Convergence of Lagrange-Newton method for control-state and pure state constrained optimal control problems

Kazimierz Malanowski

System Research Institute, Polish Academy of Sciences

ul. Newelska 6, 01-447 Warszawa, Poland.

e-mail: kmalan@ibspan.waw.pl

Abstract

The Lagrange-Newton method (LNM) is applied to a class of optimal control problems for nonlinear ODEs subject to pointwise mixed control-state and pure state constraints. Using the so called **two-norm approach**, developed in stability analysis of solutions to this class of problems, it is shown that (LNM) is locally superlinearly convergent and the rate of convergence is estimated.

Keywords

Lagrange-Newton method, sequential quadratic programming, optimal control, nonlinear ordinary differential equations, control-state and pure state constraints.

1 INTRODUCTION

The Lagrange-Newton method (LNM) is a highly effective sequential quadratic programming method, which is obtained by applying Newton's procedure to Kuhn-Tucker points of optimization problems.

In studying the local convergence of (LNM), Robinson's stability results for the solutions to generalized equations, or related concepts, have been used (see Robinson (1974), Robinson (1980), Alt (1990), Alt (1991), Alt, Malanowski (1993), Dontchev et al. (1995)). In this approach the crucial role is played by the concept of **strong regularity** (SRC). It can be shown that if (SRC) is satisfied, then (LNM) is locally quadratically convergent.

In case of optimal control problems subject to state constraints, due to the phenomenon of **two-norm discrepancy**, we are not able to assure that (SRC) is satisfied. To overcome this difficulty the so called **two-norm approach** can be used. This method is based on stability results for solutions of parametric optimization and optimal control problems developed in Malanowski (1993) and Malanowski (1995), where additional information

---

Supported by grant No.3 P403 002 05 from the State Committee for Scientific Research (KBN).

on regularity of the solutions is exploited. Using this method the conditions were derived in Alt, Malanowski (1995) under which (LNM), applied to control and state constrained optimal control problems is locally superlinearly convergent.

In the present paper the same approach is applied to more general class of optimal control problems with mixed initial-terminal boundary conditions and subject to pointwise mixed control-state and first order pure state constraints. Moreover, second order sufficient optimality conditions weaker than that in Alt, Malanowski (1995) are assumed. Using stability results for the solutions of this class of problems derived in Malanowski (1994), the local superlinear convergence of (LNM) is obtained and the rate of convergence is estimated.

Comprehensive numerical studies of application of (LNM) to state constrained optimal control problems can be found in Machielsen (1987).

2 LNM FOR ABSTRACT OPTIMIZATION PROBLEMS

In this section we recall convergence results for (LNM) applied to optimization problems with **two-norm discrepancy** in abstract Banach spaces. These results were obtained in Alt, Malanowski (1995).

Let Z and Y be two Banach spaces, the space of arguments and constraints, respectively. Moreover, two Hilbert spaces $\hat{Z}$ and $\hat{Y}$ are given, such that $Z \subset \hat{Z}$ and $Y \subset \hat{Y}$, with dense and continuous embeddings. The norms in Banach and Hilbert spaces are denoted by $\|\cdot\|$ and $\|\cdot\|$, respectively, with the appropriate subscript. By $(\cdot, \cdot)_{\hat{Z}}$ there is denoted a duality pairing between $\hat{Z}^*$ and $\hat{Z}$ or Z^* and Z . We put $\hat{Y}^* = \hat{Y}$ and by $(\cdot, \cdot)_{\hat{Y}}$ we denote the inner product in $\hat{Y}$, extended by continuity to $Y^* \times Y$. We define $X = Z \times Y$, and $\hat{X} = \hat{Z} \times \hat{Y}$. In Y there is given a closed convex cone K with the vertex at the origin. By $\hat{K}$ we denote the closure of K in $\hat{Y}$ and we put $K^+ = \{\lambda \in Y^* \mid (\lambda, y)_{\hat{Y}} \geq 0 \text{ for all } y \in K\}$.

We consider the following optimization problem

$$(P) \quad \min_{z \in Z} F(z) \quad \text{subject to } \varphi(z) \in K.$$

Assume:

- (I.1) There exists a possibly local solution $\tilde{z}$ of (P).
- (I.2) The functions $F : Z \rightarrow \mathbf{R}^1$ and $\varphi : Z \rightarrow Y$ are two times Fréchet differentiable in a neighborhood $\mathcal{O} \subset Z$ of $\tilde{z}$ and some compatibility conditions are satisfied, which assure that F_x, φ_x are well defined on $\hat{X}$. These conditions are always satisfied for optimal control problems (cf. Alt, Malanowski (1995)).
- (I.3) The local solution $\tilde{z}$ is regular in the sense of Robinson, i.e., $\varphi_z(\tilde{z})Z - K + [\varphi(\tilde{z})] = Y$, where $[y] = \{\alpha y \mid \alpha \in \mathbf{R}^1\}$.

Let us introduce the following Lagrangian associated with (P):

$$\mathcal{L} : Z \times Y^* \rightarrow \mathbf{R}^1, \quad \mathcal{L}(z, \lambda) = F(z) - (\lambda, \varphi(z))_Y.$$

By (I.3) there exists a Lagrange multiplier $\tilde{\lambda} \in Y^*$ associated with $\tilde{z}$ such that the Kuhn-Tucker conditions

$$\mathcal{L}_z(\tilde{z}, \tilde{\lambda}) := F_z(\tilde{z}) - \varphi_z(\tilde{z})^* \tilde{\lambda} = 0, \quad (\tilde{\lambda}, \varphi(\tilde{z}))_Y = 0, \quad \tilde{\lambda} \in K^+ \quad (1)$$

hold. We will assume that $\tilde{\lambda}$ is more regular:

$$(I.4) \quad \tilde{\lambda} \in Y.$$

In the Lagrange-Newton method applied to (P) the following iterative procedure is constructed:

(LNM) Having $w_k = (z_k, \lambda_k)$, compute a Kuhn-Tucker point $w_{k+1} = (z_{k+1}, \lambda_{k+1})$ of the following linear-quadratic problem:

$$(QP_{w_k}) \quad \min_{z \in Z} \tilde{F}(z, w_k) := \frac{1}{2}(\mathcal{L}_{zz}^2(z_k, \lambda_k)(z - z_k), z - z_k)_{\tilde{Z}} + (F_z(z_k), z - cz_k)_{\tilde{Z}} \\ \text{subject to } \varphi_z(z_k)(z - z_k) + \varphi(z_k) \in K.$$

If $w_k \rightarrow \bar{w} = (\bar{z}, \bar{\lambda})$, then $\bar{w}$ is a Kuhn-Tucker point of (P). If, in addition a sufficient optimality condition is satisfied at $\bar{w}$, then $\bar{z}$ becomes a solution of (P) and $\bar{\lambda}$ - an associated Lagrange multiplier.

(QP_{w_k}) can be treated as a linear-quadratic optimization problem depending on the parameter w_k . As in Alt, Malanowski (1995), to analyse convergence of (LNM) we will use stability results for parametric optimization problems with norm discrepancy, obtained in Malanowski (1993). To this end we will need the family of the following **accessory linear-quadratic problems**, depending on the perturbation $\delta = (a, b) \in \hat{X}^*$:

$$(LP_\delta) \quad \min_{z \in Z} \left\{ \frac{1}{2}(\mathcal{L}_{zz}^2(\tilde{z}, \tilde{\lambda})(z - \tilde{z}), z - \tilde{z})_{\tilde{Z}} + (F_z(\tilde{z}) - a, z - \tilde{z})_{\tilde{Z}} \right\} \\ \text{subject to } \varphi_z(\tilde{z})(z - \tilde{z}) + \varphi(\tilde{z}) - b \in \hat{K}.$$

The following assumptions constitute a weakened version of the so called **strong regularity condition** introduced in Robinson (1980).

(I.5) There exist a closed convex set $\Delta \subset \hat{X}^*$ of regular perturbations and a convex set Γ compact in X , such that $0 \in \Delta$ and for each $\delta \in \Delta$ there exists, a unique in Γ , Kuhn-Tucker point $\zeta_\delta = (x_\delta, \mu_\delta)$ of (LP_δ). Moreover, there exists $c > 0$ such that

$$\|x_{\delta_1} - x_{\delta_2}\|_{\tilde{Z}}, \|\mu_{\delta_1} - \mu_{\delta_2}\|_{\tilde{Y}} \leq c \|\delta_1 - \delta_2\|_{\hat{X}^*}, \quad \text{for all } \delta_1, \delta_2 \in \Delta. \quad (2)$$

As in Robinson (1980), Alt (1990), Alt (1991), Alt, Malanowski (1993), Alt, Malanowski (1995), for $w = (z_w, \lambda_w)$, $\xi = (z, \lambda) \in X$ we define the following functions:

$$\ell: X \times X \mapsto X^*, \quad \ell(\xi, w) = \begin{pmatrix} \ell_1(\xi, w) \\ \ell_2(\xi, w) \end{pmatrix}, \quad \text{where} \\ \ell_1(\xi, w) = F_z(\tilde{z}) + \mathcal{L}_{zz}^2(\tilde{z}, \tilde{\lambda})(z - \tilde{z}) - \varphi_z(\tilde{z})^* \lambda - \\ - F_z(z_w) - \mathcal{L}_{zz}^2(z_w, \lambda_w)(z - z_w) + \varphi_z(z_w)^* \lambda, \\ \ell_2(\xi, w) = \varphi(\tilde{z}) + \varphi_z(\tilde{z})(z - \tilde{z}) - \varphi(z_w) - \varphi_z(z_w)(z - z_w).$$

(I.6) $\ell(\xi, w) \in \Delta$ for all $\xi, w \in \Gamma$.

REMARK 21 In the original paper Robinson (1980), as well as in the other papers dealing with the convergence of the Lagrange-Newton method (cf. e.g., Alt (1990), Alt (1991), Alt, Malanowski (1993), Dontchev et al. (1995)) the so called **strong regularity condition** was used instead of (I.5). In that condition it is required that (z_δ, μ_δ) are locally Lipschitz continuous functions of δ in the stronger norm of the original space X , for all perturbations from a ball in X^* . However, in applications to nonlinear optimal control problems

subject to pure state constraints we can expect at most that the weaker condition (2) is satisfied (cf. Section 4 in Malanowski (1993)). Since Γ is a compact set, (I.5) imposes some **regularity conditions** on the Kuhn-Tucker points of (LP_δ) , whereas assumption (I.6), assures that the set of perturbations Δ is **sufficiently rich**. These assumptions were not needed in Robinson (1980).

It can be shown (cf. Lemma 2.1 in Malanowski (1993)) that, by compactness of Γ , there exists a continuous non-decreasing function $\zeta : \mathbf{R}^+ \mapsto \mathbf{R}^+$, $\zeta(0) = 0$, such that

$$\|w - \tilde{w}\|_X \leq \zeta(\|w - \tilde{w}\|_{\hat{X}}) \quad \text{for all } w \in \Gamma. \quad (3)$$

The following result is proved in Alt, Malanowski(1995) (cf. Theorem 2.5):

THEOREM 22 *If assumptions (I.1)-(I.6) are satisfied then there exists a ball $\mathcal{B}_{\hat{X}}(\tilde{w}, \rho) := \{w \in \hat{X} \mid \|w - \tilde{w}\|_{\hat{X}} \leq \rho\}$ of radius ρ about $\tilde{w}$ such that for each starting point $w_0 \in \Gamma \cap \mathcal{B}_{\hat{X}}(\tilde{w}, \rho)$ the Lagrange-Newton method generates a unique sequence $\{w_k\}$, $w_k = (z_k, \lambda_k)$ convergent to $\tilde{w}$ and*

$$\|w_{k+1} - \tilde{w}\|_{\hat{X}} / \|w_k - \tilde{w}\|_{\hat{X}} \leq c \zeta(\|w_k - \tilde{w}\|_{\hat{X}}). \quad (4)$$

REMARK 23 Since $\zeta(\cdot)$ is a non-decreasing function and $\lim_{k \rightarrow \infty} \zeta(k) = 0$, (4) shows that the Lagrange-Newton method converges superlinearly for any starting point $w_0 \in \Gamma \cap \mathcal{B}_{\hat{X}}(\tilde{w}, \rho)$. In case without two-norm discrepancy, $\zeta(\|w_k - \tilde{w}\|_{\hat{X}})$ in (4) is substituted by $\|w_k - \tilde{w}\|_{\hat{X}}$, and we obtain quadratic convergence (cf., e.g., Theorem 3.3 in Alt(1990)).

3 OPTIMAL CONTROL PROBLEM

In this section we introduce optimal control problems subject to mixed control-state and pure state constraints. We recall stability results for the solutions of these problems derived in Malanowski (1994). These results are used in the convergence analysis of (LNM).

Let us consider the following optimal control problem:

$$\begin{aligned} \text{(O)} \quad & \text{Find } (\tilde{x}, \tilde{u}) \in W^{1,\infty}(0, T; \mathbf{R}^n) \times L^\infty(0, T; \mathbf{R}^m) \text{ such that} \\ & F(\tilde{x}, \tilde{u}) = \min \{ F(x, u) := \int_0^T f^0(x(t), u(t)) dt + g(x(T)) \} \\ & \text{subject to} \\ & \dot{x}(t) - f(x(t), u(t)) = 0, \quad \text{for a.a. } t \in [0, T], \\ & \xi(x(0), x(T)) = 0, \\ & \theta(x(t), u(t)) \leq 0, \quad \text{for a.a. } t \in [0, T], \\ & v(x(t)) \leq 0, \quad \text{for all } t \in [0, T], \end{aligned}$$

where $\xi : \mathbf{R}^n \times \mathbf{R}^n \mapsto \mathbf{R}^d$, $\theta : \mathbf{R}^n \times \mathbf{R}^m \mapsto \mathbf{R}^k$, $v : \mathbf{R}^n \mapsto \mathbf{R}^l$.

Denote by $I = \{1, \dots, k\}$ and $J = \{1, \dots, l\}$ the sets of indices of constraints.

We are going to apply (LNM) to find the solution of (O) and we will use Theorem 22 to analyse convergence of the method. To do that we will use stability results obtained in Malanowski (1994). Let us recall assumptions from that paper. It is assumed that:

(II.1) There exists a possibly local solution $(\tilde{x}, \tilde{u}) \in C^1(0, T; \mathbf{R}^n) \times C(0, T; \mathbf{R}^m)$ of (O).

(II.2) Functions $f^0(\cdot, \cdot), g(\cdot), f(\cdot, \cdot), \xi(\cdot, \cdot), \theta(\cdot, \cdot), v(\cdot)$ and $v_x(\cdot)$ are two times Fréchet differentiable in all arguments, and the respective derivatives are locally Lipschitz continuous.

In order to represent (O) in the form (P) we put

$$Z = W^{1,\infty}(0, T; \mathbf{R}^n) \times L^\infty(0, T; \mathbf{R}^m), \quad \widehat{Z} = W^{1,2}(0, T; \mathbf{R}^n) \times L^2(0, T; \mathbf{R}^m).$$

Moreover we put $F(z) = F(x, u)$ and $\varphi(z) = (\dot{x} - f(x, u), \xi(x(0), x(T)), -\theta(x, u), -v(x))$. Hence it is natural to choose

$$Y = L^\infty(0, T; \mathbf{R}^n) \times \mathbf{R}^d \times L^\infty(0, T; \mathbf{R}^k) \times W^{1,\infty}(0, T; \mathbf{R}^l), \\ \widehat{Y} = L^2(0, T; \mathbf{R}^n) \times \mathbf{R}^d \times L^2(0, T; \mathbf{R}^k) \times W^{1,2}(0, T; \mathbf{R}^l),$$

with $K = K_1 \times K_2 \times K_3 \times K_4$, where

$$K_1 = \{0\} \quad K_2 = \{0\}, \\ K_3 = \{u \in L^\infty(0, T; \mathbf{R}^k) \mid u^i(t) \geq 0, \quad i = 1, \dots, k \text{ for a.a. } t \in [0, T]\}, \\ K_4 = \{x \in W^{1,\infty}(0, T; \mathbf{R}^l) \mid x^j(t) \geq 0, \quad j = 1, \dots, l \text{ for all } t \in [0, T]\}.$$

The cone $\widehat{K}$ is defined as above but with Y substituted by $\widehat{Y}$.

Note that by (II.2) conditions (I.2) are satisfied. Now we will introduce constraint qualifications. To simplify notation let us put for $w(t) = (x(t), u(t))$:

$$A(w(t)) := f_x(x(t), u(t)), \quad B(w(t)) := f_u(x(t), u(t)), \\ \Xi_0(w) := \xi_{x(0)}(x(0), x(T)), \quad \Xi_T(w) := \xi_{x(T)}(x(0), x(T)), \\ \Theta_x(w(t)) := \theta_x(x(t), u(t)), \quad \Theta_u(w(t)) := \theta_u(x(t), u(t)), \\ \Upsilon(w(t)) := v_x(x(t)).$$

Let us introduce the sets of indices of active constraints $\widehat{I}(t) = \{i \in I \mid \theta^i(\tilde{x}(t), \tilde{u}(t)) = 0\}$, $\widehat{J}(t) = \{j \in J \mid v^j(\tilde{x}(t)) = 0\}$, and denote by $\widehat{\Theta}_x(t)$, $\widehat{\Theta}_u(t)$ and $\widehat{\Upsilon}(t)$ the matrices, whose rows are the functions $\theta_u^i(\tilde{x}(t), \tilde{u}(t))$, $\theta_x^i(\tilde{x}(t), \tilde{u}(t))$ for $i \in \widehat{I}(t)$ and $v_x(\tilde{x}(t))$ for $j \in \widehat{J}(t)$, respectively. We assume that the following constraint qualifications hold:

(II.3) There exists $\beta > 0$ such that $\left| \left[B(\tilde{w}(t))^* \widehat{\Upsilon}(\tilde{w}(t))^*, \widehat{\Theta}_u(\tilde{w}(t))^* \right] \chi \right| \geq \beta |\chi|$ for all χ of appropriate dimension and all $t \in [0, T]$.

(II.4) For any $\vartheta \in \mathbf{R}^d$ there exists a solution (y, v) to the following system of equations (complete controllability condition is satisfied):

$$\dot{y}(t) - A(\tilde{w}(t))y(t) - B(\tilde{w}(t))v(t) = 0, \\ \Xi_0(\tilde{w})y(0) + \Xi_T(\tilde{w})y(T) = \vartheta, \\ \text{subject to} \\ \widehat{\Theta}_x(\tilde{w}(t))y(t) + \widehat{\Theta}_u(\tilde{w}(t))v(t) = 0, \\ \widehat{\Upsilon}(\tilde{w}(t))[A(\tilde{w}(t))y(t) + B(\tilde{w}(t))v(t)] = 0.$$

REMARK 31 It is shown in Malanowski (1994) that (II.4) can be expressed as a rank condition for a certain controllability matrix.

It follows from Lemma 3.3 in Malanowski (1994) that, by (II.3) and (II.4) constraints qualification of the type of **stable linear independence of gradients of active constraints** are satisfied (cf. condition (A3) in Malanowski (1992)). In particular it implies that (I.3) holds.

Let us introduce the following Lagrangian associated with (O):

$$\begin{aligned} \mathcal{L} : W^{1,\infty}(0, T; \mathbf{R}^n) \times L^\infty(0, T; \mathbf{R}^m) \times (L^\infty(0, T; \mathbf{R}^n))^* \times \mathbf{R}^d \times \\ \times (L^\infty(0, T; \mathbf{R}^k))^* \times (W^{1,\infty}(0, T; \mathbf{R}^l))^* , \\ \mathcal{L}(x, u, q, \rho, \kappa, \mu) = F(x, u) + \langle q, \dot{x} - f(x, u) \rangle + \langle \rho, \xi(x(0), x(T)) \rangle + \langle \kappa, \theta(x, u) \rangle + \\ + \langle \mu(0), v(x(0)) \rangle + \langle \dot{\mu}, v_x(x) f(x, u) \rangle. \end{aligned} \quad (5)$$

By Lemma 3.3 in Malanowski (1994) there exist unique Lagrange multipliers $(\tilde{q}, \tilde{\rho}, \tilde{\kappa}, \tilde{\mu}) \in \hat{X}^*$ such that the first order necessary optimality conditions in the Kuhn-Tucker form are satisfied, and in the same way as in Hager (1979), it can be shown that $\tilde{q}, \tilde{\kappa}, \tilde{\mu}$ are continuous on $[0, T]$.

For any $\alpha \geq 0$ define the sets $I_\alpha^+(t) = \{i \in \hat{I}(t) \mid \tilde{\kappa}^i(t) > \alpha\}$, and denote by $\Theta_x^\alpha(t)$ and $\Theta_u^\alpha(t)$ the matrices, whose rows are the functions $\theta_x^i(\tilde{x}(t), \tilde{u}(t))$ and $\theta_u^i(\tilde{x}(t), \tilde{u}(t))$, respectively, for $i \in I_\alpha^+(t)$. Define the following subspaces $\mathcal{D}_\alpha(t) = \{v \in \mathbf{R}^m \mid \Theta_u^\alpha(t)v = 0\}$ of $\mathbf{R}^m$ and introduce the augmented Hamiltonian

$$\mathcal{H}(t) = f^0(\tilde{x}(t), \tilde{u}(t)) - \langle \tilde{q}(t), f(\tilde{x}(t), \tilde{u}(t)) \rangle + \langle \tilde{\kappa}(t), \theta(\tilde{x}(t), \tilde{u}(t)) \rangle + \langle \dot{\tilde{\mu}}(t), v_x(\tilde{x}(t)) f(\tilde{x}(t), \tilde{u}(t)) \rangle.$$

In addition to (II.1)-(II.4), we assume :

(II.5) There exists a constant $\gamma > 0$ such that

$$\langle u, \mathcal{H}_{uu}^2(t)u \rangle \geq \gamma |u|^2, \quad \gamma > 0, \quad \text{for all } u \in \mathcal{D}_0(t) \text{ and all } t \in [0, T].$$

(II.6) There exist $\bar{\alpha} > 0$ and $\bar{\gamma} > 0$ such that the following coercivity condition holds:

$$((y, v), \begin{pmatrix} \tilde{\mathcal{L}}_{xx}^2 & \tilde{\mathcal{L}}_{xu}^2 \\ \tilde{\mathcal{L}}_{ux}^2 & \tilde{\mathcal{L}}_{uu}^2 \end{pmatrix} (y, v)) \geq \bar{\gamma} (\|y\|_2^2 + \|v\|_2^2) \quad (6)$$

for all $(y, v) \in W^{1,2}(0, T; \mathbf{R}^n) \times L^2(0, T; \mathbf{R}^n)$ such that

$$\begin{aligned} \dot{y}(t) - A(\tilde{w}(t))y(t) - B(\tilde{f}(t))v(t) &= 0, \\ \Xi_0(\tilde{w})y(0) + \Xi_T(\tilde{w})y(T) &= 0, \\ \Theta_x^{\bar{\alpha}}(t)y(t) + \Theta_u^{\bar{\alpha}}(t)v(t) &= 0. \end{aligned} \quad (7)$$

REMARK 32 Note that in (6) coercivity holds not in the norm of the space Z in which problem (O) is well defined and differentiable, but in the weaker norm of the space $\hat{Z}$ in which, in general, (O) is not well defined. This phenomenon is called **two-norm discrepancy**.

In Malanowski, Maurer (1994) condition (II.6) is expressed in terms of existence of a bounded solution to an auxiliary Riccati equation, satisfying certain boundary conditions.

In Malanowski (1994) it is shown that if assumptions (II.1)-(II.6) hold then $(\tilde{x}, \tilde{u})$ is a locally isolated, local minimizer of (O).

In Alt, Malanowski (1995) a coercivity condition stronger than (6) was used. Namely, the last equation in (7) was not present.

We are going to analyse convergence of (LNM) for optimal control problem (O). We will use abstract results presented in Section 2 in the same way as in Alt, Malanowski (1995), where less general optimal control problem was considered. The basis of our analysis will be the stability results for parametric optimal control problems derived in Malanowski (1994). Let us denote $w_k = (x_k, u_k, q_k, \rho_k, \kappa_k, \mu_k)$. For optimal control problem (O), the quadratic optimization problem (QP $_{w_k}$) used in (LNM) takes on the form:

$$\begin{aligned}
 (\text{QO}_{w_k}) \quad & \min_{(y,v) \in Z} \tilde{F}(y, v, w_k) := \int_0^T \tilde{f}^0(y(t), v(t), w_k(t)) dt + \tilde{g}(y(0), y(T), w_k) \\
 & \text{subject to} \\
 & \tilde{y}(t) - f(x_k(t), u_k(t)) - A(w_k(t))(y(t) - x_k(t)) - B(w_k(t))(v(t) - u_k(t)) = 0, \\
 & \xi(x_k(0), x_k(T)) + \Xi_0(w_k)(y(0) - x_k(0)) + \Xi_T(w_k)(y(T) - x_k(T)) = 0, \\
 & \theta(x_k(t), u_k(t)) + \Theta_x(w_k(t))(y(t) - x_k(t)) + \Theta_u(w_k(t))(v(t) - u_k(t)) \leq 0, \\
 & v(x_k(t)) + \Upsilon(w_k(t))(y(t) - x_k(t)) \leq 0,
 \end{aligned}$$

where $\tilde{F}(y, v, w_k)$ is defined as in (QP $_{w_k}$), using Lagrangian (5).

In view of the form of the mapping given by the derivatives of the Lagrangian (5), we choose as the space of perturbations the space

$$\begin{aligned}
 \hat{X}^* = & L^2(0, T; \mathbf{R}^n) \times L^2(0, T; \mathbf{R}^m) \times \mathbf{R}^n \times \mathbf{R}^n \times \\
 & \times L^2(0, T; \mathbf{R}^n) \times \mathbf{R}^d \times L^2(0, T; \mathbf{R}^k) \times W^{1,2}(0, T; \mathbf{R}^l).
 \end{aligned}$$

For $\delta = (a_1, a_2, a_3, a_4, b_1, b_2, b_3, b_4) \in \hat{X}^*$, the problem (LO $_{\delta}$), corresponding to (LP $_{\delta}$) is defined as follows:

$$\begin{aligned}
 (\text{LO}_{\delta}) \quad & \min_{(y,v) \in Z} \tilde{F}(y, v, \delta) := \\
 & := \int_0^T \left(\tilde{f}^0(y(t), v(t), \tilde{w}(t)) - \langle a_1, y(t) - \tilde{x}(t) \rangle - \langle a_2, v(t) - \tilde{u}(t) \rangle \right) dt + \\
 & + \tilde{g}(y(0), y(T), \tilde{w}) - \langle a_3, y(0) - \tilde{x}(0) \rangle - \langle a_4, y(T) - \tilde{x}(T) \rangle \\
 & \text{subject to} \\
 & \dot{y}(t) - f(\tilde{x}(t), \tilde{u}(t)) - A(\tilde{w}(t))(y(t) - \tilde{x}(t)) - B(\tilde{w}(t))(v(t) - \tilde{u}(t)) - b_1(t) = 0, \\
 & \xi(\tilde{x}(0), \tilde{x}(T)) + \Xi_0(\tilde{w})(y(0) - \tilde{x}(0)) + \Xi_T(\tilde{w})(y(T) - \tilde{x}(T)) - b_2 = 0, \\
 & \theta(\tilde{x}(t), \tilde{u}(t)) + \Theta_x(\tilde{w}(t))(y(t) - \tilde{x}(t)) + \Theta_u(\tilde{w}(t))(v(t) - \tilde{u}(t)) - b_3(t) \leq 0, \\
 & v(\tilde{x}(t)) + \Upsilon(\tilde{w}(t))(y(t) - \tilde{x}(t)) - b_4(t) \leq 0,
 \end{aligned}$$

We have to verify assumptions (1.5) and (1.6) of Theorem 22.

In the same way as in Malanowski (1994) (cf. also Malanowski (1995) and Alt, Malanowski (1995)), as the set $\Gamma \in X$ of Kuhn-Tucker points of (LO $_{\delta}$) we choose the set of functions $\mathbf{x} = (y, v, r, \sigma, \lambda, \nu)$ having the same regularity as $\tilde{\mathbf{x}} = (\tilde{x}, \tilde{u}, \tilde{\rho}, \tilde{\kappa}, \tilde{\mu})$ i.e., uniformly bounded and uniformly Lipschitz continuous, with the same bound and Lipschitz modulus. By the Arzela-Ascoli theorem the set Γ is compact in X .

It follows from Lemmas 6.6 and 6.7 in Malanowski (1995) that if the set $\Delta \subset X^*$ of regular perturbations is chosen also as the set of uniformly bounded and uniformly Lipschitz continuous functions, with sufficiently small Lipschitz modulus, then, for each $\delta \in \Delta$ there exists a unique in Γ Kuhn-Tucker point $\mathbf{x}_{\delta} := (z_{\delta}, w_{\delta}, r_{\delta}, \sigma_{\delta}, \pi_{\delta}, \nu_{\delta})$ of (LO $_{\delta}$) and

$$\|\mathbf{y}_{\delta_1} - \mathbf{y}_{\delta_2}\|_{1,2}, \|\mathbf{v}_{\delta_1} - \mathbf{v}_{\delta_2}\|_2, \|r_{\delta_1} - r_{\delta_2}\|_{1,2}, \|\sigma_{\delta_1} - \sigma_{\delta_2}\|, \|\pi_{\delta_1} - \pi_{\delta_2}\|_2, \|\nu_{\delta_1} - \nu_{\delta_2}\|_{1,2} \leq c \|\delta_1 - \delta_2\|_{\hat{X}^*}.$$

for all $\delta_1, \delta_2 \in \Delta$.

Thus all assumptions of Theorem 22 are satisfied.

Moreover, it was shown in Section 7 of Alt, Malanowski (1995) that the function ζ in (3) has the form $\zeta(r) = c r^{\frac{2}{3}}$. Hence we obtain the following convergence result for (LNM):

THEOREM 33 *If assumptions (II.1)-(II.6) are satisfied, then there exists $\rho > 0$ such that for each starting point $w_1 \in \Gamma \cap \mathcal{B}_{\hat{X}}(\tilde{w}, \rho)$ the Lagrange-Newton method applied to (O) generates a unique sequence $\{w_k\}$ convergent to $\tilde{w}$ with*

$$\|w_{k+1} - \tilde{w}\|_{\hat{X}} \leq c \|w_k - \tilde{w}\|_{\hat{X}}^{\frac{5}{3}} \quad \text{and} \quad \|w_{k+1} - \tilde{w}\|_X \leq c \|w_k - \tilde{w}\|_X^{\frac{4}{3}}.$$

REFERENCES

- Alt, W. (1990) The Lagrange-Newton method for infinite-dimensional optimization problems. *Numerical Functional Analysis and Optimization*, **11**, 201–24.
- Alt, W. (1991) Parametric programming with applications to optimal control and sequential quadratic programming. *Bayreuther Mathematische Schriften*, **35**, 1–37.
- Alt, W. and Malanowski, K. (1993) The Lagrange-Newton method for nonlinear optimal control problems. *Computational Optimization and Applications*, **2**, 77–110.
- Alt, W. and Malanowski, K. (1995) The Lagrange-Newton method for state constrained optimal control problems. *Computational Optimization and Applications*, **4**, 217–39.
- Dontchev, A.L., Hager, W.W., Poore, A.B. and Yang, B. (1995) Optimality, stability and convergence in nonlinear control. *Applied Mathematics and Optimization*, **31**, 297–326.
- Hager, W.W. (1979) Lipschitz continuity for constrained processes. *SIAM Journal on Control and Optimization*, **17**, 321–38.
- Machielsen, K.C.P. (1987) Numerical solution of optimal control problems with state constraints by sequential quadratic programming in function space *CWI Tract*, **53**, Amsterdam.
- Malanowski, K. (1992) Second order conditions and constraint qualifications in stability and sensitivity analysis of solutions in optimization problems in Hilbert spaces. *Applied Mathematics and Optimization*, **21**, 98–110.
- Malanowski, K. (1993) Two-norm approach in stability and sensitivity analysis of optimization and optimal control problems. *Advances in Mathematical Sciences and Applications*, **2**, 397–443.
- Malanowski, K. (1994) Sufficient optimality conditions in optimal control. *System Research Institute*, **WP-1-1994**, Warszawa.
- Malanowski, K. (1995) Stability and sensitivity of solutions to nonlinear optimal control problems (to appear).
- Malanowski, K. and Maurer, H. (1994) Sensitivity analysis for parametric optimal control problems with control-state constraints. *Universität Münster, Angewandte Mathematik und Informatik*, **4/94-N**, Münster.
- Robinson, S.M. (1974) Perturbed Kuhn-Tucker points and rates of convergence for a class of nonlinear-programming algorithms. *Mathematical Programming*, **7**, 1–16.
- Robinson, S.M. (1980) Strongly regular generalized equations. *Mathematics of Operations Research*, **5**, 43–62.