mu)\in\Omega\times H\times\mathbb{R}^m\ \Big|\ \mathscr{L}(x,\lambda,\mu,c)\le\gamma\Big\},\\\Omega_c(\gamma)&=\Big\{x\in\Omega\ \Big|\ f(x)+c\big(\|F(x)\|^2+|\max\{g(x),0\}^2|\big)\le\gamma\Big\}.\end{aligned} \tag{3}$$

我们来回顾一下增广拉格朗日函数全局精确性的定义。假设问题（$\mathcal{P}$）存在全局最优解。回顾

如果三元组（x， λ， μ） $\in X \times H \times \mathbb{R}^m$ 被称为问题（$\mathcal{P}$）的 KKT 点，那么 x 对此问题是可行的， $\nabla_x L$（x， λ， μ） $= 0$，并且对所有 $i \in M$，都有 $\mu_i g_i$（x） $= 0$ 且 $\mu_i \geq 0$。假设对于问题（$\mathcal{P}$）的任意全局最优解 x_*，都存在拉格朗日乘子 $\lambda_* \in H$ 和 $\mu_* \in \mathbb{R}^m$，使得三元组（x_*， λ_*， μ_*）是问题（$\mathcal{P}$）的 KKT 点。

定义 1. 若存在 $c_* > 0$，使得对于所有 $c \geq c_*$，三元组 (x_*, λ_*, μ_*) 都是该问题的全局最优解，则称增广拉格朗日函数 $\mathscr{L}(x, \lambda, \mu, c)$ 全局精确：

$$\min_{(x,\lambda,\mu)\in X\times H\times\mathbb{R}^m} \mathscr{L}(x, \lambda, \mu, c) \tag{4}$$

当且仅当 x_* 是问题 $(\mathcal{P})$ 的全局最优解，且 (x_*, λ_*, μ_*) 是该问题的 KKT点。

注 1. 值得注意的是，在某些自然假设下，增广拉格朗日函数 $\mathscr{L}(x, \lambda, \mu, c)$ 全局精确，当且仅当存在 c_*，使得对于任意 $c \geq c_*$，问题 (4) 的最优值等于问题 $(\mathcal{P})$ 的最优值（见 [14, 引理 5.2]）。

增广拉格朗日函数$\mathscr{L}(x, \lambda, \mu, c)$全局精确的充分条件可用函数来表示

$$\begin{aligned} Q(x)[\lambda, \mu] =& \frac{1}{2}\Big\| DF(x)\Big[DF(x)^*[\lambda] + \sum_{i=1}^{m} \mu_i \nabla g_i(x)\Big]\Big\|^2 \\ &+ \frac{1}{2}\Big|\nabla g(x)\Big(DF(x)^*[\lambda] + \sum_{i=1}^{m} \mu_i \nabla g_i(x)\Big) + \operatorname{diag}(g_i(x)^2)\mu\Big|^2 \end{aligned} \tag{5}$$

其中， $DF(x)^*[\cdot]$ 是线性算子 $DF(x)[\cdot]$ 的伴随算子，而$\nabla g(x)y \in \mathbb{R}^m$ 是一个向量，其第 i 个坐标对于任意 $y \in X$均为 $\langle \nabla g_i(x)\ y\rangle$ 。注意，函数 $Q(x)[\cdot]$ 关于(λ, μ)是二次的。

定义 2. 如果存在 $\sigma > 0$ 使得，则称函数 $Q(x)[\cdot]$ 在点 $x \in X$ 处是正定的。

$$Q(x)[\lambda, \mu] \geq \sigma\big(\|\lambda\|^2 + |\mu|^2\big) \quad \forall \lambda \in H,\ \mu \in \mathbb{R}^m$$

所有使上述不等式成立的 $\sigma \geq 0$ 的上确界记为 $\sigma_{\max}(Q(x))$。

在第一部分研究 [14] 中，我们得到了增广拉格朗日函数 $\mathscr{L}(x, \lambda, \mu, c)$ 的全局精确性以及所谓完全精确性这一更强性质的充分条件。这些条件在很大程度上依赖于函数 $\mathscr{L}(\cdot, c)$ 的弱序列下半连续性，以及对于适当选取的 γ 值，次水平集 $S_c(\gamma)$ 的有界性。这些假设是必要的，以确保对于任意充分大的罚参数 $c > 0$，问题 (4) 都存在全局最优解。然而，正如下文将要看到的，在某些应用中，这些假设极难验证（如果可能的话），因此使得 [14] 中的结果无法适用。为了克服这一困难，下文将给出增广拉格朗日函数 $\mathscr{L}(x, \lambda, \mu, c)$ 的全局精确性和完全精确性的充分条件，而这些条件并不依赖于此类假设。

不要求下半连续性或有界性假设。这些条件受到 Demyanov [6–11] 关于无限维空间中优化问题的精确罚函数的研究结果启发；在该研究中，明确提出了罚函数存在全局极小点这一假设。

定理 1. 设 $f_* > -\infty$ 为问题 $(\mathcal{P})$ 的最优值，并且以下假设成立：

1. $f, g_i, i \in M$ 和 F 在 Ω 上二阶连续可微 Fr´，ϕ 在其有效定义域上连续可微，且 $\phi'(0) > 0$；2. 存在 $r > 0, \varepsilon > 0$ 和 $\gamma > f_*$，使得函数 $f, g_i, i \in M$ 和 F，以及它们的一阶和二阶 Fr´ 导数，在 $\Omega_r(\gamma + \varepsilon)$ 上有界；3. 存在 $\sigma > 0$，使得对所有 $x \in \Omega_r(\gamma + \varepsilon)$，$\sigma_{\max}(Q(x)) \geq \sigma$；4. 存在一个无界递增序列 $\{c_n\}$，使得对于任意 $n \in \mathbb{N}$，函数 $\mathscr{L}(\cdot, c_n)$ 在 $X \times H \times \mathbb{R}^m$ 上取得全局最小值。

那么存在 $c_* > 0$，使得对于所有 $c \geq c_*$，增广拉格朗日函数$\mathscr{L}(x, \lambda, \mu, c)$ 在集合 $S_c(\gamma)$ 上是完全精确的，具体含义如下：

1. 问题 $(\mathcal{P})$ 与问题 (4) 的最优值相同；2. (x_*, λ_*, μ_*) 是 $\mathscr{L}(x, \lambda, \mu, c)$ 的全局最小值点，当且仅当 x_* 是问题 $(\mathcal{P})$ 的全局最优解，且 (x_*, λ_*, μ_*) 是该问题的 KKT 点；3. $(x_*, \lambda_*, \mu_*) \in S_c(\gamma)$ 是 $\mathscr{L}(x, \lambda, \mu, c)$ 的驻点，当且仅当 (x_*, λ_*, μ_*) 是问题 $(\mathcal{P})$ 的 KKT 点，且 $f(x_*) \leq \gamma$；4. 若 $(x_*, \lambda_*, \mu_*) \in S_c(\gamma)$ 是 $\mathscr{L}(x, \lambda, \mu, c)$ 的局部极小值点，则 x_*是问题 $(\mathcal{P})$ 的局部最优解，$f(x_*) \leq \gamma$，且 (x_*, λ_*, μ_*) 是该问题的 KKT 点。

该定理的证明见附录。

注2. 本文始终隐含假设，所讨论的所有优化问题都有有限的最优值和全局最优解（涉及增广拉格朗日函数的问题除外）。此外，我们还假设，对于这些问题的任意全局最优解，都存在一个拉格朗日乘子使得 KKT最优性条件成立。

2.2 L^p 与 Sobolev空间

设 $\mathscr{U}$ 为希尔伯特空间，$E \subset \mathbb{R}^n$ 为开集。对于任意 $1 \leq p \leq \infty$，我们用 $L^p(E; \mathscr{U})$ 表示由所有满足以下条件的可测函数 $u\colon E \to \mathscr{U}$ 所组成的空间

$$\|u\|_{L^p(E;\mathscr{U})} := \left(\int_E \|u(x)\|_{\mathscr{U}}^p \, dx \right)^{\frac{1}{p}} < +\infty, \quad \text{if } 1 \leq p < \infty,$$
$$\|u\|_{L^\infty(E;\mathscr{U})} := \operatorname{esssup}_{x \in E} \|u(x)\|_{\mathscr{U}} < +\infty,$$

赋予范数$\|\cdot\|_{L^p(E;\mathscr{U})}$。如果 $\mathscr{U}$ 是配备欧几里得范数 $|\cdot|$的空间 $\mathbb{R}^d$，那么我们将 $L^p(E;\mathbb{R}^d)$ 上的相应范数记为$\|\cdot\|_p$。

令 $W^{s,p}(E)$ 为 Sobole夫空间，而 $W^{s,p}(E;\mathbb{R}^d)$ 是由所有满足以下条件的向量值函数 $u\colon E\to\mathbb{R}^d$, $u=(u_1,\dots,u_d)$ 构成的空间：对所有 $i\in\{1,\dots,d\}$, 均有 $u_i\in W^{s,p}(E)$。空间 $W^{1,2}(E;\mathbb{R}^d)$ 配备内积

$$\langle u,v\rangle=\int_E\big(\langle u(x),v(x)\rangle+\langle\nabla u(x),\nabla v(x)\rangle\big)\,dx\quad\forall u,v\in W^{1,2}(E;\mathbb{R}^d).$$

相应的范数。空间 $W^{1,2}(E;\mathbb{R}^d)$ 中所有具有紧支集的无穷次连续可微函数 $u\colon E\to\mathbb{R}^d$ 所构成的空间 $C_0^\infty(E;\mathbb{R}^d)$ 的闭包，记为 $W^{1,2}_0(E;\mathbb{R}^d)$。

众所周知（例如参见 [17]），当 $E=(a,b)\subset\mathbb{R}$ 时，空间 $W^{1,2}((a,b);\mathbb{R}^d)$ 与所有满足 $u'\in L^2((a,b);\mathbb{R}^d)$ 的绝对连续函数 $u\colon[a,b]\to\mathbb{R}^d$ 所构成的空间同构，我们将其记为 $W^{1,2}([a,b];\mathbb{R}^d)$。

我们还将使用文献 [12] 中引入的一种 Sobole夫型函数空间。下面回顾该空间的定义。任意由 $\in\mathbb{Z}_+^n$ 个非负整数构成的 n 元组 $\alpha=(\alpha_1,\dots,\alpha_n)$ α_i 称为一个多重指标。其绝对值定义为 $|\alpha|=\alpha_1+\ldots+\alpha_n$。对于任意多重指标 α，记 $D^\alpha=D_1^{\alpha_1}\dots D_n^{\alpha_n}$ 为一个 $|\alpha|$ 阶微分算子，其中 $D_i=\partial/\partial x_i$。若 $\alpha=0$，则对于任意函数 u，有 $D^\alpha u=u$。对于任意 $k\in\{0,\dots,n\}$，定义

$$I_k=\Big\{\alpha\in\mathbb{Z}_+^n\;\Big|\;|\alpha|=k,\;\alpha_i=0\text{ or }\alpha_i=1\;\forall i\in\{1,\dots,n\}\Big\}.$$

最后，记 $MW^{n,2}(E)$ 为所有函数 $u\in L^2(E)$ 的集合，其中对于任意 $k\in\{1,\dots,n\}$ 和 $\alpha\in I_k$，函数 u 都存在属于 $L^2(E)$ 的弱导数 $D^\alpha u$。因此，$MW^{n,2}$ 由所有满足以下条件的函数 $u\in L^2(E)$ 构成：其所有 $k\in\{1,\dots,n\}$ 阶弱混合导数均存在，且属于 $L^2(E)$。注意，$W^{n,2}(E)\subseteq MW^{n,2}(E)$，并且当 $n=1$ 时，该包含关系成为等式。

线性空间 $MW^{n,\,2}(E)$ 配备以下内积

$$\langle u,v\rangle=\sum_{k=0}^n\sum_{\alpha\in I_k}\int_E D^\alpha u(x)D^\alpha v(x)\,dx\quad\forall u,v\in MW^{n,2}(E)$$

及其对应的范数 $\|\cdot;\;MW^{n,\,2}\|$是一个可分希尔伯特空间。$C_0^\infty(E)$ 在 $MW^{n,\,2}(E)$ 中的闭包记为 $MW^{n,\,2}_0(E)$。根据 [12,定理 2]，与内积相对应的半范数定义在 $MW^{n,\,2}_0(E)$ 上。

$$\langle u,v\rangle=\int_E D^{(1,\dots,1)}u(x)D^{(1,\dots,1)}v(x)\,dx\quad\forall u,v\in MW_0^{n,2}(E)$$

是一个与范数 $\|\cdot$等价的范数；$MW^{n,\,2}\|$。因此，下面我们假设空间 $MW^{n,\,2}_0(E)$ 配备了这一内积及其对应的范数。

在E为有界开盒的情况下，空间 $MW^{n,2}_0(E)$中的函数具有一种非常简单且方便的刻画。此时，对于任意 $v\in L^2(E)$，记

$$(\mathcal{A}v)(x)=\int_{a_1}^{x_1}\dots\int_{a_n}^{x_n}v(s)\,ds\quad\text{for a.e. }x\in E.\tag{6}$$

根据富比尼定理，函数 $\mathcal{A}v$ 定义良好，且 $\mathcal{A}$ 是一个连续线性算子，将 E映射到 E。以下结果成立（参见[12, 定理 3]）。

定理2。设 $E = \prod_{i=1}^{n}(a_i, b_i)$。那么，函数 $u\colon E \to \mathbb{R}$ 属于$MW^{n,\,20}(E)$ 当且仅当存在函数 $v \in L^2(E)$，使得：

1. $u(x) = (\mathcal{A}v)(x)$ *for a.e.* $x \in E$;
2. $\int_{a_i}^{b_i} v(x_1, \ldots, x_{i-1}, s_i, x_{i+1}, \ldots, x_n)\, ds_i = 0$ *for a.e.* $x \in E$ *and for all* $i \in \{1, \ldots, n\}$.

此外，这样的函数 v 唯一确定，并且在弱意义下 $v = D^{(1,\ldots,1)}u$ 。

3 变分法的应用

本节讨论精确增广拉格朗日函数理论在变分法约束问题中的应用。具体而言，我们为带有边界附加约束的变分问题、带有等周约束的问题，以及带有非完整等式约束的变分问题，给出新的精确增广拉格朗日函数。我们还给出关于问题数据的充分条件，以确保这些增广拉格朗日函数具有全局精确性／完全精确性。为简便起见，我们分别考察每种约束；不过，也可以为涉及上述所有类型约束的变分问题定义一个精确增广拉格朗日函数，并利用下文给出的结果推导出该增广拉格朗日函数具有全局精确性／完全精确性的充分条件。

3.1 边界处带有附加约束的问题

考虑以下在定义域 $(a, b) \subset \mathbb{R}$ 的边界处带有附加非线性等式和不等式约束的变分问题：

$$\begin{aligned} &\min_{u \in W^{1,2}([a,b];\mathbb{R}^d)} \ \mathcal{I}(u) = \int_a^b f(u(x), u'(x), x)\, dx + f_0(u(a), u(b)) \\ &\text{subject to } \ f_j(u(a), u(b)) = 0, \ \ j \in J, \ \ g_i(u(a), u(b)) \le 0, \ \ i \in M. \end{aligned} \tag{7}$$

这里，$f\colon \mathbb{R}^d \times \mathbb{R}^d \times [a, b] \to \mathbb{R}$, $f = f(u, \xi, x)$，$f_j\colon \mathbb{R}^{2d} \to \mathbb{R}$，且 $g_i\colon \mathbb{R}^{2d} \to \mathbb{R}$是给定函数，而 $J = \{1, \ldots, \ell\}$和 $M = \{1, \ldots, m\}$是有限指标集，其中任意一个都可以为空。我们假设函数 f_j 和 g_i连续可微，而 f 是 Carath´eodoary 函数（即，对于所有 $u, \xi \in \mathbb{R}^d$，函数 $f(u, \xi, \cdot)$均可测，且对于几乎处处的 $x \in [a, b]$，函数 $f(\cdot, x)$均连续），它关于 u 和 ξ 可微，并且梯度$\nabla_u f$ 和 $\nabla_\xi f$ 也是 Carath´eodory 函数。

我们还对 f 及其偏导数施加以下增长条件。也就是说，对于任意 $R > 0$，存在 $C_R > 0$ 和几乎处处非负的函数$\eta_R \in L^1(a, b)$以及 $\theta_R \in L^2(a, b)$，使得

$$\begin{aligned} &|f(u, \xi, x)| \le C_R|\xi|^2 + \eta_R(x), \quad |\nabla_u f(u, \xi, x)| \le C_R|\xi|^2 + \eta_R(x), \\ &|\nabla_\xi f(u, \xi, x)| \le C_R|\xi| + \theta_R(x). \end{aligned} \tag{8}$$

对于几乎处处的 $x \in [a,b]$ 以及任意满足 $|u| \le R$的 $u, \xi \in \mathbb{R}^d$。这些增长条件保证，对于任意 $u \in W^{1,2}([a,b];\mathbb{R}^d)$，$\mathcal{I}(u)$的值都有明确定义且有限，泛函 $\mathcal{I}$ 是盖托可微的，并且其盖托导数$\mathcal{I}'(u)[\cdot]$ 具有如下形式

$$\mathcal{I}'(u)[w] = \int_a^b \big(\langle \nabla_u f(u(x),u'(x),x), w(x)\rangle + \langle \nabla_\xi f(u(x),u'(x),x), w'(x)\rangle\big)\,dx + \langle \nabla f_0(u(a),u(b)), (w(a),w(b))\rangle \quad (9)$$

对于所有 $w \in W^{1,2}([a,b];\mathbb{R}^d)$（参见例如 [5, 第 4.3 节]）。此外，很容易看出，

$$\begin{aligned}\big|\mathcal{I}'(u)[w] - \mathcal{I}'(v)[w]\big| \le{}& \big\|\nabla_u f(u(\cdot),u'(\cdot),\cdot) - \nabla_u f(v(\cdot),v'(\cdot),\cdot)\big\|_1 \|w\|_\infty \\ &+ \big\|\nabla_\xi f(u(\cdot),u'(\cdot),\cdot) - \nabla_\xi f(v(\cdot),v'(\cdot),\cdot)\big\|_2 \|w'\|_2 \\ &+ \big|\nabla f_0(u(a),u(b)) - \nabla f_0(v(a),v(b))\big| \|w\|_\infty\end{aligned}$$

对于任意的 u、v、$w \in W^{1,2}([a,b];\mathbb{R}^d)$。因此，考虑到 $\|\cdot\|_\infty \le C\|\cdot\|_{1,2}$ 对于某个 $C>0$ 成立（这一结果可由 Sobolev 嵌入定理[1, 定理 5.4] 推出），可得

$$\begin{aligned}\|\mathcal{I}'(u) - \mathcal{I}'(v)\| \le{}& C\big\|\nabla_u f(u(\cdot),u'(\cdot),\cdot) - \nabla_u f(v(\cdot),v'(\cdot),\cdot)\big\|_1 \\ &+ \big\|\nabla_\xi f(u(\cdot),u'(\cdot),\cdot) - \nabla_\xi f(v(\cdot),v'(\cdot),\cdot)\big\|_2 \\ &+ C\big|\nabla f_0(u(a),u(b)) - \nabla f_0(v(a),v(b))\big|.\end{aligned}$$

利用这个不等式以及关于 Nemytskii 算子连续性的标准结果（参见例如 [2]），可以得出 $\mathcal{I}$ 的 Gâteaux 导数连续，因此该泛函是连续 Fr´echet 可微的。

为了为问题 (7) 定义精确增广拉格朗日函数，我们需要计算目标函数（泛函 $\mathcal{I}$）以及该问题在空间 $W^{1,2}([a,b],\mathbb{R}^d)$ 中的所有约束的梯度。为简化这一问题，我们将采用一种称为转入导数空间的技巧（变量变换），Demyanov 在其关于变分法的著作中广泛使用了这一技巧 [6–11]。

对于任意 $y \in \mathbb{R}^d$和 $v \in L^2((a,b);\mathbb{R}^d)$，记 $\mathcal{A}(y,v)(x) = y + \int_a^x v(s)\,ds$ 对所有 $x \in [a,b]$成立。线性算子 $\mathcal{A}$连续地将 $\mathbb{R}^d \times L^2((a,b);\mathbb{R}^d)$映射到 $W^{1,2}([a,b];\mathbb{R}^d)$。此外，勒贝格微分定理表明，$\mathcal{A}$ 是这两个空间之间的一一对应，其逆算子为 $\mathcal{A}^{-1}(u) = (u(a),\ u'(\cdot))$（例如参见 [17]）。因此，通过应用变量代换 $u = \mathcal{A}(y,v)$，我们可以将问题 (7) 转化为如下形式的等价变分问题：

$$\begin{aligned}&\min_{(y,v)\in\mathbb{R}^d\times L^2((a,b);\mathbb{R}^d)} \ \widehat{\mathcal{I}}(y,v) := \mathcal{I}(\mathcal{A}(y,v)) \\ &\text{subject to } \ \widehat{f}_j(y,v) = 0,\ \ j \in J,\ \ \widehat{g}_i(y,v) \le 0,\ \ i \in M.\end{aligned} \quad (10)$$

其中 $\widehat{f}_j(y,v) := f_j(y, y + \int_a^b v(s)\,ds)$ 且 $\widehat{g}_i(y,v) := g_i(y, y + \int_a^b v(s)\,ds)$。我们将在“导数空间”中，为问题 (10) 定义一个精确增广拉格朗日函数。

将 $(a,b);\mathbb{R}^d)$ 记作 $X=\mathbb{R}^d\times L^2$，并赋予空间 X 内积

$$\langle (y,v),(z,w)\rangle=\langle y,z\rangle+\int_a^b\langle v(x),w(x)\rangle\,dx\quad\forall (y,v),(z,w)\in X$$

以及相应的范数。固定某个 $(y,v)\in X$，为方便起见记作 $u=\mathcal{A}(y,v)$。由等式 (9) 可知，泛函 $\widehat{\mathcal{I}}$ 是连续 Fr´echet 可微的，其 Fr´echet 导数 $D\mathcal{I}(y,v)[\cdot]$ 的形式为

$$\begin{aligned}&D\widehat{\mathcal{I}}(y,v)[z,w]\\&=\int_a^b\Big(\Big\langle\nabla_u f(u(x),v(x),x),z+\int_a^x w(s)\,ds\Big\rangle+\langle\nabla_\xi f(u(x),v(x),x),w(x)\rangle\Big)\,dx\\&\qquad+\langle\nabla_{u_a}f_0(y,u(b))+\nabla_{u_b}f_0(y,u(b)),z\rangle+\Big\langle\nabla_{u_b}f_0(y,u(b)),\int_a^b v(x)\,dx\Big\rangle\end{aligned}$$

对于任意的 $(z,w)\in X$，其中 $f_0=f_0(u_a,u_b)$。分部积分并整理各项后，得到

$$\begin{aligned}&D\widehat{\mathcal{I}}(y,v)[z,w]\\&\quad=\Big\langle\int_a^b\nabla_u f(u(x),v(x),x)\,dx+\nabla_{u_a}f_0(y,u(b))+\nabla_{u_b}f_0(y,u(b)),z\Big\rangle\\&\qquad\qquad+\int_a^b\langle P[y,v](x),w(x)\rangle\,dx\end{aligned}$$

对于所有的 $(z,w)\in X$，其中

$$P[y,v](x)=\int_x^b\nabla_u f(u(s),v(s),s)\,ds+\nabla_\xi f(u(x),v(x),x)+\nabla_{u_b}f_0(y,u(b))$$

几乎处处对 $x\in(a,b)$ 成立。利用增长条件 (8)，可以很容易地验证 $P[y,v]\in L^2((a,b);\mathbb{R}^d)$。因此，可以得出结论，泛函 $\mathcal{I}$ 的梯度 $\nabla\mathcal{I}(y,v)\in X$ 具有如下形式。

$$\begin{aligned}&\nabla\widehat{\mathcal{I}}(y,v)\\&=\Big(\int_a^b\nabla_u f(u(x),v(x),x)\,dx+\nabla_{u_a}f_0(y,u(b))+\nabla_{u_b}f_0(y,u(b)),P[y,v]\Big)\end{aligned}\tag{11}$$

对于所有 $((y,v))\in\mathbb{R}^d\times L^2((a,b);\mathbb{R}^d)$。采用类似的方法，可以很容易地验证：

$$\begin{aligned}\nabla\widehat{f_j}(y,v)&=\Big(\nabla_{u_a}f_j(y,u(b))+\nabla_{u_b}f_j(y,u(b)),\nabla_{u_b}f_j(y,u(b))\Big)\in X,\\\nabla\widehat{g_i}(y,v)&=\Big(\nabla_{u_a}g_i(y,u(b))+\nabla_{u_b}g_i(y,u(b)),\nabla_{u_b}g_i(y,u(b))\Big)\in X\end{aligned}\tag{12}$$

对于所有的 $j\in J$ 和 $i\in M$，其中 $f_j=f_j(u_a,u_b)$ 且 $g_i=g_i(u_a,u_b)$。

备注3. 应当指出，借助 Demyanov“转入导数空间”技术，我们能够以直接明了的方式，获得目标泛函和约束的梯度的简洁且显式的解析表达式。若不采用这一技术，计算这些函数的梯度将是一项颇具挑战性的任务。

将 $F=(f_1,\dots,f_\ell)$ 和 $G=(g_1,\dots,g_m)$ 记为。问题 (10) 的经典拉格朗日函数形式为

$$L(y,v,\lambda,\mu)=\widehat{\mathcal{I}}(y,v)+\langle\lambda,\widehat{F}(y,v)\rangle+\langle\mu,\widehat{G}(y,v)\rangle\quad\forall\lambda\in\mathbb{R}^\ell,\ \mu\in\mathbb{R}^m.$$

其在（y,v）中的梯度 $\nabla_{(u,v)}L(y,v,\lambda,\mu)$ 可利用（11）和（12）轻松计算。根据等式（1）和（2），定义问题（10）的精确增广拉格朗日函数 $\mathscr{L}(y,v,\lambda,\mu,c)$，其中 $\lambda\in\mathbb{R}^\ell$和$\mu\in\mathbb{R}^m$。为简洁和方便起见，我们将直接用原问题（7）中的项来表示:

$$\begin{aligned}\mathscr{L}(u,\lambda,\mu,c)=&\int_a^b f(u(x),u'(x),x)\,dx+f_0(u(a),u(b))\\&+\langle\lambda,F(u(a),u(b))\rangle+\frac{c}{2}(1+|\lambda|^2)|F(u(a),u(b))|^2\\&+\Big\langle\mu,\max\Big\{G(u(a),u(b)),-\frac{1}{c(1+|\mu|^2)}\mu\Big\}\Big\rangle\\&+\frac{c}{2}(1+|\mu|^2)\Big|\max\Big\{G(u(a),u(b)),-\frac{1}{c(1+|\mu|^2)}\mu\Big\}\Big|^2+\eta(u,\lambda,\mu),\end{aligned}$$

其中 $F=(f_1,\dots,f_\ell)$、$G=(g_1,\dots,g_m)$，以及

$$\begin{aligned}\eta(u,\lambda,\mu)=&\frac{1}{2}\sum_{j=1}^{\ell}\Big(\langle\nabla_{u_a}f_j(u(a),u(b))+\nabla_{u_b}f_j(u(a),u(b)),\nabla_yL(u,\lambda,\mu)\rangle\\&+\int_a^b\langle\nabla_{u_b}f_j(u(a),u(b)),\nabla_vL(u,\lambda,\mu)(x)\rangle\,dx\Big)^2\\&+\frac{1}{2}\sum_{i=1}^{m}\Big(\langle\nabla_{u_a}g_i(u(a),u(b))+\nabla_{u_b}g_i(u(a),u(b)),\nabla_yL(u,\lambda,\mu)\rangle\\&+\int_a^b\langle\nabla_{u_b}g_i(u(a),u(b)),\nabla_vL(u,\lambda,\mu)(x)\rangle\,dx+g_i(u(a),u(b)^2\mu_i\Big)^2,\end{aligned}$$

and

$$\begin{aligned}\nabla_yL(u,\lambda,\mu)=&\int_a^b\nabla_uf(u(x),u'(x),x)\,dx+\nabla_{u_a}f_0(u(a),u(b))+\nabla_{u_b}f_0(u(a),u(b))\\&+\sum_{j=1}^{\ell}\lambda_j\big(\nabla_{u_a}f_j(u(a),u(b))+\nabla_{u_b}f_j(u(a),u(b))\big)\\&+\sum_{i=1}^{m}\mu_i\big(\nabla_{u_a}g_i(u(a),u(b))+\nabla_{u_b}g_i(u(a),u(b))\big),\end{aligned}$$

and

$$\nabla_vL(u,\lambda,\mu)(x)=P[u](x)+\sum_{j=1}^{\ell}\lambda_j\nabla_{u_b}f_j(u(a),u(b))+\sum_{i=1}^{m}\mu_i\nabla_{u_b}g_i(u(a),u(b)),$$

and

$$P[u](x) = \int_x^b \nabla_u f_0(u(s), u'(s), s)\, ds + \nabla_\xi f_0(u(x), u'(x), x) + \nabla_{u_b} f_0(u(a), u(b))$$

对于几乎处处的 $x \in (a,b)$。

下面给出增广拉格朗日函数 $\mathscr{L}(y, v, \lambda, \mu, c)$ 精确性的充分条件，以确保该问题

$$\min_{(y,v,\lambda,\mu)\in\mathbb{R}^d \times L^2((a,b);\mathbb{R}^d)\times\mathbb{R}^{\ell+m}} \mathscr{L}(y, \mu, \lambda, \mu, c) \tag{13}$$

在某种意义上，对于任意充分大的 c，都等价于问题（7）。我们将用函数 f_0, $f_j, j \in J$, $g_i, i \in M$ 及其导数来表述这些条件。为避免对函数 f_0 及其一阶和二阶导数作出繁琐的假设，我们假设 f_0 关于 ξ 是二次的，且关于 u 有下界，尽管下面的定理在对函数 f_0 作出较弱限制的条件下也可以证明。

为便于表述这些充分条件，回顾一下：若函数φ: $X \to \mathbb{R}$ 满足对于任意序列 $\{x_n\} \subseteq X$，只要 $\|x_n\| \to +\infty$ 在 $n \to \infty$ 时成立，就有 $\varphi(x_n) \to +\infty$，则称其为强制函数。

定理3. 设以下假设成立：

1. 函数 f 具有如下形式

$$f(u, \xi, x) = \langle \xi, h_2(x)\xi \rangle + \langle h_1(u, x), \xi \rangle + h_0(u, x) \quad \forall u, \xi \in \mathbb{R}^d,\ x \in [a, b]$$

对于某个函数 $h_2 \in L^\infty((a,b);\ \mathbb{R}^{d\times d})$ 和某些连续函数h_1: $\mathbb{R}^d \times [a,b] \to \mathbb{R}^d$，以及 h_0: $\mathbb{R}^d \times [a,b] \to \mathbb{R}$，它们在 u 上二阶连续可微，并且满足

$$\langle \xi, h_2(x)\xi \rangle \geq \varkappa |\xi|^2 \quad \forall \xi \in \mathbb{R}^d,\ x \in [a, b]$$

对于某个 $\varkappa > 0$，以及 $|h_1(u,x)| \leq \eta(x)$ 和 $h_0(u,x) \geq -\eta(x)$，对于某个几乎处处非负的函数 $\eta \in L^2(a,b)$ 以及所有 $u \in \mathbb{R}^d$ 和 $x \in [a,b]$;

2. 函数 f_j、 $j \in J \cup \{0\}$ 和 g_i、 $i \in M$ 均二阶连续可微，并且函数 $f_0(\cdot) + c(|F(\cdot)|^2 + |\max\{G(\cdot)\ 0\}|^2)$ 对某个 $c > 0$ 是强制的;

3. 对任意 $z \in \mathbb{R}^{2d}$，梯度 $\nabla f_j(z)$、 $j \in J$ 和 $\nabla g_i(z)$、 $i \in \{k \in M \mid g_k(z) = 0\}$ 均线性无关。

那么，增广拉格朗日函数 $\mathscr{L}(y, v, \lambda, \mu, c)$ 是全局精确的，并且对于任意 $\gamma \in \mathbb{R}$，都存在 $c(\gamma) > 0$，使得对于所有 $c \geq c(\gamma)$，以下结论均成立：

1. 次水平集 $S_c(\gamma)$ 是有界的；2. 若 $(y_*, v_*, \lambda_*, \mu_*) \in S_c(\gamma)$ 是问题 (13) 的局部极小值点，则$u_* = \mathcal{A}(y_*, v_*)$ 是问题 (7) 的局部最优解；3. 四元组 $(y_*, v_*, \lambda_*, \mu_*) \in S_c(\gamma)$ 是问题 (13) 的驻点，当且仅当带有 $u_* = \mathcal{A}(y_*, v_*)$ 的 (u_*, λ_*, μ_*) 是问题 (7) 的 KKT点，且满足 $\mathcal{I}(u_*) \leq \gamma$。

证明。注意，定理的第一个假设意味着增长条件 (8) 成立。此外，可以直接验证，第一个假设还确保泛函 $\mathcal{I}$ 是二次连续 Fr´echet 可微的，并且其所有一阶和二阶导数在空间 X的有界子集上有界。

进而，根据 [5, 定理 3.23]，关于函数 h_2、 h_1和 $h_0\widehat{}$的条件保证泛函 $\mathcal{I}$（因而泛函 $\mathcal{I}$ 也）是弱序列下半连续的，而映射

$$(y,v)\mapsto \nabla_u f(\mathcal{A}(y,v)(\cdot),v(\cdot),\cdot),\quad (y,v)\mapsto \nabla_\xi f(\mathcal{A}(y,v)(\cdot),v(\cdot),\cdot)$$

由于它们关于 v 是仿射的（见 [4]），因此是弱序列连续的。还需注意，若序列 $\{(y_n,v_n)\}$ 在 X中弱收敛，则相应的序列 $\{(y_n,u_n(b))\}$ 在 $\mathbb{R}^{2d}$中收敛。利用这些事实，可以很容易验证，增广拉格朗日函数 $\mathscr{L}(y,v,\lambda,\mu,c)$在$(y,v)$中弱序列下半连续。

此外，定理的第一条假设意味着

$$f(u,\xi,x)\geq \varkappa|\xi|^2-\eta(x)|\xi|-\eta(x)\quad \forall \xi,u\in\mathbb{R}^d,\ \text{a.e. } x\in(a,b)$$

因此，若序列 $\{(y_n,v_n)\}\subset X$ 满足当 $n\to\infty$时 $\|v_n\|_2\to+\infty$，则 $\int_a^b f(u_n(x)\ v_n(x)\ x)\ dx\to+\infty$ 当 $n\to\infty$时成立。进而，若序列 $\{v_n\}$在 $L^2((a,b);\mathbb{R}^d)$中有界，但当 $n\to\infty$时 $|y_n|\to+\infty$ 成立，则序列$\{\int_a^b f(u_n(x)\ v_n(x)\ x)\ dx\}$ 有下界，而

$$f_0(y_n,u_n(b))+c\Big(|F(y_n,u_n(b))|^2+|\max\{G(y_n,u_n(b)),0\}|^2\Big)\to+\infty$$

根据定理的第二个假设，$n\to\infty$。因此，对于任意 $c>0$ 和 $\gamma\in\mathbb{R}$，集合 $\Omega_c(\gamma)$（见式 (3)）是有界的。

最后，我们来验证该定理的第三个假设保证：对于任意有界集 $V\subset X$，都存在 $\sigma>0$，使得对于所有 $(y,v)\in V$，$\sigma_{\max}(Q(y,v))\geq\sigma$。然后应用 [14, 推论 3.6 以及定理 5.3 和 5.7]，即可得到所需结果。

首先，我们证明对于任意 $(y,v)\in X$，$Q(y,v)[\cdot]$ 是正定的。事实上，任取 $(y,v)\in X$。根据引理 3.3]，函数 $Q(y,v)[\cdot]$ 是正定的，当且仅当线性算子 $\mathcal{T}$: $X\to\mathbb{R}^{\ell+m(y,v)}$ 定义为

$$\mathcal{T}(z,w)=\prod_{j=1}^{\ell}\{\langle\nabla\widehat{f}_j(y,v),(z,w)\rangle\}\times\prod_{i\in M(y,v)}\{\langle\nabla\widehat{g}_i(y,v),(z,w)\rangle\}$$

是满射，其中 $M(y,v)=\{i\in M\mid \widehat{g}_i(y,v)=0\}$ 以及 $m(y,v)$ 是集合 $M(y,v)$ 的基数。

显然，线性算子 $\mathcal{T}$ 是满射，当且仅当这些向量

$$\nabla\widehat{f}_j(y,v),\quad j\in\{1,\dots,\ell\},\quad \nabla\widehat{g}_i(y,v),\quad i\in M(y,v),\tag{14}$$

线性无关。因此，只要这些向量线性无关，函数 $Q(y,v)[\cdot]$ 就是正定的。让我们检查在定理的假设下这些向量确实是线性无关的。

反设向量 (14) 线性相关。于是，结合 (12) 可知，存在实数 α_j、$j \in \{1,\dots,\ell\}$ 和 β_i、$i \in M$（y、v），它们不全为零，使得

$$\sum_{j=1}^{\ell} \alpha_j (\nabla_{u_a} + \nabla_{u_b}) f_j(y, u(b)) + \sum_{i \in M(y,v)} \beta_j (\nabla_{u_a} + \nabla_{u_b}) g_i(y, u(b)) = 0,$$
$$\sum_{j=1}^{\ell} \alpha_j \nabla_{u_b} f_j(y, u(b)) + \sum_{i \in M(y,v)} \beta_j \nabla_{u_b} g_i(y, u(b)) = 0,$$

其中 $u = \mathcal{A}(y, v)$。因此，有

$$\sum_{j=1}^{\ell} \alpha_j \nabla f_j(y, u(b)) + \sum_{i \in M(y,v)} \beta_j \nabla g_i(y, u(b)) = 0,$$

这与定理的第三个假设矛盾。因此，对于任意 $(y, v) \in X$，函数$Q(y, v)[\cdot]$是正定的。

注意，根据函数 $Q(\cdot)$ 的定义（见 (5)），在所考虑的问题中，它具有如下形式

$$Q(y,v)[\lambda,\mu] = \left\langle \begin{pmatrix} \lambda \\ \mu \end{pmatrix}, E_Q(y, u(b)) \begin{pmatrix} \lambda \\ \mu \end{pmatrix} \right\rangle \quad \forall \lambda \in \mathbb{R}^{\ell},\ \mu \in \mathbb{R}^{m}$$

对于某个矩阵 $E_Q(y, u(b))$，该矩阵由函数 G 以及 F 和 G 的偏导数定义，因此连续依赖于 $(y, u(b))$。

容易看出，$\sigma_{\max}(Q(y, v))$ 与矩阵 $E_Q(y, u(b))$ 的最小特征值一致，因为该矩阵连续依赖于 $(y, u(b))$，而矩阵$E_Q(y, u(b))$ 连续依赖于 $(y, u(b))$。因此，考虑到矩阵 $E_Q(\cdot)$ 是正定的（因为对于所有 $(y, v) \in X$，都有$\sigma_{max}(Q(y, v)) > 0$），可以得出，对于任意集合 $V \subset X$，只要集合$V(a, b) = \{(y, u(b)) \in \mathbb{R}^{2d} \mid (y, v) \in V\}$ 有界，就存在 $\sigma > 0$ 使得$\sigma_{\max}(Q(y, v)) \geq \sigma$。最后注意，若集合 $V \subset X$ 有界，则集合 $V(a, b)$ 也有界。

□

3.2 等周问题

现在来考虑一种用于带等周约束的多维变分问题的精确增广拉格朗日函数。为简单起见，我们仅考虑含一个等周等式约束的情形，不过本小节的所有结果都可以推广到含多个等周等式约束和不等式约束的变分问题。

令 $E = \prod_{i=1}^{n} (a_i, b_i) \subset \mathbb{R}^n$ 为一个有界开盒，$\overline{u} \in W^{1,2}(E; \mathbb{R}^d)$ 为给定函数。考虑以下带等周等式约束的变分问题：

$$\begin{aligned} &\min_{u \in \overline{u} + W_0^{1,2}(E;\mathbb{R}^d)} \mathcal{I}_0(u) = \int_E f_0(u(x), \nabla u(x), x)\, dx \\ &\text{subject to } \mathcal{I}_1(u) = \int_E f_1(u(x), \nabla u(x), x)\, dx - \zeta = 0. \end{aligned} \tag{15}$$

其中 f_i: $\mathbb{R}^d \times \mathbb{R}^{d \times n} \times E \to \mathbb{R}$、$f_i = f_i(u, \xi, x)$ 是 Carath´eodory 函数，它们关于 u 和 ξ 可微，其偏导数 $\nabla_u f_i$ 和 $\nabla_\xi f_i$ 具有

Carath´eodory 函数，同时 $\zeta \in \mathbb{R}$ 是一个给定常数。尽管不失一般性，可以假设 $\zeta = 0$，但在许多应用中，考虑 $\zeta \neq 0$ 的情况比重新定义函数 f_1 更为方便。最后，函数 $\overline{u} \in W^{1,2}(E;\, \mathbb{R}^d)$ 定义了边界条件，因为

$$\{\overline{u}\} + W_0^{1,2}(E;\mathbb{R}^d) = \Big\{ u \in W^{1,2}(E;\mathbb{R}^d) \;\Big|\; \mathrm{Tr}(u) = \mathrm{Tr}(\overline{u}) \Big\},$$

其中 Tr(·) 是迹算子（参见 [1, 第 5.20–5.22] 节）。

我们对函数 f_i 及其偏导数施加以下增长条件。也就是说，存在 $C > 0$ 以及几乎处处非负的函数$\eta \in L^1(E)$ 和 $\theta \in L^2(E)$，使得

$$\begin{aligned} |f_i(u,\xi,x)| &\le C\big(|u|^2 + |\xi|^2\big) + \eta(x), \\ \max\big\{|\nabla_u f_i(u,\xi,x)|, |\nabla_\xi f_i(u,\xi,x)|\big\} &\le C\big(|u| + |\xi|\big) + \theta(x). \end{aligned} \tag{16}$$

对几乎处处的 $x \in E$ 以及任意 $u \in \mathbb{R}^d$、 $\xi \in \mathbb{R}^{d\times n}$ 和 $i \in \{0,1\}$均成立。值得注意的是，利用 Sobolev 嵌入定理 [1, 定理 5.4]，可以放宽这些增长条件。特别是，在 $n = 1$ 的情况下，只需假设形式为 (8) 的增长条件成立。式 (16) 中的第一个不等式可替换为 $|f_i(u,\xi,x)| \le C\big(|u|^q + |\xi|^2\big) + \eta(x)$，其中在 $n > 2$的情况下，$q = 2n/(n-2)$；而在 $n = 2$的情况下， q 是 $[2, +\infty)$ 中的任意数，等等。

增长条件 (16) 确保泛函 $\mathcal{I}_i$, $i \in \{0,1\}$ 是 Gâteaux 可微的，且其 Gâteaux 导数具有如下形式：

$$\mathcal{I}_i'(u)[w] = \int_E \big(\langle \nabla_u f_i(u(x), u'(x), x), w(x) \rangle + \langle \nabla_\xi f_i(u(x), u'(x), x), \nabla w(x) \rangle \big)\, dx \tag{17}$$

对于所有 $w \in W^{1,2}(E;\, \mathbb{R}^d)$（例如，参见 [5, 第 3.4 节2]）。按照上一小节中的相同方法进行论证，并利用 Nemytskii 算子 [2] 连续性的标准结果，可以验证增长条件 (16) 也保证 Gâteaux 导数 $\mathcal{I}_i'(\cdot)$ 连续，因此泛函 $\mathcal{I}_i$, $i \in \{0,1\}$ 是连续 Fréchet 可微的。

为了计算泛函 $\mathcal{I}_i$ 的梯度，我们将采用作者在 [12] 中提出的“转入导数空间”技术在多维情形下的推广。为此，假设问题 (15) 的所有 KKT点 (u_*, λ_*) 都满足 $u_* \in \overline{u} + MW^{n,2}{}_0(E;\, \mathbb{R}^d)$（参见第 2.2 小节）。换言之，假设相应欧拉–拉格朗日方程的解比问题表述中所假定的更为正则，即这些解存在所有阶数为 $k \in \{2, \dots, n\}$ 的弱混合导数，且这些导数属于 $L^2(E)$。那么问题 (15) 等价于以下变分问题：

$$\begin{aligned} &\min_{u \in MW_0^{n,2}(E;\mathbb{R}^d)} \ \mathcal{I}_0(\overline{u} + u) = \int_E f_0(\overline{u}(x) + u(x), \nabla\overline{u}(x) + \nabla u(x), x)\, dx \\ &\text{s.t. } \mathcal{I}_1(\overline{u} + u) = \int_E f_1(\overline{u}(x) + u(x), \nabla\overline{u}(x) + \nabla u(x), x)\, dx - \zeta = 0. \end{aligned} \tag{18}$$

根据定理2，该问题等价于以下问题：

$$\min_{v \in X} \ \widehat{\mathcal{I}}_0(v) := \mathcal{I}_0(\overline{u} + \mathcal{A}v) \quad \text{subject to } \widehat{\mathcal{I}}_1(v) := \mathcal{I}_1(\overline{u} + \mathcal{A}v) = \zeta, \tag{19}$$

其中线性算子 $\mathcal{A}$ 在式 (6) 中定义，且

$$X = \left\{ v \in L^2(E;\mathbb{R}^d) \;\middle|\; \int_{a_i}^{b_i} v\, ds_i = 0 \text{ for a.e. } x \in E \;\; \forall i \in \{1,\dots,n\} \right\}. \quad (20)$$

应用（17）并进行分部积分（更多细节参见 [12] ）后，可以得到$\widehat{\ }$，泛函 $\mathcal{I}_i$、$i \in \{0,1\}$是连续 Fr´échet 可微的，其 Fr´échet 导数具有以下形式

$$D\widehat{\mathcal{I}}_i[v](h) = \int_E \langle P_i(v)(x), h(x)\rangle dx \quad \forall h \in X,$$

其中

$$\begin{aligned} P_i(v)(x) =& (-1)^n \int_{x_1}^{b_1} \dots \int_{x_n}^{b_n} \frac{\partial f}{\partial u}(u(s), \nabla u(s), s)\, ds \\ &+ (-1)^{n-1} \sum_{i=1}^{n} \int_{x_n}^{b_n} \dots \int_{x_{i+1}}^{b_{i+1}} \int_{x_{i-1}}^{b_{i-1}} \dots \int_{x_1}^{b_1} \\ &\qquad \frac{\partial f}{\partial \xi_i}(u(s^i), \nabla u(s^i), s^i)\, ds_1 \dots ds_{i-1} ds_{i+1} \dots ds_n \end{aligned} \quad (21)$$

对于几乎处处的 $x \in E$, $u = \overline{u} + \mathcal{A}v$, $\partial f/\partial \xi_i = (\partial f/\partial \xi_{1i}, \dots, \partial f/\partial \xi_{di})$, 以及任意 $i \in \{1,\dots,n\}$, $s^i = (s_1, \dots, s_{i-1}, x_i, s_{i+1}, \dots, s_n)$。利用增长条件 (16)，可以很容易地验证，非线性算子 $P_i(\cdot)$ 将 $L^2(E)$ 连续映射到 $L^2(E)$。因此，泛函 $\widehat{\mathcal{I}}_i$ 在希尔伯特空间 X 中点 $v \in X$ 处的梯度，是函数 $P_i(v)$ 在 $L^2(E)$ 中到子空间 X 上的正交投影 $\mathrm{Pr}_X(P_i(v))$。利用以下命题可以解析地计算该投影（参见 [12, 命题 3]）。

命题1. 对于任意 $v \in L^2(E)$，有

$$\mathrm{Pr}_X\, v = v + \sum_{k=1}^{n} \sum_{\alpha \in I_k} (-1)^{|\alpha|} c_\alpha S_\alpha v,$$

其中，对于任意 $\alpha \in I_k$，都有 $c_\alpha = \prod_{i=1}^{n} (b_i - a_i)^{-\alpha_i}$，且

$$\Big(S_\alpha v\Big)(x) = \int_{a_{\alpha_{i_1}}}^{b_{\alpha_{i_1}}} \dots \int_{a_{\alpha_{i_k}}}^{b_{\alpha_{i_k}}} v\, ds_{i_k} \dots ds_{i_1},$$

且这些指标 $1 \le i_1 < \dots < i_k \le n$ 满足：当且仅当 $\alpha_r = 0$（也就是说，i_j 是多重指标 α的非零分量的指标）。

$\widehat{\ }$因此，$\nabla \mathcal{I}_i(v) = \mathrm{Pr}_X\, P_i(v)$，$i \in \{0,1\}$。问题 (19) 的精确增广拉格朗日函数$\mathscr{L}(v,\lambda,c)$，其中 $\lambda \in \mathbb{R}$，按照等式 (1) 和 (2) 定义如下：

$$\begin{aligned} \mathscr{L}(v,\lambda,c) =& \int_E f_0(u(x), \nabla u(x), x)\, dx + \lambda \bigg(\int_E f_1(u(x), \nabla u(x), x)\, dx - \zeta \bigg) \\ &+ \frac{c}{2}(1+\lambda^2) \bigg(\int_E f_1(u(x), \nabla u(x), x)\, dx - \zeta \bigg)^2 \\ &+ \frac{1}{2} \bigg(\int_E \langle \mathrm{Pr}_X\, P_1(v)(x), \mathrm{Pr}_X\, P_0(v)(x) + \lambda\, \mathrm{Pr}_X\, P_1(v)(x) \rangle\, dx \bigg)^2 \end{aligned}$$

其中 $u = \overline{u} + \mathcal{A}v$，或者等价地，在弱意义下 $v = D^{(1,\dots,1)}(u - \overline{u})$。

下面给出增广拉格朗日函数 $\mathscr{L}(v, \lambda, c)$ 精确性的充分条件。我们将以“抽象术语”表述这些条件，然后说明在某些特定情况下，如何将这些条件改写为对函数 f_0 和 f_1 的假设。

定理 4. 假设以下条件成立：

1. 对于问题 (15) 的所有 KKT点 (u_*, λ_*)，有 $u_* \in \overline{u} + MW^{n,\,2\,0}(E;\, \mathbb{R}^d)$;

2. 增长条件 (16) 成立；

3. 函数 f_0 和 f_1 在 u 和 ξ 上二阶可微，且它们相应的所有二阶导数都是 Carath´eodory 函数，并且在 $\mathbb{R}^d \times \mathbb{R}^{d\times n} \times \mathbb{R}^n$上本质有界；

4. 增广拉格朗日函数 $\mathscr{L}(v, \lambda, c)$ 在 (v, λ) 上弱序列下半连续；

5. 对于任意 $\gamma \in \mathbb{R}$，存在 $c > 0$，使得集合

$$\Omega_c(\gamma) = \left\{ u \in \overline{u} + MW_0^{n,2}(E;\mathbb{R}^d) \;\middle|\; \mathcal{I}_0(u) + c\big(\mathcal{I}_1(u) - \zeta\big)^2 \le \gamma \right\}$$

在 $MW^{n,2}(E;\, \mathbb{R}^d)$ 中有界；

6. 对于任意有界子集 $V \subset X$，存在 $\sigma > 0$，使得不等式$\|\mathrm{Pr}_X\, P_1(v)\|_2 \ge \sigma$ 对所有 $v \in V$均成立。

那么，增广拉格朗日函数 $\mathscr{L}(v, \lambda, c)$ 是全局精确的，并且对于任意 $\gamma \in \mathbb{R}$，存在 $c(\gamma) > 0$，使得对于所有 $c \ge c(\gamma)$，以下结论均成立：

1. 次水平集 $S_c(\gamma)$ 在 X 中有界；2. 若 $(v_*, \lambda_*) \in S_c(\gamma)$ 是 $\mathscr{L}(v, \lambda, c)$ 的一个局部极小值点，则 $u_* = \overline{u} + \mathcal{A}v_*$ 是问题 (15) 的局部最优解；3. 当且仅当 (u_*, λ_*)（其中 $u_* = \overline{u} + \mathcal{A}v_*$）是问题 (15) 的一个 KKT点，且$\mathcal{I}_0(u_*) \le \gamma$ 时，二元组 $(v_*, \lambda_*) \in S_c(\gamma)$ 是 $\mathscr{L}(v, \lambda, c)$ 的一个驻点。

证明。定理的第一个假设保证，Sobole夫空间 $W^{1,2}(E;\, \mathbb{R}^d)$ 中的等周问题 (15) 等价于在空间 $MW^{n,2}(E;\, \mathbb{R}^d)$ 中表述的该问题（见 (18)）；而后者又由定理2等价于问题 (19)。

定理的第二个和第三个假设保证，泛函 $\mathcal{I}_0$ 和 $\mathcal{I}_1$ 均为二阶连续 Fr´echet 可微，且其一阶和二阶 Fr´echet 导数在空间 X 的有界子集上有界。

注意，对于问题 (19)，有 $Q(v)[\lambda] = (\|\mathrm{Pr}_X\, P_1(v)\|_2)^4 \lambda^2$（见 (5)）。因此，定理的最后一个假设意味着，函数 $Q(\cdot)$ 在空间X的任意有界子集上都一致正定。因此，应用定理的第五个假设以及 [14, 推论 3.6 和定理 5.3、5.7]，可得定理的三个结论均成立。

□

注 4. 我们来讨论前一定理的假设：(i) 利用 Sobolev 嵌入定理，在某些特定情况下，可以放宽对函数 f_0 和f_1 的二阶导数本质有界性的假设。例如，在 $n = 1$ 的情况下，只需假设，对于任意 $R > 0$，存在 $C_R > 0$ 和几乎处处非负的函数$\eta_R \in L^1(E)$ 和 $\theta_R \in L^2(E)$，使得

$$\left|\frac{\partial^2 f}{\partial u^2}(u,\xi,x)\right| \le C_R|\xi|^2 + \eta_R(x), \quad \left|\frac{\partial^2 f}{\partial u \partial \xi}(u,\xi,x)\right| \le C_R|\xi| + \theta_R(x),$$
$$\left|\frac{\partial^2 f}{\partial \xi^2}(u,\xi,x)\right| \le C_R$$

对几乎处处的 $x \in E$ 以及所有满足 $|u| \le R$的 $(u,\xi) \in \mathbb{R}^{2d}$。

(ii) 增广拉格朗日函数 $\mathscr{L}(v,\lambda,c)$ 关于 (v,λ) 弱序列下半连续，若 f_0 关于 ξ 是凸的，且关于 (u,ξ)是二次的，f_1 关于 (u,ξ)是仿射的，并且函数f_0 和 f_1 满足某些标准增长条件。在这些假设下，根据 [5, Cor. 3.24 和 Thm. 3.26]，泛函 $\mathcal{I}_0(\cdot)$ 弱序列下半连续，泛函 $\mathcal{I}_1(\cdot)$ 弱序列连续，而函数

$$(v,\lambda) \mapsto \langle \mathrm{Pr}_X P_1(v)(\cdot), \mathrm{Pr}_X P_0(v)(\cdot) + \lambda \mathrm{Pr}_X P_1(v)(\cdot)\rangle$$

关于 v 是仿射的，而且也是弱序列连续的（见 [4]）。利用这些事实，可以很容易验证 $\mathscr{L}(v,\lambda,c)$ 关于 (v,λ) 弱序列下半连续。值得一提的是，在各种特殊情形下（例如情形$n = 1$），可以在限制较少的假设下证明增广拉格朗日函数的弱序列下半连续性。

(iii) 在 $n = 1$ 的情况下，只要满足以下条件，定理的第五个假设就成立：

$$f_0(u,\xi,x) \ge C_1|\xi|^2 + C_2|u|^q + \eta(x) \quad \forall (u,\xi,x) \in \mathbb{R}^{2d} \times E \tag{22}$$

存在某些 $C_1 > 0$、$C_2 \in \mathbb{R}$、$1 \le q < 2$ 和 $\eta \in L^1(E)$，因为该不等式意味着泛函 $u \mapsto \mathcal{I}_0(\overline{u} + u)$ 在 $W^{1,20}(E;\mathbb{R}^d) = MW^{1,20}(E;\mathbb{R}^d)$ 上是强制的。在 $n > 1$ 的情况下，罚函数 $\mathcal{I}_0(\cdot) + c(\mathcal{I}_1(\cdot) - \zeta)^2$ 的次水平集 $\Omega_c(\gamma)$ 何时在 $MW^{n,20}(E;\mathbb{R}^d)$ 中有界，仍是一个颇具挑战性的开放问题，有待未来研究。

(iv) 我们指出定理的最后一个假设成立的一些特殊情形。如果函数 f_1 关于 (u,ξ) 是仿射的，即 $f_1(u,\xi,x) = \langle h_1(x)\ \xi\rangle + \langle h_2(x)\ u\rangle + h_3(x)$，那么算子 $P_1(\cdot)$ 是常值的；只要 $P_1(\cdot)$ 在 X 上的投影非零，定理的第四个假设便成立。注意，根据 [12, 命题 5]，当 h_1 可微且函数 $h_2(\cdot) - \sum_{i=1}^n D_i h_1(\cdot)$ 不恒为零时，关于 P_1 的投影的假设成立。若函数 f_1 关于 (u,ξ) 是二次的，那么只要对任意 $v \in X$ 都有 $\mathrm{Pr}_X P_1(v) \ne 0$，定理的第四个假设便成立；因为此时函数 $P_1(\cdot)$ 是仿射的，这意味着范数 $\|\mathrm{Pr}_X P_1(\cdot)\|_2$ 弱下半连续。利用这一事实，可以证明所需的 $\sigma > 0$ 存在。注意，根据 [12, 命题 5]，在某些光滑性假设下，不等式 $\mathrm{Pr}_X P_1(v) \ne 0$ 意味着泛函 $\mathcal{I}_1$ 的欧拉–拉格朗日方程无解。一般而言，定理的最后一个假设意味着泛函 $\mathcal{I}_1$ 的梯度在有界集上远离零；在某些光滑性假设下，这意味着 $\mathcal{I}_1(\cdot)$ 的欧拉–拉格朗日方程无解。

下面我们给出增广拉格朗日函数 $\mathscr{L}(v,\lambda,c)$ 具有完全精确性的充分条件，这些条件不涉及对空间 $MW^{n,20}(E;\mathbb{R}^d)$ 中次水平集 $\Omega_c(\gamma)$ 的有界性这一严格假设，也不涉及增广拉格朗日函数的弱序列下半连续性。

定理 5. 定理 4 的假设 1–3 以及以下假设成立：

1. 对任意 $\gamma\in\mathbb{R}$，存在 $r=r(\gamma)>0$，使得集合

$$Z_r(\gamma):=\left\{u\in W_0^{1,2}(E;\mathbb{R}^d)\;\middle|\;\mathcal{I}_0(\overline{u}+u)+r(\mathcal{I}_1(\overline{u}+u)-\zeta)^2\le gamma\right\}\quad(23)$$

在 $W^{1,\,2}(E;\,\mathbb{R}^d)$ 中有界；

2. 对任意子集 $V\subset X$，若集合 $\{u=\mathcal{A}v\mid v\in V\}$ 在 $W^{1,\,20}(E;\,\mathbb{R}^d)$ 中有界，则存在 $\sigma>0$，使得不等式 $\|\mathrm{Pr}_X P_1(v)\|_2\ge\sigma$ 对所有 $v\in V$ 成立；

那么，对于任意 $\gamma\in\mathbb{R}$，增广拉格朗日函数 $\mathscr{L}(v,\lambda,c)$ 在集合 $S_c(\gamma)$ 上完全精确的充要条件是：存在无界递增序列 $\{c_k\}\subset(0,+\infty)$，使得对任意 $k\in\mathbb{N}$，增广拉格朗日函数 $\mathscr{L}(\cdot,c_k)$ 在集合 $X\times\mathbb{R}$ 上取得全局最小值。

证明。定理 4 的第二个和第三个假设保证泛函 $\widehat{\mathcal{I}_0}$ 和 $\widehat{\mathcal{I}_1}$ 都是二阶连续 Fr´echet 可微的。此外，增长条件 (16) 以及在 (u,ξ) 中 f_0 和 f_1 的二阶导数的本质有界性意味着，泛函 $\mathcal{I}_0$ 和 $\mathcal{I}_1$ 及其一阶和二阶导数在任意集合 $V\subset X$ 上都是有界的，其中集合 $\{u=\mathcal{A}v\mid v\in V\}$ 在 Sobole夫空间 $W^{1,\,2}(E;\,\mathbb{R}^d)$ 中有界。特别地，对于任意 $\gamma>0$，它们在集合

$$V=\left\{v\in X\left\{v\in X\;\middle|\;\widehat{\mathcal{I}}_0(v)+r(\gamma)\big(\widehat{\mathcal{I}}_1(v)-\zeta\big)^2\le gamma\right\},\right.$$

上是有界的，因为 $\overline{u}+\{u=\mathcal{A}v\mid v\in V\}\subset Z_{r(\gamma)}(\gamma)$，并且根据我们的假设，集合 $Z_{r(\gamma)}(\gamma)$ 在 $W^{1,2}(E;\,\mathbb{R}^d)$ 中有界。

反过来，关于 $\|\mathrm{Pr}_X P_1(v)\|_2$ 的假设保证，对于任意 $\gamma>0$，存在 $\sigma(\gamma)>0$，使得对于所有 $v\in V$，$\sigma_{\max}(Q(v))\ge\sigma(\gamma)$。因此，如果 $\mathscr{L}(\cdot,c_k)$ 在集合 $X\times\mathbb{R}$ 上对于某个递增无界序列 $\{c_k\}$ 取得全局最小值，那么根据定理 1，增广拉格朗日函数 $\mathscr{L}(v,\lambda,c)$ 在集合 $S_c(\gamma)$ 上是完全精确的。

反之，如果增广拉格朗日函数是完全精确的，那么根据完全精确性的定义，对于任意充分大的 $c>0$，它必定在 $X\times\mathbb{R}$ 上取得全局最小值。 □

注5. 如果存在 $C_1>0$、$C_2>0$、$q\in[1,p)$ 和 $\eta\in L^1(E)$ 使得，则次水平集 (23) 有界。

$$f_0(u,\xi,x)\ge C_1|\xi|^2+C_2|u|^q+\eta(x)$$

对几乎处处的 $x\in E$ 以及所有的 $(u,\xi)\in\mathbb{R}^d\times\mathbb{R}^{d\times n}$，由于这一增长条件保证泛函 $\mathcal{I}_0(\cdot)$ 是强制的（例如参见 [5, 定理 3.30]）。

3.3 非完整等式约束问题

在本小节中，我们研究带有非完整等式约束（可视为一阶偏微分方程约束）的多维变分问题的精确增广拉格朗日函数。与上一节一样，设 $E = \prod_{i=1}^{n}(a_i, b_i) \subset \mathbb{R}^n$ 是一个有界开盒，且 $\overline{u} \in W^{1,2}(E; \mathbb{R}^d)$ 是一个给定函数。考虑如下带有非完整等式（一阶偏微分方程）约束的变分问题：

$$\begin{aligned} &\min_{u \in \overline{u} + W_0^{1,2}(E;\mathbb{R}^d)} \mathcal{I}(u) = \int_E f(u(x), \nabla u(x), x)\, dx \\ &\text{subject to } F(u(x), \nabla u(x), x) = 0 \text{ for a.e. } x \in E. \end{aligned} \tag{24}$$

这里，$f\colon \mathbb{R}^d \times \mathbb{R}^{d\times n} \times E \to \mathbb{R}$, $f = f(u, \xi, x)$，以及 $F\colon \mathbb{R}^d \times \mathbb{R}^{d\times n} \times E \to \mathbb{R}^\ell$,$F = F(u, \xi, x)$，均为 Carath´eodory 函数，在 u 和 ξ 上可微，且它们关于 u 和 ξ 的偏导数也都是 Carath´eodory 函数。

假设函数 f 满足形式为 (16) 的增长条件，而函数 F 满足以下增长条件。具体而言，存在 $C > 0$ 和一个几乎处处非负的函数 $\eta \in L^2(E)$，使得

$$|F(u, \xi, x)| \le C\big(|u| + |\xi|\big) + \eta(x), \quad \max\big\{|\nabla_u F(u, \xi, x)|, |\nabla_\xi F(u, \xi, x)|\big\} \le C. \tag{25}$$

这些假设保证泛函 $\mathcal{I}(\cdot)$ 连续 Fr´echet 可微，而 Nemytskii 算子 $\mathcal{F}(u) = F(u(\cdot), \nabla u(\cdot), \cdot)$ 将 Sobole夫空间 $W^{1,2}(E; \mathbb{R}^d)$ 映射到 $L^2(E; \mathbb{R}^\ell)$，且也连续 Fr´echet 可微（参见例如 [2]）。$\mathcal{I}$ 的 Fr´echet 导数具有形式 (17)，而 Nemytskii 算子 $\mathcal{F}$ 的 Fr´echet 导数定义为

$$D\mathcal{F}(u)[w] = \nabla_u F(u(\cdot), \nabla u(\cdot), \cdot) w(\cdot) + \sum_{i=1}^{n} \frac{\partial F}{\partial \xi^i}(u(\cdot), \nabla u(\cdot), \cdot) \nabla w_i(\cdot) \tag{26}$$

对于任意 $u, w \in W^{1,2}(E; \mathbb{R}^d)$，其中 $\partial/\partial\xi^i = (\partial/\partial\xi_{i1}, \ldots, \partial/\partial\xi_{in})$ 和 $w = (w_1, \ldots, w_d)$。

为了定义问题 (24) 的精确增广拉格朗日函数，我们将采用与等周约束情形相同的技巧，即对问题 (24) 的极小化解的正则性施加一个额外假设。具体而言，假设问题 (24) 的所有 KKT点 (u_*, λ_*) 都满足条件 $u_* \in \overline{u} + MW^{n,20}(E; \mathbb{R}^d)$。那么，考虑到定理2，可以得出结论：问题 (24) 等价于以下问题：

$$\min_{v \in X} \widehat{\mathcal{I}}(v) := \mathcal{I}(\overline{u} + \mathcal{A}v) \text{ subject to } \widehat{\mathcal{F}}(u) := \mathcal{F}(\overline{u} + \mathcal{A}v) = 0, \tag{27}$$

其中，希尔伯特空间 X 在 (20) 中定义。按照前一节$\widehat{\ }$的同样方法推导，可得泛函 $\mathcal{I}$ 在点 $v \in X$ 处的梯度具有形式 $\mathrm{Pr}_X P(v)$，其中 $P(v)$ 根据等式 (21) 定义，只需将 f_i 替换为 f。进而，Nemytskii算子 $\widehat{\mathcal{F}}(\cdot)$ 将 X 映射到 $L^2(E; \mathbb{R}^\ell)$，且连续 Fr´echet 可微。应用 (26) 并分部积分$\widehat{\ }$，可得 $\mathcal{F}(\cdot)$ 的 Fr´echet 导数具有形式 $D\mathcal{F}(v)[h] =$

对于任意 v、$h \in X$，$P_F(v)h$，其中

$$\begin{aligned} P_F(v)(x) =&(-1)^n \int_{x_1}^{b_1} \dots \int_{x_n}^{b_n} \nabla_u F(u(s), \nabla u(s), s)\, ds \\ &+(-1)^{n-1} \sum_{i=1}^{n} \int_{x_n}^{b_n} \dots \int_{x_{i+1}}^{b_{i+1}} \int_{x_{i-1}}^{b_{i-1}} \dots \int_{x_1}^{b_1} \\ &\qquad \frac{\partial F}{\partial \xi_i}(u(s^i), \nabla u(s^i), s^i)\, ds_1 \dots ds_{i-1} ds_{i+1} \dots ds_n \end{aligned}$$

对几乎处处的 $x \in E$，其中 $u = \overline{u} + \mathcal{A}v$。注意，增长条件 (25) 意味着，对于任意 $v \in X$，$P_F(v) \in L^\infty(E;\, \mathbb{R}^{\ell \times d})$。

问题 (27) 的精确增广拉格朗日函数 $\mathscr{L}(v, \lambda, c)$，其中 $\lambda \in L^2(E;\, \mathbb{R}^\ell)$，根据等式 (1) 和 (2) 定义如下：

$$\begin{aligned} \mathscr{L}(v, \lambda, c) = &\int_E f(u(x), \nabla u(x), x)\, dx + \int_E \langle \lambda(x), F(u(x), \nabla u(x), x) \rangle\, dx \\ &+ \frac{c}{2}(1 + \|\lambda\|_2^2) \int_E |F(u(x), \nabla u(x), x)|^2\, dx \\ &+ \frac{1}{2} \int_E \left| P_F(v)(x) \big(\operatorname{Pr}_X P(v)(x) + P_F(v)(x)^* \lambda(x) \big) \right|^2 dx \end{aligned}$$

其中 $u = \overline{u} + \mathcal{A}v$和 $P_F(v)(x)^*$是矩阵 $P_F(v)(x)$的转置。

注6. 请注意，在本节中，我们假设拉格朗日乘子 λ属于 $L^2(E;\mathbb{R}^\ell)$。然而，对于多维变分问题，即使乘子 $\lambda \in L^1(E;\mathbb{R}^\ell)$也并非总是存在（参见 [15, 第2.3] 节）。因此，为了使精确增广拉格朗日函数的一般理论能够适用于具有非完整约束的变分问题，本节仅考虑这样一类具有非完整约束的问题：对于这类问题，所有全局最优解均存在乘子 $\lambda \in L^2(E;\mathbb{R}^\ell)$（参见注2）。此外，需要强调的是，在下文定理的表述中，只有当 $\lambda_* \in L^2(E;\, \mathbb{R}^\ell)$时，一对 (u_*, λ_*) 才是问题 (24) 的KKT点。因此，如果存在一个点 u_*，它满足问题 (24) 的KKT最优性条件，且其乘子 λ_* 不属于 $L^2(E;\mathbb{R}^\ell)$，那么在我们的理论框架下，该点 u_* 不是驻点，并且一对 (u_*, λ_*) 不是问题 (24) 的KKT点。

下面给出增广拉格朗日函数 $\mathscr{L}(v, \lambda, c)$ 精确性的充分条件。为简便起见，我们假设映射 F 关于 (u, ξ) 是仿射的（也就是说，问题 (24) 是一个约束由线性偏微分方程组定义的变分问题），并在此假设下表述这些条件，因为这是使相应的 Nemytskii算子 $\mathcal{F}(\cdot)$ 二次连续 Fr´echet 可微的最简单假设。

定理 6. 假设以下条件成立：

1. 对问题 (24) 的所有 KKT点 (u_*, λ_*)，均有 $u_* \in \overline{u} + MW^{n,\,20}(E;\, \mathbb{R}^d)$；2. 函数 f 在 (u, ξ) 中二次可微，满足增长条件 (16)，且其二阶导数是 Carath´eodory 函数，在 $\mathbb{R}^d \times \mathbb{R}^{d \times n} \times \mathbb{R}^n$ 上本质有界；

3. 增广拉格朗日函数 $\mathscr{L}(v,\lambda,c)$ 在 (v,λ) 中弱序列下半连续；

4. 函数 F具有如下形式： $F(u,\xi,x) = A(x)u+\sum_{i=1}^{n} B_i(x)\xi_i + D(x)$，其中 A、B_i 和 D 是本质有界可测函数，且 $P_F(\cdot) \neq 0$（如果函数 B_i 连续可微，则只需假设 $A(x) - \sum_{i=1}^{n} D_i B_i(x) \neq 0$ 在一个正测度集上）；

5. 对任意 $\gamma \in \mathbb{R}$，存在 $c > 0$，使得集合

$$\Omega_c(\gamma) = \left\{ u \in \overline{u} + MW_0^{n,2}(E;\mathbb{R}^d) \;\middle|\; \mathcal{I}_0(u) + c\|F(u(\cdot),\nabla u(\cdot),\cdot)\|_2^2 \leq \gamma \right\}$$

在 $MW^{n,2}$中，$(E;\mathbb{R}^d)$ 有界。

则增广拉格朗日函数 $\mathscr{L}(v,\lambda,c)$ 是全局精确的，并且对任意 $\gamma \in \mathbb{R}$，存在 $c(\gamma) > 0$，使得对所有 $c \geq c(\gamma)$，下列结论成立：

1. 次水平集 $S_c(\gamma)$ 在 X 中有界；

2. 若 $(v_*,\lambda_*) \in S_c(\gamma)$ 是 $\mathscr{L}(v,\lambda,c)$ 的一个局部极小点，则 $u_* = \overline{u} + \mathcal{A}v_*$ 是问题（24）的局部最优解；3. 当且仅当 (u_*,λ_*) 连同 $u_* = \overline{u} + \mathcal{A}v_*$ 是问题（24）的一个 KKT点，且满足 $\mathcal{I}(u_*) \leq \gamma$ 时，点对 $(v_*,\lambda_*) \in S_c(\gamma)$ 是 $\mathscr{L}(v,\lambda,c)$ 的一个驻点。

该定理的证明大体重复了定理 4 的证明，内容是验证定理的假设能够确保 [14, 推论 3.6 以及定理 5.3 和 5.7]的假设成立。为简洁起见，此处略去证明。

注 7. 注意，增广拉格朗日函数 $\mathscr{L}(v,\lambda,c)$ 关于 (v,λ) 弱序列下半连续；特别是，当 $n = 1$时，函数 $F = F(u,\xi,x)$ 不依赖于 ξ （即约束是完整约束），且函数 f具有如下形式：

$$f(u,\xi,x) = \langle \xi, h_2(x)\xi \rangle + \langle h_1(x),\xi) \rangle + h_0(u,x)$$

其中 h_2、h_1和 h_0 是某些函数，且矩阵 $h_2(x)$ 对几乎处处的 $x \in E$ 都是半正定的。在这些假设下，由 [5, 推论 3.24]，可知，泛函 $\mathcal{I}(\cdot)$ 弱序列下半连续（前提是 h_2 满足适当的增长条件）；映射 $\Pr_X P(\cdot)$ 弱连续；映射 $P_F(\cdot)$ 不依赖于 v；而根据 Rellich-Kondrachov 定理（例如，参见 [1, 定理 6.2]），Nemytskii算子$u \mapsto F(u(\cdot)\ \nabla u(\cdot)\ \cdot)$ 将弱连续序列映射为强连续序列。综合这些事实可以验证，增广拉格朗日函数确实关于 (v,λ) 弱序列下半连续。增广拉格朗日函数在非完整约束情形下何时弱序列下半连续，是一个有趣的开放问题。

我们还基于定理 1 给出增广拉格朗日函数完全精确性的充分条件。应用该定理可以避免对增广拉格朗日函数的弱序列下半连续性以及空间 $MW^{n,2}(E;\mathbb{R}^d)$ 中次水平集 $\Omega_c(\gamma)$ 的有界性提出极为严格且难以验证的假设。

定理 7. 设以下假设成立：

1. 对于问题 (24) 的所有 KKT点 (u_*, λ_*)，有 $u_* \in \overline{u}+MW^{n,2\,0}(E;\,\mathbb{R}^d)$ ；

2. 函数 f 关于 (u,ξ) 二次可微，满足增长条件 (16)，且其二阶导数是 Carath ´eodory 函数，在 $\mathbb{R}^d \times \mathbb{R}^{d\times n} \times \mathbb{R}^n$ 上本质有界；3. 函数 F具有形式 $F(u,\xi,x) = A(x)u + \sum_{i=1}^n B_i(x)\xi_i + D(x)$，其中 A、 B_i 和 D是某些本质有界可测函数，且 $P_F(\cdot) \neq 0$（若函数 B_i 连续可微，则只需假设 $A(x) - \sum_{i=1}^n D_i B_i(x) \neq 0$ 在某个正测度集合上）；

4. 对于任意 $\gamma \in \mathbb{R}$，存在 $r = r(\gamma) > 0$ 使得集合

$$Z_r(\gamma) := \left\{ u \in W_0^{1,2}(E;\mathbb{R}^d) \;\middle|\; \mathcal{I}(\overline{u}+u) + r\|F(u(\cdot),\nabla u(\cdot),\cdot)\|_2^2 \le \gamma \right\}$$

在 $W^{1,2}(E;\,\mathbb{R}^d)$ 中有界。

那么，对于任意 $\gamma \in \mathbb{R}$，增广拉格朗日函数 $\mathscr{L}(v,\lambda,c)$ 在集合 $S_c(\gamma)$ 上完全精确的充要条件是：存在一个无界递增序列 $\{c_k\} \subset (0,+\infty)$，使得对于任意 $k \in \mathbb{N}$，增广拉格朗日函数 $\mathscr{L}(\cdot,c_k)$ 在集合 $X \times L^2(E;\,\mathbb{R}^\ell)$ 上达到全局最小值。

该定理的证明大体上重复了定理 5 的证明，并基于验证定理 1 的假设。因此，为简洁起见，我们略去其证明。

4 最优控制的一个应用

下面我们采用线性演化方程控制问题中的标准定义和结果，这些内容可见于 [18]。设 $\mathscr{H}$ 和 $\mathscr{U}$ 为复希尔伯特空间， $\mathbb{T}$ 是 $\mathscr{H}$ 上的一个强连续半群，其生成元为 $\mathcal{A}$: $\mathcal{D}(A) \to \mathscr{H}$，其中 $\mathcal{D}(A)$ 是 A 的定义域。再设 $\mathcal{B}$ 是 $\mathbb{T}$ 的一个可容许控制算子（见 [18, 定义 4.2.1]）。

考虑以下固定端点最优控制问题：

$$\begin{aligned} &\min_{(x,u)} \ \frac{1}{2}\int_0^T \|u(t)\|_{\mathscr{U}}^2\,dt \\ &\text{subject to } \ \dot{x}(t) = \mathcal{A}x(t) + \mathcal{B}u(t), \quad t\in[0,T], \quad x(0)=x_0, \quad x(T)=x_T. \end{aligned} \tag{28}$$

这里 $T > 0$ 和 x_0、 $x_T \in \mathscr{H}$为固定值，控制 $u(\cdot)$ 属于空间$L^2((0,T);\,\mathscr{U})$。

我们的目标是将固定端点问题 (28) 转化为一个等价的自由端点问题，其对应的精确增广拉格朗日函数如下。为定义这个增广拉格朗日函数，引入所谓的输入映射 $F_t u = \int_0^t \mathbb{T}_{t-\tau}\mathcal{B}u(\tau)d\tau$, $t \ge 0$，它对应于数对 $(\mathcal{A},\mathcal{B})$。假设 $\mathcal{B}$ 是 $\mathbb{T}$ 的可容许控制算子，这保证了对于任意 $t \in T$，输入映射 F_t 都是从 $L^2((0,T);\,\mathscr{U})$ 到 $\mathscr{H}$的有界线性算子，见 [18, 命题 4.2.2]。此外，根据 [18,命题 4.2.5]，对于任意 $x_0 \in \mathscr{H}$，初值问题

$$\dot{x}(t) = \mathcal{A}x(t) + \mathcal{B}u(t), \quad x(0) = x_0 \tag{29}$$

有唯一解 $x \in C([0,T];\mathscr{H})$，其表达式为

$$x(t) = \mathbb{T}_t x_0 + F_t u \quad \forall t \in [0,T]. \tag{30}$$

因此，问题 (28) 可以改写如下：

$$\min_{u \in L^2((0,T);\mathscr{U})} \ \frac{1}{2}\int_0^T \|u(t)\|_{\mathscr{U}}^2 \, dt \quad \text{subject to} \quad F_T u = \widehat{x}_T, \tag{31}$$

其中 $\widehat{x}_T = x_T - \mathbb{T}_T x_0$。该问题的精确增广拉格朗日函数具有如下形式

$$\begin{aligned}\mathscr{L}(u,\lambda,c) =& \frac{1}{2}\int_0^T \|u(t)\|_{\mathscr{U}}^2 \, dt + \langle \lambda, F_T u - \widehat{x}_T\rangle \\ &+ \frac{c}{2}(1+\|\lambda\|_{\mathscr{H}}^2)\big\|F_T u - \widehat{x}_T\big\|^2 + \frac{1}{2}\Big\|F_T\big(u + F_T^*\lambda\big)\Big\|^2,\end{aligned}$$

其中，$\lambda \in \mathscr{H}$是与端点约束 $x(T) = x_T$ 对应的拉格朗日乘子，而 F_T是 F_T 的伴随算子。

注 8. 注意，要计算某个 $v \in L^2((0,T);\mathscr{U})$ 的 $F_T v$，需要求解初值问题

$$\dot{x}(t) = \mathcal{A}x(t) + \mathcal{B}v(t), \quad x(0) = 0$$

并令 $F_T v = x(T)$。还需指出，对于任意 $t \in [0,T]$ 和 $\lambda \in \mathscr{H}$，均有 $(F_T^*\lambda)(t) = \mathcal{B}^*\mathbb{T}_{T-t}^*\lambda$。如果算子 $\mathcal{A}$ 是自伴的，那么对于所有 $t \geq 0$，均有 $\mathbb{T}_t = \mathbb{T}_t^*$；对于任意 $t \in [0,T]$，均有 $\mathbb{T}_{T-t}^*\lambda = y(T-t)$，其中 y 是初值问题 $\dot{y}(t) = \mathcal{A}y(t)$，$y(0) = \lambda$ 的解（见 (30)）。

下面给出增广拉格朗日函数 $\mathscr{L}(u,\lambda,c)$ 精确性的充分条件，以确保该问题

$$\min_{(u,\lambda) \in L^2((0,T);\mathscr{U})\times\mathscr{H}} \mathscr{L}(u,\lambda,c) \tag{32}$$

对于某个大于 $c_* > 0$ 的任意 c，它等价于问题 (28)。注意，问题 (32) 可以改写为以下自由端点最优控制问题：

$$\begin{aligned}\min_{(x,u,\lambda)} \ & \frac{1}{2}\int_0^T \|u(t)\|_{\mathscr{U}}^2 \, dt + \langle \lambda, x(T) - x_T\rangle + \frac{c}{2}(1+\|\lambda\|_{\mathscr{H}}^2)\|x(T) - x_T\|^2 \\ & + \frac{1}{2}\Big\|F_T\big(u + F_T^*\lambda\big)\Big\|^2 \\ \text{subject to } & \dot{x}(t) = \mathcal{A}x(t) + \mathcal{B}u(t), \quad t \in [0,T], \quad x(0) = x_0.\end{aligned}$$

为了方便地表述该问题与问题 (28) 等价的充分条件，回顾一下，若对于任意初始状态 $x_0 \in \mathscr{H}$ 和任意终态 $x_T \in \mathscr{H}$，都存在一个控制 $u \in L^2((0,T);\mathscr{U})$，使得初值问题 (29) 的相应解 $x(\cdot)$ 满足等式 $x(T) = x_T$，则称系统 (29) 是关于 L^2- 控制的精确可控系统（参见例如 [16]）。精确可控系统的具体例子可见 [3,16,18]。

定理 8. 设系统 (29) 可通过 L^2-控制实现精确可控。那么增广拉格朗日函数 $\mathscr{L}(u,\lambda,c)$ 是全局精确的，并且对于任意 $\gamma \in \mathbb{R}$，存在 $c(\gamma) > 0$，使得对于所有 $c \geq c(\gamma)$，以下陈述均成立：

1. 次水平集 $S_c(\gamma)$ 有界；

2. 如果点（u_*， λ_*） $\in S_c$（γ）是 $\mathscr{L}$（u， λ， c）的一个驻点，那么 u_* 是问题（31）的全局最优解，并且（u_*， λ_*）是该问题的 KKT 点。

证明。在问题（31）的情形下，对于任意 $u \in L^2$（（0， T）； $\mathscr{U}$），函数 Q（u）具有如下形式： Q（u）$[\lambda] = 0.5\|F_T$（$F_T^*\lambda$）$\|^2$，且不依赖于 u。注意，由初值问题（29）的解的表达式（30）可知，系统（29）的精确能控性等价于输入映射 F_T 的满射性。因此，根据定理的假设以及 $[14,$ 引理 3。3], 函数 Q（$\cdot$）在整个空间 L^2（（0， T）； $\mathscr{U}$）上是一致正定的。

由于问题（31）的目标函数显然是强制的，因此对于任意 $c \geq 0$ 和 $\gamma \in \mathbb{R}$, 次水平集 $\Omega_c(\gamma)$ 都是有界的。于是应用 $[14$,推论 3.6] 可得，对于任意充分大的 $c = c(\gamma)$，次水平集 $S_c(\gamma)$是有界的。

注意，增广拉格朗日函数 $\mathscr{L}(u, \lambda, c)$可以表示为关于$(u, \lambda)$的凸函数

$$\frac{1}{2}\int_0^T \|u(t)\|_{\mathscr{U}}^2\, dt + \langle \lambda, F_T u - \widehat{x}_T\rangle + \frac{c}{2}\|F_T u - \widehat{x}_T\|^2 + \frac{1}{2}\left\|F_T\left(u + F_T^*\lambda\right)\right\|^2$$

与函数 $\omega(u, \lambda) = \|\lambda\|_{\mathscr{H}}^2\|F_T u - \widehat{x}_T\|^2$之和。容易验证，该函数是弱序列下半连续的，因为它是两个非负凸函数（因而都是序列下半连续函数）的乘积。因此，对于任意 $c \geq 0$ ，增广拉格朗日函数 $\mathscr{L}(\cdot, c)$都是弱序列下半连续的。于是应用 $[14,$ 定理 5.3] 可得， $\mathscr{L}(u, \lambda, c)$是全局精确的。

最后，应用 $[14,$ 定理 5.7] 可得，对于任意充分大的 c驻点 $(u_*, \lambda_*) \in S_c(\gamma)$, $\mathscr{L}(u, \lambda, c)$ 恰好是问题 (31) 的 KKT 点。该问题的凸性意味着，任何这样的 u_* 都是全局最优的，也就是说，定理的最后一个结论成立。 □

注 9. 值得注意的是，关于系统 (29) 的精确能控性这一假设，已被用于与问题 (28) 的非光滑罚函数精确性密切相关的研究结果中。具体而言，在 $[16]$ 中已证明，对于任意充分大的 $c \geq 0$ ，问题 (28) 等价于自由端点罚问题。

$$\begin{aligned} &\min_{(x,u)} \ \frac{1}{2}\int_0^T \|u(t)\|_{\mathscr{U}}^2\, dt + c\|x(T) - x_T\|_{\mathscr{H}} \\ &\text{subject to } \dot{x}(t) = \mathcal{A}x(t) + \mathcal{B}u(t), \quad t \in [0, T], \quad x(0) = x_0 \end{aligned}$$

在假设系统 (29) 精确能控，且存在$C > 0$ 使得精确能控性定义中的控制满足不等式$\|u\|_{L^2((0,T);\mathscr{U})} \leq C(\|x_0\| + \|x_T\|)$ 的情况下。反过来， $[13]$ 中证明了可以去掉存在 C 这一假设；此外，精确能控性假设可以替换为较弱的假设，即输入映射 F_T的像是闭的。该映射的像闭是否足以保证增广拉格朗日函数 $\mathscr{L}(u, \lambda, c)$ 的全局精确性，仍是一个开放问题。

参考文献

[1] R. A. Adams。《Sobolev空间》。Academic Press，纽约，1975。[2] J. Appell 和 P. P. Zabrejko。《非线性叠加算子》。Cambridge University Press，纽约，1990。[3] C. Bardos、G. Lebeau 和 J. Rauch。关于从边界对波进行观测、控制和稳定化的尖锐充分条件。SIAM J. Control Opim.，30：1024–1065，1992。[4] B. Dacorogna。《非线性泛函的弱连续性与弱下半连续性》。Springer-Verlag，柏林、海德堡，1982。[5] B. Dacorogna。《变分法中的直接方法》。Springer，纽约，2007。[6] V. F. Demyanov。通过精确惩罚技术处理变分法中的约束问题。载于 P. Daniele、F. Giannessi 和 A. Maugeri 编，《均衡问题与变分模型》，第 79–108 页。Springer，波士顿，2003。[7] V. F. Dem'yanov。精确罚函数与变分法问题。Autom. Remote Control，65：280–290，2004。[8] V. F. Demyanov。一个老问题与新工具。Optim. Methods Softw.，20：53–70，2005。[9]V. F. Demyanov 和 F. Giannessi。涉及高阶导数约束的变分问题。载于 P. Daniele、F. Giannessi 和 A. Maugeri 编，《均衡问题与变分模型》，第 109–134 页。Springer，波士顿，2003。[10] V. F. Demyanov 和 G. S. Tamasyan。等周问题中的精确罚函数。Optim.，60：153–177，2011。[11] V. F. Demyanov 和 G. S. Tamasyan。参数化移动边界变分问题的直接方法。Numer. Funct. Anal. Optim.，35：934–961，2014。[12] M. V. Dolgopolik。变分法中若干多维问题的新型直接数值方法。Numer. Funct. Anal. Optim.，39：467–490，2018。[13] M. V. Dolgopolik。最优控制问题的精确罚函数 II：终端约束和逐点状态约束的精确惩罚。Optim. Control Appl. Methods，41：898–947，2020。[14] M. V. Dolgopolik。希尔伯特空间中约束优化问题的精确增广拉格朗日函数 I：理论。Optim.，73：1355–1395，2024。[15] M. Giaquinta 和 S. Hildebrandt。《变分法 I》。Springer-Verlag，Beglin，海德堡，2004。

[16] M. Gugat 和 E. Zuazua. 最优控制问题中终端约束的精确惩罚. Optim. Control Appl. Methods, 37:1329–1354, 2016.

[17] G. Leoni. Sobole夫空间入门. American Mathematical Society, Providence, RI, 2009.

[18] M. Tucsnak 和 G. Weiss. 算子半群的观测与控制. Birkh¨auser, Basel, 2009.

附录. 定理 1 的证明

我们将定理 1 的证明分为三个引理。

引理 1. 在定理 1 的假设下，存在 $c_* > 0$ 和一个有界集 $\Lambda \subset H \times \mathbb{R}^m$，使得对于任意 $c \geq c_*$，都有 $S_c(\gamma) \subset \Omega_c(\gamma+\varepsilon) \times \Lambda$。

证明。存在 $\widehat{c} > 0$，使得对于任意 $c \geq \widehat{c}$ 以及所有 $(x, \lambda, \mu) \in S_c(\gamma)$，均有 $x \in \Omega_c(\gamma+\varepsilon)$，这一事实直接由 [14, 引理 32] 得出，其证明方式与 [14, 定理 35] 的证明完全相同。

采用反证法，假设该引理的结论不成立。于是，对于任意 $n \in \mathbb{N}$，存在 $(x_n, \lambda_n, \mu_n) \in S_n(\gamma)$，使得 $\|\lambda_n\| + |\mu_n| \geq n$。正如在 [14, 定理 35], 的证明中所指出的，有

$$\eta(x, \lambda, \mu) = Q(x)[\lambda, \mu] + \langle Q_{1,\lambda}(x), \lambda \rangle + \langle Q_{1,\mu}(x), \mu \rangle + Q_0(x)$$

对于某些函数 $Q_{1,\lambda}(x)$、$Q_{1,\mu}(x)$ 以及 $Q_0(x)$，它们通过函数 f、g_i、F 及其一阶导数定义。因此，根据定理 1 的第二个假设，存在 $C > 0$，使得对于任意 $n \geq r$，均有

$$\|Q_{1,\lambda}(x_n)\| \leq C, \quad |Q_{1,\mu}(x)| \leq C, \quad |Q_0(x_n)| \leq C$$

因为 $x_n \in \Omega_n(\gamma + \varepsilon) \subseteq \Omega_r(\gamma + \varepsilon)$。因此，根据定理 1 的第三个假设，对于任意 $n \geq r$，有

$$\eta(x_n, \lambda_n, \mu_n) \geq \sigma\big(\|\lambda_n\|^2 + |\mu_n|^2\big) - C\big(\|\lambda_n\| + |\mu_n|\big) - C.$$

因此，利用 [14, 引理 3。2]可得

$$\begin{aligned}\mathscr{L}(x_n, \lambda_n, \mu_n, n) \geq f(x_n) &+ \frac{n}{2}\phi(\|F(x_n)\|^2) \\ &+ \frac{n}{2\psi(0)}\big|\max\{g(x_n), 0\}\big|^2 - \frac{1}{2n\phi'(0)} - \frac{1}{2}(1+m)\frac{\psi(0)}{n} \\ &+ \sigma\big(\|\lambda_n\|^2 + |\mu_n|^2\big) - C\big(\|\lambda_n\| + |\mu_n|\big) - C\end{aligned}$$

对所有 $n \geq r$均成立。因此，考虑到以下事实： $x_n \in \Omega_n(\gamma + \varepsilon)$，函数 f 在 $\Omega_n(\gamma + \varepsilon)$ 上有界，且 $\|\lambda_n\| + |\mu_n| \geq n$，可以得出$\mathscr{L}(x_n, \lambda_n, \mu_n, n) \to +\infty$ 当 $n \to \infty$ 时成立，这不可能，因为 $(x_n, \lambda_n, \mu_n) \in S_n(\gamma)$。 □

引理 2。在定理 1 的假设下，对于任意 $K > 0$，存在$c_* > 0$ 使得

$$\left\|\nabla \mathscr{L}(x, \lambda, \mu, c)\right\| \geq K\left(\|F(x)\| + \left|\max\left\{g(x), -\frac{p(x,\mu)}{c}\mu\right\}\right|\right) \tag{33}$$

对所有 $c \geq c_*$ 以及 $(x, \lambda, \mu) \in S_c(\gamma)$。

证明。首先注意，引理的结论与 [14,定理 4.1]的结论相同，唯一的区别在于，本引理中我们并未假设集合 $V = \Omega_r(\gamma + \varepsilon)$ 有界。我们的目的是证明，对该集合有界性的假设实际上是多余的。

事实上， [14, 定理 4.1] 的证明由三个辅助引理组成，即 [14,引理 4.2–4.4]。在 [14, 引理 4.2 和 4.3] 的证明中，并未用到关于集合 V 有界性的假设（只用到了 f、 g_i 和 F 在 V 上的有界性及其一阶和二阶导数）。因此，根据 [14,引理 4.3] ，对于任意 $K > 0$ 以及任意有界集 $\Lambda \subset H \times \mathbb{R}^m$ ，都存在 $c_* \geq r$，使得不等式 (33) 对所有 $c \geq c_*$ 及 $(x, \lambda, \mu) \in \Omega_c(\gamma + \varepsilon) \times \Lambda$均成立。因此，如有必要，可增大 c_*，并应用引理 1，从而得到不等式 (33) 对所有 $c \geq c_*$ 及 $(x, \lambda, \mu) \in S_c(\gamma)$ 均成立。

□

引理 3。定理 1 成立。

证明。设 $c_* > 0$ 来自引理 2，并令 $K = 1$ 和 $n \in \mathbb{N}$ 使得 $c_n \geq c_*$。根据我们的假设， $\mathscr{L}(\cdot, c_n)$ 在 $X \times H \times \mathbb{R}^m$ 上的某点 (x_n, λ_n, μ_n) 处取得全局最小值。由 [14, 命题 3， 1] 函数 $\mathscr{L}(\cdot, c_n)$ 在其有效定义域上 Fr´echet 可微，这意味着 $\nabla \mathscr{L}(x_n, \lambda_n, \mu_n, c_n) = 0$。因此，应用引理 2 可得，点 x_n 对问题 $(\mathcal{P})$ 是可行的，并且

$$\min_{(x,\lambda,\mu)} \mathscr{L}(x, \lambda, \mu, c_n) = \mathscr{L}(x_n, \lambda_n, \mu_n, c_n) = f(x_n) + \eta(x_n, \lambda_n, \mu_n) \geq f(x_n) \geq f_*,$$

其中， f_* 是问题 $(\mathcal{P})$ 的最优值。因此，根据 [14,引理 5， 2] 增广拉格朗日函数 $\mathscr{L}(x, \lambda, \mu, c)$ 是全局精确的，即定理 1 的第一和第二个陈述成立。

定理 1 第三和第四个陈述的证明与 [14, 定理 5.7] 的证明相同，唯一的区别在于需要应用引理 2，而不是 [14, 定理 4.1]。

□