DEMONSTRATIO MATHEMATICA
Vol. XIV No 4 1981

Zygfryd Kominek

ON THE FUNCTIONAL EQUATION $\varphi(x)=h(x, \varphi[f(x)])$

1. We shall consider the system of functional equations

$$\varphi_i(x) = h_i(x,\varphi_1[f(x)],\dots,\varphi_m[f(x)]), \quad i=1,\dots,m, \tag{1}$$

where the functions h_i and f of the type $R^{m+1} \longrightarrow R$ and $R^n \longrightarrow R^n$, respectively, are given and φ_i are unknown functions. The fundamental theorems regarding the uniqueness and the existence of solutions of the class C^r in the case m=1 are due to B.Choczewski ([1], [2]). This theory has been further extended by J.Matkowski [6]. Our theorem (see §3) generalizes also some result of the author obtained in the case of functions of real variable [4]. On the other hand, the system (1) may be treated as a generalization of Schroeder's equation. Therefore the results of this paper correspond to others contained in [3], [5] and [8].

Let $[\alpha^i_k]$, $[\beta^s_j]$, $i,k=1,\dots,m$, $s,j=1,\dots,n$ be arbitrary real matrices. By the right Kronecker product of the matrices $[\alpha^i_k]$ and $[\beta^s_j]$ we mean the matrix

$$[\alpha^i_k]\times[\beta^s_j] := \begin{bmatrix} \alpha^1_1[\beta^s_j],\dots,\alpha^1_n[\beta^s_j] \\ \vdots \qquad\qquad \vdots \\ \vdots \qquad\qquad \vdots \\ \alpha^n_1[\beta^s_j],\dots,\alpha^n_n[\beta^s_j] \end{bmatrix}. \tag{2}$$

It is well known that if $\varrho_1,\dots,\varrho_m$ and $\mu_1,\dots,\mu_n$ denote the characteristic roots of the matrices $[\alpha^i_k]$ and $[\beta^s_j]$, respectively, then $\varrho_i\mu_s$, $i=1,\dots,m$, $s=1,\dots,n$ are the characteristic roots of the matrix (2).

L e m m a 1. If $\alpha^i_k>0$, $\beta^s_j>0$, $i,k=1,\dots,m$, $s,j=1,\dots,n$ and $|\varrho_i||\mu_s|^r<1$ (r is a positive integer constant), then there exists a system of numbers $\vartheta^k_{s_1,\dots,s_r}>0$, $k=1,\dots,m$, $s_1,\dots,s_r=1,\dots,n$ satisfying the following system of inequalities

$$(3)\qquad \sum_{k=1}^{m}\sum_{s_1=1}^{n}\dots\sum_{s_r=1}^{n}\alpha^i_k\beta^{s_1}_{j_1}\dots\beta^{s_r}_{j_r}\vartheta^k_{s_1,\dots,s_r}<\vartheta^i_{j_1,\dots,j_r},$$

$i=1,\dots,m$, $j_1,\dots,j_r=1,\dots,n$.

P r o o f . This follows from Lemma 1.2 in [7], because the modulus of every characteristic root of the matrix $[\alpha^i_k]\times\underbrace{[\beta^s_j]\times\dots\times[\beta^s_j]}_{r\text{-times}}$ is less than one.

In the space R^p we introduce the norm

$$(4)\qquad \|x\|=\sum_{q=1}^{p}|x_q|,\quad x=(x_1,\dots,x_p).$$

We say that the function h defined on $G\times H$, $G\subset R^n$, $H\subset R^m$ with values in R^m is of the class C^r in $G\times H$, iff there exist a function $\tilde h$ and open sets $\tilde G$ and $\tilde H$, $\subset\tilde G\subset\tilde G$, $H\subset\tilde H$ such that $\tilde h$ is of the class C^r in $\tilde G\times\tilde H$ and the restriction $\tilde h$ to the set $G\times H$ is equal to h.

Let $f\colon G\subset R^n\to R^n$, $h\colon G\times H\subset R^{n+m}\to R^m$ be of the class C^r in G and $G\times H$, respectively. We denote $f=(f_1,\dots,f_n)$, $h=(h_1,\dots,h_m)$ and $y^\sigma=(y_{1,\underbrace{1,\dots,1}_{\sigma}},\dots,y_{1,\underbrace{1,\dots,n}_{\sigma}},\dots$ $\dots,y_{m,\underbrace{n,\dots,n,1}_{\sigma}},\dots,y_{m,\underbrace{n,\dots,n}_{\sigma}})\in R^{mn^\sigma}$ for $\sigma=1,\dots,r$. In the sequel, the following sequences of functions will be very useful:

$$(5)\quad \begin{cases} h^1_{i,j_1}(x,y,y^1) := \dfrac{\partial h_i}{\partial x_{j_1}}(x,y) + \displaystyle\sum_{k=1}^{m}\sum_{s_1=1}^{n} \dfrac{\partial h_i}{\partial y_k}(x,y)\, y^1_{k,s_1} \dfrac{\partial f_{s_1}}{\partial x_{j_1}}(x), \\ h^{\sigma}_{i,j_1,\dots,j_\sigma}(x,y,y^1,\dots,y^{\sigma}) := \dfrac{\partial h^{\sigma-1}_{i,j_1,\dots,j_{\sigma-1}}}{\partial x_{j_\sigma}}(x,y,y^1,\dots,y^{\sigma-1}) + \\ + \displaystyle\sum_{k=1}^{m}\sum_{s_1=1}^{n} \dfrac{\partial h^{\sigma-1}_{i,j_1,\dots,j_{\sigma-1}}}{\partial y_k}(x,y,y^1,\dots,y^{\sigma-1}) y^1_{k,s_1} \dfrac{\partial f_{s_1}}{\partial x_{j_\sigma}}(x) + \\ + \displaystyle\sum_{\nu=1}^{\sigma-1}\sum_{k=1}^{n}\sum_{s_1,\dots,s_\nu,s_{\nu+1}=1}^{n} \dfrac{\partial h^{\sigma-1}_{i,j_1,\dots,j_{\sigma-1}}}{\partial y^{\nu}_{k,s_1,\dots,s_\nu}}(x,y,y^1,\dots,y^{\sigma-1}) \cdot \\ \cdot\, y^{\nu+1}_{k,s_1,\dots,s_\nu,s_{\nu+1}} \dfrac{\partial f_{s_{\nu+1}}}{\partial x_{j_\sigma}}(x), \\ j_1,\dots,j_\sigma = 1,\dots,n,\quad \sigma = 2,\dots,r,\quad x\in G,\quad y\in R^{mn^{\sigma}}. \end{cases}$$

We have the following simple lemma.

Lemma 2. If $f\colon G\subset R^n \longrightarrow R^n$, $h\colon G\times H\subset R^{m+n} \longrightarrow R^m$ are of the class C^r, then the functions $h_{i,j_1,\dots,j_\sigma}$, $i=1,\dots,m$, $\sigma=1,\dots,r$, $j_1,\dots,j_\sigma=1,\dots,n$ are of class $C^{r-\sigma}$.

Now, we suppose that γ is a function defined and of the class C^r in a neighbourhood $G\subset R^n$ of zero and with the values in R^m and let

$$\gamma[f(x)] \in H \quad \text{for every } x\in G.$$

Lemma 3. Let f and h satisfy the assumptions of lemma 2. The function

(6) $$\psi(x) := h(x,\gamma[f(x)]), \quad x \in G$$

is of the class C^r in G and also we have

$$\frac{\partial^{(\sigma)}\psi_i}{\partial x_{j_1}\dots\partial x_{j_\sigma}}(x) = h^{\sigma}_{i,j_1,\dots,j_\sigma}(x,\gamma[f(x)],\dots,\gamma^{(\sigma)}[f(x)]),\ x\in G,$$

$i=1,\dots,m$, $j_1,\dots,j_\sigma=1,\dots,n$, $\sigma=1,\dots,r$.

P r o o f . Differentiating (6) with respect of the variable x_{j_1} we obtain

$$\frac{\partial\psi_i}{\partial x_{j_1}}(x) = \frac{\partial h_i}{\partial x_{j_1}}(x,\gamma[f(x)]) + \sum_{k=1}^{m}\sum_{s_1=1}^{n}\frac{\partial h_i}{\partial y_k}(x,\gamma[f(x)])\frac{\partial\gamma_k}{\partial z_{s_1}}(f(x))\frac{\partial f_{s_1}}{\partial x_{j_1}}(x) =$$

$$= h^1_{i,j_1}(x,\gamma[f(x)],\ \gamma'[f(x)]),\quad i=1,\dots,m,\ j_1=1,\dots,n,$$

which proves our assertion in the case $\sigma=1$. Assuming that

$$\frac{\delta^{(\sigma-1)}\psi_i}{\partial x_{j_1}\dots\partial x_{j_{\sigma-1}}}(x) = h^{\sigma-1}_{i,j_1,\dots,j_{\sigma-1}}(x,\gamma[f(x)],\dots,\gamma^{(\sigma-1)}[f(x)]),$$

$x\in G$, $\sigma\le r$, $i=1,\dots,m$, $j_1,\dots,j_{\sigma-1}=1,\dots,n$, similarly as above, we have

$$\frac{\partial^{(\sigma)}\psi_i}{\partial x_{j_1}\dots\partial x_{j_\sigma}}(x) = \frac{\partial h^{\sigma-1}_{i,j_1,\dots,j_{\sigma-1}}}{\partial x_{j_\sigma}}(x,\gamma[f(x)],\dots$$

$$\dots,\gamma^{(\sigma-1)}[f(x)]) + \sum_{k=1}^{m}\sum_{s_1=1}^{n}\frac{\partial h^{\sigma-1}_{i,j_1,\dots,j_{\sigma-1}}}{\partial y_k}(x,\gamma[f(x)],\dots$$

$$\dots,\gamma^{(\sigma-1)}[f(x)])\frac{\partial\gamma_k}{\partial z_{s_1}}[f(x)]\frac{\partial f_{s_1}}{\partial x_{j_\sigma}}(x) +$$

$$+\sum_{\nu=1}^{\sigma-1}\sum_{k=1}^{m}\sum_{s_1,\dots,s_\nu=1}^{n}\sum_{s_{\nu+1}=1}^{n}\frac{\partial h^{\sigma-1}_{i,j_1,\dots,j_{\sigma-1}}}{\partial y^{\nu}_{k,s_1,\dots,s_\nu}}(x,\gamma[f(x)],\dots$$

$$\ldots,\gamma^{(\sigma-1)}[f(x)])\cdot\frac{\partial^{(\nu+1)}\gamma_k}{z_{s_1}\ldots z_{s_{\nu+1}}}(f(x))\,\frac{\partial f_{s_{\nu+1}}}{\partial x_{j\sigma}}(x)=$$

$$=h^{\sigma}_{i,j_1,\ldots,j\sigma}(x,\gamma[f(x)],\ldots,\gamma^{(\sigma)}[f(x)]).$$

Thus the proof of this Lemma is finished.

L e m m a 4. The functions $h^{\sigma}_{i;j_1,\ldots,j\sigma}$ defined by (5) can be written in the form

$$(7)\quad h^{\sigma}_{i,j_1,\ldots,j\sigma}(x,y,y^1,\ldots,y^{\sigma})=Z^{\sigma}_{i,j_1,\ldots,j\sigma}(x,y,y^1,\ldots,y^{\sigma-1})+$$
$$+Q^{\sigma}_{i,j_1,\ldots,j\sigma}(x,y,y^{\sigma}),$$

where $Z^{\sigma}_{i,j_1,\ldots,j\sigma}$ are of the class $C^{r-\sigma}$ and

$$Q^{\sigma}_{i,j_1,\ldots,j\sigma}(x,y,y^{\sigma})=\sum_{k=1}^{m}\sum_{s_1=1}^{n}\ldots\sum_{s_\sigma=1}^{n}\frac{\partial h_i}{\partial y_k}(x,y)y^{\sigma}_{k,s_1,\ldots,s_\sigma}\cdot$$
$$\cdot\frac{\partial f_{s_1}}{\partial x_{j_1}}(x)\ldots\frac{\partial f_{s_\sigma}}{\partial x_{j\sigma}}(x).$$

We omit a simple proof of this lemma.

2. Necessary conditions

We assume that

(i) the function f of the type R^n into R^n is defined and of the class C^r in some neighbourhood $G\subset R^n$ of zero and for every neighbourhood $\tilde{G}_1\subset G$ of zero there exists a neighbourhood $G_1\subset\tilde{G}_1$ of zero such that $f(G_1)\subset G_1$;

(ii) the function h of the type $R^n\times R^m$ into R^m is defined and of the class C^r in the set $G\times H$, where H is open, $0\in H$ and $h(0,0)=0$.

From Lemmas 3 and 4 we have

L e m m a 5. If (i), (ii) are fulfilled and $\varphi:G\to H$ is a solution of the equation $\varphi(x)=h(x,\varphi[f(x)])$ in G fulfilling the condition $\varphi(0)=0$, then the numbers

(8) $$\eta^{\sigma}_{i,j_1,\ldots,j_\sigma} := \frac{\partial^{(\sigma)}\varphi_i}{\partial x_{j_1}\ldots\partial x_{j_\sigma}}(0),\ \sigma=1,\ldots,r$$

satisfy the system of equations

(9) $$\eta^{\sigma}_{i,j_1,\ldots,j_\sigma} = Z^{\sigma}_{i,j_1,\ldots,j_\sigma}(0,0,\eta^1,\ldots,\eta^{\sigma-1}) + \\ + Q_{i,j_1,\ldots,j_\sigma}(0,0,\eta^{\sigma}),$$

$i=1,\ldots,m$, $j_1,\ldots,j_\sigma=1,\ldots,n$, where

(10) $$\eta^{\nu} := (\eta^{\nu}_{1,1,\ldots,1},\ldots,\eta^{\nu}_{1,1,\ldots,n},\ldots,\eta^{\nu}_{m,n,\ldots,n,1},\ldots \\ \ldots,\eta^{\nu}_{m,n,\ldots,n})$$

$\nu=1,\ldots,\sigma$. Moreover, for every permutations $(j_{\alpha_1},\ldots,j_{\alpha_\sigma})$ of the system $(j_1,\ldots,j_\sigma)$ the equalities

(11) $$\eta^{\sigma}_{i,j_1,\ldots,j_\sigma} = \eta^{\sigma}_{i,j_{\alpha_1},\ldots,j_{\alpha_\sigma}}$$

hold.

The existence and the uniqueness of the numbers $\eta^{\sigma}_{i,j_1,\ldots,j_\sigma}$ satisfying the conditions (9) result from the following assumptions: $\rho_i\mu_{s_1}\ldots\mu_{s_\sigma} \neq 1$, $i=1,\ldots,m$, $\sigma=1,\ldots,r$, $s_1,\ldots,s_\sigma=1,\ldots,n$, where ρ_i denote the characteristic roots of the matrix $\left[\frac{\partial h}{\partial y}(0,0)\right]$ and $\mu_1,\ldots,\mu_n$ - the characteristic roots of the matrix $\left[\frac{\partial f}{\partial x}(0)\right]$. This follows from Cramer's theorem, because the system (9) can be written in the form

$$\eta^{\sigma} = Z^{\sigma}(0,0,\eta^1,\ldots,\eta^{\sigma-1}) + \Big\{\left[\frac{\partial h}{\partial y}(0,0)\right]\times \\ \times \underbrace{\left[\frac{\partial f}{\partial x}(0)\right]^T \times \ldots \left[\frac{\partial f}{\partial x}(0)\right]^T}_{\sigma\text{-times}}\Big\}\times \eta^{\sigma},$$

$\sigma=1,\dots,r$, where the symbol $[\cdot]^T$ denotes the transpose of $[\cdot]$. In the sequel the system of numbers satisfying conditions (9) and (11) will be called admissible (see [3]).

Now, we suppose that $\varrho^{\sigma}_{k,j_1,\dots,j_\sigma}$ is an admissible system. Without loss of generality, similarly as in [2], we can assume that $\varrho^{\sigma}_{k,j_1,\dots,j_\sigma} = 0$, $\sigma=1,\dots,r$, $k=1,\dots,n$, $j_1,\dots \dots j_r=1,\dots,n$. By (9) and Lemma 4 we get

(12) $$Z^{\sigma}_{i,j_1,\dots,j_\sigma}(0,0,\dots,0) = 0$$

and

(13) $$h^{\sigma}_{i,j_1,\dots,j_\sigma}(0,0,\dots,0) = 0,$$

$i=1,\dots,m$, $j_1,\dots,j_\sigma=1,\dots,n$, $\sigma=1,\dots,r$.

3. The existence of a solution of the class C^r

Put

$$\alpha^i_k := \left|\frac{\partial h_i}{\partial y_k}(0,0)\right|, \ \beta^s_j := \left|\frac{\partial f_s}{\partial x_j}(0)\right|, \ i,k=1,\dots,m, \ s,j=1,\dots,n,$$

and let, as above, $\varrho_1,\dots,\varrho_m$ and $\mu_1,\dots,\mu_n$ denote the characteristic roots of the matrices $[\alpha^i_k]$ and $[\beta^s_j]$, respectively.

T h e o r e m . If (i), (ii), (12) and (13) hold and for $i=1,\dots,m$, $j=1,\dots,n$

(15) $$|\varrho_i|\,|\mu_j|^r < 1,$$

then there exists a solution $\varphi = (\varphi_1,\dots,\varphi_m)$ of the class C^r of the functional equation

$$\varphi(x) = h(x,\varphi[f(x)]),$$

in a neighbourhood of zero such that $\varphi_i(0) = 0$ and $\frac{\partial^{(\sigma)}\varphi_i}{\partial x_{j_1}\ldots\partial x_{j_\sigma}}(0) = 0$ for $i=1,\ldots,m$, $j_1,\ldots,j_\sigma=1,\ldots,n$, $\sigma=1,\ldots,r$.

P r o o f . On account of (i), (ii), (14), (15) and Lemma 1 there exist neighbourhoods $U_1\subset R^n$, $V_1\subset R^m$, $0\in U_1\subset \subset \bar{U}_1\subset G$, $0\in V_1\subset\bar{V}_1\subset H$ and a constant $\varepsilon>0$ such that

$$\text{(16)}\qquad \sum_{k=1}^{m}\sum_{s_1=1}^{n}\cdots\sum_{s_r=1}^{n} a_k^i\, b_{j_1}^{s_1}\cdots b_{j_r}^{s_r}\,\vartheta^k_{s_1,\ldots,s_r} < \vartheta^i_{j_1,\ldots,j_r},\quad i=1,\ldots,m,\ j_1,\ldots,j_r=1,\ldots,n,$$

where

$$\text{(17)}\qquad a_k^i := \sup\left\{\left|\frac{\partial h_i}{\partial y_k}(x,y)\right|+\varepsilon;\ (x,y)\in\bar{U}_1\times\bar{V}_1\right\},$$

$$\text{(18)}\qquad b_j^s := \sup\left\{\left|\frac{\partial f_s}{\partial x_j}(x)\right|+\varepsilon;\ x\in\bar{U}_1\right\}.$$

The continuity of $Z^r_{i,j_1,\ldots,j_r}$ and (12) imply the existence of neighbourhoods U_2 and V_2, $0\in U_2\subset\bar{U}_2\subset U_1$, $0\in V_2\subset\bar{V}_2\subset V_1$ such that for all $x\in\bar{U}_2$, $y\in\bar{V}_2$, $y^\sigma\in\bar{V}_2^{n^\sigma} := \underbrace{\bar{V}_2\times\ldots\times\bar{V}_2}_{n^\sigma\text{-times}}$, $\sigma=1,\ldots,r-1$ the inequalities

$$\text{(19)}\qquad \left|Z^r_{i,j_1,\ldots,j_r}(x,y,y^1,\ldots,y^{r-1})\right| \le \vartheta^i_{j_1,\ldots,j_r} - \sum_{k=1}^{m}\sum_{s_1=1}^{n}\cdots\sum_{s_r=1}^{n} a_k^i\, b_{j_1}^{s_1}\cdots b_{j_r}^{s_r}\,\vartheta^k_{s_1,\ldots,s_r}$$

hold.

Of course, we may assume that

(20) $$\left\{ y \in R^m;\ \|y\| \leqslant \sup\left\{ \|x\|;\ x \in \bar{U}_2 \right\} \right\} \subset V_2$$

as well as

(21) $$\begin{cases} \displaystyle\sum_{k=1}^{m} \sum_{s_1=1}^{n} \cdots \sum_{s_r=1}^{n} \vartheta^k_{s_1,\ldots,s_r} < 1, \\ \displaystyle\sum_{k=1}^{m} \sum_{s_1=1}^{n} \cdots \sum_{s_r=1}^{n} \vartheta^k_{s_1,\ldots,s_r} \sum_{s=1}^{n} \sum_{j=1}^{n} b^s_j < 1, \end{cases}$$

because the system (16) is homogeneous.

Let $\varepsilon^i_{j_1,\ldots,j_r} > 0$, $i=1,\ldots,m$, $j_1,\ldots,j_r=1,\ldots,n$ be arbitrary numbers fulfilling the system of the inequalities

(22) $$\sum_{k=1}^{m} \sum_{s_1=1}^{n} \cdots \sum_{s_r=1}^{n} a^i_k\, b^{s_1}_{j_1} \cdots b^{s_r}_{j_r}\, \varepsilon^k_{s_1,\ldots,s_r} < \varepsilon^i_{j_1,\ldots,j_r}.$$

The functions $Z^r_{i,j_1,\ldots,j_r}$ and $\frac{\partial h_i}{\partial y_k}$ $\frac{\partial f_{s_1}}{\partial x_{j_1}} \cdots \frac{\partial f_{s_r}}{\partial x_{j_r}}$ are uniformly continuous on the sets $\bar{U}_2 \times \bar{V}_2 \times \bar{V}_2^n \times \ldots \times \bar{V}_2^{n^{r-1}}$ and $\bar{U}_2 \times \bar{V}_2$ respectively. Hence there exists a $\delta > 0$ (depending on $\varepsilon^i_{j_1,\ldots,j_r}$) such that for all $x, \bar{x} \in \bar{U}_2$, $y, \bar{y} \in \bar{V}_2$, $y^\sigma, \bar{y}^\sigma \in \bar{V}_2^{n^\sigma}$ satisfying the inequalities $\|x-\bar{x}\| \leqslant \delta$, $\|y-\bar{y}\| \leqslant \delta$, $\|y^\sigma - \bar{y}^\sigma\| \leqslant \delta$ we have

$$
(23)\begin{cases}
\left|Z^r_{i,j_1,\dots,j_r}(x,y,y^1,\dots,y^{r-1})-Z^r_{i,j_1,\dots,j_r}(\bar{x},\bar{y},\bar{y}^1,\dots,\bar{y}^{r-1})\right|\leq \\
\leq \frac{1}{2}\left[\varepsilon^i_{j_1,\dots,j_r}-\sum_{k=1}^{m}\sum_{s_1=1}^{n}\dots\sum_{s_r=1}^{n} a^i_k\, b^{s_1}_{j_1}\dots b^{s_r}_{j_r}\,\varepsilon^k_{s_1,\dots,s_r}\right], \\
\sum_{k=1}^{m}\sum_{s_1=1}^{n}\dots\sum_{s_r=1}^{n}\vartheta^k_{s_1,\dots,s_r}\left|\frac{\partial h_i}{\partial y_k}(x,y)\frac{\partial f_{s_1}}{\partial x_{j_1}}(x)\dots\frac{\partial f_{s_r}}{\partial x_{j_r}}(x)-\right. \\
\left.-\frac{\partial h_i}{\partial y_k}(\bar{x},\bar{y})\frac{\partial f_{s_1}}{\partial x_{j_1}}(\bar{x})\dots\frac{\partial f_{s_r}}{\partial x_{j_r}}(\bar{x})\right|\leq\frac{1}{2}\left[\varepsilon^i_{j_1,\dots,j_r}-\right. \\
\left.-\sum_{k=1}^{m}\sum_{s_1=1}^{n}\dots\sum_{s_r=1}^{n} a^i_k\, b^{s_1}_{j_1}\dots b^{s_r}_{j_r}\,\varepsilon^k_{s_1,\dots,s_r}\right].
\end{cases}
$$

Let $U\subset R^n$ be a neighbourhood of zero such that

$$
(24)\qquad f(\bar{U})\subset\bar{U}\subset\left\{x\in R^n;\ \|x\|\leq 1\right\}
$$

and let X_i, $i=1,\dots,m$ denote the family of all real functions φ_i of the class C^r in U fulfilling the conditions:

$$
(25)\qquad \varphi_i(0)=0;
$$

$$
(26)\qquad \frac{\partial^{(\sigma)}\varphi_i}{\partial x_{j_1}\dots\partial x_{j_\sigma}}(0)=0,\ i=1,\dots,m,\ j_1,\dots,j_\sigma=1,\dots,n,\ \sigma=1,\dots,r;
$$

$$
(27)\qquad \left|\frac{\partial^{(r)}\varphi_i}{\partial x_{j_1}\dots\partial x_{j_r}}(x)\right|\leq\vartheta^i_{j_1,\dots,j_r},\ i=1,\dots,m,\ j_1,\dots,j_r=1,\dots,n;
$$

(28) if $\varepsilon^i_{j_1,\dots,j_r}>0$, $i=1,\dots,m$, $j_1,\dots,j_r=1,\dots,n$ fulfil (22)

and $x, \bar{x} \in \bar{U}$, $\|x-\bar{x}\| \leq \delta$, where δ is chosen so that (23) holds, then

$$\left| \frac{\partial^{(r)}\varphi_i}{\partial x_{j_1}\ldots\partial x_{j_r}}(x) - \frac{\partial^{(r)}\varphi_i}{\partial x_{j_1}\ldots\partial x_{j_r}}(\bar{x}) \right| \leq \varepsilon^i_{j_1,\ldots,j_r}.$$

In the vector spaces X_i (with the usual operations "+" and "·") we define the norm

$$(29) \quad \|\varphi_i\| := \sum_{j_1,\ldots,j_r=1}^{n} \sup\left\{ \left| \frac{\partial^{(r)}\varphi_i}{\partial x_{j_1}\ldots\partial x_{j_r}}(x) \right| ; \ x \in \bar{U} \right\}$$

and in the space $X := X_1 \times \ldots \times X_m$ we put

$$(30) \qquad \|\varphi\| := \sum_{i=1}^{m} \|\varphi_i\|.$$

Note that X is a convex and compact subset of the space of all functions defined and of the class C^r in $\bar{U}$ with the values in R^m (the compactness of X follows from the conditions (25)-(28) and the theorem of Arzela). For $\varphi = (\varphi_1,\ldots,\varphi_m) \in X$ we put

$$(31) \qquad T_i(\varphi)(x) := h_i(x,\varphi[f(x)]), \ i=1,\ldots,m,$$

$$(32) \qquad T(\varphi)(x) := (T_1(\varphi)(x),\ldots,T_m(\varphi)(x)).$$

It follows from (25), (26), (i), (ii) and Lemmas 3 and 4 that for $\varphi \in X$ we have

$$(33) \ T_i(\varphi)(0)=0, \ \frac{\partial^{(\sigma)}T_i}{\partial x_{j_1}\ldots\partial x_{j_\sigma}}(0)=0, \ i=1,\ldots,m, j_1,\ldots,j_\sigma=1,\ldots,n$$

$$\sigma=1,\ldots,r.$$

By (24) and the mean-value theorem we get

$$\|\varphi[f(x)]\| = \sum_{k=1}^{m} \|\varphi_k[f(x)]\| \leqslant \sum_{k=1}^{m} \sup\left\{\|\varphi_k(x)\| ;\ x \in \overline{U}\right\} \leqslant$$

$$\leqslant \sum_{k=1}^{m} \sum_{j_1=1}^{n} \sup\left\{\left|\frac{\partial \varphi_k}{\partial x_{j_1}}(x)\right| \|x\| ;\ x \in \overline{U}\right\} \leqslant \ldots \leqslant$$

$$\leqslant \sum_{k=1}^{m} \sum_{j_1,\ldots,j_r=1}^{n} \sup\left\{\left|\frac{\partial^{(r)}\varphi_k}{\partial x_{j_1}\ldots\partial x_{j_r}}(x)\right| \|x\|^r ;\ x \in \overline{U}\right\}$$

whence, in virtue of (27), (24) and (21),

$$\|\varphi[f(x)]\| \leqslant \|x\| \tag{34}$$

which implies that

$$\varphi[f(x)] \in \overline{V}_2. \tag{35}$$

Moreover, similarly as above, we can prove that

$$\varphi^{(\sigma)}[f(x)] \in \overline{V}_2^{\,n\sigma}, \quad \sigma = 1,\ldots,r-1. \tag{36}$$

Differentiating the equality (31), along the same lines as in Lemma 3, we get

$$\frac{\partial^{(r)}T_i}{\partial x_{j_1},\ldots,\partial x_{j_r}}(x) = h^r_{i,j_1,\ldots,j_r}\left(x,\varphi[f(x)],\ldots,\varphi^{(r)}[f(x)]\right) \tag{37}$$

and using Lemma 4, (35), (36) and (19), (17), (18), (27) we have

$$\left|\frac{\partial^{(r)}T_i(\varphi)}{\partial x_{j_1}\dots\partial x_{j_r}}(x)\right| \leqslant \nu^i_{j_1,\dots,j_r}. \tag{38}$$

Suppose that the antecedent of the condition (28) is fulfilled. It follows from (37) and Lemma 4 that

$$\left|\frac{\partial^{(r)}T_i}{\partial x_{j_1}\dots\partial x_{j_r}}(x) - \frac{\partial^{(r)}T_i}{\partial x_{j_1}\dots\partial x_{j_r}}(\bar{x})\right| \leqslant \tag{39}$$

$$\leqslant \left|Z^r_{i,j_1,\dots,j_r}(x,\varphi[f(x)],\dots,\varphi^{(r-1)}[f(x)]) - \right.$$

$$\left. - Z^r_{i,j_1,\dots,j_r}(\bar{x},\varphi[f(\bar{x})],\dots,\varphi^{(r-1)}[f(\bar{x})])\right| +$$

$$+ \sum_{k=1}^{m}\sum_{s_1,\dots,s_r=1}^{n}\left(\left|\frac{\partial h_i}{\partial y_k}(x,\varphi[f(x)])\right| \left|\frac{\partial f_{s_1}}{\partial x_{j_1}}(x)\dots\frac{\partial f_{s_r}}{\partial x_{j_r}}(x)\right| \cdot\right.$$

$$\cdot\left|\frac{\partial^{(r)}\varphi_k}{\partial z_{s_1}\dots\partial z_{s_r}}[f(x)] - \frac{\partial^{(r)}\varphi_k}{\partial z_{s_1}\dots\partial z_{s_r}}[f(\bar{x})]\right| + \left|\frac{\partial^{(r)}\varphi_k}{\partial z_{s_1}\dots\partial z_{s_r}}[f(\bar{x})]\right| \cdot$$

$$\cdot\left|\frac{\partial f_{s_1}}{\partial x_{j_1}}(\bar{x})\dots\frac{\partial f_{s_r}}{\partial x_{j_r}}(\bar{x})\frac{\partial h_i}{\partial y_k}(\bar{x},\varphi[f(\bar{x})]) - \right.$$

$$\left.\left. - \frac{\partial h_i}{\partial y_k}(x,\varphi[f(x)])\frac{\partial f_{s_1}}{\partial x_{j_1}}(x)\dots\frac{\partial f_{s_r}}{\partial x_{j_r}}(x)\right|\right).$$

Putting $y = f(x)$ and $\bar{y} = f(\bar{x})$ and applying a reasoning similar to that we have used to obtain the condition (35), we get

$$\| \varphi^{(\sigma)}(y) - \varphi^{(\sigma)}(\bar{y}) \| \leq \| y - \bar{y} \| \sum_{k=1}^{m} \overline{\sum_{s_1,\ldots,s_r=1}^{n}} \vartheta^{k}_{s_1,\ldots,s_r}$$

$$\sigma = 0,1,\ldots,r-1.$$

However,

$$\| y - \bar{y} \| = \| f(x) - f(\bar{x}) \| \leq \sum_{s=1}^{n} \sum_{j=1}^{n} b^{s}_{j} \| x - \bar{x} \|$$

whence, in virtue of (21),

$$\left\| \varphi^{(\sigma)}\left[f(x)\right] - \varphi^{(\sigma)}\left[f(\bar{x})\right] \right\| \leq \| x - \bar{x} \| \leq \delta , \quad \sigma = 0,1,\ldots,r-1$$

and, on account of (39), (22), (17), (18) and (27) we infer that

$$(40) \qquad \left| \frac{\partial^{(r)} T_i(\varphi)}{\partial x_{j_1} \ldots \partial x_{j_r}}(x) - \frac{\partial^{(r)} T_i(\varphi)}{\partial x_{j_1} \ldots \partial x_{j_r}}(\bar{x}) \right| \leq \varepsilon^{i}_{j_1,\ldots,j_r}.$$

Conditions (33) (38) and (40) show that the transformation T maps the space X into itself. It is not difficult to check that T is continuous. Now, the assertion of our theorem results from Schauder's principle.

REFERENCES

[1] B. Choczewski: On differentiable solutions of a functional equation, Ann. Polon. Math. 13 (1963) 133-138.

[2] B. Choczewski: Investigation of the existence and uniqueness of differentiable solutions of a functional equation, ibidem 15 (1964) 117-141.

[3] Ph. H a r t m a n : On local homeomorphisms of Euclidean spaces, Bol. Soc. Math. Mexicana, (2) 5 (1960), 220-241.

[4] Z. K o m i n e k : Some theorems on differentiable solutions of a system of functional equations of n-th order, Ann. Polon. Math. 37 (1978) 71-91.

[5] M. K u c z m a : Note on linearlization, ibidem 29 (1974) 75-81.

[6] J. M a t k o w s k i : On the continuous dependences of C^r solutions of a functional equation on the given functions, Aequationes Math. 6 (1971) 215-227.

[7] J. M a t k o w s k i : Integrable solutions of functional equations, Diss. Math. 127 (1975) 1-68.

[8] S. S t e r n b e r g : Local contractions and a theorem of Poincare, Amer. J. Math. 79 (1957) 809-824.

INSTITUTE OF MATHEMATICS, SILESIAN UNIVERSITY, KATOWICE
Received May 21, 1980.

DEMONSTRATIO MATHEMATICA
Vol. XIV No 4 1981

Grażyna Pankiewicz

ON THE UNIQUENESS AND THE NUMBER OF THE SWITCHING POINTS OF OPTIMAL CONTROL

1. Formulation of the problem

We shall consider the time-optimal problem for the linear equation

$$\frac{dx(t)}{dt} = A\,x(t) + B\,u(t) \tag{1.1}$$

with the initial value

$$x(0) = x_1, \tag{1.2}$$

where $x(\cdot)$ is a function defined in the interval $[0,T]$ with values in a Banach space X_1; $u(\cdot)$ - control - is a function defined in the interval $[0,T]$ with values in a Banach space X_2; A, B - linear operators and:

$B:X_2 \longrightarrow X_1$ a bounded operator and the image of the unit ball by the operator B has non-empty interior,

$A:X_1 \longrightarrow X_1$ a bounded operator which generates a strongly continuous semigroup $S(t)$ of linear, bounded operators, $S(t) = e^{tA}$ (see [2]).

The time-optimal problem consists in finding the minimal time of a transfer in which the final state is attained.

Let $U \subset X_2$ be a set obtained from the open, strictly convex set by cutting off the n planes, bounded and weakly closed. The functions $u(\cdot)$ such that for all t : $u(t) \in U$

will be called the admissible controls and the set of such controls we denote by $\mathcal{U}$.

Let u_1 and u_2 be the controls passing from the zero point, $0 \in X_2$, to a certain side $R \subset U$ (U has n sides). Obviously, the vector $w = u_2 - u_1$ is placed on this side.

We shall assume that the generalized principle of generality of placing is satisfied. It consists in the fact that the system

$$Bw, ABw, A^2Bw, \dots, A^nBw, \dots \tag{1.3}$$

where w is any control placing on any side defined above, is a complete system in Banach space X_1 (i.e. the linear space spanned by these elements is dense in the space X_1). This condition has been called a generalized principle of generality of placing with respect to the infinite dimension of the space X_1, in contradistinction to the classical principle of the generality of placing for the finite dimensional space (see [3], 3.17).

Let $H = H(\psi(t), x(t), u)$ be Hamilton's function in Pontriagin's maximum principle. For the problem (1.1), (1.2) this function is of the form

$$H = H(\psi(t), x(t), u) = (\psi(t), Ax(t)) + (\psi(t), Bu),$$

where $\psi(t)$ is the solution of the equation adjoint to the equation (1.1)

$$\frac{d\psi(t)}{dt} = -A^* \psi(t). \tag{1.4}$$

Obviously, the function H (of the variable u) attains the maximum together with the function $(\psi(t), Bu)$, which we denote by $M(\psi)$. From Pontriagin's maximum principle, which is fulfilled for such problems (see [1]), it follows that if $u(t)$ is the optimal control transformed the system from the state x_1 to the state x_2, then there exists a solution $\psi(t)$ of the equation (1.4) such that

$$(\psi(t), Bu(t)) = M(\psi(t)). \tag{1.5}$$

Since the equation (1.4) does not contain the unknown functions $x(t)$ and $u(t)$, it is easy to find all its solutions, and next the solutions of the equation (1.5) among which there will be all optimal controls for the equation (1.1).

We come to the question: On how much uniquely does the condition (1.5) determine the control $u(t)$ by the function $\psi(t)$?

Theorem 2.1 quoted below gives the answer on this question.

2. Theorem about the number of the switching points

T h e o r e m 2.1. For each non-trivial solution $\psi(t)$ of the equation (1.4) the relation (1.5) uniquely determines the control $u(t)$.

P r o o f . The function $(\psi(t), Bu)$ is linear, so it is constant or it attains his maximum only on the edge of the set U. The same concerns to each side of the set U (remember that U has the finite number of sides). So this function attains its maximum in one vertex only of the set U or on whole side of this set. We shall show that by the completeness of the system (1.3) the last thesis (the achievement of maximum on the whole side of the set U) is possible only for the finite number of the values of t.

Let $\mathcal{T}$ denote an infinite set of values $t \in [0,T]$, for which the function $(\psi(t), Bu)$, where $u \in U$, attains its maximum on the wall R of the set U. We can find such a set $\mathcal{T}$ because the set U has a finite number of sides.

By the assumption of the strong continuity of the semigroup $S(t) = e^{tA}$, i.e. of the continuity of the function $S(t)x_0 = e^{tA}x_0$, the solution $\psi(t)$ of the equation (1.4) is of the form

$$\psi(t) = e^{-tA^*}\psi_0,$$

therefore it is analytic with respect to t (remember that A is a bounded operator).

For any $t \in \mathcal{T}$ the function $(\psi(t), Bu) = (e^{-tA^*}\psi_0, Bu)$ of the variable u is constant on the wall R. So we have

$$(e^{-tA^*}\psi_0, Bw) = (e^{-tA^*}\psi_0, B(u_2-u_1)) =$$

$$= (e^{-tA^*}\psi_0, Bu_2) - (e^{-tA^*}\psi_0, Bu_1) = 0.$$

Note that if $(e^{-tA^*}\psi_0, Bw) = 0$ for $t \in \mathcal{T}$, then from the analyticity of this expression it follows that it is identically equal to zero on the whole segment $t \in [0,T]$. So we have

$$(e^{-tA^*}\psi_0, Bw) \equiv 0.$$

Differentiating successively the above relation with respect to t and using the fact that $\psi(t) = e^{-tA^*}\psi_0$ is the solution of the equation (1.4) we obtain

$$\begin{cases} (A^*\psi(t), Bw) = 0 \\ (A^{*2}\psi(t), Bw) = 0 \\ \dots\dots\dots\dots\dots \\ (A^{*n}\psi(t), Bw) = 0 \\ \dots\dots\dots\dots\dots \end{cases}$$

i.e. by the equality $(x, Ay) = (A^*x, y)$ which is true for any x, y:

$$(2.1) \qquad \begin{cases} (\psi(t), ABw) = 0 \\ (\psi(t), A^2Bw) = 0 \\ \dots\dots\dots\dots\dots \\ (\psi(t), A^nBw) = 0 \\ \dots\dots\dots\dots\dots \end{cases}$$

By the assumption of completeness of the system $Bw, ABw, A^2Bw, \ldots$ in the Banach space X_1 the relations (2.1) denote that the vector $\psi(t)$ orthogonal to the vectors: $ABw, A^2Bw, \ldots$ is the zero-vector: $\psi(t) = 0$. This contradicts to the assumption about nontriviality of the solution $\psi(t) = e^{-tA^*}\psi_0$ of the equation (1.4). Hence there must be: $w = u_2 - u_1 = 0$. Therefore $u_2 = u_1 = u$.

Thus for all except the finite number of the values $t \in [0,T]$ the function $(\psi(t), Bu)$ attains on U the maximum only in one point, which is the vertex point of the set U (because U is the strictly convex set). Thus, by the relation

$$(\psi(t), Bu(t)) = \max_{u \in U} (\psi(t), Bu),$$

there follows the unique determination of the function $u(t)$, q.e.d.

D e f i n i t i o n 2.1. The discontinuity points of the optimal control are called switching points. Precisely, if Q is an discontinuity point of optimal control $u(t)$, and if $u(Q_-) = a_i$, $u(Q_+) = a_j$ (a_i, a_j - different points) then we say that for $t = Q$ the change-over of optimal control $u(t)$ from the point a_i to the point a_j has been achieved.

From the proof of Theorem 2.1 it follows that the points of segment $t_0 \leq t \leq t_1$, in which the control $u(t)$ is not uniquely determined, divide the interval $t_0 \leq t \leq t_1$ into a finite number of the segments.

By the analyticity of the solution $\psi(t)$ of the equation (1.4) (which follows from the proof of Theorem 2.1), the following result is true

T h e o r e m 2.2. On each finite segment of time the control - function $u(t)$ has a finite number of switching points.

Thus, Theorem 2.2 can be characterized shortly as a theorem about a finite number of switching points.

Observe that the above result can be obtained with the assumption that the set U has a countable number of sides. It is necessary only to note that the sum of the countable number of sets of the zero-measure has the zero-measure, which follows from the relations

$$0 \leq \left| \bigcup_{i=1}^{\infty} A_i \right| \leq \sum_{i=1}^{\infty} |A_i| ,$$

where $|A_i| = 0$ for each $i=1,2,\ldots$. Obviously, $|A_i|$ denotes here the measure of the set A_i.

REFERENCES

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

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

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

INSTITUTE OF MATHEMATICS, TECHNICAL UNIVERSITY OF WARSAW
Received November 20, 1980.