STUDIA MATHEMATICA, T. XXXV. (1970)

The minimal-norm problem and Pontriagin's maximum principle for Banach spaces, II

by

GRAŻYNA TOPOROWSKA (Warszawa)

Balakrishnan in his work [1] discusses some classes of control problems in which the state and control variables are allowed to range in Banach spaces. The system is described by the differential linear equation in a Banach space, with the right side linearly dependent on a control, of the form

$$\frac{dx(t)}{dt} = Ax(t)+Bu(t),$$

where $x(t)$ for each t belongs to a Banach space X_1, $u(t)$-control, for each t belongs to another Banach space X_2, the linear operator $A: X_1 \to X_1$ is a generator of a strongly continuous semigroup of bounded operators, and $B: X_2 \to X_1$ is a linear bounded operator.

The minimal-norm problem is considered for this system. It consists in minimizing the functional $\|x(T)-y\|_{X_1}$ for a fixed final time $t = T$ and with a fixed $y \in X_1$.

Balakrishnan has proved Pontriagin's maximum principle for a problem posed as above.

Paper [5] includes a generalization of Balakrishnan's result to the case in which the system is described by the differential equation, with the right side non-linearly dependent on a control, of the form

$$\frac{dx(t)}{dt} = Ax(t)+B\big(t, u(t)\big),$$

where $B(\cdot, \cdot)$ is a non-linear operator. Pontriagin's maximum principle is true for this problem (cf. [5]).

This work is a generalization of the results contained in [1] and [5].

Consider the equation

$$\frac{dx(t)}{dt} = Ax(t)+B\big(t, x(t), u(t)\big) \tag{1}$$

with the initial value

$$x(0) = x_1, \tag{2}$$

where:

$x(\cdot)$ is a continuous function defined in the interval $[0, T]$ with values in a Banach space X_1 (i.e. $x(\cdot)\in C([0, T]; X_1)$);

$u(\cdot)$, the control, is a bounded function defined in the interval $[0, T]$ with values in a Banach space X_2 (i.e. $u(\cdot)\in M([0, T]; X_2)$) and Bochner integrable in $[0, T]$;

A: $X_1 \to X_1$, a linear operator, is the generator of a C_0 strongly continuous semigroup $S(t) = e^{tA}$;

$B(\cdot, \cdot, \cdot)$: $[0, T]\times X_1\times X_2 \to X_1$ is a continuous mapping, differentiable by the second and the third variable on $[0, T]\times X_1\times X_2$, and the partial derivatives D_2B and D_3B (D_iB denotes the partial derivative of B with respect to the i-th variable) are continuous with respect to all the variables in the operator-norms in $\mathscr{L}(X_1; X_1)$ and $\mathscr{L}(X_2; X_1)$ respectively and are bounded.

We say that the function $x(\cdot)$, strongly continuous on the interval $[0, T]$ with the values in X_1, is the *generalized solution* of the initial value problem (1), (2) if it satisfies the integrable equation

$$x(t) = S(t)x_1 + \int_0^t S(t-\tau)B\big(\tau, x(\tau), u(\tau)\big)\,d\tau. \tag{3}$$

Under our assumptions we can prove that for each $x_1\in X_1$ and for each control $u(\cdot)\in M([0, T]; X_2)$ the initial value problem (1), (2) has the unique generalized solution. This solution may be found by the method of successive approximations.

Let Λ be a mapping assigning to every control $u(\cdot)\in M([0, T]; X_2)$ the final state $x(T)$ of the solution $x(\cdot)$ of problem (1), (2) corresponding to that control.

The functional space $M([0, T]; X_2)$, which we denote shortly by M, will be treated as a Banach space with the norm

$$\|u(\cdot)\|_M = \sup_{0\leqslant t\leqslant T} \|u(t)\|_{X_2}. \tag{4}$$

Remember that the linear operator $\Lambda'(u)\in L(M; X_1)$ such that

$$\|\Lambda(u+h) - \Lambda(u) - \Lambda'(u)h\|_{X_1} = o(\|h\|_M),$$

where $\|o(\|h\|_M)\|/\|h\| \underset{\|h\|\to 0}{\to} 0$ for each $u\in M$, is called the *strong derivative* of the mapping Λ at the point $u\in M$.

THEOREM 1. *The mapping Λ has a strong derivative at each point of* $M([0, T]; X_2)$.

Proof. We show first that the solution of the initial problem (1), (2) is a differentiable function of the control. Let us introduce a mapping

$$\mathscr{F}(\cdot, \cdot)\colon\ C\times M \to C,$$

where $C = C([0, T]; X_1)$ is defined as follows:

$$\mathscr{F}(x, u)(t) = x(t) - \int_0^t S(t-\tau)B\big(\tau, x(\tau), u(\tau)\big)\,d\tau - S(t)x_1 \tag{5}$$

for $\big(x(t), u(t)\big)\in X_1\times X_2$ and $t\in[0, T]$.

We show that the mapping $\mathscr{F}$ is strongly differentiable, and, furthermore, the following formulas are true:

$$\begin{aligned}\big(D_1\mathscr{F}(x, u)\cdot\sigma\big)(t) &= \sigma(t) - \int_0^t S(t-\tau)D_2B\big(\tau, x(\tau), u(\tau)\big)\cdot\sigma(\tau)\,d\tau,\\ \big(D_2\mathscr{F}(x, u)\cdot\Delta\big)(t) &= -\int_0^t S(t-\tau)D_3B\big(\tau, x(\tau), u(\tau)\big)\cdot\Delta(\tau)\,d\tau,\end{aligned} \tag{6}$$

where:

$D_1\mathscr{F}(x, u)\in\mathscr{L}(C; C)$ for $x\in C$, $u\in M$;

$\sigma\in C$, $\sigma(\tau)\in X_1$, $D_1\mathscr{F}(x, u)\cdot\sigma\in C$;

$\big(D_1\mathscr{F}(x, u)\cdot\sigma\big)(t)\in X_1$ and it denotes the value of the derivative $D_1\mathscr{F}(x, u)$ at the point σ, at the moment t;

D_2B: $[0, T]\times X_1\times X_2 \to \mathscr{L}(X_1; X_1)$;

$D_2\mathscr{F}(x, u)\in\mathscr{L}(M; C)$ for $x\in C$, $u\in M$;

$\Delta\in M$, $\Delta(\tau)\in X_2$, $D_2\mathscr{F}(x, u)\cdot\Delta\in M$;

$\big(D_2\mathscr{F}(x, u)\cdot\Delta\big)(t)\in X_2$ and it denotes the value of the derivative $D_2\mathscr{F}(x, u)$ at the point Δ, at the moment t;

D_3B: $[0, T]\times X_1\times X_2 \to \mathscr{L}(X_2; X_1)$.

By the definition of the strong derivative, we should investigate the norm

$$\begin{aligned}&\|\mathscr{F}(x+h, u)(t) - \mathscr{F}(x, u)(t) - \big(D_1\mathscr{F}(x, u)\cdot h\big)(t)\|_{X_1}\\ &= \|\mathscr{F}(x+h, u)(t) - \mathscr{F}(x, u)(t) - h(t) + \int_0^t S(t-\tau)D_2B\big(\tau, x(\tau), u(\tau)\big)h(\tau)\,d\tau\|_{X_1}\\ &= \|(x+h)(t) - \int_0^t S(t-\tau)B\big(\tau, (x+h)(\tau), u(\tau)\big)\,d\tau - S(t)x_1 - x(t) +\\ &\quad + \int_0^t S(t-\tau)B\big(\tau, x(\tau), u(\tau)\big)\,d\tau + S(t)x_1 - h(t) +\\ &\quad + \int_0^t S(t-\tau)D_2B\big(\tau, x(\tau), u(\tau)\big)\cdot h(\tau)\,d\tau\|_{X_1}\\ &= \|(x+h)(t) - x(t) - h(t) + \int_0^t S(t-\tau)\big[-B\big(\tau, (x+h)(\tau), u(\tau)\big) +\\ &\quad + B\big(\tau, x(\tau), u(\tau)\big) + D_2B\big(\tau, x(\tau), u(\tau)\big)\cdot h(\tau)\big]\,d\tau\|_{X_1}\\ &\leqslant K\int_0^T \eta\big(\tau, h(\tau)\big)\,d\tau\end{aligned}$$

where $K = \sup_{t\in[0,T]} \|S(t)\|$, and

$$\eta\big(\tau, h(\tau)\big) = \|B\big(\tau, x(\tau)+h(\tau), u(\tau)\big) - \\ - B\big(\tau, x(\tau), u(\tau)\big) - D_2B\big(\tau, x(\tau), u(\tau)\big)h(\tau)\|_{X_1}.$$

By the differentiability assumption of B we have

$$\eta\big(\tau, h(\tau)\big) = o\big(\|h(\tau)\|_{X_1}\big)$$

for every $\tau\in[0, T]$ and, moreover,

$$0 \leqslant \eta\big(\tau, h(\tau)\big) \leqslant 2N\|h(\tau)\|_{X_1} \leqslant 2N\|h\|_C,$$

where

$$N = \sup\{\|D_2B(t, x, u)\|:\ t\in[0, T], x\in X_1, u\in X_2\}.$$

Thus, by the Lebesgue bounded convergence theorem it follows that

$$K\int_0^T \eta\big(\tau, h(\tau)\big)d\tau = o(\|h\|_C),$$

which proves the first of the formulae (6). The proof of the second formula is analogous.

We shall prove that the partial derivatives $D_1\mathscr{F}$ and $D_2\mathscr{F}$ are continuous. In fact, if $x_n(\tau)\to x(\tau)$ and $u_n(\tau)\to u(\tau)$, then

$$\Big\|D_2B\big(\tau, x_n(\tau), u_n(\tau)\big) - D_2B\big(\tau, x(\tau), u(\tau)\big)\Big\| \to 0$$

for every $\tau\in[0, T]$, and

$$\|D_1\mathscr{F}(x_n, u_n) - D_1\mathscr{F}(x, u)\| \\ \leqslant \int_0^T \Big\|S(t-\tau)\big[D_2B\big(\tau, x_n(\tau), u_n(\tau)\big) - D_2B\big(\tau, x(\tau), u(\tau)\big)\big]\Big\|d\tau \to 0$$

by the Lebesgue theorem, because the function $\Big\|S(t-\tau)D_2B\big(\tau, x(\tau), u(\tau)\big)\Big\|$ is bounded. The proof for $D_2\mathscr{F}$ is similar.

Let $\Phi(u) = x(\cdot)$ denotes the solution of problems (1), (2) corresponding to the control $u(\cdot)$. So equality (5) is of the form

$$\mathscr{F}\big(\Phi(u), u\big) = 0, \tag{7}$$

where $\Phi(u)\in C$, $u\in M$, $\mathscr{F}\big(\Phi(u), u\big)\in C$.

The operator $D_1\mathscr{F}(x, u)$, as an integral operator of the Volterra type, is a continuous linear automorphism of the space C. Hence, by the differentiation of the implicit function [2], it follows that

$$D\mathscr{F}\big(\Phi(u), u\big) = D_1\mathscr{F}\big(\Phi(u), u\big)\circ D\Phi(u) + D_2\mathscr{F}\big(\Phi(u), u\big) = 0,$$

where $D\mathscr{F}\big(\Phi(u), u\big)\in\mathscr{L}(M; C)$, $D_1\mathscr{F}\big(\Phi(u), u\big)\in\mathscr{L}(C; C)$, $D\Phi(u)\in\mathscr{L}(M; C)$, $D_2\mathscr{F}\big(\Phi(u), u\big)\in\mathscr{L}(M; C)$. Since $\Phi(u)$ is fulfilled (3) and by (7), we have

$$\begin{aligned} \big(D\Phi(u)\cdot\delta u\big) &= \int_0^t S(t-\tau)D_2B\big(\tau, \Phi(u)(\tau), u(\tau)\big)\circ\big(D\Phi(u)\cdot\delta u\big)(\tau)\,d\tau + \\ &\quad + \int_0^t S(t-\tau)D_3B\big(\tau, \Phi(u)(\tau), u(\tau)\big)\cdot\delta u(\tau)\,d\tau. \end{aligned} \tag{8}$$

The differentiation of the mapping Φ: $M\ni u\to\Phi(u)\in C$ implies the differentiation of Λ. Thus the operator Λ has a strong derivative, q.e.d.

Now let $G\subset M([0, T]; X_2)$ be a convex, closed and bounded set. Let $x_1\in X_1$ and $y\in X_1$ be fixed. Consider the problem of minimizing the functional

$$\Psi(u) = \|\Lambda(u) - y\|_{X_1} \tag{9}$$

on the set G. The control u_0 for which

$$\Psi(u_0) = \min_{u\in G}\Psi(u)$$

will be called the *optimal control.*

THEOREM 2. *If the minimum of functional Ψ is reached at the point $u_0\in G$, then there exists a functional $x^*\in X_1^*$ such that for every $u\in G$ the inequality*

$$x^*[\Lambda'(u_0)\cdot u_0] \geqslant x^*[\Lambda'(u_0)\cdot u]$$

is fulfilled.

Proof. We shall now use the following theorem:

Let X be a Banach space, $A, B\subset X$ — convex sets, $\operatorname{Int} A\neq\emptyset$, $(\operatorname{Int} A)\cap B = \emptyset$; then there exist a functional $x^*\in X^*$ and a constant γ such that $x^*(\operatorname{Int} A) > \gamma$ and $x^*(A)\geqslant\gamma\geqslant x^*(B)$, i.e. $x^*(A)\geqslant x^*(B)$.

The equality

$$[\Lambda'(G) + x_0(T)\,]\cap\operatorname{Int} S = \emptyset, \tag{10}$$

where $S = \{\Lambda u\colon \|\Lambda u - y\|\leqslant m,\ m = \inf_{u\in G}\|\Lambda u - y\|\}$ is true.

The proof of (10) is given in [5] (see formula (11) in that work) and it is sufficient that Λ' is a weak derivative. In our case Λ' is a strong derivative, and so Λ' is a weak derivative also.

Hence in our case all the assumptions of the theorem quoted above are satisfied and, as the assertion of that theorem, we obtain

$$x^*(S)\geqslant x^*[\Lambda'(G) + x_0(T)],$$

whence after some transformulations analogous to the transmutations made in [5], we obtain

$$x^*\big(\Lambda'(u_0)\cdot u_0\big) \geqslant x^*\big(\Lambda'(u_0)\cdot u\big)$$

for all $u \in G$, q.e.d.

Theorem 2 is a generalization of Pontriagin's maximum principle. It does not give the fact that $x^*(\sigma)$ is the solution of the conjugate equation to (1), which holds for the classical Pontriagin's maximum principle [4] and which holds for the problem presented in [5].

References

[1] A. V. Balakrishnan, *Optimal control problems in Banach spaces*, J. SIAM Control, Ser. A, Vol. 3, Nr 1, USA, 1965.

[2] J. Dieudonné, *Foundations of modern analysis*, New York and London 1960.

[3] J. Kisyński, *Differentiation of solution of ordinary differential equations with respect to parameter*, Bull. Acad. Pol. Sc. 14 (1966), p. 553-557.

[4] Л. С. Понтрягин, В. Г. Болтянский, Р. В. Гамкрелидзе и Е. Ф. Мищенко, *Математическая теория оптимальных процессов*, Москва 1961.

[5] G. Toporowska, *The minimal-norm problem and Pontriagin's maximum principle for Banach spaces, I*, Studia Math., 33 (1969), p. 175-183.

Reçu par la Rédaction le 8. 2. 1969

STUDIA MATHEMATICA, T. XXXV. (1970)

Banach limits in vector lattices

by

A. L. PERESSINI (Urbana, Ill.)

S. Banach ([2], p. 34) defined generalized limits (now commonly referred to as Banach limits) as positive, linear, shift-invariant functionals on l^∞ which assign the value 1 to the constant sequence with terms equal to 1. Lorentz [4] investigated the linear subspace of l^∞ consisting of those sequences in l^∞ to which all Banach limits assigned the same value; such sequences were termed almost convergent. In this paper, we consider the extension of the concepts of Banach limit and almost convergence to vector-valued sequences. In addition to seeking generalizations of the known results for the case of real-valued sequences, we also consider a number of new questions which do not arise in the classical case. Our primary objective will be to study the geometric structure of the space of almost convergent sequences and the set of Banach limits in this more general context. Notation and terminology concerning ordered vector spaces will follow Peressini [5].

1. If E is a vector lattice, then the collection $\omega(E)$ of all sequences $\vec{x} = (x_n)$ such that $x_n \in E$ for all n is a vector lattice for the usual "coordinatewise" definitions of the linear operations and order. $l^\infty(E)$ will denote the linear subspace of $\omega(E)$ consisting of all order bounded sequences, that is, all sequences $\vec{x} = (x_n)$ for which there exist y, z in E such that $y \leqslant x_n \leqslant z$ for all n. A linear mapping $L\colon l^\infty(E) \to E$ is a *Banach limit on E* if

(1) L is positive.

(2) L is shift-invariant (i.e. $L(\vec{x}) = L(\sigma\vec{x})$, where σ is the "left-shift" on $l^\infty(E)$ defined by $\sigma\big((x_n)\big) = (x_{n+1})$.

(3) If $c \in E$ and $\vec{c}$ is the constant sequence with n^{th} term c, then $L(\vec{c}) = c$.

It follows from (2) that $L(\vec{x}) = L(\sigma^k\vec{x})$ for each natural number k, where σ^k denotes the k^{th}-iterate of σ.

If E is an order complete vector lattice and if $\vec{x} = (x_n) \in l^\infty(E)$, then

$$\underline{\lim}\, x_n = \sup_k \inf_{n\geqslant k}\{x_n\} \quad \text{and} \quad \overline{\lim}\, x_n = \inf_k \sup_{n\geqslant k}\{x_n\}$$