ON THE CONVERGENCE OF DIFFERENCE APPROXIMATIONS IN DISTRIBUTED PARAMETER OPTIMAL CONTROL PROBLEMS

HITOSHI SASAI

Department of Aeronautical Engineering

(Received November 12, 1971)

Abstract

Since optimal control problems for distributed parameter systems are intimately connected to optimal control problems described by partial differential equations and require a treatment in function spaces, there is a great difficulty to solve these problems directly. Consequently it is considered to be necessary to obtain approximating solutions by nummerical computation.

In computing a numerical solution of a partial differential equation, an approximating method in which differential operator is replaced by a difference operator is widely used.

A similar approximating procedure is commonly used in calculating an optimal control in systems with distributed parameters.

Convergency conditions of a sequence of approximating solutions for the optimal control problems of distributed parameter systems are given in this paper.

Introduction

Optimal control problems for distributed parameter systems, particularly systems described by partial differential equations, are often treated using mathematical function space techniques. As a result, there is a great difficulty to solve these problems directly. Numerical solutions are obtained by approximating the abstract operations in a computationally feasible manner. In obtaining approximating solutions, the method of expanding in eigenfunctions[1], and the finite difference method are widely used.

After having found an approximate optimal control, the question arises whether a sequence of these approximating optimal controls converges to an optimal control of the original system. A condition of convergence of a sequence of approximating solutions of initial value problems by the finite difference method in which a differential operator is replaced by a difference operator was given by H. F. Trotter[2][3]. This theorem is concerned with a homogeneous system and does not give the condition of convergence of approximating optimal controls for a distributed parameter system, but its result becomes very useful for our discussion as shown below.

The condition of convergence of a sequence of approximating solutions has been given for time-optimal control problem for a class of linear systems with distributed parameters by the author[4].

In this paper the condition of convergence of a sequence of approximating solutions is given for the optimal control problems with more general cost functions.

To discuss the convergence of approximating solutions, it is important to define a concept of metric. In the following, the space is considered to be a Hilbert space.

Definitions and Basic Theorem

For convenience, the following notations, definitions and basic theorem are introduced. Let X be a Hilbert space and X_n be a subspace of X. We consider a linear mapping P_n: $X \to X_n$ which satisfies the following conditions:

$$P_n^2 = P_n, \quad \|P_n\|_X = 1,$$
$$\lim \|P_n f - f\|_X = 0 \quad \text{for every } f \in X.$$

We denote the norm in X by $\| \|_X$ and the norm in X_n by $\| \|_{X_n}$ respectively.

Definition. The limit of a sequence of operators $\{A_n\}$, where A_n is an operator on X_n, is the operator on X whose domain consists of $f \in X$, for which $\{A_n P_n f\}$ converges, and whose value is lim $A_n P_n f$.

Let $S_n(t)$ be a sequence of strongly continuous semigroups of operators on X_n into X_n, and A_n be the sequence of associated infinitesimal generators. Then we have the following theorem, which is due to H. F. Trotter.

Theorem (*Trotter*[2]). If the range of $\lambda I - A$, denoted by $R(\lambda I - A)$, is dense in X for some $\lambda > K$ and the following conditions are satisfied, where M and K are positive constants:

$$(C) \quad A = \lim A_n \text{ and } D(A) \text{ is dense in } X,$$
$$(S) \quad \|S_n(t)\|_{X_n} \leqq M e^{Kt},$$

then a closed extension of A is the infinitesimal generator of $S(t)$, where $S(t) = \lim S_n(t)$ is a strongly continuous semigroup of operators on X.

In this theorem it is demonstrated that if conditions (C) and (S) are satisfied, approximating solutions of approximating equations starting from given appromimating initial sates converge in the sense of norm to a solution of the original homogeneous partial differential equation, which satisfies a given initial condition.

Statement of the Prbolem

Let us consider the following distributed parameter system:

$$\partial x(t, s)/\partial t = Ax(t, s) + u(t, s), \tag{0}$$

where for each t, $0 \leqq t < \infty$, $x(t, s)$ is an element of X, which is an arbitrary Hilbert space $X(\Omega)$, $s \in \Omega$, consisting of locally summable function on a bounded domain Ω; A is an unbounded linear operator from X to X, for instance A being a partial differential operator on Ω; and $u(t, s)$ is an element of X for each t and square summable in t. We denote $x(t, s)$ and $u(t, s)$ as $x(t)$ and $u(t)$, respectively. Here A is assumed to be an infinitesimal generator of the strongly continuous semigroup $S(t)$. Then (0) becomes

$$\dot{x}(t) = Ax(t) + u(t). \tag{1}$$

Along with the system (1), the following equation is considered:

$$\dot{x}_n(t) = A_n x_n(t) + u_n(t), \tag{2}$$

where $x_n(t)$ is an element of X_n for each t, $0 \leqq t < \infty$, $u_n(t)$ is an element of X_n for each t and square summable in t and A_n is the operator from X_n into X_n which is an infinitesimal generator of a strongly continuous semigroup $S_n(t)$.

In the case of A_n being the difference operator on X_n into X_n, (1) is the original system which we consider and (2) is the approximating system given by the finite difference method. In what follows A_n is assumed to be a bounded linear operator.

As an example of the mapping P_n see[2].

The solution of (1) is formally written as

$$x(t) = S(t)x(0) + \int_0^t S(t-\sigma)u(\sigma)d\sigma. \tag{3}$$

Since $u(t)$ is bounded and square integrable the integral on the right-hand side of (3) exists in the sense of Bochner. Sufficient conditions under which (3) represents the strong solution to (1) (differentiable in the strong sense) have been given by Balakrishnan[5]. But since the right-hand side of (3) exists in the sense of Bochner, we define (3) as the solution of (1) for any square inegrable function $u(t)$ and a given initial condition $x(0)$; *i.e.*, we deal with the mild solution[6] in this paper.

For simplicity's sake, one can assume without loss of generality that $x(0) = 0$ in the sequel. Through the following discussion, we assume that the conditions (C) and (S) are satisfied.

Let $B_2(X, T)$ be the space of strongly measurable functions $x(t)$ with range in X such that $\int_0^T \|x(t)\|^2 dt < \infty$.

Now, the solution $x(t)$ of (1) for any $u(t) \in B_2(X, T)$ is an element of the space $B_2(X, T)$, since $S(t)$ is strongly continuous. P_n being bounded operators, it follows that $P_n x(t) \to x(t)$ in $B_2(X, T)$ for any $x(t) \in B_2(X, T)$.

We put

$$(Lu)(t) = \int_0^T S(t-\sigma)u(\sigma)d\sigma, \quad u(t) \in B_2(X, T) \tag{4}$$

$$(L_n u_n)(t) = \int_0^T S_n(t-\sigma)u_n(\sigma)d\sigma, \quad u_n(t) \in B_2(X_n, T) \tag{5}$$

Under the hypothesis of Trotter's theorem it is shown that $S_n(t)P_n x \to S(t)x$ for any $x \in X$, and again by the condition (S) it follows that

$$(L_n P_n u)(t) \to (Lu)(t) \tag{6}$$

in $B_2(X, T)$ for any $u(t) \in B_2(X, T)$.

We consider a nonlinear function $f(x(t), u(t))$: $B_2(X, T) \times B_2(X, T) \to R_1$ and a sequence of functions $f_n(x_n(t), u_n(t))$: $B_2(X_n, T) \times B_2(X_n, T) \to R_1$ approximating $f(x(t), u(t))$ which has the following properties;

(H_1) If $x_n(t)$, $u_n(t) \in B_2(X_n, T)$ converges to $x(t)$, $u(t)$ in $B_2(X, T)$ respectively, $f_n(x_n(t), u_n(t))$ converges to $f(x(t), u(t))$.

(H_2) If $x_n(t)$, $u_n(t) \in B_2(X_n, T)$ converge weakly to $x(t)$, $u(t)$ in $B_2(X, T)$,

$$f(x(t), u(t) \leqq \underline{\lim}\, f_n(x_n(t), u_n(t)).$$

(H_3) If the norm of $x_n(t)$, $u_n(t) \in B_2(X_n, T)$ go to infinite respectively, $f_n(x_n(t), u_n(t))$ goes to infinite.

Remark. (H_1), (H_2) and (H_3) mean that the sequence of f_n is approximating f, f is weakly lower semicontinuous and radially unbounded.

In this article we discuss on convergence of approximating solutions of the following optimal control problem.

Problem

(A): The original optimal control problem in system (1) is to find the admissible control minimizing $f(x(t), u(t))$ starting from a given initial state $x(0) \in X$.

(B): The approximating optimal control problem in system (2) is to find the admissible control minimizing $f_n(x_n(t), u_n(t))$ starting from a given approximating initial state $P_n x(0)$, with $n \to \infty$.

Then our problem is reduced to, "Does a sequence of approximating optimal controls of the problem (B) converge to an optimal control of the problem (A)?".

Remark. In this article the set of admissible controls are defined such that $B_2(X, T)$ is the admissible control set for the exact problem and $B_2(X_n, T)$ is the admissible control set for the approximate problem.

Condition of Convergence

Let $u^\circ(t)$ be the optimal control of the problem (A) and $u_n^\circ(t)$ be the optimal control of the problem (B) at the n-th degree of approximation.

Now we make the assumption that the approximating optimal solutions $u_n^\circ(t)$ exist and the conditions (C), (S) and

$$(C^*): \quad \lim \|S_n^*(t) P_n f - S^*(t) f\|_X = 0, \quad \text{for fixed } t \text{ and } f \in X,$$

are satisfied, where S^* and S_n^* are the adjoint operators of S and S_n considered on the space X and X_n respectively.

Since $\|S_n\|_{X_n} = \|S_n^*\|_{X_n}$ and $\|S\|_X = \|S^*\|_X$, it can be obtained that

$$(L_n^* P_n f)(t) \to (L^* f)(t) \text{ in } B_2(X, T), \text{ where } f(t) \in B_2(X, T). \tag{7}$$

L^* and L_n^* express the adjoint operator of L and L_n considered on the space $B_2(X, T)$ and $B_2(X_n, T)$, respectively.

Now we denote the inner products of the space X, X_n and the space $B_2(X, T)$, $B_2(X_n, T)$ by $[\ ,\ ]_X$, $[\ ,\ ]_{X_n}$ and $\langle\ \rangle_X$, $\langle\ \rangle_{X_n}$ respectively. Hence the inner products of the space $B_2(X, T)$ and $B_2(X_n, T)$ are defined as follws,

$$\langle\ ,\ \rangle_X=\int_0^T[\ ,\ ]_X\,dt$$

$$\langle\ ,\ \rangle X_n=\int_0^T[\ ,\ ]_{X_n}dt.$$

Since $\{X-X_n\}\perp X_n$ and P_n is a projection operator, for $f(t)\in B_2(X, T)$, we have the following equations:

$$[(L_n u_n^\circ)(t), f(t)]_X=[(L_n u_n^\circ)(t), P_n f(t)]_X$$
$$=[(L_n u_n^\circ)(t), P_n f(t)]_{X_n}, \quad \text{for fixed } t,$$
$$[(u_n^\circ(\sigma), S_n^*(t-\sigma)(P_n f)(t)]_X=[u_n^\circ(\sigma), S_n^*(t-\sigma)P_n f(t)]_X,$$
$$\text{for fixed } t,\ \sigma.$$

Therefore we get

$$\langle(L_n u_n^\circ)(t), f(t)\rangle_X=\langle(L_n u_n^\circ)(t), P_n f(t)\rangle_{X_n}, \tag{8}$$
$$\langle u_n^\circ(t), (L_n^* P_n f)(t)\rangle_{X_n}=\langle u_n^\circ(t), (L_n^* P_n f)(t)\rangle_X. \tag{9}$$

Next we shall show the following Lemma:

Lemma 1. Let the conditions (C), (S), (C^*) and the properties (H_1), (H_2), (H_3) be satisfied. Then there exists a weakly convergent subsequence $\{u_{nj}^\circ(t)\}$ of $\{u_n^\circ(t)\}$ in $B_2(X, T)$. If wd denote its weak limit by $v(t)$, $(L_{nj}u_{nj}^\circ)(t)$ converges weakly to $(Lv)(t)$, too. Moreovere $v(t)$ is an optimal control of the problem (A).

Proof. $u_n^\circ(t)$ being optimal it is clear that the following inequality is satisfied for any $u(t)\in B_2(X, T)$:

$$f_n((L_n u_n^\circ)(t), u_n^\circ(t))\leqq f_n((L_n P_n u)(t), P_n u(t)).$$

Since $f_n((L_n P_n u)(t), P_n u(t))\to f((Lu)(t), u(t))$ by (H_1), $(L_n u_n^\circ)(t)$ and $u_n^\circ(t)$ become to be bounded in $B_2(X, T)$ by (H_3). Therefore there exists a subsequence $\{u_{nj}^\circ(t)\}$ of $\{u_n^\circ(t)\}$ converging weakly to $v(t)\in B_2(X, T)$, because $B_2(X, T)$ is a Hilbert space.

Now it can be shown that $(L_{nj}u_{nj}^\circ)(t)$ converges weakly to $(Lv)(t)$. In what follows we write n for n_j. By the equation (8) and (9), we have

$$\begin{aligned}
&\langle(L_n u_n^\circ)(t)(t)-(Lv)(t), f(t)\rangle_X\\
&\quad=\langle(L_n u_n^\circ)(t), f(t)\rangle_X-\langle(Lu_n^\circ)(t), f(t)\rangle_X+\langle(Lu_n^\circ)(t)-(Lv)(t), f(t)\rangle_X\\
&\quad=\langle(L_n u_n^\circ)(t), P_n f(t)\rangle_{X_n}-\langle(Lu_n^\circ)(t), f(t)\rangle_X+\langle(L(u_n^\circ-v))(t), f(t)\rangle_X\\
&\quad=\langle u_n^\circ(t), (L_n^* P_n f)(t)\rangle_{X_n}-\langle u_n^\circ(t), (L^* f)(t)\rangle_X+\langle u_n^\circ(t)-v(t), (L^* f)(t)\rangle_X\\
&\quad=\langle u_n^\circ(t), (L_n^* P_n f)(t)-(L^* f)(t)\rangle_X+\langle u_n^\circ(t)-v(t), (L^* f)(t)\rangle_X.
\end{aligned}$$

The first term of the right-hand side goes to zero as $n\to\infty$ by the conditions (C^*) and $u_n^\circ(t)$ being bounded. The second term goes to zero as $n\to\infty$, too because $u_n^\circ(t)$ converges weakly to $v(t)$.

Next we shall show that $v(t)$ is an optimal control of problem (A).

By the optimality of $u_n^\circ(t)$, the above results and (H_2), the following inequality can be proved:

$$f((Lv)(t),\ v(t)) \leqq \underline{\lim}\ f_n((L_n u_n^\circ)(t),\ u_n^\circ(t))$$
$$\leqq \lim f_n((L_n P_n u)(t),\ P_n u(t))$$
$$= f((Lu)(t),\ u(t)).$$

Since the above inequality is satisfied for any $u(t) \in B_2(X, T)$, it can be concluded that $v(t)$ is optimal.

Theorem 1. Let the conditions (C), (S), (C^*) and the properties (H_1), (H_2), (H_3) are satisfied. Suppose that the optimal controls $u_n^\circ(t)$ of problem (B) exist for each n. Then there exists an optimal control $u^\circ(t)$ of the problem (A) and the optimal value of the cost function of the problem (B) converges to the optimal value of the cost function of the problem (A) as $n \to \infty$. (*i.e.* $f_n((L_n u_n^\circ)(t),\ u_n^\circ(t)) \to f((Lu^\circ)(t),\ u^\circ(t))$)

Proof. The first part of the theorem is clear from the above Lemma. Any subsequence of $f_n((L_n u_n^\circ)(t),\ u_n^\circ(t))$ has its subsequence converging to $f((Lu^\circ)(t),\ u^\circ(t))$, and therefore $f_n((L_n u_n^\circ)(t),\ u_n^\circ(t))$ converges to $f((Lu^\circ)(t),\ u^\circ(t))$ as $n \to \infty$.

Remark. In order to obtain the condition (C^*), we again make use of the Trotter's theorem. Since condition (S) is always satisfied for $S_n^*(t)$, it is sufficient to check the validity of condition (C) for A_n^*.

In the sequel we confine our discussion to the following type of the cost function:

$$f(x(t),\ u(t)) = \int_0^T \{\|x(t) - y(t)\|_X^2 + \lambda \|u(t)\|_X^2\}\, dt, \tag{10}$$

$$f_n(x_n(t),\ u_n(t)) = \int_0^T \{\|x_n(t) - P_n y(t)\|_{X_n}^2 + \lambda \|u_n(t)\|_{X_n}^2\}\, dt, \tag{11}$$

where $y(t) \in B_2(X, T)$ is a given function and $\lambda > 0$.

An optimal control problem minimizing the cost function like the above (10) is called to be *the tracking problem.*

In this case it is clear that the properties (H_1), (H_2) and (H_3) are satisfied because $\|x_n\|_{X_n} = \|x_n\|_X$, for $x_n \in X_n$ and the resonance theorem[7] means that if $x_n(t) \in B_2(X, T)$ converges weakly to $x(t)$ in the space $B_2(X, T)$, $\underline{\lim} \langle x_n(t), x_n(t)\rangle_X \geqq \langle X(i), x(i)\rangle_X$.

We denote the optimal controls of problem (A) and problem (B) associated with (11) and (12) by $u^\circ(t)$ and $u_n^\circ(t)$ respectively.

Theorem 2. For the tracking problem, $u_n^\circ(t)$ converges $u^\circ(t)$ in $B_2(X, T)$, *i.e.*, $u_n^\circ(t)$ converges $u^\circ(t)$ for almast all t.

Proof. By Lemma 1, there exists a subsequence $\{u_{nj}^\circ(t)\}$ of $\{u_n^\circ(t)\}$ converging weakly to an optimal control $v(t)$ of the problem (A). One recall that $u_n^\circ(t)$ and $u^\circ(t)$ are unique[5].

Since $u^\circ(t)$ is unique, $u^\circ(t) = v(t)$ for a. a. t. By Theorem 1, $f_{nj}(x_{nj}^\circ(t),$

$u^{\circ}_{nj}(t)) \to f(x^{\circ}(t), u^{\circ}(t))$, where $x^{\circ}_{nj}(t)$ and $x^{\circ}(t)$ is the optimal trajectories for $u^{\circ}_{nj}(t)$ and $u^{\circ}(t)$ respectively (*i.e.*, $x^{\circ}_{nj}(t) = (L_{nj}u^{\circ}_{nj})(t)$, $x^{\circ}(t) = (Lu^{\circ})(t)$).

Now it can be shown (see Appendix);

$$\langle u^{\circ}_{nj}(t), u^{\circ}_{nj}(t)\rangle_X \to \langle u^{\circ}(t), u^{\circ}(t)\rangle_X.$$

Therefore we can conclude $u^{\circ}_{nj}(t)$ converges strongly $u^{\circ}(t)$ in $B_2(X, T)$, because $u^{\circ}_{nj}(t)$ converges weakly to $u^{\circ}(t)$, the norm of $u^{\circ}_{nj}(t)$ converges the norm of $u^{\circ}(t)$ and $B_2(X, T)$ is a Hilbert space.

Any subsequence of $u^{\circ}_{n}(t)$ has its subsequence converging weakly to $u^{\circ}(t)$, and therefore it has the subsequence converging strongly to $u^{\circ}(t)$ in $B_2(X, T)$ by the above facts.

Consequently $u^{\circ}_{n}(t)$ converges to $u^{\circ}(t)$ in $B_2(X, T)$.

Example

As an example we consider the tracking problem for a one-dimensional diffusion equation.

The system equation is given by

$$\begin{cases} \partial x(t, s)/\partial t = \partial^2 x(t, s)/\partial t^2 + u(t, s), & \text{for } 0 < t < \infty \text{ and } 0 < s < 1, \\ u(0, s) = 0, \\ u(t, 0) = u(t, 1) = 0. \end{cases} \tag{12}$$

Let $X = L_2(0, 1)$ and define

$$D(A) = \{f(s) \in X: \; f(0) = f(1) = 0,\ f(s) \text{ and } f'(s) \text{ are absolutely continuous, } f''(s) \in X\},$$
$$Af = f''(s) \text{ for } f(s) \in D(A).$$

Then the system (13) is phrased in following abstract notations of this paper (see, *e.g.*[8]).

$$dx(t)/dt = Ax(t) + u(t), \quad x(0) = 0 \in D(A) \tag{13}$$

It is well known that A is a infinitesimal generator of a strongly continuous semigroup[8].

Now let us construct a discretizations cheme. As an example of the mapping P_n consider the following:

$$X = L_2(0, 1), \quad x(s) \in X, \quad s \in [0, 1],$$
$$P_n x(s) = \sum_1^n \alpha_k c_k(s),$$

where

$$\alpha_k = n \int_{k/n - 1/k}^{k/n} x(s)\, ds,$$

$$C_k(s) = \begin{cases} 1, & k/n - 1/n \leqq s \leqq k/n, \\ 0, & k/n - 1/n \geqq s \text{ or } s \geqq k/n. \end{cases}$$

We put $1/n=\delta_n$ and

$$\begin{cases} P_n x(t,s) = \sum_1^n x_k(t)\, c_k(s) \\ P_n u(t,s) = \sum_1^n u_k(t) c_k(s) \\ P_n y(t,s) = \sum_1^n y_k(t) c_k(s). \end{cases} \tag{14}$$

By central difference formulas, we have

$$\partial^2 x(t,s)/\partial s^2 \cong \sum_1^n (x_{k+1}(t) - 2\,x_k(t) + x_{k-1}(t))\, c_k(s)/\delta_n^2. \tag{15}$$

By using (14) and (15), the linear differential equation which approximates the system (13) is obtained. This is

$$d(\sum_1^n x_k(t)\, c_k(s))/dt = \sum_1^n (x_{k+1}(t) - 2\,x_k(t) + x_{k-1}(t)) c_k(s)/\delta_n^2 + \sum_1^n u_k(t)\, c_k(s)$$

We may use the boundary conditions given in (12) to obtain x_0 and x_{n+1}, *i.e.*, $x_0(t)=0$ and $x_{n+1}(t)=0$.

Hence, representing this equation by the vector differential equation, we have the following approximation system (given by finite difference methods):

$$x_n(t) = A_n x_n(t) + u_n(t),$$

where

$$x_n = (x_1, \ldots, x_n), \qquad u_n = (u_1, \ldots, n_n), \qquad x_n(0) = P_n 0 = (0, \ldots, 0),$$

$$A_n = \frac{1}{\delta_n^2} \begin{pmatrix} -2, & 1, & & & \\ 1, & -2, & 1, & & \bigcirc \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ & \bigcirc & 1, & -2, & 1 \\ & & & 1, & -2 \end{pmatrix}$$

If we consider that the cost function of original problem is such that

$$f = \int_0^T \int_0^1 \{|x(t,s) - y(t,s)|^2 + \lambda\, |u(t,s)|^2\}\, ds\, dt, \tag{16}$$

the appromimating cost function becomes to be

$$f_n = \delta_n \int_0^T \{(x_n(t) - y_n(t))'(x_n(t) - y_n(t)) + \lambda u_n(t)' u_n(t)\}\, dt, \tag{17}$$

where a prime denotes a transpose.

Now let us check the conditions (C), (S) and (C^*). Condition (S) is satisfied, since A_n is a negative definite symmetric matrix. Condition (C) can be easily verified after tedious caluculations. Condition (C^*) is automatically satisfied because both A and A_n are selfadjoint.

Therefore by Theorem 2, we can conclude that a sequence of approximating optimal controls associated with cost function (17) converges to an optimal control of the original system (12).

Conclusion

In this paper we have treated the convergence of approximation solutions for distributed parameter optimal control problems with non-linear cost functions.

We have shown under stated conditions that where the sequence of approximating solutions exists, there exists an optimal control of the original problem, and the sequence of values of approximate cost functions associated with these solutions converges to the true optimal value of the original cost function (Theorem 1).

For the tracking problem, moreover, the sequence of approximating solutions converges to the true optimal control (Theorem 2).

As an example, we have dealt with a tracking problem for one-dimensional diffusion equation, constructed a discretization scheme and checked the conditions of convergence.

Acknowledgment

The author wishes to thank Prof. H. Ishigaki, the School of Education, Waseda University, for helpful suggestions and discussions.

References

1) E. I. Axelband, Function space methods for the optimal control of a class of distributed parameter control system, Proc. Joint Automatic Control Conf., (1965), pp. 374-380.
2) H. F. Trotter, Approximation of semi-groups of operators, Pacific J. Math., 8 (1958), pp. 887-919.
3) P. D. Lax, Survey of the stability of linear finite difference equations, Comm. Pure Appl. Math., **9** (1956), pp. 267-293.
4) H. Sasai and E. Shimemura, On the convergence of approximating solutions for linear distributed parameter optimal control problems, SIAM J. on Control, **9** (1971), pp. 263-273.
5) A. V. Balakrishnan, Optimal control problems in Banach space, SIAM J. on Control, **3** (1968), pp. 109-127.
6) A. Friedan, Optimal control in Banach spaces, J. Math. Anal. Appl., **19** (1968), pp. 35-55.
7) K. Yoshida, Functional Analysis, Springer-Verlag, Berlin, 1965.
8) E. I. Axelband, The structure of the optimal tracking problem for distributed parameter systems. IEEE Trans. Autmo. Control, **13** (1968), pp. 50-57.

Appendix

We put

$$\begin{aligned} f(x^\circ(t),\, u^\circ(t)) &= f((Lu^\circ)(t),\, u^\circ(t)) \\ &= \int_0^T \|(Lu^\circ)(t) - y(t)\|_X^2\, dt + \int_0^T \|u^\circ(t)\|_X^2\, dt \\ &= G(u^\circ(t)) + R(u^\circ(t)) \end{aligned}$$

$$
\begin{aligned}
f_n(x_n^\circ(t),\ u_n^\circ(t)) &= f((L_n u_n^\circ)(t),\ u_n^\circ(t)) \\
&= \int_0^T \|(L_n u_n^\circ)(t) - P_n y(t)\|_{X_n}^2 dt + \int_0^T \|u_n^\circ(t)\|_{X_n}^2 dt \\
&= G_n(u_n^\circ(t)) + R_n(u_n^\circ(t)).
\end{aligned}
$$

Now let us prove that if $f_{nj}(x_{nj}^\circ(t),\ u_{nj}^\circ(t)) \to f(x^\circ(t),\ u^\circ(t))$, $R_{nj}(u_{nj}^\circ(t)) \to R(u^0(t))$.

We write n for n_j in the following. By the resonance theorem[7] it can be shown that

$$
\begin{aligned}
&\underline{\lim}\, G_n(u_n^\circ(t)) \geqq G(u^\circ(t)) \geqq 0 \\
&\underline{\lim}\, R_n(u_n^\circ(t)) \geqq R(u^\circ(t)) \geqq 0.
\end{aligned}
$$

The following inequalities are satisfied:

$$
\begin{aligned}
G(u^\circ(t)) + R(u^\circ(t)) &= \underline{\lim}\, \{G_n(u_n^\circ(t)) + R_n(u_n^\circ(t))\} \\
&\geqq \underline{\lim}\, G_n(u_n^\circ(t)) + \underline{\lim}\, R_n(u_n^\circ(t)) \\
&\geqq G(u^\circ(t)) + R(u^\circ(t)).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\underline{\lim}\, R_n(u_n^\circ(t)) &= G(u^\circ(t)) + R(u^\circ(t)) - \underline{\lim}\, G_n(u_n^\circ(t)) \\
&\leqq G(u^\circ(t)) + R(u^\circ(t)) - G(u^\circ(t)) = R(u^\circ(t)).
\end{aligned}
$$

Therefore $\underline{\lim}\, R_n(u_n^\circ(t)) = R(u^\circ(t))$.
We have, $\underline{\lim}\, G_n(u_n^\circ(t)) = G(u^\circ(t))$, in such the same way as the above.

On the other hand,

$$
\begin{aligned}
R(u^\circ(t)) &= \underline{\lim}\, R_n(u_n^\circ(t)) \leqq \overline{\lim}\, R_n(u_n^\circ(t)) \\
&= \overline{\lim}\, \{G_n(u_n^\circ(t)) + R_n(u_n^\circ(t)) - G_n(u_n^\circ(t))\} \\
&\leqq \overline{\lim}\, \{G_n(u_n^\circ(t)) + R_n(u_n^\circ(t))\}\ \overline{\lim}\, \{-G_n(u_n^\circ(t))\} \\
&= G(u^\circ(t)) + R(u^\circ(t)) - \underline{\lim}\, G_n(u_n^\circ(t)) \\
&= R(u^\circ(t)).
\end{aligned}
$$

Consecuently it can be concluded that

$$
\lim R_n(u_n^\circ(t)) = R(u^\circ(t)).
$$