ON OPTIMALITY CONDITIONS FOR TRILEVEL DYNAMIC OPTIMIZATION PROBLEMS

Shao, Yi-Hang
Department of Mathematics, Saitama University

https://doi.org/10.5109/13485

出版情報：Bulletin of informatics and cybernetics. 31 (2), pp.117-136, 1999-12. Research Association of Statistical Sciences
バージョン：
権利関係：

Bulletin of Informatics and Cybernetics, Vol. 31, No. 2, 1999

ON OPTIMALITY CONDITIONS FOR TRILEVEL DYNAMIC OPTIMIZATION PROBLEMS

By

Yi-Hang SHAO*

Abstract

In this paper, we introduce trilevel dynamic optimization problems. Reformulating the trilevel dynamic problem as a single-level optimal control problem with state-control functional constraints, we derive the necessary optimality conditions. We also show that the necessary conditions are sufficient for optimality in a 'convex' case.

Key Words and Phrases: trilevel dynamic optimizaion problems, nonsmooth analysis, necessary conditions, sufficient conditions, constraint qualifications.

1. Introduction

In this paper, we consider the following trilevel dynamic optimization problem $(TDOP)$. In this problem, there are three players A, B and C whose controls are u, v and w, respectively. The players A, B and C minimize their cost functionals on the following manner:

For each u chosen by the player A, the player B selects a control $v = v[u] \in \mathcal{V}[u]$, where $\mathcal{V}[u]$ is the set of optimal controls v of the following optimal control problem,

$$\begin{array}{lll} P_B[u]: & Minimize: & \int_0^1 G_1(t, y(t), u(t), v(t))dt + g_1(y(1)) \\ & subject\ to: & \dot{y}(t) = \varphi(t, y(t), u(t), v(t)) \qquad a.e. \\ & & y(0) \in C_2 \\ & & v(t) \in U_2(t) \qquad a.e.. \end{array}$$

Next, the player C chooses a control $w = w[u, v] \in \mathcal{W}[u, v]$ $(v = v[u])$, where $\mathcal{W}[u, v]$ is the set of optimal controls w of the following problem,

$$\begin{array}{lll} P_C[u, v]: & Minimize: & \int_0^1 G_2(t, z(t), u(t), v(t), w(t))dt + g_2(z(1)) \\ & subject\ to: & \dot{z}(t) = \psi(t, z(t), u(t), v(t), w(t)) \qquad a.e. \\ & & z(0) \in C_3 \\ & & w(t) \in U_3(t) \qquad a.e.. \end{array}$$

---

* Department of Mathematics, Saitama University, Japan.

Among all controls $(u, v, w) = (u, v[u], w[u, v[u]])$, the player A selects a control optimizing the following optimal control problem,

$$\begin{array}{lll} P_A: & Minimize: & \int_0^1 F(t, x(t), u(t), v(t), w(t))dt + f(x(1)) \\ & subject\ to: & \dot{x}(t) = \phi(t, x(t), u(t), v(t), w(t)) \qquad a.e. \\ & & x(0) \in C_1 \\ & & u(t) \in U_1(t) \qquad a.e. \\ & & v \in \mathcal{V}[u] \\ & & w \in \mathcal{W}[u, v]. \end{array}$$

In this trilevel dynamic optimization problem $(TDOP)$,

$$\begin{array}{ll} (D): & \begin{array}{l} (x(\cdot), y(\cdot), z(\cdot)) \in AC([0,1], R^{m_1} \times R^{m_2} \times R^{m_3}) \ is\ the\ state, \\ (u(\cdot), v(\cdot), w(\cdot)) \in L^1([0,1], R^{n_1} \times R^{n_2} \times R^{n_3}) \ is\ the\ control, \\ \left.\begin{array}{l} F: [0,1] \times R^{m_1} \times \ R^{n_1} \times R^{n_2} \times R^{n_3} \to R, \\ f: R^{m_1} \to R, \\ G_1: [0,1] \times R^{m_2} \times R^{n_1} \times R^{n_2} \to R, \\ g_1: R^{m_2} \to R, \\ G_2: [0,1] \times R^{m_3} \times R^{n_1} \times R^{n_2} \times R^{n_3} \to R, \\ g_2: R^{m_3} \to R, \\ \phi: [0,1] \times R^{m_1} \times \ R^{n_1} \times R^{n_2} \times R^{n_3} \to R^{m_1}, \\ \varphi: [0,1] \times R^{m_2} \times \ R^{n_1} \times R^{n_2} \ \to R^{m_2}, \\ \psi: [0,1] \times R^{m_3} \times \ R^{m_1} \times R^{n_2} \times R^{n_3} \to R^{m_3} \end{array}\right\} \ are\ given\ functions, \\ C_i \ is\ subset\ of\ R^{m_i} \ (i = 1, 2, 3), \\ U_i(t): [0,1] \to 2^{R^{n_i}} \ is\ set\ valued\ function \ (i = 1, 2, 3), \end{array} \end{array}$$

where $AC([0,1], R^{m_1} \times R^{m_2} \times R^{m_3})$ denotes the space of absolutely continuous functions on [0,1] with value in $R^{m_1} \times R^{m_2} \times R^{m_3}$.

A control (u, v, w) of $(TDOP)$ corresponding to the state (x, y, z) is called *admissible* iff (x, u, v, w) satisfies the differential equation $\dot{x} = \phi(t, x, u, v, w)$ *a.e.* with initial condition $x(0) \in C_1$ and control constraint $u(t) \in U_1(t)$ *a.e.*, (y, v) and (z, w) are optimal solution for $P_B[u]$ and $P_C[u, v]$ respectively. An admissible control (u_*, v_*, w_*) corresponding to state (x_*, y_*, z_*) is called *optimal* for $(TDOP)$ iff (x_*, u_*, v_*, w_*) minimizes the value of cost functional of P_A over all admissible controls corresponding to the states of $(TDOP)$.

This trilevel dynamic optimization problem can be applied in various areas. For instance, in economics, the controls u are government's monetary or fiscal policies; the controls v are decisions of consumers which respond to the policies u; the controls w are decisions of firms which depend on the policies of government u and the consumer's decisions v.

Many papers have been devoted to bilevel programming problems (static optimization). Yezza (1996) studied necessary optimality conditions for multilevel programming problem. For the bilevel dynamic optimization problems, the recent results are given by Ye (1995, 1997). In these papers, under some assumptions, Ye reduced the bilevel dynamic optimization problems to single-level optimal control problems without state-control constraints. Then, Ye derived the necessary optimality conditions.

To our knowledge, there is no paper dealing with optimality conditions for general trilevel optimal control problems. The main purpose of our paper is to discuss the necessary optimality conditions for the above trilevel dynamic optimization problem in general case. This problem can not be reduced to a single-level optimal control problem which has no state-control constraints. Thus, we can not extend the results of necessary conditions in Ye (1995, 1997) to our $(TDOP)$. Moreover, we show that our necessary conditions are also sufficient for optimality under some convexity assumptions on the functions and sets in (D). Using the results given below, we can also derive the optimality conditions for another type trilevel optimization problems (see Remark (i) below). These conclusions can be generalized easily to k-level dynamic optimization problems without other additional hypothesis.

Define the value function $V_1(u) : L^1([0,1], R^{m_1}) \to R \cup \{+\infty\} \cup \{-\infty\}$ for $P_B[u]$ by:

$$\begin{aligned} V_1(u) := \quad inf \Bigg\{ & \int_0^1 G_1(t, y(t), u(t), v(t))dt + g_1(y(1)) : \\ & \dot{y}\ (t) = \varphi(t, y(t), u(t), v(t)) \ a.e., \quad y(0) \in C_2, \ v(t) \in U_2(t) \ a.e. \Bigg\}, \end{aligned}$$

the value function $V_2(u,v) : L^1([0,1], R^{m_1}) \times L^1([0,1], R^{m_2}) \to R \cup \{+\infty\} \cup \{-\infty\}$ for $P_C[u,v]$ by:

$$\begin{aligned} V_2(u,v) := \quad inf \Bigg\{ & \int_0^1 G_2(t, z(t), u(t), v(t), w(t))dt + g_2(z(1)) : \\ & \dot{z}\ (t) = \psi(t, z(t), u(t), v(t), w(t)) \ a.e., \ z(0) \in C_3, \ w(t) \in U_3(t) \ a.e. \Bigg\}. \end{aligned}$$

Throughout this paper, by convention, we assume that the infimum over empty set is $+\infty$.

Then, the above problem $(TDOP)$ is obviously equivalent to the following single-level optimal control problem,

$$\begin{aligned} P: \qquad & Minimize: && \int_0^1 F(t, x(t), u(t), v(t), w(t))dt + f(x(1)) \\ & subject\ to: && \dot{x}\ (t) = \phi(t, x(t), u(t), v(t), w(t)) \qquad a.e. \\ &&& \dot{y}\ (t) = \varphi(t, y(t), u(t), v(t)) \qquad a.e. \\ &&& \dot{z}\ (t) = \psi(t, z(t), u(t), v(t), w(t)) \qquad a.e. \\ &&& (x(0), y(0), z(0)) \in C_1 \times C_2 \times C_3 \\ &&& (u(t), v(t), w(t)) \in U_1(t) \times U_2(t) \times U_3(t) \qquad a.e. \\ &&& \int_0^1 G_1(t, y(t), u(t), v(t))dt + g_1(y(1)) - V_1(u) \le 0 \\ &&& \int_0^1 G_2(t, z(t), u(t), v(t), w(t))dt + g_2(z(1)) - V_2(u,v) \le 0, \end{aligned}$$

which contains state-control constraints in which the value functions are nonsmooth in general. We shall study the optimality conditions for such optimal control problem in the next section.

2. Nonsmooth Optimal Control Problem

In this section, we deal with the following problem:

$$\begin{array}{lll} (NOCP): & Minimize: & \int_0^1 L_0(t,x(t),u(t))dt + h_0(x(1)) \\ & subject\ to: & \dot{x}(t) = \Phi(t,x(t),u(t)) \qquad a.e. \\ & & x(0) \in C \\ & & u(t) \in U(t) \qquad a.e. \\ & & \bar{G}_i(x,u) \leq 0 \qquad i = 1,\cdots,k, \end{array}$$

in which

$$\bar{G}_i(x,u) := \int_0^1 L_i(t,x(t),u(t))dt + h_i(x(1)) - \Lambda_i(u).$$

Here, $x(\cdot) \in AC([0,1],R^m)$, $u(\cdot) \in L^1([0,1],R^n)$, $L_i : [0,1] \times R^m \times R^n \to R$, $h_i : R^m \to R$ $(i = 0,\cdots,1)$, $\Lambda_i : L^1 \to R \cup \{+\infty\} \cup \{-\infty\}$ $(i = 1,\cdots,k)$, $\Phi : [0,1] \times R^m \times R^n \to R^m$, C is a subset of R^m and $U : [0,1] \to 2^{R^n}$.

We say that $(x,u) \in AC \times L^1$ is an admissible process for $(NOCP)$ iff $L_i(\cdot,x(\cdot),u(\cdot))$ $(i = 0,\cdots,k)$ are integrable and (x,u) satisfies all constraints in $(NOCP)$. An admissible process (x_*,u_*) is called a local minimizer for $(NOCP)$ iff (x_*,u_*) minimizes the cost over all admissible processes (x,u) satisfying $\|x - x_*\|_{L^\infty} < \epsilon$ and $\|u - u_*\|_{L^1} < \epsilon$ for some $\epsilon > 0$.

2.1. Necessary conditions

Letting (x_*,u_*) be an admissible process for $(NOCP)$, we assume that $\{C,\ U(t),\ \Phi,\ L_i,\ h_i\ (i = 0,\cdots,1),\ \Lambda_i\ (i = 1,\cdots,k)\}$ satisfies (A1)-(A6) below.

(A1): C is a closed subset of R^m.

(A2): $U(\cdot)$ is a closed set-valued map, GrU is $\mathcal{L} \times \mathcal{B}$ measurable. There exists $\rho(\cdot) \in L^1$, such that $|u| \leq \rho(t)$ for any $u \in U(t)$ a.e. $t \in [0,1]$, where $|u|$ is the Euclidean norm for $u \in R^m$.

(A3): $\Phi(t,x,u)$ is measurable in t, continuously differentiable in (x,u). There exists $k(t) \in L^\infty$ and $\epsilon > 0$ such that for almost all t

$$|\Phi(t,x_1,u_1) - \Phi(t,x_2,u_2)| \leq k(t)(|x_1 - x_2| + |u_1 - u_2|)$$

for $x_1,x_2 \in x_*(t) + \epsilon B_{R^m}$, $u_1,u_2 \in R^n$. B_{R^m} is the closed unit ball of R^m.

(A4): $L_0,\cdots,L_k$ are measurable in t, and there exists $k_L(\cdot) \in L^\infty$ such that for any $i \in \{0,\cdots,k\}$,

$$|L_i(t,x_1,u_1) - L_i(t,x_2,u_2)| \leq k_L(t)(|x_1 - x_2| + |u_1 - u_2|)$$

for all $x_1,x_2 \in x_*(t) + \epsilon B_{R^m}$, $u_1,u_2 \in R^n$ a.e. $t \in [0,1]$.

(A5): There is $k_h > 0$ such that for each $i \in \{0,\cdots,1\}$,

$$|h_i(x_1) - h_i(x_2)| \leq k_h\,|x_1 - x_2|$$

for any $x_1, x_2 \in x_*(1) + \epsilon B_{R^m}$.

(A6): There exists $k_\Lambda > 0$ such that for any $i \in \{1, \cdots k\}$,

$$|\Lambda_i(u_1) - \Lambda_i(u_2)| \leq k_\Lambda \left\|u_1 - u_2\right\|_{L^1}$$

for any $u_j(\cdot)$ with $\left\|u_j - u_*\right\|_{L^1} \leq \epsilon, j = 1, 2$.

Now, we state our main theorem for necessary optimality conditions. For simplicity, we abbreviated the arguments $(t, x_*(t), u_*(t))$ to $[t]$, for instance $\Phi[t] := \Phi(t, x_*(t), u_*(t))$.

THEOREM 2.1. *Let $(x_*(\cdot), u_*(\cdot))$ be a local minimizer for (NOCP). Assume that (A1)-(A6) are satisfied. Then, there exist $\lambda_i \geq 0$ $(i = 0, \cdots, k)$ with $\sum_{i=0}^k \lambda_i = 1$, $p(\cdot) \in AC([0,1], R^m)$ and $\zeta(\cdot) \in L^\infty([0,1], R^n)$ such that*

$$(-\dot{p}(t), \zeta(t)) \in \nabla_{(x,u)} \langle p(t), \Phi[t] \rangle - \sum_{i=1}^k \lambda_i \partial_{(x,u)} L_i[t] \quad a.e. \tag{2.1}$$

$$\zeta = \zeta_1 + \zeta_2, \quad -\zeta_1 \in \sum_{i=1}^k \lambda_i \partial \Lambda_i(u_*(\cdot)), \quad \zeta_2(t) \in N_{U(t)}(u_*(t)) \ a.e. \tag{2.2}$$

$$p(0) \in N_C(x_*(0)), \quad -p(1) \in \sum_{i=1}^k \lambda_i \partial h_i(x_*(1)) \tag{2.3}$$

$$\lambda_i \left(\int_0^1 L_i[t]dt + h_i(x_*(1)) - \Lambda_i(u_*) \right) = 0 \quad (i = 1, \cdots, k). \tag{2.4}$$

In particular, assume that there exists $(x_1, u_1) \in AC \times L^1$ such that the following constraint qualifications (2.5) and (2.6) hold,

$$\int_0^1 L_i{}^\circ((x_*, u_*), (x_1, u_1))dt + h_i{}^\circ(x_*, x_1) + (-\Lambda_i)^\circ(u_*, u_1) < 0 \quad for\ i \in I_* \tag{2.5}$$

$$\dot{x}_1 = \Phi_x[t]x_1 + \Phi_u[t]u_1 \quad a.e., \quad x_1(0) \in T_C(x_*(0)) \ and \ u_1(t) \in T_{U(t)}(u_*) \quad a.e., \tag{2.6}$$

where, $I_* := \{i \in [1, \cdots, k] : \bar{G}_i[t] = 0\}$. Then, we have $\lambda_0 > 0$.

We will give the proof of this theorem in Section 5.

REMARK. (i) In Theorem 2.1, ∇ denotes the gradient in usual sense; ∂ indicates the Clarke generalized gradient; N_C and $N_{U(t)}$ are the Clarke normal cones associated with C and $U(t)$, respectively; Λ_i°, L_i°, h_i° are the Clarke generalized directional derivative; $T_C(x_*(0))$ and $T_{U(t)}$ are the Clarke tangent cones associated with C and $U(t)$, respectively (see Clarke (1983)).

(ii) To prove Theorem 2.1, we will follow the idea in the proof of Theorem 3.2.6 in Clarke (1983). By the analogous way, Ye and Zhu (1997) showed necessary optimality conditions for a bilevel perturbed differential inclusion problem (Theorem 3.1). However, in the proof of Theorem 3.1 they used a lemma (Lemma 6.2) whose proof seems insufficient. It seems that their Theorem 3.1 has not been proved.

2.2. Sufficient condition in a convex case

Pinho and Vinter (1995) pointed out that for the nonsmooth optimal control problem without mixed state-control constraints, the *weak maximum principle* is sufficient for optimality in 'normal and convex' case, while the general form of the nonsmooth *maximum principle* may fail to be sufficient. For $(NOCP)$, we show that the necessary optimality conditions in Theorem 2.1 are also sufficient in a 'convex' case. These consequences can be used in the multilevel optimization problems.

Now, let us denote by $(LOCP)$ the problem $(NOCP)$ whose state equation is given by

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) + b(t),$$

where $A(\cdot) : [0,1] \to R^{m\times m}$ and $B(\cdot) : [0,1] \to R^{m\times n}$ are integrable, $b(\cdot) : [0,1] \to R^m$ is measurable.

We will impose the following hypotheses :

(H1): C is convex set of R^m.

(H2): $U(t)$ is convex in R^n for almost all $t \in [0,1]$.

(H3): L_i $(i = 0, \cdots, k)$ are Lebesgue measurable in t, convex in (x,u).

(H4): h_i $(i = 0, \cdots, k)$ are convex functions.

(H5): The functionals Λ_i $(i = 1, \cdots, k)$ are concave.

Then, we have the following result:

THEOREM 2.2. *Let $(x_*(\cdot), u_*(\cdot))$ be an admissible process for (LOCP). Suppose that (H1)-(H5) hold. If there exist $p(\cdot) \in AC([0,1], R^n)$, $\zeta(\cdot) \in L^\infty([0,1], R^m)$, $\lambda_i \geq 0$ $(i = 1, \cdots, k)$ and $\lambda_0 > 0$ such that conditions (2.1)-(2.4) are satisfied, then $(x_*(\cdot), u_*(\cdot))$ is a minimizer for (LOCP).*

REMARK. In Theorem 2.2, the notations ∂ and N stand for the standard subdifferential and normal cone in the sense of convex analysis, respectively. The condition (2.2) is understood as

$$\zeta = \zeta_1 + \zeta_2, \quad \zeta_1 \in \sum_{i=1}^{k} \lambda_i \partial(-\Lambda_i(u_*)), \quad \zeta_2(t) \in N_{U(t)}(u_*(t)) \ a.e.. \tag{2.7}$$

PROOF. In this convex case, the condition (2.1) implies that

$$(\dot{p}(t) + p(t)A(t), -\zeta(t) + p(t)B(t)) \in \sum_{i=1}^{k} \lambda_i \partial L_i[t]. \tag{2.8}$$

Comparing the cost value of an arbitrary admissible process $(x(\cdot), u(\cdot))$ with that of $(x_*(\cdot), u_*(\cdot))$, and by (2.4) we see that

$$\begin{aligned}
&\int_0^1 L_0(t,x,u)dt + h_0(x(1)) - \int_0^1 L_0[t]dt - h_0(x_*(1)) \\
\geq\ & \frac{1}{\lambda_0}\left\{\sum_{i=1}^{k}\lambda_i \int_0^1 (L_i(t,x,u) - L_i[t])dt + \sum_{i=1}^{k}\lambda_i(h_i(x(1)) - h_i(x_*(1)))\right. \\
&\left. -\sum_{i=1}^{k}\lambda_i(\Lambda_i(u) - \Lambda_i(u_*)) + \int_0^1 (p(\dot{x} - Ax - Bu - b) - p(\dot{x}_* - Ax_* - Bu_* - b))dt\right\} \\
=\ & \frac{1}{\lambda_0}\left\{\sum_{i=1}^{k}\lambda_i \int_0^1 (L_i(t,x,u) - L_i[t])dt - \int_0^1 (\dot{p} + pA)(x - x_*)dt\right. \\
& - \int_0^1 (pB - \zeta)(u - u_*)dt + \sum_{i=1}^{k}\lambda_i(h_i(x(1)) - h_i(x_*(1))) + p(1)(x(1) - x_*(1)) \\
& + \sum_{i=1}^{k}\lambda_i(-\Lambda_i(u) + \Lambda_i(u_*)) - \int_0^1 \zeta_1(u - u_*)dt \\
&\left. -p(0)(x(0) - x_*(0)) - \int_0^1 \zeta_2(u - u_*)dt\right\} \\
=:\ & \Delta^*.
\end{aligned}$$

Recall the definitions of subdifferential of convex functions and normal cone of convex sets. From (2.3), (2.7) and (2.8), it follows that $\Delta^* \geq 0$. Thus, (x_*, u_*) is a minimizer for $(NOCP)$. $\square$

3. Value Function

The value functions $V_1(u)$ and $V_2(u,v)$ of $(TDOP)$ may be nonsmooth even if all functions in (D) are smooth. In such case, it is difficult to calculate their Clarke generalized gradients. On the differentiability, we recall the recent results given by Ye and Zhu (1997). To discuss the sufficient optimality conditions for trilevel optimization problems, we will also observe the convexity of the value functions under some assumptions.

For $(\varphi, G_1) : [0,1] \times R^{m_2} \times R^{n_1} \times R^{n_2} \to R^{m_2} \times R$ and $(\psi, G_2) : [0,1] \times R^{m_3} \times R^{n_1} \times R^{n_2} \times R^{n_3} \to R^{m_2} \times R$, we set

$$\begin{aligned}
\varphi^\dagger(t,y,u) &:= \{(\varphi(t,y,u,v), G_1(t,y,u,v)) : v \in U_2(t)\} \\
&\ : [0,1] \times R^{m_2} \times R^{n_1} \to 2^{R^{m_2} \times R} \\
H_1(t,y,p_1,u) &:= \sup_{v \in U_2(t)} \{< p_1, \varphi(t,y,u,v) > -G_1(t,y,u,v)\} \\
&\ : [0,1] \times R^{m_2} \times R^{m_2} \times R^{n_1} \to R \\
\psi^\dagger(t,z,u,v) &:= \{(\psi(t,z,u,v,w),\ G_2(t,z,u,v,w)) : w \in U_3(t)\} \\
&\ : [0,1] \times R^{m_3} \times R^{n_1} \times R^{n_2} \to 2^{R^{m_3} \times R} \\
H_2(t,z,p_2,u,v) &:= \sup_{w \in U_3(t)} \{< p_2, \psi(t,z,u,v,w) > -G_2(t,z,u,v,w)\} \\
&\ : [0,1] \times R^{m_3} \times R^{m_3} \times R^{n_1} \times R^{n_2} \to R.
\end{aligned}$$

To use the results in Ye and Zhu (1997), we assume (A7)-(A8) below.

(A7): $(G_1(t,y,u,v), \varphi(t,y,u,v), \varphi^\dagger(t,y,u), H_1(t,y,p_1,u))$ $((t,y,p_1,u,v) \in [0,1] \times R^{m_2} \times R^{m_2} \times R^{n_1} \times R^{n_2})$ satisfies the following a)-d).

a) The functions G and φ are measurable in t and continuous in (y,u,v). The multifunction $\varphi^\dagger$ has nonempty, compact, convex values.

b) There exists a nonnegative function $\theta_1(t) \in L^\infty$, such that for almost all $t \in [0,1]$,

$$\varphi^\dagger(t,y_1,u_1) \subset \varphi^\dagger(t,y_2,u_2) + \theta_1(t)(|y_1 - y_2| + |u_1 - u_2|)B_{R^{m_2}},$$

for any $y_1, y_2 \in R^{m_2}$, $u_1, u_2 \in R^{n_1}$.

c) For each $u(\cdot) \in \mathcal{U}_1 := \{u(\cdot) \in L^1 : u(t) \in U_1(t) \quad a.e.\}$, there exists a nonnegative function $\rho_u(\cdot) \in L^1[0,1]$, such that $\varphi^\dagger(t,y,u(t)) \subset \rho_u(t)B$ for any $y \in R^{m_2}$ *a.e.*.

d) The partial Clarke generalized gradients $\partial_{(y,p_1)}H_1(t,y,p_1,u)$ and $\partial_u H_1(t,y,p_1,u)$ are upper semicontinuous with respect to (t,y,p_1,u).

(A8): a)-d) above hold with $(G_1(t,y,u,v), \varphi(t,y,u,v), \varphi^\dagger(t,y,u), H_1(t,y,p_1,u))$ replaced by $(G_2(t,z,\bar{u},w), \psi(t,z,\bar{u},w), \psi^\dagger(t,z,\bar{u}), H_2(t,z,p_2,\bar{u}))$ $(\bar{u} = (u,v), (t,z,p_2,\bar{u},w) \in [0,1] \times R^{m_3} \times R^{m_3} \times R^{n_1+n_2} \times R^{n_3})$.

Let $\gamma_1(t) := \int_0^t G_1(t,y(t),u(t),v(t))dt$. Under the assumptions in Lemma 3.1 given below, according to the Filippov's Lemma (see Loewen (1993)), the optimal control problem $P_B[u]$ can be expressed by the following perturbed optimization problem with differential inclusion constraints:

$$\begin{array}{ll} \textit{Minimize}: & g(y(1)) + \gamma_1(1) \\ \textit{subject to}: & (\dot{y}(t), \dot{\gamma}_1(t)) \in \varphi^\dagger(t,y(t),u(t)) \qquad a.e. \\ & y(0) \in C_2. \end{array}$$

Recall a well-known result: if y is an optimal state (corresponding to an optimal control v) for $P_B[u]$, then there exists a Hamiltonian multiplier $p_1 \in AC$ with $(-\dot{p}_1(t), \dot{y}(t)) \in \partial_{(y,p_1)}H_1(t,y(t),p_1(t),u(t))$ *a.e.*, $p_1(0) \in N_{C_2}(y(0))$ and $-p_1(1) \in \partial g_1(y(1))$ (see Clarke (1975)).

We put

$$S_u := \{y \in AC : y \textit{ is an optimal state for } P_B[u]\}$$

$$M_u(y) := \{p_1 \in AC : (-\dot{p}_1(t), \dot{y}(t)) \in \partial_{(y,p_1)}H_1(t,y(t),p_1(t),u(t)) \quad a.e., \\ p_1(0) \in N_{C_2}(y(0)), -p_1(1) \in \partial g_1(y(1))\}$$

$$\partial_u H_1(\cdot,y(\cdot),p_1(\cdot),u(\cdot)) := \{\zeta \in L^\infty; \zeta(t) \in \partial_u H_1(t,y(t),p_1(t),u(t)) \quad a.e\}.$$

Thus, using the results of Ye and Zhu (1997), we obtain

LEMMA 3.1. *Assume that g_1 is locally Lipschitz continuous, (A1) and (A2) hold for $C = C_2$ and $U(t) = U_2(t)$, respectively, and (A7) is satisfied. Then, $V_1(u)$ is locally Lipschitz continuous and*

$$-\partial V_1(u) \subset cl^*co\{\partial_u H_1(\cdot,y(\cdot),p_1(\cdot),u(\cdot)) : \quad y \in S_u, \quad p_1 \in M_u(y)\},$$

where cl^*co denotes the weak convex closure.

Similarly, for the value function $V_2(u,v)$ we have

LEMMA 3.2. *In addition to (A8), assume that g_2 is locally Lipschitz continuous, (A1) and (A2) hold for $C = C_2$ and $U(t) = U_3(t)$, respectively. Then, $V_2(u,v)$ is locally Lipschitz continuous and*

$$-\partial V_2(u,v) \subset cl^* co\{\partial_{(u,v)} H_2(\cdot, z(\cdot), p_2(\cdot), u(\cdot), v(\cdot)) : z \in S_{(u,v)},\quad p_2 \in M_{(u,v)}(z)\},$$

where

$$\partial_{(u,v)} H_2(\cdot, z(\cdot), p_2(\cdot), u(\cdot), v(\cdot)) := \{\zeta \in L^\infty; \zeta(t) \in \partial_{(u,v)} H_2(t, z(t), p_2(t), u(t), v(t)) \;\; a.e.\},$$

$$S_{(u,v)} := \{z \in AC : z \textit{ is an optimal state of } P_C[u,v]\},$$

$$\begin{aligned} M_{(u,v)}(z) := \big\{p_2 \in AC : (-\dot p_2(t), \dot z(t)) \in \partial_{(z,p_2)} H_2(t, z(t), p_2(t), u(t), v(t)) \quad a.e., \\ p_1(0) \in N_{C_2}(y(0)), \quad -p_1(1) \in \partial g_1(y(1))\big\}. \end{aligned}$$

Next, we proceed to the convexity of the following value function:

$$\begin{aligned} V_*(u) := inf \bigg\{ \int_0^1 G(t, y(t), u(t), v(t))dt + g(y(1)) : \dot y\ (t) = A(t)y(t) + B(t)u(t) + \\ D(t)v(t) + b(t) \quad a.e,\ y(0) \in C,\ v(t) \in U(t)\ a.e. \bigg\}, \end{aligned}$$

where $(y,u,v) \in AC([0,1], R^m) \times L^1([0,1], R^{n_1}) \times L^1([0,1], R^{n_2})$, $G : [0,1] \times R^m \times R^{n_1} \times R^{n_2} \to R$, $g : R^m \to R$, $C \subset R^m$, $U(\cdot) : [0,1] \to 2^{R^{n_2}}$, $A(\cdot) : [0,1] \to R^{m\times m}$, $B(\cdot) : [0,1] \to R^{m\times n_1}$, $D(\cdot) : [0,1] \to R^{m\times n_2}$ and $b(\cdot) : [0,1] \to R^m$.

LEMMA 3.3. *Suppose that (H1) and (H2) are satisfied, $A(\cdot)$ is integrable, $B(\cdot)$, $D(\cdot)$, $b(\cdot)$, $G(\cdot, y, u, v)$ are measurable, then the following statements hold.*

(i) If $G(t,\cdot,\cdot,\cdot)$ and $g(\cdot)$ are convex, then $V_*(u)$ is convex.

(ii) If $G(t,\cdot,\cdot,v)$ and $g(\cdot)$ are concave, then $V_*(u)$ is concave.

PROOF. Let u_1, u_2 be arbitrary elements of $L^1[0,1]$, and put $\bar u = \lambda u_1 + (1-\lambda)u_2$ for any $\lambda \in (0,1)$. We define

$$N(u) := \big\{(y,v) \in AC \times L^1 : \dot y = Ay + Bu + Dv + b \quad a.e., y(0) \in C,\ v(t) \in U(t)\ a.e.\big\},$$

$$\Theta(y,u,v) := \int_0^1 G(t, y(t), u(t), v(t))dt + g(y(1)).$$

Then, we have $V_*(u) = \inf\limits_{(y,v)\in N(u)} \Theta(y,u,v)$.

(i) If $N(u_1)$ or $N(u_2)$ is empty, then $\lambda V_*(u_1) + (1-\lambda)V_*(u_2) = +\infty \geq V_*(\bar u)$. Thus we may assume that both $N(u_1)$ and $N(u_2)$ are not empty. For any $(y_1, v_1) \in N(u_1)$, $(y_2, v_2) \in N(u_2)$, let

$$\bar y = \lambda y_1 + (1-\lambda)y_2, \bar v = \lambda v_1 + (1-\lambda)v_2.$$

It is easy to see that $(\bar{y}, \bar{v}) \in N(\bar{u})$. Noting that Θ is convex, we see

$$\lambda\Theta(y_1, u_1, v_1) + (1-\lambda)\Theta(y_2, u_2, v_2) \geq \Theta(\bar{y}, \bar{u}, \bar{v}) \geq \inf_{(y,v)\in N(\bar{u})} \Theta(y, \bar{u}, v).$$

Hence, we have

$$\begin{aligned} \lambda V_*(u_1) + (1-\lambda)V_*(u_2) &= \lambda \inf_{(y,v)\in N(u_1)} \Theta(y, u_1, v) + (1-\lambda) \inf_{(y,v)\in N(u_2)} \Theta(y, u_2, v) \\ &\geq \inf_{(y,v)\in N(\bar{u})} \Theta(y, \bar{u}, v) = V_*(\bar{u}). \end{aligned}$$

(ii) If $N(\bar{u}) = \emptyset$, then $V_*(\bar{u}) = +\infty \geq \lambda V_*(u_1) + (1-\lambda)V_*(u_2)$. Thus, we may assume $N(\bar{u}) \neq \emptyset$. For every $(\tilde{y}, \tilde{v}) \in N(\bar{u})$, let $\tilde{y}_i$ be a solution of following equation,

$$\begin{cases} \dot{y}(t) = A(t)y(t) + B(t)u_i(t) + D(t)\tilde{v}(t) + b(t) \quad a.e \\ y(0) = \tilde{y}(0) \end{cases} \qquad (i = 1, 2).$$

Then, we have

$$\tilde{y} = \lambda\tilde{y}_1 + (1-\lambda)\tilde{y}_2, \ (\tilde{y}_i, \tilde{v}) \in N(u_i) \ (i = 1, 2).$$

By concavity of G and g, we see that

$$\begin{aligned} \Theta(\tilde{y}, \bar{u}, \tilde{v}) &\geq \lambda\Theta(\tilde{y}_1, u_1, \tilde{v}) + (1-\lambda)\Theta(\tilde{y}_2, u_2, \tilde{v}) \\ &\geq \lambda \inf_{(y,v)\in N(u_1)} \Theta(y, u_1, v) + (1-\lambda) \inf_{(y,v)\in N(u_2)} \Theta(y, u_2, v), \end{aligned}$$

which implies that

$$\begin{aligned} V_*(\bar{u}) = \inf_{(y,v)\in N(\bar{u})} \Theta(y, \bar{u}, v) &\geq \lambda \inf_{(y,v)\in N(u_1)} \Theta(y, u_1, v) + (1-\lambda) \inf_{(y,v)\in N(u_2)} \Theta(y, u_2, v) \\ &\geq \lambda V_*(u_1) + (1-\lambda)V_*(u_2). \end{aligned}$$

We have therefore proved this lemma. □

4. Optimality Conditions for (TDOP)

In this section we derive the optimality conditions for the trilevel dynamic optimization problem $(TDOP)$. A simple example will be given in Section 5.

In Theorem 4.1 and Corollary 4.2 given below, ∂ indicates the Clarke generalized gradient and $N_{C_i}, N_{U_i(t)}$ $(i = 1, 2, 3)$ denote the Clarke normal cones, while in Theorem 4.3, these stand for the subdifferential and the normal cones in the sense of convex analysis, respectively.

First we derive necessary optimality conditions for $(TDOP)$.

Let $(x_*, y_*, z_*; u_*, v_*, w_*) = (\tilde{x}_*; \tilde{u}_*)$ be a local optimal solution of the trilevel optimality problem $(TDOP)$, i.e. $(\tilde{x}_*; \tilde{u}_*)$ is a local minimizer for the optimal control problem P. Notice that P can be easily written in the form $(NOCP)$ with

$$(D^*):\quad \begin{array}{l} \tilde{x}(\cdot) := (x(\cdot), y(\cdot), z(\cdot)) \in AC([0,1], R^m), \quad m = m_1 + m_2 + m_3 \\ \tilde{u}(\cdot) := (u(\cdot), v(\cdot), w(\cdot)) \in L^1([0,1], R^n), \quad n = n_1 + n_2 + n_3 \\ C := C_1 \times C_2 \times C_3, \quad U(t) := U_1(t) \times U_2(t) \times U_3(t) \\ \Phi(t, \tilde{x}, \tilde{u}) := (\phi(t, x, \tilde{u}), \varphi(t, y, u, v), \psi(t, z, \tilde{u})) \\ L_0(t, \tilde{x}, \tilde{u}) := F(t, x, \tilde{u}), \quad h_0(\tilde{x}(0)) := f(x(0)) \\ L_1(t, \tilde{x}, \tilde{u}) := G_1(t, y, u, v), \quad h_1(\tilde{x}(0)) := g_1(y(0)) \\ L_2(t, \tilde{x}, \tilde{u}) := G_2(t, z, \tilde{u}), \quad h_2(\tilde{x}(0)) := g_2(z(0)) \\ \Lambda_1(\tilde{u}) := V_1(u), \quad \Lambda_2(\tilde{u}) := V_2(u, v). \end{array}$$

If the assumptions in Theorem 2.1 are satisfied, then there exist an absolutely continuous function $p(\cdot) : [0,1] \to R^{m_1} \times R^{m_2} \times R^{m_3}$, a measurable essentially bounded function $\zeta(\cdot) : [0,1] \to R^{n_1} \times R^{n_2} \times R^{n_3}$ and $\lambda_0, \lambda_1, \lambda_2 \geq 0$ with $\sum_{i=0}^{2} \lambda_i = 1$, such that

$$(-\dot{p}(t), \zeta(t)) \in \nabla_{(\tilde{x},\tilde{u})} H(t, \tilde{x}_*, \tilde{u}_*) - \partial_{(\tilde{x},\tilde{u})} E(t, \tilde{x}_*, \tilde{u}_*), \tag{4.1}$$

$$\zeta = \mu_1 + \mu_2, \ -\mu_1 \in \lambda_1 \partial_{\tilde{u}} V_1(u_*) + \lambda_2 \partial_{\tilde{u}} V_2(u_*, v_*), \ \mu_2(t) \in N_{U(t)}(\tilde{u}_*(t)) \quad a.e.,$$

$$p(0) \in N_C(\tilde{x}_*(0)), \ -p(1) \in \lambda_0 \partial_{\tilde{x}} f(x_*(1)) + \lambda_1 \partial_{\tilde{x}} g_1(y_*(1)) + \lambda_2 \partial_{\tilde{x}} g_2(z_*(1)).$$

Here, we set

$$H(t, \tilde{x}, \tilde{u})) := \langle p_1, \phi(t, x, u, v, w) \rangle + \langle p_2, \varphi(t, y, u, v) \rangle + \langle p_3, \psi(t, z, u, v, w) \rangle,$$

$$E(t, \tilde{x}, \tilde{u}) := \lambda_0 F(t, x, u, v, w) + \lambda_1 G_1(t, y, u, v) + \lambda_2 G_2(t, z, u, v, w).$$

Now we arrive at the following conclusion.

THEOREM 4.1. *Suppose that* $(x_*, y_*, z_*; u_*, v_*, w_*)$ *is a local minimizer for (TDOP). Let assumptions (A1)-(A6) hold for the data* $\{C,\ U(t),\ \Phi,\ L_i,\ h_i\ (i = 0,1,2),\ \Lambda_i\ (i = 1,2)\}$ *in* (D^*) *associated with* $(\tilde{x}_*; \tilde{u}_*)$. *Then,*

(i) there exist $p(\cdot) = (p_1(\cdot), p_2(\cdot), p_3(\cdot)) \in AC$, $\zeta(\cdot) = (\zeta_1(\cdot), \zeta_2(\cdot), \zeta_3(\cdot)) \in L^\infty$ *and* $\lambda_0, \lambda_1, \lambda_2 \geq 0$ *with* $\sum_{i=0}^{2} \lambda_i = 1$, *such that (4.1) and the following (4.2)-(4.5) hold.*

$$\begin{array}{l} (\zeta_1, \zeta_2, \zeta_3) = (\bar{\zeta}_1 + \hat{\zeta}_1 + \tilde{\zeta}_1, \hat{\zeta}_2 + \tilde{\zeta}_2, \zeta_3) \\ -\bar{\zeta}_1 \in \lambda_1 \partial_u V_1(u_*), \ -(\hat{\zeta}_1, \hat{\zeta}_2) \in \lambda_2 \partial_{(u,v)} V_2(u_*, v_*) \end{array} \tag{4.2}$$

$$(\tilde{\zeta}_1, \tilde{\zeta}_2, \zeta_3) \in N_{U_1(t)}(u_*(t)) \times N_{U_2(t)}(v_*(t)) \times N_{U_3(t)}(w_*(t)) \quad a.e. \tag{4.3}$$

$$(p_1(0), p_2(0), p_3(0)) \in N_{C_1}(x_*(0)) \times N_{C_2}(y_*(0)) \times N_{C_3}(z_*(0)) \tag{4.4}$$

$$-\,(p_1(1), p_2(1), p_3(1)) \in \lambda_0 \partial_x f(x_*(1)) \times \lambda_1 \partial_y g_1(y_*(1)) \times \lambda_1 \partial_z g_2(z_*(1)). \tag{4.5}$$

(ii) If the constraint qualifications (2.5) and (2.6) in Theorem 2.1 hold for the data $\{\Lambda_i,\ L_i,\ h_i\ \ (i = 1,2), \Phi, C, U(t)\}$ *in* (D^*), *then* $\lambda_0 > 0$.

(iii) If F, G_1, G_2 *are convex in* $(\tilde{x};\tilde{u})$, *then (4.1) implies that*

$$\begin{aligned}
&-\dot{p}_1(t) \in p_1(t)\cdot\nabla_x\varphi(t,x_*(t),\tilde{u}_*(t)) - \lambda_0\partial_x F(t,x_*(t),\tilde{u}_*(t))\\
&-\dot{p}_2(t) \in p_2(t)\cdot\nabla_y\varphi(t,y_*(t),u_*(t),v_*(t)) - \lambda_1\partial_y G_1(t,y_*(t),u_*(t),v_*(t))\\
&-\dot{p}_3(t) \in p_3(t)\cdot\nabla_z\psi(t,z_*(t),\tilde{u}_*(t)) - \lambda_2\partial_z G_2(t,z_*(t),\tilde{u}_*(t))\\
&\zeta_1(t) \in \nabla_u H(t,\tilde{x}_*(t),\tilde{u}_*(t)) - \partial_u E(t,\tilde{x}_*(t),\tilde{u}_*(t))\\
&\zeta_2(t) \in \nabla_v H(t,\tilde{x}_*(t),\tilde{u}_*(t)) - \partial_v E(t,\tilde{x}_*(t),\tilde{u}_*(t))\\
&\zeta_3(t) \in p_1(t)\cdot\nabla_w\phi(t,x_*(t),\tilde{u}_*(t)) + p_3(t)\cdot\nabla_w\psi(t,z_*(t),\tilde{u}_*(t))\\
&\qquad\qquad -\lambda_0\partial_w F(t,x_*(t),\tilde{u}_*(t)) - \lambda_2\partial_w G_2(t,z_*(t),\tilde{u}_*(t)).
\end{aligned}$$

Combining Theorem 4.1, Lemma 3.1 and Lemma 3.2, we obtain

COROLLARY 4.2. *Assume that* $(x_*,y_*,z_*;u_*,v_*,w_*)$ *be a local minimizer for the (TDOP), the assumptions (A1)-(A5) hold for the data* $\{C, U(t), \Phi, L_i, h_i,\ i=0,1,2\}$ *in* (D^*) *associated with* $(\tilde{x}_*;\tilde{u}_*)$. *Let (A7) and (A8) hold. Then there exist* $(p_1(\cdot),p_2(\cdot),p_3(\cdot)) \in AC$, $(\zeta_1(\cdot),\zeta_2(\cdot),\zeta_3(\cdot)) \in L^\infty$ *and* $\lambda_0, \lambda_1, \lambda_2 \geq 0$ *with* $\sum_{i=1}^2\lambda_i = 1$, *such that (4.1), (4.3)-(4.5) and the following (4.6) hold.*

$$\begin{aligned}
&(\zeta_1,\zeta_2,\zeta_3) = (\bar{\zeta}_1 + \hat{\zeta}_1 + \tilde{\zeta}_1, \hat{\zeta}_2 + \tilde{\zeta}_2, \zeta_3)\\
&-\bar{\zeta}_1 \in cl^*co\{\partial_u H_1(\cdot,\alpha(\cdot),q(\cdot),u_*(\cdot)) : \alpha \in S_{u_*},\ q \in M_{u_*}(\alpha)\}\\
&-(\hat{\zeta}_1,\hat{\zeta}_2) \in cl^*co\left\{\partial_{(u,v)} H_2(\cdot,\alpha(\cdot),q(\cdot),u_*(\cdot),v_*(\cdot)) : \alpha \in S_{(u_*,v_*)},\ q \in M_{(u_*,v_*)}(\alpha)\right\}
\end{aligned} \tag{4.6}$$

Now we replace the state equations in $(TDOP)$ by linear systems

$$\dot{x}(t) = A_1(t)x(t) + B_1(t)u(t) + D_1(t)v(t) + E_1(t)w(t) + b_1(t) \quad a.e.$$

$$\dot{y}(t) = A_2(t)y(t) + B_2(t)u(t) + D_2(t)v(t) + b_2(t) \quad a.e.$$

$$\dot{z}(t) = A_3(t)z(t) + B_3(t)u(t) + D_3(t)v(t) + E_3(t)w(t) + b_3(t) \quad a.e.$$

where $A_i(t) : [0,1] \to R^{m_i\times m_i}$, $B_i(t) : [0,1] \to R^{m_i\times n_1}$, $D_i(t) : [0,1] \to R^{m_i\times n_2}$ $(i = 1,2,3)$ and $E_i(t) : [0,1] \to R^{m_i\times n_3}$ $(i = 1,3)$ are integrable, $b_i(t) : [0,1] \to R^{m_i}$ $(i = 1,2,3)$ are measurable. We denote this problem by $(TDOP^*)$.

Notice that the condition (2.4) always hold for P. From Theorem 2.2 and Lemma 3.3 we get directly the following sufficient optimality conditions for $(TDOP^*)$.

THEOREM 4.3. *Assume that the control* (u_*,v_*,w_*) *corresponding to* (x_*,y_*,z_*) *is admissible for (TDOP*). Let the data* $\{C, U(t), L_i,\ i=0,1,2\}$ *in* (D^*) *satisfy (H1)-(H3), respectively,* $G_1(t,\cdot,\cdot,v)$, $G_2(t,\cdot,\cdot,\cdot,w)$ *be concave, and* $g_1(\cdot)$, $g_2(\cdot)$ *be affine. If there exist* $(p_1,p_2,p_3)(\cdot) \in AC$, $(\zeta_1,\zeta_2,\zeta_3)(\cdot) \in L^\infty$ *and* $\lambda_0 > 0$, $\lambda_1, \lambda_2 \geq 0$ *such that (4.1)-(4.5) hold, then* $(x_*,y_*,z_*;u_*,v_*,w_*)$ *is an optimal solution of (TDOP*).*

REMARK. For the trilevel optimization problem $(TDOP)$, sometimes, the player A has to guarantee that the costs of the players B and C are not too large. The problem

P_A of $(TOCP)$ with such constraints is stated as follows,

$$\begin{array}{ll} Minimize: & \int_0^1 F(t,x(t),u(t),v(t),w(t))dt + f(x(1)) \\ subject\ to: & \dot{x}(t) = \phi(t,x(t),u(t),v(t),w(t)) \qquad a.e. \\ & x(0) \in C_1 \\ & u(t) \in U_1(t) \qquad a.e. \\ & v \in \mathcal{V}[u] \\ & w \in \mathcal{W}[u,v] \\ & \int_0^1 G_1(t,y(t),u(t),v(t))dt + g_1(y(1)) \leq M_1 \\ & \int_0^1 G_2(t,z(t),u(t),v(t),w(t))dt + g_2(z(1)) \leq M_2, \end{array}$$

where M_1, $M_2 > 0$. From the results in Section 2 and 3, we can also get optimality conditions for this trilevel optimization problem.

5. Proof of Theorem 2.1

PROOF. Let us omit the variable t when this does not cause confusion. We put

$$M := \left\{(x,u) \in AC \times L^1 : \dot{x}(t) = \Phi(t,x(t),u(t)) \ a.e, \quad x(0) \in C, \quad u(t) \in U(t) \ a.e.\right\},$$

$$\Delta((x_1,u_1),(x_2,u_2)) := \|u_1 - u_2\|_{L^1} + |x_1(0) - x_2(0)|.$$

It is evident that Δ is a metric, and M is a complete metric space with respect to Δ.

For given $\epsilon > 0$, Let

$$\Gamma_\epsilon(x,u) := max\{\bar{G}_1(x,u), \cdots, \bar{G}_k(x,u), F(x,u) - F(x_*,u_*) + \epsilon^2\},$$

$$F(x,u) := \int_0^1 L_0(t,x(t),u(t))dt + h_0(x(1)).$$

Then, we see that $\Gamma_\epsilon(x,u) > 0$ for all $(x,u) \in M$. It follows that

$$\Gamma_\epsilon(x_*,u_*) \leq \inf_{(x,u)\in M} \Gamma_\epsilon(x,u) + \epsilon^2.$$

Thus, by the Ekeland Variational principle, we have

LEMMA 5.1. *There is an element* $(x_0,u_0) \in M$, *such that* (x_0,u_0) *minimizes*

$$\Gamma_\epsilon(x,u) + \epsilon\Delta((x,u),(x_0,u_0))$$

over all $(x,u) \in M$, *and*

$$\Delta((x_*,u_*),(x_0,u_0)) \leq \epsilon, \qquad \Gamma_\epsilon(x_0,u_0) \leq \epsilon^2. \tag{5.1}$$

The following Lemma 5.2 will be used to derive Lemma 5.3.

LEMMA 5.2. *Let* $\mathcal{U} = \{u \in L^1[0,1] : u(t) \in U(t) \quad a.e.\}$. *If* $u_0(\cdot) \in L^1$, *then*

$$\int_0^1 \inf_{v \in U(t)} |u_0(t) - v|\, dt = \inf_{u \in \mathcal{U}} \int_0^1 |u_0(t) - u(t)|\, dt. \tag{5.2}$$

In fact, by the measurability theorem (see Loewen (1993)), there is a sequence of Lebesgue measurable functions v_n such that $U(t) = cl\{v_n(t) : n = 1, 2, \cdots\}$.

Let $\Psi(t) = \inf_{v \in U(t)} |u_0(t) - v|$. For any $\epsilon^\circ > 0$, there exist $v_t \in U(t)$ and $n_0 \in N$, such that

$$\Psi(t) + \epsilon^\circ > |u_0(t) - v_0(t)|, \ |v_0(t) - v_{n_0}(t)| < \epsilon^\circ.$$

Then, we have

$$\begin{aligned} \inf_{n \in N} |u_0(t) - v_n(t)| &\geq \inf_{v \in U(t)} |u_0(t) - v| > |u_0(t) - v_0(t)| - \epsilon^\circ \geq |u_0(t) - v_{n_0}(t)| - 2\epsilon^\circ \\ &\geq \inf_{n \in N} |u_0(t) - v_n(t)| - 2\epsilon^\circ, \end{aligned}$$

which implies that

$$0 \leq \inf_{n \in N} |u_0(t) - v_n(t)| - \Psi(t) \leq 2\epsilon^\circ.$$

Thus, $\Psi(t) = \inf_{n \in N} |u_0(t) - v_n(t)|$. It is obvious that, for every $u(\cdot) \in \mathcal{U}$,

$$\int_0^1 \inf_{v \in U(t)} |u_0(t) - v|\, dt \leq \int_0^1 |u_0(t) - u(t)|\, dt. \tag{5.3}$$

For $\bar{\epsilon} > 0$, consider a multifunction

$$\Omega(t) := \left\{ u \in U(t); |u_0(t) - u| \leq \inf_{v \in U(t)} |u_0(t) - v| + \frac{1}{2}\bar{\epsilon} \right\},$$

which is measurable with closed nonempty values. By the Measurable selections Theorem (see Loewen (1993)), Ω admits a measurable selection $\bar{u}(\cdot)$. Since $\bar{u}$ is measurable and $\bar{u}(t) \in \Omega(t)$ a.e.. We see that $\bar{u}(\cdot) \in \mathcal{U}$, and

$$\int_0^1 |u_0(t) - \bar{u}(t)|\, dt < \int_0^1 \inf_{v \in U(t)} |u_0(t) - v|\, dt + \bar{\epsilon}. \tag{5.4}$$

Combining (5.3) and (5.4), we get (5.2).

LEMMA 5.3. *The* (x_0, u_0) *above is a local minimizer for the function.*

$$\begin{aligned} \Lambda(x,u) := \quad & \Gamma_\epsilon(x,u) + \epsilon\Delta((x,u),(x_0,u_0)) + K_2 \int_0^1 |\dot{x}(t) - \Phi(t, x(t), u(t))|\, dt \\ & + K_1 d_C(x(0)) + K_1 \int_0^1 d_{U(t)}(u(t))dt, \end{aligned}$$

where $d_C(x(0))$, $d_{U(t)}(u(t))$ *denote the distances of the points* $x(0)$ *and* $u(t)$ *to the sets* C *and* $U(t)$, *respectively, and*

$$\begin{aligned} & K_1 = \bar{K}(KK' + 1) + \epsilon), \ K_2 = \bar{K}K, \\ & \bar{K} = max\{\|k_L\|_{L^\infty} + k_h, \|k_L\|_{L^\infty} + k_\Lambda\}, \\ & K = \exp \int_0^1 k(t)dt, \quad K' = \|k\|_{L^\infty}. \end{aligned}$$

Suppose this lemma to be false. Then, there exists a sequence $\{(x_i, u_i)\}$ converging to (x_0, u_0) with $\Lambda(x_i, u_i) < \Lambda(x_0, u_0)$. Let $\Lambda(x_i, u_i) = \Lambda(x_0, u_0) - 2\epsilon_i K_1$ (where $\epsilon_i > 0$). By Lemma 5.2 there is $\tilde{u}_i(\cdot) \in \mathcal{U}$ such that

$$\|\tilde{u}_i - u_i\|_{L^1} \leq \int_0^1 \inf_{v \in U(t)} |u_i(t) - v| \, dt + \epsilon_i. \tag{5.5}$$

Since C is closed, there are $c_i \in C$ $(i = 0 \cdots k)$ such that $d_C(x_i(0)) = |x_i(0) - c_i|$. Let $\bar{x}_i(t) = x_i(t) - x_i(0) + c_i$, thus for sufficiently large i, it holds that

$$\begin{aligned} &\int_0^1 \left|\dot{\bar{x}}_i \,(t) - \Phi(t, \bar{x}_i(t), \tilde{u}_i(t))\right| dt \\ &\leq \int_0^1 |\dot{x}_i(t) - \Phi(t, x_i(t), u_i(t))| \, dt + K'(\|\tilde{u}_i - u_i\|_{L^1} + |x_i(0) - c_i|). \end{aligned} \tag{5.6}$$

Let $\tilde{x}_i(t)$ be a solution of the following integral equation

$$y(t) = \bar{x}_i(0) + \int_0^t \Phi(t, y(t), \tilde{u}_i(t)) dt.$$

We see that $(\tilde{x}_i, \tilde{u}_i) \in M$ with $\tilde{x}_i(0) = \bar{x}_i(0)$ and

$$\begin{aligned} |\tilde{x}_i(t) - \bar{x}_i(t)| &\leq \left|\int_0^t (\Phi(t, \tilde{x}_i, \tilde{u}_i) - \dot{\bar{x}}_i) dt\right| \\ &\leq \left|\int_0^t (\Phi(t, \bar{x}_i, \tilde{u}_i) - \dot{\bar{x}}_i) dt\right| + \left|\int_0^t (\Phi(t, \tilde{x}_i, \tilde{u}_i) - \Phi(t, \bar{x}_i, \tilde{u}_i)) dt\right| \\ &\leq \left|\int_0^t \Phi(t, \bar{x}_i, \tilde{u}_i) - \dot{\bar{x}}_i) dt\right| + \left|\int_0^t k(t) \, |\tilde{x}_i(t) - \bar{x}_i(t)| \, dt\right|. \end{aligned}$$

It follows that

$$\|\tilde{x}_i - \bar{x}_i\|_{L^\infty} \leq K \int_0^1 \left|\dot{\bar{x}}_i \,(t) - \Phi(t, \bar{x}_i(t), \tilde{u}_i(t))\right| dt. \tag{5.7}$$

(5.6) and (5.7) lead to

$$\begin{aligned} \|\tilde{x}_i - x_i\|_{L^\infty} \leq \; & K \int_0^1 |\dot{x}_i(t) - \Phi(t, x_i(t), u_i(t))| \, dt + KK' \, \|\tilde{u}_i - u_i\|_{L^1} \\ & + (KK' + 1) \, |x_i(0) - c_i| \,. \end{aligned} \tag{5.8}$$

From (A3), (5.8) and (5.5), we have

$$\begin{aligned} & \Gamma_\epsilon(\tilde{x}_i, \tilde{u}_i) + \epsilon\Delta((\tilde{x}_i, \tilde{u}_i), (x_0, u_0)) \\ \leq \; & \Gamma_\epsilon(x_i, u_i) + \epsilon\Delta((x_i, u_i), (x_0, u_0)) + \bar{K}(\|\tilde{u}_i - u_i\|_{L^1} + \|\tilde{x}_i - x_i\|_C) \\ & + \epsilon(\|\tilde{u}_i - u_i\|_{L^1} + |x_i(0) - c_i|) \\ \leq \; & \Gamma_\epsilon(x_i, u_i) + \epsilon\Delta((x_i, u_i), (x_0, u_0)) + \bar{K}K \int_0^1 |\dot{x}_i(t) - \Phi(t, x_i(t), u_i(t))| \, dt \\ & + (\bar{K}(KK' + 1) + \epsilon)\left(\int_0^1 \inf_{v \in U(t)} |u_i(t) - v| \, dt + \epsilon_i + |x_i(0) - c_i|\right) \\ = \; & \Lambda(x_i, u_i) + \epsilon_i K_1 \\ < \; & \Lambda(x_0, u_0) \\ = \; & \Gamma_\epsilon(x_0, u_0), \end{aligned}$$

which contradicts Lemma 5.1, so Lemma 5.3 holds.

By Lemma 5.3, we know $0 \in \partial \Lambda(x_0, u_0)$, i.e.

$$\begin{aligned} 0 \in \partial \Gamma_\epsilon(x_0, u_0) + \partial \epsilon \Delta((x_0, u_0), (x_0, u_0)) + \partial K_1 d_C(x_0(0)) + \\ \partial K_1 \int_0^1 d_{U(t)}(u_0(t)) dt + \partial K_2 \int_0^1 |\dot{x}_0(t) - \Phi(t, x_0(t), u_0(t))| \, dt. \end{aligned} \tag{5.9}$$

According to the formulas of the generalized gradients (see Clarke (1983)), we have the following.

(a) For every $\xi \in \partial \Gamma_\epsilon(x_0, u_0)$, there exist functions ξ_i, η_i $(i = 0, \cdots, k)$ with $(\xi_i, \eta_i)(t) \in \partial L_i(t, x_0(t), u_0(t))$ *a.e.* and $\nu_i \in \partial h_i(x_0(1))$ $(i = 0, \cdots, k)$, $\bar{\eta}_i \in -\partial \Lambda_i(u_0)$ $(i = 1, \cdots, k)$, $\bar{\lambda}_i \geq 0$ for $i \in I_0(x_0, u_0)$ with $\sum_{i \in I(x_0, u_0)} \bar{\lambda}_i = 1$, such that for any $(x, u) \in AC \times L^1$

$$\xi(x, u) = \sum_{i=0}^{k} \bar{\lambda}_i \int_0^1 \{\langle \xi_i, x \rangle + \langle \eta_i, u \rangle\} ds + \sum_{i=0}^{k} \bar{\lambda}_i \langle \nu_i, x(1) \rangle + \sum_{i=1}^{k} \bar{\lambda}_i \int_0^1 \langle \bar{\eta}_i, u \rangle \, ds,$$

where, $I_0(x_0, u_0) := \{i \in [0, \cdots, k] : \bar{G}_i(x_0, u_0) = \Gamma_\epsilon(x_0, u_0)\}$, $\bar{G}_0(x_0, u_0) := F(x, u) - F(x_*, u_*) + \epsilon^2$, and $\bar{\lambda}_i := 0$ for $i \notin I(x_0, u_0)$.

(b) For every $\xi \in \partial \epsilon \Delta((x_0, u_0), (x_0, u_0))$, there are function θ_2 with $\theta_2(t) \in \epsilon B_{R^n}$ a.e. and $\theta_1 \in \epsilon B_{R^m}$ such that for every $(x, u) \in AC \times L^1$

$$\xi(x(0), u) = \langle \theta_1, x(0) \rangle + \int_0^1 \langle \theta_2, u \rangle \, ds.$$

(c) Every $\xi \in \partial K_1 d_C(x_0(0))$ corresponds to a mapping $r \in K_1 \partial d_C(x_0(0))$ with

$$\xi(x) = \langle r, x(0) \rangle \qquad \textit{for any } x \in AC.$$

(d) For every $\xi \in \partial K_1 \int_0^1 d_{U(t)}(u_0(t)) dt$, there is a function $\bar{\eta}$ with $\bar{\eta}(t) \in K_1 \partial d_{U(t)}(u_0(t))$ *a.e.* such that

$$\xi(u) = \int_0^1 \langle \bar{\eta}, u \rangle \, ds \qquad \textit{for any } u \in L^1.$$

(e) Finally, for every $\xi \in \partial K_2 \int_0^1 |\dot{x}_0(t) - \Phi(t, x_0(t), u_0(t))| \, dt$, there exists

$$(\bar{p}, \mu_1, \mu_2) \in \partial_{(\dot{x}, x, u)} K_2 \, |\dot{x}_0(t) - \Phi(t, x_0(t), u_0(t))| \tag{5.10}$$

such that

$$\xi(x, u) = \int_0^1 \{\langle \bar{p}, \dot{x} \rangle + \langle \mu_1, x \rangle + \langle \mu_2, u \rangle\} dt \qquad \textit{for any } (x, u) \in AC \times L^1.$$

Since $\partial_{(\dot{x}, x, u)}(\dot{x} - \Phi(x, u)) = \{(1, -\Phi_x, -\Phi_u)\}$, by Jacobian Chain Rule, (5.10) implies that

$$(\mu_1, \mu_2) = -\nabla_{(x, u)} \langle \bar{p}, \Phi(x_0, u_0) \rangle . \tag{5.11}$$

Then, we arrive at

LEMMA 5.4. *There exists* $\bar\lambda_i$, ξ_i, η_i, ν_i $(i = 0 \cdots k)$, $\bar\eta_i$ $(i = 1 \cdots k)$, $\bar\eta$, θ_1, θ_2, r, $\bar p$, μ_1, μ_2 *stated in the above, such that for any* $(x, u) \in AC \times L^1$

$$\begin{aligned} 0 = \; & \sum_{i=0}^{k} \bar\lambda_i \int_0^1 \{\langle \xi_i, x\rangle + \langle \eta_i, u\rangle\} dt + \sum_{i=0}^{k} \bar\lambda_i \langle \nu_i, x(1)\rangle + \sum_{i=1}^{k} \bar\lambda_i \int_0^1 \langle \bar\eta_i, u\rangle \, dt \\ & + \langle \theta_1, x(0)\rangle + \int_0^1 \langle \theta_2, u\rangle \, dt + \langle r, x(0)\rangle + \int_0^1 \langle \bar\eta, u\rangle \, dt \\ & \int_0^1 \{\langle \bar p, \dot x\rangle + \langle \mu_1, x\rangle + \langle \mu_2, u\rangle\} dt. \end{aligned}$$

Separating x and u, from the above equation we get

$$\int_0^1 \left\langle \sum_{i=0}^{k} \bar\lambda_i \xi_i + \mu_1, x \right\rangle dt + \int_0^1 \langle \bar p, \dot x\rangle \, dt + \left\langle \sum_{i=0}^{k} \bar\lambda_i \nu_i, x(1) \right\rangle + \langle \theta_1 + r, x(0)\rangle = 0, \quad (5.12)$$

$$\int_0^1 \left\langle \sum_{i=0}^{k} \bar\lambda_i \eta_i + \mu_2 + \sum_{i=1}^{k} \bar\lambda_i \bar\eta_i + \bar\eta + \theta_2, u \right\rangle ds = 0 \qquad for \;\; any \;\; u \in L^1. \quad (5.13)$$

Hence, (5.13) shows that

$$\sum_{i=0}^{k} \bar\lambda_i \eta_i(t) + \mu_2(t) = -\sum_{i=1}^{k} \bar\lambda_i \bar\eta_i(t) - \bar\eta(t) - \theta_2 \quad a.e.. \quad (5.14)$$

According to Dubois-Reymond Lemma (see Hestenes (1980)), using standard variational arguments, from (5.12), we observe that

$$\bar p(t) = \int_0^t (\sum_{i=0}^{k} \bar\lambda_i \xi_i + \mu_1) ds + r + \theta_1 \quad (5.15)$$

$$\bar p(1) = -\sum_{i=0}^{k} \bar\lambda_i \nu_i \in -\sum_{i=0}^{k} \bar\lambda_i \partial h_i(x_0(1)) \quad (5.16)$$

$$\bar p(0) = r + \theta_1 \in K_1 \partial d_C(x_0(0)) + \theta_1. \quad (5.17)$$

By (5.15), we see that $\dot{\bar p}(t) - \sum_{i=0}^{k} \bar\lambda_i \xi_i = \mu_1$. Letting $\bar\zeta = -\sum_{i=0}^{k} \bar\lambda_i \eta_i - \mu_2$, from (5.11), we have

$$(-\dot{\bar p}(t), \bar\zeta) \in \nabla_{(x,u)} \langle \bar p, \Phi(x_0, u_0)\rangle - \sum_{i=1}^{k} \bar\lambda_i \partial_{(x,u)} L_i(t, x_0, u_0) \;\; a.e.. \quad (5.18)$$

Letting $\bar\zeta_1 = \sum_{i=1}^{k} \bar\lambda_i \bar\eta_i + \theta_2$ and $\bar\zeta_2 = \bar\eta$, by (5.14), (a) and (d), we get

$$\bar\zeta = \bar\zeta_1 + \bar\zeta_2, \;\; \bar\zeta_1 \in -\sum_{i=1}^{k} \lambda_i \partial\Lambda_i(u_0(\cdot)) + \theta_2, \;\; \bar\zeta_2(t) \in N_{U(t)}(u_0(t)) \;\; a.e.. \quad (5.19)$$

Notice that $\bar p$, $\bar\zeta$, $\bar\lambda_i$, $i = 0, \cdots, k$ depend on ϵ. From (5.1), we know that $(x_0, u_0) \to (x_*, u_*)$ as $\epsilon \to 0$. Letting $\epsilon \to 0$, as in the proof of Theorem 3.2.6 in Clarke (1983), from (5.16)-(5.19), it follows that there are $p \in AC$, $\zeta \in L^{\infty}$ and $\lambda_i \geq 0$. $i = 0, \cdots, k$ with $\sum_{i=0}^{k} \lambda_i = 1$ such that (2.1)-(2.3) hold.

Observe that for sufficient small $\epsilon > 0$, if $i_0 \in \left\{ i \in [1, \cdots, k] : \ \bar{G}_i(x_*, u_*) < 0 \right\}$, then $\bar{G}_{i_0}(x_0, u_0) < 0$. It is easy to see that $i_0 \notin I_0(x_0, u_0)$. Hence, we have $\lambda_{i_0} = 0$, which yields (2.4).

Finally, let (2.5) and (2.6) hold. If $\lambda_0 = 0$, then $\sum_{i=1}^{k} \lambda_i = 1$, (2.1) and (2.3) imply that

$$(\dot{p} + p\Phi_x[t], -\zeta + p\Phi_u[t]) \in \sum_{i=1}^{k} \lambda_i \partial_{(x,u)} L_i[t], \quad -p(1) \in \sum_{i=1}^{k} \lambda_i \partial h_i(x_*(1)). \tag{5.20}$$

Then, from (2.2), (2.4)-(2.6) and (5.20), it follows that

$$\begin{aligned}
0 \quad &> \sum_{i=1}^{k} \lambda_i \left(\int_0^1 L_i{}^{\circ}((x_*, u_*), (x_1, u_1))(t) dt + h_i{}^{\circ}(x_*, x_1) + (-\Lambda_i)_i^{\circ}(u_*, u_1) \right) \\
&\geq \int_0^1 (\dot{p} x_1 + p\Phi_x[t] x_1 - \zeta u_1 + p\Phi_u[t] u_1) dt - p(1) x_1(1) + \int_0^1 \zeta_1 u_1 dt \\
&= \int_0^1 (-p\dot{x}_1 + p\Phi_x[t] x_1 + p\Phi_u[t] u_1) dt - p(0) x_1(0) - \int_0^1 \zeta_2 u_1 dt \\
&= -p(0) x_1(0) - \int_0^1 \zeta_2 u_1 dt \\
&\geq 0,
\end{aligned}$$

which is a contradiction. Thus, we have $\lambda_0 \neq 0$, which completes the proof. □

Now, we give a simple example.

EXAMPLE 5.5. Let us consider the following trilevel dynamic optimization problem.

$$\begin{aligned}
P_1: \quad & \mathit{Minimize}: && x(1) + \int_0^1 (|x - u| - v - w)\, dt \\
& \mathit{subject\ to}: && \dot{x} = u - v - w \ \ a.e., \ x(0) = 0 \\
& && u \geq 0 \ \ a.e. \\
& && v \in \mathcal{V}[u], \ w \in \mathcal{W}[u, v],
\end{aligned}$$

where $\mathcal{V}[u]$ is the set of optimal controls v of the following problem,

$$\begin{aligned}
P_2[u]: \quad & \mathit{Minimize}: && y(1) + \int_0^1 y dt \\
& \mathit{subject\ to}: && \dot{y} = u + v \ \ a.e., \ y(0) = 0 \\
& && v \geq 0 \ \ a.e.,
\end{aligned}$$

$\mathcal{W}[u, v]$ is the set of optimal controls w of the following problem,

$$\begin{aligned}
P_2[u, v]: \ & \mathit{Minimize}: && z(1) + \int_0^1 z dt \\
& \mathit{subject\ to}: && \dot{z} = u + v + w \ \ a.e., \ z(0) = 0 \\
& && w \geq 0 \ \ a.e..
\end{aligned}$$

Here, $(x, y, z) \in AC([0,1], R^3)$ is the state, $(u, v, w) \in L^1([0,1], R^3)$ is the control.

For this problem, it is easy to see that

$$\begin{aligned} V_1(u) := &\quad inf\left\{y(1) + \int_0^1 ydt : \dot{y} = u + v,\ y(0) = 0,\ v \geq 0\right\} \\ = &\quad \int_0^1 \left(u + \int_0^t ud\tau\right) dt \\ V_2(u,v) := &\quad inf\left\{z(1) + \int_0^1 zdt : \dot{z} = u + v + w,\ z(0) = 0,\ w \geq 0\right\} \\ = &\quad \int_0^1 \left(u + v + \int_0^t (u+v)d\tau\right) dt \end{aligned}$$

Then, P_1 is equivalent to the following optimal control problem,

$$\begin{aligned} P_1: \qquad Minimize: &\quad x(1) + \int_0^1 (|x - u| - v - w)\, dt \\ subject\ to: &\quad \dot{x} = u - v - w \ \ a.e.,\ x(0) = 0 \\ &\quad \dot{y} = u + v \ \ a.e.,\ y(0) = 0 \\ &\quad \dot{z} = u + v + w \ \ a.e.,\ z(0) = 0 \\ &\quad u \geq 0,\ v \geq 0,\ w \geq 0 \ \ a.e. \\ &\quad y(1) + \int_0^1 \left(y - u - \int_0^t ud\tau\right) dt \leq 0 \\ &\quad z(1) + \int_0^1 \left(z - u - v - \int_0^t (u+v)d\tau\right) dt \leq 0. \end{aligned}$$

If $(x_*, y_*, z_*; u_*, v_*, w_*)$ is a local minimizer for P_1, then $v_* = w_* = 0$ and by Theorem 2.1 there are $p_1,\ p_2,\ p_3 \in AC$, $\zeta_1,\ \zeta_2,\ \zeta_3 \in L^\infty$ and $\lambda_1,\ \lambda_2,\ \lambda_3 \geq 0$ with $\sum_{i=1}^3 \lambda_i = 1$ such that (2.1)-(2.3) hold.

Here, (2.1) is equivalent to that

$$\begin{aligned} (-\dot{p}_1; \zeta_1) = (0, p_1 + p_2 + p_3) - \lambda_1\alpha(1,-1) - \lambda_2(0, \partial_u(-u - \int_0^t ud\tau)) \\ -\lambda_3(0, \partial_v(-v - \int_0^t vd\tau)) \ \ a.e., \\ for\ some\ \alpha \in [-1,1], \\ -\dot{p}_2(t) = -\lambda_2,\ \zeta_2 = -p_1 + p_2 + p_3 + \lambda_1 + \lambda_3\partial_v(v + \int_0^t vd\tau) \\ -\dot{p}_3(t) = -\lambda_2,\ \zeta_3 = -p_1 + p_3 + \lambda_1. \end{aligned} \tag{5.21}$$

(2.3) implies that

$$-p_1(1) = \lambda_1,\ -p_2(1) = \lambda_2,\ -p_3(1)) = \lambda_3. \tag{5.22}$$

Combining (5.21) and (5.22) we see that

$$p_1 = \alpha\lambda_1 t - \alpha\lambda_1 - \lambda_1,\ \ p_2 = \lambda_2 t - 2\lambda_2,\ \ p_3 = \lambda_3 t - 2\lambda_3, \tag{5.23}$$

and for any $u_0 \in L^1$

$$\langle \zeta_1, u_0 \rangle = (p_1 + p_2 + p_3)u_0 + (\alpha\lambda_1 + \lambda_2 + \lambda_3)u_0 + (\lambda_2 + \lambda_3)\int_0^t u_0 d\tau \ \ a.e.. \tag{5.24}$$

If $u_* \neq 0$, i.e. there exists $I \in [0,1]$ whose measure is not zero such that $u_*(t) > 0$ for any $t \in I$, then from (2.2) it follows that $\zeta_1(t) = 0$ for a.e. $t \in I$. Put $u_0 \equiv 1$, (5.24) implies that

$$\alpha\lambda_1 t - \lambda_1 + 2\lambda_2 t - \lambda_2 + 2\lambda_3 t - \lambda_3 = 0 \quad for \;\; a.e. \;\; t \in I,$$

which contradicts that λ_1, λ_2, $\lambda_3 \geq 0$ and $\sum_{i=1}^{3} \lambda_i = 1$. Therefore $u_* = 0$.

Here, it is not difficult to check that there exist $\lambda_1 > 0$ with p_1, p_2, p_3, ζ_1, ζ_2, ζ_3 and λ_2, λ_3 stated above such that (2.1) and (2.4) hold. Thus, by Theorem 2.2 we know that $(x_*, y_*, z_*; u_*, v_*, w_*) = 0$ is an optimal solution for P_1.

Acknowledgements

The author is greatly indebted to Professor K. Tsujioka for many helpful advices and constant encouragement. He would like to express his deep gratitude to Professor S. Koike for helpful suggestions. He also would like to thank the referee very much for useful comments.

References

Clarke, F. H. (1983). *Optimization and Nonsmooth Analysis*, Wiley-Intersicence, New York.

Clarke, F. H. (1975). Generalized gradients and applications,*Trans. Amer. Math. Soc.* **205**,247-262.

Hestenes, M. R. (1980). *Calculus of Variations and Optimal Control Theory*, John Wiley,New York.

Loewen, P. D. (1993). *Optimal Control via Nonsmooth Analysis*, CRM Proceedings Lecture Notes 0.2, American Mathematical Society, Providence.

Pinho, M. R. and Vinter, R. B. (1995). An Euler-Lagrange inclusion for optimal Control Problems,*IEEE Trans. Automat. control*, **40**, No.7,July.

Ye, J. J. (1995).Necessary conditions for bilevel dynamic optimization problems,*SIAM J, Control Optim.*, **33**,1208-1233.

Ye, J. J. (1997).Optimal strategies for bilevel dynamic problems, *SIAM J, Control Optim.*, **35**,512-531.

Ye, J. J. and Zhu, Q. J. (1997).Perturbed differential inclusion problems with nonaddictive L^1−perturbations and applications,*J. Optim. Theory Appl.*, **92**, No. 1,189-208.

Yezza, A. (1996).First-order necessary optimality conditions for general bilevel programming problems,*J. Optim. Theory Appl.*,**89**, No. 1, 189-219.

Received June 23,1998
Revised October 29,1998
Re-revised January 12,1999
Re-re-revised March 1,1999