

# ANNALES

## UNIVERSITATIS SCIENTIARUM BUDAPESTINENSIS DE ROLANDO EÖTVÖS NOMINATAE

### SECTIO MATHEMATICA

TOMUS XXXVIII.

REDIGIT  
Á. CSÁSZÁR

ADIUVANTIBUS

L. BABAI, A. BENCZÚR, M. BOGNÁR, K. BÖRÖCZKY, I. CSISZÁR,  
J. DEMETROVITS, A. FRANK, J. FRITZ, E. FRIED, S. GISBERT, A. HAJNAL,  
G. HALÁSZ, A. IVÁNYI, I. KÁTAI, P. KOMJÁTH, M. LACZKOVICH,  
L. LOVÁSZ, J. MOLNÁR, A. PRÉKOPA, A. RECKSI, F. SCHIPP,  
Z. SEBESTYÉN, L. SIMON, GY. SOÓS, J. SURÁNYI, J. SZENTHE,  
G. SZÉKLEY, L. VARGA, I. VINCZE



1995

# ANNALES

## UNIVERSITATIS SCIENTIARUM BUDAPESTINENSIS DE ROLANDO EÖTVÖS NOMINATAE

SECTIO BIOLOGICA

inceptit anno MCMLVII

SECTIO CHIMICA

inceptit anno MCMLIX

SECTIO CLASSICA

inceptit anno MCMXXIV

SECTIO COMPUTATORICA

inceptit anno MCMLXXXVIII

SECTIO GEOGRAPHICA

inceptit anno MCMLXVI

SECTIO GEOLOGICA

inceptit anno MCMLVII

SECTIO GEOPHYSICA ET METEOROLOGICA

inceptit anno MCMLXXV

SECTIO HISTORICA

inceptit anno MCMLVII

SECTIO IURIDICA

inceptit anno MCMLIX

SECTIO LINGUISTICA

inceptit anno MCMLXX

SECTIO MATHEMATICA

inceptit anno MCMLVIII

SECTIO PAEDAGOGICA ET PSYCHOLOGICA

inceptit anno MCMLXX

SECTIO PHILOLOGICA

inceptit anno MCMLVII

SECTIO PHILOLOGICA HUNGARICA

inceptit anno MCMLXX

SECTIO PHILOLOGICA MODERNA

inceptit anno MCMLXX

SECTIO PHILOSOPHICA ET SOCIOLOGICA

inceptit anno MCMLXII

## ON THE LAWS OF METABELIAN POINTED-GROUPS

By

ZAFAR ALI and ABDUL MAJEED

Department of Mathematics University of the Punjab, Lahore, Pakistan

(Received February 23, 1993)

### 1. Introduction

In [1] D. E. COHEN proved that the variety of all metabelian groups is *hereditarily* finitely based. (By a hereditarily finitely based variety we mean a variety  $\mathcal{V}$  such that every subvariety of  $\mathcal{V}$  is finitely based). Cohen's proof of this theorem introduced an important technique, the use of well-quasi-ordered sets, and this technique has subsequently been used to obtain a number of further results. (We also use the same technique to make the present work possible). Here we shall examine the analogous statement for the variety of metabelian pointed-groups, where by pointed-group we mean an ordered pair  $(G, c)$  consisting of a group  $G$  together with a distinguished element  $c$  of  $G$ . (The idea of pointed-groups comes from the category of pointed-sets, see for example [8]).

Now by a *law* of a pointed-group  $(G, c)$  we shall mean a word  $w$  of the free group on the countable set  $\{y, x_1, x_2, \dots\}$  such that  $w$  always becomes equal to the identity element of  $G$  when  $c$  is substituted for  $y$  and arbitrary elements of  $G$  are substituted for  $x_1, x_2, \dots$  (for example  $[y, x_1]$  is a law of  $(G, c)$ , iff  $c$  is a central element of  $G$ ). Included among the laws of  $(G, c)$  are the laws of the group  $G$ , or more precisely, those words in the variables  $x_1, x_2, \dots$  which are laws of the group  $G$ . Thus the idea of laws of a pointed-group generalizes the idea of laws of a group and the purpose of this paper is to generalize the result proved by D. E. COHEN in [1], that every variety of metabelian groups is finitely based. Detailed information concerning varieties of groups may be found in [3].

A pointed-group may be regarded as a group with an extra nullary operation and therefore it is an algebra in the sense of universal algebra. The more general concepts relating to universal algebras and varieties of algebras

are described in [4]. It is useful to note some of these concepts taken for pointed-groups.

As indicated above, the *factor algebra* of  $(G, c)$  is  $(G/N, cN)$  where  $N$  is a normal subgroup of  $G$  not necessarily containing  $c$ . The subalgebra of  $(G, c)$  is  $(H, c)$  where  $H$  is a subgroup of  $G$  containing  $c$ . The *cartesian product* of the family  $\{G_\lambda, c_\lambda \mid \lambda \in \Lambda\}$  is  $(G, c)$  where  $G$  is the cartesian product of  $\{G_\lambda \mid \lambda \in \Lambda\}$  and  $c$  is the element of  $G$  with value  $c_\lambda$  at  $\lambda$  for all  $\lambda \in \Lambda$ . A generating set for a pointed-group  $(G, c)$  is a subset  $S$  of  $G$  such that  $S \cup \{c\}$  generates  $G$ . The endomorphism  $\alpha$  of a group  $G$  is an endomorphism of  $(G, c)$  if  $c\alpha = c$ . A subgroup  $H$  (not necessary containing  $c$ ) of  $G$  is an *admissible subgroup* of  $(G, c)$  if  $H\alpha \leq H$  for every endomorphism  $\alpha$  of  $(G, c)$ . A variety of pointed-groups is the class of all pointed-groups in which the elements of some given set of words are all laws. Equivalently, it is a class closed under the operations of taking factor algebras, subalgebras, cartesian products and algebras isomorphic to these. (For un-explained terminology concerning pointed-groups, see [6] and [7]).

## 2. Preliminaries

We call a pointed-group  $(G, c)$  *metabelian* if  $G$  is a metabelian group. Here in this paper we shall discuss varieties of pointed-groups in which every pointed-group is metabelian. Our aim is to generalize the result proved by D. E. COHEN in [1], that the variety of all metabelian groups is hereditarily finitely based, i.e. every subvariety of the variety of metabelian groups is finitely based.

We shall need some notational definitions in the free pointed-group  $(X, y)$  generated by  $x_1, x_2, \dots$ . For any group  $G$  we write  $G^{(1)}$  to denote the derived group  $[G, G]$  and  $G^{(2)}$  to denote the second derived group  $[G^{(1)}, G^{(1)}]$ . It is easy to check that  $X^{(1)}$  and  $X^{(2)}$  are normal admissible subgroups of  $(X, y)$  which are closures of the words  $[x_1, x_2]$  and  $[[x_1, x_2], [x_3, x_4]]$  respectively in  $(X, y)$ . The variety of pointed-groups defined by  $[x_1, x_2]$  is the variety of all abelian pointed-groups (we call a pointed-group *abelian* if  $G$  is abelian) and which we shall denote by  $\mathcal{A}$ , while the variety of pointed-groups defined by  $[[x_1, x_2], [x_3, x_4]]$  is the variety of all metabelian pointed-groups, which we shall denote by  $\mathcal{M}^{PG}$ . Clearly  $\mathcal{M}^{PG}$  is finitely based.

We will prove the following theorem which is a generalization of the result in [1].

**THEOREM 2.1.** *Every subvariety of  $\mathcal{M}^{PG}$  is finitely based.*

In order to prove Theorem 2.1 we need some results which are easy generalizations of the results on varieties of groups.

LEMMA 2.2. *Let  $\mathcal{V}$  be a finitely based variety of pointed-groups. Let  $V$  be the corresponding normal admissible subgroup of  $(X, y)$ . Then every subvariety of  $\mathcal{V}$  is finitely based if and only if the maximal condition holds for the normal admissible subgroups of  $(X, y)$  which contain  $V$ .*

Now since the normal admissible subgroup of  $(X, y)$  which defines the variety  $\mathcal{M}^{PG}$  is  $X^{(2)}$ , in order to prove that every subvariety of  $\mathcal{M}^{PG}$  is finitely based, it is enough (by the above Lemma 2.2) to prove that the maximal condition holds for the normal admissible subgroups of  $(X, y)$  containing  $X^{(2)}$ .

LEMMA 2.3. *Let  $V$  and  $W$  be normal admissible subgroups of  $(X, y)$  such that  $V \subseteq W$ . Suppose that the maximal condition holds for the normal admissible subgroups of  $(X, y)$  containing  $W$  and suppose that the maximal condition holds for the normal admissible subgroups  $Z$  of  $(X, y)$  which satisfy  $V \subseteq Z \subseteq W$ . Then the maximal condition holds for the normal admissible subgroups of  $(X, y)$  containing  $V$ .*

Now by Theorem 2.1 of [6], every subvariety of the pointed-group variety  $\mathcal{A}$  is finitely based. Thus the maximal condition holds for the normal admissible subgroups of  $(X, y)$  containing  $X^{(1)}$ . Thus, to prove that every subvariety of  $\mathcal{M}^{PG}$  is finitely based, it is sufficient (by Lemma 2.3) to prove that the maximal condition holds for the normal admissible subgroups  $Z$  of  $(X, y)$  which satisfy  $X^{(2)} \subseteq Z \subseteq X^{(1)}$ .

LEMMA 2.4. *Let  $V$  and  $Z$  be normal admissible subgroups of  $(X, y)$  such that  $V \subseteq Z$ . Then  $Z/V$  is a normal admissible subgroup of  $(X/V, yV)$ .*

LEMMA 2.5. *Let  $V$  and  $W$  be normal admissible subgroups of  $(X, y)$  such that  $V \subseteq W$ . Suppose  $(X/V, yV)$  satisfies the maximal condition on the normal admissible subgroups contained in  $W/V$ . Then  $(X, y)$  satisfies the maximal condition for the normal admissible subgroups  $Z$  such that  $V \subseteq Z \subseteq W$ .*

Now let  $G = X/X^{(2)}$ . Then  $G^{(1)} = X^{(1)}/X^{(2)}$ . Thus to prove that every subvariety of  $\mathcal{M}^{PG}$  is finitely based, it is enough to show (by Lemma 2.5) that the maximal condition is satisfied for the normal admissible subgroups of  $(G, yX^{(2)})$  contained in  $G^{(1)}$ .

Notice that  $G$  is a relatively free group with free generators  $yX^{(2)}$ ,  $x_1X^{(2)}$ , .... We shall change the notation for the free generators of  $G$  and

shall write  $z_1$  to denote  $yX^{(2)}$  and  $z_{i+1}$  to denote  $x_iX^{(2)}$  for all  $i \geq 1$ . Hence  $(G, z_1)$  is a relatively free pointed-group generated by  $z_2, z_3, z_4, \dots$ .

LEMMA 2.6. *The maximal condition is satisfied for normal admissible subgroups of  $(G, z_1)$ , contained in  $G^{(1)}$ , if every normal admissible subgroup of  $(G, z_1)$ , contained in  $G^{(1)}$ , is the closure of some finite subset.*

### 3. Proof of Theorem 2.1

A method for tackling the proof of Theorem 2.1, is due to COHEN [1] and HIGMAN [2]. First we make some preliminary definitions.

DEFINITION. Let  $W$  be a set with a binary relation  $\preceq$ . We say that  $\preceq$  is a *quasi-order* if:

- i.  $s \preceq s$  for all  $s \in W$ .
- ii.  $s_1 \preceq s_2$  and  $s_2 \preceq s_3 \Rightarrow s_1 \preceq s_3$  for all  $s_1, s_2, s_3 \in W$ .

We say that  $\preceq$  is a *well-quasi-order* if it satisfies the following additional condition:

- iii. For every subset  $S$  of  $W$  there is a finite subset  $S_0$  of  $S$  such that, for each element  $s$  of  $S$ , there is an element  $s_0$  of  $S_0$  such that  $s_0 \preceq s$ .

DEFINITION. Let  $W$  be a set with a binary relation  $\leq$ . We say that  $\leq$  is a *linear order* if

- i.  $w \leq w$  for all  $w \in W$ ,
- ii.  $w_1 \leq w_2$  and  $w_2 \leq w_3 \Rightarrow w_1 \leq w_3$  for all  $w_1, w_2, w_3 \in W$ ,
- iii.  $w_1 \leq w_2$  and  $w_2 \leq w_1 \Rightarrow w_1 = w_2$  for all  $w_1, w_2 \in W$ ,
- iv. For  $w_1, w_2 \in W \Rightarrow$  either  $w_1 \leq w_2$  or  $w_2 \leq w_1$ .

DEFINITION. A linear order  $\leq$  on a set  $W$  is said to be a *well-order* if every non empty subset  $S$  of  $W$  contains a least element, i.e., an element  $s_0$  such that  $s_0 \leq s$  for all  $s \in S$ .

DEFINITION. Let  $g$  and  $h$  be elements of  $G^{(1)}$ . We say  $g$  is a *consequence* of  $h$  if  $g$  belongs to the closure of  $h$  in  $(G^{(1)}, z_1)$ . We say that  $g$  is *equivalent* to  $h$ , written as  $g \sim h$ , if  $g$  is a consequence of  $h$  and  $h$  is a consequence of  $g$ , that is, if  $\{g\}$  and  $\{h\}$  have the same closure.

The relation  $\sim$  defined above is an equivalence relation on  $G^{(1)}$ .

The key to the method of the proof of Theorem 2.1 is the following lemma:

LEMMA 3.1. *Suppose that  $W$  is a set with a well-order  $\leq$  and well-quasi-order  $\preceq$ . Suppose that  $\chi$  is a function from  $G^{(1)} \setminus \{1\}$  to  $W$  such that, whenever  $g$  and  $h$  are elements of  $G^{(1)} \setminus \{1\}$  satisfying  $\chi(g) \preceq \chi(h)$ , then there is a consequence  $g_1$  of  $g$  and an element  $h_1$  equivalent to  $h$  such that either  $h_1 = g_1$  or  $\chi(h_1 g_1^{-1}) < \chi(h)$ . Then every normal admissible subgroup of  $(G, z_1)$  contained in  $G^{(1)}$  is the closure of a finite set of elements of  $(G, z_1)$ .*

PROOF. Suppose that  $U$  is a normal admissible subgroup of  $(G, z_1)$  contained in  $G^{(1)}$ . Let  $S = \{\chi(u) \mid u \in U \setminus \{1\}\}$ . Now since  $\preceq$  is a well-quasi-order on  $W$ , so there is a finite subset  $S_0$  of  $S$  such that, for every  $s$  of  $S$ , there is an element  $s_0$  of  $S_0$  such that  $s_0 \preceq s$ .

Let  $A$  be a finite subset of  $U \setminus \{1\}$  so chosen that  $S_0 = \{\chi(u) \mid u \in A\}$ . Let  $U_0$  be the closure of  $A$  in  $(G, z_1)$ . Thus  $U_0 \subseteq U$ . The proof will be complete if we show that  $U_0 = U$ .

Now suppose otherwise and let  $T$  be the non-empty set  $T = \{\chi(u) \mid u \in U \setminus U_0\}$ . Now since  $\leq$  is a well-order on  $W$ , there is a least element of  $T$  i.e., an element  $h$  of  $U \setminus U_0$  such that  $\chi(h) \leq \chi(u)$  for all  $u \in U \setminus U_0$ . Since  $\chi(h) \in S$ , so there is an element  $g$  of  $A$  such that  $\chi(g) \preceq \chi(h)$ . Thus, by the hypothesis of the Lemma, there is a consequence  $g_1$  of  $g$  and an element  $h_1$  equivalent to  $h$  such that either  $g_1 = h_1$  or  $\chi(h_1 g_1^{-1}) < \chi(h)$ . Since  $g_1$  is a consequence of  $g$ , we have  $g_1 \in U_0$ . Also, since  $h_1$  is equivalent to  $h$ , we have  $h_1 \in U \setminus U_0$ . Hence  $h_1 \neq g_1$  and so we have  $\chi(h_1 g_1^{-1}) < \chi(h)$ . But  $h_1 g_1^{-1} \in U \setminus U_0$ . Hence  $\chi(h) \leq \chi(h_1 g_1^{-1})$  which is a contradiction. Thus the Lemma follows.

Now to prove that every subvariety of  $\mathcal{M}^{PG}$  is finitely based, it is enough to show (by Lemma 3.1) that there is a set  $W$  with a well-order  $\leq$ , a well-quasi-order  $\preceq$  and a function  $\chi : G^{(1)} \setminus \{1\} \rightarrow W$  satisfying the condition of Lemma 3.1.

To do this, first it will be necessary to describe a suitable set  $W$  with a well-order  $\leq$  and well-quasi-order  $\preceq$ .

Let  $J$  denote the set of non-negative integers. Let  $V(J)$  denote the set of infinite sequences  $\alpha = (\alpha_i) = (\alpha_1, \alpha_2, \alpha_3, \dots)$  of non negative integers in which  $\alpha_i = 0$  for all but finitely many values of  $i$ . Let  $V_3(J)$  denote the set  $P \times P \times P \times V(J)$  where  $P$  denotes the set of positive integers. Let  $\Phi$  denote the set of all functions  $\phi : J \rightarrow J$  such that  $0\phi = 0$  and  $i\phi < j\phi$  whenever  $i < j$ . Now we define a relation  $\preceq$  on  $V_3(J)$  as follows:

We write  $(r, s, t, (\alpha_i)) \preceq (r', s', t', (\alpha'_i))$  if there exists an element  $\phi \in \Phi$  such that  $r\phi = r'$ ,  $s\phi = s'$ ,  $t\phi = t'$  and  $\alpha_i \leq \alpha'_i \phi$  for all  $i \in P$ .

LEMMA 3.2.  $\preceq$  is a well-quasi-order on  $V_3(J)$ .

PROOF. Clearly  $\leq$  is a well-quasi-order on  $J$ . Hence the Lemma is a special case of the Lemma 3.2 of [5].

Next consider the subset  $V$  of  $V_3(J)$  consisting of those elements  $(r, s, t, (\alpha_i))$  in which  $t = 1$ . For brevity, we write  $(r, s, 1, (\alpha_i))$  simply as  $(r, s, (\alpha_i))$ . Now since  $V_3(J)$  has a well-quasi-order  $\preceq$ , it follows that  $V$  has a well-quasi-order  $\preceq$ . In the abbreviated notation, we have:

$$(r, s, (\alpha_i)) \preceq (r', s', (\acute{\alpha}_i))$$

if and only if there exists  $\phi \in \Phi$  such that  $1\phi = 1$ ,  $r\phi = r'$ ,  $s\phi = s'$  and  $\alpha_i \leq \acute{\alpha}_i\phi$  for all  $i \in P$ .

Now let  $W$  be the set  $V \times P$ . We define a relation  $\preceq$  on  $W$  as follows: We write  $(v, p) \preceq (v', p')$  if  $v \preceq v'$  and  $p \leq p'$ .

LEMMA 3.3.  $\preceq$  is a well-quasi-order on  $W$ .

PROOF. Since  $\leq$  is a well-quasi-order on  $P$  and  $\preceq$  is a well-quasi-order on  $V$ , the Lemma is a special case of Theorem 2.3 of [2].

We have defined above a set  $W$  with a well-quasi-order. The next task is to define a suitable well-order on  $W$ . To do this, first we need to define a well-order  $\leq$  on  $V(J)$  as follows:

We write  $(\alpha_i) \leq (\acute{\alpha}_i)$  if either  $(\alpha_i) = (\acute{\alpha}_i)$  or there exists a  $j$  such that  $\alpha_j < \acute{\alpha}_j$ , but  $\alpha_i = \acute{\alpha}_i$  for all  $i > j$ . It is easy to check that this is a well-order. Now we define a well-order  $\leq$  on  $V$  as follows:

We write  $(r, s, (\alpha_i)) \leq (r', s', (\acute{\alpha}_i))$  if either  $r < r'$  or  $r = r'$ ,  $s < s'$  or  $r = r'$ ,  $s = s'$  and  $(\alpha_i) \leq (\acute{\alpha}_i)$ . It is easy to check that  $\leq$  is well-order on  $V$ .

Now we are ready to define a well-order  $\leq$  on  $W$  as follows:

We write  $(v, p) \leq (v', p')$ , if either  $v < v'$  or  $v = v'$  and  $p \leq p'$ . It is straightforward to check that  $\leq$  is a well-order on  $W$ .

LEMMA 3.4. Suppose that  $(r, s, (\alpha_i))$  and  $(\hat{r}, \hat{s}, (\hat{\alpha}_i))$  are elements of  $V$  such that  $(\hat{r}, \hat{s}, (\hat{\alpha}_i)) < (r, s, (\alpha_i))$ . Let  $\phi$  be any element of  $\Phi$ . Then  $(\hat{r}\phi, \hat{s}\phi, (\hat{\beta}_i)) < (r\phi, s\phi, (\beta_i))$  where  $\hat{\beta}_i\phi = \hat{\alpha}_i$  and  $\beta_i\phi = \alpha_i$  for all  $i$ , while  $\hat{\beta}_j = 0$  and  $\beta_j = 0$  if  $j$  is not in the range of  $\phi$ .

PROOF. The proof is straightforward.

LEMMA 3.5. Suppose that  $(r, s, (\alpha_i))$  and  $(\hat{r}, \hat{s}, (\hat{\alpha}_i))$  are elements of  $V$  such that  $(\hat{r}, \hat{s}, (\hat{\alpha}_i)) < (r, s, (\alpha_i))$ . Let  $(\gamma_i) \in V(J)$ . Then  $(\hat{r}, \hat{s}, (\hat{\alpha}_i + \gamma_j)) < (r, s, (\alpha_i + \gamma_j))$ .

PROOF. The proof is straightforward.

Now the next task is to define a suitable function  $\chi : G^{(1)} \setminus \{1\} \rightarrow W$  so that this function  $\chi$  has the properties required in Lemma 3.1.

Let  $N$  be the normal closure in  $G$  of the set of all commutators  $[z_r, z_s]$  where  $r, s \in P$ . Then  $G/N$  is generated by the elements  $z_i N$ ,  $z_i \in G$ , and these elements commute. Hence  $G/N$  is abelian. Thus  $G^{(1)} \subseteq N$ . But  $[z_r, z_s] \in G^{(1)}$  for all  $r, s \in P$ . Hence  $N \subseteq G^{(1)}$ . Thus  $G^{(1)} = N$ . Therefore,  $G^{(1)}$  is generated by the conjugates of the elements  $[z_r, z_s]$ .

Now let  $w$  be any element of  $G$ . Then we can write  $w$  in the form:

$$w = h z_1^{\alpha_1} z_2^{\alpha_2} \dots z_n^{\alpha_n}$$

where  $h \in G^{(1)}$  and the  $\alpha_i$  are integers. Thus

$$[z_r, z_s]^w = [z_r, z_s]^{h z_1^{\alpha_1} z_2^{\alpha_2} \dots z_n^{\alpha_n}},$$

i.e.,

$$[z_r, z_s]^w = [z_r, z_s]^{z_1^{\alpha_1} z_2^{\alpha_2} \dots z_n^{\alpha_n}}$$

(because  $G^{(1)}$  is abelian and so  $[z_r, z_s]^h = [z_r, z_s]$ ).

Consequently, every element of  $G^{(1)}$  is a product of elements of the form

$$[z_r, z_s]^{z_1^{\alpha_1} z_2^{\alpha_2} \dots z_n^{\alpha_n}}$$

and their inverses. Moreover, since  $G^{(1)}$  is abelian, these elements commute. Now consider

$$([z_r, z_s]^{z_1^{\alpha_1} z_2^{\alpha_2} \dots z_n^{\alpha_n}})^{z_1^{\beta_1} z_2^{\beta_2} \dots z_n^{\beta_n}} = [z_r, z_s]^{z_1^{\alpha_1 + \beta_1} \dots z_n^{\alpha_n + \beta_n}}$$

Now if we choose  $\beta_1, \beta_2, \dots, \beta_n$  to be large enough positive integers, then  $\alpha_1 + \beta_1, \alpha_2 + \beta_2, \dots, \alpha_n + \beta_n$  will be non-negative.

Consequently, if  $g$  is any element of  $G^{(1)}$  then there is a conjugate  $g^{z_1^{\beta_1} z_2^{\beta_2} \dots z_n^{\beta_n}}$  of  $g$  such that  $g$  is a product of elements of the form  $[z_r, z_s]^{z_1^{\alpha_1} z_2^{\alpha_2} \dots z_n^{\alpha_n}}$  and their inverses where the  $\alpha_i$  are non-negative integers. We call such an element of  $G^{(1)}$  a *positive element*. Hence every element of  $G^{(1)}$  is equivalent to a positive element. Now consider the element

$$[z_r, z_s]^{z_1^{\alpha_1} z_2^{\alpha_2} \dots z_n^{\alpha_n}}$$

where the  $\alpha_i$  are non-negative integers. With this element we can associate an element of  $V$  namely  $(r, s, (\alpha_i))$ , where we take  $\alpha_i = 0$  for  $i > n$ .

Now suppose  $g_1$  is a positive element of  $\mathcal{G}^{(1)} \setminus \{1\}$ . Since the elements  $[z_r, z_s]^{z_1^{\alpha_1} \dots z_n^{\alpha_n}}$  commute, so  $g_1$  can be written as a product of powers of distinct positive elements  $[z_r, z_s]^{z_1^{\alpha_1} \dots z_n^{\alpha_n}}$ . We choose the power say

$$\left( [z_r, z_s]^{z_1^{\alpha_1} \dots z_n^{\alpha_n}} \right)^p$$

where  $p \neq 0$  in which  $(r, s, (\alpha_i))$  is the largest with respect to the well order  $\leq$  on  $V$  defined above. We call this the *leading term* of  $g_1$ . Now if  $p$  is negative we consider  $g_1^{-1}$  instead of  $g_1$  where  $g_1^{-1}$  is equivalent to  $g_1$  and if  $p$  is positive we may associate with  $g_1$  the element  $(r, s, (\alpha_i))$  of  $W$ .

Now suppose  $g \in \mathcal{G}^{(1)} \setminus \{1\}$ . As  $g_1$  varies over the set of positive elements equivalent to  $g$ , the leading terms of the elements  $g_1$  give a set of elements  $(r, s, (\alpha_i), p)$  of  $W$ .

We define  $\chi : \mathcal{G}^{(1)} \setminus \{1\} \rightarrow W$  such that  $\chi(g)$  is the least element  $(r, s, (\alpha_i), p)$  in this set of elements of  $W$  (using the well-order  $\leq$  we have defined on  $W$ ). To complete the proof, it remains to show that if  $g$  and  $h$  are elements of  $\mathcal{G}^{(1)} \setminus \{1\}$  satisfying  $\chi(g) \preceq \chi(h)$ , then there is a consequence  $g_1$  of  $g$  and element  $h_1$  equivalent to  $h$  such that either  $h_1 = g_1$  or  $\chi(h_1 g_1^{-1}) < \chi(h)$ .

Now suppose that  $g, h \in \mathcal{G}^{(1)} \setminus \{1\}$  such that  $\chi(g) \preceq \chi(h)$ . Let  $\chi(g) = (r, s, (\alpha_i), p)$  and  $\chi(h) = (r', s', (\alpha'_i), p')$ . Now by the definition of  $\chi$ , there is a positive element  $g'$  equivalent to  $g$  with leading term

$$\left( [z_r, z_s]^{z_1^{\alpha_1} z_2^{\alpha_2} \dots} \right)^p.$$

Similarly, there is a positive element  $h_1$  equivalent to  $h$  with leading term

$$\left( [z_{r'}, z_{s'}]^{z_1^{\alpha'_1} z_2^{\alpha'_2} \dots} \right)^{p'}.$$

Now  $\chi(g) \preceq \chi(h)$ . Therefore, by the definition of  $\preceq$  we have  $p \leq p'$  and there is an element  $\phi$  of  $\Phi$  such that  $1\phi = 1$ ,  $r\phi = r'$ ,  $s\phi = s'$  and  $\alpha_i\phi \leq \alpha'_i$  for all  $i$ . Now by Theorem 2.2.12 of [7] there is an endomorphism  $\theta$  of  $(\mathcal{G}, z_1)$  such that  $z_i\theta = z_i\phi$  for all  $i \geq 2$ . But since  $1\phi = 1$ , we have  $z_i\theta = z_i\phi$  for all  $i \geq 1$ .

Now clearly  $g'\theta$  is a positive element of  $\mathcal{G}^{(1)} \setminus \{1\}$  and by Lemma 3.4 the leading term of  $g'\theta$  is

$$\left( [z_{r\phi}, z_{s\phi}]^{z_1^{\alpha_1} z_2^{\alpha_2} \dots} \right)^p,$$

i.e.,

$$\left( [z_{r'}, z_{s'}]^{z_1^{\beta_1} z_2^{\beta_2} \dots} \right)^p,$$

where  $\beta_i \leq \alpha_i$  for all  $i$ .

Now let  $g_1 = (\xi\theta)^{z_1^{\alpha_1 - \beta_1} z_2^{\alpha_2 - \beta_2} \dots}$ . Then clearly  $g_1$  is a positive element and by Lemma 3.5 the leading term of  $g_1$  is

$$([z_r, z_s]^{z_1^{\alpha_1} z_2^{\alpha_2} \dots})^p,$$

Clearly  $g_1$  is a consequence of  $g$  and we have  $p \leq p'$ . Now if  $p < p'$ , then the leading term of  $h_1 g_1^{-1}$  is

$$([z_r, z_s]^{z_1^{\alpha_1} z_2^{\alpha_2} \dots})^{(p-p)},$$

while if  $p = p'$  then either  $h_1 = g_1$  or else  $h_1 g_1^{-1}$  has leading term which is a power of an element

$$[z_r'', z_s'']^{z_1^{\alpha_1} z_2^{\alpha_2} \dots},$$

in which  $(r'', s'', (\alpha_i'')) < (r', s', (\alpha_i))$ . Thus if  $h_1 \neq g_1$  then  $\chi(h_1 g_1^{-1}) < \chi(h)$  as required.

Now, by Lemma 3.1, every normal admissible subgroup of  $(G, z_1)$  contained in  $G^{(1)}$ , is the closure of a finite set of elements of  $(G, z_1)$ . Hence every subvariety of  $\mathcal{M}^{PG}$  is finitely based. This proves Theorem 2.1.

**COROLLARY 3.6.** The metabelian variety  $\mathcal{M}^{PG}$  of all metabelian pointed-groups is finitely based.

## References

- [1] D. E. COHEN, On the laws of a metabelian variety, *J. Algebra*, **5** (1967), 267–273.
- [2] G. HIGMAN, Ordering by divisibility in abstract algebras, *Proc. London Math. Soc.*, **2** (1952), 326–336.
- [3] H. NEUMANN, *Varieties of groups*, Springer Verlag, Berlin (1967).
- [4] P. M. COHN, *Universal algebra*, Harper and Row, New York, (1965).
- [5] R. H. BRYANT and M. F. NEWMANN, Some finitely based varieties of groups, *Proc. London Math. Soc.*(3), **28** (1974), 237–252.
- [6] Z. ALI and A. MAJEED, On the laws of nilpotent pointed-groups, *Bull. Korean Math. Soc.*, **28** (1991), 1–7.
- [7] Z. ALI, Finite basis problem for pointed-groups, Ph. D. Thesis submitted to the Punjab Univ. Lahore.
- [8] S. MACLANE and G. BIRKHOFF, *Modern Algebra* Macmillan, London (1976).



## A REFINEMENT OF THE ENTROPY INEQUALITY

By

H. ALZER

Waldbröl, F. R. Germany

(Received May 27, 1993)

### 1. Introduction

The well-known entropy inequality states:

PROPOSITION A. *If  $p_i$  ( $i = 1, \dots, n$ ) are non-negative real numbers with*

$$\sum_{i=1}^n p_i = 1, \text{ then}$$

$$(1.1) \quad \log(n) + \sum_{i=1}^n p_i \log(p_i) \geq 0,$$

*with equality holding if and only if  $p_1 = \dots = p_n = 1/n$ .*

(In (1.1) and in what follows we define  $0^x \log(0) := 0$  ( $x > 0$ .)

Inequality (1.1) plays a significant role in the theory of information ([2] p. 55, [3] pp. 13–14). It can be proved easily by applying Jensen's inequality for convex functions. Closely related results are given in [1], [5], [6] pp. 405–407, [8].

In 1980 K. B. STOLARSKY [10] presented an interesting extension of (1.1). By introducing a new variable  $\tilde{x}$ , inequality (1.1) can be written as

$$(1.2) \quad S(x) = \sum_{i=1}^n (np_i)^x \log(np_i) \geq 0$$

where  $x = 1$ . If the  $p_i$ 's are not all equal, then  $S$  is strictly increasing on  $(0, \infty)$ , so that the validity of (1.2) with  $x \in (0, 1)$  would provide a sharpening of the entropy inequality (1.1). The following result is proved in [10]:

PROPOSITION B. Let  $p_i$  ( $i = 1, \dots, n; n \geq 2$ ) be non-negative real numbers with  $\sum_{i=1}^n p_i = 1$ . If  $h(n) = \frac{n}{n-1} - \frac{1}{\log(n)}$ , then

$$(1.3) \quad S(h(n)) \geq 0.$$

Because of  $h(n) \in (0, 1)$ , inequality (1.3) refines the entropy inequality  $S(1) \geq 0$ . Moreover, STOLARSKY established that for  $n = 2$  inequality (1.3) is even true if  $h(2) = 0.557\dots$  is replaced by  $1/2$ , and that this result is best possible. The problem to find the best possible value  $x = x(n)$  for  $n > 2$  remains open.

STOLARSKY mentioned that for sufficiently large  $n$  inequality (1.3) “cannot be improved too much” ([10] p. 242), since there exist positive real numbers  $p_i$  ( $i = 1, \dots, n$ ) with  $\sum_{i=1}^n p_i = 1$  and a sequence  $y(n)$  with  $y(n) \rightarrow 1$  as  $n \rightarrow \infty$  such that  $S(y(n))$  is negative for sufficiently large  $n$ . In [10] it is shown that  $S(y(n)) < 0$  for large  $n$ , if

$$p_1 = \frac{1}{n}(1 + \sqrt{n-1}), \quad p_2 = \dots = p_n = \frac{1}{n}(1 - 1/\sqrt{n-1})$$

and

$$y(n) = 1 - \frac{4 \log(\log(n))}{\log(2n)}.$$

It is worth mentioning that a slight modification of the final part of Stolarsky’s proof leads to a refinement of (1.3) for all  $n \geq 2$ . It is the aim of this note to present this refinement which is even best possible in a certain sense.

## 2. A refinement of the entropy inequality

In what follows we denote by

$$L(x, y) = \frac{x - y}{\log(x) - \log(y)}, \quad x \neq y,$$

the logarithmic mean of the positive real numbers  $x$  and  $y$ .

THEOREM. Let  $p_i$  ( $i = 1, \dots, n; n \geq 2$ ) be non-negative real numbers with  $\sum_{i=1}^n p_i = 1$ . If

$$t(n) = \frac{\log L(n, 1)}{\log(n)},$$

then

$$(2.1) \quad S(t(n)) \geq 0.$$

REMARK. The function

$$t(x) = \begin{cases} \frac{\log L(x, 1)}{\log(x)}, & x > 1, \\ 1/2, & x = 1, \end{cases}$$

is continuous and increases monotonically to 1 as  $x \rightarrow \infty$ . The monotonicity of  $t$  can be proved as follows:

Let  $x > 1$ ; we have

$$(2.2) \quad x(\log(x))^2 t'(x) = \frac{x \log(x)}{x-1} - 1 - \log(x-1) + \log(\log(x)).$$

If we denote the sum on the right-hand side of (2.2) by  $q(x)$ , then we get

$$\log(x)q'(x) = (G(x, 1))^{-2} - (L(x, 1))^{-2},$$

where  $G(x, y) = \sqrt{xy}$ . Because of  $G(x, y) < L(x, y)$ ,  $x \neq y$ , (see [7] p. 272), we obtain

$$q'(x) > 0 \quad \text{and} \quad q(x) > q(1) = 0.$$

Hence,  $t'(x) > 0$  for  $x > 1$ , that is,  $t$  is strictly increasing on  $[1, \infty)$ . Furthermore, from the identity

$$h(x) - t(x) = x \log(x)t'(x)$$

we conclude

$$h(n) > t(n) \quad \text{for} \quad n = 2, 3, \dots$$

This implies

$$S(h(n)) \geq S(t(n)) \quad \text{for} \quad n = 2, 3, \dots,$$

with equality holding if and only if  $p_1 = \dots = p_n = 1/n$ .

PROOF OF THE THEOREM. Our proof is modelled after Stolarsky's proof of inequality (1.3). It is shown in [10] that (1.3) follows from three propositions which are called Theorem A, Theorem B and Theorem C. An analysis reveals that Theorem A, Theorem B and Theorem C, part (ii) remain valid if  $h(n)$  is replaced by  $t(n)$ . We show that the same is true for Theorem C, part (i). We must prove the following:

$$(2.3) \quad \begin{aligned} &\text{If } x \geq 0, y \geq 0, \Theta \geq t(n), \frac{1}{n} \leq \lambda \leq 1 - \frac{1}{n}, \text{ and} \\ &\lambda x + (1 - \lambda)y = 1, \end{aligned}$$

then for  $(x, y) \notin \{(1, 1), (1, 0), (0, 1), (0, 0)\}$  it follows

$$(2.4) \quad \lambda x^\Theta \log(x) + (1 - \lambda)y^\Theta \log(y) > 0.$$

To establish this assertion we may assume  $0 < y < 1 < x$ .

From (2.3) we get

$$(2.5) \quad x < \frac{1}{\lambda} \leq n,$$

and

$$\lambda = \frac{1 - y}{x - y},$$

so that (2.4) can be written as

$$(1 - y)x^\Theta \log(x) + (x - 1)y^\Theta \log(y) > 0$$

or

$$(2.6) \quad g(y) < g(x),$$

where

$$g(x) = \frac{x^\Theta \log(x)}{x - 1}.$$

Since  $t$  is strictly increasing on  $[1, \infty)$  we get from (2.5):  $\Theta \geq t(n) > t(x)$  which implies

$$(2.7) \quad g(x) > 1 \quad \text{for} \quad x > 1.$$

And, since  $\Theta \geq t(n) > 1/2$  we obtain

$$(2.8) \quad g(y) < \frac{\sqrt{y} \log(y)}{y - 1} = \frac{G(y, 1)}{L(y, 1)} < 1 \quad \text{for} \quad y \in (0, 1).$$

From (2.7) and (2.8) we conclude the validity of (2.6). Hence, the Theorems A, B and C in [10] are also valid if  $h(n)$  is replaced by  $t(n)$ . And, this implies that inequality (1.3) with  $t(n)$  instead of  $h(n)$  holds, too. Inequality (2.1) is proved. ■

### 3. Concluding remarks.

The logarithmic mean  $L(x, y)$  is member of the one-parameter mean value family

$$L_r(x, y) = \left[ \frac{x^r - y^r}{r(x - y)} \right]^{1/(r-1)}, \quad r \in \mathbb{R} - \{0, 1\}; \quad x, y \in \mathbb{R}^+,$$

since

$$\lim_{r \rightarrow 0} L_r(x, y) = \frac{x - y}{\log(x) - \log(y)}.$$

In the recent past  $L_r(x, y)$  has been investigated by many authors. This is in particular true for the logarithmic mean and the identric mean  $L_1(x, y) = \frac{1}{e} (x^x / y^y)^{1/(x-y)}$ . We refer to the monograph ([4] Chapter VI) and the references therein. If we define

$$t_r(n) = \frac{\log L_r(n, 1)}{\log(n)}$$

and

$$L_\infty(x, y) = \lim_{r \rightarrow \infty} L_r(x, y) = \max(x, y),$$

then the entropy inequality and inequality (2.1) can be written as

$$S(t_\infty(n)) \geq 0 \quad \text{and} \quad S(t_0(n)) \geq 0.$$

Since  $r \mapsto L_r(x, y)$  is increasing on  $\mathbb{R}$  (see [9]) it follows that  $S(t_r(n))$  is also increasing with respect to  $r$ . This leads to the question: What is the smallest real number  $r$ , such that

$$(3.1) \quad S(t_r(n)) \geq 0$$

holds for all  $n \geq 2$ ?

We assume that there exists a number  $r < 0$ , so that

$$(3.2) \quad \sum_{i=1}^n (np_i)^{t_r(n)} \log(np_i) \geq 0$$

is valid for all  $n \geq 2$  and all  $p_i \geq 0$  ( $i = 1, \dots, n$ ) with  $\sum_{i=1}^n p_i = 1$ . If we set

$$p_1 = \dots = p_{n-1} = \frac{1}{2n}, \quad p_n = \frac{n+1}{2n},$$

then (3.2) is equivalent to

$$(3.3) \quad t_r(n) \geq \frac{\log((n-1)\log(2)) - \log(\log((n+1)/2))}{\log(n+1)}.$$

And, if  $n$  tends to  $\infty$ , then (3.3) reduces to

$$1/(1-r) \geq 1 \quad \text{or} \quad r \geq 0.$$

Hence, inequality (3.1) holds if and only if  $r$  is non-negative.

### References

- [1] J. ACZÉL, Problem 3 from Reports of Meetings, *Aequat. Math.*, **1** (1968), 300.
- [2] E. F. BECKENBACH and R. BELLMAN, *Inequalities*, Springer, Berlin, (1983).
- [3] L. BRILLOUIN, Science and Information Theory, *Academic Press, New York*, (1962).
- [4] P. S. BULLEN, D. S. MITRINOVIĆ and P. M. VASIĆ, Means and Their Inequalities, *Reidel, Dordrecht*, (1988).
- [5] P. FISCHER, Sur l'inégalité  $\sum_{i=1}^n p_i f(p_i) \geq \sum_{i=1}^n p_i f(q_i)$ , *Aequat. Math.*, **2** (1969), 363.
- [6] A. W. MARSHALL and I. OLKIN, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, New York, (1979).
- [7] D. S. MITRINOVIĆ, *Analytic Inequalities*, Springer, New York, (1970).
- [8] GY. MUSZÉLY, Über die stetigen Lösungen der Ungleichung  $pf(p) + (1-p)f(1-p) \geq pf(q) + (1-p)f(1-q)$ , *Aequat. Math.*, **2** (1969), 362–363.
- [9] K. B. STOLARSKY, Generalizations of the logarithmic mean, *Math. Mag.*, **48** (1975), 87–92.
- [10] K. B. STOLARSKY, A stronger logarithmic inequality suggested by the entropy inequality, *SIAM J. Math. Anal.*, **11** (1980), 242–247.

## INTERNAL CONTROLLABILITY OF THE STRING WITH RESTRAINED CONTROLS

By

I. JOÓ and N. V. SU

Department of Analysis, Eötvös Loránd University, Budapest

*(Received October 10, 1993)*

*(Revised April 22, 1994)*

### I. Introduction

Consider a string along the segment  $[0, 1]$  fixed at the two ends 0 and 1 and controlled on the interior of the domain with a function  $h(x, t)$  described by the following wave equation

$$(1) \quad \begin{aligned} y_{tt} &= y_{xx} + h\chi_\omega, & x &\in (0, 1), & t &> 0 \\ y(0, t) &= y(1, t) = 0, & t &> 0 \\ y(x, 0) &= y_0(x), & y_t(x, 0) &= y_1(x), \end{aligned}$$

where  $\omega \subset [0, 1]$  is an open and non empty subset of  $[0, 1]$ ,  $\chi_\omega$  denotes the characteristic function of  $\omega$ .

The control problem of such system was investigated by many authors see e.g. J. L. LIONS [1], [2], [3], [4], [5], J. L. LIONS and E. MAGENES [6], A. HARAUX [7], [8], E. ZUAZUA [10], I. JOÓ [11], M. HORVÁTH [12], A. BOGMÉR [13]. Up to present, as far as we know, the authors investigate the problem of the controllability only in the case when the controls  $h$  run over full space of controls, by different methods. J. L. LIONS, E. ZUAZUA by method HUM showed that the system (1) (in general case) without restrained controls is always controllable to the origin. I. JOÓ, M. HORVÁTH, A. BOGMÉR by using method of Riesz's bases investigated the system (1) without restrained control, but in only sense of the controllability from the origin. In recent years J. LAGNESE [9] have investigated in details the exact controllability of solutions of wave equations by means of distributed control forces constrained to vanish outside of a given subset of the region in which the process evolves. Later I. JOÓ and N. V. SU [14], [15] have studied the above system with restrained controls by using discretization method (see

[16], [17]). They also obtained some conditions for approximate controllability to the origin (resp. from the origin). In this paper we will study the exact controllability (to the origin) of the system (1) with restrained controls. Note that the semigroup theory technics are used there with the results obtained in [9] or the results of method HUM. Necessary and sufficient condition for the controllability of the system (1) will be given.

## II. Definitions, Notations

Let  $\Omega_T \subset L^2(\omega \times (0, T))$  be a convex subset for which  $\text{int } \Omega_T \neq \emptyset$  in  $L^2(\omega \times (0, T))$  and  $0 \in \Omega_T$  for all  $T > 0$ . Consider the system (1) in the space  $H_0^1(0, 1) \times L^2(0, 1)$  with restrained controls in  $\Omega_T \subset L^2(\omega \times (0, T))$ . Let  $y = y(x, t)$  be the solution of the system (1) with the control  $h(x, t)$ . Define the following set

$$S_T(\Omega_T) := \{(y_0, y_1) \in H_0^1(0, 1) \times L^2(0, 1) :$$

there is  $h \in \Omega_T$  such that the solution  $y(x, t)$  of the system (1) satisfies

$$y(x, T) = 0, \quad y_t(x, T) = 0\}.$$

The set

$$S = \bigcup_{T > 0} S_T(\Omega_T)$$

is so called the controllability set after unbounded time.

DEFINITION. We say that the system (1) is locally controllable in the space  $H_0^1(0, 1) \times L^2(0, 1)$  with restrained controls if  $(0, 0) \in \text{int } S$  and globally controllable with restrained controls if

$$S \equiv H_0^1(0, 1) \times L^2(0, 1).$$

As in [18] (Lion's book, Chapter IV) we would like to rewrite the system (1) (hyperbolic type) as a system of first order equations (parabolic case).

Let us denote

$$\begin{aligned} w(x, t) &= \begin{pmatrix} y(x, t) \\ y_t(x, t) \end{pmatrix}; \\ \mathcal{H} &= H_0^1(0, 1) \times L^2(0, 1); \\ \mathcal{A} &= \begin{pmatrix} 0 & I \\ \Delta & 0 \end{pmatrix}; \end{aligned}$$

where  $I$  is the identity in  $L^2(0, 1)$ ,  $\Delta$  is the operator  $\partial/\partial x^2$ .

Let further

$$\bar{h}\chi_\omega = \begin{pmatrix} 0 \\ h\chi_\omega \end{pmatrix}.$$

Then we get formally the following first order equation in space  $\mathcal{H}$

$$(2) \quad \begin{cases} \frac{dw(x,t)}{dt} = \mathcal{A}w(x,t) + \bar{h}\chi_\omega \\ w(x,0) = w_0 \end{cases}$$

where

$$w_0 = \begin{pmatrix} y_0(x) \\ y_1(x) \end{pmatrix} \in \mathcal{H}.$$

This first order equation unfortunately is not analogous to that J. L. LIONS has studied in his book (see [18], Chapter IV, pp.306), because the operator  $\mathcal{A}$  is not coercive. But from Lemma 3.6 [3] (or see more detail in PAZY's book [24]) the "mild solution" of the abstract equation (2) in this case is represented by semigroup in the form

$$(3) \quad w(x,t) = G(t)w_0 + \int_0^t G(t-\tau)\bar{h}(x,\tau)\chi_\omega d\tau,$$

where

$$G(t) = \begin{pmatrix} I & e^{tI} - I \\ N(t) - I & I \end{pmatrix}.$$

Here  $N(t)$  is the semigroup generated by  $\Delta = \partial^2/\partial x^2$ .

REMARK. We may easily verify that

$$N(t) = \sum_{j=1}^{\infty} e^{\lambda_j t} \langle v, \sqrt{2} \sin j \pi \tau \rangle$$

where

$$\lambda_j = -j^2 \pi^2, \quad v \in D(\Delta) := \left\{ v \in L^2(0,1) : \sum_{j=1}^{\infty} |\lambda_j|^2 |\langle v, \sqrt{2} \sin j \pi \tau \rangle| < \infty \right\}.$$

(see Example 1.3, Chapter III. [18]).

So the controllability set of the system (1) has the following form

$$S_T(\Omega_T) = \left\{ w \in \mathcal{H} : G(T)w + \int_0^T G(T-\tau)\bar{h}(x,\tau)\chi_\omega d\tau = 0 \text{ for some } h \in \Omega_T \right\},$$

and

$$S = \bigcup_{T>0} S_T(\Omega_T).$$

REMARK. In general case (see J. L. LIONS [18]), the solution of the following equation

$$\frac{d^2y}{dt^2} + A(t)y = f,$$

where  $A$  is a second-order elliptic differential operator, can not be written in the form (3).

For the next part we need some crucial results which can be stated as follows.

THEOREM A. (KREIN–RUTMAN Theorem [21]). *Let  $C$  be a convex cone with non-empty interior in a Banach space  $X$  and let  $\{S(t)\}$  be a family of commutative bounded linear operators, mapping the cone  $C$  into itself, i.e.  $S(t)C \subset C$  for all  $t$ . Then there exists a positive functional  $f \in C^* \subset X^*$ , which is a common eigenvector of all dual operators:*

$$S^*(t)f = \lambda(t)f, \quad \text{for all } t,$$

where  $\lambda(t) \geq 0$ .

THEOREM B. (see Lemma 1, [22]). *Let  $M_n$  be convex subsets in an arbitrary Banach space  $X$  such that  $\text{int } M_n \neq \emptyset$ ,  $0 \in M_n$ , for all  $n \in \mathbb{N}$  and  $M_n \subset M_m$  for all  $nb < m$ . Suppose that*

$$0 \in \text{int} \left( \bigcup_{n=1}^{\infty} M_n \right),$$

then there exists some  $k \in \mathbb{N}$  such that  $0 \in \text{int } M_k$ .

### III. Main Results

In this part we will give necessary and sufficient condition for the local controllability of the system (1).

First we need the following Lemma.

LEMMA. *Consider the system (1) with the conditions that  $\Omega_T \subset L^2(\omega \times (0, T))$  is convex subset for which  $\text{int } \Omega_T \neq \emptyset$  in  $L^2(\omega \times (0, T))$  and  $0 \in \Omega_T$  for all  $T > 0$ . Then the controllability set  $S$  has nonempty interior in  $H_0^1(0, 10) \times L^2(0, 1)$ , that is  $\text{int } S \neq \emptyset$  in  $H_0^1(0, 1) \times L^2(0, 1)$ .*

PROOF. Since  $\Delta = \partial^2/\partial x^2$  it follows that the generated semigroup  $N(t)$  is unitary (see [19]), so these are surjective too. By method HUM (see [4] LION's or [10], ZUAZUA, or see [9]) we know that the system (1) without restrained controls is globally controllable after some time  $T_0$ , that is, by taking

$$\begin{aligned} \mathcal{B}_T &: \{0\} \times L^2(\omega \times (0, T)) \rightarrow \mathcal{H} \\ \mathcal{B}_T \bar{h} &= \int_0^T G(T-\tau) \bar{h}(x, \tau) \chi \omega d\tau \\ &\text{for } \bar{h} \in \{0\} \times L^2(\omega \times (0, T)) \end{aligned}$$

we have

$$(5) \quad \text{Im } G(T_0) \subset \text{Im } \mathcal{B}_{T_0}$$

where  $\text{Im } M$  denotes the image of the operator  $M$ . By the factorization theorem of DOUGLAS [20] we have that there exists a linear, bounded operator

$$C : \mathcal{H} \rightarrow \{0\} \times L^2(\omega \times (0, T_0))$$

such that

$$(6) \quad G(T_0) = \mathcal{B}_{T_0} C.$$

Since  $N(T_0)$  is surjective, it follows that  $G(T_0)$  is surjective too. From  $\text{int } \Omega_{T_0} \neq \emptyset$  in  $L^2(\omega \times (0, T_0))$  we have that the set  $\text{int } (\{0\} \times \Omega_{T_0})$  is non-empty in  $\{0\} \times L^2(\omega \times (0, T_0))$ . From (5) and the surjectivity of  $G(T_0)$  it follows that there are some  $\omega \in \mathcal{H}$  and  $h_0 \in \text{int } (\{0\} \times \Omega_{T_0})$  such that

$$G(T_0)\omega = \mathcal{B}_{T_0} \bar{h}_0.$$

From (6) we get  $\mathcal{B}_{T_0} \bar{h}_0 = \mathcal{B}_{T_0} C\omega$ , which implies that

$$(\ker \mathcal{B}_{T_0} + \text{int } (\{0\} \times \Omega_{T_0})) \cap \text{Im } C \neq \emptyset.$$

Consequently, the inverse image  $C^{-1}(\ker \mathcal{B}_{T_0} + \text{int } (\{0\} \times \Omega_{T_0}))$  has a non-empty interior in  $\mathcal{H}$ . It is easily to verify that

$$(7) \quad -C^{-1}(\ker \mathcal{B}_{T_0} + \text{int } (\{0\} \times \Omega_{T_0})) \subset S_{T_0}(\Omega_{T_0}).$$

In fact, if

$$w \in (-C^{-1}(\ker \mathcal{B}_{T_0} + \text{int } (\{0\} \times \Omega_{T_0}))),$$

then  $-Cw = b + \bar{h}_0$  for some  $b \in \ker \mathcal{B}_{T_0}$  and for some  $\bar{h}_0 \in \text{int } (\{0\} \times \Omega_{T_0})$ . Hence

$$-\mathcal{B}_{T_0} Cw = \mathcal{B}_{T_0}(b + \bar{h}_0) = \mathcal{B}_{T_0} \bar{h}_0.$$

But  $G(T_0) = \mathcal{B}_{T_0} C$  (from (6)) therefore  $G(T_0)w + \mathcal{B}_{T_0} \bar{h}_0 = 0$ .

This means that  $\omega \in S_{T_0}(\Omega_{T_0})$ , and (7) shows that  $\text{int } S_{T_0}(\Omega_{T_0}) \neq \emptyset$ , so  $\text{int } S \neq \emptyset$  in  $H_0^1(0, 1) \times L^2(0, 1)$ .

The proof of Lemma is complete. ■

**THEOREM.** *Consider the system (1) with the restrained controlled set  $\Omega_T$ . Assume that the conditions for controlled set  $\Omega_T$  in the above lemma are fulfilled. Then the system (1) with the restrained controls is locally controllable if and only if the following condition is satisfied*

(\*) *There is no "line"  $cx + d$ ,*

*( $c(t)$  and  $d(t)$  are functions of time  $t$ ,  $t \in [0, T]$ ) such that*

$$\langle cx + d, h\chi \rangle_{L^2(\omega, (0, T))} \geq 0$$

*for all  $h \in \Omega_T$  and for all  $T > 0$ .*

**PROOF. SUFFICIENCY.** Assume that the condition (\*) is fulfilled. We will show that the system (1) is locally controllable. From the above Lemma we have  $\text{int } S_{T_0}(\Omega_{T_0}) \neq \emptyset$  in  $\mathcal{H}$ .

For  $T > T_0$  we put

$$S'_T := \left\{ w \in \mathcal{H} : -G(T_0)w \in \mathcal{B}_T \begin{pmatrix} 0 \\ \Omega_T \end{pmatrix} \right\},$$

where

$$\mathcal{B}_T \begin{pmatrix} 0 \\ \Omega_T \end{pmatrix} := \left\{ \mathcal{B}_T \begin{pmatrix} 0 \\ h \end{pmatrix} : h \in \Omega_T \right\}.$$

Since  $0 \in \Omega_T$  for all  $T$  it is easy to see that  $S'_{T_1} \subset S'_{T_2}$  if  $T_1 < T_2$  and  $S'_T$  is convex for  $T > T_0$ . By definition of  $S'_T$  it follows that  $S'_{T_0} = S_{T_0}(\Omega_{T_0})$ , so this means that  $\text{int } S'_T \neq \emptyset$  for  $T \geq T_0$ .

Setting  $S' = \bigcup_{T \geq T_0} S'_T$ , we will show that  $0 \in \text{int } S'$ . Assuming the contrary,

we will show that the cone  $C = \bigcup_{\lambda > 0} \lambda S'$  is convex, not dense in  $\mathcal{H}$ . In fact,

are convexity of  $C$  is trivial. Assume that  $C$  is dense in  $\mathcal{H}$ , that  $\bar{C} = \mathcal{H}$ , so  $0 \in \text{int } \bar{C}$ . Since  $\text{int } C \neq \emptyset$  (because  $\text{int } S'_T \neq \emptyset$  for  $T \geq T_0$ ) it follows that  $\text{int } C = \text{int } \bar{C}$ . It means that  $0 \in \text{int } C$ , thus  $0 \in \text{int } S'$ . This is a contradiction. We show that  $C$  is  $G(t)$ -invariant for all  $t$ , i.e.  $G(t)C \subset C$  for all  $t$ .

In fact if  $w \in C$ , then for some  $T$  and  $\lambda$

$$-G(T_0)w \in \lambda \mathcal{B}_T \begin{pmatrix} 0 \\ \Omega_T \end{pmatrix}.$$

It follows that for all  $t$

$$-G(t)G(T_0)w \in \lambda G(t)\mathcal{B}_T \begin{pmatrix} 0 \\ \Omega_T \end{pmatrix}.$$

From  $0 \in \Omega_T$  we have that

$$\lambda G(t)\mathcal{B}_T \begin{pmatrix} 0 \\ \Omega_T \end{pmatrix} \subset \lambda \mathcal{B}_{T+t} \begin{pmatrix} 0 \\ \Omega_{T+t} \end{pmatrix}.$$

So we get

$$-G(t)G(T_0)w \subset \lambda \mathcal{B}_{T+t} \begin{pmatrix} 0 \\ \Omega_{T+t} \end{pmatrix}.$$

This means  $G(t)w \in \lambda S'_{T+t}$  so  $G(t)w \in C$ .

Applying the Krein–Rutman's theorem [21], we find in the dual cone  $C^*$  a common eigenvector  $w_0$  of all dual operators  $G^*(t)$ ,  $t \geq 0$ , with nonnegative eigenvalues  $\lambda(t)$ ,  $t > 0$ , i.e.

$$(8) \quad G^*(t)w_0 = \lambda(t)w_0, \quad \lambda(t) \geq 0 \quad \text{for all } t \geq 0.$$

Since  $\mathcal{H}$  is Hilbert space we have, by the differentiability of  $G(t)$ , that  $G^*(t)\mathcal{H} \subset \mathcal{D}(\mathcal{A}^*)$  for all  $t \geq 0$ . Thus  $w_0 \in \mathcal{D}(\mathcal{A}^*)$  and  $G^*(t)w_0$  is differentiable for  $t > 0$ . By differentiating the equality (8) we get

$$\mathcal{A}^* G^*(t)w_0 = G^*(t)\mathcal{A}^* w_0 = \lambda(t)w_0 \quad \text{for all } t > 0.$$

From the strong continuity of  $G^*(t)$  it follows that

$$\lim_{t \rightarrow +0} G^*(t)\mathcal{A}w_0 = \mathcal{A}^* w_0 = \lambda_0 w_0$$

where  $\lambda_0 = \lim_{t \rightarrow +0} \lambda(t)$ . Thus we have shown that there is the eigenvector  $w_0$  of  $\mathcal{A}^*$  with real eigenvalue  $\lambda_0$  such that  $w_0 \in C^*$ . From the definition of  $S'_T$  and  $0 \in \Omega_T$  for all  $T > 0$  it is easy to see that for some  $t > 0$

$$\int_0^t G(T-t)\bar{h}(x,\tau)\chi_\omega d\tau \in C \quad \text{for all } h \in \Omega_t.$$

This means that

$$\left\langle w_0, \int_0^t G(t-\tau)\bar{h}(x,\tau)\chi_\omega d\tau \right\rangle_{\mathcal{H}} \geq 0 \quad \text{for all } h \in \Omega_t.$$

From the strong continuity of  $G(t)$  we obtain that

$$(9) \quad \langle w_0, G(t)\bar{h}(x,\tau)\chi_\omega \rangle_{\mathcal{H}} \geq 0 \quad \text{for all } t > 0 \quad \text{and for all } h \in \Omega_t.$$

Since

$$\mathcal{A}^* = \begin{pmatrix} 0 & \Delta \\ I & 0 \end{pmatrix}$$

and the eigenvalues of  $\Delta$  are  $-j^2\pi^2, j = 1, 2, \dots$  it follows that the eigenvector  $w_0$  of  $\mathcal{A}^*$  with real eigenvalue is only form:

$$(10) \quad w_0 = \begin{pmatrix} 0 \\ c(t)x + d(t) \end{pmatrix}$$

where  $c(t)$  and  $d(t)$  are functions of  $t$ . From the formula of  $G(t)$  and from (9) and (10) we get

$$\langle c(t)x + d(t), h\chi_\omega \rangle_{L^2(\omega, (0, T))} \geq 0$$

for all  $h \in \Omega_T$  and for all  $T > 0$ . This contradicts the condition (\*) of the Theorem. Thus,  $0 \in S'$ . In view of Lemma 1 of [22], there exists  $T \geq T_0$  such that  $0 \in \text{int } S'_T$ .

If  $T = T_0$  then  $0 \in \text{int } S_{T_0}(\Omega_{T_0})$  because  $S'_{T_0} = S_{T_0}(\Omega_{T_0})$ .

If  $T > T_0$  we consider operator  $G(T - T_0) : \mathcal{H} \rightarrow \mathcal{H}$ .

Since  $0 \in \text{int } S'_T$  we have  $0 \in \text{int } (G^{-1}(T - T_0)S'_T)$ . On the other hand, it is easy to see that  $G^{-1}(T - T_0)S'_T \subset S_T(\Omega_T)$ . Therefore,  $0 \in \text{int } S_T(\Omega_T)$ , thus  $0 \in \text{int } S$ .

So the sufficiency is proved.

NECESSITY. Assume that the system (1) is locally controllable, we will show that the condition (\*) is fulfilled.

Assume the contrary, there exists some "line"  $:= cx + d$ , where  $c(t)$  and  $d(t)$  are functions of time  $t$ ,  $t \in [0, T]$  such that

$$\langle cx + d, h\chi_\omega \rangle_{L^2(\omega, (0, T))} \geq 0 \quad \text{for all } h \in \Omega_t \text{ and for all } T > 0.$$

Let

$$w_0 = \begin{pmatrix} 0 \\ cx + d \end{pmatrix}.$$

Then it is easy to verify that

$$(11) \quad \mathcal{A}^* w_0 = 0.$$

since  $0 \in \text{int } S$  and  $\text{int } S_{T_0} \neq \emptyset$  (by the above Lemma). It follows from Lemma 1 of [22] that there is some  $T_1 \geq T_0$  such that  $0 \in \text{int } S_{T_1}(\Omega_{T_1})$ . It is well known that

$$S_{T_1}(\Omega_{T_1}) = \left\{ w \in \mathcal{H} : -G(T_1)w \in \mathcal{B}_{T_1} \left( \begin{pmatrix} 0 \\ \Omega_{T_1} \end{pmatrix} \right) \right\}.$$

Let  $w \in S_{T_1}(\Omega_{T_1})$  be arbitrary, then

$$G(T_1)w + \int_0^{T_1} G(T_1 - \tau)\bar{h}(x, \tau)\chi_\omega d\tau = 0$$

for some  $h \in \Omega_{T_1}$ . From this we get

$$\langle w_0, G(T_1)w \rangle + \int_0^{T_1} \langle w_0, G(T_1 - \tau)\bar{h}(x, \tau)\chi_\omega \rangle d\tau = 0.$$

This means

$$(12) \quad \langle G^*(T_1)w_0, w \rangle = - \int_0^{T_1} \langle G^*(T_1, \tau)w_0, \bar{h}(x, \tau)\chi_\omega \rangle d\tau.$$

Since, obviously,  $w_0 \in \bigcap_{n=1}^{\infty} \mathcal{D}(\mathcal{A}^{*n})$ , the function  $G^*(T_1 - \tau)w_0$  is infinitely differentiable in  $\tau$  and, therefore by the Taylor formula (see [23]) we get the followings:

$$(13) \quad G^*(T_1 - \tau)w_0 = \sum_{k=0}^{n-1} \frac{(T_1 - \tau)^k}{k!} \mathcal{A}^{*k} w_0 + \frac{1}{(n-1)!} \int_0^{T_1} (T_1 - \tau - t)^{n-1} G^*(t) \mathcal{A}^{*n} w_0 dt$$

$$(14) \quad G^*(T_1)w_0 = \sum_{k=0}^{n-1} \frac{T_1^k}{k!} \mathcal{A}^{*k} w_0 + \frac{1}{(n-1)!} \int_0^{T_1} (T_1 - t)^{n-1} G^*(t) \mathcal{A}^{*n} w_0 dt$$

( $n = 1, 2, 3, \dots$ )

From (11), (12), (13), (14) we obtain

$$\langle w_0, w \rangle = - \int_0^{T_1} \langle w_0, \bar{h}(x, \tau)\chi_\omega \rangle d\tau = - \int_0^{T_1} \langle cx + d, h\chi_\omega \rangle d\tau.$$

But for all  $h \in \Omega_T$  and for all  $T > 0$  we have

$$\langle cx + d, h\chi_\omega \rangle \geq 0,$$

thus it follows that

$$\langle w_0, w \rangle \leq 0 \quad \text{for all } w \in S_{T_1}(\Omega_{T_1}).$$

This contradicts the proposition  $0 \in \text{int } S_{T_1}(\Omega_{T_1})$ .

So the necessity is proved.

The proof of the Theorem is complete. ■

### References

- [1] J. L. LIONS, *Contrôle des systèmes distribués singulier*, Gauthier-Villars, MMI13, Paris 1983.
- [2] J. L. LIONS, Contrôlabilité exacte systèmes distribués singulier, *C. R. Acad. Sci. Paris*, **302** (1986), 471–475.
- [3] J. L. LIONS, Exact controllability, stabilization and perturbation for distributed systems, *SIAM Rev.*, **30** (1988), 1–58.
- [4] J. L. LIONS, *Contrôlabilité, exacte perturbations et stabilisation de systèmes distribués*, Tome 1. Contrôlabilité exacte, Masson, RMA8, Paris, 1988.
- [5] J. L. LIONS, *Contrôlabilité, exacte perturbations et stabilisation de systèmes distribués*, Tome 2. Contrôlabilité exacte, Masson, RMA8, Paris, 1988.
- [6] J. L. LIONS, and E. MAGENES, *Problèmes aux limites non homogènes et applications*, Vol. 1 and 2, Dunod, Paris, 1968.
- [7] A. HARAUX. Contrôlabilité exacte d'une membrane rectangulaire au moyen d'une fonctionnelle analytique localisée, *C. R. Acad. Sci. Paris*, **306** (1988), 125–128.
- [8] A. HARAUX, On a completion problem in the theory of distributed control of wave equations, *Nonlinear partial differential equations and their applications*, Collège de France seminar 1987–1988, H. Brézis and J. L. Lions editors, Pitman research notes in mathematics.
- [9] J. LAGNESE, Control of wave processes with distributed controls supported on a subregion, *SIAM J. Control Op.*, **21** (1983), 68–85.
- [10] E. ZUAZUA, An introduction to the exact controllability for distributed systems, *CMAF, TEXTOS E NOTAS*, **44** 1990.
- [11] I. JOÓ, On the reachability set of a string, *Acta Math. Acad. Sci. Hung.*, **49** (1987), 203–211.
- [12] M. HORVÁTH, Vibrating strings with free ends, *Acta Math. Acad. Sci. Hung.*, **51** (1988), 171–180.
- [13] A. BOGMÉR, A string equation with special boundary conditions, *Acta Math. Acad. Sci. Hung.*, **53** (1989), 367–376.
- [14] I. JOÓ, and N. V. SU, On the controllability of a string with restrained controls, (to appear).

- [15] I. JOÓ, and N. V. SU, On the approximative controllability of discrete-time systems and the application to the controllability of a circular membrane, (to appear).
- [16] N. K. SON, and N. V. SU, Linear periodic control system: Controllability with restrained controls, *Appl. Math. Optim.*, **14** (1986), 173–185.
- [17] N. V. SU, Null-controllability of finite dimensional discrete-time system with restrained control, *Prob. Contr. Infor. Theory*, **20** (1991), 215–232.
- [18] J. L. LIONS, *Optimal control systems governed by partial differential equations*, Springer–Verlag, Berlin, Heidelberg, New York, 1971.
- [19] B. SZÓKEFALVY-NAGY, *Spektraldarstellung linear Transformationen des Hilbertischen Raumes*, Ergebnisse der Mathematik, V/5, Springer–Verlag (Berlin, 1942).
- [20] R. G. DOUGLAS, On majorisation, factorization and range inclusion of operator in Hilbert space, *Proc. Amer. Math. Soc.*, **17** (1966), 409–416.
- [21] M. G. KREIN, and M. A. RUTMAN, Linear operator leaving invariant a cone in a Banach space, *Uspekhi Mat. Nauk. I*, **23** (1948), 3–9.
- [22] N. K. SON, Controllability of linear discrete-time with restrained controls in Banach space, *Contr. and Cyber.*, **10** (1981), 5–16.
- [23] E. HILLE, and R. PHILLIP, *Functionalanalysis and semigroup*, Amer. Math. Soc. Providence, R. I., 1958.
- [24] A PAZY, *Semigroup of linear operators and applications to partial differential equations*, Springer–Verlag, New York, Berlin, Heidelberg, Tokyo, 1980.



## AN ORTHOGONAL TYPE PROPERTY FOR THE JACOBI POLYNOMIALS

By

S. D. BAJPAI

Department of Mathematics, University of Bahrain, Isa Town, Bahrain, and  
Institute for Basic Research, Palm Harbor, U. S. A.

*(Received October 27, 1993)*

### 1. Introduction

The Jacobi polynomials are orthogonal polynomials ([2], p. 285, (5) and (9)) over the interval  $(-1,1)$  with respect to the weight function  $(1-x)^a(1+x)^b$ , if  $Rea > -1$ ,  $Reb > -1$ .

In this paper, we present an orthogonal type property for the Jacobi polynomials over the interval  $(-1,1)$  with respect to the multiplier function  $(1-x)^{a+1}(1+x)^b$ , if  $Rea > -2$ ,  $Reb > -1$ .

The Jacobi polynomials are defined by the relation ([2], p. 170, (16)):

$$(1.1) \quad P_n^{(a,b)}(x) = \frac{(1+a)_n}{n!} {}_2F_1\left(-n, n+a+b+1; 1+a; \frac{1-x}{2}\right),$$

provided  $a$  is not a negative integer.

### 2. Orthogonal type property

The orthogonal type property to be established is

$$(2.1) \quad \int_{-1}^1 (1-x)^{a+1}(1+x)^b P_m^{(a,b)}(x) P_n^{(a,b)}(x) dx = 0, \quad \text{if } m \neq n-1, n, n+1$$

$$(2.2) = -\frac{2^{a+b+2} \Gamma(a+n+1) \Gamma(b+n+1) \Gamma(a+b+2n-1)}{(n-1)! \Gamma(a+b+n) \Gamma(a+b+2n+2)}, \quad \text{if } m = n-1,$$

$$(2.3) \quad \frac{2^{a+b+2}\Gamma(a+n+1)\Gamma(b+n+1)}{n!\Gamma(a+b+n+1)(a+b+2n+1)} \left[ \frac{(n+1)(a+n+1)}{a+b+2n+2} - \frac{n(n+a)}{a+b+2n} \right],$$

if  $m = n$ ,

where  $\operatorname{Re} a > -2$ ,  $\operatorname{Re} b > -1$ .

PROOF. In view of (1.1), the integral (2.1) can be written as

$$(2.4) \quad \frac{(1+a)_m(1+a)_n}{m!n!} \int_{-1}^1 (1-x)^{a+1}(1+x)^b {}_2F_1 \left( -m, m+a+b+1; 1+a; \frac{1-x}{2} \right) \cdot {}_2F_1 \left( -n, n+a+b+1; 1+a; \frac{1-x}{2} \right) dx =$$

$$= \frac{(1+a)_m(1+a)_n}{m!n!} \sum_{r=0}^m \frac{(-m)_r(m+a+b+1)_r}{(1+a)_r r! 2^r} \sum_{u=0}^n \frac{(-n)_u(n+a+b+1)_u}{(1+a)_u u! 2^u} \int_{-1}^1 (1-x)^{a+r+u+1}(1+x)^b dx.$$

Evaluating the last integral in (2.4) with the help of a modified form of the integral ([1], p. 9, (5)), viz

$$\int_{-1}^1 (1-x)^a(1+x)^b dx = 2^{a+b+1} \frac{\Gamma(a+1)\Gamma(b+1)}{\Gamma(a+b+2)}, \quad \operatorname{Re} a > -1, \operatorname{Re} b > -1,$$

then the right hand side of (2.4) becomes

$$(2.5) \quad \frac{(1+a)_m(1+a)_n}{m!n!} \Gamma(b+1) 2^{a+b+2} \sum_{r=0}^m \frac{(-m)_r(m+a+b+1)_r}{(1+a)_r r!} \frac{\Gamma(a+r+2)}{\Gamma(a+b+r+3)} {}_3F_2 \left[ \begin{matrix} -n, n+a+b+1, 2+a+r; 1 \\ 1+a, 3+a+b+r \end{matrix} \right].$$

Now applying Saalschutz's theorem ([1], p. 188, (3)):

$$(2.6) \quad {}_3F_2 \left[ \begin{matrix} -n, a, b; 1 \\ c, 1-c+a+b-n \end{matrix} \right] = \frac{(c-a)_n(c-b)_n}{(c)_n(c-a-b)_n}$$

to (2.5), we get

$$(2.7) \quad \frac{(1+a)_m}{m!n!} \Gamma(b+1) 2^{a+b+2} \sum_{r=0}^m \frac{(-m)_r(-r-1)_n(m+a+b+1)_r \Gamma(a+r+2)(-n-b)_n}{(1+a)_r r! \Gamma(a+b+r+3)(-n-a-b-r-2)_n}.$$

If  $r < n - 1$ , the numerator of (2.7) vanishes and since  $r$  runs from 0 to  $m$ , it follows that (2.7) vanishes when  $m < n - 1$ . By symmetry, it will be seen that (2.7) also vanishes when  $n + 1 < m$ . Now, it is clear that for  $m \neq n - 1, n, n + 1$  all terms of the series (2.7) vanish, which proves (2.1).

When  $m = n - 1$ , using the result  $(-n)_n = (-1)^n n!$ , we have

$$(2.8) \quad \int_{-1}^1 (1-x)^{a+1} (1+x)^b P_{n-1}^{(a,b)}(x) P_n^{(a,b)}(x) dx = \\ = -\frac{2^{a+b+2} \Gamma(a+n+1) \Gamma(b+n+1)}{(n-1)! \Gamma(a+b+n) \Gamma(a+b+2n+2)},$$

which proves (2.2).

When  $m = n$ , using the standard results like  $(-n)_{n-1} = (-1)^{n-1} n!$  and  $(-n-1)_n = (-1)^n (n+1)!$  and adding the resulting two terms ( $r = n - 1, n$ ) and simplifying, we have

$$(2.9) \quad \int_{-1}^1 (1-x)^{a+1} (1+x)^b \{P_n^{(a,b)}\}^2 dx = \\ = \frac{2^{a+b+2} \Gamma(a+n+1) \Gamma(b+n+1)}{n! \Gamma(a+b+n+1) (a+b+2n+1)} \left[ \frac{(n+1)(a+n+1)}{a+b+2n+2} - \frac{n(n+a)}{a+b+2n} \right],$$

which proves (2.3).

NOTE. In (2.8) replacing  $n$  by  $n + 1$ , the value of the integral for  $m = n + 1$  may be obtained.

### References

- [1] ERDELYI, A. et. al., *Higher transcendental functions*, Vol. 1, McGraw-Hill, New York (1953).
- [2] ERDELYI, A. et. al., *Higher transcendental functions*, Vol. 2, McGraw-Hill, New York (1953).
- [3] ERDELYI, A. et. al., *Tables of integral transforms*, Vol. 2, McGraw-Hill, New York (1954).



## ON ADDITIVE FUNCTIONS SATISFYING SOME CONGRUENCE PROPERTIES II

By

K. KOVÁCS\*

Department for Algebra and Number Theory, Eötvös Loránd University, Budapest

(Received January 6, 1994)

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

In [1], we proved the following theorem:

**THEOREM 1.** *If  $f$  is a completely additive integer valued function, and, for a fixed  $m \in \mathbb{Z} \setminus \{0\}$ ,  $f(n+m) \equiv f(n) \pmod n$ , then  $f = 0$ .*

In [2], P. V. CHUNG generalized my result as follows:

**THEOREM 2.** *Let  $g$  denote an additive integer valued function and  $f$  a completely additive integer valued function. If for some fixed  $a \in \mathbb{N}$ ,  $b \in \mathbb{Z} \setminus \{0\}$  and  $c \in \mathbb{Z}$   $g(an+b) \equiv f(n)+c \pmod n$ , then  $f = 0$  and  $g(n) = 0$  for all  $(n,a) = 1$ .*

**COROLLARY OF THEOREM 2.** *If  $f$  is a completely additive integer valued function and  $f(an+b) \equiv f(n)+c \pmod n$ , then  $f = 0$ .*

It is possible to generalize this corollary further. However, a slight modification can cause difficulties. We are not able to verify whether  $f(n+1) \equiv f(n-1) \pmod n$  or  $f(n^2+1) \equiv f(n^2-1) \pmod n$  imply  $f = 0$ .

In this paper we shall prove the following special cases:

**THEOREM 3.** *Let  $f$  be a completely additive integer valued function. If any of the conditions (i)–(v) are satisfied, then  $f = 0$ .*

(i)  $f(n+1) \equiv kf(n-3) \pmod n$ ,  $k \in \mathbb{Z} \setminus \{1, -3\}$ ,

(ii)  $f(n+1)+f(n-1) \equiv kf(n) \pmod{n(n-1)}$ ,  $k \in \mathbb{Z} \setminus \{2\}$ ,

---

\* Supported by the Hungarian National Foundation for Scientific Research Grant No 1901.

$$(iii) f(n^2 + n + 1) \equiv kf(n) \pmod{n(n-1) \text{ or } n(n+1)}, \quad k \in \mathbb{Z} \setminus \{0\},$$

$$(iv) f(n^2 + n + 1) \equiv f(n-1) \pmod{n(n+1)},$$

$$(v) f(n^2 - n + 1) \equiv f(n-1) \pmod{n(n-1)}.$$

PROOF OF THEOREM 3. (i) Substituting  $n^2 + 2n$  in

$$(1) \quad f(n+1) \equiv kf(n-3) \pmod{n}$$

yields that

$$(2) \quad 2f(n+1) \equiv kf(n-1) + kf(n+3) \pmod{n(n+2)}.$$

Replacing  $n$  by  $n-2$  in (2) we have

$$(3) \quad 2f(n-1) \equiv kf(n-3) + kf(n+1) \pmod{n}.$$

Using (1) and (3) we get

$$(4) \quad 2f(n-1) \equiv (k+1)f(n+1) \pmod{n}.$$

Let us replace  $n$  by  $3n$  in (1), (2) and (4), giving us

$$(5) \quad f(3n+1) \equiv kf(3n-3) \pmod{n},$$

$$(6) \quad 2f(3n+1) \equiv kf(3n-1) + kf(3n+3) \pmod{n}$$

and

$$(7) \quad 2f(3n-1) \equiv (k+1)f(3n+1) \pmod{n}.$$

Multiplying (6) by 2 and using (7) we get

$$4f(3n+1) \equiv k(k+1)f(3n+1) + 2kf(3n+3) \pmod{n}.$$

Then by (5) we have

$$[4 - k(k+1)]f(3n-3) \equiv 2f(3n+3) \pmod{\frac{n}{(n,k)}}$$

for  $k \neq 0$ . Multiplying by 2 again and using (4) we get that, for some  $l \in \mathbb{Z}$ ,

$$k[4 - (k+1)^2]f(n+1) \equiv lf(3) \pmod{\frac{n}{(n,k)}}.$$

If  $k \notin \{0, 1, -3\}$ , let us replace  $n$  by  $n^s - 1$ . Then

$$\frac{n^s - 1}{(n^s - 1, k)} \mid k[4 - (k+1)^2]sf(n) - lf(3)$$

which implies  $f(n) = 0$  as  $s \rightarrow \infty$ .

If  $k = 0$ , then  $f(n+1) \equiv 0 \pmod{n}$  implies  $f = 0$  as in the last step of the previous proof.

(ii) Substituting  $n^2$  in

$$(8) \quad f(n+1) + f(n-1) \equiv kf(n) \pmod{n(n-1)}$$

yields

$$(9) \quad f(n^2+1) + f(n^2-1) \equiv kf(n^2) \pmod{n}.$$

Replacing (8) in (9) we get

$$(10) \quad f(n^2+1) \equiv kf(n) \pmod{n}.$$

Using the identity  $(n^2+1)(4n^4+1) = (2n^3+n)^2+1$ , (10) yields

$$(11) \quad kf(2n^2+1) \equiv kf(2n^2) \pmod{n}.$$

Replacing  $n$  by  $2n^2+1$  in (8) we have

$$f(2n^2+2) + f(2n^2) \equiv kf(2n^2+1) \pmod{n}.$$

Replacing (10) and (11) in this congruence, we get

$$(k-2)[f(n) + f(2)] \equiv 0 \pmod{n}.$$

The substitution  $n^s$  for  $n$  implies  $f(n) = 0$  as  $s \rightarrow \infty$  if  $k \neq 2$ .

(iii) For the proof of cases (iii)–(v) we need the following lemma which can be easily verified:

LEMMA. Let  $h(n) = n^2 + an + b$ . Then  $h[h(n) + n] = h(n)h(n+1)$ .

Let us replace  $n$  by  $n^2 + 2n + 1$  in

$$(12) \quad f(n^2 + n + 1) \equiv kf(n) \pmod{n(n+1)}.$$

By the lemma, we have

$$(13) \quad f(n^2 + n + 1) + f[(n+1)^2 + (n+1) + 1] \equiv 2kf(n+1) \pmod{(n+1)}.$$

$$(12) \text{ implies } f[(n+1)^2 + (n+1) + 1] \equiv kf(n+1) \pmod{(n+1)(n+2)}.$$

Substituting this congruence and (12) in (13) we get

$$kf(n) \equiv kf(n+1) \pmod{(n+1)},$$

i.e.

$$f(n-1) \equiv f(n) \pmod{\frac{n}{(n,k)}}.$$

Using a modified form of Theorem 2, we get  $f = 0$ .

If the modulus is  $n(n-1)$ , the proof is similar to the one above. We derive that  $f(n+1) \equiv f(n) \pmod{\frac{n}{(n,k)}}$  and then we again use Theorem 2.

(iv) Here, the proof is similar to the above case. Replacing  $n^2 + 2n + 1$  in

$$(14) \quad f(n^2 + n + 1) \equiv f(n-1) \pmod{n(n+1)},$$

we have

$$f(n-1) + f(n) \equiv f(n^2 + 2n) \pmod{n+1}$$

i.e.

$$(15) \quad f(n-1) \equiv f(n+2) \pmod{n+1}.$$

Replacing  $n$  by  $n^2 + n - 1$  in (15), we get by (14) that

$$f(n+2) \equiv 0 \pmod{n}$$

which yields  $f = 0$ .

(v) Replacing  $n$  by  $n^2 + 1$  in

$$(16) \quad f(n^2 - n + 1) \equiv f(n-1) \pmod{n(n-1)},$$

we derive

$$f(n-1) \equiv f(n) \pmod{n}$$

through similar use of the lemma as in case (iii). Therefore,  $f = 0$  using Theorem 2.

### References

- [1] K. KOVÁCS, On additive functions satisfying some congruence properties, *Periodica Math. Hung.*, **23** (1991), 227–231.
- [2] P. V. CHUNG, Note on additive functions satisfying some congruence property III (manuscript).

## A NEW MARTINGALE EQUIVALENCE THEOREM

By

F. WEISZ

Department of Numerical Analysis, Eötvös Loránd University, Budapest.

*(Received January 31, 1994)*

*(Revised August 10, 1994)*

### 1. Introduction

The equivalence of the martingale Hardy spaces  $H_p^*$  and  $H_p^{[\ ]}$  ( $1 \leq p < \infty$ ) generated by the maximal function and by the quadratic variation, respectively, was proved by BURKHOLDER, DAVIS and GUNDY [2], [3], [4] for discrete time and later by DELLACHERIE and MEYER [6] for continuous time. This theorem is one of the fundamental theorems of martingale theory.

Usually three other martingale Hardy spaces are investigated: the space  $H_p^{\langle \rangle}$  generated by the conditional quadratic variation and the spaces  $\mathcal{P}_p$  and  $\mathcal{Q}_p$  generated by the predictions of the maximal function and of the quadratic variation, respectively. With the help of the atomic decomposition several martingale inequalities between these five Hardy spaces and a simple proof of the Burkholder–Davis–Gundy inequality are given in WEISZ [13]. As a supplement to these results it is verified in this paper that  $\mathcal{P}_p$  is equivalent to  $\mathcal{Q}_p$  for  $0 < p < \infty$ . This equivalence result was shown in the discrete time with another method in WEISZ [11].

### 2. Preliminaries and Notations

Let  $(\Omega, \mathcal{A}, P)$  be a probability measure space and  $\mathcal{F} = (\mathcal{F}_t, t \in \mathbb{R}^+)$  be a nondecreasing family of sub- $\sigma$ -algebras of  $\mathcal{A}$ . The  $\sigma$ -algebra  $\bigvee_{t \in \mathbb{R}^+} \mathcal{F}_t$  is denoted by  $\mathcal{F}_\infty$  and it is supposed that  $\mathcal{F}_\infty = \mathcal{A}$ .

---

This research was partly supported by the Hungarian Scientific Research Funds No F4189.

With the family  $(\mathcal{F}_t, t \in \mathbb{R}^+)$  the following families  $(\mathcal{F}_{t+}, t \in \mathbb{R}^+)$  and  $(\mathcal{F}_{t-}, t \in \mathbb{R}^+)$  of  $\sigma$ -algebras are associated:

$$\mathcal{F}_{t+} := \bigcap_{s > t} \mathcal{F}_s, \quad \mathcal{F}_{t-} := \bigvee_{s < t} \mathcal{F}_s.$$

For  $t = 0$  we set  $\mathcal{F}_{0-} := \mathcal{F}_0$ . In this paper it is assumed that the family  $(\mathcal{F}_t, t \in \mathbb{R}^+)$  is *right-continuous* (i.e.  $\mathcal{F}_t = \mathcal{F}_{t+}$  for every  $t$ ) and that every set  $F$  which belongs to the  $P$ -completion of the  $\sigma$ -algebra  $\mathcal{F}_\infty$  with  $P(F) = 0$  belongs to  $\mathcal{F}_0$ .

A real *stochastic process*  $X$  is a mapping from  $(\mathbb{R}^+ \times \Omega)$  into  $\mathbb{R}$  such that, for every  $t \in \mathbb{R}^+$ ,  $\omega \mapsto X_t(\omega) = X(t, \omega)$  is  $\mathcal{A}$ -measurable. A stochastic process  $X$  is *adapted* if the preceding mapping is  $\mathcal{F}_t$ -measurable for every  $t$ . A process  $X$  is *regular* if  $X$  is adapted and all the functions  $t \mapsto X(t, \omega)$  have left and right limit for every  $\omega \in \Omega$ . We use the notation  $X_{t-}$  ( $X_{t+}$ , resp.) for the left (right, resp.) limit at a point  $t$ . If these functions are right-continuous (left-continuous, resp.) then  $X$  is called *right-continuous* (*left-continuous*, resp.).

Let us denote the expectation operator by  $E$ , the conditional expectation operators relative to  $\mathcal{F}_t$  and  $\mathcal{F}_{t-}$  by  $E_t$  and  $E_{t-}$ , respectively. Let us use the shorter notation  $L_p$  for the subspace of  $L_p(\Omega, \mathcal{A}, P)$  the elements of which satisfy  $E_0 f = 0$  ( $1 \leq p \leq \infty$ ).

A stochastic process  $X$  is a *martingale* when  $X$  is an adapted process with  $E|X_t| < \infty$  ( $t \in \mathbb{R}^+$ ) and  $E_s X_t = X_s$  for every  $s < t$ . For simplicity we always suppose for a martingale  $X$  that  $X_0 = 0$ . A martingale  $X$  is said to be  *$L_p$ -bounded* if

$$\sup_{t \in \mathbb{R}^+} \|X_t\|_p < \infty.$$

A mapping  $T : \Omega \rightarrow \mathbb{R}^+ \cup \{\infty\}$  is a *stopping time* if, for every  $t \in \mathbb{R}^+$ , the subset  $\{T \leq t\}$  of  $\Omega$  belongs to  $\mathcal{F}_t$ . An adapted process  $X$  is a *local martingale* if there exists an increasing sequence of stopping times  $T_n$  such that  $\lim_{n \rightarrow \infty} T_n = \infty$  a.e. and the processes  $X_{t \wedge T_n}$  are all uniformly integrable martingales ( $n \in \mathbb{N}$ ). We can suppose from a martingale or from a local martingale that it is regular and right-continuous (see MEYER [9] p. 291). A local martingale is said to be *locally  $L_p$ -bounded* if there exists an increasing sequence of stopping times  $T_n$  such that  $\lim_{n \rightarrow \infty} T_n = \infty$  a.e. and the processes  $X_{t \wedge T_n}$  are  $L_p$ -bounded martingales ( $n \in \mathbb{N}$ ).

The following definitions will be used for a local martingale  $X$ :

$$\begin{aligned} \Delta X_t &:= X_t - X_{t-}, & (X_0 = 0), \\ X_s^* &:= \sup_{t \leq s} |X_t|, & X_\infty^* := \sup_{t \in \mathbb{R}^+} |X_t|. \end{aligned}$$

We say that a real process is *optional* resp. *predictable* if it is measurable with respect to the  $\sigma$ -algebra generated by the real regular right-continuous process resp. generated by the real regular left-continuous processes.

If  $X$  is a locally  $L_2$ -bounded local martingale then there exists a unique predictable, right-continuous and increasing process  $\langle X \rangle$  such that  $X^2 - \langle X \rangle$  is a local martingale vanishing at 0 (see DELLACHERIE and MEYER [6]).  $\langle X \rangle$  is called the *sharp bracket* or the *conditional quadratic variation* of  $X$ . Moreover, if  $X$  is a local martingale then there exists a unique right-continuous and increasing process  $[X]$  such that  $X^2 - [X]$  is a local martingale and  $\Delta[X]_t = |\Delta X_t|^2$  ( $[X]_0 = 0$ ). This process is called the *square bracket* or the *quadratic variation* of  $X$ .

Let us introduce *Hardy spaces* for  $0 < p \leq \infty$ ; denote by  $H_p^{[\ ]}$ ,  $H_p^{(\ )}$  and  $H_p^*$  the spaces of local martingales for which

$$\|X\|_{H_p^{[\ ]}} := \|[X]_\infty^{1/2}\|_p < \infty,$$

$$\|X\|_{H_p^{(\ )}} := \|\langle X \rangle_\infty^{1/2}\|_p < \infty$$

and

$$\|X\|_{H_p^*} := \|X_\infty^*\|_p < \infty,$$

respectively. It is a well-known statement in the martingale theory that if  $X \in H_p^*$ ,  $H_p^{[\ ]}$ ,  $H_p^{(\ )}$  ( $p \geq 1$ ) then there exists  $X_\infty$  such that  $X_t \rightarrow X_\infty$  a.e. and in  $L_1$  as  $t \rightarrow \infty$ . Moreover, Doob's inequality says that  $H_p^* \sim L_p$  for  $1 < p \leq \infty$  where  $\sim$  denotes the equivalence of the spaces (see e.g. DELLACHERIE and MEYER [6]).

A local martingale  $X$  is in the space  $\mathcal{P}_p$  if and only if there exists an adapted, left-continuous and increasing process  $A$  such that

$$|X_t| \leq A_t, \quad A_\infty := \sup_{t \in \mathbb{R}^+} A_t \in L_p.$$

Endow this space with the following norm:

$$\|X\|_{\mathcal{P}_p} := \inf \|A_\infty\|_p \quad (0 < p \leq \infty)$$

where the infimum is taken over all predictable processes having the property above.

If, in the previous definition, we replace the inequality  $|X_t| \leq A_t$  by

$$[X]_t^{1/2} \leq B_t$$

then the local martingale is in  $\mathcal{Q}_p$ . We define the  $\mathcal{Q}_p$  norm by

$$\|X\|_{\mathcal{Q}_p} := \inf \|B_\infty\|_p \quad (0 < p \leq \infty)$$

where the infimum is taken over all predictable processes again.

It is clear that the infimums taken in the  $\mathcal{P}_p$  and  $\mathcal{Q}_p$  norms can be achieved. Indeed, for every  $k \in \mathbb{N}$  let  $A^{(k)}$  be a left-continuous process having the above property such that  $\|A_\infty^{(k)}\|_p \rightarrow \|X\|_{\mathcal{P}_p}$  whenever  $k \rightarrow \infty$ . Set

$$A_t := \inf_{k \in \mathbb{N}} A_t^{(k)} \quad (t \in \mathbb{R}^+).$$

It is obvious that the process  $A$  is adapted, left-continuous, increasing and a majorant of  $X$  and

$$\|X\|_{\mathcal{P}_p} = \|A_\infty\|_p.$$

The proof is similar for the  $\mathcal{Q}_p$  spaces.

The  $\mathcal{P}_p$  resp.  $\mathcal{Q}_p$  spaces were introduced by GARSIA [7] and BERNARD, MAISONNEUVE [1] resp. by the author in [13]. With the help of these two spaces a simple proof of Burkholder–Davis–Gundy inequality is given in [13].

### 3. Martingale inequalities

The following theorem is proved in WEISZ [13].

THEOREM 1.

- (i)  $\|X\|_{H_p^*} \leq C_p \|X\|_{H_p^{\langle \cdot \rangle}}, \quad \|X\|_{H_p^{\square}} \leq C_p \|X\|_{H_p^{\langle \cdot \rangle}} \quad (0 < p \leq 2)$
- (ii)  $\|X\|_{H_p^{\langle \cdot \rangle}} \leq C_p \|X\|_{H_p^*}, \quad \|X\|_{H_p^{\langle \cdot \rangle}} \leq C_p \|X\|_{H_p^{\square}} \quad (2 \leq p < \infty)$
- (iii)  $\|X\|_{H_p^*} \leq \|X\|_{\mathcal{P}_p}, \quad \|X\|_{H_p^{\square}} \leq \|X\|_{\mathcal{Q}_p} \quad (0 < p < \infty)$
- (iv)  $\|X\|_{H_p^*} \leq C_p \|X\|_{\mathcal{Q}_p}, \quad \|X\|_{H_p^{\square}} \leq C_p \|X\|_{\mathcal{P}_p} \quad (0 < p < \infty)$
- (v)  $\|X\|_{H_p^{\langle \cdot \rangle}} \leq C_p \|X\|_{\mathcal{P}_p}, \quad \|X\|_{H_p^{\langle \cdot \rangle}} \leq C_p \|X\|_{\mathcal{Q}_p} \quad (0 < p < \infty),$

where the positive constants  $C_p$  depend only on  $p$ . (The symbol  $C_p$  may denote different constants in different contexts.)

The inequality due to BURKHOLDER, DAVIS and GUNDY [2], [3], [4] mentioned several times is given in the next theorem. They proved it in the discrete case. Later DELLACHERIE and MEYER [6] have extended the inequality to the continuous case (see also WEISZ [13]).

**THEOREM 2.** (BURKHOLDER–DAVIS–GUNDY INEQUALITY).  $H_p^{[\cdot]}$  and  $H_p^*$  are equivalent for  $1 \leq p < \infty$ , namely,

$$c_p \|X\|_{H_p^{[\cdot]}} \leq \|X\|_{H_p^*} \leq C_p \|X\|_{H_p^{[\cdot]}} \quad (1 \leq p < \infty).$$

This inequality is not valid for  $0 < p < 1$  (see BURKHOLDER, GUNDY [3]). However, the  $\mathcal{P}_p$  and  $\mathcal{Q}_p$  spaces that are generated by the predictions of the maximal function and of the quadratic variation, respectively, are equivalent for every  $0 < p < \infty$ . Observe that  $\mathcal{P}_\infty \sim H_\infty^* \sim L_\infty$  and  $\mathcal{Q}_\infty \sim H_\infty^{[\cdot]}$ . Hence  $\mathcal{P}_\infty$  is not equivalent to  $\mathcal{Q}_\infty$ .

**THEOREM 3.**  $\mathcal{P}_p$  and  $\mathcal{Q}_p$  are equivalent for  $0 < p < \infty$ , namely,

$$(1) \quad c_p \|X\|_{\mathcal{Q}_p} \leq \|X\|_{\mathcal{P}_p} \leq C_p \|X\|_{\mathcal{Q}_p} \quad (0 < p < \infty).$$

**PROOF.** First we prove the left hand side of (1). Let  $X \in \mathcal{P}_p$  and  $A$  be the adapted, left-continuous and increasing least majorant process of  $X$ . It is proved in DELLACHERIE, MEYER [6] that

$$(2) \quad [X]_t = X_t^2 - 2 \int_0^t X_s - dX_s$$

where the second term on the right hand side is a martingale defined by a stochastic integral. By the definitions of  $A$  and of the stochastic integral we obtain

$$(3) \quad \begin{aligned} [X]_t &\leq A_t^2 + 2 \sup_{u \leq t} \left| \int_0^u X_s - dX_s \right| \leq \\ &\leq A_t^2 + 2 \sup \left\{ \sup_{u < t} \left| \int_0^u X_s - dX_s \right|, \left| \int_0^t X_s - dX_s \right| \right\} \leq \\ &\leq A_t^2 + 2 \sup_{u < t} \left| \int_0^u X_s - dX_s \right| + 2 \left| \Delta \left( \int_0^{\cdot} X_s - dX_s \right)_t \right|. \end{aligned}$$

Since

$$(4) \quad \Delta \left( \int_0^{\cdot} X_s - dX_s \right)_t = X_t - \Delta X_t$$

(see DELLACHERIE, MEYER [6]) and  $|X_t - \Delta X_t| \leq 2A_t^2$ , we can establish that

$$[X]_t \leq 5A_t^2 + 2 \sup_{u < t} \left| \int_0^u X_s - dX_s \right| =: B_t^2.$$

Notice that  $B$  is adapted, left-continuous and increasing. Henceforth

$$\|B_\infty\|_p \leq c_p \|A_\infty\|_p + c_p \left\| \sup_{t \in \mathbb{R}^+} \left| \int_0^t X_s - dX_s \right| \right\|_{p/2}^{1/2}.$$

Assume that  $0 < p \leq 4$ . The sharp bracket of the martingale  $\int_0^\cdot X_s - dX_s$  is given by  $\int_0^\cdot |X_s -|^2 d\langle X \rangle_s$  (see DELLACHERIE, MEYER [6]). Applying Theorem 1 (i) we can conclude that

$$\begin{aligned} (5) \quad \left\| \sup_{t \in \mathbb{R}^+} \left| \int_0^t X_s - dX_s \right| \right\|_{p/2}^{1/2} &\leq \left\| \left( \int_0^\infty |X_s -|^2 d\langle X \rangle_s \right)^{1/2} \right\|_{p/2}^{1/2} \leq \\ &\leq \left\| \left( \int_0^\infty A_s^2 d\langle X \rangle_s \right)^{1/2} \right\|_{p/2}^{1/2} \leq \|A_\infty \langle X \rangle_\infty^{1/2}\|_{p/2}^{1/2}. \end{aligned}$$

From Hölder's inequality

$$(6) \quad [E(A_\infty^{p/2} \langle X \rangle_\infty^{p/4})]^{1/p} \leq E(A_\infty^p)^{1/(2p)} E(\langle X \rangle_\infty^{p/2})^{1/(2p)}.$$

Consequently, using Theorem 1 (v) we obtain

$$\begin{aligned} \|B_\infty\|_p &\leq c_p \|A_\infty\|_p + c_p \|A_\infty\|_p^{1/2} \|X\|_{H_p^{\langle \cdot \rangle}}^{1/2} \\ &\leq c_p \|X\|_{\mathcal{F}_p} + c_p \|X\|_{\mathcal{F}_p}^{1/2} \|X\|_{\mathcal{F}_p}^{1/2} \leq c_p \|X\|_{\mathcal{F}_p}. \end{aligned}$$

The left hand side of (1) can be proved in the same way for  $2 \leq p < \infty$  with the differences that we have to apply Theorem 2 instead of Theorem 1 (i) and we have to change the sharp bracket to the square bracket.

To prove the right hand side of (1) let  $X \in \mathcal{Q}_p$  and  $B$  be the adapted, left-continuous and increasing least majorant process of  $[X]^{1/2}$ . Similarly to

(3), it follows from (2) that

$$X_t^2 \leq B_t^2 + 2 \sup_{u < t} \left| \int_0^u X_s - dX_s \right| + 2 \left| \Delta \left( \int_0^{\cdot} X_s - dX_s \right) \right|_t.$$

Taking into account (4) and the fact that

$$|X_t - \Delta X_t| \leq X_{t-}^* - [X]_t^{1/2} \leq X_{t-}^* - B_t,$$

we have

$$X_t^2 \leq B_t^2 + 2X_{t-}^* - B_t + 2 \sup_{u < t} \left| \int_0^u X_s - dX_s \right| =: A_t^2.$$

Of course,  $A$  is adapted, left-continuous and increasing. So

$$\|A_\infty\|_p \leq C_p \|B_\infty\|_p + C_p E(X_\infty^{*p/2} B_\infty^{p/2})^{1/p} + C_p \left\| \sup_{t \in \mathbb{R}^+} \left| \int_0^t X_s - dX_s \right| \right\|_{p/2}^{1/2}.$$

If  $0 < p \leq 4$  then, similarly to (5) and (6), we get that

$$\begin{aligned} \left\| \sup_{t \in \mathbb{R}^+} \left| \int_0^t X_s - dX_s \right| \right\|_{p/2}^{1/2} &\leq \|X_\infty^* \langle X \rangle_\infty^{1/2}\|_{p/2}^{1/2} \leq \\ &\leq E(X_\infty^{*p})^{1/(2p)} E(\langle X \rangle_\infty^{p/2})^{1/(2p)}. \end{aligned}$$

Applying Hölder’s inequality and Theorem 1 we can see that

$$\begin{aligned} \|A_\infty\|_p &\leq \\ &\leq C_p \|B_\infty\|_p + C_p \|X\|_{H_p^*}^{1/2} \|B_\infty\|_p^{1/2} + C_p \|X\|_{H_p^*}^{1/2} \|X\|_{H_p^{\langle \cdot \rangle}}^{1/2} \leq C_p \|X\|_{\mathcal{Q}_p}. \end{aligned}$$

The case  $2 \leq p < \infty$  can be shown analogously. ■

### References

- [1] A. BERNARD, B. MAISONNEUVE, *Decomposition automatique de martingales de la classe  $H^1$* , Seminaire de Probabilités XI. (Lect. Notes Math., vol. 581, pp. 303–323) Berlin, Heidelberg, New York, Springer, 1977.
- [2] D.L. BURKHOLDER, Distribution function inequalities for martingales, *Annals of Prob.*, 1 (1973), 19–42.

- [3] D.L. BURKHOLDER, R.F. GUNDY, Extrapolation and interpolation of quasi-linear operators on martingales, *Acta Math.*, **124** (1970), 249–304.
- [4] B.J. DAVIS, On the integrability of the martingale square function. *Israel J. Math.*, **8**, (1970), 187–190.
- [5] C. DELLACHERIE, P.-A. MEYER, *Probabilities and potential A.*, North-Holland Math. Studies 29, North-Holland, 1978.
- [6] C. DELLACHERIE, P.-A. MEYER, *Probabilities and potential B.*, North-Holland Math. Studies 72, North-Holland, 1982.
- [7] A.M. GARSIA, *Martingale inequalities*, Seminar Notes on Recent Progress. Math. Lecture notes series, New York, Benjamin Inc., 1973.
- [8] M. METIVIER, *Semimartingales, a course on stochastic processes*. de Gruyter Studies in Math. 2, Walter de Gruyter, 1982.
- [9] P.-A. MEYER, *Un cours les integrales stochastiques*, Seminaire de Probabilités X. (Lect. Notes Math., vol. 511, pp. 245–400) Berlin, Heidelberg, New York, Springer, 1976.
- [10] F. WEISZ, Martingale Hardy spaces for  $0 < p \leq 1$ , *Probab. Th. Rel. Fields*, **84** (1990), 361–376.
- [11] F. WEISZ, Hardy spaces of predictable martingales, *Analysis Math.*, **20** (1994), 225–233.
- [12] F. WEISZ, Martingale Hardy spaces and their applications in Fourier-analysis. Lecture Notes in Math. vol. 1568, Berlin, Heidelberg, New York, Springer, 1994.
- [13] F. WEISZ, Martingale Hardy spaces with continuous time, *Probability Theory and Applications*, Essays to the Memory of József Mogyoródi. Eds.: J. Galambos, I. Kátai. Kluwer Academic Publishers, Dordrecht, Boston, London, 1992, 47–75.

## ON UNIQUENESS OF EXTENSION OF SUBOPERATORS

By

V. PROKAJ

Department of Applied Analysis, Eötvös Loránd University, Budapest

*(Received April 22, 1994)*

A suboperator is a restriction of a bounded operator on a Hilbert space  $\mathfrak{H}$  to a linear subspace of  $\mathfrak{H}$ . It is subpositive, subself-adjoint etc. if it is the restriction of a positive, self-adjoint etc. operator, respectively. Intrinsic characterizations of the subpositive suboperator were given by HALMOS [3] and by SEBESTYÉN in [7], [8]. Their construction lead to the smallest positive extension (see e.g. [8], [9]).

In this paper we consider the following problem: under what condition does there exist only one positive extension of a subpositive suboperator with the smallest possible norm. This problem was successfully handled by KAPOŠ and PAEZ [4] and previously in the subself-adjoint case by KREIN [5]. None of these conditions provide an easy way of constructing examples or counterexamples. I hope Corollary 1 is a step toward this direction. We need some detail of Sebestyén's construction ([10]), so for the sake of convenience we repeat the proof of this theorem in a slightly modified way.

**THEOREM 1.** *Let  $\mathfrak{D}$  be a linear subspace (not necessarily closed) of a complex Hilbert space  $\mathfrak{H}$ , and  $b : \mathfrak{D} \rightarrow \mathfrak{H}$  be a symmetric linear map. Then  $b$  is subpositive if and only if*

$$(1) \quad m := \sup \left\{ \|bx\|^2 : x \in \mathfrak{D}, \langle bx, x \rangle \leq 1 \right\} < \infty.$$

**PROOF.** Denote  $\mathfrak{R}$  the range of  $b$ . Introduce an inner product  $\langle \cdot, \cdot \rangle_{\mathfrak{R}}$  on  $\mathfrak{R}$  by definition

$$\langle bx, by \rangle_{\mathfrak{R}} := \langle bx, y \rangle \quad \text{for } x, y \in \mathfrak{D}.$$

It is well defined: if  $x, x', y, y' \in \mathfrak{D}$  and  $bx = bx', by = by'$  then

$$\langle bx, by \rangle_{\mathfrak{R}} = \langle bx, y \rangle = \langle bx', y \rangle = \langle x', by' \rangle = \langle bx', by' \rangle_{\mathfrak{R}}.$$

Finally,  $\langle \cdot, \cdot \rangle_{\mathfrak{R}}$  is positive definite since  $\|bx\|^2 \leq m\langle bx, x \rangle$  for all  $x \in \mathfrak{D}$ .

Thus  $(\mathfrak{R}, \langle \cdot, \cdot \rangle_{\mathfrak{R}})$  is a pre-Hilbert space,  $\mathfrak{H}_0$  denotes its completion. Define  $J : \mathfrak{H}_0 \rightarrow \mathfrak{H}$  as the unique continuous extension to the natural embedding of  $\mathfrak{R}(\subset \mathfrak{H}_0)$  into  $\mathfrak{H}$ . Then  $JJ^*$  is an extension of  $b$ . To prove this it is enough to show that for all  $x \in \mathfrak{D}$  the element  $bx - J^*x$  of  $\mathfrak{H}_0$  is orthogonal to the dense subspace  $\mathfrak{R}$  of  $\mathfrak{H}_0$ . But this is obvious, since

$$\langle by, J^*x \rangle_{\mathfrak{R}} = \langle Jby, x \rangle = \langle by, x \rangle = \langle by, bx \rangle_{\mathfrak{R}} \quad \text{for all } y \in \mathfrak{D}. \quad \blacksquare$$

The above extension  $JJ^*$  is proved to be the smallest positive extension of  $b$  ([9], Thm. 1). The norm of  $JJ^*$  is just  $m$  because  $\|J\|^2 = m$  by definition. The constant  $m$  defined in (1) is thus the smallest possible norm of the positive extensions for  $b$ .

There is a largest positive extension for  $b$  of smallest norm  $m$ . One can get it by replacing  $b$  with  $b_1 = mI|_{\mathfrak{D}} - b$  in the previous theorem, since for all  $x \in \mathfrak{D}$

$$\begin{aligned} \|b_1\|^2 &= m^2\|x\|^2 - 2m\langle bx, b \rangle + \|bx\|^2 = \\ &= m\langle b_1x, x \rangle + \|bx\|^2 - m\langle bx, x \rangle \leq m\langle b_1x, x \rangle. \end{aligned}$$

Let  $JJ^*$  be the smallest positive extension of  $b_1$ . The largest positive extension of  $b$  of smallest possible norm  $m$  is then  $mI - J_1J_1^*$ .

$b$  has unique positive extension with norm  $m$  if and only if the smallest and largest ones are equal or what is the same

$$(2) \quad mI = JJ^* + J_1J_1^*.$$

This condition was formulated by Kapos and Paez in the following theorem ([4], Thm. 1).

**THEOREM 2.** *Let  $b : \mathfrak{D} \rightarrow \mathfrak{H}$  be subpositive suboperator satisfying property (1).  $b$  has unique positive operator extension of smallest norm  $m$  to  $\mathfrak{H}$  if and only if*

$$\begin{aligned} m\|x\|^2 &= \sup \left\{ |\langle x, by \rangle|^2 \mid y \in \mathfrak{D}, \langle by, y \rangle \leq 1 \right\} + \\ &+ \sup \left\{ |\langle x, my - by \rangle|^2 : y \in \mathfrak{D}, \langle my - by, y \rangle \leq 1 \right\} \end{aligned}$$

*holds for each  $x$  from  $\mathfrak{H}$ .*

Identity (2) can be formulated in an other way that involves  $x$  only from  $\mathfrak{D}^\perp$ , the orthocomplement of  $\mathfrak{D}$ , and does contain only one expression on supremum.

**THEOREM 3.** *Let  $b : \mathfrak{D} \rightarrow \mathfrak{H}$  be subpositive suboperator satisfying property (1).  $b$  has unique positive operator extension of smallest norm  $m$  to  $\mathfrak{H}$  if and only if*

$$(3) \quad \|x\|^2 = \sup \left\{ |\langle x, by \rangle|^2 : y \in \mathfrak{D}, \langle by, my - Pby \rangle \leq 1 \right\}$$

*holds for each  $x$  from  $\mathfrak{D}^\perp$ , where  $P$  denotes the orthogonal projection of  $\mathfrak{H}$  onto the norm closure of  $\mathfrak{D}$ .*

**PROOF.** First we show that identity (2) is equivalent with the formally weaker

$$(2a) \quad m\langle x, x \rangle = \langle JJ^*x, x \rangle + \langle J_1J_1^*x, x \rangle \quad \text{for all } x \in \mathfrak{D}^\perp$$

and after this that this weaker condition (2a) is equivalent to (3).

To do this, let us introduce some notation: let  $\mathfrak{R}$  and  $(\mathfrak{H}_0, \langle, \rangle_{\mathfrak{R}})$  be as in the proof of Theorem 1, and  $\mathfrak{R}_1$  and  $(\mathfrak{H}_1, \langle, \rangle_1)$  be the corresponding entities when Theorem 1 is applied to  $b_1 = mI|_{\mathfrak{D}} - b$ . Denote  $(\mathfrak{K}, \langle, \rangle_{\oplus})$  the direct sum of  $\mathfrak{H}_0$  and  $\mathfrak{H}_1$ , and define  $V : \mathfrak{K} \rightarrow \mathfrak{H}$  by the formula

$$V(u, v) = Ju + J_1v \quad \text{for } u \in \mathfrak{H}_0 \text{ and } v \in \mathfrak{H}_1.$$

Since  $V^* : \mathfrak{H} \rightarrow \mathfrak{K}$  is the 'row matrix'  $(J^*, J_1^*)$ ,  $VV^* = JJ^* + J_1J_1^*$ .

For the first step we only have to prove that (2a) implies (2). Notice first that it follows from the definition of  $\mathfrak{K}$  and  $V$  that the image of  $\mathfrak{D}$  and  $\mathfrak{D}^\perp$  under  $V^*$  are orthogonal linear subspaces of  $\mathfrak{K}$ . Indeed, for  $x \in \mathfrak{D}$  and  $y \in \mathfrak{D}^\perp$  we have that

$$\langle V^*x, V^*y \rangle_{\oplus} = \langle (JJ^* + J_1J_1^*)x, y \rangle = \langle (b + mI - b)x, y \rangle = m\langle x, y \rangle = 0.$$

On the other hand identity (2) restricted to the norm closure of  $\mathfrak{D}$  holds by definitions and continuity. Thus if (2a) holds then decomposing any  $x \in \mathfrak{H}$  into  $x = y + z$ , where  $y \in \overline{\mathfrak{D}}$  and  $z \in \mathfrak{D}^\perp$  we get that

$$\langle (JJ^* + J_1J_1^*)x, x \rangle = \|V^*x\|^2 = \|V^*y\|^2 + \|V^*z\|^2.$$

Where the last term is equal to  $m\|z\|^2$  by (2a) and the previous one to  $m\|y\|^2$  as we saw above, so assuming (2a) one gets

$$\langle (JJ^* + J_1J_1^*)x, x \rangle = m\|x\|^2$$

that of course implies (2).

For the second step we have to prove that (2a) is equivalent to (3). To do that let us calculate the norm  $\|V^*x\|$  ( $x \in \mathfrak{D}^\perp$ ) as the norm of the linear functional that  $V^*x$  represents:

$$\|V^*x\|^2 = \sup \left\{ |V^*x, \xi|_{\oplus}^2 : \xi \in \mathfrak{K}, \langle \xi, \xi \rangle_{\oplus} \leq 1 \right\} =$$

$$\begin{aligned}
&= \sup \left\{ |\langle x, V\xi \rangle|^2 : \xi \in \mathfrak{K} \oplus \mathfrak{K}_1, \langle \xi, \xi \rangle_{\oplus} \leq 1 \right\} = \\
&= \sup \left\{ |\langle x, bu - bv \rangle|^2 : u, v \in \mathfrak{D}, \langle bu, u \rangle + \langle mv - bv, v \rangle \leq 1 \right\}.
\end{aligned}$$

The last condition can be then rewritten with  $\xi = u - v$ ,  $\eta = u + v$  by observing that

$$\begin{aligned}
&m \{ \langle bu, u \rangle + \langle mv - bv, v \rangle \} = \\
&= \frac{m}{4} \{ \langle b(\xi + \eta), \xi + \eta \rangle - \langle b(\eta - \xi), \eta - \xi \rangle + m \langle \eta - \xi, \eta - \xi \rangle \} = \\
&= \left\langle b\xi, \frac{m}{2}\eta \right\rangle + \left\langle \frac{m}{2}\eta, b\xi \right\rangle + \left\| \frac{m}{2}\xi \right\|^2 + \left\| \frac{m}{2}\eta \right\|^2 - \left\langle \frac{m}{2}\xi, \frac{m}{2}\eta \right\rangle - \left\langle \frac{m}{2}\eta, \frac{m}{2}\xi \right\rangle = \\
&= \left\| \frac{m}{2}\xi \right\|^2 + \left\| \frac{m}{2}\eta \right\|^2 + \left\langle \left( b - \frac{m}{2}I \right) \xi, \frac{m}{2}\eta \right\rangle + \left\langle \frac{m}{2}\eta, \left( b - \frac{m}{2}I \right) \xi \right\rangle = \\
&= \left\| \frac{m}{2}\xi \right\|^2 + \left\| \frac{m}{2}\eta + \left( b - \frac{m}{2}I \right) \xi \right\|^2 - \left\| \left( b - \frac{m}{2}I \right) \xi \right\|^2.
\end{aligned}$$

The infimum of  $m \{ \langle bu, u \rangle + \langle mv - bv, v \rangle \}$  where  $u, v \in \mathfrak{D}$  and  $\xi = u - v$  is fixed is therefore equal to

$$\begin{aligned}
&\left\| \frac{m}{2}\xi \right\|^2 - \left\| P \left( b - \frac{m}{2}I \right) \xi \right\|^2 = \\
&= \left\| \frac{m}{2}\xi \right\|^2 - \left\| \frac{m}{2}\xi \right\|^2 - \|Pb\xi\|^2 + \left\langle Pb\xi, \frac{m}{2}\xi \right\rangle + \left\langle \frac{m}{2}\xi, Pb\xi \right\rangle = \\
&= m \langle \xi, b\xi \rangle - \|Pb\xi\|^2 = \langle b\xi, m\xi - Pb\xi \rangle,
\end{aligned}$$

where  $P$  is the orthogonal projection of  $\mathfrak{H}$  to the norm closure of  $\mathfrak{D}$ . We get finally that

$$m^{-1} \|V^*x\|^2 = \sup \{ |\langle x, by \rangle|^2 : y \in \mathfrak{D}, \langle by, my - Pby \rangle \leq 1 \}.$$

This proves the equivalence of (2a) and (3) and completes the proof of the theorem.  $\blacksquare$

**COROLLARY 1.** *Let  $\mathfrak{D}$  be a closed subspace of the Hilbert space  $\mathfrak{H}$ ,  $b : \mathfrak{D} \rightarrow \mathfrak{H}$  a suboperator satisfying property (1).  $b$  has unique positive extension of smallest norm  $m$  if and only if there exists a co-isometry  $T : \mathfrak{D} \rightarrow \mathfrak{D}^\perp$  in such a way that  $C = T(mA - A^2)^{1/2}$ , where  $A = Pb$ ,  $C = (I - P)b$  and  $P : \mathfrak{H} \rightarrow \mathfrak{D}$  is orthogonal projection.*

**PROOF.** Since  $b$  satisfied (1) we have that

$$\|Cx\|^2 + \|Ax\|^2 \leq m \langle Ax, x \rangle \quad \text{for all } x \in \mathfrak{D}.$$

Therefore

$$\langle CC^*x, x \rangle \leq \langle Ax, mx - Ax \rangle \quad \text{for all } x \in \mathfrak{D}.$$

In view of Douglas factorization theorem ([1] Thm. 1, [2]) this implies that there is a contraction  $T$  from  $\mathfrak{D}$  into  $\mathfrak{D}^\perp$  such that  $C = T(mA - A^2)^{1/2}$ . We may also assume that the kernel of  $T$  contains the kernel of  $mA - A^2$ , (this additional requirement completely determines  $T$ ). For each  $x \in \mathfrak{D}^\perp$  we have

$$\begin{aligned} & \sup\{|\langle x, by \rangle|^2 : y \in \mathfrak{D}, \langle by, my - Pby \rangle \leq 1\} = \\ & = \sup\{|\langle x, by \rangle|^2 : y \in \mathfrak{D}, \langle Ay, my - Ay \rangle \leq 1\} = \\ & = \sup\{|T^*x, (mA - A^2)^{1/2}y|^2 : y \in \mathfrak{D}, \|(mA - A^2)^{1/2}y\|^2 \leq 1\}. \end{aligned}$$

By assumption the range of  $T^*$  is contained in the norm closure of the range of  $(mA - A^2)^{1/2}$ , therefore the last supremum is simply  $\|T^*x\|^2$ .

Hence, applying Theorem 3, we get that  $b$  has unique extension if and only if  $T^*$  is an isometry i.e.  $T$  is a co-isometry. ■

REMARK. According to the notation of Corollary 1

$$C^*C = (mA - A^2)^{1/2}T^*T(mA - A^2)^{1/2}.$$

If  $T^*$  is an isometry, then  $T^*T$  is an orthogonal projection and therefore  $C^*C$  is an extremal point in the operator interval  $[0, mA - A^2]$ , where 0 stands for the identically zero operator. Conversely, if  $C^*C$  is an external point in  $[0, mA - A^2]$ , then  $T^*T$  is an orthogonal projection and therefore  $T^*$  has to be a partial isometry. So the corollary can be stated as follows. *Under the above assumption  $b$  has unique positive extension with smallest norm  $m$  if and only if  $C^*C$  is an extremal point of  $[0, mA - A^2]$  and  $C$  has dense range (equivalently  $C^*$  is injective).*

We need a well n lemma about operator ranges, see e.g. [2], [6].

LEMMA 1. *Let  $D = EE^*$  be a bounded positive operator on a Hilbert space  $\mathfrak{H}$ . Then*

$$\text{ran } D^{1/2} = \text{ran } E = \{\xi \in \mathfrak{H} : \sup\{|\langle \xi, y \rangle|^2 : y \in \mathfrak{H}, \langle Dy, y \rangle \leq 1\} < \infty\}.$$

The next lemma characterizes extremal points of an operator interval.

LEMMA 2. *Let  $D, E$  be positive operators and suppose that  $0 \leq E \leq D$ .  $E$  is an extremal point in  $[0, D]$  if and only if*

$$(4) \quad \text{ran } E^{1/2} \cap \text{ran } (D - E)^{1/2} = \{0\}.$$

PROOF. First assume that  $E$  is an extremal point in  $[0, D]$ . Then  $E = D^{1/2}PD^{1/2}$  and  $E - D = D^{1/2}(I - P)D^{1/2}$ , where  $P$  is an orthogonal

projection. Therefore the range of  $E^{1/2}$  and  $(D - E)^{1/2}$  is equal to the range of  $D^{1/2}P$  and  $D^{1/2}(I - P)$  respectively. So the only if part is almost obvious.

Conversely, since  $0 \leq E \leq D$ , there is a self-adjoint contraction  $P$  in such a way that  $E = D^{1/2}PD^{1/2}$ . We can also demand that the range of  $P$  is contained in the closure of the range of  $D$  (in this case  $P$  is uniquely determined).

Since  $\text{ran } E^{1/2}$  and  $\text{ran } (D - E)^{1/2}$  is equal to  $\text{ran } D^{1/2}P^{1/2}$  and  $\text{ran } D^{1/2}(I - P)^{1/2}$  respectively, and  $D^{1/2}$  restricted to the closure of the range of  $P$  is injective we have that condition (4) implies that  $\text{ran } P^{1/2} \cap \text{ran } (I - P)^{1/2} = \{0\}$ . But then  $P(I - P)$  is identically zero,  $P = P^2$ . So  $P$  is an orthogonal projection and  $D$  is an extremal point. ■

**THEOREM 4. (KREIN'S CRITERION).** *Let  $B : \mathfrak{D} \rightarrow \mathfrak{H}$  be subpositive sub-operator satisfying property (1).  $b$  has unique positive operator extension of smallest norm  $m$  to  $\mathfrak{H}$  if and only if*

$$(5) \quad \sup\{|\langle x, by \rangle|^2 : y \in \mathfrak{D}, \langle by, my - by \rangle \leq 1\} = \infty$$

holds for each  $x \neq 0$  from  $\mathfrak{D}^\perp$ .

**PROOF.** Without loss of generality we may assume that  $\mathfrak{D}$  is closed. Write  $b$  as a column vector as in Theorem 3. According to Corollary 1, Remark and Lemma 2 we have only to prove that (5) is equivalent to the following:  $C^*$  is injective and the ranges of  $C^*$  and  $(mA - A^2 - CC^*)^{1/2}$  have zero as the only common vector.

Since for  $x \in \mathfrak{D}^\perp, y \in \mathfrak{D}$

$$\langle x, by \rangle = \langle x, Cy \rangle = \langle C^*x, y \rangle$$

and

$$\langle by, my - by \rangle = \langle mby, y \rangle - \|by\|^2 = \langle mAy, y \rangle - \langle A^2y, y \rangle - \langle C^*Cy, y \rangle.$$

(5) can be rewritten as

$$(6) \quad \sup\{|\langle C^*x, y \rangle|^2 : y \in \mathfrak{D}, \langle (mA - A^2 - C^*C)y, y \rangle \leq 1\} = \infty$$

for all  $x \neq 0$  from  $\mathfrak{D}^\perp$ .

Now (6) obviously implies that  $C^*$  is injective and  $\text{ran } C^* \cap \text{ran } (mA - A^2 - C^*C)^{1/2} = \{0\}$  (use Lemma 1).

Conversely, if  $C^*$  is injective and  $x \in \mathfrak{D}^\perp$  is not zero then  $C^*x$  is also non zero and therefore it is not an element of the range of  $(mA - A^2 - C^*C)^{1/2}$  which by virtue of Lemma 1 implies (6). ■

The name 'Krein's criterion' can be supported with the following. It was shown in [7], [9] that the extension problem of a symmetric suboperator  $s$  to a self-adjoint operator can be traced back to the positive case treated before. Indeed,  $b : \mathfrak{D} \rightarrow \mathfrak{H}$ ,  $b = \|s\|I|_{\mathfrak{D}} + s$  satisfies condition (1) with constant  $m = 2\|s\|$ .  $s$  has only one norm preserving self-adjoint extension if and only if  $b$  has unique positive extension with smallest possible norm  $m$ . Applying Theorem 4 to this special case we arrive at Krein's classical condition.

**THEOREM 5.** *Let  $s : \mathfrak{D} \rightarrow \mathfrak{H}$  be a symmetric continuous linear operator.  $s$  has unique self-adjoint norm preserving extension if and only if*

$$\sup\{|\langle x, sy \rangle|^2 : y \in \mathfrak{D}, \|s\|^2\|y\|^2 - \|sy\|^2 \leq 1\} = \infty$$

*holds true for every  $x \neq 0$  element from  $\mathfrak{D}^\perp$ .*

### References

- [1] R. G. DOUGLAS, On majorization, factorization, and range inclusion of operators in Hilbert spaces, *Proc. Amer. Math. Soc.*, **17** (1966), 413–415.
- [2] P. A. FILLMORE, J. P. WILLIAMS, On operator ranges, *Adv. in Math.*, **7** (1971), 254–281.
- [3] P. R. HALMOS, Subnormal operators and the subdiscrete topology, Anniversary vol. on. *Appr. Theory and Functional Analysis*, **65** (1984), 49–65.
- [4] L. KAPOS, J. PAEZ, Unique extremal extensions of suboperators, *Annales Univ. Sci. Budapest., Sect. Math.*, **34** (1991), 15–19.
- [5] M. G. KREIN, The theory of self-adjoint extensions of semi-bounded Hermitian operators and its application I., *Math. Sb.*, **62** (1947), 431–495.
- [6] Z. SEBESTYÉN, On ranges of adjoint operators in Hilbert space, *Acta Sci Math. (Szeged)*, **46** (1983), 295–298.
- [7] Z. SEBESTYÉN, Restrictions of positive operators, *Acta Sci Math. (Szeged)*, **46** (1983), 299–301.
- [8] Z. SEBESTYÉN, Characterization of subprojection suboperators, *Acta Math. Hung.*, **56** (1990), 115–119.
- [9] Z. SEBESTYÉN, L. KAPOS, Extremal positive and self-adjoint extensions of suboperators, *Per. Math. Hung.*, **20** (1989), 75–80.
- [10] Z. SEBESTYÉN, Operator extensions of Hilbert space, *Acta Sci. Math.*, **57** (1993), 233–248.



## CO-STABLE QUASI-UNIFORM SPACES

By

J. DEÁK\* and S. ROMAGUERA

Mathematical Institute of the Hungarian Academy of Sciences, Budapest and  
 Department of Applied Mathematics, Polytechnical University of Valencia, Valencia, Spain

(Received May 20, 1994)

FLETCHER and HUNSAKER [22] raised the following question: Does there exist a natural class of quasi-uniform spaces containing all uniform spaces and all equinormal quiet spaces in which the concepts of D-completeness, strong D-completeness and bicompleteness coincide? In particular, do the locally symmetric quiet spaces comprise such a class?\*\*\*

[7] gives a counterexample for the second part of the problem, and a solution for the first one, showing that the above notions of completeness coincide for semi-symmetric uniformly regular spaces (which are the same as the semi-symmetric quiet spaces, see [31] Theorem 4). Following the newer terminology ([10], [15]), we shall say from now on “closed-symmetric” instead of “semi-symmetric”. KÜNZI [28] proved that closed-symmetry can be replaced here by the more general notion of mixed-symmetry introduced in [10]. (According to [10], quietness means the same as uniform regularity in mixed-symmetric spaces, too.) It is the main purpose of this paper to introduce an even more general class of spaces in which the notions of completeness coincide. Calling a quasi-uniformity  $\mathcal{U}$

(A) *co-stable* if each co-D-Cauchy ( $= \mathcal{U}^{-1}$  – D-Cauchy) filter is stable ( $= \mathcal{U}$ -stable), we shall prove that the above notions of completeness coincide in co-stable quiet spaces. (Quietness cannot be replaced here by uniform regularity.) This is indeed a class larger than the one obtained by Künzi, since, as we shall see, mixed-symmetric quiet spaces are co-stable.

---

\* Supported by the Hungarian National Foundation for Scientific Research, grant No. 2114.

\*\* The notions occurring in this introduction (and a few more needed later) will be recalled in Section 1.

As a motivation for the definition, let us mention that there exist three other classes of quasi-uniformities in the literature either defined or characterized by the stability of certain filters:

- (B) A quasi-uniformity  $\mathcal{U}$  is *stable* if each D-Cauchy (=  $\mathcal{U}$ -D-Cauchy) filter is stable. (See [29] for a detailed study of stable spaces, and [2], [12] 8.6 and [20] for a generalization.)
- (C) Each (ultra)filter is stable in a quasi-uniform space iff its conjugate is hereditarily precompact. (Independently in [12] and [29], with different proofs.)
- (D) ([14], [36]; we shall not define the notions used here.) A quasi-uniformity has a completion in the sense of SMYTH [35] (see also [28]) iff each weakly hereditarily Cauchy filter [14] (“Cauchy” in [35], [36]) is stable.

In addition to the results mentioned above, we shall also give some characterizations of co-stable quiet spaces; e.g. a quiet space is co-stable iff its D-completion is point-symmetric. Several examples and counterexamples will also be included.

We would also like to call to the reader’s attention that stable filters play an important role in the theory of completeness, completions and extensions of quasi-uniform and quasi-metric spaces, see [1], [2], [4], [5], [6], [8], [12], [13], [14], [16], [19], [20], [28], [34], [35], [36], [37].

ACKNOWLEDGEMENT. The authors are indebted to PROF. Á. CSÁSZÁR for spotting a wrong counterexample in an earlier version of this paper.

## 1. Preliminaries

See [24] for basic definitions and notations. Differences in the notations: we write  $Ux$  and  $U[A]$  instead of  $U(x)$  and  $U(A)$  (where  $U$  in an entourage); for a quasi-uniformity  $\mathcal{U}$ , the uniformity  $\sup\{\mathcal{U}^{-1}, \mathcal{U}\}$  will be denoted by  $\mathcal{U}^s$ . For the convenience of the reader, we begin with recalling some definitions from [24]: a quasi-uniformity  $\mathcal{U}$  is *point-symmetric* if  $\mathcal{I}(\mathcal{U}^{-1})$  is finer than  $\mathcal{I}(\mathcal{U})$ ; *locally symmetric* if for each  $U \in \mathcal{U}$  and each point  $x$  there is a  $V \in \mathcal{U}$  with  $V^{-1}[Vx] \subset Ux$ ; *equinormal* if disjoint closed sets are far; *bicomplete* if  $\mathcal{U}^s$  is complete in the usual sense, i.e. each  $\mathcal{U}^s$ -Cauchy filter is  $\mathcal{I}(\mathcal{U}^s)$ -convergent (a filter  $\mathfrak{f}$  in a uniform space  $(X, \mathcal{V})$  is *Cauchy* if for each  $U \in \mathcal{V}$  there is an  $S \in \mathfrak{f}$  with  $S \times S \subset U$ ).

Given a property  $P$  of quasi-uniformities, we shall say that the quasi-uniformity  $\mathcal{U}$  is *doubly P* if both  $\mathcal{U}^{-1}$  and  $\mathcal{U}$  are  $P$ . A filter  $\mathfrak{f}$  in a quasi-uniform space  $(X, \mathcal{U})$  is called *stable* ([25], [1]) if for each  $U \in \mathcal{U}$ ,  $\bigcap_{S \in \mathfrak{f}} U[S] \in \mathfrak{f}$ ; *D-Cauchy* [17] if it is the second member of a Cauchy filter pair (a filter

pair  $(f^{-1}, f^1)$  is *Cauchy* ([32], [3]) if for each  $U \in \mathcal{U}$  there are  $S_i \in f^i$  with  $S_{-1} \times S_1 \subset U$ ; *co-D-Cauchy* if it is  $\mathcal{U}^{-1}$ -D-Cauchy, i.e. the first member of a Cauchy filter pair. *Sup-dense* means  $\mathcal{I}(\mathcal{U}^s)$ -dense.

A quasi-uniformity  $\mathcal{U}$  is *D-complete* ([17], [22]) if each D-Cauchy filter is convergent (=  $\mathcal{I}(\mathcal{U})$ -convergent); *strongly D-complete* [26] if each co-D-Cauchy filter has a  $\mathcal{I}(\mathcal{U})$ -cluster point; *closed-symmetric* ([7], see also [23], [10]) if  $A\delta_{\mathcal{U}}B$  implies  $B\delta_{\mathcal{U}}A$  whenever  $A$  and  $B$  are closed (=  $\mathcal{I}(\mathcal{U})$ -closed); *mixed-symmetric* [10] if  $A\delta_{\mathcal{U}}B$  implies  $B\delta_{\mathcal{U}}A$  whenever  $A$  is open and  $B$  is closed; *uniformly regular* ([1], [21]) if for each  $U \in \mathcal{U}$  there is a  $V \in \mathcal{U}$  such that  $\text{cl}^1 Vx \subset Ux$  for each point  $x$  (where  $\text{cl}^i$  denotes the  $\mathcal{I}(\mathcal{U}^i)$  closure for  $i = \pm 1$ ); *quiet* [17] if for each  $U \in \mathcal{U}$  there is a  $V \in \mathcal{U}$  such that  $x U y$  holds whenever there is a Cauchy filter pair  $(f^{-1}, f^1)$  with  $Vx \in f^1$  and  $V^{-1}y \in f^{-1}$ ; *Cauchy bounded* [26] if each ultrafilter is D-Cauchy.

Each quiet quasi-uniform space  $(X, \mathcal{U})$  has a standard *D-completion*  $(Y, \mathcal{V})$  ([17], [18], [9]); instead of the actual construction, we shall only use the following properties;  $\mathcal{V}$  is quiet and D-complete;  $(X, \mathcal{U})$  is a doubly dense subspace of  $(Y, \mathcal{V})$ ; for each  $a \in Y$ , the *trace filter pair*  $(f^{-1}(a), f^1(a))$  is minimal Cauchy (here  $f^i(a)$  denotes the trace on  $X$  of the  $\mathcal{I}(\mathcal{V}^i)$ -neighbourhood filter of  $a$ ), and each minimal Cauchy filter pair is the trace filter pair of some point.

In counterexamples, quasi-uniformities will often be induced by quasi-(pseudo)metrics. We shall use the following convention: for pairs of points not figuring in the formula defining the quasi-metric  $d$ , let  $d(x, y) = 0$  or  $1$  according as  $x = y$  or not. The entourage  $\{(x, y) : d(x, y) < \varepsilon\}$  will be denoted by  $U_{(\varepsilon)} = U_{(\varepsilon)}(d)$ .  $\mathcal{U}_{\text{so}}$  is the *Sorgenfrey quasi-uniformity* on  $\mathbb{R}$ , i.e.  $\mathcal{U}_{\text{so}} = \mathcal{U}(d_{\text{so}})$  where  $d_{\text{so}}(x, y) = y - x$  if  $x \leq y$ .

Some connections between the notions defined above: if  $\mathcal{U}$  is quiet then so is  $\mathcal{U}^{-1}$  [17]; quiet  $\Rightarrow$  (doubly) uniformly regular [21]; equinormal  $\Rightarrow$  closed-symmetric (evident); closed-symmetric  $\Rightarrow$  mixed-symmetric [10]; mixed-symmetric + regular  $\Rightarrow$  locally symmetric [10]; locally symmetric  $\Rightarrow$  point-symmetric [24]; strongly D-complete  $\Rightarrow$  D-complete [26]; D-complete + uniformly regular  $\Rightarrow$  bicomplete [22]; totally bounded  $\Rightarrow$  Cauchy bounded  $\Rightarrow$  precompact [26].

Let us also recall that the intersection of stable filters is stable. This implies that a quasi-uniformity is co-stable iff each co-D-Cauchy ultrafilter is stable. (Since a filter finer than a co-D-Cauchy filter has the same property.) For a filter  $f$  in  $(X, \mathcal{U})$ , denote by  $\mathcal{U}(f)$  the filter  $\{U[S] : U \in \mathcal{U}, S \in f\}$ . (This is the *envelope* of  $f$ , i.e. the finest one of the round filters coarser than  $f$ .)  $f$  is stable iff  $\mathcal{U}(f)$  is so (cf [4] 7.18).

## 2. Basic properties of co-stability

**2.1.** Let us begin with the observation that, although co-stability is a rather strong property (e.g. the doubly stable quiet quasi-uniformity  $\mathcal{U}_{\text{so}}$  is not co-stable), there exist co-stable spaces in abundance:

a) Uniformities are doubly co-stable for the same reason as they are doubly stable (cf [29] Example 2.1 (a)): the D-Cauchy = co-D-Cauchy filters are now the same as the Cauchy filters in the usual sense (cf [17]), and Cauchy filters in a uniform space are stable ([25] II 49).

b) If  $\mathcal{U}^{-1}$  is hereditarily precompact then  $\mathcal{U}$  is co-stable (and stable [29]) by (C). In particular, totally bounded quasi-uniformities are doubly co-stable.

c) The quasi-uniform Michael line  $(\mathbb{R}, \mathcal{M})$  defined by the quasi-metric  $d(x, y) = \min\{|x - y|, 1\}$  if  $x \in \mathbb{Q}$  is co-stable ( $\mathbb{Q}$  denotes the set of the rationals): Let  $(\mathfrak{f}^{-1}, \mathfrak{f}^1)$  be a Cauchy filter pair, and take  $S_n \in \mathfrak{f}^{-1}$ ,  $T_n \in \mathfrak{f}^1$  such that  $S_n \times T_n \subset U_{(1/n)}$ . With  $x_n \in T_n$ , we have  $S_n \subset (]x_n - 1/n, x_n + 1/n[ \cap \mathbb{Q} \cup \{x_n\})$ , so  $\mathfrak{f}^{-1}$  converges to some  $y$  in the Euclidean topology. Now either  $\mathbb{Q} \in \mathfrak{f}^{-1}$  or  $x_1 = x_2 = \dots = y \in \bigcap \mathfrak{f}^{-1}$ . Fixing some  $\varepsilon > 0$ , we have in the first case  $]y - \varepsilon, y + \varepsilon[ \cap \mathbb{Q} \subset U_{(2\varepsilon)}[S]$  for each  $S \in \mathfrak{f}^{-1}$ , since  $S$  contains a rational in  $]y - \varepsilon, y + \varepsilon[$ . In the second case,  $\mathfrak{f}^{-1}$  is the intersection of the ultrafilter fixed at  $y$  and of a filter considered in the first case, thus  $\mathfrak{f}^{-1}$  is stable again. But the D-Cauchy filter generated by the sets  $]0, \varepsilon[ \setminus \mathbb{Q}$  ( $\varepsilon > 0$ ) is not stable either in  $(\mathbb{R}, \mathcal{M})$  or in  $(\mathbb{R}, \mathcal{M}^{-1})$ , thus  $\mathcal{M}$  is neither stable nor doubly co-stable.

d) The quasi-uniformity  $\mathcal{Q}$  on  $\mathbb{R}$  induced by the quasi-pseudometric  $d(x, y) = \max\{y - x, 0\}$  is co-stable: Let  $\mathfrak{f}$  be a co-D-Cauchy filter,  $\varepsilon > 0$ . Then  $\mathfrak{f}$  has an element bounded from below. If each element of  $\mathfrak{f}$  is unbounded then  $U_{(\varepsilon)}[S] = \mathbb{R}$  ( $S \in \mathfrak{f}$ ); otherwise,  $t = \inf_{S \in \mathfrak{f}} \sup S \in \mathbb{R}$ , and  $U_{(\varepsilon)}[S] \supset ]t - \varepsilon, t + \varepsilon[ \in \mathfrak{f}$ .

$\mathcal{Q}$  is in fact doubly co-stable, since it is isomorphic with  $\mathcal{Q}^{-1}$ . But  $\mathcal{Q}$  is not stable: the sets  $]t - \varepsilon, t + \varepsilon[$  ( $x \in \mathbb{R}$ ) generate a non-stable D-Cauchy filter.

e) Co-stability is a hereditary property. (Proof: observe that a filter in a subspace is stable iff it generates a stable filter in the larger space, and also that a co-D-Cauchy filter in the subspace generates a co-D-Cauchy filter.)

f) The supremum of co-stable quasi-uniformities is so-stable. (Proof: Let the ultrafilter  $\mathfrak{u}$  be co-D-Cauchy with respect to the supremum. Then  $\mathfrak{u}$  is co-D-Cauchy in each coarser space. Thus we can apply [12] Lemma 5.9, which states that, in case of ultrafilters, it is enough to consider entourages from a subbase in the definition of stability.)

g) The co-stable spaces form a concretely reflective subcategory of the category of quasi-uniform spaces. (In view of e) and f), it is enough to check the trivial statement that co-stability is preserved by the initial structure under a surjection.) In particular, co-stability is productive.

REMARK. A partial converse to b) (similar to [29] 2.8): If  $\mathcal{U}$  is co-stable and  $\mathcal{U}^{-1}$  is Cauchy bounded then  $\mathcal{U}^{-1}$  is hereditarily precompact. (Clear from (C) and the definitions.) Analogously to [29] 2.10, it follows that if  $\mathcal{U}^{-1}$  is Cauchy bounded and  $\mathcal{U}$  is hereditarily precompact and co-stable then  $\mathcal{U}$  is totally bounded. (Cauchy boundedness is essential here: restrict  $\mathcal{U}_{\text{so}}$  to  $\{1/n : n \in \mathbb{N}\}$ .)

**2.2.** We have the following four properties of a quasi-uniformity: it is stable or co-stable, respectively its conjugate is stable or co-stable. No three of these properties imply the fourth one. For reasons of symmetry, it is enough to give two examples; both spaces are very simple, and they frequently occur in the literature:

EXAMPLES. a)  $\mathcal{U} = \mathcal{U}_{\text{so}}|_{\{0\} \cup \{1/n : n \in \mathbb{N}\}}$  is doubly stable, co-stable, but  $\mathcal{U}^{-1}$  is not co-stable.

b) Let  $X = \{-1/n : n \in \mathbb{N}\} \cup \{0\}$ ,

$$d(x, y) = \begin{cases} y - x & \text{if } x \leq y < 0, \\ -y & \text{if } x = 0, \end{cases}$$

$\mathcal{U} = \mathcal{U}(d)$ . Although  $\mathcal{U}$  is not stable (the filter consisting of the cofinite sets is D-Cauchy but not stable), it has the other three properties. ■

The space in a) is quiet; the one in b) cannot be quiet, since we shall see later that a quiet doubly co-stable space is doubly stable.

**2.3.** Compare the following result with [30] Proposition 13, which states the same for stable spaces:

PROPOSITION. *A quasi-uniform space is co-stable iff it has a sup-dense co-stable subspace.*

PROOF. Let  $Z$  be a sup-dense co-stable subspace of  $(X, \mathcal{U})$ , and take a Cauchy filter pair  $(\mathfrak{f}^{-1}, \mathfrak{f}^1)$  in  $X$ . Then  $(\mathcal{U}^{-1}(\mathfrak{f}^{-1}), \mathcal{U}(\mathfrak{f}^1))$  is also Cauchy ([4] Lemma 7.12 b)), hence so is the finer filter pair  $(\mathcal{U}^s(\mathfrak{f}^{-1}), \mathcal{U}(\mathfrak{f}^1))$ . By the sup-density, this filter pair has a trace on  $Z$ , thus  $\mathfrak{h} = \mathcal{U}^s(\mathfrak{f}^{-1})|_Z$  is co-D-Cauchy in  $Z$ . According to our assumption,  $\mathfrak{h}$  is stable, and so it generates a stable filter  $\mathfrak{k}$  in  $X$ . It is enough to prove now that  $\mathcal{U}(\mathfrak{k}) = \mathcal{U}(\mathfrak{f}^{-1})$ .

$\mathfrak{f} \supset \mathcal{U}^s(\mathfrak{f}^{-1}) \supset \mathcal{U}(\mathfrak{f}^{-1})$ , thus  $\mathcal{U}(\mathfrak{f}) \supset \mathcal{U}(\mathfrak{f}^{-1})$ . Conversely, let  $S \in \mathcal{U}(\mathfrak{f})$ . Then there are  $T \in \mathfrak{f}^{-1}$  and  $U \in \mathcal{U}$  such that  $S \supset U \left[ (U^{-1} \cap U)[T] \cap Z \right]$ . For each  $x \in T$ , there is a  $z \in Z$  with  $x U z U x$ , thus  $x \in S$ ,  $T \subset S$ ,  $S \in \mathfrak{f}^{-1}$ ,  $\mathcal{U}(\mathfrak{f}) \subset \mathfrak{f}^{-1}$ ,  $\mathcal{U}(\mathfrak{f}) \subset \mathcal{U}(\mathfrak{f}^{-1})$ . ■

**2.4.** In order to formulate some statements more flowingly, let us introduce two auxiliary notions: a quasi-uniformity  $\mathcal{U}$  is *strongly stable* if  $\mathcal{U}$  is stable and  $\mathcal{U}^{-1}$  is co-stable; it is *strongly co-stable* if  $\mathcal{U}^{-1}$  is strongly stable (i.e. if  $\mathcal{U}$  is co-stable and  $\mathcal{U}^{-1}$  is stable).

LEMMA. *The following conditions are equivalent for a quasi-uniformity  $\mathcal{U}$ :*

- (i)  $\mathcal{U}$  is strongly stable;
- (ii) each D-Cauchy ultrafilter is  $\mathcal{U}^s$ -Cauchy (=  $\mathcal{U}^s$ -stable);
- (iii) each D-Cauchy filter is  $\mathcal{U}^s$ -stable.

PROOF. (i)  $\Rightarrow$  (ii). If an ultrafilter is stable with respect to some quasi-uniformities then it is also stable with respect to their supremum ([12] Lemma 5.9). In a uniform space, an ultrafilter is stable iff it is Cauchy.

(ii)  $\Rightarrow$  (iii). The intersection of stable (ultra)filters is stable.

(iii)  $\Rightarrow$  (i). Evident. ■

We cannot add the stronger assumption that each D-Cauchy filter is  $\mathcal{U}^s$ -Cauchy (but see 4.1); On  $X = \{0, 1, 2\}$ , let  $d(0, 1) = d(0, 2) = 0$ ,  $\mathcal{U} = \mathcal{U}(d)$ ; now  $\{\{1, 2\}, X\}$  is a D-Cauchy filter, which is not  $\mathcal{U}^s$ -Cauchy. (An example with better separation properties will be given in 4.2.)

REMARK. We shall later need the following statement shown in (i)  $\Rightarrow$  (ii) of the above proof: any doubly stable ultrafilter is  $\mathcal{U}^s$ -Cauchy.

PROPOSITION. *If  $\mathcal{U}$  is a point-symmetric quasi-uniformity such that  $\mathcal{U}^{-1}$  is D-complete then  $\mathcal{U}$  is strongly co-stable.*

PROOF. Any co-D-Cauchy filter  $\mathfrak{f}$  is convergent in the topology  $\mathcal{I}(\mathcal{U}^{-1})$ , which is the same as  $\mathcal{I}(\mathcal{U}^s)$  by the point-symmetry. Hence  $\mathfrak{f}$  is  $\mathcal{U}^s$ -Cauchy, implying that (iii) in the lemma holds for  $\mathcal{U}^{-1}$ . ■

### 3. Co-stability in uniformly regular spaces

**3.1.** It is straightforward that if  $\mathcal{U}$  is point-symmetric then  $\mathcal{U}^{-1}$  is uniformly regular (e.g. [9] Remark 1.3). According to [31] Proposition 7 (combined with Lemma 4), the converse holds assuming that  $\mathcal{U}^{-1}$  is hereditarily precompact. More generally (cf (C)):

**PROPOSITION.** *A co-stable quasi-uniformity  $\mathcal{U}$  is point-symmetric iff  $\mathcal{U}^{-1}$  is uniformly regular.*

**PROOF.** Assume that  $(X, \mathcal{U})$  is co-stable and  $\mathcal{U}^{-1}$  is uniformly regular. Given  $y \in X$  and  $U \in \mathcal{U}$ , we have to show that there is a  $W \in \mathcal{U}$  with  $W^{-1}y \subset Uy$ . Take a  $V \in \mathcal{U}$  such that  $\text{cl}^{-1} V^{-1}x \subset U^{-1}x$  ( $x \in X$ ). The neighbourhood filter pair  $(f^{-1}, f^1)$  of  $y$  is Cauchy, thus  $f^{-1}$  is co-D-Cauchy, hence stable, and so there is a  $W \in \mathcal{U}$  such that  $W^{-1}y \subset V[S]$  ( $S \in f^{-1}$ ). Let  $x \in W^{-1}y$ . Then  $x \in V[S]$  ( $S \in f^{-1}$ ), i.e.  $V^{-1}x \cap S \neq \emptyset$  ( $S \in f^{-1}$ ), implying  $y \in \text{cl}^{-1} V^{-1}x \subset U^{-1}x$ , and so  $x \in Uy$ ,  $W^{-1}y \subset Uy$ . ■

Any doubly point-symmetric quasi-uniformity is doubly locally symmetric ([15] 1.1), hence a doubly co-stable quasi-uniformity is doubly uniformly regular iff it is doubly locally symmetric. But point-symmetry cannot be replaced by local symmetry in the proposition:

**EXAMPLE.** Let  $X = \{0\} \cup \mathbb{N}$ ,  $d(0, 2n) = d(2n - 1, 2n) = 1/n$  ( $n \in \mathbb{N}$ ).  $\mathcal{U}(d)$  is quiet, co-stable, but not locally symmetric. ■

**3.2. PROPOSITION.** *Any point-symmetric D-complete uniformly regular quasi-uniformity is strongly co-stable.*

**PROOF.** If  $\mathcal{U}$  is D-complete and uniformly regular then  $\mathcal{U}^{-1}$  is D-complete, too ([22] 2.1), thus Proposition 2.4 applies. ■

**COROLLARY.** *A D-complete doubly uniformly regular quasi-uniformity is point-symmetric iff it is (strongly) co-stable.*

**PROOF.** Propositions 3.1 and 3.2. ■

**REMARK.** Any D-complete doubly uniformly regular quasi-uniformity is quiet, see in [9] 1.3.

**3.3. COROLLARY.** *Any compact uniformly regular quasi-uniformity is strongly co-stable.*

**PROOF.** Compact regular quasi-uniformities are point-symmetric ([24] 2.26 and 2.24); any compact quasi-uniformity is evidently strongly D-complete (cf [26]). Thus the conditions of Proposition 3.2 are satisfied. ■

REMARKS. a) The conjugate of Example 2.2 a) shows that the compactness of  $\mathcal{U}$  cannot be replaced by that of  $\mathcal{U}^{-1}$ .

b) Although only regularity is used in the first step of the proof, uniform regularity is essential in the corollary:

EXAMPLE. On  $X = \{0\} \cup \mathbb{N}$ , let  $d(0, n) = 1/n$  ( $n \in \mathbb{N}$ ),  $d(k, n) = |1/k - 1/n|$  if  $k, n \in \mathbb{N}$ ,  $k$  is odd,  $n$  is even.  $\mathcal{U}(d)$  is compact, but it has none of the four stability properties. ■

**3.4. PROPOSITION.** *Let  $\mathcal{U}$  be a quasi-uniformity such that  $\mathcal{U}^{-1}$  is uniformly regular. Then  $\mathcal{U}$  is co-stable iff there is a doubly dense co-stable subspace.*

PROOF. Let  $Z$  be a doubly dense co-stable subset in  $(X, \mathcal{U})$ ; by Proposition 2.3, it is enough to show that if  $\mathcal{U}^{-1}$  is uniformly regular then  $Z$  is actually sup-dense. Take a point  $p \in X \setminus Z$ , and let  $(n^{-1}, n^1)$  denote the neighbourhood filter pair of  $p$  in  $Y = Z \cup \{p\}$ . We claim that  $\mathcal{U} \upharpoonright Y$  is co-stable. Let  $u$  be a co-D-Cauchy free ultrafilter. There is a filter  $f$  such that  $(u, f)$  is Cauchy; then  $(u, \mathcal{U}(f))$  is Cauchy, too, and this filter pair has a trace on  $Z$ , thus  $u \upharpoonright Z$  is co-D-Cauchy, hence stable, implying that  $u$  is stable, too (as it is the filter in  $Y$  generated by  $u \upharpoonright Z$ ). The fixed ultrafilters are evidently stable, so  $\mathcal{U} \upharpoonright Y$  is co-stable indeed.

Uniform regularity is a hereditary property, thus  $(\mathcal{U} \upharpoonright Y)^{-1}$  is uniformly regular, hence  $\mathcal{U} \upharpoonright Y$  is point-symmetric by Proposition 3.1. This means that  $n^1 \subset n^{-1}$ , thus  $n^1 \upharpoonright Z \subset n^{-1} \upharpoonright Z$ , implying that  $Z$  is sup-dense in each set  $Z \cup \{p\}$ , therefore in  $X$ . ■

In some similar situations, it is enough to know in a uniformly regular space that there is a dense subspace with the property in question (see [33] 6.4 for total boundedness, [28] Lemma 1 for Cauchy boundedness). In the case of co-stability, density in one of the topologies is not sufficient, not even in quiet spaces:

EXAMPLES. a) On  $\{0\} \cup \mathbb{N}$ , let  $d(n, 0) = 1/n$  ( $n \in \mathbb{N}$ ).  $\mathcal{U} = \mathcal{U}(d)$  is quiet. It is not co-stable, although  $\mathbb{N}$  is a  $\mathcal{G}(\mathcal{U}^{-1})$ -dense co-stable subspace.

b) On  $X = ]0, 1/2[ \cup \{(0, 0)\}$ , let

$$d((x', x''), (y', y'')) = \begin{cases} y' & \text{if } x' = 0, x'' = y'', \\ x'' + y' & \text{if } x' = 0 = y''. \end{cases}$$

$\mathcal{U} = \mathcal{U}(d)$  is quiet. It is not co-stable, but  $]0, 1/2[ \times ]0, 1/2[$  is a discrete, hence co-stable, dense subspace.

c) With  $x$  from b), let  $Y = X \cup ]-1/2, 0[ \times ]0, 1/2[$ . Using the notation  $x = (x', x'')$ ,  $y = (y', y'')$ , define

$$e(x, y) = \begin{cases} \max\{|x' - y'|, |x'' - y''|\} & \text{if } x \notin X, \\ d(x, y) & \text{if } x, y \in X. \end{cases}$$

$\mathcal{U} = \mathcal{U}(e)$  is uniformly regular, but not co-stable, since  $(X, \mathcal{U})$  from b) is a subspace of  $(Y, \mathcal{U})$ .  $Z = Y \setminus \{0\} \times ]0, 1/2[$  is a doubly dense co-stable subspace: any co-D-Cauchy filter in  $Z$  contains  $Y \setminus X$ , and  $\mathcal{U}|_{Y \setminus X}$  is a uniformity.

This shows that the uniform regularity of  $\mathcal{U}^{-1}$  cannot be replaced by that of  $\mathcal{U}$  in the proposition. ■

#### 4. Co-stability in quiet spaces

**4.1. THEOREM.** *The following conditions are equivalent for a quiet quasi-uniformity  $\mathcal{U}$ :*

- (i)  $\mathcal{U}$  is co-stable;
- (ii)  $\mathcal{U}$  is strongly co-stable;
- (iii) each co-D-Cauchy filter is  $\mathcal{U}^s$ -Cauchy;
- (iv) if  $(\mathfrak{f}^{-1}, \mathfrak{f}^1)$  is a minimal Cauchy filter pair then  $\mathfrak{f}^{-1} \supset \mathfrak{f}^1$ ;
- (v) the D-completion of  $\mathcal{U}$  is co-stable;
- (vi) the D-completion of  $\mathcal{U}$  is strongly co-stable;
- (vii) the D-completion of  $\mathcal{U}$  is point-symmetric.

REMARK. (i)  $\Rightarrow$  (ii) generalizes [31] Proposition 10; where hereditary precompactness is assumed, and it is concluded that the D-Cauchy ultrafilters are  $\mathcal{U}^s$ -Cauchy (cf. Lemma 2.4).

PROOF. (i)  $\Rightarrow$  (v). Proposition 3.4.

(v)  $\Leftrightarrow$  (vi)  $\Leftrightarrow$  (vii). Corollary 3.2.

(vii)  $\Rightarrow$  (iv). Each minimal Cauchy filter pair is the trace filter pair of a point from the D-completion (see in Section 1).

(iv)  $\Rightarrow$  (iii). Let  $(\mathfrak{f}^{-1}, \mathfrak{f}^1)$  be a Cauchy filter pair.  $\mathcal{U}$  being quiet, there is a minimal Cauchy filter pair  $(\mathfrak{m}^{-1}, \mathfrak{m}^1)$  coarser than  $(\mathfrak{f}^{-1}, \mathfrak{f}^1)$ . (This follows from [4] Lemma 7.11; we shall return to the details in 4.3.) From  $\mathfrak{m}^{-1} \supset \mathfrak{m}^1$  we have that  $(\mathfrak{m}^{-1}, \mathfrak{m}^{-1})$  is also Cauchy, i.e.  $\mathfrak{m}^{-1}$  is  $\mathcal{U}^s$ -Cauchy, and then so is  $\mathfrak{f}^{-1} \supset \mathfrak{m}^{-1}$ .

(iii)  $\Rightarrow$  (ii). Lemma 2.4.

(ii)  $\Rightarrow$  (i). Evident. ■

Another proof (of a more general result) will be given in 4.4.

**COROLLARY.** *The following conditions are equivalent for a quiet quasi-uniformity  $\mathcal{U}$ :*

- (i)  $\mathcal{U}$  is doubly co-stable;
- (ii)  $\mathcal{U}$  is doubly stable and doubly co-stable;
- (iii) if  $(f^{-1}, f^1)$  is a Cauchy filter pair then  $f^{-1} \cap f^1$  is  $\mathcal{U}^s$ -Cauchy;
- (iv) if  $(f^{-1}, f^1)$  is a minimal Cauchy filter pair then  $f^{-1} = f^1$ ;
- (v) the D-completion of  $\mathcal{U}$  is doubly point-symmetric.

**PROOF.** If  $\mathcal{U}$  is quiet then so is  $\mathcal{U}^{-1}$ , and the D-completion of  $\mathcal{U}$  can be identified with that of  $\mathcal{U}^{-1}$ , see [17], [18] or [9]. ■

**4.2.** Assume now that  $\mathcal{U}$  is not quiet, but only doubly uniformly regular. The construction of the D-completion cannot be extended to such spaces, so only the following question remains: is any of the conditions (ii) to (iv) from Theorem 4.1 equivalent to co-stability? Let us first observe that the implications (iii)  $\Rightarrow$  (ii)  $\Rightarrow$  (i) and (iii)  $\Rightarrow$  (iv) evidently hold in any space. ((iii)  $\Rightarrow$  (iv): if  $(f^{-1}, f^1)$  is minimal Cauchy and  $f^{-1}$  is  $\mathcal{U}^s$ -Cauchy then  $(f^{-1}, f^1 \cap f^{-1})$  is also Cauchy, thus  $f^1 \cap f^{-1} = f^1$  by the minimality.) The examples below show that nothing more holds. In each case, both topologies are discrete, thus the spaces are doubly uniformly regular.

**EXAMPLES.** a) (i)  $\not\Rightarrow$  (ii). On  $X = \{\pm 1/n : 1 \neq n \in \mathbb{N}\}$ , let  $d(x, y) = |y - x|$  if either  $x < y < 0$  or  $0 < x < y$  or  $y < 0 < x$ . Then  $\mathcal{U} = \mathcal{U}(d)$  is doubly co-stable, but neither  $\mathcal{U}$  nor  $\mathcal{U}^{-1}$  is stable.

b) (ii)  $\not\Rightarrow$  (iv). (A well-known space.) On  $X$  from a), let  $d(x, y) = |y - x|$  if either  $xy > 0$  or  $x < 0 < y$ . Then  $\mathcal{U} = \mathcal{U}(d)$  is totally bounded, hence doubly strongly stable.  $f^{-1}$  consisting of the cofinite subsets of  $X$  and  $f^1$  generated by the cofinite subsets of  $\{x \in X : x > 0\}$  form a minimal Cauchy filter pair for which  $f^{-1} \not\supseteq f^1$ .

c) (iv)  $\not\Rightarrow$  (i). On  $X = \mathbb{N}^2$ , let

$$d((1, k), (m, n)) = (1/k) + (1/n) \quad \text{if } 1 < m < k,$$

$\mathcal{U} = \mathcal{U}(d)$ . Assume that  $(f^{-1}, f^1)$  is a Cauchy filter pair. Then either  $f^{-1} = f^1$  is a fixed ultrafilter or  $A_1 \in f^{-1}$ ,  $X \setminus A_1 \in f^1$  where  $A_m = \{m\} \times \mathbb{N}$  ( $m \in \mathbb{N}$ ). We are going to show that  $(f^{-1}, f^1)$  cannot be minimal in the second case. Take  $S_i \in f^i$  with  $S_{-1} \times S_1 \subset U_{(1)}$ , and pick a  $k \in \mathbb{N}$  such that  $(1, k) \in S_{-1}$ . Then

$S_1 \subset \bigcup_{m=2}^{k-1} a_m$ . Now  $(f^{-1}, g)$  is a Cauchy filter pair strictly coarser than  $(f^{-1}, f^1)$  where the sets  $S \cup C$  ( $S \in f^1$ ,  $C$  is a cofinite subset of  $A_k$ ) form a base for  $g$ . Thus  $f^{-1} = f^1$  for each minimal Cauchy filter pair, but neither  $\mathcal{U}$  nor  $\mathcal{U}^{-1}$  is co-stable. ■

**4.3.** We shall consider now another generalization of quietness neither implying nor implied by double uniform regularity. A Cauchy filter pair  $(f^{-1}, f^1)$  in the quasi-uniform space  $(X, \mathcal{U})$  is *weakly concentrated* [4] if for any  $U \in \mathcal{U}$  there is a  $V \in \mathcal{U}$  such that  $Vx \in f^1$ ,  $V^{-1}y \in f^{-1}$  imply  $x U y$ ; a family of Cauchy filter pairs is *uniformly weakly concentrated* [4] if the above condition holds with  $V$  depending only on  $U$ , and not on the filter pair. Using this terminology, quietness means that the Cauchy filter pairs are uniformly weakly concentrated. Let us call  $\mathcal{U}$  *locally quiet* if all the Cauchy filter pairs are weakly concentrated (but not necessarily uniformly so). Some results for quiet spaces can be generalized to locally quiet spaces, see [9] 1.3 and [28] Footnote 13. Parts of Theorem 4.1 can be generalized similarly:

**PROPOSITION.** *For any locally quiet quasi-uniformity  $\mathcal{U}$ , Conditions (i) to (iv) from Theorem 4.1 are equivalent.*

**PROOF.** As no D-completion is available now, we have to proceed differently.

(i)  $\Rightarrow$  (ii). Assume that  $\mathcal{U}$  is co-stable, and let  $u$  be a co-D-Cauchy ultrafilter. Pick a filter  $f$  such that  $(u, f)$  is Cauchy. Given  $U \in \mathcal{U}$ , choose  $V \in \mathcal{U}$  such that  $Vx \in f$ ,  $V^{-1}y \in u$  imply  $x U y$ , and then  $S_{-1} \in u$ ,  $S_1 \in f$  with  $S_{-1} \times S_1 \subset V$ . As  $u$  is stable, we have  $T = S_{-1} \cap \bigcap_{S \in u} V[S] \in u$ . For  $z \in T$ ,  $Vz \supset S_1$ , thus  $Vz \in f$ ; on the other hand,  $z \in V[S]$  ( $S \in u$ ), thus  $V^{-1}z \cap S \neq \emptyset$  ( $S \in u$ ), and so  $V^{-1}z \in u$  (as  $u$  is an ultrafilter). Therefore  $x U y$  ( $x, y \in T$ ), implying that  $u$  is  $\mathcal{U}^s$ -Cauchy. Thus Lemma 2.4 yields that  $\mathcal{U}$  is strongly co-stable.

(ii)  $\Rightarrow$  (iv). Let  $(f^{-1}, f^1)$  be a minimal Cauchy filter pair, and take an ultrafilter  $u \supset f^{-1}$ . Then  $u$  is  $\mathcal{U}^s$ -Cauchy (Remark 2.4), thus  $(u, u)$  is a Cauchy filter pair, and so is  $(u, u \cap f^1)$ . Now  $(u, u \cap f^1)$  and  $(f^{-1}, f^1)$  are Cauchy filter pairs coarser than the weakly concentrated Cauchy filter pair  $(u, f^1)$ . According to [4] Lemma 7.11, there is always a coarsest one among the Cauchy filter pairs coarser than a given weakly concentrated Cauchy filter pair, thus  $(f^{-1}, u \cap f^1)$  is also Cauchy, i.e.  $u \supset f^1$  by the minimality of  $(f^{-1}, f^1)$ . This holds for each ultrafilter  $u \supset f^{-1}$ , therefore  $f^{-1} \supset f^1$ .

(iv)  $\Rightarrow$  (iii). Just as in the proof of Theorem 4.1, using [4] Lemma 7.11 (see above).

(iii)  $\Rightarrow$  (ii)  $\Rightarrow$  (i). See at the beginning of 4.2. ■

The doubly uniformly regular spaces in Examples 4.2 are not locally quiet (use the above proposition or check directly). On the other hand, a locally quiet, not doubly uniformly regular space is given in [9] Example 1.3 f).

**4.4.** A quasi-uniformity is *weakly quiet* [9] if it is locally quiet, and the minimal Cauchy filter pairs are uniformly weakly concentrated. The construction of the D-completion can be extended to weakly quiet spaces, and it has the properties mentioned in Section 1, except of course that the D-completion of a weakly quiet space is in general only weakly quiet, and not quiet. (See [9] 2.1 and 2.2.)

**LEMMA.** *A D-complete locally quiet quasi-uniformity is point-symmetric iff it is (strongly) co-stable.*

**PROOF.** By proposition 4.3, it is enough to prove that, for D-complete locally quiet spaces, (iv) is equivalent to point-symmetry. As the neighbourhood filter pairs are minimal Cauchy ([4] Lemma 7.14), (iv) always implies point-symmetry, and let  $(f^{-1}, f^1)$  be a minimal Cauchy filter pair. Then  $f^1$  converges to some point  $x$ . Denote by  $(n^{-1}, n^1)$  the neighbourhood filter pair of  $x$ . Now  $n^1 \subset f^1$ , thus  $(f^{-1} \cap n^{-1}, f^1)$  is Cauchy. Both  $(f^{-1} \cap n^{-1}, f^1)$  and  $(n^{-1}, n^1)$  are Cauchy filter pairs coarser than  $(n^{-1}, f^1)$ , therefore  $(f^{-1} \cap n^{-1}, n^1)$  is also Cauchy ([4] Lemma 7.11 is used in the same way as in 4.3). The minimality of  $(f^{-1}, f^1)$ , respectively of  $(n^{-1}, n^1)$ , gives that  $n^1 = f^1$ , respectively  $f^{-1} \cap n^{-1} = n^{-1}$ , thus  $f^{-1} \supset n^{-1} \supset n^1 = f^1$ . ■

**REMARK.** The spaces figuring in the above lemma are in fact weakly quiet, since the neighbourhood filter pairs are always uniformly weakly concentrated ([4] Lemma 7.16).

**THEOREM.** *All the conditions from Theorem 4.1 are equivalent for any weakly quiet quasi-uniformity  $\mathcal{U}$ .*

**PROOF.** (i) to (iv) are equivalent by Proposition 4.3, and (v) to (vii) by the lemma above. (v)  $\Rightarrow$  (i) is evident.

(i)  $\Rightarrow$  (v). As (iv) follows from (i), the original space is sup-dense in the D-completion, thus Proposition 2.3 can be applied. ■

According to Proposition 3.1, the conjugate of the D-completion of  $\mathcal{U}$  is uniformly regular in the above theorem. But  $\mathcal{U}$  is not necessarily uniformly regular:

EXAMPLE. On  $X = (\{1/n : n \in \mathbb{N}\} \cup \{-1, 0\}) \times \mathbb{N}$ , let

$$d((0, k), (1/n, k)) = 1/n, \quad d((-1, k), (1/n, k)) = 1/k$$

for  $k, n \in \mathbb{N}$ .  $\mathcal{U}(d)$  is D-complete, weakly quiet, co-stable, but not uniformly regular, since

$$(0, k+1) \in (\text{cl}^1 U_{(1/k)}(-1, k+1)) \setminus U_{(1)}(-1, k+1). \quad \blacksquare$$

## 5. Completeness in co-stable spaces

**5.1. THEOREM.** *The following conditions are equivalent for a locally quiet co-stable quasi-uniformity:*

- (i) *it is bicomplete;*
- (ii) *it is D-complete;*
- (iii) *it is strongly D-complete.*

PROOF. (i)  $\Rightarrow$  (iii). Let  $u$  be a co-D-Cauchy ultrafilter. By Proposition 4.3,  $u$  is  $\mathcal{U}^s$ -Cauchy, so it is  $\mathcal{T}(\mathcal{U}^s)$ -convergent, hence  $\mathcal{T}(\mathcal{U})$ -convergent.

(iii)  $\Rightarrow$  (ii). This implication holds for any quasi-uniformity, see [26] Theorem 3.

(ii)  $\Rightarrow$  (i). Let  $f$  be a  $\mathcal{U}^s$ -Cauchy filter, and take the (unique) minimal Cauchy filter pair  $(f^{-1}, f^1)$  coarser than  $(f, f)$ . Then  $(f^{-1}, f^1)$  is the neighbourhood filter pair of some point  $x$ : add to the reasoning in the proof of Lemma 4.4 that, by the minimality of  $(f^{-1}, f^1)$ , we also have  $f^{-1} \cap n^{-1} = f^{-1}$ , thus  $f^{-1} = n^{-1}$ . Therefore  $f$   $\mathcal{T}(\mathcal{U}^s)$ -converges to  $x$ .  $\blacksquare$

In quiet spaces, the argument in the proof of (ii)  $\Rightarrow$  (i) can be replaced by recalling that any D-complete uniformly regular quasi-uniformity is bicomplete ([22] Proposition 2.2). A bicomplete doubly uniformly regular doubly co-stable space is not necessarily D-complete, see Example 4.2 a). But all the bicomplete strongly co-stable spaces are strongly D-complete (apply Lemma 2.4), hence:

PROPOSITION. *Conditions (i) to (iii) from the above theorem are equivalent for strongly co-stable uniformly regular quasi-uniformities.*  $\blacksquare$

The class of spaces occurring in the proposition is actually smaller than the one in the theorem: any strongly D-complete uniformly regular quasi-uniformity is quiet ([22] 3.6). The theorem, but not the proposition, can be applied to  $\mathcal{U}(d)$  from Example 4.4.

**5.2.** According to [28] Proposition 6 (b), any mixed-symmetric quiet quasi-uniformity is D-complete iff it is bicomplete. The classes of spaces in Theorem and Proposition 5.1 are larger:

**PROPOSITION.** *Any mixed-symmetric quiet quasi-uniformity is co-stable.*

**PROOF.** It was established in the proof of [28] Proposition 6 (b) that if  $\mathcal{U}$  is mixed-symmetric and quiet then for each co-D-Cauchy filter there exists a finer  $\mathcal{U}^s$ -Cauchy filter. Hence each co-D-Cauchy ultrafilter is  $\mathcal{U}^s$ -Cauchy, and Lemma 2.4 applies. ■

There are not mixed-symmetric quiet spaces to which Theorem 5.1 can be applied:

**EXAMPLES.** a) On  $X = \mathbb{N}$ , let  $d(2n, 2n + 1) = 1/n$  ( $n \in \mathbb{N}$ ). Then  $\mathcal{U}(d)$  is quiet, doubly strongly stable, D-complete, but not mixed-symmetric.

b) (The quasi-metric Pixley–Roy space.) Take the metric  $e(s, t) = \min\{|s - t|, 1\}$  on  $\mathbb{R}/\mathbb{R}$ . Let  $X$  consist of the non-empty finite subsets of  $\mathbb{R}/\mathbb{R}$ , and define  $\mathcal{U} = \mathcal{U}(d)$  by

$$d(x, y) = \inf\{\varepsilon > 0; y \subset U_{(\varepsilon)}(e)[x]\} \quad \text{if } x \subset y.$$

$\mathcal{J}(\mathcal{U})$  is the well-known Pixley–Roy topology, see e.g. [24] 7.38. Each co-D-Cauchy filter is a fixed ultrafilter, thus  $\mathcal{U}$  is strongly D-complete and strongly co-stable.  $\mathcal{U}$  is uniformly regular: check that, for  $U = U_{(\varepsilon)}(d)$ , we can take  $V = U_{(\varepsilon/2)}(d)$  in the definition. Therefore, being strongly D-complete,  $\mathcal{U}$  is also quiet ([22] 3.6). But  $\mathcal{U}$  is not mixed-symmetric: if it were then, being quiet, it would be locally symmetric; however, with  $U = U_{(1)}(d)$  and  $V = U_{(\varepsilon)}(d)$ , we have  $\{\varepsilon/2\} \in V^{-1}[V\{0\}] \setminus U\{0\}$  (as  $\{0\} \in V\{0, \varepsilon/2\}$  and  $\{\varepsilon/2\} \in V\{0, \varepsilon/2\}$ ). Or recall that a topological space admitting a compatible locally symmetric quasi-metric is metrizable ([24] 2.32).

Let us also observe that  $\mathcal{U}$  is not stable and  $\mathcal{U}^{-1}$  is not co-stable: consider the D-Cauchy filter generated by the sets  $\{\{0, s\} : 0 < s < \varepsilon\}$  ( $\varepsilon > 0$ ). ■

Quietness can be replaced by uniform regularity in the proposition (because of the mixed-symmetry, see [10]); regularity, however, is not sufficient, see Example 3.3. Corollary 3.3 can be deduced from the proposition (taking (i)  $\Rightarrow$  (ii) from Theorem 4.1 into account): any compact quasi-uniformity is equinormal, hence mixed-symmetric.

## References

- [1] Á. CSÁSZÁR, Extensions of quasi-uniformities, *Acta Math. Acad. Sci. Hungar.*, **37** (1981), 121–145.
- [2] Á. CSÁSZÁR,  $D$ -complete extensions of quasi-uniform spaces, *Act Math. Hungar.*, **64** (1994), 41–54.
- [3] J. DEÁK, Extensions of quasi-uniformities for prescribed bitopologies I, *Studia Sci. Math. Hungar.*, **25** (1990) 45–67.
- [4] J. DEÁK, Extensions of quasi-uniformities for prescribed bitopologies II, *Studia Sci. Math. Hungar.*, **25** (1990) 69–91.
- [5] J. DEÁK, Quasi-uniform extensions for finer topologies, *Studia Sci. Math. Hungar.*, **25** (1990), 97–105.
- [6] J. DEÁK, Extending a quasi-metric, *Studia Sci. Math. Hungar.*, **28** (1993), 106–113.
- [7] J. DEÁK, On the coincidence of some notions of quasi-uniform completeness defined by filter pairs, *Studia Sci. Math. Hungar.*, **26** (1991), 411–413.
- [8] J. DEÁK, A survey of compatible extensions (presenting 77 unsolved problems), *Topology, theory and applications II* (Proc Sixth Colloq, Pécs, 1989), Colloq. Math, Soc. János Bolyai **55**, North-Holland (Amsterdam, 1993).
- [9] J. DEÁK, Extending and completing quiet quasi-uniformities, *Studia Sci. Math. Hungar.* (to appear)
- [10] J. DEÁK, A note on weak symmetry properties of quasi-uniformities, *Studia Sci. Math. Hungar.* (to appear)
- [11] J. DEÁK, A bitopological view of quasi-uniform completeness I, *Studia Sci. Math. Hungar.*, (to appear)
- [12] J. DEÁK, A bitopological view of quasi-uniform completeness II, *Studia Sci. Math. Hungar.*, (to appear)
- [13] J. DEÁK, A bitopological view of quasi-uniform completeness III, *Studia Sci. Math. Hungar.*, (to appear)
- [14] J. DEÁK, Quasi-uniform completeness and neighbourhood filters, *Studia Sci. Math. Hungar.*, (to appear)
- [15] J. DEÁK, Short notes on quasi-uniform spaces I. Uniform local symmetry, *Acta Math. Hungar.*, (to appear)
- [16] J. DEÁK, Short notes on quasi-uniform spaces II. Doubly uniformly strict extensions, *Acta Math. Hungar.*, (to appear)
- [17] D. DOITCHINOV, On completeness of quasi-uniform spaces, *C. R. Acad. Bulg. Sci.*, **41** (1988), 5–8.
- [18] D. DOITCHINOV, A concept of completeness of quasi-uniform spaces, *Topology Appl.*, **38** (1991), 205–217.

- [19] D. DOITCHINOV, Another class of completable quasi-uniform spaces, *C. R. Acad. Bulg. Sci.*, **44** (1991), 5–6.
- [20] D. DOITCHINOV, E-completions of quasi-uniform spaces (preprint).
- [21] P. FLETCHER, W. HUNSAKER, Uniformly regular quasi-uniformities, *Topology Appl.*, **37** (1990), 285–291.
- [22] P. FLETCHER, W. HUNSAKER, Completeness using pairs of filters, *Topology Appl.*, **44** (1992), 149–155.
- [23] P. FLETCHER, W. HUNSAKER, Symmetry conditions in terms of open sets, *Topology Appl.*, **45** (1992), 39–47.
- [24] P. FLETCHER, W. F. LINDGREN, *Quasi-uniform spaces*, Lecture Notes in Pure Appl. Math., **27**, Marcel Dekker (New York, 1982).
- [25] J. R. ISBELL, *Uniform spaces*, Math Surveys **12**, Amer Math. Soc. (Providence, 1964).
- [26] R. D. KOPPERMAN, Total boundedness and compactness for filter pairs, *Ann. Univ. Sci. Budapest, Sect. Math.*, **33** (1990), 25–30.
- [27] H-P. A. KÜNZI, Functorial admissible quasi-uniformities on topological spaces, *Topology Appl.*, **43** (1992), 27–36.
- [28] H-P. A. KÜNZI, Nonsymmetric topology, *Topology '93* (Proc. Seventh Colloq., Szekszárd, 1993), Bolyai Society Mathematical Studies **4**, János Bolyai Math. Soc. Budapest, (to appear).
- [29] H-P. A. KÜNZI, H. J. K. JUNNILA, Stability in quasi-uniform spaces and the inverse problem, *Topology Appl.*, **49** (1993), 175–189.
- [30] H-P. A. KÜNZI, A. LÜTHY, Dense subspaces of quasi-uniform spaces, *Studia Sci. Math. Hungar.*, (to appear).
- [31] H-P. A. KÜNZI, M. MRŠEVIČ, I. L. REILLY, M. K. VAMANAMURTHY, Convergence, precompactness and symmetry in quasi-uniform spaces, *Math. Japon.*, **38** (1993), 239–253.
- [32] W. F. LINDGREN, P. FLETCHER, A construction of the pair completion of a quasi-uniform space, *Canad. Math. Bull.*, **21** (1978), 53–59.
- [33] H. RENDER, Nonstandard methods of completing quasi-uniform spaces, *Topology Appl.* (to appear).
- [34] S. ROMAGUERA, Left  $K$ -complete quasi-uniform spaces, *Seminarberichte Fach. Math. Univ. Hagen*, **48** (1994), 101–114.
- [35] M. B. SMYTH, Completeness of quasi-uniform and syntopological spaces, *J. London Math. Soc.* (to appear).
- [36] P. SÜNDERHAUF, The Smyth completion of a quasi-uniform space, *Semantics of programming languages and model theory*, Gordon and Breach (New York, 1993), 189–212.
- [37] P. SÜNDERHAUF, Quasi-uniform completeness in terms of Cauchy nets, *Acta. Math. Hungar.* (to appear).

# CONVEXITY, MINIMAX THEOREMS AND THEIR APPLICATIONS

By

I. JOÓ\* and G. KASSAY

Department of Analysis, Eötvös Loránd University Budapest, and  
Faculty of Mathematics, Babeş-Bolyai University Cluj

(Received September 16, 1994)

Dedicated to Professor P. ERDŐS for his 80-th birthday

## 0. Introduction

The results concerning minimax theorems have many applications in several fields of pure and applied mathematics. Our purpose is to study the following two problems: let  $X$  and  $Y$  be nonempty sets, and let  $f, g : X \times Y \rightarrow \mathbb{R}$ ,  $\varphi : X \rightarrow Y$  be given functions such that  $f \leq g$  on  $X \times Y$ . Under which hypothesis on  $X$ ,  $Y$ ,  $f$ ,  $g$  and  $\varphi$  holds

$$(A) \quad \inf_{y \in Y} \sup_{x \in X} f(x, y) \leq \sup_{x \in X} \inf_{y \in Y} g(x, y)$$

or

$$(B) \quad \inf_{y \in Y} \sup_{x \in X} f(x, y) \leq \sup_{x \in X} g(x, \varphi(x)).$$

Results related to problem (A) (called minimax theorems, since in case  $f = g$  we have equality) have been obtained by several authors. These can be applied in game theory, mathematical economics and optimization theory. Problem (B) has been studied just in case  $X = Y$  and  $\varphi(x) = x$ . These results (called minimax inequalities) are useful for studying variational inequalities, differential equations, potential theory, etc. It is easy to see that (A) implies (B), therefore (B) can be usually stated under weaker hypothesis.

The aim of this note is to state theorems which extend some of the most important known results concerning problems (A) and (B), to study the

---

\* Research supported by OTKA No. T 007546 Hungary

connection between them and to give some applications. The first section is concerned on problem (B). Using a convexity concept introduced by JOÓ [14], Theorem 1.1 contains, in particular KY FAN [4], while Theorem 1.3 those of BRÉZIS–NIRENBERG–STAMPACCHIA [1], SIMONS [25] and KY FAN [5]. Here we use a method related to KNASTER–KURATOWSKI–MAZURKIEWICZ (KKM in short) theorem which first appeared in KASSAY–KOLUMBÁN [17]. In section 2 we apply our results from section 1 and deduce Brouwer’s fixed point theorem for this kind of convexity (pseudoconvex spaces). Further, we show that Brouwer’s fixed point theorem fails in a more general convexity structure (interval spaces, introduced by STACHÓ [28]). Although problem (A) can be treated in interval spaces (see the results of JOÓ [14], [15], STACHÓ [28], KOMORNIK [21]) our counterexample shows that problem (B) cannot be attached in interval spaces using the argument of section 1 (which uses Brouwer’s fixed point theorem). This question remains open (i.e., problem (B) in interval spaces).

In section 3 we are concerned on problem (A). It is shown how one can prove a general minimax theorem (Theorem 3.1) using three methods: the *KKM-method* used in section 1, the *level set method* discovered by JOÓ [11] and the *mixture* of the *level set* and *cone method* (discovered in the strongest form by JOÓ [13]) given by KASSAY [19].

Theorem 4.1 (section 4) is deduced using Theorem 1.3 and extends a result on variational inequalities due to BROUWER [2].

Recently, SIMONS [27] introduced the concept of *upward-downward* function in order to establish minimax theorems. He asked whether his result remains true for two functions. In section 5 we answer in the negative to both of Simons conjectures. On the other hand, we prove these conjectures under an additional hypothesis.

Finally, in section 6 we state generalized Kuhn–Tucker theorems for cone constrained and inequality constrained (nonlinear) optimization problems.

For other results concerning problems (A) and (B) see also [3], [7], [8], [9], [10], [16], [18], [29].

The results of this paper were first announced in [15].

## 1. Minimax inequalities on pseudoconvex spaces

The aim of this section is to give a minimax inequality without linear structure which, in particular implies KY FAN [5], BRÉZIS–NIRENBERG–STAMPACCHIA [1] and SIMONS [25]. Instead of the usual convexity in linear spaces, we consider a convexity concept introduced by JOÓ [14] which doesn’t need the linear structure. First we recall the following definitions:

DEFINITION 1.1 (KOMIYA [20]). Let  $X$  be a nonempty set. A mapping  $h : 2^X \rightarrow 2^X$  is said to be *convex hull operation* if it satisfies the following conditions:  $h(\emptyset) = \emptyset$ ,  $h(\{x\}) = \{x\}$ ,  $h(A) = \bigcap \{h(F) : F \subset A \text{ is a finite set}\}$ ,  $h(h(A)) = A (A \subset X)$ . One says that  $A \subset X$  is convex, if  $h(A) = A$ .

DEFINITION 1.2 (JOÓ [14]). A triple  $(X, h, \mathcal{F})$  is called *pseudoconvex space* if

- (1)  $X$  is a topological space and  $h$  is a convex hull operation on it;
- (2)  $\mathcal{F} = \{\psi_F : F \subset X \text{ is finite}\}$ ,  $\psi_F : \Delta^n \rightarrow h(F)$ , ( $n = \text{card } F - 1$ ) is a continuous mapping of  $\Delta^n$  onto  $h(F)$ , where  $\Delta^n$  denotes the standard simplex of  $\mathbb{R}^n$ ;
- (3) For each finite  $F \subset X$ ,  $\psi_F$  is convex hull preserving, i.e., if  $\Delta^n = (e_0, e_1, \dots, e_n)$  and  $F = \{x_0, x_1, \dots, x_n\}$  for each subsimplex  $(e_{i_0}, e_{i_1}, \dots, e_{i_k}) \subset (e_0, e_1, \dots, e_n)$  we have  $\psi_F((e_{i_0}, \dots, e_{i_k})) = h(x_{i_0}, \dots, x_{i_k})$ .

In [12] can be found an example for pseudoconvex space (see the next section too), where  $X = \mathbb{R}^n$  and which differs from the usual convexity. We also mention that pseudoconvex spaces include convex spaces in the sense of KOMIYA [20].

Now we state a result which extend the well known Ky Fan's intersection theorem [4] (often called the Ky Fan's lemma). For, we first need the following definition:

DEFINITION 1.3. Let  $(X, h, \mathcal{F})$  be a pseudoconvex space and let  $X_0, X_1, \dots, X_n$  be a family of subsets of  $X$ . The set  $\{x_0, x_1, \dots, x_n\} \subset X$  is said to be *KKM-selection* for  $X_0, \dots, X_n$  if for any subset  $j \subset \{0, 1, \dots, n\}$  we have

$$h(\{x_j : j \in J\}) \subset \bigcup_{j \in J} X_j.$$

Note that this concept differs from that in Ky Fan's lemma even in case when the convexity above reduces to the usual one, since the elements  $x_0, \dots, x_n$  are not necessarily distinct. For instance we could have  $x_0 = x_1 = \dots = x_n$ ; then

$$\bigcap_{i=0}^n X_i \neq \emptyset.$$

Moreover, we have the following property:  $\bigcap_{i=0}^n X_i \neq \emptyset$  if and only if there exists a KKM-selection for  $X_1, \dots, X_n$  which contains one element. The following theorem gives a sufficient condition for the existence of such a KKM-selection.

THEOREM 1.1. Let  $(X, h, \mathcal{F})$  be a pseudoconvex space,  $I$  a nonempty set and  $(X_i)_{i \in I}$  a family of compact subsets of  $X$ . Then  $\bigcap_{i \in I} X_i \neq \emptyset$  if and

only if for each finite subset  $i_0, i_1, \dots, i_n \in I$ , the family  $(X_{i_k})_{0 \leq k \leq n}$  admits a KKM-selection.

PROOF. The necessity is obvious. We prove the sufficiency. By compactness, it is enough to show the "finite intersection property" for  $(X_i)_{i \in I}$ . Let  $X_{i_0}, X_{i_1}, \dots, X_{i_n}$  ( $i_0, i_1, \dots, i_n \in I$ ) be given and let  $\{x_0, x_1, \dots, x_n\}$  be a KKM-selection of them. Without loss of generality we may suppose that  $x_0, x_1, \dots, x_n$  are distinct. Otherwise, if, say  $x_l = x_p$  for  $l \neq p$  then the corresponding sets  $X_{i_l}$  and  $X_{i_p}$  could be changed with their (nonvoid) intersection.

Let  $E_k = \psi_F^{-1}(X_{i_k} \cap h(F))$ ,  $0 \leq k \leq n$ , where  $F = \{x_0, x_1, \dots, x_n\}$ . Then for  $\{e_0, e_1, \dots, e_n\} = \Delta^n$  we have  $\text{co}\{e_j : j \in J\} \subset \bigcup_{j \in J} E_j$  for each  $J \subset \{0, 1, \dots, n\}$ .

Indeed, let  $z \in \text{co}\{e_j : j \in J\}$ . Then  $\psi_F(z) \in h(\{x_j : j \in J\}) \subset \bigcup_{j \in J} X_{i_j}$ . Let

$j_0 \in J$  such that  $\psi_F(z) \in X_{i_{j_0}}$ . Therefore,  $z \in \psi_F^{-1}(X_{i_{j_0}} \cap h(F)) = E_{j_0}$ . By a variant of the classical KKM theorem (see [17], Lemma 2.1) it follows that  $\bigcap_{k=0}^n E_k \neq \emptyset$ ; thus  $\bigcap_{k=0}^n X_{i_k} \neq \emptyset$ . This completes the proof.

In the following we give an extension to an intersection theorem due to BRÉZIS–NIRENBERG–STAMPACCHIA ([1], Lemma 1).

THEOREM 1.2. Let  $(X, h, \mathcal{F})$  be a pseudoconvex space,  $I$  a nonempty set and  $\varphi : I \rightarrow X$  a given function. Let  $(X_i)_{i \in I}$  be a family of subsets of  $X$  for which

- (4) there is an index  $i_0 \in I$  such that  $\overline{X_{i_0}}$  (the closure of  $X_{i_0}$ ) is compact;
- (5) for each  $i \in I$  and for each finite subset  $F$  of  $X$ ,  $X_i \cap h(F)$  is closed;
- (6) for each finite subset  $F$  of  $X$  we have

$$\overline{\left( \bigcap_{i \in I_F} X_i \right)} \cap h(F) = \left( \bigcap_{i \in I_F} X_i \cap h(F) \right),$$

where  $I_F := \varphi^{-1}(h(F))$  (may be empty);

- (7) for each finite subset  $J$  of  $I$  we have

$$h(\{\varphi(j) : j \in J\}) \subset \bigcap_{j \in J} X_j.$$

Then  $\bigcap_{i \in I} X_i \neq \emptyset$ .

PROOF. Let  $\Phi = \{F : F \subset X \text{ is a finite set, } \varphi(i_0) \in F\}$  and  $\mathcal{H} = \{h(F) : F \in \Phi\}$ . For simplicity, denote the family  $\mathcal{H}$  by  $(H_s)_{s \in S}$ . On  $S$  introduce the

ordering relation as follows:  $s \leq t$  ( $s, t \in S$ ) iff  $H_s \subset H_t$ . Let  $s \in S$  be fixed and  $I_s := \varphi^{-1}(H_s)$  (Since  $i_0 \in I_s$  for each  $s \in S$ ,  $I_s$  is nonempty). Define

$$E_i^s := X_i \cap H_s, \quad i \in I_s$$

Then we have

- (8)  $E_i^s$  is nonempty and closed for each  $i \in I_s$ ;
- (9)  $E_{i_0}^s$  is compact;
- (10) For each finite subset  $J$  of  $I_s$  we have

$$h(\{\varphi(j) : j \in J\}) \subset \bigcup_{j \in J} E_j^s.$$

Properties (8) and (9) are obvious. For (10) let  $J$  be a finite subset of  $I_s$ . It is clear by (7) that  $h(\{\varphi(j) : j \in J\}) \subset \bigcup_{j \in J} X_j$ . On the other hand, using the axioms of  $h$  we have  $h(\{\varphi(j) : j \in J\}) \subset H_s$ . Applying Theorem 1.1 (without loss of generality we may assume that  $E_i^s$  are compact,  $i \in I_s$ ),  $\bigcap_{i \in I_s} E_i^s \neq \emptyset$ . For each  $s \in S$ , choose an element  $u_s \in \bigcap_{i \in I_s} E_i^s$  and let  $K_s := \bigcup_{t \geq s} \{u_t\}$ . We have  $K_s \subset \overline{X_{i_0}}$ ; further for  $s_1, s_2 \in S$  there exists  $s_3 \in S$  with  $K_{s_3} \subset K_{s_1} \cap K_{s_2}$ . Using compactness of  $\overline{X_{i_0}}$  it follows that  $\bigcap_{s \in S} \overline{K_s} \neq \emptyset$ . Let  $\bar{x} \in \bigcap_{s \in S} \overline{K_s}$ . We show that  $\bar{x} \in \bigcap_{i \in I} X_i$ . By definition,  $K_s \subset \bigcap_{i \in I_s} X_i$ . Choose an index  $s_0 \in S$  such that  $\bar{x} \in H_{s_0}$ . Let  $i \in I$  be arbitrarily chosen and  $s \geq s_0$  such that  $\varphi(i) \in H_s$ . Then by (6)

$$\bar{x} \in \overline{K_s} \cap H_s \subset \overline{\left( \bigcap_{i \in I_s} X_i \right)} \cap H_s = \left( \bigcap_{i \in I_s} X_i \right) \cap H_s.$$

Thus  $\bar{x} \in X_i$ . This completes the proof.

Our next purpose is to establish a minimax inequality for pseudoconvex spaces. First we need the following definition.

A pair  $(X, h)$  is said to be *convex space* if  $X$  is a point set and  $h : 2^X \rightarrow 2^X$  is a convex hull operation (see Definition 1.1).

**THEOREM 1.3** *Let  $(X, h_1)$  be a convex space and  $(Y, h_2, \mathcal{F})$  be a pseudoconvex space. Let  $f, g : X \times Y \rightarrow \mathbb{R}$  such that  $f \leq g$  on  $X \times Y$ , let  $\varphi : X \rightarrow Y$  be a function and  $a := \sup_{x \in X} g(x, \varphi(x))$ . Suppose*

- (11) *for each finite subset  $X_0$  of  $X$ ,  $h_2(\{\varphi(x) : x \in X_0\}) \subset \varphi(h_1(X_0))$ ;*
- (12)  *$f$  is l.s.c. in its second variable on  $h_2(Y_0)$ , for each finite  $Y_0 \subset Y$ ;*

- (13)  $g$  is quasiconcave in its first variable, i.e. the set  $\{x \in X : g(x, y) > \beta\}$  is convex for each  $y \in Y$  and  $\beta \in \mathbb{R}$ ;
- (14) for each finite subset  $Y_0 \subset Y$  and for each filter  $y_\alpha \in Y$  converging to  $y \in h_2(Y_0)$ , we have:  $f(x, y_\alpha) \leq a$  for each  $x \in I_0$  implies  $f(x, y) \leq a$  for each  $x \in I_0$ , where  $I_0 := \varphi^{-1}(h_2(Y_0))$ ;
- (15) there is a compact subset  $D$  of  $Y$  and  $x_0 \in X$  such that  $f(x_0, y) > a$  for each  $y \in Y \setminus D$ .

Then there exists  $y_0 \in Y \cap D$  such that  $f(x, y_0) \leq a$  for all  $x \in X$ . In particular,  $\inf_{y \in Y} \sup_{x \in X} f(x, y) \leq \sup_{x \in X} g(x, \varphi(x))$ .

REMARK. Assumptions (14) and (15) are clearly satisfied if  $Y$  is compact and  $f$  is l.s.c. in its second variable on  $Y$ .

PROOF. For each  $x \in X$  define the sets

$$Y_x := \{y \in Y : f(x, y) \leq a\}.$$

Using Theorem 1.2, we show that  $\bigcap_{x \in X} Y_x \neq \emptyset$ . It is easy to see that properties (4), (5) and (6) follow by (15), (12) and (14) respectively, with  $I := X$ ,  $Y_x := X_i$ . We have to verify that (11) and (13) imply (7). Let  $x_1, x_2, \dots, x_n \in X$  and suppose  $h_2(\{\varphi(x_i) : i \in \{1, \dots, n\}\}) \not\subset \bigcup_{i=1}^n Y_{x_i}$ . Then for some  $y \in h_2(\{\varphi(x_i) : i \in \{1, \dots, n\}\})$  we have  $f(x_i, y) > a$  for each  $i \in \{1, 2, \dots, n\}$ . By (11),  $y \in \varphi(h_1(\{x_1, \dots, x_n\}))$ , hence there exists  $z \in h_1(\{x_1, \dots, x_n\})$  such that  $y = \varphi(z)$ . Since  $g(x_i, y) > a$  for each  $i \in \{1, \dots, n\}$ , by (13),  $g(z, y) > a$ , i.e.,  $g(z, \varphi(z)) > a$ , which contradicts the definition of  $a$ .

This completes the proof. ■

COROLLARY 1.1 (BRÉZIS–NIRENBERG–STAMPACCHIA [1]). Let  $Z$  be a closed convex subset of a Hausdorff topological vector space  $E$  and let  $f : Z \times Z \rightarrow \mathbb{R}$  such that

- (16)  $f(x, x) \leq 0$  for each  $x \in Z$ ;
- (17)  $f$  is l.s.c. in its second variable on the intersection of  $Z$  with any finite dimensional subspace of  $E$ ;
- (18)  $f$  is quasiconcave in its first variable;
- (19) Whenever  $C$  is a convex subset of  $Z$  and  $y_\alpha$  is a filter converging to  $y \in C$ , then  $f(x, y_\alpha) \leq 0$  for every  $x \in C$  implies  $f(x, y) \leq 0$  for every  $x \in C$ ;
- (20) There is a compact subset  $D$  of  $E$  and  $x_0 \in D \cap Z$  such that  $f(x_0, y) > 0$  for  $y \in Z \setminus D$ .

Then there exists  $y_0 \in D \cap Z$  such that  $f(x, y_0) \leq 0$  for all  $x \in Z$ . In particular,  $\inf_{y \in Z} \sup_{x \in Z} f(x, y) \leq 0$ .

PROOF. Apply Theorem 1.3 with  $X = Y = Z$ ,  $f = g$ ,  $\varphi(x) = x$  for each  $x \in X$  and  $h_1(A) = h_2(A) = \text{co } A$  if  $A \subset X$ .

COROLLARY 1.2 (SIMONS, [25]). Let  $Z$  be a nonempty convex subset of a topological vector space, let  $f : Z \times Z \rightarrow \mathbb{R}$  be l.s.c. in its second variable,  $g : Z \times Z \rightarrow \mathbb{R}$  quasiconcave in its first variable and  $f \leq g$  on  $Z \times Z$ . Then  $\min_{y \in Z} \sup_{x \in Z} f(x, y) \leq \sup_{x \in Z} g(x, x)$ .

COROLLARY 1.3 (KY FAN [5]). Let  $Z$  be a nonempty compact convex subset of a topological vektor space and  $f : Z \times Z \rightarrow \mathbb{R}$  be quasiconcave in its first variable and l.s.c. in its second variable. Then  $\min_{y \in Z} \sup_{x \in Z} f(x, y) \leq \sup_{x \in Z} f(x, x)$ .

## 2. Application in fixed point theory

Using the results above, in this section we prove that Brouwer's fixed point theorem holds under convexity introduced by I. LOÓ and L. L. STACHÓ [12]. This is a kind of pseudoconvexity defined in  $\mathbb{R}^n$ .

Let  $x = (x_0, \dots, x_n)$ ,  $y = (y_0, \dots, y_n) \in \mathbb{R}^{n+1}$ . The interval joining  $x$  and  $y$  will be defined as a polygon with at most  $n + 1$  pairwise orthogonal segments as follows (see also [12] and [14]). If  $x_n \geq y_n$  then let  $I_n = \{(x_0, \dots, x_{n-1}, t) : y_n \leq t \leq x_n\}$ . If  $x_n \leq y_n$  then let  $I_n = \{(y_0, \dots, y_{n-1}, t) : x_n \leq t \leq y_n\}$ . In the first case we get  $I_{n-1}$  analogously to  $I_n$ ; if, for example  $x_{n-1} \leq y_{n-1}$  then  $I_{n-1} = \{(y_0, \dots, y_{n-2}, t, y_n) : x_{n-1} \leq t \leq y_{n-1}\}$ ; if  $x_{n-1} > y_{n-1}$  then  $I_{n-1} = \{(x_0, \dots, x_{n-2}, t, y_n) : y_{n-1} \leq t \leq x_{n-1}\}$ . In the second case ( $x_n \leq y_n$ ) we construct analogously  $I_{n-1}$  and in the third step  $I_{n-2}$ , etc. The segments  $I_0, I_1, \dots, I_n$  parallel to the axis will join  $x$  and  $y$ . A set  $A \subset \mathbb{R}^n$  is convex if  $x, y \in A$  implies that the interval joining  $x$  and  $y$  is contained in  $A$ . It can be shown that  $\mathbb{R}^n$  with this convexity is a pseudoconvex space in the sense of Definition 1.2 (see JOÓ [14]).

In which follows, we use the notions of convex set, quasiconvex (quasi-concave) function in the sense of JOÓ-STACHÓ [12], described above.

LEMMA 2.1. The Chebyshev norm in  $\mathbb{R}^n$ , i.e.  $\|x\|_C = \max\{|x_1|, \dots, |x_n|\}$  for  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$  is a quasiconvex function.

PROOF. It is easy to see that each Chebyshev sphere centered at the origin is a convex set. Let  $x, y \in \mathbb{R}^n$  and  $z \in [x, y]$  (the interval joining  $x$  and  $y$ ). Let  $R = \max\{\|x\|_C, \|y\|_C\}$ . Then the sphere centered at the origin with radius  $R$  contains  $z$ . Hence  $\|z\|_C \leq R = \max\{\|x\|_C, \|y\|_C\}$ .

We mention that this statement is false for the Euclidean norm. For instance let  $x = (-2, 0)$ ,  $y = (1, -1) \in \mathbb{R}^2$ . Then  $z = (-2, -1) \in [x, y]$  and  $\|z\| > \max\{\|x\|, \|y\|\}$  where  $\|\cdot\|$  denotes the Euclidean norm.

THEOREM 2.1 (See also KY FAN [6]). *Let  $X$  be a compact convex subset of  $\mathbb{R}^n$  and  $\bar{f}, \hat{f} : X \rightarrow \mathbb{R}^n$  be two continuous mappings. Suppose*

$$\|x - \hat{f}(x)\|_C \geq \|x - \bar{f}(x)\|_C \quad \text{for each } x \in X.$$

*Then there exists a point  $x_0 \in X$  such that*

$$(21) \quad \|x - \hat{f}(x)\|_C \geq \|x_0 - \bar{f}(x_0)\|_C \quad \text{for each } x \in X.$$

PROOF. Take  $f(x, y) = \|y - \bar{f}(y)\|_C - \|x - \hat{f}(y)\|_C$ . It is clear that  $f$  is continuous in its second variable and (using Lemma 2.1) quasiconcave in its first variable. Applying Theorem 1.3 with  $X = Y$ ,  $h_1 = h_2$ ,  $f = g$ ,  $\varphi(x) = x$  for each  $x \in X$  we obtain (21).

REMARK. If in addition  $X$  is invariant under  $\hat{f}$ , i.e.  $\hat{f}(X) \subset X$ , then (21) is equivalent with  $\bar{f}(x_0) = x_0$ .

In [28] L. L. STACHÓ introduced the following convexity structure: the pair  $(X, [\cdot, \cdot])$  is called an *interval space* if  $X$  is a topological space and  $[\cdot, \cdot] : X \times X \rightarrow 2^X$  is a mapping such that  $x_0, x_1 \in [x_0, x_1]$  and  $[x_0, x_1]$  is a connected set for each  $x_0, x_1 \in X$ . A set  $A \subset X$  is called *convex* if  $x_0, x_1 \in A$  implies  $[x_0, x_1] \subset A$ . We can define the concept of quasiconvex (quasiconcave) function  $f : X \rightarrow \mathbb{R}$  in a similar way as above. It is easy to see that every pseudoconvex space  $(X, h, \mathcal{F})$  is an interval space, where the interval  $[x_0, x_1]$  is defined to be  $\psi_F(\Delta^1)$ , with  $F = \{x_0, x_1\} \subset X$ . It is natural to ask whether Brouwer's fixed point theorem remains true for interval spaces. We answer to this question in the negative. Let  $X = [0, 1] \subset \mathbb{R}$ . Introduce the following topology in  $X$  :  $A \subset X$  is closed iff  $A$  is finite or  $A = X$ . It is clear that  $X$  is a compact topological space with the topology above. For each  $x_1, x_2 \in X$  let the interval joining  $x_1$  and  $x_2$  be the whole space  $X$ . Then every interval is closed, connected and contains its endpoints. Now consider the function  $f : X \rightarrow X$  defined by  $f(x) = x + 1/2$  for  $x \in [0, 1/2]$ ,  $f(x) = x - 1/2$  for  $x \in ]1/2, 1[$  and  $f(1) = 0$ . Then  $f$  is bijective, without fixed points. It remains to prove that  $f$  is continuous. Let  $x \in [0, 1]$  be an arbitrary element and let  $V$  be a neighbourhood of  $f(x)$ . By definition  $V$  must contain each element of  $X$  except a finite number of them. Let  $V = [0, 1] \setminus \{y_1, y_2, \dots, y_n\}$

and  $x_i := f^{-1}(y_i)$ ,  $(1 \leq i \leq n)$ . Then the set  $U := [1, 0] \setminus \{x_1, x_2, \dots, x_n\}$  is a neighborhood of  $x$  for which  $f(U) \subset V$ . Thus  $f$  is continuous at  $x$ .

In section one we have seen that results concerning minimax inequalities (problem (B)) can be extended for pseudoconvex spaces. The proof is based on KKM methods which use Brouwer’s fixed point theorem. As it could be seen, these theorem fails in interval spaces. Therefore, in these spaces, the results in section one can’t be proved using the argument above. However, it would be interesting to attack problem (B) in case of interval spaces.

### 3. Minimax theorems

In this section we are concerned on problem (A). We start with the following definition.

DEFINITION 3.1. Let  $(X, h, \mathcal{F})$  be a pseudoconvex space. A function  $f : X \rightarrow \mathbb{R}$  is said to be *convex* if  $f \circ \psi_F$  is convex (in the usual sense) for each finite  $F \subset X$ .  $f$  is said to be *concave* if  $-f$  is convex. Recall that  $f$  is *quasiconcave* if  $-f$  is quasiconvex. It is clear that if  $f$  is convex (concave) then it is also quasiconvex (quasiconcave).

THEOREM 3.1. Let  $(X, h_1, \mathcal{F}_1)$  be a pseudoconvex space,  $(Y, h_2, \mathcal{F}_2)$  be a compact pseudoconvex space. Let  $f : X \times Y \rightarrow \mathbb{R}$  be u.s.c. and concave in its first variable and l.s.c. and convex in its second variable. Then

$$\min_{y \in Y} \sup_{x \in X} f(x, y) = \sup_{x \in X} \inf_{y \in Y} f(x, y).$$

We shall deduce Theorem 3.1 in three ways: the first one uses the results above concerning minimax inequalities (therefore, the proof is based on KKM methods and its generalizations); the second one uses the Hahn–Banach method (separation of convex sets), while the third one will be the method of “level sets” discovered by I. JOÓ [11] which uses neither KKM nor Hahn–Banach’s theorems.

For the beginning, we give a minimax theorem for two functions which follows by Theorem 1.3 and which clearly implies Theorem 3.1. This contains, in particular SIMONS [25], Theorem 1.4, and H. NIKAIDÓ [24].

THEOREM 3.2. Let  $(X, h_1, \mathcal{F}_1)$  and  $(Y, h_2, \mathcal{F}_2)$  be two pseudoconvex spaces with  $Y$  compact,  $f, g : X \times Y \rightarrow \mathbb{R}$  with  $f \leq g$  on  $X \times Y$  such that (22)  $f$  is l.s.c. in its second variable and quasiconcave in its first variable;

(23)  $g$  is u.s.c. in its first variable and quasiconvex in its second variable. Then

$$\min_{y \in Y} \sup_{x \in X} f(x, y) \leq \sup_{x \in X} \inf_{y \in Y} g(x, y).$$

PROOF. Suppose first that  $X$  is compact. If the result were false, we could choose  $r \in \mathbb{R}$  such that

$$(24) \quad \min_y \sup_x f(x, y) > r > \sup_x \inf_y g(x, y).$$

Let  $Z = X \times Y$ . Then  $(Z, h, \mathcal{F})$  is a compact pseudoconvex space, where  $h$  and  $\mathcal{F}$  are defined as follows: if  $F = \{(x_0, y_0), \dots, (x_n, y_n)\}$ ,  $F_1 = \{x_0, \dots, x_n\}$ ,  $F_2 = \{y_0, \dots, y_n\}$  then  $\psi_F := \psi_{1, F_1} \times \psi_{2, F_2} : \Delta^n \rightarrow Z$  (i.e.  $\psi_F(t) = (\psi_{1, F_1}, \psi_{2, F_2}(t))$  if  $t \in \Delta^n$ ) and  $\mathcal{F} = \{\psi_F : F \subset Z \text{ is finite}\}$ . The convex hull  $h(F)$  of a finite set  $F \subset Z$  is defined by  $\psi_F(\Delta^n)$  ( $n+1 = \text{card } F$ ) and the convex hull of any set  $A \subset Z$  is determined by  $h(A) = \bigcap \{h(F), F \subset A \text{ is a finite set}\}$ .

Consider the function  $\Phi : Z \times Z \rightarrow \mathbb{R}$  defined by

$$\Phi((x, y), (\hat{x}, \hat{y})) = \min\{f(x, \hat{y}) - r, r - g(\hat{x}, y)\}.$$

It is easy to see that  $\Phi$  verifies the conditions of Theorem 3.1 (with  $f = g := \Phi$ ). We have  $\Phi((x, y), (x, y)) \leq 0$  for each  $(x, y) \in Z$ . Thus, there exists  $(\hat{x}, \hat{y}) \in Z$  such that  $\Phi((x, y), (\hat{x}, \hat{y})) \leq 0$  for each  $(x, y) \in Z$ , or, in other words  $f(x, \hat{y}) \leq r$  or  $g(\hat{x}, y) \geq r$  for each  $(x, y) \in Z$ . This contradicts to (24). Now, considering the general case, it can be seen that for each  $x_1, \dots, x_n \in X \bigcap_{i=1}^n \{y \in Y : f(x_i, y) \leq \alpha\} \neq \emptyset$  where  $\alpha := \sup_x \inf_y g(x, y)$ . Using the finite intersection property for compact sets, we have  $\bigcap_{x \in X} \{y \in Y : f(x, y) \leq \alpha\} \neq \emptyset$ , as required.

In paper [19], G. KASSAY gave a simple proof for König's minimax theorem [22] based on geometrical properties on  $\mathbb{R}^2$ . König's theorem has extended by S. SIMONS [25] for two functions. We prove Simon's theorem using the method in [19].

**THEOREM 3.3** (S. SIMONS, [25]). *Let  $X$  be a nonempty set,  $Y$  a nonempty compact topological space,  $f, g : X \times Y \rightarrow \mathbb{R}$  with  $f \leq g$  on  $X \times Y$  such that*

*$f$  is l.s.c. and 1/2 convex in its second variable, i.e. for all  $y_1, y_2 \in Y$  there exists  $y_3 \in Y$  such that*

$$(25) \quad f(x, y_3) \leq \frac{f(x, y_1) + f(x, y_2)}{2} \quad \text{for each } x \in X;$$

$g$  is 1/2 concave in its first variable, i.e. for all  $x_1, x_2 \in X$  there exists  $x_3 \in X$  such that

$$(26) \quad g(x_3, y) \geq \frac{g(x_1, y) + g(x_2, y)}{2} \quad \text{for each } y \in Y.$$

Then

$$(27) \quad \min_{y \in Y} \sup_{x \in X} f(x, y) \leq \sup_{x \in X} \inf_{y \in Y} g(x, y).$$

Let  $c_* := \sup_x \inf_y g(x, y)$ . We have the following statement which has proved for one function by I. JOÓ [13].:

LEMMA 3.4. (27) holds iff  $\bigcap_{x \in X} \{y \in Y : f(x, y) \leq c\} \neq \emptyset, \quad c > c_*$ .

The proof is analogous to that of Theorem 2 [13]; we omit the details.

PROOF OF THEOREM 3.3. Since  $Y$  is compact and  $f$  is l.s.c. in its second variable, hence the sets  $H_x^c := \{y : f(x, y) \leq c\}$  ( $c > c_*$ ) are compact. Thus, it is enough to prove that the family of sets  $\{H_x^c : x \in X\}$  ( $c > c_*$ ) has the finite intersection property. We prove that any two sets of this family have nonempty intersection. Suppose the contrary, i.e. that there exists  $c > c_*$  and  $x_1, x_2 \in X$  such that  $H_{x_1}^c \cap H_{x_2}^c = \emptyset$  and define  $p : Y \rightarrow \mathbb{R}^2$  by  $p(y) = (c - f(x_1, y), c - f(x_2, y))$ . If  $K = \{(s, t) \in \mathbb{R}^2 : s \geq 0, t \geq 0\}$  then  $p(Y) \cap K = \emptyset$ . We show that  $\text{cop}(Y) \cap \text{int} K = \emptyset$  ( $\text{co}A$  denotes the standard convex hull of  $A \subset \mathbb{R}^2$ ). For this, suppose that there exist  $\lambda_1, \dots, \lambda_k \in [0, 1]$

with  $\sum_{i=1}^k \lambda_i = 1$  and  $y_0, y_2, \dots, y_k \in Y$  such that  $\sum_{i=1}^k \lambda_i p(y_i) \in \text{int} K$ . Using (25)

we can choose  $\bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_k$  with  $\bar{\lambda}_i \geq 0 \quad (1 \leq i \leq k)$  and  $\sum_{i=1}^k \bar{\lambda}_i = 1$  such that

$\sum_{i=1}^k \bar{\lambda}_i p(y_i) \in K$  and for which there exists  $\bar{y} \in Y$  such that  $p(\bar{y}) - \sum_{i=1}^k \bar{\lambda}_i p(y_i) \in$

$\in K$ . This means that  $p(\bar{y}) \in K$ , which contradicts the hypothesis. By the well-known separation theorem of Hahn-Banach in  $\mathbb{R}^2$ , there exists a line which separates the sets  $\text{cop}(Y)$  and  $K$ . That is, there exists  $b = (b_1, b_2) \in K$  with  $b_1 + b_2 = 1$  such that  $\langle u, b \rangle \leq 0$  for all  $u \in p(Y)$  or, in other words  $b_1 f(x_1, y) + b_2 f(x_2, y) \geq c$  for every  $y \in Y$ . Let  $c_1 \in \mathbb{R}$  such  $c_* < c_1 < c$  and  $d := c_1 - c$ . Then we have  $b_1 [c_1 - f(x_1, y)] + b_2 [c_1 - f(x_2, y)] \leq d$  for every  $y \in Y$ , hence the set  $p_1(Y)$  is separated from  $K$  by the line  $b_1 s + b_2 t = d$ , where  $p_1(y) = (c_1 - f(x_1, y), c_1 - f(x_2, y))$  ( $d < 0$ ). Since  $f(x_1, \cdot)$  and  $f(x_2, \cdot)$  are l.s.c. on  $Y$ , hence there exist  $\alpha, \beta > 0$  such that  $p_1(Y) \subset (-\infty, \alpha] \times (-\infty, \beta]$ . The

line  $b_1s + b_2t = d$  intersects at least one of the lines  $s = \alpha$  and  $t = \beta$ ; suppose that it intersects the second one. It is clear then, that the line  $b_1\beta s + (-d + b_2\beta)t = 0$  separates  $p_1(Y)$  and  $K$ . Let  $g_1(y) = (c_1 - g(x_1, y), c_1 - g(x_2, y))$ . Since  $p_1(y) - g_1(y) \in K$  for every  $y \in Y$ , each line which separates  $p_1(Y)$  and  $K$ , separates  $g_1(Y)$  and  $K$  too. Using (26), choose  $\mu \in [0, 1]$  and  $x_\mu \in X$  such that  $g(x_\mu, y) \geq \mu g(x_1, y) + (1 - \mu)g(x_2, y)$  for every  $y \in Y$  and such that  $\mu s + (1 - \mu)t = 0$  separates  $g_1(Y)$  and  $K$  or, in other words  $\mu[c_1 - g(x_1, y)] + (1 - \mu)[c_1 - g(x_2, y)] \leq 0$  for every  $y \in Y$ . Therefore,  $g(x_\mu, y) \geq c_1$  for every  $y \in Y$  which leads to  $\sup_x \inf_y g(x, y) \geq c_1 > c_*$ . This is a contradiction.

Hence any two sets of the family  $\{H_x^c : x \in X\}$  have nonempty intersection.

(In order to prove that for any  $c > c_*$  and  $x_1, \dots, x_n \in X$  we have  $\bigcap_{i=1}^n H_{x_i}^c \neq \emptyset$  we use induction. For the details, see [13]). This completes the proof.

Observe that Theorem 3.3 implies Theorem 3.1 (note that 1/2 convexity implies convexity in a pseudoconvex space):

Finally, Theorem 3.1 follows by a result of I. JOÓ ([14]. Theorem 3).

**THEOREM 3.4** (I. JOÓ, [14]). *Let  $X$  be an interval space,  $Y$  a compact interval space and  $f : X \times Y \rightarrow \mathbb{R}$  such that*

(28)  *$f$  is u.s.c. and quasiconvex in its first variable;*

(29)  *$f$  is l.s.c. and quasiconcave in its second variable.*

*Then*

$$\min_{y \in Y} \sup_{x \in X} f(x, y) = \sup_{x \in X} \min_{y \in Y} f(x, y).$$

#### 4. Application to variational inequalities

In this section we apply the results from section 1 to obtain an existence theorem for variational inequalities. Our statement will be an extension of F. E. BROUWER [2].

Let  $X$  be a nonempty convex subset of a topological vector space  $E$  and let  $T : X \rightarrow E^*$ , where  $E^*$  denotes dual space of  $E$ . The problem is to find  $y \in X$  such that

$$(30) \quad \langle Ty, y - x \rangle \leq 0 \quad \text{for every } x \in X.$$

This problem has been studied by many mathematicians who have solved (30) under some hypothesis on  $X$  and  $T$ . (see for instance [2], [23], [25]). Now we will study a generalized form of (30). Let  $X$  be a convex subset of a

topological vector space  $E$ ,  $(Y, h, \mathcal{F})$  a pseudoconvex space and  $E^*$  the dual space of  $E$ . Let  $T : Y \rightarrow E^*$ ,  $g : Y \rightarrow E$ .

We study the following problem: find an element  $y \in Y$  such that

$$(31) \quad \langle Ty, g(y) - x \rangle \leq 0 \quad \text{for every } x \in X.$$

**THEOREM 4.1.** *Let  $X$  be a convex subset of a topological vector space  $E$ ,  $(Y, h, \mathcal{F})$  a pseudoconvex space,  $T : Y \rightarrow E^*$ ,  $g : Y \rightarrow E$  and  $\varphi : X \rightarrow Y$ .*

*Suppose that*

$$(32) \quad \text{for each } x_1, x_2, \dots, x_n \in X, h\{\varphi(x_1), \dots, \varphi(x_n)\} \subset \varphi(\text{co}\{x_1, \dots, x_n\}).$$

$$(33) \quad \langle T(\varphi(x)), g(\varphi(x)) - x \rangle \leq 0 \text{ for every } x \in X;$$

$$(34) \quad T : Y \rightarrow E^* \text{ and } g : Y \rightarrow E \text{ are continuous};$$

$$(35) \quad \text{there is a compact subset } D \text{ of } Y \text{ and } x_0 \in X \text{ such that } \langle Ty, g(y) - x_0 \rangle > 0 \text{ for every } y \in Y \setminus D.$$

*Then there exists  $y_0 \in D$  solution of (31).*

**PROOF.** Let  $f : X \times X \rightarrow \mathbb{R}$  defined by  $f(x, y) = \langle Ty, g(y) - x \rangle$  ( $x \in X, y \in Y$ ). It is easy to see that  $f$  satisfies the conditions of Theorem 1.3 (considering  $f = g$ ). Since  $\sup_{x \in X} f(x, \varphi(x)) \leq 0$  by (33), there exists  $y_0 \in D$  such that  $f(x, y_0) \leq 0$ , as desired.

**COROLLARY 4.1.** *Suppose that (32), (33), (34) hold and  $Y$  is compact. Then there exists  $y_0 \in Y$ , solution of (31).*

**COROLLARY 4.2** (F. E. BROUWER, [2]). *Let  $X$  be a compact convex subset of a topological vector space  $E$  and  $T : X \rightarrow E^*$  a continuous map. Then there exists  $y_0 \in X$  such that*

$$\langle Ty_0, y_0 - x \rangle \leq 0 \quad \text{for every } x \in X.$$

**PROOF.** Apply Theorem 4.1 for  $X = Y$ ,  $g(x) = \varphi(x) = x$  for every  $x \in X$ .

### 5. Upward-downward functions

#### A disprove for two conjectures of S. Simons

In [27] S. SIMONS has given the following definitions. Let  $X$  and  $Y$  be nonempty sets. A function  $f : X \times Y \rightarrow \mathbb{R}$  is said to be *upward* on  $Y$  if  $\forall \varepsilon > 0 \exists \delta > 0$  such that  $\forall y_1, y_2 \in Y \exists y_3 \in Y$  such that:

$$(36) \quad \begin{cases} \forall x \in X : f(x, y_3) \leq \max\{f(x, y_1), f(x, y_2)\} & \text{and} \\ f(x, y_3) > \max\{f(x, y_1), f(x, y_2)\} - \delta \Rightarrow |f(x, y_1) - f(x, y_2)| < \varepsilon. \end{cases}$$

One says that  $f$  is *downward* on  $X$  if  $\forall \varepsilon > 0 \exists \delta > 0$  such that  $\forall x_1, x_2 \in X, \exists x_3 \in X$  such that:

$$(36') \quad \begin{cases} \forall y \in Y : f(x_3, y) \geq \min\{f(x_1, y), f(x_2, y)\} & \text{and} \\ f(x_3, y) < \min\{f(x_1, y), f(x_2, y)\} + \delta \Rightarrow |f(x_1, y) - f(x_2, y)| < \varepsilon. \end{cases}$$

SIMONS has proved in [27] that if  $f$  is upward on  $Y$ , downward on  $X$ ,  $Y$  is a compact topological space and  $f$  is l.s.c. in its second variable then  $\min_y \sup_x f(x, y) = \sup_x \min_y f(x, y)$ . The question whether this statement remains true in case of two functions remained open. Namely, [27] concludes with the following conjectures:

(37) Suppose that  $f, g : X \times Y \rightarrow \mathbb{R}$ ,  $f$  is upward on  $Y$ ,  $g$  is downward on  $X$ ,  $f \leq g$  on  $X \times Y$ ,  $X_0$  is a nonempty finite subset of  $X$ . Then  $\inf_{y \in Y} \max_{x \in X_0} f(x, y) \leq \sup_{x \in X} \inf_{y \in Y} g(x, y)$ .

(38) Suppose that  $f, g : X \times Y \rightarrow \mathbb{R}$ ,  $f$  is upward on  $Y$ ,  $g$  is downward on  $X$ ,  $f \leq g$  on  $X \times Y$ ,  $Y$  is a compact topological space and  $f$  is l.s.c. in its second variable. Then

$$\min_{y \in Y} \sup_{x \in X} f(x, y) \leq \sup_{x \in X} \inf_{y \in Y} g(x, y).$$

We shall give negative answer for both conjectures of S. SIMONS [27].

Our counterexample shows that these conjectures fail even in a special case:  $X$  and  $Y$  are both compact,  $f, g : X \times Y \rightarrow \mathbb{R}$  are continuous,  $f$  is upward on  $Y$ ,  $g$  is downward on  $X$  and  $f \leq g$  on  $X \times Y$ . However,  $\min_{y \in Y} \sup_{x \in X} f(x, y) > \sup_{x \in X} \inf_{y \in Y} g(x, y)$ .

To this end, we need the following Lemma:

LEMMA 5.1. *Let  $X$  and  $Y$  be compact subsets of  $\mathbb{R}$  and  $f : X \times Y \rightarrow \mathbb{R}$ ,  $f(x, \cdot)$  is either strictly monotone, or a constant function. If  $Y$  is convex, then  $f$  is upward on  $Y$ .*

PROOF. It is enough to show that for each  $\varepsilon > 0$  there exists  $\delta > 0$  such that for each  $x \in X$  and  $y_1, y_2 \in Y$ ,  $f\left(x, \frac{y_1 + y_2}{2}\right) \leq \max\{f(x, y_1), f(x, y_2)\}$  and  $|f(x, y_1) - f(x, y_2)| \geq \varepsilon \Rightarrow f\left(x, \frac{y_1 + y_2}{2}\right) \leq \max\{f(x, y_1), f(x, y_2)\} - \delta$ . The first relation is trivial. For the second on let's suppose that there exists  $\varepsilon > 0$  such that for all  $n \in \mathbb{N}$  there exist  $x_n \in X$ ,  $y_1^n, y_2^n \in Y$  with  $|f(x_n, y_1^n) - f(x_n, y_2^n)| \geq \varepsilon$  and  $f\left(x_n, \frac{y_1^n + y_2^n}{2}\right) > \max\{f(x_n, y_1^n), f(x_n, y_2^n)\} - \frac{1}{n}$  for all  $n \in \mathbb{N}$ . Using compactness, one can suppose  $(x_n), (y_1^n), (y_2^n)$  convergent

(if it is necessary, we choose subsequences). Since  $f$  is uniformly continuous on  $X \times Y$ , there exists  $\delta_0 > 0$  such that  $|y_1^n - y_2^n| \geq \delta_0$  for all  $n \in \mathbb{N}$ . Let  $x_n \rightarrow x, y_1^n \rightarrow y_1, y_2^n \rightarrow y_2$ . Then we obtain the following contradiction:

$$f\left(x, \frac{y_1 + y_2}{2}\right) \geq \max\{f(x, y_1), f(x, y_2)\}.$$

A similar property assures that  $g$  is downward on  $x$ .

Now let  $X = Y = [-1, 1], f, g : X \times Y \rightarrow \mathbb{R}$  given by

$$\begin{aligned} f(-1, -1) &= g(-1, -1) = f(1, 1) = g(1, 1) = 2, \\ f(-1, 1) &= g(-1, 1) = f(1, -1) = g(1, -1) = -2, \\ f(-1, 0) &= g(-1, 0) = f(1, 0) = g(1, 0) = 1, \\ f(0, -1) &= g(0, -1) = f(0, 1) = g(0, 1) = -1; \end{aligned}$$

on the segments

$$[(-1, -1), (-1, 0)], [(-1, 0), (-1, 1)], [(-1, 1), (0, 1)], \dots, [(0, -1), (-1, -1)]$$

let  $f = g$  be affine;  $f(0, y) = -1$  for each  $y \in [-1, 1]$  and  $g(x, 0) = 1$  for each  $x \in [-1, 1]$  (see figure 1). At the interior points of the square  $[-1, 1] \times [-1, 1]$  we construct  $f$  such that for each  $y \in (-1, 1), f(x, y)$  is affine for  $x \in (-1, 0)$  and  $f(x, y)$  is affine for  $x \in (0, 1)$  (see figure 2);  $g$  such that for each  $x \in (-1, 1), g(x, y)$  is affine for  $y \in (-1, 0)$  and  $g(x, y)$  is affine for  $y \in (0, 1)$  (see figure 3).

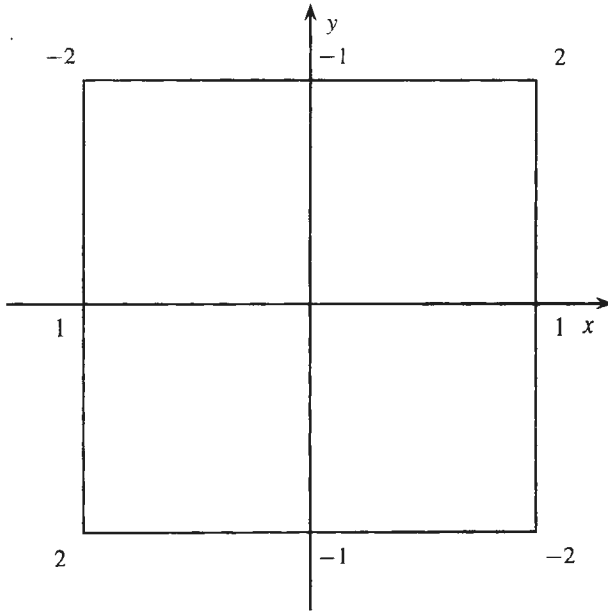


Fig. 1

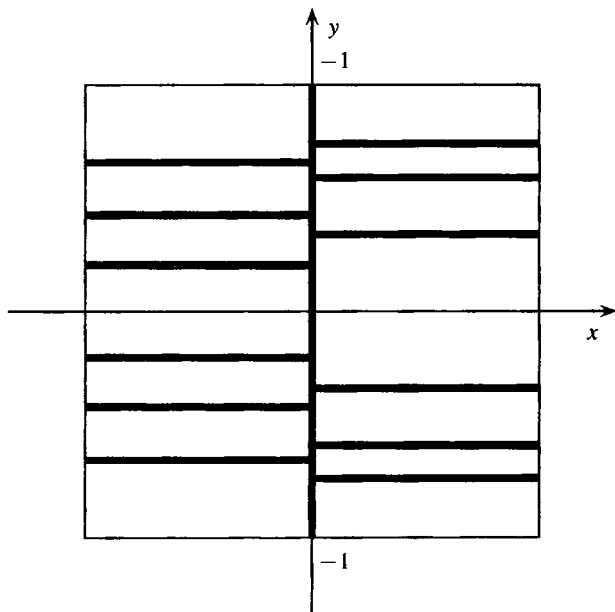


Fig. 2

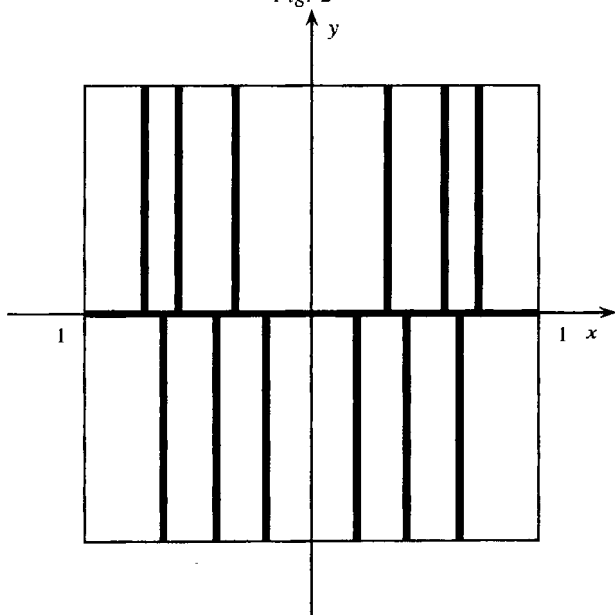


Fig. 3

Geometrically, both  $f$  and  $g$  are formed by two conoidal surfaces.

Using Lemma 5.1,  $f$  is upward on  $Y$  and  $g$  is downward on  $X$ . It is easy to verify that  $f \leq g$  on  $[-1, 1] \times [-1, 1]$ . On the other hand, we have  $\inf_y \sup_x f(x, y) = 1$  and  $\sup_x \inf_y g(x, y) = -1$ .

Therefore, the second conjecture of S. SIMONS fails (i.e. (38)). The first one (i.e. (37)) fails too: let  $X_0 = \{-1, 1\}$ .

It remains an open question: under which additional hypothesis (37) and (38) can be proved?

We answer to this question in case when one of the sets  $X$  and  $Y$  is finite. To this end we prove the following lemmas.

LEMMA 5.2. *Suppose that  $f : X \times Y \rightarrow \mathbb{R}$  is downward on  $X$  and  $X$  is a finite set. Then there exists  $x^* \in X$  such that  $f(x^*, y) \geq f(x, y)$  for each  $x \in X$  and  $y \in Y$ .*

PROOF. Let  $X = \{x_1, x_2, \dots, x_n\}$ . It is easy to see that (36) implies

$$(39) \quad \begin{aligned} & \forall x_1, x_2 \in X : \exists x_3 \in X \quad \text{such that} \\ & \left\{ \begin{array}{l} \forall y \in Y : f(x_3, y) \geq \min\{f(x_1, y), f(x_2, y)\} \quad \text{and} \\ f(x_1, y) \neq f(x_2, y) \Rightarrow f(x_3, y) > \min\{f(x_1, y), f(x_2, y)\}. \end{array} \right. \end{aligned}$$

We use induction. If  $n = 1$  the conclusion is trivial. Suppose that for  $n = k$  the property holds and prove it for  $n = k + 1$ . First we shall prove

(40) There exist

$$i, j \in \{1, 2, \dots, k + 1\} \text{ with } i \neq j \text{ such that } f(x_i, y) \geq f(x_j, y) \text{ for each } y \in Y.$$

Suppose the contrary. Then the sets  $A_1 = \{y \in Y : f(x_1, y) < f(x_2, y)\}$  and  $A_2 = \{y \in Y : f(x_1, y) > f(x_2, y)\}$  are both nonempty. Choose  $x_3$  as in (39) and let  $A_3 = \{y \in Y : f(x_3, y) < f(x_1, y)\}$ . If  $A_3 = \emptyset$ , then (40) holds. Suppose  $A_3 \neq \emptyset$ . It is easy to see by (39) that

$$\begin{aligned} f(x_3, y) &> f(x_1, y) \quad \text{for every } y \in A_1 \quad \text{and} \\ f(x_3, y) &> f(x_2, y) \quad \text{for every } y \in A_2. \end{aligned}$$

We also have  $A_3 \subset A_2$ . Let  $x_4$  be the element which corresponds to  $x_1$  and  $x_3$  by (39). Then  $f(x_4, y) > f(x_3, y)$  for every  $x \in A_3$ . Let  $A_4 := \{y \in Y : f(x_4, y) < f(x_1, y)\}$ . Then  $A_4 \subset A_3$ . If  $A_4 = \emptyset$ , we have (40). Else, we continue this procedure. It is also clear, that the elements  $x_1, x_2, x_3, \dots$  are distinct (supposing that the corresponding sets  $A_1, A_2, A_3, \dots$  are nonempty). Hence, it is impossible to continue this procedure indefinitely since  $X$  is finite. Then, for an  $i \in \{1, 2, \dots, k + 1\}$  we must have  $A_i = \emptyset$ , which implies

(40). Choose  $i, j \in \{1, 2, \dots, k+1\}$ ,  $i \neq j$  such that  $f(x_i, y) \geq f(x_j, y)$  for every  $y \in Y$ . Then (39) remains true for  $\{x_1, x_2, \dots, x_{k+1}\} \setminus \{x_j\}$ , or in other words the functions  $f(x_p, \cdot)$ ,  $p \in \{1, 2, \dots, k+1\} \setminus \{j\}$  satisfy (39). Using induction, the conclusion of Lemma 5.2 follows.

LEMMA 5.3. *Suppose  $f : X \times Y \rightarrow \mathbb{R}$  is upward on  $Y$  and  $Y$  is a finite set. Then there exists  $y^* \in Y$  such that  $f(x, y^*) \leq f(x, y)$  for each  $x \in X$  and  $y \in Y$ .*

The proof is analogous to the proof of Lemma 5.2. We omit the details.

PROPOSITION 5.1. *Suppose that  $f$  is downward on  $X$  and  $X$  is finite. Then*

$$\inf_{y \in Y} \sup_{x \in X} f(x, y) = \sup_{x \in X} \inf_{y \in Y} f(x, y).$$

PROOF. By Lemma 5.2 there exists  $x^* \in X$  such that  $f(x^*, y) \geq f(x, y)$  for every  $x \in X$  and  $y \in Y$ . Then  $\sup_{x \in X} f(x, y) = f(x^*, y)$  for every  $y \in Y$ , from which we have

$$\inf_{y \in Y} \sup_{x \in X} f(x, y) = \inf_{y \in Y} f(x^*, y) \leq \sup_{x \in X} \inf_{y \in Y} f(x, y).$$

PROPOSITION 5.2. *Suppose that  $f$  is upward on  $Y$  and  $Y$  is finite. Then*

$$\inf_{y \in Y} \sup_{x \in X} f(x, y) = \sup_{x \in X} \inf_{y \in Y} f(x, y).$$

*The proof is similar to the proof of Proposition 5.1 and uses Lemma 5.3.*

COROLLARY 5.1. *Suppose that  $f, g : X \times Y \rightarrow \mathbb{R}$  are such that  $f \leq g$  on  $X \times Y$ .*

(41) If  $f$  is upward on  $Y$  and  $Y$  is finite, then

$$\inf_{y \in Y} \sup_{x \in X} f(x, y) \leq \sup_{x \in X} \inf_{y \in Y} g(x, y);$$

(42) if  $g$  is downward on  $X$  and  $X$  is finite, then

$$\inf_{y \in Y} \sup_{x \in X} f(x, y) \leq \sup_{x \in X} \inf_{y \in Y} g(x, y).$$

### 6. Saddle points and generalized Kuhn–Tucker theorems

In this section, using König’s minimax theorem [22], we extend the Kuhn–Tucker principle (see for instance [30]) for nonconvex optimization problems with side conditions. Our results contain, in particular, those of [30]. For other extensions of Kuhn–Tucker type theorems which uses KÖNIG [22] see S. SIMONS [28].

First we discuss the case when the constrains are given in operator form.

Let  $X$  be a compact Hausdorff topological space,  $Y$  be reflexive Banach space,  $Y^*$  the dual space of  $Y$  and  $K$  a closed convex cone of  $Y$  such that its interior is nonempty. Let  $K^*$  be the dual cone of  $K$ , i.e. the set  $\{y^* \in Y^* : y^*(y) \geq 0 \text{ for all } y \in K\}$ . Consider the ordering relation  $\ll$  on  $Y$  defined by  $y_1 \ll y_2$  iff  $y_2 - y_1 \in K$ .

Let  $f : X \rightarrow \mathbb{R}$ ,  $G : X \rightarrow Y$  and  $A := \{x \in X : G(x) \ll 0\}$ . Consider the problem

$$(P) \quad \begin{cases} f(x) \rightarrow \min \\ x \in A \end{cases}$$

Then we prove

**THEOREM 6.1.** *Suppose that the following conditions hold:*

- (43)  $f$  is l.s.c. on  $X$ ,  $G$  is l.s.c. in the sense that the functionals  $x \rightarrow y^*(G(x))$  are l.s.c. on  $X$  for all  $y^* \in K^*$ ;
- (44) for each  $x_1, x_2 \in X$ , there exists  $x_3 \in X$  such that  $2f(x_3) \leq f(x_1) + f(x_2)$  and  $2G(x_3) \ll G(x_1) + G(x_2)$ ;
- (45) there exists  $\bar{x} \in X$  such that  $-G(\bar{x}) \in \text{int } K$  (Slater condition).

*Then the following two assertions are equivalent:*

- (46) (P) has a solution  $x_0$ ;  
There exists  $y_0^* \in K^*$  such that
- (47)  $f(x_0) + y_0^*(G(x_0)) = \min_{x \in X} \{f(x) + y_0^*(G(x))\}$  and  $y_0^*(G(x_0)) = 0$

**PROOF.** Define the lagrange function  $L : X \times K^* \rightarrow \mathbb{R}$  by  $L(x, y^*) = f(x) + y^*(G(x))$ . Let  $K_n^* = \{y^* \in K^* : \|y^*\| \leq n\}$ . Then we have:

For each  $x_1, x_2 \in X$ , there exists  $x_3 \in X$  such that

- (48)  $2L(x_3, y^*) \leq L(x_1, y^*) + L(x_2, y^*)$  for each  $y^* \in K_n^*$ ;
- (49) the functions  $L(x, \cdot)$  are affine and continuous on  $K_n^*$  for each  $x \in X$ ; therefore they are weakly continuous:

The functions  $L(\cdot, y^*)$  are l.s.c. on  $X$  for each  $y^* \in K^*$ . Since  $K_n^*$  is weakly compact, by König's theorem [22], there exists a saddle point  $(x_n, y_n^*)$  of  $L(x, y^*)$  on  $X \times K_n^*$ , i.e.  $L(x_n, y^*) \leq L(x_n, y_n^*) \leq L(x, y_n^*)$  for each  $x \in X$  and  $y^* \in K_n^*$ . In particular

$$(50) \quad L(x_n, 0) \leq L(x_n, y_n^*) \leq L(\bar{x}, y_n^*).$$

It is easy to see that the sequence  $(y_n^*)$  is bounded. Otherwise, by (45) one can choose a neighborhood of  $-G(\bar{x})$  which is contained in  $\text{int} K$ , that is, there exists  $r > 0$  such that

$$y^*(-G(\bar{x}) - rh) \geq 0 \quad \text{for each } y^* \in K^*$$

and  $h \in Y$  with  $\|h\| \leq 1$ . Therefore

$$(51) \quad r\|y^*\| = \sup_{\|h\|=1} y^*(rh) \leq y^*(-G(\bar{x})) \quad \text{for each } y^* \in K^*.$$

Now if  $\|y_n^*\| \rightarrow \infty$ , by (10),  $-y_n^*(G(\bar{x})) \rightarrow \infty$ . Since  $L(\cdot, 0)$  is l.s.c. on  $X$ , this contradicts (50). Without loss of generality, we may suppose  $x_n \rightarrow z \in X$ ,  $y_n^* \rightarrow y_0^* \in K^*$  (converges weakly) and  $L(x_n, y_n^*) \rightarrow \alpha$  as  $n \rightarrow \infty$ . Therefore

$$L(z, y^*) \leq \varliminf_{n \rightarrow \infty} L(x_n, y^*) \leq \alpha \leq \varliminf_{n \rightarrow \infty} L(x, y_n^*) = L(x, y_0^*)$$

for each  $x \in X$  and  $y^* \in K^*$ . In particular,  $L(x, y_0^*) = \alpha$ , hence  $(x, y_0^*)$  is a saddle point of  $L$  on  $X \times K^*$ , i.e.

$$f(z) + y^*(G(z)) \leq f(z) + y_0^*(G(z)) \leq f(x) + y_0^*(G(x))$$

for each  $x \in X$  and  $y^* \in K^*$ .

$$(52) \quad \text{In particular } z \in A \text{ and } y_0^*(G(z)) = 0.$$

Now let  $x_0$  be a solution of (P). Then  $y_0^*(G(\bar{x})) \leq 0$  (since  $x_0 \in A$  and  $f(x_0) = \inf_{x \in A} \sup_{y^* \in K^*} L(x, y^*)$ ).

Indeed,  $L(x, y^*) \leq f(x)$  for each  $x \in A$  and  $y^* \in K^*$ , hence

$$\sup_{y^* \in K^*} L(x, y^*) \leq f(x)$$

for each  $x \in A$ . Therefore

$$\inf_{x \in A} \sup_{y^* \in K^*} L(x, y^*) \leq \sup_{y^* \in K^*} L(x_0, y^*) \leq f(x_0).$$

It is easy to see that we also have  $f(x_0) \geq \inf_{x \in X} \sup_{y^* \in K^*} L(x, y^*)$  using the fact

that  $x_0$  is a solution of (P). Thus  $f(x_0) = \alpha$ , hence  $f(x_0) + y_0^*(G(x_0)) \leq f(x_0) \leq f(x) + y_0^*(G(x))$  for each  $x \in X$ . On the other hand, if we put  $x_0$  instead of

$x$ , and 0 instead of  $y^*$  in (52), we obtain  $y_0^*(G(x_0)) = 0$ . Hence (46)  $\Rightarrow$  (47). (47)  $\Rightarrow$  (46) is trivial. This completes the proof.

Now consider the case when the constraints are given in the form of inequalities. Let  $X$  be a compact Hausdorff topological space,  $f_0, f_1, \dots, f_m$  be real functions on  $X$ . Consider the problem:

$$(P') \quad \begin{cases} f_0(x) \rightarrow \min, & x \in X \\ f_1(x) \leq 0, \dots, f_m(x) \leq 0 \end{cases}$$

COROLLARY 6.1. *Suppose*

(53)  $f_i$  is l.s.c. on  $X$  for each  $i \in \{0, \dots, m\}$ ;

(54) for each  $x_1, x_2 \in X$ , there exists  $x_3 \in X$  such that

$$2f_i(x_3) \leq f_i(x_1) + f_i(x_2), \quad \text{for each } i \in \{0, \dots, m\};$$

(55) there exists  $\bar{x} \in X$  such that  $f_i(\bar{x}) < 0$  for each  $i \in \{1, 2, \dots, m\}$ .

Then the following two assertions are equivalent:

(56)  $x_0$  is a solution of  $(P')$ ;

There exist (Lagrange multipliers)  $\lambda_1^*, \lambda_2^*, \dots, \lambda_m^*$  such that

$$(57) \quad f_0(x_0) + \sum_{i=1}^m \lambda_i^* f_i(x_0) = \min_{x \in X} \{f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x)\} \quad \text{and}$$

$$\lambda_i^* f_i(x_0) = 0 \quad \text{for each } i \in \{1, \dots, m\}.$$

PROOF. Take  $Y = \mathbb{R}^m$ ,  $K = \{h = (h_1, \dots, h_m) \in \mathbb{R}^m; h_i \geq 0, i \in \{1, \dots, m\}\}$ ,  $f = f_0$ ,  $G = (f_1, \dots, f_m) : X \rightarrow \mathbb{R}^m$  and apply Theorem 6.1.

Finally we give an account on the the connection of the results considered, in Figure 4.

### References

[1] H. BRÉZIS, L. NIRENBERG and G. STAMPACCHIA, A remark on Ky Fan's minimax principle. *Boll. Un. Mat. Ital.*, **6** (1972), 293–300.  
 [2] F. E. BROUWER, The fixed point theory of multi-valued mappings in topological vector spaces, *Math. Ann.*, **177** (1968), 283–301.  
 [3] K. FAN, Fixed-point and minimax theorems in locally convex topological linear spaces, *Proc. Natl. Acad. Sci. USA*, **38** (1952), 121–126.  
 [4] K. FAN, A generalization of Tychonoff's fixed point theorem, *Math. Ann.*, **142** (1961), 305–310.

- [15] K. FAN, A minimax inequality and its applications, in "Inequalities" (O. Shisha ed.) Vol 3, 103–113. *Academic Press, New York*, 1972.
- [16] K. FAN, Some properties of convex sets related to fixed point theorems, *Math. Ann.*, **266** (1984), 519–537.
- [17] K. FAN, A survey of some results closely related to the Knaster–Kuratowski–Mazurkiewicz theorem, in *Game Theory and Applications*, 358–370, *Academic Press*, 1990.
- [18] C. HORVÁTH, Quelques theoremes en theorie des Mini-Max, *C. R. Acad. Sci. Paris*, **310** (1990), 269–272.
- [19] C. HORVÁTH, A connectivity approach to minimax inequalities, (to appear).
- [110] C. HORVÁTH, and A. SÖVEGJÁRTÓ, On convex functions, *Ann. Univ. Sci. Budapest, Sect Math.*, **29** (1986), 193–198.
- [111] I. JOÓ, A simple proof for van Neumann's minimax theorem, *Acta Sci. Math. Szeged*, **42** (1980), 91–94.
- [112] I. JOÓ, and L. L. STACHÓ, A note on Ky Fan's minimax theorem, *Acta Math. Acad. Sci. Hung.*, **39** (1982), 401–407.
- [113] I. JOÓ, Note on my paper "A simple proof for von Neumann minimax theorem", *Acta Math. Hung.*, **44** (1984), 363–365.
- [114] I. JOÓ, On some convexities, *Acta Math. Hung.*, **54** (1989), 163–172.
- [115] I. JOÓ, and G. KASSAY, Convexity, minimax theorems and their applications, Preprint, 1991.
- [116] I. JOÓ, A general minimax theorem, *Publ. Math. Debrecen*, **41** (1992) (to appear).
- [117] G. KASSAY, and I. KOLUMBÁN, On the Knaster–Kuratowski–Mazurkiewicz and Ky Fan's theorems, *Babes-Bolyai Univ. Cluj*, Preprint, **7** (1990), 87–100.
- [118] G. KASSAY, On Brézis–Nirenberg–Stampacchia's minimax principle *Babes-Bolyai Univ. Cluj*, Preprint, **7** (1991).
- [119] G. KASSAY, A simple proof for König's minimax theorem, *Acta Math. Hung.*, (1995) (to appear).
- [120] H. KOMIYA, Convexity in a topological space, *Fund. Math.*, **111** (1981), 107–113.
- [121] V. KOMORNIK, Minimax theorems for upper semicontinuous functions *Acta Math. Acad. Sci. Hung.*, **40** (1982), 159–163.
- [122] H. KÖNIG, Über das von Neumannsche Minimax-Theorem, *Arch. Math.*, **19** (1968), 482–487.
- [123] J. L. LIONS, and G. STAMPACCHIA, Variational inequalities, *Comm. Pure and Appl. Math.*, **20** (1967), 493–519.
- [124] H. NIKAIĐÓ, On von Neumann's minimax theorem, *Pacific J. Math.*, **4** (1954), 65–72.
- [125] S. SIMONS, Two-functions minimax theorems and variational inequalities for functions on compact and noncompact sets with some comments on fixed-point theorems, in *Proc. Sump. Pure Math.* (F. E. Brouwer ed.) Vol. 45, 377–392. Amer. Math. Soc., Providence, Rhode Island, 1986.

- [26] S. SIMONS, Abstract Kuhn–Tucker theorems, *J. Opt. Theory and Appl.*, **58** (1988), 148–152.
- [27] S. SIMONS, An upward-downward minimax theorem, *Arch. Math.*, **55** (1990), 275–279.
- [28] L. L. STACHÓ, Minimax theorems beyond topological vector spaces, *Acta Sci. Math. Szeged*, **42** (1980), 157–164.
- [29] F. TERKELSEN, Some minimax theorems, *Math. Scand.*, **31** (1972), 405–413.
- [30] E. ZEIDLER, *Nonlinear Functional Analysis and its Applications Part III*, Springer-Verlag, Berlin, 1985.

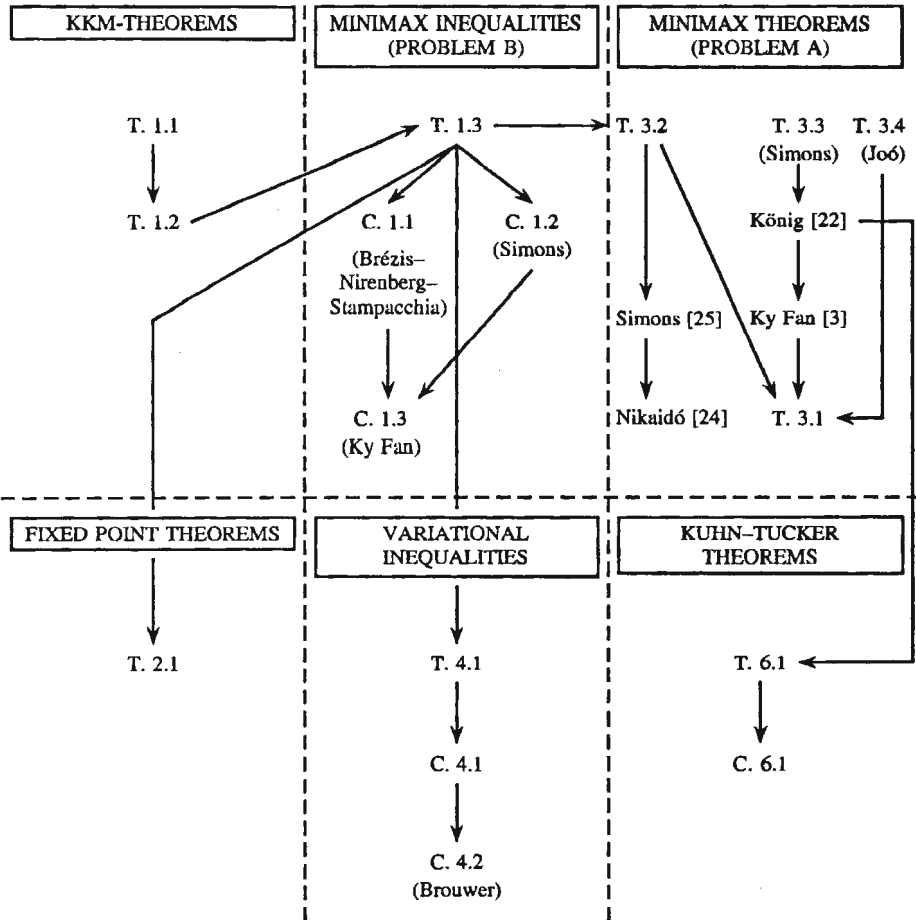


Fig. 4



## EINIGE PFLASTERUNGEN DES HYPERBOLISCHEN RAUMES MITTELS FLÄCHENTRANSITIVER BEWEGUNGSGRUPPEN\*

Von

E. MOLNÁR\*\* und J. SZIRMAI\*\*

Lehrstuhl für Geometrie, Technische Universität, Budapest

(Eingegangen am 17. Oktober 1994)

Herrn Professor Julius Strommer zu seinem 75. Geburtstag gewidmet

### 1. Einführung

In der Arbeit [1] wurden diejenigen euklidischen Pflasterungen  $(T, \Gamma)$  klassifiziert, bei denen eine Bewegungsgruppe  $\Gamma$  auf den Flächen der Pflasterung  $T$  transitiv wirkt. Der benutzte kombinatorische Algorithmus und das Computerprogramm für die Theorie der sogenannten  $D$ -Symbole (zur Ehre von B. N. DELONE, M. S. DELANEY, A. W. M. DRESS) ergeben auch solche Pflasterungen, die nur in gewissen nicht-euklidischen Räumen realisierbar sind. Die Frage nach der metrischen Realisierung einer kombinatorischen Pflasterung ist schwierig; u.a. sind beispielweise die Räume  $\mathbf{H}^2 \times \mathbb{R}^2$ ,  $\mathbf{S}^2 \times \mathbb{R}^2$  zu berücksichtigen [6] und [8].

Eine Möglichkeit für die Kennzeichnung der metrischen Realisierbarkeit der kombinatorischen Pflasterungen ist die projektiv-metrische Geometrie, wie dies in der Arbeit [2] illustriert wurde. In [9] wurde mit unseren Methoden gezeigt, daß die Verallgemeinerungen des Lambertschen Würfels im hyperbolischen Raum realisierbar sind. Im folgenden werden wir erneut Methoden der projektiven Metrik benutzen.

Wir untersuchen nochmals solche kombinatorischen Polyederpflasterungen  $(T, \Gamma)$ , bei den die Symmetriegruppe  $\Gamma$  der Pflasterung  $T$  flächentransitiv wirkt. Die Symmetriegruppe  $\Gamma$  wirkt auf der Pflasterung  $T$  flächentransitiv,

---

\* Diese Arbeit entstand in Kooperation der Geometer der TU Budapest und der Universität Potsdam

\*\* Unterstützt von der Ungarischen Wissenschaftlichen Forschungsstiftung (OTKA) No. T. 7351 (1993).

wenn es für je zwei Flächen  $f_1$  und  $f_2$  ein Element  $\gamma \in \Gamma$  gibt, das die Fläche in  $f_1$  in  $f_2 = f_1^\gamma$  so überführt, daß die ganze Pflasterung  $\mathbf{T}$  auf sich abgebildet wird. Zwei Pflasterungen  $(\mathbf{T}_1, \Gamma_1)$  und  $(\mathbf{T}_2, \Gamma_2)$  sind in derselben Klasse (äquivariant), wenn es eine bijektive inzidenttreue Abbildung  $\Phi: \mathbf{T}_1 \rightarrow \mathbf{T}_2$  gibt, für die  $\Gamma_2 = \Phi^{-1}\Gamma_1\Phi$  ist. Falls zwei Pflasterungen kombinatorisch isomorph sind ( $\mathbf{T}_1 \cong \mathbf{T}_2$ ) kann die Gruppe  $\Gamma_2$  reichhaltiger als  $\Gamma_1$  sein. In diesem Fall sagen wir, daß  $(\mathbf{T}_1, \Gamma_1)$  ein Symmetriebruch von  $(\mathbf{T}_2, \Gamma_2)$  ist. Wir interessieren uns dann für die Pflasterung  $(\mathbf{T}, \Gamma)$ , mit maximaler Bewegungsgruppe, d.h., daß die Wirkung der Gruppe  $\Gamma$  auf  $\mathbf{T}$  zur Wirkung der die Inzidenzstruktur erhaltenden Automorphismengruppe äquivariant ist:  $\Gamma \cong \text{Aut } \mathbf{T}$ .

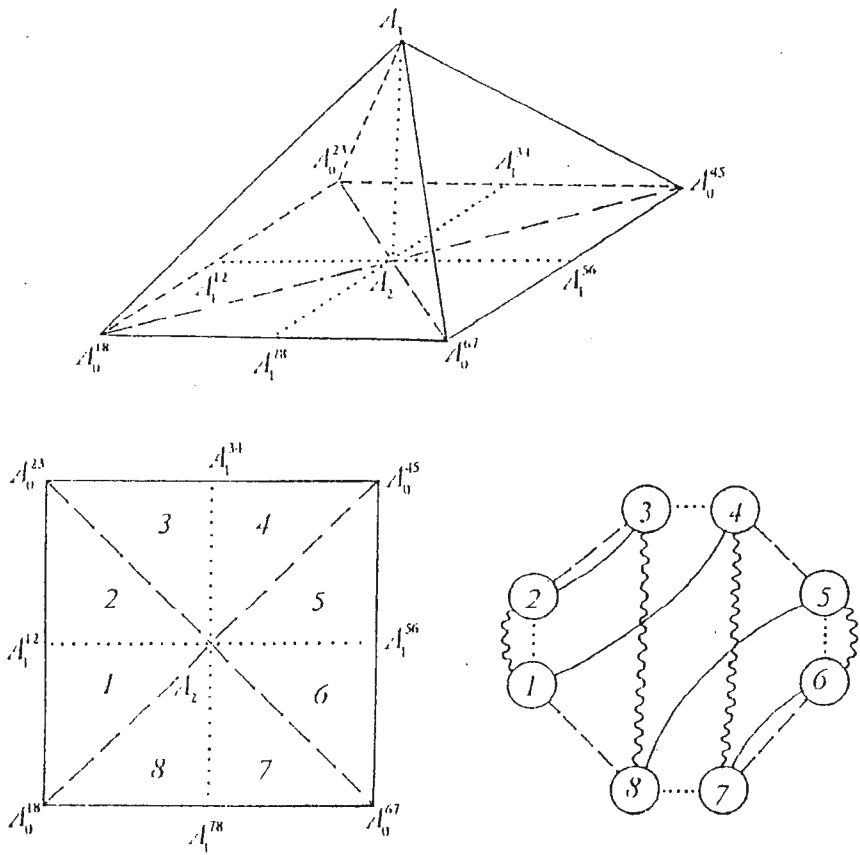


Abb. 1.

In den Sätzen 4.1, 5.1 beweisen wir, daß die Pflasterungen  $(\mathbf{T}_p, \Gamma_p)$   $p = 3, 4, 5, (6)$  — siehe den kombinatorischen Fundamentalbereich in Abb. 1 — im hyperbolischen Raum realisierbar sind. Jeder Pflasterstein  $T_p \in \mathbf{T}_p$  besitzt entsprechend 12, 24, 60,  $(\infty)$  Vierecke als Seitenflächen nach dem Stabilisator  $23 = [3, 3]^+$ ,  $432 = [4, 3]^+$ ,  $532 = [5, 3]^+$ ,  $(p6 = [6, 3]^+$ , d.h.  $\mathbf{T}_6$  ist eine Horosphärenpflasterung mit Enden als Ecken ein unendliches Polyeder).  $\mathbf{T}_3$ ,  $\mathbf{T}_4$  und  $\mathbf{T}_5$  können nur mit nicht-ebenen Seitenflächen realisiert werden.

## 2. Über die $D$ -Symbole

In der Arbeit [1] wird die Theorie der  $D$ -Symbole skizziert. In den Arbeiten [4], [5], [6], [7] finden wir einige zwei- und dreidimensionale Anwendungen der  $D$ -Symbole.

Es sei eine kombinatorische dreidimensionale Polyederpflasterung  $\mathbf{T}$  mit ihrer formalen baryzentrischen Unterteilung gegeben. Die baryzentrische Unterteilung von  $\mathbf{T}$  führt zu einer Simplexzerlegung  $\mathbf{C}$ . Die formalen Mittelpunkte eines Körpers, seine inzidenten Flächen, Kanten und Ecken bilden die Ecken eines Simplexes von  $\mathbf{C}$ . Diese Ecken markieren wir mit 3, 2, 1 und 0 ebenso wie die entsprechenden gegenüberliegenden Simplexseiten. Die Ecken eines Simplexes von  $\mathbf{C}$  werden mit  $A_i$ ,  $i \in \mathbf{I} := \{0; 1; 2; 3\}$  bezeichnet.

Es sei  $\sigma_i$  ( $i \in \mathbf{I}$ ) eine involutorische Operation (lokale topologische Spiegelung), die die bei der Seitenfläche  $i$  benachbarten Simplexe  $C_1$  und  $C_2$  von  $\mathbf{C}$  vertauscht (in Bezeichnung  $C_2 = \sigma_i C_1$ ). So werden die auf den Simplexen von  $\mathbf{C}$  von links frei wirkenden involutorischen  $\sigma_i$ -Operationen und die durch sie erzeugte zunächst freie Coxeter-Gruppe mit der Darstellung

$$\Sigma_{\mathbf{I}} = \{\sigma_i : \sigma_i^2 = 1, i \in \mathbf{I}\}$$

definiert. Wir führen ferner eine Matrixfunktion

$$\mathbf{M} : \mathbf{C} \rightarrow \mathbb{N}_{\mathbf{I} \times \mathbf{I}}, \quad C \mapsto m_{ij}(C), \quad i, j \in \mathbf{I}$$

ein, um die Wirkung der Gruppe  $\Sigma_{\mathbf{I}}$  und so die kombinatorische Struktur der Pflasterung  $\mathbf{T}$  zu beschränken. Es sei  $m_{ij}(C)$  die minimale natürliche Zahl, die die folgende Gleichung erfüllt

$$(\sigma_i \sigma_j)^{m_{ij}(C)}(C) = C.$$

Das bedeutet, daß genau  $2m_{ij}(C)$  benachbarte Simplexe sich beim Schnitt der  $i$ - und  $j$ -Seitenflächen des Simplexes  $C$  von  $\mathbf{C}$  treffen.  $\mathbf{M}$  ist die verallgemei-

nernte Coxeter-Schläffische Matrixfunktion mit den folgenden zusätzlichen Eigenschaften für alle  $C \in \mathbf{C}$  und  $i, j \in \mathbf{I}$

$$m_{ii}(C) = 1, \quad m_{ij}(C) = m_{ji}(C) = m_{ij}(\sigma_i(C)),$$

$$m_{ij}(C) = 2 \quad \text{für } |i - j| \geq 2,$$

$$m_{ij}(C) \geq 3 \quad \text{für } |i - j| = 1 \quad \text{für jedes } C \in \mathbf{C} \text{ und } i, j \in \mathbf{I}.$$

Wir nehmen an, daß  $\Gamma$  eine nicht triviale Gruppe ist, die auf  $\mathbf{T}$  wirkt. Ferner erhält jedes Element von  $\Gamma$  die baryzentrische Unterteilung von  $\mathbf{T}$  und so die  $\sigma_i$ -Operationen. Die Zerlegung  $\mathbf{C}$  und auch die Matrixfunktion  $\mathbf{M}$  werden durch die Gruppe  $\Gamma$  faktorisiert. Ein Element  $D$  von  $\mathbf{D} := \mathbf{C}/\Gamma$  ist eine Bahn (Orbit):

$$D := C^\Gamma := \{C^\gamma \in \mathbf{C} : \gamma \in \Gamma\}$$

$\mathbf{D}$  ist mit den induzierten  $\sigma_i$ -Operationen versehen. Die von  $\mathbf{C}$  induzierte Matrixfunktion ist

$$\mathbf{M} : \mathbf{D} \rightarrow \mathbb{N}_{\mathbf{I} \times \mathbf{I}}, \quad D \rightarrow (m_{ij}(C)), \quad C \in D, \quad i, j \in \mathbf{I}.$$

Nun können wir ein sogenanntes  $D$ -Symbol  $\mathbf{D}(\Sigma_{\mathbf{I}}, \mathbf{M})$  definieren, das zu einer Pflasterung  $(\mathbf{T}, \Gamma)$  gehört. Das  $D$ -Symbol besteht aus dem  $D$ -Graph und aus der entsprechenden Matrixfunktion  $\mathbf{M} = (m_{ij})$ . Der  $D$ -Graph wird durch die  $D$ -Menge  $\mathbf{D}$  und die Operationen  $\sigma_i \in \Sigma_{\mathbf{I}}$  ( $i \in \mathbf{I}$ ) definiert.

Die Bezeichnungen der Kantenlinien des  $D$ -Graphs und auch der entsprechenden  $\sigma_i$ -Operationen sind  $\sigma_0 : \dots\dots\dots$ ;  $\sigma_1 : \text{---}$ ; usw.

Von einem  $D$ -Symbol ausgehend, sehen wir, daß es die entsprechende Zerlegung (wenn sie überhaupt existiert) bis auf äquivalente Bijektion kennzeichnet (vgl. Abschnitt 3).

Von den  $D$ -Symbolen fordert man :

2.1. Der  $D$ -Graph durch  $\mathbf{D}^2$ , der von  $\mathbf{D}$  durch Streichen der  $\sigma_2$ -Operation entsteht, hat genau eine Komponente (Flächentransitivität).

2.2. Zu einem  $D$ -Symbol  $\mathbf{D}$  existiert keine surjektive Abbildung  $\Psi : \mathbf{D} \rightarrow \overline{\mathbf{D}}$  auf ein anderes  $D$ -Symbol mit  $(\sigma_i, D)^\Psi = \overline{\sigma}_i(D^\Psi)$ ,  $m_{ij}(D^\Psi) = m_{ij}(D)$  für jedes  $D \in \mathbf{D}$  alle  $i, j \in \mathbf{I}$  (Maximalität der Gruppe,  $\Gamma \cong \text{Aut } \mathbf{T}$ ).

2.3. Das Teilsymbol  $\mathbf{D}^0(\mathbf{M}^0)$ , das durch Streichen der  $\sigma_0$ -Operation und der 0-ten Zeilen und Spalten der Matrizen entsteht, hat nur solche Komponenten  $\mathbf{D}_c^0$ , für die die sogenannten Krümmungskonstanten

$$K_{\mathbf{D}_c^0} = \sum_{D \in \mathbf{D}_c^0} \left( \frac{1}{m_{12}(D)} + \frac{1}{m_{23}(D)} - \frac{1}{2} \right) > 0$$

positiv sind. So erhalten wir ausschließlich eigentliche Ecken mit endlichen Stabilisatoren  $\Gamma^0$  für die Pflasterung  $(\mathbf{T}, \Gamma)$ .

2.4. Eine analoge Forderung gilt für das Teilsymbol  $\mathbf{D}^3(\mathbf{M}^3)$  und  $\Gamma^3$ .

### 3. Die Polyederpflasterungen

$$(\mathbf{T}_p, \Gamma_p) \quad (p = 3, 4, 5)$$

3.1. Wie wir bereits erwähnt haben, ist die vollständige Aufzählung der kombinatorischen flächentransitiven Polyederpflasterungen  $(\mathbf{T}, \Gamma)$  bei der Arbeit [1] enthalten, wobei die Polyeder von  $\mathbf{T}$  eigentliche Ecken besitzen. In dieser Arbeit interessieren wir uns für diejenigen Polyederpflasterungen  $(\mathbf{T}, \Gamma)$ , die mit den folgenden  $\sigma_i$ -Operationen gelistet wurde

$$\begin{aligned} \sigma_0 &: (1,2)(3,4)(5,6)(7,8) \quad \dots\dots\dots \\ \sigma_1 &: (1,8)(2,3)(4,5)(6,7) \quad \text{-----} \\ \sigma_2 &: (1,4)(2,3)(5,8)(6,7) \quad \text{-----} \\ \sigma_3 &: (1,2)(3,8)(4,7)(5,6) \quad \text{~~~~~} \end{aligned} \quad \text{mit der Abkürzung } z \iff D_z$$

Auch die Matrixfunktion für jede Polyederpflasterung bekommen wir aus der Computerliste

$$\begin{aligned} m_{01}(D) &= 4 \quad \text{für jedes } D \in \mathbf{D} \\ m_{23}(D) &= 4 \quad \text{für jedes } D \in \mathbf{D} \\ m_{12}(D_1) &= m_{12}(D_4) = m_{12}(D_5) = m_{12}(D_8) = 4 \\ m_{12}(D_2) &= m_{12}(D_3) = 3 \\ m_{12}(D_6) &= m_{12}(D_7) = p, \quad p = 3, 4, 5. \end{aligned}$$

BEMERKUNG. Nach 2.1 sehen wir, daß die diesen  $D$ -Symbolen entsprechenden Pflasterungen  $(\mathbf{T}, \Gamma)$  flächentransitiv sind. Nach 2.3 und 2.4 besitzt die Pflasterung  $(\mathbf{T}, \Gamma)$  endliche Stabilisatoren für die Körper und für die Ecken.

In Abb. 1 sehen wir die Schemata des aus 8 Simplexen zusammengeklebten Fundamentalbereiches für die Gruppe  $\Gamma$ .

$$(3.1) \quad \mathbf{F}_\Gamma := A_0^{18} A_2 A_3 A_1^{12} \cup A_1^{12} A_2 A_3 A_0^{23} \cup A_0^{23} A_2 A_3 A_1^{34} \cup A_1^{34} A_2 A_3 A_0^{45} \cup A_0^{45} A_2 A_3 A_1^{56} \cup A_1^{56} A_2 A_3 A_0^{67} \cup A_0^{67} A_2 A_3 A_1^{78} \cup A_1^{78} A_2 A_3 A_0^{18}.$$

Die erzeugenden Bewegungen für die Gruppe  $\Gamma$  sind

$$r: \text{ ist eine Spiegelung an Achse } r = A_1^{12} A_1^{56},$$

$$r_1: \text{ ist eine Drittel-Drehung um die Achse } r_1 = A_0^{23} A_3, A_0^{18} \mapsto A_0^{45},$$

$$r_2: \text{ ist eine } p\text{-Drehung um die Achse } r_2 = A_0^{67} A_3, A_0^{45} \mapsto A_0^{18}, (p = 3, 4, 5).$$

Unter diesen Erzeugenden bestehen die folgenden definierenden Relationen in der Gruppe  $\Gamma$

$$(3.2) \quad r^2 = r_1^3 = r_2^p = (r_1 r_2)^2 = (r_1^{-1} r r_1) r (r_2^{-1} r r_2) r = 1$$

nach dem Poincaréschen Algorithmus [7].

Weiterhin bezeichnen wir die obigen kombinatorischen Pflasterungen mit  $(T_p, \Gamma_p)$  ( $p = 3, 4, 5$ ). Das Hauptproblem ist, in welchem Raum werden diese kombinatorischen Pflasterungen  $T_p$  mit einer Bewegungsgruppe  $\Gamma_p$  metrisch existieren.

**3.2.** Es sei  $V^4$  ein 4-Vektorraum über dem reellen Körper  $\mathbb{R}$ . Der duale Raum von  $V^4$  wird mit  $V_4$  bezeichnet. Wir betrachten zunächst das Tetraeder  $A_0A_1A_2A_3$  in Abb. 2. Die Formen  $b^0, b^1, b^2, b^3$  vom dualen Raum  $V_4$ , mit kursiven fetten Buchstaben, stellen die Flächenebenen dieses Tetraeders dar (Abb. 2), dessen Coxeter-Schläfli Matrix die Folgende ist:

$$(3.3) \quad (b^{ij}) = \begin{bmatrix} 1 & -\cos\beta^{01} & -\cos\beta^{02} & -\cos\beta^{03} \\ -\cos\beta^{01} & 1 & -\frac{1}{2} & 0 \\ -\cos\beta^{02} & -\frac{1}{2} & 1 & -\cos\frac{\pi}{p} \\ -\cos\beta^{03} & 0 & -\cos\frac{\pi}{p} & 1 \end{bmatrix} \quad (p = 3, 4, 5)$$

Die Matrixelemente  $b^{ij} := \begin{cases} 1, & i = j \\ -\cos\beta^{ij}, & i \neq j \end{cases}$  kennzeichnen die Keilwinkel  $\beta^{ij}$  zwischen den Ebenen  $(b^i)(b^j)$ .

Wir können mittels  $(b^{ij})$  ein Skalarprodukt

$$\langle \ ; \ \rangle : V_4 \times V_4 \rightarrow \mathbb{R}, \quad \langle u, v \rangle := \langle b^i u_i, b^j u_j \rangle = u_i b^{ij} u_j$$

und somit eine Metrik definieren, wobei wir die Einstein-Schouten Konventionen benutzen. Die Vektoren  $\mathbf{a}_i \in V^4$ ,  $i \in I = \{0, 1, 2, 3\}$ , mit stehenden fetten Buchstaben, bilden die duale Basis zu  $\{b^0, b^1, b^2, b^3\}$  mit der Gleichung  $\mathbf{a}_i b^j = \delta_i^j$  (das Kronecker Symbol). Sie kennzeichnen die Eckpunkte unseres Tetraeders  $A_0A_1A_2A_3$ . In dem Fall  $\text{Det}(b^{ij}) \neq 0$  können wir die inverse Matrix  $(a_{ij}) = (b^{ij})^{-1}$  mit  $a_{ik} b^{kj} = \delta_i^j$  bilden und ein Skalarprodukt

$$\langle \ ; \ \rangle : V^4 \times V^4 \rightarrow \mathbb{R}, \quad \langle \mathbf{x}, \mathbf{y} \rangle := \langle x^i \mathbf{a}_i, y^j \mathbf{a}_j \rangle = x^i a_{ij} y^j \quad \text{in } V^4 \text{ definieren.}$$

Die lineare Polarität und ihre Inverse

$$(\ * ) : V_4 \rightarrow V^4 : u \mapsto u_* = \mathbf{u}, \quad ( \ * ) : V^4 \rightarrow V_4 : \mathbf{x} \mapsto \mathbf{x}^* = x,$$

sind durch

$$\mathbf{u} := u_* = (b^i u_i)_* = u_i b_*^i = u_i b^{ij} \mathbf{a}_j$$

$$\mathbf{x} := \mathbf{x}^* = (x^i \mathbf{a}_i)^* = \mathbf{a}_i^* x^i = b^k a_{ki} x^i$$

definiert, wobei  $\langle \mathbf{x}, \mathbf{u} \rangle = \mathbf{xu} = \langle x, u \rangle$  gilt.

Über der Teilraumstruktur von  $V^4$  bzw.  $V_4$  gewinnen wir den projektiv-metrischen Raum  $\mathcal{P}^3(V^4, V_4(\cdot, \cdot))$ . Wenn  $x \in V^4$  und  $u \in V_4$  die Gleichungen  $0 = xu = \langle x, u \rangle$  erfüllen, sagen wir, daß  $(u)$  und  $(x)$  orthogonale Ebenen sind, oder  $(x)$  und  $(u)$  konjugierte (polare) Punkte in  $\mathcal{P}^3$  sind, oder die Ebene  $(u)$  den Punkt  $(x)$  enthält.

3.3. Weiterhin setzen wir voraus, daß wir eine hyperbolische Metrik mit der Signatur  $(+, +, +, -)$  haben. Eine Gerade wird durch die Punkte  $E(e_0), E_1(e_1)$  festgelegt. Die Geradenspiegelung an  $r$  oder die Halbdrehung um die Achse  $r = E_0E_1$  ist durch [2]

$$(3.4) \quad r : y \mapsto y^r := y - 2 \langle y, e_\alpha \rangle e^{\alpha\beta} e_\beta \quad \alpha, \beta \in \{0, 1\}$$

mit  $e_{\alpha\beta} := \langle e_\alpha, e_\beta \rangle, e^{\alpha\beta} := (e_{\alpha\beta})^{-1}$  und

$$e_0 = a_2 + \sqrt{\frac{a_{22}}{a_{33}}} a_3, \quad e_1 = a_4 + \sqrt{\frac{a_{44}}{a_{11}}} a_1$$

definiert.

Die Spiegelung  $s$  an der Ebene  $(b)$  und an seinem nicht inzidenten Pol  $(b)$  wird durch die Formel

$$(3.5) \quad x \mapsto x^s \approx x - 2 \frac{\langle x, b \rangle}{\langle b, b \rangle} b$$

für beliebigen Punkt  $(x)$  und sein Bild  $(x^s)$  definiert.

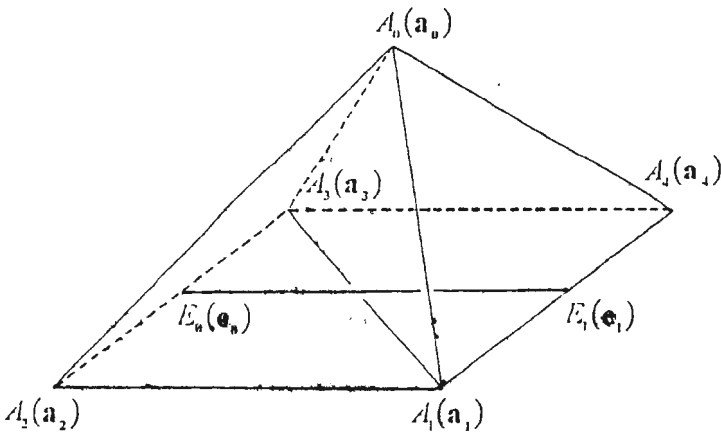


Abb. 2.

Der Abb. 2 und der Matrix  $(b^{ij})$  entnehmen wir zum Beispiel, daß die Spiegelung an der Ebene  $(b^2)$  und ihrem Pol  $(b^2)$  den Punkt  $A_2(a_2)$  in den Punkt  $A_4(a_4)$  führt :

$$a_4 \approx a_2 - 2 \frac{\langle a_2, b^2 \rangle}{\langle b^2, b^2 \rangle} b^2 = a_2 - 2b^{2i} a_i = 2 \cos \beta^{02} a_0 + a_1 - a_2 + 2 \cos \frac{\pi}{p} a_3.$$

#### 4. Die metrische Existenz der Polyederpflasterung $(T_3, \Gamma_3)$

SATZ 4.1. Die Polyederpflasterung  $(T_3, \Gamma_3)$  ist im hyperbolischen Raum  $H^3$  so metrisch realisierbar, daß die Bewegungsgruppe  $\Gamma_3$  auf den Flächen der Pflasterung  $T_3$  transitiv wirkt und jedes Polyeder eigentliche Ecken besitzt.

BEMERKUNG. In diesem Fall ist die Bewegungsgruppe  $\Gamma_3$  nicht maximal, d.h.  $\text{Aut } T_3 = \Gamma > \Gamma_3$  und  $(T_3, \Gamma_3)$  ist ein Symmetriebruch von  $(T_3, \Gamma)$ .

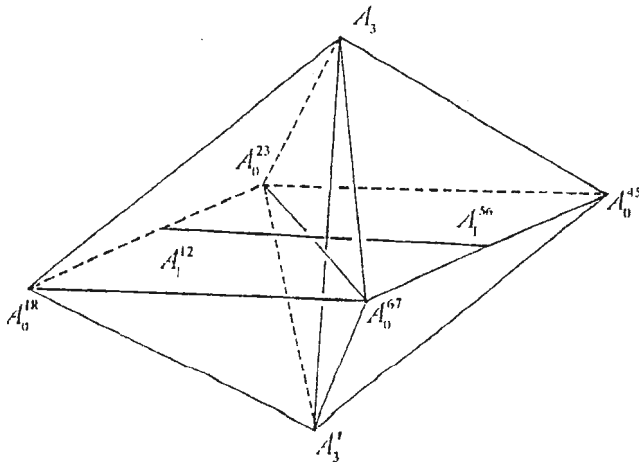


Abb. 3.

BEWEIS. Wir betrachten dasjenige Polyeder, das wir aus dem Fundamentalbereich  $F_{\Gamma_3}$  nach (3.1) der Gruppe  $\Gamma_3$  bekommen, indem wir sein  $r$ -Bild mit ihm verkleben (es formal um die Achse  $A_1^{12} A_1^{56}$  um  $180^\circ$  drehen, Abb. 3). Die Abb. 4 zeigt einen Teil von Abb. 3, das Spiegelungstetraeder

$A_0^{18} A_0^{23} A_3 A'_3$  (eines von den 9 Lannerschen Tetraedern [3]), d.h. den Fundamentalbereich einer hyperbolischen Coxeterschen Raumgruppe. Das Spiegelungstetraeder, als Hilfstetraeder  $H_0 H_1 H_2 H_3$  in Abb. 4, haben wir entsprechend in vorigem Polyeder in Abb. 3 dargestellt. Die metrische Existenz dieses Polyeders und auch der Polyederpflasterung  $(T_3, \Gamma_3)$  folgt aus der Existenz des obigen Hilfstetraeders  $A_0^{18} A_0^{23} A_3 A'_3 = H_2 H_0 H_3 H_1$ .

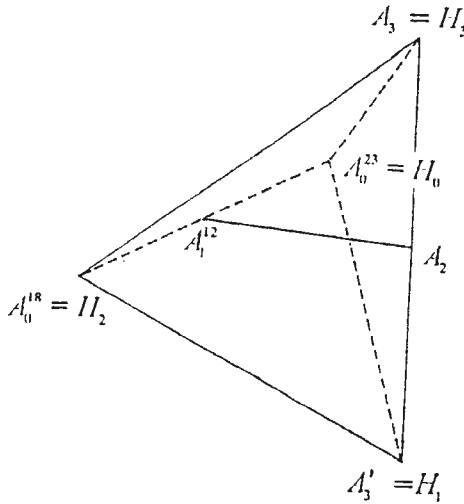
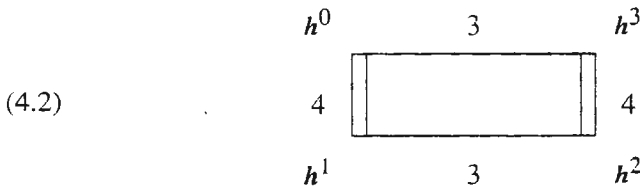


Abb. 4.

Die Coxeter-Schläfli Matrix  $(h^{ij})$  und das Coxeter-Diagramm des Hilfstetraeders  $H_0 H_1 H_2 H_3$  sind

$$(4.1) \quad (h^{ij}) = \begin{bmatrix} 1 & -\frac{1}{\sqrt{2}} & 0 & -\frac{1}{2} \\ -\frac{1}{\sqrt{2}} & 1 & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 1 & -\frac{1}{\sqrt{2}} \\ -\frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} & 1 \end{bmatrix}$$



Hier stellen die Formen  $h^0, h^1, h^2, h^3$  aus dem dualen Raum  $V_4$  die gegenüberliegenden Flächenebenen des Hilfstetraeders  $H_0H_1H_2H_3$  dar, und die Matrixelemente  $h^{ij} := -\cos(\chi^{ij})$  kennzeichnen seine Keilwinkel. Mit Hilfe der Inversen der Matrix  $(h^{ij})$  und den Formeln (3.4), (3.5) können wir die Form der Ebene  $A_0^{18}A_0^{67}A_0^{23}$  bestimmen.  $A_0^{67}$  ist das Spiegelbild von  $A_0^{23} = H_0$  an der Ebene  $H_1H_2H_3 = A_3'A_0^{18}A_3$ . (Hier sehen wir, daß die Punkte  $A_0^{18}, A_0^{67}, A_0^{23}, A_0^{45}$  in keiner Ebene liegen.) Die Ebenen  $A_0^{18}A_0^{23}A_3, A_0^{23}A_0^{67}A_3, A_0^{18}A_0^{67}A_3$  wurden im Punkt 3.2. durch die Formen  $b^1, b^2, b^3$  gekennzeichnet (vgl. Abb. 2).

Wir bestimmen die ursprünglichen Parameter  $\beta^{01}, \beta^{02}, \beta^{03}$ . Der Winkel  $\beta^{0j}$  zwischen den eigentlichen Ebenen  $(b^0), (b^j)$  ( $j = 1, 2, 3$ ).

$$(4.3) \quad \cos\beta^{0j} = \frac{-\langle b^0, b^j \rangle}{\sqrt{\langle b^0, b^0 \rangle \langle b^j, b^j \rangle}}.$$

In dem Fall  $(T_3, \Gamma_3)$  haben wir (ohne Details) die folgenden Resultate bekommen

$$\beta_3^{01} = \beta_3^{03} = 0.95531662, \quad \beta_3^{02} = 1.27795356.$$

Die Winkel sind in Radian gerechnet und werden bei den folgenden Rechnungen als wichtige Anfangsinformation benutzt.

## 5. Die metrische Existenz der Polyederpflasterungen

$$(T_p, \Gamma_p) \quad (p = 4, 5)$$

**SATZ 5.1.** *Die Polyederpflasterungen  $(T_p, \Gamma_p)$  ( $p = 4, 5$ ) sind im hyperbolischen Raum  $H^3$  metrisch realisierbar, wobei jede Bewegungsgruppe  $\Gamma_p$  ( $p = 4, 5$ ) maximal ist ( $\Gamma_p \cong \text{Aut } T_p$ ) und auf den Flächen der Pflasterung  $T_p$  transitiv wirkt und jedes Polyeder eigentliche Ecken besitzt.*

**BEWEIS.** Wir gehen von der Coxeter-Schläfli Matrix des Tetraeders  $A_0^{18}A_0^{67}A_0^{23}A_3$  aus, wie wir sie im Punkt 3.2. beschrieben haben (Abb. 5). Die Flächenebenen dieses Tetraeders werden durch die Formen  $b^i \in V_4, i \in I$  gekennzeichnet. Die Keilwinkel zwischen den Seitenflächen  $(b^i), (b^j)$  bezeichnen wir mit  $\beta^{ij}$ . Wir werden die Parameter  $\beta^{01}, \beta^{02}, \beta^{03}$  für  $p = 4, 5$  so festlegen, daß wir schließlich zu einer hyperbolischen Metrik gelangen. Die Parameter werden durch ein Gleichungssystem festgelegt, das wir mit Hilfe der erzeugenden Bewegungen und der definierenden Relationen für die Gruppe  $\Gamma_3$  ( $p = 4, 5$ ) nach 3.1. gewinnen.

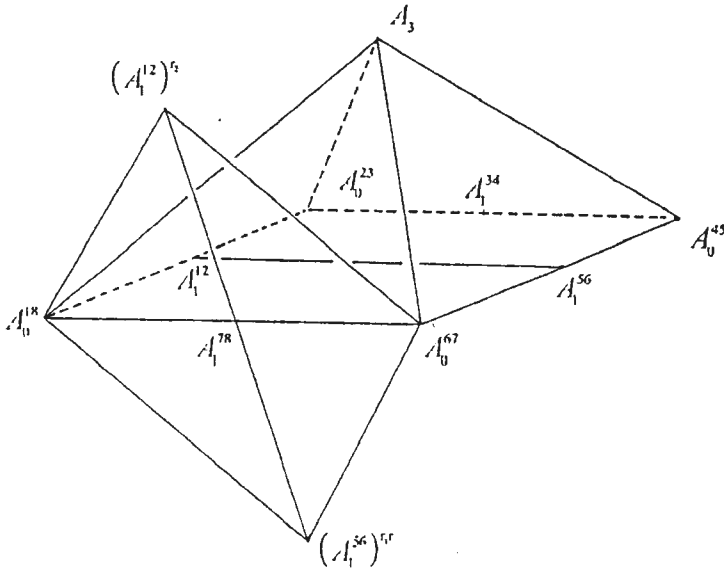


Abb. 5.

5.1. Die Gerade  $r$  wird durch die Punkte  $A_1^{12}(\mathbf{e}_0)$ ,  $A_1^{56}(\mathbf{e}_1)$  festgelegt. Die Formel der Spiegelung an der Achse  $r$  haben wir im Abschnitt 3 mit (3.4) gekennzeichnet. Die Eckpunkte des Tetraeders  $A_0^{18} A_0^{67} A_0^{23} A_0$  werden durch die Vektoren  $A_0^{18}(\mathbf{a}_2)$ ,  $A_0^{67}(\mathbf{a}_1)$ ,  $A_0^{23}(\mathbf{a}_3)$   $A_3(\mathbf{a}_0)$  festgelegt, wie sie im Punkt 3.2. und Abb. 2 beschrieben wurden.

Die erste Gleichung ergibt sich aus der Forderung  $A_0^{18} \xrightarrow{r} A_0^{23}$ :

$$\mathbf{a}_2^r = \mathbf{a}_2 - 2 \langle \mathbf{a}_2, \mathbf{e}_\alpha \rangle e^{\alpha\beta} \mathbf{e}_\beta = \sqrt{\frac{a_{22}}{a_{33}}} \mathbf{a}_3 \iff$$

$$0 = 2 \cos \beta^{02} a_{02} + \left( \sqrt{\frac{a_{22}}{a_{11}}} + 1 \right) a_{12} - \left( \sqrt{\frac{a_{13}^2}{a_{11} a_{33}}} + 1 \right) a_{22} + \left( 2 \cos \frac{\pi}{p} - \sqrt{\frac{a_{22}}{a_{33}}} \right) a_{23}.$$

5.2. Die zweite Gleichung gewinnen wir aus der Forderung  $a_0^{67} \xrightarrow{r} A_0^{45}$ :

$$0 = 2 \cos \beta^{02} a_{02} + \left( -\sqrt{\frac{a_{22}}{a_{11}}} + 1 \right) a_{12} - \left( \sqrt{\frac{a_{13}^2}{a_{11} a_{33}}} + 1 \right) a_{22} + \left( 2 \cos \frac{\pi}{p} + \sqrt{\frac{a_{22}}{a_{33}}} \right) a_{23}.$$

**5.3.** Die dritte Gleichung bekommen wir von der definierenden Relation  $(r_1^{-1} r r_1) r (r_2^{-1} r r_2) r = 1$ . Diese Relation impliziert, daß die Punkte  $(A_1^{12})r_2$ ,  $A_1^{78}$  und  $(A_1^{56})r_1 r$  (Abb. 5) in einer Gerade liegen. Diese Bedingung ist äquivalent mit der folgenden: die Punkte  $A_0^{18}$ ,  $A_0^{67}$ ,  $(A_1^{12})r_2$ ,  $(A_1^{56})r_1 r$  liegen in einer Ebene. Die Kennzeichnung der erzeugenden Bewegungen werden durch die Formel (3.4.) und (3.5) gewonnen

$$\begin{aligned} & -2A \left( 1 + \sqrt{\frac{a_{22}}{a_{11}}} \right) \left( \cos \beta^{01} + \cos \beta^{02} \right) - 4B \cos \beta^{02} + \\ & + \left( \frac{2 \cos \beta^{02}}{\sqrt{\frac{a_{22}}{a_{33}} + 2 \cos \frac{\pi}{p}} + 2 \cos \beta^{03}} \right) \cdot \\ & \cdot \left( -2A \cos \frac{\pi}{p} \left( 1 + \sqrt{\frac{a_{22}}{a_{11}}} \right) + 2C \sqrt{\frac{a_{22}}{a_{33}}} - 4B \cos \frac{\pi}{p} \right) = 0. \end{aligned}$$

wobei  $A = e_{00} e_{11} - e_{01}^2$ ,  $B = \langle e_1^{r_1}, e_0 \rangle e_{01} - \langle e_1^{r_1}, e_1 \rangle e_{00}$ ,  $C = \langle e_1^{r_1}, e_0 \rangle e_{11} - \langle e_1^{r_1}, e_1 \rangle e_{01}$  mittels der Matrix  $(a_{ij}) = (b^{ij})^{-1}$  nach (3.3) ausdrückbar sind.

Nun haben wir für die drei Unbekannten  $\beta^{01}$ ,  $\beta^{02}$ ,  $\beta^{03}$  drei komplizierte Gleichungen gefunden. Wir haben zur Lösung des Gleichungssystems zwei Computerprogramme (Eureka, Maple) benutzt. Die exakten Lösungen des Falls  $(T_3, \Gamma_3)$  haben wir bei der Computerlösung des Gleichungssystems als wichtige Anfangsinformation verwendet.

Wir haben die folgenden Winkel in Radian bekommen:

$$(T_4, \Gamma_4) \quad \beta_4^{01} \approx 0.711006493, \quad \beta_4^{02} \approx 1.43646972, \quad \beta_4^{03} \approx 1.00485229,$$

$$(T_5, \Gamma_5) \quad \beta_5^{01} \approx 0.58962576, \quad \beta_5^{02} \approx 1.52197366, \quad \beta_5^{03} \approx 1.03132995.$$

**BEMERKUNG.** Im Fall  $p = 6$  gewinnen wir auch eine hyperbolische Polyederpflasterung für das Polyeder  $A_0 A_1 A_2 A_3$  in Abb. 2, wobei die Eckpunkte des Polyeders der Polyederpflasterungen auf dem Absolutgebilde liegen. Wie es bekannt ist, sollen die gegenüberliegenden Keilwinkel von  $A_0 A_1 A_2 A_3$  dann gleich sein. Die Parameter in diesem Fall sind die folgenden:

$$(T_6, \Gamma_6) \quad \beta_6^{01} = \frac{\pi}{6} \approx 0.5235987, \quad \beta_6^{02} = \frac{\pi}{2} \approx 1.5707936, \quad \beta_6^{03} = \frac{\pi}{3} \approx 1.0471976.$$

Wir bekommen, daß die Parameter  $\beta_p^{01}$ ,  $\beta_p^{02}$ ,  $\beta_p^{03}$  ( $p = 3, 4, 5, 6$ ) monotone Folgen bilden.

$$\beta_3^{01} > \beta_4^{01} > \beta_5^{01} > \beta_6^{01}, \quad \beta_3^{02} < \beta_4^{02} < \beta_5^{02} < \beta_6^{02}, \quad \beta_3^{03} < \beta_4^{03} < \beta_5^{03} < \beta_6^{03}.$$

BEMERKUNG. Der Bahnenraum  $\mathbf{H}^3/\Gamma_p \cong \mathbf{F}_\Gamma$  in Abb. 1 kann durch das verallgemeinerte Knotendiagramm in Abb. 6 dargestellt werden. Die nummerierten Linien bezeichnen die Drehachsenklassen mit ihren Ordnungen. Die Knotenpunkte weisen auf die entsprechenden Eckenklassen von  $\mathbf{F}_\Gamma$  hin.

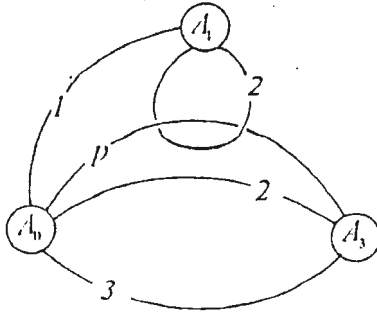


Abb. 6.

Endlich gewinnen wir, daß die Polyederpflasterungen  $(\mathbf{T}_p, \Gamma_p)$  ( $p = 3, 4, 5$ ) im Bolyai–Lobatschewskischen hyperbolischen Raum  $\mathbf{H}^3$  wirklich realisierbar sind. Die Sätze sind bewiesen.

Unsere kombinatorischen Überlegungen und Rechnungen ergeben auch die weiteren Behauptungen am Ende der Einführung.

### Literatur

- [1] A. W. M. DRESS–D. H. HUSON–E. MOLNÁR, The classification of facetransitive periodic three-dimensional tilings, Preprint, Universität Bielefeld, Combinatorics and its Applications (1990–91), ZIF Nr. 91/38. *Acta Crystallographica. A.* **49** (1993), 806–817.
- [2] E. MOLNÁR, Klassifikation der hyperbolischen Dodekaederpflasterungen von maximalen flächentransitiven Bewegungsgruppen, *Math. Pannonica*, **4/1** (1993), 113–136.
- [3] E. MOLNÁR, Projektive metrics and hyperbolic volume, *Annales Univ. Sci. Budapest. Sect. Math.*, **32** (1989), 127–157.
- [4] A. W. M. DRESS–D. H. HUSON, On tilings of the plane, *Geometriae Dedicata*, **24** (1987), 295–310.

- [5] E. MOLNÁR, Symmetry breakings of the cube tilings and the spatial chess board by  $D$ -symbols, *Beiträge Alg. Geom.*, **35** (1994), 205–238.
- [6] E. MOLNÁR–I. PROK, Classification of solid transitive simplex tilings in simply connected 3-spaces I, *Colloquia Math. Soc. J. Bolyai* **63**, *Intuitive Geometry*, Szeged (Hungary) 1991, North-Holland, 311–362 (1994).
- [7] E. MOLNÁR, Polyhedron complexes with simply transitive group actions and their realizations, *Acta Math. Hung.*, **59** (1992), 175–216.
- [8] E. MOLNÁR, On a family of four-dimensional simplex tilings and its  $d$ -dimensional variant, *Publicationes Math. Debrecen* (to appear).
- [9] J. SZIRMAI, Typen von flächentransitiven Würfelpflasterungen, *Annales Univ. Sci. Budapest. Sect. Math.* **37** (1994), 171–184.

# AN ALMOST SURE LOCAL AND GLOBAL CENTRAL LIMIT THEOREM FOR WEAKLY DEPENDENT RANDOM VARIABLES

By

GONCHIGDANZAN HURELBAATAR

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received October 20, 1994)

## 1. Introduction

In recent years the investigations of almost sure central limit theorem has attracted considerable attention in the literature. The basic results of these studies are the almost sure central limit theorems proved by SCHATTE (1988) and BROSAMLER (1988) for i.i.d. random variables having finite  $(2 + \varepsilon)$ th moments. Later LACEY and PHILIPP (1990) proved similar results assuming only finite variance.

**THEOREM 1.1.** *Let  $X_1, X_2, \dots$  be i.i.d. r.v.'s with  $\mathbf{E}X_i = 0$ ,  $\mathbf{E}X_i^2 = 1$  and set  $S_n = X_1 + \dots + X_n$ . Then*

$$(1.1) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \mathbf{I} \left\{ \frac{S_k}{\sqrt{k}} \leq x \right\} = \Phi(x) \quad \text{a.s.}$$

where  $\Phi(x)$  is the standard normal distribution function and  $\mathbf{I}$  denotes indicator function.

The same result was proved by PELIGRAD and SHAO (1995) for r.v.'s with finite second moment under the condition (2.7) with  $\xi_k = S_k / \sqrt{\mathbf{E}|S_k|^2}$  and the usual central limit theorem holds. The functional version of (1.1) was proved by HURELBAATAR (1993) for  $\varphi$ -mixing r.v.'s having finite  $(2 + \varepsilon)$ th moments.

An interesting result of this kind was proved by CSÁKI, FÖLDES, and RÉVÉSZ (1993). They proved the following theorem:

**THEOREM 1.2.** *Let  $S_1, S_2, \dots$  be sequences of sums of i.i.d. r.v.'s having expectation 0 and finite third moment. Let  $\{a_i, i \geq 1\}$  and  $\{b_i, i \geq 1\}$  be*

sequences of real numbers such that  $a_i \leq 0$ ,  $b_i \geq 0$ . Assume that

$$(1.2) \quad \sum_{k=1}^n \frac{\log k}{k^{\frac{3}{2}} \mathbf{P}(a_k \leq S_k < b_k)} = O(\log n), \quad \text{as } n \rightarrow \infty$$

Then

$$(1.3) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{\mathbf{I}\{a_k \leq S_k < b_k\}}{k \mathbf{P}(a_k \leq S_k < b_k)} = 1 \quad \text{a.s.}$$

This theorem may be called almost sure local central limit theorem, while (1.1) may be called almost sure global central limit theorem.

In this paper we give an extension to  $\rho$ -mixing case of the almost sure central limit theorem proved by HURELBAATAR (1993) for a sequence of  $\varphi$ -mixing random variables. We also establish a sufficient condition of (1.3) for  $\rho$ -mixing stationary random variables.

## 2. Definitions, notations and auxiliary results

Let  $\{X_i, i \geq 1\}$  be a sequence of r.v.'s on the probability space  $(\Omega, \mathcal{F}, \mathbf{P})$  and let  $M_a^b$  be the  $\sigma$ -algebra generated by the random variables  $\{X_i, a \leq i \leq b\}$ .

Define

$$\varphi(n) = \sup_{k \geq 1} \sup_{A \in M_1^k, B \in M_{k+n}^\infty} |\mathbf{P}(B|A) - \mathbf{P}(B)|$$

and

$$\rho(n) = \sup_{k \geq 1} \sup_{X \in L_2(M_1^k), Y \in L_2(M_{k+n}^\infty)} \frac{|\text{Cov}(X, Y)|}{(\text{Var} X)^{\frac{1}{2}} (\text{Var} Y)^{\frac{1}{2}}}.$$

The sequence  $\{X_i, i \geq 1\}$  is said to be  $\varphi$ -mixing (respectively  $\rho$ -mixing) if  $\varphi(n) \rightarrow 0$  (respectively  $\rho(n) \rightarrow 0$ ) as  $n \rightarrow \infty$ .

In the sequel we shall denote by  $C, C_1, C_2, \dots$  various positive constants and by  $\Rightarrow$  the weak convergence. Throughout this paper  $\Phi(x)$  will denote the standard normal distribution function and  $\mathbf{W}$  the standard Wiener process.

The following earlier results will be used in the proofs.

THEOREM 2.1. (Theorem 1.1 in SHAO, Q. M. (1993)). Assume that  $\mathbf{E}X_i = 0$  and  $\mathbf{E}|X_i|^t < \infty$  for some  $t \geq 2$ . Then there exists a positive constant  $K = K(t, \varrho(\cdot))$  depending only on  $t$  and  $\varrho(\cdot)$  such that for any  $k \geq 0, n \geq 1$

$$(2.1) \quad \mathbf{E} \max_{1 \leq i \leq n} |S_k(i)|^t \leq K \left( n^{\frac{t}{2}} \exp \left( K \sum_{i=0}^{[\log n]} \varrho(2^i) \right) \right) \max_{k \leq i \leq k+n} (\mathbf{E}X_i^2)^{\frac{t}{2}} + \\ + n \exp \left( K \sum_{i=0}^{[\log n]} \varrho^{\frac{2}{t}}(2^i) \right) \max_{k \leq i \leq k+n} (\mathbf{E}|X_i|)^t$$

where  $S_k(n) = \sum_{i=k+1}^{k+n} X_i$  and  $\log x$  stands for the logarithm with base 2.

THEOREM 2.2. (Theorem in ZUPAROV, T. M. (1991)). Let  $\{X_n, n \geq 1\}$  be a stationary sequence of random variables with mixing coefficient

$$(2.2) \quad \varrho(n) \leq Cn^{-\gamma}, \quad \gamma > 0$$

and

$$\mathbf{E}X_1 = 0, \quad \mathbf{E}|X_1|^s < \infty,$$

$$(2.3) \quad 2 < s < s_0(\gamma) = \frac{\gamma - 1}{\gamma} + \sqrt{\left(\frac{\gamma - 1}{\gamma}\right)^2 + \frac{4 + 2\gamma}{\gamma}}.$$

If

$$(2.4) \quad \mathbf{E}S_n^2 \geq \mu n \mathbf{E}X_1^2, \quad \mu > 0$$

then there exists a constant  $C(s, \gamma, k, \mu)$  depending only on  $s, \gamma, k$  and  $\mu$ , such that

$$(2.5) \quad \sup_x \left| \mathbf{P} \left( S_n \leq x \sqrt{\mathbf{E}S_n^2} \right) - \Phi(x) \right| \leq C(s, \gamma, k, \mu) \frac{\beta_s}{n^{\frac{s-2}{2}}}$$

where  $\beta_s = \frac{\mathbf{E}|X_1|^s}{(\mathbf{E}X_1^2)^{\frac{s}{2}}}$ .

REMARK on Theorem of ZUPAROV. By Theorem 4.1 in PELIGRAD (1982) it is clear that there exists  $n_0 \in \mathbf{N}$  such that the condition (2.4) holds for all  $n \geq n_0$ . So we can see that (2.5) remains valid for  $n \geq n_0$  under the only conditions (2.2) and (2.3). We also use the following.

LEMMA 2.1. Assume that  $\xi_1, \xi_2, \dots$  are random variables such that  $\mathbf{E}\xi_k = 1, k = 1, 2, \dots$ . Then we have

$$(2.6) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \mathbf{E} \left( \sum_{k=1}^n \frac{\xi_k}{k} \right) = 1.$$

If furthermore  $\xi_k \geq 0, k = 1, 2, \dots$  and

$$(2.7) \quad \mathbf{Var} \left( \sum_{k=1}^n \frac{\xi_k}{k} \right) \leq C \log^{(2-\varepsilon)} n, \quad n = 1, 2, \dots$$

with some  $\varepsilon > 0$ , then we have

$$(2.8) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{\xi_k}{k} = 1 \quad \text{a.s.}$$

PROOF. (2.6) is obvious. (2.8) can be proved by applying Chebyshev's inequality for a subsequence  $n_k = \exp(k^{\frac{2}{2-\varepsilon}})$  (see BERKES and DEHLING (1993) and SCHATTE (1988)).

LEMMA 2.2. Assume that we have  $\xi_k \geq 0$  and

$$(2.9) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{\xi_k}{k} = 1 \quad \text{a.s.}$$

and the sequence  $\{\eta_k\}$  is such that for any  $\varepsilon > 0$  there exists a  $k_0 = k_0(\varepsilon, \omega)$  for which

$$(2.10) \quad (1 - \varepsilon)\xi_k \leq \eta_k \leq (1 + \varepsilon)\xi_k, \quad k \geq k_0.$$

Then we have

$$(2.11) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{\eta_k}{k} = 1 \quad \text{a.s.}$$

The proof is obvious.

THEOREM 2.3. (Theorem 2.5 in PELIGRAD, M. (1982)). Let  $\{X_n, n \geq 1\}$  be a  $\rho$ -mixing sequence of random variables with  $\mathbf{E}X_i = 0$  satisfying the following three conditions

$$(2.12) \quad \lim_{n \rightarrow \infty} \frac{\mathbf{E}S_n^2}{n} = \sigma^2 > 0,$$

$$(2.13) \quad \sum_{n=1}^{\infty} \rho^{\frac{1}{2}}(2^n) < \infty,$$

$$(2.14) \quad \{X_n^2, n \geq 1\} \text{ are uniformly integrable.}$$

Then

$$(2.15) \quad \frac{S_{[nt]}}{\sigma\sqrt{n}} \Rightarrow W.$$

REMARK 2.1. By Theorem 4.1 in PELIGRAD (1982) if stationary,  $\rho$ -mixing random variables having the finite second moments satisfy

$$(2.16) \quad \sum_{n=1}^{\infty} \rho(2^n) < \infty$$

then the condition (2.12) holds.

### 3. Main results

Setting  $S_n = X_1 + \dots + X_n$  define the usual "broken line" process  $\hat{s}_n(t)$  on  $[0, 1]$  for some numerical sequence  $d_n > 0$ , by

$$(3.1) \quad \hat{s}_n(t) = \begin{cases} \frac{S_j}{d_n}, & \text{for } t = t_{nj}, \quad 0 \leq j \leq n, \\ \text{linear in between.} \end{cases}$$

Here for every  $n \geq 1$ ,  $0 = t_{n0} < t_{n1} < \dots < t_{nn} = 1$  is a partition of  $[0, 1]$  such that

$$(3.2) \quad \max_{1 \leq j \leq n} |t_{n,j-1} - t_{nj}| \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

If  $t_{nj} = \frac{j}{n}$ , then the partition is called uniform.

THEOREM 3.1. Let  $\{X_i, i \geq 1\}$  be a sequence of  $\rho$ -mixing r.v.'s, with  $\mathbf{E}X_i = 0$ , satisfying (2.16) and

$$(3.3) \quad n \max_{1 \leq j \leq n} \mathbf{E}X_j^2 \leq Cd_n^2, \quad (n \geq 1)$$

where  $d_n$  is the same as in (3.1). assume that

$$(3.4) \quad \frac{d_l}{d_k} \geq C \left( \frac{l}{k} \right)^\gamma \quad (l \geq k \geq k_0)$$

with some  $\gamma > 0$ . Then for any Borel-subset  $A$  of  $D[0, 1]$  with  $\mu(\partial A) = 0$  ( $\mu$  is a probability measure on  $D[0, 1]$ ) the following two statements are equivalent:

$$(3.5) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \mathbf{I}\{\hat{s}_k(\cdot) \in A\} = \mu(A) \quad \text{a.s.}$$

and the exceptional set of probability zero does not depend on  $A$ .

$$(3.6) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \mathbf{P}\{\hat{s}_k(\cdot) \in A\} = \mu(A).$$

The equivalence of (3.5) and (3.6) was proved by BERKES and DEHLING (1993) for independent random variables under the condition

$$\mathbf{E}f\left(\left|\frac{S_n - b_n}{d_n}\right|\right) \leq (\log \log n)^{-1-\varepsilon} f\left(e^{(\log n)^{1-\varepsilon}}\right) \quad (n \geq n_0)$$

for some function  $f$  and numerical sequences  $d_n > 0$  and  $b_n$ . The equivalence of (3.5) and (3.6) was obtained by HURELBAATAR (1993) for  $\varphi$ -mixing random variables under the conditions of Theorem 3.1 replacing  $\varrho(*)$  by  $\varphi^{1/2}(*)$  in (2.16).

**THEOREM 3.2.** *Let  $\{X_i, i \geq 1\}$  be a sequence of  $\varrho$ -mixing random variables with  $\mathbf{E}X_i = 0$ ,  $\mathbf{E}|X_i|^{2+\varepsilon} < C$ , ( $i \geq 1$ ) for some  $\varepsilon > 0$  satisfying (2.12) and (2.13). Putting*

$$(3.7) \quad d_n = \sigma\sqrt{n}$$

in (3.1) and assuming uniform partition we have

$$(3.8) \quad \frac{1}{\log n} \sum_{i=1}^n \frac{1}{i} \delta\{\hat{s}_i(\cdot)\} \Rightarrow \mathbf{W} \quad \text{a.s.}$$

Here  $\delta\{x\}$  is the point mass at  $x \in D[0, 1]$ .

Let  $\{a_k\}$  and  $\{b_k\}$  be two sequences of real numbers such that

$$(3.9) \quad -\infty \leq a_k \leq 0 \leq b_k \leq \infty, \quad k = 1, 2, \dots$$

and put

$$(3.10) \quad P_k = \mathbf{P}(a_k \leq S_k < b_k)$$

and

$$(3.11) \quad \alpha_k = \begin{cases} \frac{\mathbf{I}\{a_k \leq S_k < b_k\}}{P_k} & \text{if } P_k \neq 0, \\ 1 & \text{if } P_k = 0. \end{cases}$$

Our second result is the following

**THEOREM 3.3.** *Let  $\{X_k, k \geq 1\}$  be a stationary sequence of random variables satisfying (2.2) and (2.3). Let  $a_k, b_k$  satisfy (3.9) and assume that*

$$(3.12) \quad P_k \geq \frac{c}{\log^\varepsilon k} \quad \text{and} \quad 0 < \varepsilon < \frac{s}{3+2s}$$

where  $s$  is as in (2.3). Then we have

$$(3.13) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{\alpha_k}{k} = 1 \quad \text{a.s.}$$

Our next result is

**THEOREM 3.4.** *Let  $\{X_k, k \geq 1\}$  be a stationary sequence of  $\rho$ -mixing random variables satisfying (2.2) and (2.3). Let  $a_k, b_k$  satisfy (3.9) and assume that*

$$(3.14) \quad C_1 k^{\frac{1}{2}-\beta} \leq |a_k| \leq C_2 k^{\frac{1}{2}-\beta}$$

and

$$(3.15) \quad C_1 k^{\frac{1}{2}-\beta} \leq b_k \leq C_2 k^{\frac{1}{2}-\beta},$$

where

$$(3.16) \quad 0 < \beta < \frac{\gamma}{2(1+3\gamma)}.$$

Then we have

$$(3.17) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{\alpha_k}{k} = 1 \quad \text{a.s.}$$

For the conditions (3.14) and (3.15) see Lemma 4.3.

#### 4. Proofs

Let  $BL(D[0, 1])$  denote the set of the functions  $g : D[0, 1] \rightarrow \mathbf{R}$  such that for some  $C > 0$

$$(4.1) \quad |g(x) - g(y)| \leq Cd(x, y), \quad |g(x)| \leq C \quad \text{for all } x, y \in D[0, 1]$$

where  $d$  is the Skorohod metric. To prove Theorem 3.1 it is enough to show that for any  $g \in BL(D[0, 1])$

$$(4.2) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{i=1}^n \frac{1}{i} \xi_i = 0 \quad \text{a.s.}$$

where  $\xi_i = g(\hat{s}_i(\cdot)) - \mathbf{E}g(\hat{s}_i(\cdot))$  (see BERKES and DEHLING (1993)).

In view of Lemma 2.1, if we show that (2.7) holds for any  $k < l$ , the function  $s_{k,l}^* : [0, 1] \rightarrow \mathbf{R}$  by

$$(4.3) \quad s_{k,l}^*(t) = \begin{cases} 0, & \text{if } 0 \leq t \leq t_{l,k} \\ \frac{S_j - S_k}{d_l}, & \text{if } t_{l,j} \leq t \leq t_{l,j+1} \quad (k \leq j \leq l-1) \end{cases}$$

where  $d_n$  is as in (3.1).

It is clear that  $d(\hat{s}_l, s_{k,l}^*) \leq \|\hat{s}_l - s_{k,l}^*\|_\infty$  and thus, by (4.1)

$$(4.4) \quad |g(\hat{s}_l) - g(s_{2k,l}^*)| \leq C \frac{\max_{1 \leq j \leq 2k} |S_j|}{d_l} \quad \text{for } 2k < l.$$

Now we consider the covariance of  $\xi_k$  and  $\xi_l$  for  $2k < l$ ,  $k \geq k_0$

$$(4.5) \quad \begin{aligned} |\mathbf{E}(\xi_k \xi_l)| &= \left| \mathbf{Cov} \left( g(\hat{s}_k), g(\hat{s}_l) \right) \right| = \left| \mathbf{Cov} \left( g(\hat{s}_k), g(\hat{s}_l) - g(s_{2k,l}^*) + g(s_{2k,l}^*) \right) \right| \leq \\ &\leq \left| \mathbf{Cov} \left( g(\hat{s}_k), g(\hat{s}_l) - g(s_{2k,l}^*) \right) \right| + \left| \mathbf{Cov} \left( g(\hat{s}_k), g(s_{2k,l}^*) \right) \right|. \end{aligned}$$

By definition of  $\rho$ -mixing

$$(4.6) \quad \left| \mathbf{Cov} \left( g(\hat{s}_k), g(s_{2k,l}^*) \right) \right| \leq C \rho(k)$$

since  $g(\hat{s}_k) \in L_2(M_1^k)$ ,  $g(s_{2k,l}^*) \in L_2(M_{2k}^\infty)$ .

Further by (4.1) and (4.4) we can see that

$$(4.7) \quad \left| \mathbf{Cov} \left( g(\hat{s}_k), g(\hat{s}_l) - g(s_{2k,l}^*) \right) \right| \leq C \frac{\mathbf{E}(\max_{1 \leq j \leq 2k} |S_j|)}{d_l}.$$

Condition (3.3), Theorem 2.1 and Cauchy–Schwarz inequality jointly imply that

$$(4.8) \quad \mathbf{E} \left( \max_{1 \leq j \leq 2k} |S_j| \right) \leq \left( \mathbf{E} \max_{1 \leq j \leq 2k} |S_j|^2 \right)^{\frac{1}{2}} \leq C d_{2k}.$$

Hence from (4.6), (4.7) and (4.8) it follows that

$$(4.9) \quad |\mathbf{E}(\xi_k \xi_l)| \leq C_1 \frac{d_{2k}}{d_l} + C_2 \rho(k) \quad \text{for } 2k < l.$$

Notice that

$$(4.10) \quad \begin{aligned} \mathbf{E} \left( \sum_{i=1}^n \frac{1}{i} \xi_i \right)^2 &\leq \\ &\leq C_1 \sum_{k=k_0}^n \frac{1}{k^2} \mathbf{E} |\xi_k|^2 + C_2 \sum_{\substack{k_0 \leq k < l \leq n \\ 2k \geq l}} \frac{|\mathbf{E}(\xi_k \xi_l)|}{kl} + C_3 \sum_{\substack{k_0 \leq k < l \leq n \\ 2k < l}} \frac{|\mathbf{E}(\xi_k \xi_l)|}{kl} = \\ &= \sum_1 + \sum_2 + \sum_3 \end{aligned}$$

Now by easy calculation we can see that

$$(4.11) \quad \sum_1 \leq C_1 \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty,$$

$$(4.12) \quad \sum_2 \leq C_2 \sum_{k=1}^n \sum_{l=k+1}^{2k} \frac{1}{kl} = O(\log n).$$

$$(4.13) \quad \sum_3 \leq C_3 \sum_{l=2}^n \sum_{k=1}^{l-1} \frac{1}{k^{1-\gamma} l^{1-\gamma}} + C_3 \sum_{l=2}^n \frac{1}{l} \sum_{k=1}^{l-1} \frac{\rho(k)}{k} = O(\log n),$$

where we used (2.16), (3.4), (4.9) and boundedness of  $\xi_i$ . The bounds (4.11)–(4.13) together imply (2.7). This completes the proof of Theorem 3.1.

PROOF OF THEOREM 3.2. Since  $E|X_n|^{2+\varepsilon} < C$  for all  $n$ , (2.14) holds. Hence from Theorem 2.3 it follows that under the conditions of our Theorem the weak convergence

$$(4.14) \quad \frac{S_{[nt]}}{\sigma\sqrt{n}} \Rightarrow \mathbf{W}(t)$$

holds, which in turn implies (3.6) with  $d_n$  given in (3.7), uniform partition of  $[0, 1]$  and the Wiener measure  $\mu$ . Thus our Theorem immediately follows from Theorem 3.1.

PROOF OF THEOREM 3.3. The main point of our proof is to verify the condition (2.7). We use the following elementary inequality and lemmas.

$$(4.15) \quad |\Phi(x) - \Phi(y)| \leq C|x - y| \quad x, y \in \bar{\mathbf{R}}.$$

First we prove our Theorem in the case of

$$(4.16) \quad b_k \leq C_1\sqrt{k} \quad \text{and} \quad a_k \geq -C_2\sqrt{k}, \quad k = 1, 2, \dots$$

LEMMA 4.1. Assume (2.2), (3.12) and (4.16). Then the following bounds hold:

$$(i) \quad \sum_{l=2}^n \sum_{k=1}^l \frac{1}{lP_l} \frac{\rho(k)}{kP_k} = O(\log^{2-\varepsilon} n),$$

$$(ii) \quad \sum_{l=2}^n \sum_{k=1}^l \frac{1}{lP_l} \frac{|a_k|}{k\sqrt{l}} = O(\log^{2-\varepsilon} n),$$

$$(iii) \quad \sum_{l=2}^n \sum_{k=1}^l \frac{1}{lP_l} \frac{b_k}{k\sqrt{l}} = O(\log^{2-\varepsilon} n).$$

The proof is obvious.

LEMMA 4.2. Assume that

$$(4.17) \quad \varepsilon_k = \sqrt{k} \log^\delta k \quad \text{and} \quad \frac{3\varepsilon}{s} < \delta < 1 - 2\varepsilon$$

where  $\varepsilon$  and  $s$  are as in (3.12). If (3.12) holds, then we have

$$(i) \quad \sum_{l=2}^n \sum_{k=1}^l \frac{1}{lP_l} \frac{\mathbf{P}(|S_k| \geq \varepsilon_k)}{kP_k} = O(\log^{2-\varepsilon} n),$$

$$(ii) \quad \sum_{l=2}^n \sum_{k=1}^l \frac{1}{lP_l} \frac{\mathbf{P}(|S_{2k}| \geq \varepsilon_k)}{k} = O(\log^{2-\varepsilon} n),$$

$$(iii) \quad \sum_{l=2}^n \sum_{k=1}^l \frac{1}{lP_l} \frac{\varepsilon_k}{k\sqrt{l}} = O(\log^{2-\varepsilon} n).$$

PROOF. It is clear that (ii) follows from (i). We can see that (i) and (iii) follow from Chebyshev's inequality and Theorem 2.1. This completes the proof of Lemma 4.2.

Let  $l > 2k$ . Then we may write

$$(4.18) \quad \begin{aligned} & \mathbf{Cov}(\alpha_k, \alpha_l) \leq \\ & \leq \frac{1}{P_k P_l} \left( \mathbf{P}(a_k \leq S_k < b_k, a_l - b_k \leq S_l - S_{2k} + S_{2k} - S_k < b_l - a_k) - P_k P_l \right) \\ & \leq \frac{1}{P_k P_l} \left( \mathbf{P}(|S_{2k} - S_k| \geq \varepsilon_k) - P_k P_l + \right. \\ & \quad \left. + \mathbf{P}(a_k \leq S_k < b_k, a_l - b_k - \varepsilon_k \leq S_l - S_{2k} < b_l - a_k + \varepsilon_k) \right) \end{aligned}$$

Here  $\varepsilon_k$  is as in (4.17). By the definition of  $\varrho$ -mixing it follows that

$$(4.19) \quad \begin{aligned} & \mathbf{P}(a_k \leq S_k < b_k, a_l - b_k - \varepsilon_k \leq S_l - S_{2k} < b_l - a_k + \varepsilon_k) \leq \\ & \leq \mathbf{P}(a_k \leq S_k < b_k) \mathbf{P}(a_l - b_k - \varepsilon_k \leq S_l - S_{2k} < b_l - a_k + \varepsilon_k) + \varrho(k) \end{aligned}$$

and

$$(4.20) \quad \begin{aligned} & \mathbf{P}(a_l - b_k - \varepsilon_k \leq S_l - S_{2k} < b_l - a_k + \varepsilon_k) \leq \\ & \leq \mathbf{P}(|S_{2k}| \geq \varepsilon_k) + \mathbf{P}(a_l - b_k - 2\varepsilon_k \leq S_l < b_l - a_k + 2\varepsilon_k). \end{aligned}$$

Now applying (2.4), (4.15) and Theorem 2.2 we get

$$(4.21) \quad \begin{aligned} & \mathbf{P}(a_l - b_k - 2\varepsilon_k \leq S_l < b_l - a_k + 2\varepsilon_k) - \mathbf{P}(b_l \leq S_l < b_l) = \\ & = \mathbf{P}(a_l - b_k - 2\varepsilon_k \leq S_l < a_l) + \mathbf{P}(b_l \leq S_l < b_l - a_k + 2\varepsilon_k) \leq \\ & \leq C_1 \frac{1}{l^{\frac{(s-2)}{2}}} + C_2 \frac{(|a_k| + b_k)}{\sqrt{l}} + C_3 \frac{\varepsilon_k}{\sqrt{l}}. \end{aligned}$$

Thus, by Lemmas 4.1 and 4.2 we get

$$(4.22) \quad \sum_{\substack{1 \leq k < l < n \\ 2k < l}} \frac{\text{Cov}(\alpha_k, \alpha_l)}{kl} = O(\log^{2-\varepsilon} n).$$

We can also see that for  $l \leq 2k$

$$(4.23) \quad \begin{aligned} & \sum_{\substack{1 \leq k < l < n \\ 2k \geq l}} \frac{1}{kl} \frac{\mathbf{P}(a_k \leq S_k < b_k, a_l \leq S_l < b_l) - P_k P_l}{P_k P_l} \leq \\ & \leq \sum_{\substack{1 \leq k < l < n \\ 2k \geq l}} \frac{1}{kl} \frac{(1 - P_k)}{P_k} \leq \sum_{\substack{1 \leq k < l < n \\ 2k \geq l}} \frac{1}{k P_k} \frac{1}{l} = O(\log^{2-\varepsilon} n). \end{aligned}$$

But  $\text{Var}(\alpha_k) = 0$  if  $P_k = 0$  and

$$(4.24) \quad \text{Var}(\alpha_k) = \frac{1 - P_k}{P_k} \leq \frac{1}{P_k} \quad \text{if} \quad P_k \neq 0.$$

Now it is easy to prove Theorem 3.3. We have

$$(4.25) \quad \begin{aligned} & \text{Var} \left( \sum_{k=1}^n \frac{\alpha_k}{k} \right) = \\ & = \sum_{k=1}^n \frac{\text{Var} \alpha_k}{k^2} + 2 \sum_{\substack{1 \leq k < l \leq n \\ 2k \geq l}} \frac{\text{Cov}(\alpha_k, \alpha_l)}{kl} + 2 \sum_{\substack{1 \leq k < l \leq n \\ 2k < l}} \frac{\text{Cov}(\alpha_k, \alpha_l)}{kl}. \end{aligned}$$

So (4.22), (4.23), (4.24) and (4.25) together imply

$$(4.26) \quad \text{Var} \left( \sum_{k=1}^n \frac{\alpha_k}{k} \right) \leq C \log^{(2-\varepsilon)} n$$

and hence, applying Lemma 2.1, Theorem 3.3 is proved under condition (4.16). Now we drop the condition (4.16). Fix  $x > 0$ . By Theorem 4.1 in PELIGRAD (1982) (2.12) holds. Hence we can define that

$$(4.27) \quad a_k^* = \max(a_k - \sigma x \sqrt{k}),$$

$$(4.28) \quad b_k^* = \min(b_k, \sigma x \sqrt{k}),$$

$$(4.29) \quad P_k^* = \mathbf{P}(a_k^* \leq S_k < b_k^*).$$

Then  $P_k^* \leq P_k$ , hence assuming  $P_k^* \neq 0$ ,

$$(4.30) \quad \frac{1}{P_k} \mathbf{I}\{a_k \leq S_k < b_k\} \leq$$

$$\begin{aligned} &\leq \frac{1}{P_k^*} \mathbf{I}\{a_k^* \leq S_k < b_k^*\} + \frac{1}{P_k} (\mathbf{I}\{a_k \leq S_k < a_k^*\} + \mathbf{I}\{b_k^* \leq S_k < b_k\}) \leq \\ &\leq \frac{1}{P_k^*} \mathbf{I}\{a_k^* \leq S_k < b_k^*\} + \frac{\mathbf{I}\{S_k < -\sigma x \sqrt{k}\}}{\mathbf{P}(-\sigma x \sqrt{k} \leq S_k < 0)} + \frac{\mathbf{I}\{S_k \geq \sigma x \sqrt{k}\}}{\mathbf{P}(0 \leq S_k < \sigma x \sqrt{k})}. \end{aligned}$$

By Theorem 2.3

$$(4.31) \quad \lim_{n \rightarrow \infty} \mathbf{P}(-\sigma x \sqrt{k} \leq S_k < 0) = \Phi(0) - \Phi(-x)$$

and

$$(4.32) \quad \lim_{n \rightarrow \infty} \mathbf{P}(0 \leq S_k < \sigma x \sqrt{k}) = \Phi(x) - \Phi(0).$$

Applying Theorem 3.2 and Lemma 2.2 we get

$$(4.33) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{\mathbf{I}\{S_k < -\sigma x \sqrt{k}\}}{k \mathbf{P}(-\sigma x \sqrt{k} \leq S_k < 0)} = \frac{\Phi(-x)}{\Phi(0) - \Phi(-x)} \quad \text{a.s.}$$

and

$$(4.34) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{\mathbf{I}\{S_k \geq \sigma x \sqrt{k}\}}{k \mathbf{P}(0 \leq S_k < \sigma x \sqrt{k})} = \frac{1 - \Phi(x)}{\Phi(x) - \Phi(0)} \quad \text{a.s.}$$

Since  $a_k^*$ ,  $b_k^*$  satisfy (4.16),

$$(4.35) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{\alpha_k^*}{k} = 1 \quad \text{a.s.}$$

where

$$(4.36) \quad \alpha_k^* = \begin{cases} \frac{1}{P_k^*} \mathbf{I}\{a_k^* \leq S_k < b_k^*\} & \text{if } P_k^* \neq 0, \\ 1 & \text{if } P_k^* = 0. \end{cases}$$

(4.30), (4.33), (4.34) and (4.35) together imply that

$$(4.37) \quad \limsup_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{\alpha_k}{k} \leq 1 + \frac{2(1 - \Phi(x))}{\Phi(x) - \Phi(0)} \quad \text{a.s.}$$

On the other hand, if  $P_k^* \neq 0$

$$\frac{1}{P_k} \mathbf{I}\{a_k \leq S_k < b_k\} \geq \frac{1}{P_k^*} \mathbf{I}\{a_k^* \leq S_k < b_k^*\} \left(1 - \frac{P_k - P_k^*}{P_k}\right) \geq$$

$$(4.38) \quad \geq \frac{1}{P_k^*} \mathbf{I} \{a_k^* \leq S_k < b_k^*\} \times \left( 1 - \frac{\mathbf{P}(S_k < -\sigma x \sqrt{k}) + \mathbf{P}(S_k > \sigma x \sqrt{k})}{\min(\mathbf{P}(-\sigma x \sqrt{k} \leq S_k < 0), \mathbf{P}(0 \leq S_k < \sigma x \sqrt{k}))} \right)$$

and by Theorem 2.3

$$(4.39) \quad \lim_{k \rightarrow \infty} \frac{\mathbf{P}(S_k < -\sigma x \sqrt{k}) + \mathbf{P}(S_k > \sigma x \sqrt{k})}{\min(\mathbf{P}(-\sigma x \sqrt{k} \leq S_k < 0), \mathbf{P}(0 \leq S_k < \sigma x \sqrt{k}))} = \frac{2(1 - \Phi(x))}{\Phi(x) - \Phi(0)}.$$

Applying (4.35), (4.38) and Lemma 2.4 we get

$$(4.40) \quad \liminf_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{\alpha_k}{k} \geq 1 - \frac{2(1 - \Phi(x))}{\Phi(x) - \Phi(0)} \quad \text{a.s.}$$

So, by the arbitrariness of  $x$  and (4.37), (4.40) implies

$$(4.41) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{\alpha_k}{k} = 1 \quad \text{a.s.}$$

This completes the proof of Theorem 3.3.

**PROOF OF THEOREM 3.4.** In view of Lemma 2.1 it suffices to show that the condition (2.7) holds. We use the following inequality and lemmas.

For each  $K$  there exists  $C = C(K)$  such that

$$(4.42) \quad |\Phi(x) - \Phi(y)| \geq C|x - y|,$$

for every  $x, y \in \mathbf{R}$ ,  $|x| + |y| \leq K$ .

**LEMMA 4.3.** *If  $a_k$  and  $b_k$  satisfy (3.14) and (3.15), then we have*

$$(4.43) \quad C_1 k^{-\beta} \leq P_k \leq C_2 k^{-\beta}, \quad (k \geq k_0).$$

**PROOF.** It follows from (3.16) and (2.3) that

$$(4.44) \quad \beta < \frac{s-2}{s}.$$

Hence applying Theorem 2.2 and (4.15)

$$(4.45) \quad \begin{aligned} \mathbf{P}(a_k \leq S_k < b_k) &\leq \left| \mathbf{P}\left(\frac{S_k}{\sqrt{k}} < \frac{b_k}{\sqrt{k}}\right) \right| + \left| \mathbf{P}\left(\frac{S_k}{\sqrt{k}} < \frac{a_k}{\sqrt{k}}\right) - \Phi\left(\frac{a_k}{\sqrt{k}}\right) \right| + \\ &+ \left| \Phi\left(\frac{b_k}{\sqrt{k}}\right) - \Phi\left(\frac{a_k}{\sqrt{k}}\right) \right| \leq \frac{C_1}{k^{\frac{s-2}{s}}} + \frac{C_2}{k^\beta} \leq \frac{C_3}{k^\beta}. \end{aligned}$$

If  $k$  is large enough, then using (4.42) we get

$$(4.46) \quad \begin{aligned} & \mathbf{P}(a_k \leq S_k < b_k) = \\ & = \mathbf{P}\left(\frac{S_k}{\sqrt{k}} < \frac{b_k}{\sqrt{k}}\right) - \Phi\left(\frac{b_k}{\sqrt{k}}\right) - \mathbf{P}\left(\frac{S_k}{\sqrt{k}} < \frac{a_k}{\sqrt{k}}\right) + \Phi\left(\frac{a_k}{\sqrt{k}}\right) + \\ & \quad + \Phi\left(\frac{b_k}{\sqrt{k}}\right) - \Phi\left(\frac{a_k}{\sqrt{k}}\right) \geq -\frac{C_1}{k^{\frac{s-2}{s}}} + C_2 \frac{b_k - a_k}{\sqrt{k}} \geq \frac{C_3}{k^\beta}. \end{aligned}$$

Our Lemma immediately follows from (4.45) and (4.46).

LEMMA 4.4. Assume that

$$(4.47) \quad 2\beta = \alpha\gamma \quad \text{and} \quad \varepsilon_l = l^{\frac{\alpha}{2} + 2\beta}.$$

Then the following asymptotic relations hold as  $n \rightarrow \infty$ .

$$(4.48) \quad \sum_{\substack{1 \leq k < l \leq n \\ k < l - l^\alpha}} \frac{\varrho(l^\alpha)}{klP_lP_k} = O(\log n)$$

under (2.2)

$$(4.49) \quad \sum_{\substack{1 \leq k < l \leq n \\ k < l - l^\alpha}} \frac{1}{klP_lP_k} \frac{l^{\frac{\alpha}{2}}}{\varepsilon_l} = O(\log n),$$

$$(4.50) \quad \sum_{\substack{1 \leq k < l \leq n \\ k < l - l^\alpha}} \frac{1}{klP_l} \frac{1}{(l-k-l^\alpha)^{\frac{s-2}{s}}} = O(\log n),$$

$$(4.51) \quad \sum_{\substack{1 \leq k < l \leq n \\ k < l - l^\alpha}} \frac{1}{klP_l} \left| \Phi\left(\frac{a_l - b_k + \varepsilon_l}{\sqrt{l-k-l^\alpha}}\right) - \Phi\left(\frac{a_l}{\sqrt{l}}\right) \right| = O(\log n),$$

$$(4.52) \quad \sum_{\substack{1 \leq k < l \leq n \\ k < l - l^\alpha}} \frac{1}{klP_l} \left| \Phi\left(\frac{b_l - a_k + \varepsilon_l}{\sqrt{l-k-l^\alpha}}\right) - \Phi\left(\frac{b_l}{\sqrt{l}}\right) \right| = O(\log n).$$

PROOF.

$$\begin{aligned} \sum_{\substack{1 \leq k < l \leq n \\ k < l - l^\alpha}} \frac{\varrho(l^\alpha)}{klP_lP_k} & \leq C_1 \sum_{1 \leq l \leq n} \frac{1}{l^{1-\beta+\alpha\gamma}} \sum_{1 \leq k < l - l^\alpha} \frac{1}{k^{1-\beta}} \leq \\ & \leq C_2 \sum_{1 \leq l \leq n} \frac{1}{l^{1-\beta+\alpha\gamma}} l^\beta \leq C_2 \sum_{1 \leq l \leq n} \frac{1}{l} = O(\log n), \end{aligned}$$

proving (4.48).

$$\begin{aligned} \sum_{\substack{1 \leq k < l \leq n \\ k < l - l^\alpha}} \frac{1}{klP_l P_k} \frac{l^{\frac{\alpha}{2}}}{l^{\frac{\alpha}{2} + 2\beta}} &\leq C_1 \sum_{1 \leq l \leq n} \frac{1}{l^{1+\beta}} \sum_{1 \leq k < l - l^\alpha} \frac{1}{k^{1-\beta}} \leq \\ &\leq C_2 \sum_{1 \leq l \leq n} \frac{1}{l^{1+\beta}} l^\beta \leq C_2 \sum_{1 \leq l \leq n} \frac{1}{l} = O(\log n), \end{aligned}$$

proving (4.49). By (4.44) we get

$$\begin{aligned} \sum_{\substack{1 \leq k < l \leq n \\ k < l - l^\alpha}} \frac{1}{klP_l} \frac{1}{(l - l^\alpha - k)^{\frac{s-2}{2}}} &\leq C \sum_{1 \leq l \leq n} \frac{1}{l^{1-\beta}} \\ &\left( \sum_{1 \leq k < \frac{l-l^\alpha}{2}} \frac{1}{k(l - l^\alpha - k)^{\frac{s-2}{2}}} + \sum_{\frac{l-l^\alpha}{2} < k < l - l^\alpha} \frac{1}{k(l - l^\alpha - k)^{\frac{s-2}{2}}} \right) \leq \\ &\leq C_1 \sum_{1 \leq l \leq n} \frac{1}{l^{1-\beta}} \left( \frac{1}{(l - l^\alpha)^{\frac{s-2}{2}}} \sum_{1 \leq k < \frac{l-l^\alpha}{2}} \frac{1}{k} + \frac{1}{l - l^\alpha} \sum_{1 \leq k < \frac{l-l^\alpha}{2}} \frac{1}{k^{\frac{s-2}{2}}} \right) \leq \\ &\leq C_2 \sum_{1 \leq l \leq n} \frac{\log l}{l^{1-\beta + \frac{s-2}{2}}} \leq C_3 \sum_{1 \leq l \leq n} \frac{1}{l} = O(\log n). \end{aligned}$$

Thus (4.50) is proved. Using (4.15) we can write

$$\begin{aligned} &\sum_{\substack{1 \leq k < l \leq n \\ k < l - l^\alpha}} \frac{1}{klP_l} \left| \Phi \left( \frac{a_l - b_k + \varepsilon_l}{\sqrt{l - k - l^\alpha}} \right) - \Phi \left( \frac{a_l}{\sqrt{l}} \right) \right| \leq \\ &\leq C \sum_{\substack{1 \leq k < l \leq n \\ k < l - l^\alpha}} \frac{|a_l|}{klP_l} \left( \frac{1}{\sqrt{l - l^\alpha - k}} - \frac{1}{\sqrt{l}} \right) + C \sum_{\substack{1 \leq k < l \leq n \\ k < l - l^\alpha}} \frac{1}{klP_l} \frac{b_k}{\sqrt{l - l^\alpha - k}} + \\ &+ C \sum_{\substack{1 \leq k < l \leq n \\ k < l - l^\alpha}} \frac{1}{klP_l} \frac{\varepsilon_l}{\sqrt{l - l^\alpha - k}} = \sum_1 + \sum_2 + \sum_3. \end{aligned}$$

Now applying the same procedure, as before we can obtain that

$$\sum_1 \leq C \sum_{\substack{1 \leq k < l \leq n \\ k < l - l^\alpha}} \frac{|a_l|}{klP_l} \left( \frac{1}{l^{1-\alpha} \sqrt{l - l^\alpha - k}} + \frac{k}{l \sqrt{l - l^\alpha - k}} \right) \leq$$

$$\begin{aligned}
&\leq C_1 \sum_{1 \leq l \leq n} \frac{1}{l^{\frac{3}{2}-\alpha}} \left( \frac{1}{(l-l^\alpha)^{\frac{1}{2}}} \sum_{1 \leq k < \frac{l-l^\alpha}{2}} \frac{1}{k} + \frac{1}{l-l^\alpha} \sum_{1 \leq k < \frac{l-l^\alpha}{2}} \frac{1}{k^{\frac{1}{2}}} \right) + \\
&\quad + C_1 \sum_{1 \leq l \leq n} \frac{1}{l^{\frac{3}{2}}} \sum_{1 \leq k < l-l^\alpha} \frac{1}{\sqrt{l-l^\alpha-k}} \leq \\
(4.53) \quad &\leq C_2 \sum_{1 \leq l \leq n} \frac{\log l}{l^{2-\alpha}} \leq C_3 \sum_{1 \leq l \leq n} \frac{1}{l} = O(\log n)
\end{aligned}$$

and

$$\begin{aligned}
\sum_2 &\leq C \sum_{\substack{1 \leq k < l \leq n \\ k < l-l^\alpha}} \frac{1}{l^{1-\beta} k^{\frac{1}{2}+\beta} \sqrt{(l-l^\alpha-k)}} \leq \\
&\leq C_1 \sum_{1 \leq l \leq n} \frac{1}{l^{1-\beta}} \left( \frac{1}{\sqrt{l-l^\alpha}} \sum_{1 \leq k < \frac{l-l^\alpha}{2}} \frac{1}{k^{\frac{1}{2}+\beta}} + \frac{1}{(l-l^\alpha)^{\frac{1}{2}+\beta}} \sum_{q \leq k < \frac{l-l^\alpha}{2}} \frac{1}{\sqrt{k}} \right) \leq \\
&\leq C_2 \sum_{1 \leq l \leq n} \frac{1}{l^{1-\beta}} \left( \frac{1}{\sqrt{l-l^\alpha}} l^{\frac{1}{2}-\beta} + \frac{1}{(l-l^\alpha)^{\frac{1}{2}+\beta}} \sqrt{l-l^\alpha} \right) \leq \\
(4.54) \quad &\leq C_3 \sum_{1 \leq l \leq n} \frac{1}{l} = O(\log n).
\end{aligned}$$

By (3.16) and (4.47) we can see that  $\alpha + 6\beta < 1$ . Hence

$$\begin{aligned}
\sum_3 &\leq C \sum_{\substack{1 \leq k < l \leq n \\ k < l-l^\alpha}} \frac{1}{l^{1-3\beta-\frac{\alpha}{2}} k \sqrt{(l-l^\alpha-k)}} \leq \\
&\leq C_1 \sum_{1 \leq l \leq n} \frac{1}{l^{1-3\beta-\frac{\alpha}{2}}} \left( \frac{1}{\sqrt{l-l^\alpha}} \sum_{1 \leq k < \frac{l-l^\alpha}{2}} \frac{1}{k} + \frac{1}{(l-l^\alpha)} \sum_{1 \leq k < \frac{l-l^\alpha}{2}} \frac{1}{\sqrt{k}} \right) \leq \\
(4.55) \quad &\leq C_2 \sum_{1 \leq l \leq n} \frac{\log l}{l^{\frac{3}{2}-3\beta-\frac{\alpha}{2}}} \leq C_3 \sum_{1 \leq l \leq n} \frac{1}{l} = O(\log n).
\end{aligned}$$

(4.53)–(4.55) imply (4.51). The proof of (4.52) is similar to the proof of (4.50). This completes the proof of Lemma 4.4.

Let  $k < l - l^\alpha$  ( $k \geq k_0$ ) for  $\alpha > 0$  as in (4.47). Then we write that

$$\begin{aligned}
 (4.56) \quad & \mathbf{Cov}(\alpha_k, \alpha_l) \leq \\
 & \leq \frac{1}{P_k P_l} \left( \mathbf{P}(a_k \leq S_k < b_k, a_l - b_k \leq S_l - S_{k+l^\alpha} + S_{k+l^\alpha} - S_k < b_l - a_k) - P_k P_l \right) \leq \\
 & \leq \frac{1}{P_k P_l} \left( \mathbf{P}(|S_l| \geq \varepsilon_l) - P_k P_l + \right. \\
 & \quad \left. + \mathbf{P}(a_k \leq S_k < b_k, a_l - b_k - \varepsilon_l \leq S_l - S_{k+l^\alpha} < b_l - a_k + \varepsilon_l) \right).
 \end{aligned}$$

By the definition of  $\rho$ -mixing it follows that

$$\begin{aligned}
 (4.57) \quad & \mathbf{P}(a_k \leq S_k < b_k, a_l - b_k - \varepsilon_l \leq S_l - S_{k+l^\alpha} < b_l - a_k + \varepsilon_l) \leq \\
 & \leq \mathbf{P}(a_k \leq S_k < b_k) \mathbf{P}(a_l - b_k - \varepsilon_l \leq S_{l-k-l^\alpha} < b_l - a_k + \varepsilon_l) + \rho(l^\alpha).
 \end{aligned}$$

Now using Theorem 2.1, Theorem 2.2, Chebyshev's inequality and (4.15) we get

$$\begin{aligned}
 (4.58) \quad & \mathbf{P}(a_l - b_k - \varepsilon_l \leq S_{l-k-l^\alpha} < b_l - a_k + \varepsilon_l) - \mathbf{P}(a_l \leq S_l < b_l) \leq \\
 & \leq \Phi\left(\frac{b_l - a_k + \varepsilon_l}{\sqrt{l-k-l^\alpha}}\right) - \Phi\left(\frac{b_l}{\sqrt{l}}\right) + \Phi\left(\frac{a_l}{\sqrt{l}}\right) - \Phi\left(\frac{a_l - b_k + \varepsilon_l}{\sqrt{l-k-l^\alpha}}\right) + \\
 & \quad + \frac{C_1}{(l-k-l^\alpha)^{\frac{s-2}{2}}} + \frac{C_2}{l^{\frac{s-2}{2}}}
 \end{aligned}$$

and

$$(4.59) \quad \mathbf{P}(|S_l| \geq \varepsilon_l) \leq C_3 \frac{l^{\frac{\alpha}{2}}}{\varepsilon_l}.$$

Applying (4.56)–(4.59), Lemma 4.4 imply that

$$(4.60) \quad \sum_{\substack{1 \leq k < l \leq n \\ k < l - l^\alpha}} \frac{\mathbf{Cov}(\alpha_k, \alpha_l)}{kl} = O(\log n).$$

On the other hand we can obtain that for  $l - l^\alpha < k < l$

$$(4.61) \quad \sum_{\substack{1 \leq k < l \leq n \\ k > l - l^\alpha}} \frac{\mathbf{Cov}(\alpha_k, \alpha_l)}{kl} = O(\log n).$$

We have

$$(4.62) \quad \mathbf{Var}\left(\sum_{k=1}^n \frac{\alpha_k}{k}\right) = \sum_{k=1}^n \frac{\mathbf{Var}\alpha_k}{k^2} + 2 \sum_{\substack{1 \leq k < l \leq n \\ k > l - l^\alpha}} \frac{\mathbf{Cov}(\alpha_k, \alpha_l)}{kl} + 2 \sum_{\substack{1 \leq k < l \leq n \\ k < l - l^\alpha}} \frac{\mathbf{Cov}(\alpha_k, \alpha_l)}{kl}.$$

So (4.60)–(4.62) and (4.24) imply that

$$(4.63) \quad \text{Var} \left( \sum_{k=1}^n \frac{\alpha_k}{k} \right) = O(\log n).$$

This completes our proof.

ACKNOWLEDGEMENTS. I am much indebted to E. CSÁKI for his many comments that improved the presentation of this paper. I am also grateful to the referee for many useful comments. Research supported by Hungarian National Foundation for Scientific Research, Grant No. 1905.

### References

- [1] P. BILLINGSLEY, *Convergence of Probability Measures*, Willey, New York. (1968).
- [2] I. BERKES and H. DEHLING, Some limit theorems in log density, *Ann. Probab.*, **21** (1993), 1640–1670.
- [3] G.A. BROSAMLER, An almost everywhere central limit theorem, *Math. Proc. Camb. Phil. Soc.* **104** (1988), 561–574.
- [4] E. CSÁKI, A. FÖLDES and P. RÉVÉSZ, On almost sure local and global central limit theorems, *Probab. Theory Relat. Fields*, **97** (1993), 321–337.
- [5] G. HURELBAATAR, On the almost sure central limit theorem for  $\varphi$ -mixing random variables. *Studia Sci. Math. Hung.* (to appear)
- [6] M.T. LACEY and W. PHILIPP, A note on the almost sure central limit theorem. *Stat. Probab. Letters*, **9** (1990), 201–205.
- [7] M. PELIGRAD, Invariance principles for mixing sequences of random variables, *Ann. Probab.*, **10** (1982), 968–981.
- [8] M. PELIGRAD and Q.M. SHAO, A note on the almost sure central limit theorem for weakly dependent random variables, *Stat. Probab. Letters* **22** (1995), 131–136.
- [9] Q.M. SHAO, Maximal inequalities for partial sums of  $\rho$ -mixing sequences, *Res. Report, National University of Singapore*, (1993), N. 586.
- [10] P. SCHATTE, On strong versions of the central limit theorem, *Math. Nachr.*, **137** (1988), 249–256.
- [11] T.M. ZUPAROV, On a rate in the central limit theorem for weakly dependent random variables, *Theory Probab. Appl.*, **36** (1991), 639–644.

## ASCOLI–ARZELÁ THEOREM FOR A CLASS OF RIGHT CONTINUOUS FUNCTIONS

By

JOSÉ MENESES and RAÚL NAULIN

Departamento de Matemáticas, Universidad de Oriente, Cumaná, Venezuela

*(Received October 31, 1994)*

### 1. Introduction

Spaces of noncontinuous functions can be found in many branches of theoretical analysis. The field of applications of these functions involves the theory of stochastic processes [2], impulsive differential equations and differential equations with non continuous right hand sides [1], [7], control theory [4], etc ...

The space of noncontinuous functions we introduce in this article is motivated by the theory of impulsive equations. In this case it is necessary to work with solutions that are piecewise continuous functions, and allow points of discontinuity of the first kind [1]. Frequently these solutions have separated discontinuities, that is, if  $t_1$  and  $t_2$  are two points of discontinuity of a given solution, then they satisfy  $|t_1 - t_2| \geq m$ , where  $m$  is a fixed positive number. In this paper we show that these solutions are contained in a metric space of right continuous functions where an Ascoli–Arzelá theorem can be proven.

### 2. A class of piecewise continuous functions

In this section  $J$  will denote one of the following intervals

$$[t_0, t_1], [t_0, t_1), [t_0, \infty).$$

If the interval is finite then we denote  $\tilde{J} = J \setminus \{t_1\}$ . If  $J$  is not finite then  $\tilde{J} = J$ .

DEFINITION 1. We say that  $f : J \rightarrow \mathbb{R}$  belongs to the space  $C_+(J)$  iff

- (i) The function  $f$  is bounded.

- (ii) The interval  $\bar{J}$  can be written in the form  $\bigcup_{k \in \mathcal{F}} I_k$ , where  $\mathcal{F}$  is a subset of the set of natural numbers,  $I_k$  is an interval of the form  $I_k = [t_k, t_{k+1})$ , where  $t_k$  is a strictly increasing sequence of real numbers (this sequence depends on each function  $f$ ).
- (iii) The function  $f$  is continuous on each interval  $I_k$ .
- (iv) For each  $t_{k+1}$  the left hand side limit  $\lim_{t \rightarrow t_{k+1}^-} f(t)$  exists, and this limit will be denoted by  $f(t_{k+1}^-)$ .

In the space  $C_+(J)$  we define the norm

$$\|f\|_\infty = \sup\{|f(t)|; t \in J\}.$$

It is easy to see that the pair  $(C_+(J), \|\cdot\|_\infty)$  defines a normed space, and this space is not complete. This last observation follows from the forthcoming:

EXAMPLE. Let  $J = [0, 1]$ ; for  $n$  a natural number let us define  $f_n(x) = 1$ , for  $x \in [1/n + 1, 1/n)$  and  $f_n(x) = 0$  is  $x \in [0; 1] \setminus [1/n + 1, 1/n)$ . Then the functions

$$g_n(x) = \sum_{k=1}^n f_k(x)/2^k,$$

are functions belonging to the set  $C_+(J)$ . This sequence converges uniformly to the function

$$g(x) = \sum_{k=1}^{\infty} f_k(x)/2^k,$$

not contained in  $C_+(J)$ .

DEFINITION 2. For a positive number  $m$ , we denote by  $C_+(J, m)$  the subset of those functions of  $C_+(J)$  with the following property: if  $t_1, t_2$  are points of discontinuities of  $f$  then  $|t_1 - t_2| \geq m$ .

THEOREM 1.  $(C_+(J, m), \rho)$ ,  $\rho(g, f) = \|f - g\|_\infty$  is a complete metric space.

PROOF. Let  $(f_n)_{n=1}^\infty$  be a Cauchy sequence on the space  $(C_+(J, m), \rho)$ . Then this sequence converges uniformly to some function  $f : J \rightarrow \mathbb{R}$ . If  $f$  is not continuous at some  $t_*$ , then for some subsequence  $(f_{n_j})_{n=1}^\infty$  we have that all functions  $f_{n_j}$  are discontinuous at  $t_*$ . Since  $f_{n_j} \in C_+(J, m)$ , then all functions  $f_{n_j}$  are continuous on the interval  $[t_*, t_* + m)$ . This implies that the function  $f$  is continuous on this interval and the points of discontinuity of  $f$  are located in a distance greater than  $m$ .

Let us assume now that  $f$  is continuous on some interval  $[t^*, t_*)$ . We shall show that there exists the  $\lim_{t \rightarrow t_*^-} f(t)$ . Let  $(t_n)_{n=1}^\infty$  be a sequence such that  $t_n < t_*$  and  $\lim_{n \rightarrow \infty} t_n = t_*$ . Then the inequality

$$|f(t_n) - f(t_m)| \leq |f(t_n) - f_k(t_n)| + |f(t_m) - f_k(t_m)| + |f_k(t_n) - f_k(t_m)|,$$

and the existence of  $\lim_{t \rightarrow t_*^-} f_k(t)$ , show that for a given  $\varepsilon > 0$ , there exists a large

$N$  such that  $n, m \geq N$  imply  $|f(t_m) - f(t_n)| < \varepsilon$ . Thus there exists the limit  $\lim_{n \rightarrow \infty} f(t_n)$ . It is easy to see that this limit does not depend on the sequence  $(t_n)_{n=1}^\infty$ . ■

REMARKS.

R1: If  $J = [a, b]$ , then a function  $f \in C_+(J, m)$  has at most  $\left[ \frac{b-a}{m} \right]$  points of discontinuity, where  $[ \ ]$  denotes the greatest integer function.

R2: The space  $C_+(J, m)$  recall the space of right continuous functions  $\mathcal{D}_{[0,1]}$  with the Skorochod metric [2]. For  $J = [a, b]$  the set of functions  $C_+(J, m)$  is not in the space  $\mathcal{D}_{[0,1]}$ , because a function belonging to  $C_+(J, m)$  not necessarily is left continuous at  $t = b$ . The Skorochod metric used in the theory of stochastic processes seems to be appropriated to study spaces of discontinuous functions, but working with impulsive differential equations it is difficult to give a description of these deterministic equations with such a metric.

### 3. The Ascoli-Arzelá Theorem in $C_+(J, m)$

Let  $\mathcal{F}$  be a family of functions belonging to the space  $C_+(J)$ , and defined on the interval  $J = [a, b]$ . In order to establish the main result of the section we need the following definitions:

DEFINITION 3. We shall say that  $\mathcal{F}$  is a bounded family iff there exists a positive number  $M$  such that  $|f(t)| \leq M$  for all  $t \in J$  and all  $f \in \mathcal{F}$ .

DEFINITION 4. We shall say that  $\mathcal{F}$  is a right equicontinuous set of functions in the space  $C_+(J)$  iff for any  $\varepsilon > 0$  and any  $x$  in  $J$ , there exists a  $\delta(x, \varepsilon) > 0$  such that

(i) For any  $x \in J$ , any  $f \in \mathcal{F}$  and any  $y \in [x, x + \delta) \cap J$  one has

$$|f(x) - f(y)| < \varepsilon.$$

(ii) For any  $x \in J$ , any  $f \in \mathcal{F}$  and any  $y \in (x - \delta, x) \cap J$  one has

$$|f(x^-) - f(y)| < \varepsilon.$$

EXAMPLE. Let  $J = [0, 1]$ . Let  $q$  be a rational number in  $J$ . Let us define  $f_q(x) = 0$  if  $x \in [0, q)$ ,  $f_q(x) = 1$  if  $x \in [q, 1]$ . This sequence is right continuous at no point  $x$  belonging to  $[0, 1]$ .

DEFINITION 5. We shall say that a family of functions  $\mathbf{F} \subset C_+(J, m)$  is compact iff from any sequence of functions contained in  $\mathbf{F}$  it is possible to select a subsequence converging to a function of  $C_+(J, m)$ .

We will prove the following

THEOREM 2. Let  $J = [a, b]$ . A family  $\mathcal{F}$  contained in  $C_+(J, m)$ , is compact iff the family  $\mathcal{F}$  is bounded and right continuous.

PROOF OF THE NECESSITY. Let us assume that the family  $\mathcal{F}$  is compact. Then for any  $\varepsilon > 0$ , there exists a finite  $\varepsilon$ -net of family  $\mathcal{F}$ , say  $N = \{f_1, f_2, \dots, f_n\}$ . From here it follows the boundedness of the family  $\mathcal{F}$ . Let  $x$  be a fixed point of the interval  $[a, b]$ . Then there exists a  $\delta > 0$  such that for any  $f_i \in N$  we have

$$(1) \quad |f_i(y) - f_i(x)| < \varepsilon \quad \text{if } 0 \leq y - x < \delta, y \in [a, b],$$

and

$$(2) \quad |f_i(y) - f_i(x^-)| < \varepsilon \quad \text{if } 0 < x - y < \delta, y \in [a, b].$$

For a fixed function  $f$  of  $\mathcal{F}$  we can choose an index  $i$ , such that uniformly on all interval  $[a, b]$  we will have  $|f_i(t) - f(t)| < \varepsilon$ . This property, together with (1) and (2) give

$$|f(y) - f(x)| < 3\varepsilon, \quad \text{if } 0 \leq y - x < \delta, y \in [a, b],$$

and

$$|f(y) - f(x^-)| < 3\varepsilon, \quad \text{if } 0 \leq x - y < \delta, y \in [a, b],$$

implying the right equicontinuity of function  $f$ .

PROOF OF THE SUFFICIENCY. Let  $(f_n)_{n=1}^\infty$  be a sequence contained in  $\mathcal{F}$ .  $D_n$  denotes the set of points of discontinuity of the function  $f_n$ . Since  $f_n$  belongs to  $C_+(J, m)$ , then the set  $D_n$  is finite. Let

$$D = \bigcup_n D_n = \{d_1, d_2, d_3, \dots\}.$$

Let  $E$  denote a countable dense set in  $J \setminus D$ :  $E = \{e_1, e_2, e_3, \dots\}$ . Let us consider the following sequences:

$$(f_n(e_1))_{n=1}^\infty, \quad (f_n(d_1))_{n=1}^\infty, \quad (f_n(d_1^-))_{n=1}^\infty.$$

These sequences are bounded. Therefore there exists a subsequence  $(f_n^1)_n^\infty$  of  $(f_n)_n^\infty$  assuring the existence of the limits:

$$Y_1 = \lim_{n \rightarrow \infty} f_n^1(e_1), \quad Z_1 = \lim_{n \rightarrow \infty} f_n^1(d_1), \quad W_1 = \lim_{n \rightarrow \infty} f_n^1(d_1^-).$$

If we repeat the preceding process for the sequences

$$\left(f_n^1(e_2)\right)_{n=1}^\infty, \quad \left(f_n^1(d_2)\right)_{n=1}^\infty, \quad \left(f_n^1(d_2^-)\right)_{n=1}^\infty,$$

we can define a subsequence  $(f_n^2)_{n=1}^\infty$  of  $(f_n^1)_{n=1}^\infty$  such that there exist the limits:

$$Y_2 = \lim_{n \rightarrow \infty} f_n^2(e_2), \quad Z_2 = \lim_{n \rightarrow \infty} f_n^2(d_2), \quad W_2 = \lim_{n \rightarrow \infty} f_n^2(d_2^-).$$

Thus we can construct a family of sequences  $(f_n^j)_{n=1}^\infty$ , where  $(f_n^{j+1})_{n=1}^\infty$  is a subsequence of  $(f_n^j)_{n=1}^\infty$  and for each  $j$  we have the following property:

$$\begin{aligned} Y_i &= \lim_{n \rightarrow \infty} f_n^j(e_i), & 1 \leq i \leq j, \\ Z_i &= \lim_{n \rightarrow \infty} f_n^j(d_i), & 1 \leq i \leq j, \\ W_i &= \lim_{n \rightarrow \infty} f_n^j(d_i^-), & 1 \leq i \leq j. \end{aligned}$$

Let us define the diagonal sequence  $g_n = f_n^n$ . Then for all  $i = 1, 2, \dots$ , this sequence satisfies

$$(3) \quad Y_i = \lim_{n \rightarrow \infty} g_n(e_i), \quad Z_i = \lim_{n \rightarrow \infty} g_n(d_i), \quad W_i = \lim_{n \rightarrow \infty} g_n(d_i^-).$$

Let us define on  $E \cup D$  the function  $f$  by

$$f(e_i) = Y_i, \quad f(d_i) = Z_i.$$

By standard arguments, similar to those of the Ascoli-Arzelá theorem for continuous functions, it is possible to show that for any  $x \in J$  the sequence  $(g_n(x))_{n=1}^\infty$  is a Cauchy sequence and therefore for any  $x \in J$  we can define  $f(x)$  by means of:

$$f(x) = \lim_{n \rightarrow \infty} g_n(x).$$

We will show that  $(g_n)_{n=1}^\infty$  converges uniformly to  $f$ . By the right equicontinuity of  $(g_n)_{n=1}^\infty$  at  $x$  we have, that for a given positive number  $\varepsilon > 0$ , the existence of a  $\delta(x, \varepsilon)$  such that

$$0 \leq y - x < \delta(x, \varepsilon) \text{ implies } |g_n(y) - g_n(x)| < \varepsilon/3,$$

and

$$0 < x - y < \delta(x, \varepsilon) \text{ implies } |g_n(y) - g_n(x^-)| < \varepsilon/3.$$

But the family of open intervals

$$\{(x - \delta(x, \varepsilon), x + \delta(x, \varepsilon))\},$$

is an open covering of  $J$  from which we can select a finite subcovering, which we shall denote by

$$(x_i - \delta_i, x_i + \delta_i), \quad i = 1, 2, \dots, q.$$

Let  $x \in J$ , then  $x$  is contained in some interval  $(x_i - \delta_i, x_i + \delta_i)$ . We have two possibilities P1 and P2:

P1:  $x \in [x_i, x_i + \delta_i)$ . In this case we have the inequality:

$$\begin{aligned} |g_n(x) - g_m(x)| &\leq |g_n(x) - g_n(x_i)| + |g_m(x) - g_m(x_i)| + |g_m(x_i) - g_n(x_i)| \leq \\ &\leq \varepsilon/3 + \varepsilon/3 + |g_m(x_i) - g_n(x_i)|, \end{aligned}$$

where the convergence of  $(g_n(x_i))_{n=1}^{\infty}$  allows to fix an  $N_i$ , such that for  $n \in N_i$  we will have

$$|g_n(x) - f(x)| < \varepsilon, \quad \text{for } n \geq N_i.$$

P2:  $x \in (x_i - \delta_i, x_i)$ . In this case we will have

$$\begin{aligned} (4) \quad |g_n(x) - g_m(x)| &\leq |g_n(x) - g_n(x_i^-)| + |g_m(x) - g_m(x_i^-)| + \\ &+ |g_m(x_i^-) - g_n(x_i^-)| \leq \varepsilon/3 + \varepsilon/3 + |g_m(x_i^-) - g_n(x_i^-)|. \end{aligned}$$

If  $x_i \notin D$  then all functions  $g_n$  are continuous in  $x_i$  and therefore  $g_n(x_i^-) = g_n(x_i)$ . Because the sequence  $(g_n(x_i))_{n=1}^{\infty}$  converges, then for some positive  $N_i$  we have that

$$(5) \quad n, m \geq N_i \quad \text{implies} \quad |g_n(x) - f(x)| \leq \varepsilon.$$

Finally, if  $x_i \in D$ , then from (3) we have the convergence of the sequence  $g_n(x_i^-)$  and for some large  $N_i$  we will have

$$n, m \geq N_i \quad \text{implies} \quad |g_n(x_i^-) - g_m(x_i^-)| \leq \varepsilon,$$

and again from (4) we obtain the estimate (5).

Since  $i = 1, \dots, q$ , then for  $N = \max\{N_1, N_2, \dots, N_q\}$  we have:

$$|g_n(x) - f(x)| \leq \varepsilon \quad \text{for all } n \geq N \text{ and } x \in J = [a, b].$$

Thus the sequence  $g_n(x)$  converges uniformly to  $f$ . Since the space  $C_+(J, m)$  is complete, then  $f \in C_+(J, m)$ . ■

### 4. Examples

In this section we will consider two examples where we apply Theorem 2.

#### 4.1. Approximation of continuous functions

Let us consider a real continuous function  $f$  defined on the interval  $[a, b]$ . Then for each  $\varepsilon > 0$ , there exists a  $\delta > 0$  such that, for any pair of numbers  $x, y$  belonging to  $[a, b]$  with  $|x - y| < 2\delta$ , we have  $|f(x) - f(y)| < \varepsilon$ . Let us consider  $\mathcal{E}_{\delta, \varepsilon}$  the set of functions defined by the following properties:  $g \in \mathcal{E}_{\delta, \varepsilon}$  implies the existence of a partition  $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$  of interval  $[a, b]$  such that

- $g = \text{constant}$  on  $[x_{i-1}, x_i]$ ,  $i = 1, 2, \dots, n - 2$  and  $[x_{n-1}, 1]$ .
- $|f(x) - g(x)| < \varepsilon$  for all  $x \in [a, b]$  and all  $g \in \mathcal{E}_{\delta, \varepsilon}$ .
- The points belonging to the partition  $\mathcal{P}$  satisfy  $|x_{i+1}, x_i| \geq \delta$ .

We term the functions to the set  $\mathcal{E}_{\delta, \varepsilon}$  the  $\delta$ -step functions. Let us define

$$V(g) = |f - g|_{\infty}; \quad g \in \mathcal{E}_{\delta, \varepsilon}.$$

Because  $V(g) < \varepsilon$  for all  $g \in \mathcal{E}_{\delta, \varepsilon}$ , then there exists sequence  $(g_n)_{n=1}^{\infty}$  in  $\mathcal{E}_{\delta, \varepsilon}$  such that

$$\xi \stackrel{\text{def}}{=} \text{Inf} \{ V(g); g \in \mathcal{E}_{\delta, \varepsilon} \} = \lim_{n \rightarrow \infty} V(g_n).$$

The bounded set of functions  $\mathcal{E}_{\delta, \varepsilon}$  is right continuous, therefore using the Ascoli-Arzelá Theorem 2, we can assume that the sequence  $(g_n)_{n=1}^{\infty}$  uniformly converges to function  $g_0$ . It is easy to see that  $g_0 \in \mathcal{E}_{\delta, \varepsilon}$ . Moreover:

$$\xi = \lim_{n \rightarrow \infty} |g_n - f|_{\infty} = |g_0 - f|_{\infty}.$$

This last identity shows that the upper lower bound of the functional  $V(g)$  is attained at the set  $\mathcal{E}_{\delta, \varepsilon}$ . In other words, there exists a  $\delta$ -step function, giving the best approximation to the continuous function  $f$  in the set  $\mathcal{E}_{\delta, \varepsilon}$ . ■

## 4.2. An optimal control in impulsive systems

Let us consider a particle moving from the starting point  $(1, 0)$  by the law of the following impulse differential system:

$$(6) \quad \begin{aligned} \begin{pmatrix} x' \\ y' \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \\ \begin{pmatrix} x \\ y \end{pmatrix} (t_k) &= \begin{pmatrix} 1 & 0 \\ 0 & \lambda_k \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} (t_k^-), \end{aligned}$$

where  $0 < t_1 < t_2 < \dots < t_m < T$  is a finite sequence of impulse times contained in a finite interval  $[0, T]$ , and  $T$  is a fixed positive number. The position and the number of impulse times in this interval can vary. We only require that the following condition of separation between impulse times must be observed:

$$t_{i+1} - t_i \geq \delta, \quad \delta \geq 0.$$

The numbers  $\lambda_k$  appearing in the definition of the impulsive systems stand to change the velocity of the particle at time  $t_k$ . The impulses  $\lambda_k$  can vary in the set

$$|\lambda - 1| \leq \gamma,$$

where  $\gamma$  is a positive number. Thus, the impulsive (6) system is defined by the pair  $(\mathcal{J}, \mathcal{L})$ , where

$$\mathcal{J} = \{t_1, t_2, \dots, t_m\},$$

is the set of impulsive times on the interval  $[0, T]$ , and

$$\mathcal{L} = \{\lambda_1, \lambda_2, \dots, \lambda_m\},$$

denotes the set of changes of velocity of this particle at chosen times  $\mathcal{J}$ . Let  $M$  be a point in  $\mathbb{R}^2$ . We will assume that this particle can be controlled by the impulsive equation (6) to the point  $M$  in the interval on time  $[0, T]$ , this means the existence of a pair  $(\mathcal{J}, \mathcal{L})$ , termed an admissible pair, such that the trajectory  $x(t, \mathcal{J}, \mathcal{L})$  starting from  $(1, 0)$  satisfies  $x(T, \mathcal{J}, \mathcal{L}) = M$ . Finally we assume that each impulse  $\lambda_k$  has a cost  $c_k$  and this situation defines the cost function:

$$C(\mathcal{J}, \mathcal{L}) = \sum_i^m c_k |\lambda_k - 1|.$$

Let  $\tilde{C} = \text{Inf} \{C(\mathcal{J}, \mathcal{L})\}$ . Our question concerns the existence of an admissible pair  $(\tilde{\mathcal{J}}, \tilde{\mathcal{L}})$ , for which  $\tilde{C} = C(\tilde{\mathcal{J}}, \tilde{\mathcal{L}})$ . Certainly, such optimal pair exists. To

prove this, we can identify each pair  $(\mathcal{J}, \mathcal{L})$  with a set of functions  $\mathcal{F}$  such that a function  $g \in \mathcal{F}$  is defined by the following properties:

$$g(t) = \begin{cases} 1, & t \in [0, t_1) \\ \lambda_k, & t \in [t_k, t_{k+1}), \quad k = 1, 2, \dots, m - 1. \\ \lambda_m, & t \in [t_m, T] \end{cases}$$

Let  $(\mathcal{J}_n, \mathcal{L}_n)$  be sequence of admissible pairs such that  $\tilde{C} = \lim_{n \rightarrow \infty} C(\mathcal{J}_n, \mathcal{L}_n)$ . Let us consider the respective sequence of functions  $g_n$ . This sequence is equicontinuous and bounded, and therefore by Theorem 2, we can assume that this sequence converges to a function  $g(t)$ . Clearly, this function is contained in  $\mathcal{F}$ , and therefore defines the optimal pair  $(\tilde{\mathcal{J}}, \tilde{\mathcal{L}})$ . ■

### References

- [1] D. D. BAINOV, and P. S. SIMEONOV, *Systems with Impulse Effect (Stability, Theory and Applications)*, Ellis Horwood and John Wiley, New York, 1989.
- [2] I. I. GIXMAN, and A. B. SKOROCHOD, *Introduction to Stochastic Processes*, Nauka, Moscow, 1977 (Russian).
- [3] A. N. KOLMOGOROV, S. V. FOMIN, *Elements of the Theory of Functions and Functional Analysis*, Graylock Press, Rochester, 1957.
- [4] E. F. LEE, L. MARKUS, *Foundations of Optimal Control Theory*, John Wiley, New York, 1967.
- [5] R. NAULIN, Two optimization problems in control theory of impulse differential equations, UDO, *Departamento de Matemáticas*, 1993 (Preprint).
- [6] R. NAULIN, and C. TAPIA, Asymptotic stability of linear impulsive systems, 1993 (submitted).
- [7] A. M. SAMOILENKO, and A. N. PERESTIUK, *Differential Equations with Impulse Effect*, Kiev University, Kiev, 1980 (Russian).



**GENERAL BOUNDARY VALUE PROBLEM FOR A  
DIFFERENTIATION OPERATOR**

By

I. JOÓ and A. MINKIN

Department of Analysis of the Eötvös Loránd University, Budapest, and  
Department of Mathematics, Saratov State University, Russia

(Received November 3, 1994)

**0. Introduction\***

In the present article we study the first order boundary value problem

$$(1) \quad -idy/dx = \lambda y, \quad -1 \leq x \leq 1,$$

$$(2) \quad U(y) \equiv \int_{-a}^a y(t) d\sigma(t) = 0$$

where  $\sigma(t)$  is a function of bounded variation on  $[-a, a]$ . Set  $L(\lambda) = U(\exp(i\lambda t))$  and let  $k_L(\lambda)$  be the zero-divisor of the generating function  $L(\lambda)$ , i.e. the non-negative integer-valued function  $\lambda \rightarrow k_L(\lambda)$  with support  $\Lambda = L^{-1}(0)$  which takes the value  $m$  if  $\lambda$  is the zero of  $L$  of order  $m$  and equals zero otherwise.

Set  $e_\lambda^j(x) = x^j \cdot \exp(i\lambda x)$ ,  $-a \leq x \leq a$ . The set

$$(3) \quad R(\Lambda) := \left\{ e_\lambda^j, j = 0, \dots, k_L(\lambda) - 1 \right\}_\Lambda^{**}$$

constitutes the family of all eigen- and generalized eigenfunctions of the problem (1)–(2).

Using the concept of a sine-type function B. YA. LEVIN and V. D. GOLOVIN have established the following deep facts concerning spectral properties of the system (3) [7], [3].

---

\* The research was supported by the Hungarian National Foundation OTKA No T014244

\*\* The sign “:=” means: equal by definition.

THEOREM A (B. YA. LEVIN, V. D. GOLOVIN). *Assume that*

$$(4) \quad \sigma(a) - \sigma(a - 0) \neq 0, \quad \sigma(-a) - \sigma(-a + 0) \neq 0.$$

*Then*

$$(5) \quad \sup \{ \operatorname{Im} \lambda \mid \lambda \in \Lambda \} < \infty;$$

$$(6) \quad \sum_{s \leq \operatorname{Re} \lambda \leq s+1} k_L(\lambda) = O(1), \quad -\infty < s < \infty;$$

*there exists a certain partition*  $\{\Lambda_n\}_{n=1}^{\infty}$  *of*  $\Lambda$ ,

$$(8) \quad \bigcup_{n=1}^{\infty} \Lambda_n = \Lambda, \quad \Lambda_n \cap \Lambda_m = \emptyset, \quad n \neq m$$

*and the family of subspaces*  $E = \{E_n\}_{n=1}^{\infty}$ ,  $E_n := \operatorname{span} R(\Lambda_n)$  *constitutes an unconditional basis in*  $L^2(-a, a)$ . *Further we shall write briefly*  $E \in (UB)$ .

Afterwards V. A. MOLODENKOV and A. P. HRIMOV proved convergence and equiconvergence with a trigonometric series for corresponding eigenfunction expansions provided (4) is valid [11], [12]. Further results about convergence, equiconvergence and summability of eigenfunction expansions for this and more general boundary value problems may be found in the series of A. M. SEDLETSKII's works (see the review [17]) as well as in the important paper by G. M. GUBREEV [4].

However, nothing is known when the conditions (4) are dropped. Of course today two criteria of unconditional basisness of exponentials in  $L^2(-a, a)$  are established. One is due to S. A. AVDONIN and I. JOÓ and describes zero-divisors of the generating function as "small" perturbations of zero-sets of sine-type function provided the spectrum lies in the strip (5) [1], [2]. The zero-set of a sine-type function has also a very interesting description via its behavior in the critical points, i.e. in the zeros of the derivative [5].

Another one has been established by B. S. PAVLOV-N. K. NIKOL'SKII-S. V. HRUSCEV in the case  $\inf \{ \operatorname{Im} \lambda \mid \lambda \in \Lambda \} > -\infty$  [15], [13], [6], [16] and by A. M. MINKIN in the general case [9]. This criterion gives a *full description* on unconditional bases from exponentials in terms of their generating function  $L(\lambda)$ . Nevertheless from the view-point of boundary value problems it is important to establish a criterion in terms of a boundary condition itself. But it is obvious that such a translation from the generating function's language to the one of the boundary condition yields a new problem which seems to be nearly as difficult as the original each.

Therefore in this article we restrict ourselves to the case of the boundary condition (2) and aim to investigate a spectrality of the problem (1)–(2).

### 1. Notations and definitions

Let  $k(\lambda)$  be a divisor in the complex plane  $\mathcal{C}$  with a discrete support  $\Lambda$  which has a single limiting point at  $\infty$  and let

$$(9) \quad \Lambda = \bigcup_{n=1}^{\infty} \Lambda_n$$

be its partition into disjoint subsets  $\Lambda_n$  called clusters. Further we set

$$(10) \quad k_n(\lambda) = k(\lambda), \quad \lambda \in \Lambda_n \quad \text{and} \quad = 0 \text{ otherwise;} \quad (k(\lambda) = \sum k_n(\lambda))$$

and assume that  $\Lambda_n$  lies in a circle  $K_n$  ( $C, C_1 > 0, C_1 < 1$ )

$$(11) \quad \Lambda_n \subset K_n := \{ |z - \xi_n| \leq r_n \}; \quad r_n \leq C + C_1 \cdot |\text{Im} \xi_n|;$$

$$(12) \quad \sup_n \text{Card} \Lambda_n = N < \infty, \quad \text{Card} \Lambda_n := \sum_{\lambda} k_n(\lambda).$$

DEFINITION 1. Let  $E_n$  be the span  $(e_{\lambda}^j \mid j = 0, \dots, k(\lambda) - 1; \lambda \in \Lambda_n)$ ;  $E := \{E_n\}$ . Suppose (9)–(12) are valid and  $E \in (UB)$ . Then we shall say that  $E$  is a *block-basis* in  $L^2(-a, a)$  (notation:  $E \in (BB)$ ).

In the sequel we shall need a *block-basis* generalization of the criterion in [9].

Let  $\Lambda_+$  be the union of all  $\Lambda_n$  such that  $\sup\{\text{Im} \lambda \mid \lambda \in \Lambda\} \geq 0$ ;  $\Lambda_-$  be its complement in  $\Lambda$  and  $k_{\pm}(\lambda)$  be their divisors.

THEOREM B [10]. *Suppose the family of subspaces  $E$  is defined as above. Then  $E \in (BB)$  if and only if the following relations hold:*

- (A)  $k(\lambda)$  is a zero-divisor of some entire function  $L(z)$  of exponential type  $a$ ;
- (B)  $\text{dist}(\Lambda_n, \Lambda_m) \geq \varepsilon > 0, n \neq m$ ;
- (C)  $\Lambda_n^{\perp} := \bigcup \{ \Lambda_n \mid \inf(\pm \text{Im} K_n \geq h) \in (CV); \text{ for some fixed } h > 0;$
- (D)  $|M(\cdot - iy)|^2 \in (A_2)$ , for some  $y > 0$ .

Here  $\text{dist}(X, Y) := \inf\{|x - y| \mid x \in X, y \in Y\}$ ;  $(S_2)$  and  $(CV)$  are the well-known Muckenhoupt and Carleson–Vasjunin conditions [14] and  $M(z)$  is an entire function of the first order where zero-divisor

$$(13) \quad k_M(\lambda) := k_+(\lambda) + k_-(\lambda).$$

REMARK 1. Let  $\Theta_{\pm a} \equiv \exp(\pm iaz)$ ;  $B(z, k)$  be the Blaschke product with divisor  $k$ . Then in Theorem B one can replace (D) by

$$(14) \quad \text{dist}_{L^{\infty}} \left( \Theta_a^2 B(\cdot, k_M(\cdot - iy)), H_{\pm}^{\infty} \right) < 1.$$

Set also  $\mathcal{C}_{\pm} := \{\pm \text{Im} z \geq 0\}$  and  $\mathcal{C}_q := \{\text{Im} z \geq q\}$ .

## 2. Main results

THEOREM 1. Assume

$$(15) \quad \sigma(-a+0) - \sigma(-a) = 0, \quad \sigma(a) - \sigma(a-0) \neq 0.$$

Then

$$(16) \quad (i) \quad \inf\{\operatorname{Im}\lambda \mid \lambda \in \Lambda\} > -\infty.$$

$$(17) \quad (ii) \quad E \notin (BB).$$

An analogous statement is valid when the measure  $d\sigma$  carries a mass at the point  $-a$  while another end-point  $+a$  is mass-free.

THEOREM 2. Assume

$$(18) \quad \sigma(-a+0) - \sigma(-a) = 0, \quad \sigma(a) - \sigma(a-0) = 0$$

and let  $d\sigma$  be a sum of a purely discrete measure  $d\sigma_d$  and absolutely continuous part  $f dx$ :  $d\sigma = d\sigma_d + f dx$ ,  $f \in L^1(-a, a)$ ,

$$\int_{-a}^a y d\sigma_d(t) = \sum_j c_j y(a_j); \quad \sum |c_j| < \infty, \quad a_j \in (-a, a).$$

Then

$$(19) \quad E \notin (BB)$$

## 3. Proof of Theorem 1

3.1. Let  $c_x$  be the jump of  $d\sigma$  at the point  $x$ . Then

$$L(\lambda) - c_a \exp(+i\lambda a) + L_0(\lambda); \quad L_0(\lambda) = \int_{-a}^{a-0} \exp(i\lambda t) d\sigma(t),$$

$c_a \neq 0$ , and it is obvious that

$$|L_0(\lambda)| \leq \varepsilon \cdot |\exp(i\lambda a)|, \quad \operatorname{Im}\lambda \leq -N_\varepsilon < 0.$$

Therefore  $L(\lambda)$  has no zeros in the lower half-plane  $\{\operatorname{Im}\lambda < -N_\varepsilon\}$  and (16) is proved. Clearly

$$(20) \quad |L(\lambda)| \asymp |\exp(i\lambda a)|, \quad \operatorname{Im}\lambda \leq -N_\varepsilon^*$$

provided  $\varepsilon$  is chosen less than  $|c_a|$ .

\* Here and in the sequel  $A \asymp B$  means that  $d_1 \leq |A/B| \leq d_2$  with some positive constants  $d_1, d_2$  which don't depend on variables  $A, B$ .

3.2. Suppose now on the contrary that  $E \in (BB)$ . Without loss of generality assume that  $\inf\{\text{Im}\lambda \mid \lambda \in \Lambda\} > 0$  and that (20) is valid for  $N_\varepsilon = 0$ . Otherwise we shall replace  $L(\lambda)$  by  $L(\lambda - iN_\varepsilon)$ . Then we employ Theorem B and observe that the spectrum  $\Lambda \in (CV)$ . It is well-known that the latter implies the following relation:

$$(21) \quad |B(z, k_\Lambda)| \asymp 1, \quad z \in K_\delta$$

$$K_\delta := \{\text{Im}z \geq 0, |z - \lambda| \geq \delta \cdot |\text{Im}\lambda|, \lambda \in \Lambda\}.$$

Recall that  $L(z)$  is a function of the class Cartwright. But basicity implies completeness in  $L^2(-a, a)$  and the latter yields the fact that the indicator diagram of  $L(z)$  is an interval of length  $\geq 2a$ . At the same time this diagram lies inside the interval  $[-ia, ia]$  because  $L(z)$  is a Fourier–Laplace transform of a measure supported in  $[-a, a]$ . Therefore the diagram coincides with the interval  $[-ia, ia]$ . Hence

$$L(z) = \Pi(z) := \text{v.p.} \prod \left(1 - \frac{z}{\lambda}\right)^{k(\lambda)},$$

the product here is taken on the principal value sense over all  $\lambda \in \Lambda$ . Next

$$(22) \quad B(z, k_\Lambda) = \Pi(z) / \Pi^\sim(z),$$

where  $\Pi^\sim(z) := \overline{\Pi(\bar{z})}$ . Since we put  $N_\varepsilon = 0$  we have (see (20)) that

$$|\Pi(z)| \asymp |\Theta_a(z)|, \quad z \in \mathcal{G}_-$$

or equivalently

$$|\Pi^\sim(z)| \asymp |\Theta_{-a}(z)|, \quad z \in \mathcal{G}_+$$

whence, taking into account (21) and (22) we conclude that

$$(23) \quad |L(z)| \equiv |\Pi^\sim(z)| \cdot |B(z, k_\Lambda)| \asymp |\Theta_{-a}(z)|, \quad z \in K_\delta.$$

From the other hand

$$(24) \quad L(z) = \left( \int_{-a+\varepsilon}^a + \int_{-a}^{-a+\varepsilon} \right) e^{izt} d\sigma(t) = L_\varepsilon(z) + L^\varepsilon(z).$$

The first summand  $L_\varepsilon(z)$  is an entire function of exponential type  $\leq a - \varepsilon$  in  $\mathcal{G}_+$ . The second one admits an estimate

$$(25) \quad |L^\varepsilon(z)| \leq \exp(a \cdot \text{Im}z) \cdot \int_{-a}^{-a+\varepsilon} |d\sigma(t)|, \quad z \in \mathcal{G}_+.$$

The second factor in the right-hand side of (25) tends to zero as  $\varepsilon \rightarrow 0$  since the point  $-a$  carries no mass. Thus (23) fails and we came to a contradiction which completes the proof.

#### 4. The proof of Theorem 2

Set

$$L_d(z) := \int_{-a}^a \exp(izt) d\sigma_d(t) \quad \text{and} \quad F(z) := \int_{-a}^a \exp(izt) f(t) dt.$$

Then  $L(z) = L_d(z) + F(z)$  and its behaviour is described by the following

LEMMA 1. *In any square*

$$R(t, h) := \{z : |\operatorname{Re} z - t| \leq h, |\operatorname{Im} z| \leq h\},$$

$h$  being any fixed positive number,  $L(z)$  has  $O(1)$  zeros counting multiplicities and outside their  $\delta$ -neighborhoods  $L(z)$  is uniformly bounded from below:

$$(26) \quad |L(z)| \geq \varepsilon > 0, \quad z \in S_\delta(2h, \Lambda) := \{|\operatorname{Im} z| \leq 2h\} \setminus S_\delta(\Lambda);$$

$$S_\delta(\Lambda) := \bigcup_{\lambda \in \Lambda} C_{\lambda, \delta} := \{|\lambda z| \leq \delta\}.$$

PROOF. This statement is well-known for the almost-periodic function  $L_d(z)$  [8, Ch. VI. §2] and the proof therein exploits two facts:

1. normality of the set of functions  $\Phi_t(z) := L_d(z + t)$ ,  $z \in R(0, 2h)$ , where the parameter  $t$  runs through the real axis;
2. for any subsequence  $t_k \rightarrow +\infty$  or  $t_k \rightarrow -\infty$  as  $k \rightarrow \infty$  the limiting function  $\lim_{k \rightarrow \infty} \Phi_{t_k}$  (if exists) doesn't vanish identically.

However these properties are also satisfied for the generating function  $L(z)$ . Indeed it is clear from the definition that  $|L(z)|$  is bounded from above (and hence normal) in the strip

$$(27) \quad |\operatorname{Im} z| \leq 2h.$$

Further  $|F(z)| \rightarrow 0$  as  $z \rightarrow \infty$  remaining in the strip (27). Thus  $F(z)$  doesn't affect the limiting function  $\lim_{k \rightarrow \infty} L(z + t_k)$  at all. Then it remains to repeat considerations from [8] word for word and the lemma is proved.

Now choose in (26)  $h$  such that  $-\frac{h}{2} < \inf \operatorname{Im} \Lambda_+$ . Upon reflecting  $\Lambda_n \in \Lambda_-$  to the upper half-plane we come to a new generating function  $M(z)$  instead of  $L(z)$  with a divisor of zeros  $k_M(z)$  defined by (13). Clearly

$$(28) \quad \inf M^-(0) > -\frac{h}{2},$$

$$(29) \quad M(z)/L(z) = B\left(z, k_{\Lambda_-}^-\right).$$

The zero-divisor of  $B(z, k_{\bar{\Lambda}_-})$  satisfies the Carleson–Vasjunin condition. Therefore

$$B(z, k_{\bar{\Lambda}_-}) \asymp 1, \quad z \in S_\delta(2h, \bar{\Lambda}_-), \quad 0 \leq \text{Im} z \leq 2h.$$

Since  $B(\bar{z}) = B^{-1}(z)$  for any Blaschke product  $B(z)$  we can extend this estimate to the lower half-plane:

$$B(z, k_{\bar{\Lambda}_-}) \asymp 1, \quad z \in S_\delta(2h, \bar{\Lambda}_- \cup \Lambda_-).$$

Taking into account (29) we get that

$$M(z) \asymp L(z), \quad z \in S_\delta(2h, \bar{\Lambda}_- \cup \Lambda_-),$$

whence, employing (26) we conclude that

$$(30) \quad |M(z)| \asymp 1, \quad z \in S_\delta(2h, \bar{\Lambda}_- \cup \Lambda_-).$$

According to (28)  $M(z)$  doesn't vanish in the domain  $\text{Im} z \leq -\frac{h}{2}$ .

Let us choose now  $\delta$  in (26) such that  $(2N + 1)\delta < \frac{h}{2}$  and consider connected components of the union of circles  $C_{\lambda, \delta}$  over all  $\lambda \in \Lambda_-$ . Each component consists at most of  $N$  circles (see (12)). On account of the maximum modulus principle the estimate (30) is valid inside any component lying entirely in the domain  $-2h \leq \text{Im} z \leq -\frac{h}{2}$ .

All other components are placed outside the region

$$(31) \quad -2h + 2N\delta \leq \text{Im} z \leq -\frac{h}{2} - 2N\delta.$$

Hence

$$(32) \quad M(z) \asymp 1, \quad \text{whenever } z \text{ satisfies (31).}$$

All the more  $M(z) \asymp 1, \text{Im} z = -h$ .

Further on, let us introduce a function  $M_1(z) := M(z - ih)$ . Since  $M_1(t) \asymp 1, -\infty < t < \infty$  it is a function of the class Cartwright. Besides it has no zeros in  $\mathcal{C}_-$ . In [9, §8] it has been explained that  $M_1(z)$  might be taken as a product

$$M_1(z) = \text{v.p.} \Pi \left( 1 - \frac{z}{p} \right)^{k_1(v)},$$

over all  $v \in M + ih := \{\mu + ih \mid \mu \in M\}, k_1(v) := k_M(v - ih)$ . Moreover it has an indicator diagram  $[-ia, a]$  coinciding with that of the generating function  $L(z)$ . Here we merely put  $k = 0$  in [9, (8.9)] because a generating function is uniquely determined up to a factor  $\exp(ikz)$  which can be taken identically 1. Thus we have

$$M_1(z) \asymp \Theta_a(z), \quad z \in \mathcal{C}_-$$

whence

$$(33) \quad M_1(z) \asymp \Theta_{-a}(z), \quad z \in \mathcal{C}_+.$$

On the other hand recalling (29) with the argument  $z - ih$  we obtain that

$$(34) \quad M_1(z) \cdot \Theta_a(z) \rightarrow 0, \quad \text{Im} z \rightarrow +\infty$$

just as  $L(z - ih) \cdot \Theta_a(z - ih)$  does because the point  $-a$  is mass free. It remains to see that

$$(35) \quad M_1(z)/M_1^{\sim}(z) = B(z, k_1) \in (CV).$$

Upon collecting together (33) and (34) we conclude that

$$M_1(z)/M_1^{\sim}(z) \rightarrow 0 \quad \text{as} \quad \text{Im} z \rightarrow +\infty.$$

Thus we came to a contradiction with (35) which completes the proof.

## 5. Final remarks

Theorems 1, 2 give a background for the following main conjecture

CONJECTURE 1. *Assume that the measure  $d\sigma$  in (2) is not purely absolutely continuous,  $d\sigma \neq f dx$  and conditions (4) fail. Then*

$$E \notin (BB).$$

As a first step towards its proof it will be helpful to investigate the case  $d\sigma_d \neq \emptyset$ , i.e. when there exists a discrete component of the measure  $d\sigma$ . The situation with purely absolutely continuous measure  $d\sigma \equiv f dx, f \in L^1(-a, a)$  needs a separate treatment. We shall mention here only the papers [17], [18] and the articles cited therein.

## References

- [1] S. A. AVDONIN, I. JOÓ, Riesz bases of exponentials and sine-type functions, *Acta Math. Hung.*, **51** (1988), 3–14.
- [2] S. A. AVDONIN, I. JOÓ, M. HORVÁTH, Riesz bases from elements of the type  $x^k e^{i\lambda_n x}$ , *Vestnik LGU*, ser. 1, **4** (22) (1989), 3–7, (Russian).
- [3] V. D. GOLOVIN, On biorthogonal expansions in  $L^2$  in linear combinatorics of exponential functions, *Zap. Mekh. Mat. Fak. Khrakov, Gos. Univ. i Khrakov, Mat. Obschch*, ser. 4, **30** (1964), 18–29, (Russian).

- [14] G. M. GUBREEV, Spectral analysis of biorthogonal expansions of functions in series of exponentials, *Izv. AN USSR, ser. matem.* **53** (1989), 1236–1268, (Russian).
- [15] A. E. EREMENKO and M. L. SODIN, Parametrization of Entire Functions of Sine-Type by Their Critical Values, *Advances in Soviet Mathematics*, **1** (1992), 237–242.
- [16] S. V. HRUSCEV, Perturbation theorems for bases from exponentials and Muckenhoupt condition, *Dokl. AN USSR*, **247** (1979), 44–48, (Russian).
- [17] B. YA. LEVIN, On bases of exponential functions in  $L_2$ , *Zap. Mekh. Mat. Fak. Khrakov, Gos. Univ. i Khrakov. Mat. Obschch.*, ser. 4, **27** (1961), 39–48, (Russian).
- [18] B. YA. LEVIN, *Distribution of zeros of entire functions*, Moscow, GITTL (1956); English transl. Amer. Math. Soc., Providence, RI, 1964.
- [19] A. M. MINKIN, Reflection of exponents and unconditional exponential bases, *Algebra i Analiz*, **3** (1991), 10–135; English transl. in *St. Petersburg Math. J.*, **3** (1992), 1043–1068.
- [10] A. M. MINKIN, Projection method and unconditional bases. Algebraic and topological methods in mathematical physics (Conference held at Katziveli, Ukraine, 1–14 Sept., 1993).
- [11] V. A. MOLODENKOV, Eigenfunction expansion of one boundary value problem, *Diff. uravneniya i vytsisl. matematika*, **2** (1975), 56–65, (Russian).
- [12] V. A. MOLODENKOV, A. P. HROMOV, Eigenfunction expansion of one boundary problem for a differentiation operator, *Diff. uravneniya i vytsisl. matematika*, **1** (1972), 17–26, (Russian).
- [13] N. K. NIKOL'SKII, Bases from exponentials and from values of reproducing kernels, *Dokl. AN USSR*, **252** (1980), 1316–1320; English transl. in *Soviet Math. Dolk.*, **21** (1980).
- [14] N. K. NIKOL'SKII, *Treatise on the shift operator*, Moscow, Nauka, (1980), (Russian).
- [15] B. S. PAVLOV, The basicity of a system of exponentials and the Muckenhoupt's condition, *Dokl. Akad. Nauk USSR*, **247** (1979), 37–40; English transl. in *Soviet Math. Dolk.*, **20** (1979).
- [16] S. V. HRUSCEV, N. K. NIKOL'SKII, B. S. PAVLOV, Unconditional bases of exponentials and reproducing kernels, *Lect. Notes Math.*, **864** (1981), 214–335.
- [17] A. M. SEDLETSKII, Biorthogonal expansions of functions in series of exponentials in intervals of the real axis, *Uspekhi matem. nauk.*, **37** (1982), 51–95, (Russian).
- [18] A. M. SEDLETSKII, On the zeros of Fourier's transform of a finite measure, *Matema. Zametki*, **53** (1993), 111–120, (Russian).



## A METHOD FOR CHARACTERIZATION OF WEIGHTED K-FUNCTIONALS

By

NGUYEN XUAN KY\*

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received November 16, 1994)

### 1. Introduction

The concept of  $K$ -functional introduced by J. PETREE plays an important role in the theory of interpolation spaces and approximation theory. In this paper we present a simple method for characterization of weighted  $K$ -functionals. Let  $v(x)$  be a weight on a finite or infinite interval  $I = (a, b)$ . This means in this paper that  $v(x) > 0$  a.e. and measurable on  $(a, b)$ . Let  $1 \leq p \leq \infty$  and let  $r$  be an integer. The weighted  $K$ -functional is defined as

$$(1) \quad K_r(f, t)_X = \inf_{g \in W^{(r)}(X)} \{ \|f - g\|_X + t^r \|g^{(r)}\|_X \} \quad (f \in X, \quad t > 0),$$

where

$$X := X[I, p, v] := \{f : vf \in L^p[I]\}, \quad \|f\|_X := \|vf\|_{L^p[I]}$$

and  $W^{(r)}(X)$  denotes the class of all functions  $g$   $r$ -times locally absolutely continuous on  $(a, b)$  for which  $g^{(r)} \in X$ .

For characterization of the  $K$ -functional (1) we introduce the following moduli of smoothness of functions in the space  $X$ :

$$(2) \quad \omega_r(f, \delta)_X := \sup_{|h| \leq \delta} \|v_{r,h}(x) \Delta_h^r f(x)\|_{L^p[I_{r,h}]}, \quad 0 < \delta < |I|/r,$$

where

$$\Delta_h^r f(x) = \sum_{m=0}^r (-1)^{r+m} \binom{r}{m} f(x + mh),$$

$$v_{r,h}(x) := \min_{0 \leq m \leq r} v(x + mh), \quad I_{r,h} := \{x : x, x + rh \in I\}.$$

---

\* This research was supported by the Hungarian National Science Foundation under Grant No. T 4270.

We shall prove the equivalence between  $K$ -functionals (1) and moduli (2) for a class of weights  $\nu$  defined below. Some applications of this kind of moduli in the case  $r = 1$  were presented in [1], [2], [3], [4] and [5]. We will return to further applications in another paper.

## 2. The equivalent theorem

Denote by  $M^*(I) = M^*(a, b)$  the class of all weights  $\nu$  defined on the interval  $I = (a, b)$ , having the property that there exist two numbers  $(a <) A < B (< b)$  such that  $\nu$  is nondecreasing on  $(a, B)$  and nonincreasing on  $(A, b)$ . Let  $M(I)$  be the class of all weights  $u$  equivalent to some weight  $\nu \in M^*(I)$ , that is, there exists a weight  $\nu \in M^*(I)$  such that  $c_1 \nu(x) \leq u(x) \leq c_2 \nu(x)$  ( $x \in I$ ) with some constants  $c_1$  and  $c_2$  independent of  $x$ .

For example

$$\delta_{\alpha, \beta}(x) = (1 - \cos x)^\alpha (1 + \cos x)^\beta \in M(0, \pi) \quad (\alpha, \beta \geq 0),$$

$$w_{\beta, \gamma}(x) = x^\beta e^{-|x|^\gamma} \in M(0, \infty) \quad (\beta, \gamma \geq 0),$$

$$u_\lambda(x) = e^{-|x|^\lambda} \in M(-\infty, \infty) \quad (\lambda \geq 0).$$

The following theorem is true:

**THEOREM.** *Let  $I = (a, b)$  be a finite or infinite interval. Let  $\nu \in M(I)$ ,  $1 \leq p \leq \infty$ . Let  $K_r(f, t)_X$  and  $\omega_r(f, t)_X$  be the  $K$ -functional and the modulus defined by (1) and (2) resp. Then there exist constants  $c_1, c_2$  depending only on  $r, p, \nu$  and a constant  $c_3$  depending on  $\nu$  such that for any  $f \in X$ ,  $0 < t \leq c_3$  we have*

$$(3) \quad c_1 K_r(f, t)_X \leq \omega_r(f, t)_X \leq c_2 K_r(f, t)_X.$$

**PROOF OF THE THEOREM.** For simplicity we shall use the notations  $c(x, y, \dots)$  for constants depending only on  $x, y, \dots$ , not necessarily the same at each occurrence.

Note that if two weights are equivalent then the corresponding  $K$ -functionals (and moduli) are also equivalent. So, it is sufficient to prove the theorem for weights belonging to  $M^*(I)$ . This will be assumed throughout the proof.

A. Proof of the inequality

$$(4) \quad \omega_r(f, t)_X \leq c(r, p, \nu) K_r(f, t)_X.$$

The proof is based on the following three properties of the moduli. If  $f_1, f_2 \in X, g \in W^{(r)}(X)$ , then

$$(5) \quad \omega_r(f_1 + f_2, t)_X \leq \omega_r(f_1, t)_X + \omega_r(f_2, t)_X,$$

$$(6) \quad \omega_r(f_1, t)_X \leq 2^r \|f_1\|_X,$$

$$(7) \quad \omega_r(g, t)_X \leq t^r \|g^{(r)}\|_X.$$

Indeed, (5) and (6) follow immediately from the definition of the moduli. To prove (7) we notice that since  $v \in M^*(I)$  we have  $v \sim 1$  on every compact subinterval of  $(a, b)$ , or in other words,  $v \sim 1$  locally on  $(a, b)$ . Now for any  $g \in W^{(r)}(X)$

$$(8) \quad \Delta_h^r g(x) = \int_0^h \dots \int_0^h g^{(r)}(x + t_1 + \dots + t_r) dt_1 \dots dt_r.$$

On the other hand, since  $v \in M^*$  it is easy to see that

$$(9) \quad v_{r,h}(x) \leq cv(x+t) \quad (|t| \leq r|h| \leq c(v), \text{sgn } t = \text{sgn } h).$$

Therefore by (8) using Hölder inequality we get

$$\begin{aligned} & \|v_{r,h} \Delta_h^r g\|_{L^p[I_{r,h}]} \leq \\ & \leq \int_0^h \dots \int_0^h \|v(x+t_1+\dots+t_r) g^{(r)}(x+t_1+\dots+t_r)\|_{L^p[I_{r,h}]} dt_1 \dots dt_r \leq \\ & \leq \int_0^h \dots \int_0^h \|v g^{(r)}\|_{L^p[I]} dt_1 \dots dt_r = h^r \|v g^{(r)}\|_{L^p[I]}. \end{aligned}$$

Hence (7) follows.

Now from (5), (6), (7) we have for any  $f \in X, g \in W^{(r)}(X)$

$$\omega_r(f, t)_X \leq \omega_r(f - g, t)_X + \omega_r(g, t)_X \leq 2^r \|f - g\|_X + t^r \|g^{(r)}\|_X.$$

Hence by the definition of  $K_r$  we have (4).

**B. Proof of the converse inequality**

$$(10) \quad K_r(f, t)_X \leq c(r, p, v) \omega_r(f, t)_X.$$

We need the following further simple properties of the moduli.

a) If  $v$  is nondecreasing on  $(a, b)$  then for any  $f \in X$

$$(11) \quad \omega_r(f, h)_X = \sup_{0 < t \leq h} \|v \Delta_t^r f\|_{L^p[I_{r,t}]}.$$

b) In the case when  $v$  is nonincreasing we have

$$(12) \quad \omega_r(f, h)_X = \sup_{0 < t \leq h} \|v \Delta_{-t}^r f\|_{L^p[I_r, -t]}.$$

Indeed, in the Case a) we have

$$v_{r,h}(x) = \begin{cases} v(x) & \text{for } h > 0 \\ v(x + rh) & \text{for } h < 0. \end{cases}$$

Hence by the definition of the moduli we get

$$\begin{aligned} \omega_r(f, h)_X &= \sup_{|t| \leq h} \|v_{r,t} \Delta_t^r f\|_{L^p[I_r, t]} = \\ &= \max \left\{ \sup_{0 < t \leq h} \|v \Delta_t^r f\|_{L^p[I_r, t]}, \sup_{0 < t \leq h} \|v(x - rt) \Delta_{-t}^r f(x)\|_{L^p[I_r, -t]} \right\} =: \\ &=: \max\{\lambda_1, \lambda_2\}. \end{aligned}$$

From this, changing  $y = x - rt$  we get  $\lambda_1 = \lambda_2$  which proves (11). The proof of (12) is similar, therefore it is omitted.

Let now  $f \in X[I, p, v]$ ,  $|h| < |I|/r$ . Let

$$(13) \quad f_{[r,h]}(x) := \\ := h^{-r} \int_0^h \dots \int_0^h \sum_{s=0}^{r-1} (-1)^{r+s+1} \binom{r}{s} f\left(x + \frac{r-s}{r}(t_1 + \dots + t_r)\right) dt_1 \dots dt_r.$$

Since  $v \sim 1$  locally on  $I$ , the integral on the right hand side of (13) exists for every  $x \in I_{r,h}$ . The function  $f_{[r,h]}$  has the following properties. Let  $a < A < B < b$  be two numbers appearing in the definition of the class  $M^*$  applied to the weight  $v$ . Recall that then  $v$  is nondecreasing on  $(a, B)$ , nonincreasing on  $(A, b)$ . Choose two numbers  $A < A_1 < B_1 < B$  as  $A_1 = A + \frac{B-A}{3}$ ,  $B_1 = B - \frac{B-A}{3} = A + \frac{2(B-A)}{3}$ . Let  $I_1 := (a, B_1)$ ,  $I_2 := (A_1, b)$ . Then for  $0 < h < c(v)$  we have

$$(14) \quad \|f_{[r,h]} - f\|_{X[I_1, p, v]} \leq \omega_r(f, h)_{X[I, p, v]},$$

$$(15) \quad \|f_{[r,h]}^{(r)}\|_{X[I_1, p, v]} \leq c(r)h^{-r} \omega_r(f, h)_{X[I, p, v]},$$

$$(16) \quad \|f_{[r,-h]} - f\|_{X[I_2, p, v]} \leq \omega_r(f, h)_{X[I, p, v]},$$

$$(17) \quad \|f_{[r,-h]}^{(r)}\|_{X[I_2, p, v]} \leq c(r)h^{-r} \omega_r(f, h)_{X[I, p, v]}.$$

We prove only (14) and (15) since the proof of (16) and (17) is similar. By (11) and (13) using Hölder inequality we get

$$\begin{aligned} \|f_{[r,h]} - f\|_{X[I_1,p,v]} &\leq h^{-r} \int_0^h \dots \int_0^h \|\Delta_{(t_1+\dots+t_r)/r}^r f\|_{X[I_1,p,v]} dt_1 \dots dt_r \leq \\ &\leq \sup_{0 < t \leq h} \|\Delta_t^r f\|_{X[I_1,p,v]} \leq \omega_r(f, h)_{X[(a,B),p,v]} \leq \omega_r(f, h)_{X[I,p,v]}. \end{aligned}$$

To prove (15) we observe that for a.e.  $x \in I_{r,h}$

$$f_{[r,h]}^{(r)}(x) = h^{-r} \sum_{s=0}^{r-1} (-1)^{r+s+1} (rs) \Delta_{\frac{r-s}{r}h}^r f(x).$$

By this and (11)

$$\|f_{[r,h]}^{(r)}\|_{X[I_1,p,v]} \leq (2r)^r h^{-r} \omega_r(f, h)_{X[I,p,v]}.$$

We have (15).

Return to prove (10). With the numbers  $A_1$  and  $B_1$  defined above let  $\psi$  be the function defined as follows:  $\psi(x) = 1$  ( $a < x \leq A_1$ ),  $0$  ( $B_1 \leq x < b$ ) and  $\psi$  is  $r$ -times continuously differentiable on  $I$ ,  $|\psi^{(i)}(x)| \leq 1$  ( $i = 1, \dots, r$ ;  $A_1 \leq x \leq B_1$ ). Such a function was defined for example in [6, p. 141].

Now for  $f \in X[I,p,v]$ ,  $0 < h < c(v)$  introduce

$$(18) \quad f_{(r,h)} = \psi f_{[r,h]} + (1 - \psi) f_{[r,-h]} = f_{[r,-h]} + \psi (f_{[r,h]} - f_{[r,-h]}).$$

By (14), (16) and (18) we have

$$(19) \quad \|f - f_{(r,h)}\|_{X[I,p,v]} \leq 2\omega_r(f, h)_{X[I,p,v]}.$$

On the other hand, by (18) and the definition of  $\psi$  we get

$$(20) \quad \begin{aligned} &\|f_{(r,h)}^{(r)}\|_{X[(A_1,B_1),p,v]} \leq \\ &\leq c(r) \left\{ \|f_{[r,-h]}^{(r)}\|_{X[(A_1,B_1),p,v]} + \sum_{i=0}^r \|f_{[r,h]} - f_{[r,-h]}\|_{X[(A_1,B_1),p,v]}^{(i)} \right\}. \end{aligned}$$

Since  $v \in M^*(I)$  we have  $\|\cdot\|_{X[(A_1,B_1),p,v]} \sim \|\cdot\|_{L^p(A_1,B_1)}$ . Therefore from (20) using a well-known inequality between  $L^p$ -norm of derivatives of different orders we get (see e.g. [6, p. 18])

$$\begin{aligned} \|f_{(r,h)}^{(r)}\|_{X[(A_1,B_1),p,v]} &\leq c(r,p) \left\{ \|f_{[r,h]}^{(r)}\|_{X[(A_1,B_1),p,v]} + \right. \\ &\quad \left. + \|f_{[r,-h]}^{(r)}\|_{X[(A_1,B_1),p,v]} + \|f_{[r,h]} - f_{[r,-h]}\|_{X[(A_1,B_1),p,v]} \right\}. \end{aligned}$$

Hence by (14), (15), (16), (17) we have

$$(21) \quad \|f_{\langle r, h \rangle}^{(r)}\|_{X[I, p, v]} \leq c(p, r) h^{-r} \omega_r(f, h)_{X[I, p, v]}.$$

Combining (19) and (21) and noticing that  $f_{\langle r, h \rangle} \in W^{(r)}(X)$  we have (10). By this the proof of our theorem is complete.

### References

- [1] I. JOÓ and N. X. KY, Answer to a problem of Paul Turán, *Annales Univ. Sci. Budapest.*, Sect. Math., **31** (1988), 229–241.
- [2] N. X. KY, Jackson and Bernstein type approximation theorems in the case of approximation by polynomials in  $L^p$ -space, *Studia Sci. Math. Hung.*, **9** (1974), 405–415.
- [3] N. X. KY, On imbedding theorems for weighted polynomial approximation and modulus of continuity of functions, *Acta Sci. Math.*, **54** (1990), 313–325.
- [4] N. X. KY, On approximation by polynomials with weights, *Acta Math. Acad. Sci. Hung.*, **59** (1992), 49–58.
- [5] N. X. KY and N. Q. HOA, Uljanov-type imbedding theorems in the theory of weighted polynomial approximation, *Studia Sci. Math. Hung.*, **29** (1994), 33–42.
- [6] L. L. SHUMAKER, *Spline functions*, New York–Chichester–Brisbane–Toronto–Singapore, 1981.

## COMPRESSED AND CAUCHY FILTERS

By

B. BALAGUER

Département de Mathématiques, Université de Perpignan, Perpignan, France

(Received January 3, 1995)

### 1. Introduction

Classically compactness and completeness have similar characterizations, one by convergence of compressed filters the other by convergence of Cauchy filters. Moreover every Cauchy filter is compressed for the proximity associated with the given uniformity; the converse is true if and only if the uniformity is precompact (see [3]).

In this paper we give a definition of compressed filters relative to a quasi-proximity and we show that compactness of the induced topology is characterized by convergence of the compressed filters (see Definition 2 and Theorem 1).

We introduce a notion of Cauchy filter for a quasi-uniformity (see Definition 3) that satisfies the following properties:

- (i) coincidence with the classical definition in the uniform case,
- (ii) every convergent filter is a Cauchy filter,
- (iii) every Cauchy filter is compressed for the associated quasi-proximity,
- (iv) the converse of (iii) is true provided the quasi-uniformity is totally bounded.

These conditions lead to the well-known notion of *D-Cauchy* filter (see [6]).

However the coincidence of Cauchy filters and compressed filters does not characterize totally bounded quasi-uniformities, but an intermediate class between the classes of precompact quasi-uniformities and totally bounded quasi-uniformities (see Theorem 2). This class is that of *Cauchy bounded* quasi-uniformities, and at last we give another proof of KOPPERMAN's result [10]: the topology induced by a quasi-uniformity is compact if and only if the quasi-uniformity is *D-complete* and *Cauchy bounded* (see Theorem 3).

## 2. Preliminaries

If  $\delta$  is a binary relation on the power set  $\mathcal{P}(X)$  of  $X$ , and  $\mathcal{A}$  a subset of  $\mathcal{P}(X)$  let us set

$$\mathcal{A}^\delta = \cap \{ \delta(A) : A \in \mathcal{A} \}$$

where  $\delta(A) := \{ B \in \mathcal{P}(X) : A \delta B \}$ .

The application  $\mathcal{A} \rightarrow \mathcal{A}^\delta$  is a *polarity* (i.e turns every union into intersection) and so is decreasing; moreover for all subsets  $\mathcal{A}$  and  $\mathcal{B}$  of  $\mathcal{P}(X)$  we have:

$$\mathcal{A} \times \mathcal{B} \subseteq \delta \Leftrightarrow \mathcal{A} \subseteq \mathcal{B}^{\delta^{-1}} \iff \mathcal{B} \subseteq \mathcal{A}^\delta.$$

In particular, for the relations  $\#$  and  $*$ , respectively defined by  $A \cap B \neq \emptyset$  and  $A \neq X \setminus B$  we obtain:

$$\begin{aligned} \mathcal{A}^\# &= \{ B \in \mathcal{P}(X) : B \# A \text{ for every } A \in \mathcal{A} \} \quad (\text{grill of } \mathcal{A}), \\ \mathcal{A}^* &= \{ B \in \mathcal{P}(X) : X \setminus B \notin \mathcal{A} \} \quad (\text{conjugate of } \mathcal{A}). \end{aligned}$$

Let us note that  $\mathcal{A} \rightarrow \mathcal{A}^*$  is an involutive polarity and so turns every intersection into union.

Following [11] we say that a subset  $\mathcal{A}$  of  $\mathcal{P}(X)$  is a *stack* on  $X$  if every set that contains an element of  $\mathcal{A}$  is element of  $\mathcal{A}$ . The grill of  $\mathcal{A}$  is a stack and stacks are exactly those sets for which grill and conjugate coincide; they are also those for which  $\mathcal{A} = (\mathcal{A}^\#)^\#$ .

A *grill* (of a filter) on  $X$  is a proper stack  $\mathcal{G}$  on  $X$  that satisfies:

$$A \cup B \in \mathcal{G} \Rightarrow A \in \mathcal{G} \text{ or } B \in \mathcal{G}.$$

Equivalently  $\mathcal{G}$  is a stack and  $\mathcal{G}^\#$  is a filter.

Recall that *ultrafilters* on  $X$  are filters which coincide with their grill. The ultrafilter generated by  $\{x\}$  will be denoted by  $[x]$ .

## 3. Basic quasi-proximities on $X$ and filters on $X \times X$

We refer the reader to [8] for the basic definitions concerning quasi-uniformities and quasi-proximities.

According to the terminology of [11] a *basic quasi-proximity* on  $X$  is a binary relation  $\delta$  on  $\mathcal{P}(X)$  satisfying:  $(P_1)$   $A \delta B$  implies  $A$  and  $B$  non empty;  $(P_2)$   $A \delta B$ ,  $A \subseteq A'$ ,  $B \subseteq B'$  imply  $A' \delta B'$ ;  $(P_3)$   $A \# B$  implies  $A \delta B$ ;  $(P_4)$   $(A_1 \cup A_2) \delta (B_1 \cup B_2)$  implies  $A_i \delta B_j$  for some  $(i, j)$ . If  $\delta$  fulfils only  $(P_1)$ ,  $(P_2)$  and  $(P_4)$  we say that  $\delta$  is a *weak quasi-proximity*.

The conjugate relation of  $\delta$ , denoted  $\delta^*$ , is defined by:

$$A \delta^* B \text{ iff } B \in \delta(A)^* \text{ (i.e. } \bar{\delta}(X \setminus B) \text{ where } \bar{\delta} := \text{non } \delta).$$

Recall that the closure operator associated with the topology  $\tau(\delta)$  induced by a quasi-proximity  $\delta$  is defined by:

$$x \in \text{cl}(A) \text{ iff } \{x\} \delta A.$$

For this topology, the neighbourhood filter of  $x$  is  $\delta(x)^\#$  (since this filter is  $\delta^*(x) = \delta(x)^*$  and  $\delta(x)$  is a stack).

DEFINITION 1. For every subset  $\mathcal{R}$  of  $\mathcal{P}(X \times X)$  we denote by  $\delta(\mathcal{R})$  the binary relation on  $\mathcal{P}(X)$  defined by

$$A \delta(\mathcal{R}) B \text{ iff } (A \times B) \# R \text{ for each } R \in \mathcal{R}.$$

Clearly  $\delta(\mathcal{R})$  is a weak quasi-proximity provided  $\mathcal{R}$  is a filter, and turns into a basic quasi-proximity if and only if each element of the filter  $\mathcal{R}$  contains the diagonal of  $X \times X$ . Moreover if  $\mathcal{R}$  is a (quasi)-uniformity then  $\delta(\mathcal{R})$  is a (quasi)-proximity.

A filter  $\mathcal{R}$  on  $X \times X$  is said to be *compatible* with a weak quasi-proximity  $\delta$  if  $\delta(\mathcal{R}) = \delta$ .

Let us denote by

$$H(A, B) \text{ the relation } (X \times X) \setminus (A \times (X \setminus B)),$$

$\mathcal{R}(\delta)$  the stack generated by the finite intersections of the members of the set  $\{H(A, B) : A \delta^* B\}$ .

THEOREM ([2, (19.38)–(19.40)] and [9, Theorem 1]).

(i) *If  $\mathcal{R}$  is a filter on  $X \times X$  and  $\delta = \delta(\mathcal{R})$ , then  $\mathcal{R}(\delta)$  is the coarsest filter compatible with  $\delta$ .*

(ii) *Moreover, if  $\delta$  is a quasi-proximity, then  $\mathcal{R}(\delta)$  is the only totally bounded quasi-uniformity compatible with  $\delta$ .*

REMARKS 1. a) If  $\delta$  is a weak quasi-proximity, then  $\mathcal{R}(\delta)$  is not necessarily a filter, because a finite intersection of  $H(A_i, B_i)$  may be empty. This is not the case if  $\delta = \delta(\mathcal{R})$ , where  $\mathcal{R}$  is a filter, because any finite intersection of  $H(A_i, B_i)$  contains a finite intersection of elements of  $\mathcal{R}$ .

b) Although the first part of this theorem is not explicitly given in [2] or [9] it is implicit in the proofs given there.

c) For a basic quasi-proximity the filter  $\mathcal{R}(\delta)$  is totally bounded, but it is not the only compatible filter of this type; nevertheless (ii) can be extended to basic quasi-proximity by introducing the notion of simple filter (see [1]).

COROLLARY 1. Let  $\mathcal{R}_1$  be a filter on  $X \times X$  and  $\mathcal{R}_2$  a totally bounded quasi-uniformity on  $X$ . Then  $\delta(\mathcal{R}_1) \subseteq \delta(\mathcal{R}_2)$  if and only if  $\mathcal{R}_2 \subseteq \mathcal{R}_1$ .

PROOF. Let  $\delta_1 = \delta(\mathcal{R}_1)$  and  $\delta_2 = \delta(\mathcal{R}_2)$ . If  $\delta_1 \subseteq \delta_2$  then  $\delta_2^* \subseteq \delta_1^*$  and therefore  $\mathcal{R}(\delta_2) \subseteq \mathcal{R}(\delta_1)$ . Consequently we have by the preceding theorem:  $\mathcal{R}_2 = \mathcal{R}(\delta_2) \subseteq \mathcal{R}(\delta_1) \subseteq \mathcal{R}_1$ . The converse follows from the fact that the application  $\mathcal{R} \rightarrow \delta(\mathcal{R})$  is decreasing.

#### 4. Compressed and Cauchy filters

If  $\mathcal{A}$  and  $\mathcal{B}$  are subsets of  $\mathcal{P}(X)$ , we denote by  $\mathcal{A} \otimes \mathcal{B}$  the stack generated by  $\{A \times B : A \in \mathcal{A}, B \in \mathcal{B}\}$ .

Let us recall that a filter  $\mathcal{F}$  on  $X$  is called a *Cauchy filter* for a uniformity  $\mathcal{R}$  if  $\mathcal{R} \subseteq \mathcal{F} \otimes \mathcal{F}$ . If  $\delta$  is a proximity, a filter  $\mathcal{F}$  on  $X$  is said to be compressed if  $\mathcal{F}^\# \times \mathcal{F}^\# \subseteq \delta$ .

For the main results concerning these notions we refer the reader to [3].

##### 4.1 Compressed filters for a quasi-proximity

The standard definition of a compressed filter cannot be used in the case of quasi-proximities since a convergent filter need not be compressed: for example, the euclidian topology of  $\mathbb{R}$  is induced by the quasi-proximity  $\delta$  defined by  $A\#\text{cl}(B)$  but for the neighbourhood filter  $\delta(x)^\#$  we don't have  $\delta(x) \times \delta(x) \subseteq \delta$ .

Let us now introduce the following definition.

DEFINITION 2. Let  $\delta$  be a quasi-proximity on  $X$ . A filter  $\mathcal{F}$  on  $X$  is compressed for  $\delta$  if there exists an ultrafilter  $\mathcal{U}$  on  $X$  such that  $\mathcal{U} \times \mathcal{F}^\# \subseteq \delta$ .

LEMMA 1. Let  $\delta$  be a quasi-proximity on  $X$ ,  $\mathcal{U}$  an ultrafilter on  $X$  and  $\mathcal{A}, \mathcal{B}$  subsets of  $\mathcal{P}(X)$ . If  $\mathcal{A} \times \mathcal{U} \subseteq \delta$  and  $\mathcal{U} \times \mathcal{B} \subseteq \delta$ , then  $\mathcal{A} \times \mathcal{B} \subseteq \delta$ .

PROOF. Suppose that  $\mathcal{A} \times \mathcal{B}$  is not included in  $\delta$ . Hence there exists  $(A, B) \in \mathcal{A} \times \mathcal{B}$  such that  $A \bar{\delta} B$ , and consequently there exists a subset  $C$  such that  $A \bar{\delta} C$  and  $(X \setminus C) \bar{\delta} B$ . It follows that neither  $C$  nor  $X \setminus C$  belongs to the ultrafilter  $\mathcal{U}$  — a contradiction.

PROPOSITION 1. If  $\delta$  is a proximity, then the preceding definition is equivalent to the usual definition.

Moreover we have

**THEOREM 1.** *Let  $\delta$  be a quasi-proximity on  $X$ . Then (i) every convergent filter is compressed and (ii) the space  $(X, \tau(\delta))$  is compact if and only if every compressed filter is convergent.*

**PROOF.** (i) If  $\mathcal{F}$  converges to  $x$  then  $\delta(x)^\# \subseteq \mathcal{F}$  or equivalently  $\mathcal{F}^\# \subseteq \delta(x)$ . Clearly  $\delta(x) = [x]^\delta$ , so the convergence of  $\mathcal{F}$  to  $x$  implies  $[x] \times \mathcal{F}^\# \subseteq \delta$  i.e.  $\mathcal{F}$  is compressed. (ii) Suppose  $(X, \tau(\delta))$  is compact and let  $\mathcal{F}$  be a compressed filter. Hence there exists an ultrafilter  $\mathcal{U}$  and  $x \in X$  such that:  $\mathcal{U} \times \mathcal{F}^\# \subseteq \delta$  and  $\mathcal{U} \subseteq \delta(x)$ . As in the preceding discussion  $\mathcal{U} \subseteq \delta(x)$  is equivalent to  $[x] \times \mathcal{U} \subseteq \delta$ . So we have  $\mathcal{U} \times \mathcal{F}^\# \subseteq \delta$  and  $[x] \times \mathcal{U} \subseteq \delta$ ; by lemma 1 we obtain  $[x] \times \mathcal{F}^\# \subseteq \delta$  or equivalently  $\delta(x)^\# \subseteq \mathcal{F}$ . Hence  $\mathcal{F}$  converges to  $x$ . The converse follows from the fact that every ultrafilter is compressed.

### 4.2 Cauchy filters for a quasi-uniformity

We now introduce a notion of Cauchy filter for a quasi-uniformity which satisfies conditions (i)–(iv) from the introduction. Let us first look at condition (iv); our definition will be motivated by the following lemma.

**LEMMA 2.** (i) *If  $\mathcal{A}$  and  $\mathcal{B}$  are stacks on  $X$  then we have*

$$\mathcal{A} \times \mathcal{B} = \delta(\mathcal{A}^\# \otimes \mathcal{B}^\#).$$

(ii) *If  $\mathcal{F}_1$  and  $\mathcal{F}_2$  are filters on  $X$  and  $\mathcal{R}$  a totally bounded quasi-uniformity on  $X$  then we have*

$$\mathcal{F}_1^\# \times \mathcal{F}_2^\# \subseteq \delta(\mathcal{R}) \quad \text{iff} \quad \mathcal{R} \subseteq \mathcal{F}_1 \otimes \mathcal{F}_2.$$

**PROOF.** (i) By Definition 1 we have  $(C, D) \in \delta(\mathcal{A}^\# \otimes \mathcal{B}^\#)$  iff  $(C \times D)^\#(E \times F)$  for each  $E \in \mathcal{A}^\#$  and each  $F \in \mathcal{B}^\#$ , or equivalently  $(C, D) \in (\mathcal{A}^\#)^\# \times (\mathcal{B}^\#)^\#$  i.e.  $(C, D) \in \mathcal{A} \times \mathcal{B}$ . (ii) By the preceding result  $\mathcal{F}_1^\# \times \mathcal{F}_2^\# \subseteq \delta(\mathcal{R})$  means  $\delta(\mathcal{F}_1 \otimes \mathcal{F}_2) \subseteq \delta(\mathcal{R})$  and by corollary 1,  $\mathcal{R} \subseteq \mathcal{F}_1 \otimes \mathcal{F}_2$ .

Consequently, if  $\mathcal{R}$  is totally bounded, a filter  $\mathcal{F}$  is compressed for  $\delta(\mathcal{R})$  if and only if there exists an ultrafilter  $\mathcal{U}$  such that  $\mathcal{R} \subseteq \mathcal{U} \otimes \mathcal{F}$ . We now state the following definition.

**DEFINITION 3.** Let  $\mathcal{R}$  be a quasi-uniformity on  $X$ . A filter  $\mathcal{F}$  on  $X$  is called a Cauchy filter if there exists an ultrafilter  $\mathcal{U}$  on  $X$  such that  $\mathcal{R} \subseteq \mathcal{U} \otimes \mathcal{F}$ . The space  $(X, \mathcal{R})$  is complete if every Cauchy filter converges.

**REMARK 2.** This definition is equivalent to the definition introduced by Doitchinov to resolve the problem of completion for quiet quasi-uniform spaces (see [6]).

PROPOSITION 2. *Let  $\mathcal{R}$  be a quasi-uniformity on  $X$ .*

- (i) *If  $\mathcal{R}$  is a uniformity, then the preceding definition of Cauchy filters is equivalent to the usual definition.*
- (ii) *Every convergent filter is a Cauchy filter.*
- (iii) *Every Cauchy filter is compressed for  $\delta(\mathcal{R})$ .*
- (iv) *If  $\mathcal{R}$  is totally bounded, then every compressed filter for  $\delta(\mathcal{R})$  is a Cauchy filter.*

PROOF. As a consequence of Lemma 2 we have already obtained (iv). We let the reader verify conditions (i) and (ii) (a proof in term of sequences is given in [5]). Let us prove (iii): if  $\mathcal{R} \subseteq \mathcal{U} \otimes \mathcal{F}$  then  $\delta(\mathcal{U} \otimes \mathcal{F}) \subseteq \delta(\mathcal{R})$  and by the first point of Lemma 2 we have  $\mathcal{U} \times \mathcal{F}^\# \subseteq \delta(\mathcal{R})$ .

However the coincidence between Cauchy filters for  $\mathcal{R}$  and compressed filters for  $\delta(\mathcal{R})$  does not characterize totally bounded quasi-uniformities as we show in Theorem 2.

LEMMA 3. *Let  $\mathcal{R}$  be a quasi-uniformity and  $\delta$  the associated quasi-proximity.*

- (i) *If  $\mathcal{F}$  is a filter on  $X$ , then  $\mathcal{F}^\delta$  is a grill on  $X$ .*
- (ii) *A filter  $\mathcal{F}$  on  $X$  is a Cauchy filter if and only if  $(\mathcal{F}^\delta)^\#$  is a Cauchy filter (see also [4, Lemma 1.2]).*

PROOF. (i) We omit this straightforward proof (see [11, Theorem 3.10] where it was carried out for  $\mathcal{F}$  an ultrafilter). (ii) By the preceding point  $(\mathcal{F}^\delta)^\#$  is a filter on  $X$ . Moreover axiom  $(P_3)$  of quasi-proximities implies that  $\mathcal{F}^\# \subseteq \mathcal{F}^\delta$ , or equivalently  $(\mathcal{F}^\delta)^\# \subseteq \mathcal{F}$ . Consequently if  $(\mathcal{F}^\delta)^\#$  is a Cauchy filter so is  $\mathcal{F}$ . For the converse suppose that  $\mathcal{F}$  is a Cauchy filter, then there exists an ultrafilter  $\mathcal{U}$  such that  $\mathcal{R} \subseteq \mathcal{U} \otimes \mathcal{F}$ . We are going to show that  $\mathcal{R} \subseteq (\mathcal{U}^{\delta^{-1}})^\# \otimes (\mathcal{F}^\delta)^\#$ , and this will imply that  $(\mathcal{F}^\delta)^\#$  is a Cauchy filter. For this let us note that  $\mathcal{F}^\delta$  is a stack so  $(\mathcal{F}^\delta)^\# = (\mathcal{F}^\delta)^*$  and  $(\mathcal{F}^\delta)^* = \cup\{\delta^*(A) : A \in \mathcal{F}\}$ . Thus  $B \in (\mathcal{F}^\delta)^\#$  if there exists  $A \in \mathcal{F}$  such that  $A \bar{\delta}(X \setminus B)$ , or equivalently if there exists  $A \in \mathcal{F}$  and  $S \in \mathcal{R}$  such that  $S(A) \subseteq B$ ; in particular  $S(A) \in (\mathcal{F}^\delta)^\#$ . Then if  $R \in \mathcal{R}$  let us choose  $S \in \mathcal{R}$  such that  $S^3 \subseteq R$ ; by  $\mathcal{R} \subseteq \mathcal{U} \otimes \mathcal{F}$  there exists  $(A, B) \in \mathcal{U} \times \mathcal{F}$  such that  $A \times B \subseteq S$ . So we have  $S^{-1}(A) \times S(B) \subseteq S^3 \subseteq R$  with  $S^{-1}(A) \in (\mathcal{U}^{\delta^{-1}})^\#$  and  $S(B) \in (\mathcal{F}^\delta)^\#$ . Hence  $\mathcal{R} \subseteq (\mathcal{U}^{\delta^{-1}})^\# \otimes (\mathcal{F}^\delta)^\#$ .

Let us recall the following well known results [8, theorem 3–14]. A quasi-uniformity  $\mathcal{R}$  is (a) totally bounded if and only if every ultrafilter is a Cauchy filter for the uniformity generated by  $\mathcal{R} \cup \mathcal{R}^{-1}$ , (b) precompact if

and only if every ultrafilter is Cauchy in the sense of Pervin (i.e  $\mathcal{U}$  satisfies: for each  $R \in \mathcal{R}$  there exists  $x \in X$  such that  $R(x) \in \mathcal{U}$ ).

DEFINITION 4. A quasi-uniformity is Cauchy bounded if each ultrafilter is a Cauchy filter.

In [10], R. D. KOPPERMAN introduced the above definition, and showed that a quasi-uniform space  $(X, \mathcal{R})$  is compact iff  $\mathcal{R}$  is complete and Cauchy bounded. We give a characterization of Cauchy bounded quasi-uniformities in the following.

THEOREM 2. (i) *Given a quasi-uniformity  $\mathcal{R}$ , Cauchy filters for  $\mathcal{R}$  and compressed filters for  $\delta(\mathcal{R})$  coincide if and only if  $\mathcal{R}$  is Cauchy bounded.*

(ii) *Any totally bounded quasi-uniformity is Cauchy bounded and any Cauchy bounded one is precompact.*

PROOF. (i) If we have coincidence between Cauchy and compressed filters then every ultrafilter is a Cauchy filter, since it is clearly compressed for  $\delta(\mathcal{R})$ . For the converse suppose that every ultrafilter is a Cauchy filter. Let  $\mathcal{F}$  be a compressed filter for  $\delta(\mathcal{R})$ ; then there exists an ultrafilter  $\mathcal{U}$  such that  $\mathcal{U} \times \mathcal{F}^\# \subseteq \delta(\mathcal{R})$ , or equivalently such that  $(\mathcal{U}^\delta)^\# \subseteq \mathcal{F}$  where  $\delta = \delta(\mathcal{R})$ . By Lemma 3,  $(\mathcal{U}^\delta)^\#$  is a Cauchy filter and so is  $\mathcal{F}$ . (ii) By Proposition 2 and the preceding point if  $\mathcal{R}$  is totally bounded then  $\mathcal{R}$  is Cauchy bounded. Moreover every Cauchy filter is Cauchy in the sense of Pervin since  $A \times B \subseteq R$  implies  $B \subseteq R(x)$  for each  $x \in A$ ; so if  $\mathcal{R}$  is Cauchy bounded every ultrafilter is Cauchy in the sense of Pervin and  $\mathcal{R}$  is precompact.

As a corollary of Theorem 1 we give another proof of the fundamental result due to KOPPERMAN ([10, Theorem 6]).

THEOREM 3. *Let  $\mathcal{R}$  be a quasi-uniformity on  $X$ . Then the space  $(X, \tau(\mathcal{R}))$  is compact if and only if  $\mathcal{R}$  is complete and Cauchy bounded.*

PROOF. If  $(X, \tau(\mathcal{R}))$  is compact then by Theorem 1 every compressed filter for  $\delta(\mathcal{R})$  converges. But every Cauchy filter is compressed, so  $(X, \mathcal{R})$  is complete. Moreover by compactness every ultrafilter is convergent and so is a Cauchy filter; hence  $\mathcal{R}$  is Cauchy bounded. The converse is immediate.

I would like to thank the referee for his valuable suggestions.

## References

- [1] B. BALAGUER, Approximation relationnelle de quasi-proximités et quasi-complétion des quasi-uniformités totalement bornées, *Thèse, Université de Perpignan* (1992).
- [2] Á. CSÁSZÁR, Foundations of General Topology, *Press, Oxford*, (1963).
- [3] Á. CSÁSZÁR, General Topology, *Adam Hilger Ltd, Bristol*, (1978).
- [4] Á. CSÁSZÁR, D-complete extensions of quasi-uniform spaces, *Acta Math. Hungar.*, **64** (1994), 41–54.
- [5] D. DOITCHINOV, On completeness in quasi-metric spaces, *Topology Appl.*, **30** (1988), 127–148.
- [6] D. DOITCHINOV, On completeness of quasi-uniform spaces, *C. R. Acad. Bulgare Sci.*, **41** (1988), 5–8.
- [7] D. DOITCHINOV, A concept of completeness of quasi-uniform spaces, *Topology Appl.*, **38** (1991), 205–217.
- [8] P. FLETCHER and W. F. LINDGREN, Quasi-uniform spaces, *Lecture Notes Pure Appl. Math.*, vol.77 (Marcel Dekker, 1982).
- [9] W. HUNSAKER and W. F. LINDGREN, Construction of quasi-uniformities, *Math. Ann.*, **188** (1970), 39–42.
- [10] R. D. KOPPERMAN, Total boundedness and compactness for filter pairs, *Annales Univ. Sci. Budapest.*, Sectio Mathematica, **33** (1990), 25–30.
- [11] W. J. THRON, Proximity structures and grills, *Math. Ann.*, **206** (1973), 35–62.

## AN ALEXITS TYPE THEOREM

By

I. JOÓ\*

Department of Analysis, Eötvös Loránd University, Budapest

(Received February 17, 1995)

The classical Alexits theorem [1], [2] states that a function  $f \in L_p$  can be approximated by its Fejér means  $\sigma_n f$  in the order  $\|\sigma_n f - f\|_p = O\left(\frac{1}{n}\right)$  if and only if the derivative of its conjugate function  $\tilde{f}'$  belongs to  $L_p$ . In what follows we extend this result for Fourier integrals instead of Fourier series. Denote  $\hat{f}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ixt} dx$  the Fourier transform of  $f \in L_1(\mathbb{R})$  further let

$$D_\lambda(x) = \frac{1}{\pi} \frac{\sin \lambda x}{x}$$

be the Dirichlet kernel. It belongs to  $L_q(\mathbb{R})$  for  $1 < q \leq \infty$ , so for  $f \in L_p(\mathbb{R})$ ,  $1 \leq p < \infty$  we can define the  $\lambda$ -th partial sum of the Fourier integral expansion of  $f$  as

$$S_\lambda(f, x) = \int_{-\infty}^{\infty} f(y) D_\lambda(x - y) dy = f * D_\lambda(x).$$

If  $f \in L_1(\mathbb{R})$  then by the Fubini theorem we obtain

$$(1) \quad S_\lambda(f, x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(y) \left( \int_{-\lambda}^{\lambda} e^{i(x-y)t} dt \right) dy = \frac{1}{\sqrt{2\pi}} \int_{-\lambda}^{\lambda} e^{ixt} \hat{f}(t) dt.$$

---

\* Supported by the National Scientific Research Found. OTKA N° T 024244

The Fejér kernel is defined by

$$(2) \quad K_\lambda(x) = \frac{1}{\lambda} \int_0^\lambda D_t(x) dt = \frac{1}{\lambda} \int_0^\lambda \frac{\sin tx}{\pi x} dt = \\ = \frac{1}{\lambda} \frac{1 - \cos \lambda x}{\pi x^2} = \frac{2}{\pi} \frac{\left(\sin \frac{\lambda x}{2}\right)^2}{\lambda x^2}.$$

It belongs to all  $L_q$ ,  $1 \leq q \leq \infty$ , hence for  $f \in L_p$ ,  $1 \leq p \leq \infty$  we can define its  $\lambda$ -th Fejér means

$$\sigma_\lambda(f, x) = \int_{-\infty}^{\infty} f(y) K_\lambda(x - y) dy = f * K_\lambda(x).$$

If  $f \in L_p$ ,  $1 \leq p \leq \infty$ , then again by Fubini theorem

$$\sigma_\lambda(f, x) = \frac{1}{\lambda} \int_0^\lambda S_t(f, x) dt$$

and if  $f \in L_1$  then

$$(3) \quad \sigma_\lambda(f, x) = \frac{1}{\sqrt{2\pi}} \int_{-\lambda}^\lambda \left(1 - \frac{|t|}{\lambda}\right) e^{ixt} \hat{f}(t) dt.$$

Further define the conjugate Dirichlet kernel by

$$\tilde{D}_\lambda(x) = \frac{i}{\pi} \frac{1 - \cos \lambda x}{x}$$

and for  $f \in L_p$ ,  $1 \leq p < \infty$  let

$$\tilde{S}_\lambda(f, x) = \int_{-\infty}^{\infty} f(y) \tilde{D}_\lambda(x - y) dy = f * \tilde{D}_\lambda(x)$$

be the  $\lambda$ -th conjugate partial sum of  $f$ . Let further

$$\tilde{K}_\lambda(x) = \frac{1}{\lambda} \int_0^\lambda \tilde{D}_t(x) dt = \frac{i}{\pi} \frac{\lambda x - \sin \lambda x}{\lambda x^2}$$

be the conjugate Fejér kernel and for  $f \in L_p$ ,  $1 \leq p \leq \infty$  let  $\tilde{\sigma}_\lambda(f, x) = f * \tilde{K}_\lambda(x)$ .

If there exists  $\tilde{f} \in L_p$  such that  $\tilde{\sigma}_\lambda f \xrightarrow{L_p} \tilde{f}$  then  $\tilde{f}$  is called the conjugate function

of  $f$ . We shall see that if  $1 < p < \infty$ , every  $f \in L_p$  has a conjugate  $\tilde{f} \in L_p$  which can be characterized by the equality

$$(4) \quad \hat{\tilde{f}}(t) = \frac{|t|}{t} \hat{f}(t)$$

meant in (Schwartz-) distributional sense, in  $S'(\mathbb{R})$ . Our main aim is to prove the following version of the Alexits theorem

**THEOREM.** *Let  $1 < p < \infty$  then  $\|\sigma_\lambda f - f\| = O\left(\frac{1}{\lambda}\right) \Leftrightarrow \tilde{f}$  is locally absolutely continuous and  $f' \in L_p$ .*

■

Analogous result is obtained in D.V. GIANG [9] for  $p = 2$ .

The proof needs some auxiliary results. We begin with the widely known

**PROPOSITION 1.**

- a)  $\|\sigma_\lambda f\|_p \leq c \|f\|_p$  and  $f \in L_p(\mathbb{R})$ ,  $1 \leq p \leq \infty$ ,  $\lambda > 0$ ,
- b)  $\|\sigma_\lambda f - f\|_p \rightarrow 0$  ( $\lambda \rightarrow \infty$ ) for  $f \in L_p(\mathbb{R})$ ,  $1 \leq p < \infty$ .

For the convenience of the reader we provide a proof.

**PROOF.**

a)  $\|\sigma_\lambda f\|_p = \|f * K_\lambda\|_p \leq \|K_\lambda\|_1 \|f\|_p$  hence we have to check  $\|K_\lambda\|_1 \leq c$ . But this is easy to see: if  $|x| < \frac{1}{\lambda}$  then  $|K_\lambda(x)| \leq \frac{1}{2\pi} \lambda$  hence

$$\int_{|x| < \frac{1}{\lambda}} |K_\lambda(x)| dx \leq c$$

and analogously

$$\int_{|x| > \frac{1}{\lambda}} |K_\lambda(x)| dx \leq \int_{|x| > \frac{1}{\lambda}} \frac{c}{\lambda x^2} dx \leq c.$$

b) It is enough to find a complete set of  $f$  in  $L_p$  for which b) holds because b) follows then from a) by the Banach-Steinhaus theorem. This complete set consists of the orthonormal Hermite functions

$$\varphi_n(x) = h_n(x) e^{-\frac{x^2}{2}}$$

where  $h_n$  is the normed Hermite polynomials. The functions  $\varphi_n(x)$  form an orthonormal basis in  $L_2(\mathbb{R})$ , are complete in  $L_p(\mathbb{R})$ ,  $1 \leq p < \infty$  and are the eigenfunctions of the Fourier transformation:

$$\hat{\varphi}_n = (-i)^n \varphi_n.$$

We have then

$$\begin{aligned} S_\lambda(\varphi_n, x) - \varphi_n(x) &= \frac{1}{\sqrt{2\pi}} \int_{|t|>\lambda} e^{ixt} \hat{\varphi}_n(t) dt = \frac{(-i)^n}{\sqrt{2\pi}} \int_{|t|>\lambda} e^{ixt} \varphi_n(t) dt = \\ &= \frac{(-i)^n}{\sqrt{2\pi}} \left[ \frac{e^{ixt}}{ix} \varphi_n(t) \right]_{t=-\lambda}^{\lambda} - \frac{(-i)^n}{\sqrt{2\pi}} \int_{|t|>\lambda} \frac{e^{ixt}}{ix} \varphi_n'(t) dt = \\ &= \frac{(-i)^n}{\sqrt{2\pi}} \frac{i}{x} \left( e^{i\lambda x} \varphi_n(\lambda) - e^{-i\lambda x} \varphi_n(-\lambda) \right) + O\left( e^{-\frac{\lambda^2}{4}} |x|^{-2} \right). \end{aligned}$$

Another integration by parts gives that

$$\sigma_\lambda(\varphi_n, x) - \varphi_n(x) = \frac{1}{\lambda} \int_0^\lambda (S_t(\varphi_n, x) - \varphi_n(x)) dt = O\left( e^{-\frac{\lambda^2}{4}} |x|^{-2} \right).$$

On the other hand

$$|S_\lambda(\varphi_n, x) - \varphi_n(x)| \leq \frac{1}{\sqrt{2\pi}} \int_{|t|>\lambda} |\hat{\varphi}_n(t)| dt \leq c_n e^{-\frac{\lambda^2}{4}}$$

hence

$$\sigma_\lambda(\varphi_n, x) - \varphi_n(x) = O\left( e^{-\frac{\lambda^2}{4}} \frac{1}{1+|x|^2} \right)$$

consequently  $\|\sigma_\lambda \varphi_n - \varphi_n\|_p \rightarrow 0 (\lambda \rightarrow \infty)$ . Proposition 1 is proved. ■

**PROPOSITION 2.** *Let  $1 < p < \infty$  and define  $\tilde{f}$  by (4). Then*

a) *Every  $f \in L_p(\mathbb{R})$  has conjugate  $\tilde{f}$  and  $\|\tilde{f}\|_p \leq c_p \|f\|_p$ ,*

b)  $\tilde{\sigma}_\lambda f = \sigma_\lambda \tilde{f}$ ,

c)  $\tilde{f}$  *can be described by the other way given before (4).*

**PROOF.**

a) Follows from the Marcinkiewicz multiplier theorem because  $\frac{|t|}{t} = \text{sgn } t$  is a multiplier, see N.K. NIKOLSKII [3]. In b) we have to verify

$$(5) \quad \tilde{\sigma}_\lambda f = f * \tilde{K}_\lambda = \tilde{f} * K_\lambda = \sigma_\lambda \tilde{f}.$$

If  $f \in L_p \cap L_2$  then we can apply Fourier transform in (5), so we have to show that

$$\hat{f} \cdot \hat{K}_\lambda = \hat{\tilde{f}} \cdot \hat{K}_\lambda$$

which in turn follows from

$$(6) \quad \hat{K}_\lambda(t) = \frac{|t|}{t} \hat{K}_\lambda(t).$$

Indeed,

$$K_\lambda(x) = \frac{1}{2\pi} \int_{-\lambda}^{\lambda} \left(1 - \frac{|t|}{\lambda}\right) e^{ixt} dt = \mathcal{F}^{-1} \left( \frac{1}{\sqrt{2\pi}} \left(1 - \frac{|t|}{\lambda}\right) \chi_{|t| < \lambda} \right)$$

where  $\mathcal{F}^{-1}$  denotes the inverse Fourier transform, and

$$\tilde{K}_\lambda(x) = \frac{1}{2\pi} \int_{-\lambda}^{\lambda} \left(1 - \frac{|t|}{\lambda}\right) \frac{|t|}{t} e^{ixt} dt = \mathcal{F}^{-1} \left( \frac{1}{\sqrt{2\pi}} \left(1 - \frac{|t|}{\lambda}\right) \frac{|t|}{t} \chi_{|t| < \lambda} \right)$$

which implies (6). So (5) holds if  $f \in L_p \cap L_2$ . Now since  $K_\lambda, \tilde{K}_\lambda \in L_q$ ,  $1 < q \leq \infty$  hence by the Young inequality both sides of (5),  $f \mapsto f * \tilde{K}_\lambda$  and  $f \mapsto \tilde{f} * K_\lambda$  are continuous from  $L_p$  to  $L_r$  if  $\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1$  and  $\frac{1}{p} + \frac{1}{q} > 1$ . On the dense subset  $L_2 \cap L_p$  the two mappings are identical hence (5) extends to the whole  $L_p$ . This proves b).

Now the point c) follows from a) and b) since  $\tilde{\sigma}_\lambda f = \alpha_\lambda \tilde{f}$  tends to  $\tilde{f}$  in  $L_p$  by Proposition 1b). ■

Introduce the kernel

$$D_\lambda^*(x) = \frac{\lambda x \sin \lambda x + \cos \lambda x - 1}{\pi x^2} = \frac{1}{2\pi} \int_{-\lambda}^{\lambda} |t| e^{ixt} dt$$

and the corresponding partial sum operator

$$S_\lambda^*(f, x) = f * D_\lambda^*(x), \quad f \in L_p, \quad 1 \leq p < \infty$$

further its Fejér summation

$$K_\lambda^*(x) = \frac{1}{\lambda} \int_0^\lambda D_t^*(x) dt = \frac{2 \sin \lambda x - \lambda x (1 + \cos \lambda x)}{\pi \lambda x^3},$$

$$\sigma_\lambda^*(f, x) = f * K_\lambda^*(x) \quad f \in L_p, \quad 1 \leq p \leq \infty.$$

If  $f \in L_1$ , then changing the order of integrations we get

$$(7) \quad \begin{cases} S_\lambda^*(f, x) = \frac{1}{\sqrt{2\pi}} \int_{-\lambda}^{\lambda} |t| e^{ixt} \hat{f}(t) dt, \\ \sigma_\lambda^*(f, x) = \frac{1}{\sqrt{2\pi}} \int_{-\lambda}^{\lambda} |t| \left(1 - \frac{|t|}{\lambda}\right) e^{ixt} \hat{f}(t) dt. \end{cases}$$

The following variant of the Alexits lemma [1] holds

LEMMA 1. *If  $f \in L_p(\mathbb{R})$  with some  $1 \leq p \leq \infty$  then*

$$\|\sigma_\lambda f - f\|_p = O\left(\frac{1}{\lambda}\right) \quad \Leftrightarrow \quad \|\sigma_\lambda^* f\|_p = O(1)$$

$$(\lambda \rightarrow \infty) \qquad \qquad \qquad (\lambda \rightarrow \infty)$$

PROOF. a) Suppose first  $\|\sigma_\lambda^* f\|_p \leq K$ . A short counting gives that

$$\frac{d}{d\lambda} K_\lambda(x) = \frac{1}{\lambda^2} \frac{d}{d\lambda} (\lambda K_\lambda^*(x)) = \frac{\lambda x \sin \lambda x - (1 - \cos \lambda x)}{\pi x^2 \lambda^2}.$$

Consequently

$$\frac{d}{d\lambda} \sigma_\lambda(f, x) = \frac{1}{\lambda^2} \frac{d}{d\lambda} (\lambda \sigma_\lambda^*(f, x)).$$

Let  $0 < \lambda_1 < \lambda_2$ , then

$$\begin{aligned} \sigma_{\lambda_2}(f, x) - \sigma_{\lambda_1}(f, x) &= \int_{\lambda_1}^{\lambda_2} \frac{1}{\lambda^2} \frac{d}{d\lambda} (\lambda \sigma_\lambda^*(f, x)) d\lambda = \\ &= \left[ \frac{\sigma_\lambda^*(f, x)}{\lambda} \right]_{\lambda_1}^{\lambda_2} + 2 \int_{\lambda_1}^{\lambda_2} \frac{\sigma_\lambda^*(f, x)}{\lambda^2} d\lambda. \end{aligned}$$

Taking the  $L_p$ -norm of both sides we obtain

$$\|\sigma_{\lambda_2} f - \sigma_{\lambda_1} f\|_p \leq \frac{1}{\lambda_1} K + \frac{1}{\lambda_2} K + 2 \int_{\lambda_1}^{\lambda_2} \frac{K}{\lambda^2} d\lambda \leq \frac{4K}{\lambda}.$$

Since  $\|\sigma_\lambda f - f\|_p \rightarrow 0$ , hence  $\|\sigma_\lambda f - f\|_p \leq \frac{4K}{\lambda}$ .

b) Conversely suppose that  $\|\sigma_\lambda f - f\|_p \leq \frac{K}{\lambda}$ . Introduce the means

$$\sigma_{\lambda,2}(f, x) = \frac{2}{\lambda^2} \int_0^\lambda t \sigma_t(f, x) dt = f * K_{\lambda,2}(x)$$

where

$$\begin{aligned} K_{\lambda,2}(x) &= \frac{2}{\lambda^2} \int_0^\lambda t K_t(x) dt = \frac{2}{\pi} \int_0^\lambda \frac{1 - \cos tx}{\lambda^2 x^2} dt = \\ &= \frac{2(\lambda x - \sin \lambda x)}{\pi \lambda^2 x^3} = K_\lambda(x) - \frac{1}{\lambda} K_\lambda^*(x). \end{aligned}$$

Hence  $\sigma_{\lambda,2}f = \sigma_{\lambda}f - \frac{1}{\lambda}\sigma_{\lambda}^*f$  and then

$$\|\sigma_{\lambda,2}f - f\|_p = \left\| \frac{2}{\lambda^2} \int_0^{\lambda} t(\sigma_t f - f) dt \right\|_p \leq \frac{2}{\lambda^2} \int_0^{\lambda} t \frac{K}{t} dt = \frac{2K}{\lambda}$$

implies

$$\frac{1}{\lambda} \|\sigma_{\lambda}^*f\|_p = \|\sigma_{\lambda}f - \sigma_{\lambda,2}f\|_p \leq \|\sigma_{\lambda}f - f\|_p + \|\sigma_{\lambda,2}f - f\|_p \leq \frac{K}{\lambda} + \frac{2K}{\lambda}$$

i.e.  $\|\sigma_{\lambda}^*f\|_p \leq 3K$ . Lemma 1. is proved. ■

LEMMA 2. *If  $\|\sigma_{\lambda}^*f\|_p = O(1)$  and  $1 < p < \infty$  then  $\sigma_{\lambda}^*f$  tends weakly to some  $g \in L_p$  when  $\lambda \rightarrow \infty$ .*

PROOF. Introduce the operators

$$T_{\lambda} : L_q \rightarrow \mathbb{C}, \quad T_{\lambda}h = \int_{-\infty}^{\infty} h\sigma_{\lambda}^*f.$$

We have to show that there exists  $g \in L_p$  with

$$(8) \quad T_{\lambda}h \rightarrow \int_{-\infty}^{\infty} hg, \quad h \in L_q.$$

Prove first that the numbers  $T_{\lambda}h$  converge to some limit for every fixed  $h \in S(\mathbb{R})$  where  $S(\mathbb{R})$  is the Schwartz class of test functions (used to define tempered distributions). We know  $\hat{h} \in S$  and

$$\begin{aligned} T_{\lambda}h &= \int_{-\infty}^{\infty} h(x) \int_{-\infty}^{\infty} f(y)K_{\lambda}^*(x-y)dy dx = \\ &= \int_{-\infty}^{\infty} f(y) \int_{-\infty}^{\infty} h(x)K_{\lambda}^*(x-y)dx dy = \int_{-\infty}^{\infty} f(y)\sigma_{\lambda}^*(h,y)dy. \end{aligned}$$

Using (7) we obtain

$$S_{\lambda_2}^*(h,y) - S_{\lambda_1}^*(h,y) = \frac{1}{\sqrt{2\pi}} \left( \int_{\lambda_1}^{\lambda_2} - \int_{-\lambda_2}^{-\lambda_1} \right) e^{iyt} t \hat{h}(t) dt.$$

A simple integration by parts gives the estimate

$$S_{\lambda_2}^*(h,y) - S_{\lambda_1}^*(h,y) = o_{\lambda} \left( \frac{1}{1+|y|} \right)$$

consequently  $S_\lambda^*h$  is convergent in  $L_q$  and then  $\sigma_\lambda^*h = \frac{1}{\lambda} \int_0^\lambda S_t^*h dt$  is also convergent i.e.  $\int_{-\infty}^\infty h\sigma_\lambda^*f = \int_{-\infty}^\infty f\sigma_\lambda^*h$  is convergent in  $\mathbb{C}$ . The functionals  $T_\lambda$  are uniformly bounded:

$$|T_\lambda h| \leq \|h\|_q \|\sigma_\lambda^*f\|_p \leq c \|h\|_q.$$

In the dense subset  $S \subset L_q$  the functionals converge so they converge on the whole  $L_q$  to some continuous functional  $T$  which can be represented by some  $g \in L_p : Th = \int_{-\infty}^\infty gh$ . Lemma 2. is proved. ■

PROOF OF THE THEOREM. We can easily check that

$$(9) \quad \frac{d}{dx} \tilde{K}_\lambda(x) = \frac{1}{i\pi} \frac{(1 + \cos \lambda x)\lambda x - 2 \sin \lambda x}{\lambda x^3} = i K_\lambda^*(x)$$

and hence  $\frac{d}{dx} \tilde{\sigma}_\lambda(f, x) = i\sigma_\lambda^*(f, x)$

a) Suppose first that  $\|\sigma_\lambda f - f\|_p = O\left(\frac{1}{\lambda}\right)$ . Then by Lemma 1. and 2. there exists  $g \in L_p$  such that

$$\int_{-\infty}^\infty \left[ \frac{d}{dx} \tilde{\sigma}_\lambda(f, x) - ig(x) \right] h(x) dx = \int_{-\infty}^\infty [i\sigma_\lambda^*(f, x) - ig(x)] h(x) dx \rightarrow 0$$

if  $\lambda \rightarrow \infty$  and  $h \in L_q$ . In particular for  $h = \chi_{[a,y]}$  we obtain

$$(10) \quad \tilde{\sigma}_\lambda(f, y) - \tilde{\sigma}_\lambda(f, a) \rightarrow \int_a^y ig \quad (\lambda \rightarrow \infty, y \in \mathbb{R}).$$

Since  $\tilde{\sigma}_\lambda f \xrightarrow{L_p} \tilde{f}$ , we get that

$$(11) \quad \tilde{f}(y) = c + \int_a^y ig$$

i.e.  $\tilde{f}$  is locally absolutely continuous and  $\tilde{f}' \in L_p$ .

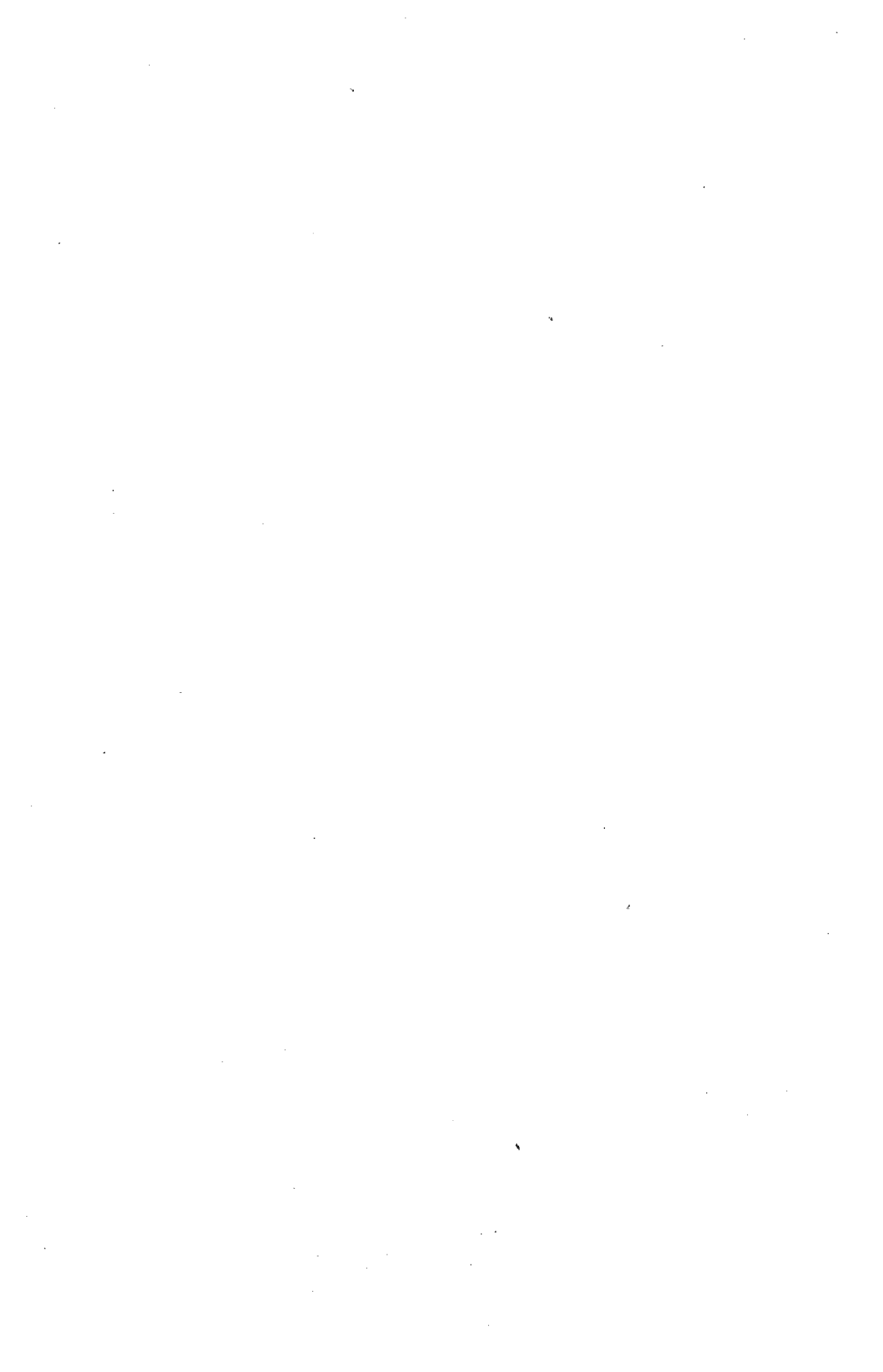
b) Suppose that (11) holds with some  $g \in L_p$ . Then

$$i\sigma_\lambda^*(f, x) = \frac{d}{dx} \tilde{\sigma}_\lambda(f, x) = \frac{d}{dx} \sigma_\lambda(\tilde{f}, x) = \sigma_\lambda(\tilde{f}', x) = i\sigma_\lambda(g, x)$$

i.e.  $\sigma_\lambda^* = \sigma_\lambda(g)$ . This implies by Proposition 1. that  $\|\sigma_\lambda^*\|_p = O(1)$ , thus  $\|\sigma_\lambda f - f\|_p = O\left(\frac{1}{\lambda}\right)$  which completes the proof. ■

### References

- [1] G. ALEXITS, On the order of approximation by Cesàro means of Fourier series, *Mat. Fiz. Lapok*, **48** (1941), 410–421. (English translation of the original Hungarian “A Fourier sor Cesàro-közepeivel való approximáció nagyságrendjéről” – see [3], pp. 41–50.)
- [2] G. ALEXITS, Sur l’ordre de grandeur de l’approximation d’une fonction périodique par les sommes de Fejér, *Acta Math. Acad. Sci. Hungar.*, **3** (1952), 29–40.
- [3] G. ALEXITS, *Approximation Theory, Selected Papers*, Akadémiai Kiadó (Budapest, 1983).
- [4] N.K. NIKOLSKII, *Approximation of functions with several variables and imbedding theorems*, (in Russian), Nauka (Moscow, 1969).
- [5] A. BOGMÉR, Generalization of a Theorem of G. Alexits, *Annales Univ. Sci. Budapest. Sect. Math.*, **31** (1988), 223–228.
- [6] I. JOÓ, Saturation theorems for Hermite Fourier series, *Acta Math. Acad. Sci. Hung.*, **57** (1991), 169–179.
- [7] M. HORVÁTH, Some saturation theorems for classical orthogonal expansions I., *Periodica Math. Hung.*, **22** (1991), 27–69.
- [8] M. HORVÁTH, Some saturation theorems for classical orthogonal expansions II., *Acta Math. Acad. Sci. Hung.*, **58** (1991), 139–169.
- [9] D.V. GIANG, Approximation on real line by Fourier transform, *Acta Sci. Math. (Szeged)*, **58** (1993), 197–209.



**ON A PROBLEM CONNECTED WITH THE WEIGHT DISTRIBUTION OF THE REED–MULLER CODE OF ORDER  $R^*$**

By

Á. G. HORVÁTH

Department of Geometry, Technical University of Budapest

(Received March 6, 1995)

**1. Introduction**

Let  $W(k)$  be the characteristic vector of a  $k$ -dimensional subspace of a vector space  $V$  of dimension  $n$  over the Galois-field  $GF(2)$  where  $k \geq 1$ . Let  $V_1, \dots, V_n$  be  $n - 1$ -dimensional subspaces of  $V$  that are generated by the subsets of  $n - 1$  elements of a fixed basis in  $V$ . It is well-known that the collection of binary vectors formed by component-wise multiplication of the  $r^{\text{th}}$  order Reed–Muller code ( $r \leq n$ ). The question is the following

*How many elements can be found in the  $r^{\text{th}}$ -order Reed–Muller code with the property that the weight of their component-wise product with the characteristic vector of  $W(k)$  is  $2^{k-1}$ ? (The weight of a vector is the number of 1's among its components.)*

In the case when  $r = 2$  this problem is solved (see in [3]), and the sought number for the case  $k = n$  is the number of codewords of weight  $2^{n-1}$ . We show that the numbers for the other values of  $k$  only depend on this one and formulate this connection. Let  $2^n$  be the dimension of the subspace  $W(k)$ . Let  $A(n, r, k)$  denote the above number. We will prove the following equalities:

$$A(n, r, k) = 2^{\binom{n}{1} + \dots + \binom{n}{r}} - \left( 1 + \binom{k}{1} + \dots + \binom{k}{r} \right) A(k, r, k) \quad \text{if } r \leq k,$$

and

$$A(n, r, k) = 2^{\sum_{i=0}^r \binom{n}{i} - (2^k - 1)} \cdot \binom{2^k}{2^k - 1} \quad \text{if } r \geq k.$$

---

\* Supported by Tempus JEP-06044-93 and Hung. Nat. Found for Sci. Research (OTKA) grant No. 7351 (1993)

## 2. Notation and Lemmas

Let  $R(r, n)$  and  $B(n, r)$  be the Reed–Muller code of order  $r$  and the corresponding group of subsets of  $V$ , respectively. We denote the group operation by  $\Delta$  which is the symmetric difference of two subsets of  $V$ . In this terminology we have that

$$A(n, r, k) = \left| \{M \in B(n, r) \mid |M \cap W(k)| = 2^{k-1}\} \right|.$$

We will call the collections of subspaces  $\{V_1, \dots, V_n\}$  and

$$\mathcal{G} = \{V, V_1, \dots, V_n, V_1 \cap V_2, \dots, V_{n-r+1} \cap \dots \cap V_n\}$$

basic generators and generators, respectively.  $\mathcal{G}$  is a linearly independent set as can be seen in Ch. 13 in [1].

LEMMA 1. *Let  $C(k)$  be a  $k$ -dimensional coset in  $V$ . Then*

$$A(n, r, k) = \left| \{M \in B(n, r) \mid |M \cap C(k)| = 2^{k-1}\} \right|.$$

PROOF. It is known that  $B(n, r)$  is generated by the collection of its minimal weight elements i.e. by those ones for which  $|M| = 2^{n-r}$ . It is also true that these elements are precisely the cosets of dimension  $n - r$  (see Th.8 in Ch. 13 and Th.12 in Ch.13 in [1]). This means that the group  $B(n, r)$  is invariant under a regular, inhomogeneous linear transformation of  $V$  so that the number  $A(n, r, k)$  is also invariant. Therefore the statement is true. ■

Let  $C(k)$  denote the coset  $(V \setminus V_n) \cap \dots \cap (V \setminus V_{k+1})$  and let  $\overline{\mathcal{G}}$  be the set of those generators which do not intersect the coset  $C(k)$ . If  $\overline{B(n, r)}$  and  $\overline{\overline{B(n, r)}}$  are the groups generated by the sets  $\overline{\mathcal{G}}$  and  $\mathcal{G} \setminus \overline{\mathcal{G}}$ , respectively, we can regard the group  $B(n, r)$  as the direct sum of these groups. This means that the rowspace  $R(r, n)$  is the direct sum of the rowspaces  $\overline{B(n, r)}$  and  $\overline{\overline{B(n, r)}}$ . (In the following we do not distinguish a set-group from the corresponding rowspace.) We can compute the number of elements of  $\overline{B(n, r)}$ .

LEMMA 2. *The number of elements of  $\overline{B(n, r)}$  is*

$$\left| \overline{B(n, r)} \right| = 2^{1 + \binom{n}{1} + \dots + \binom{n}{r}} - \left( 1 + \binom{k}{1} + \dots + \binom{k}{s} \right),$$

where  $s := \min\{k, r\}$ .

PROOF. Using the previous remark and the linear independence of the set  $\mathcal{G}$  we have:

$$\left| \overline{B(n, r)} \right| = 2^{\dim \overline{B(n, r)}} = 2^{\dim B(n, r) - \dim \overline{\overline{B(n, r)}}}.$$

But  $\overline{B(n,r)}$  is generated by those elements of  $\mathcal{G}$  which contain at least one vector from among the sets  $V_n, \dots, V_{k+1}$  as a factor. So the other generator set  $\mathcal{G} \setminus \overline{\mathcal{G}}$  is the collection of those generators which are intersections of the sets  $V, V_1, \dots, V_k$  and whose dimensions are not less than  $n - r$ . This means that the dimension of  $\overline{B(n,r)}$  is equal to  $\left(1 + \binom{k}{1} + \dots + \binom{k}{s}\right)$  where  $s := \min\{k, r\}$ . This proves the lemma. ■

### The Theorem

We now prove the following theorem:

**THEOREM.**

$$A(n, r, k) = 2^{\left(1 + \binom{n}{1} + \dots + \binom{n}{r}\right) - \left(1 + \binom{k}{1} + \dots + \binom{k}{r}\right)} \cdot A(k, r, k) \quad \text{if } r \leq k,$$

and

$$A(n, r, k) = 2^{\sum_{i=0}^r \binom{n}{i} - (2^k - 1)} \cdot \binom{2^k}{2^k - 1} \quad \text{if } r \geq k.$$

We will use a theorem about the recursive setting up of the Reed-Muller codes.

**STATEMENT.** (See Th.2 of Ch.13 in [1].) *Denote by  $|u \mid v|$  the direct sum of the vectors  $u$  and  $v$ . So it is a vector of length  $n + m$  if the length of  $u$  and  $v$  is  $n$  and  $m$ , respectively, the first  $n$  components being those of  $u$ , the last  $m$  components those of  $v$ . Then*

$$R(r, n + 1) = \{|u \mid u + v| \mid \text{where } u \in R(n, r) \text{ and } v \in R(r - 1, n)\}.$$

*provided that in the codewords the order of components is the lexicographical order of the points of  $V$ , with respect to expansion in the fixed basis, and the basis in  $GF(2)^{(n+1)}$  has first  $n$  elements those of the basis in  $GF(2)^n$ .*

**THE PROOF OF THE THEOREM.** Let  $1 \leq r \leq n$  be an arbitrary integer and consider the value  $A(k + 1, r, k)$ . First we assume that  $r \leq k$ . Without loss of generality we may also assume that  $W(k)$  is the subspace of  $V$  spanned by the first  $k$  basic vectors. So if  $M$  is an element of the set

$$H(k + 1, r, k) := \left\{ M \in B(k + 1, r) \mid |M \cap W(k)| = 2^{k-1} \right\}$$

then  $M$  has the form  $|u \mid u + v|$ , where  $u \subset W(k)$ ,  $u \in B(k, r)$  and  $v \in B(k, r - 1)$ . On the basis of the definition of  $H(k + 1, r, k)$  it can be seen

that  $u$  is in the set  $H(k, r, k)$ , too. (In the definition of  $H(k, r, k)$ ,  $W(k)$  can be taken as the  $k$ -dimensional vector space.) This means that

$$A(k+1, r, k) = |\{B(k, r-1)\}| \cdot A(k, r, k).$$

But

$$|\{B(k, r-1)\}| = 2^{1+\binom{k}{1}+\binom{k}{2}+\dots+\binom{k}{r-1}}$$

so

$$A(k+1, r, k) = 2^{1+\binom{k}{1}+\binom{k}{2}+\dots+\binom{k}{r-1}} \cdot A(k, r, k).$$

Similarly, the codewords in  $H(k+2, r, k)$  have the form

$$\left| |u \mid u+v| \mid |u \mid u+v|+w \right|$$

where  $|u \mid u+v| \in H(k+1, r, k)$ , and  $w$  is an arbitrary element of  $B(k+1, r-1)$ . Thus we get that:

$$A(k+2, r, k) = 2^{\left(1+\binom{k+1}{1}+\binom{k+1}{2}+\dots+\binom{k+1}{r-1}\right)} \cdot 2^{\left(1+\binom{k}{1}+\binom{k}{2}+\dots+\binom{k}{r-1}\right)} \cdot A(k, r, k).$$

Since we can continue, in this way we have the following equality

$$A(n, r, k) = A(k, r, k) \cdot 2^{\sum_{j=k}^{n-1} \left(1+\binom{j}{1}+\dots+\binom{j}{r-1}\right)}.$$

Add now the number  $1 + \binom{k}{1} + \dots + \binom{k}{r}$  to the exponent of 2. Then we get the expression  $1 + \binom{n}{1} + \dots + \binom{n}{r}$  so we have the more simple form of  $A(n, r, k)$ :

$$A(n, r, k) = 2^{\left(1+\binom{n}{1}+\dots+\binom{n}{r}\right) - \left(1+\binom{k}{1}+\dots+\binom{k}{r}\right)} \cdot A(k, r, k).$$

This proves the first statement of the Theorem.

Assume now that  $r \geq k$ . From Lemma 2 of the previous paragraph we can see that if  $C(k) = (V \setminus V_n) \cap \dots \cap (V \setminus V_{k+1})$ ,  $\overline{\mathcal{G}} = \{M \in \mathcal{G} \mid M \cap C(k) = \emptyset\}$  and  $\overline{\overline{B(n, r)}} := \langle \mathcal{G} \setminus \overline{\mathcal{G}} \rangle_{\Delta}$  then

$$A(n, r, k) := \left| \left\{ M \in \overline{\overline{B(n, r)}} \mid |M \cap C(k)| = 2^{k-1} \right\} \right|.$$

In the case when  $r \geq k$  the generator set  $\mathcal{G} \setminus \overline{\mathcal{G}}$  is the collection of those generators which are intersections of the vectors  $V, V_1, \dots, V_k$  of dimension at least  $n-r$  so

$$\mathcal{G} \setminus \overline{\mathcal{G}} = \{V, V_1, \dots, V_k, V_1 \cap V_2, \dots, V_{k-1} \cap V_k, \dots, V_1 \cap \dots \cap V_k\}.$$

Thus  $\overline{\overline{B(n,r)}} = \overline{\overline{B(n,k)}}$  so  $\overline{\overline{A(n,r,k)}} = \overline{\overline{A(n,k,k)}} = A(k,k,k)$ . (We used at the last equality that we have two formulas for the number  $A(n,r,k)$  if  $r = k$ .) Since

$$A(k,k,k) = \binom{2^k}{2^{k-1}} \quad \text{and} \quad \binom{k}{1} + \dots + \binom{k}{k} = 2^k - 1$$

we have proved the second statement of the Theorem, too. ■

#### 4. On the computational methods of $H(n,r,n)$

It is clear from the proved formulas that the algorithmic handling of this problem is very hard. (This is an *NP*-hard problem.) But for a small  $n$  ( $n \leq 5$ ) by the help of the **GAP** system (see [3]) the number  $A(n,r,k)$  can be determined. In particular the obtained values lead to the discovery of the above statements. However the numbers  $A(n,r,n)$  are not known if  $n \geq 6$ .

We now sketch a possible method which raises some new research problems.

In the proof of Lemma 1 we used that the automorphism group of  $B(n,r)$  contains those permutations of the elements of the vector space  $V$  which are generated by a regular affine transformation of  $V$ . However in general the group of these permutations (which is the so called *general affine group* and is denoted by  $GA(n)$ ), is equal to the full automorphism group  $\text{Aut}(B(n,r))$  if  $1 \leq r \leq n-2$  cf. Ch. 13 in [1]. (We note that in the case when  $r=0$ ,  $n-1$  or  $n$  then this group of  $S_{2^n}$ , see Th.24 of Ch.13 in [1].) Such an automorphism does not change the weight of a codeword (or set) so  $H(n,r,n)$  is invariant under the automorphisms of  $B(n,r)$ . Take now an element  $M$  in  $H(n,r,n)$  and let  $B_M$  be the subgroup of  $\text{Aut}(B(n,r))$  consisting of those automorphisms which fix the element  $M$ . If now  $\text{Aut}_M(B(n,r))$  is the equivalence class of those elements of  $H(n,r,n)$  which are isomorphic to  $M$  then we have that

$$|\text{Aut}_M(B(n,r))| = |\text{Aut}(B(n,r))/B_M|.$$

This means that  $A(n,r,n)$  is the sum of the cardinalities of the classes of the isomorphic elements of  $H(n,r,n)$ . Thus

$$A(n,r,n) = \sum_{M \in \mathcal{R}} |\text{Aut}_M(B(n,r))|$$

where  $\mathcal{R}$  is the maximal set with the property; if  $M$  and  $N$  are in  $\mathcal{R}$  then  $M$  is not isomorphic to  $N$ .

For instance if  $M := V_1$  then the cardinality of  $B_{V_1}$  can be computed easily:

$$|B_{V_1}| = |GA(n-1)| \cdot 2^{n-1}.$$

So

$$|\text{Aut}_{V_1}(B(n,r))| = \frac{|GA(n)|}{|GA(n-1)| \cdot 2^{n-1}} = 2^{n+1} - 2.$$

The sketch of a possible algorithm is the following. Take an element  $M$  from  $H(n,r,n)$  and determine the group  $B_M$ . Then we compute the cardinality of the coset space  $\text{Aut}(B(n,r))/B_M$  and those elements  $N$  of  $B(n,r)$  which are isomorphic to  $M$  are struck off the list of the possible elements. Finally we choose another element from the set  $H(n,r,n)$ . This concept raises the following (research) problems:

1. Over to find (rapidly) such an element of  $H(n,r,n)$  which is not isomorphic to the found one?
2. How can we determine the group  $B_M$  and the orbit  $\text{Aut}_M(B(n,r))$ ?

Last but not least I would like to express my thanks to Prof. J. NEUBÜSER for some discussions.

### References

- [1] F. J. MACWILLIAMS, N. J. A. SLOANE, *The Theory of Error-Correcting Codes*, North-Holland Publ. Co., Amsterdam–New York–Oxford, 1978.
- [2] ÁKOS G. HORVÁTH, Codes and Lattices, *Per. Pol. Mech. Eng.*, **36** (1992).
- [3] ÁKOS G. HORVÁTH, On the second-order Reed–Muller code, *Annalas Univ Sci. Eötvös*, Sect. Math., submitted.
- [4] MARTIN SCHÖNERT ET. AL., *GAP — Groups, Algorithms, and Programming*, Lehrstuhl D für Mathematik, RWTH, Aachen, Germany, 3<sup>th</sup> edition, 1993.

## BASES OF VECTOR EXPONENTIALS WITH A STRIP IN THE SPECTRUM

By

I. JOÓ\* and A. MINKIN

Department of Analysis, Eötvös Loránd University, Budapest and

Department of Mathematics, Saratov University, Russia

(Received March 14, 1995)

1. Unconditional bases of scalar exponentials in a finite interval  $(0, a)$  have been completely investigated during the previous 15 years [1]–[6]. In the case of a system of vector exponentials (here  $\chi_{(0,a)}$  denotes an indicator function of the interval  $(0, a)$ ,  $a > 0$ )

$$\Phi_a = \{e_\lambda^\alpha(t) \cdot \xi_\lambda\}_{\lambda \in \sigma}, \quad e_\lambda^\alpha(t) = e_\lambda \cdot \chi_{(0,a)},$$

$$e_\lambda(t) = \exp(-i\lambda t), \quad \xi_\lambda \in E, \quad \dim E = N < \infty$$

a criterion of unconditional basicity of  $\Phi_a$  in  $L^2(0, a)$  (notation:  $\Phi_a \in (UB)$ ) has been established only in the case of a half-bounded spectrum (see for instance, [3]):

$$\inf\{\operatorname{Im}\lambda \mid \lambda \in \sigma\} > 0.$$

However, it is rather difficult to extend the arguments in [6] to vector exponentials. Therefore we restrict ourselves to the case

$$(1) \quad \inf\{|\operatorname{Im}\lambda \mid \lambda \in \sigma\} > 0$$

when there is a strip in the spectrum. Our considerations basically follow those in the scalar situation [4] but a special care is needed to handle manipulations operator-valued inner functions.

Nevertheless, in the vector case not all *scalar* results are possible to obtain. We mean that there are yet no analogues of the well-known Muckenhoupt condition  $(A_2)$ . Therefore the criterion obtained in the theorem 1 requires further improvement. We refer also to the article [5] for some sufficient conditions of boundedness of the Hilbert transform in a space of vector-valued functions with a matrix weight.

---

\* The research was supported by the Hungarian National Foundation OTKA No T 014244.

2. We shall need the following definitions and notations.

$$\mathbb{C}_{\pm} = \{\pm \operatorname{Im} z > 0\}, \quad \sigma_{\pm} = \sigma \cap \mathbb{C}_{\pm}, \quad \bar{\sigma}_{-} = \{\bar{\lambda} \mid \lambda \in \sigma_{-}\};$$

$$e_{\lambda}^{+} = e_{\lambda} \cdot \chi_{(0, \infty)}, \quad e_{\lambda}^{-} = e_{\lambda} \cdot \chi_{(-\infty, 0)};$$

$L^2(\mathbb{R}, E)$  — the space of all square summable  $E$ -valued functions where  $E$  is an auxiliary  $N$ -dimensional euclidean space;

$H_{\pm}^2(E)$  — the corresponding Hardy space of  $E$ -valued functions analytic in the upper/lower half-plane;

$P_{\pm}$  — the Riesz projection onto  $H_{\pm}^2(E)$  in  $L^2(\mathbb{R}, E)$ ;

$L(E \rightarrow E)$  — the space of linear operators acting in  $E$ ;

for an inner (in  $\mathbb{C}_{\pm}$ ) operator-valued function  $\psi_{\pm}$ , taking values in  $L(E \rightarrow E)$ , set

$$H(\psi_{\pm}) = H_{\pm}^2(E) \ominus \psi_{\pm} H_{\pm}^2(E),$$

where  $\ominus$  stands for the orthogonal difference between two subspaces.

For any subspace  $M \subset L^2(\mathbb{R}, E)$  and operator-valued function  $\psi$ ,  $\psi(k) \in L(E \rightarrow E)$ , set  $\psi M = \{\psi f \mid f \in M\}$ .

Let also

$$L_a = \operatorname{span} \Phi_a, \quad L_a \subset L^2(0, a; E); \quad L_a^{\pm} = \operatorname{span} \Phi_a^{\pm}$$

where  $\Phi_a^{\pm} = \{e_{\lambda} \cdot \xi_{\lambda} \mid \lambda \in \sigma_{\pm}\}$  and set  $\Phi = \Phi^{+} \cup \Phi^{-}$ ;

$$\Phi^{+} = \{e_{\lambda^{+}} \cdot \xi_{\lambda} \mid \lambda \in \sigma_{+}\}; \quad \Phi^{-} = \{e_{\lambda^{-}}(t-a) \cdot \xi_{\lambda} \mid \lambda \in \sigma_{-}\};$$

$$X_{\lambda}^{\pm} = \frac{i}{\sqrt{2\pi}} \cdot \frac{1}{z - \lambda} \xi_{\lambda} \in H_{\pm}^2(E), \quad \lambda \in \mathbb{C}_{\pm}$$

$$X^{\pm} = \{X_{\lambda}^{\pm} \mid \lambda \in \sigma_{\pm}\}, \quad X = X^{+} \cup \Theta X^{-}, \quad \Theta X^{-} = \{\Theta X_{\lambda}^{-} \mid \lambda \in \sigma_{-}\},$$

$$\Theta = \exp(iaz) \cdot I_E,$$

$I_E$  is the unit operator in  $E$ .

Further let  $I$  be the unit operator in  $L^2(\mathbb{R}, E)$  and set

$$P_{\Theta} = I - \Theta P_{+} \Theta^{*}, \quad X_{\Theta} = P_{\Theta} X = P_{\Theta} X^{+} \cup P_{\Theta} \Theta X^{-} = X_{\Theta}^{+} \cup X_{\Theta}^{-}.$$

At last let  $\mathcal{F}^{-1}$  be the inverse Fourier transform

$$(\mathcal{F}^{-1}f)(s) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} \exp(its) f(t) dt$$

and let  $P_a$  be the orthoprojection onto  $L^2(0, a; E)$  in  $L^2(\mathbb{R}, E)$ .

For the sake of brevity we shall denote below  $L^2(\mathbb{R}, E)$  as  $H$  and let

$$G = \operatorname{span} X^{+}, \quad F = \operatorname{span} \Theta X^{-}.$$

3. Let us derive at first some needed properties of vector families introduced above. Evidently,

$$P_a \Phi^+ = \Phi_a^+, \quad P_a \Phi^- = \{P_a e_\lambda^+(t-a) \mid \lambda \in \sigma_-\} = \\ = \{e_\lambda^-(t-a) \cdot \xi_\lambda, 0 \leq t \leq a \mid \lambda \in \sigma_-\} = \{e_\lambda^a \cdot \xi_\lambda(-a) \mid \lambda \in \sigma_-\}.$$

Therefore all the properties: unconditional basicity (*UB*), uniform minimality (*UM*) and minimality (*M*) are one and the same for both systems,  $\Phi_a^-$  and  $P_a \Phi^-$  or  $P_a \Phi$  and  $\Phi_a$ . By direct calculation we find that

$$X_\lambda^\pm = \mathcal{F}^{-1} e_\lambda^\pm \xi_\lambda \in H_\pm^2(E), \quad \lambda \in \sigma_\pm.$$

Moreover,  $P_\Theta = \mathcal{F}^{-1} P_a \mathcal{F}$  is an orthoprojector onto  $H(\Theta)$  in  $H$ ;

$$\mathcal{F}^{-1} P_a \Phi = \mathcal{F}^{-1} P_a \Phi^+ \cup \mathcal{F}^{-1} P_a \Phi^- = P_\Theta X^+ \cup P_\Theta \Theta X^-$$

because

$$\mathcal{F}^{-1}(f(t-a))(s) = \exp(isa) \mathcal{F}^{-1} f(s).$$

LEMMA 1.  $X_\Theta^+ \in (M) \Rightarrow X^+ \in (M)$ .

PROOF. For a finite linear combination  $\sum_{\mu \neq \lambda} c_\mu x_\mu^+$  we have

$$\left\| x_\lambda^+ - \sum_{\mu \neq \lambda} c_\mu x_\mu^+ \right\| \geq \left\| P_\Theta \left( x_\lambda^+ - \sum_{\mu \neq \lambda} c_\mu x_\mu^+ \right) \right\| \geq \delta = \delta(\lambda) > 0$$

because  $X_\Theta^+ \in (M)$ . ■

Quite analogously, if  $X_\Theta^- \in (M)$  then  $\Theta X^- \in (M)$ . But for the families  $X^\pm$  it is known ([4], p.52) that minimality yields the so-called Blaschke condition for the portions of the spectrum (notation,  $\sigma_\pm \in (B)$ ):

$$(B) \quad \sum_{\lambda \in \sigma_\pm} \frac{|\operatorname{Im} \lambda|}{1 + |\lambda|^2} < \infty.$$

Hence, there exists the corresponding Blaschke–Potapov products  $\Pi_\pm$  ([4], p.36) such that

$$\det \Pi_\pm(z) \equiv B(z, \sigma_\pm)$$

is a Blaschke product in  $\mathbb{C}_\pm$  with zeros  $\sigma_\pm$  and

$$\ker \Pi_\pm^*(\lambda) = L_\lambda = \operatorname{span} (\xi_\lambda), \quad \lambda \in \sigma_\pm; \quad \dim L_\lambda = 1.$$

LEMMA 2.  $X_{\Theta}^+ \in (UM) \Rightarrow X^+ \in (UM)$ .

PROOF. We have only to check the double-sided inequality

$$(2) \quad \|x_{\lambda}^+\| \asymp \|x_{\Theta\lambda}^+\|$$

where  $A \asymp B$  means that  $C_1 \leq |A/B| \leq C_2$  for variables  $A, B$  and some positive absolute constants  $C_1, C_2$ . But

$$\|x_{\lambda}^+\|^2 = \frac{1}{2|\operatorname{Im}\lambda|} \cdot \|\xi_{\lambda}\|^2, \quad \lambda \in \sigma_+;$$

$$\|x_{\Theta\lambda}^+\|^2 = \|e_{\lambda}^a \cdot \xi_{\lambda}\|_{L^2(0,a;E)}^2 \asymp (1 + |\operatorname{Im}\lambda|)^{-1} \|\xi_{\lambda}\|^2.$$

It remains now to apply Pavlov's lemma ([1, p.228]) which completes the proof.  $\blacksquare$

REMARK 1. This lemma is clearly valid for the system  $X_{\Theta}^- = P_{\Theta}\Theta X^-$ , or equivalently,  $\Phi_a$ . One must only take into account that the elements  $\mathcal{F}x_{\Theta\lambda}^-$  and  $e_{\lambda}^a \cdot \xi_{\lambda}$  differ by the multiple  $e_{\lambda}^a$ . But really uniform minimality is a property of the set of subspaces  $\{\operatorname{span}(e_{\lambda}^a \cdot \xi_{\lambda}) \mid \lambda \in \sigma_-\}$  and not of their individual elements.

4. Let us state now our main results.

THEOREM 1. *The system  $\Phi_a$  constitutes an unconditional basis in  $L^2(0,a;E)$  if and only if the following two conditions are valid.*

- i)  $X^{\pm}$  is an unconditional basis in its span (in  $H^{\pm}(E)$ , respectively); notation:  $X^{\pm} \in (UB')$ .
- ii) the Toeplitz operator  $T_{\psi}$  with the symbol

$$\psi(k) = \Theta^*(k)\Pi(k), \quad k \in \mathbb{R},$$

is an isomorphism onto  $H_+^2(E)$ ;

$$T_{\psi} : H_+^2(E) \rightarrow H_+^2(E); \quad T_{\psi}f := P_+\psi f.$$

Here

$$\Pi(z) = \overline{\Pi_-(z)} \cdot \Pi_+(z), \quad \overline{\Pi_-(z)} := \Pi_-(\bar{z})^*,$$

$\Pi(z)$  is a Blaschke–Potapov product in  $\mathbb{C}_+$  with zeros  $\sigma_+ \cup \overline{\sigma_-}$ .

THEOREM 2.  $\Phi_a \in (UB')$  if and only if the following two conditions are valid:

- i)  $X^{\pm} \in (UB')$ ;
- ii)  $T_{\psi}H_+^2 = H_+^2$ .

It is known that in the theorem 1 condition i) is equivalent to the uniform minimality of the system  $X^{\pm}$  ([7], theorem 9.1). The latter has been

investigated and encoded in [7], see for instance lemma 15.3 therein, and we refer to the reader to this deep article for further details. Condition ii) may be written as follows:

$$(3) \quad \text{dist}(\psi, H_+^\infty(E)) < 1,$$

$$(4) \quad \text{dist}(\psi, H_-^\infty(E)) < 1,$$

where the distance “dist” is taken in the space  $L^\infty(E)$ . Condition ii) of the theorem 2 is equivalent to (4) (see [8], lecture VIII).

Proof of the theorems 1,2 is carried on below in sections 5–8.

5. First we prove theorem 2 provided the set  $\sigma_-$  is empty. Then

$$\Phi_a \in (UB') \Rightarrow \Phi_a \in (UM) \Rightarrow X^+ \in (UM) \Leftrightarrow X^+ \in (UB').$$

Hence, using the scheme of the projection method we come to an equivalence:

$$[\Phi_a \in (UB')] \Leftrightarrow$$

$$1) X^+ \in (UM) \text{ and}$$

$$2) P_\Theta|_{\text{span } X^+} \text{ is an isomorphism (onto span } X_\Theta).$$

If  $L$  and  $M$  are closed subspaces, we say that the algebraic sum  $L + M$  is a direct sum (notation  $L \dot{+} M$ ) if the angle of the subspaces  $L$  and  $M$  is positive; the angle is defined below just after (10).

Since  $\text{span } X^+ = H(\Pi^+)$  we conclude that 2) is equivalent to the following:

$$(5) \quad H(\Pi, E) + \ker P_\Theta \quad \text{is a direct sum.}$$

Here we employed the well-known

PROPOSITION 1. *Let  $L, M$  be closed subspaces in a Hilbert space  $H$ ;  $L^\perp, M^\perp$  be their orthogonal complements in  $H$ , and let  $P$  be the orthoprojection onto  $L$  in  $H$ . Then*

$$A) (P | M \text{ — isomorphism}) \Leftrightarrow [M + L^\perp \text{ is a direct sum}] \Leftrightarrow H = M^\perp + L;$$

$$B) (P | M \text{ — is an isomorphism onto } L) \Leftrightarrow M \dot{+} L^\perp = H \Leftrightarrow M^\perp \dot{+} L = H;$$

$$C) \text{ for any pair of closed subspaces } T, S \text{ in } H [T + S \text{ is a direct sum}] \Leftrightarrow T^\perp + S^\perp = H;$$

$$D) T \dot{+} S = H \Leftrightarrow T^\perp \dot{+} S^\perp = H.$$

For the proof see [8], p. 258; the main tool is the Banach embedding theorem.

Clearly,  $\ker P_\Theta = H_-^2(E) \oplus \Theta H_+^2(E)$  and adding the summand  $H_-^2(E)$  to the first term in (5) we come to a new direct sum:  $\Pi_+ H_-^2(E) \dot{+} \Theta H_+^2(E)$ . Going

over to orthogonal complements and exploiting part C) of the Proposition 1 we obtain

$$(6) \quad \Pi_+ H_+^2(E) + \Theta H_-^2(E) = H.$$

Therefore the Toeplitz operator  $T_\psi$  in  $H_+^2(E)$  with the symbol  $\varphi = \Theta^* \Pi_+$  is an epimorphism, i.e.

$$(7) \quad T_\varphi H_+^2(E) = H_+^2(E).$$

and this completes the proof. ■

6. Let us make further transformations. Clearly,  $X_\Theta \in (UB)$  is equivalent to the following three conditions

$$(8) \quad 1. \quad X_\Theta^\pm \in (UB').$$

$$(9) \quad 2. \quad \varphi(\text{span } X_\Theta^+, \text{span } X_\Theta^-) > 0.$$

$$(10) \quad 3. \quad \text{span}(X_\Theta^+, X_\Theta^-) = H(\Theta).$$

Here  $\varphi(M, N)$  stands for the angle between two subspaces  $M, N \subset H$ :

$$\varphi(M, N) = \inf_{\substack{x \in M \\ y \in N}} \arccos \frac{|(x, y)|}{\|x\| \|y\|}$$

LEMMA 3. *Relation (9) yields*

$$(11) \quad \varphi(G, F) > 0.$$

PROOF. First we note that

$$G = \text{span } X^+ \subset [\text{span } P_\Theta X^+] \oplus \Theta H_+^2(E) = \text{span } X_\Theta^+ \oplus \Theta H_+^2(E)$$

because  $X^+ \subset H_+^2(E)$ . Similarly  $\Theta X^- \subset \Theta H_-^2(E)$  and

$$F = \text{span } \Theta X^- \subset H_-^2(E) \oplus \text{span } P_\Theta \Theta X^- = H_-^2(E) \oplus \text{span } X_\Theta^-$$

But  $\text{span } X_\Theta^- \dot{+} \text{span } X_\Theta^+$  is already a direct sum. Adding orthogonal summands to it we obtain again the direct sum  $G \dot{+} F$  which is equivalent to (11). ■

LEMMA 4. *Relations (8)–(10) yield*

$$(12) \quad P_\Theta | (G \dot{+} F) \text{ is an isomorphism onto } H(\Theta).$$

PROOF. Recall that the sum  $G + F$  is a direct one according to the lemma 3. Let  $g \in G, f \in F$ . Then

$$\|g + f\| \asymp \|g\| + \|f\| \asymp \|P_\Theta g\| + \|P_\Theta f\|$$

according to (8) and to the part of the theorem 2 proved above in the section 5. However (9) implies that

$$\|P_\Theta g\| + \|P_\Theta f\| \asymp \|P_\Theta g + P_\Theta f\| = \|P_\Theta(f + g)\|$$

and this completes the proof. ■

LEMMA 5.  $X_\Theta \in (UB)$  if and only if

$$(13) \quad X^\pm \in (UB')$$

and conditions (11)–(12) are satisfied.

PROOF. Indeed, the necessity of (11)–(12) was established in the lemmas 3, 4. Relation (13) stems immediately from lemma 2 and the main result of [7].

Sufficiency. Clearly, (11)–(13) yield (8), (10). Let us show now (9). For  $u \in \text{span } X_\Theta^+$  and  $v \in \text{span } X_\Theta^-$  there exist  $g \in G, f \in F$  such that

$$P_\Theta(g) = u, \quad P_\Theta(f) = v, \quad \|u + v\| = \|P_\Theta(g + f)\| \asymp \|g + f\| \asymp \|g\| + \|f\|.$$

In this calculation we employed first (12), then (11) and at last the proved part of Theorem 2. ■

7. PROOF OF THEOREM 1. We shall reduce the set of three conditions (11)–(13) to a chain of equivalent relations. Applying proposition 1 we readily replace (12) with the following

$$(14) \quad [G \dot{+} F] \dot{+} \ker P_\Theta = H$$

or equivalently

$$[H(\Pi_+) \dot{+} \Theta H(\Pi_-)] \dot{+} [H_-^2(E) \oplus \Theta H_+^2(E)] = H.$$

Regrouping summands we obtain

$$H = [H(\Pi_+) \dot{+} H_-^2(E)] \dot{+} [\Theta H(\Pi_-) \dot{+} \Theta H_+^2(E)] = \Pi_+ H_-^2(E) \dot{+} \Theta \Pi_- H_+^2(E).$$

Taking advantage of the relation (B) in proposition 1 we get

$$(15) \quad H = \Pi_+ H_+^2(E) \dot{+} \Theta \Pi_- H_-^2(E).$$

Clearly multiplication by an inner function or its adjoint is a unitary transformation in  $H = L^2(\mathbb{R}, E)$ . Then, multiplying both sides of (15) by  $\Pi_-^* \Theta^*$  on the right and applying the Riesz projection  $P_+$  we conclude that  $T_\psi$  is an isomorphism onto  $H_+^2(E)$ , the symbol  $\psi = \Pi_-^* \Theta^* \Pi_+ = \Theta^* \Pi_-^* \Pi_+$  since  $\Theta$  commutes with  $\Pi_-$ . Hence, it remains to remove condition (11)

LEMMA 6. Condition (11) follows from the relation

$$(11) \quad \Pi_+ H_+^2(E) + \Theta \Pi_- H_-^2(E) = H.$$

PROOF. We have that (11)  $\Leftrightarrow$  [ $G + F$  is a direct sum]  $\Leftrightarrow G^\perp + F^\perp = H$ .  
But

$$G^\perp = H(\Pi_+)^\perp = H_-^2(E) \oplus \Pi^+ H_+^2(E),$$

$$F^\perp = [\Theta H(\Pi_-)]^\perp = \Theta [H(\Pi_-)]^\perp = \Theta \cdot [\Pi_- H_-^2(E) \oplus H_+^2(E)].$$

Hence,

$$H_-^2(E) + \Pi_+ H_+^2(E) + \Theta \Pi_- H_-^2(E) + \Theta H_+^2(E) = H.$$

Multiplying this equality by  $\Pi_-^* \Theta^*$  we come to a relation:

$$(17) \quad \Pi_-^* \Theta^* H_-^2(E) + \Pi_-^* \Theta^* \Pi_+ H_+^2(E) + H_-^2(E) + \Pi_-^* H_+^2(E) = H.$$

Multiplying now (16) by  $\Pi_-^* \Theta^*$  we see that the sum of second and third summands in (16) absorbs other two terms which completes the proof of the lemma 6. ■

Hence, the proof of the theorem 1 is finished.

**8. PROOF OF THE THEOREM 2.** All the considerations coincide with those of the theorem 1. We have only to remove (10) as well as words “onto  $H(\Theta)$ ” in (12). Then (14) means that the summands from the left constitute a direct sum and we have to replace below all direct sums by algebraic ones.

Therefore we come to the relation  $H = \Pi_+ H_+^2(E) + \Theta \Pi_- H_-^2(E)$  or, equivalently, [ $T_\psi$  is an epimorphism onto  $H_+^2(E)$ ]. The latter yields that  $(T_\psi)^* = T_{\psi^*}$  is an isomorphism which completes the proof. ■

## References

- [1] S. V. HRUSCEV, N. K. NIKOL'SKII, B. S. PAVLOV, Unconditional bases of exponentials and of reproducing kernels, *LNM*, **864** (1981), 214–335.
- [2] S. A. AVDONIN, I. JOÓ, Riesz bases of exponentials and sine-type functions, *Acta Math. Hung.*, **51** (1988), 3–14.
- [3] S. A. AVDONIN, S. A. IVANOV, and I. JOÓ, On Riesz bases from vector exponentials I., *Annales Univ. Sci. Budapest.*, Sect. Math., **32** (1989) 101–114.
- [4] S. A. AVDONIN, S. A. IVANOV, *Controllability of systems with distributed parameters and exponential families*, UMK VO, Kiev, 1989 (Russian).
- [5] S. A. IVANOV, B. S. PAVLOV, Vector systems of exponentials and zeros of entire matrixfunctions, *Vestnik LGU*, Ser. Mat. Mech. Astr., **1** (1980), 25–31 (Russian).
- [6] A. M. MINKIN, Reflection of exponents and unconditional exponential bases, *Algebra i Analiz*, **3** (1991), 110–135; English transl. in *St. Petersburg Math. J.*, **3** (1992), 1043–1068.
- [7] S. R. TEIL', Hankel operators, embedding theorems and bases of covariant subspaces of the multiple shift operator, *Algebra i Analiz*, **1** (1989), 200–234 (Russian).
- [8] N. K. NIKOL'SKII, *Treatise on the shift operator*, Nauka, 1980 (Russian).

## ON WEIGHTED (0, 2) INTERPOLATION

By

I. JOÓ

Department of Analysis, Eötvös Loránd University, Budapest

*(Received March 25, 1995)*

Dedicated to Professor JÁNOS BALÁZS on his 75th birthday

Firstly E. EGERVÁRY and P. TURÁN have started the study of (0, 2) interpolation in order to get approximate solution of the differential equation

$$y'' + f \cdot y = 0.$$

Their results had been developed by some other mathematicians. These — so called — lacunary interpolation polynomials usually cannot be determined uniquely. An other difficulty is that they have no simple explicit form and therefore the convergence theorems related to these polynomials are rather complicated.

In order to avoid these difficulties P. TURÁN suggested the following modified problem.

Let  $(a, b)$  be a finite or infinite interval,

$$-\infty \leq a < \xi_{n,n} < \dots < \xi_{1,n} < b \leq +\infty \quad (n \in \mathbb{N})$$

distinct fundamental points and  $\varrho \in C^2(a, b)$  a weight function. How can a polynomial  $R_n$  of lowest possible degree satisfying the conditions

$$R_n(\xi_{i,n}) = \alpha_{i,n}, \quad (\varrho R_n)''(\xi_{i,n}) = \beta_{i,n} \quad (i = 1, 2, \dots, n),$$

be determined where  $\alpha_{i,n}$  and  $\beta_{i,n}$  are arbitrary given real numbers? J. BALÁZS [10] was the first who investigated this so called weighted (0, 2) interpolation problem. He proved that if the fundamental points are the roots of the ultraspherical polynomial  $P_n^{(\alpha)}$  ( $\alpha > -1$ ), and the weight function is  $\varrho(x) = (1 - x^2)^{(\alpha+1)/2}$  ( $-1 \leq x \leq 1$ ), then generally there does not exist any

polynomial of degree  $\leq 2n - 1$  satisfying the requirements. But he could show that under the condition

$$R_n(0) = \sum_{i=1}^n \alpha_{i,n} l_{i,n}^2(0)$$

there exists a unique polynomial of degree  $\leq 2n$ . (If  $n$  is odd then the uniqueness is not true.) He gave the explicit form of this polynomial and proved convergence theorem. His result was generalized and sharpened for Jacobi polynomials  $p_n^{(\alpha,\beta)}$  with any  $\alpha, \beta > -1$  in [15].

In [5] L. SZILI investigated some analogous problems in that case when the fundamental points are the roots of Hermite polynomials and the weight function is

$$W(x) = e^{-x^2/2}.$$

He proved the following convergence theorem ([5], Theorem 4):

If the interpolated function  $f : \mathbb{R} \rightarrow \mathbb{R}$  is continuously differentiable,  $\lim_{|x| \rightarrow \infty} x^{2r} W(x) f(x) = 0$  ( $r = 0, 1, \dots$ ) and  $\lim_{|x| \rightarrow \infty} W(x) f'(x) = 0$ , furthermore

$$y_{i,n} = f(x_{i,n}), \quad y''_{i,n} = O\left(e^{\beta x_{i,n}^2} \sqrt{n} \cdot \omega\left(f', \frac{1}{\sqrt{n}}\right)\right) \\ \left(i = 1, 2, \dots, n; \quad 0 \leq \beta < \frac{1}{2}\right),$$

then the weighted  $(0, 2)$  interpolation polynomials  $R_n$  ( $i = 2, 4, \dots$ ) satisfy the following estimate:

$$e^{-\gamma x^2} |f(x) - R_n(x)| = O\left(\log n \omega\left(f', \frac{1}{\sqrt{n}}\right)\right),$$

here  $\omega(f, \delta)$  the so-called Freud modulus;  $\gamma > 1$ ,  $O$  does not depend on  $n$  and  $x$ , the estimate holds on the whole real line.

In this paper we prove that if  $\beta = \frac{1}{2}$  then the following estimate is true:

**THEOREM.**

$$e^{-x^2} |f(x) - R_n(x)| = O(1) \omega\left(f', \frac{1}{\sqrt{n}}\right) + O(1) \frac{1}{\sqrt{n}},$$

where  $O(1)$  does not depend on  $n, x$  and  $f$ .

This estimate is not refinable (best possible) in some sense.

For the proof we need some lemas.

We shall use the following notations:

$$A_{i,n}(x) := l_{i,n}^2(x) + \frac{H_n(x)}{H_n'(x_{i,n})} \int_0^x \frac{l_{i,n}(t)(a_{i,n}t + b_{i,n}) - l_{i,n}'(t)}{t - x_{i,n}} dt$$

where

$$a_{i,n}x_{i,n} + b_{i,n} = l_{i,n}'(x_{i,n}) = x_{i,n}, \quad a_{i,n} = -\frac{W''(x_{i,n})}{W(x_{i,n})} = 1 - x_{i,n}^2,$$

and

$$B_{i,n}(x) := \frac{H_n(x)}{2W(x_{i,n})H_n'(x_{i,n})} \int_0^x l_{i,n}(t) dt \quad (i = 1, 2, \dots, n),$$

$$l_{i,n}(x) := \frac{H_n(x)}{H_n'(x_{i,n})(x - x_{i,n})} \quad (i = 1, 2, \dots, n),$$

the roots of  $H_n$  satisfy the following relations

$$-\infty < x_{n,n} < \dots < x_{\frac{n}{2}+1,n} < 0 < x_{\frac{n}{2},n} < \dots < x_{1,n} < +\infty \quad (n = 2n),$$

$$-\infty < x_{n,n} < \dots < x_{\frac{n+1}{2}+1,n} = 0 < \dots < x_{1,n} < +\infty \quad (n = 2m + 1),$$

$$x_{i,n} = -x_{n-i+1,n} \quad \left( i = 1, 2, \dots, \left[ \frac{n}{2} \right] \right).$$

Then (see [5], Theorem 2) for even  $n$

$$R_n(x) = \sum_{i=1}^n y_{i,n} A_{i,n}(x) + \sum_{i=1}^n y_{i,n}'' B_{i,n}(x)$$

are the uniquely determined polynomials of degree  $\leq 2n$  satisfying the following requirements:

$$R_n(x_{i,n}) = y_{i,n}, \quad (W \cdot R_n)''(x_{i,n}) = y_{i,n}'' \quad (i = 1, 2, \dots, n),$$

$$R_n(0) = \sum_{i=1}^n y_{i,n} l_{i,n}^2(0).$$

We shall use the following notations:  $x_i := x_{i,n}$ ,  $l_i := l_{i,n}$ ,  $A_i := A_{i,n}$ ,  $B_i := B_{i,n}$ .

LEMMA 1 ([14], Lemma 1). *If  $\lambda_i$  ( $i = 1, \dots, n$ ) are the Christoffel-numbers on Hermite nodes, then*

$$\lambda_i \asymp e^{-x_i^2} \cdot \frac{1}{n^{1/6} \cdot i^{1/3}} \asymp e^{-x_i^2} \cdot \varphi_n(x_i), \quad i = 1, \dots, \frac{n}{2},$$

where  $\varphi_n(x_i) = x_i - x_{i+1}$  (certainly  $x_1 > x_2 > \dots > x_n$ ),  $\lambda_i = \lambda_{n-i+1}$ .

LEMMA 2. Let  $n$  be even, then

$$\sup_{x \in \mathbb{R}} e^{-x^2} \sum_{i=1}^n e^{x_i^2/2} |B_i(x)| \asymp \frac{1}{\sqrt{n}}.$$

PROOF. Without loss of generality we can assume that  $x \geq 0$ . Using Lemma 1 we get

$$\begin{aligned} |H'_n(x_i)| &= 2n |H_{n-1}(x_i)| = 2n \cdot \pi^{1/4} \cdot \sqrt{2^{n-1}!(n-1)!} \cdot |h_{n-1}(x_i)| = \\ &= \pi^{1/4} \cdot \sqrt{2} \cdot \sqrt{2^n n!} \cdot \lambda_i^{-1/2} \asymp c \cdot e^{x_i^2/2} \cdot \sqrt{2^n n!} \cdot \varphi_n^{-1/2}(x_i), \quad i = 1, \dots, n. \end{aligned}$$

We have

$$\begin{aligned} |B_i(x)| &\asymp |H_n(x)| \cdot e^{x_i^2/2} \cdot \frac{1}{|H'_n(x_i)|} \cdot \left| \int_0^x l_i(t) dt \right| = \\ &= |H_n(x)| \cdot \frac{e^{x_i^2/2}}{(H'_n(x_i))^2} \cdot \left| \int_0^x H_n(t) \frac{1}{t-x_i} dt \right|. \end{aligned}$$

$$\text{We need an estimate for } \left| \int_0^x H_n(t) \frac{1}{t-x_i} dt \right|.$$

We investigate the cases: 1.  $x_i \leq 0$ , 2.  $0 \leq x_i \leq \sqrt{n}$ , 3.  $\sqrt{n} \leq x_i \leq \sqrt{2n+1}$ .

In what follows we will use many times the result of [2], p. 700, table.

1.  $x_i \leq 0$ .

Integrating by parts and using that  $H_{n+1}(0) = 0$  because of  $n$  is even we obtain

$$\begin{aligned} (1) \quad \int_0^x H_n(t) \cdot \frac{1}{t-x_i} dt &= \\ &= \frac{1}{2(n+1)} H_{n+1}(x) \cdot \frac{1}{x-x_i} + \frac{1}{2(n+1)} \int_0^x H_{n+1}(t) \cdot \frac{1}{(t-x_i)^2} dt. \end{aligned}$$

From this we get

$$\int_0^x H_n(t) \cdot \frac{1}{t-x_i} dt = O(1) \frac{1}{n} e^{x^2/2} \frac{\sqrt{2^n(n+1)!}}{n^{1/12}} \cdot \frac{1}{|x_i|} = O(1) e^{x^2/2} \frac{\sqrt{2^n n!}}{n^{7/12}} \cdot \frac{1}{|x_i|}.$$

Hence

$$\sum_{\substack{i=1 \\ x_i \leq 0}}^n e^{x_i^2/2} |B_i(x)| = O(1)e^{x^2} \frac{1}{n^{2/3}} \sum_{\substack{i=1 \\ x_i \leq 0}}^n \frac{\varphi_n(x_i)}{|x_i|} = O(1)e^{x^2} \frac{1}{n^{2/3}} \sum_{\substack{i=1 \\ -\sqrt{n} \leq x_i \leq 0}}^n \frac{1}{\sqrt{n}} \cdot \frac{1}{|x_i|} +$$

$$(2) \quad + O(1)e^{x^2} \frac{1}{n^{2/3}} \sum_{\substack{i=1 \\ x_i \leq -\sqrt{n}}}^n \frac{1}{\sqrt{n}} \cdot \frac{1}{n^{1/6} \cdot i^{1/3}} = O(1)e^{x^2} \frac{\log n}{n^{2/3}},$$

where we used  $|H_n(x)| = O(1)e^{x^2/2} \sqrt{2^n n!} / n^{1/12}$  and  $x_i \asymp \frac{\frac{n}{2} - i + 1}{\sqrt{n}}$  if  $-\sqrt{n} \leq x_i \leq 0$ .

$$2. \quad 0 \leq x_i \leq \sqrt{n}.$$

We distinguish 6 cases: a)  $0 \leq x \leq x_i - 1$ , b)  $x_i - 1 \leq x \leq x_i - \frac{1}{\sqrt{n}}$ , c)  $x_i - \frac{1}{\sqrt{n}} \leq x \leq x_i + \frac{1}{\sqrt{n}}$ , d)  $x_i + \frac{1}{\sqrt{n}} \leq x \leq x_i + 1$ , e)  $x_i + 1 \leq x \leq (\sqrt{2} - 0.1)\sqrt{n}$ , f)  $x \geq (\sqrt{2} - 0.1)\sqrt{n}$ .

a)  $0 \leq x \leq x_i - 1$ .

In this case from (1)

$$\int_0^x H_n(t) \frac{1}{t - x_i} dt = O(1)e^{x^2/2} \frac{\sqrt{2^n n!}}{\sqrt{n} \cdot n^{1/4}}.$$

Hence

$$(3) \quad \sum_{\substack{i=1 \\ 0 \leq x_i \leq \sqrt{n} \\ 0 \leq x \leq x_i - 1}}^n e^{x_i^2/2} |B_i(x)| = O(1)e^{x^2} \cdot \frac{1}{\sqrt{n}}.$$

b)  $x_i - 1 \leq x \leq x_i - \frac{1}{\sqrt{n}}$ .

In this case

$$\int_0^x H_n(t) \frac{1}{t - x_i} dt = \int_0^{x_i - 1} + \int_{x_i - 1}^x = O(1)e^{(x_i - 1)^2/2} \frac{\sqrt{2^n n!}}{\sqrt{n} n^{1/4}} + O(1)e^{x^2/2} \frac{\sqrt{2^n n!}}{n^{1/4}}.$$

Hence

$$(4) \quad \int_0^x H_n(t) \frac{1}{t - x_i} dt = O(1)e^{x^2/2} \frac{\sqrt{2^n n!}}{n^{1/4}}$$

and

$$(5) \quad \sum_{\substack{i=1 \\ 0 \leq x_i \leq \sqrt{n} \\ x_i - 1 \leq x \leq x_i - \frac{1}{\sqrt{n}}}}^n e^{\frac{1}{2}x_i^2} |B_i(x)| = O(1)e^{x^2} \cdot \frac{1}{\sqrt{n}}.$$

$$c) \quad x_i - \frac{1}{\sqrt{n}} \leq x \leq x_i + \frac{1}{\sqrt{n}}.$$

In this case

$$\begin{aligned} & \int_0^x H_n(t) \frac{1}{t - x_i} dt = \\ &= \int_0^{x_i - \frac{1}{\sqrt{n}}} + \int_{x_i - \frac{1}{\sqrt{n}}}^x = O(1)e^{\left(x_i - \frac{1}{\sqrt{n}}\right)^2 / 2} \frac{\sqrt{2^n n!}}{n^{1/4}} + \int_{x_i - \frac{1}{\sqrt{n}}}^x \frac{H_n(t) - H_n(x_i)}{t - x_i} dt. \end{aligned}$$

Since  $H_n'(x) = 2nH_{n-1}(x)$ , therefore using the mean-value theorem we obtain

$$(6) \quad \int_0^x H_n(t) \frac{1}{t - x_i} dt = O(1)e^{x^2/2} \frac{\sqrt{2^n n!}}{n^{1/4}}.$$

Hence

$$(7) \quad \sum_{\substack{i=1 \\ 0 \leq x_i \leq \sqrt{n} \\ x_i - \frac{1}{\sqrt{n}} \leq x \leq x_i + \frac{1}{\sqrt{n}}}}^n e^{x_i^2/2} |B_i(x)| = O(1)e^{x^2} \cdot \frac{1}{\sqrt{n}}.$$

$$d) \quad x_i + \frac{1}{\sqrt{n}} \leq x \leq x_i + 1.$$

In this case

$$\int_0^x H_n(t) \frac{1}{t - x_i} dt = \int_0^{x_i + \frac{1}{\sqrt{n}}} + \int_{x_i + \frac{1}{\sqrt{n}}}^x = O(1)e^{\left(x_i + \frac{1}{\sqrt{n}}\right)^2 / 2} \frac{\sqrt{2^n n!}}{n^{1/4}} + O(1)e^{x^2/2} \frac{\sqrt{2^n n!}}{n^{1/4}}.$$

Hence

$$(8) \quad \int_0^x H_n(t) \frac{1}{t - x_i} dt = O(1)e^{x^2/2} \frac{\sqrt{2^n n!}}{n^{1/4}}$$

and

$$(9) \quad \sum_{\substack{i=1 \\ 0 \leq x_i \leq \sqrt{n} \\ x_i - \frac{1}{\sqrt{n}} \leq x \leq x_{i+1}}}^n e^{x_i^2/2} |B_i(x)| = O(1)e^{x^2} \cdot \frac{1}{\sqrt{n}}.$$

e)  $x_{i+1} + 1 \leq x \leq (\sqrt{2} - 0.1)\sqrt{n}$ .

In this case

$$\int_0^x H_n(t) \frac{1}{t - x_i} dt = \int_0^{x_i + \frac{1}{2}} + \int_{x_i + \frac{1}{2}}^x = O(1)e^{(x_i + \frac{1}{2})^2/2} \frac{\sqrt{2^n n!}}{n^{1/4}} + O(1)e^{x^2/2} \frac{\sqrt{2^n n!}}{\sqrt{n} \cdot n^{1/4}}.$$

Therefore

$$\sum_{\substack{i=1 \\ 0 \leq x_i \leq \sqrt{n} \\ x_i + 1 \leq x \leq (\sqrt{2} - 0.1)\sqrt{n}}}^n e^{\frac{1}{2}x_i^2} |B_i(x)| = O(1)e^{x^2/2} \cdot \frac{1}{n} \cdot \sum_{\substack{i=1 \\ 0 \leq x_i \leq \sqrt{n} \\ x_i + 1 \leq x \leq (\sqrt{2} - 0.1)\sqrt{n}}}^n e^{(x_i + \frac{1}{2})^2/2} + O(1)e^{x^2} \cdot \frac{1}{\sqrt{n}}.$$

Here

$$\frac{1}{4} \sum_{\substack{i=1 \\ 0 \leq x_i \leq \sqrt{n} \\ x_i + 1 \leq x}}^n e^{(x_i + \frac{1}{2})^2/2} \leq c\sqrt{n} \int_0^{x - \frac{1}{4}} e^{t^2/2} dt = O(1)\sqrt{n}e^{x^2/2},$$

where we used  $x_i - x_{i+1} \asymp \frac{1}{\sqrt{n}}$  ( $0 \leq x_i \leq \sqrt{n}$ ) and  $\int_0^x e^{t^2/2} dt = O(1) \frac{e^{x^2/2}}{1+|x|}$ .

Hence

$$(10) \quad \sum_{\substack{i=1 \\ 0 \leq x_i \leq \sqrt{n} \\ x_i + 1 \leq x \leq (\sqrt{2} - 0.1)\sqrt{n}}}^n e^{\frac{1}{2}x_i^2} |B_i(x)| = O(1)e^{x^2} \cdot \frac{1}{\sqrt{n}}.$$

f)  $x \geq (\sqrt{2} - 0.1)\sqrt{n}$ .

In this case

$$\int_0^x H_n(t) \frac{1}{t - x_i} dt =$$

$$= \int_0^{(\sqrt{2}-0.1)\sqrt{n}} + \int_{(\sqrt{2}-0.1)\sqrt{n}}^x = O(1)e^{((\sqrt{2}-0.2)\sqrt{n})^2/2} \frac{\sqrt{2^n n!}}{\sqrt{n} \cdot n^{1/4}} + O(1)e^{x^2/2} \frac{\sqrt{2^n n!}}{n \cdot n^{1/12}}.$$

Hence

$$\sum_{\substack{i=1 \\ 0 \leq x_i \leq \sqrt{n} \\ x \geq (\sqrt{2}-0.1)\sqrt{n}}}^n e^{\frac{1}{2}x_i^2} |B_i(x)| = O(1)e^{x^2} \frac{1}{\sqrt{n}}.$$

From (3), (5), (7), (9), (10) and (11) we obtain

$$(12) \quad \sum_{\substack{i=1 \\ 0 \leq x_i \leq \sqrt{n}}}^n e^{\frac{1}{2}x_i^2} |B_i(x)| = O(1)e^{x^2} \frac{1}{\sqrt{n}}.$$

3.  $\sqrt{n} \leq x_i \leq \sqrt{2n+1}$ .

For the sake of definiteness we choose an  $i_1$  index, such that  $\sqrt{n} - \frac{1}{\sqrt{n}} \leq x_{i_1} \leq \sqrt{n}$ . We distinguish 3 cases: a)  $x \geq x_1 + \frac{1}{n^{1/6}}$ , b)  $x_1 + \frac{1}{n^{1/6}} \geq x \geq x_{i_1} - \frac{1}{\sqrt{n}}$ , c)  $x \leq x_{i_1} - \frac{1}{\sqrt{n}}$ .

a)  $x \geq x_1 + \frac{1}{n^{1/6}}$ . We have to estimate the sum

$$(13) \quad S_1 := \sum_{\substack{i=1 \\ \sqrt{n} \leq x_i \leq \sqrt{2n+1}}}^n e^{\frac{1}{2}x_i^2} |B_i(x)| =$$

$$= \sum_{\substack{i=1 \\ \sqrt{n} \leq x_i \leq \sqrt{2n+1} - n^{1/3}}}^n + \sum_{\substack{i=1 \\ \sqrt{2n+1} - n^{1/3} \leq x_i \leq \sqrt{2n+1} - n}}^n +$$

$$+ \sum_{i=1}^n \quad + \sum_{i=1}^n \quad =: S_{11} + S_{12} + S_{13} + S_{14}.$$

$$\sqrt{2n+1}-n^{1/6} \leq x_i \leq \sqrt{2n+1}-1 \quad \sqrt{2n+1}-1 \leq x_i \leq \sqrt{2n+1}$$

The estimation on  $S_{11}$ :

In this case

$$\int_0^x H_n(t) \frac{1}{t-x_i} dt = \int_0^{x_i-\frac{1}{\sqrt{n}}} + \int_{x_i-\frac{1}{\sqrt{n}}}^{x_i+\frac{1}{\sqrt{n}}} + \int_{x_i+\frac{1}{\sqrt{n}}}^{x_i+\frac{1}{2}n^{1/3}} + \int_{x_i+\frac{1}{2}n^{1/3}}^x =: A_1 + A_2 + A_3 + A_4.$$

Using (1) we obtain

$$A_4 = \int_{x_i+\frac{1}{2}n^{1/3}}^x H_n(t) \frac{1}{t-x_i} dt = O(1)e^{x^2/2} \frac{\sqrt{2^n n!}}{\sqrt{n} \cdot n^{1/12}} \cdot \frac{1}{n^{1/3}}.$$

Since  $x \geq x_i + \frac{3}{4}n^{1/3}$  therefore using (1) in  $A_1, A_3$  and using the mean-value theorem in  $A_3$  we obtain

$$A_1, A_2, A_3 = O(1)e^{x^2/2} \frac{\sqrt{2^n n!}}{\sqrt{n} \cdot n^{1/12}} \cdot \frac{1}{n^{1/3}}.$$

Hence

$$(14) \quad S_{11} = \sum_{i=1}^n e^{\frac{1}{2}x_i^2} |B_i(x)| = O(1)e^{x^2} \frac{1}{n} \sum_{i=1}^n \varphi_n(x_i) = O(1)e^{x^2} \frac{1}{\sqrt{n}}.$$

$$\sqrt{n}=x_i \leq \sqrt{2n+1}-n^{1/3} \quad \sqrt{n} \leq x_i \leq \sqrt{2n+1}-n^{1/3}$$

The estimation of  $S_{12}$ :

Similarly as in the estimation of  $S_{11}$  we obtain

$$\int_0^x H_n(t) \frac{1}{t-x_i} dt = O(1)e^{x^2/2} \frac{\sqrt{2^n n!}}{\sqrt{n} \cdot n^{1/12}} \cdot \frac{1}{n^{1/6}}.$$

Hence

$$(15) \quad S_{12} = \sum_{i=1}^n e^{\frac{1}{2}x_i^2} |B_i(x)| =$$

$$\sqrt{2n+1}-n^{1/3} \leq x_i \leq \sqrt{2n+1}-n^{1/6}$$

$$= O(1)e^{x^2} \frac{1}{n^{5/6}} \sum_{i=1}^n \varphi_n(x_i) = O(1)e^{x^2} \frac{1}{\sqrt{n}}.$$

$$\sqrt{2n+1}-n^{1/3} \leq x_i \leq \sqrt{2n+1}-n^{1/6}$$

The estimation of  $S_{13}$ :

Similarly as in the estimation of  $S_{11}$  we obtain

$$\int_0^x H_n(t) \frac{1}{t - x_i} dt = O(1)e^{x^2/2} \frac{\sqrt{2^n n!}}{\sqrt{n} \cdot n^{1/12}}.$$

Hence

$$\begin{aligned} (16) \quad & \sum_{i=1}^n e^{\frac{1}{2}x_i^2} |B_i(x)| = \\ & \sum_{\sqrt{2n+1}-n^{1/6} \leq x_i \leq \sqrt{2n+1}-1} O(1)e^{x^2} \frac{1}{n^{2/3}} \sum_{i=1}^n \varphi_n(x_i) = O(1)e^{x^2} \frac{1}{\sqrt{n}}. \end{aligned}$$

The estimation of  $S_{14}$ :

We know  $\sqrt{2n+1} - x_1 \asymp \frac{1}{n^{1/6}}$ . Similarly as in the estimation of  $S_{11}$  we obtain

$$\int_0^x H_n(t) \frac{1}{t - x_i} dt = O(1)e^{x^2/2} \frac{\sqrt{2^n n!}}{\sqrt{n} \cdot n^{1/12}} \cdot n^{1/6}.$$

Hence

$$\begin{aligned} (17) \quad & \sum_{i=1}^n e^{\frac{1}{2}x_i^2} |B_i(x)| = \\ & \sum_{\sqrt{2n+1}-1 \leq x_i \leq \sqrt{2n+1}} O(1)e^{x^2} \frac{1}{\sqrt{n}} \sum_{i=1}^n \varphi_n(x_i) = O(1)e^{x^2} \frac{1}{\sqrt{n}}. \end{aligned}$$

From (13)–(17) we obtain

$$(18) \quad \sum_{\substack{i=1 \\ \sqrt{n} \leq x_i \leq \sqrt{2n+1} \\ x \geq x_1 + \frac{1}{n^{1/6}}} }^n e^{\frac{1}{2}x_i^2} |B_i(x)| = O(1)e^{x^2} \frac{1}{\sqrt{n}}.$$

b)  $x_1 + \frac{1}{n^{1/6}} \geq x \geq x_{i_1} - \frac{1}{\sqrt{n}}.$

Denote  $i_0$  the index ( $1 \leq i_0 \leq i_1$ ) for which  $|x - x_i|$  is minimal. First we estimate the sum

$$(19) \quad T_1 := \sum_{i=1}^{i_0-1} e^{\frac{1}{2}x_i^2} |B_i(x)|.$$

Using (1) we obtain

$$(20) \quad e^{\frac{1}{2}x_i^2} |B_i(x)| = O(1) \frac{|H_n(x_i)|}{2^n n!} \varphi_n(x_i) \frac{1}{n} |H_{n+1}(x)| \frac{1}{|x - x_i|} +$$

$$+ O(1) \frac{|H_n(x)|}{2^n n!} \varphi_n(x_i) \frac{1}{n} \cdot \left| \int_0^x H_{n+1}(t) \frac{1}{(t - x_i)^2} dt \right| = O(1) e^{x^2} \frac{1}{n^{2/3}} \frac{\varphi_n(x_i)}{|x - x_i|},$$

where we used  $|H_n(x)| = O(1) e^{x^2/2} \sqrt{2^n n!} / n^{1/2}$ . Using Lemma 1 we obtain

$$|x - x_i| \geq c |x_{i_0} - x_i| \asymp n^{-\frac{1}{6}} \left( i_0^{-\frac{1}{3}} + \dots + i^{-\frac{1}{3}} \right) \asymp \frac{|i^{2/3} - i_0^{2/3}|}{n^{1/6}}.$$

Hence

$$(21) \quad T_1 = O(1) e^{x^2} \frac{1}{n^{2/3}} \sum_{i=1}^{i_0-1} \frac{i^{-1/3}}{|i^{2/3} - i_0^{2/3}|} =$$

$$= O(1) e^{x^2} \frac{1}{n^{2/3}} \sum_{i=1}^{i_0/2} \frac{i^{-1/3}}{i_0^{2/3}} + O(1) e^{x^2} \frac{1}{n^{2/3}} \sum_{i=i_0/2}^{i_0-1} \frac{i_0^{-1/3}}{i_0^{-1/3} |i - i_0|} = O(1) e^{x^2} \frac{\log n}{n^{2/3}}.$$

Now we have to estimate the sum

$$(22) \quad T_2 := \sum_{\substack{i=i_0 \\ \sqrt{n} \leq x_i \leq \sqrt{2n+1}}}^n e^{\frac{1}{2}x_i^2} |B_i(x)|.$$

If  $\varphi_n(x_{i_0}) \leq \frac{1}{n^{1/3}}$  then we can write

$$T_2 = \sum_{\substack{i=i_0 \\ \sqrt{n} \leq x_i \leq \sqrt{2n+1} \\ x_i \leq x - n^{1/3}}} + \sum_{\substack{i=i_0 \\ \sqrt{n} \leq x_i \leq \sqrt{2n+1} \\ x - n^{1/3} \leq x_i \leq x - n^{1/6}}} + \sum_{\substack{i=i_0 \\ \sqrt{n} \leq x_i \leq \sqrt{2n+1} \\ x - n^{1/6} \leq x_i \leq x - 1}} +$$

$$+ \sum_{i=i_0}^n + \sum_{i=i_0}^n + \sum_{i=i_0}^n + e^{\frac{1}{2}x_{i_0}^2} |B_{i_0}(x)| =:$$

$$\begin{matrix} \sqrt{n} \leq x_i \leq \sqrt{2n+1} & \sqrt{n} \leq x_i \leq \sqrt{2n+1} & \sqrt{n} \leq x_i \leq \sqrt{2n+1} \\ x-1 \leq x_i \leq x-\frac{1}{n^{1/6}} & x-\frac{1}{n^{1/6}} \leq x_i \leq x-\frac{1}{n^{1/3}} & x-\frac{1}{n^{1/3}} \leq x_i \leq x-\frac{1}{\sqrt{n}} \end{matrix}$$

(23)  $=: T_{21} + T_{22} + T_{23} + T_{24} + T_{25} + T_{26} + T_{27}.$

If  $\varphi_n(x_{i_0}) \geq \frac{1}{n^{1/3}}$  then we can write

(24)  $T_2 = T_{21} + T_{22} + T_{23} + T_{24} + T_{25} + T_{27}.$

The estimation of  $T_{21}$ :

In this case

$$\int_0^x H_n(t) \frac{1}{t-x_i} dt = \int_0^{x_i-\frac{1}{\sqrt{n}}} + \int_{x_i-\frac{1}{\sqrt{n}}}^{x_i+\frac{1}{\sqrt{n}}} + \int_{x_i+\frac{1}{\sqrt{n}}}^{x_i+\frac{1}{2}n^{1/3}} + \int_{x_i+\frac{1}{2}n^{1/3}}^x .$$

Same as in the estimation of  $S_{11}$ , we obtain

$$\int_0^x H_n(t) \frac{1}{t-x_i} dt = O(1)e^{x^2/2} \frac{\sqrt{2^n n!}}{\sqrt{n} \cdot n^{1/12}} \cdot \frac{1}{n^{1/3}}$$

and

(25)  $T_{21} = \sum_{i=i_0}^n e^{\frac{1}{2}x_i^2} |B_i(x)| = O(1)e^{x^2} \frac{1}{\sqrt{n}}.$

$$\begin{matrix} \sqrt{n} \leq x_i \leq \sqrt{2n+1} \\ x_i \leq x-n^{1/3} \end{matrix}$$

The estimation of  $T_{22}, T_{23}, T_{24}, T_{25}$  are the same.

The estimation of  $T_{26}$ :

In this case

$$\int_0^x H_n(t) \frac{1}{t-x_i} dt = \int_0^{x_i-\frac{1}{\sqrt{n}}} + \int_{x_i-\frac{1}{\sqrt{n}}}^{x_i+\frac{1}{2}\frac{1}{\sqrt{n}}} + \int_{x_i+\frac{1}{2}\frac{1}{\sqrt{n}}}^x =: D_1 + D_2 + D_3.$$

Using (1) we obtain

$$D_1, D_3 = O(1)e^{x^2/2} \frac{\sqrt{2^n n!}}{\sqrt{n} \cdot n^{1/12}} \cdot \sqrt{n}.$$

Using the mean-value theorem

$$D_2 = O(1)e^{x^2/2} \frac{\sqrt{2^n n!}}{n^{1/12}}.$$

Hence

$$(26) \quad T_{26} = \sum_{i=i_0} e^{\frac{1}{2}x_i^2} |B_i(x)| = O(1)e^{x^2} \frac{1}{\sqrt{n}}.$$

$$\sqrt{n} \leq x_i \leq \sqrt{2n+1}$$

$$x - \frac{1}{n^{1/3}} \leq x_i \leq x - \frac{1}{\sqrt{n}}$$

The estimation of  $T_{27}$ :

$$T_{27} = e^{\frac{1}{2}x_{i_0}^2} |B_{i_0}(x)| = O(1)e^{x^2/2} \frac{\sqrt{2^n n!}}{n^{1/12}} \cdot \frac{e^{x_{i_0}^2/2}}{|H_n'(x_{i_0})|} \int_0^x e^{x_{i_0}^2/2} |l_{i_0}(t)| dt =$$

$$= O(1)e^{x^2/2} \frac{1}{n^{1/12}} \cdot \sqrt{\varphi_n(x_{i_0})} \int_0^x e^{x_{i_0}^2/2} |l_{i_0}(t)| dt.$$

Using [6], (2.36) we get

$$e^{x_{i_0}^2} l_{i_0}^2(t) \leq \sum_{i=1}^n e^{x_i^2} l_i^2(t) = O(1)e^{t^2}.$$

Therefore

$$e^{\frac{1}{2}x_{i_0}^2} |l_{i_0}(t)| = O(1)e^{t^2/2}.$$

Thus

$$(27) \quad T_{27} = O(1)e^{x^2/2} \frac{1}{n^{1/6}} \int_0^x e^{t^2/2} dt = O(1)e^{x^2} \frac{1}{\sqrt{n}}.$$

Hence from (21)–(27) we obtain

$$\sum_{i=1}^n e^{\frac{1}{2}x_i^2} |B_i(x)| = O(1)e^{x^2} \frac{1}{\sqrt{n}}.$$

$$\sqrt{n} \leq x_i \leq \sqrt{2n+1}$$

$$x_{i_1} - \frac{1}{\sqrt{n}} \leq x \leq x_1 + \frac{1}{n^{1/6}}$$

$$c) \quad x \leq x_{i_1} - \frac{1}{\sqrt{n}}.$$

In this case  $|x - x_i| \geq c|x_{i_1-1} - x_{i_1}|$  and we can repeat the method of the estimation of  $T_1$  in the case b). Hence

$$(29) \quad \sum_{i=1}^n e^{\frac{1}{2}x_i^2} |B_i(x)| = O(1)e^{x^2} \frac{1}{\sqrt{n}}.$$

$$\sqrt{n} \leq x_i \leq \sqrt{2n+1}$$

$$x \leq x_{i_1} - \frac{1}{\sqrt{n}}$$

From (18), (28) and (29)

$$(30) \quad \sum_{i=1}^n e^{\frac{1}{2}x_i^2} |B_i(x)| = O(1)e^{x^2} \frac{1}{\sqrt{n}}.$$

$$\sqrt{n} \leq x_i \leq \sqrt{2n+1}$$

From (2), (12) and (30) the upper estimate follows.

Now we prove the lower estimate.

Let  $3 \leq x^* \leq 4$  be such that  $|H_n(x^*)| \asymp e^{(x^*)^2} \cdot \frac{\sqrt{2^n n!}}{n^{1/4}}$ . We will prove that

$$(31) \quad \sum_{i=1}^n e^{x_i^2/2} |B_i(x^*)| \asymp \frac{e^{(x^*)^2}}{\sqrt{n}}.$$

$$1 \leq x_i \leq 2$$

In this case

$$\int_0^{x^*} H_n(t) \frac{1}{t - x_i} dt = \int_0^{x_i - \frac{1}{n^{1/3}}} + \int_{x_i - \frac{1}{n^{1/3}}}^{x_i + \frac{1}{n^{1/3}}} + \int_{x_i + \frac{1}{n^{1/3}}}^{x^*} =: I_1 + I_2 + I_3.$$

Using (1) we obtain

$$(32) \quad I_1, I_3 = O(1)e^{(x^*)^2/2} \frac{\sqrt{2^n n!}}{\sqrt{n} \cdot n^{1/4}} \cdot n^{1/3}.$$

For the estimate of  $I_2$  we need a sharp estimation. Obviously

$$(33) \quad I_2 = \int_{x_i - \frac{1}{n^{1/3}}}^{x_i + \frac{1}{n^{1/3}}} H_n(t) \frac{1}{t - x_i} dt = \int_{x_i - \frac{1}{n^{1/3}}}^{x_i + \frac{1}{n^{1/3}}} \frac{H_n(t) - H_n(x_i)}{t - x_i} dt =$$

$$= \int_{x_i - \frac{1}{n^{1/3}}}^{x_i + \frac{1}{n^{1/3}}} \frac{x_i \int_{x_i}^t H'_n(z) dz}{t - x_i} dt = 2n \int_{x_i - \frac{1}{n^{1/3}}}^{x_i + \frac{1}{n^{1/3}}} \frac{x_i \int_{x_i}^t H_{n-1}(z) dz}{t - x_i} dt.$$

We know ([4], (8.22.8)) that

$$\Gamma\left(\frac{n}{2} + 1\right) \Gamma(n + 1) e^{-x^2/2} H_n(x) = \cos\left(\sqrt{2n+1}x - \frac{n\pi}{2}\right) + O\left(\frac{1}{\sqrt{n}}\right),$$

( $|x| \leq c_0$ ), where the error term is uniform in  $x$ . From this we get

$$(34) \quad H_n(x) = e^{x^2/2} \left(\frac{2}{\pi}\right)^{1/4} \sqrt{2^n n!} \cdot \left(1 + O\left(\frac{1}{n}\right)\right) \frac{1}{n^{1/4}} \left(\cos\left(\sqrt{2n+1}x - \frac{n\pi}{2}\right) + O\left(\frac{1}{\sqrt{n}}\right)\right).$$

Since  $n$  is even,  $n = 2k$ , therefore

$$H_{n-1}(z) = e^{z^2/2} \left(\frac{2}{\pi}\right)^{1/4} \sqrt{2^{n-1}(n-1)!} \left(\frac{(-1)^k}{(n-1)^{1/4}} \sin(\sqrt{2n-1}z) + O\left(n^{3/4}\right)\right).$$

The remainder term gives in (33)

$$(35) \quad O(1)e^{(x^*)^2/2} \frac{\sqrt{2^n n!}}{n^{1/4+1/3}}.$$

Therefore it is enough to estimate

$$(-1)^{k+1} \left(\frac{2}{\pi}\right)^{1/4} \frac{\sqrt{2^{n-1}(n-1)!}}{(n-1)^{1/4}} 2n \int_{x_i - \frac{1}{n^{1/3}}}^{x_i + \frac{1}{n^{1/3}}} \frac{x_i \int_{x_i}^t e^{z^2/2} \sin(\sqrt{2n-1}z) dz}{t - x_i} dt.$$

Here

$$e^{z^2/2} = e^{x_i^2/2} \cdot e^{z^2/2 - x_i^2/2} = e^{x_i^2/2} \left(1 + O(1)|z^2 - x_i^2|\right) = e^{x_i^2/2} \left(1 + O(1)n^{-1/3}\right).$$

The remainder term gives in (33)

$$(36) \quad O(1)e^{(x^*)^2/2} \frac{\sqrt{2^n n!}}{n^{1/4+1/6}}.$$

Therefore it is enough to estimate

$$(37) \quad (-1)^k \left(\frac{2}{\pi}\right)^{1/4} \frac{\sqrt{2n-1}(n-1)!}{(n-1)^{1/4}} 2n \frac{e^{x_i^2/2}}{\sqrt{2n-1}} \int_{x_i - \frac{1}{n^{1/3}}}^{x_i + \frac{1}{n^{1/3}}} \frac{\cos(\sqrt{2n-1}t) - \cos(\sqrt{2n-1}x_i)}{t - x_i} dt.$$

Here

$$\begin{aligned} \cos(\sqrt{2n-1}t) - \cos(\sqrt{2n-1}x_i) &= \\ &= -2 \sin \frac{\sqrt{2n-1}}{2}(t-x_i) \sin \frac{\sqrt{2n-1}}{2}(t+x_i), \\ \sin \frac{\sqrt{2n-1}}{2}(t+x_i) &= \sin \frac{\sqrt{2n-1}}{2}(t-x_i+2x_i) = \\ &= \sin \frac{\sqrt{2n-1}}{2}(t-x_i) \cos \sqrt{2n-1}x_i + \cos \frac{\sqrt{2n-1}}{2}(t-x_i) \sin \sqrt{2n-1}x_i. \end{aligned}$$

Hence

$$\begin{aligned} \cos(\sqrt{2n-1}t) - \cos(\sqrt{2n-1}x_i) &= \\ &= -2 \cos \sqrt{2n-1}x_i \left( \sin \frac{\sqrt{2n-1}}{2}(t-x_i) \right)^2 - \\ &\quad - \sin \sqrt{2n-1}x_i \sin \sqrt{2n-1}(t-x_i). \end{aligned}$$

Using this we have

$$\begin{aligned} &\int_{x_i - \frac{1}{n^{1/3}}}^{x_i + \frac{1}{n^{1/3}}} \frac{\cos(\sqrt{2n-1}t) - \cos(\sqrt{2n-1}x_i)}{t - x_i} dt = \\ &= -2 \cos \sqrt{2n-1}x_i \int_{x_i - \frac{1}{n^{1/3}}}^{x_i + \frac{1}{n^{1/3}}} \frac{\left( \sin \frac{\sqrt{2n-1}}{2}(t-x_i) \right)^2}{t - x_i} dt - \end{aligned}$$

$$\begin{aligned}
 & -\sin \sqrt{2n-1}x_i \int_{x_i - \frac{1}{n^{1/3}}}^{x_i + \frac{1}{n^{1/3}}} \frac{\sin \sqrt{2n-1}(t-x_i)}{t-x_i} dt = \\
 & = -2 \cos \sqrt{2n-1}x_i \int_{-\frac{1}{n^{1/3}}}^{\frac{1}{n^{1/3}}} \frac{\left(\sin \frac{\sqrt{2n-1}}{2}t\right)^2}{t} dt - \\
 & -\sin \sqrt{2n-1}x_i \int_{-\frac{1}{n^{1/3}}}^{\frac{1}{n^{1/3}}} \frac{\sin \sqrt{2n-1}t}{t} dt = -\sin \sqrt{2n-1}x_i \int_{-\frac{\sqrt{2n-1}}{n^{1/3}}}^{\frac{\sqrt{2n-1}}{n^{1/3}}} \frac{\sin u}{u} du = \\
 & = -\sin \sqrt{2n-1}x_i \int_{-\infty}^{\infty} \frac{\sin u}{u} du + O(1)\frac{1}{n^{1/6}} = -\pi \sin \sqrt{2n-1}x_i + O(1)\frac{1}{n^{1/6}},
 \end{aligned}$$

where we used

$$\int_a^{\infty} \frac{\sin u}{u} du = O(1)\frac{1}{a} \quad (a \geq 1).$$

The remainder term gives in (33)

$$(38) \quad O(1)e^{x_i^2/2} \frac{\sqrt{2^n n!}}{n^{1/4+1/6}}.$$

Therefore it is enough to estimate

$$(39) \quad (-1)^{k+1} \left(\frac{2}{\pi}\right) \frac{\sqrt{2^{n-1}(n-1)!}}{(n-1)^{1/4}} 2n \frac{e^{x_i^2/2}}{\sqrt{2n-1}} \cdot \pi \cdot \sin \sqrt{2n-1}x_i.$$

We know from [2], Theorem 1, (1) that

$$\bar{x}_\nu = \frac{\nu\pi - \frac{\pi}{2}}{\sqrt{2n+1}} + O\left(\frac{1}{n}\right),$$

if  $n$  is even,  $0 < \bar{x}_1 < \bar{x}_2 < \dots < \bar{x}_\nu$  the positive zeros of  $H_n(x)$  the  $n$ -th Hermite polynomial and  $\nu = O(1)\sqrt{n}$ .

Obviously we can write

$$\bar{x}_\nu = \frac{\nu\pi - \frac{\pi}{2}}{\sqrt{2n-1}} + O\left(\frac{1}{n}\right).$$

Since  $1 \leq x_i \leq 2$  therefore  $\nu \asymp \sqrt{n}$  and  $\sqrt{2n-1}\bar{x}_\nu = \nu\pi - \frac{\pi}{2} + O\left(\frac{1}{\sqrt{n}}\right)$ . Hence  $|\sin \sqrt{2n-1}x_i| \geq \frac{1}{2}$ , if  $n \geq n_0$ ,  $n_0$  is an absolute constant. Thus we obtain from (33)–(39)

$$(40) \quad I_2 \asymp e^{(x^*)^2/2} \frac{\sqrt{2^n n!}}{n^{1/4}}.$$

From (32) and (40)

$$\int_0^{x^*} H_n(t) \frac{1}{t-x_i} dt \asymp e^{(x^*)^2/2} \frac{\sqrt{2^n n!}}{n^{1/4}}.$$

Hence

$$\sum_{\substack{i=1 \\ 1 \leq x_i \leq 2}}^n e^{x_i^2/2} |B_i(x^*)| \asymp e^{(x^*)^2} \frac{1}{n^{1/4}} \cdot \frac{1}{n^{1/4}} \sum_{\substack{i=1 \\ 1 \leq x_i \leq 2}}^n \varphi_n(x_i) \asymp e^{(x^*)^2} \frac{1}{\sqrt{n}}.$$

The (31) is proved. The Lemma 2 is completely proved.  $\blacksquare$

LEMMA 3. ([5], Lemma 2). *The “first-kind” fundamental polynomials can be written in the following form:*

$$A_i(x) = \frac{l_i^2(x)}{2} + a_i \frac{H_n(x)}{H_n'(x_i)} \int_0^x l_i(t) dt + \frac{n H_n(x)}{H_n'(x_i)} \int_0^x l_i(t) dt + \frac{H_n'(x)}{2H_n'(x_i)} l_i(x) - x \frac{H_n(x)}{H_n'(x_i)} l_i(x) - \frac{H_n(x)}{2H_n'(x_i)} l_i'(0), \quad i = 1, \dots, n$$

LEMMA 4. *Let  $n$  be even. Then we have*

$$\sum_{i=1}^n \frac{|A_i(x)|}{W(x_i)} = O\left(e^{x^2} \sqrt{n}\right).$$

PROOF. It is well-known ([6], (2.36)) that

$$(41) \quad \sum_{i=1}^n e^{x_i^2} l_i^2(x) = O(1) e^{x^2}.$$

Since  $|a_i| = |1 - x_i^2|$ , thus we obtain that

$$(42) \quad \sum_{i=1}^n \frac{|a_i| \cdot |H_n(x)|}{W(x_i) \cdot |H_n'(x_i)|} \left| \int_0^x l_i(t) dt \right| = 2 \sum_{i=1}^n |a_i| \cdot |B_i(x)| = O(1) \frac{e^{x^2}}{\sqrt{n}}.$$

Further

$$\begin{aligned}
 (43) \quad n \sum_{i=1}^n \frac{|H_n(x)|}{W(x_i) \cdot |H'_n(x_i)|} \left| \int_0^x l_i(t) dt \right| &= 2n \sum_{i=1}^n |B_i(x)| = O(1)e^{x^2} \cdot \sqrt{n}. \\
 \sum_{i=1}^n \frac{|H'_n(x)|}{W(x_i) \cdot |H'_n(x_i)|} |l_i(x)| &= O(1) \frac{n|H_{n-1}(x)|}{\sqrt{2^n n!}} \sum_{i=1}^n \varphi_n^{1/2}(x_i) |l_i(x)| = \\
 &= O(1)\sqrt{n} \cdot \frac{|H_{n-1}(x)| \cdot e^{-x^2/2}}{\sqrt{2^n (n-1)!}} \cdot e^{x^2/2} \left( \sum_{|x_i| \geq 2\sqrt{\log n}}^n + \sum_{|x_i| \leq 2\sqrt{\log n}}^n \right).
 \end{aligned}$$

Here

$$\sum_{|x_i| \geq 2\sqrt{\log n}} \varphi_n^{1/2}(x_i) \cdot |l_i(x)| \leq c \frac{1}{n^{1/12}} \quad \sum_{|x_i| \geq 2\sqrt{\log n}} |l_i(x)| \leq c \cdot \frac{e^{x^2/2}}{n^{1/12}},$$

and

$$\begin{aligned}
 \sum_{|x_i| \leq 2\sqrt{\log n}} \varphi_n^{1/2}(x_i) \cdot |l_i(x)| &\asymp \frac{1}{\sqrt[4]{n}} \sum_{|x_i| \leq 2\sqrt{\log n}} |l_i(x)| \leq \\
 &\leq \frac{1}{\sqrt[4]{n}} \sum_{i=1}^n |l_i(x)| \leq c \frac{1}{\sqrt[4]{n}} \left( \log n + e^{x^2/2} \right),
 \end{aligned}$$

where we used [1], Hilfssatz 4, Satz1.

Hence

$$\begin{aligned}
 (44) \quad \sum_{i=1}^n \frac{|H'_n(x)|}{W(x_i) \cdot |H'_n(x_i)|} |l_i(x)| &= O(1)e^{x^2} \cdot \frac{\sqrt{n}}{n^{1/6}}. \\
 |x| \cdot |H_n(x)| \sum_{i=1}^n \frac{1}{W(x_i) \cdot |H'_n(x_i)|} |l_i(x)| &\asymp |x| \cdot \frac{|H_n(x)|}{\sqrt{2^n n!}} \sum_{i=1}^n \varphi_n^{1/2}(x_i) |l_i(x)| \asymp \\
 &\asymp |x| \cdot \frac{|H_n(x)| \cdot e^{-x^2/2}}{\sqrt{2^n n!}} \cdot e^{x^2/2} \cdot \frac{e^{x^2/2}}{n^{1/12}}.
 \end{aligned}$$

If  $|c| \geq 2\sqrt{n}$ , then  $\frac{|H_n(x)| \cdot e^{-x^2/2}}{\sqrt{2^n n!}} = O(1)e^{-cn}$ , therefore we can assume that  $|x| \leq 2\sqrt{n}$ . Hence

$$(45) \quad |x| \cdot |H_n(x)| \sum_{i=1}^n \frac{1}{W(x_i) \cdot |H_n'(x_i)|} |l_i(x)| = O(1)e^{x^2} \cdot \frac{\sqrt{n}}{n^{1/6}}.$$

From the obvious equation

$$\frac{H_n(x)}{2H_n'(x_i)} l_i'(0) = \frac{-H_n(x)}{2H_n(0)} l_i'(0)$$

we obtain that

$$(46) \quad \sum_{i=1}^n \frac{|H_n(x)|}{W(x_i) \cdot |H_n'(x_i)|} |l_i'(0)| = \frac{|H_n(x)|}{|H_n(0)|} \sum_{i=1}^n l_i^2(0) = O\left(e^{x^2}\right),$$

where we used that  $\sum_{i=1}^n l_i^2(0) = O(1)$  (it will be proved later more exactly). From (41)–(46) we get the Lemma.

LEMMA 5 ([7], Theorem 4 and [8], Theorem 1). *If  $f \in C^1(\mathbb{R})$*

$$\begin{aligned} \lim_{\pm\infty} x^{2r} f(x) W(x) &= 0, & (r = 0, 1, \dots), \\ \lim_{\pm\infty} f'(x) W(x) &= 0, \end{aligned}$$

*then there exist polynomials  $p_n$  of degree  $\leq n$  such that*

$$\begin{aligned} W(x)|f(x) - p_n(x)| &= O(1) \frac{1}{\sqrt{n}} \omega\left(f', \frac{1}{\sqrt{n}}\right), & x \in \mathbb{R}, \\ W(x)|f'(x) - p_n'(x)| &= O(1) \omega\left(f', \frac{1}{\sqrt{n}}\right), & x \in \mathbb{R}. \end{aligned}$$

LEMMA 6 ([5], Lemma 4). *For  $p_n(x)$  in Lemma 5 we have*

$$\begin{aligned} W(x)|p_n(x)| &= O(1) & (x \in \mathbb{R}), \\ W(x)|p_n'(x)| &= O(1) & (x \in \mathbb{R}), \\ W(x)|p_n''(x)| &= O(1)\sqrt{n} \cdot \omega\left(f', \frac{1}{\sqrt{n}}\right), & |x| < \sqrt{2n+1}. \end{aligned}$$

Now we prove the Theorem.

Let  $n$  be even. From [5] Theorem 2 it follows, that every polynomial  $Q_n$  of degree  $\leq 2n$  satisfies the equality

$$Q_n(x) = \sum_{i=1}^n Q_n(x_{i,n})A_{i,n}(x) + \sum_{i=1}^n (W Q_n)''(x_{i,n})B_{i,n}(x) + C_n H_n(x),$$

where

$$C_n = \frac{1}{H_n(0)} \left\{ Q_n(0) - \sum_{i=1}^n Q_n(x_{i,n})l_{i,n}^2(0) \right\}.$$

Let  $p_n$  be polynomial of degree  $\leq 2n$  satisfying Lemma 5, then we have

$$\begin{aligned} e^{-x^2} |f(x) - R_n(x)| &= \\ &= O(1) \left\{ W(x)|f(x) - p_n(x)| + e^{-x^2} \left| \sum_{i=1}^n (f(x_i) - p_n(x_i)) W(x_i) \frac{A_i(x)}{W(x_i)} \right| + \right. \\ &\left. + e^{-x^2} \left| \sum_{i=1}^n [(W p_n)''(x_i) - y_i''] B_i(x) \right| + e^{-x^2} |C_n H_n(x)| \right\}. \end{aligned}$$

Using Lemma 4 and 5 we obtain

$$\begin{aligned} e^{-x^2} |f(x) - R_n(x)| &= O(1) \left\{ \omega \left( f', \frac{1}{\sqrt{n}} \right) + e^{-x^2} \sum_{i=1}^n |y_i'' B_i(x)| + \right. \\ &+ e^{-x^2} \sum_{i=1}^n W(x_i) \cdot |p_n''(x_i) \cdot B_i(x)| + e^{-x^2} \sum_{i=1}^n |W'(x_i) \cdot p_n'(x_i) \cdot B_i(x)| + \\ &\left. + e^{-x^2} \sum_{i=1}^n |W''(x_i) \cdot p_n(x_i) \cdot B_i(x)| + e^{-x^2} |C_n H_n(x)| \right\}. \end{aligned}$$

Using Lemma 2 and 6 we have

$$\begin{aligned} e^{-x^2} \sum_{i=1}^n |W''(x_i) \cdot p_n(x_i) \cdot B_i(x)| &= O(1) \frac{1}{\sqrt{n}}, \\ e^{-x^2} \sum_{i=1}^n |W'(x_i) \cdot p_n'(x_i) \cdot B_i(x)| &= O(1) \frac{1}{\sqrt{n}}, \\ e^{-x^2} \sum_{i=1}^n W(x_i) \cdot |p_n''(x_i) \cdot B_i(x)| &= O(1) \omega \left( f', \frac{1}{\sqrt{n}} \right), \end{aligned}$$

and therefore

$$(47) \quad e^{-x^2} |f(x) - R_n(x)| = O(1)\omega\left(f', \frac{1}{\sqrt{n}}\right) + e^{-x^2} |C_n H_n(x)|.$$

Here

$$\begin{aligned} e^{-x^2} |C_n H_n(x)| &= e^{-x^2} \left| \frac{H_n(x)}{H_n(0)} \right| \cdot \left| p_n(0) - \sum_{i=1}^n p_n(x_i) l_i^2(0) \right| = \\ &= O(1) \left| p_n(0) - \sum_{i=1}^n p_n(x_i) l_i^2(0) \right| = \\ &= O(1) \left| \sum_{i=1}^n (p_n(0) - p_n(x_i)) l_i^2(0) + p_n(0) \left( 1 - \sum_{i=1}^n l_i^2(0) \right) \right|. \end{aligned}$$

It is clear that

$$p_n(0) - p_n(x_i) = x_i p_n'(x_i) + O(1)x_i^2 |p_n''(z_i)|,$$

where  $O \leq z_i \leq x_i$  or  $O \geq z_i \geq x_i$ .

Therefore

$$(48) \quad \sum_{i=1}^n (p_n(0) - p_n(x_i)) l_i^2(0) = \sum_{i=1}^n x_i p_n'(x_i) l_i^2(0) + O(1) \sum_{i=1}^n x_i^2 \cdot |p_n''(z_i)| \cdot l_i^2(0).$$

Here

$$\begin{aligned} \sum_{i=1}^n x_i^2 |p_n''(z_i)| \cdot l_i^2(0) &= H_n^2(0) \sum_{i=1}^n |p_n''(z_i)| \cdot \frac{1}{(H_n'(x_i))^2} \asymp \\ &\asymp H_n^2(0) \left( \sum_{\substack{i=1 \\ |x_i| \leq \sqrt{n}}}^n + \sum_{\substack{i=1 \\ |x_i| > \sqrt{n}}}^n \right) \leq c \cdot H_n^2(0) \sqrt{n} \cdot \omega\left(f', \frac{1}{\sqrt{n}}\right) \cdot \\ &\quad \cdot \left( \sum_{\substack{i=1 \\ |x_i| \leq \sqrt{n}}}^n e^{x_i^2/2} \frac{1}{(H_n'(x_i))^2} + \sum_{\substack{i=1 \\ |x_i| \geq \sqrt{n}}}^n e^{x_i^2/2} \frac{1}{(H_n'(x_i))^2} \right) = \end{aligned}$$

$$\begin{aligned}
 &= O(1) \frac{H_n^2(0)}{2^n n!} \sqrt{n} \cdot \omega \left( f', \frac{1}{\sqrt{n}} \right) \cdot \left( \sum_{\substack{i=1 \\ |x_i| \leq \sqrt{n}}}^n \frac{1}{\sqrt{n}} \frac{1}{e^{x_i^2/2}} + \frac{1}{e^{n/2}} \sum_{\substack{i=1 \\ |x_i| \geq \sqrt{n}}}^n \varphi_n(x_i) \right) = \\
 &= O(1) \omega \left( f', \frac{1}{\sqrt{n}} \right).
 \end{aligned}$$

Therefore

$$(49) \quad \sum_{i=1}^n x_i^2 |p_n''(x_i)| \cdot l_i^2(0) \leq c \cdot \omega \left( f', \frac{1}{\sqrt{n}} \right).$$

In consequence of (48)–(49)

$$(50) \quad \sum_{i=1}^n (p_n(0) - p_n(x_i)) l_i^2(0) = \sum_{i=1}^n x_i p_n'(x_i) l_i^2(0) + O(1) \omega \left( f', \frac{1}{\sqrt{n}} \right).$$

Here using Lemma 5

$$\sum_{i=1}^n x_i p_n'(x_i) l_i^2(0) = \sum_{i=1}^n x_i f'(x_i) l_i^2(0) + O(1) \omega \left( f', \frac{1}{\sqrt{n}} \right) \cdot \sum_{i=1}^n |x_i| e^{x_i^2/2} \cdot l_i^2(0).$$

Taking [5], (5.7) and [5] p. 166 into consideration

$$(51) \quad \sum_{i=1}^n x_i p_n'(x_i) l_i^2(0) = \sum_{i=1}^n x_i f'(x_i) l_i^2(0) + O(1) \frac{\log n}{\sqrt{n}} \omega \left( f', \frac{1}{\sqrt{n}} \right).$$

In consequence of this and (50)

$$(52) \quad \sum_{i=1}^n (p_n(0) - p_n(x_i)) l_i^2(0) = \sum_{i=1}^n x_i f'(x_i) l_i^2(0) + O(1) \omega \left( f', \frac{1}{\sqrt{n}} \right).$$

Here

$$l_i^2(0) = \frac{1}{(H_n'(x_i))^2} \cdot \frac{H_n^2(0)}{x_i^2}.$$

Therefore

$$\begin{aligned}
 &\sum_{\substack{i=1 \\ |x_i| > 1}}^n x_i f'(x_i) l_i^2(0) \leq c \cdot H_n^2(0) \sum_{\substack{i=1 \\ |x_i| > 1}}^n \frac{|f'(x_i)|}{|x_i|} \cdot \frac{1}{(h_n^l(x_i))^2} \asymp \\
 &\asymp \frac{1}{\sqrt{n}} \sum_{\substack{i=1 \\ |x_i| > 1}}^n e^{-x_i^2/2} \cdot e^{-x_i^2/2} \cdot |f'(x_i)| \cdot \frac{1}{|x_i|} \cdot \varphi_n(x_i) \leq
 \end{aligned}$$

$$\leq \frac{c}{\sqrt{n}} \sum_{\substack{i=1 \\ |x_i| > 1}}^n e^{-x_i^2/2} \cdot \frac{1}{|x_i|} \cdot \varphi_n(x_i) \asymp \frac{1}{\sqrt{n}}.$$

Therefore from (52)

$$(53) \quad \sum_{i=1}^n (p_n(0) - p_n(x_i)) l_i^2(0) = \sum_{\substack{i=1 \\ |x_i| \leq 1}}^n x_i f'(x_i) l_i^2(0) + O(1) \omega \left( f', \frac{1}{\sqrt{n}} \right).$$

Since

$$H_n'(x) = 2n H_{n-1}(x),$$

([4], (5.5.10)) and  $H_{2m+1}(x) = (-1)^m \cdot 2^{2m+1} \cdot m! \cdot x \cdot L_m^{(1/2)}(x^2)$  ([4], (5.6.1)), and

$$x_{i,n} = -x_{n-i+1,n}, \quad (i = 1, \dots, n), \quad \text{therefore} \quad l_i^2(0) = l_{n-i+1}^2(0).$$

Thus

$$\begin{aligned} \sum_{\substack{i=1 \\ |x_i| \leq 1}}^n x_i f'(x_i) l_i^2(0) &= \sum_{\substack{i=1 \\ |x_i| \leq 1}}^{n/2} (x_i f'(x_i) - x_i f'(-x_i)) l_i^2(0) = \\ &= \sum_{\substack{i=1 \\ |x_i| \leq 1}}^{n/2} x_i (f(x) + f(-x))'_{x=x_i} \cdot l_i^2(0). \end{aligned}$$

Since  $f(x) + f(-x)$  is even function therefore there exists a polynomial  $r_n(x^2)$  for which it is true that

$$W(x) |f(x) + f(-x) - r_n(x^2)| = O(1) \frac{1}{\sqrt{n}} \omega \left( (f(x) + f(-x))', \frac{1}{\sqrt{n}} \right), \quad (x \in \mathbb{R})$$

(\*)

$$W(x) |(f(x) + f(-x))' - 2x r_n'(x^2)| = O(1) \omega \left( (f(x) + f(-x))', \frac{1}{\sqrt{n}} \right), \quad (x \in \mathbb{R})$$

Here

$$\omega \left( (f(x) + f(-x))', \frac{1}{\sqrt{n}} \right) = \omega \left( [f'(x) - f'(-x)], \frac{1}{\sqrt{n}} \right) \leq 2\omega \left( f', \frac{1}{\sqrt{n}} \right),$$

since  $W$  and  $r$  are even functions.

Therefore

$$\begin{aligned} & \sum_{\substack{i=1 \\ x_i < 1}}^{n/2} x_i (f(x) + f(-x))'_{x=x_i} \cdot l_i^2(0) = \\ & = \sum_{\substack{i=1 \\ x_i \leq 1}}^{n/2} 2x_i^2 r'_n(x_i^2) l_i^2(0) + O(1)\omega\left(f', \frac{1}{\sqrt{n}}\right) \sum_{\substack{i=1 \\ x_i \leq 1}}^{n/2} e^{x_i^2/2} l_i^2(0) = \\ & = 2 \sum_{\substack{i=1 \\ x_i \leq 1}}^{n/2} x_i^2 r'_n(x_i^2) l_i^2(0) + O(1)\omega\left(f', \frac{1}{\sqrt{n}}\right), \end{aligned}$$

where we used [6], (2.36) with  $x = 0$ . Therefore from (53)

$$\begin{aligned} (54) \quad & \sum_{i=1}^n (p_n(0) - p_n(x_i)) l_i^2(0) = \\ & = 2 \sum_{\substack{i=1 \\ x_i \leq 1}}^{n/2} x_i^2 r'_n(x_i^2) l_i^2(0) + O(1)\frac{1}{\sqrt{n}} + O(1)\omega\left(f', \frac{1}{\sqrt{n}}\right) = \\ & = 2H_n^2(0) \sum_{\substack{i=1 \\ x_i \leq 1}}^{n/2} r'_n(x_i^2) \cdot \frac{1}{(H'_n(x_i))^2} + O(1)\frac{1}{\sqrt{n}} + O(1)\omega\left(f', \frac{1}{\sqrt{n}}\right). \end{aligned}$$

We know ([1]), that

$$\frac{1}{(H'_n(x_i))^2} \asymp \frac{1}{e^{x_i^2} \cdot 2^n n!} \cdot \frac{1}{\sqrt{n}}, \quad |x_i| \leq \sqrt{n}.$$

Using this we obtain from (54)

$$\begin{aligned} & \sum_{i=1}^n (\tilde{p}_n(0) - p_n(x_i)) l_i^2(0) = \\ & = O(1)\frac{H_n^2(0)}{2^n n!} \cdot \frac{1}{\sqrt{n}} \sum_{\substack{i=1 \\ x_i \leq 1}}^{n/2} |e^{-x_i^2} r'_n(x_i^2)| + O(1)\frac{1}{\sqrt{n}} + O(1)\omega\left(f', \frac{1}{\sqrt{n}}\right) \asymp \end{aligned}$$

$$(55) \quad \asymp \frac{1}{n} \sum_{\substack{i=1 \\ x_i \leq 1}}^{n/2} |e^{-x_i^2} r'_n(x_i^2)| + O(1) \frac{1}{\sqrt{n}} + O(1) \omega \left( f', \frac{1}{\sqrt{n}} \right).$$

We know ([8], Lemma 3) that the polynomial

$$q_{m+1}(x) := \sum_{k=0}^m \frac{x^k}{k!}$$

satisfies

$$(**) \quad \left. \begin{array}{l} c_1 e^x \leq q_{m+1}(x) \leq c_2 e^x \\ c_1 e^x \leq q'_{m+1}(x) = q_m(x) \leq c_2 \cdot e^x \end{array} \right\} \quad -\frac{m}{4} \leq x \leq 0.$$

(Choose  $m = 100n$ ).

Using that

$$q_{m+1} \left( -\frac{x}{2} \right) r'_n(x) = \left[ q_{m+1} \left( -\frac{x}{2} \right) r_n(x) \right]' + \frac{1}{2} \cdot q'_{m+1} \left( -\frac{x}{2} \right) r_n(x),$$

we obtain

$$\left| e^{-x_i^2/2} r'_n(x_i^2) \right| \asymp \left| \left[ q_{m+1} \left( -\frac{x}{2} \right) r_n(x) \right]'_{x=x_i^2} \right| + O(1) e^{-x_i^2/2} \cdot |r_n(x_i^2)|.$$

Here (using  $(*)$ )

$$\begin{aligned} e^{-x_i^2/2} r_n(x_i^2) &= e^{-x_i^2/2} (f(x) + f(-x)) + O(1) \frac{1}{\sqrt{n}} \cdot \omega \left( f', \frac{1}{\sqrt{n}} \right) = \\ &= O(1) + O(1) \frac{1}{\sqrt{n}} \cdot \omega \left( f', \frac{1}{\sqrt{n}} \right). \end{aligned}$$

Therefore

$$\begin{aligned} \sum_{i=1}^n (p_n(0) - p_n(x_i)) l_i^2(0) &= O(1) \frac{1}{n} \sum_{\substack{i=1 \\ x_i \leq 1}}^{n/2} e^{-x_i^2/2} \left| \left[ q_{m+1} \left( -\frac{x}{2} \right) r_n(x) \right]'_{x=x_i^2} \right| + \\ &+ O(1) \frac{1}{n} \sum_{\substack{i=1 \\ x_i \leq 1}}^{n/2} e^{-x_i^2/2} \left( 1 + \frac{1}{\sqrt{n}} \cdot \omega \left( f', \frac{1}{\sqrt{n}} \right) \right) + O(1) \frac{1}{\sqrt{n}} + O(1) \omega \left( f', \frac{1}{\sqrt{n}} \right). \end{aligned}$$

From this

$$(56) \quad \sum_{i=1}^n (p_n(0) - p_n(x_i)) l_i^2(0) = \\ = O(1) \frac{1}{n} \sum_{\substack{i=1 \\ x_i \leq 1}}^{n/2} e^{-x_i^2/2} \left| \left[ q_{m+1} \left( -\frac{x}{2} \right) r_n(x) \right]'_{x=x_i^2} \right| + O(1) \frac{1}{\sqrt{n}} + O(1) \omega \left( f', \frac{1}{\sqrt{n}} \right).$$

We know ([17], Problem 20 in Ch. IV., or [18]) that if  $t(x)$  is a polynomial and its degree  $\leq n$  then

$$|t'(x)| \leq c \frac{n}{\sqrt{1-x^2}} \sup_x \left| t \left( x + \frac{c}{n} \right) - t(x) \right| \leq c \frac{n}{\sqrt{1-x^2}} \sup_{\substack{a, b \\ |a-b| \leq \frac{c}{n}}} |t(a) - t(b)|,$$

where  $|x|, |a|, |b| < 1$

The function  $s(x) := \frac{1-x}{2}$  is bijection between  $[-1, 1]$  and  $[0, 1]$ . Therefore applying this theorem for  $t \left( \frac{1-x}{2} \right)$  we obtain

$$|r'(y)| \leq c \frac{n}{\sqrt{1-y^2}} \sup_{\substack{0 \leq a, b \leq 1 \\ |a-b| \leq \frac{c}{n}}} |t(a) - t(b)|, \quad \text{where } 0 \leq y < 1.$$

From this we get

$$|t'(y)| \leq c \frac{n}{\sqrt{d^2 - y^2}} \sup_{\substack{0 \leq a, b \leq d \\ |a-b| \leq c \frac{d}{n}}} |t(a) - t(b)|, \quad \text{where } 0 \leq y < d.$$

Use this inequality in (56) for  $q_{m+1} \left( -\frac{x}{2} \right) r_n(x)$  with  $d = n$ . Then we obtain

$$\left| \left[ q_{m+1} \left( -\frac{x}{2} \right) r_n(x) \right]'_{x=x_i^2} \right| \leq \\ \leq c \sup_{\substack{0 \leq a, b \leq n \\ |a-b| \leq c}} \left| q_{m+1} \left( -\frac{a}{2} \right) r_n(a) - q_{m+1} \left( -\frac{b}{2} \right) r_n(b) \right| \leq \\ \leq c \sup_{0 \leq z \leq n} \left| q_{m+1} \left( -\frac{z}{2} \right) r_n(z) \right| \asymp \sup_{0 \leq az \leq n} \left| e^{-z/2} r_n(z) \right|,$$

in the last step we used (\*\*\*) and that if  $0 \leq z \leq n$  then  $-\frac{m}{4} \leq -\frac{z}{2} \leq 0$  (we recall that  $m = 100n$ ).

Here

$$\sup_{0 \leq z \leq n} \left| e^{-z/2} r_n(z) \right| = \sup_{0 \leq x \leq \sqrt{n}} \left| e^{x^2/2} r_n(x^2) \right|.$$

Using (\*) we have

$$\begin{aligned} e^{-x^2/2} r_n(x^2) &= e^{-x^2/2} (f(x) + f(-x)) + O(1) \frac{1}{\sqrt{n}} \cdot \omega \left( f', \frac{1}{\sqrt{n}} \right) = \\ &= O(1) + O(1) \frac{1}{\sqrt{n}} \cdot \omega \left( f', \frac{1}{\sqrt{n}} \right). \end{aligned}$$

Hence

$$\left| \left[ q_{m+1} \left( -\frac{x}{2} \right) r_n(x) \right]'_{x=x_i^2} \right| = O(1) + O(1) \frac{1}{\sqrt{n}} \cdot \omega \left( f', \frac{1}{\sqrt{n}} \right).$$

Using this estimate we get from (56)

$$\begin{aligned} \sum_{i=1}^n (p_n(0) - p_n(x_i)) l_i^2(0) &= O(1) \frac{1}{n} \sum_{\substack{i=1 \\ x_i \leq 1}}^{n/2} \left( 1 + \frac{1}{\sqrt{n}} \cdot \omega \left( f', \frac{1}{\sqrt{n}} \right) \right) + O(1) \frac{1}{\sqrt{n}} + \\ &+ O(1) \omega \left( f', \frac{1}{\sqrt{n}} \right) = O(1) \frac{1}{\sqrt{n}} + O(1) \omega \left( f', \frac{1}{\sqrt{n}} \right). \end{aligned}$$

The estimation of  $1 - \sum_{i=1}^n l_i^2(0)$ .

We know that (see: e.g. [4], (14.1.7.), (14.5.6))

$$1 = \sum_{i=1}^n (1 + 2x_i^2) l_i^2(0).$$

From this

$$\begin{aligned} 1 - \sum_{i=1}^n l_i^2(0) &= 2 \sum_{i=1}^n x_i^2 l_i^2(0) = 4 \sum_{i=1}^{n/2} x_i^2 l_i^2(0) = \\ &= 4 H_n^2(0) \cdot \sum_{i=1}^{n/2} \frac{1}{(H_n'(x_i))^2} \asymp \left[ \frac{n!}{\left(\frac{n}{2}\right)!} \right]^{2n/2} \sum_{i=1}^{n/2} \frac{1}{(H_n'(x_i))^2}. \end{aligned}$$

Here

$$\begin{aligned} \sum_{i=1}^{n/2} \frac{1}{(H'_n(x_i))^2} &= \sum_{\substack{i=1 \\ x_i \leq \sqrt{n}}}^{n/2} \frac{1}{(H'_n(x_i))^2} + \sum_{\substack{i=1 \\ x_i > \sqrt{n}}}^{n/2} \frac{1}{(H'_n(x_i))^2} \\ &\asymp \sum_{\substack{i=1 \\ x_i \leq \sqrt{n}}}^{n/2} \frac{1}{e^{x_i^2} \cdot 2^n n!} \cdot \frac{1}{\sqrt{n}} + O(1) \sum_{\substack{i=1 \\ x_i > \sqrt{n}}}^{n/2} \frac{1}{e^{x_i^2} \cdot 2^n n!} \cdot \varphi_n(x_i). \end{aligned}$$

Therefore

$$\begin{aligned} 1 - \sum_{i=1}^n l_i^2(0) &\asymp \frac{1}{n} \sum_{\substack{i=1 \\ x_i \leq \sqrt{n}}}^{n/2} e^{-x_i^2} + O(1) \frac{1}{\sqrt{n}} \sum_{\substack{i=1 \\ x_i > \sqrt{n}}}^{n/2} e^{-x_i^2} \cdot \varphi_n(x_i) \asymp \\ (57) \quad &\asymp \frac{1}{n} \sum_{\substack{i=1 \\ x_i \leq c\sqrt{\log n}}}^{n/2} e^{-x_i^2} + O\left(\frac{1}{n^2}\right) \asymp \frac{1}{\sqrt{n}}. \end{aligned}$$

Hence

$$(58) \quad e^{-x^2} |C_n H_n(x)| = O(1) \frac{1}{\sqrt{n}} + O(1) \omega\left(f', \frac{1}{\sqrt{n}}\right).$$

REMARK 1. Let be  $f(x) := x^2$ ,  $x_0 := 0$ . Then we have

$$e^{-x_0^2} |f(x_0) - R_n(x_0)| \geq c \cdot \frac{1}{\sqrt{n}}.$$

Indeed, because if we choose  $p_n(x) = x^2$ , then

$$|f(x_0) - R_n(x_0)| = |C_n H_n(x_0)| = \sum_{i=1}^n x_i^2 l_i^2(0) \asymp \frac{1}{\sqrt{n}} \asymp \omega\left(f', \frac{1}{\sqrt{n}}\right).$$

Consequently the estimation in Theorem is not refinable.

REMARK 2. The ideas and methods given above can be applied to investigate the maximal class of functions for which other interpolating processes converges. As an example we show the case of Lagrange interpolation.

The mean convergence of interpolating processes was first investigated in the famous work [19] of P. ERDŐS and P. TURÁN (see also [20]) where the weighted  $L^2$ -convergence of the Lagrange interpolation on a finite interval is proved if the nodes of the interpolation are zeros of the polynomials

orthogonal with respect to this weight, satisfy some conditions. The mean convergence of the Lagrange interpolation on infinite intervals was solved only in 1961 by J. BALÁZS and P. TURÁN in [16]. For finite intervals the Erdős–Turán theorem gives estimate for the speed of convergence by  $E_n$  (see [21]). In [16] BALÁZS and TURÁN proved only the  $L^2$ -convergence for infinite intervals. In connection to the work [16] TURÁN raised the question of the estimate of rate of the  $L^2$ -convergence of the Lagrange interpolation for infinite intervals. Seemingly, the first such result is proved in JOÓ–KY [31] answering a problem of PAUL TURÁN [32, 33; Problem 17]. In [23] JOÓ and SZABADOS, in [24] JOÓ, in [25] SZABÓ, proved an  $L^1$ -convergence theorem. In [26] NÉVAI proved  $L^p$ -convergence ( $1 < p < \infty$ ) theorem for Lagrange-interpolation based on Hermite nodes without the estimate of the rate of convergence.

We shall investigate the convergence of the Lagrange-interpolation

$$L_n(f, x) := \sum_{k=1}^n f(x_{kn}) l_{kn}(x)$$

in  $L^2(\mathbb{R})$ , where we write briefly  $x_k := x_{kn}$  and  $l_k(x) := l_{kn}(x)$ . Denote  $W(x) := e^{-x^2/2}$  and

$$\|f\|_p := \left\{ \int_{-\infty}^{\infty} |f(x)|^p dx \right\}^{1/p} \quad (1 \leq p < \infty),$$

$$\|f\|_{\infty} := \text{vrai max } |f(x)| \quad (p = \infty).$$

For a function  $f : \mathbb{R} \rightarrow \mathbb{R}$  we shall use the generalized continuity modulus  $\omega_p(f, \delta)$  (see e.g. [28]–[30])

$$\omega_p(f, \delta) := \sup_{|t| \leq \delta} \|W(x+t)f(x+t) - W(x)f(x)\|_p + \|\tau(\delta x)W(x)f(x)\|_p,$$

where

$$\tau(x) := \begin{cases} |x|, & \text{if } |x| \leq 1, \\ 1 & \text{,if } |x| > 1. \end{cases}$$

Our result is the following.

STATEMENT. *If  $Wf \in L^2(\mathbb{R}) \cap C(\mathbb{R})$ ,  $L_n(f, x)$  is the Lagrange interpolation process,  $\omega_2\left(f, \frac{1}{\sqrt{n}}\right)$  and  $\omega_{\infty}\left(f, \frac{1}{\sqrt{n}}\right)$  are the generalized continuity*

modulus, then we have

$$\left( \int_{-\infty}^{\infty} e^{-x^2} |f(x) - L_{n+1}(f, x)|^2 dx \right)^{1/2} = O(1)\omega_2 \left( f, \frac{1}{\sqrt{n}} \right) + O(1)n^{1/4}\omega_{\infty} \left( f, \frac{1}{\sqrt{n}} \right).$$

Here the factor  $n^{1/4}$  is not refinable.

For the proof we need a lemma.

LEMMA. If  $Wf \in L^p(\mathbb{R})$  and  $1 \leq p \leq \infty$  then there exist polynomials  $p_n$  of degree  $\leq n$  such that

$$\|W(f - p_n)\|_p \leq c(p)\omega_p \left( f, \frac{1}{\sqrt{n}} \right),$$

where the constant  $c(p) > 0$  depends only on  $p$ .

PROOF. See [8], [27], [28].

REMARK. The polynomials  $p_n$  do not depend on  $p$ , because in the proofs the de La Vallée-Poussin means was used which is independent on  $p$ .

PROOF OF THE STATEMENT. Using the Lemma we obtain

$$\begin{aligned} (59) \quad & \left( \int_{-\infty}^{\infty} e^{-x^2} |f(x) - L_{n+1}(f, x)|^2 dx \right)^{\frac{1}{2}} \leq \\ & \leq \left( \int_{-\infty}^{\infty} e^{-x^2} |f(x) - p_n(x)|^2 dx \right)^{\frac{1}{2}} + \left( \int_{-\infty}^{\infty} e^{-x^2} |L_{n+1}(p_n - f, x)|^2 dx \right)^{\frac{1}{2}} = \\ & = O(1)\omega_2 \left( f, \frac{1}{\sqrt{n}} \right) + O(1)\omega_{\infty} \left( f, \frac{1}{\sqrt{n}} \right) \cdot \left( \int_{-\infty}^{\infty} e^{-x^2} \sum_{k=1}^n e^{x_k^2} l_k^2(x) dx \right)^{\frac{1}{2}}, \end{aligned}$$

where we used that  $\int_{-\infty}^{\infty} e^{-x^2} l_{k_1}(x) l_{k_2}(x) dx = 0$  if  $k_1 \neq k_2$ . Using [14], (6.1)

and (6.2) we obtain

$$e^{-x^2} \sum_{k=1}^n e^{x_k^2} l_k^2(x) \asymp e^{-x^2} \frac{H_n^2(x)}{2^n n!} \sum_{k=1}^n \varphi_n(x_k) \cdot \frac{1}{(x - x_k)^2},$$

where  $\varphi_n(x_k) = x_k - x_{k+1}$ .

Since  $H_n^2(x) = H_n^2(-x)$ , the zeros are symmetrical with respect to the origin,  $x_k = -x_{n-k}$ ,  $\varphi_n(x_k) \asymp \varphi_n(x_{n-k})$  and  $\frac{1}{(x-x_k)^2} \geq \frac{1}{(x-x_{n-k})^2}$  if  $x \geq 0$ ,

thus  
(60)

$$\left( \int_{-\infty}^{\infty} e^{-x^2} \sum_{k=1}^n e^{x_k^2} l_k^2(x) dx \right)^{\frac{1}{2}} \asymp \left( \int_0^{\infty} e^{-x^2} \frac{H_n^2(x)}{2^n n!} \sum_{k=1}^{n/2} \varphi_n(x_k) \cdot \frac{1}{(x-x_k)^2} dx \right)^{\frac{1}{2}}.$$

Here

$$(61) \quad \int_0^{\infty} e^{-x^2} \frac{H_n^2(x)}{2^n n!} \sum_{k=1}^{n/2} \varphi_n(x_k) \cdot \frac{1}{(x-x_k)^2} dx = \int_0^{x_{n/2}} + \sum_{j=1}^{\frac{n}{2}-1} \int_{x_{j+1}}^{x_j} + \int_{x_1}^{\infty}.$$

We know [29]

$$e^{-x^2} \sum_{k=1}^n e^{x_k^2} l_k^2(x) \leq 1 \quad (x \in \mathbb{R});$$

this means that finitely many members of the sum (60) and (61) can be estimated by  $O(1)$ . Hence

$$(62) \quad \int_{x_{j+1}}^{x_j} e^{-x^2} \frac{H_n^2(x)}{2^n n!} \sum_{k=1}^{n/2} \varphi_n(x_k) \cdot \frac{1}{(x-x_k)^2} dx = \\ = \int_{x_{j+1}}^{x_j} e^{-x^2} \frac{H_n^2(x)}{2^n n!} \sum_{\substack{k=1 \\ k \neq j, j \pm 1, j+2}}^{\frac{n}{2}} \varphi_n(x_k) \frac{1}{(x-x_k)^2} dx + \int_{x_{j+1}}^{x_j} O(1) dx.$$

We know [2]

$$(63) \quad x_k - x_{k+1} \asymp n^{-1/6} \cdot k^{-1/3}, \quad 1 \leq k \leq \frac{n}{2}.$$

If  $x_{j+1} \leq x \leq x_j$  and  $k \neq j, j \pm 1, j+2$  then

$$(64) \quad |x - x_k| \asymp |x_j - x_k| \asymp n^{-1/6} (j^{-1/3} + \dots + k^{-1/3}) \asymp \frac{|k^{2/3} - j^{2/3}|}{n^{1/6}}.$$

We know [3], p. 700 that

$$(65) \quad \frac{e^{-x^2/2} |H_n(x)|}{\sqrt{2^n n!}} \leq c \cdot n^{-\frac{1}{8}} \left( \sqrt{2n+1} - x_j \right)^{-\frac{1}{4}}, \quad x_{j+1} \leq x \leq x_j.$$

We know from [4] that  $\sqrt{2n+1} - x_1 \asymp n^{-1/6}$ , hence by (63)

$$(66) \quad \sqrt{2n+1} - x_j \asymp \sum_{i=1}^j n^{-1/6} \cdot i^{-1/3} \asymp n^{-1/6} j^{2/3} \quad (1 \leq j \leq \frac{n}{2}).$$

Using (63), (64), (65) and (66) we obtain

$$(67) \quad \int_{x_{j+1}}^{x_j} e^{-x^2} \frac{H_n^2(x)}{2^n n!} \sum_{\substack{k=1 \\ k \neq j, j \pm 1, j+2}}^{\frac{n}{2}} \varphi_n(x_k) \frac{1}{(x-x_k)^2} dx \asymp \\ \asymp n^{1/8} \int_{x_{j+1}}^{x_j} e^{-x^2} \frac{H_n^2(x)}{2^n n!} \sum_{\substack{k=1 \\ k \neq j, j \pm 1, j+2}}^{\frac{n}{2}} \frac{k^{-1/3}}{(k^{2/3} - j^{2/3})^2} dx = \\ = O(1) n^{1/8} (x_j - x_{j+1}) n^{-1/4} (\sqrt{2n+1} - x_j)^{-\frac{1}{2}} \sum_{\substack{k=1 \\ k \neq j, j \pm 1, j+2}}^{\frac{n}{2}} \frac{k^{-1/3}}{(k^{2/3} - j^{2/3})^2} = \\ = O(1) n^{-1/6} j^{-2/3} \sum_{\substack{k=1 \\ k \neq j, j \pm 1, j+2}}^{\frac{n}{2}} \frac{k^{-1/3}}{(k^{2/3} - j^{2/3})^2}.$$

Obviously

$$\sum_{\substack{k=1 \\ k \neq j, j \pm 1, j+2}}^{\frac{n}{2}} \frac{k^{-1/3}}{(k^{2/3} - j^{2/3})^2} = \sum_{k=1}^{j/2} + \sum_{\substack{k=j/2 \\ k \neq j, j \pm 1, j+2}}^{3j/2} + \sum_{\substack{k=3j/2 \\ k \neq j, j \pm 1, j+2}}^{\frac{n}{2}}.$$

Here

$$\sum_{k=1}^{j/2} \frac{k^{-1/3}}{(k^{2/3} - j^{2/3})^2} \asymp \sum_{k=1}^{j/2} \frac{k^{-1/3}}{j^{4/3}} \asymp j^{-2/3}$$

and

$$\sum_{\substack{k=3j/2 \\ k \neq j, j \pm 1, j+2}}^{\frac{n}{2}} \frac{k^{-1/3}}{(k^{2/3} - j^{2/3})^2} \leq c \cdot \sum_{\substack{k=3j/2 \\ k \neq j+1, j+2}}^{\frac{n}{2}} \frac{k^{-1/3}}{k^{4/3}} = O(1) j^{-2/3}.$$

Finally

$$\sum_{\substack{k=j/2 \\ k \neq jj \pm 1j+2}}^{3j/2} \frac{k^{-1/3}}{(k^{2/3} - j^{2/3})^2} \asymp \sum_{\substack{k=j/2 \\ k \neq jj \pm 1j+2}}^{3j/2} \frac{j^{-1/3}}{j^{-2/3}(k-j)^2} \asymp j^{1/3}.$$

Hence

$$(68) \quad j^{-2/3} \sum_{\substack{k=1 \\ k \neq jj \pm 1j+2}}^{n/2} \frac{k^{-1/3}}{(k^{2/3} - j^{2/3})^2} = O(1)j^{-1/3} + O(1)\frac{1}{j^{4/3}}.$$

Using (68) we obtain from (67)

$$\sum_{j=1}^{\frac{n}{2}-1} \int_{x_{j+1}}^{x_j} e^{-x^2} \frac{H_n^2(x)}{2^n n!} \sum_{\substack{k=1 \\ k \neq jj \pm 1j+2}}^{n/2} \varphi_n(x_k) \frac{1}{(x - x_k)^2} dx = O(1)n^{1/2}.$$

Using the above estimation and (62) we have

$$(69) \quad \int_{x_{n/2}}^{x_1} e^{-x^2} \frac{H_n^2(x)}{2^n n!} \sum_{k=1}^{n/2} \varphi_n(x_k) \frac{1}{(x - x_k)^2} dx = O(1)n^{1/2}.$$

Similarly

$$(70) \quad \int_0^{x_{n/2}} e^{-x^2} \frac{H_n^2(x)}{2^n n!} \sum_{k=1}^{n/2} \varphi_n(x_k) \frac{1}{(x - x_k)^2} dx = O(1)n^{1/2}.$$

We know from [3], p. 700 that

$$(71) \quad \frac{e^{-\frac{x^2}{2}} |H_n(x)|}{\sqrt{2^n n!}} \leq \leq c \cdot n^{-\frac{1}{8}} \left| \sqrt{2n+1} - x \right|^{-\frac{1}{4}} \cdot \exp \left[ -\xi(2n+1)^{\frac{1}{4}} \left( x - \sqrt{2n+1} \right)^{\frac{3}{2}} \right] \leq \leq c \cdot n^{-\frac{1}{8}} (x - \sqrt{2n+1})^{-\frac{1}{4}}, \quad x_1 \leq x \leq \sqrt{2}x_1,$$

$\xi$  is some positive number. Introduce the numbers  $x_j^*$  ( $1 \leq j \leq n/2$ ) in the following way:

$$x_1^* = x_1, \quad x_j + x_j^* = 2x_1 \quad (2 \leq j \leq n/2).$$

Then we can write

$$\int_{x_1}^{\infty} \frac{H_n^2(x)}{2^n n!} \sum_{k=1}^{n/2} \varphi_n(x_k) \frac{1}{(x - x_k)^2} dx = \sum_{j=1}^{n/2-1} \int_{x_j^*}^{x_{j+1}^*} + \int_{x_{n/2}^*}^{\infty}$$

Similarly as we proved (69) we obtain

$$(72) \quad \int_{x_1^*}^{x_{n/2}^*} e^{-x^2} \frac{H_n^2(x)}{2^n n!} \sum_{k=1}^{n/2} \varphi_n(x_k) \frac{1}{(x - x_k)^2} dx = O(1)n^{1/2}.$$

We know from [3], p. 700 that

$$\frac{e^{-x^2/2} |H_n(x)|}{\sqrt{2^n n!}} = O(1) \exp(-3x^2), \quad x \geq \sqrt{2}x_1.$$

Using this we obtain easily

$$(73) \quad \int_{x_{n/2}^*}^{\infty} e^{-x^2} \frac{H_n^2(x)}{2^n n!} \sum_{k=1}^{n/2} \varphi_n(x_k) \frac{1}{(x - x_k)^2} dx = O(1).$$

From (60), (61), (69), (70), (72) and (73) we obtain

$$(74) \quad \left( \int_{-\infty}^{\infty} e^{-x^2} \sum_{k=1}^n e^{x_k^2} l_k^2(x) dx \right)^{1/2} = O(1)n^{1/4}.$$

Now we prove that the estimation (74) is sharp. From (60), (61) and (67)

$$\begin{aligned} & \int_{-\infty}^{\infty} e^{-x^2} \sum_{k=1}^n e^{x_k^2} l_k^2(x) dx \geq c \int_0^{\infty} e^{-x^2} \frac{H_n^2(x)}{2^n n!} \sum_{k=1}^{n/2} \varphi_n(x_k) \frac{1}{(x - x_k)^2} dx \geq \\ (75) \quad & \geq \sum_{j=1}^{n/2-1} \int_{x_{j+1}^*}^{x_j^*} e^{-x^2} \frac{H_n^2(x)}{2^n n!} \sum_{k \neq j, j \pm 1, j+2}^{n/2} \varphi_n(x_k) \frac{1}{(x - x_k)^2} dx \geq \\ & \geq c \cdot n^{1/6} \sum_{\substack{j=1 \\ \sqrt{n}/2 \leq x_j \leq \sqrt{n}}}^{n/2-1} j^{1/3} \int_{x_{j+1}^*}^{x_j^*} e^{-x^2} \frac{H_n^2(x)}{2^n n!} dx \geq c\sqrt{n} \int_{\sqrt{n}/2}^{\sqrt{n}} e^{-x^2} \frac{H_n^2(x)}{2^n n!} dx. \end{aligned}$$

We know [4], (8.22.12) that

$$(76) \quad e^{-x^2/2} H_n(x) = 2^{\frac{n}{2} + \frac{1}{4}} (n!)^{\frac{1}{2}} (\pi n)^{-\frac{1}{4}} (\sin \varphi)^{-\frac{1}{2}} \cdot \left\{ \sin \left[ \left( \frac{n}{2} + \frac{1}{4} \right) (\sin 2\varphi - 2\varphi) + \frac{3\pi}{4} \right] + O\left(\frac{1}{n}\right) \right\},$$

where  $x = \sqrt{2n+1} \cos \varphi$ ,  $0 < \varepsilon \leq \varphi \leq \pi - \varepsilon$ ,  $\varepsilon$  is an arbitrary but fixed absolute constant, the error term is uniform in  $\varphi$ .

Using (76) we obtain

$$(77) \quad \int_{\sqrt{n}/2}^{\sqrt{n}} e^{-x^2} \frac{H_n^2(x)}{2^n n!} dx \geq c \int_{\varphi_b}^{\varphi_a} \left\{ \sin \left[ \left( \frac{n}{2} + \frac{1}{4} \right) (\sin 2\varphi - 2\varphi) + \frac{3\pi}{4} \right] + O\left(\frac{1}{n}\right) \right\} d\varphi,$$

where  $\sqrt{n} = \sqrt{2n+1} \cos \varphi_b$ ,  $\sqrt{n}/2 = \sqrt{2n+1} \cos \varphi_a$ .

Since  $\sin^2 \alpha = \frac{1}{2}(1 - \cos 2\alpha)$  therefore

$$\begin{aligned} \int_{\varphi_b}^{\varphi_a} \left\{ \sin \left[ \left( \frac{n}{2} + \frac{1}{4} \right) (\sin 2\varphi - 2\varphi) + \frac{3\pi}{4} \right] \right\}^2 &= \\ &= \frac{\varphi_a - \varphi_b}{2} - \frac{1}{2} \int_{\varphi_b}^{\varphi_a} \cos \left[ \left( n + \frac{1}{2} \right) (\sin 2\varphi - 2\varphi) + \frac{3\pi}{2} \right] d\varphi = \\ &= \frac{\varphi_a - \varphi_b}{2} - \frac{1}{2} \int_{\varphi_b}^{\varphi_a} \sin \left[ \left( n + \frac{1}{2} \right) (\sin 2\varphi - 2\varphi) \right] d\varphi. \end{aligned}$$

Introducing the new variable  $s := \sin 2\varphi - 2\varphi$  and integrating by part we obtain

$$\int_{\varphi_b}^{\varphi_a} \sin \left[ \left( n + \frac{1}{2} \right) (\sin 2\varphi - 2\varphi) \right] d\varphi = O\left(\frac{1}{n}\right).$$

Hence from (77) we obtain

$$(78) \quad \int_{\sqrt{n}/2}^{\sqrt{n}} e^{-x^2} \frac{H_n^2(x)}{2^n n!} dx \geq c > 0.$$

From (75) and (78)

$$\int_{-\infty}^{\infty} e^{-x^2} \sum_{k=1}^n e^{x_k^2} l_k^2(x) dx \geq c\sqrt{n},$$

which means that (74) is sharp. The Statement is proved.

### References

- [1] G. FREUD, Lagrangesche interpolation über die Nullstellen der Hermiteschen Orthogonalpolynome, *Studia Sci. Math. Hung.* 4 (1969), 179–190.
- [2] I. JOÓ, On Hermite functions, *Annales Univ. Sci. Budapest, Sectio Math.*, 35 (1992), 23–35.
- [3] R. ASKEY and S. WAINGER, Mean convergence of expansions in Langerre and Hermite series, *Amer. J. Math.*, 87 (1965), 695–708.
- [4] G. SZEGŐ, *Orthogonal polynomials* (in Russian), Moscow, 1962.
- [5] L. SZILI, Weighted (0,2)-interpolation on the roots of Hermite polynomials, *Annales Univ. Sci. Budapest., Sectio Math.*, 27 (1985), 153–166.
- [6] I. JOÓ, Stabil interpolációról, *MTA III. Oszt. Közl.*, 23 (1974), 329–363.
- [7] G. FREUD, On two polynomial inequalities I., *Acta Math. Acad. Sci. Hung.*, 22 (1971), 109–116.
- [8] G. FREUD, On polynomial approximation with the weight  $\exp\left(-\frac{1}{2}x^{2k}\right)$ , *Acta Math. Acad. Sci. Hung.* 24 (1973), 363–371.
- [9] G. FREUD, On an inequality of Markov type, *Soviet Math. Dokl.*, 12 (1971), 570–573.
- [10] J. BALÁZS, Súlyozott (0,2) interpoláció ultraszférikus polinomok gyökein, *MTA III. Oszt. Közl.*, 11 (1961), 305–388.
- [11] L. FEJÉR, Über Interpolation, *Göttinger Nachrichten* (1916), 66–91.
- [12] L. PÁL, A new modification of the Hermite–Fejér interpolation, *Analysis Mathematica*, 1 (1975), 137–205.
- [13] I. JOÓ and V. E. S. SZABÓ, A generalization of Pál interpolation process, *Acta Sci. Math.* (Szeged) (to appear).
- [14] I. JOÓ, An equiconvergence theorem for Hermite–Fourier series, *Annales. Univ. Sci. Budapest., Sectio Math.*, 37 (1994), 85–108.
- [15] I. JOÓ and L. SZILI, Weighted (0,2)-interpolation on the roots of Jacobi polynomials, *Acta Math. Hungar.*, 66 (1995), 25–50.
- [16] J. BALÁZS and P. TURÁN, Notes on interpolation, VIII., *Acta Math. Acad. Sci. Hungar.*, 12 (1961), 496–474.

- [17] A. F. TIMAN, *Theory of functions of a real variable* (in Russian), Moscow, 1961.
- [18] S. B. STEČKIN, *DAN*, **60** (1948), 1511–1514.
- [19] P. ERDŐS, P. TURÁN, On interpolation I., *Annals of Math.*, **38** (1937), 142–155.
- [20] P. ERDŐS, E. FELDHEIM, Sur le mode de convergence pour l'interpolation de Lagrange, *Comptes Rendus de l'Académie des Sciences*, Paris, **203** (1936), 913–915.
- [21] I. P. NATANSON, *Constructive Function Theory*, Akadémiai Kiadó, Budapest, 1952 (in Hungarian).
- [22] J. BALÁZS, P. TURÁN, Notes on interpolation IX., *Acta Math. Acad. Sci. Hung.*, **16** (1965), 215–220.
- [23] I. JOÓ, J. SZABADOS, On the weighted mean convergence of interpolating processes, *Acta Math. Acad. Sci. Hung.*, **57** (1991), 151–159.
- [24] I. JOÓ, On the order of mean convergence of interpolating processes, *Annales Univ. Sci. Budapest.*, Sectio Math., **33** (1990), 147–160.
- [25] S. SZABÓ, Estimate of the rate of mean convergence of Hermite interpolating processes, *Publ. Math. Debrecen*, **39** (1991), 339–345.
- [26] G. P. NÉVAI, Mean convergence of Lagrange interpolation, II, *J. of Approx. Theory*, **30** (1980), 263–276.
- [27] G. FREUD, On polynomial approximation with the weight  $e^{-x^2/2}$  (in Russian), *DAN*, **201** (1971), 1292–1294.
- [28] G. FREUD, A contribution to the problem of weighted polynomial approximation, *Linear Operators and Approximation* (Proc. Oberwolfach, 1971), ISNM vol. 20, Birkhäuser, Basel, 1972, 431–447.
- [29] E. EGERVÁRY, P. TURÁN, Notes on interpolation VI., *Acta Math. Acad. Sci. Hungar.*, **10** (1959), 55–62.
- [30] K. BALÁZS, I. JOÓ, On speed of mean convergence of Lagrange and Hermite interpolation based on the roots of Laguerre polynomials, *Publicationes Math. Debrecen*, **38** (1991), 247–253.
- [31] I. JOÓ, N. X. KY, Answer to a problem of P. Turán, *Annales Univ. Sci. Budapest.*, Sectio Math., **31** (1988), 229–241.
- [32] P. TURÁN, On some open problems of Approximation Theory (in Hungarian), *Mat. Lapok*, **25** (1974), 21–75.
- [33] P. TURÁN, On some open problems of Approximation Theory, *J. Approximation Theory*, **29** (1980), 23–85.

## ON THE INFLUENCE OF LEBESGUE FUNCTIONS ON THE SUMMABILITY OF DOUBLE FUNCTION SERIES

By

S. BARON and I. JOÓ\*

Department of Mathematics and Computer Science, Bar-Ilan University, Ramat Gan and  
Department for Analysis, Eötvös Loránd University, Budapest

*(Received April 3, 1995)*

In this paper we investigate summability of double function series. First we prove a convergence theorem for arbitrary (not necessarily orthogonal) systems, after we prove a divergence theorem for orthogonal systems.

We introduce some notations. If the limits of summation are not indicated, they are 0 and  $\infty$  on each index, and the free indices assume all values 0, 1, ...;

$$\lim_{m,n} a_{mn} = a$$

means that for every  $\varepsilon > 0$  there exists indices  $m_0, n_0$  such that  $|a_{mn} - a| < \varepsilon$  if  $m > m_0$  and  $n > n_0$ , i. e. the convergence in Pringsheim's sense. Denote by  $Q$  a  $d$ -dimensional interval,  $\mu$  is a positive measure on  $Q$ . Let  $E \subseteq \mathbb{R}^d$  with  $\mu(E) < \infty$ , and let  $f = \{f_{mn}\}$  be a system of functions  $f_{mn} \in L^1_\mu(E)$ . We assume that  $\mu$  is absolutely continuous with respect to Lebesgue measure  $\lambda$ . We still prefer to retain it in order that the sets of  $\mu$ -measure zero should at the same time be sets of measure zero in the sense of Lebesgue also (cf. [1], p.10). The state of things is, viz., as follows: Without our condition a set, of  $\mu$ -measure zero is not necessarily of measure zero in the sense of Lebesgue. But if  $\mu$  is absolutely continuous with respect to  $\lambda$  (see [9], p. 124), then the Radon-Nikodym derivative  $d\mu/d\lambda$  (see [9], p. 133) exists and is uniquely determined  $\lambda$ -almost everywhere. If  $d\mu/d\lambda > 0$  is true  $\lambda$ -almost everywhere on  $E$ , then we obtain

$$\mu(D) = \int_D \frac{d\mu}{d\lambda} d\lambda$$

---

\* The authors acknowledge the support of the Minerva Foundation in Germany through the Emmy Noether Institute at Bar-Ilan University.

(see [9], p. 134, Theorem B), where  $D \subset E$  is any measurable set. Hence (see [9], p.104) if  $\mu(D) = 0$ , then we have  $\lambda(D) = 0$ .

We note, that  $\mu$  fulfills the assumptions when  $\mu(x) = \mu_1(x_1) \dots \mu_d(x_d)$ , where  $\mu_1, \dots, \mu_d$  are positive, bounded and monotone increasing in their domains of definition (see [12], p.45) and the derivative of  $\mu_j$ , where  $j = 1, \dots, d$ , equals zero only in sets with Lebesgue measure zero (cf. [1], p. 9).

Let  $T$  be a triangular matrix method of summability with  $\tau_{mnkl}$  are its elements of series to sequence transformation. We assume that the double limits

$$(1) \quad \tau_{kl} := \lim_{m,n} \tau_{mnkl}$$

exists and

$$(2) \quad \tau_{00} = 1.$$

Denote  $K_{mn}$  the  $T$ -kernels of the function system  $f$ , i.e.

$$K_{mn}(u, x) := \sum_{k,l=0}^{m,n} \tau_{mnkl} f_{kl}(u) f_{kl}(x),$$

where  $x := (x_1, \dots, x_d)$ ,  $u := (u_1, \dots, u_d)$  and  $x, u \in E$ . Let

$$(3) \quad L_{mn}(x) := \int_E |K_{mn}(u, x)| d\mu(u),$$

$$(4) \quad L'_{mn}(x) := \int_E \max_{0 \leq l \leq n} |K_{ml}(u, x)| d\mu(u),$$

$$(5) \quad L''_{mn}(x) := \int_E \max_{0 \leq k \leq m} |K_{kn}(u, x)| d\mu(u).$$

The functions  $L_{mn}$ ,  $L'_{mn}$  and  $L''_{mn}$  defined by (3), (4) and (5) are called the  $T$ -Lebesgue functions of the system  $f$ .

Let

$$(6) \quad \Phi_{MN}(u, x) := \sum_{k,l=0}^{M,N} \tau_{mnkl} \tau_{\mu\nu kl} f_{kl}(u) f_{kl}(x),$$

where

$$M := \min\{m, \mu\}, \quad N := \min\{n, \nu\}.$$

We consider the double function series

$$(7) \quad \sum_{m,n} c_{mn} f_{mn}(x),$$

where the double sequence  $c = (c_{mn})$  belongs to the Banach space  $l^2$ , i.e.

$$(8) \quad \sum_{m,n} c_{mn}^2 < \infty,$$

and denote the  $T$ -means of (7) by  $\sigma_{mn}$ , i.e.

$$(9) \quad \sigma_{mn}(x) := \sum_{k,l=0}^{m,n} \tau_{mnkl} c_{kl} f_{kl}(x).$$

In the paper [6] we proved the following

**THEOREM 1.** *Let be condition (8) is fulfilled and for (6) are exist  $\xi_{kl} \geq 0$  so that*

$$(10) \quad \Phi_{MN}(u, x) = O(1) \sum_{k,l=0}^{M,N} \xi_{kl} |K_{kl}(u, x)|$$

for  $u, x \in E$  and there exist  $\xi'_k(M), \xi''_l(N) \geq 0$  such that

$$(11) \quad \sum_{k,l=0}^{M,N} \xi_{kl} = O(1), \quad \xi_{kl} = O(1) \xi'_k(M) \xi''_l(N).$$

Suppose, further, that

$$(12) \quad L_{mn}(x) = O(1),$$

$$(13) \quad L'_{mn}(x) = O(1), \quad L''_{mn}(x) = O(1)$$

are fulfilled for  $x \in E$ . Then there exist

$$(14) \quad \lim_{m,n} \sigma_{mn}(x), \quad \lim_m \sigma_{mn}(x), \quad \lim_n \sigma_{mn}(x)$$

almost everywhere on  $E$ .

We will prove that in Theorem 1 the strong conditions (12) and (13) one can replace by that of the following weaker conditions

$$(15) \quad L_{mn}(x) = O_x(1),$$

$$(16) \quad L'_{mn}(x) = O_x(1), \quad L''_{mn}(x) = O_x(1).$$

For this proof we needed the following two theorems of Banach (see [4], p. 36 and 37, Theorems II and III) for operators  $U_n$  defined on the all Banach space  $X$ , that is  $\text{dom } U_n = X$  for each  $n$ .

Denote  $S := S(E, \mu)$  the metric space of  $\mu$ -measurable,  $\mu$ -almost everywhere finite functions on  $E$ , endowed with the complete metric of convergence in measure.

THEOREM 2. Let  $\text{dom } U_n = X$  for all  $n$  and  $U_n : X \rightarrow S$  are continuous linear operators from the Banach space  $X$  to  $S$  and

$$(17) \quad \lim_n (U_n x)(u)$$

exists for  $\mu$ -almost every  $u \in E$ . If

$$(Ux)(u) := \lim_n (U_n x)(u),$$

then the limitoperator  $U : X \rightarrow S$  is linear and continuous.

THEOREM 3. Let  $\text{dom } U_n = X$  for all  $n$  and  $U_n : X \rightarrow S$  are continuous linear operators from the Banach space  $X$  to  $S$ . Then the following two statements are equivalent:

$\alpha$ ) limit (17) for all  $x \in X$  exists for  $\mu$ -almost every  $u \in E$ ;

$\beta$ ) a dense subset  $X' \subset X$  exists such that (17) exists for all  $x \in X'$  and  $\sup_n |(U_n x)(u)| < \infty$  for all  $x \in X$  and for  $\mu$ -almost all  $u \in E$ .

In the Theorems 2 and 3 the set

$$\{u : \exists \lim_n (U_n x)(u)\}$$

may depends on  $x$ .

Now we can to prove the following

THEOREM 4. If condition (8) is fulfilled and the method  $T$  satisfies the conditions (10), (11), (15) and (16) for  $u, x \in E$ , then (14) holds.

PROOF. Let us put that (15) and (16) are satisfied. Introduce the sets

$$E_R := \{x \in E : \sup_{m,n} L_{mn}(x) \leq R, \sup_{m,n} L'_{mn}(x) \leq R, \sup_{m,n} L''_{mn}(x) \leq R\}.$$

Then

$$(18) \quad E = \bigcup_{R=1}^{\infty} E_R,$$

since, for example, from (15) follows that for every  $x \in E$  be exists some  $R$  such that  $L_{mn}(x) \leq R$ . By Theorem 1 we see that from the conditions (8) and (10)–(13) the existence of the limits (14) a.e. on each  $E_R$  follows. From (14) we obtain that

$$(19) \quad \sup_{m,n} |\sigma_{mn}(x)| < \infty$$

is fulfilled a.e. on every  $E_R$ . By (18) we obtain that (19) is satisfied a.e. on  $E$ .

We define the operators

$$U_R^{m,n} : l^2 \rightarrow S \quad (m, n, R \text{ fixed})$$

by the formula

$$(20) \quad U_R^{m,n} c = \sum_{k,l=0}^R \tau_{mnkl} c_{kl} f_{kl}.$$

The operators  $U_R^{m,n}$  are linear and continuous, since if  $c_r \rightarrow c$  in  $l^2$ , where  $c_r = (c_{kl}^r)$ , then

$$(U_R^{m,n} c_r)(x) \rightarrow (U_R^{m,n} c)(x) \quad (r \rightarrow \infty)$$

for every  $x \in E$ . By Theorem of Lebesgue (see, for example [9], p.92, [13], p.100)

$$(U_R^{m,n} c_r)(x) \Rightarrow (U_R^{m,n} c)(x) \quad (r \rightarrow \infty),$$

on  $E$ , i.e. converges in measure on  $E$ . Further, since  $T$  is a triangular method of summability, then for every  $m, n$  we obtain from (20) and (9) that

$$(21) \quad \lim_R (U_R^{m,n} c)(x) = \sigma_{mn}(x).$$

Let

$$U_{mn} : l^2 \rightarrow S$$

by the formula

$$(22) \quad (U_{mn} c)(x) = \sigma_{mn}(x).$$

From (21) and (22) we see that

$$\lim_R (U_R^{m,n} c)(x) = (U_{mn} c)(x)$$

for  $x \in E$ , therefore by Theorem 2 the operators  $U_{mn}$  are linear and continuous, since  $\text{dom } U_{mn} = l^2$  for each  $m, n$ . From (19) and (22) we obtain

$$(23) \quad \sup_{m,n} |(U_{mn} c)(x)| < \infty \quad \text{a.e. on } E.$$

Further, the set of sections  $\{c_A\}$ , where  $c_A = (c_{kl}^A)$  and  $c_{kl}^A = c_{kl}$  if  $k, l \leq A$  and  $c_{kl}^A = 0$  if  $k > A$  or  $l > A$ , is dense in  $l^2$ , and the double limit

$$(24) \quad \lim_{m,n} (U_{mn} c_A)(x) = \sum_{k,l=0}^A \tau_{kl} c_{kl} f_{kl}(x)$$

by (1) exists on  $E$ . Denoting

$$R = \min\{m, n\},$$

from (24) we get the existence of

$$(25) \quad \lim_R (U_{mn} c_A)(x) = \sum_{k,l=0}^A \tau_{kl} c_{kl} f_{kl}(x)$$

on the dense set  $\{c_A\}$ . By Theorem 3 from (23) and (25) follows, in view of (22), that

$$\lim_R (U_{mn} c)(x) = \lim_{m,n} \sigma_{mn}(x)$$

for all  $c \in l^2$  a.e. on  $E$ . The proofs of the existence of the simple limits in (14) are analogous and more simpler. Thus Theorem 4 is proved. ■

We remark that the existence of the equality

$$\lim_{m,n} (U_{mn} c)(x) = \lim_{m,n} \sigma_{mn}(x)$$

a.e. on  $E$  follows from (23) and (24), instead of Theorem 3, also by the following

**THEOREM 3'.** *Let  $T_{mn} : X \rightarrow S$  are continuous linear operators from the Banach space  $X$  to  $S$ . If*

$$\lim_{m,n} T_{mn} x$$

*exists for each  $x$  in a dense set, and if for each  $x$  in  $X$  the set  $\{T_{mn} x\}$  is bounded, then the double limit*

$$Tx = \lim_{m,n} T_{mn} x$$

*exists for each  $x$  in  $X$ , and  $T$  is continuous and linear.*

Theorem 3' is a particular case of Theorem 18 on p. 55 in [7].

Let  $\lambda = (\lambda_{mn})$  be a nondecreasing double sequence of positive numbers, i.e.  $(\lambda_{m,n_0})$  and  $(\lambda_{m_0,n})$  are nondecreasing sequences for any fixed  $m_0$  and  $n_0$ .

We define a new function system  $F = \{F_{mn}\}$ , where

$$F_{mn}(x) = \frac{f_{mn}(x)}{\sqrt{\lambda_{mn}}},$$

and consider the function series

$$(26) \quad \sum_{m,n} a_{mn} F_{mn}(x)$$

with

$$(27) \quad \sum_{m,n} a_{mn}^2 < \infty.$$

We denote by  $K_{mn}(F; \cdot, \cdot)$  the  $T$ -kernels and by  $L_{mn}(F; \cdot)$ ,  $L'_{mn}(F; \cdot)$  and  $L''_{mn}(F; \cdot)$  the  $T$ -Lebesgue functions of the function system  $F$ , i.e.

$$\begin{aligned}
 K_{mn}(F; u, x) &:= \sum_{k,l=0}^{m,n} \tau_{mnkl} F_{kl}(u) F_{kl}(x), \\
 L_{mn}(F; x) &:= \int_E |K_{mn}(F; u, x)| d\mu(u), \\
 L'_{mn}(F; x) &:= \int_E \max_{0 \leq l \leq n} |K_{ml}(F; u, x)| d\mu(u), \\
 L''_{mn}(F; x) &:= \int_E \max_{0 \leq k \leq m} |K_{kn}(F; u, x)| d\mu(u).
 \end{aligned}$$

Analogously to (6) we denote

$$\Phi_{MN}(F; u, x) := \sum_{k,l=0}^{M,N} \tau_{mnkl} \tau_{\mu\nu kl} F_{kl}(u) F_{kl}(x).$$

From Theorem 4 we now obtain the following

**THEOREM 5.** *If the condition*

$$(28) \quad \sum_{m,n} c_{mn}^2 \lambda_{mn} < \infty$$

*is fulfilled and there exist  $\xi_{kl}$ ,  $\xi'_k(M)$ ,  $\xi''_l(N) \geq 0$  such that  $T$  satisfies the conditions (11) and*

$$(29) \quad \Phi_{MN}(F; u, x) = O(1) \sum_{k,l=0}^{M,N} \xi_{kl} |K_{kl}(F; u, x)|,$$

$$(30) \quad L_{mn}(F; x) = O_x(1),$$

$$(31) \quad L'_{mn}(F; x) = O_x(1), \quad L''_{mn}(F; x) = O_x(1)$$

*for  $u, x \in E$ , then (14) for the double function series (7) holds.*

**PROOF.** Denoting  $c_{mn} = a_{mn} / \sqrt{\lambda_{mn}}$ , we obtain that (28) transfers to (27). By the Theorem 4 from (27) and other conditions of Theorem 5 follows that (14) for the double series (26) holds, since (27), (29), (30) and (31) are instead of (8), (10), (12) and (13). But (26) is the double series (7).

For simple series Theorem 5 was proved in [3], see also [2], p. 266–270 and 297–298.

Now we consider a divergence theorem, from which we see that no exist a strong relation between the magnitude of the Lebesgue functions and the summability a.e. not only of the general double function series (7), but even for orthogonal double series.

Let  $\varphi = \{\varphi_{mn}\}$  be a orthonormal system of functions  $\varphi_{mn} \in L^2_\mu(Q)$ . We say that the double sequence  $\lambda$  is a Weyl multiplier for the  $T$ -summability of the double orthogonal series

$$(32) \quad \sum_{m,n} c_{mn} \varphi_{mn}(x)$$

if the satisfying of the condition (28) implies the  $T$ -summability of the double series (32) a.e. on  $E \subseteq Q$  with  $\mu(E) < \infty$ . For simple orthogonal series EFIMOV [8] proved, that there exist orthonormal systems and regular methods of summability, for which the monotone increasing of the majorant of the Lebesgue functions (even in any power  $\alpha \geq 1$ ) is not a Weyl multiplier for this summability. A simple proof of the Efimov's theorem found ULYANOV ([15], p. 35–36). Using the idea of [15], we can generalize it for double orthogonal series (32).

We need the following Lemmas.

LEMMA 6. *If  $u_{mn} \geq 0$  and the double series*

$$\sum_{m,n} u_{mn}$$

*converges, then there exist  $0 < \gamma_m \uparrow \infty$  and  $0 < \delta_n \uparrow \infty$  such that*

$$\sum_{m,n} u_{mn} \gamma_m \delta_n < \infty.$$

PROOF. See [10], p. 94. ■

We denote

$$V_{mn} = \xi_m \eta_n \sum_{k,l=0}^{m,n} u_{kl}.$$

In the sequel we write  $r\text{-}\lim_{m,n} x_{mn} = 0$  instead of

$$\lim_{m,n} x_{mn} = \lim_n x_{mn} = \lim_m x_{mn} = 0.$$

LEMMA 7. *If  $u_{mn} \geq 0$  and the double series*

$$\sum_{m,n} \xi_m \eta_n u_{mn}$$

converges with

$$(33) \quad \xi_m \downarrow 0, \quad \eta_n \downarrow 0,$$

then

$$r\text{-}\lim_{m,n} V_{mn} = 0.$$

PROOF. See [1], p. 72, and [5], p. 177. ■

LEMMA 8. *There exists a non-factorizable double series (32), divergent everywhere on  $Q$  and its coefficients satisfy the condition (8).*

PROOF. See [5], pp. 185–186 and this is based on the results of [14], pp. 91, 101 and 104. ■

Let  $I$  the convergence method, that is method  $T$  with  $\tau_{mnkl} = 1$ .

LEMMA 9. *If (33) hold and the series  $\sum \xi_m$  and  $\sum \eta_n$  converge, then the  $I$ -Lebesgue functions  $L_{mn}(I; \cdot)$  of every orthonormal system  $\varphi$  a.e. on  $Q$  satisfy the following estimate*

$$(34) \quad r\text{-}\lim_{m,n} \sqrt{\xi_m \eta_n} L_{mn}(I; x) = 0.$$

PROOF. (Cf. [11], p. 173, Theorem 551.) For the system  $\varphi$  we have

$$L_{mn}(I; x) = \int_Q |K_{mn}(I; x)| d\mu(u) = \int_Q \left| \sum_{k,l=0}^{m,n} \varphi_{kl}(u) \varphi_{kl}(x) \right| d\mu(u).$$

Further,

$$\sum_{m,n} \xi_m \eta_n \int_Q |\varphi_{mn}(x)|^2 d\mu(x) = \sum_{m,n} \xi_m \eta_n < \infty,$$

and by B. LEVI theorem (cf. [1], Theorem 1.2.2, [5], Lemma 1, [6], Lemma 3) we obtain that the double series

$$\sum_{m,n} \xi_m \eta_n |\varphi_{mn}(x)|^2$$

converges a.e. on  $Q$  (by our assumption on the measure  $\mu$ ). Consequently, by Lemma 7 we obtain

$$(35) \quad r\text{-}\lim_{m,n} \xi_m \eta_n \sum_{k,l=0}^{m,n} |\varphi_{kl}(x)|^2 = 0$$

a.e. on  $Q$ . Using the Cauchy–Schwarz inequality we get

$$\begin{aligned} |L_{mn}(I;x)|^2 &\leq \int_Q 1^2 d\mu(u) \cdot \int_Q |K_{mn}(I;x)|^2 d\mu(u) = \\ &= \mu(Q) \int_Q \sum_{k,l=0}^{m,n} \varphi_{kl}(u)\varphi_{kl}(x) \sum_{i,j=0}^{m,n} \varphi_{ij}(u)\varphi_{ij}(x) d\mu(u) = \\ &= \mu(Q) \sum_{k,l=0}^{m,n} \sum_{i,j=0}^{m,n} \varphi_{kl}(x)\varphi_{ij}(x)\delta_{ki}\delta_{lj} = \mu(Q) \sum_{k,l=0}^{m,n} |\varphi_{kl}(x)|^2 \end{aligned}$$

and from this by (35) we obtain

$$r\text{-}\lim_{m,n} \xi_m \eta_n |L_{mn}(I;x)|^2 = 0$$

a.e. on  $Q$ , that is (34). ■

Let  $\lambda = (\lambda_{mn})$  be a nondecreasing double sequence of positive numbers.

**THEOREM 10.** *For any function  $g$  of two variables with  $0 < g(x) \uparrow +\infty$  as  $x_1, x_2 \rightarrow +\infty$  there exists:*

- 1) *a triangular factorizable method of summability  $T = A \odot B$  with regular factors  $A$  and  $B$ ,*
- 2) *a non-decreasing unbounded factorizable double sequence  $\lambda = (\lambda'_m \lambda''_n)$ ,*
- 3) *an orthonormal system  $\varphi$ ,*

*such that  $\mu$ -almost everywhere on  $Q$  the  $T$ -Lebesgue functions satisfy the estimate*

$$(36) \quad L_{mn}(T;x) = O(\lambda_{mn}),$$

*but on the other hand some double orthogonal series (32) is not  $T$ -summable a.e. on  $Q$ , although*

$$(37) \quad \sum_{m,n} |c_{mn}|^2 G(\lambda_{mn}) < \infty,$$

*where  $G(\lambda_{mn}) = g(\lambda'_m, \lambda''_n)$ .*

**PROOF.** By Lemma 8 we can choose some system  $\varphi$  and some double sequence  $c$  such that the double series (32) diverges on  $Q$ , but (8) is satisfied. By Lemma 6 from (8) follows that there exist  $0 < \gamma_m \uparrow \infty$  and  $0 < \delta_n \uparrow \infty$  such that

$$(38) \quad \sum_{m,n} |c_{mn}|^2 \gamma_m \delta_n < \infty.$$

Since  $\sum 1/n^2 < \infty$ , then by Lemma 9 for any orthonormal system  $\varphi$  the  $I$ -Lebesgue functions satisfy a.e. on  $Q$  the estimate

$$(39) \quad L_{mn}(I; x) = O(1)(m + 1)(n + 1).$$

Further, for finding  $\lambda$  we remark that the double sequences  $(g(m, n))$  and  $(\gamma_m \delta_n)$  increase, therefore we can find non-increasing sequences  $(m_i)$  and  $(n_j)$  of integers, which take the all natural values, such that  $m_i \leq i + 1$  and  $n_j \leq j + 1$ , but

$$(40) \quad g(m_i, n_j) \leq A\gamma_i \delta_j$$

for all  $i, j \geq \rho$ , where  $A$  and  $\rho \geq 1$  are some constants.

Now we choose

$$(41) \quad \lambda_{ij} = (m_i + 1)(n_j + 1)$$

and put  $\tau_{mnkl} = 1$  if  $k \leq m_i \leq m$  and  $l \leq n_j \leq n$ , but  $\tau_{mnkl} = 0$ , if  $k > m_i$  or  $l > n_j$ . This method  $T = (\tau_{mnkl})$  has regular factors and from (39) we deduce that a.e. on  $Q$  the estimate (36) hold, since for

$$L_{ij}(T; x) := \int_Q \left| \sum_{k,l=0}^{m_i, n_j} \varphi_{kl}(x) \varphi_{kl}(u) \right| d\mu(u)$$

in view of (39) and (41) we have

$$L_{ij}(T; x) = O(\lambda_{ij}).$$

On the other hand, since we has assumed that the double orthogonal series (32) diverges everywhere on  $Q$  and  $(m_i)$  and  $(n_j)$  take the all natural values, therefore the double sequence  $(s_{m_i, n_j})$  of partial sums of the double series (32) also diverges everywhere on  $Q$ . This means that (32) is not  $T$ -summable everywhere on  $Q$ .

It is remained to prove that for our  $T$  and  $\lambda$  the condition (37) holds. Indeed, in view of (41), (40) and (38) we have

$$\sum_{i,j} |c_{ij}|^2 G(\lambda_{ij}) = O(1) + \left( \sum_{i,j=\rho,0}^{\infty,\rho} + \sum_{i,j=0,\rho}^{\rho,\infty} \right) |c_{ij}|^2 G(\lambda_{ij}) + A \sum_{i,j=\rho}^{\infty} |c_{ij}|^2 \gamma_i \delta_j < \infty,$$

since, for example, by (41), (40) and (38)

$$\sum_{i,j=\rho,0}^{\infty,\rho} |c_{ij}|^2 G(\lambda_{ij}) \leq \sum_{j=0}^{\rho} \sum_{i=\rho}^{\infty} |c_{ij}|^2 G(\lambda_{i\rho}) \leq A\delta_\rho \sum_{j=0}^{\rho} \sum_{i=\rho}^{\infty} |c_{ij}|^2 \gamma_i = O(1).$$

Theorem 10 is proved.

Analogous to Theorem 10 is in fact proved in [5]. We remark, that the following question is open: May be exists the function  $g$ , a method  $T$ , a double sequence  $\lambda$  and a system  $\varphi$  from Theorem 10 such that together with the condition (36) the conditions

$$L'_{mn}(T;x) = O(\lambda_{mn}), \quad L''_{mn}(T;x) = O(\lambda_{mn})$$

are also satisfied  $\mu$ -almost everywhere on  $Q$ ?

Taking  $g(x) = (x_1 x_2)^\alpha$  with  $\alpha \geq 1$  we from Theorem 10 yields the following

**COROLLARY 11.** *There exist a method of summability  $T$  with regular factors and a monotone increasing unbounded double sequence  $\lambda$  such that the double sequence  $(|\lambda_{mn}|^\alpha)$  for none  $\alpha \geq 1$  is not a Weyl multiplier for  $T$ -summability of the double orthogonal series (32), in spite of the fact that the condition (36) holds.*

## References

- [1] ALEXITS, G., *Konvergenzprobleme der Orthogonalreihen*, Akadémiai Kiadó, Budapest, 1960.
- [2] ALEXITS, G., *Approximation Theory* (Selected papers), Akadémiai Kiadó, Budapest, 1983.
- [3] ALEXITS, G., JOÓ, I., TANDORI, K., On summability of orthogonal series, *Acta Math. Acad. Sci. Hung.*, **29** (1977), 201–205.
- [3] ALEXITS, G., JOÓ, I., TANDORI, K., On summability of orthogonal series, *Acta Math. Acad. Sci. Hung.*, **29** (1977), 201–205.
- [4] BANACH, S., Sur la convergence presque partout de fonctionnelles linéaires, *Bull. Sci. Math.*, **50** (1926), 27–32 and 36–43.
- [5] BARON, S., Lebesgue functions for summability of double orthogonal series, *Tartu Ülik. Toimetised*, (in Russian), **220** (1968), 166–189.
- [6] BARON, S., JOÓ, I., Lebesgue functions for convergence and summability of double function series, *Annal. Univ. Sci. Budapest.*, Sectio Math., **37** (1994), 73–84.
- [7] DUNFORD, N., SCHWARTZ, J. T., *Linear operators, Part I: General Theory*, Interscience Publ., New York, 1967.
- [8] EFIMOV, A. V., On nonsummability of orthogonal series by linear methods, *Dokl. Akad. Nauk SSSR*, (in Russian) **152** (1963), 31–34.
- [9] HALMOS, P. R., *Measure Theory*, Springer-Verlag, New York–Heidelberg–Berlin, 1974.

- 
- [10] KACZMARZ, S., Zur Theorie der Fourierschen Doppelreihen, *Studia Math.*, **2** (1930), 91–96.
- [11] KACZMARZ, S., STEINHAUS, H., *Theorie der Orthogonalreihen*, 2nd edition, New York, 1951.
- [12] KAMKE, E., *Das Lebesgue-Stieltjes-Integral*, 2nd Edition, Teubner Verlag, Leipzig, 1960.
- [13] RIESZ, F., SZ.-NAGY, B., *Lectures on Functional Analysis*, Dover Publ., New York, 1990.
- [14] ULYANOV, P. L., Divergent Fourier series, *Uspehi Mat. Nauk*, (in Russian) **16** (1961) no. 3, 61–142.
- [15] ULYANOV, P. L., Solved and unsolved problems in the theory of trigonometric and orthogonal series *Uspehi Mat. Nauk*, (in Russian) **19** (1964) no. 1, 3–69.



## ON CONTROL THEORY I (MAXIMUM PRINCIPLE)

By  
I. JOÓ\*

Department of Analysis, Eötvös Loránd University, Budapest

(Received July 4, 1995)

### 1. Introduction

Consider a system of equations

$$\dot{x}_i = f_i(x, u) \quad i = 1, \dots, n$$

controlled by the piecewise continuous function  $u(t)$ . The trajectory  $x(t) \in \mathbb{R}^n$  goes from  $x(t_0) = x^{(0)}$  to  $x(t_1) = x^{(1)}$ . Given  $t_0, x^{(0)}, x^{(1)}$  we intend to minimize a functional

$$J(x, u, t_1) = \int_{t_0}^{t_1} f_0(x, u) dt.$$

The triplet  $(\bar{x}, \bar{u}, \bar{t}_1)$  giving a minimum for  $J$  is called optimal,  $\bar{x}(t)$  is the optimal trajectory,  $\bar{u}(t)$  is the optimal control. Several necessary conditions for optimality are known. The first one is the Euler–Lagrange equation. If  $u(t) \in U$  a.e. for some fixed open domain  $U \subset \mathbb{R}^r$ , a necessary condition has been given by WEIERSTRASS, see e.g. [19]. In the applications arises the need of considering closed domains  $U$  for which the Weierstrass theorem can not be applied. For this case PONTRYAGIN [10] and his coworkers developed the so-called Pontryagin maximum principle stating

**THEOREM A°.** *If  $(\bar{x}, \bar{u}, \bar{t}_1)$  is optimal then there exist functions  $\Psi = (\Psi_1, \dots, \Psi_n)$ ,  $\dot{\Psi} = \Psi(t)$  and a constant  $\lambda \leq 0$  such that*

$$\dot{\Psi} = -\dot{\Psi} \cdot f'_x(\bar{x}(t), \bar{u}(t), t), \quad f = (f_1, \dots, f_n)^T$$

*and the function*

---

\* Supported by National Scientific Research grant OTKA N° 014244

$$H(x, \Psi, u) = \Psi \cdot f(x, u) + \lambda f_0(x, u)$$

takes its maximum at  $u = \bar{u}(t)$  i. e. for all  $t$

$$H(\bar{x}(t), \Psi(t), \bar{u}(t)) = 0$$

$$H(\bar{x}(t), \Psi(t), u) \leq 0 \quad \text{if } u \in U. \quad \blacksquare$$

REMARKS. 1. We can fix the time  $t_1$  by introducing a new component  $x_{n+1}$  by

$$\dot{x}_{n+1} = 1, \quad x_{n+1}(t_0) = t_0, \quad x_{n+1}(t_1) = t_1.$$

2. The original proof of PONTYAGIN uses analytical arguments based on the variation of the trajectory  $x(t)$  as the function of  $x^0$ . This idea of linearizing the change of quantities can be treated abstractly in terms of functional analysis. This functional analytic background has been developed by DUBOVITSKII and MILYUTIN [13]. The topic is discussed in details in the excellent course book of GIRSANOV [19]. Later on, most of versions of the maximum principle were proved through the Dubovitskii–Milyutin theorem. Now a very huge amount of maximum principles are known; we refer e.g. to [20], [21]. In what follows we recall a version given in [20], [21].

Namely consider the following

PROBLEM A.

$$(1.1) \quad \dot{x} = f(x, u, t) \quad f = (f_1, \dots, f_n)^T, \quad x = (x_1, \dots, x_n)^T$$

where  $x(t)$  is absolutely continuous on  $[0, T]$ , the control  $u(t) \in U$  for a.e.  $t \in [0, T]$  where  $U \subset \mathbb{R}^n$  is a fixed arbitrary subset,  $u \in L_\infty(0, T, \mathbb{R}^n)$ . Denote

$$x_0 = x(0) \quad x_T = x(T)$$

and consider the set of equality type endpoint constraints

$$(1.2) \quad K(x_0, x_T) = 0.$$

We don't specify the size  $d_k$  of the vector  $K$  (in the classical case of Pontryagin there are  $2n$  equations,  $x_0 = x^{(0)}$ ,  $x_T = x^{(1)}$ ). Introduce further some inequality type endpoint constraints

$$(1.3) \quad \varphi_i(x_0, x_T) \leq 0 \quad 1 \leq i \leq d_\varphi$$

and other constraints

$$(1.4) \quad \Phi_s(x(t), t) \leq 0 \quad 1 \leq s \leq d_\Phi.$$

We set the following assumptions

a). The functions  $K$  and  $\varphi$  are continuously differentiable in all of their variables

b) The  $\Phi_s$  are continuously differentiable in  $x$ ,  $\Phi_s$  and  $\Phi'_{s,x} (= \text{grad}_x \Phi_s)$  are continuous in  $x, t$

c)  $f$  is continuously differentiable in  $x$ , continuous in  $u$  and measurable in  $t$ .

**Problem A** is how to minimalize the functional

$$(1.5) \quad J = \varphi_0(x_0, x_T) \rightarrow \min$$

under the constraints (1.1)–(1.4) and assumptions a), b), c). The following necessary condition of optimality is given in LEVITAN, MILYUTIN, OS-MOLOVSKII [18] (in fact in [18], Theorem D.2.1 only the local version is proved, but there is a standard tool of converting local version to the above type maximum principle by making a variable substitution, see e.g. GIRSANOV [19], Chapter 12).

**THEOREM A.** Let  $\bar{x}(t)$ ,  $\bar{u}(t)$  be an optimal solution of Problem A and suppose that

$$(1.6) \quad \Phi_s(\bar{x}_0, 0) < 0, \quad \Phi_s(\bar{x}_T, T) < 0.$$

Then we can construct a function  $H = H(t, x, u)$  which is maximal in  $u$  for  $u = \bar{u}(t)$  in the sense that for almost all  $t \in [0, T]$  we have

$$(1.7) \quad H(t, \bar{x}(t), \bar{u}(t)) = \max_{u \in U} H(t, \bar{x}(t), u).$$

This function has the form

$$(1.8) \quad H = \Psi \cdot f,$$

where the function  $\Psi(t) = (\Psi_1(t), \dots, \Psi_n(t))$  is described by the following boundary value problem:

a) There exist nondecreasing functions  $\mu_s(t)$  with  $\mu_s(0) = 0$  such that

$$(1.9) \quad -\dot{\Psi} = \Psi \cdot f'_x - \sum_s \Phi'_{s,x} \cdot \frac{d\mu_s}{dt}, \quad f'_x = f'_x(\bar{x}(t), \bar{u}(t), t), \quad \Phi'_{s,x} = \Phi'_{s,x}(\bar{x}(t), t).$$

Here the rows of the matrix  $f'_x$  are the gradients of the components of  $f$  and (1.9) means that

$$(1.10) \quad \Psi(t) - \Psi(0) = - \int_0^t \Psi \cdot f'_x + \sum_s \int_0^t \Phi'_{s,x} d\mu_s.$$

b) There exist vectors  $a = (a_0, \dots, a_{d_\varphi})$ ,  $b = (b_1, \dots, b_{d_K})$  with  $a_i \geq 0$ ,  $b_j \in \mathbb{R}$  such that the boundary conditions on  $\Psi$  can be expressed by the endpoint Lagrange function

$$l = \sum_{i \geq 0} a_i \varphi_i + b \cdot K, \quad l = l(x_0, x_T)$$

by the rule

$$(1.11) \quad \Psi(0) = l'_{x_0}(\bar{x}_0, \bar{x}_T), \quad \Psi(T) = -l'_{x_T}(\bar{x}_0, \bar{x}_T).$$

c) For all  $s$  the support of the measure  $d\mu_s$  is contained in  $\{t : \Phi_s(\bar{x}(t), t) = 0\}$ . In other words, the function  $\mu_s$  is constant in every open subinterval of  $\{t : \Phi_s(\bar{x}(t), t) < 0\}$ .

d) Except for the index  $i = 0$ ,

$$(1.12) \quad a_i \varphi_i(\bar{x}_0, \bar{x}_T) = 0 \quad i \geq 1.$$

e) Nontriviality condition:

$$(1.13) \quad \sum_{i \geq 0} a_i + \sum_j |b_j| + \sum \mu_s(T) > 0.$$

REMARK 1. In the literature (1.10) is called adjoint (or co-state) equation, c) and d) is named complementary slackness (1.11) the transversality conditions,  $\lambda = (a, b, \mu(t), \Psi(t))$  is called Lagrange multiplier of the problem.

2. Like in PONTYAGIN theorem (Theorem  $A^0$ ), the cost function  $J$  does not depend only on the endpoints but on some integral involving the trajectory. E. g. in [7] the author considers

$$J = \left( \int_0^T F_1(x, u, t) dt \right)^{\alpha_1} \cdot \left( \int_0^T F_2(x, u, t) dt \right)^{\alpha_2}.$$

These functionals can easily be transformed to the form (1.5) by introducing new state variables

$$\dot{y}_1 = F_1, \quad \dot{y}_2 = F_2, \quad F_i = F_i(x(t), u(t), t)$$

and then  $J$  becomes of the form (1.5)

$$J = (y_1(T) - y_1(0))^{\alpha_1} \cdot (y_2(T) - y_2(0))^{\alpha_2}.$$

3. We formulated here the *MP* as a necessary condition for a strong local minimum. In fact [1]–[4], it is a necessary condition for the so-called Pontryagin minimum, which is a minimum with respect to variations under “Pontryagin convergence”. The latter means that

$$\|x_n - \bar{x}\|_\infty \rightarrow 0, \quad \|u_n - \bar{u}\|_\infty \leq \text{const.}, \quad \|u_n - \bar{u}\|_1 \rightarrow 0.$$

Roughly speaking, a Pontryagin minimum differs from a weak one by taking into account so-called “needle-type” variations. Obviously, a Pontryagin minimum occupies an intermediate position between a weak and strong minima.

Further we will replace the state constraints by multipoint ones with the purpose to prove a limit property. But we need (a maximum principle) necessary conditions of optimality for problems with multipoint constraints. We give them in the next Section.

4. If  $\lambda$  is a multiplier in Theorem A and  $c > 0$  then  $c\lambda$  is also a multiplier because then  $\Psi$  becomes  $c\Psi$  and  $H$  becomes  $cH$ .

## 2. A maximum principle for problems with multipoint constraints.

Consider here the following problem (**Problem B**):

$$(2.1) \quad \dot{x} = f(x, u, t), \quad u \in U$$

$$(2.2) \quad K(x_0, x_1, \dots, x_N) = 0,$$

$$(2.3) \quad \Phi_i(x_0, x_1, \dots, x_N) \leq 0, \quad i = 1, \dots, d(\Phi)$$

$$(2.4) \quad J = \Phi_0(x_0, x_1, \dots, x_N) \rightarrow \min.$$

where  $x_k = x(t_k)$ ,  $k = 1, \dots, N$ ,  $0 = t_0 < t_1 < t_2 < \dots < t_{N-1} < t_N = T$ ,  $K$ ,  $\Phi_i$  are continuously differentiable and all the other objects are the same as in Section 1.

The Problem B seemingly doesn't belong to the class of Problems A, but however, it can be easily reduced to the problem A, as we'll show below. Now we formulate necessary optimality conditions (a maximum principle) for Problem B.

**THEOREM 1.** *Let  $(x^0, u^0)$  be a Pontryagin minimum point of Problem B. Then there exist a collection of Lagrange multipliers  $(a, b, \Psi(t))$ , of the same nature as in Theorem A, such that with the Lagrange function  $l = a\phi + bK$  and Pontryagin function  $H = \Psi f$  the following conditions hold:*

a) *nontriviality condition*

$$(2.5) \quad |a| + |b| > 0,$$

b) *complementary slackness (1.10);*

c) *adjoint equation of the form:*

$$(2.6) \quad -\dot{\Psi} = \Psi f'_x$$

and

$$(2.7) \quad \Delta\Psi(t_k) := \Psi(t_k + 0) - \Psi(t_k - 0) = l_{x_k}$$

on every interval  $(t_{k-1}, t_k)$ ,  $k = 1, \dots, N - 1$ ;

d) *transversality conditions (1.13);*

e) *maximum conditions (1.14).*

PROOF OF THEOREM 1. Denote  $\Delta_k = [0, t_k]$ ,  $k = 1, \dots, N - 1$ , and introduce new state variables  $y_k$  subject to equations:

$$(2.8) \quad \dot{y}_k = f(x, u, t)\chi_{\Delta_k}(t), \quad k = 1, \dots, N - 1.$$

Then obviously  $x(t_k) = x_0 + y_k(t) - y_k(0)$ , furthermore --

$$(2.9) \quad \phi(x_0, x_1, \dots, x_T) = \phi(x_0, x_0 + y_1(T) - y_1(0), \dots, x_0 + y_k(T) - y_k(0), \dots, x_T)$$

and a similar form takes  $K$ .

Thus, the multipoint constraints with interior time points are reduced to standard endpoint constraints with respect to the old ( $x$ ) and new ( $y$ ) state variables. Therefore, we can apply Theorem A to this reduced problem.

According to Theorem A there exist Lagrange multipliers  $a$ ,  $b$ ,  $\Psi(t)$  of the above mentioned nature, and functions  $\eta_k(t)$  (related to new state variables), such that  $|a| + |b| > 0$ , and with

$$l = a\phi + bK,$$

$$\tilde{l}(x_0, y_1, \dots, y_{N-1}, x_T) = l(x_0, \dots, x_0 + y_k(T) - y_k(0), \dots, x_T),$$

$$(2.10) \quad H = \Psi f + \sum_1^{N-1} \eta_k f \chi_{\Delta_k}$$

the conditions (1.10)–(1.14) holds:

Let us consider the adjoint equations and transversality conditions. Since  $\dot{\eta}_k = -H_y = 0$ , we have  $\eta_k = \text{const}$ , and

$$(2.11) \quad \eta_k = \eta_k(0) = \frac{\partial \tilde{l}}{\partial y_k(0)} = -l_{x_k}.$$

The adjoint equation for  $\Psi$  is:

$$(2.12) \quad -\dot{\Psi} = \left( \Psi + \sum_1^{n-1} \eta_k \chi_{\Delta_k} \right) f'_x,$$

while

$$(2.13) \quad \begin{cases} \Psi(0) = \tilde{l}_{x_0} = l_{x_0} + l_{x_1} + \dots + l_{x_{N-1}}, \\ \Psi(T) = -\tilde{l}_{x_T} = -l_{x_T}. \end{cases}$$

Introduce the new function

$$(2.14) \quad \tilde{\Psi} = \Psi + \sum_1^{N-1} \eta_x \chi_{\Delta_k}.$$

Then from (2.12) it follows that on every interval  $(t_{k-1}, t_k)$ ,  $k = 1, \dots, N$

$$(2.15) \quad -\dot{\tilde{\Psi}} = \tilde{\Psi}f$$

and

$$(2.16) \quad \begin{cases} \Delta\tilde{\Psi}(t_k) = -\eta_k = l_{x_k}, \\ \tilde{\Psi}(0) = \Psi(0) + \sum \eta_k = l_{x_0}, \\ \tilde{\Psi}(T) = \Psi(T) = l_{x_T}. \end{cases}$$

Finally, since  $H$  satisfies the maximum condition with respect to  $u$ , the function  $\tilde{H} = \tilde{\Psi}f$  satisfies it too.

Theorem 1 is proved. ■

Let us now consider a special case of Problem  $B$ , which arises from the Problem  $A$  if we consider the state constraint (1.2) not at the all points  $t \in [0, T]$ , but only at some a priori chosen and fixed points  $t_1, \dots, t_{N-1}$ . That is, instead of (1.2) we have now the following constraints:

$$(2.17) \quad \Phi(x_k, t_k) \leq 0, \quad k = 1, \dots, N-1.$$

Denote this problem as **Problem  $\tilde{B}$** .

**THEOREM 2.** *Let  $(x^0, u^0)$  be a Pontryagin minimum point in Problem  $\tilde{B}$ . Then there is a collection of Lagrange multipliers  $\tilde{\lambda} = (a, b, \mu(t), \Psi(t))$  of the same nature as in Theorem A, such that all conditions (1.9)–(1.14) are valid with the additional property: function  $\mu(t)$  is constant on every interval  $(t_{k-1}, t_k)$   $k = 1, \dots, N$  (and so, it can generate only  $\delta$ -functions in the right hand side of the adjoint equation (1.12)).*

**PROOF OF THEOREM 2.** According to Theorem 1, there exist  $a \geq 0$ ,  $b \in \mathbb{R}^{d(K)}$ ,  $\xi = (\xi_1, \dots, \xi_{N-1}) \geq 0$ , such that

$$(2.18) \quad |a| + |b| + |\xi| > 0,$$

and with  $l = a\phi + bK + \sum_1^{N-1} \xi_k \Phi(x_k, t_k)$ ,  $H = \Psi f$  the corresponding conditions hold, from which we have to consider in detail only the complementary slackness condition

$$(2.19) \quad \xi_k \Phi(x_k^0, t_k) = 0, \quad k = 1, \dots, N-1$$

and the adjoint equation (2.6) with jump conditions

$$(2.20) \quad \Delta\Psi(t_k) = l_{x_k} = \xi_k \Phi'_x(x_k, t_k), \quad k = 1, \dots, N-1.$$

Denote  $\bar{\Delta}_k = [t_k, T]$  and consider the vector-function

$$(2.21) \quad \mu(t) = \sum_{k=1}^{N-1} \xi_k \chi_{\bar{\Delta}_k}(t),$$

every component of which is obviously nondecreasing.

Then (2.19) is equivalent to (1.11), and (2.6) together with (2.20) — to (1.12), whereas (2.18) is equivalent to (1.9).

Theorem 2 is proved. ■

### 3. A limit approach in the multipoint maximum principle

Let us now return to initial Problem *A*, and try to find it's solution in the following way. We fix an arbitrary finite set  $\tau$  of time points  $0 = t_0 < t_1 < \dots < t_{N-1} < t_N = T$  and consider **Problem *B*( $\tau$ )** of the form (1.1), (1.2) – (1.6), (2.9), which differs from Problem *A* by replacement of the state constraints

$$\Phi_s(x(t), t) \leq 0, \quad t \in [0, T]$$

by constraints on the state variable only at the chosen points:

$$\Phi_s(x(t_k), t_k) \leq 0, \quad k = 1, \dots, N - 1.$$

We assume that there is an optimal pair  $(x_\tau^0, u_\tau^0)$  in the Problem *B*( $\tau$ ), and that it converges in uniform and weak-\* topology (resp.) to a pair  $(x^0, u^0)$  as the number of points  $t_k$  grows in such a manner that the length of the largest interval of  $[t_{k-1}, t_k]$  tends to zero. Since  $u_\tau$  converge only in weak-\* topology, hence we should assume that the system (1.1), i.e. the function *f* is linear with respect to *u*. A limit property of the pair is then described in the next

**THEOREM 3.** *Assume that for every  $\tau$  Problem *B*( $\tau$ ) has an optimal pair  $(x_\tau^0, u_\tau^0)$ . Assume further that there are  $x^0, u^0$  such that as  $\tau$  grows*

$$(3.3) \quad x_\tau^0 \rightarrow x^0 \text{ uniformly and } u_\tau^0 \xrightarrow{\text{weak-*}} u^0.$$

*Assume that there exists a point  $t_*$  at which*

$$(3.4) \quad \Phi_s(x(t_*), t_*) < 0, \quad s = 1, \dots, d(\Phi).$$

*Then the pair  $(x^0, u^0)$  satisfies the MP for Problem *A*.*

*Along with this, if  $\lambda_\tau = (a_\tau, b_\tau, \mu_\tau, \Psi_\tau)$  are Lagrange multipliers with  $|a_\tau| + |b_\tau| + |\mu_\tau(T)| = 1$  for a pair  $(x_\tau^0, u_\tau^0)$  in the sense of Theorem 2, then it is possible to take as a collection of Lagrange multipliers for  $(x^0, u^0)$  any*

limit point of the sequence (net)  $\lambda_\tau$  regarding the convergence of  $a_\tau, b_\tau$  as finite-dimensional vectors and  $\mu_\tau, \Psi_\tau$  almost everywhere.

**PROOF OF THEOREM 3.** We apply Theorem 2 to the Problem  $B(\tau)$ .

Since all the sequences  $a_\tau, b_\tau, \mu_\tau(T)$  are uniformly bounded, hence passing if necessary to subsequence, we can consider  $a_\tau \rightarrow a, b_\tau \rightarrow b$ , where  $a, b$  are some vectors, and  $\mu_\tau(t) \rightarrow \mu(t)$  a.e., where  $\mu$  is a nondecreasing function with  $\mu(0) = 0$ .

Moreover, it is true that  $\mu_\tau(t) \rightarrow \mu(t)$  at any point of continuity of  $\mu$ . Similarly, since  $\Psi_\tau$  form a uniformly bounded sequence of function of bounded variations, we can suppose that there is a function  $\Psi(t)$  of bounded variation such that  $\Psi_\tau(t) \rightarrow \Psi(t)$  at any point of continuity of  $\Psi$  (and hence, a.e.).

It is clear that the normalization and complementary slackness conditions hold also for  $(a_0, b_0, \mu_0)$  since they hold for  $(a_\tau, b_\tau, \mu_\tau)$ .

From here and (3.4) it follows that in a neighbourhood of  $t_*$   $\mu_s(t) = \text{const.}$   $\forall s = 1, \dots, d(\Phi)$ , and hence,  $\Psi$  is continuous, which implies that

$$(3.5) \quad \Psi_\tau(t_*) \rightarrow \Psi(t_*).$$

To establish a limit property of the adjoint equation we write it in an integral form:

$$(3.6) \quad \Psi_\tau(t) = \Psi_\tau(t_*) + \int_0^t \Psi_\tau(s) f'_x(x_\tau^0(s), u_\tau^0(s), s) ds - \int_0^t \Phi'_x(x_\tau^0(s), s) d\mu_\tau(s).$$

From (3.3), (3.5), the linearity of  $f_x$  in  $u$  and the convergence a.e. of  $\Psi_\tau, \mu_\tau$  it can be shown that

$$(3.7) \quad \Psi_\tau(t) = \Psi_0(t_*) + \int_0^t \Psi(s) f'_x(x^0(s), u^0(s), s) ds - \int_0^t \Phi'_x(x^0(s), s) d\mu(s),$$

which is the adjoint equation in problem  $A$  for the trajectory  $(x^0, u^0)$ . It remains only to check the maximality condition. Since it holds for  $(x_\tau^0, u_\tau^0)$ , we know that for any measurable  $u(t) \in U$ , for all  $\tau$

$$(3.8) \quad \int_0^T \Psi_\tau(t) \left[ f(x_\tau^0(t), u(t), t) - f(x_\tau^0(t), u_\tau^0(t), t) \right] dt \leq 0.$$

From here, the linearity of  $f$  in  $u$ , the weak-\* convergence of  $u_\tau^0$  and a.e. convergence of  $\Psi_\tau$  it follows that

$$(3.9) \quad \begin{cases} \int_0^T \Psi_\tau(t) \left[ f(x^0(t), u(t), t) - f(x^0(t), u^0(t), t) \right] dt = \\ = \int_0^T \left( H(x^0(t), u(t), t) - H(x^0(t), u^0(t), t) \right) dt \leq 0. \end{cases}$$

From this and the measurable choice theorem [6, Section 8.1] it follows that for a.e.  $t$  the function  $H$  satisfies the maximality condition (1.14).

Theorem 3 is proved. ■

#### 4. Examples

In this part we provide a few examples illustrating the above theorem.

EXAMPLE 1. For fixed values  $h$ ,  $T > 0$  consider the problem

$$\begin{aligned} \dot{x} &= u, \\ x(0) &= x(T) = 0, \\ |u(t)| &\leq 1 \quad \forall t \in [0, T], \\ x(t) &\leq h \quad \forall t \in [0, T]. \end{aligned}$$

Under these assumptions maximize the area below  $x(t)$ :

$$J = \int_0^T x \rightarrow \max.$$

The direct solution of this problem is easy; the optimal trajectory can be given for  $T \geq 2h$  by

$$x(t) = \begin{cases} t & t \leq h, \\ h & h \leq t \leq T-h, \\ T-t & T-h \leq t, \end{cases}$$

and for  $T \leq 2h$  by

$$x(t) = \begin{cases} t & t \leq \frac{T}{2}, \\ T-t & t \geq \frac{T}{2}. \end{cases}$$

Now we interpret this problem as Problem A and solve it by Theorem A. To transform the cost function  $J$  we need a new state function

$$y(t) = - \int_0^t x.$$

Then the Problem A equivalent to Example 1 reads as follows

$$\dot{x} = u = f_1(t, x, y, u),$$

$$\dot{y} = -x = f_2(t, x, y, u),$$

$$K_1(x_0, x_T, y_0, y_T) = x_0 = 0,$$

$$K_2(x_0, x_T, y_0, y_T) = y_0 = 0,$$

$$K_3(x_0, x_T, y_0, y_T) = x_T = 0,$$

$$u(t) \in U = [-1, 1] \text{ for all } t \in [0, T],$$

$\Phi_1(x, y, t) = x - h$ ; i.e.  $x(t) - h \leq 0$  for all  $t \in [0, T]$ . Under these assumptions minimize  $y_T$ :

$$J = \varphi_0(x_0, x_T, y_0, y_T) = y_T \rightarrow \min.$$

Here Theorem A applies since

$$\Phi_1(x(0), y(0), 0) = \Phi_1(x(T), y(T), T) = -h < 0.$$

By Theorem A there exist  $a_0 \geq 0$ ,  $b_1, b_2, b_3 \in \mathbb{R}$ , a monotone increasing function  $\mu_1(t)$  with  $\mu_1(0) = 0$ ,  $\mu_1$  is left-continuous and there are functions  $\Psi_1, \Psi_2$  satisfying the following properties:

a)  $a_0 + |b_1| + |b_2| + |b_3| + \mu_1(T) > 0$ ,

b)  $\Phi_1(x(t), y(t), t) d\mu_1(t) \equiv 0$ ,

which means that  $\mu_1$  is constant on the open intervals where  $x(t) < h$ ,

c)  $-(\dot{\Psi}_1, \dot{\Psi}_2) = (\Psi_1, \Psi_2) \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} - (1 \ 0) \frac{d\mu_1}{dt}$  i.e.

$$-\dot{\Psi}_1 = -\Psi_2 - \frac{d\mu_1}{dt},$$

$$-\dot{\Psi}_2 = 0.$$

d) Let  $l(x_0, x_T, y_0, y_T) = a_0 y_0 + b_1 K_1 + b_2 K_2 + b_3 K_3 = a_0 y_T + b_1 x_0 + b_2 y_0 + b_3 x_T$ .

Then

$$\Psi_1(0) = l'_{x_0} = b_1, \quad \Psi_1(T) = -l'_{x_T} = -b_3,$$

$$\Psi_2(0) = l'_{y_0} = b_2, \quad \Psi_2(T) = -l'_{y_T} = -a_0.$$

By c) this implies that

$$(4.1) \quad \Psi_2 \equiv -a_0, \quad -a_0 = b_2,$$

$$(4.2) \quad \Psi_1(t) = -a_0 t + \mu_1(t) + b_1$$

and  $\Psi_1(T) = -b_3$  implies

$$(4.3) \quad b_3 = a_0 T - \mu_1(T) - b_1.$$

Consequently from a) it remains

$$(4.4) \quad a_0 + |b_1| + \mu_1(T) > 0.$$

e) The function  $H = \Psi_1 u - \Psi_2 x$  takes its maximum in  $u = u(t)$  under the assumption  $|u| \leq 1$ . This means that

$$(4.5) \quad \begin{cases} \Psi_1(t) > 0 \Rightarrow u(t) = 1, \\ \Psi_1(t) < 0 \Rightarrow u(t) = -1 \end{cases}$$

LEMMA 1.  $a_0 > 0$ .

PROOF. Suppose indirectly that  $a_0 = 0$ .

a) If  $b_1 > 0$  then  $\Psi_1(t) \geq b_1 > 0$  i.e.  $\dot{x} = u = 1$  on  $(0, T)$ ; this contradicts  $x(0) = x(T) = 0$ .

b) If  $b_1 = 0$  then by (4.4)  $\mu_1(T) > 0$ . Since  $\mu_1$  is left-continuous, there exist a minimal value  $t_0$  such that on  $(t_0, T]$   $\mu_1(t) = \Psi_1(t) > 0$ . In this segment  $\dot{x} = 1$  which implies from  $x(T) = 0$  that  $x(t_0) < 0$ . Hence in the neighbourhood of  $t_0$ ,  $\mu_1$  is constant which contradicts the definition of  $t_0$  unless  $t_0 = 0$ . But if  $t_0 = 0$  then  $\Psi_1 < 0$  on  $(0, T)$  which is impossible, see a).

c) If  $b_1 < 0$  then  $\Psi_1(t) = b_1 < 0$  until  $x(t)$  reaches  $h$ . But starting from  $t = 0$ ,  $\dot{x} = -1$  i.e.  $x(t)$  is negative, hence  $\mu_1(t) \equiv 0$  and  $\dot{x} = -1$  on  $(0, T)$  and this contradicts to  $x(0) = x(T) = 0$ . ■

Dividing all the constants and functions by  $a_0$  we can suppose that  $a_0 = 0$  hence

$$\Psi_1(t) = b_1 - t + \mu_1(t).$$

LEMMA 2. If  $x(t_1) = x(t_2) = h$  then on  $[t_1, t_2]$   $x(t) = h$ .

PROOF. Suppose the contrary. Changing the values  $t_1, t_2$  if necessary we can obtain that

$$x(t_1) = x(t_2) = h \quad \text{and on } (t_1, t_2) \quad x(t) < h.$$

This means that on  $(t_1, t_2)$   $\mu_1$  is constant and

$$\Psi_1(t) = b_1 - t + \mu_1(t_2).$$

This is a decreasing linear function. Since  $x(t) \leq h$  everywhere,  $\dot{x}(t_1) = \dot{x}(t_2) = 0$  and hence  $\Psi_1(t_1) = \Psi_1(t_2) = 0$ ; contradiction. ■

LEMMA 3. If  $T \geq 2h$  then there exists  $t$  with  $x(t) = h$ .

PROOF. Indirect:  $x < h$ . Then  $\mu_1 \equiv 0$ ,  $\Psi_1(t) = b_1 - t$ . It can not have constant sign in  $(0, T)$  since  $x(0) = x(T) = 0$  and by the same reason we must have  $b_1 = \frac{T}{2}$ . But then  $x\left(\frac{T}{2}\right) = \frac{T}{2} \geq h$ , contradiction. ■

REMARK. The above proof shows the optimal trajectory for  $T < 2h$ .

LEMMA 4. If  $T \geq 2h$  then  $b_1 = h$ .

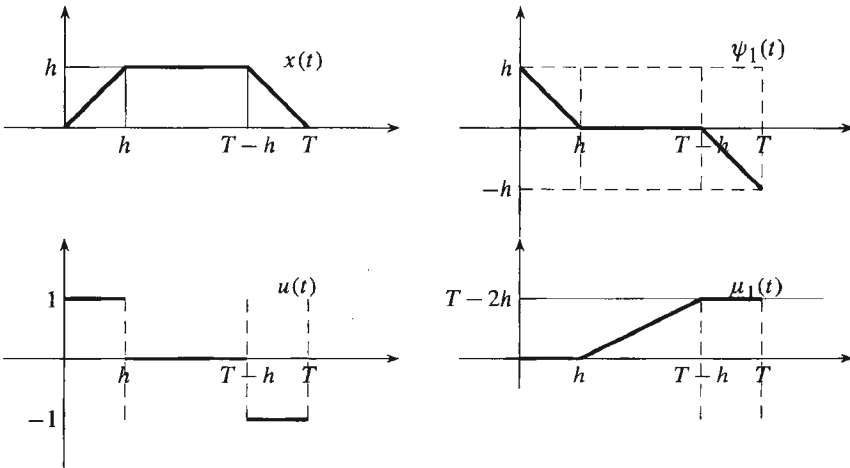
PROOF. a) If  $b_1 < h$  then on  $(0, b_1)$   $x(t)$  is growing, on  $(b_1, h)$  it is decreasing, hence can not reach  $h$ , contradiction.

b) If  $b_1 > h$  then  $\Psi_1(h) > 0$ ,  $\dot{x}(h) = 1$ ,  $x(h) = h$  which is again a contradiction. ■

LEMMA 5. If  $T \geq 2h$  then  $\mu_1(T) = T - 2h$ .

PROOF. Symmetrical to Lemma 4. ■

Now the optimal trajectory on  $[0, t]$  is  $x(t) = t$  with  $\mu(t) \equiv 0$ ,  $\Psi_1(t) = h - t$  by Lemma 4; on  $(T - h, T]$   $x(t) = T - t$ ,  $\mu(t) = T - 2h$ ,  $\Psi_1(t) = T - h - t$ ; on  $(h, T - h]$  by Lemma 2  $x(t) = h$ , hence  $\Psi_1(t) = 0$ ,  $\mu(t) = t - h$ . The results are stretched in the following figures.



EXAMPLE 2. Consider the following discretized version of Example 1. Let  $N \in \mathbb{N}$  and

$$t_k = h \frac{T}{N} \quad k = 0, \dots, N.$$

Modify Example 1 such that  $x(t) \leq h$  be assumed only at the nodes  $t = t_h$ . We give interpretation of this as a Problem  $\bar{B}$ .

By Theorem 2 all the statements (4.1)–(4.5) remain valid with the additional assumption

$$(4.6) \quad \mu_1(t) = \sum_{\substack{k=1 \\ t_k < t}}^{N-1} a_k$$

where

$$(4.7) \quad a_k \geq 0, \quad x(t_k) < h \rightarrow a_k = 0 \quad (k = 1, \dots, N - 1).$$

By immediate considerations we can describe the optimal trajectory  $x(t)$  as follows. Draw the line segments  $\left[ (0, 0), \left( \frac{T}{2}, \frac{T}{2} \right) \right]$  and  $\left[ \left( \frac{T}{2}, \frac{T}{2} \right), (T, 0) \right]$ . If strictly below this polygon there is no point  $(t_k, h)$  then this is the graph of the optimal trajectory. Otherwise let  $h_1$  and  $h_2$  be defined by

$$t_{k_1} \leq h < t_{k_1+1}, \quad t_{k_2-1} < T - h \leq t_{k_2};$$

then  $h_1 + 2 \leq h_2$  and

$$\begin{aligned} \dot{x}(t) &= 1 \text{ on } \left( 0, \frac{h+t_{k_1+1}}{2} \right) \\ &= -1 \text{ on } \left( \frac{h+t_{k_1+1}}{2}, t_{k_1+1} \right) \\ &= 1 \text{ on } \left( t_k, \frac{t_k+t_{k+1}}{2} \right) \\ &= -1 \text{ on } \left( \frac{t_k+t_{k+1}}{2}, t_{k+1} \right) \\ &= 1 \text{ on } \left( t_{k_2-1}, \frac{t_{k_2-1}+T-h}{2} \right) \\ &= -1 \text{ on } \left( \frac{t_{k_2-1}+T-h}{2}, T \right). \end{aligned} \quad \left. \vphantom{\begin{aligned} \dot{x}(t) &= 1 \text{ on } \left( 0, \frac{h+t_{k_1+1}}{2} \right) \\ &= -1 \text{ on } \left( \frac{h+t_{k_1+1}}{2}, t_{k_1+1} \right) \\ &= 1 \text{ on } \left( t_k, \frac{t_k+t_{k+1}}{2} \right) \\ &= -1 \text{ on } \left( \frac{t_k+t_{k+1}}{2}, t_{k+1} \right) \right\} k = k_1 + 1, \dots, k_2 - 2$$

Now the same result can be obtained by Theorem 2 as follows

LEMMA 6.  $a_0 > 0$ .

PROOF. The same as in Lemma 1. ■

Consequently we can suppose

$$a_0 = 1, \quad \Psi_1(t) = b_1 - t + \mu_1(t).$$

LEMMA 7. If  $x(t_{k_1}) = x(t_{k_2}) = h$  then for  $k_1 \leq k \leq k_2$ ,  $x(t_k) = h$ .

PROOF. Suppose the contrary. Redefining  $k_1$  and  $k_2$  we can obtain that  $x(t_{k_1}) = x(t_{k_2}) = h$  but  $x(t_k) < h$  for  $k_1 < k < k_2$ . Then in  $(k_1, k_2)$   $\mu_1$  is constant hence the proof of Lemma 2 applies. ■

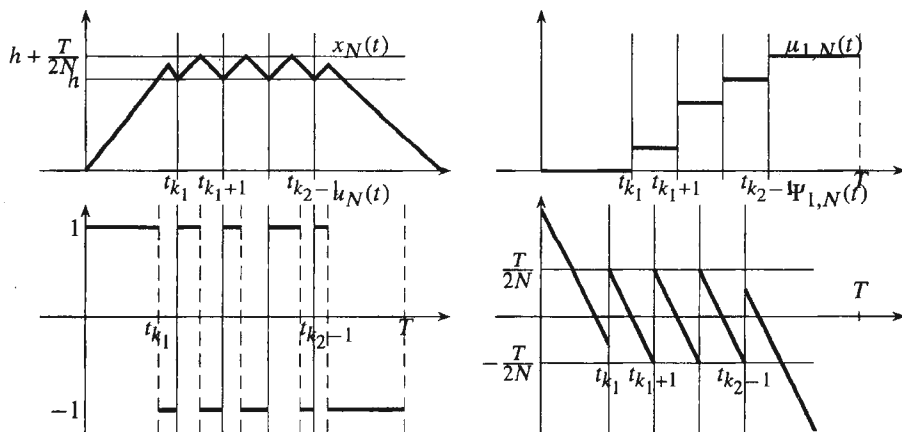
LEMMA 8.  $h \leq b_1 < t_{k_1+1}$  if  $T \geq 2h$ .

PROOF. a) If  $h > b_1$  then on  $(0, h)$   $\Psi_1(t) = b_1 - t$  hence for  $t \geq b_1$   $x(t)$  becomes decreasing before reaching  $h$ . This implies that  $\mu_1 \equiv 0$ ,  $\Psi_1 = b - t$

on  $(0, T)$  and then  $\dot{x} = 1$  on  $(0, b_1)$ ,  $\dot{x} = -1$  on  $(b_1, T)$ . From  $x(0) = x(T) = 1$  it implies that  $b_1 = \frac{T}{2}$  which contradicts  $b_1 < h$  if  $T \geq 2h$ .

b) If  $t_{k_1+1} \leq b_1$  then we argue as follows:  $\Psi_1(t) \geq b_1 - t > 0$  if  $t \in (0, b_1)$ , hence  $\dot{x} = 1$  on  $(0, b_1)$  and then  $x(t_{k_1+1}) = t_{k_1+1} > h$ , contradiction. ■

Now we can build the optimal trajectory. Suppose for the case of simplicity that  $h \neq t_{k_1}$ . Then on  $(0, t_{k_1+1})$  we have  $\Psi_1(t) = b_1 - t$ . Hence approaching  $t_{k_1+1}$  from the left  $x(t)$  becomes decreasing. If  $x(t_{k_1+1}) < h$ , then  $\Psi_1(t) = b_1 - t$  on  $(0, t_{k_1+2})$  and  $x(t)$  continues decreasing, i.e.  $\dot{x} = -1$  on  $(t_{k_1+1}, T)$ . This is possible only if such a trajectory ends at  $x(T) = 0$ . Otherwise we must have  $x(t_{k_1+1}) = h$  and on  $(t_{k_1+1}, t_{k_1+2})$   $\Psi_1(t) = b_1 + a_1 - t$  with some  $a_1 > 0$ . Now the value  $\Psi_1(t)$  in the halving point of the segment  $(t_{k_1+1}, t_{k_1+2})$  is  $\leq 0$ , otherwise  $x(t_{k_1+2}) > h$ . If this value is negative then  $x(t_{k_1+2}) < h$  hence  $x$  continues decreasing in  $(t_{k_1+2}, T)$  and this can be so only if this trajectory ends in  $x(T) = 0$ . Otherwise  $\Psi_1(t)$  will be positive in the left half and negative in the right half of  $(t_{k_1+1}, t_{k_1+2})$  and  $x(t_{k_1+2}) = h$ . Then we repeat the same arguments for the segment  $(t_{k_1+2}, t_{k_1+3})$  and so on. We can thus build up the optimal trajectory described above. In order to distinguish it from that of Example 1 we use the notation  $x_N(t)$ ,  $\Psi_{1,N}(t)$ ,  $\mu_{1,N}(t)$ ,  $u_N(t)$  instead of  $x(t)$ ,  $\Psi_1(t)$ ,  $\mu_1(t)$  and  $u(t)$ . The obtained result are drawn in the following figures.



Since  $x_N(t) \rightarrow x(t)$  uniformly and  $u_N(t) \rightarrow u(t)$  weakly, Theorem 3 can be applied. It can be seen that  $\Psi_{1,N} \rightarrow \Psi_1(t)$  and  $\mu_{1,N}(t) \rightarrow \mu_1(t)$ , both uniformly. On the other hand,  $\Psi_{2,N} \equiv -1 \equiv \Psi_2$ ,  $-1 = b_{2,N} = b_2$ ,  $b_{1,N} = \Psi_{1,N}(0) \rightarrow \Psi_1(0) = b_1$ ,  $b_{3,N} = T - \mu_{1,N}(T) - b_{1,N} \rightarrow T - \mu_1(T) - b_1 = b_3$  i.e. the Lagrange multiplier has indeed the limit property stated in Theorem 3.

EXAMPLE 3. Consider the problem

$$\begin{aligned} \dot{x} &= y, \\ \dot{y} &= u, \\ x(0) &= x(T) = 0, \\ x(t) &\leq h \quad \forall t \in [0, T], \\ |u(t)| &\leq 1 \quad \forall t \in [0, T]. \end{aligned}$$

Under these circumstances maximize the area below  $x$ :

$$J = \int_0^T x \rightarrow \max.$$

Rewrite this as Problem A: Introduce the variable  $z = -\int_0^t x$ ; then

$$\begin{aligned} \dot{x} &= y = f_1(t, x, y, z, u), \\ \dot{y} &= u = f_2(t, x, y, z, u), \\ \dot{z} &= -x = f_3(t, x, y, z, u), \\ K_1 &= x_0 = 0, \\ K_2 &= z_0 = 0, \\ K_3 &= x_T = 0, \\ u(t) &\in U = [-1, 1], \\ \Phi(x, y, z, t) &= x - h \quad \text{i.e. } x(t) - h \leq 0 \quad \forall t \in [0, T]. \end{aligned}$$

We minimize  $z_T$   $J = \varphi_0 = z_T \rightarrow \min$ .

It can be directly seen that the optimal solution is for  $T < 2\sqrt{2h}$  the concave parabola going through  $(0, 0)$  and  $(T, 0)$ ; for  $T \geq 2\sqrt{2h}$  we have

a) in  $[0, \sqrt{2h}]$   $x(t)$  is the concave parabola going through  $(0, 0)$  whose tangent line at  $t = \sqrt{2h}$  is the line  $y = h$ ,

b) in  $[\sqrt{2h}, T - \sqrt{2h}]$   $x(t) = h$ ,

c) in  $[T - \sqrt{2h}, T]$   $x(t) = x(T - t)$ .

By Theorem A, there exist  $a_0 \geq 0$ ,  $b_1, b_2, b_3 \in \mathbb{R}$ , an increasing function  $\mu_1(t)$  with  $\mu_1(0) = 0$ , which is left-continuous and functions  $\Psi_1, \Psi_2, \Psi_3$  with the properties

a)  $a_0 + |b_1| + |b_2| + |b_3| + \mu_1(T) > 0$ .

b)  $\mu_1(t)$  is constant on the open intervals where  $x(t) < h$ .

$$c) -(\dot{\Psi}_1, \dot{\Psi}_2, \dot{\Psi}_3) = (\Psi_1, \Psi_2, \Psi_3) \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} - (1, 0, 0) \frac{d\mu_1}{dt} \quad \text{i.e.}$$

$$-\dot{\Psi}_1 = -\Psi_3 - \frac{d\mu_1}{dt}, \quad -\dot{\Psi}_2 = \Psi_1, \quad -\dot{\Psi}_3 = 0.$$

d) Denote  $l = a_0 z_T + b_1 x_0 + b_2 z_0 + b_3 x_T$ ; then

$$\begin{aligned} \Psi_1(0) &= b_1 & \Psi_1(T) &= -b_3 \\ \Psi_2(0) &= 0 & \Psi_2(T) &= 0 \\ \Psi_3(0) &= b_2 & \Psi_3(T) &= -a_0. \end{aligned}$$

This implies by c) that

$$(4.8) \quad -a_0 = b_2, \quad \Psi_3 \equiv -a_0$$

$$(4.9) \quad \Psi_1(t) = b_1 - a_0 t + \mu_1(t), \quad -b_3 = \Psi_1(T) = b_1 - a_0 T + \mu_1(T)$$

$$(4.10) \quad \Psi_2(t) = a_0 \frac{t^2}{2} - b_1 t - \int_0^t \mu_1, \quad 0 = \Psi_2(T) = a_0 \frac{T^2}{2} - b_1 T - \int_0^T \mu_1$$

and hence by a),

$$(4.11) \quad a_0 + |b_1| > 0, \quad a_0 + \mu(T) > 0.$$

e) The function  $H = \Psi_1 y + \Psi_2 u - \Psi_3 x$  is maximal in  $u$  provided that  $|u| \leq 1$ , hence

$$(4.12) \quad \begin{cases} \Psi_2(t) > 0 \Rightarrow u(t) = 1, \\ \Psi_2(t) < 0 \Rightarrow u(t) = -1. \end{cases}$$

LEMMA 9.  $a_0 > 0$  if  $T > 2\sqrt{2h}$ .

PROOF. If  $a_0 = 0$  then  $\dot{\Psi}_2 = -\Psi_1$  is decreasing, hence  $\Psi_2$  is concave. Since  $\Psi_2(0) = \Psi_2(T) = 0$ , there are two possibilities

a)  $\Psi_2 \equiv 0$ . Then  $b_1 = 0$ , contradiction with (4.11).

b)  $\Psi_2 > 0$  on  $(0, T)$ ; then  $\ddot{x} = 1$ ,  $x(t)$  is convex, hence  $x(t) < 0$  on  $(0, T)$  which is obviously not optimal. ■

Suppose that  $a_0 = 1$ .

LEMMA 10. If  $x(t_1) = x(t_2) = h$  then  $x(t) = h$  on  $[t_1, t_2]$ .

PROOF. Suppose indirectly that  $x(t_1) = x(t_2) = h$  and that on  $(t_1, t_2)$   $x(t) < h$ . Then on  $[t_1, t_2]$   $\Psi_2$  is a convex parabola. Since  $x(t_1) = h$ , we have  $\ddot{x}(t_1) \leq 0$  (otherwise  $x(t) > h$  somewhere near  $t_1$ ), hence  $\Psi_2(t_1) \leq 0$ . By the

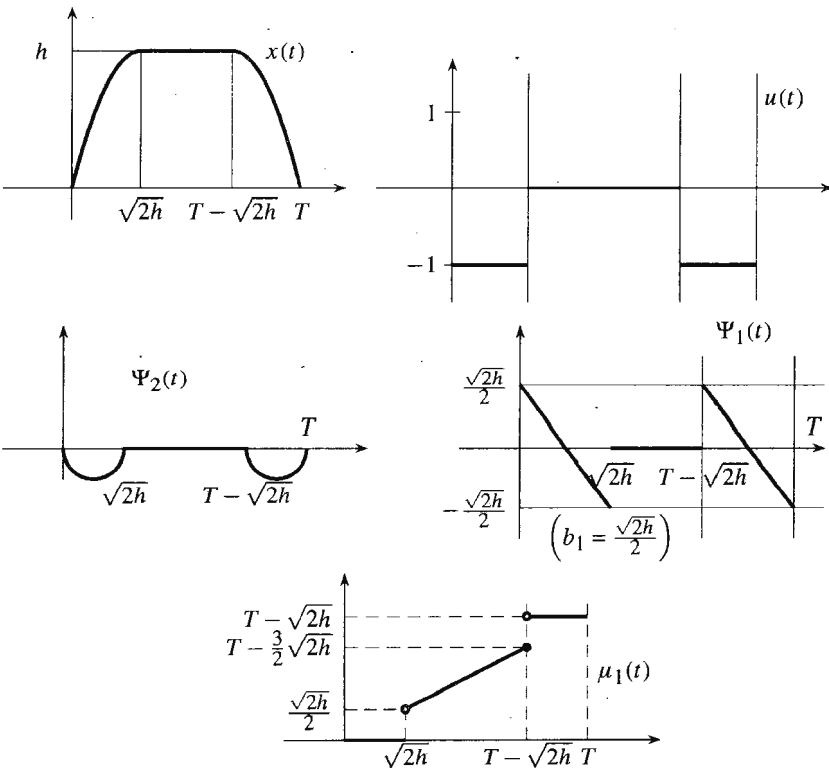
same reason  $\Psi_2(t_2) \leq 0$ . Taking into account the convexity of  $\Psi_2$ , we get that  $\Psi_2 < 0$  on  $(t_1, t_2)$  i.e.  $\ddot{x} = -1$ ,  $x$  is concave on  $[t_1, t_2]$ . But then  $x(t)$  goes over  $h$  which is not allowed. Lemma 10 is proved. ■

LEMMA 11. For  $T \geq 2\sqrt{2h}$  there exists  $t$  with  $x(t) = h$ .

PROOF. If it is not so,  $\Psi_2 = \frac{1}{2}t^2 - b_1t < 0$  on  $(0, T)$ , hence  $x(t)$  is a concave parabola through  $(0,0)$  and  $(T,0)$ . But this is impossible, since its maximal height is  $\geq h$ . ■

So if  $T > 2\sqrt{2h}$  then there exists  $0 < t_1 \leq t_2 < T$  such that in  $(0, t_1)$  and  $(t_2, T)$ ,  $x(t) < h$  and in  $[t_1, t_2]$ ,  $x(t) = h$ . In  $(0, t_1)$   $\Psi_2$  is a convex parabola,  $0 = \Psi_2(0)$ ,  $0 \geq \Psi_2(t_1)$ , hence  $0 > \Psi_2$  on  $(0, t_1)$ . This implies that  $\ddot{x}(t) = -1$ ,  $x$  is a concave parabola going through  $(0,0)$ ,  $(t_1, h)$ .

Since  $x(t)$  is maximal in  $t_1$ ,  $\dot{x}(t) = 0$ . This defines uniquely  $t_1 (= \sqrt{2h})$  and the trajectory  $x(t)$ . On  $(t_2, T)$  we can argue symmetrically. Since  $\Psi_2$  is continuous,  $\Psi_2(t_1) = 0$  (otherwise  $\ddot{x} = -1$  in a neighbourhood of  $t_1$ , which is impossible). So the optimal solution for  $T > 2\sqrt{2h}$  is



EXAMPLE 4. We discretize Example 3: let  $N \in \mathbb{N}$  be fixed and change in Example 3 the condition  $x(t) \leq h$  by

$$x(t_k) \leq h \quad k = 0, \dots, N, \quad t_k = k \frac{T}{N}.$$

By Theorem 2 the same statements (4.8)–(4.12) hold as above with the additional assumption

$$\mu_1(t) = \sum_{\substack{t_k < t \\ k \geq 1}} a_k$$

where

$$a_k \geq 0, \quad x(t_k) \neq h \Rightarrow a_k = 0 \quad (k = 1, \dots, N - 1).$$

LEMMA 12.  $a_0 > 0$ .

PROOF. The same as Lemma 9. ■

Suppose in what follows that

$$a_0 = 1.$$

LEMMA 13. For  $T > 2\sqrt{2h}$  there exists  $t_k$  with  $x(t_k) = h$ .

Proof as in Lemma 11. ■

LEMMA 14. If  $x(t_{k_1}) = x(t_{k_2}) = h$  then between  $k_1$  and  $k_2$  there are no two consecutive indices  $k$  for which  $x(t_k) < h$ .

PROOF. Let  $x(t_{k_1}) = x(t_{k_2}) = h$  and between  $k_1$  and  $k_2$ ,  $x(t_k) < h$ . We have to show that  $k_2 - k_1 \leq 2$ . Consider the parabola  $-\frac{1}{2}t^2 + \alpha t + \beta$  going through the points  $(0, h)$  and  $(\frac{T}{N}, h)$  and denote  $m^*$  its derivative at  $t = 0$ . Clearly  $m^* = o(1)$  as  $N \rightarrow \infty$ . Since the trajectory  $x(t)$  is a smooth function making piecewise from the shifted parabolas  $\frac{1}{2}t^2$  and  $-\frac{1}{2}t^2$ , hence

$$x(t_k) = h \Rightarrow |\dot{x}(t_k)| \leq m^*.$$

Indeed,  $\dot{x}(t_k) > m^*$  would imply  $x(t_{k+1}) > h$  and  $\dot{x}(t_k) < -m^*$  that  $x(t_{k-1}) > h$ . We see also that

$$x(t_k) = h, \dot{x}(t_k) = m^* \Rightarrow x(t_{k+1}) = h$$

$$x(t_k) = h, \dot{x}(t_k) = -m^* \Rightarrow x(t_{k-1}) = h$$

We investigate the sign of  $\Psi_2$  on  $(t_{k_1}, t_{k_2})$ . Since on  $(t_{k_1}, t_{k_2})$   $\Psi_2$  is a convex parabola, we have the following four cases

a)  $\Psi_2$  has constant sign on  $(t_{k_1}, t_{k_2})$ . Then  $x(t)$  is a parabola going through  $(t_{k_1}, h)$  and  $(t_{k_2}, h)$  which is possible only when  $k_2 = k_1 + 1$  (otherwise  $|\dot{x}(t_{k_1})| > m^*$  by the definition of  $m^*$ ).

b) There exists  $t_{k_1} < t^* < t_{k_2}$  such that  $\Psi_2$  is positive on  $(t_{k_1}, t^*)$  and negative on  $(t^*, t_{k_2})$ . Since  $x(t)$  is convex on  $[t_{k_1}, t^*]$  and  $x(t_{k_1}) = h, x(t_{k_1+1}) < h$  hence  $t^* < t_{k_1+1}$  (otherwise  $|\dot{x}(t_{k_1})| > m^*$ ), further we see that  $|\dot{x}(t^*)| < m^*$ . Indeed, if  $|\dot{x}(t^*)| > m^*$  then  $\dot{x}(t^*) > m^*$  and  $x(t^*) > h$ . Since on  $(t^*, t_2)$   $x(t)$  is concave,  $x(t) > h$  on  $(t^*, t^* + \frac{T}{N})$  which is a contradiction, since then  $x(t_{k_1+1}) > h$ . Since  $|\dot{x}(t^*)| < m^*$  then from  $|\dot{x}(t_{k_2})| \leq m^*$  we get that  $t_{k_2} - t^* < \frac{T}{N}$  (since in  $(t^*, t_{k_2})$   $\dot{x}$  is concave). This is possible only if  $k_2 = k_1 + 1$ .

c) There exists  $t_{k_1} < t^* < t_{k_2}$  such that  $\Psi_2$  is negative on  $(t_{k_1}, t^*)$  and positive on  $(t^*, t_{k_2})$ . This implies  $k_2 = k_1 + 1$  by the arguments symmetrical to those of b).

d) There exists  $t_{k_1} < t_1^* < t_2^* < t_{k_2}$  such that  $\Psi_2$  is positive on  $(t_{k_1}, t_1^*)$  and  $(t_2^*, t_{k_2})$  and negative on  $(t_1^*, t_2^*)$ . As in b) we see that  $t_1^* < t_{k_1+1}, t_2^* > t_{k_2-1}$  and  $|\dot{x}(t_i^*)| < m^*$ . Hence  $|\dot{x}(t_1^*)| < m^*$  and  $|\dot{x}(t_2^*)| < m^*$  and then  $t_2^* - t_1^* < \frac{T}{N}$  hence  $k_2 - k_1 \leq 2$ . Lemma 14. is proved. ■

Denote  $t_{k_1}$  and  $t_{k_2}$  the first and the last node with  $x(t_k) = h$ .

LEMMA 15. If  $T > 2\sqrt{2h}$  then

a)  $t_{k_1} = \sqrt{2h} + o(1), t_{k_2} = T - \sqrt{2h} + o(1)$  if  $N \rightarrow \infty$ ,

b)  $\Psi_2(t_{k_1}) = o(1), \Psi_2(t_{k_2}) = o(1)$ .

PROOF.  $\Psi_2$  is a convex parabola on  $[0, t_{k_1}]$  and  $\Psi_2(0) = 0$ ; hence the following cases arise.

1)  $\Psi_2 > 0$  on  $(0, t_{k_1})$ . Then  $x(t)$  is convex there and  $x(0) = 0$  which implies for large  $N$  that  $\dot{x}(t_{k_1}) > m^*$ , which is a contradiction.

2)  $\Psi_2 < 0$  on  $(0, t_{k_1})$ , then  $x(t)$  is concave there. Since  $x(0) = 0, x(t_{k_1}) = h, |\dot{x}(t_{k_1})| \leq m^* = o(1)$  hence  $x(t)$  approximates uniformly the optimal solution of Example 3 on  $(0, \sqrt{2h})$  and  $t_{k_1} = \sqrt{2h} + o(1)$ . Denote  $t^*$  the smallest value for which  $\Psi_2(t^*) = 0$ . Since on  $(0, t^*)$  the parabola  $x(t)$  can take the value  $h$  only twice, this means by Lemma 14 that  $t^* = t_{k_1} + o(1)$ . We know that  $\Psi_2(t_{k_1}-) \geq \Psi_2(t_{k_1}+)$ , hence  $\Psi_2(t_{k_1}) = o(1)$ .

3) There exists  $0 < t^* < t_{k_1}$  such that  $\Psi_2 < 0$  on  $(0, t^*)$  and  $\Psi_2 > 0$  on  $(t^*, t_{k_1})$ . Since  $x(t_{k_1}) = h$  and  $\dot{x}(t_{k_1}) = o(1)$ , it implies that  $t_{k_1} - t^* = o(1)$

(otherwise there exists  $t^* < t_k < t_{k_1}$  with  $x(t_k) > h$ ). Consequently  $x(t^*) = h + o(1)$ ,  $\dot{x}(t^*) = o(1)$  and then  $t^* = \sqrt{2h} + o(1)$ ,  $t_{k_1} = \sqrt{2h} + o(1)$  and  $\Psi_2(t_{k_1}) = o(1)$ . So in all cases  $t_{k_1} = \sqrt{2h} + o(1)$  and  $\Psi_2(t_{k_1}) = o(1)$ . The analogous statements for  $t_{k_2}$  can be proved symmetrically. Lemma 15 is proved. ■

Introduce the notations

$$x_N, u_N, \Psi_{1,N}, \Psi_{2,N}, \mu_{1,N}$$

for the functions  $x, u, \Psi_1, \Psi_2, \mu_1$  in Example 4 and the original notations means the functions from Example 3.

LEMMA 16.  $x_N \rightarrow x$  uniformly,  $u_N \rightarrow u$  weakly.

PROOF. We have seen in proving Lemma 15 that in  $[0, t_{k_1}]$   $x_N - x = o(1)$  uniformly and by symmetry the same is true in  $[t_{k_2}, T]$ . Between  $t_{k_1}$  and  $t_{k_2}$   $x_N \rightarrow h$  uniformly (see Lemma 14 and the fact that  $x(t_k) = h$  implies  $|\dot{x}(t_k)| \leq m^* = o(1)$ ). Since  $t_{k_1} \rightarrow \sqrt{2h}$ ,  $t_{k_2} \rightarrow T - \sqrt{2h}$  this means that  $x_N \rightarrow x$  uniformly. In  $(0, t_{k_1})$  and  $(t_{k_2}, T)$   $u_N = -1$ . In  $(t_{k_1}, t_{k_2})$   $\dot{x}(t) = o(1)$  uniformly. Since  $|u_N| = \pm 1$  and  $\dot{x}$  is the integral function of  $u_N$ , we get that  $u_N \rightarrow 0$  weakly on  $(t_{k_1}, t_{k_2})$ , which proves Lemma 16. ■

So the conditions of Theorem 3 are checked. To see the consequences we shall prove first that on a segment  $[t_{k_1} + o(1), t_{k_2} - o(1)]$  we have  $\Psi_{1,N}(t) = o(1)$ . Indeed,  $\Psi_{1,N}$  can have positive jumps at the nodes  $t_k$  and between the nodes it decreases linearly, its derivative being  $-1$ . Now if  $\Psi_{1,N}(t) \neq o(1)$ ,  $\Psi_{1,N}(t) > 0$  for some  $t$  from this interval then  $\Psi_{1,N} > 0$  in some interval  $[t, t + \delta]$ ,  $\delta \neq o(1)$  hence  $\Psi_{2,N}$  is strictly monotone there i.e.  $\Psi_{2,N}$  can have only one zero in  $[t, t + \delta]$  which contradicts Lemma 14. Analogously if  $\Psi_{1,N}(t) \neq o(1)$ ,  $\Psi_{1,N}(t) < 0$  for some  $t$  then  $\Psi_{1,N} < 0$  in some interval  $[t - \delta, t]$   $\delta \neq o(1)$  which is impossible as well. So in  $[t_{k_1} + o(1), t_{k_2} - o(1)]$   $\Psi_{1,N}(t) \rightarrow 0$  uniformly. Since  $\Psi_{2,N}(0) = 0$ ,  $\Psi_{2,N}(t_{k_1}) = o(1)$  hence  $\Psi_{2,N}$  decreases almost precisely on the left half of  $[0, t_{k_1}]$  and increases on the right half. This means by  $t_{k_1} = \sqrt{2h} + o(1)$  and  $\dot{\Psi}_{1,N} = -1$  that  $b_{1,N} \rightarrow \frac{1}{2}\sqrt{2h} = b_1$  and  $\Psi_{1,N} - \Psi_1 = o(1)$  uniformly on  $[0, t_{k_1})$ , on  $[t_{k_1} + o(1), t_{k_2} - o(1)]$  and because of symmetry on  $(t_{k_2}, T]$ ; finally  $\Psi_{1,N} = O(1)$  on  $[t_{k_1}, t_{k_1} + o(1)]$  and on  $[t_{k_2} - o(1), t_{k_2}]$ . Hence  $\Psi_{1,N} \rightarrow \Psi_1$  a.e. on  $[0, T]$  follows from  $\Psi_{2,N}(0) = 0$ ,  $\Psi_{2,N}(t_{k_1}) = o(1)$  that  $\Psi_{2,N} - \Psi_2 = o(1)$  uniformly in  $[0, t_{k_1}]$  and similarly in  $[t_{k_2}, T]$ . In  $[t_{k_1}, t_{k_2}]$   $\Psi_{2,N}$  have zeros at a distance  $o(1)$  by Lemma 14 and we have seen that  $\Psi_{1,N} = -\dot{\Psi}_{2,N}$  is uniformly bounded hence  $\Psi_{2,N} = o(1)$  uniformly on  $[t_{k_1}, t_{k_2}]$ . Comparing this with Lemma 15 we get that  $\Psi_{2,N} - \Psi_2 = o(1)$  uniformly on  $[0, T]$ . Finally since  $\mu_{1,N}(t) = t - b_{1,N} + \Psi_{1,N}(t)$

hence  $\mu_{1,N} - \mu = o(1)$  uniformly on  $[0, t_{k_1})$ ,  $[t_{k_1} + o(1), t_{k_2} - o(1)]$  and  $(t_{k_2}, T]$  and then  $\mu_{1,N} \rightarrow \mu_1$  a.e. on  $[0, T]$ . Thus all statement of Theorem 3 are shown in the situation of Examples 3 and 4.

### 5. Some spectral properties of the Schrödinger and Dirac operators

In the final part of the paper we investigate some spectral properties of the Schrödinger and Dirac operators. In several papers on equiconvergence of spectral expansions the authors use the boundedness of the difference of the so-called spectral functions.

If the Schrödinger operator has discrete spectrum then its spectral function is defined to be

$$\Theta(x, y, \mu) = \sum_{\sqrt{\lambda} < \mu} u_n(x) u_n(y)$$

and it turns out that if we consider Schrödinger operators  $L_1 = -u'' + q_1 u$ ,  $L_2 = -u'' + q_2 u$  then  $\Theta_1(x, y, \mu) - \Theta_2(x, y, \mu)$  is locally uniformly bounded. Here we want to remark that we know more, namely  $\Theta_1 - \Theta_2$  tends to zero as  $\mu \rightarrow \infty$  locally uniformly. Namely let

$$(5.1) \quad q \in L_1(0, \pi) \quad \int_0^\pi q = 0$$

and consider the eigenvalue problem

$$(5.2) \quad \begin{aligned} [-u_n'' + q u_n] &= \lambda_n u_n \quad (n \geq 0), \\ u_n'(0) &= u_n'(\pi) = 0. \end{aligned}$$

The corresponding problem for  $q \equiv 0$  has system of solutions

$$(5.3) \quad u_n^{(0)}(x) := \cos nx \quad n \geq 0.$$

Consider further the system

$$(5.4) \quad -\varphi'' + q\varphi = \lambda\varphi, \quad \varphi(0) = 1, \quad \varphi'(\pi) = 0.$$

It has solution for all  $\lambda$ . We choose  $\lambda_n$  from the equality

$$(5.5) \quad \varphi'(\pi) = 0$$

then the corresponding function  $\varphi(x, \lambda_n) = u_n(x) \cdot c_n$ .

LEMMA 1.

$$(5.6) \quad \varphi(x) = \cos \sqrt{\lambda}x + \frac{\sin \sqrt{\lambda}x}{2\sqrt{\lambda}} \int_0^x q + O\left(\frac{1}{\lambda}\right) \quad \lambda \geq \lambda_0.$$

PROOF. The mean value formula

$$(5.7) \quad \varphi(x) = \cos \sqrt{\lambda}x + \int_0^x \frac{\sin \sqrt{\lambda}(x-\tau)}{\sqrt{\lambda}} q(\tau) \varphi(\tau) d\tau.$$

implies

$$\|\varphi\|_\infty \leq 1 + \int_0^x \frac{1}{\sqrt{\lambda}} |q| \cdot \|\varphi\|_\infty$$

hence  $\|\varphi\|_\infty$  is bounded in  $\lambda$  for  $\lambda \geq \lambda_0$ . Consequently

$$|\varphi(x) - \cos \sqrt{\lambda}x| \leq \int_0^x \frac{1}{\sqrt{\lambda}} |q(\tau)| d\tau \cdot \|\varphi\|_\infty \leq \frac{c}{\sqrt{\lambda}} \|\varphi\|_\infty \leq \frac{c}{\sqrt{\lambda}}$$

i.e.

$$\varphi(x) = \cos \sqrt{\lambda}x + O\left(\frac{1}{\sqrt{\lambda}}\right).$$

Substitute this into (5.7) we get

$$\varphi(x) = \cos \sqrt{\lambda}x + \int_0^x \frac{\sin \sqrt{\lambda}(x-\tau) \cos \sqrt{\lambda}\tau}{\sqrt{\lambda}} q(\tau) d\tau + O\left(\frac{1}{\lambda}\right)$$

which gives (5.6). ■

LEMMA 2.

$$(5.8) \quad \varphi'(x) = -\sqrt{\lambda} \sin \sqrt{\lambda}x + \frac{\cos \sqrt{\lambda}x}{2} \int_0^x q + O\left(\frac{1}{\sqrt{\lambda}}\right).$$

PROOF. Similar to that of Lemma 1. ■

COROLLARY.

$$\sqrt{\lambda_n} = n + O\left(\frac{1}{n^2}\right)$$

$$(5.10) \quad \varphi(x, \lambda_n) = \cos nx + \frac{\sin nx}{2n} \int_0^x q + O\left(\frac{1}{n^2}\right) \quad n \geq 1$$

$$(5.11) \quad u_n(x) = \sqrt{\frac{2}{\pi}} \left[ \cos nx + \frac{\sin nx}{2n} \int_0^x q + O\left(\frac{1}{n^2}\right) \right], \quad n \geq 1$$

if  $u_n(x)$  are the normed eigenfunctions of (5.2) in  $L_2(0, T)$ .

PROOF. From  $0 = \varphi'(\pi)$  we get by (5.8) and (5.1)  $\sqrt{\lambda_n} \sin \sqrt{\lambda_n} \pi = O\left(\frac{1}{\sqrt{\lambda_n}}\right)$ . Consequently we get (5.9) Putting this into (5.6) implies (5.10). Finally,

$$\begin{aligned} \|\varphi\|_2^2 &= \int_0^\pi [\varphi(x, \lambda_n)]^2 dx = \int_0^\pi \left( \cos^2 nx + \frac{\sin 2nx}{2n} \int_0^x q \right) dx + O\left(\frac{1}{n^2}\right) = \\ &= \int_0^\pi \cos^2 nx dx + O\left(\frac{1}{n^2}\right) = \frac{\pi}{2} + O\left(\frac{1}{n^2}\right) \end{aligned}$$

which gives (5.11) since  $u_n(x) = \frac{\varphi(x, \lambda_n)}{\|\varphi\|_2}$ . ■

Consider the spectral functions

$$\Theta(x, y, N) = \sum_{n \leq N} u_n(x) u_n(y)$$

and

$$\Theta_0(x, y, N) = \sum_{n \leq N} u_n^{(0)}(x) u_n^{(0)}(y)$$

We know from (5.11) that

$$(5.12) \quad \begin{aligned} u_n(x) u_n(y) - u_n^{(0)}(x) u_n^{(0)}(y) &= \\ &= \frac{2 \cos nx \sin ny \cdot \beta(y) + \cos ny \sin nx \cdot \beta(x)}{2n} + O\left(\frac{1}{n^2}\right) \end{aligned}$$

where  $\beta(x) = \int_0^x q$ .

THEOREM 4. Add to the above assumption that  $q \in C[0, \pi]$ . Then

$$(5.13) \quad \lim_{N \rightarrow \infty} (\Theta(x, y, N) - \Theta_0(x, y, N)) = 0$$

uniformly in  $0 < \epsilon \leq x, y \leq \pi - \epsilon$  for every  $\epsilon > 0$ .

PROOF. We know from the equiconvergence theorem (see e.g. [20]) and from the orthogonality of eigenfunctions that for any  $\varphi \in D(0, \pi)$ ,

$$\sum_{n=0}^{\infty} u_n(x)(\varphi, u_n) = \varphi(x).$$

This means that for every  $x \in (0, \pi)$

$$\sum u_n(x)u_n(\cdot) = \delta_x(\cdot).$$

The same is true for  $\sum u_n^0(x)u_n^0(\cdot)$ , hence

$$(5.14) \quad \sum (u_n(x)u_n(\cdot) - u_n^0(x)u_n^0(\cdot)) = 0 \quad \text{in } D'(0, \pi).$$

From (5.12) we see that the series of (5.14) converges in  $L_2(0, \pi)$  for every fixed  $x$  and hence the equality (5.14) holds also in  $L_2$ -sense.

In ZYGMUND [9], Chapter I, Theorem 2.6 is proved that for any given  $\epsilon > 0$  the series

$$\sum \frac{\sin ny}{n}, \quad \sum \frac{\cos ny}{n}$$

are convergent uniformly in  $0 < \epsilon \leq y \leq 2\pi - \epsilon$ . Now (5.12) implies that

$$(5.15) \quad \begin{aligned} & \sum (u_n(x)u_n(y) - u_n^0(x)u_n^0(y)) = \\ & = c \cdot \sum \left( \frac{\sin(x-y)[\beta(x) - \beta(y)] + \sin n(x+y)[\beta(x) + \beta(y)]}{n} + O\left(\frac{1}{n^2}\right) \right). \end{aligned}$$

Here the uniformity of the convergence of

$$[\beta(x) + \beta(y)] \sum \frac{\sin n(x+y)}{n}$$

follows immediately from  $2\epsilon \leq x+y \leq 2\pi - 2\epsilon$ . We shall prove on the other hand that

$$(5.16) \quad \sum [\beta(x) - \beta(y)] \frac{\sin n(x-y)}{n} \quad \text{is uniformly convergent for } |x-y| < 2\pi - \epsilon.$$

Indeed, it is immediate for  $x, y$  satisfying  $|x-y| > \epsilon$ . In case  $|x-y| \leq \epsilon$  we argue as follows. Introduce the conjugate Dirichlet kernel

$$\tilde{D}_k(t) = \sum_{\nu=1}^k \sin \nu t = \frac{\cos \frac{t}{2} - \cos \left(k + \frac{1}{2}\right) t}{2 \sin \frac{t}{2}}.$$

Apply a usual Abel transformation to get

$$(5.17) \quad \sum_{k=N+1}^M \frac{\sin k(x-y)}{k} = \sum_{k=N+1}^M \frac{\tilde{D}_k(x-y) - \tilde{D}_{k-1}(x-y)}{k} = \\ = \frac{\tilde{D}_M(x-y)}{M} - \frac{\tilde{D}_N(x-y)}{N} + \sum_{k=N+1}^{M-1} \tilde{D}_k(x-y) \left( \frac{1}{k} - \frac{1}{k+1} \right).$$

Clearly:

$$\|t \tilde{D}_k(t)\|_{L_\infty(-1,1)} = O(1)$$

and

$$|\beta(x) - \beta(y)| \leq \|q\|_\infty \cdot |x - y|;$$

hence by (5.17) we get

$$\left| \sum_{k=N+1}^M \frac{\sin k(x-y)}{k} [\beta(x) - \beta(y)] \right| \leq \\ \leq c \|q\|_\infty \left( \frac{1}{M} + \frac{1}{N} + \sum_{k=N+1}^{M-1} \left( \frac{1}{k} - \frac{1}{k+1} \right) \right) = O\left(\frac{1}{N}\right).$$

So we have proved (5.16) and this means that the series (5.15) converges uniformly for  $\epsilon \leq x, y \leq \pi - \epsilon$ . We have seen on the other hand that for every fixed  $x$  this series converges to zero in  $L_{2,y}(0, \pi)$ , so the local uniform limit must be also zero. Theorem 4 is proved. ■

6. Next we shall prove the following formula for the eigenvalue problem (5.2).

**THEOREM 5.** *Let  $q \in C^1[0, \pi]$ ,  $\int_0^\pi q = 0$ . Then*

$$(6.1) \quad \sum_{n=0}^{\infty} (\lambda_n - n^2) = \frac{q(0) + q(\pi)}{4}.$$

**REMARK.** Analogous statement is given in LEVITAN, SARGSYAN [8], Chapter VIII, Section 5 for the sum  $\sum_{n=-\infty}^{\infty} \lambda_n$ , where  $\lambda_n$  are the eigenvalues of the (one-dimensional) Dirac operator. Since for large  $n$ ,  $\lambda_n$  „almost” equals to  $n$ ,  $\lambda_n + \lambda_{-n}$  is small; this is substituted in (6.1) by  $\lambda_n - n^2$ .

For the proof of Theorem 5 we introduce the parametric eigenvalue problem where in (5.2) the potential  $q$  is substituted by  $\tan(x)$ ,  $0 \leq t \leq 1$ . To this, consider the system

$$(6.2) \quad -\varphi'' + \tan(x)\varphi = \lambda\varphi, \quad \varphi = \varphi(x, \lambda, t), \quad \varphi(0) = 1, \quad \varphi'(0) = 0.$$

Since this equation is linear in  $t$  and  $\lambda$ , we get from known theorems (see e.g. HARTMAN [21], p. 100) that  $\varphi, \varphi', \varphi'' \in C^\infty$  in the variables  $(\lambda, t)$  provided that  $q \in C[0, \pi]$ . Consequently

$$F(\lambda, t) := \varphi'(\pi)$$

also belongs to  $C^\infty$  in  $(\lambda, t)$ . We know from (5.7) that

$$\varphi'(x) = -\sqrt{\lambda} \sin \sqrt{\lambda} x + \int_0^x \cos \sqrt{\lambda}(x - \tau) \cdot \tan(\tau) \varphi(\tau) d\tau,$$

Hence, using the notation of  $s = \sqrt{\lambda}$ ,

$$(6.3) \quad F = -s \sin s\pi + \int_0^\pi \cos s(\pi - \tau) \cdot \tan(\tau) \varphi(\tau) d\tau,$$

$$(6.4) \quad \frac{\delta F}{\delta s} = -s\pi \cos s\pi - \sin s\pi + \int_0^\pi (\tau - \pi) \sin s(\pi - \tau) \cdot \tan(\tau) \varphi(\tau) d\tau$$

If  $\lambda(t)$  denotes any solution of the eigenvalue problem with potential  $\tan(x)$ ,  $t$  fixed, then

$$0 = F(\lambda(t), t) = -\sin s\pi + \int_0^\pi \cos s(\pi - \tau) \cdot \tan(\tau) \varphi(\tau) d\tau$$

i.e.

$$\sin s\pi = O(1)$$

which means that

$$(6.5) \quad s = k + O\left(\frac{1}{k}\right)$$

and for these value of  $s$ ,

$$\frac{\delta F}{\delta s} = (-1)^{k+1} \pi k + O(1)$$

which is not zero, at least for large  $k$ . Hence by the implicit function theorem

$$\lambda_n(t) \in C^\infty[0, 1].$$

Now the system

$$(6.6) \quad -u_n''(x, t) + \tan(x)u_n(x, t) = \lambda_n(t)u_n(x, t)$$

has solution  $u$  where  $u$ ,  $u'$  and  $u''$  are  $C^\infty$  in  $t$ .

LEMMA 3.

$$(6.7) \quad \int_0^\pi u_n \dot{u}_n = 0$$

$$(6.8) \quad \int_0^1 \int_0^\pi q(x)u_n^2(x, t)dx dt = \lambda_n - n^2 \quad (\lambda_n = \lambda_n(1)).$$

PROOF. (6.7) is obtained by differentiating  $\int_0^\pi u_n^2 = 1$  in  $t$ . To see (6.8) we differentiate first (6.6):

$$-\dot{u}_n'' + q u_n + \tan \dot{u}_n = \dot{\lambda}_n u_n + \lambda_n \dot{u}_n.$$

Multiplying by  $u_n$  and integrating in  $(0, 1) \times (0, \pi)$  we get that

$$\begin{aligned} \int_0^1 \int_0^\pi q u_n^2 &= \int_0^1 \int_0^\pi \dot{\lambda}_n u_n^2 + \int_0^1 \int_0^\pi \lambda_n u_n \dot{u}_n - \int_0^1 \int_0^\pi \tan u_n \dot{u}_n + \int_0^1 \int_0^\pi \dot{u}_n'' u_n = \\ &= \lambda_n(1) - \lambda_n(0) - \int_0^1 \int_0^\pi \tan u_n \dot{u}_n + \int_0^1 \int_0^\pi u_n'' \dot{u}_n = \\ &= \lambda_n(1) - \lambda_n(0) - \int_0^1 \int_0^\pi \lambda_n u_n \dot{u}_n = \lambda_n(1) - \lambda_n(0). \end{aligned}$$

PROOF OF THEOREM 5 By (6.8)

$$\sum_{n=0}^N (\lambda_n - n^2) = \sum_{n=0}^N \int_0^1 \int_0^\pi q u_n^2 =$$

$$\int_0^1 \int_0^\pi q \sum_{n=0}^N (u_n^2 - u_n^{(0)2}) + \int_0^1 \int_0^\pi q \sum_{n=0}^N u_n^{(0)2} =: I_1 + I_2$$

We have seen in proving (5.11) that

$$u_n(x, t) = \sqrt{\frac{2}{\pi}} \left( \cos nx + \frac{\sin nx}{2n} \cdot t \int_0^x q + O\left(\frac{1}{n^2}\right) \right), \quad n \geq 1$$

where the implicit constant in  $O\left(\frac{1}{n^2}\right)$  is bounded in  $(x, t)$ . Consequently the series

$$\sum_1^\infty \left( u_n^2(x, t) - u_n^{(0)2}(x) \right) = \frac{2}{\pi} \sum_1^\infty \left( t\beta(x) \frac{\sin 2nx}{n} + O\left(\frac{1}{n^2}\right) \right),$$

$\beta(x) = \int_0^x q$  is convergent in  $L_2((0, \pi) \times (0, 1))$ . By Theorem 4, the series

$\sum_0^\infty (u_n^2 - u_n^{(0)2})$  tends to zero locally uniformly for any fixed  $t$ . Hence

$$\sum_0^\infty (u_n^2 - u_n^{(0)2}) = 0 \quad \text{in } L_2((0, \pi) \times (0, 1))$$

and then

$$I_1 = \int_0^1 \int_0^\pi q \sum_0^N (u_n^2 - u_n^{(0)2}) \rightarrow 0 \quad (N \rightarrow \infty).$$

In counting  $I_2$  we use several times that  $\int_0^\pi q = 0$

$$\begin{aligned} I_2 &= \int_0^1 \int_0^\pi q \sum_1^N u_n^{(0)2} = \int_0^1 \int_0^\pi q \sum_1^N \frac{2}{\pi} \cos^2 ux \, dx \, dt = \\ &= \int_0^1 \int_0^\pi q \cdot \frac{1}{\pi} \sum_1^N \cos 2ux \, dx \, dt = \frac{1}{\pi} \int_0^1 \int_0^\pi q(x) D_N(2x) \, dx \, dt \end{aligned}$$

when

$$D_N(x) = \frac{1}{2} + \sum_{n=1}^N \cos nx$$

is the usual Dirichlet kernel. By the theory of Fourier series

$$I_2 = \frac{1}{2\pi} \int_0^1 \int_0^{2\pi} q\left(\frac{x}{2}\right) D_N(x) dx dt = \frac{1}{2} \cdot \frac{1}{\pi} \int_0^{2\pi} q\left(\frac{x}{2}\right) D_N(x) dx \rightarrow \frac{1}{2} \frac{q(0) + q(\pi)}{2}$$

Theorem 5 is proved. ■

### References

- [1] A. JA. DUBOVITSKII and A. A. MILYUTIN, Extremum problems with the presence of constraints, *USSR Comput. Math. and Math. Phys.*, (1965), 5:3, 395–453.
- [2] A. JA. DUBOVITSKII and A. A. MILYUTIN, The theory of maximum principle. — in *Methods of the theory of extremal problems in economics*, (in Russian), (ed. V. L. Levin), Moscow, Nauka, (1981).
- [3] A. A. MILYUTIN, Maximum principle for regular systems. — in *The necessary condition of optimal control*, (in Russian), Moscow, Nauka, (1990).
- [4] A. A. MILYUTIN, Maximum principle for a general optimal control problem with state and regular mixed constraints — in *Optimality in controllable dynamical systems*, (in Russian), Moscow, VNIISI, (1990, no. 14.)
- [5] L. S. PONTRYAGIN, *Mathematical theory of optimal processes*, Moscow, (1961).
- [6] A. D. IOFFE and V. M. TIKHOMIROV, *Theory of extremal problems*, Moscow, Nauka, (1974).
- [7] M. B. SUBRAHMANYAM, Optimal Control with a Worst-Case Performance Criterion and Applications, *Lecture Notes in Control and Inf. Sci.*, (1990), vol. 145.
- [8] B. M. LEVITAN and I. S. SARGSYAN, Sturm–Liouville and Dirac operators, (in Russian), Nauka, Moscow, (1988).
- [9] A. ZYGMUND, *Trigonometric series*, Cambridge Univ. Press, New York, (1959).
- [10] L. S. PONTRYAGIN, Optimal processes of regulation (in Russian), in *Proceedings of the International Congress of Mathematicians 1958*, Cambridge Univ. Press, 182–202.
- [11] G. A. BLISS, The problem of Lagrange in the calculus of variations, *American Journal of Math.*, 52 (1930), 673–744.
- [12] E. J. MCSHANE, On multipliers for Lagrange problems, *American Journal of Math.*, 61 (1939), 809–819.
- [13] A. JA. DUBOVITSKII and A. A. MILYUTIN, Extremal problems in presence of constraints, (in Russian), *Doklady A. N. SSSR*, 149(4) (1963), 759–762.
- [14] A. FELDBAUM, On the synthesis of optimal systems using phase space, (in Russian), *Avtomatika i Telemekh.*, 16(2) (1955), 129–149.

- 
- [15] V. BOLTYANSKII, Maximum principle in the theory of optimal processes. (in Russian), *Doklady A. N. SSSR*, **119**(6), (1958), 1070–1073.
- [16] R. GAMKRELIDZE, On the theory of optimal processes in linear systems, (in Russian), *Doklady A. N. SSSR*, **116**(1) (1957), 9–11.
- [17] V. BOLTYANSKII and R. GAMKRELIDZE, On the theory of optimal processes, *Doklady A. N. SSSR*, **110**(1) (1956), 7–10.
- [18] E. C. LEVITAN, A. A. MILYUTIN, and N.P. OSMOLOVSKII, Higher order conditions of local minimum in the problems with constraints, (in Russian), *Uspeki Mat. Nauka*, **33**(6) (1978), 85–148.
- [19] I. V. GIRSANOV, *Lectures on the mathematical theory of extremal problems* (in Russian), Izd. Moskovsko Univ., (1970).
- [20] I. JOÓ and V. KOMORNIK, On the equiconvergence of expansions by Riesz bases formed by eigenfunctions of the Schrödinger operator, *Acta Sci. Math.*, **46** (1983), 357–375.
- [21] PH. HARTMAN, *Ordinary differential equations*.



## REMARKS ON OUR ARTICLE ON LEBESGUE FUNCTIONS

By

S. BARON and I. JOÓ

Department of Mathematics and Computer Science, Bar-Ilan University, Ramat Gan and  
Department of Analysis, Eötvös Loránd University, Budapest

In our paper “Lebesgue functions for convergence and summability of double function series”, *Annales Univ. Sci. Budapest.*, Sectio Math., **37** (1994), 73–84, for an unknown reason some references are wrong. The correct references are the following:

- [11] ALEXITS, G., *Konvergenzprobleme der Orthogonalreihen*, Akadémiai Kiadó, Budapest, 1960.
- [16] KAMKE, E., *Das Lebesgue–Stieltjes-Integral*, 2nd edition, Teubner Verlag, Leipzig, 1960.
- [20] MÓRICZ, F., On the regular summability of multiple function series and Lebesgue functions, *Anal. Math.*, **9** (1983), 57–67.
- [22] ZYGMUND, A., *Trigonometrical Series*, Warsaw–Lwow, 1935.

It must be on the page 74 line 3 from below the following:

$$s = s(m, n, x), \quad t = t(m, n, x)$$

Similar changes are in the last line on page 75.



## I N D E X

ALI, Z., MAJEED, A.: On the laws of metabelian pointed-groups .....	3
ALZER, H.: A refinement of the entropy inequality .....	13
BAJPAI, S. D.: An orthogonal type property for the Jacobi polynomials .....	31
BALAGUER, B.: Compressed and Cauchy filters .....	153
BARON, S., JOÓ, I.: On the influence of Lebesgue functions on the summability of double function series .....	223
BARON, S., JOÓ, I.: Remarks on our article on Lebesgue functions .....	269
<u>DEÁK, J.</u> , ROMAGUERA, S.: Co-stable quasi-uniform spaces .....	55
HORVÁTH, Á. G.: On a problem connected with the weight distribution of the Reed–Muller code of order $R$ .....	171
HURELBAATAR, G.: An almost sure local and global central limit theorem for weakly dependent random variables .....	109
JOÓ, I., BARON, S.: On the influence of Lebesgue functions on the summability of double function series .....	223
JOÓ, I., BARON, S.: Remarks on our article on Lebesgue functions .....	269
JOÓ, I., KASSAY, G.: Convexity, minimax theorems and their applications ....	71
JOÓ, I., MINKIN, A.: General boundary value problem for a differentiation operator .....	137
JOÓ, I., MINKIN, A.: Bases of vector exponentials with a strip in the spectrum	177
JOÓ, I., SU, N. V.: Internal controllability of the string with restrained controls	19
JOÓ, I.: An Alexits type theorem .....	161
JOÓ, I.: On weighted $(0, 2)$ interpolation .....	185
JOÓ, I.: On control theory I (Maximum principle) .....	237
KASSAY, G., JOÓ, I.: Convexity, minimax theorems and their applications ....	71

KOVÁCS, K.: On additive functions satisfying some congruence properties II .	35
KY, N. X.: A method for characterization of weighted $K$ -functionals.....	147
MAJEED, A., ALI, Z.: On the laws of metabelian pointed-groups.....	3
MENESES, J., NAULIN, R.: Ascoli–Arzelá theorem for a class of right continuous functions.....	127
MINKIN, A., JOÓ, I.: General boundary value problem for a differentiation operator.....	137
MINKIN, A., JOÓ, I.: Bases of vector exponentials with a strip in the spectrum	177
MOLNÁR, E., SZIRMAI, J.: Einige Pflasterungen des hyperbolischen Raumes mittels Flächentransitiver Bewegungsgruppen.....	95
NAULIN, R., MENESES, J.: Ascoli–Arzelá theorem for a class of right continuous functions.....	127
PROKAJ, V.: On uniqueness of extension of suboperators.....	47
ROMAGUERA, S., [DEÁK, J.]: Co-stable quasi-uniform spaces.....	55
SU, N. V., JOÓ, I.: Internal controllability of the string with restrained controls	19
SZIRMAI, J., MOLNÁR, E.: Einige Pflasterungen des hyperbolischen Raumes mittels Flächentransitiver Bewegungsgruppen.....	95
WEISZ, F.: A new martingale equivalence theorem.....	39

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: 1995. szeptember. Megjelent: 1995. december

Terjedelem: 23,3 A/4 ív. Példányszám: 500

Készült az EMT<sub>E</sub>X szedőprogram felhasználásával

az MSZ 5601–59 és 5602–55 szabványok szerint

Az elektronikus tipografálás és grafika Fried Katalin és Juhász Lehel munkája

A sokszorosítás az ELTE sokszorosító üzemében készült

Felelős vezető: Arató Tamás

**ELTE 95175**