On an Algorithm for Optimal Control Using an Augmented Hamiltonian

U. FLEMMIG and H. POPPE

An algorithm on optimal control due to Y. Sakowa and Y. Shindo is considered. It is based on Pontryagin's minimum principle and the Hamiltonian is extended by a penalty term. Using another penalty term for this algorithm, under suitable assumptions we can show that each sequence of control vectors generated by the algorithm converges with respect to the L_1-norm to an admissible control fulfilling the first order optimality condition.

Key words: optimal control, optimization algorithms, Potryagins minimum principle, global convergence of sequences of control functions
AMS subject classification: 49 D 99, 93 C 10

1. Introduction

We consider a general control problem with fixed time and Lagrange functional governed by a system of ordinary differential equations. In [4] SAKAWA and SHINDO proposed an iterative algorithm to solve problems of such kind using Pontryagins minimum principle. They defined an augmented Hamiltonian, which is the sum of the Hamiltonian and a quadratic term penalizing the difference between the new and the old controls, and the new control minimizes at each time the augmented Hamiltonian. This term contains also a real parameter. By BONNANS [1] this algorithm as considered too. The mentioned authors proved some properties of convergence for this algorithm, but they did not prove the convergence in the metric of some function spaces of a sequence of controls generated by the algorithm to an admissible control, which fulfils the (first order) optimality condition.

We want to treat the problem in the space $L_1(T, \mathbf{R}^m)$ and as a penalty term we use the L_1-norm of the difference between the new and the old controls. We use rather mild assumptions (in essential only the existence and continuity of first order partial derivatives of the relevant functions is assumed), as compared with the assumptions used by Sakawa and Bonnans, respectively. Applying these assumptions we deduce an inequality which estimates the difference of two successive values of the cost functional by the L_1-norm (Theorem 1). This inequality is analogous to inequalities which were obtained by Sakawa and Shindo and Bonnans, respectively, using their approaches. We further show that any sequence (u^k) of control functions generated by the algorithm converges to an admissible control $\hat{u}$ with respect to the L_1-norm. In order to show that $\hat{u}$ fulfils the first order optimality condition we must assume some second order conditions of differentiability for the functions L and f (see Theorem 2).

2. Problem

We consider the system of differential equations

$$\dot{x}(t) = f(x(t), u(t), t), \ t \in T = \left[t_0, t_1\right], \ x(t_0) = x_0 . \tag{1}$$

$x(t) \in X \subset \mathbf{R}^n$ and $u(t) \in U \subset \mathbf{R}^m$ denoting the state vector and control vector, respecti-

vely. The control problem is to find an admissible control $u^* \epsilon U_{ad}$ such that u^* minimizes the cost functional

$$J(u) = \int_{t_0}^{t_1} L(x(t), u(t), t) \, dt + g(x(t_1))$$ (2)

x denoting the solution of system (1) corresponding to u and

$$U_{ad} = \left\{ u : T \to \mathbf{R}^m \mid u \text{ measurable, } u(t) \epsilon U \text{ for a.a. } t \epsilon T \right\} ,$$

where U is a convex and compact subset of $\mathbf{R}^m$ and u *measurable* means Lebesgue measurability of $u^{-1}(B)$ for every Borel subset $B \subset \mathbf{R}^m$. We denote by $L_1(T, \mathbf{R}^m)$ the usual Lebesgue space. (See, e.g.,[5].) We define the augmented Hamiltonian of this problem:

$$K_\epsilon \Big(x(t), u(t), v(t), \lambda(t), t \Big)$$

$$= L\Big(x(t), u(t), t\Big) + \Big(\lambda(t), f(x(t), u(t), t)\Big) + \epsilon^{-1} \int_T \|u(t) - v(t)\| \, dt$$

$$= H\Big(x(t), u(t), \lambda(t), t\Big) + \epsilon^{-1} \int_T \|u(t) - v(t)\| \, dt .$$

We want to remark that in case of the control problem (1), (2) it is known that one can choose $\lambda_o = 1$ for the additional Lagrange multiplier.

The costate equation we can write as

$$d\lambda(t)/dt = -\partial H(x(t), u(t), \lambda(t), t)/\partial x \quad \text{a.e. on } T$$ (3)

$$\lambda(t_1) = \partial g(x(t_1))/\partial x, \text{ where } \lambda(t) \epsilon \mathbf{R}^n \quad (\text{condition of transversality}).$$

Algorithm:

Step 1: Let be given an admissible control u^o and a positive sequence $(\epsilon^k)_{k \geq 1} \subset \mathbf{R}$; set $k = 0$. Compute the state x^o associated to u^o.

Step 2: Compute the costate λ^k associated to u^k.

Step 3: Set $k = k + 1$. Compute u^k, x^k such that

$$K_{\epsilon^k}(x^k, u^k, u^{k-1}, \lambda^{k-1}, t) \leq K_{\epsilon^k}(x^k, u, u^{k-1}, \lambda^{k-1}, t)$$

for all $u(t) \epsilon U$, a.e. on T, where $u^k := u_{\epsilon^k}(x^k, u^{k-1}, \lambda^{k-1}, t) \epsilon U_{ad}$.

Step 4: If a convergence test is satisfied, then stop; else go to Step 2.

We need still the following property of U_{ad}.

Lemma 1: U_{ad} *is closed in the norm topology of* $L_1(T, \mathbf{R}^m)$.

Proof: If $u, u^k \epsilon L_1(T, \mathbf{R})$ are such that $\|u^k - u\|_{L_1} \to 0$, then by a well-known theorem we find a subsequence (u^{ki}) of (u^k) such that $u^{ki} \to u$ a.e. As is easily seen this statement also holds for the space $L_1(T, \mathbf{R}^m)$. Now let $(u^k) \subset U_{ad}$, $u \epsilon L_1(T, \mathbf{R}^m)$ and $\|u^k - u\|_{L_1} \to 0$. Let $u^{ki} \to u$ a.e., and let t be any element of T such that $u^{ki}(t) \to u(t)$. We have $u^{ki}(t) \epsilon U$ for all i and since U is compact, U is closed, too. Hence $u(t) \epsilon U$. Therefore we have $u(t) \epsilon U$ for almost all $t \epsilon T$ which proves $u \epsilon U_{ad}$ ∎

Now let us formulate the following **Hypotheses**:

(H 1) $L(\cdot, \cdot, \cdot)$ and $f(\cdot, \cdot, \cdot)$ are continuous and $L(\cdot, \cdot, t)$ and $f(\cdot, \cdot, t)$ are continuously differentiable, a.e. on T.

(H2) g is continuously differentiable from $\mathbf{R}^n$ to $\mathbf{R}$.

(H3) There exists a positive constant M such that

$$\|f(x(t), u(t), t)\| \le M(1 + \|x(t)\|) \quad \forall (x(t), u(t)) \in \mathbf{R}^n \times U, \text{ a.e. on } T.$$

(H4) $H(x(t), u(t), \lambda(t), t)$ is convex with respect to $u(t)$.

We need some properties of the solutions of the state and costate equations, respectively. At first we remark that for each $u \in U_{ad}$ equations (1) and (3) have unique solutions on T (see, e.g., [3]). Moreover, these solutions have useful boundedness properties, which we formulate in the following

Lemma 2 : *By assumption* (H3) *we obtain*

(i) *a uniformly bounded solution* $x(t)$ *of system* (1) ; *i.e.* $\|x(t)\| \le M_1$, *where* M_1 *is a positive constant not depending on* t *and* u *(see* [1]*);*

(ii) *a uniformly bounded solution* $\lambda(t) \in \mathbf{R}^n$ *of system* (3) *for any* $t \in T$; *i.e.* $\|\lambda(t)\| \le M_2$, *where* M_2 *is a positive constant not depending on* t.

We can define the compact and convex sets

$$X = \left\{ x \in \mathbf{R}^n \mid \|x\| \le M_1 \right\} \quad \text{and} \quad \Lambda = \left\{ \lambda \in \mathbf{R}^n \mid \|\lambda\| \le M_2 \right\}.$$

Proof of Lemma 2: Those of (i) and (ii) are similar. That is why we give only the proof of (ii). It is true that $d\lambda(t)/dt = -H_x(x(t), u(t), \lambda(t), t)$. If we use the definition of the Hamiltonian, then we get

$$d\lambda(t)/dt = -\lambda(t)^\tau f_x(x(t), u(t), t) - L_x(x(t), u(t), t).$$

Since f_x and L_x are continuous on the compact set $X \times U \times T$, therefore we can find positive constants c_1, c_2 such that

$$\|f_x(x(t), u(t), t)\| \le c_1 \quad \text{and} \quad \|L_x(x(t), u(t), t)\| \le c_2.$$

We get $\|\dot\lambda(t)\| \le \|\lambda(t)\| c_1 + c_2$. Using the time transformation $\tau = t_1 - t$ $(\tau \in [0, |T|], |T| = t_1 - t_0)$ we estimate

$$d\|\lambda(t)\|/d\tau \le \|d\lambda(t)/d\tau\| = \|\dot\lambda(t)\| \le \|\lambda(t)\| c_1 + c_2.$$

By integration from 0 to τ and estimation we get

$$\|\lambda(t_1 - \tau)\| - \|\lambda(t_1)\| \le c_1 \int_0^\tau \|\lambda(t_1 - \tau)\| d\tau + c_2 \int_0^{|T|} d\tau$$

and therefore $\|\lambda(t_1 - \tau)\| \le c + \int_0^\tau c_1 \|\lambda(t_1 - \tau)\| d\tau$. Now we use the Lemma of Gronwall and our assertion is proved ∎

3. Remarks about existence and uniqueness of functions u^k, u^k as generated by the algorithm

We assume that we have found functions $x^{k-1}, u^{k-1}, \lambda^{k-1}$ by the algorithm. Now we want to construct the corresponding functions x^k, u^k, λ^k.

1. Since $K_{\varepsilon k}$ is continuous in all its variables and U is compact, for fixed $t \in T$, arbitrary $x \in X$, $v \in U$, $\lambda \in \Lambda$ we find $u_{\varepsilon k}(t) \in U$ such that

$$K_{\varepsilon k}\big(x, u_{\varepsilon k}(t), v, \lambda, t\big) \leq K_{\varepsilon k}\big(x, u, v, \lambda, t\big) \quad \forall u \in U.$$

Especially for fixed $t \in T$ we find $u^k(t) \in U$ such that

$$K_{\varepsilon k}\big(x^k, u^k, u^{k-1}, \lambda^{k-1}, t\big) \leq K_{\varepsilon k}\big(x^k, u, u^{k-1}, \lambda^{k-1}, t\big) \quad \forall u \in U. \tag{4}$$

Since U is also convex and if we would assume that $K_{\varepsilon k}$ is a strictly convex function with respect to u, then we would find a unique $u^k(t) \in U$.

2. We consider a partition of the time interval $T = [t_0, t_1]$ into n (equal) parts and we want to show how to construct a pair $\big(x_n^k, u_n^k\big)$ approximating (x^k, u^k), where x^k is the state corresponding to u^k, and x^k, u^k fulfil inequality (4) (for this construction, see also [4]). Let be $t_0 = a_0 < a_1 < \ldots < a_n = t_1$. Since $x^k(t_0) = x_0$, $x^k(a_0)$ is known and hence we find $u_n^k(a_0) \in U$ which satisfies (for $t = a_0$) (4). Now, for $t \in [a_0, a_1]$ let $u_n^k(t) = u_n^k(a_0)$, and setting $u = u_n^k$ we integrate the differential equation (1) on $[a_0, a_1]$. The solution we denote by x_n^k. Starting with the new initial value $x_n^k(a_1)$ we can continue this process, yielding x_n^k and u_n^k. But we still could not prove that $x_n^k \to x^k$, $u_n^k \to u^k$ as $n \to \infty$ for some suitable notion of limit.

4. Estimation of the cost functional

Checking up the behaviour of the cost functional $J(u)$ we can formulate the following theorem.

Theorem 1 : Let the hypotheses (H1) - (H4) be fulfilled for $x(t) \in X \subset \mathbb{R}^n$, $u(t) \in U \subset \mathbb{R}^m$, $\lambda(t) \in \Lambda \subset \mathbb{R}^n$ and $t \in T$. Let $(\varepsilon^k)_{k \geq 1}$ be an increasing sequence of positive numbers. Then, with a constant $C > 0$,

$$J\big(u^k\big) - J\big(u^{k-1}\big) \leq -\big(|T|/\varepsilon^k - C\big)\| u^k - u^{k-1}\|_{L}. \tag{5}$$

Proof : By the definition of the cost functional we find

$$J(u^k) - J(u^{k-1})$$

$$= \int_T \Big[L\big(x^k, u^k, t\big) - L\big(x^{k-1}, u^{k-1}, t\big)\Big]dt + g\big(x^k(t_1)\big) - g\big(x^{k-1}(t_1)\big)$$

$$= \int_T \Big[K_{\varepsilon k}(x^k, u^k, v, \lambda^{k-1}, t) - K_{\varepsilon k-1}(x^{k-1}, u^{k-1}, v, \lambda^{k-1}, t)$$

$$+ \big(\lambda^{k-1} \cdot f(x^{k-1}, u^{k-1}, t) - f(x^k, u^k, t)\big)$$

$$+ \int_T \Big[\tfrac{1}{\varepsilon k-1}\| u^{k-1} - v\| - \tfrac{1}{\varepsilon k}\| u^{k-1} - v\|\Big]ds\Big]dt$$

$$+ g\big(x^k(t_1)\big) - g\big(x^{k-1}(t_1)\big).$$

Now we estimate the term $g(x^k(t_1)) - g(x^{k-1}(t_1))$ by means of a Lipschitzian, the inner product $\big(\lambda^{k-1} \cdot f(x^{k-1}, u^{k-1}, t) - f(x^k, u^k, t)\big)$ by means of the inequality of Schwarz and a Lipschitzian, too. We get

$$J(u^k) - J(u^{k-1}) \le C_3 \|x^k - x^{k-1}\|$$
$$+ \int_T \Big[K_{\varepsilon k}(x^k, u^k, v, \lambda^{k-1}, t) - K_{\varepsilon k-1}(x^{k-1}, u^{k-1}, v, \lambda^{k-1}, t)$$
$$+ C_1 \|x^k - x^{k-1}\| + C_2 \|u^k - u^{k-1}\|$$
$$+ \int_T \Big[\tfrac{1}{\varepsilon k-1} \|u^{k-1} - v\| - \tfrac{1}{\varepsilon k} \|u^k - v\| \Big] ds \Big] dt .$$

Further we can estimate the first integrand:

$$K_{\varepsilon k}(\ldots) - K_{\varepsilon k-1}(\ldots)$$
$$= K_{\varepsilon k}(x^k, u^k, v, \lambda^{k-1}, t) - K_{\varepsilon k}(x^k, u^{k-1}, v, \lambda^{k-1}, t)$$
$$+ K_{\varepsilon k}(x^k, u^{k-1}, v, \lambda^{k-1}, t) - K_{\varepsilon k-1}(x^k, u^{k-1}, v, \lambda^{k-1}, t)$$
$$+ K_{\varepsilon k-1}(x^k, u^{k-1}, v, \lambda^{k-1}, t) - K_{\varepsilon k-1}(x^{k-1}, u^{k-1}, v, \lambda^{k-1}, t).$$

We use the mean value theorem, the inequality of Schwarz and assumption (H1). We concentrate:

$$K_{\varepsilon k}(\ldots) - K_{\varepsilon k-1}(\ldots)$$
$$\le C_4 \|x^k - x^{k-1}\| + C_5 \|u^k - u^{k-1}\| + \Big(\tfrac{1}{\varepsilon k} - \tfrac{1}{\varepsilon k-1} \Big) \int_T \|u^k - v\| \, dt .$$

Since by the assumption on $(\varepsilon^k)_{k \le 1}$ the last term is non-positive we obtain

$$K_{\varepsilon k}(\ldots) - K_{\varepsilon k-1}(\ldots) \le C_4 \|x^k - x^{k-1}\| + C_5 \|u^k - u^{k-1}\| .$$

Summarizing we get

$$J(u^k) - J(u^{k-1})$$
$$\le \int_T \Big(C_6 \|x^k - x^{k-1}\| + C_7 \|u^k - u^{k-1}\| \Big) dt$$
$$+ \int_T \Big[\int_T \Big[\tfrac{1}{\varepsilon k-1} \|u^{k-1} - v\| - \tfrac{1}{\varepsilon k} \|u^k - v\| \Big] ds \Big] dt + C_3 \|x^k - x^{k-1}\|.$$

If we apply the algorithm at step k we can write $u^{k-1} =: v$ and the inner integratio yields $- \tfrac{1}{\varepsilon k} \|u^k - u^{k-1}\|_{L_1}$. In order to estimate $C_3 \|x^k - x^{k-1}\|$ we use the Lemma o Gronwall and thus we get $C_3 \|x^k - x^{k-1}\| \le \tilde{C} \int_T \|u^k - u^{k-1}\| \, dt$. Now we have

$$J(u^k) - J(u^{k-1}) \le \int_T \Big(C_6 \|x^k - x^{k-1}\| + C_8 \|u^k - u^{k-1}\| \Big) dt + \int_T (-\tfrac{1}{\varepsilon k}) \|u^k - v\|_{L_1} dt .$$

From the Lemma of Gronwall we can conclude

$$\int_T C_6 \|x^k - x^{k-1}\| \, dt \le C_9 \int_T \|u^k - u^{k-1}\| \, dt$$

and therewith

$$J(u^k) - J(u^{k-1}) \le \int_T C \|u^k - u^{k-1}\| \, dt - \int_T \tfrac{1}{\varepsilon k} \|u^k - u^{k-1}\|_{L_1} \, dt$$
$$= - \Big(\tfrac{|T|}{\varepsilon k} - C \Big) \|u^k - u^{k-1}\|_{L_1} .$$

Thus the theorem is proved ∎

27*

5. Results on convergence

At first we prove a lemma which we want to use in the proof of the following theorem.

Lemma 3 : *Let* (z^k) *be a sequence in a normed space* X, $(y^k) \subset \mathbf{R}$ *and* $(\alpha^k) \subset \mathbf{R}$ *decreasing,* $\alpha_k > 0$ *for each* $k \in \mathbf{N}$. *Assume further that*

$$y^k - y^{k-1} \le -\alpha_k \|z^k - z^{k-1}\| \text{ for all } k . \tag{6}$$

Then

$$y^k - y^i \le -\alpha_k \|z^k - z^i\| \text{ for } i \le k . \tag{7}$$

Proof : We give the proof with the help of a two-dimensional version of the induction principle as e.g. announced in [3]. Let $M \subset \mathbf{N} \times \mathbf{N}$ have the following properties:

(i) $(1,1) \in M$;

(ii) $\{(i,k)| i \le k \le n\} \subset M \implies \{(i, n+1)| i \le n+1\} \subset M \ (n \in \mathbf{N})$.

Then $\{(i,k)| i \le k\} \subset M$.

Obviously (7) holds if $i = k$. Now, we suppose that (7) is satisfied for $i \le k \le n$ and prove that (7) holds for $(j, n+1)$, $j \le n+1$, too. For $j = n+1$ the inequality holds. By assumption, for arbitrary j, $j \le n$, we have $y^n - y^j \le -\alpha_n \|z^n - z^j\|$. Moreover by (6) $y^{n+1} - y^n \le -\alpha_{n+1}\|z^{n+1} - z^n\|$. Adding the two inequalities we get

$$y^{n+1} - y^j \le -\alpha_{n+1}\left(\|z^{n+1} - z^n\| + \|z^n - z^j\|\right) .$$

Now $\|z^{n+1} - z^j\| \le \|z^{n+1} - z^n\| + \|z^n - z^j\|$ implies that

$$-\alpha_{n+1}\|z^{n+1} - z^j\| \ge -\alpha_{n+1}\left(\|z^{n+1} - z^n\| + \|z^n - z^j\|\right)$$

and thus showing that $y^{n+1} - y^j \le -\alpha_{n+1}\|z^{n+1} - z^j\|$ holds ∎

Theorem 2 : *Let* $(\varepsilon^k)_{k \ge 1}$ *be an increasing sequence of positive numbers with* $\varepsilon^k \le |T|/(\alpha_0 + C)$, *where* $\alpha_0 > 0$, $|T|$ *is the length of the time intervall and* $C > 0$ *is taken from inequality* (5). *Let* $(u^k)_{k \ge 1}$ *be any sequence of control functions generated by the algorithm and corresponding to the sequence* $(\varepsilon^k)_{k \ge 1}$. *We further assume that there exists an optimal control* $u^* \in U_{ad}$. *Then the following holds:*

1. $(J(u^k))$ *is a decreasing convergent sequence.*

2. *We find some* $\hat{u} \in U_{ad}$ *such that* $\|u - \hat{u}\|_{L_1} \to 0$.

3. *There exists a subsequence* (u^{kj}) *of* (u^k) *such that* $u^{kj} \to \hat{u}$ *a.e. on* T.

4. *If in addition* L_{xx} *and* f_{xu} (*or* L_{xu} *and* f_{xx}) *exist, then* $\hat{u}$ *fulfils the necessary (first order) condition of optimality and* $J(u^k) \to J(\hat{u})$.

Proof: 1. We have $\varepsilon^k \le |T|/(\alpha_0 + C)$ and hence $\varepsilon^k < |T|/C$ for all k, too. Then $|T|/\varepsilon^k - C > 0$ and hence by (5) $(J(u^k))$ decreases. If u^* is an optimal control, then $J(u^*) \le J(u^k)$ for all k and $(J(u^k))$ converges.

2. For $k \ge 1$ we set $\alpha_k = |T|/\varepsilon^k - C$, $y^k = J(u^k)$ and $z^k = u^k$. Then, applying Lemma 3 to the inequality (5), we find that

$$J(u^k) - J(u^i) \le -(|T|/\varepsilon^k - C)\|u^k - u^i\|_{L_1} \text{ for } i \le k . \tag{8}$$

The convergence of $(J(u^k))$ implies that this sequence is a Cauchy sequence, too. From

(8) we get $J(u^i) - J(u^k) \geq \alpha_k \|u^k - u^i\|_{L_1}$ for all $i \leq k$. Moreover, $\varepsilon^k \leq |T|/(\alpha_0 + C)$ implies $\alpha_k \geq \alpha_0$. These facts show that (u^k) is a Cauchy sequence with respect to the L_1-norm. Since $L_1(T, \mathbf{R}^m)$ is complete we find $\hat{u} \in L_1(T, \mathbf{R}^m)$ such that $\|u^k - \hat{u}\|_{L_1} \to 0$. By Lemma 1. $\hat{u} \in U_{ad}$.

3. Assertion 3 we obtain at once from Assertion 2 (see the proof of Lemma 1).

4. Let $\hat{x}$ and $\hat{\lambda}$ be the solutions of system (1) and (3) which correspond to the control vector $\hat{u}$. We show at first that $\|\hat{x} - x^k\|_{L_1} \to 0$ and $x^k(t) \to \hat{x}(t)$ a.e. on T. By our assumption we find constants D_1, D_2 such that

$$\frac{d}{dt}\|\hat{x}(t) - x^k(t)\| \leq D_1\|\hat{x}(t) - x^k(t)\| + D_2\|\hat{u}(t) - u^k(t)\|.$$

Integrating and applying the Lemma of Gronwall give $\|\hat{x}(t) - x^k(t)\| \leq D_3\|\hat{u} - u^k\|_{L_1}$. Integrating once more we get $\|\hat{x} - x^k\|_{L_1} \leq D_3|T|\|\hat{u} - u^k\|_{L_1}$. By statement 2 we get $\|u^k - \hat{u}\|_{L_1} \to 0$ and hence $\|\hat{x}(t) - x^k(t)\| \to 0$ a.e. on T and $\|\hat{x} - x^k\|_{L_1} \to 0$. Now we show that $\lambda^k(t) \to \hat{\lambda}(t)$ a.e. on T. Applying the additional assumption we get

$$d\|\hat{\lambda}(t) - \lambda^k(t)\|/dt$$

$$\leq \left\|\frac{d}{dt}\left(\hat{\lambda}(t) - \lambda^k(t)\right)\right\|$$

$$= \left\| - H_x(\hat{x}, \hat{u}, \hat{\lambda}, t) + H_x(x^k, u^k, \lambda^k, t)\right\|$$

$$= \left\| L_x(x^k, u^k, t) + \lambda^{kT}f_x(x^k, u^k, t) - L_x(\hat{x}, \hat{u}, t) - \hat{\lambda}^T f_x(\hat{x}, \hat{u}, t)\right\|$$

$$\leq D_4\|x^k - \hat{x}\| + \left\|\lambda^{kT}\left(f_x(x^k, u^k, t) - f_x(\hat{x}, \hat{u}, t)\right)\right\| + \left\|(\lambda^{kT} - \hat{\lambda}^T)f_x(\hat{x}, \hat{u}, t)\right\|$$

$$\leq D_4\|x^k - \hat{x}\| + D_5\|u^k - \hat{u}\| + D_6\|\lambda^k - \hat{\lambda}\|.$$

Let be $\tau = t_1 - t$ and we consider

$$d\|\hat{\lambda}(t) - \lambda^k(t)\|/d\tau \leq \|d(\ldots)/d\tau\| = \|d(\ldots)\|/dt$$

$$\leq D_4\|x^k - \hat{x}\| + D_5\|u^k - \hat{u}\| + D_6\|\lambda^k - \hat{\lambda}\|.$$

We integrate from 0 to τ and get for the left term

$$\|\hat{\lambda}(t) - \lambda^k(t)\| - \|\hat{\lambda}(t_1) - \lambda^k(t_1)\| = \|\hat{\lambda}(t) - \lambda^k(t)\| - \|g(\hat{x}(t)) - g(x^k(t_1))\|.$$

Hence we get

$$\|\hat{\lambda}(t) - \lambda^k(t)\| \leq \int_0^T D_4\|x^k(t_1 - \tau) - \hat{x}(t_1 - \tau)\|d\tau + \int_0^T D_5\|u^k(t_1 - \tau) - \hat{u}(t_1 - \tau)\|d\tau$$

$$+ \int_0^T D_6\|\lambda^k(t_1 - \tau) - \hat{\lambda}(t_1 - \tau)\|d\tau + \|g(\hat{x}(t_1)) - g(x^k(t_1))\|.$$

Application of the Lemma of Gronwall yields

$$\|\hat{\lambda}(t) - \lambda^k(t)\| \leq \left[\int_0^T D_4\|x^k(t) - \hat{x}(t)\|dt + \int_0^T D_5\|u^k(t) - \hat{u}(t)\|dt\right.$$

$$\left. + \|g(\hat{x}(t_1)) - g(x^k(t_1))\|\right]e^{\int_0^T D_6 d\tau} \leq D_7\left[\ldots\right]$$

and hence $\|\hat{\lambda}(t) - \lambda^k(t)\| \to 0$, $t \in T$. Since $u = u^k(t) \in U$ minimizes $K_{\varepsilon k}(x^k, u, u^{k-1}, \lambda^{k-1}, t)$

on the convex set U we get $\left(H_u(x^k, u^k, \lambda^{k-1}, t), u^k - u\right) \le 0$. From Assertion 3 we have a subsequence (u^{kj}) of (u^k) such that $u^{kj} \to u$ a.e. on T. Further on we have $x^{kj} \to \hat{x}$ on T. $\lambda^{kj} \to \hat{\lambda}$ on T. We obtain $\left(H_u(\hat{x}, \hat{u}, \hat{\lambda}, t), \hat{u} - u\right) \le 0$ since H_u is continuous on $X \times U \times \Lambda \times T$. By assumption (H4) the Hamiltonian is convex and hence by a well-known theorem we get $H(\hat{x}, \hat{u}, \hat{\lambda}, t) = \min \{ H(\hat{x}, u, \hat{\lambda}, t) | u \in U \}$. Moreover we find constants D_8, D_9 such that

$$\left| L(x^k, u^k, t) - L(\hat{x}, \hat{u}, t) \right| \le D_8 \| x^k - \hat{x} \| + D_9 \| u^k - \hat{u} \|$$

which yields $\left| J(u^k) - J(\hat{u}) \right| \le D_8 \| x^k - \hat{x} \|_{L_1} + D_9 \| u^k - \hat{u} \|_{L_1} \to 0$ as $k \to \infty$ ∎

LITERATUR

[1] BONNANS, J.F.: *On an algorithm for optimal control using Pontryagins Maximum Principle.* SIAM J. Control Opt. **24**(1986), 579 - 588.

[2] FATTORINI, H.O.: *A Unified Theory of Necessary Condition for Nonlinear Noncon vex Control Systems.* Appl. Math. Opt. **15**(1987), 141 - 185.

[3] POPPE, H., and A. STÄRK: *A Two-Dimensional Induction Lemma with Applications to the Proof of Combinatorial Formulas.* Math. Sem Notes, Kobe University **8**(1980), 219 - 223.

[4] SAKAWA, Y., and Y. SHINDO: *On global convergence of an algorithm of optimal con- trol.* IEEE Transl. and auto. control **AC-25**(1980), 1149 - 1153.

[5] ZEIDLER, E.: *Vorlesungen über nichtlineare Funktionalanalysis.* Teil II: *Monotone Operatoren* (Teubner Texte zur Mathematik: Bd. 9). Leipzig:B.G. Teubner Verlags- ges. 1977.

Author's address:

Dipl. Math. Ute Flemmig and Prof. Dr. Harry Poppe
Hochschule für Seefahrt Warnemünde/Wustrow
Institut für Mathematik
D(Ost) 2598 Wustrow