arXiv:2305.01781v1 [math.OC] 2 May 2023

求解特殊结构非光滑微分包含的一种方法*

Alexander Fominyh[1]

摘要。本文研究一种特殊形式的微分包含。假设包含式右端集合的支撑函数可能包含有限个关于相坐标连续可微函数的最大值。要求找到一条满足给定边界条件的微分包含轨迹，同时该轨迹还需位于给定曲面上。这类问题出现在不连续系统的实际建模以及其他应用问题中。原问题被转化为一个变分问题。本文证明，所得待极小化泛函是超可微的，并以超微分的形式给出必要极小条件。随后，采用经典形式的超微分（或最速）下降法来寻找该泛函的驻点。此外，泛函的构造方式使得能够验证所构造的驻点是否确为问题的全局最小值。本文证明了所提方法的收敛性，并通过示例说明该方法。

2020年数学主题分类。34A60, 49J52, 49K05。

.

1. Introduction

微分包含是用于建立动力系统模型的有力工具。本文研究了具有某种特殊结构的微分包含。具体而言，微分包含右端集合的支撑函数包含有限个相坐标连续可微函数的最大值函数。这类微分包含源自右端不连续的微分方程系统（考察这类系统处于滑动模式时），也源自其他一些实际问题。因此，需要寻找一条这样的微分包含的轨迹，使其同时满足边界条件并位于“不连续”曲面上。

请注意，文献中的大多数方法仅考虑右端点自由的微分包含，并采用欧拉法、龙格–库塔法等经典方法，以及各种有限差分方法等（例如，参见 [1], [2], [3], [4], [5]）。关于微分包含差分方法的综述可见 [6]。还可参见论文 [7]，其中构造了一种求解微分包含边值问题的算法。另请参见一些关于微分包含问题中最优性条件的参考文献 [8], [9], [10], [11]

Keywords and phrases: 微分包含，支撑函数，超微分

* *The main results of this paper (Sections 6, 7, 8, 9) were obtained in IPME RAS and supported by Russian Science Foundation (grant 20-71-10032).*[1] 俄罗斯科学院机械工程问题研究所，俄罗斯联邦圣彼得堡市瓦西里岛大街东段61号，邮编199178；圣彼得堡大学，俄罗斯联邦圣彼得堡市大学滨河路7–9号，邮编199034

©

。这些论文考虑了形式相当一般的微分包含，其中既有相位约束的情形，也有非光滑和非凸的情形。另一方面，所列研究更多具有理论意义，其中一些结果似乎难以在实践中应用。

本文用于求解问题的方法基于作者先前研究中提出的思想。初始问题被化为一个变分问题，而由此得到的待最小化泛函是非光滑的（但仅为超可微）。因此，需要采用非光滑优化方法来研究这一问题。在论文 [12] 中，作者提出了一种求解经典非光滑（但仅为次微分）变分问题的方法，其思路是将相位轨迹及其导数视为独立变量（并通过特殊的惩罚项考虑这些变量之间的自然联系）。论文 [13], [14], [15] 通过相应的支撑函数，将微分包含化为变分问题。论文 [16], [17] 采用类似方法研究了其他一些非光滑最优控制与变分法问题。本文还借鉴了 V. F. Demyanov 非光滑最小化科学学派的一些思想。例如，书籍 [18] 论证了一种修正的次可微（最速）下降法，用于在有限维空间中最小化有限个连续可微函数的最大值。

本文介绍了不连续系统控制的一些基本事实，并运用了微分方程和滑模理论（见第 2 节）。众所周知，当系统状态处于滑模时，可以观察到模型的一些新的有用性质。例如，从最优控制理论中的某个准则来看，这类运动可能是最优的。滑模还用于稳定系统，以及消除不需要的外部扰动。在某些情况下，系统在滑模中的运动可以通过具有特殊结构的微分包含来描述，这也是本文考虑这类微分包含的动机。有关滑模所需的基础知识以及问题陈述的动机，可参阅文献 [19], [20], [21]。

2. Basic definitions and notations

本文采用以下记号。设 $C_n[0,T]$ 为定义在 $[0,T]$ 上的 n 维连续向量函数空间。另设 $P_n[0,T]$ 为定义在 $[0,T]$ 上的 n 维分段连续有界向量函数空间。我们还需要 $L_n^2[0,T]$，即定义在 $[0,T]$ n 上的平方可和向量函数空间。若 X 是某个赋范空间，则 $||\cdot||_X$ 表示其范数，X^* 表示与给定空间共轭的空间。

我们假设每条轨迹 $x(t)$ 都是分段连续可微向量函数。令$t_0 \in [0,T)$为向量函数 $x(t)$ 的不可微点，则为明确起见，我们假设 $\dot{x}(t_0)$ 是向量函数 $x(t)$ 在点 t_0 处的右导数。类似地，我们假设 $\dot{x}(T)$ 是向量函数 $x(t)$ 在点 T 处的左导数。现在，基于这些假设和记号，我们可以认为向量函数 $x(t)$ 属于空间 $C_n[0,T]$，而向量函数 $\dot{x}(t)$ 属于空间 $P_n[0,T]$。

对于任意集合 $F \subset R^n$，定义向量 $\psi \in R^n$ 的支撑函数为 $c(F, \psi) = \sup\langle f, \psi\rangle f \in F$

其中，$\langle a, b\rangle$ 是向量 a和 $b \in R^n$的数量积。记 S_n 为 R^n 中以原点为中心的单位球面，另记 $B_r(c)$为半径为 $r \in R$ 且中心为 $c \in R^n$的球。向量 $\mathbf{e_i}$、$i = 1$、n构成 R^n中的标准基。集合 P与 $Q \subset R^n$ 的和 $P + Q$ 是它们的闵可夫斯基和，而 λP与 $\lambda \in R$的闵可夫斯基积为 λP与 $\lambda \in R$；若 q是 R^n中的某个向量，则 $Pq = \{\langle p, q\rangle \mid p \in P\}$。记 0_n为某个 n维向量函数函数空间中的零元素，记 $\mathbf{0_n}$ 为空间 R^n中的零元素。令 E_m 为单位矩阵，$\mathbf{O_{n,m}}$ 为空间 $R^n \times R^m$中的零矩阵。若 $\varphi(x) = \min f_i(x)$，其中$i=1$、M。

$f_i(x) : R^n \to R, i = \overline{1, M}$ 是一些函数，那么如果 $i \in R(x_0) = \{i \in \{1, \ldots, M\} \mid f_i(x_0) = \varphi(x_0)\}$，我们就称函数 $f_i(x)$，$i \in \{1, \ldots, M\}$ 在点 $x_0 \in R^n$ 处是活跃的。

本文将同时使用有限维空间中函数的超微分和函数空间中泛函的超微分。尽管第二个概念是对第一个概念的推广，为方便起见，我们仍分别为这两种情形，以及本文所讨论的特定函数（泛函）、其变量和空间引入定义。

考虑带有标准范数的空间 $R^n \times R^n$。设 $d = [d_1, d_2] \in R^n \times R^n$ 为任意向量。假设在点 (x, z) 处存在这样一个凸紧集 $\partial h(x, z) \subset R^n \times R^n$，使得

$$\frac{\partial h(x,z)}{\partial d} = \lim_{\alpha \downarrow 0} \frac{1}{\alpha}\big(h(x + \alpha d_1, z + \alpha d_2) - h(x,z)\big) = \min_{w \in \overline{\partial} h(x,z)} \langle w, d \rangle. \tag{1}$$

在这种情况下，函数 $h(x,\ z)$ 在点 $(x,\ z)$ 处称为超可微，集合 $\partial h(x,\ z)$ 称为函数 $h(x,\ z)$ 在点 $(\overline{x},\ z)$ 处的超微分。

从表达式（1）可以看出，以下公式

$$h(x + \alpha d_1, z + \alpha d_2) = h(x,z) + \alpha \frac{\partial h(x,z)}{\partial d} + o(\alpha, x, z, d),$$

$$\frac{o(\alpha, x, z, d)}{\alpha} \to 0,\ \alpha \downarrow 0,$$

成立。

如果函数 $\varsigma(\xi)$ 在点 $\xi_0 \in R^\ell$ 处可微，则其在该点处的超微分可表示为形式

$$\overline{\partial} \varsigma(\xi_0) = \{\varsigma'(\xi_0)\} \tag{2}$$

其中，$\varsigma'(\xi_0)$ 是函数 $\varsigma(\xi)$ 在点 ξ_0 处的梯度。还需注意，有限个超可微函数之和的超微分等于各被加项的超微分之和，即若函数 $\varsigma_k(\xi), k = 1, r$ 在点 $\xi_0 \in R^\ell$ 处超可微，则函数 $\varsigma(\xi) = \sum_{k=1}^{r} \varsigma_k(\xi)$ 在该点的超微分按公式计算。

$$\overline{\partial} \varsigma(\xi_0) = \sum_{k=1}^{r} \overline{\partial} \varsigma_k(\xi_0). \tag{3}$$

考虑带有范数 $L_n^2[0,T] \times L_n^2[0,T]$ 的空间 $C_n[0,T] \times P_n[0,T]$。令 $g = [g_1, g_2] \in C_n[0,T] \times P_n[0,T]$ 为任意向量函数。假设在点 $(x,\ z)$ 处存在这样的凸弱紧集 $\partial I(x,\ z) \subset \big(C_n[0,T] \times P_n[0,T],\ \ ||\cdot||_{L_n^2[0,T] \times L_n^2[0,T]}\big)^*$，使得

$$\frac{\partial I(x,z)}{\partial g} = \lim_{\alpha \downarrow 0} \frac{1}{\alpha}\big(I(x + \alpha g_1, z + \alpha g_2) - I(x,z)\big) = \min_{w \in \overline{\partial} I(x,z)} w(g). \tag{4}$$

在这种情况下，泛函 $I(x,\ z)$ 称为在点 $(x,\ z)$ 处超可微，而集合 $\partial I(x,\ z)$ 称为泛函 $I(x,\ z)$ 在点 $(\overline{x},\ z)$ 处的超微分。

由表达式（4）可知，下列公式

$$I(x + \alpha g_1, z + \alpha g_2) = I(x,z) + \alpha \frac{\partial I(x,z)}{\partial g} + o(\alpha, x, z, g),$$

$$\frac{o(\alpha, x, z, g)}{\alpha} \to 0,\ \alpha \downarrow 0,$$

成立。

3. Statement of the problem

考虑微分包含

$$\dot{x}_i \in A_i x + [\underline{a}_i, \overline{a}_i]|x|[-1,1] = A_i x + [-\overline{a}_i, \overline{a}_i]|x|, \quad i = \overline{1,n}, \tag{5}$$

带有初始点

$$x(0) = x_0 \tag{6}$$

并达到所需的终点

$$x_j(T) = x_{Tj}, \quad j \in J, \tag{7}$$

在公式 (5) 中， A_i 是常数 $n \times n$ 矩阵 A的第 i行， $i = 1, n$，并且 a_i、 a_i、$\overline{i} = 1, \underline{n}$是给定的非负数，且 $a_i = a_i = 0$、 $i = m+1, n$。系统在给定的有限时间区间 $[0,T]$上进行研究。我们假设， $x(t)$是相坐标的 n维连续向量函数，在区间 $[0,T]$上具有分段连续且有界的导数。在公式 (6) 中， $x_0 \in R^n$ 是给定向量；在公式 (7) 中， x_{Tj} 是与状态向量中在右端点处固定的那些坐标相对应的给定数，此处 $J \subset \{1,\dots,n\}$ 是给定的指标集合。

令 $F(x) = (f_1(x)\dots, f_n(x))'$，其中 $f_1(x)\dots, f_n(x)$ 遍历包含式（5）右端相应的集合 $F_1(x)\dots F_n(x)$，因此我们可以将给定的包含式改写为如下形式

$$\dot{x} \in F(x). \tag{8}$$

我们还考虑曲面

$$s(x) = \mathbf{0_m} \tag{9}$$

其中， $s(x)$ 是已知的连续可微的 m 维向量函数。

我们将问题表述如下：要求寻找一条轨迹 $x^* \in C_n[0,T]$（其导数为 $\dot{x}^* \in P_n[0,T]$），使其沿曲面 (9) 运动，在 $t \in [0,T]$ 时满足微分包含 (8)，并满足边界条件 (6)、(7)。假设这样的解存在。

让我们讨论一个实 际问题，它引出了上述问题表述。考虑如下系统

$$\dot{x} = Ax + Bu \tag{10}$$

并满足边界条件 (6)、(7)。在公式 (10) 中， A 是常数 $n \times n$ 矩阵， B 是常数 $n \times m$ 矩阵。为简单起见，我们假设 $B = [E_m, \mathbf{O_{n-m}}, \mathbf{m}]$。该系统在给定的有限时间区间 $[-t^*, T]$ 上考察（其中， T 是给定的终端时刻；参见下文关于时刻 t^* 的说明）。我们假设 $x(t)$ 是一个关于相坐标的 n 维连续向量函数，其导数在 $[-t^*, T]$ 上分段连续且有界； m 维控制 u 的结构将在下文说明。

另外，给定“不连续点”曲面 (9)。

考虑以下形式的控制：

$$u_i = -\alpha_i |x| \operatorname{sign}(s_i(x)), \quad i = \overline{1,m} \tag{11}$$

其中 $\alpha_i \in [\underline{a_i}, \overline{a_i}]$, $i = \overline{1,m}$是一些正数，有时称为增益因子。

在书籍 [20] 中可以看到，如果曲面 (9) 是一个超平面，那么在自然假设下，且因子 α_i, $i = 1,m$ 的取值足够大时，控制 (11) 能确保系统 (10) 从任意初始状态 (6) 出发，在有限时间 t^* 内到达该曲面的小邻域$s(x) = \mathbf{0_m}$，并在此后继续停留在该邻域内，同时满足条件 $s_i(x(t)) \Longrightarrow 0$, $i = 1,m$，在 $t \to \infty$ 下；也就是说，控制 (11) 能确保系统 (10) 的滑动模态具有“大范围”稳定性。在 [20] 中还可以找到关于时刻 t^* 的估计。这里假设所有必要条件均已满足，且数值 a_i, a_i, $i = 1,m$ 均取足够大。现在我们关注系统在“不连续”曲面上的行为（在时间区间 $[0,T]$ 上）。

我们可以看到，系统 (10) 中带有控制 (11) 的前 m 个微分方程的右端在曲面 $s_i(x) = 0,\ i = 1,m$ 上不连续。因此，必须采用已知的不连续系统解定义之一。

让我们采用这种定义的一种经典变体 [19], [23]。其要点如下。在有限时间区间 $[0,T]$ 上，考虑系统 $\dot{x} = f(x,u_1(x,t)\ldots,u_m(x,t)\ t)$，其中向量函数 $f(x,u_1,\ldots,u_m,t)$关于其所有变量连续，而向量函数 $u_i(x,t)$、 $i = 1,m$ 分别在集合 $s_i(x) = 0,\ i = 1,m$上不连续。在向量函数 $u_i(x,t),\ i = 1,m$的不连续点(x,t)处，必须定义一个闭集 $U_i(x,t),\ i = 1,m$。它是函数$f(x,u_1,\ldots,u_m,t)$的变量 u_i可能取值的集合。记 $F(x,t) = f(x,u_1,...,u_m,t)$ 为函数 $f(x,u_1,...,u_m,t)$在固定变量 x,t下的取值集合，其中 $u_1,\ldots,u_m$ 分别遍历集合$U_1(x,t)\ldots,U_m(x,t)$。于是，将该微分包含的解视为右端不连续的原微分方程的解。在物理系统中，集合 $U_i(x,t),\ i = 1,m$通常对应于不同的模块，并假定为凸集。在向量函数 $u_i(x,t),\ i = 1,m$的不连续点(x,t)处，集合 $U_i(x,t)$还必须包含所有序列 $v_k \in U_i(x_k,t_k)$的所有极限点，其中 $x_k \to x$ ，且当 $k \to \infty$时 $t_k \to t$ 。如果控制具有形式(11)，那么自然可以将 $U_i(x,t) = \mathrm{co}\{-a_i|x|,a_i|x|\},\ i = 1,m$视为这类集合。

因此，根据此定义，所考虑的不连续系统的解满足上述微分包含（5）。请注意， [23]中给出了更为详细的定义版本。该定义形式严格且相当复杂，因此这里不讨论其细节。就我们的目的而言，只需规定：按定义，采用控制（11）的系统（10）的解就是包含式（5）的解。

注3.1. 除了取自空间 $C_n[0,T]$ 且导数取自空间 $P_n[0,T]$ 的轨迹外，也可以考虑区间 $[0,T]$ 上的绝对连续轨线，其在 $[0,T]$ 上具有可测且几乎处处有界的导数；对于微分包含而言，这样的处理更为自然。本文选择该解空间，是因为它有可能被实际构造出来。

4. 化简为变分问题

为简洁起见，有时我们写作 F，而不写 $F(x)$。由于 $\forall x \in R^n$ 集合 $F(x)$ 是 R^n中的凸紧集，因此包含式（8）可改写如下 [24]:

$$\dot{x}_i(t)\psi_i(t) \le c(F_i(x(t)),\psi_i(t)) \quad \forall\psi_i(t) \in S_1, \quad \forall t \in [0,T], \quad i = \overline{1,n}.$$

计算集合 F_i 的支撑函数。为此注意，集合 F_i 是一个一维“球”，其中心为

$$c_i(x) = A_i x, \quad i = \overline{1,n},$$

且其“半径”为

$$r_i(x) = \bar{a}_i|x|, \quad i = \overline{1,m},$$
$$r_i(x) = 0, \quad i = \overline{m+1,n}.$$

因此，集合 F_i 的支撑函数可由公式 [24] 表示

$$c(F_i(x),\psi_i) = \psi_i A_i x + \bar{a}_i|x||\psi_i|, \quad i = \overline{1,m},$$

$$c(F_i(x),\psi_i) = \psi_i A_i x, \quad i = \overline{m+1,n}.$$

我们看到，当$x_i \ne 0$、 $i = 1,n$时，集合 F_i 的支撑函数在相坐标 x 中连续可微。

记 $z(t) = \dot{x}(t), \quad z \in P_n[0,T]$，则由（6）可得

$$x(t) = x_0 + \int_0^t z(\tau)d\tau. \tag{12}$$

现在让我们实现以下思路。“强制地”将点 z 和 x 视为“独立”变量。事实上，这些变量之间存在上述关系（12）（这自然意味着向量函数 $z(t)$ 是向量函数 $x(t)$ 的导数），因此，让我们使用泛函来考虑这一约束。

$$\upsilon(x,z)=\frac{1}{2}\int_0^T\left(x(t)-x_0-\int_0^t z(\tau)d\tau\right)^2dt. \tag{13}$$

此外可以看出，当 $\upsilon(x,z)=0$ 时，左端点上的条件 (6) 也成立。对于 $i=1,n$ ，令

$$\ell_i(\psi_i,x,z)=\langle z_i,\psi_i\rangle-c(F_i(x),\psi_i),$$
$$h_i(x,z)=\max_{\psi_i\in S_1}\max\{0,\ell_i(\psi_i,x,z)\},$$

然后令

$$h(x,z)=(h_1(x,z),\dots,h_n(x,z))'$$

并构造泛函

$$\varphi(x,z)=\frac{1}{2}\int_0^T h^2\big(x(t),z(t)\big)dt. \tag{14}$$

不难验证，对于泛函 (14)，关系

$$\begin{cases}\varphi(x,z)=0, & \text{if } \dot{x}_i(t)\psi_i(t)\le c(F_i(x(t)),\psi_i(t)) \quad \forall\psi_i(t)\in S_1, \quad \forall t\in[0,T], \quad i=\overline{1,n}.\\ \varphi(x,z)>0, & \text{otherwise,}\end{cases}$$

成立，即包含关系 (8) 成立当且仅当 $\varphi(x,z)=0$。

引入泛函

$$\chi(z)=\frac{1}{2}\sum_{j\in J}\left(x_{0j}+\int_0^T z_j(t)dt-x_{Tj}\right)^2. \tag{15}$$

我们看到，若 $\upsilon(x,\ z)=0$，则右端点处的条件 (7) 成立当且仅当 $\chi(z)=0$。引入泛函

$$\omega(x)=\frac{1}{2}\int_0^T s^2\big(x(t)\big)dt. \tag{16}$$

显然，轨迹 $x(t)$ 在每个 $t\in[0,T]$ 处都属于曲面 (9)，当且仅当 $\omega(x)=0$。最后构造泛函

$$I(x,z)=\varphi(x,z)+\chi(z)+\omega(x)+\upsilon(x,z). \tag{17}$$

因此，原问题已化为在空间 $C_n[0,T]\times P_n[0,T]$ 上极小化泛函 (17)。设 z^* 为该泛函的一个全局极小值点。则

$$x^*(t)=x_0+\int_0^t z^*(\tau)d\tau$$

是初始问题的一个解。

注4.1。泛函 $\varphi(x,\ z)$ 的结构是自然的，因为在每个固定的 $t\in[0,T]$ 处，值 $h_i(x(t)\ z(t))$，$i=\overline{1,\ n}$，是点 $z_i(t)$ 到集合 $F_i(x(t))$ 的欧几里得距离；泛函 (14) 是轨迹 $z_i(t)$ 与集合 $F_i(x)$、$i=1,\ n$ 在 $L_n^2[0,T]$ 范数下偏差平方和的一半。泛函 (13)、(15)、(16) 的结构含义显而易见。

尽管泛函 $I(x,z)$ 自变量的维数 n 大于初始问题的维数（即所求点 x^* 的维数），但如后文所示，它的超微分结构（在赋范空间$C_n[0,T]\times P_n[0,T]$ 中，范数为 $L_n^2[0,T]\times L_n^2[0,T]$）相当简单。这一事实将使我们能够构造一种数值方法来求解原始问题。

5. 泛函 Minimum conditions of the functional $I(x,\ z)$ 的极小条件

本节提及超微分演算规则 (2)、(3) 时，指的是它们在泛函空间 [26] 中已知的类似规则。

利用经典变分，很容易证明泛函 $\chi(z)$ 的 Gâteaux 可微性，我们有

$$\nabla\chi(z)=\sum_{j\in J}\left(x_{0j}+\int_0^T z_j(t)dt-x_{Tj}\right)\mathbf{e_j}.$$

根据超微分演算规则 (2)，可得

$$\overline{\partial}\chi(z)=\left\{\sum_{j\in J}\left(x_{0j}+\int_0^T z_j(t)dt-x_{Tj}\right)\mathbf{e_j}\right\}. \tag{18}$$

利用经典变　　分，很容易证明泛函 $\omega(x)$ 的 Gâteaux 可微性，我们有

$$\nabla\omega(x)=\sum_{i=1}^{m}s_i\big(x(t)\big)\frac{\partial s_i\big(x(t)\big)}{\partial x}.$$

根据超微分演算规则 (2)，可得

$$\overline{\partial}\omega(x)=\left\{\sum_{i=1}^{m}s_i\big(x(t)\big)\frac{\partial s_i\big(x(t)\big)}{\partial x}\right\}. \tag{19}$$

利用经典变分和分部积分，也不难检验泛函 $\upsilon(x,z)$ 的Gâteaux可微性，我们得到

$$\nabla\upsilon(x,z,t)=\begin{pmatrix}x(t)-x_0-\int_0^t z(\tau)d\tau\\ -\int_t^T\left(x(\tau)-x_0-\int_0^\tau z(s)ds\right)d\tau\end{pmatrix}.$$

根据超微分演算规则 (2)，可得

$$\overline{\partial}\upsilon(x,z)=\left\{\begin{pmatrix}x(t)-x_0-\int_0^t z(\tau)d\tau\\ -\int_t^T\left(x(\tau)-x_0-\int_0^\tau z(s)ds\right)d\tau\end{pmatrix}\right\}. \tag{20}$$

探索泛函 $\varphi(x,\ z)$ 的微分性质。为此，首先给出以下公式，用于计算点 $(x,\ z)$ 处的超微分 $\partial h^2(x,\ z)$。在 $i=1,\ m$ 处有

$$\overline{\partial}\left(\tfrac{1}{2}\,h_i^2(x,z)\right)=h_i(x,z)\left(\psi_i^*\mathbf{e_{i+n}}-\psi_i^*[A_i',\mathbf{0_n}]+\sum_{j=1}^{n}\overline{\partial}\left(-\overline{a}_i|x_j||\psi_i^*|\right)\right), \tag{21}$$

其中在 $j=\overline{1,\ n}$ 我们有

$$\overline{\partial}(-\overline{a}_i|x_j||\psi_i^*|)=\begin{cases}-\overline{a}_i|\psi_i^*|\mathbf{e_j}, & \text{if } x_j>0,\\ \overline{a}_i|\psi_i^*|\mathbf{e_j}, & \text{if } x_j<0,\\ \mathrm{co}\big\{-\overline{a}_i|\psi_i^*|\mathbf{e_j},\overline{a}_i|\psi_i^*|\mathbf{e_j}\big\}, & \text{if } x_j=0.\end{cases}$$

在 $i=\overline{m+1,\ n}$ 有

$$\overline{\partial}\left(\tfrac{1}{2}\,h_i^2(x,z)\right)=h_i(x,z)\big(\psi_i^*\mathbf{e_{i+n}}-\psi_i^*[A_i',\mathbf{0_n}]\big).$$

在所给出的公式中，ψ_i^*的取值满足 $\max\{0,\ell_i(\psi_i^*(x,\ z)\ x,\ z)\}=\max_{\psi_i\in S_1}\max\{0,\ell_i(\psi_i,\ x,\ z)\}$，$i=1,\ n$。文献 [16] 表明，如果 $h_i(x,\ z)>0$，那么 $\psi_i^*(x,\ z)$ 是唯一的，并且关于 $(x,\ z)$ 连续。文献 [25] 表明，泛函 $\varphi(x,\ z)$ 是超可微的，并且其超微分由相应的被积函数的超微分确定。

定理 5.1。区间 $[0,T]$ 可以划分为有限个区间，在每个区间内，每条相轨迹要么恒等于零，要么保持某一确定的符号。于是，泛函 $\varphi(x,z)$ 是超可微的，即

$$\frac{\partial\varphi(x,z)}{\partial g}=\lim_{\alpha\downarrow 0}\frac{1}{\alpha}\big(\varphi(x+\alpha g_1,z+\alpha g_2)-\varphi(x,z)\big)=\min_{w\in\overline{\partial}\varphi(x,z)}\int_0^T\langle w(t),g(t)\rangle dt,$$

其中 $g=[g_1,g_2]\in C_n[0,T]\times P_n[0,T]$，集合 $\partial\varphi\ (x,z)$ 定义如下

$$\overline{\partial}\varphi(x,z)=\Big\{w=[w_1,w_2]\in L_n^\infty[0,T]\times L_n^\infty[0,T]\ \Big|\tag{22}$$

$$[w_1(t),w_2(t)]\in\overline{\partial}\left(\tfrac{1}{2}\,h^2(x(t),z(t))\right)\quad for\ a.e.\ t\in[0,T]\Big\}.$$

利用得到的公式 (17)、(18)、(19)、(20)、(22) 以及超微分演算规则 (3)，可得在点 $(x,\ z)$ 处计算泛函 $I(x,\ z)$ 的超微分的最终公式

$$\overline{\partial}I(x,z)=\overline{\partial}\varphi(x,z)+\overline{\partial}\overline{\chi}(x,z)+\overline{\partial}\overline{\omega}(x,z)+\overline{\partial}\upsilon(x,z),\tag{23}$$

其中，形式上 $\chi(x,\ z):=\chi(z)$，$\overline{\omega}(x,\ z):=\omega(x)$。

根据泛函 $I(x,\ z)$ 在点 $(x^*,\ z^*)$ 处基于超微分的已知极小条件 [26]，可知下述定理成立。

定理 5.2。设区间 $[0,T]$ 可划分为有限个区间，在每个区间内，每条相位轨迹 $x_i^*\ (t)$，$i=1,n$，要么恒等于零，要么保持某一确定的符号。为了使点 (x^*,z^*) 使泛函 $I\ (x,z)$ 达到极小，必须满足

$$\overline{\partial}I(x^*(t),z^*(t))=\{\mathbf{0_{2n}}\}\tag{24}$$

在每个 $t\in[0,T]$ 处，超微分 $\partial I(x,z)$ 的表达式由公式 *(23)* 给出。若满足 $I(x^*,z^*)$，则条件 *(24)* 也是充分的。

定理 5.2 已经包含一个构造性的极小条件，因为以此为基础可以构造最速（超微分）下降方向，而且对于构造过程中出现的每个子问题，都有已知的高效算法可供求解（见下文第 6 节）。构造出最速下降方向后，就可以利用它来应用某些非光滑优化方法（见下文第 7 节）。再次注意，如果没有所提出的对变量 x 和 z 的“分离”，所考虑泛函的超微分将具有非常复杂的结构。这将使得构造该泛函的超微分下降方向成为一个极其困难（而且看来实际上不可能）的问题。

6. A more general case of the set $F(x)$ support function

现在考虑一个更一般的情形，其中相应集合 $F_i(x)$、$i=1,n$ 的支撑函数具有如下形式

$$c(F_i(x),\psi_i)=\psi_i A_i x+\sum_{j=1}^{r}\overline{a}_{ij}\max\left\{f_{i,j_1}(x)\psi_i,\ldots,f_{i,j_{k(j)}}(x)\psi_i\right\},$$

其中 $f_{i,j_1}(x)\ldots,f_{i,j_{k(j)}}(x)$、$i=1,n$、$\overline{j=1,r}$（为简化表述，我们假设 r 对每个 $i=1,n$ 都相同）是连续可微函数，而 a_{ij}、$i=\overline{1,n}$、$j=1,r$ 是一些非负数。（我们仍然只考虑在每个 $x\in R^n$ 处的凸紧集合 $F_1(x)\ldots F_n(x)$。）

由于有 $|x_i||\psi_i|=|x_i\psi_i|=\max\{x_i\psi_i,-x_i\psi_i\}$，这一情形确实更为一般。

还可以给出一个导出此类系统的实际问题。根据某些物理学方面的考虑，设物体的"速度" $\dot{x}_1$ 位于"坐标" x_1、x_2、x_3 的范围 $[\min\{x_1,x_2,x_3\},\max\{x_1,x_2,x_3\}]$ 内。3所给出的线段可写为 $\mathrm{co}\{x_1,x_2,x_3\}=\mathrm{co}\bigcup\{x_i\}$。该集合的支撑函数为 $[24]_{i=1}\max\{x_1\psi_1,x_2\psi_1,x_3\psi_1\}$。

为简单起见，考虑 $n=2$, $r=1$, $k(1)=2$ 的情形，且仅考虑函数 $\ell_1(\psi_1,x,z)$ 和 $h_1(x,z)$（这里分别记为 $\ell(\psi_1,x_1,x_2,z_1)$ 和 $h(x_1,x_2,z_1)$），以及非零长度的时间区间 $[t_1,t_2]\subset[0,T]$；一般情形可用类似方法处理。于是有 $\ell(\psi_1,x_1,x_2,z_1)=z_1\psi_1-a_1x_1\psi_1-a_2x_2\psi_1-b\max\{f_1(x)\psi_1,f_2(x)\psi_1\}$，其中 $a_1:=a_{1,1}$, $a_2:=a_{1,2}$, $b:=a_{1,1}>0$, $f_1:=f_{1,11}$, $f_2:=f_{1,12}$。固定某个点 $(x_1,x_2)\in R^2$。设 $f_1(x(t))\psi_1(t)=f_2(x(t))\psi_1(t)$ 于 $t\in[t_1,t_2]$；其他情形也可完全类似地研究。

a) 假设 $h_1(x,z)>0$，即 $h_1(x,z)=\max_{\psi_1\in S_1}\ell_1(\psi_1,x,z)>0$。

我们的目标是应用 [27]中关于方向可微性的相应定理。该书中的定理针对下确界函数，因此我们将把该定理应用于函数 $-\ell(\psi_1,x_1,x_2,z_1)$。为此，请检查函数 $h(x_1,x_2,z_1)$是否满足以下条件：

i) 函数 ℓ (ψ_1,x_1,x_2,z_1) 在 $S_1\times R\times R$ 上连续；[2]

ii) 存在一个数 β和一个紧集 $C\in R$，使得对于点 $(x_1,x_2,z_1)'$ 附近的每个$(\overline{x_1},x_2,z_1)'$，水平集

$$lev_\beta(-\ell(\cdot,\overline{x}_1,\overline{x}_2,\overline{z}_1))=\{\psi_1\in S_1\mid-\ell(\psi_1,\overline{x}_1,\overline{x}_2,\overline{z}_1)\le\beta\}$$

非空且包含于集合 C 中；

iii) 对任意固定的 $\psi_1\in S_1$，函数 $\ell(\psi_1,\cdot,\cdot,\cdot)$ 在点 $(x_1,x_2,z_1)'$ 处方向可微；iv) 若 $d=[d_1,d_2]\in\underline{R^2}\times R$、$\gamma_n\downarrow0$ 和 ψ_{1_n} 是 C 中的一个序列，则 ψ_{1_n} 有一个极限点 ψ_1，使得

$$\limsup_{n\to\infty}\frac{-\ell(\psi_{1_n},x+\gamma_nd_1,z+\gamma_nd_2)-(-\ell(\psi_{1_n},x,z))}{\gamma_n}\ge\frac{\partial(-\ell(\overline{\psi}_1,x,z))}{\partial d}$$

其中，$\underline{\partial\ell(\overline{\psi}_1,x_1,x_2,z_1)\,\partial d}$ 是函数 $\ell(\psi_1,x_1,x_2,z_1)$ 在点 $(x_1,x_2,z_1)'$ 沿方向 d 的导数。条件 i)、ii) 的验证显而易见。

为验证条件 iii)，只需注意到，由于 $b>0$，对于固定的 $\psi_1\in S_1$，函数 $-b\max\{f_1(x)\psi_1,f_2(x)\psi_1\}$ 在点 $(x_1,x_2)'$ 处超可微（因此方向可微）；此外，它在该点的超微分为 $\underline{b\,\mathrm{co}\left\{\left(\psi_1\frac{\partial f_1}{\partial x_1}(x),\psi_1\frac{\partial f_1}{\partial x_2}(x)\right)',\left(\psi_1\frac{\partial f_2}{\partial x_1}(x),\psi_1\frac{\partial f_2}{\partial x_2}(x)\right)'\right\}}$。该函数在点 $(x_1,x_2)'$ 沿方向 d_1 的导数的显式表达式为 $\underline{-b\max\left\{\left\langle\psi_1\frac{\partial f_1}{\partial x}(x),d_1\right\rangle,\left\langle\psi_1\frac{\partial f_2}{\partial x}(x),\overline{d_1}\right\rangle\right\}}$。

最后，检查条件　　关于 iv)。令 $[d_1, d_2] \in R^2 \times R$, $\gamma_n \downarrow 0$，且 ψ_{1_n} 是集合 C 中的某个序列。计算

$$\limsup_{n\to\infty} \frac{-\ell(\psi_{1_n}, x_1+\gamma_n d_{1,1}, x_2+\gamma_n d_{1,2}, z_1+\gamma_n d_2) - (-\ell(\psi_{1_n}, x_1, x_2, z_1))}{\gamma_n} =$$

$$= \limsup_{n\to\infty} \frac{1}{\gamma_n}\Big(-(z_1+\gamma_n d_2)\psi_{1_n} + a_1(x_1+\gamma_n d_{1,1})\psi_{1_n} + a_2(x_2+\gamma_n d_{1,2})\psi_{1_n} +$$
$$+ b\max\left\{f_1(x_1+\gamma_n d_{1,1}, x_2+\gamma_n d_{1,2})\psi_{1_n}, f_2(x_1+\gamma_n d_{1,1}, x_2+\gamma_n d_{1,2})\psi_{1_n}\right\} +$$
$$+ z_1\psi_{1_n} - a_1 x_1\psi_{1_n} - a_2 x_2\psi_{1_n} - b\max\left\{f_1(x_1,x_2)\psi_{1_n}, f_2(x_1,x_2)\psi_{1_n}\right\}\Big) =$$

$$= \limsup_{n\to\infty} \frac{1}{\gamma_n}\Big(-\gamma_n d_2\psi_{1_n} + \gamma_n a_1 d_{1,1}\psi_{1_n} + \gamma_n a_2 d_{1,2}\psi_{1_n} +$$
$$+ b\max\left\{\left[f_1(x) + \gamma_n\left\langle\frac{\partial f_1(x)}{\partial x}, d_1\right\rangle + o_1(\gamma_n, x, d)\right]\psi_{1_n}, \left[f_2(x) + \gamma_n\left\langle\frac{\partial f_2(x)}{\partial x}, d_1\right\rangle + o_2(\gamma_n, x, d)\right]\psi_{1_n}\right\} -$$
$$- b\max\left\{f_1(x_1,x_2)\psi_{1_n}, f_2(x_1,x_2)\psi_{1_n}\right\}\Big) \geq$$

$$\geq \limsup_{n\to\infty} \frac{1}{\gamma_n}\Big(-\gamma_n d_2\psi_{1_n} + \gamma_n a_1 d_{1,1}\psi_{1_n} + \gamma_n a_2 d_{1,2}\psi_{1_n} +$$
$$+ \gamma_n b\max\left\{\left\langle\frac{\partial f_1(x)}{\partial x}, d_1\right\rangle\psi_{1_n}, \left\langle\frac{\partial f_2(x)}{\partial x}, d_1\right\rangle\psi_{1_n}\right\} + b\min\left\{o_1(\gamma_n, x, d)\psi_{1_n}, o_2(\gamma_n, x, d)\psi_{1_n}\right\}\Big)$$

其中 $o_q(\gamma_n, x, d)/\gamma_n \to 0$, $\gamma_n \downarrow 0$, $q = \overline{1, 2}$，最后一个不等式由所作的假设 $f_1(x)\psi_1 = f_2(x)\psi_1$（在所考虑的时间区间上）以及满足该假设时两个函数的最大值的相应性质得出。

设 $\overline{\psi_1}$ 是一个 序列 ψ_{1_n} 的极限点。根据方向导数的定义，我们有

$$\frac{\partial(-\ell(\overline{\psi}_1, x, z))}{\partial d} = -d_2\overline{\psi}_1 + a_1 d_{1,1}\overline{\psi}_1 + a_2 d_{1,2}\overline{\psi}_1 + b\max\left\{\left\langle\overline{\psi}_1\frac{\partial f_1(x)}{\partial x}, d_1\right\rangle, \left\langle\overline{\psi}_1\frac{\partial f_2(x)}{\partial x}, d_1\right\rangle\right\}.$$

由最后两个关系可得，条件 iv) 得到满足。

因此，函数 $h(x_1, x_2, z_1)$ 满足条件 i)-iv)，故在点 (x_1, x_2, z_1) [27], 处方向可微，且其在该点沿方向 d 的导数由公式表示为

$$\frac{\partial h(x_1, x_2, z_1)}{\partial d} = \sup_{\psi_1 \in \mathcal{S}(x_1, x_2, z_1)} \frac{\partial \ell(\psi_1, x_1, x_2, z_1)}{\partial d}$$

其中 $\mathcal{S}(x_1, x_2, z_1) =$最大值点$_{\psi_1 \in S_1}\ell(\psi_1, x_1, x_2, z_1)$。然而，如上所示，在所考虑的问题中，集合 $\mathcal{S}(x_1, x_2, z_1)$ 仅由一个元素 $\psi_1^*(x_1, x_2, z_1)$ 构成，因此

$$\frac{\partial h(x_1, x_2, z_1)}{\partial d} = \frac{\partial \ell(\psi_1^*(x_1, x_2, z_1), x_1, x_2, z_1)}{\partial d}.$$

最后，回顾方向导数的定义可知，等式

$$\frac{\partial(-\ell(\psi_1^*, x, z))}{\partial d} = -d_2\psi_1^* + a_1 d_{1,1}\psi_1^* + a_2 d_{1,2}\psi_1^* + b\max\left\{\left\langle\psi_1^*\frac{\partial f_1(x)}{\partial x}, d_1\right\rangle, \left\langle\psi_1^*\frac{\partial f_2(x)}{\partial x}, d_1\right\rangle\right\}$$

其中我们记 $\psi_1^* := \psi_1^*(x_1, x_2, z_1)$。

由前面两个表达式，我们最终得到函数 $h(x_1,\ x_2,\ z_1)$ 在点 $(x_1,\ x_2,\ z_1)'$ 处超可微；但在所考虑的情形下它也是正函数，因此函数 $h^2(x_1,\ x_2,\ z_1)$ 在点 $(x_1,\ x_2,\ z_1)'$ 处同样超可微，并且是一个超可微正函数的平方（见 [28]）。

b) 在 $h_1(x,z)\ =0$ 的情况下，显然，函数 $h^2(x_1,x_2,z_1)$ 在点 $(x_1,x_2,z_1)'$ 处可微，且其梯度在该点处为零。

现在来探讨泛函 $\varphi(x,\ z)$ 的微分性质。为此，首先给出以下公式，用于计算点 $(x,\ z)$ 处的超微分 $\partial h^2(x,\ z)$。当 $i=1,\ n$ 时，有

$$\overline{\partial}\left(\tfrac{1}{2}\,h_i^2(x,z)\right)=h_i(x,z)\left(\psi_i^*\mathbf{e_{i+n}}-\psi_i^*[A_i',\mathbf{0_n}]+\sum_{j=1}^{r}\overline{\partial}\left(-\overline{a}_{ij}\max\left\{f_{i,j_1}(x)\psi_i^*,\ldots,f_{i,j_{k(j)}}(x)\psi_i^*\right\}\right)\right)\tag{25}$$

并且在 $j=\overline{1,r}$ 我们有

$$\overline{\partial}\left(-\overline{a}_{ij}\max\left\{f_{i,j_1}(x)\psi_i^*,\ldots,f_{i,j_{k(j)}}(x)\psi_i^*\right\}\right)=\overline{a}_{ij}\ \mathrm{co}\left\{\left[\psi_i^*\frac{\partial f_{i,j_p}(x)}{\partial x},\mathbf{0_n}\right]\right\},\quad j_p\in R_{ij}(x),\tag{26}$$

$$R_{ij}(x)=\left\{j_p\in\{j_1,\ldots,j_{k(j)}\}\ \middle|\ f_{i,j_p}(x)\psi_i^*=\max\left\{f_{i,j_1}(x)\psi_i^*,\ldots,f_{i,j_{k(j)}}(x)\psi_i^*\right\}\right\}.$$

计算函数 $\frac{1}{2}\,h_i^2(x,z)$ 沿方向 $g\ =\ [g_1,g_2]\in\ R^n\times R^n$ 的导数。根据公式 (25)、(26) 以及超微分演算规则 [18]，我们有

$$\frac{\partial\left(\frac{1}{2}\,h_i^2(x,z)\right)}{\partial g}=h_i(x,z)\left(\psi_i^*g_{2i}-\langle\psi_i^*A_i,g_1\rangle+\sum_{j=1}^{r}\min_{j_p\in R_{ij}(x)}\left\langle-\overline{a}_{ij}\psi_i^*\frac{\partial f_{i,j_p}(x)}{\partial x},g_1\right\rangle\right).\tag{27}$$

证明泛函 $\varphi(x,z)$ 是超可微的，并且其超微分由相应的被积函数的超微分确定。

定理 6.1. 设区间 $[0,T]$ 可划分为有限个区间，在每个区间内，函数 $\left\{f_{i,j_1}(x)\psi_i,\ldots,f_{i,j_{k(j)}}(x)\psi_i\right\},\ i=1,n,$ $j=1,r$ 中的一个（或多个）活跃。则泛函$\varphi(x,z)$ 是超可微的，即

$$\frac{\partial\varphi(x,z)}{\partial g}=\lim_{\alpha\downarrow0}\frac{1}{\alpha}\big(\varphi(x+\alpha g_1,z+\alpha g_2)-\varphi(x,z)\big)=\min_{w\in\overline{\partial}\varphi(x,z)}\int_0^T\langle w(t),g(t)\rangle dt\tag{28}$$

其中 $g=[g_1,g_2]\in C_n[0,T]\times P_n[0,T]$，集合 $\partial\varphi(x,z)$ 定义如下

$$\overline{\partial}\varphi(x,z)=\Big\{w=[w_1,w_2]\in L_n^\infty[0,T]\times L_n^\infty[0,T]\ \Big|\tag{29}$$

$$[w_1(t),w_2(t)]\in\overline{\partial}\left(\tfrac{1}{2}\,h^2(x(t),z(t))\right)\quad for\ a.e.\ t\in[0,T]\Big\}.$$

证明. 根据超可微泛函的定义 (4)，为证明该定理，需要验证

1) 泛函 $\varphi(x,z)$ 沿方向 g 的导数确实具有形式 (28)；
2) 同时，集合 $\partial\varphi(x,z)$ 是空间$\big(C_n[0,T]\times P_n[0,T],||\cdot||_{L_n^2[0,T]\times L_n^2[0,T]}\big)^*$的凸弱*紧子集。

下面证明结论 1）。
首先证明以下关系成立。

$$\lim_{\alpha\downarrow 0}\frac{1}{\alpha}\left|\varphi(x+\alpha g_1,z+\alpha g_2)-\varphi(x,z)-\int_0^T\min_{[w_1,w_2]\in\overline{\partial}\left(\frac{1}{2}h^2(x,z)\right)}\left(\langle w_1(t),\alpha g_1(t)\rangle+\langle w_2(t),\alpha g_2(t)\rangle\right)dt\right|=0 \quad (30)$$

or

$$\lim_{\alpha\downarrow 0}\frac{1}{\alpha}\left|\varphi(x+\alpha g_1,z+\alpha g_2)-\varphi(x,z)-\int_0^T\frac{\partial\left(\frac{1}{2}h^2(x(t),z(t))\right)}{\partial g}dt\right|=0.$$

记 $t\in[0,T]$ 处的

$$\Phi(t,\alpha)=\frac{1}{\alpha}\Big(h^2(x(t)+\alpha g_1(t),z(t)+\alpha g_2(t))-h^2(x(t),z(t))\Big)- \quad (31)$$

$$-\min_{[w_1,w_2]\in\overline{\partial}h^2(x,z)}\left(\langle w_1(t),g_1(t)\rangle+\langle w_2(t),g_2(t)\rangle\right).$$

我们的目的是通过将勒贝格控制收敛定理应用于函数$\Phi(t,\ \alpha)$（在 $\alpha\downarrow 0$处）来证明关系（30）。

首先注意，根据超微分定义（1），以及函数 $h^2(x,\ z)$的超可微性（已在本节开头证明），对每个 $t\in[0,T]$，当 $\alpha\downarrow 0$时，我们有 $\Phi(t,\ \alpha)\to 0$。

在接下来的两段中，我们证明对每个 $\alpha>0$，都有 Φ $(t,\alpha)\in L_1^\infty[0,T].$

由于 x、 $g_1\in C_n[0,T]$、 z、 $g_2\in P_n[0,T]$，且函数 $h^2(x,\ z)$因其结构 [18], 而关于其变量连续，因此我们得到，对每个 $\alpha>0$，函数 $t\ \to\ h^2(x(t)\ z(t))$和 $t\to h^2(x(t)\ +\alpha g_1(t),z(t)\ +\alpha g_2(t))$均属于空间 $L_1^\infty[0,T]$。

由于超微分映射 [22] 的上半连续性以及超微分 $\overline{\partial}h^2(x,z)$ 的结构，很容易验证映射 $t\to\partial h^2(x(t)\ z(t))$ 是上半连续的，因此它是可测的（参见 [24]）。接着，由于函数 $g_1(t)$ 的连续性、函数 $g_2(t)$ 的分段连续性，以及数量积关于其变量的连续性，我们得到映射

$$t\to\min_{[w_1,w_2]\in\overline{\partial}h^2(x(t),z(t))}\left(\langle w_1(t),g_1(t)\rangle+\langle w_2(t),g_2(t)\rangle\right) \quad (32)$$

是上半连续的 [29], ，因而也是可测的 [24]。在证明命题 2) 时将证明，在所作假设下，集合 $\partial h^2(x,z)$ 关于 $t\in[0,T]$ 一致有界；然后根据函数 $g_1(t)$ 的连续性以及函数 $g_2(t)$ 的分段连续性，容易检查映射 (32) 关于 $t\in[0,T]$ 也一致有界。因此我们最终得到，映射 (32) 属于空间 $L_1^\infty[0,T]$。

现在证明函数 $\Phi(t,\ \alpha)$ 在 $[0,T]$ 上被某个可积函数控制，对所有充分小的 $\alpha>0$ 均成立。正如上一段所示，式 (31) 中的第二项是可积函数。因此，只需考虑充分小的 $\alpha>0$ 时式 (31) 中的第一项。由中值定理 [30]（应用于超微分），对每个 $t\in[0,T]$ 和每个 $\alpha>0$，有

$$\frac{1}{\alpha}\Big(h^2(x(t)+\alpha g_1(t),z(t)+\alpha g_2(t))-h^2(x(t),z(t))\Big)\in\overline{\partial}h^2\big(v_1(\alpha,t),v_2(\alpha_2,t)\big)g(t)$$

其中 $v_1(\alpha,\ t)\ \in\mathrm{co}\{x(t)\ x(t)\ +\alpha g_1(t)\}$， $v_2(\alpha,\ t)\ \in\mathrm{co}\{z(t)\ z(t)\ +\alpha g_2(t)\}$。在证明命题 2) 时将会说明，在所作假设下，集合 $\partial h^2(x,\ z)$ 关于 $t\in[0,T]$ 一致有界；于是，根据函数 $g_1(t)$ 的连续性和函数 $g_2(t)$ 的分段连续性，对所有充分小的 $\alpha>0$，式 (31) 中的第一项都被一个分段连续函数控制。

在 [12] 中证明了

$$\int_0^T\min_{[w_1,w_2]\in\overline{\partial}\left(\frac{1}{2}h^2(x,z)\right)}\left(\langle w_1(t),g_1(t)\rangle+\langle w_2(t),g_2(t)\rangle\right)dt=\min_{w\in\overline{\partial}\varphi(x,z)}\int_0^T\langle w_1(t),g_1(t)\rangle+\langle w_2(t),g_2(t)\rangle dt. \quad (33)$$

由关系式 (30)、(33) 可得表达式 (28)。命题 2)：相应的证明可见
于 [25]。定理得证。

□

定理 6.2. 区间 $[0,T]$ 可以划分为有限个区间，在每个区间内，函数 $\left\{f_{i,j_1}(x)\psi_i,\ldots,f_{i,j_{k(j)}}(x)\psi_i\right\}$, $i=1,n$, $j=1,r$ 中有一个（或多个）活跃。为了使点(x^*,z^*) 使泛函 $I(x,z)$达到最小值，必须满足

$$\overline{\partial} I(x^*(t),z^*(t))=\{\mathbf{0_{2n}}\} \tag{34}$$

在每个 $t\in[0,T]$ 处，其中超微分 $\overline{\partial I}(x,z)$ 的表达式由公式 *(23)* 给出（其中超微分 $\partial\varphi(x,z)$采用公式 *(29)*）。若有 $I(x^*,z^*)$，则条件 *(34)* 也充分。

7. Constructing the superdifferential descent direction of the functional $I(x,\ z)$

本节仅考虑不满足定理 6.2 中极小条件的点（$x,\ z$）。我们的目标是求泛函 $I(x,\ z)$ 在点（$x,\ z$）处的超微分（或最速）下降方向。将此方向记为 $G(x,\ z)$。在此，$G=[G_1,G_2]\in L_n^2[0,T]\times L_n^2[0,T]$。为构造向量函数 $G(x,\ z)$，考虑问题

$$\max_{w\in\overline{\partial} I(x,z)}||w||^2_{L_n^2[0,T]\times L_n^2[0,T]}=\max_{w\in\overline{\partial} I(x,z)}\int_0^T w^2(t)dt. \tag{35}$$

将 w 记为该问题的解（下文将看到，这样的解确实存在）。向量函数 w 显然依赖于点（$x,\ z$），但为简洁起见，我们在记号中省略这一依赖关系。于是可以验证，向量函数

$$G\big(x(t),z(t),t\big)=-\frac{\overline{w}\big(x(t),z(t),t\big)}{||\overline{w}||_{L^2_{2n}[0,T]}} \tag{36}$$

是在点（$x,\ z$）处泛函 $I(x,\ z)$ 的超微分下降方向（参见（43）、（44））。回想一下，我们要在点（$x,\ z$）不满足极小条件（34）的情况下，求方向 $G(x,\ z)$，因此 $||w||_{L^2_{2n}[0,T]}>0$。

注意，我们有以下等式

$$\frac{\partial I(x,z)}{\partial G(x,z)}=\min_{w\in\partial\overline{I}(x,z)}\int_0^T\langle w(t),G(x(t),z(t),t)\rangle\,dt=\min_{w\in\overline{\partial} I(x,z)}\int_0^T\left\langle w(t),\frac{-\overline{w}(t)}{||\overline{w}||_{L^2_{2n}[0,T]}}\right\rangle dt=$$

$$=\frac{-1}{||\overline{w}||_{L^2_{2n}[0,T]}}\left(-\min_{w\in\overline{\partial} I(x,z)}\int_0^T\langle -w(t),\overline{w}\rangle\,dt\right)=\frac{-1}{||\overline{w}||_{L^2_{2n}[0,T]}}\max_{w\in\partial\overline{I}(x,z)}\int_0^T\langle w(t),\overline{w}(t)\rangle\,dt=-||\overline{w}||_{L^2_{2n}[0,T]}$$

结合式 (4) 和不等式 $||w||_{L^2_{2n}[0,T]}>0$，可得

$$I\left((x,z)+\alpha G(x,z)\right)<I(x,z) \tag{37}$$

当 $\alpha>0$ 足够小时。

不难验证，在这种情况下，问题 (35) 的解是多值映射$t\to\partial I\big(x(t)\ z(t)\ t\big)$ 的一个选择子，它在每个时刻 $t\in[0,T]$ 都使零点到集合 $\partial I\big(x(t)\ z(t)\ t\big)$ 中各点的距离最大。换句话说，求解问题 (35) 就意味着求解以下问题

$$\max_{w(t)\in\overline{\partial} I(x(t),z(t),t)}w^2(t) \tag{38}$$

对于每个 $t \in [0,T]$。实际上，对于每个 $t \in [0,T]$ 我们都有显然的不等式

$$\max_{w \in \overline{\partial} I(x(t),z(t),t)} w^2(t) \geq w^2(t)$$

其中， $w(t)$ 是映射 $t \to \partial I\big(x(t)\ z(t)\ t\big)$ 的可测选择子（由 (22) 可得 $w \in L_{2n}^{\infty}[0,T]$），于是得到不等式

$$\int_0^T \max_{w \in \overline{\partial} I(x(t),z(t),t)} w^2(t)dt \geq \max_{w \in \overline{\partial} I(x,z)} \int_0^T w^2(t)dt.$$

由于对每个 $t \in [0,T]$，都有

$$\max_{w \in \overline{\partial} I(x(t),z(t),t)} w^2(t) \in \Big\{ w^2(t) \mid w(t) \in \overline{\partial} I(x(t),z(t),t) \Big\}$$

由于根据超微分的定义，集合 $\partial I\big(x(t)\ z(t)\ t\big)$ 对每个固定的 t 都是闭且有界的，并且映射 $t \to \partial I\big(x(t)\ z(t)\ t\big)$ 是上半连续的 [22]，此外，范数（及其平方）关于其自变量连续，因此根据菲利波夫引理 [31]，存在映射$t \to \partial I\big(x(t)\ z(t)\ t\big)$ 的一个可测选择子 $w(t)$，使得对每个 $t \in [0,T]$，都有

$$\max_{w \in \overline{\partial} I(x(t),z(t),t)} w^2(t) = \overline{w}^2(t),$$

因此，我们找到了集合 $\partial I(x,z)$ 中的元素 $\overline{w}$，它使前一个不等式取等号。由此，最终得到

$$\int_0^T \max_{w \in \overline{\partial} I(x(t),z(t),t)} w^2(t)dt = \max_{w \in \overline{\partial} I(x,z)} \int_0^T w^2(t)dt.$$

对于每个固定的 $t \in [0,T]$，问题 (38) 都是一个有限维问题，其目标是求凸紧集（超微分）中的点到零点的最大距离。这个问题可以有效求解；下一段将介绍其解法。实际操作中，人们会对区间 $[0,T]$ 作（均匀）划分，并对划分中的每个点求解该问题，即需要计算 $G\big(x(t_i)\ z(t_i)\ t_i\big)$，其中 $t_i \in [0,T]$、$i = 1,\ N$ 是离散化点（参见下文引理 7.1 中的记号）。在某个自然的附加假设下，下文的引理 7.1 保证：当离散化阶数趋于无穷时，通过对区间 $[0,T]$ 的此类划分中每个点处求得的超微分下降方向进行分段线性插值所得到的向量函数，在空间 $L_{2n}^2[0,T]$ 中收敛于所求的向量函数 $G\big(x(t)\ z(t)\ t\big)$。

如上一段所述，为了构造泛函 $I(x,z)$ 在点 (x,z) 处的超微分下降方向，需要求出区间 $[0,T]$ 的一个（均匀）划分中各时刻，泛函 $I(x(t)\ z(t))$ 的超微分集合内各点到零点的最大距离。由公式 (23)（另见 (25)）可知，超微分 $\partial I(x(t)\ z(t))$ 是一个凸多面体 $P(t) \subset R^{2n}$。当然，集合 $P(t)$ 取决于点 (x,z)。为简便起见，本段记号中省略这一依赖关系。显然，在这种情况下，只需遍历所有顶点 $p_j(t), j = 1,s$（此处 s 是多面体 $P(t)$ 的顶点数），并从数值 $||p_j(t)||_{R^{2n}}$ 中选出最大值。记相应的顶点为 $p_j(t)$ $(j \in \{1,\ldots,s\})$；若有多个顶点达到最大范数值，则任选其一。最后，令 $w(t) = p_j(t)$。

文献 [25] 中给出了一个引理，该引理一方面具有相当自然、适用于实际应用的条件，另一方面保证：当区间 $[0,T]$ 的（均匀）划分的网格长度趋于无穷大时，通过对所求函数$v \in L_1^{\infty}[0,T]$ 作分段线性插值而得到的函数 $L(t)$ 在空间 $L_1^2[0,T]$ 中收敛到该函数。

引理 7.1. 设函数 $v \in L_1^\infty[0,T]$ 满足以下条件：对于每个 $\delta > 0$，函数$v(t)$ 在集合 $[0,T]$ 上分段连续，只有有限多个区间$\big(t_1(\delta)\ t_2(\delta)\big), \ldots, \big(t_r(\delta)\ t_{r+1}(\delta)\big)$ 除外，且这些区间的并集长度不超过 δ。

选取区间 $[0,T]$ 的（均匀）有限划分 $t_1 = 0, t_2, \ldots, t_{N-1}, t_N = T$，并计算这些点处的值 $v(t_i)$、 $i = 1$、 N。令 $L(t)$ 为通过节点 $(t_i, v(t_i))$、 $i = 1$、 N进行分段线性插值得到的函数。那么，对于每个 $\varepsilon > 0$，都存在一个这样的数 $N(\varepsilon)$，使得对于每个 $N > N(\varepsilon)$，都有 $||L - v||^2_{L_1^2[0,T]} \le \varepsilon$。

从定性角度看，引理 7.1 的条件意味着，所求函数在区间 $[0,T]$ 上的不连续点不会“太多”。如果向量函数 $w(t)$ 满足这一条件（这对应用而言很自然），那么该引理便证明了：可以用上述方式选取的离散化点处的值 $w(t_i)$、 $i = 1,\ N$，来逼近向量函数 $w(t)$，进而逼近向量函数 $G\big(x(t)\ z(t)\ t\big)$。 ——

8. On a method for finding the stationary points of the functional $I(x,\ z\ )$

构造出最速（超微分）下降方向后（见上一节），便可应用一些基于该方向的非光滑优化方法，求泛函 $I(x,\ z)$ 的驻点。

本文的数值模拟采用了最简单的最速（超微分）下降算法。为阐明主要思想，我们首先考察该方法在有限维情形下一个类似问题中的收敛性（这一问题本身也具有独立的研究价值）；随后再转向本文所讨论的更一般问题。

首先考虑一个函数最小化问题，其中该函数是有限个连续可微函数的最小值。设

$$\varphi(x) = \min_{i=\overline{1,M}} f_i(x),$$

其中 $f_i(x)$、 $i = \overline{1,\ M}$ 是定义在 R^n 上的连续可微函数。

已知 [18] 函数 $\varphi(x)$ 在任意方向 $g \in R^n$、$||g||_{R^n} = 1$上每一点 $x_0 \in R^n$ 处均可微，并且

$$\frac{\partial \varphi(x_0)}{\partial g} = \min_{i \in R(x_0)} \left\langle \frac{\partial f_i(x_0)}{\partial x}, g \right\rangle = \min_{w \in \overline{\partial}\varphi(x_0)} \langle w, g \rangle, \tag{39}$$

$$R(x_0) = \{i \in \{1, \ldots, M\} \mid f_i(x_0) = \varphi(x_0)\},$$

$$\overline{\partial}\varphi(x_0) = \operatorname{co}\left\{ \frac{\partial f_i(x_0)}{\partial x},\ i \in R(x_0) \right\}.$$

还已知，要使点 $x^* \in R^n$ 成为函数 $\varphi(x)$ 的极小点，必须满足

$$\min_{||g||_{R^n}=1} \min_{i \in R(x^*)} \left\langle \frac{\partial f_i(x^*)}{\partial x}, g \right\rangle \ge 0. \tag{40}$$

记

$$\Psi(x) = \min_{||g||_{R^n}=1} \min_{i \in R(x)} \left\langle \frac{\partial f_i(x)}{\partial x}, g \right\rangle. \tag{41}$$

那么，函数 $\varphi(x)$ 在点 x^* 处的必要极小条件 (40) 可改写为不等式$\Psi(x^*) \ge 0$。

由公式 (39)、(41) 可知，根据定义，向量 $g_0 \in R^n$、 $||g_0||_{R^n} = 1$ 是函数 $\varphi(x)$ 在点 x^0 处的最速下降方向，当我们有

$$\min_{i \in R(x_0)} \left\langle \frac{\partial f_i(x_0)}{\partial x}, g_0 \right\rangle = \Psi(x_0). \tag{42}$$

用超微分表示（见 (39)），最速下降方向 [22]是向量

$$g_0 = -\frac{w_0}{||w_0||_{R^n}} \tag{43}$$

其中， w_0 是该问题的一个解。

$$\max_{w\in\overline{\partial}\varphi(x_0)} ||w|| = ||w_0||. \tag{44}$$

对函数 $\varphi(x)$ 最小化应用超微分（最速）下降法。描述该算法在所考虑的函数和空间上的应用。固定一个任意的初始点 $x_1 \in R^n$。假设集合

$$lev_{\varphi(x_1)}\varphi(\cdot) = \{x \in R^n \mid \varphi(x) \le \varphi(x_1)\}$$

是有界的（由于初始点的任意性，实际上必须假设，对于所取的每一个初始点，集合 $lev_{\varphi(x_1)}\varphi(\cdot)$ 都是有界的）。由于函数 $\varphi(x)$ 的连续性， [22] 集合 $lev_{\varphi(x_1)}\varphi(\cdot)$ 也是闭的。设点 $x_k \in R^n$ 已经构造完成。如果 $\Psi(x_k) \ge 0$ （在实际操作中，我们仅以某个固定精度 ε 检查该条件是否满足，即 $\Psi(x_k) \ge -\varepsilon$），那么点 x_k 就是函数 $\varphi(x)$ 的驻点，且该过程终止。否则，根据以下规则构造下一个点。

$$x_{k+1} = x_k + \alpha_k g_k$$

其中，向量 g_k 是函数 $\varphi(x)$ 在点 x_k 处的最速（超微分）下降方向（见（43）、（44））， 而值 α_k 是下列一维最小化问题的解

$$\min_{\alpha\ge 0} \varphi(x_k + \alpha g_k) = \varphi(x_k + \alpha_k g_k).$$

于是 $\varphi(x_{k+1}) < \varphi(x_k)$。

定理 8.1. 在所作假设下，有如下不等式

$$\underline{\lim}_{k\to\infty} \Psi(x_k) \ge 0 \tag{45}$$

对于按照上述算法构造的序列。

Proof. 以下证明部分基于书籍 [18 中类似证明的思路].

反设不然。则存在这样的子序列 $\{x_{k_j}\}_{j=1}^{\infty}$ 和这样的数 $b > 0$，使得对于每个 $j \in \mathcal{N}$，都有如下不等式

$$\Psi(x_{k_j}) \le -b. \tag{46}$$

注意，对每个 $j \in \mathcal{N}$，集合 $R(x_{k_j})$ 都非空。

根据最速下降方向的定义（见 (42)），对每个 $j \in \mathcal{N}$，都存在一个指标$i \in R(x_{k_j})$，使得对每个 $\alpha > 0$，关系

$$f_{\bar{i}}(x_{k_j} + \alpha g_{k_j}) = f_{\bar{i}}(x_{k_j}) + \alpha\Psi(x_{k_j}) + o_{\bar{i}}(x_{k_j}, g_{k_j}, \alpha) \tag{47}$$

成立。

回忆一下，根据该指标集的定义，对集合 $\bar{R}(x_{k_j})$ 中的每个指标 i，都有 $f_i(x_{k_j}) = \varphi(x_{k_j})$。于是，由不等式 (46) 可知，由于根据假设，$f_i(x)$ 是连续可微函数，存在 [18] 这样的 $\alpha > 0$，它与数目 k_j 无关，使得对于满足 (47) 的 $i \in R(x_{k_j})$，以及每个 $\alpha \in (0, \alpha]$)，都有如下不等式

$$f_{\bar{i}}(x_{k_j} + \alpha g_{k_j}) \le \varphi(x_{k_j}) - \frac{1}{2}\alpha b.$$

利用 $R(x_{k_j})$ 对所有 $j \in \mathcal{N}$ 都有 $\neq \emptyset$ 这一事实，最终得到

$$\varphi(x_{k_j} + \overline{\alpha} g_{k_j}) \leq \varphi(x_{k_j}) - \frac{1}{2}\overline{\alpha} b = \varphi(x_{k_j}) - \frac{1}{2}\beta \tag{48}$$

关于 $j \in \mathcal{N}$ 一致。

这个不等式导致矛盾。事实上，序列 $\{\varphi(x_k)\}_{k=1}^{\infty}$ 单调递减，且以下界 min 为下界 $_{x \in lev_{\varphi(x_{(1)})}\varphi(\cdot)} \varphi(x)$，因此它有极限：

$$\{\varphi(x_k)\} \to \varphi^* \text{ at } k \to \infty,$$

因此，对每个 $k \in \mathcal{N}$，都有

$$\varphi(x_k) \geq \varphi^*. \tag{49}$$

现在选取一个足够大的数 j，使得

$$\varphi(x_{k_{\overline{j}}}) < \varphi^* + \frac{1}{2}\beta.$$

由（48）式可得

$$\varphi(x_{k_{\overline{j}}+1}) \leq \varphi(x_{k_{\overline{j}}} + \overline{\alpha} g_{k_{\overline{j}}}) \leq \varphi^* - \frac{1}{2}\beta$$

这与（49）式矛盾。 □

现在回到泛函 $I(x,\ z)$ 最小化问题。记（见公式 (35)、(36)）

$$\Psi(x,z) = \min_{||g||_{L_n^2[0,T]\times L_n^2[0,T]}=1} \frac{\partial I(x,z)}{\partial g} = \frac{\partial I(x,z)}{\partial G}. \tag{50}$$

于是，泛函 $I(x,\ z)$ 在点 $(x^*,\ z^*)$ 处的必要极小条件可写为不等式 $\Psi(x^*,\ z^*) \geq 0$。

将超微分（最陡）下降法应用于泛函 $I(x,\ z)$ 的最小化。描述应用于该泛函及所考虑空间的算法。固定任意初始点 $(x_{(1)},\ z_{(1)}) \in C_n[0,T] \times P_n[0,T]$。假设集合

$$lev_{I(x_{(1)},z_{(1)})} I(\cdot,\cdot) = \big\{(x,z) \in C_n[0,T] \times P_n[0,T] \mid I(x,z) \leq I(x_{(1)}, z_{(1)})\big\}$$

在 $L_n^2[0,T] \times L_n^2[0,T]$-范数下有界（由于初始点的任意性，实际上必须假设，对于所取的每一个初始点，集合 $lev_{I(x_{(1)},z_{(1)})} I(\cdot,\cdot)$ 都是有界的）。设点 $(x_{(k)}, z_{(k)})$ 已经构造完成。如果 $\Psi(x_{(k)}, z_{(k)}) \geq 0$ （在实际操作中，我们只需检查该条件是否在某个固定精度 ε 下成立，即 $\Psi(x_{(k)}, z_{(k)}) \geq -\varepsilon$）（换言之，如果极小条件 (34) 得到满足（在实际操作中，精度取某个固定值 ε，即 $||w\big(x_k(t)\ z_k(t)\ t\big)||_{L_{2n}^2[0,T]} \leq \varepsilon$）），那么点 $(x_{(k)}, z_{(k)})$ 就是泛函 $I(x,z)$ 的驻点，且迭代过程终止。否则，按照以下规则构造下一个点。

$$(x_{(k+1)}, z_{(k+1)}) = (x_{(k)}, z_{(k)}) + \alpha_{(k)} G\big(x_{(k)}, z_{(k)}\big),$$

其中，向量函数 $G\big(x_{(k)},\ z_{(k)}\big)$ 是泛函 $I(x,\ z)$ 在点 $(x_{(k)},\ z_{(k)})$ 处的超微分下降方向（见（35）、（36））， 而值 $\alpha_{(k)}$ 是以下一维最小化问题的解

$$\min_{\alpha \geq 0} I\Big((x_{(k)}, z_{(k)}) + \alpha G\big(x_{(k)}, z_{(k)}\big)\Big) = I\Big((x_{(k)}, z_{(k)}) + \alpha_{(k)} G\big(x_{(k)}, z_{(k)}\big)\Big).$$

于是根据（37）可得

$$I\big(x_{(k+1)}, z_{(k+1)}\big) < I\big(x_{(k)}, z_{(k)}\big).$$

现在引入集合族 $\mathcal{I}$。首先定义泛函 $I_q\, q = 1,\ \left(\prod_{j=1}^{r} k(j)\right)^n$ 如下。其被积函数与泛函 I 的被积函数相同，但对每个 $i = 1,\ n$，用函数 $f_{i,\ j_1}\psi_i,\ \ldots,\ f_{i,\ j_{k(j)}}\psi_i\ j \in \{1, \ldots, r\}$ 中的一个函数替代最大值函数 $\max\ \{f_{i,j_1}(x)\psi_i,\ \ldots,\ f_{i,\ j_{k(j)}}(x)\psi_i\}\ j = \overline{1,\ r}$。令该族 $\mathcal{I}$ 由对时间区间 $[0, T]$ 的所有可能有限划分中各区间上的积分之和组成。此时，所取和式中每个被加项的被积函数，都与某个泛函 I_q 的被积函数相同，即 $q \in \left\{1,\ \ldots,\ \left(\prod_{j=1}^{r} k(j)\right)^n\right\}$。

对于通过所述方法构造的每个点，以下假设均成立：区间 $[0, T]$ 可划分为有限个区间，在每个区间内，对于每个 $i = 1,\ n$ 要么 $h_i(x_{(k)},\ z_{(k)}) = 0$，要么函数 $\left\langle -a_{ij}\psi_i^* \frac{\partial f_{i\ \cdot\ j_p(x_{(k)})}}{\partial x},\ G_1(x_{(k)},\ z_{(k)})\right\rangle, j = 1,\ r, p = 1,\ k(j)$ 中的一个（或多个）是活跃的。

让我们通过一个例子来说明这一假设。考虑以下最简单的泛函（然而，其结构保留了一般情形的基本特征）

$$\int_0^1 \frac{1}{2}\min^2\{x(t) + 1, -x(t) + 1\}dt.$$

考虑点 $x_{(1)} = 0$ （即对所有 $t \in [0, 1],\ x_{(1)}(t) = 0$ ）。为了求出该点处泛函的最速（超微分）下降方向，根据所述理论，需要将方向导数（在该点计算）最小化，即求出这样的函数 $G \in L_1^2[0, T]$，$\int_1 G^2(t)dt = 1$，它使泛函0达到最小值

$$\int_0^1 \min\{x_{(1)}(t) + 1, -x_{(1)}(t) + 1\}\min\{g(t), -g(t)\}dt.$$

这里，$g \in L_1^2[0, T]$，$\int_0^1 g^2(t)dt = 1$。取

$$G(t) = \begin{cases} -1, & \text{if } t \in [0, 0.5], \\ 1, & \text{if } t \in (0.5, 1] \end{cases}$$

作为一个显然的解。于是，$\Psi(x_{(1)}) = \int_0^1 -1dt = -1$。我们看到，所作的假设成立。取

$$\hat{I}(x) = \int_0^{0.5} \frac{1}{2}(x(t) + 1)^2 dt + \int_{0.5}^1 \frac{1}{2}(-x(t) + 1)^2 dt.$$

于是我们得到

$$\hat{I}\left(x_{(1)} + \alpha G(x_{(1)})\right) = \hat{I}(x_{(1)}) + \alpha\Psi(x_{(1)}) + o(x_{(1)}, G(x_{(1)}), \alpha)$$

因为

$$\nabla\hat{I}(x_{(1)}) = \begin{cases} 1, & \text{if } t \in [0, 0.5], \\ -1, & \text{if } t \in (0.5, 1], \end{cases}$$

也就是说，$\nabla\hat{I}(x_{(1)})G(x_{(1)}) = \int_0^1 -1dt = -1$。显然，$\hat{I} \in \mathcal{I}$。注意，取 $\alpha_{(1)} = 1$ 时可得点

$$x_{(2)}(t) = \begin{cases} -1, & \text{if } t \in [0, 0.5], \\ 1, & \text{if } t \in (0.5, 1], \end{cases}$$

它给出了所考察泛函的全局最小值。

对于来自族 $\mathcal{I}$ 的泛函，我们作如下附加假设：存在这样的有限数 $\hat{} L$，使得对于每个 $I \in \mathcal{I}$，以及对于来自以原点为中心、半径为某个有限数 $R' + \hat{}\bar{\alpha}$ 的球内的所有 $x,\ z,\ x,\ z$（这里 $R' > \sup_{(x,\ z) \in lev_{I(x_{(1)}, z_{(1)})} I(\cdot,\ \cdot)} ||(x,\ z)||_{L_n^2[0,T] \times L_n^2[0,T]}$，且 $\hat{}\alpha$ 是某个正数），均有

$$||\nabla \hat{I}(\overline{x}, \overline{z}) - \nabla \hat{I}(\overline{\overline{x}}, \overline{\overline{z}})||_{L_n^2[0,T] \times L_n^2[0,T]} \leq L||(\overline{x}, \overline{z})' - (\overline{\overline{x}}, \overline{\overline{z}})'||_{L_n^2[0,T] \times L_n^2[0,T]}. \tag{51}$$

$\hat{}$备注 8.2. 乍看之下，要求假设中对所有 $I \in \mathcal{I}$ 同时存在 Lipschitz 常数 L 似乎负担过重。然而，只要记得在有限个 $\hat{}$ 区间 $[0,T]$ 段上，泛函 $I \in \mathcal{I}$ 的被积函数按构造（见集合 $\mathcal{I}$ 的定义）与泛函 $I_q, q \in \left\{1, \ldots, \left(\prod_{j=1}^{r} k(j)\right)^n\right\}$ 的被积函数一致，那么如果我们假设每个梯度 $\nabla I_q,\ q = 1, \left(\prod_{j=1}^{r} k(j)\right)^n$ 都满足 Lipschitz 连续性，这一假设便是自然的；而梯度的 Lipschitz 连续性条件是为可微泛函的经典优化方法提供依据时常用的假设。

引理 8.3。 设条件 *(51)* 得到满足。则对于每个泛函 $I \in \mathcal{I}$ 以及所有 $(x, z) \in lev_{I(x_{(1)}, z_{(1)})} I(\cdot, \cdot)$, $G \in C_n[0,T] \times P_n[0,T]$, $||G||_{L_n^2[0,T] \times L_n^2[0,T]} = 1$, $\alpha \in R$, $0 < \alpha \leq \hat{}\alpha$，均有不等式

$$I\left((x,z) + \alpha G\right) \leq I(x,z) + \alpha \left\langle \nabla I(x,z), G \right\rangle + \alpha^2 \frac{L}{2}.$$

证明。 证明过程可按照与 [33] 中类似命题相同的方式，通过显而易见的修改完成。□

我们假设，在方法实现过程中，对于每个 $k \in \mathcal{N}$，均有 $\alpha_{(k)} < \hat{}\alpha$，其中 $\hat{}\alpha$ 是引理 8.3 中的一个数（另见备注 8.2 前的假设）。

定理 8.4。 在上述假设下，有不等式

$$\underline{\lim}_{k \to \infty} \Psi(x_k, z_k) \geq 0 \tag{52}$$

对于按照上述规则构造的序列。

证明。反设，存在这样的子序列 $\{(x_{k_j}, z_{k_j})\}_{j=1}^{\infty}$ 以及这样的数 $b > 0$，使得对每个 $j \in \mathcal{N}$，不等式

$$\Psi(x_{k_j}, z_{k_j}) \leq -b. \tag{53}$$

由最速下降方向（见（50））和集合 $\mathcal{I}$ 定义（另见（27））可知，对每个$j \in \mathcal{N}$，都存在族 $\mathcal{I}$ 中的一个泛函 $\hat{}I$，使得对每个 $\alpha > 0$，关系

$$\hat{I}\left((x_{k_j}, z_{k_j}) + \alpha G(x_{k_j}, z_{k_j})\right) = \hat{I}(x_{k_j}, z_{k_j}) + \alpha \Psi(x_{k_j}, z_{k_j}) + o\left((x_{k_j}, z_{k_j}), G(x_{k_j}, z_{k_j}), \alpha\right) \tag{54}$$

成立；其中，

$$\Psi(x_{k_j}, z_{k_j}) = \left\langle \nabla \hat{I}(x_{k_j}, z_{k_j}), G(x_{k_j}, z_{k_j}) \right\rangle.$$

回顾一下，根据集合定义，对于族 $\mathcal{I}$ 中的每个泛函 $\hat{}I$，$\hat{}I(x_{k_j}, z_{k_j}) = I(x_{k_j}, z_{k_j})$。于是，由不等式 (53) 以及引理 8.3 可知，存在一个与数目 k_j 无关的 $\alpha > 0$，使得对于满足 (54) 的 $\hat{}\, I \in \mathcal{I}$ 以及每个 $\alpha \in (0, \alpha]$，都有如下不等式

$$\hat{I}\left((x_{k_j}, z_{k_j}) + \alpha G(x_{k_j}, z_{k_j})\right) \leq I(x_{k_j}, z_{k_j}) - \frac{1}{2}\alpha b.$$

利用集合 $\mathcal{I}$ 的定义，最终得到

$$I\left((x_{k_j}, z_{k_j}) + \alpha G(x_{k_j}, z_{k_j})\right) \le I(x_{k_j}, z_{k_j}) - \frac{1}{2}\overline{\alpha} b = I(x_{k_j}, z_{k_j}) - \frac{1}{2}\beta \tag{55}$$

关于 $j \in \mathcal{N}$ 一致地。

这个不等式导致矛盾。事实上，序列 $\left\{I(x_{(k)}, z_{(k)})\right\}_{k=1}^{\infty}$ 单调递减且下有界于零（注意，根据构造，泛函 $I(x, z)$ 是非负的），因此它有极限：

$$\left\{I(x_{(k)}, z_{(k)})\right\} \to I^* \text{ at } k \to \infty, \tag{56}$$

从而，对于每个 $k \in \mathcal{N}$，都有 $\{I(x_{(k)}, z_{(k)})\} \ge I^*$。

现在选取这样一个足够大的数 j，使得

$$I(x_{(k_{\overline{j}})}, z_{(k_{\overline{j}})}) < I^* + \frac{1}{2}\beta.$$

由（55）可得

$$I(x_{(k_{\overline{j}}+1)}, z_{(k_{\overline{j}}+1)}) \le I(x_{(k_{\overline{j}})}, z_{(k_{\overline{j}})}) + \overline{\alpha} G_{k_{\overline{j}}} \le I^* - \frac{1}{2}\beta$$

这与（56）矛盾。 □

注 8.5. 容易证明，对于所考虑的情形，公式（45）、（52）中的下极限实际上可以替换为“通常”极限，不等式也可以替换为等式。

9. Numerical examples

In t本节给出两个示例，说明所述算法的实现 n.

例 9.1. 考虑微分包含

$$\dot{x}_1 \in x_2 + [-2, 2](|x_1| + |x_2|), \quad \dot{x}_2 \in x_1 + [-2, 2](|x_1| + |x_2|)$$

在时间区间 $[0, 1]$ 上，并满足边界条件

$$x_1(0) = 0,\ x_2(0) = 1, \quad x_1(1) = 0,\ x_2(1) = 2.$$

取相位约束曲面为

$$s(x) = x_1 = 0.$$

取 $(x_{(1)},\ z_{(1)}) = (0,\ 0,\ 0,\ 0)'$ 作为初始近似（即所有考虑的函数在区间 $[0, 1]$ 上均恒等于零），则 $I(x_{(1)},\ z_{(1)}) = 1$。过程结束时，离散步长为 2×10^{-1}。图1展示了得到的轨迹。从图中可以看出，微分包含得到满足。该轨迹也位于所要求的曲面上。边界值误差不超过 5×10^{-3}。为达到这样的精度，需要进行 5 次迭代。所得轨迹上的泛函值约为 3×10^{-5}。

例9.2。 考虑微分包含

$$\dot{x}_1 \in \mathrm{co}\{x_1, x_2, x_3\}, \quad \dot{x}_2 \in \mathrm{co}\{-2x_1, x_3\}, \quad \dot{x}_3 = x_1 + x_2$$

在时间区间 $[0, 1]$ 内，满足边界条件

$$x_1(0) = -1,\ x_2(0) = -1,\ x_3(0) = 2.5, \quad x_1(1) = 1,\ x_2(1) = 1,\ x_3(1) = 2.5.$$

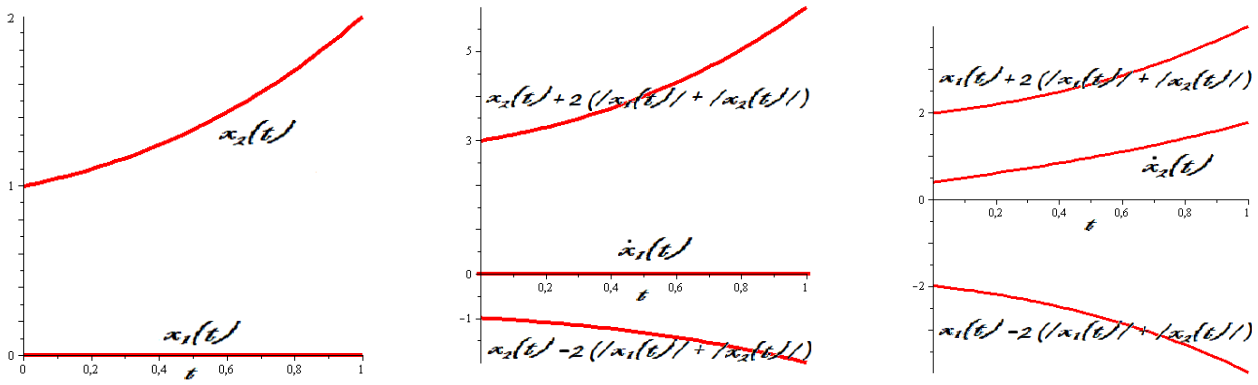

FIGURE 1. 示例 1

将相位约束曲面设为

$$s(x) = x_1 - x_2 = 0.$$

取 $(^{x_{(1)}},\ ^{z_{(1)}}) = (0,\ 0,\ 0,\ 0,\ 0,\ 0)'$ 作为初始近似（即所考虑的所有函数在区间 $[0,1]$ 上恒等于零），则 $I(^{x_{(1)}},\ ^{z_{(1)}}) = 8.125$。过程结束时，离散步长为 10^{-1}。图2展示了得到的轨迹。从图中可以看出，微分包含得到了满足。该轨迹也位于所要求的曲面上。边界值误差不超过 2×10^{-3}。为达到这样的精度，需要进行 66 次迭代。所得轨迹上的泛函值约为 2×10^{-5}。

请注意，我们特意选取这样的约束曲面，以使轨迹 $x_1(t)$和$x_2(t)$重合。因此，如果在某些迭代中，轨迹已经位于曲面 $s(x) = 0$ 上，并且 $h_1(x,\ z) > 0$，那么这两条轨迹都会在某个非零时间区间内处于活跃状态（如果在该区间上 $x_1(t) = x_2(t) > x_3(t)$）。这将导致在这些迭代中计算“完整的”超微分（而不是退化为单个点的超微分）。



FIGURE 2. 例 2

本文的主要成果（第6、7、8、9节）在俄罗斯科学院应用数学研究所完成，并得到俄罗斯科学基金会的支持（项目编号20-71-10032）。

References

[1] Sandberg M. 非凸微分包含的前向欧拉方法的收敛性 // SIAM 数值分析杂志。2008。第47卷，第1期。第308–320页。[2] Bastien J. 极大单调微分包含的隐式欧拉数值格式的收敛阶 // 应用数学与物理杂志。2013。第64卷。第955–966页。[3] Beyn W-J., Rieger J. 单侧 Lipschitz 微分包含的隐式欧拉格式 // 离散与连续动力系统 B辑。2010。第14卷，第2期。第409–428页。[4] Lempio F. 微分包含的修正欧拉方法 // 载于：集合值分析与微分包含研讨会论文集，保加利亚 Pamporovo。1990。[5] Veliov V. 强凸微分包含的二阶离散逼近 // 系统与控制快报。1989。第13卷。第263–269页。[6] Dontchev A., Lempio F. 微分包含的差分方法：综述 // SIAM 评论。1992。第34卷，第2期。第263–294页。[7] Schilling K. 求解微分包含边值问题的算法及其在最优控制中的应用 // 数值泛函分析与优化。1989。第10卷，第7期。第733–764页。[8] Cernea A., Georgescu C. 具有状态约束的微分-差分包含的必要最优性条件 // 数学分析与应用杂志。2007。第344卷。第43–53页。[9] Pappas G. S. 存在状态约束时微分包含的最优解 // 优化理论与应用杂志。1984。第44卷，第4期。第657–679页。[10] Zhu Q. J. 具有端点约束的非凸微分包含的必要最优性条件 // 微分方程杂志。1996。第124卷。第186–204页。[11] Arutyunov A. V., Aseev S. M., Blagodatskikh V. I. 具有相位约束的微分包含最优控制问题的一阶必要条件 // 俄罗斯科学院数学论文集。1994。第79卷，第1期。第117–139页。[12] Fominyh A. V. 非光滑变分问题中的次微分下降方法 // 优化通讯，2022（印刷中）。DOI 10 1007 11590 022 01897 3 。/s - - - Fominyh A. V. 求解微分包含最优解的数值方法 // 圣彼得堡大学学报，数学。[13] 2018。第51卷，第4期。第397–406页。[14] Fominyh A. V. 求解具有固定右端点的微分包含的方法 // 圣彼得堡大学学报。应用数学、计算机科学与控制过程。2018。第14卷，第4期。第302–315页。[15] Fominyh A. V. 求解线性非定常区间常微分方程系统的方法 // 圣彼得堡大学学报。应用数学、计算机科学与控制过程。2021。第17卷，第2期。第148–165页。[16] Dolgopolik, M. V., Fominyh, A. V. 最优控制问题的精确罚函数 I：主定理与自由端点问题 // 最优控制应用与方法。2019。第40卷，第6期。第1018–1044页。[17] Fominyh A. V. 由右端非光滑系统描述的对象的开环控制 // 计算数学与数学物理。2019。第59卷，第10期。第1639–1648页。[18] Demyanov V. F., Malozemov V. N. 极小极大理论导论。1990。纽约：Dover Publications Inc. 320页。[19] Filippov A. F. 具有不连续右端的微分方程。1988。多德雷赫特：Springer Science+Business Media B. V. 304页。[20] Utkin V. I. 控制与优化中的滑模。1992。柏林：Springer-Verlag。286页。[21] Levant A. 高阶滑模、微分与输出反馈控制 // 国际控制杂志。2003。第76卷，第9期。第924–941页。[22] Demyanov V. F., Vasil'ev L. V. 非可微优化。1986。纽约：Springer-Optimization Software。452页。[23] Aizerman M. A., Pyatniskii E. S. 不连续系统理论基础 I // 自动化与远程控制。1974。第7卷。第33–47页。[24] Blagodatskikh V. I., Filippov A. F. 微分包含与最优控制 // 斯捷克洛夫数学研究所论文集。1986。第169卷。第199–259页。[25] Fominyh A. V. 具有滑模的不连续系统求解方法。arXiv：2204.11266。2022。第1–17页。[26] Dolgopolik M. V. 通过余微分研究变分法中的非光滑问题 // ESAIM：控制、优化与变分法。2014。第20卷，第4期。第1153–1180页。[27] Bonnans J. F., Shapiro A. 优化问题的摄动分析。2000。纽约：Springer Science+Business Media Publ. 601页。[28] Demyanov V. F., Rubinov A. M. 非光滑分析与拟微分演算基础。1990。莫斯科：Nauka。432页。[29] Aubin J.-P., Frankowska H. 集合值分析。1990。波士顿：Birkhauser Basel。482页。[30] Makela M., Neittaanmaki P. 非光滑优化：分析、算法及其在最优控制中的应用。1992。伦敦：World Scientific Publishing。254页。[31] Filippov A. F. 关于最优控制理论中的某些问题 // 工业与应用数学学会杂志，A辑：控制。1959。第1卷，第1期。第76–84页。[32] Demyanov V. F. 极值条件与变分法。2005。莫斯科：Vysshaya shkola。335页。[33] Kantorovich L. V., Akilov G. P. 泛函分析。1977。莫斯科：Nauka 出版社。752页。