Journal of Computational Mathematics, Vol.10, No.4, 1992, 305–320.

MINIMAX METHODS FOR OPEN-LOOP EQUILIBRA IN N-PERSON DIFFERENTIAL GAMES PART II: DUALITY AND PENALTY THEORY*[1)]

Goong Chen
(*Department of Mathematics, Texas A & M University, USA*)
Zheng Quan
(*Shanghai University of Science and Technology, Shanghai, China*)

Abstract

The equilibrium strategy for N-person differential games can be obtained from a min-max problem subject to differential constraints. The differential constraints are treated here by the duality and penalty methods.

We first formulate the duality theory. This involves the introduction of $N+1$ Lagrange multipliers: one for each player and one commonly shared by all players. The primal min-max problem thus results in a dual problem, which is a max-min problem with no differential constraints.

We develop the penalty theory by penalizing $N+1$ differential constraints. We give a convergence proof which generalizes a theorem due to B.T. Polyak.

§1. Introduction

In part I[5], we have presented a new minimax approach to N-person nonzero-sum differential games. We have also seen several advantages of using this approach.

The constraint equation contains N strategy variables $u_1, \cdots, n_N$ and one state variable x. Although $2N$ auxiliary variables $v_1, \cdots, v_N$ and $x^1, \cdots, x^N$ have been added in N supplementary differential equation constraints, they play the same roles as $u_1, \cdots, u_N$ and x, respectively. The functional $F(u, v)$ depends on $u_1, \cdots, u_N, v_1, \cdots,$ and v_N; $N+1$ state variables $x, x^1, \cdots, x^N$ and $N+1$ differential constraints are eliminated by integration. Therefore, from the mathematical programming point of view, the approach taken in Part I can be classified as primal. Computationally, this involves a rather large number of quadrature evaluations[3].

It is fair for us to say that most works in the literature on minimax problems are primal in nature in the sense that their constraints are handled in an implicit way.

On the other hand, looking back at optimal control problems, we understand that the use of different mathematical programming approaches of duality and penalty (cf.

* Received December 31, 1988.
1) Supported in part by NSF grant MCS 81–01892 and NASA contract No. NAS1–15810.

[12], [19], [4]) can lead to significant insights for solutions of those problems. These approaches also have the added advantage of being very amenable to numerical computations. One may wonder what can be done for N-person differential games. Here we are interested in developing some duality and penalty theory for minimax problems as well as numerical methods for N-person differential games. Indeed, this is the main motivation of our work.

By duality or penalty, differential constraints are handled explicitly. In the duality method Lagrange multipliers are introduced which eliminate the state constraints. In the penalty method, the system dynamics equations are penalized, which again results in an unconstrained problem. Both methods invole fewer quadrature calculations, and the variational matrix equations are sparse. Thus the computation is less costly and more efficient.

In §2, we first establish the duality theory under a general setting. For N-person games, we need to introduce $N+1$ Lagrange multipliers: one for each player, and one commonly shared by all players. Under the convexity-concavity assumption, we use the Hahn-Banach separation theorem to prove that the primal inf-sup problem leads to a dual sup-inf problem.

In §3, we present the fundamental penalty theorem. $N+1$ auxiliary state equations are penalized, with $N+1$ penalty parameters. The rate of convergence with respect to the penalty parameters is determined. Our work here extends and generalizes an earlier result of B.T. Polyak [16].

The applications of duality and penalty theory to finite element and their numerical examples are given in Part III[6].

§2. Duality Theory

As in Part I, we assume the following linear dynamics:

$$\begin{aligned} &\dot{x}(t) - A(t)x(t) - \sum_{i=1}^{N} B_i(t)u_i(t) - f(t) = 0 \quad \text{on } [0,T], \\ &x(0) = x_0 \in R^n. \end{aligned} \tag{2.1}$$

For notational convenience later on, we denote the system differential equation as

$$(DE) \equiv \dot{x}(t) - A(t)x(t) - \sum_{i=1}^{N} B_i(t)u_i(t) - f(t).$$

The matrix and vector functions $A(t), f(t), B_i(t), u_i(t), i = 1, \cdots, N$, satisfy the same conditions as in Part I.

Each player wants to minimize his cost

$$J_i(x,u) = J_i(x, u_1, \cdots, u_N), \quad i = 1, 2, \cdots, N \tag{2.2}$$

which is continuous with respect to (x,u) in the $H_n^1 \times U$ norm. As before, we let

$$F(x,u;X,v) = F(x,u_1,\cdots,u_N;x^1,\cdots,x^N,v_1,\cdots,v_N)$$

$$= \sum_{i=1}^{N} [J_i(x,u) - J_i(x^i, v^i)], \tag{2.3}$$

where $X = (x^1, \cdots, x^N), v^i = (u_1, \cdots, u_{i-1}, v_i, u_{i+1}, \cdots, n_N)$ and each x^i is the solution of

$$x^i(t) - A(t)x^i(t) - \sum_{j \neq i} B_j(t)u_j(t) - B_i(t)v_i(t) - f(t) = 0, \quad \text{on } [0,T],$$
$$x^i(0) = x_0, \quad i = 1, \cdots, N. \tag{2.4}$$

Also, we denote

$$(DE)_i \equiv \dot{x}^i(t) - A(t)x^i(t) - \sum_{j \neq i} B_j(t)u_j(t) - B_i(t)v_i(t) - f(t), \quad i = 1, \cdots, N.$$

We inherit the notations $U \equiv \prod_{i=1}^{N} U_i$ and $U_i = L^2_{m_i}(0,T)$ form Part I. From now on, we signify the Sobolev space

$$H_n^k \equiv H_n^k(0,T) \equiv \left\{ y : [0,T] \to R^n \mid \|y\|_{H_n^k} \equiv \sum_{j=0}^{k} \|(d/dt)^j y\|^2_{L^2_n} < \infty \right\}.$$

For future use, we also let $H^1_{0,n} \equiv \{y \in H^1_n \mid y(T) = 0\}$ and $H^1_{n,0} \equiv \{y \in H^1_n \mid y(0) = 0\}$. Following Part I, we consider the inf-sup problem

$$\text{(P)} \qquad \inf_{x,u} \sup_{X,v} \{F(x,u;X,v) | (x,u) \in H^1_n \times U \text{ subject to (2.1)} ,$$
$$(X,v) \in [H^1_n]^N \times U \text{ subject to (2.4) for } i = 1,2,\cdots,N\}$$

which constitutes the primal problem. Associated with (P) is the dual problem

$$\text{(D)} \qquad \sup_{p_0 \in L^2_n} \inf_{p \in [L^2_n]^N} L(p_0, p),$$

where $p = (p_1, \cdots, p_N)$ and

$$L(p_0, p) = L(p_0, p_1, \cdots, p_N) \equiv \inf_{x,u} \sup_{X,v} \mathbf{L}(p_0, p; x, u; X, v)$$

with the lagrangian $\mathbf{L} : L^2_n \times [L^2_n]^N \times H^1_n \times U \times [H^1_n]^N \times U$ defined by

$$\mathbf{L}(p_0, p; x, u; X, v) \equiv F(x,u;X,v) + \langle p_0, x - Ax - \sum_{j=1}^{N} B_j u_j - f \rangle_{L^2_n}$$
$$+ \sum_{i=1}^{N} \langle p_1, \dot{x}^i - Ax^i - \sum_{j=1, j\neq i}^{N} B_j u_j - B_i v_i - f \rangle_{L^2_n} \tag{2.5}$$

for x, X safisfying $x(0) = x_0, X(0) = X_0 = (x_0, \cdots, x_0)$.

From now on we say that (x,u) or (X,v) is feasible if $(x,u) \in H^1_n \times U$ satisfies (2.1) and $(X,v) \in [H^1_n]^N \times U$ satisfies (2.4) for $i = 1, \cdots, N$ for some given u. Similarly, (p_0, p) is feasible if $(p_0, p) \in L^2_n \times [L^2_n]^N$. For any given $(x,u) \in H^1_n \times U$, $x(0) = x_0$, we define

$$\psi(x,u) = \sup_{\substack{(x,v)\\ \text{feasible}}} F(x,u;X,v) = \sum_{i=1}^{N} J_i(x,u) - \sup_{\substack{(X,v)\\ \text{feasible}}} \sum_{i=1}^{N} J_i(x^i, u_1, \cdots, v_i, u_{i+1}, \cdots, u_N). \tag{2.6}$$

We are now in a position to state the fundamental theorem in this paper.

Theorem 2.1 (Duality Theorem). *Assume that*

(A0) $\inf_{\substack{(u,x)\\ \text{feasible}}} \sup_{\substack{(X,v)\\ \text{feasible}}} F(x,u;X,v) \equiv \hat{c} < \infty.$

(A1) For any fixed $(x,u) \in H_n^1 \times U, F(x,u;X,v)$ is concave in (X,v) for any $(X,v) \in [H_n^1]^N \times U, X(0) = X_0$.

(A2) $\psi(x,u)$ as defined in (2.6) is convex in $(x,u) \in H_n^1 \times U, x(0) = x_0$.

Then

$$\sup_{p_0 \in L_n^2} \inf_{p \in [L_n^2]^N} L(p_0, p) = \max_{p_0 \in L_n^2} \inf_{p \in [L_n^2]^N} L(p_0,p) = \inf_{\substack{(x,u)\\ \text{feasible}}} \sup_{\substack{(X,v)\\ \text{feasible}}} F(x,u;X,v). \tag{2.7}$$

Consequently, if the differential game has an equilibrium strategy, then

$$\sup_{p_0 \in L_n^2} \inf_{p \in [L_n^2]^N} L(p_0,p) = \min_{\substack{(x,u)\\ \text{feasible}}} \max_{\substack{(X,v)\\ \text{feasible}}} F(x,u;X,v) = \hat{c} = 0. \tag{2.8}$$

We now prove the theorem. For any given $(x,u) \in H_n^1 \times U$, let

$$\phi(x,u,p) = \sup_{X,v} \Big\{ F(x,u;X,v) + \sum_{i=1}^{N} \langle p_i, (DE)_i \rangle | X \in [H_n^1]^N, \\ v \in U, X(0) = X_0, \quad p = (p_1, \cdots, p_N) \in [L_n^2]^N \Big\}. \tag{2.9}$$

By (A0), we know that there exists, at least, one feasible (x,u) such that

$$\sup_{\substack{(X,v)\\ \text{feasible}}} F(x,u;X,v) = \psi(x,u) < +\infty. \tag{2.10}$$

From now on we need only to study $\psi(x,u)$ and $\phi(x,u,p)$ for those (x,u) satisfying (2.10).

Lemma 2.2 (Weak Duality). *For any (x,u) satisfying* (2.10) *the functional $\phi(x,u,p)$ defined above is convex in p and the following statement holds:*

$$\inf_{p \in [L_n^2]^N} \phi(x,u,p) \geq \psi(x,u). \tag{2.11}$$

Proof. By simple verification.

Lemma 2.3 (Strong Duality). *Assume that $F(x,u;X,v)$ is concave in (X,v) for all $(X,v) \in [H_n^1]^N \times U$, $X(0) = X_0$. Then for any $(x,u) \in H_n^1 \times U, x(0) = x_0$, we have*

$$\inf_{p \in [L_n^2]^N} \phi(x,u,p) = \psi(x,u). \tag{2.12}$$

In fact, the above infimum is attained and we actually have

$$\min_{p\in[L_n^2]^N} \phi(x,u,p) = \psi(x,u). \tag{2.12$'$}$$

Proof. If $\psi(x,u) = +\infty$, then (2.12) holds trivially by Lemma 2.2. So we assume that (2.10) holds. The standard arguments such as in [12] immediately apply. We define two convex sets

$$\begin{aligned} Y =&\{(a,0) \in R \times [L_n^2]^N | \psi(x,u) \le a\}, \\ Z =&\{(a,b) \in [L_n^2]^N | F(x,u;X,v) \ge a, b = (b_1,\cdots,b_N), b_i = x^i - Ax^j - B_i v_i \\ &- \sum_{j\neq i} B_j u_j - f, x^i(0) = x_0, i = 1,\cdots,N\}. \end{aligned}$$

Then it is easily checked that Y and Z are both convex, closed and $Y \cap \operatorname{int} Z = \Phi$ since when $b = 0 \in [L_n^2]^N$,

$$a < F(x,u;X,v) \le \sup_{\substack{(X,v)\\ \text{feasible}}} F(x,u;X,v)$$

for any $(a,0) \in \operatorname{int} Z$, which is obviously nonempty. So by the separation theorem[17,18], Y and Z can be weakly separated in $R \times [L_n^2]^N$:

$$ra_1 + \sum_{i=1}^{N} \langle q_i, b_i\rangle_{[L_n^2]^N} \le r.a_2, \quad \forall (a_1,b) \in Z, \quad (a_2,0) \in Y \tag{2.13}$$

for some nontrivial $(r,q) \in R \times [L_n^2]^N$. We now argue that $r > 0$. For, if r were equal to 0, then $q \neq 0$, so we can choose $b = q$ and conclude

$$0 < \sum_{i=1}^{N} \|q_i\|^2 \le 0,$$

a contradiction. Also, if r were less than 0, we can use $b = 0, a_1 = F(x,u;X,v) - \varepsilon (\varepsilon > 0)$ and $a_2 = \psi(x,u)$ in (2.13) to get

$$r \cdot (F(x,u;X,v) - \varepsilon) + 0 \le r \cdot \psi(x,u),$$

so

$$F(x,u;X,v) - \varepsilon \ge \psi(x,u) = \sup_{\substack{(X,v)\\ \text{feasible}}} F(x,u;X,v),$$

again a contradiction.

Therefore $r > 0$ so r can be normalized to 1. Using $a_1 = F(x,u;X,v)$ and $a_2 = \psi(x,u)$ in (2.13), we get

$$F(x,u;X,v) + \sum_{i=1}^{N} \langle q_i, b_i\rangle \le \psi(x,u)$$

i.e.,

$$F(x,u;X,v) + \sum_{i=1}^{N} \langle q_i, \dot{x}^i - Ax^i - \sum_{j\neq i} B_j u_j - B_i v_i - f\rangle \le \psi(x,u). \tag{2.14}$$

Therefore, $\phi(x,u,q)\le\psi(x,u)$; thus

$$\inf_{p\in[L_n^2]^N}\phi(x,u,p)\le\phi(x,u,q)\le\psi(x,u). \tag{2.15}$$

Combining the above with (2.11), we conclude

$$\inf_{p\in[L_n^2]^N}\phi(x,u,p)=\psi(x,u).$$

It is well understood in duality theory that the 'hyperplane' separating Y and Z will define and attain the optimal dual multiplier[18]. This can be easily seen here from (2.11) and (2.15), because q satisfies

$$\phi(x,u,q)=\psi(x,u)=\inf_{p\in[L_n^2]^N}\phi(x,u,p),$$

so

$$\phi(x,u,p)=\min_{p\in[L_n^2]^N}\phi(x,u,p)=\psi(x,u).$$

Therefore (2.12)$'$ is proved.

The arguments for the following lemma are the same as those for Lemmas 2.2 and 2.3; the proofs are therefore omitted.

Lemma 2.4. *Assume the* $\psi(x,u)$ *as defined in* (2.6) *is convex in* (x,u) *for* $(x,u)\in H_n^1\times U$, $x(0)=x_0$. *Then*

$$\sup_{p\in L_n^2}\inf_{\substack{(x,u)\in H_n^1\times U\\ x(0)=x_0}}[\psi(x,u)+\langle p_0,(DE)\rangle]=\max_{p_0\in L_n^2}\inf_{\substack{(u,x)\in H_n^1\times U\\ x(0)=x_0}}[\psi(x,u)+\langle p_0,(DE)\rangle]$$
$$=\inf_{\substack{(x,u)\\ \text{feasible}}}\psi(x,u). \tag{2.16}$$

Remark 2.5. In (2.9), we have introduced N Lagrange multipliers p_i, one for each player. In (2.16) we have introduced the joint multiplier p_0 commonly shared by all players.

Proof of Theorem 2.1. From Lemmas 2.3 and 2.4 we conclude that

$$(\mathrm{P})=\inf_{x,u}\sup_{X,v}\{F(x,u;X,v)|(x,u)\text{ and }(X,v)\text{ are feasible}\}$$
$$=\inf_{\substack{(x,u)\\ \text{feasible}}}\Big[\sup_{\substack{(X,v)\\ \text{feasible}}}f(x,u;X,v)\Big]=\inf_{\substack{(x,u)\\ \text{feasible}}}\psi(x,u)$$
$$=\max_{p_0\in L_n^2}\inf_{\substack{(x,u)\in H_n^1\times U\\ x(0)=x_0}}[\psi(x,u)+\langle p_0,(DE)\rangle]\ \text{(by Lemma 2.4)}$$
$$=\max_{p_0\in L_n^2}\inf_{\substack{(x,u)\in H_n^1\times U\\ x(0)=x_0}}\Big\{\inf_{p\in[L_n^2]^N}\Big[\sup_{\substack{(X,v)\in[H_n^1]^N\times U\\ X(0)=X_0}}\Big[F(x,u;X,v)+\sum_{i=1}^N\langle p_i,(DE)_i\rangle\Big]\Big]$$
$$+\langle p_0,(DE)\rangle\Big\}=\max_{p_0\in L_n^2}\inf_{\substack{(x,u)\in H_n^1\times U\\ x(0)=x_0}}\Big\{\inf_{p\in[L_n^2]^N}\Big[\sup_{\substack{(X,v)\in[H_n^1]^N\times U\\ X(0)=X_0}}\Big[F(x,u;X,v)$$
$$+\sum_{i=1}^N\langle p_i,(DE)_i\rangle+\langle p_0,(DE)\rangle\Big]\Big\}$$

$$= \max_{p_0\in L_n^2} \inf_{\substack{(x,u)\in H_n^1\times U\\ x(0)=x_0}} \inf_{p\in[L_n^2]^N} \sup_{\substack{(X,v)\in[H_n^1]^N\times U\\ X(0)=X_0}} \Big[F(x,u;X,v)+\sum_{i=1}^N\langle p_i,(DE)_i\rangle$$

$$+\langle p_0,(DE)\rangle\Big] = \max_{p_0\in L_n^2} \inf_{p\in[L_n^2]^N} L(p_0,p). \qquad (2.17)$$

So (2.7) is proved.

Theorem 2.1 will be sufficient for the purpose of subsequent development in this paper and Part III[6]. The theory of duality can be further strengthened and improved, but the work is too general and lengthy, so we choose not to do this here.

§3. The Penalty Method for N-Person Differential Games: Rate of Convergence

For the ease of presentation, throughout this section, we assume that we have made the change of variable $x(t)\to x(t)-x_0$ in (2.1) so that $x(0)=0$. Thus the space for x will be $H_{n,0}^1$. This change of variable results in only superficial changes of J_i. We will use L^2 to denotes L_n^2.

For optimization problems with equality constraints, the penalty method approximates the original problems by penalizing the equality constraints. In the minimax problem (P), we see that $(DE)=0$ and $(DE)_i=0$ $(i=1,\cdots,N)$ are $N+1$ equality constraints. Thus, a completely natural use of penalty here is to approximate the problem by

$$\inf_{(x,u)\in H_{n,0}^1\times U} \sup_{(X,v)\in[H_{n,0}^1]^N\times U} F_e(x,u;X,v)\equiv F(x,u;X,v)+\frac{1}{e_0}\|(DE)\|_{L^2}^2$$

$$+\sum_{i=1}^N\frac{1}{e_i}\|(DE)_i\|_{L^2}^2 \qquad (3.1)$$

for some $e_0,e_1,\cdots,e_N>0$.

The most important question remains in determining the validity of the above scheme and, if it is valid, its rate of convergence. Thus, we consider the fundamental theorem of penalty for N-person differential games below. The following assumption will be needed:

(B0) $F(x,u;X,v)$ is strictly convex in (x,u) and strictly concave in (X,v) for $(x,u)\in H_{n,0}^1\times U$ and $(X,v)\in[H_{n,0}^1]^N\times U$; and $F(u,v)$ is strictly convex in u as well as strictly concave in v.

(B1) $\inf\limits_{\substack{(x,u)\in H_{n,0}^1\times U\\ (DE)=0}} \sup\limits_{\substack{(X,v)\in[H_{n,0}^1]^N\times U\\ (DE)_i=0,\, i=1,\cdots,N}} F(x,u;X,v)$ is attained by $(\hat x,\hat u;\hat X,\hat v)\in H\equiv H_{n,0}^1\times U\times[H_{n,0}^1]^N\times U$. By (B0), this point $(\hat x,\hat u;\hat X,\hat v)$ is unique.

(B2) There exist Lagrange multipliers $\hat p_0,\hat p=(\hat p_1,\cdots,\hat p_N)$ which are optimal (max-min) dual multipliers such that the following holds:

$$F(\hat{x},\hat{u};\hat{X},\hat{v}) = \min_{\substack{(x,u)\in H^1_{n,0}\times U\\ (DE)=0}} \max_{\substack{(X,v)\in[H^1_{n,0}]\times U\\ (DE)_i=0,i=1,\cdots,N}} F(x,u;X,v)$$
$$= \max_{p_0\in L^2} \min_{p\in[L^2]^N} \min_{(x,u)\in H^1_{n,0}\times U} \max_{(X,v)\in[H^1_{n,0}]^N\times U} [F(x,u;X,v)$$
$$+\langle p_0,(DE)\rangle + \sum_{i=1}^{N}\langle p_i,(DE)_i\rangle]. \tag{3.2}$$

(B3) The costs $J_i(x,u)$ are of the forms

$$J_i(x,u) = \int_0^T h_i(x(t),u(t))dt \tag{3.3}$$

so that $\hat{p}_0 \in H^1_{n,0}, \hat{p} \in [H^1_{n,0}]^N$.

(B4) The first and second derivatives F', F'' exist, and F'' satisfies the global Lipschitz condition

$$\|F''(x_1,u_1;X_1,v_1) - F''(x_2,u_2;X_2,v_2))\| \le K_1\|(x_1-x_2,u_1-u_2;$$
$$X_1-X_2,v_1-v_2)\|_{\bar{H}}, \text{ for some } K_1>0 \text{ uniformly, for } (x_1,u_1;X_1,v_1),$$
$$(X_2,u_2;X_2,v_2), \text{ where } \bar{H}\equiv L^2\times U\times[L^2]^N\times U. \tag{3.4}$$

(B5) Let $\mathbf{A}_0 \equiv \partial x^2 F, \mathbf{A}_1 \equiv \partial x^2 F$, $\mathbf{M}_0 \equiv \partial u^2 F$ and $\mathbf{M}_1 \equiv \partial v^2 F$ be second-order Frechet partial derivatives evaluated at $(\hat{x},\hat{u};\hat{X},\hat{v})$. Then $\mathbf{A}_0,\mathbf{M}_0,-\mathbf{A}_1$ and $-\mathbf{M}_1$ are positive definite linear operators on $L^2, U, [L^2]^N$ and U, respectively. Furthermore, $\mathbf{A}_0\times\mathbf{A}_1$ maps $H^1_{n,0}\times[H^1_{n,0}]^N$ into itself.

(B6) $B_1,\cdots,B_N$ are small relative to $\mathbf{A}_0,\mathbf{M}_0,-\mathbf{A}_1$ and $-\mathbf{M}_1$ (cf (3.22), e.g.).

(B7) The mixed Frechet partial derivative operators $\partial_x\partial_x^F, \partial_x\partial_u,\cdots$, etc. evaluated at (x,u;X,v) are all 0.

Remark 3.1. (i) In (B3), J_i's are assumed to be of the form (3.3), for the convenience of discussions.

(ii) Making some other assumptions, we can relax the global Lipschitz condition (3.4) to a local one.

(iii) (B7) is assumed here only for the convenience of discussions, cf. Remark 3.5 later.

Theorem 3.2. *Under conditions* (B0)–(B7), *for* $e_0,e_1,\cdots,e_N>0$ *sufficiently small, there exists a unique* $(\hat{x}_e,\hat{u}_e;\hat{X}_e,\hat{v}_e)\in\bar{H}$ *satisfying* $F'_e=0$ *such that*

(i) $\|(\hat{x}_e,\hat{u}_e;\hat{X}_e,\hat{v}_e)-(\hat{x},\hat{u};\hat{X},\hat{v})\|_H \le K_2(\max_{j=0,\cdots,N} e_j)\|(\hat{p}_0,\hat{p})\|_{L^2\times[L^2]^N}$;

(ii) $\|2/e_0(\dot{\hat{x}}_e - A\hat{x}_e - \sum_i B_i\hat{u}_{ei} - f) - \hat{p}_0\|_{L^2} + \sum_i \|2/e_i(\dot{\hat{x}}^i - A\hat{x}^i - \sum_{j\neq i} B_j\hat{u}_{ej} - B_i\hat{v}_{ei} - f) - (-\hat{p}_i)\|_{L^2} \le K_3(\max_{j=0,\cdots,N} e_j)\|(\hat{p}_0,\hat{p})\|_{L^2\times[L^2]^N}$, *for some* $K_2,K_3>0$ *independent of* $e_0,\cdots,e_N$.

Proof. We introduce the new variables

$$\xi_0 = x-\hat{x},\quad \xi_1 = \hat{X}-X,\quad \eta_0 = u-\hat{u},$$

$$\eta_1 = v - \hat{v} - \zeta_0 = 2/e_0(\dot{x}^i - Ax^i - \sum_i B_i u_i - f) - \hat{p}_0,$$

$$\zeta_1^i = -2/e_i(\dot{x}^i - Ax^i - \sum_{j\neq i} B_j u_j - B_i v_i - f) - \hat{p}_i,$$

$$\zeta_1 = (\zeta_1^1, \zeta_1^2, \cdots, \zeta_1^N). \tag{3.5}$$

In the above, we first choose $x \in H_n^2 \cap H_{n,0}^1, X \in [H_n^2 \cap H_{n,0}^1]^N, u, v \in U \cap \prod_{i=1}^N H_{m_i}^1$ and then let $(x, u; X, v)$ tend to an element in H. We further let

$$\xi = (\xi_0, \xi_1), \quad \eta = (\eta_0, \eta_1), \quad \zeta = (\zeta_0, \zeta_1).$$

For any $(\delta x, \delta u; \delta X, \delta v) \in H$, we have

$$F_e'(x, u; X, v) \cdot (\delta x, \delta u; \delta X, \delta v) = F'(x, u; X, v) \cdot (\delta x, \delta u; \delta X, \delta v)$$
$$+ 2/e_0\Big\langle \dot{x} - Ax - \sum_i B_i u_i - f, \delta\dot{x} - A(\delta x) - \sum_i B_i(\delta u_i)\Big\rangle$$
$$+ \sum_i 2/e_i\Big\langle x^i - Ax^i - \sum_{j\neq i} B_j u_j - B_i v_i - f, \delta\dot{x}^i - A(\delta x^i)$$
$$- \sum_{j\neq i} B_j \delta u_j - B_i \delta v_i\Big\rangle. \tag{3.6}$$

We can use (B4) to write

$$F'(x, u; X, v) = F'(\hat{x}, \hat{u}; \hat{X}, \hat{v}) + F''(\hat{x}, \hat{u}; \hat{X}, \hat{v})(\xi_0, \eta_0; \xi_1, \eta_1) + r(\xi, \eta), \tag{3.7}$$

where the remainder $r(\xi, \eta)$ (as a functional in H) satisfies

$$r(0, 0) = 0, \tag{3.8}$$

$$\|r'(\bar{\xi}, \bar{\eta}) - r'(\tilde{\xi}, \tilde{\eta})\| \leq \|(\bar{\xi} - \tilde{\xi}, \bar{\eta} - \tilde{\eta})\|_{(L^2 \times [L^2]^N) \times (U \times U)},$$
$$\forall (\bar{\xi}, \bar{\eta}), (\tilde{\xi}, \tilde{\eta}) \in (H_{n,0} \times [H_{n,0}^1]^N) \times (U \times U). \tag{3.9}$$

Substituting (3.7) into the first term on the RHS of (3.6) and integrating the remaining terms by parts, we get

$$\text{LHS of (3.5)} = [F'(\hat{x}, \hat{u}; \hat{X}, \hat{v}) + F''(\hat{x}, \hat{u}; \hat{X}, \hat{v})(\xi_0, \eta_0, \xi_1, \eta_1)$$
$$+ r(\xi, \eta)] \cdot (\delta x, \delta u; \delta X, \delta v) - \Big\langle [d/dt + A^*] 2/e_0(\dot{x} - Ax - \sum_i B_i u_i - f), \delta x\Big\rangle$$
$$- 2/e_0\Big\langle \dot{x}(T) - A(T)x(T) - \sum_j (B_i u_i)(T) - f(T), \delta x\Big\rangle$$
$$- \sum_i \Big\langle B_i^* \cdot 2/e_0(\dot{x} - Ax - \sum_i B_i u_i - f), \delta u_i\Big\rangle$$
$$+ \sum_i \Big\langle [d/dt + A^*] \cdot 2/e_i(\dot{x}^i - Ax^i - \sum_{j\neq i} B_j u_j - B_i v_i - f), \delta x^i\Big\rangle$$
$$- \sum_i 2/e_i\Big\langle \dot{x}(T) - A(T)x^i(T) - \sum_{j\neq i} (B_j u_j)(T) - (B_i v_i)(T) - f(T), \delta x^j(T)\Big\rangle$$

$$+\sum_i\sum_{k\neq i}\left\langle B_k^*\cdot 2/e_i\cdot\left(\dot{x}^i-Ax^i-\sum_{j\neq i}B_ju_j-B_iv_i-f\right),\delta u_k\right\rangle$$

$$+\sum_i\left\langle B_i^*\cdot 2/e_i\left(\dot{x}^iAx^i-\sum_{j\neq i}B_ju_j-B_iv_i-f\right),\delta v_k\right\rangle.$$

Let $\mathbf{L}=d/dt-A$ and $\mathbf{L}^*=d/dt+A^*$ (the formal adjoint of $-\mathbf{L}$). We now substitute (3.5) into the above and denote that

$$F'(\hat{x},\hat{u};\hat{X},\hat{v}).(\delta x,\delta u;\delta X,\delta v)-\langle\mathbf{L}^*\hat{p}_0,\delta x\rangle-\sum_i\langle B_i^*\hat{p},\delta u_i\rangle-\sum_i\langle\mathbf{L}^*\hat{p}_i,\delta x^i\rangle$$
$$-\sum_i\sum_{j\neq i}\langle B_j^*\hat{p}_i,\delta u_i\rangle-\sum_i\langle B_i^*\hat{p}_i,\delta v_i\rangle=0. \qquad (3.11)$$

We get that the solution of $F'(x,u;X,v)=0$ can be found by solving

$$[F''(\hat{x},\hat{u};\hat{X},\hat{v})(\xi_0,\eta_0;\xi_1,\eta_1)+r(\xi,\eta)]\cdot(\delta x,\delta u;\delta X,\delta v)-\langle\mathbf{L}^*\zeta_0,\delta x\rangle$$
$$-\sum_i\langle B_i^*\zeta_0,\delta u_i\rangle-\sum_i\langle\mathbf{L}^*\zeta_1^1,\delta x^i\rangle-\sum_i\sum_{j\neq i}\langle B_j^*\zeta_1^1,\delta u_j\rangle$$
$$-\sum_j\langle B_i^*\zeta_1^1,\delta v_i\rangle=0. \qquad (3.12)$$

Note that all the $\langle\ ,\ \rangle$ terms on the RHS of (3.10) disappear because to the arbitrariness of $\delta x(T)$ and $\delta x^i(T)$. By (B5) and (B7) we have

$$F''(\hat{x},\hat{u};\hat{X},\hat{v})(\xi_0,\xi_1,\eta_0,\eta_1)=\begin{bmatrix}\mathbf{A}_0&0&0&0\\0&\mathbf{A}_1&0&0\\0&0&\mathbf{M}_0&0\\0&0&0&\mathbf{M}_1\end{bmatrix}\begin{bmatrix}\xi_0\\\xi_1\\\eta_0\\\eta_1\end{bmatrix}. \qquad (3.13)$$

Therefore, from (3.12) we get, for $i=1=1,\cdots,N$,

$$\mathbf{A}_0\xi_0-L^*\zeta_0=-r_1(\xi,\eta),\quad(\mathbf{A}_1,\xi_1)^i-\mathbf{L}^*\zeta_1^i=-r(\xi,\eta),$$
$$(\mathbf{M}_0,\eta_0)^i-B_i^*\zeta_0-\sum_{j\neq i}B_j^*\zeta_1^1=-r^i(\xi,\eta),\quad(\mathbf{M}_1,\eta_1)^i-B_i^*\zeta_1^i=-r_4(\xi,\eta), \qquad (4.14)$$

where r_1,r_2,r_3, and r_4 are the respective components of $r(\xi,\eta)$ and the superscript i denotes the i-th component. Combining (3.14) with (3.5), we get the following nonlinear "matrix" equation:

$$\begin{bmatrix}\mathbf{A}_0&0&0&0&-\mathbf{L}^*&0\\0&\mathbf{A}_1&0&0&0&\begin{bmatrix}-\mathbf{L}^*&&0\\&\ddots&\\0&&-\mathbf{L}^*\end{bmatrix}\\0&0&\mathbf{M}_1&0&\mathbf{B}_1^*&\mathbf{B}_2^*\\0&0&0&\mathbf{M}_1&0&\mathbf{B}_3^*\\-\mathbf{L}&0&\mathbf{B}_1&0&e_0/2I&0\\0&\begin{bmatrix}-\mathbf{L}&&0\\&\ddots&\\0&&-\mathbf{L}\end{bmatrix}&B_2&B_3&0&\begin{bmatrix}e_{1/2}I&&\\&\ddots&\\&&e_{N/2}I\end{bmatrix}\end{bmatrix}\begin{bmatrix}\xi_0\\\xi_1\\\eta_0\\\eta_1\\\zeta_0\\\zeta_1\end{bmatrix}.$$

$$= \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -e_0/2p_0 \\ \begin{bmatrix} e_{1/2}p_1 & & \\ & \ddots & \\ & & e_{N/2}p_N \end{bmatrix} \end{bmatrix} + \begin{bmatrix} -r_1(\xi,\eta) \\ -r_2(\xi,\eta) \\ -r_3(\xi,\eta) \\ -r_4(\xi,\eta) \\ 0 \\ 0 \end{bmatrix} \tag{3.15}$$

where

$$\mathbf{B}_1 = \begin{bmatrix} -B_1 \\ -B_2 \\ \vdots \\ -B_N \end{bmatrix}, \quad \mathbf{B}_2 = \begin{bmatrix} 0 & -B_2 & \cdots & -B_{N-1} & -B_N \\ -B_1 & 0 & \cdots & -B_{N-1} & -B_N \\ \vdots & \vdots & & \vdots & \vdots \\ -B_1 & -B_2 & \cdots & -B_{N-1} & 0 \end{bmatrix},$$

$$\mathbf{B}_3 = \begin{bmatrix} -B_1 & & & 0 \\ & -B_2 & & \\ & & \ddots & \\ 0 & & & -B_N \end{bmatrix}.$$

We further abbreviate (3.15) as

$$D_e \begin{bmatrix} \xi \\ \eta \\ \zeta \end{bmatrix} = \begin{bmatrix} \Xi & 0 & -\tilde{\mathbf{L}}^* \\ 0 & M & B^* \\ -\tilde{\mathbf{L}} & B & I_e \end{bmatrix} \begin{bmatrix} \xi \\ \eta \\ \zeta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ p_e \end{bmatrix} + \begin{bmatrix} -\tilde{r}_1(\xi,\eta) \\ -\tilde{r}_2(\xi,\eta) \\ 0 \end{bmatrix}, \tag{3.16}$$

where

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_0 & 0 \\ 0 & \mathbf{A}_1 \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} \mathbf{M}_0 & 0 \\ 0 & \mathbf{M}_1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{B}_1 & 0 \\ \mathbf{B}_2 & \mathbf{B}_3 \end{bmatrix}, \quad \tilde{\mathbf{L}} = \begin{bmatrix} \mathbf{L} & & 0 \\ & \ddots & \\ 0 & & \mathbf{L} \end{bmatrix}_{(N+1)\times(N+1)}$$

$$\tilde{\mathbf{L}}^* = \begin{bmatrix} \mathbf{L}^* & & 0 \\ & \ddots & \\ 0 & & \mathbf{L}^* \end{bmatrix}_{(N+1)\times(N+1)}, \quad I_e = \begin{bmatrix} e_{0/2}I & & & 0 \\ & -e_{1/2}I & & \\ & & \ddots & \\ 0 & & & -e_{N/2}I \end{bmatrix}_{(N+1)\times(N+1)},$$

$$P_e = \begin{bmatrix} -e_{0/2}I \\ e_{1/2}I \\ \cdots \\ e_{N/2}I \end{bmatrix}, \quad \tilde{r}_1(\xi,\eta) = \begin{bmatrix} r_1(\xi,\eta) \\ r_2(\xi,\eta) \end{bmatrix}, \quad \tilde{r}_2(\xi,\eta) = \begin{bmatrix} r_3(\xi,\eta) \\ r_4(\xi,\eta) \end{bmatrix}.$$

By (B5), D_e is a closed linear operator on $(L^2 \times [L^2]^N) \times (U \times U) \times (L^2) \times [L^2]^N)$ with domain dom $(D_e) = (H^1_{n,0} \times [H_{n,0}]^N) \times (U \times U) \times (H^1_{n,0} \times [H^1_{n,0}]^N)$.

Lemma 3.3. *Under conditions* (B5), (B6) *and* (B7) *for all* $e_0, e_1, \cdots, e_N > 0$ *sufficiently small, the operator* D_e *introduced above has an inverse and*

$$\|D_e^{-1}\| \leq K_4. \tag{3.17}$$

Proof. For an arbitrarily given $(\alpha,\beta,\gamma) \in (L^2 \times [L^2]^N) \times (U \times U) \times (L^2 \times [L^2]^N)$, we wish to find some $(\bar{\xi},\bar{\eta},\bar{\zeta}) \in (H^1_{0,n} \times [H^1_{0,n}]^N) \times (U \times U) \times (H^1_{0,n} \times [H^1_{0,n}]^N)$ such that

$$D_e \begin{bmatrix} \bar{\xi} \\ \bar{\eta} \\ \bar{\zeta} \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \tag{3.18}$$

or, in detail,

$$\mathring{A}\bar{\xi} - L^*\bar{\zeta} = \alpha, \quad \mathbf{M}\bar{\eta} + \mathbf{B}^*\bar{\zeta} = \beta, \quad -\mathbf{L}\bar{\xi} + \mathbf{B}\bar{\eta} + I_e\bar{\zeta} = \gamma. \tag{3.19}$$

Let $\Phi(t,s)$ be the fundamental $n \times n$ matrix solution satisfying

$$\partial/\partial t\Phi(t,s) = A(t)\Phi(t,s), \quad 0 \le s \le t \le T,$$

$$\Phi(s,s) = I_{n\times n}$$

It is easy to see that $\tilde{\mathbf{L}}$ is invertible with inverse

$$(\tilde{\mathbf{L}})^{-1}\lambda = \int_0^t \begin{bmatrix} \Phi(t,s) & & 0 \\ & \ddots & \\ 0 & & \Phi(t,s) \end{bmatrix}_{(N+1)\times(N+1)} \lambda(s)ds.$$

Thus, we have from (3.19)

$$\bar{\xi} = (\tilde{\mathbf{L}})^{-1}(\mathbf{B}\bar{\eta} + I_e\bar{\zeta} - \gamma). \tag{3.20}$$

Substituting (3.20) into (3.19), we get

$$\mathbf{A}\tilde{L}^{-1}\mathbf{B}\bar{\eta} + (A\tilde{\mathbf{L}}^{-1}I_e - \tilde{\mathbf{L}}^*)\bar{\zeta} = \alpha + \mathbf{A}\tilde{\mathbf{L}}^{-1}\gamma.$$

The integro-differential operator $\tilde{\mathbf{L}}^* - \mathbf{A}\tilde{L}^{-1}I_e$ is easily seen to be invertible for $e = (e_0, e_1, \cdots, e_N)$ sufficiently small; thus we have

$$\bar{\zeta} = (\tilde{\mathbf{L}}^* - \mathbf{A}\tilde{\mathbf{L}}^{-1}I_e)^{-1}[\mathbf{A}\tilde{\mathbf{L}}^{-1}\mathbf{B}\bar{\eta} - (\alpha + \mathbf{A}\tilde{\mathbf{L}}^{-1}\gamma)]. \tag{3.21}$$

Substituting (3.21) into (3.19), we obtain

$$[\mathbf{M} + \mathbf{B}^*(\tilde{\mathbf{L}}^* - \mathbf{A}\tilde{\mathbf{L}}^{-1}I_e)^{-1}\mathbf{A}\tilde{\mathbf{L}}^{-1}\mathbf{B}]\bar{\eta} = \beta + \mathbf{B}^*(\tilde{\mathbf{L}}^* - A\tilde{\mathbf{L}}^* - A\tilde{\mathbf{L}}^{-1}I_e)^{-1}(\alpha + \mathbf{A}\tilde{\mathbf{L}}^{-1}\gamma).$$

Now we invoke (B6): since $\mathbf{M}$ is invertible, if $\mathbf{B}$ is relatively smaller than $\mathbf{M}$ such that

$$\mathbf{M} + \mathbf{B}^*(\tilde{\mathbf{L}}^* - A\tilde{\mathbf{L}}^{-1})^{-1}A\tilde{\mathbf{L}}^{-1}\mathbf{B} \tag{3.22}$$

is invertible for e sufficiently small, we have

$$\bar{\eta} = Q_e^{-1}[\beta + \mathbf{B}^*\tilde{\mathbf{L}}_e^{-1}(\alpha + A\tilde{\mathbf{L}}^{-1}\gamma)], \tag{3.23}$$

where

$$Q_e = \mathbf{M} + \mathbf{B}^*(\tilde{\mathbf{L}}^* - A\tilde{\mathbf{L}}^{-1}I_e)^{-1}A\tilde{\mathbf{L}}^{-1}B, \quad \tilde{\mathbf{L}}_e = \tilde{\mathbf{L}}^* - A\tilde{\mathbf{L}}^{-1}I_e. \tag{3.24}$$

Using (3.23) in (3.20) and (3.21), we obtain

$$\bar{\xi} = \tilde{\mathbf{L}}^{-1}\{[\mathbf{B}Q_e^{-1}\mathbf{B}^*\tilde{\mathbf{L}}_e^{-1} + I_e\tilde{\mathbf{L}}_e^{-1}(\tilde{\mathbf{L}}^{-1}\mathbf{B}\mathbf{B}^*\tilde{\mathbf{L}}^{-1} - I)]\alpha + [\mathbf{B}I_e^{-1} + I_e\tilde{\mathbf{L}}_e^{-1}A\tilde{\mathbf{L}}^{-1}\mathbf{B}Q_e^{-1}]\beta$$
$$+ [\mathbf{B}Q_e^{-1}\mathbf{B}^*\tilde{\mathbf{L}}_*^{-1}\mathbf{A}\tilde{\mathbf{L}}^{-1} - I + I_e\tilde{\mathbf{L}}_e^{-1}(A\tilde{\mathbf{L}}^{-1}\mathbf{B}\mathbf{B}^*\tilde{\mathbf{L}}_e^{-1} - I)A\tilde{\mathbf{L}}^{-1}]\gamma\},$$

$$\bar{\zeta} = \tilde{\mathbf{L}}_e^{-1}\{(A\tilde{\mathbf{L}}^{-1}\mathbf{B}\mathbf{B}^*\tilde{\mathbf{L}}_e^{-1} - I)\alpha + A\tilde{\mathbf{L}}^{-1}\mathbf{B}Q_e^{-1}\beta + (\mathbf{A}\tilde{\mathbf{L}}^{-1}\mathbf{B}\mathbf{B}^*\tilde{\mathbf{L}}_e^{-1} - I)\mathbf{A}\tilde{\mathbf{L}}^{-1}]\gamma\}.$$

Therefore, D_e is invertible, with

$$D_e^{-1} = \begin{bmatrix} \tilde{\mathbf{L}}^{-1}(\mathbf{B}Q_e^{-1}\mathbf{B}^*\tilde{\mathbf{L}}^{-1} + I_e\tilde{\mathbf{L}}_e^{-1}(\mathbf{A}\tilde{\mathbf{L}}^{-1}\mathbf{B}\mathbf{B}^*\tilde{\mathbf{L}}_e - I)) & \tilde{\mathbf{L}}^{-1}(\mathbf{B}Q_e^{-1} + I_e\tilde{\mathbf{L}}_e^{-1} \cdot \mathbf{A}\tilde{\mathbf{L}}^{-1}\mathbf{B}Q_e^{-1}) & \tilde{\mathbf{L}}^{-1}(\mathbf{B}Q_e^{-1}\mathbf{B}^*\tilde{\mathbf{L}}_e^{-1}\mathbf{A}\tilde{\mathbf{L}} - I + I_e\tilde{\mathbf{L}}_e^{-1}(a\tilde{\mathbf{L}}^{-1}\mathbf{B}\mathbf{B}^*\tilde{\mathbf{L}}^{-1} - I)\mathbf{A}\tilde{\mathbf{L}}^{-1}) \\ Q_e^{-1}\mathbf{B}\tilde{\mathbf{L}}^{-1} & Q_e^{-1} & Q_e^{-1}\mathbf{B}^*\tilde{\mathbf{L}}_e^{-1}\mathbf{A}\tilde{\mathbf{L}}^{-1} \\ \tilde{\mathbf{L}}_e^{-1}(\mathbf{A}\tilde{\mathbf{L}}^{-1}\mathbf{B}\mathbf{B}^*\tilde{\mathbf{L}}_e^{-1} - I) & \mathbf{A}\tilde{\mathbf{L}}^{-1}\mathbf{B}Q_e^{-1} & (\mathbf{A}\tilde{\mathbf{L}}^{-1}\mathbf{B}\mathbf{B}^*\tilde{\mathbf{L}}_e^{-1} - I)\mathbf{A}\tilde{\mathbf{L}}^{-1} \end{bmatrix}$$

Since each entry of the matrix D_e^{-1} is bounded, we have proved that D_e^{-1} is bounded for e sufficiently small.

We need the following lemma from [16]:

Lemma 3.4. *Let* $\mathbf{H}$ *be a given Hilbert space,* T *a densety defined closed linear operator from* dom $(T) \subset \mathbf{H}$ *onto* $\mathbf{H}$ *with a bounded inverse* $\|T^{-1}\| \leq c_1$, *and* $r(x)$ *a nonlinear (Frechet) differentiable operator on* $\mathbf{H}$ *such that* $r(0) = 0, \|r'(x)\| \leq c_2\|x\|$ *for all* $x \in \mathbf{H}$. *Then for any* $a \in \mathbf{H}, \|a\| \leq 1/(4c_1^2c_2)$, *the equation*

$$Tx = a - r(x)$$

has, in the sphere $\|x\| \leq 4c_1\|a\|$, *a unique solution* $x \in$ dom (T) *satisfying*

$$\|x\| \leq \frac{c_1}{2}\|a\|.$$

We note that, although in [16] it is assumed that T is bounded, a careful examination of the proof shows that the assumption is redundant.

Using $T = D_e, c_1 = k_4$ and $c_2 = K_1$ in Lemma 3.4 and applying it to (3.16), we obtain that, for

$$\|\tilde{p}_e\|_{L^2\times[L^2]^N} \leq 1/(4K_1k_4^2),$$

which is clearly satisfied if

$$\min_{j=1,\cdots,N} 2/e_j \geq 4K_1K_4^2\|(-\hat{p}_0,\hat{p})\|_{L^2\times[L^2]^N},$$

(3.18) has a solution $(\hat{\xi}_e, \hat{\eta}_e, \hat{\zeta}_e) \in [H_{n,0}^1]^{N+1} \times U^2 \times [H_{n,0}^1]^{N+1}$ satisfying

$$\|(\hat{\xi}_e, \hat{\eta}_e, \hat{\zeta}_e)\|_{[L^2]^{N+1}\times U^2\times[L^2]^{N+1}} \leq K_4/4\left(\max_{j=0,\cdots,N} e_j\right)\|(-\hat{p}_0,\hat{p})\|_{[L^2]^{N+1}}.$$

From (3.5), writing

$$\hat{x}_e = \hat{x} + \hat{\xi}_{e,0}, \quad \hat{X}_e = \hat{X} + \hat{\xi}_{e,1}, \quad \hat{u}_e = \hat{u} + \hat{\eta}_{e,0}, \quad \hat{v}_e = \hat{v} + \hat{\eta}_{e,1},$$

$$2/e_0(\dot{\hat{x}}_e - A\hat{x}_e - \sum_i B_i\hat{u}_{e,i} - f) = \hat{\zeta}_{e,0} - \hat{p}_0,$$

$$2/e_i(\dot{\hat{x}}_e^i - A\hat{x}_e^i - \sum_{j\neq i} B_j\hat{u}_{e,j} - B_i\hat{v}_{e,i} - f) = -\hat{\zeta}_{e,1}^i - \hat{p}_i, \quad i = 1,\cdots,N$$

we obtain that, for

$$\max_{j=0,\cdots,N} e_j \leq [2K_1K_4^2\|(\hat{p}_0,\hat{p})\|_{[L^2]^{N+1}\times U^2\times[L^2]^{N+1}} \leq K_4/4\left(\max_{j=0,\cdots,N} e_j\right)\|(\hat{p}_0,\hat{p})\|_{[L^2]^{N+1}}.$$

Similarly,

$$\|\hat{u}_e - \hat{u}\|_U \le K_4/4\left(\max_{j=0,\cdots,N} e_j\right)\|(\hat{p}_0,\hat{p})\|_{[L^2]^{N+1}},$$

$$\|\hat{X}_e - \hat{X}\|_{[L^2]^N} \le K_4/4\left(\max_{j=0,\cdots,N} e_j\right)\|(\hat{p}_0,\hat{p})\|_{[L^2]^{N+1}},$$

$$\|\hat{v}_e - \hat{v}\|_U \le K_4/4\left(\max_{j=0,\cdots,N} e_j\right)\|(\hat{p}_0,\hat{p})\|_{[L^2]^{N+1}},$$

$$\left\|2/e_0(\dot{\hat{x}}_e - A\hat{x}_e - \sum_i B_i\hat{u}_{e,i} - f) - (-\hat{p}_0)\right\|_{L^2} \le K_4/4\left(\max_{j=0,\cdots,N} e_j\right)\|(\hat{p}_0,\hat{p})\|_{[L^2]^{N+1}},$$

$$\left\|2/e_i\left(\dot{\hat{x}}^i_e - A\hat{x}^i_e - \sum_{j\neq i} B_j\hat{u}_{e,i} - B_i\hat{v}_{e,i} - f\right) - (-\hat{p}_i)\right\|_{L^2}$$

$$\le K_4/4\left(\max_{j=0,\cdots,N} e_j\right)\|(\hat{p}_0,\hat{p})\|_{[L^2]^{N+1}}$$

The proof of Theorem 3.2 is complete.

Remark 3.5. From the proof given above, we see that assumption (B7) can be relaxed; we need only to require that the mixed partial derivative operators $\partial_x\partial_X F, \partial_x\partial_u F$, $\cdots$, etc. be dominated by $\partial_x^2 F, \partial_u^2 F, \partial_x^2 F, \partial_v^2 F$, at $(\hat{x},\hat{u};\hat{X},\hat{v})$.

Remark 3.6. Although $F_e(x,u;X,v)$ is concave in (X,v) for all $e_0,e_1,\cdots,e_N$, in general (without assumption (B0)) it is not necessarily true that $F_e(x,u;X,v)$ is concave in (x,u). Thus $(\hat{x}_e,\hat{u}_e;\hat{X}_e,\hat{v}_e)$ need not be a saddle point for F_e.

Corollary 3.7. Under the conditions of Theorem 3.2, assume, in addition, that $F(x,u;X,v)$ is quadratic in the sense that

$$F(\tilde{x},\tilde{u};\tilde{X},\tilde{v}) = F(x,u;X,v) + 2F'(x,u;X,v).(\tilde{x}-x,\tilde{u}-u;\tilde{X}-X,\tilde{v}-v)$$
$$+\langle F'(x,u;X,v)\cdot(\tilde{x}-x,\tilde{u}-u;\tilde{X}-X,\tilde{v}-v),(\tilde{x}-x,\tilde{u}-u;\tilde{X}-X,\tilde{v}-v)\rangle$$

holds for all $(x,u;X,v),(\tilde{x},\tilde{u};\tilde{X},\tilde{v}) \in H^1_{0,n}\times U\times[H^1_{0,n}]\times U$. Then Theorem 3.2 (i) can be strengthened to

$$\|(\hat{x}_e,\hat{u}_e;\hat{X}_e,\hat{v}_e)-(\hat{x},\hat{u};\hat{X},\hat{v})\|_H \le K_2'\left(\max_{j=0,\cdots,N} e_j\right)\|(\hat{p}_0,\hat{p})\|_{[L^2]^{N+1}} \tag{3.25}$$

for all $e_0,e_1,\cdots,e_N$ sufficiently small.

Proof. Since F is quadratic, so is F_e. Therefore $r(\xi,\eta)=0$ in the proof of Theorem 3.2. By (3.16), we have

$$D_e\begin{bmatrix}\hat{\xi}_e\\ \hat{\eta}_e\\ \hat{\zeta}_e\end{bmatrix} = \begin{bmatrix}0\\0\\ \tilde{p}_e\end{bmatrix}. \tag{3.26}$$

Thus

$$\|D_e(\hat{\xi}_e,\hat{\eta}_e,\hat{\zeta}_e)\|^2_{[L^2]^{N+1}\times U^2\times[L^2]^{N+1}} \le \|\tilde{p}_e\|^2_{[L^2]^{N+1}}.$$

For $e_0, e_1, \cdots, e_N$ sufficiently small, it is easily seen that there exist $K_5, K_6 > 0$ such that

$$K_5\|(\xi,\eta,\zeta)\|^2_{[L^2]^{N+1}\times U^2\times[L^2]^{N+1}} + \|D_e(\xi,\eta,\zeta)\|^2_{[L^2]^{N+1}\times U^2\times[L^2]^{N+1}}$$
$$\geq K_6\|(\xi,\eta,\zeta)\|^2_{[H^1_{n,0}]^{N+1}\times U^2\times[H^1_{n,0}]^{N+1}} \quad (3.27)$$

for all $(\xi,\eta,\zeta)\in[H^1_{n,0}]^{N+1}\times U^2\times[H^1_{n,0}]^{N+1}$, thanks to the coercivity

$$\|\bar{L}\xi\|^2_{[H^1_{n,0}]^{N+1}} \geq K_7\|\xi\|^2_{[H^1_{n,0}]^{N+1}},$$

$$\|\bar{L}^*\xi\|^2_{[H^1_{n,0}]^{N+1}} \geq K_7\|\zeta\|^2_{[H^1_{n,0}]^{N+1}}.$$

Combining (3.26) and (3.27) with Theorem 3.2 (i), we conclude (3.25).

References

[1] I. Babusks and A.K. Aziz, The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, A.K. Aziz, ed., Academic Press, New York, 1972.

[2] W.E. Bosarge and O.G. Johnson, Error bounds of high order accuracy for the state regulator problem via piecewise polynomial approximation, *SIAM J. Control*, **9** (1971), 15–28.

[3] G. Chen and W.H. Mills, Finite elements and terminal penalization for guadratic cost optimal control problems governed by ordinary differential equations, *SIAM J. Contri Opt.*, **19** (1981), 749–764.

[4] G. Chen, W.H. Mills , S. Sun and D. Yost, Sharp error estimates for a finite element-penalty approach to a class of regulator problems, *Math. Comp.*, **37** (1983), 151–173.

[5] G. Chen, Q. Zheng and J.X. Zhou, Minimax methods for open-loop equilibra in N-person differential games, Part I, Proceedings of Royal Society of Edinburgh, 103A (1986), 15–34

[6] G. Chen W.H. Mills, W. Shaw and Q. Zheng, Minimax methods for open-loop equilibra in N-person differential games, Part III, Preprint.

[7] P.G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam, 1978.

[8] T. Eisele, Non-existence and non-uniqueness of open-loop equilibra in linear-quadratic differential games, *J. Opt. Theory Appl.*, **37** (1982), 443–468.

[9] I. Ekeland and R. Temam, Convex Analysis and Variational Problems, North-Holland, American, Elsevier, New York, 1976.

[10] K. Fan, Sur un theorem minimax, *C.R. Acad. Sci., Paris,* **259** (1964), 3925–3928.

[11] A. Friedman, Differential Games, Wiley-Interscience, New York, 1971.

[12] W.W. Hager and S.K. Mitter, Lagrange duality theory for convex control problems, *SIAM J. Control Opt.*, **14** (1976), 843–856.

[13] R. Issacs, Differential Games, Wiley, New York, 1965.

[14] D.L. Lukes and D.L. Russell, A global theory for linear quadratic differential games, *J. Math. Anal. Appl.*, **33** (1971), 96–123.

[15] F.H. Maths and G.W. Rddien, Ritz-Treftz approximation in optimal control, *SIAM J. Control Opt.*, **17** (1979), 307–310.

[16] B.T. Polyak, The convergence rate of the penalty function method, *Zh. Vychisl. Mat. Fiz.* **11** (1971), 3–11.

[17] J. Ponstein, Approaches to the Theory of Optimisation, Cambridge Univ. Press, London, 1980.

[18] R.T. Rockafellar, Convex Analysis, Princeton Univ. Press, Princeton, New Jersey, 1970.

[19] R.T. Rockafellar, Conjugate Duality and Optimization, Conference Board of Mathematical Science, Ser., no. 16, SIAM Publ., 1974.

[20] D.L. Russell, Mathematics of Finite Dimensional Control System, Theory and Design, Marcel Dekker, New York, 1979.

[21] G. Strang and G. Fix, An Analysis of the Finite Element Method, Prentice-Hall, Englewood Cliffs, New Jersey, 1973.