ANNALES
UNIVERSITATIS SCIENTIARUM BUDAPESTINENSIS DE ROLANDO EÖTVÖS NOMINATAE

SECTIO MATHEMATICA

TOMUS XXXVI.

1993

ANNALES

UNIVERSITATIS SCIENTIARUM BUDAPESTINENSIS DE ROLANDO EÖTVÖS NOMINATAE

SECTIO BIOLOGICA
incepit anno MCMLVII
SECTIO CHIMICA
incepit anno MCMLIX
SECTIO CLASSICA
incepit anno MCMXXIV
SECTIO COMPUTATORICA
incepit anno MCMLXXVIII
SECTIO GEOGRAPHICA
incepit anno MCMLXVI
SECTIO GEOLOGICA
incepit anno MCMLVII
SECTIO GEOPHYSICA ET METEOROLOGICA
incepit anno MCMLXXV
SECTIO HISTORICA
incepit anno MCMLVII
SECTIO IURIDICA
incepit anno MCMLIX
SECTIO LINGUISTICA
incepit anno MCMLXX
SECTIO MATHEMATICA
incepit anno MCMLVIII
SECTIO PAEDAGOGICA ET PSYCHOLOGICA
incepit anno MCMLXX
SECTIO PHILOLOGICA
incepit anno MCMLVII
SECTIO PHILOLOGICA HUNGARICA
incepit anno MCMLXX
SECTIO PHILOLOGICA MODERNA
incepit anno MCMLXX
SECTIO PHILOSOPHICA ET SOCIOLOGICA
incepit anno MCMLXII

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 3–4

ON ADDITIVE FUNCTIONS

By

A. BOGMÉR

Department of Analysis, Eötvös Loránd University, Budapest

(Received January 30, 1991)

Dedicated to Professor Paul Erdős on the occasion of his 80th birthday

Recently M. JOÓ [2] gave a simple nice proof for the fact that if f is real-valued number theoretical function which is monotone increasing on an arithmetical sequence $\{An+B\}$, then $f(n)=c\log n$ for each n prime to A.

The aim of the present note is to generalize this statement, namely we prove the following

THEOREM. *Let $A>0$ be any integer further B, $D(B\neq D)$ be any integers such that $A\mid B-D$ and*

(i) $(B,D)=1$ *or*

(ii) $(B,D)>1$, $(A,B)=1$

are fulfilled. Suppose f is any real-valued additive function such that

(1) $f(An+b)-f(An-D)\geq 0$ *for every natural number n.*

Then $f(n)=c\log n$ if $(n,B-D)=1$.

If f is completely additive, then independently on (i) *and* (ii) *(i.e. without these assumptions) we have*

$$f(n)=c\log n \text{ if } (n,(B-D)/(B,D))=1.$$

REMARK. For $B=D+1$ we obtain the theorem of M. JOÓ [2].

PROOF. We may suppose without loss of generality that $B>D$. Let t_0 such that $(At_0+B,B=D)=1$ and write $n+t_0$ in place of n we can write the condition (1) in the form

$$f(An+At_0+B)-f(An+At_0+D)\geq 0$$

hence writing $n(B-D)/A$ in place of n we get

$$f((B-D)n+At_0+B)-f((B-D)n+At_0+D)\geq 0$$

i.e. f is monotone increasing on the arithmetical sequence $\{(B-D)n+ +At_0+B\}$. If $(B,D)=1$ then let $t_0=0$ and if $(B,D)>1$ but $(A,B)=1$ then according to Dirichlet's theorem (there we can avoide the use of Dirichlet's theorem as in [2]) choose to so that At_0+B be a "large" prime.

Then by the method of [2] we get: $f(n)=c\log n$ if $(n,B-D)=1$. If f is totally additive, then by $t_0=0$ we have

$$f((B-D)n+B)\geq f((B-D)(n-1)+B).$$

Taking into account $((B-D)/(B-D,B),B/(B-D,B))=1$ by the result of [2] we get

$$f(n)=c\log n \text{ if } (n,(B-D)/(B,D))=1.$$

COROLLARY. *Suppose f is additive real valued and*

(2) $f(An+B)+f(An+D)$

is monotone increasing further $A\mid B-D$ and

(i) $(B,2D)=1$ or

(ii) $(B,2D)>1$, $(B,D)=1$.

Then $f(n)=c\log n$ if $(n,2(B-D)=1$.

If f is totally additive, then independently (i.e. without) of the conditions (i), (ii) *we have*

$$f(n)=c\log n \text{ if } (n,2(B-D)/(B,2D))=1.$$

For the proof it is enough to remark that we may assume without loss of generality that $B>D$. Write $n(B-D)/A$ in place of n in (2) we obtain that

$$f((B-D)n+B)+f((B-D)n+D)$$

is monotone increasing, i.e.

$$f((B-D)n+B)+f((B-D)n+D)\geq$$
$$\geq f((B-D)(n-1)+B)+f((B-D)(n-1)+D)$$

hence

$$f((B-D)n+B)\geq f((B-D)n-B+2D)$$

and the Corollary follows from our theorem.

References

[1] ERDŐS, P. On the distribution function of additive functions, *Ann. Math.*, **47** (1946), 1–20.

[2] JOÓ, M. On additive functions, *Acta Math. Acad. Sci. Hung.*, **60** (1992), 349–350.

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 5–8

ON THE ARITHMETIC OF INDEPENDENT DISCRETE DISTRIBUTIONS

By

J. ENGLÄNDER and J. G. SZÉKELY*

Department of Mathematics, Technion, Haifa, and
Department of Mathematics, Technical University, Budapest

(Received December 4, 1991)

1. Introduction

Let $p = \{p_i\}$ and $q = \{q_i\}$ be discrete probability distributions (i.e. $p_i \geq 0$, $q_j \geq 0$, $i, j = 1,2,\ldots, \sum p_i = \sum q_j = 1$). If these distributions are independent marginal distributions of a contingency table then $\{p_i q_j\}$ is the sequence of cross classification probabilities (joint distribution). For convenience suppose that $p_1 \geq p_2 \geq \ldots$ and $q_1 \geq q_2 \geq \ldots$. If we also order the sequence $\{p_i q_j\}$ nonincreasingly and denote the resulting distribution by $r = \{r_k\}$ ($r_k = p_i q_j$ for some i, j) then r can be considered the unique product of p and q. This operation makes the set of all discrete distributions p with nonincreasing p_i's a commutative semigroup D which is also a topological semigroup if endowed with the topology of pointwise convergence.

In this paper we prove some basic results on the arithmetic nature of D.

The unit element of D is $e = (1, 0, 0, \ldots)$. By definition $r \in D$ is irreducible if $r = pq$ implies that either $p = e$ or $q = e$. Every $p \in D$ concentrated on two points is clearly irreducible. The same holds if two is replaced by any prime number. Irreducible distributions do not appear in contingency tables with independent nontrivial ($\neq e$) marginals thus the problem of reducibility seems to be relevant in the analysis of contingency tables.

* Supported by Hungarian National Foundation for Science Research Nos. 1405 and 1905.

2. Main Results

THEOREM 1. *Every element of D is a finite (or countable) product of irreducible elements. This decomposition is not unique.*

THEOREM 2. *The only infinitely divisible element of D is e.*

A distribution r will be called prime if $r \mid pq$ implies $r \mid p$ or $r \mid q$.

THEOREM 3. *A distribution r concentrated on two points is prime if and only if $r = (1/2, 1/2)$. $(1/n, 1/n, \ldots, 1/n)$ is prime if and only if n is prime.*

PROBLEM 1. Is there any other prime element in D?

THEOREM 4. *The set IR of irreducible elements of D form a dense subset in D, and the Baire type of IR is G_δ.*

PROBLEM 2. Characterize the set of irreducible elements of D.

3. Proofs

We shall apply the theory of Hun semigroups discussed in RUZSAand SZÉKELY(1988).

PROPOSITION 1. *D is cancellative.*

PROOF. $pq = pr$ implies $q_1 = r_1$; omit the elements $p_i q_1 = p_i r_1$ and consider the biggest element of the remaining numbers on the left and right hand side. These numbers are equal, i.e. $p_1 q_2 = p_1 r_2$, hence $q_2 = r_2$, etc.

PROPOSITION 2. *The only idempotent of D is e.*

PROOF. The previous Proposition shows that if $x^n = x, \in D$ then $x^{n-1} = e$ and thus $x = e$ (the only divisor of e is e).

The previous observations imply

PROPOSITION 3. *D is associate-free (i.e. if $p \mid q$ and $q \mid p$ then $p = q$).*

PROPOSITION 4. *The set T_p of all divisors of $p \in D$ is compact for every $p \in D$.*

We omit the simple proof.

Recall that in RUZSAand SZÉKELY(1988) a commutative (Hausdorff) topological semigroup was called Hun if it was associate-free and T_p was compact for every element p of the semigroup. Thus our D is a Hun semigroup without nontrivial idempotents. D is also normable where the Rényi entropy (see RÉNYI(1966))

$$\phi_\alpha(p) = -1/\alpha \log \sum p_i^{1+\alpha}, \quad \alpha > 0$$

can be considered the norm of $p \in D$.

PROPOSITION 5. *D is a normable Hun semigroup.*

REMARK 1. An advantage of the norm ϕ_α is the following. If $\phi_\alpha(p) = = \phi_\alpha(q)$ for every $\alpha > 0$ then $p = q$. This immediately gives a new proof of Proposition 1.

The general theory of normable Hun semigroups implies Theorem 1 if there are no infinitely divisible elements in D (other than e). In other words in view of our theory of Hun semigroups Theorem 2 implies Theorem 1.

PROOF OF THEOREM 2. If $q^n = p$ for some $n > 1$ then $q_1^n = p_1$, $q_1^{n-1} q_2 = p_2 = p_3 = \ldots = p_{n+1}$, and this cannot hold for every $n > 1$ except the trivial case $q_2 = 0 = p_2 = p_3 = \ldots$, i.e. $p = e$.

The following example shows that the decomposition in Theorem 1 is not unique. Put $p = (1/(1+b), b/(1+b))$, $q = (1/B, b^2/B, \ldots, b^{2k-2}/B)$ where $0 < b < 1$, $B = 1 + b^2 + \ldots + b^{2k-2}$ and $p' = (1/(1+b^k), b^k/(1+b^k))$, $q' = (1/B' + b/B', \ldots, b^{k-1}/B')$ where again $0 < b < 1$ and $B' = 1 + b + + \ldots + b^{k-1}$. If k is a prime number then all the elements p, q, p', q' are irreducible in D, $pq = p'q'$ and $p \neq p', q'$.

PROOF OF THEOREM 3. Any $p \in D$ concentrated on two points has the form $p = (1/(1+b), b/(1+b))$, $0 < b \leq 1$. If $b \neq 1$, i.e. if $p \neq (1/2, 1/2)$ then the previous example shows that $p \mid p'q'$ where p', q' are irreducibles, $p \neq p'$, q', thus p is not a prime. On the other hand $p = (1/2, 1/2)$ is a divisor of q iff $q_{2j-1} = q_{2j}$, $j = 1, 2, \ldots$ which means that if $p \mid qr$ then $p \mid q$ or $p \mid r$. This observation also implies that $(1/n, 1/n, \ldots, 1/n)$ is prime iff n is prime.

PROOF OF THEOREM 4. If the support of $p \in D$ contains n points where n is a prime number then p is clearly irreducible. On the other hand the set of all distributions whose support contains prime number of points is dense in D. Thus a subset of IR is dense in D. The G_δ property of IR is a simple consequence of Theorem 19.2 of RUZSA and SZÉKELY (1988) (one can easily see that D is a stable semigroup).

References

[1] A. Rényi, *Theory of Probability with Appendix on Information Theory*(in German), VEB Deutsche Verlag der Wissenschaften, Berlin, 1966.

[2] I. Z. Ruzsa, and G. J. Székely, *Algebraic Probability Theory*, Wiley, New York, 1988.

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 9–22

ON CONTROLLABILITY OF BILINEAR SYSTEM III (AN INFINITE DIMENSIONAL CASE)

By

I. JOÓ and M. TUAN

Department of Analysis Eötvös Loránd University, Budapest

(Received February 26, 1992)

In this paper infinite dimensional bilinear systems are examined, where the state space a control spaces are infinite dimensional Banach spaces. It is known that in this case the complete controllability fails, even for the linear systems. Thus the concept of approximate controllability has been introduced. Controllability of infinite dimensional linear systems are investigated in the papers of TRIGGIANI ([21], [22]). Here we manage to get necessary and sufficient conditions for the approximate controllability of a special case of infinite dimensional bilinear systems.

We consider the following systems

$$\dot{x}(t) = Ax(t) + N(x(t), u(t)) + Bv(t), \tag{1}$$

where $0 \le t < \infty$ and $x(t) \in \mathbb{X}$, $u(t) \in \mathbb{U}$, $v(t) \in \mathbb{V}$ for all $t \ge 0$, $\mathbb{X}$, $\mathbb{U}$, $\mathbb{V}$ are complex separable infinite dimensional fixed Banach spaces. Further

$$A : \mathbb{X} \to \mathbb{X}, \qquad B : \mathbb{V} \to \mathbb{X}$$

are bounded operators,

$$N : \mathbb{X} \times \mathbb{U} \to \mathbb{X}$$

is a bounded bilinear operator.

We assume that the control functions $u(t)$, $v(t)$ are locally Bochner integrable, or equivalently that $u(t)$, $v(t)$ the strongly measurable and

$$\|u(t)\|_{\mathbb{U}},\ \|v(t)\|_{\mathbb{V}} \in L_1^{\mathrm{loc}}([0,\infty)),$$

(see HILLE, PHILIPS [4], 3.7). By a solution $x(t)$ of (1) we mean a locally absolutely continuous function satisfying (1) for almost every $t > 0$. If we require that $u(t)$, $v(t)$ be piecewise continuous control functions then (1) must hold in every continuity point. It is known (KREIN [5] II.2.1., II.3.2.) that for any given initial condition $x(0) = x_0 \in \mathbb{X}$ the system (1) has unique solution. As usual, we will also use the notations $x(t, x_0, u, v)$, $R_t(x_0)$, $R(x_0)$ (see [19], [20]).

We say that the system (1) is *approximately controllable* in $\mathbb{X}$ (in time T) if for all $x \in \mathbb{X}$, $\mathrm{cl}(R(x)) = \mathbb{X}$ (resp. $\mathrm{cl}(R_T(x)) = \mathbb{X}$). In [21] and [22] TRIGGIANI investigated the system (1) in the linear case setting $N = 0$ and gave necessary and sufficient conditions for the approximate controllability, which are the complete controllability failed in $\mathbb{X}$, even for the linear systems ($N = 0$). We have

THEOREM A. ([21]) *consider the system* (1) *with* $N = 0$

$$\dot{x}(t) = Ax(t) + Bv(t).$$

The following conditions are equivalent

(i) *The system is approximately controllable in* $\mathbb{X}$ *in time* T *for all* $T > 0$,

(ii) $\mathrm{cl\,span}\{A^n B\mathbb{V} \, : \, n \geq 0\} = \mathbb{X}$

(iii) $\bigcap\limits_{n=0}^{\infty} \{\ker\{B^*(A^*)^n\}\} = \{0\}$.

THEOREM B. ([21]) *The system* (1) *with* $N = 0$ *and* $\dim \mathbb{V} = m$:

$$\dot{x}(t) = Ax(t) + \sum_{i=1}^{m} u_i(t) b_i$$

is approximately controllable in time T *if and only if*

$$\mathrm{cl\,span}\{A^n b_i \, : \, 1 \leq i \leq m, \; n \geq 0\} = \mathbb{X}.$$

This system can never be completely controllable: $\bigcup\limits_{t>0} R_T(0) \neq \mathbb{X}$.

In this paper we consider the special case $A = 0$ of (1), namely

$$\dot{x}(t) = N(x(t), u(t)) + Bv(t). \tag{2}$$

Introduce the notations

$$\mathscr{B}(\mathbb{U}, \mathbb{X}) := \{B \, : \, \mathbb{U} \to \mathbb{X} \, : \, B \text{ is bounded linear operator}\}$$

$$\mathscr{B}([0,T], \mathbb{U}) := \{u(\cdot) \, : \, [0,T] \to \mathbb{U} \, : \, u(\cdot) \text{ is strongly measurable and } \|u(\cdot)\|_{\mathscr{B}} = \int_0^T \|u(t)\| dt < \infty\}$$

$\mathscr{B}([0,T], \mathbb{U})$ is a Banach space every the norm $\|u(\cdot)\|_{\mathscr{B}}$.

To the system (2) it corresponds the following system

$$\begin{cases} \dot{X}(t) = N(X(t), u(t)) \\ X(0) = E. \end{cases} \tag{3}$$

For fixed t, $X(t) : \mathbb{X} \to \mathbb{X}$ and $N(X(t), u(t))$ are bounded linear operators; by definition the latter one works by the rule $N(X(t), u(t))\ x :=$ $= N(X(t)x, u(t))$. If $u(\cdot) \in \mathscr{B}([0,T], \mathbb{U})$ then (3) has a unique solution, denoted by $\Phi_u(t)$. Now the solution of (2) can be given by

$$x(t, x_0, u, v) = \Phi_u(t) \left(x_0 + \int_0^t [\Phi_u(s)]^{-1} Bv(s) ds \right).$$

If $u(t) \equiv u$ then $\Phi_u(t) = e^{tN(\cdot,u)}$, thus if $u(t)$ is piecewise constant,

$$u(t), \quad \text{for} \quad t \in (t_{i-1}, t_i], \quad 0 = t_0 < t_1 < \ldots < t_p = t,$$

then

$$\Phi_u(t) = e^{(t-t_{p-1})N(\cdot,u_p)} e^{(t_{p-1}-t_{p-2})N(\cdot,u_{p-1})} \ldots e^{t_1 N(\cdot,u_1)}$$

this obeys (3) in all continuity points $t \neq t_i$.

LEMMA 1. *For the system* (2) $x \in R_T(y)$ *if and only if* $y \in R_T(x)$.

PROOF. $x \in R_T(y)$ means that $x = x(T, y, u, v)$. Let

$$\hat{u}(t) = -u(T-t), \quad \hat{v} = -v(T-t), \quad y(t) = x(T-t, y, u, v).$$

Then $\dot{y} = -\dot{x}(T-t) = N(y(t), \hat{u}(t)) + B\hat{v}(t)$ a $y(0) = x$, $y(T) = y$, hence $y \in R_T(x)$. □

LEMMA 2. *Consider the system* (2). *Let* $x_1 \in R_{t_1}(0)$, $x_2 \in R_{t_2}(0)$, α, $\beta \in \mathbb{C}$. *Then*

$$\alpha x_1 + \beta x_2 \in R_{2t_1+t_2}(0).$$

In particular, $R(0)$ *is a linear subspace of* $\mathbb{X}$.

PROOF. Since $0 \in R_{t_1}(x_1)$, $0 \in R_{t_2}(x_2)$, there exist controls u_i, v_i, $i = 1,2$ with

$$0 = x(t_i, x_i, u_i, v_i) = \Phi_{u_i}(t_i) \left(x_i + \int_0^{t_i} [\Phi_{u_i}(s)]^{-1} Bv_i(s) ds \right)$$

that is,

$$x_i = -\int_0^{t_i} [\Phi_{u_i}(s)]^{-1} Bv_i(s) ds, \qquad i = 1,2.$$

Define the controls

$$u(t) \in \mathcal{B}([0, 2t_1 + t_2], \mathbb{U}), \qquad v(t) \in \mathcal{B}([0, 2t_1 + t_2], \mathbb{V})$$

by the rule

$$(u(t), v(t)) = \begin{cases} (u_1(t), \alpha v_1(t)), & 0 \le t \le t_1 \\ (-u_1(2t_1 - t), 0), & t_1 < t < 2t_1 \\ (u_2(t - 2t_1), v_2(t - 2t_1)), & 2t_1 \le t \le 2t_1 + t_2 \end{cases}$$

Then we have

$$\Phi_u(t) = \begin{cases} \Phi_{u_1}(t), & 0 \le t \le t_1 \\ \Phi_{u_1}(2t_1 - t), & t_1 < t < 2t_1 \\ \Phi_{u_2}(t - 2t_1), & 2t_1 \le t \le 2t_1 + t_2 \end{cases}$$

in particular $\Phi_u(2t_1) = E$. Consequently

$$\alpha x_1 + \beta x_2 = -\int_0^{t_1} [\Phi_{u_1}(s)]^{-1} Bv_1(s) ds - 0 - \int_0^{t_2} [\Phi_{u_2}(s)]^{-1} Bv_2(s) ds =$$

$$= -\int_0^{2t_1 + t_2} [\Phi_u(s)]^{-1} Bv(s) ds$$

which means that

$$0 \in R_{2t_1 + t_2}(\alpha x_1 + \beta x_2), \qquad \alpha x_1 + \beta x_2 \in R_{2t_1 + t_2}(0)$$

as we asserted. □

LEMMA 3. (Gronwall inequality, see [39], p.24)*Let* $k(t)$, $\varphi(t) \ge 0$, k, $\varphi \in L_1^{\mathrm{loc}}[0, \infty)$. *Suppose that for some* $M > 0$

$$k(t) \le M + \int_0^t k(s)\varphi(s) ds \qquad \text{for all } t > 0.$$

Then

$$k(t) \le M e^{\int_0^t \varphi(s) ds} \qquad \text{for all } t > 0.$$

(Remark that equality in the assumption implies equality in the conclusion.)

THEOREM 1. *For the system* (2) *the mapping*

$$\mathcal{B}([0, T], \mathbb{U} \times \mathbb{V}) \to C([0, T], \mathbb{X}) \qquad (u(\cdot), v(\cdot)) \to x(\cdot, x_0, u, v)$$

is continuous.

PROOF. It consists of three steps.

(a) We prove first that the mapping $u(\cdot) \to \Phi_u(\cdot)$ is continuous from $\mathscr{B}([0,T],\mathbb{U})$ into $C([0,T],\mathscr{B}(\mathbb{X},\mathbb{X}))$. Indeed, let $u_n(\cdot) \to (u(\cdot)$ in $\mathscr{B}([0,T]),\mathbb{U})$-norm, i.e.

$$\int_0^T \|u_n(t) - u(t)\| dt \to 0 \qquad \text{as} \quad n \to \infty.$$

In particular we have then $\|u(\cdot)\|$, $\|u_n(\cdot)\| \leq K$. (Here $\|u(\cdot)\|$ means $\|u(\cdot)\|_{\mathscr{B}}$.) From the integral equation corresponding to (3)

$$\Phi_{u_n}(t) = E + \int_0^t N(\Phi_{u_n}(s), u_n(s))ds$$

it follows that

$$\|\Phi_{u_n}(t)\| \leq 1 + \int_0^t \|N\| \, \|\Phi_{u_n}(s)\| \, \|u_n(s)\|ds.$$

Using Lemma 3 with

$$k(t) = \|\Phi_{u_n}(t)\|, \qquad M = 1, \qquad \varphi(t) = \|N\| \, \|u_n(t)$$

we obtain

$$\|\Phi_{u_n}(t)\| \leq e^{\|N\| \int_0^t \|u_n(s)\|ds} \leq e^{K\|N\|}.$$

We have

$$\|\Phi_{u_n}(t) - \Phi_u(t)\| = \left\| \int_0^t [N(\Phi_{u_n}(s), u_n(s)) - N(\Phi_u(s), u(s))ds] \right\| \leq$$

$$\leq \left\| \int_0^t N(\Phi_{u_n}(s) - \Phi_u(s), u(s))ds \right\| + \left\| \int_0^t N(\Phi_{u_n}(s), u_n(s) - u(s))ds \right\| \leq$$

$$\leq \|N\| \int_0^t \|\Phi_{u_n}(s) - \Phi_u(s)\| \, \|u(s)\|ds + \|N\|e^{K\|N\|} \int_0^t \|u_n(s) - u(s)\|ds.$$

For large n the second member will be smaller than any prescribed $\varepsilon > 0$, hence we can apply again Lemma 3:

$$\|\Phi_{u_n}(t) - \Phi_u(t)\| \leq \varepsilon e^{\int_0^t \|N\| \, \|u_n(s)\|ds} \leq \varepsilon e^{K\|N\|}$$

which proves the point (a).

(b) We will show that if $u_n(\cdot) \to u(\cdot)$ in $\mathcal{B}$-norm, then

$$[\Phi_{u_n}(\cdot)]^{-1} \to [\Phi_u(\cdot)]^{-1} \quad \text{in} \quad C([0,T]), \mathcal{B}(\mathbb{X},\mathbb{X})).$$

For the proof we fix an $s \in (0,T]$ and set $\hat{u}(t) := -u(s-t)$, $0 \le t \le s$. Then $\Phi_{\hat{u}}(t) = \Phi_u(s-t)[\Phi_u(s)]^{-1}$ can be checked directly from (3). In particular $\Phi_{\hat{u}}(s) = [\Phi_u(s)]^{-1}$.

Now if $u_n(\cdot) \to u(\cdot)$ in $\mathcal{B}([0,T]), \mathbb{U})$, then $\hat{u}_n(\cdot) \to \hat{u}(\cdot)$ in $\mathcal{B}([0,s], \mathbb{U})$ (s is fixed). Repeating the estimates of (a) we get that for large n (independent of s)

$$\| [\Phi_{u_n}(s)]^{-1} - [\Phi_u(s)]^{-1} \| = \|\Phi_{\hat{u}_n}(s) - \Phi_{\hat{u}}(s)\| \le \varepsilon e^{K\|N\|}$$

which proves (b).

(c) We return to the statement of the theorem. Let

$$u_n(\cdot) \to u(\cdot), \qquad v_n(\cdot) \to v(\cdot)$$

in the corresponding $\mathcal{B}$-norms. Then

$$\|u(\cdot)\|, \quad \|v(\cdot)\|, \quad \|u_n(\cdot)\|, \quad \|v_n(\cdot)\| \le K$$

and from (a), (b) we see that

$$\|x(t,x_0,u_n,v_n) - x(t,x_0,u,v)\| \le$$

$$\le \|\Phi_{u_n}(t)x_0 - \Phi_u(t)x_0\| + \left\| \int_0^t \left[\Phi_{u_n}(t)\Phi_{u_n}^{-1}(s) - \Phi_u(t)\Phi_u^{-1}(s)\right] Bv_n(s)ds \right\| +$$

$$+ \left\| \int_0^t \Phi_u(t)\Phi_u^{-1}(s)B(v_n(s)-v(s))ds \right\| \le \|\Phi_{u_n}(t) - \Phi_u(t)\| \, \|x_0\| +$$

$$+ \int_0^t \left\|\Phi_{u_n}(t)\Phi_{u_n}^{-1}(s) - \Phi_u(t)\Phi_u^{-1}(s)\right\| \, \|B\| \, \|v_n(s)\|ds +$$

$$+ \int_0^t \|\Phi_u(t)\| \, \|\Phi_u^{-1}(s)\| \|B\| \, \|v_n(s)-v(s)\|ds \le \|\Phi_{u_n}(t) - \Phi_u(t)\| \, \|x_0\| +$$

$$+ \int_0^t \Big(\|\Phi_{u_n}(t) - \Phi_u(t)\| \, \|\Phi_u^{-1}(s)\| + \|\Phi_{u_n}(t)\| \, \|\Phi_{u_n}^{-1}(s) - \Phi_u^{-1}(s)\| \Big) \cdot$$

$$\cdot \|B\| \, \|v_n(s)\|ds + c \int_0^t \|v_n(s) - v(s)\|ds$$

and every member tends to zero uniformly in $0 \le t \le T$.

The proof is complete. □

LEMMA 4. *For the system (2) we have $R_t(0) = R_T(0)$ for all $0 < t \le \le T < \infty$.*

PROOF.

(a) $R_t(0) \subset R_T(0)$.

Let $x = x(t,0,u,v) \in R_T(0)$ and

$$(\tilde{u},\tilde{v})(s) := \begin{cases} 0, & 0 \le s \le T-t \\ (u,v)(s-T+t), & T-t < s \le T. \end{cases}$$

Then by (2) we have

$$x(s,0,\tilde{u},\tilde{v}) = \begin{cases} 0, & 0 \le s \le T-t. \\ x(s-T+t,0,u,v), & T-t < s \le T \end{cases}$$

in particular $x(T,0,\tilde{u},\tilde{v}) = x(t,0,u,v)$.

(b) $R_T(0) \subset R_t(0)$.

Consider first the case when u and v are piecewise constant functions, namely

$$(u(s),v(s)) = (u_i,v_i) \quad \text{for} \quad t \in (t_{i-1},t_i), \quad 0 = t_0 < t_1 < \ldots < t_p = T.$$

As we know

$$x(T,0,u,v) = \Phi_u(T)\int_0^T \Phi_u^{-1}(s)Bv(s)ds = \sum_{s=1}^{p}\int_{t_{i-1}}^{t_i} \Phi_u(T)\Phi_u^{-1}(s)Bv(s)ds =$$

$$= \sum_{i=1}^{p}\int_{t_{i-1}}^{t_i} e^{(T-t_{p-1})N(\cdot,u_p)}e^{(t_{p-1}-t_{p-2})N(\cdot,u_{p-1})}\ldots e^{t_1N(\cdot,u_1)}.$$

$$\cdot e^{-t_1N(\cdot,u_1)}\ldots e^{-(t_{i-1}-t_{i-2})N(\cdot,u_{i-1})}e^{(s-t_{i-1})N(\cdot,u_i)}Bv_i ds =$$

$$= \sum_{i=1}^{p}\int_{t_{i-1}}^{t_i} e^{(T-t_{p-1})N(\cdot,u_p)}\ldots e^{(t_i-s)N(\cdot,u_i)}Bv_i ds.$$

Take the substitutions

$$s_i := \frac{t}{T}t_i, \quad \hat{u}_i := \frac{T}{t}u_i, \quad \hat{v}_i := \frac{T}{t}v_i, \quad (\hat{u},\hat{v})(s) := (\hat{u}_i,\hat{v}_i) \quad \text{on} \quad (s_{i-1},s_i).$$

Then

$$0 < s_0 < s_1 < \ldots < s_p = t$$

and the substitution $r = \frac{t}{T}s$ in the above integral gives

$$x(T,0,u,v) = x(t,0,\hat{u},\hat{v}).$$

Now we prove (b) for general control functions. There exist piecewise constant controls u_n, v_n with

$$u_n(\cdot) \to u(\cdot) \quad \text{in} \quad \mathscr{B}([0,T],\mathbb{U}), \qquad v_n(\cdot) \to v(\cdot) \quad \text{in} \quad \mathscr{B}([0,T],\mathbb{V}).$$

Then there are $\hat{u}(\cdot)$, $\hat{v}(\cdot)$ such that

$$\hat{u}_n(\cdot) \to \hat{u}(\cdot) \quad \text{in} \quad \mathscr{B}([0,t],\mathbb{U}), \qquad \hat{v}_n(\cdot) \to \hat{v}(\cdot) \quad \text{in} \quad \mathscr{B}([0,t],\mathbb{V}).$$

(see the substitutions above). By Theorem 1 we have

$$x(T,0,u_n,v_n,) \to x(T,0,u,v) \qquad x(t,0,\hat{u}_n,\hat{v}_n,) \to x(t,0,\hat{u},\hat{v})$$

and the left numbers are equal, hence

$$x(T,0,u,v) = x(t,0,\hat{u},\hat{v})$$

which finishes the proof. □

From Lemma 2 and this Lemma we obtain the following

COROLLARY. *For the system* (2) $R(0) = R_T(0)$ *is a linear subspace of* $\mathbb{X}$ *for all* $T > 0$.

LEMMA 5. *Consider the system* (2). *If* $\mathrm{cl}R_T(0) = \mathbb{X}$ *then for all* $x \in \mathbb{X}$, $\mathrm{cl}R_T(x) = \mathbb{X}$.

PROOF. For x, $y \in \mathbb{X}$, $\varepsilon > 0$ we have to find a control (u,v) such that $\|y - x(T,x,u,v)\| < \varepsilon$. By $\mathrm{cl}R_{\frac{T}{3}}(0) = \mathbb{X}$ we have a control (u_2,v_2) such that

$$\|y - \hat{y}\| < \frac{\varepsilon}{2}, \quad \text{where} \quad \hat{y} = x\left(\frac{T}{3},0,u_2,v_2\right)$$

and there exists (u_1,v_1) such that

$$\|x - \hat{x}\| < \frac{\varepsilon}{2}\left\|\Phi_{u_2}\left(\frac{T}{3}\right)\right\|^{-1}, \quad \text{where} \quad \hat{x} = x\left(\frac{T}{3},0,u_1,v_1\right).$$

Define the control

$$(u(t),v(t)) := \begin{cases} -\left(u_1\left(\frac{T}{3}-t\right),v_1\left(\frac{T}{3}-t\right)\right), & 0 \le t \le \frac{T}{3} \\ -\left(u_1\left(t-\frac{T}{3}\right),0\right), & \frac{T}{3} < t \le \frac{2T}{3} \\ \left(u_2\left(t-\frac{2T}{3}\right),v_2\left(t-\frac{2T}{3}\right)\right), & \frac{2T}{3} < t \le T. \end{cases}$$

Using (2) and (3) we see that for $t \leq \frac{T}{3}$

$$x(t,\hat{x},u,v) = x\left(\frac{T}{3}-t,0,u_1,v_1\right) \quad \text{and} \quad x\left(\frac{T}{3},\hat{x},u,v\right) = 0$$

$$\Phi_u(t) = \Phi_{u_1}\left(\frac{T}{3}-t\right)\Phi_{u_1}^{-1}\left(\frac{T}{3}\right) \quad \text{and} \quad \Phi_u\left(\frac{T}{3}\right) = \Phi_{u_1}^{-1}\left(\frac{T}{3}\right)$$

For $\frac{T}{3} \leq t \leq \frac{2T}{3}$ we have $x(t,\hat{x},u,v) = 0$

$$\Phi_u(t) = \Phi_{u_1}\left(t-\frac{T}{3}\right)\Phi_{u_1}^{-1}\left(\frac{T}{3}\right), \qquad \Phi_u\left(\frac{2T}{3}\right) = E.$$

Finally for $\frac{2T}{3} \leq t \leq T$

$$x(t,\hat{x},u,v) = x\left(t-\frac{2T}{3},0,u_2,v_2\right), \qquad x(T,\hat{x},u,v) = \hat{y}$$

$$\Phi_u(t) = \Phi_{u_2}\left(t-\frac{2T}{3}\right).$$

Thus we prove that

$$\|y - x(T,\hat{x},u,v)\| = \|y-\hat{y}\| < \frac{\varepsilon}{2}$$

and it remains to show that

$$\|x(T,\hat{x},u,v) - x(T,x,u,v)\| < \frac{\varepsilon}{2}.$$

Indeed

$$\|x(T,\hat{x},u,v) - x(T,x,u,v)\| = \|\Phi_u(T)(\hat{x}-x)\| =$$
$$= \left\|\Phi_{u_2}\left(\frac{T}{3}\right)(\hat{x}-x)\right\| \leq \left\|\Phi_{u_2}\left(\frac{T}{3}\right)\right\| \|\hat{x}-x\| < \frac{\varepsilon}{2}.$$

Lemma is proved. □

THEOREM 2. *For the system (2) the following statements are equivalent*

(i) *Approximate controllability holds in time T, i.e.*

$$\mathrm{cl}R_T(x) = \mathbb{X} \qquad \textit{for all} \quad x \in \mathbb{X}.$$

(ii) $\mathrm{cl\,span}\{N(\cdot,u_1)\ldots N(\cdot,u_k)Bv \;:\; k \geq 0,\; u_1,\ldots,u_k \in \mathbb{U},\; v \in \mathbb{V}\} = \mathbb{X}$,

(iii) $\bigcap\{\ker(B^*N(\cdot,u_1)^*\ldots N(\cdot,u_k)^*) \;:\; k \geq 0,\; u_1,\ldots,u_k \in \mathbb{U}\} = \{0\}$.

PROOF. Denote x^* the elements of $\mathbb{X}^*$. By the Hahn–Banach theorem (i) $\Leftrightarrow$ "If $0 = \langle x^*, N(\cdot,u_1)\ldots N(\cdot,u_k)Bv\rangle$ for all $u_i \in \mathbb{U}$, $v \in \mathbb{V}$, $k \geq 0$, then $x^* = 0$".

From this and from

$$\langle x^*, N(\cdot,u_1)\dots N(\cdot,u_k)Bv\rangle = \langle B^*N(\cdot,u_k)^*\dots N(\cdot,u_1)^*x^*, v\rangle$$

we see at once that (ii) $\Leftrightarrow$ (iii). Thus by Lemma 5 it remains to prove that

$$\mathrm{cl}R_T(0) = \mathbb{X} \ \Leftrightarrow\ \text{(ii)}.$$

Introduce the notations

$$H := \{(u,v)\ :\ u,\ v \text{ are piecewise constant}\}$$

$$R_T^H(0) := \{x(T,0,u,v)\ :\ (u,v)\in H\}.$$

Since H is dense in the space of all control functions, Theorem 1 shows that

$$\mathrm{cl}R_T(0) = \mathrm{cl}R_T^H(0).$$

The statements of Lemmas 2 and 4 hold also for $R_T^H(0)$ instead of $R_T(0)$ (see the proofs there). Consequently $R_T^H(0) = R^H(0)$ is a linear subspace of $\mathbb{X}$ for all $T > 0$. Thus we have only to prove

$$\mathrm{cl}R_T^H(0) = \mathbb{X} \ \Leftrightarrow\ \text{(ii)}.$$

In fact we show that both are equivalent to the statement

(iv) *If $x^* \in \mathbb{X}^*$ a $0 = \langle x^*, \Phi_u(T)\Phi_u^{-1}(t)Bv\rangle$ for all piecewise constant function $u(\cdot)$ and all $v \in \mathbb{V}$, $0 \le t \le T$, then $x^* = 0$.*

(a) $\mathrm{cl}R_T^H(0) = \mathbb{X} \ \Rightarrow\ $ (iv).

Indeed, take an x^* satisfying the conditions of (iv). Let $x = x(T,0,u,v)$ with $(u,v) \in H$. Then

$$\langle x^*, x\rangle = \left\langle x^*, \int_0^T \Phi_u(T)\Phi_u^{-1}(s)Bv(s)ds\right\rangle =$$

$$= \int_0^T \left\langle x^*, \Phi_u(T)\Phi_u^{-1}(s)Bv(s)\right\rangle ds = 0$$

i.e. $\left\langle x^*, R_T^H(0)\right\rangle = \{0\}$. Since $R_T^H(0)$ is dense in $\mathbb{X}$, $x^* = 0$ follows.

(b) (iv) $\Rightarrow\ \mathrm{cl}R_T^H(0) = \mathbb{X}$.

Let $\left\langle x^*, R_T^H(0)\right\rangle = \{0\}$, we have to show that $x^* = 0$. Consider any $(u,v) \in H$ with

$$v(s) = \begin{cases} v, & 0 \le s \le t \\ 0, & t < s \le T. \end{cases}$$

Then

$$x(T,0,u,v) = \int_0^t \Phi_u(T)\Phi_u^{-1}(s)Bvds$$

and hence

$$a = \int_0^t \left\langle x^*, \Phi_u(T)\Phi_u^{-1}(s)Bv \right\rangle ds.$$

Taking the derivative in t we get $0 = \langle x^*, \Phi_u(T)\Phi_u^{-1}(t)Bv \rangle$ for all t, and then $x^* = 0$ by (iv).

(c) (iv) $\Rightarrow$ (ii).

Let $0 = \langle x^*, N(\cdot,u_1)\dots N(\cdot,u_k)Bv \rangle$ for all $u_1,\dots,u_k \in \mathbb{U}$, $v \in \mathbb{V}$, $k \geq 0$. By the remark at the beginning of the proof we have to show that $x^* = 0$. Let $u(\cdot)$ be arbitrary piecewise constant control,

$$u(s) = u_i \text{ for } s \in (t_i, t_{i+1}),\ t = t_1 < \dots < t_{p+1} = T,\ s_i = t_{i+1} - t_i,$$

and $v \in \mathbb{V}$. Then

$$\Phi_u(T)\Phi_u^{-1}(t)Bv = e^{s_p N(\cdot,u_p)} \dots e^{s_1 N(\cdot,u_1)} Bv =$$
$$= \sum \varrho_{i_1 \dots i_k} s_{i_1} \dots s_{i_k} N(\cdot,u_{i_1}) \dots N(\cdot,u_{i_k}) Bv$$

where $1 \leq i_k \leq \dots \leq i_1 \leq p$, $k = 1,2,\dots$. Hence

$$0 = \left\langle x^*, \Phi_u(T)\Phi_u^{-1}(t)Bv \right\rangle$$

and by (iv) we heve $x^* = 0$.

(d) (ii) $\Rightarrow$ (iv).

Suppose that x^* satisfies the conditions given in (iv). Let $u_1,\dots,u_p \in \mathbb{U}$, $v \in \mathbb{V}$, $t \in [0,T]$ be arbitrary and define

$$u(s) = \begin{cases} u_i, & t_i < s < t_{i+1},\ i = 1,\dots,p \\ 0, & s < t = t_1, \end{cases}$$

where $t = t_1 < \dots < t_{p+1}$. We repeat the counting in (c)

$$0 = \left\langle x^*, \Phi_u(T)\Phi_u^{-1}(t)Bv \right\rangle =$$
$$= \sum \varrho_{i_1 \dots i_k} s_{i_1} \dots s_{i_k} \left\langle x^*, N(\cdot,u_{i_1}) \dots N(\cdot,u_{i_k}) Bv \right\rangle = \sum d_{i_1 \dots i_k} s_{i_1} \dots s_{i_k}$$

where the sum runs over $1 \leq i_k \leq \dots \leq i_1 \leq p$, $k = 1,2,\dots$ and

$$d_{i_1 \dots i_k} = \varrho_{i_1 \dots i_k} \left\langle x^*, N(\cdot,u_{i_1}) \dots N(\cdot,u_{i_k}) Bv \right\rangle.$$

This multiple series is absolutely convergent and the only restrictions for its variables s_i are

$$(4) \qquad 0 \le s_i, \quad i = 1, \ldots, p; \qquad s_1 + \ldots + s_p \le T$$

(and denote $t = T - (s_1 + \ldots + s_p)$).

The derivate $\dfrac{\partial^p}{\partial s_1 \ldots \partial s_p}$ can be taken term by term in the interior of the domain described by (4). Letting all $s_i > 0$ tend to zero we get

$$0 = d_{p\ldots1} = \varrho_{p\ldots1} \langle x^*, N(\cdot, u_p) \ldots N(\cdot, u_1) Bv \rangle = \langle x^*, N(\cdot, u_p) \ldots N(\cdot, u_1) Bv \rangle$$

since $\varrho_{p\ldots1} = 1$. Then by (ii), $x^* = 0$.

The proof of the Theorem is complete. □

REMARK. One can see that the assumptions here are too restrictive. The question is to investigate controllability of the system (1) in the cases a) the drift term is non-zero, i.e. $A \neq 0$, b) the coefficient operators are unbounded, c) A and N do not commute (in the case A and N commute (with special unbounded bilinear operator N) the approximate controllability of (1) was studied in [40]).

References

[1] R. KALMAN, P. FALB, M. ARBIB, *Topics in Mathematical System Theory*, McGraw-Hill, New York, 1969.

[2] E. B. LEE and L. MARKUS, *Foundations of Optimal Control Theory*, Wiley, New York, 1976.

[3] V. V. AMEL'KIN, *Differential equations in applications*, Mir, Moscow, 1990.

[4] E. HILLE, R. S. PHILLIPS, Functional Analysis and Semi-Groups, *AMS Coll. Publ. vol. XXXI.*, Providence, 1957.

[5] S. G. KREIN, *Linear differential equations in Banach spaces* in Russian, Nauka, Moscow, 1967.

[6] H. J. SUSSMANN and V. JURDJEVIC, Controllability of Nonlinear Systems, *J. Diff. Equats.*, **12** (1972) 96–116.

[7] R. W. BROCKETT, System Theory on Group manifold and Coset Space *SIAM J. Control Optim.*, **10**, (1972), 265–284.

[8] R. W. BROCKETT, On the algebraic structure of Bilinear Systems in *Variable Structure Control Systems*, R. Mohler and A. Ruberti eds., Academic Press, 1972, pp. 153–168.

[9] R. W. BROCKETT, Lie algebras and Lie groups in control theory, in *Geometric Methods in System Theory*, D. Q. Mayne and R. W. Brockett Eds., Reidel, Dordrecht, 1973, pp. 43–82.

[10] R. HERMANN and A. J. KRENER, Nonlinear Controllability and Observability *IEEE Trans Aut. Control*, **22** (1977), 728–740.

[11] C. BRUNI, G. DI PILLO and G. KOCH, Bilinear Systems: An Appealing Class of "Nearly Linear" Systems in Theory and Applications, *IEEE Trans. on Aut. Control*, **19** (1974), 334–348.

[12] R. E. RINK and R. R. MOHLER, Completely Controllable Bilinear Systems, *SIAM J. Control Optim.*, **6**, (1968).

[13] D. L. LUKES, Global Controllability of Nonlinear Systems *SIAM J. Control Optim.*, **10**, (1972), 112–126.

[14] D. WILLIAMSON, Observation of Bilinear Systems with Application to Biological Control, *Automatica*, **13** (1977), 243–254.

[15] W. BOOTHBY and E. N. WILSON, Determination of the Transitivity of Bilinear Systems, *SIAM J. Contr. Optim.*, **17** (1979), 212–221.

[16] H. NIJMEIJER and A. J. VAN DER SCHAFT, *Nonlinear Dynamical Control Systems*, Springer-Verlag, 1990.

[17] U. PIECHOTTKA and P. M. FRANK, Controllability of Bilinear Systems: A survey and some new results, in *Nonlinear Control Systems design*, ed. A. Isidori, Pergamon Press, 1990, pp. 23–28.

[18] O. HÁJEK, Control Theory in the Plane, in *Lecture Notes in Control and Information Sciences*, Vol. 153, Springer Verlag, 1991.

[19] I. JOÓ and N. M. TUAN, On Controllability of Bilinear Systems I, *Annales Univ. Sci. Budapest.*, Sectio Math., **35** (1992), 189–217.

[20] I. JOÓ and N. M. TUAN, On Controllability of Bilinear Systems II, *Annales Univ. Sci. Budapest.*, Sectio Math., **35** (1992), 217–266.

[21] R. TRIGGIANI, Controllability in Banach Space with Bounded Operators, *SIAM J. Control Optim.*, **13** (1975), 462–491.

[22] R. TRIGGIANI, Extensions of Rank Conditions for Controllability a Observability to Banach Space with Unbounded Operators, *SIAM J. Contr. Optim.*, **14** (1976), 313–338.

[23] J.-L. LIONS, *Contrôlabilité Exacte, Perturbations et Stabilisation de Systèmes Distribués*, tome 1, 2, Masson Collection RMA, volumes 8 et 9, Paris, 1988.

[24] J.-L. LIONS, Exact Controllability, Stabilization and Perturbations for Distributed Parameter Systems, *SIAM Rev.*, **30** (1988), 1–68.

[25] D. L. RUSSELL, Controllability and Stabilizability Theory for Linear Partial Differential Equations, *SIAM Rev.*, **20** (1978), 639–739.

[26] D. L. RUSSELL, Nonharmonic Fourier Series in the Control Theory of Distributed Parameter Systems, *J. of Math. Anal. and Appl.*, **18** (1967), 542–560.

[27] SZ. DOLECZKI, D. L. RUSSELL, A General Theory of Observation and Control, *SIAM J. Contr. Optim.*, **15** (1977), 185–220.

[28] V. KOMORNIK, A new Method of Exact Controllability in Short Time and Appilcations, *Annales Faculté des Sciences de Toulouse*, **10** (1989), 415–464.

[29] E. ZUAZUA, V. KOMORNIK, A Direct Method for Boundary Stabilization of the Wawve Equation, *J. Math. Pures Appl.*, (1990).

[30] M. HORVÁTH, The Vibration of a Membrane in Different Points, *Annales Univ. Sci. Budapest.*, Sectio Math., to appear.

[31] I. JOÓ, Controlabilité Exacte et Propriétés d'Oscillation de l'équation des ondes par Analyse Nonharmonique, *C. R. Acad. Sci. Paris*, Serie I., **312** (1991), 119–122.

[32] S. A. AVDONIN, S. A. IVANOV, *Controllability Theory for Distributed Parameter Systems and Families of Exponentials*, Kiev, UMK, VO, 1989, (in Russian).

[33] V. KOMORNIK, Rapid Boundary Stabilization of the Wave Equation, *SIAM J. Contr. Optim.*, **29** (1991).

[34] V. I. ARNOLD, *Mathematical Methods of Classical Mechanics*, Springer, 1978.

[35] A. ISIDORI, ed. *Nonlinear Control Systems Design*, IFAC Symposia Series, 1990. Number **2**.

[36] M. FLIESS, ed. *Nonlinear Control Systems Design*, IFAC Symposia Series, 24–26 June, 1992, Bordeaux, France.

[37] H. J. ZWART, *Geometric Theory for Infinite Dimensional Systems*, in Lecture Notes in Control and Information Sciences, **115**, 1989.

[38] R. F. CURTAIN, A. J. PRITCHARD, *Infinite Dimensional Linear Systems Theory*, in Lecture Notes in Control and Information Sciences **8**, 1978.

[39] P. HARTMAN, *Ordinary Differential Equations*, Second Edition, Birkhäuser, 1982.

[40] J. M. BALL, J. E. MARSDEN and M. SLEMROD, Controllability for Distributed Bilinear Systems, *SIAM J. Contr. Opt.*, **20** (1982), 575–597.

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 23–30

ON MNP–GROUPS

By

A. SHAALAN AND M. RAMADAN

Department of Mathematics, Faculty of Science, Cairo University, Giza, Egypt

(Received February 26, 1992)

1. Introduction

Throughout this paper, let G be a finite group. We say, following GARY L. WALLS [1], that G is an MNP-group if all maximal subgroups of any Sylow subgroup of G are normal in G. In [2], S. SRINIVASAN investigated the structure of an MNP-group and proved that an MNP-group is supersolvable. Obviously, nilpotent groups are MNP-groups. The symmetric group on three letters shows that an MNP-group need not to be nilpotent. Let $G = A_5 \cong \mathrm{PSL}(2,4) \cong \mathrm{PSL}(2,5)$ or $G = \mathrm{PSL}(2,13)$. Then G is a non-abelian simple group all of whose second maximal subgroups are MNP-groups. The purpose of this paper is to give a complete classification of all non-abelian simple groups in which each second maximal subgroup is an MNP-group. The notation used in this paper is standard.

2. Preliminaries

In this section, we collect some of the results that are needed in this paper.

(2.1) (SRINIVASAN [2]). If G is an MNP-group, then G is supersolvable.

(2.2) [3, Theorem 9.3.11, p. 229]. If G is a solvable group, then G has a Sylow basis.

(2.3) (DOERK [4]; see also [5, Aufgabe 16, p. 721]). If G is a minimal non-supersolvable group (non-supersolvable group all of whose proper subgroups are supersolvable), then:

(i) G possesses an ordered Sylow tower or G is a minimal non-nilpotent group (non-nilpotent group all of whose proper subgroups are nilpotent);

(ii) G possesses a unique normal Sylow p-subgroup P for some prime p;

(iii) $P/\Phi(P)$ is a minimal normal subgroup of $G/\Phi(P)$;

(2.4) [5, Satz 5.2, p. 281]. If G is a minimal non-nilpotent group, then:

(i) $|G| = p^a q^b$, where p and q are distinct primes;

(ii) G has a normal Sylow p-subgroup P and a cyclic Sylow q-subgroup Q;

(iii) $P/\Phi(P)$ is a minimal normal subgroup of $G/\Phi(P)$.

(2.5) [6, p. 467]. PSL(3,3) has a subgroup isomorphic to S_4, where S_4 is the symmetric group on four letters.

(2.6) ([7]; see also [5, Bemerkung 7.5, p. 190.]). Any non-abelian simple group all of whose proper subgroups are solvable is isomorphic to one of the following simple groups:

(i) PSL(3,3);

(ii) PSL$(2,p)$, where p is a prime with $p > 3$ and $p^2 - 1 \not\equiv o(5)$;

(iii) PSL$(2,2^q)$, where q is a prime;

(iv) PSL$(2,3^q)$, where q is an odd prime;

(v) The Suzuki group Sz(2^q), where q is an odd prime.

(2.7) ([8, Theorem 8.2, p. 41]; see also [9, Theorem 3.3, p. 184]). If G is one of the simple groups mentioned in (2.6) other than PSL(3,3), then G is a Zassenhaus group of degree $n+1$, where $n = r$ or $n = r^2$ according as $G = \mathrm{PSL}(2,r)$ or $G = \mathrm{Sz}(r)$ and the subgroup N fixing a letter is a maximal subgroup of G. Further, N is a Frobenius group with kernel K of order n and a cyclic complement H. If $G = \mathrm{PSL}(2,r)$, then $|H| = (r-1)/d$, where $d = (r-1,2)$, and if $G = \mathrm{Sz}(r)$, then $|H| = r-1$.

(2.8) [8, Theorem 2.4, p. 178]. If A is a p'-group of automorphisms of the abelian p-group P $(p > 2)$ which acts trivially on $\Omega_1(P)$, then $A = 1$.

(2.9) [5, § 8. Die Untergruppen von PSL$(2,p^f)$, pp. 191–213].

Suppose that G is one of the following simple groups:

(i) PSL$(2,p)$, where p is a prime with $p > 11$, $p^2 - 1 \not\equiv o(5)$ and $p^2 - 1 \not\equiv o(16)$;

(ii) PSL$(2,2^q)$, where q is a prime;

(iii) PSL$(2,3^q)$, where q is an odd prime.

Then a maximal of G is one of the following groups:

(i) A dihedral group of order $2(n \pm 1)/d$, where $d = (n-1,2)$, and $n = p$ or $n = 2^q$ or $n = 3^q$.

(ii) A Frobenius group of order $n(n-1)/d$.

(iii) A_4, where A_4 is the alternating group on four letters.

3. Groups all of whose maximal subgroups are MNP-groups

In this section we study the structure of groups all of whose maximal subgroups are MNP-groups. Let $|\pi(G)|$ be the number of distinct prime divisors of the order of G. We prove the following theorem

THEOREM 3.1. *If G is a non MNP-group and each of its maximal subgroups is an MNP-group, then $|\pi(G)| = 2$ and one of the following statements is true:*

(i) *G is supersolvable*

(ii) *G has a normal Sylow p-subgroup P and a non-normal cyclic Sylow q-subgroup Q, where $p < q$ and $P/\Phi(P)$ is a minimal normal subgroup of $G/\Phi(P)$.*

(iii) *G has a normal Sylow q-subgroup Q and a non-normal cyclic Sylow p-subgroup P, where $p < q$, and $Q/\Phi(Q)$ is a minimal normal subgroup of $G/\Phi(Q)$.*

PROOF. By (2.1), all maximal subgroups of G are supersolvable. Then G is solvable by [5, Satz 9.6, p. 718]. By (2.2), G has a Sylow basis. Let H be a maximal subgroup of a Sylow subgroup of G such that H is not normal in G and let $\{P_1, P_2, \ldots, P_n\}$ be a Sylow basis of G with, say, $H < P_1$. If $|\pi(G)| \geq 3$, then P_1P_i is a proper subgroup of G, where $i = 2, 3, \ldots, n$, and $\langle P_1, P_2, \ldots, P_n\rangle = G \leq N_G(H)$, a contradiction. Therefore $|\pi(G)| = 2$.

Assume that (i) is false. We must show that either (ii) or (iii) must hold. Clearly, G is a minimal non-supersolvable group. Let P be a Sylow p-subgroup of G and let Q be a Sylow q-subgroup of G, where $p < q$. Suppose that G does not have an ordered Sylow tower. Then, by (2.3(i)) G is a minimal non-nilpotent group. It follows from (2.4) that $P \triangleleft G$ and Q is a non-normal cyclic Sylow q-subgroup of G, where $p < q$, and $P/\Phi(P)$ is a minimal normal subgroup of $G/\Phi(P)$. So (ii) holds. Now suppose that G has an ordered Sylow tower. Then $Q \triangleleft G$. By (2.3(iii)) $Q/\Phi(Q)$ is a minimal normal subgroup of $G/\Phi(Q)$. We argue that P is cyclic. If not, there exist two distinct maximal subgroups H_1 and H_2 of P. Let K be a maximal subgroup of Q. Then $K \triangleleft QH_1$ and $K \triangleleft QH_2$, by hypothesis. Since $|\pi(G)| = 2$, and $\langle QH_1, QH_2\rangle \leq N_G(K)$, it follows that $K \triangleleft G$. Hence each maximal subgroup of Q is normal in G. Let $\{Q_1, Q_2, \ldots, Q_s\}$ be the set of all maximal subgroups of Q. Then G/Q_i is supersolvable, where

$i = 1, 2, \ldots, s$. Since $G/\Phi(Q) \stackrel{\sim}{<} G/Q_1 \times G/Q_2 \times \ldots \times G/Q_s$, it follows that $G/\Phi(Q)$ is supersolvable. Hence G is supersolvable by [5, Satz 8.6a, p. 713], a contradiction. Thus P is cyclic. Hence (iii) holds.

As an immediate corollary we have:

COROLLARY 3.2. If each maximal subgroup of G is an MNP-group, then G' is nilpotent. Further, G is an MNP-group when $|\pi(G)| \geq 3$.

PROOF. Suppose that G is an MNP-group. Then Theorem (2.1) implies that G is supersolvable and hence G' is nilpotent by [3, Theorem 7.2.13, p. 157]. Now suppose that G is a non-MNP-group. Then Theorem (3.1) implies that one of statements (i), (ii), (iii) must hold. If G is as in (i), then G' is nilpotent. If G is as in (ii) or (iii), then G' is a group of prime power order and so G' is nilpotent. Clearly, Theorem (3.1) implies that G is an MNP-group, when $|\pi(G)| \geq 3$.

4. Simple groups all of whose second maximal subgroups are MNP-groups

In this section we prove several results needed for the classification of non-abelian simple groups all of whose second maximal subgroups are MNP-groups.

We need the following lemmas:

LEMMA 4.1. If G is a non-abelian simple group each of whose second maximal subgroups is an MNP-group, then

(i) $G \neq \mathrm{PSL}(3,3)$;

(ii) $G \neq \mathrm{Sz}(2^q)$, where q is an odd prime;

(iii) $G \neq \mathrm{PSL}(2,p)$, where p is a prime with $p > 11$, $p^2 - 1 \not\equiv o(5)$ and $p^2 - 1 \equiv o(16)$;

(iv) $G \neq \mathrm{PSL}(2,2^q)$, where q is a prime and $2^q - 1 \neq$ prime.

(v) $G \neq \mathrm{PSL}(2,3^q)$, where q is an odd prime and $3^q - 1 \neq 2r$, where r is an odd prime.

PROOF. (i) If $G = \mathrm{PSL}(3,3)$, then G contains a subgroup H isomorphic to S_4, where S_4 is the symmetric group on four letters, by (2.5). Let M be a maximal subgroup of G such that $H \leq M < G$. By hypothesis, all maximal subgroups of M are MNP-groups. Then Corollary 3.2 implies that M' is nilpotent and so H' is nilpotent and this is a contradiction, as A_4 is a non-nilpotent group, where A_4 is the alternating group on four letters. Therefore, $G \neq \mathrm{PSL}(3,3)$.

(ii) If $G = \mathrm{Sz}(r)$, where $r = 2^q$ and q is an odd prime, then G is a Zassenhaus group of degree $r^2 + 1$ by (2.7). Let N be the subgroup fixing a letter. Then we have that N is a maximal subgroup of G and that N is a Frobenius group with kernel K of order r^2 and a cyclic complement H by (2.7). Since K is non-abelian, we have that $Z(K)\ H$ is a proper subgroup of N. Let L be a maximal subgroup of N such that $Z(K)\ H \leq L < N$. Now L is an MNP-group by hypothesis. Then L is supersolvable by (2.1). This is a contradiction as $C_N(y) \leq K$ for all $y \in K^{\#}$. Therefore, $G \neq \mathrm{Sz}(2^q)$, where q is an odd prime.

(iii) If $G = \mathrm{PSL}(2,p)$, where p is a prime with $p > 11$, $p^2 - 1 \not\equiv o(5)$ and $p^2 - 1 \equiv o(16)$, then S_4 is a subgroup of G. Hence, as in part (i), we obtain a contradiction.

(iv) $G = \mathrm{PSL}(2,2^q)$, q is a prime and $2^q - 1 \neq$ prime. Then G is a Zassenhaus group of degree $2^q + 1$. Let N be the subgroup fixing a letter. Then N is a maximal subgroup of G. Further N is a Frobenius group with kernel K of order 2^q and a complement H of order $2^q - 1$. Since $2^q - 1 \neq$ prime, it follows that there exists a maximal subgroup L of N such that $K < L < N$. Then L is an MNP-group by hypothesis and so L is supersolvable. This is a contradiction as $C_N(y) \leq K$ for all $y \in K^{\#}$. Therefore, $G \neq \mathrm{PSL}(2,2^q)$, when q is a prime, and $2^q - 1 \neq$ prime.

(v) Since $3^{2q} - 1 \not\equiv o(16)$ and $(3^q - 1)/2 \neq$ odd prime, it follows that $3^q - 1 = 4m$ or $3^q - 1 = 2m$, , where m is a composite odd integer. Let r be a prime divisor of m. Let N be the subgroup fixing a letter. Then N is a maximal subgroup of G. Further, N is a Frobenius group with kernel K of order 3^q and a complement H of order $2m$ or m. Clearly, N has a subgroup which is the semidirect product of K by $\langle x \rangle$, where $|\langle x \rangle| = r$. Let L be a maximal subgroup of N such that $K\langle x \rangle \leq L < N < G$. By hypothesis, L is an MNP-group. Then, by (2.1), L is supersolvable and this is contradiction as $C_N(y) \leq K$ for all $y \in K^{\#}$. Therefore $G \neq \mathrm{PSL}(2,3^q)$, where q is an odd prime and $(3^q - 1)/2 \neq$ odd prime.

LEMMA 4.2. Let G be a non-nilpotent dihedral group of order $4r^J$, where $J \geq 1$ and r is an odd prime. If all maximal subgroups of G are MNP-groups, then $J = 1$.

PROOF. Assume that $J \geq 2$. Let R be a Sylow r-subgroup of G. Then $R = \langle x \rangle \lhd G$. Let R_1 be a maximal subgroup of R. Since R_1 char$R$ and $R \lhd G$, we have $R_1 \lhd G$ and so SR_1 is a maximal subgroup of G, where S is a Sylow 2-subgroup of G. Then SR_1 is an MNP-group by hypothesis. Hence $S \lhd SR_1$ and so SR_1 is a nilpotent subgroup of G. Now it follows that S is an r'-group of automorphisms of R which acts trivially of $\Omega_1(R)$.

So (2.8) implies that S acts trivially on R and this is a contradiction as G is a non-nilpotent group. Therefore, $J = 1$.

The following lemma may be considered as an improvement of Lemma 4.2.

LEMMA 4.3. Let G be a non-nilpotent dihedral group of order $4m$, where m is an odd integer. If all maximal subgroups of G are MNP-groups, then m is an odd prime.

PROOF. If $m = r^J$, where r is an odd prime, then $J = 1$ by Lemma (4.2). Thus we need only consider the case in which m is divisible by at least two distinct primes. Let S be a Sylow 2-subgroup of G. Since G is a solvable group, it follows that G has a Sylow basis. Let $\{S, P_2, \dots, P_n\}$ be a Sylow basis of G. If $|\pi(G)| \geq 3$, then SP_i are proper subgroups of G, where $i = 2, 3, \dots, n$. Clearly, SP_i are MNP-groups, where $i = 2, 3, \dots, n$. Then $S \triangleleft SP_i$, where $i = 2, 3, \dots, n$, and so $S \triangleleft G$. This is a contradiction as G is a non-nilpotent group.

We also need the following Lemma.

LEMMA 4.4. Suppose that G is one of the following groups

(a) PSL$(2, p)$, where p is a prime with $p > 11$, $p^2 - 1 \not\equiv o(5)$, $p^2 - 1 \not\equiv$ $\not\equiv o(16)$, $p - 1$ is square free and $p + 1 = 2^2 s$, where s is an odd prime, or $p - 1 = 2^2 r$, where r is an odd prime.

(b) PSL$(2, 2^q)$, where q is a prime with $2^q - 1 = r$ and r is an odd prime.

(c) PSL$(2, 3^q)$, where q is an odd prime with $3^q - 1 = 2r$, $3^q + 1 = 2^2 s$, and r and s are odd primes.

Then each second maximal subgroup of G is an MNP-group.

PROOF. By (2.9), a maximal subgroup of G is one of the following groups:

(i) A dihedral group of order $2(n \pm 1)/d$, where $d = (n - 1, 2)$, and $n = p$ or $n = 2^q$ or $n = 3^q$;

(ii) A Frobenius group N with elementary abelian kernel K of order n and a cyclic complement H of order $(n - 1)/d$, where $d = (n - 1, 2)$ and $n = p$ or $n = 2^q$ or $n = 3^q$;

(iii) A_4.

Clearly, the maximal subgroups of the groups of type (i) are MNP-groups.

If G is of type (a), then $|N| = p(p - 1)/2$. Since $p - 1$ is square free or $p - 1 = 2^2 r$, where r is an odd prime, we have that N is of square free order and so all proper subgroups of N are MNP-groups. If G is of type (b), then $|N| = 2^q(2^q - 1) = 2^q r$, where r is an odd prime. Since H acts irreducibly on K, it follows that N is a minimal non-abelian group and so all maximal subgroups of N are MNP-groups. If G is of type (c), then $|N| = 3^q(3^q - 1)/2$.

Since $(3^q - 1)/2 = r$, where r is an odd prime, it follows that $|N| = 3^q r$. Now, as before, N is a minimal non-abelian group. So all maximal subgroups of N are MNP-groups. Therefore, all maximal subgroups of type (ii) are MNP-groups.

Clearly, all maximal subgroups of A_4 are MNP-groups.

Now we can prove the following Theorem:

THEOREM 4.5. *Let G be a non-abelian simple group with the property that all its second maximal subgroups are* MNP*-groups. Then G is one of the following types:*

(a) PSL$(2,p)$, *p is a prime with $p > 11$, $p^2 - 1 \not\equiv o(5)$, $p^2 - 1 \not\equiv o(16)$, $p-1$ is square free and $p+1 = 2^2 s$ or $p-1 = 2^2 r$, where r and s are odd primes.*

(b) PSL$(2,2^q)$, *q is a prime and $2^q - 1 = r$, where r is an odd prime.*

(c) PSL$(2,3^q)$, *q is an odd prime, $3^q - 1 = 2r$ and $3^q + 1 = 2^2 s$, where r and s are odd primes.*

PROOF. by Corollary 3.2, all maximal subgroups of G are solvable. Hence we can apply (2.6) and conclude that G is one of the following types

(i) PSL$(3,3)$

(ii) PSL$(2,p)$, where p is prime with $p > 3$ and $p^2 - 1 \not\equiv o(5)$;

(iii) PSL$(2,2^q)$, where q is a prime;

(iv) PSL$(2,3^q)$, where q is an odd prime;

(v) The Suzuki group Sz(2^q), where q is an odd prime.

G cannot be PSL$(3,3)$ by (4.1(i)).

G cannot be PSL$(2,p)$, where $p > 11$, $p^2 - 1 \not\equiv o(5)$ and $p^2 - 1 \equiv o(16)$ by (4.1(iii)). Hence $p^2 - 1 \not\equiv O(16)$. Thus Sylow 2-subgroups of PSL$(2,p)$ have order 4. Using that PSL$(2,p)$ has a Frobenius subgroup N with kernel K of order p and a complement H of order $(p-1)/2$ which acts irreducibly on K, and that PSL$(2,p)$ has dihedral subgroup of order $p \pm 1$, we obtain that either $p-1$ is square free and $p+1 = 2^2 s$ or $p-1 = 2^2 r$ where s, r are odd integers. By (4.3), r and s are odd primes. So the only possibilities for G are $A_5 \cong$ PSL$(2,5) \cong$ PSL$(2,4)$ and PSL$(2,p)$, $p > 11$ with $p^2 - 1 \not\equiv o(5)$, $p^2 - 1 \not\equiv o(16)$, $p-1$ is square free and $p+1 = 2^2 s$ or $p-1 = 2^2 r$, where r and s are odd primes.

G cannot be $\mathrm{PSL}(2, 2^q)$, where q is a prime and $2^q - 1 \neq r$, where r is a prime by (4.1(iv)). So the only possibility for G is $\mathrm{PSL}(2, 2^q)$, where q is a prime and $2^q - 1$ is an odd prime.

G cannot be $\mathrm{PSL}(2, 3^q)$, where q is an odd prime and $3^q - 1 \neq 2r$, where r is an odd prime by (4.1(v)). Hence $3^q - 1 = 2r$, where r is an odd prime.

Since $3^{2q} - 1 \not\equiv o(16)$, we have $4/|G|$ and $8 \nmid |G|$. G has a dihedral subgroup of order $3^q + 1$. Then $3^q + 1 = 2^2 s$, where s is an odd integer. By (4.3), s is an odd prime. So the only possibility for G is $\mathrm{PSL}(2, 3^q)$, where q is an odd prime, $3^q - 1 = 2r$ and $3^q + 1 = 2^2 s$, where r and s are odd primes.

G cannot be $\mathrm{Sz}(2^r)$ by (4.1(ii)). But we have seen in the analysis of (a), (b) and (c) that all second maximal subgroups of G are MNP-groups; see (4.4).

The theorem is proved.

References

[1] GARY L. WALLS, Groups with maximal subgroups of Sylow subgroups normal, *Isr. J. Math.*, **43** (1982), 166–168

[2] S. SRINIVASAN, Two sufficient conditions for supersolvability of finite groups, *Isr. J. Math.*, **35** (1980), 210–214.

[3] W. R. SCOTT, *Group Theory*, Prentice-Hall, Englewood Cliffs, N. J., 1964.

[4] K. DOERK, Minimal nicht überauflösbare, endliche Gruppen. *Math. Z.*, **91** (1966), 198–205.

[5] B. HUPPERT, *Endliche Gruppen I.*, Springer, Berlin–Heidelberg–New York, 1967.

[6] A FATTAHI, On generalizations of Sylow tower groups, *Pacific J. Math.*, **45** (1973), 453–478.

[7] J. G. THOMPSON, Nonsolvable finite groups all of whose local subgroups are solvable, *Bull. Amer. Math. Soc.*, **74** (1968), 383–437.

[8] D. GORENSTEIN, *Finite Groups*, McGraw-Hill, New York, 1968.

[9] B. HUPPERT and N. BLACKBURN, *Finite Groups III.*, Springer, Berlin–Heidelberg–New York, 1982.

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 31–34

REMARKS ON ADDITIVE SET FUNCTIONS

By

Á. CSÁSZÁR

Department of Analysis, Eötvös Loránd University, Budapest

(Received March 13, 1992)

1. In a recent paper [1], the author has examined the following problem: Let $\psi : \mathcal{A} \to \mathbf{R}$ be a set function defined on a system of sets $\mathcal{A}$ belonging to a certain class of set systems; under which conditions possesses φ an additive extension $\varphi' : \mathcal{R} \to \mathbf{R}$ to the ring $\mathcal{R}$ generated by $\mathcal{A}$? The purpose of the present paper is to give some corrections and supplements to [1].

In [1], a known result was quoted for the case when $\mathcal{A}$ is a lattice $\mathcal{H}$, namely that a necessary and sufficient condition for the existence of such an extension is that φ should satisfy

$$\varphi(A \cup B) = \varphi(A) + \varphi(B) - \varphi(A \cap B) \quad \text{for} \quad A, B \in \mathcal{H}. \tag{1}$$

Unfortunately, this is not true in this form; in fact, an obvious necessary condition is

$$\varphi(\emptyset) = 0 \quad \text{if} \quad \emptyset \in \mathcal{H}, \tag{2}$$

and (2) does not follow from (1) (let $\mathcal{H} = \{\emptyset\}$, $\varphi(\emptyset) = 1$).

However, in the case $\emptyset \in \mathcal{H}$, (1) and (2) are sufficient for the existence of an additive extension to the generated ring: the method described in [2], VI.3.h.34 shows that the sets $A - B$ $(A, B \in \mathcal{H})$ constitute a semi-ring $\mathcal{P}$ and

$$\varphi'(A - B) = \varphi(A) - \varphi(A \cap B)$$

defines according to (1) an additive set function φ' on $\mathcal{P}$ that can be extended to a ring, and (2) guarantees $\mathcal{H} \subset \mathcal{P}$, $\varphi'|\mathcal{H} = \varphi$. In the case $\emptyset \notin \mathcal{H}$, it is described in [1], p. 21 that $\emptyset$ can be joined to $\mathcal{H}$ in a manner that (1) remains valid for $\mathcal{H}' = \mathcal{H} \cup \{\emptyset\}$ when $\varphi(\emptyset)$ is defined to be equal to 0.

In order to deal with the case when $\mathscr{A}$ is a meet-semi-lattice, [1] introduces the expression

$$\psi(A_1,\dots,A_n)=\sum_{k=1}^{n}(-1)^{k+1}\sum_{I\in\Gamma_k}\varphi\left(\bigcap_{i\in I}A_i\right),$$

where Γ_k is the collection of all subsets of $\{1,\dots,n\}$ with cardinality k. It is shown in [1], 1.2 that if φ has an additive extension to a ring then

(3) for $A_i\in\mathscr{A}$, $n\in\mathrm{N}$, the value of $\psi(A_1,\dots,A_n)$ depends only on the union $\bigcup_1^n A_i$,

and it is claimed in [1], that (3) suffices for the existence of an additive extension to a ring. In fact, the proof of [1], 1.4 shows that then φ can be extended to the lattice $\mathscr{H}$ generated by $\mathscr{A}$ in a manner that the extension fulfils (1). However, the above remarks imply that this assures the existence of an additive extension to a ring in the case $\emptyset\notin\mathscr{A}$ only; in the case $\emptyset\in\mathscr{A}$, (3) must be completed by the condition $\varphi(\emptyset)=0$.

In the case of a join-semi-lattice $\mathscr{A}$, [1], 1.3 and 1.5 formulate as a necessary and sufficient condition that

(4) $\chi(A_1,\dots,A_n)$ depends only on $\bigcap_1^n A_i$ for $A_i\in\mathscr{A}$, $n\in\mathrm{N}$,

where

$$\chi(A_1,\dots,A_n)=\sum_{k=1}^{n}(-1)^{k+1}\sum_{I\in\Gamma_k}\varphi\left(\bigcup_{i\in I}A_i\right).$$

It is clear that the condition

(5) $$\chi(A_1,\dots,A_n)=0 \quad\text{if}\quad A_1\in\mathscr{A},\ \bigcap_1^n A_i=\emptyset$$

has to be added to (4) in order to assure the existence of an additive extension to a ring.

2. In the case of a meet-semi-lattice $\mathscr{A}$, the condition (3) can be replaced by a simpler one:

THEOREM 1. *If $\mathscr{A}$ is a meet-semi-lattice then* (3) *is equivalent to the condition*

(6) $$\varphi\left(\bigcup_1^n A_i\right)=\psi(A_1,\dots,A_n)\quad\textit{if}\quad A_i\in\mathscr{A}\ (i=1,\dots,n)\ \textit{and}\ \bigcup_1^n A_i\in\mathscr{A}.$$

PROOF. (3) implies $\psi(A) = \psi(A_1, \dots, A_n)$ for $A = \bigcup_1^n A_i \in \mathcal{A}$, $A_i \in \mathcal{A}$, and clearly $\psi(A) = \varphi(A)$. Conversely, (6) implies

$$(7) \qquad \psi(A_1, \dots, A_n, B) = \psi(A_1, \dots, A_n) \quad \text{if} \quad A_i, B \in \mathcal{A}, \quad B \subset \bigcup_1^n A_i.$$

In fact, we have $B = \bigcup_1^n B_i$ for $B_i = B \cap A_i \in \mathcal{A}$, whenever A_i, $B \in \mathcal{A}$, $B \subset \bigcup_1^n A_i$, and it is easy to see that

$$(8) \qquad \psi(A_1, \dots, A_n, B) = \psi(A_1, \dots, A_n) + \varphi(B) - \psi(B_1, \dots, B_n),$$

because $\emptyset \neq I \subset \{1, \dots, n\}$ implies

$$\varphi\left(\bigcap_{i \in I} A_i \cap B\right) = \varphi\left(\bigcap_{i \in I} B_i\right).$$

By (6) $\varphi(B) = \psi(B_1, \dots, B_n)$ and we obtain (7). An iterated application of the latter furnishes (3), i.e.

$$\psi(A_1, \dots, A_n) = \psi(A_1, \dots, A_n, B_1, \dots, B_m) = \psi(B_1, \dots, B_m)$$

whenever $A_i, B_j \in \mathcal{A}$, $\bigcup_1^n A_i = \bigcup_1^m B_j$. □

The checking of the validity of (6) is simpler than that of (3) in consequence of

COROLLARY 2. *(6) is automatically fulfilled if* $\bigcup_1^n A_i$ *coincides with one of the sets* $A_i \in \mathcal{A}$.

PROOF. For $A = \bigcup_1^{n-1} A_i \in \mathcal{A}$, we have by (8)

$$\psi(A_1, \dots, A_{n-1}, A) = \psi(A_1, \dots, A_{n-1}) + \varphi(A) - \psi(A_1, \dots, A_{n-1}),$$

whenever A_i, $A \in \mathcal{A}$, and $B_i = A \cap A_i = A_i$ $(i = 1, \dots, n-1)$ in (8). □

Thus it suffices to check the validity of (6) in the cases only when each A_i is a proper subset of $\bigcup_1^n A_i$. It can happen that no such sets A_i exist at all in $\mathcal{A}$ and then (6) is always fulfilled (let $\mathcal{A} = \{\emptyset, \{a\}, \{b\}, \{a, b, c\}\}$ for distinct elements a, b, c).

For join-semi-lattices, we can prove in a similar way:

THEOREM 3. *If $\mathcal{A}$ is a join-semi-lattice, then* (4) *is equivalent to*

$$\varphi\left(\bigcap_1^n A_i\right) = \chi(A_1, \dots, A_n) \quad \text{for} \quad A_i, \ \bigcap_1^n A_i \in \mathcal{A}. \ \square \tag{9}$$

COROLLARY 4. (9) *is fulfilled if* $\bigcap_1^n A_i$ *coincides with one of the sets* $A_i \in \mathcal{A}$. $\square$

References

[1] Á. CSÁSZÁR, Additive extensions of set functions, *Annales Univ. Sci. Budapest.*, Sectio Math., **34** (1991), 21–26.

[2] Á. CSÁSZÁR, *Valós analízis II.*, Budapest, 1984.

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 35–46

ON THE STRONG LAW OF LARGE NUMBERS FOR LOGARITHMICALLY WEIGHTED SUMS*

By

TAMÁS F. MÓRI

Department of Probability Theory and Statistics, Eötvös Loránd University, Budapest

(Received April 30, 1992)

1. Introduction

Recently much effort has been devoted to the study of the so-called almost sure central limit theory. Probabilists are publishing several papers on extensions of classical weak limit theorems in a generalized formulation involving logarithmic average and logarithmic density. As a starting point of these studies, in 1988 BROSAMLER and SCHATTE independently proved the following version of the a.s. central limit theorem.

Suppose $X_1, X_2, \dots$ are i.i.d. random variables with mean 0 and variance 1 (in fact, stronger moment conditions were first required, but later they proved to be superfluous). As usual, let $S_n = X_1 + X_2 + \dots + X_n$. Considering that

$$\lim_{n\to\infty} \mathbf{P}(S_n/\sqrt{n} < x) = \Phi(x) = \int_{-\infty}^{x} (2\pi)^{-1/2} \exp\{-y^2/2\} dy$$

for every x, one can naturally ask whether the (random) sequence of integers n for which $S_n/\sqrt{n} < x$ holds has density a.s. In other words, does $\lim\limits_{n\to\infty} \frac{1}{n} \sum\limits_{i=1}^{n} \mathbf{I}(S_i/\sqrt{i} < x)$ exist with probability 1 (where $\mathbf{I}(.)$ denotes the indicator of the event in parentheses)? The answer is negative, but a weaker statement can still be proved, namely, the integer sequence in question has logarithmic density $\Phi(x)$ with probability 1:

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{i=1}^{n} \frac{1}{i} \mathbf{I}(S_i/\sqrt{i} < x) = \Phi(x) \quad \text{a.s}$$

---

* Research supported by the Hungarian National Foundation for Scientific Research, Grant N° 1905.

This result has been extended and generalized by several authors. A brand new paper of BERKES and DEHLING [1] provides a systematic study of logarithmic analogues of classical limit theorems. They also present a general method which can be applied to all similar problems. It is based on the fact that under very mild growth conditions on the partial sums S_n of an independent (not necessarily identically distributed) sequence $\{X_n\}$ the a.s limit behaviour of the sequences

$$\frac{1}{\log n}\sum_{i=1}^{n}\frac{1}{i}I\left(\frac{S_i-b_i}{a_i}<x\right) \quad \text{and} \quad \frac{1}{\log n}\sum_{i=1}^{n}\frac{1}{i}\mathbf{P}\left(\frac{S_i-b_i}{a_i}<x\right)$$

coincide. More precisely, defining $\xi_i = I\left(\frac{S_i-b_i}{a_i}<x\right) - \mathbf{P}\left(\frac{S_i-b_i}{a_i}<x\right)$ one can write

$$\lim_{n\to\infty}\frac{1}{\log n}\sum_{i=1}^{n}\frac{1}{i}\xi_i = 0 \quad \text{a.s.} \tag{1}$$

(cf. Remark 1 in Section 4).

This limit result shows that the almost sure central limit theorems are *not* stronger than their classical counterparts, as it was first thought to be the case. On the contrary, counterexamples have been found even in the "nice" case of i.i.d. summands where the sums had a limit distribution in the a.s. sense, but not in the ordinary sense [2]. Similar phenomenon was observed in connection with random sequences of other types.

The present paper is motivated by the increasing need of results of type (1) in the study of a.s. weak limits. We try and find conditions on the sequence $\{\xi_n, n \geq 1\}$ that are sufficient to guarantee (1). Although in the above mentioned applications the ξ_n are always bounded, it was quite natural to extend our attention to the unbounded case which proved to be even more interesting.

2. Strong laws of large numbers

Define $\ell(x) = \log x$ for $x \geq e$ and $\ell(x) = 1$ for $x < e$. Let $\ell_1(x) = \ell(x)$ and $\ell_k(x) = \ell(\ell_{k-1}(x))$ for $k \geq 2$.

Our main result is essentially an improved version of Serfling's strong law of large numbers [6].

THEOREM 1. *Let $\xi_1, \xi_2, \ldots$ be arbitrary random variables with finite variances. Suppose there exist a positive non-increasing function $h(.)$ on* *the positive numbers and a positive integer m such that*

$$\int_1^\infty h(z)\frac{\ell_m(z)}{z\ell(z)}dz<\infty \tag{2}$$

and

$$|\mathbf{E}(\xi_i\xi_j)|\le h(j/i) \quad \textit{for all } 1\le i\le j. \tag{3}$$

Then

$$\lim_{n\to\infty}\frac{1}{\ell(n)}\sum_{i=1}^n\frac{1}{i}\xi_i=0 \quad \text{a.s.} \tag{4}$$

If, in addition, the random variables $\{\xi_n, n\ge 1\}$ *are uniformly bounded,* (2) *can be weakened to require that*

$$\int_1^\infty \frac{h(z)}{z\ell(z)}dz<\infty \tag{2'}$$

PROOF. For arbitrary positive number $t>1$ let us introduce $\eta(t)= \sum_{1\le i<t}\frac{1}{i}\xi_i$. We are going to apply the standard method of subsequences (see e.g. the proof of Theorem 3.7.3. of [7]). Choosing a sufficiently sparse subsequence $\{N_k, k\ge 1\}$ we first show that $\mathbf{E}(\sum_k \eta(N_k)^2/\ell(N_k)^2)<\infty$, from which $\eta(N_k)/\ell(N_k)\to 0$ follows with probability 1. Then we only have to check if $\lim_{k\to\infty}(\max_{N_k<n<N_{k+1}}|\eta(n)-\eta(N_k)|)/\ell(N_k)=0$ a.s.

Let $a=(a_k, k\ge 1)$ be an increasing sequence of positive numbers. For estimating the expectation of $\mu^2(s,t\mid a)=\max_{t\le a_k<s}(\eta(a_k)-\eta(s))^2$, $1\le s<t$, one can use Serfling's maximal inequality [5]. Let us define $g(s,t)= 2\sum_{s\le i\le j<t}\frac{1}{ij}h\left(\frac{j}{i}\right)$. Then we clearly have

$$\mathbf{E}(\eta(t)-\eta(s))^2=\sum_{s\le i<t}\sum_{s\le j<t}\frac{1}{ij}\mathbf{E}(\xi_i\xi_j)\le g(s,t), \tag{5}$$

and for every $1\le s<t<u$

$$g(s,t)+g(t,u)\le g(s,u).$$

Hence it follows that

$$\mathbf{E}(\mu^2(s,t\mid a))\le 6\ell(\nu(s,t\mid a))^2 g(s,t), \tag{6}$$

where $\nu(s,t \mid a)$ stands for the number of a_k satisfying $s \le a_k < t$.

Let us estimate $g(s,t)$. For $1 \le i < j$ let D_{ij} denote the parallelogram with vertices (i,j), $(i,j+1)$, $(i+1,j+1)$ and $(i+1,j+2)$, i.e.,

$$D_{ij} = \{(x,y) : j \le x < j+1, j-i-1 < y-x \le j-i\}.$$

It is easy to see that for all $(x,y) \in D_{ij}$

$$\frac{1}{xy}h\left(\frac{y}{x}\right) \ge \frac{1}{(i+2)(j+1)}h\left(\frac{j}{i}\right) \ge \frac{1}{6ij}h\left(\frac{j}{i}\right),$$

hence

$$g(s,t) \le \sum_{s\le i<j<t} 12\int_{D_{ij}} \frac{1}{xy}h\left(\frac{y}{x}\right)dxdy + \sum_{s\le i<t}\frac{2}{i^2}h(1)$$

$$\le 12 \int_{\{s\le x<y<t+1\}} \frac{1}{xy}h\left(\frac{y}{x}\right)dxdy + \frac{4h(1)}{s}.$$

Here

$$\int_{\{s\le x<y<t+1\}} \frac{1}{xy}h\left(\frac{y}{x}\right)dxdy \le \int_{\{s\le x<t+1,\, 1\le \frac{y}{x}<\frac{t+1}{s}\}} \frac{1}{xy}h\left(\frac{y}{x}\right)dxdy.$$

Substituting $z = y/x$ we obtain for the right-hand side that

$$= \int_s^{t+1}\frac{dx}{x}\int_1^{\frac{t+1}{s}}\frac{h(z)}{z}dz = \log\left(\frac{t+1}{s}\right)\int_1^{\frac{t+1}{s}}\frac{h(z)}{z}dz \le$$

$$\le 2\left(\log\left(\frac{t}{s}\right)+\frac{1}{t}\right)\int_1^{t/s}\frac{h(z)}{z}dz \le 2\log\left(\frac{t}{s}\right)\int_1^{t/s}\frac{h(z)}{z}dz + \frac{2h(1)}{s}.$$

Thus

$$g(s,t) \le 24\log\left(\frac{t}{s}\right)\int_1^{t/s}\frac{h(z)}{z}dz + \frac{6h(1)}{S}. \tag{7}$$

Particularly, from (5) and (7) it follows that

$$\mathbf{E}\left(\sum_k \frac{\eta(N_k)^2}{\ell(N_k)^2}\right) \le \sum_k \frac{24}{\ell(N_k)}\int_1^{N_k}\frac{h(z)}{z}dz + \sum_k \frac{6h(1)}{\ell(N_k)^2} = \tag{8}$$

$$= \int_1^\infty \frac{h(z)}{z} \left(\sum_{N_k > z} \frac{24}{\ell(N_k)} \right) dz + \sum_k \frac{6h(1)}{\ell(N_k)^2}.$$

Similarly, if $N_{k+1}/N_k \to \infty$, then by (6) and (7) we get

$$\mathbf{E}\left(\sum_k \frac{1}{\ell(N_k)^2} \mu^2(N_k, N_{k+1} \mid a) \right) \leq \tag{9}$$

$$\leq \text{const} \cdot \sum_k \frac{\ell(\nu(N_k, N_{k+1} \mid a))^2}{\ell(N_k)^2} \left[\ell\left(\frac{N_{k+1}}{N_k} \right) \int_1^{N_{k+1}/N_k} \frac{h(z)}{z} dz + \frac{1}{N_k} \right] \leq$$

$$\leq \text{const} \cdot \int_1^\infty \frac{h(z)}{z} \left[\sum_{k: N_{k+1}/N_k > z} \frac{\ell(\nu(N_k, N_{k+1} \mid a))^2}{\ell(N_k)^2} \ell\left(\frac{N_{k+1}}{N_k} \right) \right] dz +$$

$$+ \text{const} \cdot \sum_k \frac{\ell(N_{k+1})^2}{N_k \ell(N_k)^2}.$$

Let us first deal with the case of not necessarily bounded random variables. We first note that (2) can be replaced equivalently with the following condition:

$$\int_1^\infty h(z) \frac{\ell_m(z)^2}{z\ell(z)} dz < \infty \tag{10}$$

for some integer $m > 1$. Let us introduce

$$N_k = \exp\left(\exp\left(\frac{k}{\ell_m(k)^2} \right) \right) \quad \text{and} \quad a_k = \exp\left(\exp\left(\frac{k}{\ell_{m-1}(k)^2} \right) \right).$$

Then

$$\nu(N_k, N_{k+1} \mid a) \sim \left(\frac{\ell_{m-1}(k)}{\ell_m(k)} \right)^2 \quad \text{and} \quad \ell\left(\frac{N_{k+1}}{N_k} \right) \sim \frac{\ell(N_k)}{\ell_m(k)^2}.$$

Hence

$$\sum_{k: N_{k+1}/N_k > z} \frac{\ell(\nu(N_k, N_{k+1} \mid a))^2}{\ell(N_k)^2} \ell\left(\frac{N_{k+1}}{N_k} \right) =$$

$$= O\left(\sum_{\ell(N_k) > \ell_m(k)^2 \ell(z)} \frac{1}{\ell(N_k)} \right) = O\left(\frac{1}{\ell(z)} \right)$$

as $z \to \infty$. In addition,

$$\sum_k \frac{\ell(N_{k+1})^2}{N_k \ell(N_k)^2} < \infty.$$

Thus by (9)

$$\mathbf{E}\left(\sum_k \frac{1}{\ell(N_k)^2}\mu^2(N_k, N_{k+1} \mid a)\right) \le \text{const} \cdot \int_1^\infty \frac{h(z)}{z\ell(z)} dz < \infty.$$

This implies that

$$\frac{1}{\ell(N_k)}\mu(N_k, N_{k+1} \mid a) \to 0 \quad \text{a.s.},$$

and since

$$\max\left\{\left|\frac{\eta(a_i)}{\ell(a_i)}\right| : N_k \le a_i < N_{k+1}\right\} \le \left|\frac{\eta(N_k)}{\ell(N_k)}\right| + \frac{1}{\ell(N_k)}\mu(N_k, N_{k+1} \mid a),$$

we obtain that

$$\text{(11)} \qquad \limsup_{k\to\infty} \left|\frac{\eta(a_k)}{\ell(a_k)}\right| \le \limsup_{k\to\infty} \left|\frac{\eta(N_k)}{\ell(N_k)}\right| \quad \text{with probability 1.}$$

Iterating this procedure one can conclude that (11) holds even for $a_k = \exp\left(\exp\left(\frac{k}{\ell(k)^2}\right)\right)$.

Now, let us substitute $N_k = \exp\left(\exp\left(\frac{k}{\ell_m(k)^2}\right)\right)$ into the estimation obtained in (8) for $\mathbf{E}\left(\sum_k \frac{\eta(N_k)^2}{\ell(N_k)^2}\right)$. Routine calculation shows that

$$\sum_{N_k > z} \frac{1}{\ell(N_k)} = O\left(\frac{\ell_m(z)^2}{\ell(z)}\right),$$

hence

$$\mathbf{E}\left(\sum_k \frac{\eta(N_k)^2}{\ell(N_k)^2}\right) \le \text{const} \cdot \int_1^\infty h(z)\frac{\ell_m(z)^2}{z\ell(z)} dz < \infty.$$

Consequently, $\lim_{k\to\infty} \frac{\eta(N_k)}{\ell(N_k)} = 0$ and therefore $\lim_{k\to\infty} \frac{\eta(a_k)}{\ell(a_k)} = 0$ a.s. for $a_k = \exp\left(\exp\left(\frac{k}{\ell(k)^2}\right)\right)$.

We have to prove that the fluctuation of the sequence $(\eta(n)/\ell(n), n \ge 1)$ between a_k and a_{k+1} is getting negligible as $k \to \infty$. This will be carried out in three steps.

Firstly, let $b_k = \exp(\exp(k^{1/3}))$, thus $a_k \approx b_{k^3/\log^3 k}$ and $\nu(a_k, a_{k+1} \mid b) = O(k^3)$. In addition, $\ell\left(\frac{a_{k+1}}{a_k}\right) = \exp\left(\frac{k+1}{\ell(k+1)^2}\right) - \exp\left(\frac{k}{\ell(k)^2}\right) \sim \frac{\ell(a_k)}{\ell(k)^2}$, thus

$$\sum_{k: a_{k+1}/a_k > z} \frac{\ell(\nu(a_k, a_{k+1} \mid b))^2}{\ell(a_k)^2} \ell\left(\frac{a_{k+1}}{a_k}\right) =$$

$$= O\left(\sum_{\ell(a_k) > \ell(k)^2 \ell(z)} \frac{1}{\ell(a_k)}\right) = O\left(\frac{1}{\ell(z)}\right)$$

as $z \to \infty$. Hence by (9) we have

$$\mathbf{E}\left(\sum_k \frac{1}{\ell(a_k)^2} \mu^2(a_k, a_{k+1} \mid b)\right) \leq \text{const} \cdot \int_1^\infty \frac{h(z)}{z\ell(z)} dz < \infty.$$

Consequently,

$$\frac{1}{\ell(a_k)} \mu(a_k, a_{k+1} \mid b) \to 0 \quad \text{a.s.}$$

implying

$$\lim_{k \to \infty} \frac{\eta(b_k)}{\ell(b_k)} = 0 \quad \text{a.s.} \tag{12}$$

Secondly, let $c_k = \exp(k^{1/3})$, thus $b_k \approx c_{\exp(3k^{1/3})}$ and $\nu(b_k, b_{k+1} \mid c) = O(\exp(3k^{1/3}))$. Now $\ell\left(\frac{b_{k+1}}{b_k}\right) \sim \frac{1}{3} k^{-2/3} \ell(b_k)$, therefore by (9)

$$\mathbf{E}\left(\sum_k \frac{1}{\ell(b_k)^2} \mu^2(b_k, b_{k+1} \mid c)\right) \leq$$

$$\leq \text{const} \cdot \int_1^\infty \frac{h(z)}{z} \left[\sum_{k: b_{k+1}/b_k > z} \frac{\ell(\nu(b_k, b_{k+1} \mid c))^2}{\ell(b_k)^2} \ell\left(\frac{b_{k+1}}{b_k}\right)\right] dz +$$

$$+ \text{const} \cdot \sum_k \frac{\ell(b_{k+1})^2}{b_k \ell(b_k)^2} \leq \text{const} \cdot \int_1^\infty \frac{h(z)}{z} \left(\sum_{b_{k+1}/b_k > z} \frac{1}{\ell(b_k)}\right) dz.$$

It is not so hard to see that

$$\sum_{b_{k+1}/b_k > z} \frac{1}{\ell(b_k)} = O\left(\frac{1}{\ell(z)}\right),$$

hence

$$\mathbf{E}\left(\sum_k \frac{1}{\ell(b_k)^2}\mu^2(b_k, b_{k+1} \mid c)\right) < \infty.$$

Together with (12) this implies

$$\lim_{k\to\infty} \eta(c_k)/\ell(c_k) = 0 \quad \text{a.s.} \tag{13}$$

Finally, consider the sequence $\mathbb{N}$ of positive integers. Clearly, $\nu(c_k, c_{k+1} \mid \mathbb{N}) = O(c_k)$. This time $c_{k+1}/c_k \to 1$, thus

$$\ell\left(\frac{c_{k+1}}{c_k}\right) \int_1^{c_{k+1}/c_k} \frac{h(z)}{z} dz = O\left(\left(\frac{c_{k+1}}{c_k} - 1\right)^2\right) = O(k^{-4/3})$$

From the second line of (9) it follows that

$$\mathbf{E}\left(\sum_k \frac{1}{\ell(c_k)^2}\mu^2(c_k, c_{k+1} \mid \mathbb{N})\right) \le \text{const} \cdot \sum_k \left(k^{-4/3} + \frac{1}{c_k}\right) < \infty.$$

Combining this with (12) we obtain that $\lim\limits_{n\to\infty} \eta(n)/\ell(n) = 0$ a.s.

The case of uniformly bounded random variables is much simpler. Let ε be an arbitrarily small positive number and $N_k = \exp(e^{\varepsilon k})$, $k \ge 0$. Since $\sum\limits_k \ell(N_k)^{-2} < \infty$ and

$$\sum_{N_k > z} \frac{1}{\ell(N_k)} \le \frac{1}{(1 - e^{-\varepsilon})\ell(z)},$$

from(8) we obtain that $\eta(N_k)/\ell(N_k) \to 0$ with probability 1. On the other hand, for every integer n between N_k and N_{k+1} we clearly have

$$\left|\frac{\eta(n)}{\ell(n)}\right| \le \left|\frac{\eta(n)}{\ell(N_k)}\right| \le \left|\frac{\eta(n) - \eta(N_k)}{\ell(N_k)}\right| + \left|\frac{\eta(N_k)}{\ell(N_k)}\right|.$$

The first term on the right-hand side can be estimated by

$$C\left(\frac{1}{\ell(N_k)} \sum_{N_k < i \le N_{k+1}} \frac{1}{i}\right), \qquad \text{where } \mathbf{P}(\sup_n |\xi_n| \le C) = 1.$$

This converges to $C(e^{\varepsilon} - 1)$, hence $\limsup\limits_{n\to\infty} |\eta(n)/\ell(n)| \le C(e^{\varepsilon} - 1)$ a.s. for every $\varepsilon > 0$, thus $\lim\limits_{n\to\infty} \eta(n)/\ell(n) = 0$, as claimed. □.

3. Concluding remarks

REMARK 1. The "bounded" part of Theorem 1 can be applied in a.s. central limit theory, as it will be demonstrated by the following example, borrowed from BERKES and DEHLING [1]. In fact, it is a particular case of their Theorem 1.

Let $X_1, X_2, \ldots$ be independent (but not necessarily identically distributed) random variables, $S_n = X_1 + \ldots + X_n$, and $a_n > 0$, b_n $(n \geq 1)$ numerical normalizing sequences.

THEOREM 2. *Suppose*

$$\sup_i \mathbf{E} f\left(\left|\frac{S_i - b_i}{a_i}\right|\right) < \infty, \tag{14}$$

where $f > 0$ is a Borel measurable function on $[0, \infty)$ such that both $f(z)$ and $z/f(z)$ are non-decreasing and

$$\int_1^\infty \frac{dz}{z\ell(z)f(z)} < \infty. \tag{15}$$

Assume in addition that

$$a_j/a_i \geq C(j/i)^\gamma \quad (j \geq i) \tag{16}$$

for some positive constants C and γ. Then for any distribution function G the following statements are equivalent:

(A) *For any Borel set $A \subset \mathbb{R}$ with $G(\partial A) = 0$ we have*

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{i=1}^n \frac{1}{i} \mathbf{I}\left(\frac{S_i - b_i}{a_i} \in A\right) = G(A) \quad \text{a.s.},$$

(B) *For any Borel set $A \subset \mathbb{R}$ with $G(\partial A) = 0$ we have*

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{i=1}^n \frac{1}{i} \mathbf{P}\left(\frac{S_i - b_i}{a_i} \in A\right) = G(A) \quad \text{a.s.}$$

PROOF. We shall trace the proof of the cited theorem of BERKES and DEHLING. First of all, we can suppose $b_n = 0$, since the additive normalization can be included in the random variables X_n. Then we note that the equivalence of (A) and (B) will immediately follow as soon as we show that (4) holds for $\xi_i = g(S_i/a_i) - \mathbf{E}g(S_i/a_i)$, where $g(.)$ is an arbitrary bounded Lipschitz-1 function on $\mathbb{R}$ (we can and will suppose that $0 \le g \le 1$). In order to apply our Theorem 1 we need an estimate of $\mathbf{E}(\xi_i\xi_j) = \operatorname{cov}(g(S_i/a_i), g(S_j/a_j))$. That was derived in [1] in the following way. Using the independence of the summands X_i and the properties of g we can write

$$|\operatorname{cov}(g(S_i/a_i), g(S_j/a_j))| = |\operatorname{cov}(g(S_i/a_i), g(S_j/a_j) - g((S_j - S_i)/a_j))| \le$$
$$\le \mathbf{E}|g(S_j/a_j) - g((S_j - S_i)/a_j)| \le \text{const} \cdot \mathbf{E}(\min\{|S_i/a_j|, 1\}).$$

Using the properties of f we have

$$\min\{|S_i/a_j|, 1\} = \frac{\min\{|S_i/a_i|, a_j/a_i\}}{a_j/a_i} \le \frac{f(\min\{|S_i/a_i|, a_j/a_i\})}{f(a_j/a_i)} \le \frac{f(|S_i/a_i|)}{f(a_j/a_i)}.$$

Hence by (14) and (16)

$$|\operatorname{cov}(g(S_i/a_i), g(S_j/a_j))| \le \text{const} \cdot \mathbf{E}\left[\frac{f(|S_i/a_i|)}{f(a_j/a_i)}\right] \le \text{const} \cdot \frac{1}{f(C(j/i)^\gamma)},$$

i.e., (3) fulfilled with $h(z) = \frac{\text{const}}{f(Cz^\gamma)}$. It is easy to see that the integral in (15) is finite if and only if it is finite with $f(cz^\gamma)$ in place of $f(z)$. Thus Theorem 1 can be applied to complete the proof. □

For further applications see [3].

REMARK 2. Dealing with weighted sums, why do we concentrate on logarithmic weighting instead of extending our results for a much wider class of weighted averages, although the method itself could be applied more generally? This question can also be asked in connection with the a.s. central limit theorem. However, the logarithmic weighting is intrinsic, as it has been pointed out to me by P. MAJOR.

For the sake of simplicity let us consider the example from the Introduction. Let X_1, X_2, ... be i.i.d random variables with mean 0 and variance 1 and $S_n = X_1 + \ldots + X_n$. By the invariance principle $S_i \approx W(i)$, where $W(t)$ denotes a standard Wiener process. Hence

$$\frac{1}{\log n} \sum_{i=1}^{n} \frac{1}{i} \mathbf{I}(S_i/\sqrt{i} < x) \approx \frac{1}{\log n} \int_1^n \mathbf{I}(W(t)/\sqrt{t} < x) \frac{dt}{t}.$$

Substituting $t = e^u$ into the right-hand side we obtain

$$\frac{1}{\log n} \int_0^{\log n} \mathbf{I}(e^{-u/2} W(e^u)) < x)\, du.$$

Since the integrand is stationary, ergodic theory can help. Summarizing what has happened we can see that the logarithmic weighting corresponds to a time-transform $t \mapsto e^t$ which turns the process $W(t)/\sqrt{t}$ stationary. This also explains why j/i appears on the right-hand side of (3): for $j = e^t$ and $i = e^s$ we get j/i as a function of $t - s$ (see also Remark 3 below).

PELIGRAD and RÉVÉSZ [4] has also investigated what other weighting can replace the logarithmic one so that the a.s. central limit theorem be still preserved. It turned out that essential improvement cannot be achieved: though i^{-1} can be multiplied by some logarithmic terms, it cannot be replaced by $i^{-1-\varepsilon}$, say.

REMARK 3. Since $j/i \le j - i + 1$ for $1 \le i \le j$, condition (3) of Theorem 1 can be replaced with the more familiar

$$\mathbf{E}(\xi_i \xi_j) \le h(|j - i|) \quad \text{for all } i, j. \tag{3'}$$

REMARK 4. For those interested in the a.s. convergence of the arithmetic mean of random variables the following assertion can be deduced from Theorem 1.

THEOREM 3. *Let $\xi_1, \xi_2, \ldots$ be arbitrary random variables with finite variances. Suppose there exist a positive non-increasing function $h : [0, +$ $+\infty) \to \mathbb{R}$ and a positive integer m such that*

$$\int_1^\infty \frac{h(z)}{z} \ell_m(z) dz < \infty \tag{17}$$

and

$$|\mathbf{E}(\xi_i \xi_j)| \le h(|j - i|) \quad \textit{for all } i, j. \tag{18}$$

Then

$$\lim_{n \to \infty} n^{-1} \sum_{i=1}^n \xi_i = 0 \quad \text{a.s.} \tag{19}$$

If, in addition, the random variables $\{\xi_n, n \ge 1\}$ are uniformly bounded, (14) can be weakened to require that

$$\int_1^\infty \frac{h(z)}{z} dz < \infty. \tag{17'}$$

PROOF. For $e^{k-1} \le i < e^k$ define $\xi_i' = \xi_k$. Then for every $i \le j$, $e^{k-1} \le i < e^k$, $e^{n-1} \le j < e^n$ we have

$$|\mathbf{E}(\xi_i' \xi_j')| = |\mathbf{E}(\xi_k \xi_n)| \le h(n - k) \le h'(j/i),$$

where $h'(z) = h(\ell(z) - 1)$. Hence (17) resp. (17′) imply

$$\int_e^\infty h'(z)\frac{\ell_{m+1}(z)^2}{z\ell(z)}dz = \int_e^\infty \frac{h(\ell(z)-1)}{\ell(z)}\ell_m(\ell(z))^2\frac{dz}{z} = \int_1^\infty \frac{h(t-1)}{t}\ell_m(t)^2 dt < \infty$$

and

$$\int_e^\infty \frac{h'(z)}{z\ell(z)}dz \le \int_1^\infty \frac{h(t-1)}{t}dt < \infty,$$

resp. Since

$$\frac{1}{\ell(e^k)}\sum_{i=1}^{e^k}\frac{1}{i}\xi_i' \sim \frac{1}{k}\sum_{i=1}^{k}\xi_i,$$

application of Theorem 1 to variables ξ_i' immediately completes the proof. □

This theorem slightly improves Theorem 3.7.4 of [7] on weakly stationary sequences, because the conditions of the latter imply (17) with $m = 1$.

REMARK 5. Is it true that (2′) and (17′) are sufficient even in the non-bounded case? The method applied in the proof of Theorem 1 does not seem to be suitable for this "small" improvement.

References

[1] BERKES, I., DEHLING, H., Some limit theorems in log density, *Ann. Probab.*,

[2] BERKES, I., DEHLING, H., MÓRI, T. F., Counterexamples related to the a.s. central limit theorem, *Studia Sci. Math. Hungar.*, **26** (1991), 153–164.

[3] MÓRI, T. F., The a.s. limit distribution of the longest head run, *Canad. J. Math., to appear.*

[4] PELIGRAD, M., RÉVÉSZ, P., On the almost sure central limit theorem. *In:* Bellow, A. and Jones, R. L., eds., *Almost Everywhere Convergence II.* Academic Press, New York, 1991.

[5] SERFLING, R. J., Moment inequalities for maximum cumulative sum, *Ann. Math. Statist.*, **41** (1970), 1227–1234.

[6] SERFLING, R. J., Convergence properties of S_n under moment restrictions. *Ann. Math. Statist.*, **41** (1970), 1235–1248.

[7] STOUT, W. F., *Almost Sure Convergence.* Academic Press, New York, (1974).

ANNALES UNIV. SCI. BUDAPEST., 36 (1993), 47-60

MINIMUM NUMBER OF TRAPEZOIDS PARTITIONING A POLYGONAL REGION

By

VALERIU SOLTAN and ALEXEI GORPINEVICH

Mathematical Institute of the Academy of Sciences, Kishinev, Rep. Moldova

(Received May 4, 1992)

1. Introduction

The partition problems form a well-studied field of computational geometry. This topic has a variety of applications in database systems, image processing, VLSI artwork data, etc. (see [1], [2]). One of these problems is the following: to partition the interior of a polygonal region into a minimum number of trapezoids with two horizontal bases (triangles with horizontal bases are considered as degenerate trapezoids). It is known that this problem is NP-complete [3]. Therefore, from practical needs, the investigation of some simple partition algorithms and of the estimations for the respective numbers of obtained trapezoids is actual.

In this connection, the following question appears: to determine the minimum number of trapezoids with two horizontal bases partitioning a polygonal region. A positive answer to this question gives a possibility to find good estimates for the number of partitioning trapezoids obtained by some simple algorithms.

In this paper, we determine a formula for a minimum number of trapezoids with two horizontal bases partitioning a polygonal region with arbitrary (possibly, degenerate) holes. Futher, we consider an $n \log n$ sweep-line partition algorithm, and find for this algorithm some estimates for the number of obtained trapezoids.

2. Description of polygonal regions

Let A be a closed (possibly, multiply connected) polygonal region in the plane E. The topological boundary $\mathrm{bd}\, A$ of A is assumed to be a union of a finite number of simple polygonal contours. Any two of these contours may be situated either one inside the other of mutually noninclusive; they may have common vertices but no common line segment, and cannot be crossing.

As usual, a point $x \in E$ is called interior for A provided it is contained inside an odd number of simple contours determining $\operatorname{bd} A$. All interior points of A form the topological interior $\operatorname{int} A$.

Inside A a finite family of *closed* line segments $s_1, \ldots, s_u$ and a finite family of isolated points $v_1, \ldots, v_t$ may be situated such that the following conditions are fulfilled:

1) $v_1, \ldots, v_t$ belong to $\operatorname{int} A \setminus (s_1 \cup \ldots \cup s_u)$,
2) the interior of each segment s_i is contained in $\operatorname{int} A$,
3) if some segments s_i, s_j have a common point, then it is a vertex for both s_i, s_j.

The point-set union of these points and segments is called the *ornament* of A, and denoted by $\operatorname{Or} A$.

In order to consider the degenerate holes, we introduce a new topology in the plane E. Since the description of the respective constructions is sometimes rather formal, we use below some illustrative examples. The first of them is shown in Fig. 1.

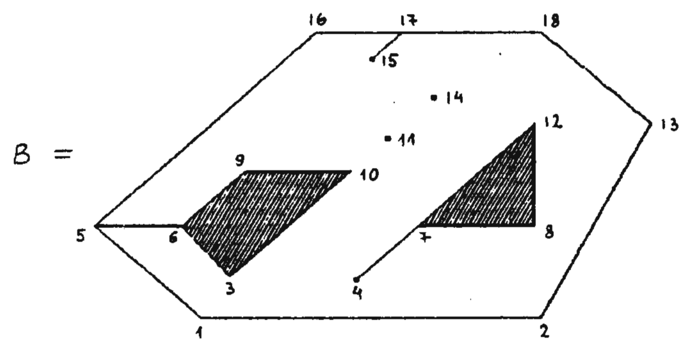

Fig. 1

Everywhere below the sets

$$\operatorname{Bd} A := \operatorname{Or} A \cup \operatorname{bd} A, \qquad \operatorname{Int} A := A \setminus \operatorname{Bd} A$$

are called, respectively, the *formal boundary* and the *formal interior* of A. A bounded component of $E \setminus \operatorname{Int} A$ is called a *formal* hole of A. As usual, a topological hole of A is a bounded component of $E \setminus \operatorname{int} A$.

EXAMPLE. For region B represented in Fig. 1, put $\mathrm{Or}\, B = [4,7] \cup \cup [5,6] \cup [15,17] \cup \{11\} \cup \{14\}$. There are three formal holes: $\{11\}$, $\{14\}$, and $[4,7] \cup \mathrm{conv}(7,8,12)$.

A point $x \in E$ is named a *vertex* of A if it is either a vertex of a segment forming $\mathrm{Bd}\, A$ or an isolated point in $\mathrm{Bd}\, A$. The set of all vertices of A is denoted by $V(A)$.

A closed line segment $[x,z]$, $x \neq z$, is called an *elementary segment* of $\mathrm{Bd}\, A$ if $[x,z] \subset \mathrm{Bd}\, A$ and $[x,z] \cap V(A) = \{x,z\}$. The set of all elementary segments of $\mathrm{Bd}\, A$ is denoted by $S(A)$.

Below a polygonal region A is considered to be determined if the sets $\mathrm{int}\, A$, $\mathrm{Bd}\, A$ and $V(A)$ are known. It is easily seen that $\mathrm{Bd}\, A$ and $V(A)$ uniquely determine $S(A)$.

Now let us make more precise the expression "to partition into trapezoids". A polygonal region A is called *partitioned into trapezoids* $T_1, \ldots, T_t$ if and only if

$$\bigcup_{i=1}^{t} \mathrm{int}\, T_i \subset \mathrm{Int}\, A \subset \bigcup_{i=1}^{t} T_i, \quad \mathrm{int}\, T_i \cap \mathrm{int}\, T_j = \emptyset, \quad i \neq j.$$

3. Measure of vertices

Any vertex x (if it is not an isolated point in $\mathrm{Bd}\, A$) is the apex of at least one inner angle of A. An inner angle of A is formed by two elementary segments of $\mathrm{Bd}\, A$ and contains no other inner angle (these segments may coincide and form in this case an angle of size 2π).

DEFINITION 1. An inner angle α of A with the apex x is called *divisible* if it is different from a straight angle (i.e., $\alpha \neq \pi$) and there is a horizontal segment $[x,y] \subset \mathrm{Int}\, A$ whose addition to $\mathrm{Bd}\, A$ divides α into two inner angles of A.

EXAMPLE. In Fig. 1, the inner angles with apexes 3, 4, 8, 9, 10, 12, 13, 15, respectively, are divisible.

OBSERVATION 1. It is easily seen that for any divisible angle with apex x there are at most two horizontal segments such that after their addition to $\mathrm{Bd}\, A$ all inner angles with apex x become indivisible.

OBSERVATION 2. For any vertex x of A, there are at most two divisible inner angles with apex x. If such two angles exist, then it is possible to draw inside each of them one segment with a vertex x making all the obtained inner with apex x indivisible.

OBSERVATION 3. For any isolated point $x \in \operatorname{Bd} A$, it is necessary to draw inside $\operatorname{Int} A$ two collinear segments of the type $[v, x]$, $[x, w]$ such that the obtained inner angles with apex x become indivisible.

DEFINITION 2. For any vertex x of A, denote by $m(x)$ the minimum number of horizontal segments in $\operatorname{Int} A$ with vertex x such that after drawing them (i.e., after the addition of these segments to $\operatorname{Bd} A$) all angles with apex x become indivisible. The number $m(x)$ is called the *measure* (of divisibility) of A at x.

COROLLARY. By Observations 1–3, $0 \leq m(x) \leq 2$ for any vertex x. Obviously, $m(x) = 2$ for any isolated point $x \in \operatorname{Bd} A$.

Let $m(A) := \sum m(x)$, where the sum is considered over the set $V(A)$ of all vertices of A.

EXAMPLE. In Fig. 1, $m(B) = 16$: the measure of B at each of the vertices 8, 9, 10, 13 is equal to one, at each of the vertices 3, 4, 11, 12, 14, 15 it is equal to two, and at any other vertex it is equal to zero.

4. Families of intervals

Let $v_i =]x_i, y_i[$, $i = 1, \ldots, l$ be some family of pairwise disjoint *open* line intervals belonging to $\operatorname{Int} A$. Put $L := \bigcup_{i=1}^{l} v_i$. If the ends of all the intervals $v_1, \ldots, v_l$ belong to $\operatorname{Bd} A \cup L$, it is possible to consider the region A_L obtained from A by the addition of L to $\operatorname{Or} A$:

$$\operatorname{Bd} A_L = \operatorname{Bd} A \cup L, \quad \operatorname{Int} A_L = \operatorname{Int} A \setminus L,$$
$$V(A_L) = V(A) \cup \{x_1, y_1, \ldots, x_l, y_l\}.$$

DEFINITION 3. A family $\mathscr{L} = \{v_1, \ldots, v_l\}$ of pairwise disjoint *open* line intervals in $\operatorname{Int} A$ is called *complete* (relative to A) if the following conditions are fulfilled:

1) the ends of all the intervals $v_1, \ldots, v_l$ belong to $\operatorname{Bd} A \cup L$,
2) for every vertex $x \in V(A_L) \setminus V(A)$, all the inner angles of A_L with apex x are indivisible.

For any complete family $\mathscr{L} = \{v_1, \ldots, v_l\}$, define its *weight* $g(\mathscr{L})$ by the formula

$$g(\mathscr{L}) := m(A) - m(A_L) - l. \tag{1}$$

DEFINITION 4. A complete family $\mathscr{L}$ of intervals is called *admissible* if the following conditions 1)–3) are fulfilled:

1) the deletion from Or A_L of any horizontal interval $v =]a, b[\in \mathscr{L}$ implies that one of the cases (a)–(c) occurs:

(a) the measure of A_L increases by one at both points a, b,

(b) the measure of A_L increases by one at one of the points a, b, and there is at least one nonhorizontal interval in $\mathscr{L}$ having an end in $]a, b[$,

(c) the measure of A_L remains unchanged at both points a, b and there are at most two nonhorizontal intervals in $\mathscr{L}$ each having an end in $]a, b[$;

2) The deletion from Or A_L of any nonhorizontal interval $v \in \mathscr{L}$ increases the measure $m(A_L)$ by two or more units;

3) the deletion from Or A_L of any two nonhorizontal collinear intervals of the form $]a, b[$, $]b, c[\in \mathscr{L}$ increases the measure $m(A_L)$ by three or more units.

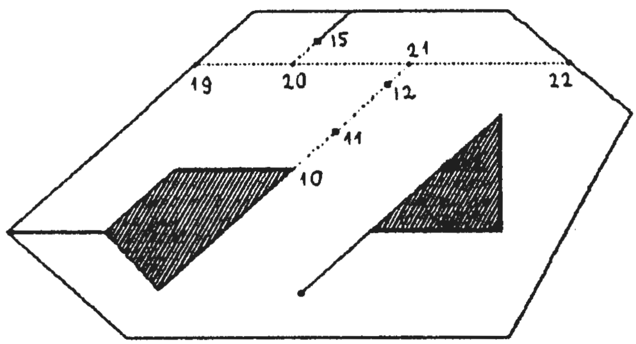

Fig. 2

EXAMPLE. For region B in Fig. 1, the family of open line intervals

$$\mathscr{L} = \{]10, 11[,\ ]11, 12[,\ ]12, 21[,\ ]15, 20[,\ ]19, 22[\}$$

is admissible (see Fig. 2). The region B_L has new (relative to B) vertices 19, 20, 21, 22. The weight of $\mathscr{L}$ equals

$$g(\mathscr{L}) = m(B) - m(B_L) - l = 16 - 9 - 5 = 2.$$

Note that segment $[19, 22]$ is the union of three elementary segments $[19, 20]$, $[20, 21]$, $[21, 22] \in \operatorname{Or} B_L$.

5. Some preliminary results

Denote by $\alpha_0 = \alpha_0(A)$ the number of vertices of A and by $\alpha_1(A)$ the number of elementary segments of Bd A.

LEMMA 1 (Generalized Euler's relation). For a polygonal region A,

$$\alpha_0 - \alpha_1 + s_0 = s + h - h_0, \tag{2}$$

where

s_0 is the number of connected components of $\operatorname{Int} A$,

s is the number of connected components of A,

h is the number of formal holes of A,

h_0 is the number of topological holes of A.

PROOF. Suppose that $h > 0$ and let C be any formal hole of A. Then C is separated from other formal holes or from the exterior of A by some connected component D of $\operatorname{Int} A$. It is possible to draw inside D a simple polygonal line $[a_1, \dots, a_k]$ such that a_1 is a vertex of A belonging to C, and a_k is vertex of A belonging either to another formal hole or to the exterior of A. In both cases, the addition of $[a_1, \dots, a_k]$ to $\operatorname{Bd} A$ preserves the value $s_0(A)$ and reduces the difference $\alpha_0 - \alpha_1$ by one.

Executing this procedure for each formal hole of A we obtain a polygonal region for which the canonical Euler's relation is valid:

$$\alpha_0 - \alpha_1 - h + s_0 + h_0 = s. \ \square$$

LEMMA 2. For any complete family $\mathcal{L} = \{v_1, \dots, v_l\}$,

$$\alpha_1(A_L) - \alpha_1(A) - [\alpha_0(A_L) - \alpha_1(A)] = l. \tag{3}$$

PROOF. The closure $\overline{v}_i$ of each open line interval $v_i \in \mathcal{L}$ is uniquely represented in the form

$$\overline{v}_i = [a_1, a_2] \cup \dots \cup [a_{k_i - 1}, a_{k_i}], \qquad k_i \geq 2,$$

where $a_2, \dots, a_{k_i - 1}$ are new (relative to A) vertices of A_L and

$$[a_1, a_2], \dots, [a_{k_i - 1}, a_{k_i}] \in \operatorname{Or} A.$$

If we delete v_1 from $\operatorname{Or} A_L$, then $k_1 - 1$ segments from $\operatorname{Or} A_L$ and $k_1 - 2$ vertices from $V(A_L)$ vanish (the ends of v_1 remain the vertices of $A_{L \setminus v_1}$). Therefore by the deletion from $\operatorname{Or} A_L$ of all intervals $v_1, \dots, v_l$ there vanish $(k_1 - 1) + \dots + (k_l - 1)$ elementary segments and $(k_1 - 2) + \dots + (k_l - 2)$ vertices of A_L.

After the deletion of $v_1, \dots, v_l$ from $\operatorname{Or} A$, some vertices of A_L which are not vertices of A can remain in $\operatorname{Bd} A$. We arrange them into groups such that the vertices of each group belong to the same elementary segment of $\operatorname{Bd} A$, and vertices of different groups belong to different elementary segments.

Denote by $[c_i, d_i]$, $i = 1, \dots, r$ all segments containing the above discussed vertices, and let δ_i be the number of vertices of $V(A_L)$ placed in $[c_i, d_i]$. Then

$$\alpha_1(A_L) - \alpha_1(A) = \sum_{i=1}^{l}(k_i - 1) + \sum_{i=1}^{r}(\delta_i + 1) - r,$$

$$\alpha_0(A_L) - \alpha_0(A) = \sum_{i=1}^{l}(k_i - 2) + \sum_{i=1}^{r}\delta_i.$$

Hence

$$\alpha_1(A_L) - \alpha_1(A) - [\alpha_0(A_L) - \alpha_0(A)] = l. \ \square$$

THEOREM 1. *Let some complete family* $\mathscr{L} = \{v_1, \dots, v_l\}$ *of open line intervals form together with* $\operatorname{Bd} A$ *a partition of* A *into trapezoids with horizontal bases. Then the number* $t(A)$ *of these trapezoids equals*

$$t(A) = m(A) + s_0(A) - h(A) - g(\mathscr{L}), \tag{4}$$

where

$s_0(A)$ *is the number of connected components of* $\operatorname{int} A$,

$h(A)$ *is the number of formal holes of* A.

PROOF. As above, A_L denotes the region obtained from A by the addition of the set $L := \bigcup_{i=1}^{l} v_i$ to $\operatorname{Bd} A$. By Lemma 1,

$$\alpha_0(A) - \alpha_1(A) + s_0(A) = s(A) + h(A) - h_0(A),$$
$$\alpha_0(A_L) - \alpha_1(A_L) + s_0(A_L) = s(A_L) + h(A_L) - h_0(A_L).$$

Due to relations

$$s(A) = s(A_L), \quad s_0(A_L) = t(A), \quad h(A_L) = 0, \quad h_0(A) = h_0(A_L),$$

one has $\alpha_0(A) - \alpha_1(A) + s_0(A) = \alpha_0(A_L) - \alpha_1(A_L) + t(A) + h(A)$.

By Lemma 2,

$$t(A) = [\alpha_1(A_L) - \alpha_1(A)] - [\alpha_0(A_L) - \alpha_0(A)] + s_0(A) - h(A) =$$
$$= l + s_0(A) - h(A).$$

Now, due to (1) and to the relation $m(A_L) = 0$, we receive (4).

6. Minimum number of trapezoids

Let $A \subset E$ be any polygonal region, and let $\mathscr{M} = \{w_1, \dots, w_k\}$ be a family of *open* line intervals in $\operatorname{Int} A$ such that: 1) any two segments of $\mathscr{M}$ have at most one point in common, 2) $w_1 \cup \ldots \cup w_k$ partitions A together with $\operatorname{Bd} A$ into trapezoids with horizontal bases.

LEMMA 3. The set $M := \bigcup_{i=1}^{k} w_i$ can be represented as a union of pairwise disjoint open line intervals forming a complete family for A.

PROOF. The absence of divisible angles of the region A_M is obvious. Hence it remains to show that M can be represented as a union of pairwise disjoint open intervals all of whose ends belong to $\mathrm{Bd}\, A \cup M$. This representation can be organized by the repeated implementation of the following

Procedure: If some vertex x of A_M belongs to the intersection of at least two intervals from $\mathcal{M}$, then one of these intervals remains unchanged and each other interval $]y, z[$ is replaced in $\mathcal{M}$ by two open intervals $]y, x[$, $]x, z[$.

Since $\mathcal{M}$ is finite, we obtain the desired representation. □

THEOREM 2. *If some complete family $\mathcal{L} = \{v_1, \ldots, v_l\}$ of open line intervals partitions together with* $\mathrm{Bd}\, A$ *a polygonal region A into a minimum number of trapezoids, then there exists an admissible family $\mathcal{N}$ satisfying the equality $g(\mathcal{N}) = g(\mathcal{L})$.*

PROOF of theorem 2 is realized in two steps:

— The family $\mathcal{L}$ is changed by another complete family $\mathcal{M}$,

— the desired family $\mathcal{N}$ is selected in $\mathcal{M}$.

For the realization of these steps, we need the following transformations of the partition of A into trapezoids:

(d) if some two partitioning trapezoids are placed as it is shown in Fig. 3a and $]a, b[\subset L$, they should be replaced by two trapezoids shown in Fig. 3b.

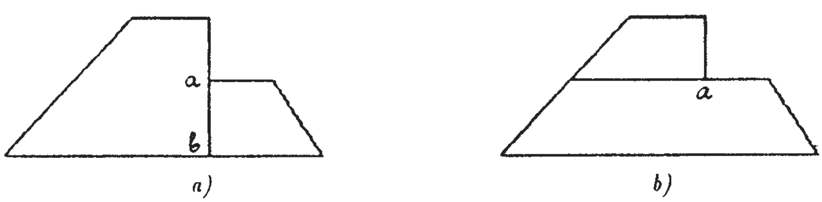

Fig. 3

(e) if some two partitioning trapezoids are placed as it is shown in Fig. 4a and $]a, b[\cup]b, c[\subset L$, they should be replaced by two trapezoids shown in Fig. 4b.

(f) if some three partitioning trapezoids are placed as it is shown in Fig. 5a and $]a,b[\cup]c,d[\subset L$, they should be replaced by three trapezoids shown in Fig. 5b.

a) b)

Fig. 4

a) b)

Fig. 5

Transformations of (d)–(f)-type are applied to A_L in any order until it is possible. Observe that any transformation of (d)–(f)-type has the following properties: 1) it deletes at most one non-horizontal elementary segment of the initial region A_L whose ends are placed on two horizontal levels determined by some vertices of A_L, 2) no new non-horizontal elementary segment appears in $\operatorname{Bd} A$. Hence these transformations can be applied no more than a finite number of times.

As the result, we obtain some new region A_M, where M is the relative interior of the union of all parts of all sides of the partitioning trapezoids which do not belong to $\operatorname{Bd} A$. According to Lemma 3, the set M can be represented as the union of elements of some complete family $\mathcal{M}$ of pairwise disjoint open line intervals.

OBSERVATION 4. The above obtained family $\mathcal{M}$ contains no interval $]a,b[$ with the following properties: 1) there are no vertices of A_M in $]a,b[$, 2) the deletion of $]a,b[$ from $\operatorname{Or} A_M$ preserves $m(A_M)$ at both points a, b. Indeed, otherwise the removing of $]a,b[$ from $\operatorname{Or} A_M$ reduces the number of trapezoids.

Obviously, the numbers of partitioning trapezoids for the domains A_L and A_M coincide. Hence from Theorem 1, $g(\mathcal{L}) = g(\mathcal{M})$.

Now an admissible family $\mathcal{N}$ will be selected in $\mathcal{M}$. For this purpose, we remove from M all horizontal intervals which contain no vertex of A_M and whose deletion from Or A_M increases $m(A_M)$ by one. Denote by r the number of all such deleted intervals.

The remaining family $\mathcal{N} = \{w_1, \ldots, w_n\}$ is admissible. Indeed, the completeness of $\mathcal{N}$ follows from the method of the interval deletion from $\mathcal{M}$, taking into consideration the transformations (d) and (e), and the minimality of the number of trapezoids. Now we shall verify the validity of conditions 1)–3) from Definition 4. Below N denotes the union of the segments belonging to $\mathcal{N}$.

1°. Assume that some horizontal segment $]a, b[\in \mathcal{N}$ does not satisfy condition 1). Then one of the following cases can occur:

(g) $]a, b[$ contains no vertex of A_N and the deletion of $]a, b[$ from Or A_N does not change the measure of A_N at any of a, b,

(h) $]a, b[$ contains no vertex of A_N and the deletion of $]a, b[$ from Or A_N increases by one the measure of A_N at one of the points a, b only,

(i) there exists in $\mathcal{N}$ only one nonhorizontal interval having a vertex in $]a, b[$ and the removal of $]a, b[$ from Or A_N does not change the measure of A_N at any of a, b.

Case (g) is impossible because of the minimality of the number of trapezoids. Case (h) is impossible because of the removal of this type of intervals in the process of forming of $\mathcal{M}$. Case (i) is impossible because of the previous (f)-type transformations (see Fig. 5).

2°. Suppose that some nonhorizontal interval $]a, b[\in \mathcal{N}$ does not satisfy condition 2). Then one of the following cases can occur:

(j) the deletion of $]a, b[$ from Or A_N changes the measure of A_N at none of the points a, b,

(k) the deletion of $]a, b[$ from Or A_N increases the measure of A_N at one of the points a, b only.

Case (j) is impossible either because of the minimality of the number of trapezoids or because of the previous (d)-type transformations.

3°. Suppose the existence of two nonhorizontal collinear intervals $]a, b[$, $]b, c[\in L$ such that their deletion from Or A_N increases the measure $m(A_N)$ by no more than two units.

Because of 2°, it is sufficient to verify the case that $m(A_N)$ increases by two units. Due to previous (d)- and (f)-type transformations, this situation is impossible.

Hence $\mathcal{N}$ is an admissible family. By all the said above,

$$g(\mathcal{N}) = m(A) - m(A_N) - n = m(A) - r - n = m(A) - \operatorname{card}\mathcal{M} = g(\mathcal{M}) = g(\mathcal{L}). \square$$

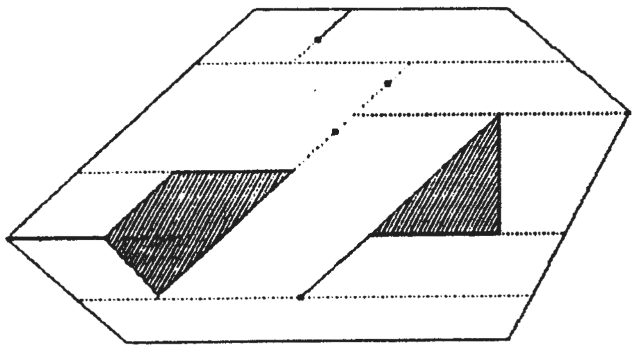

Fig. 6.

EXAMPLE. For region B in Fig. 1, a partition into a minimum number of trapezoids is shown in Fig. 6.

With the help of Theorem 2, it is possible to indicate the following partition algorithm into a minimum number of trapezoids.

Let $A \subset E$ be any polygonal region.

Algorithm 1.

1. Construct an admissible family $\mathcal{L} = \{v_1, \ldots, v_l\}$ of open intervals having a maximum weight $g(\mathcal{L})$,
2. Put $L := \bigcup_{i=1}^{l} v_i$ and $\operatorname{Or} A_L := \operatorname{Or} A \cup L$,
3. For every vertex x of A, draw in $\operatorname{Int} A_L m_L(x)$ open horizontal segments maximal under inclusion removing the divisibility of every divisible angle α of A_L with the apex x.

From Theorem 1, it follows that for a polygonal region A, the number of trapezoids obtained by the application of Algorithm 1 is equal to $t(A) = m(A) + s_0(A) - h(A) - g(\mathcal{L})$.

7. Some estimates for an approximation algorithm

In the section an $O(n \log n)$ sweep-line partition algorithm consisting of step 3 of Algorithm 1 is considered.

Algorithm 2.

For each divisible angle α of A with the apex x, draw inside $\operatorname{Int} A$ $m(x)$ open horizontal segments maximal under inclusion removing the divisibility of α.

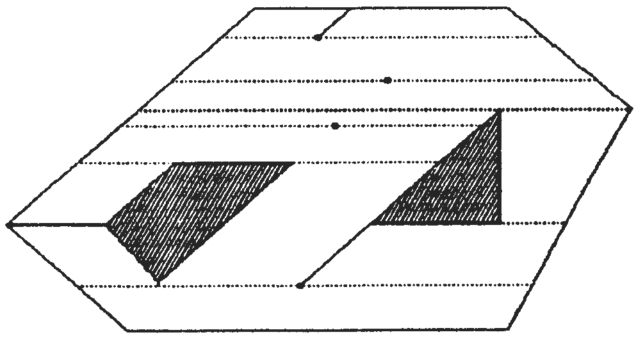

Fig. 7.

EXAMPLE. The partitioning trapezoids obtained for region B represented in Fig. 1 by the application of Algorithm 2 are shown in Fig. 7.

Denote by $t_1(A)$ and $t_2(A)$, respectively, the numbers of partitioning trapezoids obtained for a polygonal region A by the application of Algorithm 1 and Algorithm 2.

Below we find some estimates for the ratio

$$\gamma(A) := \frac{t_2(A)}{t_1(A)}.$$

Let $\mathcal{L}_1 = \{v_1, \dots, v_p\}$ be the family of open intervals obtained for A by the implementation of Algorithm 1 (the set $L := \bigcup_{i=1}^{p} v_i$ partitions together with $\operatorname{Bd} A$ the region A into the minimum number of trapezoids with horizontal bases), and let $\mathcal{L}_2 = \{w_1, \dots, w_q\}$ be the complete family of open horizontal intervals obtained for A by the implementation of Algorithm 2.

LEMMA 4. The following relations hold:

1) $h(A) \le p - 1$ if $h > 0$,

2) $h(A) \le p/2 - 1$ if $h > 0$ and each formal hole of A has nonempty interior.

3) $q \le m(A) \le 4p$,

4) $q \le m(A) \le 2p$ if $\operatorname{Or} A = \emptyset$.

PROOF. a) For any formal hole H of A, denote by $\gamma(H)$ the number of all open intervals in $\mathcal{L}_1$ incident to H. Similarly define the number $\gamma(C)$, where C means the topological boundary $\mathrm{bd}\,A$ of A. Obviously, $\gamma(C) \geq 2$ and $\gamma(H) \geq 2$ for any hole H of A (in case $h > 0$). Then

$$h(A) \leq \frac{1}{2}\left(\sum \gamma(H)\right) \leq \frac{1}{2}(2p - \gamma(C)) \leq p - 1.$$

2) If each formal hole H of A has nonempty interior, then, obviously, $\gamma(C) \geq 4$ and $\gamma(H) \geq 4$ (in the case $h > 0$). As above,

$$h(A) \leq \frac{1}{4}\left(\sum \gamma(H)\right) \leq \frac{1}{4}(2p - \gamma(C)) \leq p/2 - 1.$$

3) The inequality $q \leq m(A)$ is obvious. In order to prove the inequality $m(A) \leq 4p$, put

$$V_i = \{x \in V(A) \ : \ m(x) = i\}, \quad i = 1, 2,$$

and denote by $\gamma(x)$ the number of all open intervals in $\mathscr{L}_1$ incident to x. Since $m(A) = \mathrm{card} V_1 + 2 \cdot \mathrm{card} V_2$, one has

$$4p \geq 2 \cdot \sum \{\gamma(x) \ : \ x \in V\} \geq 2 \cdot \sum \{\gamma(x) \ : \ x \in V_1\} +$$

$$+2 \cdot \sum \{\gamma(x) : x \in V_2\} \geq 2 \cdot (\mathrm{card} V_1 + \mathrm{card} V_2) = m(A) + \mathrm{card} V_1 \geq m(A).$$

4) If $\mathrm{Or}\, A = \emptyset$ then, obviously, $\gamma(x) \geq 2$ for each vertex $x \in V_2$. In this case

$$2p \geq \sum \{\gamma(x) \ : \ x \in V\} \geq \sum \{\gamma(x) \ : \ x \in V_1\} +$$

$$+ \sum \{\gamma(x) : x \in V_2\} \geq \mathrm{card} V_1 + \mathrm{card} V_2 = m(A).$$

THEOREM 3. *For a polygonal region A, the following inequalities are fulfilled:*

1)	$\gamma(A) \leq \frac{3}{8} m(A) + 1$	*if A has formal holes,*
2)	$\gamma(A) < 7$	*if A has formal holes and each formal hole of A has nonempty interior*
3)	$\gamma(A) < 4$	*if A has no formal holes,*
4)	$\gamma(A) < 3$	*if A has formal holes and $\mathrm{Or}\, A = \emptyset$,*
5)	$\gamma(A) < 2$	*if A has no formal holes and $\mathrm{Or}\, A = \emptyset$.*

PROOF. By Theorem 1 and (1),

$$t_1(A) = m(A) + s_0(A) - h(A) - g(\mathscr{L}_1) = p + s_0(A) - h(A),$$
$$t_2(A) = m(A) + s_0(A) - h(A) - g(\mathscr{L}_2) = q + s_0(A) - h(A).$$

Hence

$$\gamma(A) = \frac{t_2(A)}{t_1(A)} = 1 + \frac{q - p}{p + s_0(A) - h(A)}.$$

Obviously, $s_0(A) \geq 1$. Put $m = m(A)$.

3) If A has no formal hole (i.e., $h(A) = 0$), then, by Lemma 4, $p + s_0(A) \geq m/4 + 1$, and

$$\frac{q-p}{p+s_0(A)} \leq \frac{m - \frac{m}{4}}{\frac{m}{4}+1} = \frac{3m}{m+4} < 3.$$

5) If A has no formal hole and $\operatorname{Or} A = \emptyset$, then $p \geq m/2$, and

$$\frac{q-p}{p+s_0(A)} \leq \frac{m - \frac{m}{2}}{\frac{m}{2}+1} = \frac{m}{m+2} < 1.$$

Assume that $h > 0$. By Lemma 4, one has $p + s_0(A) - h(A) \geq 2$, and

$$\frac{q-p}{p+s_0(A)-h(A)} \leq \frac{m - \frac{m}{4}}{2} = \frac{3}{8}m$$

in case 1) of the theorem. Since $p + s_0(A) - h(A) = s_0(A) + (p/2 - h(A)) + p/2$, one has

$$\frac{q-p}{p+s_0(A)-h(A)} \leq \frac{m - \frac{m}{4}}{2+\frac{m}{8}} = \frac{6m}{m+16} < 6$$

in case 2) of the theorem, and

$$\frac{q-p}{p+s_0(A)-h(A)} \leq \frac{m - \frac{m}{2}}{2+\frac{m}{4}} = \frac{2m}{m+8} < 2$$

in case 4) of the theorem (since $\operatorname{Or} A = \emptyset$ implies that the interior of each formal hole of A is not empty). □

References

[1] N. Sugiyam, K. Saitoh, Electron-beam exposure system AMDES, *Computer Aided Design*, **11** (1979), 59–65.

[2] L. E. Nosov, E. G. Nosova, E. B. Rabinovich, V. Z. Feinberg, Optimal algorithms for solving the placing and covering problems for the plane polygonal figures, *IZV. AN BSSR.* Ser. Fiz.-Mat. Nauk., **4** (1985), 28–36 (Russian).

[3] T. Asano, R. Asano, Minimum partition of polygonal regions into trapezoids, *24th Annual Symp. on Found, of Comp. Sci.*, Tucson, Ariz., (1983), 233–241.

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 61–66

ON THE CONTROL OF SOME DISTRIBUTED PARAMETER SYSTEMS I

By

ABDUL SAMEE ABDUL RAZZAK AL-JANABI*

Dept. of Mathematics, College of Science, Al-Mustansiriya University, Baghdad, IRAQ

(Received September 7, 1992)

Introduction

In this paper and in the next papers we consider systems described by differential equations:

$$(1) \qquad u_t(x,t) = \sum_{k=0}^{2n} p_k(x) \frac{\partial^{2n-k}}{\partial x^{2n-k}} u(x,t) \qquad (0 < t < T,\ 0 < x < 1)$$

and boundary equations:

$$(2) \quad \sum_{k=0}^{2n-1} [\alpha_{jk}^0 u^{(k)}(0,t) + \alpha_{jk}^1 u^{(k)}(1,t)] \equiv \tau_j u(.,t) = \sum_{\mu=1}^{l} \beta_j^\mu f_\mu(t)$$
$$(j = 1,2,3,\ldots,2n).$$

Where $u^{(k)}(0,t)$ and $u^{(k)}(1,t)$ denote the k^{th} derivative of $u(x,t)$ with respect to x in the points $(0,t)$, $(1,t)$ respectively.

$p_k(.) \in H^{2n-k}(0,1)$ and suppose $p_0(x) > 0$ $(0 \leq x \leq 1)$ if $n \equiv 0 \pmod 2$. Further $p_0(x) < 0$ $(0 \leq x \leq 1)$ if $n \equiv 1 \pmod 2$. The control functions f_μ are chosen from the class $H^1(0,T)$. It is supposed that the boundary conditions $\tau_j u = 0$ $(j = 1,2,3,\ldots,2n)$ are "regular" [1].

The initial conditions of (1), (2) are given in the form:

$$(3) \qquad u(x,0) = u_0(x), \qquad u_0 \in L^2(0,1).$$

Under the conditions above which are mode on u_0 and $f_\mu(t)$ $(\mu = 1,2,\ldots,l)$ the problem (1), (2) and (3) have no classical solution with 2nd partial

* This work was written while the author was visiting at the Eötvös Loránd University Budapest.

derivatives with respect to x. The generalized solutions of (1), (2) and (3) which exist and are unique and defined (as it will be seen later) for any fixed u_0 and f_μ $(\mu = 1, 2, \ldots, l)$ (from classes given above).

Moment problem

Denote the differential operator by σu on the right hand side of (1) $\sigma^* \nu$, $\nu \in H^{2n}(0,1)$ and its adjoint by $\sigma^* \nu$ (see [1]): i.e.

$$\{\sigma^* \nu\}(x) = \sum_{k=0}^{2n} (-1)^k \frac{\partial^{2n-k}}{\partial x^{2n-k}} (\overline{P}_k(x) \nu_k(x)). \tag{4}$$

For any u, $\nu \in H^{2n}(0,1)$ we have the Lagrange-formula:

$$\int_0^1 (\sigma u)\overline{\nu} \mathrm{d}x = \int_0^1 u(\overline{\sigma^* \nu}) \mathrm{d}x + \sum_{j=1}^{2n} \left[(\tau_j u)(\overline{\tau_j^* \nu}) + (\tilde{\tau}_j u)(\overline{\tau_j^* \nu}) \right]. \tag{5}$$

Here $\tau_j^* \nu$ $(j = 1, 2, \ldots, 2n)$ denotes the linear form of quantities $\nu^{(k)}(0)$ and $\nu^{(k)}(1)$, adjoint to $\tau_j u$ $(k = 0, 1, 2, \ldots, 2n-1)$; the linear form $\tilde{\tau}_j u$ and $\tilde{\tau}_j^* \nu$ complete $\tau_j u$, $\tau_j^* \nu$ respectively to a complete linearity independent system. We use the spaces $H^s(\Omega)$, $H^{s,p}(\Omega \times [0,T])$ in the paper as given in the fundamental book [6].

Denote ν the set of functions $\nu(x,t)$ from the class $H^{n,1}(Q)$, $Q = [0,1] \times$ $\times [0,T]$ for which $(\tau_j^* \nu)(.,t) = 0$, $j = 1, 2, \ldots, 2n$; $\nu(.,T) = 0$.

DEFINITION. The solution of (1), (2) and (3) is called a function $u \in$ $\in L^2(Q)$ for which:

$$\iint_Q u \left[-\frac{\partial \overline{\nu}}{\partial t} - \overline{\sigma^* \nu} \right] \mathrm{d}x \mathrm{d}t = \tag{6}$$

$$= \int_0^1 u_0(x) \overline{\nu(x,0)} \mathrm{d}x + \int_0^T \sum_{j=1}^{2n} \sum_{\mu=1}^{l} \beta_j^\mu f_\mu(t) \overline{(\tilde{\tau}_j^* \nu)(.,\tau)} \mathrm{d}t$$

holds for every $\nu \in V$.

If u_0 and f_μ $(\mu = 1, 2, \ldots, l)$ are smooth enough, then the classical solutions of (1), (2) and (3) exist, and are unique ([1]). By (5) it can be easily checked that they coincide with the previously defined generalized solutions. We now prove the following theorem:

THEOREM. *For any $u \in L^2(Q)$ the generalized solution exists and is unique (even in the case when $u_0^{(k)} \in L^2(0,1)$ and $f_\mu(t) \in H^1(0,T)$). Further for any fixed t, $u(.,t) \in L^2(0,1)$ and $t \to u(.,t)$ is continuous as $\mathbb{R} \to L^2(0,1)$.*

We shall investigete the following problems:

1. The description of $L^2(0,1)$ the set $D(u_0,T)$ of states u_1 which we can reach from the state u_0 in time less than T.

2. The description of $L^2(0,1)$ the set $B(u_1,T)$ of states u_0 from which we can arrive at u_1 in time less than T.

We shall reduce these problems to moment problems in the space $L^2(0,T)$. For this the following operator is introduced.

(7) $$(Au)(x) = -\sum_{k=0}^{2n} p_k \frac{\partial^{2n-k}}{\partial x^{2n-k}} u(x) \quad (= -\sigma u) \quad \text{with domain:}$$

$$\operatorname{dom} A := \{u \in H^{2n}(0,1) \ : \ \tau_j u = 0; \quad j = 1,2,\ldots,2n\}.$$

The adjoint operator of A is $A^* = -\sigma^* \nu$ with domain:

$$\operatorname{dom} A^* := \{\nu \in V \ : \ \tau_j v = 0; \ j = 1,2,\ldots,2n\}.$$

It is well known (see [1]) that A has discrete spectrum. Further for the eigenvalues $\{\lambda_k\}$ if A we have the asymptotic formula

(8) $$\lambda_k = \left(\frac{\pi k}{L}\right)^{2n} \left[1 + O\left(\frac{1}{k}\right)\right]; \quad L := \int_0^1 \frac{\mathrm{d}x}{\sqrt[2n]{|p_0(x)|}}.$$

It is known also (see [1]) that according to the Kesselmen–Mihajlov theorem, the eigenfunctions of A form Riesz basis in $L^2(0,1)$, moreover the multiplicity of the eigenvalues is uniformly bounded.

Denote $\{\psi_{mp}\}$ $(p = 1,2,3,\ldots,r_m)$ the chain of eigenfunctions with eigenvalues λ_m. Further denote $\{\phi_{ks}\}$ corresponding to the biorthogonal system $\{\psi_{mp}\}$; i.e.:

(9) $$\int_0^1 \phi_{ks}(x)\overline{\psi_{mp}}(x)\mathrm{d}x = \begin{cases} 1 & \text{if } k = m,\ s = p \\ 0 & \text{otherwise.} \end{cases}$$

It can be shown (see [1]) that $\{\phi_{ks}\}$ are the eigenfunctions of A^*, i.e.

(10) $$A\phi_{k1} = \lambda_k \psi_{k1}, \quad A\phi_{ks} = \lambda_k \phi_{ks} + \phi_{ks-1} \quad (s = 2,3,\ldots,r_k)$$
$$A^*\psi_{mr_m} = \overline{\lambda_m}\psi_{mr_m}, \quad A^*\psi_{mp} = \overline{\lambda_m}\psi_{mp} + \psi_{mp+1} \quad (p = 2,3,\ldots,r_m).$$

We can write the solution of (1), (2), and (3) in the form:

$$u(x,t)=\sum_{k,s} a_{ks}(t)\phi_{ks}(x), \qquad \sum_{k,s}\int_0^T |a_{ks}(t)|^2 \mathrm{d}t<\infty$$

and for the initial condition $u_0 \in L^2(0,1)$ we have:

(11) $$u_0(x)=\sum_{k,s} a^0_{ks}\phi_{ks}(x), \qquad \sum_{k,s}|a^0_{ks}|^2<\infty.$$

To determine the coefficients $a_{ks}(t)$, let $\nu(x,t) := \overline{W(t)}\phi_{ks}(x)$, where $W(t)C^1(0,T)$ is such that $W(T)=0$, and substitute this function $\nu(x,t)$ into (6).

Taking into account (9), (10) we obtain:

(12) $$\int_0^T \left[-a_{ks}(t)\frac{\mathrm{d}W(t)}{\mathrm{d}t}+\lambda_k a_{ks}(t)W(t)+a_{ks+1}(t)W(t)\right]\mathrm{d}t=$$

$$=a^0_{ks}W(0)+\sum_{j=1}^{2n}\sum_{\mu=1}^{l}\beta^{\mu}_j\overline{\bar{\tau}^*_j(\psi_{ks})}\int_0^T f_\mu(t)W(t)\mathrm{d}t.$$

This is equivalent to the following:

(13) $$\frac{\mathrm{d}a_{ks}(t)}{\mathrm{d}t}+\lambda_k a_{ks}(t)+a_{ks+1}(t)=\sum_{j=1}^{2n}\sum_{\mu=1}^{l}\beta^{\mu}_j\gamma^j_{ks}f_\mu(t),$$

$a_{ks}(0)=a^0_{ks}$, $(s=1,2,\ldots,r_k;\ k=1,2,\ldots)$, where $\gamma^j_{ks}=\overline{\bar{\tau}^*_j(\psi_{ks})}$, $a_k r_k := 0$.

It is easily shown that the solution of (12) has the form:

(14) $$a_{ks}(t)=e^{-\lambda_k t}\sum_{p=0}^{r_k-s} a^0_{ks+p}\frac{(-1)^p t^p}{p!}+$$

$$+\sum_{s=1}^{2k}\sum_{\mu=1}^{l}\sum_{p=0}^{r_k-s}\beta^{\mu}_j\gamma^j_{ks+p}\int_0^T f_\mu(\tau)\frac{(-1)^p(t-\tau)^p}{p!}e^{-\lambda_k(t-\tau)}\mathrm{d}\tau.$$

Suppose that u_1 has the exposition:

(15) $$u_1(x)=\sum_{k,s} a^1_{ks}\phi_{ks}(x), \qquad \sum|a^1_{ks}|^2<\infty.$$

Obviously, the equality $u(.,T) = u_1$ holds if and only if:

$$a_{ks}(T) = a^1_{ks} \qquad (s = 1,2,\dots,r_k; \quad k = 1,2,\dots).$$

Hence the controllability of the system considered is equivalent to the following moment problem: there exists $f_\mu \in H^1(0,T)$ $(\mu = 1,2,\dots,l)$ such that:

$$(16)\quad a^1_{ks} - e^{-\lambda_k T}\sum_{p=0}^{r_k-s} a^0_{ks+p}\frac{(-1)^p T^p}{p!} =$$

$$= \sum_{j=1}^{2n}\sum_{\mu=1}^{l}\sum_{p=0}^{r_k-s}\beta^\mu_j\gamma^j_{ks+p}\int_0^T f_\mu(\tau)\frac{(-1)^p(T-\tau)^p}{p!}e^{-\lambda_k(T-\tau}\mathrm{d}\tau$$

$(s = 1,2,\dots,r_k;\ k = 1,2,3,\dots)$.

Hence the problem of describing the sets D and B is equivalent to that of the set of coefficients $\{a^0_{ks}\}$ and $\{a^1_{ks}\}$ for which the moment problem (16) has the solutions $f_\mu \in H^{-1}(0,T)$.

Note that from (14) it follows that the mapping $t \to u(.,T)$ as $[0,T] \to$ $\to L^2(0,1)$ is continuous. Indeed, taking into account the Lagrange formula (5) and the (special) form of the boundary conditions (3) it follows that $\tilde{\tau}^*_j(\nu)$ contains the derivatives of ν with respect to x of order $\leq 2n-1$. Now applying the asymptotic formulas for the eigenfunctions of the operator A^* (see [1]) we obtain:

$$\tilde{\tau}^*_j(\psi_{ks}) = O(k^{2n-1}) \qquad (k \to \infty).$$

Hence for every j we have: $\gamma^j_{ks} = O(k^{2n-1})$. Using (14) and applying the Cauchy inequality we get:

$$|a^1_{ks}|^2 \leq C_1(t)e^{-\lambda_k t}\max_s|a^0_{ks}|^2 +$$

$$+C_2(t)\max_{\mu,j,s}\left[\frac{|\gamma^j_{ks}|^2}{|\lambda_k|^2}|f_\mu(T)|^2 + \frac{\|f_\mu\|_{L^2(0,T)}}{|\lambda_k|^2}|\gamma^j_{ks}|^2\right]$$

where $C_1(t)$ and $C_2(t)$ are continuos and are independent on k and s. By the asymptotic formula (8) we get:

$$|\gamma^j_{ks}|^2|\lambda_k|^{-2} = O(k^{4n-2})\cdot O(k^{-4n}) = O(k^{-2}) \quad (k \to \infty).$$

Hence $|a_{ks}(t)|^2 = O(k^{-2})$ and consequently $\sum_{k,s}|a_{ks}(t)|^2 < \infty$. Because easily calculations give

$$\sum_{k,s}|a_{k,s}(t)|^2 \asymp \|u(.,t)\|^2_{L^2(0,1)}.$$

Hence $u(.,t) \in L^2(0,1)$ for every $t \in [0,1]$. We can similarly prove that $u(.,t)$ depends continuously on t in $L^2(0,1)$.

References

[1] HARTMAN, P., *Ordinary Differential Equations*, Wiley and sons, New York, 1964.

[2] RUSSEL, D. L., Controllability and Stabilizability Theory for Linear Partial Differential Equations, Recent Progress and Open Questions, *SIAM Review*, **20** (1978), 639–739.

[3] LIONS, J.-L., Exact Controllability, Stabilizations and Perturbations for Distributed Systems, *SIAM Review*, **30** (1988), 1–68.

[4] HORVÁTH, M., Vibrating Strings With Free Ends, *Acta Math Hung.*, **51** (1988), 171–180.

[5] JOÓ, I., On the Reachability Set of a String, *Acta Math.* **49** (1987), 203–211.

[6] LIOUS, J.-L., MENEGES, E., *Non-homogeneous boundary value problem and applications*, Paris, 1970.

ANNALES UNIV. SCI. BUDAPEST., 36 (1993), 67–70

ON THE CONTROL OF DISTRIBUTED PARAMETER SYSTEMS II (MOMENT PROBLEM)

By

ABDUL SAMEE ABDUL RAZZAK AL-JANABI

Dept. of Mathematics, College of Science, Al-Mustansiriya University, Baghdad, IRAQ

(Received September 7, 1992)

Introduction

In this paper we investigate the moment problem of the form

$$C_{kp} = \int_0^T g(t)t^p e^{-\lambda_k t}\mathrm{d}t \quad (p=0,\dots,r_k-1;\ k=1,2,\dots). \tag{1}$$

As we have seen in [1], this problem plays an important role in the control of distributed parameter systems.

Moment Problem

Assume that for $\{\lambda_k\}$ the following asymptotic formula is valid

$$\lambda_k = \left(\frac{\pi k}{L}\right)^{2n}\left[1+O\left(\frac{1}{k}\right)\right]; \quad L>0 \quad (k=1,2,\dots). \tag{2}$$

Also suppose that $r_k=0$ except for a finite number of indices k. Denote by K_i the set of exceptional indices. Our plan is the following: we investigate the moment problem (1) in $L^2(0,T)$, we prove that the family of functions $\{t^p e^{-\lambda_k t}\}$ is minimal in $L^2(0,T)$ and hence it has a biorthogonal system $\{\Theta_{kp}\}$.

Consequently, the formal solution of (1) has the form:

$$g(t)=\sum_{k,p} C_{kp}\overline{\Theta_{kp}(t)}. \tag{3}$$

After we estimate the norm of the functions Θ_{kp} in the space $L^2(0,T)$ and describe the set of coefficients $\{C_{kp}\}$.

In this way we obtain the solution $g \in L^2(0,T)$ of the moment problem (1).

It is easy to check that $E_0 := \{e^{-\lambda_k t}\}$ is minimal in $L^2(0,T)$. Denote by $S_0(T)$ the minimal closed subspace in $L^2(0,T)$ which contains E_0. The functions $\{t^p e^{-\lambda_k t}\}$ $(p = 1,\dots,r_k - 1;\ k \in K_i)$ do not belong to E_0, hence

$$E := \{t^p e^{-\lambda_k t}\ :\ p = 0,1,\dots,r_k - 1;\ k = 1,2,\dots\}$$

is minimal in $L^2(0,T)$.

We construct the biorthogonal system $\{\Theta_{kp}]$ of E as follows. Denote

$$S(T) := V_{L^2(0,1)}E,\quad S_{k,p}(T) := V_{L^2(0,1)}(E \setminus t^p e^{-\lambda_k t}).$$

Let $\zeta_{kp}(t)$ be the projection of $t^p e^{-\lambda_k t}$ onto $S_{kp}(T)$ and

$$(4)\quad \Theta_{kp}(t) := \frac{t^p e^{-\lambda_k t} - \zeta_{kp}(t)}{\|t^p e^{-\lambda_k t} - \zeta_{kp}(t)\|^2_{L^2(0,T)}},\quad (p = 0,1,\dots,r_k - 1;\ k = 1,2,\dots).$$

It is easy to see that Θ_{kp} is biorthogonal to E, the linear hull of E_0. Any biorthogonal system $\Theta^1_{kp}(t)$ to E has the form $\Theta^1_{kp}(t) = \Theta_{kp}(t) + \hat{\Theta}_{kp}(t)$ where $\hat{\Theta}_{kp}$ belongs to the orthogonal complement of $S(T)$ in $L^2(0,1)$. Hence $\|\Theta_{kp}\| \le \|\Theta^1_{kp}\|$ i.e. the system $\{\Theta_{kp}\}$ biorthogonal to E consists of elements with minimal norm.

The series (3) is a formal solution of the moment problem (1), hence we investigate the set of coefficients $\{C_{kp}\}$ at which the series (3) converges absolutely, i.e.

$$(5)\qquad \sum_{k,p} |C_{k,p}| \|\Theta_{k,p}\|_{L^2(0,T)} < \infty.$$

For this we estimate the norm $\|\Theta_{kp}\|_{L^2(0,T)}$. As we see, this depends on the asymptotic behaviour of $\{\lambda_k\}$.

Since $\operatorname{Re}\lambda_k \ge C > -\infty$, we can assume without loss of generality that $\operatorname{Re}\lambda_k > 0$, $(k = 1,2,\dots)$. Indeed, otherwise it is enough to set $g(t) = {}=: e^{-\lambda t}h(t)$, $\lambda > -\inf_k \operatorname{Re}\lambda_k$ and to consider the moment problem for the function $h(t)$:

$$\int_0^T e^{-(\lambda_k+\lambda)t} t^p h(t)\,dt = C_{kp}.$$

We need the following

LEMMA. *Let $\{e_k\}_1^\infty$ be any minimal system in a Hilbert space and $\{\psi_k\}_1^\infty$ $\{\varphi_k\}_{N+1}^\infty$ be biorthogonal to $\{e_k\}_1^\infty$, resp. $\{e_k\}_{N+1}^\infty$ constructed as in (3). Then there exists $C \in (0,\infty)$ such that*

$$\|\varphi_k\| \le |\psi_k\| \le C|\varphi_k\| \quad (k = N+1, N+2, \ldots .$$

PROOF. Let S_0, S_1 and S denote the closed linear hull of $\{e_k\}_{N+1}^\infty$, $\{e_k\}_1^\infty$, $\{e_k\}_1^\infty$ respectively. Then from (3) we have $\psi_k = \varphi_k + \chi_k$ ($k = N+1, N+2, \ldots$), where $\psi_k \in S$, $\varphi_k \in S_0$, $\chi_k \in S \ominus S_0$ and hence the inequality $\|\psi_k\| \le \|\varphi_k\|$ follows.

Now prove $\|\psi_k\| \le C\|\varphi_k\|$. If this is not true, then there exist sequences $\{\varphi_j\}$, $\{\psi_j\}$ such that $\|\varphi_j\|\|\psi_j\|^{-1} \to 0$ $(j \to \infty)$. Consequently

$$\psi_j\|\psi_j\|^{-1} - \chi_j\|\psi_j\|^{-1} \to 0 \quad (j \to \infty).$$

Thus, the sequence $\chi_j\|\psi_j\|^{-1}$ is bounded because

$$\chi_j\|\psi_j\|^{-1} \in S \ominus S_0, \quad \dim S \ominus S_0 < \infty.$$

Hence there exists a subsequence $\{\chi_{js}\|\chi_{js}\|^{-1}\}_{s=1}^\infty$ which converges to some

$$\chi \in S \ominus S_0, \quad \|\chi\| = 1.$$

Hence $\psi_k \perp \{e_i\}_1^N$ $(k > N)$, because $\chi \perp S_1$. This contradiction completes the proof of the Lemma. □

It is known that

$$\operatorname{dist}_{L^2(0,\infty)}(e^{-\lambda_k t}, E_0 \mid e^{-\lambda_k t}) = (2\operatorname{Re}\lambda_k)^{-\frac{1}{2}} \prod_{j=1}^{\infty} \left| \frac{\lambda_j - \lambda_k}{\overline{\lambda_j} + \lambda_k} \right|.$$

On the other hand, we obtain from (3) that $\|\Theta_k\| = [d_k(T)]^{-1}$.

As L. SCHWARTZ proved [3] the operator $S_0(\infty) \to S_0(T)$, associating with every function its restriction to $[0,T]$, has bounded inverse. Hence, for every $T > 0$ there exists $C_T > 0$ such that

$$\|\Theta_k\|_{L^2(0,T)} \asymp C_T|\lambda_k|^{\frac{1}{2}} \prod_{j=1}^{\infty} \left| \frac{\overline{\lambda_j} + \lambda_k}{\lambda_j - \lambda_k} \right|.$$

Now apply Lemma to the systems $\{e^{-\lambda_k t}\}$ and $\{t^p e^{-\lambda_k t}\}$, we get $f(x) \asymp g(x)$ because $C_1 f(x) \le g(x) \le C_2 f(x)$ for every x.

$$\|\Theta_{kp}\|_{L^2(0,T)} \asymp K_T|\lambda_k|^{\frac{1}{2}} \prod_{j=1}^{\infty} \left| \frac{\overline{\lambda_j} + \lambda_k}{\lambda_j - \lambda_k} \right|,$$

where $K_T > 0$ is a constant, depending on T. Hence we get

$$\|\Theta_{kp}\|_{L^2(0,T)} = K_T \exp\{M_{2n}k + O(k)\},$$

where

$$M_\beta = 2v \cdot p \cdot \int_0^\infty \frac{\tau^{1/\beta}}{\tau^2 - 1} d\tau.$$

Thus we have proved the following

THEOREM. *The moment problem* (1) *has an absolutely integrable solution, if for some* $\varepsilon > 0$

$$\sum_{k,p} |C_{kp}| \exp\{M_{2n}k + \varepsilon k\} < \infty.$$

References

[1] AL-JANABI, A. S. A., On the control distributed parameter systems I., *Annales Univ. Sci. Budapest.*, Sect. Math., **36** (1993), 67–70.

[2] YOUNG, M. R., *An introduction to nonharmonic Fourier series*, Acad. Paris, New York, 1980.

[3] SCHWARTZ, L., *Étude des sommes, d'exponentielles*, Paris, Hermann, 1959. 152 p.

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 71–74

ON THE CONTROL OF DISTRIBUTED PARAMETER SYSTEMS III (CONTROLLABILITY)

By

ABDUL SAMEE ABDUL RAZZAK AL-JANABI

Dept. of Mathematics, College of Science, Al-Mustansiriya University, Baghdad, IRAQ

(Received September 7, 1992)

Introduction

In this paper we investigate the complete controllability of a time-delayed system of the form

(1) $$y_t = (py_x)_x + qy$$

(2) $$y_x(0,t) = 0, \qquad y_x(l,t) = by(0,t-\tau) + u(t)$$

(3) $$y(x,t) = y_0(x,t), \qquad (x,t) \in (0,l) \times [-\tau, 0]$$

$$y(0,t) = \psi(t), \qquad \tau \in [-\tau, 0]$$

where $0 < p \in C^1[0,l]$, $b \in \mathbb{R}$ (real) $\tau > 0$, $u \in L^2(0,T)$, $y_0 \in C([-\tau,0], L^2(0,l))$; $\psi \in L^2(-\tau,0)$ and $q \in C[0,l]$.

y is called a solution of (1)–(3) if

(4) $$\int_0^l \int_0^T y[z_t + (pz_x)_x + qz]\mathrm{d}x\mathrm{d}t =$$

$$= -\int_0^l z(x,0)y_0(x,0)\mathrm{d}x - \int_0^\tau p(l)z(l,t)[b\psi(t-\tau) + \nu(t)]\mathrm{d}t -$$

$$- \int_\tau^T p(l)z(l,t)[by(0,t-\tau) + u(t)]\mathrm{d}t$$

holds for any $z \in C^2([0,l] \times [0,T])$ with $z_x(0,t) = z_x(l,t)$, $0 \le t \le T$, $z(.,T) = 0$.

Complete Controllability

We can easily prove that by using the "step by step" method and the theory of nonhomogeneous boundary value problems developed in [1], that for any $T > 0$ and any fixed control function $u \in L^2(0,T)$ the problem (1)–(3) has a unique generalized solution y (defined above) and $y \in C([0,T], L^2(0,l))$. The proof uses the same idea as in I, [4] so we omit it. Our aim is to prove the following Theorem.

THEOREM. *The system* (1)–(3) *is strongly controllable for any* $T > 0$ *i.e. for any initial condition* (3) *there is a* $\nu \in L^2(0,T)$ *such that* $y(.,t) = 0$ *for* $t > T$.

PROOF. Let $T = T_0 + m\tau$ ($m \in \mathbb{N}$, $0 < T_0 \leq \tau$). As in I, we can write down the condition of (1)–(3) in the form

$$y(x,t) = \sum_{n=0}^{\infty} C_n(t)\varphi_n(x), \tag{5}$$

further

$$y_0(x,t) = \sum_{n=0}^{\infty} C_n^0(t)\varphi_n(x) \tag{6}$$

where (φ_n) are the eigenfunctions of the boundary-value problem

$$-(p\varphi_x)_x + q\varphi = \lambda\varphi \qquad (x \in (0,l)), \quad \varphi_x(0) = \varphi_x(\ell) = 0 \tag{7}$$

$$\dot{C}_n(t) + \lambda_n C_n(t) = \gamma_n[by(0,t-\tau) + u(t)] \tag{8}$$

$$C_n(0) = C_n^0, \quad \gamma_n = p(l)\cdot\varphi_n(l), \quad y(0,t) = \psi(t) \quad (t \in [-\tau, 0])$$

and λ_n are the eigenvalues of the boundary-value problem (7).

We prove first that we can arrive to the case $C_n(T_0) = 0$ ($n = 0,1,2,\ldots$). In the case of $T > \tau$ we can write the solution of (8) in the form:

$$C_n(t) = C_n^0 e^{-\lambda_n t} + \int_0^t \gamma_n[b\psi(t-\tau) + u(t)]e^{-\lambda_n(t-s)}\,ds. \tag{9}$$

Hence the conditions $C_n(T_0) = 0$ ($n = 0,1,2,\ldots$) the equivalent to the following:

$$-C_n^0 e^{-\lambda_n T_0} = \int_0^{T_0} \gamma_n[b\psi(t-\tau) + u(t)]e^{-\lambda_n(T_0-t)}\,dt \qquad (n = 0,1,2,\ldots). \tag{10}$$

We look for $u(t)$ in the form

$$u(t) = -b\psi(t-\tau) + \nu(t) \quad (0 \le t \le T_0),$$

where $\nu(t)$ will be chosen approximately from $L^2(0,T_0)$ normaly so that

$$-C_n^0 e^{-\lambda_n T_0} = \gamma_n \int_0^{T_0} \nu(t) e^{-\lambda_n (T_0 - t)} \mathrm{d}t \tag{11}$$

be fulfilled.

Hence we reduced our problem T_0 to the moment problem of the form (II) [3]. The solvability of (II) [3] depends on the properties of the system $\{e^{-\lambda_n(T_0-t)}\}_0^\infty$; in our case $\lambda_n \sim n^2$ i.e. $\sum \frac{1}{\lambda_n} < \infty$ is fulfilled, (see [2]). This system is minimal in $L^2(0,T_0)$. Consequently the system $\{e^{-\lambda_n(T_0-t)}\}_0^\infty$ is a biorthogonal one $\{\sigma_n(t)\}$, i.e.

$$\int_0^{T_0} \sigma_n(t) e^{-\lambda_n(T_0-t)} \mathrm{d}t = \delta_{n,m} = \begin{cases} 0, & \text{if } n \neq m \\ 1, & \text{if } n = m \end{cases}$$

and we can write down the solution of (11) in form

$$\nu(t) = \sum_{n=0}^{\infty} \tilde{C}_n \sigma_n(t) \tag{12}$$

where $\tilde{C}_n = -\frac{C_n^0 e^{-\lambda_n T_0}}{\gamma_n}$.

Taking into account the estimate ([2]) we have

$$\|\sigma_n\|_{L^2(0,T_0)} \le M(T_0) e^{(L+\varepsilon)n}, \quad \left(n = 0,1,2,\ldots,\ \varepsilon > 0,\ L = \int_0^l \frac{\mathrm{d}x}{\sqrt{p(x)}} \right).$$

We can check that the series (12) converges absolutely, i.e.

$$\sum_{n=0}^{\infty} |C_n| \|\sigma_n\|_{L^2(0,T_0)} < \infty$$

(we use here $\lambda_n \sim n^2$ and $0 < C_1 < |\gamma_n| < C_2$ with absolute constant C_1, C_2). Consequently, the moment-problem (11) has a solution in $L^2(0,T_0)$ for any $T_0 > 0$ and $\{C_n^0\} \in l^2$. Taking into account that the function φ belongs to $L^2(0,T_0)$ we obtain a control function u in $L^2(0,T_0)$ such that $C_n(T_0) = 0$ $(n = 0,1,\ldots)$.

Now take $u(t) = -by(0, t-\tau)$ if $t > T_0$ because $y(0,\cdot) \in L^2(0,T)$ hence $u \in L^2(0,T)$ and it follows from the solution of (8) that $C_n(t) = 0$ for $t \geq T_0$.

If $t \geq T_0 + \tau$ then $y(0, t-\tau) = 0$ and hence $u(t) = 0$ if $t \geq T_0 + \tau$ and our theorem is proved. □

References

[1] LIONS, J.-L., *Optimal control of systems governed by partial differential equations*, Dunod, Paris, 1968.

[2] RUSSEL, D. L., FATTORINI, H. O., Exact controllability theorems for linear parabolic equations in one space dimensions, *Arch. for Rational Mech. and Anal.*, **43** (1971).

[3] AL-JANABI, A. S. A., On the control distributed parameter systems I., *Annales Univ. Sci. Budapest.*, Sect. Math., **36** (1993), 61–66.

[4] AL-JANABI, A. S. A., On the control distributed parameter systems II., *Annales Univ. Sci. Budapest.*, Sect. Math., **36** (1993), 67–70.

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 75–79

ON THE CONTROL OF DISTRIBUTED PARAMETER SYSTEMS IV (CONTROLLABILITY)

By

ABDUL SAMEE ABDUL RAZZAK AL-JANABI

Dept. of Mathematics, College of Science, Al-Mustansiriya University, Baghdad, IRAQ

(Received September 7, 1992)

Introduction

In this paper we consider and investigate the controllability of a hyperbolic system

$$(1)\qquad u_{tt} = \sum_{k=0}^{2n} q_k(x) \frac{\partial^{2n-k}}{\partial x^{2n-k}} u = (\sigma u(.,t))(x)$$
$$(0 < t < T < \infty, \qquad 0 < x < X < \infty).$$

with boundary conditions

$$(2)\qquad \sum_{k=0}^{2n-1} \left[\alpha_{jk}^0 u^{(k)}(0,t) + \alpha_{jk}^1 u^{(k)}(l,t)\right] \equiv \tau_j u(.,t) = \sum_{\mu=1}^{l} \beta^\mu f_\mu(t)$$
$$(j = 1, 2, \ldots, 2n),$$

where we suppose that $P_k \in H^{2n-k}(0,X)$, $P_0(x) > 0$; $f_\mu \in H^1(0,T)$ are the control functions and we suppose that τ_j implies strongly regular homogeneous boundary condition [1].

Our initial conditions are

$$(3)\qquad \begin{aligned} u(x,0) &= u_0(x), \quad u_0 \in L^2(0,X) \\ u_t(x,0) &= \hat{u}_0(x), \quad \hat{u}_0 \in H^{-n}(0,X). \end{aligned}$$

The Controllability of a Hyperbolic System

We say that $u \in L^2((0,X)\times[0,T])$ is a solution of (1)–(3) if

$$\int_0^X\int_0^T u\left[\overline{\nu_{tt}} - \overline{\sigma^*\nu}\right]\mathrm{d}x\mathrm{d}t = \tag{4}$$

$$= \int_0^X\left[\hat{u}_0(x)\overline{\nu(x,0)} - u_0(x)\overline{\nu_t(x,0)}\right]\mathrm{d}x + \sum_{j=1}^{2n}\sum_{\mu=1}^{l}\beta_j^\mu\int_0^T f_\mu(t)\overline{\tilde{\tau}_j^*\nu(.,t)}\mathrm{d}t$$

holds for any $\nu \in H^{n,2}([0,X]\times[0,T])$ for which

$$\nu(.,T) = \nu_t(.,T) = 0, \quad \tau_j^*\nu(.,t) = 0, \qquad j = 1,2,\ldots,2n.$$

Here the boundary operators τ_j^*, $\tilde{\tau}_j$, $\tilde{\tau}_j^*$, $\sigma^*\nu$ and σu are the same as in [1]. In this paper we reduce the control to a moment problem, but according to the hyperbolicity of the system, we have to apply a different approach. We look for the solution of (1)–(3) in the form

$$u(x,t) = \sum_{k,s} a_{k,s}(t)\varphi_{ks}(x) \qquad (s = 1,2,\ldots,r_k;\quad k = 1,2,\ldots), \tag{5}$$

where φ_{ks} are the generalized eigenfunctions of the operator A defined by $-\sigma\varphi$ and by the boundary conditions $\tau_j\varphi = 0$ ($j = 1,2,\ldots,2n$).

We can write u_0 and $\hat{u}_0$ in the form

$$u_0(x) = \textstyle\sum_{ks} a_{ks}^0\varphi_{ks}, \quad \sum_{ks}|a_{ks}^0|^2 < \infty$$

$$\hat{u}_0(x) = \textstyle\sum_{ks} \hat{a}_{ks}^0\varphi_{ks}, \quad \sum_{ks}|\hat{a}_{ks}^0|^2|k|^{-2k} < \infty.$$

We apply (4) for $\nu(x,t) = \overline{w(t)}\psi_{ks}(x)$ where $w(t) \in \mathrm{C}^2(0,T)$ is any such function for which $w(T) = w'(T) = 0$ and (ψ_{ks}) is the dual system of the eigenfunctions (φ_{ks}) of A^*. Substituting this particular $\nu(x,t)$ into (4), for $a_{ks}(t)$ we get the following differential equations:

$$\ddot{a}_{ks}(t) + \lambda_k a_{ks}(t) + a_{ks+1}(t) = \sum_{j=1}^{2n}\sum_{\mu=1}^{l}\beta_j^\mu\gamma_{ks}^j f_\mu(t) \tag{7}$$

$$a_{ks}(0) = a_{ks}^0, \quad \dot{a}_{ks}(0) = \hat{a}_{ks}^0 \qquad (s = 1,2,\ldots,r_k;\ k = 1,2,\ldots)$$

where

$$\gamma_{ks}^j = \overline{\tau_j^*\psi_{ks}}, \qquad a_{kr_k+1} = 0.$$

Now apply the Laplace transform to (7) get its solution $a_{ks}(t)$ in terms of Bessel functions

$$
\begin{aligned}
a_{ks}(t) = &\sum_{j=1}^{2n}\sum_{\mu=1}^{l}\beta_j^{\mu}\int_0^t f_\mu(\tau)\times \\
&\times\sum_{p=0}^{2k-s}\gamma_{ks+p}^{j}\frac{\sqrt{\pi}(-1)^p}{p!}\left(\frac{t-\tau}{2\omega_k}\right)^{p+\frac{1}{2}} J_{p+\frac{1}{2}}(\omega_k(t-\tau))\mathrm{d}\tau+ \\
&+\sum_{p=0}^{2k-s}(-1)^p a_{ks+p}^{0}\frac{\sqrt{\pi}\omega_k}{p!}\left(\frac{t}{2\omega_k}\right)^{p+\frac{1}{2}} J_{p-\frac{1}{2}}(\omega_k t)+ \\
&+\sum_{p=0}^{2k-s}(-1)^p \hat{a}_{ks+p}^{0}\frac{\sqrt{\pi}}{p!}\left(\frac{t}{2\omega_k}\right)^{p+\frac{1}{2}} J_{p+\frac{1}{2}}(\omega_k t),
\end{aligned}
\tag{8}
$$

where $\omega_k^2 = \lambda_k$ define ω_k.

For large k we have $r_k = 1$ i.e. the eigenvalues of A have multiplicity 1, by strong regularity. In this case (7) has the simpler form

$$
\ddot{a}_{ks}(t) + \lambda_k a_{ks}(t) = \sum_{j=1}^{2n}\sum_{\mu=1}^{l}\beta_j^{\mu}\gamma_{ks}^{j} f_\mu(t)
\tag{7'}
$$

and its solution is

$$
\begin{aligned}
a_{ks}(t) &= \\
&=\sum_{j=1}^{2n}\sum_{\mu=1}^{l}\beta_j^{\mu}\gamma_{ks}^{j}\omega_k^{-1}\int_0^t f_\mu(\tau)\sin\omega_k(t-\tau)\mathrm{d}\tau + a_{ks}^0\cos\omega_k t + \frac{\hat{a}_{ks}^0}{\omega_k}\sin\omega t_k t.
\end{aligned}
\tag{8'}
$$

For convenience we will use the notations

$$
\begin{aligned}
&\omega_{-k} = -\omega_k, \qquad r_{-k} = r_k, \qquad \gamma_{-ks}^{j} = \gamma_{ks}^{j} \qquad k = 1, 2, \ldots \\
&b_{ks}(t) = i\omega_k a_{|k|s}(t) + \dot{a}_{|k|s}(t) \\
&b_{ks}^0 = i\omega_k a_{|k|s}^0 + \hat{a}_{|k|s}^0 \qquad k = \pm1, \pm2, \pm3, \ldots
\end{aligned}
$$

There is (obviously) a one-to-one correspondence between $\{a_{ks}(t), \dot{a}_{ks}(t)\}_k$ and $\{b_{ks}(t), b_{-ks}(t)\}_k$, and also between $\{a_{ks}^0, \hat{a}_{ks}^0\}$ and $\{b_{ks}^0, b_{-ks}^0\}$. Consequently there is a one-to-one correspondence between the states

$$
\{u(.,t), u_t(.,t)\} \qquad (0 \le t \le T)
$$

of the system (1)–(3) and between

$$\{b_{ks}(t)\ :\ s=1,2,\dots,r_k;\quad k=\pm1,\pm2,\dots\}.$$

It is easy to see (taking into account the expressions of $J_{p+\frac{1}{2}}$ in terms of trigonometric function) that (8) is equivalent to

$$(8^*)\qquad b_{ks}(t)=\sum_{p=0}^{r_k-s} b^0_{ks+p}\mathcal{H}^{0k}_{sp}t^p e^{i\omega_k t}+$$
$$+\sum_{j=1}^{2n}\sum_{\mu=1}^{l}\sum_{p=0}^{r_k-s}\beta^\mu_j\mathcal{H}^{jk}_{sp}\int_0^t f_\mu(\tau)(t-\tau)^p e^{i\omega_k(t-\tau)}\mathrm{d}\tau$$
$$(s=1,2,\dots,r_k;\quad k=\pm1,\pm2,\pm3,\dots).$$

Here we do not need the complicated form of $\mathcal{H}^{0k}_{sp}$ and $\mathcal{H}^{jk}_{sp}$, hence we do not write down them explicitly (but in our following papers their investigation will play a central role). Hence we need only the fact that $\mathcal{H}^{0k}_{sr_k-s}\neq 0$ and $\mathcal{H}^{jk}_{sr_k-s}$ is proportional to $\gamma^j_{kr_k}$. For large value of k, taking into account $(8')$ we have

$$(8^{**})\qquad b_{ks}(t)=b^0_{ks}e^{i\omega_k t}+\sum_{j=1}^{2n}\sum_{\mu=1}^{l}\beta^\mu_j\gamma^j_{ks}\int_0^t f_\mu(\tau)e^{i\omega_k(t-\tau)}\mathrm{d}\tau.$$

Now we are in the position to prove that the function $t\to(u(.,t),u_t(.,t))$ is a continuous mapping of $[0,T]$ into $L^2(0,X)\times H^{-n}(0,X)$.

Taking into account $\omega_k\sim k^n$

$$L=\int_0^X\frac{\mathrm{d}x}{\sqrt[n]{P_0(x)}},\qquad \sin\omega_k=O(1),\qquad \gamma^j_{ks}=O(k^{2n-1})$$

we obtain from (8^*) and (8^{**}) that

$$\sum_{k,s}|k|^{-2k}|b_{ks}(t)|^2<\sum_{k>s}|k|^{-2k}|b^0_{ks}|^2+\sum_{\mu=1}^{l}\|f_\mu\|^2_{H^1(0,T)},$$

and hence, taking into account

$$\sum_{k,s}|k|^{-2k}|b_{ks}(t)|^2\sim\sum_{k,s}|k|^{-2k}\left(|\omega_k a_{ks}(t)|^2+|\dot a_{ks}(t)|^2\right)\sim$$
$$\sim\|u(.,t)\|^2_{L^2(0,X)}+\|u_t(.,t)\|^2_{H^{-n}(0,X)},$$

we obtain

$$\|u(.,t)\|^2_{L^2(0,X)} + \|u_t(.,t)\|^2_{H^{-n}(0,X)} \leq$$
$$\leq \|u_0\|^2_{H^2(0,X)} + \|\hat{u}_0\|^2_{H^{-n}(0,X)} + \sum_{\mu=1}^{l} \|f_\mu\|^2_{H^1(0,T)}.$$

One can prove similarly, that the sum $\|u(.,t)\|_{L^2(0,X)} + \|u_t(.,t)\|^2_{H^{-n}(0,X)}$ is continuous in t.

Now we investigate the control of our system (1)–(3). Let $(u_1, \hat{u}_1)$ be any given state in $L^2(0,X) \times H^{-n}(0,X)$ of the system considered. Expand u_1 and $\hat{u}_1$ in the form

$$u_1(x) = \sum_{k,s} a^1_{ks}\varphi_{ks}(x) \qquad \hat{u}_1(x) = \sum_{k,s} \hat{a}^1_{ks}\varphi_{ks}(x).$$

Then the system is in the state $(u_1, \hat{u}_1)$ in time T if and only if

(9) $$b^1_{ks} = b_{ks}(T) \qquad (s = 1,2,\ldots,r_k; \quad k = 1,2,\ldots).$$

This is fulfilled, where $b_{ks}(T)$ is defined in (8^*), (8^{**}) and

$$b^1_{ks} = i\omega_k a^1_{|k|s} + \hat{a}^1_{|k|s}.$$

So we have proved the equivalence of our control problem to that of prescribing the set $[b_{ks}, b^1_{ks}\}$ for which the moment problem (9) has a solution $f_\mu \in H^1(0,T)$ $(\mu = 1,2,\ldots)$.

We have proved

THEOREM. *The system* (1)–(3) *has unique generalized solution* u *in* $L^2([0,X] \times [0,X])$ *and this solution* $(u(.,t), u_t(.,t))$ *depends continuously on* t *in* $L^2(0,X) \times H^{-n}(0,X)$.

References

[1] HARTMAN, P., *Ordinary Differential Equations*, Wiley and sons, New York, 1964.

[2] AL-JANABI, A. S. A., On the control distributed parameter systems I., *Annales Univ. Sci. Budapest.*, Sect. Math., **36** (1993), 61–66.

[3] LIONS, J.-L., Controle Optimal des Systemes Gourernés par des Equations, Dérivées Partielles, Dunod, Paris, 1968.

[4] LIONS, J.-L., Contrôlabilité exacte et stabilization de Systémes distributés, Contrôlabilité exacte, Massen, Paris, 1988.

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 81–84

ON THE CONTROL OF DISTRIBUTED PARAMETER SYSTEMS V (MOMENT PROBLEM)

By

ABDUL SAMEE ABDUL RAZZAK AL-JANABI

Dept. of Mathematics, College of Science, Al-Mustansiriya University, Baghdad, IRAQ

(Received September 7, 1992)

Introduction

Consider the moment problem

$$C_{kp} = \int_0^T g(t) t^p e^{i\omega_k t} \mathrm{d}t \tag{1}$$

$$(p = 1, 2, \ldots, r_k - 1; \quad k = \pm 1, \pm 2, \ldots; \quad \sum_{k,p} |C_{kp}|^2 < \infty).$$

This problem played a central role in our paper IV [1]. Assume that $\sin \omega_k = O(1)$ and $\omega_k \sim k^n$ $(n > 1)$ are satisfied. (This is the most important case for control of hyperbolic systems.)

Moment Problem

Now we consider only the case $n > 1$ because if $n = 1$, then the problem seems to be harder and we do not know the solution of the case $n = 1$ yet.

First we investigate the moment problem (1) in $L^2(0,T)$ and after in $H^m(0,T)$. For this we have to show that we can complete $\{t^p e^{i\omega_k t}\}$ so that the resulting system is a Riesz basis in $L^2(0,T)$. To this consider the orthonormal basis

$$E_0 = \left\{ e^{i\frac{2\pi}{T}j} \right\}; \qquad (j = 0, \pm 1, \pm 2, \ldots)$$

in $L^2(0,T)$. Starting from E_0 we shall construct a family $\tilde{E}_0$, which contains the family

$$E = \{t^p e^{i\omega_k t}\}; \qquad (p = 0, 1, 2, \ldots, r_k - 1, \quad k = \pm 1, \pm 2, \ldots).$$

We do this as follows. Let N be large enough such that $r_k = 1$ and

$$|\mathrm{Re}\omega_n - \mathrm{Re}\omega_s| > \frac{2\pi}{T} \quad \text{if} \quad |s| > |n| > N.$$

Denote by E_1 the corresponding family of functions $\{e^{i\omega_k t}\}$. According to $\omega_k \sim k^n$ the set $E_2 := E \setminus E_1$ is finite for any $e^{i\omega_k t} \in E_1$. Choose $e^{i\frac{2\pi}{T}j}$ such that $\frac{2\pi}{T}j$ is at minimal distance on the real line from $\mathrm{Re}\omega_k$ (we can choose any of the two such points). Pick functions from E_1 into E_0 instead of the corresponding function of the form $e^{i\frac{2\pi}{T}j}$ and the resulting system of functions E_{01} write down in the form

$$\left\{e^{i\frac{2\pi}{T}(j+\delta_j)}\right\} \qquad (j = 0, \pm1, \pm2, \ldots);$$

where $\delta_j = 0$ or $j + \delta_j$ coincides with some of ω_k, according to the construction we have $|\mathrm{Re}\delta_j| \leq \frac{1}{2}$. In this way we obtain that E_{01} is a Riesz basis in $L^2(0,T)$ (see [2]).

With any function from E_2 we associate a function from E_{01} for which $\delta_j = 0$ (with different functions we associate different functions).

After we change resulting functions $e^{i\frac{2\pi}{T}j}$ (form E_{01}) to functions form E_2. Accordingly to [2] the resulting family $\tilde{E}_0$ of functions remains a Riesz basis in $L^2(0,T)$ and by the construction, it contains the set E. Thus, we have proved that $\{t^p e^{i\omega_k t}\}$ is minimal and forms a Riesz basis in its linear hull in $L^2(0,T)$. Consequently, there exists a biorthogonal system

$$\{\Theta_{kp} \ : \ p = 0, 1, 2, \ldots, r_k - 1; \quad k = \pm1, \pm2, \pm3, \ldots\}$$

to it. The closure of the linear hull of E is an eigensubspace in $L^2(0,T)$, hence the biorthogonal system $\{\Theta_{kp}\}$ is not unique. Thus we consider the functions Θ with minimal norm (we choose this biorthogonal system) and according to a known theorem of BARI [2], this family belongs to E. Now the solution of our moment problem has the form

$$g(t) = \sum_{k,p} C_{kp} \overline{\Theta_{kp}(t)}$$

and this solution has minimal norm, further we have

$$\|g\|^2_{L^2(0,T)} \sim \sum_{k,p} |C_k p|^2.$$

Thus, we have proved the following

THEOREM. *The moment problem* (1) *has a solution* $g \in L^2(0,T)$ *for any* $T > 0$ *and for any* $\{C_{kp}\} \in \ell^2$.

Now we have to consider its solution on $H^m(0,T)$. For this it is enough to consider the m-th integral of the system $\{t^p e^{i\omega_k t}\}$, elements of which consists from the linear combinations of the functions $t^s e^{i\omega_k t}$: $(s = 0, 1, \ldots, p)$. Denote the elements of this linear hull by $\varphi_{kp}^{2n}(t)$. Now consider the moment problem

$$(2)\qquad \begin{aligned} 0 &= \int_0^T h(t)(T-t)^q \mathrm{d}t; \qquad q = 0, 1, 2, \ldots, m-1, \\ C_{kp} &= \int_0^T h(t)\varphi_{kp}^{2n}(t)\mathrm{d}t; \quad p = 0, 1, \ldots, r_k - 1; \quad k = \pm 1, \pm 2, \ldots . \end{aligned}$$

Using the explicit form of $\varphi_{kp}^{2n}(t)$, it is easy to check that (2) is equivalent to

$$(3)\qquad \tilde{C}_{ks} = \int_0^T h(t) t^s e^{i\omega_k t} \mathrm{d}t; \qquad (s = 0, 1, \ldots, r_k - 1, \quad k = \pm 1, \pm 2, \ldots),$$

where the coefficients $\{\tilde{C}_{ks}\}$ are uniquely determined by $\{C_{ks}\}$, further, for large k we have

$$\tilde{C}_{ks} = (i\omega_k)^m C_{ks}.$$

It follows, that if we add to the family E the system $(T-t)^q$; $q = 0, 1, \ldots, m-1$ then the resulting system will be a Riesz basis in its linear hull (in its closure) and this means that the moment problem (3) has solution $h \in L^2(0,T)$ such that

$$\|h\|^2_{L^2(0,T)} \sim \sum_{k,s} |\tilde{C}_{ks}|^2 \sim \sum_{k,p} |k|^{2mn} |C_{kp}|^2.$$

Integrating by parts we get that the function

$$g(t) = \frac{1}{(m-1)} \int_0^t (t-\tau)^{m-1} h(\tau) \mathrm{d}\tau$$

is a solution of the moment problem (2) and $\|g\|^2_{H^m(0,T)} < \infty$. □

References

[1] AL-JANABI, A. S. A., On the control distributed parameter systems IV., *Annales Univ. Sci. Budapest.*, Sect. Math., **36** (1993), 75–79.

[2] YOUNG, M. R., *An introduction to nonharmonic Fourier series*, Acad. Paris, New York, 1980.

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 85–102

SOME SERIES OF HYPERBOLIC SPACE GROUPS

By

MILICA STOJANOVIĆ

University of Belgrade, Belgrade

(Received September 7, 1992)
(Revised January 26, 1992)

1. Introduction

The isometry groups, acting discontinuously on the hyperbolic 3-space with compact fundamental domain (briefly called hyperbolic space group), has not been classified yet, so investigations in this direction are very actual now. One possibility to do that is to look for fundamental domains of these groups. Face pairing identifications of a given polyhedron may give us generators and relations for a space group by Poincaré theorem [1], [4].

The simplest fundamental domains are simplices and truncated simplices by polar planes of vertices when they are out of the absolute. There are 64 combinatorially different face pairings of fundamental simplices [7], [8], [6], furthermore 35 solid transitive simplex identifications [6]. I. K. ZHUK [7], [8] has classified Euclidean and hyperbolic fundamental simplices of finite volume up to congruence. Some completing cases and algorithmic procedure are discussed by E. MOLNÁR [3] and by E. MOLNÁR and I. PROK [5]. In [6] they summarize all these results, arranging identified simplices into 32 families. Each family is characterized by so-called maximal series of simplex tilings. Besides spherical, Euclidean, hyperbolic realizations there exist also other metric realizations in 3-dimensional simply connected homogeneous Riemann spaces.

Each identified simplex, considered in this paper, has five edges in the same equivalence class, i.e. they are mapped onto each other under the group generated by face pairing of the simplex. One edge, denoted by A_2A_3 in figures, is in another equivalence class. That means, we may have two different lengths for edges. All the vertices are in the same equivalence class. There are 11 different face pairings to investigate in this paper [6], [7].

When vertices are out of the absolute, the simplex is not compact and then we truncate it with polar planes of the vertices. The new compact polyhedron obtained in that way is fundamental domain of some larger group. It has new triangular faces whose pairing gives new generators. Dihedral angles around new edges are $\pi/2$. That means there are four congruent polyhedra around them in the fundamental space filling.

The result of this paper will be collected in the figures and tables. The summarizing Theorem 2 will be formulated in Section 6.

I am very grateful to my mentor DR EMIL MOLNÁR, who guides my studies in Budapest, for helping me in preparing this manuscript.

2. Projective metrics, spherical and hyperbolic spaces

To prepare our method of construction we consider the 4-dimensional real vector space V^4, whose dual space, i.e. the space of its linear forms, is denoted by $\mathcal{V}_4^*$. In the usual way the projective 3-space $P^3(V^4, \mathcal{V}_4^*)$ can be introduced. The 1-dimensional subspaces of V^4 (or the 3-subspaces of $\mathcal{V}_4^*$) represent the points of P^3, and the 1-subspaces of $\mathcal{V}_4^*$ (or the 3-subspaces of V^4) represent the planes of P^3. The point $X(x)$ and the plane $\alpha(a)$ are incident iff $xa = 0$, i.e. the value of the linear form a on the vector x is equal to zero ($x \in V^4 \setminus \{0\}$, $a \in \mathcal{V}_4^* \setminus \{0\}$). The straight lines of P^3 are characterized by 2-subspaces of V^4 or of $\mathcal{V}_4^*$, respectively. If $\{e_i\}$ is a basis on V^4 and $\{e^j\}$ is its dual basis on $\mathcal{V}_4^*$, i.e. $e_i e^j = \delta_i^j$ (the Kronecker symbol), then the form $a = e^j a_j$ takes the value $xa = x^i a_i$ on the vector $x = x^i e_i$. We use the summation convention for the same upper and lower indices.

We can introduce projective metric in P^3 by giving a bilinear form

$$\langle\ ;\ \rangle :\ \mathcal{V}_4^* \times \mathcal{V}_4^* \to \mathbb{R}, \quad \langle b^i u_i; b^j v_j \rangle = u_i b^{ij} v_j$$

where $(\langle b^i; b^j \rangle) = (b^{ij})$ is a Schäfli matrix, and the basis $\{b^i\}$ in $\mathcal{V}_4^*$ represents planes containing simplex faces opposite to the vertices A_i, respectively. Vectors a_j of the dual basis $\{a_j\}$ in V^4, defined by $a_j b^i = \delta_j^i$, represent the vertices of A_j of the simplex. The induced bilinear form

$$\langle\ ;\ \rangle :\ V^4 \times V^4 \to \mathbb{R} \quad \langle x^i a_i; y^j a_j \rangle = x^i a_{ij} y^j$$

is defined by the matrix $(\langle a_i; a_j \rangle) = a_{ij}$ inverse to (b_{ij}).

We assume, that the bilinear form $\langle\ ;\ \rangle$ is either of signature $(+,+,+,-)$ which characterizes the hyperbolic metric, or $(+,+,+,+)$, this will be the elliptic (spherical) metric. Signature $(+,+,+,0)$ would describe the Euclidean geometry that will not occur in our considerations.

It is well-known that the bilinear form induces the distance and the angle measure of the 3-space. Let $X(x)$ and $Y(y)$ be two points in the projective space P^3. Then their distance $d(x,y)$ is determined by

$$(1)\qquad \cos(d(x,y)) = \frac{\langle x;y\rangle}{\sqrt{\langle x;x\rangle\langle y;y\rangle}} \quad\text{and}\quad \operatorname{ch}(d(x,y)) = \frac{-\langle x;y\rangle}{\sqrt{\langle x;x\rangle\langle y;y\rangle}}$$

for elliptic and hyperbolic case, respectively.

Mention the next lemma which we need later.

LEMMA. *For any $(r+1)$-minor determinant of a regular matrix (a_{ij}) and complementary $(n-r)$-minor of its inverse (b^{ij}) holds the following equality*

$$\begin{vmatrix} a_{i_0j_0} & \dots & a_{i_0j_r} \\ \vdots & & \vdots \\ a_{i_rj_0} & \dots & a_{i_rj_r} \end{vmatrix} = \det(a_{ij}) \begin{vmatrix} b^{i_{r+1}j_{r+1}} & \dots & b^{i_{r+1}j_n} \\ \vdots & & \vdots \\ b^{i_nj_{r+1}} & \dots & b^{i_nj_n} \end{vmatrix} \sigma.$$

Here $\sigma = \operatorname{sign}(i_0,\dots,i_r,i_{r+1},\dots,i_n)\cdot\operatorname{sign}(j_0,\dots,j_r,j_{r+1},\dots,j_n)$ denotes the sign product of the corresponding permutations of the elements $0,1,\dots,n$.

All simplices (Table 2) considered here have a Schäfli matrix

$$(2)\qquad B = (b_{ij}) = \begin{bmatrix} 1 & p & q & q \\ p & 1 & q & q \\ q & q & 1 & r \\ q & q & r & 1 \end{bmatrix}, \quad\text{where}\quad \begin{aligned} p &= -\cos 2\pi/\overline{a} \\ q &= -\cos\beta \\ r &= -\cos(2\pi/\overline{b} - 4\beta) \end{aligned}$$

with the natural parameters $\overline{a} \geq 3$, $\overline{b} \geq 1$ and $0 < \beta < \pi/2\overline{b}$.

Since $d(A_0,A_1) = d(A_0,A_2)$, by the symmetry of our simplex, we have a necessary condition. Using the inverse matrix (a_{ij}) in (1) we can express any side length (or its function) by the angles, so

$$\frac{a_{01}^2}{a_{00}a_{11}} = \frac{a_{02}^2}{a_{00}a_{22}} \quad\text{or}\quad \frac{a_{00}a_{11} - a_{01}^2}{a_{00}a_{11}} = \frac{a_{00}a_{22} - a_{02}^2}{a_{00}a_{22}}.$$

Applying our lemma, we get

$$a_{22}(b_{22}b_{33} - b_{23}^2) = a_{11}(b_{11}b_{33} - b_{13}^2)$$

and using notation in (2)

$$(3)\quad f(\beta) = 0, \text{ for } f(\beta) = (1-p)(1+r)(1+p-2q^2) - (1+r-2q^2)(1-q^2)$$

If $\overline{b} = 1$ that equality becomes $4q^2(1-2p) = 1-4p^2$ and it has one solution for q and so β in cases $3 < \overline{a} < 6$. For $\overline{a} = 3$, $\overline{b} = 1$ we get a spherical extra case, i.e. a degenerate simplex $\mathcal{T}_4$ where the „angle" is just π at the „edge" A_0A_1, and $\beta = \pi/4$ (see at Table 2: $\Gamma_{55}(a, 5b)$, $a = 3$, $b = 1$). The computation agrees with K_{mn} ($m = \overline{a}$, $n = 2$) of Zhuk's cases [7], [8].

It is possible to check that if $\overline{b} \geq 2$ the function $f(\beta)$ is increasing, $f(0) < 0$ and $f(\pi/2\overline{b}) > 0$. So, the equation (3) has exactly one solution for $0 < \beta < \pi/2\overline{b}$, for any value $\overline{a} \geq 3$.

By finding the eigenvalues of matrix B we can establish in which space our simplices are realizable. Since

$$\det(B - \lambda I) =$$
$$= (\lambda - (1-p))(\lambda - (1-r))\left(\lambda - \left(\frac{p+r+2}{2} + \frac{\sqrt{(p-r)^2 + 16q^2}}{2}\right)\right) \cdot$$
$$\cdot \left(\lambda - \left(\frac{p+r+2}{2} - \frac{\sqrt{(p-r)^2 + 16q^2}}{2}\right)\right)$$

has three positive roots, we have hyperbolic. spherical or Euclidean space, which depends on sign of the fourth root. It is easy to check that for $\overline{b} = 1$ ($\overline{a} = 4, 5$) it is positive — spherical case, and for $\overline{b} \geq 2$ it is negative — hyperbolic case. As extra cases $\overline{b} = 1$, $\overline{a} = 6$ leads to a Nil-group, while $\overline{b} = 1$, $\overline{a} \geq 7$ just provide groups in $SL_2(\mathbb{R})$ as indicated also in [6].

For hyperbolic simplices it is interesting to investigate the cases when the vertices are proper, or they lie on the absolute or out of the absolute. Therefore, we need any submatrix B_{ii} of B corresponding to the vertex A_i which we obtain by excluding i-th row and i-th column. By symmetry arguments we may consider B_{00}. From the chain of the principal minor determinants of the B_{00}

$$1 > 0, \qquad 1 - q^2 > 0, \qquad (1-r)(1+r-2q^2) < 0 \qquad (\overline{b} \geq 2)$$

we can see that the signature of the partial bilinear form with matrix B_{00} is $(+,+,-)$, these vertices are out of the absolute.

3. Construction of discontinuously acting isometry groups

With any simplex $\mathcal{T}$ of angles satisfying (3) we can fill the spherical space S^3 for $\overline{b} = 1$, $\overline{a} = 3$, 4, 5, or H^3 for $\overline{b} \geq 2$ using identifications which are indicated in the figures and tables.

Identifications on the simplex $\mathcal{T}$ are face pairings by isometries, satisfying the following conditions

a) For each face $f_{g^{-1}}$ of $\mathcal{T}$ there is another face f_g and an identifying isometry g if the space S^3 (H^3), which maps $f_{g^{-1}}$ onto f_g and $\mathcal{T}$ onto $\mathcal{T}^g \not\equiv \mathcal{T}$, the neighbour of $\mathcal{T}$ along f_g.

b) The isometry g^{-1} maps the face f_g onto $f_{g^{-1}}$ and $\mathcal{T}$ onto $\mathcal{T}^{g^{-1}}$, joining the simplex $\mathcal{T}$ along $f_{g^{-1}}$.

The face paring identifications of $\mathcal{T}$ generate an isometry group G.

These generators induce subdivision of the edges into oriented segments such that a segment does not contain two equivalent points in its interior. An equivalence class consisting of edge segments $e_1, e_2, \ldots, e_r$ with dihedral angles $\varepsilon(e_1)$, $\varepsilon(e_2)$, ..., $\varepsilon(e_r)$, respectively, is defined as follows.

We consider an edge segment, say e_1, and choose one of the faces denoted by $f_{g_1^{-1}}$ whose boundary contains e_1. The isometry g_1 maps e_1 and $f_{g_1^{-1}}$ onto e_2 and f_{g_1}. There exists exactly one other face $f_{g_2^{-1}}$ with e_2 on the boundary, furthermore the isometry g_2 maps e_2 and $f_{g_2^{-1}}$ onto e_3 and f_{g_2}, and so on. We obtain a cycle of isometries $g_1, g_2, \ldots, g_r$ according to the scheme

(4)
$$(e_1, f_{g_1^{-1}} \xrightarrow{g_1} (e_2, f_{g_1});\ (e_2, f_{g_2^{-1}} \xrightarrow{g_2} (e_3, f_{g_2});\ \ldots;\ (e_r, f_{g_r^{-1}} \xrightarrow{g_r} (e_1, f_{g_r});$$

where the symbols are not necessarily distinct. More precisely, we have two essentially different cases for the scheme (4)

1) if a plane reflection $m_i = g_i$ occurs then $e_{i+1} = e_i$, and we turn back to e_1, then, say, e_{-1} comes. Furthermore, another plane reflection $m_{-j} = g_{-j}$ shall appear in the cycle. Then each edge segment comes two times in the scheme (4), and the cycle transformation is of the form

$$c = g_1 g_2 \cdots g_r = (g_1 \cdots g_{i-1} m_i g_{i-1}^{-1} \cdots g_1^{-1})(g_{-1}^{-1} \cdots g_{-j+1}^{-1} m_{-j} g_{-j+1} \cdots g_{-1})$$

2) there is no plane reflection in the cycle, this will be the simpler (general) case. (In dimension 3 we have 5 subcases for full edges at all [4].)

In other words the segment e_1 is successively surrounded by simplices

$$\mathcal{T},\ \mathcal{T}^{g_1^{-1}},\ \mathcal{T}^{g_2^{-1} g_1^{-1}},\ \ldots,\ \mathcal{T}^{g_r^{-1} \cdots g_2^{-1} g_1^{-1}}$$

which fill an angular region of measure $2\pi/\nu$. In above case 1) holds

(5) $$\varepsilon(e_1) + \ldots + \varepsilon(e_i) + \varepsilon(e_{-1} + \ldots + \varepsilon(e_{-j+1}) = \pi/\nu.$$

In case 2) we have

(6) $$\varepsilon(e_1) + \ldots + \varepsilon(e_r) = 2\pi/\nu.$$

Finally, the cycle transformation $c = g_1 g_2 \cdots g_r$ belonging to the edge segment class $\{e_i\}$ is a rotation, say, of order ν. Thus we have the cycle relation in both cases

$$(7) \qquad (g_1 g_2 \cdots g_r)^\nu = 1$$

c) Assume that (5) or (6) holds for the face angles at $\{e_1\}$ in each segment equivalence class.

We need the specified Poincaré theorem:

THEOREM 1. *Let $\mathcal{T}$ be a simplex, or a truncated simplex in a space $\mathcal{S}^3$ of constant curvature (now S^3 or H^3) and G be the group generated by the face identifications, satisfying conditions* a)–c) *and* (3) *also holds. Then G is discontinuously acting group on $\mathcal{S}^3$, $\mathcal{T}$ is a fundamental domain for G and the cycle relations of type* (7) *for every equivalence class of edge segments form a complete set of relations for G, if we also add the relations $g_i^2 = 1$ to the occasional involutive generators $g_i = g_i^{-1}$.*

4. The isometry groups for simplex $\mathcal{T}_1$

In case of simplex $\mathcal{T}_1 \equiv \overline{\mathcal{T}}$ (Fig. 1) there are four half-turns pairing the faces with themselves

$$r_0 : \begin{bmatrix} A_1A_2A_3 \\ A_1A_3A_2 \end{bmatrix}; \quad r_1 : \begin{bmatrix} A_0A_2A_3 \\ A_0A_3A_2 \end{bmatrix}; \quad r_2 : \begin{bmatrix} A_0A_1A_3 \\ A_0A_3A_1 \end{bmatrix}; \quad r_3 : \begin{bmatrix} A_1A_0A_2 \\ A_1A_2A_0 \end{bmatrix};$$

These half-turns induce subdivision of edges each into two oriented segments. There are two classes, denoted by a and b, of the edge segments and so two cycle relations obtained by algorithm (4). The relators are

$$a \longrightarrow (r_0r_1)^a \quad (a \geq 2)$$
$$b \longrightarrow (r_0r_3r_2r_1r_3r_1r_2r_3r_0r_2)^b \quad (b \geq 1)$$

Condition c) is fulfilled for $\overline{a} = 2a$, $\overline{b} = 2b$ if in addition the angular conditions (3) are satisfied. Then the isometry group is

$$G(\overline{\mathcal{T}}, a, b) = (r_0, r_1, r_2, r_3 - r_0^2 = r_1^2 = r_2^2 = r_3^2 = (r_0r_1)^a$$
$$= (r_0r_3r_2r_1r_3r_1r_2r_3r_0r_2)^b = 1; \quad a \geq 2,\ b \geq 1).$$

All the vertices of the simplex $\overline{\mathcal{T}}$ are in the same class. Considering vertex figures on a 2-dimensional surfaces around the vertices we can obtain a fundamental domain, e.g., for the stabilizer group $\overline{G}_{A_3}$. Transformation r_2 is mapping A_3 onto A_1 and $\overline{\mathcal{T}}_{A_3}$ onto $\overline{\mathcal{T}}_{A_1}^{r_2}$. That means the vertex figures $\overline{\mathcal{T}}_{A_3}$ a $\overline{\mathcal{T}}^{r_2}_{A_1}$ have joint edge corresponding to the joint face f_{r_2} of the simplices $\overline{\mathcal{T}}$ and $\overline{\mathcal{T}}^{r_2}$. Similarly, vertex figures $\overline{\mathcal{T}}_{A_3}$ and $\overline{\mathcal{T}}^{r_1}_{A_2}$ have joint edge corresponding to f_{r_1} and $\overline{\mathcal{T}}^{r_1}_{A_2}$ and $\overline{\mathcal{T}}^{r_3r_1}_{A_0}$ to $f^*_{r_3} := (f_{r_3})^{r_1}$. One of the domains for $\overline{G}_{A_3}$ (Fig. 1) is

$$\overline{\mathcal{T}}^{r_3r_1}_{A_0} \cup \overline{\mathcal{T}}^{r_1}_{A_2} \cup \overline{\mathcal{T}}_{A_3} \cup \overline{\mathcal{T}}^{r_2}_{A_1} := \overline{\mathcal{P}}_{A_3}.$$

In the diagram for $\overline{\mathcal{P}}_{A_3}$ the minus sign in notations a^- and b^- means that segments in these two classes are directed to vertices according to vertex figure (plus means opposite direction).

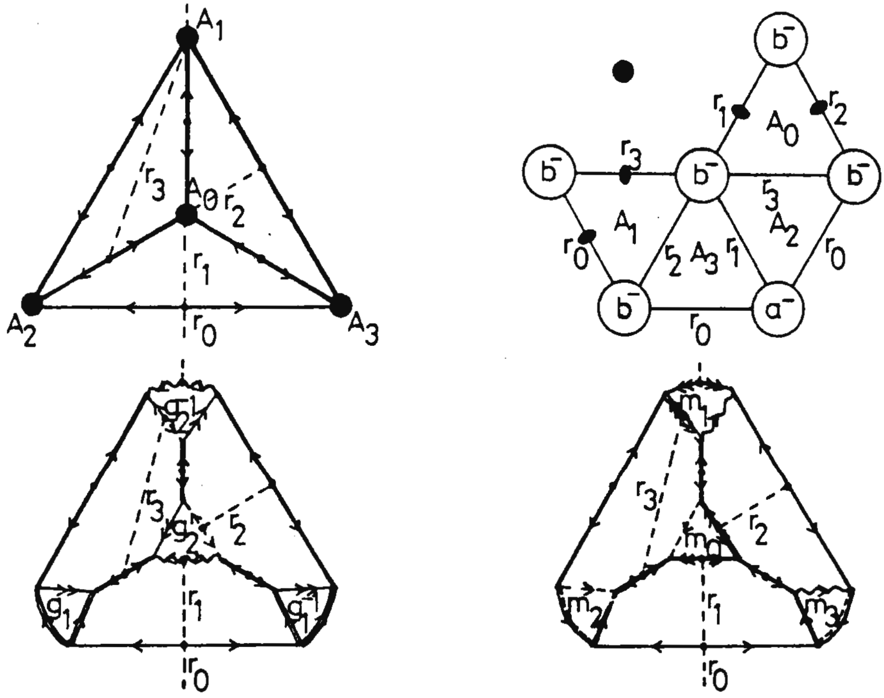

Fig. 1

The generators for $\overline{G}_{A_3}$, obtained from $\overline{\mathcal{P}}_{A_3}$, are

$$r_2r_0r_2 : (f_{r_0})^{r_2} \to (f_{r_0})^{r_2} \quad r_2r_3r_2 : (f^{r_2}_{r_3} \to (f_{r_3})^{r_2};$$
$$r_0r_1 : f_{r_0} \to (f_{r_0})^{r_1}; \quad (r_1r_3)r_1(r_3r_1) : (f_{r_1})^{r_3r_1} \to (f_{r_1})^{r_3r_1};$$
$$(r_1r_3)r_2(r_3r_1) : (f_{r_2})^{r_3r_1} \to (f_{r_2})^{r_3r_1}.$$

Discussion in Section 2 shows that simplex $\overline{\mathcal{T}}$ is hyperbolic with vertices out of the absolute for all values of parameters a, b. So if we truncate the simplex by the polar planes of the vertices we get a compact polyhedron $\overline{\mathcal{O}}$. If we equip it with additional face pairing isometries, it will be a fundamental domain for the group $G(\overline{\mathcal{O}}, a, b)$ which will be a supergroup for $G(\overline{\mathcal{T}}, a, b)$. There are more possibilities for face pairing.

1) The trivial one, which is always possible, is with plane reflection m_0, m_1, m_2, m_3 in polar planes of the vertices. That means the new triangular faces of $\overline{\mathcal{O}}^1$ are paired with themselves. The new group is

$$\begin{aligned} G(\overline{\mathcal{O}}^1, a, b) &= (r_0, r_1, r_2, r_3, m_0, m_1, m_2, m_3, -r_0^2 = r_1^2 = r_2^2 = r_3^2 = \\ &= m_0^2 = m_1^2 = m_2^2 = m_3^2 = (r_0 r_1)^a = (r_0 r_3 r_2 r_1 r_3 r_1 r_2 r_3 r_0 r_2)^b = \\ &= r_0 m_1 r_0 m_1 = r_1 m_0 r_1 m_0 = r_2 m_0 r_2 m_0 = r_3 m_1 r_3 m_1 = \\ &= r_0 m_3 r_0 m_2 = r_1 m_3 r_1 m_2 = r_2 m_0 r_2 m_3 = r_3 m_2 r_3 m_0 = 1; \quad a \geq 2,\ b \geq 1). \end{aligned}$$

2) New additional face pairings of $\overline{\mathcal{O}}^2$ have to satisfy the next criteria. Polar plane of A_3 and so group $\overline{G}_{A_3}$ will be invariant under these new transformations, fixing A_3, and exchanging half-spaces obtained by the polar plane. Thus, fundamental domain $\overline{\mathcal{P}}_{A_3}$ is divided into two parts, and the new stabilizer of the polar plane will be supergroup for $\overline{G}_{A_3}$, namely of index two. Symmetries of the $\overline{\mathcal{P}}_{A_3}$-tiling give us the idea how to indicate the new generators. If r is the new rotatory reflection mapping $\overline{\mathcal{T}}_{A_3}$ and $\overline{\mathcal{T}}_{A_1}^{r_2}$ onto vertex figures $\overline{\mathcal{T}}_{A_2}^{r_1}$ and $\overline{\mathcal{T}}_{A_0}^{r_3 r_1}$, but exchanging the half spaces, such that $rr = r_0 r_1$ is an old generator of $\overline{G}_{A_3}$, then the new generators for $G(\overline{\mathcal{O}}^2, a, b)$ will be the glide reflections

$$g_1 = r r_1, \qquad g_2 = r_2 r r_1 r_3$$

and the new relators are

——»——	$r_3 g_1^{-1} r_2 g_2$	∿∿»∿∿	$r_0 g_2 r_1 g_2^{-1}$
━━>━━	$r_1 g_1 r_0 g_1$	- - > - -	$r_3 g_2 r_2 g_2$

So we obtain the new group by presentation

$$\begin{aligned} G(\overline{\mathcal{O}}^2, a, b) &= (r_0, r_1, r_2, r_3, g_1, g_2 - r_0^2 = r_1^2 = r_2^2 = r_3^2 = (r_0 r_1)^a = \\ &= (r_0 r_3 r_2 r_1 r_3 r_1 r_2 r_3 r_0 r_2)^b = r_3 g_1^{-1} r_2 g_2 = r_1 g_1 r_0 g_1 = \\ &= r_0 g_2 r_1 g_2^{-1} = r_3 g_2 r_2 g_2^{-1} = 1; \qquad a \geq 2,\ b \geq 1). \end{aligned}$$

The $\overline{\mathcal{P}}_{A_3}$-tiling ini the polar plane of A_3 do not allow other identifications on the truncated simplex $\overline{\mathcal{O}}$, because of the Poincaré theorem, since angles at truncations being right ones.

Groups $G(\overline{\mathcal{T}},a,b)$, $G(\overline{\mathcal{O}}^i,a,b)$ $(i=1,2)$ are not maximal, i.e. they are subgroups of certain groups $G(\overline{\mathcal{S}},a,b)$ and $G(\overline{Q}^i,a,b)$, respectively, which leave invariant the tilings with simplices $\overline{\mathcal{T}}$, resp. truncated simplices $\overline{\mathcal{O}}^i$. Each of the domains $\overline{\mathcal{T}}$, $\overline{\mathcal{O}}^i$ can be divided into smaller parts, each with a face pairing to be fundamental domain of a larger group, i.e. supergroup of starting one. In our cases it is possible to halve $\overline{\mathcal{T}}$ or $\overline{\mathcal{O}}^i$ with plane through edge A_0A_1 and midpoint M of A_2A_3. Equipping the new simplex $\overline{\mathcal{S}}$, resp. the truncated simplex $\overline{Q}$ with the face pairing, as indicated in Fig. 2 and

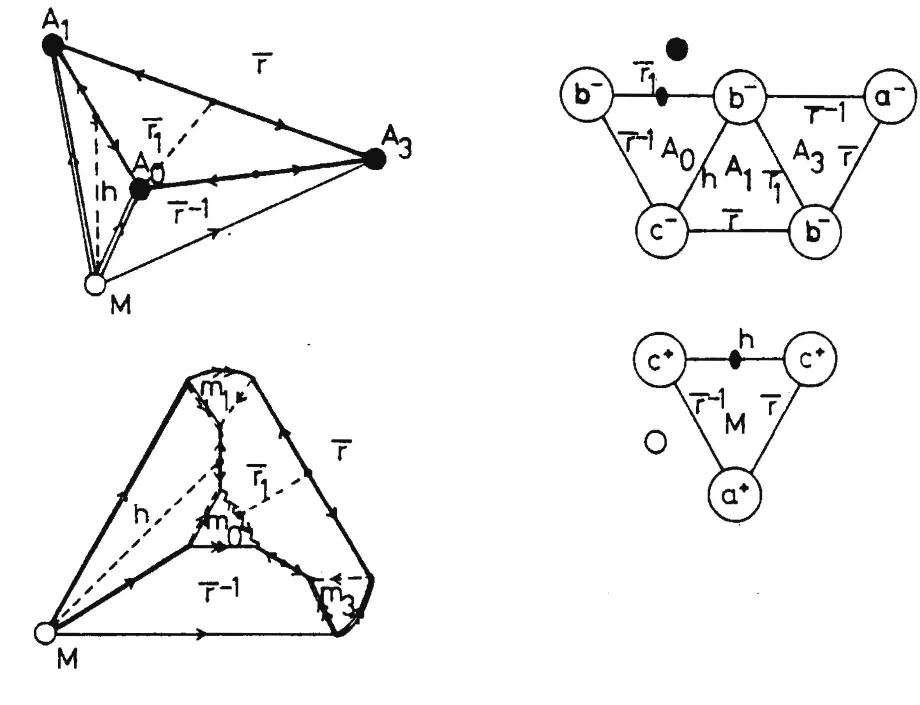

Fig. 2

Table 1 ($\overline{\mathcal{S}}=\mathcal{S}_1$, $\overline{Q}=Q_1$) we obtain the groups

$$G(\overline{\mathcal{S}},a,b)=(\bar r,h,\bar r_1 - h^2=\bar r_1^2=\bar r^{2a}=(h\bar r_1\bar r\bar r_1\bar r^{-1}\bar r_1)^{2b}=(\bar r h)^2=1;\quad a\geq 2,\ b\geq 1)$$

and

$$G(\overline{Q},a,b)=(\bar r,h,\bar r_1,m_0,m_1,m_3 - h^2=\bar r_1^2=m_0^2=m_1^2=m_3^2=\bar r^{2a}=$$
$$=(h\bar r_1\bar r\bar r_1\bar r^{-1}\bar r_1)^{2b}=(\bar r h)^2=(\bar r_1 m_0)^2=\bar r m_1\bar r^{-1}m_0=$$
$$=hm_1hm_0=\bar r_1 m_1\bar r_1 m_3=\bar r m_3\bar r^{-1}m_3=1;\quad a\geq 2,\ b\geq 1).$$

Half-turns for $\overline{\mathcal{T}}$ and $\overline{\mathcal{O}}^i$ expressed by the generators $\overline{r}$, h, $\overline{r}_1$ are

$$r_0 = \overline{r}^{-1}h, \qquad r_1 = \overline{r}h, \qquad r_2 = \overline{r}_1, \qquad r_3 = h\overline{r}_1 h$$

and the isometries, pairing faces of $\overline{\mathcal{O}}^2$, are $g_1 = m_3 h$, $g_2 = \overline{r}_1 m_3 \overline{r}_1 h$.

So, $G(\overline{\mathcal{S}}, a, b)$ os supergroup of $G(\overline{\mathcal{T}}, a, b)$ and $G(\overline{Q}, a, b)$ is supergroup of groups $G(\overline{\mathcal{O}}^i, a, b)$. The groups $G(\overline{\mathcal{S}}, a, b)$ and $G(\overline{Q}, a, b)$ have already been maximal ($a \neq 5b$), because it is not possible to find supergroups of them, leaving invariant the $\overline{\mathcal{T}}$-tiling, resp. the $\overline{\mathcal{O}}$-tiling.

If $a = 5b$ then simplex $\overline{\mathcal{T}}$ and truncated simplices $\overline{\mathcal{O}}^i$ may have more symmetries and their maximal groups have as fundamental domains simplex $\mathcal{S}$ and truncated simplex Q, respectively (Table 1). Such a smaller simplex $\mathcal{S}$ would be a characteristic simplex of a regular $\overline{\mathcal{T}}$ and $\mathcal{S}$ would have only one vertex out of absolute and the other vertices would be proper. Faces of $\mathcal{S}$ a Q would be identified by plane reflections with themselves (for similar cases see [6], Family 1).

5. The isometry groups for simplex $\mathcal{T}_5$

Rotatory reflection z and rotation r

$$z: \begin{bmatrix} A_0A_2A_1 \\ A_0A_1A_3 \end{bmatrix}; \qquad r: \begin{bmatrix} A_1A_2A_3 \\ A_0A_2A_3 \end{bmatrix};$$

are pairing the faces of the simplex $\overline{\mathcal{T}} \equiv \mathcal{T}_5$ (Fig. 3). There are two classes, a and b of edges in this simplex. To class a belongs only the edge A_2A_3. All other edges are in the class b, moreover, the face pairing isometries divide those edges into two equivalent oriented segments. Condition c) is fulfilled and if angular conditions (3) hold for $\overline{a} = a$ and $\overline{b} = 2b$ then the isometry group, with relations obtained by algorithm (4), is

$$G(\overline{\mathcal{T}}, a, b) = (r, z - (r)^a = (z^2 r^{-1} z^{-1} r)^{2b} = 1; \qquad a \geq 3,\ b \geq 1).$$

As for the simplex in Section 4, we consider vertex figures on a 2-dimensional surfaces around the equivalent vertices and glue a fundamental domain, e.g. for the stabilizer group $\overline{G}_{A_1}$ of A_1. One of the domains for $\overline{G}_{A_1}$ is

$$\overline{\mathcal{T}}_{A_0}^{r^{-1}} \cup \mathcal{T}_{A_1} \cup \overline{\mathcal{T}}_{A_2}^{z} \cup \mathcal{T}_{A_3}^{z^{-1}} =: \overline{\mathcal{P}}_{A_1}$$

where $\overline{\mathcal{T}}_{A_0}^{r^{-1}}$, $\overline{\mathcal{T}}_{A_1}$, $\overline{\mathcal{T}}_{A_2}^{z}$, $\overline{\mathcal{T}}_{A_3}^{z^{-1}}$ are vertex figures.

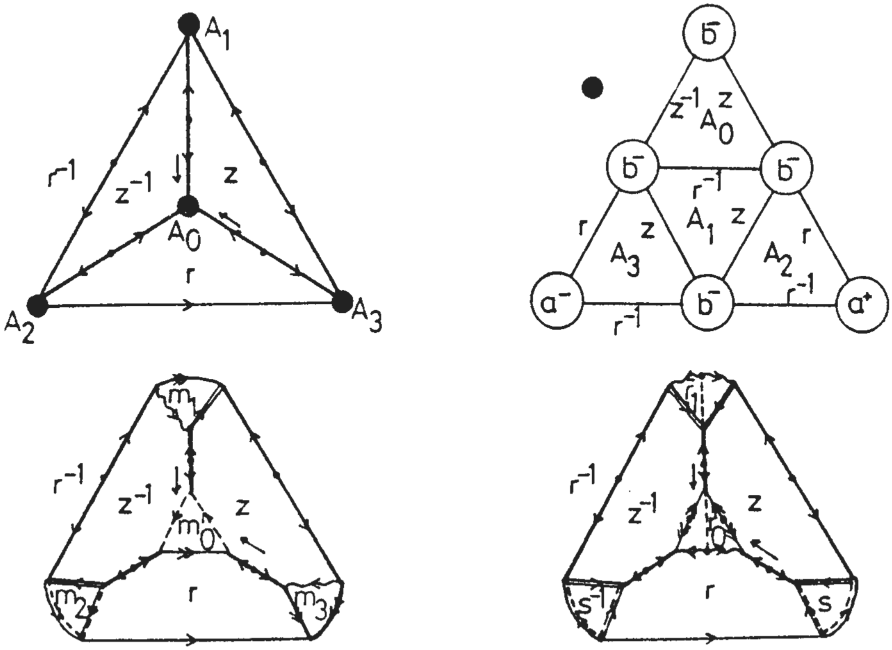

Fig. 3

In diagram for $\overline{\mathcal{P}}_{A_1}$ edges are noted with a^-, a^+ and b^- as before. The generators for $\overline{G}_{A_1}$, obtained from $\overline{\mathcal{P}}_{A_1}$, are

$$z^{-1}rz : (f_{r^{-1}})^z \to (f_r)^z; \quad zrz^{-1} : (f_{r^{-1}})^{z^{-1}} \to (f_r)^{z^{-1}};$$
$$zrz^{-1} : (f_{z^{-1}})^{r^{-1}} \to (f_z)^{r^{-1}}.$$

By the discussions in Section 2 simplex $\overline{\mathcal{T}}$ is hyperbolic with vertices out of the absolute for all parameters a, b. So, if we truncate the simplex by the polar planes of the vertices we get a compact polyhedron $\overline{\mathcal{O}} = \mathcal{O}_5$, which serves as a fundamental domain for some supergroup $G(\overline{\mathcal{O}}, a, b)$ of the above group $G(\overline{\mathcal{T}}, a, b)$. Besides the trivial case

$$\begin{aligned} G(\overline{\mathcal{O}}^1, a, b) &= (r, z, m_0, m_1, m_2, m_3 = m_0^2 = m_1^2 = m_2^2 = m_3^2 = r^a = \\ &= (z^2r^{-1}z^{-1}r)^{2b} = m_0zm_0z^{-1} = m_2zm_1z^{-1} = m_1zm_3z^{-1} = \\ &= m_1rm_0r^{-1} = m_2rm_2r^{-1} = m_3rm_3r^{-1} = 1, \quad a \geq 3,\ b \geq 1), \end{aligned}$$

there is only one more possibility.

Half-turn r_1 maps vertex figures $\overline{\mathcal{T}}^{r^{-1}}_{A_0}$ and $\overline{\mathcal{T}}_{A_1}$ onto themselves and vertex figure $\overline{\mathcal{T}}^{z}_{A_2}$ onto $\overline{\mathcal{T}}^{z^{-1}}_{A_3}$ but it exchanges the half spaces. The new generators for $G(\overline{\mathcal{O}}^2, a, b)$ are r_1, $r_0 = r^{-1}r_1r$ and $s = zrz$, so the new group is

$$G(\overline{\mathcal{O}}^2, a, b) = (r, z, r_0, r_1, s - r_0^2 = r_1^2 = r^a = (z^2r^{-1}z^{-1}r)^{2b} = \\ = (zr_0)^2 = r^{-1}r_1rr_0 = zs^{-1}zr_1 = rsrs^{-1} = 1; \quad a \geq 3, \ b \geq 1).$$

It is possible to divide domains $\overline{\mathcal{T}}$ and $\overline{\mathcal{O}}^i$ $(i = 1, 2)$ of groups $G(\overline{\mathcal{T}}, a, b)$ a $G(\overline{\mathcal{O}}^i, a, b)$ into smaller parts, each with a face pairings, to be fundamental domains of their supergroups $G(\overline{\mathcal{S}}, a, b)$ and $G(\overline{Q}, a, b)$. It may be done with the plane through edge A_0A_1 and midpoint M of A_2A_3 (Fig. 4). Face

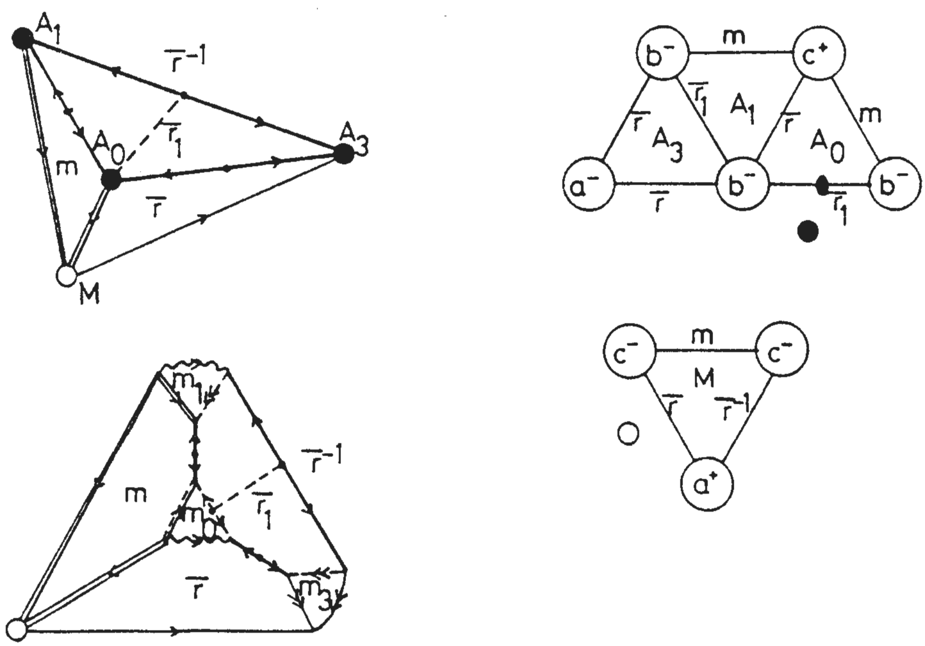

Fig. 4

pairings of domains $\overline{\mathcal{S}} = \mathcal{S}_2$ and $\overline{Q} = Q_2$ $(a \neq 10b)$ give maximal groups

$$G(\overline{\mathcal{S}}, a, b) = (\bar{r}, \bar{r}_1, m - m^2 = \bar{r}_1^2 = \bar{r}^a = (\bar{r}_1\bar{r}^{-1}\bar{r}_1\bar{r}\bar{r}_1m\bar{r}_1\bar{r}^{-1}\bar{r}_1\bar{r}\bar{r}_1m)^b = \\ = \bar{r}m\bar{r}^{-1}m = 1, \quad a \geq 3, \ b \geq 1)$$

and

$$G(\overline{Q},a,b)=(\overline{r},\overline{r}_1,m,m_0,m_1,m_3-m^2=\overline{r}_1^2=m_0^2=m_1^2=m_3^2=\overline{r}^a=$$
$$=(\overline{r}_1\overline{r}^{-1}\overline{r}_1\overline{r}\overline{r}_1m\overline{r}_1\overline{r}^{-1}\overline{r}_1\overline{r}\overline{r}_1m)^b=\overline{r}m\overline{r}^{-1}m=m_0\overline{r}_1m_0\overline{r}_1=m_1\overline{r}_1m_3\overline{r}_1=$$
$$=m_3\overline{r}m_3\overline{r}^{-1}=m_1\overline{r}m_0\overline{r}^{-1}=(m_1m)^2=(m_0m)^2=1;\quad a\geq 3,\ b\geq 1).$$

Generators of groups $\overline{\mathcal{T}}$ and $\overline{O}^2$ expressed by the generators of groups $\overline{\mathcal{S}}$ and $\overline{Q}$ are

$$z=m\overline{r}_1,\qquad r=\overline{r},\qquad r_0=mm_0,\qquad r_1=mm_1,\qquad s=mm_3.$$

If $a=10b$ maximal supergroups are $G(\mathcal{S},a,b)$ and $G(Q,a,b)$ (Table 1).

6. The isometry groups for simplices $\mathcal{T}_1-\mathcal{T}_{11}$

For simplices $\mathcal{T}_1-\mathcal{T}_{11}$ the face pairings are given in the first rows of Table 2, respectively. In the second row parameters $\overline{a}$, $\overline{b}$ are expressed by natural parameters a, b. Then presentation of the group with one of these simplices or truncation of them as a fundamental domain follows, together with notations given in [6] for $Aut\,\mathcal{T}_i$ (e.g. $\frac{2}{2}\Gamma_5$) and numberings of families including $\mathcal{T}_i$ (e.g. F.11). Finally, we give the generators of $G(\mathcal{T}_i,a,b)$ and $G(\mathcal{O}_i^j,a,b)$ ($i\in\{1,\dots,11\}$, $j=1,2$) by generators of these supergroups $G(\mathcal{S}_k,\overline{a},\overline{b})$ and $G(Q_k,\overline{a},\overline{b})$, respectively, if they exist.

In Table 1 there are summarized the presentations of possible supergroups of groups $G(\mathcal{T}_i,a,b)$ and $G(\mathcal{O}_i,a,b)$ and face pairings of their fundamental domains.

In both tables m_i means the plane reflection in the truncating plane of the vertex A_i.

Summarizing results for simplices $\mathcal{T}_1-\mathcal{T}_{11}$ are given in

THEOREM 2. a) *If $b=1$, then*

— *simplices $\mathcal{T}_2$ for $a=2$ a $\mathcal{T}_4$ for $a=3,4,5$ are spherical;*

— *simplices $\mathcal{T}_2$ for $a=3$ and $\mathcal{T}_4$ for $a=6$ are realizable in Nil-space;*

— *simplices $\mathcal{T}_2$ for $a\geq 4$ and $\mathcal{T}_4$ for $a\geq 7$ are realizable in* SL(ℝ)*-space;*

Simplices $\mathcal{T}_2$ and $\mathcal{T}_4$ for $b\geq 2$ and simplices $\mathcal{T}_1$, $\mathcal{T}_3$, $\mathcal{T}_5$, $\mathcal{T}_6$, $\mathcal{T}_7$, $\mathcal{T}_8$, $\mathcal{T}_9$, $\mathcal{T}_{10}$, $\mathcal{T}_{11}$ for all values of parameters a, b are hyperbolic with vertices out of the absolute.

b) *Groups $G(\mathcal{T}_i,a,b)$ ($i=1,2,3,4$) are subgroups of group $G(\mathcal{S}_1,\overline{a},\overline{b})$ and groups $G(\mathcal{O}_i^j,a,b)$ ($i=1,3,4;\ j=1,2$), $G(\mathcal{O}_2,a,b)$ are subgroups of group $G(Q_1,\overline{a},\overline{b})$ (with $\overline{a}$, $\overline{b}$ expressed by a, b as it is indicated in Table 2).*

c) *Groups* $G(\mathcal{I}_i, a, b)$ $(i = 5,6,7,8)$ *are subgroups of group* $G(\mathcal{S}_2, \bar{a}, \bar{b})$, *and groups* $G(\mathcal{O}_i^j, a, b)$ $(i = 5,6,7;\ j = 1,2)$, $G(\mathcal{O}_8, a, b)$ *are subgroups of group* $G(Q_2, \bar{a}, \bar{b})$.

d) *Groups* $G(\mathcal{I}_9, a, b)$, $G(\mathcal{O}_9, a, b)$, $G(\mathcal{I}_{10}, a, b)$, $G(\mathcal{O}_{10}, a, b)$, $G(\mathcal{I}_{11}, a, b)$, $G(\mathcal{O}_{11}, a, b)$ $(\bar{a} \neq 5\bar{b})$ *are maximal groups. Supergroups* $G(\mathcal{S}_1, \bar{a}, \bar{b})$, $G(Q_1, \bar{a}, \bar{b})$, $G(\mathcal{S}_2, \bar{a}, \bar{b})$, $G(Q_2, \bar{a}, \bar{b})$ $(\bar{a} \neq 5\bar{b})$ *are also maximal groups. For other parameters* $\bar{a} = 5\bar{b}$ *of these groups, their maximal supergroups are* $G(\mathcal{S}, \bar{a} = 5\bar{b})$ *for groups* $G(\mathcal{I}_i, a, b)$ $(i = 1, \ldots, 11)$, $G(\mathcal{S}_j, \bar{a}, \bar{b})$ $(j = 1,2)$, *moreover,* $G(Q, \bar{a} = 5\bar{b})$ *are maximal for groups* $G(\mathcal{O}_i, a; b)$ $G(Q_j, \bar{a}, \bar{b})$.

References

[1] E. MOLNÁR, Two hyperbolic football manifolds, *Differential Geometry and Applications*, Proceedings of the Conference, June 26–July 3, 1988, Dubrovnik, Yugoslavia.

[2] E. MOLNÁR, Projective metrics and hyperbolic volume, *Annales Univ. Sci. Budapest.*, Sect. Math., **32** (1989), 127–157.

[3] E. MOLNÁR, Eine Klasse von hyperbolischen Raumgruppen, *Beiträge zur Algebra und Geometrie*, **30** (1990), 79–100.

[4] E. MOLNÁR, Polyhedron complexes with simply transitive group actions and their realizations, *Acta Math. Hung.*, **59** (1992), 175–216.

[5] E. MOLNÁR and I. PROK, A polyhedron algorithm for finding space groups. *Proceedings of Third Int. Conf. on Engineering Graphics and Descriptive Geometry*, Vienna 1988, Vol. 2, 37–44.

[6] E. MOLNÁR and I. PROK, Classifications of solid transitive simplex tilings in 3-dimensional homogeneous simply connected Riemann spaces, I, Preprint 92–003, Universität Bielefeld, SFB 343 "Diskrete Strukturen in der Mathematik", In: Colloquia Math. Soc. J. Bolyai 63. Intuitive Geometry, Szeged (Hungary) 1991, North-Holland Co. Amsterdam–Oxford–New York (to appear).

[7] И. К. ЖУК, *Правильные разбиения пространств постоянной кривизны на конгруэнтные тетраэдры I-II.*, Академия наук Белорусской ССР, Институт Математики, Минск 1980, 1983, Препринт.

[8] I. K. ZHUK, Fundamental tetrahedra in Euclidean and Lobachevsky spaces, *Soviet Math. Dokl.*, **270** (183), 540–543.

Table 1

$\mathscr{S}_1$ Fig. 2	$\bar{r} := \begin{bmatrix} A_0MA_3 \\ A_1MA_3 \end{bmatrix}, \quad h: \begin{bmatrix} A_0MA_1 \\ A_1MA_0 \end{bmatrix}, \quad \bar{r}_1: \begin{bmatrix} A_1A_0A_3 \\ A_3A_0A_1 \end{bmatrix}$
${}^{2}_{2}\Gamma_5(\bar{a}, 5\bar{b})$ F.11	$G(\mathscr{S}_1, \bar{a}, \bar{b}) = (\bar{r}, h, \bar{r}_1 - h^2 = \bar{r}_1^2 = (\bar{r})^{\bar{a}} = (h\bar{r}_1\bar{r}\bar{r}_1\bar{r}^{-1}\bar{r}_1)^{\bar{b}} = (\bar{r}h)^2 = 1,$ $\bar{a} \geq 3, \bar{b} \geq 1)$ $G(Q_1, \bar{a}, \bar{b}) = (\bar{r}, h, \bar{r}_1, m_0, m_1, m_3 - h^2 = \bar{r}_1^2 = m_0^2 = m_1^2 = m_3^2 = \bar{r}^{\bar{a}} =$ $= (h\bar{r}_1\bar{r}\bar{r}_1\bar{r}^{-1}\bar{r}_1)^{\bar{b}} = (\bar{r}h)^2 = (\bar{r}_1m_0)^2 = \bar{r}m_1\bar{r}^{-1}m_0 = hm_1hm_0 =$ $= \bar{r}_1m_1\bar{r}_1m_3 = \bar{r}m_3\bar{r}^{-1}m_3 = 1; \quad a \geq 2, b \geq 1)$
$\mathscr{S}_2$ Fig. 4	$\bar{r}: \begin{bmatrix} A_1MA_3 \\ A_0MA_3 \end{bmatrix}, \quad m: \begin{bmatrix} A_0MA_1 \\ A_0MA_1 \end{bmatrix}, \quad \bar{r}: \begin{bmatrix} A_1A_0A_3 \\ A_3A_0A_1 \end{bmatrix}$
${}^{m}_{2}\Gamma_4(\bar{a}, 5\bar{b})$ F.7	$G(\mathscr{S}_2, \bar{a}, \bar{b}) = (\bar{r}, \bar{r}_1, m - m^2 = \bar{r}_1^2 = \bar{r}^{\bar{a}} = (\bar{r}_1\bar{r}^{-1}\bar{r}_1\bar{r}\bar{r}_1m)^{\bar{b}} = \bar{r}m\bar{r}^{-1}m = 1,$ $\bar{a} \geq 3, \bar{b} \geq 1)$ $G(Q_2, \bar{a}, \bar{b}) = (\bar{r}, \bar{r}_1, m, m_0, m_1, m_3 - m^2 = \bar{r}_1^2 = m_0^2 = m_1^2 = m_3^2 = \bar{r}^{\bar{a}} =$ $= (\bar{r}_1\bar{r}^{-1}\bar{r}_1\bar{r}\bar{r}_1m)^{\bar{b}} = \bar{r}m\bar{r}^{-1}m = m_0\bar{r}_1m_0\bar{r}_1 = m_1\bar{r}_1m_3\bar{r}_1 =$ $= m_3\bar{r}m_3\bar{r}^{-1} = m_1\bar{r}m_0\bar{r}^{-1} = (m_1m)^2 = (m_0m)^2 = 1; \quad \bar{a} \geq 3, \bar{b} \geq 1)$
$\mathscr{S}$	$\overline{m}_0: \begin{bmatrix} \overline{A}_1\overline{A}_2\overline{A}_3 \\ \overline{A}_1\overline{A}_2\overline{A}_3 \end{bmatrix}, \overline{m}_1: \begin{bmatrix} \overline{A}_0\overline{A}_2\overline{A}_3 \\ \overline{A}_0\overline{A}_2\overline{A}_3 \end{bmatrix}, \overline{m}_2: \begin{bmatrix} \overline{A}_0\overline{A}_1\overline{A}_3 \\ \overline{A}_0\overline{A}_1\overline{A}_3 \end{bmatrix}, \overline{m}_3: \begin{bmatrix} \overline{A}_0\overline{A}_1\overline{A}_2 \\ \overline{A}_0\overline{A}_1\overline{A}_2 \end{bmatrix}$
${}^{\bar{4}3m}_{24}\Gamma(\bar{a} = 5\bar{b})$ F.1	$G(\mathscr{S}, \bar{a}) = (\overline{m}_0, \overline{m}_1, \overline{m}_2, \overline{m}_3 - \overline{m}_0^2 = \overline{m}_1^2 = \overline{m}_2^2 = \overline{m}_3^2 = (\overline{m}_0\overline{m}_1)^3 =$ $= (\overline{m}_0\overline{m}_2)^2 = (\overline{m}_0\overline{m}_3)^2 = (\overline{m}_1\overline{m}_2)^3 = (\overline{m}_1\overline{m}_3)^2 = (\overline{m}_2\overline{m}_3)^{\bar{a}} = 1, \quad \bar{a} \geq 3$ $G(Q, \bar{a}) = (\overline{m}_0, \overline{m}_1, \overline{m}_2, \overline{m}_3, m_0 - \overline{m}_0^2 = \overline{m}_1^2 = \overline{m}_2^2 = \overline{m}_3^2 = m_0^2 =$ $= (\overline{m}_0\overline{m}_1)^3 = (\overline{m}_0\overline{m}_2)^2 = (\overline{m}_0\overline{m}_3)^2 = (\overline{m}_1\overline{m}_2)^3 = (\overline{m}_1\overline{m}_3)^2 =$ $= (\overline{m}_2\overline{m}_3)^{\bar{a}} = (\overline{m}_1m_0)^2 = (\overline{m}_2m_0)^2 = (\overline{m}_3m_0)^2 = 1, \quad \bar{a} \geq 3)$

Table 2

$\mathcal{J}_1$ Fig. 1	$r_0: \begin{bmatrix} A_1A_2A_3 \\ A_1A_3A_2 \end{bmatrix}$, $r_1: \begin{bmatrix} A_0A_2A_3 \\ A_0A_3A_2 \end{bmatrix}$, $r_2: \begin{bmatrix} A_0A_1A_3 \\ A_0A_3A_1 \end{bmatrix}$, $r_3: \begin{bmatrix} A_1A_0A_2 \\ A_1A_2A_0 \end{bmatrix}$
	$\bar{a} = 2a, \quad \bar{b} = 2b$
$\Gamma_{14}(2a, 10b)$ F.11	$G(\mathcal{J}_1, a, b) = (r_0, r_1, r_2, r_3 - r_0^2 = r_1^2 = r_2^2 = r_3^2 = (r_0r_1)^a =$ $= (r_0r_3r_2r_1r_3r_1r_2r_3r_0r_2)^b = 1; \quad a \geq 2,\ b \geq 1)$ $G(\mathcal{O}_1^1, a, b) = (r_0, r_1, r_2, r_3, m_0, m_1, m_2, m_3, -r_0^2 = r_1^2 = r_2^2 = r_3^2 = m_0^2 =$ $= m_1^2 = m_2^2 = m_3^2 = (r_0r_1)^a = (r_0r_3r_2r_1r_3r_1r_2r_3r_0r_2)^b = r_0m_1r_0m_1 =$ $= r_1m_0r_1m_0 = r_2m_0r_2m_0 = r_3m_1r_3m_1 = r_0m_3r_0m_2 = r_1m_3r_1m_2 =$ $= r_2m_1r_2m_3 = r_3m_2r_3m_0 = 1; \quad a \geq 2,\ b \geq 1)$ $G(\mathcal{O}_1^2, a, b) = (r_0, r_1, r_2, r_3, g_1, g_2 - r_0^2 = r_1^2 = r_2^2 = r_3^2 = (r_0r_1)^a =$ $= (r_0r_3r_2r_1r_3r_1r_2r_3r_0r_2)^b = r_3g_1^{-1}r_2g_2 = r_1g_1r_0g_1 = r_0g_2r_1g_2^{-1} =$ $= r_3g_2r_2g_2^{-1} = 1; \quad a \geq 2,\ b \geq 1)$
Fig. 2	$r_0 = \bar{r}^{-1}h, \quad r_1 = \bar{r}h, \quad r_2 = \bar{r}_1, \quad r_3 = h\bar{r}_1h;$ $g_1 = m_3h, \quad g_2 = m_1h$
$\mathcal{J}_2$	$r_1: \begin{bmatrix} A_1A_2A_3 \\ A_1A_3A_2 \end{bmatrix}$, $r_2: \begin{bmatrix} A_0A_2A_3 \\ A_0A_3A_2 \end{bmatrix}$, $s: \begin{bmatrix} A_0A_2A_1 \\ A_3A_1A_0 \end{bmatrix}$
	$\bar{a} = 2a, \quad \bar{b} = b$
$\Gamma_{36}(2a, 5b)$ F.11	$G(\mathcal{J}_2, a, b) = (r_1, r_2, s - r_1^2 = r_2^2 = (r_1r_2)^a = (s^2r_2sr_1)^b; \quad a \geq 2,\ b \geq 1)$ $G(\mathcal{O}_2, a, b) = (r_1, r_2, s, m_0, m_1, m_2, m_3 - r_1^2 = r_2^2 = m_0^2 = m_1^2 = m_2^2 =$ $= m_3^2 = (r_1r_2)^a = (s^2r_2sr_1)^b = (m_0r_2)^2 = (m_1r_1)^2 = m_0sm_3s^{-1} =$ $= m_1sm_0s^{-1} = m_2sm_1s^{-1} = m_2r_2m_3r_2 = m_2r_1m_3r_1 = 1;\ a \geq 2,\ b \geq 1)$
Fig. 2	$r_1 = \bar{r}^{-1}h, \quad r_2 = \bar{r}h, \quad s = h\bar{r}_1$
$\mathcal{J}_3$	$r_1: \begin{bmatrix} A_0A_1A_3 \\ A_0A_3A_1 \end{bmatrix}$, $r_2: \begin{bmatrix} A_1A_0A_2 \\ A_1A_2A_0 \end{bmatrix}$, $r: \begin{bmatrix} A_1A_2A_3 \\ A_0A_2A_3 \end{bmatrix}$
	$\bar{a} = a, \quad \bar{b} = 2b$
$\Gamma_{45}(a, 10b)$ F.11	$G(\mathcal{J}_3, a, b) = (r_1, r_2, r - r_1^2 = r_2^2 = (r)^a = (r_2rr_2r^{-1}r_2r_1r^{-1}r_1rr_1)^b =$ $= 1; \quad a \geq 3,\ b \geq 1)$ $G(\mathcal{O}_3^1, a, b) = (r_1, r_2, r, m_0, m_1, m_2, m_3 - r_1^2 = r_2^2 = (r)^a = m_0^2 = m_1^2 =$ $= m_2^2 = m_3^2 = (r_2rr_2r^{-1}r_2r_1r^{-1}r_1rr_1)^b = m_0r_1m_0r_1 =$ $= m_1r_2m_1r_2 = m_1r_1m_3r_1 = m_0r_2m_2r_2 = m_1rm_0r^{-1} = m_2rm_2r^{-1} =$ $= m_3rm_3r^{-1} = 1; \quad a \geq 3,\ b \geq 1)$ $G(\mathcal{O}_3^2, a, b) = (r_1, r_2, r, z_1, z_2 - r_1^2 = r_2^2 = (r)^a = (r_2rr_2r^{-1}r_2r_1r^{-1}r_1rr_1)^b =$ $r_1z_1^{-1}r_2z_1 = r_2z_2r_1z_1 = rz_1^{-1}rz_1^{-1} = rz_2rz_2^{-1} = 1; \quad a \geq 3,\ b \geq 1)$
Fig. 2	$r_1 = \bar{r}_1,\ r_2 = h\bar{r}_1h,\ r = \bar{r}^{-1}; \quad z_1 = hm_0,\ z_2 = hm_3$

$\mathcal{T}_4$	$r: \begin{bmatrix}A_0A_2A_3\\A_1A_2A_3\end{bmatrix}, \quad s: \begin{bmatrix}A_1A_2A_0\\A_0A_1A_3\end{bmatrix}$
	$\bar{a}=a, \quad \bar{b}=b$
$\Gamma_{55}(a,5b)$ F.11	$G(\mathcal{T}_4,a,b)=(r,s-(r)^a=(s^2rs^{-1}r)^b=1; \quad a\geq 3,\ b\geq 1$ $G(\mathcal{O}_4^1,a,b)=(r,s,m_0,m_1,m_2,m_3-m_0^2=m_1^2=m_2^2=m_3^2=(r)^a=$ $=(s^2rs^{-1}r)^b=m_0sm_3s^{-1}=m_1sm_0s^{-1}=m_2sm_1s^{-1}=m_0rm_1r^{-1}=$ $=m_2rm_1r^{-1}=m_3rm_3r^{-1}=1; \quad a\geq 3,\ b\geq 1)$ $G(\mathcal{O}_4^2,a,b)=(r,s,g_1,g_2-(r)^a=(s^2rs^{-1}r)^b=g_1s^{-1}g_1s^{-1}=$ $=g_2s^{-1}g_1^{-1}s^{-1}=g_1rg_1r=g_2rg_2^{-1}r=1; \quad a\geq 3,\ b\geq 1)$
Fig. 2	$s=h\bar{r}_1,\ r=\bar{r}; \quad g_1=hm_0,\ g_2=hm_3$
$\mathcal{T}_5$ Fig. 3	$z: \begin{bmatrix}A_0A_2A_1\\A_0A_1A_3\end{bmatrix}, \quad r: \begin{bmatrix}A_1A_2A_3\\A_0A_2A_3\end{bmatrix}$
	$\bar{a}=a, \quad \bar{b}=2b$
$\Gamma_{56}(a,10b)$ F.7	$G(\mathcal{T}_5,a,b)=(r,z-(r)^a=(z^2r^{-1}z^{-1}r)^{2b}=1; \quad a\geq 3,\ b\geq 1$ $G(\mathcal{O}_5^1,a,b)=(r,z,m_0,m_1,m_2,m_3-m_0^2=m_1^2=m_2^2=m_3^2=r^a=$ $=(z^2r^{-1}z^{-1}r)^{2b}=m_0zm_0z^{-1}=m_2zm_1z^{-1}=m_1zm_3z^{-1}=$ $=m_1rm_0r^{-1}=m_2rm_2r^{-1}=m_3zm_3r^{-1}=1; \quad a\geq 3,\ b\geq 1)$ $G(\mathcal{O}_5^2,a,b)=(r,z,r_0,r_1,s-r_0^2=r_1^2=r^a=(z^2r^{-1}z^{-1}r)^{2b}=(zr_0)^2=$ $=r^{-1}r_1rr_0=zs^{-1}zr_1=rsr^{-1}s^{-1}=1; \quad a\geq 3,\ b\geq 1)$
Fig. 4	$z=m\bar{r}_1,\ r=\bar{r}; \quad r_0=mm_0,\ r_1=mm_1,\ s=mm_3$
$\mathcal{T}_6$	$r_1: \begin{bmatrix}A_0A_1A_3\\A_0A_3A_1\end{bmatrix}, \quad r_2: \begin{bmatrix}A_0A_1A_2\\A_0A_2A_1\end{bmatrix} \quad r: \begin{bmatrix}A_0A_2A_3\\A_1A_2A_3\end{bmatrix}$
	$\bar{a}=a, \quad \bar{b}=2b$
$\Gamma_{50}(a,10b)$ F.7	$G(\mathcal{T}_6,a,b=(r_1,r_2,r-r_1^2=r_2^2=(r)^a=(r_1rr_1r^{-1}r_1r_2rr_2r^{-1}r_2)^b=1;$ $a\geq 3,\ b\geq 1)$ $G(\mathcal{O}_6^1,a,b)=(r_1,r_2,r,m_0,m_1,m_2,m_3-r_1^2=r_2^2=(r)^a=m_0^2=m_1^2=$ $=m_2^2=m_3^2=(r_1rr_1r^{-1}r_1r_2rr_2r^{-1}r_2)^b=m_0r_1m_0r_1=m_0r_2m_0r_2=$ $=m_0rm_1r^{-1}=m_1r_2m_2r_2=m_1r_1m_3r_1=m_2rm_2r^{-1}=m_3rm_3r^{-1}=$ $=1; \quad a\geq 3,\ b\geq 1)$ $G(\mathcal{O}_6^2,a,b)=(r_1,r_2,r,r_3,r_4,s-r_1^2=r_2^2=r_3^2=r_4^2=(r)^a=$ $=(r_1rr_1r^{-1}r_1r_2rr_2r^{-1}r_2)^b=r^{-1}r_4rr_3=r_2sr_1r_3=r_2r_4r_1r_4=$ $=r^{-1}srs^{-1}=1; \quad a\geq 3,\ b\geq 1)$
Fig. 4	$r_1=\bar{r}_1,\ r_2=m\bar{r}_1m,\ r=\bar{r}^{-1}; \quad r_3=mm_1,\ r_4=mm_0,\ s=mm_3$

$\mathcal{J}_8$	$z_1:\begin{bmatrix}A_2A_1A_3\\A_3A_0A_2\end{bmatrix},\quad z_2:\begin{bmatrix}A_0A_3A_1\\A_0A_1A_2\end{bmatrix}$
	$\bar{a}=2a,\quad \bar{b}=2b$
$\Gamma_{63}(2a,10b)$ F.7	$G(\mathcal{J}_8,a,b)=(z_1,z_2-(z_1)^{2a}=(z_2^2z_1^{-1}z_2z_1)^{2b}=1;\quad a\geq 2,\ b\geq 1)$ $G(\mathcal{O}_8(a,b)=(z_1,z_2,m_0,m_1,m_2,m_3=(z_1)^{2a}=m_0^2=m_1^2=m_2^2=m_3^2==(z_2^2z_1^{-1}z_2z_1)^{2b}=m_0z_2m_0z_2^{-1}=m_3z_2m_1z_2^{-1}=m_1z_2m_2z_2^{-1}==m_2z_1m_3z_1^{-1}=m_1z_1m_0z_1^{-1}=m_3z_1m_2z_1^{-1}=1;\quad a\geq 2,\ b\geq 1)$
Fig. 4	$z_1=m\bar{r},\quad z_2=\bar{r}_1m$
$\mathcal{J}_9$	$r_0:\begin{bmatrix}A_1A_2A_3\\A_1A_3A_2\end{bmatrix},\ r_1:\begin{bmatrix}A_0A_2A_3\\A_0A_3A_2\end{bmatrix},\ r_2:\begin{bmatrix}A_3A_0A_1\\A_3A_1A_0\end{bmatrix},\ r_3:\begin{bmatrix}A_1A_0A_2\\A_1A_2A_0\end{bmatrix}$
	$\bar{a}=2a,\quad \bar{b}=2b$
$\Gamma_{15}(2a,10b)$ F.23	$G(\mathcal{J}_9,a,b)=(r_0,r_1,r_2,r_3-r_0^2=r_1^2=r_2^2=r_3^2=(r_0r_1)^a==(r_3r_1r_2r_0r_3r_2r_3r_0r_2r_1)^b=1;\quad a\geq 2,\ b\geq 1)$ $G(\mathcal{O}_9,a,b)=(r_0,r_1,r_2,r_3,m_0,m_1,m_2,m_3,-r_0^2=r_1^2=r_2^2=r_3^2=m_0^2==m_1^2=m_2^2=m_3^2=(r_0r_1)^a=(r_3r_1r_2r_0r_3r_2r_3r_0r_2r_1)^b=m_0r_1m_0r_1==m_1r_0m_1r_0=m_3r_2m_3r_2=m_1r_3m_1r_3=m_1r_2m_0r_2=m_0r_3m_2r_3==m_2r_1m_3r_1=m_2r_0m_3r_0=1;\quad a\geq 2,\ b\geq 1)$
$\mathcal{J}_{10}$	$r:\begin{bmatrix}A_0A_2A_3\\A_0A_3A_2\end{bmatrix},\quad s:\begin{bmatrix}A_0A_2A_1\\A_1A_0A_3\end{bmatrix},\quad m:\begin{bmatrix}A_1A_2A_3\\A_1A_2A_3\end{bmatrix}$
	$\bar{a}=4a,\quad \bar{b}=2b$
$\Gamma_{25}(4a,10b)$ F.25	$G(\mathcal{J}_{10},a,b)=(r,s,m-r^2=m^2=(rm)^{2a}=(s^2ms^{-2}rs^{-1}msr)^b==1;\quad a\geq 1,\ b\geq 1)$ $G(\mathcal{O}_{10},a,b)=(r,s,m,m_0,m_1,m_2,m_3-r^2=m^2=m_0^2=m_1^2=m_2^2==m_3^2=(rm)^{2a}=(s^2ms^{-2}rs^{-1}msr)^b=m_0rm_0r=m_0sm_1s^{-1}=m_1sm_3s^{-1}=m_2sm_0s^{-1}=m_2rm_3r=(m_1m)^2=(m_2m)^2=(m_3m)^2==1;\quad a\geq 1,\ b\geq 1)$
$\mathcal{J}_{11}$	$r_1:\begin{bmatrix}A_3A_0A_1\\A_3A_1A_0\end{bmatrix},\quad r_2:\begin{bmatrix}A_0A_1A_2\\A_0A_2A_1\end{bmatrix},\quad z:\begin{bmatrix}A_3A_0A_2\\A_2A_1A_3\end{bmatrix}$
	$\bar{a}=2a,\quad \bar{b}=2b$
$\Gamma_{44}(2a,10b)$ F.32	$G(\mathcal{J}_{11},a,b)=(r_1,r_2,z-r_1^2=r_2^2=(z)^{2a}=(zr_2z^{-1}r_1z^{-1}r_2r_1r_2zr_1)^b==1;\quad a\geq 2,\ b\geq 1)$ $G(\mathcal{O}_{11},a,b)=(r_1,r_2,z,m_0,m_1,m_2,m_3-r_1^2=r_2^2=(z)^{2a}=m_0^2=m_1^2==m_2^2=m_3^2=(zr_2z^{-1}r_1z^{-1}r_2r_1r_2zr_1)^b=m_0r_2m_0r_2=m_3r_1m_3r_1==m_0r_1m_1r_1=m_1r_2m_2r_2=m_3zm_2z^{-1}=m_0zm_1z^{-1}=m_2zm_3z^{-1}==1;\quad a\geq 2,\ b\geq 1)$

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 103–127

A BASIS THEOREM FOR ORDINARY DIFFERENTIAL OPERATOR OF n-TH ORDER

By

I. JOÓ, M. B. TAHIR and S. SZABÓ

Department of Analysis, Eötvös Loránd University, Budapest

(Received September 8, 1992)

The equiconvergence theorems play an important role in the theory of expansions. HORVÁTH, JOÓ and KOMORNIK proved a general theorem of this nature for the one dimensional Schrödinger operator in [1] and their result was extended for the operator $Lu := u^{(4)}$ in [4]. The aim of the present paper is to prove a similar theorem for the operator $Lu := u^{(n)} + q_2(x)u^{(n-2)} + \ldots + q_n(x)u$.

Let G be an open interval (finite or infinite) on the real line, n a natural number, $q_s \in L_1^{\text{loc}}(G)$ $(s = 2, \ldots, n)$ complex functions and consider the differential operator:

$$Lu := u^{(n)} + q_2(x)u^{(n-2)} + \ldots + q_n(x)u \qquad (n \geq 2)$$

defined on $H^n_{\text{loc}}(G)$. (Recall that, by definition, $H^k_{\text{loc}}(G)$ is the set of all complex functions $v \in L^2_{\text{loc}}(G)$ having distributional derivatives in $L^2_{\text{loc}}(G)$ of order up to k.) Given a complex number λ, the function $u\ G \to \mathbb{C}$, $u \equiv 0$ is called an eigenfunction of order -1 of the operator L with the eigenvalue λ. Furthermore, a function $u\ G \to \mathbb{C}$, $u \not\equiv 0$ is called an eigenfunction of order k $(k = 0, 1, \ldots)$ of the operator L with the eigenvalue λ if the function $u^* := Lu - \lambda u$ is an eigenfunction of order $(k-1)$ with the same eigenvalue λ.

Let us now be given a complete and minimal system $(u_\alpha) \subset L^2(G)$ of eigenfunctions of the operator L, denote by λ_α (resp. o_α) the eigenvalue (resp. order) of u_α and assume

$$\sup_\alpha o_\alpha < \infty, \tag{1}$$

$$\text{in case } o_\alpha > 0, \quad \lambda_\alpha u_\alpha - Lu_\alpha = u_{\alpha-1}. \tag{2}$$

We introduce some notations:

Index the n-th roots of λ_α such that

$$\mathrm{Re}\mu_{1,\alpha} \geq \ldots \geq \mathrm{Re}\mu_{n,\alpha}, \quad \mathrm{Im}\mu_{j,\alpha} > \mathrm{Im}\mu_{j+1,\alpha} \text{ in case } \mathrm{Re}\mu_{j,\alpha} = \mathrm{Re}\mu_{j+1,\alpha}$$

and put $\mu_\alpha := \mu_{m,\alpha}$, $\varrho_\alpha := |\mathrm{Re}\mu_\alpha|$, $\nu_\alpha := |\mathrm{Im}\mu_\alpha|$, where $m = \left[\frac{n+1}{2}\right]$;

$$\delta(\nu, \nu_\alpha) := \begin{cases} 1, & \nu > \nu_\alpha \\ \frac{1}{2}, & \nu = \nu_\alpha \\ 0, & \nu < \nu_\alpha; \end{cases}$$

$$W_R(t) := \begin{cases} \frac{\sin \nu(x-t)}{\pi(x-t)}, & |x-t| \leq R \\ 0, & |x-t| > R, \end{cases}$$

where $x \in K$, K is an arbitrary fixed compact interval $K \subset G$ and $R \in \in (0, \mathrm{dist}(K, \partial G))$;

$$D_{R_0} f := \frac{2}{R_0} \int_{\frac{R_0}{2}}^{R_0} f(R) dR, \quad 0 < R_0 < \mathrm{dist}(K, \partial G);$$

$W(t) := D_{R_0}(W_R)$;

$$\sigma_\nu(f, x) := \sum_{\nu_\alpha < \nu} \langle f, v_\alpha \rangle u_\alpha(x) + \sum_{\nu_\alpha = \nu}{}^{*} c_\alpha \langle f, v_\alpha \rangle u_\alpha(x),$$

where c_α are arbitrary constants, $|c_\alpha| \leq C$, and $\sum^*$ denotes the sum for any subset of $\{\alpha : \nu_\alpha = \nu\}$, $f \in L^2(G)$, $\nu > 0$, $x \in G$, (v_α) is the dual system of (u_α) (i.e. $(v_\alpha) \subset L^2(G)$ and $\langle v_k, u_j \rangle = \delta_{k,j}$);

$$S_\nu(f, x) := \int_{x-R}^{x+R} \frac{\sin \nu(y-x)}{\pi(y-x)} f(y) dy,$$

where $f \in L^2(G)$, $\nu > 0$, $x \pm R \in G$;

$$K_b := \{x \in G : \ \mathrm{dist}(x, K) \leq b\},$$

where $K \subset G$ is a compact interval and $0 < b < \mathrm{dist}(K, \partial G)$. We prove the following

THEOREM *Assume that (1), (2), $q_2 \equiv 0$, $u_\alpha^* \equiv 0$ and*

$$\sup_{t>0} \sum_{t \leq \nu_\alpha \leq t+1} 1 < \infty \tag{3}$$

are fulfilled. Then the following three statements are equivalent:

(a) *For any compact interval* $K \subset G$

$$\sup_{\alpha} \|v_\alpha\|_{L^2(G)} \cdot \|u_\alpha\|_{L^2(K)} < \infty.$$

(b) *For any compact interval* $K \subset G$ *and any subsum* $\sum^*$

$$\lim_{\nu\to\infty} \sup_{x\in K} |S_\nu(f,x) - \sigma_\nu(f,x)| = 0$$

for every $f \in L^2(G)$ *and every* $0 < R < \operatorname{dist}(K, \partial G)$.

(c) *For every compact interval* $K \subset G$ *and every subsum* $\sum^*$

$$\lim_{\nu\to\infty} \|f - \sigma_\nu(f)\|_{L^2(K)} = 0$$

for every $f \in L^2(G)$.

In the case of $|\mathrm{Im}\sqrt{\lambda_n}| \le c$, this statement was proved by V. A. Il'in in [8].

Proof of the Theorem (a)$\Longrightarrow$(b). Denote $w_1, \ldots, w_n$ the n-th roots of unity. Now let

$$\hat{u}_\alpha(y) := u_\alpha(y) + \int_x^y \sum_{p=1}^n \frac{w_p}{n\mu_\alpha^{n-1}} e^{\mu_\alpha w_p(y-\tau)} \cdot Q(\tau)d\tau$$

where

$$Q(\tau) := \sum_{s=2}^n q_s(\tau) u_\alpha^{(n-s)}(\tau) - u_\alpha^*(\tau).$$

According to the definition of $\hat{u}_\alpha(y)$ and the equations

$$0 = \sum_{p=1}^n w_p = \sum_{p=1}^n w_p^2 = \ldots = \sum_{p=1}^n w_p^{n-1}, \quad n = \sum_{p=1}^n w_p^n$$

we have

$$\hat{u}_\alpha(x) = u_\alpha(x), \quad \hat{u}'_\alpha(x) = u'_\alpha(x), \ldots, \hat{u}_\alpha^{(n-1)}(x) = u_\alpha^{(n-1)}(x)$$

and

$$\hat{u}_\alpha^{(n-1)}(y) = u_\alpha^{(n-1)}(y) + \mu_\alpha^{n-1} \int_x^y \sum_{p=1}^n \frac{w_p^n e^{\mu_\alpha w_p(y-\tau)}}{n\mu_\alpha^{n-1}} Q(\tau)d\tau$$

further

$$\hat{u}_\alpha^{(n)}(y) = u_\alpha^{(n)}(y) + Q(y) + \mu_\alpha^n \int_x^y \sum_{p=1}^n \frac{w_p}{n\mu_\alpha^{n-1}} e^{\mu_\alpha w_p(y-\tau)} Q(\tau)d\tau =$$

$$= u_\alpha^{(n)}(y) + Q(y) + \mu_\alpha^n(\hat{u}_\alpha(y) - u_\alpha(y)) = u_\alpha^{(n)}(y) + Q(y) +$$
$$+ \lambda_\alpha \hat{u}_\alpha(y) - \lambda_\alpha u_\alpha(y) = \lambda_\alpha \hat{u}_\alpha(y) + [u_\alpha^{(n)}(y) + Q(y) - \lambda_\alpha u_\alpha(y)]$$

i.e.

$$\hat{u}_\alpha^{(n)}(y) = \lambda_\alpha \hat{u}_\alpha(y).$$

Consequently $\hat{u}_\alpha(y)$ is a linear combination of the functions $e^{\mu_\alpha w_1 y}, \ldots, e^{\mu_\alpha w_n y}$. We shall distinguish two cases:

a) n is even i.e. $n = 2m$, $m \geq 2$;

b) n is odd i.e. $n = 2m - 1$, $m \geq 2$.

a) We have for any $0 \leq t \leq R \leq s \leq 2R$, $R > 0$:

$$0 = \begin{vmatrix} \hat{u}_\alpha(x) & \begin{matrix}\hat{u}_\alpha(x+t)+ \\ +\hat{u}_\alpha(x-t)\end{matrix} & \begin{matrix}\hat{u}_\alpha(x+2s)+ \\ +\hat{u}_\alpha(x-2s)\end{matrix} & \ldots & \begin{matrix}\hat{u}_\alpha(x+ms)+ \\ +\hat{u}_\alpha(x-ms)\end{matrix} \\ 1 & 2\mathrm{ch}\mu_{1,\alpha}t & 2\mathrm{ch}2\mu_{1,\alpha}s & \ldots & 2\mathrm{ch}m\mu_{1,\alpha}s \\ 1 & 2\mathrm{ch}\mu_{2,\alpha}t & 2\mathrm{ch}2\mu_{2,\alpha}s & \ldots & 2\mathrm{ch}m\mu_{2,\alpha}s \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 2\mathrm{ch}\mu_{m,\alpha}t & 2\mathrm{ch}2\mu_{m,\alpha}s & \ldots & 2\mathrm{ch}m\mu_{m,\alpha}s \end{vmatrix}.$$

Hence, multiplying the first column by $2\mathrm{ch}\mu_\alpha t$ and subtracting from the second one:

$$0 = \begin{vmatrix} \hat{u}_\alpha(x) & \begin{matrix}\hat{u}_\alpha(x+t)+\hat{u}_\alpha(x-t)- \\ -2\hat{u}_\alpha(x)\mathrm{ch}\mu_\alpha t\end{matrix} & \begin{matrix}\hat{u}_\alpha(x+2s)+ \\ +\hat{u}_\alpha(x-2s)\end{matrix} & \ldots & \begin{matrix}\hat{u}_\alpha(x+ms)+ \\ +\hat{u}_\alpha(x-ms)\end{matrix} \\ 1 & 2(\mathrm{ch}\mu_{1,\alpha}t - \mathrm{ch}\mu_\alpha t) & 2\mathrm{ch}2\mu_{1,\alpha}s & \ldots & 2\mathrm{ch}m\mu_{1,\alpha}s \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 2\mathrm{ch}2\mu_\alpha s & \ldots & 2\mathrm{ch}m\mu_\alpha s \end{vmatrix}.$$

Expanding this determinant according to the first row we get

$$[\hat{u}_\alpha(x+t) + \hat{u}_\alpha(x-t) - 2\hat{u}_\alpha(x)\mathrm{ch}\mu_\alpha t]d(\mu_\alpha, s) = \tag{4}$$
$$= \sum_{\substack{0 \leq k \leq m \\ k \neq 1}} d_k(\mu_\alpha, s, t)[\hat{u}_\alpha(x+ks) + \hat{u}_\alpha(x-ks)].$$

Taking into account

$$\hat{u}_\alpha(x+a) + \hat{u}_\alpha(x-a) =$$
$$= u_\alpha(x+a) + u_\alpha(x-a) + \int_x^{x+a} \sum_{p=1}^{n} \frac{w_p}{n\mu_\alpha^{n-1}} e^{\mu_\alpha w_p(x+a-\tau)} Q(\tau) d\tau -$$

$$-\int_{x-a}^{x}\sum_{p=1}^{n}\frac{w_p}{n\mu_\alpha^{n-1}}e^{\mu_\alpha w_p(x-a-\tau)}Q(\tau)d\tau$$

we obtain

(5)
$$[u_\alpha(x+t)+u_\alpha(x-t)-2u_\alpha(x)\mathrm{ch}\mu_\alpha t]d(\mu_\alpha,s)=$$
$$=\sum_{\substack{0\le k\le m\\ k\ne 1}} d_k(\mu_\alpha,s,t)[u_\alpha(x+ks)+u_\alpha(x-ks)]-$$
$$-d(\mu_\alpha,s)\int_{x}^{x+t}\sum_{p=1}^{n}\frac{w_p e^{\mu_\alpha w_p(x+t-\tau)}}{n\mu_\alpha^{n-1}}\cdot Q(\tau)d\tau+$$
$$+d(\mu_\alpha,s)\int_{x-t}^{x}\sum_{p=1}^{n}\frac{w_p e^{\mu_\alpha w_p(x-t-\tau)}}{n\mu_\alpha^{n-1}}\cdot Q(\tau)d\tau+\sum_{\substack{0\le k\le m\\ k\ne 1}} d_k(\mu_\alpha,s,t)\cdot$$
$$\cdot\Big[\int_{x}^{x+ks}\sum_{p=1}^{n}\frac{w_p e^{\mu_\alpha w_p(x+ks-\tau)}}{n\mu_\alpha^{n-1}}Q(\tau)d\tau-\int_{x-ks}^{x}\sum_{p=1}^{n}\frac{w_p e^{\mu_\alpha w_p(x-ks-\tau)}}{n\mu_\alpha^{n-1}}Q(\tau)d\tau\Big].$$

Denote
$$Q(\mu_\alpha,s):=e^{(m\mu_{1,\alpha}+\ldots+2\mu_{m-1,\alpha})s}.$$

Now we estimate $d_0(\mu_\alpha,s,t)$. In this determinant the minor corresponding to the element $2(\mathrm{ch}\mu_{1,\alpha}t-\mathrm{ch}\mu_\alpha t)$ is in absolute value smaller than or equal to
$$e^{\mathrm{Re}(m\mu_{2,\alpha}+(m-1)\mu_{3,\alpha}+\ldots+2\mu_\alpha)s},$$

further the minor corresponding to the element $2(\mathrm{ch}\mu_{j,\alpha}t-\mathrm{ch}\mu_\alpha t)$ $(1<j\le$ $\le m-1)$ is in absolute value smaller than or equal to
$$e^{\mathrm{Re}(m\mu_{1,\alpha}+\ldots+(m-j+2)\mu_{j-1,\alpha}+(m-j+1)\mu_{j+1,\alpha}+(m-j)\mu_{j+2,\alpha}+\ldots+3\mu_{m-1,\alpha}+2\mu_\alpha)s}.$$

This means that for the order of the terms dividing by $Q(\mu_\alpha,s)$ we obtain the following orders respectively:
$$|ch\mu_{1,\alpha}t-\mathrm{ch}\mu_\alpha t|\cdot e^{\mathrm{Re}2\mu_\alpha s-\mathrm{Re}(m\mu_{1,\alpha}-\mu_{2,\alpha}-\mu_{3,\alpha}-\ldots-\mu_{m-1,\alpha})s}\le$$
$$\le|\mathrm{ch}\mu_{1,\alpha}t-\mathrm{ch}\mu_\alpha t|\cdot e^{\mathrm{Re}2\mu_\alpha s-\mathrm{Re}2\mu_{1,\alpha}s},$$

$$|\mathrm{ch}\mu_{j,\alpha}t-\mathrm{ch}\mu_\alpha t|\cdot e^{\mathrm{Re}2\mu_\alpha s-\mathrm{Re}((m-j+1)\mu_{j,\alpha}-\mu_{j+1,\alpha}-\ldots-\mu_{m-1,\alpha})s}\le$$
$$\le|\mathrm{ch}\mu_{j,\alpha}t-\mathrm{ch}\mu_\alpha t|\cdot e^{\mathrm{Re}2\mu_\alpha s-\mathrm{Re}2\mu_{j,\alpha}s}\quad(1<j\le m-1).$$

On the other hand

$$|\mathrm{ch}\mu_{j,\alpha}t - \mathrm{ch}\mu_\alpha t| \le \begin{cases} c|\mu_\alpha|t, & 0 \le t \le \frac{1}{|\mu_\alpha|} \\ ce^{\mathrm{Re}\mu_{j,\alpha}t}, & t > \frac{1}{|\mu_\alpha|}. \end{cases}$$

Using these estimates we obtain

$$\text{(6)}\quad |d_0(\mu_\alpha,s,t)| \le c|Q(\mu_\alpha,s)|e^{\mathrm{Re}(2\mu_\alpha - 2\mu_{m-1,\alpha})s}\cdot|\mu_\alpha|t,\ 0 \le t \le \frac{1}{|\mu_\alpha|};$$

$$\text{(7)}\quad |d_0(\mu_\alpha,s,t)| \le c|Q(\mu_\alpha,s)|e^{\mathrm{Re}(2\mu_\alpha - 2\mu_{m-1,\alpha})s+\mathrm{Re}\mu_{m-1,\alpha}t},\ t > \frac{1}{|\mu_\alpha|}.$$

Now we estimate $d_k(\mu_\alpha,s,t)$ for $2 \le k \le m$. Repeating the method of estimate of $d_0(\mu_\alpha,s,t)$ we obtain

$$\text{(8)}\quad |d_k(\mu_\alpha,s,t)| \le c|Q(\mu_\alpha,s)|e^{-\mathrm{Re}2\mu_{m-1,\alpha}s}\cdot|\mu_\alpha|t,\quad 0 \le t \le \frac{1}{|\mu_\alpha|};$$

$$\text{(9)}\quad |d_k(\mu_\alpha,s,t)| \le c|Q(\mu_\alpha,s)|e^{-\mathrm{Re}2\mu_{m-1,\alpha}s+\mathrm{Re}\mu_{m-1,\alpha}t},\ t > \frac{1}{|\mu_\alpha|},$$

$(2 \le k \le m)$. Another estimate of d_0, d_k can be found in [2] without proof, therefore we proved here a little more exact one. We have that

$$\text{(10)}\qquad \left|\int_R^{2R} \frac{d(\mu_\alpha,s)}{Q(\mu_\alpha,s)}ds\right| > \frac{R}{2}$$

if $R_0 \ge R \ge R_{0/2} > 0$ and $|\mu_\alpha| \ge A(R_0) \ge 2$ where $A(R_0)$ is a constant depending on R_0 only (see [2]). (It can be proved if we use that

$$\text{(11)}\qquad \left|\frac{1}{R}\int_R^{2R} e^{\mu t}dt\right| \le \frac{4}{|R\mu|}e^{R\cdot\mathrm{Re}\mu},$$

if $\mu \in \mathbb{C}$, $\mathrm{Re}\mu \le 0$, $R > 0$.)

Now we estimate $|\langle u_\alpha, w\rangle - \delta(\nu,\nu_\alpha)u_\alpha(x)|$. We have

$$\text{(12)}\qquad \langle u_\alpha, w\rangle - \delta(\nu,\nu_\alpha)u_\alpha(x) =$$

$$= D_{R_0}\Big(\int_0^R \frac{\sin\nu t}{\pi t}[u_\alpha(x+t) + u_\alpha(x-t) - 2u_\alpha(x)\mathrm{ch}\mu_\alpha t]dt\Big) +$$

$$+ D_{R_0}\Big(\int_0^R \Big[\frac{2\sin\nu t\,\mathrm{ch}\mu_\alpha t}{\pi t} - \delta(\nu,\nu_\alpha)\Big]dt\Big)u_\alpha(x).$$

Here ([3], Lemma 3.2)

$$\left|D_{R_0}\Big(\int_0^R \frac{2\sin\nu t\operatorname{ch}\mu_\alpha t}{\pi t} - \delta(\nu,\nu_\alpha)dt\Big)\right| \le c\frac{(1+\varrho_\alpha^2)e^{|\varrho_\alpha|R_0}}{1+(\nu-\nu_\alpha)^2}. \tag{13}$$

We know that $w_{p+m} = -w_p$ $(p=1,\dots,m)$ hence

$$\sum_{p=1}^{n}\frac{w_p}{n\mu_\alpha^{n-1}}e^{w_p\mu_\alpha b} = \sum_{p=1}^{m}\frac{w_p}{m\mu_\alpha^{2m-1}}\cdot\frac{e^{w_p\mu_\alpha b}-e^{-w_p\mu_\alpha b}}{2} = \sum_{p=1}^{m}\frac{w_p\operatorname{sh}w_p\mu_\alpha b}{m\mu_a^{2m-1}},$$

so we obtain from (5)

$$\begin{aligned}(5^*)\qquad &[u_\alpha(x+t)+u_\alpha(x-t)-2u_\alpha(x)\operatorname{ch}\mu_\alpha t]d(\mu_\alpha,s) = \\ &= \sum_{\substack{0\le k\le m\\ k\ne 1}} d_k(\mu_\alpha,s,t)[u_\alpha(x+ks)+u_\alpha(x-ks)] - \\ &-d(\mu_\alpha,s)\int_{x-t}^{x+t}\sum_{p=1}^{m}\frac{w_p\operatorname{sh}\mu_\alpha w_p(t-|x-\tau|)}{m\mu_\alpha^{2m-1}}Q(\tau)d\tau + \\ &+\sum_{2\le k\le m} d_k(\mu_\alpha,s,t)\int_{x-ks}^{x+ks}\sum_{p=1}^{m}\frac{w_p\operatorname{sh}\mu_\alpha w_p(ks-|x-\tau|)}{m\mu_\alpha^{2m-1}}Q(\tau)d\tau = \\ &= \sum_{\substack{0\le k\le m\\ k\ne 1}} d_k(\mu_\alpha,s,t)[u_\alpha(x+ks)+u_\alpha(x-ks)] + \int_{x-ms}^{x+ms} D(\mu_\alpha,s,t,|x-\tau|)Q(\tau)d\tau\end{aligned}$$

where

$$D(\mu_\alpha,s,t,|x-\tau|) = \tag{14}$$

$$= \begin{cases} \displaystyle\sum_{2\le k\le m} d_k(\mu_\alpha,s,t)\sum_{p=1}^{m}\frac{w_p\operatorname{ch}\mu_\alpha w_p(ks-|x-\tau|)}{m\mu_\alpha^{2m-1}} - \\ \qquad \displaystyle -d(\mu_\alpha,s)\sum_{p=1}^{m}\frac{w_p\operatorname{sh}\mu_\alpha w_p(t-|x-\tau|)}{m\mu_\alpha^{2m-1}}, & \text{if } |x-\tau|\le t; \\ \displaystyle\sum_{2\le k\le m} d_k(\mu_\alpha,s,t)\sum_{p=1}^{m}\frac{w_p\operatorname{sh}\mu_\alpha w_p(ks-|x-\tau|)}{m\mu_\alpha^{2m-1}}, & \text{if } t\le|x-\tau|\le 2s; \\ \displaystyle\sum_{j\le k\le m} d_k(\mu_\alpha,s,t)\sum_{p=1}^{m}\frac{w_p\operatorname{sh}\mu_\alpha w_p(ks-|x-\tau|)}{m\mu_\alpha^{2m-1}}, & \\ \qquad\qquad \text{if } (j-1)s\le|x-\tau|\le js,\ 3\le j\le m. \end{cases}$$

For $D(\mu_\alpha, s, t, |x-\tau|)$ we have ([2], p. 277, without proof)

$$(15) \qquad |D(\mu_\alpha, s, t, |x-\tau|)| \le c|Q(\mu_\alpha, s)| \cdot |\mu_\alpha|^{1-2m} \cdot \min\{1, |\mu_\alpha t|\} \cdot e^{\varrho_\alpha s}.$$

For the sake of completeness we give here the estimate of $D(.)$. The calculation is similar to the method of estimate of $d_0(\mu_a, s, t)$. We shall distinguish two cases

$\boxed{1}$ $|x-\tau| \le t$, $\boxed{2}$ $t < |x-\tau| \le 2s$ or $(j_0-1)s < |x-\tau| \le j_0 s$, $3 \le j_0 \le m$.

$\boxed{1}$ $|x-\tau| \le t$.

We can write $D(\mu_\alpha, s, t, |x-\tau|)$ as a determinant, modified the determinant on page 4 in the following way: for $k = 2, 3, \ldots, m$ we change $\hat{u}_\alpha(x+ks) + \hat{u}_\alpha(x-ks)$ to $\sum_{p=1}^{m} \frac{w_p \mathrm{sh}\mu_\alpha w_p(ks-|x-\tau|)}{m\mu_\alpha^{2m-1}}$, we change $\hat{u}_\alpha(x+t) +$ $+ \hat{u}_\alpha(x-t) - 2\hat{u}_\alpha(x) ch\mu_\alpha t$ to $\sum_{p=1}^{m} \frac{w_p \mathrm{sh}\mu_\alpha w_p(t-|x-\tau|)}{m\mu_\alpha^{2m-1}}$, we change $\hat{u}_\alpha(x)$ to 0. We can write this determinant as the sum of m determinants in the following way:

$$D(\mu_\alpha, s, t, |x-t|) = \frac{1}{m\mu_\alpha^{2m-1}} \sum_{p=1}^{m} \omega_p *$$

$$* \begin{vmatrix} 0 & \mathrm{sh}\mu_\alpha w_p(t-|x-\tau|) & \mathrm{sh}\mu_\alpha w_p(2s-|x-\tau|) & \ldots & \mathrm{sh}\mu_\alpha w_p(ms-|x-\tau|) \\ 1 & 2(\mathrm{ch}\mu_{1,\alpha}t - \mathrm{ch}\mu_\alpha t) & 2\mathrm{ch}2\mu_{1,\alpha}s & \ldots & 2\mathrm{ch}m\mu_{1,\alpha}s \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 2\mathrm{ch}2\mu_\alpha s & \ldots & 2\mathrm{ch}m\mu_\alpha s \end{vmatrix}.$$

In this determinant there are many cancellations: e.g. if a product contains $e^{\mu_{q,\alpha}(t-|x-\tau|)}$ and $e^{k'\mu_{q,\alpha}s}$, then it is cancelled by the one where they are replaced by $e^{\mu_{q,\alpha}(k's-|x-\tau|)}$ and $e^{\mu_{q,\alpha}t}$, where $\mu_{q,\alpha} := \pm\mu_\alpha w_p$. (We recall that $2\mathrm{sh}z = e^z - e^{-z}$, $2\mathrm{ch}z = e^z + e^{-z}$.)

We shall prove the above estimate for all the m determinants, hence we can suppose $\mu_{q,\alpha} = w_p\mu_\alpha$. We denote $\mu_q = \mu_{q,\alpha}$. Expand the determinant as the signed sum of $m+1$-member products when the members are taken from different rows and columns of the matrix. Obviously we can suppose that the last 1 digit is chosen from the first column.

i) $1 \le q \le m-1$, $|\mu_\alpha|t \ge 1$.

We know that μ_q occurs in the first and $q+1$-th row. Hence all the products in the determinant can be pairwise added such that the sum of any pair has the following form:

i$_1$) $\mathrm{sh}\mu_q(t-|x-\tau|)\mathrm{ch}k\mu_q s-(\mathrm{ch}\mu_q t-\mathrm{ch}\mu_\alpha t)\mathrm{sh}\mu_q(ks-|x-\tau|)$ multiplied by a summand of the corresponding subdeterminant $(2\le k\le m)$

or

i$_2$) $\mathrm{sh}\mu_q(ks-|x-\tau|)\mathrm{ch}\ell\mu_q s-\mathrm{sh}\mu_q(\ell s-|x-\tau|)\mathrm{ch}k\mu_q s$ multiplied by a summand of the corresponding subdeterminant $(2\le k<\ell\le m)$.

In case i$_1$) we have

$$\mathrm{sh}\mu_q(t-|x-\tau|)\mathrm{ch}k\mu_q s-(\mathrm{ch}\mu_q t-\mathrm{ch}\mu_\alpha t)\mathrm{sh}\mu_q(ks-|x-\tau|)=$$
$$=e^{\mu_q(t-|x-\tau|)}e^{k\mu_q s}-e^{\mu_q t}e^{\mu_q(ks-|x-\tau|)}+O(e^{k\mu_q s})+O(e^{\varrho_\alpha R}e^{k\mu_q s})=$$
$$=O(e^{\varrho_\alpha R}e^{k\mu_q s}).$$

The corresponding subdeterminant does not contain μ_q in the exponents and neither contains any $e^{k\mu_j s}$, $1\le j\le m$. Hence this two-member sum can be estimated as

$$O(e^{\varrho_\alpha R}e^{k\mu_q s})O\Big(\prod_{\substack{\ell\ne k\\ j(\ell)\ne q}} e^{\ell\mu_{j(\ell)}s}\Big)=O(e^{\varrho_\alpha R}Q(\mu_\alpha,s)).$$

Analogously in i$_2$)

$$\mathrm{sh}\mu_q(ks-|x-\tau|)\mathrm{ch}\ell\mu_q s-\mathrm{sh}\mu_q(\ell s-|x-\tau|)\mathrm{ch}k\mu_q s=O(e^{\ell\mu_q s})$$

and in the corresponding subdeterminant there is no $e^{j\mu_q s}, e^{\mu_q t}, e^{k\mu_j s}, e^{\ell\mu_j s}$ $(j=1,\dots,m)$. Hence the two-member sum can be estimated as

$$O(e^{\ell\mu_q s})O\Big(e^{\mu_r t}\prod_{\substack{j\ne k,\ell\\ r(j)\ne r,q}} e^{j\mu_{r(j)}s}\Big)=O(Q(\mu_\alpha,s))$$

ii) $1\le q\le m-1$, $|\mu_\alpha|t\le 1$.

Then the proof of i) can be repeated word for word, only in i$_1$) we apply the trivial estimate

$$\mathrm{sh}\mu_q(t-|x-\tau|)\mathrm{ch}k\mu_q s-(\mathrm{ch}\mu_q t-\mathrm{ch}\mu_\alpha t)\mathrm{sh}\mu_q(ks-|x-\tau|)=$$
$$=O(|\mu_\alpha|t\cdot e^{k\mu_q s})$$

and in i$_2$) the estimate $\mathrm{ch}\mu_r t-\mathrm{ch}\mu_\alpha t=O(e^{\mu_r t})$ is substituted by $O(|\mu_\alpha|t)$.

iii) $q=m$.

We expand the determinant as the sum of $m+1$-member product. We can suppose that $\mathrm{sh}\mu_\alpha(t-|x-\tau|)$ is chosen from the second column and the last 1 from the first column, and then the corresponding subdeterminant can be trivially estimated by $O(Q(\mu_\alpha,s))$. Since

$$|\mathrm{sh}\mu_\alpha(t-|x-\varrho|)|\le\begin{cases} c|\mu_\alpha|t & \text{if } |\mu_\alpha|t\le 1\\ ce^{\varrho_\alpha R} & \text{if } |\mu_\alpha|t\ge 1\end{cases}$$

the desired estimate is proved.

[2] $t \le |x-\tau| \le 2s$ or $(j_0-1)s \le |x-\tau| \le j_0 s$ for some $3 \le j_0 \le m$.

In case $t \le |x-\tau| \le 2s$ the determinants defining D in case [1] has to be modified such that $\mathrm{sh}\mu_a w_p(t-|x-\tau|)$ is substituted by zero, in case $(j_0-1)s \le |x-\tau| \le j_0 s$, we write also zeros instead of $\mathrm{sh}\mu_\alpha w_p(2s-|x-\tau|)$, ..., $\mathrm{sh}\mu_\alpha w_r((j_0-1)s-|x-\tau|)$. In case $q=m$ all summands of the determinant have been estimated by $O(e^{\varrho_\alpha R}Q(\mu_\alpha,s))$. If some elements of the matrix are substituted by zero, the same estimate a fortiori holds. Hence we can suppose $q \le m-1$.

i) $|\mu_\alpha| t \ge 1$.

The first nonzero element in the first row is $\mathrm{sh}\mu_q(j_0 s-|x-\tau|)$, where $j_0=2$ in case $\tau \le |x-\tau| \le 2s$. The first and the $q+1$-th row contain exponents μ_q. As in [1], we take pairwise the products, and we can again suppose that the last 1 is chosen in the first column. We have the following cases

i$_1$) $\mathrm{sh}\mu_q(j_0 s-|x-\tau|)(\mathrm{ch}\mu_q t-\mathrm{ch}\mu_\alpha t)$ multiplied by a product, where there is not, $e^{j_0\mu_r s}$ and $e^{r\mu_q s}$ for $1 \le r \le q-1$. In case $t \le |x-\tau| \le 2s$ we have (by $j_0=2$)

$$\mathrm{sh}\mu_q(j_0 s-|x-\tau|)(\mathrm{ch}\mu_q t-\mathrm{ch}\mu_\alpha t)=O(e^{\mu_q(2s-t)}e^{\mu_q t})=O(e^{j_0 s\mu_q})$$

and then the two-member sum is estimated as

$$O(e^{j_0 s\mu_q})O\Big(\prod_{\substack{j\ne j_0\\ \tau(j)\ne q}} e^{j\mu_{\tau(j)}s}\Big)=O(Q(\mu_a,s)).$$

The same trick goes (even easier) if $(j_0-1)s \le |x-\tau| \le j_0 s$.

i$_2$) $\mathrm{sh}\mu_q(j_0 s-|x-\tau|)\mathrm{ch}j_1\mu_q s$ (where $2 \le j_1 < j_0$) multiplied by a product without $e^{r\mu_q s}$, $e^{j_0\mu_r s}$, $e^{j_1\mu_r s}$ $(r=1,\dots,m-1)$. Then

$$\mathrm{sh}\mu_q(j_0 s-|x-\tau|)\mathrm{ch}j_1\mu_q s=O(e^{\mu_q s}e^{j_1\mu_q s})=O(e^{j_0\mu_q s})$$

hence the whole product can be estimated as

$$O(e^{j_0\mu_q s})O\Big(e^{\mu_r t}\prod_{\substack{j\ne j_0,j_1\\ \tau(j)\ne q,r}} e^{j\mu_{\tau(j)}s}\Big)=O(Q(\mu_\alpha,s)).$$

i$_3$) $\mathrm{sh}\mu_q(j_1 s-|x-\tau|)\mathrm{ch}j_2\mu_q s-\mathrm{sh}\mu_q(j_2 s-|x-\tau|)\mathrm{ch}j_1\mu_q s$ (where $j_0 \le j_1 < j_2 \le m$) multiplied by a product without $e^{r\mu_q s}$, $e^{j_1\mu_r s}$, $e^{j_2\mu_r s}$ $(r=1,\dots,m-1)$. This is exactly the case i$_2$) of [1].

The above counting proves the estimate of $\underline{D}(\mu_\alpha,s,t,|x-\tau|)$ for n even.

From (5*) we obtain

$$|u_\alpha(x+t)+u_\alpha(x-t)-2u_\alpha(x)\mathrm{ch}\mu_\alpha t|\cdot|d(\mu_\alpha,s)|\leq$$
$$\leq \sum_{\substack{0\leq k\leq m\\ k\neq 1}} |d_k(\mu_\alpha,s,t)|\cdot|u_\alpha(x+ks)+u_\alpha(x-ks)|+$$
$$+\int_{x-ms}^{x+ms} |D(\mu_\alpha,s,t,|x-\tau|)|\cdot|Q(\tau)|d\tau.$$

Using (6), (7), (8), (9), (15) we get

$$|u_\alpha(x+t)+u_\alpha(x-t)-2u_\alpha(x)\mathrm{ch}\mu_\alpha t|\cdot\frac{|d(\mu_\alpha,s)|}{|Q(\mu_\alpha,s)|}\leq \tag{16}$$
$$\leq ce^{\mathrm{Re}(2\mu_\alpha-2\mu_{m-1,\alpha})s}\cdot|\mu_\alpha|t\cdot\|u_\alpha\|_{L^\infty(K_{2mR})}+$$
$$+c|\mu_\alpha|^{2-2m}t\cdot e^{\varrho_\alpha s}\int_{x-ms}^{x+ms}|Q(\tau)|d\tau,\ \text{if } 0\leq t\leq\frac{1}{|\mu_\alpha|}$$

and

$$|u_\alpha(x+t)+u_\alpha(x-t)-2u_\alpha(x)\mathrm{ch}\mu_\alpha t|\cdot\frac{|d(\mu_\alpha,s)|}{|Q(\mu_\alpha,s)|}\leq \tag{17}$$
$$\leq ce^{\mathrm{Re}(2\mu_\alpha-2\mu_{m-1,\alpha})s+\mathrm{Re}\mu_{m-1,\alpha}t}\cdot\|u_\alpha\|_{L^\infty(K_{2mR})}+$$
$$+c|\mu_\alpha|^{1-2m}\cdot e^{\varrho_\alpha s}\cdot\int_{x-ms}^{x+ms}|Q(\tau)|d\tau,\quad \text{if}\quad t>\frac{1}{|\mu_\alpha|}.$$

If $|\mu_\alpha|>\max\left\{1,\frac{1}{R}\right\}$ then we obtain from (16) and (17)

$$\frac{|d(\mu_\alpha,s)|}{|Q(\mu_\alpha,s)|}\int_0^R\left|\frac{u_\alpha(x+t)+u_\alpha(x-t)-2u_\alpha(x)\mathrm{ch}\mu_\alpha t}{t}\right|dt\leq$$

$$\leq ce^{\mathrm{Re}(2\mu_\alpha-2\mu_{m-1,\alpha})s}\cdot\|u_\alpha\|_{L^\infty(K_{2mR})}+c|\mu_\alpha|^{1-2m}e^{\varrho_\alpha s}\int_{x-ms}^{x+ms}|Q(\tau)|d\tau+$$

$$+c\frac{|\mu_\alpha|}{\mathrm{Re}\mu_{m-1,\alpha}}e^{\mathrm{Re}(2\mu_\alpha-2\mu_{m-1,\alpha})s+\mathrm{Re}\mu_{m-1,\alpha}R}\cdot\|u_\alpha\|_{L^\infty(K_{2mR})}+$$

$$+c|\mu_\alpha|^{1-2m}\log|\mu_\alpha|e^{\varrho_\alpha s}\int\limits_{x-ms}^{x+ms}|Q(\tau)|d\tau\le$$

$$\le ce^{\mathrm{Re}(2\mu_\alpha-2\mu_{m-1,\alpha})s+\mathrm{Re}\mu_{m-1,\alpha}R}\cdot\|u_\alpha\|_{L^\infty(K_{2mR})}+$$

$$+c|\mu_\alpha|^{1-2m}(1+\log|\mu_\alpha|)e^{\varrho_\alpha s}\int\limits_{x-ms}^{x+ms}|Q(\tau)|d\tau,$$

where we used $\mathrm{Re}\mu_{m-1,\alpha}\ge c_1|\mu_\alpha|$ ($c_1>0$ is a constant). Using (10) we have

$$(18)\qquad \int\limits_0^R\left|\frac{u_\alpha(x+t)+u_\alpha(x-t)-2u_\alpha(x)\mathrm{ch}\mu_\alpha t}{t}\right|dt\le ce^{-\mathrm{Re}\mu_{m-1,\alpha}R}.$$

$$\cdot\|u_\alpha\|_{L^\infty(K_{2mR})}e^{\mathrm{Re}2\mu_\alpha R}+c|\mu_\alpha|^{1-2m}\cdot\log|\mu_\alpha|e^{2\varrho_\alpha R}\int\limits_{x-2mR}^{x+2mR}|Q(\tau)|d\tau,$$

if $|\mu_\alpha|\ge A(R_0)\ge 2$.

We know ([2], (3)) that for any compact intervals K, K' with $K\subset K'\subset$ $\subset G$, $K\subset\mathrm{int}\,K'$ there exists $\varepsilon_0=\varepsilon_0(K,K')>0$ such that

$$(19)\qquad \|u_\alpha\|_{L^\infty(K)}e^{\varepsilon_0|\varrho_\alpha|}\le\frac{1}{\varepsilon_0}\|u_\alpha\|_{L^2(K')}.$$

We know also ([6], Theorem 2) that for any compact subinterval K of G there exists a constant $c>0$ such that

$$(20)\qquad \|u_\alpha^{(i)}\|_{L^\infty(K)}\le c(1+|\mu_\alpha|)^i\|u_\alpha\|_{L^\infty(K)},\ (i=0,1,\ldots).$$

Hence

$$(21)\qquad \int\limits_{x-2mR}^{x+2mR}|Q(\tau)|d\tau\le c(1+|\mu_\alpha|)^{2m-3}\|u_\alpha\|_{L^\infty(K_{4mR})}.$$

Using (21) we obtain from (18)

$$(22)\qquad \int\limits_0^R\left|\frac{u_\alpha(x+t)+u_\alpha(x-t)-2u_\alpha(x)\mathrm{ch}\mu_\alpha t}{t}\right|dt\le\cdot ce^{-\mathrm{Re}\mu_{m-1,\alpha}R}$$

$$\cdot\|u_\alpha\|_{L^\infty(K_{4mR})}e^{\mathrm{Re}2\mu_\alpha R}+c\frac{\log|\mu_\alpha|}{|\mu_\alpha|^2}\|u_\alpha\|_{L^\infty(K_{4mR})}e^{2\mathrm{Re}\mu_\alpha R},$$

if $|\mu_\alpha| \geq A(R_0) \geq 2$.

We recall that $R_0 \geq R \geq \frac{R_0}{2} > 0$. If we choose $R_0 > 0$ such that $2R_0 < \varepsilon_0$, then according to (12), (13), (19), (22) we have for $|\mu_\alpha| \geq A(R_0) \geq 2$

(23) $$|\langle u_\alpha, w\rangle - \delta(\nu, \nu_\alpha)u_\alpha(x)| \leq$$
$$\leq c\left(\frac{1}{1+(\nu-\nu_\alpha)^2} + e^{-\mathrm{Re}\mu_{m-1,\alpha}\frac{R_0}{2}} + \frac{\log|\mu_\alpha|}{|\mu_\alpha|^2}\right)\cdot \|u_\alpha\|_{L^2(K')}.$$

Now we prove the estimate analogous to (23) for $|\mu_\alpha| \leq A(R_0)$. We know that

$$\langle u_\alpha, W\rangle = D_{R_0}\Big(\int_0^R [u_\alpha(x+t)+u_\alpha(x-t)]\frac{\sin\nu t}{\pi t}dt\Big) =$$

$$= D_{R_0}\Big([u_\alpha(x+R)+u_\alpha(x-R)]\int_0^R \frac{\sin\nu t}{\pi t}dt\Big) -$$

$$-D_{R_0}\Big(\int_0^R u'_\alpha(x+t)-u'_\alpha(x-t)]\cdot\Big(\int_0^t \frac{\sin\nu\tau}{\pi\tau}d\tau\Big)dt\Big).$$

Here

$$\Big|\int_0^R \frac{\sin\nu t}{\pi t}dt\Big| = \Big|\int_0^{\nu R}\frac{\sin\mu}{\pi\mu}d\mu\Big| \leq c.$$

According to (20) we have

(24) $$|\langle u_\alpha, W\rangle| \leq c\|u_\alpha\|_{L^\infty(K_{R_0})} \quad \text{if} \quad |\mu_\alpha| \leq A(R_0).$$

Now we prove (a)$\Longrightarrow$(b).

From (23), (24), (3) follows

$$\sum_{\alpha=1}^{\infty}|\langle u_\alpha, W\rangle - \delta(\nu,\nu_\alpha)u_\alpha(x)|\cdot\|v_\alpha\|_{L^2(G)} \leq$$

$$\leq \sum_{\alpha=1}^{\infty}\frac{|\langle u_\alpha, W\rangle - \delta(\nu,\nu_\alpha)u_\alpha(x)|}{\|u_\alpha\|_{L^2(K')}} \leq$$

$$\leq c\sum_{\alpha=1}^{\infty}\Big(\frac{1}{1+(\nu-\nu_\alpha)^2} + e^{-c_1|\mu_\alpha|\frac{R}{2}} + \frac{\log|\mu_\alpha|}{|\mu_\alpha|^2}\Big) \leq c_3,$$

where c_3 is a constant independent of ν. Hence for any fixed x and ν the series

$$F(x,y) := \sum_{\alpha=1}^{\infty} [\langle u_\alpha, W\rangle - \delta(\nu,\nu_\alpha)u_\alpha(x)]\overline{v_\alpha(y)}$$

is absolutely convergent in $L^2_y(G)$ and

$$\int_G u_\alpha(y)F(x,y)dy = \langle u_\alpha, W\rangle - \delta(\nu,\nu_\alpha)u_\alpha(x).$$

Since (u_α) is complete and minimal therefore the Fourier-expansion is unique. Hence

$$F(x,y) = W(y) - \sum_{\nu_\alpha<\nu} u_\alpha(x)\overline{v_\alpha(y)} - \frac{1}{2}\sum_{\nu_\alpha=\nu} u_\alpha(x)\overline{v_\alpha(y)}$$

and

$$\sup_{\nu>0}\sup_{x\in K}\left\|W(y) - \sum_{\nu_\alpha<\nu} u_\alpha(x)\overline{v_\alpha(y)} - \frac{1}{2}\sum_{\nu_\alpha=\nu} u_\alpha(x)\overline{v_\alpha(y)}\right\|_{L^2_y(G)} = M < \infty.$$

Thus for every $f \in L^2(G)$ we have

$$\sup_{\nu>0}\sup_{x\in K}|S_\nu(f,x) - \sigma_\nu(f,x)| \le M\|f\|_{L^2(G)}$$

for the variant of σ_ν where $\sum^* = \frac{1}{2}\sum_{\nu_\alpha=\nu}\langle f, v_\alpha\rangle u_\alpha(x)$. The same estimate holds for any other σ_ν as well, since

$$|\langle f, v_\alpha\rangle||u_\alpha(x)| \le \|f\|_{L_2(G)}\|v_\alpha\|_{L_2(G)}\|u_\alpha\|_{L_\infty(K)} \le$$
$$\le c\|f\|_{L_2(G)}\|v_\alpha\|_{L_2(G)}\|u_\alpha\|_{L_\infty(K')} \le c\|f\|_{L_2(G)}.$$

Now it suffices to show that

$$\lim_{\nu\to\infty}\sup_{x\in K}|S_\nu(f,x) - \sigma_\nu(f,x)| = 0$$

for any f from a dense subset of $L^2(G)$. But this last property is satisfied for any finite linear combination f of the eigenfunctions u_α (lin (u_α) is a dense subspace, because then f is continuously differentiable and $\sigma_\nu(f,x) \equiv f(x)$ for ν sufficiently large, therefore one can apply a classical result of the theory of Fourier-series ([8], p. 95). Hence in the case of $n = 2m$, $m = 2, 3, \dots$ (a)$\Longrightarrow$(b) is proved.

b) We have for any $0 \le t \le R \le s \le 2R$, $R > 0$:

$$0 = \begin{vmatrix} \hat{u}_\alpha(x-ms)..\hat{u}_\alpha(x-2s) & \begin{matrix}\hat{u}_\alpha(x-t)+\\ +\hat{u}_\alpha(x+t)-\\ -2\hat{u}_\alpha(x)\mathrm{ch}\mu_\alpha t\end{matrix} & \hat{u}_\alpha(x+2s)..\hat{u}_\alpha(x+(m+1)s) \\ e^{-m\mu_{1,\alpha}s}\dots e^{-2\mu_{1,\alpha}s} & 2(\mathrm{ch}\mu_{1,\alpha}t - \mathrm{ch}\mu_\alpha t) & e^{2\mu_{1,\alpha}s}\dots e^{(m+1)\mu_{1,\alpha}s} \\ \vdots & \vdots & \vdots \\ e^{-m\mu_{m-1,\alpha}s}\dots e^{-2\mu_{m-1,\alpha}s} & 2(\mathrm{ch}\mu_{m-1,\alpha}t - \mathrm{ch}\mu_\alpha t) & e^{2\mu_{m-1,\alpha}s}\dots e^{(m+1)\mu_{m-1,\alpha}s} \\ e^{-m\mu_\alpha s}\dots e^{-2\mu_\alpha s} & 0 & e^{2\mu_\alpha s}\dots e^{(m+1)\mu_\alpha s} \\ e^{-m\mu_{m+1,\alpha}s}\dots e^{-2\mu_{m+1,\alpha}s} & 2(\mathrm{ch}\mu_{m+1,\alpha}t - \mathrm{ch}\mu_\alpha t) & e^{2\mu_{m+1,\alpha}s}\dots e^{(m+1)\mu_{m+1,\alpha}s} \\ \vdots & \vdots & \vdots \\ e^{-m\mu_{n,\alpha}s}\dots e^{-2\mu_{n,\alpha}s} & 2(\mathrm{ch}\mu_{n,\alpha}t - \mathrm{ch}\mu_\alpha t) & e^{2\mu_{n,\alpha}s}\dots e^{(m+1)\mu_{n,\alpha}s} \end{vmatrix}.$$

Expanding this determinant according to the first row we get

$$(4') \qquad [\hat{u}_\alpha(x-t) + \hat{u}_\alpha(x+t) - 2\hat{u}_\alpha(x)\mathrm{ch}\mu_\alpha t]d(\mu_\alpha, s) = \\ = \sum_{\substack{-m\le k\le m+1\\ |k|\ge 2}} d_k(\mu_\alpha, s, t)\hat{u}_\alpha(x+ks).$$

Taking into account the definition of $\hat{u}_\alpha$ we obtain

$$(5') \qquad [u_\alpha(x-t) + u_\alpha(x+t) - 2u_\alpha(x)\mathrm{ch}\mu_\alpha t]d(\mu_\alpha, s) =$$

$$= \sum_{\substack{-m\le k\le m+1\\ |k|\ge 2}} d_k(\mu_\alpha, s, t)u_\alpha(x+ks) - d(\mu_\alpha, s)\int_x^{x+t}\sum_{p=1}^n \frac{w_p e}{n\mu_\alpha^{n-1}} e^{\mu_\alpha w_p(x+t-\tau)}Q(\tau)d\tau +$$

$$+d(\mu_\alpha, s)\int_{x-t}^{x}\sum_{p=1}^n \frac{w_p e^{\mu_\alpha w_p(x-t-\tau)}}{n\mu_\alpha^{n-1}}Q(\tau)d\tau +$$

$$+ \sum_{\substack{-m\le k\le m+1\\ |k|\ge 2}} d_k(\mu_\alpha, s, t)\int_x^{x+ks}\sum_{p=1}^n \frac{w_p e^{\mu_\alpha w_p(x+ks-\tau)}}{n\mu_\alpha^{n-1}}Q(\tau)d\tau.$$

Denote

$$Q(\mu_\alpha, s) := e^{((m+1)\mu_{1,\alpha}+\dots+2\mu_{m,\alpha}-2\mu_{m+1,\alpha}-\dots-m\mu_{n,\alpha})s}.$$

Now we estimate $d_k(\mu_\alpha, s, t)$. In this determinant the minor corresponding to the element $2(\mathrm{ch}\mu_{1,\alpha}t - \mathrm{ch}\mu_\alpha t)$ is in absolute value smaller than

$$e^{\mathrm{Re}((m+1)\mu_{2,\alpha}+m\mu_{3,\alpha}+\dots+2\mu_{m+1,\alpha}-2\mu_{m+2,\alpha}-\dots-(m-1)\mu_{n,\alpha})s},$$

further the minor corresponding to the element $2(\mathrm{ch}\mu_{j,\alpha}t - \mathrm{ch}\mu_\alpha t)$, $(1 < < j \le m-1,\ m+1 \le j \le n)$ is in absolute value smaller than

$$e^{\mathrm{Re}((m+1)\mu_{1,\alpha}+\dots+(m+1-j+2)\mu_{j-1,\alpha}+(m+1-j+1)\mu_{j+1,\alpha}+(m+1-j)\mu_{j+2,\alpha}+\dots}.$$

$$\cdot e^{\dots+3\mu_{m,\alpha}+2\mu_{m+1,\alpha}-2\mu_{m+2,\alpha}-\dots-(m-1)\mu_{n,\alpha})s},$$

if $1<j\leq m-1$;

$$e^{\mathrm{Re}((m+1)\mu_{1,\alpha}+\dots+2\mu_{m,\alpha}-2_{m+2,\alpha}-\dots-(m-1)\mu_{n,\alpha})s}, \text{ if } j=m+1;$$

$$e^{\mathrm{Re}(m+1)\mu_{1,\alpha}+\dots+2\mu_{m,\alpha}-2\mu_{m+1,\alpha}-3\mu_{m+2,\alpha}-\dots-(j-m)\mu_{j-1,\alpha}-(j-m+1)\mu_{j+1,\alpha}}\cdot$$
$$\cdot e^{-(j-m+2)\mu_{j+2,\alpha}-\dots-(m-1)\mu_{n,\alpha})s}, \text{ if } m+1<j\leq n.$$

This means that for the order of the terms dividing by $Q(\mu_\alpha, s)$ we obtain the following orders respectively:

$$|\mathrm{ch}\mu_{1,\alpha}t-\mathrm{ch}\mu_\alpha t|\cdot$$
$$\cdot e^{\mathrm{Re}(\mu_{2,\alpha}+\dots+\mu_{m,\alpha}-(m+1)\mu_{1,\alpha}+4\mu_{m+1,\alpha}+\mu_{m+2,\alpha}+\dots+\mu_{n,\alpha})s}\leq$$
$$\leq|\mathrm{ch}\mu_{1,\alpha}t-\mathrm{ch}\mu_\alpha t|\cdot e^{-\mathrm{Re}2\mu_{1,\alpha}s};$$

$$|ch\mu_{j,\alpha}t-\mathrm{ch}\mu_\alpha t|\cdot$$
$$\cdot e^{\mathrm{Re}(-(m+1-j+1)\mu_{j,\alpha}+\mu_{j+1,\alpha}+\dots+\mu_{m,\alpha}+4\mu_{m+1,\alpha}+\mu_{m+2,\alpha}+\dots+\mu_{n,\alpha})s}\leq$$
$$\leq|\mathrm{ch}\mu_{j,\alpha}t-\mathrm{ch}\mu_\alpha t|\cdot e^{-\mathrm{Re}2\mu_{j,\alpha}s},\quad \text{if } 1<j\leq m-1;$$

$$|\mathrm{ch}\mu_{m+1,\alpha}t-\mathrm{ch}\mu_\alpha t|\cdot e^{\mathrm{Re}(2\mu_{m+1,\alpha}+\mu_{m+2,\alpha}+\dots+\mu_{n,\alpha})s}\leq$$
$$\leq|\mathrm{ch}\mu_{m+1,\alpha}t-\mathrm{ch}\mu_\alpha t|\cdot e^{\mathrm{Re}2\mu_{m+1,\alpha}s};$$
$$|\mathrm{ch}\mu_{j,\alpha}t-\mathrm{ch}\mu_\alpha t|\cdot e^{\mathrm{Re}((j+1-m)\mu_{j,\alpha}+\mu_{j+1,\alpha}+\dots+\mu_{n,\alpha})s}\leq$$
$$\leq|\mathrm{ch}\mu_{j,\alpha}t-\mathrm{ch}\mu_\alpha t|\cdot e^{\mathrm{Re}2\mu_{j,\alpha}s},\quad \text{if } m+1<j\leq n.$$

On the other hand

$$|\mathrm{ch}\mu_{j,\alpha}t-\mathrm{ch}\mu_\alpha t|\leq\begin{cases} c|\mu_\alpha|t, & 0\leq t\leq\frac{1}{|\mu_\alpha|}\\ ce^{|\mathrm{Re}\mu_{j,\alpha}|t}, & t>\frac{1}{|\mu_\alpha|}\end{cases}$$

(where we used $|\mathrm{Re}\mu_\alpha|\leq|\mathrm{Re}\mu_{k,\alpha}|$, $k\neq m$). Obviously $\mathrm{Re}\mu_{m-1,\alpha}>0$ and $\mathrm{Re}\mu_{m+1,\alpha}<0$, therefore

$$|\mathrm{ch}\mu_{1,\alpha}t-\mathrm{ch}\mu_\alpha t|\cdot e^{-\mathrm{Re}2\mu_{1,\alpha}s}\leq\begin{cases} c|\mu_\alpha|te^{-\mathrm{Re}2\mu_{1,\alpha}s}, & 0\leq t\leq\frac{1}{|\mu_\alpha|}\\ ce^{\mathrm{Re}\mu_{1,\alpha}(t-2s)}, & t>\frac{1}{|\mu_\alpha|};\end{cases}$$

if $1 < j \le m-1$ we have

$$|\mathrm{ch}\mu_{j,\alpha}t - \mathrm{ch}\mu_\alpha t| \cdot e^{-\mathrm{Re}2\mu_{j,\alpha}s} \le \begin{cases} c|\mu_\alpha|te^{-\mathrm{Re}2\mu_{j,\alpha}s}, & 0 \le t \le \frac{1}{|\mu_\alpha|} \\ ce^{\mathrm{Re}\mu_{j,\alpha}(t-2s)}, & t > \frac{1}{|\mu_\alpha|}; \end{cases}$$

if $j = m+1$ we have

$$|\mathrm{ch}\mu_{m+1,\alpha}t - \mathrm{ch}\mu_\alpha t| \cdot e^{\mathrm{Re}2\mu_{m+1,\alpha}s} \le \begin{cases} c|\mu_\alpha|te^{-|\mathrm{Re}2\mu_{m+1,\alpha}|s}, & 0 \le t \le \frac{1}{|\mu_\alpha|} \\ ce^{|\mathrm{Re}\mu_{m+1,\alpha}|\cdot(t-2s)}, & t > \frac{1}{|\mu_\alpha|}; \end{cases}$$

if $m+1 < j \le n$ we have

$$|\mathrm{ch}\mu_{j,\alpha}t - \mathrm{ch}\mu_\alpha t| \cdot e^{\mathrm{Re}2\mu_{j,\alpha}s} \le \begin{cases} c|\mu_\alpha|te^{-|\mathrm{Re}2\mu_{j,\alpha}|s}, & 0 \le t \le \frac{1}{|\mu_\alpha|} \\ ce^{|\mathrm{Re}\mu_{j,\alpha}|\cdot(t-2s)}, & t > \frac{1}{|\mu_\alpha|}. \end{cases}$$

Using these estimates we obtain

(8′) $$|d_k(\mu_\alpha, s, t)| \le c|Q(\mu_\alpha, s)| \cdot \left(e^{-\mathrm{Re}2\mu_{m-1,\alpha}s} + e^{\mathrm{Re}2\mu_{m+1,\alpha}s}\right) \cdot |\mu_\alpha|t, \quad 0 \le t \le \frac{1}{|\mu_\alpha|};$$

(9′)
$$|d_k(\mu_\alpha, s, t)| \le c|Q(\mu_\alpha, s)| \cdot \left(e^{\mathrm{Re}\mu_{m-1,\alpha}(t-2s)} + e^{|\mathrm{Re}\mu_{m+1,\alpha}|\cdot(t-2s)}\right), \ t > \frac{1}{|\mu_\alpha|}.$$

Now we have that

(10′) $$\left|\int_R^{2R} \frac{d(\mu_\alpha, s)}{Q(\mu_\alpha, s)} ds\right| > \frac{R}{2}$$

if $R_0 \ge R \ge R_0/2 > 0$ and $|\mu_\alpha| \ge A(R_0) \ge 2$.

Next we estimate $|\langle u_\alpha, W\rangle - \delta(\nu, \nu_\alpha)u_\alpha(x)|$. We have

(12′) $$\langle u_\alpha, W\rangle - \delta(\nu, \nu_\alpha)u_\alpha(x) =$$
$$= D_{R_0}\Big(\int_0^R \frac{\sin \nu t}{\pi t}[u_\alpha(x+t) + u_\alpha(x-t) - 2u_\alpha(x)\mathrm{ch}\mu_\alpha t]dt\Big) +$$
$$+ D_{R_0}\Big(\int_0^R \Big[\frac{2\sin \nu t \mathrm{ch}\mu_\alpha t}{\pi t}dt - \delta(\nu, \nu_\alpha)\Big]dt\Big)u_\alpha(x).$$

We obtain from (5′)

$$(25)\qquad [u_\alpha(x-t)+u_\alpha(x+t)-2u_\alpha(x)\operatorname{ch}\mu_\alpha t]d(\mu_\alpha,s)=$$

$$=\sum_{\substack{-m\le k\le m+1\\ |k|\ge 2}} d_k(\mu_\alpha,s,t)u_\alpha(x+ks)+\int\limits_{x-ms}^{x+(m+1)s} D(\mu_\alpha,s,t,x-\tau)Q(\tau)d\tau,$$

where

$$D(\mu_\alpha,s,t,x-\tau)=$$

$$=\begin{cases}
-d(\mu_\alpha,s)\sum\limits_{p=1}^{n}\dfrac{w_p e^{\mu_\alpha w_p(x+t-\tau)}}{n\mu_\alpha^{n-1}}+\sum\limits_{2\le k\le m+1} d_k(\mu_\alpha,s,t)\sum\limits_{p=1}^{n}\dfrac{w_p e^{\mu_\alpha w_p(x+ks-\tau)}}{n\mu_\alpha^{n-1}}, & \text{if } 0\le\tau-x\le t;\\[2ex]
\sum\limits_{2\le k\le m+1} d_k(\mu_\alpha,s,t)\sum\limits_{p=1}^{n}\dfrac{w_p e^{\mu_\alpha w_p(x+ks-\tau)}}{n\mu_\alpha^{n-1}}, & \text{if } t\le\tau-x\le 2s;\\[2ex]
\sum\limits_{j\le k\le m+1} d_k(\mu_\alpha,s,t)\sum\limits_{p=1}^{n}\dfrac{w_p e^{\mu_\alpha w_p(x+ks-\tau)}}{n\mu_\alpha^{n-1}}, & \text{if } (j-1)s\le\tau-x\le js; 3\le j\le m+1;\\[2ex]
d(\mu_\alpha,s)\sum\limits_{p=1}^{n}\dfrac{w_p e^{\mu_\alpha w_p(x-t-\tau)}}{n\mu_\alpha^{n-1}}-\sum\limits_{-m\le k\le -2} d_k(\mu_\alpha,s,t)\sum\limits_{p=1}^{n}\dfrac{w_p e^{\mu_\alpha w_p(x+ks-\tau)}}{n\mu_\alpha^{n-1}}, & \text{if } -t\le\tau-x\le 0;\\[2ex]
-\sum\limits_{-m\le k\le -2} d_k(\mu_\alpha,s,t)\sum\limits_{p=1}^{n}\dfrac{w_p e^{\mu_\alpha w_p(x+ks-\tau)}}{n\mu_\alpha^{n-1}}, & \text{if } -2s\le\tau-x\le -t;\\[2ex]
-\sum\limits_{-m\le k\le j} d_k(\mu_\alpha,s,t)\sum\limits_{p=1}^{n}\dfrac{w_p e^{\mu_\alpha w_p(x+ks-\tau)}}{n\mu_\alpha^{n-1}}, & \text{if } (j-1)s\le\tau-x\le js,\ -m+1\le j\le -2.
\end{cases}$$

Here we give the proof of the estimate

$$(15') \qquad |D(\mu_\alpha, s, t, x-\tau)| \le c|Q(\mu_\alpha, s)||\mu_\alpha|^{1-n} \min\{1, |\mu_\alpha|t\} \cdot e^{n\varrho_\alpha s}.$$

This is a slightly modified version of [2], p. 275 ($e^{2\varrho_\alpha s}$ has been changed to $e^{n\varrho_\alpha s}$). With $e^{2\varrho_\alpha s}$ the version of (15′) does not hold for $n = 3$.

Begin the proof of (15′) with

$\boxed{1}$ $x \le \tau \le x + t$.

Then D can be expressed in the following form

$$D = \sum_{p=1}^{n} \frac{w_p}{n\mu_\alpha^{n-1}} \cdot$$

$$\cdot \begin{vmatrix} 0 & \dots & 0 & e^{\mu_\alpha w_p (t+x-\tau)} & e^{\mu_\alpha w_p (2s+x-\tau)} & \dots & e^{\mu_\alpha w_p ((m+1)s+x-\tau)} \\ e^{-m\mu_{1,\alpha} s} & \dots & e^{-2\mu_{1,\alpha} s} & 2(\mathrm{ch}\mu_{1,\alpha} t - \mathrm{ch}\mu_n t) & e^{2\mu_{1,\alpha} s} & \dots & e^{(m+1)\mu_{1,\alpha} s} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ e^{-m\mu_{n,\alpha} s} & \dots & e^{-2\mu_{n,\alpha} s} & 2(\mathrm{ch}\mu_{n,\alpha} t - \mathrm{ch}\mu_\alpha t) & e^{2\mu_{n,\alpha} s} & \dots & e^{(m+1)\mu_{n,\alpha} s} \end{vmatrix}.$$

Let again $\mu_q := \mu_{q,\alpha} := \mu_\alpha w_p$. Following the ideas used in case of n even, we have the following cases in expanding the determinants

i) $e^{\mu_q(t+x-\tau)} \cdot e^{-k\mu_q s}$ multiplied by a product where there is no t, μ_q and $e^{-k\mu_r s}$ for any r. If $|\mu_\alpha| t \le 1$ then

$$e^{\mu_q(t+x-\tau)} = O(|\mu_\alpha| t), \quad e^{-k\mu_q s} \prod_{\substack{j \ne -k \\ r(j) \ne q}} e^{j\mu_{r(j)} s} = O(Q(\mu_\alpha, s)).$$

If $|\mu_\alpha| t \ge 1$, then in case $q \ge m$ we have $e^{\mu_q(t+x-\tau)} = O(e^{\varrho_\alpha R})$ and if $q \le m-1$, then there exists $j \ge 2$ with $r(j) \ge m+1$, hence $e^{\mu_q(t+x-\tau)} e^{-k\mu_q s} e^{j\mu_{r(j)} s} = = O(e^{-k\mu_{r(j)} s} e^{j\mu_q s})$ which shows that the whole product is $O(Q(\mu_\alpha, s))$.

ii) $\left[e^{\mu_q(t+x-\tau)} e^{k\mu_q s} - e^{\mu_q(ks+x-\tau)} \cdot 2(\mathrm{ch}\mu_q t - \mathrm{ch}\mu_\alpha t)\right] \cdot \prod_{\substack{j \ne k \\ r(j) \ne q}} e^{j\mu_{r(j)} s}$.

If $q \le m-1$, then this is $O\Big(e^{k\mu_q s} e^{\varrho_\alpha R} \cdot \prod_{\substack{j \ne k \\ r(j) \ne q}} e^{j\mu_{r(j)} s}\Big) = O(e^{\varrho_\alpha R} Q(\mu_\alpha s))$

in case $|\mu_\alpha| t \ge 1$, and $O(|\mu_\alpha| t \cdot e^{k\mu_q s} \cdot \prod e^{j\mu_{r(j)} s}) = O(|\mu_\alpha| t \cdot Q(\mu_\alpha, s))$ in case $|\mu_\alpha| t \le 1$. If $q = m$, then $e^{\mu_q(t+x-\tau)} e^{k\mu_q s} = O(e^{\varrho_\alpha R} e^{k\mu_q s})$ resp. $O(|\mu_\alpha| t e^{k\mu_q s})$ if $|\mu_\alpha| t \ge 1$ resp. ≤ 1.

If $q \ge m+1$ we estimate the expression in brackets by $O(e^{\mu_q(ks-2R)})$ resp. $O(e^{\mu_q(ks-2R)} |\mu_\alpha| t)$ if $|\mu_\alpha| t \ge 1$ resp. ≤ 1. Since k is paired with μ_q which has negative real part, there exists $j_0 \le -2$ and $r(j_0) \le m$

such that the factor $e^{j_0\mu_{r(j)}s}$ occurs in the product ii). Now we have $e^{\mu_q(ks-2R)}e^{j_0\mu_{r(j)}s} = O(e^{j_0\mu_{r(j)}s}\cdot e^{2\varrho_\alpha R})$. Indeed, if $r(j_0)\le m-1$, then $\mathrm{Re}\mu_{r(j_0)}>0$, $\mathrm{Re}\mu_q<0$, hence $e^{\mu_q((k-j_0)s-2R)}=O(1)=O(e^{(k-j_0)\mu_{r(j_0)}})$ and if $r(j_0)=m$ then $e^{\mu_q((k-j_0)s-2R)}e^{\mu_{r(j_0)}(j_0-k)s}=e^{(\mu_q-\mu_\alpha)((k-j_0)s-2R)}\cdot$ $\cdot\, e^{-2R\mu_\alpha}=O(e^{-2R\mu_\alpha})=O(e^{2\varrho_\alpha R})$. Consequently we can estimate the products ii) by

$$O\Big(\min\{1,|\mu_\alpha|t\}\cdot e^{2\varrho_\alpha R}e^{j_0\mu_q s+k\mu_{r(j_0)}s}\prod_{\substack{j\ne k,j_0\\ r(j)\ne q}} e^{j\mu_{r(j)}s}\Big)=$$

$$=O(\min\{1,|\mu_\alpha|t\}\cdot e^{2\varrho_\alpha r}\cdot Q(\mu_\alpha,s)).$$

iii) $e^{\mu_q(ks+x-\tau)}e^{\mu_q\ell s}-e^{\mu_q(\ell s+x-\tau)}e^{\mu_q ks}=0$ does not give new members.

iv) $e^{\mu_q(ks+x-\tau)}e^{-\mu_q\ell s}\cdot 2(\mathrm{ch}\mu_r t-\mathrm{ch}\mu_\alpha t)\prod\limits_{\substack{j\ne k,-\ell\\ r(j)\ne q,r}} e^{j\mu_{r(j)}s}$.

If ($q\le m-1$ and $r\ge m+1$) resp. ($q\ge m+1$ and $r\le m-1$) then

$$e^{\mu_q((k-\ell)s+x-\tau)}(\mathrm{ch}\mu_r t-\mathrm{ch}\mu_\alpha t)=O(e^{\mu_q ks}e^{-\mu_r\ell s}\min\{1,|\mu_\alpha|t\}),$$

resp. $O(e^{-\mu_q\ell s}e^{\mu_r ks}\min\{1,|\mu_\alpha|t\})$, hence the whole product is estimated by $O(\min\{1,|\mu_\alpha|t\}Q(\mu_\alpha,s))$. If ($q=m$ and $r\ge m+1$) resp. ($q=m$ and $r\le m-1$) then we get $e^{\mu_q((k-\ell)s+x-\tau)}(\mathrm{ch}\mu_r t-\mathrm{ch}\mu_\alpha t)=O(e^{\mu_q ks}-$ $-\,e^{\mu_r\ell s}e^{2\varrho_\alpha R}\cdot\min\{1,|\mu_\alpha|t\})$, resp. $O(e^{-\mu_q\ell s}e^{\mu_r ks}e^{2\varrho_\alpha R}\min\{1,|\mu_\alpha|t\})$. Let now $q\le m-1$ and $r\le m-1$. Then there exists $j\ge 2$ with $r(j)\ge m+1$ and then

$$e^{\mu_q((k-\ell)s+x-\tau)}(\mathrm{ch}\mu_r t-\mathrm{ch}\mu_\alpha t)e^{j\mu_{r(j)}s}=$$

$$=O(\max\{1,|\mu_\alpha|t\}\cdot e^{\mu_q ks}e^{\mu_r js}e^{-\ell\mu_{r(j)}s})$$

hence the whole product is $O(\max\{1,|\mu_\alpha|t\}Q(\mu_\alpha,s))$. Finally let $q\ge m+1$ and $r\ge m+1$. Then there exists $j\le -2$ with $r(j)\le m$, hence

$$e^{\mu_q((k-\ell)s+x-\tau)}(\mathrm{ch}\mu_r t-\mathrm{ch}\mu_\alpha t)e^{j\mu_{r(j)}s}=$$

$$O(\min\{1,|\mu_\alpha|t\}\cdot e^{-\mu_q\ell s}e^{j\mu_r s}e^{k\mu_{r(j)}s}e^{2\varrho_\alpha R}),$$

because after dividing with the exponentials we get

$$e^{\mu_q(ks+x-\tau)}e^{\mu_r(-js-t)}e^{(j-k)\mu_{r(j)}s}=e^{(\mu_q-\mu_{r(j)})(ks+x-\tau)}\cdot$$

$$\cdot e^{(\mu_r-\mu_{r(j)})(-js-t)}\cdot e^{\mu_{r(j)}(x-\tau-t)}=O(1)\cdot O(1)\cdot O(e^{2\varrho_\alpha R}).$$

Thus the whole product iv) is $O(\min\{1,|\mu_\alpha|t\}e^{2\varrho_\alpha R}Q(\mu_\alpha,s))$.

[2] $x+t\le\tau\le x+2s$ or $x+(j_0-1)s\le\tau\le x+j_0s$, $3\le j_0\le m+1$.

If $x+t\le\tau\le x+2s$ then in the determinants defining D we substitute $e^{\mu_\alpha w_p(t+x-\tau)}$ by zero: if $x+|j_0-1|s\le\tau\le x+j_0s$ then we substitute $e^{\mu_\alpha w_p(t+x-\tau)}$, $e^{\mu_\alpha w_p(2s+x-\tau)}$, ..., $e^{\mu_\alpha w_p((j_0-1)s+x-\tau)}$ by zeros. We deal only with the products which are new with respect to the case 1).

i) $e^{\mu_q(ks+x-\tau)}2(\mathrm{ch}\mu_q t-\mathrm{ch}\mu_\alpha t)\prod\limits_{\substack{j\neq k\\ r(j)\neq q}} e^{j\mu_{r(j)}s}$.

If $q\le m-1$, then $e^{\mu_q(ks+x-\tau)}(\mathrm{ch}\mu_q t-\mathrm{ch}\mu_\alpha t)=O(e^{\mu_q ks}\min\{1,|\mu_\alpha|t\})$, hence the whole product is $O(\min\{1,|\mu_\alpha|t\}Q(\mu_\alpha,s))$. If $q\ge m+1$ then there exists $j\le -2$ with $r(j)\le m$. Then

$$e^{\mu_q(ks+x-\tau)}e^{j\mu_{r(j)}s}(\mathrm{ch}\mu_q t-\mathrm{ch}\mu_\alpha t)=$$
$$=O(\min\{1,|\mu_\alpha|t\}e^{\mu_q(ks+x-\tau-t)}e^{j\mu_{r(j)}s})=$$
$$=O(\min\{1,|\mu|t\}e^{j\mu_q s}e^{k\mu_{r(j)}s}\cdot e^{(m+2)\varrho_\alpha s}),$$

hence the whole product is $O(\min\{1,|\mu_\alpha|t\}Q(\mu_\alpha,s)\cdot e^{(m+2)\varrho_\alpha s})$.

ii) $e^{\mu_q(ks+x-\tau)}e^{\ell\mu_q s}2(\mathrm{ch}\mu_r t-\mathrm{ch}\mu_\alpha t)\cdot\prod\limits_{\substack{j\neq k,\ell\\ r(j)\neq q,r}} e^{j\mu_{r(j)}s}$ where $2\le\ell<j_0\le$ $\le k$. Let $r\le m-1$. Then in case $\mathrm{Re}\mu_q\ge 0$ resp. $\mathrm{Re}\mu_q\le 0$ we have

$$e^{\mu_q((k+\ell)s+x-\tau)}(\mathrm{ch}\mu_r t-\mathrm{ch}\mu_\alpha t)=O(\min\{1,|\mu_\alpha|t\}e^{k\mu_q s}e^{\ell\mu_r s}),$$

resp. $O(\min\{1,|\mu_\alpha|t\}\cdot e^{k\mu_r s}e^{\ell\mu_q s})$ (here we used $\ell s+x-\tau\le 0$, $ks+x-\tau\ge 0$). If $r\ge m+1$ then there exists $j\le -2$ with $r(j)\le m$. Hence if $\mathrm{Re}\mu_q\ge 0$ resp. $\mathrm{Re}\mu_q\le 0$, we have

$$e^{\mu_q((k+\ell)s+x-\tau)}(\mathrm{ch}\mu_r t-\mathrm{ch}\mu_\alpha t)e^{j\mu_{r(j)}s}=$$
$$=O(\min\{1,|\mu_\alpha|t\}\cdot e^{k\mu_q s}\cdot e^{\ell\mu_{r(j)}s}e^{j\mu_r s}e^{n\varrho_\alpha s}),$$

resp. $$O(\min\{1,|\mu_\alpha|t\}\cdot e^{\ell\mu_q s+k\mu_{r(j)}s+j\mu_r s}\cdot e^{n\varrho_\alpha s}),$$

so the whole product is $O(\min\{1,|\mu_\alpha|t\}e^{n\varrho_\alpha s}Q(\mu_\alpha,s))$. The estimate of D is completely proved for $x\le\tau$. The case $x\ge\tau$ can be dealt with similarly, so the proof of (15′) is complete.

From (25) we obtain

$$|u_\alpha(x+t)+u_\alpha(x-t)-2u_\alpha(x)\mathrm{ch}\mu_\alpha t|\cdot|d(\mu_\alpha,s)|\le$$
$$\le\sum_{\substack{-m\le k\le m+1\\ |k|\ge 2}}|d_k(\mu_\alpha,s,t)|\cdot|u_\alpha(x+ks)|+$$

$$+\int_{x-ms}^{x+(m+1)s} |D(\mu_\alpha, s, t, x-\tau)| \cdot |Q(\tau)| d\tau.$$

Using (8′), (9′), (15′) we get

$$|u_\alpha(x+t)+u_\alpha(x-t)-2u_\alpha(x)\mathrm{ch}\mu_\alpha t| \cdot \frac{|d(\mu_\alpha, s)|}{|Q(\mu_\alpha, s)|} \leq \tag{16′}$$

$$\leq c(e^{-\mathrm{Re}2\mu_{m-1,\alpha}s} + e^{\mathrm{Re}2\mu_{m+1,\alpha}s}) \cdot |\mu_\alpha| t \cdot \|u_\alpha\|_{L^\infty(K_{(2m+2)R})} +$$

$$+c|\mu_\alpha|^{2-n} t \cdot e^{|\varrho_\alpha|s} \cdot \int_{x-ms}^{x+(m+1)s} |Q(\tau)| d\tau, \quad \text{if } 0 \leq t \leq \frac{1}{|\mu_\alpha|}$$

and

$$|u_\alpha(x+t)+u_\alpha(x_t)-2u_\alpha(x)\mathrm{ch}\mu_\alpha t| \cdot \frac{|d(\mu_\alpha, s)|}{|Q(\mu_\alpha, s)|} \leq \tag{17′}$$

$$\leq c(e^{\mathrm{Re}\mu_{m-1,\alpha}(t-2s)} + e^{|\mathrm{Re}\mu_{m+1,\alpha}| \cdot (t-2s)}) \cdot \|u_\alpha\|_{L^\infty(K_{(2m+2)R})} +$$

$$+c|\mu_\alpha|^{1-n} \cdot e^{|\varrho_\alpha|s} \cdot \int_{x-ms}^{x+(m+1)s} |Q(\tau)| d\tau, \quad \text{if } t > \frac{1}{|\mu_\alpha|}.$$

If $|\mu_\alpha| > \max\left\{1, \frac{1}{R}\right\}$ then we obtain from (16′) and (17′)

$$\frac{|d(\mu_\alpha, s)|}{|Q(\mu_\alpha, s)|} \int_0^R \left| \frac{u_\alpha(x+t)+u_\alpha(x-t)-2u_\alpha(x)\mathrm{ch}\mu_\alpha t}{t} \right| dt \leq$$

$$\leq c(e^{-\mathrm{Re}2\mu_{m-1,\alpha}s} + e^{\mathrm{Re}2\mu_{m+1,\alpha}s}) \cdot \|u_\alpha\|_{L_\infty(K_{(2m+2)R})} +$$

$$+c|\mu_\alpha|^{1-n} e^{|\varrho_\alpha|s} \int_{x-ms}^{x+(m+1)s} |Q(\tau)| d\tau +$$

$$+c|\mu_\alpha| \left(\frac{e^{\mathrm{Re}\mu_{m-1,\alpha}(R-2s)}}{\mathrm{Re}\mu_{m-1,\alpha}} + \frac{e^{\mathrm{Re}\mu_{m+1,\alpha}(R-2s)}}{|\mathrm{Re}\mu_{m+1,\alpha}|} \right) \cdot \|u_\alpha\|_{L^\infty(K_{(2m+2)R})} +$$

$$+c|\mu_\alpha|^{1-n} \log|\mu_\alpha| e^{|\varrho_\alpha|s} \int_{x-ms}^{x+(m+1)s} |Q(\tau)| d\tau \leq$$

$$\leq c\left(e^{\mathrm{Re}\mu_{m-1,\alpha}(R-2s)}+e^{|\mathrm{Re}\mu_{m+1,\alpha}|\cdot(R-2s)}\right)\cdot\|u_\alpha\|_{L^\infty(K(2m+2)R)}+$$

$$+c|\mu_\alpha|^{1-n}(1+\log|\mu_\alpha|)e^{|\varrho_\alpha|s}\int\limits_{x-ms}^{x+(m+1)s}|Q(\tau)|d\tau,$$

where we used $\mathrm{Re}\mu_{m-1,\alpha}\geq c_4|\mu_\alpha|$ and $\mathrm{Re}\mu_{m+1,\alpha}\leq -c_4|\mu_\alpha|$ ($c_4>0$ is a constant).

Using (10′) we have

$$(18')\qquad \int\limits_0^R\left|\frac{u_\alpha(x+t)+u_\alpha(x-t)-2u_\alpha(x)\mathrm{ch}\mu_\alpha t}{t}\right|dt\leq$$

$$\leq c\left(e^{-\mathrm{Re}\mu_{m-1,\alpha}R}+e^{\mathrm{Re}\mu_{m+1,\alpha}R}\right)\cdot\|u_\alpha\|_{L^\infty(K_{(2m+2)R})}+$$

$$+c|\mu_\alpha|^{1-n}\log|\mu_\alpha|e^{2|\varrho_\alpha|R}\int\limits_{x-2mR}^{x+2(m+1)R}|Q(\tau)|d\tau$$

if $|\mu_\alpha|\geq A(R_0)\geq 2$.

Using (20) we have

$$(21')\qquad \int\limits_{x-2mR}^{x+2(m+1)R}|Q(\tau)|d\tau\leq c(1+|\mu_\alpha|)^{n-3}\|u_\alpha\|_{L^\infty(K_{4mR})}.$$

Using (21′) we obtain from (18′)

$$(22')\qquad \int\limits_0^R\left|\frac{u_\alpha(x+t)+u_\alpha(x-t)-2u_\alpha(x)\mathrm{ch}\mu_\alpha t}{t}\right|dt\leq$$

$$\leq c\left(e^{-\mathrm{Re}\mu_{m-1,\alpha}R}+e^{\mathrm{Re}\mu_{m+1,\alpha}R}\right)\cdot\|u_\alpha\|_{L(K_{4mR})}+$$

$$+c\frac{\log|\mu_\alpha|}{|\mu_\alpha|^2}\|u_\alpha\|_{L^\infty(K_{4mR})}e^{2|\varrho_\alpha|R},$$

if $|\mu_\alpha|\geq A(R_0)\geq 2$.

We recall that $R_0\geq R\geq\frac{R_0}{2}>0$. If we choose $R_0>0$ such that $2R_0<\varepsilon_0$ then according to (12′), (13), (19), (22′) we have for $|\mu_\alpha|\geq A(R_0)\geq 2$

$$(23')\qquad |\langle u_\alpha,W\rangle-\delta(\nu,\nu_\alpha)u_\alpha(x)|\leq c\Big(\frac{1}{(1+(\nu-\nu_\alpha)^2}+$$

$$+e^{-\mathrm{Re}\mu_{m-1,\alpha}\frac{R_0}{2}}+e^{\mathrm{Re}\mu_{m+1,\alpha}\frac{R_0}{2}}+\frac{\log|\mu_\alpha|}{|\mu_\alpha|^2}\Big)\cdot\|u_\alpha\|_{L^2(K)}.$$

Now we prove (a)⇒(b).

From (23′), (24), (3) follows

$$\sum_{\alpha=1}^{\infty}|\langle u_\alpha,W\rangle-\delta(\nu,\nu_\alpha)u_\alpha(x)|\cdot\|v_\alpha\|_{L^2(G)}\le$$

$$\le c\sum_{\alpha=1}^{\infty}\frac{|\langle u_\alpha,W\rangle-\delta(\nu,\nu_\alpha)u_\alpha(x)|}{\|u_\alpha\|_{L^2(K')}}\le$$

$$\le c\sum_{\alpha=1}^{\infty}\left(\frac{1}{(1+(\nu-\nu_\alpha)^2}+e^{-c_4|\mu_\alpha|\frac{R_0}{2}}+\frac{\log|\mu_\alpha|}{|\mu_\alpha|^2}\right)\le c_5,$$

where c_5 is a constant independent of ν.

Introduce the function $F(x,y)$ in the same way as we did in the case a).

Following the method applied for n even we obtain the statement (a)⟹(b). Hence (a)⟹(b) is proved completely.

(b)⟹(c): It follows from (b) that

$$\lim_{\nu\to\infty}\|S_\nu(f)-\sigma_\nu(f)\|_{L^2(K)=0}.$$

On the other hand it is a classical result that

$$\lim_{\nu\to\infty}\|f-S_\nu(f)\|_{L^2(K)}=0.$$

The statement is proved.

(c)⟹(a): For every $f\in L^2(G)$ $\|\sigma_\nu(f)\|_{L^2(K)}$, $\nu=1,\dots$ is bounded.

Taking into account that $\sum^*$ can be arbitrary, we get from here that for all α,

$$\|\langle f,v_\alpha\rangle u_\alpha(\cdot)\|_{L_2(K)}\le c\|f\|_{L_2(G)},\quad f\in L_2(G).$$

Since $\sup\limits_{f\in L_2(G)}\frac{|\langle f,v_\alpha\rangle|}{\|f\|_{L_2(G)}}=\|v_\alpha\|_{L_2(G)}$, we obtain (a).

The proof of the Theorem is complete.

References

[1] M. Horváth, I. Joó and V. Komornik, An equiconvergence theorem, *Annales Univ. Sci. Budapest. Sect. Math.*, **31** (1988), 19–26.

[2] V. KOMORNIK, On the equiconvergence of eigenfunction expansions associated with ordinary linear differential operators, *Acta Math. Acad. Sci. Hungar.*, **47** (1986), 261–280.

[3] I. JOÓ, On the divergence of eigenfunction expansions, *Annales Univ. Sci. Budapest., Sect. Math.*, **32** (1988), 3–36.

[4] M. B. TAHIR, An equiconvergence theorem, *Studia Sci. Math. Hungar.* (to appear).

[5] S. TIMOSHENKO, D. H. YOUNG, W. WEAVER, *Vibration problems in engineering*, John Wiley & Sons 1974.

[6] V. KOMORNIK, Upper estimates for the eigenfunctions of higher order of a linear differential operator, *Acta Sci. Math.*, **45** (1983), 261–271.

[7] A. ZYGMUND, *Trigonometrical series*, Cambridge Univ. Press, (New York, 1959).

[8] V. A. IL'IN, The necessary and sufficient condition of basis property and that of equiconvergence with trigonometric series I, II (in Russian), *Differential Equations*, **16** (1980), 771–794, 980–1009.

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 129–131

SHORT PROOF TO A STRONG PARROTT THEOREM

By

ZOLTÁN SEBESTYÉN

Department of Applied Analysis, Eötvös Loránd University, Budapest

(Received September 21, 1992)

THEOREM 1. (Parrott's theorem). *Let H, K be Hilbert spaces, $H_2 \subset \subset H$, $K_2 \subset K$ be closed subspaces. Then if the operators $T_2 : H_2 \to K$, $T_2' : H \to K_2$ are any contractions satisfying*

$$P_{K_2} T_2 = T_2' | H_2, \tag{1}$$

there exists a contraction operator $T : H \to K$ such that

$$T_2 = T | H_2 \quad \text{and} \quad T_2' = P_{K_2} T. \tag{2}$$

THEOREM 2. (Strong Parrott theorem). *Let H, K be Hilbert spaces, H_0, $H_1 \subset H$, $K_0, K_1 \subset K$ be closed subspaces. Let $T_0 : H_0 \to K$, $T_0' : H \to K_0$ be any contractions satisfying*

$$P_{K_0} T_0 = T_0' | H_0. \tag{3}$$

Suppose moreover that $V : H_1 \to K_1$ is unitary operator and satisfies two following identities

$$P_{K_0} V = T_0' | H_1, \tag{4}$$

$$P_{K_1} T_0 = V P_{H_1} | H_0. \tag{5}$$

Then there exists a contraction $T : H \to K$ such that

$$T_0 = T | H_0 \quad \text{and} \quad T_0' = P_{K_0} T, \tag{6}$$

$$V = T | H_1 \quad \text{and} \quad V^* = T^* | K_1. \tag{7}$$

PROOF. (reduction to Theorem 1). Let $H_2 : \overline{H_0 + H_1}$, $K_2 := \overline{K_0 + K_1}$ and $T_2 : H_2 \to K$, $T_2' : H \to K_2$ be operators such that T_2 operates on $H_0 + H_1 = \{h_0 + h_1 : h_0 \in H_0,\ h_1 \in H_1\}$ as follows

$$T_2(h_0 + h_1) := T_0 h_0 + V h_1,$$

the adjoint of T_2', R_2 is defined on

$$K_0 + K_1 = \{k_0 + k_1 : k_0 \in K_0,\ k_1 \in K_1\}$$

by $R_2(k_0 + k_1) := (T_0')^* k_0 + V^* k_1$.
What we have to check is first that these are well-defined contractions, second that T_2 and $T_2' = R_2^*$ satisfy identity (1).

Theorem 1 implies that (6) and (7) follow from (2) since $T_2 = T^*|K_2$ is just included in (2).
The first claim follows by the following estimation

$$\begin{aligned}\|T_0h_0+Vh_1\|^2 &= \|T_0h_0\|^2+(P_{K_1}T_0h_0,Vh_1)+(Vh_1,P_{K_1}T_0h_0)+\|Vh_1\|^2 = \\ &\overset{(5)}{=} \|T_0h_0\|^2+(VP_{H_1}h_0,Vh_1)+(Vh_1,VP_{H_1}h_0)+\|Vh_1\|^2 = \\ &= \|T_0h_0\|^2+(P_{H_1}h_0,h_1)+(h_1,P_{H_1}h_0)+\|h_1\|^2 \le \\ &\le \|h_0\|^2+(h_0,h_1)+(h_1,h_0)+\|h_1\|^2 = \|h_0+h_1\|^2.\end{aligned}$$

R_2 is well-defined contraction by a similar (dual) reason. Finally to check identity (1) is enough to prove that

$$(8) \qquad (T_2(h_0+h_1), k_0+k_1) = (h_0+h_1, R_2(k_0+k_1)).$$

The left-hand side in (8) term to term use of (3), (5) and (4), is the very same as the opposite:

$$\begin{aligned}&(P_{K_0}T_0h_0,k_0)+(P_{K_1}T_0h_0,k_1)+(P_{K_0}Vh_1,k_0)+(Vh_1,k_1) = \\ &= (T_0'h_0,k_0)+(VP_{H_1}h_0,k_1)+(T_0'h_1,k_0)+(Vh_1,k_1) = \\ &= (h_0+h_1,(T_0')^*k_0+V^*k_1).\end{aligned}$$

REMARK. A recent construction for a proof of Parrott's theorem is given [6] as an application of (smallest) positive and self-adjoint operator extensions [3], [4], ([5]). As a result this process yields [7], using middle self-adjoint extension, a compact operator T in Theorem 1 whenever T_2 and T_2' are both compact.

References

[1] C. FOIAS, A. TANNENBAUM, A strong Parrott theorem, *Proc. Amer. Math. Soc.*, **106** (1989), 777–784.

[2] S. PARROTT, On the quotient norm and the Sz. Nagy–Foias lifting theorem, *J. Functional Analysis*, **30** (1978), 311–328.

[3] Z. SEBESTYÉN, Restrictions of positive operators, *Acta Sci. Math. (Szeged)*, **46** (1983), 299–301.

[4] Z. SEBESTYÉN, On the existence of certain semibounded self-adjoint operators in Hilbert space, *Acta Math Hung.*, **47** (1986), 29–32.

[5] Z. SEBESTYÉN, L. KAPOS, Extremal positive and self-adjoint extensions of suboperators, *Periodica Math. Hung.*, **20** (1989), 75–80.

[6] Z. SEBESTYÉN, A proof of Parrott's theorem on quotient norms, *Periodica Math. Hung.*, **20** (1989), 85–87.

[7] Z. SEBESTYÉN, Á. MAGYAR, Compact self-adjoint extension, *Lin. Alg. and Its Appl.*, (submitted 1989).

ANNALES UNIV. SCI. BUDAPEST., 36 (1993), 133–137

A CLASS OF TWO DIMENSIONAL VECTOR FIELDS WHICH ARE FOKKER–PLANCK INTEGRABLE

By

ANGEL DOMINGO RUEDA

Faculty of Sciences, University of Zulia, Maracaibo, Venezuela(*)

(Received September 29, 1992)

1. Introduction

Besides some problems where fluctuations are considered and Fokker–Planck equations naturally appear [4] we may be interested in a kind of quantitative analysis of the dynamics of the flow associated to a given vector field $\vec{V}$ by the system $\dot{x} = \vec{V}(x)$. We may be interested in the intensity of its attractors and whether this intensity does not vary much under perturbations which occur in many problems derived from mechanics, population dynamics etc. [4], [5]. It is useful to know the Zeeman Stability of a vector field because this stability means the persistence of the intensity of the attractors of its flow. This way we arrive at a Fokker–Planck equation with ε-diffusion whose steady state (sufficient conditions for existence and uniqueness are given in [6]) is an ε-smoothing of a measure invariant with respect to the flow associated to $\vec{V}$, with support over the attractors of this flow. Then the existence of a neighbourhood $B(\vec{V})$ for which every vector field $\vec{W} \in B(\vec{V})$ has a steady state diffeomorphic to the steady state of $\vec{V}$, means the persistence of the intensity of attraction and therefore a kind of robustness of the domain of attraction, instead of the persistence of the complete topological structure as in case of structural stability in the Andronov–Pontriagin sense. In [6] the case of a gradient vector field $\vec{V} = -\text{grad}(\Phi)$ is studied by the Morse stability of its steady state $u(x,y) = N \cdot \exp(-\Phi(x,y)/\varepsilon)$ where $\int u$ may be chosen equal to 1 if certain conditions on $\vec{V}$ or compacity over its domain hold. On the other hand, because of the monotonicity of the exponential function we can reduce the problem by Lemma 1 of [6] to the Morse stability of Φ. If the vector field is

(*) Present address: Department of Mathematics, Faculty of Mechanical Engineering, Budapest, 1521 not a gradient, in other words, if the potential conditions are not fulfilled the situation is usually more complicated. A class of vector fields will be given such that the steady state of the associated Fokker–Planck equation can be computed. We note that this class of vector fields contains some interesting models of biology.

2. The Zeeman stability

Given a vector field $\vec{V}$ over a compact manifold M and given an $\varepsilon > 0$ our Fokker–Planck equation (FPE) with ε-diffusion, $\varepsilon > 0$, associated, to $\vec{V}$ is defined as

$$\frac{\partial u}{\partial t} = \varepsilon \Delta u - \operatorname{Div}(u, \vec{V}).$$

We say that $\vec{V}$ is ε-Zeeman stable if for every vector field belonging to some C^{∞}-neighbourhood of $\vec{V}$ the steady states of the respective FPE associated are diffeomorphic. We say that $\vec{V}$ is Zeeman stable if it is ε-Zeeman stable for every small $\varepsilon > 0$. finally, we say that the system $\dot{x} = \vec{V}(x)$ is Zeeman stable iff $\vec{V}$ is.

We note that a system of ordinary differential equations which has only one finite non-degenerate equilibrium point and has a repeller at infinity is Zeeman-stable. This follows from the existence Theorem 3 of [6] and from the characterization of stable Morse functions of [1]. Clearly, these statements are structurally stable in the Andronov–Pontriagin sense if the equilibrium point is hyperbolic. This occurs in the following case: $\dot{x} = -x - ax^2y$, $\dot{y} = -y$ where $a > 0$ which represents two species in the process of extinction, the second of them killing of the first one.

3. A class of vector fields

Our main result is the following Theorem:

THEOREM 1. *Let $\vec{P} = (p_1, p_2)$ be a two dimensional vector field such that $\operatorname{Div}(\vec{P}) = 0$. Let Φ be a first integral of $\vec{P}$ and consider the vector field*

$$\vec{V} = (\beta_x + \alpha p_1,\ \beta_y + \alpha p_2)$$

where β, $\alpha \in C^1(\mathbb{R}^2)$. If $\alpha(x,y) = w(\Phi(x,y))$ and $\beta(x,y) = s(\Phi(x,y))$ for arbitrary functions s, $w \in C^1$ then

$$\frac{\partial u}{\partial t} = \varepsilon \Delta u - \operatorname{Div}(u.\vec{V})$$

has a steady state $u(x,y) = N\exp(\Phi(x,y)/\varepsilon)$.

PROOF. (For simplicity in the proof we assume $\beta(\Phi(x,y)) = \Phi(x,y)$). We can put $\vec{V} = \vec{G} + \vec{R}$ with $G = \text{grad}\Phi$, $\vec{R} = (r_1, r_2)$ and $r_1 = w(\Phi(x,y)).p_1(x,y)$, $r_2 = w(\Phi(x,y)) \cdot p_2(x,y)$ so that

$$\varepsilon\Delta u - \text{Div}\left\{(\vec{G}+\vec{R})u\right\} = \varepsilon\Delta u - \text{Div}(\vec{G}.u) - \text{Div}(\vec{R}u) =$$
$$= \varepsilon\Delta u - \text{Div}(\vec{G}.u) - u\text{Div}\,\vec{R} - \vec{R}.\text{Grad}(u) = 0$$

is equivalent to $\varepsilon\Delta u - \text{Div}(\vec{G}.u) = 0$ because

$$\text{Div}(\vec{R}) = w'(\Phi(x,y)).\left\{\Phi_x(x,y).p_1(x,y) + \Phi_y(x,y).p_2(x,y)\right\} +$$
$$+ w(\Phi(x,y)).\text{Div}(\vec{P}) = 0$$

and we have $\vec{R} \perp \text{Grad}(u)$.

COROLLARY 1. Let f be a stable Morse function and suppose $\int \exp(f/\varepsilon) < \infty$. Then the system $\dot{x} = f_x - w(f)f_y$, $\dot{y} = f_y + w(f)f_x$ is Zeeman stable for every function $w \in C^1$.

PROOF. It follows from the Morse stability of $u(x,y) = N\exp(f/\varepsilon)$.

4. Examples

a) In the system $\dot{x} = -x + y(x^2+y^2)$, $\dot{y} = -y - x(x^2+y^2)$ we can take $\Phi(x,y) = -(x^2+y^2)/2$ and $w(r) = r$. Then the steady state of the corresponding Fokker–Planck equation with ε-diffution is $u(x,y) = N\exp(-$ $-(x^2+y^2)/2\varepsilon)$. This system is structurally stable and Zeeman stable.

b) Consider the system $\dot{x} = -2x(1+x^2)y^2$, $\dot{y} = 2y(y^2-1)x^2$. It has a steady state for its FPE with ε-diffusion $u(x,y) = N\exp(-x^2y^2/\varepsilon)$. Its critical points cover the coordinate axes $x = 0$ and $y = 0$, hence they are degenerate, thus, they are neither structurally nor Zeeman stable. The following examples have applications in population dynamics.

c) The Kolmogorov system $\dot{x} = x(1-y^2)$, $\dot{y} = -y(x^2-1)$ has only one rest point which is hyperbolic, so that it is structurally stable but isn't Zeeman stable because the steady state attached is $u(x,y) = N \cdot$ $\cdot \exp(-x^2y^2/2\varepsilon)$.

d) Consider $\dot{x} = x(y^2 - (1+x^2y^4))$, $\dot{y} = y(x^2+1-x^4y^3)$. The corresponding steady state $u(x,y) = N\exp(-(x^2y^2-1)^2/4\varepsilon)$ is not a Morse function, thus the system is not Zeeman stable but structurally stable.

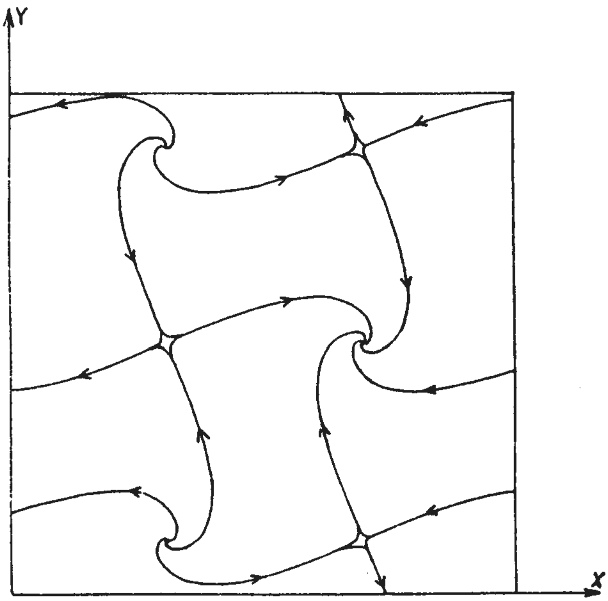

Fig. a)

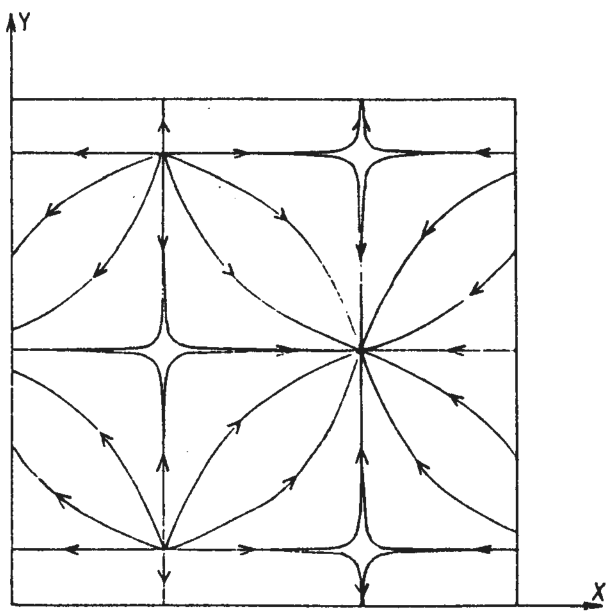

Fig. b)

e) The following system which is defined on the torus: $\dot{x} = \cos(x) + \sin(y)$, $\dot{y} = -\sin(y) + \cos(x)$ is structurally stable but not Zeeman stable because the steady state has the same critical value in the two saddle points. In Figures a) and b) the phase portraits are showed for the original system and for the system associated to $\mathrm{Grad}(\exp([\sin(x) + \cos(y)]/\varepsilon))$.

References

[1] GOLUBITSKY, M. and GUILLEMIN, V., *Stable mappings and their singularities*, Berlin, Springer, 1973.

[2] HIRSCH, MORRIS W., Systems of differential equations which are competitive or cooperative. I: Limit sets, *SIAM J. Math. Anal.*, **13** (1982), 167–179.

[3] PEIXOTO, M. M. Structural stability in two-dimensional manifolds, *Topology* **1** (1962), 102–120.

[4] RISKEN, H., *The Fokker–Planck equation*, Berlin, Springer, 1984.

[5] RUEDA, A. D., *Proceeding of Equadiff 91*, Universitat autonoma de Barcelona, Spain (1991).

[6] ZEEMAN, E. C., Stability of Dynamical Systems, *Nonlinearity*, **1** (1988), 115–155.

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 139–146

ON THE TIME-MONOTONICITY OF THE SOLUTIONS OF LINEAR SECOND ORDER HOMOGENEOUS PARABOLIC EQUATIONS

By

T. PFEIL

Department of Applied Analysis, Eötvös Loránd University, Budapest

(Received October 9, 1992)

Let $n \in \mathbb{N}^+$, $\Omega \subset \mathbb{R}^n$ be a bounded domain with a smooth boundary and L the following symmetric second order linear differential operator

$$Lu := \sum_{i,j=1}^{n} \partial_i(a_{ij}\partial_j u) + du,$$

where $a_{ij} \in C^1(\overline{\Omega})$, $a_{ij} = a_{ji}$, $i, j = 1,\ldots,n$; $d \in C(\overline{\Omega})$, $d \leq 0$ and suppose that L is uniformly elliptic in Ω. Let $Q := (0,+\infty) \times \Omega$, $\Gamma := [0,+\infty) \times \partial\Omega$, $\Omega_0 := \{0\} \times \Omega$.

Let us denote the eigenvalues of the homogeneous eigenvalue problem with homogeneous Dirichlet boundary condition

$$\text{(1)} \qquad \begin{cases} Lw + \lambda w = 0 & \text{in } \Omega \\ w\Big|_{\partial\Omega} = 0 \end{cases}$$

by λ_k, $k \in \mathbb{N}^+$ (numbered in such a way that they form a monotone nondecreasing sequence) and the corresponding eigenfunctions by w_k. λ_1 is called the principal eigenvalue of (1).

Consider the following generalized initial-boundary value problem

$$\text{(2)} \qquad \begin{cases} \partial_0 u - Lu = 0 & \text{in } Q \\ u\Big|_{\Gamma} = 0 \\ u\Big|_{\Omega_0} = \tilde{\varphi} \end{cases}$$

where $\tilde{\varphi}(0,x) := \varphi(x)$ for $x \in \Omega$ with a given $\varphi \in L^2(\Omega)$. Let ξ_k, $k \in \mathbb{N}^+$ be defined by

$$\varphi = \sum_{k=1}^{\infty} \xi_k w_k \quad \text{in } \Omega.$$

Under the previous assumptions there exists a unique solution $u \in H^{0,1}(Q)$ of (2) which can be written in the form of a series (it is convergent in the norm of $H^{0,1}(Q)$):

$$u(t,x) = \sum_{k=1}^{\infty} \xi_k e^{-\lambda_k t} w_k(x) \quad \text{for} \quad (t,x) \in Q. \tag{3}$$

Let $l := 1 + \left[\frac{n}{2}\right]$, and suppose that Ω has an l times continuously differentiable boundary if $n > 1$; if $n = 1$ then let $l := 2$ and Ω be a finite open interval, $a_{ij} \in C^{l-1}(\overline{\Omega})$, $d \in C^{l-2}(\overline{\Omega})$. In this case the eigenfunctions of L with the homogeneous Dirichlet boundary condition are continuous in $\overline{\Omega}$.

Moreover, the solution of (2) given by the series (3) is C^1 in t due to the fact that there exist positive numbers C_1 and C_2 such that

$$C_1 k^{2/n} \leq \lambda_k \leq C_2 k^{2/n} \quad \text{for} \quad k \in \mathbb{N}^+;$$

and

$$\partial_0 u(t,x) = -\sum_{k=1}^{\infty} \xi_k \lambda_k e^{-\lambda_k t} w_k(x) \quad \text{for} \quad (t,x) \in Q \tag{4}$$

(for these results see e.g. SIMON and BADERKO [8]).

Results of NARASIMHAN [7] and FRIEDMAN [4] show that the solution of the associated classical initial-boundary value problem with homogeneous Dirichlet boundary condition tends to zero uniformly in Ω as t tends to infinity.

We will examine the monotonicity with resepct to time of the solution of (2).

Under the given assumptions the principal eigenvalue of L is simple and the corresponding eigenfunction w_1 can be chosen as a positive function in Ω (see KREIN and RUTMAN [6] and CHICCO [1]).

THEOREM 1. *Let u be the (unique) solution of the initial-boundary value problem (2) with all the conditions given previously. Suppose that the first Fourier coefficient of φ is not equal to zero. Then for every $x \in \Omega$ there exists a positive number T such that the function $t \mapsto u(t,x)$, $t > T$ is strictly decreasing if $\xi_1 > 0$, and strictly increasing if $\xi_1 < 0$.*

Furthermore, if $K \subset \Omega$ is compact then for $x \in K$ T can be chosen independently of x.

PROOF. Using (4) we have

$$\partial_0 u(t,x) = -e^{-\lambda_1 t} w_1(x) \left(\xi_1 \lambda_1 + e^{-(\lambda_2-\lambda_1)t} \sum_{k=2}^{\infty} \xi_k \lambda_k e^{-(\lambda_k-\lambda_2)t} \cdot \frac{w_k(x)}{w_1(x)} \right)$$

$$\text{for} \quad (t,x) \in Q. \tag{5}$$

According to Theorem 8.15 in GILBARG and TRUDINGER [5] the weak solution of the elliptic boundary value problem

$$\begin{cases} Lw = 0 & \text{in } \Omega \\ w\Big|_{\partial\Omega} = 0 \end{cases}$$

is essentially bounded under our assumptions without the sign condition for d. The bound only depends on the dimension of the space, the volume of Ω, the uniform ellipticity constant of L and the essential suprema of the coefficients of L. Applying the proof of the above mentioned theorem to the eigenvalue problem (1) with $d+\lambda$ instead of d, it is easy to see that the constant C in formula (8.36) in [5] is $O(k^{1/2})$. Hence, following the proof of formula (8.34) in the proof of Theorem 8.15 in [5], we actually prove the existence of positive numbers M, $s \in \mathbb{R}^+$ such that

$$\operatorname*{ess\,sup}_{\Omega} |w_k| \le M k^s, \quad k \in \mathbb{N}^+ \tag{6}$$

(for the detailed proof see Theorem 3 in the Appendix).

Let $x \in \Omega$ be fixed. The series in (5) can be estimated as follows:

$$\left| \sum_{k=2}^{\infty} \xi_k \lambda_k e^{-(\lambda_k-\lambda_2)t} \cdot \frac{w_k(x)}{w_1(x)} \right| \le \frac{M_1}{w_1(x)} \sum_{k=2}^{\infty} |\xi_k| |\lambda_k|^{v+1} e^{-(\lambda_k-\lambda_2)t} \tag{7}$$

with $v := \frac{n}{2}s$ and $M_1 := M/c_1^v$. Estimating the function $x \mapsto x^{v+1} e^{-\frac{1}{2}xt}$ by its maximum on $\mathbb{R}^+$, we obtain

$$|\lambda_k|^{v+1} e^{-\frac{1}{2}\lambda_k t} \le \left(\frac{2v+2}{et} \right)^{v+1} \tag{8}$$

Furthermore, there exist a positive integer k_0 and a positive real number τ such that for every $k > k_0$ and for all $t \ge \tau$

$$\lambda_k - 2\lambda_2 > 0 \quad \text{and} \quad t > 2\frac{\log k}{\lambda_k - 2\lambda_2}. \tag{9}$$

Thus estimating the first $k_0 - 1$ terms on the right-hand side of (7) by using $\lambda_k - \lambda_2 > 0$ and the remaining part by employing

$$|\lambda_k|^{v+1} e^{-(\lambda_k-\lambda_2)t} = |\lambda_k|^{v+1} e^{-\frac{1}{2}\lambda_k t} e^{-\frac{1}{2}(\lambda_k - 2\lambda_2)t}$$

and inequalities (8) and (9), we obtain for $t \geq \tau$ and $x \in \Omega$

$$(10)\qquad \left|\sum_{k=2}^{\infty} \xi_k \lambda_k e^{-(\lambda_k-\lambda_2)t} \cdot \frac{w_k(x)}{w_1(x)}\right| \leq$$

$$\leq \frac{M_1}{w_1(x)} \left(\sum_{k=2}^{k_0} |\xi_k||\lambda_k|^{v+1} + \left(\frac{2v+2}{e\tau}\right)^{v+1} \sum_{k=k_0+1}^{\infty} |\xi_k| \frac{1}{k} \right).$$

The term in the right-hand side is independent of t thus the second term in the bracket in (5) tends to zero, hence for t large enough

$$\operatorname{sign}\{\partial_0 u(t,x)\} = \operatorname{sign}\{-\xi_1 \lambda_1\}.$$

If $K \subset \Omega$ is compact then let $\delta := \min\{w_1(x) : x \in K\}$. Replacing $w_1(x)$ with δ in the right hand-side of (10) shows that there exists an appropriate T independent of x.

Theorem 1 is proved.

COROLLARY 1. *Suppose that the conditions of Theorem 1 are fulfilled and suppose in addition that there exist positive numbers N and r such that*

$$(11)\qquad \left|\frac{w_k(x)}{w_1(x)}\right| \leq N k^r, \qquad k \in \mathbb{N}^+$$

(i.e. $\left|\frac{w_k(x)}{w_1(x)}\right| \leq N^* \lambda_k^{r^*}$ for some N^*, $r^* \in \mathbb{R}^+$).

Then there exists an appropriate T independent of x, only depending on Ω.

PROOF. The series in (5) can be estimated similarly to (7):

$$\left|\sum_{k=2}^{\infty} \xi_k \lambda_k e^{-(\lambda_k-\lambda_2)t} \cdot \frac{w_k(x)}{w_1(x)}\right| \leq N^* \sum_{k=2}^{\infty} |\xi_k||\lambda_k|^{r^*+1} e^{-(\lambda_k-\lambda_2)t},$$

therefore an argument similar to the proof of Theorem 1 leads to a T independent of x.

Corollary 1 is proved.

REMARK 1. Condition (11) and the continuity of the eigenfunctions is true for the heat conduction equation defined in an n dimensional interval, hence the assertion of Corollary 1 is also valid.

REMARK 2. Let us denote the outward normal direction by ν. Suppose that $w_k \in C^2(\overline{\Omega})$ then $\partial_\nu w_1 < 0$ holds in $\partial\Omega$. (The condition is fulfilled if we assume in addition that $\partial\Omega \in C^l$, $a_{ij} \in C^{l-1}(\overline{\Omega})$ and $d \in C^{l-2}(\overline{\Omega})$, where $l := 3 + \left[\frac{n}{2}\right]$ and $l := 4$ if $n = 1$). To ensure (11) we only need an estimate for the essential supremum of $\partial_\nu w_k$ on $\partial\Omega$ similar to (6) for w_k in Ω.

Now assume that the first Fourier coefficient of φ in the initial-boundary value problem (2) is zero. Denote the first non-zero Fourier coefficient by ξ_p, $p > 1$. Further let the zeros of w_p be denoted by Z_p.

Let u be the (unique) solution of problem (2).

THEOREM 2. *For every $x \in \Omega$ there exists a $T > 0$ such that the function $t \mapsto u(t,x)$ is monotone on $[T, +\infty)$.*

Assuming the eigenvalue λ_p corresponding to the eigenfunction w_p to be simple and $x \in \Omega \setminus Z_p$, T can be given such that the function $t \mapsto u(t,x)$ is strictly monotone on $[T, +\infty)$. If $\xi_p w_p(x) > 0$ then the function is decreasing, and it is increasing if $\xi_p w_p(x) < 0$.

Moreover, if λ_p is simple and K is a compact subset of a connected component of $\Omega \setminus Z_p$ then a positive number T can be chosen for the whole K.

PROOF. If $w_p(x) \neq 0$ then using $\partial_0 u(t,x) =$

$$= -e^{-\lambda_p t} w_p(x) \left(\xi_p \lambda_p + e^{-(\lambda_{p+1} - \lambda_p)t} \sum_{k=p+1}^{\infty} \xi_k \lambda_k e^{-(\lambda_k - \lambda_p)t} \cdot \frac{w_k(x)}{w_p(x)} \right)$$

the statement can be proved similarly to Theorem 1.

If either $w_p(x) = 0$ or λ_p is a multiple eigenvalue then the function $t \mapsto u(t,x)$ is constant zero or its monotonicity can be proved similarly, factoring out $e^{-\lambda_q t}$ for an appropriate q.

Theorem 2 is proved.

REMARK 3. $\Omega \setminus Z_p$ only consists of a finite number of connected components, which is the last proposition in CHICCO [1]. It is worth mentioning that Theorem 2 is also true for $p = 1$, so it is a generalization of Theorem 1.

REMARK 4. A whole connected component of $\Omega \setminus Z_p$ cannot be taken in Theorem 2 instead of a compact subset of a component as the following example shows. Consider the one dimensional classical heat conduction equation

$$(12) \qquad \begin{cases} \partial_0 u - \partial_1^2 u = 0 & \text{in } (0, +\infty) \times (0, \pi) \\ u(t, 0) = u(t, \pi) = 0 & \text{for } t \in \mathbb{R}_0^+ \\ u(0, x) = \sin 2x - \sin 3x & \text{for } x \in (0, \pi) \end{cases}$$

The unique solution of (12) is

$$u(t,x)=e^{-4t}\sin 2x-e^{-9t}\sin 3x, \qquad (t,x)\in\mathbb{R}^+\times(0,\pi).$$

According to Theorem 2 for any $\delta>0$ there exists T such that $\partial_0 u(t,x)>0$, if $t>T$ and $x\in\left(\frac{\pi}{2}+\delta,\pi-\delta\right)$. However for every $t\in\mathbb{R}^+$ there exists $x_0\in\left(\frac{\pi}{2},\pi\right)$ such that $\partial_0 u(t,x_0)<0$.

REMARK 5. For the heat conduction equation in the one dimensional case there are estimates for T given for the whole interval (see [3]).

REMARK 6. Several results are known for the monotonicity in time of the numerical solution of the one dimensional heat conduction problem. Certain conditions are given on the choice of the discretization parameters to ensure the monotonicity in time for the finite difference method in SAMARSKII [9], for the finite element method in FARAGÓ [2].

Appendix

Here we give a possible proof for formula (6), i.e. an upper bound for the essential supremum of the eigenfunctions w_k of (1) depending on k. This proof was obtained by supplementing the proof of Theorem 8.15 in [5].

THEOREM 3. *Let L be a uniformly elliptic operator with its principal part in divergence form on a bounded domain Ω and a_{ij}, $d\in L^\infty(q)$. There exist positive numbers M, $s\in\mathbb{R}^+$ such that for the normed eigenfunctions w_k of* (1)

$$\operatorname*{ess\,sup}_{\Omega}|w_k|\le Mk^s, \qquad k\in\mathbb{N}^+ \tag{13}$$

(i.e. $\operatorname*{ess\,sup}_{\Omega}|w_k|\le M^*\lambda_k^{s^*}$ for some M^*, $s^*\in\mathbb{R}^+$).

PROOF. Let q be a real number satisfying $q>n$, further $w_k^+:=$ $=\max\{w_k,0\}$, $w_k^-:=\min\{w_k,0\}$ and $d_k:=d+\lambda_k$.

For $\beta\ge 1$ and $N\in\mathbb{N}^+$ define a function H on $[0,+\infty)$ by

$$H(z):=\begin{cases} z^\beta & \text{if } z\in[0,N]\\ \beta N^{\beta-1}(z-N)+N^\beta & \text{if } z\in(N,+\infty)\end{cases}$$

In the proof of Theorem 8.15 in [5] the existence of a positive constant C_1 is proved fulfilling the following inequality:

$$\|H\circ w_k^+\|_{L^{2\hat{n}/(\hat{n}-2)}(\Omega)}\le C_1\|d_k\|^{1/2}_{L^{q/2}(\Omega)}\|(H'\circ w_k^+)w_k^+\|_{L^{2q/(q-2)}(\Omega)},$$

where $\hat{n} = n$ if $n > 2$ and $\hat{n}$ is between 2 and q for $n = 2$. C_1 depends on n, the volume of the domain and the uniform ellipticity constant.

Using

$$\|d + \lambda_k\|_{L^{q/2}(\Omega)} \leq \|d\|_{L^{q/2}(\Omega)} + \lambda_k \mathrm{vol}(\Omega)^{2/q} \leq C_2 \lambda_k \mathrm{vol}(\Omega)^{2/q} = C_3^2 \lambda_k$$

(C_2 and C_3 are appropriate constants) we have a modification of formula (8.36) in [5]:

$$\|H \circ w_k^+\|_{L^{2\hat{n}/(\hat{n}-2)}(\Omega)} \leq C_3 \lambda_k^{1/2} \|(H' \circ w_k^+) w_k^+\|_{L^{2q/(q-2)}(\Omega)}.$$

Proceeding as in the original proof we obtain a version of formula (8.37) in [5] as $N \to +\infty$:

$$\|w_k^+\|_{L^{\beta\chi q^*}(\Omega)} \leq (C_3 \beta)^{1/\beta} \lambda_k^{1/2\beta} \|w_k^+\|_{L^{\beta q^*}(\Omega)},$$

where $q^* := 2q/(q-2)$ and $\chi := \hat{n}(q-2)/q(\hat{n}-2) > 1$. The iteration of L^p-norms ($\beta = \chi^m$, $m = 0, 1, 2, \ldots$) leads to the estimate

$$\|w_k^+\|_{L^{\chi^M q^*}(\Omega)} \leq \prod_{m=0}^{M} (C_3 \chi^m \lambda_k^{1/2})^{\chi^{-m}} \|w_k^+\|_{L^{q^*}(\Omega)} \leq C_4 \lambda_k^{\sigma/2} \|w_k^+\|_{L^{q^*}(\Omega)},$$

where $\sigma := \sum_{m=0}^{\infty} \chi^{-m}$ and $C_4 := C_3^\sigma \chi^{\sum_{m=0}^{\infty} m\chi^{-m}}$. Therefore by letting $M \to +\infty$ we obtain

$$\operatorname*{ess\,sup}_{\Omega} w_k^+ \leq C_4 \lambda_k^{\sigma/2} \|w_k^+\|_{L^{q^*}(\Omega)}.$$

Further, by using the interpolation inequality (7.10) in [5] we have

$$\operatorname*{ess\,sup}_{\Omega} w_k^+ \leq C_5 \lambda_k^{\sigma(\mu+1)/2} \|w_k^+\|_{L^2(\Omega)},$$

where μ is an appropriate constant only depending on q^* (C_5 is also a constant).

A similar inequality can be proved for w_k^-, so we obtain (13).

References

[1] M. Chicco, Some properties of the first eigenvalue and the first eigenfunction of linear second order elliptic partial differential equations in divergence form, *Boll. Un. Mat. Ital.*, **5** (1972), 245–254.

[2] I. FARAGÓ, Galerkin's semidiscretization method for solving of parabolic problems, *Proc. of AIO '91 'Method of lines and surfaces'*, London, (to appear).

[3] I. FARAGÓ, H. A. HARITON, N. KOMÁROMI, T. PFEIL, Qualitative properties of the heat condition (in Hungarian, accepted for publication by *Alk. Mat. Lapok*).

[4] A. FRIEDMAN, *Partial Differential Equations of Parabolic Type*, Prentice Hall, N. J., 1964.

[5] D. GILBARG, N. S. TRUDINGER, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, Berlin–Heidelberg–New York, 1977.

[6] M. G. KREIN, M. A. RUTMAN, Linear operators leaving invariant a cone (in Russian), *Uspekhi Mat. Nauk* **3** (1948), 3–95. English translation in *Amer. Math. Soc. Transl.*, **10** (1962), 199–325.

[7] R. NARASIMHAN, On the asymptotic stability of solutions of parabolic differential equations, *J. Rat. Mech. Anal.*, **3** (1954), 303–319.

[8] SIMON L., E. A. BADERKO, *Second Order Linear Partial Differential Equations*, Tankönyvkiadó, Budapest, 1983, (in Hungarian).

[9] A. A. SAMARSKII, *Theory of Difference Schemes*, Moscow, Nauka, 1977, (in Russian).

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 147–150

DESCRIPTION OF THE CLOSURE IN N^N OF THE SET OF DIGIT SEQUENCES OF THE GREEDY EXPANSIONS OF ALL x, $0 \leq x < 1$

By

I. JOÓ

Department of Analysis, Eötvös Loránd University, Budapest

(Received October 28, 1992)

Dedicated to Professor Béla Sz.-Nagy for his 80^{th} birthday

Let $\Theta > 1$ be an arbitrary real number, $0 \leq x < 1$ and consider the greedy expansion of x:

$$x = \frac{x_1}{\Theta} + \frac{x_2}{\Theta^2} + \dots$$

where x_i are integers and $0 \leq x_i < \Theta$ $(i = 1, 2, \dots)$.

Denote D_Θ the set of digit sequences (x_i) of the greedy expansions of all $x \in [0,1)$ furthermore denote S_Θ the closure of D_Θ in the topology of N^N. The convergence in this topology means that more and more coordinates become constant, hence $(s_1, s_2, \dots) \in S_\Theta$ if and only if for every $n \in N$ (natural number) there exists an element of D_Θ beginning with $s_1, \dots, s_n$. In this case the sequence $s_1, \dots, s_n$ itself is the greedy expansion of

$$\frac{s_1}{\Theta} + \dots + \frac{s_n}{\Theta^n} < 1,$$

hence

$$(1) \qquad (S_n)_1^\infty \in S_\Theta \Longleftrightarrow (s_1, \dots, s_n) \in D_\Theta \quad \text{for every} \quad n \in N.$$

Denote $d_\Theta(1)$ the greedy expansion of 1. We have

THEOREM 1. *If $d_\Theta(1)$ is infinite i.e. contains infinitely many nonzero digits, then S_Θ consists of exactly the following elements*

1. *all elements of D_Θ,*
2. *$d_\Theta(1)$,*
3. *the sequences of the following form: take a finite sequence of D_Θ, diminish the last digit by 1, and shift $d_\Theta(1)$ such that it begins at the position after the diminished digit (in other words, the*

sequences $(y_1,\dots,y_{k-1},y_k-1,t_1,t_2,\dots)$ *if* $(y_1,\dots,y_k)\in D_\Theta$ *and* $d_\Theta(1)=(t_1,t_2,\dots)$ *and* $y_k\ge 1$.

If $d_\Theta(1)$ *is finite,* $d_\Theta(1)=(t_1,\dots,t_r)$ *then let*

$$d^*_\Theta(1)=(t_1,\dots,t_{r-1},t_r-1)^\omega$$

and the same statements holds if we change $d_\Theta(1)$ *by* $d^*_\Theta(1)$*:*

THEOREM 2. *If* $d_\Theta(1)$ *is finite then* S_Θ *consists of exactly the following sequences*

1. *all elements of* D_Θ,
2. $d^*_\Theta(1)$,
3. *all sequences* $(y_1,\dots,y_{k-1},y_k-1,(t_1,\dots,t_{r-1},t_r-1)^\omega)$ *where* $(y_1,\dots,y_k)\in D_\Theta$ *and* $y_k\ge 1$.

In particular; $d_\Theta(1)\notin S_\Theta$ *in this case.*

PROOF OF THEOREM 1. a) The indicated sequences belong to S_Θ. To see this, use (1). It is clear for the elements of D_Θ and also for $d_\Theta(1)$. Take a greedy expansion

$$y=\frac{y_1}{\Theta}+\dots+\frac{y_k}{\Theta^k} \tag{1}$$

then

$$y=\frac{y_1}{\Theta}+\dots+\frac{y_{k-1}}{\Theta^{k-1}}+\frac{y_k-1}{\Theta^k}+\frac{t_1}{\Theta^{k+1}}+\frac{t_2}{\Theta^{k+2}}+\dots=y.$$

The first N digits will belong to D_Θ. Indeed, the corresponding sum is $<y<1$ and ask for the first digit which is not the largest possible. If we increase first one of the first $k-1$-th digit, the sum will be greater than y because (1) is greedy. The k-th digit can not be increased since our sum is $<y$ which contradicts to (1). On the other hand

$$\frac{t_1}{\Theta^{k+1}}+\dots+\frac{t_s}{\Theta^{k+s}}$$

is a greedy expansion, hence the later digits can not be increased as well.

b$_1$) If $(s_n)\in S_\Theta$ and $\frac{s_1}{\Theta}+\frac{s_2}{\Theta^2}+\dots<1$ is not greedy, then

$$(s_n)=(s_1,\dots,s_{k-1},s_k,t_1,t_2,\dots)\quad\text{for some } k \quad\text{and}\quad (s_1,\dots,s_{k-1},s_{k+1})\quad\text{is greedy.} \tag{2}$$

Indeed, since (s_n) is not greedy, there is a first index k where s_k can be increased. Since s_k can not be increased in the greedy expansions $\frac{s_1}{\Theta}+\dots+\frac{s_N}{\Theta^N}$, $N\ge k$, i.e.

$$\frac{s_1}{\Theta}+\dots+\frac{s_{k-1}}{\Theta^{k-1}}+\frac{s_k+1}{\Theta^k}>\frac{s_1}{\Theta}+\dots+\frac{s_N}{\Theta^N},\forall N,$$

the only possibility is that s_k can be increased by 1 and

$$\frac{s_1}{\Theta}+\ldots+\frac{s_{k-1}}{\Theta^{k-1}}+\frac{s_k+1}{\Theta^k}=\frac{s_1}{\Theta}+\ldots+\frac{s_{k-1}}{\Theta^{k-1}}+\frac{s_k}{\Theta^k}+\frac{s_{k+1}}{\Theta^{k+1}}+\ldots$$

where the left hand side is greedy. Since $(s_1,\ldots,s_N)$ is greedy for every N, it follows that

$$1=\frac{s_{k+1}}{\Theta}+\frac{s_{k+2}}{\Theta^2}+\ldots$$

is greedy, i.e. $s_{k+1}=t_1, s_{k+2}=t_2,\ldots$.

b$_2$) If $(s_n)\in S_\Theta$ and $\sum\limits_1^\infty s_n/\Theta^n=1$, then $(s_n)=(t_n)=d_\Theta(1)$. Indeed, suppose indirectly that (s_n) is not greedy i.e. that there exists a first k with

$$\frac{s_1}{\Theta}+\ldots+\frac{s_{k-1}}{\Theta^{k-1}}+\frac{s_k+1}{\Theta^k}\le 1 \tag{3}$$

and then (3) is greedy, consequently we have <1 (strong inequality) in (3). Since $\sum\limits_1^\infty s_n/\Theta^n=1$, then there exists N with

$$\frac{s_1}{\Theta}+\ldots+\frac{s_{k-1}}{\Theta^{k-1}}+\frac{s_k+1}{\Theta^k}<\sum_1^\infty\frac{s_n}{\Theta^n}$$

which contradicts to the fact that $\sum\limits_1^\infty s_n/\Theta^n$ is greedy. Theorem 1 is proved. □

Proof of Theorem 2. a) The indicated sequences belong to S_Θ. Consider first $(s_n)=d^*_\Theta(1)$. We have to prove that $(s_1,\ldots,s_N)\in D_\Theta$ for every N. The first $r-1$ digits can not be increased since $1=\frac{t_1}{\Theta}+\ldots+\frac{t_r}{\Theta^r}$ is greedy; the r-th digit can not be increased since

$$\frac{s_1}{\Theta}+\ldots+\frac{s_N}{\Theta^N}<1=\frac{t_1}{\Theta}+\ldots+\frac{t_r}{\Theta^r}.$$

Now $s_{r+i}=s_i$, hence we see by induction on N that $(s_1,\ldots,s_N)\in D_\Theta$ indeed. Next consider a sequence

$$(s_n)=(y_1,\ldots,y_{k-1},y_k-1,(t_1,\ldots,t_{r-1},t_r-1)^\omega)$$

with some $(y_1,\ldots,y_k)\in D_\Theta$. From

$$\sum_1^\infty s_n/\Theta^n<\frac{y_1}{\Theta}+\ldots+\frac{y_k}{\Theta^k}=\sum_1^\infty s_n/\Theta^n$$

we get that the first k digits in $(s_1, \dots, s_N)$ can not be increased. Hence we arrive to the statement (already proved) that the finite sequences cut from $d^*_\Theta(1)$ are greedy.

b$_1$) If $(s_n) \in S_\Theta$ and $\sum_1^\infty s_n/\Theta^n = 1$ then $(s_n) = d^*_\Theta(1)$. Indeed, since $d_\Theta(1)$ is finite, by $(*)$ $(s_n) \neq d_\Theta(1)$ hence is not greedy. Let k be the first digit for which

$$\frac{s_1}{\Theta} + \dots + \frac{s_{k-1}}{\Theta^{k-1}} + \frac{s_k + 1}{\Theta^k} \leq 1$$

Since $(s_1, \dots, s_N)$ is greedy, hence

$$\frac{s_1}{\Theta} + \dots + \frac{s_{k-1}}{\Theta^{k-1}} + \frac{s_k + 1}{\Theta^k} > \sum_1^N s_n/\Theta^n, \forall n$$

and then $\frac{s_1}{\Theta} + \dots + \frac{s_{k-1}}{\Theta^{k-1}} + \frac{s_k+1}{\Theta^k} = 1$ which is greedy. Consequently $k = r$ and $s_1 = t_1, \dots, s_r = t_r - 1$. From here we get $(s_{r+1}, s_{r+2}, \dots) \in S_\Theta$ and $\sum_1^\infty s_{r+n}/\Theta^n = 1$, so we get repeating the argument that $(s_n) = d^*_\Theta(1)$.

b$_2$) If $(s_n) \in S_\Theta$ is not greedy and $\sum_1^\infty s_n/\Theta^n < 1$ then

$$(s_n) = (y_1, \dots, y_{k-1}, y_k - 1, (t_1, \dots, t_{r-1}, t_r - 1)^\omega)$$

with some $(y_1, \dots, y_k) \in D_\Theta$. Indeed, let k be the first digit with

$$\frac{s_1}{\Theta} + \dots + \frac{s_{k-1}}{\Theta^{k-1}} + \frac{s_k + 1}{\Theta^k} \leq \sum_1^\infty s_n/\Theta^n (< 1). \tag{4}$$

Since $(s_1, \dots, s_N)$ is greedy, $\frac{s_1}{\Theta} + \dots + \frac{s_{k-1}}{\Theta^{k-1}} + \frac{s_k+1}{\Theta^k} > \sum_1^\infty s_n/\Theta^n$, hence in (4) equality holds and on the left we have a greedy expansion; we denote $y_1 = s_1, \dots, y_{k-1} = s_{k-1}$, and $y_k = s_k + 1$. From (4) we get that $1 = \sum_1^\infty s_{k+n}/\Theta^n$ and all finite sequences $(s_{k+1}, \dots, s_{k+N})$ belong to D_Θ. From b$_1$) we get that $(s_{k+n}) = d^*_\Theta(1)$. Theorem 2 is proved. □

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 151–159

SOLVING LINEAR FUNCTIONAL EQUATIONS FOR AN UNBOUNDED OPERATOR WITH THE HELP OF THE GRADIENT METHOD

By

J. KARÁTSON

Department of Applied Analysis, Eötvös Loránd University, Budapest

(Received November 30, 1992)

1. Introduction

The gradient method is commonly used for the approximate solution of linear functional equations. KANTOROVICH, L. V. proved (see [1] and [2], Ch. XV.) that if we have an equation $Ax = y$ in a Hilbert space H with a bounded self-adjoint operator $A \in \mathscr{B}(H)$ having bounds $0 < m < M < \infty$, the method yields a sequence of approximations converging to the solution with the speed of the geometric sequence. In this paper the method is extended to the case of unbounded linear operators which are symmetric and have a lower bound $m > 0$; the result for the speed of convergence remains similar to the bounded case. Finally some applications are shown.

2. The method in a Hilbert space

Let us consider the equation $Ax = y$ $(y \in R(A))$ with a densely defined symmetric operator A in a Hilbert space H having a positive lower bound m_0 (this also implies that the solution is unique.) We introduce another injective operator B and determine a new Hilbert space on which $B^{-1}A$ will prove to be the restriction of a bounded operator.

Therefore let B be a symmetric operator in H with the following properties:

(1) $D(B) = D(A)$, $R(B) \subset R(A)$

(2) there exist $M > m > 0$ such that for every $x \in D(B)$

$$m\langle Bx, x\rangle \leq \langle Ax, x\rangle \leq M\langle Bx, x\rangle.$$

This operator also has a positive lower bound $p = m_0/M$.

There always exists such a B, e.g. A itself would do. As a third property of B, we also expect the equations of the kind $Bx = y$ to be much easier to solve — without iteration, possibly with a formula — than $Ax = y$. The existence of such operators is proved in some special cases (see the applications in section 3).

THEOREM 1. *For an arbitrary symmetric operator B in H with properties (1) and (2) there exists a Hilbert space H_B (containing $D(B)$ as a subspace in vector space sense) and a bounded, positively definite and self-adjoint operator $C \in \mathcal{B}(H_B)$ such that $C_{|D(B)} = B^{-1}A$.*

PROOF. For $x, y \in D(B)$ let $[x, y] := \langle Bx, y\rangle$. The corresponding norm is denoted by $|x|$. The symmetricity and the positive lower bound of B imply the scalar product properties for $[\ ,\ ]$ and also the inequality

(3) $$|x| \geq \sqrt{p}\|x\|$$

between the two norms. Hence a Cauchy sequence in the new norm is a Cauchy sequence in the original norm as well, therefore it converges to an element of H. Thus the completion of the space $(D(B), |\ |)$ can obviously be given as

$$H_B := (x \in H : \text{there exists a Cauchy sequence } (x_n) \subset D(B) \text{ in the norm } |\ | \text{ which converges to } x \text{ in the norm } \|\ \|.)$$

In addition, inequality (3) can be extended from $D(B)$ to H_B.

Now let us consider the bilinear form $(x, y) \mapsto \langle Ax, y\rangle$ on $D(B) \times D(B)$. Using the Cauchy–Schwarz inequality for this form thought of as a scalar product and property (2) of B, we receive $|\langle Ax, y\rangle| \leq M|x||y|$, that is, this bilinear form is continuous with respect to the $|\ |$ norm topology. It can be easily extended to a form $\varphi : H_B \times H_B \mapsto \mathbb{C}$ which remains continuous with the same bound: $|\varphi(x, y)| \leq M|x||y|$.

Therefore we may apply the Riesz representation theorem in order to obtain a bounded linear operator $C : H_B \mapsto H_B$ so that for every $x, y \in H_B$ $\varphi(x, y) = [Cx, y]$ is valid. This operator has useful properties: since φ was generated by A on $D(B)$ and defined by continuous extension on H_B, C inherits the symmetricity of A, which means that it is even self-adjoint, being defined on the whole space H_B. Furthermore, relation (2) between A and B is equivalent to

(4) $$m|x|^2 \leq [Cx, x] \leq M|x|^2$$

for any $x \in D(B)$ and, owing to the continuity of C, for every $x \in H_B$ as well. This means

(5) $m \leq m_C \leq M_C \leq M$ for the bounds m_C and M_C of C.

Now it is left to prove that C is the extension of $B^{-1}A$ (the latter makes sense because of the assumption $R(B) \subset R(A)$). By definition, for every $x, y \in D(B)$

$$[Cx, y] = \varphi(x, y) = \langle Ax, y\rangle = \langle BB^{-1}Ax, y\rangle = [B^{-1}Ax, y],$$

which, as $D(B)$ is dense in H_B, means that

$$Cx = B^{-1}Ax \quad (x \in D(B)).$$

The theorem is proved.

Having received the bounded extension of $B^{-1}A$, we can now turn to the way in which the gradient method can be used. Owing to (6) we can write the equation $Ax = y$ in the form $BCx = y$. Now first we solve $Bz = y$, which was supposed to be easy ($y \in R(B)$ follows from $y \in R(A)$ and $R(A) \subset R(B)$), then, having got z, we have to find the solution of $Cx = z$. Inequalities (4) and (5) show that C is just of the type on H_B for which the gradient method for bounded operators can be used in order to solve equations of the kind $Cx = z$ (see [2], Ch. XV. 1.2–3). The iteration in this setting will be as follows: beginning with $x_0 \in D(B)$ and supposing to have found x_{n-1},

(7) $$x_n := x_{n-1} + t_n z_n \quad \text{with}$$

(8) $$z_n := -(Cx_{n-1} - z) \quad \text{and}$$

(9) $$t_n := \frac{|z_n|^2}{[Cz_n, z_n]}.$$

The sequence (x_n) tends to the solution x^* with the speed given by the estimate

(10) $$|x_n - x^*| \le \frac{1}{m_C}|Cx_0 - z| \left[\frac{M_C - m_C}{M_C + m_C}\right]^n.$$

However, these formulae are not really useful in this form since they contain the new norm and the hardly determinable operator C together with its unknown bounds. We now get round this inconvenience.

The sequence (x_n) remains in $D(B)$: supposing $x_{n-1} \in D(B)$, $Cx_{n-1} = B^{-1}Ax_{n-1}$ implies $Cx_{n-1} \in R(B^{-1}) = D(B)$, hence $z \in D(B)$ implies $z_n \in D(B)$; this leads to $x_n \in D(B)$.

This allows us to express the formulae of the iteration with the help of the original operators and scalar product. First, since $z_n \in D(B)$, B can be applied to (8), thus $Bz_n = BCx_{n-1} - Bz = Ax_{n-1} - y =: w_n$. This means that the equation

$$Bz_n = w_n$$

has to be solved in each step (an easy task by assumption). The constant t_n will be of the form

$$t_n := \frac{|z_n|^2}{[Cz_n, z_n]} = \frac{\langle Bz_n, z_n \rangle}{\langle BCz_n, z_n \rangle} = \frac{\langle Bz_n, z_n \rangle}{\langle Az_n, z_n \rangle}.$$

We still have to see that the convergence can also be expressed with the help of the original norm.

THEOREM 2. *Let $x_0 \in D(B)$ and $x_n := x_{n-1} + t_n z_n$, supposed that z_n is the solution of the equation $Bz_n = Ax_{n-1} - y$ and $t_n := \frac{\langle Bz_n, z_n \rangle}{\langle Az_n, z_n \rangle}$. Then*

$$\|x^* - x_n\| \le \frac{\|Ax_0 - y\|}{mp} \left[\frac{M-m}{M+m} \right]^n . \tag{11}$$

PROOF. Inequality (10) has to be used. For the left hand side $\|x^* - x_n\| \le \frac{1}{\sqrt{p}} |x^* - x_n|$ is valid, since $x^* - x_n \in D(A) = D(B)$. Thus we have to give an upper estimate for the right hand side.

First, $|Cx_0 - z|^2 = |B^{-1}(Ax_0 - y)|^2 = \langle Ax_0 - y, B^{-1}(Ax_0 - y) \rangle \le \|B^{-1}\| \|Ax_0 - y\|^2$, and here we have $\|B^{-1}\| \le \frac{1}{p}$ from $\langle Bx, x \rangle \ge p\|x\|^2$ for $x \in D(B)$. Hence $|Cx_0 - z| \le \frac{1}{\sqrt{p}} \|Ax_0 - y\|$. On the other hand, $m \le m_C \le M_C \le M$ implies that $\left[\frac{M_C - m_C}{M_C + m_C}\right]^n \le \left[\frac{M-m}{M+m}\right]^n$. These estimates lead directly to the desired inequality.

REMARK. At the beginning of this section we assumed that $y \in R(A)$. This can be omitted if we use Friedrichs's theorem (see [3]): every symmetric operator S in H with a positive lower bound has a self-adjoint extension $\hat{S}$, also with a positive lower bound, for which $D(\hat{S}) \subset H_S$ (a space constructed similarly to H_B) and $R(\hat{S}) = H$. Applying this theorem to A, we see that for arbitrary $y \in H$ the equation $Ax = y$ always has a generalized solution, i.e. $\hat{x} \in D(\hat{A})$ with $\hat{A}\hat{x} = y$. If we restrict ourselves to $y \in R(B)$ then the sequence (x_n) will remain in $D(B)$ just like before (the arguments used there remain valid), but this result will be lost if an arbitrary $y \in H$ is allowed. Then at the n-th step of the iteration we only have a generalized solution for $Bz_n = Ax_{n-1} - y$ (considering the Friedrichs extension $\hat{B}$ of B). In this case it can be seen that (x_n) remains in $D(\hat{B})$.

3. Some applications to differential equations

The extension of the gradient method to unbounded operators enables us to apply it to some types of symmetric differential equations. We show in this section that in the case of second order ordinary and elliptic equations there exists an appropriate operator B of the kind required in the abstract setting.

a) Ordinary case.

Let $I := [a,b] \subset \mathbb{R}$ and let p and q be real-valued functions in $C^1(I)$ and $C(I)$, respectively, with $p(x) \geq m > 0$ and $q(x) \geq 0$ for $x \in I$. Define L as the following differential operator of Sturm form: $D(L) := C_0^2(I) :=$ $= \{y \in C^2(I) : y(a) = y(b) = 0\} \subset L^2(I)$ and $Ly := -(py')' + qy$. Then L is a densely defined operator in $L^2(I)$, it is symmetric and the lower bounds for p and q imply a positive lower bound for L. (To see this, we can use the following formulae: integrating by parts and using the boundary conditions, we get

$$\int_a^b (Ly)\overline{y} = \int_a^b (p|y'|^2 + q|y|^2). \tag{12}$$

Further, for $y \in C^1(I)$ with $y(a) = y(b) = 0$, hence in $D(L)$ too

$$\int_a^b |y|^2 \leq \left(\frac{b-a}{\pi}\right)^2 \int_a^b |y'|^2 \tag{13}$$

is valid, which is easily proved from the inequality $0 \leq \int_0^\pi |y'(x) - y(x)\mathrm{ctg}\, x|^2 dx$ and integration by parts.)

Besides, it is known that for every $f \in C(I)$ the equation $Ly = f$ has a solution in $D(L)$, so $C(I) \subset R(L)$; this means (with $R(L) \subset C(I)$, which is obvious) that $R(L) = C(I)$.

The solution of $Ly = f$ is always unique owing to the positive lower bound of L. Wishing to use the gradient method for determining y, we introduce $By := -y''$ with $D(B) := D(L)$.

The operator B is also of the type of L with $p \equiv 1$ and $q \equiv 0$, therefore it is symmetric and $R(B) = R(L) = C(I)$; these facts mean that only property (2) for B is left to be proved.

Applied to the operator B, equation (12) means

$$\int_a^b (By)\overline{y} = \int_z^b |y'|^2. \tag{14}$$

Using this together with (12) and (13) and letting $M := \max_I p + \frac{b-a}{\pi} \max_I q$, we have $m\langle By, y\rangle \le \langle Ly, y\rangle \le M\langle By, y\rangle$, which means that B is of the desired type indeed.

The nth step of the iteration is $y_n = y_{n-1} + t_n z_n$, where z_n is the solution of $Bz_n = g_n := Ly_{n-1} - f$ (the smoothness assumptions for f and the coefficients of L imply that $g_n \in C(I)$). Thus our task is to solve $-z_n'' = g_n$ in $C_0^2(I)$, which is easily done by the formula

$$z_n(x) = -\int_a^x \int_a^t g_n + \frac{x-a}{b-a} \int_a^b \int_a^t g_n.$$

Finally, supposed we started from $y_0 :\equiv 0$, (10) now means

$$\|y_n - y^*\|_{L^2(I)} \le c \left[\frac{M-m}{M+m}\right]^n$$

with $c := \frac{\|f\|_{L^2(I)}}{m\pi}(b-a)M$, using the fact that in this case $p = m_0/M = m\pi/(b-a)M$.

b) Elliptic case

Let $0 < \alpha < 1$ and let $\Omega \subset \mathbb{R}^n$ be a bounded domain with a boundary $\partial\Omega \subset C^{2,\alpha}$ (that is, the 2nd partial derivatives are Lipschitz-α-continuous). Now our operator will be (densely) defined as

$$D(A) := (u \in C^{2,\alpha}(\overline{\Omega}) : \quad u_{|\partial\Omega} = 0)$$

and

$$Au := -\sum_{i,j=1}^n \partial_i(a_{ij}\partial_j u) + qu,$$

where $a_{ij} \in C^{1,\alpha}(\overline{\Omega})$, $a_{ij} = a_{ji}$ $(i,j = 1,\dots,n)$ and the operator is uniformly elliptic:

$$m\sum_{i=1}^n |z_i|^2 \le |\sum_{i,j=1}^n a_{ij}(x) z_i \overline{z}_j| \le m' \sum_{i=1}^n |z_i|^2 \quad (x \in \Omega,\, z \in \mathbb{C}^n)$$

with some $m, m' > 0$; finally, let $q \in \mathrm{Lip}^\alpha(\overline{\Omega})$ and $q \ge 0$. Then $D(A)$ is dense in $L^2(\Omega)$; further, A is symmetric and has m as a positive lower bound.

The reason for always assuming Lipschitz-α property on $\overline{\Omega}$ instead of simple continuity on Ω is that analogue of the existence theorem in the ordinary case does not hold, i.e. the equation $Au = f$ with $f \in C(\Omega)$ (and coefficients of A in $C^1(\Omega)$ and $C(\Omega)$) cannot always be solved in $C^2(\Omega)$. However, we have a similar theorem with $f \in \mathrm{Lip}^{\alpha}(\overline{\Omega})$ and with our stronger assumptions for the coefficients of A which yields a solution in $C^{2,\alpha}(\overline{\Omega})$ (see MIRANDA, C. [4]). This solution is unique, since A is injective.

REMARK. For $n = 2$, $\Omega = S(0,1)$ and a diagonal matrix $[a_{ij}]$ (that is, for an operator of the form $-\partial_1(a\partial_1 u) - \partial_2(b\partial_2 u) + qu$ with $a, b \in C^1(\Omega)$ with positive lower bounds on $\overline{\Omega}$, $q \in C(\Omega)$) KANTOROVICH, L. V. and AKILOV, G. P. [2] reduce the problem to the bounded case, using the fact that in this case $\Delta^{-1}A$ is bounded, self-adjoint and has a positive lower bound on the space $H_0^2(\Omega)$. Now we apply the modified method to an arbitrary uniformly elliptic operator.

The role of B is played by $Bu =: -\Delta u$ with the domain $D(B) := D(A)$. Similarly to the ordinary case, B is the special case of the type of A ($[a_{ij}] = I, q \equiv 0$), thus it is symmetric and has a positive lower bound. To prove property (2), we use the following formulae:

$$\int_{\Omega} (Au)\overline{u} = \int_{\Omega} \left[\sum_{i=1}^{n} a_{ij}(\partial_i u)(\partial_j \overline{u}) + q|u|^2 \right], \quad \int_{\Omega} (Bu)\overline{u} = \int_{\Omega} \sum_{i=1}^{n} |\partial_i u|^2$$

(consequences of the homogeneous boundary conditions), and

$$\int_{\Omega} |u|^2 \leq c_1 \int_{\Omega} \sum_{i=1}^{n} |\partial_i u|^2$$

with $c_1 > 0$ (see [5], Ch.11. §A). Then uniform ellipticity and $q \geq 0$ imply $m\langle Bu, u\rangle \leq \langle Au, u\rangle \leq M\langle Bu, u\rangle$ with $M := K + c_1 \max_{\Omega} q$. Thus we are allowed to use our method since the conditions for the domains and ranges of A and B are fulfilled, the first by definition, the latter because of the theorem quoted after defining A, implying $R(A) = R(B) = \mathrm{Lip}^{\alpha}(\overline{\Omega})$ (for this we also need the fact that $R(A) \subset \mathrm{Lip}^{\alpha}(\overline{\Omega})$, which is obvious from the smoothness of the coefficients).

In order to determine $u_n := u_{n-1} - t_n z_n$, we now need to solve $-\Delta z_n = g_n := Au_n - f$ in $C^{2,\alpha}(\overline{\Omega})$ with the boundary condition $z_{n|\partial\Omega} = 0$.

The smoothness assumptions for f and the coefficients of A imply that $g_n \in \mathrm{Lip}^{\alpha}(\overline{\Omega})$. The solution can be given by an integral formula if the Green function is known for Ω, e.g. in the case of a ball or a rectangular domain etc. The solution can be simplified further (avoiding the use of Green functions) if the coefficients are functions of special form. For polynomial coefficients, KANTOROVICH and AKILOV (see [2], Ch. XV.) show an easy method by looking for the solution of $\Delta p = q$ in $S(0,1) \subset \mathbb{R}^2$ in the form $p(x,y) := (x^2+y^2-1)r(x,y)$ with $\operatorname{gr} r = \operatorname{gr} q$.

The estimate for the speed of convergence (starting again from $u_0 :\equiv 0$) is

$$\|u_n - u^*\|_{L^2(\Omega)} \le c \left[\frac{M-m}{M+m}\right]^n$$

with $c := \frac{\|f\|_{L^2(\Omega)}}{m^2} c_1 M$, using that now $p = m_0/M = m/c_1 M$.

REMARK. Similar considerations show the applicability of the method for higher order symmetric equations both in the ordinary and elliptic case, namely, for the following types:

a) a $2n$-th order ordinary equation of the form

$$Ly := \sum_{k=0}^{n} (-1)^k (a_k y^{(k)})^{(k)} = f$$

on $I := [a,b]$ with $D(L) := \{y \in C^{(2n)}(I) : y^{(j)}(a) = y^{(j)}(b) = 0$ for $j = 0,1,\dots,n-1\}$, $f \in C(I)$ and coefficients $a_k \in C^{(k)}(I)$, $a_k \ge 0$ $(k = 0,1,\dots,n-1)$ and $a_n \ge m > 0$. Now a suitable operator B can be defined by $D(B) := D(L)$ and $By := (-1)^n y^{(2n)}$ for $y \in D(B)$, bearing in mind that $Ly = f$ can be solved in $C^{(2n)}(I)$ for any $f \in C(I)$, including the case $L = B$.

b) a $2n$-th order elliptic equation of the form

$$Au := \sum_{k=0}^{n} (-1)^k \sum_{|\alpha|=|\beta|=k} \partial_\alpha (a_{\alpha\beta} \partial_\beta u) = f$$

in a bounded domain $\Omega \subset \mathbb{R}^N$ with $\partial\Omega \subset C^{2n,\nu}$ $(0 < \nu < 1)$, defining $D(A) := \{u \in C^{2n,\nu}(\overline{\Omega}) : \partial_\alpha u_{|\partial\Omega} = 0$ for multiindices $|\alpha| \le n-1\}$ and supposing

$$\sum_{|\alpha|=|\beta|=k} a_{\alpha\beta} x_\alpha \overline{x}_\beta \ge 0 \quad (k = 0,1,\dots,n-1)$$

and

$$\sum_{|\alpha|=|\beta|=n} a_{\alpha\beta} x_\alpha \overline{x}_\beta \ge m \sum_{|\alpha|=n} |x_\alpha|^2 \quad (m > 0)$$

for every array $x \in \mathbb{C}^{[n^k]} := \mathbb{C}^{n \times n \times \dots \times n}$; further, we suppose that $a_{\alpha\beta} \in C^{k,\nu}(\overline{\Omega})$ $(|\alpha| = |\beta| = k;\ k = 0,1,\dots,n)$ and $f \in \operatorname{Lip}^\nu(\overline{\Omega})$. In this case $Bu := (-1)^n \Delta^n u$ $(u \in D(B) := D(A))$ will be suitable. In order to assure the existence of a solution $u \in C^{2n,\nu}(\overline{\Omega})$ and the solvability of $\Delta^n u = g_n$ in each step of the iteration, we may refer to [6], 12.1.

References

[1] KANTOROVICH, L. V., On an effective method of solution of extremal problems for quadratic functionals, *Dokl. Akad. Nauk. SSSR*, **48** (1945), 483–487.

[2] KANTOROVICH, L. V., AKILOV, G. P., *Functional Analysis*, Pergamon Press, 1982.

[3] FRIEDRICHS, K., Spektraltheorie halbbeschränkten Operatoren und Anwendung auf die Spektralzerlegung von Differentialoperatoren, *Math. Annalen*, **109** (1934), 465–487.

[4] MIRANDA, C., *Equazioni alle derivati parziali di tipo ellittico*, Springer–Verlag, 1955.

[5] SMOLLER, J., *Shock waves and reaction-diffusion equations*, Springer–Verlag, 1982.

[6] AGMON, S., DOUGLIS, A., NIRENBERG, L., Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions, *Communications on Pure and Applied Mathematics*, **12** (1959), 623–727.

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 161–170

NOTES ON MINIMAX THEOREMS

By

I. JOÓ

Department of Analysis, Eötvös Loránd University, Budapest

(Received December 4, 1992)

Dedicated to Professor János Balázs for his 70^{th} birthday

A famous two-function minimax theorem is the Nikaido–Isoda one, connected with classical convexity. M. HORVÁTH asked, whether this theorem remains valid for KY FAN concave functions or not, and negative answer is given in [5]. After we have constructed a counterexample from C^∞-functions with M. HORVÁTH in [3] and I remarked that if the x_3 in the Ky Fan concavity depends "regularly" on the parameters x_1, x_2 and λ then the answer is positive. Namely, let $\psi : [0,1]\times[0,1]$ be a contraction in the sense that for some $0<\varepsilon<1$

$$|\psi(x_1,x_1)-\psi(x_2,x_2)|\leq\varepsilon|x_1-x_2| \qquad (x_1,x_2\in[0,1]).$$

THEOREM 1. *Let $f : [0,1]\times[0,1]\to\mathbb{R}$ be continuous and suppose that $f(\psi(x_1,x_2),y)\geq 0.5\circ[f(x_1,y)+f(x_2,y)]$ for all $x_1,x_2,y\in[0,1]$. Then there exists $x^*\in[0,1]$ with $f(x^*,x^*)\geq f(z,x^*)$ for all $z\in[0,1]$.*

The proof given below uses only Banach fixed point theorem, valid in any complete metric space, so very general Nikaido–Isoda type theorems can be obtained in this way.

PROOF. For $x_1=x_2=x$ we obtain from our assumption $f(\psi(x,x),y)\geq$ $\geq f(x,y)$ and hence if $B : x\to\psi(x,x)$, then B^nx tends to the same x^*, starting from any x, further $f(B^nx,y)\nearrow f(x^*,y)$, consequently $f(x,y)\leq$ $\leq f(x^*,y)\forall x,y$. This is more than what we need, namely, our statement follows in the special case of $y=x^*$ (with $z=x$). ☐

In the next part of this paper we prove some minimax theorems for two functions. These studies were motivated by the nice result of IRLE [4] and by the interesting, stimulating papers [11–13] of S. SIMONS. The author

wishes to thank Professor S. SIMONS for sending him his new unpublished and published papers.

Let $\varrho_i : R^2 \to R$ (R denotes the real line, $i = 1,2$) be any continuous functions such that

(1) $$x < y \Rightarrow x < \varrho_i(x,y) < y, \quad x < \varrho_i(y,x) < y$$

(2) $$x = y \Rightarrow x = y = \varrho_i(x,y),$$

(3) for any $a_{ij} \in R \quad (i,j = 1,2)$

$$\varrho_2(\varrho_1(a_{11},a_{12}),\varrho_1(a_{21},a_{22})) \geq \varrho_1(\varrho_2(a_{11},a_{21}),\varrho_2(a_{12},a_{22}))$$

hold. First we prove

THEOREM 2. *Let X, Y be any non-empty sets (without any topology!), $f,g : X \times Y \to R$ any functions defined on $X \times Y$ such that*

(4) $$f \leq g \quad \text{on} \quad X \times Y, \quad f \text{ is bounded}$$

(5) f *is* ϱ_2 *— convex on* Y, *i.e.* $\forall y_1, y_2 \in Y \exists y_3 \in Y$: $$f(x,y_3) \leq \varrho_2(f(x,y_1), f(x,y_2)) \forall x \in X,$$

(6) g *is* ϱ_1 *— concave on* X, *i.e.* $\forall x_1, x_2 \in X \exists x_3 \in X$: $$g(x_3,y) \geq \varrho_1(g(x_1,y), g(x_2,y)) \forall y \in Y.$$

(7) $x_i \geq x_i' \Rightarrow \varrho_i(x_1,x_2) \geq \varrho_i(x_1',x_2')$ $(i = 1,2)$ *i.e.* ϱ_1 *and* ϱ_2 *are monotone. Then for any finite set* $H \subset X$ *we have*

(8) $$\inf_Y \sup_H f(x,y) \leq \sup_X \inf_Y g(x,y).$$

The concept of "averaging functions" was introduced in [4]. For special ϱ_1 and ϱ_2 Theorem 2 was proved in the earlier papers [6–9] and their proof works also in the present general case.

PROOF. We need a series of lemmas. Denote $\mathcal{F} = \mathcal{F}(X,R)$ the set of all functions $f : X \to R$ defined on R and let $|f| := \sup\{|f(x)| : x \in X\}$. $Z \subset \mathcal{F}$ is said to be ϱ — convex with respect to some $\varrho : R^2 \to R$ if $f, g \in Z \Rightarrow \varrho(f,g) \in Z$ $(\varrho(f,g)(x) := \varrho(f(x),g(x)))$. Denote $Co\varrho(Z)$ the ϱ-convex hull of any $Z \subset \mathcal{F}$, i.e. the intersection of all ϱ-convex sets containing Z (which is ϱ-convex).

LEMMA 1. *If Z is any bounded ϱ-convex set in $\mathcal{F}$ and ϱ is any continuous $R^2 \to R$ function, then $\overline{Z}$ (the closure of Z) is also ϱ-convex.*

PROOF OF LEMMA 1. Let $f,g \in \overline{Z} \Rightarrow \exists f_n, g_n \in Z$: $f_n \to f$, $g_n \to g$ uniformly on X. The boundedness of Z means that $\exists C \in R$: $|f| \leq C$

$\forall f \in Z$ hence $\forall f \in \overline{Z}$: $|f| \leq C$. Consider the closed ball $S(0,C)$ in R^2 centered at 0 and of radius C. Because (we assumed) $\varrho : R^2 \to R$ is continuous, hence $\varrho \mid S(0,C)$ is uniformly continuous, i.e.

$$\forall \varepsilon > 0 \exists \delta > 0 : \left\{(x-u)^2 + (y-v)^2\right\}^{1/2} < \delta \Rightarrow |\varrho(x,y) - \varrho(u,v)| < \varepsilon.$$

Because $f_n \to f$, $g_n \to g$ uniformly, hence

$$\exists T : |f_n - f| < \delta/\sqrt{2}, \quad |g_n - g| < \delta/\sqrt{2} \quad \text{if } n > T,$$

consequently

$$\left[(f_n(x) - f(x))^2 + (g_n(x) - g(x))^2\right]^{1/2} < \delta \Rightarrow |\varrho(f_n(x), g_n(x)) -$$
$$-\varrho(f(x), g(x))| < \varepsilon \quad (x \in X,\ n > T) \quad \text{i.e. } |\varrho(f_n, g_n) - \varrho(f,g)| \to 0.$$

But thus $f_n, g_n \in Z \Rightarrow \varrho(f_n, g_n) \in Z \Rightarrow \varrho(f,g) \in \overline{Z}$. □

For the Proof of Theorem 2 we use induction in the cardinality of H. If H has 1 element, then (8) is true; suppose we know (8) for any H with $|H| < n$ and prove it for the case when $|H| = n$. Suppose indirectly that

$$\alpha > \beta \tag{9}$$

where $\alpha := \inf_Y \sup_H f$, $\beta := \sup_X \inf_Y g$, denote $H := \{x_1, \ldots, x_n\}$, $Z := = \{(z_1, \ldots, z_n) \in R^n \colon \exists y,\ z_i = f(x_i, y),\ i = 1, \ldots, n\}$.

Lemma 2. *For any $z \in Co\varrho_2(Z) \exists i$: $z_i \geq \alpha$.*

Proof of Lemma 2. Denote $Z_0 := Z$, $Z_{i+1} := \{\varrho_2(z', z'') : z', z'' \in \in Z_i\}$. Obviously, $Z_0 \subset Z_1 \ldots$ and $Co\varrho_2(Z) = \bigcup_0^\infty Z_i$. Hence it is enough to prove Lemma 2 for every Z_i. To this, it is enough to show that

$$\forall z \in Z_n \exists v \in Z : \ v \leq z \quad (\text{i.e. } v_i \leq z_i \text{ for } i = 1, \ldots, n). \tag{10}$$

We use induction in n. The statement is true for $n = 1$ i.e. for Z_1, because $z \in Z_1 \Rightarrow z = \varrho_2(z_1, z_2)$; $z_1, z_2 \in Z$

$$z_1 = (f(x_1, y_1), \ldots, f(x_n, y_1)), \qquad z_2 = (f(x_1, y_2), \ldots, f(x_n, y_2)).$$

According to (5) for $y_1, y_2 \exists y_3$: $f(x_i, y_3) \leq \varrho_2(f(x_i, y_1), f(x_i, y_2))$ $(1 \leq i \leq n)$, i.e. for $z_3 = (f(x_1, y_3), \ldots, f(x_n, y_3))$ we have $z_3 \leq \varrho_2(z_1, z_2) \Rightarrow z_3 \leq z$.

Now suppose our statement (10) is valid for Z_n and prove it for Z_{n+1}. Let $z \in Z_{n+1}$. Then $z = \varrho_2(z_1, z_2)$ and $z_1, z_2 \in Z_n$. The induction hypothesis implies: there exists $z_1^*, z_2^* \in Z$ with $z_1^* \leq z_1$, $z_2^* \leq z_2$, and hence by (7) $z = \varrho_2(z_1, z_2) \geq \varrho_2(z_1^*, z_2^*) \geq z_3$ for some $z_3 \in Z$. The second inequality is true because $z_1^*, z_2^* \in Z$ and we know the statement for Z_1 $(\varrho_2(z_1^*, z_2^*) \in Z_1)$.

Thus we have proved the property (10). For any $v \in Z \exists i \ : \ v_i \geq \alpha$, because if $v_i < \alpha$ for every i, then $v_i = f(x_i, y^*) < \alpha$ for some $y*$ and then $\alpha = \inf_Y \sup_H f(x,y) < \sup_H f(x,y^*) < \alpha$ which is a contradiction. Thus for any $v \in Z \exists i$: $v_i \geq \alpha$. But we know that $\forall z \in Z_n \exists v \in Z$: $v \leq z$, hence $\forall z \in Z_n \exists i$: $z_i \geq \alpha$. Taking into account $Co\varrho_2(Z) = \bigcup_0^\infty Z_n$, Lemma 2 follows. □

REMARK. Taking limits we immediately obtain that

$$\forall v \in \overline{Co\varrho_2(Z)} \exists i \ : \ v_i \geq \alpha. \tag{11}$$

LEMMA 3. *$\overline{Co\varrho_2(Z)}$ is bounded.*

PROOF OF LEMMA 3. By (4) Z is bounded, hence $\exists c \in R$: $z \in Z \Rightarrow |z| < c$, hence $|z_i| < c$ for every i, so $\varrho_2(z_i^{(1)}, z_i^{(2)}) < c \forall i$. Thus we have $v_i < c$ for every $v \in Z_1$. We can obtain the same for $Z_2, Z_3, \ldots$ by induction. Thus $v \in Z_j \Rightarrow v_i < c$ for every i, so $|v| \leq \sqrt{n}c$ hence any element of $Co\varrho_2(Z) = \bigcup_1^\infty Z_j$ is smaller than $\sqrt{n}c$ and Lemma 3 is proved. □

Until now we have proved:

$$K := \overline{Co\varrho_2(Z)} \text{ is compact, } \varrho_2\text{-convex and } \forall z \in K \exists i : z_i \geq \alpha. \tag{12}$$

Denote $h_i \ : \ K \to R$ the i-th projection, i.e. $h_i(z) := z_i$ for any $z \in K$. Let

$$L := \overline{Co\varrho_1\{h_1, \ldots, h_n\}} \qquad (\subset K_R).$$

LEMMA 4. *We have for every $h \in L$*

$$\inf_{z \in K} h(z) \leq \beta. \tag{13}$$

PROOF OF LEMMA 4. (13) is fulfilled for h_i, $1 \leq i \leq n$ because $\beta = \sup_X \inf_Y g$ by definition, so $\forall x_i \in X$: $\inf_Y g(x_i, y) \leq \beta$ but because of $f(x_i, y) \leq g(x_i, y)$ we have $\inf_{z \in K} h_i(z) \leq \inf_{y \in Y} f(x_i, y)$. $Co\varrho_1(\{h_1, \ldots, h_n\}) = \bigcup_0^\infty L_n$, where $L_0 := \{h_1, \ldots, h_n\}$ and $L_{i+1} := \{\varrho_1(h', h'') : h', h'' \in L_i\}$. Now we show that $\forall h \in L_k \exists x' \in X : h(z) \leq g(x', y)$, where y corresponds to z. Use induction in k. For $k = 1$ this is true, because in this case $h(z) = \varrho_1(h_1(z), h_j(z)) = \varrho_1(f(x_i, y), f(x_j, y)) \leq \varrho_1(g(x_i, y), g(x_j, y)) \leq g(x', y)$. The first inequality follows from the monotonity of ϱ_1 (see (7)) and the second one from the ϱ_1-concavity of g (x' corresponds to x_i and x_j). Suppose we know the statement for k and prove it for $k+1$. So let $h \in L_{k+1}$. $h(z) =$

$= \varrho_1(h'(z), h''(z))$, where $h', h'' \in L_n$. It follows from the induction hypothesis that

$$\exists x', x'' \in X : h'(z) \leq g(z', y), \quad h''(z) \leq g(x'', y)$$
$$h''(z) \leq \varrho_1(h'(z), h''(z)) \leq \varrho_1(g(x', y), g(x'', y))$$

because of the monotonity of ϱ_1.

Because g is ϱ_1-concave and ϱ_1 is monotone increasing, for $x', x'' \in \in X \exists x'''$:

$$g(x''', y) \geq \varrho_1(g(x', y), g(x'', y)) \forall y \in Y$$

and

$$h(z) \leq \varrho_1(g(x', y), g(x'', y)) \leq g(x''', y),$$

whence

$$\forall h \in L_n \exists x^* \in X : h(z) \leq g(x^*, y),$$

where y corresponds to z. Consequently, for any

$$h \in \bigcup_1^\infty L_n = Co\varrho_1(\{h_1, \ldots, h_n\}) \exists x^* \in X : h(z) \leq g(x^*, y)$$

and hence $\inf_{x \in K} h(z) \leq \inf_{y \in Y} g(x^*, y)$. Now let $h \in \overline{Co\varrho_1\{h_1, \ldots, h_n\}} =: L$ be arbitrary. Then $\exists h^{(n)} \in L : h^{(n)} \to h$ in norm. Suppose indirectly that $\gamma := \inf_{z \in K} h(z) = \min_{z \in K} h(z) > \beta$. Then $\gamma > \beta + \varepsilon$ with $\varepsilon = \frac{\gamma - \beta}{3}$. Because $h^{(n)} \to h$ uniformly, hence $\exists N$: for every $n > N$ we have $\sup |h - h^{(n)}| < \frac{\varepsilon}{2}$. But because of $h > \beta + \varepsilon$ we have $h^{(n)}(z) > \beta + \frac{\varepsilon}{2}$ for $n > N$ and $z \in K$, hence $\inf_{z \in K} h^{(n)}(z) \geq \beta + \frac{\varepsilon}{2}$ which is a contradiction. Lemma 4 is proved. □

Denote F the set of those ϱ_2-convex sets D of $\mathbb{R}^N$ for which the following three conditions are fulfilled:

$$z \in D \Rightarrow \exists i : z_i \geq \alpha, \tag{14}$$

$$\inf_{z \in D} h(z) = \min_{z \in D} h(z) \leq \beta \quad \text{for every } h \in L, \tag{15}$$

$$D \subset K. \tag{16}$$

Remark that (16) implies (14).

Obviously, $F \neq \emptyset$ because $K \in F$. Our aim now is to show that the conditions of Zorn's Lemma are fulfilled for the system F. To this let D_i, $i \in I$ be any chain, ordered by inclusion. Since the D_i are compact,

$D^* = \bigcap_{i \in I} D_i \neq \emptyset$ and (14), (16) obviously holds for D^*. The index set I is confinal with a cardinal number, so we can suppose that I itself is a cardinality κ (and $i < j \Leftrightarrow D_i \supset D_j$). Since the D_i satisfy (15), there exist for every fixed $h \in L$ points $z_i \in D_i$ with $h(z_i) \leq \beta$. Denote z^* a condensation point of the net z_i. From the continuity of h we get $h(z^*) \leq \beta$. If $z^* \notin D^*$ then $z^* \notin D_i$ for some $i < \kappa$, hence the neighbourhood D_i^c of z^* can only contain the points z_j, $j < i$ whose cardinality is $< \kappa$. This contradicts the condensation point property, hence $z^* \in D^*$. We proved that $D^* \in F$ hence the Zorn-lemma can be applied. Applying Zorn's Lemma we have: there exists at least one minimal element R in F. By (16) we have $R \subset K$. Introduce the notations:

$$N_\xi(h) := \{z \in R : h(z) \leq \xi\}, \quad N_\xi^*(h) := \{z \in R : h(z) < \xi\}.$$

LEMMA 5. *If $\beta < \gamma < \alpha$ then $N_\gamma(h) \neq \emptyset$, closed, ϱ_2-convex and for some $h^* \in L$ we have $N_\gamma(h^*) \subsetneq R$.*

PROOF OF LEMMA 5. (In the formulation of Lemma 5 we used the indirect assumption $\alpha > \beta$.) Obviously,

$$\inf_{z \in R} h(z) = \min_{z \in R} h(z) \leq \beta < \gamma \Rightarrow N_\gamma(h) \neq \emptyset \text{ and } N_\gamma^*(h) \neq \emptyset.$$

We have $N_\gamma(h_i) \subsetneq R$ for some $1 \leq i \leq n$ because if $N_\gamma(h_i) = R$ for $i = 1, \ldots, n$, then $h_i(z) \leq \gamma$ for every $z \in R$ and this contradicts the fact that $\forall z \in R \exists i : z_i \geq \alpha$, namely $\alpha > \gamma$.

We have to prove at last that $N_\gamma(h)$ is ϱ_2-convex for every $h \in L$. To this we need the

SUBLEMMA. *For any $h \in L$ and $z_1, z_2 \in R$ we have*

$$h \cdot (\varrho_2(z_1, z_2)) \leq \varrho_2(h(z_1), h(z_2)). \tag{17}$$

PROOF OF THE SUBLEMMA. Consider the usual partition

$$Co\varrho_1(\{h_1, \ldots, h_n\}) = \bigcup_{k=0}^{\infty} L_k.$$

It is enough to prove (17) for L_k, $k = 0, 1, 2, \ldots$. We use induction in k. The statement is true for L_0, because e.g. for h_1 we have:

$$h_1\left(\varrho_2(z^{(1)}, z^{(2)})\right) = h_1\left(\varrho_2(z_1^{(1)}, z_1^{(2)}), \ldots \varrho_2(z_n^{(1)}, z_n^{(2)})\right) =$$
$$= \varrho_2(z_1^{(1)}, z_1^{(2)}) = \varrho_2\left(h_1(z^{(1)}), h_1(z^{(2)})\right).$$

Suppose we know the statement for L_k and prove it for L_{k+1}. Let $h', h'' \in$ $\in L_k$, $h := \varrho_1(h', h'') \in L_{k+1}$. We have to prove that

$$h(\varrho_2(z_1, z_2)) \leq \varrho_2(h(z_1), h(z_2)), \qquad \text{i.e.}$$
$$\varrho_1(h', h'')(\varrho_2(z_1, z_2)) \leq \varrho_2(\varrho_1(h', h'')(z_1), \varrho_1(h', h'')(z_2))$$

which is equivalent (after some transformation) to

$$\varrho_1(h'(\varrho_2(z_1, z_2)), h''(\varrho_2(z_1, z_2))) \leq$$
$$\leq \varrho_2\left(\varrho_1(h'(z_1), h''(z_1)), \varrho_1(h'(z_2), h''(z_2))\right).$$

According to our induction hypothesis

$$h'(\varrho_2(z_1, z_2)) \leq \varrho_2(h'(z_1), h'(z_2)), \quad h''(\varrho_2(z_1, z_2)) \leq \varrho_2(h''(z_1), h''(z_2)),$$

whence by the monotonity of ϱ_2 and using also (3) we obtain

$$\varrho_1\left(h'(\varrho_2(z_1, z_2)), h''(\varrho_1(z_1, z_2))\right) \leq$$
$$\leq \varrho_1\left(\varrho_2(h'(z_1), h'(z_2)), \varrho_2(h''(z_1), h''(z_2))\right) \leq$$
$$\leq \varrho_2\left(\varrho_1(h'(z_1), h''(z_1)), \varrho_1(h'(z_2), h''(z_2))\right).$$

Thus the Sublemma is proved for every $h \in Co\varrho_1(\{h_1, \ldots, h_n\}) = \bigcup_{k=0}^{\infty} L_k$.

Now, if

$$h \in \overline{Co\varrho_1(\{h_1, \ldots, h_n\})} \Rightarrow \exists h^{(k)} \in Co\varrho_1(\{h_1, \ldots, h_n\}) \ : \ h^{(k)} \to h$$

uniformly. According to the considerations above we have

$$h^{(k)}(\varrho_2(z_1, z_2)) \leq \varrho_2(h^{(k)}(z_1), h^{(k)}(z_2)),$$

hence using the continuity of ϱ_2, we obtain:

$$h^{(k)}(\varrho_2(z_1, z_2)) \to h(\varrho_2(z_1, z_2))$$

and

$$\varrho_2(h^{(k)}(z_1), h^{(k)}(z_2)) \to \varrho_2(h(z_1), h(z_2))$$

whence taking the limits as $k \to \infty$: $h(\varrho_2(z_1, z_2)) \leq \varrho_2(h(z_1), h(z_2))$ follows. Thus the Sublemma is proved. Now we can prove that $N_\gamma(h)$ is ϱ_2-convex, i.e. $z_1, z_2 \in N_\gamma(h) \Rightarrow \varrho_2(z_1, z_2) \in N_\gamma(h)$. Because of $z_1, z_2 \in N_\gamma(h)$ we have $h(z_1) \leq \gamma$, $h(z_2) \leq \gamma$ and hence $\varrho_2(h(z_1), h(z_2)) \leq \gamma$. Using the Sublemma we obtain: $h(\varrho_2(z_1, z_2)) \leq \varrho_2(h(z_1), h(z_2)) \leq \gamma$ follows, thus Lemma 5 is proved. □

LEMMA 6. *If $T \subset K$ is any ϱ_2-convex, compact set then it is not separable with disjoint closed sets M, N.*

PROOF OF LEMMA 6. Suppose indirectly that T is separable, then there exist $T_1, T_2 \subset K$ such that $T_1 \cup T_2 = T$, $T_1 \subset N$, $T_2 \subset M$, T_1 and T_2

are compact. Hence $d(T_1,T_2)=:a>0$ and $\exists t_1\in T_1,\ t_2\in T_2: d(t_1,t_2)=a$. But $d(t_1,\varrho_2(t_1,t_2))<a,\ d(t_2,\varrho_2(t_1,t_2))<a \Rightarrow \varrho(t_1,t_2)\in T_1$ and $\varrho_2(t_1,t_2)\in T_2$ and this contradicts $T_1\cap T_2=\emptyset$. Lemma 6 is proved. □

LEMMA 7. *There exist* $h^{(1)},h^{(2)}\in L$ *and* $\delta\in(\beta,\gamma)$ *such that* $N_\delta(h^{(1)})\cap \cap N_\delta(h^{(2)})=\emptyset$.

PROOF OF LEMMA 7. We know that $\exists h^*\in L: N_\gamma(h^*)\subsetneq R$. But R is minimal in F hence $N_\gamma(h^*)\notin F$. Consequently some of conditions in the definition of F are not fulfilled. (14) and (16) are fulfilled, hence (15) does not hold, i.e. $\exists h'\in L: \inf\limits_{z\in N_\gamma(h^*)} h'(z)>\beta$. Choose any δ such that $\inf\limits_{z\in N_\gamma(h^*)} h'(z)>\delta>\beta$. Then $N_\gamma(h^*)\cap N_\delta(h^*)=\emptyset$. Because $\delta<\gamma$, hence $N_\delta(h^*)\subset N_\gamma(h^*)$ and we obtain: $N_\delta(h')\cap N_\delta(h^*)=\emptyset$. Thus Lemma 7 is proved with $h^{(1)}:=h',\ h^{(2)}:=h^*$. □

Let $h^{(1)},h^{(2)}\in L$ such that for $N_\delta(h^{(1)})=N,\ N_\delta(h^{(1)})=N',\ N\cap N'=\emptyset$ is fulfilled. Let $S:=\{h\in L: N_\delta(h)\subset N'\}$, $P:=\{h\in L: N_\delta(h)\subset N\}$. $S\neq\emptyset$ because $h^{(2)}\in S$ and $P\neq\emptyset$ because $h^{(1)}\in P$. Define $s:=\sup\limits_{h\in S}\inf\limits_{z\in K} h(z)$, $p:=\sup\limits_{h\in P}\inf\limits_{z\in K} h(z)$. Obviously $s\le\beta,\ p\le\beta$. Let $r:=\min\limits_{\substack{|x-y|\ge\delta-\beta\\ |x|,|y|<c}}\frac{\varrho_1(x,y)-x\wedge y}{|x-y|}$ where c is a bound for f, i.e. $|f|<c$. Let $\varepsilon>0$ be such that $\frac{\varepsilon}{\delta-\beta}<r$. Then $\exists h'\in S$: $\inf\limits_{z\in K} h'(z)>s-\varepsilon$ and $\exists h''\in P$: $\inf\limits_{z\in K} h''(z)>p-\varepsilon$. Let $h=\varrho_1(h',h'')$. Then $N_\delta(h)\subset N_\delta(h')\cup N_\delta(h'')\subset N\cup N'$. But $N_\delta(h)\neq\emptyset$ is connected, hence, because of $N\cap N'=\emptyset$ we have $N_\delta(h)\subset N$ or $N_\delta(h)\subset N'$. Suppose $N_\delta(h)\subset N'$. Then $h\in S$. Let z_0 be such that $h(z_0)=\min\limits_{z\in K} h(z)$. Obviously, $z_0\in N'$ and $h'(z_0)<h(z_0)\le\beta<\delta<h''(z_0)$. Because

$$|h'(z_0)-h''(z_0)|>\delta-\beta,$$

hence according to the definition of r we have

$$r\le\frac{\varrho_1(h'(z_0),h''(z_0))-h'(z_0)}{h''(z_0)-h'(z_0)}=\frac{h(z_0)-h'(z_0)}{h''(z_0)-h'(z_0)},$$

$$\varepsilon<r(\delta-\beta)<r(h''(z_0)-h'(z_0))\le h(z_0)-h'(z_0),$$

$$h'(z_0)+\varepsilon\le h(z_0)=\inf_{z\in K} h(z).$$

But $h'(z_0)+\varepsilon \le \inf_{z\in K} h(z)+\varepsilon > s$ whence $\inf_{z\in K} h(z) > s$ but this contradicts $h \in S$. That is $N_\delta(h) \subset N'$ is not fulfilled. Similarly, we obtain a contradiction if we assume $N_\delta(h) \subset N$. □

At last we show that the statement of Theorem 2 remains valid if we omit condition (3) but instead of (5) and (6) we assume the following "mixed" ones.

By the method of the proof of Theorem 2 we can prove also the following theorem.

THEOREM 3. *Suppose the condition of Theorem 2 are fulfilled, but (5) and (6) are replaced by*

$$(5') \qquad \forall y_1, y_2 \in Y \exists y_3 \in Y : g(x,y_3) \le \varrho_2(f(x,y_1), g(x,y_2)) \quad \forall x \in X,$$

$$(6') \qquad \forall x_1, x_2 \in X \exists x_3 \in X : f(x_3,y) \ge \varrho_1(f(x_1,y), g(x_2,y)) \quad \forall y \in Y.$$

Then for any finite set $H \subset X$ we have (8).

REMARK. Condition (3) is necessary in proving Theorem 2. Indeed, define the means

$$\varrho_1(a,b) := \frac{3}{4}(a \vee b) + \frac{1}{4}(a \wedge b), \quad \varrho_2(a,b) := \frac{1}{4}(a \vee b) + \frac{3}{4}(a \wedge b).$$

Then taking the functions f, g constructed in [11] as a counterexample we can easily see that all conditions of Theorem 2 (without (3)) fulfil and the conclusion (8) does not hold. On the other hand it is not clear whether the monotonity (1) of the means is necessary in theorems 1 and 2 or not.

References

[1] K. FAN, Minimax Theorems, *Proc. Nat. Acad. Sci. U.S.A.* **39** (1953), 42–47.

[2] M. HORVÁTH – A. SÖVEGJÁRTÓ, On convex functions, *Ann. Univ. Sci. Budapest.*, Sect. Math., **29** (1986), 193–198.

[3] M. HORVÁTH – I. JOÓ, On Ky Fan-convexity (in Hungarian), *Mat. Lapok*, **34** (1987), 137–140.

[4] I. IRLE, A General Minimax Theorem, *Zeitschrift für Operations Research.* **29** (1985), 229–247.

[5] I. JOÓ, Answer to a problem of M. Horváth and A. Sövegjártó, *Ann. Univ. Sci. Budapest.* Sect. Math., **24** (1986), 203–207.

[6] I. JOÓ, A general minimax theorem, *Publ. Math. Debrecen*, **41** (1992), 1–4.

[7] I. JOÓ, On some two functions minimax theorems, to appear.

[8] I. Joó, Minimax theorems for König-convex functions, to appear.

[9] I. Joó, On some minimax theorems of S. Simons, to appear.

[10] G. Kassay, A simple proof for König's minimax theorem *Ann. Univ. Sci. Budapest.*, Sect. Math., to appear.

[11] I. Joó and G. Kassay, Convexity, minimax thoerems and their applications, to appear.

[12] H. König, Über das von Neumannsche Minimax-Theorem, *Acta Math.*, **19** (1968), 482–487.

[13] S. Simons, *Minimax and Variational inequalities.* Are they of fixed-point or Hahn–Banach type?, Game Theory and Mathematical Economics, 379–389. North-Holland Publishing Company, 1981.

[14] S. Simons, Two-function minimax theorems and variational inequalities for functions on compact and non-compact sets, with some comments on fixed-point theorems, *Proceedings of Symposia in Pure Mathematics* **45** (1986), 377–392.

[15] S. Simons, Minimax theorems with staircases, to appear.

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 171–178

ON A PROBLEM OF B. PIOCHI ON PERMUTABLE SEMIGROUPS

By

C. GUTAN and M. GUTAN

Faculty of Mathematics "Al. I. Cuza" University, Iasi

(Received December 9, 1992)

1. Introduction

In 1902 W. BURNSIDE formulated the following question: "Are the finitely generated and periodic groups finite?", question called in the group theory the general Burnside problem. The answer to this question is generally negative and this was proved by E. GOLOD in 1964. But a first important result for this problem was obtained by I. SCHUR (1911) who showed that any periodic and finitely generated group of matrices over the field of the complex numbers $\mathbb{C}$ is finite. More than that, in 1965 I. KAPLANSKI proved that the above statement is also true when the field $\mathbb{C}$ is replaced by an arbitrary field. In 1975 MCNAUGHTON and Y. ZALCSTEIN showed that any semigroup of matrices over a field, periodic and finitely generated, is finite. We also mention a result established by S. I. ADJAN, P. S. NOVIKOV (1968) and J. L. BRITTON (1973). They proved that, for each natural number n, $n \geq 72$, there exists infinite groups with two generators satisfying the relation $x^n = 1$ for every element x of the group.

In connection with the general Burnside problem (for groups or semigroups) a question is raised about the properties P which have to be added to the conditions of being periodic and finitely generated in order to obtain the finiteness of the group (semigroup). Such a property for semigroups was introduced by A. RESTIVO and C. REUTENAUER ([21]) namely the permutability property.

If n is a natural number, $n \geq 2$, the semigroup S is called *n-permutable* if, for each $(x_1, \ldots, x_n) \in S^n$, there exists a permutation σ of $\{1, \ldots, n\}$, $\sigma \neq \mathrm{id}$, such that $x_1 x_2 \ldots x_n = x_{\sigma(1)} x_{\sigma(2)} \cdots x_{\sigma(n)}$.

If S is n-permutable we denote this by $S \in P_n$. The semigroup S is called *permutable* if $S \in \bigcup_{n\geq 2} P_n$. In [21] A. RESTIVO and C. REUTENAUER prove that a finitely generated and periodic semigroup is finite if and only if it is permutable.

The study of permutability property for groups was initiated by M. CURZIO, P. LONGOBARDI and M. MAJ in [2] and continued by many other authors ([1], [3], [4], [7], [8], [9], [10], [11], [12], [19]).

In [3] it is proved that a group G is permutable if and only if there exists in G a normal subgroup N of finite index such that N' is finite.

The groups G with the property P_3 are characterized by the condition $|G'| \leq 2$. In [10] the groups with the property P_4 are classified.

The *index of permutability* of a permutable semigroup S is the natural number

$$k(S) = \min\{n \mid S \in P_n\}.$$

Let $\mathbf{C}$ be a class of finite semigroups having the property that for each non-zero natural number m there exists in $\mathbf{C}$ a semigroup with m elements. We define the sequence of natural numbers $(k_{\mathbf{C}}(m)_{m\geq 1}$ by

$$k_{\mathbf{C}}(m) = \max\{k(S) \mid S \in \mathbf{C} \text{ and } S \text{ has } m \text{ elements}\}.$$

For $\mathbf{C} = \mathbf{G} =$ the class of finite groups, the numbers $k_{\mathbf{G}}(m)$, $1 \leq m \leq 32$, are known ([19]). It is also known that $\lim_{m\to\infty} \frac{1}{m} k_{\mathbf{G}}(m) = 0$. In [16] B. PIOCHI tackles an analogous problem for $\mathbf{C} = \mathbf{S} =$ the class of all finite semigroups. It is known that the sequence $(k_{\mathbf{S}}(m))_{m\geq 1}$ is increasing and if the sequence $\left(\frac{1}{m} k_{\mathbf{S}}(m)\right)_{m\geq 1}$ has a limit then this limit is less than or equal to 4/5 ([16]). In [6] it is established that $k_{\mathbf{S}}(5) = 4$ and $k_{\mathbf{S}}(6) = 5$.

In this paper it is proved that $k_{\mathbf{S}}(7) = k_{\mathbf{S}}(8) = 6$ and that if the sequence $\left(\frac{1}{m} k_{\mathbf{S}}(m)\right)_{m\geq 1}$ has a limit then this limit is less than or equal to 2/3.

2. Conditions under which the product of four elements of a semigroup cannot be reordered

Let S be a semigroup. For any non-zero natural number n and $(a_1, \ldots, a_n) \in S^n$ we denote

$$F(a_1, \ldots, a_n) = \{a_i a_{i+1} \cdots a_j \mid 1 \leq i \leq j \leq n\}.$$

Consider a, b, c, d four distinct elements of a finite semigroup such that the product $abcd$ cannot be reordered. In the sequel we shall study this situation when the supplementary condition that $F(a,b,c,d)$ has exactly five elements is satisfied.

2.1. PROPOSITION ([6]). *If* a, b, c *are elements of a finite semigroup such that the product* abc *cannot be reordered and* $F(a,b,c) \subset \{a,b,c\}$ *then* $ab = b = bc$. □

2.2. PROPOSITION ([16]). *Let* a, b, c, d *be four elements of a finite semigroup such that the product* $abcd$ *cannot be reordered. Then at least one of the products* ab, bc, cd, abc, bcd *is not in* $\{a,b,c\}$. □

We now suppose that a, b, c, d are elements of a finite semigroup such that the product $abcd$ cannot be reordered. Suppose in addition that $F(a,b,c,d) \subset \{a,b,c,d\}$. We shall prove that only the following situations are possible:

(α)	$ab = bc = b$,	$cd = d$,	$bcd = e$;
(α')	$ab = a$,	$bc = cd = c$,	$abc = e$;
(β)	$ab = a$,	$bc = e$, $cd = c$,	$abc = c$;
(β')	$ab = b$,	$bc = e$, $cd = d$,	$bcd = b$;
(γ)	$ab = a$,	$bc = e$, $cd = c$,	$abc = e$;
(γ')	$ab = b$,	$bc = e$, $cd = d$,	$bcd = e$;
(δ)	$ab = a$,	$bc = e$, $cd = d$,	$abc = bcd = e$;
(ε)	$ab = b$,	$bc = e$,	$cd = c$.

2.3. LEMMA. *Let* a, b, c, d, e *be elements of a finite semigroup such that the product* $abcd$ *cannot be reordered. If* $F(a,b,c,d) \subset \{a,b,c,d,e\}$ *and* $bc \in \{b,c\}$ *then only one of the cases* (α) *or* (α') *can exist.*

PROOF. We shall analyse the case $bc = b$. The case $bc = c$ can be obtained by symmetry. As the product $abcd$ cannot be reordered it follows that $ab \in \{a,b,e\}$ and $cd \neq ab$. By Proposition 2.1 we deduce that $ab \neq a$. We prove now that $ab \neq e$. Suppose $ab = e$. By hypothesis and Proposition 2.1 we have $cd = d$ and $bcd = bd = e$. Thus $bcd = e = ec = bdc$, a contradiction. Hence $ab = b = bc$. By virtue of Proposition 2.1, $cd \neq c$ and if we assume $cd = d$ then $bcd = e$. Thus case (α) holds as required (similarly if $bc = c$ we obtain case (α')). To finish it remains to prove that $cd \neq e$. Assume that $cd = e$. Then $bcd = bd = be \neq a$. If $bd = be = b$ then $a(bcd) = b = (a(bd))c$ and we have supposed that is impossible. Consider $bd = be = c$. We may deduce that $c = be = abe = ac$ from which it immediately follows that $abc = b = bac$, in contradiction to this choice of the elements a, b, c, d. As $abcd = abd = abe$ and $ab = b$ once again by Proposition 2.1 it follows that $bd = be \notin \{d,e\}$. □

2.4. LEMMA. *Let a, b, c, d, e be elements of a finite semigroup such that either $ab=e$ or $cd=e$ and $F(a,b,c,d)\subset\{a,b,c,d,e\}$. Then the product $abcd$ can be reordered.*

PROOF. If $ab=e$ let us suppose that the product $abcd$ cannot be reordered. As a direct consequence of Lemma 2.3 it follows that $bc=e$. Then because the product cannot be reordered we have that $cd\in\{c,d\}$. If $cd=c$ we deduce $e=bc=bcd=ed$ hence $abcd=abc=((ab)d)c$ and this contradicts the earlier assumption. Hence $cd=d$ and as a consequence it follows that $bcd=bd=ed\neq a$. It remains to analyse four situations. First if $bd=ed=b$ then $b=ed=abd=ab=e$. If we consider $bd=ed=c$ we immediately obtain $c=ed=abd=ac$ hence $abcd=abd=ac=c=bd=bcd=$ $=bacd$ which is a contradiction. Now $bd=ed=d$ implies $bcd=d=cbd$ and finally if $bd=ed=e$ then $e=ed=abd=ae=abc=ec$ so $abcd=ed=e=abdc$.

We conclude that if $ab=e$ and $F(a,b,c,d)\subset\{a,b,c,d,e\}$ then the product can be reordered. A similar argument establishes the statement if the case $cd=e$. □

2.5. PROPOSITION. *Let a, b, c, d, e be five distinct elements of a finite semigroup such that the product $abcd$ cannot be reordered and $F(a,b,c,d)\subset\{a,b,c,d,e\}$. Then only the case (α), (α'), (β), (β'), (γ), (γ'), (δ), (ε) can exist.*

PROOF. By Lemma 2.4 we obtain $ab\in\{a,b\}$ and $cd\in\{c,d\}$. By virtue of Lemma 2.3 we have to analyse only the situation $bc=e$. We deduce that it is necessary to study the following cases:

I) $ab=a$, $cd=d$; III) $ab=b$, $cd=d$;

II) $ab=a$, $cd=c$; IV) $ab=b$, $cd=c$.

But we notice that IV) is exactly (ε) and that III) may be deduced from II) with a symmetrical argument. Thus only the situations I) and II) have to be considered.

I) $ab=a$, $cd=d$. We shall show that $abc=bcd=e$ (i.e. (δ) is obtained). Clearly $abc=ac=ae\neq d$. Also, by Lemma 2.1, as $abcd=acd$, it follows $ac\neq c$. We prove that $ac\neq a$ and $ac\neq b$. First we suppose that $ac=c$. We deduce that $acb=a=abc$, a contradiction because the product can be reordered. If we assume that $ac=b$ then $a^2c=a$. Now if we consider that the monogenic semigroup generated by c has the index m and the period r then we have $c^m=c^{m+r}$. It follows that $a=a^2c=a^3c^2=\ldots=a^{m+1}c^m=$ $=\ldots=a^{m+r+1}c^{r+1}=a^r a^{m+1}c^m=a^{r+1}$ so $a=a^2c=\ldots=a^{r+1}c^r=ac^r$. But $b=ac=a^2c^2=\ldots=a^rc^r=a^r$. As a consequence we obtain that $ab=ba$, once again a contradiction. Hence $abc = ac = ae = e$ and by symmetry we have $bcd = e$.

II) $ab = a$, $cd = c$. It remains to prove that $abc \in \{c, e\}$ (i.e. the cases (β) or (γ) are obtained. Obviously $bcd = bc = e = ed$ and $abc = ac = ae$. Because the product $abcd$ cannot be reordered we immediately deduce that $abc \neq d$. Also by Proposition 2.1, as $abcd = aed$, we deduce that $ae \neq a$. Finally, if we consider that $ac = b$ we have $acd = bd = ac = b$ hence $bcd = bdc$, which is impossible. So we have $abc \in \{c, e\}$ as required. □

3. The maximum index of permutability for semigroups with seven of eight elements

3.1. PROPOSITION. *Let a, b, c, d, e, f, g, h be eight distinct elements of a finite semigroup such that $F(a,b,c,d,e,f) \subset \{a,b,c,d,e,f,g,h\}$. Then the product $abcdef$ can be reordered.*

PROOF. If we assume that the product $abcdef$ cannot be reordered it results $F(a,b,c) \cap F(d,e,f) = \emptyset$. Therefore we can consider $F(a,b,c) \subset$ $\subset \{a,b,c,g\}$ and $F(d,e,f) \subset \{d,e,f,h\}$. As $cd \in \{c,d,g,h\}$, because of symmetry, it is sufficient to analyse the following two situations:

I) $cd = g$; II) $cd = c$;

I) If we consider the former case $cd = g$ we immediately deduce $ab \in$ $\in \{a,b\}$. Considering $ab = a$ we obtain $bc \in \{c,g\}$. But $bc = c$ implies $bcd = cd = g$ and also by Proposition 2.1 we have $abc = ac = g$ which is impossible. If $bc = g$ then $a(bcd) = abc = ac = ag = (ab)(cd)$. So we have to study the situations $ac \in \{a,b,c,g\}$ and it immediately follows that these are impossible. If on the contrary $ab = b$ then $F(a,b,c,d) \subset \{a,b,c,d,g\} \cup \{bcd\}$ and according to Proposition 2.5 we deduce $bcd = abcd = h = ah$. It follows $h \notin F(c,d,e,f)$ and this contradicts Proposition 2.5 because $cd = g$.

II) We examine now the latter case $cd = c$. We have that $F(a,b,c,d) \subset$ $\subset \{a,b,c,d,g\}$ and by Proposition 2.5 we deduce that $gd = g$. Also by Proposition 2.1 it results $de \in \{e,h\}$.

II.1. Consider $de = e$. Using once again Proposition 2.1 we obtain $cde = ce \notin \{c,d,e\}$. By hypothesis $abcdef$ cannot be reordered. This assures us that $ce \notin \{a,b,f\}$. More of that $ce \neq g$ (if $ce = g$ then $cde = g = ced$). We conclude that $ce = h$ from which it immediately follows $cdef = cef \neq g$. By virtue of Proposition 2.5 we have $ef = e$ and as a consequence it results $hf = h$. We also deduce (according to Proposition 2.1) that $bc \in \{c,g\}$ and with a similar reasoning we obtain in both of these cases a contradiction.

II.2. If $de = h$, by Proposition 2.1 and Proposition 2.5, we have $bc = g$ hence $abc = ag \in \{c, g\}$. Let us consider $ag = g$. Then according to Lemma 2.4 $gde = gh = ge \notin \{a, d, e, g, h\}$ and obviously $ge \neq f$. It follows that $ge \in \{b, c\}$ and studying these two situations we obtain a contradiction. Similarly it is impossible to have $abc = c$. □

3.2. COROLLARY. *Let S be a finite semigroup having the index of permutability $m+1$, where $m \geq 6$. Then the set $S \setminus \mathrm{Cen}S$ has at least $m + 3[m/6]$ elements.*

PROOF. As $k(S) = m+1$ it follows that there exists $(a_1, \dots, a_m) \in S^m$ such that the product cannot be reordered. We consider $m = 6k + i$, where $k \in \mathbb{N}^*$ and $i \in \{0, 1, \dots, 5\}$. Then

$$a_1 a_2 \dots a_m = (a_1 \dots a_6)(a_7 \dots a_{12}) \dots (a_{6k-5} \cdots a_{6k}) a_{6k+1} \cdots a_{6k+i}.$$

We denote $X_1 = F(a_1, \dots, a_6)$, $X_2 = F(a_7, \dots, a_{12})$, $\dots$, $X_k = F(a_{6k-5}, \dots, a_{6k})$ and $X_{k+1} = F(a_{6k+1}, \dots, a_{6k+i})$. By proposition 3.1 we obtain that $|X_1| \geq 9$, $|X_2| \geq 9$, $\dots$, $|X_k| \geq 9$, $|X_{k+1}| \geq i$. It is obvious that $X_r \cap X_s = \emptyset$ if $r \neq s$ and $X_r \subset S \setminus \mathrm{Cen}S$, where r, s are elements of $\{1, \dots, k+1\}$. So we immediately obtain

$$|S - \mathrm{Cen}S| \geq \sum_{i=1}^{k+1} |X_i| = 9k + i = m + 3[m/6]. \ \square$$

3.3. THEOREM. $k_S(7) = k_S(8) = 6$.

PROOF. It is known that $6 \leq k_S(7) \leq k_S(8)$ (see [6]). But by Proposition 3.2 $k_S(8) < 7$. Hence it results that $k_S(7) = k_S(8) = 6$. □

3.4. PROPOSITION. *If the sequence $\left(\frac{1}{n} k_S(n)\right)_{n \geq 1}$ has a limit for n tending to ∞ then this limit is less than or equal to $2/3$.*

PROOF. Because $\lim_{n \to \infty} k_S(n) = \infty$ we may assume that $k_S(n) = m \geq 7$. As $\frac{1}{n} k_S(n) \leq m(m - 1 + 3[(m-1)/6])^{-1}$ and $\lim_{m \to \infty} m(m - 1 + 3[(m-1)/6])^{-1} = 2/3$ we obtain the conclusion. □

In [16] B. PIOCHI proves that if the sequence $\left(\frac{1}{n} k_S(n)\right)_{n \geq 1}$ has a limit for n tending to ∞ then this limit is less than or equal to $4/5$. In spite of the fact that Proposition 3.4 improves the result established by B. PIOCHI we suppose that this result can be further improved.

In the end we formulate some open problems:

PROBLEM 1: $\lim_{n\to\infty} \frac{1}{n} k_S(n) = 0$.

PROBLEM 2: $k_S(n+1) - k_S(n) \leq 1$ for each natural number n.

PROBLEM 3: Let n be a natural number and let $k_S(n) = m$. Then, for each natural number m' with $2 \leq m' \leq m$, there exists a semigroup with n elements which has the index of permutability m'.

We notice that an affirmative answer to Problem 2 implies an affirmative answer to Problem 3. For $n \in \{1,\ldots,8\}$ we notice that the inequality given by Problem 2 is verified.

We also proposed the study of the congruences $(\varrho_n)_{n\geq 2}$ on a fixed semigroup S, with $\varrho_n = \cap\{\sigma \mid S/\sigma \in P_n\}$ = the "P_n-radical" of S. Information about the congruence ϱ_2 when S is a Clifford semigroup, a Brandt semigroup, a Reilly semigroup or even generally, for inverse semigroups has been obtained by many authors (C. BONZINI, A. CHARUBINI, J. D. P. MELDRUM, W. D. MUNN, N. R. REILLY, B. PIOCHI).

ACKNOWLEDGEMENTS. The authors thank the referee for various suggestions on this paper.

References

[1] M. BIANCHI, R. BRANDL, A. GILLIO BERTA MAURI, On the 4-th permutational property, *Arch. Math.*, **48** (1987), 281–285.

[2] M. CURZIO, P. LONGOBARDI, M. MAJ, Su di un problema combinatorio in teoria dei gruppi, *Atti. Acad. Naz. dei Lincei*, **74** (1983), 136–142.

[3] M. CURZIO, P. LONGOBARDI, M. MAJ, D. J. S. ROBINSON, A permutational property of groups, *Arch. Math.*, 44 (1985), 385–389.

[4] M. CURZIO, P. LONGOBARDI, A. RHEMTULLA, M. MAJ, Groups with many rewriteable products, Preprint Univ. di Napoli, 39 (1990).

[5] M. GARZON, Y. ZALCSTEIN, On permutation properties in groups and semigroups, *Semigroup Forum*, **35** (1987), 337–351.

[6] M. GUTAN, The index of permutability for semigroups with five or six elements, *Pure Mathematics and Applications*, ser A, **2** (1991), 45–53.

[7] J. JUSTIN, B. PIOCHI, G. PIRILLO, *On fixed-point-free permutation properties in groups and semigroups* (to appear *Ann. Mat. Pura Appl.*).

[8] P. LONGOBARDI, M. MAJ, A. RHEMTULLA, H. SMITH, Periodic groups with many permutable subgroups, Preprint Univ. di Napoli, 28 (1990).

[9] P. LONGOBARDI, M. MAJ, On groups in which every product of four elements can be reordered, *Arch. Math.*, **49** (1987), 273–276.

[10] P. LONGOBARDI, M. MAJ, S. STONEHEWER, The classification of groups in which every product of four elements can be reordered, Preprint.

[11] M. MAJ, On the derived length of groups with some permutational property, Preprint, Univ. di Napoli, 17 (1988).

[12] M. MAJ, Some remarks on groups with permutable subgroup products, *Comm. in Algebra*, **17** (10) (1989), 2539–2555.

[13] F. MIGLIORINI, J. SZÉP, Semigroups with periodic generators, *Pure Mathematics and Applications*, ser. A, **1** (1990), 39–44.

[14] F. MIGLIORINI, J. SZÉP, On finitely generated semigroups with permutation property, *Pure Mathematics and Applications*, ser. A, **1** (1990), 33–37.

[15] F. MIGLIORINI, J. SZÉP, Finitely generated semigroups and the problem of Burnside, *Pure Mathematics and Applications*, ser. A, **1** (1990), 29–32.

[16] B. PIOCHI, Two problems on permutation property in semigroups, *C. R. Acad. Sci. Paris*, **309** (1989), série I, 969–974.

[17] B. PIOCHI, A finiteness condition for finitely generated semigroups with periodic generators, *Pure Mathematics and Applications*, ser. A, **1** (1990), 49–57.

[18] B. PIOCHI, Permutability of center-by-finite groups, (to appear in *Atti. Acad. Naz. dei Lincei*).

[19] B. PIOCHI, G. PIRILLO, Sur une propriété de permutabilité des groupes finis, *C. R. Acad. Sci. Paris*, **307** (1988), série I, 115–117.

[20] B. PIOCHI, Solvability and permutation property in inverse semigroups, *Semigroup Forum*, **42** (1992), 113–116.

[21] A. RESTIVO, C. REUTENAUER, On the Burnside problem for semigroups, *J. of Algebra*, **89** (1984), 102–104.

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 179–191

THE HAUSDORFF DIMENSION OF THE BOUNDARY OF SETS

$$\left\{ z \in \mathbb{C},\ z = \sum_{k=1}^{\infty} \frac{a_k}{\alpha^k},\ 0 \le a_k < N(\alpha),\ \deg(\alpha) = 2 \right\}$$

By

I. KÖRNYEI

Computer Center, Eötvös Loránd University, Budapest

(Received December 9, 1992)

Let α be a non-real quadratic integer. The minimal polynomial of α has the form

$$x^2 + cx + N, \tag{1}$$

where $N = (c/2)^2 + m$, if c is even and $N = (c/2)^2 + m/4$, if c is odd and in both cases m is an integer at least one and $m \equiv 3\ (4)$ in the second case. N is the norm of α. The rational integers r with $0 \le r < N$ form a complete set of residues mod α in the ring $Z[\alpha]$. An element $\eta \in Z[\alpha]$ is called to be representable if $\eta = \sum_{j=0}^{K} a_j \alpha^j, \quad 0 \le a_j < N$. The digits a_j are uniquely determined for representable numbers. The elements

$$\sum_{j=0}^{k-1} a_j \alpha^j; \quad 0 \le a_j < N$$

form a complete system of residues modulo α^k in $Z[\alpha]$.

Thus we have for every $\eta \in Z[\alpha]$ and for every k, $\eta = \xi_k + \varrho_k$, where ξ_k is representable and ϱ_k is divisible by α^k.

Let H be the set of complex numbers z, representable in the form with rational integers a_k

$$z = \sum_{k=1}^{\infty} \frac{a_k}{\alpha^k}; \quad 0 \le a_k < N.$$

It is easy to see that H is compact. We shall denote the boundary of H by ∂H. The absolute values of the elements of H are at most $\sqrt{N} + 1$.

The aim of this paper is to prove the following

THEOREM. *If $c \neq 0$ and $N > 1$, then the Hausdorff dimension of ∂H is*

$$\frac{\log \xi}{\log \sqrt{N}},$$

where ξ is the greatest real root of the polynomial

$$x^3 - (|c| - 1)x^2 - (N - |c|)x - N. \tag{2}$$

The Pontrjagin–Schnirelman dimension of ∂H is equal to the Hausdorff dimension.

REMARKS. 1. The Hausdorff dimension of ∂H is one, if $\alpha = \pm i\sqrt{N}$, because in that case the set H is a rectangle.

2. The polynomial (2) has a root between $\sqrt{N}$ and N because of the equalities

$$N\sqrt{N} - (|c| - 1)N - (N - |c|)\sqrt{N} - N = -|c|(N - \sqrt{N})$$

and

$$N^3 - (|c| - 1)N^2 - (N - |c|)N - N = N^2(N - |c|) + (|c| - 1)N,$$

and of the inequalities

$$N - |c| = (|c|/2 - 1)^2 + m - 1 \geq 0$$

or

$$N - |c| = (|c|/2 - 1)^2 + m/4 - 1 \geq 0.$$

Thus we have

$$1 < \frac{\log \xi}{\log \sqrt{N}} < 2,$$

that is, ∂H is a fractal curve. The other roots of the polynomial are non-real or are negative. Their absolute values are smaller than $\sqrt{N}$.

3. S. ITO proved the above theorem, in the case, when α is a base in $Z[\alpha]$ [6]. He used the properties of some endomorphisms on free groups, adopting Dekking's method [1]. I. KÁTAI investigated the existence of the base in quadratic fields with J. SZABÓ and B. KOVÁCS in [7], [8] and [9]. The proof of the present paper is based on linear recursive forms, following a method of I. KÁTAI, which was used in the paper [5].

First we prove that all complex numbers z are representable in the form $z = \gamma + \delta$, where γ is an element of the ring $Z[\alpha]$ and δ is an element of H. The generalization of this assertion is contained in [10]. Let η_k be for every k that element of $Z[\alpha]$, which is nearest to $\alpha^k z$. We have

$\eta_k = \xi_k + \varrho_k$, where ξ_k is representable, ϱ_k is divisible by α^k and we have $\frac{\eta_k}{\alpha^k} \to z$, if $k \to \infty$. But

$$\frac{\eta_k}{\alpha^k} = \sum_{j=1}^{k} \frac{a_j^{(k)}}{\alpha^j} + \theta_k,$$

Where $\theta_k \in Z[\alpha]$. We get

$$|z - \theta_k| \leq \sqrt{N} + 1 + \delta_k,$$

where $\delta_k \to 0$, if $k \to \infty$. We have at least one $\gamma \in Z[\alpha]$, for which $\theta_{k_1} = \gamma$ holds for infinitely many k_1, because $Z[\alpha]$ is discrete, in this way z is a limit point of $H + \gamma$, that is the assertion is proved. Therefore, $\mathbb{C}$ is covered by countable many translated copy of H. BY Baire's category theorem this implies that H is of second category. As H is compact, it follows that int$H \neq \emptyset$. This immediately implies that H equals the closure of its interior. Indeed, if $\eta \in H$, then $\eta = \sum_{k=1}^{\infty} \frac{a_k}{\alpha^k}$ and the sets $\sum_{k=1}^{m} \frac{a_k}{\alpha^k} + \alpha^{-m}H$ have non-empty interior for every m.

We show that if $z \in \partial H$, then $z = \gamma + \delta$, where $\gamma \in Z[\alpha]$, $\gamma \neq 0$, and $\delta \in H$. Indeed, if $z \in \partial H$, then $z = \lim z_k$, $z_k \notin H$. Thus $z_k = \gamma_k + \delta_k$, where $\gamma_k \in Z[\alpha] \setminus \{0\}$ and $\delta_k \in H$ for every k. Since $Z[\alpha]$ is discrete, one of the sets $H + \gamma$ contains infinitely many z_k-s, it follows that $z = \gamma + \delta$, where $\gamma \neq 0$ and $\gamma \in Z[\alpha]$, $\delta \in H$, because $H + \gamma$ is compact.

For a fixed $z \in \partial H$ has a representation

$$z = \gamma + \sum_{k=1}^{\infty} \frac{b_k}{\alpha^k}; \quad 0 \leq b_k < N$$

with a non-zero γ in $Z[\alpha]$ and a representation

$$z = \sum_{k=1}^{\infty} \frac{a_k}{\alpha^k}; \quad 0 \leq a_k < N.$$

First we try to find those elements γ-s of $Z[\alpha]$, for which

$$H \cap (H + \gamma) \neq \emptyset.$$

In that case the number γ is representable in the form

$$\gamma = \sum_{k=1}^{\infty} \frac{a_k - b_k}{\alpha^k}, \tag{3}$$

where

$$0 \le a_k < N, \quad 0 \le b_k < N. \tag{4}$$

It follows from (3) and (4) that

$$|\gamma| \le \sqrt{N} + 1 \tag{5}$$

and we have a rational integer d between $-(N-1)$ and $N-1$ that

$$|\alpha\gamma - d| \le \sqrt{N} + 1. \tag{6}$$

For example, $d = a_1 - b_1$ is a possible choice. The members of the sequence, defined by the recursive formula

$$\gamma_{k+1} = \alpha\gamma_k - (a_k - b_k) \tag{7}$$

with $\gamma_1 = \gamma$, have the properties (5) and (6).

Conversely, if the members of the sequence have the properties (5) and (6) with suitable $(a_k - b_k)$-s, the intersection $H \cap (H + \gamma)$ is not empty.

We shall prove that in cases of the polynomials

$$x^2 \pm 3x + 3, \quad x^2 \pm 4x + 5, \quad x^2 \pm 5x + 7$$

we have ten possibilities and in other cases only six possibilities for the choice of the γ-s, and the order of the γ-s is not optional.

The number γ has the form

$$\gamma = a + b\alpha, \tag{8}$$

with rational integers a, b and

$$|\gamma|^2 = (a - bc/2)^2 + b^2(N - (c/2)^2) = (a - bc/2)^2 + b^2 m_1. \tag{9}$$

It follows from (1) that

$$(a + b\alpha)\alpha = -bN - (bc - a)\alpha. \tag{10}$$

By (6), we can choose a rational integer d, for which

$$|-bN - (bc - a)\alpha - d| \le \sqrt{N} + 1 \tag{11}$$

and $-N < d < N$.

It is easy to see that

$$(\sqrt{N} + 1)^2 = N + 2\sqrt{N} + 1 \le (|c|/2 + 1)^2 + (\sqrt{m_1} + 1)^2 - 1, \tag{12}$$

where $m_1 = m$ or $m_1 = m/4$ with $m \equiv 3 \bmod (4)$ and $m > 0$. As the set of the possible γ-s is symmetric about the origin, it is enough to consider the cases when $b \ge 0$.

If $b \geq 2$, it follows from (5), (9) and (12) that the inequality

$$|a - bc/2| < |c|/2 + 1, \tag{13}$$

is valid, because $4m_1 > (\sqrt{m_1} + 1)^2 - 1$ holds in our cases.

We shall prove that $b \geq 3$ is impossible. We have from (13)

$$|bc - a| \geq \frac{(b-1)|c|}{2} - \frac{1}{2}.$$

In this way $|bc - a| \geq |c| - \frac{1}{2}$ for $b \geq 3$. We have

$$|bc - a| \geq |c|. \tag{14}$$

It follows from (10) and (14) that

$$|(a + b\alpha)\alpha - d|^2 \geq (bc - a)^2 m_1 \geq c^2 m_1 = (c^2 - 3)m_1 + 3m_1. \tag{15}$$

But the inequality

$$3m_1 \geq m_1 + 2\sqrt{m_1}$$

is also true for $m_1 \geq 1$ and also

$$\frac{3}{4}(c^2 - 3) > \left(\frac{|c|}{2} + 1\right)^2,$$

for $|c| \geq 4$. At this time (11) is inconsistent with (12), if both of the inequalities $|c| \geq 4$ and $m_1 \geq 1$ are satisfied. For $m_1 = 3/4$ we make a comparison between (15) and (12). The inequality

$$c^2 \cdot \frac{3}{4} > \left(\frac{|c|}{2} + 1\right)^2 + \left(\frac{\sqrt{3}}{2} + 1\right)^2 - 1$$

holds for $|c| \geq 4$. It follows from (9) that $|\gamma|^2$ is at least $9(N - 1/4)$ for $|c| = 1$, $9(N - 1)$ for $|c| = 2$, and $9(N - 9/4)$ for $|c| = 3$. These values are greater than $(\sqrt{N} + 1)^2$ in our cases, except the last one if $N = 3$ and $b = 3$. But in the last case, by (9), $a = 5\mathrm{sg}(c)$ or $4\mathrm{sg}(c)$ and the coefficient of α in (11) has an absolute value greater than 3, so thus we have in all the cases that $b \leq 2$.

At present, we select the cases for which $b = 2$ is maybe possible. For $b = 2$ we have from (5), (6), (7) and (10) that $|a - 2c| \leq 2$.

But from (13) it follows that $|a - c| < |c|/2 + 1$, so thus $|c| - 2 < |c|/2 + 1$. Thus we have $|c| < 6$.

If $|c| \leq 2$, we make a comparison between the last term of (9) and $(\sqrt{N} + 1)^2$. For $|c| = 1$, the inequality $4N - 1 > N + 2\sqrt{N} + 1$, is valid for

$N \geq 2$. If $|c| = 2$, it holds for $N \geq 3$ that $4N - 4 > N + 2\sqrt{N} + 1$. We shall discuss later the case of the polynomials $x^2 \pm 2x + 2$.

If $|c| = 3$, we have $N = (9 + m)/4$, $a = 5\mathrm{sg}(c)$ or $4\mathrm{sg}(c)$.

If $a = 5\mathrm{sg}(c)$, then it follows from (9) that

$$|5\mathrm{sg}(c) + 2\alpha|^2 = 4 + m$$

and there is a contradiction with (5) for $m \geq 7$. $(4 + m \leq (9 + m)/4 + \sqrt{9 + m} + 1)$.

If $a = 4\mathrm{sg}(c)$ it follows from (9) that

$$|4\mathrm{sg}(c) + 2\alpha|^2 = 1 + m,$$

and it is inconsistent with (5) for $m \geq 11$. $(1 + m < (9 + m)/4 + \sqrt{9 + m} + 1)$.

Now we have two extra cases $x^2 \pm 3x + 3$ and $x^2 \pm 3x + 4$, but we shall see that the second one belongs to the general case.

If $|c| = 4$, then we have $N = 4 + m$ and

$$|2(c - \mathrm{sg}(c)) + 2\alpha|^2 = 4 + 4m$$

and there is a contradiction with (5) for $m > 1$. $(4 + 4m \leq 4 + m + 2\sqrt{4 + m} + 1)$.

We shall discuss the case

$$x^2 \pm 4x + 5$$

later.

If $|c| = 5$, then $N = (25 + m)/4$ and

$$|2(c - \mathrm{sg}(c)) + 2\alpha|^2 = 9 + m,$$

and that value contradicts (5) for $m \geq 7$. $(9 + m \leq (25 + m)/4 + \sqrt{25 + m} + 1)$. In this way $x^2 \pm 5x + 7$ is an extra case.

The general case. Except the extra polynomials, we can assume that $|b| \leq 1$. It is enough to consider those γ-s, which have the form $\gamma = a + \alpha$, if $b \neq 0$.

We have

$$(a + \alpha)\alpha = -N + (a - c)\alpha.$$

That is $|c - a| \leq 1$, because

$$a = c - \mathrm{sg}(c), \quad c, \quad c + \mathrm{sg}(c).$$

If $b = 0$, then $a = \pm 1$, because $a \cdot \alpha = a\alpha$. We have eight possibilities for the γ-s

$$1, \quad c - \mathrm{sg}(c), \quad c + \mathrm{sg}(c) + \alpha,$$

$$-1, \quad -c + \mathrm{sg}(c) - \alpha, \quad -c - \alpha, \quad -c - \mathrm{sg}(c) - \alpha,$$

but the numbers $\pm(c+\mathrm{sg}(c)+\alpha)$ are not suitable, because $(c+\mathrm{sg}(c)+ +\alpha)\alpha = -N+\mathrm{sg}(c)\alpha$, and it is impossible to add an integer with absolute value less than N, to have one number among the listed numbers, so we have only six points.

$$(16) \qquad A_1 = 1, \quad A_2 = c+\alpha, \quad A_3 = c-\mathrm{sg}(c)+\alpha,$$
$$A_4 = -1, \quad A_5 = -c-\alpha, \quad A_6 = -c+\mathrm{sg}(c)-\alpha.$$

The extra cases. We have two γ-s with $b=2$,

$$2c-\mathrm{sg}(c)+2\alpha, \quad 2c-2\mathrm{sg}(c)+2\alpha.$$

There are five numbers to discuss with $b=1$,

$$c-2\mathrm{sg}(c)+\alpha, \quad c-\mathrm{sg}(c)+\alpha, \quad c+\alpha, \quad c+2\mathrm{sg}(c)+\alpha, \quad c+\mathrm{sg}(c)+\alpha,$$

because of the equality

$$(a+\alpha)\alpha = -N+(a-c)\alpha.$$

The number $c+2\mathrm{sg}(c)+\alpha$ is not suitable, because

$$|c+2\mathrm{sg}(c)+\alpha|^2 = 2|c|+4+N > N+2\sqrt{N}+1.$$

Note that we have in all the cases under discussion

$$2|c|+3 > [2\sqrt{N}].$$

The number $c+\mathrm{sg}(c)+\alpha$ is also not possible, because

$$(c+\mathrm{sg}(c)+\alpha)\alpha = -N+\mathrm{sg}(c)\alpha,$$

and it is impossible to add an integer with absolute value less than N, to have one number among the listed ones and their products by -1.

Now, we can see that the number $2c-\mathrm{sg}(c)+2\alpha$ is not suitable, because

$$2c-\mathrm{sg}(c)+2\alpha)\alpha = -2N-\mathrm{sg}(c)\alpha$$

and the inequality

$$2N-|c| \geq N$$

is valid.

If $b=0$, then $a=\pm1$, because $a\cdot\alpha = a\alpha$, $|a|\neq 2$. Since the following inequality $|2c-2\mathrm{sg}(c)(c)| = 2(|c|-1) \geq N$ is valid in these cases, we have only ten possible numbers:

$$(17) \quad A_1 = 1, \quad a_2 = c+\alpha, \quad A_3 = c-\mathrm{sg}(c)+\alpha, \quad A_4 = c-2\mathrm{sg}(c)+\alpha,$$
$$A_5 = 2(c-\mathrm{sg}(c))+2\alpha, \quad A_6 = -1, \quad A_7 = -c-\alpha,$$
$$A_8 = -c+\mathrm{sg}(c)-\alpha, \quad A_9 = -c+2\mathrm{sg}(c)-\alpha, \quad A_{10} = -2(c-\mathrm{sg}(c))-2\alpha.$$

We return to the examination of the general cases. The order of the γ-s in the sequence (7) is not optional. If for instance $\gamma_k = A_1$, then $\gamma_{k+1} = c + \alpha$ or $c - \mathrm{sg}(c) + \alpha$ and we must choose in these cases as in place of $a_k - b_k$ the numbers $-c$, $-c + \mathrm{sg}(c)$ accordingly.

There we give a table of the possibilities

γ_k		γ_{k+1}		$a_k - b_k$
A_1		A_2		$-c$
		A_3		$-c + \mathrm{sg}(c)$
A_2		A_4		$-(N-1)$
	$c > 0$		$c < 0$	
A_3	A_5		A_2	$-(N - \lvert c\rvert)$
	A_6		A_3	$-(N - \lvert c\rvert + 1)$

because

$$A_1\alpha = \alpha, \quad A_2\alpha = -N, \quad A_3\alpha = -N - \mathrm{sg}(c)\alpha.$$

If $|c| = N$, then A_2 does not follow A_1, but the coefficient of $M(k-1, A_2)$ in the latter recurrence relation is zero. The situation of A_3 is the same for $|c| = 1$.

Let $M(k, A_j)$ denote the number of the possible values of the k-th partial sums of the points of $H \bigcap (H + A_j)$ for the sums

$$\sum_{j=1}^{\infty} \frac{a_j}{\alpha^j}.$$

Note that formally different partial sums have different values, because the representation of (the representable) elements are unique.

If the difference $a_k - b_k$ is a fixed number u with $N > u \geq 0$, then $a_k = u$, $b_k = 0$, $a_k = u + 1$, $b_k = 1, \ldots, a_k = N - 1, b_k = N - 1 - u$ are the possible values, there are $N - u$ possibilities. If $u < 0$, we have $N - |u|$ different cases.

From the table we get the following recurrence relations.

$$M(k, A_1) = (N - |c|)M(k-1, A_2) + (N - |c| + 1)M(k-1, A_3),$$
$$M(k, A_2) = M(k-1, A_1),$$
$$M(k, A_3) = |c|M(k-1, A_2) + (|c| - 1)M(k-1, A_3),$$

that is,

$$M(k, A_1) = (N - |c|)M(k-2, A_1) + (N - |c| + 1)M(k-1, A_3),$$
$$M(k, A_3) = |c|M(k-2, A_1) + (|c| - 1)M(k-1, A_3).$$

Let $F(z)$ and $G(z)$ be the generating functions

$$F(z)=\sum_{k=1}^{\infty} M(k,A_1)z^k; \quad G(z)=\sum_{k=1}^{\infty} M(k,A_3)z^k.$$

Then we have the equations,

$$F(z)(1-(N-|c|)z^2)-G(z)(N-|c|+1)z=a_1+a_2z,$$

$$-F(z)|c|z^2+G(z)(1-(|c|-1)z)=b_1+b_2z,$$

where a_1, a_2, b_1, b_2 are depending on the first two members of the $M(k,A_j)$-s. It follows that

$$F(z)=z\cdot\frac{p_1(z)}{1-(|c|-1)z-(N-|c|)z^2-Nz^3}, \tag{18}$$

$$G(z)=z\cdot\frac{p_2(z)}{1-(|c|-1)z-(N-|c|)z^2-Nz^3}, \tag{19}$$

where the polynomials p_1 and p_2 are of degree at most three.

In the extra cases, for the numbers in (17) we have the following table for the possibilities

γ_k	γ_{k+1}		a_k-b_k
A_1	A_2		$-c$
	A_3		$-c+\text{sg}(c)$
	A_4		$-c+2\text{sg}(c)$
A_2	A_6		$-(N-1)$
	$c>0$	$c<0$	
A_3	A_7	A_2	$-(N-\|c\|)$
	A_8	A_3	$-(N-\|c\|+1)$
	A_9	A_4	$-(N-\|c\|+2)$
A_4	A_{10}	A_5	$-(N-2(\|c\|-1))$
A_5	A_{10}	A_5	$-2(N-\|c\|+1)$.

The numbers in the last column have absolute values less than N, except the cases $|c|=2$, $N=2$ and $|c|=3$, $N=4$. The last number in the last column has absolute value two or four respectively, which is the norm, therefore these cases are contained in the general case.

Because of the symmetry of the γ-s in the extra cases, we get from the table the following recursive formulas

$$M(k,A_1) = (N-|c|)M(k-1,A_2) + (N-|c|+1)M(k-1,A_3) + \\ + (N-|c|+2)M(k-1,A_4),$$
$$M(k,A_2) = M(k-1,A_1),$$
$$M(k,A_3) = |c|M(k-1,A_2) + (|c|-1)M(k-1,A_3) + (|c|-2)M(k-1,A_4),$$
$$M(k,A_4) = (N-1)M(k-1,A_5),$$
$$M(k,A_5) = M(k-1,A_5).$$

It follows from the last formula that $M(k,A_5)=1$, $M(k,A_4)=N-1$, in this way H and $H+A_5$ have only one common point, and H and $H+A_4$ have common points of number $N-1$.

Let $F(z)$ and $G(z)$ be the generator function for $M(k,A_1)$ and $M(k,A_3)$, then we have

$$F(z)(1-(N-|c|)z^2) - G(z)(N-|c|+1)z = \\ = \frac{2(|c|-1)(N-|c|+2)}{1-z}z^3 + z(a_1+a_2z),$$

$$-F(z)|c|z^2 + G(z)(1-|c|+1)z = \frac{2(|c|-1)(|c|-2)}{1-z}z^3 + z(b_1+b_2z),$$

in this way we have

$$F(z) = z\cdot\frac{p_1(z)}{(1-(|c|-1)z-(N-|c|)z^2-Nz^3)(1-z)}, \tag{20}$$

$$G(z) = z\cdot\frac{p_2(z)}{(1-(|c|-1)z-(N-|c|)z^2-Nz^3)(1-z)}, \tag{21}$$

where p_1 and p_2 are polynomials of degree at most four.

It is true for polynomials that

$$x^n p(1/x) = \prod_{j=1}^{n}(1-\xi_j x),$$

where the ξ_j-s are the roots of the polynomial. We get with partial fraction expansion from (18)-(21) that

$$M(k,A_j) = c_j\xi^k + o(\xi^k), \tag{22}$$

for $j=1,2,3$, while ξ has the greatest absolute value among the roots. It is easy to prove that the c_j-s are positive.

Let E_j be $H\bigcap(H+A_j)$. We shall prove that for the κ-dimensional outer measure (see K.J. FALCONER [2] 1.2.) where $\kappa = \frac{\log\xi}{\log\sqrt{N}}$, the following inequality holds;

$$0 < \mathcal{H}^k(E_j) < \infty,$$

if $j \neq 4,5,9,10$ in the extra cases. Let us consider the discs, centered around the values of the k-th partial sums of the points of E_j, with diameter

$$2(\sqrt{N}+1)/(\sqrt{N})^k.$$

E_j is covered by these discs. Then it follows from (22) that

$$\mathcal{H}_\delta^\kappa(E_j) \leq \frac{2^\kappa(\sqrt{N}+1)^\kappa M(k,A_j)}{(\sqrt{N})^{\kappa k}} \leq O(1)$$

with $\delta = 2(\sqrt{N}+1)/(\sqrt{N})^k$, and in this way we have

$$\mathcal{H}^\kappa(E_j) < \infty. \tag{23}$$

For all k one can define an outer measure on $\mathbb{C}$. Let U be a set. Let $P_k(U)$ denote the number of open discs, which have a common point with U and whose centers are the k-th partial sum of $\sum_{k=1}^{\infty} \frac{a_k}{\alpha^k}$ for the points of E_j with diameter $2(\sqrt{N}+1)/(\sqrt{N})^k$. We define

$$\mu_k(U) = P_k(U)2^\kappa/(\sqrt{N})^{\kappa k}.$$

We now have for E_j

$$d_1 < \mu_k(E_j) < d_2$$

with positive d_1 and d_2, independent of k. If S is a disc, we have

$$\mu_k(S \cap E_j) \leq d_3 \mathrm{diam}(S)^\kappa,$$

if k is large enough.

Let l denote the integer l, for which

$$1/(\sqrt{N})^{l+1} < \mathrm{diam}(S)/2 < 1/(\sqrt{N})^l$$

holds. Let us consider the open discs, having at least one common point with S, and whose centers are an element of $Z[\alpha]/\alpha^l$, with diameter $2(\sqrt{N}+1)/(\sqrt{N})^l$. The number of these discs is at most v, where v is the maximal number of the elements of $Z[\alpha]$ in a disc of diameter $2(\sqrt{N}+1)$. Among these centers there are the values of the l-th partial sums of the representations $\sum_{k=1}^{\infty} \frac{a_k}{\alpha^k}$ of points in E_j. Consider those $l+u$-th partial sums of the points whose l-th partial sum is a given center. The number of these values is $M(u, A_w)$ with a fixed w. In this way we have

$$\mu_{l+u}(S) = P_{l+u}(S)2^\kappa/(\sqrt{N})^{(l+u)\kappa} \leq v \cdot d_3 \cdot \xi^u/(\sqrt{N})^{(l+u)\kappa} \leq$$

$$\leq v \cdot d_3 \cdot (\sqrt{N})^\kappa \cdot \mathrm{diam}(S)^\kappa.$$

We shall apply the Frostman's lemma (see [3]). If there exists an outer measure with the properties

$$\mu(E_j) > 0; \quad \mu(S \cap E_j) \leq d_4 \cdot \mathrm{diam}(S)^\kappa,$$

then

$$\mathcal{H}^\kappa(E_j) > 0,$$

because we have for all cover of E_j

$$0 < \mu(E_j) \leq \sum \mu(S_t) \leq d_4 \sum \mathrm{diam}(S_t)^\kappa.$$

The measures μ_k-s have at least one weak $*$ limit μ with these properties.

We have

$$\dim(\cup_j(H \cap (H + A_j))) = \max \dim(E_j). \tag{24}$$

This implies that

$$\partial H = \cup_j(H \cap (H + A_j)). \tag{25}$$

Indeed, if $\eta \in H \cap (H + A_j)$ and $\eta \neq \partial H$, then η is the centre of a disc D_1 being in H. There is a sequence $\eta_k \in \mathrm{int}(H + A_j)$ convergent to η because $(H + A_j)$ is the closure of its own interior, so thus we have a new disc, contained in $D \cap (H + A_j)$, but it is impossible, because $\dim E_j < 2$. We have proved above that $\partial H \subset \cup_j(H \cap (H + A_j))$ and hence (25) holds. Therefore the first part of the theorem is proved.

REMARK. Some numbers in (17) and (18) were examined by W. J. GILBERT in [4].

The Pontrjagin-Schnirelman dimension of a set is

$$\liminf_{\varepsilon \to 0} \frac{\log N_\varepsilon}{\log(1/\varepsilon)},$$

where N_ε denotes the minimal number of discs of a cover by discs of radius ε.

We have for the E_j-s that if $\varepsilon = (\sqrt{N} + 1)/\sqrt{N}^k$, then the minimal number of discs of radius ε is at most $M(k, E_j)$ and at least $1/v_1 M(k)$, where v_1 is the maximal number of elements of $Z[\alpha]$ in a discs of radius $3(\sqrt{N} + 1)$. Because the Pontrjagin-Schnirelman dimension is not smaller than the Hausdorff one, this dimension of the E_j is the same as the another one. Because (24) is valid for the Pontrjagin–Schnirelman dimension, the second part of the theorem is proved.

References

[1] F.M. DEKKING, Recurrent sets, *Adv. Math.*, **44** (1982), 78–104.

[2] K. J. FALCONER, *The geometry of fractal sets*, Cambridge Univ. Press, London, New York, 1985.

[3] O. FROSTMAN, Potential d'équilibre et capecité des ensembles avec quelques applications á la théorie des fonctions, *Meddel. Lunds Univ. Math. Sem.*, **3** (1935), 1–118.

[4] W. J. GILBERT, Complex numbers with three radix expansions, *Can. J. Math.*, **34** (1982), 335–1348.

[5] K. H. INDLEKOFER, I. KÁTAI, P. RACSKÓ, Some remarks on generalized numbers systems. (to appear.)

[6] S. ITO, On the curves induced from the complex radix expansion, *Tokyo J. Math.*, **12** (1989), 299–320.

[7] I. KÁTAI, J. SZABÓ, Canonical number systems for complex integers, *Acta Sci. Math.* (Szeged), **37** (1975), 255–260.

[8] I. KÁTAI, B. KOVÁCS, Canonical number systems in imaginary quadratic fields, *Acta Math. Acad. Sci. Hung.*, **37** (1981), 159–164.

[9] I. KÁTAI, B. KOVÁCS, Kanonische Zahlensysteme in der Theory der quadratischen Zahlen, *Acta Sci. Math.* (Szeged), **42** (1980), 99–107.

[10] I. KÁTAI, I. KÖRNYEI, On number systems in algebraic number fields, *Publ. Math. Debrecen*, **41** (1992), 1–6.

ANNALES UNIV. SCI. BUDAPEST., 36 (1993), 193–228

A CLASS OF NONLINEAR POPULATION EVOLUTION EQUATIONS IN BANACH SPACES

By

I. JOÓ and ZHANG SHENGHAI(*)

Department of Analysis, Eötvös Loránd University, Budapest

(Received December 16, 1992)

1. Introduction

In recent years there has been a rapid growth of interest in the application of the theory of functional analysis and semigroups of operators to population dynamics, very nice results can be found in J. SONG and J.-Y. YU [13], WEBB [15], METZ and DIEKMANN [9], HEIJMANS [5], ZHANG [16] and ZHANG and ZHU [17], to name just a few. On the other hand, the MCKENDRICK equation of population dynamics [15] is one of the most important age-dependent population models. It has been used to study a great many biological and physical phenomena. Its application to human population has been reported in [13], [12]. These models involve first order partial differential equation with nonlocal boundary conditions.

The model which used to study human population can be written as follows [13]:

$$(1)\qquad \begin{cases} \frac{\partial p(r,t)}{\partial t} + \frac{\partial p(r,t)}{\partial r} = -\mu(r)p(r,t), & 0 < r < r_m,\ t > 0, \\ p(r,0) = p_0(r), & 0 \le r \le r_m, \\ p(0,t) = \beta \int\limits_{r_1}^{r_2} k(r)h(r)p(r,t)dr, & t > 0 \end{cases}$$

where $p(r,t)$ is the age density distribution of the population, r denotes age, t is the time, $0 \le r \le r_m$, $t \ge 0$, r_m is the maximal age ever attained by individuals of the population. Here we assume that the specific fertility rate of females is a constant β, the female sex ratio $k(r)$ and fertility pattern

(*) On leaving from: Shaanxi Institute of Finance and Economics, Xian, China.

$h(r)$ are independent of time, $h(r)$ satisfies

$$\int_{r_1}^{r_2} h(r)dr = 1.$$

$[r_1, r_2]$ denotes the fecundity period of female, the relative mortality rate denoted by $\mu(r)$, which is a function dependent only on age and satisfies

$$\int_0^r \mu(\varrho)d\varrho < +\infty, \ r < r_m, \ \int_0^{r_m} \mu(\varrho)d\varrho = +\infty.$$

$p_0(r)$ is the initial density distribution of population.

Introducing $\Xi = L^p(0, r_m)$ $(1 \leq p < \infty)$ as the state space and the population evolution operator on Ξ as follows:

$$A : D(A) \subset \Xi \to \Xi$$

$$(A\Phi)(r) = -\frac{d\Phi(r)}{dr} - \mu(r)\Phi(r)$$

$$D(A) = \left\{ \Phi \mid \Phi, A\Phi \in \Xi, \Phi(0) = \beta \int_{r_1}^{r_2} k(r)h(r)\Phi(r)dr \right\}.$$

With the operator A at hand, equation (1) can be written as an abstract evolution equation in the Banach space Ξ:

$$\begin{cases} \frac{dp(\cdot,t)}{dt} = Ap(\cdot,t) \\ p(\cdot,0) = p_0(\cdot). \end{cases} \tag{2}$$

We have the following result [13]:

LEMMA 1. *The operator A is the infinitesimal generator of a one parameter strongly continuous semigroup $\{T(t)\}_{t\geq 0}$ of bounded operators in the Banach space and has the following asymptotic expansion*

$$T(t)\Phi(r) = C_\Phi e^{-\lambda_0 r - \int_0^r \mu(\varrho)d\varrho} e^{\lambda_0 t} + o(e^{(\lambda_0 - \varepsilon)t}) \text{ as } t \to +\infty \tag{3}$$

where λ_0 is the real eigenvalue of A, which has the maximum real part in the spectral of A, C_Φ is a constant depending only on Φ, $\varepsilon > 0$ is any positive number such that $\sigma(A) \cap \{\lambda_0 - \varepsilon < Re\lambda < \lambda_0\} = \emptyset$. About spectral of A, we have

(i) $\sigma(A) = \sigma_p(A)$;

(ii) *A has only one real eigenvalue λ_0, its algebraic multiplicity is* 1;

(iii) *$\sigma(A)$ is infinite set.*

The linear model can be used simply for forecasting and simulation or, in some extent, demonstrate exactly some aspects properties of population in the population statistics. [See Appendix.] On the other hand, for its simplicity, more complicated phenomena usually being in ecosystem cannot be explained in linear model, and some conclusion deduced from linear model is not appropriate to the practice. By this, starting from an age-dependent logistic nonlinear population model, we study, in this paper, the solution of the model in Banach space, the stability of population system and periodic and oscillation behaviour usually presenting in the ecosystem and provide more exact theoretical foundation of population analysis.

2. The Classical Solution of the System

We are interested in the following logistic growth model of population dynamics.

(4)

$$\begin{cases} \frac{\partial p(r,t)}{\partial t} + \frac{\partial p(r,t)}{\partial r} = -\mu(r)p(r,t) - Kf(N(t))p(\cdot,t), & 0 < r < r_m, t > 0 \\ p(r,0) = p_0(r) & 0 \le r \le r_m \\ p(0,t) = \beta \int_{r_1}^{r_2} k(r)h(r)p(r,t)dr, & t > 0 \end{cases}$$

where $N(t) = \int_0^{r_m} p(r,t)dr$ is the total number of population, $f(N(t))$ is called logistic growth term and $f(\xi)$ is a real function on $\mathbb{R}^+ = [0,\infty)$, and satisfies

$$f(0) = 0, \ f(\xi) \ge 0, \text{ when } \xi > 0.$$

If $K = 0$, (i.e. independent of habitant) then system (4) becomes system (1) which is well known age-dependent linear model of population evolution. We assume, in this section, that $K > 0$.

In Banach space $\Xi = L^p(0, r_m)$ $(1 \le p < \infty)$ we have the following nonlinear abstract equation:

$$
\begin{cases} \frac{dp(\cdot,t)}{dt} = Ap(\cdot,t) - Kf(N(t))p(\cdot,t) \\ p(\cdot,0) = p_0(\cdot) \\ N(t) = \int\limits_0^{r_m} p(r,t)dr. \end{cases} \tag{5}
$$

Here, assume that the population parameters in equation (5) are the same as before. μ, k, h are bounded, nonnegative and measurable functions.

We have had some results (see [16] and [17]) about the existence and uniqueness of solution of equation (5) under the conditions of the function $f(\xi)$ satisfying

$$f(0) = 0; \; f(\xi) > 0, \; \forall \xi > 0; \qquad f(\xi) \quad \text{is continuously differentiable.}$$

REMARK. ZHANG and ZHU studied the equation (5) in 1987. The results were published in [17]. GUO and CHAN use the ideas of ZHANG and ZHU studied the time delay population dynamics [23].

We will discuss, in this paper, the existence and uniqueness of solution of equation (5) under the condition of the function $f(\xi)$ satisfying

$$f(0) = 0; \; f(\xi) \ge 0, \; \forall \xi > 0; \quad f(\xi) \text{ is continuous.} \tag{6}$$

First, we should understand the concept of solution. We call $p(\cdot,t) \in \Xi$ $(0 \le t < \infty)$ a global classical solution [10] of equation (5), if

(i) $p(\cdot,t) \in C([0,\infty];\Xi)$ and $p(\cdot,t) \in C^1((0,\infty);\Xi)$;

(ii) $p(\cdot,t) \in D(A)$, $0 < t < \infty$

(iii) equation (5) is satisfied on $[0,\infty)$.

In order to prove the existence and uniqueness of solution of equation (5) under conditions (6), we need the following lemma.

LEMMA 2.1. *Let the initial density distribution of population $p_0(r) \in$ $\in D(A)$; the function f satisfies conditions (6), then equation (5) has a global classical solution if and only if there is a continuous function $N(t)$ on $[0,\infty)$ satisfying*

$$N(t) = N_q(t) e^{-K\int\limits_0^t f(N(\varrho))d\varrho}, \quad 0 \le t < \infty$$

where $N_q(t) = \int_0^{r_m} q(r,t)dr$ *and* $q(\cdot,t)$ *satisfies following linear evolution equation*

$$(7) \qquad \begin{cases} \frac{dq(\cdot,t)}{dt} = Aq(\cdot,t) \\ q(\cdot,0) = p_0(\cdot). \end{cases}$$

PROOF. It is easy to check that $p(\cdot,t)$ satisfies equation (5) if and only if $q(\cdot,t)$ satisfies equation (7), and

$$q(\cdot,t) = p(\cdot,t)e^{K\int_0^t f(N(\varrho))d\varrho}, \quad 0 \le t < \infty.$$

So, the lemma has been proved.

Now, we can prove the main theorem in this section.

THEOREM 2.2. *For nonlinear population evolution equation* (5), *if initial density distribution of population* $p_0(\cdot) \in D(A)$; *and the function* f *satisfies condition* (6), *then equation* (5) *has a global classical solution. Moreover, if the function* f *also satisfies the following condition*

$$(8) \quad |f(\xi_1) - f(\xi_2)| \le L(M)|\xi_1 - \xi_2|, \quad |\xi_i| \le M \quad (i = 1,2) \text{ for any } M > 0$$

where $L(M)$ *is a constant related with* M, *then the solution is unique.*

PROOF. Let $q(\cdot,t) = T(t)p_0(\cdot)$, where $\{T(t)\}_{t\ge0}$ is strongly continuous semigroup generated by operator A; then it is known (see [10]) that $q(\cdot,t)$ is a classical solution of equation (7). Moreover,

$$N_q(t) = \int_0^{r_m} q(r,t)dr$$

is a continuous function on $[0,\infty)$.

Let

$$Y_0 = \{\omega(t) \mid \omega \in C[0,1]; 0 \le \omega(t) \le |N_q(t)|, \ 0 \le t \le 1\}$$

then Y_0 is a convex and bounded subset of $C[0,1]$.

Define a mapping F: $C[0,1] \longrightarrow Y_0 \subset C[0,1]$

$$(9) \qquad (FU)(t) = N_q(t)e^{-\int_0^t Kf(U(\varrho))d\varrho}.$$

Since the function f is nonnegative and continuous, we have $F(Y_0) \subset Y_0$ and the mapping F is continuous.

We will show that $F(Y_0)$ is a relatively compact set.

Let $0 \le t_0 < t \le 1$, then

$$(FU)(t) - (FU)(t_0) = (N_q(t) - N_q(t_0))e^{-K\int_0^t f(U(\varrho))d\varrho} + \\ +N_q(t_0)\left(e^{-K\int_0^t f(U(\varrho))d\varrho} - e^{-K\int_0^{t_0} f(U(\varrho))d\varrho}\right).$$

It follows that

$$(10)\quad |(FU)(t) - (FU)(t_0)| \le |N_q(t) - N_q(t_0)| + |N_q(t_0)|\left|1 - e^{-M_0|t-t_0|}\right|,$$

where $M_0 = \max\limits_{0 \le \xi \le M_1} Kf(\xi)$, and $M_1 = \max\limits_{0 \le t \le 1} |N_q(t)|$.

For $0 \le t < t_0 \le 1$, (10) is also valid. From (10) $F(Y_0)$ is equicontinuous, on $[0,1]$. According to Schauder's second fixed point theorem, F has a fixed point, which is denoted by $N_1(t)$. Now we get

$$N_1(t) = N_q(t)e^{-\int_0^t Kf(N_1(\varrho))d\varrho}.$$

In the following, we extend $N_1(t)$ to $[0,\tau]$, in which τ is an integer and greater than 1.

Define a mapping on $C[1,\tau]$: $\tilde{F} : C[1,\tau] \to \tilde{Y}_0 \subset C[1,\tau]$. Here,

$$\tilde{Y}_0 = \left\{\omega(t) \mid \omega \in C[1,\tau]; 0 \le \omega(t) \le N_q(t)e^{-K\int_0^1 f(N_1(\varrho))d\varrho}, 1 \le t \le \tau\right\}$$

$$(\tilde{F}U)(t) = N_q(t)e^{-K\int_0^1 f(N_1(\varrho))d\varrho - K\int_1^t f(U(\varrho))d\varrho}, \quad 1 \le t \le \tau.$$

Similarly to above $\tilde{F}$ has a fixed point in $\tilde{Y}_0$. We denote the fixed point by $R(t)$. So, we have

$$R(t) = N_q(t)e^{-K\int_0^1 f(N_1(\varrho))d\varrho - K\int_1^t f(R(\varrho))d\varrho}, \quad 1 \le t \le \tau.$$

Let

$$N_\tau(t) = \begin{cases} N_1(t), & 0 \le t < 1 \\ R(t), & 1 \le t \le \tau \end{cases}$$

then $N_\tau(t)$ is continuous on $[0,\tau]$, and

$$N_\tau(t) = N_q(t)e^{-K\int_0^t f(N_\tau(\varrho))d\varrho}.$$

Similarly, a series of functions $\{N_n(t)\}$ can be obtained, they satisfy the following equalities:

$$N_{n+1}(t) = N_n(t), \qquad 0 \le t \le n$$

$$N_n(t) = N_q(t)e^{-K\int_0^t f(N_n(\varrho))d\varrho}, \quad 0 \le t \le n.$$

Let $N(t) = \lim_{n\to\infty} N_n(t)$, then $N(t)$ is continuous on $[0,\infty)$ and

$$N(t) = N_q(t)e^{-\int_0^t Kf(N(\varrho))d\varrho}, \quad 0 \le t < \infty.$$

Hence equation (5) has a global classical solution $p(\cdot,t)$, by virtue of Lemma 2.1, we have for $t \in [0,\infty)$

$$p(\cdot,t) = q(\cdot,t)e^{-K\int_0^t f(N(\varrho))d\varrho},$$

i.e.

$$p(\cdot,t) = T(t)p_0(\cdot)e^{-\int_0^t Kf(N(\varrho))d\varrho}. \tag{11}$$

We finished the proof of first part of the theorem. In order to show the uniqueness of the solutions, setting T is a positive number and $\overline{p}_T(\cdot,t)$ is a classical solution on $[0,T]$. Define

$$g(s) = T(t-s)\overline{p}_T(\cdot,s) \tag{12}$$

then $g \in C^1([0,T];\Xi)$.

Differentiating (12) we have

$$\frac{dg(s)}{ds} = -AT(t-s)\overline{p}_T(\cdot,s) + T(t-s)\left[A\overline{p}_T(\cdot,s) - Kf(\overline{N}_T(s))\overline{p}_T(\cdot,s)\right]$$

and so

$$\frac{dg(s)}{ds} = -T(t-s)Kf(\overline{N}_T(s))\overline{p}_T(\cdot,s) \tag{13}$$

where

$$\overline{N}_T = \int_0^{r_m} \overline{p}_T(r,t)dr.$$

From 0 to t, integrating both sides of (13), we have the following equality

$$\overline{p}_T(\cdot,t) = T(t)p_0(\cdot) - K\int_0^t T(t-s)f(\overline{N}_T(s))\overline{p}_T(\cdot,s)ds \quad (0 \leq t \leq T).$$

In the same way, we also have

$$p(\cdot,t) = T(t)p_0(\cdot) - K\int_0^t T(t-s)f(N(s))p(\cdot,s)ds \quad (0 \leq t \leq T).$$

It is clear that there exists a constant depending on T, which satisfies

$$|\overline{N}_T(s)| \leq C_1; \quad |N(s)| \leq C_1,$$

when $0 \leq s \leq T$.

From the conditions of the function f, we have

$$\|\overline{p}_T(\cdot,t) - p(\cdot,t)\|_\Xi \leq K\int_0^t \|T(t-s)f(\overline{N}_T(s))(\overline{p}_T(\cdot,s) - p(\cdot,s))\|_\Xi ds +$$

$$+ K\int_0^t \|T(t-s)(f(\overline{N}_T(s)) - f(N(t))p(\cdot,s)\|_\Xi ds \leq$$

$$\leq c\int_0^t \|\overline{p}_T(\cdot,s) - p(\cdot,s)\|_\Xi ds, \qquad 0 \leq t \leq T$$

where c is a constant depending on T, from Gronwall's inequality we get

$$p(\cdot,t) = \overline{p}_T(\cdot,t), \qquad 0 \leq t \leq T.$$

Thus, the arbitrariness of T asserts that the solution is unique.

Now, we would like to ask a question. If the initial density distribution of population $p_0 \in \Xi^+$, where $\Xi^+ = \{\Phi \in \Xi \mid \Phi(r) \geq 0 \text{ a.e. on } (0, r_m)\}$, is the solution of equation (5) also positive? The answer is affirmative. In fact, we have the following corollary:

COROLLARY 2.3. *When equation* (5) *has a positive initial density distribution* $p_0 \in \Xi^+$, *the solution of equation* (5) *is non-negative.*

PROOF. It is known [17], [2] that the strongly continuous semigroup $\{T(t)\}_{t\geq 0}$ generated by the operator A is a positive semigroup on the order Banach space $(\Xi, \Xi^+, \|\cdot\|)$. So, it is deduced that the corollary is correct.

3. The Stability, Periodic and Oscillator Behaviour of the System

The major goal in this section is to discuss the large time behaviour of the nonlinear population system (5). We assume, in this section, the function f satisfies conditions (6). We suppose that the initial density distribution $p_0(\cdot) \in D(A)$ and $p_0 \in \Xi^+$.

By virtue of (3) and (11) we have

$$p(r,t) = C_{p_0} e^{-\lambda_0 r - \int_0^r \mu(\varrho)d\varrho} \cdot e^{\int_0^t (\lambda_0 - Kf(N(\varrho))d\varrho} + \\ + o\left(e^{\int_0^t (\lambda_0 - \varepsilon - Kf(N(\varrho)))d\varrho} \right) \quad \text{as } t \to +\infty \tag{14}$$

where C_{p_0} is a constant not related with time t and $\varepsilon > 0$ satisfies

$$\sigma(A) \cap \{\lambda \mid \mathrm{Re}\lambda \geq \lambda_0 - \varepsilon\} = \{\lambda_0\}.$$

Let $\beta_{cr} = \left[\int_{r_1}^{r_2} k(r)h(r)e^{-\int_0^r \mu(\varrho)d\varrho} \right]^{-1}$ be the critical fertility rate of females. When the specific fertility rate β of females satisfies $\beta \leq \beta_{cr}$ then from [13] we have $\lambda_0 \leq 0$ in (14). And so

$$N(t) = Me^{\int_0^t (\lambda_0 - Kf(N(\varrho)))d\varrho} + o(e^{-\varepsilon t}), \qquad t \to +\infty \tag{15}$$

by virtue of (14), where M is some constant. The first term in (15) is a decreasing function of t, therefore, $\lim_{t\to\infty} N(t) = c$ exists. We observe that when $Kf(\xi) > 0$ for $\xi > 0$, we have $c = 0$. And so

PROPOSITION 3.1. *When $\beta \leq \beta_{cr}$, the solution $p(\cdot, t)$ of equation (5) is global asymptotic stable*

$$\lim_{t\to\infty} \|p(\cdot,t)\|_{L^1(0,r_m)} = \lim_{t\to\infty} N(t) = c < +\infty.$$

In the following we discuss mainly the case in which $\beta > \beta_{cr}$. In this case, we have $\lambda_0 > 0$ in (14). If for any $\xi \geq 0$, $Kf(\xi) < \lambda_0$, then by virtue of the asymptotic expansion

$$N(t) = Me^{\int_0^t [\lambda_0 - Kf(N(\varrho))]d\varrho} + o(e^{-\varepsilon t}) \cdot e^{\int_0^t [\lambda_0 - Kf(N(\varrho))]d\varrho}, \quad t \to \infty \tag{16}$$

$\lim_{t\to\infty} N(t)$ exists. If in addition we assume $\text{mes}\{r \mid p_0(r) \neq 0,\ r \in [r_1, r_2]\} > 0$ then $M > 0$ in (16), and

$$\lim_{t\to\infty} \|p(\cdot,t)\|_{L^1(0,r_m)} = \lim_{t\to\infty} N(t) = +\infty.$$

Generally, in mathematics there is a possibility that

$$\lim_{\xi\to\infty} Kf(\xi) = +\infty \tag{17}$$

But in practice, the situation (17) is impossible. Later we shall discuss it in detail. For the continuity reason we suppose that there exists a $\xi_0 > 0$ such that

$$Kf(\xi_0) = \lambda_0 \quad \text{and} \quad \int_0^{r_m} |T(t)p_0(r)|dr \neq 0.$$

If we write $q(r,t) = T(t)p_0(r)$, $N_q(t) = \|q(\cdot,t)\|_{L^1(0,r_m)}$, then

$$\begin{cases} p(r,t) = q(r,t)e^{-K\int_0^t f(N(\varrho))d\varrho} \\ N(t) = N_q(t)e^{-K\int_0^t f(N(\varrho))d\varrho} \end{cases} \tag{18}$$

Let $g(t) = \dfrac{N_q'(t)}{N_q(t)} - \lambda_0$, then we notice the fact that

$$\lim_{t\to\infty} g(t) = 0. \tag{19}$$

We adopt the idea of [15] and discuss the limit set of equation (18). Let

$$\Omega(N) = \{N^* \mid \exists t_n \to \infty \text{ such that } N(t_n) \to N^* \quad (n \to \infty)\}.$$

Take arbitrary $N^* \in \Omega(N)$, (N^* may be infinite) then there exists $t_n \to \infty$ such that $\lim_{n\to\infty} N(t_n) = N^*$. First we consider the situation for which $N^* < +\infty$. If $Kf(N^*) < \lambda_0$, then there are a small positive number $\varepsilon > 0$ and S_ε, a neighbourhood of N^* such that $Kf(\xi) < \lambda_0 - \varepsilon$, and $\lambda_0 - \varepsilon > 0$ provided $\xi \in S_\varepsilon(N^*)$. Suppose that $N(t_n) \in S_\varepsilon(N^*)$ for $n > N_\varepsilon$, then there is a $t_0 > N_\varepsilon$ such that $g(t) > -\varepsilon$ for $t \geq t_0$. Choose $N > N_\varepsilon$ such that $t_n \geq t_0$ for $n \geq N$ and then consider the equation

$$\frac{d\underline{N}(t)}{dt} = [\lambda_0 - Kf(\underline{N}(t))]\underline{N}(t) - \varepsilon\underline{N}(t), \quad t_n \leq t$$

$$\underline{N}(t_n) = N(t_n), \qquad n > N.$$

Since

$$\frac{dN(t)}{dt} = [\lambda_0 - Kf(N(t))]N(t) + g(t)N(t),$$

$$\frac{d}{dt}[N(t) - \underline{N}(t)]\Big|_{t=t_n} > 0 \quad \text{and} \quad [N(t) - \underline{N}(t)]\Big|_{t=t_n} = 0$$

in some neighbourhood of t_n, along which t is increasing, $N(t) > \underline{N}(t)$. We assert at this moment that

$$N(t) > \underline{N}(t), \quad \text{for } t > t_n. \tag{20}$$

Indeed, if there is a $\hat{t} > t_n$, such that $N(\hat{t}) = \underline{N}(\hat{t})$ but $N(t) > \underline{N}(t)$ for any $t \in (t_n, \hat{t})$, then it will lead to a contradiction since $\frac{d}{dt}[N(t) - \underline{N}(t)]\Big|_{t=\hat{t}} > 0$, and thus (20) holds. On the other hand, $\underline{N}(t)$ satisfies

$$\underline{N}(t) = N(t_n)e^{\int_{t_n}^{t}[\lambda_0 - \varepsilon - Kf(\underline{N}(\varrho))]d\varrho}, \quad t \geq t_n. \tag{21}$$

If the function f satisfies following condition

$$\begin{cases} \text{For any } M > 0, \text{ there is a constant } L(M) \text{ depending on } M \text{ such that} \\ |f(\xi_1) - f(\xi_2)| \leq L(M)|\xi_1 - \xi_2|, \qquad |\xi_i| \leq M, \quad i = 1,2 \end{cases}$$

then we knew that $\underline{N}(t)$ satisfying equation (21) is uniquely determined. Since $Kf(\underline{N}(t)) < \lambda_0 - \varepsilon$ in a neighbourhood of t_n, along which t is increasing, $\underline{N}(t)$ is a strictly monotone increasing function. Assume that there is a t^* such that

$$Kf(\underline{N}(t^*)) = \lambda_0 - \varepsilon, \qquad Kf(\underline{N}(t)) < \lambda_0 - \varepsilon \text{ for } t \in t_n t^*)$$

then there is a $t_1 < t^*$ such that $N(t_1) = \underline{N}(t^*)$ thus $Kf(N(t_1)) = \lambda_0 - \varepsilon$, which is impossible. Thus for all $t \geq t_n$, $Kf(\underline{N}(t)) \neq \lambda_0 - \varepsilon$ and hence $\lim_{t\to\infty} \underline{N}(t) = +\infty$, and this leads to contradiction (see (20)). Similarly, $Kf(N^*) > \lambda_0$ is also impossible, and hence we get

$$\Omega(N) \setminus \{\infty\} \subset \{\xi \mid Kf(\xi) = \lambda_0\}. \tag{22}$$

Let $\hat{a} = \underline{\lim}_{t\to\infty} N(t)$, $\hat{b} = \overline{\lim}_{t\to\infty} N(t)$, if $\hat{a} \neq \hat{b}$, it is clear that for any $N^* \in (\hat{a}, \hat{b})$, there exists a sequence $\{t_n\}$, $t_n \to \infty$, such that $N(t_n) = N^*$. (Notice that $N(t)$ is continuous.) There must be an interval $[a, b]$, $a \neq b$, $b < +\infty$ and

$$[a, b] \subset \Omega(N) \setminus \{\infty\}. \tag{23}$$

Since $b-a>0$, we can take $\varepsilon>0$ and $\xi_0\in(a,b)$ such that $[\xi_0-\varepsilon,\xi_0+\varepsilon]\subset\subset(a,b)$ and then take t_n, such that $t_n\to\infty$ $(n\to\infty)$ and $N(t_n)=\xi_0$ (this is possible). We notice that

$$(24)\qquad N(t)=[M+o(e^{-\varepsilon t})]e^{\int\limits_0^t[\lambda_0-Kf(N(\varrho))]d\varrho}=\overline{q}(t)e^{\int\limits_0^t[\lambda_0-Kf(N(\varrho))]d\varrho}$$

where $\overline{q}(t)=N_q(t)e^{-\lambda_0 t}$. From (24) we have $\lim_{t\to\infty}\overline{q}(t)=M$ and we knew that for $t\ge t_n$

$$N(t)=\xi_0\cdot\frac{\overline{q}(t)}{\overline{q}(t_n)}\cdot e^{\int\limits_{t_n}^t[\lambda_0-Kf(N(\varrho))]d\varrho}$$

if n is large enough such that for any $t\ge t_n$

$$\xi_0\cdot\frac{\overline{q}(t)}{\overline{q}(t_n)}\in(\xi_0-\varepsilon,\ \xi_0+\varepsilon).$$

We want to show for any $t\ge t_n$

$$(*)\qquad f(N(t))=\lambda_0.$$

If $(*)$ is not right then there exists t^* and t_1 such that $f(N(t))=\lambda_0,\quad t\in\in[t_n,t^*]$ and

$$f(n(t))>\lambda_0,\quad t\in(t^*,t_1)\quad(\text{or } f(N(t))<\lambda_0,\quad t\in(t^*,t_1)).$$

We can take $\overline{t}\in(t^*,t_1)$ such that

$$a<(\xi_0-\varepsilon)e^{\int\limits_{t^*}^{\overline{t}}(\lambda_0-f(N(\varrho)))d\varrho}\le N(\overline{t})\le(\xi_0+\varepsilon)e^{\int\limits_{t^*}^{\overline{t}}(\lambda_0-f(N(\varrho)))d\varrho}<b,$$

hence $f(N(\overline{t}))=\lambda_0$, this is a contradiction. Then for any $t\ge t_n$ we have

$$N(t)=\xi_0\cdot\frac{\overline{q}(t)}{\overline{q}(t_n)}$$

and so $\lim_{t\to\infty}N(t)=\xi_0$ this is contradiction, thus $\hat{b}-\hat{a}=0$, i.e. $\Omega(N)==\{N^*\}$.

Summarising, we get

THEOREM 3.2. *For any* $Kf(\xi)$ *which satisfies conditions* (6) *and* (8) *and* $p_0(\cdot)\in D(A)\cap\Xi^+$ *and* $\|p_0(\cdot)\|_{L^1}\ne 0$, *the limit of the solution of equation* (5) *as* t *goes to infinite exists, i.e.*

$$(25)\qquad \lim_{t\to\infty}p(r,t)=ce^{-\lambda_0 r-\int\limits_0^r\mu(\varrho)d\varrho}$$

where $c \geq 0$ may be infinite.

In the following, we want to find some conditions which make $N(t)$ bounded. Suppose that

$$\overline{\lim_{\xi\to\infty}}\, Kf(\xi) < \lambda_0, \tag{26}$$

taking $p_0(r) = ce^{-\lambda_0 r - \int_0^r \mu(\varrho)d\varrho}$, where $c > 0$ such that $Kf(\xi) < \lambda_0$ for $\xi > M$ in which

$$M = Cp_0 \int_0^{r_m} e^{-\lambda_0 r - \int_0^r \mu(\varrho)d\varrho} dr$$

then

$$N(t) = Me^{\int_0^t [\lambda_0 - Kf(N(\varrho))]d\varrho} + o(e^{-\varepsilon t}) \cdot e^{\int_0^t [\lambda_0 - Kf(N(\varrho))]d\varrho}. \tag{27}$$

The first term corresponding to $p_0(r)$ of (27) is a strictly monotone increasing function and $\lim_{t\to\infty} N(t) = \infty$. Therefore, in order $N(t)$ to be bounded, it must be $\overline{\lim}_{\xi\to\infty} Kf(\xi) \geq \lambda_0$.

If $\overline{\lim\limits_{\xi\to\infty}}\, Kf(\xi) > \lambda_0$ and $\lim\limits_{t\to\infty} N(t) = \infty$, then take $\varepsilon > 0$ and $t_n \to \infty$ such that $Kf(N(t_n)) > \lambda_0 + \varepsilon$, for this ε there is a $t_0 \geq 0$ such that the corresponding $g(t) < \varepsilon$ for all $t \geq t_0$, suppose that $t_n \geq t_0$ for all $n > N$, considering the limiting equation

$$\begin{cases} \frac{d\overline{N}(t)}{dt} = [\lambda_0 - Kf(\overline{N}(t))]\overline{N}(t) + \varepsilon\overline{N}(t), & t \geq t_n \\ \overline{N}(t_n) = N(t_n), & n \geq N. \end{cases}$$

Similar argument to (20), we can get $N(t) \leq \overline{N}(t)$ for all $t \geq t_0$, but $\overline{N}(t)$ is a monotone decreasing function, it is obvious that $N(t) \leq \overline{N}(t) \leq N(t_n)$, hence $N(t)$ is bounded, this is a contradiction. If $\overline{\lim\limits_{\xi\to\infty}}\, Kf(\xi) = \lambda_0$ and $Kf(\xi) \geq \lambda_0$ for all ξ large enough, it can also be deduced that $N(t)$ is bounded.

If $\overline{\lim\limits_{\xi\to\infty}}\, Kf(\xi) = \lambda_0$ and $Kf(\xi) < \lambda_0$ for all sufficiently large ξ, then $N(t)$ is unbounded like (26). Summarising the above, we have

THEOREM 3.3. *Assume that the conditions in Theorem 3.2 are satisfied. The necessary condition to make the solution of equation (5) be bounded is that*

$$\overline{\lim_{\xi\to\infty}}\, Kf(\xi) \geq \lambda_0.$$

Furthermore, if $\overline{\lim\limits_{\xi\to\infty}} Kf(\xi) > \lambda_0$, *no unbounded solution of equation* (5) *exists. In the case* $\overline{\lim\limits_{\xi\to\infty}} Kf(\xi) = \lambda_0$ *if*

$$Kf(\xi) \geq \lambda_0, \quad \text{for each sufficiently large } \xi, \tag{28}$$

then no unbounded solution of (5) *exists; if*

$$Kf(\xi) < \lambda_0, \quad \text{for each sufficiently large } \xi, \tag{29}$$

then there exists an unbounded solution for (5).

REMARK The methods used in the proof of Theorem 3.2. make us to demonstrate simply the stability of nonnegative equilibrium state of system (5). The conclusions are stronger than those of [17], meanwhile, no linearization is involved.

Now we discuss the locally asymptotic stable in $D(A)\cap\Xi^+$ about the nonnegative equilibrium state of system (5). We have the following result.

THEOREM 3.4. *Let* $\xi_0 > 0$ *be the nonnegative equilibrium state of system* (5) *(iff* $Kf(\xi_0) = \lambda_0$*). If* $Kf(\xi)$ *is strictly monotone increasing in a neighbourhood of* ξ_0*, then the system* (5) *is asymptotic stable in* $D(A)\cap\Xi^+$ *about* ξ_0*; if* $Kf(\xi)$ *is strictly monotone decreasing in a neighbourhood of* ξ_0*,* ξ_0 *is not stable.*

PROOF. We only give out the proof for $N(t)$ because the relation between $N(t)$ and the solution $p(r,t)$ of equation (5) tells us that it is equivalent to discuss $p(r,t)$.

We say system (5) is asymptotic stable in $D(A)\cap\Xi^+$ about ξ_0, if for any $\varepsilon > 0$ there exists $\delta > 0$ when $|N_0 - \xi_0| < \delta$ the solution $N(t)$ of the equation

$$\begin{cases} N(t) = (N_0 + o(e^{-\varepsilon t}))e^{\int_0^t (\lambda_0 - Kf(N(\varrho)))d\varrho} \\ N(0) = N_0 \end{cases}$$

satisfy $|N(t) - \xi_0| < \varepsilon$ (for $t >$ some t_1).

First, let $Kf(\xi)$ be strictly monotone decreasing in $[\xi_0 - \delta, \xi_0 + \delta]$, here $\delta > 0$, ξ_0 corresponds to the initial state

$$p_0(r) = C_0 e^{-\lambda_0 r - \int_0^r \mu(\varrho)d\varrho}$$

where C_0 is a constant such that $\|p_0(r)\|_{L^1(0,r_m)} = \xi_0$. Take $0 < c < 1$ and $\hat{p}(r) = cp_0(r)$ such that $\|\hat{p}(r)\|_{L^1(0,r_m)} \in (\xi_0 - \delta, \xi_0)$ and the total number $\hat{N}(t)$ corresponding to initial state satisfies

$$\hat{N}(t) = \left(\|\hat{p}(r)\|_{L^1(0,r_m)} + o(e^{-\varepsilon t})\right) e^{\int_0^t [\lambda_0 - Kf(\hat{N}(\varrho))]d\varrho}.$$

We denote $\|\hat{p}(r)\|_{L^1(0,r_m)} e^{\int_0^t [\lambda_0 - Kf(\hat{N}(\varrho))]d\varrho}$ by $N^*(t)$. Since $N^*(t)$ is a monotone function at $t = 0$, and $Kf(\|\hat{p}\|_{L^1(0,r_m)}) > \lambda_0$, so $N^*(t)$ monotone decreasing to

$$\sup_{\xi \leq \|\hat{p}\|_{L^1(0,r_m)}} \{\xi \mid Kf(\xi) = \lambda_0\} < \xi_0 - \delta.$$

But $0 < c < 1$ can be chosen such that $1 - c$ is small enough, hence ξ_0 is not stable.

Suppose $Kf(\xi)$ is monotone increasing in $[\xi_0 - \delta, \xi_0 + \delta]$, choose $\varepsilon > 0$ such that $Kf(\xi) = \lambda_0 + \varepsilon$ has a solution in $[\xi_0, \xi_0 + \delta]$. It is not difficult to prove that there exists a neighbourhood $\hat{O}_{\xi_0}$ of ξ_0, such that $\hat{O}_{\xi_0} \subset \subset [\xi_0 - \delta, \xi_0 + \delta]$ and for any $N_0 \in \hat{O}_{\xi_0}$ $|g(t)| < \varepsilon$, $\forall t \geq t_0$ where $t_0 > 0$ is a positive number independent of N_0. Here

$$\begin{cases} N(t) = N_q(t) e^{-K\int_0^t f(N(\varrho))d\varrho} \\ N(0) = N_0 \\ g(t) = N_q'(t)/N_q(t)^{-\lambda_0}. \end{cases}$$

By the continuous dependence of solution to initial value, we knew that there is a neighbourhood $O_{\xi_0} \subset \hat{O}_{\xi_0}$, such that for any $N_0 \in O_{\xi_0}$, the corresponding solution $N(t)$ satisfies $\exists t_1 > t_0$ $\{N(t), t \geq t_1\} \subset \hat{O}_{\xi_0} \subset \subset (\xi_0 - \delta, \xi_0 + \delta)$. The solution of $Kf(\xi) = \lambda_0 + \varepsilon$ in $(\xi_0 - \delta, \xi_0 + \delta)$, denoted by $\hat{\xi}_\varepsilon$, does not belong to $\hat{O}_{\xi_0}$.

For any $N_0 \in O_{\xi_0}$, consider the following limiting equation

$$\begin{cases} \dfrac{d\overline{N}(t)}{dt} = [\lambda_0 - Kf(\overline{N}(t))]\overline{N}(t) + \varepsilon \overline{N}(t), \quad t \geq t_0 \\ \overline{N}(t_0) = N_0 \end{cases}.$$

By compare principle, $N(t) \leq \overline{N}(t)$ for all $t \geq t_0$, but $\overline{N}(t) \leq \hat{\xi}_\varepsilon$, hence when $t \geq t_1$, we have $N(t) \leq \hat{\xi}_\varepsilon$. Similar arguments to lower solution, we can get a neighbourhood of ξ_0 and ξ_ε such that when N_0 belongs to this neighbourhood, the corresponding solution $N(t) \geq \xi_\varepsilon$, and there exists t_2 when $t \geq t_2$.

Because $\xi_\varepsilon, \hat{\xi}_\varepsilon$ depend continuously on ε, $\lim_{\varepsilon \to 0} \xi_\varepsilon = \xi_0$ and $\lim_{\xi \to 0} \hat{\xi}_\varepsilon = \xi_0$, so, ξ is stable. The asymptotic stability can also be deduced by Theorem 3.2.

REMARK. We have already known that mathematical stability theory has a lot of definitions of stability. One of the basic ones is the stability in Lyapunov's sense. It is this notion that is used in this paper.

However, other definitions of stability have been used in biological population [14]. For instance, stability is interpreted as boundedness of solutions from below and above (it is often called stability in Lagrange's sense).

Next, we shall discuss the periodic and oscillator behaviour of system (5). We would rather give a direct proof than rely on Theorem 3.2, although the periodic problem can be demonstrated directly by it.

THEOREM 3.5. *No non-trivial periodic solution exists for system* (5).

PROOF. In fact, if $N(t) = N(t+T)$, $\forall t \geq 0$, for some fixed $T > 0$, then by the asymptotic expansion

$$N(t)e^{-\int_0^t [\lambda_0 - Kf(N(\varrho))]d\varrho} = N^* + o(e^{-\varepsilon t}).$$

From $N(t) = N(t+T)$, we have

$$\begin{aligned} N(t) &= (N^* + o(e^{-\varepsilon t}))e^{\int_0^t [\lambda_0 - Kf(N(\varrho))]d\varrho} \\ &= \left(N^* + o(e^{-\varepsilon(t+nT)})\right) e^{\int_0^t [\lambda_0 - Kf(N(\varrho))]d\varrho} \cdot e^{\int_t^{nT} [\lambda_0 - Kf(N(\varrho))]d\varrho} \end{aligned}$$

for $n = 1, 2, 3, \ldots$ so

$$N(t) = N^* e^{\int_0^\infty [(\lambda_0 - Kf(N(\varrho))]d\varrho}$$

moreover $N(o) = N(T)$. So, we must have

$$N(t) = N^*, \qquad Kf(N^*) = \lambda_0.$$

It means no non-trivial periodic solution for system (5).

Now, we discuss the oscillation of the system (5). We start with definition given in [15].

DEFINITION. The population system is called oscillation about its positive equilibrium position, if for any $\lim_{t\to\infty} N(t) = N^* > 0$, and interval $[\alpha, \infty)$, there is at least one zero point of $N(t) - N^*$ in $[\alpha, \infty)$ for any $\alpha > 0$. In other words, there are $t_n \to \infty$ $(n \to \infty)$, such that $N(t_n) = N^*$; if there is $\lim_{t\to\infty} N(t) = N^* > 0$ but $N(t) - N^* \neq 0$ for each sufficiently large t, then population system is called non-oscillation about its positive equilibrium state.

In [3], we have proved that for linear system

$$\begin{cases} \frac{dp(\cdot,t)}{dt} = Ap(\cdot,t) \\ p(\cdot,0) = p_0(\cdot) \end{cases} \tag{29}$$

there is the following result

THEOREM 3.6. *When $\beta = \beta_{cr}$, i.e. $\lambda_0 = 0$, the system (29) oscillates about its equilibrium state*

$$N^* = \left\| c_0 e^{-\int_0^r \mu(\varrho)d\varrho} \right\|_{L^1(0,r_m)}.$$

However, for nonlinear system (5) we have

THEOREM 3.7. *Assume that $f \in C^1[0,\infty)$ and $\lim_{\xi\to\infty} f(\xi) = +\infty$. If $\lambda_0 > 0$ and $Kf(\xi) = \lambda_0$ has solution, then nonlinear population system (5) is not oscillation about its positive equilibrium state.*

PROOF. Take initial distribution

$$p_0(r) = ce^{-\lambda_0 r - \int_0^r \mu(\varrho)d\varrho}$$

$c > 0$, then the corresponding solution of system (5)

$$N(t) = \int_0^{r_m} p(r,t)dr$$

satisfies

$$N(t) = \left(M + o(e^{-\varepsilon t})\right) e^{\int_0^t [\lambda_0 - Kf(N(\varrho))]d\varrho}. \tag{30}$$

From Theorem 3.2 $\lim_{t\to\infty} N(t)$ exists (it may be infinite). Choose $c > 0$ small enough such that

$$M < \min_{\xi\geq 0}\{\xi \mid Kf(\xi) = \lambda_0\} \tag{31}$$

then for any $t \geq 0$, $N(t) \in [M, N^*]$, and

$$N^* = \min_{\xi\geq 0}\{\xi \mid Kf(\xi) = \lambda_0\}.$$

Actually, there exists $p_0(r)$ such that $\|p_0\|_{L^1} = M$ and $q(r,t) = p_0(r)e^{\lambda_0 t}$ is a solution of the equation

$$\begin{cases} \frac{dq(r,t)}{dt} = Aq(r,t) \\ q(r,0) = p_0(r). \end{cases}$$

The $N(t)$ is the solution of the equation

$$\begin{cases} N(t) = N_q(t)e^{-\int_0^t Kf(N(\varrho))d\varrho} \\ N(0) = M \\ N_q(t) = \int_0^{r_m} |q(r,t)|dr. \end{cases}$$

We will show that $N(t) \in [M, N^*]$, if not, $\exists t^* > 0$, $t_1 > t^*$ such that

$$N(t^*) = N^*, \quad N(t_1) > N^*. \tag{29'}$$

Since

$$N^* = N(t^*) = N_q(t^*)e^{-\int_0^{t^*} Kf(N(\varrho))d\varrho} = Me^{\lambda_0 t^*}e^{-\int_0^{t^*} Kf(N(\varrho))d\varrho}$$

$$Me^{-\int_0^{t^*} Kf(N(\varrho))d\varrho} = N^* e^{-\lambda_0 t^*}.$$

We have

$$N(t) = N^* e^{\lambda_0(t-t^*)} \cdot e^{-\int_{t^*}^{t} Kf(N(\varrho))d\varrho}, \quad t \geq t^*.$$

Hence, $N(t)$ satisfy

$$\begin{cases} \overline{N}(t) = \overline{N}_q(t)e^{-\int_{t^*}^{t} f(N(\varrho))d\varrho}, \quad t > t^* \\ \overline{N}(t^*) = N^* \end{cases} \tag{30'}$$

where $\overline{N}_q(t) = \int\limits_0^{r_m} |\overline{q}(r,t)|dr$, and $\overline{q}(r,t) = p_0(r)e^{\lambda_0 t^*} e^{\lambda_0(t-t^*)}$.

It is easy to see $\overline{q}(r,t)$ satisfy

$$\begin{cases} \frac{dq(r,t)}{dt} = Aq(r,t) \\ q(r,t^*) = p_0(r)e^{\lambda_0 t^*}. \end{cases}$$

But $\overline{N}(t) = N^*$ is also the solution of equation (30′), so $N(t) = N^*$, $(t \geq t^*)$, this is a contradiction to (29′).

We write the Taylor expression of $Kf(\xi)$ at $\xi = N^*$ as

$$Kf(\xi) = \lambda_0 + Kf'(N^*)(\xi - N^*) + o(\xi - N^*) \tag{32}$$

for $\xi \in (N^* - \delta, N^*)$, $\delta > 0$. Hence there are constants $\delta_0 > 0$ and $c_0 > 0$ such that

$$\frac{Kf(\xi) - \lambda_0}{\xi - N^*} \doteq Kf'(N^*) + o(1) < c_0$$

for any $\xi \in (N^* - \delta_0, N^*)$, i.e. $Kf(\xi) > \lambda_0 + c_0(\xi - N^*)$ for $\xi \in (N^* - \delta_0, N^*)$. We define the function

$$K\hat{f}(\xi) = \begin{cases} U(\xi), & \xi \in [0, N^* - \hat{\delta}_0] \\ \lambda_0 + c_0(\xi - N^*), & \xi \in (N^* - \hat{\delta}_0, \infty) \end{cases}$$

where $\hat{\delta}_0 < \delta_0$, $U(\xi)$ is a function such that

$$K\hat{f}(\xi) > 0, \quad \text{for } \xi > 0; \qquad K\hat{f}(0) = 0 \qquad K\hat{f} \in C^1[0, \infty)$$

and $K\hat{f}(\xi) \neq \lambda_0$ for $\xi \in [0, N^*)$. So $Kf(\xi) > K\hat{f}(\xi)$, $\forall \xi \in [N^* - \hat{\delta}_0, N^*]$. Take the initial distribution such that $M \in (N^* - \hat{\delta}_0, N^*)$, denote $\hat{N}(t)$ the solution of (5) corresponding to $K\hat{f}(\xi)$. Since $N(t)$, $\hat{N}(t) \in (N^* - \hat{\delta}_0, N^*)$, we have

$$N(t) < \hat{N}(t), \quad \forall t \geq 0. \tag{33}$$

but $\hat{N}(t) \leq N^*$. It follows that $\lim_{t \to \infty} N(t) = N^*$, $N(t) < N^*$, $\forall t \geq 0$. By definition, $N(t)$ does not oscillate about N^*. The Theorem is proved.

REMARK. Theorem 3.5 and 3.6 tell us the difference between linear and nonlinear models. Furthermore, generally speaking, the solutions of nonlinear equation (5) are bounded (under certain conditions). This avoids the possible situation of

$$\lim_{t \to \infty} N(t) = +\infty$$

which is not possible in practice.

4. Non-linear Age-dependent Population Dynamics with Immigration

If we consider immigration, we have the following nonlinear population evolution equation

$$(34)\qquad \begin{cases} \frac{\partial p(r,t)}{\partial t} + \frac{\partial p(r,t)}{\partial r} = -\mu(r)p(r,t) - K\tilde{f}(N(t))p(r,t) + g(r,t) \\ \qquad\qquad 0 < r < r_m,\ t > 0, \\ p(r,0) = p_0(r) \\ p(0,t) = \beta \int\limits_{r_1}^{r_2} k(r)h(r)p(r,t)dr,\ t > 0 \end{cases}$$

in which $N(t) = \int\limits_0^{r_m} |p(r,t)|dr$; $g(r,t)$ is densitive distribution of immigration population at time t. Generally speaking, the relation of the $g(r,t)$, $p(r,t)$ and $N(t)$ is very complex. We discuss, in this section, the case when $g(r,t)$ and $p(r,t)$ are in following relation $g(r,t) = \omega(N(t))p(r,t)$, here ω is a real function which is not required to be positive.

So, the system (34) can be written as

$$(35)\qquad \begin{cases} \frac{\partial p(r,t)}{\partial t} + \frac{\partial p(r,t)}{\partial r} = -\mu(r)p(r,t) - f(N(t))p(r,t),\ 0 < r < r_m,\ t > 0, \\ p(r,0) = p_0(r), \quad 0 \le r \le r_m \\ p(0,t) = \beta \int\limits_{r_1}^{r_2} k(r)h(r)p(r,t)dr, \quad t > 0 \end{cases}$$

in which f is a real function which is not required to be positive. The difference between system (4) and system (35) is only the requirement for function f.

In state space Ξ we have the following abstract evolution equation

$$(36)\qquad \begin{cases} \frac{dp(\cdot,t)}{dt} = Ap(\cdot,t) - f(N(t))p(\cdot,t), \quad t > 0, \\ p(\cdot,0) = p_0(\cdot), \\ N(t) = \int\limits_0^{r_m} |p(r,t)|dr, \quad t \ge 0. \end{cases}$$

About the existence of equation (36), we have

THEOREM 4.1. *For equation* (36), *if f is continuous on* $[0,\infty)$, $f(0)=0$ *and*

$$|f(\xi_1) - f(\xi_2)| \le \frac{c(\xi')}{\xi'}|\xi_1 - \xi_2|,$$

in which $\xi' = \max\{\xi_1, \xi_2\}$, $c(\xi) \in L^1(0,\infty)$; *and* $p_0 \in D(A)$. *Then equation* (36) *has unique global classical solution in state space* Ξ. *Moreover if* $p_0 \geq 0$, *the solution is nonnegative.*

PROOF. Let $Y = D(A)$, $\|g\|_Y = \|g\|_\Xi + \|Ag\|_\Xi$ for any $g \in Y$, then $[Y, \|\cdot\|_Y]$ become a Banach space, denoted by Y.

Since $T(t)Y \subset Y$ and

$$\lim_{t\to 0} \|T(t)\varphi - \varphi\|_Y = 0, \quad \forall \varphi \in Y$$

if $S(t) = T(t)\Big|_Y$, i.e. $S(t)$ is the restriction of $T(t)$ on Y, then $\{S(t)\}_{t\geq 0}$ is a strongly continuous semigroup on Y. The infinitesimal generator of $\{S(t)\}_{t\geq 0}$ is denoted by B, then $D(A^2) = D(B)$ and $B\varphi = A\varphi$, for any $\varphi \in D(\bar{B})$. Define mapping

$$F : L^2([0,T]);Y) \to L^2([0,T);Y),$$

$$(F\varphi)(\cdot,t) = S(t)p_0(\cdot) - \int_0^t S(t-\tau)f(N(\tau))\varphi(\cdot,\tau)d\tau,$$

here $\varphi \in L^2([0,T);Y)$ and $N(t) = \int_0^{r_m} |\varphi(r,t)|dr$. It is clear that the mapping F has meaning.

Let $\varphi_i(\cdot,t) \in Y$, $i = 1,2$, then

$$\|(F\varphi_1)(\cdot,t) - (F\varphi_2)(\cdot,t)\|_\Xi \leq c_1 \int_0^t \|\varphi_1(\cdot,\tau) - \varphi_2(\cdot,\tau)\|_\Xi d\tau \tag{37}$$

and

$$\|F(A\varphi_1)(\cdot,t) - F(A\varphi_2)(\cdot,t)\|_\Xi \leq c_2 \int_0^t \|A\varphi_1(\cdot,\tau) - A\varphi_2(\cdot,\tau)\|_\Xi d\tau \tag{38}$$

where c_i, $i = 1,2$, are constants. Since operator A and F can be exchanged, we have

$$\|(F\varphi_1)(\cdot,t) - (F\varphi_2)(\cdot,t)\|_Y \leq c_3 \int_0^t \|\varphi_1(\cdot,\tau) - \varphi_2(\cdot,\tau)\|_Y d\tau$$

by virtue of (37) and (38), where c_3 is a constant.

So we have

$$\|(F\varphi_1)(\cdot,t)-(F\varphi_2)(\cdot,t)\|_Y \le ct^{\frac{1}{2}}\|\varphi_1-\varphi_2\|_{L^2(0,T:Y)}. \tag{39}$$

According to (39), there are the following inequalities

$$\|F^2\varphi_1 - F^2\varphi_2\|_Y \le c\int_0^t \|F\varphi_1 - F\varphi_2\|_Y d\tau \le$$

$$\le c^2\|\varphi_1-\varphi_2\|_{L^2([0,T);Y)}\cdot\int_0^t \tau^{\frac{1}{2}}d\tau \le c^2\cdot\frac{t^{\frac{1}{2}+1}}{\frac{1}{2}+1}\cdot\|\varphi_1-\varphi_2\|_{L^2(0,T;Y)}.$$

It is obvious that for any positive integer n we have

$$\|(F^n\varphi_1)(\cdot,t)-(F^n\varphi_2)[\cdot,t)\|_Y \le$$

$$\le c^n\cdot\frac{t^{\frac{1}{2}+n-1}}{\left(\frac{1}{2}+1\right)\left(\frac{1}{2}+2\right)\cdots\left(\frac{1}{2}-n-1\right)}\cdot\|\varphi_1-\varphi_2\|_{L^2(0,T;Y)}.$$

Hence

$$\|F^n\varphi_1 - F^n\varphi_2\|_{L^2(0,T;Y)} \le$$

$$\le c^n\cdot\frac{T^{\frac{1}{2}+n}}{\left(\frac{1}{2}+1\right)\left(\frac{1}{2}+2\right)\cdots\left(\frac{1}{2}+n\right)}\cdot\|\varphi_1-\varphi_2\|_{L^2(0,T;Y)}$$

and there exists a positive number $\varrho<1$ and an integer n_0 such that

$$\|F^{n_0}\varphi_1 - F^{n_0}\varphi_2\|_{L^2(0,T;Y)} \le \varrho\cdot\|\varphi_1-\varphi_2\|_{L^2(0,T;Y)}$$

for any $\varphi_i \in L^2(0,T;Y)$, $i=1,2$, by virtue of

$$\lim_{n\to\infty}\frac{C^n}{n!}=0.$$

By the principle of Banach Fixed Point, the mapping F has a unique fixed point, denoted by $\varphi(\cdot,t)$, in $L^2(0,T;Y)$, i.e.

$$\varphi(\cdot,t)=S(t)p_0(\cdot)-\int_0^t S(t-\tau)f(N(\tau))\varphi(\cdot,\tau)d\tau. \tag{40}$$

Next, we consider inhomogeneous equation in Ξ.

$$(41)\qquad \begin{cases} \frac{dp(\cdot,t)}{dt} = Ap(\cdot,t) - f(N(t))\varphi(\cdot,t), \\ p(\cdot,0) = p_0(\cdot), \end{cases}$$

in which $N(t)=\int_0^{r_m}|\varphi(r,t)|dr$. Let $g(t)=f(N(t))\varphi(\cdot,t)$ then $g(t)\in L^1(0,T;\Xi)$, $g\in C([0,T);\Xi)$, $g(t)\in D(A)$ and $Ag(t)\in L^1(0,T;\Xi)$. By [10], equation (41) has a classical solution $p(\cdot,t)$, and

$$p(\cdot,t) = T(t)p_0(\cdot) - \int_0^t T(t-\tau)f(N(\tau))\varphi(\cdot,\tau)d\tau \qquad 0\le t<T.$$

So $p(\cdot,t)=\varphi(\cdot,t)$ by virtue of (40) and the definition of semigroup $\{S(t)\}_{t\ge 0}$. The classical solution of equation (36) in Ξ is:

$$(42)\qquad p(\cdot,t) = T(t)p_0(\cdot) - \int_0^t T(t-\tau)f(N(\tau))p(\cdot,\tau)d\tau, \qquad 0\le t<T.$$

In the following we will show when $p_0 \in D(A)\cap\Xi^+$, the $p(\cdot,t)$ defined by (42) is the solution of the equation

$$(43)\qquad \frac{dp(\cdot,t)}{dt} = Qp(\cdot,t) + (f_\infty - f(N(t)))\,p(\cdot,t)$$

and the solution is nonnegative. Here the operator is defined as follows

$$\begin{cases} D(Q) = D(A) \\ Q\varphi(r) = -\frac{d\varphi}{dr} - (\mu(r)+f_\infty)\varphi(r), \end{cases}$$

where $f_\infty = \|f\|_{L^\infty(0,\infty)}$.

Let e^{Qt} be the strongly continuous semigroup generated by operator Q defining a sequence of function in $C([0,T);\Xi)$

$$(44)\qquad \begin{cases} p_0(\cdot,t) = e^{Qt}p_0(\cdot) \\ p_n(\cdot,t) = e^{Qt}p_0(\cdot) + \int_0^t e^{Q(t-\tau)}(f_\infty - f(N_{n-1}(\tau)))p_{n-1}(\cdot,\tau)d\tau. \end{cases}$$

It is easy to see that $p_n(\cdot,t)\in C([0,T);\Xi)$ and $\{p_n(\cdot,t)\}$ converge uniformly on $[0,T)$. Let

$$g(\cdot,t) = \lim_{n\to\infty} p_n(\cdot,t) \qquad \overline{N}(t) = \int_0^{r_m} |g(r,t)|dr.$$

From (44)

$$g(\cdot,t)=e^{Qt}p_0(\cdot)+\int_0^t e^{Q(t-\tau)}(f_\infty-f(\overline{N}(\tau)))g(\cdot,\tau)d\tau. \tag{45}$$

It is not difficult to show that $g(\cdot,t)\geq 0$. So

$$p(\cdot,t)=e^{Qt}p_0(\cdot)+\int_0^t e^{Q(t-\tau)}(f_\infty-f(N(\tau)))p(\cdot,t)d\tau \tag{46}$$

by virtue of [10] and (43), in which

$$N(t)=\int_0^{r_m}|p(r,t)|dr.$$

From (45) and (46) together with Gronwall's inequality $p(r,t)=g(r,t)\geq 0$.

It follows that equation (36) has a unique nonnegative classical solution on $[0,T)$ for any $T>0$, therefore, it has a unique classical solution on $[0,\infty)$. The proof has been completed.

In many cases, requiring the existence of a classical solution (i.e. strong solution) to (36) is too stringent, and so we apply the variation of constants formula to (36) to obtain the equation

$$p(\cdot,t)=T(t)p_0(\cdot)+\int_0^t T(t-s)f(N(s))p(\cdot,s)ds, \quad t\geq 0. \tag{47}$$

A continuous function $p(\cdot,\tau):[0,T)\to\Xi$ that satisfies (47) for $t\in[0,T)$ is called a mild solution to (36) on $[0,T)$ (see [10]). Any classical solution to (36) is also a mild solution; however, the converse is not true in general. In this paper, we give better results than in [16].

We have the following basic existence result for (47) by virtue of [8], [7].

THEOREM 4.2. *If for each $R>0$ there is an $L(R)>0$ such that*

$$|f(\xi_1)-f(\xi_2)|\leq L(R)|\xi_1-\xi_2| \tag{48}$$

for all $\xi_1,\xi_2\in[0,\infty)$, $|\xi_1|,|\xi_2|\leq R$, then (47) has a unique solution $p(\cdot,t)=$ $=p_{p_0}(\cdot,t)$ on $[0,b_{p_0})$ for each $p_0\in D(A)$. Furthermore, if $b_{p_0}<\infty$ then $\|p_{p_0}(\cdot,t)\|_\Xi\to\infty$ as $t\to b_{p-_0}$.

Theorem 4.2 does not guarantee the continuity of the solution, i.e. Theorem 4.2 does not assert the existence of mild solution for (47). However, we have the following existence result of mild solution for (47).

THEOREM 4.3. *If the function f satisfies condition* (48) *on* $[0,\infty)$, *then* (47) *has a unique mild solution* $p(\cdot,t) = p_{p_0}(\cdot,t)$ *on* $[0, b_{p_0})$ *for each* $p_0 \in D(A)$. *Furthermore, if* $b_{p_0} < \infty$ *then* $\|p_{p_0}(\cdot,t)\|_\Xi \to \infty$ *as* $t \to b_{p_0}^-$.

PROOF. From Theorem 4.2, there is a function $p(\cdot,t)$ satisfies (47) on $[0, b_{p_0})$. Only if we can show that $p \in C([0, b_{p_0}))$, then the proof will be completed.

Let

$$W(t) = \int_0^t T(t-s)f(N(s))p(\cdot,s)ds, \quad 0 \le t < b_{p_0}.$$

For $h > 0$

$$\begin{aligned} W(t+h) - W(t) &= \int_0^{t+h} T(t+h-s)f(N(s))p(\cdot,s)ds - \\ &- \int_0^t T(t-s)f(N(s))p(\cdot,s)ds = \\ &= (T(h)-I)W(t) + \int_t^{t+h} T(t+h-s)f(N(s))p(\cdot,s)ds \end{aligned}$$

so

$$\|W(t+h) - W(t)\|_\Xi \le$$
$$\le \|(T(h)-I)W(t)\|_\Xi + \|T(h)\| \int_t^{t+h} \|T(t-s)f(N(s))p(\cdot,s)\|_\Xi ds.$$

By virtue of the strongly continuous of semigroup $\{T(t)\}_{t\geq 0}$, $W(t)$ is right-continuous. Again, let $0 < h < \varepsilon < t$, then

$$\begin{aligned}
&W(t) - W(t-h)\\
&= \int_0^t T(t-s)f(N(s))p(\cdot,s)ds - \int_0^{t-h} T(t-h-s)f(N(s))p(\cdot,s)ds\\
&= (T(\varepsilon) - T(\varepsilon - h))W(t-\varepsilon) + \int_{t-\varepsilon}^{t-h} T(t-s)f(N(s))p(\cdot,s)ds\\
&= \int_{t-\varepsilon}^{t-h} T(t-h-s)f(N(s))p(\cdot,s)ds.
\end{aligned}$$

It is not difficult to know that $\|W(t) - W(t-h)\|_{\Xi} \to 0$, $(h \to 0^+, \varepsilon \to 0^+)$ hence $W(t)$ is continuous in Ξ. It is obvious that $T(t)p_0$ is continuous, so $p(\cdot,t)$ is continuous.

We will show that the global mild solution exists under certain conditions for the function f.

COROLLARY 4.4. *In addition to the function f satisfies the condition of Theorem 4.3, suppose that f is bounded on $[0,\infty)$, then (47) has global mild solution.*

PROOF. If system (47) has no global mild solution, then $b_{p_0} < +\infty$, and

$$\|p(\cdot,t)\|_{\Xi} \to +\infty \quad \text{as} \quad t \to b_{p^-_0},$$

by virtue of Theorem 4.3.

On the other hand, we have

$$\|p(\cdot,t)\|_{\Xi} \leq \|T(t)p_0(\cdot)\|_{\Xi} + M\int_0^t |f(n(s))| \cdot e^{\omega(t-s)} \cdot \|p(\cdot,s)\|_{\Xi} ds$$

for $t \in [0, bp_0)$. Here, we suppose $\|T(t)\| \leq Me^{\omega t}$. Since $\|p(\cdot,t)\|_{\Xi}$ is continuous on $[0, b_{p_0})$, from Gronwall's inequality it follows that

$$\|p(\cdot,t)\|_{\Xi} \leq M'\|p_0\|_{\Xi} e^{M'\int_0^t |f(N(s))|ds}$$

in which $M' = \max\left\{1, Me^{\omega b p_0}\right\}$. Let $M'' = M'\|p_0\|_\Xi e^{M'\|f\|_{L^1(0,\infty)} \cdot b_{p_0}}$, then $\|p(\cdot,\tau)\|_\Xi \leq M''$, $t \in [0, b_{p_0})$ this is a contradiction, therefore the corollary has been proved.

It is very interesting that we have found that the mild solution of system (36) and the generalized solution are equivalent under weak condition. We start with the definition of generalized solution.

A generalized solution (see [7]), or weak solution (see [1]) of the initial value problem (36) can be defined as follows: a function $p(\cdot,t) \in C([0,T);\Xi)$ is a generalized solution of (36) if

(i) $p(\cdot, o) = p_0(\cdot)$;

(ii) for every $q \in C_0^\infty(0, r_m)$, $\langle p(\cdot,t), q\rangle$ is absolutely continuous on any finite interval, here $\langle p(\cdot,t), q\rangle = \int_0^{r_m} p(r,t)q(r)dr$;

(iii) $\frac{d}{dt}\langle p(\cdot,t), q\rangle = \langle p(\cdot,t), A^*q\rangle - \langle f(N(t))p(\cdot,t), q\rangle$ a.e. on $[0,T)$ where A^* is the adjoint of A.

The result is as follows.

THEOREM 4.5. *Assume that state space $\Xi = L^p(0, r_m)$ $(1 < p < \infty)$; $p_0 \in D(A)$, the function f satisfies*

$$f \in C[0,\infty) \cap L^\infty(0,\infty).$$

Then the mild solution and generalized solution of equation (36) are equivalent.

PROOF. First, we show that the mild solution $p(\cdot,t)$ will be generalized solution.

Let $V(t) = \langle p(\cdot,t), q\rangle$ where $q \in C_0^\infty(0, r_m)$, then when $t_1 > t_2$, we have

$$|V(t_1) - V(t_2)| \leq |\langle T(t_1)p_0 - T(t_2)p_0, q\rangle| + \tag{49}$$
$$+ \left|\left\langle \int_{t_1}^{t_2} T(t_1 - s)f(N(s))p(\cdot,s)ds, q \right\rangle\right| +$$
$$+ \left|\left\langle (T(t_1 - t_2) - I) \int_0^{t_2} T(t_2 - s)f(N(s))p(\cdot,s)ds, q \right\rangle\right| \leq$$

$$\le M_1|t_1-t_2| + M_2 \int_{t_2}^{t_1} |f(N(s))| \cdot \|p(\cdot,s)\|_\Xi ds +$$

$$+ M_3 \int_0^{T_1} |f(N(s))| \cdot \|p(\cdot,s)\|_\Xi \cdot (\|A^*q\|_{\Xi^*} + C)\, ds \cdot |t_1-t_2|$$

in which M_i $(i=1,2,3,)$ and C are constants which are independent of t, here $t, t_i \in [0,T_1]$, T_1 is any positive number. So, $V(t)$ is absolutely continuous on any finite interval, by virtue of (49).

We will show the differentiability.

Let $h>0$, $q \in C_0^\infty(0,r_m)$, then $q \in D(A^*)$, we have

$$\frac{1}{h}\langle p(\cdot,t+h) - p(\cdot,t), q\rangle = \tag{50}$$

$$= \frac{1}{h}\left\langle (T(h)-I)p(\cdot,t) - \int_t^{t+h} T(t+h-s)f(N(s))p(\cdot,s)ds, q\right\rangle =$$

$$= \frac{1}{h}\langle p(\cdot,t), (T^*(h)-I)q\rangle -$$

$$- \left\langle \frac{1}{h}\int_t^{t+h} T(t+h-s)f(N(s))p(\cdot,s)ds, q\right\rangle.$$

Since

$$\lim_{h\to 0+} \left\langle p(\cdot,t), \frac{1}{h}(T^*(h)-I)q\right\rangle = \langle p(\cdot,t), A^*q\rangle \tag{51}$$

and

$$\lim_{h\to 0} \frac{1}{h}\int_t^{t+h} \|f(N(s))p(\cdot,s)\|_\Xi ds$$

exists a.e. on $[0,T_1]$, hence

$$\left|\frac{1}{h}\int_t^{t+h} \langle f(N(s))p(\cdot,s), T^*(t+h-s)q - q\rangle\, ds\right| \le$$

$$\le \sup_{t\le s\le t+h} \|T^*(t+h-s)q - q\|_{\Xi^*} \cdot \frac{1}{h}\int_t^{t+h} \|f(N(s))p(\cdot,s)\|_\Xi ds \longrightarrow 0$$

a.e. on $[0, T_1]$, as $h \to 0$. By virtue of (50) and (51), we have

$$\lim_{h\to 0^+} \frac{1}{h}\langle p(\cdot, t+h) - p(\cdot, t), q\rangle = \langle p(\cdot, t), A^* q\rangle - \langle f(N(t))p(\cdot, t), q\rangle$$

a.e. on $[0, T_1]$. Similarly, we also have

$$\lim_{h\to 0^+} -\frac{1}{h}\langle p(\cdot, t) - p(\cdot, t-h), q\rangle = \langle p(\cdot, t), A^* q\rangle - \langle f(N(t))p(\cdot, t), q\rangle$$

a.e. on $[0, T_1]$. Therefore, $p(\cdot, t)$ is a generalized solution (i.e. weak solution) of the equation (36).

We now prove the second part of the theorem.

Let $p(r, t)$ is a generalized solution of equation (36) on $[0, t)$, then

$$T(t)f(N(t))p(\cdot, t) \in L^1([0, T); \Xi).$$

Let

$$\overline{p}(\cdot, t) = T(t)p_0(\cdot) - \int_0^t T(t-s)f(N(s))p(\cdot, s)ds$$

for every $q \in C_0^\infty(0, r_m)$. Similarly to the first part of the proof, we have

$$\frac{d}{dt}\langle \overline{p}(\cdot, t), q\rangle = \langle \overline{p}(\cdot, t), A^* q\rangle - \langle f(N(t))p(\cdot, t), q\rangle, \quad \text{a.e. on } [0, T)$$

Let $Q(r, t) = p(r, t) - \overline{p}(r, t)$ then

$$\frac{d}{dt}\langle Q(\cdot, t), q\rangle = \langle Q(\cdot, t), A^* q\rangle \quad \text{a.e. on } [0, T).$$

By virtue of [11], we have

$$\langle Q(\cdot, t), q\rangle = \left\langle \int_0^t Q(\cdot, s)ds, A^* q\right\rangle \quad \text{a.e. on } [0, T). \tag{52}$$

Since

$$\int_0^t Q(\cdot, s)ds \in C([0, T); \Xi)$$

(52) holds for every $t \in [0, T)$.

Take $\lambda \in \sigma(A^*)$, let $q = (\lambda - A^*)^{-1} g$, by virtue of (9) and [13], we have

$$\left\langle (\lambda - A)^{-1} Q(\cdot, t), g\right\rangle = \left\langle \int_0^t Q(\cdot, s)ds, A^* R(\lambda, A^*)g\right\rangle$$

i.e.

$$\langle R(\lambda,A)Q(\cdot,t),g\rangle = \left\langle \int_0^t Q(\cdot,s)ds, -g \right\rangle + \lambda \left\langle \int_0^t Q(\cdot,s)ds, R(\lambda,A^*)g \right\rangle$$

therefore

$$\left\langle R(\lambda,A)(Q(\cdot,t) - \lambda \int_0^t Q(\cdot,s)ds; g \right\rangle = -\left\langle \int_0^t Q(\cdot,s)ds, g \right\rangle.$$

From [13], we know that $\{\theta(r) \mid \theta(r) = (\lambda - A^*)q, q \in C_0^\infty(0,r_m)\}$ is dense in Ξ^*, it implies

$$R(\lambda,A)\left(Q(\cdot,t) - \lambda \int_0^t Q(\cdot,s)ds\right) = -\int_0^t Q(\cdot,s)ds$$

and

$$\int_0^t Q(\cdot,s)ds \in D(A) \text{ for } t \in [0,T).$$

By (52), we have

$$A\int_0^t Q(\cdot,s) = Q(\cdot,t) \text{ for any } t \in [0,T).$$

It is obvious that $\int_0^t Q(\cdot,s)ds$ is the solution of the following equation

(53) $$\begin{cases} \frac{du(t)}{dt} = Au(t), & 0 \le t < T, \\ u(0) = 0. \end{cases}$$

So

$$\int_0^t Q(\cdot,s)ds = 0, \quad 0 \le t < T,$$

i.e.

$$\overline{p}(\cdot,t) = p(\cdot,t), \quad 0 \le t < T.$$

The proof has been completed.

Appendix

The model of the population evolution is the foundation of analysis of population evolution process, population forecasting and population optimal control. We can obtain many good results in studying population evolution process with mathematical model in the light of modern science technology and methods, such as optimal control theory, system science, technology of system engineering and computer science. For instance, in order to control population growth in a planned way, we can get optimal female birth rate by applying optimal control theory (see [18]). If we want to stipulate the fertility level to stop the population growth or the population drop eventually, we can find the formula to calculate the critical fertility rate by applying the theory of stability (see [20]).

There are many kinds of mathematical models to describe the population evolution process. The model studied in our paper is very useful in human population theory, we will show it with some data.

We start with introducing the population parameter used in the following evolution equation

$$\begin{cases} \frac{\partial p(r,t)}{\partial t} + \frac{\partial p(r,t)}{\partial r} = -\mu(r,t)p(r,t) \\ p(r,0) = p_0(r) \\ p(0,t) = \beta(t) \int_{r_1}^{r_2} k(r,t)h(r,t)p(r,t)dr. \end{cases}$$

The distribution density of population $p(r,t)$ is defined by

$$p(r,t) = \frac{\partial N(r,t)}{\partial r}.$$

Here, the $N(r,t)$ denotes the total population aged r at time t.

The function $\mu(r,t)$ of relative mortality rate is a very important parameter in population evolution process. It is defined in the following way:

$$\mu(r,t) = \lim_{\Delta r \to 0} \frac{M(r,\Delta r,t)}{p(r,t)\Delta r}$$

where $M(r,\Delta r,t)$ denotes the total mortality population in unit time between age r and age $r+\Delta r$ at time t. The total living population in $[r, r+\Delta r]$ at same time is $p(r,t)\Delta r$.

Now, we give the definition of the specific fertility rate.

We denote the total population fertilized in unit time at a time t by $\varphi(t)$, it is the absolute fertility rate. The ratio of the $\varphi(t)$ to the total population $N(t)$ at time t is denoted by $U(t)$, called relative fertility rate. Table 1 tells us the absolute fertility rate and relative fertility rate in China in recent years (see [19]).

t (year)	1975	1976	1978	1979	1980	1982
$\varphi(t)$ (million/year)	21.09	18.53	17.86	17.27	14.90	20.689
$U(t)$ (%)	2.293	1.987	1.889	1.790	1.535	2.091

Table 1.

The $k(r,t)$ denotes the female ratio at time t, $[r_1, r_2]$ denotes the fecundity period of females.

Let the total population between age r and age $r+\Delta r$ be $p(r,t)$ (taking $\Delta r = 1$), the total population of female at time t at same age interval be $k(r,t)p(r,t)$. Let $f(r,t)$ denote the number of parity of each female aged r at time t, then the total number $F(r,t)$ of parity of female aged r is

$$F(r,t) = f(r,t)k(r,t)p(r,t).$$

Hence, the total number of parity of female aged between r_1 and r_2 at time t is $\Phi(t)$, and

$$\Phi(t) = \int_{r_1}^{r_2} F(r,t)dr$$

i.e.

$$\Phi(t) = \int_{r_1}^{r_2} f(r,t)k(r,t)p(r,t)dr. \tag{A1}$$

The $f(r,t)$ reflects the fertility levels at different age and at different time.

Now, we normalize the $f(r,t)$. Let

$$f(r,t) = \beta(t)h(r,t) \tag{A2}$$

the $h(r,t)$ be called the fertility pattern which is required to satisfy

$$\int_{r_1}^{r_2} h(r,t)dr = 1 \tag{A3}$$

the (A3) is called the condition of normalization.

Integrating both sides of (A2) from r_1 to r_2 for r, we have

$$\beta(t) = \int_{r_1}^{r_2} f(r,t)dr. \tag{A4}$$

The (A4) means that $\beta(t)$ is the number of parity of each female in $[r_1, r_2]$ at time t. It is called specific fertility rate. Replacing $f(r,t)$ in (A1) by (A2), we have

$$\Phi(t) = \beta(t) \int_{r_1}^{r_2} h(r,t)k(r,t)p(r,t)dr.$$

Table 2 is taken from the fertility pattern data of Tianjing city in China (see [19]).

Age	Fertility rate f (1978) ‰	Fertility pattern h (1978) ‰
23	0.55	0.46
24	7.53	6.28
25	39.14	32.62
26	115.95	96.63
27	143.16	119.30
28	175.91	146.59
29	212.32	176.93
30	205.75	171.45
31	101.35	84.46
32	85.21	71.01
33	57.26	47.72
34	27.03	22.35
35	17.64	14.70
36	8.07	6.73
37	2.02	1.68

Table 2.

It is easy to know $\beta(1978) = 1.20$.

After having observed census data, we have found that the fertility pattern $h(r,t)$ can be approximated by Gamma distribution density.

During short time, for example, during one year, we suppose that $h(r,t)$ be independent on t, in this case, we have

$$h(r,t)=\begin{cases}\frac{(r-r_1)^{\alpha-1}e^{-\frac{r-r_1}{\theta}}}{\theta^{\alpha}\Gamma(\alpha)}, & r>r_1,\\ 0, & r\le r_1.\end{cases}$$

Here $\Gamma(\alpha)$ is Gamma function

$$\Gamma(\alpha)=\int_0^{\infty} e^{-x}e^{\alpha-1}d\chi.$$

The parameters α and θ can be determined by the statistical data. For example, by computer data processing for Jiling Province's census data (see [2]), these are the following results

$$\alpha=\frac{n}{2},\ \theta=\frac{4}{5}$$

$$(A5)\qquad h(r,t)=\begin{cases}\frac{1}{\left(\frac{4}{5}\right)^{\frac{n}{2}}\Gamma(\frac{n}{2})}\cdot(r-r_1)^{\frac{n}{2}-1}e^{-\frac{5}{4}(r-r_1)}, & r>r_1,\\ 0, & r\le r_1,\end{cases}$$

the n satisfy the equality $n=\frac{5}{2}(r_m-r_1)+2$.

The following table (Table 3) shows that the fertility pattern determined by (A5) can approximate the real fertility data very well (see [19]).

Province	Year	The number of babies born during the year		$\lvert(1)-(2)\rvert$	Relative error %
		Sensus data (1)	calculated data (2)		
Inner Mon.	1981	18639	18900	261	1.4
Jiling	1980	29945	29973	28	0.1

Table 3.

Now, we have understood the parameters in the population evolution equations. According to the numerical methods for partial differential equation, we can use the equation to forecast the population (see [21]).

We can compare the calculated data with sensus data. It is shown in the following table.

Year	Population increased during the year (million)			The total population (million)		
	Sensus data (1)	Calculated data (2)	(3) ((1)/(2))%	Sensus data (4)	Calculated data (5)	(6) ((4)/(5))%
1975	14.38	14.40	99.9	919.70	918.50	100.1
1976	11.78	11.80	99.8	932.67	930.29	100.3
1977	11.38	11.37	100.1	945.23	941.66	100.4
1978	11.47	11.46	100.1	958.09	953.11	100.5

Table 4.

It follows that the model we have studied can approximate the real population evolution process very well.

From above we know that the population evolution equation is very successful in analysis of population. It is that reason we want to study this kind of population evolution model further on.

References

[1] BREZIS, H. and LIONS, J. L., *Nonlinear Partial Differential Equations and their Applications*, Research Notes in Mathematics, 84. Boston, London, Melbourne, 1983.

[2] GUO, B.-Z. and ZHU, G.-T., The Properties of C_0-semigroup $\{T(t)\}_{t\geq 0}$ Generated by the Population Operator, *J. Sys. Sci. & Math. Scis.*, **6** (1986).

[3] GUO, B.-Z., JIANHUA, YUAN and SHENGHAI, ZHANG, Asymptotic Analysis and Oscillatory Properties of the Logistic Model of Age-dependent Population Growth, *J. Sys. Sci. & Math. Scis.*, **8**, 1988.

[4] HARAUX, A., *Nonlinear Evolution Equation Global Behaviour of Solutions*, Lecture Notes in Math. 841, Springer-Verlag, 1981.

[5] HEIJMANS, H. J. A. M., Structured Population, Linear Semigroups and Positivity, *Math. Z.* **191** (1986).

[6] LIONS, J. L. and STRAUSS, W. A., Some Nonlinear Evolution Equations, *Bull. Soc. Math. France*, **93** (1965).

[7] MARTIN, R. H., JR., *Applications of Semigroup Theory to Reaction–Diffusion Systems*, Lecture Notes in Math., 1248, Springer-Verlag, Berlin, 1987.

[8] MARTIN, R. H., JR., Nonlinear Perturbation of Linear Evolution System, *J. Math. Soc. Japan*, **29** (1977).

[9] METZ, J. A. J. and DIEKMAN, O. (EDS.), *The Dynamics of Physiologically Structured Populations*, Springer Lecture Notes in Biomath., 68, 1986.

[10] PAZY, A., *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, New York, 1983.

[11] RUDIN, W., *Real and Complex Analysis*, New York, 1972.

[12] SONG, J. and YU, J.-Y., *Control Theory of Human Population*, Science Press, Beijing, 1985.

[13] SONG, J. and YU, J.-Y., *Population System Control*, Springer-Verlag, New York, 1987.

[14] SVIREZHEV, YU. M. and LOGOFET, D. O., *Stability of Biological Communities*, Mir Publishers, Moscow, 1983.

[15] WEBB, G. F., *Theory of Nonlinear Age-dependent Population Dynamics*, Marcel Dekker, New York, 1985.

[16] ZHANG, SHENGHAI, On the Mild Solution and Weak Solution of Nonlinear Population Evolution Equation, *Applied Mathematics* (Journal of Chinese Universities), **5**, (1990).

[17] ZHANG, SHENGHAI and ZHU, G.-T., The Solution and Asymptoticity of Nonlinear Population Evolution Equation, *Acta Mathematicae Applicatae Sinica*, **14**, (1991).

[18] CHAN, W. L. and GUO BAO ZHU, Optimal Birth Control of Population Dynamics, *Journal of Mathematical Analysis and Applications*, **144** (1989), 532–552.

[19] SONG, J. and YU, J. Y., *Control Theory of Human Population Science Press*, Beijing, 1985.

[20] SONG, JIAN and YU, JING-YUAN, *Stability Theory of Population System and Critical Fertility Rates*, Mathematical Modeling, Vol. 2, 1981.

[21] SONG, JIAN and YU, JING-YUAN, The Forecasting of Population Evolution Process, *China Science*, No. 9, 1980.

[22] JOÓ, I. and TUAN, N. M., On Controllability of Bilinear System III (An infinite dimensional case), *Annales Univ. Sci. Budapest.*, Sectio Math., **36** (1993) 9–22.

[23] GUO, B.-Z., CHAN, W. L., A semigroup approach to age-dependent population dynamics with delay, *Commun. in partial equations*, **14** (1989), 809–832.

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 229–233

ON THE NUMBER OF EXPANSIONS $1=\sum q^{-n_i}$, II

By

P. ERDŐS and I. JOÓ

Mathematical Institute of the Hungarian Academy of Sciences and
Department of Analysis, Eötvös Loránd University, Budapest

(Received January 7, 1993)

In this paper we continue the investigations [1]–[3] about the expansions in non-integer bases. Recall that for a given value $1 < q < 2$ an expansion of 1 is a series

$$1=\sum_{i=1}^{\infty}\varepsilon_i q^{-i} \tag{1}$$

where the digits ε_i can be 0 or 1. Such a series always exists. Sometimes we substitute it by the sequence of digits, writing $1=\varepsilon_1\varepsilon_2\ldots$ (and $\varepsilon_1\ldots\varepsilon_n$ means $\varepsilon_1\ldots\varepsilon_n000\ldots$). We prove the following

THEOREM. *There are $2^{\aleph_0}$ many values q for which 1 has precisely $\aleph_0$ many expansions in basis q. The following expansions define such values q:*

$$\underbrace{1\ \ldots\ 1}_{19}\,[\underbrace{0\ \ldots\ 0}_{19}\ 1\,0\,a\,0\,0\,1\,0\,a\,0\,0\,1]^{\omega}=(\varepsilon_i^*) \tag{2}$$

where the letter "a" denotes either 0 or 1 which may be different at each occurrences. The other $\aleph_0$ expansions are $\varepsilon_1^*,\ldots,\varepsilon_{k-1}^*,0,(\varepsilon_i^*+\varepsilon_{i-1}^*)_{i=k+1}^{\infty}$, where $\varepsilon_k^*=1$ is followed by nine zeros in (2).

Remark that this answers Problems 3 and 7 in [1].

For the proof the following statements will be useful

LEMMA 1. *Suppose that 1 has an expansion (ε_i^*) beginning with n digits ($n\geq 2$), and let (ε_i) be another expansion of 1. Suppose that there exists an index $k\geq 1$ such that $\varepsilon_k=1$ and*

a) $\varepsilon_{k+1}+\ldots+\varepsilon_{k+n-1}>0$, *or*

b) *there exists* $\ell \in \mathbb{N}$ *with* $\varepsilon_{k+\ell n+1} + \ldots + \varepsilon_{k+(\ell+1)n-1} > 0$ *but* $\varepsilon_{k+jn+1} = \ldots = \varepsilon_{k+(j+1)n-1} = 0$, $\varepsilon_{k+(j+1)n} = 1$ $(j = 0, 1, \ldots, \ell - 1)$.
Then there is no expansion of 1 *beginning with the* k *digits* $\varepsilon_1 \ldots \varepsilon_{k-1} 0$.

PROOF. The existence of (ε_i^*) implies that

$$r\frac{1-r^n}{1-r} = r + \ldots + r^n \leq 1, \qquad r := \frac{1}{q}. \tag{3}$$

Denote $\alpha = \varepsilon_1 \ldots \varepsilon_{k-1}$. Then in case a) we have $\alpha + r^k + r^m \leq 1$ for some $k+1 \leq r \leq k+n-1$, and in case b) $\alpha + r^k(1 + r^n + \ldots + r^{\ell n}) + r^m \leq 1$ for some $k + \ell n + 1 \leq m \leq k + (\ell+1)n - 1$. In both cases we obtain

$$\alpha + r^k \frac{1 - r^{(\ell+1)n}}{1 - r^n} + r^{k+(\ell+1)n-1} \leq 1 \quad \text{for some } \ell \geq 0. \tag{4}$$

We have to prove that

$$\alpha + r^{k+1} + r^{k+2} + \ldots = \alpha + \frac{r^{k+1}}{1-r} < 1 \tag{5}$$

since this implies the impossibility of beginning an expansion with $\varepsilon_1 \ldots \varepsilon_{k-1} 0$. Now (4) implies (5) if we prove that

$$\frac{r^{k+1}}{1-r} < r^k \frac{1 - r^{(\ell+1)n}}{1 - r^n} + r^{k+(\ell+1)n-1}$$

i.e. that

$$r\frac{1-r^n}{1-r} < 1 - r^{(\ell+1)n} + (1 - r^n) r^{(\ell+1)n-1} = 1 + r^{(\ell+1)n-1}(1 - r - r^n).$$

Here the left side is ≤ 1 by (3) and the right is > 1 since $1 - r - r^n > 0$ again by (3). So (5) follows from (4) which finishes the proof. □

LEMMA 2. *Let* (ε_i^*) *be an expansion of* 1 *beginning with* n *consecutive digits* 1, *and suppose that there are no* n *consecutive digits* 0 *in* (ε_i^*). *Then in any other expansion of* 1 *the first modified digit cannot be an* $\varepsilon_k^* = 1$.

PROOF. As above, we have (3). On the other hand, if $\varepsilon_k^* = 1$ then denoting $\alpha = \varepsilon_1^* \ldots \varepsilon_{k-1}^*$, we have

$$1 = \alpha + r^k + \sum_{i=1}^{\infty} \varepsilon_{k+i}^* \cdot r^{k+i} \geq \alpha + r^k(1 + r^n + r^{2n} + \ldots) = \alpha + \frac{r^k}{1-r^n}. \tag{6}$$

If (ε_i) is another expansion where $\varepsilon_k^* = 1$ is the first modified digit $(\varepsilon_k = 0)$ then

$$1 = \alpha + \sum_{i=1}^{\infty} \varepsilon_{k+i} \cdot r^{k+i} \leq \alpha + r^{k+1} + r^{k+2} + \ldots = \alpha + \frac{r^{k+1}}{1-r}.$$

Taking (3) into account, we can continue by

$$1 \le \alpha + r^k \frac{r}{1-r} \le \alpha + r^k \frac{1}{1-r^n}$$

in contradiction with (6). This shows the impossibility of the expansion (ε_i). Lemma 2 is proved. □

Now we give the corresponding modification test for the digits 0. Since the above method seems to be difficult in applying to the situation below, we choose another idea.

LEMMA 1'. *Suppose that* 1 *has an expansion* (ε_i^*) *starting with* $n \ge 2$ *consecutive digits* 1 *and let* (ε_i^*) *be any expansion of* 1. *Suppose that there exists* $k \ge 1$ *with* $\varepsilon_k = 0$ *and*

a) $\varepsilon_{k+1} + \ldots + \varepsilon_{k+n-1} < n-1$ *or*

b) *there exists* $\ell \in \mathbb{N}$ *with* $\varepsilon_{k+\ell n+1} + \cdots + \varepsilon_{k+(\ell+1)n-1} < n-1$, *and* $\varepsilon_{k+jn+1} = \cdots = \varepsilon_{k+(j+1)n-1} = 1$, $\varepsilon_{k+(j+1)n} = 0$ *for* $j = 0, \ldots, \ell-1$.
Then there is no expansion of 1 *beginning with the digits* $\varepsilon_1 \ldots \varepsilon_{k-1} 1$.

PROOF. It is enough to show that

$$1 < \varepsilon_1 \ldots \varepsilon_{k-1} 1. \tag{7}$$

In case a) we change $\varepsilon_k = 0$ to 1 and compensate it by adding $-\varepsilon_{i-k}^*$ at each places $i \ge k+1$; in other words

$$1 = \varepsilon_1^*, \ldots, \varepsilon_{k-1}, 1, \varepsilon_{k+1} - \varepsilon_1^*, \varepsilon_{k+2}^* - \varepsilon_2^*, \ldots \tag{8}$$

Here $\varepsilon_{k+i} - \varepsilon_i^* \le 0$ $(i = 1, \ldots, n)$ and by a) there exists $1 \le i \le n-1$ with $\varepsilon_{k+i} - \varepsilon_i^* = -1$. We expand this digit -1, i.e. we substitute it by 0 and we add $-\varepsilon_j^*$ to the $k+i+j$-th digit, $j = 1, 2, \ldots$. Then we have ≤ -1 at the $k+n$-th place, and ≤ 0 at $k+n+1$. Analogously expand one -1 at $k+n$, after one at $k+n+1$; then at $k+2n$ we have ≤ -1, at $k+2n+1$ we have ≤ 0. Thus we expand one -1 at $k+2n$, after one -1 at $k+2n+1$, then at $k+3n$, $k+3n+1$ etc. Finally, we obtain an expansion of 1 (the convergence follows from the fact that at the i-th place stands a digit between 1 and $-i$). This expansion of 1 begins with $\varepsilon_1 \ldots \varepsilon_{k-1} 1$ and all other digits are ≤ 0, some of them are ≤ -1. This proves (7) in case a). In case b) we change as well $\varepsilon_k = 0$ to 1 and get the expansion (8). Then $\varepsilon_{k+i} - \varepsilon_i^* = 0$ $(i = 1, \ldots, n-1)$, $\varepsilon_{k+n} - \varepsilon_n^* = -1$. Expand then $\varepsilon_{k+n} - \varepsilon_n^* = -1$, after one -1 at the $k+2n$-th,…,$k+\ell n$-th places. This last step makes a situation almost identical (and even simpler to handle) as in case a), and the proof can be finished by the method of a).

LEMMA 2'. *Let* (ε_i^*) *be an expansion of* 1 *starting with* n *digits* 1 *such that* $\varepsilon_{n+1}^* = 0$ *and suppose that there are no* n *consecutive digits* 1 *after this position. Then in any other expansion of* 1 *the first modified digit cannot be an* $\varepsilon_k^* = 0$.

PROOF. Apply Lemma 1' (with $\varepsilon_i := \varepsilon_i^*$). If one of the conditions a), b) holds for (ε_i^*), then the digit $\varepsilon_k^* = 0$ cannot be modified. So it remains to investigate the case $(\varepsilon_i^*) = \varepsilon_1^* \dots \varepsilon_{k-1}^* 0(\underbrace{1}_{1} \dots \underbrace{1}_{n-1} 0)^\omega$. In this case we write 1 to the k-th position and add $-\varepsilon_i^*$ to the $k+i$-th positions. Then -1 appears at $k+n$ which we expand; after that we expand one -1 at $k+2n$, $k+3n$ etc. Finally we get an expansion of 1 starting with $\varepsilon_1^* \dots \varepsilon_{k-1}^* 1$, and after that position stand only nonpositive digits. Among them negative digits will appear since $(\varepsilon_i^*) \neq \varepsilon_1^* \dots \varepsilon_n^*$. As in Lemma 1' we conclude that $1 < \varepsilon_1^* \dots \varepsilon_{k-1}^*$ which finishes the proof. □

As an immediate corollary of Lemma 2 and Lemma 2' we obtain the point 3) of Theorem 1 in [1] on the uniqueness of expansions starting with n 1 digits if further there are no n consecutive 0 or 1 in the expansion.

PROOF OF THEOREM. We put two expansions different from (2) into the same equivalence class if the first digit different from (2) has the same index. The first modifiable digit in (2) cannot be a zero since there are no 8 consecutive 1 digits following a zero. By Lemma 2 the digits followed by 9 zeros can only be the first modified digits. Let $\varepsilon_k^* = 1$ be such a digit; change it by 0 and add ε_{i-k}^* to ε_i^*, $i \geq k+1$. Then at the places $k+1, \dots, k+9, k+10$ stand 1 digits and at $i \geq k+1$ there is no '2" digit because in (2) for $i \geq k+1$ there may be 1 digits only in the places $5\ell+1$ or $5\ell+4$ while in the shifted sequence the 1 digits may be at 5ℓ or $5\ell+3$. Thus we obtained an expansion which we denote by $(2)_k$. We show that there is no other expansion (ε_i) in the equivalence class of $(2)_k$. If (ε_i) is another element of this class, then it is a modification of $(2)_k$, where the first modified digit is at a plce $i \geq k+1$. But for $i \geq k+1$ there are no 8 consecutive 1 following a zero and 8 consecutive zeros occur only when two 1 digits follows (since in (2), digits 1 stand at all places $10\ell+9$, and in ε_{i-k}^*, in all places $i = 10\ell+8$). Hence by Lemma 1 and 2, the expansion $(2)_k$ can not be first modified at any places $i \geq k+1$. So the equivalence class of $(2)_k$ consists of one element: $(2)_k$ itself. Then all expansions of 1 are (2) and the $(2)_k$, which is $\aleph_0$ expansions indeed.

Consider the set A of the positions where the symbol "a" occurs in (2). Any subset will code such a value q for which (2) is an expansion of 1 where we write 1 instead of exactly those "a" whose position belongs to this subset. There are $2^{\aleph_0}$ many subsets of A such that any two subsets are comparable: the one is larger than the another. Then the corresponding values q are different: to larger subset belongs a larger q. Hence there are $2^{\aleph_0}$ many "good" q which produce (2)-type expansions of 1. Theorem 1 is proved. □

We formulate the open

PROBLEM 1. *If in the expansion* $1=\sum q^{-n_i}$, $1\le n_1<n_1<\dots$ *we have* $\lim(n_{i+1}-n_i)=\infty$, *then there are* $2^{\aleph_0}$ *many expansions of* 1.

PROBLEM 2. *There exists an expansion with* $\sup(n_{i+1}-n_i)=\infty$ *such that there exist precisely* $\aleph_0$ *expansions of* 1.

References

[1] P. ERDŐS, M. HORVÁTH and I. JOÓ, On the uniqueness of the expansions, *Acta Math. Hung.*, **58** (1991), 333–342.

[2] P. ERDŐS and I. JOÓ, On the expansion $1=\sum q^{-n_i}$, *Periodica Math. Hung.*, **23** (1991), 25–28.

[3] P. ERDŐS and I. JOÓ, On the number of expansions $1=\sum q^{-n_i}$, *Annales Univ. Sci. Budapest., Sectio Math.*, **35** (1992), 129–132.

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 235–246

LACUNARY SPLINE INTERPOLATION AND TWO BOUNDARY VALUE PROBLEMS

By

N. A. A. RAHMAN

Department of Numerical Analysis, Eötvös Loránd University, Budapest

(Received January 18, 1993)

1. Introduction

In this paper we give an approximation solution with lacunary spline function for the Liouville type second order differential equation, if the boundary values are given. Boundary value problems are investigated with spline functions by many authors (see[1], [8]). Our spline function is of lacunary (0,2) type.

The differential equation is

$$y''(x) + A(x)y(x) = F(x), \qquad x \in I := [0, b], \tag{1.1}$$

with the following boundary condition

$$y(0) = \alpha, \quad y(b) = \beta, \tag{1.2}$$

where $A(x)$, $F(x)$ are given continuous functions in $[0, b]$.

In our paper, we give not only an approximating solution of the problem (1.1) with (1.2), if the exact solution exist, but a criterion is also given under which the solution of this problem uniquely exists.

Several authors investigated these problems and a number of approximating methods exist. For example, it is known that, if $y \in C^4(I)$,

$$A(x), F(x) \in C(I), \qquad A(x) \geq 0 \quad \text{for all } x \in I,$$

then this problem has a unique solution see [10]. Here the condition $A(x) \geq 0$ for all $x \in I$ was not necessary. It is enough that $y(x) \in C^2(I)$.

2. The Definition of the (0,2)-interpolational Spline Functions

In the interval $I=[0,b]$, let the system of grid points

$$(2.1)\qquad \Delta:=\left\{x_i=x_{i,n}=i\frac{b}{n}\right\},$$

$$0=x_0<x_1<\ldots<x_{n-1}<x_n=b,\quad i=\overline{0,n},\quad n=2,3,\ldots$$

be given. Let $\overline{y}_i=\overline{y}_{i,n}$, $\overline{y}_i''=\overline{y}_{i,n}''$ denote the approximate values of the exact values $y_i=y_{i,n}=y(x_{i,n})$, $y_i''=y_{i,n}''=y''(x_{i,n})$, respectively to the grid points system (2.1). We define the (0,2)-interpolational spline function $S_\Delta(x,y)$ corresponding to the function $y(x)$ as follows:

$$(2.2)\qquad S_\Delta(x,y)\equiv S_\Delta(x)\equiv S_i(x)=$$

$$=\overline{y}_i+a_1^{(i)}(x-x_i)+\overline{y}_i''\frac{(x-x_i)^2}{2}+a_2^{(i)}(x-x_i)^3,$$

where $x\in I_i=[x_i,x_{i+1}]\subset I$, $i=\overline{0,n-1}$, $n=2,3,4,\ldots$.

In (2.2) the coefficients $a_1^{(i)}$, $a_2^{(i)}$ are chosen such that the following equations will be satisfied

$$(2.3)\qquad \begin{array}{lll}
\text{(a)} & S_\Delta(x_i,y)=S_\Delta(x_i)=S_i(x_i)=\overline{y}_i & i=\overline{0,n-1};\\
\text{(b)} & S_{n-1}(x_n)=\overline{y}_n; & \\
\text{(c)} & S_i(x_{i+1})=S_{i+1}(x_{i+1})=\overline{y}_{i+1}, & i=\overline{0,n-2};\\
\text{(d)} & S_i''(x_i)=\overline{y}_i'', & i=\overline{0,n-1};\\
\text{(e)} & S_{n-1}''(x_n)=\overline{y}_n''; & \\
\text{(f)} & S_i''(x_{i+1})=S_{i+1}''(x_{i+1})=\overline{y}_{i+1}'', & i=\overline{0,n-2}.
\end{array}$$

The equations (2.3) (a) and (d) are obviously satisfied by (2.2). Furthermore, the equations (2.3) (b), (2.3) (c), (2.3) (e) are satisfied if

$$(2.4)\qquad \begin{array}{lll}
(\alpha) & a_1^{(i)}+a_2^{(i)}h^2=\dfrac{1}{h}(\overline{y}_{i+1}-\overline{y}_i)-\dfrac{h}{2}\overline{y}_i'', & i=\overline{0,n-1},\\
(\beta) & a_2^{(i)}=\dfrac{1}{6h}(\overline{y}_{i+1}''-\overline{y}_i''), & i=\overline{0,n-1},
\end{array}$$

where $h=x_{i+1}-x_i$, $i=\overline{0,n-1}$, that is our grid points system (2.1) is equidistant. From (2.2) and the equation (2.3) (c), (2.4) (α) it follows obviously that the $S_\Delta(x,y)\equiv S_i(x,y)$ spline functions are continuous in I, and from (2.2), (2.3) (f) and (2.4) (β) it follows that the $S_\Delta''(x,y)\equiv S_i''(x,y)$ spline functions are continuous in I, too. We call the spline functions $S_\Delta(x,y)$ interpolational functions of (0,2)-type (see [2], [4], [6], [9]) by the equations (2.3) (a), (b), (d) and (e), that is, they interpolate the values $\overline{y}_i$, $\overline{y}_i''$ at the grid points, but they don't interpolate the function $\overline{y}_i'$. We note that the function $S_\Delta(x,y)$ is identical with a polynomial of degree 3 by (2.2) on each subinterval $I_i \subset I$, $i=\overline{0,n-1}$, and by (2.4) its degree is minimal.

From the previous result we have the following

THEOREM 1. *The (0,2)-interpolational spline functions* $S_\Delta(x,y) \in$ $\in C(I)$ *corresponding to the function* $y(x)$, *satisfying the equations* (2.3) *for the system of grid points* (2.1), *uniquely exist.*

We note that we don't need an equidistant subdivision in I, the Theorem 1 remains valid for any nonuniform division, where $x_{i+1}-x_i=h_i$, $i=\overline{0,n-1}$ i.e. the h_i's are different.

3. First Approximational Method

For the construction of $S_\Delta(x)$ in (2.2) we need an iterative method to calculate the approximate values $\overline{y}_{i,n}$, $\overline{y}_{i,n}''$, $i=\overline{0,n-1}$, $n=2,3,\ldots$. To get the values $\overline{y}_i$ one can use different kind of numerical algorithms applicable for the numerical solution of problem (1.1) and (1.2). Here we give an algorithm which in several cases can be used for numerical calculation too, but the main result is an existence and uniqueness theorem for problem (1.1) with (1.2) under very general conditions. With this algorithms first we need to calculate the approximate value $\overline{y}_0' \approx y_0' = y'(0)$ and then we can calculate the approximate values $\overline{y}_{i,n}$, $\overline{y}_{i,n}''$, $i=\overline{0,n-1}$, $n=2,3,\ldots$.

$$\overline{y}_{0,n}=y_{0,n}=y(0)=\alpha, \qquad \overline{y}_{n,n}=y_{n,n}=y(b)=\beta.$$

Since $y(x)\in C^{(2)}(I)$, by using (1.1), for $y_i=y_{i,n}=y(x_{i,n})$, $i=\overline{0,n-1}$ we have;

$$y_i=y(x_i)=y_0+y_0'x_i+\int_0^{x_i}\int_0^{t_1} y''(t_2)dt_2dt_1=y_0+y_0'x_i-$$

$$-\int_0^{x_i}\int_0^{t_1} A(t_2)y(t_2)dt_2dt_1+\int_0^{x_i}\int_0^{t_1} F(t_2)dt_2dt_1=y_0+y_0'x_i-$$

$$-\int_0^{x_i}\int_0^{t_1} A(t_2)\left[y_0+y_0't_2+\int_0^{t_2}\int_0^{t_3} y''(t_4)dt_4dt_3\right]dt_2dt_1+\int_0^{x_i}\int_0^{t_1} F(t_2)dt_2dt_1=$$

$$=y_0\left[1-\int_0^{x_i}\int_0^{t_1} A(t_2)dt_2dt_1\right]+y_0'\left[x_i-\int_0^{x_i}\int_0^{t_1} A(t_2)t_2dt_2dt_1\right]+$$

$$+\int_0^{x_i}\int_0^{t_1} F(t_2)dt_2dt_1+\int_0^{x_i}\int_0^{t_1} A(t_2)\left[\int_0^{t_2}\int_0^{t_3} y''(t_4)dt_4dt_3\right]dt_2dt_1=$$

$$=y_0B_1^{(i)}+y_0'C_1^{(i)}+D_1^{(i)}+I_2^{(i)}(y''(t_4)).$$

Repeating this process k-times we have the following relation

$$y_i=y_{i,n}=y_0B_k^{(i)}+y_0'C_k^{(i)}+D_k^{(i)}+I_{2k+2}^{(i)}(y''(t_{2k+2})),$$

where

$$\tag{3.1} \begin{aligned} B_k^{(i)}&=1+\sum_{r=1}^{k}(-1)^r I_{2r}(A(t_{2r})),\\ C_k^{(i)}&=x_i+\sum_{r=1}^{k}(-1)^r I_{2r}(A(t_{2r})t_{2r})\\ D_k^{(i)}&=\sum_{r=1}^{k}(-1)^{r+1} I_{2r}(F(t_{2r})) \end{aligned}$$

and

$$I_{2r}^{(i)}(W)=\int_0^{x_i}\int_0^{t_1} A(t_2)\int_0^{t_2}\int_0^{t_3} A(t_4)\times\ldots\times$$

$$\times A(t_{2r-2})\int_0^{t_{2r-2}}\int_0^{t_{2r-1}} W dt_{2r}dt_{2r-1}\ldots dt_1.$$

The symbol $\times$ denotes the double integral and W in $B_k^{(i)}$, $C_k^{(i)}$, $D_k^{(i)}$ is $A(t_{2r})$, $A(t_{2r})t_{2r}$, $F(t_{2r})$ respectively, $b\geq t_{2r}\geq 0$, $r=1,2,\ldots$.

It is easy to show that these sequences C_k, B_k, D_k satisfy the Cauchy convergent criterion, i.e. $C_k^{(i)}$, $B_k^{(i)}$, $D_k^{(i)}$ are convergent in $x\in I$, if $k\to\infty$.

Now we give the proof of the convergence of the series $C_k^{(i)}$, $k=1,2,\dots$. For the series B_k, D_k the proof is similar. In the proof the following well-known facts will be applied, if $a>0$ is an arbitrary real number then $a^{\frac{1}{N}} \to 1$, if $N\to\infty$, and $N^{\frac{1}{N}} \to 1$, if $N\to\infty$. Moreover if γ is an arbitrary fixed real number, $[\gamma]=$ greatest integer less than or equal γ, then if q is a given fixed real number $0<q<1$ then there exists $N_0(q)$ such that

$$\left(\frac{[\gamma]+1}{N}\right)^N \le q^N, \qquad N \ge N_0(q).$$

Let m_1, m_2 be arbitrary natural numbers, $m_2 \ge m_1$. From (3.1) and since $t_{2n} \le x_i \le b$, $i=\overline{0,n-1}$, let $M=\max_{x\in I}|A(x)|$, we have

$$|C_{m_2}^{(i)} - C_{m_1}^{(i)}| = \left|\sum_{r=m_1+1}^{m_2} (-1)^r I_{2r}(A(t_{2r})t_{2r})\right| \le$$

$$\le |I_{2m_1+2}(A(t_{2m_1+2})t_{2m_1+2})| + |I_{2m_1+4}(A(t_{2m_1+4})t_{2m_1+4})| + \dots$$
$$\dots + |I_{2m_2}(A(t_{2m_2})t_{2m_2})| \le$$
$$\le b\left(\frac{M^{m_1+1}b^{2m_1+2}}{(2m_1+2)!} + \frac{M^{m_1+2}b^{2m_1+4}}{(2m_1+4)!} + \dots + \frac{M^{m_2}b^{2m_2}}{(2m_2)!}\right).$$

By using the well-known Stirling formula (see[7])

$$(3.2) \qquad (2m+2)! = 2\sqrt{\pi(m+1)}(2m+2)^{2m+2}e^{-(2m+2)}(1+\omega_{2m+2})$$

where $0<\omega_{2n+2} \le e^{\frac{1}{2(n+1)}} - 1$, we obtain

$$|C_{m_2}^{(i)} - C_{m_1}^{(i)}| \le b\left[\frac{1}{\sqrt{4\pi(m_1+1)}}\left(\frac{M^{\frac{1}{2}}eb}{2m_1+2}\right)^{2m_1+2} + \right.$$
$$\left. + \frac{1}{\sqrt{4\pi(m_1+2)}}\left(\frac{M^{\frac{1}{2}}eb}{2m_1+4}\right)^{2m_1+4} + \dots + \frac{1}{\sqrt{2\pi(m_2)}}\left(\frac{M^{\frac{1}{2}}eb}{2m_2}\right)^{2m_2}\right] \le$$
$$\le \frac{bq^{2m_1+2}}{\sqrt{4\pi(m_1+1)}}(1+q^2+q^4+\dots+q^{2(m_2-m_1)-2}+\dots) \le$$
$$\le \frac{b}{\sqrt{4\pi(m+1)}} q^{2m_1+2}\left(\frac{1}{1-q^2}\right),$$

where $q = \frac{[M^{\frac{1}{2}}eb]+1}{2m_1+2}$, $0<q<1$ is an arbitrary fixed number for $m_1 \geq m_1^*(q)$ and $m_1^*(q)$ is aldo fixed, i.e. $|C_{m_1}^{(i)} - C_{m_2}^{(i)}| \to 0$ if $m_1, m_2 \to \infty$, i.e. the sequence $C_k^{(i)}$ satisfies the Cauchy convergent criterion.

Now we shall calculate the approximate value $\overline{y}'_{0,n} = \overline{y}'_0$. If $i = n$, i.e. $x_n = b$, we have by using (1.2), (2.1)

$$\beta = \alpha B_k^{(n)} + y'_0 C_k^{(n)} + D_k^{(n)} + I_{2k}^{(n)}(y''(t_{2k+2})).$$

From this equation the exact value $y'_0 = y'(0)$ is

$$y'_0 = \frac{\beta - \alpha B_k^{(n)} - D_k^{(n)} - I_{2k+2}^{(n)}(y''(t_{2k+2}))}{C_k^{(n)}}, \quad k = 1,2\ldots$$

Now we define the approximate values $\overline{y}'_0 = \overline{y}'_{0,n}$ with the following formula

$$\overline{y}'_0 = \overline{y}'_{0,n} = \frac{\beta - \alpha B_k^{(n)} - D_k^{(n)}}{C_k^{(n)}}, \quad n = 2,3,\ldots, \quad k = 1,2,\ldots$$

We proved that the sequences $B_k^{(n)}$, $C_k^{(n)}$, $D_k^{(n)}$ are convergents, i.e. $i = n_1$, therefore if $\lim_{n\to\infty} C_k^{(n)} = c \neq 0$, then the approximate values $\overline{y}'_0 = \overline{y}'_{0,n}$ exist.

So by using (3.1), (3.2) it is easy to see from these equalities that

$$(3.3)\quad |y'_0 - \overline{y}'_0| = \left|\frac{I_{2k+2}^{(n)}(y''(t_{2k+2}))}{C_k^{(n)}}\right| = \frac{1}{|C_k^{(n)}|}\left|\int_0^b\int_0^{t_1} A(t_2)\int_0^{t_2}\int_0^{t_3} A(t_4)\times\ldots\right.$$

$$\left.\ldots\times A(t_{2k})\int_0^{t_{2k}}\int_0^{t_{2k+1}} y''(t_{2k+2})dt_{2k+2}\ldots dt_1\right| \leq$$

$$\leq \frac{M^{k+1}M^* b^{2k+2}}{|C_k^{(n)}|(2k+2)!} \leq \frac{K_1}{\sqrt{k+1}}\left(\frac{|C_k^{(n)}|^{-\frac{1}{2k+2}} M^{\frac{1}{2}} eb}{2k+2}\right)^{2k+2} = \frac{K_1 q^{2k+2}}{\sqrt{k+1}},$$

where

$$M^* = \max_{x\in I}|y''(x)|, \quad M = \max_{x\in I}|A(x)|, \quad K_1 = \frac{M^*}{\sqrt{4\pi}}, \quad 0<q<1,$$

q is an arbitrary fixed number for $k \geq k_1(q)$ and $k_1(q)$ is also fixed. K_1 is constant independent from n. We note that the constants K_i, $i = 1,2,\ldots$ are always independent from n.

From (3.1) define $\overline{y}_i$ by the following

$$\overline{y}_i = B_k^{(i)} y_0 + C_k^{(i)} \overline{y}_0' + D_k^{(i)}. \tag{3.4}$$

So by using (3.1), (3.3), (3.4) and triangular inequality we have by simple calculation we get

$$|y_i - \overline{y}_i| \leq \frac{|C_k^{(i)}| K_1 q^{2k+2}}{\sqrt{k+1}} + |I_{2k+2}^{(i)}(y''(t_{2k+2}))| \leq \frac{K_2 q^{2k+2}}{\sqrt{k+1}},$$

where $0 < q < 1$ fix number if $k \geq k_2(q)$, $k_2(q)$ is also fixed.

By (3.4) we have the following

THEOREM 2. *If $y(x) \in C^2(I)$, $A(x)$, $F(x) \in C(I)$ and if y_i, $\overline{y}_i$ are the exact and the approximate values of $y(x)$ respectively in the grid points (2.1) x_i, $i = \overline{1, n-1}$, then the following inequalities hold:*

$$|y_i - \overline{y}_i| \leq K_2 \frac{q^{2k+2}}{\sqrt{k+1}}, \quad i = \overline{0, n-1}. \tag{3.5}$$

So from (3.5) it follows that $\overline{y}_i = \overline{y}_{i,k} \to y_i$ very fast, if $k \to \infty$.

From (1.1), (3.4) we define $\overline{y}_i'' = -A(x_i)\overline{y}_i + F(x_i)$ in the grid points x_i, $i = \overline{1, n-1}$. From (2.2)

$$\overline{y}_0'' = y_0'' = -A(0)\alpha + F(0) \quad \text{and} \quad \overline{y}_n'' = y_n'' = -A(b)\beta + F(b).$$

By using (3.5) it is easy to see that for $\overline{y}_i''$, y_i'', $i = \overline{1, n-1}$, i.e. for the approximate and the exact values of $y''(x)$ in the grid points x_i, $i = \overline{1, n-1}$, the following inequality holds:

$$|y_i'' - \overline{y}_i''| \leq |A(x_i)(y_i - \overline{y}_i)| \leq M K_2 \frac{q^{2k+2}}{\sqrt{k+1}}, \tag{3.6}$$

where $M = \max_{x \in I} |A(x)|$, $0 < q < 1$, for $k \geq k_2(q)$ fixed number.

So we can formulate the following

THEOREM 3. *The differential equation (1.1) with boundary conditions (1.2) has a unique solution in I, if $\lim_{k \to \infty} C_k^{(n)} = c \neq 0$.*

The proof is a simple consequence of (1.1), (3.3), (3.5), (3.6).

We note that there exist several numerical algorithms for finding the approximate values of the exact values y_i, y_i'', for example finite difference method or different order finite element methods (see [10]).

We shall show that the spline approximation given in (2.2) is very useful to approximate the exact solution of the problem (1.1) because by getting the approximate values $\overline{y}_i$ we can calculate approximate values $\overline{y}''_i$, too. Moreover, this approximation will approximate $y(x)$, $y'(x)$, $y''(x)$. For the proof of the approximation theorems we need to prove the following

LEMMA 1. *For $y'_i = y'(x_i)$ the exact values of $y'(x)$ and $a_1^{(i)}$ in (2.4) (α) the following inequality holds*

$$|y'_i - a_1^{(i)}| \le \frac{2}{3}\omega(h;y'')h + \left(2 + \frac{5Mh}{6}\right)\frac{K_2 q^{2k+2}}{\sqrt{k+1}}, \tag{3.7}$$

where q in (3.4), $\omega(h;y'')$ is the modulus of continuity of y'', and $M = \max\limits_{x\in I}|A(x)|$.

PROOF. By using (2.4) (α), (3.4), (3.5), (3.6) and applying the triangular inequality we have

$$|y'_i - a_1^{(i)}| = |y'_i - \frac{\overline{y}_{i+1} - \overline{y}_i}{h} + \frac{\overline{y}''_i}{2}h + \frac{1}{6}(\overline{y}''_{i+1} - y''_i)h| =$$

$$= \Big|y'_i - \frac{\overline{y}_{i+1} - \overline{y}_i}{h} + \frac{\overline{y}''_i}{2}h + \frac{1}{6}(\overline{y}''_{i+1} - \overline{y}''_i)h + \frac{y_{i+1} - y_i}{h} + \frac{y''_i}{2}h +$$

$$+\frac{1}{6}(y''_{i+1} - y''_i)h - \frac{y_{i+1} - y_i}{h} - \frac{y_i}{2}h - \frac{1}{6}(y''_{i+1} - y''_i)h\Big| \le$$

$$\le \left|y'_i - \frac{y_{i+1} - y_i}{h} + \frac{y''_i h}{2}\right| + \frac{h}{6}|y''_{i+1} - y''_i| + \left|\frac{y_{i+1} - \overline{y}_{i+1}}{h}\right| +$$

$$+\left|\frac{y_i - \overline{y}_i}{h}\right| + |\overline{y}''_i - y''_i|\frac{h}{2} + |y''_{i+1} - \overline{y}''_{i+1}|\frac{h}{6} + |y''_i - \overline{y}''_i|\frac{h}{6}.$$

Taylor's formula yields

$$\left|y'_i - \frac{y_{i+1} - y_i}{h} + y''_i\frac{h}{2}\right| = \left|\frac{y''_i h}{2} - y''(\xi)\frac{h}{2}\right| \le \omega(h;y'')\frac{h}{2}$$

where $x_i < \xi < x_{i+1}$.

Therefore

$$|y'_i - a_1^{(i)}| \le \frac{2}{3}\omega(h;y'')h + \left(2 + \frac{5Mh}{6}\right)\frac{K_2 q^{2k+2}}{\sqrt{k+1}}. \ \square$$

Now by using (2.2), (3.1), (2.3), (3.2) for $x \in I_i$, $i = \overline{0, n-1}$ and applying the triangular inequality we have the following convergence theorem.

THEOREM 4. *Let $y(x) \in C^{(2)}(I)$ be the exact solution of (1.1) with boundary conditions (1.2), for $x \in I_i \subset I$, $i = \overline{0, n-1}$, $S_\Delta(x)$ be the spline function in (2.2) defined on the grid points system (2.1). Then the following inequalities hold*

$$(3.8) \quad \text{(i)} \qquad |y''(x) - S''_\Delta(x)| = |y''(x) - S''_i(x)| \leq \leq 2\omega(h; y'') + 3M \frac{K_2 q^{2k+2}}{\sqrt{k+1}},$$

$$(3.9) \quad \text{(ii)} \qquad |y'(x) - S'_\Delta(x)| = |y'(x) - S'_i(x)| \leq \leq \frac{8}{3}\omega(h; y'')h + \left(2 + \frac{23Mh}{6}\right) \frac{K_2 q^{2k+2}}{\sqrt{k+1}},$$

$$(3.10) \quad \text{(iii)} \qquad |y(x) - S_\Delta(x)| = |y(x) - S_i(x)| \leq \frac{11}{3}\omega(h; y'')h^2 + \left(3 + \frac{16}{3}Mh^2\right) \frac{K_2 q^{2k+2}}{\sqrt{k+1}},$$

where $M = \max_{x \in I} |A(x)|$, $\omega(h; y'')$ is the modulus of continuity of y''. That is

$$|y^{(s)}(x) - S_\Delta^{(s)}(x)| \equiv |y^{(s)}(x) - S_i^{(s)}(x)| \leq K\omega(h; y'')h^{2-s} + K' \frac{K_2 q^{2k+2}}{\sqrt{k+1}}$$

$s = 0, 1, 2$, K, K' are constants independent of n.

PROOF. (i) By using (2.2), (2.4) (β), (3.6) and applying the triangular inequality for $x \in I_i \subset I$, we have

$$|y''(x) - S''_i(x)| = \left| y''(x) - \overline{y}''_i - \frac{1}{h}(\overline{y}''_{i+1} - \overline{y}''_i)(x - x_i) \right| =$$
$$= \left| y''(x) - \left[y''_i + \frac{1}{h}(y''_{i+1} - y''_i)(x - x_i) \right] + \left[y''_i + \frac{1}{h}(y''_{i+1} - y''_i)(x - x_i) \right] - \right.$$
$$\left. - \left[\overline{y}''_i + \frac{1}{h}(\overline{y}''_{i+1} - \overline{y}''_i)(x - x_i) \right] \right| \leq 2\omega(h; y'') + 3M \frac{K_2 q^{2k+2}}{\sqrt{k+1}}.$$

(ii) By using (2.2), (3.7), (3.10) and applying the Taylor formula and the triangular inequality for $x \in I_i \subset I$, we have

$$|y'(x) - S'_i(x)| = |(y'_i - a_1^{(i)}) + (y''(\eta_i) - S''_i(\eta_i))(x - x_i)| \leq$$

where $x_i < \eta_i < x_{i+1}$

$$\leq |y'_i - a_1^{(i)}| + |y''(\eta_i) - S''_i(\eta_i)|h \leq$$
$$\leq \frac{2}{3}\omega(h; y'')h + \left(2 + \frac{5Mh}{6}\right) \frac{K_2 q^{2k+2}}{\sqrt{k+1}} + \omega(h; y'')h + 3Mh \frac{K_2 q^{2k+2}}{\sqrt{k+1}} \leq$$

$$\leq \frac{8}{3}\omega(h;y'')h + \left(2 + \frac{23Mh}{6}\right)\frac{K_2 q^{2k+2}}{\sqrt{k+1}}.$$

(iii) By using (2.2), (3.8), (3.9) and applying the triangular inequality for all $x \in I_i \subset I$, we have

$$|y(x) - S_i(x)| = |(y_i - \overline{y}_i) + (y'(x_i) - S_i'(x_i))(x - x_i) + $$
$$+ (y''(\eta_i^*) - S_i''(\eta_i^*))\frac{(x - x_i)^2}{2}\Big| \leq$$

where $x_i < \eta_i^* < x_{i+1}$,

$$\leq \frac{K_2 q^{2k+2}}{\sqrt{k+1}} + \frac{8}{3}\omega(h;y'')h^2 + \left(2 + \frac{23Mh^2}{6}\right)\frac{K_2 q^{2k+2}}{\sqrt{k+1}} + \omega(h;y'')h^2 +$$
$$+ \frac{3}{2}Mh^2\frac{K_2 q^{2k+2}}{\sqrt{k+1}} \leq \frac{11}{3}\omega(h;y'')h^2 + \left(3 + \frac{16}{3}Mh^2\right)\frac{K_2 q^{2k+2}}{\sqrt{k+1}}. \square$$

Now we shall give the approximation of the solution of the (1.1) satisfying (1.2) with $S_\Delta(x,y)$ spline function in (2.2) in $x \in I_i \subset I$, $i = \overline{0, n-1}$.

Indeed, the spline function $S_\Delta(x)$ in (2.2) satisfies (1.2), i.e.

$$S_\Delta(0,y) = S_\Delta(0) = \alpha, \quad S_\Delta(b,y) = S_{n-1}(b) = \beta$$

boundary conditions, while $\overline{y}_0 = y_0 = y(0) = \alpha$, $\overline{y}_n = y_n = y(b) = \beta$.

THEOREM 5. *Let $y(x) \in C^{(2)}(I)$ be the exact solution of (1.1) with boundary condition (1.2). The spline function $S_\Delta(x)$ in (2.2) defined on the grid points system (2.1), for each $x \in I_i \subset I$ satisfies the following inequality*

(3.11)
$$|S''_\Delta(x,y) + A(x)S_\Delta(x,y) - F(x)| \leq$$
$$\leq \left(2 + \frac{11Mh}{3}\right)\omega(h;y'') + \left(6 + \frac{16}{3}Mh^2\right)M\frac{K_2 q^{2k+2}}{\sqrt{k+1}},$$

where $M = \max_{x\in I}|A(x)|$, $\omega(h;y'')$ is the modulus of continuity of y''.

PROOF. By using (2.2), (1.1), (3.8), (3.10), and applying the triangular inequality we have

$$|S_i''(x,y) - A(x)S_i(x,y) - F(x)| \leq$$
$$\leq |S_i''(x,y) - y''(x)| + |A(x)||S_i(x,y) - y(x)| \leq$$
$$\leq 2\omega(h;y'') + 3M\frac{K_2 q^{2k+2}}{\sqrt{k+1}} + \frac{11M}{3}\omega(h;y'')h + \left(3 + \frac{16}{3}Mh^2\right)M\frac{K_2 q^{2k+2}}{\sqrt{k+1}} =$$
$$= \left(2 + \frac{11Mh}{3}\right)\omega(h;y'') + \left(6 + \frac{16}{3}Mh^2\right)M\frac{K_2 q^{2k+2}}{\sqrt{k+1}}. \square$$

REMARK 1. We note that our approximational method is stable in the sence that, if we construct functions $S^*_\triangle(x,y)$, similarly as in (2.2) with the values y^*_i, $y^{*\prime\prime}_i$, $y^{*\prime\prime}_i = -A(x_i)y^*_i + F(x_i)$ instead of $\overline{y}_i$, $\overline{y}''_i = -A(x_i)\overline{y}_i + F(x_i)$ then we can prove theorems to similar theorems 4 and 5, if the inequality

$$|\overline{y}'_0 - y^{*\prime}_0| \le \frac{K_1 q^{2k+2}}{\sqrt{k+1}}, \quad |\overline{y}_i - y^*_i| \le \frac{K_2 q^{2k+2}}{\sqrt{k+1}}$$

for $i = \overline{1, n-1}$ holds and $\overline{y}^*_0 = y_0 = \alpha$, $\overline{y}^*_n = y(b) = \beta$. This statement is easy to prove.

REMARK 2. If $y(x) \in C^{(3)}(I)$, then the spline function $S_\triangle(x,y)$ in (2.2) approximates with the best order, i.e.

$$|y^{(s)}(x) - S^{(s)}_\triangle(x)| = \tag{3.12}$$

$$= |y^{(s)}(x) - S^{(s)}_i(x,y)| \le K\omega(h; y'')h^{3-s} \quad s = 0,1,2,3.$$

This inequality is easy to prove.

So by means of (3.12) with $S_\triangle(x,y)$ in I we can give an approximation of the solution of the third order differential equation

$$y'''(x) + A_1(x)y''(x) + A_2(x)y'(x) + A_3(x)y(x) = G(x)$$

if the initial values $y(0) = \alpha_1$, $y'(0) = \alpha_2$, $y''(0) = \alpha_3$ are given.

REMARK 3. We can generalize our approximational method if $y(x) \in$ $\in C^{(r)}(I)$, $r \ge 2$, but in this case the $S_i(x)$ polynomials $i = \overline{0, n-1}$ are of degree $r+1$. We have the following inequalities;

$$|y^{(s)}(x) - S^{(s)}_i(x,y)| \le K\omega(h; y^{(r)})h^{r-s} + R^*, \quad s = \overline{0,r},$$

i.e. we approximate $y^{(s)}(x)$ with *best order*, $s = \overline{0,r}$.

Numerical examples are given in [5].

References

[1] J. H. AHLBERG, E. N. NILSON, J. IL. WALSH, *Theory of Splines and Their Applications*, Acad. Press, New York and London, 1967.

[2] S. DEMKO, Lacunary Polynomial Spline Interpolation, *SIAM. J. Numer. Anal.*, **13** (1976), 369–381.

[3] TH. FAWZY, F. S. HOLAIL, W. A. EL-GANAYNI, *A Spline Solution for the* $(0,m)$ *Lacunary Interpolation Problem.*

[4] J. GYŐRVÁRI, Lakunare Interpolation mit Spline-Funktion, *Acta Math. Hung.*, **42** (1983) 23–33.

[5] J. KŐHEGYI, N. A. A. RAHMAN, *Numerical Solution of the Two Boundary Value Problems with Lacunary Interpolation Spline Functions* (to appear).

[6] A. MEIR, A. SHARMA, Lacunary Interpolation by Splines, *SIAM J. Numer. Anal.*, **10** (1973), 433–442.

[7] I. P. NATANSOM, *Constructive Function Theory*, Ungar, New York, Vol. **3**, 1965.

[8] L. SCHUMAKER, *Spline Functions (Basic Theory)*, John Wiley and Sons, New York, 1981.

[9] B. K. SCHWARTZ, R. S. VARGA, A Note on Lacunary Interpolation by Splines, *SIAM J. Numer. Anal.*, **10** (1973), 443–447.

[10] J. STOER, R. BULIRSCH, *Introduction to Numerical Analysis*, Springer-Verlag, New York Inc., 1980.

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 247–252

UNICITY OF SOLUTION OF THE MAXWELL EQUATIONS WITH DISCONTINUOUS COEFFICIENTS IN INFINITE DOMAIN

By

RUSZLÁN H. FARZAN

Department of Numerical Analysis, Eötvös Loránd University, Budapest

(Received February 19, 1993)

1. The system of the Maxwell equations written in frequency variables in $\mathbb{R}^3$ is ([1])

$$\text{rot}\mathbf{E} = i\omega\mu\mathbf{H}, \qquad \text{rot}\mathbf{H} = i\omega\epsilon'\mathbf{E} + \mathbf{j}, \tag{1}$$

where **E**, **H** are vectors of electric and magnetic fields and **j** is the density distribution of electric sources with compact support. Furthermore,

$$\epsilon' = \epsilon + i\sigma/\omega, \tag{2}$$

where ϵ, ω and μ are constant positive parameters. On σ we suppose that $\mathbb{R}^3$ can be decomposed into a finite number of domains in which σ is a continuous positive function or $\sigma \equiv 0$; the surfaces of these domains are of Ljapunov's type (or they consists of a finite number of such surfaces).

So σ is a piecewise continuous function. Denote by S the surface of discontinuity of σ. On S the following boundary conditions must hold: the tangential components of the vectors **E** and **H**:

$$E_\tau, H_\tau \quad \text{are continuous.} \tag{3}$$

The uniqueness of the solution of the problem (1) has been discussed under the following two assumptions on σ.

— There exists a ball V_R of a large radius R such that outside V_R σ is a continuous positive function. Then at infinity the vectors of the electromagnetic field satisfy the following conditions ([1])

$$\lim_{R\to\infty} R|\mathbf{E}| = 0, \quad \lim_{R\to\infty} R|\mathbf{H}| = 0, \quad R = |\mathbf{R}|. \tag{4}$$

— Outside the ball V_R σ is equal to zero. Then the radiation conditions can be used for $R \to \infty$ (see [1,2])

$$\begin{gathered} |\mathbf{E}| = O(R^{-1}), \quad |\mathbf{H}| = O(R^{-1}), \\ |\mathbf{E} + \sqrt{\mu/\epsilon}(\mathbf{n}\times\mathbf{H})| = o(R^{-1}), \end{gathered} \tag{5}$$

where $\mathbf{n}$ is the external normal to the boundary S_R of V_R.

In these cases the unicity theorems for the solution of the Maxwell equations have been proved in [2]–[9] etc.

In some problems of electrodynamics however there doesn't exist a ball V_R so that outside of it σ is continuous and either strongly positive or equals to zero. The model of the stratified Earth under electromagnetic sounding can be example of this situation ([10]). The boundaries between layers — the surfaces of discontinuity of σ, — are infinite planes parallel to each other. In some of the layers σ can be zero. For example, this is approximately true in the air, which is the upper half-space in the model. It's clear that now σ is not continuous outside any large ball V_R and it can vanish on some subsets outside V_R.

In this paper we shall consider such cases.

It will be proved that system (1) may have at most one solution, satisfying (5) in those subdomains where $\sigma \equiv 0$ and (4), i.e.

$$|\mathbf{E}| = o(R^{-1}), \quad |\mathbf{H}| = o(R^{-1}), \tag{6}$$

where $\sigma > 0$.

2. The system (1) can be rewritten as one vector equation

$$\mathrm{rotrot}\mathbf{E} = k^2\mathbf{E} + i\omega\mu\mathbf{j}, \quad k^2 = \omega^2\mu\epsilon'. \tag{7}$$

At first we prove the unicity theorem in the case of positive coefficient.

THEOREM 1. *If σ doesn't vanish anywhere in $\mathbb{R}^3$ the solution of equation* (7) *is unique.*

PROOF. Let V_R be decomposed into subdomains G_i such that in every G_i k is continuous. Let equation (7) have two solutions: $\mathbf{E}^{(1)}$ and $\mathbf{E}^{(2)}$. Then their difference $\mathbf{E}^r$ satisfies equation

$$\mathrm{rotrot}\mathbf{E}^r = k^2\mathbf{E}^r \tag{8}$$

in all G_i with boundary conditions (3) and conditions (6) at infinity. Because k is continuous in G_i $\mathbf{E}^r$ is continuous with its second derivatives ([11]). Let

$$\mathbf{B} := \mathrm{rot}\mathbf{E}^r \tag{9}$$

Furthermore, let us multiply the equation (8) by $\overline{\mathbf{E}^r}$ (complex conjugate) and integrate over G_i. Now we have

$$\int_{G_i} \overline{\mathbf{E}^r}\mathrm{rot}\mathbf{B}d\mathbf{R} = \int_{G_i} k^2|\mathbf{E}^r|^2 d\mathbf{R}. \tag{10}$$

By Green's formula (see [1]) and (9)

$$\int_{G_i} \overline{\mathbf{E}^r} \operatorname{rot} \mathbf{B} d\mathbf{R} = \int_{S_i} (\mathbf{B} \times \overline{\mathbf{E}^r}) \mathbf{n} dS + \int_{G_i} |\mathbf{B}|^2 d\mathbf{R}, \tag{11}$$

where S_i is the boundary of G_i and $\mathbf{n}$ is the external normal to S_i.

Let (ξ, η, ζ) be the local coordinate system in a given point of S_i such that the direction of ζ is given by $\mathbf{n}$. Then the integrand of the surface integral in (11) can be written as

$$B_\xi \overline{\mathbf{E}^r_\eta} - B_\eta \overline{\mathbf{E}^r_\xi}. \tag{12}$$

It means that (12) is a combination of the tangential components of vectors $\mathbf{B}$ and $\mathbf{E}^r$. Because of (1) $\mathbf{B}$ differs from $\mathbf{H}^r$ only in a constant multiplier. Therefore, by (3), the expression (12) remains continuous across the surface of discontinuity S_i.

Let us denote by $S_i^{(1)}$ the part of S_i which is inside V_R and by $S_i^{(2)}$ the part of S_i which lies on the sphere S_R. Let G_i and G_j be neighbouring subdomains and suppose that $S_i^{(1)}$ coincides with $S_j^{(1)}$. On these surfaces the external normals have opposite directions. On the other hand expression (12) is continuous across S_i (and S_j). Therefore the sum of the surface integrals over $S_i^{(1)}$ and $S_j^{(1)}$ equals to zero.

Now if we sum up the equations in (11) for all G_i then the sum of all the surface integrals over the surfaces $S_i^{(1)}$ equals to zero and so only the integral over the sphere S_R remains ([12]). Therefore (10) implies

$$\int_{S_R} (\mathbf{B} \times \overline{\mathbf{E}^r}) \mathbf{n} dS + \int_{V_R} |\mathbf{B}|^2 d\mathbf{R} = \int_{V_R} k^2 |\mathbf{E}^r|^2 d\mathbf{R}. \tag{13}$$

The integrand of the surface integral (which is the same as (12)) can be estimated by $o(R^{-2})$ as $R \to \infty$ because of (6) and (1), (9). So if R tends to infinity the surface integral tends to zero, and from (13) we obtain

$$\int_{\mathbb{R}^3} |\mathbf{B}|^2 d\mathbf{R} = \int_{\mathbb{R}^3} k^2 |\mathbf{E}^r|^2 d\mathbf{R}. \tag{14}$$

k^2 is a complex function with strongly positive real and imaginary parts in $\mathbb{R}^3$, namely

$$\operatorname{Re}(k^2) = \omega^2 \mu \epsilon > 0, \quad \operatorname{Im}(k^2) = \omega \mu \sigma > 0. \tag{15}$$

Therefore the equality (14) is true only if $\mathbf{E}^r \equiv 0$. □

3. Let us consider the general case when σ can vanish in some subdomains of $\mathbb{R}^3$.

THEOREM 2. *The solution of the equation* (7) *remains unique if* σ *vanishes in some subdomains of* $\mathbb{R}^3$.

PROOF. It is easy to show that if $\sigma = 0$ the equalities (10) and (13) remain true.

Let σ vanish on the subset $S_R{}^*$ of S_R. Then on $S_R{}^*$ the following holds

$$\mathrm{Re}(k^2) = \omega^2\mu\epsilon > 0, \quad \mathrm{Im}(k^2) = 0, \tag{16}$$

and on $S_R \setminus S_R{}^*$ the relations (15) are true.

Let

$$\mathbf{D} := \mathbf{E}^r + \sqrt{\mu/\epsilon}(\mathbf{n} \times \mathbf{H}^r). \tag{17}$$

According to (5) on $S_R{}^*$

$$|\mathbf{D}| = o(R^{-1}). \tag{18}$$

Furthermore, multiplying $\mathbf{D}$ by $\overline{\mathbf{E}^r}$ and using (9) we get

$$\mathbf{D}\overline{\mathbf{E}^r} = |\mathbf{E}^r|^2 + (ik)^{-1}(\mathbf{n} \times \mathbf{B})\overline{\mathbf{E}^r},$$

from which we have

$$(\mathbf{B} \times \overline{\mathbf{E}^r})\mathbf{n} = ik\mathbf{D}\overline{\mathbf{E}^r} - ik|\mathbf{E}^r|^2. \tag{19}$$

Substituting (19) into (13) we obtain

$$\int_{S_R{}^*} ik(\mathbf{D}\overline{\mathbf{E}^r})dS + \int_{S_R \setminus S_R{}^*} (\mathbf{B} \times \overline{\mathbf{E}^r})\mathbf{n}dS + \int_{V_R} |\mathbf{B}|^2 d\mathbf{R} = \int_{S_R{}^*} ik|\mathbf{E}^r|^2 dS + \int_{V_R} k^2|\mathbf{E}^r|^2 d\mathbf{R}. \tag{20}$$

By (6) and (18) the surface integrals on the left hand side of (20) vanish if R tends to infinity, and the integral over V_R is real and strongly positive.

Since on $S_R{}^*$ k is real and positive: $k = \omega\sqrt{\mu\epsilon}$ thus the surface integral on the right hand side is imaginary with a nonnegative imaginary part. The integral over V_R on the right hand side is complex with nonnegative imaginary part.

Consequently, in those subdomains G_j where $\sigma > 0$ we obtain $\mathbf{E}^r \equiv 0$ since $\mathrm{Im}(k^2) > 0$ there. This implies $\mathbf{H}^r \equiv 0$ in G_j.

Let us show that in the domain G_i where $\sigma = 0$, $\mathbf{E}^r \equiv 0$ will hold. The domain G_i is enclosed by domains G_j where $\sigma \neq 0$. (If for two neighbouring domains $\sigma = 0$ then k =const and because of our definition of G_j they make one domain.) Note that the domain G_i can be unbounded.

Now if in G_j $\sigma \neq 0$, in this domain $\mathbf{E}^r \equiv 0$, $\mathbf{H}^r \equiv 0$ and according to (3) on the boundary of G_i the tangential components of these vectors equal to zero: $E^r_\tau = H^r_\tau = 0$. Furthermore,

$$\frac{\partial H^r_\eta}{\partial \xi} - \frac{\partial H^r_\xi}{\partial \eta} = i\omega\epsilon E^r_\zeta.$$

Therefore on this boundary $E^r_\zeta = 0$ and analogously $H^r_\zeta = 0$. So, on the boundaries of the domain G_i with $\sigma = 0$

$$\mathbf{E}^r = \mathbf{H}^r = 0. \tag{21}$$

Since k =const, from (1) we obtain $\operatorname{div}\mathbf{E}^r \equiv 0$ in G_i. Therefore in this domain the equation (8) can be rewritten as

$$\Delta \mathbf{E}^r + k^2 \mathbf{E}^r = 0, \tag{22}$$

i.e. we get three independent homogeneous scalar Helmholtz equations. It is known ([13]), that these equations with boundary conditions (21) and conditions (5) at infinity have only trivial solutions: $\mathbf{E}^r \equiv 0$. Thus $\mathbf{E}^r$ must be equal to zero in the domains where $\sigma = 0$ as well. □

ACKNOWLEDGEMENT. I would like to express my thanks to Prof. LÁSZLÓ SIMON for usefull duscussion.

References

[1] В. И. ДИМИТРИЕВ, Е. В. ЗАХАРОВ, Интегральные уравнения в краевых задачах электродинамики, Изд. МГУ, М., 1987.

[2] Л. А. ВАЙНШТЕЙН, *Электромагнитные волны*, Сов. Радио, М., 1957.

[3] P. WERNER, On the exterior boundary value problem of perfect reflection for stationary electromagnetic wave fields, *J. Math. Analysis and Applications*, **7** (1963).

[4] C. MÜLLER, Foundations of the mathematical theory of electromagnetic waves, Springer, Berlin, 1969.

[5] A. J. POGGIO, E. K. MILLER, Solution of three dimetional scattering problem for electromagnetic waves by using an integral equation, *Computer techniques for electromagnetics*, Ed.: R. Mitra, Pergamon Press, Oxford, 1973.

[6] Е. В. Захаров, О единственности и существовании решения интегральных уравнений электродинамики неоднородных сред, *Вычислительные методы и программирование*, XXIV. Изд. МГУ, М. 1984.

[7] А. С. Ильинский, А. Г. Свешников, Развитие методов математического моделирования в электродинамике, *Вестник МГУ, сер. Вычисл. математика и кибернетика,* **3** (1981).

[8] Б. Р. Вайнберг, *Асимптотические методы в уравнениях математической физики,* Изд. МГУ, М., 1982.

[9] V. Bezvoda, K. Segeth, Mathematical modelling in electromagnetic prospecting methods, Universitet Karlova, Praha, 1982.

[10] R. H. Farzan, *Propagation of electromagnetic waves in stratified media with local inhomogenity. Mathematical models in physics and chemistry.* Teubner-Texte für Mathematik, Band 61, 1984.

[11] S. Agmon, A. Douglis, L. Nirenberg, Estimate near boundary for solutions of elliptic PDE satisfying general boundary conditions. Part II., *Comm. Pure Appl. Math.*, **17** (1964).

[12] R. H. Farzan, Application of integral equation methods to the solution of Maxwell equations, *Colloquia Mathematica Societatis János Bolyai, 59*, Numerical methods, Miskolc'90. North Holland Publ., 1991.

[13] A. J. Burton, Numerical solution of scalar diffraction problems. In: *Numerical solution of integral equations*, L. M. Delves, J. Walsh, Oxford, 1974.

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 253–267

THE PROPERTIES OF A FAMILY OF POPULATION EVOLUTION OPERATORS

By

I. JOÓ and ZHANG SHENGHAI(*)

Department of Analysis, Eötvös Loránd University, Budapest

(Received March 21, 1993)

1. Introduction

We are concerned with a class of non-autonomous evolution equation [6], [10] in the Banach space $\Xi = L^p(0, r_m)$ $(1 \le p < \infty)$,

$$\text{(1)} \qquad \begin{cases} \frac{dp(\cdot,t)}{dt} = A(t)p(\cdot,t), \\ p(\cdot,0) = p_0(\cdot) \end{cases}$$

where the linear operators $A(t) : D(A(t)) \subset \Xi \to \Xi$, are defined as follows:

$$A(t)\varphi(\cdot) := -\varphi'(\cdot) - \mu(\cdot)\varphi(\cdot)$$

for any $\varphi(\cdot) \in D(A(t))$, and

$$D(A(t)) := \Big\{\varphi \in \Xi \,\Big|\, \varphi' \in \Xi, \mu(\cdot)\varphi(\cdot) \in \Xi, \\ \varphi(0) = \beta(t)\int_{r_1}^{r_2} k(r)h(r)\varphi(r)dr\Big\}.$$

In the equation (1), the $p(r,t)$ is the population density, r denotes age, t represents time, r_m is the maximum age, $\beta(t)$ is the specific fertility rate of females at time t, $k(r)$ and $h(r)$ denote, respectively, the female ratio and the fertility pattern; $[r_1, r_2]$ is the fertility interval with

$$\int_{r_1}^{r_2} h(r)dr = 1.$$

(*) On leaving from: Shaanxi Institute of Finance and Economics, Xian, China.

The $p_0(r)$ is an initial population density and the mortality rate $\mu(r)$ satisfies

$$\int_0^r \mu(\varrho)d\varrho < +\infty, \quad r < r_m, \quad \int_0^{r_m} \mu(\varrho)d\varrho = +\infty.$$

Generally speaking, the population parameters $\mu(r)$, $k(r)$ and $h(r)$ are time dependent. Here we assume that they are time independent functions, because they are changed slowly in certain time interval (see [11]). In order to forecast population evolution more exactly, the assumption that the specific fertility rate of females β is dependent on time t is necessary.

There are some results on stationary case of equation (1) (as [6], [5] and [13]) that means they assume that β is time independent.

It is well-known that many non-autonomous equations are very difficult to deal with [7], [2] and [1]. As for equation (1), it is more difficult to deal with, because the intersection of population evolution operators $A(t)$ is empty, moreover, for different time t_1, t_2, may be $A(t_1) \cap A(t_2) = \emptyset$.

In order to discuss the solution of equation (1), it is necessary to discuss the property of a family of population evolution operators first.

2. Preliminaries

The following lemmas will deal with the spectral properties of the family of population evolution operators.

The spectrum of operator $A(t)$ is denoted as $\sigma(A(t))$ and $\sigma_p(A(t))$ is the point spectrum of operator $A(t)$: the set of λ's such that $(\lambda I - A(t))$ is not injective.

We start with a fundamental result, which plays a crucial role when discussing the solution of equation (1).

LEMMA 2.1. *For any $t \geq 0$, there are following properties*

(i). $\sigma(A(t)) = \left\{\lambda \,\middle|\, 1 - \beta(t) \int_{r_1}^{r_2} k(r)h(r)e^{-\lambda r - \int_0^r \mu(\varrho)d\varrho} dr = 0\right\}.$

(ii). *There is a unique $\lambda_0(t) \in R$ such that*

$$1 - \beta(t) \int_{r_1}^{r_2} k(r)h(r)e^{-\lambda_0(t)r - \int_0^r \mu(\varrho)d\varrho} dr = 0.$$

Furthermore, if $\lambda \in \mathbb{C}$ *such that*

$$1-\beta(t)\int_{r_1}^{r_2} k(r)h(r)e^{-\int_0^r \mu(\varrho)d\varrho}\cdot e^{-\lambda r}dr=0$$

and $Im(\lambda)\neq 0$ *then* $Re(\lambda)<\lambda_0(t)$.

PROOF. The proof is immediate from [10].

The complementary set $\varrho(A(t))=\mathbb{C}\setminus\sigma(A(t))$ denotes the resolvent set of $A(t)$.

From [10], we also have

LEMMA 2.2. *If* $\lambda\in\varrho(A(t))$, *then for any* $\Psi(\cdot)\in\Xi$, *we have*

$$R(\lambda,A(t))\Psi(r)=$$

$$=\frac{1}{F(\lambda)}\cdot e^{-\lambda r-\int_0^r\mu(\varrho)d\varrho}\cdot\beta(t)\int_{r_1}^{r_2}k(r)h(r)\int_0^r\Psi(s)e^{-\lambda(r-s)-\int_s^r\mu(\varrho)d\varrho}\cdot dsdr+$$

$$+\int_0^r\Psi(s)e^{-\lambda(r-s)-\int_s^r\mu(\varrho)d\varrho}\cdot ds$$

where

$$F(\lambda)=1-\beta(t)\int_{r_1}^{r_2}k(r)h(r)e^{-\int_0^r\mu(\varrho)d\varrho}\cdot e^{-\lambda r}dr.$$

From [4], we know that for any $t\geq 0$ $A(t)$ is a generator of a strongly continuous semigroup of linear operators on Ξ.

In fact, the function $\beta(t)$ of fertility rate is bounded. In this paper, we assume that

$$\beta_1\leq\beta(t)\leq\beta_2,\quad \text{for any } t\in[0,\infty) \tag{2}$$

here, β_1 and β_2 are constants.

Let A_i be population evolution operator associated to β_i, i.e.

$$\begin{cases} D(A_i)=\left\{\varphi(\cdot)\in\Xi\;\middle|\;A_i\varphi\in\Xi,\quad \varphi(0)=\beta_i\int_{r_1}^{r_2}k(r)h(r)\varphi(r)dr\right\}\\ (A_i\varphi)(\cdot):=-\varphi'(\cdot)-\mu(\cdot)\varphi(\cdot)\end{cases}$$

here $i=1,2$.

From Lemma 2.1, we have

LEMMA 2.3. *If* $\lambda_0(t) \in \sigma(A(t))$ *such that* $Im(\lambda_0(t)) = 0$, *and* $\lambda_i \in \sigma(A_i)$ *such that* $Im(\lambda_i) = 0$, $i = 1, 2$, *then* $\lambda_1 \le \lambda_0(t) \le \lambda_2$.

Let $\{S_t(q)\}_{q \ge 0}$ denote the semigroup of linear operators generated by $A(t)$, and $\{S_{\beta_i}(q)\}_{q \ge 0}$ denote the semigroup of linear operators generated by A_i.

It is well-known [9] that there are constants ω^*, ω_t, M_{β_2} and M_t such that

$$(3) \qquad \|S_{\beta_2}(q)\| \le M_{\beta_2} e^{\omega^* q} \quad (q \ge 0). \qquad \|S_t(q)\| \le M_t e^{\omega_t q} \quad (q \ge 0).$$

We would like to ask if there exist constants M and ω which are independent of time t such that the following inequality is satisfied:

$$\|S_t(q)\| \le M e^{\omega q}, \quad \text{for any } q, t \ge 0.$$

In order to answer the question, we start with the following theorem (see [8]):

THEOREM 2.4. *Let A be a closed linear operator in Ξ, then the following are equivalent*

(i) *A is the generator of an exponentially bounded C-semigroup $\{S(t)\}_{t \ge 0}$ with $\|S(t)\| \le M e^{at}$ for $t \ge 0$;*

(ii) *A satisfies the conditions*

(a) *for some $\lambda \in R$, $\lambda - A$ is injective and $D((\lambda - A)^{-1}) \supset R(C)$;*

(b) $C^{-1} A C = A$,

(c) *for every initial value $\chi \in (\lambda - A)^{-1} C(\Xi)$ the following equation*

$$(4) \qquad \begin{cases} \frac{du(t)}{dt} = Au(t) \\ u(0) = \chi \end{cases}$$

has a unique solution $u(t; \chi)$ such that $\|u(t; \chi)\|$ and $\left\|\frac{d}{dt} u(t; \chi)\right\|$ are $O(e^{at})$ as $t \to \infty$.

From [1, 2] we know that for any $t \in [0, \infty)$ and $\chi \in D(A(t))$ the initial value problem

$$(5) \qquad \begin{cases} \frac{du(s)}{ds} = A(t)u(s) \\ u(0) = \chi \end{cases}$$

has a unique solution $u_t(s)$ such that

$$u_t(s)(r) = (S_t(s)\chi)(r) =$$

$$(6) \qquad = C_t(\chi)e^{\lambda_0(t)s - \int_0^r [\lambda_0(t)-\mu(\varrho)]d\varrho} + o\left(e^{(\lambda_0(t)-\varepsilon)s}\right), \quad \text{as } s \to \infty.$$

Let ω satisfy $\omega > \lambda_2$ from (6) we have

$$\|u_t(s)\| \le \left\|C_t(\chi)e^{-\int_0^r [\lambda_0(t)-\mu(e)]d\varrho}\right\| \cdot e^{\omega s} + \|o(e^{\lambda_0(t)s-\varepsilon})\|, \quad \text{as } s \to \infty.$$

Hence, we deduce that $\|u_t(s)\|$ is $o(e^{\omega s})$, as $s \to \infty$.

On the other hand, since $u_t(s) = S_t(s)\chi$ we have

$$\frac{du_t(s)}{ds} = S_t(s)A(t)\chi.$$

We deduce that

$$\left\|\left(\frac{d}{ds}\right)u_t(s)\right\| \quad \text{is } O(e^{\omega s}), \text{ as } s \to \infty.$$

From Theorem 2.4, it follows that $\|S_t(q)\| \le M_t e^{\omega q}$, $(q \ge 0)$ for any $t \ge 0$.

Moreover, from [1] and [6] the semigroup $\{S_t(q)\}_{q\ge 0}$ has the following asymptotic expansion

$$(S_t(q)\Phi)(r) = C_\Phi e^{-\lambda_0(t)r - \int_0^r \mu(\varrho)d\varrho} \cdot e^{\lambda_0(t)s} + o(e^{(\lambda_0(t)-\varepsilon)s}), \quad \text{as } s \to \infty.$$

Actually, C_Φ is independent on t. It follows that there exists a constant M such that

$$(7) \qquad \|S_t(q)\| \le Me^{\omega q} \qquad (q \ge 0)$$

for any $t \ge 0$.

REMARK. The inequality can also be proved by the definition of the semigroups generated by the population operators.

The rest of this section is devoted to show that the family of operators $\{A(t)\}_{t\ge 0}$ be a stable family of infinitesimal generators in Ξ.

First, we have the following result:

THEOREM 2.5. *Assume that* (2) *holds, then for any finite sequence*

$$0 \le t_1 \le t_2 \le \cdots \le t_k < \infty, \qquad k = 1, 2, \ldots$$

following inequality holds

$$\left\|\prod_{j=1}^{k} S_{t_j}(q_j)\right\| \le M_{\beta_2}\exp\left\{\omega^*\sum_{j=1}^{k} q_j\right\} \tag{8}$$

for $q_j \ge 0$ $(j = 1, 2, \ldots, k)$, *in which* M_{β_2} *and* ω^* *satisfy* (3).

PROOF. From the definition [10] of the semigroup for any $p(\cdot) \in \Xi$ we have $|S_{t_1}(q_1)p(r)| \le S_{\beta_2}(q_1)|p(r)|$ it follows that $\|S_{t_1}(q_1)\| \le \|S_{\beta_2}(q_1)\|$.

For any t_j, $\{S_{t_j}(q)\}_{q\ge 0}$ is positive semigroup [12] associated with the positive cone

$$\Xi^+ = \{p(\cdot) \mid p(\cdot) \in \Xi; p(r) \ge 0, \quad \text{a.e. on} \quad [0, r_m]\}$$

therefore

$$|S_{t_2}(q_2)S_{t_1}(q_1)p(r)| \le S_{t_2}(q_2)S_{\beta_2}(q_1)|p(r)|$$

again, from the definition of the population semigroup of operators, we have

$$|S_{t_2}(q_2)S_{t_1}(q_1)p(r)| \le S_{\beta_2}(q_2)S_{\beta_2}(q_1)|p(r)|$$

hence $\|S_{t_2}(q_2)S_{t_1}(q_1)\| \le \|S_{\beta_2}(q_2)S_{\beta_2}(q_1)\|$. If

$$|S_{t_{k-1}}(q_{k-1})S_{t_{k-2}}(q_{k-2})\cdots S_{t_2}(q_2)S_{t_1}(q_1)p(r)| \le$$
$$\le S_{t_{k-1}}(q_{k-1})S_{\beta_2}(q_{k-2})\cdots S_{\beta_2}(q_2)S_{\beta_2}(q_1)|p(r)|$$

then it is clear that

$$\left|\prod_{j=1}^{k-1} S_{t_j}(q_j)p(r)\right| \le S_{\beta_2}\Big(\sum_{j=1}^{k-1} q_j\Big)|p(r)|.$$

It follows that

$$S_{t_k}(q_k)\left|\prod_{j=1}^{k-1} S_{t_j}(q_j)p(\cdot)\right| \le S_{\beta_2}(q_k)S_{\beta_2}\Big(\sum_{j=1}^{k-1} q_j\Big)|p(\cdot)|$$

i.e.

$$\left|\prod_{j=1}^{k} S_{t_j}(q_j)p(r)\right| \le S_{\beta_2}\Big(\sum_{j=1}^{k} q_j\Big)|p(r)|$$

so

$$\left\|\prod_{j=1}^{k} S_{t_j}(q_j)\right\| \le M_{\beta_2}e^{\omega^*\left(\sum_{j=1}^{k} q_j\right)}, \tag{8}$$

the proof has been completed.

DEFINITION 6 [9]. Let Ξ be a Banach space. A family $\{A(t)\}_{t\geq 0}$ of infinitesimal generators of strongly continuous semigroups on Ξ is called stable if there are constants $M \geq 1$ and ω such that

$$\varrho(A(t)) \supset (\omega, \infty), \tag{9}$$

and

$$\left\| \prod_{j=1}^{k} R(\lambda, A(t_j)) \right\| \leq M(\lambda - \omega)^{-k} \quad \text{for } \lambda > \omega, \tag{10}$$

and every finite sequence $0 \leq t_1 \leq t_2, \cdots t_k \leq T$, $k = 1, 2, \cdots$.

Note that under assumption (2), the population operators $\{A(t)\}_{t\geq 0}$ satisfy (9). Actually, the family of generators $\{A(t)\}_{t\geq 0}$ of population semigroup is stable, it follows that Theorems 2.5 and the following result:

THEOREM 2.7 [9]. *For $t \in [0, \infty)$, let $A(t)$ be the infinitesimal generator of a strongly continuous semigroup $S_t(q)$ on a Banach space Ξ. The family of generators $\{A(t)\}_{t\geq 0}$ is stable if and only if there are constants $M \geq 1$ and ω such that $\varrho(A(t)) \supset (\omega, \infty)$ for $t \in [0, \infty)$ and either one of the following conditions is satisfied*

$$\left\| \prod_{j=1}^{k} S_{t_j}(q_j) \right\| \leq M \exp\left\{ \omega \sum_{j=1}^{k} q_j \right\} \quad \text{for } q_j \geq 0$$

and any finite sequence $0 \leq t_1 \leq t_2 \leq \cdots \leq t_k < \infty$, $k = 1, 2, \ldots$ or

$$\left\| \prod_{j=1}^{k} R(\lambda_j; A(t_j)) \right\| \leq M \prod_{j=1}^{k} (\lambda_j - \omega)^{-1} \quad \text{for } \lambda_j > \omega$$

and any finite sequence $0 \leq t_1 \leq t_2 \leq \cdots \leq t_k < \infty$, $k = 1, 2, \ldots$.

Summarizing the above, we have

THEOREM 2.8. *If the function $\beta(t)$ of fertility rate is bounded, then the family $\{A(t)\}_{t\geq 0}$ of infinitesimal generators of population semigroup is stable.*

Now, we discuss the solution of equation (1).

3. The Approximating Solution

In this section, we adopt the method which was used in [3] and based on an interpolation of the equation as a functional equation in the space $L^p(t_0, T; \Xi)$ and on the use of the Yosida approximations of the time derivative considered as an operator in that space.

We will use the Banach spaces $L^p(0,\infty;\Xi)$ and $W^{1,p}(0,\infty;\Xi)$. $L^p(0,\infty;\Xi)$ is the space of all Bochner measurable functions $U:[0,\infty)\to\Xi$ such that $\|U(\cdot)\|\in L^p(0,\infty)$, $1\le p<\infty$. The $W^{1,p}(0,\infty;\Xi)$ is the set of all functions $U:[0,\infty)\to\Xi$ such that there exists $v\in L^p(0,\infty;\Xi)$ such that

$$U(t)=U(0)+\int_0^t v(s)ds,\quad t\in[0,\infty).$$

We want to write the problem (1) as a functional equation in the space $L^p(0,\infty;\Xi)$ by considering the time derivative as an operator in $L^p(0,\infty;\Xi)$ with a domain which takes into account the initial condition $U(0)=U_0$ and which is also dense in $L^p(0,\infty;\Xi)$.

Define

$$\begin{aligned} D(B) &: = \{U\in W^{1,p}(0,\infty;\Xi) \mid U(o)=0\} \\ B &: D(B)\subset L^p(0,\infty;\Xi)\to L^p(0,\infty;\Xi) \\ BU &: = -U' \end{aligned}$$

and

$$D(A):=\{U\in L^p(0,\infty;\Xi) \mid U(t)\in D(A(t)),\quad \text{a.e.}\quad A(t)U(t)\in L^p(0,\infty;\Xi)\}$$
$$(AU)(t)=A(t)U(t),\qquad t\in[0,\infty),\quad \text{a.e.}$$

It is easy to check that B is the generator of a strongly continuous semigroup of contractions in $L^p(0,\infty;\Xi)$, i.e. $\|\lambda R(\lambda,B)\|_{\mathcal{L}(L^p(0,\infty;\Xi))}\le 1$ for $\lambda>0$ and $\varrho(B)=\mathbb{C}$ moreover for each $\lambda\in\mathbb{C}$ and $U\in L^p(0,\infty;\Xi)$

$$(R(\lambda,B)U)(t)=\int_0^t e^{-\lambda(t-s)}U(s)ds,\quad t\in[0,\infty).$$

The Yosida approximation of B is $B_n=nBR(n,B)=n^2R(n,B)-nI$. It is well-known that for any $U\in D(B)$

$$\lim_{n\to\infty}\|B_nU-BU\|_{L^p(0,\infty;\Xi)}=0$$

the B_nU has the following expression

$$(B_nU)(t) = n^2 \int_0^t e^{-n(t-s)}U(s)ds - nU(t).$$

For each $U \in D(B)$, the following equalities hold

$$B_nU = -nR(n,B)U' = -n(R(n;B)U)', \quad (B_nU)' = B_nU'.$$

Let $v(t) = R(\lambda, A(t))U(t)$. Until now, we do not know if $v \in L^p(0,\infty;\Xi)$, but, we have the following lemma:

LEMMA 3.1. *If $\beta \in C(0,\infty)$; then $R(\lambda, A(t)) \in C(0,\infty;\Xi)$ moreover, if $\beta \in C^k(0,\infty)$, then $R(\lambda, A(t)) \in C^k(0,\infty;\Xi)$ $k = 1,2,\ldots$*

PROOF. From [10], we have the following expression of $R(\lambda, A(t))$,

$$R(\lambda; A(t))\Psi(r) = \frac{\beta(t)}{F(\lambda)} \int_{r_1}^{r_2} k(r)h(r) \int_0^r \Psi(s)e^{-\lambda(r-s)-\int_0^r \mu(\varrho)d\varrho} \cdot dsdr \cdot$$

$$\cdot e^{-\lambda r - \int_0^r \mu(\varrho)d\varrho} + \int_0^r \Psi(s)e^{-\lambda(r-s)-\int_s^r \mu(\varrho)d\varrho} ds,$$

Hence, for any $t_1, t_2 > 0$, $\Psi \in \Xi$.

$$[R(\lambda; A(t_1)) - R(\lambda; A(t_2))]\Psi(r) = (\beta(t_1) - \beta(t_2)) \cdot \frac{1}{F(\lambda)} \cdot$$

$$\cdot e^{-\lambda r - \int_0^r \mu(\varrho)d\varrho} \cdot \int_{r_1}^{r_2} k(r)h(r) \int_0^r \Psi(s)e^{-\lambda(r-s)-\int_s^r \mu(\varrho)d\varrho} dsdr$$

so

$$\|R(\lambda; A(t_1)) - R(\lambda; A(t_2))\| \leq M \cdot \frac{|\beta(t_1) - \beta(t_2)|}{|F(\lambda)|} \tag{11}$$

in which M is a constant. From (11) the lemma follows.

When $\beta \in C(0,\infty)$, from Lemma 1, if $\nu(t) = R(\lambda, A(t))u(t)$, in which $u \in L^p(0,\infty;\Xi)$, then $\nu \in L^p(0,\infty;\Xi)$ and

$$(R(\lambda, A)u)(t) = R(\lambda, A(t))u(t), \quad t \in [0,\infty), \quad \text{a.e.}$$

Now, equation (1) can be written as an equation in $L^p(0,\infty;\Xi)$:

$$B(p(\cdot,\cdot) - p_0(\cdot)) + Ap(\cdot,\cdot) = 0, \qquad p(\cdot,\cdot) \in D(B). \tag{12}$$

Now we want to solve a problem which approximates (12).

$$B_n(v_n - u_0) + \Lambda v_n = 0. \tag{13}$$

LEMMA 3.2. *Let $n > \omega$, where ω is stability constant, defining a bounded linear operator being dependent on t: $H(t) := n^2 R(n; A(t))$ then*

$$\left\| \prod_{j=1}^{k} e^{H(t_j)S_j} \right\| \le \exp\left\{ \frac{n^2 M}{n-\omega} \Big(\sum_{j=1}^{k} S_j \Big) \right\}, \quad \text{for } s_j \ge 0$$

and any finite sequence $0 \le t_1 \le t_2 \le \cdots \le t_k < \infty$, in which $\{e^{H(t)s}\}_{s \ge 0}$ is a strongly continuous semigroup generated by the operator $H(t)$.

Moreover, if $\beta \in C(0,\infty)$, then $H \in (0,\infty; \mathcal{L}(\Xi))$.

PROOF. It is well-known that

$$e^{H(t)s} = \sum_{k=0}^{\infty} \frac{(H(t)s)^k}{k}$$

because $H(t)$ is bounded linear operator.

From Theorem 2.7

$$\|(H(t))^k\| \le n^{2k} \cdot \frac{M}{(n-\omega)^k}$$

hence

$$\left\| \prod_{j=1}^{k} e^{H(t_j)s_j} \right\| \le \prod_{j=1}^{k} \|e^{H(t_j)s_j}\| \le \prod_{j=1}^{k} e^{\|H(t_j)\| s_j} \le e^{\frac{n^2 M}{n-\omega} \left(\sum_{j=1}^{k} s_j \right)}.$$

The second part of the lemma can be checked by Lemma 1.

In order to solve the approximating problem we need the following result:

THEOREM 3.3. *Assume that β is continuous uniformly, then there exists a unique evolution system $U(t,s)$, $0 \le s \le t < \infty$, in Ξ satisfying*

(E_1). $\|U(t,s)\| \le e^{\frac{Mn^2}{n-\omega}(t-s)}$

for $0 \le s \le t < \infty$.

(E_2). $\frac{\partial^t}{\partial t} U(t,s)v \Big|_{t=s} = H(s)v$

for $v \in \Xi$, $0 \le s < \infty$.

(E_3). $\frac{\partial}{\partial s} U(t,s)v = -U(t,s)H(s)v$

for $v \in \Xi$, $0 \le s \le t < \infty$, where the derivative from the right in (E_2) and the derivative in (E_3) are in strong sense in Ξ.

PROOF. We use the method introduced by Pazy [9] by approximating the family $\{H(t)\}_{t \ge 0}$ by piecewise constant families $\{H_n(t)\}_{t \in [0,\infty)}$, $n = 1, 2, \ldots$, defined as follows:

Let $\delta_n = \frac{1}{n}$ and let $t_k^n = k\delta_n$, $k = 0, 1, 2, \ldots$ and let

$$H_n(t) = H(t_k^n), \quad t_k^n \le t < t_{k+1}^n, \quad k = 0, 1, 2, \ldots.$$

Since $\beta(t)$ is continuous uniformly in $\mathscr{L}(\Xi)$, it follows that

$$\|H(t) - H_n(t)\|_{\mathscr{L}(\Xi)} \to 0, \quad (n \to \infty)$$

uniformly for $t \in [0, \infty)$.

Now, we define a two-parameter family of operator $U_n(t,s)$, $0 \le s \le t < \infty$ by

$$U_n(t,s) = \begin{cases} \exp(H(t_j^n)(t-s)), \text{ for } t_j^n \le s \le t \le t_{j+1}^n, \\ \exp(H(t_k^n)(t - t_k^n)) \left[\prod_{j=\ell+1}^{k-1} \exp(H(t_j^n)\left(\frac{1}{n}\right)\right] \cdot \\ \cdot \exp(H(t_\ell^n)(t_{\ell+1}^n - s)), \quad \text{for } k > \ell,\ t_k^n \le t \le t_{k+\ell}^n, \\ \quad t_\ell^n \le s \le t_{\ell+1}^n. \end{cases}$$

It is easy to verify that $U_n(t,s)$ is satisfied with the following conditions:

$$U_n(s,s) = I, \quad U_n(t,s) = U_n(t,r)U_n(r,s)$$

for $0 \le s \le r \le t \le T$ and $(t,s) \to U_n(t,s)$ is uniformly continuous on $0 \le s \le t \le \infty$

$$\|U_n(t,s)\| \le e^{\frac{M\lambda^2}{\lambda - \omega}(t-s)}$$

for $0 \le s \le t < \infty$ in which λ is the same with n in Lemma 2.

The definition of $U_n(t,s)$ implies that for $v \in \Xi$

$$\frac{\partial}{\partial t} U_n(t,s)v = H_n(t)U_n(t,s)v,$$

for $t \ne t_j^n$, $j = 0, 1, \ldots$,

$$\frac{\partial}{\partial s} U_n(t,s)v = -U_n(t,s)H_n(s)v$$

for $s \ne t_j^n$, $j = 0, 1, \ldots$. It follows that except for a finite number of values of r the map $r \to U_n(t,r)U_m(r,s)v$ is differentiable in r, $s \le r \le t$ and

$$U_n(t,s)v - U_m(t,s)v = -\int_s^t \frac{\partial}{\partial r} U_n(t,r)U_m(r,s)v\,dr =$$

$$= \int_s^t U_n(t,r)(H_n(r) - H_m(r))U_m(r,s)v\,dr$$

so

$$\|U_n(t,s)v - U_m(t,s)v\| \le \exp\left\{\frac{M\lambda^2}{\lambda-\omega}(t-s)\right\} \cdot \int_s^t \|H_n(r) - H_m(r)\|dr.$$

Let

$$U(t,s)\chi = \lim_{n\to\infty} U_n(t,s)\chi$$

for $\chi \in \Xi$, $0 \le s \le t < \infty$, we have $U(t,s)$ is a two-parameter bounded operator

$$\|U(t,s)\| \le e^{\frac{M\lambda^2}{\lambda-\omega}(t-s)}$$

for $0 \le s \le t < \infty$.

Other conclusions can be proved by the way used in [9].

Let $f \in C(0,\infty;\Xi)$, and consider the initial value problem

$$\left\{\begin{array}{l} \frac{du(t)}{dt} = H(t)u(t) + f(t) \\ u(0) = v. \end{array}\right. \tag{14}$$

We have the following result:

THEOREM 3.4. *If $f \in C[0,\infty;\Xi)$, $v \in \Xi$, then the initial value problem* (14) *has a unique solution u, i.e. $u \in C^1(0,\infty;\Xi)$ and* (14) *is satisfied in Ξ, and this solution is unique and moreover*

$$u(t) = U(t,s)v + \int_s^t U(t,r)f(r)dr \tag{15}$$

where $U(t,s)$ is given by Theorem 3.3.

PROOF. Because

$$U(t,s)\chi = \lim_{n\to\infty} U_n(t,s)\chi$$

for $\chi \in \Xi$, $0 \le s \le t < \infty$, in which $U_n(t,s)$ was given in Theorem 3.3.

Since $(t,s) \to U_n(t,s)$ is uniformly continuous in $\mathscr{L}(\Xi)$, for each $v \in \Xi$, $U(t,s)v$ must be continuous in Ξ, for $0 < s \le t < \infty$. From [9], we have completed the proof.

Now we can solve the problem which approximates problem (12).

THEOREM 3.5. *Given $u_0 \in \Xi$, there exists for each $n > \omega$ a unique $u_n \in D(\Lambda)$ being a solution of*

$$B_n(u_n - u_0) + \Lambda u_n = 0, \tag{16}$$

and the following estimate holds for $0 \leq t < \infty$, a.e.

$$\|u_n(t)\| \leq \frac{nM}{n-\omega} e^{\frac{n^2 M}{n-\omega}t} \cdot \|u_0\|$$

and $u_n \in \{v \in C(0,\infty;\Xi) \mid v(t) \in D(A(t)),\ 0 \leq t < \infty\}$.

PROOF. Assume that $u_n \in D(\Lambda)$ is a solution of the problem (16). Since

$$(B_n u)(t) = n^2 \int_0^t e^{-n(t-s)} u(s)ds - nu(t),\ \ 0 \leq t < \infty$$

for $u \in L^p(0,\infty;\Xi)$, we have

$$n^2 \int_0^t e^{-n(t-s)} (u_n(s) - u_0)ds - n(u_n(t) - u_0) + A(t)u_n(t) = 0,$$

for $t \in [0,\infty)$ a.e., i.e.

(17)

$$n^2 \int_0^t e^{-n(t-s)} u_n(s)ds - n^2 \int_0^t e^{-n(t-s)} u_0 ds - (n - A(t))u_n(t) + nu_0 = 0.$$

By applying $R(n, A(t))$ to both sides of (17), we get

$$n^2 e^{-nt} R(n, A(t)) \int_0^t e^{ns} u_n(s)ds-$$

$$-n^2 e^{-nt} R(n, A(t)) \int_0^t e^{ns} ds u_0 - u_n(t) + R(n, A(t))nu_0 = 0$$

i.e.

$$n^2 e^{-nt} R(n; A(t)) \int_0^t e^{ns} u_n(s)ds+$$

$$+e^{-nt} nR(n; A(t))u_0 - nR(n; A(t))u_0 - u_n(t) + R(n; A(t))nu_0 = 0.$$

Eventually, we have

$$n^2R(n;A(t))\int_0^t e^{ns}u_n(s)ds+nR(n;A(t))u_0-u_n(t)e^{nt}=0. \tag{18}$$

Let

$$H(t)=n^2R(n;A(t)),\quad \omega_n(t)=\int_0^t e^{ns}u_n(s)ds,\quad f(t)=nR(n;A(t))u_0,\quad v=0.$$

From (18), we have

$$\begin{cases}\omega_n'(t)=H(t)\omega_n(t)+f(t)\\ \omega_n(0)=0,\end{cases} \tag{19}$$

hence, $\omega_n(t)$ is the solution of equation (19). From Theorem 3.4 it follows that

$$\omega_n(t)=\int_0^t U(t,s)nR(n;A(s))u_0ds$$

where $U(t,s)$ is the evolution system given in Theorem 3.3.

From (18)

$$\begin{aligned}u_n(t)=&n^2R(n;A(t))e^{-nt}\omega_n(t)+nR(n;A(t))e^{-nt}u_0\\ u_n(t)=&R(n;A(t))[n^2e^{-nt}\omega_n(t)+ne^{-nt}u_0]\end{aligned}$$

hence,

$$\begin{aligned}\|u_n(t)\|&\le\frac{nM}{n-\omega}e^{-nt}\|u_0\|+\frac{n^2M}{n-\omega}e^{-nt}\|\omega_n(t)\|\le\\ &\le\frac{nM}{n-\omega}e^{-nt}\|u_0\|+\frac{n^3M}{n-\omega}e^{-nt}\|u_0\|\int_0^t\|U(t,s)\|\cdot\|R(n;A(s))\|ds\le\\ &\le\frac{nM}{n-\omega}e^{-nt}\|u_0\|+\frac{n^3M}{n-\omega}e^{-nt}\|u_0\|\int_0^t e^{\frac{n^2M}{n-\omega}(t-s)}\cdot\frac{M}{n-\omega}ds\le\\ &\le\frac{nM}{n-\omega}e^{-nt}\|u_0\|+\frac{nM}{n-\omega}\|u_0\|e^{-nt}\left(e^{\frac{n^2M}{n-\omega}t}-1\right)\le\\ &\le\frac{nM}{n-\omega}e^{\frac{n^2M}{n-\omega}t}\|u_0\|.\end{aligned}$$

Conversely, given $u_0 \in \Xi$, from Theorem 3.4 there exists a solution $\omega_n \in$ $\in C^1(0,\infty;\Xi)$ of equation (14), with

$$H(t) = n^2 R(n; A(t)), \qquad v = 0, \qquad f(t) = R(n; A(t)) n u_0$$

therefore, $u_n(t) = e^{-nt}\omega_n'(t)$ is a solution of equation (16). The proof is completed.

References

[1] ACQUISTAPACE, P. and TERRENI, B., A Unified Approach to Abstract Linear Non-autonomous Parabolic Equations, *Differential Integral Equations*, **1** (1988), 433–457.

[2] BUTTU, ANNA, On the Evolution Operator for a Class of Non-autonomous Abstract Parabolic Equations, *Journal of Mathematical Analysis and Applications*, **170** (1992), 115–137.

[3] DA PRATO, G. and SINESTRARI, E., Differential Operators with Nondense Domain, *Ann. Sci. Norm. Sup. Pisa*, **14** (1987), 285–300.

[4] GUO, B.-Z. and ZHU, G.-T., The Properties of C_0-semigroup $\{\tau(t)\}_{t\geq 0}$ Generated by the Population Operator, *J. Sys. Sci. & Math. Sci.*, **6** (1986).

[5] GUO, B.-Z. and CHAN, W. L., A Semigroup Approach to Age-dependent Population Dynamics with Time Delay, Comm. *Partial Differential Equations* **14** (1989), 809–832.

[6] JOÓ, I. and ZHANG SHENGHAI, A Class of Nonlinear Population Evolution Equation in Banach Spaces, Mathematical Institute of the Hungarian Academy of Sciences, Preprint, No. 39/1992.

[7] KATO, T., Linear Evolution Equations of "Hyperbolic Type" II, *J. Math. Soc. Japan*, **25** (1973), 648–666.

[8] NAOKI TANAKA and ISAO MIYADERA, C-Semigroups and the Abstract Cauchy Problem, *Journal of Mathematical Analysis and Applications*, **170** (1992), 196–206.

[9] PAZY, A., *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, New York, 1983.

[10] SONG, J. and YU, J.-Y., *Population System Control*, Springer-Verlag, New York, 1987.

[11] SONG, J. and YU, J.-Y., *Control Theory of Human Population*, Science Press, Beijing, 1985.

[12] ZHANG, SHENGHAI and ZHU, G.-T., The Solution and Asymptoticity of Nonlinear Population Evolution Equation, *Acta Mathematicae Applicatae Sinica*, **14**, (1991).

[13] WEBB, G. F., *Theory of Nonlinear Age-dependent Population Dynamics*, Marcel Dekker, New York, 1985.

ANNALES UNIV. SCI. BUDAPEST., **36** (1993), 269–274

ON ADDITIVE FUNCTIONS

By

I. JOÓ

Department of Analysis, Eötvös Loránd University, Budapest

(Received April 13, 1993)

Dedicated to Professor Paul Erdős for his 80^{th} *birthday*

In this paper we give a characterization theorem. The first such characterization is apparently that of P. ERDŐS [1]. He proved that if a real-valued additive function is non-decreasing, or satisfies $f(n+1)-$ $-f(n)\to 0$ as $n\to\infty$, then it must have the form $A\cdot\log n$ for some constant A. He had a separate argument for each case.

DEFINITION. We say, that $f:\mathbb{N}\to\mathbb{R}$ is completely additive mod 1, if for every $n,m\in\mathbb{N}$ holds $\{f(n\cdot m)\}=\{f(n)+f(m)\}$, where $\{x\}$ is the fractional part of x.

THEOREM. *If f is completely additive* mod *1, $a,b\in\mathbb{N}$, $(a,b)=1$,*

$$(*)\qquad \{f(a(n+1)+b)-f(an+b)\}\to 0,$$

as $n\to\infty$, then $\{r(B)\cdot f(an+B)\}=\{r(B)\cdot c_0\cdot\log(an+B)\}$, where $(a,B)=$ $=1$, c_0 is a constant independent of n, $r(B)$ is the smallest positive integer for wich $B^{r(B)}\equiv 1$ (mod a).

We remark that if we change $(*)$ to the condition $\{f(n+1)\}-\{f(n)\}\to 0$ then corresponding theorem is an open problem. Futhermore it is possible to give the monotone version of the theorem above. We shall give it in a subsequent paper.

PROOF. From the assumption of Theorem we have

$$(1)\qquad f(a(n+1)+b)-f(an+b)=k(n+1)+r(n+1),$$

where $k(n+1)\in\mathbb{Z}$, $0\le r(n+1)<1$ and $r(n+1)\to 0$ if $n\to\infty$. Using (1) with $n:=n+1,\ n+2,\dots,n+b$ and summarizing them we obtain

$$f(an+ab+b)-f(an+b)=k(n+1)+\ldots+k(n+b)+r(n+1)+\ldots+r(n+b).$$

Hence

$$\lim_{n\to\infty}\{f(an+ab+b)-f(an+b)\}=0.$$

In this relation writing $b\cdot n$ in place of n we get

$$\lim_{n\to\infty}\{f(a(n+1)b+b)-f(anb+b)\}=0.$$

Here we have

$$\{f(a(n+1)b+b)-f(anb+b)\}=\{\{f(a(n+1)b+b)-f(anb+b)\}\}=$$
$$=\{\{f(b)+f(a(n+1)+1)\}-\{f(b)+f(an+1)\}\}=$$
$$=\{f(a(n+1)+1)-f(an+1)\}.$$

Hence

(2) $$\lim_{n\to\infty}\{f(a(n+1)+1)-f(an+1)\}=0.$$

Introduce F in the following way:
$F(a+1):=\{f(a+1)\}$, $\{F(an+1)\}:=\{f(an+1)\}$, and choose the values of F such away, that $F(a(n+1)+1)\in\left[F(an+1)-\frac{1}{2},F(an+1)+\frac{1}{2}\right)$. Then F is uniquely determined on the set $(an+1:\ n=1,2,\dots)$. Denote
(3)
$$q(n_1,n_2):=F((an_1+1)(an_2+1))-F(an_1+1)-F(an_2+1),\ n_1,n_2\in\mathbb{N}.$$
Since f is completely additive mod 1, therefore $\{q(n_1,n_2)\}=0$, i.e. $q(n_1,n_2)\in\mathbb{Z}$. Obviously

(4) $$q(n_1,n_2+1)-q(n_1,n_2)=F((an_1+1)(an_2+a+1))-$$
$$-F((an_1+1)(an_2+1))-(F(an_2+a+1)-F(an_2+1)).$$

Since

$$\{F(a(n+1)+1)-F(an+1)\}=\{\{F(a(n+1)+1)\}-\{F(an+1)\}\}=$$
$$=\{\{f(a(n+1)+1)\}-\{f(an+1)\}\}=\{f(a(n+1)+1)-f(an+1)\},$$

therefore

$$\lim_{n\to\infty}\{F(a(n+1)+1)-F(an+1)\}=0.$$

Since

$$|F(a(n+1)+1)-F(an+1)|\le\frac{1}{2},\ \text{therefore}$$

(5) $$\lim_{n\to\infty}(F(a(n+1)+1)-F(an+1))=0$$

and

(6) $$F(a(n+1)+1)\ge F(an+1),\quad \text{if } n>n_0.$$

Obviously

$$F((an_1+1)(an_2+a+1))-F((an_1+1)(an_2+1))=$$
$$=F((an_1+1)(an_2+1)+a(an_1+1))-F((an_1+1)(an_2+1))$$

Here $(an_1+1)(an_2+1)=a(an_1n_2+n_1+n_2)+1$. Since n_1 is an arbitrary, but fixed number, therefore

(7) $$\lim_{n_2\to\infty}(F((an_1+1)(an_2+a+1))-F((an_1+1)(an_2+1)))=0.$$

From (5),(7) we obtain

(8) $$\lim_{n_2\to\infty}(q(n_1,n_2+1)-q(n_1,n_2))=0.$$

But $q(n_1,n_2+1)-q(n_1,n_2)$ is an integer, therefore

(9) $$q(n_1,n_2)=q(n_1,n_2+1)=q(n_1,n_2+2)=\ldots,\text{ if } n_2>N(n_1).$$

Introduce the function G,

(10) $$G(am+1):=F(am+1)+\lim_{n\to\infty}q(m,n)=$$
$$=\lim_{n\to\infty}(F((am+1)(an+1))-F(an+1)).$$

Now we prove, that G is completely additive on the set $(an+1:n=1,2,\ldots)$. Let be $m_1,m_2\in\mathbb{N}$.

(11) $$G((am_1+1)(am_2+1))=$$
$$=\lim_{n\to\infty}(F((am_1+1)(am_2+1)(an+1))-F(an+1))=$$
$$=\lim_{n\to\infty}(F((am_1+1)(am_2+1)(an+1))-F((am_2+1)(an+1))+$$
$$+F((am_2+1)(an+1))-F(an+1)).$$

Because of (10)

(12) $$\lim_{n\to\infty}(F((am_2+1)(an+1))-F(an+1))=G(am_2+1).$$

Further $(am_2+1)(an+1)=a(am_2n+m_2+n)+1$. Using this we obtain taking into account of (10),

(13)
$$\lim_{n\to\infty}(F((am_1+1)(am_2+1)(an+1))-F((am_2+1)(an+1)))=G(am_1+1).$$

From (11), (12), (13) we get G is completely additive on the set $(an+1:1,2,\ldots)$. Following the calculations of proof of Theorem 1 in [3], we obtain $G(an+1)=c_0\cdot\log(an+1)$, where c_0 is a constant, independent of n. Indeed assume that $\frac{G(an+1)}{\log(an+1)}$ is not constant. Then exist $a_1\neq b_1$, $1<$ $<a_1,b_1\in\mathbb{N}$, such that e.g.

(14) $$c_1:=\frac{G(aa_1+1)}{\log(an+1)}<\frac{G(ab_1+1)}{\log(ab_1+1)}.$$

Denote

(15) $$g(an+1) := G(an+1) - c_1 \cdot \log(an+1).$$

First we prove $g(an+1)$ is bounded and after this that $g(an+1)$ is unbounded. From (1) $g(aa_1+1)=0$. Since G is completely additive on the set $(an+1 : n=1,2,\ldots)$ therefore

(16) $$g[(aa_1+1)^k] = 0, \ (k=1,2,\ldots).$$

Consider now $g(ab_1+1)$.

$$g(ab_1+1) = G(ab_1+1) - c_1 \cdot \log(ab_1+1) =$$
$$= \log(ab_1+1)\left(\frac{G(ab_1+1)}{\log(ab_1+1)} - \frac{G(aa_1+1)}{\log(aa_1+1)}\right) > 0.$$

Hence $g(ab_1+1) > 0$.

Let $n \geq 1$ be arbitrary. Then there exists $l \geq 0$, $l \in \mathbb{N}$ such that

$$(aa_1+1)^l \leq an+1 < (aa_1+1)^{l+1}.$$

Thus we have

$$G(an+1) \leq G((aa_1+1)^{l+1}) \text{ and } \log(aa_1+1)^l \leq \log(an+1).$$

Therefore

$$g(an+1) \leq (l+1)\cdot G(aa_1+1) - c_1 \cdot \log(aa_1+1) =$$
$$= G(aa_1+1) + l\cdot(G(aa_1+1) - c_1 \cdot \log(aa_1+1)) = G(aa_1+1).$$

Hence $g(an+1) \leq G(aa_1+1)$, i.e. $g(an+1)$ is bounded. Now let be $n \geq 1$ integer. Then there exists $s \geq 0$ integer, such that

$$(ab_1+1)^s \leq an+1 \leq (ab_1+1)^{s+1}.$$

Thus

$$g(an+1) = G(an+1) - c_1 \cdot \log(an+1) \geq$$
$$\geq G((ab_1+1)^s) - c_1 \cdot \log(ab_1+1)^{s+1} =$$
$$= s \cdot G(ab_1+1) - c_1 \cdot (s+1) \cdot \log(ab_1+1) =$$
$$= s \cdot (G(ab_1+1) - c_1 \cdot \log(ab_1+1)) - c_1 \cdot \log(ab_1+1) =$$
$$= s \cdot g(ab_1+1) - c_1 \cdot \log(ab_1+1).$$

Here $g(ab_1+1) > 0$. If $n \to \infty$ then $s \to \infty$, therefore $s \cdot g(ab_1+1) \to \infty$, hence $g(an+1) \to \infty$ $(n \to \infty)$. Thus $g(an+1)$ is unbounded. Hence we have obtained a contradiction, thus $G(an+1) = c_0 \cdot \log(an+1)$. Hence from (10) and the definition of F we have

$$\{c_0 \cdot \log(an+1)\} = \{G(an+1)\} = \{F(an+1)\} = \{f(an+1)\}.$$

Now let B be $B \neq 1$, $(a, B) = 1$. Then we have $(an+B)^{r(B)} = am+1$, where m is a suitable positive integer. Hence

$$\{f((an+B)^{r(B)})\} = \{f(am+1)\} = \{c_0 \cdot \log(am+1)\} = \\ = \{c_0 \cdot \log(an+B)^{r(B)}\}.$$

From this

$$\{r(B) \cdot f(an+B)\} = \{r(B) \cdot c_0 \cdot \log(an+B)\},$$

which implies the theorem. In the theorem $r(B)$ cannot be eliminated e.g. take $a = 3$, $b = 2$, $f(p) := logp$ if $p \equiv 1 \pmod 3$, $f(p) := \log p + \frac{1}{2}$ if $p \equiv 2 \pmod 3$, p is a prime, f is completely additive mod 1.

REMARK. If $f : \mathbb{N} \to \ell_2$ is completely additive and $\|f\|$ (the ℓ_2 norm of f) is monotonic, then $f(n) = c \cdot \log n$, where $c \in \ell_2$ is a constant vector.

PROOF. Let a be an arbitrary, $a \geq 2$ integer. If $n \in \mathbb{N}$, $n \geq 2$ then we can write $a^k \leq n \leq a^{k+1}$ with suitable k, $k \in \mathbb{N}$. Then we have

$$k \cdot \|f(a)\| \leq \|f(n)\| < (k+1) \cdot \|f(a)\|. \tag{17}$$

Since $k \cdot \log a \leq \log n < (k+1) \cdot \log a$ therefore from (17)

$$\frac{\|f(a)\|}{\log a} - \frac{\|f(a)\|}{\log n} \leq \frac{\|f(n)\|}{\log n} \leq \frac{\|f(a)\|}{\log a} + \frac{\|f(a)\|}{\log n}. \tag{18}$$

Hence

$$\left| \frac{\|f(n)\|}{\log n} - \frac{\|f(a)\|}{\log a} \right| \leq \frac{\|f(a)\|}{\log n}.$$

Letting n (and so k) $\to \infty$ we see that

$$A = \lim_{n \to \infty} \frac{\|f(n)\|}{\log n}$$

exists and is finite, and that for every positive integer a, $\|f(a)\| = A \cdot \log a$. Whitout loss of generality we may assume that $A = 1$. Using this for $n, m, n \cdot m$, where $n, m \in \mathbb{N}$ we obtain

$$\|f(n)\|^2 = \log^2 n, \tag{19}$$

$$\|f(m)\|^2 = \log^2 m, \tag{20}$$

$$\|f(n \cdot m)\|^2 = log^2 nm. \tag{21}$$

Using (19),(20) and that f is completely additive we get from (21)

$$\langle f(m), f(n) \rangle = \log m \cdot \log n. \tag{22}$$

From this

$$\left\langle \frac{f(m)}{\log m}, \frac{f(n)}{\log n} \right\rangle = 1, \quad (m, n \geq 2).$$

It is well-known, that if H is a Hilbert space, $(x_n) \subset H$ is a bounded sequence, then there exist $(x_{n_k}) \subset H$ subsequence and $x \in H$ such that for every $y \in H$ holds $\langle y, x_{n_k} \rangle \to \langle y, x \rangle$. Using this we obtain that there exist $(n_k) \subset \mathbb{N}$ and $\tau \in \ell_2$ such that

$$\left\langle \frac{f(m)}{\log m}, \frac{f(n_k)}{\log n_k} \right\rangle \to \left\langle \frac{f(m)}{\log m}, \tau \right\rangle.$$

But the left-hand is equal to 1 for every $m, n_k \geq 2$, thus

$$\left\langle \frac{f(m)}{\log m}, \tau \right\rangle = 1.$$

Here similarly as in above, we obtain

$$\left\langle \frac{f(m_k)}{\log m_k}, \tau \right\rangle \to \langle \tau, \tau \rangle = \|\tau\|^2.$$

Hence $\|\tau\| = 1$. From (22)

$$\left\langle \frac{f(m)}{\log m}, f(n) \right\rangle = \log n$$

which implies $\langle f(n), \tau \rangle = \log n$, so

$$f(n) = \tau \cdot \log n + g(n), \qquad \text{where} \quad \langle g(n), \tau \rangle = 0. \tag{23}$$

From this $\|f(n)\|^2 = \log^2 n + \|g(n)\|^2$. But we have seen $\|f(n)\|^2 = \log^2 n$, thus $g(n) = 0$. Hence from (23) follows the statement.

References

[1] P. Erdős, On the distribution function of additive functions, *Ann. Math.*, **47** (1946), 1–20.

[2] A. Bogmér, On additive functions, *Annales Univ. Sci. Budapest.*, Sect. Math., **36** (1993), 3–4.

[3] M. Joó, On additive functions, *Acta Math. Acad. Sci. Hung.*, **60** (1992), 349–350.

INDEX

Address:
MATHEMATICAL INSTITUTE, EÖTVÖS LORÁND UNIVERSITY
Budapest, Múzeum krt. 6–8.
H - 1088
ISSN 0524–9007

Technikai szerkesztő:
Dr. SCHARNITZKY VIKTOR

A kiadásért felelős: az Eötvös Loránd Tudományegyetem rektora
A kézirat a nyomdába érkezett: 1993. szeptember. Megjelent: 1993. december
Terjedelem: 24 A/4 ív. Példányszám: 500
Készült az EMTEX szedőprogram felhasználásával
az MSZ 5601–59 és 5602–55 szabványok szerint
Az elektronikus tipografálás Juhász Lehel és Fried Katalin munkája
A sokszorosítás az ELTE sokszorosító üzemében készült
Felelős vezető: Arató Tamás
ELTE 93197