Diese eigentlich selbstverständliche Voraussetzung ist in [1] vergessen worden, man findet sie aber bei Ditkin und Prudnikov [2]. Man könnte jetzt zwar daran denken, die Vermutung von Mikusiński durch eine einfache Abänderung des Beispiels (3) zu widerlegen, doch scheint dies unmöglich zu sein, da für dieses Beispiel die Existenz von Nullteilern (bei der gewöhnlichen Multiplikation) wesentlich ist.

Aus der Existenz des Gegenbeispiels von Frau Sändig (bezüglich der gewöhnlichen Multiplikation) geht noch hervor, daß die Beweise der Spezialfälle I und III in [1] nur einen lokalen Charakter besitzen, die Aussagen also nur für hinreichend kleine λ-Intervalle bewiesen sind. Will man globale Aussagen für alle vorkommenden λ haben, so sind noch Zusatzbetrachtungen erforderlich.

Außerdem ist im Fall I noch die Voraussetzung hinzuzufügen, daß die dort vorkommenden Laplace-Integrale bezüglich λ gleichmäßig existieren.

Literaturnachweis

[1] L. Berg, *Asymptotische Auffassung der Operatorenrechnung*, Studia Math. 21 (1962), S. 215-229.

[2] V. A. Ditkin u. A. P. Prudnikov, *Integral transforms and operational calculus*, Oxford 1965.

[3] J. Mikusiński, *Operatorenrechnung*, Berlin 1957.

[4] A.-M. Sändig, *Verallgemeinerte Mittelwertformeln*, Diplomarbeit Univ. Rostock 1968.

Reçu par la Rédaction le 15. 6. 1968

STUDIA MATHEMATICA, T. XXXIII. (1969)

The minimal norm problem and Pontriagin's maximum principle for Banach spaces (I)

by

GRAŻYNA TOPOROWSKA (WARSZAWA)

1. Introduction. Balakrishnan [1] discusses some classes of control problems, in which the state and input or control variables are allowed to range in Banach spaces. The equation considered in a general case is of the form

$$\dot{x}(t) = f(x(t), u(t), t),$$

where, for each t, $u(t)$ and $x(t)$ are values in a Banach space. For the linear case we have

$$\dot{x}(t) = f(x(t), u(t), t) = A(t)x(t)+B(t)u(t)+z(t), \tag{1}$$

where $x(t)$ and $z(t)$ for each t belong to a Banach space X_1, $u(t)$ — a control — for each t belongs to another Banach space X_2, $A(t)$ and $B(t)$ are linear operators for each t, $B(t)\colon X_2 \to X_1$ for each t is a bounded operator, and $A(t)\colon X_1 \to X_1$ for each t is a closed operator (not necessarily bounded) with domain dense in X_1.

Balakrishnan in his work discusses the minimal norm problem for a particular case of equation (1), i.e. for

$$\dot{x}(t) = Ax(t)+Bu(t), \tag{2}$$

where A and B are constant operators and, besides, the operator A is the generator of a strongly continuous semigroup $S(t)$ of bounded operators (see [3] and [5]). The solution of equation (2) is of the following form:

$$x(t) = S(t)x(0)+\int_0^t (t-\sigma)Bu(\sigma)d\sigma.$$

The minimal norm problem consists in finding a control $u_0(t)$ with a corresponding state $x_0(t)$ such that

$$\|x_0(T)-y\| = \min \|x(T)-y\|$$

for a fixed final time $t = T$ and with a fixed $y \in X_1$. Besides, the following constraint is imposed on the control $u(t)$: consider the space $B_p(X_2; [0, T])$ of strongly measurable functions $u(\cdot)\colon [0, T] \to X_2$ such that

$$\int_0^T \|u(t)\|^p dt < \infty \quad \text{for a fixed } p \ (1 < p < \infty).$$

We assume that the control $u(\cdot)$ belongs to a closed bounded convex set $C \subset B_p(X_2; [0, T])$. Balakrishnan proved Pontriagin's maximum principle (see [4]) for a problem posed as above for equation (2).

This paper is concerned with the minimal norm problem for an equation with the right-hand side not necessarily linearly dependent on a control, of form

$$\dot{x}(t) = Ax(t) + B(t, u(t))$$

with the initial value $x(0) = x_1$ and shows the validity of Pontriagin's maximum principle for such problems.

2. The minimal norm problem and Pontriagin's maximum principle. Consider the equation

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

with the initial value

$$x(0) = x_1,$$

where $x(\cdot)$ is a function defined in the interval $[0, T]$ with values in a Banach space X_1, $u(\cdot)$ — a control — is a function defined in the interval $[0, T]$ with values in a Banach space X_2 and Bochner integrable in $[0, T]$, $A\colon X_1 \to X_1$ — a linear operator — is a generator of a C_0 strongly continuous semigroup $S(t)$ (see [5], chap. III), and $B(\cdot, \cdot)$ is a mapping of $[0, T] \times X_2$ into X_1 such that

1^0 for every fixed $u \in X_2$ the element $B(t, u)$ of X_1 is a function of the variable t, strongly measurable on the interval $[0, T]$,

2^0 it is strongly differentiable by the second variable on $[0, T] \times X_2$ and the partial derivative $B'_u(t, u)$ depends on the variable u in the operator norm in $\mathscr{L}(X_2, X_1)$ and is bounded in the operator norm.

We shall discuss the problem of minimizing the functional

$$\|x(T) - y\|, \tag{2}$$

where $y \in X_1$ is a fixed element.

Let the pair $(x_0(t), u_0(t))$ satisfy equation (1) and let the state $x_0(t)$ corresponding to the control $u_0(t)$ minimize the given functional (2).

The solution of equation (1) is of the form

$$x(t) = S(t)x(0) + \int_0^t S(t-s)B(s, u(s))\,ds \tag{3}$$

since the solution of non-homogeneous equation

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

is

$$x(t) = S(t)x(0) + \int_0^t S(t-s)v(s)\,ds.$$

In our case we have a Bochner integrable $B(s, u(s))$ instead of $v(s)$.

We now disturb the control $u(t)$ by putting $u_0(t) + \varepsilon\delta u(t)$. Let the state $\bar{x}(t)$ correspond to this disturbed control. We shall prove that

$$\bar{x}(t) = x_0(t) + \varepsilon\delta x(t) + o_t(\varepsilon), \tag{4}$$

where $\delta x(t)$ satisfies the equation:

$$\frac{dx(t)}{dt} = Ax(t) + B'_u(t, u_0(t))\,\delta u(t). \tag{5}$$

By assuming that the state $\bar{x}(t)$ corresponds to the control $u_0(t) + \varepsilon\delta u(t)$ we have

$$\frac{d\bar{x}(t)}{dt} = A\bar{x}(t) + B(t, u_0(t) + \varepsilon\delta u(t))$$

and $\bar{x}(t)$, from (3), is of the form

$$\begin{cases} \bar{x}(t) = S(t)\bar{x}(0) + \int_0^t S(t-s)B(s, u_0(s) + \varepsilon\delta u(s))\,ds, \\ \bar{x}(0) = x_1. \end{cases} \tag{6}$$

For every fixed $t \in [0, T]$ the element $B(t, u_0(t) + \sigma\delta u(t))$ from X_1 is a function of the variable σ in $[0, \varepsilon]$ with a continuous derivative

$$\frac{d}{d\sigma}[B(t, u_0(t) + \sigma\delta u(t))] = B'_u(t, u_0(t) + \sigma\delta u(t))\,\delta u(t)$$

by differentiation of a composite function. Hence

$$\begin{aligned} B(t, u_0(t) + \varepsilon\delta u(t)) - B(t, u_0(t)) &= \int_0^\varepsilon B'_u(t, u_0(t) + \sigma\delta u(t))\,\delta u(t)\,d\sigma \\ &= \Big[\int_0^\varepsilon B'_u(t, u_0(t) + \sigma\delta u(t))\,d\sigma\Big]\,\delta u(t). \end{aligned} \tag{7}$$

Since the solution of equation (1) is the state $x_0(t)$ corresponding to the control $u_0(t)$, it is of the form

$$\begin{cases} x_0(t) = S(t)x_0(0)+\int\limits_0^t S(t-s)B\big(s, u_0(s)\big)ds, \\ x_0(0) = x_1, \end{cases}$$

and hence, from (6) and the above equality

$$\bar{x}(t)-x_0(t) = \int\limits_0^t S(t-s)\big[B\big(s, u_0(s)+\varepsilon\delta u(s)\big)-B\big(s, u_0(s)\big)\big]ds,$$

$$\bar{x}(0) = x_0(0) = x_1$$

and then from (7) we have

$$\bar{x}(t)-x_0(t) = \int\limits_0^t S(t-s)\Big[\int\limits_0^\varepsilon B'_u\big(s, u_0(s)+\sigma\delta u(s)\big)d\sigma\Big]\,\delta u(s)ds.$$

On the other hand, since $\delta x(t)$ satisfies equation (5) and $\delta x(0) = 0$,

$$\delta x(t) = \int\limits_0^t S(t-s)B'_u\big(s, u_0(s)\big)\delta u(s)ds.$$

Therefore

$$\begin{aligned} &\bar{x}(t)-x_0(t)-\varepsilon\delta x(t) \\ &= \int\limits_0^t S(t-s)\Big[\int\limits_0^\varepsilon B'_u\big(s, u_0(s)+\sigma\delta u(s)\big)d\sigma-\varepsilon B'_u\big(s, u_0(s)\big)\Big]\,\delta u(s)ds \\ &= \int\limits_0^t S(t-s)\Big\{\int\limits_0^\varepsilon \big[B'_u\big(s, u_0(s)+\sigma\delta u(s)\big)-B'_u\big(s, u_0(s)\big)\big]d\sigma\Big\}\,\delta u(s)ds. \end{aligned}$$

Now, let

$$M = \sup_{[0,T]}\|S(t)\|,$$

$$K_\varepsilon(s) = \sup_{\sigma\in[0,\varepsilon]}\|B'_u\big(s, u_0(s)+\sigma\delta u(s)\big)-B'_u\big(s, u_0(s)\big)\|_{\mathscr{L}(X_2,X_1)}.$$

With the above denotations, for $t\in[0, T]$

(8) $$\|\bar{x}(t)-x_0(t)-\varepsilon\delta x(t)\|_{X_1} \leqslant M\varepsilon\int\limits_0^T K_\varepsilon(s)\|\delta u(s)\|_{X_2}ds.$$

From the assumed continuity of the partial derivative $B'_u(t, u)$ by u, we have

(9) $$\lim_{\varepsilon\to 0}K_\varepsilon(s) = 0$$

for every $s\in[0, T]$.

$K_\varepsilon(s)$ is a bounded function, since we assumed that $B'_u(t, u)$ is bounded, and therefore $K_\varepsilon(s)\|\delta u(s)\|$ is an integrable function. Let us note that it is non-decreasing by ε. Hence from (9) and Lebesgue's theorem we have that

$$\lim_{\varepsilon\to 0}\int\limits_0^T K_\varepsilon(s)\|\delta u(s)\|_{X_2}ds = 0.$$

From this equality and from (8) we can see that

$$\|\bar{x}(t)-x_0(t)-\varepsilon\delta x(t)\|_{X_1} \leqslant o(\varepsilon)$$

for all $t\in[0, T]$, where $o(\varepsilon)$ does not depend on t. The above inequality can be written as a condition equivalent to (4), that is

$$\bar{x}(t) = x_0(t)+\varepsilon\delta x(t)+o_t(\varepsilon)$$

where

$$\lim_{\varepsilon\to 0}\frac{\|o_t(\varepsilon)\|}{\varepsilon} = 0$$

uniformly with respect to t in $[0, T]$.

Consider now space $B_p(X_2; [0, T])$ of strongly measurable functions $u(t)$ with range in X_2 and such that

$$\int\limits_0^T \|u(t)\|^p dt < \infty \quad \text{for a fixed } p\ (1 < p < \infty).$$

Now let in our case the convex bounded set $C \subset B_p(X_2; [0, T])$ be the set of controls. Let us introduce an operator A, mapping the set C into X_1 and assigning to every control $u(\cdot)$ the final state $x(T)$ of the solution $x(\cdot)$ corresponding to that control. This operator shall be defined as follows:

(10) $$A\big(u(\cdot)\big) = S(T)x(0)+\int\limits_0^T S(T-s)B\big(s, u(s)\big)ds = x(T).$$

$S(T)x(0)$ is a fixed point in X_1. Therefore practically our attention in this discussion shall be concentrated on the integral operator.

From our assumptions we know that $\big(x_0(t), u_0(t)\big)$ satisfies equation (1) and minimizes the given functional (2). This implies that the point $x_0(T) = A\big(u_0(\cdot)\big)$ is a boundary point of $A(C)$. Let us consider a set of disturbed controls $\{u_0(\cdot)+\varepsilon\delta u(\cdot)\}$, where $\delta u(\cdot)$ is an admissible control, i.e. there exists an $\varepsilon_0 > 0$ such that $u_0(\cdot)+\varepsilon\delta u(\cdot)$ is also an element of C provided that $0 < \varepsilon < \varepsilon_0$. It can be seen that the set K of admissible controls forms a cone with vertex in 0. To make it clear, we have to prove that if $\delta_1 u$ and $\delta_2 u$ are admissible controls and λ_1 and λ_2 are non-negative numbers, then

$$\delta_3 u = \lambda_1\delta_1 u+\lambda_2\delta_2 u$$

is also an admissible control.

If $\delta_1 u$ and $\delta_2 u$ are admissible controls, then ,by definition, there exist $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$ such that $u_0 + \varepsilon' \delta_1 u \in C$ for $0 < \varepsilon' < \varepsilon_1$ and $u_0 + \varepsilon'' \delta_2 u \in C$ for $0 < \varepsilon'' < \varepsilon_2$. When $\lambda_1 = \lambda_2 = 0$, $\delta_3 u = 0$ is obviously an admissible control. If $\lambda_1 + \lambda_2 > 0$, then $\varepsilon_0 = (\lambda_1/\varepsilon_1 + \lambda_2/\varepsilon_2)^{-1}$ is a well-defined positive number. Let $\alpha_i = \lambda_i/\varepsilon_i \varepsilon_0$; then $\alpha_i \geqslant 0$ and $\alpha_1 + \alpha_2 = 1$. When $0 < \varepsilon < \varepsilon_0$, $u_0 + \varepsilon \varepsilon_0^{-1} \varepsilon_i \delta_i u$ is an element of C for $i = 1, 2$, and this implies, because of the convexity of C, that

$$u_0 + \varepsilon \delta_3 u = \alpha_1 (u_0 + \varepsilon \varepsilon_0^{-1} \varepsilon_1 \delta_1 u) + \alpha_2 (u_0 + \varepsilon \varepsilon_0^{-1} \varepsilon_2 \delta_2 u) \in C.$$

Therefore $\delta_3 u$ is an admissible control. Obviously $C \subset u_0(\cdot) \cup K$.

The functional (2) defines a certain sphere S with centre in y,

$$S = \{z \colon \|z - y\| \leqslant m\},$$

where

$$m = \inf_{u \in C} \|\Lambda u - y\|.$$

Suppose $m > 0$. By the assumption that $x_0(T) = \Lambda u_0(\cdot)$ minimizes, the functional (2) we know that this point is a boundary point of S and — by the definition of Λ — a boundary point of the set $\Lambda(C)$. According to (4) we know that operator Λ has a weak differential in $u_0(\cdot)$, which will be denoted by Λ'. With this denotation we have, for every admissible control δu,

$$\Lambda\big(u_0(\cdot) + \varepsilon \delta u(\cdot)\big) = \Lambda\big(u_0(\cdot)\big) + \Lambda \cdot \varepsilon \delta u(\cdot) + o(\varepsilon),$$

where $o(\varepsilon) = o_T(\varepsilon)$ and Λ' is a linear operator assigning to the admissible control δu the element

$$\Lambda' \delta u = \int_0^T S(T-s) B_u'\big(s, u_0(s)\big) \delta u(s)\, ds.$$

We shall now prove that

$$[\Lambda'(K) + x_0(T)] \cap \operatorname{Int} S = \emptyset. \tag{11}$$

Λ' is a linear operator, whence it transforms a cone into a cone and a vertex into a vertex. Suppose (11) is not true. Then there exists a direction $\delta x(T)$ corresponding to the control $\delta u(\cdot)$ and entering the interior of S. The curve corresponding to that direction, $x_0(T) + \varepsilon \delta x(T)$ $\big(\varepsilon \to \Lambda(u_0 + \varepsilon \delta u)\big)$ is tangent to the direction $\delta x(T)$. Thus the corresponding point

$$x(T) = x_0(T) + \varepsilon \delta x(T) + o_T(\varepsilon)$$

belongs to the interior of S (Int S) provided that $\varepsilon > 0$ is small enough. But, on the other hand, $x(T) \in \Lambda(C)$. Therefore

$$\Lambda(C) \cap \operatorname{Int} S \neq \emptyset,$$

which is a contradiction, since if it were so, the couple $(x_0(t), u_0(t))$ would not be the optimal solution. Thus we must have

$$[\Lambda'(K) + x_0(T)] \cap \operatorname{Int} S = \emptyset.$$

We shall now use the following theorem (see [3], theorem 2.6.3):

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. $x^*(A) \geqslant x^*(B)$.

In our case, it follows from (11) that there exists a functional $x^* \in X_1^*$ such that. $x^*(S) \geqslant x^*\big(\Lambda'(K) + x_0(T)\big)$, i.e.

$$x^*\Big(\Lambda\big(u_0(\cdot)\big)\Big) \geqslant x^*\big(\Lambda'(K) + x_0(T)\big),$$

but, since x^* is linear and $x^*\Big(\Lambda\big(u_0(t)\big)\Big) = x^*\big(x_0(T)\big)$, we have $x^*\big(\Lambda'(K)\big) \leqslant 0$, and hence, since $\delta u = u - u_0$ is a permissible control for every $u \in C$ and the operator Λ' and the functional x^* are linear,

$$x^*\Big(\Lambda'\big(u_0(\cdot)\big)\Big) \geqslant x^*\Big(\Lambda'\big(u(\cdot)\big)\Big) \quad \text{for } u \in C. \tag{12}$$

Note that

$$\Lambda' = \Lambda_1 \Lambda_2 \colon B\ (X_2; [0, T]) \to X_1,$$

where

$$(\Lambda_2 u)(t) = B_u'\big(t, u_0(t)\big) u(t)$$

for $u(\cdot) \in B_p(X_2; [0, T])$;

$$\Lambda_1 v = \int_0^T S(T-s) v(s)\, ds \in X_1,$$

$$\Lambda_1 \colon B_p(X_1; [0, T]) \to X_1.$$

Let us now define y^* as

$$y^* = \Lambda_1^* x^*, \tag{13}$$

where

$$\begin{aligned} &\Lambda_2^* \colon B_q(X_1^*; [0, T]) \to B_q(X_2^*; [0, T]), \\ &\Lambda_1^* \colon X_1^* \to B_q(X_1^*; [0, T]) \end{aligned} \qquad \left(\frac{1}{p} + \frac{1}{q} = 1\right),$$

and

$$\Lambda'^* = \Lambda_2^* \Lambda_1^* \colon X_1^* \to B_q(X_2^*; [0, T]);$$

thus $y^* \in B_q(X_1^*; [0, T])$.

From (12) we have $x^*\big(\Lambda_1 \Lambda_2 (u_0)\big) \geqslant x^*\big(\Lambda_1 \Lambda_2 (u)\big)$, whence

$$\Lambda_1^* x^*\big(\Lambda_2(u_0)\big) \geqslant \Lambda_1^* x^*\big(\Lambda_2(u)\big),$$

that is, from (13),

(14) $$y^*(A_2(u_0)) \geqslant y^*(A_2(u)).$$

Note that (see [3], the form of the linear functional on $B_p(X; [0, T])$

$$y^*(A_2(u)) = \int_0^T [y^*(\sigma), (A_2(u))(\sigma)]\, d\sigma,$$

where

(15) $$y^*(\sigma) = S^*(T-\sigma)x^* \in X_1,$$

for every σ. The value of the functional $y^*(\sigma)$ on the element $(A_2(u)(\sigma))$ is $[y^*(\sigma), (A_2(u))(\sigma)]$.

Then (14) is of the form

$$\int_0^T [y^*(\sigma), (A_2 u)(\sigma)]\, d\sigma \leqslant \int_0^T [y^*(\sigma), (A_2 u_0)(\sigma)]\, d\sigma \quad \text{for } u \in C.,$$

that is

(*) $$\int_0^T [y^*(\sigma), B'_u(\sigma, u_0(\sigma))u(\sigma)]\, d\sigma \leqslant \int_0^T [y^*(\sigma), B'_u((\sigma, u_0(\sigma))u_0(\sigma)]\, d\sigma.$$

From (15) we know that $y^*(\sigma)$ is continuous and that for every $x \in X_1$ we have

$$y^*(\sigma)(x) = x^*(S(T-\sigma)x)$$

and

$$\frac{d}{d\sigma} y^*(\sigma)(x) = -x^*[AS(T-\sigma)x] = -A^*x^*[S(T-\sigma)x] = -A^*y^*(\sigma)x$$

for $0 \leqslant \sigma \leqslant T$ and $x \in D(A)$.

Thus formula (*) may be considered as a *maximum principle.*

Note that inequality (14) may also be written in the form

(16) $$\int_0^T [z^*(\sigma), u(\sigma)]\, d\sigma \leqslant \int_0^T [z^*(\sigma), u_0(\sigma)]\, d\sigma \quad \text{for } u \in C,$$

where

$$z^*(\sigma) = \{B'_u(\sigma, u_0(\sigma))\}^* S^*(T-\sigma)x^*.$$

Besides, if we assume that $(\sigma, u) \to B'_u(\sigma, u)$ is a continuous mapping of $[0, T] \times X_2$ into $\mathscr{L}(X_2, X_1)$ with a topology of operator norms, then if the function u_0 is bounded and strongly measurable, the function z^* is also bounded and strongly measurable. In particular, by these assumptions of regularity of B'_u it follows from (16) that if C is the set of all strongly measurable functions with values in X_2, defined on $[0, T]$ and satisfying the inequality

$$\operatorname*{ess\,sup}_{\sigma \in [0,T]} \|u(\sigma)\| \leqslant M,$$

and in the set C so defined there exists an optimal control u_0, then $\|u(\sigma)\| = M$ for almost all $\sigma \in [0, T]$.

Note that Pontriagin's maximum principle can be proved in the same way for the minimal-time problem, when the set S which we want to attain is any convex closed set with an interior.

I wish to express my gratitude to Professor S. Rolewicz for his valuable help when this work has been written.

References

[1] A. V. Balakrishnan, *Optimal control problems in Banach spaces*, J. SIAM Control, Ser. A, 3. I (1965).

[2] N. Dunford and J. T. Schwartz, *Linear operators*, New York — London 1958.

[3] E. Hille and R. S. Phillips, *Functional analysis and semigroups*, Amer. Math. Coll. Publ., Providence 1957.

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

[5] Н. Я. Виленкин и Е. А. Горин, *Функциональный анализ*, Москва 1964.

Reçu par la Rédaction le 30. 6. 1968